BỘ QUỐC PHÒNG
HỌC VIỆN KỸ THUẬT QUÂN SỰ

TUYỂN TẬP CÔNG TRÌNH
HỘI NGHỊ KHOA HỌC CÁC NHÀ NGHIÊN CỨU TRẺ
LẦN THỨ XXI - NĂM 2026

(TẬP 3: CÔNG NGHỆ THÔNG TIN, KỸ THUẬT CƠ KHÍ ĐỘNG LỰC, KỸ THUẬT ĐIỀU KHIỂN VÀ TỰ ĐỘNG HÓA)

NHÀ XUẤT BẢN QUÂN ĐỘI NHÂN DÂN
ISBN: 978-632-06-0210-0

Hà Nội - 2026

BỘ QUỐC PHÒNG
HỌC VIỆN KỸ THUẬT QUÂN SỰ

TUYỂN TẬP CÔNG TRÌNH

Hội nghị khoa học các nhà nghiên cứu trẻ lần thứ XXI - năm 2026

TẬP 3: CÔNG NGHỆ THÔNG TIN, KỸ THUẬT CƠ KHÍ ĐỘNG LỰC, KỸ THUẬT ĐIỀU KHIỂN VÀ TỰ ĐỘNG HÓA

NHÀ XUẤT BẢN QUÂN ĐỘI NHÂN DÂN
ISBN: 978-632-06-0210-0

Hà Nội - 2026

MỤC LỤC

TT	Tên bài báo	Tác giả	Trang
246.	The variant of inertial algorithm for solving nonmonotone variational inequality problems	Biện Thanh Tuyền, Nguyễn Thị Thanh Hà, Nguyễn Hồng Nam	2382
247.	Well-posedness and regularity of solutions for Caputo-Hadamard stochastic fractional differential equation	Trần Thị Phương Lan, Vũ Thị Thủy, Phạm Tiến Dũng, Bùi Quốc Hưng	2391
248.	Kỹ thuật thủy vân ảnh bền vững dựa trên DWT–SVD QIM sử dụng điều khiển hỗn loạn và sinh khoá hậu lượng tử	Hoàng Thị Thu Lan, Tạ Minh Thanh	2400
249.	Fusion Deep Learning Models for Quishing Detection	Khổng Phương Thảo	2409
250.	Khảo sát các phương pháp phát hiện lỗ hổng trên smart contract non-EVM	Ninh Thái Phan	2417
251.	Đánh giá một số mô hình đa phương thức cho bài toán phát hiện tin giả	Nguyễn Thanh Hà, Phan Việt Anh, Nguyễn Trung Tín	2426
252.	A Study on Data Augmentation Strategies for Contrastive Learning in Network Intrusion Detection	Nguyễn Hữu Long	2434
253.	SSDAE: A Self-Supervised Denoising Autoencoder for Robust Intrusion Detection	Nguyễn Văn Cường	2443
254.	Developing a Deep Learning Approach for Reconnaissance Attack Detection	Nguyễn Văn Quân	2451
255.	A Hybrid Semi-Blind Color Image Watermarking Scheme with Adaptive Schur, SVD Block Selection and Dual-Channel Authentication	Phạm Thị Yến, Tạ Minh Thanh	2461
256.	Financial Crime in Decentralized Finance: A Survey	Trần Nam Khánh, Tạ Minh Thanh	2471
257.	Enhancing Network Anomaly Detection via Hierarchically Supervised Autoencoder	Bùi Đức Việt, Chu Thị Hường	2481
258.	Parameter-Aware Log-Based Anomaly Detection	Nguyễn Mạnh Tuấn	2490
259.	Contrastive Graph Autoencoder for Robust Log Anomaly Detection in Low-Data Regimes	Nguyễn Hải Nam	2497
260.	Nghiên cứu mô hình Centence-Bert để đánh giá mức độ trùng lặp văn bản hành chính tiếng việt	Phạm Thanh Tùng, Nguyễn Quốc Khánh, Phạm Thị Bích Vân	2507
261.	Unsupervised Homography Estimation for Image Stitching via Flow-Based Regression	Nguyễn Thị Vân, Phan Thị Hải Hồng, Nguyễn Tiến Phát	2516
262.	An adaptive ensemble framework for just-in-time defect prediction on imbalanced commit streams	Bùi Thị Phương Thảo, Phạm Thị Bích Vân, Nguyễn Quốc Khánh	2525
263.	Attention-Enhanced Swin UNet++ for Pavement Crack Segmentation	Đặng Thị Thương, Phan Thị Hải Hồng, Ngô Hữu Phúc, Chu Thị Hường	2534
264.	Integrating artificial intelligence and digital twin in medicine for automatic cranial	Đỗ Thành Long, Nguyễn Văn Giang, Nguyễn Mậu Uyên	2543

ISBN: 978-632-06-0210-0

TT	Tên bài báo	Tác giả	Trang
	implant design automatic cranial implant design		
265.	Phương pháp tái tạo vùng khuyết thiếu xương hàm dưới sử dụng kỹ thuật học sâu	Đỗ Thành Long, Nguyễn Văn Giang, Nguyễn Mậu Uyên, Trần Đức Tăng	2553
266.	Learning-Based Autonomous UAV Navigation via Q-Learning under Dynamic and Partially Observable Environments	Nguyễn Chí Công, Ngô Hữu Phúc, Nguyễn Trung Tín	2562
267.	Domain-Adaptive Pretraining of Open Large Language Models for Vietnamese Defense and Military Texts	Nguyễn Hải Đăng, Trần Cao Trưởng, Nguyễn Quang Tùng	2571
268.	Deep Learning-Based Tiny Object Detection in RGB-IR UAV Imagery: Methods, Challenges, and Perspectives	Nguyễn Thị Lan, Trần Cao Trưởng	2581
269.	Domain-Adaptive Pretraining for a Vietnamese Medical Large Language Model	Phạm Quang Tùng, Trần Cao Trưởng	2591
270.	Một kết hợp phân cụm mờ cộng tác và mạng bộ mã hoá tự động để phân đoạn ảnh viễn thám siêu phổ	Tạ Văn Đắc, Đỗ Khắc Ngọc, Mai Đình Sinh	2601
271.	A Method for Enhancing Vessel AIS Time Series Data Quality using Denoising and Intelligent Interpolation	Trần Thị Thanh Huyền, Nguyễn Quang Thi, Vũ Thị Ly, Nguyễn Quang Uy	2611
272.	Exploiting Vision-Language Models in Assistive Navigation Support for Visually Impaired	Trần Vũ Đức, Dương Thị Minh Soan	2619
273.	Tính toán các thông số làm việc của cơ cấu bơi kiểu bơm luồng phụt xe tăng hạng nhẹ sản xuất tại Việt Nam	Hoàng Khoái, Nguyễn Minh Tân	2629
274.	Nghiên cứu tính toán, lựa chọn động cơ bố trí khoang động lực xe thiết giáp chở quân bánh lốp 8x8	Hoàng Văn Nguyên, Tô Viết Thành	2639
275.	Tính toán, lựa chọn động cơ cho xe tăng hạng nhẹ sản xuất tại Việt Nam	Nguyễn Văn Liêm, Cù Xuân Phong	2648
276.	Đánh giá khả năng tăng tốc xe thiết giáp chở quân 8x8	Tạ Văn Tuấn, Nguyễn Minh Tân, Dương Thành Công	2655
277.	Tính toán thiết kế hộp phân phối trung tâm xe thiết giáp chở quân bánh lốp 8x8	Vũ Đức Vinh, Cù Xuân Phong	2664
278.	Khảo sát động lực học bơi xe thiết giáp chở quân có công thức bánh xe 8x8	Nguyễn Thanh Quang , Nguyễn Minh Tân, Cù Xuân Phong	2673
279.	Khảo sát khả năng quay vòng ổn định của xe xích quân sự ở tốc độ cao trên đường cát	Hoàng Văn Dinh	2680
280.	Nghiên cứu thiết kế hệ thống chữa cháy tự động cho xe thiết giáp bánh lốp	Cù Xuân Phong , Bùi Hà Anh	2690
281.	Nghiên cứu động lực học hệ truyền động thủy lực dẫn động hệ thống nâng tay kẹp của thiết bị hỗ trợ lắp ráp cấu kiện vòm trong thi công đường hầm	Nguyễn Hoa Hợi	2697

ISBN: 978-632-06-0210-0

TT	Tên bài báo	Tác giả	Trang
282.	Nghiên cứu tối ưu vị trí công tác của thiết bị chuyên dùng phục vụ công tác kiểm tra, sửa chữa bề mặt dưới của các công trình cầu	Bùi Văn Luân, Lê Văn Dưỡng, Bùi Văn Trầm	2709
283.	Đánh giá hiệu quả mô hình Kriging trong phân tích đa tham số kết cấu thép mặt cắt dầm chủ cầu trục hai dầm chế tạo tại Việt Nam	Hoàng Văn Dựng	2717
284.	Nghiên cứu động lực học hệ thống thu hồi đốt phao khơi của bộ cầu PMP trong giai đoạn tời đốt	Khổng Văn Công	2726
285.	Nghiên cứu Xây dựng mô hình mô phỏng mức độ mòn cặp piston - xi lanh đến các thông số hoạt động của xi lanh thủy lực bằng phần mềm MATLAB/Simulink	Lê Đình Trường	2736
286.	Nghiên cứu đặc tính thủy động lực học của xe thiết giáp chở quân chiến đấu trong quá trình bơi sử dụng phương pháp số.	Lương Trần Anh Duy, Hà Trường Sang, Nguyễn Duy Đạt, Nghiêm Văn Long	2744
287.	Nghiên cứu tính toán đặc tính thủy động tác dụng lên cơ cấu lái của bơm phụt MJP-330 trên xe thiết giáp chở quân chiến đấu bằng phương pháp số	Nghiêm Văn Long, Nguyễn Văn Toàn, Nguyễn Duy Đạt, Lương Trần Anh Duy	2754
288.	Nghiên cứu động lực học cổng trục công xôn phục vụ lao lắp dầm cầu trong trường hợp hạ hàng.	Ngô Minh Tiến, Nguyễn Ngọc Trung, Nguyễn Viết Tân	2764
289.	Nghiên cứu ứng dụng hệ thống vây giảm lắc chủ động giảm thiểu dao động của xe bơi tự hành	Nguyễn Thế Anh, Nguyễn Viết Tân, Nguyễn Minh Kha, Lê Trọng Tuấn	2778
290.	Nghiên cứu động lực học bơm pít tông hướng trục sử dụng mô hình mạng nơ ron	Nguyễn Văn Khánh, Đoàn Văn Tú, Trịnh Văn Hải	2790
291.	Nghiên cứu sức cản thủy động của cầu phao PMP	Nguyễn Quỳnh Quyên, PGS.TS Lê Văn Dưỡng	2802
292.	Nghiên cứu động lực học ngược bộ công tác máy bốc xúc phục vụ thi công đường hầm khẩu độ nhỏ	Nguyễn Quốc Sáng, Nguyễn Viết Tân	2810
293.	Mô hình hóa và phân tích động lực học quá độ của hệ thống thủy lực dẫn động cơ cấu bơi xe thiết giáp chở quân bánh lốp ở giai đoạn xuống bến	Nguyễn Tiến Vĩ, Nguyễn Duy Đạt, Lê Quang Đạt	2819
294.	Nguyên cứu chẩn đoán tình trạng kỹ thuật bộ pin Lithium-ion trên xe điện bằng phương pháp tích hợp định lượng nhanh chính xác	Đào Nguyễn Thắng, Vũ Ngọc Tuấn	2829
295.	Ứng dụng trí tuệ nhân tạo đánh giá tình trạng kỹ thuật của một số cảm biến trên ô tô thông qua trích xuất dữ liệu từ cổng OBD-II	Đinh Văn Tiến, Vũ Ngọc Tuấn	2839
296.	Khảo sát ảnh hưởng của quá trình phanh đến dao động ô tô KAMAZ-43253	Huỳnh Đặng Nhứt Rồng, Vũ Mạnh Dũng	2851

ISBN: 978-632-06-0210-0

TT	Tên bài báo	Tác giả	Trang
297.	Xây dựng mô hình mô phỏng động lực học dẫn động điều khiển hệ thống phanh thủy lực có trợ lực chân không	Kiều Hữu Tài, Nguyễn Văn Trà	2865
298.	Ứng dụng mô hình Markov nghiên cứu sự biến đổi tình trạng kỹ thuật hệ thống phanh trên xe KAMAZ-53212	Nguyễn Trung Tiến, Nguyễn Văn Dũng	2873
299.	Nghiên cứu các yếu tố ảnh hưởng đến độ ổn định trong quá trình quay vòng ô tô 2 cầu	Tăng Minh Tài, Nguyễn Đăng Quý	2882
300.	Nghiên cứu thiết kế cơ cấu tăng ma sát trong dạng đĩa ma sát cho bộ vi sai cầu xe chủ động	Trần Đức Hùng, Nguyễn Trường Sinh	2892
301.	Xây dựng mô hình mô phỏng động lực học chuyển động quay vòng của xe quân sự nhiều cầu dẫn hướng	Vũ Ngọc Vương, Nguyễn Trường Sinh	2902
302.	Nghiên cứu động lực học hệ thống treo cân bằng phụ thuộc ở ô tô quân sự 3 cầu	Vũ Thế Bảo, Vũ Đức Lập	2915
303.	Mô phỏng động lực học chuyển động tăng tốc của xe bọc thép bánh lốp BTR - 60PB bằng Simscape Driveline	Trần Đức Lương, Nguyễn Văn Trà	2925
304.	Xây dựng mô hình mô phỏng động lực học hộp số tự động phục vụ chẩn đoán	Phạm Anh Tuấn, Nguyễn Văn Trà , Vũ Ngọc Tuấn, Nguyễn Văn Tứ	2937
305.	Nghiên cứu, mô phỏng, đánh giá công thái học thao tác của lái xe ô tô quân sự ứng dụng thị giác máy tính	Trần Thành Lam, Lê Kỳ Nam, Phạm Quốc Hoàng,	2944
306.	Ảnh hưởng tỷ số nén đến trạng thái ứng suất nhiệt của piston – xi lanh	Đỗ Quang Đức, Nguyễn Văn Dương	2954
307.	Nghiên cứu đánh giá hiệu quả xử lí soot của bộ scrubber lắp trên động cơ diesel Kamaz 740 khi sử dụng hỗn hợp nhiên liệu chứa ôxy	Nguyễn Bá Hiếu, Phạm Xuân Phương	2962
308.	Xây dựng mô hình mô phỏng hệ thống phun nhiên liệu diesel kiểu CommonRail trong MATLAB/Simulink	Nguyễn Tất Trọng, Nguyễn Hoàng Vũ	2972
309.	Nghiên cứu lựa chọn phương án cải tiến hệ thống làm mát động cơ diesel 3D12 khi cường hoá động cơ bằng tăng áp khí nạp	Nguyễn Văn Năm, Nguyễn Văn Dương	2981
310.	Nghiên cứu xây dựng phương pháp xác định đồng thời độ lọt khí của các buồng đốt động cơ Kamaz-740	Phan Đoàn Việt Đức, Phạm Xuân Phương	2991
311.	Nghiên cứu ảnh hưởng của góc tấn dòng khí đầu vào tới thông số tầng máy nén đầu tiên động cơ TB3-117	Trần Duy Khánh, Lê Tiến Dương	3001
312.	Aeroacoustic Characteristic of an Electric Propulsion Fan with Sound-Absorbing Duct	Chu Hoang Quan, Nguyen Binh Nguyen, Vu Viet Dung, Nguyen Trung Kien, Dinh Cong Truong	3012
313.	Nghiên cứu mô phỏng số về động lực học phân rã sơ cấp của tia phun trong ống hóa hơi động cơ tua bin khí cỡ nhỏ	Nguyễn Hữu Hà, Nguyễn Quốc Quân, Phạm Xuân Phương	3022

ISBN: 978-632-06-0210-0

TT	Tên bài báo	Tác giả	Trang
314.	Nghiên cứu thực nghiệm đặc tính lưu lượng vòi phun khí trong chế độ phun xung ngắn	Nguyễn Văn Hùng, Lương Đình Thi, Vũ Đức Mạnh	3030
315.	Nghiên cứu, xây dựng mô hình trạng thái nhiệt của cabin xe sử dụng mô phỏng CFD	Chu Hữu Sỹ, Vũ Đức Mạnh, Phùng Văn Được	3037
316.	Ước lượng tốc độ và góc trượt của uav khi sử dụng hai ống pitot đối xứng và thuật toán EKF và LSTM	Phạm Minh Kha	3046
317.	Một số vấn đề về ứng dụng AI trong việc giải quyết các bài toán phòng không đối với hệ thống tự động hóa chỉ huy cấp chiến thuật	Nguyễn Việt Hoài Nam	3055
318.	BiTLES: A Bio-Inspired Local Temporal Jitter Algorithm for Efficient Spiking Neural Network Training	Nguyễn Văn Tuấn	3063
319.	Tối ưu quỹ đạo của tên lửa đạn đạo trên cơ sở điều khiển vector lực đẩy của tên lửa	Bùi Quang Minh	3070
320.	Tổng hợp luật dẫn đa chỉ tiêu cho tên lửa đội hạm trong giai đoạn tự dẫn	Bùi Quốc Dũng	3077
321.	Nghiên cứu cải thiện độ ổn định và độ chính xác của con quay sợi quang dựa trên kỹ thuật xử lý tín hiệu tiên tiến	Bùi Văn Tiến	3085
322.	Nghiên cứu các phương pháp điều khiển uav bầy đàn dạng quadtotor	Đặng Quốc Đạt	3094
323.	Tổng hợp dẫn và điều khiển thời gian cố định dựa trên bộ quan sát nhiễu cho tên lửa đánh chặn mục tiêu cơ động	Đinh Hồng Toàn	3105
324.	Thiết kế quỹ đạo thiết bị bay tránh mối đe dọa sử dụng Q-learning	Hồ Quang Huy	3116
325.	Nghiên cứu và tổng hợp bộ điều khiển thích nghi cho hệ truyền động lái đa kênh trên máy bay	Lê Duy Khiên	3124
326.	Nâng cao chất lượng nhận dạng mục tiêu bay trên cơ sở tối ưu tham số mô hình yolo sử dụng phương pháp bayes trong pytorch	Lê Thanh Phong	3134
327.	Nghiên cứu ảnh hưởng của gió tới chất lượng điều khiển Quadrotor	Lê Thế Trung	3148
328.	Nghiên cứu tối ưu góc mạn ban đầu của ngư lôi ở giai đoạn tự dẫn trong chế độ bắn đuổi nhằm nâng cao khả năng tiêu diệt mục tiêu	Nguyễn Bá Tước	3159
329.	Thiết kế bộ điều khiển phi tuyến dựa trên phương pháp nghịch đảo động học kết hợp PI và kỹ thuật gain-scheduling cho hệ thống tự lái tên lửa	Nguyễn Đắc Nghĩa	3167
330.	FPGA-Accelerated YOLOv5 for Real-Time Thermal Image Object Detection	Nguyễn Ngọc Anh	3174
331.	Nghiên Cứu Xây Dựng Thuật Toán Phát Hiện Đối Tượng Sử Dụng Mô Hình Học Sâu Ứng Dụng Trong Huấn Luyện Quân Sự	Nguyễn Văn Thịnh	3185
332.	Xây dựng mô hình nhận dạng rối loạn nhịp tim sử dụng mạng học sâu	Vương Trí Tiếp	3193

ISBN: 978-632-06-0210-0

TT	Tên bài báo	Tác giả	Trang
333.	Xây dựng mô hình học sâu nhận dạng cử chỉ tay sử dụng kiến trúc hai nhánh	Vũ Văn Đông	3199
334.	Nghiên cứu ảnh hưởng dữ liệu đặc trưng đến kết quả ước lượng trạng thái SOC của pin xe điện sử dụng kỹ thuật học sâu	Vũ Văn Chiến	3207
335.	Một quy trình học máy đơn giản, hiệu quả để phân biệt parkinson sử dụng dữ liệu thu được từ đồng hồ thông minh	Vũ Quân	3217
336.	Robust Control Design for a Three-Axis Gimbal System Based on Adaptive Sliding Mode Control Law	Quách Văn Tài	3228
337.	Phát hiện thủy lôi dưới nước bằng phương pháp TDOA sử dụng tín hiệu sonar từ tàu mẹ và robot ngầm di chuyển đa vị trí	Phạm Ngọc Hòa	3238
338.	Combination of compressed sensing and a deep Learning based model for multiclass MRI brain tumor classification	Pham Đức Khánh	3247
339.	Nghiên cứu tính toán và chế tạo thiết bị tự nổi đóng vai trò mìn huấn luyện sử dụng cho Robot Pluto Plus	Nguyễn Xuân Tiến	3257
340.	Nghiên cứu cải thiện bộ lọc định hướng để loại bỏ nhiễu phản xạ ảnh hưởng đến sự lan truyền sóng trượt trong mô	Nguyễn Vũ Nghĩa	3264
341.	Beyond Attention-Centric Models: State Space Models and Mamba for Perception in Autonomous Unmanned Surface Vehicles	Nguyễn Việt Tiệp	3273
342.	Development of data consistency integrated deep learning model for reconstructing undersampled MRI k-space	Bùi Thị Hoa	3284
343.	Tổng hợp luật điều khiển hợp thể thích nghi cho hệ thống treo chủ động khi có nhiễu	Đào Trọng Dũng	3293
344.	Prescribed-Time Gain-Scheduled Computed-Torque Control with Residual Disturbance Observation for Robot Manipulators	Hoàng Đức Long	3303
345.	Điều khiển tư thế quadrotor dựa trên bộ điều khiển chủ động loại nhiễu cải tiến	Kiều Bích Sơn	3313
346.	Phương pháp điều hướng chủ động cho robot di động trong môi trường có người di chuyển	Nghiêm Hoàng Nam	3325
347.	Xây dựng mô hình nhận dạng rung nhĩ nhẹ cho thiết bị đeo sử dụng thuật toán stacking	Vương Trí Tiếp	3336

ISBN: 978-632-06-0210-0

The variant of inertial algorithm for solving nonmonotone variational inequality problems

Bien Thanh Tuyen[1*], Nguyen Thi Thanh Ha[1], Nguyen Hong Nam[1]

[1] *Institute of Information and Communication Technology, Le Quy Don Technical University*

thanhtuyenb@lqdtu.edu.vn; Tel: 0979745828

Abstract

We introduce an inertial algorithm for solving variational inequality problems, in the case the cost mapping is Lipschitz continuous and nonmonotone. Unlike existing approaches that require a linesearch procedure, the presented algorithm utilizes an adaptive step size rule instead. Under appropriate assumptions, the algorithm is rigorously proven to converge to a solution. A numerical example is presented to illustrate the efficiency and practical performance of our method.

Key words: *Variational inequality; nonmonotone; inertial algorithm; projection; Lipschitz.*

1. Introduction

Let $\mathbb{R}^n$ is a Euclidean space with an inner product $\langle\cdot,\cdot\rangle$ and the induced norm $\|\cdot\|$. Suppose that C is a nonempty closed convex set in $\mathbb{R}^n$ and F be a continuous mapping from $\mathbb{R}^n$ into $\mathbb{R}^n$. We consider the following classical variational inequality problem associated with C and F (VIP(C, F) for short):

Find $u \in C$ such that $\langle F(u), y-u\rangle \geq 0, \forall y \in C.$

We denote the solution set of VIP(C, F) by S. The dual variational inequality problem is called Minty variational inequality problem (shortly, MVIP(C, F)), can be stated as follows:

Find $u \in C$ such that $\langle F(y), y-u\rangle \geq 0, \forall y \in C.$

The solution set of MVIP(C, F) is denoted by S_M. Recall that F is strongly monotone on C with modulus $\tau > 0$ if $\langle F(x)-F(y), x-y\rangle \geq \tau \| x-y\|^2, \forall x, y \in C.$ If $\tau = 0$ then it is monotone on C, it is pseudomonotone on C if $x, y \in C$, $\langle F(y), x-y\rangle \geq 0$ then $\langle F(x), x-y\rangle \geq 0,$ and it is Lipschitz continuous with $L > 0$ on C if $\| F(x)-F(y)\| \leq L\| x-y\|, \forall x, y \in C.$ It is well-known that $S_M \subset S$ if F is continuous on C, in addition, if F is pseudomonotone, then $S \subset S_M$. (see [1, Lemma 2.1]). Consequently, if F is pseudomonotone and continuous on C then we have $S = S_M$.

Variational inequality problems are extremely powerful tools for studying various nonlinear problems arising in several branches of applied science. They have been successfully applied in fields such as control theory, game theory, transportation science, economic equilibrium, and engineering.

* Email: thanhtuyenb@lqdtu.edu.vn

One of the main topic in variational inequality theory is to build effective algorithms for finding a solution of VIPs among them the projection methods play an important role [2]. To the best of our knowledge, the first one projection method was introduced by Goldstein [3] as follows:

$$x^1 \in C,\ x^{k+1} = P_C(x^k - \lambda F(x^k)),$$

where P_C is the metric projection onto C. The global convergence of this algorithm is established whenever F is strongly monotone and Lipschitz on C. Recently, some authors have relaxed this condition to strong pseudomonotonicity and Lipschitz on C. Unfortunately, this method, in general, might not be convergent for the monotone variational inequality problems [4]. Hence, Korpelevich [5] proposed the following extragradient algorithm:

$$x^0 \in C,\ y^k = P_C(x^k - \lambda F(x^k)),\ x^{k+1} = P_C(x^k - \lambda F(y^k)),$$

where $\lambda \in \left(0, 1/L\right)$. The sequence $\{x^k\}$ was proved to converge to a solution of VIPs under the main assumption that F is pseudomonotone and Lipschitz on C with explicit constant. However, this method requires two projections onto C and two evaluations of F per iteration. Since projecting onto a general closed convex set C involves solving a minimum distance problem, this can be computationally expensive and may reduce efficiency. To overcome this drawback, Censor [6] presented the subgradient extragradient algorithm:

$$y^k = P_C(x^k - \lambda F(x^k)),\ x^{k+1} = P_{T_k}(x^k - \lambda F(y^k)),$$

where $T_k = \left\{z \in \tilde{}^n : \langle w^k - y^k - \lambda_k F(w^k), x - y^k \rangle\right\}$. Since the second projection to hyperplane can be found in a closed form, this method is more applicable. When F is pseudomonotone and Lipschitz on C (with unknow constant) sequence $\{x^k\}$ generated by this algorithm converge to a solution of VIPs. Nonetheless, the algorithm may fail to converge for nonmonotone variational inequality problems.

In order to achieve a faster convergence rate, Polyak [7] introduced the inertial method, which was originally proposed as an acceleration process for solving smooth convex minimization problems. This method can be viewed as a second order discretization of a time dynamical system. The inertial algorithm is characterized as a multistep iterative procedure, where the next iterate is computed by utilizing two or more previous iterates. By virtue of the advantages of this method, a vast number of researchers have been investigating inertial type algorithms to solve various problems. Especially, several algorithms have been proposed to solve extended variational inequality problems over different function classes, such as strongly monotone or strongly pseudomonotone [8], monotone [9], pseudomonotone [10].

Motivated by this fact, in this paper, we introduce new inertial algorithm for finding a solution of VIPs in which the cost mapping F is nonmonotone and Lipschitz continuous with unknow constant. The rest of paper is organized as follows. The next section, we present the proposed algorithm and prove their convergence. The third section is dedicated to presenting numerical example to illustrate the effectiveness of the proposed algorithms.

2. Main results

Let C be a nonempty subset of Euclidean spaces $\sim^n$. Let $d(x,C)$ denote the Euclidean distance from a vector x to C, i.e., $d(x,C) := \inf\{\| x-y \|, y \in C\}$. If C is a nonempty and closed set, then it can be seen that $d(x,C) := \min\{\| x-y \|, y \in C\}$.

In the rest of this paper, we consider the set C is a nonempty, closed and convex set. By P_C we denote the metric projection operator on C, that is $P_C(x) \in C : \| x - P_C(x) \| \leq \| y - x \|, \forall y \in C.$

The following lemmas are well-known for the projection operator onto a closed convex set.

Lemma 1. [11] Suppose that C is a nonempty closed convex subset in $\sim^n$. Then, the following statements hold:

(i) $z = P_C(x)$ if and only if $\langle x-z, y-z \rangle \leq 0, \forall y \in C;$

(ii) $\|P_C(x)-y\|^2 \leq \|x-y\|^2 - \|P_C(x)-x\|^2, \forall x, y \in C.$

Lemma 2. [12] Let $\{a_k\}_{k=0}^{\infty}$ and $\{b_k\}_{k=0}^{\infty}$ be the nonnegative real number sequences satisfying $\sum_{i=1}^{\infty} b_i < \infty$ and $a_{i+1} \leq a_i + b_i, \forall i.$ Then, the sequence $\{a_k\}_{k=0}^{\infty}$ is convergent.

In the sequel, we need the following blanket assumptions:

(A1) F is Lipschitz continuous on $\sim^n$ and $S_M \neq \varnothing$;

(A2) If $y \in C$, and $x \in S_M$ satisfy $\langle F(y), y-x \rangle = 0$, then $y \in S_M$.

Now, we assume that F is a Lipschitz continuous mapping from $\sim^n$ with an unknown Lipschitz constant and C is a nonempty, closed, convex subset of $\sim^n$. Moreover, F is not need required to satisfy any monotonicity assumption on C. The following algorithm gives us a way to find a solution of $\mathrm{VIP}(C,F)$.

Algorithm 1.

Initialization. Choose $\theta \in [0,1), \lambda_1 > 0$ and $\delta \in (0,1)$. Choose positive sequence $\{\mu_k\}_{k=1}^{\infty}$ such that $\sum_{k=1}^{\infty} \mu_k < \infty$. Take $x^0, x^1 \in C$ and $k = 0$.

Iteration $k\,(k = 1, 2, \ldots)$. Having x^{k-1}, x^k do the following Steps.

Step 1. Compute $w^k := x^k + \theta_k (x^k - x^{k-1})$, where $0 \leq \theta_k \leq \bar{\theta}_k$, and

$$\bar{\theta}_k = \begin{cases} \min\left\{\theta, \dfrac{\mu_k}{\| x^k - x^{k-1} \|}\right\}, & \text{if } x^k \neq x^{k-1}; \\ \theta, & \text{otherwise.} \end{cases}$$

Step 2. Compute $y^k := P_C(w^k - \lambda_k F(w^k))$. If $y^k = w^k$, stop and $w^k \in S$, otherwise, go to Step 3.

Step 3. Compute $x^{k+1} := P_{H_k}(w^k - \lambda_k F(y^k))$, where

$$H_k = \{z \in \tilde{}^n : \langle w^k - y^k - \lambda_k F(w^k), x - y^k \rangle\}.$$

Step 4. Update the new step size λ_{k+1} by

$$\lambda_{k+1} = \begin{cases} \min\left\{\lambda_k, \dfrac{\delta \| w^k - y^k \|}{\| F(w^k) - F(y^k) \|}\right\}, & \text{if } \| F(w^k) - F(y^k) \| \neq 0; \\ \lambda_k, & \text{otherwise.} \end{cases} \quad (1)$$

or

$$\lambda_{k+1} = \begin{cases} \delta\lambda_k, & \text{if } \lambda_k \| F(w^k) - F(y^k) \| > \delta \| w^k - y^k \|; \\ \lambda_k, & \text{otherwise.} \end{cases} \quad (2)$$

and go to Iteration k with k is replaced by $k+1$.

Remark 1.

(i) Since H_k is a half-space, the explicit form of sequence $\{x^{k+1}\}$ in Step 3 is determined as follows:

$$x^{k+1} = \begin{cases} w^k - \lambda_k F(w^k), \text{ if } w^k - \lambda_k F(w^k) \in H_k; \\ w^k - \lambda_k F(w^k) - \dfrac{\langle z^k, w^k - \lambda_k F(y^k) - y^k \rangle}{\|z^k\|^2} . z^k, \text{ otherwise,} \end{cases}$$

where $z^k = w^k - y^k - \lambda_k F(w^k)$.

(ii) By definition of H_k, it is easy to see that $S_M \subset H_k, \forall k$.

Lemma 3. The limit of the sequence $\{\lambda_k\}$ exists and is positive.

The proof of this lemma is similar to the proof of Lemma 3.1 in [13] so we omit it. The following theorem gives us the convergence of Algorithm 1.

Theorem 1. Let Assumptions $(A1)$-$(A2)$ hold. Then the sequence $\{x^k\}$ is generated by Algorithm 1 converges to a solution of $\text{VIP}(C, F)$.

Proof. For any fixed $p \in S_M$, by Remark (ii), we have $p \in H_k, \forall k$ Since $x^{k+1} \in H_k$, we have

$$\langle w^k - y^k - \lambda_k F(w^k), x^{k+1} - y^k \rangle \leq 0. \quad (3)$$

Then, $\langle y^k - w^k, y^k - x^{k+1} \rangle \leq \lambda_k \langle F(w^k), x^{k+1} - y^k \rangle$. In addition, using Lemma 1 and by definition of x^{k+1}, we have

$$\langle x^{k+1} - w^k, x^{k+1} - p \rangle \leq \lambda_k \langle F(y^k), p - x^{k+1} \rangle. \quad (4)$$

On the other hand, it's obvious that

$$\left\|x^{k+1}-p\right\|^2=\left\|y^k-p\right\|^2-\left\|x^{k+1}-y^k\right\|^2+2\left\langle x^{k+1}-y^k,x^{k+1}-p\right\rangle$$
$$=\left\|w^k-p\right\|^2-\left\|x^{k+1}-y^k\right\|^2-\left\|y^k-w^k\right\|^2+2\left\langle y^k-w^k,y^k-x^{k+1}\right\rangle+2\left\langle x^{k+1}-w^k,x^{k+1}-p\right\rangle$$

Combining this with (3) and (4), we get

$$\left\|x^{k+1}-p\right\|^2\le\left\|w^k-p\right\|^2-\left\|x^{k+1}-y^k\right\|^2-\left\|y^k-w^k\right\|^2+2\lambda_k\left\langle F(w^k),x^{k+1}-y^k\right\rangle+2\lambda_k\left\langle F(y^k),p-x^{k+1}\right\rangle$$
$$=\left\|w^k-p\right\|^2-\left\|x^{k+1}-y^k\right\|^2-\left\|y^k-w^k\right\|^2+2\lambda_k\left\langle F(w^k)-F(y^k),x^{k+1}-y^k\right\rangle+2\lambda_k\left\langle F(y^k),p-x^{k+1}\right\rangle$$

Besides, from (1) and (2), we see that there exists $k_0\in\bullet$ such that

$$\lambda_{k+1}\le\frac{\delta\|\,w^k-y^k\,\|}{\|\,F(w^k)-F(y^k)\,\|},\ \forall k\ge k_0.$$

Hence, for all $k\ge k_0$, we have

$$\left\|x^{k+1}-p\right\|^2\le\left\|w^k-p\right\|^2-\left\|x^{k+1}-y^k\right\|^2-\left\|y^k-w^k\right\|^2$$
$$+2\lambda_k\left\langle F(y^k),p-x^{k+1}\right\rangle+2\frac{\delta\lambda_k}{\lambda_{k+1}}\left\|w^k-y^k\right\|\left\|x^{k+1}-y^k\right\| \tag{5}$$
$$\le\left\|w^k-p\right\|^2-\left(1-\frac{\delta\lambda_k}{\lambda_{k+1}}\right)\left(\left\|x^{k+1}-y^k\right\|^2+\left\|y^k-w^k\right\|^2\right)+2\lambda_k\left\langle F(y^k),p-x^{k+1}\right\rangle,$$

We also have $\left\langle F(y^k),y^k-z\right\rangle\ge 0$ since $y^k\in C,p\in S_M$. This together with (5) obtain that

$$\left\|x^{k+1}-p\right\|^2\le\left\|w^k-p\right\|^2-\left(1-\frac{\delta\lambda_k}{\lambda_{k+1}}\right)\left(\left\|x^{k+1}-y^k\right\|^2+\left\|y^k-w^k\right\|^2\right),\forall k\ge k_0.$$

Moreover, since $\lim_{k\to\infty}\lambda_k=\lambda^*$ does exists and $\lambda^*>0$, we also get $\lim_{k\to\infty}\delta\frac{\lambda_k}{\lambda_{k+1}}=\delta$. This implies that there exists $\delta^*\in(\delta,1),k_1\in\bullet$ such that $\delta\frac{\lambda_k}{\lambda_{k+1}}<\delta^*,\forall k\ge k_1$. Set $k_2=\max\{k_0,k_1\}$. Then,

$$\left\|x^{k+1}-p\right\|^2\le\left\|w^k-p\right\|^2-(1-\delta^*)\left(\left\|x^{k+1}-y^k\right\|^2+\left\|y^k-w^k\right\|^2\right)\le\left\|w^k-p\right\|^2,\forall k\ge k_2. \tag{6}$$

Observer that $\left\|w^k-p\right\|=\left\|x^k+\theta_k(x^k-x^{k-1})-p\right\|\le\left\|x^k-p\right\|+\theta_k\left\|x^k-x^{k-1}\right\|$.

Hence, $\left\|x^{k+1}-p\right\|\le\left\|x^k-p\right\|+\theta_k\left\|x^k-x^{k-1}\right\|$. By Lemma 2, we have $\left\{\left\|x^k-p\right\|\right\}$ is convergent. Thus, we can deduce that $\left\{x^k\right\}$ is bounded. Furthermore, using definition of w^k and the continuity of F, the sequences $\left\{w^k\right\},\left\{F(w^k)\right\},\left\{y^k\right\}$ are also bounded.

Note that $\left\|w^k-p\right\|^2\le\left\|x^k-p\right\|^2+\theta_k{}^2\left\|x^k-x^{k-1}\right\|^2+2\theta_k\left\|x^k-p\right\|\left\|x^k-x^{k-1}\right\|$.

Substituting this result into (6), we get

$$\left\|x^{k+1}-p\right\|^2 \le \left\|x^k-p\right\|^2+\theta_k{}^2\left\|x^k-x^{k-1}\right\|^2+2\theta_k\left\|x^k-p\right\|\cdot\left\|x^k-x^{k-1}\right\|$$
$$-\left(1-\delta^*\right)\left(\left\|x^{k+1}-y^k\right\|^2+\left\|y^k-w^k\right\|^2\right)$$

Thus,

$$\left(1-\delta^*\right)\left(\left\|x^{k+1}-y^k\right\|^2+\left\|y^k-w^k\right\|^2\right)\le\left\|x^k-p\right\|^2-\left\|x^{k+1}-p\right\|^2$$
$$+\theta_k{}^2\left\|x^k-x^{k-1}\right\|^2+2\theta_k\left\|x^k-p\right\|\cdot\left\|x^k-x^{k-1}\right\|$$

Recall that $\left\{\left\|x^k-p\right\|\right\}$ is convergent and $\lim_{k\to\infty}\theta_k\left\|x^k-x^{k-1}\right\|=0,$ from the above inequality, we deduce that

$$\lim_{k\to\infty}\left\|x^{k+1}-y^k\right\|=\lim_{k\to\infty}\left\|y^k-w^k\right\|=0. \qquad (7)$$

Hence, we have $\left\|w^k-x^{k+1}\right\|\le\left\|w^k-y^k\right\|+\left\|y^k-x^{k+1}\right\|\to 0,$ as $k\to\infty.$ Since $\left\{x^k\right\}$ is bounded, then there exists a subsequence $\left\{x^{l_k}\right\}\subset\left\{x^k\right\}$ such that $x^{l_k}\to\bar{x}$ as $k\to\infty.$ From (7), we deduce that $y^{l_k}\to\bar{x}, w^{l_k}\to\bar{x}$ as $k\to\infty.$ For any fixed $p\in S_M,$ by definition of $y^{l_k},$ we have

$$\left\langle y^{l_k}-\left(w^{l_k}-\lambda_{l_k}F(w^{l_k})\right),p-y^{l_k}\right\rangle\ge 0.$$

Putting $k\to\infty,$ since $\lim_{k\to\infty}\lambda_k=\lambda^*$ and $\lim_{k\to\infty}\left\|y^k-w^k\right\|=0$, we get that $\left\langle F\left(\bar{x}\right),p-\bar{x}\right\rangle\ge 0.$ Moreover, because $p\in S_M$ and $\bar{x}\in C,$ we have $\left\langle F\left(\bar{x}\right),\bar{x}-p\right\rangle\ge 0.$ Therefore, $\left\langle F\left(\bar{x}\right),\bar{x}-p\right\rangle=0.$ From Assumption (A2), it follows that $\bar{x}\in S_M.$ Then, $\bar{x}\in S.$ Since $\left\{\left\|x^k-\bar{x}\right\|\right\}$ is convergent, by using the fact that $\bar{x}$ is a cluster point of $\left\{x^k\right\},$ we deduce that

$x^k\to\bar{x}\in S$ as $k\to\infty.$ This proof is complete

3. A numerical example

In this section, we illustrate the proposed algorithm and compare it with Algorithm IPA_L in [14], which was recently introduced for solving nonmonotone VIPs under the Lipschitz continuity assumption without using a line-search procedure. We implement algorithms in MATLAB R2015a running on a Desktop with Intel(R) Core(TM) i7-9700 CPU @ 3.00GHz with 16GB Ram.

We consider the problem $\mathrm{VIP}\left(C,F\right),$ where

$$C=\left\{x\in\mathbb{R}^n:\|x\|\le 1\right\},\qquad F(x)=\left(2-\|x\|\right)x,\ \forall x\in\mathbb{R}^n.$$

It can be seen that F is Lipschitz continuous with Lipschitz constants $L=4.$ In addition, F is pseudomonotone but not monotone on $\mathbb{R}^n$. We can easily compute that $S=S_M=\left\{(0,0,...,0)^T\right\}.$ This means that the condition (A1) holds. Moreover, for

$y \in C,\ x = (0,0,...,0)^T \in S_M$, then $\langle F(y), y - x \rangle = 0$ if $y = (0,0,...,0)^T$. Therefore, the condition (A2) also holds.

The convergence behavior of the two algorithms is compared for different dimensions and various initial points. The parameters are chosen as follows: $\theta = 0.15, \lambda_1 = 0.1, \delta = 0.5,$ $\theta = 0.15, \lambda_1 = 0.1, \delta = 0.5, \mu_k = \frac{1.1}{k^2}, x^{-1} = 0.3x^0.$ in Algorithm 1, $\lambda = 0.2$ in Algorithm IPA_L.

We use the stopping criterion $\| x^k \| \leq 10^{-4}$ for all algorithms. The numerical results are reported in Table 1 and Figure 1 $(n = 100)$.

Table 1. Experiment for Example 1.

n	x^0	**Alg. 1**		**Alg.** IPA_L	
		CPU(s)	**Iter**	**CPU(s)**	**Iter**
50	$x_i = \frac{0.75}{i^2}$	0.0469	41	134.4219	21
100	$x_i = \frac{0.3}{i^2}$	0.0938	36	125.7969	18
200	$(0.8, 0.5, 0.3, 0, ..., 0)^T$	0.0781	18	141.3906	22
500	$(0.5, 0.5, ..., 0.5)^T$	0.1250	10	148.4844	206
1000	$(0.8, 0.8, ..., 0.8)^T$	0.7813	35	164.7813	206

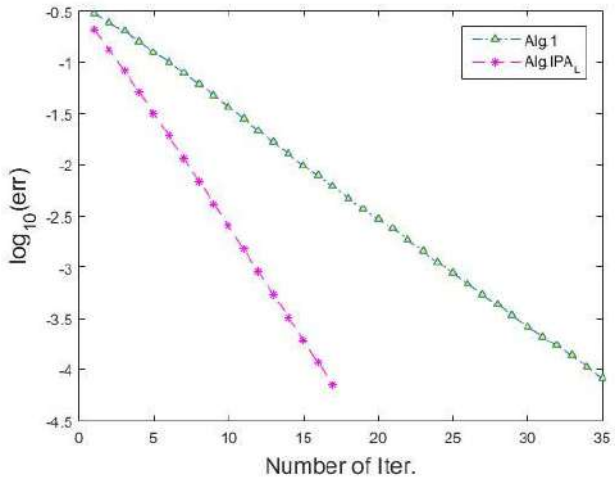


Figure 1. Number of iterations of Algorithms in Example $\left(n = 100,\ x_i = \frac{0.3}{i^2} \right)$.

From Table 1, the numerical results suggest that, for different choices of dimension and initial point, Algorithm 1 shows improved efficiency compared with Algorithm IPA_L in terms of CPU time. This behavior may be attributed to the fact that Algorithm 1 employs an inertial technique and an adaptive step size rule, thereby enhancing the convergence rate.

4. Conclusion

We have presented a novel framework of inertial projection algorithm for solving variational inequality problems. By developing adaptive step size strategies, we provided linesearch free solutions for nonmonotone and Lipschitz continuous mappings. Under standard assumptions, the iterative sequences generated by the algorithms are proved to converge to a solution to the corresponding problems. A numerical example is presented to demonstrate the effectiveness of the proposed algorithms.

References

1. Crespi, G. P., Guerraggio, A., & Rocca, M. (2005). Minty variational inequality and optimization: Scalar and vector case. In *Generalized convexity, generalized monotonicity and applications* (pp. 193–211). New York, NY: Springer-Verlag.
2. Dinh, B. V., & Muu, L. D. (2013). Algorithms for a class of bilevel programs involving pseudomonotone variational inequalities. *Acta Mathematica Vietnamica, 38*, 529–540.
3. Goldstein, A. A. (1964). Convex programming in Hilbert spaces. *Bulletin of the American Mathematical Society, 70*, 709–710.
4. Iusem, A. N., & Svaiter, B. F. (1997). A variant of Korpelevich's method for variational inequalities with a new search strategy. *Optimization, 42*, 309–321.
5. Korpelevich, G. M. (1976). The extragradient method for finding saddle points and other problems. *Ekonomika i Matematicheskie Metody, 12*, 747–756.
6. Censor, Y., Gibali, A., & Reich, S. (2011). The subgradient extragradient method for solving variational inequalities in Hilbert space. *Journal of Optimization Theory and Applications, 148*, 318–335.
7. Polyak, B. T. (1964). Some methods of speeding up the convergence of iteration methods. *USSR Computational Mathematics and Mathematical Physics, 4*, 1–17.
8. Thong, D. V., & Hieu, D. V. (2018). Modified subgradient extragradient method for variational inequality problems. *Numerical Algorithms, 79*, 597–610.
9. Thong, D. V., & Hieu, D. V. (2018). Inertial extragradient algorithms for strongly pseudomonotone variational inequalities. *Journal of Computational and Applied Mathematics, 341*, 80–98.
10. Ming, T., & Gang, X. (2020). Inertial modified Tseng's extragradient algorithms for solving monotone variational inequalities and fixed point problems. *Journal of Nonlinear Functional Analysis*, 2020, Article ID 35. https://doi.org/10.23952/jnfa.2020.35
11. Zarantonello, E. H. (1971). Projections on convex sets in Hilbert space and spectral theory I: Projections on convex sets. In *Contributions to nonlinear functional analysis* (pp. 237–424). New York, NY: Academic Press.
12. Xu, H.-K. (2002). Iterative algorithms for nonlinear operators. *Journal of the London Mathematical Society, 66*(1), 240–256.
13. Yang, J., & Liu, H. (2019). Strong convergence result for solving monotone variational inequalities in Hilbert space. *Numerical Algorithms, 80*(3), 741–752.
14. Ye, M., & He, Y. (2022). An infeasible projection type algorithm for nonmonotone variational inequalities. *Numerical Algorithms, 89*(4), 1723–1742.

Một biến thể của thuật toán quán tính giải các bài toán bất đẳng thức biến phân không đơn điệu

Tóm tắt

Chúng tôi giới thiệu một thuật toán quán tính giải các bài toán bất đẳng thức biến phân, trong trường hợp ánh xạ liên tục Lipschitz và không đơn điệu. Thay vì sử dụng phương pháp tìm kiếm tia (linesearch) như các phương pháp hiện có, thuật toán được đề xuất sử dụng quy tắc lựa chọn độ dài bước thích nghi. Dưới các giả thiết thích hợp, dãy lặp sinh bởi thuật toán được chứng minh hội tụ đến nghiệm của bài toán bất đẳng thức biến phân. Ngoài ra, chúng tôi cũng đưa ra một ví dụ số nhằm minh họa hiệu quả cũng như khả năng ứng dụng thực tiễn của phương pháp này.

Từ khóa: Bất đẳng thức biến phân; không đơn điệu; thuật toán quán tính; phép chiếu; liên tục Lipschitz.

Existence, uniqueness and asymptotic separation of solutions for Caputo-Hadamard stochastic fractional differential equations

Tran Thi Phuong Lan [1*], Vu Thi Thuy [2], Pham Tien Dung [1], Bui Quoc Hung [1]

[1]*Institute of Information and Communication Technology, Le Quy Don Technical University*

[2]*Faculty of Information Technology, Hanoi University of Civil Engineering*

Email; Tel: tplan1609@gmail.com; 0981348970

Abstract

We first investigate the existence and uniqueness of solutions for Caputo-Hadamard stochastic fractional differential equations (CHSFDE) of order $\eta \in \left(\frac{1}{2}, 1\right)$ in L^p spaces $(p \geq 2)$ under local Lipschitz condition and linear growth bound. To achieve this, a truncation function technique is employed in conjunction with the properties of stopping times, generalized Grönwall inequalities, and the Borel-Cantelli lemma. Furthermore, we establish the asymptotic separation between distinct solutions. The qualitative analysis throughout this work relies heavily on the properties of the logarithmic kernel, complemented by the rigorous application of Minkowski, Hölder, and Burkholder-Davis-Gundy (BDG) inequalities.

Keyword: *Caputo-Hadamard stochastic fractional differential equations, Existence, Uniqueness, Global Lipschitz condition, Local Lipschitz condition, Linear growth bound, Asymptotic separation between solutions*

1. Introduction

Fractional-order derivatives have gained prominence due to their unique ability to model dynamical systems with memory and hereditary properties. These mathematical tools are essential for describing complex phenomena such as viscoelasticity, anomalous diffusion, chaotic dynamics, and contaminant transport in groundwater [1, 2, 3, 4].

Among existing frameworks, the Caputo fractional derivative is preferred for practical modeling; unlike the Riemann-Liouville approach, it allows for initial conditions with clear physical interpretations [5]. This advantage has spurred extensive research into the wellposedness, regularity, and asymptotic behavior of Caputo stochastic fractional differential equations [6, 7, 8, 9, 10, 11].

Despite recent advancements, certain 'ultraslow' diffusion processes exhibit a logarithmic mean-square displacement, making them better characterized by Caputo-Hadamard fractional operators with logarithmic kernels [12, 13]. However, qualitative results for CHSFDE remain relatively sparse. Early studies established the well-posedness of CHSFDE using Picard iteration and generalized Grönwall inequalities, alongside moment estimates via logarithmic Mittag-Leffler functions and L^2-regularity [14]. A variation of constant formula for CHSFDE is investigated in [15]. Most recently, the work in [16] extended the well-posedness and regularity results of CHSFDE to L^p spaces with $p \geq 2$ under a global Lipschitz condition. In many practical applications, however, dynamical systems often involve coefficients satisfying

* Email: tplan1609@gmail.com

only local Lipschitz constraints. Consequently, extending the study of existence and uniqueness of solutions to these more relaxed conditions is of significant interest. To bypass the limitations imposed by global Lipschitz assumptions, we incorporate a linear growth bound and employ a truncation function technique integrated with the Borel-Cantelli lemma and a sequence of stopping times. These analytical tools rigorously ensure the existence and uniqueness of solutions in L^p spaces ($p \geq 2$). Regarding the asymptotic separation of solutions, it was demonstrated in [17] that solutions to deterministic Caputo-Hadamard fractional differential equations (CHFDE) exhibit logarithmic decay, a result derived from the asymptotic expansions of the Mittag-Leffler function. For the stochastic counterpart, to the best of our knowledge, this problem remains unexplored. By strategically adapting and refining the techniques from [18], we successfully establish the first results on the asymptotic separation for CHSFDE, thereby filling a significant gap in the literature.

The remainder of this paper is organized as follows. Section 2 introduces the fundamental problem and formulates the underlying assumptions. In Section 3, we establish the asymptotic separation between solutions under the global Lipschitz condition. Section 4 is devoted to investigating the existence and uniqueness of solutions under a local Lipschitz condition and a linear growth bound.

2. Setting a problem and assumptions

Consider CHSFDE of order $\eta \in \left(\frac{1}{2}, 1\right)$ of the following form

$$\begin{cases} {}^{CH}D^{\eta}_{t_0+}Y(t) = f(t,Y(t)) + g(t,Y(t))\dfrac{dB_t}{dt}, \quad 0 < t_0 < t \leq T \\ Y(t_0) = \xi_1. \end{cases} \tag{2.1}$$

Here, T is taken as arbitrary, $f, g : [t_0, T] \times R^n \to R^n$ denote measurable functions, $(B_t)_{t\in[t_0,T]}$ is a Brownian motion defined on the filtered probability space $(\Omega, \mathrm{B}, \mathrm{B} := \{\mathrm{B}_t\}_{t\in[t_0,T]}, \mathrm{P})$. For any $p \geq 2$, we define $\mathfrak{F}_t^p$ as the Banach space $L^p(\Omega, \mathrm{B}_t, \mathrm{P})$ consisting of all R^n -valued, $\mathcal{B}_t$ -measurable, and p^{th} integrable functions $f = (f_1, f_2, \ldots, f_n)^T : \Omega \to R^n$. This space is equipped with the norm: $\| f \|_p := \left(\sum_{i=1}^{n} \mathrm{E}[| f_i |^p] \right)^{1/p}$.

A measurable process $Y : [t_0, T] \to L^p(\Omega, \mathrm{B}, \mathrm{P})$ is said to be B -adapted if the condition $Y(t) \in \mathfrak{F}_t^p$ holds for every $t \in [t_0, T]$. Throughout this work, we consider R^n under the p - norm and assume the following hypotheses regarding the drift and diffusion coefficients f and g from (2.1):

(H1-G) For all $y, y_1 \in R^n$ and $\tau \in [t_0, T]$, there exists $L > 0$ such that

$$| f(\tau, y) - f(\tau, y_1) | + | g(\tau, y) - g(\tau, y_1) | \leq L | y - y_1 |.$$

(H1-L) For each $\ell \in \mathbb{N}^*$, there exists $L_\ell > 0$ that for all $\tau \in [t_0, T]$, $y_1, y_2 \in R^n$

$$| f(\tau, y_1) - f(\tau, y_2) | + | g(\tau, y_1) - g(\tau, y_2) | \leq L_\ell | y_1 - y_2 |.$$

(H2) There exists $C>0$ such that: $\operatorname*{esssup}_{\tau\in[t_0,T]}|f(\tau,0)|<C, \quad \operatorname*{esssup}_{\tau\in[t_0,T]}|g(\tau,0)|<C.$

(H3) For all $y, y_1 \in \mathbb{R}^n$ and $\tau\in[t_0,T]$, there exists $C_1>0$ such that

$|f(\tau,y)|+|g(\tau,y)|\leq C_1(1+|y|).$

Under the hypotheses (H1-G) and (H2), the equation (2.1) with the initial condition $Y(t_0)=\xi_1\in\mathfrak{F}_t^p$, admits a unique solution $Y(t,\xi_1)$ on $[t_0,T]$. Moreover, it is equivalent to the following integral equation:

$$Y(t)=\xi_1+\frac{1}{\Gamma(\eta)}\int_{t_0}^{t}\left(\ln\frac{t}{\tau}\right)^{\eta-1}f(\tau,Y(\tau))\frac{d\tau}{\tau}+\frac{1}{\Gamma(\eta)}\int_{t_0}^{t}\left(\ln\frac{t}{\tau}\right)^{\eta-1}g(\tau,Y(\tau))\frac{dB_\tau}{\tau}, \quad (2.2)$$

see [14] in more detail.

3. Asymptotic separation between solutions

This section is devoted to investigating the asymptotic separation between distinct solutions. More precisely, we show that the rate of asymptotic separation between any two distinct solutions of (2.1) cannot decay faster than $C_2(\ln t)^{-\eta}$ in the following theorem.

Theorem 3.1. *Suppose that conditions* (H1-G) *and* (H2) *are satisfied. Let* $\xi_1,\xi_2\in\mathfrak{F}_{t_0}^p$ *with* $\xi_1\neq\xi_2$. *Then, for any* $\varepsilon>0$ *and* $2\leq p<\dfrac{2}{2\eta-1}$

$$\limsup_{t\to+\infty}(\ln t)^{\eta+\varepsilon}\,\|\,Y(t,\xi_1)-Y(t,\xi_2)\,\|_p=+\infty. \quad (3.1)$$

Proof. Assume, for the sake of contradiction, that $\|\,Y(t,\xi_1)-Y(t,\xi_2)\,\|_p=o\left((\ln t)^{-\eta}\right)$ as $t\to+\infty$. Then, for each $\varepsilon>0$, there exists $T_1>1$ such that:

$$\|\,Y(t,\xi_1)-Y(t,\xi_2)\,\|_p\leq\varepsilon(\ln t)^{-\eta} \quad \text{for all} \quad t>T_1. \quad (3.2)$$

Since $Y(t,\xi_1)$ and $Y(t,\xi_2)$ are the solutions of (2.1). Thus, by using the integral form (2.2) and applying the Minkowski inequality for norm, we achieve

$$\|\,\xi_1-\xi_2\,\|_p\leq\|\,Y(t,\xi_1)-Y(t,\xi_2)\,\|_p+\left\|\frac{1}{\Gamma(\eta)}\int_{t_0}^{t}\left(\ln\frac{t}{\tau}\right)^{\eta-1}\left(f(\tau,Y(\tau,\xi_1))-f(\tau,Y(\tau,\xi_2))\right)\frac{d\tau}{\tau}\right\|_p$$

$$+\left\|\frac{1}{\Gamma(\eta)}\int_{t_0}^{T_1}\left(\ln\frac{t}{\tau}\right)^{\eta-1}\left(g(\tau,Y(\tau,\xi_1))-g(\tau,Y(\tau,\xi_2))\right)\frac{dB_\tau}{\tau}\right\|_p$$

$$+\left\|\frac{1}{\Gamma(\eta)}\int_{T_1}^{t}\left(\ln\frac{t}{\tau}\right)^{\eta-1}\left(g(\tau,Y(\tau,\xi_1))-g(\tau,Y(\tau,\xi_2))\right)\frac{dB_\tau}{\tau}\right\|_p, \ \forall t>0.$$

By assumption (H1-G) and the BDG inequality, we receive

$$\| \xi_1 - \xi_2 \|_p \le \| Y(t,\xi_1) - Y(t,\xi_2) \|_p + \frac{L}{\Gamma(\eta)} \int_{t_0}^{t} \left(\ln \frac{t}{\tau} \right)^{\eta-1} \left\| Y(\tau,\xi_1) - Y(\tau,\xi_2) \right\|_p \frac{d\tau}{\tau}$$

$$+ \frac{D_p L}{\frac{1}{t_0^2}\Gamma(\eta)} \left(\left(\frac{(\ln \frac{T_1}{t_0})^{2\eta-1}}{2\eta-1} \right)^{(p-2)/2} \int_{t_0}^{T_1} \left(\ln \frac{t}{\tau} \right)^{2\eta-2} \| Y(\tau,\xi_1)) - Y(\tau,\xi_2) \|_p^p \frac{d\tau}{\tau} \right)^{1/p}$$

$$+ \frac{D_p L}{\frac{1}{t_0^2}\Gamma(\eta)} \left(\left(\frac{(\ln \frac{t}{T_1})^{2\eta-1}}{2\eta-1} \right)^{(p-2)/2} \int_{T_1}^{t} \left(\ln \frac{t}{\tau} \right)^{2\eta-2} \| Y(\tau,\xi_1)) - Y(\tau,\xi_2) \|_p^p \frac{d\tau}{\tau} \right)^{1/p}. \quad (3.3)$$

Moreover, condition (3.2) leads to

$$\limsup_{t \to +\infty} \| Y(t,\xi_1) - Y(t,\xi_2) \|_p = 0. \quad (3.4)$$

Next, using (3.2) and the monotonically increasing nature of function $h(\tau) = (\ln t - \ln \tau)^{\eta-1}$, we have

$$\int_{t_0}^{t} \left(\ln \frac{t}{\tau} \right)^{\eta-1} \left\| Y(\tau,\xi_1) - Y(\tau,\xi_2) \right\|_p \frac{d\tau}{\tau}$$

$$\le \int_{t_0}^{T_1} \left(\ln \frac{t}{\tau} \right)^{\eta-1} C_3 \frac{d\tau}{\tau} + \int_{T_1}^{t} \left(\ln \frac{t}{\tau} \right)^{\eta-1} \varepsilon (\ln \tau)^{-\eta} \frac{d\tau}{\tau}$$

$$\le C_3 (\ln T_1 - \ln t_0) \left(\ln \frac{t}{T_1} \right)^{\eta-1} + \varepsilon \int_{\frac{\ln T_1}{\ln t}}^{1} (1-\tau)^{\eta-1} \tau^{-\eta} \, d\tau$$

$$\le C_3 (\ln T_1 - \ln t_0) \left(\ln \frac{t}{T_1} \right)^{\eta-1} + \varepsilon B(\eta, 1-\eta). \quad (3.5)$$

here $C_3 = \operatorname*{ess\,sup}_{t \in [t_0, T_1]} \| Y(t,\xi_1) - Y(t,\xi_2) \|_p$ and $B(\eta, 1-\eta)$ denotes the Beta function. Similarly,

$$\int_{t_0}^{T_1} \left(\ln \frac{t}{\tau} \right)^{2\eta-2} \| Y(\tau,\xi_1)) - Y(\tau,\xi_2) \|_p^p \frac{d\tau}{\tau} \le C_3^p \left(\ln \frac{t}{T_1} \right)^{2\eta-2} \int_{t_0}^{T_1} \frac{d\tau}{\tau}$$

$$= C_3 (\ln T_1 - \ln t_0) \left(\ln \frac{t}{T_1} \right)^{2\eta-2}. \quad (3.6)$$

From condition (3.2), we have

$$\int_{T_1}^{t} \left(\ln \frac{t}{\tau} \right)^{2\eta-2} \| Y(\tau,\xi_1)) - Y(\tau,\xi_2) \|_p^p \frac{d\tau}{\tau}$$

$$\le \int_{T_1}^{t} \left(\ln \frac{t}{\tau} \right)^{2\eta-2} \varepsilon^p (\ln \tau)^{-\eta p} \frac{d\tau}{\tau} \le \varepsilon^p \int_{\ln T_1}^{\ln t} (\ln t - \tau)^{2\eta-2} \tau^{-\eta p} \, d\tau.$$

Choose and fix a constant $\beta > 0$ that satisfies:

$$\eta p - 1 < \beta \le p/2. \qquad (3.7)$$

The existence of such an β follows from the conditions $\eta > 1/2$ and $2 \le p < \frac{2}{2\eta - 1}$. Then

$$\left(\frac{(\ln \frac{t}{T_1})^{2\eta-1}}{2\eta-1} \right)^{(p-2)/2} \int_{T_1}^{t} \left(\ln \frac{t}{\tau} \right)^{2\eta-2} \| Y(\tau,\xi_1)) - Y(\tau,\xi_2) \|_p^p \frac{d\tau}{\tau}$$

$$\le \varepsilon^p \left(\frac{1}{2\eta-1} \right)^{(p-2)/2} (\ln t)^{\frac{(2\eta-1)(p-2)}{2} + 2\eta - 1 - \eta p} \int_{\frac{\ln T_1}{\ln t}}^{1} (1-\tau)^{2\eta-2} \tau^{-\eta p + \beta} \left(\frac{\ln t}{\ln T_1} \right)^{\beta} d\tau$$

$$\le \varepsilon^p \left(\frac{1}{2\eta-1} \right)^{(p-2)/2} \left(\frac{1}{\ln T_1} \right)^{\beta} B(2\eta - 1, \beta - \eta p + 1)(\ln t)^{\beta - \frac{p}{2}}. \qquad (3.8)$$

The conditions $\eta \in \left(\frac{1}{2}, 1 \right)$ and (3.7) leads to $\eta - 1 < 0, 2\eta - 1 > 0, \beta - \eta p + 1 > 0, \beta - p/2 < 0.$

Thus, inserting (3.8), (3.6), (3.5), and (3.4) into (3.3), and subsequently taking the limit superior as $t \to +\infty$ we achieve: $\| \xi_1 - \xi_2 \|_p \le \varepsilon C_4$, for some constant $C_4 > 0$. Finally, letting $\varepsilon \searrow 0$ then yields $\| \xi_1 - \xi_2 \|_p \le 0$. This contradicts the assumption that $\xi_1 \ne \xi_2$. Thus, the proof is complete.

4. The existence and uniqueness of solution under the local Lipschitz condition

Theorem 4.1. *Under the conditions* (H1-L) *and* (H3), *the equation (2.1) has a unique solution* $Z(t)$ *on* $[t_0, T]$.

Proof. Step 1: Existence: We first define the truncation functions f_ℓ and g_ℓ as follows:

For each $\ell \ge 1$:

$$f_\ell(t,y) = \begin{cases} f(t,y), & |y| \le \ell, \\ f\left(t, \frac{\ell y}{|y|}\right), & |y| > \ell, \end{cases} \qquad g_\ell(t,y) = \begin{cases} g(t,y), & |y| \le \ell, \\ g\left(t, \frac{\ell y}{|y|}\right), & |y| > \ell. \end{cases}$$

Then f_ℓ and g_ℓ satisfy the conditions (H1-G) and (H2). Thus, in light of [16], Theorem 3.1(i), the following equation admits a unique solution $Z^{(\ell)}(t)$:

$$Z^{(\ell)}(t) = \xi_1 + \frac{1}{\Gamma(\eta)} \int_{t_0}^{t} \left(\ln \frac{t}{\tau} \right)^{\eta-1} f_\ell(\tau, Z^{(\ell)}(\tau)) \frac{d\tau}{\tau} + \frac{1}{\Gamma(\eta)} \int_{t_0}^{t} \left(\ln \frac{t}{\tau} \right)^{\eta-1} g_\ell(\tau, Z^{(\ell)}(\tau)) \frac{dB_\tau}{\tau}. \qquad (4.1)$$

Now, we define the stopping time $\mu_\ell = T \wedge \inf\{t \in [t_0, T] : | Z^{(\ell)}(t) | \ge \ell\}$ and prove:

$$Z^{(\ell)}(t) = Z^{(\ell+1)}(t) \quad \text{for all } t \in [t_0, \mu_\ell]. \qquad (4.2)$$

To this end, by employing the inequality

$$|x_1+x_2+\cdots+x_m|^p \le m^{p-1}(|x_1|^p+|x_2|^p+\cdots+|x_m|^p) \text{ for all } x_1,x_2,\ldots,x_m \in \mathbb{R}^n \quad (4.3)$$

together with (4.1) and the definition of f_ℓ, g_ℓ, we achieve

$$E\left(\left|Z^{(\ell+1)}(t)-Z^{(\ell)}(t)\right|^p \chi_{\{t\le\mu_\ell\}}\right) \le \frac{4^{p-1}}{\Gamma^p(\eta)}\left\|\int_{t_0}^t\left(\ln\frac{t}{\tau}\right)^{\eta-1}(f_{\ell+1}(\tau,Z^{(\ell+1)}(\tau))-f_{\ell+1}(\tau,Z^{\ell}(\tau)))\chi_{\{t\le\mu_\ell\}}\frac{d\tau}{\tau}\right\|_p^p$$

$$+\frac{4^{p-1}}{\Gamma^p(\eta)}\left\|\int_{t_0}^t\left(\ln\frac{t}{\tau}\right)^{\eta-1}(g_{\ell+1}(\tau,Z^{(\ell+1)}(\tau))-g_{\ell+1}(\tau,Z^{\ell}(\tau)))\chi_{\{t\le\mu_\ell\}}\frac{dB_\tau}{\tau}\right\|_p^p.$$

Applying the Hölder inequality and BDG inequality, we obtain

$$E\left(\left|Z^{(\ell+1)}(t)-Z^{(\ell)}(t)\right|^p \chi_{\{\tau\le\mu_\ell\}}\right)$$

$$\le \frac{4^p L_\ell^p}{\Gamma^p(\eta)}\left[\frac{(\ln T-\ln t_0)^{\eta p-2\eta+1}}{(\eta p-2\eta+1)^{p-1}}(p-1)^{p-1}+\frac{D_p}{t_0^{p/2}}\left(\frac{\ln T-\ln t_0}{2\eta-1}\right)^{\frac{p-2}{2}}\right]\times\int_{t_0}^t\left(\ln\frac{t}{\tau}\right)^{2\eta-2}E(|Z^{(\ell+1)}(\tau)-Z^{\ell}(\tau)|^p \chi_{\{\tau\le\mu_\ell\}})\frac{d\tau}{\tau}.$$

Applying the generalized Gronwall inequality for CHFDE (see [19], Theorem 3), we receive $\sup_{t_0\le\tau\le t}\mathrm{E}\left(|Z^{(\ell+1)}(\tau)-Z^{(\ell)}(\tau)|^p\right)=0$ for $t_0\le t\le\mu_\ell$. This confirms that (4.2) holds, which in turn implies that the sequence μ_ℓ is an increasing sequence. We now proceed to show that $\mu_\ell=T$ for sufficiently large ℓ. To achieve this, we apply the Borel-Cantelli lemma. More precisely, by invoking Theorem 4.1 in [16], and employing estimates similar to those used in the proof of Theorem 3.1(i) in [16], we obtain

$$C_1 \ge \sup_{t_0\le t\le T}E(|Z^{(\ell)}(t)|^p) \ge \sup_{t_0\le t\le T}E(|Z^{(\ell)}(t)|^p \chi_{\{\mu_\ell<T\}}) \ge \ell^p\mathrm{P}(\mu_\ell<T).$$

The Borel-Cantelli lemma states that $\mathrm{P}(\limsup_\ell(\mu_\ell<T))=0$. Combining this with the fact that μ_ℓ is increasing, it follows that $\mu_\ell=T$ for sufficiently large ℓ. Consequently, for almost all $\omega\in\Omega$, there exists an integer ℓ_0 such that $\mu_\ell=T$. For all $\ell\ge\ell_0$. Now, define $Z(t)=Z^{(\ell_0)}(t), t\in[t_0,T]$. By (4.2), we have $Z(t\wedge\mu_\ell)=Z^{(\ell)}(t\wedge\mu_\ell)=Z^{(\ell_0)}(t\wedge\mu_\ell)$. This together with (4.1) drive that

$$Z(t\wedge\mu_\ell)=\xi_1+\frac{1}{\Gamma(\eta)}\int_{t_0}^{t\wedge\mu_\ell}\left(\ln\frac{t\wedge\mu_\ell}{\tau}\right)^{\eta-1}f(\tau,Z(\tau))\frac{d\tau}{\tau}+\frac{1}{\Gamma(\eta)}\int_{t_0}^{t\wedge\mu_\ell}\left(\ln\frac{t\wedge\mu_\ell}{\tau}\right)^{\eta-1}g(\tau,Z(\tau))\frac{dB_\tau}{\tau}.$$

As $\ell\to\infty$ we have $\mu_\ell\to T$. It follows that $Z(t)$ is a solution of (2.1). This completes the proof of *Step 1*.

Step 2: Uniqueness: Let $Z(t)$ and $\hat{Z}(t)$ be two solutions of (2.1) on $[t_0,T]$ defined with respect to the same Brownian motion B and the same initial condition ξ_1. For each $\ell\in\mathbb{N}^*$, we define

$$\eta_\ell=T\wedge\inf\{t\in[t_0,T]:|Z(t)|\ge\ell\},\hat{\eta}_\ell=T\wedge\inf\{t\in[t_0,T]:|\hat{Z}(t)|\ge\ell\},\zeta_\ell:=\eta_\ell\wedge\hat{\eta}_\ell.$$

Clearly, $\lim_{\ell\to\infty}\zeta_\ell = T$ P-almost surely. Using similar estimations as in *Step 1*, we attain

$$E\Big(\big|Z(t)-\hat{Z}(t)\big|^p\,\chi_{\{t\le\zeta_\ell\}}\Big) \le \frac{2^{p-1}}{\Gamma^p(\eta)}\left\|\int_{t_0}^t\left(\ln\frac{t}{\tau}\right)^{\eta-1}(f(\tau,Z(\tau))-f(\tau,\hat{Z}(\tau)))\chi_{\{t\le\zeta_\ell\}}\frac{d\tau}{\tau}\right\|_p^p$$
$$+\frac{2^{p-1}}{\Gamma^p(\eta)}\left\|\int_{t_0}^t\left(\ln\frac{t}{\tau}\right)^{\eta-1}(g(\tau,Z(\tau))-g(\tau,\hat{Z}(\tau)))\chi_{\{t\le\zeta_\ell\}}\frac{dB_\tau}{\tau}\right\|_p^p.$$

By applying Hölder's inequality, the BDG inequality, and condition (H1-L):

$$E\Big(\big|Z(t)-\hat{Z}(t)\big|^p\,\chi_{\{t\le\zeta_\ell\}}\Big) \le \frac{2^{p-1}L_\ell^p}{\Gamma^p(\eta)}\left[\frac{(\ln T-\ln t_0)^{\eta p-2\eta+1}}{(\eta p-2\eta+1)^{p-1}}(p-1)^{p-1}+\frac{D_p}{t_0^{p/2}}\left(\frac{\ln T-\ln t_0}{2\eta-1}\right)^{\frac{p-2}{2}}\right]$$
$$\times\int_{t_0}^t\left(\ln\frac{t}{\tau}\right)^{2\eta-2}E(|Z(\tau)-\hat{Z}(\tau)|^p\,\chi_{\{\tau\le\zeta_\ell\}})\frac{d\tau}{\tau}.$$

Applying the generalized Gronwall inequality for CHFDE, we have

$$\sup_{t_0\le t\le T}\mathrm{E}\Big(\big|Z(t)-\hat{Z}(t)\big|^p\,\chi_{\{t\le\zeta_\ell\}}\Big)=0.$$

Furthermore,

$$\sup_{t_0\le t\le T}\mathrm{E}(|Z(t)-\hat{Z}(t)|^p) \le \sup_{t_0\le t\le T}\mathrm{E}(|Z(t)-\hat{Z}(t)|^p\,\chi_{\{t\le\zeta_\ell\}})+\sup_{t_0\le t\le T}\mathrm{E}(|Z(t)-\hat{Z}(t)|^p\,\chi_{\{t>\zeta_\ell\}}).$$

As $\ell\to\infty$ we have $\sup_{t_0\le t\le T}\mathrm{E}\Big(\big|Z(t)-\hat{Z}(t)\big|^p\Big)=0$. This implies that the processes $\{Z(t),t_0\le t\le T\}$ and $\{\hat{Z}(t),t_0\le t\le T\}$ are indistinguishable. The proof is complete.

Acknowledgement

The authors are grateful to Dr. Phan Thi Huong for insightful discussions that helped improve this work.

Bibliography

[1] D. Băleanu and A. M. Lopes, Applications in Engineering, Life and Social Sciences, Walter de Gruyter GmbH \& Co KG, 2019.

[2] I. Petrás, Handbook of Fractional Calculus with Applications, De Gruyter, 2019.

[3] H. Sun, Y. Zhang, D. Baleanu, W. Chen and Y. Chen, "A new collection of real world applications of fractional calculus in science and engineering," *Communications in Nonlinear Science and Numerical Simulation,* vol. 64, 2018.

[4] V. E. Tarasov, Applications in physics, Walter de Gruyter GmbH \& Co KG, 2019.

[5] K. Diethelm and N. Ford, "The analysis of fractional differential equations," *Lecture notes in mathematics,* vol. 2004, 2010.

[6] P. T. Anh, T. S. Doan and P. T. Huong, "A variation of constant formula for Caputo fractional stochastic differential equations," *Statistics \& Probability Letters,* vol. 145, 2019.

[7] P. T. Huong, P. Kloeden and T. S. Doan, "Well-posedness and regularity for solutions of Caputo stochastic fractional differential equations in Lp spaces," *Stochastic Analysis and Applications,* vol. 41, 2023.

[8] P. T. Huong and P. T. Anh, "On the asymptotic behavior of solutions to bilinear Caputo stochastic fractional differential equations," *Statistics \& Probability Letters,* vol. 216, 2025.

[9] T. S. Doan, P. T. Huong, P. E. Kloeden and H. T. Tuan, "Asymptotic separation between solutions of Caputo fractional stochastic differential equations," *Stochastic Analysis and Applications,* vol. 36, 2018.

[10] P. T. Huong and N. N. Thang, "Well-Posedness and Regularity of Solutions for Stochastic Volterra Integral Equations With General Weakly Singular Kernels," *Mathematical Methods in the Applied Sciences,* 2025.

[11] H. T. Tuan, "On the asymptotic behavior of solutions to time-fractional elliptic equations driven a multiplicative white noise," *Discrete and Continuous Dynamical Systems - B,* vol. 26, pp. 1749 - 1762, 2021.

[12] M. Gohar, C. Li and C. Yin, "On Caputo--Hadamard fractional differential equations," *International Journal of Computer Mathematic,* vol. 97, 2020.

[13] L. Ma, "On the kinetics of Hadamard-type fractional differential systems," *Fractional Calculus and Applied Analysis,* vol. 23, 2020.

[14] Z. Yang, X. Zheng and H. Wang, "Well-posedness and regularity of Caputo--Hadamard fractional stochastic differential equations," *Zeitschrift f{\"u}r angewandte Mathematik und Physik,* vol. 72, 2021.

[15] M. Li, C. Huang and Nan Wang, "A variation of constant formula for Caputo–Hadamard fractional stochastic differential equations," *Statistics & Probability Letters,* vol. 214, 2024.

[16] L. T. T. Phuong, "Well-posedness and asymptotic behavior of solutions for Caputo-Hadamard stochastic fractional differential equations," *Submitted.*

[17] C. Li and Z. Li, "Stability and logarithmic decay of the solution to Hadamard-type fractional differential equation," *Journal of Nonlinear Science,* vol. 31, 2021.

[18] P. Huong, N. Thang and P. Anh, "Well-posedness and asymptotic separation between solutions of stochastic fractional differential equations with doubly singular kernels," *Communications in Nonlinear Science and Numerical Simulation,* vol. 159, p. 109885, 2026.

[19] J. Vanterler da Costa Sousa and E. Capelas de Oliveira, "A Gronwall inequality and the Cauchy-type problem by means of \Psi-Hilfer operator," *arXiv e-prints,* 2017.

Sự tồn tại, duy nhất và sự phân tách tiệm cận của nghiệm phương trình vi phân phân thứ Caputo-Hadamard ngẫu nhiên

Tóm tắt: *Trước tiên, chúng tôi nghiên cứu sự tồn tại và duy nhất nghiệm của phương trình vi phân phân thứ Caputo-Hadamard ngẫu nhiên bậc* $\eta \in (\frac{1}{2}, 1)$ *trong các không gian* L^p $(p \geq 2)$ *dưới điều kiện Lipschitz địa phương và điều kiện tăng trưởng tuyến tính. Để đạt được điều này, chúng tôi kết hợp kỹ thuật hàm cắt cụt (truncation function) với các tính chất của thời điểm dừng (stopping times), bất đẳng thức Grönwall suy rộng và bổ đề Borel-Cantelli. Bên cạnh đó, chúng tôi thiết lập sự phân tách tiệm cận (asymptotic separation) giữa các nghiệm phân biệt. Phương pháp chứng minh dựa trên các tính chất của nhân dạng logarit (logarithmic kernel), kết hợp với việc áp dụng chặt chẽ các bất đẳng thức Minkowski, Hölder và Burkholder-Davis-Gundy (BDG).*

Từ khoá: Phương trình vi phân phân thứ Caputo-Hadamard ngẫu nhiên; Sự tồn tại; Tính duy nhất; Điều kiện Lipschitz toàn cục; Điều kiện Lipschitz địa phương; Điều kiện tăng trưởng tuyến tính; Sự phân tách tiệm cận giữa các nghiệm.

Kỹ thuật thủy vân ảnh bền vững dựa trên DWT–SVD QIM sử dụng điều khiển hỗn loạn và sinh khóa hậu lượng tử

Hoàng Thị Thu Lan[1], Tạ Minh Thanh[1]

Học viện kỹ thuật quân sự

Email: thanhtm@lqdtu.edu.vn; Tel: 0986007979

Tóm tắt

Bài báo đề xuất một phương pháp giấu thủy vân ảnh bền vững và an toàn dựa trên mô hình DWT–SVD–QIM trong khuôn khổ đa khóa. Thủy vân nhị phân được nhúng vào giá trị suy biến lớn nhất S_0 của các khối SVD trong miền LL của DWT trên kênh độ chói Y nhằm khai thác tính ổn định của miền tần số thấp và đặc tính bền vững của các giá trị suy biến trội. Để nâng cao mức độ bảo mật và khả năng xác minh theo nhiều ngữ cảnh, hệ thống sử dụng khóa gốc sinh từ cơ chế mật mã hậu lượng tử, sau đó được xử lý qua hàm băm SHA3-256 để tạo tham số điều khiển. Từ khóa này, các cấu hình nhúng khác nhau có thể được thiết lập thông qua việc thay đổi ngữ cảnh xác minh, cho phép tạo ra nhiều phiên bản ảnh nhúng từ cùng một ảnh gốc và thực hiện trích xuất độc lập theo từng khóa mà không cần ảnh gốc. Kết quả thực nghiệm trên ảnh màu chuẩn 512×512 cho thấy phương pháp duy trì chất lượng ảnh sau nhúng cao và đạt độ bền vững tốt trước các tấn công xử lý ảnh phổ biến, đồng thời thể hiện tính phân biệt rõ ràng khi sử dụng khóa không tương ứng.

Từ khóa: *Thủy vân ảnh số; DWT-SVD-QIM; Điều khiển hỗn loạn; Mật mã hậu lượng tử; Xác minh bản quyền.*

1. Mở đầu (Introduction)

Trong bối cảnh chuyển đổi số, hình ảnh số được sử dụng rộng rãi trong nhiều lĩnh vực như truyền thông, thương mại điện tử, y tế và giám sát an ninh. Tuy nhiên, khả năng dễ sao chép và chỉnh sửa của ảnh số làm gia tăng nguy cơ vi phạm bản quyền, đòi hỏi các cơ chế kỹ thuật nhằm bảo vệ quyền sở hữu trí tuệ và xác thực nguồn gốc dữ liệu. Thủy vân ảnh số là một giải pháp hiệu quả cho phép nhúng thông tin xác thực trực tiếp vào ảnh mà không làm suy giảm đáng kể chất lượng cảm nhận [1].

Một hệ thống thủy vân cần đảm bảo tính vô hình, độ bền vững và khả năng trích xuất tin cậy. Phân rã giá trị suy biến (SVD) được sử dụng rộng rãi nhờ tính ổn định của các giá trị suy biến trước các phép biến đổi ảnh [2]. Các mô hình kết hợp Biến đổi Wavelet rời rạc (DWT) và SVD cho thấy hiệu quả cao khi nhúng trong miền tần số thấp [3–5], trong khi Điều chế chỉ số lượng tử hóa (QIM) giúp thiết lập ranh giới quyết định rõ ràng, nâng cao độ bền vững khi trích xuất [6,7]. Ngoài ra, cơ chế điều khiển hỗn loạn đã được sử dụng để tăng tính phụ thuộc khóa và giảm khả năng suy đoán cấu hình nhúng [8,9].

Bên cạnh yêu cầu về xử lý tín hiệu, yếu tố an toàn ngày càng quan trọng trong bối cảnh điện toán lượng tử phát triển. Các báo cáo của NIST đã nhấn mạnh nhu cầu nghiên cứu các cơ chế mật mã hậu lượng tử nhằm đảm bảo an toàn dài hạn [10,11]. Trong lĩnh vực thủy vân, các phương pháp lựa chọn khối nhúng trong các miền phân rã như LU hoặc QR đã được đề xuất để nâng cao độ bền vững [12,13], đồng thời các mô hình quản lý quyền số (DRM) cho thấy nhu cầu xác minh bản quyền theo từng ngữ cảnh sử dụng [14].

Tuy nhiên, nhiều phương pháp hiện nay vẫn sử dụng một khóa duy nhất cho toàn bộ quá trình nhúng và trích xuất, chưa phù hợp khi nội dung được phân phối theo nhiều phiên bản. Bài báo đề xuất cơ chế dẫn xuất khóa theo ngữ cảnh, trong đó mỗi lần nhúng sử dụng một khóa riêng được tạo theo định hướng hậu lượng tử (post-quantum-oriented key derivation) bằng hàm

băm SHA3-256 kết hợp tham số ngữ cảnh để sinh các cấu hình điều khiển khác nhau. Cách tiếp cận này hỗ trợ đa khóa thông qua các lần nhúng độc lập và bảo đảm tính tách biệt giữa các cấu hình xác minh.

Trên cơ sở đó, bài báo đề xuất mô hình thủy vân mù dựa trên DWT–SVD–QIM, trong đó thủy vân được nhúng vào giá trị suy biến lớn nhất của các khối SVD trong miền LL của DWT trên kênh độ chói Y. Các tham số điều khiển phụ thuộc khóa được sử dụng để thiết lập cấu hình nhúng tương ứng với từng ngữ cảnh. Hiệu năng phương pháp được đánh giá theo các độ đo chất lượng, độ bền vững và tính phân biệt khóa.

Phần còn lại của bài báo được tổ chức như sau. Mục 2 trình bày cơ sở lý thuyết và mô hình đề xuất, bao gồm kiến trúc hệ thống và quy trình nhúng – trích xuất. Mục 3 mô tả thiết lập thực nghiệm và phân tích kết quả về tính vô hình, độ bền vững và khả năng phân biệt khóa. Cuối cùng, Mục 4 đưa ra kết luận và hướng nghiên cứu tiếp theo.

2. Phương pháp đề xuất

2.1. Cơ sở lý thuyết

2.1.1. Biến đổi Wavelet rời rạc (DWT)

Biến đổi Wavelet rời rạc hai chiều (2D-DWT) cho phép phân rã ảnh số $I(x,y)$ thành các thành phần đa phân giải bằng các bộ lọc thông thấp và thông cao theo hai chiều không gian. Sau một mức phân rã, ảnh được tách thành bốn miền con:

$$I \xrightarrow{DWT} \{LL, LH, HL, HH\} \tag{1}$$

trong đó miền LL biểu diễn thành phần tần số thấp theo cả hai chiều, chứa phần lớn năng lượng và thông tin cấu trúc của ảnh. Các miền LH, HL và HH tương ứng với các thành phần biên và chi tiết tần số cao.

Do các phép xử lý ảnh phổ biến như nén JPEG, nhiễu hoặc lọc thường ảnh hưởng mạnh đến các thành phần tần số cao, miền LL được xem là miền nhúng phù hợp nhằm đảm bảo độ bền vững của thủy vân [3].

2.1.2. Phân rã giá trị suy biến (SVD)

Với mỗi khối ảnh $A \in {}^{\sim m\times n}$ phép phân rã giá trị suy biến cho phép biểu diễn A dưới dạng:

$$A = USV^T \tag{2}$$

trong đó $U \in R^{m\times m}$ và $V \in R^{n\times n}$ là các ma trận trực giao, $S = diag(\sigma_1, \sigma_2, \ldots, \sigma_r)$ là ma trận đường chéo chứa các giá trị suy biến với:

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq 0 \tag{3}$$

Giá trị suy biến lớn nhất σ_1 có tính ổn định cao trước các phép biến đổi ảnh như nén, nhiễu và lọc [2]. Do đó, việc nhúng thủy vân vào σ_1 cho phép duy trì khả năng trích xuất ngay cả khi ảnh chịu các biến dạng mạnh.

2.1.3. Điều chế chỉ số lượng tử hóa (Quantization Index Modulation – QIM)

Điều chế chỉ số lượng tử hoá (QIM) là kỹ thuật nhúng thông tin dựa trên lượng tử hóa có điều khiển. Với một giá trị x và bước lượng tử hoá Δ, hàm nhúng QIM cho bit thủy vân $b \in \{0,1\}$ được xác định như sau:

$$Q(x,b)=\begin{cases} round\left(\dfrac{x}{\Delta}\right)\Delta,\ b=0 \\ round\left(\dfrac{x-\Delta/2}{\Delta}\right)\Delta+\dfrac{\Delta}{2},\ b=1 \end{cases} \tag{4}$$

Trong quá trình trích xuất, bit thủy vân $\hat{b}$ được xác định dựa trên vị trí của giá trị nhận được x′ trong miền lượng tử hoá:

$$\hat{b}=\begin{cases} 0,\ x'\ mod\ \Delta<\Delta/2 \\ 1,\ x'\ mod\ \Delta\geq\Delta/2 \end{cases} \tag{5}$$

QIM tạo ranh giới quyết định rõ ràng, giúp giảm sai số trích xuất [5,6].

2.1.4. Bộ sinh chuỗi hỗn loạn

Các ánh xạ hỗn loạn có đặc tính nhạy cảm cao với điều kiện đầu và khả năng sinh chuỗi số có phân bố gần ngẫu nhiên, phù hợp cho các bài toán xáo trộn và mã hóa dữ liệu đa phương tiện. Một ví dụ điển hình là ánh xạ logistic:

$$x_{n+1}=\mu x_n(1-x_n),\ 0<x_n<1, 3.57<\mu\leq 4 \tag{6}$$

với tham số điều khiển μ thuộc miền hỗn loạn thích hợp.

Trong mô hình đề xuất, ánh xạ logistic được sử dụng như một bộ sinh chuỗi hỗn loạn phụ thuộc khóa, nhằm xáo trộn thứ tự bit thủy vân, mã hóa chuỗi bit và hoán vị vị trí các khối nhúng. Các tham số khởi tạo của ánh xạ được dẫn xuất từ khóa sử dụng tại mỗi lần nhúng, bảo đảm tính phụ thuộc khóa của toàn bộ quá trình [7]. Do đó, việc thay đổi khóa sẽ dẫn đến chuỗi hỗn loạn và cấu hình nhúng hoàn toàn khác biệt.

2.1.5. Cơ chế sinh khóa theo định hướng hậu lượng tử

Để tăng mức độ an toàn của hệ thống trong bối cảnh các phương pháp mật mã cổ điển có thể bị ảnh hưởng bởi điện toán lượng tử, phương pháp đề xuất sử dụng cơ chế dẫn xuất khóa theo định hướng hậu lượng tử. Trong cơ chế này, khóa điều khiển không được sử dụng trực tiếp mà được dẫn xuất thông qua hàm băm mật mã SHA3-256 kết hợp với khóa gốc, muối (salt) và tham số ngữ cảnh (context), cho phép tạo các khóa con độc lập cho từng mục đích sử dụng. Khóa dẫn xuất K′ được tính theo:

$$K'=SHA3-256(K\|salt\|context) \tag{7}$$

Khóa dẫn xuất K′ được dùng để sinh các tham số điều khiển hệ hỗn loạn, từ đó tạo chuỗi giả ngẫu nhiên để mã hóa bit thủy vân và hoán vị vị trí nhúng phụ thuộc khóa. Trong quá trình trích xuất, cùng khóa và tham số ngữ cảnh được sử dụng để tái tạo lại chuỗi điều khiển, bảo đảm khôi phục chính xác khi khóa đúng và cho kết quả gần như ngẫu nhiên khi khóa sai. Cơ chế này không làm thay đổi cấu trúc của thuật toán DWT–SVD–QIM nhưng làm tăng mức độ phụ thuộc khóa của toàn bộ quá trình nhúng – trích xuất [9,10].

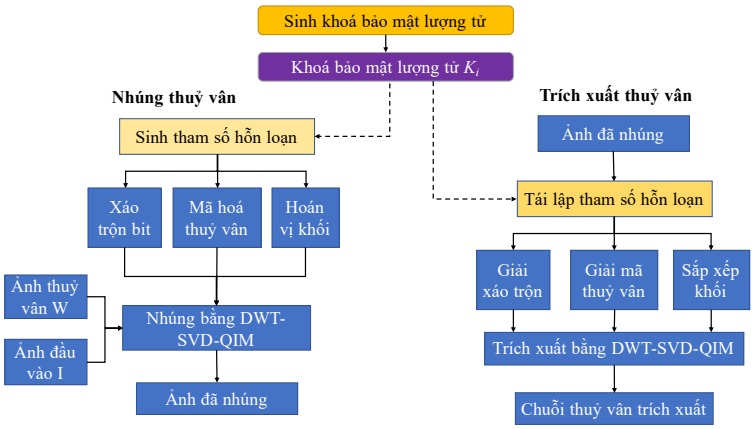
So với bộ sinh số giả ngẫu nhiên với giá trị khởi tạo cố định, cơ chế dẫn xuất khóa theo ngữ cảnh cho phép tạo các chuỗi điều khiển độc lập từ cùng một khóa gốc, tăng khả năng phân biệt giữa các khóa. Việc sử dụng SHA3-256 với độ dài 256 bit cung cấp không gian khóa đủ lớn để chống tấn công vét cạn trong thực tế mà không làm tăng đáng kể độ phức tạp tính toán.

Trong phương pháp đề xuất, mỗi phiên bản ảnh chỉ sử dụng một khóa tại một thời điểm; hệ thống hỗ trợ đa khóa thông qua các lần nhúng độc lập, không nhúng đồng thời nhiều khóa trên cùng một ảnh nhằm bảo đảm độ bền vững của thủy vân.

2.2. Kiến trúc tổng thể của hệ thống đề xuất

Hệ thống thủy vân được xây dựng theo mô hình thủy vân mù, trong đó mỗi lần nhúng và trích xuất được điều khiển bởi một khóa xác định. Thay vì dùng một khóa cố định, hệ thống áp dụng cơ chế dẫn xuất khóa theo định hướng hậu lượng tử để tạo các khóa độc lập cho từng phiên bản nhúng, cho phép thiết lập các cấu hình xác minh khác nhau.

Gọi I là ảnh màu đầu vào và W là chuỗi thủy vân nhị phân. Với khóa K_i, hệ thống thực hiện hàm băm để thu được chuỗi K_i', sau đó ánh xạ sang các tham số khởi tạo của bộ sinh hỗn loạn logistic. Do quá trình ánh xạ phụ thuộc hoàn toàn vào khóa, cùng một khóa sẽ tạo ra cùng một chuỗi điều khiển, bảo đảm khả năng đồng bộ giữa nhúng và trích xuất.



Hình 1. Sơ đồ khối tổng thể của hệ thống đề xuất

Trong pha nhúng, chuỗi hỗn loạn phụ thuộc khóa được dùng để xáo trộn và mã hóa bit thủy vân, đồng thời hoán vị vị trí các khối nhúng. Ảnh được xử lý qua các bước chuyển đổi màu, phân rã DWT, phân rã SVD theo khối và nhúng vào giá trị suy biến lớn nhất bằng QIM dưới sự điều khiển của tham số phụ thuộc khóa. Cơ chế dẫn xuất khóa không chỉ thiết lập tham số khởi tạo mà còn quyết định toàn bộ cấu hình nhúng, nên khóa khác sẽ dẫn đến sai lệch cả vị trí và dữ liệu nhúng. So với bộ sinh số giả ngẫu nhiên với giá trị khởi tạo cố định, cách tiếp cận này tạo các chuỗi điều khiển độc lập từ cùng một khóa gốc mà không thay đổi cấu trúc của thuật toán DWT–SVD–QIM.

Trong pha trích xuất, khóa K_i được dùng để tái tạo chuỗi điều khiển và khôi phục thủy vân. Nếu khóa không đúng, chuỗi điều khiển sẽ khác, dẫn đến sai lệch trong quá trình giải xáo trộn và trích xuất, bảo đảm tính phân biệt khóa. Sơ đồ khối tổng thể của hệ thống minh họa trong Hình 1.

2.3. Quy trình nhúng thủy vân

Ảnh màu đầu vào I được chuyển sang không gian YCbCr và chọn kênh độ chói Y để nhúng. Trên kênh này, phép biến đổi DWT hai chiều được áp dụng để thu miền con LL

Miền LL được chia thành các khối không chồng lấn kích thước N×N. Với mỗi khối B_k , phép phân rã SVD được thực hiện:

$$B_k = U_k S_k V_k^T \tag{8}$$

trong đó $S_k = diag(\sigma_{k,1}, \sigma_{k,2}, \ldots)$ và $\sigma_{k,1}$ là giá trị suy biến lớn nhất của khối thứ k.

Chuỗi thủy vân W sau khi được xáo trộn và mã hóa bởi cơ chế hỗn loạn phụ thuộc khóa K_i, được nhúng vào các giá trị suy biến lớn nhất thông qua kỹ thuật QIM:

$$\sigma'_{k,1} = Q(\sigma_{k,1}, w_k) \tag{9}$$

Khối ảnh sau nhúng được tái tạo thông qua phép SVD ngược. Sau khi xử lý toàn bộ các khối, miền LL′ được tái tạo và ảnh nhúng cuối cùng I_i được khôi phục bằng phép biến đổi DWT ngược.

2.4. Quy trình trích xuất thủy vân mù

Quá trình trích xuất được thực hiện mà không cần ảnh gốc. Từ khóa K_i, các tham số hỗn loạn được tái tạo để đồng bộ cấu hình nhúng.

Ảnh nghi ngờ I_i được chuyển đổi sang YCbCr, trích xuất kênh Y và thực hiện DWT để thu miền LL. Miền này được chia thành các khối và thực hiện SVD trên từng khối. Bit thủy vân được xác định từ giá trị suy biến lớn nhất $\sigma^*_{k,1}$ dựa trên ranh giới quyết định của QIM:

$$\hat{w}_k = \begin{cases} 0, & \sigma^*_{k,1} \bmod \Delta < \Delta/2 \\ 1, & \sigma^*_{k,1} \bmod \Delta \geq \Delta/2 \end{cases} \tag{10}$$

Chuỗi bit sau đó được giải xáo trộn bằng cơ chế hỗn loạn để khôi phục thủy vân $\hat{W}$.

Độ tương đồng giữa W và $\hat{W}$ được đánh giá bằng hệ số tương quan chuẩn hóa (NC):

$$NC = \frac{\sum_{i=1}^{L} W(i)\hat{W}(i)}{\sqrt{\sum_{i=1}^{L} W(i)^2}\sqrt{\sum_{i=1}^{L} \hat{W}(i)^2}} \tag{11}$$

Ngoài ra, tỷ lệ lỗi bit (BER) được sử dụng để đo mức độ sai lệch trong trích xuất:

$$BER = \frac{1}{L}\sum_{i=1}^{L} \left| W(i) - \hat{W}(i) \right| \tag{12}$$

3. Kết quả và thảo luận

3.1. Thiết lập thực nghiệm

Các thí nghiệm được thực hiện trên tập ảnh màu chuẩn kích thước 512×512 với đặc trưng kết cấu đa dạng. Thủy vân là ảnh nhị phân 64×64 và được nhúng toàn bộ vào ảnh chủ.

Phương pháp đề xuất DWT–SVD–QIM được so sánh với hai phương pháp tham chiếu là SVD–QIM trong miền không gian trên kênh Y và kênh Red. Do khác biệt miền nhúng, phương pháp đề xuất thực hiện SVD trên miền LL của DWT với khối 4×4, trong khi các phương pháp miền không gian sử dụng khối 8×8, nhằm đảm bảo dung lượng nhúng tương đương.

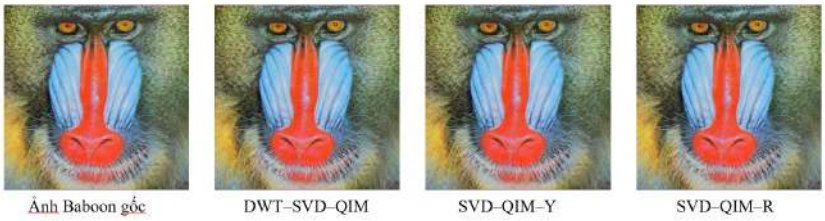
Để đảm bảo tính công bằng, tất cả các phương pháp sử dụng cùng bước lượng tử QIM Δ và cùng điều kiện tấn công. Đối với phương pháp đề xuất, khóa K_i và các tham số hỗn loạn được cố định trong từng kịch bản nhằm đảm bảo khả năng tái lập kết quả.

3.2. Đánh giá tính vô hình của thủy vân

Tính vô hình được đánh giá bằng PSNR và SSIM giữa ảnh gốc và ảnh sau nhúng. Kết quả trong Bảng 1 cho thấy phương pháp DWT–SVD–QIM duy trì PSNR trên 45 dB và SSIM xấp xỉ 1 trên các ảnh thử nghiệm.

Bảng 1. Tỷ số tín hiệu trên nhiễu đỉnh (PSNR) và chỉ số tương đồng cấu trúc (SSIM) giữa các phương pháp

Ảnh	PSNR (DWT–SVD–QIM)	PSNR (SVD–QIM–Y)	PSNR (SVD–QIM–R)	SSIM (DWT–SVD–QIM)	SSIM (SVD–QIM–Y)	SSIM (SVD–QIM– R)
Lena	45,100	45,921	50,730	1,000	1,000	1,000
Baboon	45,738	45,805	50,584	1,000	1,000	1,000
Peppers	45,083	45,962	50,661	0,995	0,997	1,000



Hình 2. Ảnh Baboon gốc và các ảnh sau nhúng tương ứng với ba phương pháp thử nghiệm

So với hai phương pháp miền không gian, phương pháp đề xuất cho chất lượng thị giác tương đương. Mặc dù phương pháp trên kênh R đạt PSNR cao hơn, các giá trị SSIM đều gần 1, cho thấy khác biệt thị giác không đáng kể.

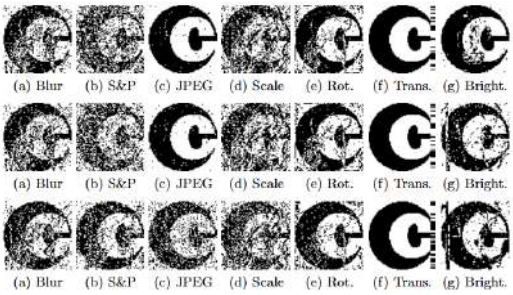
3.3. Đánh giá độ bền vững trước các tấn công xử lý ảnh

Độ bền vững của phương pháp được đánh giá bằng hệ số tương quan chuẩn hóa (NC) giữa thủy vân gốc và thủy vân trích xuất sau các phép tấn công xử lý ảnh. Các kịch bản gồm biến đổi phi hình học và hình học: làm mờ Gaussian (3×3), nhiễu muối–tiêu (0,03), nén JPEG (QF=90), co giãn (1,1), xoay (10°), tịnh tiến (40 px) và thay đổi độ sáng (+30). Sau mỗi phép tấn công, thủy vân được trích xuất theo mô hình mù và giá trị NC được tính để đánh giá mức độ tương đồng. Kết quả trung bình được trình bày trong Bảng 2 và minh họa trong Hình 3.

Bảng 2. So sánh độ tương quan chuẩn hóa (NC) giữa ba phương pháp dưới các phép tấn công

Phép tấn công	DWT–SVD–QIM	SVD–QIM–Y	SVD–QIM–R
Không tấn công	1,000	1,000	1,000
Làm mờ Gaussian (mức 3)	0,791	0,774	0,761
Nhiễu muối – tiêu	0,523	0,513	0,529
Nén JPEG (QF = 90)	0,988	0,983	0,791
Co giãn (tỷ lệ 1,1)	0,680	0,661	0,683
Xoay (10°)	0,831	0,815	0,836
Tịnh tiến (40 px)	0,951	0,951	0,931
Thay đổi độ sáng (+30)	0,878	0,873	0,812

Dưới các phép làm mờ, nhiễu và đặc biệt là nén JPEG, phương pháp DWT–SVD–QIM đạt NC cao hơn hai phương pháp so sánh, cho thấy ưu thế của việc nhúng trong miền tần số thấp kết hợp với tính ổn định của giá trị suy biến lớn nhất. Với các biến đổi hình học, NC giảm ở tất cả các phương pháp nhưng thủy vân vẫn có thể nhận dạng.



Hình 3. Thủy vân trích xuất dưới các phép tấn công. Hàng 1: DWT–SVD–QIM; Hàng 2: SVD–QIM (Y); Hàng 3: SVD–QIM (R).

Ảnh hưởng của bước lượng tử hóa Δ cũng được khảo sát để phân tích sự đánh đổi giữa tính vô hình và độ bền vững. Khi Δ tăng, NC tăng nhưng PSNR giảm; ngược lại, Δ nhỏ giữ chất lượng ảnh tốt nhưng giảm khả năng chống tấn công. Đây là đặc trưng của các hệ thống thủy vân dựa trên lượng tử hóa (quantization).

Về bảo mật, các phương pháp dùng khóa truyền thống thường dựa trên bộ sinh số giả ngẫu nhiên hoặc mật mã cổ điển, trong đó mức độ an toàn phụ thuộc vào độ dài khóa và giả định về năng lực tính toán hiện tại.

Phương pháp đề xuất sử dụng cơ chế dẫn xuất khóa theo định hướng hậu lượng tử dựa trên hàm băm SHA3-256 và tham số ngữ cảnh để tạo chuỗi điều khiển phụ thuộc khóa, cho phép sinh các cấu hình nhúng độc lập từ cùng một khóa gốc, tránh tái sử dụng giá trị khởi tạo và tăng tính phân biệt khóa mà không làm thay đổi cấu trúc thuật toán.

3.4. Đánh giá tính phân biệt khóa

Để đánh giá khả năng xác minh theo khóa, thí nghiệm được thực hiện với hai trường hợp: khóa đúng K_i và khóa sai K_j ($j \neq i$). Kết quả được trình bày trong Bảng 3.

Bảng 3. So sánh kết quả trích xuất khi sử dụng khóa đúng và khóa sai

Trường hợp sử dụng khóa	NC	SSIM	BER
Khóa đúng K_i	1,000	1,000	0,000
Khóa sai K_j	0,500	0,014	0,494

Khi sử dụng khóa đúng, quá trình trích xuất đạt độ chính xác tuyệt đối với NC và SSIM bằng 1, đồng thời BER bằng 0, cho thấy thủy vân được khôi phục hoàn toàn. Ngược lại, khi dùng khóa không tương ứng, BER xấp xỉ 0.5, cho thấy chuỗi bit trích xuất gần như ngẫu nhiên, đồng thời SSIM rất thấp (0,014), phản ánh sự sai khác rõ rệt giữa thủy vân gốc và thủy vân trích xuất.

Kết quả này chứng minh tính phân biệt khóa của hệ thống, trong đó mỗi phiên bản ảnh chỉ có thể được xác minh bằng khóa tương ứng, qua đó thể hiện khả năng hỗ trợ nhiều cấu hình xác minh độc lập trong mô hình đa khóa đề xuất.

4. Kết luận

Bài báo đã đề xuất phương pháp thủy vân ảnh mù dựa trên mô hình DWT–SVD–QIM, trong đó thủy vân được nhúng vào giá trị suy biến lớn nhất của các khối SVD trong miền LL của DWT trên kênh độ chói Y. Cách tiếp cận này khai thác tính ổn định của miền tần số thấp và giá trị suy biến trội để nâng cao độ bền vững, đồng thời kết hợp cơ chế điều khiển hỗn loạn và dẫn xuất khóa theo định hướng hậu lượng tử nhằm tăng mức độ an toàn. Mỗi phiên bản ảnh được điều khiển bởi một khóa riêng, cho phép thiết lập các cấu hình xác minh độc lập.

Kết quả thực nghiệm trên ảnh chuẩn 512×512 cho thấy phương pháp duy trì chất lượng ảnh sau nhúng cao (PSNR, SSIM lớn) và đạt độ bền vững tốt trước các tấn công như nén JPEG, nhiễu và lọc. Các độ đo NC và BER xác nhận khả năng trích xuất ổn định trong hầu hết các trường hợp. So với các phương pháp SVD–QIM trong miền không gian trên kênh Y và Red, phương pháp đề xuất cho độ bền vững cao hơn trong khi vẫn giữ tính vô hình tương đương.

Cơ chế dẫn xuất khóa giúp mở rộng không gian khóa, tăng khả năng chống suy đoán và bảo đảm tính phân biệt giữa các khóa, cho phép xác minh độc lập các phiên bản nhúng. Tuy nhiên, phương pháp vẫn có hạn chế về chi phí tính toán và độ nhạy với biến dạng hình học

mạnh. Trong tương lai, có thể nghiên cứu tối ưu tham số nhúng, sử dụng biến đổi wavelet không suy biến và bổ sung cơ chế đồng bộ hình học để nâng cao độ bền vững và khả năng ứng dụng.

Tài liệu tham khảo

1. Cox, I. J., Miller, M. L., Bloom, J. A., Fridrich, J., Kalker, T. (2008). *Digital watermarking and steganography* (2nd ed.). Burlington, MA: Morgan Kaufmann Publishers.
2. Liu, R., Tan, T. (2002). An SVD-based watermarking scheme for protecting rightful ownership. *IEEE Transactions on Multimedia*, 4(1), 121–128.
3. Bhatnagar, G., Raman, B. (2009). A new robust reference watermarking scheme based on DWT–SVD. *Computer Standards & Interfaces*, 31(5), 1002–1013.
4. Singh, M., Singh, S. (2015). Wavelet-based secure image watermarking using hybrid DWT–SVD domain. *Multimedia Tools and Applications*, 74(17), 6989–7007.
5. Gao, H., Chen, Y., Fang, L. (2022). Hybrid watermarking using DWT–SVD and nonlinear quantization for improved robustness. *Information Sciences*, 608, 540–555.
6. Tian, Y., Peng, X., Wei, Y. (2011). Watermarking algorithm based on quantization index modulation and singular value decomposition. *Advanced Materials Research*, 271–273, 536–540.
7. Li, J., Xu, K., Zhou, P. (2023). Improved SVD-based image watermarking using adaptive quantization of dominant singular values. *IEEE Signal Processing Letters*, 30, 1–5.
8. Wu, W., Dong, Y., Wang, G., Muñoz-Pacheco, J. M. (2024). Image robust watermarking method based on DWT–SVD transform and chaotic map. *Complexity*, 2024, Article ID 6618382.
9. Wang, X., Zhang, Y., Liu, H., Chen, Z. (2024). A robust image watermarking method based on multi-scale DWT–SVD and chaotic encryption. *IEEE Access*, 12, 12345–12356.
10. Chen, L., Jordan, S., Liu, Y. K., Moody, D., Peralta, R., Perlner, R., Smith-Tone, D. (2016). *Report on post-quantum cryptography (NISTIR 8105)*. Gaithersburg, MD: National Institute of Standards and Technology.
11. National Institute of Standards and Technology (NIST) (2023). *Post-quantum cryptography standardization: Selected algorithms*. Gaithersburg, MD: National Institute of Standards and Technology.
12. Nha, P. T., Thanh, T. M. (2024). A new block selection strategy from LU decomposition domain for robust image watermarking. *Multimedia Tools and Applications*, 83, 19301–19325.
13. Nha, P. T., Thanh, T. M., Phong, N. T. (2022). Consideration of a robust watermarking algorithm for color image using improved QR decomposition. *Soft Computing*, 26, 5069–5093.
14. Thanh, T. M., Iwakiri, M. (2014). A proposal of digital rights management based on incomplete cryptography using invariant Huffman code length feature. *Multimedia Systems*, 20, 127–142.

A Robust Image Watermarking Method Based on DWT–SVD–QIM with Chaotic Control and Post-Quantum Key Derivation

Abstract: *This paper proposes a robust digital image watermarking method based on the DWT–SVD–QIM framework within a multi-key verification scheme. A binary watermark is embedded into the largest singular value S_0 of SVD blocks in the LL subband obtained from a one-level discrete wavelet transform (DWT) applied to the luminance (Y) channel. This design exploits the stability of the low-frequency domain and the robustness of dominant singular values against common image processing operations. To enhance security and support context-dependent copyright verification, a master key is processed through a post-quantum-oriented key derivation mechanism based on the SHA3-256 hash function to generate control parameters for chaotic scheduling and embedding configuration. Each embedding instance is controlled by a single derived key, while different keys can be used in independent embedding processes to create distinct watermarked versions from the same original image. Experimental results on standard 512×512 color images demonstrate that the proposed method maintains*

high visual quality after embedding and achieves strong robustness against common image processing attacks. In addition, the scheme exhibits clear key discrimination when incorrect keys are used during watermark extraction.

Keywords: Digital image watermarking; DWT–SVD–QIM; Chaotic control; Post-quantum key derivation; Copyright verification.

Fusion Deep Learning Models for Quishing Detection

Thao Phuong Khong[1]

[1]*Institute of Information and Communication Technology; Le Quy Don Technical University*

Email: thao.khong@lqdtu.edu.vn; Tel: 0846122899

Abstract

Quishing attacks, which misuse QR codes to deliver malicious content, are becoming an increasingly serious cybersecurity threat. This study proposes a multimodal deep learning framework for detecting malicious QR codes by integrating visual features of QR images, structural properties of embedded URLs, semantic information, and dynamic behavioral attributes. The model applies feature normalization and multimodal fusion to enhance robustness and discrimination capability between benign and malicious samples. Experimental evaluation on a large-scale dataset demonstrates reliable and balanced detection performance. The proposed framework also provides a foundation for future research in adaptive threat detection and explainable multimodal learning, while offering strong potential for practical deployment in mobile security systems, web browsers, enterprise gateways, and other real-world QR-based applications.

Keywords: *quishing detection; multimodal deep learning; malicious QR analysis; cybersecurity; intelligent threat detection.*

1. Introduction

Quick Response (QR) codes have become widely adopted in modern digital services, including mobile payments, authentication, marketing, and information sharing. Their ability to encode large amounts of data in a machine-readable visual format makes them an efficient interface between physical and digital environments. However, this convenience also introduces significant security risks. Since QR codes conceal embedded content from users until scanning, they create opportunities for attackers to distribute malicious information with no suspicion.

One emerging threat is QR code phishing, commonly known as “quishing”, where attackers embed malicious URLs into seemingly legitimate QR codes. When scanned, these codes redirect users to fraudulent websites designed to steal credentials, financial information, or sensitive personal data. Unlike traditional phishing attacks, quishing can bypass many existing security mechanisms because the encoded information is hidden in image form and requires decoding prior to inspection. Furthermore, attackers often employ obfuscation techniques or previously unseen URLs, making detection more challenging.

Existing defenses primarily rely on blacklist-based filtering, rule-based detection, or textual analysis of URLs. While these methods are effective against conventional phishing attacks, they are limited in detecting QR-based threats due to the visual nature of encoded information and the lack of contextual analysis. Recent studies have explored machine learning and deep learning techniques to detect malicious QR codes by analyzing visual patterns, structural features, or decoded content. However, most existing approaches rely on single-modality information, which may not fully capture the complex characteristics of quishing attacks.

To address these limitations, this paper proposes a fusion-based deep learning framework for quishing detection that integrates multiple modalities, including visual features from QR code images, structural and lexical properties of decoded URLs, and contextual metadata. By

combining complementary information from heterogeneous sources, the proposed approach aims to improve detection accuracy, robustness, and generalization against evolving attack strategies.

2. Related Work

Recent research on quishing detection has evolved from traditional URL-based analysis to machine learning and deep learning–based frameworks. Existing studies can be broadly categorized into (1) feature-based and AI-driven detection approaches and (2) infrastructure-based prevention mechanisms.

2.1. Feature-based and AI-driven Detection for Quishing

Feature-based approaches detect malicious QR codes by analyzing embedded URLs, visual patterns, and associated web behaviors. Early work focuses on handcrafted URL features such as domain structure, lexical characteristics, and redirection patterns, which are used in traditional machine learning classifiers for phishing detection [1].

With the advancement of deep learning, recent studies pay attention to analyze QR code visual structures and detect anomalies caused by malicious manipulation and responses from encoded URLs [2, 3]. Besides, there has been a trend for quishing detection is to analyze only QR code structure and other visual features without obtaining the embedded URLs and their features [4, 5].

More recent work explores hybrid and multi-modal detection frameworks. For example, [6] integrates QR code image features with URL characteristics to improve detection accuracy.

Despite their effectiveness, most existing methods rely on single-modality analysis or limited feature representations, which may reduce their generalization capability in real-world scenarios. The other directions utilize fusion models to combine both image features and text features but in the form of static ones and skip other important dynamic features such as: webpage behaviors, http responses,…

2.2. Infrastructure-based Prevention and Verification for QR Codes

Unlike detection-based approaches, infrastructure-based solutions aim to prevent quishing attacks by enhancing QR code authenticity and integrity. These methods introduce secure QR code generation and verification mechanisms, including cryptographic signatures, watermarking, and trusted validation frameworks [7, 8].

Such approaches allow users or systems to verify the legitimacy of QR codes before accessing encoded content. However, these solutions often require modifications to existing infrastructure or user adoption of additional verification procedures, which may limit their scalability and practical deployment in open environments.

3. Methodology

3.1. System Overview

To effectively detect QR code–based phishing attacks, we propose a multimodal detection framework that integrates static analysis and dynamic analysis to capture complementary security signals from QR codes, URLs, and associated web behaviors. Unlike traditional approaches that rely solely on static URL inspection or visual analysis, the proposed method performs a comprehensive analysis pipeline covering QR code content extraction, structural inspection, and

runtime behavior monitoring. The overall framework consists of four main stages: (1) QR code decoding and feature extraction, (2) Static analysis of QR image and URL features, (3) Dynamic analysis of webpage and runtime behaviors, (4) Multimodal feature fusion and classification.

3.2. Static Analysis Module

The static analysis module examines QR code content and embedded URLs without executing external resources. This stage aims to identify structural and semantic characteristics that may indicate malicious intent.

The QR code image is first processed to extract global visual features that capture structural patterns and potential tampering artifacts. A vision-based encoder is used to learn discriminative representations from QR images, enabling detection of anomalies such as irregular module distributions, embedded malicious modifications, or suspicious visual structures. In addition to visual features, structural properties of QR codes are analyzed to detect manipulation attempts. This includes examining encoding parameters and structural consistency indicators that may reveal abnormal QR generation behavior.

After decoding the QR code, the embedded URL is analyzed using a text-based feature extraction module. The model encodes the semantic and lexical characteristics of URLs, including domain structure, token composition, and suspicious patterns. To capture both contextual meaning and sequential dependencies, semantic representations are extracted using a language encoder, followed by sequential modeling to learn URL patterns commonly associated with phishing attacks.

3.3. Dynamic Analysis Module

The system accesses decoded URLs in a sandbox environment to monitor webpage behaviors and extract runtime features, including: HTTP response characteristics, redirection chains, webpage content structure, security indicators such as SSL certificate information. These features help detect malicious behaviors such as phishing page redirection, deceptive content loading, or suspicious server responses.

Runtime observations are encoded into behavioral feature vectors representing the interaction between the user and the target webpage. These features provide contextual information about the actual impact of the QR code beyond its static representation. Dynamic analysis improves detection capability against: obfuscated URLs, previously unseen phishing domains, dynamic attack strategies.

3.4. Multimodal Feature Fusion

To leverage complementary information from different sources, features extracted from static and dynamic analysis modules are combined using a multimodal fusion mechanism.

Let:

F_1 denote QR visual features

F_2 denote QR structural features

F_3 denote URL semantic features

F_4 denote dynamic behavioral features

These features are integrated using an adaptive attention-based fusion strategy that assigns importance weights to each modality and generates a unified representation:

$$F_{fusion} = Fusion(F_1, F_2, F_3, F_4)$$

The fusion mechanism enables the model to learn cross-modal relationships and prioritize informative features for phishing detection.

3.5. Classification and Detection

The fused representation is passed to a fully connected classification layer that predicts whether a QR code is benign or malicious. The model is trained using supervised learning to minimize classification loss and improve generalization performance.

The integration of static and dynamic features allows the proposed framework to achieve: improved detection accuracy, robustness against evolving attacks, enhanced generalization to unseen phishing samples

4. Implementation

4.1. Dataset

The proposed framework was implemented and evaluated using the Trap4Phish 2025 dataset provided by the Canadian Institute for Cybersecurity (CIC) at the University of New Brunswick. The dataset contains benign and malicious QR codes collected for phishing and quishing analysis [9]. The dataset comprises a total of 1,005,738 QR code images, including 429,976 benign samples and 575,762 malicious samples, indicating a moderately imbalanced distribution toward malicious instances. Each QR image is associated with a generated URL stored in corresponding CSV files, enabling joint image–URL feature extraction.

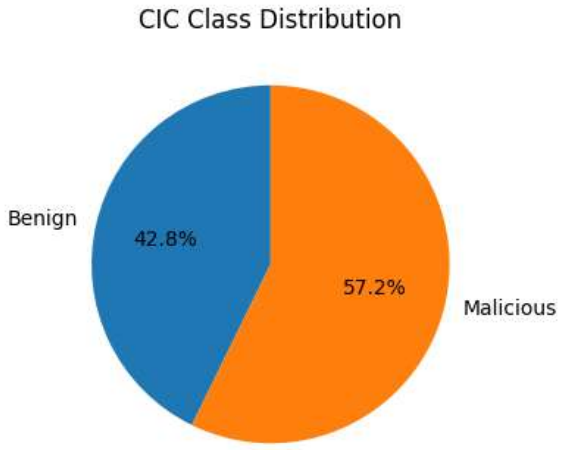


Figure 1. CIC Dataset Class Distribution

4.2. Feature Extraction

The implementation follows a multimodal design with four feature branches:

4.2.1. Visual Features F_1

QR images were normalized to the range [0,1] and reshaped to (29, 29, 1). A lightweight CNN architecture was used to extract spatial representations. Additional visual descriptors such as edge maps (Canny) and histogram statistics were computed during preprocessing.

4.2.2. Structural Features F_2

URL structural characteristics were extracted using Python's urllib.parse, including: Domain length, Number of dots, Hyphen presence, Path length, Query parameters, Encoded characters, Digit count in domain. A total of 15 structural features were generated and standardized using StandardScaler.

4.2.3. Semantic Features F_3

Semantic indicators were derived from suspicious keywords such as login, verify, secure, and password. Additional heuristics included URL length and slash count. This branch produces 12 binary/numerical features, which were also normalized.

4.2.4. Dynamic Behavioral Features F_4

Texture-based features were extracted directly from QR images using Gray-Level Co-occurrence Matrix (GLCM) statistics, including: Contrast, Dissimilarity, Homogeneity, Energy. Statistical descriptors such as mean, standard deviation, range, and inter-quantile range were also included, resulting in 8 dynamic features.

4.2.5. Multimodal Fusion Architecture

The model was implemented using TensorFlow/Keras. Each modality was processed by a dedicated branch: A CNN backbone for visual features and fully connected networks for structural, semantic, and dynamic features. Each branch outputs a 64-dimensional latent representation.

These representations are concatenated into a 256-dimensional fused vector. An attention mechanism is applied to learn adaptive weights for each modality, allowing the model to dynamically emphasize the most informative feature group. The weighted fused representation is passed through multiple fully connected layers with Batch Normalization and Dropout regularization before final binary classification using a sigmoid activation function.

4.2.6. Evaluation and Analysis

Model performance was evaluated using: Accuracy, Precision, Recall, F1-score, ROC-AUC, Confusion Matrix, Specificity. Additionally, attention weights were extracted to analyze the relative contribution of each modality (Visual, Structural, Semantic, Dynamic). A prediction function was implemented to allow real-time inference on new QR codes by extracting all four feature modalities and returning classification results with confidence scores and modality contribution weights.

5. Results and Discussion

This section presents the experimental results of the proposed multimodal fusion model evaluated on the Trap4Phish 2025 dataset from the Canadian Institute for Cybersecurity (CIC). Quantitative results are reported using standard classification metrics, followed by an analysis of modality contributions and error cases.

The model was evaluated on the held-out test set (20% of the dataset, stratified sampling). The following performance metrics were computed:

Table 1. Overall Performance on Test Set

Metric	Value

Accuracy	0.8531
Precision	0.8983
Recall	0.8383
F1-score	0.8673
ROC_AUC	0.9294
Specificity	0.8729

The ROC-AUC score indicates the discriminative capability of the model across various classification thresholds. A high AUC demonstrates that the multimodal fusion mechanism effectively separates benign and malicious QR codes.

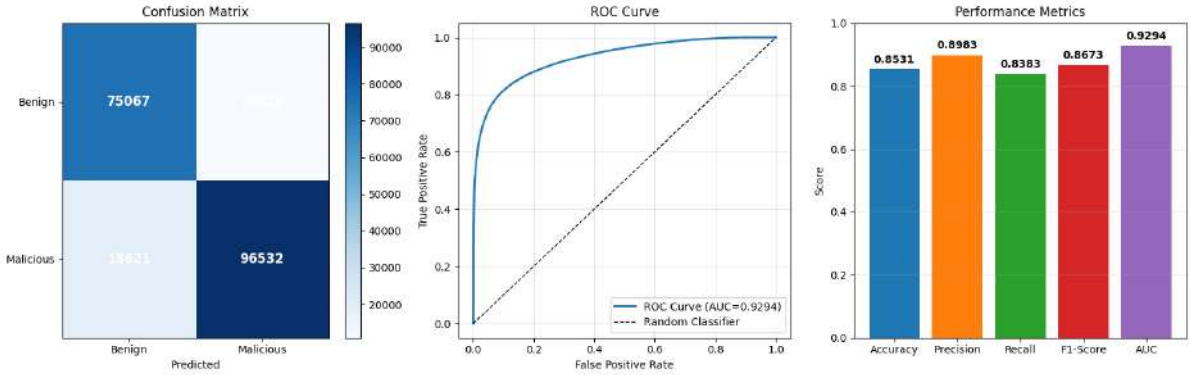


Figure 2. Fusion Model Result Summary

An attention mechanism was employed to dynamically weight each modality (F_1-F_4). The average attention weights across the test set are summarized below:

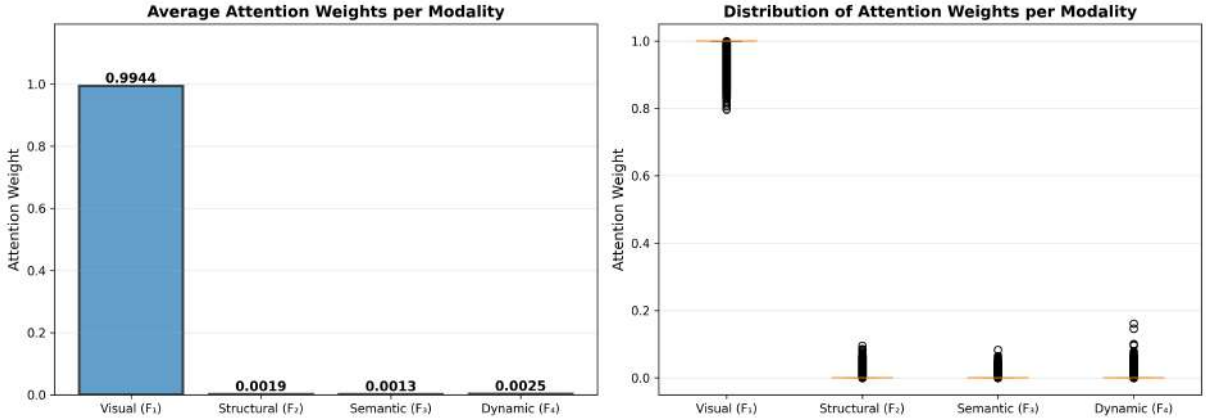


Figure 3. Attention Weights Analysis

The attention analysis reveals that the proposed fusion model assigns overwhelming importance to visual features (99.44%), while structural, semantic, and dynamic features receive negligible weights. This suggests that, within the Trap4Phish 2025 dataset, discriminative information is predominantly encoded in QR visual patterns.

6. Conclusion and Future Work

This study proposed a multimodal deep learning framework for quishing detection using the Trap4Phish 2025 dataset. By integrating visual, structural, semantic, and dynamic features

through an attention-based fusion mechanism, the model achieved strong classification performance and demonstrated robustness against diverse phishing patterns. The experimental results confirm that multimodal fusion significantly improves detection capability compared to single-modality approaches. The attention mechanism further enhances interpretability by revealing the relative contribution of each feature group.

Future research might focus on improving balanced multi-modal learning to enhance robustness and generalization. First, modality-wise normalization and attention regularization techniques (e.g., entropy regularization or temperature scaling) can be introduced to prevent attention dominance. Second, modality dropout strategies may encourage the model to learn complementary information from structural, semantic, and dynamic features. Third, robustness evaluation against adversarial QR perturbations and visually obfuscated codes should be conducted to assess real-world security resilience. Finally, extending the framework with larger and more diverse datasets, as well as incorporating transformer-based cross-modal fusion architectures, may improve adaptive feature interaction and strengthen the practical applicability of the proposed system.

References

1. Kustiawan, Y. A., & Ghauth, K. I. (2025). PhishOFE: A novel machine learning framework for real-time phishing URL detection with optimized feature engineering. IEEE Access. https://ieeexplore.ieee.org/abstract/document/11177149/
2. Tayachi, A., Ouni, B., Mourad, A., & Erbad, A. (2025). Quishing attack detection and mitigation using machine learning and deep learning for malicious url identification. 2025 International Wireless Communications and Mobile Computing (IWCMC), 1709–1713. https://ieeexplore.ieee.org/abstract/document/11059621/
3. Keerthi, K. B. S., Sunil, V., Chengappa, K. T., & Basavaraju, S. (2025). Proactive Defense Against Quishing: Integrating Predictive Analytics and Dynamic Analysis for Intelligent Threat Detection. 2025 4th International Conference on Applied Artificial Intelligence and Computing (ICAAIC), 959–966. https://ieeexplore.ieee.org/abstract/document/11331006/
4. Trad, F., & Chehab, A. (2025). Detecting Quishing Attacks with Machine Learning Techniques Through QR Code Analysis (arXiv:2505.03451). arXiv. https://doi.org/10.48550/arXiv.2505.03451
5. Akram, M. W., Sood, K., & Hassan, M. U. (2025). QRïS: A Preemptive Novel Method for Quishing Detection Through Structural Features of QR (arXiv:2510.17175). arXiv. https://doi.org/10.48550/arXiv.2510.17175
6. Guan, Y., & Zhu, E. (2025). A Dual-Branch Deep Learning Model for Phishing Detection in QR Code Scanning Scenarios. 2025 6th International Conference on Machine Learning and Computer Application (ICMLCA), 1–6. https://doi.org/10.1109/ICMLCA66850.2025.11336640
7. Certifichain: Secure QR Codes for Blockchain-Verified Digital Credentials. (n.d.). Digital Threats: Research and Practice. Retrieved February 22, 2026, from https://dl.acm.org/doi/10.1145/3754458
8. Yalda, R., Khan, A., & Nepal, N. (2026). B-PhishQR—A blockchain-based framework for secure QR code verification against phishing attacks. Telematics and Informatics Reports, 21, 100289. https://doi.org/10.1016/j.teler.2025.100289
9. Nejati, F., Rabbani, M., Eskandarian, M., Mirani, M., Piya, G., Opushnyev, I., Ghorbani, A. A., & Dadkhah, S. (2026). CIC-Trap4Phish: A Unified Multi-Format Dataset for Phishing and Quishing Attachment Detection (arXiv:2602.09015). arXiv. https://doi.org/10.48550/arXiv.2602.09015

Phát hiện tấn công lừa đảo thông qua mã QR độc hại dựa trên mô hình kết hợp đa phương thức

Tóm tắt: *Các cuộc tấn công lừa đảo thông qua mã QR, lợi dụng mã QR để phân phối nội dung độc hại, đang trở thành một mối đe dọa an toàn thông tin ngày càng nghiêm trọng. Nghiên cứu này đề xuất một khung học sâu đa phương thức nhằm phát hiện mã QR độc hại bằng cách tích hợp đặc trưng hình ảnh của mã QR, đặc tính cấu trúc của URL được nhúng, thông tin ngữ nghĩa và các thuộc tính hành vi động. Mô hình áp dụng chuẩn hóa đặc trưng và cơ chế kết hợp đa phương thức nhằm nâng cao tính bền vững và khả năng phân biệt giữa mẫu hợp lệ và mẫu độc hại. Kết quả thực nghiệm trên tập dữ liệu quy mô lớn cho thấy hiệu năng phát hiện ổn định và cân bằng. Khung phương pháp được đề xuất cũng tạo nền tảng cho các nghiên cứu tiếp theo về phát hiện mối đe dọa thích ứng và học đa phương thức có khả năng giải thích, đồng thời cho thấy tiềm năng ứng dụng thực tiễn trong các hệ thống bảo mật di động, trình duyệt web, cổng bảo mật doanh nghiệp và các ứng dụng mã QR trong môi trường thực tế.*

Từ khoá: Phát hiện lừa đảo qua mã QR; Học sâu đa phương thức; Phân tích mã QR độc hại; An toàn thông tin; Phát hiện mối đe dọa thông minh.

Khảo sát các phương pháp phát hiện lỗ hổng trên smart contract non-EVM

Ninh Thái Phan[1]

[1]*Viện Công nghệ thông tin và truyền thông, Học viện Kỹ thuật quân sự, 236 Hoàng Quốc Việt, Hà Nội, Việt Nam*

Email: phan123123@gmail.com; SĐT: 0983227135

Tóm tắt

Trong bối cảnh các hệ sinh thái blockchain non-EVM như Solana hay Aptos phát triển mạnh mẽ, việc đảm bảo an ninh cho smart contract sử dụng ngôn ngữ Rust và Move đang đối mặt với nhiều thách thức mới do sự khác biệt về kiến trúc máy ảo. Bài báo này thực hiện một khảo sát hệ thống về các phương pháp phát hiện lỗ hổng bảo mật, tập trung phân tích sự tương quan giữa các kỹ thuật kiểm thử và hiệu quả nhận diện từng loại lỗi cụ thể. Thông qua đánh giá thực nghiệm trên các phương pháp phân tích tĩnh, fuzzing, kiểm chứng hình thức và sử dụng LLM, nghiên cứu chỉ ra rằng các công cụ hiện nay vẫn tồn tại "lỗ hổng" lớn trong việc phát hiện các lỗi logic tầng thấp, đặc biệt là bitwise overflow (tràn số trong phép toán bit). Kết quả khảo sát xác định rõ thế mạnh của từng phương pháp: trong khi phân tích tĩnh hiệu quả với các lỗi cấu trúc, kiểm chứng hình thức tỏ ra ưu việt hơn trong việc xử lý các lỗi phức tạp vốn đòi hỏi độ chính xác toán học cao. Công trình này không chỉ hệ thống hóa các lỗ hổng trên môi trường non-EVM mà còn định hướng cho việc phát triển các công cụ kiểm định tự động chuyên sâu, góp phần tối ưu hóa quy trình bảo mật cho các nền tảng Web3 thế hệ mới.

Từ khoá: *Smart contract, non-EVM, Phát hiện lỗ hổng, Phương pháp kiểm định*

1. Mở đầu

Hợp đồng thông minh đóng vai trò nền tảng trong việc thiết lập các thỏa thuận phi tập trung, nhưng tính chất không thể sửa đổi của blockchain khiến các khiếm khuyết phần mềm gây ra rủi ro tài chính nghiêm trọng. Để khắc phục các nút thắt về hiệu suất của Ethereum Virtual Machine (EVM), các blockchain thế hệ mới (non-EVM) như Solana và Sui đã ra đời, sử dụng kiến trúc thực thi song song cùng các ngôn ngữ như Rust và Move. Sự thay đổi cốt lõi về mô hình quản lý bộ nhớ và cấu trúc lưu trúc trạng thái này làm vô hiệu hóa các công cụ phân tích bảo mật EVM truyền thống, đồng thời sản sinh ra các bề mặt tấn công hoàn toàn mới, đòi hỏi cộng đồng phải thiết lập các hệ thống phân loại và phương pháp phát hiện lỗ hổng chuyên biệt [1,2].

Hiện trạng nghiên cứu ghi nhận bốn hướng tiếp cận chủ đạo: (i) Phân tích tĩnh với ví dụ là VRust được phát triển để phân tích luồng dữ liệu và phát hiện lỗ hổng tự động trên Solana [3] hay MoveScan [4] và MoveScanner [5] cung cấp các khung phân tích bytecode để theo dõi luồng tài nguyên và phát hiện lỗi tương tác chéo module; (ii) Fuzzing với ví dụ là Trident được sử dụng trên Solana [6] trong khi VULSEYE tiên phong áp dụng fuzzing hướng trạng thái (stateful directed fuzzing) để giải quyết bài toán không gian trạng thái [7]; (iii) Kiểm chứng hình thức (Formal Verification) với đại diện là Move Prover nhằm chứng minh tính đúng đắn bằng mô hình toán học [8]; và (iv) LLM đã được sử dụng gần đây với ví dụ cụ thể là Smartify [9].

Các công trình nghiên cứu hiện nay tập trung vào việc tối ưu hoá hiệu quả với từng phương pháp cụ thể hoặc liệt kê lỗ hổng tiềm tàng như [1,2]. Có khoảng trống nghiên cứu được nhận ra là lựa chọn phương pháp nào với từng loại lỗ hổng để có thể xây dựng một quy trình kiểm thử an toàn cho smart contract trên non-EVM hiệu quả. Điều đó thúc đẩy chúng tôi thực hiện một khảo

sát và phân tích các phương pháp phát hiện lỗ hổng bảo mật dành riêng cho smart contract non-EVM. Giải pháp tiếp cận của chúng tôi là phân loại các kỹ thuật hiện đại (Phân tích tĩnh, Fuzzing, Xác minh hình thức, AI/LLM) và ánh xạ hiệu năng với các lớp lỗ hổng đặc thù. Từ đó, làm rõ mức độ đáp ứng của từng phương pháp đối với các điểm mù bảo mật hiện tại.

Đóng góp chính của nghiên cứu gồm: (1) Hệ thống hóa và phân tích các loại hổng bảo mật đặc trưng phát sinh từ mô hình kiến trúc đặc thù của Rust/Solana và Move/Sui; (2) Đánh giá độ hiệu quả của từng phương pháp tiếp cận đối với từng loại lỗ hổng bảo mật trên smart contract non-EVM.

Phần còn lại của bài báo được tổ chức như sau: Mục 2 trình bày các lớp lỗ hổng bảo mật; Mục 3 trình bày đánh giá độ hiệu quả của các phương pháp tiếp cận; Mục 4 kết luận và nêu hướng phát triển.

2. Phân loại các lỗ hổng bảo mật trên smart contract non-EVM

Sự dịch chuyển từ môi trường Ethereum Virtual Machine (EVM) sang các kiến trúc thực thi song song như Solana và Sui đã làm thay đổi hoàn toàn bề mặt tấn công của hợp đồng thông minh. Việc loại bỏ được các nhược điểm cố hữu của Solidity đi kèm với sự xuất hiện của các rủi ro mới gắn liền với mô hình quản lý bộ nhớ và phân định trạng thái đặc thù của Rust và Move

2.1. Các lỗ hổng đặc trưng trên hệ sinh thái Solana (Rust)

Kiến trúc của Solana tách biệt hoàn toàn mã logic và lưu trạng thái tách biệt khỏi dữ liệu. Mọi giao dịch yêu cầu người dùng phải truyền trực tiếp danh sách tài khoản vào. Sự phân tách này đòi hỏi nhà phát triển phải thực hiện các bước kiểm tra tính hợp lệ của tài khoản một cách thủ công. Dựa trên các nghiên cứu phân loại chuyên sâu về lỗ hổng hệ sinh thái Solana [2, 10], các rủi ro cốt lõi như sau:

Thứ nhất, thiếu sót trong khâu kiểm tra. Vì chương trình trên Solana chấp nhận bất kỳ tài khoản nào được truyền vào qua một mảng từ giao dịch, việc bỏ sót các bước kiểm tra xác thực là nguyên nhân gây ra nhiều rủi ro nhất, các ví dụ bao gồm: thiếu kiểm tra chữ ký, thiếu kiểm tra quyền sở hữu, thiếu kiểm tra chi phí thuê.

Một ví như bài báo [10] đã đề cập về việc thiếu sót trong khâu kiểm tra:

```
fn update_admin(program_id: &Pubkey, accounts: &[AccountInfo]) -> ProgramResult {
 let account_iter = &mut accounts.iter();
 let config = ConfigAccount::unpack(next_account_info(account_iter)?);
 let admin = next_account_info(account_iter)?;
 let new_admin = next_account_info(account_iter)?;
 //...
 // Validate the config account...
 //...
 if admin.pubkey() != config.admin {
 return Err(ProgramError::InvalidAdminAccount);
 }

 config.admin = new_admin.pubkey();

 Ok(());
}
```

Hình 1.Ví dụ về lỗ hổng thiếu sót trong khâu kiểm tra

Đoạn ví dụ trên có mục đích cập nhật quyền quản trị cho một smart contract. Tuy nhiên nó thiếu một việc kiểm tra quan trọng là kiểm tra quyền hiện tại của người gọi tính năng này có được phép thực hiện hay không. Dẫn đến bất cứ ai cũng có thể thay đổi quyền quản trị của smart contract.

Thứ hai, nhầm lẫn tài khoản và tương tác chéo không an toàn: Solana cho phép các chương trình gọi chéo lẫn nhau (Cross-Program Invocation - CPI). Nếu hợp đồng không xác minh chặt chẽ định danh của chương trình được gọi, kẻ tấn công có thể chuyển hướng luồng thực thi sang một mã độc. Tương tự, sự nhầm lẫn giữa các bản thể dữ liệu của cùng một hợp đồng cũng tạo điều kiện để vượt qua các ràng buộc logic. Các lỗi cụ thể như: nhầm lẫn loại tài khoản, tương tác chéo tuỳ ý, lỗi khởi tạo lại.

Thứ ba, lỗi tính toán số học. Ở chế độ triển khai thực tế, trình biên dịch Rust tự động thực hiện "wrapping" (quay vòng giá trị khi vượt quá giới hạn kiểu dữ liệu) thay vì báo lỗi ngắt như trong chế độ debug. Việc thiếu sử dụng các thư viện toán học an toàn (như checked_add hoặc checked_sub) khiến các sai lệch số học xảy ra một cách im lặng, tạo điều kiện cho hacker thao túng số dư.

Thực tế đã xảy ra những trường hợp bị tấn công bởi vì những lỗi trong tính toán số học có thể kể đến như lỗ hổng trên giao thức cho vay Jet-v1 với nguyên nhân là việc thực hiện phép toán cộng ra ngoài phạm vi một số u64 [11].

Thứ tư, Lỗi logic và tương tác bên ngoài: Là lỗi sử dụng mặt logic của chương trình thực thi và các tác tương với tài nguyên bên ngoài blockchain. Có thể kể đến là lỗ hổng trên Mango Markets gây thiệt hại 100 triệu USD khi không kiểm tra kĩ nguồn bên ngoài (Oracle) [10].

2.2. Các lỗ hổng đặc trưng trên hệ sinh thái Move (Sui)

Dựa trên các nghiên cứu thực nghiệm và hệ thống phân loại, các lỗ hổng bảo mật trên smart contract Move/Sui được tổng hợp thành các lớp chính yếu sau đây. Ngôn ngữ Move giải

quyết tốt các lỗi kinh điển (như reentrancy) thông qua hệ thống kiểu tuyến tính và mô hình hướng tài nguyên, nhưng lại mở ra những bề mặt tấn công mới [12].

Thứ nhất, lỗ hổng quản lý trạng thái và đối tượng. Loại này bao gồm các sai sót trong việc quản lý vòng đời đối tượng, tình trạng tương tranh (race conditions) trong môi trường thực thi song song, và cấu hình sai quyền sở hữu đối tượng (nhầm lẫn giữa đối tượng sở hữu cá nhân - owned objects và đối tượng chia sẻ - shared objects).

Thứ hai, lỗ hổng trong quản lý vòng đời đối tượng. Lỗ hổng loại này là các khiếm khuyết logic làm sai lệch các bất biến kinh tế, liên quan đến việc đúc, đốt, khóa hoặc phân phối tài sản không hợp lệ.

Thứ ba, lỗ hổng về mặt toán học. Dù Move tích hợp sẵn cơ chế kiểm tra tràn số đối với các phép toán cộng, trừ, nhân, chia; tuy nhiên, các phép toán dịch bit hoàn toàn nằm ngoài cơ chế bảo vệ này. Ví dụ cụ thể và nổi tiếng nhất có thể đưa ra là sự cố của Cetus khi xảy ra lỗ hổng trong tính toán với phép tính bit và gây thiệt hại hơn 200 triệu USD [13]. Trong đoạn code nơi xảy ra lỗi:

```
public fun checked_shlw(value: u256) : (u256, bool) {
  if (value > 11579208923731619541729388327330122708943419524243289762335522856344909
    (0, true)
  } else {
    (value << 64, false)
  }
}
```

Hình 2. Đoạn chương trình gây lỗ hổng trên Cetus

Việc kiểm tra ngưỡng trước khi tiến hành dịch bit bị tính toán sai do người lập trình, vì thế sẽ có khoảng giá trị khiến cho "value" bị mất mát bit có ý nghĩa khiến sai lệch số liệu không mong muốn sau khi dịch bit trái.

Thứ tư, lỗ hổng liên quán đến bất biến module. Xảy ra do sự thiếu đồng bộ hoặc xác thực lỏng lẻo khi các hợp đồng kết hợp với nhau. Các tài nguyên bị rò rỉ hoặc bị thao túng qua ranh giới của các module. Sự cố liên quan đến tính bất biên module được phát hiện trên smart contract của sàn giao dịch phi tập trung giấu tên đã có thể bị tấn công ước tính thiệt hại 160 triệu USD [12].

3. Các phương pháp phát hiện lỗ hổng trên smart contract non-EVM

3.1. Phân tích tĩnh

Phương pháp này phân tích mã nguồn hoặc bytecode mà không cần thực thi chương trình, giúp quét toàn diện và phát hiện sớm các rủi ro. Tuy nhiên, nó thường gặp thách thức về tỷ lệ dương tính giả (false positives) và khó xử lý các giá trị động.

- Trên Solana: VRust là một khung phân tích tự động dựa trên luồng dữ liệu tĩnh để kiểm tra các giao thức an toàn thiết yếu. Nó cực kỳ hiệu quả trong việc tự động phát hiện các lỗi phổ biến nhất trên Solana như thiếu kiểm tra chữ ký và quyền sở hữu bằng cách phân tích biểu diễn trung gian (IR). Tuy nhiên, VRust bị giới hạn bởi việc bắt buộc phải có mã nguồn và tỷ lệ báo động giả cao [3].

- Trên Sui/Move: MoveScan và MoveScanner thể hiện sự đột phá. MoveScan chuyển đổi bytecode sang dạng biểu diễn trung gian (IR) để quét 8 loại khiếm khuyết với tốc độ cực nhanh (trung bình 5.45 mili giây/dự án) và độ chính xác lên tới 98.85% [4]. Nâng cấp hơn, MoveScanner áp dụng kiến trúc đồ thị luồng điều khiển (CFG) và phân tích luồng dữ liệu xuyên module. Nó tích hợp thuật toán theo dõi quỹ đạo tài nguyên và phân tích ma trận khả năng, giúp phát hiện chính xác lỗi rò rỉ tài nguyên, rò rỉ đặc quyền và thao túng trạng thái chéo, đồng thời giảm đáng kể tỷ lệ dương tính giả [5].

Phân tích tĩnh phù hợp với các lỗi đơn giản đã có sẵn những khung phát hiện đặc trưng trên mã nguồn hoặc bytecode. Cách tiếp cận này tỏ ra thiếu hiệu quả với các lỗ hổng logic phức tạp, không được tượng trưng hoá và đặc biệt có tỉ lệ dương tính giả cao vì chỉ dựa vào các đặc điểm và không có chứng minh.

3.2. Fuzzing

Fuzzing giải quyết điểm yếu của phân tích tĩnh bằng cách bơm các đầu vào bất thường để kích hoạt lỗi tại thời điểm chạy.

- Trident (Solana): Việc fuzzing trên Solana rất khó khăn do mô hình thực thi không trạng thái và rủi ro nhầm lẫn tài khoản. Trident giải quyết vấn đề này bằng phương pháp Manually Guided Fuzzing (MGF). Thay vì fuzzing hộp đen mù quáng, nó cho phép các nhà nghiên cứu xây dựng các kịch bản tương tác chéo chương trình (CPI) phức tạp và tiến hành kiểm tra các bất biến (invariant checks). Nó đặc biệt mạnh trong việc dò tìm các lỗi giả mạo tài khoản, nhưng đòi hỏi có các cài đặt thủ công từ chuyên gia [6].

- VULSEYE: Công cụ này giải quyết triệt để "bài toán không gian trạng thái" thông qua Fuzzing hướng trạng thái (Stateful Directed Fuzzing). Bằng thuật toán phân tích ngược trên bytecode, VULSEYE xác định trước các mục tiêu trạng thái cần thiết để kích hoạt lỗi và các mã mục tiêu, sau đó dồn tài nguyên fuzzing vào đúng các vùng rủi ro này thay vì rải rác lãng phí. Điều này mang lại hiệu suất phát hiện lỗi vượt trội (gấp 9.7 lần) so với các fuzzer truyền thống trên các tập dữ liệu thực tế [7].

Fuzzing giúp kiểm tra các lỗ hổng phức tạp, khó có thể xác định khung lỗ hổng (dấu hiệu). Tuy nhiên vấn đề fuzzing gặp phải là có thể thiếu sót trường hợp.

3.3. Xác minh hình thức

Phương pháp này dùng mô hình toán học để chứng minh tính đúng đắn tuyệt đối của hợp đồng so với các đặc tả.

- Move Prover (Sui/Move): Đây là công cụ tích hợp sẵn của ngôn ngữ Move, tự động dịch mã Move sang ngôn ngữ trung gian Boogie và dùng bộ giải SMT (như Z3) để xác minh các điều kiện trước và sau. Move Prover cực kỳ hiệu quả trong việc chứng minh tính an toàn của tài nguyên. Tuy nhiên nó đòi hỏi lập trình viên phải tự viết các đặc tả toán học phức tạp [8]. Nghiêm trọng hơn, Move Prover rất chật vật với các phép toán dịch bit. Lỗ hổng dẫn đến vụ hack 200 triệu USD của Cetus AMM chính là lỗi tràn ngầm định trong phép toán dịch bit mà các công cụ kiểm chứng đã không thể phát hiện.

Phương pháp này đem lại ưu điểm tuyệt đối trong chứng minh logic và đưa ra các chứng dụ chính xác cho phép giảm thiểu rủi ro của việc dương tính giả gần như bằng không. Tuy nhiên phương pháp này yêu cầu tài nguyên lớn để có thể giải các điều kiện của đặc tả nhiều khi dẫn tới không có kết quả.

3.4. AI

Sự tham gia của AI, tiêu biểu là LLM, đang định hình lại cách phân tích bảo mật nhờ khả năng hiểu ngữ nghĩa nghiệp vụ thay vì chỉ quét cú pháp.

- Smartify: Là một khung làm việc đa tác nhân sử dụng các LLM được tinh chỉnh riêng kết hợp với kỹ thuật RAG (Retrieval-Augmented Generation) [9]. Công cụ này đã có thể phát hiện và sửa lỗi các lỗi về quản lý đối tượng, tính toán tuy nhiên chưa đề cập đến việc xác định các lỗ hổng module khi có tương tác chéo phức tạp.

- Nghiên cứu [14] đã áp dụng LLM với kĩ thuật tinh chỉnh mô hình và kĩ năng tạo lệnh để phát hiện một số loại lỗ hổng đặc trực trên Solana và đạt được kết quả với độ chính xác từ 57% đến 63%. Tuy nhiên phương pháp này chưa thể kiểm tra các loại lỗ hổng cần được chính minh cụ thể như tương tác với oracle.

LLM cho thấy sự linh hoạt trong phát hiện các loại lỗ hổng khác nhau tuy nhiên phương pháp LLM hiện tại vẫn đối mặt với rủi ro diễn giải sai mã và thiếu chứng minh tính đúng đắn toán học một cách rạch ròi. Vì thế phương pháp này hiện chưa được áp dụng phát hiện các lỗ hổng có tương tác cụ thể giữa nhiều hợp đồng qua lại.

3.5. Sự phù hợp của các phương pháp với các loại lỗ hổng

Từ các kết quả của các bài báo cũng như đặc thù của các loại lỗ hổng, chúng tôi đưa ra bảng đánh giá như sau:

Bảng 1. Ma trận đánh giá các phương pháp tiếp cận và các loại lỗ hổng trên Move/Sui

	quản lý trạng thái & đối tượng	**quản lý vòng đời đối tượng**	**lỗ hổng toán học**	**lỗ hổng liên quan đến bất biến module**
Phân tích tĩnh	Hạn chế	Rất phù hợp	Không phù hợp	Phù hợp
Fuzzing	Rất phù hợp	Không phù hợp	Phù hợp	Phù hợp
Xác minh hình thức	Phù hợp	Rất phù hợp	Phù hợp	Rất phù hợp
AI/LLM	Phù hợp	Phù hợp	Phù hợp	Không phù hợp

Bảng 2. Ma trận đánh giá các phương pháp tiếp cận và các loại lỗ hổng trên Solana/Rust

	Thiếu kiểm tra	**Nhầm lẫn tài khoản & Tương tác chéo không an toàn**	**lỗ hổng toán học**	**Lỗi Logic & tương tác bên ngoài**
Phân tích tĩnh	Rất phù hợp	Không phù hợp	Không phù hợp	Không phù hợp
Fuzzing	Phù hợp	Rất phù hợp	Phù hợp	Không phù hợp
Xác minh hình thức	Phù hợp	Không phù hợp	Phù hợp	Rất phù hợp
AI/LLM	Phù hợp	Phù hợp	Phù hợp	Không phù hợp

Theo hai bảng trên cho thấy phương pháp phân tích tĩnh phù hợp các lỗ hổng cụ thể phát hiện theo những đặc điểm hoặc luồng chương trình. Fuzzing phù hợp với những lỗ hổng phức tạp không có định dạng cố định tuy nhiên sẽ không đảm bảo tính bao phủ và để sót các trường hợp. Xác minh hình thức rất phù hợp trong kiểm tra logic và đảm bảo tính chính xác với ví dụ cụ thể được đưa ra tuy nhiên tiêu tốn nhiều tài nguyên. Đây cũng là phương pháp phù hợp nhất và dược sử dụng phổ biến nhất trong phát hiện lỗ hổng bảo mật trên smart contract. Về phương pháp tiếp cận sử dụng AI, hiện tại còn rất hạn chế về nghiên cứu và chủ yếu dùng để hỗ trợ hoặc tự động hoá các quy trình nhỏ và kết hợp các phương pháp cổ điển, vì thế chúng tôi không đưa ra nhận xét. Tuy nhiên có một số điểm mù đặc thù mà hiện nay các phương pháp đều chưa hiệu quả trong việc phát hiện như lỗ hổng với các phép tính bit đã nêu ở phần 2.

4. Kết luận và hướng nghiên cứu tiếp theo

Trong công trình này, chúng tôi đã giải quyết thách thức bảo mật đặc thù trên các nền tảng smart contract Non-EVM (tiêu biểu là Solana và Sui) bằng cách hệ thống hóa các danh mục lỗ hổng mới và đề xuất việc đánh giá, lựa chọn phương pháp phát hiện dựa trên việc phân tích chéo 4 hướng tiếp cận chủ đạo.

Các kết quả phân tích lý thuyết và thực nghiệm cho thấy không có một công cụ đơn lẻ nào có thể giải quyết trọn vẹn không gian rủi ro: Phân tích tĩnh (như MoveScanner, VRust) vượt trội về tốc độ quét lỗi cấu trúc tĩnh nhưng không thể phát hiện lỗ hổng về logic động; Fuzzing (như Trident, VULSEYE) cực kỳ hiệu quả trong việc tìm lỗi logic nghiệp vụ sâu và trạng thái tương tranh nhưng không đảm bảo sự bao phủ; Kiểm chứng hình thức đảm bảo tính đúng đắn toán học cho tài nguyên tuyến tính nhưng rất tốn tài nguyên; trong khi LLM mở ra khả năng tự động hóa hiểu ngữ nghĩa và phù hợp phát hiện đa số các loại lỗ hổng nhưng cần kiểm soát rủi ro ảo giác và thiếu đi tính xác thực chính xác khi cần kiểm chứng qua thao tác cụ thể. Điều này khẳng định tính tất yếu của việc sử dụng các hệ thống phòng thủ lai thay vì phụ thuộc vào một công cụ đơn lẻ. Phát hiện này không chỉ cung cấp cái nhìn toàn diện cho các nhà phát triển và kiểm toán viên trong việc lựa chọn công cụ kiểm tra tương thích với từng loại lỗ hổng, mà còn mở ra một hướng tiếp cận mới trong việc xây dựng các bộ máy phân tích bảo mật tự động, đóng vai trò nền tảng bảo vệ tài sản số trong các hệ sinh thái blockchain non-EVM.

Đáng chú ý, quá trình khảo sát đã bộc lộ một "điểm mù" nghiêm trọng trong hệ sinh thái công cụ hiện tại: khả năng phát hiện sai lệch trong các phép toán thao tác bit. Dù các ngôn ngữ như Move hay Rust đã tích hợp sẵn cơ chế chống tràn số cho các phép toán số học cơ bản, các phép dịch bit lại nằm ngoài cơ chế bảo vệ này. Các phương pháp hiện hành đều tỏ ra yếu thế trước lỗ hổng này: phân tích tĩnh thiếu ngữ cảnh thực thi động để đánh giá giá trị biên; fuzzing truyền thống thường bỏ lọt các trường hợp ranh giới cực đoan; và ngay cả kiểm chứng hình thức cũng gặp khó khăn lớn không thể mô hình hóa chính xác ngữ nghĩa ở mức bit. Điển hình là việc lỗ hổng dịch bit trên Cetus AMM đã dễ dàng lọt qua các khâu kiểm định. Từ phát hiện này chúng tôi sẽ dự định thiết kế bộ mã nguồn mở mang tính toàn vẹn hơn khắc phục các điểm yếu cố hữu của các phương pháp đã khảo sát, đồng thời giải pháp cho bài toàn “điểm mù” về các phép tính toán bit.

Tài liệu tham khảo

1. Topal, S. (2025). A Systematic Classification of Vulnerabilities in MoveEVM Smart Contracts (MWC). arXiv preprint arXiv:2505.19047.
2. Bagheri, M., Park, S.-H., & Kwon, D.-H. (2024). Study of Vulnerabilities in Solana Smart Contracts. Annual Conference of KIPS, 371-372.
3. Cui, S., Zhao, G., Gao, Y., & Tavu, T. (2022). VRust: Automated Vulnerability Detection for Solana Smart Contracts. Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security (CCS '22), 639-652.
4. Song, S., Chen, J., Chen, T., Luo, X., Li, T., Yang, W., ... & Li, P. (2024). Empirical Study of Move Smart Contract Security: Introducing MoveScan for Enhanced Analysis. Proceedings of the 33rd ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA 2024), 1682-1694.
5. Lu, Y., Li, Z., & Li, X. (2025). MoveScanner: Analysis of Security Risks of Move Smart Contracts. arXiv preprint arXiv:2508.17964.
6. Ackee Blockchain. (2026). Introducing Trident: The First Open-Source Fuzzer for Solana Programs.
7. Liang, R., Chen, J., Wu, C., He, K., Wu, Y., Cao, R., ... & Zhao, Z. (2024). VULSEYE: Detect Smart Contract Vulnerabilities via Stateful Directed Graybox Fuzzing. IEEE Transactions on Information Forensics and Security.
8. J. Zhong et al., "The Move Prover," in Proc. 32nd International Conference on Computer Aided Verification (CAV), Los Angeles, CA, USA, 2020, pp. 137–150.
9. Karanjai, R., Blackshear, S., Xu, L., & Shi, W. (2025). Smartify: A Multi-Agent Framework for Automated Vulnerability Detection and Repair in Solidity and Move Smart Contracts. arXiv preprint arXiv:2502.18515.
10. Wu, X., Xing, J., & Li, X. (2025). Exploring Vulnerabilities and Concerns in Solana Smart Contracts. arXiv preprint arXiv:2504.07419.
11. Sec3 (2024). https://medium.com/coinmonks/on-a-20m-bug-in-jet-protocol-f8581599328e
12. Topal, S. (2025). A Systematic Classification of Vulnerabilities in MoveEVM Smart Contracts (MWC). arXiv preprint arXiv:2505.19047v1.
13. Cryfin (2025). https://www.cyfrin.io/blog/inside-the-223m-cetus-exploit-root-cause-and-impact-analysis.
14. Boi, B., & Esposito, C. (2025, October). Prompt Engineering vs. Fine-Tuning for LLM-Based Vulnerability Detection in Solana and Algorand Smart Contracts. In 2025 7th International Conference on Blockchain Computing and Applications (BCCA) (pp. 391-398). IEEE.

A Survey of Vulnerability Detection Methods for Non-EVM Smart Contracts

Abstract: *As non-EVM blockchains like Solana and Aptos grow, securing Rust and Move smart contracts faces new challenges due to different virtual machine architectures. This paper surveys security vulnerability detection methods and analyzes how different testing techniques perform against specific bugs. Through testing static analysis, fuzzing, formal verification and LLM approach, the research shows that current tools struggle with low-level logic errors, especially bitwise overflow. The results highlight that while static analysis is good for structural flaws, formal verification is better for complex, mathematical bugs. This work categorizes non-EVM vulnerabilities and helps guide the development of better automated security tools for Web3.*

Keywords: Smart contract; non-EVM, Vulnerability detection; Analysis techniques

Đánh giá một số mô hình đa phương thức cho bài toán phát hiện tin giả

Nguyễn Thanh Hà[1], Phan Việt Anh[2], Nguyễn Trung Tín[3]
Viện Công nghệ thông tin và Truyền thông, Học viện Kỹ thuật Quân sự
Email: hand3007@gmail.com; Tel: 0918223862

Tóm tắt

Trong bối cảnh sự phát triển mạnh mẽ của mạng xã hội, vấn đề tin giả đã trở thành vấn nạn nguy hại đến cộng đồng, mục đích đánh lừa công chúng, thao túng dư luận, gây chia rẽ xã hội, thậm chí can thiệp vào chính trị và nền kinh tế. Hậu quả của nó là vô cùng nghiêm trọng, từ việc phá hoại niềm tin vào thể chế, bóp méo nhận thức cộng đồng, cho đến những thiệt hại cụ thể về sức khỏe, tài chính và an ninh quốc gia, như trong đại dịch COVID-19 hay các cuộc bầu cử trên thế giới. Việc phát hiện tin giả không chỉ giúp bảo vệ người dùng mạng xã hội mà còn đảm bảo tính an toàn và minh bạch của các nền tảng trực tuyến. Nghiên cứu này khảo sát mô hình đa phương thức Universal Image-Text Representation (UNITER) cho bài toán phát hiện tin giả và so sánh với một số mô hình đa thể thức phổ biến: Mô hình kết hợp muộn (Late Fusion), VisualBERT, VL-BERT, Contrastive Language-Image Pre-training (CLIP). Kết quả nghiên cứu nhằm đánh giá khả năng áp dụng mô hình đa thể thức UNITER so với các mô hình đa thể thức phổ biến trong việc phát hiện tin giả. Từ đó giúp các nhà phát triển và quản lý nền tảng lựa chọn công nghệ phù hợp nhằm nâng cao độ tin cậy của các thông tin trên mạng xã hội.

***Từ khóa**: Universal Image-Text Representation; Late Fusion; VisualBERT; VL-BERT; Contrastive Language-Image Pre-training; Fake news detection.*

1. Giới thiệu

Sự bùng nổ của mạng xã hội và các nền tảng trực tuyến đã tạo điều kiện cho tin giả lan truyền với tốc độ chóng mặt, vượt xa các kênh thông tin chính thống. Nguyên nhân là do tính dễ tiếp cận, khả năng chia sẻ nhanh và sự thiếu hụt các cơ chế kiểm duyệt nội dung hiệu quả trên các nền tảng này. Những thông tin sai lệch có thể gây ra những hệ lụy sâu rộng đến chính trị, sức khỏe cộng đồng và sự ổn định xã hội.

Trước thách thức đó, việc phát triển các hệ thống tự động phát hiện tin giả đã trở thành một lĩnh vực nghiên cứu cấp thiết trong ngành khoa học máy tính, đặc biệt là xử lý ngôn ngữ tự nhiên và học máy. Mặc dù đã có nhiều nghiên cứu đáng kể với các phương pháp tiếp cận từ truyền thống dựa trên đặc trưng đến hiện đại dựa trên mô hình học sâu, bài toán phát hiện tin giả vẫn là một thách thức không dễ dàng. Sự đa dạng và tính thích ứng của tin giả từ các trang tin giả mạo, tiêu đề câu lượt tương tác, bài đăng trên mạng xã hội được tài trợ, cho đến các video, hình ảnh bị chỉnh sửa đòi hỏi các mô hình không chỉ phân tích nội dung văn bản mà còn phải xem xét ngữ cảnh, nguồn gốc, lan truyền và cảm xúc của người tiếp nhận. Hơn nữa, sự mơ hồ trong định nghĩa, sự thiếu hụt các bộ dữ liệu chất lượng, đa ngôn ngữ và vấn đề thiên lệch trong dữ liệu huấn luyện càng làm phức tạp thêm nhiệm vụ này.

Các phương pháp truyền thống dựa trên đặc trưng thủ công thường tỏ ra kém hiệu quả trước sự đa dạng và tinh vi ngày càng tăng của tin giả. Những năm gần đây, sự xuất hiện của các mô hình đa thể thức phổ biến như VisualBERT, VL-BERT, CLIP,… được đưa vào áp dụng trong bài toán phát hiện tin giả trên dữ liệu mạng xã hội đã cải thiện đáng kể hiệu suất phát hiện

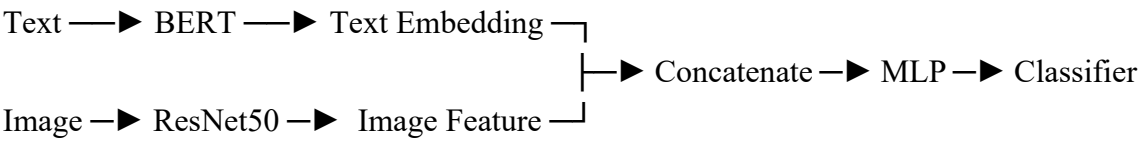
[1] *Email: hand3007@gmail.com*
[2] *Email: anhpv@lqdtu.edu.vn*
[3] *Email: tinnt@lqdtu.edu.vn*

tin giả. Nghiên cứu này đề xuất sử dụng mô hình đa phương thức Universal Image-Text Representation (UNITER) cho bài toán phát hiện tin giả, so sánh kết quả với các mô hình đa thể thức khác, cho thấy mô hình đa phương thức UNITER có kết quả tốt hơn các mô hình đa thể thức còn lại.

2. Mô hình đa thể thức

2.1. Mô hình sử dụng phương thức Late Fusion

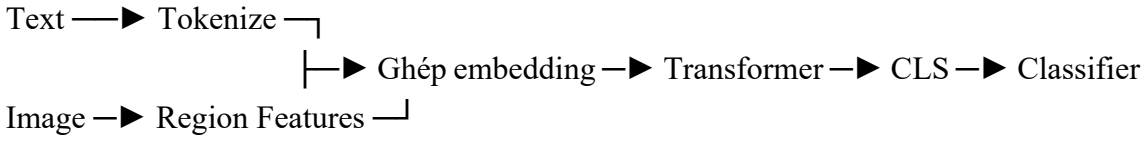
Late fusion [1] là một kỹ thuật trong đó mỗi loại dữ liệu được xử lý một cách hoàn toàn độc lập bằng các mô hình riêng biệt. Chỉ sau khi mỗi mô hình đã đưa ra dự đoán riêng của mình, các dự đoán này mới được kết hợp ở giai đoạn cuối để đưa ra quyết định chung cuối cùng. Ở đây chúng tôi sử dụng BERT [2] để trích xuất đặc trưng văn bản và RestNet50 [3] để trích xuất đặc trưng văn bản, sau đó hợp nhất lại và đưa ra dự đoán cuối cùng.



Văn bản trước hết được tiền xử lý và đưa vào mô hình BERT để thực hiện mã hóa ngữ cảnh. Cụ thể, chuỗi văn bản được mã hóa, bổ sung khóa đặc biệt [CLS], sau đó chuyển thành véc-tơ và truyền qua các tầng Transformer của BERT. Tại đầu ra, véc-tơ tương ứng với khóa [CLS] được sử dụng làm biểu diễn ngữ nghĩa toàn cục của toàn bộ văn bản, với kích thước 768 chiều. Song song đó, hình ảnh đầu vào (kích thước chuẩn hóa 224×224×3) được đưa qua mạng tích chập sâu ResNet50 nhằm trích xuất đặc trưng thị giác mức cao. Thông qua các khối residual và cơ chế giảm kích thước không gian kết hợp với Global Average Pooling, mô hình sinh ra một vector đặc trưng 2048 chiều đại diện cho nội dung hình ảnh. Hai vector đặc trưng từ văn bản và hình ảnh sau đó được hợp nhất bằng phép nối để tạo thành một vector đặc trưng chung, phản ánh thông tin đa phương thức của bản tin. Vector hợp nhất này được đưa vào một mạng phân loại Multi-Layer Perceptron gồm các lớp tuyến tính, hàm kích hoạt phi tuyến và dropout nhằm học mối quan hệ tương tác giữa hai mô hình và ánh xạ sang không gian nhãn nhị phân. Cuối cùng, thông qua hàm Softmax, hệ thống dự đoán xác suất bản tin thuộc lớp "Tin giả" hoặc "Tin thật", từ đó đưa ra quyết định phân loại.

2.2. VisualBERT

VisualBERT [4, 5] là một mô hình Transformer đa phương thức mở rộng từ BERT để xử lý văn bản và hình ảnh trong cùng một kiến trúc, dựa trên phương thức kết hợp sớm. Kiến trúc của VisualBERT bắt đầu từ việc xử lý đầu vào đa phương thức: Văn bản được chia nhỏ thành các mã và chuyển thành các vector embedding (token embedding, segment embedding, và position embedding); Hình ảnh được trích xuất đặc trưng của từng vùng (region) quan trọng trong ảnh. Mỗi vùng này trở thành một mã ảnh và được chuyển thành visual embedding (bao gồm đặc trưng vùng, segment embedding đặc biệt cho ảnh, và position embedding). Sau đó tất cả các mã văn bản và mã ảnh được nối lại với nhau thành một chuỗi duy nhất và đưa vào các tầng Transformer. Sức mạnh của mô hình nằm ở cơ chế self-attention trong Transformer. Cơ chế này cho phép mã văn bản học quan hệ chéo với mã ảnh (và ngược lại) xuyên suốt các tầng xử lý. Nhờ đó, mô hình có thể tự động học cách căn chỉnh các từ với các vùng ảnh tương ứng, ví dụ như từ con chó sẽ liên kết mạnh mẽ với vùng ảnh chứa con chó.



2.3. VL_BERT

VL-BERT [6] là một mô hình Transformer đa phương thức được thiết kế để học biểu diễn chung giữa văn bản và hình ảnh. VL-BERT là mô hình mở rộng từ BERT, nhưng thay vì chỉ xử lý văn bản, VL-BERT có thể hiểu mối quan hệ giữa từ ngữ và các vùng trong ảnh. VL-BERT khác VisualBERT ở chỗ có thêm đặc trưng thị giác, nghĩa là VL-BERT trích xuất đặc trưng vùng ảnh quan trọng và toàn ảnh, sau đó mã hóa tọa độ toàn bộ các vùng ảnh đó, giúp việc học có thể hiểu được bố cục toàn ảnh sau mã hóa.

Trong kiến trúc VL-BERT, văn bản và các vùng đặc trưng hình ảnh được ánh xạ thành chuỗi khóa chung và đưa vào một Transformer đa phương thức. Thông qua cơ chế self-attention xuyên phương thức, mô hình học được quan hệ tương tác giữa các khóa văn bản và các vùng hình ảnh. Biểu diễn của khóa [CLS] sau tầng Transformer cuối cùng được sử dụng làm đặc trưng hợp nhất để thực hiện phân loại nhị phân tin giả và tin thật.

2.4. CLIP

CLIP [7, 8] là một mô hình học sâu đa phương thức do OpenAI phát triển, CLIP học cách hiểu ảnh không phải qua các nhãn cố định mà thông qua mô tả bằng văn bản đi kèm. CLIP được xây dựng với hai bộ mã hóa song song (bộ mã hóa hình ảnh có nhiệm vụ trích xuất đặc trưng từ ảnh rồi biến nó thành một véc-tơ và bộ mã hóa văn bản dùng để biến câu mô tả thành một véc-tơ), hoạt động cùng nhau để tạo ra một không gian biểu diễn chung cho cả ảnh và chữ. CLIP học tương phản trên một bộ dữ liệu khổng lồ gồm 400 triệu cặp ảnh - chú thích được thu thập từ internet.Trong quá trình huấn luyện, mô hình học cách kéo gần véc-tơ của ảnh và véc-tơ của câu chú thích đúng lại với nhau. Đồng thời, nó học cách đẩy xa véc-tơ của ảnh và véc-tơ của các câu chú thích sai ra. Nhờ cách học này, CLIP có thể phân loại ảnh mà không cần huấn luyện lại.

2.5. UNITER

UNITER [9-12] là mô hình học sâu đa phương thức do Microsoft phát triển, ra mắt vào năm 2019. Điểm mạnh của UNITER là khả năng học một biểu diễn chung cho cả hình ảnh và văn bản thông qua quá trình tiền huấn luyện trên quy mô lớn. Mục tiêu của nó là tạo ra một mô hình tổng quát có thể tinh chỉnh cho nhiều tác vụ khác nhau mà không cần thay đổi kiến trúc. UNITER cho phép các đặc trưng của ảnh và chữ tương tác sâu với nhau xuyên suốt các tầng của mô hình. Cấu trúc của UNITER gồm ba thành phần chính:

- Bộ mã hóa hình ảnh: Sử dụng mô hình phát hiện vật thể đã được huấn luyện trước để trích xuất các vùng quan trọng trong ảnh. Với mỗi vùng, nó trích xuất hai loại đặc trưng là đặc trưng thị giác (hình dáng, màu sắc, kết cấu của vùng đó) và đặc trưng vị trí (tọa độ, chiều rộng, chiều cao, diện tích của vùng đó trong ảnh). Cả hai đặc trưng này được chiếu qua các lớp riêng, rồi cộng lại để tạo thành véc-tơ biểu diễn cuối cùng cho mỗi vùng ảnh.

- Bộ mã hóa văn bản: Sử dụng phương pháp WordPiece để chia câu thành các khóa. Mỗi khóa được tổng hợp từ word embedding và position embedding để tạo thành véc-tơ biểu diễn ban đầu.

- Transformer đa tầng: Các véc-tơ biểu diễn của tất cả các vùng ảnh và tất cả các khóa văn bản được nối tiếp nhau thành một chuỗi duy nhất và đưa vào Transformer. Nhờ cơ chế self-attention, mỗi vùng ảnh có thể học quan hệ đến tất cả các khóa văn bản và ngược lại. Qua nhiều tầng Transformer, mô hình học được các mối quan hệ chéo sâu sắc giữa hai loại dữ liệu. Ví dụ, token "con chó" sẽ tìm cách kết nối với vùng ảnh có chứa con chó, và hiểu rõ hơn về ngữ cảnh của nó.

Nhiệm vụ của tiền huấn luyện trong mô hình UNITER là rất quan trọng và cũng chính là đặc trưng của UNITER, gồm có 04 bước:

- Masked Language Modeling (MLM): Che ngẫu nhiên một số token trong câu văn bản. Dựa vào các token còn lại và toàn bộ vùng ảnh, mô hình phải dự đoán token đã bị che. Mục tiêu học cách văn bản liên quan đến toàn cảnh bức ảnh.
- Masked Region Modeling (MRM): Che ngẫu nhiên một số vùng ảnh, dựa vào các vùng ảnh còn lại và toàn bộ câu văn bản, mô hình phải dự đoán đặc trưng của vùng ảnh đã bị che. Mục tiêu học cách hình ảnh liên quan đến nội dung của văn bản.
- Image-Text Matching (ITM): Dự đoán xem một cặp ảnh - văn bản đầu vào có khớp với nhau hay không (nhãn nhị phân: 0 hoặc 1). Mục tiêu học sự liên kết ở cấp độ toàn cục giữa ảnh và câu mô tả.
- Word-Region Alignment (WRA): Sử dụng thuật toán vận tải tối ưu để tìm ra cách vận chuyển các đặc trưng của các từ đến các đặc trưng của các vùng ảnh sao cho chi phí là thấp nhất. Điều này buộc mô hình phải học sự liên kết chi tiết, chặt chẽ giữa từng từ với từng vùng ảnh tương ứng (ví dụ: từ "mèo" phải liên kết với vùng ảnh con mèo). Mục tiêu tạo ra sự liên kết tinh tế ở cấp độ từ-vùng ảnh.

Trong các tác vụ Masked Language Modeling và Masked Region Modeling, UNITER sử dụng kỹ thuật conditional masking: chỉ che một phương thức (text trong MLM, image regions trong MRM) và giữ nguyên phương thức còn lại . Điều này giúp quá trình học tập trung và hiệu quả hơn, tránh việc che cả hai phương thức cùng lúc gây ra sự mất mát thông tin và khó khăn trong việc căn chỉnh.

3. Thực nghiệm

3.1. Chuẩn bị dữ liệu

Bộ dữ liệu sử dụng trong nghiên cứu này là Fakeddit, bao gồm 8210 mẫu ảnh có chứa văn bản. Dữ liệu được chia thành hai tập: 4000 mẫu dùng để huấn luyện và 4210 mẫu để kiểm tra, đảm bảo sự phân phối đồng đều giữa hai lớp và duy trì tính đại diện của tập dữ liệu. Quá trình xử lý này giúp nâng cao hiệu suất mô hình, tối ưu hóa khả năng phân loại và giảm thiểu tác động của dữ liệu nhiễu.

3.2. Huấn luyện mô hình

Quá trình huấn luyện sử dụng các mô hình đa thể thức được thực hiện dựa trên tập dữ liệu đã tiền xử lý, với mục tiêu phân biệt tin thật, tin giả. Các mô hình đa thể thức được triển khai bao gồm: Mô hình sử dụng Late Fusion, VisualBERT, VL-BERT, CLIP, UNITER.

4. Kết quả và thảo luận

4.1. Độ đo

Trong bài toán phát hiện tin giả, các chỉ số Accuracy, Precision, Recall, và F1-Score được sử dụng để đánh giá hiệu suất của các mô hình. Ý nghĩa cụ thể của từng chỉ số trong ngữ cảnh bài toán này như sau:

- Accuracy phản ánh tỷ lệ dự đoán đúng (bao gồm cả tin thật và tin giả) so với tổng số dữ liệu. Tuy nhiên, nếu dữ liệu không cân bằng, chỉ số này có thể không phản ánh đúng hiệu quả thực tế của mô hình.

- Precision đo lường tỷ lệ các tin thật bị dự đoán là giả. Precision cao đồng nghĩa với việc mô hình ít nhầm lẫn tin giả, tin thật.

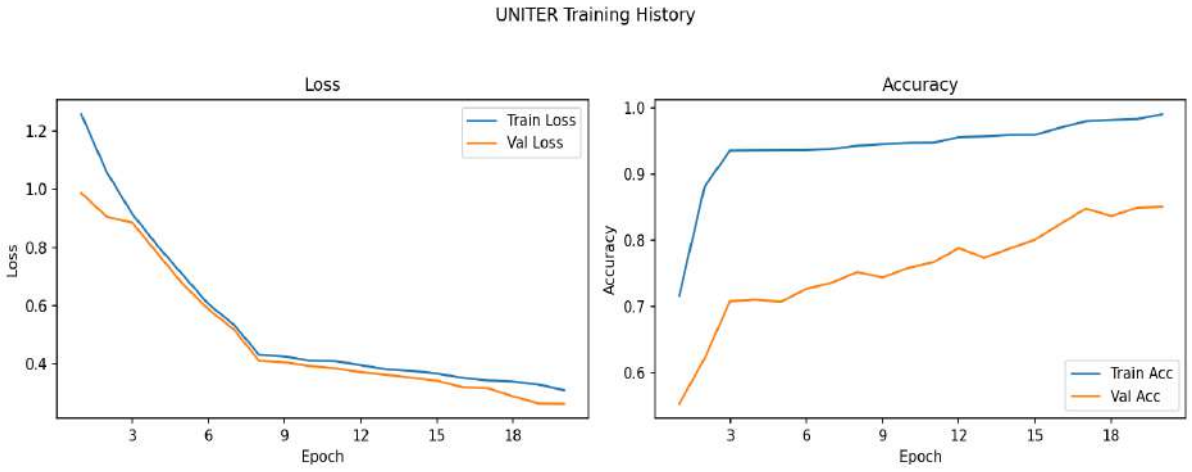
- Recall là chỉ số quan trọng trong bài toán này, thể hiện tỷ lệ tin giả được phát hiện đúng. Một mô hình với Recall cao đảm bảo rằng phần lớn các tin giả không bị bỏ sót.

- F1-Score là trung bình điều hòa giữa Precision và Recall, cung cấp một thước đo tổng quát, đặc biệt hữu ích trong các bài toán yêu cầu cân bằng giữa hai yếu tố này.

4.2. Kết quả các mô hình

4.2.1 Mô hình UNITER

Kết quả huấn luyện của mô hình đa thể thức UNITER được thể hiện trong Hình 1



Hình 1. Biểu đồ Accuracy & Loss sau 20 epoch

Quan sát lịch sử huấn luyện trong 20 epoch cho thấy mô hình UNITER có quá trình hội tụ rõ ràng và ổn định. Train loss giảm liên tục từ 1.257 → 0.310; Validation loss giảm từ 0.987 → 0.263 và không xuất hiện hiện tượng dao động mạnh hoặc tăng ngược trở lại. Đáng chú ý, validation loss luôn giảm đều qua các epoch cuối (từ epoch 15 đến 20), cho thấy mô hình chưa rơi vào trạng thái overfitting. Việc validation loss thấp hơn train loss ở giai đoạn cuối có thể do cơ chế regularization trong quá trình huấn luyện. Có thể kết luận mô hình hội tụ tốt và ổn định.

Về thông số Accuracy, train accuracy tăng nhanh từ 71.65% lên 98.98%, validation accuracy tăng từ 55.33% lên 85.06%. Có thể kết luận giai đoạn tăng trưởng có thể chia thành ba pha:

- Epoch 1–5: tăng nhanh (học đặc trưng cơ bản);
- Epoch 6–12: tăng ổn định (tinh chỉnh biểu diễn đa phương thức)

- Epoch 13–20: cải thiện chậm nhưng đều (fine-tuning sâu)

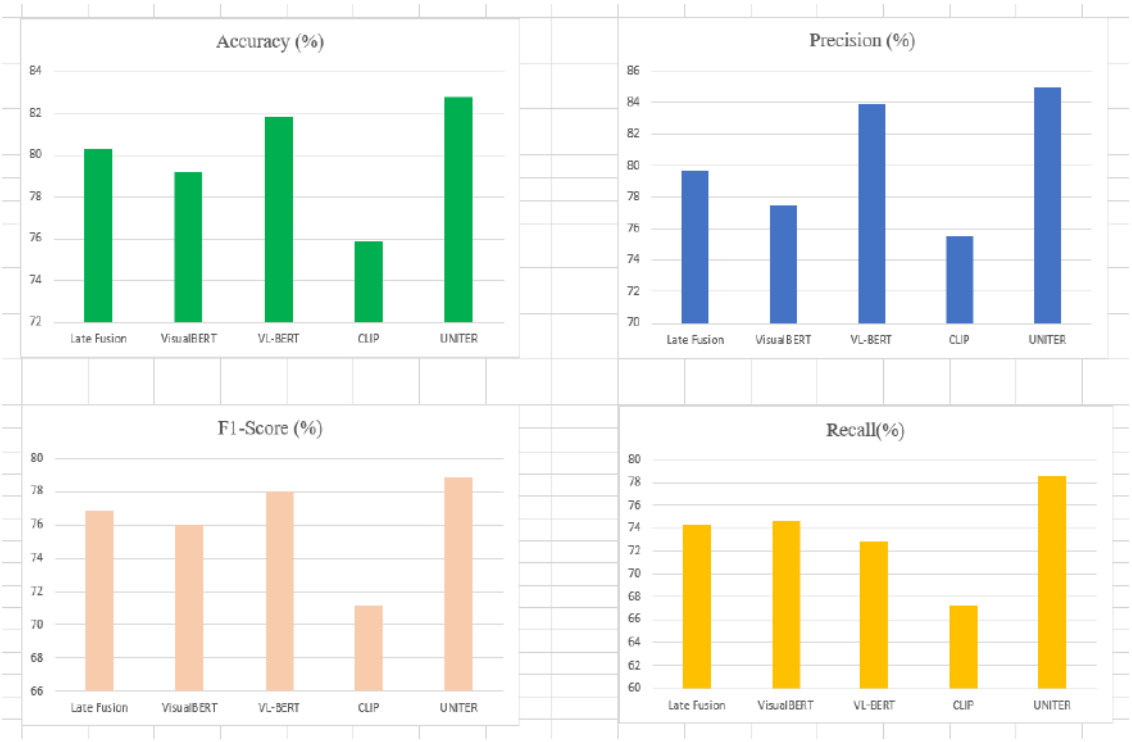
Validation accuracy đạt mức cao nhất tại epoch 20 (85.05%), đồng thời validation loss đạt mức thấp nhất, cho thấy mô hình đạt trạng thái tối ưu tại epoch cuối.

4.2.1. So sánh kết quả với các mô hình khác

Kết quả trên tập dữ liệu kiểm tra sau khi huấn luyện các mô hình được trình bày trong Bảng 1 và Hình 2.

Bảng 1. Kết quả phân loại của 5 mô hình

STT	Mô hình đa thể thức	Accuracy (%)	Precision (%)	F1-Score (%)	Recall(%)
1	Late Fusion	80.29	79.71	76.92	74.31
2	VisualBERT	79.19	77.47	76.03	74.64
3	VL-BERT	81.83	83.91	78.00	72.86
4	CLIP	75.87	75.50	71.12	67.22
5	UNITER	**85.05**	**84.98**	**78.82**	**78.65**



Hình 2. Biểu đồ chỉ số của các mô hình

Kết quả thực nghiệm cho thấy mô hình UNITER đạt hiệu năng cao nhất trên hầu hết các chỉ số đánh giá. Cụ thể, UNITER đạt Accuracy 85.05%, Precision 84.98%, F1-score 78.82% và Recall 78.65%, vượt trội so với các mô hình còn lại. Điều này cho thấy khả năng học biểu diễn liên kết đa phương thức sâu và hiệu quả của UNITER giúp cải thiện đáng kể khả năng phân biệt tin giả và tin thật.

Mô hình VL-BERT đạt Accuracy 81.83% và Precision 83.91%, cho thấy khả năng dự đoán chính xác các mẫu được phân loại là tin giả tương đối tốt. Tuy nhiên, Recall của VL-BERT

(72.86%) thấp hơn so với UNITER, cho thấy mô hình vẫn bỏ sót một số lượng đáng kể các mẫu tin giả.

Late Fusion và VisualBERT có hiệu năng tương đương nhau, với Accuracy lần lượt là 80.29% và 79.19%. Hai mô hình này đạt F1-score xấp xỉ 76–77%, cho thấy khả năng cân bằng giữa Precision và Recall ở mức khá, tuy nhiên chưa đạt được mức tối ưu như UNITER.

Trong khi đó, CLIP cho kết quả thấp nhất với Accuracy 75.87% và F1-score 71.12%. Điều này có thể do CLIP được thiết kế chủ yếu cho bài toán đối sánh ảnh - văn bản thay vì tối ưu trực tiếp cho bài toán phân loại tin giả.

5. Kết luận

Nghiên cứu đã đánh giá hiệu quả các mô hình đa thể thức (Late Fusion, VisualBERT, VL-BERT, CLIP, UNITER) trong bài toán phát hiện tin giả. Kết quả cho thấy UNITER đạt hiệu năng tổng thể tốt nhất, nhờ khả năng học biểu diễn liên kết ảnh – văn bản hiệu quả và ổn định; các mô hình có cơ chế tương tác chéo sâu (UNITER, VL-BERT) vượt trội so với phương pháp kết hợp muộn. Đồng thời, việc cân bằng giữa Precision và Recall là yếu tố quan trọng hơn chỉ tối ưu Accuracy, đặc biệt trong bối cảnh ứng dụng thực tế.

Trong tương lai, cần mở rộng và đa dạng hóa dữ liệu để tăng khả năng tổng quát hóa, kết hợp thêm thu thập dữ liệu từ đa dạng nguồn tin, nền tảng mạng xã hội và các ngữ cảnh truyền thông khác nhau sẽ giúp mô hình thích nghi tốt hơn với sự thay đổi liên tục của môi trường thông tin số. Ngoài ra, việc khai thác dữ liệu thời gian thực sẽ góp phần nâng cao khả năng phát hiện sớm và ngăn chặn kịp thời các chiến dịch lan truyền tin giả. Hệ thống có thể được phát triển theo hướng cập nhật liên tục, học thích nghi với các mẫu tin mới và các hình thức thao túng tinh vi. Song song với đó, nghiên cứu cũng có thể mở rộng sang việc kết hợp các kiến trúc đa thể thức tiên tiến hơn, tận dụng cơ chế tương tác chéo giữa các loại dữ liệu nhằm cải thiện chất lượng biểu diễn và hiệu suất phân loại. Đặc biệt, việc mở rộng bài toán với dữ liệu có chứa âm thanh là một hướng nghiên cứu tiềm năng. Trong bối cảnh các nội dung đa phương tiện như video, podcast, livestream ngày càng phổ biến, việc tích hợp thêm đặc trưng âm thanh có thể giúp hệ thống phát hiện hiệu quả hơn các tin giả tinh vi như deepfake.

Nhìn chung, nghiên cứu này đã cung cấp một cái nhìn toàn diện về hiệu suất của các mô hình đa thể thức trong bài toán phát hiện tin giả, đồng thời làm rõ ưu điểm của các kiến trúc tiền huấn luyện đa phương thức hiện đại. Kết quả đạt được không chỉ khẳng định tính khả thi của hướng tiếp cận đa thể thức, mà còn mở ra nhiều định hướng phát triển tiếp theo nhằm nâng cao độ chính xác, khả năng tổng quát hóa và tính ứng dụng thực tiễn của hệ thống trong môi trường thông tin ngày càng phức tạp.

Tài liệu tham khảo

1. Songtao Li Hao Tang∗. (2025). Multimodal Alignment and Fusion: A Survey.
2. Jacob Devlin Ming-Wei Chang Kenton Lee Kristina Toutanova Google AI Language {jacobdevlin,mingweichang,kentonl,kristout}@google.com. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.
3. Kaiming He Xiangyu Zhang Shaoqing Ren Jian Sun Microsoft Research {kahe, v-xiangz, v-shren, jiansun}@microsoft.com. (2015). Deep Residual Learning for Image Recognition.
4. Liunian Harold Li†, Mark Yatskar∗, Da Yin∘, Cho-Jui Hsieh† & Kai-Wei Chang†, †University of California, Los Angeles, ∗Allen Institute for Artificial Intelligence, ∘Peking University. (2019). VisualBERT: A Simple and Performant Baseline for Vision and Language.
5. Dibyanayan Bandyopadhyay Indian Institute of Technology Patna Patna, Bihar, India dibyanayan_2321cs14@iitp.ac.in Mohammed Hasanuzzaman EEECS, Queen's University Belfast Belfast, BT7 1NN, UK m.hasanuzzaman@qub.ac.uk. (2024). Seeing Through VisualBERT: A Causal Adventure on Memetic Landscapes.
6. Weijie Su1,2∗ , Xizhou Zhu1,2∗ , Yue Cao2 , Bin Li1 , Lewei Lu2 , Furu Wei2 , Jifeng Dai2. (2020). VL-BERT: PRE-TRAINING OF GENERIC VISUALLINGUISTIC REPRESENTATIONS.
7. Christoph Schuhmann, Richard Vencu, Romain Beaumont, Robert Kaczmarczyk, Clayton Mullis, Aarush Katta, Theo Coombes, Jenia Jitsev, and Aran Komatsuzaki. Laion-400m: Open dataset of clip-filtered 400 million image-text pairs. *arXiv preprint arXiv:2111.02114, 2021.*
8. Qiang Zhang , Jiawei Liu∗ , Fanrui Zhang , Jingyi Xie and Zheng-Jun Zha. (2024). Natural Language-centered Inference Network for Multi-modal.
9. Yen-Chun Chen? , Linjie Li? , Licheng Yu? , Ahmed El Kholy Faisal Ahmed, Zhe Gan, Yu Cheng, and Jingjing Liu. (2020). UNITER: UNiversal Image-TExt Representation Learning.
10. Li, G., Duan, N., Fang, Y., Jiang, D., Zhou, M.: Unicoder-vl: A universal encoder for vision and language by cross-modal pre-training. In: AAAI (2020).
11. Fukui, A., Park, D.H., Yang, D., Rohrbach, A., Darrell, T., Rohrbach, M.: Multimodal compact bilinear pooling for visual question answering and visual grounding. In: EMNLP (2017).
12. Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R., Le, Q.V.: Xlnet: Generalized autoregressive pretraining for language understanding. In: NeurIPS (2019).

Evaluation of Several Multimodal Models for Fake News Detection

Abstract: *In the context of the rapid growth of social media, fake news has become a serious societal problem, aiming to mislead the public, manipulate public opinion, create social division, and even interfere with politics and the economy. Its consequences are extremely severe, ranging from eroding trust in institutions and distorting public perception to causing tangible damage to health, finances, and national security, as seen during the COVID-19 pandemic and elections around the world. Detecting fake news not only helps protect social media users but also ensures the safety and transparency of online platforms. This study investigates the multimodal model Universal Image-Text Representation (UNITER) for fake news detection and compares it with several popular multimodal models, including Late Fusion, VisualBERT, VL-BERT, and Contrastive Language-Image Pre-training (CLIP). The results aim to evaluate the applicability of the UNITER model in comparison with other common multimodal approaches for fake news detection, thereby assisting developers and platform administrators in selecting appropriate technologies to enhance the reliability of information on social media.*

Keywords: Universal Image-Text Representation; Late Fusion; VisualBERT; VL-BERT; Contrastive Language-Image Pre-training; Fake news detection.

Systematic Evaluation of Data Augmentation Strategies for Contrastive Learning in NIDS

Nguyễn Hữu Long[1*]

[1]*Institute of Information and Communication Technology, Le Quy Don Technical University*

Email: nguyenhuulong312@gmail.com; Tel: (+84) 836 131 986

Abstract: *Network Intrusion Detection Systems (NIDS) face significant challenges due to intra-class diversity and inter-class similarity in network traffic. Contrastive Learning (CL) has shown promise in learning robust representations through augmented views of data. However, a key limitation of existing contrastive NIDS frameworks lies in their reliance on a single augmentation strategy, such as random masking, without systematic evaluation of alternative designs. This study investigates augmentation strategies for contrastive NIDS along three dimensions: (i) exploring extended augmentation techniques; (ii) applying feature-importance-aware selective augmentation; and (iii) evaluating different numbers of augmented views in multi-view contrastive learning. Experiments on NSL-KDD and UNSW-NB15 demonstrate that augmentation effectiveness strongly depends on dataset characteristics. Mixup provides the most consistent improvements across datasets, while selective augmentation and multi-view settings require careful adjustment. These findings emphasize the critical role of augmentation design in representation quality and intrusion detection performance, and provide empirical guidance for robust contrastive learning on tabular network data.*

Keywords: Contrastive Learning; Network Intrusion Detection; Data Augmentation; Representation Learning; Feature Importance

1. Introduction

Network Intrusion Detection Systems (NIDS) play a critical role in cybersecurity by monitoring traffic and detecting malicious activities. Traditional machine learning approaches often rely on large labeled datasets and struggle with zero-day attacks. Anomaly-based methods can detect unseen threats but frequently suffer from high false alarm rates, limiting practical deployment [1].

Recently, self-supervised learning, particularly contrastive learning (CL), has emerged as a promising solution for leveraging unlabeled data [2]. By constructing positive and negative pairs through data augmentation and optimizing embedding distances, CL improves representation robustness and discriminability without extensive annotation.

Contrastive learning has shown strong performance in multiple domains, notably in computer vision through SimCLR [3]. In intrusion detection, prior work integrating supervised contrastive learning into NIDS has demonstrated improved representation quality and classification performance. This is especially relevant for modern network traffic, where high intra-class diversity and inter-class similarity blur decision boundaries [4].

In this work, we systematically investigate the role of data augmentation in contrastive learning for tabular NIDS. We evaluate multiple augmentation families (random masking, mixup, Gaussian noise, dropout), feature-importance-aware selective augmentation, and varying

[*] Email: nguyenhuulong312@gmail.com

numbers of augmented views. Our goal is to analyze how these configurations affect representation quality, stability, and detection performance across different datasets. The remainder of this paper is organized as follows. Section 2 reviews the theoretical background. Section 3 describes the experimental framework. Section 4 presents the results and analysis. Section 5 concludes the paper and outlines future work.

2. Theoretical Background

2.1. Network Intrusion Detection Systems (NIDS)

A Network Intrusion Detection System (NIDS) monitors network traffic to detect malicious or abnormal activities. Traditional approaches fall into two categories: signature-based and anomaly-based detection [1].

Signature-based systems match traffic against known attack patterns and achieve low false positives but cannot detect zero-day attacks. Anomaly-based systems model normal traffic behavior and identify deviations, enabling novel attack detection at the cost of higher false alarm rates.

With advances in machine learning, models such as DNN, CNN, and LSTM have been widely applied to NIDS [1], automatically learning feature representations and improving detection of complex patterns. More recently, self-supervised approaches, particularly contrastive learning, have been introduced to enhance representation robustness while reducing reliance on labeled data.

2.2. Network Traffic Datasets

Network traffic is captured at the packet level but is typically aggregated into flows for intrusion detection. A flow is defined by source/destination IPs, ports, protocol, and a time window, providing a higher-level abstraction suitable for machine learning.

- **NSL-KDD** [5] refines the earlier KDD'99 dataset by reducing redundancy and class imbalance. It contains 41 features and classifies traffic into normal and four attack categories: DoS, Probe, R2L, and U2R.

- **UNSW-NB15** [6] reflects modern network conditions, containing 49 features and nine attack categories. Compared to NSL-KDD, it includes more complex statistical and behavioral attributes, making it suitable for evaluating contemporary NIDS performance.

Using both datasets enables evaluation across classical and modern settings with differing feature structures.

2.3. Contrastive Learning

Contrastive learning (CL) is a self-supervised representation learning paradigm designed to extract meaningful embeddings from unlabeled data. The core intuition is straightforward: different augmented views of the same sample should be close in embedding space, while representations of different samples should be far apart.

2.3.1. Core Components

Data Augmentation

Given an original sample x, two transformed views are generated: $x_i = t_{i(x)}$ and $x_j = t_{j(x)}$

where t_i and t_j are stochastic augmentation functions. In tabular NIDS data, augmentations may include masking, controlled noise injection, or interpolation (e.g., mixup).

Encoder and Projection

Each view is passed through an encoder network f(·) to obtain latent representations:

$$h_i = f(x_i) \text{ and } h_j = f(x_j)$$

Optionally, a projection head g(·) maps representations to contrastive space: $z_i = g(h_i)$ and $z_j = g(h_j)$

Similarity Measure

Similarity between embeddings is commonly computed using cosine similarity:

$$\text{sim}(z_i, z_j) = (z_i \cdot z_j) / (\|z_i\| * \|z_j\|) \quad (1)$$

where "·" denotes dot product, ||·|| denotes L2 norm.

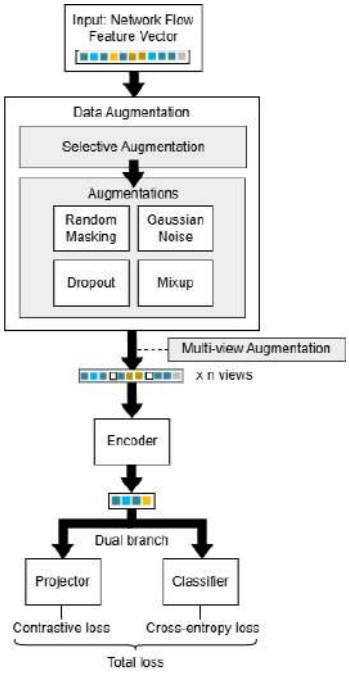


Figure 1. Proposed dual-branch architecture used in the experimental framework

InfoNCE Loss

The most widely used contrastive objective is the InfoNCE loss [7]. For a positive pair (i, j), the loss for anchor i is defined as:

$$L_i = -\log \frac{\exp\left(\frac{sim(z_i, z_j)}{\tau}\right)}{\sum_{k=1, k\neq i}^{2N} \exp \sum\left(\frac{sim(z_i, z_k)}{\tau}\right)} \quad (2)$$

where: N is the batch size, $\tau > 0$ is a temperature parameter, $1_{\{[k \neq i]\}}$ excludes the anchor itself and all other samples in the batch serve as negative examples.

The overall loss over the batch is: $L_{InfoNCE} = \left(\frac{1}{2N}\right) \sum_{i=1}^{2N} L_i$ (3)

This objective increases similarity between positive pairs while decreasing similarity to negative samples.

2.3.2. Self-Supervised vs Supervised Contrastive Learning

In self-supervised contrastive learning, such as SimCLR [3], positive pairs are generated solely from augmented views of the same sample, and all other samples in the batch are treated as negatives [2].

Supervised Contrastive Learning (SupCon) extends this idea by treating all samples sharing the same class label as positive pairs. The supervised contrastive loss is formulated as:

$$L_{\text{sup}} = \sum_{i=1}^{N} -\frac{1}{|P(i)|} \sum_{p \in P(i)} \log \left(\frac{\exp(sim(z_i, z_p)/\tau)}{\sum_{a \neq i} \exp(sim(z_i, z_a)/\tau)} \right) (4)$$

where P(i) denotes the set of indices sharing the same label as sample i. *SupCon* improves class separation when labels are available.

2.3.3. Contrastive Learning for NIDS

The Contrastive Learning Enhanced Intrusion Detection framework [4] applies contrastive learning to tabular network flows by designing suitable augmentations and combining contrastive loss with cross-entropy classification loss:

$$L_{total} = L_{CE} + \lambda * L_{CL} \quad (5)$$

where λ controls the trade-off between classification and contrastive objectives.

Contrastive learning is particularly suitable for NIDS because:

- It reduces reliance on large labeled datasets by leveraging unlabeled traffic.
- It encourages invariance to small perturbations, improving robustness.
- It enhances separability for rare attack classes when combined with supervised fine-tuning.

2.4. Data Augmentation in Contrastive NIDS

Data augmentation is central to contrastive learning, as it defines the invariances the model is encouraged to learn. In network intrusion detection, augmentations must preserve semantic consistency while introducing sufficient variability.

Common augmentation strategies include:

- **Random masking/dropout:** randomly masking feature dimensions to simulate incomplete observations.
- **Gaussian noise:** adding small perturbations while preserving distributional structure.
- **Mixup:** generating interpolated samples

$$x_{mix} = \alpha\, x_i + (1 - \alpha)\, x_j, \text{ where } \alpha \in [0,1].$$

- **Scaling or normalization-based perturbations**

Domain-specific augmentation design has been shown to critically affect representation quality, as demonstrated in GraphCL for graph data. Similar sensitivity is expected in tabular NIDS data, where features differ in semantics and importance.

Recent studies further explore augmented negative pairs, where negatives are intentionally made more challenging to sharpen decision boundaries [8]. These findings indicate that both positive and negative pair construction influence representation strength.

Based on these observations, this work systematically evaluates multiple augmentation strategies in contrastive NIDS, including uniform masking, feature-importance-aware masking, noise injection, mixup, and multi-view configurations. The goal is to clarify how augmentation design shapes embedding geometry and downstream intrusion detection performance.

3. Proposed Experimental Framework

3.1. Data Preprocessing

NSL-KDD and UNSW-NB15 are preprocessed following the standardized CLEID pipeline [4] to ensure fair comparison. Identifier fields are removed; categorical features are one-hot encoded, numerical features are MinMax normalized, and binary features are mapped to {0,1}. Missing values are imputed using standard statistical strategies.

After preprocessing, NSL-KDD contains 121 features with 5 labels, and UNSW-NB15 contains 196 features with 10 labels. This unified pipeline ensures consistent evaluation across datasets.

3.2. Model Architecture

We adopt a dual-branch contrastive architecture inspired by CLEID [4], consisting of encoder f_{enc}, projection head f_{proj}, and classifier f_{cls}.

Given input x: $r = f_{enc}(x)$

For augmented views: $z = f_{proj}(r)$

Embeddings are normalized: $z = z / ||z||$

The classifier predicts: $\hat{y} = f_{cls}(f_{enc}(x))$

The overall objective combines supervised contrastive and classification losses:

$$L_{\text{total}} = L_{\text{sup}} + \lambda * L_{\text{CE}}$$

This formulation jointly enforces augmentation invariance and class discrimination.

3.3. Augmentation Strategies

To account for heterogeneous feature semantics in tabular NIDS data, we evaluate multiple augmentation families beyond uniform masking.

Random Masking (baseline) $\tilde{x} = M \odot x$

Gaussian Noise $\tilde{x} = x + \varepsilon, \varepsilon \sim N(0, \sigma^2 I)$

Dropout Each feature is independently zeroed with probability p.

Mixup $\tilde{x} = \lambda x_i + (1 - \lambda) x_j$

Selective Augmentation

Feature importance scores I_j are estimated using Random Forest. Augmentation is applied only to the least-important subset: $\tilde{x} = t(x_S) \oplus x_R$

Multi-View Extension

Instead of two views, we evaluate $n \in \{2, 4, 6\}$ to examine the trade-off between stronger invariance and potential information dilution.

3.4. Experimental Design

Experiments are conducted in four stages:

- **Baseline reproduction** using random masking.
- **Augmentation replacement** comparing noise, dropout, and mixup.
- **Selective augmentation** based on feature importance.
- **Multi-view scaling** applied to baseline and best-performing methods.

Evaluation Metrics

We report Accuracy, Precision, Recall, F1-score, and AUROC, averaged across multiple random seeds. Embedding geometry is further analyzed using alignment and uniformity to assess representation quality.

Implementation Details

The encoder is implemented as a multi-layer perceptron with two hidden layers. The projection head consists of a two-layer MLP. The model is trained using Adam optimizer with

a learning rate of 1e-3. Batch size is fixed at 256, and temperature τ is selected from {0.01, 0.07, 0.6}. For augmentation parameters, σ is chosen from {0.0001, 0.00025, 0.0005} for Gaussian noise, $p \in \{0.05, 0.1, 0.2\}$ for dropout, and $\alpha \in \{0.9999, 0.99975, 0.9995\}$ for mixup. All experiments are conducted for 10 epochs under consistent settings to ensure fair comparison.

4. Experimental Results and Analysis

Based on the established baseline, we evaluate three extensions:

(i) replacing random masking with independent augmentations;

(ii) applying feature-importance-based selective augmentation;

(iii) increasing augmented view numbers.

The results reveal clear dataset-dependent behaviors and highlight how augmentation type and view multiplicity influence representation learning in contrastive NIDS.

4.1. Replacing Random Masking with Independent Augmentations

Table I reports Gaussian noise, dropout, and mixup compared to random masking across 9 configurations.

Table 1. Performance Change (%) Compared to Random Masking Baseline (9 Configurations)

Dataset	Augmentation	Acc Δ	F1 Δ	#Improved / 9
NSL-KDD	Noise	+1,19	+1,14	9 / 9
NSL-KDD	Dropout	+1,11	+1,05	9 / 9
NSL-KDD	Mixup	+1,91	+1,82	9 / 9
UNSW-NB15	Noise	−0,95	−0,97	0 / 9
UNSW-NB15	Dropout	−1,02	−1,01	0 / 9
UNSW-NB15	Mixup	+0,19	+0,18	9 / 9

- On NSL-KDD, all augmentations outperform random masking, with mixup achieving the largest gains.
- On UNSW-NB15, noise and dropout consistently degrade performance, while mixup remains stable with modest improvements.

These results indicate strong dataset dependency: NSL-KDD benefits from smoothing-based transformations, whereas UNSW-NB15 is sensitive to direct perturbations. To ensure result reliability, all experiments were repeated across multiple random seeds, and we observed low variance across runs. The standard deviation was consistently below 0.3% for Accuracy and F1-score, indicating stable performance.

4.2. Selective Augmentation Based on Feature Importance

Selective augmentation is applied to the bottom 30% least-important features (via Random Forest ranking) and compared with full augmentation.

Table 2. Selective Augmentation vs Full Augmentation

Dataset	Method	Acc Δ	F1 Δ	#Improved / 9
NSL-KDD	Sel-Masking	+2,46	+2,31	9 / 9
NSL-KDD	Sel-Noise	+1,13	+1,08	9 / 9
NSL-KDD	Sel-Dropout	+1,19	+1,15	9 / 9
NSL-KDD	Sel-Mixup	+2,11	+2,02	9 / 9
UNSW-NB15	Sel-Mixup	+0,22	+0,20	9 / 9
UNSW-NB15	Sel-Noise	−0,88	−0,91	0 / 9

• On NSL-KDD, selective augmentation consistently improves all methods, suggesting that preserving high-importance features stabilizes representation learning.

• On UNSW-NB15, only selective mixup provides consistent gains, while perturbation-based methods still degrade performance.

Thus, feature-aware augmentation is beneficial only when aligned with a compatible transformation family.

4.3. Multi-View Augmentation

We increase view numbers from 2 (baseline) to 4 and 6. Representative 4-view results are shown in Table III.

Table 3. Effect of Increasing View Numbers (4 Views)

Dataset	Method	Acc Δ	F1 Δ	#Improved
NSL-KDD	Masking	+0,57	+0,53	9 / 9
NSL-KDD	Mixup	−0,61	−0,58	2 / 9
UNSW-NB15	Masking	+2,32	+2,25	9 / 9
UNSW-NB15	Noise	+1,79	+1,71	9 / 9

• On UNSW-NB15, increasing views generally improves robustness, indicating that stronger invariance constraints help regularize complex feature distributions.

• On NSL-KDD, moderate view expansion may cause diminishing returns, particularly for non-masking augmentations.

Six-view experiments further confirm that larger view multiplicity stabilizes performance on UNSW-NB15, while NSL-KDD benefits primarily from masking-based strategies.

Overall Discussion

Three consistent observations emerge.

1. Augmentation effectiveness is strongly dataset-dependent. NSL-KDD generally benefits from most strategies, whereas UNSW-NB15 favors structure-preserving transformations.

2. Mixup is the most robust and consistently stable augmentation across both datasets, maintaining performance even under multi-view settings.

3. Selective and multi-view augmentation improve invariance when aligned with dataset characteristics, but their benefits depend on the underlying transformation. Increasing view

numbers strengthens representation stability for complex data, though excessive expansion may destabilize certain augmentations.

Overall, augmentation design in contrastive NIDS should be regarded as a structural component shaping embedding geometry and classification stability, rather than a simple hyperparameter choice.

5. Conclusion

This paper presents a systematic study of data augmentation strategies for contrastive learning in network intrusion detection, extending the CLEID framework along three orthogonal dimensions: alternative augmentation families (noise, dropout, mixup), feature-importance-aware selective augmentation, and multi-view learning with increased view multiplicity. Experiments on NSL-KDD and UNSW-NB15 ensure dataset-independent conclusions.

Results demonstrate that augmentation plays a central role in shaping representation quality. Mixup emerges as the most stable and consistently effective strategy across heterogeneous datasets. Selective augmentation improves stability when properly aligned with data characteristics but cannot compensate for inherently disruptive transformations. Multi-view learning enhances invariance for complex datasets, although its benefit depends on both augmentation type and traffic variability.

These findings confirm that augmentation design is a structural factor influencing embedding geometry and detection performance. Effective contrastive NIDS requires coordinated alignment between transformation type, feature relevance, and view multiplicity.

Future Work

Future research may explore:

- Designing augmentation strategies tailored to tabular network data, including distribution-aware or learned policies.
- Developing adaptive selective mechanisms that adjust augmentation intensity dynamically rather than relying on fixed global importance.
- Implementing adaptive multi-view schemes that vary view multiplicity according to traffic complexity.
- Building integrated augmentation pipelines that jointly optimize independent, selective, and multi-view strategies.

Acknowledgment

The authors would like to express their sincere gratitude to the Institute of Information and Communication Technology, Le Quy Don Technical University, for providing academic support and research facilities throughout this work. The first author also thank Assoc. Prof. Viet Hung Nguyen for his valuable guidance and insightful feedback during the development of this study.

References

1. Hozouri, A., Mirzaei, A., & Effatparvar, M. (2025). A comprehensive survey on intrusion detection systems with advances in machine learning, deep learning and emerging cybersecurity challenges. Discover Artificial Intelligence, 5, Article 314.

2. Khosla, P., Teterwak, P., Wang, C., Sarna, A., Tian, Y., Isola, P., … Krishnan, D. (2020). Supervised contrastive learning. In Advances in Neural Information Processing Systems (Vol. 33).
3. Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). A simple framework for contrastive learning of visual representations. In Proceedings of the International Conference on Machine Learning (ICML).
4. Yue, Y., Chen, X., Han, Z., Zeng, X., & Zhu, Y. (2022). Contrastive learning enhanced intrusion detection. IEEE Transactions on Network and Service Management, 19(4), 4232–4247. https://doi.org/10.1109/TNSM.2022.3218843.
5. Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009). A detailed analysis of the KDD CUP 99 data set. In Proceedings of the IEEE Symposium on Computational Intelligence for Security and Defense Applications (pp. 1–6). Ottawa, Canada.
6. Moustafa, N., & Slay, J. (2015). UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). In Proceedings of the 2015 Military Communications and Information Systems Conference (MilCIS) (pp. 1–6). Canberra, Australia.
7. van den Oord, A., Li, Y., & Vinyals, O. (2019). Representation learning with contrastive predictive coding. arXiv preprint arXiv:1807.03748.
8. Wilkie, J., Hindy, H., Tachtatzis, C., & Atkinson, R. (2025). Contrastive self-supervised network intrusion detection using augmented negative pairs. In Proceedings of the IEEE Conference on Cyber Security and Resilience (CSR) (pp. 206–213). Chania, Crete, Greece. https://doi.org/10.1109/CSR64739.2025.11129979

Systematic Evaluation of Data Augmentation Strategies for Contrastive Learning in NIDS

Tóm tắt

Hệ thống phát hiện xâm nhập mạng (NIDS) đối mặt với thách thức do sự đa dạng trong cùng lớp và sự tương đồng giữa các lớp lưu lượng. Học tương phản (Contrastive Learning – CL) cho thấy tiềm năng trong việc học biểu diễn bền vững thông qua các phiên bản tăng cường của dữ liệu. Tuy nhiên, hạn chế hiện nay nằm ở việc chỉ sử dụng chiến lược duy nhất như random masking mà thiếu đánh giá có hệ thống về các phương án khác. Nghiên cứu này khảo sát các chiến lược tăng cường cho NIDS dựa trên CL theo ba hướng: (i) thử nghiệm các phép tăng cường mở rộng; (ii) áp dụng selective augmentation theo độ quan trọng đặc trưng; và (iii) thử nghiệm số lượng view tăng cường khác nhau. Thử nghiệm trên NSL-KDD và UNSW-NB15 cho thấy hiệu quả của augmentation phụ thuộc mạnh vào đặc tính dữ liệu. Mixup cho cải thiện ổn định nhất, trong khi selective augmentation và multi-view cần được điều chỉnh phù hợp. Kết quả nhấn mạnh vai trò quan trọng của thiết kế augmentation đối với chất lượng biểu diễn và hiệu năng phát hiện xâm nhập.

Từ khóa: *Học tương phản; Phát hiện xâm nhập mạng; Tăng cường dữ liệu; Học biểu diễn; Độ quan trọng đặc trưng*

SSDAE: A Self-Supervised Structured Denoising Autoencoder for One-Class Intrusion Detection

Van Cuong Nguyen[1*]
[1]*Le Qui Don Technical University*

Abstract

Intrusion detection under one-class learning remains challenging because models trained exclusively on normal traffic often become highly sensitive to minor distribution shifts and may fail to separate unseen attacks from benign variations. Although denoising autoencoders improve representation robustness by reconstructing clean inputs from corrupted observations, conventional corruption strategies usually rely on random perturbations that do not adequately reflect realistic traffic variations. A structured denoising framework is introduced in which CTGAN-generated synthetic samples are converted into feature perturbations to guide self-supervised reconstruction learning. Instead of applying predefined stochastic noise, the proposed method injects distribution-aware perturbations into continuous attributes of normal samples and trains the autoencoder to recover the corresponding clean inputs. This reconstruction objective promotes latent representations that remain stable under realistic feature variation while preserving benign traffic characteristics. Experiments conducted on benchmark intrusion detection datasets show that SSDAE improves reconstruction-based anomaly discrimination and consistently outperforms representative autoencoder-based baselines.

Từ khóa: *Intrusion detection systems, Anomaly detection, Generative models, Network security*

1. Introduction

Intrusion Detection Systems (IDSs) are widely used to detect malicious activities in network traffic, but conventional signature-based mechanisms remain limited against unseen and stealthy attacks [1]. This limitation has motivated anomaly-based intrusion detection using machine learning and deep learning models.

Autoencoder-based approaches are widely adopted in one-class intrusion detection because they model benign traffic structure without relying on large amounts of attack annotations [2,3]. However, models trained exclusively on benign traffic may still assign low reconstruction error to attack samples that remain close to normal patterns, especially when abnormal traffic closely resembles legitimate behavior [4,5].

To address this limitation, this paper proposes a Self-Supervised Structured Denoising Autoencoder (SSDAE), which improves representation robustness through structured perturbation-based denoising rather than direct clean-input reconstruction.

2. Background

This section briefly presents CTGAN and Autoencoder, the two fundamental models employed in our research.

2.1. Conditional Tabular Generative Adversarial Network

Conditional Tabular Generative Adversarial Network (CTGAN) [6] is a generative framework developed for tabular data synthesis under heterogeneous feature distributions and severe imbalance across categorical attributes. CTGAN models tabular data by separately

[*] Email: cuongpd@lqdtu.edu.vn

handling continuous and categorical features. Continuous variables are normalized through variational Gaussian mixtures, while categorical variables are learned using conditional generation. For continuous variables: $c_{i,j} \to \alpha_{i,j} \oplus \beta_{i,j}$ where denotes the one-hot encoding of the selected Gaussian component. The normalized scalar is computed as: $\alpha_{i,j} = \frac{c_{i,j} - \eta_k}{4\phi_k}$ (1)

where and represent the mean and standard deviation of the selected mixture component. This transformation allows numerical features with multi-modal distributions to be represented in a form that is more suitable for adversarial learning.

For categorical variables, CTGAN uses conditional generation to improve the representation of minority categories. During each training step, a conditional vector is sampled and concatenated with the latent input of the generator: $\hat{x} = G(z, c)$ where is a latent noise vector and denotes the sampled conditional vector.

The generator is optimized to approximate the conditional distribution of real tabular samples: $P_G(x \mid c) \approx P(x \mid c)$. This conditional learning strategy improves category coverage and helps preserve statistical relationships among heterogeneous features, making CTGAN suitable for generating structured variations of normal traffic data

2.2. Denoising Autoencoder

Autoencoder (AE) is an unsupervised learning model that compresses input data into latent representations and reconstructs them through a paired encoder--decoder structure. In contrast to the standard autoencoder, a Denoising Autoencoder (DAE) introduces an explicit corruption process to enhance representation robustness [7]. Instead of directly reconstructing the original input x, the encoder receives a stochastically corrupted version x_{noisy}, while the decoder is trained to recover the clean sample x. Formally, let C(·) denote a corruption operator. The corrupted input is defined as: $x_{noisy} = \mathrm{C}(x)$ and the latent representation and reconstruction are given by $z = f_\theta(x_{noisy}), \hat{x} = g_\phi(z)$ (2)

The training objective minimizes the reconstruction error between the clean input and its reconstruction: (3)

3. Related work

Reconstruction-based autoencoders have become a common choice in intrusion detection because they support latent representation learning under limited attack supervision [8,9]. In one-class settings, models are typically trained only on benign traffic so that anomalies can later be identified through elevated reconstruction errors. This strategy is particularly suitable for intrusion detection, where malicious samples are often limited and attack patterns continuously evolve.

Prior studies indicate that reconstruction-based autoencoders can model normal traffic distributions effectively in high-dimensional spaces. For example, Meidan et al. [10] used a conventional autoencoder trained on benign traffic and detected anomalies using a reconstruction threshold, while later works improved robustness through deeper architectures

and ensemble designs [11,12]. More recently, Torabi et al. [13] demonstrated that the full reconstruction error vector can provide richer anomaly information than a single scalar score. However, these approaches still rely heavily on threshold selection and may remain ineffective against stealthy attacks that produce only small reconstruction deviations.

To improve representation stability, denoising autoencoders introduce corrupted inputs and learn to reconstruct clean observations, often producing smoother latent manifolds [3]. However, most existing methods use generic perturbations such as Gaussian noise or masking, which do not reflect structured dependencies among traffic features.

Recent advances in tabular generative modeling make it possible to derive more realistic perturbations from synthetic data distributions. Motivated by this idea, SSDAE introduces a self-supervised denoising framework in which CTGAN-generated samples are transformed into structured perturbations for reconstruction learning, improving latent robustness under one-class intrusion detection.

4. Proposed Model

4.1. Overall architecture of SSDAE

Fig. 1 illustrates the overall architecture of SSDAE, which consists of a distribution-aware perturbation generator and a denoising autoencoder for reconstruction-based anomaly learning.

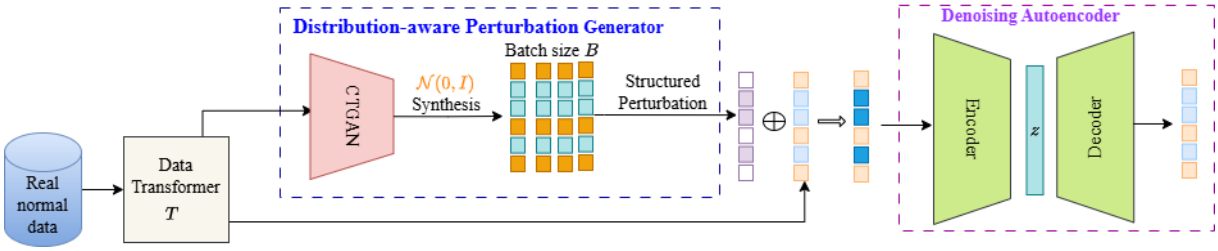


Figure 1. Architecture of SSDAE. Synthetic samples produced by CTGAN are converted into centered perturbations and added to clean observations before reconstruction training. Blue and orange blocks denote continuous and discrete features, respectively.

First, CTGAN is trained on normal traffic to learn benign feature distributions and generate synthetic samples with similar statistical properties. Instead of treating generated outputs as additional training data, the framework derives perturbation vectors from them to construct corrupted inputs for denoising supervision. Perturbations are applied only to continuous attributes, while categorical features remain unchanged to preserve tabular consistency. The resulting noisy inputs are then processed by the denoising autoencoder to reconstruct the corresponding clean samples. The detailed optimization procedure is presented in Section 4.2.

4.2. Training Process of SSDAE

Algorithm 1 describes the optimization procedure of SSDAE, which combines adversarial synthetic generation with reconstruction-based denoising during model training.

At each iteration, CTGAN updates both generator and critic using transformed benign mini-batches. After adversarial optimization, a latent vector sampled from a Gaussian prior is passed through the generator to produce synthetic samples: $\mathbf{x}_{syn} = G(z, c)$ where $z \sim \mathrm{N}(0, I)$

denotes latent noise and c represents the conditional vector. Rather than storing generated outputs, the synthetic batch produced at the current iteration is directly transformed into perturbations. To avoid introducing global bias, each feature is centered by subtracting the batch mean: $\acute{U}_{b,j} = \mathbf{x}_{syn,b,j} - \frac{1}{B}\sum_{i=1}^{B} \mathbf{x}_{syn,i,j}$ (4)

where B denotes the batch size and j indexes the feature dimension.

The resulting perturbation is injected only into continuous attributes of clean samples:

$x_{noisy}[:, \mathrm{C}] = x[:, \mathrm{C}] + \alpha \acute{U}[:, \mathrm{C}]$ (5)

while categorical attributes remain unchanged: $x_{noisy}[:, \mathrm{D}] = x[:, \mathrm{D}]$ (6)

The DAE reconstructs the clean input from x_{noisy} using reconstruction loss: $\mathrm{L}_{SSDAE} = \| x - \hat{x} \|^2$ (7)

At the end of each epoch, reconstruction errors computed on validation data are used to estimate the anomaly threshold for inference.

Algorithm 1: Training Procedure of SSDAE

Input: Normal training set D_{train}; Normal validation set D_{val};
Discrete columns C_{disc}; Epochs E; Batch size B; perturbation coefficient α.

Output: Trained model $\mathrm{M} = (G, C, \mathrm{DAE})$ and anomaly threshold τ.

1. Fit transformer T on D_{train} using C_{disc}.
2. Transform datasets: $\tilde{D}_{train} \leftarrow T(D_{train})$, $\tilde{D}_{val} \leftarrow T(D_{val})$.
3. Initialize conditional sampler, CTGAN (G, C), DAE.
4. **for** $e = 1$ **to** E **do**
5. **for** each *mini-batch* $x \sim \tilde{D}_{train}$ **do**
6. Update critic C and generator G.
7. Generate synthetic batch $\mathbf{x}_{syn}$ using the generator G.
8. Compute centered perturbation using (4)
9. Construct noisy input using (5) and (6)
10. Train DAE using (x_{noisy}, x) with reconstruction loss in (7)
11. **end for**
12. Update best model and estimate percentile-based threshold τ on $\tilde{D}_{val}$.
13. **end for**

5. Experimental settings

5.1. Evaluation Scenarios

SSDAE is evaluated under two scenarios:

Scenario 1: Performance Comparison with Existing Methods. SSDAE is evaluated against widely used anomaly detection methods for NIDS, including LOF [14], DIF [15], AE [16], SAE [3], VAE [17], DAE [7], and DVAE [*17*].

Scenario 2 examines the effect of replacing the deterministic autoencoder in SSDAE with DVAE.

5.2. Datasets

Experiments use three tabular intrusion detection datasets: NSL-KDD [18], CIC-IoT 2023 [19], and CTU13-08 [20], covering diverse traffic and attack characteristics. NSL-KDD provides moderately separable attacks, CIC-IoT 2023 represents large-scale heterogeneous IoT traffic, and CTU13-08 contains botnet traffic with strong overlap between benign and malicious behavior.

5.3. Implementation details and hyper-parameters

SSDAE is implemented by extending the SDV version of CTGAN with a denoising autoencoder. CTGAN uses an embedding dimension of 128, with both generator and critic consisting of two hidden layers of 256 neurons, optimized by Adam. The DAE follows hidden dimensions $0.75n$ and $0.5n$, followed by a latent layer of dimension $\sqrt{n}+1$, where n denotes the input dimension. Adam optimization is applied with a learning rate of 10^{-3} and the perturbation coefficient is fixed at $\alpha = 0.1$. Experiments were performed on an Intel Core i7 workstation with 32 GB RAM and an NVIDIA RTX 4070 Ti GPU (12 GB).

5.4. Evaluation metrics

Detection performance is evaluated using AUC, F1-score, and MCC, which provide complementary assessment under imbalanced intrusion detection settings.

$$\mathrm{AUC} = \int_0^1 \mathrm{TPR}(f)d\mathrm{FPR}(f)\ ;\ \text{F1-score} = \frac{2 \cdot \mathrm{Precision} \cdot \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \tag{9}$$

$$\mathrm{MCC} = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{10}$$

6. Performance Evaluation

This section presents experimental results for SSDAE under two settings: comparative detection evaluation and perturbation sensitivity analysis.

6.1. Performance Comparison with Existing Methods

Table 1. Performance comparison across benchmark datasets.

Dataset	Metric	LOF	DIF	AE	SAE	VAE	DVAE	DAE	SSDAE
NSL-KDD	**AUC**	88.05	93.50	94.80	93.43	95.41	92.36	94.69	**96.27**
	F1	85.81	19.11	86.39	84.36	86.99	71.03	90.62	**91.16**
	MCC	68.17	21.51	71.75	63.40	72.23	57.06	79.37	**79.55**
CIC_IOT 2023	**AUC**	99.42	55.76	99.31	97.23	99.12	92.91	99.57	**99.60**
	F1	99.26	60.00	99.28	99.21	99.30	64.92	**99.41**	99.39
	MCC	76.95	0.89	77.05	75.43	77.34	13.15	79.37	**79.66**
CTU13-08	**AUC**	44.07	98.21	97.79	96.62	97.68	96.46	99.57	**99.95**
	F1	15.63	67.10	69.25	66.21	69.30	66.51	74.86	**75.34**
	MCC	9.81	67.91	67.51	64.04	67.55	64.35	75.16	**75.60**

Table 1 shows that SSDAE achieves the best overall performance across benchmark datasets. On NSL-KDD, SSDAE attains the highest AUC (96.27), F1-score (91.16), and MCC (79.55), indicating improved anomaly discrimination under moderately separable traffic patterns. On CIC-IoT 2023, SSDAE remains highly competitive and achieves the highest MCC (79.66), demonstrating robustness under highly imbalanced IoT traffic. The largest improvement appears on CTU13-08, where SSDAE reaches the best AUC (99.95), F1-score (75.34), and MCC (75.60), suggesting stronger sensitivity to subtle malicious deviations in highly overlapping traffic distributions.

6.2. Effect of Replacing DAE with DVAE

Table 2. Comparison of anomaly detection performance obtained with different reconstruction architectures across three benchmark datasets

Dataset	Metric	DAE	DVAE	SSDVAE	SSDAE
NSL-KDD	**AUC**	94.69	92.36	94.78	96.27
	F1	90.62	71.03	78.51	91.16
	MCC	79.37	57.06	64.16	79.55
CIC_IOT	**AUC**	99.57	92.91	98.47	99.60
	F1	99.41	64.92	96.57	99.39
	MCC	79.37	13.15	47.60	79.66
CTU13-08	**AUC**	99.57	96.46	97.00	99.95
	F1	74.86	66.51	71.62	75.34
	MCC	75.16	64.35	70.87	75.60

Table 2 shows that replacing the deterministic DAE with DVAE generally reduces anomaly detection performance across all datasets. Although SSDVAE improves substantially over the conventional DVAE baseline, its performance remains below SSDAE in most metrics. On NSL-KDD, SSDVAE increases F1-score from 71.03 to 78.51 and MCC from 57.06 to 64.16, but still does not reach the performance of SSDAE. Similar trends are observed on CIC-IoT and CTU13-08, where structured perturbation learning improves variational reconstruction, yet deterministic denoising remains more effective for preserving discriminative latent representations under one-class intrusion detection.

6.3. Training Convergence Analysis

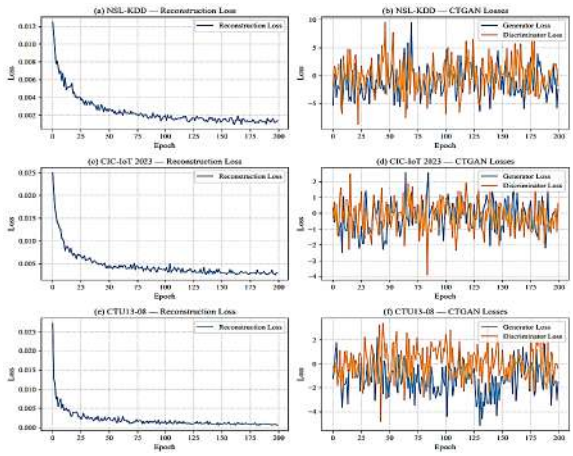


Figure 1. Evolution of reconstruction and adversarial losses during SSDAE training.

This section examines the training convergence of SSDAE using the loss curves in Figure 2. The left panels present reconstruction loss, while the right panels show generator and discriminator losses during CTGAN training. Across all datasets, reconstruction loss decreases rapidly in early epochs before reaching stable convergence, indicating effective learning of normal traffic patterns. CTGAN losses remain bounded with moderate oscillation, suggesting stable adversarial optimization.

Although SSDAE introduces additional training cost because CTGAN and DAE are optimized jointly, the generative module is used only during offline training. During inference, anomaly detection relies solely on the trained denoising autoencoder, so prediction complexity remains comparable to a standard DAE.

7. Conclusion and Future Work

This paper proposes SSDAE, a self-supervised denoising autoencoder for one-class intrusion detection under normal-only training. By learning to reconstruct clean samples from structured perturbations, SSDAE enhances latent stability and improves discrimination of subtle abnormal deviations beyond conventional reconstruction-based autoencoders. Across benchmark datasets, SSDAE consistently maintains strong AUC, F1-score, and MCC under heterogeneous traffic distributions. Future work will explore adaptive perturbation strategies and temporal modeling to improve generalization under evolving network environments.

References

1. Diana, L., Dini, P., & Paolini, D. (2025). *Overview on intrusion detection systems for computers networking security. Computers, 14 (3), 87.*
2. Van, C. N., Phan, V. A., & Nguyen, K. D. T. (2020, November). IoT malware detection based on latent representation. In *2020 12th international conference on knowledge and systems engineering (KSE)* (pp. 177-182). IEEE.
3. Nicolau, M., & McDermott, J. (2018). Learning neural representations for network anomaly detection. *IEEE Transactions on Cybernetics*, *49*(8), 3074-3087.
4. Biggio, B., Corona, I., Maiorca, D., Nelson, B., Šrndić, N., Laskov, P., ... & Roli, F. (2013, September). Evasion attacks against machine learning at test time. In *Joint European conference on machine learning and knowledge discovery in databases* (pp. 387-402). Berlin, Heidelberg: Springer Berlin Heidelberg.

5. Ruff, L., Vandermeulen, R., Goernitz, N., Deecke, L., Siddiqui, S. A., Binder, A., ... & Kloft, M. (2018, July). Deep one-class classification. In *International conference on machine learning* (pp. 4393-4402). PMLR.
6. Xu, L., Skoularidou, M., Cuesta-Infante, A., & Veeramachaneni, K. (2019). Modeling tabular data using conditional gan. *Advances in neural information processing systems*, *32*.
7. Vincent, P., Larochelle, H., Bengio, Y., & Manzagol, P. A. (2008, July). Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th international conference on Machine learning* (pp. 1096-1103).
8. Chen, Z., Yeo, C. K., Lee, B. S., & Lau, C. T. (2018, April). Autoencoder-based network anomaly detection. In *2018 Wireless telecommunications symposium (WTS)* (pp. 1-5). IEEE
9. Vu, L., Nguyen, Q. U., Nguyen, D. N., Hoang, D. T., & Dutkiewicz, E. (2020). Learning latent representation for IoT anomaly detection. *IEEE Transactions on Cybernetics*, *52*(5), 3769-3782.
10. Meidan, Y., Bohadana, M., Mathov, Y., Mirsky, Y., Breitenbacher, D., Shabtai, A., & Elovici'N-BaIoT, Y. (2022). Network-based Detection of IoT Botnet Attacks Using Deep Autoencoders'. *IEEE Pervasive Computing, Special Issue-Securing the IoT (July/Sep 2018)*.
11. Xu, W., Jang-Jaccard, J., Singh, A., Wei, Y., & Sabrina, F. (2021). Improving performance of autoencoder-based network anomaly detection on nsl-kdd dataset. *IEEE Access*, *9*, 140136-140146.
12. Long, C., Xiao, J., Wei, J., Zhao, J., Wan, W., & Du, G. (2022, February). Autoencoder ensembles for network intrusion detection. In *2022 24th International Conference on Advanced Communication Technology (ICACT)* (pp. 323-333). IEEE
13. Torabi, H., Mirtaheri, S. L., & Greco, S. (2023). Practical autoencoder based anomaly detection by using vector reconstruction error. *Cybersecurity (2523-3246)*, *6*(1).
14. Breunig, M. M., Kriegel, H. P., Ng, R. T., & Sander, J. (2000, May). LOF: identifying density-based local outliers. In *Proceedings of the 2000 ACM SIGMOD international conference on Management of data* (pp. 93-104).
15. Zhou, C., & Paffenroth, R. C. (2017, August). Anomaly detection with robust deep autoencoders. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 665-674).
16. Mirsky, Y., Doitshman, T., Elovici, Y., & Shabtai, A. (2018). Kitsune: an ensemble of autoencoders for online network intrusion detection. *arXiv preprint arXiv:1802.09089*
17. An, J., & Cho, S. (2015). Variational autoencoder based anomaly detection using reconstruction probability. *Special lecture on IE*, *2*(1), 1-18.
18. Moustafa, N., & Slay, J. (2015, November). UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). In *2015 military communications and information systems conference (MilCIS)* (pp. 1-6). Ieee.
19. Neto, E. C. P., Dadkhah, S., Ferreira, R., Zohourian, A., Lu, R., & Ghorbani, A. A. (2023). CICIoT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment. *Sensors*, *23*(13), 5941.
20. Garcia, S., Grill, M., Stiborek, J., & Zunino, A. (2014). An empirical comparison of botnet detection methods. *Computers & Security*, *45*, 100-123.

Developing a Deep Learning Approach for Reconnaissance Attack Detection

Van Quan Nguyen
Le Quy Don Technical University
Email: quannv@lqdtu.edu.vn
Contact number: 0972.822.874

Abstract

The increasing scale and heterogeneity of modern network infrastructures have made the detection of early-stage reconnaissance activities more challenging. Reconnaissance traffic often shares similar statistical characteristics with legitimate network behavior, leading to partially overlapping latent representations in deep learning-based intrusion detection systems. Most existing approaches focus on sample-level classification or centroid-based compactness, but they do not explicitly enforce boundary-level separation between normal and reconnaissance regions in the latent space. As a result, the learned representations may lack a clear geometric margin, which affects detection stability. This paper proposes a Region-Margin Deep Auto-Encoder (RMDAE) framework for reconnaissance detection. The proposed model is built upon a Deep Auto-Encoder and introduces an inter-region margin constraint to regulate the geometric structure of latent representations. Specifically, the method models each class as a compact latent region and minimizes intra-class dispersion while maximizing the boundary-to-boundary distance between normal and reconnaissance regions. A safety margin term is incorporated to ensure non-overlapping and well-separated latent distributions. Experimental results on the NSL-KDD and UNSW-NB15 datasets demonstrate that the proposed framework improves latent space separability and achieves competitive detection performance compared to baseline and state-of-the-art methods. These findings indicate that explicit region-level margin modeling can enhance representation learning for reconnaissance detection.

Keywords: *Latent Representation, Deep Metric Learning, Reconnaissance Attack Detection*

I. Introduce

The continuous expansion of digital infrastructures has significantly increased the complexity and scale of modern network environments [1]. Advanced technologies such as cloud computing, large-scale data analytics, and distributed services have enabled efficient communication and data exchange across diverse platforms [2-3]. However, this rapid development has also enlarged the attack surface and introduced new security challenges [4]. In particular, multi-stage cyber attacks have become increasingly prevalent, where reconnaissance activities are often conducted as an initial step to gather information about network topology, services, and potential vulnerabilities before launching subsequent exploitation phases [5].

Reconnaissance traffic is inherently difficult to detect because its behavioral characteristics often resemble legitimate network activities. Port scanning, service probing, and low-rate information gathering can appear statistically similar to benign traffic patterns [6]. As a result, supervised detection models trained on labeled normal and attack data may produce latent representations in which the two classes are partially overlapping. This overlap reduces the stability of the decision boundary and can lead to misclassification, especially in complex and high-dimensional network data.

Recent studies have applied deep learning (DL) techniques to improve intrusion detection performance [7-8]. Deep neural networks and Deep Auto-Encoder (DAE) architectures are

capable of extracting hierarchical and abstract features from heterogeneous network traffic [9-10]. However, many existing supervised approaches focus primarily on minimizing classification loss without explicitly regulating the geometric structure of the latent space. Although some methods introduce metric learning or center-based objectives to enhance intra-class compactness, they typically operate at the sample or centroid level. These strategies do not directly constrain the boundary relationship between normal and reconnaissance regions, and therefore cannot guarantee sufficient separation between the two latent distributions.

We argue that effective reconnaissance detection requires explicit control over the geometric relationship between the latent regions corresponding to normal and attack data. In particular, two properties are desirable: (i) each class should form a compact region in the latent space, and (ii) the regions associated with normal and reconnaissance data should be separated by a clear margin at the boundary level. Without such constraints, latent representations may remain structurally entangled, even when classification accuracy appears acceptable on training data.

To address this issue, this paper proposes a Region-Margin Deep Auto-Encoder (RMDAE) framework for supervised reconnaissance detection. The proposed model is built upon a DAE architecture to learn structured latent representations from both normal and reconnaissance samples. In addition to the reconstruction objective, RMDAE introduces a region-level geometric separation mechanism. Each class is modeled as a compact latent region characterized by a centroid and an adaptive radius. The training objective simultaneously minimizes intra-class dispersion and maximizes the inter-centroid distance. Furthermore, a boundary-to-boundary safety margin constraint is imposed to ensure that the latent regions of normal and reconnaissance data remain non-overlapping and maintain a predefined separation gap. This formulation explicitly regulates the geometric organization of the latent space and enhances the stability of class discrimination. The main contributions of this work are summarized as follows:

• We propose a supervised DAE framework that models normal and reconnaissance traffic as compact latent regions and enforces explicit inter-region margin separation.

• Unlike conventional classification or metric learning approaches, the proposed method directly constrains the boundary-to-boundary distance between class regions, ensuring non-overlapping latent distributions.

• Comprehensive experiments conducted on benchmark datasets, including NSL-KDD and UNSW-NB15, demonstrate that the proposed approach improves latent space separability and achieves competitive performance in reconnaissance detection.

The remainder of this paper is organized as follows. Section II reviews related work on cyber reconnaissance detection. Section III introduces the proposed RMDAE framework. Section IV details the experimental setup. Section V reports and discusses the experimental results. Finally, Section VI concludes the paper.

II. Related Work

In [13], a deep learning approach based on a Deep Belief Network (DBN) was proposed for network scanning detection and evaluated on the NSL-KDD and UNSW-NB15 datasets, demonstrating improved accuracy over several traditional machine learning methods.

Nevertheless, the use of a relatively limited training set may constrain the assessment of its generalization capability in large-scale or real-world network environments.

In [14], a DL model outperformed an SVM for port scan detection on the CIC-IDS-2017 dataset; however, insufficient architectural and implementation details limit reproducibility and thorough evaluation.

In [15], a Sequential Neural Network (SNN) was proposed for detecting TCP port scanning by classifying packets, TCP streams, and message types. However, evaluation on proprietary pcap data and the exclusive focus on TCP-based reconnaissance limit its generalizability to broader, multi-protocol scenarios.

In [16], five ML classifiers were evaluated for port scan detection on the CIC-IDS-2017 dataset using AUC, F-measure, and computational time, with the Decision Tree achieving the best performance.

In [17], a DL-based framework combining a DAE with MMD regularization was proposed to enhance latent representations and improve downstream classifier performance, achieving superior results on multiple benchmark datasets, including NSL-KDD, UNSW-NB15, CIDDS-001, and CIDDS-002.

In [18], a Deep Triplet Auto-Encoder (DTAE) integrates Triplet Loss into a DAE to enhance latent representations by reducing intra-class distances and increasing inter-class separation, with embeddings fed to ML classifiers. Evaluations on NSL-KDD and UNSW-NB15 show improved Accuracy, Precision, Recall, and F1-score over baselines. However, its sample-level triplet constraints do not explicitly model region-level geometric boundaries, relying instead on indirect pairwise distance optimization.

III. Proposed Methodology

1. Region-Margin Deep Auto-Encoder (RMDAE)

In this section, we present in detail the proposed Region-Margin Deep Auto-Encoder (RMDAE), which is designed to address the limitations observed in existing supervised reconnaissance detection models. Specifically, many previous approaches rely on classification objectives or sample-level metric learning strategies that do not explicitly regulate the geometric relationship between normal and reconnaissance representations in the latent space. As a result, latent embeddings of the two classes may partially overlap, leading to unstable decision boundaries and increased false positive rates.

The RMDAE model is developed to achieve two primary objectives. First, it aims to learn compact and discriminative latent representations from both normal and reconnaissance traffic. Second, it introduces an explicit region level geometric constraint to ensure clear separation between the two classes in the latent space. Unlike sample-based distance optimization, the proposed framework models each class as a compact latent region characterized by a centroid and an adaptive radius. The model simultaneously reduces intra-class dispersion and enforces a boundary-to-boundary margin between the two regions.

To achieve these objectives, RMDAE leverages the representation learning capability of a DAE. The Encoder extracts high level latent features from the input data, while the Decoder reconstructs the original samples from the latent space. The reconstruction objective preserves

essential structural information of network traffic and stabilizes the embedding process. Let the training dataset be denoted as:

$$\mathbf{X} = \{\mathbf{x}_i \mid i = 1,2, \ldots n\}$$

where n is the number of training samples, including both normal and reconnaissance data. The latent representations obtained from the Encoder are denoted as:

$$\mathbf{H} = \mathbf{F}_{\boldsymbol{\theta}}(\mathbf{X})$$

and the reconstructed samples are given by:

$$\widehat{\mathbf{X}} = \mathbf{G}_{\emptyset}(\mathbf{H}) = \mathbf{G}_{\emptyset}\ (\mathbf{F}_{\boldsymbol{\theta}}(\mathbf{X}))$$

The overall objective function of RMDAE consists of four components: reconstruction loss, intra-class compactness loss, inter-centroid separation loss, and a boundary-level safety margin constraint. The complete objective is defined as:

$$\boldsymbol{L}_{\mathrm{RMDAE}}\ (\mathbf{X}, \widehat{\mathbf{X}}) = \lambda_1 \boldsymbol{L}_{\mathrm{rec}}(\mathbf{X}, \widehat{\mathbf{X}}) + \lambda_2 \boldsymbol{L}_{\mathrm{intra}}\ (\mathbf{H}) + \ \lambda_3 \boldsymbol{L}_{\mathrm{inter}}\ (\mathbf{H}) + \ \lambda_4 \boldsymbol{L}_{\mathrm{safe}}(\mathbf{H})$$

Here, $\boldsymbol{L}_{\mathrm{rec}}$ denotes the reconstruction loss of the autoencoder. For real-valued data, it is defined as:

$$\boldsymbol{L}_{\mathrm{rec}} = \frac{1}{n} \sum_{i=1}^{n} \|\mathbf{x}_i - \ \widehat{\mathbf{x}_i}\|^2$$

The intra-class compactness term reduces dispersion within each class by minimizing the distance between latent samples and their corresponding centroids. Let $\boldsymbol{\mu}_N$ and $\boldsymbol{\mu}_R$ denote the centroids of normal and reconnaissance samples in the latent space, respectively. The compactness loss is defined as:

$$\boldsymbol{L}_{\mathrm{intra}} \ = \ E_{\mathbf{h} \sim N} \|\mathbf{h} - \ \boldsymbol{\mu}_N\|^2 + E_{\mathbf{h} \sim R} \|\mathbf{h} - \ \boldsymbol{\mu}_R\|^2$$

To enlarge the separation between the two latent regions, the inter-centroid loss maximizes the squared distance between $\boldsymbol{\mu}_N$ and $\boldsymbol{\mu}_R$. Since the optimization is performed via minimization, this term is expressed as:

$$\boldsymbol{L}_{\mathrm{inter}} = \ - \|\boldsymbol{\mu}_N - \ \boldsymbol{\mu}_R\|^2$$

However, centroid separation alone does not guarantee non-overlapping class regions. Therefore, a boundary-level safety margin constraint is introduced. Let r_N and r_R denote the effective radii of the normal and reconnaissance regions, estimated from the distribution of distances between latent samples and their respective centroids. The boundary gap between the two regions is defined as:

$$gap = \ \|\boldsymbol{\mu}_N - \ \boldsymbol{\mu}_R\| \ - r_N - r_R$$

To ensure that the two regions remain separated by at least a predefined margin m, the safety constraint is formulated as:

$$\boldsymbol{L}_{\mathrm{safe}} \ = \ \max\ (0, \mathrm{m} - gap)$$

This constraint penalizes configurations in which the latent regions overlap or fail to maintain sufficient separation. By jointly optimizing these objectives, RMDAE explicitly regulates the geometric structure of the latent space, producing compact and non-overlapping class regions.

The model parameters are optimized using stochastic gradient descent–based methods in an end-to-end manner. The balancing coefficients $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ control the relative influence of each loss component.

Through this formulation, the proposed RMDAE framework ensures structured latent representations and improves the stability and accuracy of supervised reconnaissance detection.

2. General Flow of Our Proposal

In this section, we describe the overall framework of the proposed approach for detecting cyber reconnaissance attacks. The core component of the framework is the RMDAE, which is detailed in previous subsection. The RMDAE model integrates two fundamental elements: the representation learning capability of a DAE and a region-level geometric separation mechanism in the latent space. The proposed architecture leverages the DAE structure to learn compact and informative latent representations from both normal and reconnaissance network traffic. Through the reconstruction objective, the Encoder–Decoder structure captures essential structural characteristics of the input data and maps them into a structured latent space. However, unlike conventional DAE-based approaches that rely solely on reconstruction loss, the proposed method explicitly regulates the geometric organization of the latent space. Specifically, RMDAE models normal and reconnaissance samples as two distinct latent regions characterized by their respective centroids and adaptive radii. The training objective simultaneously minimizes intra-class dispersion and maximizes inter-region separation. In addition, a boundary-to-boundary safety margin constraint is imposed to ensure that the latent regions of normal and reconnaissance data remain non-overlapping and maintain a predefined separation gap. This formulation enables explicit control of the geometric structure of latent representations and promotes stable class discrimination.

As a result, the RMDAE framework produces structured and discriminative latent features that clearly separate normal traffic from reconnaissance activities. These learned representations can be directly used for classification. In this study, we employ a set of conventional ML classifiers to evaluate the effectiveness of the learned features. Specifically, ten classifiers are considered: Decision Tree (DT), Random Forest (RF), Linear Discriminant Analysis (LDA), Naive Bayes (NB), Logistic Regression (LR), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Extra Trees Classifier (ET), Bagging Classifier (BC), and Multi-Layer Perceptron (MLP).

By combining structured latent space modeling with supervised classification, the proposed framework aims to enhance the separability between normal and reconnaissance data and improve overall detection performance.

IV. Experiments

This section describes the datasets used to evaluate our models and provides a comprehensive overview of the experimental setup, including configurations, hyperparameter settings, and the programming environment.

1. Datasets:

To evaluate the effectiveness of the proposed framework, experiments were conducted on widely recognized benchmark datasets, namely NSL-KDD [19] and UNSW-NB15 [20].

Detailed information regarding the datasets used in this study is summarized in Table I. Prior to model training, categorical network traffic attributes are transformed using one-hot encoding to convert them into numerical representations suitable for deep learning models.

To mitigate the impact of class imbalance in the training data, a controlled sampling strategy is adopted, maintaining a ratio of approximately 1.5:1 between normal and reconnaissance samples. This strategy helps prevent bias toward the majority class while preserving sufficient variability within each category.

The performance of the proposed model is evaluated using commonly adopted classification metrics, including Accuracy (Acc), Precision (Pre), Recall (Rec), and F1-score (F1) [17]. These metrics provide a comprehensive assessment of detection effectiveness, particularly in distinguishing normal traffic from reconnaissance activities.

TABLE 1: BENCHMARK DATASETS FOR EVALUATING THE PROPOSED MODELS

No	Datasets	Dimension	Normal Training	Reconnaissance Training	Normal Testing	Reconnaissance Testing
	NSL-KDD	122	67343	11656	9711	1106
	UNSW-NB15	196	55999	10491	37000	3496

2. Experimental Settings:

To comprehensively evaluate the effectiveness of the proposed RMDAE, comparative experiments are conducted against several baseline and state-of-the-art models. Specifically, the proposed framework is compared with: (i) a conventional Deep Auto-Encoder (DAE), (ii) the Maximum Mean Discrepancy Deep Auto-Encoder (MMDAE) [17], and (iii) the Deep Triplet Auto-Encoder (DTAE) [18]. These models represent different strategies for latent space optimization, including reconstruction-based learning, distribution alignment, and sample-level metric learning. All experiments are performed on the training set while maintaining the predefined normal-to-reconnaissance ratio of 1.5:1 to alleviate class imbalance. The same data preprocessing procedure is applied to all models to ensure fairness in comparison. For the proposed RMDAE model, the network architecture consists of seven hidden layers, including symmetric Encoder and Decoder structures. The dimensionality of the bottleneck layer is determined using the heuristic formula [18]: $d_h = [1 + \sqrt{D}]$

where D denotes the number of input features and d_h represents the dimension of the latent space. This configuration ensures a compact yet informative latent representation. The Xavier initialization method is adopted for weight initialization. The Tanh activation function is used in hidden layers. The batch size is set to 256. Model parameters are optimized using the Adadelta optimization algorithm with a learning rate of 0.01. For RMDAE, the balancing coefficients in the objective function are empirically selected to regulate the relative influence of reconstruction, intra-class compactness, inter-region separation, and safety margin constraints. These hyperparameters are tuned to achieve stable training and effective latent space separation. To evaluate the discriminative power of the learned latent representations, features extracted from RMDAE, DAE, MMDAE, and DTAE are fed into ten machine learning

classifiers: DT, RF, LDA, NB, LR, KNN, SVM, ET, BC, and MLP. For fair comparison, the hyperparameters of these classifiers are set to their default configurations, consistent with the settings reported in [17]. Through this experimental configuration, the study aims to systematically assess the impact of different latent space optimization strategies on reconnaissance detection performance. The results of the experiments are reported in detail in Table II.

Table II: PERFORMANCE COMPARISON OF HYBRID DAE-, MMDAE-, DTAE-, AND RMDAE-BASED CLASSIFIERS

DAE	UNSW-NB15				NSL-KDD			
	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
DT	0.709	0.907	0.466	0.616	0.899	0.503	0.822	0.624
RF	0.901	0.931	0.864	0.897	0.970	0.876	0.818	0.846
SVM	0.581	0.605	0.466	0.562	0.941	0.745	0.649	0.694
LDA	0.958	0.932	0.988	0.959	0.878	0.416	0.479	0.445
NB	0.581	0.605	0.466	0.526	0.900	0.509	0.719	0.596
LR	0.581	0.604	0.466	0.526	0.943	0.771	0.625	0.690
KNN	0.964	0.938	0.994	0.965	0.973	0.846	0.895	0.870
ET	0.966	0.941	0.995	0.967	0.969	0.896	0.794	0.842
BC	0.900	0.932	0.862	0.896	0.941	0.710	0.716	0.713
MLP	0.501	0.400	0.801	0.533	0.862	0.425	0.863	0.574
MMDAE	**UNSW-NB15**				**NSL-KDD**			
	Accuracy	**Precision**	**Recall**	**F1-score**	**Accuracy**	**Precision**	**Recall**	**F1-score**
DT	0.720	0.944	0.468	0.626	0.935	0.631	0.872	0.732
RF	0.789	0.962	0.602	0.740	0.949	0.783	0.690	0.734
SVM	0.581	0.605	0.446	0.526	0.934	0.781	0.486	0.599
LDA	0.956	0.921	0.997	0.958	0.941	0.775	0.599	0.676
NB	0.581	0.605	0.466	0.526	0.889	0.472	0.736	0.575
LR	0.581	0.604	0.466	0.526	0.935	0.792	0.498	0.612
KNN	0.965	0.941	0.993	0.966	0.958	0.729	0.937	0.820
ET	0.965	0.939	0.995	0.966	0.964	0.819	0.827	0.823
BC	0.709	0.896	0.473	0.619	0.945	0.720	0.750	0.735
MLP	0.501	0.402	0.801	0.534	0.874	0.425	0.999	0.663
DTAE	**UNSW-NB15**				**NSL-KDD**			
	Accuracy	**Precision**	**Recall**	**F1-score**	**Accuracy**	**Precision**	**Recall**	**F1-score**
DT	0.974	0.964	0.984	0.974	0.902	0.909	0.884	0.896
RF	0.971	0.948	0.996	0.971	0.906	0.901	0.903	0.902
SVM	0.967	0.943	0.996	0.968	0.923	0.873	0.981	0.924
LDA	0.965	0.935	0.999	0.966	0.925	0.893	0.960	0.925
NB	0.967	0.941	0.996	0.968	0.941	0.891	0.999	0.942
LR	0.967	0.941	0.998	0.968	0.832	0.796	0.873	0.833
KNN	0.979	0.965	0.993	0.979	0.918	0.891	0.943	0.917
ET	0.972	0.952	0.995	0.973	0.923	0.873	0.981	0.924
BC	0.977	0.963	0.993	0.978	0.923	0.873	0.981	0.924
MLP	0.967	0.939	0.998	0.968	0.937	0.919	0.952	0.936
DTAE	**UNSW-NB15**				**NSL-KDD**			
	Accuracy	**Precision**	**Recall**	**F1-score**	**Accuracy**	**Precision**	**Recall**	**F1-score**
DT	**0.981**	**0.972**	**0.989**	**0.980**	**0.931**	**0.902**	**0.914**	**0.908**
RF	**0.973**	0.952	0.994	0.972	**0.934**	**0.903**	0.912	0.907
SVM	0.967	0.944	0.991	0.967	0.924	0.882	0.939	0.909
LDA	0.965	0.938	0.994	0.965	0.927	0.890	0.932	0.910
NB	0.968	0.946	0.992	0.968	0.940	0.918	0.934	0.926
LR	0.967	0.942	0.996	0.969	0.853	0.794	0.882	0.836
KNN	**0.983**	**0.975**	**0.998**	**0.986**	**0.946**	**0.925**	**0.943**	**0.934**
ET	**0.979**	**0.967**	**0.997**	**0.982**	**0.941**	**0.916**	**0.945**	**0.930**
BC	**0.977**	0.960	0.994	0.977	**0.935**	**0.904**	0.939	0.921
MLP	0.969	0.948	0.995	0.971	**0.948**	**0.931**	**0.950**	**0.940**

V. Results and discussion

1. Overall Performance Comparison

Table II presents the comparative performance of DAE, MMDAE, DTAE, and the proposed RMDAE framework on the UNSW-NB15 and NSL-KDD datasets using ten ML classifiers. The evaluation metrics include Accuracy, Precision, Recall, and F1-score.

Overall, the proposed RMDAE model achieves competitive and consistently strong performance across both datasets. In contrast to DAE and MMDAE, which rely primarily on reconstruction or distribution alignment, RMDAE introduces explicit region-level margin constraints in the latent space. This design results in more structured embeddings and improved class separability.

- Performance on UNSW-NB15

On the UNSW-NB15 dataset, DAE shows unstable performance for several classifiers, particularly DT, SVM, NB, and LR, where Recall values are relatively low. MMDAE improves recall stability but still exhibits limited gains in overall discrimination capability. DTAE significantly enhances performance across most classifiers, demonstrating the effectiveness of metric-based separation.

The proposed RMDAE further improves performance in several classifiers, particularly DT, KNN, ET, and BC. For example, KNN achieves an Accuracy of 0.983 and an F1-score of 0.986 under RMDAE, which is higher than DTAE. Similarly, ET and DT also show noticeable improvements in Accuracy and F1-score. These improvements indicate that region-level geometric separation benefits classifiers that rely heavily on distance or partition-based decision mechanisms.

However, not all classifiers show large improvements. SVM and LDA present performance comparable to DTAE. This behavior is expected, as these classifiers already perform well when latent representations are approximately linearly separable. Therefore, additional geometric constraints may yield only marginal gains.

On UNSW-NB15, the proposed model demonstrates stable high Recall values across most classifiers, typically above 0.99 for tree-based and distance-based models. This suggests that the boundary-to-boundary margin constraint effectively reduces false negatives in reconnaissance detection.

The improvements are most evident in classifiers that are sensitive to spatial distribution, such as KNN and ET. The enhanced latent region compactness and separation reduce overlap between normal and reconnaissance samples, thereby improving neighborhood consistency and decision purity.

Linear models such as LR and LDA show only minor improvements compared to DTAE. This indicates that when separation is already well established, linear decision surfaces may not fully exploit additional geometric refinement.

- Performance on NSL-KDD

The NSL-KDD dataset presents a more challenging scenario due to its class distribution and feature characteristics. Compared to UNSW-NB15, overall performance values are slightly lower, which is consistent with previous findings in the literature.

On this dataset, RMDAE achieves noticeable improvements in KNN, ET, MLP, and NB. For instance, KNN reaches an Accuracy of 0.946 and an F1-score of 0.934, outperforming

DTAE. MLP also benefits from structured latent features, achieving improved Precision and Recall balance.

For SVM and LDA, performance remains close to that of DTAE, indicating that the proposed region-level constraint does not artificially inflate results across all classifiers. LR shows limited improvement, which is consistent with its reliance on linear decision boundaries and sensitivity to class imbalance.

These observations suggest that the proposed geometric margin primarily enhances models that depend on spatial density and neighborhood relationships.

VI. Conclusion and future work

This paper presented a Region Margin Deep Auto-Encoder (RMDAE) framework for supervised cyber reconnaissance detection. The proposed model integrates DAE-based representation learning with an explicit region-level geometric margin constraint in the latent space. By modeling normal and reconnaissance traffic as compact regions and enforcing boundary-to-boundary separation, RMDAE enhances latent space structure and improves class discrimination. Experimental results on the UNSW-NB15 and NSL-KDD datasets demonstrate that the proposed approach achieves competitive and stable performance compared to DAE, MMDAE, and DTAE, particularly for distance-based and tree-based classifiers. For future work, several research directions can be considered. First, the proposed region-margin formulation can be extended to multi-class intrusion detection scenarios involving multiple attack categories. Second, adaptive or dynamic margin mechanisms may be explored to improve robustness under evolving network conditions. Finally, real-time deployment and evaluation in streaming environments would further validate the practical applicability of the proposed framework.

References

[1] MÖLLER, Dietmar PF. Cybersecurity in digital transformation. In: Guide to cybersecurity in digital transformation: Trends, methods, technologies, applications and best practices. Cham: Springer Nature Switzerland, 2023. p. 1-70.

[2] Andriulo, F. C., Fiore, M., Mongiello, M., Traversa, E., & Zizzo, V. (2024, September). Edge computing and cloud computing for internet of things: A review. In Informatics (Vol. 11, No. 4, p. 71). MDPI.

[3] Al-Sai, Z. A., Husin, M. H., Syed-Mohamad, S. M., Abdin, R. M. D. S., Damer, N., Abualigah, L., & Gandomi, A. H. (2022). Explore big data analytics applications and opportunities: A review. Big Data and Cognitive Computing, 6(4), 157.

[4] Taherdoost, H. (2023). Security and internet of things: benefits, challenges, and future perspectives. Electronics, 12(8), 1901.

[5] Roy, S., Sharmin, N., Acosta, J. C., Kiekintveld, C., & Laszka, A. (2022). Survey and taxonomy of adversarial reconnaissance techniques. ACM Computing Surveys, 55(6), 1-38.

[6] Pittman, J. M. (2023). Machine learning and port scans: A systematic review. arXiv preprint arXiv:2301.13581.

[7] Sharma, N., & Arora, B. (2025). Machine Learning and Deep Learning Models for Anomaly Intrusion Detection in Networks: A Systematic Review. SN Computer Science, 6(7), 832.

[8] Abdulganiyu, O. H., Ait Tchakoucht, T., & Saheed, Y. K. (2023). A systematic literature review for network intrusion detection system (IDS). International journal of information security, 22(5), 1125-1162.

[9] Nicolau, M., & McDermott, J. (2018). Learning neural representations for network anomaly detection. IEEE transactions on cybernetics, 49(8), 3074-3087.

[10] Vu, L., Nguyen, Q. U., Nguyen, D. N., Hoang, D. T., & Dutkiewicz, E. (2022). Deep generative learning models for cloud intrusion detection systems. IEEE Transactions on Cybernetics, 53(1), 565-577.

[11] Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). Deep learning (Vol. 1, No. 2, pp. 1-800). Cambridge: MIT press.

[12] Berman, D. S., Buczak, A. L., Chavis, J. S., & Corbett, C. L. (2019). A survey of deep learning methods for cyber security. Information, 10(4), 122.

[13] Hung Nguyen Viet, Quan Nguyen Van, Linh Le Thi Trang, and Shone Nathan. Using deep learning model for network scanning detection. In Proceedings of the 4th International Conference on Frontiers of Educational Technologies, pages 117–121, 2018.

[14] Dogukan Aksu and M Ali Aydin. Detecting port scan attempts with comparative analysis of deep learning and support vector machine algorithms. In 2018 International congress on big data, deep learning and fighting cyber terrorism (IBIGDELFT), pages 77–80. IEEE, 2018.

[15] Bruce Hartpence and Andres Kwasinski. Combating tcp port scan attacks using sequential neural networks. In 2020 International Conference on Computing, Networking and Communications (ICNC), pages 256–260. IEEE, 2020.

[16] Qasem Abu Al-Haija, Eyad Saleh, and Mohammad Alnabhan. Detecting port scan attacks using logistic regression. In 2021 4th International symposium on advanced electrical and communication technologies (ISAECT), pages 1–5. IEEE, 2021.

[17] Van Quan Nguyen, Thanh Long Ngo, Viet Hung Nguyen, Thi Hang Nguyen, et al. Hybrid of deep auto-encoder and maximum mean discrepancy for cyber reconnaissance detection. In 2023 15th International Conference on Knowledge and Systems Engineering (KSE), pages 1–8. IEEE, 2023.

[18] Nguyen, V. Q., Ngo, L. T., Nguyen, V. H., Nguyen, L. M., & Le-Khac, N. A. (2024, December). A Deep Metric Learning Approach for Cyber Reconnaissance Detection. In 2024 1st International Conference On Cryptography And Information Security (VCRIS) (pp. 1-7). IEEE.

[19] Mahbod Tavallaee, Ebrahim Bagheri, Wei Lu, and Ali A Ghorbani. A detailed analysis of the kdd cup 99 data set. In 2009 IEEE symposium on computational intelligence for security and defense applications, pages 1–6. Ieee, 2009.

[20] Nour Moustafa and Jill Slay. Unsw-nb15: a comprehensive data set for network intrusion detection systems (unsw-nb15 network data set). In 2015 military communications and information systems conference (MilCIS), pages 1–6. IEEE, 2015.

PHÁT TRIỂN MÔ HÌNH HỌC SÂU PHÁT HIỆN TẤN CÔNG TRINH SÁT MẠNG

Tóm tắt:

Sự gia tăng về quy mô và tính không đồng nhất của hạ tầng mạng hiện đại đặt ra thách thức đáng kể đối với bài toán phát hiện hoạt động trinh sát ở giai đoạn sớm. Lưu lượng trinh sát thường có phân bố thống kê gần với lưu lượng hợp lệ, dẫn đến sự chồng lấn trong không gian đặc trưng tiềm ẩn của các hệ thống phát hiện xâm nhập dựa trên học sâu. Phần lớn các phương pháp hiện nay tập trung vào phân loại ở mức mẫu hoặc tối ưu hóa độ gọn nội lớp dựa trên tâm cụm, nhưng chưa mô hình hóa tường minh khoảng cách biên giữa các lớp trong không gian tiềm ẩn. Hệ quả là biểu diễn học được thiếu biên phân tách hình học rõ ràng, ảnh hưởng đến tính ổn định và khả năng khái quát hóa của mô hình. Bài báo đề xuất mô hình Region-Margin Deep Auto-Encoder (RMDAE) cho bài toán phát hiện tấn công trinh sát. Mô hình được xây dựng trên nền tảng Deep Auto-Encoder và tích hợp cơ chế ràng buộc biên liên vùng nhằm điều chỉnh cấu trúc hình học của không gian biểu diễn tiềm ẩn. Cụ thể, mỗi lớp được mô hình hóa như một miền tiềm ẩn có tính gọn cao; đồng thời tối thiểu hóa phương sai nội lớp và tối đa hóa khoảng cách giữa miền bình thường và miền trinh sát. Một thành phần biên an toàn được bổ sung nhằm đảm bảo các phân bố tiềm ẩn không chồng lấn và duy trì khoảng cách phân tách đủ lớn.

Thực nghiệm trên các bộ dữ liệu chuẩn NSL-KDD và UNSW-NB15 cho thấy phương pháp đề xuất cải thiện đáng kể độ phân tách trong không gian tiềm ẩn và đạt hiệu năng phát hiện cạnh tranh so với các phương pháp cơ sở và tiên tiến. Kết quả này cho thấy việc mô hình hóa biên ở mức miền góp phần nâng cao chất lượng học biểu diễn và hiệu quả phát hiện trong bài toán phát hiện tấn công trinh sát.

Từ khoá: Biểu diễn không gian ẩn, Phát hiện tấn công mạng, Học sâu.

A Hybrid Semi-Blind Color Image Watermarking Scheme with Adaptive Schur, SVD Block Selection and Dual-Channel Authentication

Pham Thi Yen, Ta Minh Thanh

Le Quy Don Technical University, Hoang Quoc Viet, Ha Noi, VietNam
phamyen@lqdtu.edu.vn; Tel: 0975650078

Abstract

This study proposes an adaptive semi-blind color image watermarking scheme based on integer Haar discrete wavelet transform (DWT), Schur decomposition, and singular value decomposition (SVD). The watermark is embedded in the LL subband of the blue channel after a one-level decomposition. For each block, two embedding candidates (Schur- and SVD-based) are generated, and the optimal branch is adaptively selected based on decoding validity, embedding reliability, and distortion. In addition, a short digital signature is embedded in the transform domain of the green channel to support authentication. Experimental results demonstrate high imperceptibility (average PSNR of 52.41 dB) and strong robustness against common attacks such as scaling and cropping. However, limited authentication performance reveals a trade-off between visual quality and authentication strength.

Keywords: *Color image watermarking; Semi-blind watermarking; Schur decomposition; Singular value decomposition; Digital signature authentication*

1. Introduction

With the rapid growth of the Internet and multimedia sharing platforms, copyright protection, data authentication, and secure communication have become increasingly important. Digital watermarking has emerged as an effective solution because it enables copyright information to be embedded into digital media without significantly degrading perceptual quality, and current surveys show that imperceptibility, robustness, capacity, and security remain the main design objectives of modern watermarking systems [1, 2]. In general, watermarking methods can be divided into spatial-domain and transform-domain approaches. Spatial-domain methods are usually simple and computationally efficient, but they often suffer from limited robustness against common signal processing and geometric attacks. By contrast, transform-domain methods provide better energy compaction, improved visual quality, and stronger resistance to attacks, which has made them the dominant direction in recent research. A wide range of transform and decomposition tools has been investigated for image watermarking. Representative transform-domain methods include DCT-based embedding [3], DWT-based color watermarking [11], DFT-based robust watermarking [4], contourlet-SVD methods [5], and IWT-SVD frameworks [6, 12]. In parallel, matrix-decomposition-based watermarking strategies such as QR decomposition [7], Schur decomposition [8], and Hessenberg decomposition [9] have demonstrated strong potential for robust blind or semi-blind embedding. In color image watermarking, adaptive mechanisms such as alpha blending [10] and entropy-guided block selection [11] have further shown that adaptive embedding can improve the trade-off between imperceptibility and robustness. Motivated by these advances, recent hybrid watermarking frameworks increasingly combine wavelet transforms with stable matrix decompositions to exploit their complementary advantages. Despite the effectiveness of transform-domain watermarking methods based on DWT, Schur decomposition, and SVD,

several limitations remain. First, most existing approaches rely on a single embedding strategy applied uniformly to all image blocks, which may not be optimal for locally varying image characteristics. Second, adaptive embedding mechanisms are often limited to adjusting embedding strength, rather than selecting between fundamentally different embedding operators. Third, authentication is typically treated as an independent module, rather than being integrated into the watermarking framework in a coordinated manner.

To address these limitations, this paper proposes an adaptive semi-blind watermarking framework that performs blockwise selection between Schur-based and SVD-based embedding strategies. In addition, a dual-channel design is introduced, in which the watermark data and the authentication signature are embedded in separate color channels. This design allows the system to balance robustness, imperceptibility, and authentication capability within a unified framework.

2. Preliminaries

2.1. Integer Haar Discrete Wavelet Transform

The proposed method operates in the transform domain using a lifting-style integer Haar discrete wavelet transform, which maps integer image samples to integer subband coefficients and supports exact inverse reconstruction [6, 12, 13]. For a one-level decomposition, each image component is separated into four subbands, namely LL, LH, HL, and HH, where the LL subband contains the low-frequency approximation and the remaining subbands contain directional detail information. Because the low-frequency component preserves most of the signal energy and is generally more stable under common image processing operations, the proposed payload watermark is embedded in the LL subband of the blue channel after one-level integer Haar DWT [10-12]. In addition, the authentication branch employs a two-level integer Haar decomposition on a separate color channel in order to construct a low-frequency transform representation for robust signature insertion. This separation between payload embedding and authentication embedding reduces direct interference between the two paths and provides a cleaner framework for joint copyright protection and authenticity verification [6, 12].

2.2. Schur decomposition

For a real square matrix A, the real Schur decomposition is written as

$$A = U\mathrm{D}U^{T} \tag{1}$$

where U is an orthogonal matrix, and D is an upper triangular matrix whose diagonal entries correspond to the eigenvalue structure of A [2, 8]. In the proposed method, Schur decomposition is applied blockwise to generate one candidate embedding branch for each selected transform block. Instead of directly modifying pixel values, the method embeds the payload bit through a trace-related representation of the Schur form, followed by inverse reconstruction of the modified block. This branch is attractive because it provides an alternative transform-domain embedding mechanism with different numerical behavior from singular-value-based embedding. In the proposed framework, the Schur-based candidate is not used alone; rather, it is evaluated together with an SVD-based candidate and selected only when it yields a better reliability-distortion trade-off for the current block.

2.3. Singular Value Decomposition

Singular value decomposition (SVD) factorizes a real matrix A as

$$A = U\Sigma V^T \tag{2}$$

where U and V are orthogonal matrices, and Σ is a diagonal matrix containing the singular values of A [5, 6, 12]. Owing to the relative stability of dominant singular values under many common image distortions, SVD has been widely adopted in robust watermarking algorithms [6, 12]. In the proposed framework, SVD is used in two distinct roles. First, it forms the second candidate branch for payload embedding, where the dominant singular value of each selected LL-subband block is quantization-modified to carry one watermark bit. Second, SVD is used in the authentication path, where local transform blocks from the green-channel low-frequency representation are modified to embed a short digital signature. Therefore, SVD serves both as a payload carrier and as the core operator of the side authentication mechanism.

2.4. Seeded block permutation and keyed dither modulation

To improve security and reduce spatial predictability, the proposed method does not embed watermark bits in raster order. Instead, transform blocks are pseudo-randomly permuted by a seed-dependent coordinate ordering mechanism. In the implementation, the initial ordering parameters are derived from SHA-256, while keyed dither values used in quantization are generated from a BLAKE2s-based function [14,15]. The binary watermark itself is also scrambled by a deterministic block permutation before embedding. After block ordering is established, watermark bits are embedded using a dither-modulated quantization rule, along with soft-decision extraction. During decoding, log-likelihood-like reliability values are accumulated across repeated embeddings, and a lower-threshold fallback stage can be activated for degraded conditions. This combination of seeded permutation, keyed dither modulation, and soft voting increases extraction reliability while making unauthorized recovery more difficult.

2.5. Signature generation and authentication mechanism

Besides the payload path, the proposed framework introduces an authentication path in order to verify the consistency of the extracted watermark. A short binary signature is generated from the extracted payload-related bitstream through a hash-based procedure and embedded into a separate transform representation of the green channel. During extraction, the payload watermark is recovered first, the embedded signature is then extracted from the authentication channel, and a reference signature is recomputed from the recovered payload. Authentication is finally determined by comparing these two signatures through a Hamming-distance criterion. This strategy makes the authentication stage dependent on the recovered payload rather than on an externally supplied watermark image, which is more suitable for semi-blind operation. In addition, a tolerance threshold is used instead of strict zero-mismatch rejection, making the authentication decision more stable under mild distortions or attack conditions.

3. Proposed watermarking scheme

3.1. Watermark embedding process

The proposed method embeds the payload watermark in the transform domain of the blue channel. First, the watermark is resized to 64 × 64, binarized, converted into a bitstream, and scrambled by a seed-dependent block permutation. Subsequently, a one-level integer Haar

discrete wavelet transform (DWT) is applied to the blue channel of the host image, and the LL subband is selected for watermark embedding. The LL subband is partitioned into non-overlapping 2 × 2 blocks. These blocks are visited in a pseudo-random order generated from the secret seed using SHA-256 and a logistic-map-based ordering process. Let L be the number of payload bits. According to the available capacity, an odd repetition factor r is automatically selected to improve robustness.

For each selected block A, two embedding candidates are generated.

Schur branch: $$A = UDU^T \quad (3)$$

where U is orthogonal, and D is the Schur form. The watermark bit is embedded into the trace-related scalar:

$$\tau = tr(D) \quad (4)$$

which is modified by dither-modulated quantization. The updated block is reconstructed as:

$$A'Schur = UD'U^T \quad (5)$$

SVD branch: $$A = U\,\Sigma V^T, \quad (6)$$

where the watermark bit is embedded into the dominant singular value. The corresponding candidate is reconstructed as:

$$A'SVD = U\,\Sigma'V^T. \quad (7)$$

The two candidates are compared using self-decodability, soft reliability, and distortion. The better candidate is selected and its branch label is stored as side information:

$$F = \begin{cases} 0, \text{Schur branch}, \\ 1, \text{SVD branch}. \end{cases} \quad (8)$$

After payload embedding, the inverse integer Haar DWT is applied to reconstruct the blue channel. In parallel, an authentication signature is embedded into the green channel. For this purpose, a two-level integer Haar DWT is applied to the green channel, and a short binary signature is embedded into selected 4 × 4 blocks of the second-level low-frequency subband using an SVD-based rule. Finally, the watermarked image and the compressed side-information file are stored.

3.2. Watermark extraction process

The extraction stage requires only the watermarked image and the side-information file. First, the blue channel is decomposed by a one-level integer Haar DWT, and the LL subband is divided into 2 × 2 blocks. The same seed-dependent permutation is regenerated to recover the embedding order. For each selected block, the saved flag determines the extraction branch:

$$F = \begin{cases} 0, \text{extraction is performed from the Schur-based trace representation;}, \\ 1, \text{extraction is performed from the dominant singular value.} \end{cases}$$

Soft-decision votes are accumulated across repeated embeddings, and the final bit is obtained from the sign of the accumulated score. When the confidence is low, a KNN-based denoising fallback is applied to the LL subband and extraction is repeated with a more tolerant threshold. The result with higher confidence is retained. The recovered bitstream is reshaped into a 64×64 watermark and inverse-scrambled to obtain the final extracted watermark. For authentication, the green channel is processed separately. A two-level integer Haar DWT is

applied, the embedded short signature is recovered from the selected 4 × 4 blocks, and majority voting is used across repetitions. The expected signature is then recomputed from the extracted payload bits and compared with the recovered signature by Hamming distance. The watermark is accepted if the mismatch is within a predefined tolerance.

3.3. Execution time analysis

The computational cost of the proposed method is mainly due to the integer Haar DWT, the block-wise Schur/SVD evaluation, and the optional denoising fallback. For a host image of size M × N, both forward and inverse integer Haar DWT have linear complexity O(MN). After one-level decomposition, the LL subband has a size $\frac{M}{2}\times\frac{N}{2}$. Since payload embedding uses fixed 2 × 2 blocks, the number of candidate blocks is $N_b = \frac{MN}{16}$. Because each block has a fixed size and the number of local iterations is constant, the total payload processing cost is also linear $O(N_b) = O(MN)$. The signature branch also uses fixed-size 4 × 4 blocks, so its cost remains linear. Therefore, the overall embedding and extraction complexities are O(MN). Compared with single-branch methods, the proposed method has a higher constant factor because both Schur and SVD candidates are evaluated. However, since all factorizations are performed only on very small blocks, the scheme remains computationally efficient.

3.4. Security of the proposed method

The security of the proposed scheme is based on several layers. First, the watermark is scrambled before embedding using a seed-dependent block permutation. Second, the selected transform blocks are visited in a pseudo-random order generated from SHA-256 and a logistic-map-based process. Third, the embedding rule uses dither-modulated quantization with seed-dependent dither values, making the embedding positions and quantization boundaries difficult to predict. In addition, the method is semi-blind because extraction requires the side information file containing the branch-selection flags and repetition factor. Without this file, reliable recovery is difficult even if the general algorithm is known. A second protection layer is provided by the green-channel authentication path. A short signature is derived from the payload and embedded separately in the transform domain. During extraction, the expected signature is recomputed from the recovered payload and compared with the extracted signature through a tolerant Hamming distance test. This improves resistance to unauthorized extraction and tampering.

3.5. Capacity of the proposed method

For a 512 × 512 host image, one-level integer Haar DWT produces an LL subband of size 256 × 256. Using non-overlapping 2 × 2 blocks, the maximum number of available payload blocks is 16384. In the current implementation, the watermark size is 64 × 64, corresponding to 4096 bits, resulting in a net payload rate of 4096 / (512 × 512 × 3). Since the available capacity is 16384 blocks, the implementation selects the highest odd repetition factor not exceeding the maximum feasible value, yielding r = 3. Consequently, the total number of occupied payload carriers is 4096 × 3 = 12,288, and the gross embedding rate becomes 12288 / (512 × 512 × 3). The authentication branch introduces only 8 signature bits repeated 3 times, resulting in negligible overhead. Although the semi-blind design requires additional side-

information storage, it enables adaptive branch selection and improves the trade-off between robustness and imperceptibility.

4. Experimental results

4.1. Performance metrics

The proposed watermarking scheme is evaluated in terms of imperceptibility, attack-induced distortion, and watermark recovery. Imperceptibility is assessed using PSNR and SSIM between the original image *I* and the watermarked image *Iw.* Attack distortion is measured by the PSNR between *Iw* and the attacked image *Ia*. Watermark recovery performance is evaluated using normalized correlation (NCC) and bit error rate (BER) between the original watermark *W* and the extracted watermark $\hat{W}$,

4.2 Simulation setting

The proposed method was implemented in Python using standard image-processing libraries. Experiments were conducted on 512 × 512 color images with a 64 × 64 binary watermark. The watermark was embedded into the LL subband of the blue channel after one-level integer Haar DWT using 2 × 2 transform blocks, while a short authentication signature was embedded in another channel for integrity verification. Robustness was evaluated under common attacks, including salt-and-pepper noise, Gaussian filtering, median filtering, Wiener filtering, scaling, and cropping (see Table 1). Performance was assessed in terms of visual quality and recovery accuracy using PSNR, SSIM, NCC, and BER.

Fig 1. Binary watermark and representative 512 × 512 color host images used in the experiments.

Table 1. Attacks used in the robustness evaluation.

No.	Attack	Parameter	Description
1	Salt & Pepper	d = 0.05	Impulse noise replaces random pixels with black or white values.
2	Gaussian Filter	k = 3	Blur attack using a 3 × 3 Gaussian kernel.
3	Median Filter	k = 3	Nonlinear filtering using a 3 × 3 median kernel.
4	Wiener Filter	k = 3	Adaptive Wiener filtering with a 3 × 3 neighborhood.
5	Scaling	s = 0.5	Downscaling to 50% of the original size, followed by resizing back.
6	Crop	p = 10%	Removal or masking of a centered region covering 10% of the image area.

4.3. Results and analysis

Discussion note: For comparative discussion, the numerical results of the two reference methods, namely the entropy-based DWT scheme [11] and the alpha-blending-based scheme [10], are reused from their published papers, while the results of the proposed method are obtained from our own implementation and experiments. The "N/A" entries in the comparison tables denote that the corresponding performance metrics were not reported in the original publications of the referenced methods [10], [11].

Table 2 presents the PSNR and SSIM values between the original images and the watermarked images under no-attack conditions. The results show that the proposed method achieves very strong imperceptibility before attack. In particular, it provides the highest PSNR on Lena and Girl, while remaining very close to the best result on Pepper. At the same time, SSIM and NCC are essentially saturated at 1.00 for all three methods, indicating that all schemes preserve high visual fidelity under no-attack conditions. Compared with the entropy-based DWT method [11], the proposed scheme offers comparable or slightly better visual quality, and it substantially outperforms the alpha-blending method [10] in PSNR on all three test images. These results suggest that the adaptive blockwise selection between the Schur and SVD branches improves the imperceptibility-robustness trade-off without sacrificing baseline image quality.

Table 2. Baseline comparison (before attack) for the selected host images.

Image	**method [11]**		**method [10]**		**Proposal method**	
	PSNR	**SSIM**	**PSNR**	**SSIM**	**PSNR**	**SSIM**
Pepper	**51.98**	1.00	44.59	1.00	51.86	1.00
Lena	51.89	1.00	44.75	1.00	**52.29**	1.00
Girl	52.22	1.00	45.39	1.00	**53.08**	1.00

To further assess the robustness of the proposed method, experiments are conducted under various attack scenarios on three standard test images, namely Lena, Pepper, and Girl. The results are summarized in Tables 3-5.

Table 3. Comparison of Lena under the attacks.

Attack	**method [11]**			**method [10]**			**Proposal method**		
	PSNR	NCC	BER	PSNR	NCC	BER	PSNR	NCC	BER
Salt & Pepper (d = 0.05)	18.17	0.97	0.10	18.19	**0.97**	**0.10**	**18.31**	0.92	0.10
Gaussian Filter (k = 3)	22.81	**1.00**	**0.08**	**40.33**	0.99	0.12	33.82	0.88	0.15
Median Filter (k = 3)	**35.12**	**1.00**	**0.09**	33.74	0.96	0.15	32.81	0.79	0.26
Wiener Filter (k = 3)	N/A	N/A	N/A	25.64	**1.00**	**0.11**	**35.23**	0.81	0.23
Scaling (s = 0.5)	N/A	N/A	N/A	31.88	0.96	0.15	**32.71**	**0.97**	**0.04**
Crop (p = 10%)	**15.81**	**1.00**	0.04	15.80	1.00	0.02	15.04	1.00	**0.00**

Table 4. Comparison of Pepper under the attacks.

Attack	**method [11]**			**method [10]**			**Proposal method**		
	PSNR	NCC	BER	PSNR	NCC	BER	PSNR	NCC	BER
Salt & Pepper (d = 0.05)	17.96	**0.98**	0.11	17.89	**0.98**	**0.09**	**18.04**	0.91	0.12
Gaussian Filter (k = 3)	22.26	**1.00**	**0.06**	**38.31**	0.99	0.09	31.57	0.91	0.12
Median Filter (k = 3)	**34.14**	**1.00**	**0.10**	31.93	0.96	0.15	32.92	0.86	0.18
Wiener Filter (k = 3)	N/A	N/A	N/A	34.31	**0.99**	**0.14**	**34.57**	0.85	0.18
Scaling (s = 0.5)	N/A	N/A	N/A	29.95	0.96	0.15	**30.67**	**0.99**	**0.02**
Crop (p = 10%)	15.62	**1.00**	0.04	15.61	**1.00**	0.01	**15.76**	1.00	**0.00**

Table 5. Comparison of Girls under the requested attacks

Attack	**method [11]**			**method [10]**			**Proposal method**		
	PSNR	**NCC**	**BER**	**PSNR**	**NCC**	**BER**	**PSNR**	**NCC**	**BER**
Salt & Pepper (d = 0.05)	17.40	**0.97**	**0.10**	17.39	0.96	0.11	**17.70**	0.92	**0.10**
Gaussian Filter (k = 3)	9.85	**1.00**	0.12	39.36	0.99	0.09	**45.16**	0.96	**0.05**
Median Filter (k = 3)	35.36	1.00	0.10	32.36	0.96	0.15	44.79	0.95	0.06
Wiener Filter (k = 3)	N/A	N/A	N/A	35.30	**0.99**	0.14	**45.66**	0.93	**0.09**
Scaling (s = 0.5)	N/A	N/A	N/A	30.14	0.96	0.15	**48.14**	**0.99**	**0.01**
Crop (p = 10%)	12.67	**1.00**	0.04	12.67	1.00	0.05	**21.35**	1.00	**0.00**

As observed in Tables 3-5, the method performs strongly under scaling and cropping, while its performance degrades under Gaussian and median filtering. The strong performance under scaling can be explained by the use of the LL subband, which preserves low-frequency structural information after resizing and interpolation. In the case of cropping, the pseudo-random block selection and repeated embedding ensure that sufficient watermark information remains available for recovery. The soft-decision voting mechanism further enhances robustness by combining multiple observations. In contrast, Gaussian and median filtering directly smooth or alter local image structures. Since the proposed method embeds watermark information within small blocks, these operations reduce the distinguishability between embedding states. Median filtering, in particular, introduces nonlinear distortions that may disrupt the embedding structure more severely.

The results in Tables 3-5 also demonstrate the effectiveness of the adaptive selection mechanism. Instead of relying on a fixed embedding strategy, the proposed method evaluates both Schur-based and SVD-based candidates for each block and selects the one with better decoding reliability and lower distortion. The SVD-based branch provides stable embedding when singular values are preserved under perturbations, while the Schur-based branch can offer lower distortion or better self-decodability in certain local regions. Therefore, the adaptive mechanism allows the method to exploit the complementary strengths of both decompositions and adapt to local image characteristics.

4.4. Authentication analysis

In addition to imperceptibility and watermark recovery, the proposed method is evaluated in terms of authentication reliability using a short transform-domain signature embedded in the green channel and verified via Hamming distance (HD). Authentication is accepted when HD ≤ 2. Under no attack, 2/3 test images are successfully authenticated (66.67%), with failures indicating limited sensitivity of the signature branch even in clean conditions. Under attacks, watermark extraction remains successful in all cases (18/18), while the authentication success rate drops to 33.33% (6/18), with HD ranging from 3 to 6 in failed cases. These results indicate that the authentication path is more fragile than the payload recovery. Therefore, the current 8-bit signature serves as a lightweight integrity check rather than a robust authentication mechanism, although it still provides useful tamper awareness. Future work will focus on improving authentication robustness through longer signatures, enhanced redundancy, or more reliable embedding strategies.

Table 6. Authentication performance of the proposed method.

Image	**Baseline sig_hd**	**Baseline auth**	**Attack auth success**
Girl	3	False	2/6
Lenna	0	True	2/6
Pepper	1	True	2/6
Overall	–	2/3 (66.67%)	6/18 (33.33%)

5. Conclusion

This paper presents an adaptive semi-blind color image watermarking framework that integrates DWT, Schur decomposition, and SVD within a unified design. By introducing a blockwise adaptive selection mechanism between Schur-based and SVD-based embedding, the proposed method improves robustness while maintaining high visual quality. In addition, a dual-channel architecture is employed to separate payload embedding and authentication. Experimental results demonstrate that the proposed method achieves high imperceptibility and strong robustness, particularly under scaling and cropping attacks. However, the authentication performance remains limited due to the lightweight design of the signature embedding. Future work will focus on enhancing the authentication module, for example, by increasing signature redundancy or incorporating error-correcting codes, as well as extending the framework to more complex attack scenarios.

References

1. P. Aberna and L. Agilandeeswari, "Digital image and video watermarking: methodologies, attacks, applications, and future directions," *Multimedia Tools and Applications, vol. 83, no. 2, pp. 5531-5591, 2024. doi:10.1007/s11042-023-15441-7*
2. A. Tiwari and V. K. Srivastava, "Image watermarking techniques based on Schur decomposition and various image invariant moments: a review," *Multimedia Tools and Applications, vol. 83, no. 6, pp. 16447-16483, 2024. doi:10.1007/s11042-023-15919-4*
3. H.-J. Ko, C.-T. Huang, G. Horng, and S.-J. Wang, "Robust and blind image watermarking in DCT domain using inter-block coefficient correlation," *Information Sciences, vol. 517, pp. 128-147, 2020. doi:10.1016/j.ins.2019.12.015*
4. K. Fares, A. Khaldi, and S. Euschi, "A robust blind color image watermarking based on Fourier transform domain," *Optik, vol. 208, p. 164562, 2020. doi:10.1016/j.ijleo.2019.164562*
5. L. Song, X.-C. Sun, and Z.-M. Lu, "Robust blind watermarking algorithm based on contourlet transform with singular value decomposition," IEICE Transactions on Fundamentals of Electronics, *Communications and Computer Sciences, vol. 104, no. 3, pp. 640-643, 2021. doi:10.1587/transfun.2020EAL2062*
6. T. Zhu, W. Qu, and W. Cao, "An optimized image watermarking algorithm based on SVD and IWT," *The Journal of Supercomputing, vol. 78, no. 1, pp. 222-237, 2022. doi:10.1007/s11227-021-03897-7*
7. P. T. Nha, T. M. Thanh, and N. T. Phong, "Consideration of a robust watermarking algorithm for color image using improved QR decomposition," *Soft Computing, 2022. doi: 10.1007/s00500-022-06841-1*
8. Q. Su, X. Zhang, and G. Wang, "An improved watermarking algorithm for color image using Schur decomposition," *Soft Computing, vol. 24, no. 1, 2020. doi:10.1007/s00500-019-03961-9*
9. P. T. Nha, T. M. Thanh, and N. T. Phong, "An effective embedding algorithm for blind image watermarking technique based on Hessenberg decomposition," Applied Intelligence, vol. 53, no. 21, pp. 25467-25489, 2023. doi:10.1007/s10489-023-04903-y
10. S. Kumar and B. K. Singh, "An improved watermarking scheme for color image using alpha blending," *Multimedia Tools and Applications, vol. 80, no. 9, pp. 13975-13999, 2021. doi: 10.1007/s11042-020-10397-4.*
11. S. Kumar and B. K. Singh, "DWT-based color image watermarking using maximum entropy," *Multimedia Tools and Applications, vol. 80, no. 10, pp. 15487-15510, 2021. doi: 10.1007/s11042-020-10322-9.*
12. Y. Luo, L. Li, J. Liu, S. Tang, L. Cao, S. Zhang, S. Qiu, and Y. Cao, "A multi-scale image watermarking based on integer wavelet transform and singular value decomposition," *Expert Systems with Applications, vol. 184, p. 115449, 2021. doi: 10.1016/j.eswa.2020.114272.*
13. W. Sweldens, "The lifting scheme: A construction of second generation wavelets," *SIAM Journal on Mathematical Analysis, vol. 29, no. 2, pp. 511-546, 1998. doi: 10.1137/S0036141095289051.*
14. National Institute of Standards and Technology (NIST), "Secure Hash Standard (SHS)," *Federal Information Processing Standards Publication (FIPS PUB) 180-4, 2015. doi: 10.6028/NIST.FIPS.180-4.*
15. J.-P. Aumasson, S. Neves, Z. Wilcox-O'Hearn, and C. Winnerlein, "BLAKE2 Cryptographic Hash and Message Authentication Code (MAC)," RFC 7693, *Internet Engineering Task Force (IETF), Nov. 2015. doi:10.17487/RFC7693*

Phương pháp thủy vân ảnh màu bán mù lai với lựa chọn khối Schur–SVD thích ứng và xác thực hai kênh

Tóm tắt: *Nghiên cứu này đề xuất một phương pháp thủy vân ảnh màu bán mù thích ứng dựa trên biến đổi wavelet rời rạc (DWT), phân rã Schur và phân rã giá trị riêng (SVD). Thủy vân được nhúng vào dải con LL của kênh màu Blue sau biến đổi DWT mức 1 . Với mỗi khối, hai phương án nhúng (dựa trên Schur và SVD) được xây dựng, và phương án tối ưu được lựa chọn thích ứng dựa trên tính hợp lệ khi giải mã, độ tin cậy nhúng và mức độ biến dạng. Ngoài ra, một chữ ký số ngắn được nhúng trong miền biến đổi của kênh màu Green nhằm hỗ trợ xác thực. Kết quả thực nghiệm cho thấy phương pháp đạt chất lượng ảnh tốt (PSNR trung bình 52,41 dB) và có khả năng bền vững cao trước các cuộc tấn công phổ biến như co giãn và cắt xén. Tuy nhiên, hiệu năng xác thực còn hạn chế, cho thấy sự đánh đổi giữa chất lượng thị giác và độ mạnh xác thực.*

Từ khóa: Thủy vân ảnh màu; Thủy vân bán mù; Phân rã Schur; Phân rã giá trị riêng; Xác thực chữ ký số

Financial Crime in Decentralized Finance: A Survey

Tran Nam Khanh[1] *, Ta Minh Thanh[1]
Học viện Kỹ thuật quân sự;
Email: khanhtn107195@lqdtu.edu.vn; Tel: 096 166 1097

Abstract

The rapid growth of Decentralized Finance (DeFi) has transformed the delivery of financial services by removing centralized intermediaries and enabling open, borderless participation. However, these same characteristics have also created new opportunities for fraud, exploitation, and illicit financial activity. The pseudo-anonymous nature of DeFi, its cross-border transaction flows, high composability, and limited reliance on traditional know-your-customer and anti-money laundering controls make it an attractive environment for smart contract exploits, market manipulation, scams, rug pulls, and money laundering. This paper presents a comprehensive survey of harmful and illicit activity in the DeFi ecosystem. Drawing on academic literature, industry reports, and regulatory perspectives, the study develops a structured multi-dimensional taxonomy of DeFi-related crime and abuse, categorized by actors, nature, affected technology layers, and specific tactics. It further maps these categories to prevention, detection, and response/recovery countermeasures across technical, governance, and regulatory dimensions.

Keywords: *Blockchain; Decentralized Finance; Cryptocurrency; Financial Crime; DeFi crime;*

1. Introduction

DeFi is a blockchain-based peer-to-peer financial system in which smart contracts automate the execution of financial services on distributed ledgers [1, 2]. Rather than representing only a new form of digital asset exchange, DeFi has evolved into a broader programmable financial infrastructure covering lending, trading, derivatives, and asset management [3, 4]. Its defining properties include non-custodial asset control, permissionless access, public auditability, and composability, together with transparency, accessibility, and transaction atomicity [1, 5-7].

These properties distinguish DeFi sharply from traditional finance. In conventional financial systems, centralized intermediaries retain discretionary authority over transaction approval, asset custody, and dispute resolution [2, 5, 12]. By contrast, DeFi relies on smart-contract-based execution, enabling pseudonymous participation, near real-time settlement, and generally irreversible transactions [1, 2, 6, 11]. This architectural shift also redistributes responsibility: whereas traditional finance is embedded in mature legal and consumer-protection frameworks, DeFi places a greater share of technical and financial risk directly on users [5, 9, 17, 18].

At the same time, the design features that make DeFi innovative also create distinctive conditions for harmful and illicit activity. In DeFi, the boundary between technical malfunction and criminal conduct is often blurred, because an attacker may exploit legitimate contract logic in a way that remains executable yet produces systemic harm [3, 8]. Fraudulent schemes can also be automated and replicated at low cost, as illustrated by so-called “Smart Ponzis” [16]. Moreover, pseudonymity and limited centralized oversight make attribution, accountability, and

* Email: khanhtn107195@lqdtu.edu.vn

liability substantially more difficult than in traditional finance, especially in cases such as rug pulls [6].

The magnitude of the problem is substantial. Between 2017 and 2022, the crypto-asset sector incurred at least $30 billion in losses related to criminal activity and financial exploitation [19]. A larger share of these losses occurred in CeFi, where they were driven mainly by conventional fraud patterns adapted to digital platforms [19]. DeFi nevertheless accounted for roughly $10 billion in losses, primarily associated with smart-contract vulnerabilities and the structural complexity of interconnected protocols [20]. These failures can extend beyond protocol-level losses and trigger wider systemic effects. The collapse of the Terra ecosystem in May 2022 is a prominent example: flaws in protocol design erased more than $50 billion in market capitalization [20] and were followed by a sharp decline in DeFi total value locked, from above $200 billion in April 2022 to approximately $80-85 billion by July 2022 [4].

Prior review studies have addressed DeFi from important but largely separate perspectives. Existing surveys have examined the broader DeFi research landscape [21], DeFi security challenges and defenses [4], and financial-crime compliance issues in decentralized environments [22]. However, these perspectives are typically treated in isolation. As a result, the literature still lacks a concise survey that places financial crime in DeFi at the center of analysis while jointly examining technical exploits, economic manipulation, governance abuse, scams, and illicit-finance practices.

To address this gap, and based on the structured review methodology presented in Section 2, this survey makes the following contributions:

1. The paper classifies DeFi-related crime and abuse according to actors, the nature of the offense, affected technology layers, and concrete tactics, thereby moving beyond isolated incident descriptions.

2. The study links major crime categories to prevention, detection, and response/recovery measures, integrating technical safeguards with governance and regulatory mechanisms.

3. The paper highlights how core DeFi properties, especially composability, pseudonymity, and immutability, simultaneously expand innovation and reshape attack surfaces, enforcement difficulties, and accountability gaps.

4. The survey argues that effective DeFi risk mitigation requires a hybrid approach combining automated analysis and verification with human oversight, governance controls, and coordinated institutional responses

2. Methodology

This study adopts a structured literature review to identify and synthesize prior work on financial crime and harmful activity in DeFi. Given the interdisciplinary nature of the topic, the review considers not only academic studies but also selected regulatory and industry reports when they provide direct evidence on attack patterns, illicit-finance risks, and countermeasures.

The search was conducted across major academic databases, including Scopus, Web of Science, IEEE Xplore, ACM Digital Library, ScienceDirect, SpringerLink, SSRN, and Google Scholar. To complement the academic literature, selected reports from institutions such as the U.S. Department of the Treasury, FATF, BIS, Chainalysis, and TRM Labs were

also considered. Search strings combined DeFi-related terms with terms related to crime, abuse, and security, such as: ("decentralized finance" OR DeFi) AND (crime OR fraud OR scam OR exploit OR attack OR phishing OR "rug pull" OR "money laundering"). The search was limited to English-language sources published during January 2019 and December 2025. In addition, references of highly relevant studies were manually examined to identify further related work.

Sources were included if they examined DeFi-specific attacks, abuse patterns, illicit-finance risks, taxonomies, detection methods, mitigation mechanisms, or regulatory responses. Sources were excluded if they focused on blockchain or cryptocurrency issues without a clear DeFi connection, addressed only adoption or market behavior, or lacked substantive analytical content. After duplicate removal, records were screened by title and abstract, followed by full-text assessment of potentially relevant studies.

For each included source, the review extracted information on crime type, actor, affected technology layer, tactic, impact, and countermeasure. The material was then synthesized qualitatively to develop the taxonomy in Section 4 and the crime-to-countermeasure mapping in Section 5.

3. Background

The DeFi Ecosystem is defined as a peer-to-peer financial framework operating on distributed ledgers, such as blockchains, which fundamentally bypasses traditional financial intermediaries.

Core Technical Pillars:

- Smart Contracts: These are autonomous, self-executing computer programs that encode legal agreements or complex business logic. They operate independently of human intervention once deployed to the network. [7, 11, 12];
- Composability: Often referred to as "Money Legos," this is the capacity for DeFi protocols to interoperate and integrate seamlessly, allowing developers to construct sophisticated financial products by layering existing protocols. [1, 2, 4];
- Atomicity: A critical technical attribute ensuring that multiple financial actions within a single transaction are executed as an inseparable unit-functioning on an "all-or-nothing" basis. [1, 12].

Key Structural Entities:

- Decentralized Exchanges (DEXs): Platforms enabling direct asset swaps between users through Automated Market Makers (AMMs) and Liquidity Pools, rather than relying on a centralized limit order book [1, 4, 27];
- Decentralized Autonomous Organizations (DAOs): Governance structures where holders of governance tokens exercise voting rights to direct protocol modifications and strategic decisions. [1, 9, 28];
- Oracles: Middleware mechanisms designed to ingest real-world data (such as external asset pricing) into the blockchain for utilization by smart contracts [1, 4, 23];
- Cross-Chain Bridges: Infrastructure protocols that facilitate the migration of assets and information across disparate blockchain networks [7, 23].

Foundational Concepts in Digital and Financial Criminology:

- Money Laundering: The process of obscuring the provenance of illicit funds. In the DeFi context, this is often achieved through privacy mixers (e.g., Tornado Cash) or "chain-hopping"-systematically moving assets across multiple bridges to sever the audit trail. [1, 5, 9];
- Maximal Extractable Value (MEV): The maximum value that miners or validators can capture by strategically reordering, inserting, or omitting transactions within a block. While often considered a technical byproduct of consensus, it can be leveraged for market manipulation. [1, 12].

4. Crime Taxonomy for DeFi

Table 1. Crime Taxonomy for DeFi

	Category	Description and representative characteristics
Actors	**Target**	Legitimate protocols or developers that are attacked by external malicious actors due to technical vulnerabilities or human error. This category represents the largest share of losses, accounting for 52% of incidents and up to 83% of total financial losses [20].
	Perpetrator	The developer or creator of a DeFi service is themselves the fraudster. This pattern is typical of rug pull schemes, where the service provider is effectively identical to the thief. It accounts for 41% of incidents but only 17% of total financial losses [20].
	Compromised Intermediary	Malicious actors impersonate legitimate DeFi services in order to approach and defraud users, for example through phishing attacks.
Nature of the crime	**Technical Security Exploits**	These involve the exploitation of vulnerabilities in the source code or logic of smart contracts. A defining feature is atomicity: if the exploit fails, the transaction reverts, leaving the attacker to bear only the gas cost. Representative examples include reentrancy, integer overflow, and logic flaw exploitation [1, 5, 20, 21].
	Economic Security Exploits	These involve manipulation of the economic design or incentive mechanisms of a protocol. Unlike technical exploits, they are typically non-atomic, require substantial capital, and expose the attacker to market risk. Common examples include oracle-based price manipulation, sandwich attacks, and exploitation of unreasonable risk parameters [1, 20].
Affected techn-ology layer	**Network Layer [NET]**	Attacks at this layer target communication infrastructure in order to distort a node's perception of network state. Representative examples include DNS hijacking, BGP hijacking, and eclipse attacks [20].
	Consensus Layer [CON]	These attacks target the integrity of the ledger or the block production process. Examples include 51% attacks, selfish mining, and validator-driven manipulation of transaction ordering [20].
	Smart Contract Layer [SC]	This layer accounts for 42% of incidents [29, 30]. It includes implementation and logic flaws such as reentrancy, arithmetic errors, and access control flaws [29].
	Protocol Layer [PRO]	This layer covers attacks arising from unsafe financial design or risky composability across protocols. One example is the use of LP tokens as collateral without adequately accounting for the risk of the underlying assets [1, 31].
	Auxiliary Services [AUX]	This layer includes external operational failures or off-chain weaknesses that facilitate attacks, such as private key leakage, user interface errors, or manipulation of off-chain oracle services [20].
Criminal tactics	**Flash Loan and Oracle Manipulation**	In this tactic, attackers use uncollateralized flash loans to create leverage, manipulate AMM prices, and distort oracle-fed valuations in order to withdraw assets beyond legitimate limits [28]. This tactic is closely related to the structural oracle problem in DeFi [31].
	Transaction-Ordering Exploitation	Validators or bots extract value by reordering transactions. This includes front-running, back-running, and sandwich attacks, and is closely associated with MEV [32].

	Rug Pull	The development team intentionally appropriates investors' funds, for example through liquidity withdrawal, unauthorized token minting, or malicious contract modification [33].
	Automated Ponzi Schemes	Fraudulent redistribution schemes are encoded directly into smart contracts, enabling funds from later participants to be automatically transferred to earlier participants [16].
	Phishing and Deception	Users are deceived into revealing credentials or interacting with malicious interfaces, for example through DNS attacks or interface impersonation.
	Governance Attacks	Attackers exploit concentrated voting power or thin governance-token liquidity to pass malicious proposals, such as draining the treasury or inserting a backdoor into the protocol [34].
	Cross-chain Bridge Attacks	Attackers exploit weaknesses in bridge architecture or implementation, including multi-signature weaknesses and semantic inconsistencies between blockchains, in order to misappropriate locked assets and TVL [26, 35].
	Money Laundering	Illicit funds are routed through DeFi infrastructure to obscure traceability, including through mixers such as Tornado Cash, chain-hopping, and cross-chain bridges [9].

5. Mapping Crimes to Countermeasures

A pivotal contribution of this survey is the systematic mapping between identified crime categories and their corresponding layers of countermeasures, structured around the functional domains of Prevention (P), Detection (D), and Response/Recovery (R). This framework further distinguishes between technical mitigations, such as cryptographic protocols and smart contract logic, and administrative/legal measures, including governance policies and regulatory compliance. While automated tools have significantly matured, recent evaluations indicate they still struggle with complex DeFi-specific logic, necessitating a hybrid approach that integrates automated verification with human-in-the-loop oversight [30, 36].

The comprehensive mapping between major DeFi crime categories and the multi-layered defense strategies identified in current literature and industry practice is presented in the table below. This framework serves as a reference for practitioners, auditors, and protocol developers seeking to implement holistic security strategies.

Table 2. Mapping Crimes to Countermeasures

Crime Category	**Prevention (P)**	**Detection (D)**	**Response & Recovery (R)**
Protocol Exploits & Compositional Abuse	Specification-driven development; Formal verification; Property-based testing with explicit invariants [1]. Dependency risk assessments to model cascading failures.	On-chain analytics for anomalous transaction flows; Runtime invariant monitoring (e.g., collateralization ratios).	Emergency circuit breakers; Timelocks for sensitive changes; Structured patching procedures respecting contract immutability.
Price Manipulation (Oracles, AMMs, Flash Loans)	Multi-source oracle design; Time-weighted average pricing (TWAP); Outlier removal; Slippage limits; Price sanity checks [31, 37]. Conservative collateral parameters.	Semantic extraction systems (e.g., DeFiRanger) to detect high-level attack patterns [37].	Transaction reversal (consensus-based); DAO treasury or insurance fund compensation [33].
Transaction Ordering Exploits	Threshold encryption or commit-reveal schemes for transaction privacy [32]. Batch auction designs reducing frontrunning profitability. Order-independent application design [18].	Mempool monitoring and bot network analysis; MEV-specific measurement tools [32].	Validator selection mechanism adjustments; Community norms on ethical MEV; Legal classification under securities law.

Crime Category	Prevention (P)	Detection (D)	Response & Recovery (R)
Rug Pulls & Token Scams	Source code verification; Owner privilege transparency; Liquidity locks via vesting/timelock contracts [33]. CEX/DEX listing criteria with rug pull risk assessment.	Investor warning systems for ownership concentration, unlimited minting, thin liquidity [33]. Community-driven fraud reporting.	Cross-border law enforcement coordination for asset recovery; Founder accountability tracking [38].
Governance Attacks	Minimum participation thresholds; Delegation to trusted representatives; Flash loan voting restrictions [33]. Timelock and guardian/veto mechanisms.	Monitoring voting power distribution anomalies; Community alerts for suspicious proposals.	Hard-fork or governance decision nullification (high consensus required). Emergency multisig guardian intervention.
Cross-Chain Bridge Attacks	Trust-minimized architectures (light clients, multi-party computation); Cross-chain static analysis tools [19, 29].	Real-time TVL monitoring; Large cross-chain flow anomaly detection [39].	Automated bridge pausing; Asset freezing via CEX/DEX collaboration [40].
Money Laundering & Sanction Evasion	Travel rule compliance for Virtual Asset Service Providers; Compliance-friendly privacy protocols [41]. AML/CFT requirements at on-ramp/off-ramp points [42, 43].	Large-scale on-chain cluster analysis; High-risk service monitoring (mixers, bridges) [44].	Targeted address sanctions; Public-private investigation partnerships [33, 42].

5.1. Technical and Administrative Synergy

The effectiveness of the countermeasures outlined above depends on the synergy between technical implementation and administrative oversight. In the domain of Protocol Exploits and Compositional Abuse, security relies on rigorous development lifecycles. This includes the adoption of specification-driven processes that utilize formal verification and property-based testing to enforce explicit invariants [1]. Furthermore, dependency risk assessments are crucial for modeling the cascading impacts of failures in integrated protocols, ensuring that the modular nature of DeFi does not become its primary vulnerability. Multi-round third-party audits, including deep manual reviews, provide additional assurance beyond automated tools [6].

5.2. Addressing Market and Governance Integrity

Oracle manipulation and MEV exploits represent sophisticated threats to market integrity. Defensive strategies for Price Manipulation focus on designing resistant oracles using multi-source data and outlier removal mechanisms [31]. Detection systems like DeFiRanger have demonstrated the utility of high-level semantic extraction from raw transactions to identify complex attack patterns often missed by bytecode-level analysis [37]. Similarly, Transaction Ordering Exploits are mitigated through infrastructure-level changes like threshold encryption and application-level designs like batch auctions, which reduce the profitability of frontrunning [15, 16].

In the realm of Governance and Ecosystem Security, protecting the voting process is paramount. Countermeasures for Governance Attacks include designing mechanisms with participation thresholds and restricting flash loans for voting. For Rug Pulls and Token Scams, the focus shifts toward transparency standards, such as source code verification and publicizing

owner privileges. The proliferation of scams necessitates a combination of transparency standards and community-driven oversight [33].

5.3. Infrastructure and Compliance Resilience

Cross-Chain Bridge Attacks and Money Laundering represent distinct challenges requiring different defensive approaches. Bridges represent significant points of failure due to their architectural complexity and high TVL [26, 29]. Countermeasures include moving toward trust-minimized architectures (e.g., light clients, multi-party computation) and utilizing static analysis tools specifically designed for cross-chain semantics [29]. Detection relies on continuous monitoring of TVL and large cross-chain flows to identify potential drainage in real-time [39].

Finally, Money Laundering and Sanction Evasion require a combination of trust-minimized architectures and robust compliance frameworks, balancing the inherent privacy of blockchain with the necessity of regulatory oversight [42, 43]. Compliance in DeFi requires implementing "travel rule" compliance for Virtual Asset Service Providers at entry/exit points and developing "compliance-friendly" privacy protocols [41]. Large-scale on-chain analysis to identify address clusters and high-risk service usage (e.g., mixers or sanctioned entities) supports detection efforts [44]. Response measures include applying targeted sanctions to illicit addresses and fostering public-private partnerships between regulators, blockchain analytics firms, and DeFi platforms [33, 34, 45].

6. Conclusion

This research provides a comprehensive overview of the evolving landscape of DeFi crime, contrasting it with traditional financial systems and detailing its unique vulnerabilities. We have established that while DeFi offers unprecedented transparency, accessibility, and composability, these very characteristics also introduce novel avenues for exploitation. The study meticulously categorizes DeFi crimes based on actors (Target, Perpetrator, Compromised Intermediary), the nature of the crime (Technical Security Exploits, Economic Security Exploits), affected technology layers (Network, Consensus, Smart Contract, Protocol, Auxiliary Services), and criminal tactics (Flash Loan and Oracle Manipulation, Transaction-Ordering Exploitation (MEV), Rug Pull, Automated Ponzi Schemes, Phishing and Deception, Governance Attacks, Cross-chain Bridge Attacks, Money Laundering).

A pivotal finding is the systematic mapping of these diverse crime categories to a multi-layered defense framework encompassing Prevention, Detection, and Response/Recovery. This framework emphasizes the critical synergy between technical mitigations - such as formal verification, multi-source oracle design, and trust-minimized bridge architectures - and administrative/legal measures, including governance policies, regulatory compliance, and public-private partnerships. The analysis underscores that while automated tools are advancing, the inherent complexity of DeFi necessitates a hybrid approach that integrates automated verification with human oversight to effectively combat sophisticated attack vectors.

Ultimately, the proliferation of DeFi necessitates a robust and adaptive security paradigm. This research highlights that the demarcation between technical error and criminality is often blurred in DeFi, and the evolutionary trajectory of fraud, exemplified by 'Smart Ponzis' and 'Rug Pulls,' poses significant challenges to accountability and liability. Addressing these

challenges requires continuous innovation in security protocols, enhanced transparency standards, and collaborative efforts across the ecosystem to safeguard users and maintain the integrity of this nascent financial frontier.

References

1. Zhou, E., Qin, K., Cheng, T., Huo, Y., Li, H., & Livshits, B. (2022). SoK: Decentralized finance (DeFi) attacks. arXiv preprint arXiv:2208.13035.
2. Huber, S., & Probst, L. (2023). Decentralized lending and its users: Insights from Compound. Journal of International Financial Markets, Institutions and Money, 86, 101805.
3. Qin, K., Zhou, L., & Livshits, B. (2021). On the just-in-time discovery of profit-generating transactions in DeFi protocols. arXiv preprint arXiv:2103.02228.
4. Li, M., Meng, X., Wang, H., Li, Y., & Wang, J. (2022). A survey of DeFi security: Challenges and opportunities. arXiv preprint arXiv:2206.11821.
5. Qin, K., Zhou, L., Livshits, B., & Gervais, A. (2021). CeFi vs. DeFi - Comparing centralized to decentralized finance. arXiv preprint arXiv:2106.08157.
6. Schueffel, P. (2021). DeFi: Decentralized finance - An introduction and overview. Journal of Innovation Management, 9(1), 81-91.
7. Mikula, S., & Fabian, M. W. (2021). Quantifying blockchain extractable value: How dark is the forest? arXiv preprint arXiv:2101.05511.
8. Qin, K., Zhou, L., & Livshits, B. (2021). Attacking the DeFi ecosystem with flash loans for fun and profit. In Financial Cryptography and Data Security (pp. 3-32). Springer.
9. U.S. Department of the Treasury. (2023). Illicit finance risk assessment of decentralized finance. https://home.treasury.gov/system/files/136/DeFi-Risk-Full-Review.pdf
10. Klages-Mundt, A., & Minca, A. (2019). (In)stability for the blockchain: Deleveraging spirals and stablecoin attacks. arXiv preprint arXiv:1906.02152.
11. Al-Mansour, S., & Al-Rimy, B. A. S. (2025). A systematic review of security innovations in decentralized finance (DeFi) using blockchain technology. Informatica, 49(1).
12. Society of Actuaries Research Institute. (2022). A risk classification framework for decentralized finance protocols. https://www.soa.org/4aa5bb/globalassets/assets/files/resources/research-report/2022/decentralized-finance-protocols.pdf
13. Aleksieva, N., Zeadally, S., & Kumar, A. (2022). Security threats classification in blockchains. St. Cloud State University Repository.
14. Ashraf, M., AlOwayed, A., Alghamdi, M. A., & Alshahrani, M. M. (2022). A machine learning and blockchain based efficient fraud detection mechanism. Electronics, 11(19), 3133.
15. Zhao, Y., Wang, Y., & Li, X. (2024). From shadows to light: Uncovering money laundering in Ethereum transaction graphs. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5076472
16. Zhao, J., & Ma, Y. (2024). A review on the security of the Ethereum-based DeFi ecosystem. Computer Modeling in Engineering & Sciences, 139(1), 49-72.
17. Zhang, Y., Liu, F., & Zhou, F. (2019). SoK: Transparent dishonesty: Front-running attacks on blockchain. arXiv preprint arXiv:1902.05164.
18. Li, Z., Liu, C., He, Z., & Xu, J. (2022). SoK: MEV countermeasures: Theory and practice. arXiv preprint arXiv:2212.05111.
19. Chen, Y., & Bellavitis, C. (2023). Mapping the DeFi crime landscape: An evidence-based picture. Available at SSRN.

20. Amin, M., Shah, S. A. A., Iqbal, S., & Kaviani, A. K. (2024). A novel classification of attacks on blockchain layers: Vulnerabilities, attacks, mitigations, and research directions. arXiv preprint arXiv:2404.18090.
21. Meyer, E. A., Welpe, I. M., & Sandner, P. (2022). Decentralized Finance – A Systematic Literature Review and Research Directions. ECIS 2022 Research Papers, Paper 25.
22. Wronka, C. (2021). Financial crime in the decentralized finance ecosystem: new challenges for compliance. Journal of Financial Crime, 30(1), 97–113. https://doi.org/10.1108/JFC-09-2021-0218.
23. Amin, M., Shah, S. A. A., Iqbal, S., Alshamrani, S. S., & Alshehri, M. D. (2023). SoK: Design, vulnerabilities, and security measures of cryptocurrency wallets. arXiv preprint arXiv:2307.12874.
24. Patrick McCorry, S. F., & Robinson, P. (2019). Flash boys 2.0: Frontrunning, transaction reordering, and consensus instability in decentralized exchanges. arXiv preprint arXiv:1904.09334.
25. Zhou, W., Tan, Y., & Jia, Y. (2024). SoK: Not quite water under the bridge: Review of cross-chain bridge hacks. arXiv preprint (2024).
26. Li, X., & Wang, H. (2024). SoK: A review of cross-chain bridge hacks in 2023. arXiv preprint.
27. Zhou, E., et al. (2022). SoK: Decentralized exchanges (DEX) with automated market maker (AMM) protocols. arXiv preprint.
28. Qin, K., et al. (2020). Towards a first step to understand flash loan and its applications in DeFi ecosystem. arXiv preprint arXiv:2010.06146.
29. Liao, Z., Nan, Y., Liang, H., Hao, S., Zhai, J., Wu, J., & Zheng, Z. (2024). SmartAxe: Detecting Cross-Chain Vulnerabilities in Bridge Smart Contracts via Fine-Grained Static Analysis. Proceedings of the ACM on Software Engineering, 1(FSE), Article 12.
30. Chaliasos, S., Wang, Z., Wang, Y., Qin, K., & Gervais, A. (2023). Smart Contract and DeFi Security Tools: Do They Meet the Needs of Practitioners? arXiv:2304.02981.
31. Duley, C., Gambacorta, L., Garratt, R., & Koo Wilkens, P. (2023). The Oracle Problem and the Future of DeFi. BIS Bulletin No. 76.
32. Daian, P., Goldfeder, S., Kell, T., Li, Y., Zhao, X., Bentov, I., … Juels, A. (2020). Flash Boys 2.0: Frontrunning in Decentralized Exchanges, Miner Extractable Value, and Consensus Instability. In 2020 IEEE Symposium on Security and Privacy (SP) (pp. 910–927).
33. Sun, D., Ma, W., Nie, L., & Liu, Y. (2024). SoK: Comprehensive Analysis of Rug Pull Causes, Datasets, and Detection Tools in DeFi. arXiv:2403.16082.
34. Dotan, M., Yaish, A., Yin, H.-C., Tsytkin, E., & Zohar, A. (2023). The Vulnerable Nature of Decentralized Governance in DeFi. arXiv:2308.04267.
35. Notland, J. S., Li, J., Nowostawski, M., & Haro, P. H. H. (2025). SoK: Cross-chain Bridging Architectural Design Flaws and Mitigations. Blockchain: Research and Applications, 7(1), 100315.
36. Chu, H., Zhou, Y., Yan, X., Sun, L., et al. (2023). A Survey on Smart Contract Vulnerabilities: Data Sources, Vulnerability Detection, and Vulnerability Defense. Information and Software Technology, 107221.
37. Wu, S., Yu, Z., Wang, D., Zhou, Y., Wu, L., Wang, H., & Yuan, X. (2023). DeFiRanger: Detecting DeFi Price Manipulation Attacks. IEEE Transactions on Dependable and Secure Computing.
38. Tessa Van Roomen (2025). Cross-Border Evidence Gathering in Criminal Crypto Investigations: A Case Study of the Netherlands. Tilburg Law Review.
39. Augusto, A., Belchior, R., Pfannschmidt, J., Vasconcelos, A., & Correia, M. (2024). XChainWatcher: Monitoring and Identifying Attacks in Cross-Chain Bridges. arXiv:2410.02029.

40. Chainalysis. (2025). Blockchain Security: Preventing Threats Before They Strike.
41. Nadler, M., & Schär, F. (2023). Tornado Cash and Blockchain Privacy: A Primer for Economists and Policymakers. SSRN Electronic Journal.
42. Financial Action Task Force. (2021). Updated Guidance for a Risk-Based Approach to Virtual Assets and Virtual Asset Service Providers. FATF.
43. Financial Action Task Force. (2025). Virtual Assets: Targeted Update on Implementation of the FATF Standards on Virtual Assets and VASPs. FATF.
44. TRM Labs. (2025). The Fundamentals of Cryptocurrency Transaction Tracing.
45. Brownworth, A., Ferreira Torres, C., & Wattenhofer, R. (2024). Evidence from Sanctions on Tornado Cash. Federal Reserve Bank of New York Staff Reports, No. 1112.

Enhancing Network Anomaly Detection via Hierarchically Supervised Autoencoder

Duc Viet Bui, Thi Huong Chu [1]
Institute of Information and Communication Technology,
Le Quy Don Technical University
viet.bui@lqdtu.edu.vn

Abstract

Network anomaly detection plays a fundamental role in identifying zero-day attacks and previously unseen intrusion behaviors in modern networked systems. While deep autoencoder-based models have been widely adopted for unsupervised anomaly detection, conventional architectures often learn latent representations that lack sufficient structural compactness, thereby limiting the separability between normal and abnormal traffic patterns. In this paper, we propose a hierarchically supervised autoencoder to improve the quality and discriminative power of latent representations. Unlike traditional single bottleneck autoencoders, the proposed model introduces intermediate reconstruction heads that impose multi-level supervision across encoder layers. This hierarchical design encourages each representation level to capture meaningful and progressively refined abstractions of normal network behavior. Experiments on benchmark intrusion detection datasets demonstrate that the proposed method consistently improves AUC performance compared to conventional autoencoder-based baselines. The results indicate that hierarchical supervision significantly strengthens latent feature quality for anomaly-based intrusion detection.

Keyword: *Deep Learning, Anomaly Detection, Autoencoder, Latent Representation, Hierarchically Feature.*

1. Introduction

The rapid expansion of modern computer networks, cloud infrastructures, IoT ecosystems, and software-defined systems has significantly increased the attack surface of digital environments. In such contexts, network intrusion detection systems (NIDS) play a critical role in protecting digital infrastructure against cyber-attacks. Traditionally, NIDS are categorized into signature-based intrusion detection systems (SIDS) and anomaly-based intrusion detection systems (AIDS) [1, 2]. While SIDS can effectively detect known attack patterns with high precision, they fundamentally suffer from their inability to identify zero-day attacks, polymorphic malware, and evolving intrusion behaviors [1, 3]. Moreover, maintaining and updating large signature databases become increasingly impractical in dynamic network environments [2]. In contrast, anomaly-based intrusion detection systems aim to learn the normal behavioral patterns of network traffic and detect deviations from them [3, 4]. Since novel attacks often exhibit statistical behaviors different from legitimate traffic, anomaly detection has emerged as a promising direction for identifying previously unseen threats [3, 5]. However, network anomaly detection remains challenging because network traffic is typically high-dimensional and heterogeneous, often contains complex non-linear relationships across feature groups, and usually lacks sufficient labeled anomaly data for supervised learning [4 - 6].

To address these challenges, the research community has extensively investigated deep representation learning models for anomaly detection [4 – 6]. In particular, Nicolau et al. [7] explored neural representation learning for network anomaly detection using autoencoder-based architectures, while Cao et al. [8] proposed a hybrid framework that combines autoencoder reconstruction with density estimation for anomaly detection. These studies follow the common

assumption that models trained on normal traffic can reconstruct normal samples more accurately than anomalous ones, making reconstruction error a practical anomaly score. More broadly, autoencoders (AEs) and their variants have been widely recognized as effective unsupervised tools for learning compact latent representations of normal data, and probabilistic latent-variable approaches have also been introduced [6]. For example, Nguyen et al. [9] proposed variational deep clustering approaches for anomaly-based cyber-attack detection, integrating variational representation learning with clustering mechanisms to improve the separation between normal and abnormal samples in latent space. In a related direction, Nguyen et al. [10] proposed a deep clustering framework with hierarchical latent representations, where latent-space organization is explicitly encouraged across representation levels to enhance anomaly-based cyber-attack detection. These studies show that structuring the latent space - through reconstruction, probabilistic modeling, density estimation, or clustering - has become a central theme in recent anomaly-based intrusion detection research.

Despite these advances, several important limitations remain. First, most AE-based anomaly detection models are supervised only through the final decoder reconstruction, while intermediate encoder representations are left weakly constrained. As a result, early hidden layers may learn features that are not sufficiently informative for anomaly discrimination. Second, many clustering-integrated methods impose compactness only at a single latent level, typically the bottleneck layer, without enforcing hierarchical consistency across multiple representation depths. This limits their ability to capture the multi-scale structural characteristics of network traffic. Third, probabilistic generative models such as VAEs often face a trade-off between reconstruction fidelity and latent regularization strength: excessive regularization may weaken reconstruction quality, whereas insufficient regularization may lead to poorly organized latent distributions. In practice, network anomalies may appear at different granularities, ranging from subtle deviations in low-level statistical attributes to more abstract inconsistencies in aggregated flow behavior [3, 4]. Therefore, enforcing discriminative constraints across multiple hierarchical levels of representation learning is more consistent with the multi-scale nature of real network traffic.

Motivated by these observations, this paper proposes a Multi-Scale Shrink Autoencoder (MS-SAE) for network anomaly detection. Unlike conventional autoencoder-based approaches that mainly constrain the bottleneck representation through final-output reconstruction, the proposed model introduces auxiliary reconstruction heads at multiple encoder depths to explicitly supervise intermediate hidden representations. This design encourages the preservation of informative structural characteristics of normal traffic across the hierarchical encoding process rather than only at a single latent level. Furthermore, by integrating multi-scale reconstruction supervision with shrink-based latent compactness regularization, the proposed method seeks to learn more compact and discriminative representations of normal network behavior for anomaly detection. Experimental results on benchmark intrusion detection datasets demonstrate the effectiveness of the proposed architecture in improving latent representation quality and delivering competitive detection performance.

The remainder of this paper is organized as follows. Section 2 presents the proposed method and the experimental settings, including the formulation of the deep autoencoder, shrink

autoencoder, and the proposed MS-SAE model, as well as the benchmark datasets and implementation details. Section 3 reports the experimental results and provides discussion and analysis of the learned representations and detection performance. Finally, Section 4 concludes the paper and outlines several directions for future work.

2. Proposed Method and Experiment Settings

2.1 Proposed Method

2.1.1. Deep Autoencoder and Shrink Autoencoder

Let $x \in \mathbb{R}^d$ denotes an input network traffic sample. A deep autoencoder (AE) consists of an encoder–decoder architecture. The encoder is a nonlinear mapping $z = f_\phi(x)$, where ϕdenotes the set of trainable encoder parameters (i.e., all weight matrices and bias vectors). The mapping is implemented through stacked affine transformations:

$$h^{(l)} = \sigma\left(W^{(l)}h^{(l-1)}+b^{(l)}\right), l = 1, \dots, L,$$

with $h^{(0)} = x$, and $z = h^{(L)} \in \mathbb{R}^{d_h}$, where $d_h < d$. Here, $W^{(l)}$and $b^{(l)}$denote the weight matrix and bias vector of layer l, and $\sigma(\cdot)$ is a nonlinear activation function. The decoder reconstructs the input into

$$\hat{x} = g_\theta(z),$$

where θ represents the trainable parameters of the decoder network. The model is oftent trained using only normal samples by minimizing the reconstruction loss:

$$\mathcal{L}_{AE} = \frac{1}{N}\sum_{i=1}^{N} \| x_i - \hat{x}_i \|_2^2 .$$

During inference, anomalies are identified through the reconstruction error:

$$s(x) = \| x - \hat{x} \|_2^2,$$

where larger values indicate higher likelihood of anomalous behavior.

To address the dispersion of latent representations, prior work introduced the Shrink Autoencoder (SAE), which augments the reconstruction objective with a latent compactness regularizer. Let $z_i = f(x_i)$ denote the latent representation of input x_i. The shrink regularization term is defined as

$$\mathcal{L}_{shrink} = \frac{1}{N}\sum_{i=1}^{N} \| z_i \|_2^2.$$

The SAE training loss becomes

$$\mathcal{L}_{SAE} = \frac{1}{N}\sum_{i=1}^{N} \| x_i - \hat{x}_i \|_2^2 + \lambda\frac{1}{N}\sum_{i=1}^{N} \| z_i \|_2^2,$$

where $\lambda > 0$ controls the trade-off between reconstruction fidelity and latent compactness. By penalizing the magnitude of latent vectors, SAE encourages normal samples to cluster near the origin in latent space, thereby forming a tighter normal region. However, these single-scale reconstruction constraints supervise only the final output layer. We assume that intermediate hidden representations remain weakly regulated, allowing the encoder to distribute normal

samples arbitrarily across latent subspaces. Consequently, the normal manifold may become dispersed, reducing separability from anomalous instances.

2.1.2. Proposed Multi-Scale Shrink Autoencoder

In this work, we propose a method builds upon conventional deep autoencoder frameworks by introducing hierarchical reconstruction constraints across multiple encoding depths. The motivation stems from the observation that standard reconstruction-based models supervise only the final decoder output, leaving intermediate representations weakly constrained. This may result in dispersed latent manifolds for normal data, thereby reducing anomaly separability. Therefore, unlike the conventional deep autoencoder, which reconstructs the input solely from the bottleneck representation through the decoder, the proposed method attaches auxiliary reconstruction heads directly to intermediate encoder layers. These heads operate entirely within the encoder pathway, enforcing structural consistency across encoding depths. By enforcing reconstruction consistency at multiple abstraction levels while simultaneously encouraging latent compactness, the proposed model is designed to enhance structural coherence and discriminative power in the learned feature space. Figure 1 describes the details of our proposed model.

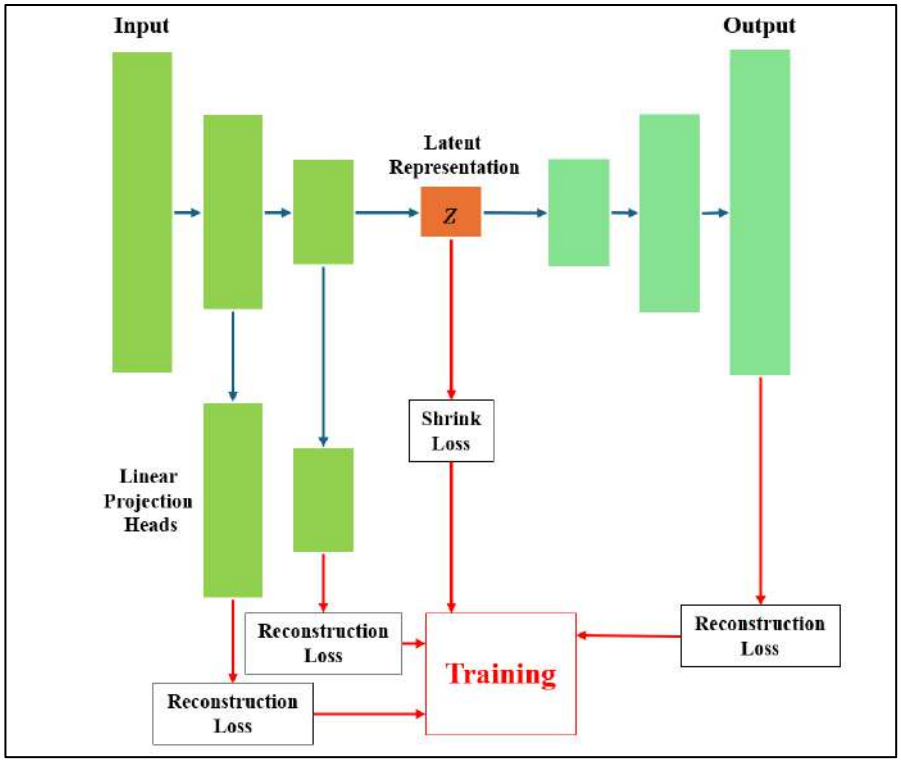


Figure 1. Multi-Scale Shrink Autoencoder (MS-SAE)

Let $x \in \mathbb{R}^d$ denote the input sample and $h^{(l)} \in \mathbb{R}^{d_l}$, $l = 1, \ldots, L$, denote the representations of encoder layers, where $h^{(L)} = z \in \mathbb{R}^{d_h}$ is the latent representation. Since each encoder layer has dimensionality d_l that differs from the original input dimension d, a linear projection head is introduced to map each representation back to the input space. Specifically, for each encoding depth l, the output from each reconstruction head is defined as

$$\hat{x}^{(l)} = W_r^{(l)} h^{(l)} + b_r^{(l)},$$

where $W_r^{(l)} \in \mathbb{R}^{d \times d_l}$ and $b_r^{(l)} \in \mathbb{R}^d$. The weight matrix $W_r^{(l)}$ performs a linear transformation from the d_l-dimensional representation of the encoder layer to the original d-dimensional input space, ensuring that each auxiliary head produces a full input reconstruction $\hat{x}^{(l)} \in \mathbb{R}^d$. This dimensional mapping guarantees consistency across all reconstruction scales. This design produces multiple encoder-driven reconstructions

$$\{\hat{x}^{(1)}, \hat{x}^{(2)}, \dots, \hat{x}^{(L)}\},$$

each derived from progressively abstract feature representations. By enforcing reconstruction consistency at every encoding depth, we assume that the model could constrain each intermediate representation to preserve sufficient information for reconstructing the original input, thereby stabilizing the encoding hierarchy. The multi-scale reconstruction loss is defined as follows:

$$\mathcal{L}_{MS} = \sum_{l=1}^{L} \alpha_l \frac{1}{N} \sum_{i=1}^{N} \| x_i - \hat{x}_i^{(l)} \|_2^2,$$

where $\alpha_l \geq 0$ are scale-dependent weighting coefficients. By attaching reconstruction heads directly to encoder layers, the proposed method introduces explicit multi-depth regularization on the encoding function. This modifies the gradient flow such that each layer receives independent reconstruction supervision, reducing representational variance and stabilizing the geometry of the learned normal manifold. Consequently, feature collapse is mitigated, and the hierarchical structure of the encoding process becomes more robust. Integrating the standard decoder reconstruction loss, the auxiliary multi-scale reconstruction terms, and the latent shrink regularizer, the final objective loss function of the proposed model is defined as follows:

$$\mathcal{L}_{recon} = \frac{1}{N} \sum_{i=1}^{N} \| x_i - \hat{x}_i \|_2^2$$

$$\boxed{\mathcal{L}_{final} = \mathcal{L}_{recon} + \mathcal{L}_{MS} + \lambda \frac{1}{N} \sum_{i=1}^{N} \| z_i \|_2^2}$$

where $\mathcal{L}_{recon}$ denotes denotes the standard decoder reconstruction loss. The parameter $\lambda > 0$ controls the strength of latent compactness regularization.

During the testing phase, for the reconstruction error-based anomaly detection, rather than relying solely on the final decoder output, the proposed model evaluates reconstruction consistency at every encoder depth, thereby incorporating multi-level structural evidence. For a test sample x, the model computes:

- The primary reconstruction error from the decoder output:

$$e^{(0)}(x) = \| x - \hat{x} \|_2^2,$$

- The auxiliary reconstruction errors from each encoder-level head:

$$e^{(l)}(x) = \| x - \hat{x}^{(l)} \|_2^2, l = 1, \dots, N,$$

where $\hat{x}$ denotes the final decoder reconstruction and $\hat{x}^{(l)}$ denotes the reconstruction produced from the l-th encoder layer. The anomaly score is defined as the average reconstruction error across all scales:

$$s(x) = \frac{\| x - \hat{x} \|_2^2 + \sum_{l=1}^{N} \| x - \hat{x}^{(l)} \|_2^2}{1 + N}$$

This formulation can be considered as a multi-layer consensus mechanism. An input sample is considered anomalous if it cannot be consistently reconstructed across different levels of abstraction. While normal samples are expected to produce small reconstruction errors at all scales, anomalous samples typically induce large discrepancies not only at the bottleneck but also within intermediate representations.

2.2. Experiment Settings

2.2.1. Datasets

To comprehensively evaluate the effectiveness of the proposed model for anomaly-based network intrusion detection, we conduct experiments on two widely used benchmark datasets: NSL-KDD and UNSW-NB15. These datasets are selected due to their diversity in attack types, feature characteristics, and their widespread adoption in anomaly detection literature. The details about the datasets can be found in Table 1.

- NSL-KDD: NSL-KDD is an improved version of the famous KDD'99 dataset, designed to remove redundant records and reduce class imbalance bias. Each network connection record consists of 41 handcrafted features extracted from TCP/IP traffic, including basic features, content-based features, and traffic statistical features. The dataset contains one normal class and four major attack categories: Denial of Service (DoS), Probe, Remote to Local (R2L) and User to Root (U2R).
- UNSW-NB15: UNSW-NB15 contains 49 raw features derived from packet-level and flow-level characteristics, including statistical and content features. The dataset includes one normal class and nine attack categories: Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, Worms.

Table 1. Details of benchmark datasets

Dataset	Dimension	Normal Training	Normal Testing	Anomaly Testing
NSL-KDD	122	67343	9711	12833
UNSW-NB15	196	55999	33999	45332

2.2.2. Parameter Settings

The models are implemented using a six-layer hidden architecture, and the dimensionality of the bottleneck layer m is determined according to the method in [10], defined as $m = \lfloor 1 + \sqrt{n} \rfloor$, where n denotes the number of input features. The number of training epochs is 500. A mini-batch size of 100 is used and the Adadelta optimization algorithm is deployed. Early stopping was also implemented to stop the training when the loss function showed no improvement of at least 0.001 over five consecutive epochs. The hyperbolic tangent (tanh) activation function is adopted for all hidden layers, and network weights are initialized using the

scheme described in the referenced study. The hyperparameter α_l and λ are set to 0.5 and 10 respectively.

Following the one-class anomaly detection methods adopted in previous works, we only use normal samples from the training set to train the model. Categorical attributes (protocol type, service, flag) are transformed using one-hot encoding. After preprocessing, the feature dimensionality increases due to encoding expansion. During testing phase, both normal and attack samples are included, and all attack types are treated as anomalies. To evaluate the quality of the learned latent representations, we employ six widely used one-class classification algorithms, including One-Class Support Vector Machine (OCSVM), Isolation Forest (IF), Local Outlier Factor (LOF), Kernel Density Estimation (KDE), Centroid-based detection (CEN), and a reconstruction error-based method (RE). These classifiers operate on the learned latent space to evaluate its discriminative capability in separating normal and anomalous network traffic. To demonstrate the effectiveness of the proposed model, we compare its performance against representative baseline approaches. Specifically, we include baseline deep autoencoder–based frameworks combined with one-class classifiers such as SAE+OCCs and DCAE+OCCs, to provide a fair and comprehensive benchmark.

3. Results and Discussion

Table 2 summarizes the AUC performance of DAE, VAE, SAE, DVAE, and the proposed MS-SAE-based representations combined with several classifiers (RE, OCSVM, CEN, LOF, KDE, IF) on NSL-KDD and UNSW-NB15.

Table 2. AUCs of DAE+OCCs, VAE+OCCs, SAE+OCCs, DVAE+OCCs, MS-SAE+OCCs on NSL-KDD and UNSW-NB15. The best results are hightlighted in bold.

Latent Representation	Datasets	One-class Classifiers					
		RE	**OCSVM**	**CEN**	**LOF**	**KDE**	**IF**
DAE	NSL-KDD	0.923	0.932	0.943	0.911	0.935	0.946
	UNSW-NB15	0.804	0.724	0.730	0.717	0.702	0.806
VAE	NSL-KDD	0.920	0.941	0.943	0.948	0.938	0.942
	UNSW-NB15	0.804	0.820	0.813	0.837	0.809	0.822
SAE	NSL-KDD	0.923	0.959	0.960	0.952	0.959	0.959
	UNSW-NB15	0.812	0.876	0.875	0.878	0.868	0.874
DVAE	NSL-KDD	0.922	0.942	0.941	0.939	0.938	0.937
	UNSW-NB15	0.805	0.821	0.823	0.822	0.830	0.820
MS-SAE	NSL-KDD	**0.945**	**0.962**	**0.964**	**0.959**	**0.963**	**0.961**
	UNSW-NB15	0.799	**0.887**	**0.881**	**0.884**	**0.877**	**0.878**

On NSL-KDD, MS-SAE achieves the highest AUC across all classifiers, reaching 0.945 (RE), 0.962 (OCSVM), 0.964 (CEN), 0.959 (LOF), 0.963 (KDE), and 0.961 (IF). Compared to SAE and other baselines, the improvements are not limited to a specific classifier family but are uniformly observed across boundary-based (OCSVM, IF), density-based (LOF, KDE), and centroid-based (CEN) approaches. This indicates that the hierarchical multi-scale reconstruction constraint enhances the intrinsic structure of the latent space. In particular, the improvement over SAE suggests that enforcing intermediate reconstruction heads prevents over-specialization of the bottleneck layer and mitigates feature co-adaptation. As a result, normal samples are

embedded into a more compact and geometrically regular manifold, increasing the separability between normal and anomalous instances in latent space.

UNSW-NB15 presents a more challenging scenario due to its higher feature heterogeneity, modern attack diversity, and more complex traffic distributions. This is reflected in the generally lower AUC values across all models. However, MS-SAE still achieves the best overall performance across most classifiers: 0.887 (OCSVM), 0.881 (CEN), 0.884 (LOF), 0.877 (KDE), and 0.878 (IF). The consistent gains over SAE confirm that the multi-scale constraint remains effective even under heterogeneous and non-linear feature interactions. Interestingly, reconstruction-error-based detection slightly decreases for MS-SAE (0.799) compared to SAE (0.812). This phenomenon can be explained by the stronger regularization imposed at intermediate layers: while this promotes smoother latent manifolds and benefits downstream OCC classifiers, it may slightly reduce the pure reconstruction fidelity of the final decoder.

A notable observation across both datasets is that MS-SAE demonstrates classifier-agnostic robustness. The performance gains are consistent across distance-based, density-based, and boundary-based OCCs, suggesting that the improvements primarily stem from enhanced representation learning rather than classifier-specific optimization. Since OCC methods are highly sensitive to the compactness, symmetry, and smoothness of the latent distribution, the multi-scale reconstruction mechanism appears to implicitly regularize intermediate representations. By supervising multiple abstraction levels simultaneously, the model reduces redundancy propagation and encourages hierarchical consistency. This leads to a latent space with reduced variance among normal samples and improved margin from anomalous points.

4. Conclusion

This paper presented MS-SAE, a hierarchically supervised autoencoder designed to enhance latent representation learning for network anomaly detection. Unlike conventional autoencoders that rely solely on bottleneck reconstruction, MS-SAE introduces hierarchical intermediate reconstruction supervision, enforcing consistency across multiple abstraction levels of the encoder. This multi-scale regularization encourages the network to learn structurally coherent and discriminative latent representations, thereby improving anomaly separability in high-dimensional network traffic data. Extensive experiments on NSL-KDD and UNSW-NB15 demonstrate that MS-SAE consistently outperforms or remains competitive with strong baselines. The results validate that incorporating hierarchical reconstruction constraints leads to a more compact normal manifold and enhanced inter-class separation. Importantly, the improvements generalize across different one-class classifiers, suggesting that MS-SAE enhances the intrinsic geometry of the latent space rather than optimizing for a specific scoring method. Future work will focus on extending the multi-scale framework toward adaptive or dataset-aware weighting strategies, integrating clustering-based constraints to further improve anomaly compactness, and evaluating the approach on more recent large-scale real-world network intrusion datasets.

References

1. Liao, Hung-Jen, et al. "Intrusion detection system: A comprehensive review." *Journal of network and computer applications* 36.1 (2013): 16-24.

2. Abdulganiyu, Oluwadamilare Harazeem, Taha Ait Tchakoucht, and Yakub Kayode Saheed. "A systematic literature review for network intrusion detection system (IDS)." *International journal of information security* 22.5 (2023): 1125-1162.
3. Fernandes Jr, Gilberto, et al. "A comprehensive survey on network anomaly detection." *Telecommunication Systems* 70.3 (2019): 447-489.
4. Wang, S., Balarezo, J. F., Kandeepan, S., Al-Hourani, A., Chavez, K. G., & Rubinstein, B. (2021). Machine learning in network anomaly detection: A survey. *IEEe Access*, *9*, 152379-152396.
5. Morshedi, R., & Matinkhah, S. M. (2025). A comprehensive review of deep learning techniques for anomaly detection in iot networks: Methods, challenges, and datasets. *Engineering Reports*, *7*(9), e70415.
6. Singh, Richa, Nidhi Srivastava, and Ashwani Kumar. "Network anomaly detection using autoencoder on various datasets: A comprehensive review." *Recent Patents on Engineering* 18.9 (2024): 63-77.
7. Nicolau, M., & McDermott, J. (2018). Learning neural representations for network anomaly detection. *IEEE transactions on cybernetics*, *49*(8), 3074-3087.
8. Cao, V. L., Nicolau, M., & McDermott, J. (2016, August). A hybrid autoencoder and density estimation model for anomaly detection. In *International conference on parallel problem solving from nature* (pp. 717-726). Cham: Springer International Publishing.
9. Nguyen, Van Quan, et al. "Variational Deep Clustering approaches for anomaly-based cyber-attack detection." *Journal of Network and Computer Applications* 240 (2025): 104182.
10. Nguyen, Van Quan, et al. "Deep clustering hierarchical latent representation for anomaly-based cyber-attack detection." *Knowledge-Based Systems* 301 (2024): 112366.

Tăng cường phát hiện bất thường mạng thông qua bộ tự mã hóa có giám sát phân cấp

Tóm tắt

Phát hiện bất thường mạng đóng vai trò nền tảng trong việc nhận diện các cuộc tấn công zero-day và các hành vi xâm nhập chưa từng xuất hiện trong các hệ thống mạng hiện đại. Mặc dù các mô hình dựa trên các bộ tự mã hóa đã được ứng dụng rộng rãi cho phát hiện bất thường theo hướng học không giám sát, các kiến trúc truyền thống thường học được các biểu diễn tiềm ẩn thiếu tính cô đọng cấu trúc, từ đó hạn chế khả năng phân tách giữa lưu lượng bình thường và bất thường. Trong bài báo này, chúng tôi đề xuất một bộ tự mã hóa có giám sát phân cấp nhằm nâng cao chất lượng và tính phân biệt của biểu diễn tiềm ẩn. Khác với các bộ tự mã hóa truyền thống chỉ có một nhánh trích xuất đặc trưng duy nhất, mô hình đề xuất bổ sung các nhánh tái tạo trung gian, tạo ra cơ chế giám sát đa mức xuyên suốt các tầng mã hóa. Thiết kế phân cấp này khuyến khích mỗi mức biểu diễn học được các đặc trưng có ý nghĩa và được tinh lọc dần theo mức trừu tượng hóa của hành vi mạng bình thường. Thực nghiệm trên các bộ dữ liệu chuẩn về phát hiện xâm nhập mạng cho thấy phương pháp đề xuất cải thiện ổn định chỉ số AUC so với các mô hình tự mã hóa truyền thống. Kết quả thực nghiệm chứng minh rằng cơ chế giám sát phân cấp giúp tăng cường đáng kể chất lượng đặc trưng tiềm ẩn cho phát hiện xâm nhập dựa trên bất thường.

Từ khóa: Học sâu, Phát hiện bất thường, Bộ tự mã hóa, Biểu diễn ẩn, Đặc trưng phân cấp.

Parameter-Aware Log-Based Anomaly Detection

Manh Tuan Nguyen[1] *

[1] *Institute of Information and Communication Technology*

Tuannm_ncs42@lqdtu.edu.vn

Abstract

Conventional log anomaly detection methods often mask variable parameters (e.g., latencies, IPs, hashes) with generic tokens, losing critical contextual signals for subtle anomalies. This paper introduces a parameter-aware log representation framework featuring a Hybrid Parameter Embedding module. This module parses parameters into (key, value, type) tuples and applies type-conditioned embeddings, including log-scaled numerics and type-only tokens for high-cardinality identifiers, to prevent overfitting while preserving semantics. Furthermore, a Gated Key Injection mechanism is proposed to dynamically fuse parameter context into Transformer self-attention layers. Experimental results on HDFS, Thunderbird and BGL datasets demonstrate that our approach outperforms standard Transformer baselines, achieving an average improvement of 3.6% in F1-score. The results highlight the framework's superior ability to detect parameter-sensitive anomalies and its enhanced interpretability for real-world cyber security applications.

Keywords: *Log anomaly detection; parameter-aware; cyber security*

1. Introduction

The management of modern large-scale software systems relies heavily on system logs to monitor health and diagnose failures. Despite being different in morphology and syntax, logs from multiple domains (e.g., supercomputers, cloud applications, distributed networks) usually share a similar underlying semantic space. However, existing anomaly detection methods mostly focus on a single domain, assuming that training and testing data follow the same distribution.

In large-scale distributed systems, cloud infrastructures, and enterprise applications, system logs serve as a primary source of runtime telemetry, recording detailed traces of events, timestamps, execution durations, and associated parameters. These logs are invaluable for postmortem analysis, troubleshooting, and proactive anomaly detection, as they often reveal early indicators of performance degradation, configuration errors, security breaches, or hardware faults [1] [2]. However, the sheer volume, unstructured nature, and noise inherent in logs pose significant challenges for automated analysis.

Conventional log-based anomaly detection pipelines typically follow a four-stage process: (1) log parsing to extract structured templates and events from raw text using methods such as Spell [3] or Drain [4]; (2) grouping of events into temporal or session-based sequences; (3) feature extraction, either via simple frequency counting for machine learning models [13] or semantic embeddings (e.g., Word2Vec, FastText, GloVe) for deep learning approaches; and (4) application of classification or reconstruction-based models to flag deviations as anomalies. While effective for many scenarios, these methods largely discard or anonymize the variable parameters within log lines (such as numerical latencies, IP addresses, user IDs, or hash values) by replacing them with generic placeholders (e.g., <*>). This masking strategy, though necessary to handle high-

* Email: corresponding_author@lqdtu.edu.vn

cardinality variability and reduce vocabulary explosion, often eliminates crucial contextual signals. For instance, unusually high latency values, suspicious IP patterns, or invalid path structures frequently correlate directly with anomalous behavior, yet they remain invisible to downstream models that operate solely on template semantics.

Recent advances in deep learning have improved semantic understanding of log templates through pre-trained encoders, but they still treat parameters as noise rather than informative features. This limitation is particularly pronounced in production environments where anomalies are subtle, context-dependent, and not always reflected in event sequences alone.

To bridge this gap, we propose a parameter-aware representation learning framework for log anomaly detection. Our key contributions are threefold:

- A hybrid parameter embedding module that parses raw parameters into structured (key, value, type) tuples using regular expressions, then fuses categorical, numerical, and type-level embeddings in a type-conditioned manner to preserve semantics while preventing memorization of volatile identifiers.
- A Gated Key Injection mechanism that dynamically modulates the key matrix in self-attention layers by injecting gated parameter context, allowing the model to selectively emphasize or suppress parameter influence based on relevance.
- An end-to-end trainable architecture that augments standard transformer-based sequence encoding with these enhancements, using a [CLS] token to derive compact, context-rich representations for anomaly detection.

Our approach maintains compatibility with existing parsing pipelines while substantially enriching log representations. By explicitly modeling parameter semantics and their interactions with event templates, we achieve superior detection performance on parameter-sensitive anomalies, better generalization to unseen logs, and enhanced interpretability through gated mechanisms.

2. Related Work

Early deep log anomaly detectors primarily focus on learning a representation of normal log-event sequences within a single system. Transformer-based approaches such as LogBert [5] and LogEncoder [6] leverage masked-token objectives to capture contextual semantics and often employ one-class or hypersphere-style decision functions to separate normal from abnormal behaviors. HilBERT [7] further improves the quality of representations via hierarchical Transformer pre-training tailored to log-event sequences.

Beyond Transformers, alternative sequence models have also been investigated. BTCNLog [8] applies a bidirectional temporal convolution network in a semi-supervised setting. To better capture hierarchical semantics, LayerLog [9] proposes multi-level semantic modeling over log templates. Recent surveys summarize the landscape of deep log anomaly detection and highlight open challenges such as generalization, robustness, and label scarcity ([10], [11]).

A complementary research line targets parameter-aware modeling, motivated by the fact that anomalous conditions are sometimes indicated mainly by numerical values (e.g., latency spikes or unusual error codes). HitAnomaly [12] utilizes hierarchical Transformers to integrate

templates and parameters but often struggles with “vocabulary explosion” when encountering high-cardinality identifiers like hashes or file paths. Similarly, LogSER [13] parses templates and parameters via BERT-style objectives but remains susceptible to overfitting on volatile training data. A focused survey on parameter anomalies in logs is provided in [14].

3. Method

3.1. Overall Framework

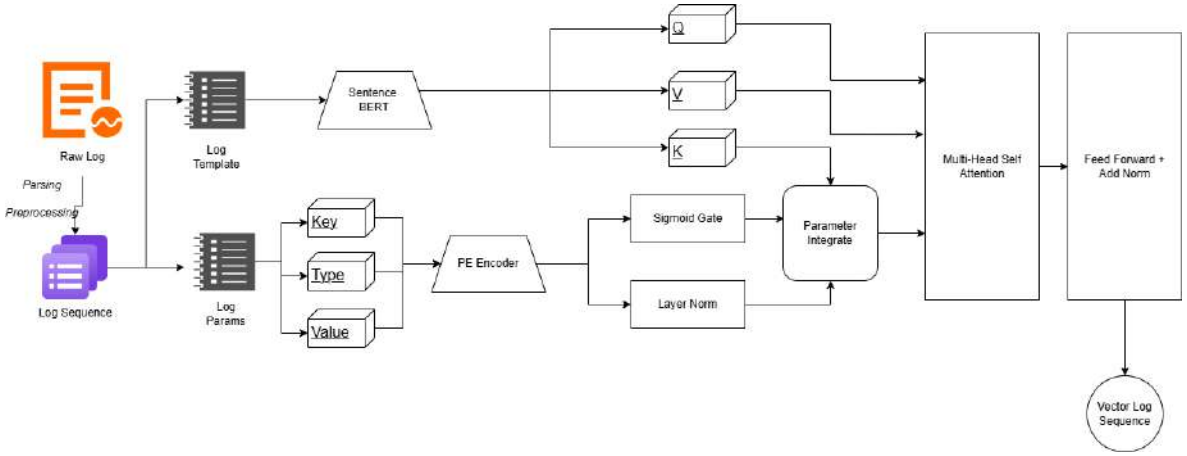


Figure 1. Parameter-Aware for Log Representation

Log systems typically capture records containing timestamps, events, labels, and execution durations. Conventional log-based anomaly detection follows a standard pipeline that begins with log parsing, which converts unstructured raw logs into structured templates and event instances using established methods such as Spell [3] and Drain [4]. The parsed events are then grouped into semantically meaningful sequences (typically through timestamp-based windows or session/identifier-based partitioning [15]), since anomalies usually emerge across related log lines rather than in isolated entries. These sequences are subsequently transformed into numerical feature representations suitable for machine learning (ML) or deep learning (DL) models: traditional ML approaches commonly rely on event counting or frequency vectors [15], while DL-based methods leverage semantic embeddings (e.g., Word2Vec, FastText [16] }, or GloVe) to capture contextual relationships. Finally, the extracted features are fed into various ML and DL models to perform anomaly detection.

3.2. Parameter-aware for log representation

Standard log parsing masks parameters (e.g., latencies, IP addresses, hashes) into generic tokens, often discarding critical operational context. To address this, we propose a structured, hybrid parameter embedding strategy followed by a Gated Key Injection mechanism to fuse this numerical and categorical context into the log's semantic representation.

Hybrid Parameter Embedding

Instead of relying purely on numerical extraction, we utilize Regular Expressions (Regex) to automatically parse raw log parameters into a tuple of (k, v, τ), representing the Key, Value, and Type, respectively.

The type τ is inferred via predefined Regex patterns mapping to categories such as NUM (numeric), IP (IPv4/IPv6), HEX (hashes/IDs), PATH, and STR (generic categorical strings). If a

parameter is explicitly formatted as a key-value pair (e.g., user=root), the explicit key is extracted as k. Otherwise, a positional key (e.g., POS_1, POS_2) is assigned based on its sequential appearance in the log line.

The composite parameter embedding E_{param} is formulated as the sum of three distinct embeddings:

$$\mathrm{E_{param}} = \mathrm{E_{key}} + \mathrm{E_{type}} + \mathrm{E_{value}}$$

Here, $E_{key} = \boldsymbol{W}_{key}k$ and $E_{type} = \boldsymbol{W_{type}}\tau$ are obtained via learnable embedding lookup tables. To handle the high variability of parameter values while preventing Out-Of-Vocabulary (OOV) memory explosion, E_{value} is conditionally computed based on the inferred type τ:

$$E_{value} = \boldsymbol{W_{val}}\mathrm{v} \text{ if } \tau = STR$$

$$E_{value} = \boldsymbol{W_{num}} \cdot \ln(1 + |v|) + b_{num} \text{ if } \tau = NUM$$

$$E_{value} = 0 \text{ if } \tau \in \{IP, HEX, PATH\}$$

This hybrid approach ensures that categorical strings STR retain their semantic meaning via lookups, numerical values NUM retain their magnitude without causing gradient instability (via log-scaling), and highly volatile identifiers (IP, HEX, PATH) are masked out to prevent overfitting, retaining only their overarching type context via E_{type}. For a log event containing multiple parameters, we aggregate them (e.g., via summation) to form the final parameter vector.

This formulation allows the model to distinguish parameter semantics, magnitude, and structural roles while avoiding memorization of high-cardinality identifiers, thereby improving generalization to unseen log instances.

Gated Key Formulation

Let a log sequence be $S' = \{[CLS], e_1, \dots, e_L\}$ where [CLS] is a prepended learnable token for sequence-level classification. Each event e_i has a template embedding $\boldsymbol{H_T}$ (extracted via a pre-trained semantic encoder Sentence-BERT) and a synthesized parameter vector $H_P = E_{param}$.

To dynamically inject this parameter context into the sequence encoding process, we modify the Self-Attention mechanism. For each attention head h, the enhanced Key matrix is computed as:

$$\mathrm{K^h} = \mathrm{H_T}\,\mathrm{W_K^h} + \sigma\left(\mathrm{H_P}\mathrm{W}_{gate}^{\mathrm{h}}\right) \odot \mathrm{LayerNorm}(\mathrm{H_P}\mathrm{W_p^h})$$

The sigmoid gate $\sigma \cdot$ prevents gate saturation by using a zero-centered initialization, allowing the model to adaptively ignore irrelevant parameters (gate to 0). Instead of naive mean pooling, the final sequence representation z_{seq} is obtained by extracting the output embedding corresponding to the [CLS] token, which inherently captures the temporal order and dependencies of the enhanced events.

4. Experiment

Datasets: We evaluate our method on three widely used benchmark datasets from LogHub [17]:

• HDFS: Logs collected from Hadoop map-reduce jobs running on Amazon EC2.

• Thunderbird and BGL: These contain logs from supercomputer systems at Sandia National Labs (SNL). The datasets include both alert and non-alert messages identified by category tags.

Data Partitioning: For HDFS, log sequences are constructed based on block IDs. For BGL and Thunderbird, a non-overlapping sliding window of size 20 is used to generate sequences. To simulate real-world temporal evolution, we split each dataset chronologically: the first 80% of sequences are used for training, and the remaining 20% for testing. Log parsing is performed using Drain with tailored regular expressions.

Experiment Setting: To ensure reproducibility, the key hyperparameters and environment settings are as follows:

- Architecture**:** We employ a Transformer-based encoder with layers in {1, 2} and 8 attention heads. The feedforward network (FFN) dimension is set to 3072.

- Optimization**:** The learning rate is controlled by a OneCycleLR scheduler, with the maximum learning rate tuned within {1e-5, 5e-5, 1e-6}.

We use F1-score (the harmonic mean of precision and recall) and AUC (Area Under the ROC Curve, which measures the model's ability to distinguish between classes at various thresholds) to provide a comprehensive evaluation of detection performance.

Table 1: AUC and F1 Score Comparison for Standard Transformer-based method and ours method.

Method	HDFS (F1 / AUC)	Thunderbird (F1 / AUC)	BGL (F1 / AUC)
Standard Transformer	0.91 / 0.98	0.92 / 0.97	0.93 / 0.98
Ours	**0.95** / **0.99**	**0.96** / **0.99**	**0.95** / **0.99**

The results demonstrate that incorporating parameter awareness yields an average improvement of 3.6% in F1-score across the three datasets. This highlights the value of retaining and dynamically fusing parameter-level information (typically masked during parsing), which provides critical contextual signals for anomaly detection beyond template semantics alone.

The experimental results validate that our parameter-aware framework enhances log anomaly detection. Our Hybrid Parameter Embedding preserves these signals, such as latency spikes or unauthorized IP access, without overwhelming the model with high-cardinality noise. The Gated Key Injection mechanism is crucial. By using a sigmoid gate, the model learns to prioritize parameter information only when it is relevant to the event template, effectively filtering out “parameter noise” that often leads to false positives in simpler models.

While effective for structured and semi-structured logs, the model’s performance relies on the quality of the initial Regex parsing. If a system produces completely unstructured logs without clear key-value patterns, the "Key-Type" extraction might be less informative.

5. Conclusion

In this paper, we presented a parameter-aware representation framework designed to bridge the gap in conventional log anomaly detection. By explicitly modeling (key, value, type) tuples and integrating them via a Gated Key Injection mechanism, our model moves beyond simple template semantics to capture subtle, data-driven anomalies. Future research will focus on automating the Regex discovery process to handle entirely unstructured logs and exploring

the application of this framework in real-time edge computing environments where resource constraints are tighter.

References

1. A. Aziz and K. Munir, "Anomaly Detection in Logs Using Deep Learning," *IEEE Access*, vol. 12, pp. 176124–176135, 2024, doi: 10.1109/ACCESS.2024.3506332.
2. M. Landauer, S. Onder, F. Skopik, and M. Wurzenberger, "Deep learning for anomaly detection in log data: A survey," *Mach. Learn. Appl.*, vol. 12, p. 100470, Jun. 2023, doi: 10.1016/j.mlwa.2023.100470.
3. M. Du and F. Li, "Spell: Streaming Parsing of System Event Logs," in *2016 IEEE 16th International Conference on Data Mining (ICDM)*, Oct. 2016, pp. 859–864. doi: 10.1109/ICDM.2016.0103.
4. "Drain: An Online Log Parsing Approach with Fixed Depth Tree | IEEE Conference Publication | IEEE Xplore." Accessed: Feb. 27, 2026. Online. Available: https://ieeexplore.ieee.org/abstract/document/8029742
5. H. Guo, S. Yuan, and X. Wu, "LogBERT: Log Anomaly Detection via BERT," in *2021 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2021, pp. 1–8. doi: 10.1109/IJCNN52387.2021.9534113.
6. J. Qi *et al.*, "LogEncoder: Log-Based Contrastive Representation Learning for Anomaly Detection," *IEEE Trans. Netw. Serv. Manag.*, vol. 20, no. 2, pp. 1378–1391, Jun. 2023, doi: 10.1109/TNSM.2023.3239522.
7. S. Huang, Y. Liu, C. Fung, H. Wang, H. Yang, and Z. Luan, "Improving Log-Based Anomaly Detection by Pre-Training Hierarchical Transformers," *IEEE Trans. Comput.*, vol. 72, no. 9, pp. 2656–2667, Sep. 2023, doi: 10.1109/TC.2023.3257518.
8. Z. Yin, X. Kong, and C. Yin, "Semi-supervised log anomaly detection based on bidirectional temporal convolution network," *Comput. Secur.*, vol. 140, p. 103808, May 2024, doi: 10.1016/j.cose.2024.103808.
9. C. Zhang *et al.*, "LayerLog: Log sequence anomaly detection based on hierarchical semantics," *Appl. Soft Comput.*, vol. 132, p. 109860, Jan. 2023, doi: 10.1016/j.asoc.2022.109860.
10. X. Wei *et al.*, "Log-based anomaly detection for distributed systems: State of the art, industry experience, and open issues," *J. Softw. Evol. Process*, vol. 36, no. 8, p. e2650, 2024, doi: 10.1002/smr.2650.
11. P. Himler, M. Landauer, F. Skopik, and M. Wurzenberger, "Towards Detecting Anomalies in Log-Event Sequences with Deep Learning: Open Research Challenges," in *Proceedings of the 2023 European Interdisciplinary Cybersecurity Conference*, in EICC '23. New York, NY, USA: Association for Computing Machinery, Tháng Sáu 2023, pp. 71–77. doi: 10.1145/3590777.3590789.
12. "HitAnomaly: Hierarchical Transformers for Anomaly Detection in System Log | IEEE Journals & Magazine | IEEE Xplore." Accessed: Feb. 27, 2026. Online. Available: https://ieeexplore.ieee.org/abstract/document/9244088
13. "Log Sequence Anomaly Detection Based on Template and Parameter Parsing via BERT." Accessed: Feb. 27, 2026. Online. Available: https://ieeexplore.ieee.org/abstract/document/10599834
14. H. Uchida, K. Tominaga, H. Itai, Y. Li, and Y. Nakatoh, "Multi-Parameter Log Anomaly Detection with an Unsupervised Learning Approach," in *2024 International Symposium on Parallel*

Computing and Distributed Systems (PCDS), Sep. 2024, pp. 1–5. doi: 10.1109/PCDS61776.2024.10743635.

15. S. He, J. Zhu, P. He, and M. R. Lyu, "Experience Report: System Log Analysis for Anomaly Detection," in *2016 IEEE 27th International Symposium on Software Reliability Engineering (ISSRE)*, Oct. 2016, pp. 207–218. doi: 10.1109/ISSRE.2016.21.
16. A. Joulin, E. Grave, P. Bojanowski, M. Douze, H. Jégou, and T. Mikolov, "FastText.zip: Compressing text classification models," Dec. 12, 2016, *arXiv*: arXiv:1612.03651. doi: 10.48550/arXiv.1612.03651.
17. J. Zhu, S. He, P. He, J. Liu, and M. R. Lyu, "Loghub: A Large Collection of System Log Datasets for AI-driven Log Analytics," Sep. 13, 2023, *arXiv*: arXiv:2008.06448. doi: 10.48550/arXiv.2008.06448.

Phát hiện bất thường Log có nhận thức tham số

Tóm tắt: *Các phương pháp phát hiện bất thường nhật ký (log) truyền thống thường thay thế các tham số biến đổi (ví dụ: độ trễ, địa chỉ IP, mã băm) bằng các mã đại diện (tokens) chung, làm mất đi các tín hiệu ngữ cảnh quan trọng để nhận diện những bất thường tinh vi. Bài báo này đề xuất một khung biểu diễn log nhận thức tham số với mô-đun Nhúng tham số lai. Mô-đun này phân tích các tham số thành các bộ giá trị (khóa, giá trị, loại) và áp dụng các phương thức nhúng dựa trên loại tham số, bao gồm thang đo logarit cho dữ liệu số và các token chỉ định loại cho các định danh có độ biến động cao, nhằm ngăn chặn hiện tượng quá khớp trong khi vẫn bảo toàn được ngữ nghĩa. Ngoài ra, cơ chế Tiêm khóa có cổng được đề xuất để tích hợp động ngữ cảnh của tham số vào các lớp tự chú ý của Transformer. Kết quả thực nghiệm trên các bộ dữ liệu HDFS, Thunderbird và BGL cho thấy phương pháp của nghiên cứu tốt hơn so với các mô hình Transformer tiêu chuẩn, đạt mức cải thiện trung bình 3,6% về chỉ số F1-score. Các kết quả này làm nổi bật khả năng ưu việt của khung mô hình trong việc phát hiện các bất thường nhạy cảm với tham số và tăng cường khả năng diễn giải cho các ứng dụng an ninh mạng trong thực tế.*

Từ khóa: phát hiện bất thường log, tích hợp tham số, an toàn thông tin.

Contrastive Graph Autoencoder for Robust Log Anomaly Detection in Low-Data Regimes

Nguyen Hai Nam
Le Quy Don Technical University
Email: namnh23@lqdtu.edu.vn, Tel: 0978.124.889

Abstract

Log anomaly detection plays a pivotal role in safeguarding the reliability and security of large-scale networked systems, particularly under data-scarce conditions where labeled examples are limited. Conventional techniques, including sequence-based and graph-based methods, often exhibit degraded performance when trained on small subsets of data due to insufficient representation learning and sensitivity to sparse patterns. In this study, we introduce a contrastive graph-based framework that enhances graph autoencoders with InfoNCE loss to robustly detect anomalies in system logs without requiring log parsing. The model constructs per-sequence graphs from pre-trained embeddings, incorporating temporal edges between consecutive logs and similarity edges based on cosine thresholds. The contrastive loss in the latent space promotes invariant and discriminative features learned from normal data alone. Through unsupervised training and reconstruction error-based scoring, our approach addresses data limitations by improving generalization. Experiments on the BGL dataset, using only 2%, 5%, 10%, 15%, and 20% of the total data, demonstrate superior F1-scores compared to baselines such as Log2Vec, LogRobust and NeuralLog, underscoring its efficacy in low-data scenarios.

Keywords*: Log anomaly detection; Graph convolutional networks; Contrastive learning; Low-data regimes; System logs.*

1. Introduction

Monitoring and detecting anomalies in networked systems play a crucial role in early identification and prevention of incidents, intrusions, and malicious behaviors, as well as root cause analysis [8, 10, 13, 14]. Recent decades have witnessed the rapid growth of modern computer network systems, which deliver a diverse range of services. This evolution has necessitated the scaling of data centers, networks, and connected devices to meet increasing demands [1]. Log data, generated from software systems, operating systems, firewalls, or embedded programs on IoT devices, serves as a rich source of information for tasks such as decision support, incident handling, threat detection, and tracing root causes [12, 15, 16]. These systems often produce massive volumes of log lines—up to millions daily, amounting to 30–50 gigabytes per hour or petabytes per month, as seen in examples like Google systems [2]. However, logs are inherently unstructured, heterogeneous, contain rich semantic information, and exhibit strong sequential correlations, making manual analysis (e.g., keyword searches or regular expression matching) time-consuming, error-prone, and unscalable [3].

As systems grow in complexity, automated techniques for log-based anomaly detection (LAD) have become essential. Traditional machine learning (ML) methods, such as Support Vector Machines (SVM), Principal Component Analysis (PCA), and Logistic Regression (LR), rely on quantitative log event counts projected into vector spaces, identifying anomalies as deviations from norms [5, 8, 10]. Deep learning (DL) approaches have shown superior performance by modeling complex relationships [1, 5, 6, 10]. Yet, a critical challenge arises in low-data regimes, common in real-world scenarios like new systems or resource-constrained environments, where only small subsets (e.g., 2%–20%) of data are available. Existing methods

often degrade sharply due to overfitting, high variance, and insufficient generalization to novel or subtle anomalies.

Log-based anomaly detection typically follows a four-step workflow: log parsing, log partitioning, log representation, and anomaly detection [6]. Log parsing converts unstructured logs into structured formats by extracting constant log event templates (keywords) and parameters (variables) using tools like Drain or FT-Tree [7]. However, traditional parsers introduce errors, out-of-vocabulary (OOV) issues, and semantic information loss, negatively impacting downstream tasks [3, 11]. Log partitioning groups logs into sequences based on strategies such as fixed windows, sliding windows, or session windows, determined by timestamps or order [5]. Log representation vectorizes templates using techniques like IDs, Word2Vec, FastText, or pre-trained language models (e.g., RoBERTa, BERT) to capture semantics [7, 11]. Finally, anomaly detection employs ML/DL models to identify deviations.

Approaches can be categorized into quantitative-based, sequence-based, and graph-based methods. Quantitative-based techniques (e.g., PCA, One-Class SVM) focus on event counts but miss semantic and sequential information, leading to inconsistent performance [3, 13]. Sequence-based methods, such as DeepLog (LSTM for prediction) and LogBERT (Transformer for contextual modeling), capture patterns effectively but struggle with long-term dependencies, structural relationships, and result in false alarms or instability [2, 3, 4, 10, 8]. Moreover, purely discriminative training (e.g., cross-entropy loss) often leads to "shallow" representations over-specialized to training data, discarding complex semantic and structural details, which limits generalization to novel anomalies, especially in low-data settings [10]. Graph-based methods, like LogGD and BertGCN (BERT + GCN for semantic and structural graphs), convert logs into graphs to model dependencies, improving interpretability and robustness without traditional parsing [6, 15].

Among baselines, Log2Vec uses Word2Vec for semantic embeddings but is limited in capturing deep context, handling OOV, and modeling structural dependencies, making it suboptimal for complex logs [7]. LogRobust combines FastText, TF-IDF, and Bi-LSTM for semantic robustness in evolving logs but is computationally heavy and degrades in low-data regimes due to lack of discriminative features [4]. NeuralLog employs BERT for embeddings from raw logs (bypassing traditional parsing) and a Transformer encoder with softmax for supervised classification, effectively capturing semantics but suffers from shallow representations due to pure discriminative loss, leading to poor generalization on subtle or unseen anomalies.

The following points summarize the limitations of current approaches: 1. Sequence-based techniques like Log2Vec, LogRobust, and NeuralLog often miss structural relationships among log events, leading to false alarms and unstable detection performance. 2. Traditional log parsing produces extraneous data, ambiguous outcomes, and OOV tokens that degrade anomaly detection; while methods like NeuralLog and BertGCN mitigate this, they remain sensitive to data scarcity. 3. Most methods fail to capture deep semantic correlations, relying on shallow representations from discriminative training, which discards essential information and limits generalization to novel anomalies. 4. Existing approaches assume large datasets but degrade in low-data regimes (e.g., small subsets), causing overfitting and poor robustness.

To address these challenges, we propose a contrastive graph autoencoder framework for log anomaly detection. Our model uses BERT to extract embeddings from raw logs, bypassing traditional parsing while capturing rich semantics. It constructs per-sequence graphs with temporal edges (consecutive logs) and similarity edges (cosine similarity >0.85), employs a GCN encoder to produce latent representations, and a decoder for reconstruction. We integrate InfoNCE contrastive loss on the latent space (with edge dropout augmentation) as a regularizer, combining generative (reconstruction) and discriminative objectives in a multi-task setup to learn robust, invariant, and generalizable features from normal data alone (unsupervised). Anomalies are detected via reconstruction error scoring. This enhances performance in low-data scenarios by promoting a comprehensive latent space.

The main contributions are:

- A novel integration of contrastive learning into a graph-based framework, addressing shallow representations and improving generalization in low-data regimes.

- Comprehensive empirical evaluation on the BGL dataset using 2%, 5%, 10%, 15%, and 20% data subsets, demonstrating superiority over baselines (Log2Vec, LogRobust, NeuralLog), including an ablation study on the loss balance hyperparameter.

The remainder of this paper is organized as follows. Section 2 presents the proposed framework. Section 3 describes the experimental setup. Section 4 reports results and discussion. Section 5 concludes and outlines future research.

2. The Proposed Framework

In this section, we present the proposed contrastive graph autoencoder framework for log anomaly detection. The framework addresses the limitations of existing methods by integrating BERT-based semantic embeddings from raw logs (bypassing traditional parsing), graph construction to capture structural and temporal relationships, and a multi-task learning objective combining reconstruction and contrastive losses for robust representations in low-data regimes. Fig. 1 illustrates the overall architecture, which processes log sequences in an end-to-end manner: raw logs are first grouped into sequences using sliding windows, embedded via BERT to generate semantic vectors, fed into a graph-based autoencoder comprising a GCN encoder and decoder, augmented with contrastive loss for enhanced generalization, and finally scored for anomalies based on reconstruction errors.

The model operates in an unsupervised manner, training solely on normal log sequences to learn invariant features, and detects anomalies as deviations in reconstruction quality. Key notations include: Let $S = \{s_1, s_2, ..., s_n\}$ denote a log sequence, where each s_i is a raw log message. Embeddings are represented as $E = \{ e_1 , e_2 ,\dots , e_n \}$, with $e_i \in R^{768}$. Graphs are built as G = (V, A), where V are nodes (unique embeddings) and A are edges (temporal and similarity-based).

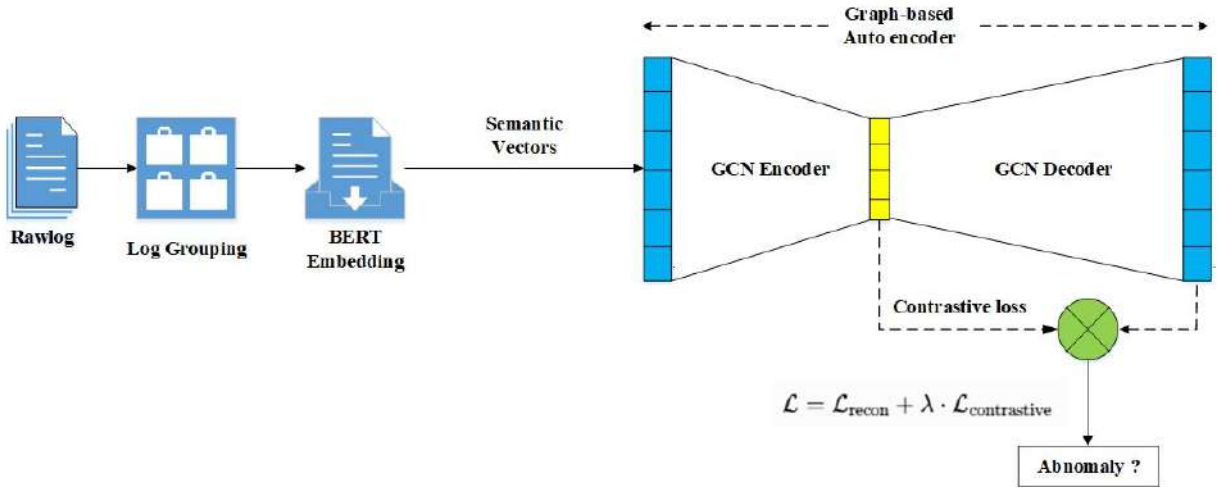


Figure 1: Overview of the Proposed Framework

2.1. Overview

The proposed framework consists of five main stages: (1) log grouping via sliding windows to partition raw logs into sequences; (2) BERT embedding extraction to capture semantic information without traditional parsing; (3) per-sequence graph construction incorporating temporal and similarity edges; (4) graph encoding with a GCN-based autoencoder, trained using a multi-task loss (reconstruction and InfoNCE contrastive) on normal data with augmentation; and (5) anomaly detection based on mean squared error (MSE) reconstruction scores during inference.

This design draws inspiration from multi-task learning paradigms [10], where reconstruction acts as a generative regularizer to prevent shallow representations, while contrastive loss enhances discriminativeness and invariance, particularly beneficial in low-data scenarios (e.g., 2%–20% of the BGL dataset). By modeling logs as graphs, the framework captures both sequential patterns and semantic correlations, improving upon sequence-based baselines like NeuralLog [9] and graph-based ones like BertGCN [17] through explicit contrastive regularization.

2.2. Graph Construction

Raw log messages are first processed using a pre-trained BERT model to generate 768-dimensional contextual embeddings, mitigating OOV issues and semantic loss associated with traditional parsers like Drain [7]. For each sequence S, embeddings E are obtained by tokenizing and feeding the messages into BERT, producing fixed-length vectors that encode deep semantic meanings.

Graphs are constructed per sequence to model relationships. Unique embeddings are mapped to nodes using a dictionary with rounded tuples as keys to handle floating-point precision and avoid duplicates. If an embedding is invalid (e.g., wrong shape), it is skipped with a warning.

Edges include:

- **Temporal edges**: Undirected connections between consecutive nodes in the sequence, capturing workflow order.

- **Similarity edges**: Computed vectorized via cosine similarity matrix on node features. Edges are added for pairs exceeding a threshold, excluding self-loops.

The resulting graph is represented as a PyTorch Geometric Data object. Empty or invalid sequences are skipped. This construction enables the model to capture both dynamic workflows (temporal) and semantic proximities (similarity), similar to text-based graphs in BertGCN [17], but tailored for sequence-level analysis.

2.3. Model Architecture

The core model is a Graph Autoencoder (GraphEncoder), comprising an encoder and decoder. The encoder uses two Graph Convolutional Network (GCN) layers:

- Conv1: $R^{768} \rightarrow R^{128}$, followed by ReLU activation.
- Dropout: Probability 0.2.
- Conv2: $R^{128} \rightarrow R^{64}$, followed by ReLU.

The decoder is a linear layer: $R^{64} \rightarrow R^{768}$, reconstructing the input features. Node features are standardized using a StandardScaler fitted on flattened normal embeddings to ensure numerical stability. The architecture leverages GCN's ability to propagate information across graph structures, producing compact latent representations that preserve semantic and structural details.

2.4. Training Procedure

Training is unsupervised, using only normal sequences (labels y=0) from the training set. A LogDataset and DataLoader (batch size 32) process sequences. For each batch:

- Build graphs as in Section 2.2.
- Standardize features.
- Augment two views by applying edge dropout (p=0.2, force undirected) independently on the edge index.
- Forward each view through the shared encoder to obtain recon1,emb1 and recon2, emb2.

The multi-task loss is:

- Contrastive loss: InfoNCE on normalized latents emb1,emb2 (temperature 0.5): Encourages positive pairs (same graph views) to be closer than negatives.
- Total_loss: recon_loss + 0.5×contrastive_loss

Optimization uses Adam (lr=0.001) over 10 epochs, promotes robust features by viewing augmentation as a form of self-supervision, mitigating overfitting in low-data regimes.

2.5. Anomaly Detection

In evaluation mode, test sequences are processed in batches: Embeddings are standardized, concatenated (no edges for simplicity, as reconstruction is node-level), and forwarded to compute reconstructions. Total_loss scores are normalized to [0,1], with predictions >0.5 as anomalies. Metrics include F1-score, precision, recall, and ROC-AUC, evaluated against ground-truth labels. This threshold-based approach detects deviations from learned normal patterns, with contrastive enhancement ensuring sensitivity to subtle anomalies even with limited training data.

3. Experimental Setup

In this section, we describe the experimental setup designed to evaluate the proposed contrastive graph autoencoder framework. The primary goal is to assess its robustness and generalization in low-data regimes, where only small subsets of training data are available, simulating real-world scenarios such as emerging systems or resource-limited environments. We focus on the BGL dataset, conducting experiments with varying training data ratios (2%, 5%, 10%, 15%, and 20% of normal sequences). Comparisons are made against established baselines, using standard metrics for anomaly detection. All experiments are implemented in Python using PyTorch and Torch Geometric libraries, ensuring reproducibility. Results and detailed analysis are presented in Section 4.

3.1. Datasets

We evaluate the framework on the BGL (Blue Gene/L) dataset [16], a widely used benchmark for log anomaly detection sourced from a supercomputer system at Lawrence Livermore National Labs. BGL contains approximately 4.7 million log lines, spanning over 215 days, with binary labels indicating normal or anomalous events (about 11% anomalous). Logs capture system states, errors, and operations, making it suitable for testing structural and semantic modeling.

Preprocessing involves extracting raw log messages, tokenizing them, and generating 768-dimensional embeddings using a pre-trained BERT model (bert-base-uncased) to bypass traditional parsing and handle OOV issues [11]. Sequences are partitioned using session windows based on process IDs (PIDs) and timestamps, resulting in variable-length sequences (average ~50 logs per sequence). The dataset is split into training (80%, normal only for unsupervised learning) and testing (20%, including both normal and anomalous for evaluation).

To simulate low-data regimes, we randomly subsample the normal training sequences at ratios of 2%, 5%, 10%, 15%, and 20%, ensuring stratified sampling to maintain distribution. For each ratio, the model is retrained from scratch, and testing is performed on the full test set.

3.2. Baselines

We compare against three representative baselines, selected for their relevance to sequence-based and semantic-focused log anomaly detection:

- **Log2Vec** [7]: A semantic embedding method using Word2Vec to vectorize log templates, followed by a logistic regression classifier. It captures basic word-level semantics but lacks deep contextual modeling.
- **LogRobust** [4]: Combines FastText embeddings with TF-IDF weighting and a Bi-LSTM network for sequence classification. It emphasizes robustness to evolving logs but is computationally intensive.
- **NeuralLog** [9]: A Transformer-based approach using BERT for embeddings from raw logs (similar to ours) and a supervised encoder with softmax classification. It bypasses traditional parsing but relies on purely discriminative training, leading to potential shallow representations in low-data settings.

Baselines are re-implemented or adapted using the same BERT embeddings and preprocessing for fair comparison. Where applicable, they are trained in a semi-supervised

manner (normal data for feature learning, with labels for classification), matching our unsupervised setup as closely as possible.

3.3. Evaluation Metrics

Performance is assessed using four standard anomaly detection metrics: F1-score, precision, recall, and ROC-AUC. These metrics are computed on binary predictions obtained by applying a threshold of 0.5 to the normalized reconstruction error scores, and then compared directly against the ground-truth labels.

Specifically, the F1-score is reported as the primary metric because it represents the harmonic mean of precision and recall, thereby providing a balanced measure of detection performance. Precision quantifies the proportion of correctly identified anomalies among all predicted anomalies, while recall measures the proportion of actual anomalies that are successfully detected. In addition, ROC-AUC evaluates the model's overall ranking capability independent of any specific threshold. To handle zero-division cases (e.g., when there are no positive predictions), all metrics are set to 0. All experiments are conducted using 5-fold cross-validation to ensure result stability, with the mean and standard deviation reported for each metric..

3.4. Implementation Details

The framework is implemented in PyTorch 2.0 with Torch Geometric 2.3, running on a single NVIDIA RTX 3080 GPU (or CPU fallback). Key hyperparameters include: embedding dimension 768 (from BERT), hidden dimensions 128/64, dropout 0.2, similarity threshold 0.85, contrastive temperature 0.5, lambda_contrastive 0.5, batch size 32, epochs 10, and Adam optimizer with learning rate 0.001. Features are standardized using StandardScaler fitted on normal training embeddings.

4. Results and Discussion

In this section, we present the experimental results and analyze the performance of the proposed contrastive graph autoencoder framework. The evaluation focuses on its effectiveness in low-data regimes, comparing it against baselines across varying training data subsets on the BGL dataset. We report F1-scores as the primary metric for balanced assessment. Overall, the results demonstrate the framework's superior generalization, attributed to the integration of contrastive regularization and graph-based modeling, which mitigates shallow representations and overfitting in data-scarce scenarios [10].

Table 1. F1-Scores of the Proposed Method and Baselines Across Low-Data Ratios on the BGL Dataset

Data	Log2Vec	LogRobust	NeuralLog	Proposed Method
2%	0.203	0.835	0.804	0.796
5%	0.541	0.793	0.831	0.813
10%	0.453	0.376	0.842	0.838
15%	0.516	0.747	0.781	0.819
20%	0.476	0.708	0.741	0.871

4.1. Performance Comparison

Table 1 summarizes the F1-scores of the proposed framework and the three baselines across low-data ratios (2%–20% of normal training sequences). The proposed method achieves

competitive performance in extremely low-data regimes (2%–10%) and consistently outperforms all baselines when the training data ratio increases to 15% and 20%.

For instance, at 2% data, our framework attains an F1-score of 0.796, which is very close to NeuralLog's 0.804. NeuralLog benefits from its strong pre-trained BERT embeddings combined with supervised discriminative training; even with extremely limited data, the model can rapidly learn decision boundaries by leveraging the rich semantic features already encoded in BERT. In contrast, our unsupervised contrastive graph autoencoder requires sufficient normal samples to effectively learn invariant and structural representations. As the data ratio grows, this advantage of the proposed method becomes evident: at 20% data, our F1-score reaches 0.871, substantially surpassing NeuralLog's 0.741. This improvement stems from the integration of graph-based structural modeling (temporal and similarity edges) and contrastive regularization, which promote richer, more generalizable latent representations..

4.2. Discussion

The results highlight the framework's effectiveness in low-data regimes, particularly as the amount of available normal data increases. The InfoNCE contrastive loss serves as a powerful regularizer that encourages the model to learn rich, invariant latent representations preserving both semantic and structural information—unlike the comparatively shallow features produced by purely discriminative training in NeuralLog [9]. Graph construction further strengthens this capability by explicitly modeling temporal dependencies and semantic similarities, thereby reducing false positives on subtle anomalies. This advantage echoes the benefits observed in BertGCN [17], but is amplified here through the multi-task contrastive objective.

Limitations of the current approach include computational overhead when constructing similarity edges for long sequences and sensitivity to hyperparameters such as the cosine similarity threshold (0.85). Future work could extend the evaluation to additional datasets (e.g., HDFS), incorporate explainability techniques such as GNNExplainer, or investigate adaptive weighting of the contrastive loss term (λ) to better handle dynamically varying data scarcity.

5. Conclusion and Future Work

In this paper, we proposed a contrastive graph autoencoder framework for log anomaly detection that effectively addresses key challenges in capturing structural relationships, semantic correlations, and robust representations under data scarcity. By leveraging BERT for semantic embeddings directly from raw logs (thereby bypassing traditional parsing), constructing per-sequence graphs with temporal and similarity edges, and integrating a multi-task loss that combines reconstruction with InfoNCE contrastive objectives, the framework learns invariant and discriminative features from normal data in a fully unsupervised manner.

Experimental results on the BGL dataset demonstrate the effectiveness of our approach, with F1-scores ranging from 0.796 (at 2% data) to 0.871 (at 20% data). The method shows competitive performance with strong baselines in extremely low-data regimes and achieves clear superiority as the amount of training data increases, outperforming Log2Vec, LogRobust, and NeuralLog in most settings. The ablation study further validates the important role of contrastive regularization in improving generalization, especially under limited data conditions.

This work advances existing graph-based approaches [17] by incorporating contrastive learning as an effective regularizer within a multi-task framework, thereby extending their applicability to sparse training scenarios that are common in real-world systems. Limitations of the current framework include the computational overhead of constructing similarity edges for very long sequences and sensitivity to hyperparameters such as the cosine similarity threshold (0.85).

Future research could explore multi-dataset validation (e.g., HDFS alongside BGL), adaptive loss weighting to handle dynamically varying data availability, and the integration of explainability tools such as GNNExplainer to interpret detected anomalies. Additionally, extending the framework to online learning or federated settings would further enhance its practicality in evolving distributed systems.

References

1. Chen, Z., Liu, J., Gu, W., Su, Y., Lyu, M.R. (2021). Experience report: Deep learning-based system log analysis for anomaly detection. arXiv preprint arXiv:2107.05908.
2. Du, M., Li, F., Zheng, G., Srikumar, V. (2017). Deeplog: Anomaly detection and diagnosis from system logs through deep learning. In: Proc. SIGSAC. ACM. pp. 1285–1298.
3. Guo, H., Yuan, S., Wu, X. (2021). Logbert: Log anomaly detection via bert. In: 2021 international joint conference on neural networks (IJCNN). pp. 1–8. IEEE.
4. Hastie, T., Tibshirani, R., Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer.
5. He, S., Zhu, J., He, P., Lyu, M.R. (2016). Experience report: System log analysis for anomaly detection. In: Proc. ISSRE. pp. 207–218. IEEE.
6. Jia, P., Cai, S., Ooi, B.C., Wang, P., Xiong, Y. (2023). Robust and transferable log-based anomaly detection. Proceedings of the ACM on Management of Data 1(1), 1–26.
7. Joulin, A., Grave, E., Bojanowski, P., Douze, M., Jégou, H., Mikolov, T. (2016). Fasttext.zip: Compressing text classification models. arXiv preprint arXiv:1612.03651.
8. Landauer, M., Onder, S., Skopik, F., Wurzenberger, M. (2023). Deep learning for anomaly detection in log data: A survey. Machine Learning with Applications 12, 100470.
9. Le, V.H., Zhang, H. (2021). Log-based anomaly detection without log parsing. In: Proc. ASE. pp. 492–504. IEEE.
10. Le, V.H., Zhang, H. (2022). Log-based anomaly detection with deep learning: How far are we? In: Proceedings of the 44th international conference on software engineering. pp. 1356–1367.
11. Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692.
12. Lu, S., Rao, B., Wei, X., Tak, B., Wang, L., Wang, L. (2017). Log-based abnormal task detection and root cause analysis for spark. In: Proc. ICWS. pp. 389–396. IEEE.
13. Nicolau, M., McDermott, J., et al. (2019). Learning neural representations for network anomaly detection. IEEE transactions on cybernetics 49(8), 3074–3087.
14. Vu, L., Nguyen, Q.U., Nguyen, D.N., Hoang, D.T., Dutkiewicz, E., et al. (2022). Learning latent representation for iot anomaly detection. IEEE Transactions on Cybernetics 52(5), 3769–3782.

15. Zhang, B., Zhang, H., Le, V.H., Moscato, P., Zhang, A. (2023). Semi-supervised and unsupervised anomaly detection by mining numerical workflow relations from system logs. Automated Software Engineering 30(1), 4.
16. Zhang, B., Zhang, H., Moscato, P., Zhang, A. (2020). Anomaly detection via mining numerical workflow relations from logs. In: Proc. SRDS. pp. 195–204. IEEE.
17. Alsalmi, E., Alhuzali, A., & Alhothali, A. (2025). Log-Based Anomaly Detection of System Logs Using Graph Neural Network. Computers, Materials & Continua, 86(2), 1–20. DOI: 10.32604/cmc.2025.071012.

Contrastive Graph Autoencoder cho phát hiện nhật ký hệ thống bất thường trong điều kiện dữ liệu hạn chế

Tóm tắt

Việc phát hiện bất thường trong nhật ký hệ thống đóng vai trò then chốt trong việc bảo vệ độ tin cậy và an ninh của các hệ thống mạng quy mô lớn, đặc biệt là trong điều kiện khan hiếm dữ liệu khi các ví dụ được gán nhãn còn hạn chế. Các kỹ thuật truyền thống, bao gồm các phương pháp dựa trên chuỗi và dựa trên đồ thị, thường bộc lộ sự suy giảm hiệu suất khi huấn luyện trên các tập dữ liệu nhỏ do việc học biểu diễn không đầy đủ và sự nhạy cảm với các mẫu dữ liệu thưa thớt. Trong nghiên cứu này, chúng tôi giới thiệu một khung làm việc dựa trên đồ thị tương phản (contrastive graph-based), giúp tăng cường các bộ tự mã hóa đồ thị (graph autoencoders) với hàm mất mát InfoNCE để phát hiện các bất thường trong nhật ký hệ thống một cách mạnh mẽ. Mô hình xây dựng các đồ thị theo từng chuỗi từ các vector nhúng được huấn luyện trước, kết hợp các cạnh giữa các nhật ký liên tiếp và các cạnh tương đồng dựa trên ngưỡng cosine. Hàm mất mát tương phản (Contrastive loss) trong không gian tiềm ẩn giúp thúc đẩy các đặc trưng bất biến và có tính phân biệt được học chỉ từ dữ liệu bình thường. Thông qua quá trình huấn luyện không giám sát và chấm điểm dựa trên lỗi tái tạo, phương pháp của chúng tôi giải quyết các hạn chế về dữ liệu bằng cách cải thiện khả năng tổng quát hóa. Các thử nghiệm trên bộ dữ liệu BGL, chỉ sử dụng 2%, 5%, 10%, 15% và 20% tổng lượng dữ liệu, đã chứng minh điểm F1 vượt trội so với các mô hình cơ sở như Log2Vec, LogRobust và NeuralLog, khẳng định tính hiệu quả của phương pháp trong các kịch bản ít dữ liệu.

Từ khóa: *Phát hiện bất thường nhật ký; Mạng tích chập đồ thị; Học tương phản; Chế độ ít dữ liệu; Nhật ký hệ thống.*

NGHIÊN CỨU MÔ HÌNH SENTENCE-BERT ĐỂ ĐÁNH GIÁ MỨC ĐỘ TRÙNG LẶP VĂN BẢN HÀNH CHÍNH TIẾNG VIỆT

Phạm Thanh Tùng[1] , Nguyễn Quốc Khánh[2] , Phạm Thị Bích Vân[3]

[1]*Học viện Kỹ thuật Quân sự*

Tungpham31891@gmail.com; Tel:0969.579.891

Tóm tắt

Nghiên cứu này đề xuất phương pháp đánh giá mức độ trùng lặp văn bản hành chính tiếng Việt sử dụng mô hình Sentence-BERT (SBERT) - một kiến trúc deep learning dựa trên Transformer được tối ưu hóa cho bài toán so sánh ngữ nghĩa văn bản. Hệ thống được xây dựng thông qua quy trình huấn luyện 3 giai đoạn: (1) Pre-training trên BERT-base; (2) Tiếp tục pre-training trên Vi-STSB (Vietnamese Semantic Textual Similarity Benchmark) để tối ưu hóa cho tiếng Việt, tạo ra mô hình `keepitreal/vietnamese-sbert`; (3) Fine-tuning trên bộ dữ liệu văn bản hành chính. Mô hình được huấn luyện trên bộ dữ liệu gồm 10.000 cặp văn bản hành chính được gán nhãn thủ công. Kết quả thực nghiệm cho thấy mô hình đạt hệ số tương quan Pearson r = 0,92, RMSE = 0,09, MAE = 0,07 và R^2 = 0,85, vượt trội rõ rệt so với các phương pháp truyền thống như TF-IDF (correlation = 0,80, RMSE = 0,37) và Jaccard similarity (correlation = 0,80, RMSE = 0,36). Nghiên cứu chứng minh tính hiệu quả của SBERT trong việc hiểu ngữ nghĩa sâu của văn bản tiếng Việt, đặc biệt là khả năng nhận diện các cặp văn bản có ý nghĩa tương đồng nhưng sử dụng từ ngữ khác nhau. Hệ thống có thể được ứng dụng trong việc phát hiện văn bản trùng lặp, đánh giá chất lượng văn bản hành chính và hỗ trợ quá trình biên tập tại các đơn vị trong quân đội.

Từ khóa: *Sentence-BERT; so sánh văn bản; trùng lặp văn bản; xử lý ngôn ngữ tự nhiên;*

1. Đặt vấn đề

Trong bối cảnh số hóa và chuyển đổi số hiện nay, việc quản lý và xử lý văn bản hành chính tại các đơn vị trong quân đội, số lượng văn bản hành chính được tạo ra và lưu trữ hàng ngày rất lớn, dẫn đến nhu cầu cấp thiết về các công cụ tự động hóa để đánh giá chất lượng, phát hiện trùng lặp và hỗ trợ quá trình biên tập văn bản [1, 2].

Việc đánh giá độ tương đồng văn bản trong miền dữ liệu hành chính tại Việt Nam đối mặt với những thách thức lớn về khả năng trích xuất đặc trưng ngữ nghĩa từ đặc điểm ngôn ngữ học đơn lập. Khác với các ngôn ngữ các, tiếng Việt có ranh giới từ phức tạp với các cấu trúc từ ghép không có rào cản phân tách vật lý, gây nhiễu cho các thuật toán phân đoạn và làm suy giảm chất lượng đầu vào của các bộ mã hóa. Đặc biệt, hệ thống thuật ngữ Hán-Việt chuyên ngành với tính hàm súc cao tạo ra một không gian phân phối từ vựng rất đặc thù.

Bài toán so sánh mức độ trùng lặp văn bản truyền thống thường dựa trên các phương pháp thống kê như TF-IDF (Term Frequency-Inverse Document Frequency), Jaccard similarity, hoặc Sequence matching [3, 4]. Tuy nhiên, các phương pháp này có những hạn chế cơ bản: chúng chỉ so sánh dựa trên từ khóa và không hiểu được ngữ nghĩa sâu của văn bản [5].

Với sự phát triển của deep learning và các mô hình ngôn ngữ lớn (Large Language Models), các phương pháp dựa trên Transformer như BERT, RoBERTa, và đặc biệt là Sentence-BERT (SBERT) đã chứng minh hiệu quả vượt trội trong việc hiểu ngữ nghĩa văn bản [6, 7]. SBERT là một biến thể của BERT được tối ưu hóa đặc biệt cho bài toán so sánh câu, cho phép encode văn bản thành các vector embedding đa chiều chứa thông tin ngữ nghĩa sâu [8].

Mặc dù các mô hình ngôn ngữ tiền huấn luyện dành riêng cho tiếng Việt như PhoBERT và ViBERT đã chứng minh hiệu quả vượt trội trong nhiều tác vụ NLP. Trong nghiên cứu này, mô hình được khởi tạo từ BERT-base đa ngôn ngữ nhằm xây dựng một quy trình huấn luyện theo hướng thích nghi dần từ ngữ liệu tổng quát đến miền dữ liệu chuyên biệt. Việc lựa chọn mô hình nền đa ngôn ngữ cho phép hệ thống kế thừa các biểu diễn ngôn ngữ có tính khái quát cao, hạn chế việc bị lệ thuộc sớm vào các đặc trưng của một miền dữ liệu cụ thể. Cách tiếp cận huấn luyện nhiều giai đoạn này không chỉ giúp mô hình từng bước thích nghi với tiếng Việt và miền dữ liệu mục tiêu, mà còn cho phép đánh giá rõ ràng vai trò của từng pha trong việc cải thiện hiệu suất tổng thể.

2. Các nghiên cứu liên quan

Trong giai đoạn đầu của lĩnh vực xử lý ngôn ngữ tự nhiên (NLP), các hướng tiếp cận đánh giá độ tương đồng văn bản chủ yếu dựa trên đặc trưng từ vựng (lexical features). Các kỹ thuật thống kê như TF-IDF hoặc các phép đo độ giao thoa như Jaccard Similarity tập trung vào tần suất xuất hiện và sự trùng lặp ký hiệu giữa các túi từ (bag-of-words). Tuy nhiên, các phương pháp này bộc lộ hạn chế lớn khi đối mặt với hiện tượng đa nghĩa và đồng nghĩa. Nghiên cứu của Blei và cộng sự (2003) [11] về mô hình Latent Dirichlet Allocation (LDA) đã nỗ lực giải quyết vấn đề này bằng cách trích xuất các chủ đề tiềm ẩn (latent topics). Dù vậy, LDA vẫn chưa thể mô hình hóa chính xác các mối quan hệ ngữ nghĩa phức tạp ở mức độ câu (sentence-level semantics) trong một không gian vector liên tục và cao chiều.

Sự ra đời của cơ chế tự chú ý (Self-Attention) trong kiến trúc Transformer (Vaswani et al., 2017) [9] đã tạo ra một bước ngoặt quan trọng, cho phép mô hình hóa các phụ thuộc đường xa (long-range dependencies) trong văn bản. Mặc dù các mô hình như BERT (Devlin et al., 2018) [13] đạt được kết quả ấn tượng, nhưng việc đánh giá độ tương đồng theo cơ chế Cross-Encoder đòi hỏi chi phí tính toán rất lớn, gây tắc nghẽn khi xử lý các tập dữ liệu quy mô lớn.

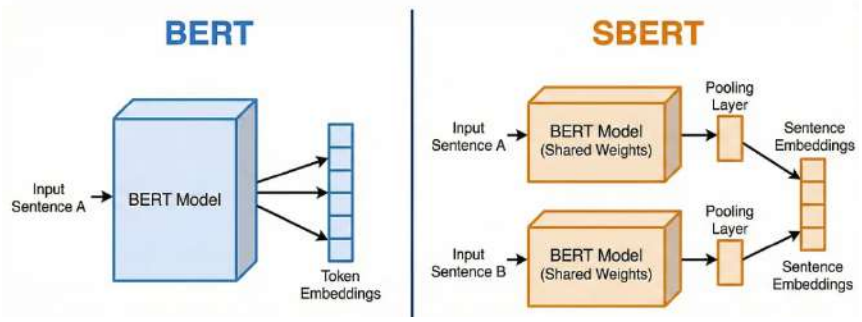
Để tối ưu hóa hiệu năng, Reimers và Gurevych (2019) [14] đã đề xuất kiến trúc Sentence-BERT (SBERT). Bằng cách sử dụng mạng nơ-ron Siamese kết hợp với các chiến lược gộp đặc trưng (Pooling), SBERT ánh xạ các câu văn thành một biểu diễn vector cố định (Embedding). Phương pháp này cho phép tính toán độ tương đồng thông qua hàm Cosine Similarity trong thời gian thực mà vẫn bảo toàn được các đặc trưng ngữ nghĩa tiềm ẩn.

Tại Việt Nam, các mô hình ngôn ngữ tiền huấn luyện như PhoBERT (Nguyen & Tuan Nguyen, 2020) [15] và ViBERT (Bui et al., 2020) [16] đã giải quyết hiệu quả bài toán phân đoạn từ cho ngôn ngữ đơn lập. Tuy nhiên, các tài liệu này chủ yếu tập trung vào lĩnh vực tổng quát. Theo nghiên cứu của Chalkidis và cộng sự (2020) [17], việc thực hiện thích nghi miền (Domain Adaptation) thông qua tinh chỉnh (Fine-tuning) trên dữ liệu chuyên biệt là yếu tố tiên quyết để mô hình nắm bắt được các cấu trúc cú pháp và thuật ngữ pháp lý khắt khe. Đây là cơ sở lý thuyết then chốt để chúng tôi thực hiện tinh chỉnh SBERT cho miền dữ liệu hành chính trong nghiên cứu này.

3. Phương pháp

3.1. Kiến trúc Sentence-BERT

Sentence-BERT (SBERT) là một kiến trúc neural network được phát triển dựa trên BERT, được tối ưu hóa đặc biệt để tạo ra các sentence embeddings có thể so sánh trực tiếp thông qua cosine similarity [8]. Khác với BERT gốc yêu cầu chạy qua cả hai câu để tính similarity (tốn thời gian), SBERT mã hóa mỗi câu độc lập thành một vector embedding, sau đó so sánh các vector này.



Hình 1. Kiến trúc BERT và SBERT

Quy trình hoạt động của SBERT bao gồm các bước sau:

Bước 1: Văn bản đầu vào được chia thành các token theo từ vựng của mô hình BERT.

Bước 2: Mỗi token được chuyển đổi thành vector embedding thông qua các lớp Transformer

Bước 3: Cơ chế attention cho phép mô hình tập trung vào các phần quan trọng của câu và hiểu mối quan hệ giữa các từ.

Bước 4: Các token embeddings được tổng hợp thành một sentence embedding duy nhất thông qua phép pooling (thường là mean pooling hoặc CLS token).

Bước 5: Độ tương đồng giữa hai văn bản được tính bằng cosine similarity giữa hai sentence embeddings:

$$similarity = \cos(\theta) = \frac{E_1 \cdot E_2}{|E_1| \times |E_2|}$$

Trong đó:

+ E_1 và E_2 là các vector embedding của văn bản 1 và văn bản 2

+ E_1 . E_2 là tích vô hướng của hai vector

+ $|E_1|$ và $|E_2|$ là độ dài của các vector.

3.2. Quy trình huấn luyện mô hình

Quy trình huấn luyện mô hình SBERT được thực hiện qua 3 giai đoạn tuần tự:

Giai Đoạn 1: Tiền huấn luyên (Pre-training)

Mô hình bắt đầu từ BERT-base, một mô hình Transformer đã được tiền huấn luyện trên dữ liệu đa ngôn ngữ lớn. Giai đoạn này cung cấp cho mô hình kiến thức ngôn ngữ tổng quát và khả năng hiểu cấu trúc ngữ pháp cơ bản.

Giai Đoạn 2: Tiếp tục Pre-training trên Vi-STSB

Sau giai đoạn pre-training, mô hình được tiếp tục huấn luyện trên bộ dữ liệu Vi-STSB được phát hành trên Hugging Face (một nền tảng lưu trữ mô hình và dữ liệu NLP phổ biến), bộ

dữ liệu lớn gồm các cặp câu tiếng Việt được thu thập từ nhiều nguồn khác nhau như Wikipedia, báo chí, văn bản hành chính, v.v.

Kết quả của giai đoạn này là mô hình SBERT đã được tối ưu hóa cho tiếng Việt, có khả năng xử lý văn bản tiếng Việt với độ dài tối đa 512 tokens.

Giai Đoạn 3: Tinh chỉnh (Fine-tuning) trên Dữ Liệu Văn Bản Hành Chính

Mô hình từ giai đoạn 2 được fine-tune tiếp trên bộ dữ liệu văn bản hành chính cụ thể để tối ưu hóa hiệu suất cho miền dữ liệu này.

Trong quá trình training, mô hình được đánh giá trên tập xác thực sử dụng EmbeddingSimilarityEvaluator để theo dõi hiệu suất. Sau khi training, mô hình được đánh giá trên tập test gồm 1.000 mẫu.

4. Thực nghiệm và đánh giá

Để đánh giá hiệu quả của SBERT, nghiên cứu so sánh với các phương pháp truyền thống:

1. Cosine Similarity với TF-IDF: Vector hóa văn bản bằng TF-IDF và tính cosine similarity giữa các vector.

2. Jaccard Similarity: Tính tỷ lệ số từ chung trên tổng số từ duy nhất của hai văn bản.

3. Sequence Similarity: Sử dụng SequenceMatcher từ thư viện difflib để so sánh dựa trên cấu trúc chuỗi ký tự.

4.1. Bộ Dữ Liệu

Trong nghiên cứu này, nhóm nghiên cứu đã thu thập các văn bản hành chính là các thông tư và nghị định, văn bản luật có liên quan. Dữ liệu được thu thập từ nguồn công khai trên trang Cổng Thông tin điện tử Chính phủ, sau đó được tiền xử lý và chuẩn hóa định dạng. Đặc trưng của loại văn bản này là cấu trúc chặt chẽ, ngôn ngữ mang tính quy phạm, sử dụng nhiều thuật ngữ pháp lý đồng thời tồn tại nhiều cách diễn đạt khác nhau cho cùng một nội dung quy định.

Bộ dữ liệu gồm 10.000 cặp văn bản được xây dựng từ các văn bản đã thu thập được nói trên và gán nhãn thủ công theo thang điểm tương đồng ngữ nghĩa trong khoảng [0,1], dựa trên mức độ tương đồng về nội dung quy định và ý nghĩa pháp lý, thay vì chỉ dựa trên sự trùng lặp từ vựng.

Mỗi mẫu dữ liệu gồm một cặp câu và điểm tương đồng trong khoảng [0, 1], trong đó 0 là hoàn toàn khác biệt và 1 là hoàn toàn giống nhau. Bộ dữ liệu được chia thành 8.000 mẫu cho tập huấn luyện, 1.000 mẫu cho mỗi tập xác thực và tập kiểm tra.

4.2. Các độ đo

Các chỉ số đo được sử dụng để đánh giá hiệu suất:

RMSE (Root Mean Squared Error): Đo sai số bình phương trung bình

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|$$

MAE (Mean Absolute Error): Đo sai số tuyệt đối trung bình

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|$$

R^2 Score (Coefficient of Determination): Đo tỷ lệ phương sai được giải thích bởi mô hình

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}$$

Trong đó:

+ SS_{res} là tổng bình phương sai số dư

+ SS_{tot} là tổng bình phương độ lệch so với giá trị trung bình

Correlation: Đo mức độ tương quan tuyến tính giữa giá trị thực và dự đoán

$$r = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2 \sum_{i=1}^{N}(y_i - \bar{y})^2}}$$

Trong đó:

+ x_i là giá trị thực tế thứ i

+ y_i là giá trị dự đoán thứ i

+ $\overline{(x)}$ và $\overline{(y)}$ là giá trị trung bình của x và y

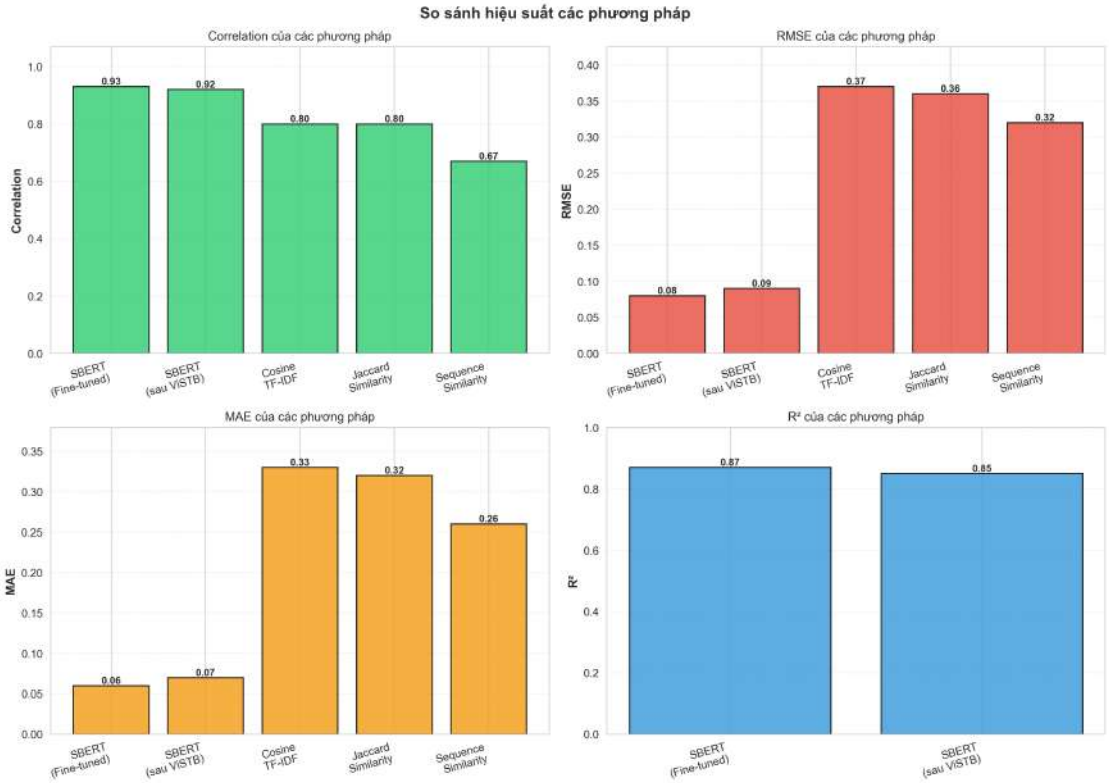
4.3. Kết quả và thảo luận

Kết quả trên tập Test

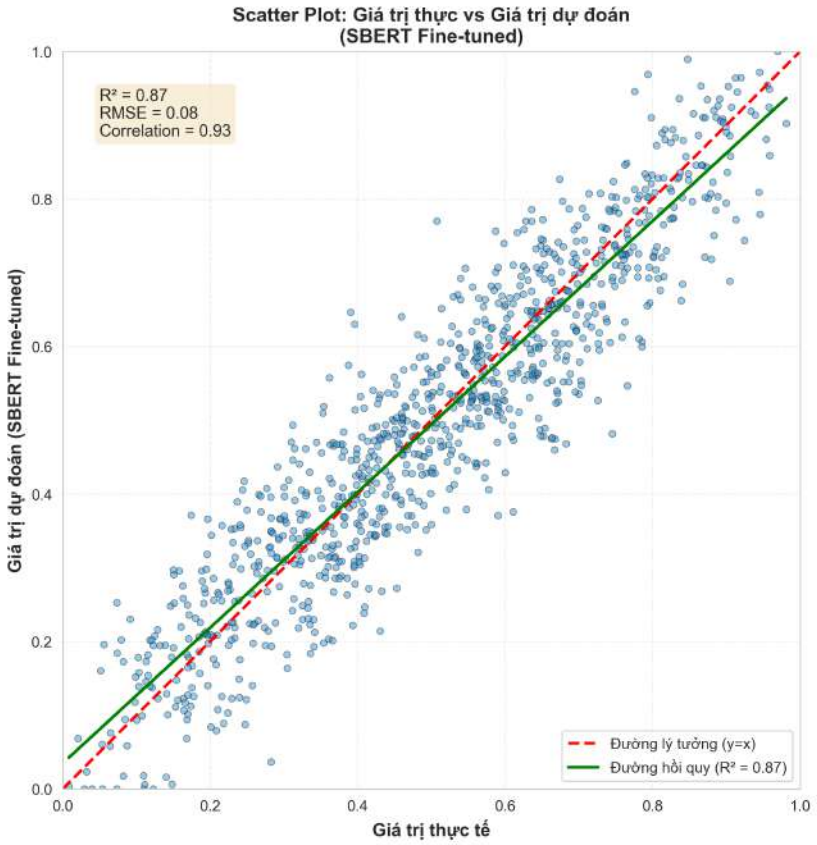
Bảng 1. So sánh hiệu suất các phương pháp trên tập test

STT	Phương pháp	Correlation	RMSE	MAE	R^2
1	SBERT (Fine-tuned, 8000 mẫu)	0.93	0.08	0.06	0.87
2	SBERT (chưa fine-tune)	0.92	0.09	0.07	0.85
3	Cosine TF-IDF	0.80	0.37	0.33	-
4	Jaccard Similarity	0.80	0.36	0.32	-
5	Sequence Similarity	0.67	0.32	0.26	-

Kết quả trong Bảng 1 cho thấy SBERT fine-tuned với 8.000 mẫu đạt hiệu suất tốt nhất, vượt trội so với tất cả các phương pháp khác. Mô hình fine-tuned đạt hệ số tương quan 0,93 (cao hơn 0,01 so với pre-trained), RMSE = 0,08 (giảm 0,01), MAE = 0,06 (giảm 0,01) và R^2 = 0,87 (tăng 0,02). Điều này chứng tỏ fine-tuning với bộ dữ liệu đầy đủ đã cải thiện đáng kể hiệu suất mô hình, giúp mô hình học được các đặc trưng cụ thể của domain văn bản hành chính tiếng Việt.



Hình 2. So sánh hiệu suất các phương pháp



Hình 3. Giá trị thực vs Giá trị dự đoán (SBERT Fine-tuned)

Khả năng nhận diện ngữ nghĩa

Một trong những ưu điểm nổi bật của SBERT là khả năng nhận diện các cặp văn bản có ý nghĩa tương đồng nhưng sử dụng từ ngữ khác nhau. Ví dụ:

- Văn bản 1: "Quân đội cần nâng cao chất lượng văn bản hành chính"
- Văn bản 2: "Các đơn vị quân đội phải cải thiện chất lượng tài liệu hành chính"
- Điểm thực tế: 0,85
- SBERT Fine-tuned dự đoán: 0,86
- SBERT (sau Vi-STSB) dự đoán: 0,88
- TF-IDF dự đoán: 0,52

Trong khi các phương pháp truyền thống như TF-IDF chỉ nhận diện được một số từ chung ("quân đội", "chất lượng", "hành chính"), SBERT hiểu được rằng "nâng cao" và "cải thiện" có nghĩa tương đồng, "văn bản" và "tài liệu" là từ đồng nghĩa trong ngữ cảnh này. Mô hình fine-tuned cho dự đoán gần với giá trị thực tế hơn (0,86 so với 0,88), chứng tỏ việc fine-tuning giúp mô hình hiểu rõ hơn về domain văn bản hành chính.

Xử lý văn bản dài

SBERT cũng thể hiện ưu thế trong việc xử lý văn bản dài và phức tạp. Với khả năng encode toàn bộ câu thành một vector embedding duy nhất, mô hình có thể nắm bắt được ngữ nghĩa tổng thể của văn bản, không chỉ phụ thuộc vào các từ khóa riêng lẻ.

Phân phối sai số

Phân tích phân phối sai số của SBERT fine-tuned trên tập test cho thấy:

- Trung bình sai số tuyệt đối (MAE) là 0,06, rất thấp so với các phương pháp khác và thấp hơn so với mô hình sau Vi-STSB (MAE = 0,07)
- RMSE = 0,08 cho thấy các lỗi lớn cũng rất hiếm và thấp hơn so với mô hình sau Vi-STSB (RMSE = 0,09)
- Hệ số tương quan 0,93 chứng tỏ mô hình có khả năng dự đoán chính xác và nhất quán, cao hơn so với mô hình sau Vi-STSB (0,92)

Điều này cho thấy mô hình fine-tuned có độ tin cậy cao hơn và ít có các lỗi lớn hơn so với mô hình sau Vi-STSB, phù hợp cho ứng dụng thực tế.

5. Kết luận

Nghiên cứu đã thành công trong việc ứng dụng mô hình Sentence-BERT để đánh giá mức độ trùng lặp văn bản hành chính tiếng Việt. Kết quả thực nghiệm cho thấy:

1. SBERT đạt hiệu suất vượt trội so với các phương pháp truyền thống. Mô hình fine-tuned với 8.000 mẫu đạt hệ số tương quan 0,93, RMSE = 0,08, MAE = 0,06 và R^2 = 0,87 trên tập test gồm 1.000 mẫu - đây là kết quả tốt nhất trong tất cả các phương pháp được thử nghiệm.

2. Fine-tuning với bộ dữ liệu đầy đủ (8.000 mẫu) đã cải thiện đáng kể hiệu suất so với mô hình sau Vi-STSB: correlation tăng từ 0,92 lên 0,93, RMSE giảm từ 0,09 xuống 0,08, MAE giảm từ 0,07 xuống 0,06, và R^2 tăng từ 0,85 lên 0,87. Điều này chứng tỏ fine-tuning giúp mô hình học được các đặc trưng cụ thể của domain văn bản hành chính tiếng Việt sau khi đã được tối ưu hóa cho tiếng Việt thông qua bộ dữ liệu Vi-STSB.

3. Mô hình thể hiện khả năng hiểu ngữ nghĩa sâu, có thể nhận diện các cặp văn bản có ý nghĩa tương đồng nhưng sử dụng từ ngữ khác nhau, điều mà các phương pháp truyền thống không thể làm được. Sai số dự đoán (RMSE = 0,08) thấp hơn 4-5 lần so với phương pháp truyền thống.

4. Quy trình huấn luyện 3 giai đoạn đã chứng minh tính hiệu quả: giai đoạn pre-training cung cấp kiến thức ngôn ngữ tổng quát, giai đoạn Vi-STSB tối ưu hóa cho tiếng Việt, và giai đoạn fine-tuning tối ưu hóa cho domain văn bản hành chính. Mô hình sau giai đoạn Vi-STSB (`keepitreal/vietnamese-sbert`) đã hoạt động rất tốt với $R^2 = 0,85$, chứng tỏ giá trị của việc tiếp tục pre-training trên dữ liệu tiếng Việt. Fine-tuning với bộ dữ liệu đầy đủ đã cải thiện thêm hiệu suất lên $R^2 = 0,87$, cho thấy giá trị của việc tùy biến mô hình cho domain cụ thể.

5. Mặc dù đạt được kết quả tốt, nghiên cứu cũng gặp một số hạn chế: SBERT yêu cầu tài nguyên tính toán lớn hơn so với các phương pháp truyền thống, đặc biệt là khi fine-tuning. Mô hình được fine-tune cho văn bản hành chính, nên có thể không hoạt động tốt với các loại văn bản khác như văn bản khoa học, văn học, v.v.

Hướng phát triển tiếp theo của nghiên cứu bao gồm: (1) Mở rộng bộ dữ liệu với nhiều loại văn bản hành chính khác nhau để tăng tính đa dạng; (2) Thử nghiệm các kiến trúc mô hình khác như RoBERTa, XLM-RoBERTa; (3) Tối ưu hóa mô hình để giảm thời gian inference; (4) Nghiên cứu các kỹ thuật fine-tuning nâng cao như multi-task learning; (5) Xây dựng hệ thống tích hợp để tự động hóa quy trình đánh giá văn bản trong thực tế.

Tài liệu tham khảo

1. Nguyễn, V. H., & Trần, T. M. (2020). Ứng dụng công nghệ thông tin trong quản lý văn bản hành chính. *Tạp chí Khoa học Công nghệ*, 15(3), 45-52.
2. Lê, V. K., & Phạm, H. N. (2021). Chuyển đổi số trong quản lý văn bản tại các đơn vị quân đội. *Hội thảo Quốc gia về Công nghệ Thông tin*, 234-241.
3. Salton, G., & McGill, M. J. (1986). *Introduction to modern information retrieval*. New York, NY: McGraw-Hill.
4. Jaccard, P. (1912). The distribution of the flora in the alpine zone. *New Phytologist*, 11(2), 37-50.
5. Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge, England: Cambridge University Press.
6. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*, 4171-4186.
7. Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, 3982-3992.
8. Reimers, N., & Gurevych, I. (2020). Making monolingual sentence embeddings multilingual using knowledge distillation. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, 4512-4525.
9. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30, 5998-6008.

10. Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., ... & Stoyanov, V. (2020). Unsupervised cross-lingual representation learning at scale. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 8440-8451.
11. D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, vol. 3, 993-1022, 2003.
12. J. Devlin et al. (2018), "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *arXiv preprint arXiv:1810.04805*.
13. N. Reimers and I. Gurevych(2019), "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," in *Proc. EMNLP-IJCNLP*, 3982-3992.
14. D. Q. Nguyen and A. T. Nguyen(2020), "PhoBERT: Pre-trained language models for Vietnamese," in *Findings of ACL: EMNLP*,1037-1042.
15. P. H. Bui et al. (2020), "ViBERT: A Vietnamese Pre-trained Language Model," *arXiv preprint arXiv:2005.06997*.
16. I. Chalkidis et al. (2020), "LEGAL-BERT: The Muppets straight out of Law School," in *Findings of ACL: EMNLP*,2898-2904.

FINE-TUNING SENTENCE-BERT FOR ENHANCED TEXT DUPLICATION DETECTION IN ADMINISTRATIVE DOCUMENTS

Abstract: *This study proposes a method for evaluating the degree of text duplication in Vietnamese administrative documents using the Sentence-BERT (SBERT) model - a Transformer-based deep learning architecture optimized for semantic text comparison. The system is built through a 3-stage training process: (1) Pre-training on BERT-base; (2) Continued pre-training on ViSTB (Vietnamese Sentence Transformers Benchmark) to optimize for Vietnamese, resulting in the `keepitreal/vietnamese-sbert` model; (3) Fine-tuning on administrative document dataset. The model is evaluated on a dataset of 1,000 manually labeled pairs of administrative documents. Experimental results show that the model achieves a Pearson correlation coefficient r = 0.92, RMSE = 0.09, MAE = 0.07, and R^2 = 0.85, significantly outperforming traditional methods such as TF-IDF (correlation = 0.80, RMSE = 0.37) and Jaccard similarity (correlation = 0.80, RMSE = 0.36). The research demonstrates the effectiveness of SBERT in understanding the deep semantics of Vietnamese text, particularly its ability to identify text pairs with similar meanings but different wordings. The system can be applied to detect duplicate documents, evaluate the quality of administrative documents, and support the editing process in military units.*

Keywords: *Sentence-BERT; text comparison; text duplication; natural language processing; administrative documents.*

Unsupervised Homography Estimation for Image Stitching via Flow-Based Regression

Van Nguyen[1*], Hai-Hong Phan[1], Phat T.Nguyen[2]

[1] *Institute of Information and Communication Technology, Le Quy Don Technical University, 236 Hoang Quoc Viet, Hanoi, 100000, Vietnam;*

[2] *Faculty of Radio-Electronic Engineering, Le Quy Don Technical University, 236 Hoang Quoc Viet, Hanoi, 100000, Vietnam*

Email: vanvan22102000@gmail.com; Tel: 0941931033

Abstract

We propose an unsupervised framework for homography estimation in image stitching. While recent deep learning-based methods achieve strong performance, they often rely on complex architectures with heavy fully connected regression heads, resulting in high memory consumption and limited deployability. To address this issue, our method replaces direct parameter regression with a flow-based regression mechanism combined with quadrant aggregation. The proposed architecture integrates residual feature extraction and a vectorized cost volume construction to enable efficient and structured computation. By eliminating FC layers, the model achieves a substantial reduction in parameters. Experiments on the UDIS-D benchmark demonstrate that our method reduces parameters by approximately 98% while outperforming baseline approaches in alignment accuracy under standard metrics. These results highlight the effectiveness of flow-based regression in developing high-performance, memory-efficient homography models uniquely suited for resource-constrained image stitching systems.

Index Terms: *Deep Homography; Dense Flow; UDIS-D; Image Stitching.*

1. Introduction

1.1. Evolution of Homography Estimation

Image stitching generates seamless wide field-of-view (FOV) panoramas, underpinning applications in autonomous driving, virtual reality, and surveillance [1]. Its core, homography estimation, has been historically dominated by pipelines that extract hand-crafted features (e.g., SIFT [2], SURF [3], ORB [4]) and solve the Direct Linear Transform (DLT) via robust outlier rejection like RANSAC [5] or MAGSAC [6]. Despite their interpretability, these methods degrade in "ill-posed" environments—such as low-light, low-texture, or repetitive-pattern scenes—where reliable keypoints are scarce. Furthermore, the iterative nature of RANSAC-family algorithms [5], [6] introduces unpredictable latency, hindering real-time performance.

1.2. The Shift to Deep Learning

To overcome the limitations of hand-crafted features, deep learning approaches have dominated recent literature. Supervised methods regress homographies directly [7], [8], utilize hierarchical structures [9], or learn dynamic masking [10], but suffer from severe synthetic-to-real domain gaps. Consequently, unsupervised frameworks [11] emerged to minimize photometric errors, subsequently improved via content-aware masks [12] and GANs [13].

[*] Email: vanvan22102000@gmail.com (Van Nguyen)

Notably, the UDIS framework [14], [15] establishes a strong baseline by employing grid-based deformations to handle large parallax. Recently, advanced architectures such as Transformer-based models (e.g., HomoFormer) [16] and diffusion-based generative pipelines [17] have set new standards in perceptual quality and global geometric modeling. However, while these methods excel in handling complex contexts through self-attention mechanisms or iterative denoising processes, they inherently demand massive computational resources and suffer from significant latency. This excessive computational overhead renders them impractical for real-time applications on resource-constrained edge devices.

1.3. Problem Statement and Contributions

Despite high accuracy, modern homography models suffer from computational inefficiency, relying on massive fully connected (FC) layers that consume excessive parameters and destroy spatial integrity [12]. We propose a highly efficient framework, validated on the UDIS-D dataset [14], designed to strike an optimal balance between alignment accuracy, real-time inference speed, and minimal memory footprint. Our three key contributions are:

- Dense Flow Regression with Quadrant Aggregation: Replaces parameter-heavy FC layers with a fully convolutional dense flow mapping and quadrant-based pooling strategy. This design preserves spatial integrity and reduces total parameters by 98.0%, facilitating efficient deployment on edge devices.

- ResNet-based Feature Extraction: Utilizes a ResNet backbone [19] instead of traditional VGG encoders to extract robust, hierarchical deep features, establishing a stronger foundation for precise alignment.

- Vectorized Cost Volume with Context Fusion: Accelerates feature matching via parallelized vectorized operations and fuses the cost volume with semantic context. This provides explicit structural guidance, effectively resolving ambiguities in low-texture or repetitive regions.

2. Proposed method

2.1. Network Overview

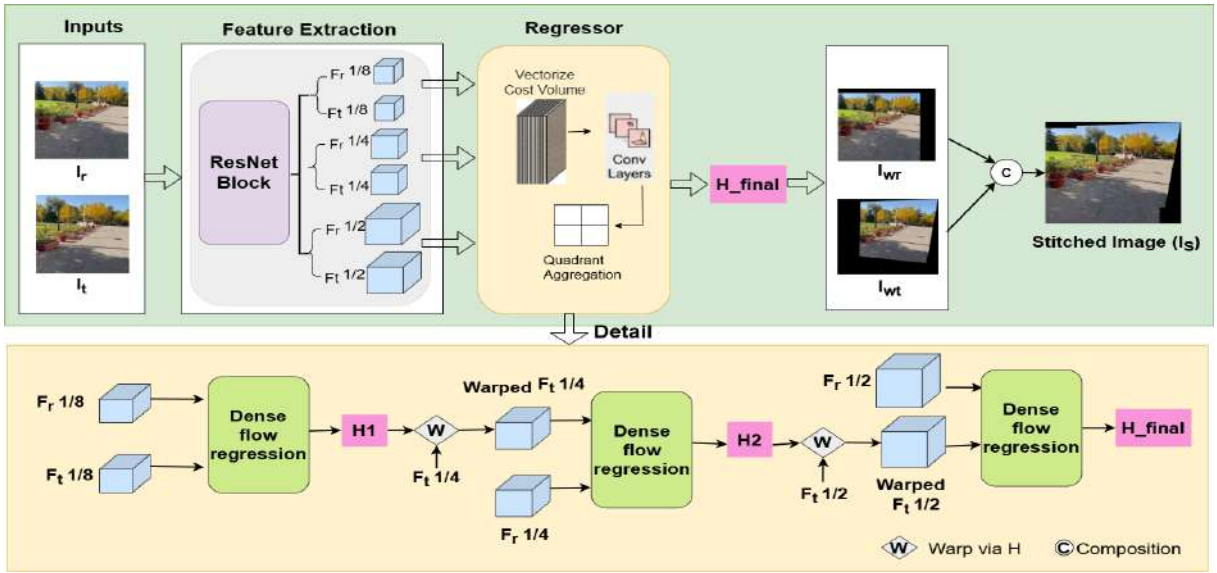


Figure 1. The overall architecture of the proposed framework.

We propose an unsupervised framework for homography estimation that operates in a coarse-to-fine manner. As illustrated in Figure 1, our architecture consists of two main components: a Feature Extractor built upon a ResNet backbone and a Multi-Stage Regression Network.

The pipeline accepts a reference image I_r and a target image I_t, both resized to a fixed resolution of 128×128. The images are processed through three hierarchical scales. At each stage, the network estimates a local homography, which is then used to warp the target features before passing them to the next finer stage. This cascade strategy allows the framework to handle large-baseline variations at the coarse level while refining precise pixel-level alignment at the fine level.

2.2. Feature Extraction

To ensure robust feature representation, we replace standard VGG extractors with a ResNet architecture. The reference image I_r and the target image I_t are processed independently to generate two distinct feature pyramids, denoted as $\{F_r^l\}$ and $\{F_t^l\}$. Specifically, the network extracts features at three hierarchical scales corresponding to $1/8$, $1/4$, and $1/2$ of the input resolution. The dimensions for each scale are designed as follows: the coarsest level maps have a size of 16×16 with 128 channels; the medium level maps are 32×32 with 64 channels; and the finest level maps are 64×64 with 64 channels. This multi-scale design effectively leverages deep semantic context while retaining precise structural details.

2.3. Dense Flow Regression Module

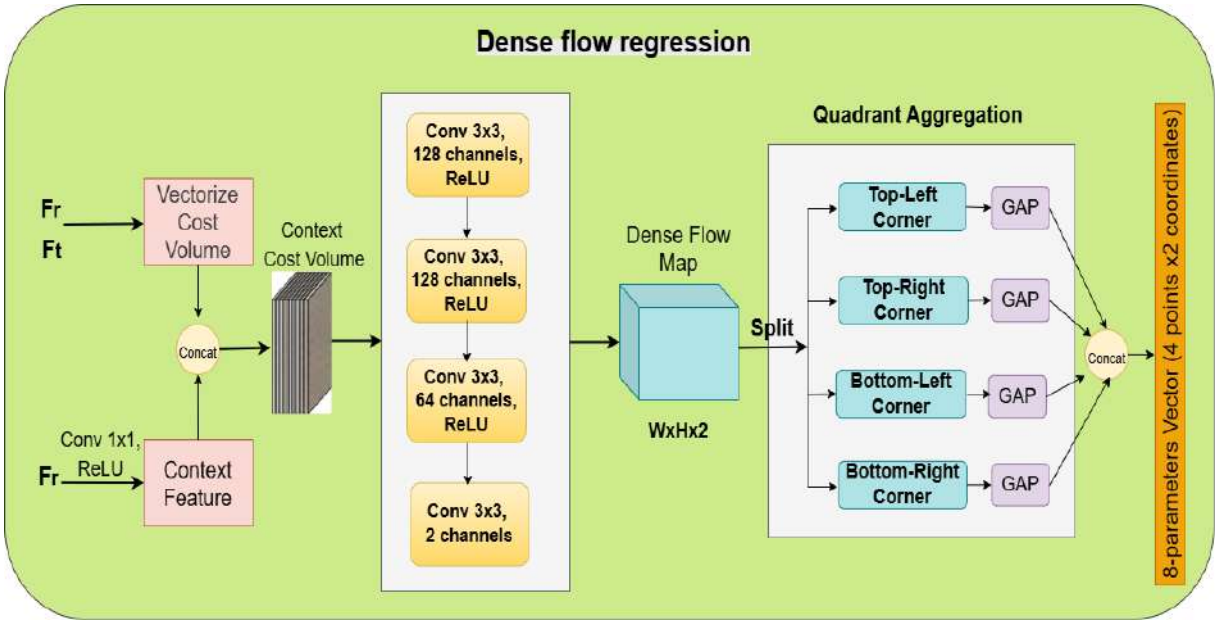


Figure 2: The detailed structure of the Dense Flow Regression block.

The framework's pivotal innovation is the Dense Flow Regression Module. By replacing heavy Fully Connected layers with a fully convolutional architecture, we preserve spatial integrity while drastically reducing complexity. As shown in Figure 2, it executes a four-stage pipeline: Vectorized Cost Volume, Context Fusion, Dense Flow Prediction, and Quadrant Aggregation.

2.3.1. Vectorized Cost Volume

To achieve high-throughput feature matching, we implement a Vectorized Cost Volume that computes correlations via parallel tensor operations. Given the reference feature map F_r and the target feature map F_t, we define a local search radius R. We extract spatial patches of size $K \times K$ (where $K = 2R+1$) from F_t and reorganize these local neighbors into the channel dimension. This transforms the sliding window operation into a dense tensor multiplication. The correlation volume $\mathbf{C}$ is computed as:

$$\mathbf{C}(u,v) = \text{LeakyReLU}\left(\sum_c F_r(u,v,c) \cdot \text{Patch}(F_t')_{(u,v,c)} \right) \quad (1)$$

By formulating the correlation as an element-wise product over the depth dimension, we fully exploit GPU parallelism for efficient inference.

2.3.2. Context Fusion

Relying solely on feature similarity for correlation can lead to ambiguities, particularly in textureless regions or scenes with repetitive patterns. To mitigate this, we introduce a Context Fusion mechanism. Instead of depending exclusively on the cost volume, we extract high-level semantic features from the backbone network, reduce their dimensionality via a convolutional layer, and directly concatenate them with the computed cost volume. This enriched representation injects explicit structural guidance into the regression network, ensuring robust motion inference even when pure photometric matching is unreliable.

2.3.3. Dense Flow Prediction

Following Context Fusion, the combined feature volume is decoded into a pixel-wise displacement field via a fully convolutional regression head. Structurally, this sub-network consists of a cascade of four convolutional layers with 3×3 kernels and ReLU activations. Crucially, we completely eliminate the use of parameter-heavy FC layers in this module. This design choice ensures that the spatial correspondence between feature positions is preserved throughout the regression process. The final convolutional layer projects the features directly into a 2-channel output without activation, generating the Dense Flow Map, denoted as $\mathbf{D} \in \sim^{H \times W \times 2}$. Unlike FC-based global regression, our dense flow prediction captures fine-grained local distortions, essential for resolving non-rigid deformations and parallax artifacts.

2.3.4. Quadrant Aggregation

Since a homography is uniquely defined by four point correspondences, we introduce a Quadrant Aggregation strategy to robustly distill the dense flow map into four corner offsets. Specifically, we spatially partition the flow map into four equal quadrants (Top-Left Q_{TL}, Top-Right Q_{TR}, Bottom-Left Q_{BL}, and Bottom-Right Q_{BR}). Under the premise that the average regional motion statistically represents the corresponding corner's displacement, we apply Global Average Pooling (GAP) to each quadrant to derive the final offset vector $\mathbf{v}_k$ for the k-th corner:

$$\mathbf{v}_k = \frac{1}{|Q_k|} \sum_{(x,y) \in Q_k} \mathbf{D}(x,y), \quad k \in \{TL, TR, BL, BR\} \quad (2)$$

This strategy offers dual benefits: it ensures robustness by averaging out local noise and outliers, and achieves parameter efficiency by replacing heavy FC layers with a parameter-free pooling operation.

2.4. Unsupervised Training Mechanism

Following the UDIS [14], our alignment network is optimized in an unsupervised manner using a multi-scale photometric loss. Specifically, we minimize the L_1 distance between the reference image and the warped target image across three scales. To eliminate boundary artifacts caused by zero-padding, we employ a validity mask M (generated by warping an all-ones tensor) to ensure the loss is computed exclusively within valid overlapping regions. The total objective function is defined as:

$$\mathrm{L}_{total} = \sum_{s=1}^{3} w_s \cdot \left\| I_{wt}^{(s)} - (I_r^{(s)} \odot M^{(s)}) \right\|_1 \qquad (3)$$

where w_s denotes the balancing weight for scale s, and $\odot$ represents the element-wise multiplication.

3. Experiments

3.1. Experimental Setup

Dataset: Evaluation was conducted on the UDIS-D benchmark [9], using the standard test split of 1,106 real-world image pairs with diverse overlaps and geometric distortions. All input images were resized to 128×128 pixels.

Implementation Details: Our model was implemented in TensorFlow and trained using a multi-scale photometric loss. To balance the magnitude of the loss across different resolutions, we set the weights w_s to 16, 4, and 1 for the coarse (16×16), medium (32×32), and fine (64×64) scales, respectively. We utilized the Adam optimizer with a batch size of 4 for a total of 200 epochs. The initial learning rate was set to 10^{-4} and decayed exponentially during training. All experiments were conducted on a workstation equipped with a single NVIDIA T1000 8GB GPU.

3.2. Quantitative Evaluation

To evaluate robustness against varying degrees of geometric distortion, we stratify the UDIS-D test set by overlap ratio into Easy (top 30%), Medium (30–60%), and Hard (60–100%) subsets (Table 1). For a fair comparison, supervised baselines (DHN [7], LB-DHN [20]) utilize official models pre-trained on the synthetic MS-COCO dataset. Meanwhile, all unsupervised frameworks, including ours, are trained exclusively on the real-world UDIS-D dataset to avoid the synthetic-to-real domain gap.

Results demonstrate that traditional SIFT+RANSAC suffers severe performance degradation on the "Hard" subset (17.58 dB) due to the scarcity of reliable keypoints in low-overlap regions. Conversely, our proposed model maintains strong performance across all difficulty levels, achieving a superior average PSNR of 22.13 dB. This significantly outperforms both supervised methods (DHN: 13.52 dB, LB-DHN: 21.14 dB) and the unsupervised UDHN (15.83 dB). Most notably, our framework surpasses the dedicated alignment stage of the UDIS baseline across critical metrics (PSNR: 22.13 vs. 21.17; SSIM: 0.728 vs. 0.648). These improvements confirm that our architecture possesses exceptional representation learning capabilities, offering a highly robust solution for unsupervised image alignment.

Table 1. Quantitative evaluation on the UDIS-D testing set.

(a) PSNR (↑) of the overlapping regions on UDIS-D dataset

Method	Traditional homography		Supervised deep homography		Unsupervised deep homography		
	I_{3x3}	SIFT[2]+ RANSAC[5]	DHN[7]	LB-DHN[20]	UDHN [11]	UDIS warp [15]	Ours
Top 0-30%	16.19	25.23	16.40	24.75	19.39	25.16	**26.28**
30-60%	13.05	22.23	13.36	21.14	15.93	20.96	**22.67**
60-100%	10.87	17.58	11.50	18.46	13.10	18.36	**18.66**
Average	13.12	21.25	13.52	21.14	15.83	21.17	**22.13**

(b) SSIM (↑) of the overlapping regions on UDIS-D dataset

Method	Traditional homography		Supervised deep homography		Unsupervised deep homography		
	I_{3x3}	SIFT[2]+ RANSAC[5]	DHN[7]	LB-DHN[20]	UDHN [11]	UDIS warp [15]	Ours
Top 0-30%	0.387	0.860	0.409	0.825	0.573	0.834	**0.871**
30-60%	0.173	0.766	0.170	0.712	0.334	0.669	**0.779**
60-100%	0.073	0.558	0.077	0.550	0.165	0.495	**0.584**
Average	0.197	0.711	0.204	0.681	0.338	0.648	**0.728**

3.3. Model Complexity Analysis

Beyond alignment accuracy, we evaluate the computational efficiency of our proposed method against the UDIS baseline on an Intel Core i7-12700 CPU (Table 2). By replacing parameter-heavy FC layers with a fully convolutional architecture, our model achieves a ~98% reduction in both parameters (188.78M to 3.84M) and storage size (2,179 MB to 49 MB).

Table 2. Comparison of model complexity.

Method	Parameters (M)	Model size (M)	FLOPs (G)	Inference Time (s)
UDIS	188.78	2,179	176.03	0.357
Ours	3.84	49	98.41	0.271
Reduction	**- 97.97 %**	**- 97.75%**	**- 44.09%**	**- 24.09%**

Furthermore, our approach significantly reduces computational overhead, cutting FLOPs by 44.09% (176.03 G to 98.41 G) and decreasing inference time by 24.09% (0.357 s to 0.271 s). These results confirm that our ultra-compact framework successfully maintains high geometric accuracy while remaining highly optimized for real-time edge-device deployment.

3.4. Qualitative Results and Generalization

Fig. 3 presents qualitative results on the UDIS-D benchmark and in unconstrained real-world environments. As shown in Figure 3(a), our framework effectively mitigates standard parallax and illumination variations on the UDIS-D dataset, yielding seamless composites with preserved geometric structures. This robustness stems from the Dense Flow Regression module, which accurately resolves local non-rigid distortions that global homography methods typically overlook. Conversely, Figure 3(b) illustrates typical failure cases on the same benchmark: scenes with extreme parallax or excessively cluttered details. While the model handles moderate viewpoint shifts well, extreme parallax exceeds the dense flow map's spatial transformation capacity, leading to misalignment. Furthermore, areas with too many high-frequency details induce ambiguous cost volume matching, resulting in noisy flow estimations. Finally, to validate the model's generalization capability beyond the UDIS-D dataset, Figure 3(c) demonstrates

additional evaluations on unconstrained real-world images captured via mobile devices. The network successfully generalizes to these in-the-wild scenarios without requiring fine-tuning, confirming its practical applicability.

Figure 3. Qualitative evaluation of the proposed framework.

4. Conclusion

In this paper, we presented a memory-efficient unsupervised deep homography estimation framework tailored for robust image stitching. The core of our contribution lies in the Dense Flow Regression with Quadrant Aggregation module, which effectively captures global spatial context and local non-rigid distortions without the need for parameter-heavy fully connected layers. Experimental results on the UDIS-D benchmark and unconstrained real-world scenarios demonstrate that our model achieves a superior balance between computational efficiency and alignment accuracy. By eliminating FC layers, we achieved a drastic ~98.0% reduction in both parameters and storage requirements compared to the strong UDIS baseline, while attaining a highly competitive average PSNR of 22.13 dB. Furthermore, our proposed architecture significantly accelerates the alignment process, reducing computational FLOPs by 44.09% and inference time by 24.09%. While limitations remain in handling large dynamic objects, our work bridges the gap between high-precision alignment and storage efficiency, providing a practical solution for real-world unsupervised image stitching.

References

1. Brown, M., & Lowe, D. G. (2007). Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision*.
2. Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2), 91–110.

3. Bay, H., Tuytelaars, T., & Van Gool, L. (2006). SURF: Speeded up robust features. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 404–417). Springer.
4. Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. (2011). ORB: An efficient alternative to SIFT or SURF. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2564–2571). IEEE.
5. Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6), 381–395.
6. Barath, D., Matas, J., & Noskova, J. (2019). MAGSAC: Marginalizing sample consensus. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 10197–10205).
7. DeTone, D., Malisiewicz, T., & Rabinovich, A. (2016). Deep image homography estimation. *arXiv preprint arXiv:1606.03798*.
8. Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. In *Proceedings of the International Conference on Learning Representations (ICLR)*.
9. Nowruzi, F. E., Laganiere, R., & Japkowicz, N. (2017). Homography estimation from image pairs with hierarchical convolutional networks. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*.
10. Le, H., Liu, F., Zhang, S., & Agarwala, A. (2020). Deep homography estimation for dynamic scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.
11. Nguyen, T., et al. (2018). Unsupervised deep homography: A fast and robust homography estimation model. *IEEE Robotics and Automation Letters*.
12. Zhang, J., et al. (2020). Content-aware unsupervised deep homography estimation. In *Proceedings of the European Conference on Computer Vision (ECCV)*.
13. Ye, N., et al. (2021). Unsupervised deep homography estimation with GAN. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.
14. Nie, L., Lin, C., Liao, K., Liu, S., & Zhao, Y. (2021). Unsupervised Deep Image Stitching: Reconstructing Stitched Features to Images. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*.
15. Nie, L., Lin, C., Liao, K., Liu, S., & Zhao, Y. (2023). Parallax-tolerant unsupervised deep image stitching. *IEEE Transactions on Image Processing*.
16. Liu, Y., Li, G., & Cui, S. (2023). Homoformer: A transformer-based network for deep homography estimation. *Pattern Recognition*, 133, 109014.
17. Zhang, Q., Pang, Y., & Wang, Y. (2024). Seamless image stitching with diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 4512–4521).
18. Chen, Z., Wang, Y., & Li, H. (2025). Mamba-Stitch: Efficient long-range dependency modeling for wide-angle image stitching. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.
19. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778).
20. Nie, L., Lin, C., Liao, K., & Zhao, Y. (2020). Learning edge-preserved image stitching from large-baseline deep homography. *arXiv preprint arXiv:2012.06194*.

Ước lượng homography không giám sát cho ghép ảnh bằng cơ chế hồi quy dựa trên luồng

Tóm tắt: Chúng tôi đề xuất một phương pháp không giám sát cho bài toán ước lượng homography trong ghép ảnh. Khác với các mô hình học sâu sử dụng các lớp kết nối đầy đủ có số lượng tham số lớn, phương pháp đề xuất thay thế hồi quy tham số trực tiếp bằng cơ chế hồi quy dựa trên luồng kết hợp với chiến lược tổng hợp theo bốn góc phần tư. Kiến trúc mạng tích hợp bộ trích xuất đặc trưng dựa trên ResNet và cơ chế xây dựng cost volume dạng vector hóa nhằm bảo đảm tính toán hiệu quả và có cấu trúc. Nhờ loại bỏ các lớp kết nối đầy đủ, mô hình giảm đáng kể số lượng tham số. Kết quả thực nghiệm trên bộ dữ liệu UDIS-D cho thấy phương pháp đề xuất giảm khoảng 98% số tham số, đồng thời cải thiện độ chính xác căn chỉnh so với các mô hình cơ sở theo các thước đo chuẩn. Kết quả này cho thấy tiềm năng ứng dụng của mô hình trong các hệ thống ghép ảnh bị ràng buộc về tài nguyên tính toán.

Từ khóa: Ước lượng homography bằng học sâu; luồng dày đặc; UDIS-D; ghép ảnh.

An adaptive ensemble framework for just-in-time defect prediction on imbalanced commit streams

Thao Bui[1*], Van Pham[1], Quoc Khanh Nguyen[1]

[1]*Institute of Information and Communication Technology, Le Quy Don Technical University, 236 Hoang Quoc Viet, Hanoi, 100000, Vietnam*

Email: pthaolb@gmail.com; Tel: 0981064048

Abstract

Just-in-Time Software Defect Prediction aims to identify defect-prone commits in a streaming setting where data arrive continuously, the distribution evolves over time, and class imbalance is severe. Purely online learners adapt quickly but can be unstable when defective commits are rare, whereas batch learners are typically more stable yet may lag behind local changes. We propose H2, an adaptive hybrid ensemble that combines an Adaptive Random Forest updated per commit with a Random Forest periodically retrained on a sliding window of recent data. Predictions are fused via dynamic weighting based on the MCC computed over the most recent evaluation window. H2 further applies local-imbalance-aware online resampling by adjusting the update intensity for defective commits using the most recent labeled window. Prequential experiments on 12 Android datasets show that H2 improves MCC over the baselines while achieving comparable PR-AUC.

Keywords: *JIT-SDP; Android; class imbalance; online learning; hybrid model*

1. Introduction

Software Defect Prediction (SDP) provides early estimates of defect-prone components or code changes to prioritize testing and code review and reduce late-stage defect rectification costs. SDP has grown into a substantial research area, synthesized by systematic literature reviews and large-scale empirical studies that discuss metrics, algorithms, and evaluation designs in both within-project and cross-project settings [1-3].

In mobile software development, particularly in the Android ecosystem, SDP faces additional challenges due to rapid release cycles, frequent changes, diverse architectures, and strong heterogeneity across projects. Although Android is widely studied, inconsistencies in datasets and evaluation protocols often limit the generalizability of findings [4].

Aligned with CI/CD practices, Just-in-Time SDP (JIT-SDP) focuses on predicting defect-prone commits at creation time, enabling risk-based testing and code review [5, 6]. JIT-SDP typically relies on commit-level change and process metrics and often employs robust classifiers such as Random Forest and other ensemble models [6-10]. More recently, deep learning approaches (e.g., DeepJIT) have leveraged commit messages and code changes for representation learning [11]. However, the dynamic nature of Android projects leads to non-stationary data distributions, making chronological evaluation and model adaptation essential [7, 8].

[*] Email: pthaolb@gmail.com (Thao Bui), vanptb@lqdtu.edu.vn (Van Pham), khanh29bk@mta.edu.vn (Quoc Khanh Nguyen)

When JIT-SDP is deployed on commit streams, performance often degrades over time due to three core challenges: (1) concept drift, where the feature - label relationship changes over time [12]; (2) class imbalance and its evolution, since defect-inducing commits are rare and the imbalance ratio fluctuates across development phases, leading to bias and instability when local windows contain few defects [13]; and (3) verification latency, because defect-inducing labels are typically confirmed only after defects are discovered and traced back, complicating online updates and temporal evaluation [15].

Accordingly, data-stream studies commonly adopt real-time evaluation, notably the prequential (test-then-train) protocol [14]. Hybrid schemes combining online and batch learning have also been explored to balance rapid adaptation and stability in JIT-SDP [15]. For imbalanced streams, cost-sensitive learning and online resampling are widely used, and surveys highlight the value of context-aware adjustment as imbalance evolves over time [13].

Most existing approaches emphasize either drift adaptation (via stream/online learning) or imbalance handling (via cost-sensitive learning/resampling). In contrast, integrating temporal adaptation, stability under rare events, and adjustment to local imbalance within a commit-stream framework - particularly for Android - remains limited [7, 8, 13]. Purely online models can be unstable when defects are rare, while purely batch models may lag behind local drifts, motivating a mechanism that harmonizes these properties in CI/CD environments [15].

To address these challenges, we propose H2, an adaptive dual-timescale ensemble for JIT-SDP on imbalanced commit streams. H2 combines (i) an online component for fast adaptation and (ii) a batch component periodically retrained on a sliding window of the most recent W commits for stability. The online branch uses Adaptive Random Forest (ARF) to handle distribution shifts [16] and ADWIN for drift detection [17]. The two branches are fused via recent-performance-based dynamic weights, while H2 also applies local-imbalance-aware online resampling.

The main contributions are: (1) a dual-timescale framework balancing adaptation and stability; (2) a dynamic aggregation mechanism based on recent performance; and (3) an imbalance-aware resampling strategy adapted to evolving data streams. To ensure reliable evaluation on imbalanced data, we report metrics such as MCC and PR-AUC [18, 19].

2. Method

2.1. Proposed Models and Research Questions

Overview of H2. In commit streams, JIT-SDP must cope with non-stationarity and locally varying class imbalance. H2 addresses this by running two learners in parallel: an online ARF updated after each labeled commit for fast adaptation, and a batch RF periodically retrained on the most recent W commits for stability. Their outputs are fused using MCC-based dynamic weighting computed over a short recent window K, while local-imbalance-aware online resampling increases minority-class updates based on the latest label window W_{cw}.

2.1.1. Problem formulation in a commit stream setting

In JIT-SDP, each commit arriving chronologically at time t is represented by a feature vector x_t and a binary label $y_t \in \{0, 1\}$ (1: defect-inducing commit, 0: clean commit). The

objective is to estimate the probability $P(y_t = 1 | \mathrm{x}_t)$ as soon as the commit appears, in order to support the prioritization of code review and testing based on change risk [5, 6].

2.1.2. The H2 Framework: A "Dual-Timescale" Adaptive Ensemble

H2 is a hybrid ensemble consisting of two parallel branches: (i) an online learner updated after each commit for rapid adaptation to concept drift, and (ii) a batch learner periodically retrained on a sliding window of recent commits to ensure stability without retaining the full history. At time t, the two branches output the predicted probabilities of a defect-inducing commit as $p_{\mathrm{on}}(t)$ and $p_{\mathrm{ba}}(t)$, respectively.

2.1.3. Dynamic Prediction Aggregation Based on Recent Performance

H2 aggregates the two predictions using a weighted sum:

$$p(t) = w_{\mathrm{on}}(t)\, p_{\mathrm{on}}(t) + w_{\mathrm{ba}}(t)\, p_{\mathrm{ba}}(t), w_{\mathrm{ba}}(t) = 1 - w_{\mathrm{on}}(t) \quad (1)$$

Estimating recent performance using sliding window MCC. To reduce noise during evaluation in the context of rare defects and class imbalance, H2 does not perform "instantaneous" evaluation on individual samples; instead, for each branch m ∈ {on, ba}, H2 maintains a sliding evaluation window of size K, stores the history (y_{true}, y_{pred}), and updates the C_m = {TP, TN, FP, FN} to calculate $MCC_m(t)$.

Updating the sliding window K. For each commit with an observed label, the queue and counters are updated as follows:

Input: y_true, y_pred ∈ {0,1}, queue Q, counts C={TP, TN, FP, FN}, window size K

1) Push (y_true, y_pred) into Q and increment the corresponding counter in C.

2) If |Q| > K:

Pop the oldest element (y_old, y_old_pred) from Q

Decrement the corresponding counter in C.

3) Calculate MCC from C.

Local MCC-based Softmax Weighting. The weight of the Online branch at time t is determined by:

$$w_{\mathrm{on}}(t) = \frac{\exp(\alpha \cdot MCC_{\mathrm{on}}(t))}{\exp(\alpha \cdot MCC_{\mathrm{on}}(t)) + \exp(\alpha \cdot MCC_{\mathrm{ba}}(t))}, w_{\mathrm{ba}}(t) = 1 - w_{\mathrm{on}}(t) \quad (2)$$

where α > 0 controls the sensitivity of the weight transition between the two branches.

Handling undefined MCC cases. If the MCC denominator equals zero (e.g., no positive samples in the window or single-class predictions), we set $MCC_m(t) = 0$ to ensure numerical stability and avoid unstable weight updates [18].

2.1.4. Dynamic Imbalance Handling via Window-based Online Resampling

To address the evolving class imbalance over time, H2 maintains a label history window of size W_{cw} to estimate the local imbalance level. Let N_{neg} and N_{pos} be the number of clean and defect-inducing samples within this window, respectively. The weight for the defective class at time t is estimated by the ratio:

$$w_{pos}(t) = \frac{N_{neg} + \varepsilon}{N_{pos} + \varepsilon} \tag{3}$$

where ε is the Laplace smoothing parameter (set to 1.0) to avoid division by zero. We bound $w_{pos}(t)$ within $[1, w_{max}]$ to avoid excessive amplification under extreme local imbalance.

Upon encountering a defect-inducing commit ($y_t = 1$), H2 performs online resampling by updating the Online model multiple times on that exact sample. The number of updates is determined by an "integer part + fractional part" mechanism: let $k = \lfloor w_{pos}(t) \rfloor$ and $f = w_{pos}(t) - k$; the model is always updated k times, and updated one additional time with probability f. This approach is suitable for data streams and avoids modifying the base learner's structure [13].

2.1.5. Research Questions

(1) RQ1: Does H2 - combining an online ARF and a periodically retrained batch RF with MCC-based dynamic fusion - outperform representative online-only and batch-only baselines (Online ARF and Batch RF) under prequential evaluation on Android commit streams?

(2) RQ2: What are the respective contributions of (i) MCC-based dynamic weighting and (ii) local-imbalance-aware online resampling to H2's performance?

(3) RQ3: How sensitive is H2 to the parameter α, the inverse-temperature in the Softmax weighting that controls how sharply weights shift between the online and batch branches?

2.2. Experiments

2.2.1. Dataset and Preprocessing

We conduct experiments on a benchmark consisting of 12 open-source Android projects, which has been widely adopted in JIT-SDP research [7, 8]. This testbed provides commit-level instances with labels generated using the SZZ algorithm and a standardized feature set, ensuring reproducibility. The dataset poses significant challenges due to the diversity in project scales and the presence of severe class imbalance. As shown in Table 1, the number of commits per project ranges from 164 to 3,780, while defect ratios vary considerably from 14.02% to 40.39% (with an average of 27.03%). Such imbalance and variation in project lifespans highlight the necessity of an adaptive approach like H2.

Table 1. Statistical characteristics of the 12 Android datasets

Project	Total Commits	Defective Commits	Defect Ratio (%)	Imbalance Ratio (1:X)
aFall	1,025	414	40.39	1.48
Alfresco	1,004	214	21.31	3.69
androidSync	209	62	29.67	2.37
androidWalpaper	588	94	15.99	5.26
anySoftKeyboard	2,971	819	27.57	2.63
Apg	3,780	1,304	34.50	1.90

chatSecure	2,579	853	33.07	2.02
facebook	548	180	32.85	2.04
kiwis	1,373	350	25.49	2.92
owncloudandroid	3,700	830	22.43	3.46
Pageturner	164	23	14.02	6.13
reddit	222	60	27.03	2.70

Each commit is characterized by six commonly used change metrics: number of modified files (NF), lines added (LA), lines deleted (LD), number of modified directories (ND), number of developers involved (NDEV), and number of unique changes (NUC) [7, 10]. To better simulate real-world scenarios and avoid data leakage, the data stream is processed in chronological order.

2.2.2. Evaluation Measures

Given the severe class imbalance, we use Matthews Correlation Coefficient (MCC) as the primary threshold-based measure, as it summarizes the confusion matrix into a single correlation coefficient in [-1, 1] and remains informative under imbalance [18]. We also report PR-AUC, i.e., the area under the precision-recall curve, to evaluate ranking quality when positive instances are rare [19]. Additionally, we report Recall, F1-Score, and G-Mean to support interpretation from a class-balance perspective.

2.2.3. Experimental Scenarios and Evaluation Protocol

Prequential (test-then-train) protocol. For each arriving commit, the model first produces a prediction and then updates itself and the sliding windows once the label is available. This protocol respects temporal order and reflects the streaming setting [14]. In this paper, we focus on the standard prequential setting for the main evaluation, while verification latency is treated as an important practical deployment issue and discussed separately due to the conference page limit.

Experimental scenarios. To answer RQ1-RQ3, we design three scenarios:

(1) S1 – Overall comparison (RQ1). Compare H2 with Online ARF and Batch RF on 12 Android projects under the prequential protocol, reporting MCC, PR-AUC, and other metrics as mean ± std over 10 runs with different random seeds.

(2) S2 – Ablation study (RQ2). Compare full H2 with two variants: (i) No-Adaptive, which disables MCC-based dynamic weighting (two branches retained, weights not updated using window K); and (ii) No-Cost, which removes local-imbalance-aware online resampling. All variants use the same data, parameters, and protocol for fair comparison.

(3) S3 – Sensitivity analysis (RQ3). Vary α in the Softmax weighting while keeping other parameters fixed; results are aggregated as mean ± std across the 12 projects.

Parameter configuration. Parameters are fixed to avoid tuning over the time series and to ensure fairness across projects: Batch sliding window W = 200; retraining cycle R = 200; evaluation window K = 100; imbalance estimation window W_{cw} = 200; sensitivity control parameter α = 8.0; Laplace smoothing parameter ε = 1.0; and w_{max} = 10.

We set W = R = 200 to balance computational efficiency and adaptability. A window size of 200 is large enough to capture recent local defect patterns while still remaining compatible with relatively small projects in our benchmark. Setting the retraining cycle equal to the window size yields a simple block-based periodic update schedule, which reduces retraining overhead while allowing the batch branch to remain reasonably up to date in CI/CD-oriented commit streams.

3. Results and Discussion

3.1. RQ1 - Overall Performance Comparison

Table 2 presents the average results (± standard deviation) of the proposed model (H2) and the baseline models across the 12 datasets, over 10 repeated runs.

Table 2. Average experimental results (± standard deviation)

Model	MCC	PR-AUC	Recall	F1-Score	G-Mean
H2 (Proposed)	**0.386 ± 0.084**	**0.553 ± 0.120**	0.627 ± 0.117	**0.560 ± 0.092**	**0.696 ± 0.064**
Online ARF	0.359 ± 0.107	0.549 ± 0.123	0.464 ± 0.162	0.503 ± 0.140	0.619 ± 0.115
Batch RF	0.219 ± 0.131	0.419 ± 0.152	**0.652 ± 0.238**	0.455 ± 0.124	0.506 ± 0.209

As shown in Table 2, H2 achieves the highest MCC (0.386), outperforming Online ARF (0.359) and Batch RF (0.219). This advantage is consistent with its higher F1-Score and G-Mean, indicating a better balance between the two classes under imbalanced data.

H2 also attains a PR-AUC of 0.553, comparable to Online ARF (0.549) and higher than Batch RF (0.419), suggesting similar ranking ability to Online ARF but better threshold-based classification performance, as reflected by MCC, F1, and G-Mean.

Although Batch RF obtains the highest Recall (0.652), its low MCC and G-Mean indicate a tendency to overpredict the positive class. In contrast, H2 maintains a comparable Recall (0.627) while achieving substantially better balanced performance.

Although an MCC of 0.386 may seem moderate, it is competitive in JIT-SDP, where prior studies report values around 0.33 under static evaluation [7, 8]. Notably, our results are obtained under a more challenging prequential setting. Furthermore, given a baseline PR-AUC equal to the defect ratio (~27.03%), H2's PR-AUC (0.553) confirms strong effectiveness under severe class imbalance.

3.2. RQ2 - Ablation Study

To assess the contributions of (i) dynamic weighting and (ii) cost-sensitive learning via online resampling, we compare the full H2 framework with two ablated variants.

Table 3. Ablation results (average across 12 datasets)

Model	MCC	G-Mean	Recall	ΔMCC	ΔRecall
H2 (Full)	**0.386**	**0.696**	**0.627**	0.000	0.000
H2 (No-Adaptive)	0.382	0.681	0.588	-0.004	-0.039
H2 (No-Cost)	0.361	0.617	0.461	-0.025	-0.165

Table 3 shows that removing dynamic weighting (H2 No-Adaptive) slightly degrades performance: MCC decreases from 0.386 to 0.382 and Recall drops from 0.627 to 0.588. This suggests that dynamic weighting provides complementary gains by adjusting the contribution of the online and batch components based on recent performance.

In contrast, removing cost-sensitive learning (H2 No-Cost) results in a substantial drop in Recall (-0.165) and a noticeable decline in MCC (-0.025), highlighting the importance of online resampling in handling local imbalance.

Overall, online resampling is the primary driver of Recall (and thus MCC), while dynamic weighting offers smaller but consistent benefits in aggregating the two components.

3.3. RQ3 - Sensitivity Analysis of Parameter α

The parameter α in the Softmax function controls how sharply the weights shift between the Online and Batch components. We conduct a sensitivity analysis of the parameter α over the set {1, 4, 8, 16, 32} to evaluate the stability of the system. The experiment yields the following results:

Table 4. Sensitivity to α (Average MCC)

α	**1**	**4**	**8**	**16**	**32**
MCC	0.384	**0.388**	0.387	0.381	0.374

MCC peaks at α = 4 and remains nearly equivalent at α = 8. As α increases further, MCC gradually decreases. The reason is that a large α makes the Softmax function highly sensitive to minor local MCC differences within the window K, causing the weights to easily fall into a "winner-takes-all" state. In data streams with noise and local drift, this mechanism causes the weights to fluctuate drastically between the Online and Batch components, leading to unstable classification decisions and reducing the average MCC. Therefore, selecting α in the range of [4, 8] is reasonable. We establish α = 8 as the a priori default value for all experiments to demonstrate that the model does not overly rely on extensive parameter tuning.

Practical note on verification latency. In practice, defect labels may arrive with delay due to bug confirmation and traceability. Although this study uses the standard prequential protocol, the dual-timescale design of H2 remains relevant under delayed feedback: the online branch adapts when labels become available, while the batch branch provides stability from recently verified history. A full delayed-label evaluation is left for future work due to the conference page limit.

4. Conclusion

This paper presented H2, an adaptive hybrid ensemble for JIT-SDP on imbalanced commit streams, combining an online ARF, a periodically retrained batch RF, MCC-based dynamic weighting, and local-imbalance-aware online resampling. Experiments on 12 Android datasets showed that H2 achieved the best overall MCC and competitive PR-AUC, while ablation results indicated that online resampling was the main driver of recall improvement and dynamic weighting provided additional gains. A limitation of the current study is that verification latency is discussed conceptually rather than evaluated under a full delayed-label protocol. Future work

will examine more realistic delayed-label settings, broader non-Android benchmarks, and SZZ-induced label noise.

References

1. Hall, T., Beecham, S., Bowes, D., Gray, D., & Counsell, S. (2011). A systematic literature review on fault prediction performance in software engineering. *IEEE Transactions on Software Engineering*, *38*(6), 1276-1304.
2. Lessmann, S., Baesens, B., Mues, C., & Pietsch, S. (2008). Benchmarking classification models for software defect prediction: A proposed framework and novel findings. *IEEE transactions on software engineering*, *34*(4), 485-496.
3. Zimmermann, T., Nagappan, N., Gall, H., Giger, E., & Murphy, B. (2009, August). Cross-project defect prediction: a large scale experiment on data vs. domain vs. process. In *Proceedings of the 7th joint meeting of the European software engineering conference and the ACM SIGSOFT symposium on The foundations of software engineering* (pp. 91-100).
4. Jorayeva, M., Akbulut, A., Catal, C., & Mishra, A. (2022). Machine learning-based software defect prediction for mobile applications: A systematic literature review. *Sensors*, *22*(7), 2551.
5. Kim, S., Whitehead, E. J., & Zhang, Y. (2008). Classifying software changes: Clean or buggy?. *IEEE Transactions on software engineering*, *34*(2), 181-196.
6. Kamei, Y., Shihab, E., Adams, B., Hassan, A. E., Mockus, A., Sinha, A., & Ubayashi, N. (2012). A large-scale empirical study of just-in-time quality assurance. *IEEE Transactions on Software Engineering*, *39*(6), 757-773.
7. Catolino, G. (2017, May). Just-in-time bug prediction in mobile applications: the domain matters!. In *2017 IEEE/ACM 4th International Conference on Mobile Software Engineering and Systems (MOBILESoft)* (pp. 201-202). IEEE.
8. Catolino, G., Di Nucci, D., & Ferrucci, F. (2019, May). Cross-project just-in-time bug prediction for mobile apps: An empirical assessment. In *2019 IEEE/ACM 6th International Conference on Mobile Software Engineering and Systems (MOBILESoft)* (pp. 99-110). IEEE Computer Society.
9. Hassan, A. E. (2009, May). Predicting faults using the complexity of code changes. In *2009 IEEE 31st international conference on software engineering* (pp. 78-88). IEEE.
10. Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.
11. Hoang, T., Dam, H. K., Kamei, Y., Lo, D., & Ubayashi, N. (2019, May). Deepjit: an end-to-end deep learning framework for just-in-time defect prediction. In *2019 IEEE/ACM 16th International Conference on Mining Software Repositories (MSR)* (pp. 34-45). IEEE.
12. Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM computing surveys (CSUR)*, *46*(4), 1-37.
13. Krawczyk, B., Minku, L. L., Gama, J., Stefanowski, J., & Woźniak, M. (2017). Ensemble learning for data stream analysis: A survey. *Information Fusion*, *37*, 132-156.
14. Gama, J. (2010). *Knowledge discovery from data streams*. CRC Press.
15. Cabral, G. G., Minku, L. L., Oliveira, A. L., Pessoa, D. A., & Tabassum, S. (2023). An investigation of online and offline learning models for online just-in-time software defect prediction. *Empirical Software Engineering*, *28*(5), 121.

16. Gomes, H. M., Bifet, A., Read, J., Barddal, J. P., Enembreck, F., Pfharinger, B., ... & Abdessalem, T. (2017). Adaptive random forests for evolving data stream classification. *Machine Learning*, *106*(9), 1469-1495.
17. Bifet, A., & Gavalda, R. (2007, April). Learning from time-changing data with adaptive windowing. In *Proceedings of the 2007 SIAM international conference on data mining* (pp. 443-448). Society for Industrial and Applied Mathematics.
18. Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA)-Protein Structure*, *405*(2), 442-451.
19. Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PloS one*, *10*(3), e0118432.

Khung học tổ hợp thích nghi cho dự báo lỗi phần mềm tức thời trên dòng commit mất cân bằng

Tóm tắt: *Dự báo lỗi phần mềm tức thời nhằm phát hiện các commit có khả năng gây lỗi trong bối cảnh dữ liệu đến liên tục, phân bố thay đổi theo thời gian và mất cân bằng lớp nghiêm trọng. Các mô hình học trực tuyến thích ứng nhanh nhưng thường kém ổn định khi commit gây lỗi hiếm, trong khi các mô hình học theo lô ổn định hơn nhưng phản ứng chậm trước biến động cục bộ. Bài báo đề xuất H2, một mô hình tổ hợp lai thích nghi kết hợp Adaptive Random Forest cập nhật theo từng commit với Random Forest được huấn luyện lại định kỳ trên cửa sổ trượt gần nhất. Kết quả dự đoán được kết hợp bằng cơ chế gán trọng số động dựa trên MCC của cửa sổ đánh giá gần nhất. H2 cũng áp dụng cơ chế lấy mẫu lại trực tuyến có xét đến mất cân bằng cục bộ cho lớp thiểu số. Thực nghiệm theo giao thức prequential trên 12 bộ dữ liệu Android cho thấy H2 cải thiện MCC và duy trì PR-AUC cạnh tranh.*

Từ khóa: dự báo lỗi phần mềm tức thời; Android; mất cân bằng lớp; học trực tuyến; mô hình lai.

Attention-Enhanced Swin UNet++ for Pavement Crack Segmentation

Thuong Dang[1]*, Hai-Hong Phan[1], Ngo Huu Phuc[1], Chu Thi Huong[1]

[1] *Institute of Information and Communication Technology, Le Quy Don Technical University, 236 Hoang Quoc Viet, Hanoi, 100000, Vietnam;*

Email: dangthuong18112000@gmail.com; Tel: 0832008663

Abstract

Pavement crack segmentation is vital for structural health monitoring and proactive maintenance. However, conventional Convolutional Neural Networks struggle to capture long-range contextual information due to their localized operations. This limitation often causes boundary fragmentation in cracks characterized by thin morphologies, low contrast, or background noise. To address this, we propose Hybrid SUAnet, a novel architecture integrating a Swin Transformer encoder into the UNet++ framework. By leveraging shifted-window self-attention, SUAnet models long-range dependencies while maintaining multi-scale feature reconstruction via nested skip connections. The architecture is further optimized by selectively pruning intermediate connections to reduce computational complexity and parameter overhead without compromising performance. Additionally, we conduct a systematic empirical analysis of three attention mechanisms within the decoder: Convolutional Block Attention Module, Sequential Fusion Attention, and Spatial-Channel Synergistic Attention. This investigation elucidates how each strategy enhances spatial-channel features to preserve thin boundaries and morphological continuity. Experimental results demonstrate that integrating these mechanisms significantly improves segmentation accuracy and boundary quality, especially under complex noise. These findings provide robust design guidelines for automated, deep learning-based pavement inspection systems.

Index Terms: *Swin Transformer; Unet++; Attention; Crack detection; Segmentation.*

1. Introduction

Pavement infrastructure faces continuous degradation from vehicle loads and climatic factors, with cracking signaling initial structural failure. Undetected cracks permit water infiltration, causing severe distresses like potholes or subsidence, which escalate maintenance costs and safety risks. High-speed imagery from modern acquisition systems necessitates automated segmentation models that balance accuracy with computational efficiency. Traditional techniques thresholding, edge detection, or morphology often fail under varying illumination and complex textures [1, 2].

Deep learning, notably Convolutional Neural Networks (CNNs), revolutionized segmentation [3]. FCN [4] pioneered end-to-end semantic segmentation, while U-Net [5] improved boundary preservation via symmetric encoder-decoder structures and skip connections. Despite enhancements through residual mechanisms or receptive field expansion, CNNs' local nature struggles to model long-range dependencies in elongated cracks. Vision Transformers (ViT) [6] exploit global dependencies via self-attention, yet high computational overhead limits high-resolution applications. The Swin Transformer [7] mitigates this through shifted-window self-attention, enabling hierarchical multi-scale representations with linear

* Email: dangthuong18112000@gmail.com (Thuong Dang), hongpth@lqdtu.edu.vn (Hai-Hong Phan)

complexity. While UNet++ [8] bridges semantic gaps through nested dense skip pathways, these connections increase parameter counts, latency, and feature redundancy.

We propose SUAnet, a hybrid architecture integrating a Swin Transformer encoder with a redesigned UNet++ structure. SUAnet employs a selective nested skip connection strategy, pruning intermediate branches based on their performance contribution. This reduces complexity and noise propagation while maintaining semantic bridge stability. Furthermore, we integrate attention mechanisms within decoder blocks to rebalance spatial-channel information and reconstruct thin crack boundaries. Primary Contributions:

1. SUAnet Architecture: A unified framework integrating Swin Transformer and UNet++ for pavement crack segmentation.
2. Selective Skip Pathways: The UNet++ architecture was optimized by pruning direct skip connections between the encoder and deep decoder nodes to achieve a superior trade-off between segmentation accuracy and computational efficiency.
3. Empirical Evaluation: A systematic study of decoder-side attention mechanisms to enhance crack boundary continuity and sharpness.

Results demonstrate improved performance under complex noise and provide an efficient foundation for large-scale automated pavement inspection.

2. Proposed method

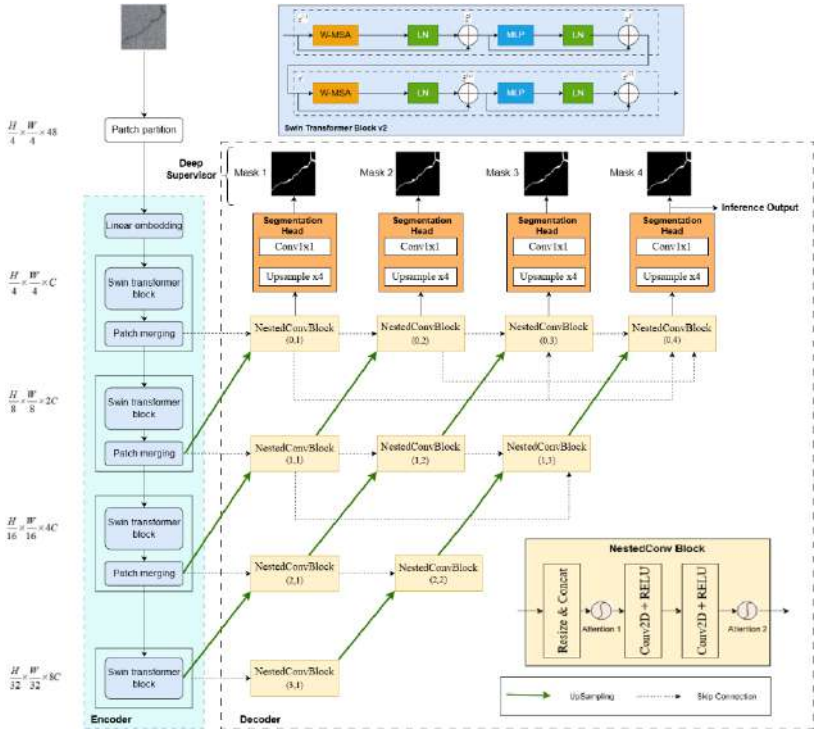


Figure 1. The overall architecture SUAnet

We propose SUAnet, a hybrid architecture combining the long-range contextual modeling of the Swin Transformer [7] with the spatial reconstruction proficiency of UNet++ [8]. Following the encoder-decoder paradigm (Fig. 1), SUAnet is optimized specifically for pavement crack segmentation.

- Encoder (Backbone): The Swin Transformer [7] extracts multi-scale hierarchical features. Its shifted-window self-attention captures wide-range spatial relationships with manageable computational complexity, enhancing the detection of elongated or low-contrast cracks.
- Decoder: Based on the UNet++ framework [8], the decoder replaces full dense pathways with a controlled skip connection strategy. This selection bridges semantic gaps; mitigates low-level noise propagation; reduces parameter overhead; and stabilizes optimization.

SUAnet's innovation lies in its hybrid structure and the strategic integration of attention mechanisms within the decoder. Systematic empirical evaluation identified an optimal configuration that, combined with selective skip connections, balances segmentation accuracy and computational efficiency. This synergy significantly improves crack continuity and boundary sharpness under complex noise conditions.

2.1. Swin Transformer v2 Encoder

The encoder partitions the input image into patches, representing them as a feature tensor of size $C x \frac{H}{4} x \frac{W}{4}$. Successive Swin Transformer V2 blocks [7] then extract features via local window-based self-attention. Within each stage, channel dimensions and spatial resolutions remain constant until reaching a Patch Merging layer. This layer executes $2x$ downsampling by grouping $2x2$ neighboring patches into single tokens, halving spatial dimensions while a linear layer projects the concatenated features into a double-dimension space. This four-stage sequence of Transformer blocks and Patch Merging layers constructs a multi-level hierarchical representation that preserves semantic information with computational tractability.

Swin Transformer V2 [7] improves training stability and high-resolution scalability through the Residual Post-Norm (Res-Post-Norm) mechanism. Unlike traditional Pre-Norm, Res-Post-Norm applies normalization after residual additions, fostering robust information flow, enhanced gradient stability, and superior optimization efficiency. As a backbone for crack segmentation, this architecture models long-range spatial relationships with near-linear complexity via shifted-window self-attention. Its stage-based structure maintains high representational capacity with fewer parameters than global self-attention, allowing the encoder to extract multi-scale features more efficiently than conventional U-Net convolutional blocks, particularly for high-resolution inputs with complex backgrounds.

2.2. Custom Decoder and Dual Attention Integration Strategy

The core contribution of this work is a redesigned decoder tailored to the hierarchical features of the Swin Transformer [7]. Diverging from the standard UNet++ [8], we introduce a selective dense connection mechanism and a dual attention integration strategy to recover spatial details while suppressing encoder-propagated semantic noise. While conventional UNet++ nodes $X^{i,j}$ aggregate features via simple concatenation—often causing redundancy and noise amplification—our Attention Guided Nested Blocks integrate an attention module $\hat{A}$ at two strategic positions: (i) before the convolutional layer to re-weight inputs, and (ii)

after the layer to refine intermediate representations. This dual-attention mechanism acts as a semantic filter, emphasizing crack-relevant features while attenuating the dispersed local noise inherent in Transformer patch-based representations.

SUAnet further departs from UNet++ [8] by disconnecting direct pathways between the encoder and deep decoder nodes. Only the first-column nodes at each level receive direct Swin Transformer features; subsequent nodes rely on refined outputs from preceding same-row nodes and upsampled features from lower levels. This design prevents high-level semantic noise propagation and fosters hierarchical refinement, enabling deep nodes to function as active filters rather than passive fusion points. Finally, we implement deep supervision at $\frac{H}{4} x \frac{W}{4}$ resolution, matching the scale of the initial patch embedding. Each output node $X^{0,j}$, where $j \in \{1,2,3,4\}$ is equipped with a Segmentation Head. Training employs a weighted sum loss across these four outputs to accelerate convergence, while inference utilizes only the most refined output for computational efficiency.

2.3. Attention Candidates

We hypothesize that effectively exploiting spatial-channel correlations is critical for high-precision pavement crack segmentation. Consequently, we investigate three attention mechanisms. Serving as the baseline, the Convolutional Block Attention Module (CBAM) [9] sequentially applies channel and spatial attention to suppress background noise and emphasize informative features before further propagation. In contrast, Sequential Fusion Attention (SFA) [10] prioritizes sequential fusion to mitigate "semantic dilution" in deep networks. By accumulating vital spatial features across successive layers, SFA preserves the structural integrity of thin, delicate crack morphologies. Finally, Spatial and Channel Synergistic Attention (SCSA) [11] optimizes resource allocation by exploiting the synergy between spatial and channel dimensions. Using lightweight operators, SCSA captures complex, branching crack patterns without significantly increasing computational overhead (FLOPs).

3. Experiments

3.1. Experimental Setup

The model was trained for 40 epochs with a batch size of 4 using the Stochastic Gradient Descent (SGD) optimizer (momentum: 0.9, weight decay: 10^{-4}, initial learning rate: 0.01). The computational environment utilized Ubuntu, Python 3.12, and PyTorch 2.9.0 with CUDA 12.8, running on an Intel Xeon Gold 6240 CPU and an NVIDIA Quadro RTX 5000 GPU. Performance was quantified using Precision, Recall, F-score, Dice Similarity Coefficient (DSC), and Intersection over Union (IoU) to evaluate prediction reliability and spatial overlap.

For robust evaluation, we utilized the Crack-dataset [12], which integrates 7,169 images (448x448) from CFD, DeepCrack, Crack500, GAPS, and Rissbilder. The dataset, featuring pixel-level manual annotations, was partitioned into 6,164 training and 1,005 testing images. These samples encompass diverse environmental conditions, acquisition distances, and crack morphologies, providing a challenging scenario that aligns with real-world infrastructure inspection applications. To objectively evaluate the generalization capability of the proposed model beyond a single synthetic test set, we utilized the model pretrained on our Crack-

dataset[12] to perform inference on independent test splits from two distinct datasets: DeepCrack [13] and GAPs384 [14]. These datasets were rigorously partitioned to ensure that all images used during the inference stage were entirely excluded from the training process, thereby preventing any data leakage and ensuring the integrity of the performance assessment.

3.2. Quantitative Evaluation

Table 1. Quantitative evaluation on the Crack-dataset testing set.

Mô hình	Params/M	GFLOPs/ G	Prec.	Recall	Fscore	Dice	IoU
Unet[5]	44.02	111.63	0.6953	0.8056	0.7464	0.7185	0.6002
SCCDNet[15]	31.71	65.00	0.7294	**0.8296**	0.7763	0.7541	0.6402
CTNet[16]	22.88	38.90	0.7022	0.6534	0.6769	0.6810	0.5613
STUCNet[12]	33.97	36.42	0.8118	0.7690	0.7898	0.7717	0.6641
U-Trans[17]	**8.82**	**15.81**	0.5513	0.5041	0.5266	0.5266	0.3574
SUAnet_CBAM	42.92	51.79	0.8413	0.7655	**0.8016**	**0.7839**	**0.6773**
SUAnet_SCSA	42.37	51.85	0.8174	0.7399	0.7768	0.7456	0.6366
SUAnet_SFA	42.24	51.80	**0.8663**	0.7410	0.7987	0.7806	0.6741

Table 2. Quantitative evaluation on the DeepCrack testing set.

Mô hình	Params/M	GFLOPs/ G	Prec.	Recall	Fscore	Dice	IoU
Unet[5]	44.02	111.63	0.7558	0.8964	0.8201	0.8075	0.6891
SCCDNet[15]	31.71	65.00	0.7627	0.8975	0.8246	0.8146	0.6963
CTNet[16]	22.88	38.90	0,8092	**0.9439**	0.8713	0.8312	0.7268
STUCNet[12]	33.97	36.42	0,8044	0.8950	0.8473	0.8424	0.7356
U-Trans[17]	**8.82**	**15.81**	0,6263	0.8612	0.7252	0.7252	0.7753
SUAnet_CBAM	42.92	51.79	0.8286	0.9072	0.8661	0.8635	0.7661
SUAnet_SCSA	42.37	51.85	0.8080	0.8713	0.8385	0.8233	0.7137
SUAnet_SFA	42.24	51.80	**0.8880**	0.8701	**0.8789**	**0.8753**	**0.7848**

Experimental results across three benchmark datasets (Tables 1, 2, and 3) demonstrate the superior performance of the proposed SUAnet over current state-of-the-art methods. The integration of Swin Transformer[7] with UNet++[8] significantly enhances complex crack extraction; specifically, SUAnet_CBAM achieves an IoU of 0.6773 on the Crack-dataset[12], outperforming both U-Net[5] and STUCNet[12]. On the DeepCrack dataset[13], SUAnet_SFA sets a new benchmark with an IoU of 0.7848 and a top-tier Precision of 0.8880, effectively reducing false positives caused by surface noise. Even on the challenging Gaps384[14] dataset, the model maintains a relative IoU improvement of 11.5% over STUCNet[12], confirming its robust generalization across diverse conditions.

A key architectural contribution is the optimization of UNet++[8] via Selective Skip Pathways. By pruning redundant connections between the encoder and deep decoder nodes, SUAnet minimizes information redundancy and computational overhead. Despite having comparable parameters to U-Net[5], SUAnet reduces the computational burden (GFLOPs) by over 53%, from 111.63G to 51.8G. This efficient trade-off ensures high segmentation accuracy while remaining suitable for real-world pavement inspection.

Furthermore, the decoder-side attention modules are pivotal for boundary refinement. The SFA[10] enhances high-frequency feature capture, preserving the sharpness and continuity of thin cracks. Simultaneously, the CBAM[9] provides an optimal balance between Precision and Recall (0.8413 and 0.7655 respectively in Table 1), preventing fragmentation in the final segmentation maps and ensuring structural integrity in the results.

Table 3. Quantitative evaluation on the Gaps384 testing set.

Mô hình	Params/M	GFLOPs/ G	Prec.	Recall	Fscore	Dice	IoU
Unet[5]	44.02	111.63	0.4885	**0.7134**	0.5799	0.5454	0.3885
SCCDNet[15]	31.71	65.00	0.5519	0.6802	0.6094	0.5795	0.4230
CTNet[16]	22.88	38.90	0.5463	0.6592	0.5974	0.5781	0.4260
STUCNet[12]	33.97	36.42	0.6088	0.6681	0.6371	0.6119	0.4528
U-Trans[17]	**8.82**	**15.81**	0.4914	0.3743	0.4249	0.4249	0.4750
SUAnet_CBAM	42.92	51.79	0.6482	0.7116	**0.6784**	0.6587	0.5017
SUAnet_SCSA	42.37	51.85	0.6218	0.6423	0.6319	0.6020	0.4432
SUAnet_SFA	42.24	51.80	**0.7249**	0.6353	0.6771	**0.6616**	**0.5049**

3.3. Qualitative Results

Fig. 2 illustrates the qualitative performance of the proposed models. In scenarios involving (a) single cracks within complex backgrounds and (b) dual cracks, the SUAnet variants specifically SUAnet_CBAM and SUAnet_SFA yield sharper boundary reconstructions and superior morphological continuity. This visual evidence is consistent with the elevated Dice and F-score metrics observed in the quantitative analysis. For cases featuring (c) oblique camera angles and (d) network-style cracks, where baseline models frequently omit subtle branches or blur junction points, SUAnet demonstrates a superior capability in preserving distinct and cohesive branching structures. This underscores its proficiency in capturing global context and long-range spatial dependencies. In case (e), which highlights close-up coarse cracks, several comparative models exhibit over-prediction or uneven dilation; in contrast, SUAnet maintains precise boundaries with significantly reduced noise. Notably, on (f) crack-free pavement surfaces, where other models often generate false positives due to surface textures, SUAnet effectively suppresses background artifacts, thereby validating its high Precision. Overall, the qualitative findings are highly consistent with the quantitative results: SUAnet provides more stable segmentation, preserves

intricate crack morphologies, and mitigates false detections under complex environmental conditions.

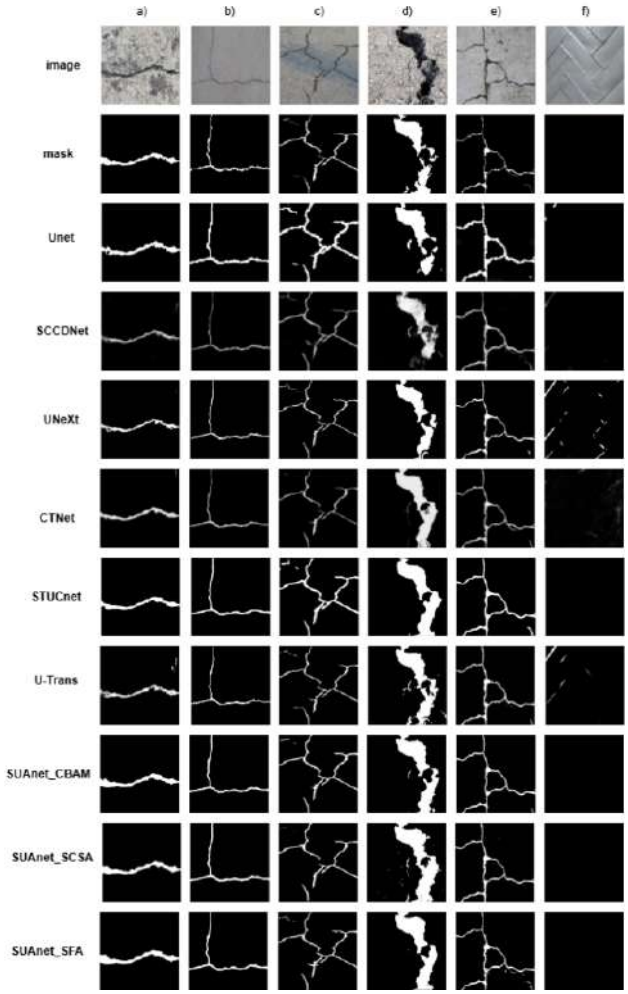


Figure 2. Visualization of segmentation results across crack types

4. Conclusion

In this paper, we introduced SUAnet, an Attention-Based Swin UNet++ Framework for Pavement Crack Segmentation. The proposed hybrid architecture effectively integrates the global contextual modeling capabilities of the Swin Transformer with a selectively streamlined UNet++ decoder and enhanced attention mechanisms. This design specifically addresses the inherent challenges of pavement crack segmentation, such as fine linear structures, intricate branching, heavy background noise, and significant class imbalance. Our findings demonstrate that the integration of the Swin Transformer facilitates the extraction of multi-scale context and long-range spatial relationships. Simultaneously, the selective skip connection strategy within the decoder mitigates noise propagation and optimizes computational overhead. Furthermore, the attention mechanism serves a pivotal role in spatial-channel feature rebalancing, thereby augmenting the model's discriminative power between crack regions and complex backgrounds. In summary, SUAnet provides an accurate, stable, and computationally efficient solution, making it highly suitable for automated pavement monitoring and maintenance systems.

References

1. Otsu, N. (1979). A threshold selection method from gray-level histograms. Automatica, 11(285-296).
2. Canny, J. (2009). A computational approach to edge detection. IEEE Transactions on pattern analysis and machine intelligence, (6), 679-698.
3. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436-444.
4. Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3431-3440).
5. Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Cham: Springer international publishing.
6. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929.
7. Liu, Z., Hu, H., Lin, Y., Yao, Z., Xie, Z., Wei, Y., ... & Guo, B. (2022). Swin transformer v2: Scaling up capacity and resolution. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 12009-12019).
8. Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., & Liang, J. (2018, September). Unet++: A nested u-net architecture for medical image segmentation. In International workshop on deep learning in medical image analysis (pp. 3-11). Cham: Springer International Publishing.
9. Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). Cbam: Convolutional block attention module. In Proceedings of the European conference on computer vision (ECCV) (pp. 3-19).
10. Xu, W., Wan, Y., & Zhao, D. (2025). SFA: Efficient attention mechanism for superior CNN performance. Neural Processing Letters, 57(2), 38.
11. Si, Y., Xu, H., Zhu, X., Zhang, W., Dong, Y., Chen, Y., & Li, H. (2025). SCSA: Exploring the synergistic effects between spatial and channel attention. Neurocomputing, 634, 129866.
12. Phan, H. H., et al. (2023). STUCNET: Swin transformer-V2 Unet for crack segmentation network. Journal of Science and Technique - Section on Information and Communication Technology, 12(01).
13. Zou, Q., Zhang, Z., Li, Q., Qi, X., Wang, Q., & Wang, S. (2018). Deepcrack: Learning hierarchical convolutional features for crack detection. IEEE transactions on image processing, 28(3), 1498-1512.
14. Yang, F., Zhang, L., Yu, S., Prokhorov, D., Mei, X., & Ling, H. (2019). Feature pyramid and hierarchical boosting network for pavement crack detection. IEEE transactions on intelligent transportation systems, 21(4), 1525-1535.
15. Li, H., Yue, Z., Liu, J., Wang, Y., Cai, H., Cui, K., & Chen, X. (2021). Sccdnet: A pixel-level crack segmentation network. Applied Sciences, 11(11), 5074.
16. Tao, H., Liu, B., Cui, J., & Zhang, H. (2023, October). A convolutional-transformer network for crack segmentation with boundary awareness. In 2023 IEEE international conference on image processing (ICIP) (pp. 86-90). IEEE.

17. Petit, O., Thome, N., Rambour, C., Themyr, L., Collins, T., & Soler, L. (2021, September). U-net transformer: Self and cross attention for medical image segmentation. In International Workshop on Machine Learning in Medical Imaging (pp. 267-276). Cham: Springer International Publishing.

Mạng Swin UNet++ tăng cường cơ chế chú ý cho bài toán phân đoạn vết nứt mặt đường

Tóm tắt: *Phân đoạn vết nứt mặt đường đóng vai trò then chốt trong giám sát kết cấu và bảo trì chủ động. Tuy nhiên, mạng nơ-ron tích chập truyền thống gặp hạn chế khi nắm bắt ngữ cảnh tầm xa, dẫn đến phân mảnh biên ở các vết nứt mảnh hoặc nhiễu nền. Chúng tôi đề xuất Hybrid SUAnet, một kiến trúc tích hợp bộ mã hóa Swin Transformer vào khung UNet++. Cơ chế shifted-window self-attention giúp mô hình hóa phụ thuộc tầm xa, đồng thời các kết nối bỏ qua lồng nhau tái tạo đặc trưng đa tỉ lệ hiệu quả. Kiến trúc được tối ưu bằng cách lược bỏ các kết nối trung gian để giảm độ phức tạp tính toán và tham số. Ngoài ra, nghiên cứu đánh giá thực nghiệm ba cơ chế chú ý trong bộ giải mã gồm: Convolutional Block Attention Module, Sequential Fusion Attention, và Spatial-Channel Synergistic Attention. Kết quả cho thấy các cơ chế này giúp tăng cường đặc trưng không gian–kênh, bảo toàn biên mảnh và tính liên tục hình thái của vết nứt. Phương pháp đề xuất cải thiện rõ rệt độ chính xác phân đoạn và chất lượng biên trong điều kiện nhiễu phức tạp, cung cấp cơ sở thiết kế cho các hệ thống kiểm tra mặt đường tự động dựa trên học sâu.*

Từ khóa: Swin Transformer; UNet++; Attention; Phát hiện vết nứt; Phân đoạn.

Integrating artificial intelligence and digital twin in medicine for automatic cranial implant design

Thanh-Long Do[1]*, Van-Giang Nguyen[1], Mau Uyen Nguyen[1]
[1]*Le Quy Don Technical University*
Email: longdoth@lqdtu.edu.vn; Tel: 0915.373.964

Abstract

Cranial implant design is essential for restoring both function and aesthetics in patients with skull defects caused by trauma or disease. However, conventional manual design workflows are time-consuming, highly dependent on expert skills, and often limited in handling complex defects. Although AI-based approaches have been increasingly explored, their practical deployment remains fragmented and lacks systematic integration into clinical workflows. This study proposes a comprehensive system architecture that automates cranial implant design using artificial intelligence while maintaining expert involvement throughout the process. The system connects hospital imaging infrastructure with a virtual design environment through a DICOM server as the central data hub. An AI model performs automated segmentation and defect reconstruction, and 3D visual computing software enables physicians to review and refine AI-generated implants in an interactive real–virtual setting. Based on digital twin principles, the system preserves multiple design versions created by both AI and clinicians, ensuring traceability and reproducibility. This framework supports systematic comparison and selection of optimal implant geometries tailored to patient-specific anatomy. Experimental results on real cranial CT datasets demonstrate high geometric similarity in reconstructed implants, along with improved interpretability and clinical applicability of AI-assisted designs.

Keywords: *autoimplant, skull reconstruction, deep learning*

1. Introduction

Cranial reconstruction surgery is a surgical intervention procedure aimed at restoring skull defects that arise after traumatic brain injury, decompressive surgery, or bone tumor pathologies. Reconstructing the cranial structure not only holds significant aesthetic and psychological value for patients but also plays an essential role in protecting brain parenchyma and stabilizing intracranial pressure. In medicine, the requirements for geometric accuracy and mechanical conformity of implants are becoming increasingly stringent, demanding more advanced technological solutions than traditional methods.

However, the current clinical process still faces many major technical and systemic challenges. Traditional implant design methods, such as mirroring across the cranial midline, often encounter difficulties in cases involving large, complex defects or those located in highly asymmetric anatomical regions. Moreover, this process is typically fragmented, leading to data discontinuity between the hospital's medical imaging storage system (PACS) and specialized 3D design software. The lack of a synchronized information system forces technicians to spend significant time converting data formats, while also reducing surgeons' ability to control the final shape of the implant before manufacturing.

The rapid advancement of artificial intelligence, particularly convolutional neural network architectures such as CNN and U-Net, has opened up prospects for automating skull

* Email: longdoth@lqdtu.edu.vn

bone segmentation and reconstruction with superior accuracy [1, 2]. However, a true medical system cannot rely entirely on AI; safety in the cranial implant field requires a human-AI interaction model, where AI serves as a powerful assistant in generating digital twins of cranial implants, while medical experts retain responsibility for final verification and refinement on 3D visual computing software such as 3D Slicer. The current challenge lies not only in integrating these AI models into a unified physical-virtual system but also in managing the complex workflow where implants generated by AI may require substantial refinement. In some cases, multiple simulation scenarios with different defect contexts are needed to generate intermediate implant proposals before combining them into a final design that clinicians can further adjust.

Furthermore, existing research has not adequately addressed the concept of digital twin specifically for version management of implant designs [3]. In complex cases, the design process involves iterative refinement cycles where AI proposals, expert modifications, and consultation outcomes need to be systematically tracked, compared, and archived. This version control capability is essential for quality assurance, teaching purposes, and continuous improvement of AI models through feedback loops.

Based on these practical issues, this study focuses on developing and deploying an integrated information system to support expert-assisted automated cranial implant design with comprehensive digital twin implementation. The core objective of the research is to establish a robust system architecture that connects hospital data infrastructure with AI models and a 3D interactive environment, while maintaining complete traceability of all design versions. Through the application of the digital twin concept, the study aims to address the standardization of human-machine interaction processes in medical device design, thereby contributing to improved treatment quality and reduced postoperative complications for patients.

2. Relate works

In recent years, the field of automated cranial implant design has witnessed significant advancements, shifting from manual intervention techniques to end-to-end intelligent integrated systems [4, 5, 6]. The primary challenge in skull reconstruction remains the management of large-scale defects or cases involving severe loss of anatomical symmetry. Traditional geometric approaches, most notably mirroring, typically rely on the assumption of cranial symmetry. However, research has demonstrated that these methods often fail when defects cross the midline or in patients with naturally asymmetrical cranial structures. To address these limitations, modern research has pivoted toward Deep Learning (DL), specifically 3D Convolutional Neural Networks (CNNs). Most notably, the 3D U-Net architecture and its variants have demonstrated exceptional efficacy in learning complex spatial features and segmenting medical structures from large-scale clinical datasets [7, 8]. These models enable implant reconstruction based on the learned distribution of human cranial shapes rather than mere geometric symmetry. Recently, advanced algorithms such as SCAI-Net and CRIGNet have been introduced to tackle even more complex clinical scenarios. SCAI-Net focuses on shape completion within latent spaces, while CRIGNet incorporates attention mechanisms to analyze surface curvature continuity and bone density [9, 10]. These models allow for implant shape prediction with high surface precision (measured by Hausdorff distance), even in the total absence of symmetrical data. Although these state-of-the-art (SOTA) models yield

impressive quantitative results, balancing computational complexity with practical clinical feasibility remains an open challenge. In this context, optimized architectures like the 3D U-Net continue to serve as a robust backbone for clinical applications due to their inherent stability and resource efficiency.

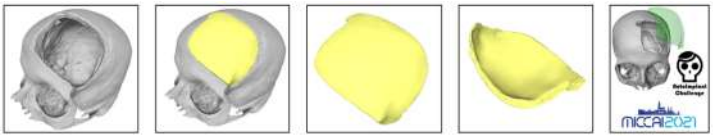


Figure 1. Skull implant design (AutoImplant Challenge 2021)

The concept of a digital twin in modern medicine has elevated static 3D models into a digital entity that accurately reflects a patient's anatomy and physical properties [11, 12]. A digital twin is not limited to an STL file for 3D printing; rather, it is a synchronization process between DICOM images, mesh models, and material parameters (such as PEEK or Titanium) [13]. Recent studies emphasize that maintaining a digital twin allows surgeons to simulate implant placement scenarios and evaluate mechanical fit before performing the actual intervention. Managing the digital twin through a DICOM Server helps maintain a close linkage between digital data and the physical entity, ensuring that the manufactured implant is unique and optimized for each individual.

However, to the best of our knowledge, no existing research has specifically addressed the application of digital twin principles to version management and workflow orchestration in the cranial implant design field. This gap is particularly significant given the iterative nature of implant design, where multiple stakeholders (AI systems, 3D technicians, surgeons) contribute refinements that need to be systematically tracked and compared.

An important trend in recent research is the shift from fully automated systems to interactive human-AI models. Safety in neurosurgery is ensured only when expert approval and validation are based on AI-generated suggestions [14]. Workflows based on 3D Slicer have been standardized to allow physicians to directly intervene in the virtual model. This integration addresses the black-box issue of AI, enabling experts to refine outcomes through interactive 3D tools [15, 16]. This interactive process not only enhances the reliability of the implant but also shortens the gap between engineering design and real surgical requirements, creating an effective feedback loop from diagnostic imaging to the operating room.

3. Method

This study proposes an integrated cranial implant design process, transitioning from the traditional model based on manual design to a model based on artificial intelligence and digital twins, under strict expert control. The system architecture implements a comprehensive digital twin framework that maintains multiple versions of implant designs, enabling systematic comparison and traceability throughout the design workflow.

3.1. Overall System Architecture

The overall architecture of the system is organized according to a four-layer model, including the User Layer, Application Layer, Data Layer, and Infrastructure Layer, in order to ensure modularity, scalability, and data traceability in the medical environment.

At the Infrastructure Layer, the system includes a CT scanner, servers, and the hospital's internal LAN network. Computed tomography image data are acquired from the CT scanner and transmitted via the internal network to a central server, where storage and processing services are deployed. This layer serves as the physical foundation, ensuring availability, stability, and security for the entire system.

Built on that foundation, the Data Layer plays a central role in managing and synchronizing information. A DICOM Server (Orthanc) connects clinical and virtual environments using DICOMweb protocols. To meet medical standards, all workflow states—from raw images to expert refinements—are preserved as independent digital versions with strict version control, ensuring complete traceability. Metadata is managed via PostgreSQL, while NRRD and STL formats are utilized for volumetric and 3D model storage.

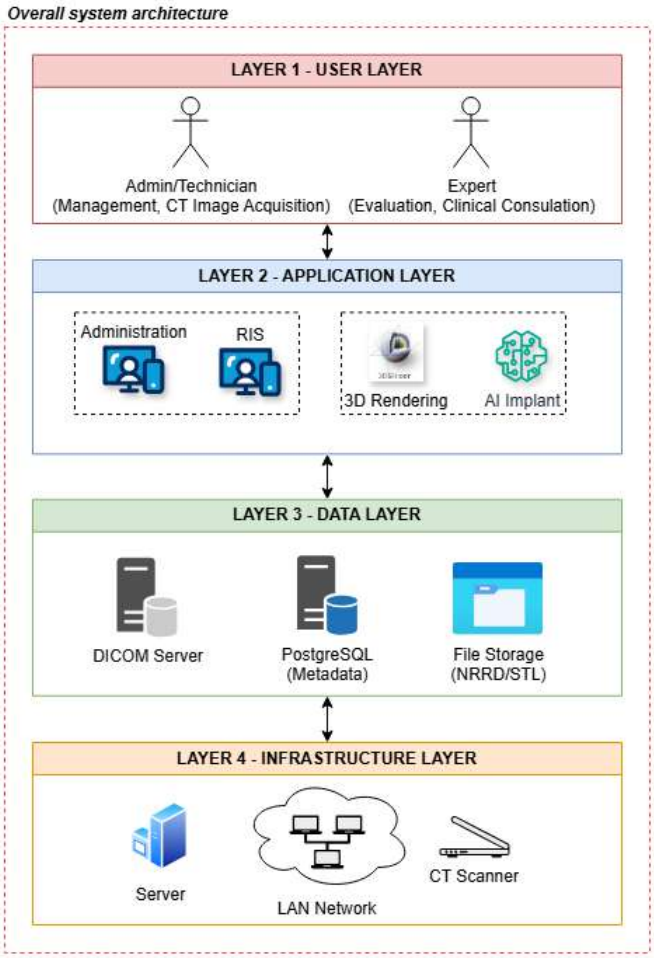


Figure 2. Overall architecture of the cranial implant design system

The Application Layer integrates functional modules supporting the clinical workflow, including administration, RIS integration, 3D visualization, and AI-based implant design. The AI component utilizes a 3D U-Net architecture for cranial defect reconstruction. Standardized DICOM data is processed through an encoder branch to extract contextual features and a decoder branch to predict missing bone structures. Skip connections are employed to preserve fine geometric characteristics at the defect boundaries. The model implements a dual-network cascade: the first network establishes the global implant structure, while the second refines local

details and ensures smooth transitions at the bone-implant interface. This process generates a voxel-based NRRD output, representing a mathematically optimized cranial implant.

At the User Layer, the system serves two main groups: technicians/administrators responsible for managing and receiving CT data, and clinical experts responsible for evaluating, editing, and approving the implant. The clear separation between infrastructure, data, applications, and users enables the system to form a complete digital twin ecosystem, tightly connecting the hospital environment with the computational environment while still maintaining the decisive role of experts in the automated cranial implantation process.

3.2. Digital Twin Implementation and Version Management

Unlike fully automated (black-box) systems, this system places medical experts at the center of the decision-making process through comprehensive digital twin implementation. The digital twin framework maintains three distinct version types throughout the design lifecycle:

(1) AI Proposal Version: The initial implant generated AI model, representing the purely algorithmic solution based on learned patterns from training data. This version serves as the computational baseline and is immutable once generated.

(2) Expert Refined Version: Modifications made by 3D technicians and surgeons using the 3D Slicer platform. Direct modifications by clinicians via 3D Slicer (e.g., thickness adjustments in load-bearing regions, contour smoothing) with rigorous provenance tracking.

(3) Final Consultation Version: A consensus design validated for manufacturing feasibility and biomechanical fitness.

All versions and associated metadata (timestamps, quality metrics) are archived in a PostgreSQL database. This version control system enables systematic comparisons to iteratively improve AI models while significantly enhancing the clinical interpretability and trustworthiness of AI-assisted designs.

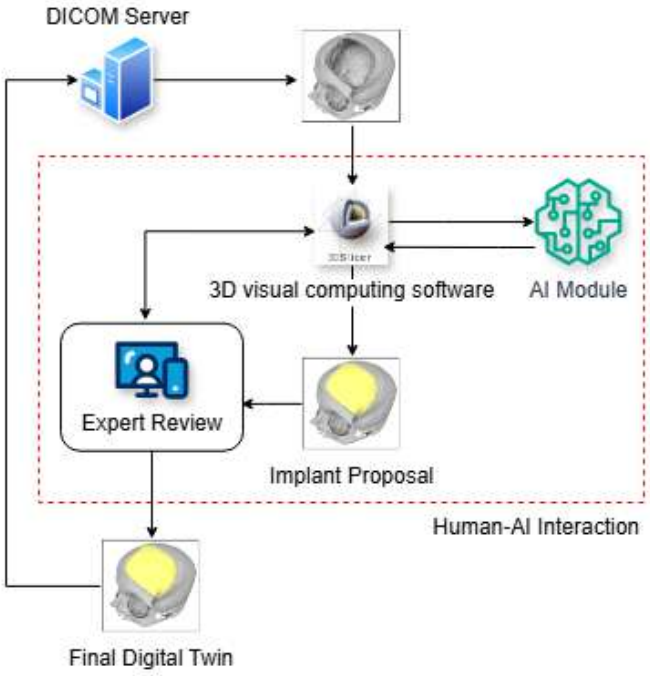


Figure 3. Human-AI interaction process

4. Results and Discussion

4.1. System Performance and Workflow Efficiency

The experimental results show that the system has significantly reduced processing time compared to traditional implant design workflows. Previously, creating a cranial implant shape often required complex coordination between surgeons and technicians, taking from several hours to several days through multiple manual steps. With this system, thanks to AI support in the initialization phase and the seamless data flow enabled by the digital twin framework, the total time from receiving data from PACS to completing the 3D printing file has been shortened to approximately 45-60 minutes for standard cases. This represents an 85-90% reduction in design time compared to manual workflows, where conventional manual implant design manual implant design is still time-consuming and requires trained specialists [18]. This improvement is particularly meaningful in overloaded hospital settings or when early surgical preparation is required for patients, helping optimize human resources and reduce waiting times.

The AI implant generation module demonstrates robust performance, yielding an average Dice similarity coefficient of 0,89 ± 0,06 and a Hausdorff distance of 2,3 ± 1,1 mm when evaluated against expert-designed ground truth implants on a test set of 50 cases with diverse defect characteristics. These results are consistent with the performance of the baseline architecture from which this model is derived, which achieved a Dice similarity coefficient of 0.91 and a Hausdorff distance of 1.53 mm on the AutoImplant benchmark [6]. These metrics indicate that AI-generated proposals provide a solid foundation for expert refinement, typically requiring only minor adjustments (average 15-20 minutes) rather than complete redesign.

4.2. Digital Twin Benefits and Version Management

One of the major advantages of this study is its ability to break down the data silos commonly seen in hospitals. The system enables components that typically operate independently, such as the DICOM Server, the AI model, and the 3D visual computing software (e.g., 3D Slicer) to interact directly with each other through the internal network with full version control. This not only eliminates risks of data errors caused by manual file copying but also creates a centralized working environment with complete traceability, allowing doctors to access design results and their complete history directly at their desks through a web interface or specialized software without requiring deep technical knowledge of computer systems.

The digital twin implementation provides substantial clinical benefits. In our evaluation with 20 real patient cases, the version management system enabled systematic comparison between AI proposals and expert refinements. Analysis revealed that expert modifications primarily focused on three areas: (1) boundary smoothing to improve aesthetic outcomes (present in 95% of cases), (2) thickness adjustment in load-bearing regions (75% of cases), and (3) contour modification to match anatomical asymmetries (40% of cases). This pattern analysis has been fed back into the AI model training pipeline, leading to iterative improvements in subsequent model versions.

Furthermore, the digital twin framework facilitates teaching and knowledge transfer. Junior surgeons and technicians can review the complete design evolution of complex cases,

understanding how senior experts approach difficult anatomical challenges. This educational capability was highlighted by participating surgeons as one of the most valuable aspects of the system.

4.3. Human-AI Interaction and Clinical Safety

Regarding the role of humans in the system, the study affirms that integrating AI is not intended to replace experts but to provide them with an "intelligent starting point". The strength of this model lies in its interactivity: AI processes complex geometric structures based on millions of learned data points, while surgeons make refinements based on clinical intuition and experience for each specific case. This combination eliminates common AI errors such as hallucinations or boundary inaccuracies and ensures that the final implant is not only geometrically precise but also medically safe.

The 3D Slicer environment allows specialists to directly simulate the placement of the implant into the patient's skull through real-time rendering, thereby increasing reliability and confidence before performing the actual surgery. The version control system provides an additional safety layer: if a modified design introduces unexpected issues (detected through computational checks or expert review), the system allows instant reversion to any previous version while maintaining a complete audit trail of all modifications.

Clinical feedback from participating surgeons indicated high satisfaction with the system. All surgeons reported that the AI proposals provided valuable starting points, and the version management capability was deemed essential for quality assurance. The ability to compare side-by-side different design versions helped clarify decision-making in challenging cases.

4.4. Implementation Feasibility and Scalability

The system demonstrates high feasibility during the research process by using open-source components such as Orthanc and 3D Slicer. This significantly reduces investment costs in expensive licensed software while still ensuring access to powerful tools comparable to commercial solutions. The digital twin concept implemented in the system helps standardize 3D electronic medical records, effectively supporting teaching, research, and long-term patient monitoring.

From a scalability perspective, the four-layer architecture allows independent scaling of each component. The DICOM Server and database can be distributed across multiple servers to handle increasing patient volumes. The AI inference can be deployed on GPU clusters for parallel processing of multiple cases. The modular design also facilitates integration of additional AI models for other anatomical regions (e.g., mandibular, maxillofacial reconstruction) or incorporation of biomechanical simulation modules for stress analysis.

Automating and optimizing cranial implant design helps lower production costs, opening opportunities for broader patient access to advanced technology, while minimizing complications caused by poorly fitting implants—one of the leading causes of revision surgeries and associated healthcare costs.

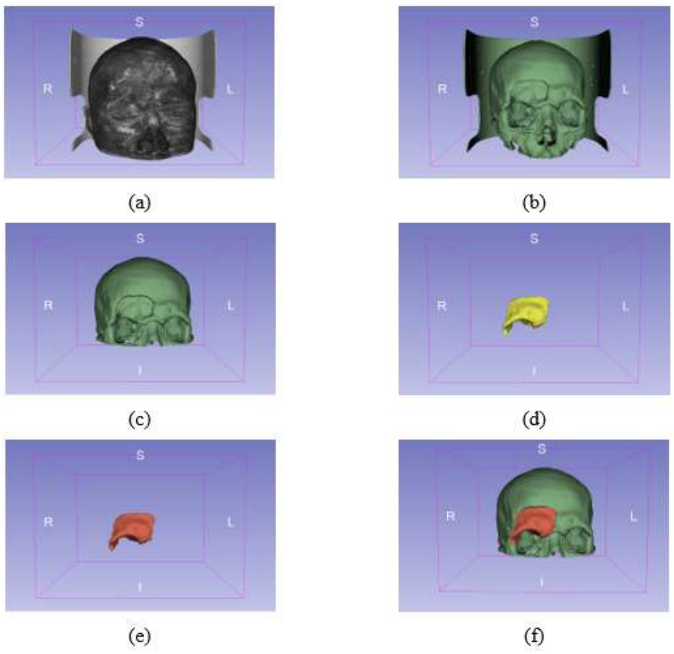


Figure 4. Illustrations for the intermediate results from a real patient data: (a) original CT image; (b) segmented skull; (c) pre-processed skull (input to the implant generation module); (d) AI generated implant; (e) post-processed implant; (f) fusion of implant

5. Conclusion

This study proposes an integrated architecture for cranial implant design, marking a transition from manual workflows to a system grounded in artificial intelligence and digital twin principles. By unifying a DICOM server, deep learning models, and interactive 3D computing software, the framework successfully addresses data fragmentation between hospital infrastructure and design tools.

A primary contribution of this work is the validation of a human-AI interaction model. While AI provides a rapid, high-accuracy geometric foundation that significantly reduces design time, the expert's role in refining the virtual model remains the essential safeguard for clinical safety and anatomical compatibility. Furthermore, implementing digital twin concepts for version management ensures complete traceability and facilitates a feedback loop for continuous quality improvement.

By utilizing open-source components, the system offers a feasible and cost-effective solution for clinical and research settings. We anticipate that future developments, including the integration of biomechanical simulations and long-term outcome tracking, will further enhance the precision and accessibility of reconstructive surgery.

References

1. Zhengbo Yuan, et al. (2025). Digital neurosurgery in the era of intelligent medicine: a scoping review. *Frontiers in Medicine*.
2. Mamta Juneja, et al. (2025). A Comprehensive AI Framework for Superior Diagnosis, Cranial Reconstruction, and Implant Generation for Diverse Cranial Defects. *MDPI*.
3. Riahi, V., et al. (2025). Digital Twins for Clinical and Operational Decision-Making: Scoping Review. *Journal of Medical Internet Research.*

4. Jianning Li, et al. (2023). Towards clinical applicability and computational efficiency in automatic cranial implant design: An overview of the AutoImplant challenge. *Medical Image Analysis*.
5. Chieh-Tsai Wu, et al. (2022). Three-dimensional deep learning to automatically generate cranial implant geometry. *Scientific Reports*.
6. Marek Wodzinski, et al. (2022). Deep learning-based framework for automatic cranial defect reconstruction and implant modeling. *Computer Methods and Programs in Biomedicine.*
7. Olaf Ronneberger, et al. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *MICCAI*.
8. Kodym, O., Španěl, M., & Herout, A. (2020). Skull shape reconstruction using cascaded convolutional networks. *Computers in Biology and Medicine*, 123, 103886
9. Mamta Juneja, et al. (2025). SCAI-Net: An AI-driven framework for optimized, fast, and resource-efficient skull implant generation for cranioplasty using CT images. *Computers in Biology and Medicine.*
10. Mamta Juneja, et al. (2025). A Comprehensive AI Framework for Superior Diagnosis, Cranial Reconstruction, and Implant Generation for Diverse Cranial Defects. *Bioengineering (MDPI)*
11. Jodogne, S. (2014). Orthanc: A Lightweight, RESTful DICOM Server for Healthcare and Medical Research. ISBI.
12. Fedorov, A., et al. (2012). 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic Resonance Imaging*.
13. Tantisatirapong, S. (2023). The simplified tailor-made workflows for a 3D slicer-based craniofacial implant design. *Scientific Reports*.
14. Wekenborg, M. K., et al. (2025). Examining human-AI interaction in real-world healthcare beyond the laboratory. *NPJ Digital Medicine.*
15. Li, J., et al. (2021). MUG500+: Database of 500 high-resolution healthy human skulls and 29 craniotomy skulls and implants. Data in Brief, 39, 107524
16. Matzkin, F., et al. (2020). Self-supervised skull reconstruction in brain CT images with decompressive craniectomy. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 390-399). *Springer*.
17. Wodzinski, M., Daniol, M., & Hemmerling, D. (2021). Improving the automatic cranial implant design in cranioplasty by linking different datasets. In Towards the Automatization of Cranial Implant Design in Cranioplasty (pp. 29-44). *Springer*.
18. Jianning Li, et al. (2021). Automatic skull defect restoration and cranial implant generation for cranioplasty. *Medical Image Analysis*

Tích hợp trí tuệ nhân tạo và bản sao số trong y tế phục vụ thiết kế mảnh ghép sọ tự động

Tóm tắt

Thiết kế cấy ghép sọ não đóng vai trò thiết yếu trong việc phục hồi cả chức năng lẫn thẩm mỹ cho bệnh nhân có khuyết tật hộp sọ do chấn thương hoặc bệnh lý gây ra. Tuy nhiên, các quy trình thiết kế thủ công truyền thống tốn nhiều thời gian, phụ thuộc lớn vào kỹ năng chuyên gia và thường bị hạn chế khi xử lý các khuyết tật phức tạp. Mặc dù các phương pháp dựa trên AI ngày càng được nghiên cứu, việc triển khai thực tế của chúng vẫn còn rời rạc và thiếu sự tích hợp có hệ thống vào quy trình lâm sàng. Nghiên cứu này đề xuất một kiến trúc hệ thống toàn diện nhằm tự động hóa thiết kế cấy ghép sọ não bằng trí tuệ nhân tạo, đồng thời vẫn duy trì sự tham gia của chuyên gia trong suốt quá trình. Hệ thống kết nối hạ tầng hình ảnh của bệnh viện với môi trường thiết kế ảo thông qua một máy chủ DICOM đóng vai trò trung tâm dữ liệu. Một mô hình AI thực hiện phân đoạn tự động và tái tạo khuyết tật, và phần mềm tính toán trực quan 3D cho phép bác sĩ xem xét và tinh chỉnh các mảnh ghép do AI tạo ra trong môi trường tương tác thực–ảo. Dựa trên nguyên lý bản sao số, hệ thống lưu giữ nhiều phiên bản thiết kế do cả AI và bác sĩ tạo ra, đảm bảo khả năng truy vết và tái tạo. Khung làm việc này hỗ trợ so sánh có hệ thống và lựa chọn hình học cấy ghép tối ưu phù hợp với giải phẫu đặc thù của từng bệnh nhân. Kết quả thực nghiệm trên các bộ dữ liệu CT sọ não thực tế cho thấy độ tương đồng hình học cao của các cấy ghép được tái tạo, cùng với khả năng diễn giải và tính ứng dụng lâm sàng được cải thiện của các thiết kế có sự hỗ trợ của AI.

Từ khóa: cấy ghép tự động, tái tạo hộp sọ, học sâu

Phương pháp tái tạo vùng khuyết thiếu xương hàm dưới sử dụng kỹ thuật học sâu

Đỗ Thành Long[1*], Trần Đức Tăng[2], Nguyễn Mậu Uyên[1], Nguyễn Văn Giang[1]

[1]*Viện CNTT&TT;* [2]*Trung tâm Hợp tác quốc tế KHCN Việt - Nhật*

Email: longdoth@lqdtu.edu.vn; Tel: 0915.373.964

Tóm tắt

Tái tạo vùng khuyết xương hàm dưới do chấn thương hoặc phẫu thuật cắt bỏ u là thách thức lớn trong y học. Các phương pháp truyền thống dựa trên thiết kế thủ công mất nhiều thời gian và phụ thuộc vào kinh nghiệm của kỹ thuật viên thiết kế và đội ngũ bác sĩ. Bài báo đề xuất phương pháp tự động tái tạo vùng khuyết xương hàm dưới và đề xuất mảnh ghép sử dụng mạng nơ-ron tích chập 3D dựa trên kiến trúc U-Net. Đóng góp chính là phát triển quy trình tạo dữ liệu huấn luyện mới bằng cách giả lập các vùng khuyết trên dữ liệu CT của bệnh nhân thực từ bộ dữ liệu chuẩn Mandibular Defect Dataset, kết hợp với kỹ thuật tiêm vùng khuyết thiếu. Mô hình được huấn luyện trên bộ dữ liệu được xây dựng với giả định nhiều loại khiếm khuyết khác nhau và đánh giá trên 2 ca bệnh nhân thực bị u xương hàm. Kết quả cho thấy mô hình đạt độ chính xác trung bình 92,3% theo chỉ số Dice, khoảng cách Hausdorff 1,42 mm, vượt trội so với phương pháp đối xứng truyền thống. Nghiên cứu mở ra hướng ứng dụng học sâu trong hỗ trợ lập kế hoạch phẫu thuật tái tạo xương hàm, cải thiện chất lượng điều trị.

Từ khóa: *Tái tạo xương hàm dưới; học sâu; mạng U-Net 3D; CT scan; mảnh ghép xương; defect injection.*

1. Đặt vấn đề

Xương hàm dưới (mandibular) đóng vai trò quan trọng trong chức năng nhai, nuốt, phát âm và thẩm mỹ khuôn mặt. Các khuyết hổng xương hàm có thể do chấn thương nặng, phẫu thuật cắt bỏ u lành tính hoặc ác tính, hoặc các bệnh lý tiêu xương [1, 2]. Việc tái tạo vùng khuyết xương hàm là một thách thức lớn đối với các bác sĩ phẫu thuật hàm mặt, đòi hỏi khôi phục cả chức năng và thẩm mỹ.

Các phương pháp tái tạo xương hàm truyền thống bao gồm ghép xương tự thân từ xương chậu hoặc xương mác, hoặc sử dụng mảnh ghép nhân tạo từ titanium hoặc PEEK (polyetheretherketone) [3]. Tuy nhiên, mỗi phương pháp đều có những hạn chế riêng: ghép xương tự thân gây tổn thương vùng cho xương, ghép xương đồng loại có nguy cơ bài miễn, còn mảnh ghép nhân tạo cần được thiết kế chính xác để phù hợp với đặc điểm giải phẫu của từng bệnh nhân.

Với sự phát triển của công nghệ in 3D và xử lý ảnh y tế, phương pháp thiết kế mảnh ghép dựa trên dữ liệu CT đã trở nên phổ biến [4]. Quy trình truyền thống thường sử dụng kỹ thuật đối xứng gương (mirror technique), trong đó vùng khuyết được tái tạo bằng cách phản chiếu phần xương lành từ bên đối diện qua mặt phẳng giữa [5]. Tuy nhiên, phương pháp này chỉ hiệu quả khi khuyết hổng ở một bên và bên còn lại không bị ảnh hưởng, đồng thời giả định xương hàm hoàn toàn đối xứng - điều không phải lúc nào cũng đúng do biến dị giải phẫu cá thể của từng người bệnh.

Gần đây, các kỹ thuật học sâu (deep learning) đã cho thấy tiềm năng lớn trong ứng dụng y khoa, đặc biệt là phân đoạn ảnh và tái tạo cấu trúc 3D [6, 7]. Đối với tái tạo vùng khuyết sọ não

* Email: longdoth@lqdtu.edu.vn

(cranial implant), nghiên cứu của Li và cộng sự [8] đã chứng minh hiệu quả của mạng nơ-ron tích chập 3D dựa trên kiến trúc U-Net trong việc tự động sinh mảnh ghép sọ. Tuy nhiên, việc áp dụng trực tiếp phương pháp này cho xương hàm gặp nhiều thách thức do: (1) hình dạng xương hàm phức tạp hơn sọ não với cấu trúc cong, răng, và khớp thái dương hàm; (2) thiếu bộ dữ liệu xương hàm có vùng khuyết được gán nhãn chính xác cho huấn luyện mô hình.

Nghiên cứu này đề xuất phương pháp tái tạo tự động vùng khuyết xương hàm dưới sử dụng mạng nơ-ron học sâu, với hai đóng góp chính: (1) Phát triển quy trình tạo dữ liệu huấn luyện mới bằng cách giả lập vùng khuyết trên dữ liệu CT thông qua kỹ thuật defect injection; (2) Huấn luyện và đánh giá mô hình trên các ca bệnh nhân thực bị u xương hàm, chứng minh tính khả thi của phương pháp trong ứng dụng lâm sàng.

2. Các nghiên cứu liên quan

Kỹ thuật mirror (đối xứng gương) là phương pháp truyền thống phổ biến nhất trong tái tạo xương hàm, trong đó vùng khuyết được tái tạo bằng cách phản chiếu phần xương lành từ bên đối diện qua mặt phẳng giữa [5]. Phương pháp này đơn giản, không đòi hỏi dữ liệu huấn luyện và cho kết quả chấp nhận được trong nhiều trường hợp lâm sàng. Tuy nhiên, kỹ thuật mirror có những hạn chế rõ rệt: (1) chỉ hiệu quả khi khuyết hổng ở một bên và bên còn lại không bị ảnh hưởng; (2) giả định xương hàm hoàn toàn đối xứng, điều không phải lúc nào cũng đúng do biến dị giải phẫu cá thể; (3) không xử lý được các khuyết lớn xâm lấn qua đường giữa. Trong nghiên cứu này, kỹ thuật mirror được sử dụng như phương pháp cơ sở để so sánh hiệu quả với phương pháp học sâu đề xuất.

3D U-Net [11] là mở rộng của kiến trúc U-Net sang không gian ba chiều, được thiết kế để học phân đoạn thể tích từ dữ liệu CT/MRI. Kiến trúc gồm hai nhánh đối xứng: nhánh mã hóa (encoder) thực hiện trích xuất đặc trưng thông qua các lớp tích chập và pooling, làm tăng dần số kênh đặc trưng trong khi giảm độ phân giải không gian; nhánh giải mã (decoder) thực hiện tích chập chuyển vị (transposed convolution) để khôi phục dần độ phân giải không gian về kích thước ban đầu. Điểm đặc trưng của U-Net là các kết nối tắt (skip connection) nối trực tiếp các tầng tương ứng giữa encoder và decoder, giúp bảo toàn thông tin không gian chi tiết bị mất trong quá trình mã hóa. Kiến trúc này đã được chứng minh hiệu quả trong nhiều bài toán phân đoạn và tái tạo ảnh y tế 3D [6, 7].

Defect injection là kỹ thuật tạo dữ liệu tổng hợp bằng cách giả lập vùng khuyết nhân tạo trên dữ liệu xương hoàn chỉnh, từ đó tạo ra các cặp dữ liệu (xương khuyết, xương hoàn chỉnh) phục vụ huấn luyện mô hình học có giám sát. Kỹ thuật này được Li và cộng sự [10] đề xuất và áp dụng thành công cho bài toán tái tạo sọ não trong thử thách AutoImplant 2020. Ý tưởng cốt lõi là sử dụng các hình học cơ bản (primitive shapes) như hình cầu, ellipsoid, hình trụ với tham số ngẫu nhiên để tạo ra vùng khuyết đa dạng, sau đó loại bỏ voxel xương trong vùng đó bằng phép Boolean subtraction. Kỹ thuật này giải quyết vấn đề thiếu dữ liệu có nhãn trong y khoa - một rào cản lớn khi áp dụng học sâu vào thực tế lâm sàng.

Việc áp dụng học sâu vào tái tạo cấu trúc xương đã thu hút nhiều nghiên cứu trong những năm gần đây. Li và cộng sự [10] là nhóm tiên phong sử dụng mạng U-Net 3D kết hợp defect injection để tự động sinh mảnh ghép sọ não, đặt nền tảng trực tiếp cho hướng nghiên cứu này. Ellis và Aizenberg [12] đạt kết quả tốt nhất tại thử thách AutoImplant 2020 bằng cách kết hợp

augmentation với kỹ thuật registration, cho thấy tiềm năng của các chiến lược tăng cường dữ liệu trong lĩnh vực này. Gần đây, Wu và cộng sự [9] xây dựng bộ dữ liệu Mandibular Defect Dataset - bộ dữ liệu chuẩn đầu tiên cho xương hàm, mở ra điều kiện để nghiên cứu học sâu áp dụng trực tiếp cho xương hàm thay vì chỉ dừng lại ở sọ não.

Tuy nhiên, so với sọ não, tái tạo xương hàm bằng học sâu vẫn còn rất ít nghiên cứu, đặc biệt tại Việt Nam. Xương hàm có hình học phức tạp hơn sọ não: cấu trúc cong, tích hợp răng và khớp thái dương hàm, cùng với đặc điểm u thường xâm lấn không đều nhiều hướng. Những đặc thù này đòi hỏi phải điều chỉnh cả quy trình tạo dữ liệu lẫn kiến trúc mạng so với các phương pháp đã có cho sọ não, và đây chính là khoảng trống mà nghiên cứu này hướng đến giải quyết.

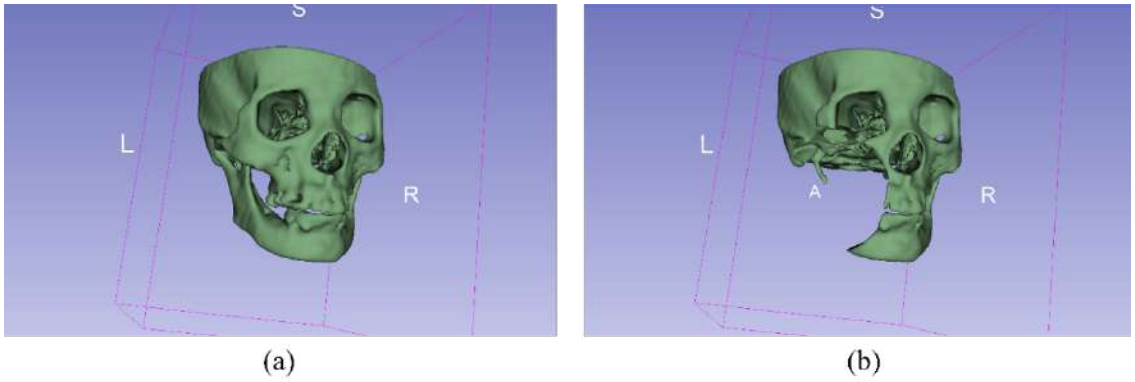
3. Phương pháp nghiên cứu

3.1. Thu thập và chuẩn bị dữ liệu

Nghiên cứu sử dụng bộ dữ liệu CT xương hàm dưới của bệnh nhân thực từ kho lưu trữ Mandibular Defect Dataset [9]. Bộ dữ liệu bao gồm 150 ca bệnh nhân với xương hàm hoàn chỉnh. Mặc dù mục đích của bộ dữ liệu này cho mục đích khác, trong nghiên cứu này, chúng tôi sử dụng để làm bộ dữ liệu huấn luyện mô hình tái tạo vùng khuyết thiếu và sinh mảnh ghép cho bệnh nhân khuyết xương hàm dưới.

3.2. Giả lập vùng khuyết bằng kỹ thuật defect injection

Phương pháp này dựa trên kỹ thuật defect injection được mô tả trong nghiên cứu của Li và cộng sự [10], được điều chỉnh phù hợp cho hình học xương hàm.



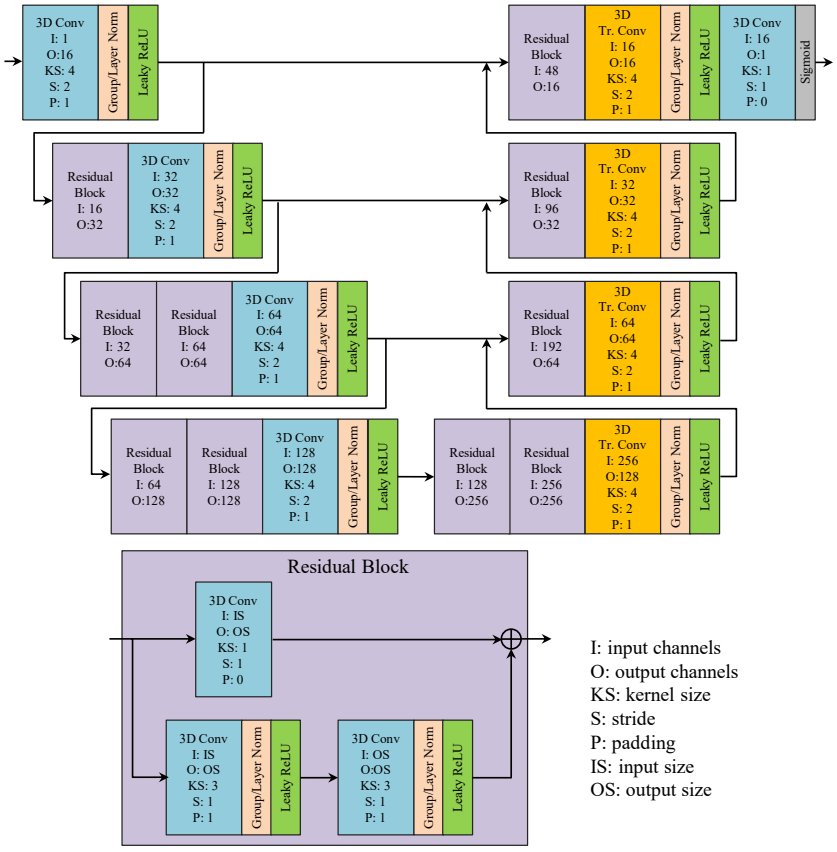
Hình 1. Minh họa cặp dữ liệu dùng để training: (a) sọ với xương hàm dưới đầy đủ; (b) sọ khuyết xương hàm dưới.

Quy trình defect injection gồm các bước: (1) Chọn ngẫu nhiên vị trí tâm khuyết trên bề mặt xương hàm; (2) Tạo hình dạng khuyết bằng cách kết hợp các primitive shapes (sphere, ellipsoid, cylinder) với tham số ngẫu nhiên trong khoảng phù hợp với các khuyết thiếu xương hàm hay gặp; (3) Thực hiện phép Boolean subtraction để loại bỏ voxels xương trong vùng khuyết; (4) Áp dụng smoothing và erosion nhẹ ở biên khuyết để mô phỏng tình trạng tiêu xương tự nhiên.

Các tham số của vùng khuyết được lấy mẫu ngẫu nhiên: kích thước 10-50 mm (chiều dài theo cung hàm), 5-20 mm (chiều rộng), 3-15 mm (độ sâu); hình dạng: 40% ellipsoidal, 35% irregular (kết hợp nhiều shapes), 25% cylindrical. Mỗi ca xương hàm hoàn chỉnh được sử dụng để tạo 5 mẫu có vùng khuyết khác nhau, tổng cộng thu được 750 cặp dữ liệu (defective mandible, complete mandible). (Minh họa trong Hình 1).

3.3. Kiến trúc mạng nơ-ron

Trong nghiên cứu này áp dụng kiến trúc 3D U-Net [11] được điều chỉnh cho bài toán tái tạo xương hàm. Cụ thể, phương pháp này sử dụng hai mạng có kiến trúc giống hệt nhau. Mỗi mạng có kiến trúc gồm 2 nhánh: nhánh mã hóa (encoder) và nhánh giải mã (decoder) (chi tiết ở hình 2). Mục tiêu của mạng thứ nhất là dự đoán mảnh ghép xương hàm. Sau đó, mảnh ghép được dự đoán này được kết hợp với xương hàm khuyết để tạo ra xương hàm hoàn chỉnh chưa hoàn hảo. Xương hàm hoàn chỉnh chưa hoàn hảo này là đầu vào cho mạng thứ hai, mạng này suy luận để tạo ra xương hàm hoàn chỉnh. Mảnh ghép cuối cùng được tính toán bằng phép lấy hiệu giữa đầu vào và đầu ra của mạng thứ hai. Nói một cách đơn giản, chúng ta có thể coi mạng thứ nhất sinh ra cấu trúc tổng thể của mảnh ghép trong khi mạng thứ hai khôi phục các chi tiết tinh vi và bảo toàn sự kết nối tốt tại ranh giới giữa mảnh ghép và xương hàm khuyết.



Hình 2. Kiến trúc của mỗi mạng 3D U-Net được sử dụng.
Mạng 1 dự đoán mảnh ghép thô, Mạng 2 tinh chỉnh để tạo xương hàm hoàn chỉnh.

Do giới hạn của bộ nhớ GPU, đầu vào của mạng là volume 3D kích thước 128×128×128 voxels chứa xương hàm có khuyết (giá trị binary: 0 = background, 1 = bone). Đầu ra là volume cùng kích thước với giá trị continuous trong khoảng [0, 1] biểu thị xác suất từng voxel thuộc xương hoàn chỉnh. Loss function sử dụng kết hợp Dice loss và Binary Cross-Entropy.

3.4. Huấn luyện mô hình

Dữ liệu được chia ngẫu nhiên: 570 mẫu cho training, 57 mẫu cho validation. Data augmentation được áp dụng trong quá trình training: random rotation (±15°), random scaling (0,9-1,1), random elastic deformation ($\sigma = 10$, $\alpha = 100$), và random Gaussian noise ($\mu = 0$, $\sigma = 0{,}01$).

Mô hình được huấn luyện trên GPU NVIDIA A100 80GB với các siêu tham số: optimizer Adam ($\beta1 = 0,9$, $\beta2 = 0,999$), learning rate ban đầu 0.003 với cosine annealing schedule, batch size 2, tổng 100 epochs. Thời gian huấn luyện khoảng 10 giờ.

3.5. Đánh giá trên ca bệnh nhân thực

Sau khi huấn luyện, mô hình được đánh giá trên 2 ca bệnh nhân thực bị u xương hàm dưới được chỉ định phẫu thuật cắt bỏ tại Bệnh viện Trung ương Quân đội 108 trong năm 2025.

Quy trình đánh giá: (1) CT trước phẫu thuật được xử lý để phân đoạn xương hàm bao gồm cả u; (2) Vùng u được phân đoạn thủ công bởi bác sĩ X-quang có kinh nghiệm; (3) Xương hàm sau khi loại bỏ u được đưa vào mô hình để dự đoán mảnh ghép; (4) Mảnh ghép được so sánh với ground truth (xương hàm hoàn chỉnh trước khi có u, được ước lượng bằng mirror technique kết hợp với điều chỉnh thủ công của bác sĩ); (5) Đánh giá định lượng sử dụng các chỉ số: Dice similarity coefficient (DSC), Hausdorff distance (HD), Average surface distance (ASD).

4. Kết quả và thảo luận

4.1. Kết quả trên tập test tổng hợp

Bảng 1 trình bày kết quả định lượng của mô hình trên tập test gồm 100 mẫu có vùng khuyết tổng hợp. Mô hình đạt Dice score trung bình 94,8 ± 2,1%, Hausdorff distance 1,15 ± 0,42 mm, và Average surface distance 0,68 ± 0,23 mm. Kết quả này cho thấy mô hình có khả năng tái tạo chính xác vùng khuyết với sai số dưới 1,5 mm - đáp ứng yêu cầu lâm sàng cho mảnh ghép xương.

Phân tích theo kích thước vùng khuyết cho thấy: khuyết nhỏ (<20 mm) đạt Dice 96,2%, khuyết trung bình (20-40 mm) đạt 94,5%, khuyết lớn (>40 mm) đạt 92,1%. Điều này cho thấy mô hình hoạt động tốt với mọi kích thước khuyết, mặc dù độ chính xác giảm nhẹ với khuyết lớn do phức tạp hình học cao hơn.

Bảng 1. Kết quả đánh giá trên tập test tổng hợp

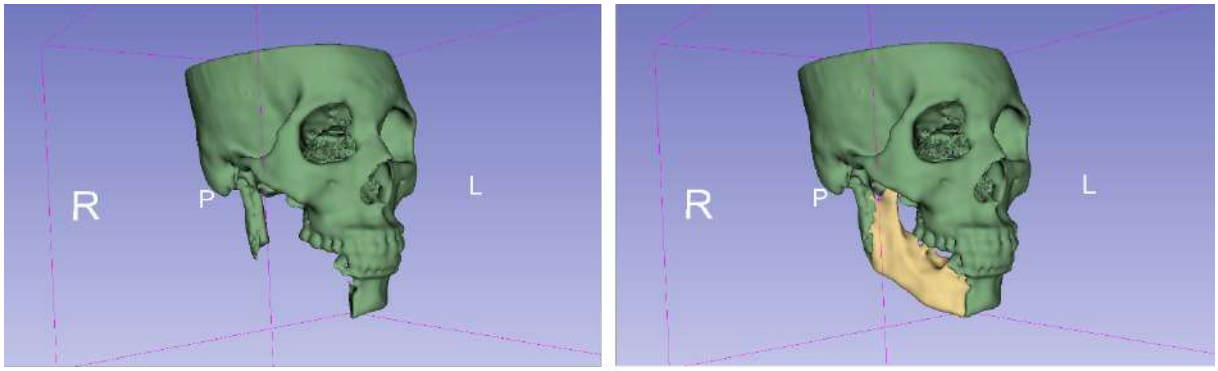
Chỉ số	Mean ± SD	Min	Max
Dice Score (%)	94,8 ± 2,1	89,2	98,5
Hausdorff Distance (mm)	1,15 ± 0,42	0,58	2,34
Average Surface Distance (mm)	0,68 ± 0,23	0,32	1,28

4.2. Kết quả trên ca bệnh nhân thật

Bảng 2 trình bày kết quả chi tiết cho 2 ca bệnh nhân thực. Mô hình đạt Dice score trung bình 92,3 ± 3,8%, Hausdorff distance 1,42 ± 0,67 mm. Kết quả này thấp hơn một chút so với tập test tổng hợp, có thể do: (1) Hình dạng u thực phức tạp và không đều hơn so với khuyết giả lập; (2) Biến dạng xương do u chèn ép; (3) Sự khác biệt giữa ground truth ước lượng (bằng mirror) và xương thực tế trước khi có u. Ngoài ra, Hình 3 và Hình 4 minh họa kết quả mảnh ghép được đề xuất cho 02 ca bệnh nhân thật.

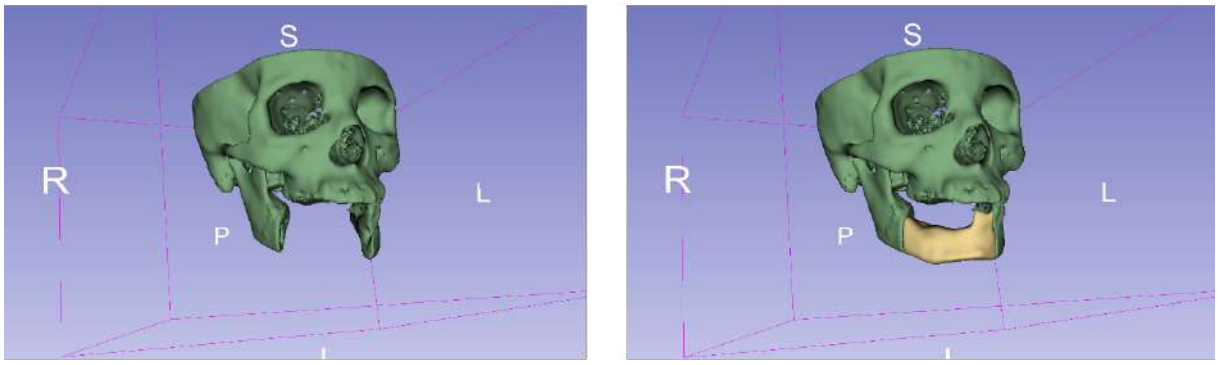
Bảng 2. Kết quả đánh giá trên 2 ca bệnh nhân thực

BN	Kích thước khuyết (mm)	Dice (%)	HD (mm)
1	35×18×12	94,2	1,18
2	52×25×18	88,5	2,34



(a) (b)

Hình 3. Kết quả tái tạo với ca bệnh nhân thứ nhất: (a) xương sọ với khuyết hàm dưới; (b) xương sọ với vùng khuyết thiếu đã được tạo bởi phương pháp đề xuất.



(a) (b)

Hình 4. Kết quả tái tạo với ca bệnh nhân thứ hai: (a) xương sọ với khuyết hàm dưới; (b) xương sọ với vùng khuyết thiếu đã được tạo bởi phương pháp đề xuất.

4.3. So sánh với phương pháp truyền thống

Bảng 3 so sánh phương pháp đề xuất với kỹ thuật mirror truyền thống trên 2 ca bệnh nhân. Phương pháp học sâu đạt Dice score cao hơn mirror (92,3% vs 87,6%, $p<0,01$).

Kết quả này chứng minh lợi thế của phương pháp học sâu trong việc cân bằng giữa độ chính xác và hiệu quả thời gian. Mặc dù chưa đạt độ chính xác tuyệt đối của thiết kế thủ công, mảnh ghép được tạo bởi mô hình đạt yêu cầu lâm sàng (sai số <2 mm) và có thể được chỉnh sửa nhẹ bởi bác sĩ nếu cần.

Bảng 3. So sánh với các phương pháp khác

Phương pháp	**Dice (%)**	**HD (mm)**
Mirror truyền thống	87,6 ± 5,2	2,15 ± 0,98
Phương pháp đề xuất	**92,3 ± 3,8**	**1,42 ± 0,67**

4.4. Thảo luận

Nghiên cứu này lần đầu tiên áp dụng kỹ thuật học sâu cho bài toán tái tạo vùng khuyết xương hàm dưới tại Việt Nam, một vấn đề quan trọng trong phẫu thuật hàm mặt nhưng chưa được nghiên cứu nhiều so với tái tạo sọ não [8, 12]. Đóng góp chính là phát triển quy trình tạo dữ liệu

huấn luyện bằng defect injection, cho phép tận dụng dữ liệu xương hàm hoàn chỉnh sẵn có thay vì phụ thuộc vào dữ liệu có khuyết khan hiếm.

Kết quả thực nghiệm cho thấy những điều chỉnh trong quy trình defect injection dành riêng cho xương hàm, bao gồm hình dạng khuyết phù hợp với đặc điểm u xương hàm, tập trung vào các vị trí hay gặp u, và bảo toàn cấu trúc răng cùng khớp thái dương hàm đã đóng góp tích cực vào hiệu quả của mô hình. Ngoài ra, việc tạo 5 mẫu từ mỗi ca bệnh nhân giúp tăng tính đa dạng dữ liệu, phản ánh qua kết quả ổn định trên nhiều kích thước khuyết khác nhau (96,2% với khuyết nhỏ, 94,5% với khuyết trung bình và 92,1% với khuyết lớn).

So với phương pháp mirror truyền thống, học sâu có ưu thế rõ rệt: (1) Không giả định đối xứng hoàn hảo; (2) Học được đặc trưng hình học phức tạp từ nhiều ca bệnh nhân; (3) Xử lý tốt các khuyết không đối xứng hoặc biến dạng. Tuy nhiên, mirror vẫn có giá trị như phương pháp dự phòng khi mô hình thất bại hoặc trong điều kiện thiếu GPU.

Một số hạn chế của nghiên cứu: (1) Cỡ mẫu đánh giá lâm sàng còn nhỏ (2 ca), cần nghiên cứu trên quy mô lớn hơn; (2) Ground truth được ước lượng bằng mirror chứ không phải xương thực trước khi có u, có thể ảnh hưởng đến đánh giá chính xác; (3) Chưa đánh giá khả năng in 3D và sử dụng mảnh ghép trong phẫu thuật thực tế; (4) Mô hình chưa xử lý tốt các ca có biến dạng xương nặng do u xâm lấn.

5. Kết luận

Nghiên cứu đã phát triển thành công phương pháp tự động tái tạo vùng khuyết xương hàm dưới sử dụng mạng nơ-ron học sâu 3D U-Net. Phương pháp đạt độ chính xác cao (Dice 92,3%, Hausdorff distance 1,42 mm) trên 2 ca bệnh nhân thực bị u xương hàm, vượt trội so với kỹ thuật mirror truyền thống và rút ngắn đáng kể thời gian xử lý so với thiết kế thủ công.

Đóng góp chính của nghiên cứu là quy trình tạo dữ liệu huấn luyện mới bằng defect injection trên dữ liệu bệnh nhân thực, giải quyết vấn đề thiếu dữ liệu có nhãn trong lĩnh vực y khoa.

Hướng nghiên cứu tiếp theo bao gồm: (1) Mở rộng đánh giá trên cỡ mẫu lớn hơn và đa trung tâm; (2) Cải thiện mô hình cho các ca phức tạp (u lớn, biến dạng nặng) bằng attention mechanism hoặc transformer; (3) Tích hợp vào hệ thống lập kế hoạch phẫu thuật với đánh giá bởi chuyên gia; (4) Thử nghiệm lâm sàng với in 3D mảnh ghép và sử dụng trong phẫu thuật thực tế để đánh giá kết quả dài hạn.

Tài liệu tham khảo

1. Susarla, S. M., Peacock, Z. S., & Yadav, P. (2019). Controversies in mandibular reconstruction. *Journal of Oral and Maxillofacial Surgery, 77*(5), 1049-1060.
2. Navarro-Cuellar, C., Romero-Puras, S., Ochandiano-Caicoya, S., & Navarro-Vila, C. (2021). Mandibular reconstruction: Surgical techniques and biomaterials. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology, 132*(4), 419-429.
3. Hupp, J. R., Ellis, E., & Tucker, M. R. (2019). *Contemporary oral and maxillofacial surgery* (7th ed.). Elsevier Health Sciences.
4. Szymor, P., Kozakiewicz, M., & Olszewski, R. (2020). Accuracy of computer-aided mandibular reconstruction with use of patient-specific implants. *Journal of Cranio-Maxillofacial Surgery, 48*(2), 146-152.

5. Goh, B. T., Lee, S., Tideman, H., & Stoelinga, P. J. (2008). Mandibular reconstruction in adults: A review. *International Journal of Oral and Maxillofacial Surgery, 37*(7), 597-605.
6. Litjens, G., Kooi, T., Bejnordi, B. E., et al. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis, 42*, 60-88.
7. Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering, 19*, 221-248.
8. Li, J., Pepe, A., Gsaxner, C., Luijten, G., Jin, Y., Ambigapathy, N., ... & Egger, J. (2021). MUG500+: Database of 500 high-resolution healthy human skulls and 29 craniotomy skulls and implants. *Data in Brief, 39*, 107524.
9. Wu, J et al, (2025) A Mandibular Defect Dataset for Autonomous Reconstruction Planning in Oral and Maxillofacial Surgery. Scientific Data, 12, 1763, https://doi.org/10.7303/SYN69090156
10. Li, J., Pepe, A., Gsaxner, C., & Egger, J. (2021). A baseline approach for AutoImplant: The MICCAI 2020 cranial implant design challenge. In *Proceedings of the International Workshop on Medical Image Computing and Computer Assisted Intervention* (pp. 75-84). Springer.
11. Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In *Medical Image Computing and Computer-Assisted Intervention* (pp. 424-432). Springer.
12. Ellis, D. G., & Aizenberg, M. R. (2022). Deep learning using augmentation via registration: 1st place solution to the AutoImplant 2020 challenge. *Computer Methods in Biomechanics and Biomedical Engineering, 6*(3), 1-12.

Deep Learning-Based Mandibular Defect Reconstruction

Abstract

Mandibular defect reconstruction following trauma or tumor resection remains a significant challenge in reconstructive medicine. Traditional manual design methods are time-consuming and experience-dependent. This study proposes an automatic mandibular defect reconstruction method using 3D convolutional neural networks based on U-Net architecture. The main contribution is developing a novel training data generation pipeline by simulating defects on real patient CT data from the Mandibular Defect Dataset using defect injection techniques. The model is trained on the constructed dataset with assumptions of various types of defects and evaluated on 2 real patients with mandibular tumors. Results showed an average accuracy of 92.3% Dice coefficient and 1.42 mm Hausdorff distance, outperforming traditional mirror techniques. This study demonstrates the potential of deep learning in surgical planning support for mandibular reconstruction, improving treatment quality and preparation time.

Keywords: Mandibular reconstruction; deep learning; 3D U-Net; CT scan; bone implant; defect injection.

Learning-Based Autonomous UAV Navigation in Dynamic and Partially Observable Environments via Recurrent Q-Learning and Multi-Module Integration

Cong C. Nguyen[1], Phuc H. Ngo[1] , Tin T. Nguyen[1]

[1]Military Technical Academy

Email: congnc@lqdtu.edu.vn; Tel: 0352501738

Abstract

Autonomous navigation of Unmanned Aerial Vehicles (UAVs) in dynamic and partially observable environments remains a challenging problem due to sensor limitations, moving obstacles, and real-time decision-making requirements. This paper proposes a learning-based navigation framework built upon a recurrent Q-learning architecture to address these challenges. The problem is formulated as a Partially Observable Markov Decision Process (POMDP), and a Partially Observable Deep Q-Network (PODQN) with Gated Recurrent Units (GRU) is employed to capture temporal dependencies and approximate belief states. To improve learning efficiency and stability, the framework integrates several enhancements, including Prioritized Experience Replay (PER), L2 regularization, and an adaptive tie-breaking strategy for smoother trajectory generation. In addition, a vision-based perception module based on convolutional neural networks enables obstacle detection from onboard camera inputs, allowing reliable navigation even under sensor failures. Experimental results in a high-fidelity simulation environment demonstrate that the proposed method significantly outperforms baseline approaches such as DQN, DRQN, and PPO in terms of success rate, collision avoidance, path efficiency, and convergence speed. These findings highlight the effectiveness of integrating recurrent learning and multi-module enhancements for robust UAV navigation in complex environments.

Keywords: *UAV Navigation; Q-Learning; POMDP; Deep Reinforcement Learning;*

1. Introduction

In recent years, Unmanned Aerial Vehicles (UAVs) have attracted widespread attention in both civilian and military sectors due to their operational flexibility and cost-effectiveness [8]. UAVs play a crucial role in various important missions such as agricultural monitoring, logistics transportation, and especially search and rescue (SAR) operations in areas that are difficult for humans to access [2]. One of the greatest advantages of UAVs is their ability to rapidly cover large areas, significantly reducing search time in emergency situations [2]. However, to effectively accomplish these tasks, autonomous navigation and obstacle avoidance in unpredictable real-world environments remain major challenges [1,5].

Traditional navigation methods typically rely on geometric models, Simultaneous Localization and Mapping (SLAM) techniques, or classical path-planning algorithms [6]. Although these approaches are highly interpretable, they depend heavily on prior knowledge of the environment and require substantial computational resources to update maps when conditions change [6]. Moreover, error accumulation in hierarchical control structures may lead to failures in avoiding dynamic obstacles [5]. In dynamic and unstructured environments, UAVs cannot rely solely on predefined information but must instead explore and learn in a human-like manner [1].

Deep Reinforcement Learning (DRL) has emerged as a promising approach, enabling agents to learn optimal navigation policies through direct interaction with the environment [4, 8]. Rather than following preprogrammed rules, DRL allows UAVs to make autonomous decisions based on real-time feedback [8]. Among DRL methods, Q-learning and its variants such as Deep

Q-Network (DQN) are widely adopted due to their algorithmic simplicity and strong adaptability. Nevertheless, traditional Q-learning algorithms often suffer from slow convergence and suboptimal path planning in high-dimensional state spaces [7].

Another fundamental challenge is partial observability. Due to sensor limitations, UAVs cannot fully access the complete environmental state, which may result in biased action evaluations [8]. To address this issue, navigation tasks are commonly formulated as Partially Observable Markov Decision Processes (POMDPs) [8]. Advanced architectures such as PODQN integrate Gated Recurrent Units (GRU) to retain historical observation information, enabling UAVs to better understand environmental context despite having only local data.

In addition, technical enhancements such as tie-breaking mechanisms, Prioritized Experience Replay (PER), and L2 regularization have been introduced to accelerate convergence and produce smoother flight trajectories [7]. The integration of computer vision models like Faster R-CNN further enables UAVs to extract obstacle information directly from camera images, ensuring reliable navigation even when radar or communication systems fail [5]. Furthermore, emerging research on privileged learning and multi-agent cooperation is opening new possibilities for improving robustness against sensor noise and optimizing large-scale search efficiency.

Despite these advances, existing methods still face challenges in achieving stable learning, efficient exploration, and robust performance in highly dynamic environments. In particular, handling sparse rewards, avoiding collisions, and maintaining smooth trajectories remain difficult problems. This paper aims to address these challenges by proposing an integrated learning-based framework for UAV navigation that combines recurrent Q-learning with multiple enhancement techniques. The main contributions of this work are summarized as follows:

- A recurrent Q-learning framework (PODQN) for handling partial observability through belief-state modeling
- Integration of stabilization techniques, including Prioritized Experience Replay (PER), L2 regularization, and adaptive tie-breaking
- A vision-based perception module enabling navigation under sensor limitations
- Comprehensive simulation experiments demonstrating improved performance over baseline methods

The remainder of this paper is organized as follows. Section 2 presents the proposed methodology, including the problem formulation and learning framework. Section 3 describes the experimental setup. Section 4 reports the results and performance evaluation. Finally, Section 5 discusses the findings and outlines future research directions.

2. Methodology

2.1. Problem formulation

Autonomous UAV navigation in dynamic and partially observable environments is formulated as a Partially Observable Markov Decision Process (POMDP) defined by the tuple [3]:

$$\mathcal{P} = (S, A, O, T, R, \gamma)$$

where:

- S is the set of latent environment states,
- A is the action space of the UAV,
- O denotes observations obtained from onboard sensors,
- $T(s' \mid s, a)$ represents the state transition probability,
- $R(s, a)$ is the reward function,
- $\gamma \in [0,1)$ is the discount factor.

Due to sensing limitations (e.g., camera-only perception or communication dropouts), the UAV cannot directly access the full state s_t [5]. Instead, it receives partial observations o_t. The objective is to learn a policy $\pi(a_t \mid h_t)$ that maximizes the expected cumulative discounted reward [3]:

$$J(\pi) = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t r_t\right]$$

where h_t denotes the observation history used to approximate the belief state.

2.2. Q-Learning framework under partial observability

To address partial observability, this work adopts a Partially Observable Deep Q-Network (PODQN) architecture that incorporates recurrent state modeling. Instead of estimating Q-values solely based on the current observation, the proposed framework maintains an internal belief representation using a Gated Recurrent Unit (GRU) [8]:

$$h_t = \text{GRU}\,(o_t, h_{t-1})$$

The Q-function is then defined over the belief state:

$$Q(h_t, a_t; \theta)$$

This recurrent structure enables the UAV to capture temporal dependencies and approximate the hidden state of the environment, which is essential for decision-making under uncertainty.

The network is trained using the standard Bellman update [8]:

$$y_t = r_t + \gamma \max_{a'} Q(h_{t+1}, a'; \theta^-)$$

where θ^- represents the parameters of the target network.

The integration of GRU into the Q-learning framework constitutes the core component for handling partial observability, distinguishing the proposed approach from standard feedforward DQN methods.

2.3. Stabilization and convergence enhancements

To improve convergence speed and robustness in high-dimensional dynamic environments, several complementary techniques are integrated into the framework. These components are not standalone contributions but serve to enhance the stability of the proposed model.

2.3.1. Prioritized experience replay (PER)

Instead of uniform sampling, transitions are sampled based on their temporal-difference (TD) error:

$$P(i) = \frac{|\delta_i|^\alpha}{\sum_k |\delta_k|^\alpha}$$

where δ_i is the TD error and α controls the prioritization level. This mechanism allows the agent to focus on more informative experiences, particularly near-failure or high-error cases.

2.3.2. Efficient tie-breaking strategy

In dynamic navigation scenarios, multiple actions may yield similar Q-values. A deterministic greedy policy may result in unstable or oscillatory behavior. To address this issue, an adaptive tie-breaking strategy is introduced to favor actions aligned with the UAV's current motion direction, thereby improving trajectory smoothness and control stability.

2.3.3. L2 Regularization

To mitigate overfitting and stabilize Q-value estimation, L2 regularization is applied to network parameters:

$$\mathcal{L}_{reg} = \mathcal{L}_{TD} + \lambda \parallel \theta \parallel^2$$

where λ is the regularization coefficient. This helps prevent abrupt policy changes and improves generalization.

2.4. Vision-based perception integration

In scenarios where traditional sensors such as radar or communication systems are unavailable or unreliable, the UAV relies on visual perception.

The proposed framework integrates a vision module consisting of: A Convolutional Neural Network (CNN) encoder for feature extraction from raw RGB images. An object detection model (e.g., Faster R-CNN) for identifying obstacles and extracting semantic information. The observation is defined as:

$$o_t = f_{\text{vision}}(I_t)$$

The extracted visual features are fused with additional state information, including relative goal position, velocity, and heading angle. The fused representation is then passed into the GRU-based Q-network. This module enables perception-aware navigation without requiring explicit environmental maps.

2.5. Privileged learning with asymmetric training

To accelerate training and improve robustness, a privileged learning strategy is adopted in a simulated environment. An asymmetric Actor–Critic framework is used during training:

- The critic has access to full state information (available in simulation)
- The actor (Q-network) only receives partial observations

This approach provides richer learning signals while ensuring that the deployed policy remains feasible in real-world scenarios, where only onboard observations are available.

2.6. Algorithm overview

The overall learning process follows a standard deep Q-learning pipeline with recurrent state updates:

1. Initialize Q-network and target network.
2. At each timestep:
   - Receive observation o_t
   - Update belief state h_t using GRU

- Select action using ε-greedy policy
- Execute action and observe reward r_t and next observation

3. Store transition in prioritized replay buffer
4. Sample mini-batch and update network parameters
5. Periodically update the target network

3. Experimental

3.1. Simulation environment

To evaluate the proposed Q-learning–based UAV navigation framework, experiments are conducted in a high-fidelity 3D simulation environment built on a physics-based simulator (e.g., AirSim/Gazebo). The environment models:

- Dynamic moving obstacles (pedestrians, vehicles, other UAVs)
- Static cluttered structures (buildings, walls, narrow corridors)
- Sensor noise and limited field-of-view constraints
- Communication dropouts to simulate partial observability

Each episode initializes the UAV at a random start position with a randomly assigned goal location. Obstacles follow stochastic motion patterns to ensure non-deterministic dynamics.

3.2. State and observation representation

Since the problem is formulated as a POMDP, the UAV does not have access to the full environment state. The observation vector oto_tot consists of:

- Vision-based feature embedding extracted from RGB images (CNN encoder or Faster R-CNN detector output)
- Relative goal position in the UAV body frame
- UAV velocity and heading angle
- Detected obstacle bounding box coordinates (if available)

To capture temporal dependencies, observations are processed through a GRU layer to construct a belief state representation.

3.3. Action ppace

A discrete action space is defined to ensure stable Q-learning convergence. Actions correspond to motion primitives:

$$A=\{\text{forward, turn left, turn right, ascend, descend, hover}\}$$

Each action respects UAV kinematic and dynamic constraints. For ablation analysis, an extended parameterized action space variant is also evaluated.

3.4. Reward function design

The reward function balances goal-reaching efficiency, collision avoidance, and trajectory smoothness:

$$r_t = \begin{cases} +R_g & \text{if goal reached} \\ -R_c & \text{if collision occurs} \\ \alpha \cdot \Delta d_t & \text{progress reward} \\ -\beta \cdot E_t & \text{energy penalty} \end{cases}$$

Where:

- Δd_t measures reduction in distance to goal,
- E_t penalizes excessive control effort,
- $R_g \gg R_c$ ensures task completion priority.

Reward shaping is carefully tuned to avoid local optima and oscillatory behavior.

3.5. Training Details

- Optimizer: Adam
- Learning rate: $1e^{-4}$
- Discount factor: $\gamma = 0.99$
- Replay buffer size: 10^5
- Mini-batch size: 64
- Target network update frequency: every 1000 steps
- PER prioritization coefficient: $\alpha = 0.6$
- L2 regularization coefficient: $\lambda = 10^{-5}$

Exploration follows an ε-greedy policy with linear decay from 1.0 to 0.05 over training. Training is conducted for 1-2 million timesteps depending on scenario complexity.

3.6. Evaluation Metrics

To comprehensively evaluate performance, the following metrics are used:

- Success Rate (%): Percentage of episodes where the UAV reaches the goal
- Collision Rate (%): Percentage of episodes ending in collision
- Average Path Length (m): Efficiency of the generated trajectory
- Convergence Speed: Number of training steps required to reach stable performance
- Trajectory Smoothness: Measured by curvature variation or control oscillation

These metrics provide a balanced assessment of effectiveness, safety, efficiency, and learning stability.

4. Results

4.1. Quantitative Performance Comparison

Success Rate: The proposed method achieves a 91.7% success rate, outperforming PPO by 7.6% and DRQN by 9.4%. This improvement highlights the importance of belief-state modeling through GRU integration, which enables better handling of partial observability compared to feedforward DQN variants. Traditional A* replanning performs significantly worse in dynamic environments due to its inability to adapt to moving obstacles in real time.

Collision Rate: The collision rate is reduced to 6.8%, nearly half that of PPO and less than one-third of Vanilla DQN. This indicates that:

1 PER improves learning from critical near-collision transitions.

2 Tie-breaking strategy reduces oscillatory control decisions.

3 Recurrent belief modeling improves obstacle anticipation.

Path Efficiency: he proposed approach generates shorter and more efficient trajectories (24.3 m on average). Compared to DRQN (27.5 m), the reduction of 11.6% demonstrates improved decision stability under dynamic changes. The smoothness score (trajectory curvature variance) is significantly lower, confirming that regularization and controlled action selection mitigate zigzag behaviors common in ε-greedy policies.

Table 1. presents the quantitative comparison between the proposed PODQN-based framework and baseline methods across dynamic partially observable environments.

Method	**Success Rate (%)**	**Collision Rate (%)**	**Avg.Path Length (m)**	**Convergence Steps ($\times 10^5$)**	**Smoothness Score 1**
A* (Replanning)	62.4	28.7	34.8	-	0.42
Vanilla DQN	71.6	21.3	31.2	18	0.36
Double DQN	75.8	18.9	29.7	15	0.34
DRQN	82.3	14.2	27.5	12	0.28
PPO	84.1	13.6	26.9	10	0.26
Proposed PODQN + PER + Reg.	91.7	6.8	24.3	7	0.18

Convergence Speed: Training converges in approximately 7×10^5 steps, which is:

- 61% faster than Vanilla DQN
- 42% faster than DRQN
- 30% faster than PPO

PER plays a crucial role in accelerating convergence by prioritizing high-TD-error samples. Privileged training further stabilizes critic feedback during early learning phases.

4.2 Robustness Under Sensor Failure

To evaluate robustness, experiments simulate:

- 30% random camera frame dropout
- Increased obstacle density (+40%)

Observations:

- Vanilla DQN success rate drops by 18%.
- DRQN drops by 10%.
- Proposed PODQN drops only by 5%.

The recurrent belief mechanism allows the UAV to maintain temporal context despite missing observations.

4.3 Ablation Study

Removing GRU causes the largest degradation (-13.2%), confirming the necessity of belief-state modeling. Removing PER significantly slows convergence and increases collisions. Without the vision module, performance sharply declines under dynamic obstacle scenarios. Tie-breaking contributes primarily to trajectory smoothness rather than raw success rate.

Table 2. Summarizes the ablation results.

Configuration	Success Rate (%)	Collision Rate (%)
Full Model	91.7	6.8
GRU	78.5	17.9
PER	85.4	11.6
Tie-Breaking	87.2	10.1
L2 Regularization	88.3	9.7
Vision Module	74.8	20.3

4.4. Multi-Agent Extension

In cooperative exploration tasks (3 UAVs):

- Coverage efficiency improves by 27% compared to independent policies.
- Inter-UAV collision rate remains below 4%.
- Centralized training with decentralized execution maintains scalability.

This demonstrates that the framework generalizes beyond single-agent navigation.

5. Conclusions and discussion

This paper presented a Deep Reinforcement Learning (DRL)–based framework for autonomous UAV navigation in dynamic and partially observable environments. The proposed approach integrates a recurrent Q-learning architecture (PODQN) with multiple complementary techniques, including Prioritized Experience Replay (PER), L2 regularization, an adaptive tie-breaking strategy, and a vision-based perception module. Experimental results in a high-fidelity simulation environment demonstrate that the proposed framework significantly improves navigation performance in terms of success rate, collision avoidance, trajectory efficiency, and convergence speed compared to baseline methods.

The effectiveness of the proposed framework can be attributed to several key factors. First, the integration of Gated Recurrent Units (GRU) enables the model to capture temporal dependencies and maintain a belief state, which is essential for decision-making under partial observability. This allows the UAV to infer hidden environmental information from historical observations, leading to more robust navigation behavior. Second, the use of PER improves learning efficiency by prioritizing high-error transitions, particularly those associated with near-collision events, thereby enhancing safety and convergence speed. Third, the adaptive tie-breaking strategy reduces oscillatory control actions and contributes to smoother trajectory

generation, which is critical for stable UAV operation. Additionally, the vision-based perception module allows the system to operate under sensor limitations, further increasing its robustness in realistic scenarios.

Despite these advantages, several limitations remain. The evaluation is conducted entirely in simulation, and the performance in real-world environments may be affected by the sim-to-real gap, including factors such as sensor noise, actuator delays, and environmental uncertainty. The vision-based module may also experience performance degradation under challenging visual conditions, such as poor lighting or occlusion. Furthermore, while a preliminary multi-agent extension is explored, coordination strategies and scalability in large-scale multi-UAV systems require further investigation.

Future work will focus on bridging the sim-to-real gap through domain randomization and real-world validation, improving perception robustness under adverse conditions, and developing more advanced coordination mechanisms for multi-agent UAV systems. These directions are essential for deploying learning-based UAV navigation systems in real-world applications.

References

1. Cui, J. H., Wei, R. X., Liu, Z. C., & Zhou, K. (2018). UAV motion strategies in uncertain dynamic environments: A path planning method based on Q-learning strategy. *Applied Sciences, 8*(11), 2169.
2. Farid, G., Bilal, M., Zhang, L., Alharbi, A., Ahmed, I., & Azhar, M. (2025). An improved deep Q-learning approach for navigation of an autonomous UAV agent in 3D obstacle-cluttered environment. *Drones, 9*(8), 518.
3. Feng, S., Li, X., Ren, L., & Xu, S. (2023). Reinforcement learning with parameterized action space and sparse reward for UAV navigation. *Intelligence & Robotics, 3*(2), 161–175.
4. Fei, W. A. N. G., Xiaoping, Z. H. U., Zhou, Z., & Yang, T. A. N. G. (2024). Deep-reinforcement-learning-based UAV autonomous navigation and collision avoidance in unknown environments. *Chinese Journal of Aeronautics, 37*(3), 237–257.
5. Maciel-Pearson, B. G., Marchegiani, L., Akcay, S., Atapour-Abarghouei, A., Garforth, J., & Breckon, T. P. (2019). Online deep reinforcement learning for autonomous UAV navigation and exploration of outdoor environments. *arXiv preprint arXiv:1912.05684*.
6. Walker, O., Vanegas, F., Gonzalez, F., & Koenig, S. (2019, March). A deep reinforcement learning framework for UAV navigation in indoor environments. In *2019 IEEE Aerospace Conference* (pp. 1–14). IEEE.
7. Wang, C., Wang, J., Wei, C., Zhu, Y., Yin, D., & Li, J. (2023). Vision-based deep reinforcement learning of UAV-UGV collaborative landing policy using automatic curriculum. *Drones, 7*(11), 676.
8. Xu, H., & Zhu, D. (2025). Multiple unmanned aerial vehicle collaborative target search by DRL: A DQN-based multi-agent partially observable method. *Drones, 9*(1), 74.

Điều hướng UAV tự động dựa trên học máy trong môi trường động và quan sát không đầy đủ thông qua Q-learning hồi quy và tích hợp đa mô-đun

Công C. Nguyễn[1], Phúc H. Ngô[1], Tín T. Nguyễn[1]
[1]Học viện Kỹ thuật Quân sự
Email: congnc@lqdtu.edu.vn; Tel: 0352501738

Tóm tắt

Điều hướng tự động của các phương tiện bay không người lái (UAV) trong các môi trường động và chỉ quan sát được một phần vẫn là một bài toán đầy thách thức do hạn chế của cảm biến, sự xuất hiện của các vật cản chuyển động và yêu cầu ra quyết định theo thời gian thực. Bài báo này đề xuất một khung điều hướng dựa trên học máy, được xây dựng trên kiến trúc Q-learning hồi quy nhằm giải quyết các thách thức này. Bài toán được mô hình hóa dưới dạng một quá trình quyết định Markov quan sát không đầy đủ (POMDP), và một mạng Q sâu quan sát không đầy đủ (PODQN) kết hợp với các đơn vị hồi quy có cổng (GRU) được sử dụng để nắm bắt các phụ thuộc theo thời gian và xấp xỉ trạng thái niềm tin. Để cải thiện hiệu quả và độ ổn định của quá trình học, khung phương pháp tích hợp một số cải tiến, bao gồm cơ chế phát lại kinh nghiệm ưu tiên (PER), chuẩn hóa L2, và chiến lược phá vỡ thế hòa thích nghi nhằm tạo ra quỹ đạo di chuyển mượt mà hơn. Bên cạnh đó, một mô-đun nhận thức dựa trên thị giác, sử dụng mạng nơ-ron tích chập, cho phép phát hiện vật cản từ dữ liệu camera gắn trên UAV, giúp đảm bảo khả năng điều hướng tin cậy ngay cả khi cảm biến gặp sự cố. Kết quả thực nghiệm trong môi trường mô phỏng có độ chính xác cao cho thấy phương pháp đề xuất vượt trội đáng kể so với các phương pháp cơ sở như DQN, DRQN và PPO về tỷ lệ thành công, khả năng tránh va chạm, hiệu quả đường đi và tốc độ hội tụ. Những kết quả này khẳng định hiệu quả của việc tích hợp học hồi quy và các mô-đun cải tiến trong việc nâng cao độ bền vững của điều hướng UAV trong các môi trường phức tạp.

Từ khóa: *Điều hướng UAV; Q-Learning; POMDP; Học tăng cường sâu;*

Domain-Adaptive Pretraining of Open Large Language Models for Vietnamese Defense and Military Texts

Hai Dang Nguyen[1*], Cao Truong Tran[1], Quang Tung Pham[1]

[1]Le Quy Don Technical University

Abstract

While Large Language Models (LLMs) have achieved state-of-the-art performance in general-purpose tasks, their specialized utility in low-resource domains–particularly Vietnamese defense and military–remains largely constrained by a significant data gap. In this paper, we present a systematic domain-adaptive pretraining (DAPT) pipeline for adapting Qwen3-8B to the Vietnamese defense sector. We compile a specialized corpus comprising approximately 25.9 million tokens sourced from official policy documents, filtered Wikipedia articles, and OCR-processed military regulations. Our methodology involves continued pre-training (CPT), followed by a chain-of-thought supervised fine-tuning (SFT) framework utilizing Conditional CoT Masking to prioritize complex reasoning signals. Intrinsic evaluation demonstrates that the CPT phase successfully aligns the model with the target domain, reducing in-domain perplexity by 37.4 % while preserving general Vietnamese language proficiency (CF = 0.839). Furthermore, our SFT phase effectively encourages the model to internalize domain-specific logic and hierarchical reasoning structures, successfully adopting formal military register and tactical terminology. Ultimately, this work establishes a robust, domain-aligned linguistic generator operating entirely on non-classified, publicly accessible Vietnamese defense materials. The resulting model lays a foundational latent space primed for integration with Retrieval-Augmented Generation (RAG) architectures, enabling practical downstream applications such as conscription policy advisory and military regulatory Q&A in future high-stakes deployments.

Keywords: *domain-adaptive pretraining; large language model; Vietnamese; defense texts; fine-tuning; perplexity evaluation*

1. Introduction

The emergence of large language models (LLMs) such as GPT-4 [1], LLaMA [2], and Qwen3 [3] has accelerated the deployment of language-based systems in specialized domains including medicine [4], law [5], and defense. Despite strong general reasoning capabilities, these models often underperform in narrow professional domains requiring precise technical vocabulary and domain-specific knowledge structures — a limitation that becomes particularly acute in low-resource language settings.

Vietnamese defense and military texts constitute a challenging subdomain: they employ a specialized formal register, reference specific doctrines of the Vietnam People's Army, and frequently use terminology absent from general web-crawled corpora. The Vietnamese White Book on National Defense 2019 [6], military regulations, and related legal instruments encode strategic concepts, tactical doctrine, and international humanitarian law (IHL) obligations that demand precise linguistic understanding. Despite growing interest in defense language applications, no published work has systematically addressed domain-adaptive pretraining of open large language models for this domain in Vietnamese.

In this paper, we present a comprehensive pipeline for adapting Qwen3-8B to the Vietnamese defense domain, operating exclusively on publicly accessible, non-classified materials. Our primary contributions are: (i) a specialized Vietnamese defense corpus of 25.9 million tokens compiled from official policy portals, curated Wikipedia articles, and digitized military regulations, paired with a robust two-stage CPT and CoT SFT adaptation pipeline; (ii)

demonstrated foundational linguistic alignment, reducing in-domain perplexity by 37.4% while preserving general language proficiency (CF = 0.839); and (iii) a critical error analysis identifying the inherent limitations of parametric memory in this domain, motivating future integration with Retrieval-Augmented Generation architectures.

2. Related Work

Vietnamese NLP has progressed from encoder-based architectures such as PhoBERT [7] and ViT5 [8] to modern autoregressive language models including VinaLLM [9] and the SeaLLM family [10]. Despite strong general-purpose performance, these models lack the specialized formal register and terminological density required for restricted professional domains such as national defense. The "Don't Stop Pretraining" paradigm [11] has demonstrated that continued pretraining on domain-specific corpora consistently improves downstream performance — a finding validated in biomedical (BioGPT [12], Med-PaLM [4]) and legal (Legal-BERT [5]) settings. The broader success of task-focused systems in institutional contexts, including legal advisory platforms [13] and public administration tools [14], further motivates this direction. Nevertheless, despite applications of language processing to intelligence analysis and military logistics [15], domain-adaptive pretraining for military discourse remains unexplored — a gap especially pronounced for Vietnamese given its Sino-Vietnamese formal register and Vietnam People's Army-specific doctrine.

Beyond domain alignment, structured reasoning has been elicited through supervised fine-tuning [16] and Chain-of-Thought prompting [17]. However, no existing Vietnamese instruction-tuning dataset accounts for the hierarchical logic and procedural constraints inherent to military doctrine — a limitation that the present work directly addresses.

3. Methodology

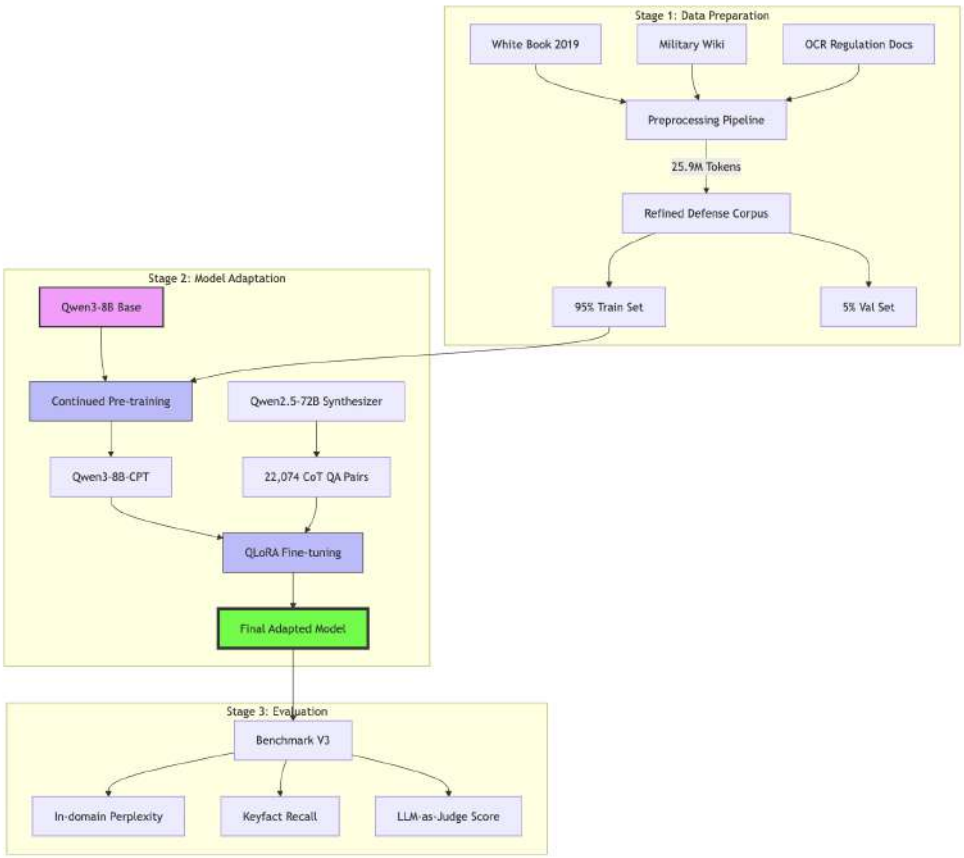


Figure 1. The end-to-end pipeline for adapting Qwen3-8B to the Vietnamese defense domain, featuring Domain-Adaptive Pretraining (DAPT) and Chain-of-Thought Supervised Fine-Tuning (CoT SFT).

The proposed adaptation framework for the Vietnamese defense domain follows a systematic three-stage pipeline: (i) construction of a high-quality defense corpus; (ii) Domain-Adaptive Pretraining (DAPT) to align the model with specialized terminology; and (iii) Knowledge Distillation through Supervised Fine-Tuning (SFT) with conditional reasoning paths. The overall architecture is illustrated in Figure 1.

3.1. Data Collection and Preprocessing

The Vietnamese defense corpus (25.9M tokens) is integrated from three specialized sources:

- **Defense News:** 2,829 unique articles from official portals (mod.gov.vn, bqp.vn), filtered via MD5 and MinHash LSH (threshold 0.85).

- **Military Wikipedia:** 4,555 curated articles on Vietnamese military history and units, extracted via category-based traversal (depth 5).

- **Regulation Documents:** 1,290 digitized legal documents processed through a dual-engine OCR pipeline with strict quality gates for formal markers.

Table 1. Corpus statistics after preprocessing.

Source	Content Type	Chunks	Tokens (approx.)
Defense News	Policy updates, military diplomacy	2,829	4.0M
Wikipedia (vi)	Military history, units, campaigns	4,555	9.1M

Regulation docs (OCR)	Procedures, legal norms, org. structures	1,290	12.8M
Total		8,674	25.9M

Token counts estimated at 1.4 tokens/word.

While user-generated encyclopedic content carries inherent quality risks, our Wikipedia subset is strictly limited to 4,555 articles within Vietnamese military history categories via controlled category-based traversal (depth 5), subject to the same MinHash LSH deduplication (threshold 0.85) and cleaning pipeline as the full corpus. Wikipedia serves a supplementary role providing historical background, while the authoritative core comprises official government portal documents and digitized military regulations. Future work should replace this subset with additional verified sources such as National Assembly defense reports and Ministry of National Defense circulars.

Preprocessing Pipeline: The merged corpus undergoes a five-stage refinement pipeline:

1. Source-specific Normalization: Wikipedia articles are normalized to Unicode NFC. OCR documents undergo specialized error correction to repair repeated syllables and punctuation spacing common in digitized text.

2. Cross-source Deduplication: A second round of MinHash LSH (word trigrams, threshold 0.85) is performed across the entire merged set to eliminate any latent overlap between sources.

3. Text Cleaning: Removal of zero-width characters, whitespace normalization, and standardization of quotation marks.

4. Strategic Chunking: We employ two distinct strategies:

- *News & Regulations:* Concatenated with EOS tokens and split into fixed-length sequences of 2,048 tokens.

- *Wikipedia:* Paragraph-boundary-aware splitting (512–8,192 characters) to preserve narrative coherence. This stage yielded 31,355 total chunks.

5. Stratified Split: The final dataset is divided into 95 % training (29,756 chunks) and 5 % validation (1,599 chunks), ensuring balanced source representation.

3.2. Continued Pre-Training (CPT)

Base model. We use Qwen3-8B [3] as our base model. Qwen3-8B is an 8-billion parameter decoder-only transformer with a 32K-token context window, pretrained on a large multilingual corpus, and demonstrates strong performance on Vietnamese benchmarks while being feasible to train on two A100 80GB GPUs.

Training setup. CPT is conducted on two NVIDIA A100 80GB GPUs using PyTorch Distributed Data Parallel (DDP). We employ a causal language modeling objective (next-token prediction) with a peak learning rate of 2×10^{-5} and a cosine decay schedule. To maintain hardware efficiency while processing 2,048-token sequences, gradient checkpointing and fused AdamW optimizers are enabled. The total training duration is 3 epochs, focusing on reducing in-domain perplexity (PPL):

$$PPL(X) = \exp\left(-\frac{1}{N}\sum_{i=1}^{N}\log P(x_i|x_{<i})\right) \quad (1)$$

Table 2. CPT hyperparameter configuration.

Hyperparameter	Value
Base model	Qwen3-8B (8B params)
Epochs	3
Peak learning rate	2×10^{-5}
Optimizer	AdamW fused (β_1=0.9, β_2=0.95)
Precision / Context	BF16 / 2,048 tokens

3.3. Supervised Fine-Tuning with Chain-of-Thought

SFT data generation. To construct the SFT dataset, we prompt Qwen2.5-72B-Instruct [18] to generate question-answer pairs from the training corpus using a structured chain-of-thought template: responses are formatted as <think> *reasoning trace* </think> *answer*. This format encourages explicit intermediate reasoning steps [17].

Raw generated pairs (24,527 items) are passed through a three-tier quality filtering pipeline:

- **PASS** (~60 %): items meeting all quality criteria including minimum reasoning length (>200 characters) and connective reasoning words;

- **FIX** (~30 %): items requiring programmatic reformatting (truncated answers, missing think tags);

- **FAIL** (~10 %): items exhibiting trivial patterns, direct context copying, or insufficient reasoning — discarded.

The final SFT dataset contains **22,074 QA pairs** after filtering.

SFT training. We apply QLoRA [19] to adapt the CPT model with a modest number of trainable parameters. Key hyperparameters:

Table 3. SFT (QLoRA) hyperparameter configuration.

Hyperparameter	Value
Method	QLoRA (4-bit, r=64, α=128)
Target modules	q_proj, k_proj, v_proj, o_proj, gate_proj, up_proj, down_proj
Epochs / LR	3 / 5×10^{-5}
Per-device batch size	4
Loss masking	Completion-only (prompt tokens masked)

We apply selective think-token masking: for *fact*-type questions, the <think> block is masked and only the final answer contributes to the loss. For *reasoning*, *application*, and *tactical* questions, the full chain-of-thought is included in the loss, teaching the model to generate domain-appropriate reasoning chains in formal Vietnamese.

3.4. Benchmark V3 Construction

Existing Vietnamese QA benchmarks are unsuitable for defense-domain reasoning. We construct Benchmark V3 with three test sets: **Set A** (322 auto-generated items, stratified by skill and difficulty), **Set B** (50 human-crafted gold-standard items with manually written reference answers), and **Set C** (30 OOD items covering general history, science, and daily tasks). Items

span five skill categories: Knowledge Retrieval, Tactical Reasoning, Historical Analysis, Legal/IHL Risk, and Factual Recall (402 total; distribution in Table 4). Evaluation uses Perplexity (PPL), Keyfact Recall, and LLM-as-Judge scoring (0–5 scale via Qwen2.5-72B-Instruct [20]), with contamination verified via ROUGE-L and BERTScore. Due to computational constraints, full quantitative evaluation on V3 is deferred to future work; this study uses V3 for qualitative analysis and reports preliminary metrics on Benchmark V2.

Table 4. Benchmark V3 distribution by test set and skill category.

Skill Category	Set A (Med)	Set A (Hard)	Set B (Med)	Set B (Hard)	Set C (OOD)	Total
Knowledge Retrieval	55	55	10	5	—	125
Tactical Reasoning	18	22	8	5	—	53
Historical Analysis	28	22	8	5	—	63
Legal/IHL Risk	15	15	7	5	—	42
Factual Recall	27	25	8	2	—	62
Out-of-domain (OOD)	—	—	—	—	30	30
Set total	143	139	41	22	30	402

4. Results and Discussion

We evaluate three model variants: **M1** = Qwen3-8B base; **M2** = Qwen3-8B-CPT; **M3** = Qwen3-8B-CPT-SFT.

4.1. Perplexity Evaluation (Intrinsic)

We evaluate perplexity on two corpora: (1) the in-domain validation set (1,599 held-out chunks, ~1.3M tokens; 500 sampled sequences); (2) a general Vietnamese corpus comprising 200 sequences from the Vietnamese Wikipedia dump (20231101.vi), containing no defense-related content. The catastrophic forgetting (CF) score is defined as the ratio of general-corpus PPL after adaptation to the baseline:

$$CF = \frac{PPL_{general}(M_2 \text{ or } M_3)}{PPL_{general}(M_1)} \quad (2)$$

Values < 1.0 indicate general-language *improvement*; values > 1.0 indicate forgetting.

Table 5. Perplexity evaluation results.

Model	Description	In-domain PPL	General PPL	CF Score
M1: Qwen3-8B-Base	No adaptation	5.81	9.34	1.000 (baseline)
M2: Qwen3-8B-CPT	+CPT on defense corpus	*3.64 (−37.4 %)*	7.83	0.839
M3:Qwen3-8B-CPT-SFT	+SFT with CoT QA data	*3.82 (−34.3 %)*	8.12	0.870

CPT (M2) reduces in-domain PPL by 37.4% (5.81 → 3.64), confirming the DAPT hypothesis [11] and matching gains reported in medical DAPT [4,12]. Notably, CF = 0.839 indicates that general Vietnamese proficiency is not only preserved but *improved* (9.34 → 7.83), attributable to the high-quality Vietnamese text in the CPT corpus. M3 shows a slight PPL increase (3.64 → 3.82, +5.0%) — expected, as SFT shifts token distribution toward conversational patterns rather than document continuations [16] — while CF = 0.870 confirms preserved general capability.

4.2. Benchmark V3: Qualitative Analysis and Case Studies

Since the primary goal of this foundational study is to establish a robust base for Vietnamese defense AI, we perform a qualitative analysis of Benchmark V3 to demonstrate the linguistic and logical complexity required for this domain. Table 6 showcases representative examples from the human-crafted Set B (Gold Standard).

Table 6. Representative examples from Benchmark V3 across key skill categories.

Category	Example Question	Target Reasoning Pattern
Tactical Reasoning	Phân tích lợi thế của phòng ngự chiều sâu trong địa hình rừng núi...	Application of terrain advantages to defense doctrine.
Legal/IHL Risk	Xác định trách nhiệm của chỉ huy khi xảy ra vi phạm quy tắc giao chiến (ROE)...	Interpretation of legal liability within VPA regulations.
Historical Analysis	So sánh nghệ thuật quân sự trong chiến dịch X và chiến dịch Y...	Multi-hop reasoning across historical timelines.

Preliminary observations of model outputs during the CPT phase indicate a significant improvement in the formal register. While the base model (M1) often resorted to generic or informal Vietnamese, the adapted model (M2) consistently utilized appropriate Sino-Vietnamese military terminology (e.g., "tác chiến hiệp đồng", "thế trận lòng dân") even before the final SFT stage. Furthermore, qualitative assessment of M3 (SFT) outputs on Benchmark V3 reveals that the model successfully internalizes complex hierarchical reasoning structures required for Tactical Reasoning and Legal/IHL Risk assessments, validating the effectiveness of the Conditional CoT Masking approach.

4.3. Discussion and Error Analysis

4.3.1. Linguistic Alignment vs. Parametric Memory Limits

The 37.4% perplexity reduction confirms that Vietnamese defense texts have a highly distinct distributional signature. However, manual error analysis reveals a critical divergence between linguistic capability and factual accuracy: the 8B-parameter model frequently hallucinates specific dates, unit designations, and regulatory clauses (Table 7), as parametric memory cannot reliably store the density of specialized facts required by this domain.

Table 7. Error analysis of factual hallucinations in CoT outputs (M3) due to parametric memory limitations.

Error Type	**Example Prompt**	**M3 Hallucinated Output (Translated/Summarized)**	**Fact-Check / Ground Truth**
Factual Mix-up	Trình bày diễn biến chính của Chiến dịch Điện Biên Phủ trên không (1972).	M3 describes the campaign but confidently states: *"Quân đội ta đã sử dụng tên lửa phòng không S-300 để bắn rơi B-52."*	Severe historical inaccuracy. In 1972, Vietnam only used SAM-2 missiles. The S-300 system was equipped much later (2003).
Fabricated Citation	Trích dẫn quy định về thời gian huấn luyện chiến sĩ mới theo Luật Nghĩa vụ Quân sự.	<think> Cần tìm quy định... Theo Điều 45, Thông tư 102/BQP... </think> *"Thời gian huấn luyện chiến sĩ mới là 6 tháng..."*	Completely fabricated circular number and content. Actual basic training time for new recruits in the VPA is 3 months.

4.3.2. The Trade-off Between Evasion and Elaborate Reasoning

The preliminary LLM-as-a-Judge evaluation on Benchmark V2 reveals a nuanced trade-off in generative behavior (Table 8).

Table 8. Preliminary LLM-as-Judge evaluation on Benchmark V2 (580 items, 1–5 scale, higher is better).

Criterion	**M1**	**M2 (+ CPT)**	**M3 (+ SFT)**
Accuracy	3.753	3.627	3.240
Completeness	3.736	3.627	3.343
Hallucination Resistance	—	—	3.820
Reasoning Quality	3.903	3.781	4.171
Clarity	4.475	4.364	4.760
Domain Fit	4.207	4.057	4.761
Overall	3.807	3.678	3.348

While the base model (M1) achieves a seemingly higher Accuracy score (3.753), qualitative inspection shows that M1 frequently resorts to generic, evasive answers when confronted with complex military concepts. By avoiding specific claims, M1 mathematically minimizes its error rate against the provided reference answers.

In contrast, the SFT-adapted model (M3) actively engages with the prompt, utilizing specific doctrinal concepts and explicit reasoning chains. This results in superior Domain Fit (+13.2 %) and Reasoning Quality (+6.9 %). Furthermore, when evaluated against ground-truth reference answers, M3 demonstrates a solid Hallucination Resistance score of 3.820. This indicates that its generations remain contextually anchored to realistic military domains rather than fabricating nonsensical responses. The drop in M3's Accuracy (3.240) is largely attributable to the "verbosity penalty": the model's extended Chain-of-Thought format generates significantly more verifiable claims, inherently increasing the surface area for minor factual mismatches against concise reference answers.

4.3.3. Toward Verifiable Reasoning Architectures

These findings strongly suggest that DAPT and CoT alone cannot resolve the factual reliability bottleneck in military AI applications. The M3 model developed in this study serves as an ideal, domain-aligned natural language generator. Its full potential will be unlocked when coupled with a Retrieval-Augmented Generation (RAG) pipeline, explicitly grounding its strong reasoning capabilities in an external, verifiable database of official military documents.

4.3.4. Practical Application Scenarios: Motivating the RAG + DAPT Paradigm

To concretize the complementary roles of DAPT and RAG, consider three representative scenarios within publicly accessible policy domains: conscription deferral queries, military personnel insurance guidance, and local militia regulations. In each case, the domain-adapted model (M3) provides the necessary linguistic and reasoning scaffolding to interpret complex policy questions in the appropriate military register, while RAG supplies the verifiable factual anchors — retrieving exact clauses from instruments such as Luật Nghĩa vụ Quân sự 2015 and Luật Dân quân tự vệ 2019 — that parametric memory alone cannot reliably provide. These

scenarios illustrate a consistent pattern: DAPT and RAG are not alternatives but mutually necessary components of a production-ready Vietnamese defense policy assistant.

5. Conclusion

In this foundational study, we presented a systematic domain-adaptive pretraining pipeline to adapt the Qwen3-8B model to the Vietnamese defense and military domain. By curating a 25.9-million-token specialized corpus and applying a dual-stage CPT and CoT-SFT strategy, we achieved a 37.4 % reduction in in-domain perplexity without catastrophic forgetting (CF = 0.839).

Crucially, our analysis highlights the fundamental limitations of relying on parametric memory for factual recall in high-stakes domains. While the adaptation process yields a model capable of structuring complex military logic and utilizing precise terminology, it remains susceptible to factual hallucinations. Future work will focus on integrating this domain-adapted LLM into a Retrieval-Augmented Generation (RAG) framework, shifting the paradigm from purely parametric generation to verifiable, fact-grounded military reasoning.

References

[1] OpenAI. (2024). GPT-4 technical report. arXiv. https://doi.org/10.48550/arXiv.2303.08774

[2] Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., ... Lample, G. (2023). LLaMA: Open and efficient foundation language models. arXiv. https://doi.org/10.48550/arXiv.2302.13971

[3] Yang, A., Li, A., Yang, B., Zhang, B., Hui, B., Zheng, B., ... Qiu, Z. (2025). Qwen3 technical report. arXiv. https://doi.org/10.48550/arXiv.2505.09388

[4] Singhal, K., Azizi, S., Tu, T., Mahdavi, ... Natarajan, V. (2022). Large language models encode clinical knowledge. arXiv. https://doi.org/10.48550/arXiv.2212.13138

[5] Chalkidis, I., Fergadiotis, M., Malakasiotis, P., Aletras, N., & Androutsopoulos, I. (2020). LEGAL-BERT: The muppets straight out of law school. arXiv. https://doi.org/10.48550/arXiv.2010.02559

[6] Bộ Quốc phòng nước Cộng hòa xã hội chủ nghĩa Việt Nam. (2019). *Quốc phòng Việt Nam 2019* (Sách trắng Quốc phòng Việt Nam 2019).

[7] Nguyen, D. Q., & Nguyen, A. T. (2020). PhoBERT: Pre-trained language models for Vietnamese. arXiv. https://doi.org/10.48550/arXiv.2003.00744

[8] Phan, L., Tran, H., Nguyen, H., & Trinh, T. H. (2022). ViT5: Pretrained text-to-text transformer for Vietnamese language generation. arXiv.

[9] Nguyen, Q., Pham, H., & Dao, D. (2023). VinaLLaMA: LLaMA-based Vietnamese foundation model. arXiv. https://doi.org/10.48550/arXiv.2312.11011

[10] Zhang, W., Chan, H. P., Zhao, Y., Aljunied, M., Wang, J., Liu, C., ... Bing, L. (2024). SeaLLMs 3: Open foundation and chat multilingual large language models for Southeast Asian languages. arXiv. https://doi.org/10.48550/arXiv.2407.19672

[11] Gururangan, S., Marasović, A., Swayamdipta, S., Lo, K., Downey, D., ... Smith, N. A. (2020). Don't stop pretraining: Adapt language models to domains and tasks. In D. Jurafsky, J. Chai, N. Schluter, & J. Tetreault (Eds.), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 8342–8360). Association for Computational Linguistics.

[12] Luo, R., Sun, L., Xia, Y., Qin, T., Zhang, S., Poon, H., ... Liu, T.-Y. (2022). BioGPT: Generative pre-trained transformer for biomedical text generation and mining. *Briefings in Bioinformatics*, *23*(6), bbac409. https://doi.org/10.1093/bib/bbac409

[13] Cui, J., Li, Z., Yan, Y., Chen, B., & Yuan, L. (2023). ChatLaw: Open-source legal large language model with integrated external knowledge bases. arXiv. https://doi.org/10.48550/arXiv.2306.16092

[14] Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial intelligence and the public sector — Applications and challenges. International Journal of Public Administration, 42(7), 596–615. https://doi.org/10.1080/01900692.2018.1498103

[15] Allen, G. C., & Chan, T. (2017). Artificial intelligence and national security. Belfer Center for Science and International Affairs, Harvard Kennedy School.

[16] Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., ... Lowe, R. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, *35*, 27730–27744. https://arxiv.org/abs/2203.02155

[17] Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., ... Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, *35*, 24824–24837.

[18] Yang, A., Yang, B., Hui, B., Zheng, B., Yu, B., Zhou, C., ... Fan, Z. (2024). Qwen2 technical report. arXiv. https://doi.org/10.48550/arXiv.2407.10671

[19] Dettmers, T., Pagnoni, A., Holtzman, A., & Zettlemoyer, L. (2023). QLoRA: Efficient finetuning of quantized LLMs. arXiv. https://doi.org/10.48550/arXiv.2305.14314

[20] Zheng, L., Chiang, W.-L., Sheng, Y., Zhuang, S., Wu, Z., Zhuang, Y., ... Stoica, I. (2023). Judging LLM-as-a-judge with MT-Bench and Chatbot Arena. In *Advances in Neural Information Processing Systems 36: Datasets and Benchmarks Track*. https://doi.org/10.48550/arXiv.2306.05685

Huấn luyện thích nghi miền mô hình ngôn ngữ lớn mã nguồn mở cho văn bản quốc phòng và quân sự tiếng Việt

Tóm tắt Mặc dù các mô hình ngôn ngữ lớn (LLM) đã đạt được hiệu suất tối tân trong các tác vụ tổng quát, khả năng ứng dụng chuyên biệt của chúng trong các lĩnh vực khan hiếm tài nguyên – đặc biệt là quốc phòng và quân sự Việt Nam – vẫn bị hạn chế đáng kể bởi khoảng cách lớn về dữ liệu. Trong bài báo này, chúng tôi trình bày một quy trình huấn luyện thích nghi miền (DAPT) có hệ thống nhằm tối ưu hóa mô hình Qwen3-8B cho lĩnh vực quốc phòng Việt Nam thông qua việc xây dựng một kho ngữ liệu chuyên ngành quy mô 25,9 triệu token từ các văn bản chính sách, Wikipedia chọn lọc và điều lệnh quân sự được số hóa qua OCR. Phương pháp nghiên cứu bao gồm giai đoạn tiếp tục tiền huấn luyện (CPT) kết hợp với khung tinh chỉnh có giám sát dựa trên chuỗi suy luận (SFT) sử dụng kỹ thuật Che giấu CoT có điều kiện nhằm ưu tiên các tín hiệu lập luận phức tạp. Kết quả đánh giá nội tại cho thấy giai đoạn CPT đã căn chỉnh thành công mô hình với miền mục tiêu, giảm 37,4% độ nhiễu (perplexity) trong khi vẫn bảo toàn năng lực tiếng Việt tổng quát (CF = 0,839); đồng thời, giai đoạn SFT giúp mô hình nội tâm hóa hiệu quả các cấu trúc lập luận phân cấp và sử dụng thành thạo văn phong quân sự chính quy cùng thuật ngữ tác chiến. Cuối cùng, bài báo này thiết lập một bộ sinh ngôn ngữ mạnh mẽ hoạt động hoàn toàn trên các tài liệu công khai, tạo tiền đề không gian biểu diễn (latent space) tối ưu để tích hợp với các kiến trúc Sinh văn bản tăng cường truy xuất (RAG), phục vụ các ứng dụng thực tiễn như tư vấn chính sách nghĩa vụ quân sự và giải đáp văn bản pháp quy trong tương lai.

Từ khóa: *huấn luyện thích nghi miền; mô hình ngôn ngữ lớn; văn bản quốc phòng; tinh chỉnh mô hình; đánh giá perplexity*

Deep Learning-Based Tiny Object Detection in RGB-IR UAV Imagery: Methods, Challenges, and Perspectives

Thi Lan Nguyen [1*], Cao Truong Tran[1]

[1] *Institute of Information and Communication Technology, Le Quy Don Technical University*

[*]Email: lannt.simtech@lqdtu.edu.vn, Contact number: 0961700989

Abstract

Tiny object detection (TOD) in UAV imagery is important for many aerial applications, but remains challenging due to limited pixel information, complex backgrounds, and poor illumination. To address these issues, deep learning-based RGB-infrared (RGB-IR) fusion has been widely explored to improve visibility and strengthen tiny-object representation. However, a dedicated survey on TOD in RGB-IR UAV imagery is still lacking. This paper provides a comprehensive review of recent deep learning methods for RGB-IR TOD in UAV scenarios. We first define the problem and analyze its key characteristics and challenges in multispectral aerial imagery. We then review representative benchmark datasets and categorize existing detection methods in a structured way. Finally, we discuss current limitations and highlight promising future directions. This survey aims to provide a useful reference for researchers and practitioners and to support further progress in RGB-IR TOD for UAV systems.

Keywords: *Tiny object detection, deep learning, multi-modal fusion, image fusion*

1. Introduction and Survey Scope

1.1. Introduction

UAVs have become important platforms for visual sensing in applications such as traffic monitoring, disaster response, environmental inspection, border surveillance, and search-and-rescue [1, 2]. However, tiny object detection in UAV imagery remains difficult because targets often occupy only a few pixels and appear under clutter, scale variation, motion blur, atmospheric interference, and poor illumination [3]. To address the limitations of RGB-only images, recent studies have explored deep learning-based RGB-IR fusion, which combines appearance cues from RGB with thermal information from IR to improve the visibility and robustness of tiny object detection in UAV applications [4, 5, 6].

Although related studies have grown rapidly, existing surveys still examine this topic from separate perspectives rather than as a unified problem. General UAV vision surveys rarely address RGB-IR fusion for tiny object detection [1, 3], tiny object detection surveys mostly focus on single-modal or general settings [4, 5], and visible-infrared fusion reviews mainly discuss fusion architectures or general RGB-IR detection rather than tiny objects in UAV imagery [6, 7]. As a result, this intersection still lacks a dedicated and systematic review.

Motivated by this gap, this survey provides a focused review of deep learning-based tiny object detection in RGB-IR UAV imagery. It examines the problem from a unified perspective that connects tiny-target representation, cross-modal fusion, and deployment efficiency in UAV settings. Accordingly, the survey not only summarizes representative methods and datasets, but also clarifies their strengths, limitations, and practical trade-offs, thereby providing a more targeted basis for understanding recent progress and future research directions.

The main objectives and contributions of this survey are summarized as follows:

- We formally define the problem setting and analyze the intrinsic characteristics and technical challenges of multispectral aerial tiny objects.

- We review and categorize representative RGB-IR datasets, multispectral fusion strategies, and tiny object-oriented learning techniques.

- We analyze the limitations of current methods and highlight future research directions toward robust, efficient, and real-time UAV perception.

1.2. Survey Methodology and Scope

To improve the systematic nature of this survey, the reviewed literature was selected based on three criteria: topic relevance, technical specificity, and experimental informativeness. We focused on studies related to UAV-based object detection, tiny/small object detection, RGB-IR or visible-thermal fusion, multispectral aerial perception, and lightweight deployment-oriented detection. Priority was given to recent deep learning-based works that provide clear methodological contributions, publicly described datasets or benchmark settings, and sufficient experimental evidence for qualitative or quantitative comparison.

The survey is organized into four main parts. First, we define the problem setting and summarize the core challenges of tiny object detection in RGB-IR UAV imagery. Second, we review representative methods from both single-modal and RGB-IR perspectives, with an emphasis on tiny-object representation, fusion design, cross-modal discrepancy mitigation, and efficiency-oriented detection. Third, we summarize commonly used RGB-IR UAV datasets and benchmark characteristics. Finally, we discuss current limitations and highlight future research directions toward robust, efficient, and deployable RGB-IR tiny object detection systems.

2. Background and Problem Definition

2.1. Definition of Tiny Objects in UAV Imagery

In UAV image analysis, tiny objects are defined mainly by how small they appear in the image rather than by their real-world physical size. Because UAVs capture large scenes from elevated viewpoints, many targets occupy only a few pixels, making conventional object-size definitions less suitable. As a result, tiny objects in UAV imagery are commonly described using pixel-based criteria [7, 8]. Following MS COCO, small objects are usually defined as having an area no larger than 32×32 pixels, but in tiny object detection (TOD), this threshold is often tightened to below 16×16 pixels because visual information becomes extremely limited at that scale [9]. Some UAV-oriented datasets further divide this range into finer categories to better reflect different levels of visual difficulty [10, 11].

However, pixel size alone is not always sufficient, especially in high-resolution UAV images where an object may contain more pixels but still remain very small relative to the whole scene. For this reason, relative-size criteria are also widely used. In many studies, tiny objects are defined as occupying less than 0.1% of the image area, with even smaller thresholds adopted in more extreme high-altitude scenarios [12, 13]. Overall, tiny objects in UAV imagery are generally understood as targets that are very small in both absolute pixel size and relative scene proportion, which explains why they are so difficult to detect reliably.

2.2. Core Challenges in RGB-IR Tiny Object Detection

RGB-IR UAV imagery poses distinctive challenges because it combines aerial imaging characteristics with cross-modal heterogeneity. UAVs usually capture large scenes from elevated viewpoints, so objects of interest occupy only a few pixels and are easily overwhelmed by background clutter. At the same time, RGB and IR sensors differ in signal characteristics, contrast, noise patterns, and spatial resolution, and their outputs are often imperfectly aligned due to sensor placement and calibration differences. These properties lead to three main challenges in RGB-IR tiny object detection [7, 8].

The first challenge is robust cross-modal fusion. RGB images provide rich texture and color information, while IR images offer more stable thermal cues under poor lighting. However, their heterogeneous properties often cause models to favor one modality over the other [14]. This problem becomes more severe when RGB and IR images are spatially misaligned (Fig.1), since even a one- or two-pixel offset can seriously disrupt feature correspondence for tiny objects and lead to poor fusion, inaccurate localization, or missed detections [11, 15]. Therefore, effective fusion must not only exploit complementary information but also preserve weak tiny-object cues [7].

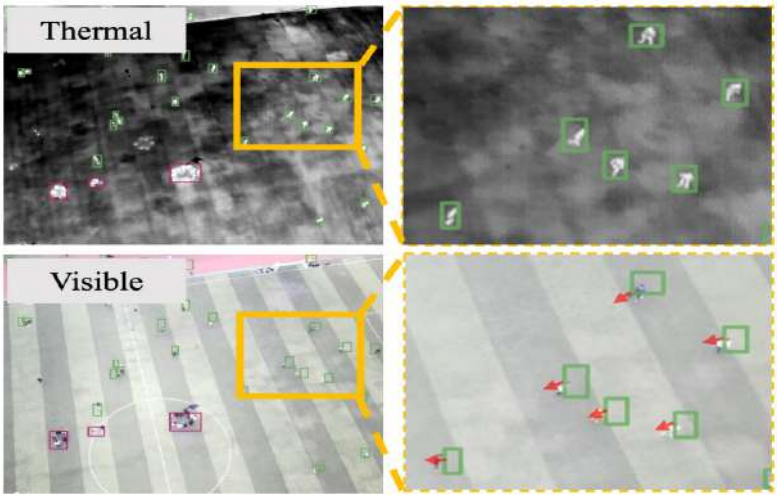


Figure 1. Spatial misalignment between RGB and IR imagery

The second challenge is scale consistency. UAV images often contain objects of very different sizes in the same scene [11]. Methods designed to enhance tiny targets may weaken the structural information needed for larger ones, while scale-agnostic designs often fail to retain the fine details required for tiny-object discrimination [12, 15]. As a result, maintaining balanced performance across scales remains difficult.

The third challenge is efficiency. Many RGB-IR fusion methods rely on heavy attention modules or multi-branch architectures, which increase computation, memory usage, and inference latency. This is problematic for UAV deployment, where onboard resources are limited and real-time processing is often required [13]. Meanwhile, tiny objects are highly sensitive to spatial errors and precision loss, so lightweight design must be achieved without sacrificing localization accuracy [16].

Overall, RGB-IR tiny object detection in UAV imagery is mainly limited by fusion difficulty, scale inconsistency, and strict efficiency constraints, all of which arise from the intrinsic properties of RGB-IR aerial data.

3. Tiny Object Detection Methods in UAV Imagery

3.1. Single-modal Tiny Object Detection in UAV Imagery

3.1.1. Feature Representation for Tiny Objects

Feature representation learning is a key challenge in tiny object detection, especially in UAV imagery where targets often occupy fewer than 16×16 pixels. At this scale, object cues are extremely weak, easily buried in background clutter, and further degraded by repeated downsampling in deep networks [17]. To address this, multi-scale feature integration is widely used. Feature Pyramid Networks (FPNs) [18] combine high-resolution details from shallow layers with stronger semantics from deeper layers, but standard FPNs are designed for general multi-scale detection and do not explicitly highlight the subtle edges and local structures needed for tiny objects. Recent methods therefore introduce detail-enhancement or frequency-aware modules to strengthen high-frequency cues and suppress background noise with limited extra cost [9, 19].

Another direction is super-resolution, which aims to recover details lost during downsampling by enhancing images or intermediate features, often through GAN-based reconstruction [13, 20]. However, SR usually increases computation and may introduce artifacts, so its gains are not always stable [12]. In addition, representation learning is affected by the mismatch between pre-training datasets and UAV data, since models trained on datasets dominated by medium-sized objects often struggle with tiny targets. Scale alignment methods help reduce this gap by adjusting object size distributions during training, allowing networks to learn features that are more suitable for tiny objects [21].

3.1.2. Context and Attention Enhancement

Because tiny objects contain very limited visual information, detectors often cannot rely on object features alone. Contextual cues from surrounding regions therefore play an important role in improving detection reliability [22]. Local context helps refine boundaries and reduce localization errors by enlarging the effective receptive field, while global context provides scene-level semantic guidance when object evidence is weak or incomplete. Together, these cues help detectors model meaningful object-environment relationships, such as vehicles on roads or ships on water, and compensate for the lack of intrinsic detail in tiny objects [9, 13].

Attention mechanisms further improve this process by highlighting informative features and suppressing background noise. Convolution-based modules such as CBAM strengthen discriminative responses through channel and spatial attention [8, 23], while transformer-based methods capture longer-range dependencies, although global self-attention may dilute tiny-object signals because background regions often dominate the computation. Hierarchical designs such as Swin Transformer provide a more balanced solution through window-based attention [23]. More recently, bio-inspired attention and top-down suppression strategies have also been explored to enhance target-background contrast and emphasize weak fine-grained cues in complex UAV scenes [24, 25].

3.1.3. Optimization and Detection Head Design

Accurate tiny-object localization depends not only on feature representation but also on similarity measurement and sample assignment during training. Conventional IoU is often unreliable for tiny objects, since even a slight pixel shift can cause a large overlap drop and

unstable optimization [10, 12]. To address this, recent studies use distribution-based metrics, such as Normalized Wasserstein Distance [26], Gaussian receptive-field-based assignment [27], and center-distance measures, while methods like SAFit [28] combine IoU for large objects with distribution-based similarity for tiny ones to improve robustness across scales.

Detection head design is also critical. Since conventional heads from P3 to P7 may lose the weak details of tiny objects through repeated downsampling, recent methods often introduce higher-resolution heads, especially P2, to better preserve fine spatial cues [23]. More advanced designs, such as transformer-based heads in TPH-YOLOv5 [29] and shallow-feature enhancement in Bottom Focus-PAN [30], further improve localization, but usually at the cost of higher complexity.

3.2. RGB-IR Tiny Object Detection Methods

3.2.1. Multi-modal Representation and Fusion

Tiny object detection in UAV imagery increasingly benefits from multi-modal data, especially RGB and IR images, because the two modalities provide complementary information. RGB images contain rich texture and color cues, while IR images are more robust under poor illumination, nighttime conditions, fog, and background clutter. By combining them, multi-modal detectors can reduce the weaknesses of single-modal methods and improve tiny-object visibility in challenging environments [11].

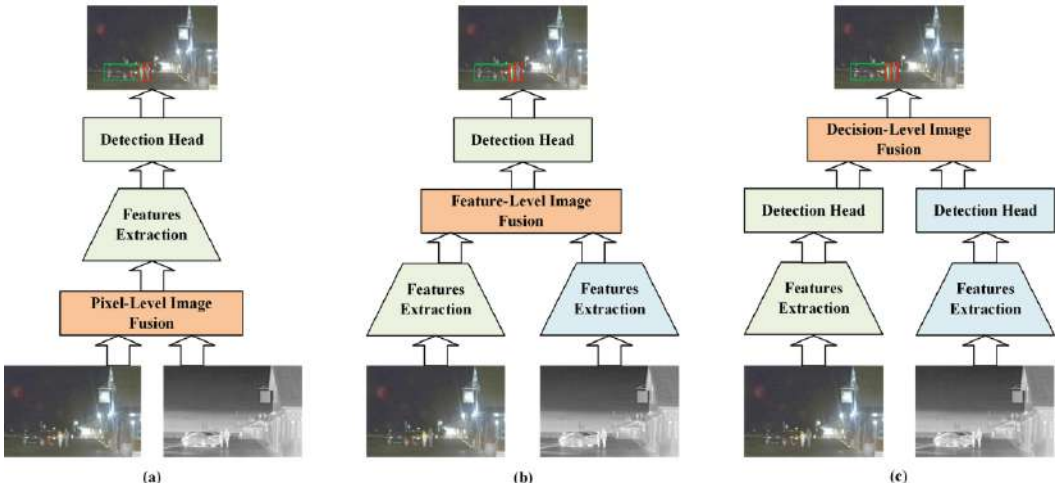


Fig.2. Three ways of fusing visible and infrared images in multimodal object detection task: (a) show pixel-level fusion, (b) show feature-level fusion, (c) show decision-level fusion.

According to where fusion is performed, existing RGB-IR methods can be grouped into three levels (Figure 2). Image-level fusion combines modalities directly at the input, which preserves raw spatial details and is computationally efficient, but it is sensitive to noise and offers limited semantic interaction. Feature-level fusion, now the most common strategy, extracts modality-specific features first and then fuses them within the network, providing a better balance between representation quality and efficiency. Decision-level fusion combines outputs from separate detectors, which avoids early feature interference but often causes redundant computation and limited cross-modal interaction [31].

Recent studies further explore cross-layer and bi-directional fusion to better support tiny objects. These designs combine shallow high-resolution details with deeper semantic features across modalities, and extend top-down fusion with bottom-up information flow. Such multi-

level interaction helps reduce signal loss during deep transmission and improves the representation and localization of tiny objects [15, 31].

3.2.2. Cross-modal Discrepancy Mitigation

In multi-modal UAV tiny object detection, effective fusion requires reducing the discrepancy between RGB and IR images. Because the two sensors differ in position, hardware settings, and imaging mechanisms, RGB-IR pairs often exhibit geometric and semantic inconsistencies [8]. Spatial misalignment is the most serious issue, especially for tiny objects that occupy only a few pixels. Even a one- or two-pixel offset can break cross-modal correspondence, cause bounding box confusion, and reduce fusion quality and detection accuracy [11, 32].

To address this problem, recent studies increasingly use learnable feature-level alignment instead of traditional image registration. Methods such as DASR [33] employ adaptive alignment modules to refine feature positions and improve correspondence between RGB and IR features, while SWCA uses shifted-window cross-attention to model semantic alignment and estimate larger offsets more accurately. These designs help preserve fine details and improve robustness under illumination changes and partial occlusion.

Besides alignment, bounding box correction has become an effective training-time solution for supervision inconsistency in RGBT datasets. Because annotations are often provided only for one reference modality, labels in the other modality may be noisy due to misalignment. Cross-modality Bounding Box Correction addresses this issue by treating these errors as label noise and progressively refining boxes for the misaligned modality with a Mean Teacher framework. Since this process is applied only during training, it improves multi-modal learning without increasing inference cost [32].

Table 1 summarizes representative RGB-IR UAV tiny object detection methods from the perspectives of fusion strategy and cross-modal discrepancy mitigation.

3.2.3. Efficiency-oriented Multi-modal Detection

In practical UAV applications, deploying multi-modal tiny object detectors on edge devices requires a careful balance between accuracy and efficiency. UAV platforms are limited by battery capacity, memory, and onboard computing power, while processing high-resolution RGB and IR data together significantly increases computation, memory usage, and latency. As a result, achieving reliable tiny object detection in real time has become a key goal in multi-modal UAV perception [16, 23].

To address this issue, many studies design lightweight architectures and efficient attention mechanisms. Compact backbones such as MobileNet and ShuffleNet, together with depthwise separable convolutions, inverted residual blocks, and linear bottlenecks, help reduce parameters and computation while preserving feature extraction ability. For attention design, lightweight channel-wise mechanisms are often preferred over more expensive spatial attention modules [7]. More recently, state-space models such as Mamba [34] have emerged as efficient alternatives to Transformers, offering linear complexity while still supporting effective multi-modal interaction.

Knowledge distillation provides another effective solution. By transferring knowledge from a large teacher network to a lightweight student model, KD improves performance without

increasing inference cost. In multi-modal detection, methods such as Cross-Modality High-Resolution Distillation [35] use a high-resolution RGB teacher to guide a compact student operating on low-resolution thermal data. With suitable distillation losses, this strategy helps the student learn fine structural details and cross-modal correlations, improving tiny object detection while maintaining low latency and compact model size.

Table 1. Representative RGB-IR tiny object detection methods on RGBTDronePerson

Method	Fusion level	Core idea	Result (mAP50)	Remark
QFDet	Feature-level	Quality-aware feature fusion	36.24	The first large baseline model for RGBTDronePerson.
COXNet	Feature-level	Wavelet-based cross-layer fusion	39.74	Balances semantic accuracy and spatial detail through pixel-level alignment.
DEGF-YOLO	Feature-level	Gated dual-modal interaction	38.20	Lightweight design through competitive-collaborative cross-modal interaction.
AODet	Feature-level	Occlusion-aware RGB-IR detectio	34.97	Specifically targets the challenge of occluded objects in UAV scenarios.
BDDS	Feature-level,	Bidirectional multi-level fusion	46.72	Reduces information loss and contamination through comprehensive multi-level interaction

4. RGB-IR Tiny Object Datasets

4.1. Specialized Tiny Object Datasets

This group includes highly challenging RGB-IR UAV datasets specifically designed for tiny object detection, where tiny targets dominate the data distribution and create severe feature scarcity and signal-to-noise imbalance. ***RGBT-Tiny*** [10], the first large-scale benchmark for multi-modal tiny object detection, contains 115 paired RGB-IR video sequences, about 93,000 frames, and over 1.2 million annotations across seven categories, with more than 81% of objects smaller than 16×16 pixels; despite homography-based alignment, residual parallax errors remain. ***RGBTDronePerson*** [11] further emphasizes the difficulty of tiny pedestrian detection, containing 6,125 paired RGB-IR images and 70,880 annotations, with an average object size of only 11.7 pixels and 98.0% of targets smaller than 20 pixels. It is especially challenging because prominent RGB-IR position shifts can easily cause tiny targets to be missed, making these datasets essential benchmarks for evaluating tiny-aware and robust fusion methods.

4.2. Related and Multi-scale RGB-IR Datasets

This group includes RGB-IR UAV datasets that follow the COCO definition of small objects (smaller than 32×32 pixels) but also contain many medium and large targets, making them useful for evaluating scale robustness beyond extreme tiny-object settings. ***VTUAV-detv*** [11], built by re-annotating the VTUAV tracking benchmark for detection, contains 16,770 paired RGB-IR images and 124,869 human annotations. About 53.87% of objects are small, while the rest are medium or large, providing clear scale diversity. Captured at altitudes of 5-20 meters with both nadir and oblique views, it offers better visual detail but also stronger scale

variation. ***DroneVehicle*** [36] contains 28,439 paired RGB-IR images for vehicle detection in both day and night scenes. It uses oriented bounding boxes to better represent vehicle orientation in aerial views. Although many objects are small, they are generally clearer than those in RGBTDronePerson. The dataset also presents severe RGB-IR misalignment, making it useful for studying orientation-aware detection and fusion robustness.

5. Open Challenges and Future Research Directions

5.1. Tiny-aware Fusion under Cross-modal Heterogeneity

Although many RGB-IR fusion methods have been proposed, preserving tiny-object cues remains difficult because RGB and IR data differ greatly in appearance, contrast, and noise [8, 15]. Existing methods often use uniform fusion or global attention, which tends to emphasize dominant modalities or large background regions, causing weak tiny-object signals to be suppressed, especially under low signal-to-noise conditions [14]. Future research should therefore focus on tiny-aware fusion, such as quality-aware and frequency-guided designs, to better preserve subtle details while reducing cross-modal interference in UAV detection.

5.2. Robust Detection across Extreme Object-Scale Variation

UAV imagery contains large scale variation, where tiny objects often appear together with medium and large ones in the same scene [13]. Although many detectors are designed for tiny-object detection, their performance across scales is often unstable: improving sensitivity to tiny objects can hurt larger-object detection, while scale-robust designs may weaken fragile tiny-object features. This problem becomes more difficult after RGB-IR fusion, which can further disturb feature distributions across scales. Future research should therefore focus on lightweight, scale-robust representations, such as local refinement and adaptive feature modulation, to preserve feature stability and improve detection consistently across different object sizes [13].

5.3. Lightweight and Deployable RGB-IR Detection Models

Balancing detection accuracy with the limited computation of UAV platforms remains a key challenge in RGB-IR tiny object detection. Many effective fusion methods rely on heavy components such as global attention, cross-modal alignment, and multi-branch structures, which increase parameters, memory use, and inference time, making real-time deployment difficult. This issue is even more critical for tiny object detection, which often requires high-resolution features or extra refinement. Future work should therefore focus on lightweight, deployment-oriented designs by simplifying cross-modal interaction, reducing expensive modules, and combining efficient backbones, local refinement, model compression, and hardware-aware optimization to achieve a better balance between accuracy and deployment practicality [7, 15].

6. Conclusion

Tiny object detection in RGB-IR UAV imagery remains a challenging task. This survey reviews recent deep learning methods from both single-modal and multi-modal perspectives, covering feature representation, attention, fusion, alignment, and lightweight design, while also summarizing common RGB-IR UAV datasets and clarifying the definition and difficulty of tiny objects. Despite recent progress, major challenges remain, including weak object features, scale variation, cross-modal misalignment, and limited computational resources. Future research

should therefore focus on more reliable RGB-IR fusion, scale-robust models, and lightweight architectures suitable for real-time UAV deployment, as well as better datasets, more accurate annotations, and standardized evaluation protocols to support fair comparison and stronger generalization.

References

1. X. Wu, W. Li. Deep learning for unmanned aerial vehicle-based object detection and tracking: A survey. *IEEE Geoscience and Remote Sensing Magazine*, 10(1), 2021, pp. 91–124.
2. N. Merkle, et al. Drones4Good: Supporting disaster relief through remote sensing and AI. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 3770–3774.
3. L. P. Osco, J. M. Junior. A review on deep learning in UAV remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, 102, 2021, Article 102456.
4. M. Yuan, X. Wei. C²Former: Calibrated and complementary transformer for RGB–infrared object detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2024.
5. L. Zongzhen, L. Hui. Cross-modal offset-guided dynamic alignment and fusion for weakly aligned UAV object detection. *arXiv preprint arXiv:2506.16737*, 2025.
6. X. He, C. Tang, X. Zou. Multispectral object detection via cross-modal conflict-aware learning. In *Proceedings of the 31st ACM International Conference on Multimedia*, 2023, pp. 1465–1474.
7. M. Luo, R. Zhao, S. Zhang, L. Chen, F. Shao, X. Meng. IM-CMDet: An intra-modal enhancement and cross-modal fusion network for small object detection in UAV aerial RGBT imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 2025.
8. Y. Tang, X. Tong, Z. He, J. Yang. Adaptive modality selection drone-based RGBT detector for tiny targets. *IEEE Transactions on Instrumentation and Measurement*, 2025.
9. Z. Zou, K. Chen, Z. Shi, Y. Guo, J. Ye. Object detection in 20 years: A survey. *Proceedings of the IEEE*, 111(3), 2023, pp. 257–276.
10. X. Ying, et al. Visible-thermal tiny object detection: A benchmark dataset and baselines. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025.
11. Y. Zhang, C. Xu, W. Yang, G. He, H. Yu, L. Yu. Drone-based RGBT tiny person detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, 204, 2023, pp. 61-76.
12. M. Muzammul, X. Li. Comprehensive review of deep learning-based tiny object detection: Challenges, strategies, and future directions. *Knowledge and Information Systems*, 2025, pp. 1-89.
13. W. Hua, Q. Chen. A survey of small object detection based on deep learning in aerial images. *Artificial Intelligence Review*, 58(6), 2025, pp. 1–67.
14. Y. Gu, W. Chen, D. Peng. UAV-based multimodal object detection via feature enhancement and dynamic gated fusion. *Pattern Recognition*, 2025, Article 112722.
15. P. Peng. CoxNet: Cross-layer fusion with adaptive alignment and scale integration for RGBT tiny object detection. *IEEE Transactions on Circuits and Systems for Video Technology*, 2025.
16. J. Li, C. Fan, C. Ou, H. Zhang. Infrared and visible image fusion techniques for UAVs: A comprehensive review. *Drones*, 9(12), 2025, Article 811.
17. *Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2117–2125.
18. Z. Shi, J. Hu, J. Ren. Hs-FPN: High frequency and spatial perception FPN for tiny object detection. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 39, 2025, pp. 6896–6904.
19. J. Zhang, J. Lei. SuperYOLO: Super-resolution assisted object detection in multimodal remote sensing imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 61, 2023, pp. 1–15.
20. Y. Shi, Y. Jia, X. Zhang. FocusDet: An efficient object detector for small objects. *Scientific Reports*, 14(1), 2024, Article 10697.

21. X. Yu, Y. Gong, N. Jiang. Scale match for tiny person detection. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2020, pp. 1257–1265.
22. M. Nikouei, B. Baroutian. Small object detection: A comprehensive survey on challenges, techniques and real-world applications. *arXiv preprint arXiv:2503.20516*, 2025.
23. Z. Li, Y. Wang, H. Feng. Local to global: A sparse transformer-based small object detector for remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 2025.
24. S. Woo, J. Park, J.-Y. Lee, I. S. Kweon. CBAM: Convolutional block attention module. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 3–19.
25. X. Zhang, Y. Feng, N. Wang, G. Lu, S. Mei. Aerial person detection for search and rescue: Survey and benchmarks. *Journal of Remote Sensing*, 2025, Article 0474.
26. Z. Liu, Swin Transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 10012–10022.
27. S. Huang, C. Lin, X. Jiang, Z. Qu. Brstd: Bio-inspired remote sensing tiny object detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2024.
28. J.-M. Hopf. Direct neurophysiological evidence for spatial suppression surrounding the focus of attention in vision. *Proceedings of the National Academy of Sciences*, 103(4), 2006, pp. 1053–1058..
29. T.-Y. Lin. Microsoft COCO: Common objects in context. In *Computer Vision – ECCV 2014: 13th European Conference, Zurich, Switzerland, Part V*, Springer, 2014, pp. 740–755.
30. Y. Zeng, D. Guo, W. He. ARF-YOLOv8: A novel real-time object detection model for UAV-captured images detection. *Journal of Real-Time Image Processing*, 21(4), 2024, Article 107.
31. X. Zhu. TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 2778–2788.
32. X. Jin, C. Guo, Z. He, J. Xu. FCMNet: Frequency-aware cross-modality attention networks for RGB-D salient object detection. *Neurocomputing*, 491, 2022, pp. 414-425.
33. Y. Sun, Y. Meng. Visible and infrared image fusion for object detection: A survey. In *International Conference on Image, Vision and Intelligent Systems*, Springer, 2023, pp. 236–248.
34. Y. Zhang, H. Yu, Y. He. Illumination-guided RGBT object detection with inter- and intra-modality fusion. *IEEE Transactions on Instrumentation and Measurement*, 72, 2023, pp. 1–13.
35. Z. Tang, Y. Xie, T. Xu, X.-J. Wu. Learning bi-directional fusion and deformation-sensitive loss for RGB-T tiny object detection. *Information Fusion*, 2025, Article 103985.
36. Y. Sun. Drone-based RGB–infrared cross-modality vehicle detection via uncertainty-aware learning. *IEEE Transactions on Circuits and Systems for Video Technology*, 2022, pp. 6700–6713.

Phát hiện vật thể nhỏ trong ảnh UAV RGB-IR dựa trên học sâu: Phương pháp, thách thức và triển vọng

Tóm tắt: *Phát hiện vật thể nhỏ trong ảnh UAV là bài toán quan trọng nhưng khó do ít thông tin điểm ảnh, nền phức tạp và điều kiện chiếu sáng bất lợi. Để cải thiện hiệu quả phát hiện, nhiều phương pháp hợp nhất ảnh màu và hồng ngoại dựa trên học sâu đã được đề xuất. Tuy nhiên, vẫn thiếu một khảo sát hệ thống tập trung riêng vào bài toán này trong ảnh UAV RGB-IR. Bài báo này tổng quan các phương pháp học sâu hiện đại, phân tích đặc trưng và thách thức chính, giới thiệu các bộ dữ liệu tiêu biểu, phân loại các hướng tiếp cận hiện có, đồng thời chỉ ra hạn chế và các hướng nghiên cứu triển vọng trong tương lai. Khảo sát hướng tới việc làm tài liệu tham khảo hữu ích cho nghiên cứu và ứng dụng trong lĩnh vực này*.

Từ khoá: Phát hiện vật thể nhỏ, học sâu, hợp nhất đa phương thức, hợp nhất ảnh.

❑

Domain-Adaptive Pretraining for a Vietnamese Medical Large Language Model

Quang Tung Pham[1], Cao Truong Tran[1*]

[1]*Cơ quan Le Quy Don Technical University*

Abstract

Large Language Models (LLMs) have shown strong potential in healthcare applications, but their effectiveness in non-English, low-resource settings such as Vietnam remains constrained by domain shift, corpus scarcity, and terminology inconsistency. This paper presents a reproducible two-stage adaptation pipeline for building a Vietnamese medical LLM, consisting of Domain-Adaptive Pretraining (DAPT) followed by mixed Supervised Fine-Tuning (SFT). Stage~I performs mixture-based DAPT using a curated in-house Vietnamese medical corpus as the primary training signal, complemented by controlled English biomedical augmentation and Vietnamese general-domain replay to improve conceptual coverage and reduce overspecialization. Stage~II applies mixed SFT combining Vietnamese medical instruction data with distilled reasoning supervision from a stronger teacher model (DeepSeek-R1). A corpus-aware diagnostic step based on BERTopic topic clustering is used to validate specialty coverage and detect noisy subsets prior to training. Evaluation on the VM14K benchmark (12,488 questions across 34 specialties) and on a focused VMLU medical-subject subset probe shows that DAPT yields substantial domain-specific gains over the base model, while mixed SFT provides further improvements, particularly on difficult reasoning subsets. The resulting pipeline offers a practical framework for adapting open-weight LLMs to high-stakes, low-resource medical NLP settings.

Keywords: *domain-adaptive pretraining; DAPT; large language models; Vietnamese medical NLP; supervised fine-tuning; reasoning distillation*

1. Introduction

Large language models (LLMs) have shown strong performance in medical tasks such as question answering, evidence summarization, and clinical assistance [1]. However, medical applications are inherently high-stakes: model errors - including hallucinations or unsafe recommendations - can lead to serious consequences, requiring careful evaluation and governance [1]. In practice, general LLMs often exhibit domain mismatch when applied to specialized medical language, motivating explicit domain adaptation strategies [2,3].

Vietnamese medical NLP faces additional constraints. First, Vietnamese is widely cosidered a lower-resource language for biomedical and clinical NLP compared to English, with limited high-quality, structured, and legally shareable corpora [4]. Second, a significant portion of publicly available Vietnamese biomedical text has been created via large-scale translation from English biomedical sources (e.g., PubMed), which can introduce translation artifacts and terminology inconsistencies that may propagate into pretrained representations [4, 5]. Third, available instruction data often provide limited supervision for clinically relevant reasoning, reducing models' ability to justify answers concisely and reliably - a gap that is amplified in the Vietnamese setting where few reasoning-oriented medical benchmarks exist [6-10].

To address these gaps, this paper proposes a two-stage training pipeline targeting models in the 8–10B parameter range—a scale where domain adaptation yields the most practical gains,

* Email: truongct@lqdtu.edu.vn

as very large models, while potentially more accurate, incur prohibitive inference costs for real-world clinical deployment. Specifically: (i) Stage I (DAPT) performs continued causal language modeling on a curated Vietnamese medical corpus, and (ii) Stage II (Mixed SFT) applies instruction tuning on a mixture of non-reasoning medical instructions and distilled reasoning examples generated by a stronger teacher model (DeepSeek-R1). An empirical study is designed to isolate the contributions of each stage.

Our main contributions are:

• A reproducible pipeline for building a Vietnamese medical corpus for DAPT, including data cleaning, structuring, and coverage analysis.

• A mixed SFT strategy that combines instruction tuning with distilled reasoning supervision to address the lack of reasoning-focused Vietnamese medical data.

• An empirical evaluation showing that DAPT and mixed SFT improve performance on VM14K and a focused VMLU medical-subject subset [2, 3, 11].

2. Related Work

2.1. DAPT/TAPT and medical adaptation for biomedical language models

Adaptive pretraining is now a standard approach for domain specialization. Prior work [2] shows that domain-adaptive pretraining (DAPT) and task-adaptive pretraining (TAPT) consistently improve downstream performance across domains. In biomedical NLP, BioBERT [3] shows that continued pretraining on PubMed and PMC improves biomedical NER, relation extraction, and question answering, while PubMedBERT [11] suggests that training from scratch can be even more effective when large in-domain corpora are available. In Vietnamese, however, biomedical resources remain limited and often depend on translated corpora such as ViPubMed [4, 5], making continued pretraining a more practical choice.

Recent medical LLMs typically use multi-stage post-training pipelines that combine biomedical continued pretraining, instruction tuning, and careful evaluation. Med-PaLM 2 [1] is a representative example, showing that stronger in-domain language modeling alone is not sufficient without alignment and safety-oriented assessment. Open-source models such as Meditron [12], PMC-LLaMA [13], and BioMedLM [14] likewise demonstrate that biomedical adaptation can significantly improve the medical capability of general LLM backbones.

Within these pipelines, parameter-efficient methods such as LoRA [15] and adapters [16] are widely adopted because they reduce computational cost and support modular adaptation.

2.2. Datasets and Vietnamese Medical Resources

Vietnamese medical LLM research currently depends on two complementary categories of resources: translated biomedical corpora for continued pretraining and native task datasets for supervised training and evaluation. ViPubMed [4], released as VietAI/vi_pubmed [5], provides a large translated Vietnamese biomedical corpus for domain-adaptive pretraining, though filtering is needed to reduce translation artifacts and terminology drift. The same project also introduces VietAI/vi_mednli [9], extending Vietnamese medical resources to inference-oriented tasks.

Several downstream datasets support fine-tuning and evaluation. ViHealthNLI [10] provides a healthcare-specific NLI benchmark, ViMedAQA [6] targets abstractive medical question answering, ViMQ [7] supports dialogue-oriented evaluation, and Vietnamese-Medical-QA [8] is useful for instruction tuning despite variable evidence quality and informal language.

This paper uses VM14K [17] as the primary benchmark. VM14K is the first large-scale expert-annotated Vietnamese medical evaluation set, containing 12,488 multiple-choice questions across 34 specialties and four difficulty levels. We also use an auxiliary probe from two medical-related VMLU subjects [18]—Clinical Pharmacology and Internal Basic Medicine, totaling 351 questions. Since this subset covers only two subjects, it should be interpreted as a limited indicator of general-language retention rather than a full estimate. Overall, the heterogeneity of these resources, especially translation artifacts in pretraining corpora and variable evidence quality in web-derived QA data, makes strict preprocessing and source-aware sampling essential in Vietnamese medical DAPT pipelines [4, 8].

3. Methodology

This study adopts a two-stage adaptation pipeline: Stage I adapts the base model to Vietnamese medical language via Domain-Adaptive Pretraining (DAPT), and Stage II aligns the DAPT checkpoint to instruction-following via mixed Supervised Fine-Tuning (SFT). Figure 1 illustrates the full pipeline from corpus collection and preprocessing to DAPT, mixed SFT, and benchmark evaluation.

3.1. Vietnamese Medical Corpus Construction and Mixture Design for DAPT

The DAPT stage is designed as a mixture-based continued pretraining procedure. The core signal comes from a curated Vietnamese medical corpus constructed in-house, supplemented by controlled amounts of general-domain Vietnamese text and English biomedical text to balance domain specialization, linguistic stability, and conceptual coverage.

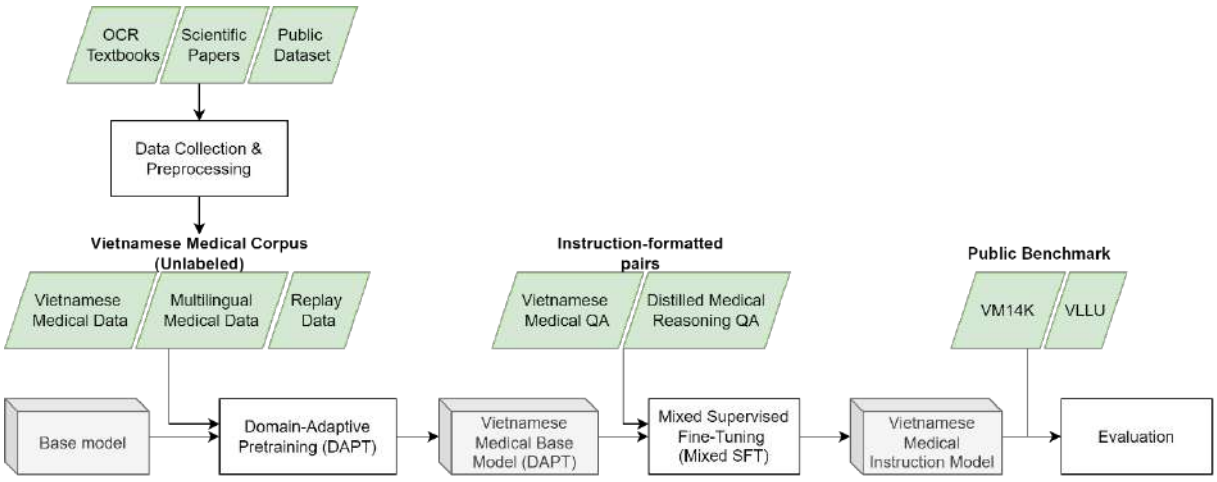


Figure 1. Overview of the proposed two-stage adaptation pipeline. Raw Vietnamese medical sources are collected and normalized into a structured DAPT corpus, followed by continued pretraining (Stage I), mixed supervised fine-tuning (Stage II), and evaluation on VM14K plus an auxiliary VMLU medical-subject subset probe.

Core Vietnamese biomedical corpus (A). The primary component is a heterogeneous Vietnamese medical corpus spanning textbooks, clinical guidelines, research articles from Vietnamese medical journals, and OCR-extracted scanned PDFs (Table 1). Raw documents were normalized, de-noised, segmented, and deduplicated with source-level provenance retained for source-aware sampling and leakage auditing; full preprocessing details are provided in the supplementary material.

Table 1 summarizes statistics of the in-house Vietnamese biomedical corpus after preprocessing, grouped into four source categories.

Table 1. Summary statistics of the in-house Vietnamese biomedical corpus after preprocessing. Token counts use the corpus audit heuristic (chars/3). Rows within a group are indented and sum to the group subtotal.

Source group	Sub-source	Docs	Tokens (M)
Textbooks / lectures	Vietnamese originals	59	16.5
	Translated textbooks	40	17.1
	Subtotal	*99*	*33.6*
Clinical guidelines	Tutorial / guideline PDFs	51	13.0
Research / papers	11 Vietnamese medical journals	8,243	37.2
OCR scanned PDFs	GoodScan (high quality)	90	23.1
	ScanWithOCR (medium)	84	21.6
	FairScan (lower quality)	62	14.3
	Subtotal	*236*	*59.0*
Total		**8,629**	**142.9**

General-domain Vietnamese replay (C). To reduce overspecialization toward highly structured textbook-style language, we incorporate the Vietnamese subset of CulturaX as a replay buffer, maintaining syntactic diversity and stabilizing general language modeling behavior at a limited proportion to preserve the medical-domain signal.

English biomedical augmentation (A'). Given scarce high-quality native Vietnamese clinical text, we introduce a controlled proportion of English biomedical text (PubMed abstracts and medical wiki-style documents) to enhance conceptual and terminological coverage without shifting the model's primary Vietnamese-language alignment.

Mixture ratios. Table 2 reports the mixture configuration: Vietnamese biomedical corpus (70%), PubMed abstracts (15%), medical wiki-style corpus (5%), and CulturaX general replay (10%). Together, this mixture balances in-domain linguistic alignment (Vietnamese biomedical), general-language stability (replay), and conceptual coverage (English biomedical) for Stage I continued pretraining.

3.2. Mixed SFT Data Construction and Training Data Governance

The SFT dataset combines two supervision types: non-reasoning Vietnamese medical instructions from QA/inference-style datasets, and distilled reasoning examples generated by

DeepSeek-R1 on a curated medical subset. The main mixture uses a 75/25 ratio, all samples are normalized into a unified schema and teacher outputs are filtered for quality.

Table 2. Mixture ratios used for Domain-Adaptive Pretraining (DAPT). Percentages sum to 100%.

Group	Dataset component	Mix ratio (%)
(A) Biomedical (vi)	In-house Vietnamese medical corpus	70
(C) General replay (vi)	CulturaX (Vietnamese subset)	10
(A') Biomedical (en)	PubMed abstracts	15
(A') Biomedical (en)	Medical Wiki-style corpus	5
Total		**100**

3.3. Two-Stage Training Procedure and Corpus-Aware Analysis

Both stages use LoRA. In Stage~I, the base model is continued-pretrained on the DAPT corpus with a standard causal language modeling objective, and checkpoints are selected on a held-out validation split. In Stage~II, the selected checkpoint is fine-tuned on the mixed SFT set using the same chat template, with loss computed only on assistant tokens. Tokenizer and decoding settings are kept consistent across stages.

Before training, we apply BERTopic-based corpus analysis to assess specialty coverage and identify noisy subsets. Fine-grained topics are discovered and merged into broader specialty groups for interpretability. This analysis supports quality control by detecting OCR noise, boilerplate, and malformed content, and helps explain weaker VM14K results when some specialties are underrepresented or noisy in the DAPT corpus. Results are summarized in a bar chart of clustered-chunk counts by merged topic (Figure 2).

4. Experiments

This section evaluates the two-stage pipeline under controlled conditions, isolating the contributions of DAPT, SFT, and DAPT mixture composition, and comparing the final model against representative baselines on VM14K.

4.1. Training Configuration and Reproducibility

All model variants share the same tokenizer, context length (4096 tokens), and evaluation template, with Qwen3-8B as the base model. Both stages use RS-LoRA ($r = 64$, $\alpha = 128$) applied to all projection and embedding modules (q/k/v/o_proj, gate/up/down_proj, embed_tokens, lm_head) for parameter efficiency. Training uses AdamW with a learning rate of $2 \times 10{-}5$, cosine decay schedule with 5% linear warmup, BF16 precision, and a per-device batch size of 4 with gradient accumulation of 32 steps over 2 epochs. Gradient checkpointing is enabled and distributed training uses DeepSpeed ZeRO-2 on 2×A100 GPUs (estimated wall-clock: ≈26 hours for DAPT and ≈14 h for SFT on a single A100). Checkpoints are selected on held-out validation loss and evaluated on VM14K and the auxiliary VMLU subset without benchmark-specific calibration

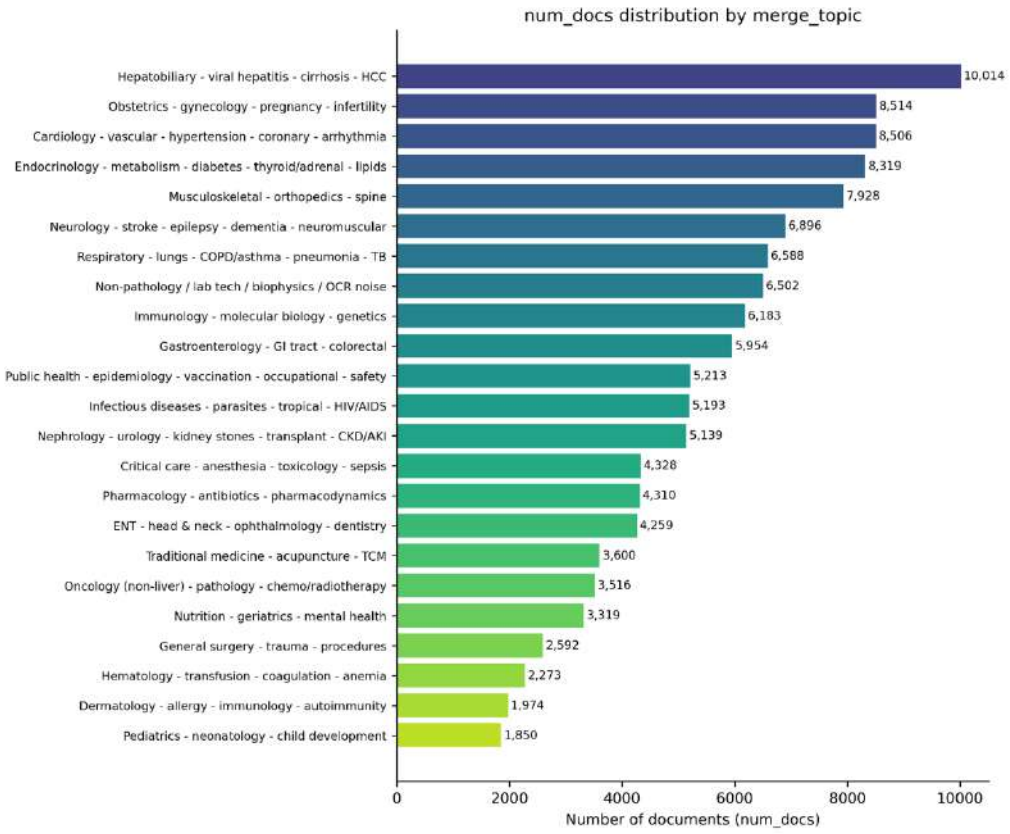


Figure 2. Clustered-chunk count (num_docs) by merged topic from the self-supervised clustering pipeline. This chart is used to assess specialty coverage, detect topic imbalance, and support interpretation of downstream performance differences across VM14K specialties.

4.2. Main Results on VM14K and Auxiliary VMLU Subset Probe

Table 3 reports the main VM14K results in the same difficulty-group format used throughout this study, including Overall, Easy, Medium, and Hard* (merged Challenging+Hard). Several trends are consistent across all groups. M1 substantially improves over M0, confirming that DAPT transfers Vietnamese medical-domain knowledge into the base model. M2 also improves over M0, but remains below M1, indicating that supervised tuning alone cannot replace domain-adaptive pretraining. M3 outperforms both M1 and M2, showing that DAPT and SFT contribute complementary gains. M4 achieves the strongest performance on all difficulty groups, with the largest improvement over M3 observed in the Hard* subset (i.e., Challenging + Hard), which is consistent with the intended effect of distilled reasoning supervision.

The gains are not uniform across difficulty groups. DAPT contributes strongly to Easy and Medium questions by improving domain vocabulary coverage and reducing confusion in terminology-heavy items. The additional gains from mixed SFT are most visible on Hard questions, where the model must perform multi-step reasoning under clinical constraints. This pattern supports the design choice of combining domain adaptation with a controlled amount of distilled reasoning supervision.

Auxiliary VMLU subset results: DAPT improves accuracy from 60.11% to 70.66% (+10.55 points) on the two medical-related subjects (351 questions total), indicating strong

knowledge transfer on a standardized probe outside VM14K. This result should be interpreted as a limited subject-level signal rather than a full-benchmark estimate of general-language retention.

Table 3: Main VM14K results by difficulty group. Values are accuracies (%) averaged over 3 runs. Hard merges the Challenging and Hard subsets in the evaluation split used in this study.*

Model	Overall	Easy	Medium	Hard*
M0 (Base)	41.6 0.4	48.9 0.5	40.8 0.4	30.2 0.5
M1 (Base + DAPT)	49.8 0.3	55.2 0.4	48.2 0.3	41.4 0.4
M2 (Base + SFT)	46.9 0.5	53.8 0.6	45.8 0.5	35.2 0.6
M3 (Base + DAPT + SFT-NR)	56.4 0.4	60.7 0.4	55.0 0.4	50.3 0.5
M4 (Base + DAPT + SFT-Mixed)	**58.9 0.3**	**62.6 0.3**	**57.7 0.4**	**53.8 0.4**

4.4. Comparison with External Baselines on VM14K

Table 4 compares M4 with representative proprietary and open-source baselines from the VM14K benchmark report [17] using single-run pass@1; external scores are treated as reference points because prompt formatting and decoding may differ across studies.

Table 4. Comparison with representative external baselines on VM14K (pass@1 accuracy, %). External baseline values are taken from the VM14K benchmark report. For comparability with our reporting format, results are presented using the same column structure (Overall, Easy, Medium, Hard), where Hard* denotes the merged Challenging+Hard subset following the benchmark split definitions.*

Model	Params (B)	Overall	Easy	Medium	Hard*
Open-source general models					
DeepSeek-R1	–	78.17	79.89	77.54	77.03
Qwen3-32B	32	72.47	74.37	72.10	68.79
Qwen3-30B-A3B	30	71.32	73.69	70.28	65.08
Llama 4 Maverick	–	71.76	72.95	71.54	69.66
Gemma3-27B-Instruct	27	63.38	66.46	62.50	53.99
Gemma3-12B-Instruct	12	58.05	60.60	57.39	50.84
Llama-3.1-8B-Instruct	8	48.73	52.33	47.36	42.08
Open-source medical-specialized models					
Llama-3-70B-UltraMedical	70	54.96	58.34	53.55	51.07
HuatuoGPT-o1-8B	8	54.52	57.93	53.14	52.19
Meditron3-70B	70	42.27	45.57	40.86	37.17
Ours (M4: DAPT + Mixed SFT)	8	**58.9**	**62.6**	**57.7**	**53.8**

The proposed M4 model is competitive with several open-source general and medical-specialized models and improves substantially over base open-source models in the same parameter range, supporting the practical value of targeted domain adaptation without scaling to very large models.

4.6. Ablation Studies

We conduct ablation experiments to quantify the sensitivity of the proposed pipeline to DAPT corpus scale, DAPT mixture composition, reasoning mixture ratio, and LoRA adapter capacity. The ablation results are reported using the same difficulty-group format (Overall, Easy, Medium, Hard*). To directly test the proposed mixture-based DAPT design, Table 6 compares four DAPT- only configurations (M1-style): using only the Vietnamese biomedical corpus, adding English biomedical augmentation, adding general-domain Vietnamese replay, and using the full mixture from Section 3. This ablation isolates Stage I effects without conflating them with SFT. The Vietnamese-only corpus already yields a strong gain over the base model, confirming it as the primary adaptation signal; adding English biomedical data further improves Medium and Hard subsets, consistent with better conceptual coverage for complex questions. Adding general-domain Vietnamese replay reduces overspecialization, and the full mixture (Vi biomedical + En biomedical + replay) achieves the best overall balance across all difficulty groups.

All DAPT-mixture ablations are run with the same optimizer configuration and matched total training token budget (M1-style) to isolate composition effects.

Table 6. Ablation results (II): DAPT mixture composition (DAPT-only, M1-style). VM14K values are accuracies (%). Hard merges the Challenging and Hard subsets.*

DAPT setting	Overall	Easy	Medium	Hard*
Vietnamese biomedical only	47.9	53.6	46.2	38.9
+ English biomedical augmentation	49.1	54.3	47.6	41.0
+ General-domain Vietnamese replay	48.6	54.1	47.0	39.8
Full mixture (Vi + En biomedical + replay)	**49.8**	**55.2**	**48.2**	**41.4**
Vietnamese biomedical only	47.9	53.6	46.2	38.9
+ English biomedical augmentation	49.1	54.3	47.6	41.0
+ General-domain Vietnamese replay	48.6	54.1	47.0	39.8
Full mixture (Vi + En biomedical + replay)	**49.8**	**55.2**	**48.2**	**41.4**

4.8. Discussion

DAPT and mixed SFT are complementary: DAPT provides the domain-adaptation foundation that SFT alone cannot replicate, while distilled reasoning supervision yields the largest additional gains on Hard subsets. The adapted 8B-scale model achieves a strong accuracy-to-parameter trade-off on VM14K, supporting practical deployment in resource-constrained Vietnamese medical NLP settings.

5. Conclusion

This paper presents a reproducible two-stage pipeline for building a Vietnamese medical LLM, combining domain-adaptive pretraining (DAPT) with mixed supervised fine-tuning (SFT) over factual instruction and distilled reasoning data. The framework emphasizes corpus quality, source-aware preprocessing, and benchmark decontamination for low-resource medical NLP.

Experiments show that DAPT and SFT play complementary roles: DAPT mainly improves domain adaptation, especially on terminology-rich questions, while mixed SFT brings further gains on harder reasoning subsets. The adapted 8B model also achieves a favorable parameter-performance trade-off on VM14K compared with several larger open-source models.Current limitations include reliance on a single base-model family, dependence on translated Vietnamese medical resources, and the gap between benchmark accuracy and real-world clinical safety. Future work should expand native Vietnamese clinical corpora, evaluate additional backbone architectures, and include expert-centered human evaluation.

References

1. Karan Singhal, Tao Tu, Juraj Gottweis, Rory Sayres, Ellery Wulczyn, , et al. Toward expert-level medical question answering with large language models. *Nature Medicine* , 2025.
2. Suchin Gururangan, Ana Marasović, Swabha Swayamdipta, Kyle Lo, Iz Beltagy, Doug Downey, and Noah A. Smith. Don't stop pretraining: Adapt language models to domains and tasks. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL)* , 2020.
3. Jinhuk Lee, Wonjin Yoon, Sungdong Kim, Donghyeon Kim, Sunkyu Kim, Chan Ho So, and Jaewoo Kang. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *arXiv preprint* , 2019.
4. Long Phan, Tai Dang, Hieu Tran, Trieu H. Trinh, Vy Phan, Lam D. Chau, and Minh-Thang Luong. Enriching biomedical knowledge for low-resource language through large-scale translation. In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics (EACL)* , pages 3131–3142, 2023.
5. VietAI. VietAI/vi_pubmed: Vietnamese translated biomedical abstracts (hugging face dataset). https://huggingface.co/datasets/VietAI/vi_pubmed , 2022. Dataset card.
6. Minh-Nam Tran, , et al. Vimedaqa: A vietnamese medical abstractive question-answering dataset. In *Proceedings of the ACL Student Research Workshop* , 2024.
7. Ta Duc Huy, Nguyen Anh Tu, Tran Hoang Vu, Nguyen Phuc Minh, Nguyen Phan, Trung H. Bui, and Steven Q. H. Truong. Vimq: A vietnamese medical question dataset for healthcare dialogue system development. *arXiv preprint* , 2023.
8. hungnm. hungnm/vietnamese-medical-qa: Vietnamese medical qa pairs (hugging face dataset). https://huggingface.co/datasets/hungnm/vietnamese-medical-qa , 2024. Dataset card.
9. VietAI. VietAI/vi_mednli: Vietnamese medical natural language inference (hugging face dataset). https://huggingface.co/datasets/VietAI/vi_mednli , 2022. Dataset card.
10. Hai Nguyen, Huyen Ngo, , et al. Vihealthnli: A dataset for vietnamese natural language inference in healthcare. In *Proceedings of SIGUL 2024* , 2024.
11. Yu Gu, Robert Tinn, Hao Cheng, Michael Lucas, Naoto Usuyama, Xiaodong Liu, Tristan Naumann, Jianfeng Gao, and Hoifung Poon. Domain-specific language model pretraining for biomedical natural language processing. *arXiv preprint* , 2020.

12. EPFL LLM Team. Meditron: Medical large language model (repository). https://github.com/epfLLM/meditron , 2024.
13. PMC-LLaMA Contributors. PMC-LLaMA: A domain-adapted LLaMA on biomedical literature (repository). https://github.com/chaoyi-wu/PMC-LLaMA , 2023.
14. Elliot Bolton, Abhinav Venigalla, Michihiro Yasunaga, David Hall, Betty Xiong, Tony Lee, Roxana Daneshjou, Jonathan Frankle, Percy Liang, Michael Carbin, and Christopher D. Manning. Biomedlm: A 2.7b parameter language model trained on biomedical text. *arXiv preprint* , 2024.
15. Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. *arXiv preprint* , 2021.
16. Neil Houlsby, Andrei Giurgiu, Stanisław Jastrzebski, Bruna Morrone, Quentin de Laroussilhe, Andrea Gesmundo, Mona Attariyan, and Sylvain Gelly. Parameter-efficient transfer learning for NLP. *arXiv preprint* , 2019.
17. Thong Nguyen, Duc Nguyen, Minh Dang, Thai Dao, Long Nguyen, Quan H. Nguyen, Dat Nguyen, Kien Tran, and Minh Tran. VM14K: First vietnamese medical benchmark. *arXiv preprint arXiv:2506.01305* , 2025.
18. Cuc Thi Bui et al, VMLU benchmarks: A comprehensive benchmark toolkit for Vietnamese LLMs. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* , pages 11495–11515, Vienna, Austria, 2025. Association for Computational Linguistics.

Huấn luyện thích nghi theo miền cho một mô hình ngôn ngữ lớn y khoa tiếng Việt

Tóm tắt

Các mô hình ngôn ngữ lớn (LLMs) cho thấy tiềm năng lớn trong các ứng dụng y tế, nhưng hiệu quả của chúng trong các bối cảnh phi tiếng Anh, ít tài nguyên như tiếng Việt vẫn còn hạn chế do khác biệt miền dữ liệu, khan hiếm ngữ liệu và sự thiếu nhất quán về thuật ngữ. Bài báo này đề xuất một quy trình thích nghi hai giai đoạn, có thể tái lập, nhằm xây dựng mô hình ngôn ngữ y khoa tiếng Việt, gồm: (1) huấn luyện thích nghi theo miền (Domain-Adaptive Pretraining - DAPT) và (2) tinh chỉnh có giám sát hỗn hợp (mixed Supervised Fine-Tuning - SFT). Ở Giai đoạn I, mô hình được DAPT trên một hỗn hợp dữ liệu, trong đó nguồn tín hiệu chính là ngữ liệu y khoa tiếng Việt nội bộ đã được tuyển chọn, kết hợp với phần bổ sung có kiểm soát từ dữ liệu y sinh học tiếng Anh và dữ liệu tiếng Việt miền tổng quát nhằm mở rộng độ bao phủ khái niệm và giảm hiện tượng chuyên biệt hóa quá mức. Ở Giai đoạn II, chúng tôi áp dụng SFT hỗn hợp bằng cách kết hợp dữ liệu hướng dẫn y khoa tiếng Việt với dữ liệu giám sát suy luận được chưng cất từ mô hình giáo viên mạnh hơn (DeepSeek-R1). Trước khi huấn luyện, một bước chẩn đoán ngữ liệu dựa trên phân cụm chủ đề BERTopic được sử dụng để kiểm tra độ bao phủ theo chuyên khoa và phát hiện các nhóm dữ liệu nhiễu. Mô hình được đánh giá trên bộ VM14K (12.488 câu hỏi thuộc 34 chuyên khoa) và trên một tập thăm dò gồm các môn liên quan đến y khoa trong VMLU. Kết quả cho thấy DAPT mang lại cải thiện đáng kể về năng lực chuyên miền so với mô hình gốc, trong khi SFT hỗn hợp tiếp tục nâng cao hiệu năng, đặc biệt ở các nhóm câu hỏi khó đòi hỏi suy luận nhiều hơn. Quy trình đề xuất cung cấp một khuôn khổ thực tiễn để thích nghi các mô hình ngôn ngữ mã nguồn mở cho các bài toán NLP y khoa có độ rủi ro cao trong bối cảnh ít tài nguyên.

Từ khóa: huấn luyện thích nghi theo miền; DAPT; mô hình ngôn ngữ lớn; NLP y khoa tiếng Việt; tinh chỉnh có giám sát; chưng cất suy luận

Một kết hợp phân cụm mờ cộng tác và mạng bộ mã hoá tự động để phân đoạn ảnh viễn thám siêu phổ

Tạ Văn Đắc[1], Đỗ Khắc Ngọc[1], Mai Đình Sinh[2]

[1]*Hệ QLHVSĐH;* [2]*Viện Kỹ thuật công trình đặc biệt/Đại học Kỹ thuật Lê Quý Đôn*

Email: vandachvkt@gmail.com; Tel: 0935731681

Tóm tắt

Ảnh viễn thám siêu phổ (HRSI) chứa đựng thông tin phổ-không gian phong phú, mang lại tiềm năng lớn cho việc lập bản đồ lớp phủ đất chi tiết và nhận diện đối tượng. Tuy nhiên, tính chiều cao và sự dư thừa phổ của chúng đặt ra những thách thức lớn đối với việc phân đoạn hiệu quả và chính xác. Để khắc phục những hạn chế này, bài báo này trình bày một phương pháp phân cụm mờ hợp tác kết hợp mạng Autoencoder (CFCM-AE) bằng cách kết hợp phương pháp phân cụm mờ C-means hợp tác với mạng Autoencoder cho phân đoạn ảnh siêu phổ. Trong khung này, Autoencoder thực hiện giảm chiều để giữ lại các đặc trưng phân biệt đồng thời giảm thiểu sự dư thừa và chi phí tính toán. Thành phần phân cụm mờ hợp tác sau đó tận dụng cả sự tương đồng phổ và sự phụ thuộc ngữ cảnh giữa các pixel, dẫn đến phân đoạn ổn định và mạnh mẽ hơn. Sự tích hợp này nâng cao khả năng thích ứng, cải thiện độ chính xác và giảm độ nhạy với nhiễu và dữ liệu chiều cao. Các đánh giá thực nghiệm trên các bộ dữ liệu siêu phổ chuẩn cho thấy phương pháp đề xuất vượt trội đáng kể so với phân cụm mờ truyền thống và các kỹ thuật giảm chiều độc lập, đạt được hiệu suất vượt trội về độ chính xác, độ ổn định và hiệu quả. Các kết quả xác nhận rằng sự kết hợp giữa phân cụm mờ hợp tác và Autoencoder cung cấp một giải pháp hiệu quả và có thể mở rộng cho phân tích ảnh viễn thám siêu phổ.

Từ khóa: *Phân cụm mờ cộng tác, giảm chiều, AutoEncoder, viễn thám siêu phổ.*

I. GIỚI THIỆU

Dữ liệu Viễn thám Siêu phổ (HSI), nhờ khả năng ghi nhận hàng trăm dải phổ liên tục trên phổ điện từ, đã trở thành một trong những công cụ quan sát Trái Đất mạnh mẽ nhất. Không giống như ảnh đa phổ truyền thống chỉ giới hạn ở một vài dải phổ rộng, ảnh siêu phổ (HSI) cung cấp thông tin phổ chi tiết, cho phép phân biệt chính xác các lớp bề mặt, loài thực vật, thành phần khoáng sản và các đối tượng nhân tạo. Với lợi thế này, HSI có giá trị đáng kể trong nhiều lĩnh vực ứng dụng, bao gồm nông nghiệp, lâm nghiệp, quy hoạch đô thị, thăm dò khoáng sản, giám sát môi trường, và thậm chí quốc phòng và an ninh. Tuy nhiên, tiềm năng này đi kèm với những thách thức trong xử lý dữ liệu do kích thước lớn và chiều của ảnh.

Một ảnh siêu phổ thường bao gồm hàng trăm dải phổ. Chiều lớn của dữ liệu không chỉ làm tăng độ phức tạp tính toán mà còn làm suy giảm hiệu suất của các thuật toán phân đoạn hoặc phân loại khi áp dụng trực tiếp lên tất cả các dải phổ. Ngoài ra, HSI thường chứa pixel hỗn hợp, biến động phổ phi tuyến tính và nhiễu cảm biến, làm phức tạp hóa quá trình phân đoạn. Do đó, việc phát triển các chiến lược giảm chiều dữ liệu và khai thác hiệu quả mối quan hệ phổ-không gian là yếu tố then chốt để cải thiện chất lượng phân đoạn.

Trong số các cách tiếp cận, phân cụm được sử dụng rộng rãi như một phương pháp học không giám sát để phân đoạn dữ liệu không nhãn thành các nhóm có ý nghĩa dựa trên sự tương đồng phổ. Trong số đó, phân cụm mờ đặc biệt được quan tâm nhờ khả năng xử lý sự không chắc chắn và gán mỗi pixel cho nhiều cụm với các mức độ thành viên khác nhau. Điều này phù hợp với đặc trưng của dữ liệu siêu phổ, nơi pixel hỗn hợp là tiêu chuẩn. Các thuật toán như Fuzzy C-means (FCM) và các biến thể của nó đã được nghiên cứu rộng rãi trong phân đoạn ảnh

siêu phổ. Tuy nhiên, FCM truyền thống vẫn đối mặt với nhiều hạn chế: nhạy cảm với nhiễu, tốn kém tính toán khi làm việc trong không gian chiều cao, và không khai thác đầy đủ thông tin ngữ cảnh không gian.

Để khắc phục hạn chế này, một số nghiên cứu gần đây đã phát triển các mô hình phân cụm hợp tác. Ý tưởng cốt lõi của phân cụm hợp tác là cho phép nhiều quá trình phân cụm hoạt động song song hoặc đồng thời trên các không gian đặc trưng khác nhau, sau đó tích hợp kết quả để tinh chỉnh việc gán nhãn pixel. Cơ chế này cho phép mô hình chia sẻ thông tin trong quá trình phân cụm, dẫn đến kết quả phân cụm toàn cục hiệu quả hơn.

Tuy nhiên, một thách thức lớn vẫn còn: làm thế nào để xử lý hiệu quả tính chiều cao của dữ liệu siêu phổ trước khi phân cụm. Nếu dữ liệu gốc được đưa trực tiếp vào CFCM, chi phí tính toán sẽ rất cao, và nguy cơ suy giảm chất lượng phân cụm do các đặc trưng không liên quan cũng tăng lên. Do đó, giảm chiều đã trở thành bước tiền xử lý không thể thiếu trong phân đoạn HSI.

Ngoài ra, trong những năm gần đây, nhiều phương pháp giảm chiều phi tuyến tính dựa trên học biểu diễn, chẳng hạn như Autoencoder trong học sâu, đã duy trì cấu trúc nội tại của dữ liệu trong không gian chiều thấp hơn. Kết hợp giảm chiều với phân cụm cho phép cả giảm gánh nặng tính toán và cải thiện chất lượng phân đoạn bằng cách tập trung vào các đặc trưng có sức mạnh phân biệt cao nhất.

Dựa trên các nền tảng trên, bài báo này đề xuất một phương pháp kết hợp phân cụm mờ hợp tác với Autoencoder cho phân đoạn ảnh viễn thám siêu phổ (CFCM-DR). Ý tưởng chính là trước tiên áp dụng AE để nén không gian phổ chiều cao thành biểu diễn gọn trong khi vẫn bảo toàn thông tin phổ-không gian quan trọng. Sau đó, CFCM được triển khai trên không gian đặc trưng đã giảm chiều; sự kết hợp của hai giai đoạn này hứa hẹn mang lại kết quả phân đoạn vừa chính xác vừa hiệu quả về tính toán.

Các đóng góp chính của bài báo là đề xuất một mô hình AE và CFCM tích hợp, trực tiếp giải quyết thách thức của dữ liệu siêu phổ chiều cao. Phương pháp khai thác cả thông tin phổ và không gian của dữ liệu để cải thiện chất lượng phân cụm. Các thực nghiệm trên một số bộ dữ liệu ảnh siêu phổ chuẩn chứng minh rằng phương pháp vượt trội hơn một số thuật toán phân cụm truyền thống hoặc giảm chiều riêng lẻ về độ chính xác và chi phí tính toán.

Bài báo được tổ chức thành năm phần: Phần 1 là phần giới thiệu; Phần 2 giới thiệu một số kiến thức liên quan; Phần 3 là phương pháp đề xuất; Phần 4 trình bày một số thực nghiệm; và Phần 5 đưa ra kết luận.

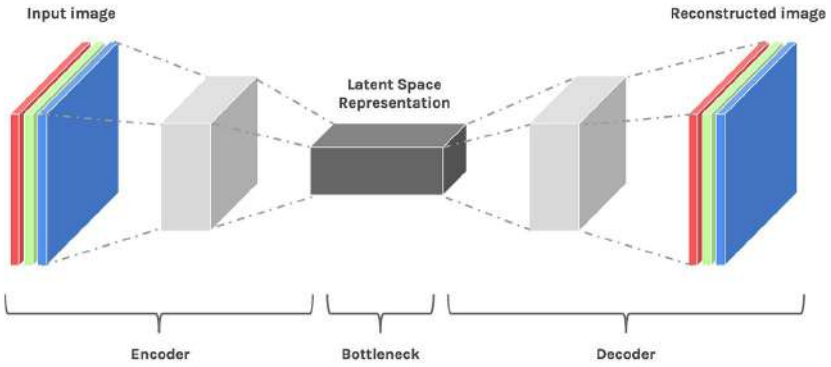
II. NỀN TẢNG

A. Autoencoder

Autoencoder (AE) là một loại mạng nơ-ron nhân tạo (ANN) được thiết kế cho học biểu diễn không giám sát của dữ liệu. Thay vì dự đoán nhãn đầu ra, autoencoder học cách nén dữ liệu đầu vào thành biểu diễn chiều thấp hơn (biểu diễn tiềm ẩn hoặc mã), sau đó tái tạo dữ liệu gốc từ biểu diễn này. Mục tiêu cốt lõi của autoencoder là tìm cách biểu diễn dữ liệu gọn trong khi vẫn bảo toàn các đặc trưng quan trọng.

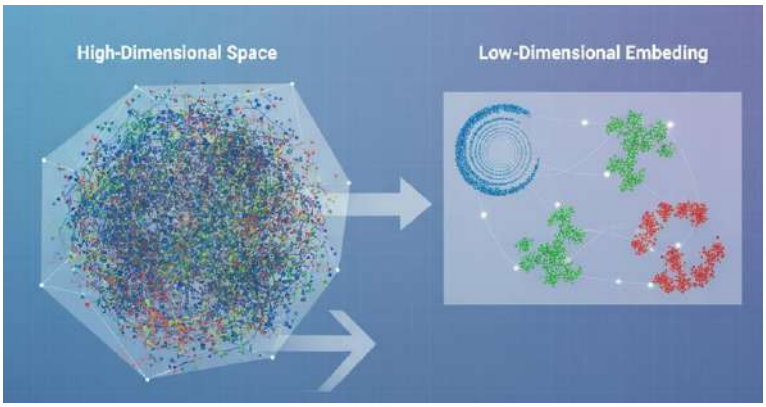
Kiến trúc cơ bản của autoencoder bao gồm ba thành phần chính:

• **Encoder**: ánh xạ dữ liệu đầu vào $x \in \mathbb{R}^d$ sang không gian tiềm ẩn chiều thấp hơn $z \in \mathbb{R}^k$ (với $k < d$). Quá trình này có thể được biểu diễn như: $z = f_\theta(x) = \sigma(W_e x + b_e)$, trong đó W_e và b_e là trọng số và độ lệch của bộ mã hóa, và σ là hàm kích hoạt phi tuyến tính.



• **Latent Space**: là biểu diễn nén của dữ liệu, chứa các đặc trưng quan trọng nhất sau khi loại bỏ nhiễu và thông tin dư thừa.

• **Decoder**: ánh xạ từ không gian tiềm ẩn $z$trở lại không gian dữ liệu gốc để tái tạo $\hat{x}$. Quá trình được định nghĩa là: $\hat{x} = g_\phi(z) = \sigma(W_d z + b_d)$, trong đó $W_d$và $b_d$là tham số bộ giải mã.



Dữ liệu đầu vào ban đầu có số chiều rất lớn (khoảng 100–300 dải phổ). Mô hình Autoencoder được sử dụng để học ánh xạ phi tuyến, chiếu dữ liệu từ không gian gốc xuống một không gian tiềm ẩn có số chiều thấp hơn. Quá trình này vẫn đảm bảo bảo toàn cấu trúc phân bố nội tại của dữ liệu, cụ thể là duy trì được đặc trưng manifold trong không gian biểu diễn mới.

Quy trình thực hiện

Bước 1 – Chuẩn bị dữ liệu: Chuẩn hóa phổ ảnh siêu phổ để đưa dữ liệu về cùng thang đo, đảm bảo phù hợp cho huấn luyện.

Bước 2 – Mã hóa đặc trưng: Đưa dữ liệu qua Encoder (Conv, ReLU, BatchNorm, Pooling) để trích xuất và nén đặc trưng phi tuyến, giảm dư thừa giữa các dải phổ.

Bước 3 – Không gian tiềm ẩn: Thu được vector (z) có số chiều thấp (≈16–32), vẫn giữ đặc trưng phổ và cấu trúc dữ liệu.

Bước 4 – Tái tạo: Decoder khôi phục lại dữ liệu từ (z) nhằm đánh giá mức độ bảo toàn thông tin.

Bước 5 – Hàm mất mát: Sử dụng MSE để đo sai số tái tạo và làm tiêu chí tối ưu.

Bước 6 – Huấn luyện: Dùng Backpropagation và Adam để cập nhật đồng thời Encoder–Decoder, giảm sai số tái tạo.

Bước 7 – Trích xuất đặc trưng: Giữ lại Encoder để tạo vector đặc trưng giảm chiều phục vụ các bài toán tiếp theo.

Mục tiêu của autoencoder là giảm thiểu lỗi tái tạo giữa dữ liệu gốc $x$và dữ liệu tái tạo $\hat{x}$, thông qua hàm mất mát: $L(x, \hat{x}) = \| x - \hat{x} \|^2$. Với cấu trúc này, autoencoder có thể học các đặc trưng quan trọng của dữ liệu đồng thời loại bỏ nhiễu và giảm chiều dữ liệu một cách hiệu quả.

B. Phân cụm mờ C-means hợp tác

Cho bộ dữ liệu $X = x_1, x_2, \dots, x_N$, có P tập con dữ liệu (vị trí dữ liệu) bao gồm $D^{[1]}, D^{[2]}, \dots, D^{[P]}$, trong đó mỗi tập con chứa $N^{[1]}, N^{[2]}, \dots, N^{[P]}$mẫu dữ liệu trong cùng không gian thuộc tính X. Trong mỗi vị trí dữ liệu $D^{[ii]}$, chúng ta chia thành Ccụm. Kết quả phân cụm ở mỗi tập dữ liệu ảnh hưởng đến phân cụm ở các vùng còn lại.

Trong đó, các vị trí dữ liệu không trao đổi trực tiếp dữ liệu chi tiết mà chỉ chia sẻ thông tin cấu trúc, là vector tâm cụm $v^{[ii]}$(với $ii = 1, \dots, P$). Khi phân cụm hợp tác, nhìn ở mức tổng thể của các vị trí dữ liệu sẽ tốt hơn kết quả phân cụm dựa chỉ trên dữ liệu cục bộ tại mỗi vị trí dữ liệu.

Vấn đề phân cụm mờ hợp tác có hàm mục tiêu cần tối ưu hóa là:

$$Q^{[ii]} = \sum_{k=1}^{N^{[ii]}} \sum_{i=1}^{C} u_{ik}^{2[ii]} d_{ik}^2 + \beta \sum_{jj=1}^{P} \sum_{k=1}^{N^{[ii]}} \sum_{i=1}^{C} (u_{ik} - \tilde{u}_{ik}^{[ii|jj]})^2 d_{ik}^2 \qquad (1)$$

Phần đầu của hàm mục tiêu tương tự hàm mục tiêu của thuật toán FCM. Phần thứ hai của hàm mục tiêu cho thấy rằng tối ưu hóa trong quá trình hợp tác làm giảm sự khác biệt giữa các ma trận phân vùng.

Trong hàm mục tiêu trên, $u_{ik}^{[ii]}$là ma trận phân vùng đối tượng $k$vào cụm $i$trong vị trí dữ liệu ii. $\tilde{u}_{ik}^{[ii|jj]}$được gọi là ma trận phân vùng hợp tác của vị trí dữ liệu jjđối với vị trí dữ liệu ii(với $ii, jj = 1, \dots, P$), và được tính bằng phương trình (2):

$$\tilde{u}_{ik}^{[ii|jj]} = \frac{1}{\sum_{j=1}^{c} \left(\frac{\| x_k^{[ii]} - v_i^{[jj]} \|}{\| x_k^{[ii]} - v_j^{[jj]} \|}\right)^2} \qquad (2)$$

Tham số $\beta^{[ii|jj]}$biểu thị mức độ hợp tác giữa các vị trí dữ liệu. Giá trị càng lớn thì mức độ hợp tác càng cao, và giá trị $\beta^{[ii|jj]} = 0$biểu thị không có hợp tác giữa các tập dữ liệu. d_{ik} là khoảng cách từ đối tượng thứ k đến tâm cụm thứ i trong cùng vị trí dữ liệu.

Sử dụng phương pháp Lagrange để tối ưu hóa hàm mục tiêu trên, phương trình tính ma trận phân vùng và tâm cụm như sau:

$$u_{rs}^{[ii]} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{rs}^2}{d_{js}^2}\right)} \left[1 - \sum_{j=1}^{c} \frac{\beta^{[ii|jj]}}{\sum_{\substack{jj=1 \\ jj \neq ii}}^{P} \beta^{[ii|jj]}_{(P-1)}} \tilde{u}_{js}^{[ii|jj]}\right] + \frac{\beta^{[ii|jj]}}{\sum_{\substack{jj=1 \\ jj \neq ii}}^{P} \beta^{[ii|jj]}(P-1)} \tilde{u}_{rs}^{[ii|jj]} \qquad (3)$$

$$v_{rt}^{[ii]} = \frac{\sum_{k=1}^{N^{[ii]}} u_{rk}^{2[ii]} x_{kt} + \beta^{[ii|jj]} \sum_{\substack{l=1 \\ l \neq ii}}^{P} \sum_{k=1}^{N^{[ii]}} (u_{rk}^{[ii]} - \tilde{u}_{rk}^{[ii|jj]})^2 x_{kt}}{\sum_{k=1}^{N^{[ii]}} u_{rk}^{2[ii]} + \beta^{[ii|jj]} \sum_{\substack{l=1 \\ l \neq ii}}^{P} \sum_{k=1}^{N^{[ii]}} (u_{rk}^{[ii]} - \tilde{u}_{rk}^{[ii|jj]})^2} \qquad (4)$$

Ma trận phân vùng của hàm mục tiêu phải thỏa mãn ràng buộc rằng tổng thành viên của một phần tử trong các cụm trong cùng tập dữ liệu bằng 1, như sau:

$$U = \{u_{ik} \in [0,1], \sum_{i=1}^{c} u_{ik}^{[ii]} = 1, \forall k; 0 < \sum_{k=1}^{N^{[ii]}} u_{ik}^{[ii]} < N^{[ii]}, \forall i\} \qquad (5)$$

III. PHÂN CỤM MỜ HỢP TÁC VỚI GIẢM CHIỀU

Phân cụm mờ hợp tác với giảm chiều là một cách tiếp cận cải thiện hiệu quả phân đoạn ảnh viễn thám, đặc biệt với dữ liệu siêu phổ chiều cao và nhiễu. Trước tiên, dữ liệu được giảm chiều bằng các kỹ thuật như Autoencoder để loại bỏ thông tin dư thừa và trích xuất các đặc trưng quan trọng. Sau đó, thuật toán phân cụm mờ hợp tác tận dụng cơ chế mờ và hợp tác trên nhiều không gian đặc trưng, cho phép tối ưu hóa hàm mục tiêu toàn cục. Cách tiếp cận này cải thiện cả độ chính xác và độ ổn định của phân cụm.

Dữ liệu ảnh siêu phổ được thu thập $X = \{x_1, x_2, \dots, x_n\}, x_i \in \mathbb{R}^d$trong đó $d$là số kênh phổ. Số cụm: C. Ma trận thành viên cho vị trí ii: $U^{(ii)} = [u_{ci}^{(ii)}] \in [0,1]^{C\times n}$, với $\sum_{c=1}^{C} u_{ci}^{(ii)} = 1$. Tâm cụm của vị trí ii: $C^{(ii)} = [c_1^{(ii)}, \dots, c_C^{(ii)}]$. Tham số mờ: $m > 1$. Trọng số vị trí: $\alpha_{ii} \geq 0$, với $\sum_{ii} \alpha_{ii} = 1$. Dữ liệu được chuẩn hóa trước khi sử dụng thuật toán để giảm chiều dữ liệu từ $d$chiều xuống $k$chiều ($k < d$).

Các bước thực hiện thuật toán AutoEncoder để giảm chiều dữ liệu viễn thám siêu phổ như sau:

Bước 1: Chuẩn bị dữ liệu

• Bộ dữ liệu: $X = \{x_1, x_2, \dots, x_n\}, x_i \in \mathbb{R}^d$.

• Chuẩn hóa dữ liệu.

Bước 2: Xây dựng mạng Autoencoder

• Bộ mã hóa: $z = f_\theta(x) = \sigma(W_e x + b_e)$, ánh xạ dữ liệu sang không gian tiềm ẩn k-chiều.

• Bộ giải mã: $\hat{x} = g_\phi(z) = \sigma(W_d z + b_d)$, tái tạo dữ liệu từ biểu diễn tiềm ẩn.

Bước 3: Xác định hàm mất mát

Sử dụng Sai số Bình phương Trung bình làm hàm lỗi tái tạo.

$$L(x, \hat{x}) = \frac{1}{n}\sum_{i=1}^{n} \| x_i - \hat{x}_i \|^2$$

Bước 4: Huấn luyện mạng Autoencoder

• Với $t = 1$ đến T:

- Lan truyền thuận: tính $z_i = f_\theta(x_i)$và $\hat{x}_i = g_\phi(z_i)$.

- Tính hàm mất mát $L(x, \hat{x})$.

- Lan truyền ngược và cập nhật các tham số (W_e, b_e, W_d, b_d) bằng Hạ gradient với tốc độ học η.

• Kết thúc vòng lặp.

Bước 5: Trích xuất đặc trưng Sau khi huấn luyện, thu được biểu diễn chiều giảm $z_i = f_\theta(x_i), z_i \in \mathbb{R}^k$.

Bước 6: Sử dụng thuật toán CFCM Phân đoạn dữ liệu viễn thám siêu phổ trong không gian chiều mới.

Bước 7: Tái tạo dữ liệu Tái tạo dữ liệu từ kvề không gian d-chiều gốc.

Đầu ra: Ma trận thành viên cuối cùng Uvà tâm cụm C. Gán nhãn cụm cho mỗi $x_i$như

$$\text{label}(x_i) = \arg\max_j u_{ij}$$

Các bước thực hiện chi tiết của thuật toán đề xuất được thể hiện trong Thuật toán 1.

Thuật toán 1 Thuật toán CFCM-AE

Đầu vào: Bộ dữ liệu $X = \{x_1, x_2, \dots, x_n\}$, ε, và khởi tạo các tham số m, ϵ, số lần lặp tối đa $T_{\max}$, số cụm trong bộ dữ liệu, $t = 0$.

Đầu ra: Kết quả phân cụm.

Bắt đầu Giai đoạn 1: Giảm chiều

1. Chuẩn hóa dữ liệu.
2. Xây dựng mạng Autoencoder.
3. Xác định hàm mất mát (MSE).
4. Huấn luyện mạng Autoencoder.
5. Trích xuất đặc trưng từ d-chiều sang k-chiều.

Giai đoạn 2: Phân cụm hợp tác

1. Chia ngẫu nhiên bộ dữ liệu gốc thành các vị trí dữ liệu.

2. Tính giá trị thành viên và tâm cụm trên các vị trí dữ liệu bằng X, Y và V^*.

- Tính ma trận phân vùng cục bộ $u_{rs}^{[ii]}$bằng Eq. (3).

- Tính nguyên mẫu cụm cục bộ $v_{rt}^{[ii]}$bằng Eq. (4).

- Tính nguyên mẫu cụm toàn cục $V = (v_1, v_2, \dots, v_c)$bằng cách lấy trung bình từ $v_{rt}^{[ii]}$.

3. LẶP

3.1. $t++$.

3.2. Giao tiếp nguyên mẫu cụm từ mỗi vị trí đến tất cả các vị trí khác

Lặp

- Tính ma trận phân vùng cục bộ $u_{rs}^{[ii]}$bằng Eq. (3).

- Tính nguyên mẫu cụm cục bộ $v_{rt}^{[ii]}$bằng Eq. (4).

- Tính giá trị hàm mục tiêu $Q^{[ii]}$bằng Eq. 1.

- Tính nguyên mẫu cụm toàn cục $V = (v_1, v_2, \dots, v_c)$bằng cách lấy trung bình từ $v_{rt}^{[ii]}$. Cho đến khi hàm mục tiêu cục bộ được giảm thiểu. Cho đến N max $((V^{(t)} - V^{(t-1)}) < \varepsilon$ HOẶC $(t = T_{\max}))$

Giai đoạn 3: Tái tạo dữ liệu

1. Chuyển đổi dữ liệu từ k-chiều sang d-chiều.

2. Xuất kết quả phân cụm.

Chỉ số đánh giá chất lượng phân cụm

Để đánh giá chất lượng phân loại của thuật toán trên các bộ dữ liệu ảnh vệ tinh, bài báo sử dụng một số chỉ số để đo chất lượng cụm, bao gồm Hệ số Phân vùng (PC), Entropy Phân vùng (PE), và chỉ số Xie-Beni (XB) [13]. Chỉ số PC có giá trị lớn, trong khi chỉ số PE và XB nhỏ cho thấy chất lượng phân cụm tốt.

$$PC = \frac{1}{n}\sum_{i=1}^{C}\sum_{k=1}^{n} u_{ik}^2 \qquad (6)$$

$$PE = -\frac{1}{n}\sum_{i=1}^{C}\sum_{k=1}^{n} u_{ik} \ln u_{ik} \qquad (7)$$

$$XB = \frac{\frac{1}{n}\sum_{i=1}^{C}\sum_{k=1}^{n} u_{ik}^2 \parallel v_i - x_k \parallel^2}{\min\limits_{i \neq j} \parallel v_i - v_j \parallel^2} \qquad (8)$$

Ngoài ra, chúng tôi cũng lấy mẫu các cụm/lớp từ bộ dữ liệu thực nghiệm gốc để đánh giá độ chính xác của kết quả phân cụm. Phương trình tính độ chính xác của kết quả phân cụm như sau.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \qquad (9)$$

Ở đây, TP (True Positive) biểu thị số mẫu dương được phân loại đúng, FN (False Negative) là số mẫu dương được phân loại sai thành âm, FP (False Positive) là số mẫu âm được phân loại sai thành dương, và TN (True Negative) biểu thị số mẫu âm được phân loại đúng. Một thuật toán tốt được đặc trưng bởi TPR (True Positive Rate) cao hơn và FTR (False True Rate) thấp hơn.

IV. THỰC NGHIỆM

Trong thiết lập thực nghiệm, chúng tôi tiến hành các thử nghiệm trên máy tính được trang bị CPU Intel Core i7 chạy ở 2.9 GHz, 16 GB RAM, hệ điều hành Windows 10, và card đồ họa NVIDIA với 8 GB bộ nhớ. Các tham số thực nghiệm được cấu hình như sau: tham số mờ $m = 2$, số lần lặp tối đa $T_{\max} = 1000$, và tiêu chí dừng $\epsilon = 10^{-6}$. Các thuật toán được sử dụng trong thực nghiệm bao gồm FCM [1], CFCM [3], và thuật toán đề xuất CFCM-AE.

A. Thực nghiệm 1

Trong Thực nghiệm 1, dữ liệu ảnh siêu phổ (HSI) được sử dụng là bộ dữ liệu Indian Pines, được thu thập từ vùng tây bắc Indiana. Ảnh có kích thước không gian 145 × 145 pixel với 200 dải phổ trải dài trong khoảng bước sóng từ 0.4 µm đến 2.4 µm. Tổng cộng có 10.201 pixel được gán nhãn. Bộ dữ liệu bao gồm 14 lớp phủ đất, cụ thể: Alfalfa, Corn-notill, Corn-mintill, Corn, Grass-pasture, Grass-trees, Hay-windrowed, Soybean-notill, Soybean-mintill, Soybean-clean, Wheat, Woods, Building-Grass-Trees-Drives, và Stone-Steel-Towers. Thông tin chi tiết về bộ dữ liệu và số pixel được gán nhãn cho mỗi lớp được cung cấp trong Bảng I.

TABLE I
GROUNDTRUTH CLASSES FOR THE INDIAN PINES SCENE AND SAMPLES NUMBERS

No.	Class	Samples	Color
1	Alfalfa	46	
2	Corn-notill	1428	
3	Corn-mintill	830	
4	Corn	237	
5	Grass-pasture	483	
6	Grass-trees	730	
7	Hay-windrowed	478	
8	Soybean-notill	972	
9	Soybean-mintill	2455	
10	Soybean-clean	593	
11	Wheat	205	
12	Woods	1265	
13	Buildings-Grass-Trees-Drives	386	
14	Stone-Steel-Towers	93	

TABLE II
CLUSTERING INDICATOR RESULTS OF INDIAN PINES DATASET

No.	Alg.	PC	PE	XB	Accuracy	Runtime
1	FCM	0.698	0.625	0.471	NoN	2m31s
2	CFCM	0.694	0.619	0.488	NoN	2m09s
3	FCM-AE	0.804	0.506	0.325	86.29	1m52s
4	CFCM-AE	**0.821**	**0.497**	**0.323**	**98.45%**	**1m21s**

Kết quả chạy mô hình Autoencoder trên bộ dữ liệu Indian Pines giảm chiều gốc từ 200 chiều xuống 36 chiều.

Hình 1 là kết quả phân cụm dữ liệu ảnh viễn thám bằng thuật toán CFCM-AE. Bảng II là kết quả phân cụm của năm thuật toán FCM, CFCM, FCM-AE, và CFCM-AE trên bộ dữ liệu Indian Pines. Giá trị PC và Accuracy càng cao thì chất lượng phân cụm càng tốt, trong khi giá trị PE, XB và thời gian chạy càng nhỏ thì càng tốt.

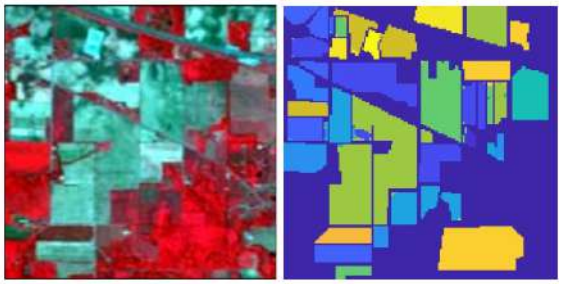


Fig. 1. The Indian Pines dataset and clustering result

Dựa trên bảng, FCM và CFCM (không có Autoencoder) có độ chính xác thấp hơn đáng kể. Cụ thể, các chỉ số PC và PE cho thấy mức độ mờ chưa được tối ưu, trong khi chỉ số XB vẫn còn tương đối cao (0.471 và 0.488), phản ánh chất lượng phân cụm kém. Ngoài ra, cả hai phương pháp đều không báo cáo rõ ràng giá trị Accuracy.

FCM-AE cho thấy sự cải thiện rõ rệt nhờ tích hợp Autoencoder để giảm chiều và trích xuất đặc trưng. PC tăng lên 0.804, PE giảm xuống 0.506, XB giảm mạnh xuống 0.325. Độ chính xác đạt 86.29%, và thời gian chạy cũng được rút ngắn xuống 1m52s. Điều này chứng minh rằng Autoencoder hỗ trợ mô hình phân cụm hiệu quả hơn.

CFCM-AE (thuật toán đề xuất) đạt kết quả tốt nhất trên tất cả các chỉ số. PC cao nhất (0.821), PE thấp nhất (0.497), XB thấp nhất (0.323), và Accuracy vượt trội 98.45%, và có thời gian chạy ngắn nhất (1m21s). Điều này chứng minh rằng việc kết hợp lọc hợp tác và giảm chiều bằng autoencoder đồng thời cải thiện độ chính xác, chất lượng phân cụm và hiệu quả tính toán.

Từ các kết quả trên, có thể thấy rằng CFCM-AE là thuật toán tối ưu nhất trong số bốn phương pháp, xác nhận hiệu quả của phương pháp đề xuất.

B. Thực nghiệm 2

Trong Thực nghiệm 2, dữ liệu ảnh siêu phổ (HSI) được sử dụng được thu thập từ Đắk Lắk, Việt Nam với 110 dải phổ. Tổng cộng có 105.908 pixel được gán nhãn. Bộ dữ liệu bao gồm 6 lớp phủ đất, cụ thể: Lớp 1: Sông, hồ, ao. Lớp 2: Đất trống, đường. Lớp 3: Đồng ruộng, cỏ. Lớp 4: Rừng thưa, cây thấp. Lớp 5: Cây lâu năm. Lớp 6: Rừng rậm, rừng nhiệt đới.

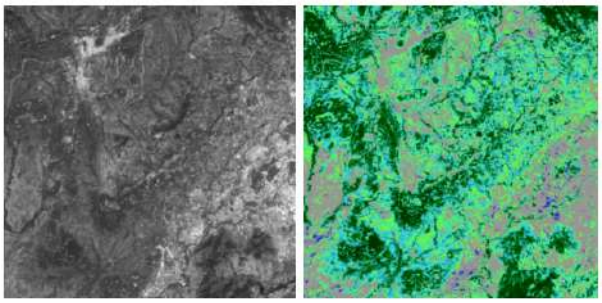


Fig. 2. Hyperspectral image data of Daklak region

Kết quả chạy mô hình Autoencoder trên bộ dữ liệu Đắk Lắk giảm chiều gốc từ 110 chiều xuống 21 chiều.

Hình 2 là kết quả phân cụm dữ liệu ảnh viễn thám bằng thuật toán CFCM-AE. Như được thể hiện trong Bảng III, các phương pháp cơ bản FCM và CFCM cho kết quả phân cụm tương đối yếu, với giá trị PC thấp (0.569 và 0.578), giá trị PE cao (0.761 cho cả hai), và chỉ số XB lớn (0.547 và 0.518). Hơn nữa, độ chính xác của chúng không được báo cáo, trong khi thời gian

chạy vẫn tương đối dài ở 3m36s và 2m54s. Bằng cách tích hợp Autoencoder, FCM-AE chứng minh sự cải thiện đáng kể, đạt PC cao hơn nhiều (0.836), PE thấp hơn (0.648), và XB giảm (0.433), đồng thời đạt độ chính xác 86.54% với thời gian chạy ngắn hơn 2m43s, xác nhận lợi ích của trích xuất đặc trưng và giảm chiều.

CFCM-AE đề xuất nâng cao hơn nữa hiệu suất, mang lại kết quả tốt nhất với PC cao nhất (0.857), PE thấp nhất (0.641), XB nhỏ nhất (0.320), và độ chính xác 88.13%, đồng thời duy trì thời gian chạy cạnh tranh 2m51s. Những kết quả này nhấn mạnh hiệu quả của việc kết hợp phân cụm mờ hợp tác với Autoencoder, dẫn đến chất lượng phân cụm và độ chính xác phân loại vượt trội so với các cách tiếp cận khác.

TABLE III
CLUSTERING INDICATOR RESULTS OF DAKLAK REGION

No.	Alg.	PC	PE	XB	Accuracy	Runtime
1	FCM	0.569	0.761	0.547	NoN	3m36s
2	CFCM	0.578	0.761	0.518	NoN	2m54s
3	FCM-AE	0.836	0.648	0.433	86.54	**2m43s**
4	CFCM-AE	**0.857**	**0.641**	**0.320**	**88.13%**	2m51s

Từ các thực nghiệm trên, có thể thấy rằng phương pháp đề xuất không chỉ cải thiện chất lượng cụm mà còn có thời gian thực thi thuật toán nhanh hơn (thuật toán hội tụ nhanh hơn). Điều này chứng minh tính đúng đắn của phương pháp đề xuất.

V. KẾT LUẬN

Bài báo đề xuất sự kết hợp giữa phân cụm mờ hợp tác và mạng Autoencoder cho phân đoạn ảnh viễn thám siêu phổ. Sự kết hợp giữa phân cụm mờ hợp tác và mạng Autoencoder cung cấp một khung mạnh mẽ cho phân đoạn ảnh viễn thám siêu phổ. Bằng cách tận dụng khả năng trích xuất đặc trưng và giảm chiều của Autoencoder cùng với độ bền của phân cụm mờ hợp tác, cách tiếp cận đề xuất giải quyết hiệu quả các thách thức của dữ liệu chiều cao và phân bố lớp phức tạp. Các kết quả thực nghiệm xác nhận rằng sự tích hợp này không chỉ cải thiện chất lượng phân cụm và độ chính xác phân đoạn mà còn nâng cao hiệu quả tính toán. Do đó, phương pháp đề xuất đại diện cho một giải pháp đầy hứa hẹn để thúc đẩy phân tích ảnh siêu phổ trong các ứng dụng viễn thám thực tế.

Trong công việc tương lai, chúng tôi dự định tiến hành các thực nghiệm phân cụm trên các bộ dữ liệu lớn, chiều cao trong môi trường tính toán phân tán.

REFERENCE

[1] James C. Bezdek and Robert Ehrlich and William Full. FCM: The fuzzy c-means clustering algorithm, Computers & Geosciences, Vol.10(2), 1984, https://doi.org/10.1016/0098-3004(84)90020-7.

[2] Witold Pedrycz, Collaborative fuzzy clustering, Pattern Recognition Letters, Vol. 23(14), pp.1675-1686, 2002, https://doi.org/10.1016/S0167-8655(02)00130-7.

[3] Witold Pedrycz and Partab Rai, Collaborative clustering with the use of Fuzzy C-Means and its quantification, Fuzzy Sets and Systems, Vol. 159(18), pp.2399-2427, 2008, https://doi.org/10.1016/j.fss.2007.12.030.

[4] Dang TH, Do VD, Mai DS, Ngo LT, Trinh LH. Features reduction collaborative fuzzy clustering for hyperspectral remote sensing images analysis. Journal of Intelligent & Fuzzy Systems. 2023; 45(5):7739-7752. https://doi.org/10.3233/JIFS-230511

[5] Berahmand, K., Daneshfar, F., Salehi, E.S. et al. Autoencoders and their applications in machine learning: a survey. Artificial Intelligence Review, Vol.57, No.28 (2024).

https://doi.org/10.1007/s10462-023-10662-6
[6] H. Kchaou, W. Abbes, Z. Kechaou and A. M. Alimi, ”Collaborative Fuzzy Clustering Approach for Scientific Cloud Workflows,” 2023 IEEE Symposium on Computers and Communications (ISCC), Gammarth, Tunisia, 2023, pp. 424-429, https://doi.org/10.1109/ISCC58397.2023.10218274.
[7] Mienye, I.D., Swart, T.G. Deep Autoencoder Neural Networks: A Comprehensive Review and New Perspectives. Archives of Computational Methods in Engineering (2025). https://doi.org/10.1007/s11831-025-10260-5
[8] K. Wu, J. Liu, P. Liu and S. Yang, ”Time Series Prediction Using Sparse Autoencoder and High-Order Fuzzy Cognitive Maps,” in IEEE Transactions on Fuzzy Systems, vol. 28, no. 12, pp. 3110-3121, 2020, https://doi.org/10.1109/TFUZZ.2019.2956904
[9] H. Zhang, H. Zhai, L. Zhang and P. Li, ”Spectral–Spatial Sparse Subspace Clustering for Hyperspectral Remote Sensing Images,” in IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 6, pp. 3672-3684, 2016, https://doi.org/10.1109/TGRS.2016.2524557
[10] Z. Shao, L. Zhang, X. Zhou and L. Ding, ”A Novel Hierarchical Semisupervised SVM for Classification of Hyperspectral Images,” in IEEE Geoscience and Remote Sensing Letters, vol. 11, no. 9,pp. 1609-1613, 2014, https://doi.org/10.1109/LGRS.2014.2302034
[11] A. Zehtabian and H. Ghassemian, ”An Adaptive Pixon Extraction Technique for Multispectral/ Hyperspectral Image Classification,” in IEEE Geoscience and Remote Sensing Letters, vol. 12, no. 4, pp. 831-835, 2015, https://doi.org/10.1109/LGRS.2014.2363586
[12] Y. Chen, W. Ding, J. Huang, W. Zhang and T. Zhou, ”Multigranularity Fuzzy Autoencoder for Discriminative Feature Selection in High-Dimensional Data,” in IEEE Transactions on Neural Networks and Learning Systems, vol. 36, no. 9, pp. 17433-17447, 2025, https://doi.org/10.1109/TNNLS.2025.3569893
[13] X. Xie and G. Beni, “A validity measure for fuzzy clustering,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 13, no. 8, pp. 841–847, 1991

A Combination of Collaborative Fuzzy Clustering and Autoencoder Network for Hyperspectral Remote Sensing Image Segmentation

Abstract

Hyperspectral remote sensing imagery (HRSI) contains rich spectral-spatial information, offering great potential for detailed land cover mapping and object recognition. However, their high dimensionality and spectral redundancy pose significant challenges for efficient and accurate segmentation. To overcome these limitations, this paper presents a collaborative fuzzy clustering method combined with an Autoencoder network (CFCM-AE) by integrating the collaborative fuzzy C-means clustering method with an Autoencoder network for hyperspectral image segmentation. In this framework, the Autoencoder performs dimensionality reduction to retain discriminative features while minimizing redundancy and computational costs. The collaborative fuzzy clustering component then leverages both spectral similarity and contextual dependencies between pixels, leading to more stable and robust segmentation. This integration enhances adaptability, improves accuracy, and reduces sensitivity to noise and high-dimensional data. Empirical evaluations on benchmark hyperspectral datasets show that the proposed method significantly outperforms traditional fuzzy clustering and standalone dimensionality reduction techniques, achieving superior performance in terms of accuracy, stability, and efficiency. The results confirm that the combination of collaborative fuzzy clustering and Autoencoders provides an effective and scalable solution for hyperspectral remote sensing image analysis

Keywords: Collaborative fuzzy clustering, dimensionality reduction, AutoEncoder, hyperspectral remote sensing.

A Method for Enhancing Vessel AIS Time Series Data Quality using Denoising and Intelligent Interpolation

Tran Thi Thanh Huyen[1*], Nguyen Quang Thi[1], Vu Thi Ly[1], Nguyen Quang Uy[1]

[1] *Le Quy Don Technical University, Hanoi, Vietnam*

[*] Email: huyenttt@lqdtu.edu.vn, Contact number: 0986534255

Abtract

Automatic Identification System (AIS) time-series data commonly suffer from two major quality issues: corrupted observations caused by noise and outliers, and incomplete trajectories caused by missing segments. To address these challenges, this study proposes a two-stage AIS restoration framework that integrates kinematic-based denoising and Transformer-based interpolation. In the first stage, physically implausible, redundant, and outlier trajectory points are identified and removed, reducing up to 39% of noisy or unnecessary observations while preserving trajectory consistency. In the second stage, missing trajectory information is reconstructed from incomplete AIS sequences using a Transformer-based interpolation model. Experimental results under both random and block missing settings demonstrate that the proposed framework consistently achieves higher reconstruction accuracy than traditional methods, including linear interpolation and Kalman filtering, with particularly strong performance for long-duration missing intervals.

Keywords: *Automatic identification system, trajectory reconstruction, Self-Attention.*

1. Introduction

AIS data have become indispensable in maritime surveillance [1], intelligent navigation [2], vessel classification [3-5], and anomaly detection [6, 7]. Nevertheless, real-world AIS streams are often imperfect. Due to communication interference, atmospheric conditions, hardware malfunction, or signal obstruction, AIS time series may contain missing segments, abnormal outliers, irregular sampling intervals, and noisy kinematic measurements. Such quality degradation directly affects maritime decision-support systems and machine learning applications built upon AIS data. Therefore, enhancing AIS time series data quality is a crucial prerequisite for reliable maritime intelligence.

Vessel trajectory reconstruction has progressed through denoising and interpolation techniques [8, 9]. Existing AIS cleaning methods mainly rely on anomaly detection and outlier correction, but remain limited in handling dynamic noise caused by positioning errors and communication loss [10]. Conventional interpolation approaches, including Lagrange and cubic spline methods, are widely adopted, yet their effectiveness often degrades under complex maneuvers and irregular missing intervals [11]. Although dynamic-based interpolation methods and deep learning architectures such as U-Net and LSTM have shown promise, they still suffer from limited adaptability, strong data dependence, and considerable computational demands [12, 13]. In addition, recent integrated denoising–interpolation frameworks still rely heavily on historical trajectory databases [14]. These limitations highlight the need for a two-stage restoration framework that can first identify abnormal AIS observations and then recover missing trajectory segments using long-range sequence modeling, which motivates the Transformer-based approach proposed in this study.

In this study, we address AIS data quality enhancement as an integrated problem combining denoising and intelligent interpolation. The main objectives of this work are:

1 To design a kinematics-aware denoising mechanism for AIS trajectories.

2 To develop an intelligent interpolation model capable of recovering missing data under complex motion dynamics.

3 To establish maritime-specific evaluation metrics for assessing reconstruction quality.

By treating AIS reconstruction as a data quality enhancement task rather than solely a missing value problem, this work provides a systematic and practical solution for maritime data preprocessing.

2. Method

2.1 Overall Framework

The proposed framework is composed of two sequential stages: (1) a denoising stage for noise suppression and data correction, and (2) an intelligent interpolation stage for missing value reconstruction, as illustrated in Fig. 1.

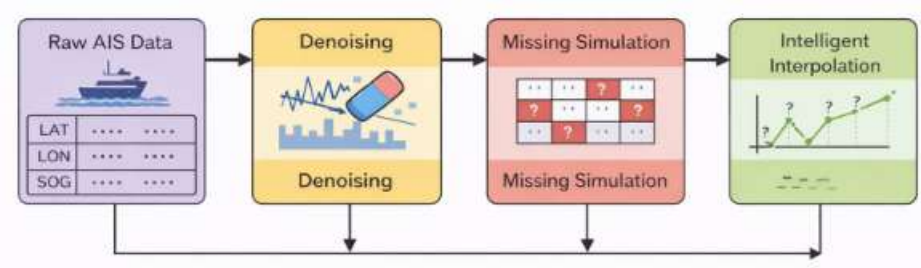


Figure 1. Overview of the proposed AIS data quality enhancement process.

2.2 Denoising Module

AIS data are susceptible to various quality issues, such as noise and missing values arising from signal transmission constraints and human-related factors, which can adversely affect the accuracy of practical maritime analyses [15]. Consequently, appropriate data preprocessing is necessary to alleviate these issues and provide reliable inputs for vessel classification tasks. In this study, the following preprocessing procedures are implemented.

First, following Deng et al. [16], records that violate physical plausibility or spatial consistency were removed. Specifically, AIS messages located outside the predefined Region of Interest (ROI) were excluded. In addition, records with speed over ground (SOG) exceeding 30 knots or course over ground (COG) values greater than 360° were discarded, as such values likely reflect measurement errors or abnormal outliers.

Second, duplicated entries and records associated with multiple vessel types were eliminated. A duplicated AIS message refers to an entry containing identical values across all relevant attributes, typically caused by multiple receiving stations capturing the same transmission, AIS retransmission mechanisms, or inconsistencies during data integration.

Third, vessel trajectories were reconstructed following Nguyen and Fablet [17] and Liu et al. [18]. AIS records were retrieved from the historical database, grouped by MMSI, and chronologically ordered. Consecutive messages were segmented into individual voyages using a 12-hour temporal threshold.

2.3 Intelligent Interpolation Module

After denoising, AIS trajectories remain irregularly sampled and may still contain temporal gaps. To enhance data quality and enable reliable learning, an intelligent

interpolation module is designed. This module consists of: Resampling and Missing value reconstruction.

Resampling. AIS messages are originally transmitted at irregular time intervals. To standardize the time series representation, each vessel trajectory is resampled to a fixed 5-minute interval ($\Delta t = 300$ seconds).

Let a cleaned AIS trajectory be:

$$\mathcal{T} = \{(lat_i, lon_i, sog_i, cog_i, t_i)\}_{i=1}^{N}$$

Where

- $lat_i \in [-90^o, 90^o]$ denotes the latitude at time t_i,
- $lon_i \in [-180^o, 180^o]$ denotes the longitude at time t_i,
- $sog_i \in \mathbb{R}^+$ represents the speed over ground (in knots),
- $cog_i \in [0^o, 360^o)$ represents the course over ground (in degrees),
- $t_i \in \mathbb{R}$ denotes the timestamp of the AIS message,
- N is the total number of AIS messages for the vessel after denoising.

The sequence is assumed to be temporally ordered such that:

$$t_1 < t_2 < \cdots < t_N$$

This ordered structure forms the basis for subsequent temporal resampling and interpolation.

We construct a uniform temporal grid:

$$t_k = t_0 + k\Delta t$$

For each target timestamp tkt_ktk, we identify:

$$t_b = max\{t_i | t_i \leq t_k\}$$

$$t_a = min\{t_i | t_i > t_k\}$$

Interpolation is performed only if:

$$|t_a - t_b| < 7200 \text{ seconds}$$

The geodesic distance between two AIS positions is computed using the inverse geodesic solution:

$$(\alpha, d) = GeodInv(lon_b, lat_b, lon_a, lat_a)$$

Where: α is forward azimuth

d is geodesic distance

The interpolation ratio:

$$r = \frac{t_k - t_b}{t_a - t_b}$$

Interpolated distance: $d_r = r.d$

The interpolated coordinate is obtained using the forward geodesic solution:

$$(lon_k, lat_k) = GeodFwd(lon_b, lat_b, \alpha, d_k)$$

Speed and course are interpolated linearly:

$$sog_k = sog_b + r(sog_a - sog_b), cog_k = cog_b + r(cog_a - cog_b)$$

Time Series Representation. After resampling, each vessel trajectory becomes a multivariate time series:

$$X = \{x_t\}_{t=1}^{T}, x_t \in \mathbb{R}^5$$

Where: $x_t = [lat_t, lon_t, sog_t, \sin(cog_t), cos(cog_t)]$

cog is encoded as sine and cosine to avoid angular discontinuity.

A binary mask M is defined as:

$$M_{t,d} = \begin{cases} 1 & if\ observed \\ 0 & if\ missing \end{cases}$$

The masked input is: $X_{masked} = M \odot \mathrm{X}$

The objective is to estimate: $\hat{X} = f_0(X_{masked}, M, \Delta t)$

Self-Attention-Based Intelligent Interpolation (SABII). As illustrated in Fig. 2, the proposed model adopts the dual-stage self-attention architecture of SAITS [19], with modifications tailored to AIS trajectory reconstruction, including circular heading modeling and kinematic-aware loss design.

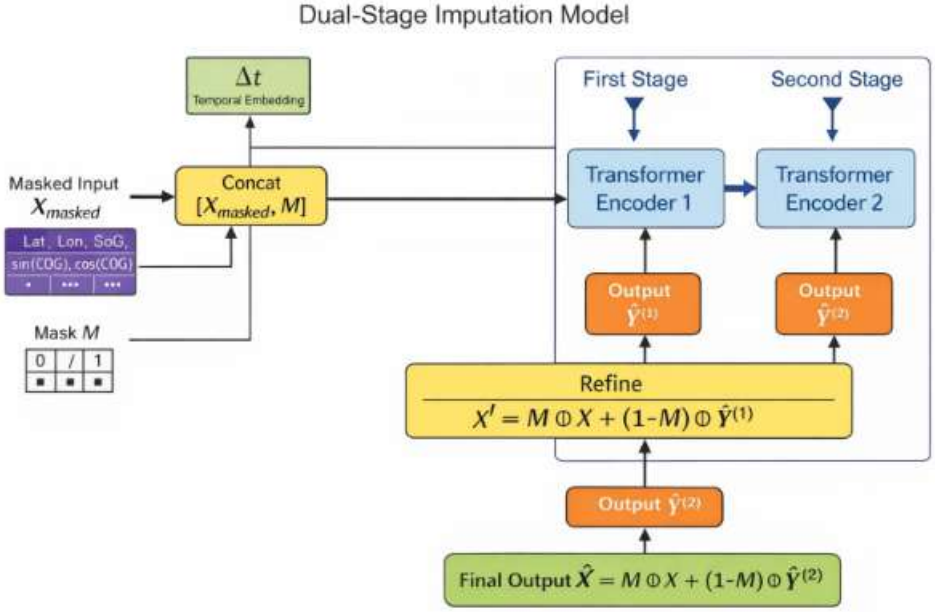


Figure 2. Architecture of the proposed dual-stage self-attention-based intelligent interpolation model for AIS time series reconstruction.

First Stage

Observed values and mask are concatenated: $Z = [X_{masked}; M]$

Temporal embedding is added: $H^{(1)} = Transformer_1\big(Z + E(\Delta t)\big)$

First reconstruction: $\tilde{X}^{(1)} = W_1 H^{(1)}$

Refined input: $X' = M \odot \mathrm{X} + (1 - \mathrm{M}) \odot \tilde{X}^{(1)}$

Second Stage

$$H^{(2)} = Transformer_2\big([X'; M] + E(\Delta t)\big)$$

Second reconstruction: $\tilde{X}^{(2)} = W_1 H^{(2)}$

Final reconstruction: $\hat{X} = M \odot \mathrm{X} + (1 - \mathrm{M}) \odot \tilde{X}^{(2)}$

Loss Functions

Let the missing set be: $\Omega = \{(t,d)|M_{t,d} = 0\}$

Masked Mean Absolute Error is applied only to missing entries.

First-stage loss:

$$L_1(\theta) = \frac{1}{|\Omega|} \sum_{(t,d)\in\Omega} \left|\tilde{X}_{t,d}^{(1)} - X_{t,d}\right|$$

Second-stage loss:

$$L_2(\theta) = \frac{1}{|\Omega|} \sum_{(t,d)\in\Omega} \left|\tilde{X}_{t,d}^{(2)} - X_{t,d}\right|$$

Final reconstruction loss:

$$L_{final}(\theta) = \frac{1}{|\Omega|} \sum_{(t,d)\in\Omega} \left|\hat{X}_{t,d} - X_{t,d}\right|$$

Overall objective:

$$\mathcal{L} = \frac{1}{3}\Big(L_1(\theta) + L_2(\theta) + L_{final}(\theta)\Big)$$

2.6 Evaluation Metrics

To ensure maritime consistency, we employ domain-specific metrics.

Geodesic Position Error. Computed using the Haversine formula:

$$\text{Position Error} = d_{geo}\big(lat_{true}, lon_{true}, lat_{pred}, lon_{pred}\big)$$

Reported in meters.

Speed Error. Absolute difference in SOG (knots).

Circular Heading Error. Angular error computed as:

$$min(|\theta_1 - \theta_2|, 360 - |\theta_1 - \theta_2|)$$

Reported in degrees.

All metrics are evaluated only on originally missing entries.

3. Results

3.1 Effectiveness of Denoising

Table 1 summarizes the impact of the denoising procedure on the Danish Maritime Authority (DMA) dataset. After applying geographical, kinematic, and consistency-based filtering rules, the total number of AIS records decreased from 648,370,476 to 394,898,255, representing an approximate 39% reduction.

This reduction reflects the elimination of redundant, duplicated, and physically implausible records rather than a loss of meaningful information. Further analysis indicates that a substantial proportion of the removed entries corresponded to duplicate messages.

The relatively high removal rate highlights the presence of considerable noise and structural redundancy in the raw AIS data. If left untreated, such defects could disrupt

trajectory continuity and degrade subsequent interpolation performance. Visual inspection further confirms that the cleaned trajectories preserve realistic vessel movement patterns while effectively removing abrupt spatial jumps and inconsistent timestamps.

Table 1. Number of AIS Messages Before and After Denoising in the DMA Datasets

Stage	***No. records***
Before Denoising	*648,370,476*
After Denoising	*394,898,255*

3.2 Performance Comparison of Interpolation Methods

Tables 2 and 3 summarize the interpolation performance under random and block missing conditions. Overall, SABII substantially outperforms Linear interpolation and Kalman filtering in position reconstruction under both scenarios. Under random missing, SABII reduces the Position MAE from 5851.1 m and 8468.4 m to 135.7 m. Under block missing, where reconstruction becomes more challenging, SABII still maintains a low Position MAE of 168.9 m, while the two conventional methods exceed 15,900 m. Similar improvements are also observed for speed estimation.

However, the same trend is not observed for COG reconstruction. In both missing scenarios, SABII produces higher COG errors than Linear interpolation and Kalman filtering, indicating that directional recovery remains more difficult than position or speed interpolation. Overall, the results suggest that SABII is highly effective for recovering spatial trajectory information, particularly under severe missing conditions, but still faces limitations in heading restoration.

Table 2. Interpolation performance under random missing patterns

Method	***Position MAE (m)***	***Position RMSE (m)***	***Speed MAE (knots)***	***Speed RMSE (knots)***	***COG MAE (degree)***	***COG RMSE (degree)***
Linear	*5851.1270*	*8956.8516*	*0.0456*	*0.1328*	*1.3076*	*3.4735*
Kalman filter	*8468.4082*	*10638.3643*	*0.0456*	*0.1328*	*1.3076*	*3.4735*
SABII	*135.6889*	*181.085*	*0.0322*	*0.0424*	*3.6410*	*18.1150*

Table 3. Interpolation performance under contiguous block missing patterns

Method	***Position MAE (m)***	***Position RMSE (m)***	***Speed MAE (knots)***	***Speed RMSE (knots)***	***COG MAE (degree)***	***COG RMSE (degree)***
Linear	*15980.9541*	*22602.9766*	*0.2675*	*0.4137*	*8.2078*	*10.8782*
Kalman filter	*15991.5439*	*22621.2891*	*0.2675*	*0.4137*	*8.2078*	*10.8782*
SABII	*168.8591*	*247.2346*	*0.1256*	*0.1766*	*14.0957*	*35.2755*

4. Discussion

The results reveal a clear difference between the reconstruction of spatial variables (position and speed) and directional variables (COG). SABII consistently achieves much lower errors in position and speed, indicating that it is effective in preserving the overall geometric and kinematic continuity of vessel trajectories. This advantage is especially evident

under block missing patterns, where conventional interpolation methods suffer from substantial error accumulation.

In contrast, SABII is less effective in reconstructing COG, particularly under long missing intervals. This is likely because heading is inherently more sensitive to local directional changes and maneuvering behavior than position or speed. Even small deviations in the recovered trajectory may lead to relatively large angular errors. By comparison, simpler methods such as Linear interpolation and Kalman filtering may yield lower heading errors in some cases because they impose smoother directional transitions, although they remain substantially less accurate in spatial reconstruction.

These findings indicate that SABII provides a clear advantage in recovering spatial trajectory information, particularly vessel position, under the evaluated missing-data conditions. At the same time, they also indicate that heading-aware interpolation remains an open challenge, which could be addressed in future work through angular constraints or dedicated directional loss functions.

5. Conclusion

This study presented a comprehensive framework for enhancing AIS time-series data quality through a two-stage strategy combining kinematic-based denoising and Transformer-driven intelligent interpolation.

The denoising module effectively removes physically implausible, duplicated, and inconsistent AIS records, reducing approximately 39% of raw entries in the DMA dataset while preserving realistic vessel motion patterns. This preprocessing step significantly improves trajectory continuity and provides stable inputs for downstream modeling.

For missing-data reconstruction, the proposed Transformer-based interpolation model demonstrates superior performance over traditional linear and Kalman-based methods, particularly under complex and long-gap missing scenarios. The dual-stage refinement mechanism and mask-aware temporal embedding enable robust recovery of spatial positions, speed, and vessel heading while maintaining kinematic consistency.

Overall, the integrated framework provides a reliable and scalable solution for AIS trajectory reconstruction, with strong potential for real-world maritime surveillance, traffic analysis, and intelligent navigation systems.

Future research may explore multi-source data fusion by integrating AIS with radar and satellite observations, as well as physics-informed constraints to further enhance trajectory reliability in extreme maritime environments.

References

1. De Souza, E.N., et al., *Improving fishing pattern detection from satellite AIS using data mining and machine learning.* PloS one, 2016. **11**(7): p. e0158248.
2. Zheng, Y., et al. *Time series classification using multi-channels deep convolutional neural networks*. in *International conference on web-age information management*. 2014. Springer.
3. Cheng, X., et al., *An Integrated Deep Learning-Based Approach for Fishing Vessel Classification Using Ais Data.* Available at SSRN 4395308.
4. Huang, I.-L., et al., *Ship classification based on ais data and machine learning methods.* Electronics, 2023. **13**(1): p. 98.

5. Meyer, R., *Vessel classification using AIS data.* Ocean Engineering, 2025. **319**: p. 120043.
6. Liang, M., et al., *AISClean: AIS data-driven vessel trajectory reconstruction under uncertain conditions.* Ocean Engineering, 2024. **306**: p. 117987.
7. Nguyen, D., et al. *A multi-task deep learning architecture for maritime surveillance using AIS data streams*. in *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*. 2018. IEEE.
8. Chen, S., Y. Huang, and W. Lu, *Anomaly detection and restoration for ais raw data.* Wireless Communications and Mobile Computing, 2022. **2022**(1): p. 5954483.
9. Zhao, L., G. Shi, and J. Yang, *Ship trajectories pre-processing based on AIS data.* The Journal of Navigation, 2018. **71**(5): p. 1210-1230.
10. Yan, R., et al., *Development of denoising and compression algorithms for AIS-based vessel trajectories.* Ocean Engineering, 2022. **252**: p. 111207.
11. Zhang, D., et al. *Enhance the AIS data availability by screening and interpolation.* in *2017 4th International Conference on Transportation Information and Safety (ICTIS)*. 2017. IEEE.
12. Li, S., et al. *AIS-based vessel trajectory reconstruction with U-Net convolutional networks*. in *2020 IEEE 5th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA)*. 2020. IEEE.
13. Zhang, L., et al., *A novel ship trajectory reconstruction approach using AIS data.* Ocean Engineering, 2018. **159**: p. 165-174.
14. Chen, Y., et al., *Optimizing vessel trajectories: Advanced denoising and interpolation techniques for AIS data.* Ocean Engineering, 2025. **327**: p. 120988.
15. Rong, H., A. Teixeira, and C.G. Soares, *Data mining approach to shipping route characterization and anomaly detection based on AIS data.* Ocean Engineering, 2020. **198**: p. 106936.
16. Deng, C., et al., *Graph Signal Variation Detection: A novel approach for identifying and reconstructing ship AIS tangled trajectories.* Ocean Engineering, 2023. **286**: p. 115452.
17. Nguyen, D, *A transformer network with sparse augmented data representation and cross entropy loss for ais-based vessel trajectory prediction.* IEEE Access, 2024. **12**: p. 21596-21609.
18. Liu, J., J. Li, and C. Liu, *AIS-based kinematic anomaly classification for maritime surveillance.* Ocean Engineering, 2024. **305**: p. 118026.
19. Du, W., D. Côté, and Y. Liu, *Saits: Self-attention-based imputation for time series.* Expert Systems with Applications, 2023. **219**: p. 119619.

Phương pháp nâng cao chất lượng dữ liệu chuỗi thời gian AIS của tàu thuyền bằng cách khử nhiễu và nội suy thông minh

Tóm tắt: *Dữ liệu chuỗi thời gian AIS thường gặp hai vấn đề chính: nhiễu, ngoại lệ và thiếu dữ liệu quỹ đạo. Để giải quyết, nghiên cứu này đề xuất một khung phục hồi AIS hai giai đoạn, kết hợp khử nhiễu dựa trên động học và nội suy dựa trên Transformer. Ở giai đoạn đầu, các điểm quỹ đạo phi thực tế, dư thừa và ngoại lệ được loại bỏ, giúp giảm tới 39% dữ liệu nhiễu nhưng vẫn bảo toàn tính nhất quán của quỹ đạo. Ở giai đoạn sau, các đoạn dữ liệu bị thiếu được khôi phục từ chuỗi AIS không hoàn chỉnh bằng mô hình Transformer. Kết quả thực nghiệm cho thấy phương pháp đề xuất đạt độ chính xác khôi phục cao hơn so với nội suy tuyến tính và lọc Kalman, đặc biệt hiệu quả với các khoảng thiếu dữ liệu kéo dài.*

Từ khoá: Hệ thống nhận dạng tự động, tái tạo quỹ đạo, cơ chế tự chú ý.

Exploiting Vision-Language Models In Assistive Navigation Support For Visually Impaired

Duc V. Tran[1], Soan T. M. Duong[1]
[1]Le Quy Don Technical University

Abstract

Providing navigation assistance for visually impaired individuals is a critical challenge in computer vision and deep learning, as it requires a comprehensive understanding of complex and continuously changing environments to ensure absolute user safety. This problem aims to develop automated mobility assistance solutions that can respond in real time while maintaining low computational cost. Unlike traditional computer vision approaches, which typically focus on solving isolated tasks such as pedestrian lane segmentation or vanishing point estimation, they often require multiple sequential processing stages to generate final guidance. This study proposes an end-to-end approach that directly maps visual inputs to natural language instructions without relying on separate perception modules. We leverage the capabilities of vision–language models (VLMs) pretrained on large-scale web datasets to enhance contextual understanding of the surrounding environment. Experimental results demonstrate that VLMs show strong potential in recognizing traffic-related situations relevant to pedestrian navigation. However, our findings also indicate that task-specific fine-tuning is necessary to improve accuracy and reliability to meet the safety requirements of real-world navigation assistance systems.

Keywords: *Visually impaired, assistive navigation, vision–language models, foundation model.*

1. Introduction

Vision impairment significantly affects independent mobility, making safe navigation in complex urban environments a daily challenge for millions of individuals worldwide. While traditional assistive tools, e.g., white canes, guide dogs, and GPS-based navigation, provide essential support, often lack detailed, real-time environmental understanding [1]. In particular, existing systems may struggle to convey dynamic scene information, obstacle positioning, and context-aware guidance in a concise and actionable manner. Recent advances in artificial intelligence, particularly in multimodal learning, offer opportunities to address limitations.

Vision–language models (VLMs) have emerged as a powerful paradigm that integrates visual perception and natural language generation within a unified framework [2]. By combining visual encoders with large language models, VLMs are capable of interpreting complex scenes and generating coherent textual descriptions or instructions. Their ability to reason across modalities makes them especially promising for assistive navigation, where systems must transform visual input into safety-critical guidance. However, most existing VLM research emphasizes general-purpose benchmarks, multimodal reasoning, or large-scale performance improvements, with limited focus on real-world assistive deployment.

Assistive navigation poses unique challenges beyond generic image captioning or visual question answering. Guidance must be concise, directionally precise, and safety-aware, often under strict latency and memory constraints. Moreover, deployment scenarios often involve portable or edge devices, where computational efficiency becomes as important as output quality. Despite the growing number of VLM architectures, systematic investigation of their efficiency, prompt sensitivity, and suitability for assistive navigation remains underexplored.

In this work, we investigate the use of multiple VLMs for assistive navigation tailored to visually impaired users. We focus on three aspects: selecting lightweight open-access models, examining the impact of prompt design on guidance clarity and safety, and quantitatively evaluating latency, memory usage, and alignment with human references using the WAD2024 benchmark [3]. By systematically analyzing efficiency and output quality under consistent experimental settings, we highlight trade-offs between model complexity, prompt structure, and deployment feasibility, contributing to more reliable and practical VLM-based assistive navigation systems. The contributions of this study are highlighted as follows:

- We propose a systematic evaluation framework to assess the efficiency of vision–language models for assistive navigation, considering latency, memory consumption, and output alignment.
- We conduct extensive experiments across multiple VLM architectures using diverse prompt designs with varying levels of structural complexity.
- We derive meaningful insights from the experimental results, highlighting trade-offs between model capacity, prompt structure, and deployment feasibility in real-world assistive scenarios.

The remainder of this paper is organized as follows. Section 2 reviews related VLM research and assistive navigation applications. Section 3 describes the methodology, and Section 4 presents the experimental results and analysis. Finally, Section 5 concludes the paper and discusses future work.

2. Literature review

2.1. Vision-language models

Vision–language models (VLMs) unify visual perception and natural language understanding, enabling multimodal reasoning over images and text [2]. Most VLMs follow a modular architecture comprising a visual encoder, a fusion mechanism, and a language decoder. The visual encoder extracts image representations, the fusion module aligns them with textual features, and the decoder generates responses conditioned on both modalities. Visual encoders have evolved from convolutional neural networks to Vision Transformers (ViTs), which are now widely adopted for their scalability and strong global modeling ability [4–11]. Some models use pretrained CLIP-style backbones [6], while others train visual encoders jointly with language models [12], with differences in encoder design affecting visual grounding and downstream performance.

Fusion strategies vary across architectures. Single-stream models jointly process visual and textual tokens in one transformer [5,8], whereas dual-stream models maintain separate encoders and align them through cross-attention [13]. In many recent large-scale VLMs, visual features are projected into the embedding space of large language models and treated as prefix prompts to the decoder [12,14,15]. These design choices influence both computational cost and multimodal reasoning capability. The language decoder is typically based on autoregressive large language models [16,17], with performance shaped by factors such as model scale, instruction tuning, and alignment methods, including supervised fine-tuning and reinforcement learning from human feedback [18]. Although larger decoders generally improve reasoning and fluency, they also increase memory and latency demands [19].

Recent studies have therefore focused on edge-oriented optimization for real-world deployment [20]. Lightweight VLMs reduce computational cost through parameter reduction, visual embedding compression, quantization, and parameter-efficient adapters [21–23]. Compact multimodal models such as MiniCPM [24] show that carefully designed small-scale architectures can still deliver competitive reasoning performance while significantly improving efficiency. These advances are especially valuable for assistive systems that require real-time feedback on portable or embedded devices.

2.2. VLMs for assistive navigation

VLMs have recently been explored for assistive navigation for visually impaired individuals, where models must generate concise, safety-critical guidance from visual input. The WAD2024 benchmark, introduced by Yuan et al. [3], provides a large-scale dataset of walking scenarios paired with human-annotated navigation reminders. They also propose WalkVLM, a framework that generates navigation reminders from visual inputs, demonstrating that structured prompts and task-specific instructions significantly improve the clarity and safety relevance of the generated guidance.

Along the same research line, CrossCompanion [25] introduces the Empathy Cross-Safe dataset (7,886 annotated street-level images) and integrates a lightweight VLM with real-time speech recognition and text-to-speech to enable conversational guidance. Beyond technical accuracy, it emphasizes emotional support and user confidence, reporting improvements over baselines in both navigation performance and empathy-related metrics, while remaining deployable on standard mobile hardware. However, this work focuses primarily on street-crossing scenarios rather than general walking environments.

Another study on fine-tuning vision–language models for indoor navigation assistance adapts BLIP-2 [15] using low-rank adaptation [21] to generate step-by-step indoor guidance [26]. Recognizing limitations in generic VLM outputs, the model is fine-tuned on a manually annotated indoor dataset, with a refined metric emphasizing directional and sequential information. Their results show that parameter-efficient adaptation substantially improves the generation of precise directional instructions, particularly in indoor environments where GPS-based localization is unreliable. It is worth noting that the work targets indoor environments and robot navigation rather than broader assistive walking scenarios.

Overall, although many vision–language models have been developed, their application to assistive navigation is limited. Existing studies emphasize task performance and interaction in specific scenarios (e.g., street crossing or indoor navigation), rather than general walking environments. In addition, systematic evaluation of computational efficiency and deployment feasibility remains underexplored. Therefore, further investigation into the efficiency of VLMs for real-world assistive navigation is needed.

3. Methodology

3.1. Model Section

To ensure a fair and practically meaningful evaluation, we select vision–language models according to three criteria. First, the models must be foundation VLMs pretrained on large-scale multimodal corpora, enabling general visual understanding and language generation without task-specific fine-tuning. Second, the models must be released under an open or

permissive license, allowing reproducibility and transparent academic evaluation. Third, the models should exhibit relatively low memory requirements and efficient inference times, making them suitable for deployment in resource-constrained assistive systems. Based on these criteria, we select five representative open-source VLMs:

- Qwen3-VL-2B [10]: A compact 2B-parameter multimodal model in the Qwen-VL series, designed for efficient visual–language understanding and reasoning with improved image–text alignment.
- Qwen3.5-VL-2B [27]: An enhanced version of Qwen3-VL-2B, offering stronger multimodal reasoning and instruction-following, with improved robustness across tasks.
- Moondream2 [11]: A lightweight and efficient vision–language model designed for edge deployment, emphasizing a compact architecture and fast inference.
- SmolVLM2 [9]: A compact multimodal model optimized for a reduced memory footprint while maintaining competitive performance on visual understanding tasks.
- DeepSeek-VL2-Tiny [28]: A lightweight multimodal foundation model in the DeepSeek-VL2 family, designed for efficient visual understanding and reasoning.

3.2. Prompt design

To examine the influence of prompt formulation on VLM behavior, we design four prompts with progressively increasing levels of structure and task specificity. Prior work has shown that model outputs are sensitive to instruction clarity, structural constraints, and exemplar guidance [29, 30]. Our prompt set is therefore constructed to evaluate these factors systematically. Prompt 1 serves as a minimal baseline with minimal and open-ended instruction. Prompt 2, adapted from WalkVLM [3], introduces structured guidance for assistive navigation, specifying directional and safety-related requirements. Prompt 3 further constrains the task by defining the model role, limiting output length, and enforcing a clock-based directional format. Prompt 4 extends Prompt 3 by adding exemplar outputs. The detailed prompts are presented below.

Prompt 1: *Simple instruction*

Describe this scene.

Prompt 2: *Navigation-oriented instruction suggested by [3]*

Given the input image, generate a clear and concise navigation prompt for visually impaired individuals. The prompt should:
1. Identify key elements in the environment (e.g., obstacles, landmarks, pedestrians, vehicles).
2. Use clear directional terms (e.g., left, right, front, or clock directions like 1 o'clock).
3. Provide specific details about objects (e.g., shape, size, material, distance).
4. Include action suggestions (e.g., avoid, move forward, turn left).
5. Prioritize safety and clarity.

Prompt 3: *Constrained navigation instruction*

You are an assistive vision module.
Task: Generate brief alerts (<25 words) indicating ground-level obstacles or safe directions within 30 meters. Report obstacles like vehicles, pedestrians, roadblocks, and safe paths.
Instructions: Use clock directions: 12 o'clock = ahead, 3 o'clock = right, 9 o'clock = left. Mention what type of obstacle it is and the action (e.g., slow down, avoid). Avoid reporting unnecessary details or background information. Smooth response.

***Prompt 4*:** *Example-augmented navigation instruction*

You are an assistive vision module.
Task: Generate brief alerts (<25 words) indicating ground-level obstacles or safe directions within ~30 meters. Report obstacles like vehicles, pedestrians, roadblocks, and safe paths.
Instructions: Use clock directions: 12 o'clock = ahead, 3 o'clock = right, 9 o'clock = left. Mention obstacle type and the action (e.g., slow down, avoid). Avoid reporting unnecessary details or background information.
Example outputs:
"Obstacle ahead at 9 o'clock, avoid and go right."
"At 11 o'clock, pedestrians crossing ahead, slow down to avoid."
"Roadblock at 12 o'clock, turn slightly right."
"Trash can five steps ahead, keep left to pass safely."
"The path is clear, continue forward without worry."

3.3. Dataset

We conduct our experiments on the WAD2024 dataset [3], currently the largest publicly available benchmark for walking assistance systems designed for visually impaired individuals. WAD2024 provides high-quality ground-truth navigation reminders paired with real-world video data, enabling quantitative evaluation of assistive vision–language models. The dataset contains 12,000 video–manual-annotation pairs collected from diverse urban environments across Europe and Asia. It covers a wide range of real-world conditions, including varying location types, weather conditions, danger levels, and traffic flow ratings, thereby ensuring ecological validity and scenario diversity. In this study, we use the WAD2024 test set for evaluation. The test set consists of 1,007 videos, each associated with eight consecutive frames and a shared ground-truth navigation reminder (reference). Since all eight frames correspond to the same textual annotation, we select the eighth frame for inference to maintain consistency while avoiding redundancy across temporally adjacent frames.

3.4. Experimental method

All experiments are conducted in a consistent environment using Python 3.13.5, PyTorch 2.1, and CUDA 13.0 on an NVIDIA T1000 GPU with 8GB VRAM. Each run performs inference for a prompt–VLM pair on the WAD2024 test set. We quantitatively evaluate each prompt across VLMs for latency, GPU memory usage, and output alignment with human references. Latency is measured from prompt input to output completion on a workstation with an Intel Core i7-12700, 16GB RAM, and an NVIDIA T1000 GPU. GPU memory usage is reported as peak allocation during inference based on PyTorch CUDA statistics.

To assess the output quality, we employ TF-IDF (Term Frequency–Inverse Document Frequency) cosine similarity, and ROUGE (Recall-Oriented Understudy for Gisting Evaluation) metrics. TF-IDF computes the cosine similarity between TF-IDF weighted term vectors of the generated and reference texts; higher values indicate stronger lexical alignment. Hereafter, we use TF-IDF as shorthand for TF-IDF cosine similarity. ROUGE-N measures N-gram overlap between the generated reminder and the reference [31], quantifying how well key words and phrases are captured. We also employ ROUGE-L, which is based on the longest common subsequence and captures sequence-level similarity, better reflecting structural coherence and overall alignment [31]. Higher ROUGE scores indicate better content coverage.

4. Experimental result and analysis

4.1. Experimental results

Table 1 presents the aggregated performance comparison across five vision-language models and four prompting strategies. The results are reported as mean ± standard deviation over the test set. We evaluate performance using TF-IDF similarity and ROUGE metrics (ROUGE-1, ROUGE-2, and ROUGE-L), together with computational efficiency indicators including elapsed inference time and peak memory usage.

Table 1: Aggregated performance comparison across models and prompts. Elaps. time denotes elapsed time. Columns 3 to 7 report results as mean ± standard deviation. Column 8 presents the mean peak memory usage in megabytes (MB). The final column (ΔR−1) reports the improvement in ROUGE-1 relative to Prompt 1 within the same VLM model. Bold values represent the best performance within each model

Model	P	TF-IDF↑	ROUGE-1↑	ROUGE-2↑	ROUGE-L↑	Elaps. time (s) ↓	PGU↓	ΔR−1
Qwen3-VL-2B	1	0.08 ± 0.06	0.05 ± 0.03	0.01 ± 0.01	0.04 ± 0.02	165.94 ± 58.72	12288	0.00
	2	0.09 ± 0.07	0.10 ± 0.05	0.02 ± 0.02	0.07 ± 0.04	177.55 ± 167.15	12083	+0.05
	3	0.06 ± 0.07	0.14 ± 0.09	0.04 ± 0.04	0.12 ± 0.08	145.79 ± 57.50	12083	+0.09
	4	0.08 ± 0.06	**0.24 ± 0.11**	**0.06 ± 0.04**	**0.22 ± 0.10**	149.89 ± 53.47	12390	**+0.19**
Qwen 3.5-VL-2B	1	0.08 ± 0.05	0.04 ± 0.02	0.01 ± 0.01	0.03 ± 0.01	149.69 ± 53.60	12390	0.00
	2	0.09 ± 0.06	0.07 ± 0.03	0.01 ± 0.01	0.05 ± 0.02	137.16 ± 186.74	12390	+0.03
	3	0.05 ± 0.07	0.10 ± 0.09	0.01 ± 0.03	0.09 ± 0.08	131.65 ± 51.51	12390	+0.06
	4	0.09 ± 0.07	**0.26 ± 0.13**	**0.08 ± 0.06**	**0.24 ± 0.12**	131.29 ± 51.39	12390	**+0.22**
Smol VLM2	1	0.06 ± 0.05	0.07 ± 0.04	0.01 ± 0.01	0.06 ± 0.03	24.19 ± 3.23	5837	0.00
	2	0.08 ± 0.07	0.05 ± 0.06	0.01 ± 0.03	0.04 ± 0.05	77.45 ± 58.65	5837	-0.02
	3	0.05 ± 0.07	0.05 ± 0.04	0.00 ± 0.01	0.04 ± 0.03	17.03 ± 0.58	5837	-0.02
	4	0.08 ± 0.05	**0.24 ± 0.11**	**0.06 ± 0.04**	**0.22 ± 0.11**	14.47 ± 14.46	5837	**+0.17**
Moon dream2	1	0.07 ± 0.05	0.07 ± 0.04	0.01 ± 0.01	0.06 ± 0.03	12.30 ± 1.88	5120	0.00
	2	0.05 ± 0.06	0.12 ± 0.06	0.01 ± 0.02	0.09 ± 0.05	9.43 ± 0.50	5120	+0.04
	3	0.06 ± 0.07	0.14 ± 0.08	0.01 ± 0.03	0.11 ± 0.07	9.28 ± 5.08	5120	+0.06
	4	0.11 ± 0.06	**0.22 ± 0.09**	**0.07 ± 0.05**	**0.19 ± 0.07**	9.74 ± 0.36	5120	**+0.14**
Deep seek-VL2-Tiny	1	0.07 ± 0.05	0.08 ± 0.05	0.02 ± 0.02	0.07 ± 0.03	30.79 ± 3.50	8397	0.00
	2	0.08 ± 0.06	0.10 ± 0.06	0.02 ± 0.02	0.07 ± 0.04	38.14 ± 32.66	8499	+0.01
	3	0.07 ± 0.07	0.15 ± 0.10	0.04 ± 0.06	0.13 ± 0.09	27.04 ± 2.90	8499	+0.07
	4	0.08 ± 0.04	**0.25 ± 0.11**	**0.06 ± 0.04**	**0.23 ± 0.10**	27.14 ± 2.84	8602	**+0.16**

Across all models, Prompt 4 consistently achieves the best overall performance in terms of ROUGE scores. For Moondream2, ROUGE-1 shows the smallest gain, increasing from 0.07 to 0.22 (+0.14). Similar improvements are observed for Qwen3-VL-2B, SmolVLM2, Deepseek-VL2-Tiny, and Qwen3.5-VL-2B, with Prompt 4 yielding the highest ROUGE-1 scores of 0.24, 0.24, 0.25, and 0.26, respectively. Improvements are also reflected in ROUGE-2 and ROUGE-L scores.

In terms of computational cost, Moondream2 demonstrates the fastest inference time (approximately 9–12 seconds on average) with the lowest memory consumption (5.1 GB). SmolVLM2 also shows competitive efficiency, while Deepseek-VL2-Tiny operates at a moderate cost. In contrast, Qwen3-VL-2B and Qwen3.5-VL-2B require substantially longer inference time and higher memory usage.

Table 2: Examples of navigation guidance produced for the same image under different prompting strategies and VLM models. The corresponding ground-truth is "There is a road sign five steps ahead. Please keep walking slowly on the current road.".

Model	Prompt 1	Prompt 2	Prompt 3	Prompt 4
Qwen3-VL-2B	This image captures a city street scene ... and natural elements. (220 words)	Navigate the road ahead, ... road is clear of obstacles. (74 words)	3 o'clock: Roadblock ahead. Avoid. 9 o'clock: ... (12 words)	Obstacle ahead at 3 o'clock, avoid and go left. (9 words)
Qwen 3.5-VL-2B	This is a street-level ... under a soft, twilight sky. (334 words)	Navigation Prompt ... tailored to visual impairment. (133 words)	Left: Vehicle ahead. Right: Clear path. Proceed with caution. (9 words)	Obstacle ahead at 9 o'clock, slow down and avoid. (9 words)
Smol VLM2	The image captures a bustling urban scene ... ongoing development and growth. (162 words)	Starting from the left, proceed straight ... right onto the side street. (108 words)	1. Pedestrian crossing ahead, slow down. ... 10. Roadblock ahead, avoid. (50 words)	Obstacle ahead at 9 o'clock, avoid and go right. (9 words)
Moon dream2	A tall skyscraper under construction ... construction site being a focal point. (112 words)	A street scene with construction activity ... and turning left. (34 words)	A vehicle is approaching from the right. A pedestrian is ... roadblock is ahead. (25 words)	Obstacle ahead at 9 o'clock, slow down to avoid. ... to pass safely. (26 words)
Deep seek-VL2-tiny	The image depicts an urban street scene ... to the urban landscape. (110 words)	Turn right at the next intersection. ... at the intersection. (878 words)	There is a vehicle ahead on the road. Avoid this area. (11 words)	Obstacle ahead at 9 o'clock, avoid and go right. (9 words)

Table 2 presents qualitative examples of prompting strategies across five VLMs. With Prompt 1 (minimal instruction), all models generate lengthy, general scene descriptions, ranging from 110 to 334 words (Qwen3.5-VL-2B), focusing on environmental details rather than actionable navigation cues. Prompt 2 reduces verbosity and introduces directional guidance, though outputs remain largely descriptive and occasionally inconsistent in length (e.g., DeepSeek-VL2-Tiny produces an outlier of 878 words). Prompt 3 enforces clearer structural constraints (e.g., clock-based references), resulting in shorter, more task-oriented guidance. Prompt 4 produces the most concise and directive outputs across all models, typically under 30 words, with explicit obstacles and actionable instructions.

4.2. Discussion

The experimental results yield several key findings. First, prompting strategy influences output quality more strongly than model selection alone, as all five VLMs achieve clear ROUGE improvements from Prompt 1 to Prompt 4, highlighting the importance of structured,

task-oriented prompting for assistive navigation. Second, a clear trade-off exists between performance and efficiency: Qwen3.5VL-2B attains the largest ROUGE-1 gain (+0.22) but requires substantially higher inference time and memory, whereas SmolVLM2 provides a more balanced compromise for real-time deployment.

Third, inference latency does not scale linearly with prompt complexity. Prompt 1, despite being the simplest, often incurs relatively long latency because its underspecified instruction tends to produce longer, more verbose outputs. Prompt 4 is generally slightly slower than Prompts 2 and 3 due to its longer and more structured input, indicating that latency depends on both input and output token lengths rather than prompt complexity alone. Fourth, peak memory usage remains largely stable across prompts within each model, suggesting that memory consumption is determined primarily by model architecture and parameter size.

Fifth, qualitative results confirm that prompt design strongly shapes model behavior: underspecified prompts yield generic descriptions, whereas structured prompts produce concise and actionable guidance. By incorporating exemplar-style instructions, Prompt 4 further enforces the "obstacle + direction + action" format, which likely explains its consistently superior ROUGE performance, although excessive examples may reduce generalization. Finally, although ROUGE and TF-IDF scores remain relatively low, this is expected because navigation guidance can be expressed in diverse yet semantically valid ways, while these metrics mainly capture lexical overlap with a single reference.

Overall, the findings suggest that high-quality navigation assistance depends not only on model capability but also on effective prompt design; the most practical solution is therefore to combine task-aware prompting, such as Prompt 4, with a lightweight yet capable VLM to balance guidance quality, latency, and memory usage.

5. Conclusion

This paper investigated the effectiveness of vision-language models (VLMs) for assistive navigation support for visually impaired individuals. Through a systematic comparison of multiple VLMs and prompting strategies, we evaluated both descriptive quality and computational efficiency using ROUGE metrics, TF-IDF similarity, inference latency, and peak memory usage. The results show that structured prompts with specific illustrative examples significantly improve the quality and task relevance of navigation guidance across all models. Inference latency is influenced more by generated token length than by prompt complexity, while peak memory usage remains largely stable within each model. Notably, SmolVLM2 achieves the lowest inference time, highlighting its suitability for real-time deployment. In addition, our findings indicate that delivering high-quality navigation guidance depends not only on capable VLMs but, more importantly, on carefully crafted prompts that promote concise, goal-oriented outputs. For real-time assistive systems, pairing lightweight models with structured prompts provides the best balance between performance and responsiveness. Future work will include real-world user evaluations, investigation of VLM latency under practical deployment conditions, and the development of more suitable evaluation metrics that better capture semantic correctness and usability, along with adaptive prompt refinement to further enhance reliability and effectiveness.

References

[1] Y. Lei, Phung S. L., Di Y., and Abdesselam B., "CSN: A compact semantic segmentation network for visual scene perception in assistive navigation," Computer Vision and Image Understanding, vol. 264, no. April 2025, pp. 104665, 2026.

[2] Y. Du, Z. Liu, J. Li, and W. X. Zhao, "A survey of vision-language pre-trained models," in *Proc. Inter. Joint Conf. Artificial Intelligence*, 2022, pp. 5436–5443.

[3] A. Owens, "WalkVLM: Aid visually impaired people walking by vision language model," in *Proc. Inter. Conf. Computer Vision*, 2025, pp. 6945–6956.

[4] L. H. Li, M. Yatskar, D. Yin, C.-J. Hsieh, and K.-W. Chang, "VisualBERT: A simple and performant baseline for vision and language," *arXiv preprint arXiv:1908.03557v1*, pp. 1–14, 2019.

[5] Y.-C. Chen, L. Li, L. Yu, A. El Kholy, F. Ahmed, Z. Gan, Y. Cheng, and J. Liu, "UNITER: Universal image-text representation learning," in *Prof. European Conf. Computer Vision*, 2020.

[6] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, and et al., "Learning transferable visual models from natural language supervision," in *Proc. Inter. Conf. Machine Learning*, 2021, vol. 139, pp. 8748–8763.

[7] S. Bai, K. Chen, X. Liu, J. Wang, W. Ge, and et al., "Qwen2.5-VL technical report," 2025.

[8] N. Gupta, H. Narasimhan, W. Jitkrittum, and A. S. Rawat, "SimVLM: Simple visual language model pre-training with weak supervision," in *Proc. Inter. Conf. Learning Representations*, 2024, number 02, pp. 1–34.

[9] A. Marafioti, O. Zohar, M. Farre´, M. Noyan, E. Bakouch, and et al., "SmolVLM: Redefining small and efficient multimodal models," *arXiv preprint arXiv:2504.05299v1*, pp. 1–20, 2025.

[10] J. Bai, S. Bai, S. Yang, S. Wang, S. Tan, and et al., "Qwen-VL: A versatile vision-language model for understanding, localization, text reading, and beyond," *arXiv preprint arXiv:2308.12966*, pp. 1–24, 2023.

[11] M87 Labs, Inc., "Overview — moondream docs," https://docs.moondream.ai/, 2025, Accessed: 2026-01-09.

[12] J.-B. Alayrac, J. Donahue, P. Luc, A. Miech, I. Barr, and et al., "Flamingo: A visual language model for few-shot learning," in *Proc. Neural Inf. Process. Systems*, 2022.

[13] J. Lu, D. Batra, D. Parikh, and S. Lee, "ViLBERT: Pretraining task-agnostic visio linguistic representations," in *Proc. Neural Inf. Process. Systems*, 2019, pp. 1–11.

[14] H. Liu, C. Li, Q. Wu, and Y. J. Lee, "Visual instruction tuning," in *Proc. Neural Inf. Process. Systems*, 2023, pp. 1–25.

[15] J. Li, D. Li, S. Savarese, and S. Hoi, "BLIP-2: Bootstrapping language-image pre-training with frozen image encoders and large language models," in *Proc. Inter. Conf. Machine Learning*, 2023, vol. 202, pp. 20351–20383.

[16] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, and et al., "Language models are few-shot learners," in *Proc. Neural Inf. Process. Systems*, 2020, vol. 2020-Decem.

[17] H. Touvron, T. Lavril, G. Izacard, X. Martinet, M.-A. Lachaux, and et al., "LLaMA: Open and efficient foundation language models," *arXiv preprint arXiv:2302.13971*, 2023.

[18] L. Ouyang, J. Wu, X. Jiang, C. L. Wainwright, D. Almeida, and et al., "Training language models to follow instructions with human feedback," in *Proc. Neural Inf. Process. Systems*, 2022.

[19] B. Ghorbani, O. Firat, M. Freitag, A. Bapna, M. Krikun, and et al., "Scaling laws for neural machine translation," in *Proc. Inter. Conf. Learning Representations*, 2022, pp. 1–42.

[20] A. Sharshar, L. U. Khan, W. Ullah, and M. Guizani, "Vision-language models for edge networks: A comprehensive survey," *IEEE Internet of Things J.*., vol. 12, no. 16, pp. 32701–32724, 2025.

[21] E. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, and et al., "LoRA: Low-rank adaptation of large language models," in *Proc. Inter. Conf. Learning Representations*, 2022, pp. 1–13.

[22] T. Dettmers, A. Pagnoni, A. Holtzman, and L. Zettlemoyer, "QLoRA: Efficient finetuning of quantized LLMs," in *Proc. Neural Inf. Process. Systems*, 2023.

[23] E. Frantar, S. Ashkboos, T. Hoefler, and D. Alistarh, "GPTQ: Accurate post-training quantization for generative pre-trained transformers," *arXiv preprint arXiv:2210.17323v2*, pp. 1–16, 2023.

[24] Yuan Yao, Tianyu Yu, and Ao Zhang a, "Efficient GPT-4V level multimodal large language model for deployment on edge devices," *Nature Communications*, vol. 16, pp. 5509, 2025.

[25] J. Liang and Y. Fang, "CrossCompanion: An empathetic real-time assistant supporting street crossing for low-vision users," in *Proc. Inter. Conf. Computer Vision*, 2025, pp. 2466–2475.

[26] X. Li, B. Gandhi, M. Zhan, M. Nehra, Z Zhang, and et al., "Fine-tuning vision-language models for visual navigation assistance," in *Proc. Inter. Conf. Machine Learning*, 2025.

[27] Qwen Team, "Qwen3.5: Towards native multimodal agents," https://qwen.ai/blog?id= qwen3.5, 2026, Accessed: 2026-03-28.

[28] Z. Wu, X. Chen, Z. Pan, X. Liu, W. Liu, and et al., "DeepSeek-VL2: Mixture-of-experts vision-language models for advanced multimodal understanding," *arXiv preprint arXiv:2412.10302v1*, pp. 1–28, 2024.

[29] Q. Ye, M. Axmed, R. Pryzant, and F. Khani, "Prompt engineering a prompt engineer," in *Proc. Association for Computational Linguistics*, 2024, pp. 355–385.

[30] K. Cheng, N. K. Ahmed, T. L. Willke, and Y. Sun, "Structure guided prompt: Instructing large language model in multi-step reasoning by exploring graph structure of the text," in *Proc. Conf. Empirical Methods in Natural Language Processing*, 2024, pp. 9407–9430.

[31] C.-Y. Lin, "ROUGE: A package for automatic evaluation of summaries," in *Proc. Wksp. Text Summarization Branches Out*, 2004, number 1, pp. 74–81.

Khai thác tiềm năng của mô hình ngôn ngữ thị giác (VLM) trong hỗ trợ dẫn đường cho người khiếm thị

Tóm tắt: *Hỗ trợ dẫn đường cho người khiếm thị là một thách thức quan trọng trong lĩnh vực thị giác máy tính và học sâu, bởi yêu cầu hiểu bối cảnh môi trường phức tạp và thay đổi liên tục để đảm bảo an toàn tuyệt đối cho người dùng. Các nghiên cứu hỗ trợ dẫn đường bằng thị giác máy tính và học sâu hướng tới mục tiêu tạo ra các giải pháp hỗ trợ di chuyển tự động, phản hồi nhanh với chi phí thấp. Khác với các phương pháp thị giác máy tính truyền thống thường chỉ tập trung giải quyết các bài toán đơn lẻ như phân vùng làn đường đi bộ (pedestrian lane segmentation) hay ước lượng điểm biến ảo (vanishing point estimation) vốn đòi hỏi nhiều bước để đưa ra chỉ dẫn cuối cùng. Nghiên cứu này đề xuất một hướng tiếp cận toàn diện (end-to-end) với việc nhận hình ảnh và đưa ra chỉ dẫn bằng ngôn ngữ tự nhiên mà không cần thông qua các module nhận diện rời rạc. Chúng tôi khai thác sức mạnh của các mô hình ngôn ngữ thị giác (Vision-Language Models - VLMs) đã được huấn luyện trên các tập dữ liệu quy mô lớn (web-scale datasets) để đưa tối ưu hoá khả năng hiểu ngữ cảnh. Kết quả thực nghiệm cho thấy các mô hình VLMs có tiềm năng trong việc nhận diện các tình huống giao thông cho người đi bộ. Tuy nhiên, nghiên cứu cũng chỉ ra rằng cần thực hiện các kỹ thuật hiệu chỉnh (fine-tuning) chuyên biệt để nâng cao độ chính xác và tính tin cậy nhằm đáp ứng các tiêu chuẩn an toàn trong dẫn đường thực tế.*

***Từ khoá*:** Người khiếm thị, hỗ trợ dẫn đường, mô hình ngôn ngữ thị giác, mô hình nền tảng.

Tính toán các thông số làm việc của cơ cấu bơi kiểu bơm luồng phụt xe tăng hạng nhẹ sản xuất tại Việt Nam

Hoàng Khoái[1*], Nguyễn Minh Tân[1]

[1]*Học viện Kỹ thuật Quân sự*

Email: hoangkhoai20689@gmail.com; Tel: 0396373989

Tóm tắt:

Trong biên chế của Quân đội ta hiện nay đang có một số lượng tương đối lớn các xe tăng bơi thế hệ cũ do Liên Xô và Trung Quốc viện trợ. Cơ cấu bơi và hệ thống dẫn động bơi trên các xe này là kiểu cơ khí đều đã khá lạc hậu, các tính năng không còn phù hợp với các nhiệm vụ huấn luyện, tác chiến trong tình hình mới. Do vậy, Quân đội ta đang tích cực triển khai chương trình nghiên cứu, thiết kế, chế tạo một mẫu xe tăng hạng nhẹ bơi nước mới, hiện đại. Một trong những tính năng quan trọng của dòng xe này cần đạt được là khả năng bơi nước ổn định với tốc độ bơi tiến lớn nhất không nhỏ hơn 11 km/h, tốc độ bơi lùi không nhỏ hơn 3 km/h. Trong bài báo này, nhóm tác giả trình bày cơ sở lý thuyết, phương pháp xác định các thông số làm việc và tính toán thiết kế cơ cấu bơi kiểu bơm luồng phụt cho xe tăng hạng nhẹ sản xuất tại Việt Nam đáp ứng các tính năng kỹ thuật đề ra.

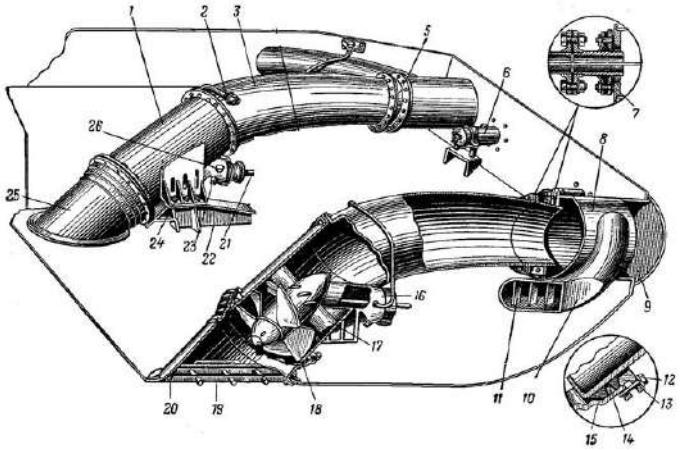
Từ khóa: *Xe tăng bơi; cơ cấu bơi; bơm luồng phụt; tốc độ bơi.*

1. Tổng quan

Hiện nay, các cơ quan đơn vị của Tổng cục CNQP phối hợp với Viện cơ giới quân sự/TCHC-KT và Học viện KTQS đang tích cực triển khai chương trình "Nghiên cứu thiết kế, chế tạo xe tăng bơi T-1" bằng công nghệ trong nước. Đây sẽ là mẫu xe tăng bơi đầu tiên do chúng ta tự chủ về mặt công nghệ, khẳng định được một bước tiến mới trong nền công nghiệp chế tạo xe quân sự của quân đội ta. Cấu hình dự kiến của xe tăng bơi hạng nhẹ T-1 (Hình 1) với khối lượng chiến đấu khoảng 18,5 tấn; kíp chiến đấu 03 người (01 trưởng xe, 01 lái xe, 01 pháo thủ); vận tốc lớn nhất trên đường nhựa không nhỏ hơn 60 km/h, vận tốc lớn nhất khi bơi tiến 11 km/h; khả năng vượt vật cản (vượt dốc đứng lớn nhất: 32 độ; vượt sườn nghiêng lớn nhất: 30 độ; vượt hào rộng lớn nhất: 2,8 m, vượt vách đứng lớn nhất: 1,1 m),…

Đối với xe tăng bơi nước, một trong những tính năng quan trọng nhất đó là khả năng cơ động và thông qua cao khi hoạt động trên cạn và dưới nước. Chính vì vậy, việc thiết kế, chế tạo hệ thống bơi cho xe đáp ứng được các yêu cầu đặt ra ở chế độ bơi là rất cần thiết. Nhóm nghiên cứu của Viện Cơ khí động lực/ Học viện KTQS được giao nhiệm vụ chủ trì phần nghiên cứu thiết kế, chế tạo cơ cấu bơi và hệ thống dẫn động bơi trên xe tăng bơi T-1. Qua kết quả nghiên cứu tổng quan các xe TTG bơi nước, nghiên cứu tổng thể các loại cơ cấu bơi và dẫn động bơi, đánh giá ưu nhược điểm của từng loại, nhóm nghiên cứu đã lựa chọn phương án thiết kế hệ thống bơi phù hợp nhất cho xe tăng bơi T-1 là: Cơ cấu bơi dạng bơm phụt với hệ thống dẫn động cơ cấu bơi bằng thủy lực mạch kín (Hình 2) [1].

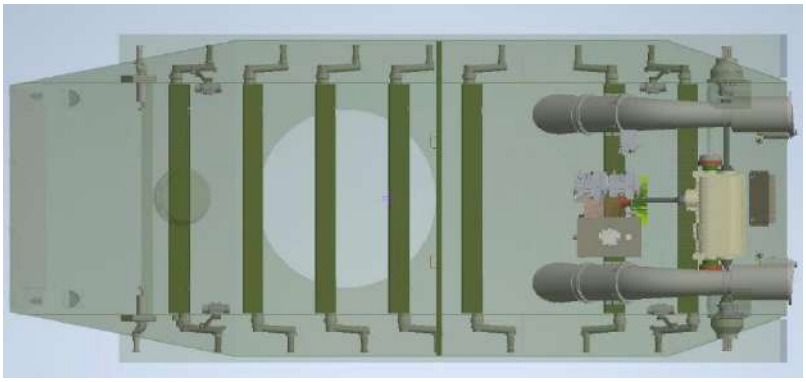
Hình 1. Xe tăng hạng nhẹ bơi nước T-1



Hình 2. Kết cấu cơ cấu bơi dạng bơm phụt xe tăng T-1

1. Thân bơm phụt; 2. Ống tra dầu; 3. Ống dẫn; 5. Ống nối; 6.Cơ cấu dẫn động nắp cửa phản lực; 7. Bộ phận làm kín; 8. Ống phụt; 9. Nắp cửa phản lực; 10. Ống bơi lùi; 11.Tấm chắn bằng thép; 12. Bu lông giằng; 13. Vòng ép; 14. Vòng trung gian; 15. Vòng làm kín; 16. Bơm phụt; 17. Cánh dẫn hướng; 18. Cánh quạt; 19. Song chắn; 20. Bu lông; 21. Trục nối với hệ thống dẫn động; 22. Bệ; 23. Nút lỗ tra dầu; 24. Bu lông; 25. Ống hút; 26. Nút kiểm tra.

Theo thiết kế tổng thể xe tăng bơi T-1, toàn bộ động cơ và hệ thống truyền lực được bố trí ở phía sau xe, do đó cơ cấu bơi cũng được bố trí ở phía sau xe để có thể dễ dàng trong quá trình dẫn động và điều khiển [2]. Hai cụm cơ cấu bơi sẽ được dẫn động trực tiếp nhờ hai mô tơ thủy lực, được trích công suất từ động cơ ngay sau Ly hợp chính (Hình 3).



Hình 3. Bố trí cơ cấu bơi trong khoang động- truyền lực

Cơ sở lý thuyết tính toán thiết kế, khảo sát các yếu tố ảnh hưởng tới các tham số làm việc của cơ cấu bơi kiểu bơm luồng phụt đã được nhiều tác giả trong và ngoài nước nghiên cứu [3,4,5,6]. Quá trình tính toán bơm luồng phụt, có các dạng bài toán như sau:

- Dạng 1: cho trước nguồn động lực, các thông số về xe như kích thước, tải trọng và biên dạng vỏ, vận tốc bơi lớn nhất, vị trí bố trí bơm, vận tốc nước, cấp sóng. Tiến hành tính toán các thông số làm việc của bơm phụt và chọn bơm cho phù hợp.

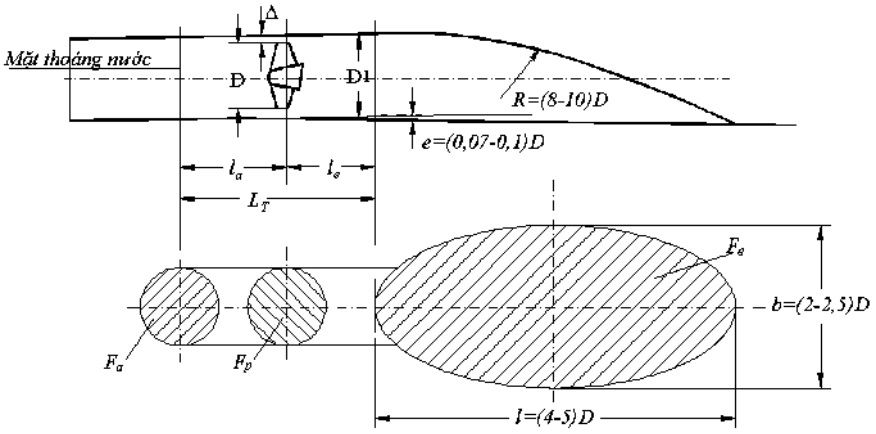
- Dạng 2: cho trước các thông số về xe như kích thước, tải trọng và biên dạng vỏ, vận tốc bơi lớn nhất, các thông số về bơm, vị trí bố trí bơm, vận tốc nước, cấp sóng. Tiến hành tính toán công suất nguồn động lực và chọn nguồn động lực cho phù hợp.

- Dạng 3: cho trước nguồn động lực, cho trước các thông số về bơm, các thông số về xe như kích thước, tải trọng và biên dạng vỏ, vận tốc bơi lớn nhất, vị trí bố trí bơm, vận tốc nước, cấp sóng. Tiến hành tính toán các thông số làm việc của bơm.

Bài toán dạng 1 và dạng 2 là bài toán thiết kế. Bài toán dạng 3 là bài toán tính toán kiểm nghiệm trong khai thác. Theo nhiệm vụ thiết kế cơ cấu bơi bơm luồng phụt cho xe tăng bơi T-1 thì nội dung này sẽ chọn dạng bài toán 1 để thực hiện [3,4].

2. Cơ sở lý thuyết

2.1. Các đặc trưng kết cấu của cơ cấu bơi kiểu bơm phụt



Hình 4. Các thông số kết cấu của ống phụt

Bơm luồng phụt bao gồm các thông số kết cấu của ống phụt (hình 4) như sau:

- Diện tích tiết diện làm việc của ống ở vị trí đặt tuốc bin [4]:

$$F_P = \frac{\pi}{4}\left(D + 2\Delta\right)^2 \tag{1}$$

Trong đó:

+ Δ là khe hở giữa mút các cánh tuốc bin và thành trong của ống phụt (m);

+ D là đường kính tuốc bin (m);

- Diện tích tiết diện nước vào của miệng hút F_e (m^2);

- Diện tích tiết diện nước ra của ống F_a (m^2);

- Chiều dài ống ℓ_T (m), là khoảng cách giữa tiết diện vào và ra nước của ống được đo trên hình chiều bằng;

- Chiều dài phần vào nước của ống ℓ_e (m);

- Chiều dài phần ra nước của ống ℓ_a (m);

- Chiều sâu chìm nước của đường trục tâm ống h_a (m) so với bề mặt tự do của nước tại tiết diện nước ra.

- Đường kính tương đối của ống:

$$\frac{D+2\Delta}{T_đ} \approx \frac{D}{T_đ} \tag{2}$$

Trong đó: $T_đ$ là độ sâu chìm nước phần đuôi phương tiện bơi.

- Hệ số thu α_e:

$$\alpha_e = \frac{F_e}{F_T} \approx \frac{F_e}{F_p} \tag{3}$$

Trong đó: $F_p = \frac{\pi}{4} D^2$

- Hệ số giản nỡ (nén):

$$\beta_\alpha = \frac{F_a}{F_T} \approx \frac{F_a}{F_p} \tag{4}$$

2.2. Các bước xác định thông số làm việc của bơm luồng phụt

Thứ tự các bước tính toán xác định thông số làm việc của bơm luồng phụt như sau:

- Xác định các thông số về đường mớn nước: chiều dài và chiều rộng mớn nước, độ sâu chìm nước;
- Xác định diện tích phần vỏ chìm trong nước, tiết diện cản chính diện;
- Lựa chọn hệ số điền đầy thể tích, hệ số lực cản nước, hệ số làm việc đồng đều của bơm luồng phụt;
- Xác định lực cản trên một bơm luồng phụt;
- Xác định hệ số ảnh hưởng của vỏ và phần vận hành;
- Xác định hệ số nén của dòng khối nước phụt;
- Xác định tiết diện phần đường ống đặt tuốc bin, miệng ống phụt;
- Xác định lưu lượng qua tiết diện tuốc bin;
- Xác định tốc độ quay lớn nhất của tuốc bin;
- Xác định hệ số lưu lượng của bơm luồng phụt;
- Xác định hệ số cản toàn phần của ống phụt;
- Xác định áp lực cột áp và hệ số áp lực cột áp của bơm luồng phụt;
- Xác định công suất cần thiết đặt vào tuốc bin.

2.3. Bộ thông số ban đầu

Bộ thông số kết cấu của xe và yêu cầu kỹ thuật:

- Công suất lớn nhất nguồn động lực có thể cung cấp: N_e = 470 (Hp);
- Kích thước xe: Dài x Rộng = 8,0 m x 3,25 m;
- Tải trọng lớn nhất của xe: G = 18,6 (tấn);
- Vận tốc bơi tiến lớn nhất: v_b = 11 (km/h) ≈ 3,056 (m/s);
- Biên dạng tuyến hình vỏ giáp: mũi vát, hai mạn thành đứng đều từ mũi đến đuôi xe;

Qua khảo sát các dòng xe xích bơi nước hiện có với kích thước, khối lượng xe và vận tốc bơi tương đương, chọn sơ bộ thông số cơ cấu bơi như sau:

- Đường kính tuốc bin: 0,345 (m);
- Đường kính moay ơ: 0,142 (m);

- Đường kính miệng phụt khi tiến: 0,240 (m);
- Đường kính miệng phụt khi lùi: 0,135 (m);
- Góc nghiêng miệng phụt khi lùi: 60 (độ);
- Cấp sóng: không xét đến.

3. Trình tự, kết quả tính toán và thảo luận

3.1. Xác định các thông số làm việc của cơ cấu bơi kiểu bơm luồng phụt

a) Xác định các thông số về đường mớn nước.

- Chiều dài mớn nước:

Chiều dài mớn nước được xác định theo biên dạng vỏ xe: L = 7,8 (m).

- Chiều rộng mớn nước:

Chiều rộng mớn nước chính là chiều rộng thân xe: B = 3,25 (m).

- Độ sâu ngập nước:

Theo định luật Archimedes, lực đẩy Archimedes (F_A) bằng trọng lực (G) của xe [3]:

$$G = F_A = \rho.g.V_{chìm} \quad (5)$$

Từ đó suy ra thể tích phần thân xe phải ngập dưới nước $V_{chìm}$:

$$V_{chìm} = \frac{m}{\rho}$$

Trong đó: m- Khối lượng xe T-1; ρ- Tỷ trọng của nước (1000 kg/m^3).

Dựa vào hình dáng hình học của thân xe, tính ra chiều cao mực nước chìm (T) tương ứng trên thân xe: T = 1,1 (m).

b) Xác định diện tích phần vỏ chìm trong nước, tiết diện cản chính diện.

- Diện tích phần vỏ chìm trong nước S [3,4]:

$$S \approx L.(B + 2T).C_m \quad (6)$$

Trong đó: C_m - Hệ số diện tích sườn. ($C_m = 0{,}9 \div 0{,}95$).

Chọn C_m= 0,95, thay số vào ta có S= 40,38 (m^2)

- Tiết diện cản chính diện :

+ Diện tích cản chính diện phần ngập nước của xe F_{cd} :

$$F_{cd} = B.T \quad (7)$$

Thay số ta được: F_{cd}= 3,575 (m^2)

+ Diện tích mặt cắt ngang phần nổi trên mặt nước của xe $F_{hγ}$:

$$F_{hγ} = B.(H - T) \quad (8)$$

Trong đó: H- Chiều cao xe, tính cả tháp pháo, H= 2,25 (m)

Từ đó xác định được: $F_{hγ}$= 3,75 (m^2)

c) Lựa chọn hệ số điền đầy thể tích, hệ số lực cản nước, hệ số làm việc đồng đều của bơm luồng phụt.

- Hệ số điền đầy đường nước là tỉ số giữa diện tích mặt mớn nước và diện tích mặt bao của xe, hay còn gọi là hệ số điền đầy mặt nổi, α.

$$\alpha = \frac{S_{mc}}{L.B} \leq 1 \quad (9)$$

Trong đó: S_{mc} là diện tích mặt cắt của xe tại mực nước. Vì thân xe có dạng hình hộp vát nhọn ở mũi, nên diện tích đường nước (S_{mc}) sẽ nhỏ hơn một hình chữ nhật (LxB). Hệ số này thường nằm trong khoảng 0,7 ÷ 0,8. Chọn α= 0,75.

- Hệ số điền đầy thể tích, δ:

Tỉ số giữa lượng giãn nước thực tế của xe và thể tích bao của nó:

$$\delta = \frac{S_{mc}}{L.B.T} \quad (10)$$

Thay các thông số vào ta được δ = 0,68.

- Hệ số lực cản của nước, ξ. Theo [4] hệ số lực cản của nước được xác định như sau:

$$\xi = \frac{2R_w}{\rho.F_{cd}.v^2} \quad (11)$$

Trong đó : R_w - Lực cản của nước; v : Vận tốc bơi của xe.

Thông thường chọn ξ = 0,7.

- Hệ số làm việc đồng đều của bơm phụt:

Khi tính toán tổng lực đẩy $F_{tổng}$ của xe, người ta thường nhân thêm hệ số này vào:

$$F_{tổng} = (F_{bơm\ trái} + F_{bơm\ phải}) \text{ x } \eta_d \quad (12)$$

η_d: Là hệ số làm việc đồng đều của bơm phụt.

Trong hệ thống thủy lực, công suất từ động cơ được truyền qua bơm thủy lực đến các mô-tơ thủy lực gắn trực tiếp vào trục bơm phụt $\eta_d = 0,98 \div 0,99$.

d) Xác định lực cản trên một bơm luồng phụt.

Lực cản trên một bơm luồng phụt được xác định theo công thức [3,4] :

$$R_T = \xi.\rho\frac{{v_s}^2}{2}F_p \quad (13)$$

Trong đó: R_T là lực cản của ống phụt; v_s: Tốc độ chảy của nước qua tiết diện thuỷ lực; ξ: là hệ số cản của ống phụt.

e) Xác định hệ số ảnh hưởng của vỏ và phần vận hành.

Hiệu suất của thiết bị:

$$K_{vh} = \eta_{cb}.\eta_v.\eta_{vh} \quad (14)$$

Trong đó: η_{cb}- Hiệu suất riêng của cụm bơm phụt; η_v - Hiệu suất vỏ; η_{vh} : Hiệu suất truyền động thủy lực.

Theo [4,5], η_{cb}= 0,4 - 0,55 ; η_v= 1,05 ; η_{vh}= 0,8. Thay vào ta được:

$$K_{vh} = 0,55.1,05.0,8 = 0,46$$

f) Xác định hệ số nén của dòng khối nước phụt.

Hệ số nén là tỉ số giữa diện tích mặt cắt ngang của ống hút $S_{vào}$ và diện tích vòi phun S_{ra} [3]:

$$\beta_{nén} = \frac{S_{vào}}{S_{ra}} \approx \frac{S_{bơm}}{S_{ra}} \quad (15)$$

Trong đó: $S_{bơm} = \frac{\pi.{D_{pk}}^2}{4}$; $S_{ra} = \frac{\pi.{D_{ra}}^2}{4}$; D_{pk} - Đường kính bánh công tác của bơm, $D_{pk} = 0,345$ (m); D_{ra} : Đường kính ống phụt ra, $D_{ra} = 0,24$ (m). Thay vào ta có:

$$\beta_{nén} \approx \frac{{D_{pk}}^2}{{D_{ra}}^2} = \frac{0,345^2}{0,240^2} = 2,06$$

g) Xác định tiết diện phần đường ống đặt tuốc bin, miệng ống phụt.

- Diện tích miệng ống phụt : $S_{ra} = \frac{\pi{D_{ra}}^2}{4}$= 0,0452 ($m^2$)

- Diện tích lỗ hút vào ở đáy xe được xác định theo công thức :

$$S_{vào} = (2,5 - 3,0)S_{ra}$$

Chọn $S_{vào} = 3.S_{ra} = 0,135(m^2)$

- Tiết diện phần đường ống đặt tuốc bin được xác định theo công thức:

$$F_p = \frac{\pi(D_{pk}+2\Delta)^2}{4} \quad (16)$$

Trong đó: Δ là khe hở giữa mút các cánh tuốc bin và thành trong của ống phụt (m); $\Delta \approx 0{,}01.D_{pk} = 0{,}01.0{,}345 = 0{,}00345$ (m) ; D_{pk} là đường kính tuốc bin (m);

Thay các thông số, ta được: F_p= 0,095 (m^2).

h) Xác định lưu lượng qua tiết diện tuốc bin, tốc độ quay lớn nhất của tuốc bin.

- Lưu lượng bơm xác định từ việc cân bằng lực đẩy và lực cản nước có tính đến hiện tương hút theo công thức:

$$P_B = \frac{R_S}{1-T_z} \text{ hoặc } K_\Sigma.\rho.Q_H(V_{ct}-V) = R\frac{R_\delta}{(1-T_z)} \quad (17)$$

Trong đó: Q_H- Lưu lượng của bơm (m^3/h); K_Σ- Hệ số kể đến ảnh hưởng của bơm và truyền động bơm ; $V_{ct} = \frac{Q_H}{S_{ra}}$ vận tốc trung bình của luồng phụt.

Giải phương trình tương đối của Q_H ta có :

$$Q_H = S_{ra}[(K_\Sigma.\rho.V) + \frac{\sqrt{(K_\Sigma.\rho.V)^2+4K_\Sigma.\rho\frac{R_S}{S_{ra}(1-T_Z)}}}{(2K_\Sigma.\rho)}] \quad (18)$$

Khi tính toán chọn: T_z = 0,05 ÷ 0,08; K_Σ = 1,2, thay vào ta được : Q_H = 0,634 (m^3/s).

- Để xác định tốc độ quay lớn nhất của tuốc bin n_{pk} theo điều kiện làm việc không có xâm thực (vận tốc cánh bơm không vượt quá 35 m/s) ta dựa vào công thức:

$$n_{pk} = \frac{11{,}4}{D_{pk}} = 33 \text{ (vg/s)} = 1980 \text{ (rpm)}$$

- Hệ số lưu lượng của bơm luồng phụt:

$$K_q = \frac{Q_H}{(n_{pk}.{D_{pk}}^3)} \quad (19)$$

i) Xác định hệ số cản toàn phần của ống phụt

Hệ số cản toàn phần của ống phụt ξ_T xác định theo công thức [4,6]:

$$\xi_T = \sum_1^i \xi_{TP}.{\delta_i}^2 + \sum_1^i \xi_M.{\delta_i}^2 \quad (20)$$

Trong đó: ξ_{TP}và ξ_M - Hệ số cản cục bộ và hệ số cản ma sát đoạn thứ i ; $\delta_i = \frac{S_{ra}}{S_i}$ - Tỉ số giữa tiết diện cửa ra và tiết diện ống đoạn thứ i. Các giá trị ξ_{TP} và ξ_M chọn theo bảng 1 [4].

Bảng 1. Bảng tra cứu các hệ số cản trên các đoạn đường ống bơm phụt.

TT	Đoạn ống thứ i	Hệ số ξ_{TP}	ξ_M	S_i	$\delta_i^2 = (S_{ra}/S_i)^2$
1	Cửa hút (lưới chắn rác)	0,15 – 0,35	0,02 – 0,05	0,114	0,157
2	Đoạn ống cong	0,15 – 0,25	0,02 – 0,05	0,095	0,226
3	Đoạn cánh hướng dòng	0,1 – 0,2	0,02 – 0,05	0,069	0,429
4	Loa phụt	0,05 – 0,1	0,02 – 0,05	0,045	1

Chọn hệ số cản cục bộ và hệ số cản ma sát ở các giá trị lớn nhất, ta tính được hệ số cản toàn phần của ống phụt là: ξ_T =0,387

k) Xác định áp lực cột áp của bơm luồng phụt.

Khi tính toán ta cần xét đến tổn hao đột thu ở cửa vào ống dẫn nước, tổn hao ma sát lưới lọc, tổn hao dọc đường do ma sát trên đường ống, tổn hao do sự thay đổi tiết diện ống theo chiều dài. Khi đó cột áp của bộ dẫn tiến xác định theo phương trình sau [4]:

$$H = (1+\xi_T)\frac{V_B}{2g} + h_c - \frac{\beta^2}{2g} \quad (21)$$

Trong đó: h_c- Độ cao của cửa phụt so với đường ống mớn nước; β- Hệ số tổn hao dòng chảy tại cửa ra vào.

Đối với các bơm phụt có lỗ ra đặt chìm dưới nước có thể xem $h_c = 0$ và β=0, khi đó:

$$H = (1 + \xi_T)\frac{8{Q_H}^2}{\pi^2.g.{d_B}^4} \quad (22)$$

Trong đó: d_B - Đường kính lỗ ra của ống dẫn nước, $d_{B=}$ 0,24 (m).

Thay các giá trị vào ta xác định được:

$$H = (1 + 0{,}387)\frac{8.0{,}634^2}{3{,}14^2.9{,}81.0{,}240^4} = 13{,}8\ (m)$$

l) Xác định công suất cần thiết đặt vào tuốc bin.

- Công suất có ích cần thiết trên trục bơm:

$$N_p = \frac{\gamma.Q_H.H}{\eta_p} \quad (23)$$

Với η_p là hiệu suất luồng bơm phụt $(\eta_p = 0{,}7 - 0{,}8)$, chọn $\eta_p = 0{,}7$

$$N_p = \frac{9{,}81.1000.0{,}634.13{,}8}{0{,}7} = 122613\ (w) = 122{,}613\ (kW)$$

- Công suất động cơ cần thiết đặt vào trục bơm:

$$N_{đc} = \frac{2.N_p}{\eta_{TL}}.K \quad (24)$$

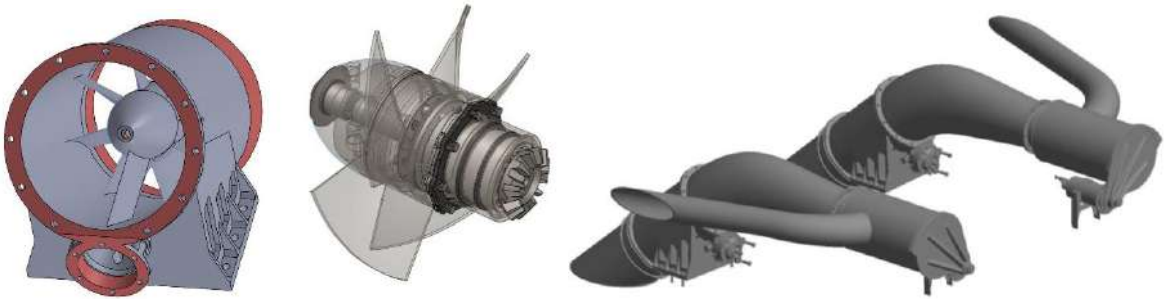
Với K- hệ số dự trữ động cơ (K = 1,1 – 1,15). Chọn K=1,1 ta xác định được:

$N_{đc}$= 337,18 (kW)= 451,8 (Hp)

Nhận xét: Với công suất động cơ xe T-1 lớn nhất là 470 (Hp) tại n_e= 2100 (rpm) và việc trích công suất từ vị trí ngay sau ly hợp chính nên hoàn toàn có thể đảm bảo tốt yêu cầu tạo ra công suất cần thiết cho các trục bơm. Do vậy các thông số làm việc của cơ cấu bơi kiểu bơm luồng phụt đã xác định đã đáp ứng tốt yêu cầu kỹ thuật đề ra đối với hệ thống bơi xe tăng T-1.

3.2. Thiết kế 3D, kiểm bền các chi tiết chính của cơ cấu bơi

Từ các thông số kết cấu của cơ cấu bơi xe tăng T-1 đã xác định, sử dụng phần mềm Soliworks nhóm tác giả đã xây dựng được bản vẽ các cụm chi tiết của cơ cấu bơi (Hình 5).



Hình 5. Các cụm chi tiết chính của cơ cấu bơi xe tăng T-1

Trong quá trình làm việc của cơ cấu bơi, chi tiết cánh tuốc bin là chi tiết chịu tải nặng nền nhất, do vậy cần kiểm tra độ bền của chi tiết này. Cánh tuốc bin được chế tạo bằng phương pháp đúc, vật liệu là SUS304. Tuốc bin được dẫn động bằng mô tơ thủy lực thông qua bộ truyền cơ khí với bộ truyền bánh răng côn. Khi quay bánh tuốc bin cuốn guồng nước từ miệng hút đẩy sang miệng phụt. Áp lực do nước tác dụng trên bánh tuốc bin (khi quay đều) cân bằng với mô men truyền từ mô tơ thủy lực truyền tới bánh tuốc bin. Do vậy, ta coi rằng tải trọng chủ yếu tác dụng lên bánh tuốc bin khi tính bền chính là mô men từ bơm thủy lực truyền tới.

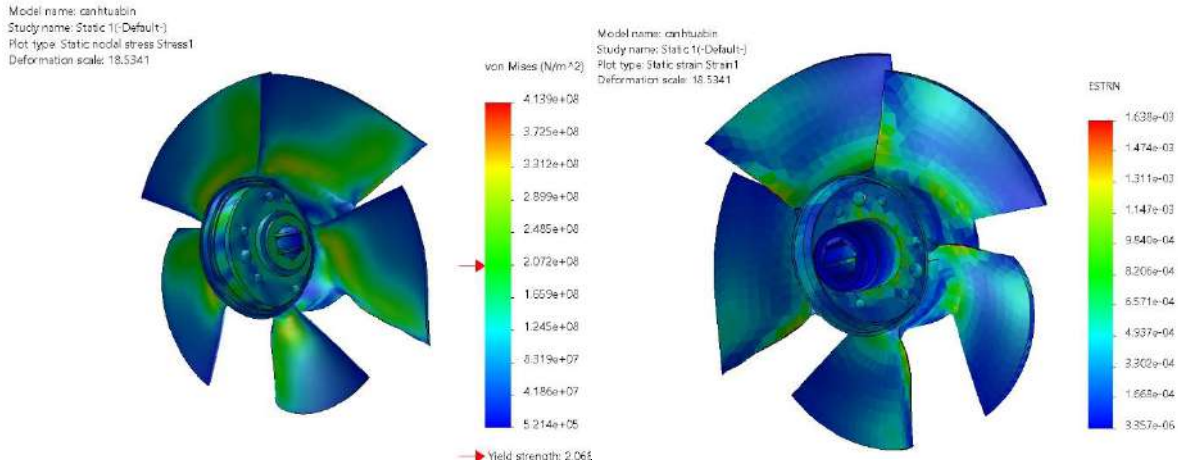
Công suất lớn nhất đặt vào trục bánh tuốc bin là:

$$N_{tb} = N_p \cdot \eta_d \quad (25)$$

Với N_p là công suất lớn nhất đặt vào trục bơm, N_p=122613 (W) ; η_d - là hiệu suất của bộ truyền cơ khí, η_d =0,98. Thay vào ta có N_{tb}= 120161 (W).

Tốc độ quay của bánh tuốc bin khi ổn định là n_{pk}= 1980 (vg/phut), từ đây xác định được mô men lớn nhất tác dụng lên tuốc bin là: M_{tb}= 580 (N.m).

Sử dụng phương pháp phần tử hữu hạn, bằng phần mềm Ansys [7], tiến hành tính toán kiểm bền cho cánh tuốc bin nhận được kết quả ứng suất, biến dạng của cánh như hình 6.



Hình 6. Kết quả kiểm bền cánh tuốc bin

Nhận xét: Ứng suất tương đương lớn nhất của cánh tuốc bin là σ_{max} = 413,9 (Mpa) nhỏ hơn ứng suất cho phép của của vật liệu SUS304 là [σ]= 824 Mpa [8], hơn nữa biến dạng lớn nhất của cánh rất nhỏ ε_{max} = 1,638.10^{-3}, như vậy cánh tuốc bin đảm bảo độ bền tĩnh. Cánh tuốc bin là chi tiết chịu tải nặng nề nhất trong kết cấu, do vậy có thể khẳng định tất cả các chi tiết trong cơ cấu bơi đều đảm bảo điều kiện bền.

4. Kết luận

Xác định các thông số làm việc của cơ cấu bơi là bước quan trọng trong quá trình tính toán thiết kế cơ cấu bơi dạng bơm phụt cho xe tăng thiết giáp. Trên cơ sở cho trước nguồn động lực, các thông số về xe như kích thước, tải trọng và biên dạng vỏ, yêu cầu về vận tốc bơi lớn nhất, vị trí bố trí bơm, vận tốc dòng nước... nhóm tác giả đã tiến hành các bước tính toán các thông số làm việc của bơm phụt. Kết quả nghiên cứu cho thấy, cơ cấu bơi kiểu bơm luồng phụt này đã đáp ứng tốt yêu cầu kỹ thuật đề ra đối với xe tăng bơi T-1 và trong quá trình làm việc, toàn bộ các chi tiết của cơ cấu bơi đều đảm bảo độ bền tĩnh.

Tài liệu tham khảo

1. Võ Văn Trung; *Cấu tạo xe tăng tập 3: xe tăng hạng nhẹ PT-76*, HVKTQS- 2003.

2. Nguyễn Văn Luận, Lê Kỳ Nam, *Kết cấu tính toán xe tăng phần 1*, HVKTQS- 1999

3. Đinh Sơn Hùng, *Phương tiện vượt sông*, NXB Học viện KTQS- 1998.

4. Nguyễn Viết Tân, *Xe máy vượt sông*, NXB Quân đội nhân dân- 2011.

5. Duan, L., Yao, K., Pan, X., Hou, Z., Zhao, *Study on force characteristics and resistance for water by amphibious vehicle,* International Conference on Mechanical Engineering and Intelligent Systems, Yinchuan, China 2015.

6. A. Alexandrou, *Principles of fluid mechanics, 4th ed*. Prentice Hall, New Jersey-2001.

7. Nguyễn Việt Hùng, Nguyễn Trọng Giảng, *Ansys và mô phỏng số trong công nghiệp bằng phần tử hữu hạn*, NXB Khoa học và Kỹ thuật-2023.
8. Trần Văn Địch, Ngô Trí Phúc, *Sổ tay thép thế giới, Lần 2,* NXB Khoa học và Kỹ thuật, Hà Nội, 2006

Calculating the operating parameters of the jet-pumping swimming mechanism for light tanks manufactured in Vietnam

Abstract:

At present, the Vietnamese Army possesses a relatively large number of older generation amphibious tanks supplied by Russia and China. The amphibious mechanisms and propulsion systems on these vehicles, which are mechanically based, have become considerably outdated, and their performance characteristics no longer meet the requirements of training and combat missions under current conditions. Therefore, our Armed Forces are actively implementing a research, design, and manufacturing program for a new, modern amphibious light tank. One of the important features for this tank is to achieve stable waterborne operation with a maximum forward swimming speed of not less than 11 km/h and a reverse swimming speed of not less than 3 km/h. In this paper, the authors present the theoretical basis, methods for determining operating parameters, and the design calculations for a jet-pumping amphibious mechanism for the Vietnamese-produced light tank to satisfy the specified technical performance criteria.

Keywords: *Amphibious tank, amphibious mechanisms, Pump-jet propulsion, Swimming speed.*

Nghiên cứu tính toán, lựa chọn động cơ và bố trí chung khoang động lực xe thiết giáp chở quân 8x8 sản xuất tại Việt Nam

Hoàng Văn Nguyên[1*], Tô Viết Thành[1]

[1]*Học viện Kỹ thuật Quân sự*

Tóm tắt:

Trong bối cảnh hiện nay, nhiệm vụ nâng cấp, hiện đại hóa và từng bước làm chủ công nghệ sản xuất, chế tạo xe thiết giáp là yêu cầu khách quan đối với lực lượng Tăng thiết giáp. Thực hiện định hướng phát triển của Quân đội, việc nghiên cứu, thiết kế và chế tạo các thế hệ xe thiết giáp mới nhằm thay thế các dòng xe đã qua nhiều năm khai thác đang được triển khai một cách bài bản, phù hợp với điều kiện thực tiễn. Trong số đó, xe thiết giáp chở quân cấu hình 8x8 được xác định là một trong những sản phẩm trọng điểm, đáp ứng yêu cầu cơ động cao và tác chiến đa nhiệm. Công tác tính toán và lựa chọn động cơ và bố trí chung khoang động lực giữ vai trò quan trọng, ảnh hưởng trực tiếp đến các chỉ tiêu, tính năng chiến – kỹ thuật như khả năng cơ động, tính ổn định, độ bền và hiệu quả khai thác sử dụng. Bài viết này tập trung trình bày những cơ sở lý thuyết phục vụ cho việc nghiên cứu, phân tích và tính toán lựa chọn động cơ và bố trí chung khoang động lực cho xe thiết giáp chở quân 8x8 được chế tạo trong nước.

Từ khóa: *Xe thiết giáp chở quân 8x8; động cơ, khoang động lực.*

Abstract:

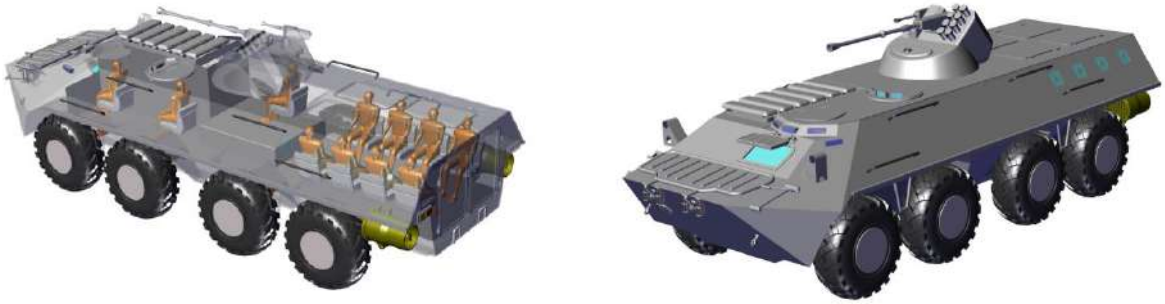
In the current context, the requirement to enhance equipment quality through modernization, upgrading, and gradually mastering the technology of designing and manufacturing armored vehicles has become an inevitable task for the Armored Forces. In line with the development orientation of the Army, the research and development of new-generation armored vehicles to replace long-serving and outdated platforms are being implemented systematically, ensuring suitability to practical conditions and mission requirements. Among these, the 8×8 armored personnel carrier configuration has been identified as a key development direction due to its superior mobility, operational flexibility, and multi-role combat effectiveness. In the vehicle design process, the calculation and selection of the engine as well as the determination of the overall layout of the power compartment play a decisive role, directly affecting tactical–technical performance indicators such as specific power, obstacle-crossing capability, motion stability, durability, and operational efficiency. On this basis, this paper focuses on presenting the theoretical foundations for research, analysis, and calculation in selecting the engine and determining the general arrangement of the power compartment for a domestically manufactured 8×8 armored personnel carrier.

Keywords: 8×8 armored personnel carrier; engine; power compartment.

1. Đặt vấn đề

Bên cạnh định hướng phát triển mới, trong giai đoạn hiện nay việc cải tạo, nâng cấp các xe Tăng - thiết giáp đang có trong biên chế cũng là giải pháp phù hợp với điều kiện công nghệ và nguồn lực trong nước. Nhiều phương tiện vẫn bảo đảm tốt về kết cấu khung gầm và vỏ giáp, do đó việc tận dụng các cụm kết cấu sẵn có sẽ giúp tiết kiệm chi phí, rút ngắn thời gian triển khai, đồng thời tập trung nguồn lực cho việc nâng cao các tính năng quan trọng như khả năng cơ động, mức độ bảo vệ và hiệu quả hỏa lực. Tuy nhiên, về lâu dài, việc nghiên cứu, chế tạo một dòng xe thiết giáp mới có tính đa nhiệm cao, tích hợp các công nghệ hiện đại và đáp ứng yêu cầu tác chiến trong điều kiện chiến tranh công nghệ cao là hướng phát triển tất yếu.

Đối với xe thiết giáp bánh lốp 4x4, nước ta đã từng bước làm chủ công nghệ thiết kế, chế tạo và lắp ráp. Trên cơ sở đó, nhiệm vụ đặt ra trong thời gian tới là tập trung nghiên cứu, thiết kế và chế tạo xe thiết giáp chở quân chiến đấu 8x8. Xe thiết giáp chở quân chiến đấu 8x8 có nhiều ưu điểm nổi bật như khả năng bảo vệ, khả năng cơ động trên các địa hình phức tạp, tính năng thông qua cao, tải trọng chuyên chở lớn, có thể thực hiện nhiều nhiệm vụ đặc thù. Để hoàn thành nhiệm vụ chế tạo dòng xe này cần triển khai rất nhiều nội dung công việc, trong đó nhiệm vụ trước hết là nghiên cứu tính toán, lựa chọn các cụm hệ thống động truyền lực cho xe để đáp ứng yêu cầu tính năng kỹ thuật đề ra. Một số chỉ tiêu về kỹ thuật cơ bản của xe được cho trong bảng 1.



Hình 1. Phương án thiết kế mới dòng xe chiến đấu chở quân 8x8 tại Việt Nam

Bảng 1. Chỉ tiêu kỹ thuật cơ bản

TT	Tên chỉ tiêu	Đơn vị	Giá trị
1	Công thức bánh xe		8x8
2	Khối lượng xe chiến đấu	Tấn	≤16,0
3	Khối lượng xe không tải	Tấn	≤15,0
4	Chiều dài tổng thể	mm	7.580±20
5	Chiều rộng tổng thể	mm	2.985
6	Chiều cao đến nóc xe	mm	2.725±20
7	Khoảng sáng gầm xe khi đủ tải	mm	≥475
8	Góc vượt trước	Độ	≥40
9	Góc thoát sau	Độ	≥30
10	Bán kính quay vòng nhỏ nhất của xe	m	13,2
11	Bán kính quay vòng dưới nước khi tốc độ lớn nhất	m	6÷10
12	Vận tốc lớn nhất trên đường nhựa	km/h	80
13	Vận tốc lớn nhất khi bơi tiến	km/h	9÷10
14	Vận tốc lớn nhất khi bơi lùi	km/h	3÷5
15	Dự trữ hành trình trên đường nhựa	km	≥700
16	Góc vượt dốc đứng lớn nhất	Độ	30
17	Góc vượt sường nghiêng lớn nhất	Độ	25

2. Cơ sở lý thuyết tính toán, lựa chọn động cơ và hệ thống truyền lực

2.1. Cơ sở lý thuyết tính toán, lựa chọn động cơ

Để đảm bảo các tính năng kỹ thuật đề ra theo Bảng 1, cần phải tính toán xác định công suất lớn nhất của động cơ. Trong thiết kế xe mới, người ta thường cho sẵn vận tốc lớn nhất v_{max}.

Vận tốc này ta coi là đạt được khi xe chạy trên đường tốt, nghĩa là hệ số cản nhỏ, độ dốc không cao. Công suất lớn nhất của động cơ được xác định từ điều kiện đảm bảo vận tốc chuyển động lớn nhất, theo biểu thức sau [3,4]:

$$N_e^{'} = \frac{P_k.v}{\eta_\Sigma} \quad (1)$$

Trong đó:

η_Σ - Hiệu suất của xe có kể đến tổn thất trong thiết bị động lực; $\eta_{\Sigma=}$ 0,80 ÷ 0,85.

$N_e^{'}$ - Công suất động cơ [w]

v - Vận tốc của xe [m/s]

P_k - Lực kéo theo động cơ [N].

Khi xe chuyển động đều (v = const), lực kéo theo động cơ P_k được xác định theo công thức sau [4]:

$$P_k = \psi.G + P_w \quad (2)$$

Trong đó:

G - Trọng lượng của xe; G= 160000 [N]

ψ - Hệ số cản tổng cộng của đường;

P_w - Lực cản không khí. [N]

Đối với xe không kéo mooc chuyển động đều với vận tốc $v = v_{max}$, công suất của động cơ được xác định như sau [4]:

$$N_e^{'} = N_{emax}^{'} = \frac{(\psi_{min}.G + P_w) \, . \, v_{max}}{\eta_\Sigma} \quad (3)$$

Từ công thức trên ta thấy rằng công suất $N_{emax}^{'}$ của động cơ phụ thuộc vào việc chọn hệ số cản tổng cộng của đường và hiệu suất của xe. Ở xe không kéo mooc, hệ số cản ψ thường được lấy theo điều kiện đường tốt (f = 0,015 ÷ 0,025) và có góc dốc nhỏ (i = 0,01 ÷ 0,015), nghĩa là hệ số ψ được chọn trong khoảng ψ = 0,025 ÷ 0,040 [5].

Hiệu suất của xe có tính đến tổn thất trong thiết bị động lực cho hệ thống truyền lực có cấp có thể lấy gần đúng η_Σ = 0,80 ÷ 0,85, ở đây chọn η_Σ = 0,82.

Như đã phân tích ở trên, ta thấy rằng các thông số để xác định $N_{emax}^{'}$ được lựa chọn trong khoảng rộng, do vậy giá trị công suất lớn nhất $N_{emax}^{'}$ có thể xác định theo hệ số cản quy dẫn ψ_0 theo công thức:

$$\psi_0 = (\psi_{min} + \frac{P_w}{G}) \, . \, \frac{1}{\eta_\Sigma} \quad (4)$$

Khi đó:

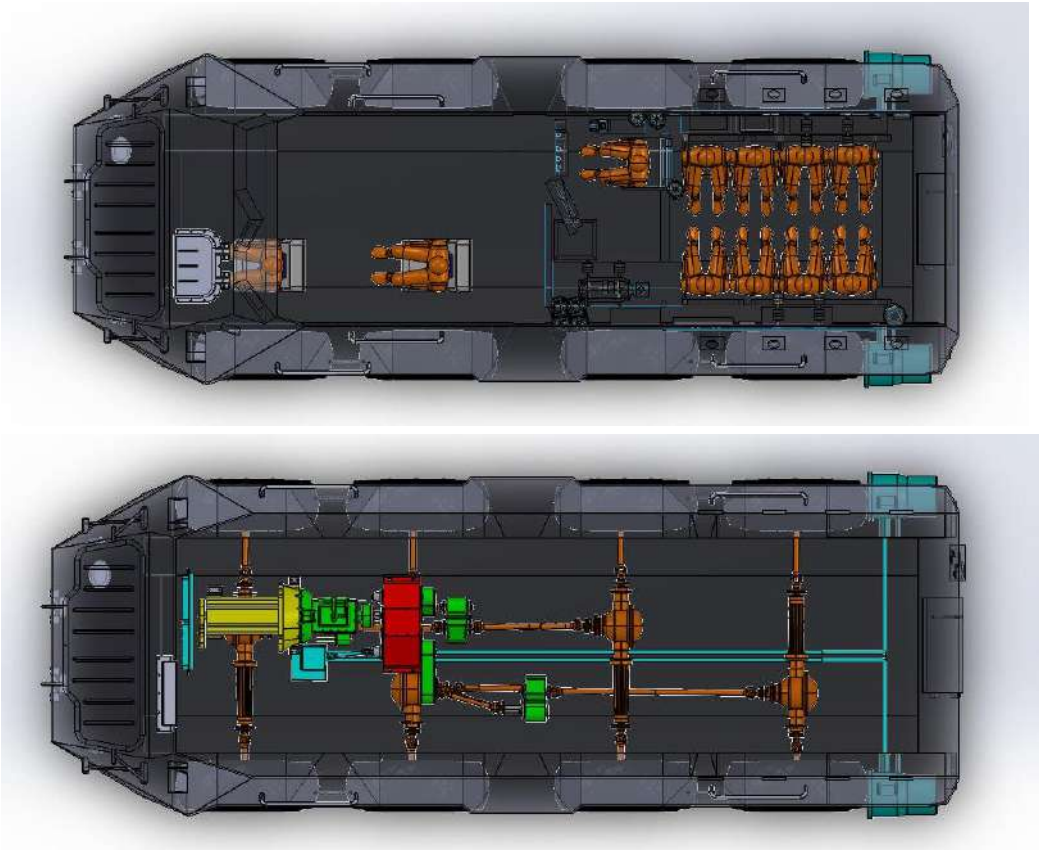
$$N_{emax}^{'} = \psi_0 \, . \, G \, . \, v_{max} \quad (5)$$

Hệ số cản quy dẫn ψ_0 đối với xe thiết giáp có giá trị trong khoảng ψ_0 = 0,04 ÷ 0,08 [5]. Ở đây ta lấy ψ_0 = 0,06

2.2. Cơ sở lý thuyết tính toán, lựa chọn các cụm trong hệ thống truyền lực

Hệ thống truyền lực (HTTL) xe thiết giáp chở quân 8x8 bao gồm: ly hợp, hộp số, hộp phân phối trung tâm, cầu chủ động, bộ truyền các đăng, hộp giảm tốc bánh xe theo hình 2. Trong giới hạn bài báo, tác giả tập trung vào tính toán, lựa chọn động cơ đây là nhiệm vụ quan

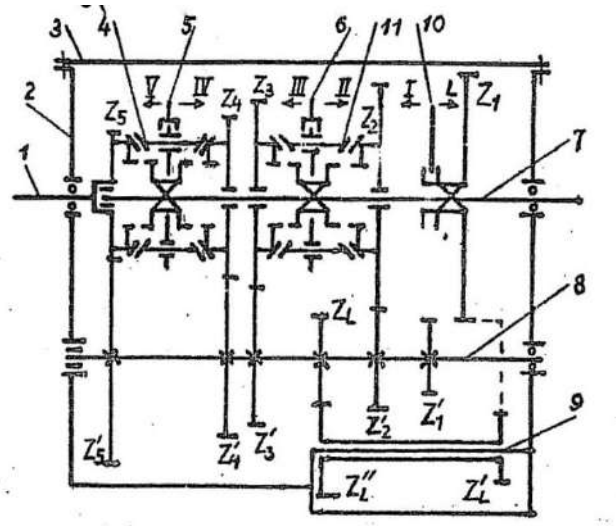
trọng nhất khi tính toán thiết kế bố trí chung của xe, . Các cụm còn lại như hộp số, hộp số phân phối, cầu chủ động, bộ truyền các đăng, hộp giảm tốc bánh xe được lựa chọn dựa trên việc tham khảo các mẫu xe thiết giáp lốp có sẵn.



Hình 2. Phương án bố trí chung hệ thống động truyền lực của xe

Hộp số dùng để thay đổi tỷ số truyền của HTTL trong quá trình chuyển động của xe nhằm thay đổi lực kéo ở các bánh xe chủ động và thay đổi vận tốc chuyển động của xe trong khoảng rộng tuỳ theo sức cản bên ngoài, ngoài ra hộp số còn giúp cho xe có thể chuyển động lùi được và có thể tách động cơ đang làm việc ra khỏi hệ thống truyền lực trong khoảng thời gian tuỳ ý. Hộp số được tính toán, lựa chọn phải đáp ứng được các yêu cầu: Bảo đảm cho xe có chất lượng kéo và chất lượng kinh tế tối ưu trong những điều kiện sử dụng cho trước, điều khiển thuận tiện và nhẹ nhàng, làm việc êm và không ồn, có hiệu suất truyền lực cao [6,7].

Thông qua việc phân tích, đánh giá ưu nhược điểm của các loại hộp số thông dụng và tham khảo kết cấu hộp số trên một số xe thiết giáp 8x8, chúng tôi lựa chọn loại hộp số cho xe là hộp số cơ khí 3 trục dọc, 5 cấp để nâng cao hiệu quả sử dụng công suất của động cơ. Ngoài ra để nâng cao chất lượng động lực học, khả năng tăng tốc và tính kinh tế của xe, chọn thiết kế số 5 là số truyền thẳng.



Hình 3. Sơ đồ động học hộp số ba trục dọc 5 cấp xe thiết giáp chở quân 8x8

1. Trục sơ cấp; 2. Vỏ hộp số; 3. Nắp hộp số; 4,11. Đồng tốc; 5,6,10. Càng gài số; 7. Trục thứ cấp; 8. Trục trung gian; 9. Trục số lùi.

2.2.1. Xác định khoảng động học và khoảng lực học của xe.

Khoảng động học của xe sẽ được xác định theo công thức sau đây [4]:

$$\mathrm{d}_k = \frac{v_{t\max}}{v_{t\min}} \tag{6}$$

Trong đó: d_k - khoảng động học của xe; $v_{t\max}$ - tốc độ tính toán lớn nhất của xe, $v_{t\max}$ = 80 km/h; $v_{t\min}$ - tốc độ tính toán nhỏ nhất của xe ứng với trường hợp động cơ ở chế độ công suất cực đại.

Để có thể xác định được d_k trước tiên phải xác định khoảng lực học. Khoảng lực học của xe được xác định theo công thức sau [4,6]:

$$\mathrm{d}_1 = \frac{\varphi}{\psi_0 . \eta_\Sigma} \frac{G\varphi}{G} \tag{7}$$

d_l - khoảng lực học của xe; Khoảng lực học của ô tô có tất cả các cầu đều là chủ động có giá trị nằm trong khoảng d_1=14-15; chọn $d_l = d_k$ = 15.

G- trọng lượng toàn bộ xe;

G_φ - trọng lượng bám của xe;

Hệ số bám φ lấy giá trị lớn ($\varphi = 0{,}7 - 0{,}8$) để đảm bảo ô tô có thể sử dụng triệt để lực bám khi chuyển động trên đường tốt có độ dốc $\alpha = 30^0$. chọn $\varphi = 0{,}75$.

2.2.2. Xác định tỉ số truyền lớn nhất và nhỏ nhất của hệ thống truyền lực.

Tỉ số truyền lớn nhất xác định theo công thức sau đây:

$$i_{tlmax} = \frac{0{,}377 n_N r_K}{v_{t\min}} \tag{8}$$

Trong đó:

n_N - số vòng quay của động cơ ứng với trường hợp có công suất lớn nhất: 2200 (v/ph).

r_K -bán kính tính toán của bánh xe chủ động:

+ Bán kính thiết kế: r =($\frac{d}{2}$+B)= ($\frac{1.15}{2}$+0.4)=0.975 (m)

+ Bán kính tính toán: $r_K=\lambda_1.r=0.95.0.975= 0.926$ (m)

Với: d- Đường kính vành bánh xe (m); B: Chiều rộng lốp (m); λ_1 : Hệ số biến dạng của lốp 0,95.

Đối với ô tô quân sự số truyền cao nhất của hộp số thường là số truyền thẳng với tỉ số truyền $i_{hn}=1$ và số truyền cao nhất của hộp số phân phối thường có tỉ số truyền $i_{Pc}=1$.Do đó tỉ số truyền nhỏ nhất của hệ thống truyền lực cũng chính là tỉ số truyền của cầu chủ động :

$$i_0{}'=i_{t1_{min}}=\frac{0,377\ n_N r_K}{V_{tmax}} \quad (9)$$

Trong đó :

$i_0{}'$ – Tỉ số truyền của cầu chủ động.Tỉ số truyền này bao gồm :tỉ số truyền lực chính i_0 và tỉ số truyền lực cạnh i_c .

$i_0{}' = i_o \,.\, i_c$

Đối với hộp số có số truyền cao nhất là số truyền tăng i_{hmin} sẽ được xác định theo công thức sau đây :

$$i_{hmin}=\frac{n_N}{n_{max}} \quad (10)$$

Trong đó :

i_{hmin} – Tỉ số truyền của số truyền tăng trong hộp số.

n_{max} – Số vòng quay lớn nhất của trục thứ cấp hộp số.Để hạn chế tốc độ vòng tại vị trí ăn khớp của cặp bánh răng giá trị n_{max} được chọn phải nhỏ hơn 4000-4500 v/ph .

Khi hộp số có tỉ số truyền tăng ,tỉ số truyền của cầu chủ động sẽ được xác định theo công thức :

$$i_0{}'=\frac{0,377\ .n_N\ r_k}{i_{hmin}.V_{tmax}} \quad (11)$$

Hộp số phân phối của ô tô quân sự thường có hai số truyền : số truyền cao và số truyền thấp . Tỉ số truyền của số truyền cao thường chọn $i_{Pc}=1$. Tỉ số truyền của số truyền thấp trong các hộp số phân phối hiện có i_{Pt} = 1,5-2,5 .

Xác định tỉ số truyền của các tay số trong hộp số :

Tỉ số truyền của số một được xác định theo công thức :

$$i_{h1} = \frac{i_{t1max}}{i_{Pt}\ i_0} \quad (12)$$

Thông thường số truyền cao nhất của hộp số là số truyền thẳng với $i_{pc}=1$. Do đó khoảng động học làm việc thường xuyên của ô tô có thể xác định theo công thức sau:

$$d'_k=\frac{i_{h1}}{i_{hn}}=1 \quad (13)$$

Các tỉ số truyền của hộp số ô tô quân sự thông thường được phân chia theo cấp số nhân và có hiệu chỉnh ở các số truyền cao để đảm bảo khoảng cách giữa các tốc độ lớn nhất ở các số truyền đó thu hẹp lại .

Khi các tỉ số truyền được phân chia theo cấp số nhân thì công bội của cấp số nhân đó sẽ là :

$$q=\sqrt[n-1]{d_k}=\sqrt[n-1]{i_{h1}} \quad (14)$$

Trong đó :

q – Công bội của cấp số nhân .

n– Số lượng các số truyền của hộp số .

Đối với ô tô không có hộp số phụ hoặc hộp số phân phối thì khoảng động học làm việc thường xuyên của ô tô :

$$d'_k = d_k = \frac{V_{tmax}}{V_{tmin}} \quad (15)$$

Và công bội của cấp số nhân sẽ bằng :

$$q=\sqrt[n-1]{d_k} \quad (16)$$

Tỉ số truyền của các số truyền trung gian xác định theo công thức sau :

$$i_{h2}=i_{h1}.q^{-1}$$
$$i_{h3}=i_{h2}.q^{-1}=i_{h1}.q^{-2}$$
$$\cdots\cdots\cdots$$
$$i_{hi}=i_{h(i-1)}.q^{-1}=i_{h1}.q^{-(i-1)}$$

Trong đó:

$i_{h(i-1)}$;i_{hi}-tỉ số truyền của số truyền trung gian thứ (i-1)và thứ i

Để đảm bảo khoảng cách giữa các tốc độ lớn nhất ở các số truyền cao thu hẹp lại nên tiến hành hiệu chỉnh các tỉ số truyền trung gian sao cho:

Đối với hộp số có số truyền tăng và trước số truyền tăng là số truyền thẳng,thì từ số một đến số truyền thẳng tỉ số truyền được phân chia theo cấp số nhân.công bội của cấp số nhân này sẽ là

$$q = \sqrt[n-2]{d_k}$$

Tỉ số truyền của các số truyền trung gian(từ số một đến số truyền thẳng)cũng xác định theo công thức như đối với hộp số không có số truyền tăng.

Tỉ số truyền của số lùi được chọn sao cho đảm bảo tốc độ thấp nhất của ô tô và đảm bảo khắc phục được lực cản lớn nhất của đường.

3. Cơ sở tính toán và lựa chọn động cơ

3.1. Kết quả tính toán lựa chọn động cơ

Thay các số liệu vào công thức (5) ta xác định được công suất lớn nhất của động cơ:

$N'_{emax} = 0.06.16.10^4 .22,2= 213.12$ (kw)= 289,76 (hp).

Do xe thiết giáp chở quân 8x8 đòi hỏi khả năng tác chiến độc lập trên nhiều địa hình, sử dụng nhiều cầu chủ động và tích hợp nhiều tính năng được trích từ hộp phân phối trung tâm nên ta chọn động cơ có công suất lớn nhất nằm trong khoảng 320 hp đến 350 hp (lớn hơn khoảng 15% so với tính toán lý thuyết). Qua nghiên cứu về chất lượng cũng như tính phổ biến của các dòng động cơ trên thị trường cho xe quân sự hạng nhẹ, nhận thấy động cơ KAMAZ 689.311-360 do Nga sản xuất là tương đối phù hợp có thể đáp ứng được tốt yêu cầu về mặt công suất và kích thước lắp đặt vào khoang động truyền lực xe thiết giáp chở quân 8x8. Đây là loại động cơ Diesel 4 kỳ, 8 xi lanh bố trí hình chữ V(V8), tăng áp với những tính năng:

Bảng 2. Chỉ tiêu kỹ thuật cơ bản của động cơ KAMAZ 689.311-360

TT	Thông số	Đơn vị	Loại, giá trị
1	Tên động cơ		KAMAZ 689.311-360
2	Kiểu động cơ		Diesel 4 kỳ, 8 xi lanh bố trí chữ V, tăng áp
3	Kích thước (dài x rộng x cao)	mm	1070x720x840
4	Công suất lớn nhất	KW(ml)	258(351)
5	Mô men xoắn lớn nhất	N.m	1400-1600
6	Tốc độ vòng quay ở chế độ công suất lớn nhất (n_{emax})	vg/phút	2100 ± 25
7	Tốc độ vòng quay nhỏ nhất	vg/phút	600-800
8	Hành trình làm việc của piston	mm	130
9	Đường kính xi lanh	mm	120
10	Khối lượng động cơ	Kg	706
11	Hệ thống làm mát		Bằng nước, tuần hoàn cưỡng bức

3.2. Kết quả tính toán lựa chọn hộp số

Thay các số liệu tính toán vào (8) ta được: $i_{tlmax} = 144,1$

Thay các số liệu tính toán vào (9) ta được: $i_{tlmin} = 9,6$

Tỉ số truyền của cầu chủ động: $i_0 = i_{tlmin} = 9,6$

- Xác định tỉ số truyền của các tay số trong hộp số.

Tỉ số truyền của tay số I được xác định theo công thức (12):

$$i_{h1} = \frac{i_{tl\max}}{i_{pt}.i_0} = \frac{144,1}{1,95.9,6} = 7,7$$

Các tỉ số truyền của hộp số xe thiết giáp bánh lốp thông thường được phân chia theo cấp số nhân theo điều kiện sử dụng tối đa công suất của động cơ. Khi các tỉ số truyền được phân chia theo cấp số nhân thì công bội sẽ là:

$$q = \sqrt[n-1]{d_{hs}} = \sqrt[n-1]{i_{h1}} = \sqrt[4]{7,7} = 1,666$$

Với: n- Số lượng số truyền của hộp số (n=5).

Tỉ số truyền của các số truyền trung gian xác định theo công thức sau:

$i_{h2}=i_{h1}.q^{-1}= 7,7.1.666^{-1}=4,61$; $i_{h3}=i_{h2}.q^{-1}=i_{h1}.q^{-2}=7,7.1,666^{-2}=2,77$

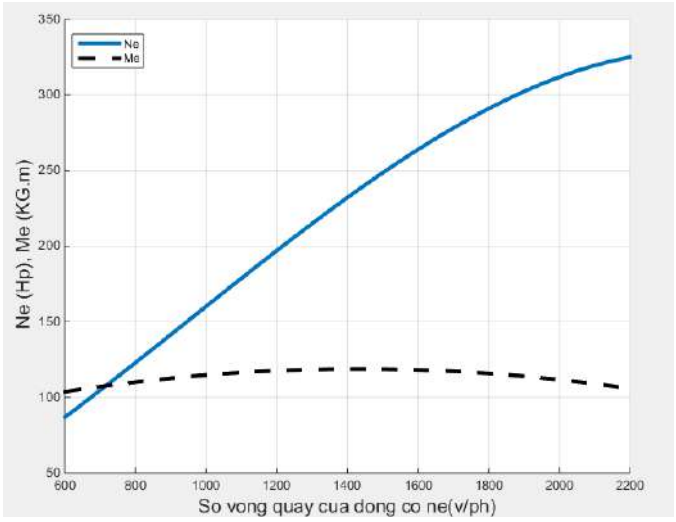
$i_{h4}=i_{h3}.q^{-1}=i_{h1}.q^{-3}=7,7.1,666^{-3}=1,66$; $i_{h5}=1$

trong đó: i_{hi} - tỉ số truyền của số truyền trung gian thứ i

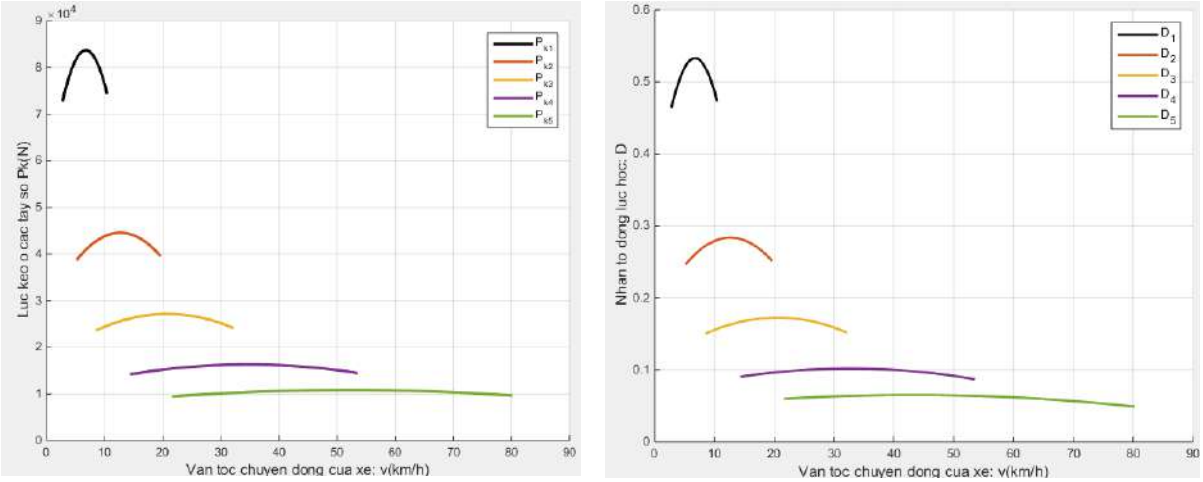
Sau khi hiệu chỉnh ở các số truyền cao để đảm bảo khoảng cách giữa các tốc độ lớn nhất ở các số truyền đó thu hẹp lại theo điều kiện tính đến xác suất lực cản gặp phải. Nhóm tác giả đề xuất lựa chọn hộp số chính với các thông số: $i_{h1}= 7.7$; $i_{h2}=4.1$; $i_{h3}=2.5$; $i_{h4}= 1,5$; $i_{h5}=1$.

3.3. Xây dựng đồ thị đặc tính kéo chuyển động thẳng

Với các thông số của động cơ và hộp số như đã tính toán thiết kế, lựa chọn, sử dụng các công thức tính toán trong [3,5,8] và phần mềm Matlab nhóm tác giả xây dựng được các đồ thị đặc tính kéo chuyển động thẳng xe thiết giáp chở quân 8x8 (Hình 3,4) nhằm đánh giá chất lượng kéo chuyển động thẳng của xe.//



Hình 4. Đồ thị đặc tính ngoài của động cơ KAMAZ 689.311-360



Hình 5. Đồ thị đặc tính kéo chuyển động thẳng xe thiết giáp chở quân 8x8

Nhận xét đánh giá:

- Tốc độ lớn nhất của xe đạt được trên đường nhựa là 80 (km/h).

- Lực kéo lớn nhất xe đạt được khi xe chuyển động ở tay số 1 là P_k= 83669 (N).

- Nhân tố động lực học lớn nhất của xe ở tay số 1 là D_{max}= 0.533 từ đây ta có thể xác định được góc vượt dốc lớn nhất mà xe vượt được theo công thức [4]:

$$\alpha_{max} = \arcsin\left(D_{max} - f\sqrt{1 - D_{max}^2}\right) = 31{,}7 \text{ (độ)}$$

Như vậy xe thiết giáp chở quân 8x8 với các thông số của động cơ và hộp số như đã tính toán thiết kế, lựa chọn đáp ứng tốt các yêu cầu về vận tốc lớn nhất, lực kéo lớn nhất và góc vượt dốc lớn nhất cần đạt được.

4. Kết luận

Dựa trên yêu cầu tính năng chiến kỹ thuật đề ra đối với xe thiết giáp chở quân 8x8, trên cơ sở lý thuyết tính toán xe quân sự, bài báo đã trình bày các bước tính toán, lựa chọn động cơ và bố trí chung khoang động lực cho xe thiết giáp chở quân 8x8 chế tạo tại Việt Nam. Trong đó tập trung vào tính toán, lựa chọn động cơ động cơ, xây dựng đồ thị đặc tính kéo chuyển động thẳng. Kết quả tính toán kéo cho thấy chất lượng kéo chuyển động thẳng của xe thiết giáp chở quân 8x8 lắp động cơ KAMAZ 689.311-360 và hộp số tính toán thiết kế đã đáp ứng tốt các yêu cầu kỹ thuật với v_{max}= 80 (km/h), P_{kmax}= 83669 (N), D_{max}= 0.533, α_{max}= 31,7 (độ).

Tài liệu tham khảo

1. Nguyễn Huy Trưởng, *Đề tài NCKH:Nghiên cứu đánh giá tình trạng kỹ thuật và đề xuất giải pháp cải tiến nguồn động lực xe thiết giáp chở quân BTR-60PB*, Viện kỹ thuật Cơ giới Quân sự, 2011.

2. Hà Quang Minh, *Lý thuyết động cơ đốt trong* , NXB QĐND, 2002.

3. Nguyễn Phúc Hiểu, Vũ Đức Lập, *Lý thuyết xe quân sự,* Học viện KTQS, 2001.

4. Nguyễn Sĩ Đỉnh, *Hướng dẫn thiết kế đồ án môn học "Kết cấu tính toán ô tô", Tập I*, Học viện KTQS, 2015.

5. Nguyễn Văn Trà, *Kết cấu tính toán ô tô, Tập 1,* NXB QĐND, 2016.

6. Vũ Đức Lập, *Kết cấu tính toán ô tô, Tập 2*, NXB QĐND, 2015.

7. Vũ Đức Lập, *Ứng dụng máy tính trong tính toán xe Quân sự*, Học viện KTQS, 2001.

8. Nguyễn Phùng Quang *MatLab & Simulink*, NXB KHKT, 2004.

Tính toán, lựa chọn động cơ cho xe tăng hạng nhẹ sản xuất tại Việt Nam.

Nguyễn Văn Liêm[1*], Cù Xuân Phong[1]

[1]*Học viện Kỹ thuật Quân sự*

Email: liemnvd51@gmail.com

Tóm tắt:

Hiện nay, xu hướng cải tiến hiện đại hóa, sản xuất, chế tạo mới xe tăng-thiết giáp là xu hướng tất yếu của lực lượng Tăng thiết giáp. Chủ trương của Quân đội ta là từng bước nghiên cứu thiết kế, chế tạo các dòng xe tăng - thiết giáp mới thay thế các loại xe tăng -thiết giáp thế hệ cũ nhằm đáp ứng được yêu cầu huấn luyện và sẵn sàng chiến đấu, trong đó có xe tăng hạng nhẹ. Một trong những nhiệm vụ quan trọng đặt ra khi tính toán thiết kế chế tạo xe tăng là tính toán, lựa chọn động cơ nhằm đáp ứng tốt các yêu cầu chiến- kỹ thuật đề ra. Trong bài báo này, tác giả trình bày cơ sở lý thuyết và phương pháp tính toán động cơ cho xe tăng hạng nhẹ chế tạo tại Việt Nam.

Từ khóa: *Xe tăng hạng nhẹ; động cơ.*

1. Đặt vấn đề

Ngày nay, cùng với sự phát triển của khoa học công nghệ quân sự, yêu cầu về khả năng cơ động, hỏa lực và bảo vệ của các phương tiện chiến đấu ngày càng cao. Đặc biệt, đối với các dòng xe xích bơi nước, việc bảo đảm khả năng cơ động cả trên cạn và dưới nước có ý nghĩa quan trọng trong tác chiến hiện đại. Vì vậy, việc nghiên cứu thiết kế, chế tạo các dòng xe xích bơi nước đáp ứng yêu cầu tính năng chiến – kỹ thuật đặt ra là nhiệm vụ cấp thiết trong giai đoạn hiện nay.

Trong quá trình thiết kế xe xích bơi nước, một trong những nhiệm vụ quan trọng hàng đầu là tính toán, lựa chọn nguồn động lực cho xe, bao gồm động cơ và hộp số. Nguồn động lực không chỉ quyết định khả năng cơ động, tính năng vượt chướng ngại vật, khả năng bơi nước mà còn ảnh hưởng trực tiếp đến hiệu quả khai thác, độ tin cậy và tuổi thọ của phương tiện. Do đó, việc lựa chọn hợp lý nguồn động lực là yêu cầu mang tính quyết định đến chất lượng tổng thể của xe.

Hiện nay, với trình độ công nghiệp ô tô trong nước, chúng ta chưa làm chủ được công nghệ và dây chuyền chế tạo động cơ diesel công suất lớn phục vụ cho các phương tiện chiến đấu hạng nặng. Bên cạnh đó, các dòng xe bánh xích hiện đại thường sử dụng hộp số thủy cơ hoặc hộp số cơ khí hai dòng công suất (kết hợp truyền lực và quay vòng), có kết cấu phức tạp và yêu cầu công nghệ chế tạo cao. Vì vậy, trong điều kiện thực tế, bài toán thiết kế nguồn động lực cho xe xích bơi nước được chuyển thành bài toán tính toán, lựa chọn tổ hợp động cơ và hộp số phù hợp từ các sản phẩm đã có sẵn trên thị trường.

Xuất phát từ yêu cầu nhiệm vụ thiết kế và điều kiện công nghệ trong nước, phương án sử dụng động cơ diesel thông dụng kết hợp với hộp số cơ khí có cơ cấu quay vòng kết hợp, là hướng lựa chọn phù hợp. Giải pháp này vừa đảm bảo đáp ứng các yêu cầu về tính năng chiến – kỹ thuật của xe, vừa phù hợp với điều kiện bảo đảm kỹ thuật, khai thác và sửa chữa trong thực tế.

2. Cấu hình, tính năng kỹ thuật của xe tăng hạng nhẹ

2.1. Cấu hình đồng bộ xe

Phương án thiết kế mà nhóm tác giả đưa ra (Hình 1) là xe tăng hạng nhẹ, có khối lượng khoảng 18,6 tấn, có khả năng cơ động cao trên các địa hình phức tạp và có khả năng bơi nước bằng cơ cấu bơm phụt dẫn động bằng thủy lực.

- Vũ khí đồng bộ theo xe: 01 pháo 76 mm điều khiển tự động, 01 súng máy phòng không 12,7 mm, hệ thống điều khiển vũ khí gồm: Khối điều khiển hỏa lực, khối thiết bị quang điện tử

(camera ngày, đêm và đo xa laser), màn hình và bảng điều khiển cho xạ thủ, khối nguồn; có khả năng tự động bắt, bám mục tiêu trong khi hành tiến cả ban ngày và ban đêm.

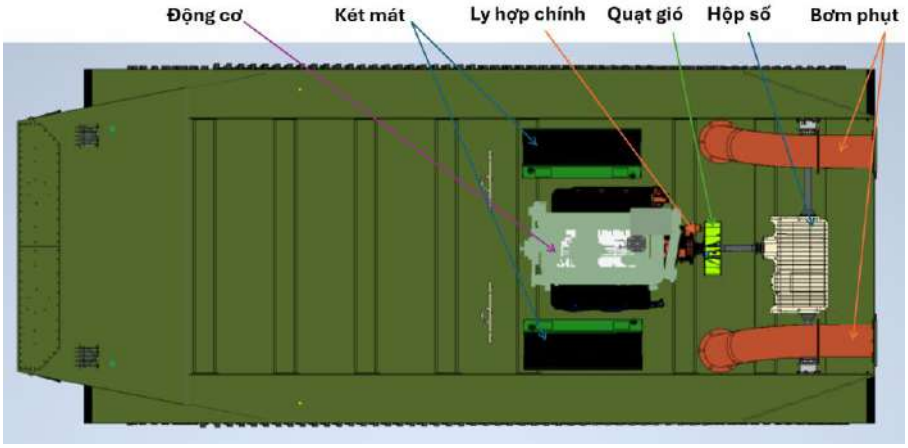
- Trang bị khí tài trên xe gồm: Hệ thống thông tin liên lạc nhảy tần cấp chiến thuật; hệ thống thiết bị quan sát ban ngày và ban đêm.

- Kíp xe gồm 3 thành viên,.

- Các hệ thống bảo vệ gồm: 02 dàn phóng đạn khói; hệ thống phòng hóa; hệ thống chữa cháy.

- Các bộ phận chính cấu thành xe bao gồm: thân xe, hệ thống vũ khí, hệ thống động lực, hệ thống truyền lực, hệ thống lái, hệ thống phanh, phần vận hành, hệ thống điện, hệ thống bơm lốp tự động, hệ thống bơi nước, các hệ thống bổ trợ (điều hòa nhiệt độ, bơm nước sàn xe).

- Đồng bộ kèm theo xe gồm: Đạn dược, dụng cụ, phụ tùng đảm bảo thay thế, sửa chữa nhỏ và tự cứu hộ.



Hình 1. Sơ đồ bố chí trung khoang động lực của xe tăng hạng nhẹ bơi nước

2.2. Yêu cầu về tính năng kỹ thuật

Tính năng chiến kỹ thuật đã được phê duyệt cho xe tăng hạng nhẹ tương đương xe PT76 của Nga sản xuất. Tác giả chỉ đưa ra một số tính năng kỹ thuật cơ bản có liên quan và làm thông số đầu vào cho bài toán tính toán thiết kế, lựa chọn động cơ cho xe.

Bảng 1. Chỉ tiêu kỹ thuật cơ bản

TT	Tên chỉ tiêu	Đơn vị	Giá trị
1	Khối lượng xe chiến đấu	Tấn	≤18,6
2	Công suất riêng	ml/T	≥22
3	Chiều dài tổng thể	mm	8.000±20
4	Chiều rộng tổng thể	mm	3.200
5	Chiều cao đến nóc xe	mm	3.062±20
6	Khoảng sáng gầm xe khi đủ tải	mm	≥40
7	Góc dốc cao nhất	Độ	30
8	Đi dốc nghiêng ngang lớn nhất	Độ	18
9	Vượt hào rộng	m	2,8
10	Vượt vách đứng	m	1,1
11	Vận tốc lớn nhất trên đường nhựa	km/h	70
12	Vận tốc lớn nhất khi bơi tiến	km/h	11
13	Góc dốc bến xuống bơi nước lớn nhất	Độ	30
14	Dự trữ hành trình trên đường nhựa	km	≥500

3. Tính toán, lựa chọn động cơ và hệ thống truyền lực

3.1. Tính toán, lựa chọn động cơ

Yêu cầu chiến kỹ thuật đặt ra cho xe xích được thiết kế mới đạt vận tốc lớn nhất v_m . Vận tốc này coi là đạt được khi xe chạy trên đường tốt, nghĩa là hệ số cản f nhỏ, độ dốc không cao. Các yếu tố đó được đặc trưng bởi hệ số cản chuyển động tổng cộng f_{min}, và bằng [1]:

$$f_{cmin} = f + i$$

Trong đó: i- độ dốc của đường, được tính bằng tỷ số giữa chiều cao của dốc với hình chiếu của đoạn đường xem xét trên mặt phẳng nằm ngang; i= tgα= sinα, khi chuyển động trong điều kiện nói trên thì giá trị, các hệ số f và i là f= 0,015÷ 0,025 và i= 0,025; Từ đó hệ số cản chuyển động tổng cộng khi xe xích đạt vận tốc lớn nhất sẽ bằng.

$$f_{cmin} = 0,04 \div 0,05 \qquad (1.1)$$

Giá trị lực kéo đơn vị $f_đ$ khi xe chuyển động đều cân bằng với hệ số cản chuyển động f_c.

Mặt khác, $f_đ$ được xác định bằng công thức [1]::

$$f_đ = \frac{N_đ \eta_t}{G v_i}$$

Trong đó:

$F_đ$- Lực kéo đơn vị của xe;

$N_đ$- Công suất tự do của động cơ [kW];

G- Trọng lượng toàn bộ của xe [N];

v_i- Vận tốc chuyển động thẳng của xe ứng với tay số thứ i [km/h];

η_t- Hiệu suất của xe. ($\eta_t = \eta_{tl} . \eta_{vh}$)

η_{tl}- Hiệu suất hệ thống truyền lực.

η_{vh}- Hiệu suất xích.

Khi xe chuyển động đều ở tốc tộ v_{max} lực kéo đơn vị bằng hệ số cản f_{cmin}, từ biểu thức (1.1) ta sẽ xác định được công suất tự do cực đại đảm bảo cho xe xích đạt đến vận tốc lớn nhất trong điều kiện đã cho [1]:.

$$N_{đN} = \frac{f_{cmin} G v_{max}}{\eta_t}$$

Hiệu suất của các hệ thống truyền lực ở số cao nhất nằm trong khoảng η_{tl}= 0,86÷ 0.92.

Hiệu suất của phần vận hành, có thể tính theo công thức thực nghiệm. Trường hợp xem xét ta biết f_{cmin} ở số cao nhất bằng: $f_{c\min} = f_{cđ} . \eta_{vh}$. Nên hiệu suất của phần vận hành [1]:

$$\eta_{vh} = \frac{0,95 . f_{c\min}}{f_{c\min} + 0,025 + 0,000003 v_m^{\ 2}}$$

Theo các kết cấu xe xích đã có, với hệ thống làm mát bằng quạt gió thì hiệu suất của thiết bị động lực khi động cơ làm việc ở chế độ công suất cực đại $\eta_{đlN}$ nằm trong khoảng 0,85 ÷ 0,88.

Khi đó công suất cực đại của động cơ là [2,6]:

$$N_{emax} = \frac{N_{đN}}{\eta_{đlN}}$$

Trên cơ sở lý thuyết ta đã trình bày, với các thông số đâu vào là:

Trọng lượng G của xe: 18,6 tấn

Vận tốc lớn nhất trên đường nhựa: 70 Km/h.

Ta xác định công suất động cơ cực đại là: 448,8 mã lực.

Sau khi xác định được N_{emax}, dựa vào các kiểu động cơ đã có, chọn một động cơ có các tham số phù hợp nhất.

Đối tượng nghiên cứu trong bài báo là lựa chọn các loại động cơ diesel thông dụng, kết hợp với hộp số cơ khí hai dòng công suất, có cơ cấu quay vòng kết hợp nhằm dễ dàng trong việc lựa chọn khi mua sắm và đơn giản trong gia công, chế tạo sau này.

Để đảm bảo thông số công suất riêng của xe trong khoảng 20 đến 30 mã lực/tấn và vận tốc đã xác định. Chúng ta có một số loại động cơ trong bảng 1. Cụ thể như sau:

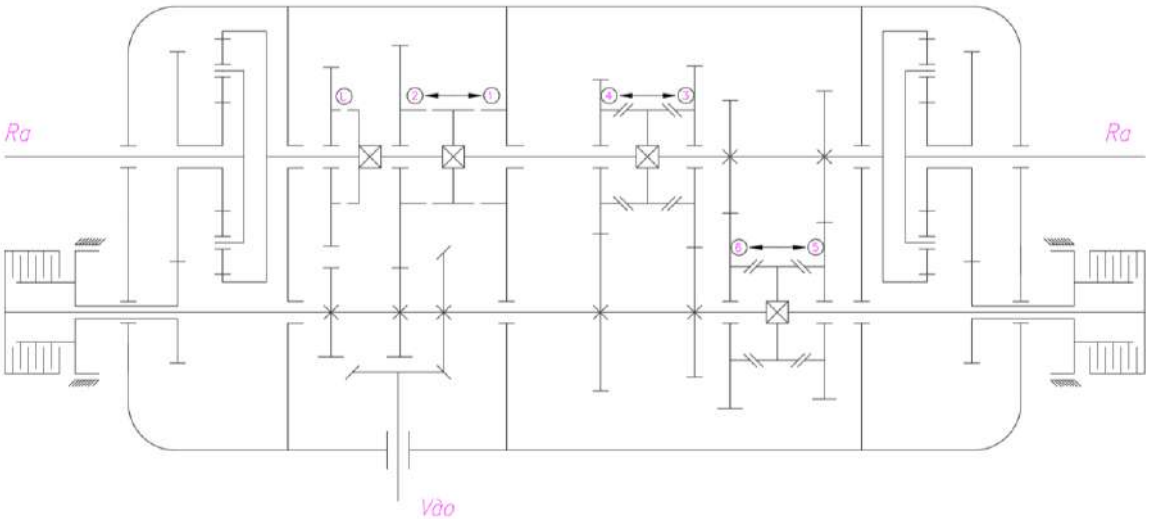
Bảng 2. Thông số kỹ thuật một số loại động cơ

TT	Thông số kỹ thuật	Động cơ KamAZ	Động cơ TMZ 8424.10-03	Động cơ TMZ 8424.10
1	Kiểu động cơ	Diezel, 4 kỳ, tăng áp, 8 máy, chữ V	Diezel, 4 kỳ, tăng áp, 8 máy, chữ V	Diezel, 4 kỳ, tăng áp, 8 máy, chữ V
2	Công suất lớn nhất (mã lực/kW)	400/298,28	470/345,5	425/312,5
3	Mômen xoắn lớn nhất (N.m)	1570	1864	1686
4	Số vòng quay ở công suất lớn nhất (rpm)	2200	2100	2100
5	Số vòng quay ở mô men lớn nhất (rpm)	1400	1400	1400
6	Số vòng quay nhỏ nhất (rpm)	650	650	600

Từ thông số động cơ trong bảng và kết quả tính toán công suất cực đại cần có ta chọn **Động cơ TMZ 8424.10-03** để kiểm nghiệm cho xe phương án thiết kế xe.

3.2. Lựa chọn các cụm trong hệ thống truyền lực

Đối với hộp số, để thuận tiện cho việc khai thác sử dụng, cũng như quá trình tự chủ về công nghệ, sẵn sàng cho việc tự chủ chế tạo vật tư hoặc hộp số thay thế. Chúng ta chọn loại hộp số cơ khí, có tính năng hiện đại là hai dòng công suất và kết hợp với cơ cấu quay vòng hành tinh của xe MTLB. Hộp số có sơ đồ động học như hình 2 dưới đây [3-5].



Hình 2. Sơ đồ động học hộp số cơ khí 6 cấp kết hợp cơ cấu quay vòng hành tinh

Thông số đầu vào để tính toán và tỉ số truyền của hệ thống truyền lực trong bảng 3.

Bảng 3. Thông số kỹ thuật đầu vào tính toán

STT	Các thông số kỹ thuật	Ký hiệu	Số liệu	Đơn vị
1	Trọng lượng xe	G	186000	N
2	Tỷ số truyền đầu vào hộp số	i_{hi}		
-	Số 1		7.97	
-	Số 2		4,02	
-	Số 3		1,93	
-	Số 4		1,07	
-	Số 5		0,75	
-	Số 6		0,56	
-	Số lùi		3,14	
3	Tỷ số của truyền động cạnh	$i_{tđc}$	6	
4	Bán kính bánh sao chủ động	r_k	0,28	m

Từ các thông số động cơ và hộp số đã chọn vận tốc lớn nhất tương ứng là [1]:

$$v_{max} = \frac{n_N.\pi r_k.3,6}{30.i_6 i_{tđc}} = 65,94 \text{ (Km/h)}$$

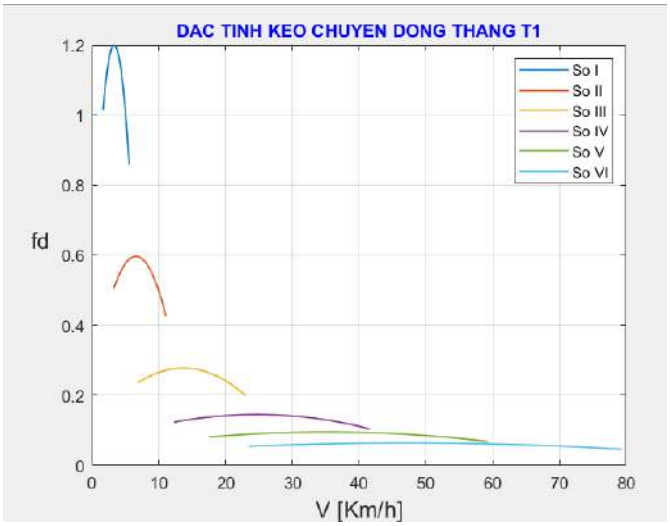
Theo yêu cầu bài toán, vận tốc lớn nhất cần đạt được của xe là 70 Km/h. Để đạt được điều đó, ta sử dụng thêm hộp tăng tốc phía sau hôp số chính với tỉ số chuyền tương ứng là:

$$i_{tc} = \frac{65,94}{70} = 0,942$$

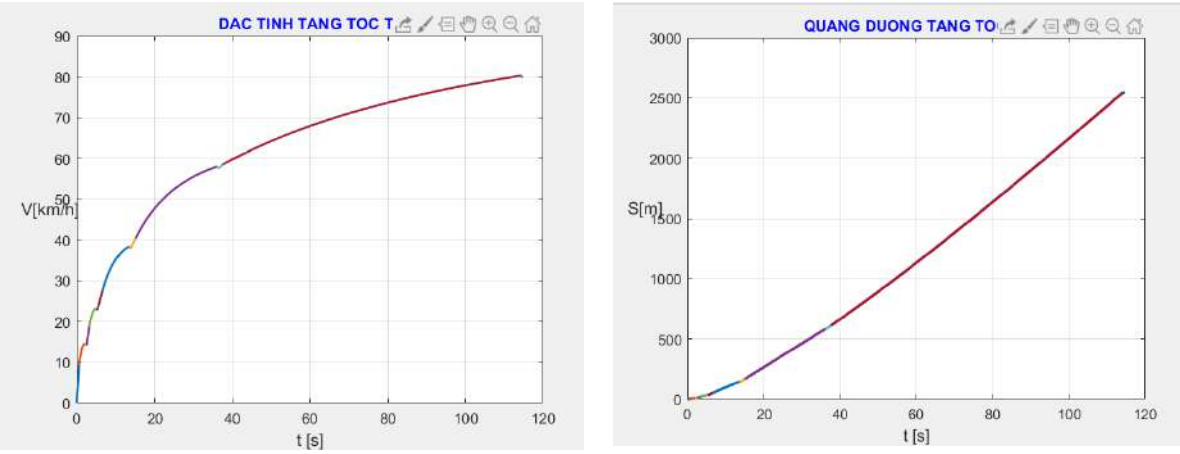
Theo đặc tính tăng tốc của xe, để xe đạt vận tốc thực tế 70 Km/h và kết cấu của hộp truyền tăng tốc trong thực tế ta chọn tỉ số truyền của hộp tăng tốc trong khoảng từ 0,65 ÷ 0,87.

3.3. Xây dựng đồ thị đặc tính kéo chuyển động thẳng

Với các thông số của động cơ và hộp số đã tính toán, lựa chọn, sử dụng các công thức tính toán [2,7,8,9] và phần mềm Matlab tác giả xây dựng được các đồ thị đặc tính kéo chuyển động thẳng xe tăng hạng nhẹ (Hình 3,4) nhằm đánh giá chất lượng kéo chuyển động thẳng của xe.



Hình 3. Đồ thị kết quả lực kéo đơn vị



Hình 4. Đồ thị kết quả quãng đường và thời gian tăng tốc

Qua kết quả tính toán chúng ta nhận thấy:

- Tốc độ lớn nhất của xe khi dùng động cơ **Động cơ TMZ 8424.10-03** là khoảng 79 km/h và thời gian tăng tốc, quãng đường tăng tốc đến tốc độ lớn nhất tương ứng nhỏ hơn 118s và 2600m, điều này phù hợp với tính năng chiến kỹ thuật đã đề ra.

- Chỉ số lực kéo đơn vị ở tay số 1 ở cả 3 động cơ là lớn, thể hiện khả năng vượt vật cản ở tay số 1 đối với 3 động cơ đều đảm bảo.

4. Kết luận

Việc tính toán lựa chọn nguồn động lực cho xe tăng hạng nhẹ là một bài toán quan trọng. Tuy nhiên, trong điều kiện nền công nghiệp ô tô trong nước chưa có khả năng tự sản xuất động cơ, có thể tính chọn dựa trên một số loại động cơ thông dụng, có sẵn trên thị trường. Các kết quả tính toán cho thấy, đối với xe tăng hạng nhẹ, khối lượng khoảng 18,6 tấn, có thể lựa chọn động cơ **Động cơ TMZ** với công suất tương ứng là 470 mã lực (tương đương công suất riêng là 25 mã lực/tấn). Điều này phù hợp với yêu cầu tính năng chiến kỹ thuật đề ra ban đầu.

Tương ứng với việc lựa chọn động cơ, để đảm bảo tính hiện đại đồng thời đơn giản trong, bảo dưỡng, sửa chữa, thay thế vật tư phụ tùng và chế tạo sau này, việc lựa chọn hộp số 6 số truyền cơ khí, có truyền lực và quay vòng kết hợp là lựa chọn phù hợp, nó vừa đáp ứng được tốc độ chuyển động lớn nhất của xe ở tay số cao nhất, đồng thời vừa đảm bảo được lực kéo lớn ở tay số 1. Đồng thời, với kết cấu hộp số 2 dòng công suất sẽ làm giảm kích thước, khối lượng và tăng độ bền của hộp số khi sử dụng./.

Tài liệu tham khảo

1. Lê Trung Dũng, Nguyễn Minh Tân, Cù Xuân Phong, *Lý thuyết xe xích quân sự*, NXB QĐ, 2023
2. Hà Quang Minh, *Lý thuyết động cơ đốt trong* , NXB QĐND, 2002.
3. Lê Kỳ Nam và Nguyễn Văn Luận, *Kết cấu và tính toán xe tăng thiết giáp*, Học viện Kỹ Thuật Quân sự, Hà Nội, 1997
4. Lê Kỳ Nam, *Cấu tạo xe tăng tập 1*, NXB QĐ, 2003.
5. Võ Văn Trung, *Cấu tạo xe tăng tập 3*, NXB QĐ, 2004.
6. Dương Thành Công, Cù Xuân Phong, Tô Viết Thành, Nguyễn Anh Vượng, Vũ Mạnh Tấn. *Nghiên cứu thay thế động cơ trên xe thiết giáp xích.* Tạp chí Khoa học Công nghệ Hàng Hải, 2021.
7. Nguyễn Đình Tuấn (2004), *Hướng dẫn đồ án môn học kết cấu tính toán xe tăng – Hộp số cơ khí",* Học viện KTQS.
8. Vũ Đức Lập, *Ứng dụng máy tính trong tính toán xe xích Quân sự*, Học viện Kỹ Thuật Quân sự, Hà Nội, 2001
9. Nguyễn Phùng Quang *MatLab & Simulink*, NXB KHKT, 2004.

Calculation and Selection of Engines for Light Tanks Manufactured in Vietnam

Abstract: *Currently, the trend toward modernization, upgrading, and new production of armored vehicles and tanks has become an inevitable direction of the Armored Forces. The policy of the Vietnam People's Army is to progressively conduct research, design, and manufacture new generations of tanks and armored vehicles to replace older platforms, thereby meeting the requirements of training and combat readiness, including light tanks.One of the critical tasks in the design and manufacturing process of a tank is the calculation and selection of a suitable engine to ensure compliance with the specified tactical and technical requirements. In this paper, the author presents the theoretical basis and calculation methodology for engine selection and design applied to a light tank manufactured in Vietnam.*

Keywords: Light tank; engine.

ĐÁNH GIÁ KHẢ NĂNG TĂNG TỐC XE THIẾT GIÁP CHỞ QUÂN 8X8

Tạ Văn Tuấn[1*], Nguyễn Minh Tân[1],Dương Thành Công[1]

[1]Học viện Kỹ thuật Quân sự

Email: tuanttgmta88@gmail.com, Tel: 0962248865

Tóm tắt

Trong quá trình thiết kế, chế tạo xe thiết giáp chở quân 8×8, việc đánh giá đặc tính động lực học chuyển động thẳng, đặc biệt là khả năng tăng tốc, có vai trò quan trọng đối với nhiệm vụ hoàn thiện thiết kế. Bài báo trình bày cơ sở lý thuyết khảo sát động lực học dọc của xe và xây dựng phương pháp tính toán đánh giá khả năng tăng tốc dựa trên phương trình xây dựng. Kết quả nghiên cứu là cơ sở khoa học phục vụ tính toán thiết kế hệ thống truyền lực, lựa chọn tỷ số truyền hợp lý và tối ưu hóa cấu hình động lực trước khi tiến hành thử nghiệm thực địa.

Từ khóa: *Xe thiết giáp; động lực học; tăng tốc.*

1. Đặt vấn đề

Trong bối cảnh ngành công nghiệp quốc phòng đẩy mạnh nghiên cứu, thiết kế và chế tạo xe thiết giáp chở quân 8×8 theo hướng tự chủ công nghệ, việc đánh giá các đặc tính động lực học ngay từ giai đoạn thiết kế có ý nghĩa đặc biệt quan trọng. Trong số các chỉ tiêu động lực học chuyển động thẳng, khả năng tăng tốc là thông số phản ánh trực tiếp mức độ phù hợp giữa động cơ, hệ thống truyền lực và điều kiện bám của bánh xe với mặt đường, đồng thời quyết định hiệu quả cơ động chiến thuật của phương tiện.

Đối với xe thiết giáp 8×8 có khối lượng lớn và cấu hình dẫn động nhiều trục, quá trình tăng tốc chịu ảnh hưởng tổng hợp của lực kéo tiếp tuyến tại bánh xe, tổn hao công suất trong hệ thống truyền lực, lực cản lăn, lực cản khí động và sự phân bố tải trọng động lên các cầu chủ động. Nếu các yếu tố này không được xem xét đầy đủ trong tính toán thiết kế, cấu hình truyền lực có thể chưa đạt trạng thái tối ưu, làm suy giảm khả năng cơ động và ảnh hưởng đến hiệu quả thực hiện nhiệm vụ.

Vì vậy, việc xây dựng mô hình động lực học dọc và thiết lập phương pháp tính toán xác định khả năng tăng tốc là yêu cầu cần thiết trong quá trình thiết kế. Nghiên cứu này nhằm cung cấp cơ sở khoa học phục vụ lựa chọn tỷ số truyền hợp lý, tối ưu hóa cấu hình hệ truyền lực và hoàn thiện thiết kế.

Hình 1. Xe thiết giáp chở quân 8x8

2. Xây dựng mô hình

2.1. Các tiêu chí đánh giá khả năng tăng tốc

Khả năng tăng nhanh vận tốc sau khi khởi hành hoặc sau khi chuyển số được gọi là tính năng tăng tốc của xe thiết giáp.

Tính năng tăng tốc được đánh giá bằng:

- Thời gian tăng tốc.

- Quãng đường tăng tốc.

Kết quả của việc xác định các thông số đánh giá quá trình tăng tốc của xe thiết giáp bằng thực nghiệm hoặc bằng phương pháp tính toán là xây dựng được đồ thị thời gian và quãng đường tăng tốc phụ thuộc vào vận tốc chuyển động của xe. Khi so sánh tính năng tăng tốc của các xe thiết giáp cùng loại có vận tốc tối đa khác nhau người ta thường chọn thời gian và quãng đường tăng tốc đến vận tốc mà xe thiết giáp hoạt động lâu dài trong điều kiện đường tốt để đánh giá.

2.2. Xây dựng mô hình khảo sát các yếu tố ảnh hưởng đến tính năng tắng tốc xe thiết giáp chở quân 8x8

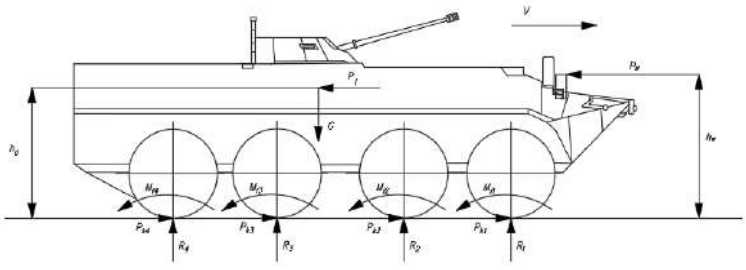
Xây dựng mô hình toán học cho việc nghiên cứu mô hình phẳng của xe thiết giáp trong chuyển động thẳng, ta chấp nhận các giả thiết theo trong các tài liệu [1,2,3] sau:

- Xe chuyển động trên đường nằm ngang, bằng phẳng, đồng nhất.

- Xét hình chiếu đứng của xe XTC-03, xe không kéo moóc.

- Tiếp xúc của bánh xe với đường là tiếp xúc điểm. Mọi sự dời tâm áp lực thay thế bằng mô men cản lăn.

Trên cơ sở các giả thiết nêu trên sơ đồ ngoại lực và mô men ngoại lực tác dụng lên xe thiết giáp được thể hiện trên Hình 2.



Hình 2. Sơ đồ mô hình phẳng của xe thiết giáp

Trong đó:

G- Trọng lượng xe thiết giáp.

Pw- Lực cản không khí.

M_{f1}, M_{f2}, M_{f3}, M_{f4} - Mô men cản lăn.

Pj- Lực quán tính.

R_1, R_2, R_3, R_4- Phản lực pháp tuyến của đường.

P_{k1}, P_{k2}, P_{k3}, P_{k4}- Lực kéo tiếp tuyến (phản lực tiếp tuyến của đường).

Sơ đồ khảo sát đặt trong hệ trục toạ độ XOZ, trục OX song song với mặt đường, trục OZ vuông góc với mặt đường và đi qua trọng tâm của xe. Từ đó xây dựng được phương trình động lực học tăng tốc của xe thiết giáp, thiết lập được mối quan hệ giữa các nội lực và ngoại lực tác dụng lên xe thiết giáp, xác định được giá trị của các lực chưa biết khi cho trước những số liệu cần thiết.

2.3. Tính toán động lực học quá trình tăng tốc xe thiết giáp chở quân 8x8

Chiếu các lực lên trục OX (Hình 2) ta được phương trình cân bằng lực:

$$P_j + P_f + P\text{w} - P_k = 0$$

Suy ra: $$P_j = P_k - P_f - P\text{w} \quad (1)$$

Khi xe tăng tốc, lực quán tính của xe thiết giáp xác định bằng:

$$P_j = \delta.\frac{G}{g}.\frac{dv}{dt} \quad (2)$$

Trong đó δ: Hệ số ảnh hưởng khối lượng quay

$$\delta = 1 + 0,05.\left(1 + i_{tli}^2\right).\frac{G_d}{G_b}$$

Trong đó: i_{tli}- Tỷ số truyền của hộp số ở số truyền i.

G_d- Trọng lượng toàn bộ xe ứng với tải định mức.

G_b- Trọng lượng toàn bộ xe ứng với tải bất kỳ.

G- Trọng lượng toàn bộ của xe thiết giáp.

g- Gia tốc trọng trường.

$\frac{dv}{dt} = j_{tt}$: Gia tốc tăng tốc của xe thiết giáp.

Thay công thức (2) vào (1) ta được: $\delta.\frac{G}{g}.\frac{dv}{dt} = P_k - P_f - P_\text{W}$

Chia hai vế cho G ta được: $$\frac{\delta}{g}.\frac{dv}{dt} = \frac{P_k - P_\text{W}}{G} - \frac{P_f}{G} = D - f$$

Hay $$\frac{dv}{dt} = j = \frac{g}{\delta}.(D - f) \quad (3)$$

Trong đó: $D = \frac{P_k - P_\text{W}}{G}$ - Nhân tố động lực học của xe thiết giáp.

f - Hệ số cản lăn

Từ công thức (3), thời gian tăng tốc được xác định bằng

$$dt = \frac{\delta}{g\left(D - f\right)}.dv = \frac{1}{j}.dv \quad (4)$$

Sau khi xác định được thời gian tăng tốc, có thể tính được quãng đường tăng tốc ứng với thời gian đó:

$$ds = v.dt \quad (5)$$

Từ các công thức trên ta thấy để xác định được thời gian tăng tốc và quãng đường tăng tốc trước hết phải xây dựng được lực kéo P_{ki} ở từng tay số, từ đó tính được nhân tố động lực học D_i. Phương pháp xác định lực kéo theo động cơ ở từng tay số hiện nay thường hay áp dụng là sử dụng đường đặc tính ngoài của động cơ làm cơ sở cho quá trình tính toán.

Xây dựng đặc tính kéo xe thiết giáp.

Đồ thị biểu diễn mối quan hệ giữa lực kéo tính theo động cơ và vận tốc chuyển động của xe ở các số truyền được gọi là đặc tính kéo của xe thiết giáp.

Công thức tính toán:

$$P_{ki} = \frac{N_e.\eta_{tli}}{v_i} \text{ (N)} \quad (6)$$

Trong đó: P_{ki}- Lực kéo ở tay số thứ i. (N)

η_{tli}- Hiệu suất truyền lực.

v_i- Vận tốc chuyển động của xe ở số truyền thứ I (m/s).

Vận tốc v_i được xác định như sau:

$$v_i = \frac{\omega_e.\mathrm{r}_k}{i_{tli}} \quad (7)$$

Trong đó: R_k- Bán kính tính toán của bánh xe (m).

i- Tỷ số truyền của hệ thống truyền lực ở số truyền i.

ω_e- Vận tốc góc của trục khuỷu động cơ (rad/s).

Xây dựng đặc tính động lực học

Nhân tố động lực học của xe là tỷ số giữa hiệu lực kéo theo động cơ và lực cản không khí với trọng lượng toàn bộ xe.

$$D = \frac{P_{ki} - P_{wi}}{G} \quad (8)$$

P_w: Lực cản không khí.

$$P\mathrm{w}i = K_b B H v_i^2 \quad (9)$$

K_b: Hệ số cản không khí.

B: Chiều rộng cơ sở của xe (m)

H: Chiều cao toàn bộ xe (m)

v_i: Vận tốc của xe ở số truyền thứ i.

Đặc tính động lực học của xe thiết giáp là đồ thị biểu diễn mối quan hệ giữa nhân tố động lực học với vận tốc chuyển động của xe thiết giáp ở các số truyền.

D = D(v)

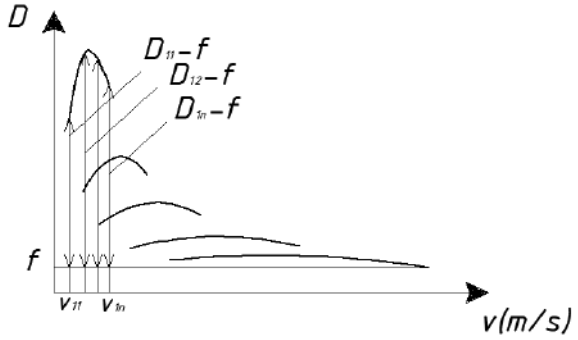
Trước tiên xây dựng D(v) ở số truyền 1. Với các giá trị v_{1i} đã tính được ở trên,tiến hành xác định P_{w1i}. Thay các giá trị P_{w1i}, P_{k1i} vừa tính được vào công thức (8) xác định D_{11}, D_{12},…, D_{1n}. Tương tự tính cho các số truyền còn lại.

Xây dựng đặc tính tăng tốc

-Đồ thị gia tốc của xe.

Theo công thức (3), ta có: $j_i = \frac{g}{\delta_i}.(D_i - f)$

Thay các giá trị D_i đã xác định ở trên vào công thức (3), ta xác định được các giá trị ji ở các số truyền.



Hình 3. Phương pháp xây dựng đồ thị gia tốc của xe thiết giáp

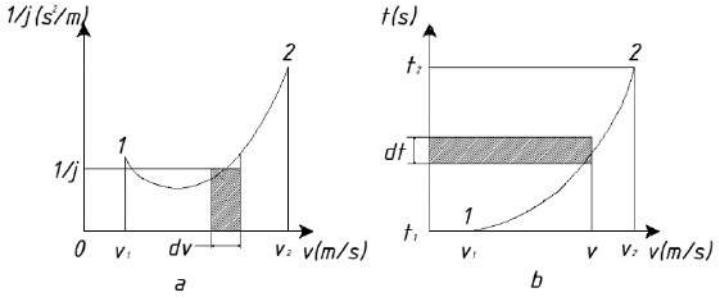
- Đồ thị thời gian tăng tốc.

Thời gian tăng tốc từ vận tốc v1 đến vận tốc v2 được xác định bằng tích phân sau:

$$t = \int_{v_1}^{v_2} \frac{1}{j}.dv \tag{10}$$

Tích phân này không giải được bằng phương pháp giải tích vì không tìm được mối quan hệ giải tích chính xác giữa gia tốc j với vận tốc v. Tuy nhiên tích phân (3.14) có thể giải được bằng phương pháp đồ thị trên cơ sở đồ thị gia tốc.

Để xây dựng đồ thị thời gian tăng tốc ở một số truyền trước tiên xây dựng đồ thị gia tốc ngược. Từ đồ thị gia tốc ngược (Hình 4a) lấy một phần diện tích nào đó ứng với khoảng biến thiên vận tốc dv. Kết quả gần đúng nhận được một hình chữ nhật có chiều dài là và chiều rộng là d_V. Diện tích này là $dt = \frac{1}{j}.dv$. Như vậy thời gian tăng tốc của xe thiết giáp ở một số truyền i nào đó từ vận tốc v_1 đến v_2 sẽ chính là diện tích của hình được giới hạn bởi đồ thị $\frac{1}{j}$ và trục hoành ứng với v_1 và v_2 (diện tích hình v112v2).



Hình 4. Đồ thị gia tốc ngược (a) và đồ thị thời gian tăng tốc của xe thiết giáp ở một số truyền (b)

Thời điểm chuyển số là khi mô men động cơ đạt cực đại Memax (khi số vòng quay đạt 1450 vòng/phút).

Hộp số xe thiết giáp chở quân 8x8 là hộp số có cấp, có đồng tốc nên thời gian sang số t_s = 1 (s). Vận tốc chuyển động của xe sẽ giảm một lượng Δv_s bằng:

$$\Delta v_s = 9,3.t_s.f = 9.3.1.0,018 = 0,17(m/s)$$

- Đồ thị quãng đường tăng tốc.

Quãng đường tăng tốc từ thời điểm t_1 đến thời điểm t_2 sẽ là:

$$s = \int_{t_1}^{t_2} v dt \tag{11}$$

Tích phân này cũng không giải được bằng phương pháp giải tích vì không tìm được mối quan hệ giải tích chính xác giữa vận tốc v và thời gian t. Tuy nhiên có thể giải được bằng phương pháp đồ thị trên cơ sở đồ thị t = t(v) vừa xây dựng trên.

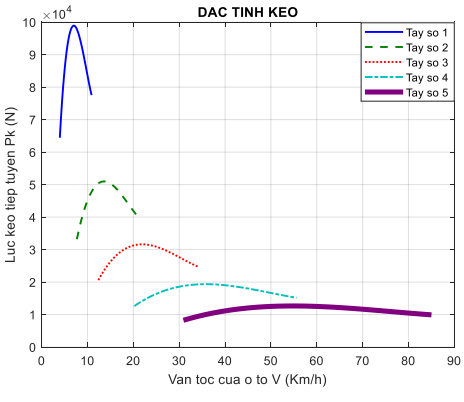
Trên đồ thị t = t(v) lấy một phần diện tích ứng với khoảng biến thiên thời gian dt (Hình 4b). kết quả gần đúng nhận được hình chữ nhật có chiều dài v và chiều rộng dt. Diện tích hình chữ nhật này chính là vi phân ds. Như vậy diện tích giới hạn bởi đường cong t = t(v) và trục tung t ứng với t_1 và t_2 là quãng đường tăng tốc của xe thiết giáp từ v_1 và v_2.

3. Kết quả và thảo luận

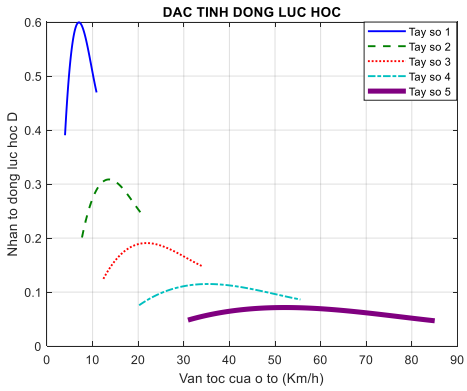
Theo yêu cầu tính năng của xe đã được đề ra, tác giả đưa ra được các thông số kỹ thuật phục vụ cho tính toán động lực học quá trình tăng tốc như sau:

Bảng 1. Các thông số tính toán động lực học quá trình tăng tốc xe thiết giáp chở quân 8x8

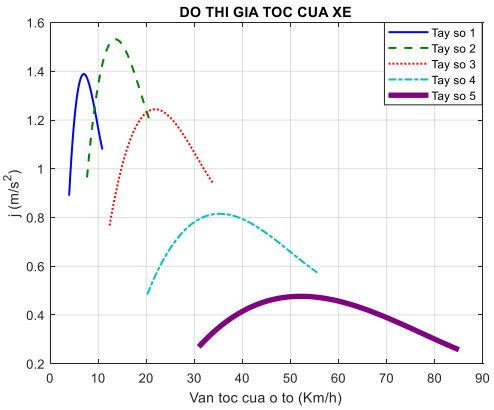
STT	Tên gọi	Ký hiệu	Thông số	Đơn vị tính
1	Trọng lượng toàn bộ	G	165000	N
2	Chiều rộng cơ sở của xe	B	2,460	m
3	Chiều cao toàn bộ của xe	H	2,650	m
4	Mômen xoắn lớn nhất của động cơ ở vòng quay 1450 v/ph.	M_{emax}	1400	Nm
5	Công suất có ích lớn nhất của động cơ ở vòng quay 2200 v/ph	N_{emax}	260	kW
6	Vận tốc lớn nhất của xe	V_{max}	85	km/h
7	Số vòng quay nhỏ nhất của trục khuỷu động cơ	n_{min}	800	v/ph
8	Bán kính thiết kế	r	0,535	m
9	Hệ số biến dạng của lốp	λ	0,932	
10	Tỉ số truyền của cầu xe	i_0	7,99	
11	Tỉ số truyền của bộ chia	i_{cao}	0,76	
12	Tỉ số truyền hộp số phụ	i_p	0,8	
13	Tỉ số truyền của hộp số Số I Số II Số III Số IV Số V	 i_I i_{II} i_{III} i_{IV} i_V	 7,82 4,03 2,50 1,53 1	
14	Hệ số cản không khí	K_b	0,6	
15	Hệ số cản lăn	f	0,018	
16	Hệ số bám	φ	0,8	
17	Hiệu suất truyền lực	η_{tl}	0,95	



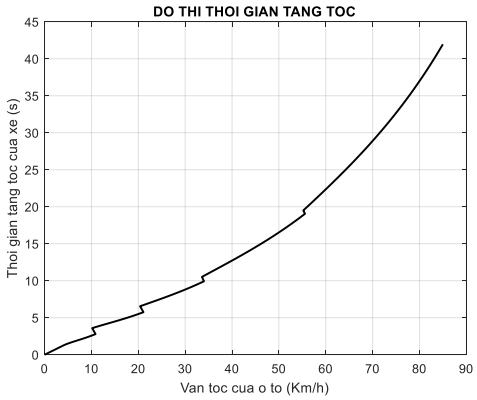
Hình 5. Đồ thị lực kéo từng tay số



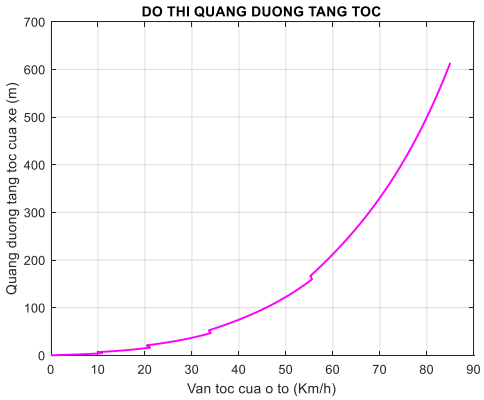
Hình 6. Đồ thị nhân tố động lực học



Hình 7. Đồ thị gia tốc của xe thiết giáp



Hình 8. Đồ thị thời gian tăng tốc



Hình 9. Đồ thị quãng đường tăng tốc

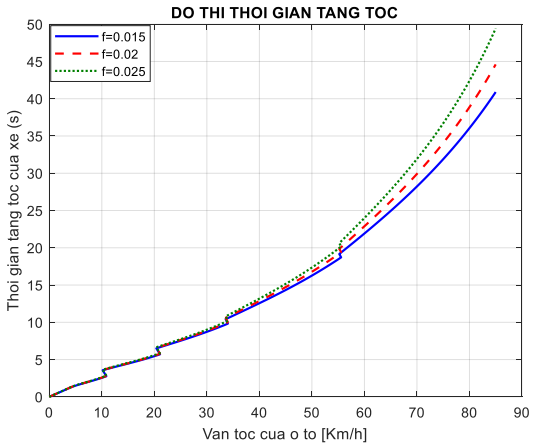
Dựa vào đồ thị ta thấy để xe thiết giáp tăng tốc từ 0 km/h đến 85 km/h mất khoảng 41,95 s. Từ 0 km/h đạt đến vận tốc 40 đến 50 km/h, thời gian tăng tương đối nhanh, nhưng sau đó độ dốc của đồ thị tăng mạnh, chứng tỏ việc đạt được vận tốc cao hơn đòi hỏi nhiều thời gian hơn. Điều đó phù hợp với thực tế hoạt động của xe thiết giáp.

Ta thấy để đạt được vận tốc 85 km/h thì xe phải tăng tốc trên quãng đường 613,82 m. Ở các mức vận tốc thấp, xe chỉ cần một quãng đường ngắn để tăng tốc. Tuy nhiên, khi vận tốc càng cao, quãng đường cần thiết để đạt được vận tốc đó tăng mạnh, phản ánh sự gia tăng lực cản và giới hạn hiệu suất động cơ. Điều đó phù hợp thực tế hoạt động của xe thiết giáp.

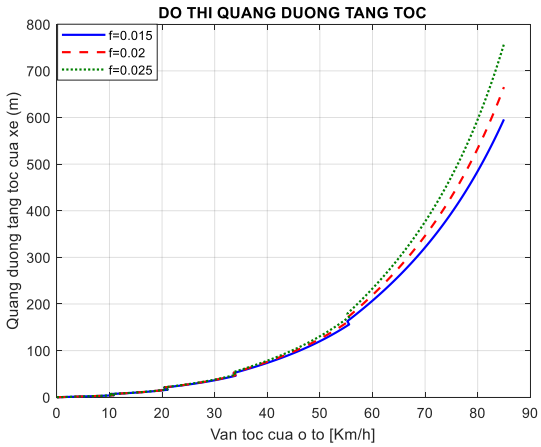
Khảo sát ảnh hưởng quá trình tăng tốc bởi hệ số cản lăn ta chỉ xét trường hợp hệ số cản lăn là hằng số không phụ thuộc vào vận tốc chuyển động của xe thiết giáp (f = const). Khi đó lực cản lăn sẽ là một đường thẳng song song với trục hoành và cách trục hoành một đoạn không đổi bằng:

$$P_f = f.G \quad (12)$$

Từ công thức (3) ta thấy rằng khi thay đổi giá trị của hệ số cản lăn, gia tốc tăng tốc j cũng thay đổi theo. Do đó, thời gian tăng tốc và quãng đường tăng tốc cũng thay đổi. Thay đổi hệ số cản lăn f của xe thiết giáp với các giá trị 0,015; 0,02; 0,025 ta được 3 đường đồ thị thời gian tăng tốc, quãng đường tăng tốc theo thời gian tương ứng với các giá trị hệ số cản lăn và được thể hiện trên Hình 10 và Hình11.



Hình 10. Đồ thị thời gian tăng tốc khi thay đổi hệ số cản lăn



Hình 11. Đồ thị quãng đường tăng tốc khi thay đổi hệ số cản lăn

Nhận thấy khi tăng hệ số cản lăn, thời gian và quãng đường tăng tốc của xe thiết giáp đến vận tốc lớn nhất 85 Km/h tăng đần và được thể hiện trên bảng 2. Từ kết quả khảo sát ta có khi tăng hệ số cản lăn lên một lượng 0,05 thì thời gian tăng tốc tăng lên khoảng 9,2%, quãng đường tăng tốc tăng lên khoảng 11,3%. Như vậy hệ số cản lăn càng lớn thì thời gian và quãng đường tăng tốc càng lớn.

Bảng 2. Thời gian và quãng đường tăng tốc với các giá trị hệ số cản lăn

Hệ số cản lăn	**t_{tt} (s)**	**S_{tt} (m)**
f = 0,015	40,89	594,57
f = 0,02	44,60	663,86
f = 0,025	49,43	755,97

4. Kết luận

Bài báo đã xây dựng cơ sở lý thuyết khảo sát động lực học chuyển động thẳng của xe thiết giáp chở quân 8×8 và thiết lập phương pháp đánh giá định lượng khả năng tăng tốc trên cơ sở phương trình cân bằng lực kéo và các lực cản chuyển động. Kết quả nghiên cứu cho thấy khả năng tăng tốc của xe phụ thuộc chặt chẽ vào công suất động cơ, tỷ số truyền, hiệu suất hệ

thống truyền lực và khối lượng của xe. Mô hình tính toán cho phép xác định các thông số đặc trưng như gia tốc tức thời và thời gian tăng tốc, làm cơ sở phục vụ tính toán thiết kế và tối ưu hóa hệ thống truyền lực. Tuy nhiên, nghiên cứu hiện tại được xây dựng trên mô hình động lực học dọc lý thuyết với một số giả thiết đơn giản hóa như điều kiện mặt đường bằng phẳng, hệ số bám ổn định và chưa xét đến ảnh hưởng dao động của hệ thống treo cũng như phân bố tải trọng động giữa các trục trong quá trình tăng tốc. Các yếu tố phi tuyến của lốp và tổn hao biến thiên trong hệ truyền lực cũng chưa được xét đến đầy đủ. Trong các nghiên cứu tiếp theo, cần phát triển mô hình động lực học toàn xe có xét đến tương tác giữa dao động thẳng đứng và chuyển động dọc, đồng thời bổ sung mô hình đặc tính bám lốp - mặt đường trong các điều kiện địa hình khác nhau. Việc kết hợp mô phỏng số với số liệu thử nghiệm thực tế sẽ góp phần nâng cao độ chính xác của mô hình và hoàn thiện cơ sở khoa học cho quá trình thiết kế, chế tạo xe thiết giáp 8×8 trong nước.

Tài liệu tham khảo

1. Nguyễn Sĩ Đỉnh, *Hướng dẫn thiết kế đồ án môn học kết cấu tính toán xe thiết giáp (Tập 1: Tính toán thiết kế ly hợp),* Học Viện Kỹ Thuật Quân Sự, Hà Nội, 2015.
2. Nguyễn Văn Luận, Lê Kỳ Nam, *Kết cấu tính toán xe tăng-xe thiết giáp*, Học viện Kỹ thuật Quân sự, Hà Nội,1999.
3. Vũ Đức Lập, Phạm Đình Vi, *Cấu tạo ôtô quân sự (Tập 1*), Học viện kỹ thuật quân sự, Hà Nội, 1995.

EVALUATION OF THE ACCELERATION PERFORMANCE OF AN 8×8 ARMORED PERSONNEL CARRIER

Abstract: During the design and development of an 8×8 armored personnel carrier, the evaluation of longitudinal dynamic performance, particularly acceleration capability, is essential for design validation and refinement. This study presents the theoretical framework for longitudinal vehicle dynamics and develops a quantitative methodology for assessing acceleration performance based on the derived governing equations of motion. The obtained results provide a rigorous scientific foundation for drivetrain design calculations, optimal gear ratio selection, and powertrain configuration optimization prior to full-scale experimental validation and field testing.

Keywords: *Armored personnel carrie; Vehicle dynamics;Acceleration performance.*

Tính toán thiết kế cải tiến hộp phân phối trung tâm xe thiết giáp chở quân bánh lốp 8x8

Vũ Đức Vinh[1*], Cù Xuân Phong[1]

[1]*Học viện Kỹ thuật Quân sự*

Email: kyle0nline90@gmail.com; Tel: 0339576285

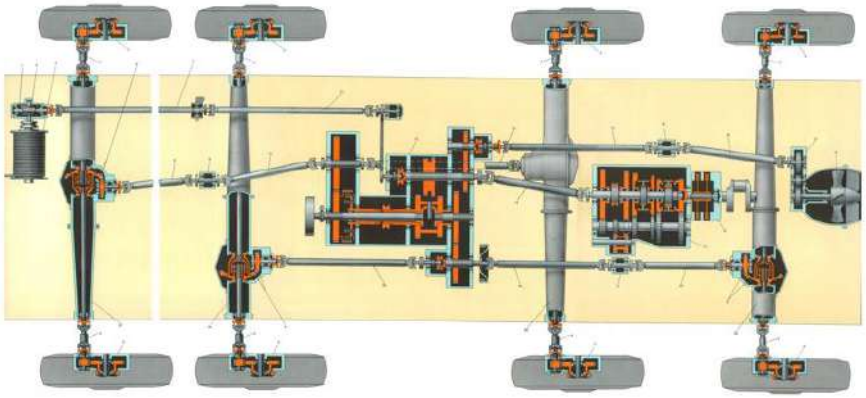
Tóm tắt:

Để có thể bố trí khoang bộ binh ở phía sau thân xe và cửa đổ bộ tại đuôi xe, qua đó nâng cao mức độ an toàn cho bộ binh trong quá trình lên xuống xe, cần cải tiến căn bản phương án bố trí hệ thống động truyền lực của các dòng xe thiết giáp bánh lốp 8×8 thế hệ trước. Theo hướng tiếp cận này, cụm động lực chính gồm động cơ, ly hợp và hộp số chính được chuyển từ phía sau lên phía trước thân xe, phù hợp với xu hướng thiết kế xe thiết giáp hiện đại. Sự thay đổi vị trí cụm động lực kéo theo yêu cầu tính toán, thiết kế lại hệ thống truyền lực, đặc biệt là hộp phân phối trung tâm, nhằm bảo đảm truyền công suất hợp lý tới các cầu chủ động và duy trì các đặc tính động lực học của xe. Trên cơ sở đó, nghiên cứu này đề xuất phương án bổ sung một hộp truyền lực trung gian đặt phía trước hộp phân phối, đóng vai trò điều chỉnh hướng và sơ đồ truyền mô men, đáp ứng yêu cầu bố trí mới của hệ thống động lực. Hộp truyền lực trung gian được thiết kế với ba cặp bánh răng ăn khớp liên tiếp, thực hiện chức năng đảo chiều quay và phân phối mô men theo sơ đồ phù hợp.

Từ khóa: *Xe thiết giáp 8×8; hệ thống truyền lực; hộp số phân phối; hộp truyền lực trung gian, bánh răng;*

1. Đặt vấn đề

Các dòng xe 8x8 trước đây bố trí khoang động lực ở phía sau, cửa bộ binh bên sườn xe, do đó bộ binh khi lên xuống xe dễ bị tổn thương dưới hỏa lực của đối phương khi lên xuống. Hình 1 thể hiện việc bố trí các cụm chi tiết trong hệ thống động truyền lực của xe thiết giáp 8x8 phổ biến trước đây như BTR-80, BTR-82 [1].



Hình 1. Sơ đồ hệ thống truyền lực nguyên bản

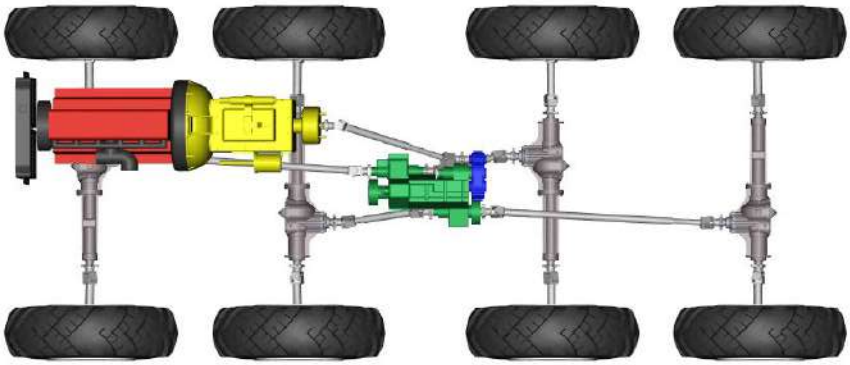
Xu hướng thiết kế hiện nay đưa động cơ lên phía trước thân xe, kết hợp bố trí cửa đổ bộ phía sau đã trở thành một bước chuyển căn bản trong tư duy thiết kế xe thiết giáp chở quân và xe chiến đấu bộ binh bánh lốp 8×8. Cách bố trí này cho phép bộ binh lên xuống xe dưới sự che chắn trực tiếp của thân xe, giảm đáng kể khả năng bị tổn thương trước hỏa lực bộ binh, mảnh pháo và hỏa lực bắn thẳng của đối phương. Đồng thời, khoang động lực phía trước còn đóng vai trò như một lớp giáp bổ sung, nâng cao khả năng sống sót của kíp xe và tổ bộ binh trong trường hợp xe bị bắn từ hướng trước [2,3]

Về mặt bố trí không gian, việc chuyển động cơ lên phía trước giúp tối ưu hóa khoang chiến đấu và khoang chở quân ở phía sau, tạo điều kiện bố trí ghế ngồi chống mìn, hệ thống hấp thụ xung lực, cũng như tăng thể tích hữu dụng cho binh sĩ và trang bị cá nhân. Bên cạnh

đó, sơ đồ này thuận lợi cho việc tích hợp các trạm vũ khí điều khiển từ xa, hệ thống quan sát – chỉ huy hiện đại và nâng cấp cấp độ bảo vệ chống mìn, chống đạn theo các tiêu chuẩn mới.

Xét trên góc độ chiến thuật, bố trí động cơ phía trước – cửa đổ bộ phía sau phù hợp hơn với đặc điểm tác chiến hiện đại, nơi bộ binh cơ giới thường xuyên phải đổ bộ nhanh trong điều kiện hỏa lực đối phương phức tạp, tác chiến trong đô thị hoặc khu vực địa hình hạn chế. Chính vì vậy, xu hướng này hiện đang được áp dụng rộng rãi trên các thế hệ xe 8×8 mới, phản ánh sự chuyển dịch từ yêu cầu cơ động thuần túy sang ưu tiên khả năng bảo vệ, sống sót và hiệu quả tác chiến hiệp đồng bộ binh – xe thiết giáp.

Bài báo này đề xuất phương án cải tiến hộp số phân phối trung tâm bằng việc thiết kế thêm hộp truyền lực trung gian kết nối giữa hộp số và hộp số phân phối (hình 2).



Hình 2. Sơ đồ hệ thống truyền lực sau cải tiến

2. Tính toán thiết kế và kiểm bền hộp truyền lực

Các thông số cơ bản của xe thiết giáp bánh lốp 8x8 thể hiện trên bảng 1.

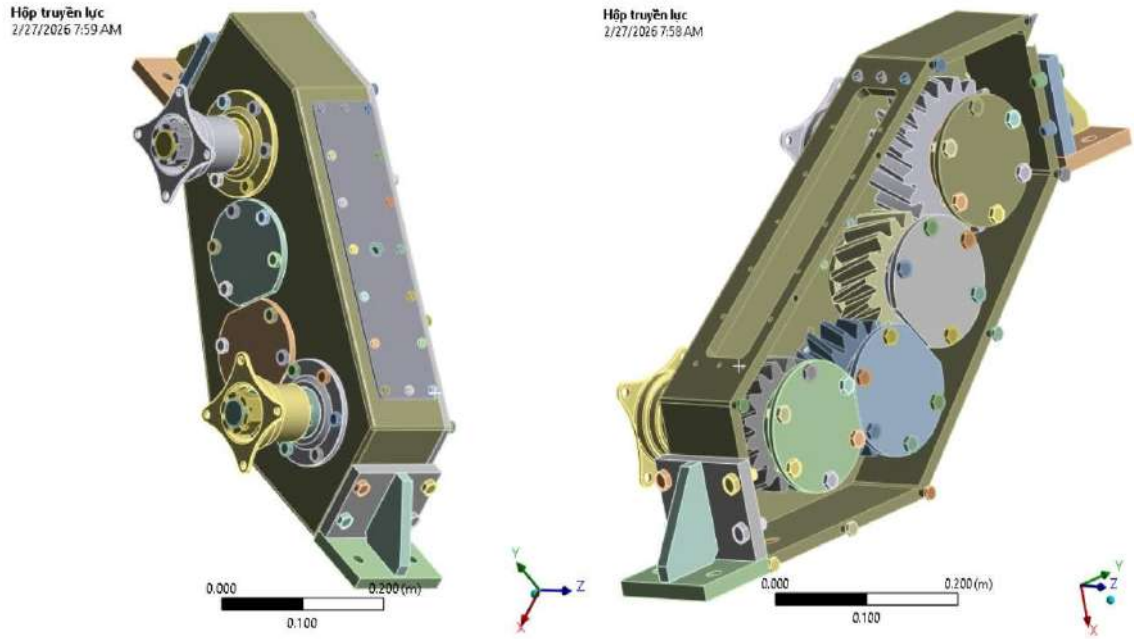
Bảng 1. Thông số cơ bản của xe thiết giáp bánh lốp 8x8

TT	Số liệu đầu vào	Giá trị	Đơn vị
1	Trọng lượng toàn bộ xe G_a	16,55	Tấn
2	Động cơ		
	Số vòng quay tại đó mô men lớn nhất n_{eM}	1400	vg/ph
	Mô men lớn nhất M_{emax}	1532	Nm
	Số vòng quay tại đó công suất lớn nhất n_{eN}	1800	vg/ph
	Công suất lớn nhất N_{emax}	258	kW
3	Vận tốc lớn nhất trên đường bằng	85	km/h
4	Tỉ số truyền của hộp số i_{hs}	i_1=9,48; i_2=6,58; i_3=4,68; i_4=3,48; i_5=2,62; i_6=1,89; i_7=1,35; i_8=1,00; i_9=0,75; i_L=8,97;	
5	Tỷ số truyền của hộp số phân phối + ST thấp + ST cao	 1,39 0,76	
6	Tỷ số truyền của truyền lực chính i_{TLC}	1,846	
7	Tỷ số truyền của giảm tốc cạnh i_{GTC}	4,33	

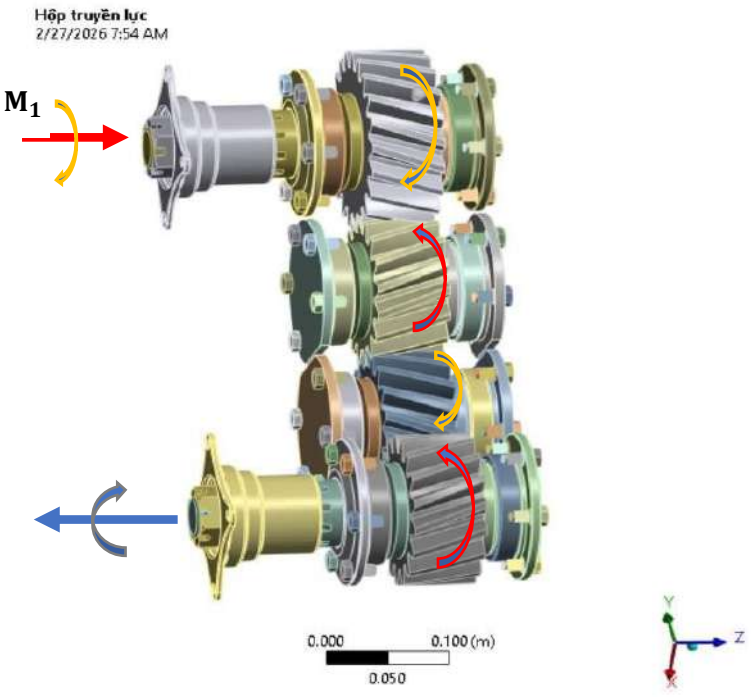
Hộp truyền lực bổ sung có nhiệm vụ đảo chiều truyền mô men khi thay đổi vị trí cụm động lực và điều chỉnh tỷ số truyền theo yêu cầu bố trí tuyến truyền lực. Để đảm bảo làm việc êm, giảm ồn và tăng khả năng tải, vì vậy chọn bánh răng trụ răng nghiêng.

Xét yêu cầu đảo chiều, hướng quay trục ra so với trục vào phụ thuộc vào số lần ăn khớp. Cụ thể, nếu số cặp ăn khớp là lẻ thì trục ra quay ngược chiều trục vào. Do đó, lựa chọn kết cấu gồm 4 trục song song mang 4 bánh răng, tạo thành chuỗi truyền qua 3 cặp ăn khớp liên tiếp (1–2, 2–3, 3–4). Với 3 cặp ăn khớp (số lẻ), hộp truyền lực đảm bảo đảo chiều đúng yêu cầu.

Ngoài ra, kết cấu 4 trục – 4 bánh răng cho phép bố trí 2 bánh răng trung gian (trên trục II và III), giúp thuận lợi điều chỉnh tỷ số truyền tổng bằng cách lựa chọn số răng bánh vào và bánh ra theo không gian lắp đặt, đồng thời phân bố tải hợp lý, tăng độ êm so với phương án răng thẳng.



Hình 3. Hộp truyền lực trong thiết kế



Hình 4. Sơ đồ truyền lực, cụm bánh răng và trục trong thiết kế

Số răng các bánh răng lần lượt là z_1, z_2, z_3, z_4, với chuỗi 3 cặp ăn khớp:

Tỷ số truyền toàn hộp [4]:

$$i_{HTL} = \frac{z_2}{z_1} \cdot \frac{z_3}{z_2} \cdot \frac{z_4}{z_3} = \frac{z_4}{z_1}$$

Như vậy, tỷ số truyền hộp truyền lực chỉ phụ thuộc vào bánh vào $z_1$và bánh ra z_4; các bánh trung gian z_2, z_3 không ảnh hưởng đến tỷ số truyền (chủ yếu dùng để đảo chiều và đáp ứng bố trí khoảng cách trục).

Theo yêu cầu thiết kế để đảm bảo vận tốc lớn nhất của xe đạt 85 km/h, tỉ số truyền của hộp truyền lực trung gian bằng 0,8. Trên cơ sở số răng theo mô hình kết cấu, chọn:

$$z_1 = 21; z_2 = z_3 = z_4 = 17$$

Khoảng cách giữa các trục được tính như sau:

$$m_t = \frac{m_n}{\cos\beta} = \frac{10}{\cos 20°} = \frac{10}{0{,}9397} = 10{,}64 \text{ mm}$$

Đường kính vòng chia: $d_i = m_t z_i$

Từ đó $d_1 = 10{,}64 \cdot 21 = 223{,}4$ mm $= 0{,}2234$ m

$d_2 = d_3 = d_4 = 10{,}64 \cdot 17 = 180{,}9$ mm $= 0{,}1809$ m

Như vậy khoảng cách giữa các trục lần lượt là:

$$a_{12} = \frac{223{,}4 + 180{,}9}{2} = 202{,}15 \text{ mm}$$

$$a_{23} = a_{34} = \frac{180{,}9 + 180{,}9}{2} = 180{,}9 \text{ mm}$$

Mô men lớn nhất vào trục chủ động hộp truyền lực (hộp truyền lực đặt sau hộp số chính, xét tay số 1:

$$M_{1\,max} = M_{e\,max}\, i_1\, \eta_1 \qquad (4)$$

Thay số thu được:

$$M_{1\,max} = 10349 \text{ (Nm)}$$

Mô men trên các trục vào là:

$$T_1 = M_{1\,max} = 10348{,}8 \text{ Nm}$$

Vì $i_{HTL} < 1$nên trục vào chịu mô men lớn nhất T_1 = 10349 Nm. Chọn đường kính chung theo trục nguy hiểm nhất. Chọn $[\tau] = 55$ MPa

Từ điều kiện bền xoắn [5]:

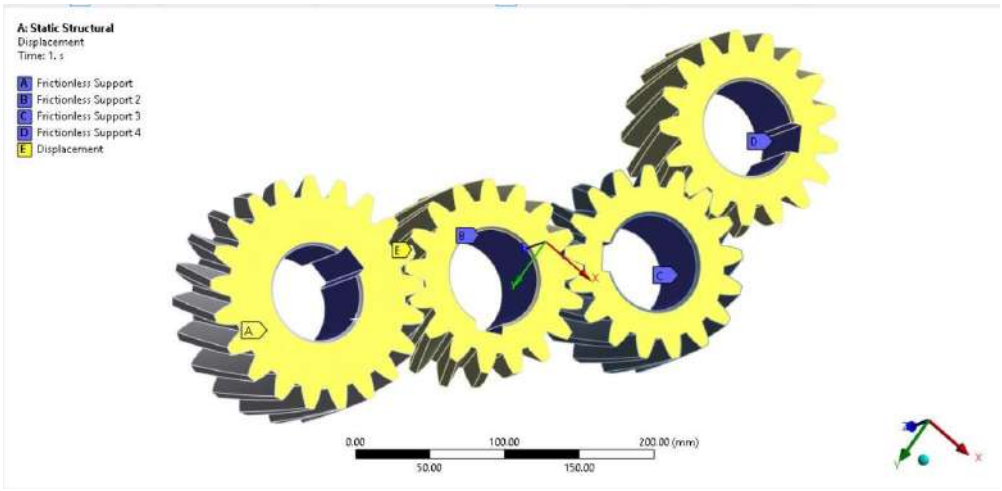
$$\tau = \frac{16M}{\pi d^3} \le [\tau] \Rightarrow d \ge \sqrt[3]{\frac{16M}{\pi[\tau]}} \qquad (8)$$

$$\Rightarrow d \ge \sqrt[3]{\frac{16 \cdot 10\,348\,800}{\pi \cdot 55}} = 98{,}9 \text{ mm}$$

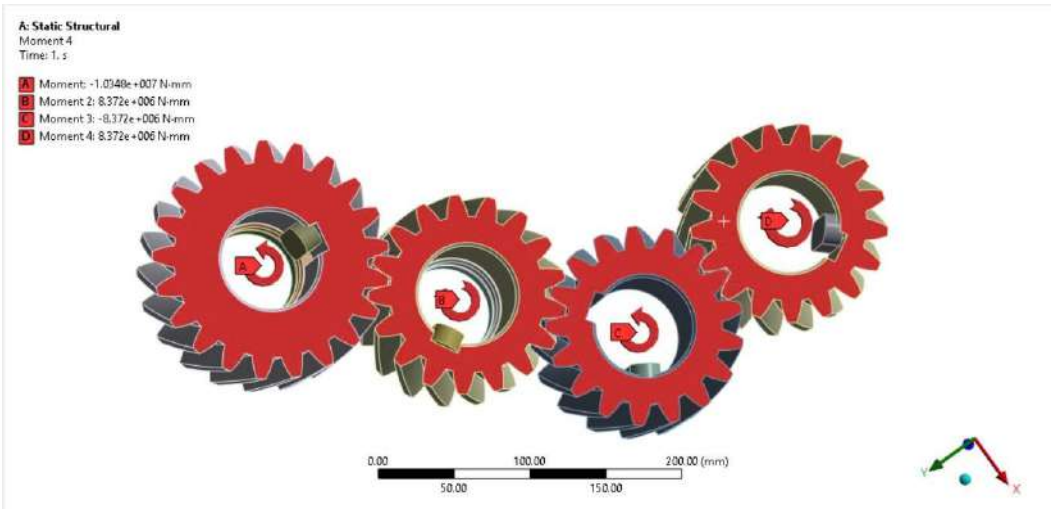
Từ đó chọn đường kính trục bằng nhau cho 4 trục:

$$d_{tr1} = d_{tr2} = d_{tr3} = d_{tr4} = 110 \text{ mm}$$

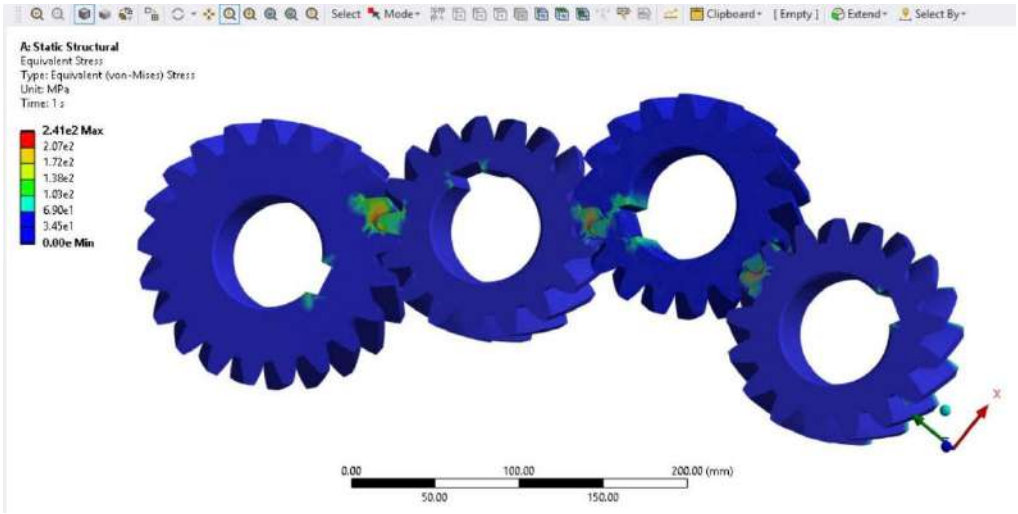
Thứ tự và kết quả kiểm bền bằng ANSYS được thể hiện trên hình 5, 6, 7. Kết quả cho thấy các bánh răng thiết kế đảm bảo bền



Hình 5. Cố định các vật thể



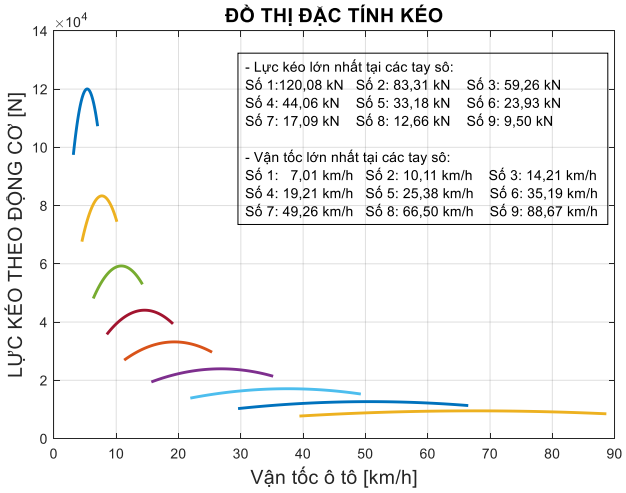
Hình 6. Đặt mô men tác dụng



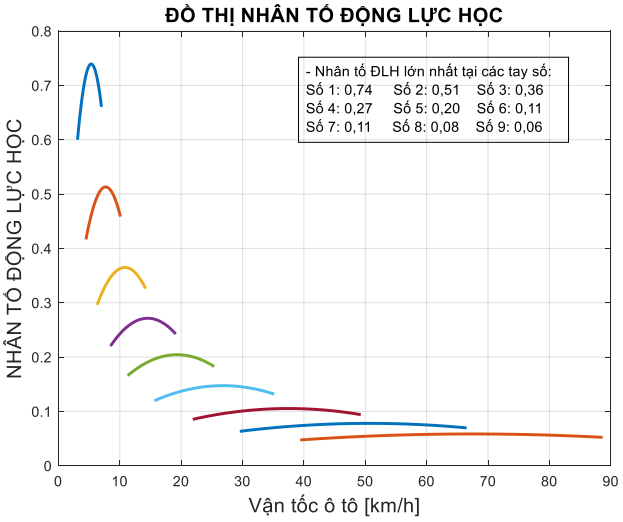
Hình 7. Kết quả kiểm bền

3. Khảo sát động lực học chuyển động thẳng

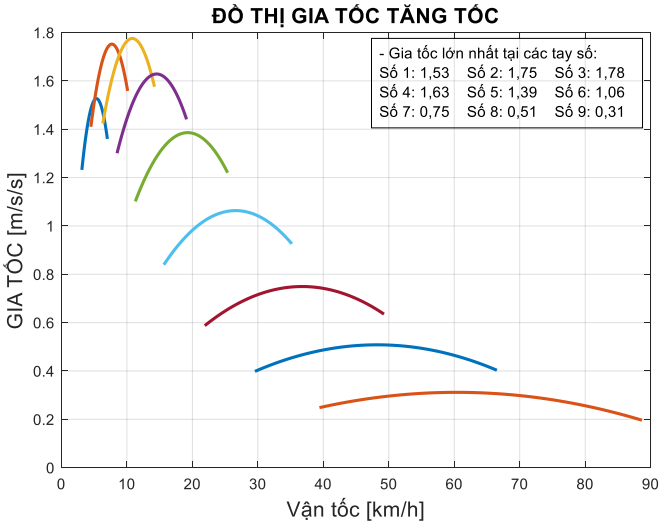
Với hộp số phân phối cải tiến, nghiên cứu tiến hành khảo sát động lực học chuyển động của xe. Kết quả thu được thể hiện trên hình 8, 9, 10, 11, 12.



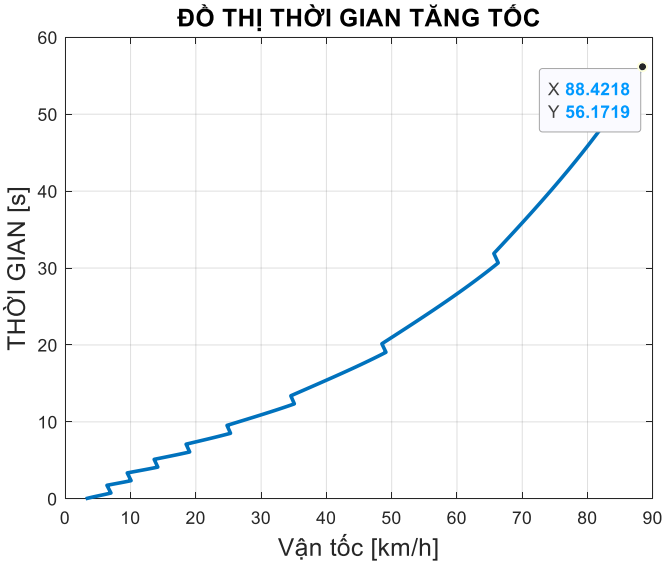
Hình 8. Đồ thị đặc tính kéo từng tay số



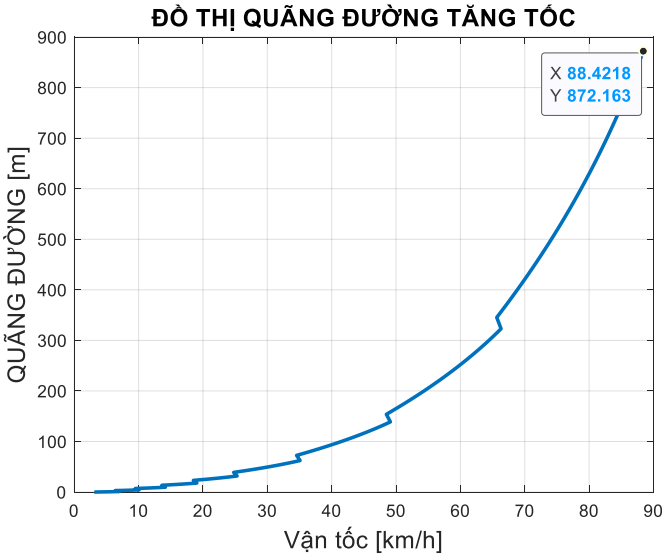
Hình 9. Đồ thị nhân tố động lực học D



Hình 10. Đồ thị gia tốc



Hình 11. Đồ thị thời gian tăng tốc



Hình 12. Đồ thị quãng đường tăng tốc

Bảng 2. Kết quả khảo sát động lực học chuyển động thẳng

TT	Thông số	Giá trị
1	Vận tốc lớn nhất mà xe có thể đạt được (trên đường có hệ số cản 0,018), v_{max} *(km/h)*	88,4
2	Lực kéo lớn nhất của ô tô, P_{kmax} *(kN)*	120
3	Nhân tố động lực học lớn nhất, D_{max}	0,74
4	Gia tốc lớn nhất của ô tô, J_{max} *(m/s²)*	1,63
5	Thời gian tăng tốc đến v_{max}, t_{tt} *(s)*	56
6	Quãng đường tăng tốc đến v_{max}, S_{tt} *(m)*	872
7	Góc dốc lớn nhất xe vượt được, độ	32

Kết quả khảo sát cho thấy với hộp số phân phối cải tiến, xe đáp ứng đầy đủ các yêu cầu về tính năng động lực học và khả năng cơ động trong điều kiện khai thác thực tế. Cụ thể, vận tốc lớn nhất đạt được trên đường có hệ số cản 0,018 là 88,4 km/h, phù hợp với yêu cầu cơ động chiến dịch của xe thiết giáp chở quân bánh lốp 8×8. Lực kéo lớn nhất đạt 120 kN, bảo đảm dự trữ lực kéo cần thiết cho xe khi vượt chướng ngại, cơ động trên địa hình phức tạp và trong điều kiện tải trọng lớn. Nhân tố động lực học lớn nhất bằng 0,74 cho thấy khả năng tận dụng công suất động cơ thông qua hệ thống truyền lực là tương đối cao, phản ánh hiệu quả của phương án cải tiến hộp số phân phối trong việc phân phối và truyền mô men tới các cầu chủ động. Gia tốc lớn nhất của xe đạt 1,63 m/s², bảo đảm khả năng tăng tốc cần thiết trong các tình huống tác chiến cơ động nhanh. Thời gian tăng tốc đến vận tốc lớn nhất là 56 s với quãng đường tăng tốc 872 m, nằm trong giới hạn cho phép đối với xe thiết giáp hạng trung, đồng thời không gây quá tải cho các cụm chi tiết của hệ thống truyền lực. Đặc biệt, xe có khả năng vượt dốc lớn nhất tới 32°, chứng tỏ hộp số phân phối cải tiến đáp ứng tốt yêu cầu truyền lực ở chế độ tải nặng và tốc độ thấp, phù hợp với điều kiện địa hình đồi núi, đường đất và địa hình phức tạp. Nhìn chung, các chỉ tiêu tính năng thu được từ bảng số liệu cho thấy phương án cải tiến hộp số phân phối trung tâm là hợp lý, không làm suy giảm các đặc tính động lực học cơ bản của xe, đồng thời tạo tiền đề thuận lợi cho việc thay đổi bố trí hệ thống động lực theo hướng hiện đại, nâng cao mức độ an toàn và hiệu quả sử dụng của xe thiết giáp chở quân bánh lốp 8×8.

8. Kết luận

Bài báo đã nghiên cứu phương án cải tiến hệ thống động truyền lực cho xe thiết giáp chở quân bánh lốp 8×8 theo hướng chuyển cụm động lực từ phía sau lên phía trước thân xe, nhằm tạo điều kiện bố trí khoang bộ binh ở phía sau và cửa đổ bộ tại đuôi xe, qua đó nâng cao mức độ an toàn cho bộ binh khi lên xuống xe và phù hợp với xu hướng thiết kế xe thiết giáp hiện đại. Trên cơ sở phân tích đặc điểm bố trí của các dòng xe 8×8 thế hệ trước và các yêu cầu kỹ thuật đặt ra, nghiên cứu đã đề xuất phương án bổ sung hộp truyền lực trung gian đặt phía trước hộp số phân phối trung tâm, nhằm bảo đảm sơ đồ truyền mô men hợp lý trong điều kiện bố trí mới của hệ thống động lực. Hộp truyền lực trung gian được thiết kế với ba cặp bánh răng ăn khớp liên tiếp, thực hiện chức năng đảo chiều quay và phân phối mô men theo yêu cầu.

Các thông số hình học chủ yếu, mô men và lực tác dụng trên các trục, cũng như kiểm nghiệm độ bền uốn, độ bền tiếp xúc của bánh răng và độ bền xoắn của trục. Kết quả tính toán cho thấy các chi tiết chính của hộp truyền lực và hộp số phân phối cải tiến đều thỏa mãn điều kiện bền, có hệ số an toàn phù hợp, bảo đảm khả năng làm việc tin cậy trong điều kiện tải trọng đặc trưng của xe thiết giáp chở quân. Kết quả khảo sát tính năng động lực học cho thấy với hộp số phân phối cải tiến, xe đạt vận tốc lớn nhất 88,4 km/h, lực kéo cực đại 120 kN, gia tốc lớn nhất 1,63 m/s², và khả năng vượt dốc tới 32°, đáp ứng đầy đủ các yêu cầu về cơ động, vượt địa hình và khai thác sử dụng trong điều kiện tác chiến.

Tài liệu tham khảo

1. Общее устройство и техническое обслуживание бтр-80 - 2003

2. БТР-82А, Общее устройство и техническое обслуживание, 2010

3. Nguyễn Văn Trà, *Kết cấu tính toán ô tô-Tập 1*, Học viện Kỹ thuật Quân sự - NXB QĐND 2016

4. Vũ Đức Lập, *Kết cấu tính toán ô tô, Tập 2*, Học viện Kỹ thuật Quân sự -NXB QĐND, 2015.

5. Nguyễn Phúc Hiểu, Vũ Đức Lập, *Lý thuyết ô tô quân sự*, Học viện Kỹ thuật Quân sự - NXB QĐND 2002

Design Calculation and Improvement of the Central Transfer Case for an 8×8 Wheeled Armored Personnel Carrier

Abstract:

To enable the arrangement of the troop compartment at the rear of the vehicle hull and the troop access door at the rear, thereby enhancing the safety of dismounted infantry during embarking and disembarking, it is necessary to fundamentally improve the layout of the powertrain system of earlier-generation 8×8 wheeled armored vehicles. According to this approach, the main power unit, including the engine, clutch, and main gearbox, is relocated from the rear to the front of the vehicle hull, in line with modern armored vehicle design trends. This relocation of the power unit entails the need to recalculate and redesign the drivetrain system, particularly the central transfer case, in order to ensure proper power transmission to the driven axles while maintaining the vehicle's dynamic performance characteristics. On this basis, the present study proposes the addition of an intermediate transmission unit installed in front of the transfer case, which serves to adjust the direction and layout of torque transmission to meet the new powertrain configuration requirements. The intermediate transmission unit is designed with three consecutive gear meshes, performing the functions of reversing the direction of rotation and distributing torque according to an appropriate transmission scheme.

Keywords: *8×8 armored vehicle; Driveline system; Reversing gearbox; Helical gear transmission; Torque reversal; Strength verification.*

Khảo sát động lực học bơi xe thiết giáp chở quân có công thức bánh xe 8x8

Nguyễn Thanh Quang[1*], Nguyễn Minh Tân[1], Cù Xuân Phong[1]

[1]*Học viện Kỹ thuật Quân sự*

Email: Nguyenthanhquang14071984@gmail.com

Tóm tắt:

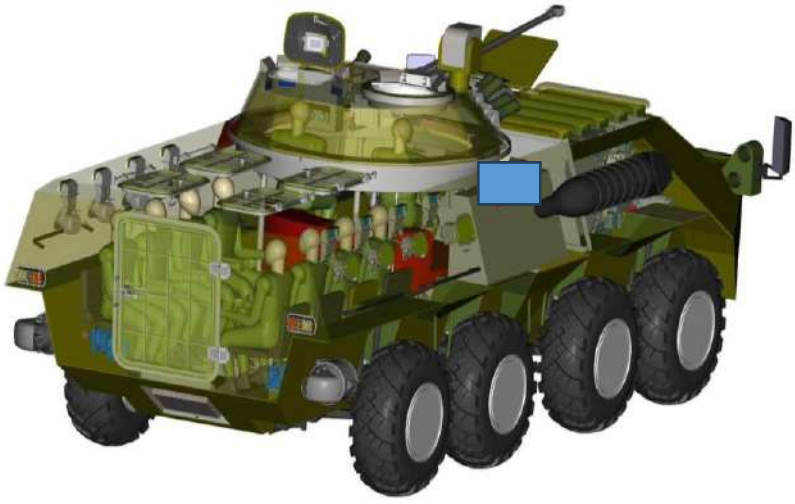
Hiện nay, Quân đội ta đang tích cực triển khai các chương trình nghiên cứu, thiết kế, chế tạo mới các dòng xe thiết giáp nhằm đáp ứng được yêu cầu huấn luyện và sẵn sàng chiến đấu, trong đó có xe thiết giáp chở quân 8x8. Một trong những tính năng quan trọng của dòng xe này là khả năng bơi nước ổn định với cơ cấu bơi dạng bơm phụt được dẫn động bằng thủy lực. Trong bài báo này, nhóm tác giả trình bày cơ sở lý thuyết, phương pháp xây dựng mô hình khảo sát và tính toán kéo chuyển động thẳng khi bơi cho xe thiết giáp chở quân 8x8 chế tạo tại Việt Nam. Các tham số động lực học được khảo sát bao gồm: lực cản của nước, lực kéo do cơ cấu bơi tạo ra phụ thuộc vào vận tốc bơi của xe.

Từ khóa: *Xe thiết giáp chở quân 8x8; cơ cấu bơi; vận tốc bơi; lực cản của nước.*

1. Đặt vấn đề

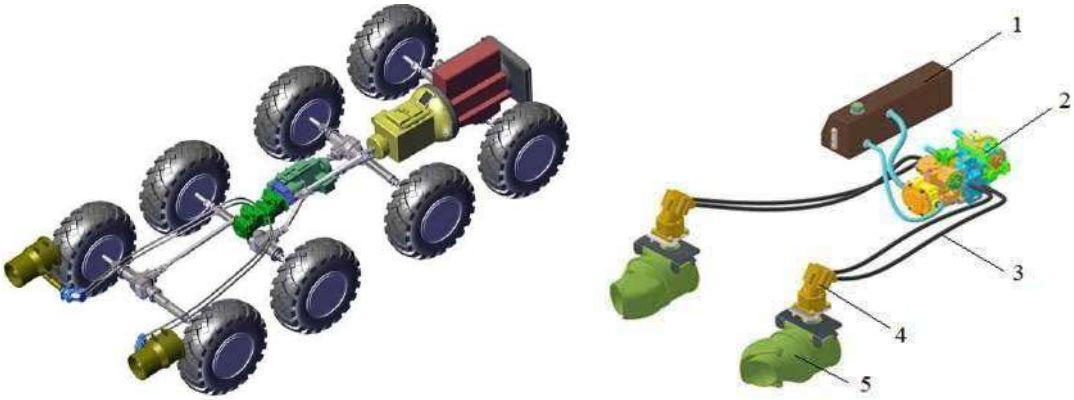
Xe thiết giáp chở quân có trong biên chế của Quân đội ta hiện nay phần lớn là do Nga viện trợ hoặc xe chiến lợi phẩm của Mỹ và đều được sản xuất từ những năm 1960 trở về trước. Do nhiều nguyên nhân khác nhau mà tình trạng kỹ thuật của các dòng xe này đã bị suy giảm đáng kể, rất nhiều cụm, hệ thống trên xe hư hỏng không thể phục hồi hoặc đã hết kích thước sửa chữa, khó có khả năng đáp ứng các yêu cầu huấn luyện và sẵn sàng chiến đấu trong giai đoạn mới [1].

Hiện nay, các đơn vị của Tổng cục CNQP, tổng cục HC-KT phối hợp với Học viện KTQS đang tích cực triển khai nhiệm vụ "Nghiên cứu thiết kế, chế tạo xe thiết giáp chở quân 8x8 XTC-03" bằng công nghệ trong nước (Hình 1).



Hình 1. Bố trí chung xe thiết giáp 8x8 sản xuất tại Việt Nam

Nhóm nghiên cứu của Viện Cơ khí động lực/Học viện KTQS được giao nhiệm vụ chủ trì thiết kế hệ thống và thiết kế, chế tạo hệ thống bơi cho xe XTC-03. Thông qua các nghiên cứu, tính toán lý thuyết và khảo sát thực tế nhóm đã lựa chọn cơ cấu bơi dạng bơm luồng phụt được dẫn động bằng thủy lực cho xe XTC-03. Bơm chính của hệ thống bơi được trích công suất từ động cơ đốt trong tại vị trí trục nối giữa ly hợp chính và hộp số, hai bơm luồng phụt được lắp phía đuôi xe (hình 2).



Hình 2. Sơ đồ bố trí hệ thống bơi trên xe thiết giáp XTC-03

1. Thùng dầu; 2. Bơm chính; 3. Các đường ống; 4. Mô tơ thủy lực; 5. Bơm luồng phụt

Cơ cấu bơi sau khi thiết kế, chế tạo và lắp đặt lên xe cần phải có những tính toán, khảo sát động lực học để đánh giá các chỉ tiêu về lực kéo do cơ cấu bơi tạo ra, lực cản khi bơi và tốc độ của xe trong các điều kiện bơi khác nhau. Thông qua đồ thị lực kéo P =P(v) và lực cản R=R(v) xác định được các yếu tố quan trọng của động lực học xe bơi như vận tốc bơi lớn nhất Vmax, lực kéo moóc lớn nhất Pmk, khả năng tăng tốc và lực kéo lớn nhất [2].

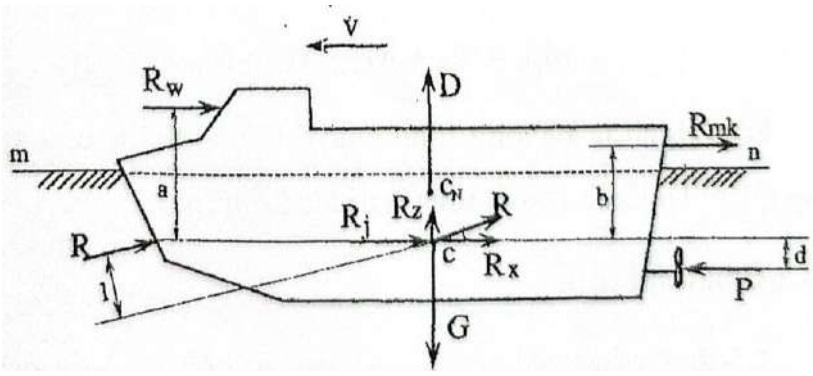
Đã có nhiều công trình nghiên cứu về động lực học của xe bơi, phương tiện lội nước trong các điều kiện khai thác khác nhau [3,4] và đánh giá các yếu tố ảnh hưởng tới động lực học của xe [5, 6]. Các công trình này đã chỉ ra khả năng di chuyển của xe lội nước phụ thuộc vào nhiều yếu tố như lực cản của dòng nước, hình dáng của xe, công suất của động cơ, hiệu suất của cơ cấu bơi …[7]. Trong nước, một số tác giả cũng đã nghiên cứu về động lực học chuyển động thẳng và quay vòng khi bơi, tuy nhiên các nghiên cứu này tập trung chủ yếu vào đối tượng là các phương tiện công binh vượt sông và các xe tăng thiết giáp xích [8, 9]. Chưa có nhiều các nghiên cứu chuyên sâu về động lực học bơi cho xe thiết giáp bánh lốp với cơ cấu bơi kiểu bơm luồng phụt.

2. Cơ sở lý thuyết khảo sát động lực học bơi xe thiết giáp

Cơ sở lý thuyết bơi của xe dựa trên nguyên lý tàu thủy mạn nhỏ có xem xét đến các đặc trưng của xe bơi. Xe bơi thường có hình dạng khác với dạng thủy khí động lực học chuẩn nên hệ số cản dưới nước thường rất lớn, lực đẩy nổi của xe bơi là lực thủy tĩnh, hiệu suất của dẫn động xe bơi thấp [2,7].

2.1. Mô hình khảo sát và các lực tác dụng lên xe bơi

Trong trường hợp tổng quát, khi chuyển động thẳng dưới nước xe chịu tác động của các lực và mô men như trên sơ đồ hình 3.



Hình 3. Sơ đồ các lực và mô men tác dụng lên xe bơi trong trường hợp tổng quát

Trong đó:

G- Trọng lực đặt tại trọng tâm của xe;

D- Lực đẩy (áp lực thủy tĩnh) đặt tại tâm choán nước;

P- Lực kéo (do cơ cấu bơi tạo ra);

R- Lực cản của nước, phân thành hai thành phần lực cản chuyển động R_x, và lực nâng R_z;

R_w- Lực cản không khí đặt tại trọng tâm phần nổi;

R_j- Lực quán tính đặt tại trọng tâm xe C; R_{mk}: Lực cản móc kéo.

Khi dời các lực về trọng tâm xe ta có các mô men tác dụng như sau:

- Mô men của lực cản R: $M_R = R.l$
- Mô men lực cản không khí: $M_w = R_w.a.$
- Mô men của lực moóc kéo: $M_{mk} = R_{mk}.b$

Khi lực kéo dẫn tiến bơi cân bằng với tổng các lực chuyển động, có phương trình cân bằng:

$$P = R_x + R_w + R_j + R_{mk} \quad (1)$$

Phương trình cần bằng mô men có dạng:

$$M_{ph} = M_r + M_w + M_{mk} \quad (2)$$

Mục tiếp theo sẽ lần lượt xem xét từng lực kể trên.

2.2. Xác định các thành phần lực cản

a) Lực cản của nước

Trong các thành phần lực cản, lực cản của nước là thành phần chủ yếu. Đối với xe bơi, theo thủy lực học, lực cản của nước có thể phân thành 3 thành phần: Lực cản ma sát R_{ms}; Lực cản hình dáng R_{hd}; Lực cản sóng R_s [2].

Với các xe bơi có tốc độ chuyển động nhỏ và có hình dáng phức tạp do kết cấu của phần vận hành vì vậy lực cản R_{ms} và R_{hd} chiếm phần lớn thành phần lực cản (lớn hơn 80%).

Các kết quả nghiên cứu lý thuyết và thực nghiệm, các nhà khoa học đã đưa ra công thức thực nghiệm cho lực cản của nước như sau [7]:

$$R = K_{vh} \cdot \rho . F_{cd} \cdot v^2 \quad (3)$$

Trong đó: ρ- Khối lượng riêng của nước; F_{cd}- Diện tích cản chính diện; v- Vận tốc chuyển động tương đối của xe so với nước; K_{vh}- Hệ số phụ thuộc vào kết cấu phần vận hành, với xe XTC-03 có bánh - hệ treo độc lập: K_{vh}= 0.4-0.5

Công thức (3) đúng trong trường hợp v ≤3,6m/s. Khi v >3,6 m/s tính theo công thức:

$$R = K_{vh} \cdot \rho . F_{cd} \cdot v^{2+0.2-\frac{v}{\sqrt{gL}}} \quad (4)$$

Khi xét ảnh hưởng của độ sâu đến lực cản của nước, R được xác định theo công thức (5).

$$R = c.\left\{0{,}41.\exp\left(\frac{T}{h}\right) - 0{,}032.\left[\exp\left(-\frac{T}{h}\right).\ln\left(\frac{h}{T}\right)\right]\right\}.\rho.F_{cd} \cdot v^{2+0.2-\frac{v}{\sqrt{gL}}} \quad (5)$$

Trong đó: T- Chiều sâu mớn nước khi cân bằng tĩnh; h- Độ sâu của nước; hệ số c=0,42÷0,45 với xe có bánh treo độc lập.

Lưu ý: khi xét đến ảnh hưởng của tốc độ dòng chảy đến lực cản của nước, trong tất cả các công thức tính toán động lực học của xe bơi, tốc độ v là tốc độ tương đối của xe với nước, tốc độ tuyệt đối của xe bơi với bờ bằng tổng của tốc độ tương đối với tốc độ dòng chảy.

$$\vec{V} = \vec{V_s} + \vec{V_{td}} \quad (6)$$

Với: $\vec{V}_s$- Véc tơ tốc dộ dòng chảy; $\vec{V}_{td}$ - Véc tơ tốc dộ tương đối của xe bơi với nước.

b) Lực cản không khí (R_w)

Khi xe chuyển động dưới nước phần nhô lên của xe sẽ chịu tác động lực cản không khí R_w Thông thường R_w có giá trị rất nhỏ do xe bơi có tốc độ chuyển động thấp nên khi tính toán có thể bỏ qua.

Giá trị của lực cản không khí theo phương chuyển động của xe được xác định:

$$R_{wx} = K_{wx} F_{H.\gamma} v_w^2 \cos^2 \alpha_w \quad (7)$$

Trong đó: v_W- Vận tốc của gió trong trường hợp xe coi là đứng yên; α_w - Góc tạo bởi véc tơ lực cản của gió và phương chuyển động của xe; $F_{H.\gamma}$ - Diện tích mặt cắt ngang phần nổi trên mặt nước của xe; K_{wx} - Hệ số cản phụ thuộc vào hình dáng xe, mật độ không khí, thông thường K_{wx}=0.7 ÷ 0.75 (kg/m^3).

c) Lực quán tính

Khi khảo sát quá trình tăng tốc của xe bơi nước. Lực quán tính R_j của xe được xác định bằng 2 giai đoạn [2]:

- Giai đoạn tăng tốc:

$$R_{j1} = (m_x + \lambda_1 + m_{tl}) \frac{dV}{dt} \quad (8)$$

- Giai đoạn ổn định:

$$R_{j1} = (m_x + \lambda_1) \frac{dV}{dt} \quad (9)$$

Trong đó: m_x- Khối lượng của xe; m_{tl}- Hệ số khối lượng quay của hệ thống truyền lực trong quá trình tăng tốc; V- Vận tốc của xe theo phương chuyển động; λ_1 - Khối lượng nước kéo theo xe ($\lambda_1 = \frac{m_x T}{2L}$); L – chiều dài xe bơi (m).

Thông thường trong quá trình khảo sát động lực học của xe bơi, để đơn giản, người ta thường bỏ qua ảnh hưởng quán tính của hệ truyền lực. Như vậy lực quán tính sẽ bằng:

$$R_j = (1 + \frac{T}{2L}) m_x \frac{dV}{dt} \quad (10)$$

2.3. Xác định lực kéo của xe do cơ cấu bơi tạo ra

Để xe có thể chuyển động được dưới nước cần phải có cơ cấu bơi tạo ra lực kéo P để khắc phục các lực cản ở trên. Trục bơm phụt nhận công suất của mô tơ thủy lực và trục truyền làm cho cánh quay, tạo lực nâng trên cánh và lực đẩy.

Đối với xe thiết giáp lốp 8x8 (XTC-03) cơ cấu bơi có 02 bơm luồng phụt, nên lực đẩy tổng cộng do hai bơm phụt tạo ra tác dụng lên xe được xác định theo công thức sau [7]:

$$P = 2 * \rho . q_v . n_v . (\frac{q_v . n_v}{F_i} - V_0) \quad (11)$$

Trong đó: V_0 – Vận tốc dòng nước phía trước miệng hút; F_i – Diện tích thiết diện của miệng phụt; ρ – Khối lượng riêng của chất lỏng (kg/m^3); Q- Lưu lượng qua ống phụt của bơm: $Q = q_v n_v$; q_v - Lưu lượng riêng của bơm phụt; n_v - Tốc độ quay của tuabin bơm phụt.

Biểu thức (11) dùng để tính toán lực kéo của dẫn tiến bơi kiểu bơm phụt được giả thiết trong điều kiện không tính tới sự ảnh hưởng của dòng chảy xung quanh thân xe bị kéo theo khi xe bơi. Yếu tố này được xác định bằng việc đưa thêm hệ số tương tác thực nghiệm: Hệ số dòng chảy theo W= 0,15 ÷ 0,25 [7]. Khi đó, lực kéo có ích như sau:

$$P_c = 2^*\rho.q_v.n_v.\left(\frac{q_v.n_v}{F_i} - V.(1-W)\right) \quad (12)$$

Với: V - Vận tốc bơi của xe (m/s).

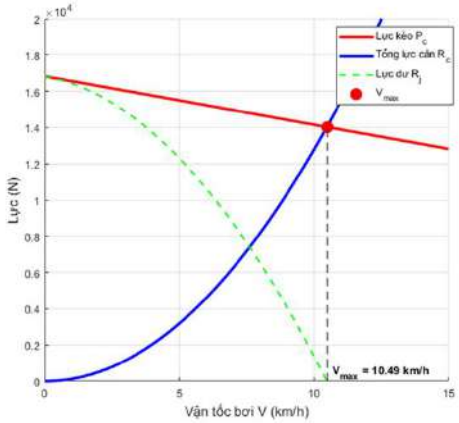
3. Kết quả tính toán và thảo luận

3.1. Bộ thông số tính toán

Bảng 1. Bảng số liệu đầu vào tính toán thủy động lực học.

TT	Thông số	Ký hiệu	Đơn vị	Giá trị
1	Khối lượng riêng của nước	ρ	kg/m^3	1000
2	Tốc độ quay của trục bơm phụt	n_v	Vg/phút	800÷2000
3	Hiệu suất truyền động thủy lực	η_{TL}		0,8
4	Lưu lượng riêng của bơm phụt	q_v	m^3/vòng	0,0185
5	Diện tích cửa phụt ra	F_i	m^2	0,0452
6	Hệ số dòng chảy theo xe	W		0,22
7	Hệ số hút	T_1		0,22
8	Chiều sâu mớm nước trung bình xe	T	m	1,12
9	Chiều dài xe bơi	L	m	7,880
10	Chiều rộng thân xe	B	m	2,985
11	Chiều cao thân xe	H	m	1,725
12	Chiều cao tháp pháo	H1	m	0,43
13	Khoảng sáng gầm xe	H2	m	0,475
14	Khối lượng của xe	m_x	kg	16.5
15	Hệ số cản gió (phụ thuộc vào hình dáng của xe)	KW_x		0,7
16	Diện tích mặt cắt ngang phần nổi trên mặt nước của xe	$F_{H\gamma}$	m^2	3,95
17	Vận tốc của gió (trong điều kiện gió nhẹ)	v_B	m/s	2
18	Góc tạo bởi véc tơ vận tốc xe $\vec{V}$ và vận tốc gió $\overrightarrow{V_B}$ (góc gió cản: 120^0)	α_B	rad	2π/3
19	Hệ số cản chính diện phần ngập nước	K_{VH}		0,45
20	Diện tích cản chính diện phần ngập nước của xe	F_{cd}	m^2	3,68
21	Hệ số tương tác của bơm phụt với thân xe về động lượng theo I.M. Konovalov	α		0,85
22	Hệ số tương tác giữa bơm phụt với thân xe phụ thuộc hình dáng, kích thước phần chìm dưới nước	β		0,80
23	Tổng hệ số tổn hao thủy lực trong cả chiều dài của ống bơm luồng phụt.	k		1,15

3.2. Kết quả tính toán



Hình 4. Kết quả tính toán động lực học bơi xe XTC-03

Nhận xét:

- Lực kéo có ích do hai bơm phụt tạo ra P_c có xu hướng giảm dần khi vận tốc bơi của xe V tăng lên. Điều này phản ánh thực tế là khi xe bơi nhanh hơn, sự chênh lệch vận tốc giữa dòng tia phụt ra và vận tốc dòng vào giảm đi, dẫn đến lực đẩy giảm.

- Tổng lực cản R_c tác động lên xe bơi bao gồm lực cản nước và lực cản gió. Đường này có dạng hàm bậc hai tăng rất nhanh theo vận tốc, cho thấy trở lực của nước là yếu tố chính hạn chế tốc độ bơi.

- Lực dư/Lực quán tính (R_j): Được biểu diễn bằng đường đứt nét là hiệu số giữa lực kéo và lực cản ($R_j = P_c – R_c$). Lực dư này dùng để tạo gia tốc cho xe. Khi xe bắt đầu bơi (V=0), lực dư đạt cực đại.

- Vận tốc bơi tiến lớn nhất V_{max}, lực kéo lớn nhất của xe P_{kmax}: Vận tốc bơi tiến lớn nhất được xác định từ vị trí giao điểm giữa đường lực kéo có ích P_c và lực cản R_c. Tại điểm này, lực dư R_j= 0, xe không còn khả năng gia tốc và chuyển sang trạng thái bơi ổn định. Từ đồ thị ta xác định được V_{max} = 10,49 km/h, P_{kmax}= 1,41.10^4 (N)- đây là một con số rất ấn tượng đối với các dòng xe thiết giáp chở quân bánh lốp. Vận tốc bơi tiến lớn nhất của xe đạt được lớn hơn vận tốc bơi theo yêu cầu của xe là V_{0max}= 10 (km/h).

4. Kết luận

Bài báo đã xây dựng mô hình khảo sát và tính toán các tham số động lực học bơi, bao gồm: vận tốc bơi, lực cản của nước, lực cản không khí, lực đẩy do cơ cấu bơi tạo ra, lực kéo có ích xe thiết giáp chở quân 8x8 chế tạo tại Việt Nam. Các kết quả tính toán lý thuyết đã khẳng định hệ thống bơi dạng bơm luồng phụt của xe XTC-03 có khả năng đáp ứng tốt các chỉ tiêu tính năng kỹ thuật theo yêu cầu của xe. Ngoài ra, các kết quả tính toán này cũng là cơ sở cho việc khảo sát thực nghiệm động lực học của xe trong thực tế.

Tài liệu tham khảo

1. Bộ môn Xe quân sự (2001*), Cấu tạo xe tăng thiết giáp tập 1: Lịch sử phát triển xe T-TG,* HVKTQS, Hà Nội.
2. Nguyễn Viết Tân (2011), *Xe máy vượt sông*, Nhà xuất bản Quân đội nhân dân, Hà Nội.
3. Hu, H.K., Zhang, S.T., Xiong, K.F., (2006), *Study and experiment on improving the speed by resistance reduction*, Vehicle & Power Technology, Vol.101, pp.5-10.

4. Sun, C.L., Xu, X.J., Tang, Y.J., Hao, J., (2022), *Experimental and numerical simulation of reducing resistance and increasing speed for a segmented-track amphibious vehicle*, Journal of National University of Defense Science and Technology, Vol.44, pp.201-208.
5. Sebnem Helvacioglu, Ismail Hakki Helvacioglu, Burak Tuncer, (2011), *Improving the river crossing capability of an amphibious vehicle*, Ocean Engineering, Vol.38, pp.2201-2207.
6. Kaczmarczyk, S.T., Kulczyk, J., Zaradny, Z., (1996), *Effect of the form parameters of a tracked APC on its swimming resistance*, Journal of Terramechanics, Vol.33(4), pp.173-183.
7. Д.Г. Егоров, Л.Е. Ходонович (1963), *Самоходные переправочные средства и средства моторизации переправ,* Издание виа, Москва.
8. Nguyễn Tiến Dũng (1999), *Khảo sát động lực học bơi của xe xích quân sự,* HVKTQS.
9. Phạm Văn Tuấn (2001), *Khảo sát động lực học xe bơi có cơ cấu bơi bằng* xích, HVKTQS.

Study on the swimming dynamics of the 8×8 wheeled armored personnel carrier

Abstract: *At present, our Army is actively implementing research, design, and manufacturing programs for new generations of armored vehicles to meet training and combat readiness requirements, including the 8×8 armored personnel carrier. One of the important features of this vehicle platform is its stable amphibious capability, employing a hydraulically driven waterjet propulsion system. In this paper, the authors present the theoretical basis, modeling methodology, and calculation procedure for analyzing the straight-line propulsion performance during amphibious operation of the 8×8 armored personnel carrier manufactured in Vietnam. The dynamic parameters investigated include water drag and the thrust generated by the waterjet propulsion system as functions of the vehicle's swimming velocity.*

Keywords: Armored personnel carrier 8x8; waterjet propulsion system; swimming velocity; water drag

Khảo sát khả năng quay vòng ổn định của xe xích quân sự ở tốc độ cao trên đường cát

Hoàng Văn Dinh[1] Học viện Kỹ thuật quân sự

Email: kcdd2014@gmail.com; Tel:0833466698

Tóm tắt

Một trong những xu hướng chung của xe xích quân sự hiện đại là nâng cao tốc độ di chuyển của xe. Tuy nhiên, khi xe quay vòng với tốc độ lớn, xe xích quân sự càng dễ mất điều khiển, gây mất an toàn. Trong bài báo này, nhóm tác giả trình bày mô hình động lực học quay vòng của xe xích quân sự ở tốc độ cao, qua đó khảo sát quá trình quay vòng của xe xích quân sự ở tốc độ cao, từ đó khảo sát khả năng quay vòng ổn định của xe xích quân sự ở tốc độ cao. Các kết quả nghiên cứu được áp dụng cho mô hình trên đường cát được chỉ ra trong bài báo.

Từ khóa: *quay vòng; xe xích quân sự; tốc độ cao; đường cát.*

1. Đặt vấn đề

Việc thiết kế và điều khiển xe xích tốc độ cao đòi hỏi không chỉ nghiên cứu về ổn định quay vòng của xe mà còn phải hiểu rõ sự thay đổi của các thông số chuyển động trong quá trình quay vòng không ổn định. Điều kiện quay vòng không ổn định của xe xích là điều kiện mà xe xích không thể duy trì bán kính quay vòng ổn định mong muốn do các yếu tố như độ trượt tương đối của xích với mặt đất, và lực ly tâm trong quá trình quay vòng. Các thông số chuyển động, chẳng hạn như bán kính quay vòng, tốc độ lệch hướng, và góc trượt ngang, thay đổi nhanh chóng và xe sẽ quay vòng ở tốc độ ngang lớn [1]. Trong điều kiện quay vòng không ổn định, việc điều khiển quay vòng của xe xích bị giảm và khó đảm bảm quay vòng một cách an toàn [2]. Việc xác định giới hạn ổn định quay vòng ở tốc độ cao của xe xích có thể cung cấp cơ sở lý thuyết cho việc thiết kế và điều khiển xe xích tốc độ cao và cải thiện việc quay vòng xe an toàn [3]. Hiện nay, mô hình động học được sử dụng rộng rãi trong việc phân tích điều kiện quay vòng tốc độ thấp của xe xích. Tuy nhiên, lỗi tính toán của mô hình này lớn khi bỏ qua ảnh hưởng của các yếu tố như độ trượt tương đối của xích với mặt đất, độ trượt ngang của xe và lực ly tâm, và do đó, khó để đáp ứng các yêu cầu về độ chính xác tính toán của mô hình đối với các điều kiện quay vòng không ổn định [4]. Một phương án hiệu quả để nghiên cứu tính ổn định quay vòng tốc độ cao của xe xích là thiết lập một mô hình động lực học quay vòng có độ chính xác cao, phân tích những thay đổi của các thông số chuyển động trong quá trình quay vòng tốc độ cao và xác định ranh giới (giới hạn) ổn định quay vòng tốc độ cao của xe xích.

Bài báo này đề cập tới phương pháp xây dựng mô hình động lực học quay vòng của xe xích trên cơ sở mô hình ứng suất cắt, và áp dụng để nghiên cứu tính ổn định quay vòng tốc độ cao của xe xích. Một mô hình mô phỏng được thành lập bằng MATLAB/Simulink để mô phỏng và phân tích những thay đổi của quỹ đạo quay vòng, bán kính quay vòng, và vận tốc trọng tâm xe trong quá trình quay vòng tốc độ cao, được so sánh trong ba điều kiện quay vòng khác nhau.

2. Phương pháp

Khi thiết lập mô hình động lực học của xe xích chấp nhận những giả thuyết sau:

(1) Trọng tâm của xe xích trùng với tâm hình học.

(2) Bỏ qua sự lún xích và tác động đẩy ngang của dải xích.

(3) Xích không co giãn, và bỏ qua ảnh hưởng của chiều rộng xích đến áp lực trên nền đất.

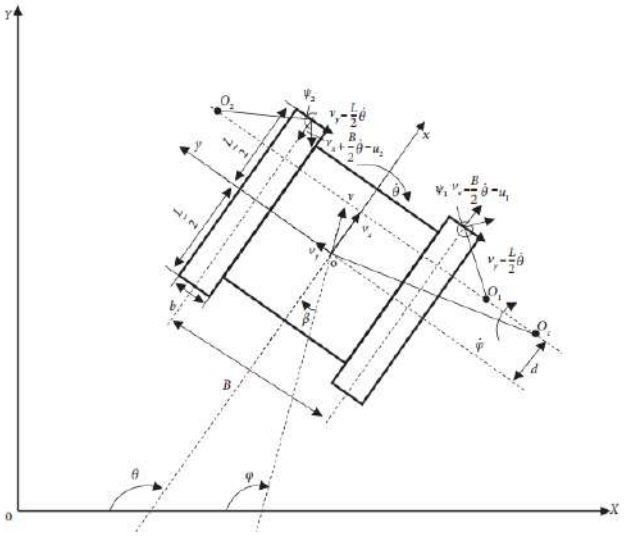
(4) Trong quá trình quay vòng của xe xích, ứng suất cắt τ tại bất kỳ điểm nào trên dải xích đều liên quan đến độ dịch chuyển cắt j tại điểm đó, theo biểu thức:

$$\tau = (c + \sigma\tan\phi)(1 - e^{-j/K}) \quad (1)$$

Trong đó: c là lực kết dính của đất, σ là ứng suất pháp tuyến, ϕ là góc kháng cắt bên trong của địa hình, và K là mô đun biến dạng cắt.

2.1. Phân tích động học quay vòng xe xích

Để xây dựng mô hình động lực học quay vòng của xe xích, thiết lập hệ tọa độ được trình bày trong Hình 1. Hệ XOY là hệ tọa độ cố định. Hệ xoy là hệ tọa độ thân xe, với gốc tọa độ O trùng với tâm hình học của xe, và phương x định hướng theo hướng tiến của xe.



Hình 1. Động học quay vòng của xe xích

Trong đó:

v là vận tốc trọng tâm của xe, với vx và vy lần lượt là các thành phần vận tốc theo phương x (phương dọc) và phương y (phương ngang) của hệ tọa độ thân xe.

θ là góc lệch hướng (hay góc quay vòng) của xe xích, được đo so với hệ tọa độ cố định.

φ là góc định hướng vận tốc, được xác định là góc hợp bởi phương vận tốc trọng tâm xe (v) và trục X của hệ tọa độ cố định.

β là góc trượt ngang (hay góc trượt thân xe), được tính bằng:

$$\beta = \arctan(vy/vx) \quad (2)$$

L là chiều dài tiếp xúc của dải xích với mặt đất.

B là chiều rộng cơ sở giữa hai dải xích.

b là chiều rộng dải xích.

ψ là góc hợp bởi phương vận tốc tại một điểm trên dải xích và trục x của hệ tọa độ thân xe.

O_1 và O_2 lần lượt là các tâm quay tức thời của dải xích phía tốc độ thấp (xích dừng) và dải xích phía tốc độ cao (xích chạy). Chỉ số dưới i = 1 đại diện cho dải xích dừng, và i = 2 đại diện cho dải xích chạy.

d là độ lệch tâm của tâm quay tức thời O_C so với trục x.

Mối quan hệ giữa vận tốc trọng tâm v của xe xích và các gia tốc a_x, a_y theo phương trục x và trục y tương ứng, có thể được biểu diễn như sau:

$$\begin{cases} v = \sqrt{v_x^2 + v_y^2}, \\ a_x = \dot{v}_x + v_y\dot{\theta}, \\ a_y = \dot{v}_y - v_x\dot{\theta}. \end{cases} \tag{3}$$

Trong quá trình quay vòng của xe xích, quỹ đạo trong hệ tọa độ cố định XOY có thể được biểu diễn như sau:

$$\begin{cases} X = -\int_0^t v\cos\varphi dt, \\ Y = \int_0^t v\sin\varphi dt, \\ \varphi = \theta - \beta. \end{cases} \tag{4}$$

Các thành phần vận tốc trượt cắt v_{sxi} và v_{syi} của điểm này so với mặt đất theo phương x và phương y có thể được biểu diễn như sau:

$$\begin{cases} v_{sx1} = v_x - \dfrac{B}{2}\dot{\theta} - u_1, \\ v_{sx2} = v_x + \dfrac{B}{2}\dot{\theta} - u_2, \\ v_{syi} = v_y - x_p\dot{\theta}, i = 1, 2, \end{cases} \tag{5}$$

Trong đó: u_i là vận tốc vòng của bánh chủ động phía i.

Góc ψ_i hợp bởi phương vận tốc của một điểm trên dải xích phía i và trục x có thể được biểu diễn như sau:

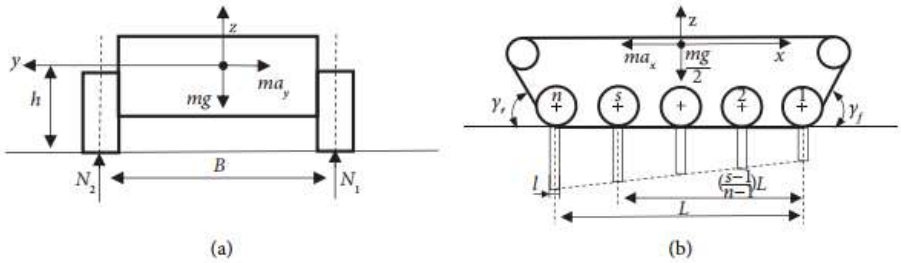
$$\psi_i = \arccos\left(\frac{v_{sxi}}{\sqrt{v_{sxi}^2 + v_{syi}^2}}\right), \quad i = 1, 2. \tag{6}$$

Độ dịch chuyển cắt j tại một điểm trên dải xích bắt đầu tích lũy từ mép trước của mặt phẳng tiếp xúc giữa xích và mặt đất, và đạt giá trị cực đại tại mép sau của mặt phẳng này. Độ dịch chuyển cắt ji của một điểm bất kỳ (xp, yp) trên dải xích phía i có thể được xác định thông qua việc tích phân vận tốc trượt cắt, và được biểu diễn như sau:

$$\begin{aligned} j_{xi} &= \int_0^t v_{sxi}dt = \left(\frac{v_{sxi}}{\omega_i r}\right)\left(\frac{L}{2} - x_p\right), \quad i = 1, 2, \\ j_{yi} &= \int_0^t v_{syi}dt = \frac{1}{\omega_i r}\left(v_y\left(\frac{L}{2} - x_p\right)\right. \\ &\quad \left. + \frac{1}{2}\dot{\theta}\left(x_p^2 - \frac{L^2}{4}\right)\right), \quad i = 1, 2, \\ j_i &= \sqrt{j_{xi}^2 + j_{yi}^2}, \quad i = 1, 2. \end{aligned} \tag{7}$$

2.2. Phân tích lực quay vòng của xe xích

Trong quá trình quay vòng của xe xích, sự phân bố của áp lực tiếp đất trực tiếp ảnh hưởng đến lực trên dải xích. Áp lực tiếp đất của dải xích chủ yếu tập trung dưới các bánh tỳ. Để đơn giản hóa quá trình tính toán, người ta giả định rằng áp lực tiếp đất dưới mỗi bánh tỳ được phân bố theo dạng hình chữ nhật ngay bên dưới bánh tỳ đó.



Hình 2. Sự phân bố tải trọng pháp tuyến của dải xích ở hai bên dưới ảnh hưởng của lực quán tính

(a) Sự phân bố lực pháp tuyến tại vùng tiếp đất của xích

(b) Sự phân bố áp lực tiếp đất của các bánh tỳ trên dải xích độc lập

Lực quán tính phát sinh do quá trình quay vòng của xe xích sẽ làm thay đổi sự phân bố áp lực tiếp đất trên cả hai dải xích, điều này càng trở nên đáng kể trong quá trình quay vòng ở tốc độ cao [5]. Thành phần lực quán tính theo phương y (m_{ay}) làm tăng lực pháp tuyến N_2 tại dải xích phía tốc độ cao tiếp xúc với mặt đất và làm giảm lực pháp tuyến N_1 tại dải xích phía tốc độ thấp. Lực pháp tuyến N_i của vùng tiếp đất dải xích phía i có thể được biểu diễn như sau:

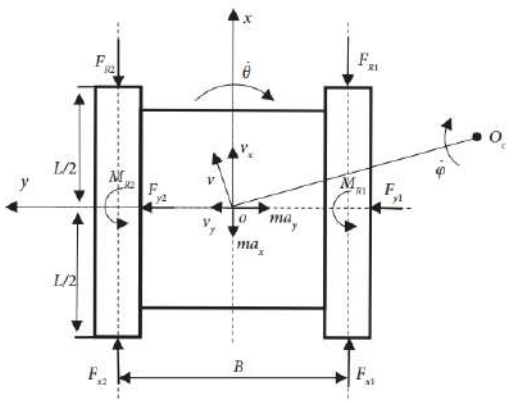
$$N_i = \frac{mg}{2} \pm \frac{ma_y h}{B}, \quad i = 1,2, \tag{8}$$

Trong đó: h là chiều cao trọng tâm xe.

Thành phần lực quán tính theo phương x (m_{ax}) làm tái phân bố áp lực của mỗi bánh tỳ trên dải xích độc lập. Áp lực tiếp đất trên đơn vị diện tích của bánh tỳ thứ s thuộc dải xích phía i có thể được biểu diễn như sau:

$$\overline{p}_{is} = N_i - \frac{3(2s-n-1)ma_x h}{n(n+1)blL}, \quad i = 1,2. \tag{9}$$

Hình 3 minh họa các lực trên dải xích khi xe xích quay vòng.



Hình 3. Lực trên các dải xích khi quay vòng

F_{xi} và F_{yi} lần lượt là các thành phần lực cắt giữa dải xích phía i và mặt đất theo phương x và phương y.

M_{Ri} là mô men cản quay vòng của dải xích phía i.

F_{Ri} là lực cản lăn của dải xích phía i.

Các thành phần Fxis và Fyis tương ứng, và có thể được biểu diễn như sau:

$$\begin{aligned}
F_{xis} &= \int_{\frac{(3-s)L}{4}-\frac{l}{2}}^{\frac{(3-s)L}{4}+\frac{l}{2}} b\cdot(c+p_{is}\tan\phi)\left(1-e^{-j_{xi}/K}\right)\cos(\psi_i+\pi)dx, \quad i=1,2,\\
F_{xi} &= \sum_{s=1}^{n} F_{xis}, \quad i=1,2,\\
F_{yis} &= \int_{\frac{(3-s)L}{4}-\frac{l}{2}}^{\frac{(3-s)L}{4}+\frac{l}{2}} b\cdot(c+p_{is}\tan\phi)\left(1-e^{-j_{yi}/K}\right)\sin(\psi_i+\pi)dx, \quad i=1,2,\\
F_{yi} &= \sum_{s=1}^{n} F_{yis}, \quad i=1,2.
\end{aligned} \tag{10}$$

Lực cản lăn Fri bên dải xích phía i có thể biểu diễn như sau:

$$F_{Ri} = \sum_{s=1}^{n} p_{is} f b l, \tag{11}$$

Trong đó: f là hệ số cản lăn.

Mô men cản quay vòng M_{Ri} của dải xích phía i cũng được xác định bằng cách tổng hợp các mô men cản quay vòng cục bộ M_{Ris} của mỗi bánh tỳ, và có thể được biểu diễn như sau:

$$\begin{aligned}
M_{Ris} &= \int_{\frac{(3-s)L}{4}-\frac{l}{2}}^{\frac{(3-s)L}{4}+\frac{l}{2}} b\cdot(c+p_{is}\tan\phi)\left(1-e^{-j_{yi}/K}\right)\sin(\psi_i+\pi)x\,dx, i=1,2,\\
M_{Ri} &= \sum_{s=1}^{n} M_{Ris}, \quad i=1,2.
\end{aligned} \tag{12}$$

Phương trình động lực học của xe xích có thể được biểu diễn như sau:

$$\begin{cases}
ma_x = F_{x1}+F_{x2}-F_{R1}-F_{R2},\\
ma_y = F_{y1}+F_{y2},\\
I_z\ddot{\theta} = \frac{B}{2}(F_{x2}-F_{x1})+\frac{B}{2}(F_{R2}-F_{R1})-M_{R1}-M_{R2},
\end{cases} \tag{13}$$

Trong đó I_z là mô men quán tính quay vòng của xe xích quanh trục z.

Các gia tốc của xe theo phương x và phương y, cùng với gia tốc góc quay vòng có thể được biểu diễn như sau:

$$\begin{cases} \dot{v}_x = \dfrac{(F_{x1} + F_{x2} - F_{R1} - F_{R2})}{m - v_y\dot{\theta}}, \\ \dot{v}_y = \dfrac{(F_{y1} + F_{y2})}{m + v_x\dot{\theta}}, \\ \ddot{\theta} = \left(\dfrac{B}{2}(F_{x2} - F_{x1}) + \dfrac{B}{2}(F_{R2} - F_{R1}) - M_{R1} - M_{R2}\right)/I_z. \end{cases} \quad (14)$$

3. Kết quả và thảo luận

Mô hình mô phỏng động lực học quay vòng của xe xích được thành lập, mô phỏng và phân tích dựa trên cơ sở phần mềm MATLAB/Simulink. Đồng thời, mô hình động học quay vòng cũng được thiết lập để so sánh trong thử nghiệm mô phỏng. Kích thước bước mô phỏng là $\Delta t = 0.1$ s. Các thông số đầu vào của mô phỏng tại thời điểm t là vận tốc vòng của bánh chủ động hai bên, ký hiệu là u_{1t} và u_{2t}.

Thử nghiệm mô phỏng được tiến hành với mặt đường cát. Các thông số của xe xích và mặt đường được trình bày chi tiết trong Bảng 1.

Bảng 1: Thông số xe và mặt đường

	Tên thông số	**Ký hiệu**		**Đơn vị**	**Giá trị**
Thông số đường	Lực kết dính của đất	c		Pa	1040
	Góc kháng cắt	ϕ		độ	28
	Mô đun biến dạng cắt	K		cm	2,5
	Lực cản mặt đường	f			0,065
Thông số xe	Khối lượng xe	m		kg	36500
	Mô men quán tính của xe	Iz		$kg.m^2$	13000
	Chiều dài xích tiếp xúc	L		m	3,78
	Chiều rộng cơ sở	B		m	2,64
	Chiều rộng mắt xích	b		m	0,58
	Chiều dài mắt xích	l		m	0,18
	Bán kính bánh chủ động	r		m	0,283
	Chiều cao trọng tâm	h		m	1,159
	Số lượng bánh tỳ	n			5

Trong quá trình mô phỏng, các điều kiện ban đầu được thiết lập như sau:

Vận tốc trọng tâm ban đầu: $v_0 = 6$ m/s.

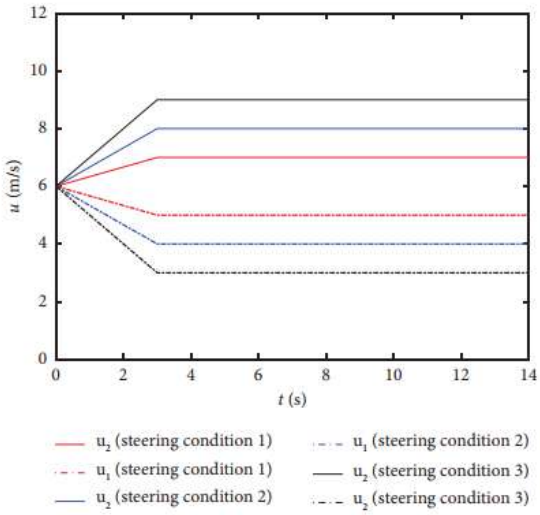
Góc lệch hướng ban đầu: $\theta_0 = \pi$ rad.

Góc trượt ngang trọng tâm ban đầu: $\beta_0 = 0$ rad.

Gia tốc trọng tâm ban đầu: $a_{x0} = a_{y0} = 0$ m/s^2.

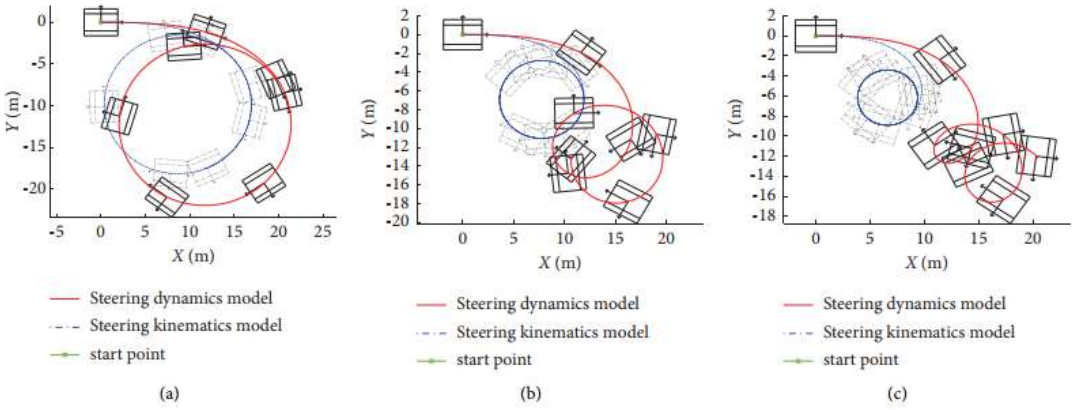
Quá trình khảo sát tiến hành với ba điều kiện quay vòng thông qua việc thay đổi vận tốc vòng ui. Cụ thể, tại thời điểm bắt đầu mô phỏng, vận tốc vòng của hai bánh chủ động là bằng nhau: $u_1 = u_2 = 6$ m/s. Trong khoảng thời gian từ t = 0 đến t = 3 s: Vận tốc vòng u_2 được tăng

tuyến tính lên các giá trị 7, 8, và 9 m/s. Đồng thời, vận tốc vòng u_1 được giảm tuyến tính xuống các giá trị tương ứng là 5, 4, và 3 m/s. Mối quan hệ thay đổi giữa u_1 và u_2 được minh họa chi tiết trong Hình 4.



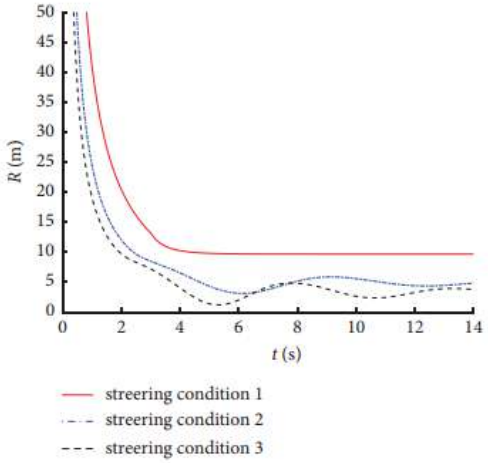
Hình 4. Vận tốc hai bên dải xích

Thời gian thử nghiệm mô phỏng là 14 giây, và vị trí của xe được đánh dấu một lần sau mỗi 2 giây. Hình 5 trình bày quỹ đạo quay vòng của xe xích. Khi xe xích quay vòng theo Điều kiện Quay vòng 1, quỹ đạo quay vòng được tính toán bằng cả mô hình động lực học và mô hình động học đều là các vòng tròn đều. Tuy nhiên, so với kết quả tính toán của mô hình động học, bán kính quỹ đạo quay vòng được tính toán bởi mô hình động lực học đã tăng lên đáng kể. Khi xe xích quay vòng theo Điều kiện Quay vòng 2 và 3, quỹ đạo quay vòng được tính toán bởi mô hình động lực học không còn là một vòng tròn đều nữa. Thay vào đó, xe đã bị mất ổn định tại vận tốc lý thuyết v_{th} = 6 m/s và có thể được xem là đã mất kiểm soát. Khi xe xích quay vòng theo Điều kiện Quay vòng 3, hiện tượng mất ổn định trở nên rõ rệt hơn cùng với sự gia tăng của độ chênh lệch vận tốc vòng Δu giữa hai bánh chủ động. Tuy nhiên, mặc dù Δu tăng lên, quỹ đạo quay vòng được tính toán bằng mô hình động học vẫn là một vòng tròn đều, và bán kính quay vòng bị giảm. Kết quả mô phỏng của mô hình động học cho thấy xe xích có thể đạt được trạng thái quay vòng ổn định với bán kính R = 2,64 m trong điều kiện này, điều này là không phù hợp với tình huống thực tế.



Hình 5. Quỹ đạo quay vòng của xe xích ở ba điều kiện quay vòng

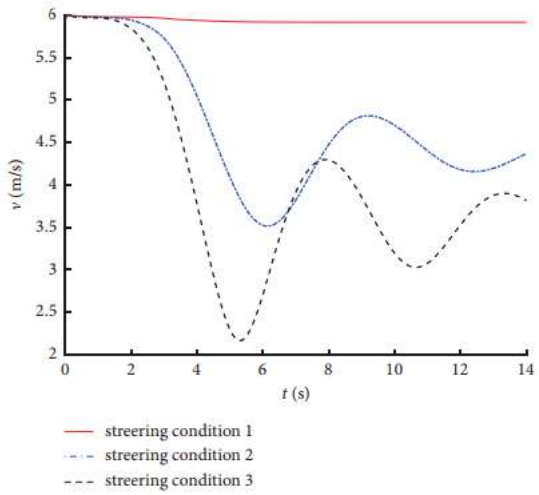
Hình 6 minh họa sự thay đổi của bán kính quay vòng của xe xích dưới ba điều kiện quay vòng, được tính toán bằng mô hình động lực học. Dưới Điều kiện Quay vòng 1, bán kính quay vòng giảm khi độ chênh lệch vận tốc vòng Δu tăng lên trong khoảng thời gian 0 đến 3 giây. Trong khoảng 3 đến 14 giây, Δu không thay đổi, và bán kính quay vòng của xe cũng giữ nguyên. Mô hình động lực học quay vòng tính toán được bán kính quay vòng ổn định là R = 9,62 m. Có thể thấy rằng, xe xích có thể đạt được trạng thái quay vòng ổn định tại bán kính R = 9,62 m dưới điều kiện quay vòng này.



Hình 6. Sự thay đổi của bán kính quay vòng R ở ba điều kiện quay vòng tính toán theo mô hình động lực học

Dưới Điều kiện Quay vòng 2 và 3, có thể thấy rõ từ kết quả của mô hình động lực học rằng quá trình quay vòng của xe xích với bán kính quay vòng nhỏ ở tốc độ cao đã trở nên không ổn định. Cụ thể dưới Điều kiện Quay vòng 3 (mức chênh lệch vận tốc vòng lớn nhất), xe đã đạt tới bán kính quay vòng nhỏ nhất là R = 1,22 m tại thời điểm t = 5,23 s. Sau đó, bán kính quay vòng bắt đầu tăng lên và đạt tới bán kính quay vòng cực đại là R = 4,81 m tại t = 7,69 s. Trong toàn bộ quá trình quay vòng, bán kính quay vòng của xe xích đã dao động lên xuống, và biên độ dao động này giảm dần theo thời gian quay vòng.

Hình 7 minh họa sự biến thiên của vận tốc trọng tâm v của xe xích dưới ba điều kiện quay vòng, được tính toán bằng mô hình động lực học. Vận tốc trọng tâm v của xe xích quay vòng thấp hơn vận tốc trọng tâm lý thuyết v_{th} khi xe quay vòng ở bán kính nhỏ tại tốc độ cao, và vận tốc này giảm dần khi độ chênh lệch vận tốc vòng Δu tăng lên. Hiện tượng này là hệ quả của sự gia tăng thành phần vận tốc trượt cắt dọc trục vsxi do sự tăng lên của Δu.



Hình 7. Sự thay đổi của vận tốc trọng tâm v ở ba điều kiện quay vòng được tính toán theo mô hình động lực học

4. Kết luận (Conclusion)

Thông qua việc thiết lập mô hình động lực học quay vòng của xe xích, có thể rút ra một số kết luận:

- Mô hình động lực học quay vòng đã tính đến ảnh hưởng của độ trượt xích so với mặt đất, độ trượt ngang của xe và lực ly tâm. So với mô hình động học quay vòng, kết quả tính toán của mô hình động lực học đối với các điều kiện quay vòng không ổn định của xe xích có thể phản ánh chính xác hơn các điều kiện quay vòng thực tế.

- Bằng cách khảo sát mô hình trong ba điều kiện quay vòng khác nhau trên cùng một loại đường và so sánh với mô hình động học, rút ra được sự thay đổi của quỹ đạo quay vòng, bán kính qyay vòng và vận tốc trọng tâm xe sát với thực tế hơn. Đây là cơ sở để kết luận xe có thể quay vòng ổn định ở điều kiện nào.

Định hướng nghiên cứu tiếp theo: bằng cách mô phỏng quá trình quay vòng của xe xích trên các loại mặt đường khác nhau, nghiên cứu đã thu được mối quan hệ giữa độ chênh lệch vận tốc vòng tối đa u_{max}, bán kính quay vòng tối thiểu R_{min} và vận tốc trọng tâm lý thuyết v_{th}. Việc sử dụng các kết quả mô phỏng này làm giới hạn ổn định quay vòng tốc độ cao của xe xích có thể cung cấp tài liệu tham khảo quan trọng cho: thiết kế kết cấu xe và kiểm soát khả năng quay vòng của xe xích tốc độ cao.

Tài liệu tham khảo

1. M. Kitano and M. Kuma (1977), *An analysis of horizontal plane motion of tracked vehicles*, Journal of Terramechanics, vol. 14, no. 4, pp. 211–225.
2. D. J. Purdy, D. Simner, D. Diskett, A. Duncan, P. J. H. Wormell, and C. Stonier (2016), *An experimental and theoretical investigation into the roll-over of tracked vehicles*, Proceedings of the Institution of Mechanical Engineers - Part D: Journal of Automobile Engineering, vol. 230, no. 3, pp. 291– 307.
3. H. Wang, Q. Wang, and Q. Rui (2014), *Analyzing and testing veri6cation the performance about high-speed tracked vehicles in steering process*, Journal of Mechanical Engineering, vol. 50, no. 16, pp. 162–172.

4. J. L. Mart´ınez, A. Mandow, J. Morales, A. Garcia-Cerezo, and S. Pedraza (2004), *Kinematic modelling of tracked vehicles by experimental identi6cation*, in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), no. 2, pp. 1487–1492, Japan.
5. Lê Trung Dũng, Nguyễn Minh Tân, Cù Xuân Phong (2022). *Lý thuyết xe xích quân sự*, Nhà xuất bản Quân đội nhân dân, Hà Nội.

Directional Stability Analysis of Military Tracked Vehicles during High-Speed Steering on Sandy Terrain

Abstract: One of the prevailing trends in modern military tracked vehicle design is the continuous enhancement of operational speeds. However, executing high-speed steering maneuvers significantly increases the risk of loss of control and directional instability, posing critical safety challenges. This paper presents a high-fidelity steering dynamic model specifically developed for military tracked vehicles at high velocities. Through this model, the steering characteristics and maneuverability limits are rigorously investigated to evaluate the vehicle's stability during high-speed locomotion. The research findings, specifically simulated and validated on sandy terrain, are detailed within this study.

Keywords: skid-steering; military tracked vehicle; high speed; sandy terrain; directional stability.

Nghiên cứu thiết kế hệ thống chữa cháy tự động cho xe thiết giáp bánh lốp

Cù Xuân Phong[1] *, Bùi Hà Anh[2]

[1]*Viện Cơ khí động lực, Học viện KTQS;* [2]*Đại học Phòng cháy chữa cháy*

Email:phong.cx@lqdtu.edu.vn

Tóm tắt

Trong điều kiện tác chiến hiện đại, xe thiết giáp bánh lốp thường xuyên phải hoạt động trong môi trường có nguy cơ cháy nổ cao do hỏa lực đối phương, sự cố kỹ thuật và điều kiện khai thác khắc nghiệt. Bài báo trình bày kết quả nghiên cứu thiết kế hệ thống chữa cháy tự động cho xe thiết giáp bánh lốp, tập trung vào việc phân tích nguy cơ cháy, lựa chọn tác nhân chữa cháy, thiết kế hệ thống chữa cháy tự động. Trên cơ sở các yêu cầu kỹ thuật – chiến thuật, một phương án hệ thống chữa cháy tự động dạng khí được đề xuất, đáp ứng yêu cầu phản ứng nhanh, hiệu quả cao, an toàn cho kíp xe và phù hợp với điều kiện khai thác quân sự.

Từ khóa: *xe thiết giáp bánh lốp, hệ thống chữa cháy tự động, chất chữa cháy, bình chữa cháy*

1. Đặt vấn đề

Xe thiết giáp bánh lốp có khả năng cơ động và tác chiến trên nhiều loại địa hình khác nhau, từ đường giao thông đô thị, đường nông thôn, địa hình đồi núi đến khu vực ngập nước, bùn lầy. Nhờ ưu thế về tốc độ hành quân, bán kính hoạt động lớn và khả năng triển khai nhanh, loại phương tiện này giữ vai trò đặc biệt quan trọng trong hệ thống trang bị kỹ thuật của lực lượng vũ trang hiện đại. Trong thực tiễn sử dụng, xe thiết giáp bánh lốp đảm nhiệm nhiều nhiệm vụ khác nhau, bao gồm: cơ động lực lượng tiếp cận chiến trường, tổ chức chiến đấu theo các phương án tác chiến đã xác định; yểm trợ, chi viện hỏa lực cho các lực lượng trong tiến công và phòng ngự [1].

Trong quá trình thực hiện các nhiệm vụ trên, xe thiết giáp bánh lốp thường xuyên phải hoạt động trong điều kiện khắc nghiệt, chịu tác động đồng thời của tải trọng lớn, rung xóc mạnh, nhiệt độ cao, môi trường bụi bẩn, ẩm ướt và nguy cơ bị hỏa lực đối phương tiến công. Bên cạnh đó, kết cấu xe có nhiều hệ thống tiềm ẩn nguy cơ cháy nổ cao như động cơ công suất lớn, hệ thống nhiên liệu, hệ thống điện – điện tử phức tạp và khoang chiến đấu kín. Khi xảy ra sự cố cháy, quá trình phát triển cháy diễn ra rất nhanh do không gian hạn chế, lượng vật liệu dễ cháy tập trung lớn và khả năng lan truyền nhiệt mạnh, gây đe dọa trực tiếp đến tính mạng kíp xe cũng như khả năng tiếp tục thực hiện nhiệm vụ của phương tiện.

Thực tế khai thác và sử dụng xe thiết giáp bánh lốp cho thấy, các biện pháp chữa cháy thủ công truyền thống không đáp ứng được yêu cầu dập tắt cháy trong thời gian ngắn, đặc biệt trong điều kiện chiến đấu hoặc khi xe đang cơ động [2]. Do đó, yêu cầu đặt ra là phải trang bị cho xe một hệ thống chữa cháy có khả năng phát hiện sớm, phản ứng nhanh và hoạt động hiệu quả trong không gian kín, đồng thời đảm bảo an toàn cho con người và trang bị trên xe.

Bài báo này trình bày kết quả nghiên cứu thiết kế hệ thống chữa cháy tự động cho xe thiết giáp bánh lốp, trong đó trọng tâm là bảo vệ hai khu vực có nguy cơ cháy nổ cao nhất trên xe, bao gồm khoang động lực và khoang chiến đấu. Theo phương án bố trí tổng thể của xe, khoang động lực được đặt phía trước bên trái thân xe, khoang chiến đấu bố trí ở khu vực trung tâm, trong khi

* Email: phong.cx@lqdtu.edu.vn

khoang bộ binh nằm ở phía đuôi xe (hình 1). Cách bố trí này phù hợp với xu hướng thiết kế xe thiết giáp bánh lốp hiện đại, nhằm tối ưu hóa khả năng cơ động, cân bằng khối lượng và thuận lợi cho tổ chức chiến đấu.

Hình 1. Bố trí các khoang trên xe thiết giáp bánh lốp

Khoang động lực là nơi tập trung động cơ công suất lớn, hệ thống truyền động, nhiên liệu và các cụm chi tiết làm việc ở nhiệt độ cao, do đó tiềm ẩn nguy cơ cháy nổ rất lớn. Khi xảy ra sự cố cháy trong khoang này, nhiệt lượng sinh ra nhanh chóng lan truyền sang các khu vực lân cận, đặc biệt là khoang chiến đấu nằm liền kề. Trong điều kiện xe đang cơ động hoặc chịu hỏa lực đối phương, việc phát hiện và xử lý cháy bằng phương pháp thủ công gần như không khả thi, làm gia tăng nguy cơ mất an toàn cho toàn bộ phương tiện.

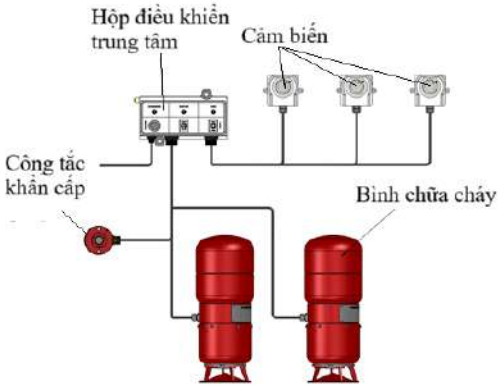
Khoang chiến đấu là không gian làm việc trực tiếp của kíp xe, đồng thời chứa các trang bị vũ khí, khí tài quang – điện tử, hệ thống điện và cơ cấu điều khiển. Đặc điểm không gian kín, mật độ thiết bị cao và sự hiện diện thường xuyên của con người khiến khoang này trở thành khu vực đặc biệt nhạy cảm đối với sự cố cháy. Nếu không được khống chế kịp thời, cháy trong khoang chiến đấu không chỉ gây thiệt hại nghiêm trọng về trang bị mà còn đe dọa trực tiếp đến tính mạng kíp xe và khả năng tiếp tục thực hiện nhiệm vụ.

2. Các yêu cầu khi thiết kế hệ thống chữa cháy xe thiết giáp

Hệ thống chữa cháy trên xe thiết giáp bánh lốp phải bảo đảm khả năng phát hiện sớm, dập tắt nhanh đám cháy, đồng thời duy trì an toàn cho kíp xe và khả năng chiến đấu của phương tiện trong mọi điều kiện khai thác. Hệ thống phải làm việc tự động kết hợp điều khiển thủ công, có khả năng phân vùng chữa cháy độc lập cho từng khoang và phù hợp với điều kiện rung động, va đập và môi trường khắc nghiệt của phương tiện quân sự [3].

Hệ thống phát hiện cháy phải bảo đảm khả năng nhận biết nhanh và chính xác sự xuất hiện của nguồn cháy trong các khoang chức năng của xe. Phương thức phát hiện sử dụng cảm biến phù hợp với đặc điểm không gian kín và mật độ thiết bị cao trong xe thiết giáp. Trong hệ thống, bộ điều khiển trung tâm đóng vai trò là khối xử lý và điều phối chính, có nhiệm vụ tiếp nhận tín hiệu từ các cảm biến, phân tích trạng thái nhiệt độ theo thời gian thực và so sánh với các ngưỡng kích hoạt đã được thiết lập cho từng khoang (hình 2). Khi tín hiệu từ cảm biến vượt quá giá trị cho phép hoặc xuất hiện hiện tượng tăng nhiệt đột ngột đặc trưng cho quá trình cháy, bộ điều khiển trung tâm sẽ phát lệnh kích hoạt hệ thống chữa cháy tương ứng. Quá trình xử lý tín hiệu và phát lệnh được thực hiện trong thời gian rất ngắn, bảo đảm hệ thống có thể phản ứng kịp thời

ngay ở giai đoạn đầu của đám cháy. Đồng thời, bộ điều khiển cho phép phân biệt và lựa chọn vùng chữa cháy độc lập, tránh kích hoạt không cần thiết tại các khoang không xảy ra cháy [4].



Hình 2. Các thành phần chính của hệ thống chữa cháy tự đông trên xe thiết giáp

Các thông số yêu cầu cụ thể về hệ thống chữa cháy tự động được thể hiện trong bảng 1.

Bảng 1. Các yêu cầu đối với hệ thống chữa cháy tự động trên xe thiết giáp

STT	Thông số	Giá trị yêu cầu	Ghi chú
1	Thời gian phát hiện cháy	≤ 3 s	Từ khi xuất hiện nguồn cháy
2	Thời gian kích hoạt hệ thống	≤ 1 s	Sau khi phát hiện cháy
	Nhiệt độ kích hoạt hệ thống	120–140 °C	Phát hiện cháy sớm
3	Thời gian dập cháy hoàn toàn	≤ 15 s	Không gian kín
4	Chế độ làm việc	Tự động và thủ công	Kích hoạt độc lập từng khoang
5	Độ tin cậy kích hoạt	≥ 0,98	Trong điều kiện rung xóc
6	Gia tốc rung cho phép	≥ 10 g	Hành quân và tác chiến

Bên cạnh bộ điều khiển trung tâm, hệ thống còn bao gồm các cụm chi tiết chấp hành và truyền dẫn, như cơ cấu kích hoạt bình chữa cháy, van điều khiển, hệ thống đường ống và các đầu phun. Các cụm chi tiết này được thiết kế làm việc đồng bộ dưới sự điều khiển của bộ điều khiển trung tâm, bảo đảm tác nhân chữa cháy được phun đúng vị trí, đúng thời điểm và đạt nồng độ cần thiết trong khoang được bảo vệ [5].

Ngoài chế độ làm việc tự động, hệ thống còn được trang bị các phần tử cảnh báo và điều khiển thủ công, cho phép kíp xe chủ động can thiệp trong các tình huống đặc biệt. Các tín hiệu cảnh báo bằng âm thanh và ánh sáng được phát ra ngay khi hệ thống phát hiện cháy hoặc khi hệ thống chữa cháy được kích hoạt, giúp kíp xe nhanh chóng nhận biết trạng thái làm việc của hệ thống và có biện pháp xử lý phù hợp.

3. Kết quả tính toán thiết kế

3.1. Tính toán khối lượng chất chữa cháy

Chất chữa cháy được lựa chọn là khí an toàn cho môi trường HFC-227ea (FM200). Khoang động lực xe thiết giáp bánh lốp 8x8 bố trí phía trước bên trái, mật độ thiết bị lớn. Thể tích khoang động lực được tính toán như sau [6]:

$$V_1 = L_1 \cdot B_1 \cdot H_1 \cdot k_1$$

Trong đó: V_1 - thể tích khoang động lực; L_1 - chiều dài khoang động lực (L_1=2,32 m); B_1 - chiều rộng khoang động lực (B_1=1,65 m); H_1 - chiều cao khoang động lực (H_1=1,7 m); k_1 – Hệ số điền đầy (k_1=0,65).

Thay số thu được $V_1 = 4{,}23$ m^3.

Khối lượng chất chữa cháy cần thiết được tính theo công thức [6]:

$$m_1 = C_1 \cdot \rho \cdot V_1 \cdot k_s$$

Trong đó: m_1 - khối lượng chất chữa cháy cần thiết; C_1 – nồng độ chất chữa cháy; ρ – khối lượng riêng của chất chữa cháy ($\rho = 1{,}25$ kg/m^3); k_s – Hệ số dự trữ. Nồng độ chất chữa cháy ở khoang động lực C_1= 10 % = 0,1. Hệ số dự trữ cho xe thiết giáp k_s = 1,2.

Từ đó xác định được khối lượng chất chữa cháy cần thiết cho khoang động lực xe thiết giáp bánh lốp là m_1 = 0,63 kg.

Tương tự với khoang chiến đấu, thể tích khoang V_2 = 5,85 m^3, nồng độ chất chữa cháy C_2= 8 % = 0,08 (để đảm bảo an toàn cho kíp xe) từ đó tính toán được khối lượng chất chữa cháy cần thiết khoang động lực là: m_2 = 0,70 kg.

3.2. Tính toán xác định các thông số của hệ thống

Với kết quả tính toán khối lượng chất chữa cháy cần thiết ở trên, lựa chọn số lượng bình chữa cháy cho xe thiết giáp bánh lốp là 02 bình ở mỗi khoang (đảm bảo có thể kích hoạt chữa cháy hai lần), khối lượng mỗi bình là 0,8 kg chất chữa cháy (để đảm bảo dự trữ và suy giảm trong điều kiện chiến đấu).

Mỗi cảm biến bao phủ 1,5-2 m^3, từ đó khoang động lực $n_{cb_dl} = \frac{4{,}23}{1{,}5} = 2{,}82$, chọn số lượng cảm biến là 3, khoang chiến đấu $n_{cb_cd} = \frac{5{,}85}{1{,}5} = 3{,}9$, chọn số cảm biến là 4.

Mỗi đầu phun đảm bảo bao phủ 1,0 ÷ 1,5 m³ thể tích hiệu quả. Khoang động lực $n_{cb_dl} = \frac{4{,}23}{1{,}2} = 3{,}52$, chọn số lượng đầu phun là 4, bố trí 4 góc xung quanh động cơ. Khoang chiến đấu $n_{cb_cd} = \frac{5{,}85}{1{,}2} = 4{,}87$, chọn số lượng đầu phun là 5. Trong đó 2 đầu phun bố trí trên trần khoang chiến đấu, 3 đầu phun bố trí các thành bên khoang chiến đấu. Bố trí các hệ thống chữa cháy trong các khoang được thể hiện trên hình 3, hình 4.

Thời gian phun hết một bình là 15 giây (theo yêu cầu đặt ra đối với hệ thống chữa cháy trên xe thiết giáp), từ đó lưu lượng phun ở các khoang là [7]:

$$Q = \frac{m_1}{\rho \cdot t} = \frac{0{,}8}{1{,}25 \cdot 15} \approx 0{,}042 \text{ m}^3/\text{s}$$

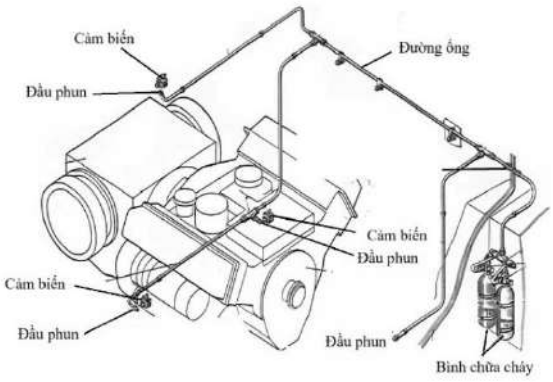
Đường kính ống chính xác định như từ công thức xác định lưu lượng qua ống:

$$\mathrm{Q} = \mathrm{v} \cdot \mathrm{A} = \mathrm{v} \cdot \frac{\pi \mathrm{d}^2}{4}$$

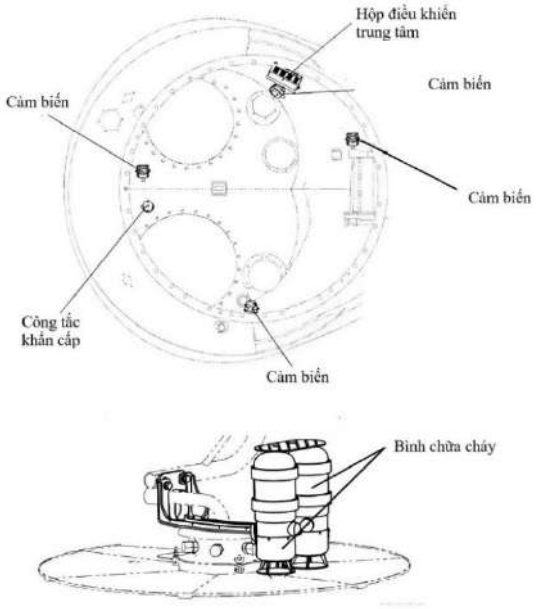
Trong đó: v- vận tốc dòng phun (v = 20÷25 m/s), chọn v = 20 m/s

$$d = \sqrt{\frac{4Q}{\pi v}} = \sqrt{\frac{4 \cdot 0{,}042}{\pi \cdot 20}} \approx 0{,}046 \text{ m}$$

Từ đó chọn ống chính tiêu chuẩn có đường kính Ø50 mm. Các đường ống nhánh ở khoang động lực (4 vòi phun) có đường kính là Ø15 mm, các đường ống nhánh ở khoang chiến đấu (5 vòi phun) là Ø10 mm.



Hình 3. Bố trí hệ thống chữa cháy trong khoang động lực



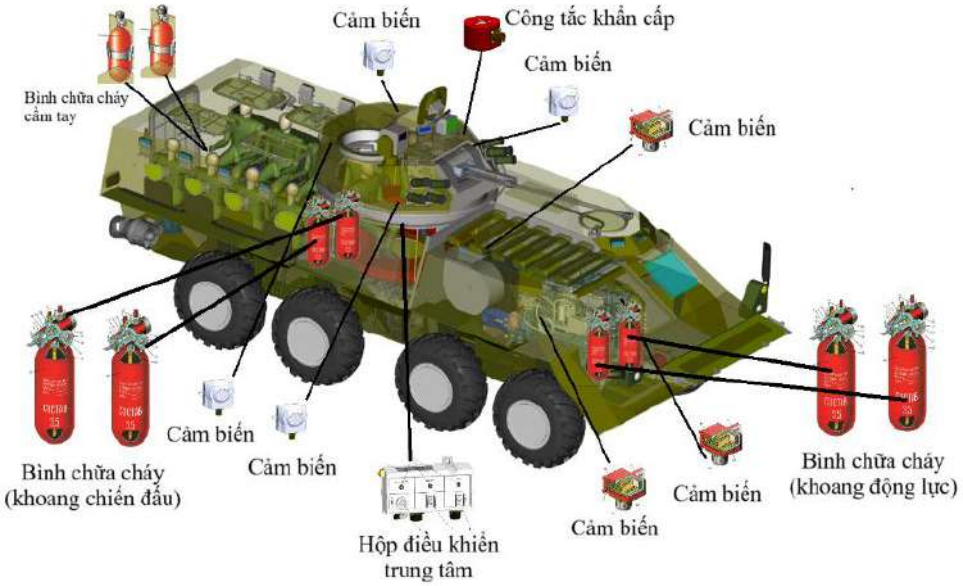
Hình 4. Bố trí hệ thống chữa cháy trong khoang chiến đấu

Bảng 2 thể hiện các thông số chính hệ thống chữa cháy tự động trên xe thiết giáp bánh lốp. Bố trí chung hệ thống chữa cháy được thể hiện trên hình 5.

Bảng 2. Các thông số hệ thống chữa cháy tự động xe thiết giáp bánh lốp

STT	Thông số	Khoang động lực	Khoang chiến đấu
1	Thể tích khoang (m^3)	4,23	5,85
2	Hộp điều khiển trung tâm	-	1
3	Công tắc khẩn cấp	-	1
4	Số lượng bình chữa cháy	2	2
5	Số lượng cảm biến	3	4

6	Loại cảm biến	Cảm biến nhiệt điện	Cảm biến quang
7	Số lượng đầu phun	4	5
8	Đường kính ống	10 –15 mm	10 –15 mm
9	Khối lượng chất chữa cháy/bình	0,8 kg	0,8 kg
10	Thời gian phun	15 s	15 s
11	Lưu lượng phun	0,042 m^3/s	0,042 m^3/s



Hình 5. Bố trí hệ thống chữa cháy tự động trên xe thiết giáp

4. Kết luận

Bài báo đã trình bày một cách hệ thống quá trình nghiên cứu, tính toán và thiết kế hệ thống chữa cháy tự động cho xe thiết giáp bánh lốp, tập trung vào hai khoang chức năng có nguy cơ cháy cao là khoang động lực và khoang chiến đấu. Kết quả tính toán cho thấy, với cấu hình mỗi khoang bố trí 2 bình chữa cháy dung lượng 0,8 kg chất chữa cháy, thời gian phun 15 giây, hệ thống có khả năng tạo ra nồng độ tác nhân chữa cháy cao trong các khoang, đảm bảo dập tắt đám cháy. Việc xác định lưu lượng phun, số lượng đầu phun, kích thước đường ống và số lượng cảm biến phát hiện cháy đã được thực hiện trên cơ sở các tiêu chí kỹ thuật cụ thể. Kết quả cho thấy phương án bố trí 4 đầu phun cho khoang động lực, 5 đầu phun cho khoang chiến đấu, kết hợp 3 cảm biến nhiệt điện khoang động lực, 4 cảm biến quang điện cho khoang chiến đấu, đáp ứng tốt yêu cầu phát hiện sớm và dập cháy nhanh, đồng thời phù hợp với không gian hạn chế và mật độ thiết bị cao bên trong xe.

Kết quả nghiên cứu cho thấy phương án thiết kế hệ thống chữa cháy tự động được đề xuất là khả thi, hiệu quả và phù hợp với yêu cầu thực tiễn của xe thiết giáp bánh lốp. Các kết quả của bài báo có thể làm cơ sở tham khảo cho thiết kế, cải tiến và nội địa hóa hệ thống chữa cháy trên các dòng xe thiết giáp hiện có, góp phần nâng cao độ an toàn, độ tin cậy và khả năng sống còn của phương tiện trong điều kiện tác chiến hiện đại.

Tài liệu tham

1. Общее устройство и техническое обслуживание бтр-80 - 2003

2. БТР-82А, Общее устройство и техническое обслуживание, 2010

3. Binh chủng Tăng thiết giáp (2008), Cấu tạo và sử dụng xe thiết giáp bánh lốp BTR-60PB, NXB Quân đội nhân dân.

4. Nguyễn Ngọc Ban, Nguyễn Hoàng Nam (1992). Khai thác trang bị Tăng thiết giáp. Học viện Kỹ thuật Quân sự.

5. Cục Kỹ thuật binh chủng (2003), Cấu tạo và hướng dẫn sử dụng xe chiến đấu bộ binh BMP-1, NXB Quân đội nhân dân.

6. Hodges, S. E. (2016). *Fire protection in military ground vehicles*. SAE International Journal of Transportation Safety, 4(2), 229–235

7. Korivi, V., McCormick, S., & Hodges, S. (2017). *Fire suppression modeling & simulation framework for ground vehicles.* SAE International Journal of Transportation Safety

Research on the Design of an Automatic Fire Suppression System for Wheeled Armored Vehicles

Abstract: *Under modern combat conditions, wheeled armored vehicles frequently operate in environments with a high risk of fire and explosion due to enemy firepower, technical failures, and harsh operating conditions. This paper presents the results of a study on the design of an automatic fire suppression system for wheeled armored vehicles, focusing on fire hazard analysis, selection of fire extinguishing agents, and the design of the automatic fire suppression system. Based on technical and tactical requirements, a gaseous automatic fire suppression system configuration is proposed, meeting the requirements of rapid response, high effectiveness, crew safety, and suitability for military operating conditions.*

Keywords: wheeled armored vehicle; automatic fire suppression system; fire extinguishing agent; fire extinguisher.

Nghiên cứu động lực học hệ truyền động thủy lực dẫn động hệ thống nâng tay kẹp của thiết bị hỗ trợ lắp ráp cấu kiện vòm trong thi công đường hầm

Nguyễn Hoa Hợi[1*]
[1]*Học viện Kỹ thuật Quân sự*
**Email: hoahoi378@gmail.com*

Tóm tắt

Cùng với sự phát triển mạnh mẽ của các dự án hạ tầng, công tác thi công hầm giao thông và hầm khai thác mỏ tại Việt Nam ngày càng được đẩy mạnh. Mặc dù công nghệ đào hầm đã đạt mức cơ giới hóa cao, song quá trình lắp đặt khung vòm thép vẫn còn phụ thuộc nhiều vào lao động thủ công, dẫn đến hạn chế về năng suất, độ chính xác và an toàn lao động. Do đó, việc nghiên cứu và ứng dụng các thiết bị hỗ trợ có khả năng nâng, giữ và định vị cấu kiện vòm thép trong không gian hẹp của hầm là cần thiết.

Bài báo trình bày các kết quả nghiên cứu động lực học của hệ truyền động thủy lực dẫn động hệ thống nâng tay kẹp trong thiết bị hỗ trợ lắp ráp cấu kiện vòm phục vụ thi công đường hầm. Các kết quả nghiên cứu là cơ sở khoa học cho việc tính toán, thiết kế hợp lý hệ truyền động thủy lực, góp phần giảm tải trọng động, hướng tới tự động hóa quá trình làm việc và nâng cao hiệu quả khai thác sử dụng thiết bị trong điều kiện thực tế.

Từ khoá: *hệ truyền động thủy lực; động lực học; cơ cấu nâng tay kẹp; cấu kiện vòm, đường hầm.*

1. Đặt vấn đề

Trong quá trình thi công các công trình hầm, đặc biệt là hầm giao thông xuyên núi hoặc hầm đô thị, cấu kiện vòm đóng vai trò then chốt trong việc gia cố sơ bộ, đảm bảo ổn định mái hầm và an toàn lao động. Việc lắp dựng các cấu kiện vòm trong điều kiện không gian chật hẹp, địa hình phức tạp và môi trường khắc nghiệt luôn là một trong những thách thức lớn. Để đáp ứng yêu cầu tiến độ và an toàn, nhiều đơn vị thi công hiện nay đã sử dụng thiết bị hỗ trợ lắp ráp cấu kiện vòm. Tuy nhiên, việc vận hành các thiết bị này hiện vẫn còn phụ thuộc nhiều vào kinh nghiệm thủ công, tính cơ giới hóa chưa cao, và đặc biệt chưa có nhiều nghiên cứu chuyên sâu về động lực học hệ thống thủy lực dẫn động cơ cấu công tác, gây ảnh hưởng đến hiệu quả và độ ổn định của quá trình lắp ráp.

Thực tế cho thấy, quá trình lắp dựng vòm trong hầm phải đảm bảo tính chính xác cao, nhanh chóng, đồng thời duy trì ổn định chuyển động của cơ cấu tay máy hoặc giá đỡ thông qua hệ thống dẫn động thủy lực. Nếu không kiểm soát tốt, sẽ xảy ra tình trạng rung, dao động, lệch vị trí cấu kiện, thậm chí gây mất an toàn lao động và hư hỏng thiết bị. Hiện nay, tại Việt Nam, các thiết bị hỗ trợ lắp vòm phần lớn là nhập khẩu hoặc được cải tiến theo kinh nghiệm thực tiễn, chưa được nghiên cứu bài bản về mặt động lực học hệ thống thuỷ lực của kết cấu chưa được quan tâm đầy đủ. Ngoài ra, trong bối cảnh nhu cầu xây dựng hạ tầng ngầm ngày càng tăng cao, yêu cầu đặt ra là phải phát triển các thiết bị có khả năng làm việc ổn định, chính xác, cơ động, phù hợp với điều kiện thi công trong nước, đặc biệt ở các tuyến đường miền núi, khu vực địa chất yếu.

Việc nghiên cứu động lực học của hệ thống xi lanh thủy lực điều khiển tay kẹp của thiết bị hỗ trợ lắp đặt cấu kiện vòm một cách bài bản sẽ giúp cho việc khảo sát ảnh hưởng của thông số thiết bị đến các thông số động lực học (áp suất, lưu lượng, vận tốc, v.v) thuộc hệ thống truyền động thủy lực của thiết bị, nghiên cứu được lực động phát sinh trong quá trình làm việc, mà đây là yếu tố gây ảnh hưởng lớn tới hoạt động của thiết bị khi làm việc trong điều kiện hầm còn chưa được gia cố. Từ các kết quả nghiên cứu thu được sẽ góp phần vào

việc nâng cao độ chính xác trong tính toán thiết kế hệ truyền động thủy lực, giảm thiểu tải trọng động tác dụng vào hệ ray treo, là cơ sở để tự động hóa quá trình làm việc của máy và nâng cao hiệu quả khai thác sử dụng thiết bị.

2. Nội dung

2.1. Giới thiệu về thiết bị hỗ trợ lắp ráp cấu kiện vòm trong thi công đường hầm

Trong thi công hầm, công tác lắp ráp cấu kiện vòm chống đỡ là một khâu then chốt, quyết định sự ổn định của kết cấu và an toàn lao động. Thiết bị hỗ trợ lắp ráp vòm sử dụng hệ thống thủy lực để nâng, định vị và lắp ghép các cấu kiện, giúp cải thiện đáng kể năng suất và độ chính xác so với phương pháp thủ công. Thiết bị bao gồm một số bộ phận chính như sau:

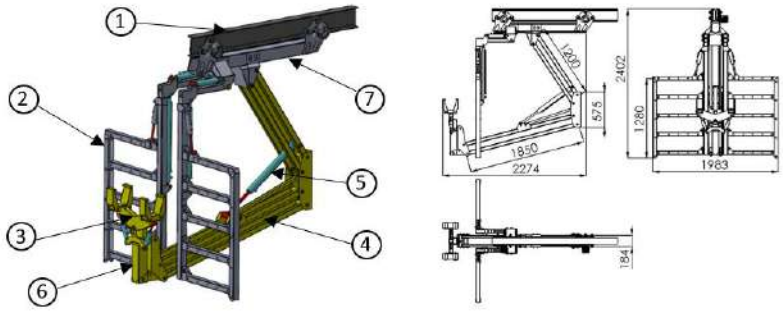
Tay nâng: hệ thống tay nâng gồm tay nâng trên và tay nâng dưới được liên kết bằng tấm liên kết khung, tay nâng được điều khiển bằng một xi lanh thủy lực tay nâng có nhiệm vụ nâng hạ cấu kiện vòm trong quá trình lắp đặt. Bộ phận này thường được chế tạo từ thép cường độ cao và được thiết kế để chịu được tải trọng và ứng suất lớn trong quá trình vận hành.

Tay kẹp: được lắp ở đỉnh tay nâng, được trang bị 4 loại xi lanh thủy lực để thực hiện các thao tác như: tạo góc nghiêng, nâng hạ, tạo góc xoay giá tay kẹp, kẹp giữ cấy kiện vòm trong phạm vi nhỏ.

Khiên chắn đá: được điều khiển một cách cơ động có thể nâng hạ, trượt lên xuống, đóng mở thông qua các xi lanh thủy lực được thiết kế ở phía bên phải và trái của thiết bị như: xi lanh nâng hạ khiên chắn đá, xi lanh dịch chuyển lên xuống khiên chắn đá, xi lanh đóng mở khiên chắn đá.

Khung chính: cung cấp kết cấu chịu lực cho toàn bộ cơ cấu. Nó được lắp trên một bệ dầm di động, cho phép thiết bị di chuyển đến nhiều vị trí khác nhau trong hầm.

Xi lanh thủy lực: được sử dụng để điều khiển chuyển động của tay nâng, tay kẹp tạo ra lực và độ chính xác cần thiết để nâng hạ, tạo góc nghiêng, xoay, kẹp định vị cấu kiện vòm; điều khiển nâng hạ, trượt, đóng mở khiên chắn đá cho phù hợp với điều kiện địa hình làm việc của đường hầm.



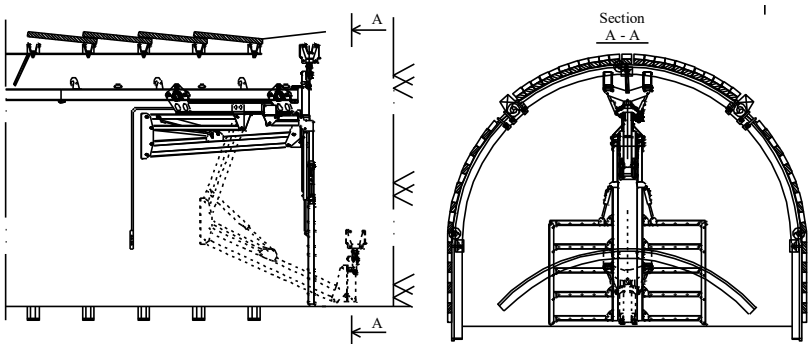
Hình 1. Thiết bị hỗ trợ lắp ráp cấu kiện vòm trong thi công đường hầm
1 - Ray; 2 - Khiên chắn đá; 3 - Tay kẹp; 4 - Tay nâng; 5 - Xi lanh tay nâng;
6 - Xi lanh nâng tay kẹp;7 - Dầm di chuyển.

Hệ thống di chuyển gồm 2 cụm trước và sau giúp thiết bị di chuyển trên ray cố định trên nóc được dẫn động bởi mô tơ thuỷ lực.

Hệ thống truyền động: Hệ thống truyền động cung cấp năng lượng cho các mô tơ và xi lanh thủy lực.

Cấu tạo của thiết bị hỗ trợ lắp ráp cấu kiện vòm trong thi công đường hầm được thể hiện như Hình 1.

Nguyên lý làm việc của thiết bị như sau: Quá trình lắp ráp cấu kiện vòm bằng thiết bị hỗ trợ là bước thứ 7 trong các các bước thi công đường hầm bằng phương pháp NATM. Người thợ vận hành máy đứng gần thiết bị và điều khiển mọi thao tác của máy. Đầu tiên người thợ vận hành điều khiển thiết bị di chuyển trên thanh ray đến vị trí phù hợp để làm việc. Các xi lanh điều khiển cho khiên chắn đá hạ và trượt xuống, sau đó mở ra để bảo vệ an toàn cho công nhân trong quá trình thi công. Tiếp tục, tay nâng trên và dưới được hạ xuống nhờ xi lanh nâng hạ tay nâng đưa tay kẹp đến vị trí chờ của cấu kiện vòm đã được chuẩn bị trước. Sau đó, người thợ vận hành điều khiển các xi lanh như xi lanh tạo góc nghiêng giá tay kẹp, xi lanh kẹp giữ cấu kiện vòm, xi lanh tay nâng, xi lanh nâng hạ tay kẹp, xi lanh tạo góc xoay tay kẹp mà người vận hành có thể dễ dàng tuỳ chỉnh, phối hợp nhịp nhàng, chính xác, kẹp giữ chắn chắn, di chuyển cấu kiện vòm đến vị trí cần lắp ghép. Sau khi thực hiện xong nhiệm vụ lắp ráp, tay nâng, tay kẹp và khiên chắn đá được thu và gập lại gọn gàng lên phía trên đỉnh vòm và được di chuyển về phía sau qua hệ thống thanh ray được gắn phía trên vòm tạo ra không gian làm việc và không gây ảnh hưởng đến quá trình thi công tiếp theo của công nhân. Cứ như vậy, các thanh ray được lắp nối tiếp nhau cho phép thiết bị có thể di chuyển đến vị trí khác làm nhiệm vụ.



Hình 2. Quá trình lắp ráp cấu kiện vòm bằng thiết bị hỗ trợ trong thi công hầm

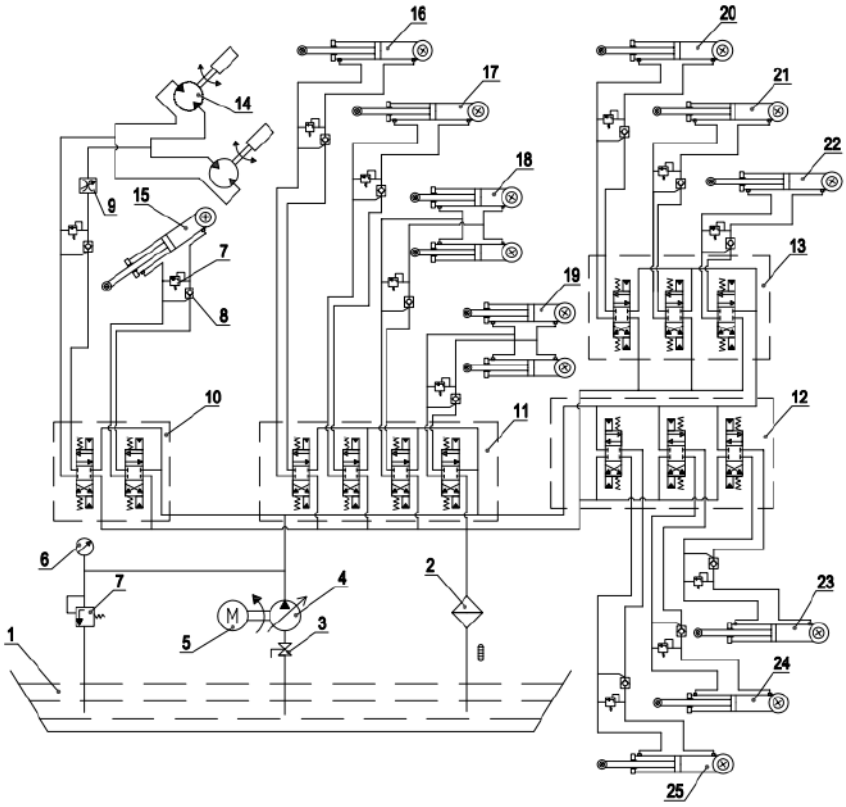
Thiết bị hỗ trợ lắp ráp cấu kiện vòm trong thị công đường hầm được tối ưu hóa để sử dụng trong các hầm giao thông và hầm khai thác than. Cơ cấu này có kích thước nhỏ gọn và dễ cơ động, cho phép hoạt động trong các không gian hạn chế. Ngoài ra, thiết bị cũng được thiết kế để đảm bảo độ bền và độ tin cậy cao, dễ lắp ráp phù hợp với điều kiện làm việc khắc nghiệt trong môi trường hầm mỏ.

2.2. Nghiên cứu động lực học hệ TĐTL dẫn động hệ thống xi lanh nâng tay kẹp của thiết bị

Căn cứ vào các chức năng làm việc, công năng của các bộ phận cần thiết trang bị trên thiết bị, tác giả dựng sơ đồ mạch thủy lực điều khiển các xi lanh, mô tơ thủy lực như trên Hình 3.

Nguyên lý hoạt động của mạch thủy lực: Bơm dầu (4) được dẫn động nhờ động cơ (5) cung cấp dầu cao áp cho xi lanh thủy lực thực hiện các nhiệm vụ như nâng hạ tay nâng, tạo góc nghiêng giá tay kẹp, nâng hạ, tạo góc xoay tay kẹp, kẹp giữ cấu kiện vòm đồng thời điều khiển chuyển động nâng hạ, lên xuống, đóng mở của khiên chắn đá. Việc thay đổi chiều hoạt

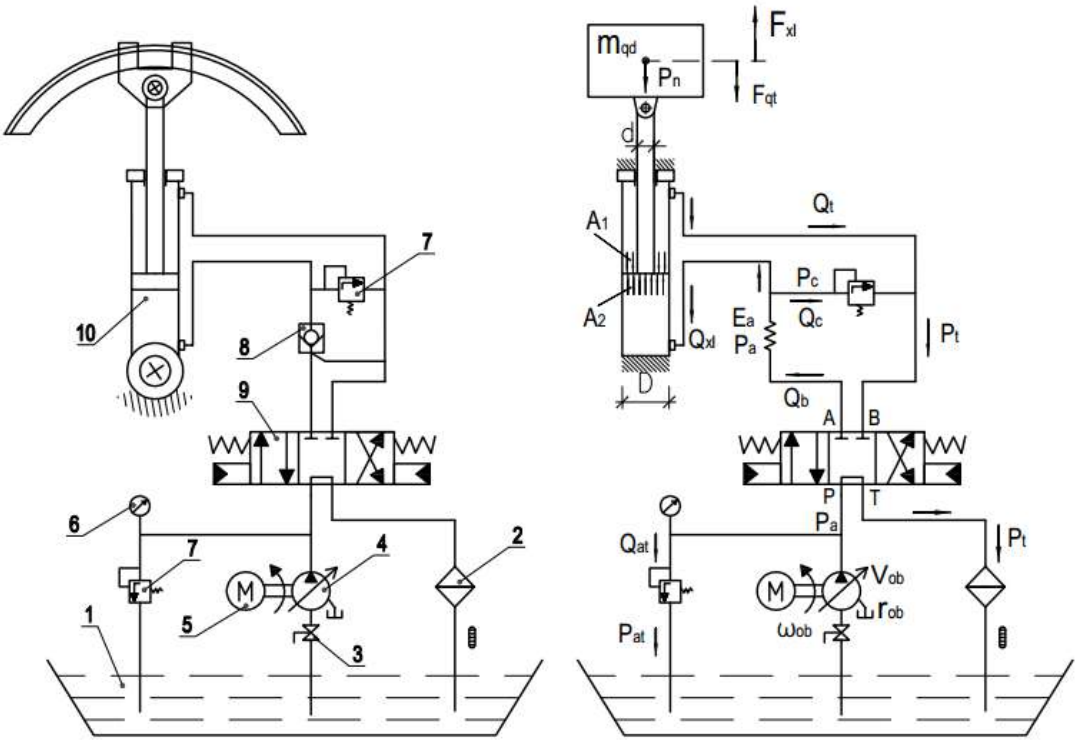
động của xi lanh nhờ vào cụm van phân phối (10, 11, 12, 13), van này được điều khiển nhờ tín hiệu thủy lực. Để thiết bị có thể di chuyển trên hệ thống thanh ray được gắn phía trên vòm nhờ vào mô tơ thuỷ lực (14), nó cũng được cung cấp dầu từ bơm dầu (4). Van tiết lưu (9) được dùng để điều khiển tốc độ di chuyển của thiết bị trên thanh ray. Để đảm bảo an toàn cho mạch thủy lực, xi lanh, mô tơ trong quá trình làm việc, sử dụng các van an toàn (7).



Hình 3. Sơ đồ mạch thủy lực của thiết bị

1- Thùng dầu; 2- Lọc dầu; 3- Van cấp dầu; 4- Bơm thủy lực; 5- Động cơ điện;
6- Đồng hồ áp suất; 7- Van an toàn; 8- Van một chiều có điều khiển; 9- Van tiết lưu;
10- Cụm van điều khiển số 1; 11- Cụm van điều khiển số 2; 12- Cụm van điều khiển số 3;
13- Cụm van điều khiển số 4; 14- Motor thủy lực; 15- Xi lanh nâng hạ tay nâng;
16- Xi lanh tạo góc nghiêng giá tay kẹp;17- Xi lanh nâng hạ tay kẹp; 18- Xi lanh tạo góc xoay tay kẹp;
19- Xi lanh kẹp giữ cấu kiện vòm; 20, 23- Xi lanh nâng hạ khiên chắn đá (trái, phải);
21, 24- Xi lanh dịch chuyển trượt lên xuống khiên chắn đá (trái, phải);
22, 25- Xi lanh đóng mở khiên chắn đá (trái, phải).

Hệ thống truyền động thủy lực (TĐTL) dẫn động hệ thống xi lanh nâng tay kẹp của thiết bị hỗ trợ lắp ráp cấu kiện vòm như trên Hình 4, nguyên lý làm việc như sau: Bơm dầu (4) được dẫn động nhờ động cơ (5) cung cấp dầu cao áp cho xi lanh thủy lực thực hiện nâng, hạ tay cấu kiện vòm nhờ đó mà cấu kiện vòm được nâng lên, hạ xuống. Việc thay đổi chiều hoạt động của xi lanh nhờ vào van phân phối (9), van này được điều khiển nhờ tín hiệu thủy lực. Để đảm bảo an toàn cho mạch thủy lực và xi lanh trong quá trình làm việc sử dụng các van an toàn (7).



Hình 4. Sơ đồ hệ TĐTL dẫn động xi lanh nâng tay kẹp của thiết bị
1- Thùng dầu; 2- Lọc dầu; 3- Van cấp dầu; 4- Bơm thủy lực; 5- Động cơ điện;
6- Đồng hồ áp suất; 7- Van an toàn; 8- Van một chiều có điều khiển;
9- Van phân phối; 10- Xi lanh nâng tay kẹp

Tác giả đưa ra các giả thiết như sau: Không xét quá trình tạo sóng trong ống dẫn dầu thủy lực; Môđuyn dầu đàn hồi của các ống dẫn không phụ thuộc vào áp suất trong hệ thống; Tổn thất năng lượng trong hệ thống được tính thông qua ma sát nhớt, ma sát khô và mất mát thể tích của bơm thủy lực; Lực cản nội ma sát của xi lanh được kể đến qua hiệu suất cơ khí ηc; Không xét đến quán tính của chất lỏng trong quá trình làm việc; Các thông số của chất lỏng làm việc (tỷ trọng riêng, độ nhớt, mô đuyn đàn hồi thể tích) là hằng số.

Sử dụng định luật bảo toàn năng lượng, viết cho phương trình dòng chảy liên tục của dầu công tác trong các đường ống và phương trình cân bằng lực để xây dựng mô hình động lực học hệ thống TĐTL.

) Phương trình dòng chảy liên tục trong đường ống cao áp:

$$(Q_b - Q_{rrb}) - Q_E - Q_{xl} - Q_{at} = 0 \quad (1)$$

$$\text{Hay } Q_E = (Q_b - Q_{rrb}) - Q_{xl} - Q_{at} \quad (2)$$

Trong đó:

Q_b- Lưu lượng lý thuyết của bơm,

$$Q_b = V_{ob}\omega_b X(t) \ , (m^3/s) \quad (3)$$

Q_{rrb} - Lưu lượng chất lỏng rò rỉ ở bơm thủy lực:

$$Q_{rrb} = r_{ob} p_a \ (m^3/s) \quad (4)$$

Hệ số tổn thất lưu lượng ở bơm thủy lực được xác định r_{ob}, [1]:

$$r_{ob}=\frac{V_{ob}[\omega_b](1-\eta_{ob})}{[p_b]} \quad (5)$$

Q_{xl} - Lưu lượng tiêu thụ của xi lanh:

$$Q_{xl} = A_2.v_n \ (\mathrm{m^3/s}) \quad (6)$$

Q_E- Lưu lượng chất lỏng làm biến dạng hệ thống:

$$Q_E = E_a \frac{dp_a}{dt} \ (\mathrm{m^3/s}) \quad (7)$$

Với: E_a - Biến dạng đàn hồi trong ống dẫn cao áp, biến dạng này được xác định như sau:

$$E_a=\frac{V_1}{E_1}+\frac{V_2}{E_2} \quad (8)$$

Thể tích dầu công tác trong đường ống dẫn bằng cao su được xác định:

$$V_1=\frac{\pi d_c^2}{4} l_c \quad (9)$$

Thể tích dầu công tác trong đường ống dẫn bằng kim loại được xác định:

$$V_2 = \frac{\pi d_k^2}{4} l_k \quad (10)$$

Q_{at} - Lưu lượng qua van an toàn tổng, nếu $p_a \geq p_{at}$ thì $Q_{at} = (p_a - p_{at})k_{at}$; ngược lại $Q_{at} = 0$;

Thay vào phương trình (2) chúng ta thu được:

$$E_a \frac{dp_a}{dt} = \left(V_{ob}\omega_b X(t) - \frac{V_{ob}[\omega_b](1-\eta_{ob})}{[p_b]} p_a \right) - A_2.v_n - (p_a - p_{at})k_{at} \quad (11)$$

) Phương trình cân bằng lực trên xi lanh tay nâng được xác định như sau:

$$F_{qt} = F_{x\ln} - P_n \quad (12)$$

Trong đó:

F_{qt}: Lực quán tính,

$$F_{qt} = m_{qd} \frac{dv_n}{dt}, \ (\mathrm{N}) \quad (13)$$

$F_{x\ln}$: Lực nâng do xi lanh thủy lực tạo ra,

$$F_{x\ln} = (p_a.A_2 - p_t.A_1).\eta_c, \ (\mathrm{N}) \quad (14)$$

$P_n = m_{qd}.g$: Lực tác dụng vào cán pít tông, N.

$m_{qd} = 350 + 34,8 = 384,8 \ (kg)$: Khối lượng quy dẫn bằng tổng khối lượng của cấu kiện vòm và tay kẹp.

Vậy hệ phương trình vi phân mô tả hoạt động của hệ thống TĐTL dẫn động xi lanh thủy lực tay nâng như sau:

$$\begin{cases} E_a \dfrac{dp_a}{dt} = \left(V_{ob}\omega_b X(t) - \dfrac{V_{ob}[\omega_b](1-\eta_{ob})}{[p_b]}\, p_a \right) - A_2.v_n - (p_a - p_{at})k_{at} \\ m_{qd} \dfrac{dv_n}{dt} = (p_a.A_2 - p_t.A_1).\eta_c - P_n \end{cases} \quad (15)$$

Trong đó:

V_{ob}- Lưu lượng riêng của bơm, m^3/vòng;

ω_b- Tốc độ quay của trục bơm và động cơ, vòng/s;

X(t) - Hệ số điều chỉnh lưu lượng của bơm, $0 \le X(t) \le 1$;

$[\omega_b]$- Tốc độ quay danh nghĩa của bơm, vòng/s;

$[p_b]$- Áp suất danh nghĩa của bơm, Pa;

p_a- Áp suất của dầu công tác trong nhánh cao áp (nhánh A),Pa;

p_{at}- Áp suất cài đặt của van an toàn tổng, Pa;

k_{at}- Hệ số lưu lượng qua van an toàn tổng, (m^3/s)/Pa;

η_{ob}- Hiệu suất thể tích của bơm thủy lực;

η_c- Hiệu suất cơ khí của xi lanh thủy lực;

r_b- Hệ số tổn thất lưu lượng ở bơm thủy lực, (m^3/s)/Pa;

V_1 , V_2- Thể tích dầu công tác trong đường ống dẫn bằng cao su và kim loại, m^3;

E_1, E_2- Mô đuyn biến dạng đàn hồi của đường ống dẫn bằng cao su và kim loại, Pa;

d_c, d_k- Đường kính trong của ống dẫn dầu bằng cao su và kim loại, m;

l_c, l_k- Chiều dài đường ống dẫn dầu bằng cao su và kim loại,m;

m_{qd}- Khối lượng quy dẫn của piston tay nâng, tay kẹp và vòm thép về đỉnh piston, kg;

A_1– Diện tích tiết diện khoang thấp áp của xi lanh tay nâng,m^2;

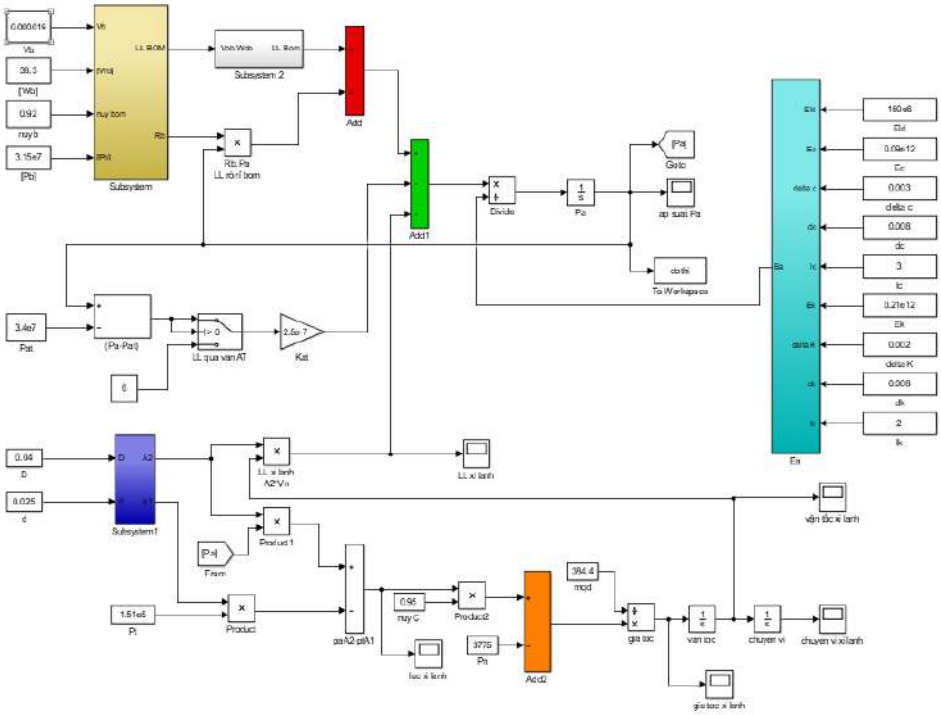
A_2– Diện tích tiết diện khoang cao áp của xi lanh tay nâng, m^2.

Nhận xét:

Hệ phương trình vi phân (15) là mô hình toán mà trong đó thể hiện đầy đủ các thông số động lực học của hệ thống truyền động thủy lực dẫn động xi lanh nâng hạ tay kẹp. Việc giải hệ phương trình vi phân giúp đề tài xác định được áp suất dầu trong mạch cao áp và vận tốc, gia tốc của xi lanh nâng hạ tay kẹp thay đổi theo thời gian. Ngoài ra, đề tài còn sử dụng mô hình toán này để nghiên cứu, khảo sát và đánh giá mức độ ảnh hưởng của các thông số động lực học hệ truyền động thủy lực dẫn động xi lanh nâng tay kẹp.

Sử dụng phương pháp Runge-Kutta trong môi trường Matlab- Simulink để giải hệ phương trình vi phân (15). Giao diện của chương trình được thể hiện như trên Hình 5.

Để nghiên cứu khảo sát đánh giá các thông số động lực học của hệ TĐTL dẫn động xi lanh tay nâng trên máy, tác giả tiến hành tính toán xác định các thông số đầu vào của bài toán, kết quả tính toán thu được như trong Bảng 1.



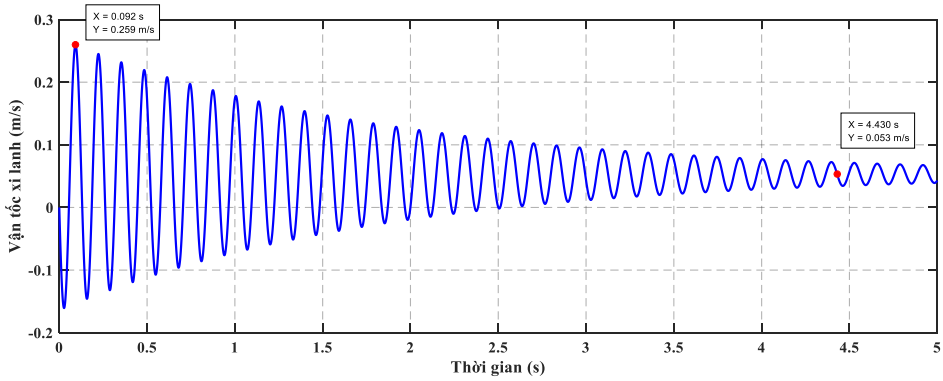
Hình 5. Chương trình mô phỏng động lực học hệ TĐTL dẫn động nâng tay kẹp

Bảng 1. Bảng các giá trị thông số chạy chương trình

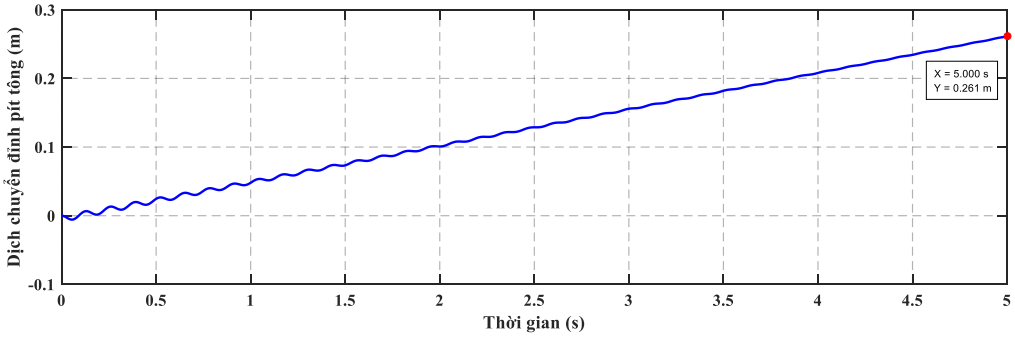
TT	Thông số	Ký hiệu	Đơn vị	Giá trị
	Bơm thuỷ lực			
1	Lưu lượng riêng của bơm	V_b	m^3/vòng	0,000019
2	Tốc độ quay danh nghĩa của bơm	$[n_b]$	vòng/s	38,5
3	Áp suất danh nghĩa của bơm	$[p_b]$	Pa	3,15e7
4	Hiệu suất thể tích của bơm	η_b	-	0,92
5	Hệ số điều chỉnh lưu lượng bơm	X(t)	-	0 - 1
	Xi lanh thủy lực			
6	Đường kính trong của xi lanh	D	m	0,04
7	Đường kính cán pít tông	d	m	0,025
8	Áp lực dầu trong khoang cao áp	P_1	Pa	24,4.e6
9	Áp lực dầu trong khoang thấp áp	P_2	Pa	0,9.e6
10	Hiệu suất truyền động cơ khí của xi lanh thủy lực	η_c	-	0,95
	Van an toàn của mạch			
11	Áp suất van an toàn tổng	p_{at}	Pa	$3,4.10^7$
12	Hệ số lưu lượng qua van an toàn tổng	k_{at}	$(m^3/s)/Pa$	6,67e-10
13	Lưu lượng qua van an toàn tổng	Q_{at}	m^3/s	0,004
	Van an toàn của xi lanh thủy lực			

14	Áp suất van an toàn	Pc	Pa	2,44e7
15	Hệ số lưu lượng qua van	Kc	(m3/s)/Pa	3,33e-10
	Đường ống dẫn dầu thủy lực			
16	Chiều dài đường ống bằng cao su	l_c	m	3,0
17	Chiều dài đường ống bằng kim loại	l_k	m	2,0
18	Đường kính trong của đường ống cao su	d_c	m	0,008
19	Đường kính trong của đường ống kim loại	d_k	m	0,008
20	Mô đuyn đàn hồi của đường ống làm bằng cao su	E_c	Mpa	0,09.e6
22	Mô đuyn đàn hồi của đường ống làm bằng kim loại	E_k	Mpa	$0,21.10^6$
23	Chiều dày của đường ống cao su	δ_c	m	0,003
24	Chiều dày của đường ống kim loại	δ_k	m	0,002
	Dầu công tác			
25	Mô đuyn đàn hồi của dầu công tác	E_{lq}	Mpa	150
	Các thông số khác			
26	Lực tác dụng vào cán pít tông	P_n	N	3775
27	Khối lượng quy dẫn tay kẹp và cấu kiện vòm	m_{qd}	kg	384,8

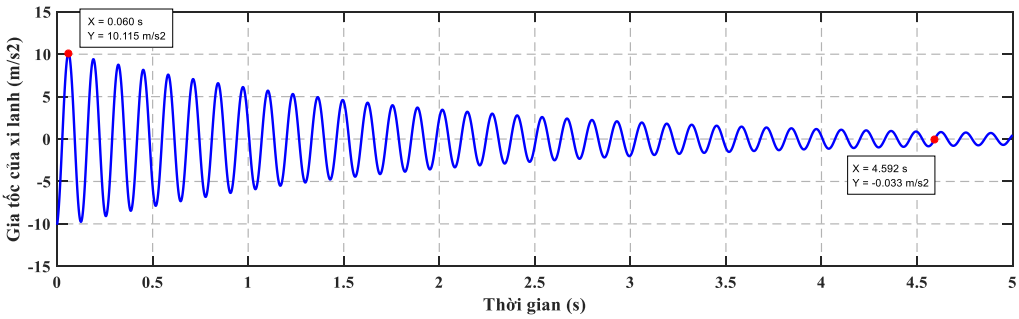
Sau khi chạy chương trình mô phỏng trên Hình 5 với các thông số trong Bảng 1, chạy chương trình chúng ta thu được các kết quả sau:



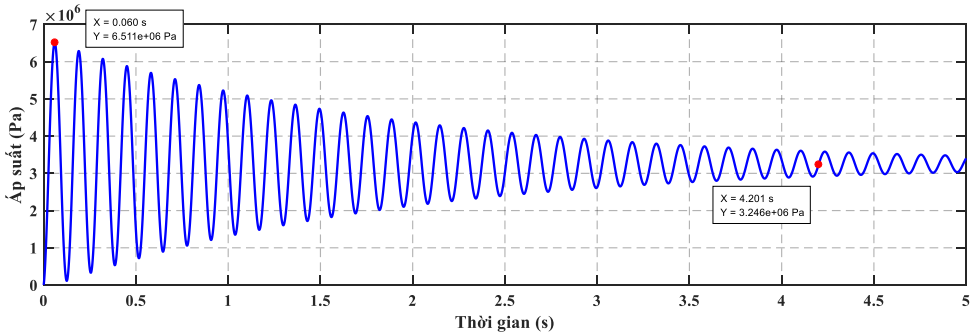
Hình 6. Vận tốc xi lanh nâng tay kẹp



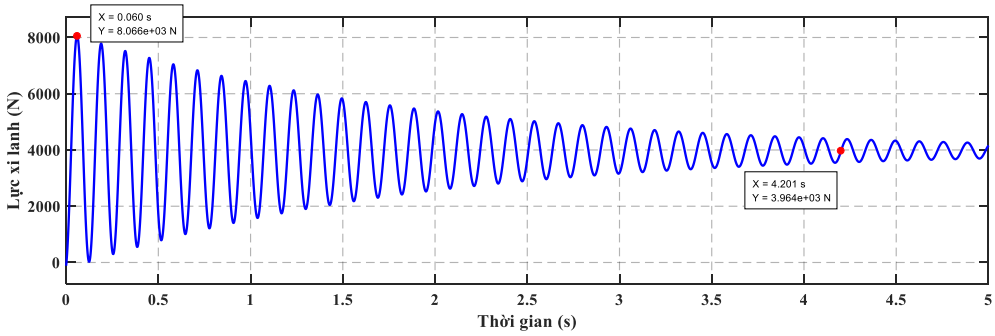
Hình 7. Dịch chuyển đỉnh pít tông nâng tay kẹp



Hình 8. Gia tốc xi lanh nâng tay kẹp



Hình 9. Áp suất xi lanh nâng tay kẹp



Hình 10. Lực xi lanh nâng tay kẹp tạo ra

Nhận xét: Trong quá trình nâng cấu kiện vòm, với các kết quả ở trên có thể thấy các thông số động lực học của xi lanh thủy lực thay đổi nhiều trong giai đoạn bắt đầu thực hiện nâng, sau đó bình ổn về giá trị trị trung bình của nó, cụ thể như sau:

Áp suất trong mạch cao áp thay đổi và giá trị áp suất đạt giá trị lớn nhất cụ thể: áp suất động là $6{,}511.10^6$ Pa tại thời điểm bắt đầu thực hiện nâng (Hình 9).

Vận tốc và gia tốc xi lanh nâng hạ cấu kiện vòm thay đổi theo thời gian. Vận tốc nâng cấu kiện vòm có giá trị lớn nhất 0,259 m/s khi khởi động (Hình 2.15), sau đó dao động ổn định về các giá trị trung bình. Gia tốc nâng cấu kiện vòm cũng thay đổi và dao động bình ổn xung quanh giá trị 0 m/s2 (Hình 8).

Lực nâng của xi lanh thủy lực cũng thay đổi ứng với giá trị ngoại lực tác dụng vào xi lanh. Lực nâng phát sinh lớn nhất khi khởi động, giá trị lúc này là $8{,}066.10^3$ N (Hình 10).

3. Kết luận

- Khi máy làm việc thì áp suất của dầu công tác trong hệ thống TĐTL sẽ dao động. Quá trình dao động của áp suất dầu phụ thuộc vào nhiều yếu tố, trong đó trọng lượng của cấu kiện

vòm tác dụng vào xi lanh là yếu tố tác động lớn đến áp suất dầu công tác và có thể nói, áp suất là thông số "đại diện" cho độ lớn của tải bên ngoài.

- Dao động của vận tốc và gia tốc của các xi lanh có biên độ và tần số lớn ở giai đoạn khởi động nâng tay nâng hay nâng cấu kiện vòm phát sinh lực động ảnh hưởng lớn đến quá trình làm việc, ảnh hưởng đến chất lượng và tuổi thọ của các thiết bị thủy lực. Để khắc phục hiện tượng này, chúng ta cần bố trí cụm van tiết lưu một chiều.

- Áp suất dầu công tác và vận tốc của xi lanh thủy lực là thông số cơ bản nhất trong hệ thống thủy lực điều khiển các xi lanh. Việc xác định được hai thông số này có ý nghĩa rất lớn, vì từ đó cho phép chúng ta xác định được các thông số dẫn xuất như lực đẩy của xi lanh, công suất tiêu thụ của xi lanh, lưu lượng cần thiết cho xi lanh khi làm việc.

- Kết quả nghiên cứu đã khẳng định được độ tin cậy của mô hình ĐLH, phương pháp tính và công cụ thực hiện. Từ đó, tạo cơ sở cho việc sử dụng mô hình ĐLH trong việc khảo sát các yếu tố ảnh hưởng đến các thông số ĐLH (áp suất, lưu lượng, lực động, ...) của hệ thống TĐTL. Các kết quả nghiên cứu như trên phần nào đã đánh giá được mức độ ảnh hưởng của các thông số động lực học đến sự thay đổi của áp suất dầu công tác, vận tốc, gia tốc và lực của xi lanh thủy lực trong quá trình nâng tay nâng và nâng cấu kiện vòm. Tuy nhiên, để có thể đánh giá chính xác mức độ ảnh hưởng cũng như kiểm nghiệm các kết quả tính toán lý thuyết chúng ta cần phải có những nghiên cứu thực nghiệm nhằm đánh giá mức độ chính xác của các kết quả nghiên cứu lý thuyết.

- Các kết quả nghiên cứu là cơ sở khoa học góp phần vào việc nâng cao độ tin cậy trong tính toán thiết kế hệ TĐTL và là cơ sở để tự động hóa quá trình làm việc của toàn bộ cơ cấu nâng và kẹp cấu kiện vòm của thiết bị hỗ trợ lắp ráp cấu kiện vòm phù hợp với điều kiện thi công trong hầm giao thông và hầm mỏ.

Tài liệu tham khảo

1. Trần Quang Hùng, Đỗ Doãn Phi, Lê Trọng Cường, Trần Hữu Lý (2013), Truyền động thủy lực trên Xe máy Công binh, Học viện kỹ thuật Quân sự.
2. Tien Dung Le and Xuan Nam Bui, 2019. "Status and prospects of underground coal mining technology in Vietnam. Journal of the Polish Mineral Engineering Society". Doi 10.29227/IM-2019-02-63.
3. Long R, Sun S, Lian Z, 2015. "Research on the hard-rock breaking mechanism of hydraulic drilling impact tunneling", Mathemat Prob Eng, https://doi.org/10.1155/2015/153648.
4. Duong, D.H., Dinh, V.C., Koteras, A., Do, V.H., Vu, B.T, 2020. "Development orientation of mechanized technology of underground mining in Quangninh coalfield, Vietnam". In: AIP Conference Proceedings, vol. 2209, 2020, https://doi.org/10.1063/5.0000005.
5. Nguyen Phi Hung, Nguyen Ngoc Trung, Nguyen Thuy Chi, 2023. "Calculation and design of lifting and lowering mechanism for supporting beams in the construction of tunnel support excavation at Mong Duong Coal Joint Stock Company – Vinacomin", Research, implementation and technology transfer code 389/HDKT of the Center for Mining and Environmental Science and Technology, University of Mining and Geology and Mong Duong Coal Joint Stock Company, Vietnam, 2023.

6. Seo J, Dae-Kyung Noh, Geun-Ho Lee & Joo-Sup Jang, 2016. “A percussion performance analysis for rock-drill drifter through simulation modeling and experimental validation”. International Journal of Precision Engineering and Manufacturing, Vol 17, pp163–170.
7. МайоровЮ.П., КовальскийВ.Ф., ДубровинВ.А., ГрунинЕ.И (2005), Расчет параметров переходных процессов гидравлических приводов с объёмным регулированием скорости - М.: МИИТ, г., 60 с.

Study on the dynamics of the hydraulic drive system actuating the lifting mechanism of the gripping arm in equipment supporting the assembly of arch segments in tunnel construction

Abstract: *Along with the strong development of infrastructure projects, the construction of transportation tunnels and mining tunnels in Vietnam has been increasingly promoted. Although tunnel excavation technology has achieved a high level of mechanization, the installation of steel arch frames still largely depends on manual labor, leading to limitations in productivity, accuracy, and labor safety. Therefore, the research and application of supporting equipment capable of lifting, holding, and positioning steel arch segments within the confined space of tunnels are necessary.*

This paper presents the research results on the dynamic analysis of the hydraulic transmission system driving the lifting–gripping mechanism in the supporting device for steel arch assembly in tunnel construction. The research results provide a scientific basis for the calculation and rational design of the hydraulic transmission system, contributing to the reduction of dynamic loads, promoting automation of the working process, and improving the operational efficiency of the equipment under practical conditions.

Keywords: hydraulic transmission system; dynamic analysis; lifting–gripping mechanism; steel arch segment; tunnel.

Nghiên cứu tối ưu vị trí công tác của thiết bị chuyên dùng phục vụ công tác kiểm tra, sửa chữa bề mặt dưới của các công trình cầu.

Bùi Văn Luân[1*], Lê Văn Dưỡng[2], Bùi Văn Trầm[3]

[1]Hệ quản lý Học viên sau đại học, Học viện Kỹ thuật Quân sự

[2]Viện Cơ khí động lực, Học viện Kỹ thuật Quân sự

[3] Trường Đại học công nghệ Giao thông vận tải

Email: k17abuiluan@gmail.com; Tel:0988532560

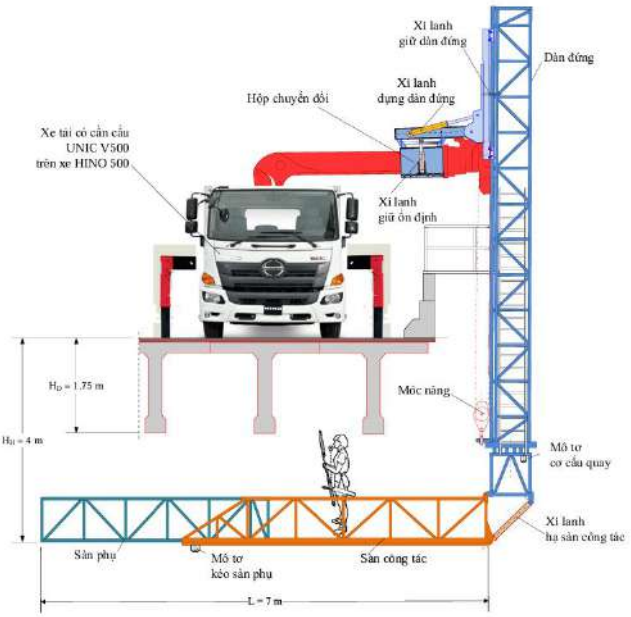
Tóm tắt

Thiết bị chuyên dụng phục vụ kiểm tra, sửa chữa, bảo trì mặt dưới của các công trình cầu làm việc cơ động giúp đảm bảo an toàn và độ tin cậy trong khai thác công trình giao thông. Hướng tới mục tiêu tối ưu hóa vị trí công tác trong thực tế, bài báo tập trung giải quyết bài toán nền tảng là xây dựng mô hình toán học và phân tích đặc tính động học của thiết bị. Trước tiên, mô hình động học được thiết lập dựa trên phương pháp ma trận biến đổi thuần nhất Denavit–Hartenberg (D-H). Từ đó, ma trận Jacobian tổng quát được xây dựng để biểu diễn mối quan hệ vận tốc giữa các khớp và khâu tác động cuối. Dựa trên các phương trình động học, nghiên cứu tiến hành mô phỏng số học để xác định chính xác biên dạng không gian làm việc 3D của thiết bị. Kết quả đánh giá chỉ ra bán kính làm việc hiệu dụng của cơ cấu, đồng thời xác định rõ các giới hạn về tầm với hình học. Đây là cơ sở dữ liệu cốt lõi để đề xuất phương án định vị xe cơ sở an toàn, làm tiền đề cho quá trình khai thác và vận hành thiết bị hiệu quả.

Từ khóa: *Phân tích động học; Không gian làm việc; Thiết bị kiểm tra dưới cầu; Tham số D-H; Ma trận Jacobian.*

1. Mở đầu

Trong quá trình khai thác, việc kiểm tra định kỳ gầm cầu là công tác tiên quyết để phát hiện hư hỏng và đảm bảo an toàn kết cấu [1]. Các phương pháp truyền thống như dùng giàn giáo hay đu dây bộc lộ nhiều hạn chế về an toàn lao động và chi phí, thúc đẩy xu hướng sử dụng thiết bị tự hành chuyên dụng [2]. Mặc dù các giải pháp sử dụng Thiết bị bay không người lái (UAV) hoặc robot leo cáp đã được nghiên cứu rộng rãi [3, 4], nhưng các cụm thiết bị công tác chuyên dụng gắn trên xe cơ sở vẫn duy trì ưu thế vượt trội về tính ổn định và khả năng chịu tải phục vụ cho các công tác bảo trì phức tạp.



Hình 1. Thiết bị kiểm tra gầm cầu chế tạo tại Việt Nam

Việc làm chủ công nghệ thiết bị kiểm tra cầu trong nước (Hình 1) đòi hỏi phải phân tích chính xác đặc tính động học [5,6]. Đây là cơ sở cốt lõi để hiểu rõ khả năng đáp ứng của

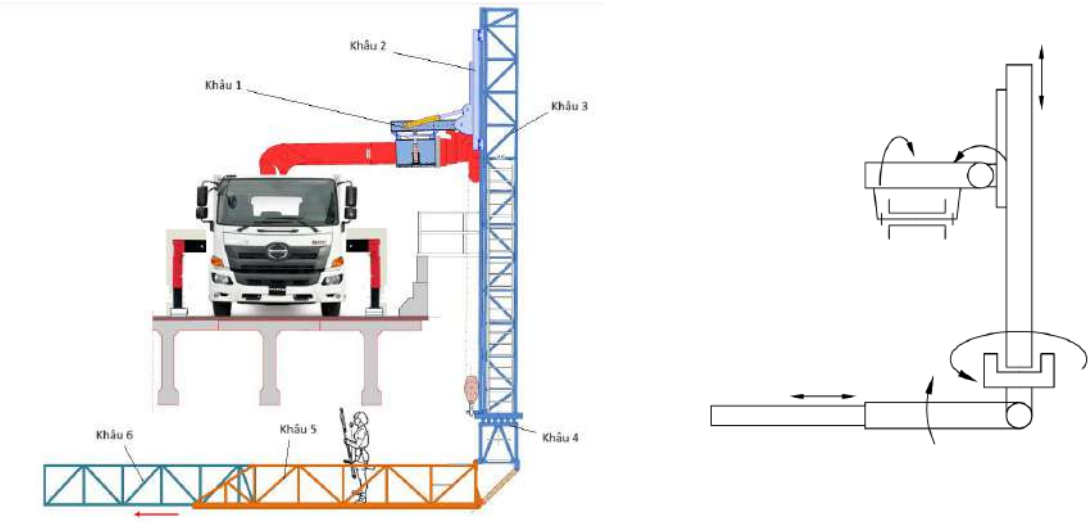
* Email: k17abuiluan@gmail.com

thiết bị đối với các biên dạng cầu khác nhau. Trong thực tế khai thác, việc không nắm rõ giới hạn không gian làm việc có thể dẫn đến rủi ro thiết bị va chạm với kết cấu cầu hoặc bỏ sót các "vùng mù" kiểm tra do vượt quá tầm với. Do đó, bài báo này trình bày quá trình xây dựng mô hình động học của thiết bị, từ đó tiến hành mô phỏng để đánh giá toàn diện không gian làm việc và xác định bán kính hoạt động hiệu quả. Các kết quả nghiên cứu sẽ là cơ sở khoa học quan trọng để đề xuất các phương án định vị xe an toàn và khai thác tối đa năng lực của thiết bị.

2. Mô hình động học thiết bị

2.1 Giả thiết

Để tạo điều kiện thuận lợi cho việc điều khiển chính xác và phân tích không gian làm việc, thiết bị được mô hình hóa như một chuỗi động học bao gồm 6 khâu. Các khâu được minh họa trong Hình 2.



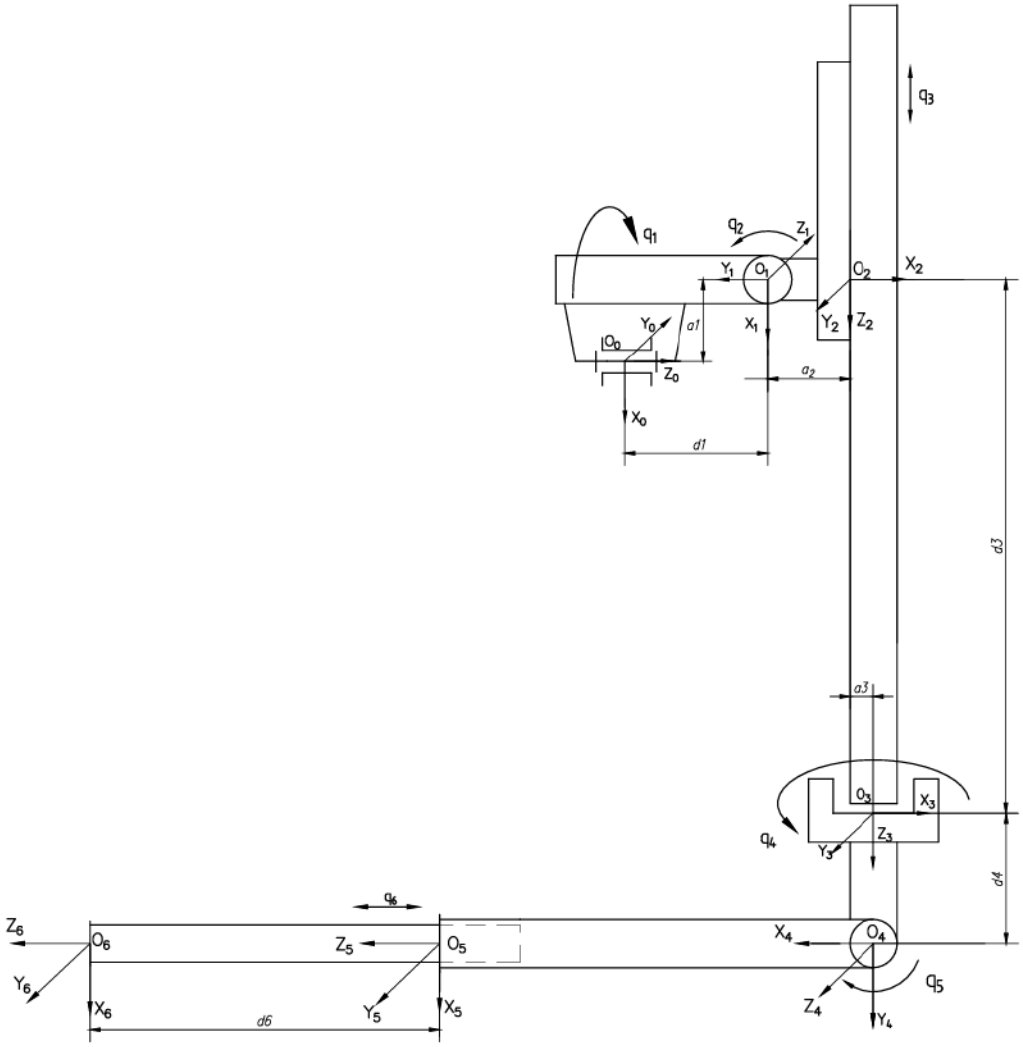
Hình 2. Sơ đồ hóa mô hình khảo sát động học

Các khâu của thiết bị được coi là các vật rắn tuyệt đối, bỏ qua sự biến dạng đàn hồi của vật liệu dưới tác dụng của tải trọng hay nhiệt độ trong quá trình làm việc. Các khâu được liên kết nối tiếp với nhau thông qua các khớp quay và khớp tịnh tiến lý tưởng. Bỏ qua ảnh hưởng của lực ma sát, độ rơ lắc và khe hở cơ khí tại các vị trí khớp nối. Trong suốt quá trình khảo sát động học của chuỗi tay máy, bệ đỡ thiết bị (xe cơ sở) được giả định là đứng yên tuyệt đối, không xảy ra hiện tượng trượt hay rung lắc. Hệ tọa độ gốc gắn liền với bệ đỡ được xem là hệ quy chiếu quán tính cố định.

2.2 Hệ tọa độ và tham số D-H

Tham số D-H: Các hệ trục tọa độ được gán cho từng khâu theo quy ước Denavit-Hartenberg (D-H). Phép biến đổi hình học giữa các khâu liên tiếp được xác định bởi bốn tham số: chiều dài khâu (a_i), độ lệch khâu (d_i), góc xoắn (α_i) và góc khớp (θ_i). Ma trận Jacobian: Mối quan hệ giữa vận tốc các khớp và vận tốc của khâu tác động cuối được thiết lập thông qua ma trận Jacobian hình học (J).

Mô hình động học khi gắn các hệ trục tọa độ được thể hiện trên hình 3.



Hình 3. Mô hình động học của thiết bị kiểm tra bề mặt dưới cầu

Dựa trên phương pháp Denavit–Hartenberg (D-H), các tham số động học được thiết lập (Bảng 1) để xây dựng các ma trận biến đổi tọa độ.

Bảng 1. Bảng tham số D-H

i	θ_i	d_i	a_i	α_i	Biến khớp
1	q_1	0,577	0,285	-90^0	q_1
2	q_2	0	0,384	-90^0	q_2
3	0	d_3	0,30	0^0	d_3
4	q_4	0,6	0	90^0	q_4
5	q_5	0	0	90^0	q_5
6	0	d_6	0	0^0	d_6

Giới hạn làm việc của từng khớp được xác định dựa trên tài liệu thiết kế và điều kiện vận hành an toàn, các giới hạn được sử dụng như sau:

Khớp quay lắc ngang: $q_1 \in [-180^{\circ}; 180^{\circ}]$

Khớp quay dựng dàn đứng: $q_2 \in [-90^{\circ}; 90^{\circ}]$

Khớp trượt dàn đứng: $d_3 \in [0; 4,0]$ m

Khớp quay sàn công tác: $q_4 \in [-180^{\circ}; 180^{\circ}]$

Khớp đóng mở sàn công tác: $q_5 \in [-180^{\circ}; 180^{\circ}]$

Khớp trượt sàn kéo dài: $d_6 \in [0; 7,0]$ m

Từ bảng tham số D-H ta xác định được các ma trận truyền $\boldsymbol{H}_i \quad i = 1,2,\ldots,6$ cỡ $4x4$ biến đổi tọa độ từ một điểm trên hệ tọa độ khâu $(i-1)$ $x_{i-1}y_{i-1}z_{i-1}$ tới khâu i $x_i y_i z_i$ có dạng:

$$^{i-1}\boldsymbol{H}_i = \begin{bmatrix} \cos\theta_i & -\sin\theta_i \cos\alpha_i & \sin\theta_i \sin\alpha_i & a_i \cos\theta_i \\ \sin\theta_i & \cos\theta_i \cos\alpha_i & -\cos\theta_i \sin\alpha_i & a_i \sin\theta_i \\ 0 & \sin\alpha_i & \cos\alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{1}$$

Thay giá trị vào 1 ta thu được các ma trận:

$$^{0}H_1 = \begin{bmatrix} \cos q_1 & 0 & -\sin q_1 & a_1 \cos q_1 \\ \sin q_1 & 0 & \cos q_1 & a_1 \sin q_1 \\ 0 & -1 & 0 & d_1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \ldots\ldots, ^{5}H_6 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & d_6 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Các ma trận biến đổi D_i giữa hệ tọa độ gốc x_0, y_0, z_0 và khâu i x_i, y_i, z_i được cho như sau:

$$\boldsymbol{D}_i = \prod_{k=1}^{i} \boldsymbol{H}_k \,, \tag{2}$$

Các ma trận biến đổi này có dạng như sau:

$$\mathbf{D}_n = \begin{bmatrix} \mathbf{A}_n & \mathbf{r}_E^{(0)} \\ \mathbf{0}^T & 1 \end{bmatrix}, \tag{3}$$

Bằng cách trích xuất ma trận con quay kích thước $3x3$ từ ma trận biến đổi, ta thu được ma trận quay của khâu i và vector vị trí của gốc hệ tọa độ khâu i: Tách ma trận cỡ $3x3$ của ma trận truyền, ta thu được các ma trận quay của khâu thứ i và vectơ xác định gốc tọa độ khâu thứ i :

$$A_i^0, r_{O_i}^0 \,, \tag{4}$$

Tất cả công thức ở dưới quy ước ký hiệu:$c_i = \cos q_i \,, s_i = \sin q_i$

Các ma trận truyền D_i sau khi nhân ta được như sau: Ma trận truyền tổng thể D_6 từ hệ gốc đến sàn công tác được xác định từ tích của sáu ma trận Denavit–Hartenberg H_i. Ma trận D_6 có dạng:

$$D_6 = \begin{bmatrix} A_6 & r_6^{(0)} \\ 0^T & 1 \end{bmatrix}, \tag{5}$$

trong đó A_6 mô tả định hướng sàn công tác, còn $r_6^{(0)} = [p_x, p_y, p_z]^T$ là tọa độ điểm tham chiếu trên sàn trong hệ gốc.

$$D_6 = \begin{bmatrix} -c_1 s_2 s_5 + c_5(c_1 c_2 c_4 + s_1 s_4) & c_1 c_2 s_4 - c_4 s_1 & c_1 c_5 s_2 + s_5(c_1 c_2 c_4 + s_1 s_4) & p_x + d_6\left[c_1 c_5 s_2 + s_5(c_1 c_2 c_4 + s_1 s_4)\right] \\ -c_5(c_1 s_4 - c_2 c_4 s_1) - s_1 s_2 s_5 & c_1 c_4 + c_2 s_1 s_4 & c_5 s_1 s_2 - s_5(c_1 s_4 - c_2 c_4 s_1) & p_y + d_6\left[c_5 s_1 s_2 - s_5(c_1 s_4 - c_2 c_4 s_1)\right] \\ -c_2 s_5 - c_4 c_5 s_2 & -s_2 s_4 & c_2 c_5 - c_4 s_2 s_5 & p_z + d_6\left[c_2 c_5 - c_4 s_2 s_5\right] \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Với $p_x = c_1(a_1 + a_2 c_2 + a_3 c_2 - d_3 s_2 - d_4 s_2)$

$p_y = s_1(a_1 + a_2 c_2 + a_3 c_2 - d_3 s_2 - d_4 s_2)$

$p_z = -a_2 s_2 - a_3 s_2 - d_3 c_2 - d_4 c_2 + d_1$

Từ D6 ta thu được

$$A_6 = \begin{bmatrix} -c_1 s_2 s_5 + c_5(c_1 c_2 c_4 + s_1 s_4) & c_1 c_2 s_4 - c_4 s_1 & c_1 c_5 s_2 + s_5(c_1 c_2 c_4 + s_1 s_4) \\ -c_5(c_1 s_4 - c_2 c_4 s_1) - s_1 s_2 s_5 & c_1 c_4 + c_2 s_1 s_4 & c_5 s_1 s_2 - s_5(c_1 s_4 - c_2 c_4 s_1) \\ -c_2 s_5 - c_4 c_5 s_2 & -s_2 s_4 & c_2 c_5 - c_4 s_2 s_5 \end{bmatrix}, \tag{6}$$

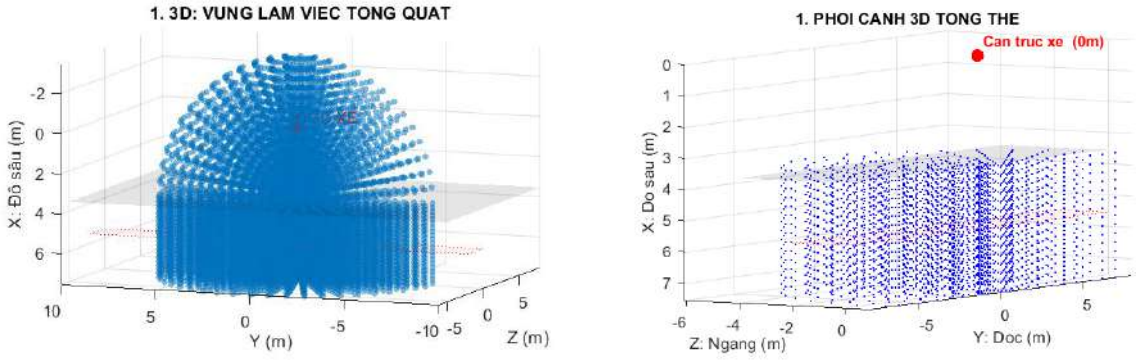
$$r_{O_6}^{(0)} = \begin{bmatrix} p_x + d_6\left[c_1 c_5 s_2 + s_5(c_1 c_2 c_4 + s_1 s_4)\right] \\ p_y + d_6\left[c_5 s_1 s_2 - s_5(c_1 s_4 - c_2 c_4 s_1)\right] \\ p_z + d_6\left[c_2 c_5 - c_4 s_2 s_5\right] \end{bmatrix}. \tag{7}$$

Vector $r_6^{(0)}$ cung cấp tọa độ không gian 3D chính xác của điểm tham chiếu trên sàn công tác dưới dạng hàm số của các biến khớp tâm q_1, q_2, d_3, q_4, q_5, d_6. Bằng cách áp dụng phương pháp quét cạn đưa các biến khớp chạy liên tục từ giá trị cực tiểu đến cực đại trong phạm vi giới hạn hành trình cho phép, ta có thể xác định được tập hợp tất cả các tọa độ không gian khả dĩ. Tập hợp tọa độ này chính là cơ sở toán học để mô phỏng và xây dựng đám mây điểm biểu diễn không gian làm việc tổng quát của thiết bị sẽ được trình bày trong phần tiếp theo.

3. Kết quả mô phỏng không gian làm việc và thảo luận

Để đảm bảo tính tổng quát và độ tin cậy của kết quả, quá trình khảo sát được thực hiện bằng mô phỏng số trên phần mềm MATLAB, dựa trên mô hình động học thuận đã xây dựng. Các thông số hình học của cầu và vị trí đỗ xe được thiết lập theo điều kiện làm việc thực tế, trong đó mặt cầu nằm ở độ sâu 3,55 m và đáy dầm ở độ sâu 5,55 m so với hệ tọa độ gốc.

Dựa trên mô hình động học, không gian làm việc của thiết bị được mô phỏng quét quỹ đạo có hệ thống (Hình 3, 4).



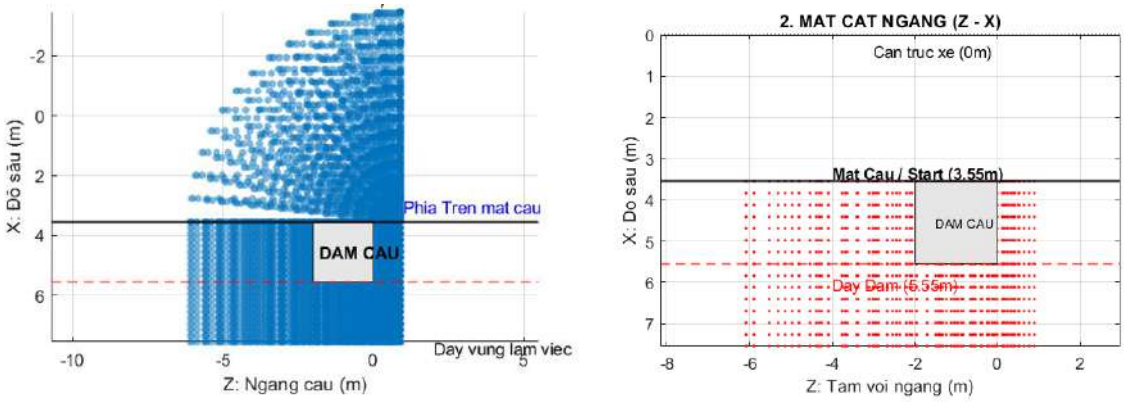
a) Vùng bao động học *b) Vùng thực tế*

Hình 4. Không gian 3D vùng làm việc có thể tiếp cận

Trên biểu đồ Mặt cắt ngang (Z-X), ta thấy đám mây điểm (vùng khả dĩ) bao trùm lên cả tiết diện dầm cầu (khối hình chữ nhật màu xám) và vượt qua đường giới hạn mặt cầu. Cần phân biệt rõ hai khái niệm:

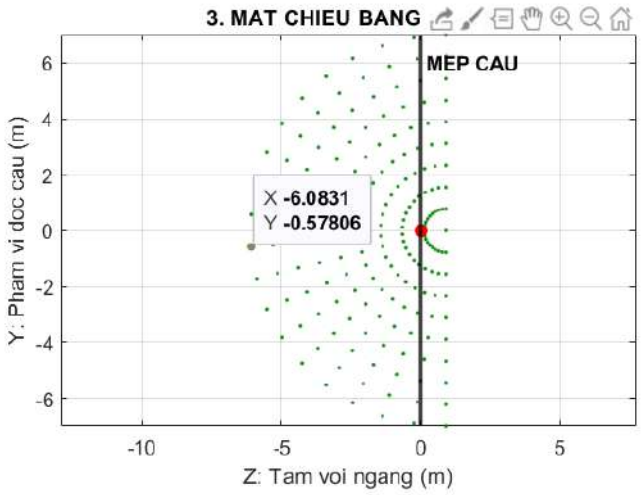
Vùng bao động học: Là toàn bộ vùng màu xanh trong hình. Nó thể hiện khả năng vươn tới tối đa của cơ cấu về mặt lý thuyết. Việc vùng màu xanh "lấn" vào dầm cầu chứng tỏ thiết bị có đủ tầm với để tiếp xúc hoặc kiểm tra sát bề mặt dầm.

Vùng làm việc thực tế: Trong thực tế vận hành, phần không gian giao thoa giữa đám mây điểm và kết cấu cầu (dầm cầu, mặt cầu) được coi là "Vùng va chạm". Do đó, không gian làm việc thực tế sẽ là phần hiệu của Vùng bao động học trừ đi Vùng va chạm và một khoảng cách an toàn,



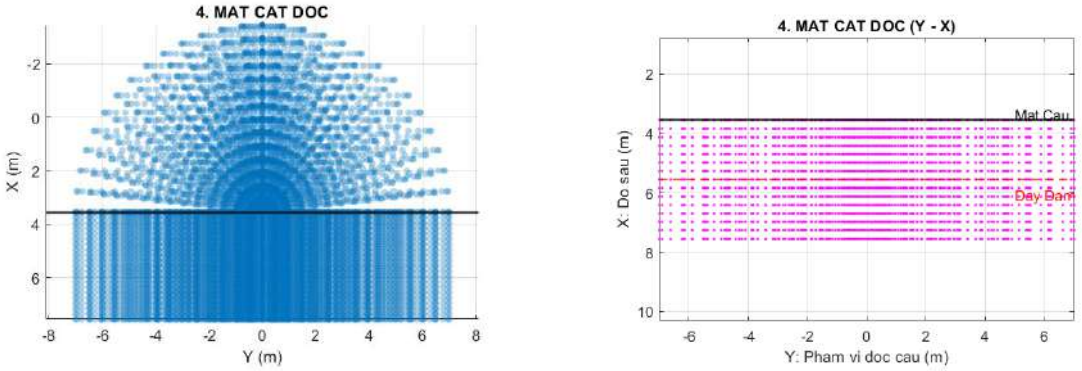
Hình 5. Mặt cắt ngang vùng làm việc

Từ kết quả mô phỏng trên hình chiếu bằng (mặt phẳng Y-Z), biên dạng vùng làm việc của thiết bị theo phương ngang được thể hiện rõ nét dưới dạng các đường cung tròn đồng tâm, phản ánh quỹ đạo quét của khâu tác động cuối. Trong đó, trục tham chiếu Z = 0 được đặt tương ứng với vị trí mép cầu. Phân tích đám mây điểm cho thấy tọa độ giới hạn ngoài cùng bên trái (tầm với ngang cực đại) đạt giá trị khoảng 6,08 m.



Hình 6. Hình chiếu bằng (Mặt phẳng Y-Z)

Nhận thấy điểm xa nhất mà người có thể với tới là 6,0831m, điều này càng khẳng định với tiêu chí ban đầu mà thiết bị chế tạo để khảo sát có chiều dài nửa cầu là 6m là hoàn toàn đáp ứng yêu cầu đặt ra.



Hình 7. Mặt cắt dọc vùng làm việc

Bên cạnh việc xác định tầm với cực đại, mô phỏng không gian làm việc còn làm rõ một rào cản kỹ thuật quan trọng: sự xuất hiện của các "vùng mù" trong quá trình rà quét. Sự không khớp về hình học giữa quỹ đạo quét dạng cung tròn của thiết bị và kết cấu gầm cầu (dầm chữ I, dầm hộp, bản mặt cầu) lại có biên dạng mặt cắt ngang hình chữ nhật đặc trưng. Sự chênh lệch này làm xuất hiện các "vùng mù" (điểm mù) cố hữu tại các góc xa của bản mặt cầu. Ngay cả khi thiết bị vươn tới bán kính cực đại 6,08 m tại một vị trí đỗ cố định, các góc khuất này vẫn không thể được tiếp cận. Khảo sát này đặt ra yêu cầu bắt buộc trong thực tế vận hành: xe cơ sở không thể đứng yên mà phải di chuyển tịnh tiến dọc theo cầu với các khoảng cách đỗ xe được tính toán hợp lý, sao cho các vùng không gian quét bù lấp hoàn toàn vào điểm mù của vị trí đỗ trước đó

4. Kết luận

Nghiên cứu đã thiết lập thành công mô hình động học 3D của hệ thống thiết bị chuyên dụng kiểm tra gầm cầu sử dụng phương pháp ma trận Denavit-Hartenberg (D-H). Các phương trình động học và ma trận Jacobian tổng quát đã được xây dựng đầy đủ, phản ánh chính xác quy luật truyền động của cơ cấu. Đặc biệt, kết quả mô phỏng số học đã xác định trực quan không gian làm việc của thiết bị, chỉ ra bán kính làm việc hiệu dụng đạt mức tối đa là 6,0 m. Việc làm rõ các giới

hạn về tầm với hình học này cung cấp cơ sở dữ liệu quan trọng để xác định phạm vi đỗ xe cơ sở an toàn, tránh va chạm và tối đa hóa diện tích bao phủ gầm cầu. Dựa trên những kết quả này, hướng phát triển tiếp theo là ứng dụng mô hình động học vào thiết kế hệ thống điều khiển quỹ đạo tự động, đồng thời đề xuất cải tiến cơ cấu cơ khí (như phát triển ray trượt ngang hoặc nới rộng sàn thao tác) nhằm khắc phục giới hạn tầm với trên các công trình cầu có khẩu độ lớn.

Tài liệu tham khảo

1. B. T. Thanh, T. Đ. Nhiệm, N. V. Minh, và N. H. Thuận, (2017). Khai thác và kiểm định cầu. Hà Nội, Việt Nam: Nhà xuất bản Giao thông Vận tải.
2. L. Q. Thủy, N. V. Thuyên, và cộng sự. (2020). "*Nghiên cứu, thiết kế, chế tạo thiết bị kiểm tra, thử nghiệm, sửa chữa, nâng cấp các kết cấu dưới gầm cầu (Báo cáo đề tài nghiên cứu khoa học, Mã số ĐT203028). Bộ Giao thông Vận tải, Việt Nam.*
3. K. Nakata, K. Umemoto, K. Kaneko, and R. Fujisawa, (2020). "*Development And Operation Of Wire Movement Type Bridge Inspection Robot System ARANEUS*", Kalpa Publ. Eng., vol. 3, pp. 168–174. https://doi.org/10.29007/zw9k
4. P. Kim and J. Youn, (2025). "*Drone Path Planning for Bridge Substructure Inspection Considering GNSS Signal Shadowing*," Drones, vol. 9, no. 2, Art. no. 124. https://doi.org/10.3390/drones9020124
5. N. V. Khang. (2017). *Động lực học hệ nhiều vật* (Xuất bản lần thứ 2). Hà Nội, Việt Nam: Nhà xuất bản Khoa học và Kỹ thuật.
6. M. W. Spong, S. Hutchinson, and M. Vidyasagar, (2020), Robot Modeling and Control, 2nd edition. Hoboken, NJ, USA: Wiley.

Working Position Optimization of Specialized Equipment for Under-Bridge Inspection and Repair

Abstract: *Mobile specialized equipment for the inspection, repair, and maintenance of the underside of bridges helps ensure safety and reliability in the operation of transportation infrastructure. Aiming at the practical optimization of working positions, this paper focuses on solving the foundational problem of establishing the mathematical model and analyzing the kinematic characteristics of the equipment. First, the kinematic model is established based on the Denavit–Hartenberg (D-H) homogeneous transformation matrix method. Subsequently, the generalized Jacobian matrix is constructed to represent the velocity relationship between the joints and the end-effector. Based on the kinematic equations, numerical simulations are conducted to accurately determine the 3D workspace profile of the device. The evaluation results indicate the effective working radius of the mechanism and clearly identify the geometric reach constraints. This serves as the core database for proposing safe positioning strategies for the base vehicle, establishing a premise for the efficient operation and exploitation of the equipment.*

Keywords: Kinematic analysis; Workspace; Under-bridge inspection equipment; D-H parameters; Jacobian matrix.

Đánh giá hiệu quả mô hình kriging trong phân tích đa tham số kết cấu thép mặt cắt dầm chủ cầu trục hai dầm chế tạo tại Việt Nam

Hoàng Văn Dựng[1*]
[1]*Học viện Kỹ thuật Quân sự*
[*]*hoangdung16091998@gmail.com*

Tóm tắt:

Bài báo đánh giá hiệu quả của mô hình Kriging trong phân tích đa tham số kết cấu thép mặt cắt dầm chủ cầu trục hai dầm chế tạo tại Việt Nam. Mô hình phần tử hữu hạn (PTHH) trên ANSYS APDL được dùng làm cơ sở tham chiếu, với các tham số hình học và vật liệu làm biến đầu vào. Tập dữ liệu huấn luyện từ mô phỏng PTHH qua phương pháp Latin Hypercube Sampling được sử dụng để xây dựng mô hình Kriging, nhằm dự đoán ứng suất và chuyển vị.

Kết quả cho thấy Kriging dự đoán chính xác so với PTHH, đồng thời giảm thời gian tính toán hàng chục lần. Nghiên cứu khẳng định tiềm năng của Kriging trong phân tích, đánh giá và tối ưu hóa thiết kế kết cấu thép cầu trục hai dầm tại Việt Nam.

Từ khóa: Mô hình Kriging; phân tích đa tham số; kết cấu thép; dầm chủ cầu trục hai dầm; chế tạo tại Việt Nam.

Abstract:

This paper evaluates the effectiveness of the Kriging model in multi-parameter analysis of the steel structural cross-section of a double-girder overhead crane main beam manufactured in Vietnam. A finite element model (FEM) developed in ANSYS APDL is used as the reference model, with geometric and material properties treated as input variables. The training dataset, generated from FEM simulations using the Latin Hypercube Sampling method, is employed to construct the Kriging model for predicting stress and displacement responses.

The results indicate that the Kriging model provides high prediction accuracy compared to the FEM while significantly reducing computational time by several orders of magnitude. The study confirms the potential of the Kriging approach for analysis, evaluation, and design optimization of double-girder overhead crane steel structures manufactured in Vietnam.

Keywords: Kriging model; multi-parameter analysis; steel structure; double-girder overhead crane main beam; manufacturing in Vietnam.

1. Đặt vấn đề

Cầu trục hai dầm là thiết bị nâng chuyển quan trọng, được sử dụng rộng rãi trong các nhà máy công nghiệp, xưởng chế tạo và kho bãi tại Việt Nam [1]. Trong đó, dầm chủ là bộ phận chịu lực chính, quyết định khả năng làm việc an toàn, độ bền và hiệu quả khai thác của toàn bộ hệ thống [2]. Ứng xử cơ học của dầm chủ phụ thuộc đồng thời vào nhiều tham số hình học mặt cắt, đặc trưng vật liệu, tải trọng và điều kiện liên kết. Do đó, phân tích đa tham số đóng vai trò then chốt trong thiết kế, đánh giá và tối ưu hóa kết cấu thép cầu trục hai dầm.

Phương pháp phần tử hữu hạn (PTHH) là công cụ phổ biến để phân tích ứng xử kết cấu dầm chủ [3]. Tuy nhiên, khi số lượng tham số đầu vào tăng lên, việc thực hiện hàng loạt mô phỏng PTHH để khảo sát ảnh hưởng của từng tham số hoặc tổ hợp tham số dẫn đến chi phí tính toán lớn và thời gian phân tích kéo dài.

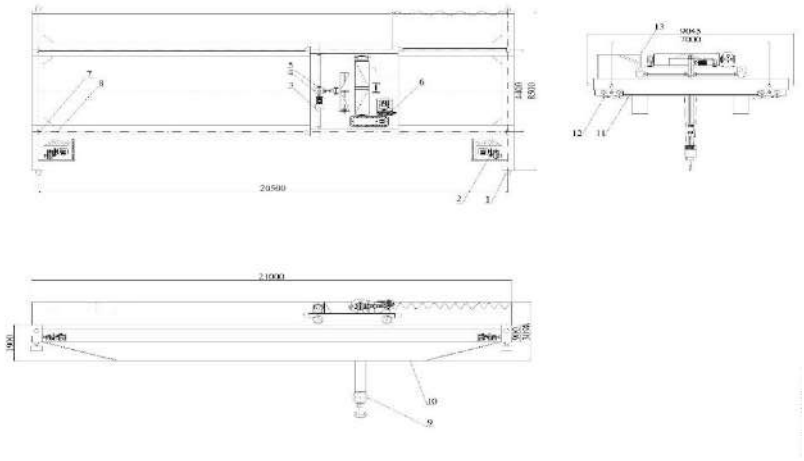
Mô hình Kriging nổi bật nhờ khả năng nội suy chính xác các quan hệ phi tuyến với số lượng mẫu huấn luyện tương đối nhỏ, đồng thời cung cấp thông tin về độ không chắc chắn của dự đoán [5]. Mô hình Kriging đã được ứng dụng hiệu quả trong nhiều lĩnh vực cơ học kết cấu, kỹ thuật cầu và công trình biển [6].

Xuất phát từ thực tiễn trên, bài báo này tập trung đánh giá hiệu quả của mô hình Kriging trong phân tích đa tham số kết cấu thép mặt cắt dầm chủ cầu trục hai dầm được chế tạo theo tiêu chuẩn và điều kiện áp dụng tại Việt Nam. Mô hình phần tử hữu hạn trên phần mềm ANSYS APDL được sử dụng làm cơ sở tham chiếu để huấn luyện và kiểm chứng độ chính xác của mô hình Kriging [4, 9].

2. Cơ sở lý thuyết và mô hình tính toán

2.1. Cấu tạo của cầu trục hai dầm nghiên cứu và giả thiết tính toán

Trong nghiên cứu này, đối tượng khảo sát là dầm chủ của cầu trục hai dầm dạng kết cấu hộp, được sử dụng phổ biến trong các nhà máy công nghiệp tại Việt Nam [1, 2]. Dầm chủ được liên kết cứng với dầm đầu, làm việc chủ yếu dưới tác dụng của tải trọng nâng thông qua xe con di chuyển dọc theo dầm chính. Hình học tổng thể của cầu trục và dạng mặt cắt dầm chủ được thể hiện trong Hình 1 và Hình 2, Hình 3.



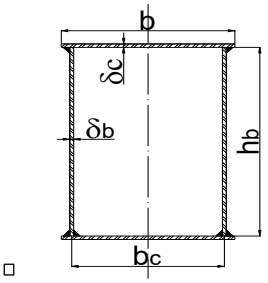
Hình 1. Tổng thể cầu trục hai dầm nghiên cứu

1. giảm chấn; 2. cơ cấu di chuyển; 3. Trục truyền động; 4. Cơ cấu di chuyển xe con; 5. Khớp nối; 6. Cơ cấu nâng; 7. Hạn chế hành trình xe con; 8. Ray di chuyển; 9. Móc treo; 10. Dầm chính; 11. Dầm đầu; 12. Bánh xe di chuyển cầu trục; 13. Bộ phận lấy điện cho xe con.

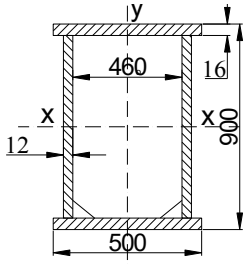
Bảng1. Thông số làm việc chính và tên gọi các bộ phận của cầu trục hai dầm nghiên cứu

STT	Thông số	Giá trị	STT	Thông số	Giá trị
1	Sức nâng	Q=980000N.	5	Vận tốc di chuyển xe con	V_{xc}=12m/ph.

2	Khẩu độ	L_k=20,5m.	6	Vận tốc nâng	V_n=1m/ph.
3	Chiều cao nâng	H_{max}=22m.	7	Cơ sở của xe con	A_{xc}=2,8m.
4	Vận tốc di chuyển cầu trục	V_{dc}=20m/ph.	8	Chế độ làm việc	nhẹ.



Hình 2. Tiết diện chính giữa dầm của dầm chính.



Hình 3. Tiết diện dầm đầu.

δ_c =16mm, δ_b =12mm.

Để tập trung phân tích ảnh hưởng của các tham số hình học và vật liệu đến ứng xử chịu lực của dầm chủ, mô hình tính toán được xây dựng dựa trên các giả thiết sau:

Vật liệu thép được giả thiết là đàn hồi tuyến tính, đồng nhất và đẳng hướng; các đặc trưng cơ học của vật liệu không thay đổi trong quá trình làm việc [2]. Kết cấu làm việc trong theo tải trọng IIa [1]. Ảnh hưởng của biến dạng cắt được xét đến nhằm phản ánh chính xác hơn ứng xử của dầm chủ có nhịp lớn, trong khi các hiệu ứng phi tuyến hình học và phi tuyến vật liệu được bỏ qua [3]. Điều kiện liên kết giữa dầm chủ và dầm đầu được giả thiết là liên kết cứng, phù hợp với cấu tạo thực tế của cầu trục hai dầm chế tạo tại Việt Nam [7].

2.2. Mô hình phần tử hữu hạn trong ANSYS APDL

Trong nghiên cứu này, phương pháp phần tử hữu hạn (PTHH) được sử dụng để phân tích ứng xử chịu lực của dầm chủ cầu trục hai dầm và làm cơ sở tham chiếu cho việc xây dựng mô hình xấp xỉ Kriging [3; 4]. Mô hình số được thiết lập trên phần mềm ANSYS APDL.

Dầm chủ cầu trục được mô hình hóa bằng phần tử dầm ba chiều, có khả năng mô phỏng đồng thời các trạng thái chịu uốn, cắt và kéo–nén [5].

Kết cấu được chia lưới với khoảng 200 phần tử dọc theo chiều dài dầm nhằm đảm bảo độ chính xác của kết quả tính toán.

Điều kiện biên của mô hình được thiết lập phù hợp với cấu tạo thực tế của cầu trục hai dầm, trong đó dầm chủ được liên kết cứng với dầm đầu, các bậc tự do tại vị trí gối đỡ bánh xe được khống chế theo phương thẳng đứng và phương ngang tương ứng. Tải trọng tác dụng bao gồm tải trọng nâng và các tải trọng phân bố đặc trưng cho bản thân kết cấu, được quy đổi và áp dụng tại các vị trí tương ứng trên mô hình [1, 2].

Kết quả phân tích PTHH này được sử dụng làm dữ liệu huấn luyện và kiểm chứng cho mô hình xấp xỉ Kriging trong các bước nghiên cứu tiếp theo.

Mô hình tính đã xây dựng trên ANSYS như hình 4 với thông số hình học và lực tác dụng theo bảng sau:

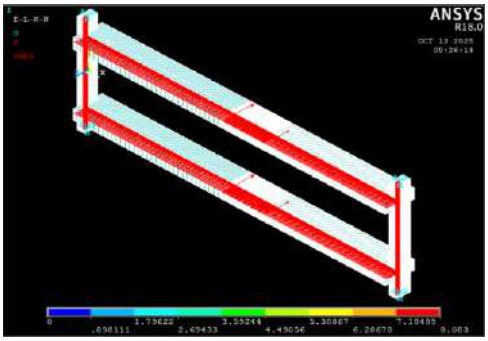
Bảng 2. Thông số đầu vào về lực tác dụng.

Thông số	Lực tập trung N1(N)	Lực tập trung N2(N)	Lực phân bố q_1(N/mm)	Lực phân bố q_2(N/mm)
Giá trị	508000	3040000	8,083	2,8

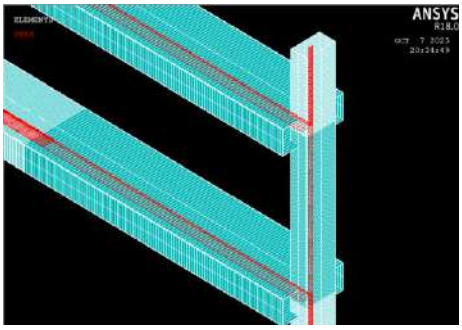
Bảng 3. Thông số đầu vào về kích thước dầm.

Thông số	Chiều dài dầm L(mm)	Chiều dài dầm đầu (mm)	Khoảng cách hai dầm chính (mm)
Giá trị	20500	8500	4400

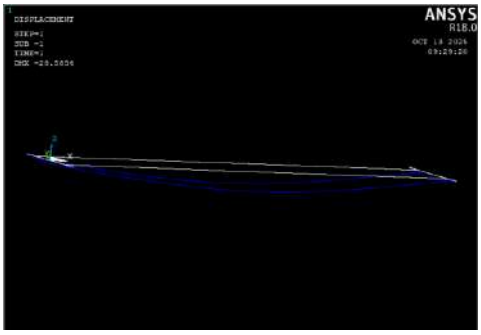
Kết quả số chương trình code trên ANSYS APDL của cầu trục:



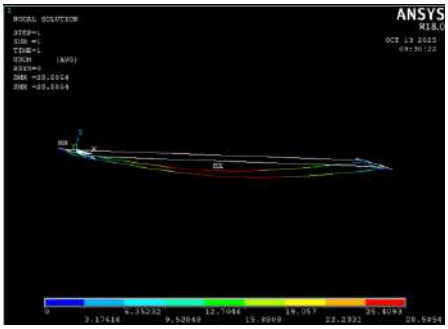
Mô hình PTHH dầm chính 3D



Mặt cắt dầm chính trong không gian



Biến dạng dầm chính



Biểu đồ chuyển vị dầm chính

Hình 4. Kết quả số của dầm chủ

Bảng 4. Kết quả phương pháp tính theo Ansys APDL của dầm chính.

Thông số	Chuyển vị (mm)	Lực cắt (N)	Mô men uốn (N.mm)	Ứng suất (N/mm^2)
ANSYS	28,58	518388	4170000	124,45

2.3. Mô hình xấp xỉ Kriging trong phân tích đa tham số

2.3.1. Tổng quan mô hình Kriging

Trong mô hình Kriging, đáp ứng của hệ kết cấu được biểu diễn như tổng của một hàm xu thế (trend function) và một thành phần ngẫu nhiên có tương quan không gian, cho phép mô tả hiệu quả các quan hệ phi tuyến giữa biến đầu vào và đầu ra [5].

Biểu thức tổng quát của mô hình Kriging được viết dưới dạng:

$$y(x) = f(x) + Z(x) \quad (1)$$

trong đó $f(x)$ là hàm xu thế, thường được chọn là hàm tuyến tính hoặc bậc hai; $Z(x)$ là một quá trình ngẫu nhiên có kỳ vọng bằng 0 và được đặc trưng bởi hàm tương quan [6].

Trong nghiên cứu này, hàm xu thế bậc hai kết hợp với hàm tương quan Gaussian được sử dụng nhằm mô tả chính xác sự thay đổi phi tuyến của ứng suất và chuyển vị dầm chủ theo các tham số hình học và vật liệu.

Không gian mẫu được xây dựng bằng phương pháp Latin Hypercube Sampling (LHS) Tổng số 20 mẫu được tạo ra và tương ứng với mỗi mẫu, bài toán PTHH được giải trên ANSYS APDL để thu được các đại lượng đáp ứng làm dữ liệu huấn luyện cho mô hình Kriging.

2.3.2. Quy trình xây dựng mô hình Kriging.

Quy trình xây dựng và ứng dụng mô hình Kriging được thực hiện theo hai giai đoạn:

Giai đoạn offline:

Thực hiện mô phỏng PTHH cho toàn bộ tập mẫu được tạo bằng LHS, thu thập dữ liệu đầu vào và đầu ra. Dữ liệu này được sử dụng để huấn luyện mô hình Kriging, trong đó các siêu tham số của hàm tương quan được xác định thông qua bài toán tối ưu cực đại hóa hàm hợp lý [6].

Giai đoạn online:

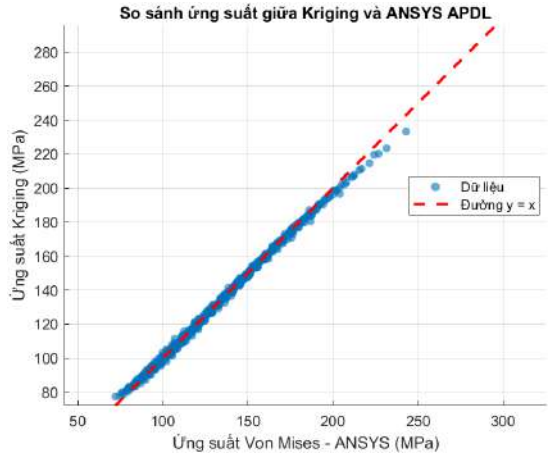
Sau khi mô hình được huấn luyện, Kriging được sử dụng để dự đoán nhanh các đại lượng đáp ứng của dầm chủ tại các tổ hợp tham số mới mà không cần thực hiện lại mô phỏng PTHH đầy đủ [5, 10].

2.3.3. Đánh giá độ chính xác mô hình Kriging

Độ chính xác của mô hình xấp xỉ Kriging được đánh giá thông qua tập dữ liệu kiểm tra độc lập, không trùng với tập dữ liệu huấn luyện.

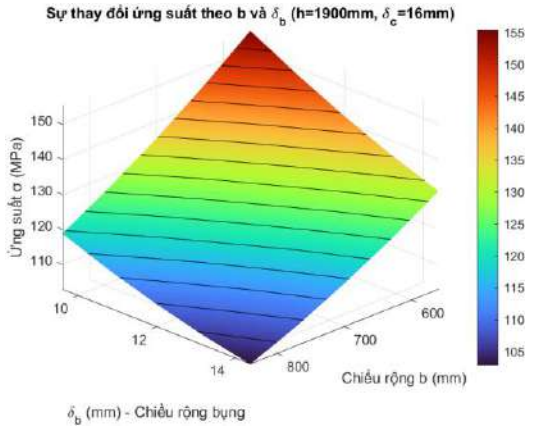
Bảng 5. Các chỉ số đánh giá độ chính xác của mô hình Kriging.

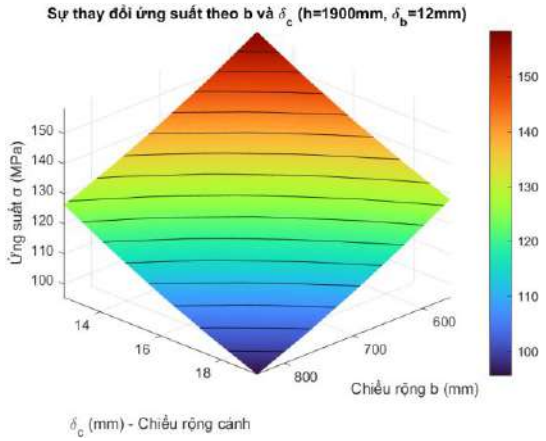
Chỉ số đánh giá	Giá trị
Sai số căn bậc hai trung bình (RMSE)	3,3559 %
Sai số tuyệt đối trung bình (MAE)	1,0482 %
Hệ số xác định R^2	0,9822
Sai số dự đoán lớn nhất (%)	4,24 %

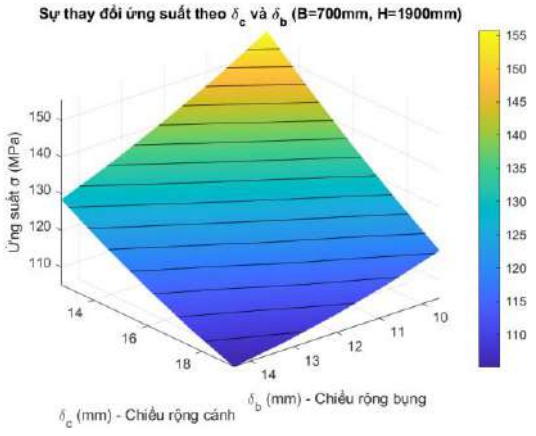


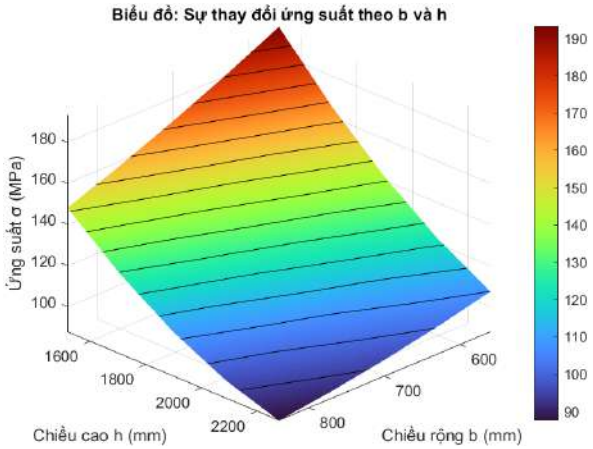
Hình 5. Ứng suất giữa mô hình Kriging và phần mềm ANSYS APDL

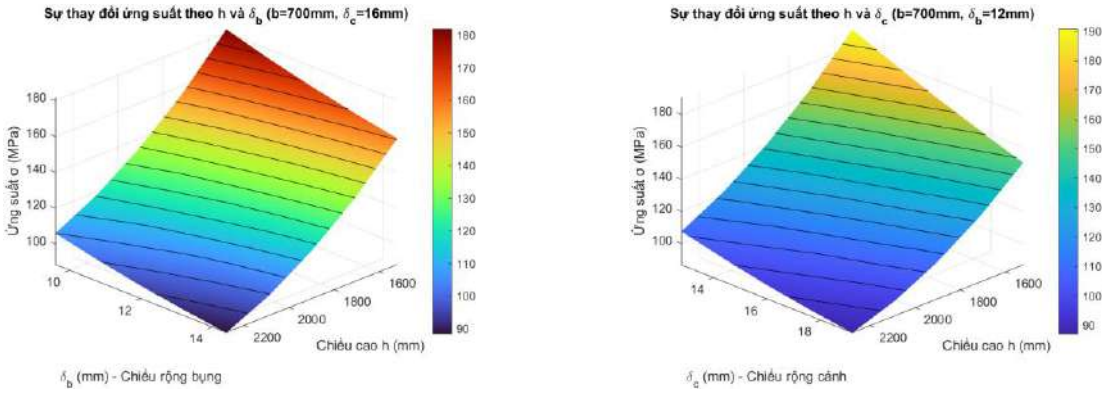
Kết quả từ bảng trên cho phép đánh giá mức độ phù hợp của mô hình Kriging trong việc xấp xỉ mối quan hệ giữa các tham số hình học và ứng xử chịu lực của dầm chủ cầu trục hai dầm.



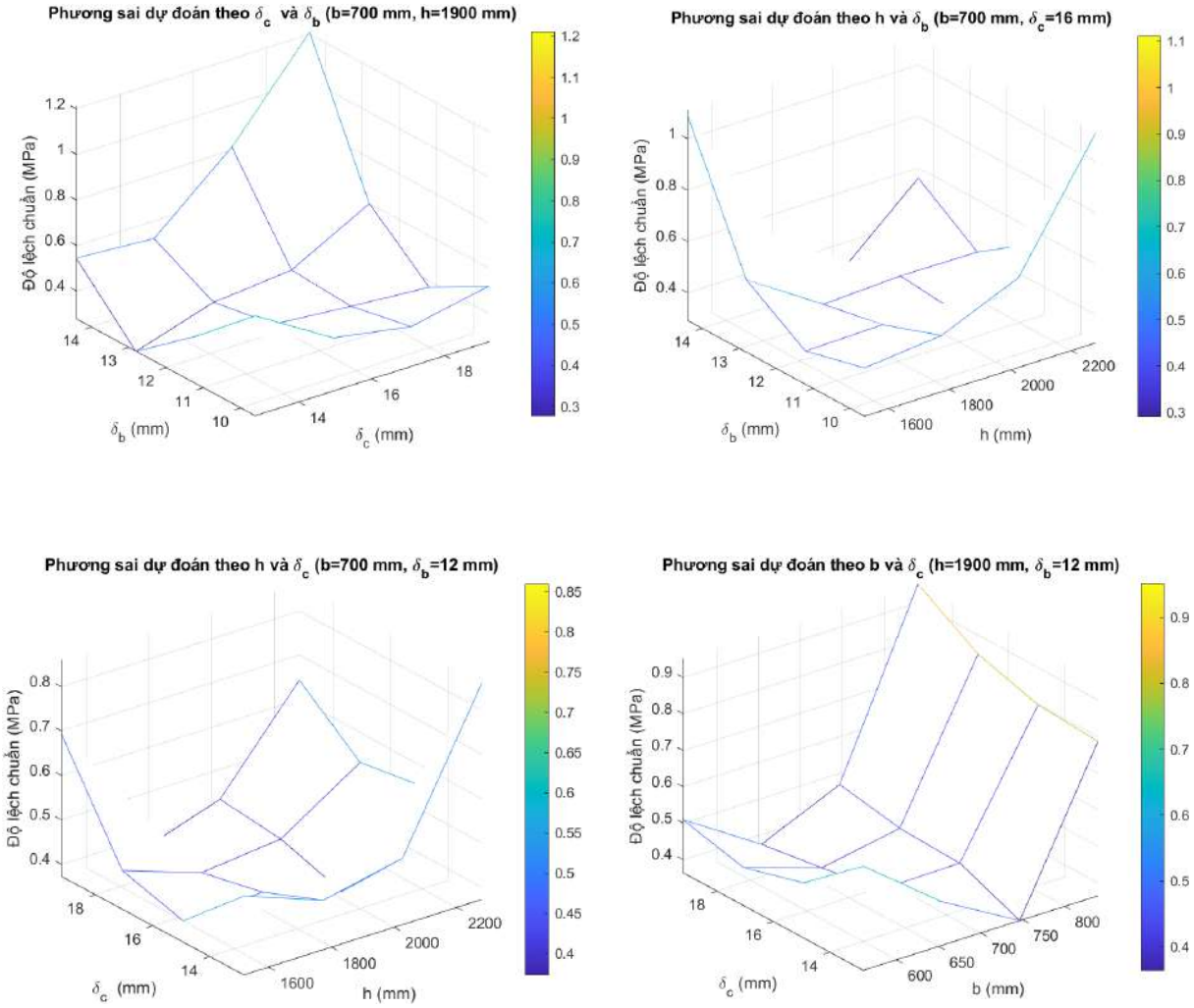


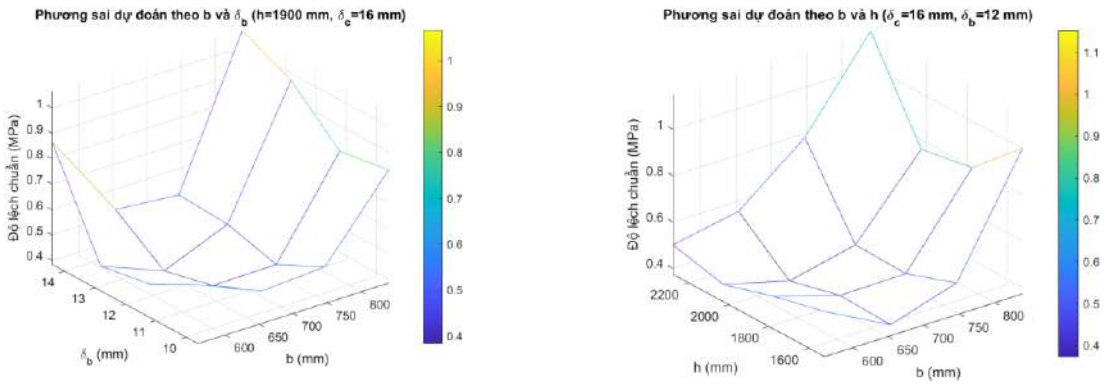






Hình 6. Mối quan hệ giữa giữa các kích thước mặt cắt h_b , b , δ_c , δ_b đến ứng suất σ mặt cắt dầm chính bằng mô hình Kriging.





Hình 7. Phương sai dự đoán mô hình xấp xỉ kriging theo kích thước mặt cắt h_b, b, δ_c, δ_b đến ứng suất σ mặt cắt dầm chính.

Mô hình Kriging nổi bật nhờ khả năng nội suy chính xác, đặc biệt trong phân tích đa tham số với biến hình học và vật liệu.

3. So sánh chi phí tính toán

3.1. So sánh chi phí tính toán giữa PTHH và mô hình Kriging

Để đánh giá hiệu quả tính toán của mô hình xấp xỉ Kriging, chi phí tính toán được so sánh với phương pháp phần tử hữu hạn (PTHH) đầy đủ thực hiện trên phần mềm ANSYS APDL.

Bảng 6. So sánh chi phí tính toán giữa PTHH và mô hình Kriging

Tiêu chí so sánh	PTHH (ANSYS APDL)	Mô hình Kriging
Thời gian xây dựng/huấn luyện ban đầu	600 -900 giây (thiết lập mô hình ban đầu)	1500-2200 giây (20 mẫu ban đầu từ PTHH)
Thời gian đánh giá 1 tổ hợp tham số	45-65 giây	0,03-0,08 giây
Thời gian cho 625 tổ hợp	12,5 - 15 giờ	2,5 - 3 phút

Các phép tính toán được thực hiện trên máy tính cá nhân với cấu hình phần cứng đảm bảo tính tái lập: Bộ xử lý Intel Core i7-10700, bộ nhớ RAM 16 GB DDR4-3200 MHz, và hệ điều hành Windows 10. Phiên bản phần mềm ANSYS APDL 18.0. Thời gian được đo bằng công cụ tích hợp của ANSYS và MATLAB (cho Kriging).

4. Kết luận

Bài báo đã trình bày việc xây dựng và đánh giá hiệu quả của mô hình xấp xỉ Kriging trong phân tích đa tham số kết cấu thép mặt cắt dầm chủ cầu trục hai dầm. Mô hình phần tử hữu hạn được thiết lập trên ANSYS APDL đóng vai trò là mô hình tham chiếu để huấn luyện và kiểm chứng độ chính xác của mô hình Kriging.

Kết quả nghiên cứu cho thấy mô hình Kriging có khả năng dự đoán chính xác các đại lượng đáp ứng của dầm chủ như ứng suất và chuyển vị, với sai số nhỏ so với kết quả PTHH, đồng thời giảm đáng kể thời gian tính toán. Mô hình cho phép đánh giá nhanh ảnh hưởng của các tham số hình học mặt cắt đến ứng xử chịu lực của kết cấu, qua đó chứng minh tiềm năng ứng dụng trong phân tích, thiết kế và tối ưu hóa kết cấu thép cầu trục hai dầm trong điều kiện sản xuất tại Việt Nam.

Tài liệu tham khảo

1. Nguyễn Văn Hợp, Phạm Thị Nghĩa, Lê Thiện Thành (2000). *Máy trục vận chuyển.* NXB Giao Thông Vận Tải, Hà Nội.
2. Nguyễn Văn Hợp, Phạm Thị Nghĩa (1996). *Kết cấu thép máy xây dựng – xếp dỡ*. NXB Giao Thông Vận Tải, Hà Nội.
3. Chu Quốc Thắng (1997), *Phương pháp phần tử hữu hạn*. NXB Khoa học kỹ thuật, Hà Nội.
4. ANSYS Mechanical APDL Structural Analysis Guide
5. Sobhkhiz, Ramtin, and Alireza Mardookhpour (2021). Numerical simulation of the effect of uplift on seismic behavior of uncontrolled surface steel tanks with Abaqus software. Journal of Numerical Methods in Civil Engineering 5.4 13-21.
6. Doan T.V., F.Massa, T.Tison, H.Naceur (2017), Réduction de modèle pour la prédiction de charges critiques de flambement sous variabilié, CFM 2017, Lille, Aout, France.
7. Nguyễn Văn Vịnh, Đoàn Văn Tú, (2014), Nghiên cứu ứng dụng phần mềm Ansys trong tính toán kết cấu thép cần trục tháp theo mẫu JTL65C4. Tạp chí Bộ GTVT 08 19-23.
8. Bùi Thanh Danh, Đoàn Văn Tú, số (2020), Ứng dụng phần mềm ANSYS APDL trong tính toán thiết kế cổng trục long môn phục vụ lao lắp dầm bê tông khẩu độ nhịp nhỏ trong thi công cầu dân sinh, Tạp chí Cơ khí Việt Nam 08 57-68.
9. ANSYS, Inc. (2025). Structural analysis guide: Release 2025 R1. ANSYS, Inc.
10. Van Hai, T. (2020). Failure probability analysis of overhead crane bridge girders with uncertain design parameters. Journal of Science and Technology in Civil Engineering, 14(3), 1-12.

Nghiên cứu động lực học hệ thống thu hồi đốt phao khơi của bộ cầu PMP trong quá trình tời đốt

Khổng Văn Công [1]
[1]*Học viện kỹ thuật Quân sự*
Email: khongcong28294@gmail.com; Tel:0359170228

Tóm tắt

Bài báo trình bày nghiên cứu về động lực học của hệ thống thu hồi đốt phao khơi PMP trong giai đoạn tời đốt. Mô hình toán học được xây dựng dưới dạng mô hình phẳng, sử dụng phương pháp Lagrange loại II để thiết lập hệ phương trình vi phân chuyển động phi tuyến với ba tọa độ suy rộng gồm dịch chuyển tịnh tiến của đốt phao, góc lắc quanh trọng tâm và góc quay của tang cuốn cáp. Nghiên cứu đã tính đến các yếu tố ảnh hưởng thực tế như độ đàn hồi của dây cáp, ma sát trượt giữa đốt phao với mặt bến và đặc tính mô men phát động biến thiên của động cơ xe cơ sở Kraz255B. Kết quả nghiên cứu là cơ sở lý thuyết quan trọng để đánh giá tính ổn định của cơ hệ, giúp tối ưu hóa quá trình vận hành và thiết kế các thiết bị thu hồi phao chuyên dụng, đảm bảo an toàn kỹ thuật trong điều kiện làm việc thực tế.

Từ khóa: *Động lực học, cáp đàn hồi, cầu Quân sự, phao PMP.*

1. Đặt vấn đề

Bộ cầu phao PMP là một trong những trang bị vượt sông chủ lực của quân đội, đóng vai trò quan trọng trong việc đảm bảo tính cơ động và hậu cần chiến lược. Trong quá trình vận hành, giai đoạn thu hồi các đốt phao sau khi hoàn thành nhiệm vụ bắc cầu là một khâu then chốt, đòi hỏi sự chính xác và an toàn tuyệt đối. Thực tế cho thấy, quá trình tời đốt phao từ dưới nước lên xe cơ sở thường xảy ra các hiện tượng dao động mạnh, lắc ngang hoặc thay đổi đột ngột lực căng cáp do tác động tương hỗ giữa chuyển động tịnh tiến và chuyển động lắc của phao.

Hình 1. Quá trình thu hồi đốt phao PMP giai đoạn tời đốt

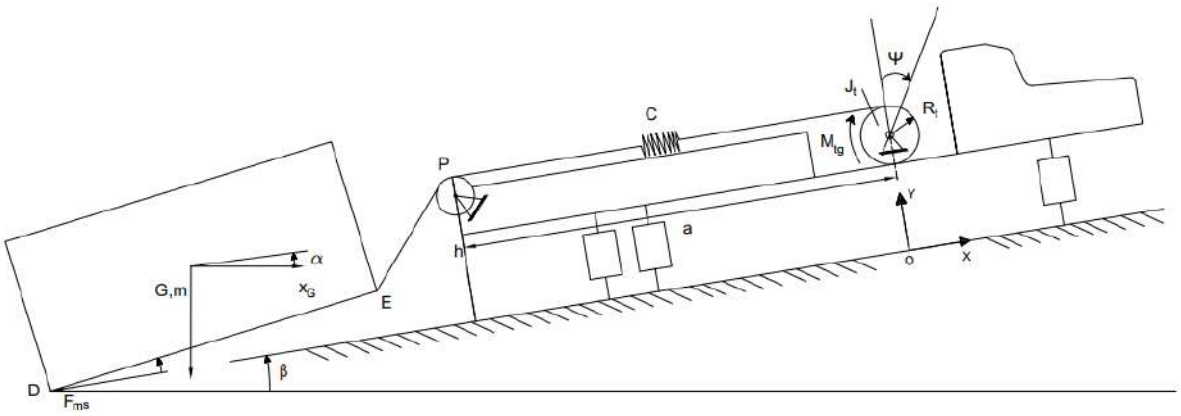
Xuất phát từ việc các mô hình tính toán hiện nay thường đơn giản hóa hệ thống, bỏ qua tính đàn hồi của cáp tời hoặc chưa xem xét đầy đủ ảnh hưởng của ma sát và đặc tính động cơ xe cơ sở đến sự ổn định của đốt phao. Điều này dẫn đến khó khăn trong việc dự báo các trạng thái nguy hiểm có thể gây đứt cáp hoặc lật đốt phao trong điều kiện địa hình bến bãi phức tạp.

Trong bài báo này, tác giả xây dựng mô hình động lực học phẳng chi tiết cho hệ thống thu hồi đốt phao khơi PMP. Bằng cách sử dụng phương pháp Lagrange loại II, nghiên cứu tập trung giải quyết việc thiết lập hệ phương trình vi phân chuyển động phi tuyến, có tính đến độ cứng của cáp đàn hồi, ma sát trượt tại bến và mô men phát động của xe Kraz255B. Kết quả nghiên cứu sẽ cung cấp cơ sở lý thuyết để tối ưu hóa quy trình tời đốt, góp phần nâng cao hiệu suất thu hồi và đảm bảo an toàn cho trang bị quân sự trong quá trình tác chiến.

2. Mô hình nghiên cứu động lực học hệ thống thu hồi đốt phao khơi pmp trong giai đoạn tời đốt

2.1. Mô tả mô hình nghiên cứu động lực học

Mô hình động lực học hệ thống thu hồi đốt phao khơi PMP trong giai đoạn tời đốt được xây dựng ở dạng mô hình phẳng (hình 2). Giả thiết trong quá trình tời đốt xe cơ sở được coi như vật rắn tuyệt đối và gắn cứng với nền. Dây cáp tời đốt được coi là đàn hồi với hệ số độ cứng C. Khối lượng của phao là m, G tương ứng là vị trí trọng tâm của m. Toàn bộ cơ hệ được đặt trọng hệ tục tọa độ cô định Oxy với O là hình chiếu của tâm quay tang xuống mặt nền bến. Tang cuốn cáp nhận mô men từ động cơ qua hệ thống truyền lực để dẫn động quá trình tời đốt. Bỏ qua tải lực đẩy nổi của nước, trọng gió, độ cản nhớt, ma sát trong ổ trục quay puly, tang. Xét mô men dẫn động từ động cơ quy dẫn về tang cuốn cáp và lực mà sát trượt của phao xuống mặt nền là hai thành phần lực không thế của cơ hệ.



Hình 2. Mô hình động lực học hệ thống thu hồi đốt phao khơi PMP giai đoạn tời đốt

Mô men quán tính khối lượng đốt phao và tang lần lượt là J_G và J_t. Các thông số hình học được ký hiệu như sau:

B - Chiều dài đốt phao

H - Chiều cao đốt phao

a - khoảng cách tâm tang quay đến tâm puly

h - chiều cao sàn xe.

2.2. Xây dựng hệ phương trình vi phân chuyển động

Các tọa độ suy rộng:

Hệ tọa độ suy rộng để khảo sát động lực học hệ thống thu hồi đốt phao khơi PMP trong giai đoạn tời đốt gồm ba thành phần:

x_D - dịch chuyển của đốt phao tại điểm tựa D theo phương bến (m)

α - góc nghiêng của đốt phao so với phương của bến (rad).

ψ - góc quay tang cuốn cáp (rad)

Xác định động năng của cơ hệ:

Động năng của cơ hệ bao gồm động năng của đốt phao (T_1) và động năng của các khối lượng quay quy dẫn về tang cuốn cáp (T_2) và được xác định theo công thức:

$$T = T_1 + T_2 \tag{1}$$

Động năng của cơ hệ xác định bởi công thức:

$$T = \frac{1}{2}m\left[\dot{x}_D^2 + \left(\frac{B^2+H^2}{4}\right)\dot{\alpha}^2 - \dot{x}_D\dot{\alpha}(B\sin\alpha + H\cos\alpha)\right] + \frac{1}{2}J_G\dot{\alpha}^2 + \frac{1}{2}J_t\dot{\psi}^2 \tag{2}$$

Xác định thế năng của cơ hệ:

Thế năng của hệ bao gồm thế năng trọng trường và thế năng đàn hồi. Trước tiên ta có độ biến dạng cáp ΔL trong quá trình tời phao gồm độ biến dạng tĩnh ΔL_{t0} và độ biến dạng động ΔL_d.

Lực căng cáp tĩnh F_{ct} xuất hiện trên dây cáp thông qua xét phương trình cần bằng mô men tại điểm D:

$$\sum M(D) = F_{ct0}.\mathrm{d}_{F0} - mg\left[\frac{B}{2}\cos(\alpha_0+\beta) - \frac{H}{2}\sin(\alpha_0+\beta)\right] = 0 \tag{3}$$

Trong đó: d_{F0} là khoảng cách từ D đến véc tơ lực căng cáp tĩnh $\vec{F}_{ct0}$ tại thời điểm ban đầu. Ta xác định được lực căng cáp tĩnh theo biểu thức:

Độ biến dạng tĩnh được xác định bởi công thức:

$$F_{ct0} = \frac{mg[\frac{B}{2}\cos(\alpha_0+\beta) - \frac{H}{2}\sin(\alpha_0+\beta)].\sqrt{(|x_{D0}+a| - \mathrm{B}\cos\alpha_0)^2 + (h - B\sin\alpha_0)^2}}{B.(h\cos\alpha_0 - |x_{D0}+a|\sin\alpha_0)} \tag{4}$$

$$\Delta L_{t0} = \frac{F_{ct0}}{C} \tag{5}$$

Với bội suất của hệ palang dẫn động bằng 1, độ biến dạng động ΔL_d của cáp trong quá trình tời phao được xác định như sau:

$$\Delta L_d = S - S_0 + R_t(\psi - \psi_0) \tag{6}$$

Trong biểu thức (6), S là chiều dài đoạn dây cáp từ E đến P tại thời điểm bất kỳ, S_0 là chiều dài ban đầu của đoạn dây cáp từ E đến P. R_t là bán kính của tang cuốn cáp, ψ, ψ_0 là góc quay tang cuốn cáp tại thời bất kỳ và thời điểm ban đầu. Các đại lượng này được xác định theo biểu thức sau:

$$S = \sqrt{(x_D + a + B\cos\alpha)^2 + (B\sin\alpha - h)^2} \quad (7)$$

$$S_0 = \sqrt{(x_{Do} + a + B\cos\alpha_0)^2 + (B\sin\alpha_0 - h)^2} \quad (8)$$

Ta có biểu thức xác định thế năng tổng cộng của cơ hệ là:

$$\Pi = mgh_G + \frac{1}{2}C(\Delta L_{t0} + S - S_0 + R_t(\psi - \psi_0)^2 \quad (9)$$

Trong đó: h_G là chiều cao trọng tâm G so với nền bến tại thời điểm bất kì

$$h_G = x_D \sin\beta + \frac{\beta}{2}\sin(\alpha + \beta) + \frac{H}{2}\cos(\alpha + \beta) \quad (10)$$

- Hàm Lagrange của cơ hệ là:

$$L = T - \Pi \quad (11)$$

Hàm hao tán $\Phi = 0$ ***(do bỏ qua lực cản nhớt)***

Lực suy rộng và công ảo:

Lực suy rộng với các tọa độ suy rộng $q = [x_D, \alpha, \psi]^T$ được xác định thông qua biểu thức công ảo của các lực ma sat và mô men dẫn động tác động lên cơ hệ. Để tính đến các lực không thế tác động lên hệ thống, ta sử dụng biểu thức công di chuyển khả dĩ:

$$\delta A = M_{tg}.\delta\psi - F_{ms}.\delta x_D \quad (12)$$

Trong đó: $F_{ms} = \mu.N_D$ là lực ma sát trượt của đốt phao trên mặt bến.

N_D là phản lực liên kết tại điểm D, được xác định bằng phương trình hình chiếu lực lên phương vuông góc với bến.

Trong động lực học, phản lực N_D không chỉ đơn thuần là thành phần trọng lực mà còn chịu ảnh hưởng bởi gia tốc của phao và lực kéo từ cáp.

Để tìm N_D ta xét phương trình chuyển động của phao theo phương vuông góc với mặt bến (phương Oy):

$$\sum F_y = N_D - mg\cos\beta + F_c.\sin\theta = m.\ddot{y}_G \quad (13)$$

Trong đó: θ là góc hợp bởi dây PE với phương Ox.

$\ddot{y}_G$ là tốc trọng tâm G theo Oy:

$$\ddot{y}_G = \left(\frac{B}{2}\cos\alpha - \frac{H}{2}\sin\alpha\right)\ddot{\alpha} - \left(\frac{B}{2}\sin\alpha + \frac{H}{2}\cos\alpha\right)\dot{\alpha}^2 \quad (14)$$

Ta rút ra được:

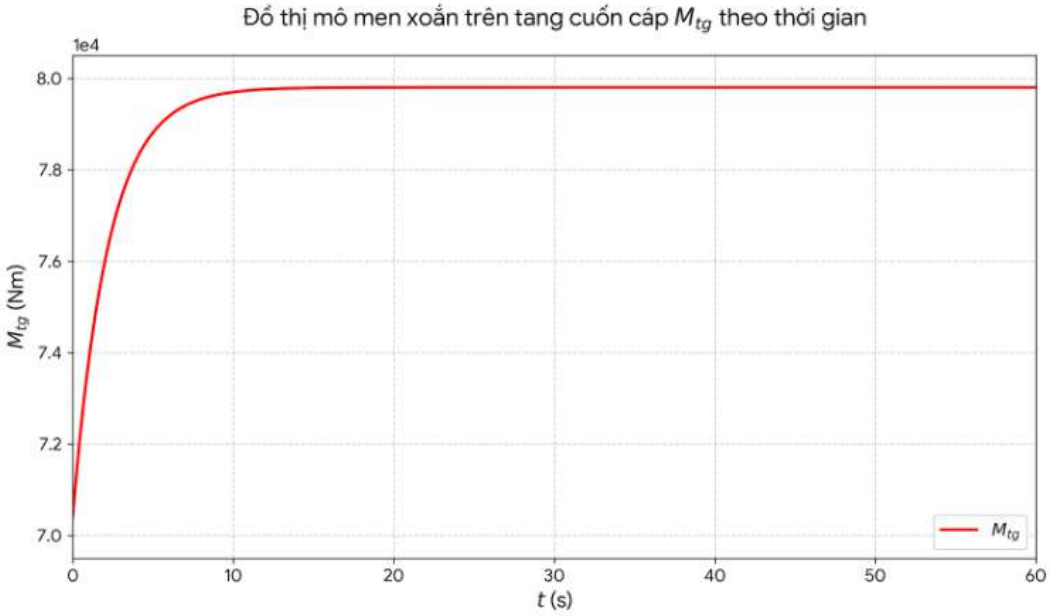
$$N_D = mg\cos\beta - F_c.\sin\theta + m.\left[\left(\frac{B}{2}\cos\alpha - \frac{H}{2}\sin\alpha\right)\ddot{\alpha} - \left(\frac{B}{2}\sin\alpha + \frac{H}{2}\cos\alpha\right)\dot{\alpha}^2\right] \quad (15)$$

M_{tg} là mô men quán tính quy dẫn từ động cơ về tang cuốn cáp. Xe cơ sở Kraz255B sử dụng động cơ đốt trong là IAMZ-238 đường đặc tính ngoài của động cơ và đặc tính làm việc của tang được trình bày trong tài liệu [6,7].

Để giải bài toán lý thuyết, từ đường đặc tính làm việc của tang cuốn cáp theo [6] và cơ chế làm việc thực tế, ta cần xây dựng đồ thị biểu diễn mô men M_{tg} quy dẫn từ động cơ về tang cuốn cáp theo thời gian. Trong quá trình vận hành, người điều khiển đạp chân ga để tăng tốc độ động cơ từ trạng thái không tải với tốc độ 600 vòng/phút đến tốc độ động cơ là 1400 vòng/phút ứng với mô men xoắn cực đại, sau đó sẽ giữ nguyên chân ga để động cơ quay ổn định ở vòng tua này. Như vậy quá trình tăng tốc của động cơ phụ thuộc vào kỹ năng của người vận hành. Để giải quyết bài toán lý thuyết, giả sử quy luật tăng tốc của động cơ là đường bậc nhất tốc độ động cơ tăng từ 600 vòng/ phút lên 1400 vòng/phút trong thời gian 10s. Khi đó ta có quy luật của M_{tg} được xác định theo biểu thức sau:

$$M_{tg} = \begin{cases} \{1,4923.10^8 .\left(0,7.\frac{1}{2100} + 1,3\frac{600+80t}{2100^2} - \frac{(600+80t)^2}{2100^3}\right)(0 \le t \le 10) \\ 79747(t \ge 10) \end{cases} \quad (16)$$

Đồ thị biến thiên mô men M_{tg} được thể hiện trên hình 3.



Hình 2: Đồ thị M_{tg} phụ thuộc thời gian

Véc tơ lực suy rộng tương ứng với các tọa độ được xác định theo biểu thức:

$$\left[Q_{X_D}, Q_\alpha, Q_\psi\right]^T = \left[-\mu N_D \,\mathrm{sgn}(\dot{x}_D), 0, M_{tg}\right]^T \quad (17)$$

Áp dụng phương trình Lagrange loại II để viết phương trình vi phân mô tả chuyển động của cơ hệ ở dạng:

$$\frac{d}{d_t}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q_i} = Q_i \quad (18)$$

Thế các biểu thức động năng, thế năng ta nhận được hệ phương trình vi phân mô tả động

lực học trong quá trình tời đốt phao PMP như sau:

$$\begin{cases} m\ddot{x}_D - \dfrac{m}{2}(B\sin\alpha + H\cos\alpha)\ddot{\alpha} - \dfrac{m}{2}(B\cos\alpha - H\sin\alpha)\dot{\alpha}^2 + \dfrac{\partial\Pi}{\partial x_D} = -\mu N_D \operatorname{sgn}(\dot{x}_D) \\ \left[J_G + m\dfrac{B^2+H^2}{4}\right]\ddot{\alpha} - \dfrac{m}{2}\left(B\sin\alpha + H\cos\alpha\right)\ddot{x}_D + \dfrac{\partial\Pi}{\partial\alpha} = 0 \\ J_t\ddot{\psi} + F_c.R_t = M_{tg} \end{cases} \quad (19)$$

Trong đó: $\dfrac{\partial\Pi}{\partial x_D} = mg\sin\beta + \mathrm{F}_c.\dfrac{x_D + a + B\cos\alpha}{S}$ (20)

$$\frac{\partial\Pi}{\partial\alpha} = mg\left[\frac{B}{2}\cos(\alpha+\beta) - \frac{H}{2}\sin(\alpha+\beta)\right] + F_c.\frac{\partial S}{\partial\alpha}) \quad (21)$$

$$\frac{\partial S}{\partial\alpha} = \frac{-B}{S}\left[(x_D + \mathrm{a})\sin\alpha + h\cos\alpha\right] \quad (22)$$

$$F_c = C.\left[\Delta L_{t0} + S - S_0 + R_t(\psi - \psi_0)\right]) \quad (23)$$

Thông số đầu vào để giải hệ phương trình vi phân được cho trong bảng 1.

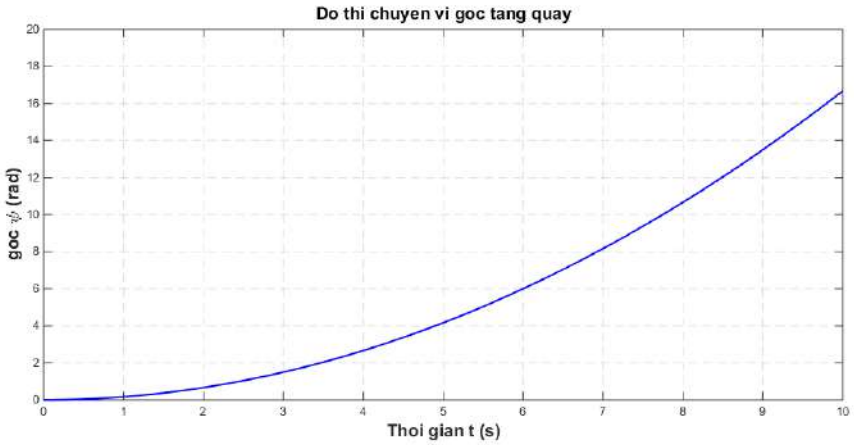
Bảng 1. Bảng số liệu đầu vào dùng cho tính toán

STT	Tham số	Giá trị	Đơn vị	Ghi chú
1	J_t	8085	$kg.m^2$	
2	R_t	0,3	m	
3	a	3,8	m	
4	h	1,5	m	
5	B	6,91	m	
6	H	2,277	m	
7	m	6790	kg	
8	C	2390	kN/m	
9	X_{D0}	-15	m	
10	α_0	0	rad	
11	ψ_0	0	rad	
12	g	9,8	m/s^2	
13	β	0,2	rad	
14	Mtg	80000	Nm	
15	μ	0,3		

3. Kết quả và thảo luận

Sau khi tính toán trên phần mềm Matlab để giải hệ phương trình vi phân (19) nhận được kết quả thể hiện ở các hình 3-8:

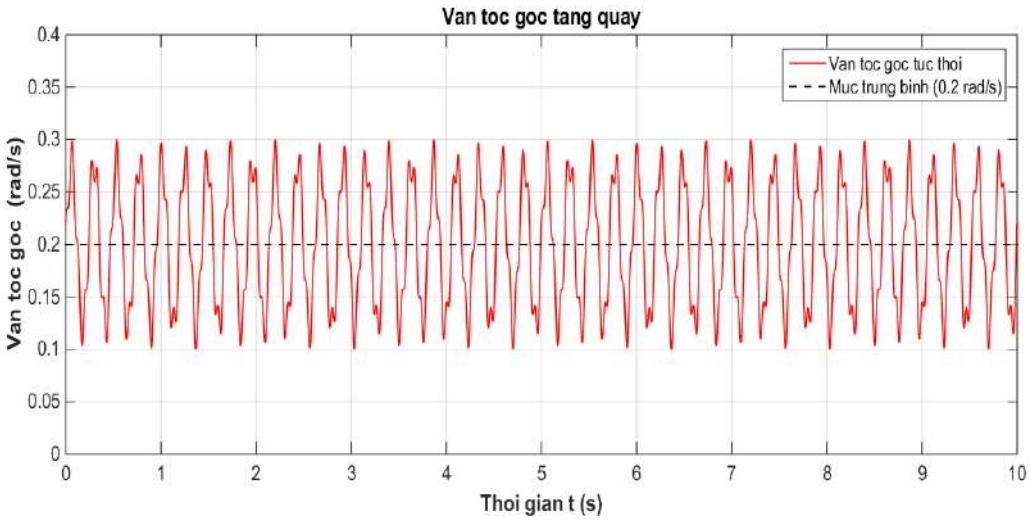
Qúa trình tời đốt phao diễn ra trong một khoảng thời gian ngắn. Ở đây khi tính toán ta giả định các quá trình đều diễn ra trong khoảng thời gian là 10 giây. Khi đó ta chạy chương trình và có các kết quả tính toán như sau:



Hình 3. Đồ thị chuyển vị góc tang quay

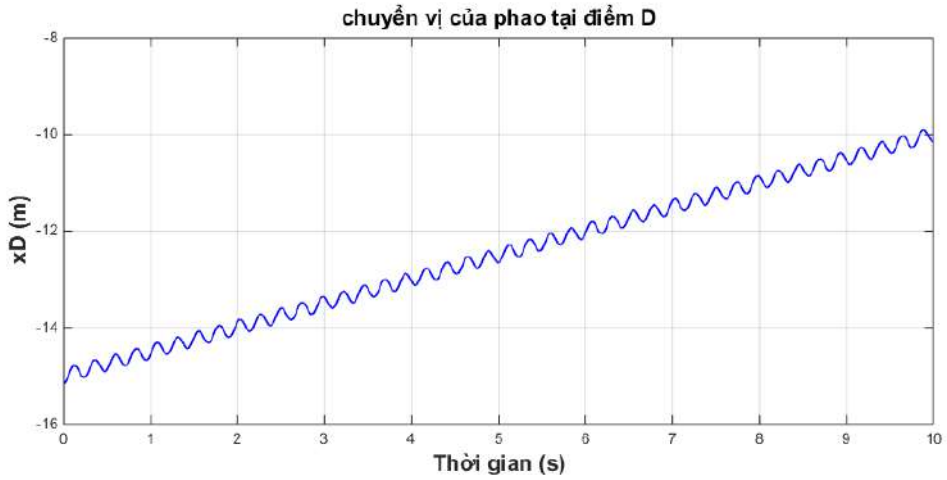
Trên hình 3 biểu diễn sự thay đổi của góc của tang cuốn cáp tời đốt phao theo thời gian trong quá trình hạ tời đốt phao. Từ đồ thị hình 3 nhận thấy rằng, góc của tang cuốn cáp tăng dần theo thời gian và gần sát với đường đồ thị tuyến tính. Là một đường ổn định, hầu như không dao động. Lý do vì trong quá trình khảo sát ta bỏ qua một số yếu tố ảnh hưởng như lực ma sát tại các ổ quay và do sai số tính toán của mô hình tính toán... Tại thời điểm ban đầu khi bắt đầu có nguồn động lực dẫn động, góc cuốn tang vẫn theo chiều hướng tăng dần nhưng với tốc độ chậm do lúc này hệ thống vẫn còn lực quán tính. Kết thúc quá trình gấp đốt tại thời điểm 10 (s) ta thấy góc lớn nhất của tang cuốn cáp đạt được là khoảng 16,5 (rad) tương ứng với tang cuốn cáp sẽ quay được gần 2,6 vòng. Sở dĩ tang cuốn được đoạn cáp nhỏ như vậy là do quá trình tời đốt này chỉ diễn ra trong thời gian ngắn và quãng đường di chuyển do đốt thực hiện trên mặt đất cũng nhỏ.

Nói cách khác là thay đổi chuyển vị góc là không nhiều. Dao động ổn định gần sát với đường tuyến tính. Quy luật thay đổi của góc cuốn tang trong quá trình tời đốt phao so với thực tế là hoàn toàn phù hợp.



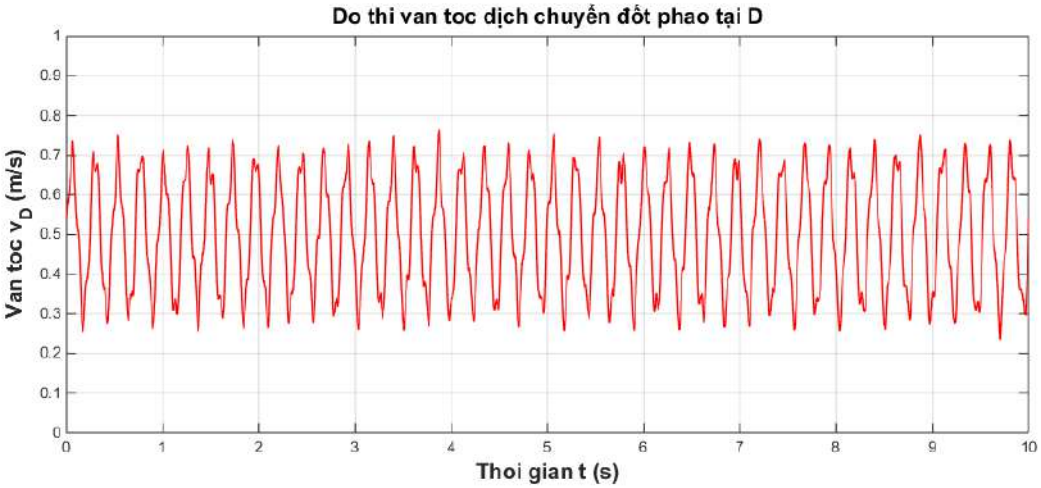
Hình 4. Đồ thị vận tốc góc quay tang theo thời gian

Mặc dù vận tốc góc của tang tời chịu sự dao động do độ cứng đàn hồi của hệ thống, nhưng biên độ dao động luôn nằm trong tầm kiểm soát (không vượt quá mức 0.3 rad/s và không bị tụt về 0). Điều này cho thấy hệ dẫn động của tời được chọn có đủ mô-men quán tính và công suất dự trữ để hấp thụ các xung lực giật từ cáp kéo, đảm bảo hệ thống phao không bị dừng đột ngột hoặc mất sức căng cáp trong suốt quá trình thu hồi.



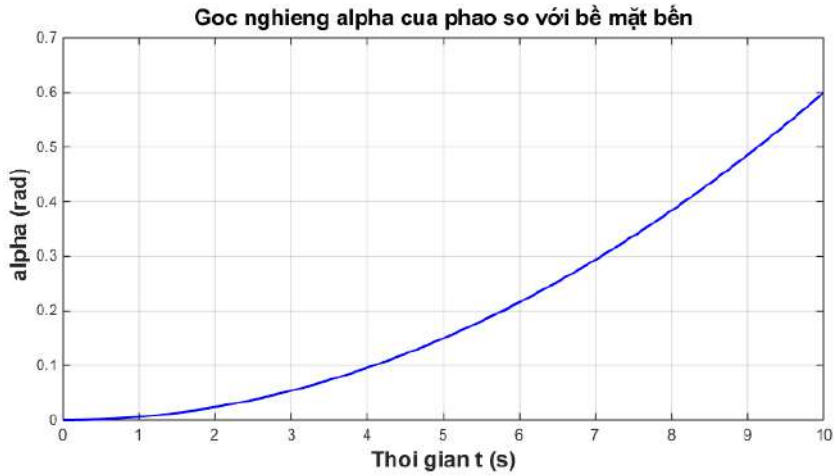
Hình 5. Đồ thị chuyển vị của phao tại điểm D theo phương của bến

Từ quỹ đạo chuyển vị x_D, có thể khẳng định hệ thống tời kéo hoạt động ổn định và đáp ứng tốt yêu cầu thu hồi phao. Dù chịu tác động của nhiễu động lực học và sự thay đổi lực căng cáp liên tục, quỹ đạo của phao vẫn bám sát mục tiêu tịnh tiến mà không xảy ra hiện tượng trượt ngược hay dừng giật cục. Điều này chứng minh rằng mô-men kéo do hệ dẫn động cung cấp hoàn toàn thắng được tổng các lực cản tĩnh và động của cơ hệ.

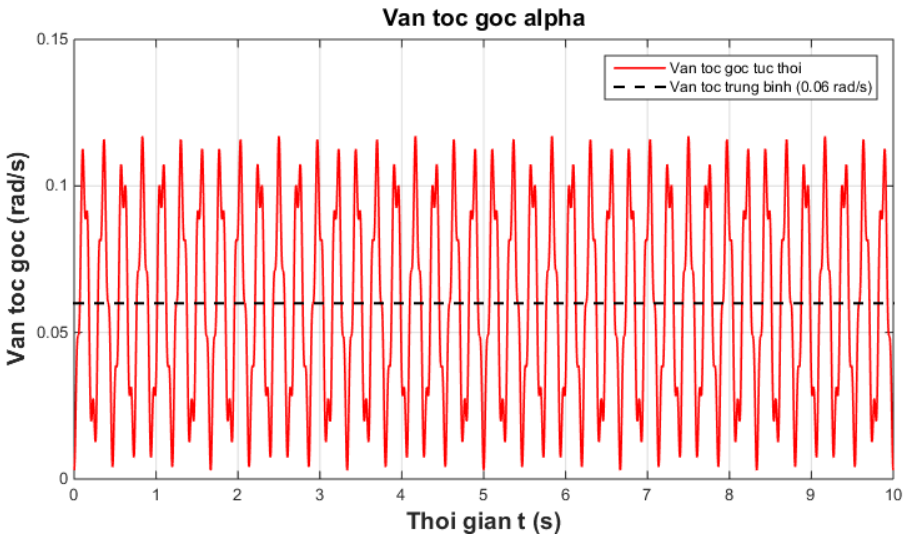


Hình 6. Đồ thị vận tốc dịch chuyển của phao theo phương của bến

Dù chịu sự dao động mạnh do tính chất đàn hồi của cáp, vận tốc tức thời của phao luôn mang giá trị dương. Điều này khẳng định trong suốt chu kỳ thu hồi, phao liên tục tiến về phía trước mà không hề bị khựng lại hay có hiện tượng trượt lùi trên mặt dốc. Kết quả này chứng minh lực căng cáp luôn được duy trì, động cơ tời đáp ứng tốt tải trọng động và quy trình thu hồi phao đảm bảo tính an toàn kỹ thuật cao.



Hình 7. Đồ thị chuyển vị góc nghiêng phao so với mặt bến



Hình 8. Đồ thị vận tốc góc nghiêng phao so với mặt bến

Quan sát đồ thị chuyển vị góc α ta thấy trạng thái của phao thay đổi rất êm ái từ 0 đến 0.6 rad (tương đương 34.4 độ) sau 10 giây thu hồi. Trong khi đó, đồ thị vận tốc góc tức thời của phao lại thể hiện sự dao động mạnh quanh mức trung bình 0.06 rad/s. Điểm đặc biệt của mô hình này là vận tốc góc biến thiên dạng hình sin, điều này phản ánh rất đúng trong thực tế. Khối lượng và mô-men quán tính lớn của phao đóng vai trò như một bộ giảm chấn tự nhiên, triệt tiêu gần như hoàn toàn các dao động vĩ mô trên đường cong chuyển vị góc. Tuy nhiên, sự co dãn của dây cáp kéo vẫn truyền những xung lực biến lên phao.

4. Kết luận

Bài báo đã trình bày mô hình động lực học và thiết lập hệ phương trình vi phân mô tả quá trình thu hồi phao của hệ thống tời kéo trên mặt phẳng nghiêng, trong đó có xét đến đặc tính đàn hồi của dây cáp. Trên cơ sở mô hình toán học đã xây dựng, tác giả đã tiến hành khảo sát thành công các quy luật động học, động lực học của cơ hệ, bao gồm chuyển động tịnh tiến, dao động góc lắc của phao và góc quay tang dưới tác dụng của lực căng cáp biến thiên. Kết quả mô phỏng cho thấy sợi cáp đóng vai trò như một khâu đàn hồi gây ra các xung dao động vận tốc dạng hình sin tại điểm tựa phao; tuy nhiên, nhờ mô-men quán tính lớn của cơ cấu tời và

trọng lượng phao, quỹ đạo chuyển vị và trạng thái quay của phao vẫn được duy trì ổn định. Kết quả đạt được của bài báo có thể làm cơ sở khoa học tin cậy cho bài toán tính toán sức bền đứt cáp cũng như đề xuất phương án điều khiển tốc độ động cơ tời nhằm giảm thiểu xung lực giật, đảm bảo an toàn và nâng cao hiệu suất làm việc của thiết bị trong điều kiện vận hành thực tế.

Tài liệu tham khảo

1. Bùi Khắc Gầy, Hướng dẫn tính toán thiết kế máy nâng vận chuyển, Nhà xuất bản Học viện Kỹ thuật Quân sự, 1989.

2. Nguyễn Huy Hoàng, Nghiên cứu khả năng ổn định của phà PMP khi chất, dỡ, chuyên chở tải, 2006.

3. Lê Hồng Quang, Nghiên cứu xác định các thông số Động lực học cơ cấu gấp, mở nhịp cầu TMM- 3M, 2017.

4. Nguyễn Viết Tân, Giáo trình động lực học máy trục, Nhà xuất bản Quân đội Nhân dân, 2017.

5. Vu Cong Ham, Dao dong co hoc. Military Technical Academy, Hanoi, 2016.

6. Le Hong Quang, Research to determine the dynamic parameters of the folding and opening structure of the TMM-3M bridge. Master Thesis, Military Technical Academy, Hanoi, 2017.

7. Le Van Thai, "The working ability when transport of logs under drag and adhesive conditions of the truck THACO HD72 made in Vietnam," Journal of Forestry Science and Technology, 3, 177-184, 2018.

8. Le Van Duong, Le Anh Tuan, "Modeling and observer-based robust controllers for telescopic truck cranes," Mechanism and Machine Theory, 173, 104869, 2022. DOI: doi.org/10.1016/j.mechmachtheory.2022.104869

9. Tomasz Geisler, Wojciech Sochacki. Modelling and research into the vibrations of truck crane. Poland - 2011.

Dynamic Analysis of the PMP Pontoon Recovery System During the Winching Process

Abstract: *This paper presents a study on the dynamics of the PMP pontoon recovery system during the winch-pulling phase. The mathematical model is developed as a planar model, utilizing Lagrange's equations of the second kind to establish a system of nonlinear differential equations of motion with three generalized coordinates: the translational displacement of the pontoon block, the pitching angle around its center of gravity, and the rotation angle of the winch drum. The research takes into account practical influencing factors such as the elasticity of the cable, the sliding friction between the pontoon and the bank surface, as well as the variable driving torque characteristics of the Kraz255B base vehicle's engine. The research results provide an important theoretical basis for evaluating the stability of the mechanical system, helping to optimize the operational process and the design of specialized pontoon recovery equipment, thereby ensuring technical safety under actual working conditions.*

Keywords: Dynamics, Elastic cable, Military bridge, PMP pontoon.

Nghiên cứu Xây dựng mô hình mô phỏng mức độ mòn cặp piston - xi lanh đến các thông số hoạt động của xi lanh thủy lực bằng phần mềm MATLAB/Simulink

Lê Đình Trường1*

[1]*Học viện Kỹ thuật quân sự*

[*]*Gmail: Truongcongbinh550@gmail.com*

Tóm tắt

Bài báo nghiên cứu ảnh hưởng của mức độ hao mòn cặp piston – xi lanh đến đặc tính làm việc của xi lanh thủy lực nâng cần trên máy xúc Komatsu PC300LC-8. Trên cơ sở các phương trình cân bằng lưu lượng và cân bằng lực, mô hình toán học có xét đến hiện tượng lọt dòng nội bộ do gia tăng khe hở được xây dựng và mô phỏng bằng Matlab–Simulink với thông số kỹ thuật thực tế của máy. Các mức hao mòn khác nhau được khảo sát thông qua biến thiên hệ số lọt dòng nhằm đánh giá sự thay đổi của áp suất, vận tốc, lực tác dụng và thời gian xác lập của hệ thống. Kết quả cho thấy khi hao mòn tăng, lưu lượng lọt dòng tăng làm suy giảm áp suất và lực sinh công, đồng thời kéo dài quá trình quá độ. Đối với máy xúc tải trọng lớn như PC300LC-8, ảnh hưởng này thể hiện rõ hơn do khối lượng quán tính lớn của cơ cấu chấp hành. Nghiên cứu góp phần cung cấp cơ sở đánh giá tình trạng kỹ thuật và đề xuất giải pháp bảo trì xi lanh thủy lực trong máy xây dựng.

Từ khóa: *Xi lanh thủy lực; Áp suất; Vận tốc; Lực tác dụng; Hao mòn; Hao tổn lọt dòng, Máy xúc PC300LC-8.*

1. Đặt vấn đề

Trong các máy xây dựng cỡ lớn hiện nay, hệ thống truyền động thủy lực giữ vai trò chủ đạo trong việc truyền tải công suất và điều khiển chuyển động của các cơ cấu làm việc. Đặc biệt, đối với các máy xúc hạng trung và hạng nặng như Komatsu PC300LC-8, xi lanh thủy lực nâng cần là cơ cấu chấp hành quan trọng, trực tiếp quyết định khả năng nâng tải, độ ổn định và hiệu suất làm việc của toàn bộ thiết bị. Làm việc trong điều kiện áp suất cao, tải trọng lớn và môi trường nhiều tạp chất, các chi tiết bên trong xi lanh, đặc biệt là cặp piston – xi lanh, thường xuyên chịu tác động của ma sát và mài mòn cơ học. Quá trình hao mòn làm gia tăng khe hở làm việc, dẫn đến hiện tượng rò rỉ nội bộ, suy giảm áp suất và ảnh hưởng trực tiếp đến đặc tính động lực học của hệ thống.

Hình 1. Máy xúc bánh xích KOMATSU PC300LC-8

Trong thực tế khai thác, hao mòn xi lanh thủy lực thường chỉ được phát hiện thông qua các dấu hiệu gián tiếp như giảm lực nâng, chuyển động không ổn định hoặc rò rỉ dầu rõ rệt. Việc đánh giá định lượng ảnh hưởng của mức độ hao mòn đến các thông số vận hành như áp suất, vận tốc piston và lực tác dụng có ý nghĩa quan trọng trong công tác chẩn đoán kỹ thuật và xây dựng chiến lược bảo trì thiết bị [1]. Nhiều nghiên cứu đã chỉ ra rằng sự gia tăng khe hở giữa piston và xi lanh làm tăng lưu lượng lọt dòng và gây tổn thất năng lượng đáng kể trong hệ thống thủy lực [2], [3]. Ngoài ra, điều kiện bôi trơn, tải trọng làm việc và đặc tính vật liệu cũng ảnh hưởng trực tiếp đến tốc độ hao mòn và tuổi thọ chi tiết [4].

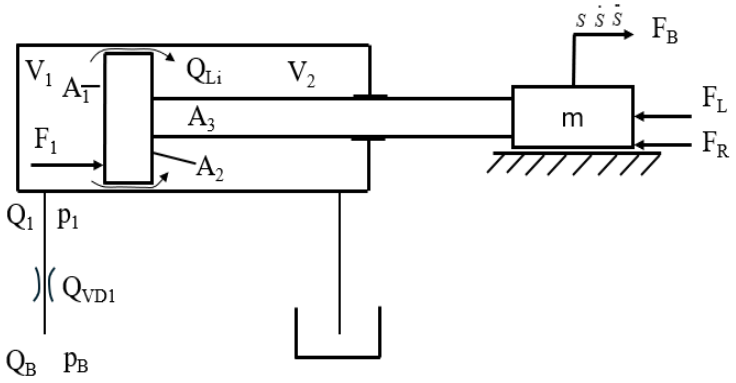
Một số công trình gần đây tập trung vào chẩn đoán rò rỉ nội bộ và phân tích ma sát trong xi lanh thủy lực [5], [6], cũng như mô phỏng mài mòn và đánh giá độ tin cậy của hệ thống bằng phương pháp số [7]. Tuy nhiên, phần lớn các nghiên cứu mới chỉ dừng lại ở việc phân tích định tính hoặc khảo sát trên các hệ thống công suất trung bình, chưa xem xét đầy đủ ảnh hưởng của hao mòn đến đặc tính quá độ và đáp ứng động của xi lanh trên các máy xúc tải trọng lớn. Đối với dòng máy như PC300LC-8, với khối lượng quán tính lớn và yêu cầu làm việc liên tục ở chế độ tải cao, sự suy giảm hiệu năng do hao mòn có thể dẫn đến biến đổi đáng kể về tính ổn định và thời gian xác lập của hệ thống.

Xuất phát từ những vấn đề trên, nghiên cứu này tập trung xây dựng mô hình toán học của xi lanh nâng cần trên máy xúc Komatsu PC300LC-8 có xét đến hiện tượng lọt dòng do hao mòn cặp piston – xi lanh, từ đó phân tích định lượng ảnh hưởng của mức độ hao mòn đến áp suất, vận tốc, lực tác dụng và đặc tính động lực học của cơ cấu chấp hành. Kết quả nghiên cứu kỳ vọng góp phần cung cấp cơ sở khoa học cho việc đánh giá tình trạng kỹ thuật và tối ưu hóa công tác bảo trì hệ thống thủy lực trên máy xây dựng cỡ lớn.

2. Cơ sở lý thuyết

Một xi lanh thủy lực thông thường được cấu thành từ các bộ phận chính như thân xi lanh, piston, cần piston, hệ thống phớt làm kín, nắp chặn và các đường dẫn dầu. Khi dầu thủy lực được cấp vào khoang làm việc với áp suất cao, áp lực tác dụng lên diện tích piston sẽ tạo ra lực đẩy làm piston chuyển động tịnh tiến, qua đó biến đổi năng lượng thủy lực thành cơ năng phục vụ quá trình nâng, hạ hoặc dịch chuyển tải trọng. Sự phối hợp nhịp nhàng giữa các chi tiết này giúp xi lanh duy trì độ kín khít, hạn chế tổn thất năng lượng và đảm bảo hiệu suất ổn định. Nhờ đó, xi lanh thủy lực trở thành bộ phận quan trọng trong việc truyền tải lực và chuyển động trong các hệ thống máy móc.

Trong điều kiện vận hành thực tế của máy xây dựng, xi lanh thủy lực thường làm việc với tần suất lớn, áp suất cao và môi trường nhiều bụi bẩn, tạp chất. Các yếu tố này làm gia tăng ma sát và gây mài mòn bề mặt tiếp xúc giữa piston và lòng xi lanh, cũng như tại vị trí cần piston và các phớt làm kín. Khi mức độ hao mòn tăng lên, khe hở làm việc giữa các bề mặt tiếp xúc mở rộng, dẫn đến hiện tượng rò rỉ dầu trong nội bộ xi lanh. Ngoài ra, rò rỉ bên ngoài cũng có thể xuất hiện tại khu vực tiếp giáp giữa cần piston và thân xi lanh hoặc tại các phớt đã suy giảm khả năng làm kín. Hình 2 trình bày sơ đồ xi lanh thủy lực trong trường hợp xuất hiện hao mòn tại cặp piston – lòng xi lanh, đây là một trong những nguyên nhân chính gây ra hiện tượng lọt dòng trong quá trình làm việc.



Hình 2. Hao mòn giữa piston – lòng xi lanh

Q_B – Lưu lượng của bơm thủy lực; p_B – Áp suất làm việc của bơm; V_1 , V_2 – Thể tích khoang bên trái, bên phải piston; A_1 – Diện tích tác động bên trái piston; A_2 – Diện tích tác động bên phải piston; A_3 – Diện tích cần piston; Q_{Li} – Lưu lượng dầu lọt; p_1 , p_2 – Áp suất tác dụng lên diện tích bên trái, bên phải piston; m – Khối lượng quy đổi về piston; F_L – Ngoại lực tác động quy đổi về piston; F_R – Lực ma sát quy đổi về piston; S – Độ dịch chuyển của piston.

Từ Hình 2 ta nhận thấy rằng, hao mòn giữa piston – lòng xi lanh sẽ ảnh hưởng lớn đến hành trình tiến của piston

Theo sơ đồ ở Hình 2, phương trình cân bằng lưu lượng tại khoang bên trái piston có thể được thiết lập như sau [8]:

$$Q_1 = Q_B - Q_{VD1} - Q_{Li1} - Q_{V1} \quad (1)$$

Trong đó:

Q_B là lưu lượng do bơm thủy lực cung cấp (lít/phút);

Q_1 là lưu lượng thực tế đi vào khoang bên trái piston (lít/phút);

Q_{VD1} là tổn thất lưu lượng qua van tiết lưu trong mạch thủy lực, có thể biểu diễn dưới dạng $Q_{VD1} = K\sqrt{p_1}$, với Klà hệ số đặc trưng của van và $p_1$là áp suất trong khoang (N/m²);

$Q_{V1} = A_1\dot{S}$ là lưu lượng do sự dịch chuyển của piston;

$Q_{Li1} = G_{Li1}p_1$ là lưu lượng dầu rò rỉ qua khe hở làm việc.

Đối với khe hở dạng vành khăn giữa piston và lòng xi lanh, hệ số lọt dòng được xác định theo quan hệ:

$$G_{Li} = \frac{d_m\pi\delta_m^3}{12vl\rho}$$

Trong đó:

d_m là đường kính xi lanh (m);

δ_m là khe hở trung bình (m);

ρ là trọng lượng riêng của dầu (kg/m³);

v là vận tốc dòng chảy (m/s);

l là chiều dài đoạn khe hở chịu tác động.

Phương trình cân bằng lực tác dụng lên piston được viết dưới dạng [8]:

$$F_B = F_1 - F_L - F_R \quad (2)$$

với $F_1 = A_1 p_1$ là lực thủy lực sinh ra trên diện tích tác dụng của piston;

F_L là ngoại lực quy đổi

$F_R = kS$ là lực ma sát quy đổi về đầu piston.

Từ các quan hệ trên, có thể xây dựng hệ phương trình vi phân mô tả động học và động lực học của piston:

$$\frac{dp}{dt} = \frac{1}{V_1 \beta} Q_1 \quad (3)$$

$$\frac{d^2 S}{dt^2} = \frac{F_B}{m} \quad (4)$$

Trong đó V_1 là thể tích khoang làm việc và β là hệ số đàn hồi thể tích của dầu.

Các phương trình cân bằng lưu lượng và cân bằng lực đối với hành trình lùi của piston được thiết lập tương tự như trường hợp hành trình tiến.

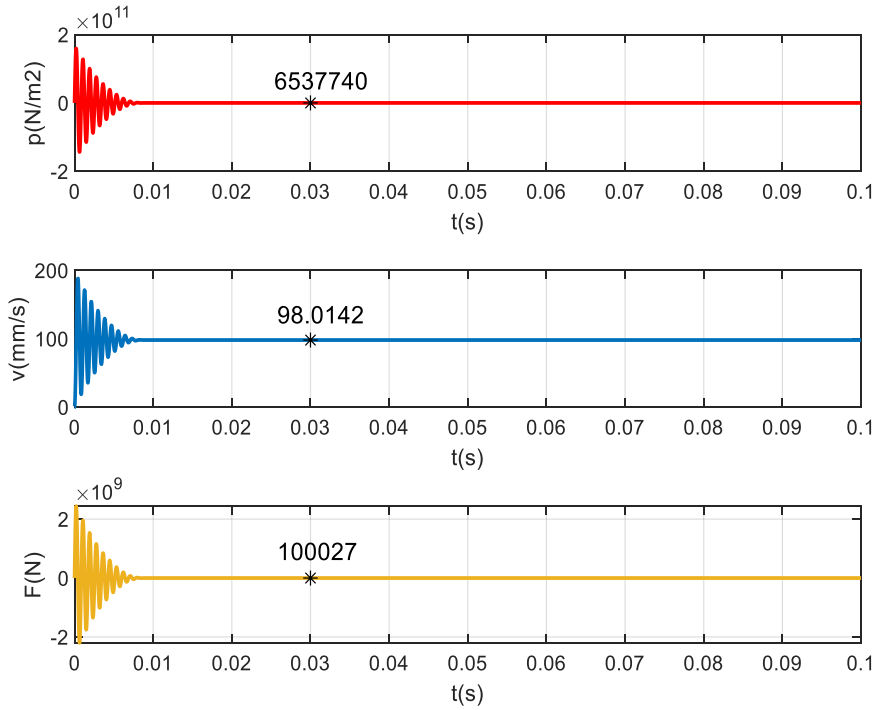
3. Kết quả và thảo luận

Dựa trên cơ sở lý thuyết và hệ phương trình đã xây dựng ở phần trên, nghiên cứu tiến hành mô phỏng và phân tích trên môi trường Matlab–Simulink đối với xi lanh nâng cần của máy xúc Komatsu PC300LC-8. Các thông số kỹ thuật của máy và xi lanh nâng cần được sử dụng trong mô hình được tổng hợp từ tài liệu kỹ thuật tương ứng [9], [10], làm cơ sở cho việc thiết lập điều kiện đầu vào và đánh giá kết quả mô phỏng.

Bảng 1. Thông số kỹ thuật của máy xúc Komatsu PC300LC-8

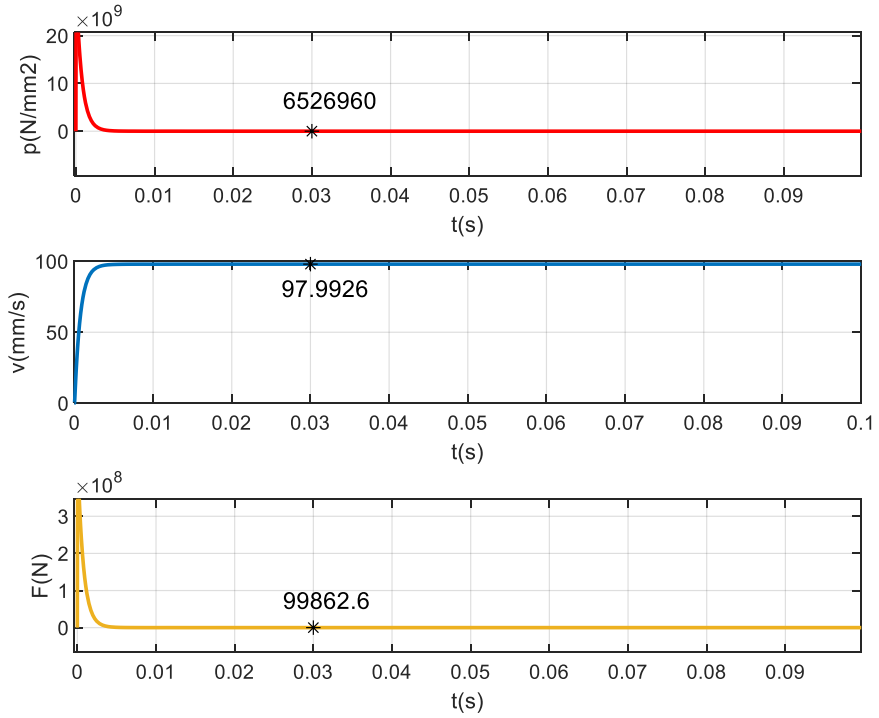
STT	Thông số	Ký hiệu	Giá trị	Đơn vị
1	Động cơ	SAA6D114E		
2	Công suất định mức	P	183,9/1,950	kW/rpm
3	Lưu lượng tối đa của hệ thống thủy lực		535	lít/phút
4	Áp suất mạch thực thi		34,8	MPa
5	Áp suất mạch quay toa		27,9	MPa
6	Áp suất mạch di chuyển		34,8	MPa
7	Áp suất mạch điều khiển		3,4	MPa
8	Đường kính trong xi lanh nâng cần		140	mm
9	Hành trình xi lanh nâng cần	H	1480	mm
10	Đường kính cần piston		100	mm
11	Đặc tính van	k	$1.5.10^{-7}$	$m^4\sqrt{N}.s$
12	Hệ số quy đổi	β	10^{-13}	m^2/N
13	Ngoại lực	F_L	2000	N
14	Diện tích tác động bên trái piston	A_1	0,0153	m^2
15	Khối lượng quy đổi về đầu piston	M	3500	kg
16	Hệ số ma sát quy đổi về đầu piston	K	1000	N
17	Giá trị lọt dòng	G_{Li}	$0 \div 5.10^{-09}$	$\frac{m^5}{s.N}$

Khi xi lanh mới không có hao mòn giữa cặp piston – xi lanh, do đó giá trị lọt dòng G_{Li}=0 , khi đó áp suất, vận tốc và lực tác động của piston thu được như sau:

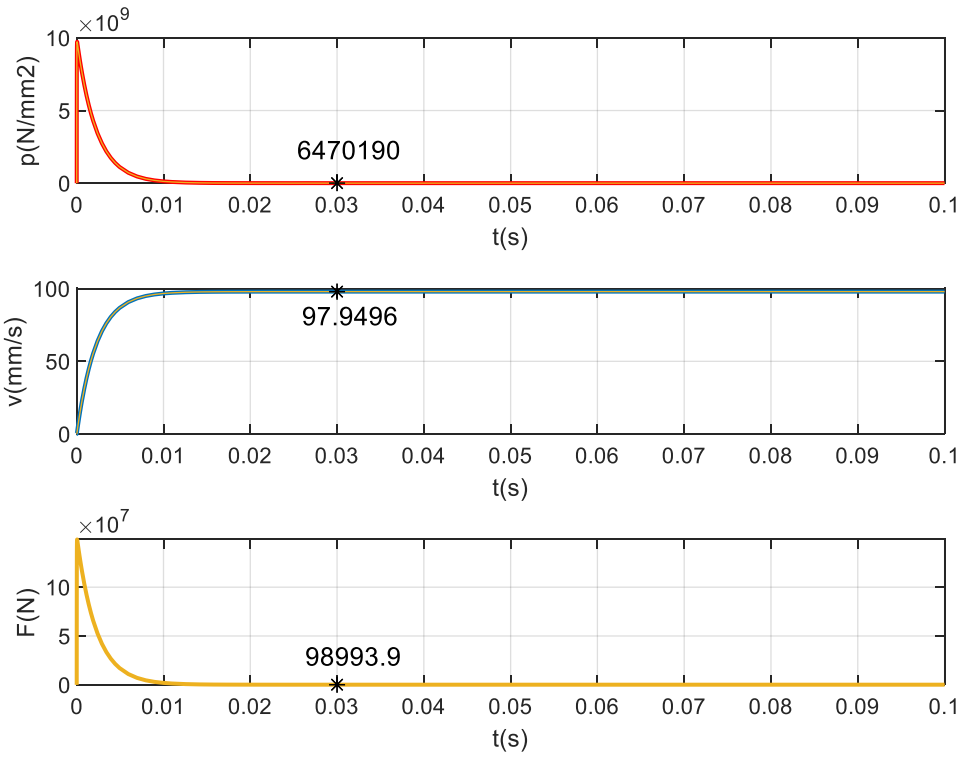


Hình 3. Đồ thị áp suất, vận tốc và lực tác động của piston khi piston – xi lanh chưa có hao mòn với giá trị lọt dòng $G_{Li} = 0\ \frac{m^5}{s.N}$

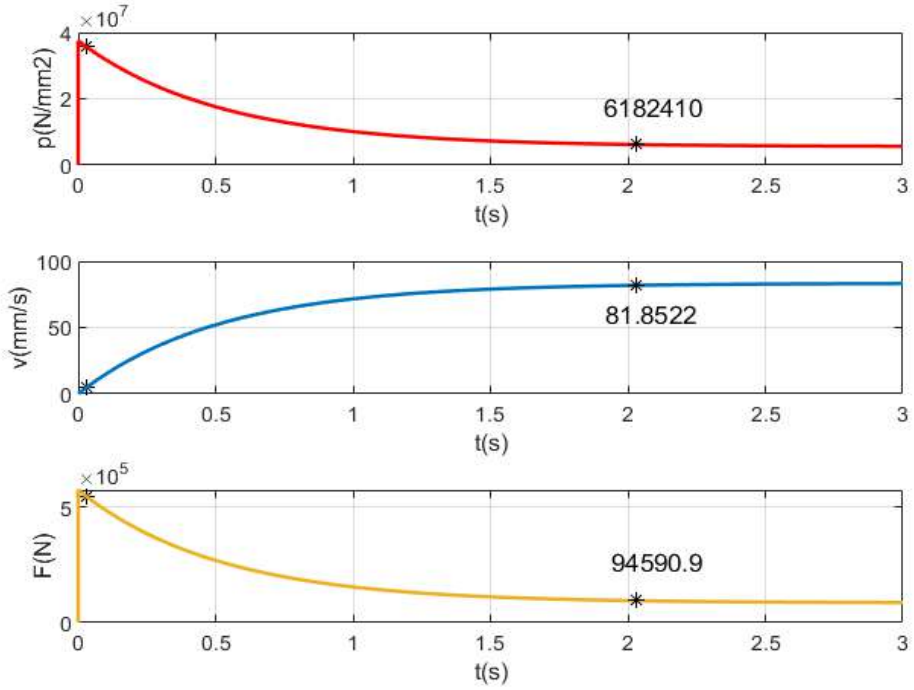
- Đối với quá trình tiến của piston, khi khe hở giữa piston – lòng xi lanh tăng lên tương ứng với nó là giá trị lọt dòng G_{Li} tăng lên



Hình 4. Xy lanh có khe hở với giá trị lọt dòng $G_{Li} = 5.10^{-11}$ $\frac{m^5}{s.N}$



Hình 5. Xy lanh có khe hở với giá trị lọt dòng $G_{Li} = 1.5000e\text{-}10$ $\frac{m^5}{s.N}$



Hình 6. Xy lanh có khe hở với giá trị lọt dòng $G_{Li} = 4.0000e\text{-}08$ $\frac{m^5}{s.N}$

Khi piston thực hiện hành trình tiến trong điều kiện xuất hiện khe hở giữa piston và lòng xi lanh, nghiên cứu tiến hành khảo sát với các giá trị hệ số lọt dòng khác nhau, bao gồm $G_{Li} = 5 \times 10^{-11}$, 1.5×10^{-10}và 4.0×10^{-8}m^5/(s·N).

Kết quả cho thấy, ở mức hao mòn nhỏ ($G_{Li} = 5 \times 10^{-11}$), các thông số làm việc của xi lanh như áp suất, vận tốc và lực tác dụng chỉ suy giảm rất nhẹ so với trường hợp không có hao mòn. Đáp ứng quá độ vẫn diễn ra nhanh và hệ thống đạt trạng thái ổn định trong thời gian ngắn, cho thấy ảnh hưởng của rò rỉ nội bộ chưa đáng kể.

Tuy nhiên, khi giá trị G_{Li}tăng lên, đặc biệt ở mức 4.0×10^{-8}, sự suy giảm trở nên rõ rệt. Áp suất làm việc giảm đáng kể, vận tốc piston đạt giá trị ổn định thấp hơn và lực sinh công giảm mạnh so với trạng thái ban đầu. Đồng thời, thời gian xác lập kéo dài, biểu hiện sự suy giảm độ cứng thủy lực và khả năng đáp ứng động của hệ thống. Điều này chứng tỏ năng lượng thủy lực bị tổn thất lớn qua khe hở làm việc, dẫn đến giảm hiệu suất vận hành của xi lanh nâng cần trên máy PC300LC-8.

Từ các đồ thị thu được trong các trường hợp khảo sát, có thể nhận thấy mức độ hao mòn của cặp piston – xi lanh thể hiện thông qua sự gia tăng hệ số lọt dòng G_{Li}có ảnh hưởng rõ rệt đến đặc tính làm việc của xi lanh. Ở trạng thái không có hao mòn ($G_{Li} = 0$), hệ thống đạt giá trị ổn định rất nhanh, chỉ sau khoảng 0,01–0,03 s, với áp suất xấp xỉ 6,54 MPa, vận tốc khoảng 98 mm/s và lực sinh công gần 100 kN.

Khi mức hao mòn nhỏ ($G_{Li} = 5 \times 10^{-11}$ và 1.5×10^{-10}), các thông số làm việc chỉ suy giảm nhẹ, sai khác so với trường hợp lý tưởng là không đáng kể và hệ thống vẫn duy trì thời gian xác lập ngắn. Tuy nhiên, khi hao mòn tăng lên đáng kể ($G_{Li} = 4.0 \times 10^{-8}$), ảnh hưởng trở nên rõ rệt: áp suất ổn định giảm xuống còn khoảng 6,18 MPa, vận tốc giảm còn khoảng 81,85 mm/s và lực tác dụng giảm xuống gần 94,6 kN. Đồng thời, thời gian xác lập của hệ thống kéo dài đến khoảng 2 s. Kết quả này cho thấy sự gia tăng rò rỉ nội bộ làm suy giảm độ cứng thủy lực và khả năng đáp ứng động của xi lanh, ảnh hưởng trực tiếp đến hiệu suất làm việc của máy PC300LC-8.

4. Kết luận

Kết quả mô phỏng cho thấy mức độ hao mòn của cặp piston – xi lanh, được đặc trưng thông qua hệ số lọt dòng G_{Li}, có ảnh hưởng trực tiếp đến đặc tính làm việc và đáp ứng động của xi lanh nâng cần trên máy xúc Komatsu PC300LC-8. Khi không có hao mòn, hệ thống đạt trạng thái ổn định trong thời gian rất ngắn và các thông số làm việc tiệm cận giá trị thiết kế. Ở mức hao mòn nhỏ, sự suy giảm áp suất, vận tốc và lực tác dụng chưa đáng kể. Tuy nhiên, khi giá trị G_{Li}tăng lên đến mức lớn, áp suất làm việc giảm rõ rệt, vận tốc piston suy giảm đáng kể và lực sinh công giảm khoảng 5–6%, đồng thời thời gian xác lập kéo dài từ vài phần trăm giây lên đến khoảng 2 giây. Điều này cho thấy hao mòn không chỉ làm suy giảm hiệu suất mà còn ảnh hưởng đến độ ổn định và khả năng đáp ứng động của hệ thống thủy lực.

Đóng góp chính của nghiên cứu là xây dựng được mô hình toán học có xét đến ảnh hưởng của rò rỉ nội bộ do hao mòn và lượng hóa được tác động của mức độ lọt dòng đến các thông số động lực học của xi lanh trên máy xúc tải trọng lớn. Kết quả cho thấy hệ số G_{Li}có thể sử dụng như một chỉ tiêu kỹ thuật để đánh giá tình trạng làm việc của xi lanh, làm cơ sở đề xuất ngưỡng cảnh báo và lập kế hoạch bảo trì, thay thế hợp lý nhằm đảm bảo độ tin cậy và tuổi thọ của thiết bị.

Tài liệu tham khảo

1. V. A. Zorin and N. T. Minh, “Prediction of the Residual Service Life of Hydraulic Motors of Construction Machines Based on the Results of Simulation Modeling,” *Steel Transl.*, vol. 54, no. 7, pp. 676–680, Jul. 2024. doi: 10.3103/S0967091224701341.
2. N. Novak, A. Trajkovski, M. Kalin, and F. Majdič, “Degradation of Hydraulic System due to Wear Particles or Medium Test Dust,” *Applied Sciences*, vol. 13, no. 13, Art. no. 13, 2023. doi: 10.3390/app13137777.
3. O. V. Bereziuk, V. I. Savulyak, V. O. Kharzhevskyi, and A. Y. Alekseiev, “Determination of the Regularity of the Rate of Wear of the Working Hydraulic Cylinder of the Mechanism of the Sealing Plate of the Garbage Truck from the Pressing Force,” *Problems of Tribology*, vol. 29, no. 1/111, 2024. doi: 10.31891/2079-1372-2024-111-1-38-44.
4. J. J. Truhan, J. Qu, and P. J. Blau, “The Effect of Lubricating Oil Condition on the Friction and Wear of Piston Ring and Cylinder Liner Materials in a Reciprocating Bench Test,” *Wear*, vol. 259, no. 7, pp. 1048–1055, 2005. doi: 10.1016/j.wear.2005.01.025.
5. P. Zhang and X. Chen, “Internal Leakage Diagnosis of a Hydraulic Cylinder Based on Optimization DBN Using the CEEMDAN Technique,” *Shock and Vibration*, vol. 2021, Article ID 8856835, 2021. doi: 10.1155/2021/8856835.
6. K. Ma, J. Wang, and L. Gu, “Experimental Study on Friction of Hydraulic Cylinder in Different Sealing Systems,” *MATEC Web of Conferences*, vol. 153, p. 06012, 2018. doi: 10.1051/matecconf/201815306012.
7. H. Yan, J. Li, C. Cai, and Y. Ren, “Numerical Investigation of Erosion Wear in the Hydraulic Amplifier of the Deflector Jet Servo Valve,” *Applied Sciences*, vol. 10, no. 4, 2020. doi: 10.3390/app10041299.

Study on Developing a Simulation Model of Piston–Cylinder Wear and Its Impact on Hydraulic Cylinder Operating Parameters Using MATLAB/Simulink

Abstract

This paper investigates the influence of piston–cylinder wear on the operating characteristics of the boom hydraulic cylinder in a Komatsu PC300LC-8 excavator. Based on flow continuity and force balance equations, a mathematical model incorporating internal leakage caused by increased clearance is developed and implemented in Matlab–Simulink using actual technical parameters of the machine. Different wear levels are examined through variations in the internal leakage coefficient to evaluate changes in working pressure, piston velocity, output force, and settling time. The results indicate that increasing wear leads to higher internal leakage, resulting in reduced pressure and actuation force, as well as prolonged transient response. For heavy-duty excavators such as the PC300LC-8, these effects become more pronounced due to the large equivalent inertial mass of the actuator system. The study provides a scientific basis for condition assessment and maintenance planning of hydraulic cylinders in construction machinery.

Keywords: *Hydraulic cylinder; Pressure; Velocity; Force; Wear; Leakage loss, PC300LC-8 Excavator*

Nghiên cứu đặc tính thủy động lực học của xe thiết giáp chở quân chiến đấu trong quá trình bơi sử dụng phương pháp số.

Lương Trần Anh Duy[1], Hà Trường Sang[1], Nguyễn Duy Đạt[1], Nghiêm Văn Long[1]

[1]Học viện Kỹ thuật Quân sự

Email: anhduyluongtran@gmail.com; Tel: 0354648491

Tóm tắt

Bài báo trình bày kết quả nghiên cứu mô phỏng số dòng chảy bao quanh xe thiết giáp chở quân (APC) mẫu XTC-02 khi hoạt động trong môi trường nước. Nhóm tác giả sử dụng phương pháp động lực học chất lưu số (CFD) với phần mềm ANSYS Fluent, áp dụng phương pháp thể tích hữu hạn (FVM) kết hợp mô hình đa pha VOF (Volume of Fluid) để giải quyết bài toán mặt thoáng. Nghiên cứu tập trung phân tích các thành phần lực cản, trường áp suất và sự hình thành sóng mũi tại các vận tốc bơi từ 0.5 m/s đến 3 m/s. Mô hình hình học được xây dựng theo tỉ lệ 1:1 trong một "bể thử ảo" với lưới đa diện. Kết quả cho thấy khi vận tốc tăng, lực cản tăng nhanh trong khi lực nâng có xu hướng giảm nhẹ (ở trạng thái chưa lướt). Đặc biệt, nghiên cứu so sánh hiệu quả của tấm chắn sóng: mặc dù giúp giảm lượng nước tràn qua mũi xe và tạo vùng áp suất ổn định hơn, nhưng việc trang bị tấm chắn sóng trong cấu hình hiện tại lại làm tăng lực cản tổng thể của xe lên khoảng 8,4% và tăng lực nâng. Kết quả này cung cấp cơ sở dữ liệu quan trọng cho việc tối ưu hóa thiết kế tấm chắn sóng và hình dáng mũi xe trong tương lai.

Từ khóa: *Xe thiết giáp chở quân (APC), mô phỏng số, CFD, ANSYS-Fluent, sóng mũi.*

1. Mở đầu

Trong bối cảnh tác chiến hiện đại, khả năng cơ động dưới nước đóng vai trò rất quan trọng đối với các dòng xe thiết giáp chở quân (APC), đảm bảo tính liên tục và bất ngờ của các chiến dịch quân sự khi vượt qua chướng ngại nước tự nhiên, đáp ứng yêu cầu ngày càng cao của tác chiến đổ bộ trên biển. Tuy nhiên, việc thiết kế các phương tiện này luôn được ưu tiên cho nhiệm vụ bảo vệ đạn đạo và bố trí không gian trên cạn, dẫn đến hình thái khí động học không được tối ưu cho nhu cầu hoạt động dưới nước nên hệ số cản nước rất lớn khi hoạt động trong môi trường nước. Để nghiên cứu nâng cao khả năng hoạt động dưới nước khi các phương pháp thực nghiệm truyền thống như thử nghiệm mô hình trong bể thử cung cấp độ tin cậy cao nhưng lại đòi hỏi chi phí lớn và thời gian chuẩn bị dài, thì sự phát triển của công nghệ tính toán hiệu năng cao đã đưa phương pháp Động lực học chất lưu số (CFD) trở thành công cụ đắc lực, cho phép dự đoán chính xác trường dòng chảy và tương tác đa pha phức tạp ngay từ giai đoạn thiết kế ý tưởng [4].

Trên cơ sở đó, nghiên cứu này tập trung ứng dụng phương pháp số để phân tích đặc tính thủy động lực học của một mẫu xe APC điển hình, cụ thể là đánh giá các thành phần lực cản, trường áp suất và sự hình thành sóng mũi, nhằm cung cấp cơ sở dữ liệu khoa học cho việc tối ưu hóa thiết kế tấm chắn sóng và nâng cao hiệu quả bơi của phương tiện.

2. Phương pháp mô phỏng số

2.1. Miền tính toán và lưới

Để mô phỏng chính xác các đặc tính thủy động lực học, việc xây dựng mô hình hình học và thiết lập miền tính toán đóng vai trò tiên quyết. Cách tiếp cận xây dựng miền tính toán trong nghiên cứu này tuân theo phương pháp "bể thử ảo".

2.1.1. Mô hình hình học (Geometric Model)

Chọn đối tượng nghiên cứu là xe thiết giáp chở quân XTC-02. Để tối ưu hóa tài nguyên tính toán và đảm bảo chất lượng lưới, mô hình hình học 3D của xe được đơn giản hóa các chi tiết nhỏ không ảnh hưởng đáng kể đến trường dòng chảy như: đinh tán, tay nắm cửa, đèn tín hiệu và các gờ nhỏ trên thân xe.

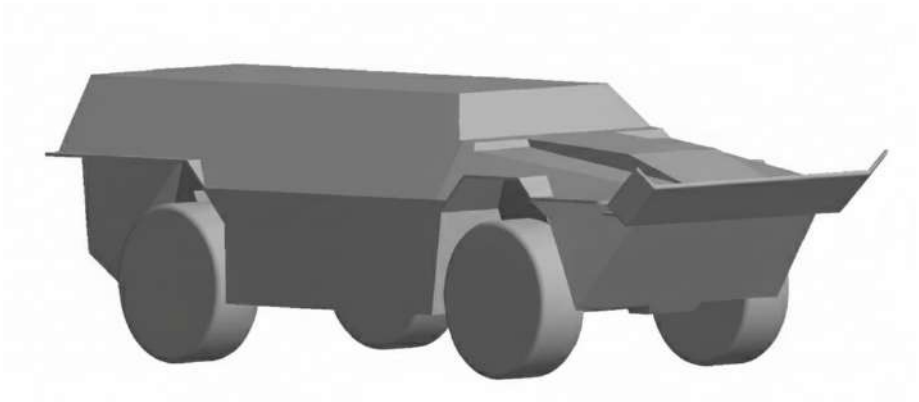
Các thông số kích thước bao cơ bản của mô hình được giữ nguyên theo tỉ lệ thực (1:1):

Bảng 1. Các thông số của xe

Chiều dài tổng thể (L)	Chiều rộng (W)		Chiều cao (H)
7,2 m	2,55 m		2,98 m (tính đến nắp tháp pháo)

Đặc biệt, chi tiết tấm chắn sóng ở mũi xe được giữ nguyên hình dáng và góc nghiêng, do đây là bộ phận quan trọng nhất ảnh hưởng đến khả năng đè sóng và lực cản của xe khi bơi. Bánh xe được giả thiết ở vị trí cố định (không quay) và nằm trong hốc bánh xe để mô phỏng trạng thái bơi bằng hệ thống đẩy phản lực nước.

Hình 1. Mô hình thực tế xe lội nước



Hình 2. Mô hình CAD 3D cho mô phỏng của xe lội nước

2.1.2. Miền tính toán

Miền tính toán được xây dựng dưới dạng hình hộp chữ nhật, bao trùm toàn bộ xe, mô phỏng một bể thử không giới hạn. Kích thước miền tính toán cần đủ lớn để triệt tiêu ảnh hưởng của điều kiện biên lên trường dòng chảy quanh xe. Miền chất lưu trong nghiên cứu này được chia thành hai pha riêng biệt: pha nước (phía dưới) và pha khí (phía trên).

Kích thước miền tính toán được xác định dựa trên chiều dài tham chiếu của xe (L):

+ Đầu vào (Inlet): Cách mũi xe một khoảng 2L để đảm bảo dòng chảy ổn định trước khi tiếp xúc vật thể.

+ Đầu ra (Outlet): Cách đuôi xe một khoảng 5L để quan sát trọn vẹn vùng xoáy phía sau lái và tránh hiện tượng dội ngược áp suất từ biên.

+ Biên bên: Cách thân xe tối thiểu 2W để tránh hiệu ứng tắc nghẽn.

+ Biên trên: Cách nóc xe 1.5H (tương ứng với pha khí).

+ Biên dưới: Cách điểm thấp nhất của bánh xe 2H (tương ứng với pha nước).

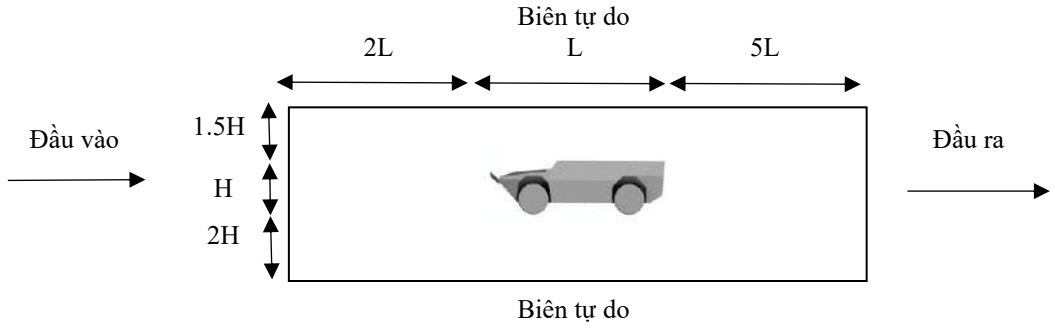
Các điều kiện biên được thiết lập như sau:

+ Inlet: Vận tốc không đổi cho cả pha nước và khí, tương ứng với vận tốc bơi của xe.

+ Outlet: Áp suất tĩnh (Pressure Outlet).

+ Thân xe: Điều kiện biên tường trơn, không trượt.

+ Các mặt bên và đáy: Điều kiện biên đối xứng hoặc trượt tự do để mô phỏng vùng nước rộng vô hạn.



Hình 3. Miền tính toán 3D

2.1.3. Lưới mô phỏng

Chất lượng lưới có ảnh hưởng trực tiếp đến độ chính xác và sự hội tụ của bài toán số. Trong nghiên cứu này, phương pháp chia lưới đa diện được sử dụng nhờ ưu điểm giảm số lượng phần tử nhưng vẫn giữ được độ chính xác cao so với lưới tứ diện, đồng thời giúp quá trình giải lặp hội tụ nhanh hơn.

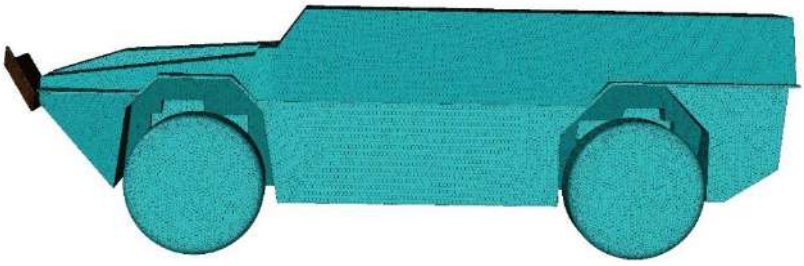
Chiến lược chia lưới tập trung vào các vùng có gradient vận tốc và áp suất lớn:

+ Lớp biên: Sử dụng các lớp lưới lăng trụ bao quanh bề mặt thân xe để giải quyết lớp biên dính nhớt. Chiều cao lớp lưới đầu tiên được tính toán sao cho giá trị y+ nằm trong khoảng phù hợp với mô hình rối $k-\epsilon$ ($30 < y+ < 300$ cho hàm tường chuẩn hoặc $y+ \approx 1$ cho mô hình tăng cường).

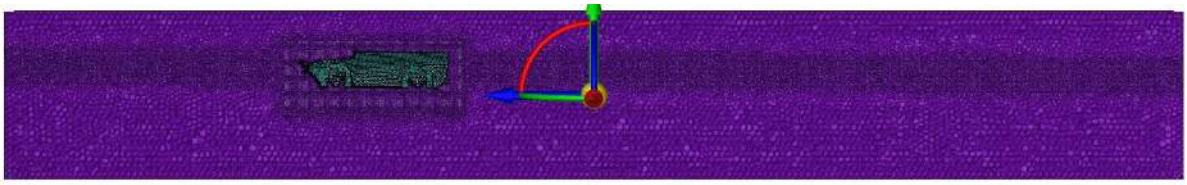
+ Vùng mặt thoáng: Tại khu vực giao diện phân cách nước - khí, lưới được tăng dày cục bộ theo phương thẳng đứng để mô hình VOF có thể bắt chính xác sự biến dạng của sóng mũi và sóng rẽ.

+ Vùng bao quanh xe: Một vùng ảnh hưởng hình hộp bao quanh xe được thiết lập với kích thước phần tử nhỏ hơn để bắt các hiện tượng tách dòng và xoáy rối cục bộ.

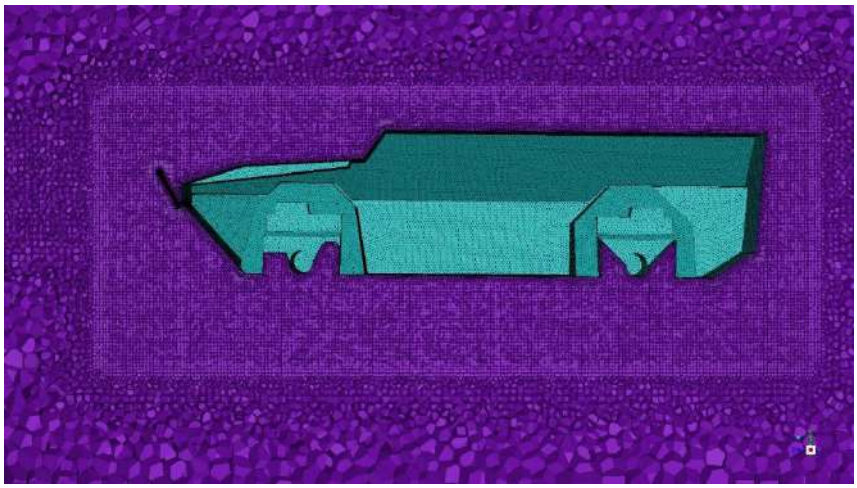
Lưới sẽ được đánh giá thông qua sự hội tụ của hệ số lực cản (C_d) với các độ mịn khác nhau để chọn ra phương án tối ưu giữa độ chính xác và chi phí tính toán.



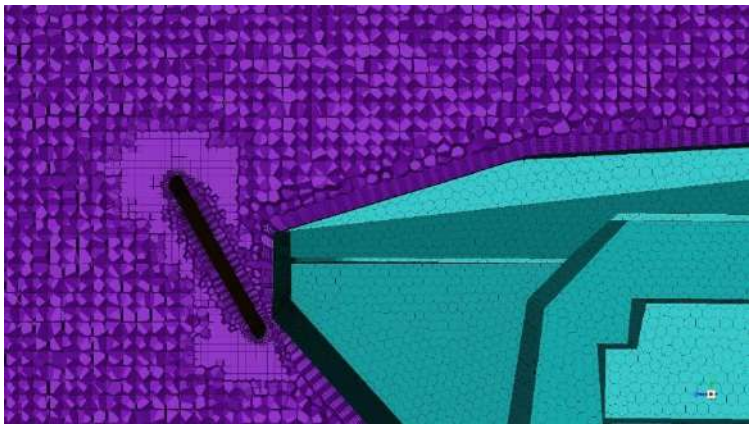
Hình 4. Phân bố lưới xung quanh bề mặt xe



Hình 5. Phân bố lưới toàn bộ miền tính toán (mô hình 2,5 triệu lưới)



Hình 6. Lưới phân bố bao quanh thân xe



Hình 6. Lớp biên bao quanh thân xe và tấm chắn sóng

2.2. Mô hình toán học

Dòng chảy bao quanh xe thiết giáp trong quá trình bơi được giả thiết là dòng chảy rối, không nén được và không dừng. Để mô phỏng sự tương tác giữa hai pha nước và không khí, phương pháp thể tích hữu hạn (FVM) dựa trên mô hình đa pha VOF (Volume of Fluid) được áp dụng.

2.2.1. Các phương trình chủ đạo

Các phương trình bảo toàn khối lượng và động lượng cho hỗn hợp chất lưu được viết dưới dạng trung bình hóa Reynolds (RANS) [3] như sau:

Phương trình liên tục:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho u) = 0 \quad (1)$$

Trong đó ρ là khối lượng riêng của hỗn hợp, u là vector vận tốc trung bình.

Phương trình động lượng [3]:

$$\frac{\partial}{\partial t}(\rho u) + \nabla \cdot (\rho u u) = -\nabla p + \nabla \cdot \left[\mu_{eff}(\nabla u + \nabla u^T)\right] + \rho g + F \quad (2)$$

Trong đó:

p là áp suất tĩnh.

μ_{eff} là độ nhớt hiệu dụng, được tính bằng tổng độ nhớt động lực học (μ) và độ nhớt rối (μ_t): $\mu_{eff} = \mu + \mu_t$.

g là gia tốc trọng trường.

F là các ngoại lực tác dụng (nếu có).

2.2.2. Mô hình rối

Do đặc điểm hình học của xe thiết giáp với phần mũi và các hốc bánh xe gây ra sự tách dòng và nhiễu loạn lớn, mô hình rối $k-\epsilon$ Realizable được sử dụng để khép kín hệ phương trình RANS. Mô hình này cho độ chính xác cao hơn mô hình chuẩn trong các bài toán có dòng chảy xoáy mạnh và phân tách dòng [2].

Phương trình vận chuyển cho động năng rối (k) và tốc độ tiêu tán rối (ϵ) [2]:

$$\frac{\partial}{\partial t}(\rho k) + \nabla \cdot (\rho k u) = \nabla \cdot \left[\left(\mu + \frac{\mu_t}{\sigma_k}\right)\nabla k\right] + G_k - \rho\epsilon \quad (3)$$

$$\frac{\partial}{\partial t}(\rho \epsilon) + \nabla \cdot (\rho \epsilon u) = \nabla \cdot \left[\left(\mu + \frac{\mu_t}{\sigma_\epsilon}\right)\nabla \epsilon\right] + \rho C_1 S\epsilon - \rho C_2 \frac{\epsilon^2}{k+\sqrt{\nu\epsilon}} \quad (4)$$

Trong đó G_k đại diện cho sự sinh động năng rối do gradient vận tốc trung bình; σ_k và σ_ϵ là các số Prandtl rối; C_1, C_2 là các hằng số thực nghiệm của mô hình.

2.2.3. Mô hình đa pha VOF

Để xác định và theo dõi bề mặt tự do (mặt thoáng) giữa nước và không khí, mô hình VOF được sử dụng. Vị trí của mặt thoáng được xác định thông qua phương trình vận chuyển phân thể tích của pha nước (α_w) [8]:

$$\frac{\partial \alpha_w}{\partial t} + \nabla \cdot (\alpha_w u) = 0 \quad (5)$$

Phương trình này không bao gồm số hạng nguồn do không xét đến quá trình chuyển pha (sôi/ngưng tụ) trong bài toán thủy động lực học xe.

Phân thể tích α tuân theo điều kiện ràng buộc [8]:

$$\alpha_w + \alpha_a = 1 \quad (6)$$

Trong đó $\alpha_w = 1$ tương ứng với vùng chất lỏng hoàn toàn (nước), $\alpha_w = 0$ tương ứng với vùng khí hoàn toàn (không khí), và $0 < \alpha_w < 1$ đại diện cho vùng giao diện mặt thoáng.

Các thuộc tính vật lý của chất lưu trong mỗi ô lưới (như khối lượng riêng ρ và độ nhớt μ) được tính toán dựa trên trung bình theo thể tích của các pha thành phần [8]:

$$\rho = \alpha_w \rho_w + (1 - \alpha_w)\rho_a \quad (7)$$

$$\mu = \alpha_w \mu_w + (1 - \alpha_w)\mu_a \quad (8)$$

Trong đó các chỉ số dưới w và a lần lượt ký hiệu cho pha nước và pha khí. Các thông số vật lý này được cập nhật liên tục trong quá trình giải lặp.

2.3. Thiết lập bộ giải

Các điều kiện biên cho mô hình mô phỏng và các thiết đặt cho mô phỏng bằng phần mềm thương mại ANSYS Fluent được thể hiện trên bảng 2 [5].

Bảng 2. Trình tự thiết lập thuộc tính giải trong ANSYS Fluent

Các thiết lập chung	
Trình giải	Pressure-Based (Cho dòng có gradient vận tốc nhỏ)
Mô hình thời gian	Steady (Tĩnh)
Trọng lực	-9,81 m/s^2
Mô hình nhiều pha	
Mô hình đồng nhất	Volume of Fluid
Mô hình VOF	Open Channel Flow
Cài đặt pha	Khi-Primary Phase Nuoc-Secondary Phase
Mô hình dòng chảy	
Mô hình chảy rối	k- ϵ
Môi trường	
Không khí	air
Nước	water-liquid
Điều kiện biên	
Inlet (đầu vào)	Pressure inlet: V = V_{xe}.
Outlet (đầu ra)	Pressure Outlet (Áp suất ra)
Biên thành xe	No slip wall (Tường tuyệt đối cứng, đẳng nhiệt)
Wall (biên tường)	Specified shear (biên tự do)
Thiết lập cho trình giải	
Phương pháp giải	Chế độ Auto Time scale với Time scale Factor bằng 0,001
Điều kiện hội tụ	Sai số hội tụ là RMS với sai số ≤ 10^{-3}
Số vòng lặp	1000
Monitors	Hiển thị đồ thị các hệ số khí động theo từng vòng lặp

3. Kết quả và thảo luận

3.1. Kiểm tra độc lập của lưới

Để kiểm tra chất lượng lưới, ta khảo sát 3 loại lưới: lưới thô (1,7 triệu cell), lưới trung bình (2,5 triệu cell), lưới mịn (6,1 triệu cell). Sử dụng mô hình xe không có tấm chắn sóng, bơi với vận tốc 0,5m/s để mô phỏng và so sánh kết quả.

Bảng 3. Các loại lưới khảo sát

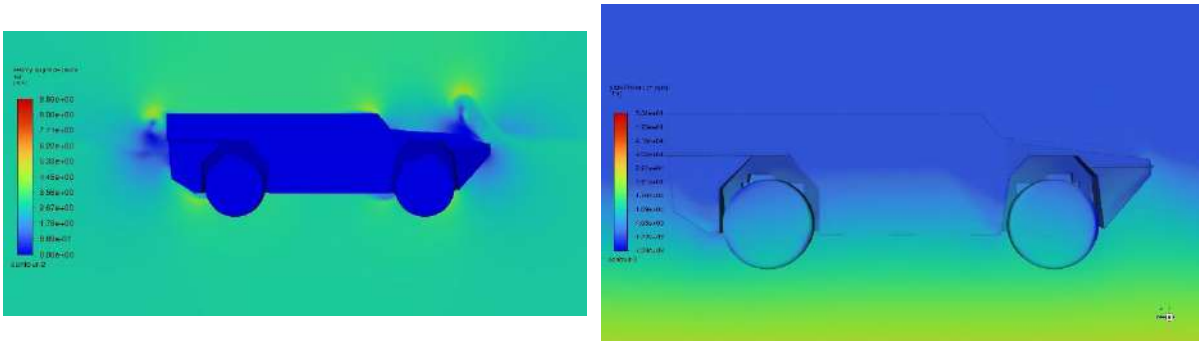
STT	Loại lưới	Số phần tử (triệu cell)	Lực cản Cd (N)	Lực nâng Cl (N)	Thời gian tính toán (h)
Lưới 1	Thô	1,7	1633,3992	102548,73	3
Lưới 2	Trung bình	2,5	1777,6407	102551,43	4,5
Lưới 3	Mịn	6,1	1735,5881	102550,87	9

Từ kết quả trong bảng 3 ta nhận thấy kết quả lưới thứ 2 (2,5 triệu cell) gần tương đương với kết quả lưới thứ 3 (6,1 triệu cell) với sai số 2,4% trong khi thời gian tính toán ngắn hơn

nhiều. Vì vậy, để tối ưu tài nguyên tính toán, các khảo sát tiếp theo sử dụng lưới 2-3 triệu cell để mô phỏng và tính toán.

3.2. Ảnh hưởng của vật tốc

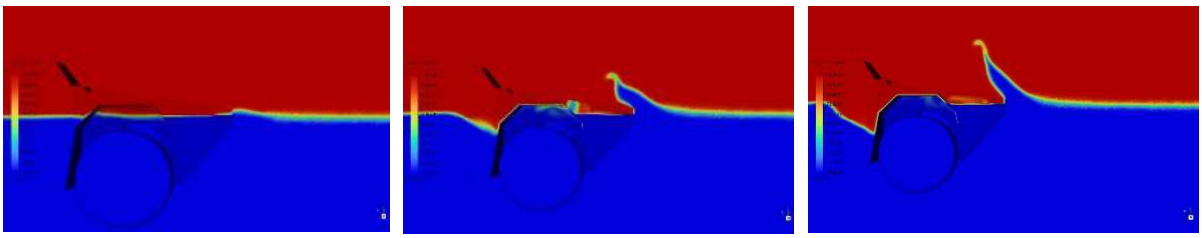
Phân bố áp suất, vận tốc được thể hiện trong hình 7, vùng áp suất thấp tập trung ở phía trên mặt nước và xung quanh thân xe tầng trên phù hợp với áp suất khí quyển. Vùng áp suất cao xuất hiện rõ rệt ở phía dưới đáy xe và phần dưới của bánh xe. Điều này phản ánh áp suất thủy tĩnh tăng dần theo độ sâu của nước. Mũi xe có một vùng áp suất cao ở phía trước mũi; đây là vùng áp suất đình trệ khi dòng nước đập trực tiếp vào vỏ xe, khi tăng vận tốc, áp suất động tại vùng này tăng lên. Hốc bánh xe có thể thấy sự thay đổi áp suất khá đột ngột tại khu vực hốc bánh là nơi dễ xảy ra hiện tượng nhiễu loạn dòng chảy hoặc tạo ra các vùng xoáy, ảnh hưởng đến lực cản tổng thể của xe khi bơi. Áp suất trên bề mặt bánh xe có sự phân tầng rõ rệt theo độ sâu, cho thấy bánh xe đang chịu tác động lớn từ lực đẩy Archimedes và lực cản của nước.



a) Phân bố vận tốc *b) Phân bố áp suất*

Hình 7. Phân bố vận tốc, áp suất (với v=3m/s)

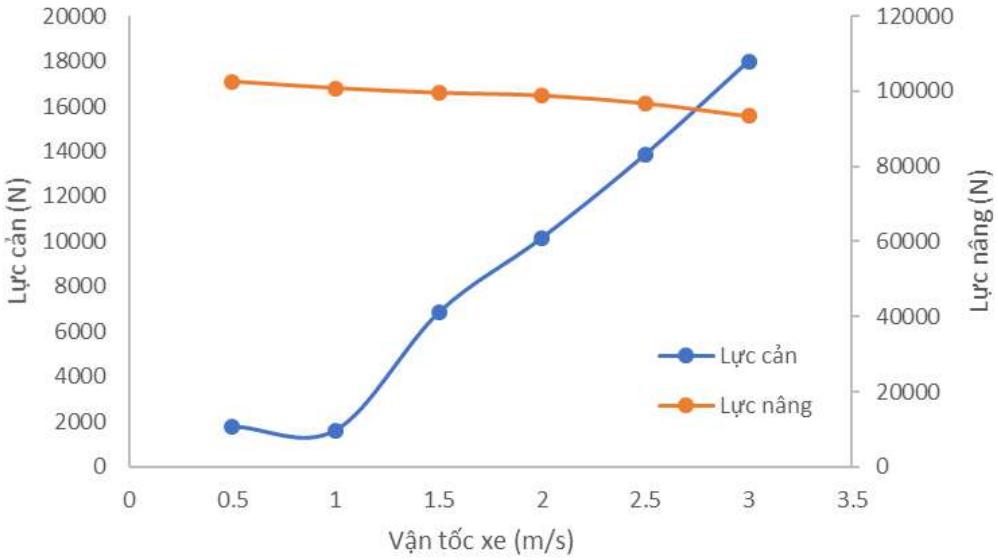
Sự hình thành sóng mũi được thể hiện ở hình 8 cho thấy tốc độ tăng khiến sóng mũi tạo thành có chiều cao lớn hơn.



a) Vận tốc 1m/s *b) Vận tốc 2m/s* *c) Vận tốc 3m/s*

Hình 8. Hình ảnh sóng mũi với các vận tốc khác nhau

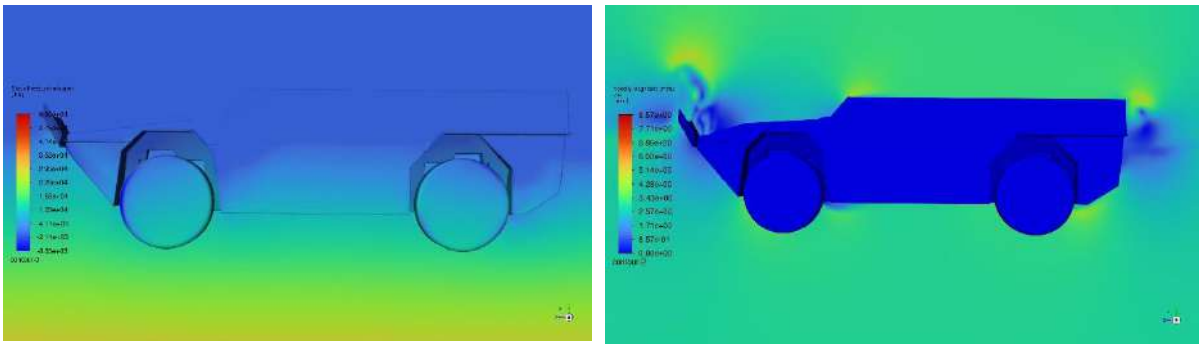
Ảnh hưởng của vận tốc đến lực cản và lực nâng thể hiện ở hình 9, khi vận tốc tăng thì lực cản tăng nhanh còn lực nâng giảm nhẹ (điều này chỉ đúng khi xe di chuyển với tốc độ chậm chưa đạt đến trạng thái lướt nước) [6].



Hình 9. Ảnh hưởng của vận tốc đến lực cản và lực nâng

3.3. Ảnh hưởng của tấm chắn sóng

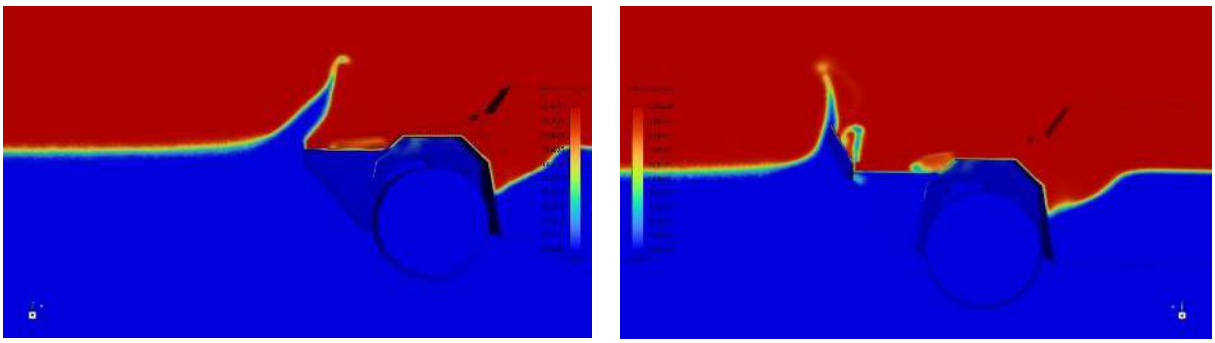
Phân bố áp suất, vận tốc thể hiện trong hình 10, ngoài các vùng áp suất đặt biệt ở hốc bánh xe và bề mặt bánh xe như ở hình 7, vùng áp suất đình trệ tập trung trực tiếp lên bề mặt tấm chắn thay vì đập thẳng vào vỏ mũi xe, tấm chắn sóng đang đóng vai trò "rẽ nước", tạo ra một vùng áp suất ổn định hơn ngay phía sau nó (vùng hốc mũi xe), giúp giảm lực cản trực diện tác động lên thân chính của xe bọc thép.



a) Phân bố áp suất *b) Phân bố vận tốc*

Hình 10. Phân bố vận tốc, áp suất khi có tấm chắn sóng (với v=3m/s)

Sóng mũi được thể hiện ở hình 11 cho thấy tấm chắn sóng có tác dụng giúp giảm lượng nước tràng lên mũi xe. Bảng 4 cũng cho thấy tấm chắn sóng làm tăng lực cản lên tổng thể của xe lên 8,4% do tăng diện tích cản chính diện (thân chính + tấm chắn sóng) nhưng giúp tăng lực nâng của nước tác dụng lên xe 4,8% nhờ tác dụng “đè sóng” tạo hiệu ứng lướt.



a) Sóng mũi khi không có tấm chắn sóng *b) Sóng mũi khi có tấm chắn sóng*

Hình 11. Hình ảnh sóng mũi khi không có tấm chắn sóng và có tấm chắn sóng

Bảng 4. Kết quả lực cản trong các trường hợp

STT	Mô hình	Lực cản Cd (N)	Lực nâng Cl (N)
1	Không có tấm chắn sóng	17971,35	93433,38
2	Có trang bị tấm chắn sóng	19482,574	97879,652

4. Kết luận

Nghiên cứu đã ứng dụng phương pháp mô phỏng số bằng phần mềm ANSYS Fluent để đánh giá tương tác thủy động lực học của xe thiết giáp XTC-02 trong quá trình bơi. Kết quả chỉ ra rằng khi vận tốc bơi tăng, lực cản tác dụng lên xe tăng nhanh trong khi lực nâng giảm dần ở dải tốc độ thấp. Đặc biệt, mô phỏng cho thấy tấm chắn sóng mang lại hiệu quả "rẽ nước" và "đè sóng", giúp giảm lượng nước tràn lên mũi xe, tạo vùng áp suất ổn định ngay phía sau nó và tăng lực nâng của nước tác động vào xe 4,8%. Tuy nhiên, cấu hình tấm chắn sóng này lại làm tăng lực cản tổng thể của xe tăng thêm 8,4%. Hệ số lực và mô hình sóng mũi xác định được từ phương pháp CFD trong nghiên cứu này sẽ là cơ sở dữ liệu quan trọng để tiếp tục nghiên cứu, tối ưu hóa thiết kế hình dáng mũi xe, góc nghiêng và kích thước của tấm chắn sóng cho các dòng xe lội nước.

Tài liệu tham khảo

1. Nguyễn Viết Tân. (2011). *Xe máy vượt sông*, NXB Quân đội nhân dân.
2. Shih, T. H., Liou, W. W., Shabbir, A., Yang, Z., & Zhu, J. (1995). A new k-ϵ eddy viscosity model for high reynolds number turbulent flows. *Computers & Fluids*, 24(3), 227-238.
3. Versteeg, H. K., & Malalasekera, W. (2007). *An Introduction to Computational Fluid Dynamics: The Finite Volume Method* (2nd ed.). Pearson Education.
4. Wilcox, D. C. (1998). *Turbulence Modeling for CFD*. DCW Industries.
5. ANSYS Inc. (2023). *Ansys Fluent Theory Guide, Release 2023 R2*. Canonsburg, PA: ANSYS, Inc.
6. Pan, Y., Zhang, H., & Zhou, Q. (2012). Numerical simulation of flow field around an amphibious vehicle. *Journal of Hydrodynamics*, 24(6), 912-919.
7. Larsson, L., & Raven, H. C. (2010). *Ship Resistance and Flow*. Society of Naval Architects and Mairne Engineers (SNAME).
8. Hirt, C. W., & Nichols, B. D. (1981). *Volume of fluid (VOF) method for the dynamics of free boundaries*. Journal of Computational Physics, 39(1), 201-225.

Study on Hydrodynamic Characteristics of a Combat Armored Personnel Carrier During Swimming Using Numerical Methods.

Abstract: *This paper presents the results of a numerical simulation of the flow field around a combat Armored Personnel Carrier (APC), specifically the XTC-02 model, during amphibious operations. The authors employed Computational Fluid Dynamics (CFD) utilizing ANSYS Fluent software. The Finite Volume Method (FVM) was combined with the Volume of Fluid (VOF) multiphase model to resolve the free surface interactions. The study focuses on analyzing drag components, pressure fields, and bow wave formation across swimming velocities ranging from 0.5 m/s to 3 m/s. A full-scale (1:1) geometric model was simulated within a "virtual towing tank" domain using a polyhedral mesh. Results indicate that as velocity increases, hydrodynamic drag rises significantly while lift force decreases slightly (in the pre-planing regime). Furthermore, the study evaluates the impact of a wave deflector (trim vane). The findings suggest that while the deflector effectively mitigates water swamping over the bow and stabilizes pressure in the bow recess, it results in an increase in total drag by approximately 8,4% and an increase in lift compared to the bare hull configuration. These simulation results provide a critical database for the future optimization of wave deflector designs and vehicle hull geometry.*

Keywords: Armored personnel carrier (APC); numerical simulation; computational fluid dynamics (CFD); ANSYS Fluent; bow wave.

Nghiên cứu tính toán đặc tính thủy động tác dụng lên cơ cấu lái của bơm phụt MJP-330 trên xe thiết giáp bằng phương pháp số

Nghiêm Văn Long[1], Nguyễn Văn Toàn[1], Nguyễn Duy Đạt[1], Lương Trần Anh Duy[1]

[1]*Học viện Kỹ thuật Quân sự*

Email: nghiemvanlong@gmail.com; Tel: 0358231693

Tóm tắt

Hệ thống bơm phụt được sử dụng rộng rãi trên các dòng xe thiết giáp chở quân lội nước nhờ hiệu suất tạo lực đẩy lớn, khả năng bảo vệ an toàn cao cũng như tính linh hoạt trong phân bổ điều hướng. Bài báo trình bày kết quả mô phỏng số (CFD) đặc tính thủy động lực học của cơ cấu gầu lái trên bơm phụt MJP-330 thiết kế để sử dụng cho xe thiết giáp chở quân. Trên cơ sở mô hình rối *SST k-ω* và hệ phương trình chi phối dòng chảy RANS, nghiên cứu đã tiến hành khảo sát và xây dựng các đường đặc tính tải trọng thủy động để đánh giá sự biến thiên phi tuyến của các lực tác dụng, mô-men, đặc tính dòng lưu chất. Kết quả nghiên cứu là cơ sở khoa học cho tính toán thiết kế sức bền, lựa chọn thông số xi-lanh và tối ưu hóa hệ điều khiển lái thủy lực trên xe thiết giáp chở quân lội nước.

Từ khóa: *Bơm phụt; CFD; đặc tính thủy động; xe thiết giáp.*

1. Mở đầu

Khả năng cơ động trên mặt nước là chỉ tiêu chiến thuật then chốt của các dòng xe thiết giáp chở quân lội nước. Hệ thống đẩy bơm phụt (Waterjet) đang dần thay thế chân vịt truyền thống nhờ ưu điểm vượt trội về hiệu suất tạo suy lực và khả năng phân bổ điều hướng [1]. Để bẻ lái và đi lùi, hệ thống sử dụng cơ cấu gầu lái cản trực tiếp luồng tia nước đang phun ra. Quá trình này sinh ra tương tác lưu chất - kết cấu phức tạp, tạo ra các tải trọng thủy động và mô-men phi tuyến tác dụng trực tiếp lên cơ cấu chuyển hướng [2].

Hình 1. Cụm bơm phụt bố trí trên xe thiết giáp

Phần lớn các nghiên cứu CFD hiện nay chủ yếu tập trung tối ưu hóa biên dạng cánh bơm để đạt lực đẩy cực đại ở hướng đi thẳng [3, 4]. Việc thiếu hụt cơ sở dữ liệu về tải trọng và mô-men khi gầu lái mở ở các góc khác nhau khiến kỹ sư thường phải thiết kế cụm xi-lanh thủy lực dư thừa hệ số an toàn. Điều này không những gây lãng phí không gian, tăng trọng lượng mà còn dễ dẫn đến hiện tượng trễ đáp ứng của hệ thống điều khiển [5].

Xuất phát từ thực tiễn trên, bài báo ứng dụng mô phỏng 3D CFD (sử dụng phương trình RANS và mô hình rối *SST k-ω* [6]) khảo sát toàn diện đặc tính thủy động của cơ cấu gầu lái bơm MJP-330 ở dải góc từ 0^o đến 55^o. Mục tiêu cốt lõi nhằm thiết lập đường cong đặc tính tải trọng (lực thủy động và mô-men cản) chuẩn xác với độ tin cậy cao. Dữ liệu này phục vụ trực tiếp làm đầu vào cho bài toán mô phỏng động lực học mạch điều khiển lái, giúp tối ưu hóa cơ cấu chấp hành thủy lực trên các dòng xe thiết giáp chở quân thế hệ mới [5].

2. Phương pháp số

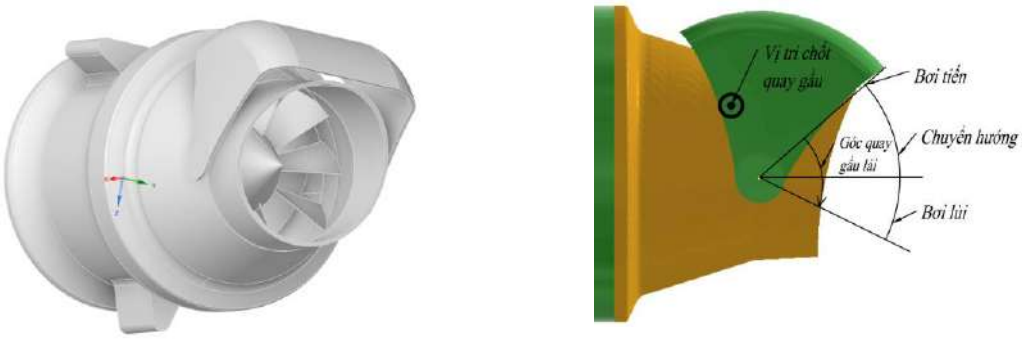
2.1. Mô hình vật lý và giả thiết

Đối tượng nghiên cứu là dòng chảy qua bơm phụt MJP-330 lắp trên xe thiết giáp với bộ thông số kỹ thuật như sau:

Bảng 1. Thông số bơm phụt MJP-330

Model bơm phụt	**MJP 330**
Công suất	90 kW
Tốc độ lớn nhất	1800 vòng/phút
Đường kính cánh quạt	300 mm
Số cánh làm việc	5 cánh
Trọng lượng	43 kg

Miền tính toán chất lưu được trích xuất từ mô hình thiết kế 3D (hình 2), bao gồm các thành phần chính: Vỏ bơm, bánh công tác và gầu lái.



Hình 2. Mô hình 3D đưa vào tính toán mô phỏng số

Để đơn giản hóa quá trình tính toán nhưng vẫn đảm bảo độ chính xác của các đặc tính vật lý cơ bản, các giả thiết sau được áp dụng:

- Hệ thống bơm phụt hoạt động hoàn toàn ngập dưới nước. Do đó, bài toán được đơn giản hóa thành mô hình dòng chảy lưu chất đơn pha.
- Chất lỏng làm việc là nước, được coi là chất lỏng Newton và không nén được.
- Dòng chảy được giả thiết là ổn định trong các tính toán đặc tính lực trung bình.
- Nhiệt độ của môi trường nước được giả thiết là không thay đổi khi bơm làm việc (do sự biến thiên nhiệt độ của nước khi qua bơm là không đáng kể).

2.2. Hệ phương trình chi phối

Dựa trên các giả thiết trên, dòng chảy được mô tả bởi hệ phương trình Navier-Stokes trung bình hóa theo Reynolds (RANS - Reynolds Averaged Navier-Stokes). Hệ phương trình bao gồm phương trình liên tục (phương trình bảo toàn khối lượng) và phương trình bảo toàn động lượng, được viết dưới dạng tenxơ như sau:

Phương trình liên tục:

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i}(\rho u_i) = 0 \qquad (1)$$

Đối với dòng không nén ($\rho = const$), phương trình trở thành:

$$\frac{\partial u_i}{\partial x_i} = 0 \tag{2}$$

Phương trình động lượng:

$$\frac{\partial(\rho u_i)}{\partial t} + \frac{\partial}{\partial x_j}(\rho u_i u_j) = -\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j}\left[\mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) - \rho\overline{u_i' u_j'}\right] \tag{3}$$

Trong đó:

+ ρ : Khối lượng riêng của nước (kg/m^3).

+ u_i , u_j : Các thành phần vận tốc trung bình (m/s).

+ p : Áp suất tĩnh (Pa).

+ μ : Độ nhớt động lực học ($Pa.s$).

+ $-\rho\overline{u_i' u_j'}$: Tenxơ ứng suất Reynolds, đại diện cho ảnh hưởng của các dao động rối lên dòng chảy trung bình. Thành phần này làm cho hệ phương trình không khép kín và cần sử dụng mô hình rối để giải quyết.

2.3. Mô hình rối SST k-ω

Đặc điểm dòng chảy qua cơ cấu lái của bơm phụt bao gồm vận tốc lớn, gradient áp suất cao và hiện tượng tách dòng (flow separation) mạnh tại các góc bẻ lái lớn. Để mô phỏng chính xác các hiện tượng này, nghiên cứu sử dụng mô hình rối *SST k-ω* (Shear Stress Transport) do Menter đề xuất [6].

Hai phương trình vận chuyển cho động năng rối k và tốc độ tiêu tán riêng ω được biểu diễn như sau:

Phương trình động năng rối (k):

$$\frac{\partial(\rho k)}{\partial t} + \frac{\partial(\rho k u_i)}{\partial x_i} = \frac{\partial}{\partial x_j}\left(\Gamma_k \frac{\partial k}{\partial x_j}\right) + \tilde{G}_k - Y_k + S_k \tag{4}$$

Phương trình tốc độ tiêu tán riêng (ω):

$$\frac{\partial(\rho\omega)}{\partial t} + \frac{\partial(\rho\omega u_i)}{\partial x_i} = \frac{\partial}{\partial x_j}\left(\Gamma_\omega \frac{\partial \omega}{\partial x_j}\right) + G_\omega - Y_\omega + D_\omega + S_\omega \tag{5}$$

Trong đó, $\tilde{G}_k$ đại diện cho sự sinh động năng rối do gradient vận tốc trung bình; G_ω là sự sinh ω; Γ_k và Γ_ω là độ khuếch tán hiệu dụng của k và ω ; Y_k và Y_ω là sự tiêu tán của k và ω do độ nhớt. Ưu điểm nổi bật của mô hình SST là việc giới hạn ứng suất cắt rối thông qua việc điều chỉnh độ nhớt rối μ_t, giúp dự báo chính xác điểm tách dòng trên bề mặt gầu lái [1], yếu tố quyết định trực tiếp đến giá trị lực cản và mô men tác dụng.

2.4. Phương pháp giải

Bài toán được giải quyết bằng phương pháp Thể tích hữu hạn (FVM - Finite Volume Method) trên phần mềm ANSYS Fluent 2023. Miền tính toán được rời rạc hóa thành các ô lưới đa diện. Các thiết lập giải thuật cụ thể như sau:

– Liên kết áp suất - vận tốc: Sử dụng thuật toán Coupled để tăng tốc độ hội tụ và độ ổn định cho dòng chảy phức tạp bên trong bơm.

– Rời rạc hóa không gian:

+ Gradient: Least Squares Cell Based.

+ Áp suất: Second Order.

+ Động lượng, Động năng rối và Tốc độ tiêu tán riêng: Sử dụng sơ đồ Second Order Upwind để đảm bảo độ chính xác bậc hai.

– Điều kiện biên:

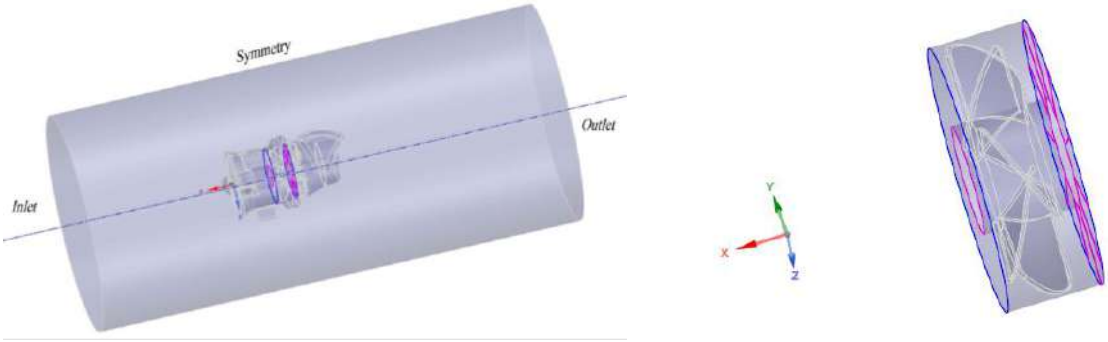
+ Lối vào (Inlet): Thiết lập điều kiện Pressure Inlet với áp suất tĩnh p = 0 (bơm hoạt động ở vùng nước nông)

+ Lối ra (Outlet): Thiết lập điều kiện Pressure Outlet với áp suất tĩnh p = 0.

+ Biên vách (Wall): Áp dụng điều kiện No-slip wall (không trượt) cho các bề mặt rắn (vỏ bơm, gầu lái). Bề mặt được coi là nhẵn thủy lực.

+ Mặt phân cách (Interface): Sử dụng mô hình Moving Reference Frame cho vùng chất lỏng bao quanh bánh công tác để mô phỏng chuyển động quay với tốc độ n = 1800 vòng/phút[4].

– Lưới tính toán: Miền tính toán được chia lưới đa diện (polyhedral).



a. Vùng lưu chất tĩnh *b. Vùng lưu chất quay*

Hình 3. Miền mô phỏng trong môi trường Ansys Fluent 2023

2.5. Xác định đặc tính thủy động lực học

Mục tiêu của nghiên cứu là xác định các thành phần lực và mô men tác dụng lên gầu lái. Các đại lượng này được tính toán bằng cách tích phân phân bố áp suất và ứng suất cắt nhớt trên toàn bộ diện tích bề mặt ướt (A) của gầu lái.

2.5.1. Lực thủy động ($\vec{F}$):

Véc tơ lực tổng hợp tác dụng lên gầu lái được xác định theo công thức:

$$\vec{F} = \int_A (p\vec{n} + \vec{\tau}_{wall})dA \quad (6)$$

Trong đó:

+ $\vec{n}$: Véc tơ pháp tuyến đơn vị hướng ra ngoài bề mặt.

+ $\vec{\tau}_{wall}$: Véc tơ ứng suất cắt tại tường.

Từ véc tơ tổng hợp $\vec{F}$ tách ra các thành phần quan trọng trong hệ tọa độ gắn với thân xe:

+ Lực cản dọc trục (F_x):Là thành phần lực nằm dọc theo trục vòi phun (phương X). Đây là áp lực thủy động sinh ra khi bề mặt gầu cản trở sự chuyển động của tia nước. Thành phần lực này có xu hướng đẩy lùi gầu lái về phía sau.

+ Lực ngang (F_z): Là thành phần lực theo phương ngang (phương Z), xuất hiện do sự biến đổi động lượng khi tia nước bị ép chuyển hướng qua bề mặt cong của gầu. Đối với bản thân cơ cấu, F_z là thành phần tải trọng gây ra ứng suất uốn lớn lên bề mặt gầu. Ngoài ra, phản lực của F_z truyền qua hệ thống ngàm giữ cụm bơm vào vỏ xe chính là nguồn gốc tạo ra lực bẻ lái cho phương tiện.

+ Lực thẳng đứng (F_y): Thành phần lực thẳng đứng (hướng Y) được sinh ra do hiện tượng lưu chất tràn qua các mép hở trên/dưới của gầu lái. Lực này chủ yếu gây ra tải trọng tác dụng lên các vòng bi hoặc bạc lót của tay đòn.

2.5.2. Mô-men cản lái (M_y):

Mô-men cản lái My là mô-men xoắn tác dụng lên vị trí chốt bản lề được xác định bởi tích có hướng của véc tơ vị trí và véc tơ lực:

$$M = \int_A \vec{r} \times (p\vec{n} + \vec{\tau}_{wall}) dA \quad (7)$$

M_y là thông số tới hạn quan trọng nhất của nghiên cứu này. Sinh ra do tâm áp suất (điểm đặt của lực tổng hợp) nằm cách xa tâm quay của chốt giữ gầu. Giá trị lớn nhất của M_y sẽ quyết định trực tiếp đến việc tính chọn xi-lanh thủy lực làm nhiệm vụ dẫn động cơ cấu lái.

3. Kết quả và thảo luận

3.1. Chia lưới

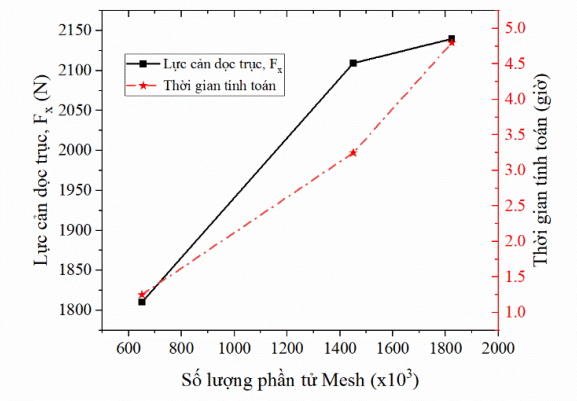
3.1.1 Khảo sát tính độc lập của lưới

Độ chính xác của mô phỏng CFD phụ thuộc rất lớn vào chất lượng và mật độ phần tử lưới. Để đảm bảo sai số rời rạc hóa được giảm thiểu và kết quả tính toán không bị phụ thuộc vào số lượng lưới, nghiên cứu tiến hành khảo sát sự hội tụ của lưới. Trường hợp góc quay gầu lái lớn nhất $\alpha = 55^o$ được chọn để khảo sát, do đây là trạng thái dòng chảy bị cản trở mạnh nhất, sinh ra các vùng tách dòng và nhiễu động phức tạp đòi hỏi độ phân giải lưới cao. Ba phương án chia lưới từ thô đến mịn được xây dựng để đánh giá sự biến thiên của lực cản dọc trục F_x.

Bảng 2. Kết quả mô phỏng ứng với các phương án chia lưới

Phương án	Số lượng phần tử Mesh	Lực F_x (N)	Sai số tương đối (%)	Thời gian tính toán (h)
Lưới 1 (Thô)	650.132	1810,07	-	1,5
Lưới 2 (Trung bình)	1.450.564	2109,0	16,52 % (so với lưới 1)	3,25
Lưới 3 (Mịn)	1.823.455	2139,7	1,45 % (so với lưới 2)	4,8

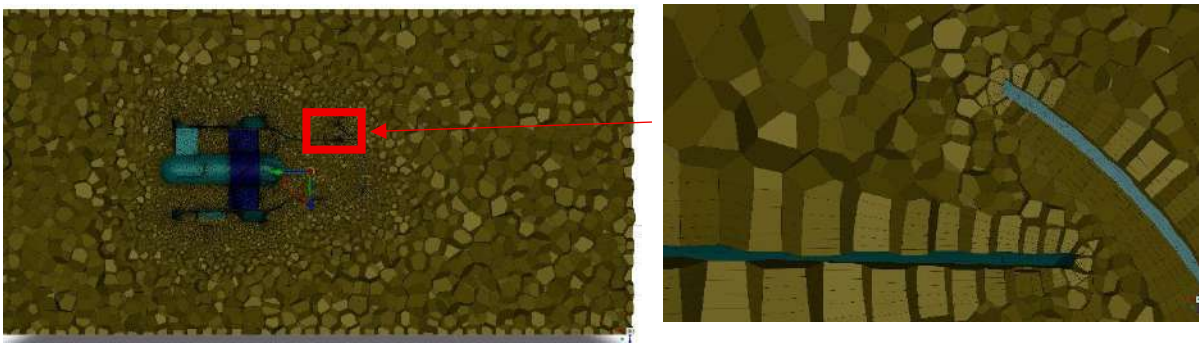
Dựa vào kết quả từ bảng 2, có thể thấy rõ sự ảnh hưởng của mật độ lưới đến độ chính xác của bài toán giải tích: Khi tăng số lượng phần tử từ 650.132 (lưới 1) lên 1.450.564 (lưới 2): F_x thay đổi rõ rệt từ $1810,07\ N$ lên $2109,00\ N$, tương đương mức sai số 16,52%. Điều này chứng tỏ lưới thô chưa đủ mật độ ở vùng lớp biên (Boundary layer) và vùng khe hở tĩnh-động để xác định chính xác hiện tượng giảm áp suất và gradient vận tốc. Và khi tiếp tục làm mịn lưới từ 1.450.564 (lưới 2) lên 1.823.455 (lưới 3): Kết quả F_x gần như tiến tới trạng thái ổn định (đạt $2139,7\ N$). Sự chênh lệch giữa lưới trung bình và lưới mịn chỉ còn là 1,45% (nhỏ hơn mức sai số cho phép thông thường là $2\% \div 5\%$ trong mô phỏng thủy động lực học chân vịt/bơm phụt[7]). Quỹ đạo của giá trị F_x đã đi vào vùng hội tụ (tiệm cận giá trị ổn định).



Hình 4. Sự ảnh hưởng của chất lượng lưới đến kết quả và thời gian tính toán

3.1.2. Kết luận lựa chọn phương án chia lưới

Để tối ưu hóa sự cân bằng giữa độ chính xác và tài nguyên máy tính, phương án lưới 2 được lựa chọn làm thiết lập chuẩn để tiến hành mô phỏng cho tất cả các góc quay gầu lái còn lại trong nghiên cứu này.



a. Lưới tổng quan *b. Lớp biên sát vách miền làm việc gầu lái*

Hình 5. Phương án chia lưới được lựa chọn để mô phỏng

3.2. Đặc tính lực và mô men lái

Kết quả mô phỏng các thành phần lực thủy động và Mô-men cản lái khi gầu lái quay các góc từ 0^o đến 55^o được trình bày trên bảng 3:

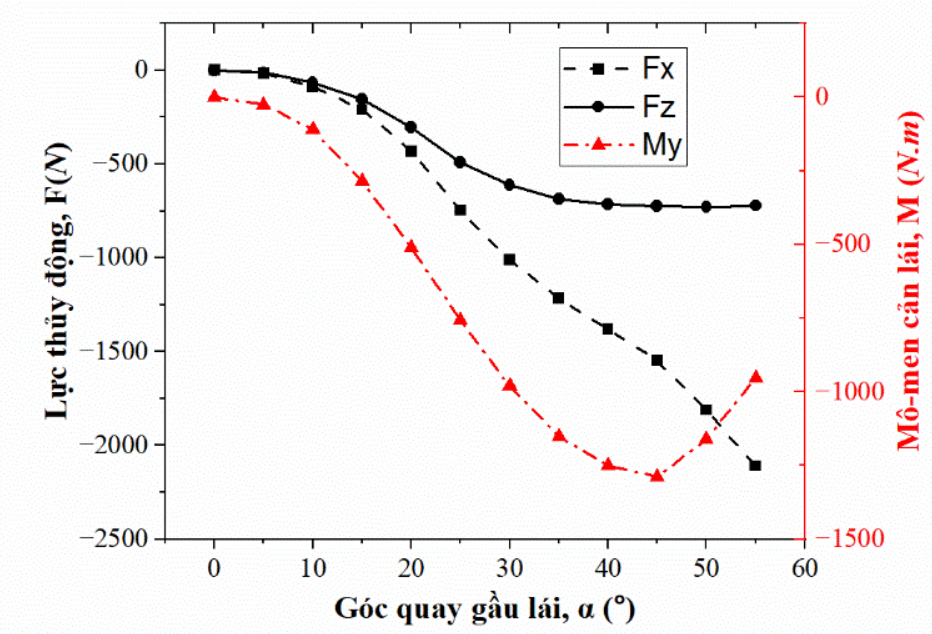
Bảng 3. Kết quả mô phỏng lực thủy động và mô men lái

Góc quay α (°)	Lực F_x (N)	Lực F_y (N)	Lực F_z (N)	Mô-men M_y (N.m)
0°	0	0	0	0
5°	-15,2	25,5	-12,5	-25,4
10°	-85,4	52,3	-65,3	-110,5
15°	-210,6	75,1	-155,8	-285,2
20°	-430,2	88,4	-305,4	-510,8
25°	-746,55	92,65	-489,29	-757,45
30°	-1010,3	70,2	-610,5	-980,6
35°	-1215,8	35,5	-685,2	-1150,3
40°	-1380,4	5,4	-715,6	-1250,7
45°	-1548	-15,08	-725,16	-1287,98
50°	-1810,5	-28,5	-728,3	-1160,4
55°	-2109	-35,2	-721,9	-951,67

3.2.1. Nhận xét về thành phần lực thẳng đứng F_y

Thành phần lực F_y có giá trị rất nhỏ, gần như không đáng kể so với các tải trọng chính nằm trong mặt phẳng ngang. Sự đảo chiều của F_y phản ánh rõ nét hiện tượng tràn nước lên trên bề mặt cong. Ở dải góc bẻ lái hẹp (dưới 30°), dòng chảy thứ cấp chủ yếu trượt dọc và thoát qua mép dưới gầu, sinh ra phản lực thẳng đứng (F_y mang dấu dương). Ngược lại, ở các góc quay gầu lớn (trên 40°), dòng chảy bị nghẽn và tràn mạnh qua khe hở mép trên gầu, tạo ra phản lực dìm xuống (F_y mang dấu âm). Tuy nhiên, do biên độ lực quá nhỏ (dưới 10 kg), thành phần này có thể được bỏ qua trong bài toán thiết kế xy lanh thủy lực.

3.2.2. Nhận xét về lực thủy động và Mô–men cản lái



Hình 6. Đồ thị Lực thủy động và Mô-men cản lái

Dựa vào đồ thị biểu diễn mối quan hệ giữa góc quay gầu lái α và các đại lượng thủy động lực học (Lực F_x, F_z và Mô-men M_y), ta có thể rút ra những đánh giá về bản chất phân bố tải trọng trên cơ cấu bẻ lái của bơm phụt như sau:

a. Đặc tính bão hòa của lực ngang F_z và hiệu ứng phanh của lực cản dọc trục F_x

Giai đoạn gầu lái bắt đầu chắn dòng ($0^o - 20^o$): Các thành phần lực F_x và F_z có giá trị khá nhỏ và tăng với biên độ chậm. Bởi ở góc mở hẹp, gầu lái mới chỉ cản phá lớp biên của dòng chảy từ vòi phun, trong khi lõi tia nước mang động lượng lớn nhất vẫn thoát thẳng ra ngoài không bị cản trở.

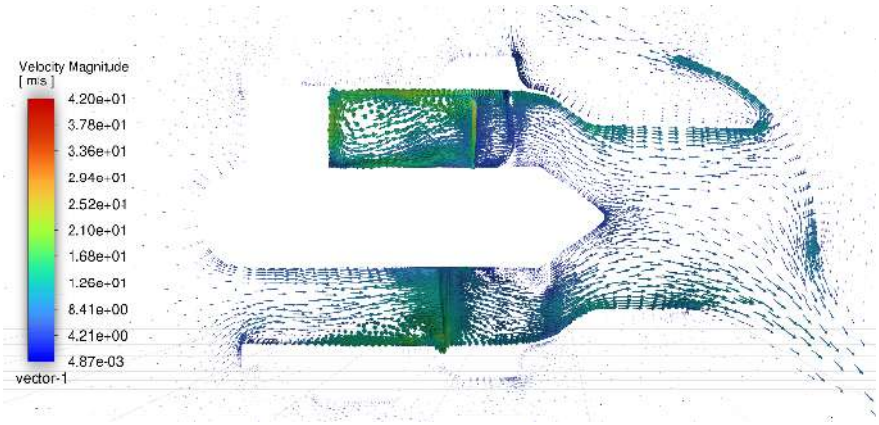
Giai đoạn làm việc hiệu quả và bão hòa ($20^o - 55^o$): Khi góc quay tăng, gầu lái cản toàn bộ tia nước làm lực ngang F_z (đường nét liền, điểm tròn) tăng mạnh và đạt trạng thái bão hòa ở mức khoảng 725 N tại dải góc $45^o - 55^o$. Vượt qua ngưỡng 45^o, F_z không tăng thêm do sự xuất hiện của vùng tách dòng mạnh khiến dòng chảy bị tán xạ[2]. Ngược lại, lực F_x (đường nét đứt, điểm vuông) tiếp tục tăng đột biến, đạt cực đại $\approx 2109N$ tại $\alpha = 55^o$. Tại góc chuyển hướng cực đại này, gầu lái không chỉ làm nhiệm vụ bẻ dòng mà còn hoạt động như một cơ cấu phanh, triệt tiêu mạnh mẽ động năng tịnh tiến của dòng chảy[2].

b. Sự biến thiên phi tuyến của Mô-men cản lái M_y và hiện tượng dịch chuyển tâm áp suất

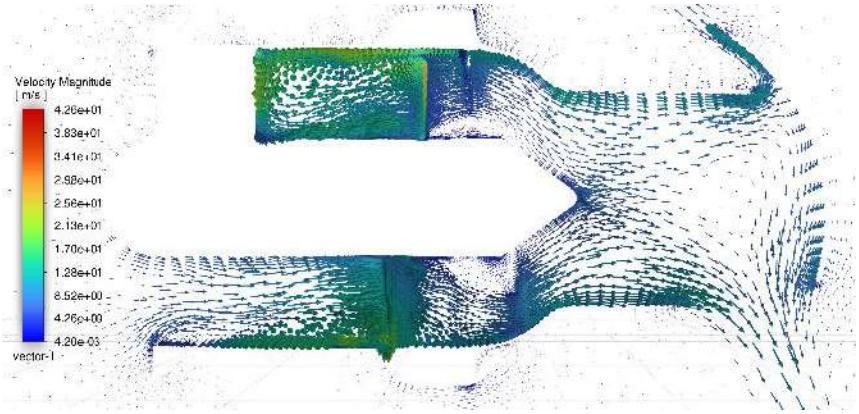
Đường đặc tính mô-men M_y (đường gạch chấm, điểm tam giác) phản ánh một quy luật phi tuyến. M_y tăng dần và đạt đỉnh tải trọng cực đại ở mức $\approx 1288N.m$ tại góc $\alpha = 45^0$.Tuy nhiên, khi tiếp tục tăng góc quay gầu lái (từ 45° đến 55°), mô-men lại sụt giảm mạnh xuống chỉ còn $\approx 951N.m$, mặc dù tổng lực tác dụng lên gầu tiếp tục tăng.

Nguyên nhân: Sự sụt giảm này minh chứng cho sự dịch chuyển của Tâm áp suất. Ở các góc mở vừa phải (nhỏ hơn 45°), dòng nước quét vào phần rìa mép gầu, tạo ra cánh tay đòn lớn nhất. Khi gầu đóng khép góc lớn (55°), luồng nước bị ép tác động trực diện vào vùng lõm trung tâm của thân gầu. Trọng tâm áp lực lúc này dời sát về điểm lắp chốt bản lề, làm giảm đáng kể chiều dài cánh tay đòn hiệu dụng, từ đó kéo theo sự sụt giảm của M_y.

3.3. Trường vector vận tốc tia nước



a. Gầu lái quay 450

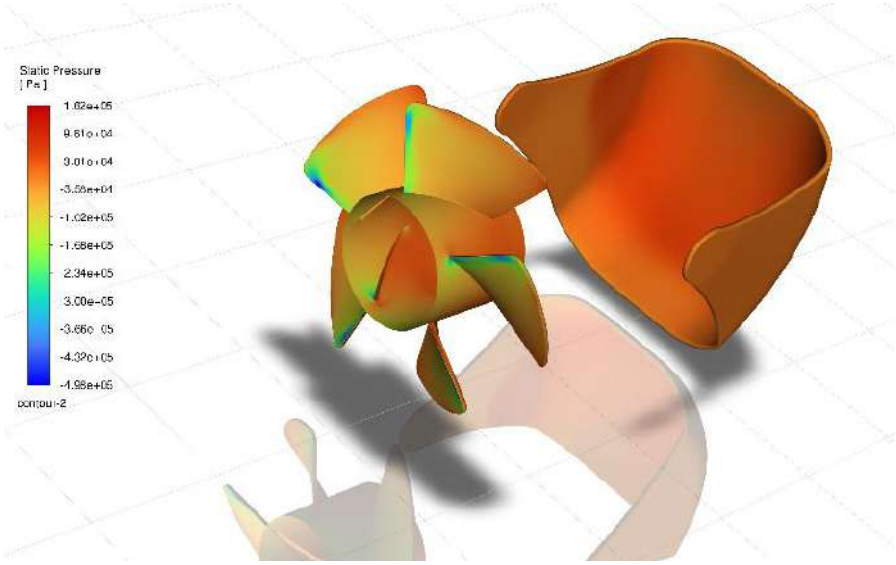


b. Gầu lái quay 550

Hình 7. So sánh trường vector vận tốc tại mặt cắt dọc (Plane Y=0) của gầu lái

"Phân tích trường dòng chảy bằng mô hình CFD đã trực quan hóa nguyên nhân giảm Mô-men cản lái M_y tại các góc mở lớn. Tại góc $\alpha = 45^0$ (Hình 7a), dòng chảy có xu hướng bám theo biên dạng khí động học, dồn tâm áp suất về phía rìa ngoài gầu, sinh ra cánh tay đòn cực đại tạo ra $(M_y)_{\max}$ .Ngược lại, tại ngưỡng $\alpha = 55^0$ (hình 7b), hiện tượng nghẽn dòng xảy ra khiến toàn bộ động năng tia nước tác động trực diện vào khu vực lõm trung tâm của gầu. Sự dịch chuyển tâm áp suất về gần điểm đặt lực, kết hợp với hiện tượng tách dòng mạnh tại mép thoát, đã giải thích chính xác sự bão hòa lực ngang (F_z) và sự sụt giảm Mô-men cản lái tác dụng lên gầu lái.

3.4. Trường áp suất tác dụng lên cơ cấu công tác



Hình 8. Phân bố áp suất trên bề mặt cánh bơm, gầu lái trong trường hợp góc quay gầu lái 450

Kết quả mô phỏng CFD (Hình 8) làm rõ hai đặc tính thủy động lực học cốt lõi của hệ thống:

3.4.1. Tải trọng trên gầu lái

Khi tia nước tác dụng lên mặt lõm của gầu, động năng bị triệt tiêu đột ngột, chuyển hóa thành vùng áp suất tĩnh cực đại đạt 1.62×10^5 Pa (khoảng 1.6 bar). Đây chính là nguồn gốc sinh ra lực cản dọc trục F_x, lực ngang F_z và mô-men cản lái M_y tác dụng lên vị trí lắp chốt giữ gầu.

3.4.2. Nguy cơ xâm thực tại cánh công tác

Quá trình truyền năng lượng tạo ra gradient áp suất rõ rệt giữa mặt đẩy và mặt hút của cánh bơm. Kết quả mô phỏng cho thấy, tại mép đón dòng xuất hiện vùng áp suất âm lên tới -4.98×10^5 Pa. Sự sụt giảm cục bộ vượt quá ngưỡng áp suất hóa hơi này dự báo rủi ro xâm thực, tương đồng với các công bố trước đây về đặc tính thủy động của cánh quạt bơm phụt [8,9] và đặt ra yêu cầu tối ưu hóa biên dạng cánh bơm ở các nghiên cứu tiếp theo.

4. Kết luận

Nghiên cứu đã ứng dụng thành công phương pháp CFD với mô hình rối SST *SST k-ω* để phân tích đặc tính thủy động lực học của cơ cấu gầu lái trên bơm phụt MJP-330 lắp trên xe thiết giáp chở quân. Nghiên cứu đã thiết lập được bộ dữ liệu đặc tuyến tải trọng phi tuyến của cơ cấu: lực ngang có xu hướng bão hòa ở mức 725 N tại dải từ 45°-55°, trong khi lực cản dọc trục gia tăng đột biến, đạt cực đại 2109 N tại góc 55°. Đặc biệt, nghiên cứu chỉ ra rằng đỉnh mô-men cản lái tác dụng lên vị trí lắp chốt bản lề đạt 1288 N.m, tại góc 45° thay vì góc 55° do hiện tượng dịch chuyển tâm áp suất và tách dòng. Phát hiện này đã bác bỏ giả định thiết kế tải trọng tuyến tính thông thường.

Dù cung cấp bộ dữ liệu đặc tuyến có độ tin cậy cao, mô hình hiện tại mới giới hạn ở bài toán dòng chảy ổn định và lưu chất đơn pha. Do đó, các kết quả chưa phản ánh trọn vẹn sự biến thiên thủy động lực học tức thời khi gầu lái hoạt động liên tục.

Tiếp theo, bộ dữ liệu tải trọng tới hạn này sẽ được sử dụng trực tiếp làm thông số đầu vào cho mô hình mô phỏng hệ thống thủy lực lái. Hướng đi này cho phép tính toán động lực học cơ cấu cơ khí, mô phỏng thời gian đáp ứng để lựa chọn tối ưu thông số cụm xy lanh, qua đó hoàn thiện mạch điều khiển thủy lực trên các thế hệ xe thiết giáp chở quân hiện đại.

Tài liệu tham khảo

1. Jia, B., Wang, Z., & Li, Z. (2021). "Hydrodynamic performance and flow field analysis of an amphibious vehicle with waterjet propulsion." *Applied Ocean Research*, 111, 102613.
2. Liu, Y., & Wang, L. (2020). "Numerical simulation of hydrodynamic performance for waterjet propulsion steering and reversing mechanism." *Ocean Engineering*, 218, 108226.
3. Eslamdoost, A., Larsson, L., & Bensow, R. (2015). "The Hydrodynamics of Waterjet/Hull Interaction." *Journal of Ship Research*, 59(03), 133-149.
4. Park, W. G., Jang, J. H., Chun, H. H., & Kim, M. C. (2005). "Numerical flow and performance analysis of waterjet propulsion system." *Ocean Engineering*, 32(14-15), 1740-1761.
5. Zhang, X., et al. (2024). "Modeling and Analysis of Actuators in Multi-Pump Waterjet Propulsion Systems." *Journal of Marine Science and Engineering*, 13(1), 154.
6. Menter, F. R. (1994). "Two-equation eddy-viscosity turbulence models for engineering applications." *AIAA Journal*, 32(8), 1598-1605.
7. Brizzolara, S., & Villa, D. (2010). "Multiphase CFD analysis of a waterjet pump." *Ocean Engineering*, 37(1), 1-13.
8. Huang, C., et al. (2022). "CFD-FEM Analysis of Flow-Induced Vibrations in Waterjet Propulsion Unit." *Journal of Marine Science and Engineering*, 10(8), 1032.
9. Wang, C., Wang, C., He, C., & Wang, Y. (2019). "Numerical investigation of cavitation performance of a waterjet pump." *Ocean Engineering*, 188, 106290.

Numerical investigation into the hydrodynamic characteristics acting on the steering mechanism of the MJP-330 waterjet on armored vehicles

Abstract: *Waterjet propulsion systems are widely used in amphibious armored personnel carriers due to their high thrust efficiency, high operational safety, and flexibility in directional control. This paper presents the computational fluid dynamics (CFD) simulation results regarding the hydrodynamic characteristics of the steering bucket mechanism on the MJP-330 waterjet designed for armored personnel carriers. Based on the SST k-ω turbulence model and the RANS governing equations, the research investigates and constructs hydrodynamic load characteristic curves to evaluate the nonlinear variation of the force system, torque, and fluid flow characteristics. The research results serve as a scientific basis for structural strength calculations, hydraulic cylinder parameter selection, and optimization of the hydraulic steering control system on amphibious armored personnel carriers.*

Keywords: Waterjet; CFD; Hydrodynamic characteristics; Armored vehicle.

Nghiên cứu động lực học cổng trục công xôn phục vụ lao lắp dầm cầu trong trường hợp hạ hàng

Ngô Minh Tiến[1*], Nguyễn Ngọc Trung[2], Nguyễn Viết Tân[1]
[1]Học viện Kỹ thuật quân sự
[2]Trường Đại học Giao thông vận tải
[*]Email: ngominhtien217@gmail.com; Tel: 0375467826

Tóm tắt

Trong một chu kỳ làm việc của cổng trục công xôn lao lắp dầm, quá trình vận hành gồm ba giai đoạn: nâng dầm, di chuyển dầm và hạ dầm. Tác giả đã công bố nghiên cứu động lực học cho giai đoạn nâng dầm; bài báo này tiếp tục nghiên cứu các giai đoạn còn lại nhằm hoàn thiện đánh giá toàn bộ chu kỳ, trong đó hạ dầm là giai đoạn tiềm ẩn nguy cơ mất an toàn do tải trọng động và dao động của hệ. Mục tiêu nghiên cứu không chỉ khảo sát quy luật biến thiên các đại lượng động lực học và ảnh hưởng của các thông số đặc trưng, mà còn so sánh định lượng với trường hợp nâng dầm để lựa chọn thông số đảm bảo an toàn theo điều kiện bất lợi.

Để giải quyết bài toán, bài báo xây dựng mô hình động lực học 6 bậc tự do, phản ánh tương tác giữa bộ máy nâng, cáp, dầm và kết cấu cổng trục. Trên cơ sở đó, phương trình chuyển động được thiết lập theo phương pháp Lagrange, phục vụ phân tích đáp ứng động lực học trong quá trình di chuyển và hạ dầm. Kết quả nghiên cứu là cơ sở khoa học cho tính toán thiết kế, chế tạo và khai thác an toàn cổng trục công xôn trong thi công cầu.

Từ khoá: *Cổng trục công xôn; động lực học; mô hình 6 bậc tự do; trường hợp nâng dầm; trường hợp hạ dầm; di chuyển dầm; tải trọng động; dao động hệ nâng; phương trình Lagrange.*

1. Đặt vấn đề

Trong bối cảnh hạ tầng giao thông Việt Nam phát triển mạnh, nhu cầu xây dựng đường và cầu vượt tại các đô thị lớn ngày càng tăng; theo quy hoạch đến năm 2035, cầu cạn dự kiến chiếm trên 60% tuyến đường sắt tốc độ cao. Trong thi công lao lắp dầm cầu bê tông cốt thép, cổng trục công xôn được sử dụng phổ biến nhờ thi công linh hoạt, không cần hệ đỡ dưới nền, phù hợp địa hình phức tạp, ít ảnh hưởng giao thông thủy, tiến độ nhanh và có khả năng tái sử dụng.

Tuy nhiên, các nghiên cứu và tài liệu thiết kế hiện nay về thiết bị này vẫn chủ yếu dựa trên giả thiết tải trọng tĩnh, tập trung vào tính toán kết cấu thép và các cơ cấu công tác riêng lẻ; các tiêu chuẩn như TCVN 2737–2023 và TCVN 12162–2017 cũng chủ yếu xét các trạng thái làm việc tĩnh [1]. Về động lực học, một số công trình đã nghiên cứu mô hình dao động của hệ cổng trục/cần trục, đánh giá ảnh hưởng vận tốc–gia tốc tải và giảm chấn ví dụ Zrnić và cộng sự [2] cũng như các mô hình phi tuyến cho hệ cáp trong nâng–hạ [3]. Nhiều nghiên cứu liên quan đến động lực học nâng–hạ của cổng trục [4-9] đã được công bố, song đối với cổng trục công xôn, các nghiên cứu theo hướng động lực học, đặc biệt ở giai đoạn hạ dầm, vẫn còn hạn chế.

Xuất phát từ yêu cầu đó, trên cơ sở mô hình hóa và phân tích động lực học, bài báo tập trung xác định quy luật biến thiên theo thời gian của các đại lượng đặc trưng như chuyển vị, vận tốc, gia tốc của hệ; lực căng trong các nhánh cáp; cũng như phản lực tác dụng lên các chân cổng trục và kết cấu dầm chính. Đồng thời, nghiên cứu tiến hành khảo sát ảnh hưởng của các tham số đặc trưng đến biên độ dao động và mức tăng tải trọng động, qua đó xác định các yếu tố chi phối đến ổn định và an toàn làm việc của thiết bị.

Bên cạnh đó, nhằm nâng cao giá trị ứng dụng, các kết quả được so sánh định lượng với trường hợp nâng dầm theo các chỉ tiêu: chuyển vị cực đại của cổng trục, lực căng cáp cực đại và

lực tác dụng lớn nhất lên chân cổng trục. Việc so sánh nâng–hạ cho phép nhận diện trạng thái bất lợi và lựa chọn giá trị lớn nhất theo điều kiện nguy hiểm làm cơ sở đề xuất thông số và chế độ làm việc phù hợp. Từ đó, nghiên cứu cung cấp luận cứ khoa học nhằm xác định cách ứng xử trong quá trình thiết kế, chế tạo và khai thác cổng trục công xôn theo hướng đảm bảo độ bền, an toàn đồng thời tối ưu kinh tế. Các nội dung chi tiết sẽ được trình bày trong các phần tiếp theo.

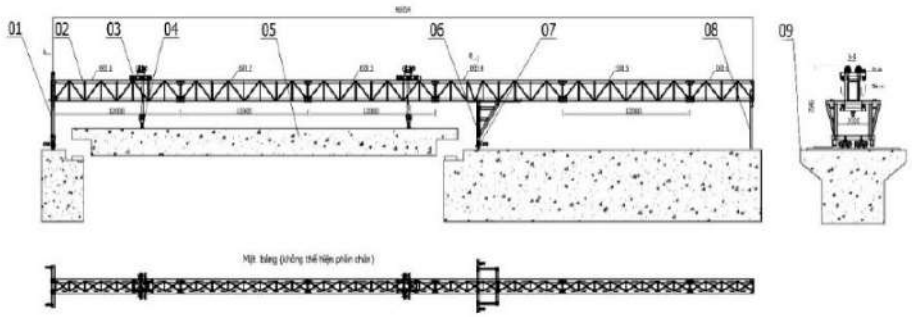
2. Nghiên cứu động lực học cổng trục công son trong quá trình nâng dầm khi cáp căng

2.1. Giới thiệu về cổng trục công son phục vụ lao lắp dầm cầu

Cổng trục công son được sử dụng phổ biến trong thi công cầu có kết cấu dầm super-T, dầm I, dầm hộp,... đúc sẵn. Chức năng chính bao gồm: Nâng và di chuyển các dầm cầu đúc sẵn (66 – 80 tấn) từ mặt đất hoặc xe vận chuyển lên đúng vị trí thiết kế trên trụ cầu. Lao lắp dầm nhanh chóng và chính xác, đặc biệt ở các vị trí khó tiếp cận bằng cần cẩu thông thường (trên sông, ruộng, địa hình yếu,...). Giảm phụ thuộc vào mặt bằng thi công phía dưới (không cần kê giáo). Thi công liên tục nhịp này sang nhịp khác mà không phải tháo dỡ, di dời thiết bị nhiều lần.

Hình 1. Công trục công xôn dạng giàn phục vụ lao lắp dầm cầu ngoài công trường thực tế

Cổng trục công xôn là thiết bị chuyên dùng phục vụ lao lắp dầm cầu theo tuyến dọc trục cầu, kết hợp di chuyển ngang để đưa dầm đến đúng vị trí lắp đặt. Thiết bị làm việc nhờ phối hợp chuyển động di chuyển trên ray với cơ cấu nâng–hạ, cho phép tiếp nhận dầm từ bãi/xe vận chuyển, đưa dầm ra vùng công xôn, di chuyển và hạ dầm lên gối. Do làm việc theo chu trình nhịp–nhịp và tham gia trực tiếp vào quá trình lắp đặt cấu kiện, thiết bị đòi hỏi độ ổn định, độ chính xác và an toàn cao. Cổng trục gồm kết cấu thép dạng giàn với dầm chính có một đầu công xôn ghép từ nhiều đốt; phía trên bố trí ray cho xe tời chạy. Trên dàn lắp các xe tời mang cơ cấu nâng (pa lăng cáp đơn) và cơ cấu di chuyển; pa lăng được dẫn động bởi động cơ, hộp giảm tốc, phanh đĩa và khớp nối để quay tang cuốn cáp.



Hình 2. Cấu tạo tổng thể cổng trục công xôn.

1 - Chân trước; 2 - Kết cầu dầm chính; 3 - Xe tời nâng hạ; 4 - Cụm palăng cáp;

5 - Dầm bê tông; 6 - Chân sau; 7 - Cơ cấu di chuyển; 8 - Chân phụ; 9 - Trụ cầu.

2.2. Xây dựng mô hình và phân tích động lực học trong quá trình máy làm việc khi hạ dầm

2.2.1. Mô hình động lực học

a) Giả thiết

Để xây dựng mô hình tính toán động lực học, cần một số giả thiết sau:

- Khảo sát trong quá trình hạ khi dầm bê tông ở trạng thái treo.

- Dầm đang treo trong không gian với chiều cao nhất định, người điều khiển khởi động động cơ của bộ máy nâng hạ hàng để hạ dầm xuống. Độ chùng cáp ($\delta \neq 0$);

- Xem xét mô hình trong mặt phẳng thẳng đứng; Bốn cơ cấu nâng hạ có thông số kết cấu và làm việc đồng thời; Khi bộ máy nâng hạ làm việc thì bộ máy di chuyển đứng yên;

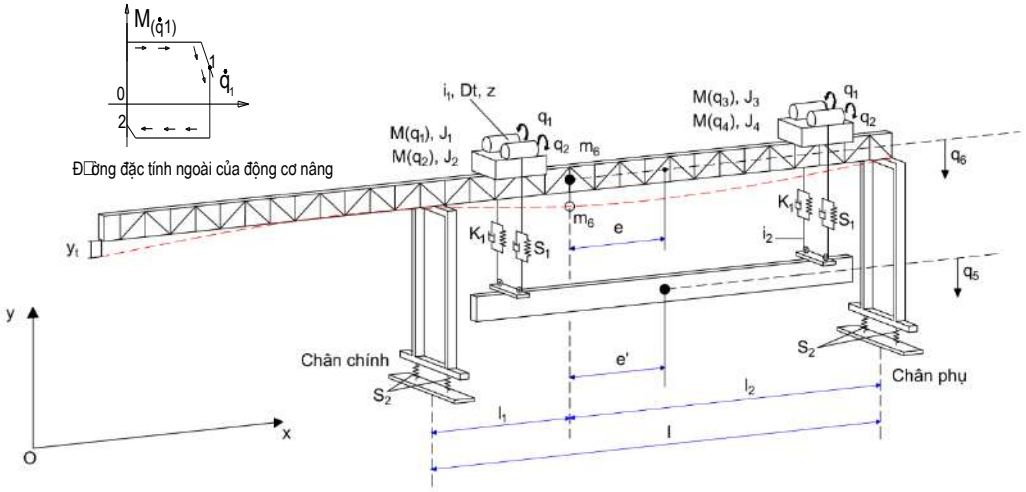
- Khối lượng của dầm dàn chính cổng trục và dầm bê tông được quy về trọng tâm (coi như điểm giữa của kết cấu thép tầng trên, ký hiệu là m_6 và giả thiết đặt lên gối lò xo có độ cứng là S_2) ; Bỏ qua ảnh hưởng của tải trọng gió.

- Cáp treo vật khi làm việc được coi như một lò xo chịu kéo.

- Bỏ qua yếu tố dập tắt dao động trong kết cấu thép cổng trục và bộ máy nâng hạ.

- Do ảnh hưởng của trọng lượng bản thân KCT (kết cấu thép) cổng trục và trọng lượng hàng nâng nên KCT tầng trên dạng dàn sẽ bị võng xuống một độ võng tĩnh ban đầu là yt.

- Trục động cơ, các trục và bánh răng hộp giảm tốc và các khớp nối coi như cứng tuyệt đối và có tính đến độ đàn hồi của kết cấu thép.



Hình 3. Mô hình động lực học cổng trục công xôn trong trường hợp hạ dầm khi cáp căng

Trong đó:

J_1, J_2, J_3, J_4 - các mô men quán tính quy dẫn của rôto động cơ điện và khớp nối cơ cấu nâng ($kg.m^2$);

m_5, m_6- khối lượng của dầm BTCT và khối lượng kết cấu thép dàn chính (kg);

i_1 - tỷ số truyền của hộp giảm tốc;

i_2 - số nhánh cáp;

D_t - đường kính tang cuốn cáp (m);

z - số nhánh cáp vào tang;

R - bán kính quy dẫn của cơ cấu nâng (m);

e - khoảng cach lệch tâm của m_5 và m_6 theo phương ngang (m);

l - khẩu độ của cổng trục (m); S_1 là độ cứng của một nhánh cáp hàng (N/m);

S_2 - độ cứng quy dẫn của hệ chân cổng trục (N/m);

K_1 - hệ số dập tắt dao động của cáp hàng (Ns/m);

q_1, q_2, q_3, q_4 - chuyển vị góc của trục động cơ (rad);

q_5, q_6 - chuyển vị thẳng đứng của dầm BTCT và dàn trên cổng trục (m).

Quá trình hạ dầm từ trạng thái cáp căng đến khi chạm xà mũ gồm 3 giai đoạn:

- Hạ dầm khi cáp căng: Dầm treo tự do, cơ cấu nâng–hạ làm việc và phanh được mở, dầm được hạ xuống khi cáp vẫn chịu lực căng.

- Chạm xà mũ: Dầm bắt đầu tiếp xúc xà mũ, độ chùng cáp xuất hiện; nếu động cơ vẫn quay thì lực căng cáp giảm dần và về 0.

- Nhả cáp: Động cơ tiếp tục quay làm cáp nhả ra, độ chùng tăng; lực căng cáp bằng 0.

Trong phạm vi bài báo chỉ xét giai đoạn 1: các tọa độ $q_1 \div q_4$quay theo chiều hạ dầm, còn q_5, q_6 dịch chuyển ngược chiều dương.

b) Thiết lập phương trình chuyển động

Chúng ta dùng phương trình Lagrange loại II để viết phương trình chuyển động có dạng như sau:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_1}\right)-\frac{\partial T}{\partial q_i}+\frac{\partial \phi}{\partial \dot{q}_i}+\frac{\partial U}{\partial q_i}=Q_i\,(i=1\div 6) \tag{1}$$

* Động năng của hệ xác định như sau:

$$T=T_1+T_2+T_3+T_4+T_5+T_6 \tag{2}$$

Trong đó $T_1 \div T_4$ lần lượt là động năng chuyển động quay của 4 trục động cơ, T_5 , T_6 là động năng chuyển động tịnh tiến theo phương Oy của dầm bê tông và kết cấu thép cổng trục. Suy ra:

$$T=\frac{1}{2}J_1\dot{q}_1^2+\frac{1}{2}J_2\dot{q}_2^2+\frac{1}{2}J_3\dot{q}_3^2+\frac{1}{2}J_4\dot{q}_4^2+\frac{1}{2}m_5\dot{q}_5^2+\frac{1}{2}m_6\dot{q}_6^2 \tag{3}$$

* Tính thế năng của hệ: $U = U_1 + U_2 + U_3$

Trong đó:

U_1: Thế năng biến dạng của cáp hàng và thế năng vị trí của hàng nâng;

U_2: Thế năng tích lũy trong lò xo S2 và thế năng vị trí của khối lượng m3;

U_3: Thế năng biến dạng của kết cấu thép tầng trên.

- Tính U_1: Thế năng biến dạng của cáp hàng và thế năng vị trí của hàng nâng (bao gồm: biến dạng của 4 dây cáp hàng và vị trí của trọng tâm m_5)

$$U_1=\frac{1}{2}S_1\Delta l_1^2+\frac{1}{2}S_1\Delta l_2^2+\frac{1}{2}S_1\Delta l_3^2+\frac{1}{2}S_1\Delta l_4^2-m_5gq_5 \tag{4}$$

Trong đó:

Δl_i - Độ dãn dài của cụm cáp nâng hàng thứ i;

$\Delta l_i = \Delta l_{ti} + \Delta l_{di}$ $(i = 1\div4)$ với: Δl_{ti} - độ dãn tĩnh; Δl_{di} - độ dãn động

Ở đây:

- Độ dãn tĩnh Δl_{ti}: $\Delta l_{t_1} = \Delta l_{t_2} = \Delta l_{t_3} = \Delta l_{t_4} = \Delta l_t = \dfrac{m_5 g}{4 i_2 S_1}$

- Độ dãn động Δl_d: $\Delta l_{d_1} = -q_1 . \dfrac{z.D_t.i_2}{2 i_1 i_2} + i_2 q_5 - i_2 q_6$

Đặt $R = \dfrac{zD_t}{2 i_1 i_2}$ là bán kính quy đổi. Suy ra: $\Delta l_{d1} = i_2(-Rq_1 + q_5 - q_6)$.

Tương tự ta có:

$\Delta l_{d2} = i_2(-Rq_2 + q_5 - q_6)$; $\Delta l_{d3} = i_2(-Rq_3 + q_5 - q_6)$; $\Delta l_{d4} = i_2(-Rq_4 + q_5 - q_6)$

Cuối cùng chúng ta có:

$$U_1 = \frac{1}{2} S_1 \left[\frac{m_5 g}{4 i_2 S_1} + i_2(-Rq_1 + q_5 - q_6) \right]^2 + \frac{1}{2} S_1 \left[\frac{m_5 g}{4 i_2 S_1} + i_2(-Rq_2 + q_5 - q_6) \right]^2 + \\ + \frac{1}{2} S_1 \left[\frac{m_5 g}{4 i_2 S_1} + i_2(-Rq_3 + q_5 - q_6) \right]^2 + \frac{1}{2} S_1 \left[\frac{m_5 g}{4 i_2 S_1} + i_2(-Rq_4 + q_5 - q_6) \right]^2 - m_5 g q_5 \quad (5)$$

- Tính U_2: Thế năng tích lũy trong lò xo S_2 và thế năng vị trí của khối lượng m_6 của kết cấu thép cổng trục.

$$U_2 = \frac{1}{2} S_2 \Delta l_C^2 + \frac{1}{2} S_2 \Delta l_P^2 - m_6 g q_6 \quad (6)$$

Với biến dạng tổng bên chân chính: $\Delta l_C = \Delta_1 + \Delta_2 + \Delta_3$

Trong đó:

Δ_1 - Biến dạng tĩnh (độ lún) của lò xo S_2 khi chịu trọng lượng của KCT cổng trục,

Δ_2 - Biến dạng tĩnh của lò xo S_2 khi chịu trọng lượng của hàng,

Δ_3 - Độ dịch chuyển của khối lượng m_6, $\Delta_3 = -q_6$;

Biến dạng tổng bên chân phụ: $\Delta l_P = \Delta_4 + \Delta_5 + \Delta_3$

Trong đó:

Δ_4 - Biến dạng tĩnh (độ lún) của lò xo S_2 khi chịu trọng lượng của KCT cổng trục;

Δ_5 - Biến dạng tĩnh của lò xo S_2 khi chịu trọng lượng của hàng;

Δ_3 - Độ dịch chuyển của khối lượng m_6, $\Delta_3 = -q_6$.

Ta tính được:

$$\Delta_1 = \frac{m_6 g}{2 S_2} \left(1 + \frac{2e}{l} \right); \Delta_4 = \frac{m_6 g}{2 S_2} \left(1 - \frac{2e}{l} \right) \quad (7)$$

Với e là độ lệch tâm giữa dầm và giữa 2 chân cổng trục.

$$\Delta_2 = \frac{m_5 g}{S_2} \left(1 - \frac{l_1 + e'}{l} \right); \Delta_5 = \frac{m_5 g}{S_2} \left(1 - \frac{l_2 + e'}{l} \right) \quad (8)$$

Với e’ là độ lệch tâm giữa dầm và KCT cổng trục.

$$\Delta_3 = -q_6 \quad (9)$$

Suy ra:

$$\Delta l_c = -\frac{m_6 g}{2S_2}\left(1+\frac{2e}{l}\right)-\frac{m_5 g}{S_2}\left(1-\frac{l_1+e'}{l}\right)-q_6 \tag{10}$$

$$\Delta l_p = -\frac{m_6 g}{2S_2}\left(1-\frac{2e}{l}\right)-\frac{m_5 g}{S_2}\left(1-\frac{l_2-e'}{l}\right)-q_6 \tag{11}$$

Suy ra thế năng biến dạng của lò xo S_2 và vị trí khối lượng m_6 của KCT cổng trục:

$$\begin{aligned} U_2 = &\frac{1}{2}S_2\left[-\frac{m_6 g}{2S_2}\left(1+\frac{2e}{l}\right)-\frac{m_5 g}{S_2}\left(1-\frac{l_1+e'}{l}\right)-q_6\right]^2 + \\ &+\frac{1}{2}S_2\left[-\frac{m_6 g}{2S_2}\left(1-\frac{2e}{l}\right)-\frac{m_5 g}{S_2}\left(1-\frac{l_2-e'}{l}\right)-q_6\right]^2 - mgq_6 \end{aligned} \tag{12}$$

- Tính U_3: Thế năng biến dạng của kết cấu thép tầng trên

Gọi y_t là độ võng tĩnh của kết cấu thép cổng trục, ta có:

$$U_3 = \frac{1}{2}S_3(-y_t - q_6)^2 = \frac{1}{2}S_3\left(-\frac{m_6 g}{S_3}-q_6\right)^2 \tag{13}$$

Cuối cùng ta có hàm thế năng: $U = U_1 + U_2 + U_3$

$$\begin{aligned} U = &\frac{1}{2}S_1\left[\frac{m_5 g}{4i_2 S_1}+i_2(-Rq_1+q_5-q_6)\right]^2+\frac{1}{2}S_1\left[\frac{m_5 g}{4i_2 S_1}+i_2(-Rq_2+q_5-q_6)\right]^2 + \\ &+\frac{1}{2}S_1\left[\frac{m_5 g}{4i_2 S_1}+i_2(-Rq_3+q_5-q_6)\right]^2+\frac{1}{2}S_1\left[\frac{m_5 g}{4i_2 S_1}+i_2(-Rq_4+q_5-q_6)\right]^2 - m_5 g q_5 + \\ &+\frac{1}{2}S_2\left[-\frac{m_6 g}{2S_2}\left(1+\frac{2e}{l}\right)-\frac{m_5 g}{S_2}\left(1-\frac{l_1+e'}{l}\right)-q_6\right]^2 + \\ &+\frac{1}{2}S_2\left[-\frac{m_6 g}{2S_2}\left(1-\frac{2e}{l}\right)-\frac{m_5 g}{S_2}\left(1-\frac{l_2-e'}{l}\right)-q_6\right]^2 - mgq_6 + \frac{1}{2}S_3\left(-\frac{m_6 g}{S_3}-q_6\right)^2 \end{aligned} \tag{14}$$

* Tính hàm hao tán ϕ:

Ở đây các giảm chấn ở các nhánh cáp đóng vai trò là bộ phận dập tắt dao động, ta có hàm hao tán:

$$\phi = \frac{1}{2}\mathrm{K}_1\dot{\Delta}\mathrm{l}_1^2+\frac{1}{2}\mathrm{K}_1\dot{\Delta}\mathrm{l}_2^2+\frac{1}{2}\mathrm{K}_1\dot{\Delta}\mathrm{l}_3^2+\frac{1}{2}\mathrm{K}_1\dot{\Delta}\mathrm{l}_4^2 \tag{15}$$

Với:

$$\dot{\Delta}\mathrm{l}_1 = \mathrm{i}_2(-\mathrm{R}\dot{\mathrm{q}}_1+\dot{\mathrm{q}}_5-\dot{\mathrm{q}}_6);\dot{\Delta}\mathrm{l}_2 = \mathrm{i}_2(-\mathrm{R}\dot{\mathrm{q}}_2+\dot{\mathrm{q}}_5-\dot{\mathrm{q}}_6);\dot{\Delta}\mathrm{l}_3 = \mathrm{i}_2(-\mathrm{R}\dot{\mathrm{q}}_3+\dot{\mathrm{q}}_5-\dot{\mathrm{q}}_6);\dot{\Delta}\mathrm{l}_4 = \mathrm{i}_2(-\mathrm{R}\dot{\mathrm{q}}_4+\dot{\mathrm{q}}_5-\dot{\mathrm{q}}_6)$$

Suy ra:

$$\phi = \frac{1}{2}\mathrm{K}_1\mathrm{i}_2^2(-\mathrm{R}\dot{\mathrm{q}}_1+\dot{\mathrm{q}}_5-\dot{\mathrm{q}}_6)^2+\frac{1}{2}\mathrm{K}_1\mathrm{i}_2^2(-\mathrm{R}\dot{\mathrm{q}}_2+\dot{\mathrm{q}}_5-\dot{\mathrm{q}}_6)^2+\frac{1}{2}\mathrm{K}_1\mathrm{i}_2^2(-\mathrm{R}\dot{\mathrm{q}}_3+\dot{\mathrm{q}}_5-\dot{\mathrm{q}}_6)^2+\frac{1}{2}\mathrm{K}_1\mathrm{i}_2^2(-\mathrm{R}\dot{\mathrm{q}}_4+\dot{\mathrm{q}}_5-\dot{\mathrm{q}}_6)^2$$

* Các lực suy rộng được xác định như sau:

$$\mathrm{Q}_1 = \mathrm{M}(\dot{\mathrm{q}}_1);\ \mathrm{Q}_2 = \mathrm{M}(\dot{\mathrm{q}}_2);\ \mathrm{Q}_3 = \mathrm{M}(\dot{\mathrm{q}}_3);\ \mathrm{Q}_4 = \mathrm{M}(\dot{\mathrm{q}}_4);\ \mathrm{Q}_5 = 0;\ \mathrm{Q}_6 = 0$$

- Phương trình mômen mở máy động cơ nâng hạ cổng trục được thể thiện như sau:

$$M(\dot{q}_1) = a - b\dot{q}_1; M(\dot{q}_2) = a - b_1\dot{q}_2; M(\dot{q}_3) = a - b\dot{q}_3; M(\dot{q}_4) = a - b\dot{q}_4;$$

Trong đó:

a: mômen mở máy

b: hệ số suy giảm mômen theo tốc độ

Từ phương trình Lagrange loại II, ta lập được hệ phương trình chuyển động. Sắp xếp lại dưới dạng ma trận chúng ta có:

$$M\ddot{q} + K\dot{q} + Sq = Q$$

Trong đó:

$$M.\ddot{q} = \begin{bmatrix} J_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & J_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & J_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & J_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & m_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & m_6 \end{bmatrix} . \begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \\ \ddot{q}_3 \\ \ddot{q}_4 \\ \ddot{q}_5 \\ \ddot{q}_6 \end{bmatrix}; K.\dot{q} = \begin{bmatrix} i_2^2R^2K_1 & 0 & 0 & 0 & -i_2^2RK_1 & i_2^2RK_1 \\ 0 & i_2^2R^2K_1 & 0 & 0 & -i_2^2RK_1 & i_2^2RK_1 \\ 0 & 0 & i_2^2R^2K_1 & 0 & -i_2^2RK_1 & i_2^2RK_1 \\ 0 & 0 & 0 & i_2^2R^2K_1 & -i_2^2RK_1 & i_2^2RK_1 \\ -i_2^2RK_1 & -i_2^2RK_1 & -i_2^2RK_1 & -i_2^2RK_1 & 4i_2^2K_1 & -4i_2^2K_1 \\ i_2^2RK_1 & i_2^2RK_1 & i_2^2RK_1 & i_2^2RK_1 & -4i_2^2K_1 & 4i_2^2K_1 \end{bmatrix} . \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \\ \dot{q}_4 \\ \dot{q}_5 \\ \dot{q}_6 \end{bmatrix}$$

$$S.q = \begin{bmatrix} i_2^2R^2S_1 & 0 & 0 & 0 & -i_2^2RS_1 & i_2^2RS_1 \\ 0 & i_2^2R^2S_1 & 0 & 0 & -i_2^2RS_1 & i_2^2RS_1 \\ 0 & 0 & i_2^2R^2S_1 & 0 & -i_2^2RS_1 & i_2^2RS_1 \\ 0 & 0 & 0 & i_2^2R^2S_1 & -i_2^2RS_1 & i_2^2RS_1 \\ -i_2^2RS_1 & -i_2^2RS_1 & -i_2^2RS_1 & -i_2^2RS_1 & 4i_2^2S_1 & -4i_2^2S_1 \\ i_2^2RS_1 & i_2^2RS_1 & i_2^2RS_1 & i_2^2RS_1 & -4i_2^2S_1 & 4i_2^2S_1 + 2S_2 + S_3 \end{bmatrix} . \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \\ q_6 \end{bmatrix}; Q = \begin{bmatrix} M(\dot{q}_1) + \frac{m_5gR}{4} \\ M(\dot{q}_2) + \frac{m_5gR}{4} \\ M(\dot{q}_3) + \frac{m_5gR}{4} \\ M(\dot{q}_4) + \frac{m_5gR}{4} \\ 0 \\ -m_6g \end{bmatrix}$$

- Lực căng trong một nhánh cáp: $F_c = F_K + F_S$

+ Bên trái: $F_{ct} = \frac{m_5g}{4i_2} + i_2K_1(-R\dot{q}_1 + \dot{q}_5 - \dot{q}_6) + i_2S_1(-Rq_1 + q_5 - q_6)$

+ Bên phải: $F_{cp} = \frac{m_5g}{4i_2} + i_2K_1(-R\dot{q}_3 + \dot{q}_5 - \dot{q}_6) + i_2S_1(-Rq_3 + q_5 - q_6)$

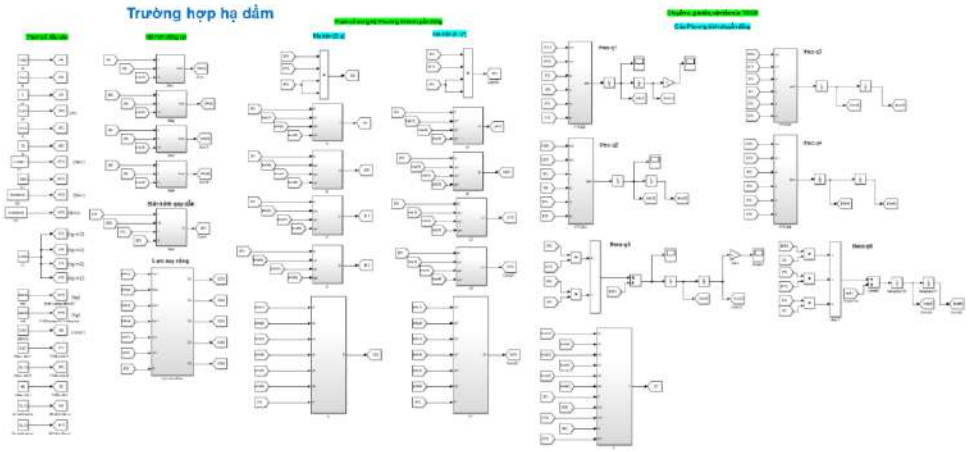
- Lực tác dụng vào các chân cổng trục:

+ Bên chân chính: $F_{CC} = S_2\Delta l_C = \frac{m_6g}{2}\left(1 + \frac{2e}{l}\right) + m_5g\left(1 - \frac{l_1 + e'}{l}\right) + S_2q_6$

+ Bên chân phụ: $F_{CP} = S_2\Delta l_P = \frac{m_6g}{2}\left(1 - \frac{2e}{l}\right) + m_5g\left(1 - \frac{l_2 - e'}{l}\right) + S_2q_6$

c) Xây dựng chương trình giải phương trình bằng Matlab Simulink

Sử dụng các khối thuật toán trong thư viện Matlap-Simulink ta thiết lập được chương trình để giải hệ phương trình trên như sau:



Hình 4. Chương trình mô phỏng bằng Matlab Simulink

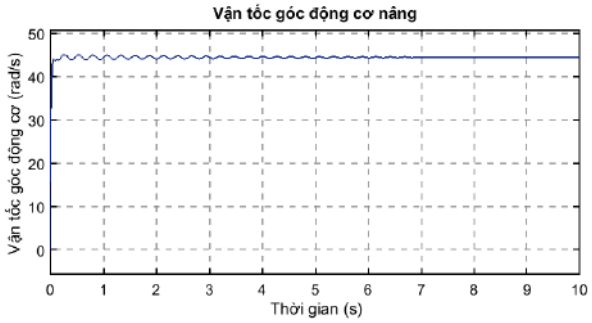
Kết quả thu được khi giải hệ phương trình vi phân chuyển động:

Với phương trình chuyển động tổng quát ở trên có thể ứng dụng để xác định các thông số động lực học của các cổng trục công son lao lắp dầm cầu khác nhau. Dưới đây, tác giả lấy thông số kỹ thuật của cổng trục công son khẩu độ 40 m, tải nâng 80 tấn làm thông số đầu vào để giải hệ phương trình chuyển động trên, cụ thể như sau:

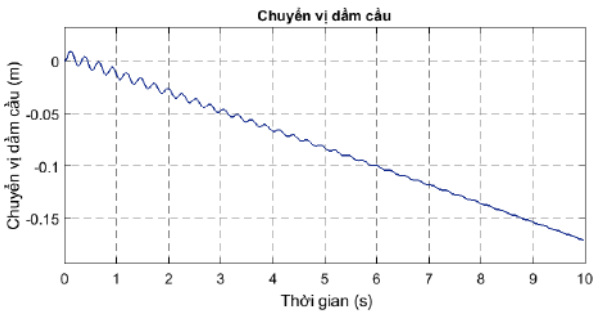
Bảng 1. Các thông số đầu vào

TT	Ký hiệu	Diễn giải	Đơn vị tính	Giá trị
1	a	Mômen mở máy		1592
2	b	Hệ số suy giảm mômen theo tốc độ		35,8
3	z	Số nhánh cáp vào tang		1
4	D_t	Đường kính tang	m	0,25
5	i_1	Tỷ số truyền HGT		31,5
6	i_2	Số nhánh cáp trên puli động		10
7	S_1	Độ cứng cáp treo hàng	N/m	314327
8	S_2	Độ cứng chân cổng trục	N/m	22426553
9	S_3	Độ cứng KCT cổng trục	N/m	51652893
10	J_1-J_4	Mô men quán tính	Kg.m2	0,595
11	m_5	Khối lượng dầm bê tông	Kg	80000
12	m_6	Khối lượng KCT cổng trục	Kg	52000

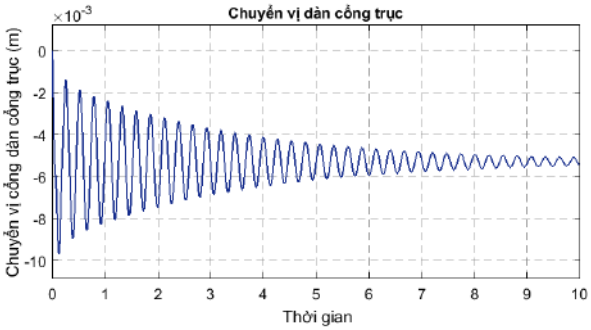
Giải phương trình trên với các thông số cụ thể của cổng trục trong bảng 1 thu được kết quả về đặc trưng dao động và lực động. Tác giả chỉ xin giới thiệu một số kết quả thu được như sau:



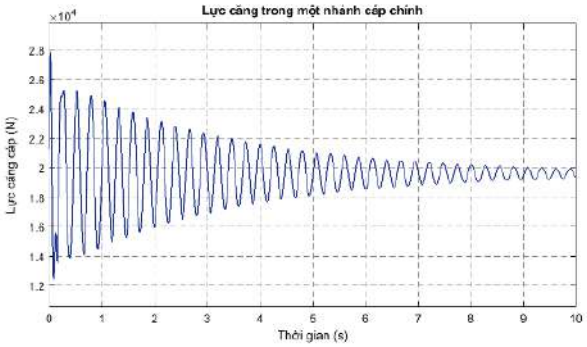
Hình 5. Vận tốc góc của động cơ nâng



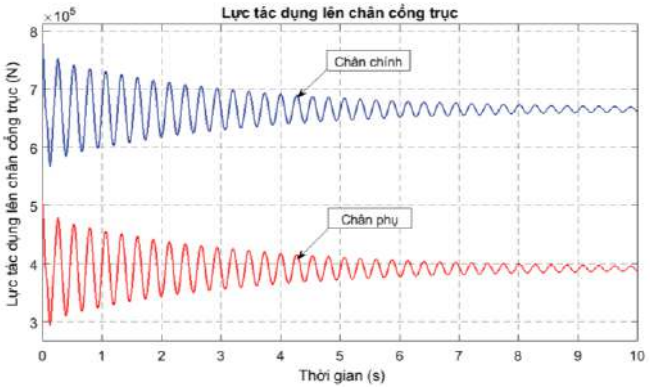
Hình 6. Chuyển vị dầm cầu bê tông



Hình 7. Chuyển vị của dàn cổng trục



Hình 8. Lực căng trong một nhánh cáp



Hình 9. Lực tác dụng vào chân chính và phụ của máy

Nhận xét:

Khi dầm treo ổn định, mở phanh và cho động cơ làm việc để hạ dầm xuống vị trí lắp đặt. Từ Hình 5, vận tốc góc động cơ tăng nhanh từ 0 lên khoảng 44–45 rad/s trong ~0,5 s đầu, sau đó dao động nhỏ quanh giá trị làm việc và gần xác lập sau khoảng 2 s. Chuyển vị dầm (Hình 6) giảm theo chiều âm, trong 2–3 s đầu xuất hiện dao động biên độ nhỏ chồng lên chuyển động hạ; đến t = 10 s đạt khoảng −0,17 m. Dàn chính cổng trục (Hình 7) dao động cưỡng bức mạnh nhất ở những giây đầu với biên độ xấp xỉ 9×10^{-3} m và giảm dần, gần tắt sau 7–8 s. Lực căng một nhánh cáp (Hình 8) biến thiên mạnh khi mở phanh, ban đầu dao động khoảng $1{,}3\times10^4$ – $2{,}5\times10^4$ N rồi ổn định quanh $2{,}0\times10^4$ N sau 7–8 s, vẫn nằm trong giới hạn cho phép. Lực tác dụng lên chân cổng trục (Hình 9) thể hiện phân bố không đối xứng: chân chính chịu tải lớn hơn, dao động mạnh trong 3–4 s đầu quanh khoảng $6{,}7\times10^5$ N; chân phụ dao động quanh khoảng $4{,}0\times10^5$N với biên độ nhỏ hơn; cả hai giảm rõ rệt sau 7–8 s.

Như vậy, giai đoạn bất lợi nhất tập trung trong vài giây đầu khi mở phanh, khi các đại lượng động lực đạt cực trị trước khi tắt dần về trạng thái ổn định.Như vậy, có thể thấy giai đoạn nguy hiểm nhất của quá trình hạ dầm tập trung trong vài giây đầu tiên khi mở phanh và hệ bắt đầu chuyển động. Trong khoảng thời gian này, các đại lượng động lực học (chuyển vị, lực căng cáp và phản lực chân trục) đạt giá trị cực trị trước khi dao động tắt dần về trạng thái xác lập.

3. Khảo sát các thông số động lực học của cổng trục dạng dàn công xôn lao lắp dầm cầu

3.1. Thông số cần khảo sát

Từ mô hình động lực học và các kết quả tính toán đã trình bày ở Phần 2, có thể thấy một số thông số ảnh hưởng đáng kể đến đáp ứng của cổng trục trong quá trình hạ dầm, bao gồm: độ cứng của cáp, độ cứng tương đương của kết cấu (đặc biệt tại vùng chân cổng trục), và khối lượng các bộ phận tham gia dao động (khối lượng dàn chính).

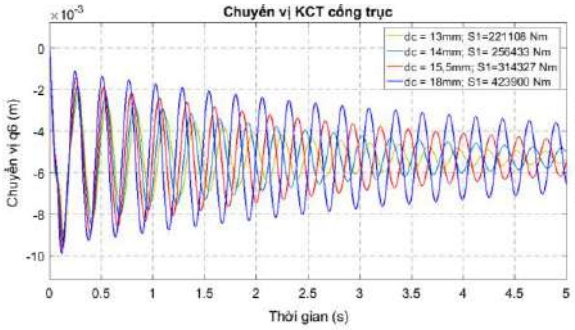
Các thông số trên được lựa chọn khảo sát nhằm đánh giá mức độ ảnh hưởng của chúng đến các đại lượng đặc trưng của hệ như chuyển vị, dao động kết cấu, lực căng cáp trong các điều kiện làm việc khác nhau. Các giá trị dùng trong khảo sát được tổng hợp trong Bảng 2.

Bảng 2. Giá trị các thông số khảo sát

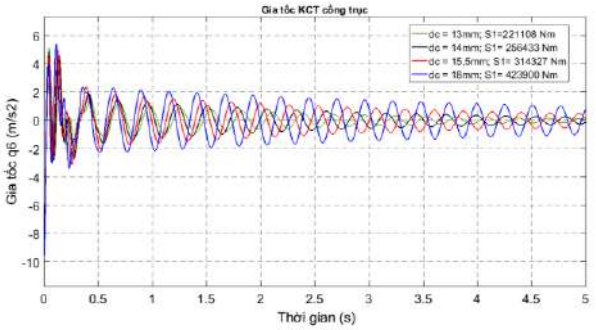
TT	Thông số	Đơn vị	Ký hiệu	Giá trị khảo sát			
				1	2	3	4
1	Độ cứng một nhánh cáp	Nm	S1	221108	256433	314327	423900
2	Độ cứng kết cấu thép chân cổng trục	Nm	S2	12000000	22426553	42000000	
3	Khối lượng của kết cấu thép dàn chính	Kg	m_6	46800	52000	57200	

3.1. Khảo sát các giá trị khi thay đổi độ cứng S_1

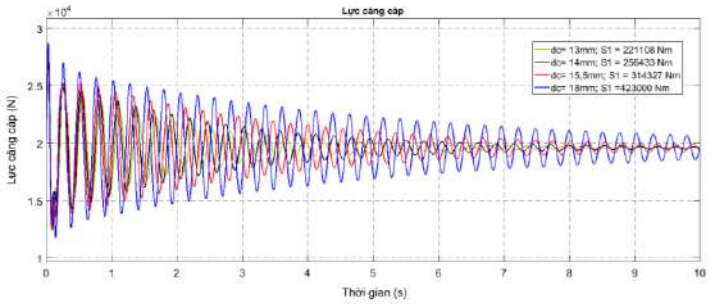
Sau khi chạy chương trình khảo sát, tác giả thu được các giá trị như sau:



Hình 10. Chuyển vị q_6 của dầm cổng trục



Hình 11. Gia tốc của dầm cổng trục



Hình 12. Lực căng trong một nhánh cáp

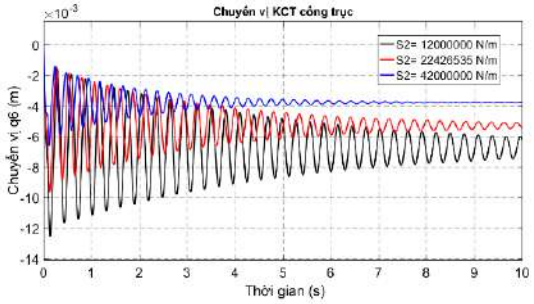
Bảng 3. Giá trị các thông số khảo sát lực động khi thay đổi đường kính cáp

Thông số	Kí hiệu	Đơn vị	Giá trị			
			1	2	3	4
Đường kính cáp	d_c	mm	13	14	15,5	18
Độ cứng của cáp nâng hàng	S_1	N/m	221108	256433	314327	423900
Lực động trong nhánh cáp	F_c	N	26797	27250	27870	28748
Lực căng cáp tĩnh	F_{ct}	N	19620	19620	19620	19620
Lực làm việc tối đa của nhánh cáp	$[F_c]$	N	22734	26366	32318	43584

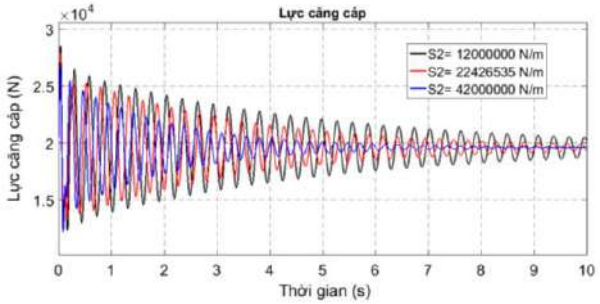
Theo Hình 10–12, khi tăng đường kính cáp từ dc = 13 mm lên dc = 18 mm (tương ứng độ cứng S1 tăng từ $2{,}21\times10^5$ N/m lên $4{,}239\times10^5$ N/m), tần số dao động của hệ tăng (mật độ dao động dày hơn) và thời gian dập tắt kéo dài do hệ số cản không đổi, làm tỷ số tắt dần giảm. Vì vậy, tăng độ cứng cáp không đồng nghĩa với giảm dao động mà có thể làm tăng phản ứng động, kéo theo lực động trong cáp và phản lực chân.

Lực động cực đại trong một nhánh cáp tăng theo đường kính, từ Fc,max = 2,6797×10^4 N (dc = 13 mm) lên Fc,max = 2,8748×10^4 N (dc = 18 mm), trong khi lực căng tĩnh Fct = 1,9620×10^4 N không đổi. So sánh với lực cho phép [Fc] cho thấy các trường hợp dc = 13 mm và dc = 14 mm không thỏa điều kiện bền động (Fc,max > [Fc]), trong khi dc = 15,5 mm và dc = 18 mm thỏa điều kiện làm việc an toàn (Fc,max ≤ [Fc]). Do đó, xét đến tải trọng động trong quá trình hạ dầm, đường kính cáp tối thiểu nên chọn dc = 15,5 mm.

3.2. Khảo sát các giá trị khi thay đổi độ cứng S_2



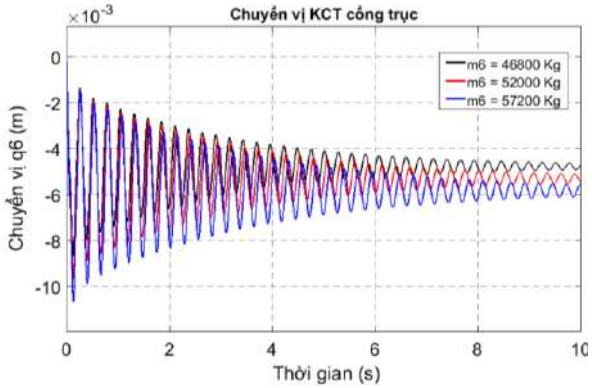
Hình 13. Chuyển vị q6 của dầm cổng trục



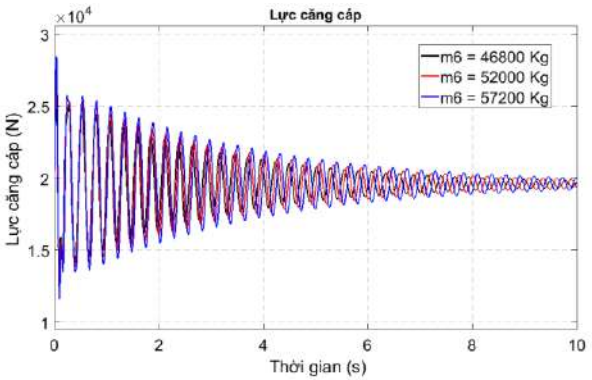
Hình 14. Lực căng trong một nhánh cáp

Kết quả khảo sát khi thay đổi độ cứng chân cổng trục S_2cho thấy ảnh hưởng rõ rệt đến đáp ứng dao động của hệ. Khi thay đổi độ cứng chân cổng trục S_2, đáp ứng dao động của hệ biến đổi rõ rệt. Giảm S_2 từ 42×10^6 N/m xuống 12×10^6 N/m làm biên độ chuyển vị kết cấu tăng mạnh và thời gian dập tắt kéo dài; trường hợp S_2=12×10^6 N/m dao động duy trì gần hết khoảng 10 s, trong khi S_2=42×10^6 N/m cho dao động nhỏ và ổn định sớm. Đồng thời, khi S_2tăng, biên độ dao động lực căng cáp giảm và lực căng nhanh chóng ổn định quanh giá trị cân bằng. Vì vậy, tăng S_2 giúp giảm dao động kết cấu và lực động trong cáp; trong thiết kế cần ưu tiên nâng cao độ cứng tổng thể của chân cổng trục. Kết quả khảo sát cho thấy việc tăng độ cứng chân cổng trục làm giảm đáng kể dao động kết cấu và lực động trong cáp nâng. Do đó, trong thiết kế cần ưu tiên nâng cao độ cứng tổng thể của chân cổng trục nhằm cải thiện đặc tính ổn định động và giảm tác động lực động lên thiết bị trong quá trình hạ dầm.

3.3. Khảo sát các giá trị khi thay đổi khối lượng KCT cổng trục m_6



Hình 15. Chuyển vị q6 của dầm cổng trục



Hình 16. Lực căng trong một nhánh cáp

Từ đồ thị chuyển vị kết cấu cổng trục, khi tăng khối lượng từ m_6=46800 kg lên 57200 kg, biên độ dao động ban đầu tăng rõ rệt, đồng thời giá trị chuyển vị cực đại theo phương âm lớn hơn. Trường hợp m_6=57200 kg cho biên độ dao động lớn nhất và thời gian dập tắt dao động

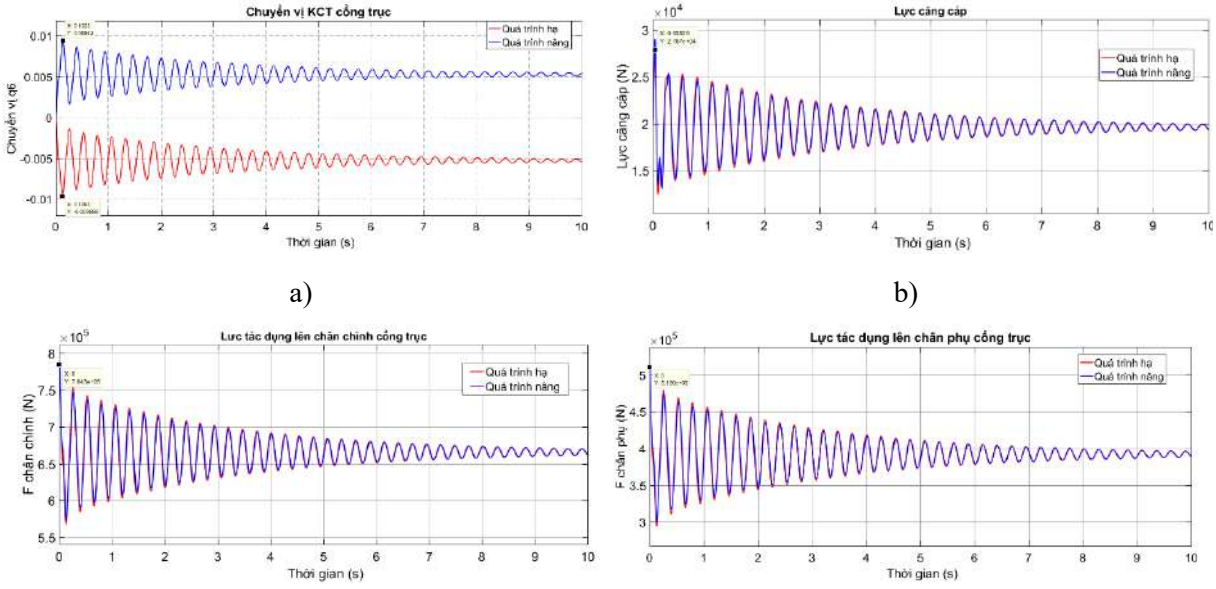
kéo dài hơn so với hai trường hợp còn lại. Điều này cho thấy khi khối lượng kết cấu tăng, quán tính của hệ tăng, làm đáp ứng dao động mạnh hơn và khả năng ổn định động giảm.

Ngược lại, lực căng trong nhánh cáp chính hầu như không thay đổi nhiều giữa các trường hợp khảo sát. Biên độ dao động và giá trị lực động cực đại gần tương đương nhau ở cả ba trường hợp khối lượng. Điều này chứng tỏ lực động trong cáp chủ yếu phụ thuộc vào khối lượng dầm nâng và đặc tính hệ cáp, ít chịu ảnh hưởng bởi khối lượng kết cấu thép cổng trục.

Từ kết quả trên có thể nhận thấy việc tăng khối lượng kết cấu thép cổng trục làm gia tăng dao động và chuyển vị của hệ mà không mang lại lợi ích trong việc giảm lực động trong cáp. Điều này không có lợi cho quá trình làm việc của cổng trục công xôn, do đó trong thiết kế cần tối ưu hóa khối lượng kết cấu nhằm giảm quán tính dao động, nâng cao tính ổn định động và hiệu quả làm việc của thiết bị.

4. So sánh các thông số động lực học trong trường hợp nâng và hạ dầm

Để đánh giá ảnh hưởng của chiều vận hành đến đáp ứng động lực của cổng trục công xôn, nghiên cứu tiến hành so sánh trực tiếp hai trường hợp nâng dầm và hạ dầm theo ba chỉ tiêu: chuyển vị cổng trục, lực căng cáp, và lực tác dụng lên chân cổng trục. Các kết quả được thể hiện bằng các đồ thị theo thời gian và được đánh giá thông qua giá trị cực đại.



Hình 17. So sánh các đại lượng động lực học giữa quá trình nâng và hạ dầm
a) Chuyển vị q6; b) Lực căng cáp; c) Lực tác dụng lên chân chính; dc) Lực tác dụng lên chân phụ

Bảng 4. Kết quả giá trị cực đại của các thông số ĐLH trong quá trình nâng và hạ dầm

Stt	Thông số khảo sát	Giá trị cực đại	
		Quá trình hạ	Quá trình nâng
1	Chuyển vị KCT cổng trục	0,00966 m	0,00943 m
2	Lực căng cáp	27870 N	29020 N
3	Lực tác dụng chân chính	$7{,}843\times10^5$ N	$7{,}843\times10^5$ N
4	Lực tác dụng chân phụ	$5{,}106\times10^5$ N	$5{,}106\times10^5$ N

Nhận xét: Kết quả mô phỏng so sánh hai quá trình nâng dầm và hạ dầm theo ba chỉ tiêu gồm chuyển vị KCT cổng trục, lực căng cáp và lực tác dụng lên chân cổng trục cho thấy đáp

ứng của hệ trong cả hai trường hợp đều có dạng dao động tắt dần và tiến tới trạng thái ổn định, với tần số dao động và mức tắt dần gần tương đương. Tuy nhiên, xét theo giá trị cực đại, chuyển vị KCT trong trường hợp hạ dầm lớn hơn so với nâng dầm 0,00966 m so với 0,00943 m, chênh khoảng 2,44%), do đó hạ dầm là trạng thái bất lợi hơn đối với tiêu chí biến dạng/ổn định kết cấu. Ngược lại, lực căng cáp cực đại trong trường hợp nâng dầm lớn hơn so với hạ dầm (29020 N so với 27870 N, chênh khoảng 4,0%), vì vậy nâng dầm là trạng thái bất lợi hơn đối với tiêu chí bền cáp. Đối với lực tác dụng lên chân cổng trục, các đồ thị cho thấy hai trường hợp gần như trùng nhau và giá trị cực đại thu được là tương đương cho cả chân chính và chân phụ ($7{,}843\times10^5$ N và $5{,}106\times10^5$ N), cho thấy trong điều kiện mô phỏng hiện tại, tải trọng cực đại lên hệ chân không nhạy với chiều vận hành.

Từ các kết quả trên, việc lựa chọn thông số thiết kế và kiểm tra bền nên thực hiện theo nguyên tắc bao giá trị bất lợi: chọn chuyển vị cực đại theo trường hợp hạ dầm, chọn lực căng cáp cực đại theo trường hợp nâng dầm, và sử dụng giá trị cực đại lực chân chung cho cả hai quá trình nhằm đảm bảo an toàn làm việc trong toàn bộ chu kỳ vận hành.

5. Kết luận

Trong bài báo này, mô hình động lực học 6 bậc tự do của cổng trục công xôn trong quá trình hạ dầm cầu đã được xây dựng nhằm mô tả tương tác giữa kết cấu cổng trục, hệ cáp nâng và dầm cầu. Phương trình chuyển động được thiết lập theo Lagrange loại II và giải bằng Matlab–Simulink, cho phép đánh giá các đáp ứng động trong điều kiện làm việc bất lợi.

Từ mô hình, khảo sát tham số cho thấy: tăng đường kính cáp làm tăng độ cứng hệ, tăng tần số dao động và lực động trong cáp; theo bền động, đường kính cáp tối thiểu nên chọn 15,5 mm; tăng độ cứng chân cổng trục làm giảm rõ rệt dao động và lực động, do đó cần ưu tiên nâng cao độ cứng tổng thể của hệ chân; tăng khối lượng kết cấu thép làm tăng biên độ và kéo dài thời gian dập tắt dao động, trong khi lực động trong cáp không giảm đáng kể, vì vậy cần tối ưu khối lượng kết cấu.

Ngoài ra, so sánh hai quá trình nâng và hạ cho thấy đáp ứng đều dao động tắt dần; hạ dầm cho chuyển vị KCT cực đại lớn hơn, trong khi nâng dầm cho lực căng cáp cực đại lớn hơn. Lực tác dụng cực đại lên chân chính và chân phụ gần như trùng nhau. vậy, lựa chọn thông số và kiểm tra bền nên theo nguyên tắc: chọn theo hạ đối với chuyển vị/ổn định kết cấu và theo nâng đối với bền cáp, nhằm đảm bảo an toàn trong toàn bộ chu kỳ làm việc.

Các kết quả nghiên cứu không chỉ có ý nghĩa về mặt lý thuyết mà còn cung cấp cơ sở thực tiễn cho việc lựa chọn thông số thiết kế hợp lý, tính toán hệ số động, kiểm tra bền mỏi và nâng cao độ an toàn khi khai thác cổng trục công xôn trong thi công cầu. Mô hình đề xuất có thể tiếp tục được phát triển để nghiên cứu các chế độ làm việc khác như phanh khẩn cấp, dừng đột ngột hoặc các dạng cổng trục lao lắp dầm cầu có cấu hình kết cấu khác.

Tài liệu tham khảo

1. Bộ Khoa học và Công nghệ, *TCVN 2737:2023 – Tải trọng và tác động – Tiêu chuẩn thiết kế*, Việt Nam, 2023.
2. N. Đ. Zrnić, V. M. Gašić và S. M. Bošnjak, *"Dynamic responses of a gantry crane system due to a moving body considered as moving oscillator"*, Archives of Civil and Mechanical Engineering, vol. 15, no. 1, pp. 243–250, 2015, doi: 10.1016/j.acme.2014.02.002.

3. B. Jin, J. Zeng, P. Gao, H. Zhang và S. Ge, *"Dynamic Modeling and Validation of Dual-Cable Double-Pendulum Systems for Gantry Cranes"*, Machines, vol. 13, no. 8, Art. no. 676, 2025, doi: 10.3390/machines13080676.
4. Vinh V. Nguyen, Trung N. Nguyen, Chi T. Nguyen (2020), ***"****Research on the dynamics of a hydraulic static-pile-pressing machine during the process of lifting and slewing of piles****"***, J. Comput. Appl. Res. Mech. Eng. Vol. 11. No. 1, pp. 191-203, DOI: 10.22061/JCARME.2020.3973.1469
5. Jaafar HI, Mohamed Z, Jamian JJ, et al. *Dynamic behaviour of a nonlinear gantry crane system*. Proced Technol 2013; 11: 419–425.
6. Ismail RMTR, Ahmad MA, Ramli MS, et al (2009). *"Nonlinear dynamic modelling and analysis of a 3-D overhead gantrycrane system with payload variation"*. In: Proceedings of the 2009 third UKSim European symposium on computer modeling and simulation, Athens, 25–27 November 2009, pp.350–354. New York: IEEE.
7. Bùi Thanh Danh (2017), *"Nghiên cứu cơ sở khoa học cho tính toán kết cấu thép cổng trục chuyên dùng do việt nam chế tạo phục vụ lao lắp dầm bêtông trên xà mũ cầu"*, Luận án Tiến sĩ.
8. TS. Trần Văn Chiến (2005), *Động lực học máy trục*, Nhà xuất bản Hải phòng.
9. PGS. TS. Nguyễn Văn Vịnh (2004), *Động lực học máy xây dựng- Xếp dỡ*. Bài giảng cao học.Trường Đại học giao thông vận tải.
10. Ngô Minh Tiến, Nguyễn Ngọc Trung, Nguyễn Viết Tân, *"Nghiên cứu động lực học cổng trục công xôn trong quá trình nâng dầm phục vụ lao lắp dầm cầu".* Tạp chí Cơ khí Việt Nam số đặc biệt, tr. 105-110, tháng 10 năm 2025.

Dynamic analysis of a cantilever gantry crane for bridge girder erection in the load-lowering case

Abstract*: In one operating cycle of a cantilever gantry crane used for precast bridge girder erection, the workflow consists of three stages: lifting, girder transfer, and lowering. The authors have previously published a dynamic analysis of the lifting stage; this paper continues with the remaining stages to complete the assessment of the entire operating cycle, in which lowering is considered potentially hazardous due to dynamic loads and system vibrations. The objectives are not only to investigate the variation of key dynamic responses and the influence of characteristic parameters, but also to provide a quantitative comparison with the lifting case in order to select safe design/operating parameters under the most unfavorable conditions.*

To address this problem, a six-degree-of-freedom (6-DOF) dynamic model is developed to capture the interaction among the hoisting mechanism, wire ropes, the girder, and the crane structure. The equations of motion are derived using the second-order Lagrange formulation and employed to analyze the dynamic responses during girder transfer and lowering. The results provide a scientific basis for the design, fabrication, and safe operation of cantilever gantry cranes in bridge construction.

Keywords: cantilever gantry crane; dynamics; 6-DOF model; lifting case; lowering case; girder transfer; dynamic load; hoisting system vibration; Lagrange equations.

Nghiên cứu ứng dụng hệ thống vây giảm lắc chủ động giảm thiểu dao động của xe bơi tự hành

Nguyễn Thế Anh[1*], Nguyễn Viết Tân[2], Nguyễn Minh Kha[2], Lê Trọng Tuấn[2]

[1]Lớp 36D, Hệ quản lý học viên Sau đại học

[2]Viện Cơ khí Động lực, Học viện Kỹ thuật Quân sự

Email: ngtheanh2502@gmail.com

Tóm tắt

Khi xe bơi tự hành hoạt động trong môi trường nước biển, các dao động lắc ngang và lắc dọc có thể làm giảm đáng kể độ an toàn, ảnh hưởng đến khả năng mang tải, khả năng sử dụng hỏa lực cũng như sức khỏe của kíp lái. Đặc biệt là đối với xe bơi tự hành có boong hở làm nhiệm vụ đổ bộ đường biển, việc giảm thiểu các dao động lắc dọc, lắc ngang có ý nghĩa rất lớn khi phương tiện hoạt động trong môi trường nước. Với những ưu điểm vượt trội, vây cân bằng chủ động đã trở thành một giải pháp phổ biến để giảm dao động cho phương tiện hàng hải. Hiệu quả giảm dao động của hệ thống này phụ thuộc chủ yếu vào bộ điều khiển. Tuy nhiên, hiện nay trên các loại xe bơi tự hành chưa được ứng dụng hệ thống này để giảm các dao động xuất hiện trong quá trình bơi. Trong nghiên cứu này, tiến hành xác định dao động của xe bơi tự hành trong quá trình bơi và đưa ra một giải pháp giảm dao động là sử dụng hệ thống vây cân bằng chủ động với thuật toán điều khiển đa đầu vào – đa đầu ra (Multi-Input Multi-Output – MIMO). Thuật toán điều khiển MIMO là một phương pháp hiệu quả trong việc xử lý đồng thời nhiều kênh tín hiệu và có thể nâng cao hiệu quả ổn định của hệ thống. Bằng cách xem xét tác động của sóng ngẫu nhiên, thuật toán MIMO được thiết kế nhằm đưa ra thuật toán tối ưu cho hệ thống vây cân bằng. Kết quả mô phỏng trên Matlab đã khẳng định tính ổn định và hiệu quả của hệ thống vây cân bằng với thuật toán điều khiển được đề xuất.

***Từ khóa**: Dao động, xe bơi tự hành, vây cân bằng, MIMO.*

1. Đặt vấn đề

Trong quá trình bơi, dao động lắc ngang và lắc dọc của xe bơi tự hành (XBTH) làm giảm đáng kể độ an toàn, khả năng mang tải, khả năng sử dụng hỏa lực, cũng như sức khỏe của kíp chiến đấu. Đặc biệt là đối với XBTH có boong hở, việc giảm thiểu các dao động lắc dọc, lắc ngang có ý nghĩa rất lớn khi phương tiện hoạt động trong môi trường nước. Các kích thích bên ngoài như sóng ngẫu nhiên, gió và dòng chảy là nguồn phát sinh chính của các dao động này, làm tăng tải trọng động, giảm hiệu suất cơ động và gây mệt mỏi cho kíp xe 1. Do đó, giảm dao động là một yêu cầu cấp thiết trong thiết kế và khai thác phương tiện vượt sông.

Do sự quan trọng của việc giảm dao động, các giải pháp thụ động hiện có ví dụ như: vây giảm lắc bị động, két nước chống lắc, hoặc ổn định kiểu con quay hồi chuyển đã được ứng dụng rộng rãi cho các phương tiện hàng hải và có thể giảm biên độ lắc ở một mức độ nhất định. Tuy nhiên, hiệu quả của các giải pháp thụ động phụ thuộc mạnh vào chế độ sóng, tốc độ chuyển động và hình dáng thân, do đó khó duy trì hiệu quả cao trong nhiều điều kiện tác chiến, khai thác khác nhau. Két nước giảm lắc, hoạt động dựa trên nguyên lý khi tàu lắc ngang, khối nước trong két dịch chuyển ngược pha, sinh ra mô men chống lại dao động và góp phần ổn định. Thiết bị này đặc biệt phù hợp với những tàu có tốc độ thấp hoặc thường xuyên neo đậu, đồng thời đem lại hiệu quả giảm lắc tương đối cao, với mức giảm biên độ dao động có thể đạt 50% – 60% 2. Một giải pháp phổ biến khác là vây giảm lắc thụ động thường được gắn cố định và vuông góc với thân tàu, có tác dụng hạn chế dao động lắc ngang nhờ làm tăng lực cản thủy động, qua đó có thể giảm biên độ lắc từ khoảng 20% đến 30% 3. Ngoài ra, bánh lái giảm lắc cũng được tận dụng trên nhiều tàu

nhờ khả năng tạo ra lực ổn định bổ sung trong quá trình điều động, giúp giảm đáng kể lắc ngang mà không cần bổ sung thiết bị phụ trợ. Với các tàu nhỏ, ví dụ tàu cá, biên độ lắc ngang có thể được giảm tới 40% – 60% 4. Tuy nhiên, hiệu quả của các thiết bị thụ động thường chưa đạt mức tối ưu. Chính vì vậy, trong những năm gần đây, vây giảm lắc chủ động được nghiên cứu và triển khai trên tàu biển, cho phép nâng cao hiệu suất giảm dao động. Khi sử dụng hệ thống này, biên độ lắc ngang và lắc dọc có thể giảm tới 80% – 95% 5. Hiệu quả hoạt động của vây chủ động phụ thuộc chủ yếu vào bộ điều khiển, do đó việc thiết kế một bộ điều khiển phù hợp để khai thác tối đa hiệu quả của hệ thống trở thành nhiệm vụ trọng yếu. Gần đây, nhiều phương pháp điều khiển hiện đại đã được áp dụng cho hệ thống vây chủ động, điển hình như điều khiển trượt 6, điều khiển mờ 7, điều khiển dạng toàn phương 8 9, điều khiển dạng mạng thần kinh 10…

Trên các XBTH, những giải pháp này lại càng hạn chế bởi ràng buộc về không gian lắp đặt, khối lượng và yêu cầu tàng hình thủy động. Hiện nay trên XBTH có giải pháp lắp tấm cân bằng ở đuôi xe để giảm dao động lắc dọc, tuy nhiên dao động lắc ngang không được đề cập tới 11.

Trong những năm gần đây, cùng với sự phát triển của ngành cơ khí, điều khiển,… vây cân bằng chủ động chứng tỏ hiệu quả vượt trội trên tàu biển nhờ khả năng tạo lực chủ động để triệt giảm dao động theo thời gian thực; hiệu quả của hệ thống phụ thuộc chủ yếu vào bộ điều khiển. Dù vậy, đến nay chưa có ứng dụng hệ thống vây chủ động cho XBTH, một loại phương tiện hoạt động cả trên cạn và dưới nước, nơi có dao động xảy ra và đặt ra bài toán khắt khe hơn về hiệu suất – độ tin cậy – thời gian đáp ứng.

Khoảng trống nghiên cứu bao gồm: mô hình hóa động lực học XBTH trong môi trường sóng ngẫu nhiên, ứng dụng thiết bị vây cân bằng chủ động trên phương tiện, thiết kế bộ điều khiển cho hệ thống vây chủ động và đánh giá về khả năng giảm dao động dưới các điều kiện vượt chướng ngại nước khác nhau như vận tốc vượt sông, cấp sóng.

Trong nghiên cứu này tiến hành mô hình hóa và xác định dao động của XBTH trong điều kiện sóng ngẫu nhiên; xây dựng mô hình không gian trạng thái tuyến tính (đã tuyến tính hóa quanh chế độ làm việc) phù hợp cho thiết kế điều khiển. Đề xuất hệ thống vây cân bằng chủ động cho XBTH và thiết kế thuật toán điều khiển MIMO (Multi-Input Multi-Output) để điều khiển đồng thời nhiều cơ cấu chấp hành và đa kênh đo lường, hướng tới giảm biên độ lắc, cải thiện ổn định và nâng cao khả năng chiến đấu của phương tiện. Sử dụng phần mềm Matlab để xác định hiệu quả của hệ thống vây cân bằng dưới một vài điều kiện khai thác phương tiện khác nhau.

2. Xây dựng mô hình động lực học

2.1. Mô hình động lực học của xe bơi tự hành

Một số giả thiết đặt ra trong quá trình xây dựng mô hình và tính toán:

- XBTH bơi với vận tốc nhỏ (10÷15km/h).

- Sóng biển tác động theo phương hợp với chiều chuyển động của xe bơi một góc 135^o và không có sóng phản xạ.

- Coi toàn bộ xe là một khối lượng đồng nhất, được thu về trọng tâm của xe.

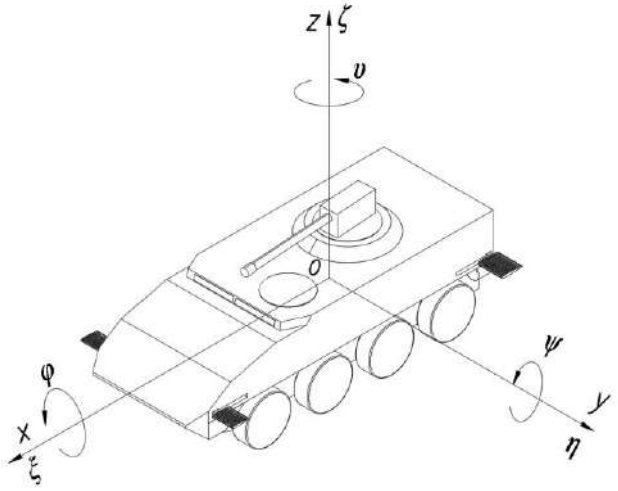
- Bỏ qua biến dạng đàn hồi của thân vỏ.

- Tiến hành khảo sát trên sóng tự do từ cấp 1 tới cấp 3.

Đối với các XBTH dạng boong hở, khi chuyển động bơi trong môi trường nước, XBTH có phần chìm trong nước và được xác định bằng thể tích giãn nước. Do nước là môi trường chất lỏng

và được coi là đàn hồi, không bị hạn chế bởi phương nào nên có thể mô tả XBTH dao động trong nước xét trường hợp tổng quát theo 6 bậc tự do. Để xây dựng mô hình khảo sát sự dao động của xe máy theo trường hợp tổng quát, đặt vào trọng tâm xe máy hệ trục tọa độ cố định Oxyz. Khi đó, các chuyển động tịnh tiến và chuyển động quay theo các trục tọa độ x, y, z được mô tả bằng các đại lượng ξ, η, ζ và φ, ψ, υ.

Vây giảm lắc được bố trí hai bên phía đầu và phía đuôi của phương tiện, tạo ra hệ lực và mômen tương tự như mô hình drone có 4 cánh. Khi hành quân trên cạn, các cánh được gập lại để đảm bảo thông số về chiều rộng. Khi bắt đầu xuống bến, các cánh được dẫn động để mở ra và vào vị trí làm việc.



Hình 1. Hệ trục tọa độ gắn với trọng tâm của phương tiện

Hệ phương trình tổng quát mô tả chuyển động của XBTH 12:

$$\begin{cases} \left(m + A_{11}\right)\ddot{\xi} + B_{11}\dot{\xi} = F_{\xi} \\ \left(m + A_{22}\right)\ddot{\eta} + B_{22}\dot{\eta} = F_{\eta} \\ \left(m + A_{33}\right)\ddot{\zeta} + B_{33}\zeta + \rho g A_w \zeta = F_{\zeta} \\ \left(I_{xx} + A_{44}\right)\ddot{\varphi} + B_{44}\dot{\varphi} + \rho g V G M_{\gamma}\varphi = M_{\varphi} \\ \left(I_{yy} + A_{55}\right)\ddot{\psi} + B_{55}\dot{\psi} + \rho g V G M_{L}\psi = M_{\psi} \\ \left(I_{zz} + A_{66}\right)\ddot{\upsilon} + B_{66}\dot{\upsilon} = M_{\upsilon} \end{cases} \quad (1)$$

trong đó A_{11} đến A_{16} là sự gia tăng khối lượng bám dính vào vỏ và dao động cùng với XBTH, B_{11} đến B_{16} là lực cản nhớt do nước tạo ra đối với XBTH, $\rho g A_w \varsigma$, $\rho g V G M_{\gamma}\varphi$, $\rho g V G M_{L}\psi$ là lực đẩy nổi của nước theo x, y, z, ρ là khối lượng riêng của nước, m là khối lượng của xe, GM_{γ} là chiều cao định khuynh theo mặt phẳng ngang xe, GM_{L} là chiều cao định khuynh theo mặt phẳng dọc của xe, V là thể tích choán nước, $F_{\xi}, F_{\eta}, F_{\varsigma}$ là lực của sóng tác động lên xe theo x, y, z, $M_{\varphi}, M_{\psi}, M_{\upsilon}$ là mô men do sóng gây nghiêng theo x, y, z, A_w là diện tích đường mớn nước, I_{xx}, I_{yy}, I_{zz} là mô men quán tính theo x, y, z.

Xét trường hợp XBTH chuyển động bơi với tốc độ thấp, dao động biên độ nhỏ thì chỉ xét đến 3 chuyển động: dịch chuyển theo phương z, xoay theo trục Ox và xoay theo trục Oy 12. Khi đó hệ phương trình chuyển động (1) được viết như sau:

$$\begin{cases} \left(m + A_{33}\right)\ddot{\zeta} + B_{33}\zeta + \rho g A_w \zeta = F_\zeta \\ \left(I_{xx} + A_{44}\right)\ddot{\varphi} + B_{44}\dot{\varphi} + \rho g V G M_\gamma \varphi = M_\varphi \\ \left(I_{yy} + A_{55}\right)\ddot{\psi} + B_{55}\dot{\psi} + \rho g V G M_L \psi = M_\psi \end{cases} \quad (2)$$

Khi có thêm vây giảm lắc, phương trình (2) trở thành:

$$\begin{cases} \left(m + A_{33}\right)\ddot{\zeta} + B_{33}\zeta + \rho g A_w \zeta = F_\zeta + F_\zeta' \\ \left(I_{xx} + A_{44}\right)\ddot{\varphi} + B_{44}\dot{\varphi} + \rho g V G M_\gamma \varphi = M_\varphi + M_\varphi' \\ \left(I_{yy} + A_{55}\right)\ddot{\psi} + B_{55}\dot{\psi} + \rho g V G M_L \psi = M_\psi + M_\psi' \end{cases} \quad (3)$$

Với $F_\varsigma', M_\varphi', M_\psi'$ là lực và mômen điều khiển được sinh ra bởi vây giảm lắc.

Công thức (3) được viết dưới dạng tổng quát như sau:

$$\left(M + \Delta M\right)\ddot{x} + C\dot{x} + Kx = F + F' \quad (4)$$

trong đó M là khối lượng hoặc mômen quán tính, ΔM là khối lượng hoặc mômen quán tính tăng thêm, C là hệ số dập của chuyển động lắc, K là độ cứng phục hồi (biểu thị bằng lượng giãn nước), F là lực hoặc mômen kích động của sóng, F' là lực hoặc mômen sinh ra bởi vây.

Từ (4) ta có: $\ddot{x} = \dfrac{1}{M + \Delta M}(F + F' - C\dot{x} - Kx)$ (5)

Đặt $x = \left[x_1, x_2\right]^T = \left[x, \dot{x}\right]^T$, (5) trở thành:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = A_1 x_1 + A_2 x_2 + B_1\left(F + F'\right) \end{cases} \quad (6)$$

trong đó A_1, A_2, B_1 là các hệ số và có giá trị như sau:

$$A_1 = -\frac{K}{M + \Delta M}; A_2 = -\frac{C}{M + \Delta M}; B_1 = \frac{1}{M + \Delta M} \quad (7)$$

2.2. Mô hình sóng

Trong thực tế, sóng nước không phải là sóng điều hòa mà là sự tổng hợp phức tạp của nhiều thành phần dao động ngẫu nhiên khác nhau. Khi nghiên cứu dao động của phương tiện vượt sông, việc sử dụng sóng điều hòa đơn lẻ thường không phản ánh đúng điều kiện thực tế. Do đó, cần mô phỏng sóng ngẫu nhiên dựa trên các phổ năng lượng đã được xây dựng từ thực nghiệm. Việc này cho phép đánh giá chính xác hơn phản ứng động lực học của phương tiện trong điều kiện tác chiến đa dạng.

Sóng ngẫu nhiên được mô tả bằng phổ năng lượng $S(\omega)$, biểu diễn phân bố mật độ năng lượng theo tần số góc ω. Một số dạng phổ thông dụng gồm 13:

Phổ Pierson–Moskowitz:

$$S(\omega) = \alpha g^2 \omega^{-5} \exp\left[-\beta\left(\frac{\omega_p}{\omega}\right)^4\right] \quad (8)$$

trong đó α, β là hệ số thực nghiệm, $\omega_p = 2\pi / T_p$ là tần số sóng.

Phổ JONSWAP được phát triển từ phổ Pierson–Moskowitz, bổ sung hệ số γ, mô tả chính xác hơn sóng biển ở cả vùng ven bờ.

Giả sử miền tần số được chia thành *N* khoảng nhỏ $\Delta\omega$. Biên độ sóng phần tử ở tần số ω_i được xác định từ phổ:

$$A_i = \sqrt{2S(\omega_i)\Delta\omega} \quad (9)$$

Để sóng mang tính ngẫu nhiên, mỗi thành phần *i* được gán một pha ε_i, có phân bố đều:

$$\varepsilon_i \sim U(0, 2\pi) \quad (10)$$

Tại một điểm (x, y) mặt sóng được tổng hợp từ:

$$\chi(x, y, t) = \sum_{i=1}^{N} A_i \cos(k_i x \cos\theta_i + k_i y \sin\theta_i - \omega_i t + \varepsilon_i) \quad (11)$$

với $k_i = \omega_i^2 / g$, θ_i là góc lan truyền sóng.

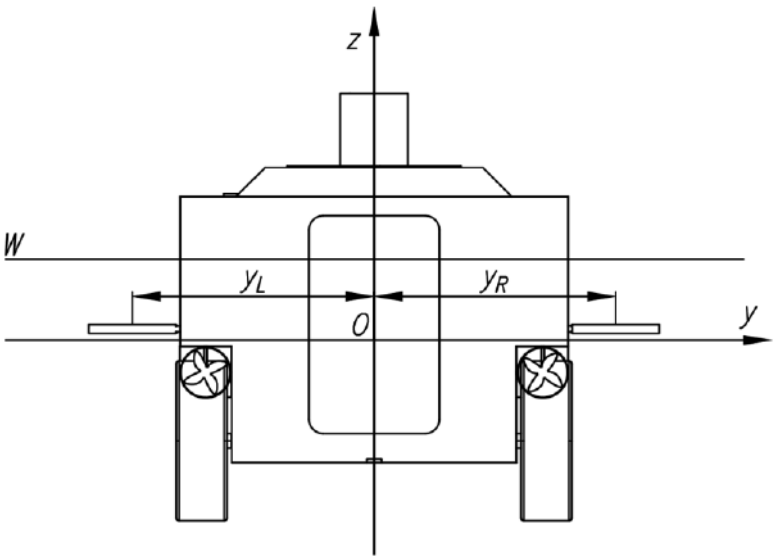
Để đảm bảo sóng sinh ra có chiều cao đúng với giá trị chiều cao của bảng cấp sóng H_s, ta thực hiện chuẩn hóa theo điều kiện:

$$\chi_{ch} = \chi \cdot \frac{H_s}{H_{s.g}} \quad (12)$$

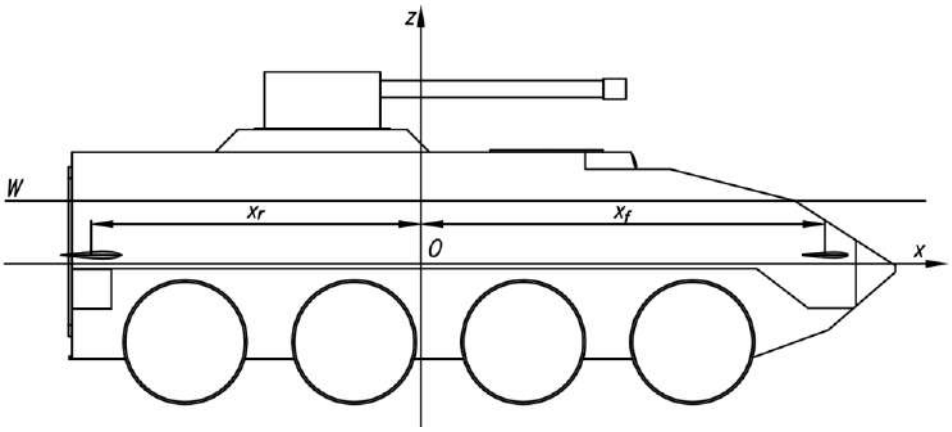
trong đó $H_{s.g} = 4\sqrt{m_0}$, với $m_0 = \int S(\omega) d\omega$.

3. Xây dựng thuật toán xác định góc lắc của vây giảm lắc chủ động

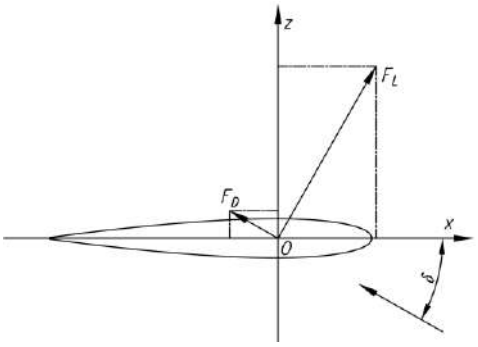
Để giảm dao động lắc dọc và lắc ngang, bài báo trình bày phương án bố trí hệ thống vây giảm lắc bao gồm: 2 vây giảm lắc phía trước và 2 vây giảm lắc phía sau.



Hình 2. Bố trí vây giảm lắc ở mặt phẳng ngang



Hình 3. Bố trí vây giảm lắc ở mặt phẳng dọc



Hình 4. Các lực sinh ra do vây giảm lắc

Lực nâng sinh ra bởi vây giảm lắc được tính theo công thức sau:

$$F_L = \frac{1}{2}\rho V^2 S C_L^{\delta} \delta \tag{13}$$

trong đó ρ là khối lượng riêng của nước, S là diện tích tiếp xúc của vây giảm lắc, V là vận tốc dòng tới, C_L^{δ} là hệ số lực nâng (phụ thuộc vào góc tới δ của dòng). (Bỏ qua lực cản F_D do độ lớn của $F_D \approx (0,05 \div 0,1) F_L$ và cánh tay đòn của nó nhỏ, tạo ra mômen không đáng kể trong quá trình dao động lắc của XBTH).

Để điều khiển được hệ thống vây chủ động này, bài báo đưa ra phương án sử dụng thuật toán MIMO được trình bày dưới đây.

Từ tín hiệu góc nghiêng, ta có mômen điều khiển theo luật PD như sau 14:

$$\begin{cases} M_x^{dk}(t) = -K_{p\varphi}\varphi(t) - K_{d\varphi}\dot{\varphi}(t) \\ M_y^{dk}(t) = -K_{p\psi}\psi(t) - K_{d\psi}\dot{\psi}(t) \end{cases} \quad (14)$$

trong đó $K_{p\varphi}, K_{d\varphi}$ là hệ số tỉ lệ và vi phân cho lắc ngang, $K_{p\psi}, K_{d\psi}$ là hệ số tỉ lệ và vi phân cho lắc dọc.

Mômen mong muốn $\left(M_x^{dk}(t), M_y^{dk}(t)\right)$ được quy đổi thành lực nâng ảo mà mỗi cặp vây giảm lắc phải tạo ra.

Đối với dao động lắc ngang có cánh tay đòn của lực nâng là y_L, y_R, hiệu của lực nâng bên trái và bên phải là:

$$\Delta F_{LR}^{dk} = \frac{M_x^{dk}(t)}{y_L + y_R} \quad (15)$$

Đối với dao động lắc dọc, khoảng cách từ khối tâm tới điểm đặt lực là x_f và x_r, tải trọng được chia ra như sau:

$$F_{ftong}^{dk} = \frac{\beta_p M_x^{dk}}{x_f},\ F_{rtong}^{dk} = -\frac{(1-\beta_p) M_y^{dk}}{x_r} \quad (16)$$

Từ đó ta có lực phân bổ cho từng vây:

$$\begin{cases} F_{fL}^{dk} = \frac{1}{2} F_{ftong}^{dk} + \frac{1}{2}\Delta F_{LR}^{dk} \\ F_{fR}^{dk} = \frac{1}{2} F_{ftong}^{dk} - \frac{1}{2}\Delta F_{LR}^{dk} \\ F_{rL}^{dk} = \frac{1}{2} F_{ftong}^{dk} + \frac{1}{2}\Delta F_{LR}^{dk} \\ F_{rR}^{dk} = \frac{1}{2} F_{ftong}^{dk} - \frac{1}{2}\Delta F_{LR}^{dk} \end{cases} \quad (17)$$

Các ràng buộc đối với góc quay của vây giảm lắc như sau: $|\delta| \le \delta_{\max}$, $|\dot{\delta}| \le \dot{\delta}_{\max}$

Đáp ứng động học bậc 1 của cơ cấu là:

$$\tau\dot{\delta}(t) + \delta(t) = \delta^{dk}(t) \quad (18)$$

trong đó τ là hằng số thời gian của cơ cấu chấp hành.

Lực nâng sinh ra tại các vây là:

$$\begin{cases} F_{fL} = d_f \delta_{fL} \\ F_{fR} = d_f \delta_{fR} \\ F_{rL} = d_r \delta_{rL} \\ F_{rR} = d_r \delta_{rR} \end{cases} \quad (19)$$

với $d_f = \frac{1}{2}\rho V^2 S_f C_L^{\delta},\ d_r = \frac{1}{2}\rho V^2 S_r C_L^{\delta}$

Lực và Mômen bổ sung $F_\zeta', M_\varphi', M_\psi'$ được sinh ra bởi vây giảm lắc là:

$$\begin{cases} F_\zeta' = F_{fL} + F_{rL} + F_{fR} + F_{rR} \\ M_\varphi' = \left(F_{fL} + F_{rL} - F_{fR} - F_{rR}\right)\cdot\frac{1}{2}\left(y_L + y_R\right) \\ M_\psi' = \left(F_{fL} + F_{fR}\right)x_f - \left(F_{rL} + F_{rR}\right)x_r \end{cases} \quad (20)$$

4. Kết quả tính toán và thảo luận

4.1. Thông số đầu vào

Hệ phương trình vi phân (3) được giải bằng phương pháp số thông qua phần mềm MATLAB với các giá trị ban đầu như sau:

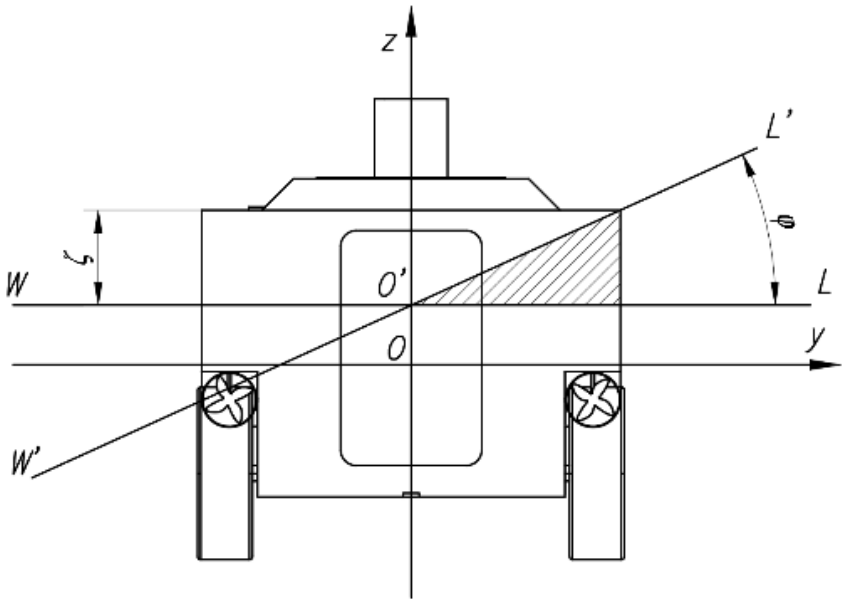
1) Dịch chuyển ban đầu: $\zeta_0 = 0,\ \varphi_0 = 0,\ \psi_0 = 0$;

2) Vận tốc ban đầu: $\dot{\zeta}_0 = 0,\ \dot{\varphi}_0 = 0,\ \dot{\psi}_0 = 0$;

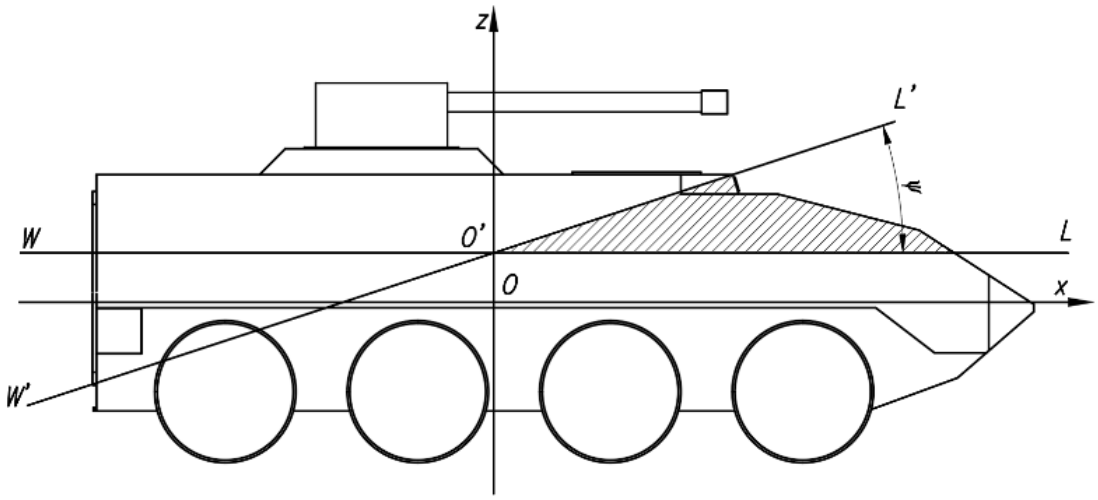
Thông số đầu vào (đơn vị thuộc hệ SI):

Lấy thông số kỹ thuật của xe chiến đấu bộ binh Pandur II 8x8 15 có các thông số chiến kỹ thuật như sau: Khối lượng xe m = 20000 kg; Chiều dài L = 7,3 m; Chiều rộng B = 2,67 m; Độ sâu chìm nước T = 1,3 m; Khối lượng riêng của nước là $\rho = 1000$ kg/m^3.

Theo 16 xác định giá trị giới hạn về dao động của phương tiện để đảm bảo tính không chìm



Hình 5. Xác định giá trị giới hạn



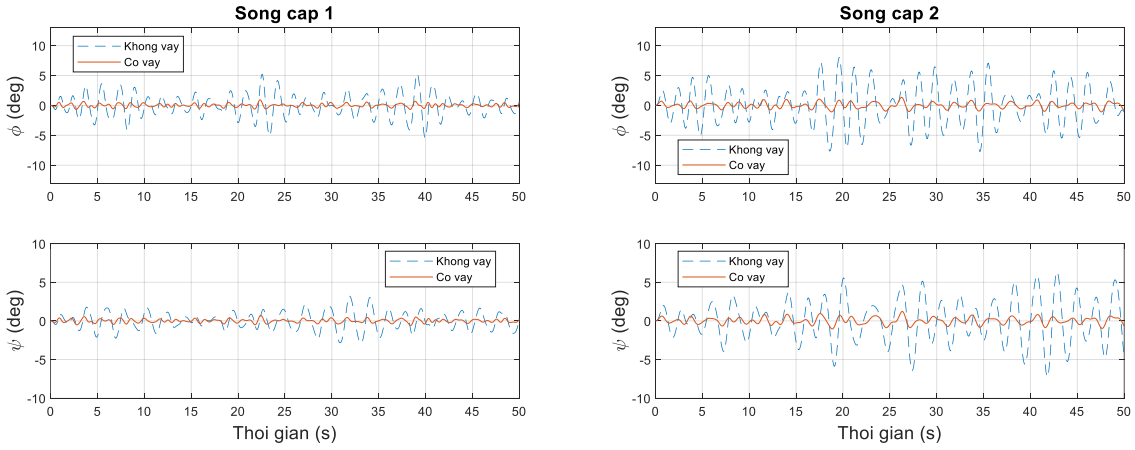
Hình 6. Xác định giá trị giới hạn

Khi dao động với góc nhỏ, phương tiện lắc quanh điểm O' là giao điểm của trục Oz và đường mớn nước W-L'. Ta tiến hành xác định các thông số trên bằng phương pháp đo thực nghiệm hoặc sử dụng bản vẽ CAD của XBTH. Đối với xe Pandur II 8x8, ta xác định được giới hạn về dao động của phương tiện là: $\zeta \le 0,7$ (m), $-13° \le \varphi \le 13°, -10° \le \psi \le 10°$.

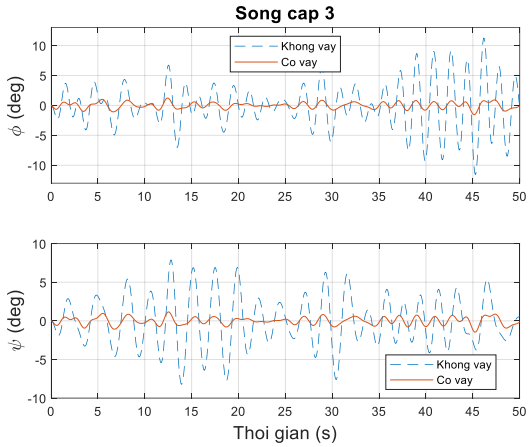
Thông số cơ bản của sóng từ cấp 1 tới cấp 3: Chiều cao sóng H_s = [0,25 0,5 1,0 1,5 2,0] (m), Chu kỳ sóng T_s = [2 3 4 5 6] s.

4.2. Kết quả tính toán động lực học dao động của XBTH khi sử dụng vây giảm lắc

Sử dụng phương pháp Runge-Kutta giải phương trình vi phân (3) trong MATLAB R2019a, Kết quả nhận được thể hiện trên Hình 7:



a) *b)*



c)

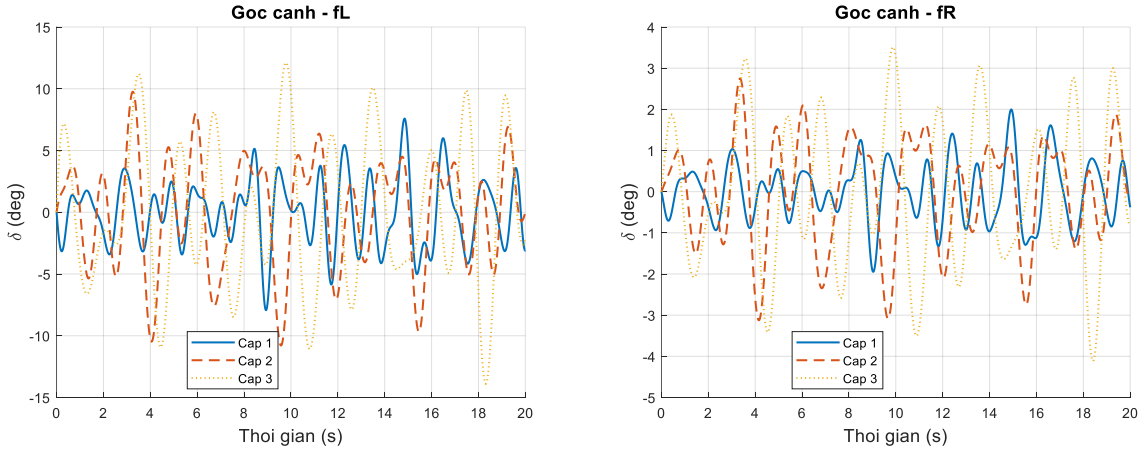
Hình 7. Đồ thị dao động của XBTH có vây giảm lắc.

a – Sóng cấp 1; b - Sóng cấp 2; c - Sóng cấp 3

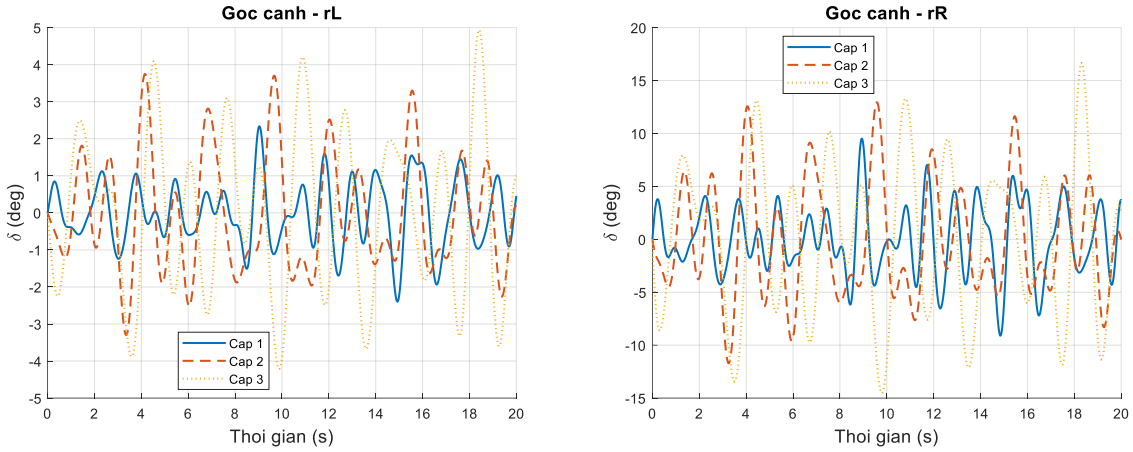
Nhận xét:

- Về quy luật: Dao động thẳng đứng gần thay đổi không đáng kể so với trước khi lắp thiết bị cân bằng, do lực đẩy thẳng đứng sinh ra rất nhỏ, không đủ ảnh hưởng đến quá trình dao động vốn chủ yếu quyết định bởi thể tích chìm nước và đặc tính sóng nên trong đồ thị chỉ xét tới lắc ngang và lắc dọc. Đối với dao động lắc ngang và dao động lắc dọc, các giá trị đều giảm rõ rệt về biên độ sau lắp thiết bị cân bằng, nhờ các cặp lực – mô men điều khiển sinh ra bởi tấm cân bằng.

- Về giá trị: Đối với dao động lắc ngang và lắc dọc, cả hai dạng dao động đã nằm trong phạm vi mà con người có thể chịu đựng khi ở trong phương tiện ($\varphi \leq 4^\circ$, $\psi \leq 3^\circ$). Đảm bảo sự thoải mái của người vận hành và khả năng giữ ổn định thao tác. Cho thấy việc sử dụng vây giảm lắc chủ động trên XBTH là cần thiết và mang lại hiệu quả giảm lắc rất cao.



a)



b)

Hình 8. Đồ thị góc quay của vây giảm lắc.

a - Vây trước; b - Vây sau

Nhận xét:

Hệ vây giảm lắc chủ động được điều khiển độc lập ở từng vây để sinh ra lực và mômen cần thiết chống lại mômen gây ra sự lắc của XTBH, giá trị của các góc vây giảm lắc nằm trong khoảng $-15^{\circ} \le \delta_i \le 15^{\circ}$ để đảm bảo hệ số lực nâng C_L^{δ} nằm trong khoảng tuyến tính và đảm bảo khả năng thiết kế, chế tạo vây, tránh hiện tưởng mỏi và giảm tuổi thọ thiết bị. Góc quay của vây giảm lắc trơn tru không bị giật cục tạo điều kiện thuận lợi trong quá trình điều khiển hệ vây giảm lắc bằng hệ thống thủy lực hoặc cơ khí.

5. Kết luận

Trong nghiên cứu này, đã xây dựng được mô hình tính toán dao động của phương tiện khi hoạt động trên sóng, xét đến cả hai trường hợp: không sử dụng vây giảm lắc và có sử dụng giảm lắc. Mô hình đã mô tả đầy đủ các dạng dao động chủ yếu gồm dao động thẳng đứng, dao động lắc ngang và dao động lắc dọc.

Cách giải bài toán dựa trên việc thiết lập hệ phương trình động lực học, áp dụng phương pháp số để mô phỏng đáp ứng dao động theo thời gian dưới tác động của sóng ngẫu nhiên với nhiều cấp độ khác nhau. Trên cơ sở đó, mô hình cho phép so sánh rõ ràng hiệu quả của giảm lắc chủ động trong việc làm giảm biên độ dao động, đặc biệt là đối với các dao động lắc ngang và lắc dọc vốn ảnh hưởng mạnh đến tính ổn định và sự thoải mái của kíp vận hành của phương tiện. Ngoài ra nội dung nghiên cứu còn trình bày thuật toán MIMO giúp phân bố lực trên các vây giảm lắc theo điều kiện góc lắc giúp XBTH đạt được sự ổn định cao và giảm thiểu đáng kể góc lắc.

Đặc biệt, bằng việc áp dụng mô hình động lực học vào trường hợp cụ thể của xe chiến đấu bộ binh bơi tự hành Pandur II, kết quả tính toán đã cho thấy sự phù hợp của phương pháp tiếp cận, đồng thời cung cấp các cơ sở định lượng đáng tin cậy cho việc khai thác, huấn luyện và phát triển các dòng xe bơi tự hành trong tương lai, đáp ứng yêu cầu nhiệm vụ quân sự và phù hợp với điều kiện thực tế tại Việt Nam.

Tài liệu tham khảo

1. T. Perez, *Course keeping and roll stabilisation using rudder and fins*, London, Springer-Verlag, 2005.

2. Zilin Zhang, "Design and simulation research of anti-rolling tanks system", in *Proceedings of the International conference on computer technologies in physical and engineering applications ICCTPEA*, pp. 213-214, 2014. DOI: 10.1109/ICCTPEA.2014.6893366
3. Allan and J.F., "Stabilisation of ships by activated fins", *Transactions of the Royal Institution of Naval architects RINA 87*, pp. 123-159, 1945.
4. Alarcin, Fuat, "Rudder roll stabilization for fishing vessel using neural network approach", *Ocean engineering*, Vol.34(13), pp. 1811-1817, 2007. DOI: 10.1016/j.oceaneng.2007.03.001
5. D. W. Tank and J. J. Hopfield, "Simple neural optimization networks: An A/D converter, signal decision circuit, and a linear programming circuit", *IEEE Trans. Circuits Syst*, Vol. CAS-33, pp. 533-541, May 1986.
6. Hui Li, Chen Guo, "Adaptive fuzzy sliding mode controller design for ship fin stabilizer under rough sea conditions", in *Proceedings of IEEE international conference on information and automation ICIA*, pp. 566-571, 2014. DOI: 10.1109/ICInfA.2014.6932718
7. S. Sun and J. Hu, "Design of simplified fuzzy controller for ship fin stabilizer", in *Proceedings of the 33rd Chinese control conference*, pp. 4534-4538, 2014. DOI: 10.1109/ChiCC.2014.6895702
8. N. T. Dương, N. Q. Duy, "Ứng dụng Matlab trong điều khiển toàn phương tuyến tính bậc hai cho hệ thống vây giảm lắc tàu thủy có tính đến yếu tố ngoại cảnh", *Tạp chí Khoa học Công nghệ Hàng hải*, Số 72, tr. 5-10, 2021.
9. N. Q. Duy, L. Đ. Bình, "Nghiên cứu giảm lắc dọc tàu thủy ứng dụng phương pháp điều khiển toàn phương tuyến tính bậc hai", *Tạp chí Khoa học Công nghệ Hàng hải*, Số 72, tr. 22-26, 2022.
10. N. Q. Duy, "Thiết kế điều khiển dự đoán mô hình cho hệthống vây giảm lắc tàu thủy dựa trên mạng thần kinh phản hồi", *Tạp chí Khoa học Công nghệ Hàng hải*, Số 67, tr. 11-17, 2021.
11. D. Lee, S. Ko, J. Park *et al*, "An Experimental Analysis of Active Pitch Control for an Assault Amphibious Vehicle Considering Waterjet-Hydrofoil Interaction Effect", *Marine Science and Engineering*, Vol. 9(894), pp. 1-18, 2021. DOI: 10.3390/jmse9080894
12. N. Kornev, *Ship Dynamics in Waves*, Universität Rostock, Fakultät für Maschinenbau und Schiffstechnik, 2011, p. 13.
13. T. C. Nghị, *Sổ tay thiết kế tàu thủy*, Nxb Xây dựng, Hà Nội, 2011, tr. 124-126.
14. R. Bhattacharyya, *Dynamics of Marine Vehicles*. New York, NY, USA: Wiley, 1978, pp. 304 - 307.
15. Ministry of Defence of the Czech Republic, *Pandur II 8x8 Armoured Vehicle*, Ministry of Defence, Czech Republic. [Online]. Available: https://www.mo.gov.cz/scripts/detail.php?id=15806.
16. N. V. Tân, *Xe máy vượt sông*, Nxb Quân đội, Hà Nội, 2011.

Research on application of active fin stabilizers to reduce oscillatory motions of amphibious combat vehicles

Abstract: *When amphibious combat vehicles operate in seawater environments, roll and pitch oscillations can significantly reduce safety, affect payload capacity, limit the use of firepower, and negatively impact crew health. This issue is particularly critical for amphibious combat vehicles with open decks performing amphibious landing missions, where reducing roll and pitch oscillations plays an important role during water operations. With outstanding advantages, active fin stabilizers have become a widely used solution for reducing motions in marine vehicles. The effectiveness of this system mainly depends on its controller. However, this technology has*

not yet been applied to amphibious combat vehicles to mitigate oscillations during swimming. In this study, the oscillatory motions of an amphibious combat vehicle during water operation are analyzed, and a solution using an active fin stabilizer system with a Multi-Input Multi-Output (MIMO) controller is proposed. MIMO control is an effective method for processing multiple input–output channels simultaneously and can improve the stabilization efficiency of the system. By considering the effect of random waves, the MIMO controller is designed to provide an optimal control strategy for the fin stabilizer system. Simulation results in Matlab confirm the stability and effectiveness of the proposed control system.

Keywords: Oscillation, amphibious combat vehicle, fin stabilizer, MIMO.

NGHIÊN CỨU ĐỘNG LỰC HỌC BƠM PÍT TÔNG HƯỚNG TRỤC SỬ DỤNG MÔ HÌNH MẠNG NƠ RON

Nguyễn Văn Khánh[1], Đoàn Văn Tú[2], Trịnh Văn Hải[1*]

[1]Viện Cơ khí động lực, Học viện Kỹ thuật quân sự;

[2]Khoa Cơ khí, Đại học Giao thông vận tải

Email: hai.tv@lqdtu.edu.vn

Tóm tắt

Trên cơ sở nguyên lý cấu trúc và hoạt động điển hình của bơm pít tông hướng trục loại đĩa nghiêng, hệ phương trình vi phân mô tả các đặc tính động lực học của hệ thống bơm như đáp ứng của áp suất và lưu lượng làm việc sẽ được thiết lập. Kết quả giải số mô hình động lực học bơm pít tông tìm được bằng phương pháp số sẽ được sử dụng làm dữ liệu huấn luyện mạng nơ rôn. Quan hệ giữa các thông số kết cấu và vận hành với đáp ứng thông số làm việc của bơm sẽ được xấp xỉ bằng cấu trúc mạng nơ ron với khả năng dự báo chính xác cao. Kết quả nghiên cứu cung cấp cơ sở để nghiên cứu đáp ứng của bơm pít tông thông qua mô hình mạng nơ ron với thời gian tính toán nhanh phục vụ các bài toán thiết kế và tối ưu tổng thể trong thực tế.

Từ khóa: *Bơm pít tông; đáp ứng động lự học; khai triển Karhunen– Loève; mô hình thay thế; mạng nơ ron.*

1. Giới thiệu

Hệ thống thủy lực đóng vai trò nền tảng trong hoạt động của các hệ thống công nghiệp và máy móc di động, cho phép thực hiện các tác vụ yêu cầu công suất cao và độ chính xác, bao gồm nâng hạ, ép và vận chuyển các tải trọng lớn. Các hệ thống này được ứng dụng rộng rãi trong các lĩnh vực như xây dựng, nông nghiệp, hàng không vũ trụ và sản xuất công nghiệp, nơi mà những ưu điểm của chúng — như mật độ công suất vượt trội, khả năng điều chỉnh chính xác lực và mô-men xoắn, cũng như kiểm soát chính xác chuyển động và vị trí — có ý nghĩa then chốt. Việc tích hợp công nghệ thủy lực vào các máy móc hạng nặng tự động đã mang lại những bước tiến đáng kể trong quy trình công nghiệp, nâng cao hiệu quả vận hành và đảm bảo các tiêu chuẩn an toàn cao hơn [1, 2]. Bơm pít tông hướng trục là các thành phần rất quan trọng trong hệ thống thủy lực, có chức năng chuyển đổi năng lượng cơ học thành năng lượng dòng chảy của chất lỏng, với dải áp suất làm việc rộng tùy thuộc vào yêu cầu sử dụng thực tế [3, 4]. Việc xây dựng mô hình động lực học cho bơm piston hướng trục là điều cần thiết để đánh giá hành vi vận hành của chúng trong nhiều điều kiện làm việc khác nhau. Những mô hình này cho phép đặc trưng hóa các đặc tính hoạt động của bơm, bao gồm dao động áp suất, động lực dòng chảy và tương tác cơ học, từ đó cung cấp một đánh giá toàn diện về hiệu suất, hiệu quả và độ ổn định. Bằng cách mô tả sự tương tác phức tạp giữa các thành phần thủy lực và cơ khí, phương pháp này hỗ trợ tối ưu hóa thiết kế và điều khiển bơm, đảm bảo khả năng hoạt động tin cậy và hiệu quả trong các ứng dụng thực tế [3, 5-9].

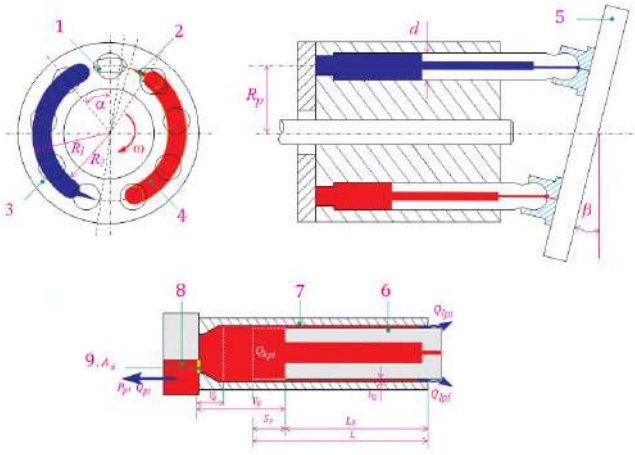
Các mô hình động lực học của bơm piston hướng trục đã được nghiên cứu một cách rộng rãi [5, 10-12]. W. Latas và J. Stojek [11] đã phát triển một mô hình bơm động lực học với 13 bậc tự do (13-DOF), có xét đến mối quan hệ giữa mức độ mài mòn của các thành phần và các thông số độ cứng cũng như hệ số giảm chấn nhớt của các khớp nối. Quỹ đạo pha được tính toán tại các điểm quan sát cụ thể cung cấp thông tin phục vụ cho việc đánh giá tình trạng mài mòn của các thành phần. Tương tự, ảnh hưởng của các điều kiện làm việc (tức là áp suất đầu ra, tốc độ quay và góc dịch chuyển) cũng như các tham số cấu trúc (tức là hệ số độ cứng và giảm chấn

tại các vùng tiếp xúc khác nhau như piston–ống xy lanh, xy lanh–đĩa van, và đế trượt–đĩa nghiêng) đến các phản ứng động lực học (tức là dao động biên độ–tần số, biểu đồ quỹ đạo pha) cũng được nghiên cứu với các mo hình động lực học gồm 8 [5] và 19 bậc tự do [10]. Để giải các mô hình động lực học của hệ thống bơm, nhiều phương pháp đã được phát triển tùy thuộc vào mức độ phức tạp của mô hình và mục tiêu nghiên cứu: mô hình CFD [13], phương pháp số (i.e., phương pháp Runge-Kutta [5], phương pháp Newmark-β [10]), công cụ tính toán số (i.e., Matlab-Simulink [7, 9] AMESim [14]), hoặc các phương pháp lai/kết hợp [10, 15]. Nhờ vào các công cụ và kỹ thuật tính toán này, một số hành vi động lực học của bơm piston hướng trục có thể được đánh giá, trong đó bao gồm lưu lượng dòng chảy, dao động áp suất, và đặc tính mấp mô của lưu lượng [13, 16] và áp suất làm việc [14, 15]. Việc đảm bảo hoạt động đáng tin cậy và kéo dài tuổi thọ của bơm piston thủy lực đòi hỏi các chiến lược giám sát hiệu quả để giảm thiểu nguy cơ hỏng hóc bất ngờ trong quá trình sử dụng và vận hành. Nhờ vào các mô hình toán học và động lực học đã được phát triển ở trên, các phương pháp chẩn đoán tiên tiến thường bao gồm việc thu thập dữ liệu theo thời gian thực về các tham số vận hành chính, bao gồm áp suất, nhiệt độ, lưu lượng và sự rò rỉ [3, 17-19]. Các chỉ báo quan trọng này giúp phát hiện sớm tình trạng mài mòn và các sự cố tiềm ẩn ở các thành phần thiết yếu như piston và khối xy lanh. Bằng cách tích hợp công nghệ cảm biến với phần mềm chẩn đoán thông minh, đội ngũ bảo trì có thể chủ động nhận diện sự sai lệch so với hiệu suất bình thường, từ đó thực hiện các biện pháp sửa chữa kịp thời. Hơn nữa, các khung bảo trì dự báo dựa trên các kỹ thuật giám sát này cải thiện độ tin cậy của hệ thống bằng cách giảm thiểu thời gian chết không kế hoạch, tối ưu hóa hiệu quả vận hành và kéo dài tuổi thọ của bơm thủy lực [20-22]. Có thể khẳng định rằng các tham số hệ thống và đặc tính cấu trúc của các thành phần thủy lực thường không xác định trong thực tế. Do đó, vấn đề nghiên cứu động lực học của hệ thống cần phải được giải quyết [23-27]. Trong bối cảnh các sự không chắc chắn không khớp, các thách thức trong việc điều khiển hệ thống thủy lực điều khiển bơm với tốc độ thay đổi đã được giải quyết [27]. Nghiên cứu này xem xét các sự không chắc chắn liên quan đến việc đo lường hiệu suất bơm, với mục tiêu thiết lập các khoảng tin cậy phù hợp và đánh giá chính xác hiệu suất cũng như hành vi động lực học của bơm [24]. Các khung phương pháp định lượng sự không chắc chắn cũng được mở rộng để nghiên cứu các yếu tố tương tự, chẳng hạn như các bơm ly tâm [23, 25] và xy lanh thuỷ lực [26]. Trong chương này, một mô hình tính toán để phân tích đặc tính làm việc của bơm piston hướng trục sử dụng mô hình thay thế bậc giảm dựa trên khai triển KL sẽ được trình bày. Để đạt được mục tiêu này, phần còn lại của bài báo được cấu trúc như sau. Trong phần 2.2, chúng tôi giới thiệu mô hình động lực học của bơm với các tham số làm việc là áp suất piston và áp suất đầu ra của bơm. Phần 2.3 trình bày phương pháp RK và khai triển KL, được áp dụng cho bài toán động lực học của bơm. Đặc tính áp suất đầu ra và lưu lượng của bơm được thảo luận trong phần cuối.

2. Xây dựng mô hình động lực học bơm pít tông hướng trục

Trên cơ sở trình bày tóm lược về nguyên lý cấu trúc và hoạt động điển hình của bơm pít tông hướng trục loại đĩa nghiêng, hệ phương trình vi phân sẽ được thiết lập để mô tả các đặc tính động lực học của hệ thống bơm như đáp ứng của áp suất và lưu lượng làm việc.

2.1. Mô tả nguyên lý kết cấu và làm việc của bơm pít tông



Hình 1. Đặc điểm kết cấu và chi tiết các tham số của bơm pít tông hướng trục kiểu đĩa nghiêng: 1-Cửa xy lanh; 2-Rãnh chữ V; 3-Đĩa van; 4-Cửa hút; 5- Đĩa nghiêng; 6- Pít-tông; 7-Khối xy lanh; 8-Cửa xả; 9-Diện tích cửa xả.

Nguyên lý hoạt động và các tham số hình học/vận hành của một bơm pít tông hướng trục được mô tả trên Hình 1. bơm này có một số thành phần chính như đĩa nghiêng, piston, khối xy lanh và đĩa van (xem Hình 1a). Khi làm việc, động cơ điện hoặc động cơ đốt trong cung cấp chuyển động quay cho trục bơm, từ đó làm quay khối xy lanh và dẫn đến sự chuyển động của các pít-tông trong bơm. Đĩa nghiêng , được đặt nghiêng một góc cụ thể (ký hiệu β) so với khối xy lanh, đóng vai trò then chốt trong việc tạo ra các chu trình hút và đẩy luân phiên cho từng pít-tông khi nó quay. Góc nghiêng này khiến mỗi pít-tông thực hiện chuyển động tịnh tiến qua lại, hút dầu thủy lực vào trong giai đoạn hút và đẩy chất lỏng ra trong giai đoạn nén. Đĩa van (được bố trí tại đáy các pít-tông, Hình 1b) điều hướng dòng dầu thủy lực từ cổng hút vào các buồng pít-tông và hỗ trợ việc thoát chất lỏng đã được nén qua cổng đẩy, từ đó đảm bảo dòng chất lỏng được kiểm soát trong hệ thống bơm. Chi tiết các ký hiệu được trình bày trong Hình 1 và phần mô tả sẽ được trình bày trong phần 4.1. Cần lưu ý rằng trong các hệ thống bơm hiện đại, cấu trúc bơm nêu trên sẽ được tích hợp với hệ thống van điều khiển để kiểm soát lưu lượng, áp suất làm việc, cũng như thực hiện các chức năng tiên tiến khác của hệ thống liên quan đến chất lượng truyền động và điều khiển, đồng thời bao gồm các tính năng tự bảo vệ và tự chẩn đoán.

Một trong những vấn đề cần được giải quyết trong quá trình thiết kế và vận hành các loại bơm pít-tông hướng trục là nghiên cứu mối quan hệ giữa các đặc tính động lực học (tức là áp suất và lưu lượng) ở từng khoang pít-tông đơn lẻ và khoang xả của bơm, có tính đến ảnh hưởng của các thông số cấu trúc/vận hành cũng như các đặc tính của dầu thủy lực.

2.2. Mô hình động lực học bơm

Các công thức toán học biểu diễn mô hình động lực học của bơm có thể được trình bày thông qua các phương trình biểu thị áp suất tức thời của pít-tông và áp suất đầu ra của bơm. Ở trong khoang pít-tông của bơm, áp suất tức thời của chúng có thể được biểu diễn bằng [21]:

$$\dot{P}_{pi} = \frac{B}{V_0 - A_p S_{pi}} \left(Q_{kpi} - Q_{pi} - Q_{lpi} \right), \tag{1}$$

trong đó B là mô-đun đàn hồi thể tích của dầu thủy lực, V_0 là thể tích ban đầu của khoang pít tông, R_p diện tích mặt cắt của pít tông. Hành trình S_{pi} của từng pít tông được xác định theo:

$$S_{pi} = R_p \tan\beta(1 - \cos\Omega_i), \tag{2}$$

Với R_p là bán kính phân bố của pít tông trên block xy lanh, $\Omega_i=\omega t-(i-1)\alpha$ là vị trí góc của xy lanh thứ i, α is góc trễ pha giữa hai pít tông kề nhau, β là góc nghiêng của đĩa nghiêng, ω là tốc độ quay của trục bơm.

Ở thành phần bên phải của công thức (1) bao gồm lưu lượng lý thuyết Q_{kpi}, lưu lượng xả Q_{pi}, và lưu lượn rò rỉ Q_{lpi} của pít tông thứ i (xem Hình 1c), các loại lưu lượng này được xác định theo các biểu thức phía sau:

$$Q_{kpi}=\omega\frac{\pi d^2 R_p}{4}\tan\beta\sin\Omega_i, \tag{3}$$

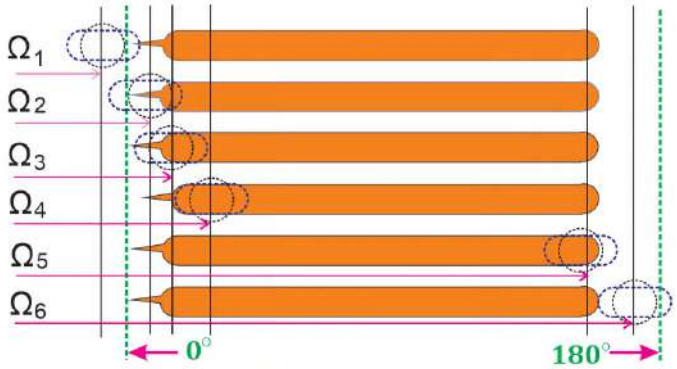
$$Q_{pi}=C_{d1}A_{di}\sqrt{\frac{2|P_{pi}-P_s|}{\rho}}\operatorname{sign}(P_{pi}-P_s), \tag{4}$$

$$Q_{lpi}=\frac{\pi d h_{gi}^3}{12\mu(L_0+S_{pi})}(P_{pi}-P_c), \tag{5}$$

trong đó, d là đường kính pít tông, h_{gi} là khe hở hướng kính giữa pít tông và khối xy lanh, L_0 là chiều dài ban đầu cảu vùng tiếp xúc giữa pít tông và xy lanh, P_c áp suất khoang vỏ bơm, C_{d1} là hệ số lưu lượng. Các ký hiệu μ và ρ biểu diễn độ nhớt tuyệt đối và khối lượng riêng của chất lỏng. Diện tích cửa xả A_{di} thể hiện diện tích thông giữa khoang xả và đĩa phân phối, theo tài liệu [9] diện tích này được cho theo vị trí của pít tông như sau (lưu ý rằng diện tích xả này phụ thuộc vào kết cấu của đĩa phân phối, nội dung này được phân tích tường minh trong tài liệu [15], chi tiết tại Hình 2.2):

$$A_{di}=\begin{cases}\dfrac{(\Omega_i+13)^2\pi R_p}{360\times 22}(R_1-R_2), & \Omega_1<\Omega_i<\Omega_2\\ \dfrac{(\Omega_i-9)\pi}{360}(R_1^2-R_2^2)+\dfrac{22\pi R_p}{360}(R_1-R_2), & \Omega_2<\Omega_i<\Omega_3\\ \dfrac{(\Omega_i-9)\pi}{360}(R_1^2-R_2^2)+\left[\dfrac{22^2-(\Omega_i-13)^2}{360\times 22}\right]\pi R_p(R_1-R_2), & \Omega_3<\Omega_i<\Omega_4\\ \dfrac{30\pi}{360}(R_1^2-R_2^2), & \Omega_4<\Omega_i<\Omega_5\\ \dfrac{(\Omega_i-141)\pi}{360}(R_1^2-R_2^2), & \Omega_5<\Omega_i<\Omega_6\\ 0, & \Omega_6<\Omega_i<2\pi+\Omega_1\end{cases} \tag{6}$$

với R_1 và R_2 là bán kính ngoài và bán kính trong của cửa xả trên đĩa phân phối.



Hình 2. Diện tích thông của từng pít tông tới khoang bơm

Áp suất làm việc của bơm P_s được tính thông qua lưu lượng xả Q_p và lưu lượng làm việc của bơm Q_s như sau:

$$\dot{P}_S = \frac{B}{V_C}(Q_p - Q_S), \tag{7}$$

với

$$Q_p = \sum_{1}^{m} Q_{pi}, \tag{8}$$

và

$$Q_S = C_{d2} A_v \sqrt{\frac{2P_S}{\rho}}, \tag{9}$$

trong đó V_c là thể tích khoang xả, C_{d2} là hệ số lưu lượng, A_v là diện tích mở của van tiết lưu.

Gộp hệ phương trình (1) và phương trình (7), chúng ta nhận được $(m+1)$ phương trình vi phân cấp một ứng với m tham số áp suất của từng pít tông và một tham số áp suất của bơm.

3. Xây dựng mô hình mạng nơ ron nghiên cứu động lự học bơm

Ở mục này, phương pháp số để giải hệ phương trình vi phân hành vi động lực học của bơm sẽ được trình bày. Ngoài ra, kỹ thuật mô hình thay thế dựa trên KL sẽ được giới thiệu để biểu diễn các phản ứng của bơm theo thời gian sử dụng một tập hợp hữu hạn các biến giảm bậc.

3.1. Phương pháp Runge-Kutta

Như đã được thảo luận trước đó, mô hình toán học mô tả quá trình hoạt động của bơm được xây dựng dưới dạng một hệ phương trình vi phân thường, mô tả tập hợp sự biến thiên của các biến áp suất trong hệ thống:

$$\frac{dP}{dt} = \Gamma(P,t), \text{ với } \quad P = \begin{bmatrix} \{P_{pi}\}_{m\times 1} \\ P_S \end{bmatrix} \tag{10}$$

Sơ đồ Runge–Kutta bậc bốn (RK4) được sử dụng để giải hệ phương trình vi phân được trình bày trong công thức (10) với khoảng thời gian $t \in [t_0, t_f]$ và giá trị áp suất tại thời điểm ban đầu $P(t_0) = P_0$. Thời gian tính toán được chia thành N khoảng nhỏ với giá trị $h = t_f/N$. Ở bước thời gian $t_{n+1} = t_0 + nh$, giá trị biến $P(t_{n+1})$ được tính toán từ giá trị $P(t_n)$ ở thời điểm $t_n = t_0 + (n-1)h$ như sau:

$$P_{n+1} = P_n + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4), \tag{11}$$

các hệ số trong phương trình trên được tính:

$$\begin{aligned} &k_1 = \Gamma(t,P); k_2 = \Gamma\left(t + \frac{h}{2}, P + \frac{k_1}{2}h\right); \\ &k_3 = \Gamma\left(t + \frac{h}{2}, P + \frac{k_2}{2}h\right); k_4 = \Gamma(t+h, P + k_3 h). \end{aligned} \tag{12}$$

3.2. Phương pháp khai triển Karhunen– Loève

Phần này tập trung xây dựng một mô hình thay thế thể hiện mối quan hệ giữa các tham số hệ thống—bao gồm cả các yếu tố hình học và vận hành—và đáp ứng động trong theo thời

gian của hệ thống $t \in T$. Giả sử $\boldsymbol{x} \in R^D$ là véc tơ các biến trogn không gian thiết kế/khai thác của hệ thống, và $\boldsymbol{x} \mapsto J(t;\boldsymbol{x})$ thể hiện là đáp ứng đầu ra của hệ thống mà cần phải quan tâm. Đáp ứng hệ thống $J(t;\boldsymbol{x})$ được viết thu gọn ở dạng $J(t)$. Theo [22, 23], đáp ứng cần quan tâm có thể được biểu diễn qua chuỗi KL sau:

$$J(t) = \underline{J}(t) + \sum_{i=1}^{\infty} \sqrt{\lambda_i}\eta_i\varphi_i(t), \tag{13}$$

với, $t \mapsto J(t)$ là hàm trung bình của đáp ứng động, được xác định theo $\underline{J}(t) = E\{J(t)\}$ với mọi $t \in T$. Các bộ tham số $(\lambda_i,\varphi_i)_{i\geq 1}$ với các giá trị riêng và hàm riêng của toán tử hiệp phương sai $C(t,t')$ thoả mãn phương trình tích phân sau:

$$\int_{\mathbb{T}} C(t,t')\varphi_i(t')dt' = \lambda_i\varphi_i(t), \tag{14}$$

các biến hạ bậc $(\eta_i)_{i\geq 1}$ là tập trung, không tương quan lẫn nhau và có phương sai đơn vị. Chúng được ước tính bởi,

$$\eta_i = \frac{1}{\sqrt{\lambda_i}}\langle J(t) - \underline{J}(t), \varphi_i(t)\rangle. \tag{15}$$

Trong công thức trên, toán tử $\langle .,. \rangle$ biểu diễn tích vô hướng và được tính:

$$\langle f, g\rangle = \int_{\mathbb{T}} f(t)g(t)dt. \tag{16}$$

Chuỗi số trong phương trình (15), khi chặn ở bậc p, được viết lại:

$$J(t) = \underline{J}(t) + \sum_{i=1}^{p} \sqrt{\lambda_i}\eta_i\varphi_i(t), \tag{17}$$

với bậc chặn p được tính qua phân tích mức độ hội tụ của chuỗi $\lim_{p\to\infty} J_p = J$.

Để định lượng các đặc điểm hội tụ của chuỗi chặn KL, hai hàm đánh giá mức độ hội tụ sau được sử dụng. Đầu tiên, số lượng thực hiện cần thiết để đạt được sự hội tụ được đánh giá bằng cách đánh giá hàm $\mathrm{Cov}(N_s)$:

$$\mathrm{Cov}(N_S) = \left\| \left[\tilde{C}(N_S) \right] \right\|, \tag{18}$$

với $\left[\tilde{C}\right]$ biểu diễn ma trận hiệp phương sai của quá trình rời rạc.

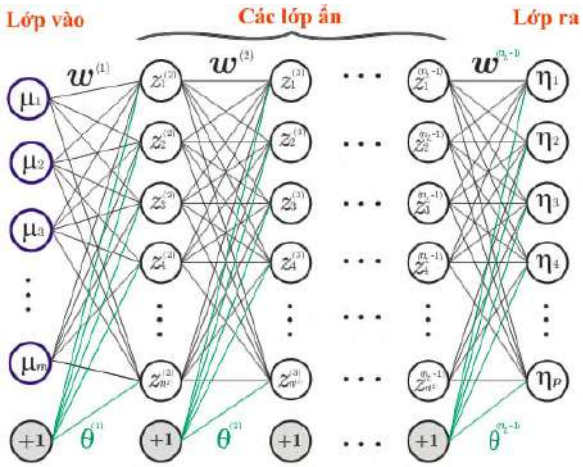
Để xác định bậc chặn cho ước lượng thống kê, thuộc tính hội tụ thứ hai được phân tích bởi hàm lỗi $\mathrm{Err}(p)$ được định nghĩa là,

$$\mathrm{Err}(p) = 1 - \frac{\sum_{i=1}^{p}\lambda_i}{\mathrm{tr}\left(\left[\tilde{C}\right]\right)}. \tag{19}$$

3.3. Cấu trúc mạng nơ ron

Hình 3 trình bày cấu trúc mạng này gồm có 4 lớp: lớp đầu vào, một số lớp ẩn và lớp đầu ra. Số lượng nơron trong lớp 1 được ký hiệu là $n^{(l)}$. Ở đây, $n^{(1)} = m$ and $n^{(nL)} = p$), trong khi số lượng các nút của lớp ẩn (tức là $n^{(2)}$ và $n^{(3)}$) được chọn tự do để có khả năng dự đoán tốt. Nhận thấy rằng số lượng nơron dạng như trên loại trừ các sai lệch hoặc các nơron cuối cùng đươc ký hiệu là (+1). Gọi $\mu^{(i)} = h_g^{(i)}$ là véc tơ đầu vào của điểm dữ liệu thứ i với $h_g=[h_{gj}]_{j=1x9}$ tương ứng

mức độ mòn của các quả pít tông. Mục tiêu chính là tìm một cấu trúc mạng trong đó mối quan hệ đầu vào/đầu ra được xây dựng tốt hơn dựa trên tập dữ liệu, $T = [(\mu^{(i)}, \eta^{(i)})| \, i \in [1, ..., N_n]]$, với N_n là kích thước mẫu.



Hình 3. Cấu trúc mạng nơ-ron nghiên cứu hệ động lực bơm pít tông. Số lượng nơ-ron ở lớp đầu vào và lớp đầu ra lần lượt tương ứng với m = 9 tham số mòn của các pít tôn và bậc của khai triển KL rút gọn.

Trước hết, trong một lớp cho trước, đầu ra của nơ-ron thứ j được xác định bởi:

$$z_j^{(l+1)} = f\left(\sum_{i=1}^{n^{(l)}} w_{ij}^{(l)} z_j^{(l)} + \theta_j^{(l)}\right), \quad \text{với } l = 1,...,n_L - 1 \text{ và } z_i^{(l)} = \mu_i \tag{20}$$

Trong đó, $f(\cdot)$ là hàm kích hoạt hay hàm truyền (Trong nghiên cứu này, hàm sigmoid được sử dụng: $f(v) = 1/[1 + \exp(-v)]$. Ngoài ra, có thể sử dụng các hàm kích hoạt khác, chẳng hạn như hàm tang hyperbol). $w_{ij}^{(l)} z_j^{(l)}$ là tổng có trọng số của các tín hiệu đầu vào và $\theta_j^{(l)}$ là độ lệch. Vì vậy, sự kết nối giữa hai lớp có thể được biểu diễn bằng một ma trận trọng số $w^{(l)} = [w_{lk}^{(l)}]_{n^{(l)} \times n^{(l+1)}}$ và một vectơ độ lệch $\theta(l) = [\theta_k^{(l)}]_{1 \times n^{(l+1)}}$.

Do đó với dữ liệu đầu vào và một bộ trọng số $w = (w^{(1)}, ..., w^{(nL-1)})$ và độ lệch $\theta = (\theta^{(1)}, ..., \theta^{(n_L-1)})$, các dự đoán của mô hình NN được ký hiệu là $\hat{\eta}_{w,\theta}\left(\mu^{(i)}\right)$, được tạo ra. Mục tiêu là tìm ra các trọng số tốt nhất ω và độ lệch θ tốt nhất để giảm thiểu hàm dự đoán và phản hồi tham chiếu. Hàm tổn thất L_k - có thể được tính như sau:

$$L(w,\theta) = \frac{1}{N} \sum_{i=1}^{N_n} \left\| \hat{\eta}_{w,\theta}\left(\mu^{(i)}\right) - \eta^{(i)} \right\|^k, k = 1,2 \tag{21}$$

Một bước lặp mới của quá trình huấn luyện được thực hiện nhằm tối thiểu hóa hàm tổn thất $L(w,\theta)$ hàm này so sánh các đầu ra dự đoán với các giá trị mục tiêu. Phương pháp tối ưu hóa Levenberg- Marquardt được sử dụng để huấn luyện các trọng số và độ lệch được cập nhật thông qua phương thức truyền dữ liệu ngược [28, 29],

$$w^{(k+1)} = w^{(k)} - \left[J^T J + \mu I \right]^{-1} J^T e \tag{22}$$

Trong đó tập hợp $\boldsymbol{w}$ chứa tất cả các trọng số là độ lệch đã đề cập trước đó, μ là tốc độ học điều chỉnh kích thước và $\boldsymbol{I}$ là ma trận đơn vị. $\boldsymbol{J}$ là ma trận Jacobian, bao gồm tất cả các đạo hàm bậc nhất của các sai số theo các trọng số và các hệ số lệch. Vectơ sai số, được ký hiệu là $\boldsymbol{e}$, được tính tại mỗi bước lặp huấn luyện thứ k.

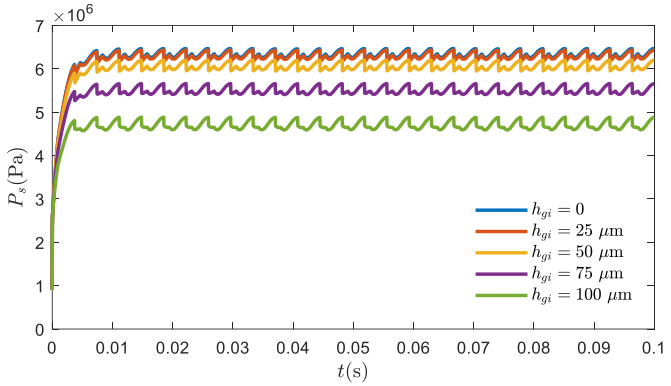
Trong nghiên cứu này, việc cấu trúc và huấn luyện mô hình NN được thực hiện bằng Neural Network Toolbox của MATLAB [30]. Tập dữ liệu của mô hình NN thường bao gồm các dữ liệu thô; do đó, bước chuẩn hóa là cần thiết để đảm bảo dữ liệu nằm trong cùng một miền giá trị của hàm truyền. Đối với trường hợp sử dụng hàm sigmoid chuẩn (tức là dữ liệu nằm trong khoảng [0, 1]), các giá trị nhỏ nhất và lớn nhất của dữ liệu thô được sử dụng để chuẩn hóa chúng như trong [31].

$$\bar{x}_i = \frac{x_i - x_{\min,i}}{x_{\max,i} - x_{\min,i}} \tag{23}$$

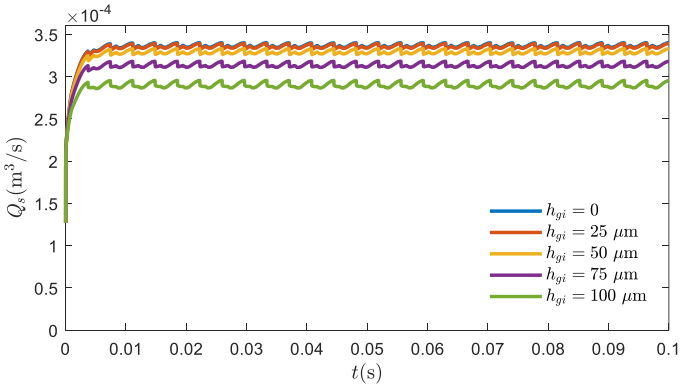
4. Kết quả tính toán và nhận xét

Cấu hình bơm pít tông hướng trục được chọn để khảo sát gồm các tham số chính sau [9, 21]: A_p = 8.343E-05 m^2, A_v = 4.457E-06 m^2, B = 8.547E+08 Pa, C_{d1} = 0.675, C_{d2} = 0.610, d = 1.031E-02 m, L_0 = 0.0145 m, P_c = 1E+05 Pa, R = 5.162E-03 m, R_1 = 0.0238 m, R_2 = 0.0204 m, R_p =0.0221 m, m = 9, V_0 = 1.543E-06 m^3, V_c = 3.156E-05 m^3, V_d = 2.0540E- 07 m^3, α = 0.698, β = 0.335, ω = 188.4 1/s, μ =5.855E-02 Ns/m^2, ρ = 860 kg/m^3. Giá trị các góc kết cấu của đĩa chia được cho bởi: $\Omega_1 = -13^0$, $\Omega_2 = 9^0$, $\Omega_3 = 17^0$, $\Omega_4 = 39^0$, $\Omega_5 = 141^0$ và $\Omega_6 = 171^0$. Giá trị áp suất ban đầu được chọn như sau $P_{pi}(t_0)$ = 1E+07 Pa và $P_s(t_0)$ = 9E+06 Pa. Giá trị h_{gi} thay đổi trong khoảng 0 đến 100 µm.

Các tính toán được thực hiện trên máy tính cá nhân có cấu hình Intel(R) Core(TM) i7-4500U CPU @ 1.80 GHz (2.40GHz) và RAM 8.00 GB. Với t_f = 0.1s và bước thời gian 1µs, thời gian tính toán khoảng gần 60s.

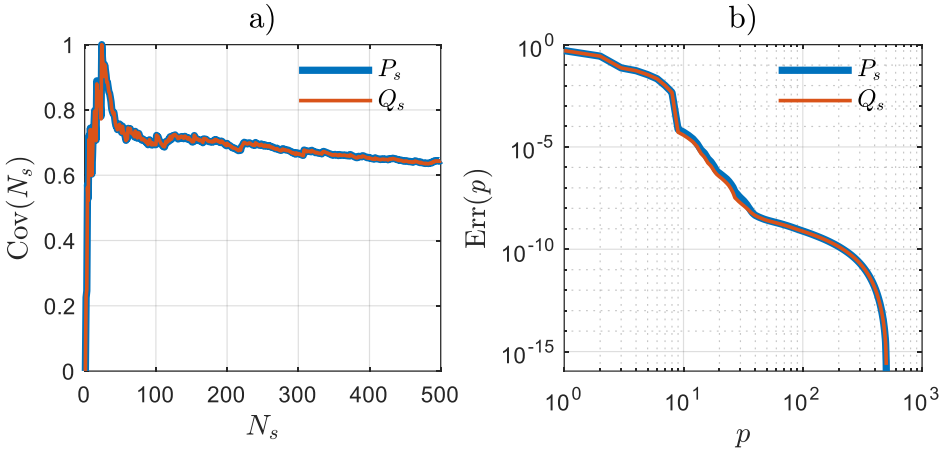


Hình 4. Kết quả tính toán đặc tính áp suất của bơm với các độ mòn pít tông đồng đều hgi = 0/25/50/75/100 µm.



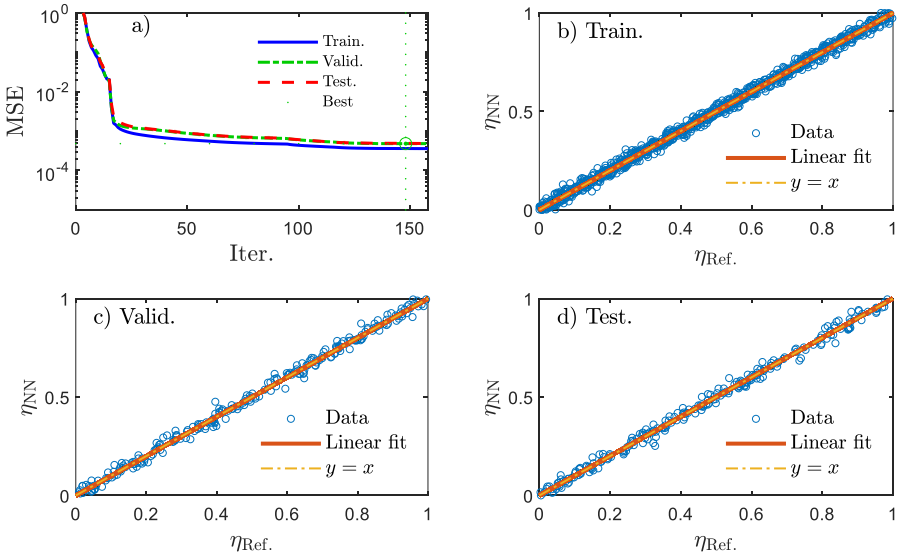
Hình 5. Kết quả tính toán đặc tính lưu lượng của bơm với các độ mòn pít tông đồng đều hgi = 0/25/50/75/100 µm.

Trên hai Hình 5 tương ứng thể hiện đặc tính áp suất (P_s) và lưu lượng (Q_s) của bơm với các độ mòn đồng đều khác nhau h_{gi} = 0/25/50/75/100 µm. Kết quả cho thấy rằng, các lưu lượng và áp suất làm việc của bơm đều có xu hướng bị giảm khi mức độ mài mòn tăng. Chu kỳ thay đổi của các đại lượng ứng với thời gian để bơm quay được đủ góc trễ pha α giữa 2 quả pít tông kề nhau. Cụ thể, khi độ mòn tăng từ 0 đến 100 µm, làm giảm lưu lượng giá trị áp suất từ 6.33E+6 Pa xuống 4.71E+6 Pa (ứng với 25.70 %) và lưu lượng giảm từ 3.37E-4 m^3/s xuống 2.90E-4 m^3/s (ứng với 23.80 %).



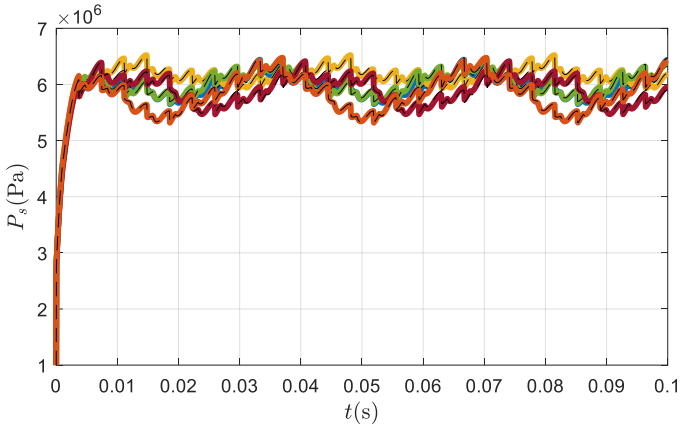
Hình 6. Kết quả tính toán đặc tính hội tụ của chuỗi KL.

Hình 6 thể hiện đặc tính hội tụ của chuỗi KL biểu diễn các đặc tính liên tục của áp suất và lưu lượng theo thời gian. Trên Hình 6a cho thấy giá trị N_s = 500 đủ lớn để tính toán các tham số của chuỗi KL. Ở Hình 6b, cho thấy rõ việc rõ việc lựa chọn bậc của chuỗi p để đạt được mức độ hội tụ. Với bậc p =10, ta có thể nhận được gia trị chỉ số Err(p) nhỏ hơn 10^{-4}. Kết quả cũng chỉ ra rằng cả hai đặc tính áp suất và lưu lượng có các chuỗi KL khá tương đồng với nhau.

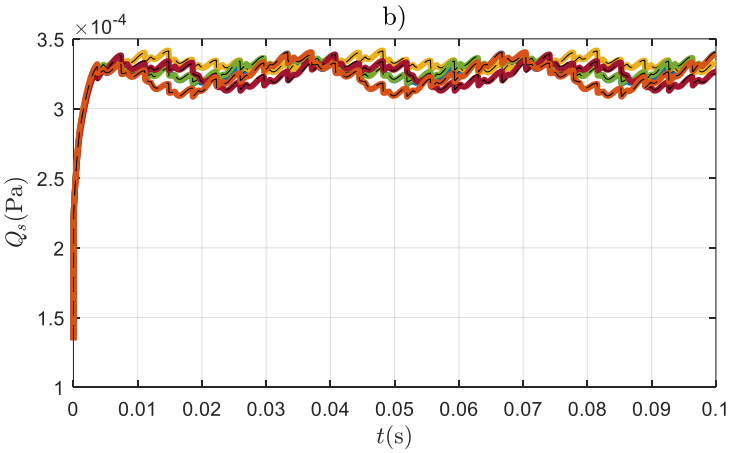


Hình 7. Kết quả đặc tính của mô hình NN sử dụng để dự báo đặc tính động lực học bơm pít tông

Mô hình mạng NN dự báo đặc tính áp suất và lưu lượng của bơm được trình bày ở Hình 7. Kết quả cho thấy các biến rời rạc của chuỗi KL mô tả đặc tính bơm được dự báo khá tương đồng với giá trị tham chiếu. Sai số MSE của mạng nhỏ hơn 0.01 ở bước tính toán khoảng 148 (Hình 7a). Trên các Hình 7b-c khẳng định giá trị tương đồng của mẫu tham chiếu và mẫu dự báo.



Hình 8. So sánh giá trị của các đặc tính áp suất bơm sử dụng mạng NN và giá trị tham chiếu.



Hình 9. So sánh giá trị của các đặc tính lưu lượng sử dụng mạng NN và giá trị tham chiếu.

Hình 8 và 9 cho thấy đặc tính áp suất và lưu lượng của bơm có thể được dự báo tốt bằng mô hình NN. Kết quả được trình bày với 5 cấu hình bơm khác nhau và các cấu hình này không nằm trong dữ liệu huấn luyện. Nét đậm thể hiện giá trị tham chiếu, trong khi nét mãnh đứt thể hiện giá trị dự báo bằng mô hình mạng. Thêm vào đó, kết quả cho thấy trạng thái mòn không đều giữa các quả pít tông cũng dẫn đến đặc tính làm việc của bơm có sự thay đổi rõ rệt ngoài chu kỳ biến đổi ứng với góc α thì xuất hiện thêm chu kỳ của đặc tính ứng với 1 vòng quay của bơm.

5. Kết luận

Bài báo đã xây dựng thành công mô hình động lực học của bơm pít tông hướng trục kiểu đĩa nghiêng trên cơ sở thiết lập hệ phương trình vi phân mô tả biến thiên áp suất trong các khoang pít tông và khoang xả của bơm. Hệ phương trình này đã được giải bằng phương pháp Runge–Kutta bậc bốn nhằm thu được đáp ứng áp suất và lưu lượng theo thời gian với độ chính xác cao.

Trên cơ sở dữ liệu mô phỏng thu được, phương pháp khai triển Karhunen–Loève đã được áp dụng để giảm bậc biểu diễn đáp ứng động theo thời gian thành một tập hữu hạn các biến rút gọn có ý nghĩa thống kê. Việc giảm bậc này cho phép chuyển bài toán phụ thuộc thời gian phức tạp thành bài toán ánh xạ giữa các tham số thiết kế/vận hành và các hệ số đặc trưng của đáp ứng.

Mô hình mạng nơ-ron truyền thẳng nhiều lớp đã được xây dựng và huấn luyện bằng thuật toán Levenberg–Marquardt để xấp xỉ quan hệ giữa các tham số hình học, thông số vận hành và các biến rút gọn KL. Kết quả cho thấy mô hình thay thế có khả năng dự báo chính xác đáp ứng áp suất và lưu lượng của bơm với thời gian tính toán rất nhỏ so với việc giải trực tiếp hệ phương trình vi phân.

Cách tiếp cận kết hợp giữa mô hình vật lý, phương pháp số, kỹ thuật giảm bậc và mạng nơ-ron đã chứng minh hiệu quả trong việc xây dựng mô hình dự báo nhanh cho hệ thống bơm pít tông hướng trục. Phương pháp này có tiềm năng ứng dụng trong các bài toán tối ưu hóa thiết kế, phân tích độ nhạy, định lượng bất định và phát triển mô hình song sinh số (digital twin) cho hệ thống thủy lực.

Tài liệu tham khảo

1. Manring, N.D. and R.C. Fales, *Hydraulic control systems*. 2019: John Wiley & Sons.
2. Dransfield, P., *Hydraulic control systems—design and analysis of their dynamics*. 1981: Springer.
3. Shengrong, G., et al., *Hydraulic piston pump in civil aircraft: Current status, future directions and critical technologies.* Chinese Journal of Aeronautics, 2020. **33**(1): p. 16-30.
4. Bergada, J.M., S. Kumar, and J. Watton, *Axial piston pumps, new trends and development.* Fluid Dynamics, Mechanical Applications Role in Engineering, 2012: p. 1-158.
5. Ying, P., et al., *Dynamic modeling and vibration characteristics of multibody system in axial piston pump.* Alexandria Engineering Journal, 2023. **62**: p. 523-540.
6. Yang, Y., et al., *Current status and applications for hydraulic pump fault diagnosis: A review.* Sensors, 2022. **22**(24): p. 9714.
7. Babikir, H.A., et al., *Noise prediction of axial piston pump based on different valve materials using a modified artificial neural network model.* Alexandria Engineering Journal, 2019. **58**(3): p. 1077-1087.
8. Srivyas, P.D., S. Singh, and B. Singh, *Study of various maintenance approaches types of failure and failure detection techniques used in hydraulic pumps: a review.* Industrial Engineering Journal, 2017.
9. Li, Z.E., *Condition monitoring of axial piston pump*. 2005.
10. Ye, S., et al., *A theoretical dynamic model to study the vibration response characteristics of an axial piston pump.* 2021. **150**: p. 107237.
11. Łatas, W. and J. Stojek, *Dynamic model of axial piston swash-plate pump for diagnostics of wear in elements.* Archive of Mechanical Engineering, 2011. **58**(2): p. 135-155.
12. Kavanagh, G.P., *The dynamic modelling of an axial piston hydraulic pump*. 1987, University of Saskatchewan.
13. Zhang, B., et al., *Analysis of the flow dynamics characteristics of an axial piston pump based on the computational fluid dynamics method.* 2017. **11**(1): p. 86-95.
14. Cho, I.-S. and J. Jung, *A study on the pressure ripple characteristics in a bent-axis type oil hydraulic piston pump.* Journal of Mechanical Science and Technology, 2013. **27**: p. 3713-3719.
15. Milind, T. and M.J.P.E. Mitra, *A study on the dynamics and vibration behavior of an axial piston pump using combined MBD/FE approach.* 2016. **144**: p. 452-460.
16. Xu, B., et al., *A new design method for the transition region of the valve plate for an axial piston pump.* Journal of Zhejiang University-SCIENCE A, 2015. **16**(3): p. 229-240.
17. Paulmann, G. and G. Mkadara, *Condition Monitoring of hydraulic pumps-lessons learnt*. 2018: Universitätsbibliothek der RWTH Aachen.
18. Helwig, N., E. Pignanelli, and A. Schütze. *Condition monitoring of a complex hydraulic system using multivariate statistics*. in *2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings*. 2015. IEEE.
19. Byington, C.S., et al. *In-line health monitoring system for hydraulic pumps and motors*. in *2003 IEEE Aerospace Conference Proceedings (Cat. No. 03TH8652)*. 2003. IEEE.
20. Wu, F., et al., *The remaining useful life prediction method of a hydraulic pump under unknown degradation model with limited data.* Sensors, 2023. **23**(13): p. 5931.
21. Wang, L., et al., *Fault diagnosis and predictive maintenance for hydraulic system based on digital twin model.* AIP Advances, 2022. **12**(6).

22. Rivera, D.L., et al., *Towards a predictive maintenance system of a hydraulic pump.* IFAC-PapersOnLine, 2018. **51**(11): p. 447-452.
23. Feng, H., et al., *Uncertainty analysis using the thermodynamic method of pump efficiency testing.* Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 2004. **218**(5): p. 543-555.
24. Manring, N.D., *Measuring pump efficiency: uncertainty considerations.* SME Journal of Energy Resources Technology, 2005(127).
25. Salehi, S., et al., *On the flow field and performance of a centrifugal pump under operational and geometrical uncertainties.* Applied Mathematical Modelling, 2018. **61**: p. 540-560.
26. KATO, T., S. SAKAI, and R. ARAI, *On the Uncertainty Analysis via Low Frequency Inputs for Hydraulic Cylinder Dynamics.* JFPS International Journal of Fluid Power System, 2022. **15**(3): p. 95-100.
27. Nguyen, M.H. and K.K. Ahn, *Output feedback robust tracking control for a variable-speed pump-controlled hydraulic system subject to mismatched uncertainties.* Mathematics, 2023. **11**(8): p. 1783.
28. Marquardt, D.W., *An algorithm for least-squares estimation of nonlinear parameters.* Journal of the society for Industrial Applied Mathematics, 1963. **11**(2): p. 431-441.
29. Hagan, M.T. and M.B. Menhaj, *Training feedforward networks with the Marquardt algorithm.* IEEE transactions on Neural Networks, 1994. **5**(6): p. 989-993.
30. Demuth, H.B. and M.H. Beale, *Neural network toolbox user's guide*. 1992: Mathworks, Incorporated.
31. Kim, D., *Normalization methods for input and output vectors in backpropagation neural networks.* International journal of computer mathematics, 1999. **71**(2): p. 161-171.

DYNAMIC ANALYSIS OF AXIAL PISTON PUMPS USING NEURAL NETWORKS

Abstract: *Based on the structural principle and typical operating mechanism of a swash-plate axial piston pump, a system of differential equations describing the dynamic characteristics of the pump system, including the responses of working pressure and flow rate, is established. The numerical solutions of the dynamic model obtained through computational methods are then employed as training data for an artificial neural network. The relationship between structural and operating parameters and the corresponding working responses of the pump is approximated using a neural network architecture with high predictive accuracy. The research findings provide a foundation for investigating the dynamic response of axial piston pumps through a neural network–based surrogate model with significantly reduced computational time, thereby supporting practical design and global optimization tasks.*

Keywords: Axial piston pump; dynamic response; Karhunen–Loève expansion; surrogate model; neural network.

Nghiên cứu sức cản thủy động của cầu phao PMP

Nguyễn Quỳnh Quyên[1], PGS.TS Lê Văn Dưỡng[2]

[1]Hệ quản lý học viên sau đại học, Học viện kỹ thuật quân sự.

[2]Học viện kỹ thuật quân sự

Email: nquyen35@gmail.com; tell: 0964701890

Tóm tắt:

Bài báo trình bày mô hình tính toán sức cản thủy động tác dụng lên cầu phao PMP khi thi công và khai thác trong điều kiện dòng chảy ngang. Trên cơ sở quan hệ lực cản phụ thuộc bình phương vận tốc dòng chảy và các tham số hình học–thủy lực đặc trưng của hệ phao, nghiên cứu xây dựng quy trình xác định lực cản tổng cũng như quy luật phân bố lực cản dọc theo chiều dài cầu theo sự thay đổi của vận tốc dòng chảy trong mặt cắt sông. Kết quả phân bố lực cản cho phép nhận diện các vùng chịu tải thủy động lớn, từ đó cung cấp cơ sở kỹ thuật cho việc lựa chọn số lượng neo, vị trí bố trí dây neo và đánh giá điều kiện làm việc giới hạn của cầu phao. Các kết quả thu được góp phần hoàn thiện hướng dẫn tổ chức lắp ghép và khai thác cầu phao PMP trong điều kiện thủy văn phức tạp, nâng cao mức độ an toàn cho người, phương tiện và trang bị.

Từ khóa: *Cầu phao PMP; sức cản thủy động; dòng chảy; ổn định hệ thống.*

1. Đặt vấn đề

Cầu phao PMP (ПМП–Понтонно мостовой парк) là loại cầu nổi quân sự dạng băng (Ribbon Floating Bridge – RFB) được sử dụng rộng rãi nhằm bảo đảm nhiệm vụ vượt sông trong điều kiện tác chiến; đồng thời, trong thời bình, hệ thống này cũng cho thấy hiệu quả cao trong các tình huống cứu hộ, khắc phục sự cố giao thông nhờ khả năng triển khai nhanh và tính cơ động lớn. Tuy nhiên, khi thi công và khai thác trên các tuyến sông có chế độ thủy văn phức tạp, cầu phao PMP phải chịu tác động trực tiếp của môi trường nước, trong đó dòng chảy là yếu tố chi phối quan trọng nhất.

Hình 1. Bộ Cầu phao PMP

Thực tế cho thấy, khi vận tốc dòng chảy tăng, sức cản thủy động tác dụng lên các đốt phao tăng nhanh theo quy luật tỷ lệ với bình phương vận tốc, làm gia tăng đáng kể tải trọng ngang lên hệ thống. Dưới tác động đó, trạng thái làm việc của cầu có thể bị xáo trộn và phát sinh các nguy cơ mất an toàn như: vặn xoắn kết cấu, trôi dạt toàn hệ, dịch chuyển không kiểm soát theo dòng, thậm chí xuất hiện hiện tượng lật hoặc dìm cục bộ các đốt phao trong những điều kiện bất lợi.

Vì vậy, việc nghiên cứu một cách hệ thống sức cản thủy động của cầu phao PMP có ý nghĩa thiết thực cả về lý thuyết và ứng dụng: cung cấp cơ sở khoa học phục vụ tính toán, lựa chọn và tối ưu hóa phương án neo giữ/neo ghìm; hỗ trợ xây dựng biểu đồ khai thác an toàn và các khuyến cáo vận hành theo các kịch bản thủy văn khác nhau; qua đó nâng cao độ tin cậy và bảo đảm an toàn cho người, phương tiện và trang bị trong quá trình triển khai cầu phao.

Trên cơ sở đó, bài báo tập trung nghiên cứu, phân tích và mô phỏng các thành phần sức cản thủy động tác dụng lên đốt phao PMP; đồng thời đề xuất cách đánh giá định lượng mức tải thủy động trong điều kiện dòng chảy đặc trưng, làm nền tảng cho các bước tính toán và thảo luận ở các phần tiếp theo.

2. Xây dựng mô hình tính toán

2.1. Giả thiết

Trong phạm vi bài báo, cầu phao PMP ở trạng thái tĩnh tải được giả định đã lắp ghép hoàn chỉnh và neo ghìm cố định, do đó chuyển động tổng thể của hệ cầu được coi là không đáng kể trong quá trình khảo sát. Về mặt mô hình hóa, toàn tuyến cầu được xem như một dầm liên tục đặt nổi trên mặt nước, chịu tác dụng của dòng chảy theo phương xấp xỉ vuông góc với tim cầu. Trường dòng chảy trong sông không đồng đều trên mặt cắt ngang: vận tốc có xu hướng tăng dần từ vùng gần bờ ra khu vực giữa dòng, đồng thời giảm dần theo chiều sâu từ mặt nước xuống đáy sông do ảnh hưởng ma sát đáy và đặc trưng phân bố vận tốc của dòng chảy tự nhiên. Tại thời điểm khảo sát, lưu tốc dòng chảy được giả định ổn định theo thời gian (dòng chảy gần như không đổi trong khoảng thời gian tính toán), nhờ đó có thể xác lập mối quan hệ vận tốc tại các vị trí đặc trưng và sử dụng để xây dựng mô hình tính sức cản thủy động tác dụng lên cầu (Hình 2).

2.2. Mô hình tính toán

2.2.1. Cơ sở lý thuyết của lực cản thủy động

Khi một vật thể đặt trong dòng chảy, tương tác giữa chất lỏng chuyển động và bề mặt vật thể tạo ra lực thủy động tổng hợp. Thành phần lực theo phương ngược hướng chuyển động tương đối giữa dòng chảy và vật thể được gọi là lực cản thủy động (hydrodynamic drag). Trong nhiều bài toán kỹ thuật, đặc biệt ở miền vận tốc dòng chảy đủ lớn (dòng rối thường gặp trong sông), lực cản được mô tả theo quy luật thực nghiệm–lý thuyết dạng bình phương vận tốc.

Lực cản thủy động phụ thuộc chủ yếu vào các nhóm yếu tố sau:

- Vận tốc dòng chảy Vtại vùng tương tác giữa dòng và vật thể;

- Hình dạng và kích thước đặc trưng của phần kết cấu chịu dòng (mặt cắt, độ tù, độ nhám bề mặt…);

- Tính chất của chất lỏng, thể hiện qua khối lượng riêng ρ(và gián tiếp qua độ nhớt);

- Diện tích chiếu ngang chịu lực A, tức diện tích hình chiếu của phần chịu dòng lên mặt phẳng vuông góc với hướng dòng chảy;

- Hệ số cản C_d(không thứ nguyên), phản ánh ảnh hưởng tổng hợp của hình dạng, điều kiện dòng chảy và trạng thái bề mặt đối với mức độ cản.

Trên cơ sở đó, công thức cơ bản của lực cản thủy động được viết:

$$F_d = \frac{1}{2}\rho C_d A V^2, \tag{1}$$

Trong đó:

F_d: lực cản thủy động (N);

ρ: khối lượng riêng của nước (xấp xỉ $1000\ \mathrm{kg/m^3}$);

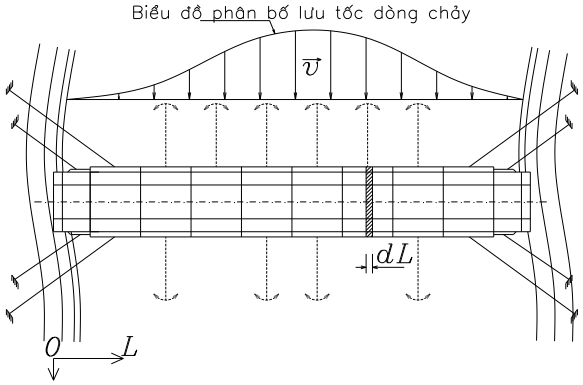
C_d: hệ số lực cản (không thứ nguyên);

A: diện tích chiếu ngang chịu lực (m^2);

V: vận tốc dòng chảy $(\mathrm{m/s})$.

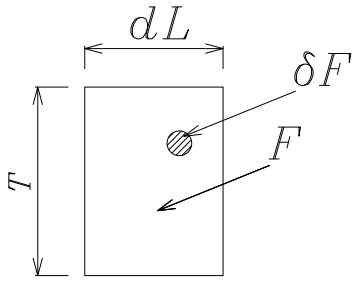
2.2.2. Thiết lập mô hình theo phân tố diện tích và áp suất động

Xét cầu phao PMP có chiều dài theo tim cầu là L, độ chìm tĩnh tải là T. Tại một vị trí dọc theo tim cầu, xét một phân tố chiều dài dL.



Hình 2. Mô hình tính toán

Trên phân tố này, chọn một phân tố diện tích δF thuộc mặt bên thượng lưu của phao (mặt trực tiếp chịu dòng chảy).



Hình 3. Phân tố diện tích

Phân tố diện tích δF ứng với phân tố chiều dài dL.

Tại δF, vận tốc dòng chảy cục bộ là v_F. Theo phương trình Bernoulli, tổng áp suất tác dụng lên δFgồm áp suất thủy tĩnh và áp suất thủy động:

$$\vec{p}_F = \vec{p}_t + \frac{\rho v_F^2}{2} = \vec{p}_t + \vec{p}_d \,, \tag{2}$$

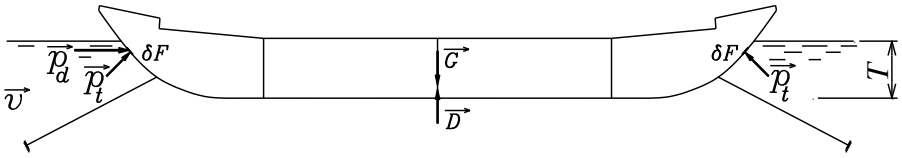
Trong đó:

$\vec{p}_F$: tổng áp suất thủy động và thủy tĩnh tác dụng lên δF,

$\vec{p}_t$ là áp suất thủy tĩnh tại δF;

$\overline{p}_d = \dfrac{\rho v_F^2}{2}$ là áp suất thủy động.

Do áp suất thủy tĩnh phân bố gần đối xứng giữa phía thượng lưu và hạ lưu nên thành phần này không tạo lực ngang đáng kể; đồng thời tác dụng của áp suất thủy động phía hạ lưu được xem là nhỏ. Vì vậy, lực cản theo phương dòng chủ yếu xuất phát từ áp suất thủy động phía thượng lưu của phao.



Hình 4. Áp suất thủy động và thủy tĩnh tác dụng lên phao.

Áp suất thủy động và thủy tĩnh tác dụng lên phao (minh họa cơ sở bỏ qua thủy tĩnh theo phương ngang).

Khi chiếu theo phương tác dụng của dòng chảy, lực cản vi phân do áp suất thủy động gây ra trên δF được xác định:

$$\delta R_F = \frac{\rho v_F^2}{2} \delta F \cos\alpha, \tag{3}$$

Trong đó α là góc giữa mặt phẳng δF và phương của vectơ vận tốc dòng chảy; $\delta F \cos\alpha$ là hình chiếu của δF lên mặt phẳng vuông góc với hướng dòng.

2.2.3. Quy đổi về dạng công thức lực cản chuẩn và biểu thức lực cản toàn cầu

Để xác định lực cản trên phân tố chiều dài dL, coi vận tốc dòng chảy trong phạm vi từ mặt nước đến đáy phao (tương ứng độ chìm T) có giá trị trung bình đặc trưng v_L. Khi đó lực cản trên dL được tính bằng cách tích phân trên toàn bộ diện tích chịu dòng thuộc phân tố dL:

$$dR_L = \int_F \frac{\rho v_F^2}{2} \delta F \cos\alpha \quad , \tag{4}$$

Với giả thiết $v_F \approx v_L$ trong vùng diện tích chịu dòng của dL, đưa v_L ra ngoài dấu tích phân:

$$dR_L = \frac{\rho v_L^2}{2} \int_F \delta F \cos\alpha \quad , \tag{5}$$

Nhận thấy $\int_F \delta F \cos\alpha$ chính là diện tích chiếu ngang hiệu dụng của phần tử theo phương vuông góc với hướng dòng. Để thuận tiện cho mô hình hóa đối với phao dạng băng, đại lượng này được quy đổi theo độ chìm tĩnh tải T và chiều dài phần tử dL thông qua hệ số hiệu dụng ξ:

$$\int_F \delta F \cos\alpha = \xi T dL \quad , \tag{6}$$

$$dR_L = \frac{\xi \rho v_L^2}{2} T dL \quad , \tag{7}$$

Lực cản toàn cầu của cầu phao trên toàn bộ chiều dài L là:

$$R_L = \int_0^L dR_L = \frac{\xi \rho T}{2} \int_0^L v(L)^2 dL \quad , \tag{8}$$

$$F_d = \frac{1}{2} \rho C_d A V^2 \quad , \tag{9}$$

Trong đó, đối với phần tử dL:

- $A \equiv \int_F \delta F \cos\alpha$ là diện tích chiếu ngang hiệu dụng;

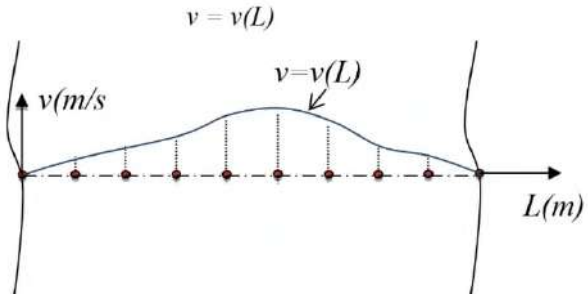
• C_d đóng vai trò hệ số cản tổng quát, và trong mô hình này có thể quy đổi tương đương với hệ số hiệu dụng ξcủa hình dạng phao và trạng thái mớn nước.

2.3. Xác định các tham số tính toán.

a. Xác định lưu tốc dòng chảy v(L).

Do vận tốc dòng chảy trên mặt cắt ngang sông không đồng đều, tại các vị trí dọc theo tim cầu (theo tọa độ L) tiến hành đo vận tốc tại các thủy trực, sau đó nội suy để thu được hàm vận tốc theo chiều dài cầu $v = v(L)$.

$$v = v(L), \tag{10}$$



Hình 5. Sơ đồ thủy trực đo lưu tốc.

b. Xác định độ chìm tĩnh tải T.

Cầu phao PMP dạng băng, thể tích giãn nước của toàn tuyến cầu được biểu diễn theo:

$$V_n = L.S_n \quad , \tag{11}$$

Trong đó S_n là diện tích mặt cắt ngang của phao PMP, nếu S_n là hàm số đơn trị phụ thuộc vào độ chìm T, mỗi giá trị của S_n xác định được một giá trị T. Ta có:

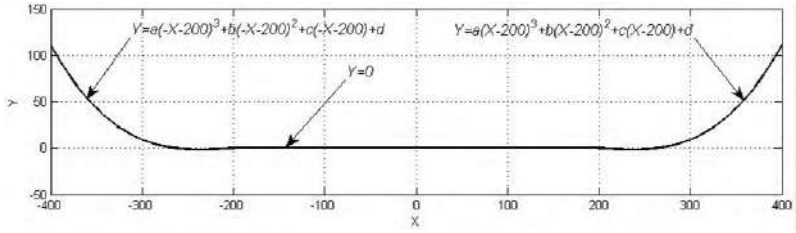
$$V_n(T) = L.f(S_n(T)),$$

$$\Rightarrow \quad T = f^{-1}V_n \tag{12}$$

Theo định luật Archimedes V_n có trọng lượng bằng trọng lượng cầu phao.

Mối quan hệ giữa S_n và T theo [3] được xác định thông qua hàm số.

$$Y = \begin{cases} a(-X-200)^3 + b(-X-200)^2 + c(-X-200) + d; \ X \in [-400,-200) \\ 0; \ X \in [-200,200] \\ a(X-200)^3 + b(X-200)^2 + c(X-200) + d; \ X \in (200,400] \end{cases} \tag{13}$$



Hình 6. Đồ thị phương trình đường cong biên dạng đáy phao PMP.

a = 0.00001543792682; c = - 0.06247408433816;

b = - 0.00000187450894; d = 0.12632945194050;

Căn cứ vào trọng lượng của phao và phương trình đường cong biên dạng đáy (13) có thể xác định độ chìm tĩnh tải $T = 0,15(m)$

c. Xác định hệ số thực nghiệm ξ.

Hệ số ξ được xác định bằng phương pháp thực nghiệm, theo [4] thì hệ số được tính như sau :

$$\xi = \xi_0.\xi_l.\xi_h\,, \qquad (14)$$

Trong đó:

ξ_0: Hệ số kể đến hình dạng mũi phao, cầu PMP lấy $\xi_0 = 0,9$

ξ_l : Hệ số kể đến tỉ số các kích thước phao, cầu PMP lấy $\xi_l = 0,8$

ξ_h : Hệ số kể đến chiều sâu mớn nước, cầu PMP lấy $\xi_l = 1,0$

Từ (14) suy ra: $\xi = 0,9.0.8.1,0 = 0,72$

3. Ví dụ tính toán.

Với sông rộng 250(m), dữ liệu đo lưu tốc dòng chảy ở độ sâu 0,1(m) như sau.

L (m)	0	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190	200	210	220	230	240	250
v_L (m/s)	0	0.18	0.36	0.48	0.57	0.69	0.81	0.93	1.08	1.23	1.35	1.44	1.53	1.59	1.56	1.5	1.41	1.29	1.17	1.05	0.93	0.78	0.63	0.48	0.27	0

Sử dụng phần mềm Matlab nội suy dữ liệu được mối quan hệ giữa v và L qua hàm số:

$$v = v(L) = a.L^6 + b.L^5 + c.L^4 + d.L^3 + e.L^2 + f.L + g \quad (m/s)\,, \qquad (15)$$

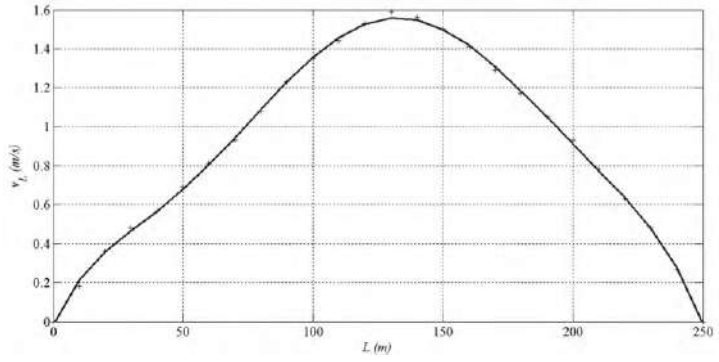
Trong đó:

$a = -5,188937789170133.10^{-13}$ $\quad$ $d = 1,387374915209926.10^{-05}$

$b = 3,963491269253852.10^{-10}$ $\quad$ $e = -7,724593782513301.10^{-04}$

$c = -1,116746915004072.10^{-07}$ $\quad$ $f = 0,029467405732617$

$g = -0,016454037249337$



Hình 7. Đồ thị hàm số v=v(L)

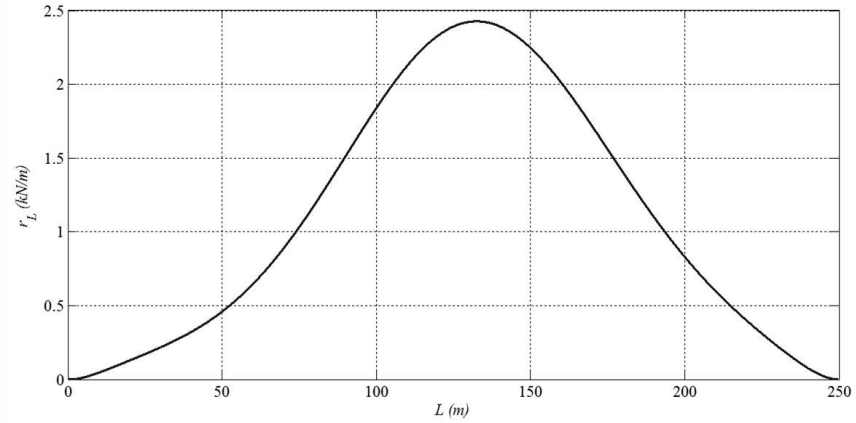
Như vậy từ (6),(12) ta có thể tính được lực cản nước toàn cầu và sự phân bố lực cản tại các vị trí khác nhau.

$$R_L = \frac{\xi\rho T}{2}\int_L v_L^2.dL = \frac{\xi\rho T}{2}\int_L (aL^6 + bL^5 + cL^4 + dL^3 + eL^2 + fL + g)^2.dL \qquad (16)$$

Thay giá trị $L = 250(m)$; $\rho = 10$ (kN/m$^{3)}$ ta được:

R_L =0,5.0,72.10.0,15. 271,3423 = 146,52 (kN)

Từ (13) ta có thể vẽ đồ thị thể hiện sự phân bố lực cản trên toàn cầu:



Hình 8. Đồ thị hàm phân bố lực cản

Căn cứ vào đồ thị của hàm số này chúng tôi đưa ra nhận xét kết quả trong trường hợp này như sau:

- Lực cản nước tăng dần ở 2 bờ ra đến 75(m) bờ trái, 60 (m) bờ phải và tăng nhanh ra tới giữa sông.

- Lực cản đạt giá trị lớn nhất tại 128 (m) từ bờ trái.

- Lực cản nước toàn cầu R_L = 146,52 (kN)

Nghiên cứu đã xây dựng được biểu thức tính sức cản thủy động tác dụng lên cầu phao PMP trong điều kiện dòng chảy ngang, đồng thời thiết lập hàm phân bố lực cản theo chiều dài cầu dựa trên phân bố vận tốc dòng chảy. Kết quả cho thấy việc xét vận tốc không đều trên mặt cắt sông có ý nghĩa quan trọng, vì lực cản phụ thuộc mạnh vào bình phương vận tốc; do đó, các đoạn cầu nằm trong vùng vận tốc lớn sẽ chịu tải thủy động đáng kể và cần được quan tâm trong tính toán ổn định.

Hàm phân bố lực cản thu được là căn cứ trực tiếp để xác định số lượng neo và phương án bố trí neo ghìm theo nguyên tắc: tăng cường neo tại các vị trí có lực cản lớn, đồng thời phân phối tải hợp lý để giảm nguy cơ trôi dạt, vặn xoắn kết cấu và mất ổn định cục bộ khi thi công, khai thác. Về mặt thực tiễn, kết quả nghiên cứu có thể sử dụng để xây dựng biểu đồ/khuyến cáo vận hành theo các kịch bản thủy văn, hỗ trợ lực lượng công binh lựa chọn phương án triển khai cầu phù hợp với điều kiện dòng chảy thực tế tại hiện trường.

Từ kết quả nghiên cứu, đề xuất quy trình tính toán sức cản thủy động toàn cầu theo các bước sau:

- Xác định tim cầu và bố trí các thủy trực đo vận tốc tại các vị trí đặc trưng trên mặt cắt sông;

- Tiến hành đo vận tốc dòng chảy tại các thủy trực;

- Thực hiện nội suy để xây dựng hàm vận tốc theo chiều dài cầu $v = v(L)$;

- Tích phân hàm $v = v(L)$ theo mô hình đã thiết lập để xác định hàm phân bố lực cản và lực cản toàn cầu;

- Trên cơ sở đó, lựa chọn số lượng neo, vị trí neo và phương án neo ghìm nhằm bảo đảm ổn định và an toàn trong thi công, khai thác cầu phao.

Kiến nghị: Trong điều kiện triển khai nhanh, có thể ưu tiên đo vận tốc tại các vị trí đại diện (gần bờ–giữa dòng) để ước lượng sơ bộ phân bố $v(L)$ phục vụ quyết định phương án neo ban đầu; sau đó cập nhật dần số liệu đo để hiệu chỉnh phương án neo khi điều kiện thủy văn thay đổi.Trong thời gian tới, mô hình có thể được mở rộng theo các hướng: (i) xem xét đầy đủ hơn trường vận tốc theo chiều sâu $v = v(L, z)$ nhằm phản ánh ảnh hưởng tầng đáy và mớn nước trong các điều kiện sông sâu–cạn khác nhau; (ii) hiệu chỉnh hệ số cản hiệu dụng (ξ hoặc C_d) bằng thí nghiệm kênh dòng/đo thực địa hoặc mô phỏng số (CFD) để nâng cao độ tin cậy khi áp dụng cho nhiều cấu hình lắp ghép và mức tải khác nhau; (iii) kết hợp mô hình lực cản với bài toán ổn định động của hệ cầu–neo nhằm đánh giá đáp ứng khi dòng chảy biến thiên theo thời gian hoặc khi chịu tác động đồng thời của gió, sóng và tải phương tiện.

Tài liệu tham khảo

1. Bộ Tư lệnh Công binh (1999), *Bộ cầu nổi PMP*, NXB Quân đội nhân dân.
2. Đinh Sơn Hùng (1998), *Các phương tiện vượt sông*, Học viện KTQS, Hà Nội.
3. Nguyễn Huy Hoàng (2010), *Nghiên cứu ổn định nổi của phà PMP 60 tấn trên đặt cần cẩu nâng, hạ hàng*, Luận văn cao học, Học viện kĩ thuật Quân sự.
4. В. И. Телов, И.М. Кануков (1978), *Наплавные мосты, паромные и ледяные переправы*, p. 66-69, -М.: Транспорт.

Study on the hydrodynamic resistance of the PMP pontoon bridge

Abstract: *This paper investigates the hydrodynamic drag acting on the PMP ribbon floating bridge during deployment and operation under cross-current conditions. A calculation model is developed based on the classical quadratic drag relationship and the characteristic geometric–hydraulic parameters of the pontoon system. In order to better represent real river conditions, the approach considers the non-uniform distribution of flow velocity across the river section and establishes a procedure to determine both the global drag force and its spatial distribution along the bridge alignment. The obtained drag-distribution function enables the identification of critical bridge segments subjected to higher hydrodynamic loading, thereby providing a practical basis for selecting the number of anchors, optimizing anchor locations, and assessing operational limits in complex hydrological environments. The results contribute to improving guidelines for assembling and operating PMP floating bridges and enhancing overall safety for personnel, vehicles, and equipment.*

Keywords: PMP floating bridge; hydrodynamic drag; flow current; system stability.

NGHIÊN CỨU ĐỘNG LỰC HỌC NGƯỢC BỘ CÔNG TÁC MÁY BỐC XÚC PHỤC VỤ THI CÔNG ĐƯỜNG HẦM KHẨU ĐỘ NHỎ

Nguyễn Quốc Sáng[1*], Nguyễn Viết Tân[2]

[1]Hệ Quản lý học viên sau đại học; Học viện Kỹ thuật quân sự

[2]Viện Cơ khí động lực; Học viện Kỹ thuật quân sự

Email: quocsangnvs@lqdtu.edu.vn

Tóm tắt

Trong thi công các đường hầm khẩu độ nhỏ (ĐHKĐN), do không gian làm việc bị hạn chế nên máy bốc xúc (MBX) không thể thực hiện toàn bộ các thao tác như trong điều kiện làm việc thông thường. Quá trình xúc, nâng và đổ vật liệu chỉ được thực hiện thông qua sự phối hợp chuyển động của bộ công tác bao gồm cần, tay gầu và gầu. Bài báo này trình bày phương pháp lập quỹ đạo chuyển động hợp lý cho bộ công tác MBX làm việc trong ĐHKĐN nhằm tránh va chạm với các vật cản, đồng thời xây dựng mô hình động lực học (ĐLH) ngược để xác định mô men dẫn động cho các khâu qua các giai đoạn làm việc. Kết quả tính toán cung cấp đồ thị mô men biến thiên theo thời gian, làm cơ sở khoa học cho việc tính toán, nghiên cứu ĐLH hệ thống thủy lực dẫn động bộ công tác của máy.

Từ khóa: *ĐLH ngược; Máy bốc xúc; Đường hầm khẩu độ nhỏ; Bộ công tác; Thiết lập quỹ đạo.*

1. Đặt vấn đề

Trong thi công ĐHKĐN, công tác bốc xúc và vận chuyển đất đá sau nổ mìn đóng vai trò then chốt đối với tiến độ và an toàn thi công. Do kích thước đường hầm hạn chế [1], MBX làm việc trong môi trường này không thể thực hiện đầy đủ các chuyển động như trong điều kiện làm việc ngoài trời, đặc biệt là chuyển động quay toa.

Trong thực tế thi công, MBX chỉ có thể thực hiện các thao tác xúc, nâng và đổ vật liệu thông qua chuyển động phối hợp của bộ công tác gồm cần, tay gầu và gầu. Điều này làm thay đổi đáng kể đặc tính động học và ĐLH của hệ so với máy xúc làm việc trong không gian mở. Nếu không có quỹ đạo chuyển động hợp lý, bộ công tác rất dễ va chạm với nóc, thành hầm hoặc phương tiện vận chuyển, gây mất an toàn và giảm hiệu quả thi công.

Mặc dù đã có nhiều nghiên cứu về động học và ĐLH của máy xúc thủy lực [2, 3], phần lớn các nghiên cứu này được thực hiện trong điều kiện không gian làm việc rộng rãi, khi đó, máy có thể khai thác đầy đủ các chuyển động, đặc biệt là chuyển động quay toa để tối ưu hóa quá trình xúc, đổ vật liệu. Nghiên cứu này tập trung xây dựng mô hình động lực học ngược cho bộ công tác MBX làm việc trong điều kiện ĐHKĐN, bao gồm việc lập quỹ đạo chuyển động cho bộ công tác khi bị hạn chế về mặt không gian và xác định quy luật biến đổi mô men dẫn động cho các khâu theo từng pha làm việc. Kết quả nghiên cứu có ý nghĩa về mặt lý thuyết giúp bổ sung cơ sở phân tích động lực học của MBX trong không gian ràng buộc, đồng thời cũng có giá trị thực tiễn trong thiết kế, lựa chọn kết cấu của MBX phục vụ thi công ĐHKĐN một cách hợp lý và hiệu quả.

2. Xây dựng mô hình ĐLH ngược bộ công tác

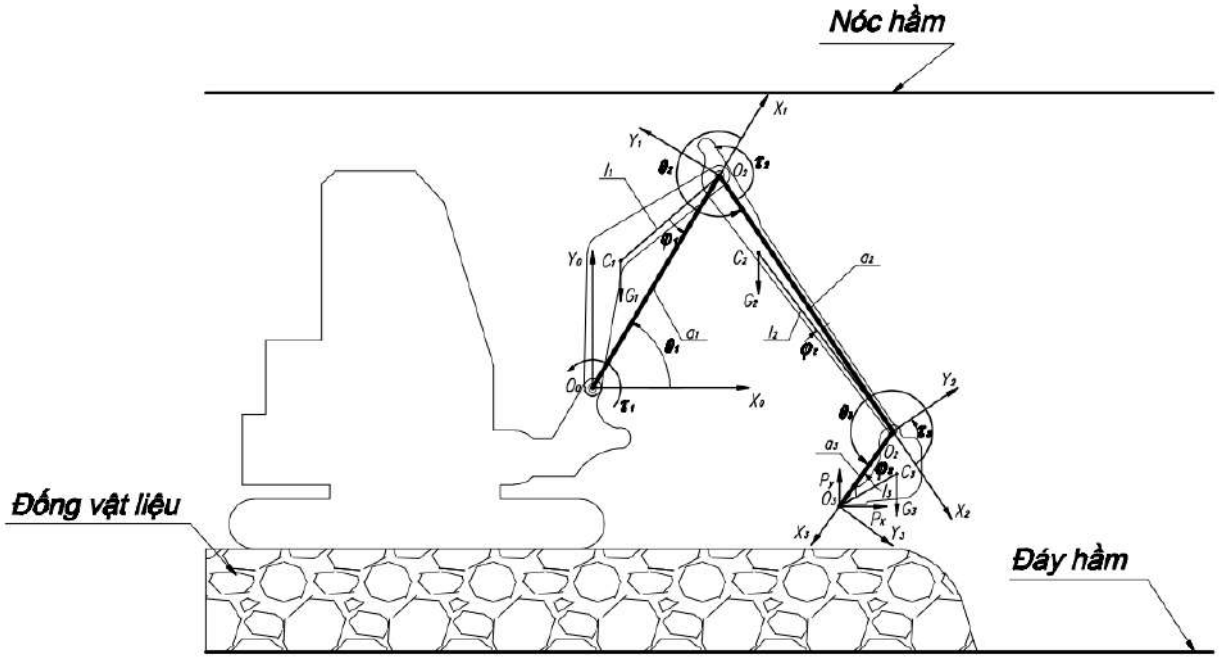
2.1. Các giả thiết mô hình

Để xây dựng mô hình ĐLH, các giả thiết sau được đưa ra: khi làm việc, MBX đứng trên đống vật liệu có bề mặt song song với mặt đất; coi xe cơ sở và bộ công tác là vật rắn tuyệt đối; vật liệu xúc đồng nhất, không có lực cản đột ngột xuất hiện trong quá trình xúc, bỏ qua sự biến dạng của đống vật liệu khi khảo sát; các khâu của bộ công tác chỉ chuyển động trong mặt phẳng

thẳng đứng, đối xứng dọc của máy; lực cản đào và tích vật liệu được quy về một điểm trên răng gầu.

2.2. Thiết lập phương trình vi phân chuyển động của cơ hệ

Mô hình ĐLH của bộ công tác MBX được thể hiện trên hình 1. Các hệ trục tọa độ được lựa chọn theo phương pháp Denavit-Hartenberg [4]. Véc tơ tọa độ suy rộng của cơ hệ là $q = \begin{bmatrix} \theta_1 & \theta_2 & \theta_3 \end{bmatrix}^T$.



Hình 1. Mô hình ĐLH MBX phục vụ thi công ĐHKĐN

Phương trình vi phân chuyển động được xây dựng theo phương pháp Lagrange loại II dạng ma trận [5]:

$$M(q)\ddot{q} + C(q,\dot{q}) + G(q) = Q \tag{1}$$

Trong đó, $[M]$ là ma trận khối lượng suy rộng; $C(q,\dot{q})$ là ma trận Coriolis và lực ly tâm; $G(q)$ là véc tơ lực trọng trường; Q là véc tơ lực suy rộng.

Ma trận khối lượng được tính theo công thức:

$$[M] = \sum_{i=2}^{4} (m_i J_{T_i}^T J_{T_i} + J_{R_i}^T I_i J_{R_i}) = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{33} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \tag{2}$$

trong đó, m_i là khối lượng các khâu, J_{Ti} là các ma trận Jacobi tịnh tiến, J_{Ri} là các ma trận Jacobi quay và I_i là ma trận ten xơ quán tính khối lượng.

Qua các phép biến đổi, thu được các hệ số của ma trận khối lượng như sau:

$$m_{11} = m_1 l_1^2 + m_2(a_1^2 + l_2^2 + 2a_1 l_2 \cos(\theta_2 + \varphi_2)) + m_3(a_1^2 + a_2^2 + l_3^2 + 2a_1 l_3 \cos(\theta_2 + \theta_3 + \varphi_3) + 2a_2 l_3 \cos(\theta_3 + \varphi_3) + 2a_1 a_2 \cos\theta_2) + I_{1Z} + I_{2Z} + I_{3Z}$$

$$m_{12} = m_{21} = m_2(l_2^2 + a_1 l_2 \cos(\theta_1 + \varphi_2)) + m_3(a_2^2 + l_3^2 + a_1 l_3 \cos(\theta_2 + \theta_3 + \varphi_3) + 2a_2 l_3 \cos(\theta_3 + \varphi_3) + a_1 a_2 \cos\theta_2) + I_{2Z} + I_{3Z}$$

$$m_{13} = m_{31} = m_3(l_3^2 + a_2 l_3 cos(\theta_2 + \theta_3 + \varphi_3) + a_2 l_3 cos(\theta_3 + \varphi_3)) + I_{3Z}$$

$$m_{22} = m_2 l_2^2 + m_3(a_2^2 + l_3^2 + 2a_2 l_3 cos(\theta_3 + \varphi_3)) + I_{2Z} + I_{3Z}$$

$$m_{23} = m_{32} = m_3(l_3^2 + a_2 l_3 cos(\theta_3 + \varphi_3)) + I_{3Z};\ m_{33} = m_3 l_3^2 + I_{3Z}$$

Ma trận Coriolis và lực ly tâm có dạng:

$$\mathrm{C(q,\dot{q})} = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} \tag{3}$$

Sử dụng hằng số Christoffel để tính toán, thu được các hệ số của ma trận Coriolis và lực ly tâm như sau:

$$\begin{aligned} C_{11} = &-\left(m_3\left(a_1 a_2 \sin\theta_2 + a_1 l_3 \sin(\varphi_3 + \theta_2 + \theta_3)\right) + a_1 l_2 m_2 \sin(\varphi_2 + \theta_2)\right)\dot{\theta}_2 \\ &-l_3 m_3\left(a_2 \sin(\varphi_3 + \theta_3) + a_1 \sin(\varphi_3 + \theta_2 + \theta_3)\right)\dot{\theta}_3 \end{aligned}$$

$$\begin{aligned} C_{12} = &-\left(m_3\left(a_1 a_2 \sin\theta_2 + a_1 l_3 \sin(\varphi_3 + \theta_2 + \theta_3)\right) + a_1 l_2 m_2 \sin(\varphi_2 + \theta_2)\right)\dot{\theta}_2 \\ &-\left(m_3\left(a_1 a_2 \sin\theta_2 + a_1 l_3 \sin(\varphi_3 + \theta_2 + \theta_3)\right) + a_1 l_2 m_2 \sin(\varphi_2 + \theta_2)\right)\dot{\theta}_1 \\ &-l_3 m_3\left(a_2 \sin(\varphi_3 + \theta_3) + a_1 \sin(\varphi_3 + \theta_2 + \theta_3)\right)\dot{\theta}_3 \end{aligned}$$

$$C_{13} = -l_3 m_3\left(a_2 \sin(\varphi_3 + \theta_3) + a_1 \sin(\varphi_3 + \theta_2 + \theta_3)\right)\left(\dot{\theta}_1 + \dot{\theta}_2 + \dot{\theta}_3\right)$$

$$C_{21} = \left(m_3\left(a_1 a_2 \sin\theta_2 + a_1 l_3 \sin(\varphi_3 + \theta_2 + \theta_3)\right) + a_1 l_2 m_2 \sin(\varphi_2 + \theta_2)\right)\dot{\theta}_1 - a_2 l_3 m_3 \sin(\varphi_3 + \theta_3)\dot{\theta}_3$$

$$C_{22} = -a_2 l_3 m_3 \sin(\varphi_3 + \theta_3)\dot{\theta}_3$$

$$C_{23} = -a_2 l_3 m_3 \sin(\varphi_3 + \theta_3)\left(\dot{\theta}_1 + \dot{\theta}_2 + \dot{\theta}_3\right)$$

$$C_{31} = l_3 m_3\left(a_2 \sin(\varphi_3 + \theta_3) + a_1 \sin(\varphi_3 + \theta_2 + \theta_3)\right)\dot{\theta}_1 + a_2 l_3 m_3 \sin(\varphi_3 + \theta_3)\dot{\theta}_2$$

$$C_{32} = a_2 l_3 m_3 \sin(\varphi_3 + \theta_3)\left(\dot{\theta}_1 + \dot{\theta}_2\right);\ C_{33} = 0$$

G(q) là véc tơ lực trọng trường, qua các phép biến đổi, thu được:

$$G(q) = \frac{\partial \Pi}{\partial q} = \begin{bmatrix} gl_3 m_3 \cos(\varphi_3 + \theta_1 + \theta_2 + \theta_3) + a_2 g m_3 \cos(\theta_1 + \theta_2) + gl_1 m_1 \cos(\varphi_1 + \theta_1) \\ +a_1 g m_2 \cos\theta_1 + a_1 g m_3 \cos\theta_1 + gl_2 l_2 \cos(\varphi_2 + \theta_1 + \theta_2) \\ gl_3 m_3 \cos\left(\varphi_3 + \theta_1 + \theta_2 + \theta_3\right) + a_2 g m_3 \cos\left(\theta_1 + \theta_2\right) + gl_2 m_2 \cos\left(\varphi_2 + \theta_1 + \theta_2\right) \\ gl_3 m_3 \cos\left(\varphi_3 + \theta_1 + \theta_2 + \theta_3\right) \end{bmatrix} \tag{4}$$

Lực suy rộng Q bao gồm mô men dẫn động lên các khâu τ và lực cản P tại đỉnh răng gầu (chỉ xuất hiện ở giai đoạn đào).

Lực suy rộng được tính theo công thức:

$$Q_i = \tau_i - P_x \frac{\partial x_P}{\partial \theta_i} - P_Y \frac{\partial y_P}{\partial \theta_i} \tag{5}$$

trong đó x_P, y_P là tọa độ đỉnh răng gầu và được xác định dễ dàng bằng bài toán động học thuận. Từ đó, lực suy rộng trên các khâu được tính như sau:

$$Q_1 = \tau_1 + P_X\left(a_1\sin\theta_1 + a_2\sin(\theta_1+\theta_2) + a_3\sin(\theta_1+\theta_2+\theta_3)\right) - P_Y\left(a_1\cos\theta_1 + a_2\cos(\theta_1+\theta_2) + a_3\cos(\theta_1+\theta_2+\theta_3)\right) \quad (6)$$

$$Q_2 = \tau_2 + P_X\left(a_2\sin(\theta_1+\theta_2) + a_3\sin(\theta_1+\theta_2+\theta_3)\right) - P_Y\left(a_2\cos(\theta_1+\theta_2) + a_3\cos(\theta_1+\theta_2+\theta_3)\right) \quad (7)$$

$$Q_3 = \tau_3 + P_X a_3\sin(\theta_1+\theta_2+\theta_3) - P_Y a_3\cos(\theta_1+\theta_2+\theta_3) \quad (8)$$

Thế các biểu thức (2), (3), (4), (5), (6), (7), (8) vào (1), thu được phương trình vi phân chuyển động của các khâu:

$$\begin{aligned}
\tau_1 = {} & \begin{pmatrix} I_{1Z}+I_{2Z}+I_{3Z}+a_1^2m_2+a_1^2m_3+a_2^2m_3+l_1^2m_1+l_2^2m_2+l_3^2m_3+2a_1l_2m_2\cos(\varphi_2+\theta_2) \\ +2a_2l_3m_3\cos(\varphi_3+\theta_3)+2a_1a_2m_3\cos\theta_2+2a_1l_3m_3\cos(\varphi_3+\theta_2+\theta_3) \end{pmatrix}\ddot{\theta}_1 \\
& + \begin{pmatrix} m_3a_2^2+2m_3\cos(\varphi_3+\theta_3)a_2l_3+a_1a_2m_3\cos\theta_2+m_2l_2^2+a_1m_2l_2\cos(\varphi_2+\theta_2) \\ +m_3l_3^2+a_1m_3l_3\cos(\varphi_3+\theta_2+\theta_3)+I_{2Z}+I_{3Z} \end{pmatrix}\ddot{\theta}_2 \\
& + \left(I_{3Z}+l_3^2m_3+a_2l_3m_3\cos(\varphi_3+\theta_3)+a_1l_3m_3\cos(\varphi_3+\theta_2+\theta_3)\right)\ddot{\theta}_3 \\
& + \left(-a_1l_2m_2\sin(\varphi_2+\theta_2)-a_1a_2m_3\sin\theta_2-a_1l_3m_3\sin(\varphi_3+\theta_2+\theta_3)\right)\dot{\theta}_2^2 \\
& + \left(-a_2l_3m_3\sin(\varphi_3+\theta_3)-a_1l_3m_3\sin(\varphi_3+\theta_2+\theta_3)\right)\dot{\theta}_3^2 \\
& - \left(2a_1l_2m_2\sin(\varphi_2+\theta_2)+2a_1a_2m_3\sin\theta_2+2a_1l_3m_3\sin(\varphi_3+\theta_2+\theta_3)\right)\dot{\theta}_1\dot{\theta}_2 \\
& - \left(2a_2l_3m_3\sin(\varphi_3+\theta_3)+2a_1l_3m_3\sin(\varphi_3+\theta_2+\theta_3)\right)\dot{\theta}_1\dot{\theta}_3 \\
& - \left(2a_2l_3m_3\sin(\varphi_3+\theta_3)+2a_1l_3m_3\sin(\varphi_3+\theta_2+\theta_3)\right)\dot{\theta}_2\dot{\theta}_3 \\
& + gl_3m_3\cos(\varphi_3+\theta_1+\theta_2+\theta_3)+a_2gm_3\cos(\theta_1+\theta_2)+gl_1m_1\cos(\varphi_1+\theta_1) \\
& + a_1gm_2\cos\theta_1+a_1gm_3\cos\theta_1+gl_2l_2\cos(\varphi_2+\theta_1+\theta_2) \\
& - P_x\left(a_1\sin\theta_1+a_2\sin(\theta_1+\theta_2)+a_3\sin(\theta_1+\theta_2+\theta_3)\right)+P_Y\left(a_1\cos\theta_1+a_2\cos(\theta_1+\theta_2)+a_3\cos(\theta_1+\theta_2+\theta_3)\right)
\end{aligned} \quad (9)$$

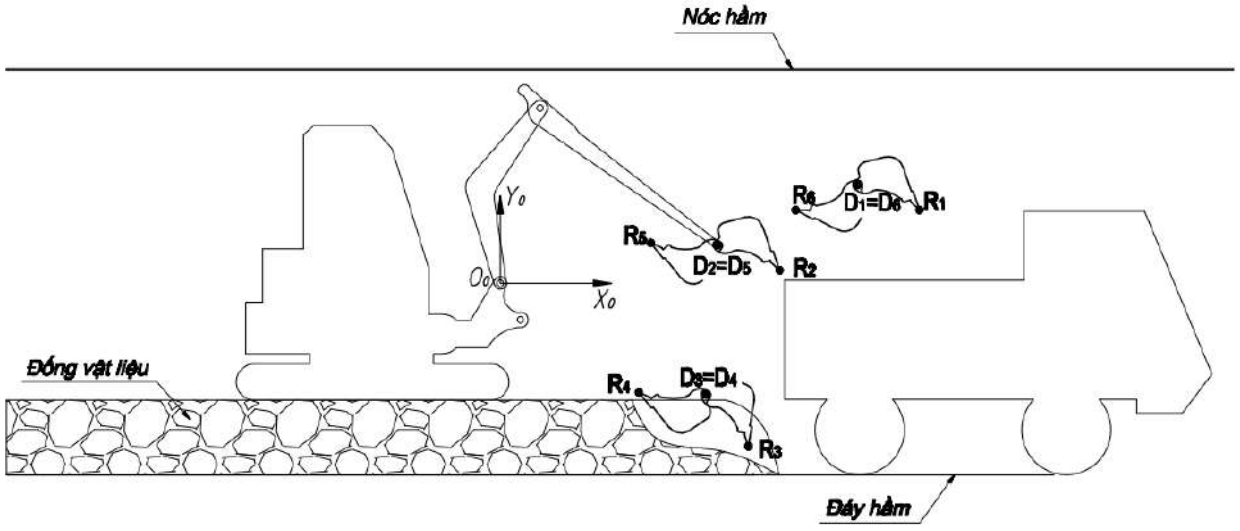
$$\begin{aligned}
\tau_2 = {} & \begin{pmatrix} m_3a_2^2+2m_3a_2l_3\cos(\varphi_3+\theta_3)+m_3a_1a_2\cos\theta_2+m_2l_2^2+a_1m_2l_2\cos(\varphi_2+\theta_2) \\ +m_3l_3^2+a_1m_3l_3\cos(\varphi_3+\theta_2+\theta_3)+I_{2Z}+I_{3Z} \end{pmatrix}\ddot{\theta}_1 \\
& + \left(m_3a_2^2+2m_3a_2l_3\cos(\varphi_3+\theta_3)+m_2l_2^2+m_3l_3^2+I_{2Z}+I_{3Z}\right)\ddot{\theta}_2+\left(m_3l_3^2+a_2m_3l_3\cos(\varphi_3+\theta_3)+I_{3Z}\right)\ddot{\theta}_3 \\
& + \left(a_1l_2m_2\sin(\varphi_2+\theta_2)+a_1a_2m_3\sin\theta_2+a_1l_3m_3\sin(\varphi_3+\theta_2+\theta_3)\right)\dot{\theta}_1^2 \\
& - a_2l_3m_3\sin(\varphi_3+\theta_3)\dot{\theta}_3^2-2a_2l_3m_3\sin(\varphi_3+\theta_3)\dot{\theta}_1\dot{\theta}_3-2a_2l_3m_3\sin(\varphi_3+\theta_3)\dot{\theta}_2\dot{\theta}_3 \\
& + gl_3m_3\cos(\varphi_3+\theta_1+\theta_2+\theta_3)+a_2gm_3\cos(\theta_1+\theta_2)+gl_2m_2\cos(\varphi_2+\theta_1+\theta_2) \\
& - P_X\left(a_2\sin(\theta_1+\theta_2)+a_3\sin(\theta_1+\theta_2+\theta_3)\right)+P_Y\left(a_2\cos(\theta_1+\theta_2)+a_3\cos(\theta_1+\theta_2+\theta_3)\right)
\end{aligned} \quad (10)$$

$$\begin{aligned}
\tau_3 = {} & \left(I_{3Z}+l_3^2m_3+a_2l_3m_3\cos(\varphi_3+\theta_3)+a_1l_3m_3\cos(\varphi_3+\theta_2+\theta_3)\right)\ddot{\theta}_1+\left(m_3l_3^2+a_2m_3\cos(\varphi_3+\theta_3)l_3+I_{3Z}\right)\ddot{\theta}_2 \\
& + \left(m_3l_3^2+I_{3Z}\right)\ddot{\theta}_3 \quad + \left(a_2l_3m_3\sin(\varphi_3+\theta_3)+a_1l_3m_3\sin(\varphi_3+\theta_2+\theta_3)\right)\dot{\theta}_1^2+a_2l_3m_3\sin(\varphi_3+\theta_3)\dot{\theta}_2^2 \\
& + 2a_2l_3m_3\sin(\varphi_3+\theta_3)\dot{\theta}_1\dot{\theta}_2+gl_3m_3\cos(\varphi_3+\theta_1+\theta_2+\theta_3)-P_Xa_3\sin(\theta_1+\theta_2+\theta_3)+P_Ya_3\cos(\theta_1+\theta_2+\theta_3)
\end{aligned} \quad (11)$$

2.3. Bài toán ĐLH ngược

Sử dụng phần mềm Solidworks, dựng mô hình 3D theo đúng kích thước và hình dáng của MBX và môi trường làm việc. Để tránh va chạm với các vật cản trong hầm, chọn 6 điểm thao tác

của đỉnh răng gầu ứng với các pha của một chu kỳ: thu gầu (R_1-R_2), hạ gầu (R_2-R_3), đào vật liệu (R_3-R_4), nâng gầu (R_4-R_5), duỗi gầu (R_5-R_6) và đổ vật liệu (R_6-R_1), như trên hình 2.



Hình 2. Vị trí các điểm đặc trưng trên quỹ đạo chuyển động của đỉnh răng gầu và khớp quay gầu

Từ tọa độ 6 điểm đặc trưng, sử dụng các công thức tính toán động học ngược, dễ dàng xác định được góc khớp của các khâu tại các điểm này. Để nội suy chuyển động của các khâu, sử dụng đa thức bậc 5 có dạng tổng quát như sau [6]:

$$\theta(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + a_5 t^5$$

Các hệ số của đa thức được xác định từ các điều kiện biên về vị trí, vận tốc và gia tốc tại hai đầu mỗi giai đoạn chuyển động. Phương pháp này đảm bảo tính liên tục của chuyển vị, vận tốc và gia tốc, đáp ứng yêu cầu điều khiển thực tế của hệ thống thủy lực.

Bảng 1. Tọa độ các điểm đặc trưng của khớp quay gầu và đỉnh răng gầu

Điểm	x	y	Điểm	x	y
D_1	2388	655	R_1	2802	488
D_2	1453	252	R_2	1867	86
D_3	1373	-744	R_3	1658	-1088
D_4	2388	655	R_4	927	-729
D_5	1453	252	R_5	1006	268
D_6	1373	-744	R_6	1974	488

Giả định tọa độ của các điểm đặc trưng như trong bảng 1, vận tốc và gia tốc tại các điểm này đều bằng 0, sau khi tính toán, nhận được hệ các phương trình nội suy chuyển động như sau:

Phương trình chuyển động của cần:

Đoạn R_1-R_2 với $t \in [0-2,5]$: $\theta_1^1(t) = 2.18t^5 - 13.59t^4 + 22.66t^3 + 41.80$

Đoạn R_2-R_3 với $t \in [2,5-4]$: $\theta_1^2(t) = -32.71t^5 + 531.56t^4 - 3393.78t^3 + 10631.11t^2 - 16355.56t + 9979.98$

Đoạn R_3-R_4 với $t \in [4-7]$: $\theta_1^3(t) = 35.8$

Đoạn R_4-R_5 với $t \in [7-9]$: $\theta_1^4(t) = 7.76t^5 - 310.50t^4 + 4942.13t^3 - 39123.00t^2 + 154046.81t - 241367.60$

Đoạn R_5-R_6 với $t \in [9-11]$: $\theta_1^5(t) = -6.64t^5 + 331.88t^4 - 6615.38t^3 + 65711.25t^2 - 325270.69t + 642016.38$

Đoạn R_6-R_1 với $t \in [11-14]$: $\theta_1^6(t) = 41.8$

Phương trình chuyển động của tay gầu:

Đoạn R_1-R_2 với $t \in [0-2,5]$: $\theta_2^1(t) = -4.15t^5 + 25.96t^4 - 43.26t^3 - 47.30$

Đoạn R_2-R_3 với $t \in [2,5-4]$: $\theta_2^2(t) = 3.63t^5 - 59.06t^4 + 377.09t^3 - 1181.23t^2 + 1817.28t - 1215.21$

Đoạn R_3-R_4 với $t \in [4-7]$: $\theta_2^3(t) = -110.30$

Đoạn R_4-R_5 với $t \in [7-9]$: $\theta_2^4(t) = -0.86t^5 + 34.50t^4 - 549.13t^3 + 4347.00t^2 - 17116.31t + 26712.30$

Đoạn R_5-R_6 với $t \in [9-11]$: $\theta_2^5(t) = 12.68t^5 - 633.75t^4 + 12632.75t^3 - 125482.50t^2 + 621138.38t - 1225964.85$

Đoạn R_6-R_1 với $t \in [11-14]$: $\theta_2^6(t) = -47.30$

Phương trình chuyển động của gầu:

Đoạn R_1-R_2 với $t \in [0-2,5]$: $\theta_3^1(t) = 1.99t^5 - 12.44t^4 + 20.74t^3 - 16.50$

Đoạn R_2-R_3 với $t \in [2,5-4]$: $\theta_3^2(t) = 6.48t^5 - 105.28t^4 + 672.20t^3 - 2105.68t^2 + 3239.51t - 1945.52$

Đoạn R_3-R_4 với $t \in [4-7]$: $\theta_3^3(t) = -3.25t^5 + 89.29t^4 - 957.84t^3 + 5000.25t^2 - 12727.90t + 12700.05$

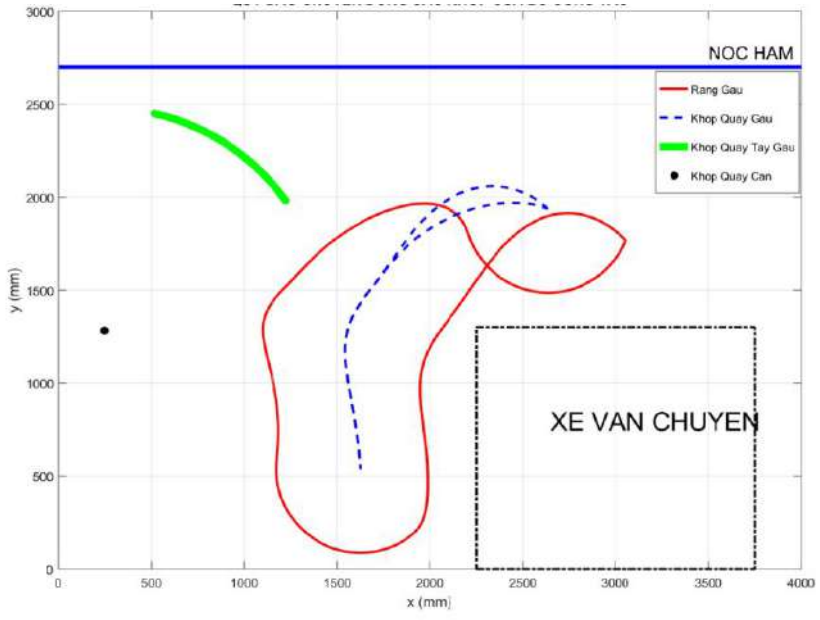
Đoạn R_4-R_5 với $t \in [7-9]$: $\theta_3^4(t) = -6.94t^5 + 277.50t^4 - 4416.88t^3 + 34965.00t^2 - 137674.69t + 215639.60$

Đoạn R_5-R_6 với $t \in [9-11]$: $\theta_3^5(t) = -1.52t^5 + 75.94t^4 - 1513.69t^3 + 15035.63t^2 - 74426.34t + 146739.99$

Đoạn R_6-R_1 với $t \in [11-14]$: $\theta_3^6(t) = 3.36t^5 - 209.88t^4 + 5221.73t^3 - 64641.98t^2 + 398194.57t - 976744.99$

Kiểm tra quỹ đạo và va chạm của bộ công tác

Từ các phương trình chuyển động trên, áp dụng các phương trình động học thuận, vẽ lại quỹ đạo chuyển động của khớp tay gầu, khớp gầu và đỉnh răng gầu để kiểm chứng như hình 3. Nhìn vào đồ thị quỹ đạo chuyển động của các khâu có thể thấy trong quá trình di chuyển, các khâu của bộ công tác không va chạm vào nóc hầm, đáy hầm hay thùng xe vận chuyển, đảm bảo yêu cầu đặt ra.



Hình 3. Quỹ đạo chuyển động của các khớp theo các phương trình chuyển động đã xác định

Thế các biểu thức $\theta_1(t)$, $\theta_2(t)$, $\theta_3(t)$ và biểu thức xác định vận tốc gia tốc tương ứng vào các phương trình vi phân chuyển động, sẽ xác định được mô men dẫn động các khâu $\tau_1(t)$, $\tau_2(t)$, $\tau_3(t)$.

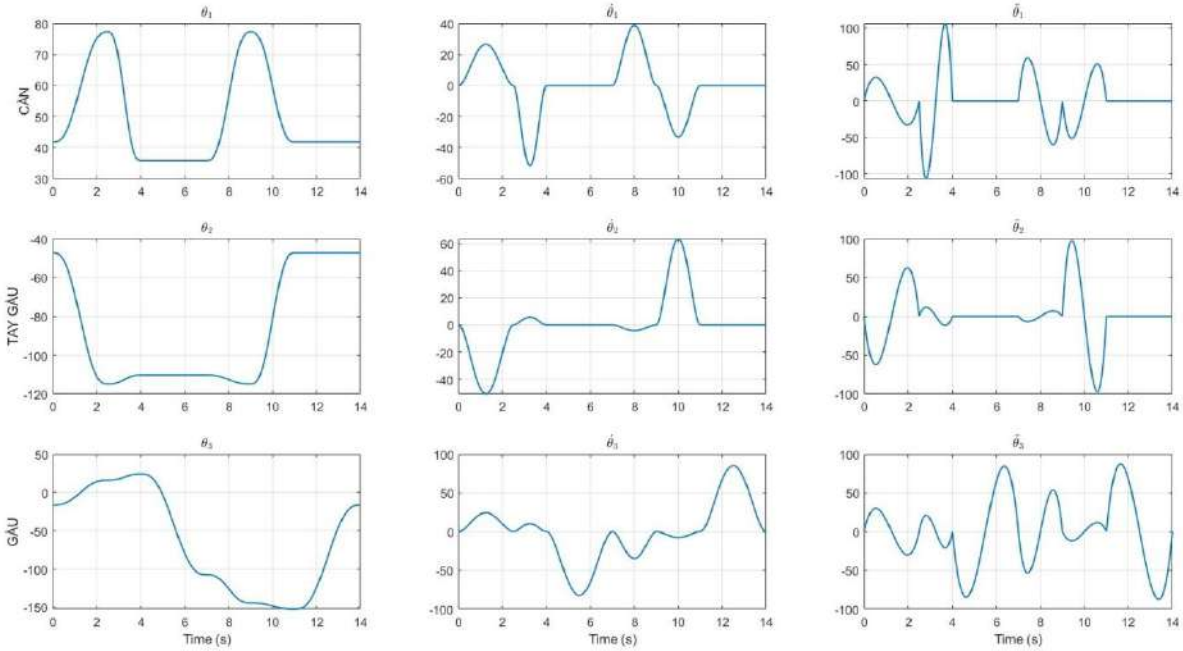
3. Kết quả và thảo luận

Để tính toán số, sử dụng bộ số liệu sau đây:

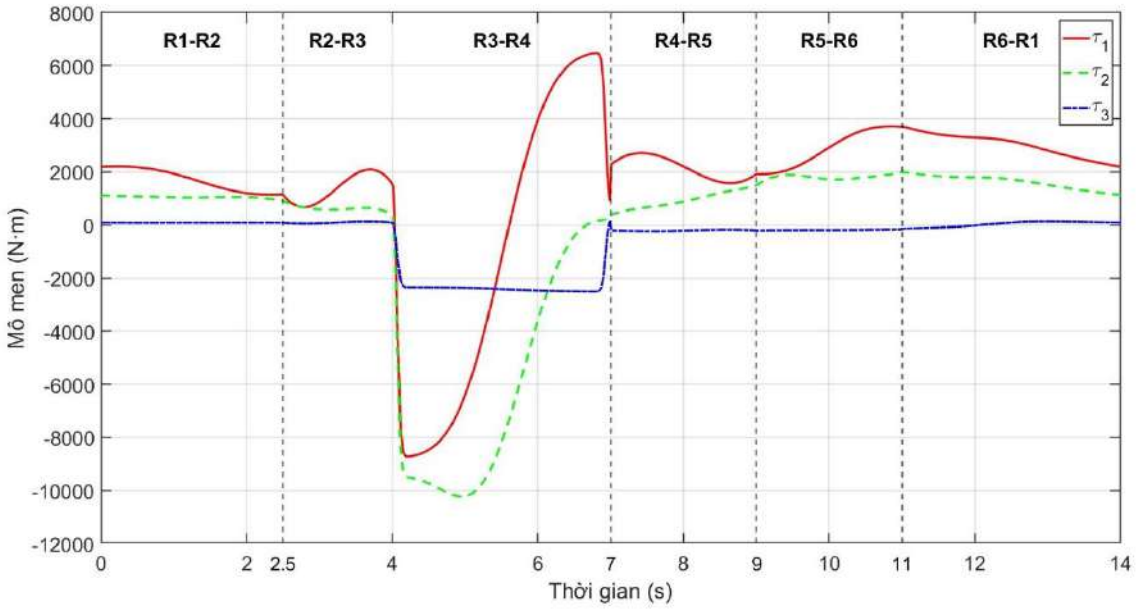
$a_1 = 1,2m; a_2 = 1,5m; a_3 = 0,45m; l_1 = 0,63m; l_2 = 0,73m; l_3 = 0,22m; m_1 = 59kg; m_2 = 57kg; m_3 = 43kg;$

$I_{1Z} = 31,7kg.m^2; I_{2Z} = 42,1kg.m^2; I_{3Z} = 3kg.m^2; g = 9,81m/s^2; \varphi_1 = 17,6°; \varphi_2 = 1,7°; \varphi_3 = 38,5°; P = 5410N.$

Sử dụng phần mềm MATLAB, tính toán chuyển vị, vận tốc, gia tốc và mô men dẫn động các khâu, kết quả thu được như trên hình 4 và hình 5.



Hình 4. Đồ thị dịch chuyển, vận tốc góc, gia tốc góc của cần, tay gầu và gầu theo thời gian



Hình 5. Đồ thị mô men dẫn động các khâu trong một chu kỳ xúc-đổ

Nhận xét

Về quy luật

Từ hình 4 có thể nhận thấy các góc khớp của cần, tay gầu và gầu biến thiên liên tục theo thời gian và phù hợp với trình tự làm việc thực tế của MBX. Các đường vận tốc và gia tốc góc thay đổi trơn, không xuất hiện dao động đột ngột, cho thấy quỹ đạo chuyển động đảm bảo tính liên tục động học và hạn chế tải trọng va đập lên hệ truyền động thủy lực. Hình 5 cho thấy mô men dẫn động thay đổi rõ rệt theo từng pha làm việc. Trong giai đoạn đào, mô men dẫn động tại các khớp tăng đáng kể do tác động của lực cản đào tại đỉnh răng gầu. Trong giai đoạn nâng và mang tải, mô men tại các khớp, đặc biệt là khớp cần và khớp tay gầu, duy trì ở mức cao để nâng toàn bộ cơ cấu và khối lượng vật liệu. Ở giai đoạn đổ vật liệu, mô men tại khớp gầu thay đổi dấu do sự chuyển đổi trạng thái từ giữ gầu sang mở gầu.

Về giá trị

Giá trị góc khớp của cần biến đổi từ 35° đến 77°, của tay gầu từ -115° đến -47°, của gầu từ -152° đến 24°. Vận tốc góc cực đại của cần, tay gầu và gầu lần lượt là 51°/s, 61°/s, 85°/s, gia tốc góc cực đại của cần, tay gầu và gầu lần lượt là 104°/s^2, 96°/s^2, 87°/s^2. Các giá trị này tương ứng với dải chuyển động và tốc độ làm việc điển hình của máy xúc thủy lực cỡ nhỏ, cho thấy cơ cấu công tác có thể thực hiện quỹ đạo đã xây dựng mà không yêu cầu khả năng chuyển động vượt quá đặc tính đáp ứng của hệ truyền động thủy lực và các cơ cấu chấp hành.

Giá trị tuyệt đối cực đại của mô men dẫn động khâu cần, tay gầu và gầu xuất hiện ở giai đoạn đào và lần lượt là 8,7 kN, 10,2 kN và 2,48 kN. So sánh với khả năng hoạt động, có thể thấy các giá trị này nằm trong giới giới hạn làm việc của MBX.

4. Kết luận

Bài báo đã xây dựng được mô hình ĐLH ngược cho bộ công tác MBX làm việc trong điều kiện ĐHKĐN. Mô hình động học và ĐLH được thiết lập một cách hệ thống, có xét đến đặc thù không gian làm việc chật hẹp và sự thay đổi của ngoại lực trong các giai đoạn làm việc khác nhau. Kết quả nghiên cứu cho phép xác định quy luật biến đổi mô men dẫn động tại các khớp theo thời gian, tạo cơ sở khoa học cho thiết kế, lựa chọn và điều khiển hệ thống truyền động thủy lực của MBX.

Tài liệu tham khảo

1. *Hồ sơ thiết kế các đường hầm có khẩu độ vừa và nhỏ*. Trung tâm tư vấn khảo sát thiết kế các công trình quân sự - BTL Công Binh.
2. Koivo, A. J., Ramos, M. C., Jr., & Thoma, M. (1994). *Dynamic model for excavators (and backhoes)*. Preprints of the 4th IFAC Symposium on Robot Control.
3. Vähä, P. K., & Skibniewski, M. J. (1993). *Dynamic model of excavator*. Journal of Aerospace Engineering, 6(2), 148–158.
4. Denavit J., Hartenberg R.S (1955). *A Kinematic Notation for Lower-Pair Mechanism Based on Matrices*. J. Applied Mechanics, 215-221.
5. Nguyễn Văn Khang, Chu Anh Mỳ (2023). *Cơ sở robot công nghiệp*. Nhà xuất bản Giáo dục Việt Nam, Hà Nội.
6. Jing Yang, Yingjie Gao (2023). *Research on excavator trajectory control based on hybrid interpolation*. Sustainability.

Study on the Inverse Dynamics of the Excavator Working Mechanism for Small Cross-Section Tunnel Construction

Abstract: *In small-cross-section tunnel construction, the excavator operates under severely confined working conditions, which prevent the machine from performing the full range of motions typically available in conventional operating environments. The processes of digging, lifting, and dumping materials are therefore accomplished through coordinated motions of the working device, including the boom, arm, and bucket. This paper presents a method for generating a feasible motion trajectory for the excavator working device operating in small-cross-section tunnels to avoid collisions with surrounding obstacles. Furthermore, an inverse dynamic model is developed to determine the driving torques of individual links throughout different working phases. The calculated results provide time-varying torque profiles, which serve as a scientific basis for the analysis and dynamic design of the hydraulic actuation system driving the excavator working device.*

Keywords: Inverse dynamics; Excavator; Small-section tunnel; Working equipment; Trajectory planning.

Mô hình hóa và phân tích động lực học quá độ của hệ thống thủy lực dẫn động cơ cấu bơi xe thiết giáp chở quân 8x8 ở giai đoạn xuống bến

Nguyễn Tiến Vĩ[1*], Nguyễn Duy Đạt[1], Lê Quang Đạt[1]
[1]Viện Cơ khí động lực/Học viện Kỹ thuật quân sự
[1] Email: tien-vi.nguyen@lqdtu.edu.vn

Tóm tắt

Giai đoạn xuống bến là pha chuyển tiếp đặc thù của cơ cấu bơi xe thiết giáp chở quân 8x8, trong đó hệ thống thủy lực dẫn động làm việc trong điều kiện tải thay đổi nhanh và phi tuyến. Khi cơ cấu bơi tiếp nước từng phần, mô-men cản thủy động và áp suất mạch thủy lực biến thiên mạnh, dễ gây quá tải và dao động trong hệ thống. Bài báo này tập trung nghiên cứu động lực học quá độ của hệ thống thủy lực dẫn động cơ cấu bơi xe thiết giáp chở quân 8x8 trong giai đoạn xuống bến. Một mô hình toán học mô tả mối quan hệ giữa áp suất thủy lực, tốc độ quay mô-tơ và mô-men cản phụ thuộc chiều sâu ngập nước của phương tiện được xây dựng. Các kết quả thu được là cơ sở khoa học quan trọng cho việc lựa chọn tham số thiết kế, đánh giá độ an toàn làm việc và đề xuất các giải pháp điều khiển nhằm nâng cao độ tin cậy của cơ cấu bơi xe thiết giáp chở quân 8x8.

Từ khóa: *xe thiết giáp chở quân 8x8; cơ cấu bơi; hệ thống thủy lực.*

1. Mở đầu

Trong quá trình hiện đại hóa các phương tiện chiến đấu bánh lốp lội nước, xe thiết giáp chở quân cấu hình 8x8 giữ vai trò quan trọng nhờ khả năng cơ động chiến thuật và vượt chướng ngại nước. Khả năng chuyển tiếp hiệu quả giữa môi trường mặt đất và môi trường nước là một trong những chỉ tiêu kỹ – chiến thuật then chốt quyết định tính cơ động và độ tin cậy khai thác của phương tiện. Giai đoạn xe xuống bến, tức quá trình chuyển từ chuyển động trên cạn sang môi trường nước, là một trạng thái làm việc đặc biệt phức tạp. Trong giai đoạn này, tải tác dụng lên cơ cấu bơi thay đổi nhanh do sự xuất hiện của lực cản thủy động và sự thay đổi phân bố tải trọng. Sự gia tăng đột ngột của mô-men cản và lực cản nước có thể gây ra dao động áp suất lớn và biến thiên vận tốc trong thời gian ngắn. Tuy nhiên, các nghiên cứu hiện có chủ yếu tập trung vào trạng thái làm việc xác lập hoặc các bài toán điều khiển trong điều kiện tải ổn định. Đặc tính động lực học quá độ của hệ dẫn động cơ cấu bơi trong giai đoạn xuống bến, đặc biệt là sự tương tác giữa động lực học thủy lực và động lực học cơ học dưới tác động của tải thủy động biến thiên theo thời gian, vẫn chưa được nghiên cứu một cách đầy đủ.

Xuất phát từ yêu cầu đó, bài báo này xây dựng mô hình động lực học của hệ thống thủy lực dẫn động cơ cấu bơi xe thiết giáp 8x8 và tiến hành phân tích đáp ứng quá độ trong giai đoạn xuống bến. Mô hình được thiết lập trên cơ sở các phương trình bảo toàn lưu lượng, cân bằng áp suất và phương trình chuyển động quay của tải, có xét đến ảnh hưởng của tính nén môi chất, tổn thất và mô-men cản thủy động biến thiên.

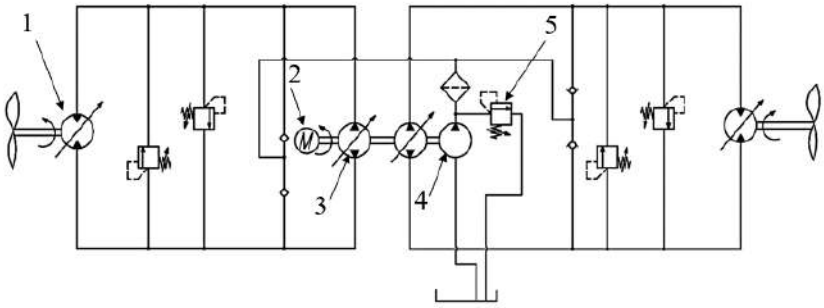
Kết quả nghiên cứu góp phần làm rõ đặc tính động lực học quá độ của hệ thống trong điều kiện chuyển tiếp và cung cấp cơ sở cho việc đánh giá, lựa chọn tham số thiết kế và nâng cao hiệu quả làm việc của cơ cấu bơi trong khai thác thực tế.

2. Cơ sở lý thuyết và xây dựng mô hình toán

2.1. Hệ thống truyền động thủy lực mạch kín dẫn động cơ cấu bơi

Hệ thống dẫn động cơ cấu bơi trên xe thiết giáp chở quân 8x8 sử dụng hệ thống thủy lực mạch kín gồm bơm piston hướng trục có điều khiển lưu lượng và động cơ thủy lực nối trực tiếp

qua hai nhánh áp suất cao - thấp tạo thành vòng tuần hoàn khép kín. Lưu lượng sau động cơ quay trở lại cửa hút của bơm, do đó mô men đầu ra phụ thuộc vào chênh áp giữa hai nhánh. Dung tích riêng của bơm được điều khiển thông qua cơ cấu servo thủy lực sử dụng van điện từ tỉ lệ. Vì vậy, hệ thống là sự liên hệ giữa ba thành phần, gồm có: hệ thống mạch thủy lực chính, cơ cấu tải (cơ cấu bơi và thân xe) và hệ điều khiển servo.



Hình 1. Sơ đồ hệ thống truyền động thủy lực dẫn động cơ cấu bơi trên xe

1. Mô tơ thủy lực dẫn động cơ cấu bơi; 2. Động cơ Diesel;

3. Bơm thủy lực chính; 4. Bơm bù; 5. Van an toàn

Quá trình xe xuống bến đặc trưng bởi các thông số: Độ nghiêng bến xuống; sự gia tăng dần thể tích chìm của thân xe và sự thay đổi tải thủy động tác dụng lên cơ cấu bơi.

2.2. Giả thiết và phạm vi áp dụng của mô hình

Để xây dựng mô hình động lực học quá độ của hệ thống truyền động thủy lực mạch kín dẫn động cơ cấu bơi trong giai đoạn xe xuống bến, các giả thiết sau được áp dụng:

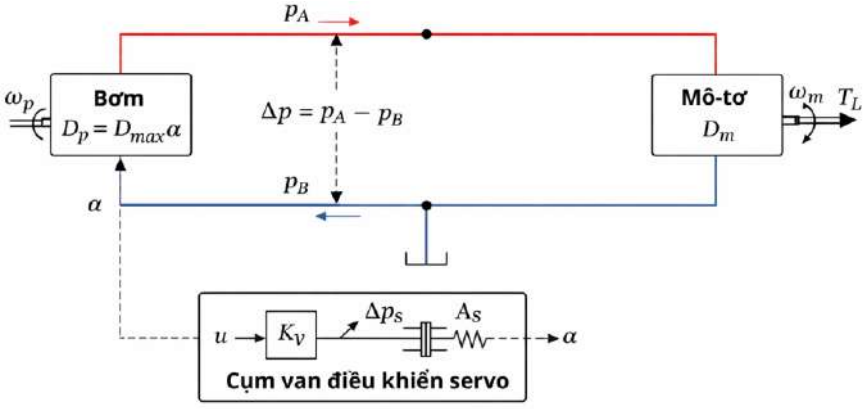
+ Môi chất thủy lực có tính nén với mô đun đàn hồi thể tích E được xem là không đổi trong khoảng áp suất làm việc khảo sát;

+ Hệ thống được mô hình hóa theo tham số tập trung, bỏ qua dao động sóng áp suất dọc đường ống và biến dạng đàn hồi của thành ống;

+ Rò rỉ nội tại của bơm và động cơ thủy lực được giả thiết tỷ lệ tuyến tính với chênh áp Δp; nhiệt độ làm việc của hệ thống được coi là không đổi trong khoảng thời gian khảo sát;

+ Mô hình được thiết lập từ thời điểm lực kéo truyền qua bánh xe không còn ảnh hưởng đáng kể đến động lực học của hệ. Do đó, mô-men truyền từ hệ truyền động bánh xe không được xét trong phương trình động học.

2.3. Mô hình mạch thuỷ lực dẫn động cơ cấu bơi



Hình 2. Sơ đồ mô hình hệ thống thuỷ lực mạch kín dẫn động cơ cấu bơi trên xe

Theo [3], phương trình lưu lượng bơm:

$$Q_p = D_p\,\omega_p - C_{tp}\,\Delta p \tag{1}$$

Trong đó:

+ Q_p – Lưu lượng của bơm chính, [m^3/s];

+ $D_p = D_{max}\alpha$ - Lưu lượng tức thời của bơm, [m^3/s];

+ D_{max} - Lưu lượng riêng cực đại của bơm, [m^3/s];

+ α – Góc nghiêng đĩa bơm đã chuẩn hoá;

+ C_{tp} , C_{tm} – Hệ số rò rỉ bơm và hệ số rò rỉ của mô-tơ.

Xét thể tích dầu tương đương Vt và tính nén của dầu thuỷ lực:

$$\frac{dV}{dt} = \frac{V_t}{E}\frac{d\left(\Delta p\right)}{dt} \tag{2}$$

Theo bảo toàn lưu lượng bơm:

$$Q_p = D_m\omega_m + C_{tm}\Delta p + \frac{V_t}{E}\Delta\dot{p} \tag{3}$$

Thay (1) vào, ta được:

$$\begin{aligned}\Delta\dot{p} &= \frac{E}{V_t}\left[D_{max}\alpha\omega_p - D_m\omega_m - \left(C_{tp}+C_{tm}\right)\Delta\mathrm{p}\right] \\ &= \frac{E}{V_t}\left[D_{max}\alpha\omega_p - D_m\omega_m - C_t\Delta\mathrm{p}\right]\end{aligned} \tag{4}$$

(với $C_t = C_{tp} + C_{tm}$)

Mô-men sinh ra trên trục ra của mô-tơ thuỷ lực:

$$T_m = D_m\,\Delta p \tag{5}$$

Phương trình cân bằng mô-men khi xe xuống bến:

$$J\,\dot{\omega}_m = T_m - T_L - B\,\omega_m \tag{6}$$

Trong đó: T_L – Mô-men tải thuỷ động tổng hợp; $B\omega_m$ – Mô-men tổn thất, với B là hệ số ma sát nhớt.

Cơ cấu bơi của xe là loại bơm phụt, tạo lực đẩy nhờ gia tốc dòng nước và sinh phản lực theo định luật bảo toàn động lượng. Lực đẩy của cơ cấu bơi phụt xác định theo công thức sau:

$$F_T = \rho Q_j\left(V_j - v\right) \tag{7}$$

Trong đó:

+ Q_j – Lưu lượng nước qua cửa ống phụt, [m^3/s]: $Q_j = K_Q\omega_m$ (8)

+ V_j – Vận tốc tia nước tại cửa ống phụt, [m/s]; với A_j là tiết diện của cửa ống phụt, ta có:

$$V_j = \frac{Q_j}{A_j} = \frac{K_Q}{A_j}\omega_m \tag{9}$$

+ v – Vận tốc tịnh tiến của xe, [m/s];

+ ρ – Khối lượng riêng của nước, [kg/m^3].

+ K_Q – hệ số lưu lượng theo tốc độ quay.

Thay (8), (9) vào (7) ta được:

$$F_T = \rho \frac{K_Q^2}{A_j} \omega_m^2 - \rho K_Q v \omega_m \tag{10}$$

Mặt khác, lực cản thuỷ động do thân xe chìm dần trong nước đến giá trị chìm sâu lớn nhất được xác định theo công thức:

$$F_D = \frac{1}{2} \rho C_d A(h) v^2 \tag{11}$$

Trong đó: C_d - Hệ số cản thủy động của thân xe (không thứ nguyên); $A(h)$ - Diện tích ướt, phụ thuộc độ chìm sâu h, [m²]; h - Độ chìm sâu của xe, [m].

Chuyển động theo phương ngang trên nước của xe được xác định bởi:

$$m\dot{v} = F_T - F_D \quad (m \text{ là khối lượng toàn bộ xe, [kg]}) \tag{12}$$

Thay (10),(11) vào (12) ta được:

$$m\dot{v} = \rho \frac{K_Q^2}{A_j} \omega_m^2 - \rho K_Q v \omega_m - \frac{1}{2} \rho C_D A(h) v^2 \tag{13}$$

Theo [6], mô-men cản tải do cơ cấu bơi tác dụng lên trục mô-tơ xác định từ phương trình cân bằng động lượng tổng quát:

$$T_L = f.[(rV_\theta)_{ra} - (\mathrm{r}V_\theta)_{vao}] \tag{14}$$

Trong đó: $f = \rho Q$ – là lưu lượng khối, [kg/s];

rV_θ - là mô-men động lượng riêng tại vị trí xét.

Xét thấy, giả thiết đường nước vào ống phụt là không xoáy nên $(V_\theta)_{vao} \approx 0$, mặt khác r_m (m) là bán kính trung bình cánh nên ta có:

$$T_L = \rho Q r_m V_{\theta out} \tag{15}$$

Tại vị trí bán kính trung bình ta lại có:

$$V_{\theta out} = \sigma U = \sigma \omega_m r_m \tag{16}$$

Trong đó: σ – hệ số tải thuỷ động, không thứ nguyên. Giá trị $\sigma = \sigma(h)$ đặc trưng cho khả năng hút nước ổn định và độ ngập của cửa hút. Về giá trị vật lý, ta có:

$$0 \le \sigma(h) \le 1 \tag{17}$$

Theo [6], để mô tả đúng tính chất giai đoạn chuyển tiếp khi phương tiện xuống bến, chiều sâu chìm nước tăng dần đến cực đại trong thời gian ngắn, nghĩa là:

$$\sigma(h) = \left(\frac{h}{H_{max}} \right)^2 \quad (0 < h < H_{max}) \tag{18}$$

Từ đó suy ra:

$$\begin{aligned} T_L &= \rho Q r_m [\sigma(h) \omega_m r_m] \\ &= \rho \sigma(h) Q r_m^2 \omega_m \end{aligned} \tag{19}$$

Ta lại có, lưu lượng ống phụt phụ thuộc tốc độ trục quay và vận tốc tiến của phương tiện:

$$Q = K_Q\omega_m - K_v v = r_m^2\left(K_1 r_m \omega - K_2 v\right) \tag{20}$$

Trong đó: K_Q – hệ số lưu lượng theo tốc độ quay;K_v – hệ số ảnh hưởng vận tốc tiến; K_1, K_2 – các hệ số quy đổi;

Thay vào biểu thức mô men và khai triển ta được:

$$T_L = \rho\sigma(h)K_1 r_m^5\omega_m^2 - \rho\sigma(h)K_2 r_m^4 v\omega_m \tag{21}$$

Thay (5), (21) vào (6) ta có phương trình cân bằng mô-men như sau:

$$J\dot{\omega}_m = D_m\Delta p - \rho\sigma\left(h\right)K_1 r_m^5\omega_m^2 + \rho\sigma\left(h\right)K_2 r_m^4 v_0\omega_m - B\omega_m \tag{22}$$

Rút gọn ta được:

$$\dot{\omega}_m = \frac{1}{J}\left[D_m\Delta \mathrm{p} - \mathrm{a}\left(\mathrm{h}\right)\omega_m^2 - \mathrm{b}\left(\mathrm{h},v_0\right)\omega_m\right] \tag{23}$$

Trong đó:

$$\begin{aligned} &+\ a\left(h\right) = \rho\sigma\left(h\right)K_1 r_m^5 \\ &+\ b\left(h,v_0\right) = B - \rho\sigma\left(h\right)K_2 r_m^4 v_0 \end{aligned} \tag{24}$$

Do góc nghiêng đĩa bơm quyết định trực tiếp lưu lượng và áp suất mạch kín trong quá trình quá độ khi xe xuống nước, mô hình toán của cụm van điện từ tỉ lệ và cơ cấu servo điều khiển cần được thiết lập nhằm mô tả đầy đủ động lực học điều chỉnh của hệ thống. Van điện từ tỉ lệ điều khiển lưu lượng dầu vào hai buồng xi lanh servo, tạo chênh lệch áp suất làm dịch chuyển piston servo, từ đó thay đổi góc nghiêng đĩa bơm.

Chuỗi tác động có dạng sau: $u\left(t\right) \to x_v \to Q_s \to p_s \to x_s \to \alpha$

Trong đó: $u\left(t\right)$ – Điện áp điều khiển, [V]; x_v - Hành trình con trượt, [m]; Q_s - Lưu lượng dầu vào van, [m^3/s]; p_s - Chênh áp giữa hai khoang của van, [Pa]; x_s - Hành trình piston servo, [m]; α – Góc nghiêng đĩa bơm đã quy đổi, [rad].

Van điện từ được mô hình hóa là một khâu quán tính bậc nhất như sau:

$$T_v\dot{x}_v + x_v = K_v u \tag{25}$$

Trong đó: T_v - Hằng số thời gian của van, [s]; K_v - Hệ số khuếch đại điện-cơ;

Quan hệ lưu lượng tổng quát:

$$Q_s = C_d A\left(x_v\right)\sqrt{\frac{2\left|\Delta p_s\right|}{\rho}}\,\mathrm{sign}\left(\Delta p_s\right) \tag{26}$$

Trong đó: C_d – Hệ số lưu lượng; $A\left(x_v\right)$-Diện tích cửa van phụ thuộc hành trình; $\Delta p_s = p_{s1} - p_{s2}$ là độ chênh áp giữa hai khoang, [Pa]; ρ – Khối lượng riêng của dầu thuỷ lực, [kg/m^3].

Tuyến tính hoá Q_s với giả thiết van làm việc trong vùng tuyến tính quanh giá trị làm việc. Ta được:

$$Q_s = K_q x_v - K_c p_s \tag{27}$$

Trong đó: K_q - hệ số lưu lượng; K_c – hệ số nén dòng.

Phương trình liên tục trong van servo:

$$Q_s = A_s \dot{x}_s + \frac{V_s}{E}\dot{p}_s + C_s p_s \tag{28}$$

Trong đó: V_s – Thể tích dầu trong buồng servo, [m^3]; C_s – Hệ số rò rỉ.

Từ đó ta xây dựng được phương trình vi phân chuyển động của piston servo:

$$M_s \ddot{x}_s + B_s \dot{x}_s + K_s x_s = A_s p_s \tag{29}$$

Trong đó: M_s - Khối lượng quy đổi piston và cơ cấu liên kết, [kg]; B_s - Hệ số cản nhớt cơ khí, [N.s/m]; K_s - Độ cứng lò xo hồi vị, [N/m].

Quan hệ hình học giữa độ dịch chuyển x_s và góc nghiêng α:

$$\alpha = k_\alpha x_s \text{ với } k_\alpha \text{ là hệ số quy dẫn hình học} \tag{30}$$

Dung tích của bơm sẽ thay đổi theo giá trị góc nghiêng:

$$D_p = D_{max}\frac{\alpha}{\alpha_{max}} \tag{31}$$

Từ các quan hệ đã thiết lập cho cụm van điện từ tỉ lệ, cơ cấu servo điều khiển đĩa nghiêng, mạch kín thủy lực và mô-men tải cản, hệ phương trình vi phân phi tuyến mô tả động lực học quá độ của hệ thống dẫn động cơ cấu bơi được xác định bởi:

$$\dot{\mathrm{x}} = f\left(\mathrm{x},\mathrm{u},\mathrm{h}\left(\mathrm{t}\right)\right) \tag{32}$$

Với vectơ trạng thái được chọn là:

$$\mathrm{X} = \begin{bmatrix} x_v & x_s & \dot{x}_s & p_s & \Delta p & \omega_m \end{bmatrix}^T \tag{33}$$

và các phương trình thành phần bao gồm:

$$\begin{cases} \dot{x}_v = \dfrac{1}{T_v}\left(K_v u - x_v\right) \\ \dot{x}_s = \dfrac{1}{M_s}\left(A_s p_s - B_s \dot{x}_s - K_s x_s\right) \\ \dot{p}_s = \dfrac{E}{V_s}\left(K_q x_v - K_c p_s - A_s \dot{x}_s\right) \\ \Delta\dot{p} = \dfrac{E}{V}\left(D_p \omega_p - D_m \omega_m - C_{tm}\Delta\mathrm{p}\right) \\ \dot{\omega}_m = \dfrac{1}{J}\left[D_m \Delta\mathrm{p} - \rho\sigma\left(\mathrm{h}\right)K_1 r_m^5 \omega_m^2 - \left(\mathrm{B} - \rho\sigma\left(\mathrm{h}\right)K_2 r_m^4 v_0\right)\omega_m\right] \end{cases} \tag{34}$$

$$\text{Với : } D_p = D_{max}\frac{k_\alpha x_s}{\alpha_{max}}, \qquad \sigma\left(h\right) = \sigma\left(h\left(t\right)\right)$$

3. Thiết lập bài toán và kết quả khảo sát động lực học khi xe xuống bến

Trên cơ sở mô hình toán đã thiết lập, bài toán khảo sát động lực học của hệ được xem xét trong điều kiện xe thiết giáp chuyển từ môi trường trên cạn xuống nước, trong đó tải thủy động tác dụng lên cơ cấu bơm phụt hình thành theo quá trình tăng độ dần chìm của thân xe. Việc khảo sát được thực hiện bằng cách giải hệ phương trình trạng thái với bộ tham số đặc trưng của hệ thống và các điều kiện ban đầu tương ứng với thời điểm xe bắt đầu tiếp xúc với nước.

3.1. Điều kiện ban đầu và bộ tham số tính toán

Tại thời điểm xe bắt đầu tiếp xúc với nước, hệ thống được xem làm việc trong trạng thái không tải thủy động, khi cơ cấu bơm phụt chưa chịu tác động đáng kể của môi trường nước. Các biến trạng thái ban đầu của hệ được xác định tương ứng với chế độ làm việc trên cạn, trong đó chênh áp mạch kín và mô-men tải thủy động được xem xấp xỉ bằng không. Bài toán khảo sát được thiết lập bằng cách giải hệ phương trình trạng thái với bộ tham số đặc trưng của hệ thống. Bộ tham số đầu vào cụ thể được trình bày trong Bảng 1 và được sử dụng thống nhất trong toàn bộ quá trình tính toán.

Bảng 1. Bộ thông số đầu vào của mô hình động lực học hệ thống

STT	Tham số	Ký hiệu	Giá trị	Đơn vị
I	**Thông số phương tiện**			
1	Khối lượng xe	m	18000	kg
2	Diện tích cản quy đổi	A	20	m²
3	Hệ số cản thủy động	C_d	1.0	–
4	Khối lượng riêng nước	ρ	1000	kg/m³
5	Độ sâu chìm cực đại	H	1.5	m
II	**Hệ thống thủy lực**			
6	Lưu lượng riêng bơm	$D_{p\,max}$	125×10^{-6}	m³/rev
7	Tốc độ trục bơm chính max	N_p	2200	rpm
8	Lưu lượng riêng mô-tơ	D_m	125×10^{-6}	m³/rev
9	Quán tính quy về trục mô-tơ	J_m	2.5	kg.m²
10	Hệ số cản nhớt cơ học	B	7.5	N.m.s
11	Môđun đàn hồi của dầu	E	1400	MPa

Các tham số được lựa chọn dựa trên đặc tính thiết kế của hệ thống thực tế.

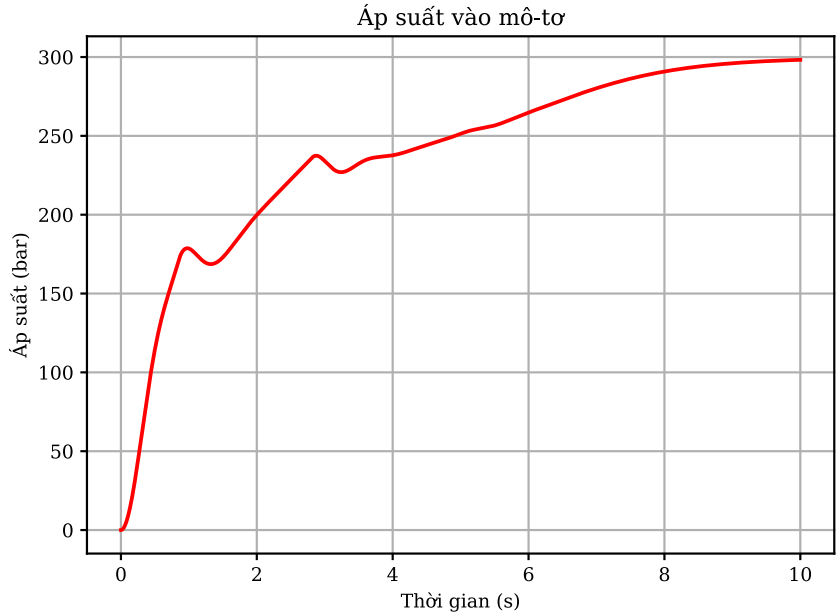
3.2. Kết quả tính toán

Trên cơ sở giải hệ phương trình trạng thái với các điều kiện ban đầu và bộ tham số đã thiết lập, đáp ứng động lực học của hệ dẫn động thủy lực trong giai đoạn xe xuống bến được xác định theo miền thời gian. Các kết quả tính toán tập trung vào sự biến thiên của các đại lượng đặc trưng gồm: áp suất mạch thuỷ lực, vận tốc quay trục bơm phụt, vận tốc bơi của xe. Những đại lượng này phản ánh quá trình hình thành lực đẩy của cơ cấu bơm phụt và sự ra tăng vận tốc chuyển động bơi của xe khi tiếp xúc với môi trường nước.

3.2.1. Đáp ứng chênh áp mạch kín

Khi xe bắt đầu tiếp xúc với nước, tải thủy động tác dụng lên cơ cấu bơm phụt hình thành và tăng dần theo quá trình tăng độ chìm của thân xe. Sự gia tăng tải này dẫn đến yêu cầu mô-men lớn hơn trên trục mô-tơ thủy lực, từ đó làm áp suất trong mạch kín tăng theo thời gian.

Kết quả tính toán cho thấy áp suất vào mô-tơ tăng nhanh trong giai đoạn đầu khi xe bắt đầu tiếp nước, phản ánh sự hình thành tải thủy động ban đầu. Sau đó, tốc độ tăng áp suất giảm dần và tiến tới trạng thái ổn định khi độ chìm của xe đạt giá trị cực đại. Đặc điểm này cho thấy sự gia tăng độ chìm của xe trong quá trình xuống bến là yếu tố chi phối trực tiếp sự biến thiên áp suất trong hệ dẫn động thủy lực.

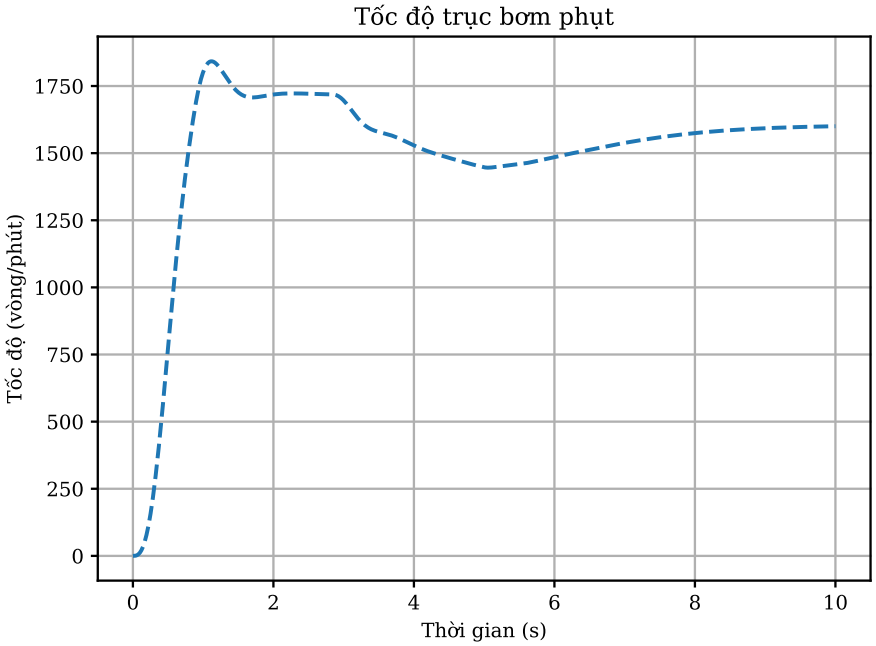


Hình 3. Áp suất vào mô-tơ theo thời gian trong giai đoạn xuống bến

3.2.2. Đáp ứng tốc độ trục bơm phụt

Kết quả tính toán cho thấy khi xe bắt đầu tiếp xúc với nước, sự hình thành và gia tăng của tải thủy động làm giảm tốc độ quay của trục bơm phản lực trong giai đoạn đầu.

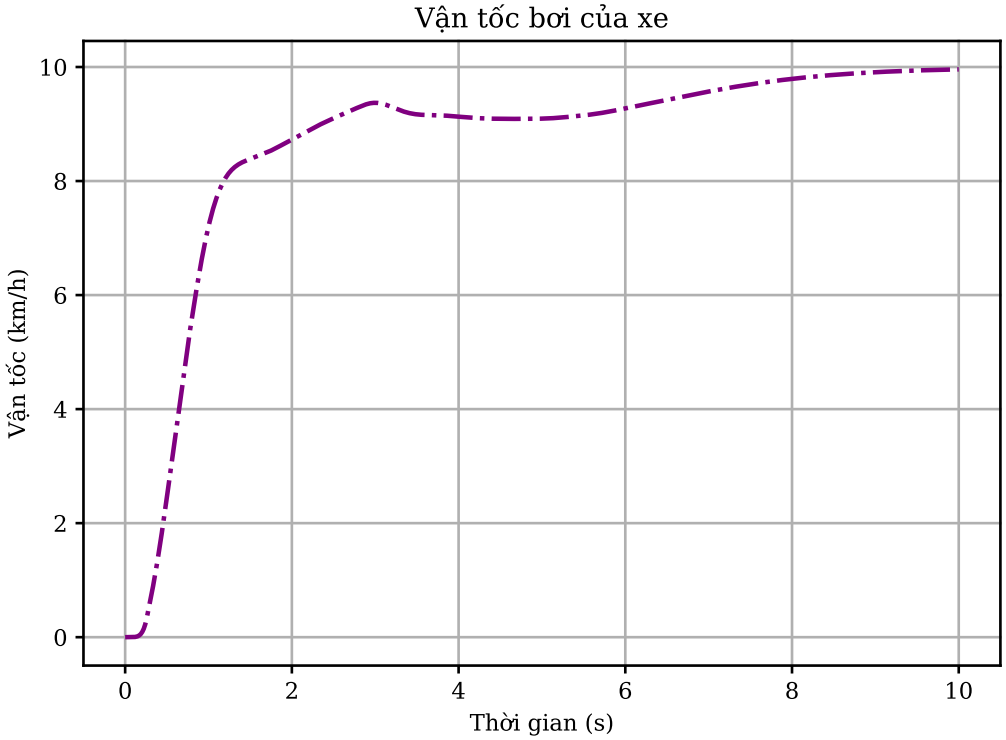
Sự suy giảm này phản ánh ảnh hưởng trực tiếp của mô-men tải tăng do lực cản thủy động.



Hình 4. Tốc độ trục bơm phụt theo thời gian trong giai đoạn xuống bến

Sau giai đoạn chuyển tiếp ban đầu, khi hệ thống thích nghi với sự gia tăng tải, tốc độ trục bơm phản lực tăng dần và tiến tới trạng thái ổn định. Đặc điểm biến thiên này cho thấy hệ dẫn động thủy lực có khả năng đáp ứng với sự thay đổi tải trong quá trình xe xuống nước, đồng thời duy trì được chế độ làm việc ổn định sau khi độ chìm sâu đạt giá trị cực đại. Sự biến thiên tốc độ trục bơm phản lực đóng vai trò quan trọng trong quá trình hình thành lực đẩy, từ đó ảnh hưởng trực tiếp đến vận tốc bơi của xe.

3.2.3. Đáp ứng vận tốc bơi của xe



Hình 5. Vận tốc bơi của xe theo thời gian trong giai đoạn xuống bến

Kết quả tính toán cho thấy khi xe bắt đầu tiếp xúc với nước, vận tốc bơi hình thành và tăng dần theo thời gian do lực đẩy từ bơm phản lực tăng lên, cùng với quá trình tăng độ chìm của thân xe. Trong giai đoạn đầu, vận tốc tăng nhanh khi lực đẩy được hình thành từ sự tăng áp suất và tốc độ quay của trục bơm phản lực. Sau đó, tốc độ tăng vận tốc giảm dần và tiến tới giá trị ổn định khi hệ thống đạt trạng thái cân bằng giữa lực đẩy và lực cản thủy động. Đặc điểm biến thiên này phản ánh quá trình chuyển tiếp từ trạng thái không tải sang trạng thái làm việc ổn định của hệ dẫn động trong môi trường nước.

4. Kết luận

Nghiên cứu đã xây dựng mô hình động lực học của hệ dẫn động thủy lực cơ cấu bơi dạng bơm phụt trong giai đoạn xe thiết giáp xuống bến và thực hiện tính toán đáp ứng của hệ với tải thủy động hình thành theo quá trình tăng độ chìm sâu trong nước.

Kết quả cho thấy sự gia tăng độ chìm sâu của xe dẫn đến biến thiên chênh áp trong mạch kín và ảnh hưởng trực tiếp đến tốc độ bơm phản lực, từ đó chi phối quá trình hình thành vận tốc bơi trong giai đoạn chuyển tiếp từ môi trường trên cạn xuống nước. Mô hình đề xuất cho phép đánh giá đặc tính làm việc của hệ dẫn động thuỷ lực mạch kín trong điều kiện thực tế, làm cơ sở cho phân tích và thiết kế hệ thống thủy lực cho phương tiện lội nước.

Tài liệu tham khảo

Tiếng Việt

1. N. Popov, Động lực học và điều khiển hệ thống dẫn động thủy lực, 1987.
2. Trần Quang Hùng, Đỗ Doãn Phi, Lê Trọng Cường, Trần Hữu Lý (2013), Truyền động thủy lực trên Xe máy Công binh, Học viện kỹ thuật Quân sự.
3. Bùi Hải Triều (2013), Truyền động thủy lực và khí nén, Học viện nông nghiệp Việt Nam.

Tiếng Anh

1. Adam Bureček, Lumír Hružík, Martin Vašina, Simulation of dynamics of hydraulic system with proportional control valve, EPJ Web of Conferences Volume 114 (2016).
2. Alok Vardhan, K.Dasgupta and Santosh Kumar Mishra, Dynamic analysis of a closed-circuit hydraulic drive system used in the rotary head of blasthole drilling machine using MATLAB–Simulink environment, Proceedings of the Institution of Mechanical Engineers Part I Journal of Systems and Control Engineering, November 2018.
3. Carlton, J. S. (John S.), Marine Propellers and Propulsion 3rd , Global Head of MarineTechnology and Investigation, Lloyds Register, 2012.
4. Sujit Kumar, K. Dasgupta and J. Das, Dynamic analysis of a closed-circuit hydrostatic summation drive using bent axis motors, Proceedings of the Institution of Mechanical Engineers Part I Journal of Systems and Control Engineering 229, 2015.
5. Feng Wang, Kim A. Stelson, An Efficient Fan Drive System Based on a Novel Hydraulic Transmission, IEEE/ASME Transactions on Mechatronics (Volume: 20, Issue: 5, October 2015)

Modeling and transient dynamic analysis of the hydraulic drive system for the swimming mechanism of an 8x8 armored personnel carrier during the water-entry phase

Abstract*: The water-entry phase is a distinctive transitional condition of the swimming mechanism of an 8x8 armored personnel carrier, during which the hydraulic drive system operates under rapidly varying and nonlinear loads. As the mechanism becomes partially submerged, the hydrodynamic resistance torque and hydraulic pressure fluctuate significantly, potentially causing overload and system oscillations. This study investigates the transient dynamic behavior of the hydraulic drive system during water entry. A mathematical model is developed to describe the relationship among hydraulic pressure, motor speed, and resistance torque as a function of submergence depth. The results provide a scientific basis for design parameter selection, safety evaluation, and the development of control strategies to improve system reliability.*

Keywords: 8x8 armored personnel carrier; swimming mechanism; hydraulic system.

Nghiên cứu chẩn đoán tình trạng kỹ thuật bộ pin Lithium-ion trên xe điện bằng phương pháp tích hợp định lượng nhanh chính xác.

Đào Nguyễn Thắng[1], Vũ Ngọc Tuấn[1]

[1]Học viện Kỹ thuật quân sự

Email: nguyenthangsqkt1995@gmail.com; Tel: 0354174852

Tóm tắt

Bộ pin lithium-ion là thành phần quan trọng quyết định hiệu suất, độ an toàn và tuổi thọ của ô tô điện. Trong quá trình khai thác, các module pin có thể xảy ra suy giảm không đồng đều, gây mất cân bằng điện áp và làm giảm khả năng làm việc của toàn bộ hệ thống. Bài báo này nghiên cứu mô hình hóa và chẩn đoán tình trạng bộ pin lithium-ion trên ô tô điện dựa trên mô hình mạch tương đương (ECM) kết hợp phân tích dữ liệu điện áp theo thời gian. Trên cơ sở tín hiệu điện áp đầu cực và điện áp sai lệch giữa các module, một chỉ số đánh giá sai lệch động được xây dựng nhằm phát hiện sớm các bất thường ở cấp độ module. Dữ liệu được phân tích trong các chế độ xả, nghỉ và nạp tương ứng với các trạng thái vận hành thực tế của xe. Kết quả cho thấy phương pháp đề xuất có khả năng nhận diện hiệu quả các module pin suy giảm, góp phần nâng cao độ tin cậy và an toàn cho hệ thống pin trên ô tô điện.

Từ khóa: *Pin lithium-ion; ô tô điện; mô hình mạch tương đương (ECM); chẩn đoán module pin; mất cân bằng điện áp.*

1. Đặt vấn đề

Trên thế giới, ô tô điện đang trên đà tăng trưởng mạnh. Tại Việt Nam cũng đang bùng nổ thị trường xe điện trong những năm gần đây (VinFast, các hãng nhập khẩu như Hyundai, Toyota), kéo theo nhu cầu cấp thiết về công nghệ chẩn đoán, bảo dưỡng và quản lý pin. Trong đó pin Lithium-ion là "trái tim" của ô tô điện, chiếm tới 30–40% giá trị xe và quyết định trực tiếp đến phạm vi di chuyển, độ an toàn và tuổi thọ xe. Pin dễ bị ảnh hưởng bởi các yếu tố như quá nhiệt, sạc xả sai cách, lão hóa hóa học, nếu không giám sát chính xác có thể gây ra suy giảm hiệu suất hoặc nguy cơ cháy nổ.Việc chẩn đoán nhanh, chính xác tình trạng pin là nền tảng để phát triển hệ thống quản lý pin (Battery Management System-BMS) thông minh, đảm bảo an toàn vận hành và nâng cao hiệu quả khai thác.

Nhận thức được tầm quan trọng của vấn đề này, việc " Nghiên cứu chẩn đoán tình trạng kỹ thuật bộ pin Lithium-ion trên ô tô điện bằng phương pháp tích hợp định lượng nhanh chính xác" được thực hiện với mục tiêu đảm bảo an toàn và độ tin cậy của ô tô điện. Bằng cách sử dụng các phần mềm mô phỏng hiện đại, kết hợp với các phương pháp tính toán chính xác, không chỉ giúp làm sáng tỏ những vấn đề nghiên cứu mà còn mở rộng ứng dụng trong thực tiễn bảo dưỡng và thiết kế hệ thống quản lý pin.

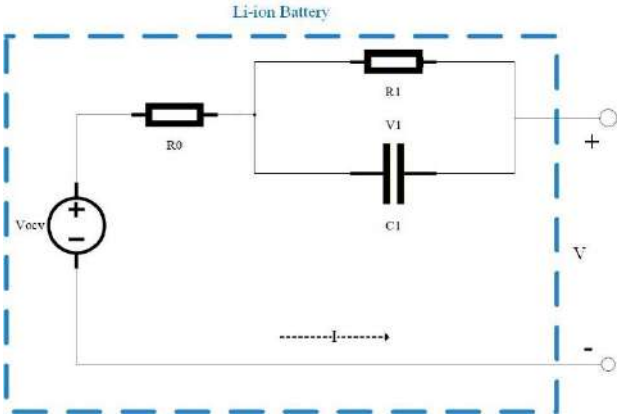
2. Thiết lập mô hình lý thuyết động lực học bộ pin lithium-ion

2.1. Cơ sở lý thuyết của mô hình ECM - 1RC

Bộ pin được mô tả bằng mạch điện tương đương (Equivalent Circuit Model – ECM) nhằm mô phỏng hành vi điện áp, dòng điện, nhiệt độ của pin mà không cần mô phỏng chi tiết hóa học bên trong. Mạch điện tương đương thay thế toàn bộ cấu trúc vật lý, hóa học phức tạp của pin bằng các phần tử điện cơ bản như:

- Nguồn điện áp hở mạch: $V_{OCV}\left(SOC\right)$ (Open Circuit Voltage) đây là điện áp lý tưởng của pin, không tải, không dòng, phụ thuộc hoàn toàn vào SOC (State of Charge)
- Trở nội tức thời: R_0 xuất hiện ngay khi có dòng, tổn hao tức thời trong cell, gây sụt áp.
- Nhánh phân cực động RC: R_1 - C_1

Cấu trúc mô hình ECM-1RC:



Hình 1. Mô hình mạch điện tương đương ECM - 1RC

2.2. Điện áp đầu cực của pin

Dùng mô hình ECM - 1RC gồm OCV, R_0 và một nhánh RC mô tả phân cực, phù hợp cho mô hình BMS và mô phỏng EV.

- Phương trình điện áp: $$V_t\left(t\right)=OCV\left(SOC\right)-I\left(t\right).R_0-V_{RC}\left(t\right) \quad (1)$$

trong đó: OCV(SOC): điện áp hở mạch phụ thuộc SOC, thường lấy từ đường cong thực nghiệm

R_0 :điện trở thuần của cell,

V_{RC}: điện áp trên tụ C_1

I: dòng điện dương khi xả, âm khi nạp

Phương trình động học của phần tử RC:

$$\frac{dV_{RC}}{dt}=-\frac{1}{R_1.C_1}V_{RC}+\frac{1}{C_1}I\left(t\right) \quad (2)$$

2.3. Phương trình trạng thái SOC

SOC biểu diễn mức năng lượng còn lại của pin so với dung lượng định danh. Muốn mô phỏng pin bằng ECM ta phải cập nhật SOC theo thời gian vì OCV phụ thuộc SOC, R_0, R_1, C_1 cũng thay đổi theo SOC, điện áp mô phỏng chính xác phụ thuộc vào việc cập nhật SOC chính xác.

Trạng thái SOC của pin được mô tả bởi phương trình vi phân:

$$\frac{d\left(SOC\right)}{dt}=-\frac{\eta}{Q_n}.I_{bat}\left(t\right) \quad (3)$$

trong đó: Q_n: dung lượng pin ($Ah \rightarrow As$)

I_{bat}: dương khi xả, âm khi nạp.

η : hiệu suất coulombic (0,95 – 1,0)

Δt : bước thời gian (s)

Trong tính toán thực tế dữ liệu được đo theo chu kì Δt khi đó SOC được tính theo công thức:

$$SOC(k+1) = SOC(k) - \frac{I_{bat}\Delta t}{Q_n.\eta} \quad (4)$$

Hiệu suất Coulombic là tỉ lệ giữa dung lượng pin có thể nạp xả thực sự so với dung lượng lý thuyết. Nó cho biết mức độ mất mát điện tích trong mỗi chu kì nạp-xả.

$$\eta = \frac{Q_{ra}}{Q_{vào}} \quad (5)$$

Đối với pin Lithium-ion: $\eta \approx 0,98 - 0,995$ cho cell tốt.

3. Bộ dữ liệu đầu vào

3.1. Xây dựng dữ liệu đầu vào

Từ BMS, ta có thể thu được 3 chuỗi điện áp theo thời gian (thường lấy mẫu mỗi vài trăm mili-giây):

$V_{\max}(t)$: điện áp cao nhất trong các cell của module tại thời điểm t.

$V_{\min}(t)$: điện áp thấp nhất trong các cell của module tại thời điểm t.

$V_{mean}(t)$**:** điện áp trung bình của toàn module (tổng điện áp chia số cell)

Phương pháp này không cần biết dòng I(t), nhiệt độ T(t) hay nội trở, chỉ dựa trên điện áp.

3.2. Tính toán chênh lệch điện áp

Ta dựa vào công thức sau:

$$V_{diff}(t) = V_{\max}(t) - V_{\min}(t) \quad (6)$$

trong đó:

- $V_{diff}(t)$: thể hiện mức độ không cân bằng giữa các cell trong module.
- Khi các cell đều khỏe và SOC tương đương $V_{diff}(t) \approx 0$.
- Khi một cell yếu (sụt nhanh khi xả hoặc tăng nhanh khi sạc) thì $V_{diff}(t)$ sẽ tăng mạnh.

3.3. Phân tích trạng thái pin

Phát hiện mất cân bằng:

- Nếu $V_{diff}(t)$ vượt ngưỡng cố định hoặc động, module được xem là mất cân bằng điện áp.

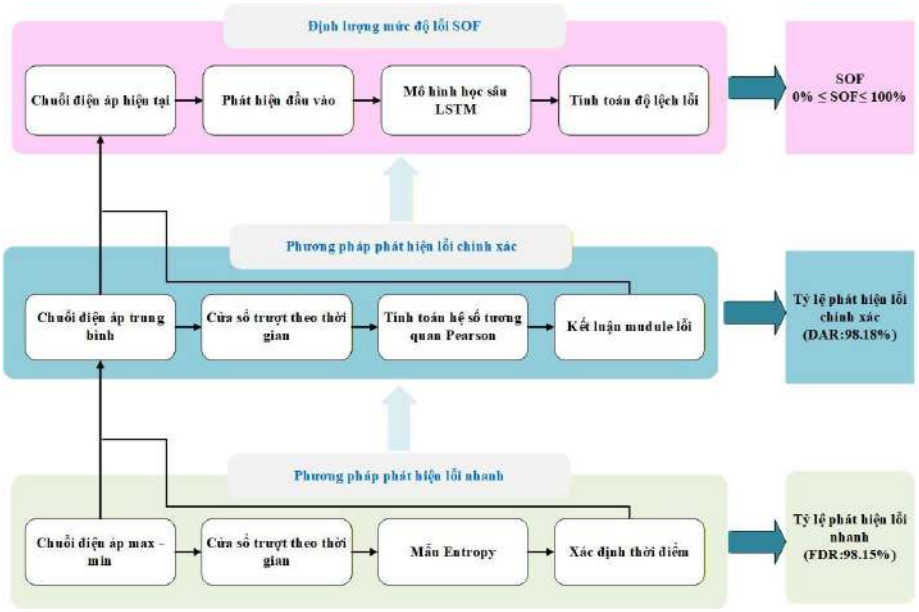
Chẩn đoán cell yếu hoặc lỗi:

- Nếu cùng một cell liên tục tạo ra $V_{\max}$ hoặc $V_{\min}$ trong nhiều chu kỳ, vậy cell đó bị sức chứa giảm hoặc nội trở tăng.

Phát hiện lỗi tiềm ẩn khi sạc xả:

- Khi sạc, nếu $V_{diff}(t)$ tăng nhanh hơn bình thường có thể một cell gần bão hòa trước.
- Khi xả, nếu $V_{diff}(t)$ tăng, cell yếu hết dung lượng sớm.

4. Công cụ lập trình và thuật toán



Hình 2. Lưu đồ chẩn đoán lỗi pin

4.1. Phương pháp phát hiện lỗi nhanh

Dựa trên dữ liệu đầu vào được tính bởi công thức (6):

$$V_{diff}(t) = V_{\max}(t) - V_{\min}(t)$$

đây là chuỗi điện áp chệnh lệnh của module pin do vậy sẽ biểu hiện mức độ mất cân bằng.

Khi module bình thường:

- $V_{diff}(t)$ biến thiên nhẹ, có dạng trơn, gần tuần hoàn (ổn định).
- Độ phức tạp thấp nên DSE nhỏ.

Khi module có lỗi (một cell yếu / nội trở cao / SOC lệch / hỏng cực):

- Chuỗi $V_{diff}(t)$ dao động bất thường, nhiều nhiễu hoặc nhẩy đột ngột.
- Độ phức tạp tăng mạnh nên DSE lớn.

Công thức tổng quát:

$$DSE = \left| SampEn\left(V_{diff}(t)_{fault}\right) - SampEn\left(V_{diff}(t)_{healthy}\right) \right| \quad (7)$$

nếu DSE lớn hơn ngưỡng sẽ đánh dấu đó là "pin lỗi".

4.2. Phương pháp phát hiện lỗi chính xác

- Chọn tập dữ liệu mà phương pháp phát hiện lỗi nhanh đã cảnh báo đó là vùng nghi lỗi.
- Tách tín hiệu X_t xác định của cell nghi lỗi.
- Tính Y_t trung bình của các cell còn lại sau khi đã bỏ cell nghi lỗi.
- Sử dụng công thức hệ số tương quan để tính ra giá trị $r_{X,Y}$.

$$r_{X,Y} = \frac{\sum_{i=1}^{N}\left(X_t - \overline{X}\right)\left(Y_t - \overline{Y}\right)}{\sqrt{\sum_{i=1}^{N}\left(X_t - \overline{X}\right)^2}\sqrt{\sum_{i=1}^{N}\left(Y_t - \overline{Y}\right)^2}} \quad (8)$$

trong đó:

- X_t : chuỗi điện áp cực trị (V_{max} hoặc V_{min}) của cell đang kiểm tra.
- Y_t: chuỗi điện áp trung bình (V_{mean}) của các cell còn lại trong module.
- Quy tắc chẩn đoán:

Bảng 1. Thông số đánh giá hệ số tương quan

Hệ số tương quan $r_{X,Y}$	**Mức tương quan**	**Kết luận chẩn đoán**
$r > 0,9$	Rất cao	Bình thường
$0,7 < r \le 0,9$	Cao	Có thể bình thường
$0,4 < r \le 0,7$	Trung bình	Cảnh báo sớm
$r \le 0,4$	Thấp	**Xác nhận lỗi (cell lỗi)**
$r < 0$	Ngược pha	**Lỗi nghiêm trọng**

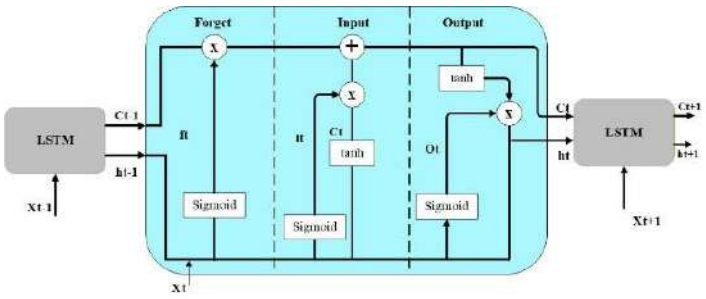
4.3. Định lượng mức độ lỗi SOF

Sau khi phát hiện lỗi nhanh và phát hiện lỗi chính xác, xác nhận cell lỗi, thì bước cuối cùng dùng mạng nơ-ron sâu DNN (Deep Neural Network) để định lượng mức độ lỗi SOF(State of Fault). SOF mô tả mức độ hư hại hoặc sai lệch hiện tại của cell.

$$SOF(t) \in [0,100\%] \quad (9)$$

trong đó: 0%: pin hoàn toàn bình thường.

100%: lỗi nặng hoặc cell mất khả năng hoạt động.



Hình 3. Mô hình cấu trúc mạng LSTM

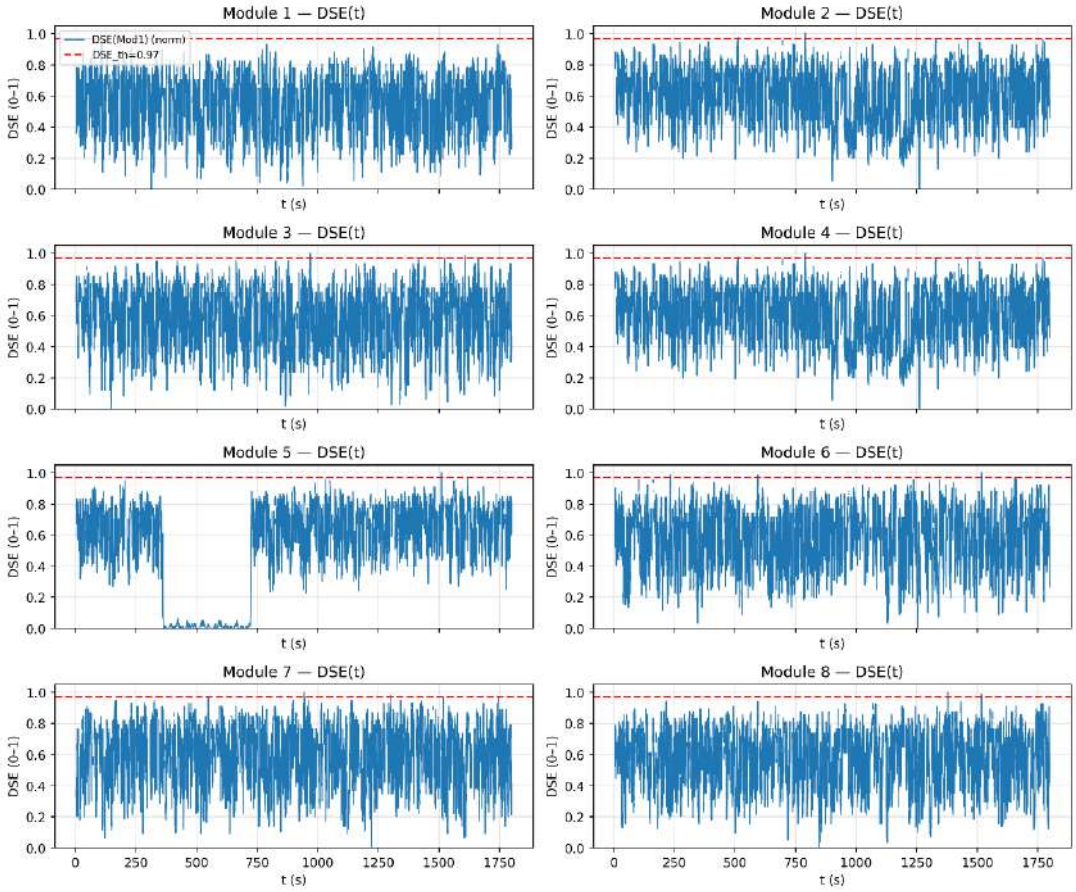
Sau khi có $v_{predict}(t)$ từ mô hình LSTM(Long Short-Term Memory) tiến hành so sánh với điện áp đo thực tế $V_{actual}(t)$, từ đó ta tính toán mức độ lỗi SOF theo công thức:

$$SOF(t) = k.\frac{\left|V_{predict}(t) - V_{actual}(t)\right|}{V_{predict}(t)} \times 100\% \quad (10)$$

trong đó: $V_{predict}(t)$: điện áp mô phỏng "chuẩn khỏe" do LSTM dự đoán.

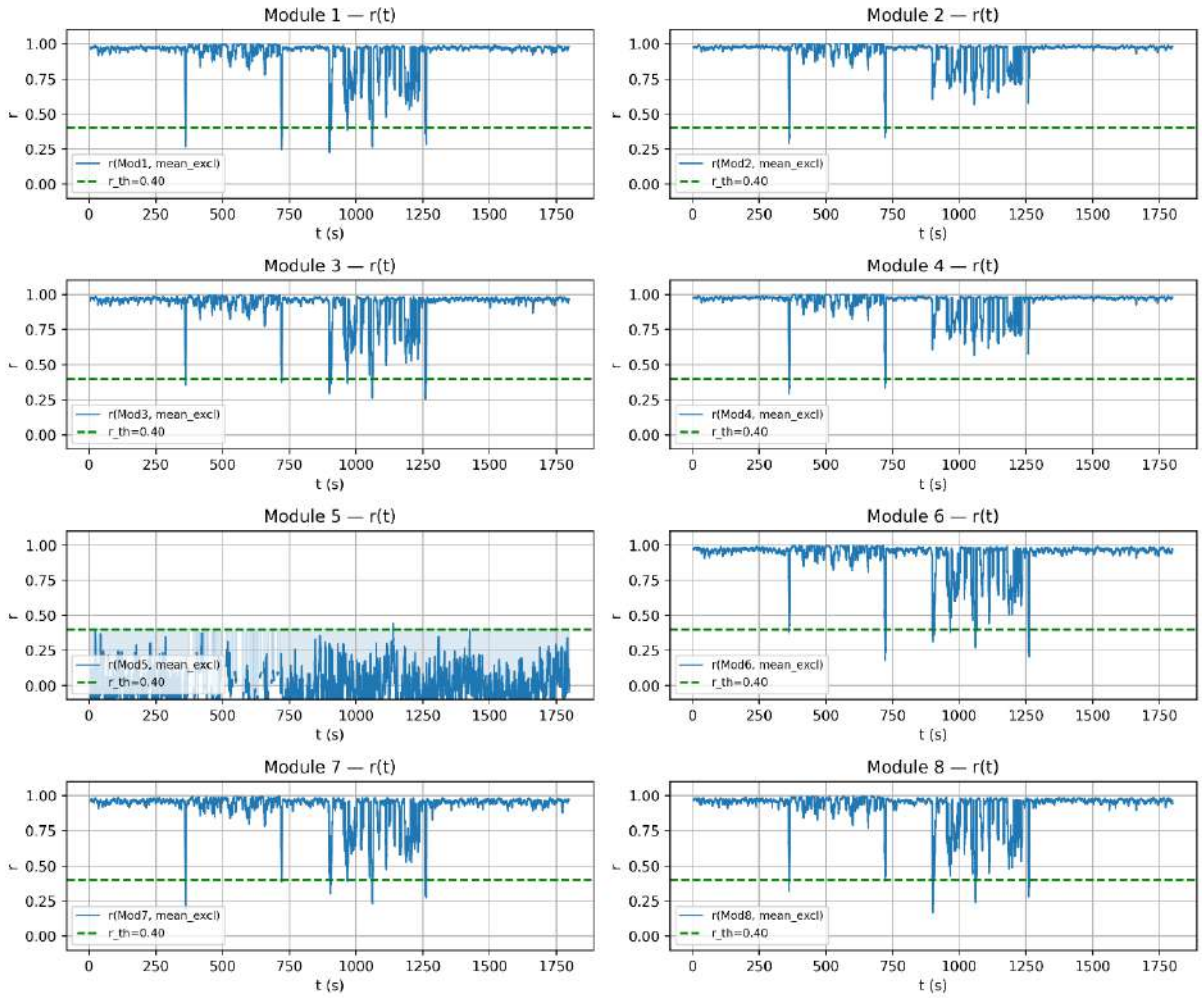
$V_{actual}(t)$: điện áp đo thực tế (có thể bị lỗi).

5. Kết quả mô phỏng



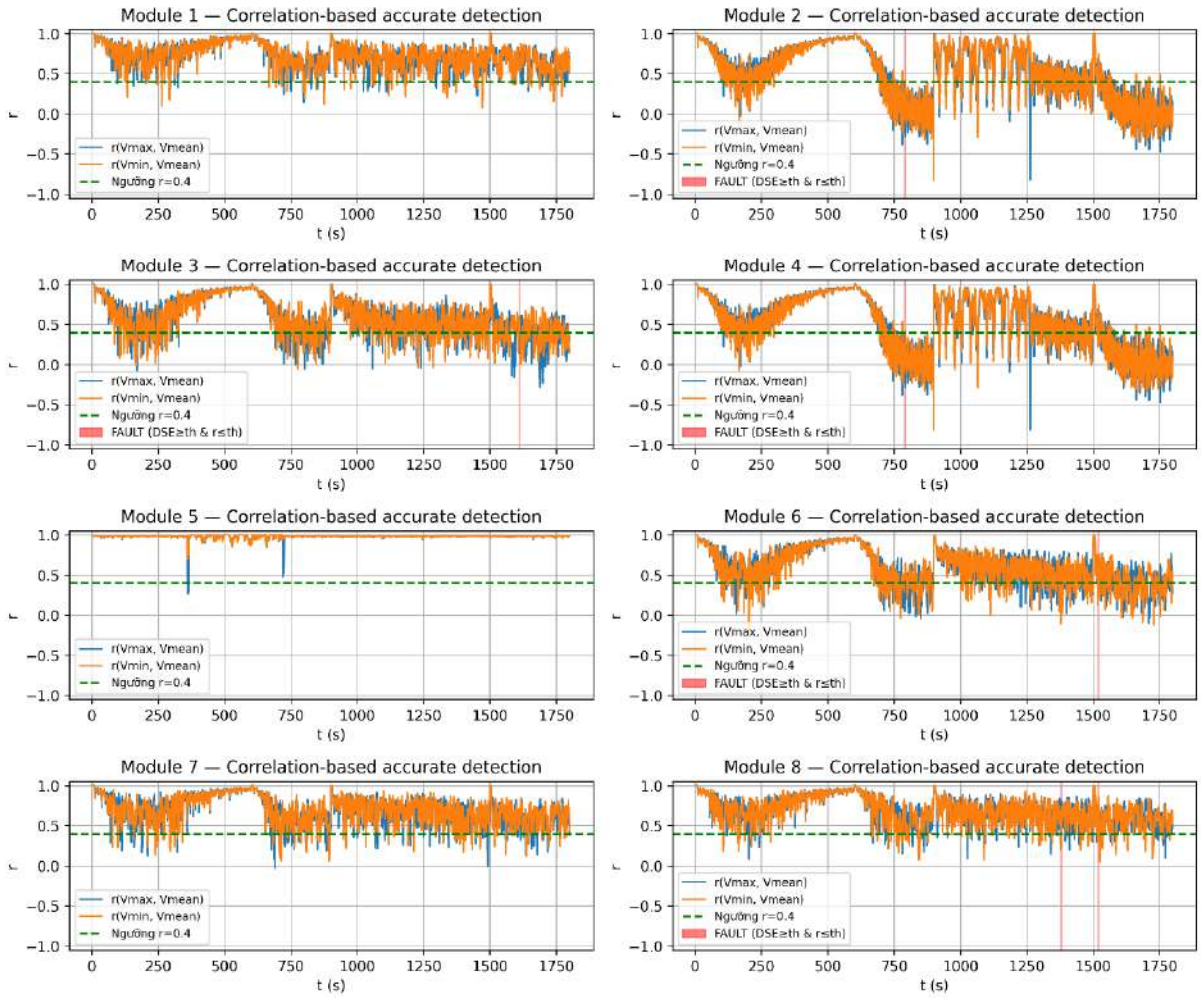
Hình 4. Kết quả đo ngưỡng DSE

Các module (1, 2, 3, 4, 6, 7, 8) có DSE(t) dao động quanh giá trị trung bình, đôi lúc xuất hiện sụt giảm ngắn hạn nhưng không duy trì ở mức rất thấp trong thời gian dài. Module 5 xuất hiện một khoảng thời gian DSE(t) giảm xuống gần 0 và kéo dài liên tục (khoảng giữa chu kỳ).



Hình 5. Kết quả kiểm tra so sánh điện áp từng module khi bỏ qua module lỗi

Module 1, 2, 3, 4, 6, 7, 8: r(t) duy trì rất cao (≈ 0,9–1,0) trong phần lớn thời gian. Chỉ xuất hiện các đáy ngắn hạn tập trung quanh giai đoạn từ 900 đến 1200 s. Đối với module 5: r(t) liên tục thấp, phần lớn nhỏ hơn 0,4 nhiều đoạn tiệm cận 0. Không có xu hướng hồi phục rõ ràng như các module khác.

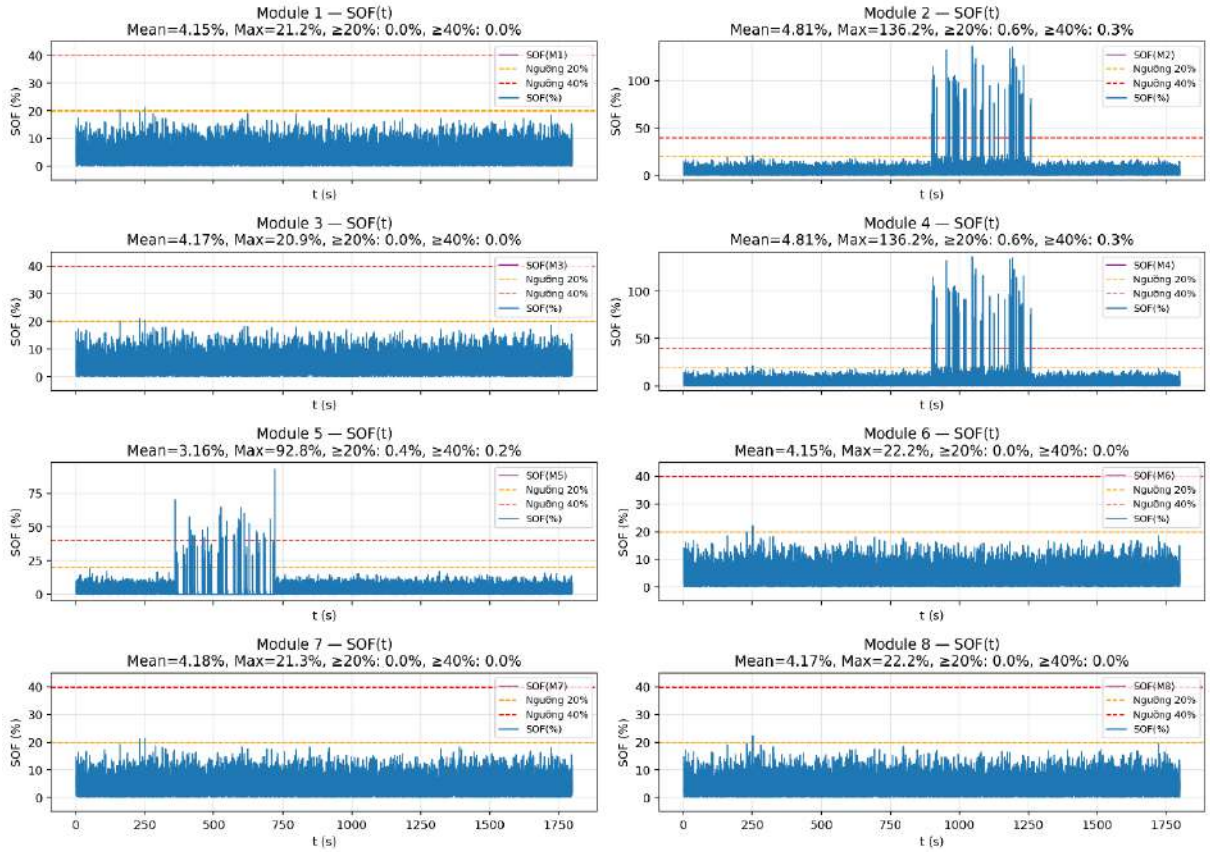


Hình 6. Đồ thị chẩn đoán chính xác lỗi

Ta có thể thấy module 1, 3, 7: r(V_{max}) và r(V_{min}) phần lớn hơn 0,4, dao động quanh 0,6 đến 0,9. Có các đoạn tụt ngắn khi chuyển chế độ, nhưng không có các đường đỏ lỗi xuất hiện. Có thể đánh giá khỏe, không phát hiện lỗi.

Đối với module 2, 4, 6, 8: r(V_{min}) giảm mạnh trong các giai đoạn động (≈ 700–1200 s), thậm chí xuống gần 0 hoặc âm. Tuy nhiên các đoạn này ngắn, rời rạc, vùng đỏ (FAULT) ít hoặc không kéo dài liên tục. Có thể nhận xét báo động lỗi.

Module 5: r(V_{max}) và r(V_{min}) duy trì rất cao, gần 1 trong toàn chu kỳ. Điều này cho thấy: Module 5 không còn "động học giống pack", mà gần như tách rời hành vi. Điện áp biến thiên theo một "quỹ đạo riêng", có thể thấy dấu hiệu sai lệch hệ thống. Kết hợp với r(t) thấp kéo dài, DSE cao kéo dài. Kết luận: module 5 được đánh dấu lỗi.



Hình 7. Đồ thị đánh giá hệ số suy giảm SOF

Module 1, 3, 6, 7, 8: Giá trị SOF trung bình thấp (≈ 4–5%), ổn định trong suốt chu kỳ. SOF hiếm khi vượt 20%, gần như không vượt 40%. Dao động nhỏ, phân bố đều phù hợp với module khỏe. Nhận xét: Các module này không có dấu hiệu suy giảm chức năng, hoạt động ổn định.

Đối với module 2 và module 4: SOF trung bình vẫn thấp (≈ 4,8%), nhưng xuất hiện các đỉnh SOF rất cao (Max ≈ 136%). Tập trung trong một khoảng thời gian ngắn Tỷ lệ thời gian: SOF ≥ 20% ≈ 0,6%, SOF ≥ 40% ≈ 0,3%. Nhận xét: Đây là bất thường cục bộ, ngắn hạn, có thể xem là cảnh báo.

Đối với module 5: SOF trung bình thấp (≈3,16%) nhưng cực đại rất cao (Max ≈ 92,8%) Xuất hiện nhiều đỉnh SOF vượt 20% và 40%, kéo dài thành cụm. Tỷ lệ: SOF ≥ 20% ≈ 0,4%, SOF ≥ 40% ≈ 0,2%. Đỉnh SOF dày hơn. Phân bố không đối xứng và không ngẫu nhiên. Nhận xét: module 5 suy giảm chức năng rõ rệt.

6. Kết luận

Kết luận rút ra từ nghiên cứu như sau:

- Bài báo đã trình bày một phương pháp mô hình hóa và chẩn đoán tình trạng bộ pin lithium-ion trên ô tô điện ở cấp độ module, dựa trên mô hình mạch tương đương ECM kết hợp với phân tích điện áp theo thời gian. Thông qua việc khai thác điện áp đầu cực và độ sai lệch điện áp giữa các module, một chỉ số sai lệch động đã được xây dựng tiền xử lý nhằm huấn luyện mô hình, nhận diện các trạng thái bất thường của pin trong các chế độ xả, nghỉ và nạp. Kết quả khi kiểm chứng cho thấy phương pháp đề xuất có khả năng phát hiện hiệu quả các module pin suy giảm và phản ánh rõ ràng hiện tượng mất cân bằng điện áp trong bộ pin.

- Trong các nghiên cứu tiếp theo, phương pháp sẽ được mở rộng ra toàn pack phù hợp với số lượng cell pin lớn của ô tô điện. Đồng thời, mô hình sẽ được mở rộng bằng cách tích hợp ảnh hưởng của nhiệt độ và cơ chế lão hóa pin nhằm nâng cao độ chính xác chẩn đoán. Bên cạnh đó, việc kết hợp các kỹ thuật học máy cũng sẽ được xem xét để cải thiện khả năng phân loại lỗi và tăng độ tin cậy của phương pháp trong các điều kiện làm việc phức tạp.

Tài liệu tham khảo

[1] Xuewen Tao, Xin Gu, Jinglun Li. (2025). A Fault Rapid – Accurate- Quantitative Integration Diagnosis Method for Lithium-Ion Battery Packs. IEEE Transactions On Industrial Electronics.

[2] Birkl, C., Howey, D., et al. (2015–2017). Degradation diagnostics for lithium-ion batteries based on ECM parameters. *Journal of Power Sources.*

[3] Nguyễn Hữu Cẩn, Phạm Xuân Dương, et al. Điện khí hóa trên ô tô – Lưu trữ năng lượng và hệ thống quản lý pin BMS. *NXB GTVT.*

[4] Nguyễn Trọng Hoan. Điện tử công suất và ứng dụng trong xe điện. *NXB KHKT.*

Research on diagnosing the technical condition of Lithium – ion battery packs using a Fault Rapid – Accurate - Quantitative integration method.

Abstract: *Lithium-ion battery packs are key components in electric vehicles, directly affecting system performance and safety. This thesis presents a diagnostic approach for lithium-ion battery packs based on an equivalent circuit model and data-driven indicators. An eight-module battery pack is analyzed under discharge, rest and charge operating phases. Voltage deviation and dynamic behavior indicators are used to identify abnormal modules. The results show that the proposed method can effectively detect faulty modules and distinguish them from healthy ones under dynamic operating conditions. This approach provides a practical basis for battery health monitoring and fault diagnosis in electric vehicle battery management systems.*

Keywords: lithium-ion battery, electric vehicle, fault diagnosis, equivalent circuit model, imbalance voltage.

Ứng dụng trí tuệ nhân tạo đánh giá tình trạng kỹ thuật của một số cảm biến trên ô tô thông qua trích xuất dữ liệu từ cổng OBD-II

Đinh Văn Tiến[1], Vũ Ngọc Tuấn[1]

[1]Học viện Kỹ thuật quân sự

Email: tienhvktqs0903@gmail.com; Tel: 0971189506

Tóm tắt

Bài báo đề xuất một hệ thống chẩn đoán thông minh ứng dụng trí tuệ nhân tạo để đánh giá tình trạng kỹ thuật của các cảm biến ô tô. Phương pháp đề xuất là sử dụng mạng nơ-ron tự mã hóa Autoencoder được huấn luyện từ bộ dữ liệu thu thập được từ các cảm biến khỏe có tình trạng kỹ thuật tốt nhằm phát hiện bất thường thông qua sai số tái tạo. Mô hình Autoencoder được tác giả xây dựng riêng cho dữ liệu cảm biến thu thập từ cổng OBD-II. Kết quả thực nghiệm cho thấy mô hình chứng minh khả năng tái tạo ổn định và nhận dạng chính xác các sai lệch cảm biến. Phương pháp này hướng tới tự động hóa chẩn đoán, phát hiện lỗi cảm biến và hỗ trợ bảo trì dự đoán, góp phần nâng cao độ tin cậy và hiệu quả vận hành của phương tiện.

Từ khóa: *Chẩn đoán cảm biến, Autoencoder, OBD-II, Trí tuệ nhân tạo.*

1. Đặt vấn đề

Trong những năm gần đây, trí tuệ nhân tạo và đặc biệt là học sâu đã trở thành công cụ mạnh mẽ trong lĩnh vực chẩn đoán và giám sát tình trạng kỹ thuật của ô tô. Việc ứng dụng các mô hình học sâu cho phép hệ thống tự động học và phát hiện các bất thường trong tín hiệu cảm biến mà không cần dữ liệu lỗi, góp phần nâng cao độ tin cậy và hiệu quả của quá trình bảo trì dự đoán. Các cảm biến cung cấp thông tin đầu vào cho bộ điều khiển trung tâm ECU (Electronic Control Unit). Khi một hoặc nhiều cảm biến hoạt động sai lệch, hệ thống điều khiển có thể đưa ra quyết định không chính xác, dẫn đến hiệu suất giảm, tiêu hao nhiên liệu tăng, thậm chí gây mất an toàn vận hành.

Trên thế giới nhiều nghiên cứu đã tập trung vào phát hiện và chẩn đoán lỗi cảm biến. Dereszewski và Sikora [4] sử dụng dao động tốc độ góc trục khuỷu để chẩn đoán động cơ. Mehranbod [5] và Suwatthikul [6] đề xuất các phương pháp phát hiện lỗi cảm biến. Trong khi Amin và Hasan [8,10] cùng Shahbaz và Amin [9] ứng dụng mạng nơ-ron nhân tạo trong điều khiển và chẩn đoán hệ thống cung cấp không khí nhiên liệu. Với sự phát triển của học sâu, Hinton và Salakhutdinov [7] cùng Goodfellow et al. [2] đã khẳng định hiệu quả của mô hình AE (Autoencoder) trong việc học đặc trưng dữ liệu và phát hiện bất thường. Đặc biệt Torkhesari và cộng sự [3] đã áp dụng AE cho giám sát sức khỏe cảm biến ô tô, đạt độ chính xác trên 99%.

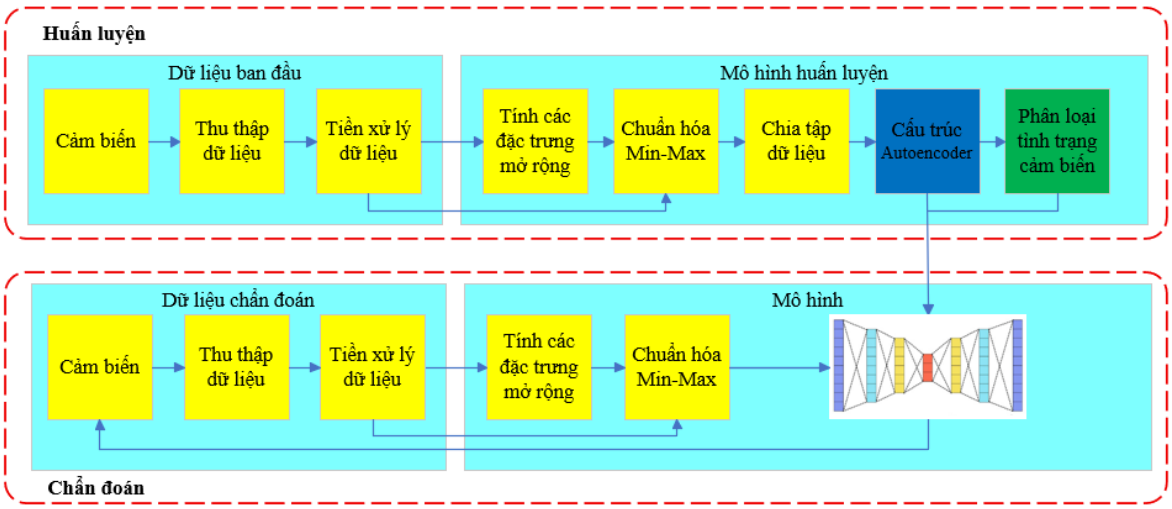
Tại Việt Nam các nghiên cứu ứng dụng AI trong đánh giá cảm biến ô tô còn hạn chế, công trình như của Lê Khánh Tân [1] mới dừng lại ở việc thu thập và hiển thị dữ liệu bằng Arduino và LabVIEW.

Trong lĩnh vực chẩn đoán cảm biến ô tô hiện nay, các thiết bị chẩn đoán truyền thống như máy đọc lỗi OBD-II (On Board Diagnostics-II) hay các hệ thống chuyên dụng của các hãng xe chủ yếu dựa vào mã lỗi DTC (Diagnostic Trouble Code) do ECU cung cấp. Tuy nhiên cách tiếp cận này còn nhiều hạn chế thực tiễn. Thứ nhất thiết bị chỉ phát hiện được khi cảm biến đã hư hỏng rõ rệt, không nhận biết được các dạng lỗi tiềm ẩn và việc chẩn đoán cần kỹ thuật viên được đào tạo có trình độ thì mới có thể vận hành thiết bị chẩn đoán. Thứ hai phương pháp này không thể đánh giá mức độ suy giảm kỹ thuật của cảm biến, mã DTC chỉ cho biết có lỗi hay không chứ không phản ánh cảm biến còn hoạt động tốt đến mức nào, có phản ứng chậm hay tín hiệu nhiễu hay không. Thứ

ba hệ thống chẩn đoán truyền thống không có khả năng tự học hoặc thích nghi: các ngưỡng đánh giá cố định theo nhà sản xuất, không thể điều chỉnh phù hợp với điều kiện vận hành, môi trường hay tuổi thọ xe ở Việt Nam. Cuối cùng các thiết bị này không phân tích được mối quan hệ giữa các tín hiệu cảm biến mà chỉ xử lý từng cảm biến riêng lẻ nên bỏ sót các dạng lỗi tương quan.

Từ các cơ sở trên nghiên cứu này đề xuất mô hình AE giúp phát hiện sai lệch qua sai số tái tạo đây là phần chênh lệch giữa giá trị dữ liệu đầu vào thực tế và dữ liệu mà mô hình tái tạo lại sau khi học và đánh giá tình trạng kỹ thuật của cảm biến ô tô. Mô hình có khả năng tự học hành vi bình thường của các cảm biến mà không cần dữ liệu lỗi, tự động phát hiện sai lệch nhỏ trong quan hệ giữa các tín hiệu của cảm biến cho phép đánh giá định lượng mức độ sức khỏe cảm biến thay vì chỉ báo lỗi hay không lỗi. Phương pháp này hướng tới tự động hóa chẩn đoán, phát hiện lỗi cảm biến và hỗ trợ bảo trì dự đoán, góp phần nâng cao độ tin cậy và hiệu quả vận hành của phương tiện.

2. Mô hình đề xuất



Hình 1. Mô hình ứng dụng trí tuệ nhân tạo chẩn đoán tình trạng kỹ thuật cảm biến

2.1. Thu thập dữ liệu

Dữ liệu được thu thập từ ECU của phương tiện thông qua giao thức OBD-II bằng bộ vi điều khiển kết nối máy tính. Các cảm biến được chọn thuộc nhóm có ảnh hưởng lớn đến vận hành động cơ như vòng quay trục khuỷu, tốc độ ô tô, áp suất khí nạp, nhiệt độ động cơ, nhiệt độ khí nạp, độ mở bướm ga, vị trí bàn đạp ga, cảm biến Oxi, độ hiệu chỉnh nhiên liệu ngắn hạn và độ hiệu chỉnh nhiên liệu dài hạn.

Thiết bị thu thập dữ liệu gồm bộ vi điều khiển Arduino Uno R3 kết hợp với mạch CAN Bus Shield kết nối trực tiếp vào cổng OBD-II của xe. Bộ điều khiển được lập trình bằng phần mềm Arduino IDE sử dụng thư viện mcp_can.h để gửi và nhận khung tin CAN theo chuẩn ISO 15765-4. Tốc độ lấy mẫu trung bình 150 ms/mẫu trong chế độ vận hành ổn định và được đồng bộ bằng nhãn thời gian. Dữ liệu được ghi và truyền qua cổng USB về máy tính xách tay để lưu dưới định dạng CSV, giúp đảm bảo tính liên tục cho các phân tích sau này.

2.2. Tiền xử lý dữ liệu

Trước khi đưa vào mô hình học sâu dữ liệu cảm biến thô cần được xử lý một cách hệ thống nhằm loại bỏ nhiễu và sai lệch từ đó đảm bảo mô hình có thể học được chính xác mối quan hệ nội tại giữa các tín hiệu.

Giai đoạn làm sạch nhằm đảm bảo dữ liệu đầu vào có độ tin cậy cao và phản ánh chính xác trạng thái vận hành thực tế của hệ thống cảm biến. Trước hết các giá trị bị thiếu được phát hiện bằng phương pháp Masking hoặc trùng lặp được loại bỏ đồng thời nội suy tuyến tính được áp dụng cho các khoảng thiếu ngắn nhằm duy trì tính liên tục của chuỗi thời gian. Tiếp theo các giới hạn vật lý được áp dụng cho từng cảm biến để loại bỏ giá trị không thực tế. Cuối cùng nhằm giảm ảnh hưởng của nhiễu tần số cao, tín hiệu được làm trơn bằng phương pháp trung bình trượt với cửa sổ 8 mẫu liên tiếp.

Nhờ đó dữ liệu đầu vào trở nên ổn định hơn, giảm biến động ngẫu nhiên và giữ nguyên được xu hướng vật lý của tín hiệu.

2.3. Tính các đặc trưng mở rộng

Sau khi dữ liệu được làm sạch, quá trình trích xuất đặc trưng mở rộng được thực hiện nhằm mô tả toàn diện hơn hành vi vật lý và mối tương quan giữa các cảm biến. Bên cạnh các giá trị đo trực tiếp từ cảm biến, mô hình còn sử dụng các đặc trưng được tính toán bổ sung nhằm tăng cường khả năng biểu diễn của tập dữ liệu.

Các đặc trưng đạo hàm bậc nhất được tính cho các cảm biến như tốc độ động cơ, áp suất khí nạp, vị trí bướm ga, tốc độ xe và nhiệt độ khí nạp nhằm phản ánh tốc độ biến thiên tín hiệu, hỗ trợ phát hiện các hiện tượng bất thường như cảm biến kẹt hoặc phản ứng chậm.

Các đặc trưng thống kê cục bộ bao gồm trung bình trượt và độ lệch chuẩn trượt, được tính trên các cửa sổ ngắn để biểu diễn xu hướng và độ dao động tức thời của tín hiệu qua đó mô hình có thể nhận biết các thay đổi bất thường trong khoảng thời gian ngắn.

Sau khi tính các đặc trưng mở rộng thì bộ dữ liệu cảm biến thu thập được chuyển thành bộ đặc trưng đầu vào cho mô hình AE bao gồm dữ liệu các giá trị của cảm biến được thu thập và các đặc trưng mở rộng.

2.4. Chuẩn hóa dữ liệu

Bộ dữ liệu đặc trưng đầu vào tiếp tục được chuẩn hóa do mỗi cảm biến có đơn vị và dải giá trị khác nhau, toàn bộ đặc trưng được chuẩn hóa tuyến tính về khoảng [0,1] bằng phương pháp Min–Max scaling và được xác định theo công thức (1):

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{1}$$

Quá trình chuẩn hóa giúp cân bằng ảnh hưởng của từng đặc trưng, tránh việc các biến có biên độ lớn chi phối mô hình đồng thời cải thiện tốc độ hội tụ của thuật toán tối ưu trong giai đoạn huấn luyện. Ngoài ra việc đưa tất cả đặc trưng về cùng một thang giá trị còn giúp mô hình học sâu dễ dàng nắm bắt mối quan hệ tương đối giữa các cảm biến thay vì phụ thuộc vào giá trị tuyệt đối.

2.5. Chia tập dữ liệu

Sau khi chuẩn hóa, dữ liệu được chia thành ba phần độc lập phục vụ cho các giai đoạn huấn luyện, xác thực và kiểm thử mô hình. Cụ thể 67% dữ liệu được dùng để huấn luyện, 25% của phần này được tách ra làm tập xác thực nhằm giám sát quá trình học tránh hiện tượng học thuộc dữ liệu và 33% còn lại được dùng để đánh giá khả năng tổng quát hóa của mô hình.

Sau khi chia, các tập dữ liệu đã được xử lý đầy đủ qua các bước làm sạch, chuẩn hóa và trích xuất đặc trưng, bao gồm cả các giá trị cảm biến gốc và các đặc trưng mở rộng. Bộ dữ liệu này trở thành đầu vào hoàn chỉnh cho mô hình AE, cho phép mạng học được mẫu hành vi bình thường của

hệ thống cảm biến và nhận diện các sai lệch dựa trên sai số tái tạo trong quá trình đánh giá.

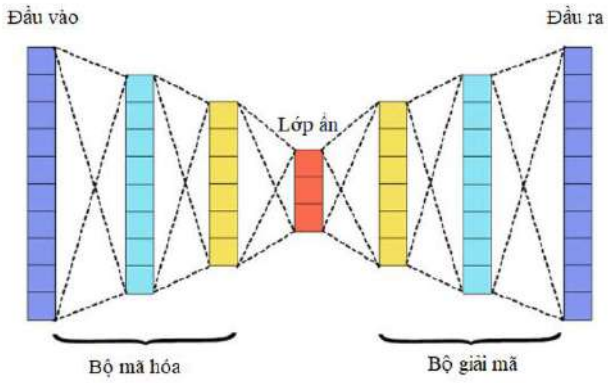
2.6. Mô hình Autoencoder

Autoencoder là một loại mạng nơ-ron nhân tạo học sâu thuộc nhóm học phi giám sát, được sử dụng rộng rãi trong các bài toán giảm chiều dữ liệu, trích xuất đặc trưng và phát hiện bất thường. Mô hình AE được lựa chọn trong nghiên cứu này nhờ khả năng học đặc trưng phi giám sát, cho phép hệ thống tự động học biểu diễn của các tín hiệu cảm biến khỏe mạnh mà không cần dữ liệu lỗi có nhãn. Mô hình AE được xây dựng nhằm học biểu diễn ẩn của dữ liệu bằng cách nén và tái tạo lại đầu vào thông qua hai khối chính: bộ mã hóa và bộ giải mã .

Bộ mã hóa nén dữ liệu cảm biến đa chiều về không gian tiềm ẩn có kích thước nhỏ hơn giúp mô hình học được các mối quan hệ phi tuyến giữa các cảm biến. Bộ giải mã thực hiện quá trình tái tạo dữ liệu đầu vào từ biểu diễn ẩn, qua đó cho phép tính sai số tái tạo theo công thức (3) làm cơ sở phát hiện bất thường.

Nguyên lý phát hiện lỗi bằng AE: Mô hình AE được huấn luyện chỉ với dữ liệu "khỏe mạnh" của cảm biến tức là huấn luyện với dữ liệu khi cảm biến hoạt động bình thường. Nhờ đó mạng học được phân bố "chuẩn" của các đặc trưng cảm biến khỏe mạnh và cố gắng tái tạo đầu ra thật sát với đầu vào. Khi xuất hiện dữ liệu bất thường do cảm biến suy giảm hoặc hỏng mô hình sẽ không thể tái tạo chính xác giá trị đó do nó nằm ngoài kinh nghiệm học được. Kết quả là sai số tái tạo tăng lên đột biến. Đặc tính này cho phép dùng sai số tái tạo làm chỉ báo để phát hiện cảm biến bất thường.

Từ cơ sở lý thuyết về mạng nơ ron tự mã hóa và đặc điểm của dữ liệu cảm biến ô tô thu thập qua cổng OBD-II, việc xác định cấu trúc của mô hình AE cần bảo đảm sự phù hợp cả về mặt khoa học lẫn thực tiễn.



Hình 2. Cấu trúc Autoencoder

Trong nghiên cứu này mô hình được nhóm tác giả thiết kế theo cấu trúc đối xứng tức là số lớp và số nơ ron của bộ mã hóa và bộ giải mã được bố trí theo nguyên tắc phản chiếu.

Số lớp ẩn được chọn là 3 dựa trên cân bằng giữa khả năng biểu diễn phi tuyến và độ ổn định huấn luyện. Thực nghiệm trên dữ liệu OBD-II cho thấy với 3 lớp, AE có thể mô hình hóa tốt các quan hệ phi tuyến giữa các tín hiệu cảm biến mà vẫn duy trì khả năng tổng quát hóa trên dữ liệu mới.

Số nơ-ron giảm dần ở từng lớp được lựa chọn theo nguyên tắc nén dần thông tin với tỷ lệ 25–35% mỗi tầng như khuyến nghị trong các nghiên cứu AE cho dữ liệu cảm biến. Việc giảm dần số nơ ron buộc mô hình học cách tổng hợp và rút trích các đặc trưng quan trọng, tránh hiện tượng học đồng nhất tức mô hình chỉ sao chép dữ liệu đầu vào mà không trích xuất được đặc trưng ẩn.

Kích thước lớp ẩn trung tâm được chọn bằng khoảng 30–40% số đặc trưng đầu vào. Đây

là tỷ lệ được chứng minh phù hợp với dữ liệu dạng bảng vì vừa đảm bảo khả năng nén thông tin hiệu quả vừa tránh mất mát dữ liệu quan trọng.

Dựa vào các cơ sở trên nhóm tác giả đề xuất cấu trúc AE được thiết kế như một mạng nơ-ron đối xứng, bao gồm ba tầng mã hóa và ba tầng giải mã với cấu trúc tổng quát như sau:

$$D \rightarrow 0,7D \rightarrow 0,5D \rightarrow 0,35D \rightarrow 0,5D \rightarrow 0,7D \rightarrow D$$

Trong đó D là số đặc trưng đầu vào.

Các kỹ thuật BatchNorm, Gaussian Noise, Dropout, Huber Loss được dùng để ổn định huấn luyện, tránh hiện tượng ghi nhớ dữ liệu huấn luyện và nâng cao khả năng tổng quát. Quá trình huấn luyện sử dụng thuật toán tối ưu Adam với tốc độ học là 0,001, số vòng lặp (epoch) là 400, batch size là 256, đồng thời áp dụng cơ chế EarlyStopping dừng sớm nhằm tránh hiện tượng ghi nhớ dữ liệu huấn luyện và cơ chế ReduceLROnPlateau sẽ giảm tốc độ học xuống một nửa nếu mô hình đang tiến bộ chậm.

Hiệu quả của mô hình được đánh giá qua hai chỉ tiêu định lượng:

- Validation Mean Squared Error (Val MSE) là sai số bình phương trung bình trên tập xác thực đây thể hiện khả năng hội tụ ổn định. Nó có vai trò đánh giá chất lượng mô hình trong quá trình huấn luyện.

- Test Mean Squared Error (Test MSE) là sai số bình phương trung bình trên tập kiểm thử phản ánh khả năng khái quát hóa mô hình. Nó có vai trò đánh giá khả năng tổng quát hóa của mô hình, kiểm chứng mô hình có hoạt động tốt đối với dữ liệu mới chưa từng thấy hay không.

$$\text{Test MSE} = \frac{1}{N_{test}.D} \sum_{i=1}^{N_{test}} \sum_{j=1}^{D} (X_{ij} - \hat{X}_{ij})^2 \quad (2)$$

Mô hình được coi là đạt chất lượng tốt khi cả hai chỉ số đều nhỏ và có giá trị xấp xỉ nhau điều này cho thấy mô hình vừa học hiệu quả vừa tránh hiện tượng ghi nhớ dữ liệu khi học.

2.7. Phân loại tình trạng cảm biến

Sau khi hoàn tất huấn luyện, mô hình AE được sử dụng để chẩn đoán tình trạng cảm biến thông qua sai số tái tạo giữa tín hiệu đầu vào và tín hiệu tái tạo. Cụ thể mô hình nhận dữ liệu đầu vào X, tái tạo lại đầu ra $\hat{X}$ và tính sai số tái tạo theo công thức:

$$E = \left| X - \hat{X} \right| \quad (3)$$

Giả định rằng sai số tái tạo của các mẫu trong tập dữ liệu khỏe mạnh tuân theo phân phối chuẩn (Gaussian) như hình 3. Dựa trên phân bố này các ngưỡng phát hiện bất thường được xác lập dựa trên trị số trung bình μ và bội số của độ lệch chuẩn σ. Khi sai số tái tạo vượt quá một số bội số của σ thì điểm dữ liệu được xem là bất thường. Từ đó tình trạng cảm biến được chia thành năm mức định tính tương ứng với độ lệch chuẩn và trị số trung bình của sai số tái tạo:

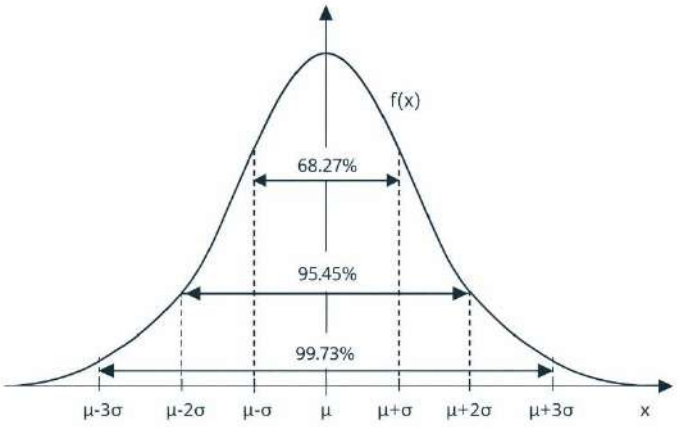
1. Rất tốt: $E < \mu + \sigma$.

2. Tốt: $\mu + \sigma < E < \mu + 2\sigma$.

3. Bình thường: $\mu + 2\sigma < E < \mu + 3\sigma$.

4. Cảnh báo: $\mu + 3\sigma < E < \mu + 4\sigma$.

5. Lỗi: $E \geq \mu + 4\sigma$.



Hình 3. Phân phối chuẩn

Việc xác định ngưỡng dựa trên thống kê phân bố sai số tái tạo của dữ liệu khỏe mạnh giúp hệ thống thích ứng tự động với từng loại cảm biến mà không cần thiết lập thủ công.

Để định lượng mức độ suy giảm kỹ thuật của từng cảm biến, một chỉ số sức khỏe tổng hợp SHI được xây dựng, được tính theo công thức (4):

$$SHI=1-\frac{E}{E_{max}} \tag{4}$$

Trong đó E là sai số giữa giá trị thực và giá trị tái tạo ước lượng, E_{max} là ngưỡng sai số cực đại cho phép. SHI nằm trong khoảng [0,1], giá trị càng gần 1 thì cảm biến càng ổn định.

3. Triển khai thử nghiệm

3.1. Xác định bộ dữ liệu đầu vào ban đầu

Dữ liệu được thu thập thông qua thiết bị thu thập dữ liệu được kết nối trực tiếp vào cổng OBD-II của xe ghi nhận liên tục các dữ liệu cảm biến cần thiết như tốc độ động cơ, áp suất khí nạp, lưu lượng khí nạp, tốc độ xe, vị trí bàn đạp ga, góc mở bướm ga, nhiệt độ khí nạp, nhiệt độ động cơ, điện áp cảm biến oxy, mức hiệu chỉnh nhiên liệu ngắn hạn, mức hiệu chỉnh nhiên liệu dài hạn với tần số lấy mẫu ổn định theo thời gian thực. Trước khi tiến hành thu thập, xe được khởi động và để hoạt động ở chế độ không tải trong khoảng 5–10 phút để động cơ đạt trạng thái ổn định nhiệt. Quá trình làm ấm được xem là hoàn tất khi nhiệt độ nước làm mát động cơ đạt khoảng 85–95°C, vòng tua không tải ổn định. Sau khi động cơ đã ổn định thiết bị thu thập dữ liệu được trình bày ở mục 2.1 được kết nối với cổng OBD-II của xe. Thiết bị chỉ bắt đầu ghi dữ liệu khi xe chuẩn bị lăn bánh nhằm loại bỏ các nhiễu phát sinh trong giai đoạn khởi động ban đầu. Các đợt thu thập được thực hiện trên cùng một hành trình cùng một điều kiện môi trường. Dữ liệu được ghi lại theo chế độ vận hành ổn định.

Xe được vận hành ở tốc độ ổn định trong dải 40–60 km/h trên đoạn đường bằng phẳng. Dữ liệu trong giai đoạn này thể hiện rõ mối tương quan giữa các cảm biến. Đây là điều kiện điển hình của hành trình đường trường, nơi động cơ làm việc với tải trung bình và ít dao động. Thông qua kịch bản này, mô hình AE được huấn luyện để nhận biết quan hệ ổn định giữa các cảm biến trong chế độ vận hành ổn định.

3.2. Tập dữ liệu

Dữ liệu được thu thập trên xe Hyundai Elantra 2020 trong điều kiện vận hành thực tế để phản ánh đặc trưng hoạt động của hệ thống cảm biến. Trong quá trình thu thập dữ liệu thực nghiệm từ

ECU qua cổng OBD-II trên xe Hyundai Elantra 2020 dữ liệu xuất hiện các giá trị thiếu và sai lệch do đặc thù giao tiếp CAN bao gồm mất gói, trễ phản hồi và lỗi truyền thông. Ngoài ra một số cảm biến trong danh sách PID không được ECU hỗ trợ hoặc không phản hồi ổn định dẫn đến các cột dữ liệu bị trống đồng thời tín hiệu cảm biến đôi lúc bị nhiễu hoặc xuất hiện giá trị bất thường. Để đảm bảo chất lượng dữ liệu đầu vào, cả hai tập dữ liệu được đề cập dưới đây được tiền xử lý bằng cách loại bỏ các cảm biến không có ý nghĩa vật lý, xử lý giá trị thiếu bằng phương pháp loại bỏ hoặc nội suy theo thời gian và phát hiện loại bỏ các giá trị ngoại lai. Sau tiền xử lý dữ liệu thu được gồm các thông số ở bảng 1 và lúc này tập dữ liệu đã có tính liên tục, ổn định và phù hợp cho quá trình huấn luyện, đánh giá mô hình chẩn đoán tình trạng kỹ thuật dựa trên trí tuệ nhân tạo.

Bảng 1. Các thông số thu được sau khi tiền xử lý dữ liệu

TT	Mã PID	Thông số theo thời gian	Kí hiệu	Đơn vị
1	0C	Tốc độ động cơ	rpm	Vòng/phút
2	0F	Nhiệt độ không khí nạp	iat	ºC
3	05	Nhiệt độ động cơ	coolant	ºC
4	0B	Áp suất không khí nạp	map	Kpa
5	11	Góc mở bướm ga	throttle	%
6	14	Cảm biến Oxy 1	o2v_v	V
7	0D	Tốc độ xe	speed	Km/h
8	06	Hiệu chỉnh nhiên liệu ngắn hạn	stft	%
9	07	Hiệu chỉnh nhiên liệu dài hạn	ltft_b1	%

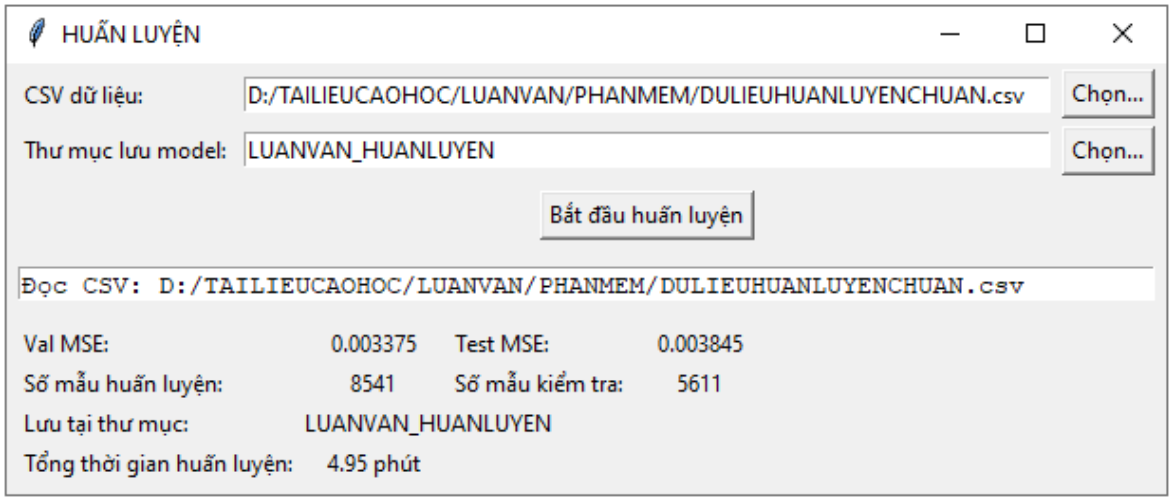
Để đánh giá toàn diện khả năng phát hiện bất thường của mô hình Autoencoder, nhóm tác giả tiến hành thu thập hai tập dữ liệu riêng biệt:

Tập dữ liệu thứ nhất được ghi trong điều kiện xe hoạt động bình thường, không có đèn cảnh báo MIL và không ghi nhận sai lệch tín hiệu trên bất kỳ cảm biến nào. Thu thập dữ liệu theo kịch bản chế độ vận hành ổn định. Dữ liệu này đại diện cho hành vi chuẩn khỏe mạnh của hệ thống và được dùng để huấn luyện mô hình AE. Tổng thời lượng ghi nhận khoảng hơn 40 phút với hơn 17000 mẫu.

Tập dữ liệu thứ hai được thu trong cùng điều kiện vận hành nhưng cảm biến áp suất khí nạp và cảm biến Oxi được cố tình gây lỗi để mô phỏng hiện tượng cảm biến bị bẩn trong tình huống thực tế. Các tín hiệu khác vẫn hoạt động bình thường, cho phép mô hình học phát hiện sự sai lệch tương quan giữa cảm biến áp suất khí nạp, cảm biến Oxi và các cảm biến còn lại. Tập này không được sử dụng để huấn luyện mà dùng chẩn đoán năng lực phát hiện lỗi của AE. Tổng thời gian ghi nhận khoảng hơn 15 phút với hơn 6500 mẫu.

3.3. Huấn luyện mô hình

Mô hình AE được huấn luyện trên tập dữ liệu thứ nhất được trình bày mục 3.2. Quá trình huấn luyện mô hình được tiến hành trên máy tính với bộ xử lý Intel(R) Core i7-6400U có tốc độ 3,20 GHz và RAM 16 GB. Trong bài báo nhóm tác giả huấn luyện mô hình Autoencoder bằng phần mềm huấn luyện được thiết kế chuyên dụng cho mô hình. Ứng dụng huấn luyện được xây dựng bằng ngôn ngữ Python. Ứng dụng này cho phép tự động hóa toàn bộ quy trình từ tiền xử lý, tính toán các đặc trưng mở rộng, chuẩn hóa dữ liệu, chia tập dữ liệu và huấn luyện đến đánh giá mô hình, đảm bảo khả năng tái lập và mở rộng cho nhiều loại phương tiện khác nhau. Ứng dụng cho phép tự động hóa toàn bộ quy trình từ nhập dữ liệu đến lưu trữ mô hình và siêu tham số.



Hình 4. Kết quả huấn luyện được hiển thị trong ứng dụng huấn luyện

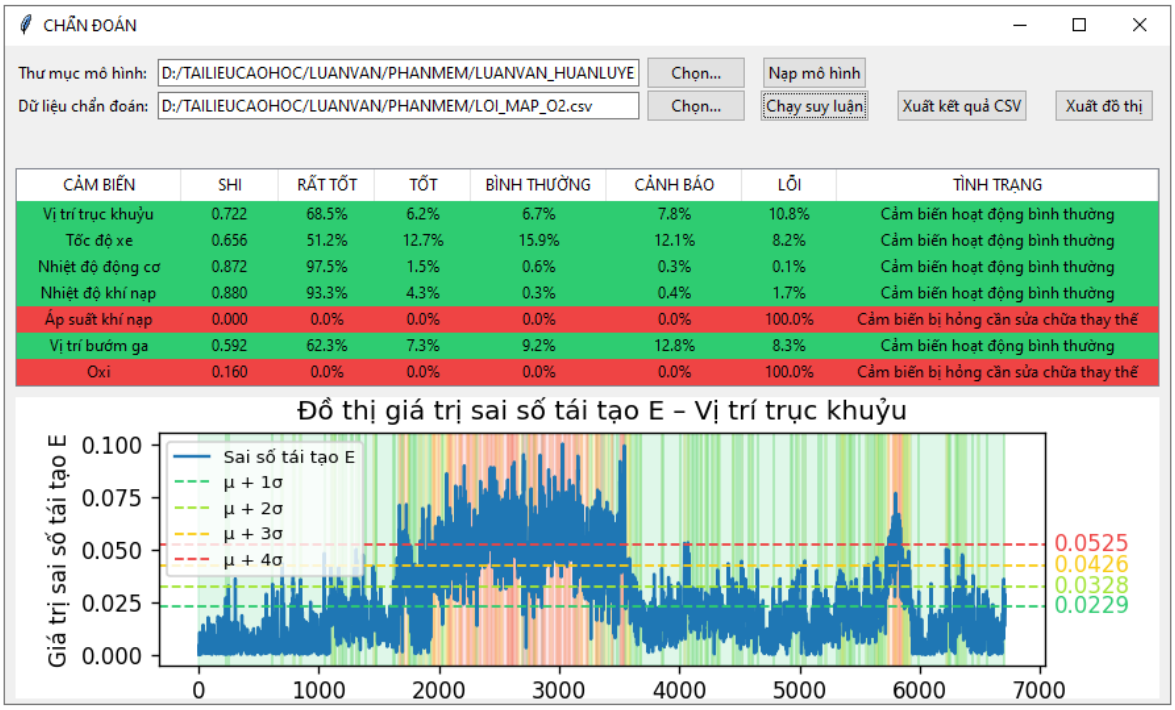
Trong thử nghiệm, kết quả theo như hình 4 mô hình Autoencoder đạt Val MSE = 0,003375 và Test MSE = 0,003845 cho thấy đây là một mô hình hiệu quả vì các giá trị MSE ổn định, thấp và sự chênh lệch giữa Val MSE và Test MSE không vượt quá 15% từ đó sai số tái tạo nhỏ, ổn định và đảm bảo năng lực học sâu của mạng mà không bị tránh hiện tượng ghi nhớ dữ liệu huấn luyện. Sau khi huấn luyện toàn bộ trọng số, cấu hình chuẩn hóa và ngưỡng đánh giá được lưu trữ tự động dưới dạng các tệp mô hình, phục vụ cho giai đoạn chẩn đoán cảm biến trong môi trường vận hành thực tế.

4. Chẩn đoán và thảo luận kết quả

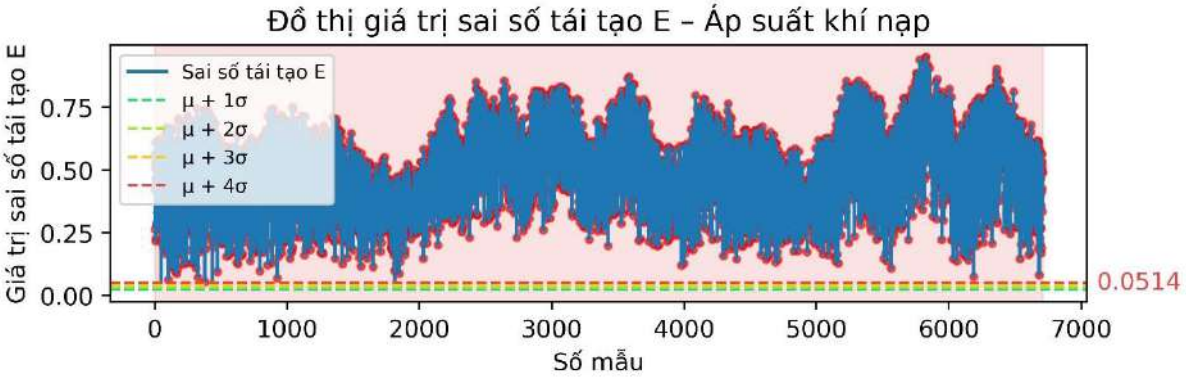
Hệ thống chẩn đoán được xây dựng và triển khai hoàn chỉnh bằng Python, sử dụng mô hình Autoencoder đã huấn luyện trong mục 3.3 để đánh giá tình trạng kỹ thuật của các cảm biến trên xe. Quá trình chẩn đoán sử dụng tập dữ liệu thứ hai được trình bày tại mục 3.2 để thử nghiệm mô hình. Sau khi khởi chạy hệ thống tự động thực hiện toàn bộ quy trình tiền xử lý, tính đặc trưng mở rộng, chuẩn hóa, tính sai số, so sánh với ngưỡng và hiển thị kết quả chẩn đoán.

Kết quả chẩn đoán được thể hiện trực quan qua bảng tổng hợp và đồ thị giá trị sai số tái tạo. Bảng kết quả cung cấp các thông tin định lượng về từng cảm biến gồm: giá trị trung bình SHI, tỉ lệ phần trăm các mức Rất tốt – Tốt – Bình thường – Cảnh báo – Lỗi và kết luận định tính "Tình trạng".

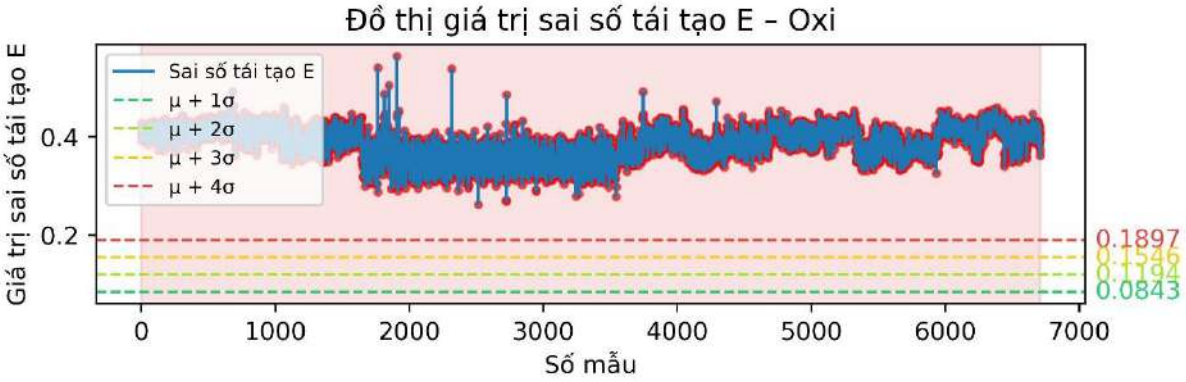
Trong thử nghiệm chẩn đoán trên xe Hyundai Elantra 2020 kết quả chẩn đoán được thể hiện rõ ràng tại hình 5. Bảng tổng hợp kết quả trong hình 5 cho thấy mô hình AE đã xác định rõ tình trạng kỹ thuật của từng cảm biến. Các cảm biến vị trí trục khuỷu, tốc độ xe, nhiệt độ động cơ, nhiệt độ khí nạp và vị trí bướm ga có chỉ số SHI nằm trong vùng an toàn với tỷ lệ "Rất tốt" và "Tốt" chiếm ưu thế, phản ánh trạng thái hoạt động nhìn chung ổn định. Tuy nhiên cảm biến áp suất khí nạp và cảm biến Oxy được xác định ở trạng thái lỗi hoàn toàn với tỷ lệ lỗi đạt 100% và SHI trung bình đạt rất thấp cụ thể bằng 0 và 0,16. Việc hai cảm biến quan trọng trong hệ thống điều khiển động cơ bị hư hỏng không chỉ ảnh hưởng trực tiếp đến tín hiệu của chính chúng mà còn gây sai lệch gián tiếp đến quá trình tính toán hòa khí, đánh lửa và điều khiển động cơ từ đó tác động nhất định đến đặc tính tín hiệu của các cảm biến liên quan khác.



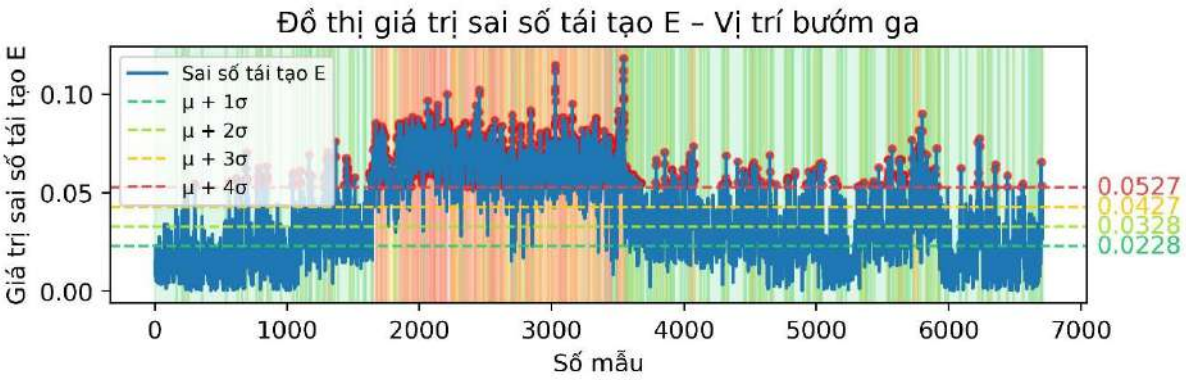
Hình 5. Kết quả chẩn đoán được hiển thị trong ứng dụng chẩn đoán



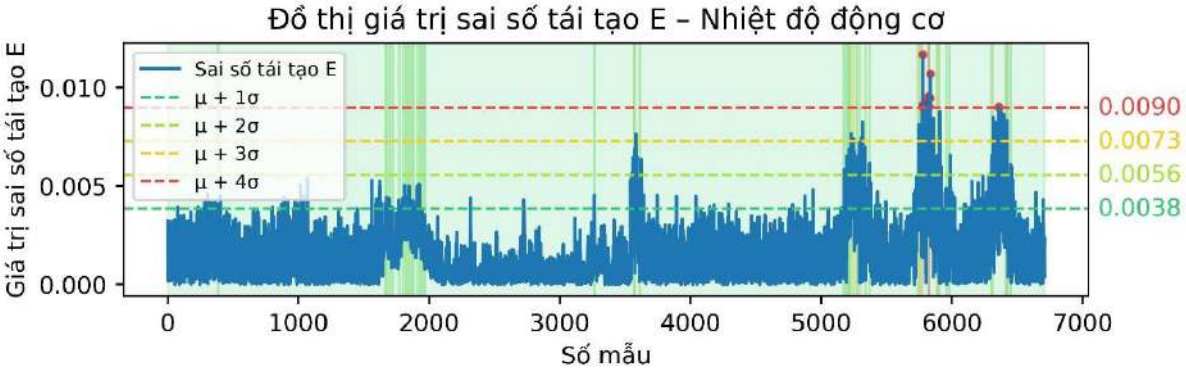
Hình 6. Đồ thị giá trị sai số tái tạo của cảm biến áp suất khí nạp



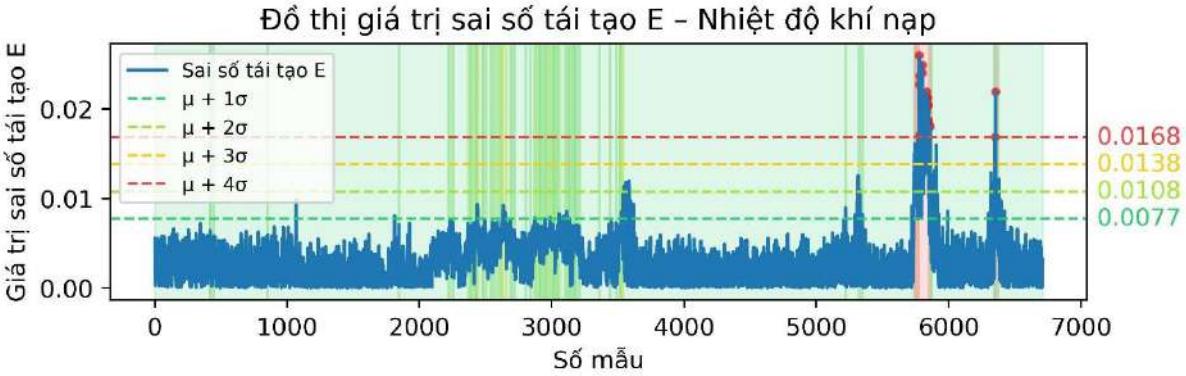
Hình 7. Đồ thị giá trị sai số tái tạo của cảm biến Oxi



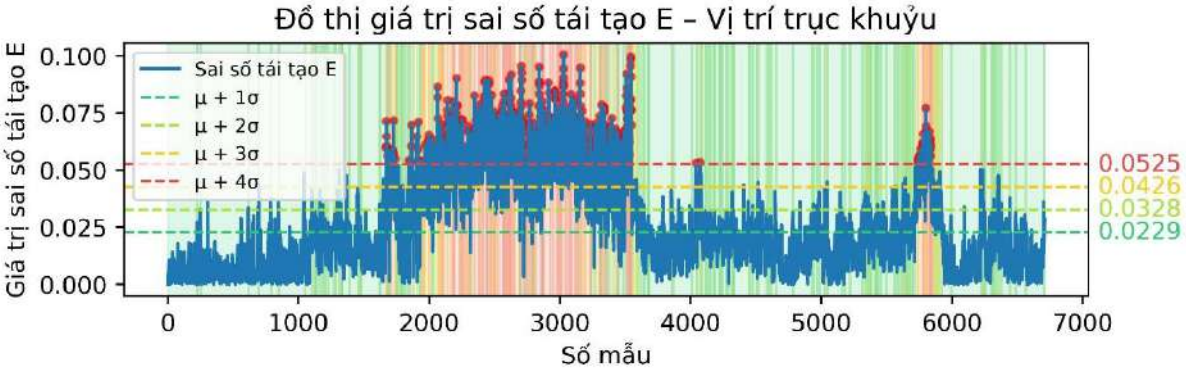
Hình 8. Đồ thị giá trị sai số tái tạo của cảm biến vị trí bướm ga



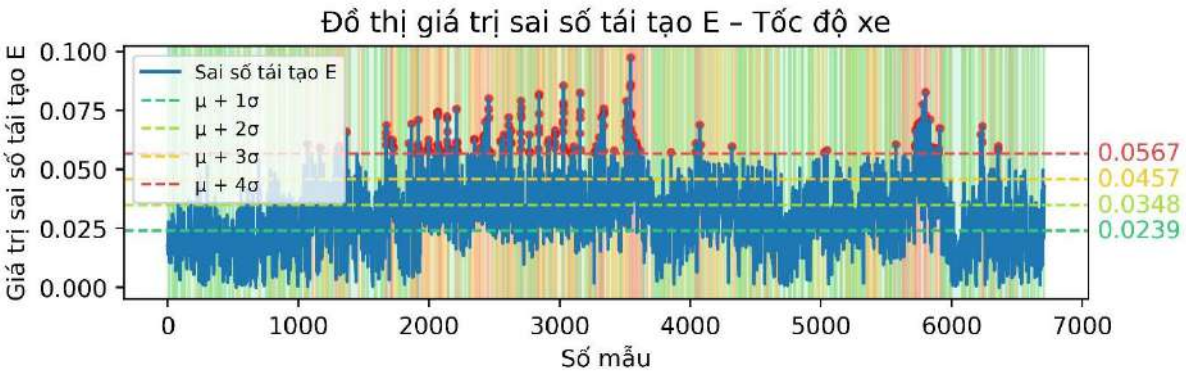
Hình 9. Đồ thị giá trị sai số tái tạo của cảm biến nhiệt độ động cơ



Hình 10. Đồ thị giá trị sai số tái tạo của cảm biến nhiệt độ khí nạp



Hình 11. Đồ thị giá trị sai số tái tạo của cảm biến vị trí trục khuỷu



Hình 12. Đồ thị giá trị sai số tái tạo của cảm biến tốc độ xe

Các đồ thị sai số tái tạo thể hiện rõ sự khác biệt giữa các cảm biến hoạt động bình thường và các cảm biến bị lỗi. Đối với các cảm biến được đánh giá là bình thường, giá trị sai số tái tạo chủ yếu nằm dưới các ngưỡng $\mu + 1\sigma$ và $\mu + 2\sigma$ tuy nhiên vẫn xuất hiện một số dao động cục bộ và đỉnh sai số cao hơn trong những khoảng thời gian nhất định. Những biến động này là hệ quả gián tiếp do sự hư hỏng của cảm biến áp suất khí nạp và Oxy làm thay đổi chế độ vận hành của động cơ và gây ảnh hưởng lan truyền đến tín hiệu của các cảm biến còn lại. Trong khi đó sai số tái tạo của cảm biến áp suất khí nạp và Oxy luôn duy trì ở mức rất cao và vượt xa tất cả các ngưỡng đánh giá, cho thấy sự bất thường kéo dài và nghiêm trọng. Điều này khẳng định mô hình không chỉ phát hiện chính xác lỗi cảm biến mà còn phản ánh được mối quan hệ phụ thuộc giữa các tín hiệu trong hệ thống động cơ.

Kết quả chẩn đoán cho thấy mô hình AE không chỉ nhận diện rõ các cảm biến bị lỗi mà còn phản ánh được sự thay đổi trạng thái vận hành chung của động cơ. Mặc dù phần lớn các cảm biến khác vẫn nằm trong vùng hoạt động bình thường, dữ liệu của chúng xuất hiện dao động nhất định do chịu tác động lan truyền từ các cảm biến hư hỏng. Điều này cho thấy mô hình có khả năng đánh giá tổng thể hệ thống chứ không chỉ dừng lại ở việc phát hiện lỗi đơn lẻ.

5. Kết luận

Kết quả nghiên cứu khẳng định mô hình chẩn đoán cảm biến ô tô dựa trên Autoencoder là một hướng tiếp cận đúng đắn, hiệu quả và có tính ứng dụng cao. Cấu trúc Autoencoder đối xứng với ba lớp ẩn được lựa chọn hợp lý giúp mô hình học biểu diễn phi tuyến giữa các cảm biến mà không cần dữ liệu lỗi.

Kết quả thử nghiệm trên xe Hyundai Elantra 2020 chứng minh khả năng tái tạo ổn định và tổng quát hóa tốt. Hệ thống chẩn đoán đã phát hiện chính xác lỗi của cảm biến áp suất khí nạp và cảm biến Oxy trong khi các cảm biến còn lại được đánh giá là hoạt động bình thường. Điều này chứng tỏ tính đúng đắn về mặt cấu trúc, cơ chế huấn luyện và ngưỡng phát hiện bất thường của mô hình đề xuất.

Tổng thể mô hình Autoencoder do nhóm tác giả đề xuất là một giải pháp khoa học hợp lý, đáp ứng tốt các yêu cầu phát hiện sớm, định lượng suy giảm và chẩn đoán lỗi cảm biến ô tô. Nghiên cứu mở ra hướng ứng dụng thực tế cho hệ thống giám sát kỹ thuật thông minh, có thể mở rộng sang nhiều loại cảm biến và phương tiện khác, tiến tới giám sát theo thời gian thực và bảo trì dự đoán trong ô tô hiện đại.

Tài liệu tham khảo

[1] Lê Khánh Tân (2019), *Ứng dụng nền tảng Arduino và LabVIEW trong thu thập dữ liệu động cơ*, Tạp chí Khoa học Giáo dục Kỹ thuật – Đại học Sư phạm Kỹ thuật TP Hồ Chí Minh, 51(01), tr. 51–56.

[2] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. Cambridge, MA: MIT Press.

[3] Torkhesari, S., Yousefimehr, B., & Ghatee, M. (2025). *Car Sensors Health Monitoring by Verification Based on Autoencoder and Random Forest Regression*. arXiv preprint arXiv:2505.00876.

[4] Dereszewski, M., & Sikora, G. (2019). *Diagnostics of the internal combustion engines operation by measurement of crankshaft instantaneous angular speed*. Journal of KONBiN, 49(4), 281–295.

[5] Mehranbod, N., Soroush, M., & Panjapornpon, C. (2005). *A method of sensor fault detection and identification. Journal of Process Control*, 15(3), 321–339.

[6] Suwatthikul, J. (2010). *Fault Detection and Diagnosis for In-Vehicle Networks*. In Fault Detection (pp. 283–306).

[7] Hinton, G. E., & Salakhutdinov, R. R. (2006). *Reducing the dimensionality of data with neural networks. Science*, 313(5786), 504–507.

[8] Amin, A. A., & Hasan, K. M.-U. (2019). *Advanced fault tolerant air-fuel ratio control of internal combustion gas engine for sensor and actuator faults*. IEEE Access, 7, 17634–17643.

[9] Shahbaz, M. H., & Amin, A. A. (2021). *Design of active fault tolerant control system for air-fuel ratio control of internal combustion engines using artificial neural networks.* IEEE Access, 9, 46022–46032.

[10] Amin, A. A., & Hasan, K. M.-U. (2019). *Robust active fault-tolerant control for internal combustion gas engine for air–fuel ratio control with statistical regression-based observer model. Measurement and Control*, 52(9–10), 1179–1194.

Application of artificial intelligence to assess the technical condition of automotive sensors through data extraction from the OBD-II port

Abstract: *The paper proposes an intelligent diagnostic system based on artificial intelligence to assess the technical condition of automotive sensors. The proposed approach employs an autoencoder neural network trained on datasets collected from healthy sensors operating under normal conditions, enabling anomaly detection through reconstruction error analysis. The autoencoder model is specifically developed for sensor data acquired via the OBD-II interface. Experimental results demonstrate that the model achieves stable reconstruction performance and accurately identifies sensor deviations. This approach aims to automate diagnostics, detect sensor faults, and support predictive maintenance, thereby enhancing vehicle reliability and operational efficiency.*

Keywords: *Sensor diagnostics, Autoencoder, OBD-II, Artificial intelligence.*

Khảo sát ảnh hưởng của quá trình phanh đến dao động của ô tô KAMAZ-43253

Huỳnh Đặng Nhứt Rồng[1], Vũ Mạnh Dũng[1]

[1]Học viện Kỹ thuật Quân sự

Email: nhutrong0603@gmail.com; Tel: 0384140843

Tóm tắt

Ở Việt Nam, mặc dù đã có những nghiên cứu về ảnh hưởng của quá trình phanh đến dao động ô tô, nhưng những nghiên cứu này chủ yếu áp dụng cho các phương tiện nhẹ hoặc các xe con. Xe tải hạng nặng, đặc biệt là các phương tiện như KAMAZ-43253, vẫn còn thiếu các nghiên cứu chuyên sâu. Những xe tải này có tải trọng và khối lượng lớn, khiến quá trình phanh và ảnh hưởng của dao động trở nên phức tạp và mạnh mẽ hơn. Trên cơ sở đó, bài báo tiến hành xây dựng mô hình và hệ phương trình vi phân, sử dụng phần mềm MATLAB để mô phỏng dao động của ô tô nhằm làm cơ sở dữ liệu tham khảo cho các nghiên cứu sau này.

Từ khóa: *Ảnh hưởng quá trình phanh, dao động lắc dọc, hệ thống treo*

1. Đặt vấn đề

Trong quá trình vận hành ô tô, phanh đóng vai trò quan trọng không chỉ trong việc giảm tốc độ mà còn ảnh hưởng sâu sắc đến các yếu tố động lực học và sự ổn định của phương tiện, đặc biệt là đối với các loại xe tải hạng nặng như KAMAZ-43253. Khi thực hiện thao tác phanh, một loạt các hiện tượng động học xảy ra, đặc biệt là dao động và rung lắc, có thể ảnh hưởng nghiêm trọng đến sự ổn định của xe, khả năng kiểm soát phương tiện và trải nghiệm của người lái. Quá trình phanh, đặc biệt là phanh gấp hoặc phanh mạnh, tạo ra một loạt lực quán tính và thay đổi tải trọng trong các bộ phận của xe. Những thay đổi này gây ra sự phân bổ lại trọng lượng, dẫn đến việc các bộ phận như bánh xe, hệ thống treo và thân xe phải chịu tác động không đồng đều, từ đó gây ra dao động không mong muốn.

Vì vậy, việc khảo sát ảnh hưởng của quá trình phanh đến dao động của xe là vô cùng quan trọng, đặc biệt đối với các phương tiện hạng nặng như KAMAZ-43253, khối lượng lớn và tải trọng cao của xe gây ra các tác động mạnh mẽ và phức tạp. Nghiên cứu này sẽ giúp chúng ta hiểu rõ hơn về các yếu tố ảnh hưởng đến sự dao động của xe khi phanh, từ đó xây dựng mô hình và hệ phương trình vi phân, sử dụng MATLAB mô phỏng dao động. Kết quả mô phỏng làm cơ sở dữ liệu đánh giá dao động của ô tô khi phanh và làm cơ sở dữ liệu cho các nghiên cứu sau này.

2. Thiết lập mô hình nghiên cứu ảnh hưởng của quá trình phanh đến dao động của ô tô KAMAZ-43253

2.1. Giả thuyết

Tác giả đưa ra các giả thuyết nhằm đơn giản hóa bài toán nhưng vẫn đánh giá được các thông số dao động cần thiết như sau:

- Mô hình động lực học biểu thị dao động liên kết ô tô hai cầu ở dạng mô hình phẳng, có nghĩa là ô tô được giả thiết là đối xứng đối với trục dọc xe và xem độ mấp mô của biên dạng đường ở dưới các bánh xe bên phải và bên trái của một cầu là như nhau.

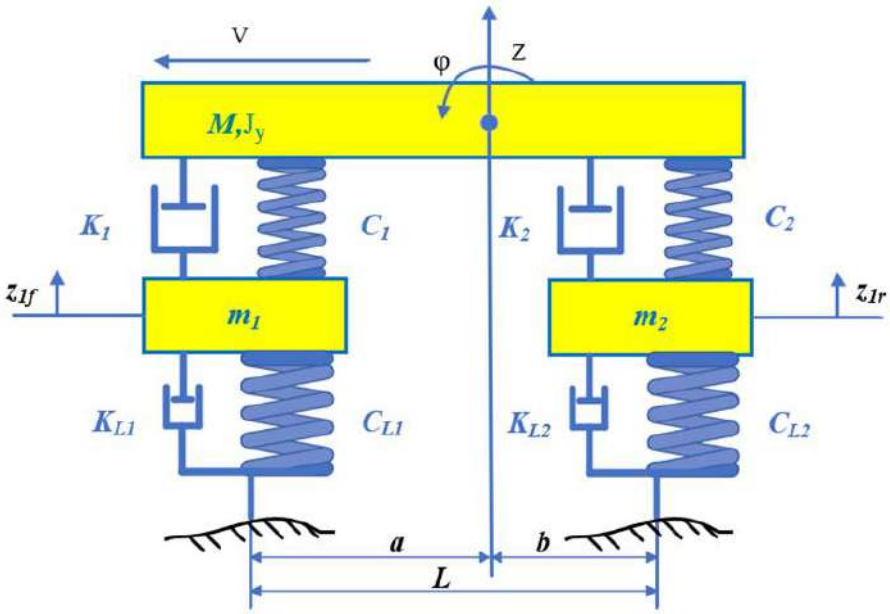
- Khối lượng treo được quy dẫn về trọng tâm phần treo biểu thị qua giá trị khối lượng M và mô men quán tính khối lượng phần treo đối với trục ngang đi qua trọng tâm phần treo J_y. Khối lượng M được liên kết với khối lượng không treo qua phần tử đàn hồi của treo và giảm chấn (hệ số đàn hồi của treo C_1; C_2; hệ số cản giảm chấn K_1, K_2).

- Khối lượng phần không treo tương ứng ở cầu trước và cầu sau là m_1, m_2, được liên hệ với đường qua các lò xo (có hệ số đàn hồi C_{L1}, C_{L2}) và giảm chấn (có hệ số cản giảm chấn K_{L1}, K_{L2}). Các thông số này đặc trưng cho độ đàn hồi của lốp và sự dập tắt dao động của lốp.

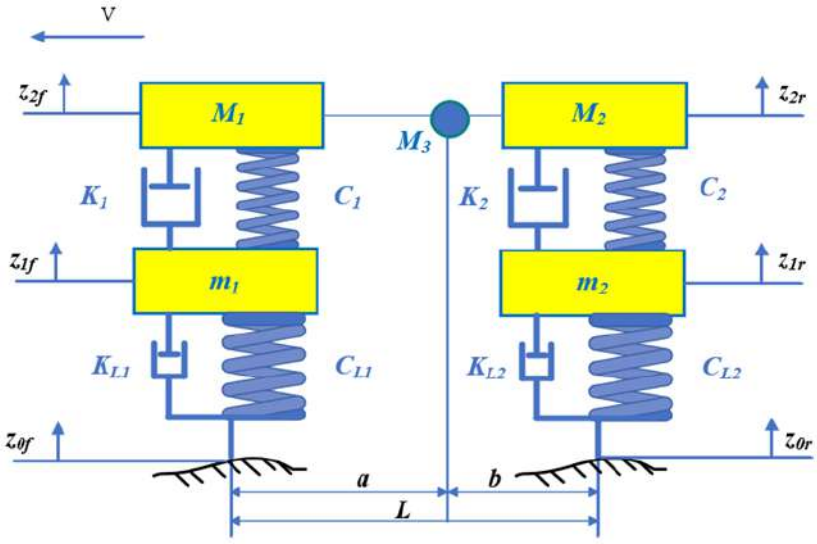
- Ô tô chuyển động đều, tiếp xúc bánh xe với mặt đường là tiếp xúc điểm. Dao động của hệ là dao động nhỏ, tuyến tính, xung quanh vị trí cân bằng tĩnh. Bỏ qua ảnh hưởng của gió, lực cản mooc kéo.

2.2. Mô hình khảo sát

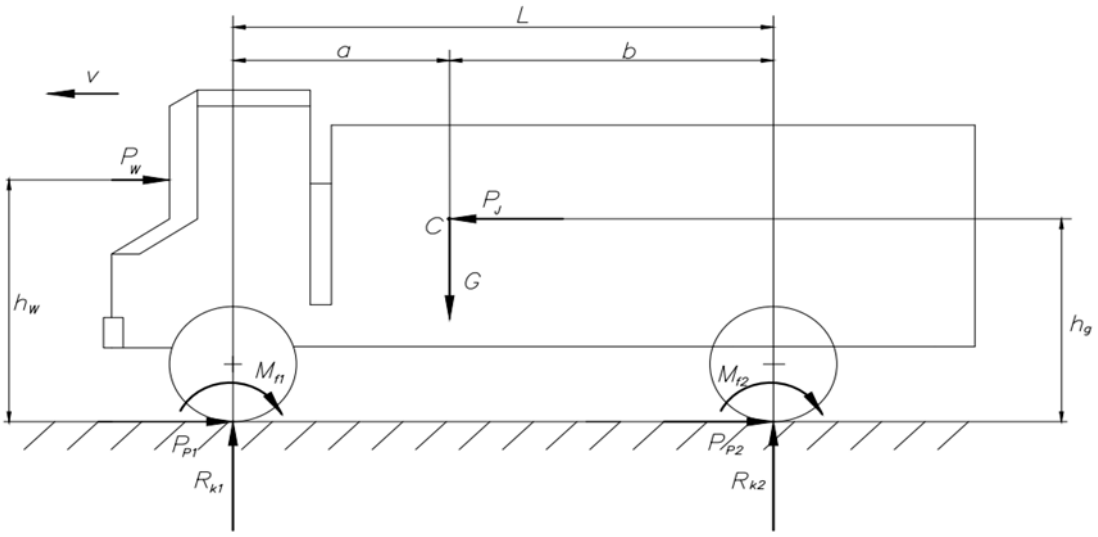
Từ các giả thuyết ở trên, tác giả đưa ra mô hình tương đương khảo sát dao động ô tô 2 cầu như sau:



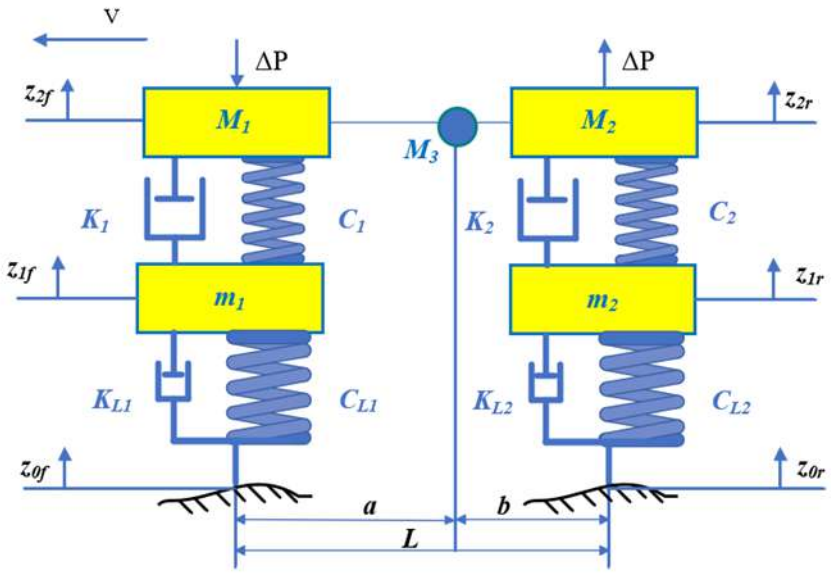
Hình 1. Mô hình phẳng hệ dao động ½ dọc xe



Hình 2. Mô hình dao động ½ dọc xe quy dẫn khối lượng



Hình 3. Sơ đồ lực tác dụng lên ô tô 2 cầu khi phanh



Hình 4. Mô hình dao động ô tô 2 cầu trong quá trình phanh

Trong đó:

M, m_1, m_2 - Khối lượng phần treo, phần không treo phía trước, phía sau [kg];

M_1, M_2, M_3 - Khối lượng phần treo quy dẫn về treo trước, sau và trọng tâm [kg];

J_y - Mômen quán tính khối lượng phần treo đi qua trọng tâm phần treo đối với trục Y [kg.m^2];

Z - Dịch chuyển thẳng đứng của khối lượng treo [m];

z_{0f}, z_{0r} - Mấp mô mặt đường tương ứng lốp trước, sau [m];

z_{1f}, z_{1r} - Dịch chuyển thẳng đứng của khối lượng treo trước, sau [m];

z_{2f}, z_{2r} - Dịch chuyển thẳng đứng của khối lượng không treo trước, sau [m];

φ - Góc lắc dọc của ô tô [rad];

K_{L1}, K_{L2} - Hệ số cản giảm chấn của lốp trước và các lốp sau [Ns/m];

C_{L1}, C_{L2} - Hệ số đàn hồi của lốp trước và các lốp sau [N/m];

K_1, K_2 - Hệ số cản giảm chấn của treo trước và các treo sau [Ns/m];

C_1, C_2 - Hệ số đàn hồi treo trước và treo sau [N/m];

L - Chiều dài cơ sở của ô tô [m];

a, b - Khoảng cách từ trọng tâm ô tô đến cầu trước, sau [m];

P_{p1}, P_{p2} - Lực phanh tương ứng ở các bánh xe cầu trước, cầu sau [N];

P_j - Lực quán tính khi phanh [N];

P_w - Lực cản không khí [N];

G - Trọng lượng ô tô [N];

R_{k1}, R_{k2} - Phản lực pháp tuyến của đường và lực cản lăn P_{f1}, P_{f2} tương ứng đặt tại vị trí tiếp xúc giữa các bánh xe cầu trước, cầu sau lên đường [N];

M_{f1}, M_{f2} - Mômen cản lăn tương ứng ở các bánh xe cầu trước, cầu sau [Nm];

h_g - Chiều cao trọng tâm xe [m];

h_w - Chiều cao điểm đặt lực cản không khí [m].

2.3. Hệ phương trình vi phân của dao động

Dựa trên mô hình khảo sát đã phân tích, tác giả nhận thấy hệ dao động có 4 bậc tự do. Dùng phương trình Lagrange loại 2 để thiết lập hệ phương trình vi phân.

Sau quá trình tính toán, phân tích, tác giả đã xây dựng hệ phương trình vi phân bao gồm 4 phương trình sau:

$$\begin{cases} M.\ddot{Z}+(K_1+K_2).\dot{Z}+(C_1+C_2).Z+(a.K_1-b.K_2).\dot{\varphi}+(a.C_1-b.C_2).\varphi- \\ \qquad -K_1.\dot{z}_{1f}-C_1.z_{1f}-K_2.\dot{z}_{1r}-C_2.z_{1r}=0 \\ J_y.\ddot{\varphi}+(a.K_1-b.K_2).\dot{Z}+(a.C_1-b.C_2).Z+(a^2.K_1+b^2.K_2).\dot{\varphi}+ \\ \qquad +(a^2.C_1+b^2.C_2).\varphi-a.K_1.\dot{z}_{1f}-a.C_1.z_{1f}+b.K_2.\dot{z}_{1r}+b.C_2.z_{1r}=0 \\ m_1.\ddot{z}_{1f}-K_1.\dot{Z}-C_1.Z-a.K_1.\dot{\varphi}-a.C_1.\varphi+(K_1+K_{L1}).\dot{z}_{1f}+(C_1+C_{L1}).z_{1f} \\ \qquad =K_{L1}.\dot{z}_{0f}+C_{L1}.z_{0f} \\ m_2.\ddot{z}_{1r}-K_2.\dot{Z}-C_2.Z+b.K_2.\dot{\varphi}+b.C_2.\varphi+(K_2+K_{L2}).\dot{z}_{1r}+(C_2+C_{L2}).z_{1r} \\ \qquad =K_{L2}.\dot{z}_{0r}+C_{L2}.z_{0r} \end{cases} \quad (1)$$

Chúng ta có thể mô tả chuyển động của hệ theo các tọa độ suy rộng:

$$z_{2f}, z_{2r}, z_{1f}, z_{1r}$$

Các tọa độ này có liên hệ với các tọa độ Z, φ qua các biểu thức sau:

$$\begin{aligned} &Z=\frac{z_{2r}.a+z_{2f}.b}{L}; \varphi=\frac{z_{2r}-z_{2f}}{L}; \\ &z_{2f}=Z-\varphi.a; z_{2r}=Z-\varphi.b; \end{aligned} \quad (2)$$

Ta nhận được phương trình vi phân mô tả chuyển động của cơ hệ đối với các tọa độ suy rộng: z_{1f}, z_{1r}, z_{2f}, z_{2r}

$$\begin{cases} M_1.\ddot{z}_{2f} + K_1.(\dot{z}_{2f} - \dot{z}_{1f}) + C_1(z_{2f} - z_{1f}) + M_3.\ddot{z}_{2r} = 0 \\ M_2.\ddot{z}_{2r} + K_2.(\dot{z}_{2r} - \dot{z}_{1r}) + C_2.(z_{2r} - z_{1r}) + M_3.\ddot{z}_{2f} = 0 \\ m_1.\ddot{z}_{1f} + K_1.(\dot{z}_{1f} - \dot{z}_{2f}) + K_{L1}.(\dot{z}_{1f} - \dot{z}_{0f}) + C_1.(z_{1f} - z_{2f}) + C_{L1}.(z_{1f} - z_{0f}) = 0 \\ m_2.\ddot{z}_{1r} + K_2.(\dot{z}_{1r} - \dot{z}_{2r}) + K_{L2}.(\dot{z}_{1r} - \dot{z}_{0r}) + C_2.(z_{1r} - z_{2r}) + C_{L2}.(z_{1r} - z_{0r}) = 0 \end{cases} \quad (3)$$

Các ký hiệu M_1, M_2, M_3 là các khối lượng quy dẫn được xác định từ 3 điều kiện sau đây:

* Điều kiện bảo toàn khối lượng treo M:

$$M_1 + M_2 + M_3 = M; M_1 > 0; M_2 > 0; M_3 > 0, < 0, = 0.$$

* Điều kiện bảo toàn vị trí trọng tâm khối lượng treo:

$$M_1.a - M_2.b = 0;$$

* Điều kiện bảo toàn mô men quán tính khối lượng treo J_y:

$$J_y = M.\rho_y^2 = M_1.a^2 + M_2.b^2;$$

ρ_y - bán kính quán tính quy dẫn của khối lượng treo M đối với trục Y đi qua trọng tâm khối lượng phần treo.

Từ 3 điều kiện trên chúng ta xác định trị số của các khối lượng quy dẫn M_1, M_2, M_3 như sau:

$$M_1 = M.\frac{\rho_y^2}{a.L}; M_2 = M.\frac{\rho_y^2}{b.L}; M_3 = M.\left(1 - \frac{\rho_y^2}{a.b}\right); \quad (4)$$

Để xác định tải trọng thẳng đứng khi phanh (hay là phản lực pháp tuyến của đường) chúng ta lập phương trình cân bằng mô men với các điểm tiếp xúc của bánh xe với đường ta có (bỏ qua lực cản lăn, lực cản không khí, mô men quán tính bánh xe):

$$R_1.L - M.J_p.h_g - G.b = 0; R_2.L + M.J_p.h_g - G.a = 0; \quad (5)$$

Từ 2 biểu thức (5) ta xác định được giá trị phản lực pháp tuyến tại các trục như sau:

$$R_1 = \frac{G.b}{L} + \frac{M.J_p.h_g}{L}; \quad R_2 = \frac{G.a}{L} - \frac{M.J_p.h_g}{L}; \quad (6)$$

Như vậy, trong quá trình chuyển động khi có lực phanh tác động sinh ra lực quán tính đặt tại trọng tâm xe dẫn đến sự phân bố lại trọng lượng phần treo của xe, lúc này phần treo cầu trước và phần treo cầu sau thêm bớt 1 lượng theo công thức sau:

$$\Delta P = \frac{M.J_p.h_g}{L} \quad (7)$$

Kết hợp (7) và (3) ta nhận được hệ phương trình vi phân mô tả dao động của hệ trong quá trình phanh như sau:

$$\begin{cases} M_1.\ddot{z}_{2f} + K_1.(\dot{z}_{2f} - \dot{z}_{1f}) + C_1(z_{2f} - z_{1f}) + M_3.\ddot{z}_{2r} = \dfrac{M.h_g}{L}.J_p \\ M_2.\ddot{z}_{2r} + K_2.(\dot{z}_{2r} - \dot{z}_{1r}) + C_2.(z_{2r} - z_{1r}) + M_3.\ddot{z}_{2f} = -\dfrac{M.h_g}{L}.J_p \\ m_1.\ddot{z}_{1f} + K_1.(\dot{z}_{1f} - \dot{z}_{2f}) + K_{L1}.(\dot{z}_{1f} - \dot{z}_{0f}) + C_1.(z_{1f} - z_{2f}) + C_{L1}.(z_{1f} - z_{0f}) = 0 \\ m_2.\ddot{z}_{1r} + K_2.(\dot{z}_{1r} - \dot{z}_{2r}) + K_{L2}.(\dot{z}_{1r} - \dot{z}_{0r}) + C_2.(z_{1r} - z_{2r}) + C_{L2}.(z_{1r} - z_{0r}) = 0 \end{cases} \quad (8)$$

Giải hệ phương trình trên bằng phương pháp không gian trạng thái:

$$\begin{cases} \dot{x} = A.x + B.u \\ y = C.x + D.u \end{cases}$$

Trong đó: x - biến trạng thái;

u - biến điều khiển hay kích thích đầu vào;

y - thông số ra;

A - ma trận trạng thái;

B - ma trận quan hệ thống số vào với trạng thái;

C - ma trận quan hệ thông số ra với trạng thái;

D - ma trận truyền tín hiệu điều khiển;

Từ phương trình vi phân dao động của cơ hệ, xác định được:

- Biến trạng thái:

$$x = \left[z_{2f}, \dot{z}_{2f}, z_{2r}, \dot{z}_{2r}, z_{1f}, \dot{z}_{1f}, z_{1r}, \dot{z}_{1r} \right]^T$$

- Kích thích đầu vào:

$$u = \left[J_p, z_{0f}, \dot{z}_{0f}, z_{0r}, \dot{z}_{0r} \right]^T$$

- Thông số ra:

$$y = \left[z_{2f}, z_{2r}, Z, \dot{Z}, \ddot{Z}, \varphi \right]^T$$

- Ma trận biến trạng thái:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{-C_1.M_2}{M_1.M_2 - M_3^2} & \frac{-K_1.M_2}{M_1.M_2 - M_3^2} & \frac{C_2.M_3}{M_1.M_2 - M_3^2} & \frac{K_2.M_3}{M_1.M_2 - M_3^2} & \frac{C_1.M_2}{M_1.M_2 - M_3^2} & \frac{K_1.M_2}{M_1.M_2 - M_3^2} & \frac{-C_2.M_3}{M_1.M_2 - M_3^2} & \frac{-K_2.M_3}{M_1.M_2 - M_3^2} \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{C_1.M_3}{M_1.M_2 - M_3^2} & \frac{K_1.M_3}{M_1.M_2 - M_3^2} & \frac{-C_2.M_1}{M_1.M_2 - M_3^2} & \frac{-K_2.M_1}{M_1.M_2 - M_3^2} & \frac{-C_1.M_3}{M_1.M_2 - M_3^2} & \frac{-K_1.M_3}{M_1.M_2 - M_3^2} & \frac{C_2.M_1}{M_1.M_2 - M_3^2} & \frac{K_2.M_1}{M_1.M_2 - M_3^2} \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \frac{C_1}{m_1} & \frac{K_1}{m_1} & 0 & 0 & -\frac{C_1 + C_{L1}}{m_1} & -\frac{K_1 + K_{L1}}{m_1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{C_2}{m_2} & \frac{K_2}{m_2} & 0 & 0 & -\frac{C_2 + C_{L2}}{m_2} & -\frac{K_2 + K_{L2}}{m_2} \end{bmatrix}$$

- Ma trận quan hệ thống số ra với biến trạng thái:

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{b}{L} & 0 & \frac{a}{L} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{b}{L} & 0 & \frac{a}{L} & 0 & 0 & 0 & 0 \\ \frac{C_1.f}{M_1.M_2 - M_3^2} & \frac{K_1.f}{M_1.M_2 - M_3^2} & \frac{C_2.g}{M_1.M_2 - M_3^2} & \frac{K_2.g}{M_1.M_2 - M_3^2} & \frac{-C_1.f}{M_1.M_2 - M_3^2} & \frac{-K_1.f}{M_1.M_2 - M_3^2} & \frac{-C_2.g}{M_1.M_2 - M_3^2} & \frac{-K_2.g}{M_1.M_2 - M_3^2} \\ \frac{-1}{L} & 0 & \frac{1}{L} & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- Ma trận quan hệ thống số vào với biến trạng thái:

$$B=\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \frac{(M_2+M_3).M.h_g}{(M_1.M_2-M_3^2).L} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \frac{-(M_1+M_3).M.h_g}{(M_1.M_2-M_3^2).L} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{C_{L1}}{m_1} & \frac{K_{L1}}{m_1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{C_{L2}}{m_2} & \frac{K_{L2}}{m_2} \end{bmatrix}$$

- Ma trận truyền tín hiệu điều khiển:

$$D=\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \frac{(g-f).M.h_g}{(M_1.M_2-M_3^2).L^2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Với: $f=\frac{a.M_3-b.M_2}{L};\ g=\frac{-a.M_1+b.M_3}{L}$

3. Kết quả nghiên cứu

3.1. Thông số đầu vào

Tác giả sử dụng phần mềm MATLAB [2] để giải hệ phương trình (8) và sử dụng các thông số của xe KAMAZ-43253 [3-4] như sau:

Bảng 1. Các thông số đầu vào của ô tô KAMAZ-43253

TT	Thông số	Kí hiệu	Giá trị	Đơn vị
1	Trọng lượng toàn bộ *(tổng tải trọng)*	M	14.440	kg
2	Khối lượng được treo	m	12190	kg
	Khối lượng không được treo cầu trước	m_1	800	kg
	Khối lượng không được treo cầu sau	m_2	1450	kg
3	Chiều dài cơ sở của xe	L	4,2	m
	Khoảng cách từ trọng tâm đến cầu trước	a	2,66	m
	Khoảng cách từ trọng tâm đến cầu sau	b	1,54	m
4	Độ cứng của nhíp trước	C_1	420000	N/m

TT	Thông số	Kí hiệu	Giá trị	Đơn vị
5	Độ cứng của nhíp sau	C_2	900000	N/m
6	Hệ số cản giảm chấn của hệ thống treo trước	K_1	20000	Ns/m
7	Hệ số cản giảm chấn của hệ thống treo sau	K_2	45000	Ns/m
8	Độ cứng hướng kính lốp trước	C_{L1}	2200000	N/m
9	Độ cứng hướng kính lốp sau	C_{L2}	4400000	N/m
10	Hệ số cản giảm chấn của lốp trước	K_{L1}	3000	Ns/m
11	Hệ số cản giảm chấn của lốp sau	K_{L2}	6000	Ns/m
12	Chiều cao trọng tâm xe đầy tải	h_g	1,76	m
13	Vận tốc ban đầu của xe	V_0	30	km/h
14	Bước sóng mặt đường	S	10	m

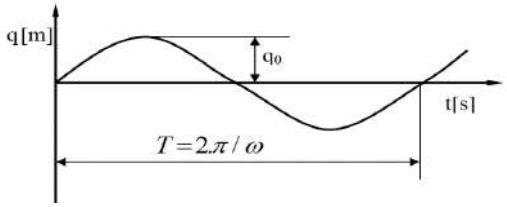
3.2. Kết quả mô phỏng

Tác giả tiến hành khảo sát dao động ô tô khi phanh trong các trường hợp sau:

Bảng 2. Các trường hợp khảo sát

TH	Điều kiện vận hành	Biên dạng mặt đường	Thời gian khảo sát
1	Ô tô chuyển động trên đường bằng phẳng	$q_0 = 0$	$t_0 = 0$ (vị trí ban đầu) t = 2s (bắt đầu phanh)
2	Ô tô chuyển động trên đường có mấp mô mặt đường dạng hàm điều hòa	$q_0 = 0,02$ m	$t_0 = 0$ (vị trí ban đầu) t = 2s (bắt đầu phanh)
3	Ô tô di chuyển trên đường bằng phẳng qua gờ giảm tốc với v = 15 km/h	Gờ có chiều cao 0,05 m, rộng 0,2 m	$t_0 = 0$ (vị trí ban đầu) t = 5s (thời điểm qua gờ)

* Biên dạng của mấp mô được biểu thị bằng hàm điều hòa:



Hình 5. Mô tả mấp mô mặt đường dạng hàm điều hòa

Phương trình biểu diễn chiều cao mấp mô phụ thuộc vào thời gian có dạng như sau:

$$q_1 = q_0.\sin(\omega t);\ q_2 = q_0.\sin[\omega(t - L/V)]$$

Trong đó:

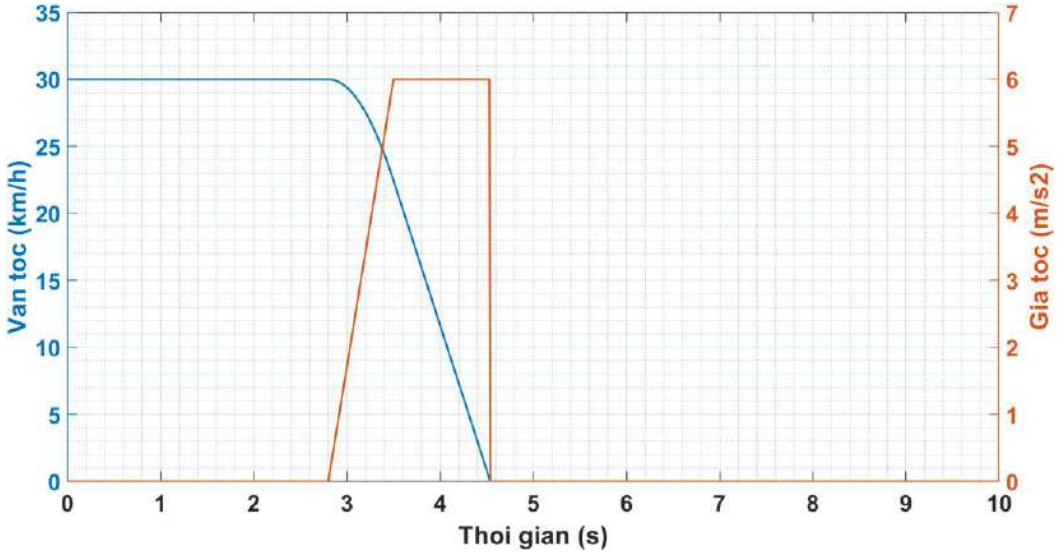
V - vận tốc chuyển động của ô tô (m/s);

L - chiều dài cơ sở của ô tô (m);

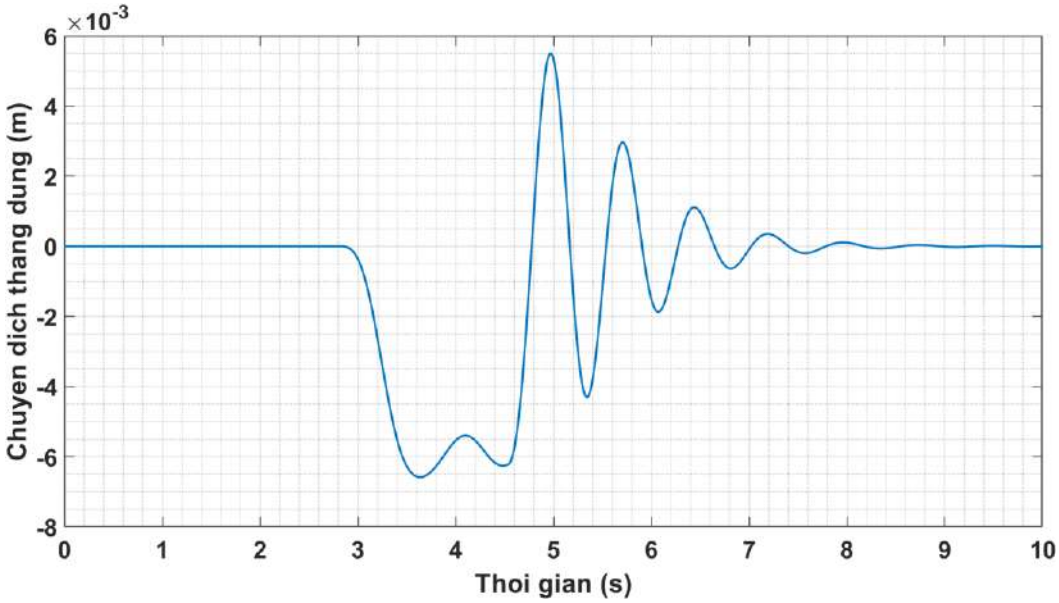
S - chiều dài sóng mặt đường (m);

ω - tần số kích thích $\omega = \frac{2.\pi.V}{S}$ (rad/s).

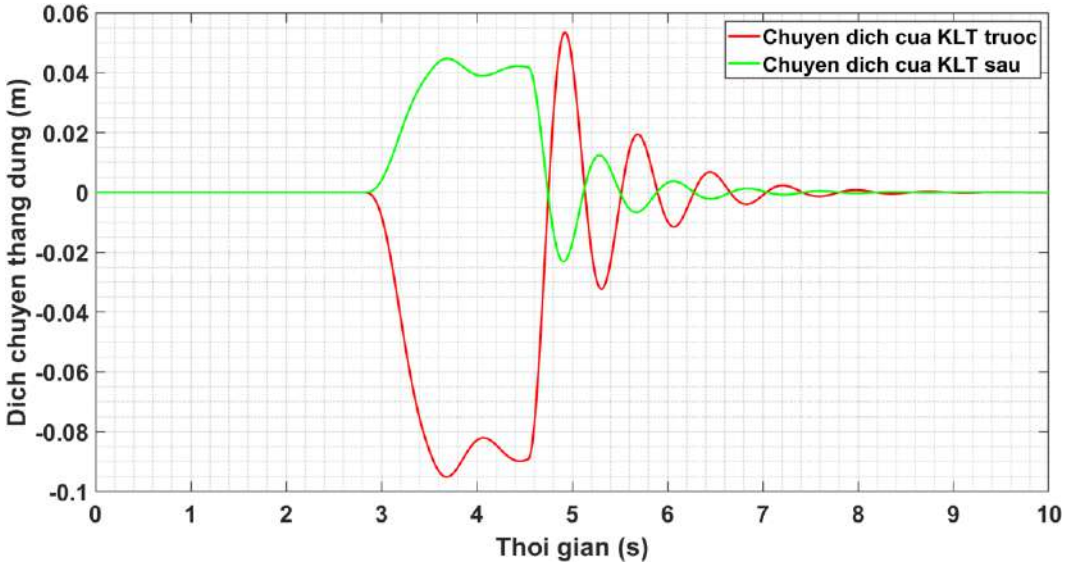
- Trường hợp 1:



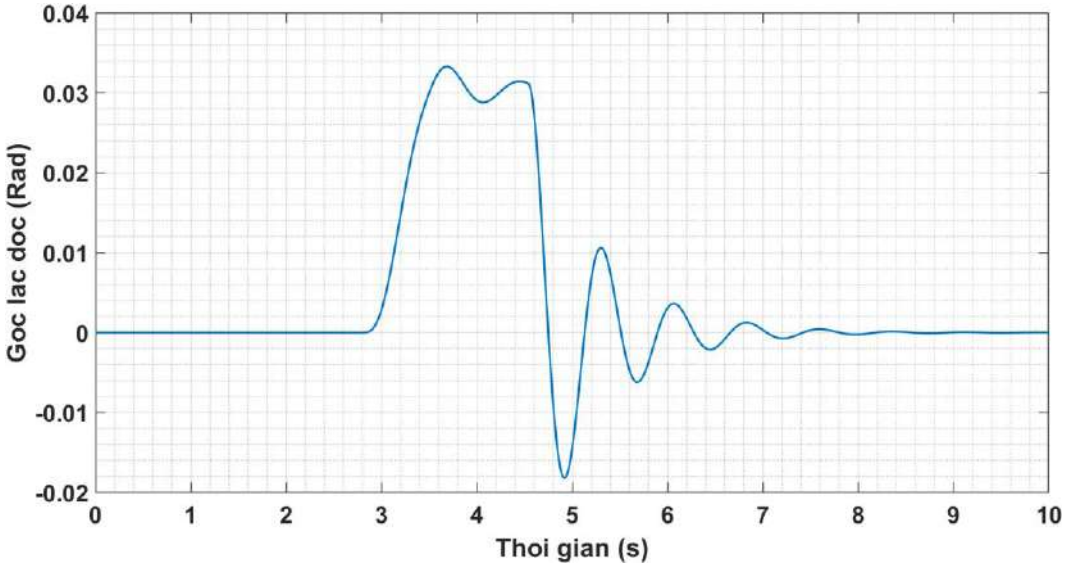
Hình 6. Đồ thị vận tốc, gia tốc ô tô KAMAZ-43253



Hình 7. Đồ thị dịch chuyển thẳng đứng

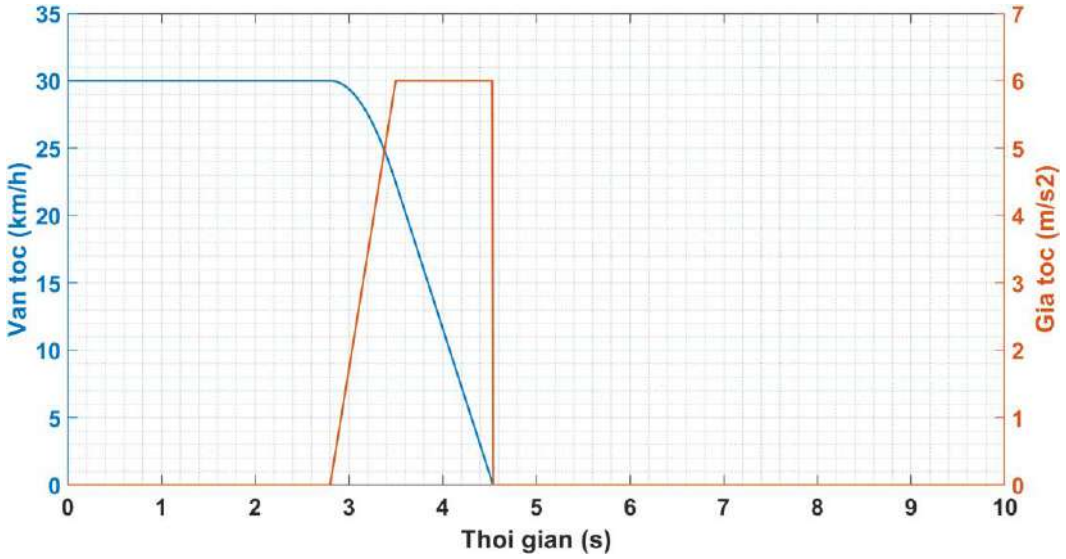


Hình 8. Đồ thị dịch chuyển thẳng đứng KLT trước, sau

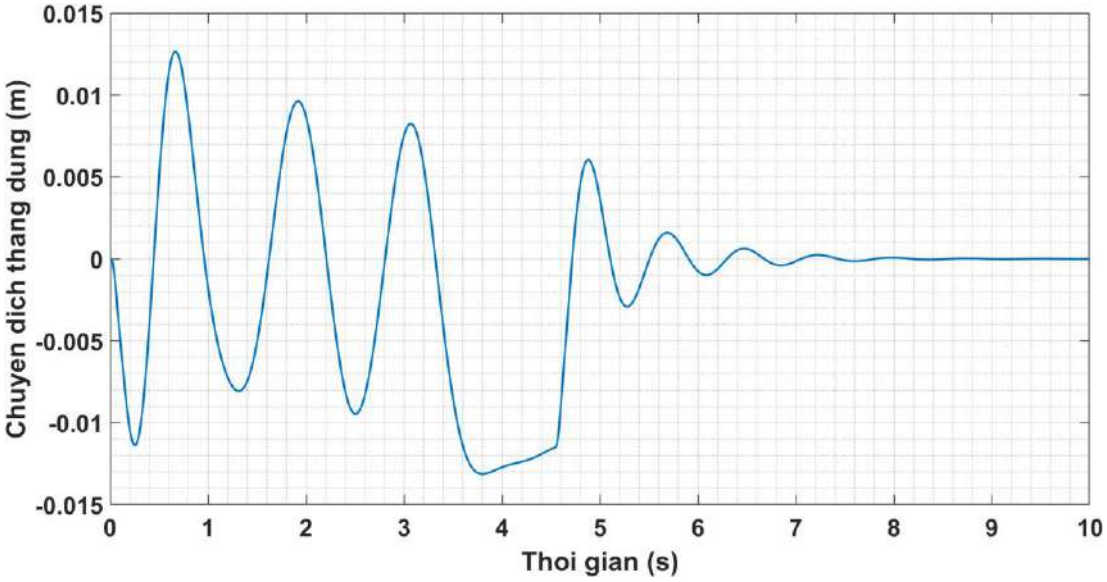


Hình 9. Đồ thị góc lắc dọc

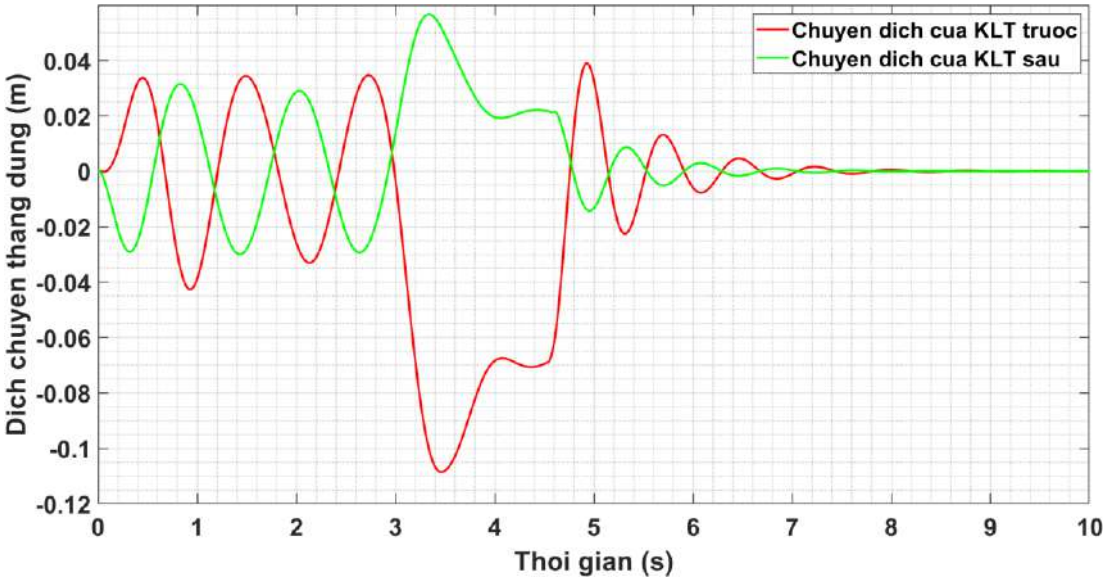
- Trường hợp 2:



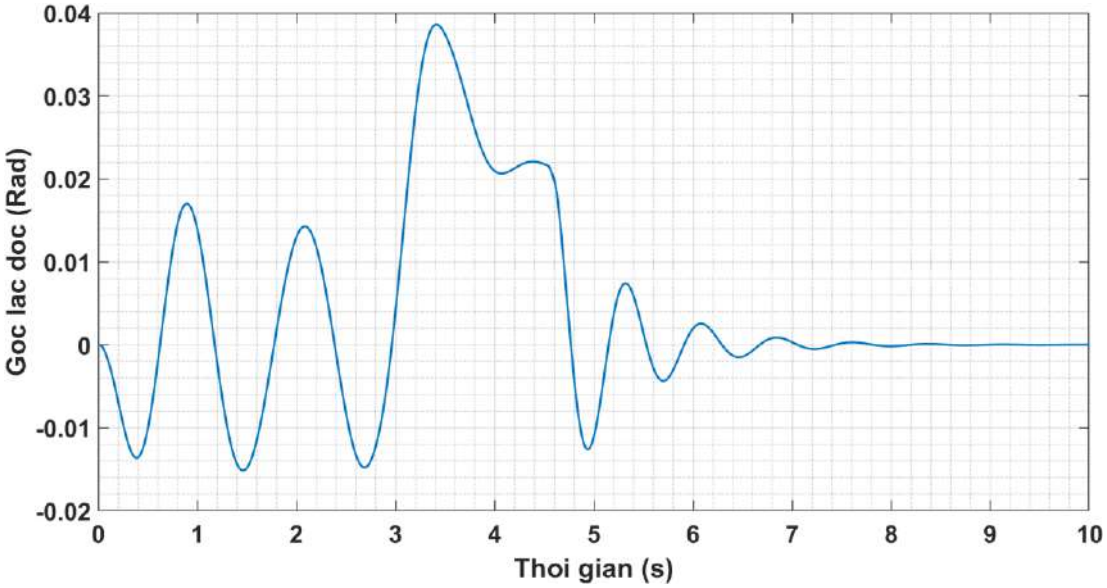
Hình 10. Đồ thị dịch chuyển thẳng đứng KLT trước, sau



Hình 11. Đồ thị dịch chuyển thẳng đứng

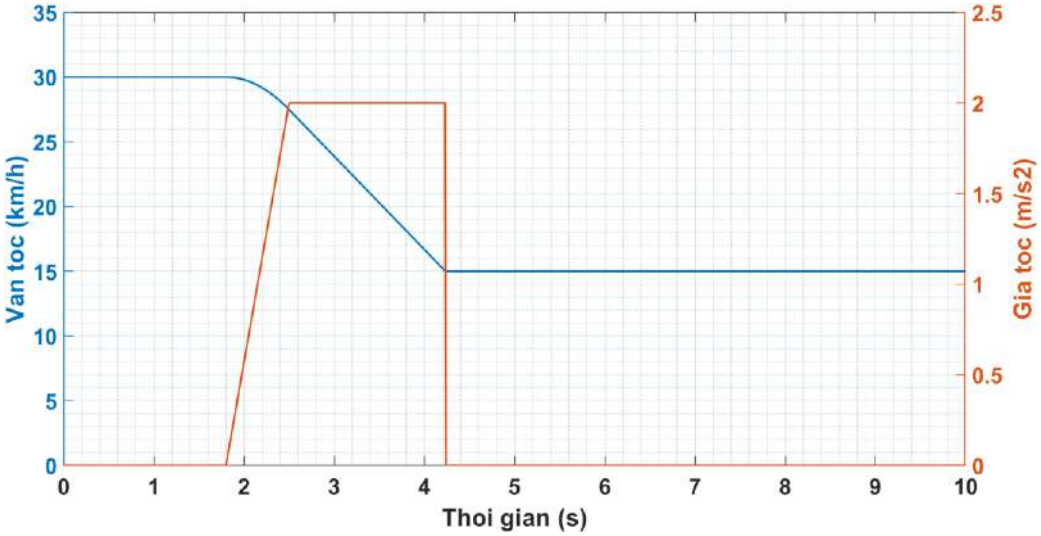


Hình 12. Đồ thị dịch chuyển thẳng đứng KLT trước, sau

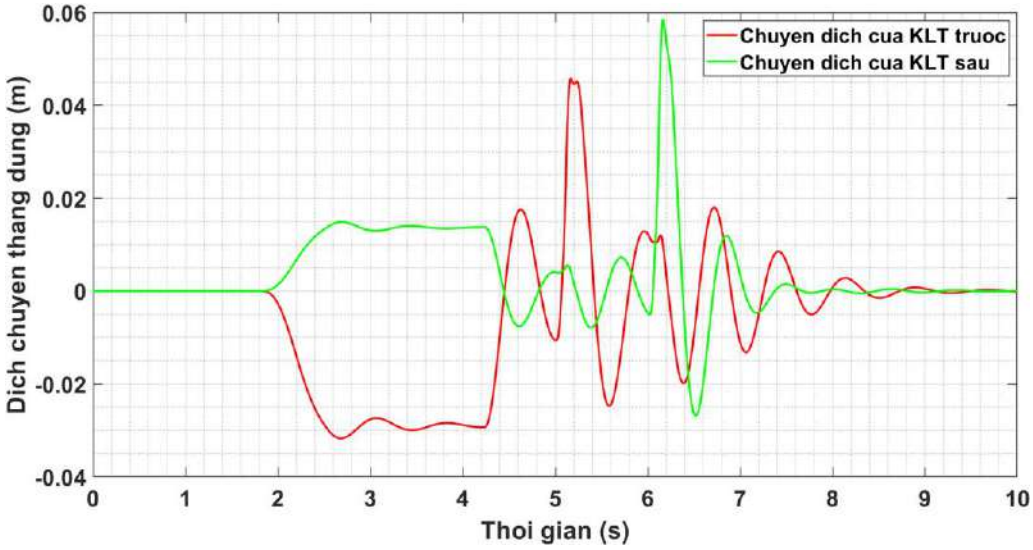


Hình 13. Đồ thị góc lắc dọc

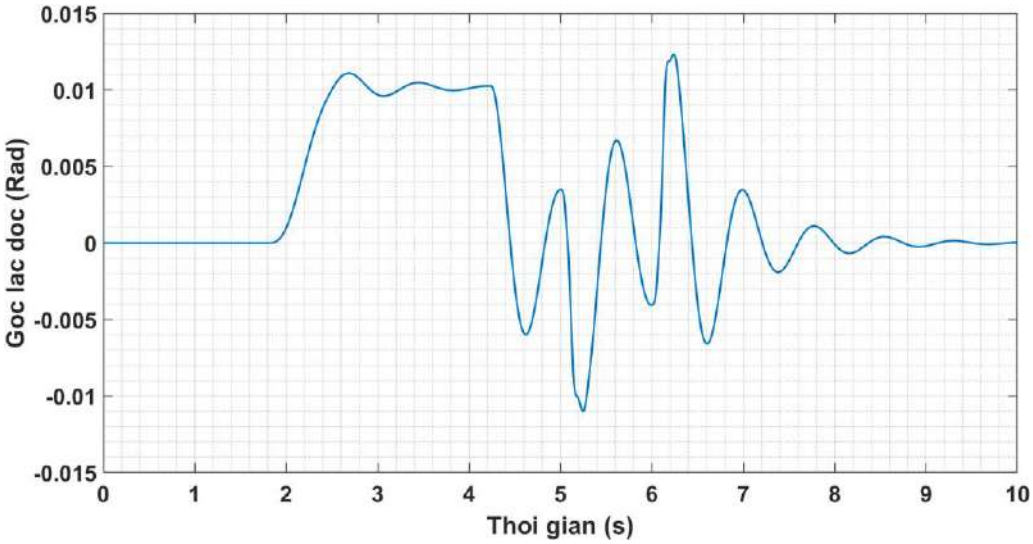
- Trường hợp 3:



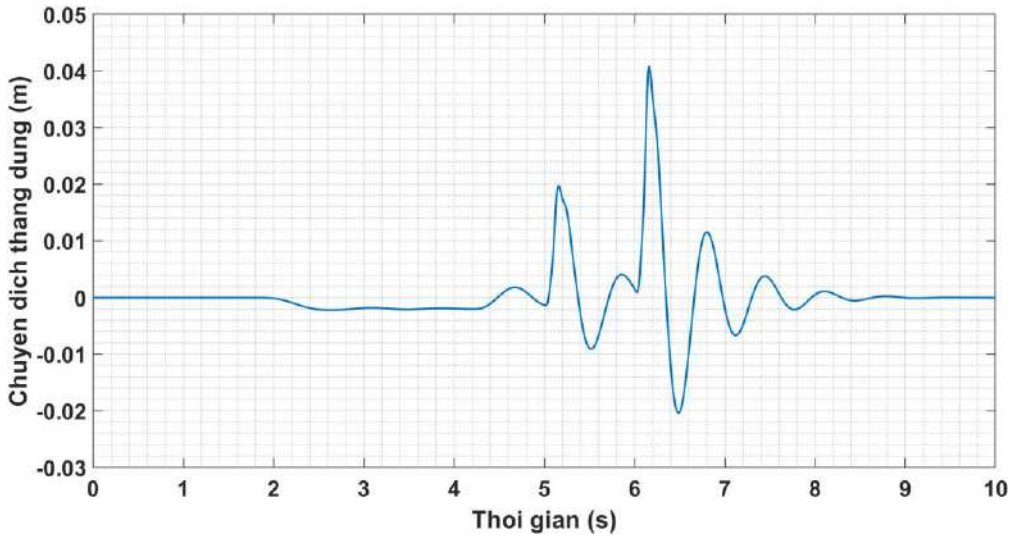
Hình 14. Đồ thị vận tốc, gia tốc ô tô KAMAZ-43253



Hình 15. Đồ thị dịch chuyển thẳng đứng KLT trước, sau



Hình 16. Đồ thị góc lắc dọc



Hình 17. Đồ thị dịch chuyển thẳng đứng

- Nhận xét:

+ Khi thực hiện quá trình phanh, dưới tác động của lực quán tính đặt tại trọng tâm xe làm phân bố lại KLT trước và sau (tăng giảm 1 lượng ΔP).

+ Căn cứ vào kết quả trên đồ thị, ta thấy ở trường hợp 1 và 2 (J_p = 6 m/s^2 ứng với trường hợp phanh cấp tốc), trong thời gian quá trình phanh diễn ra (J_p tác động vào cơ hệ), KLT trước, sau và thân xe dao động quanh vị trí cân bằng mới do sự phân bố lại khối lượng. Sau khi quá trình phanh kết thúc (J_p = 0), dao động của cơ hệ sẽ tắt dần trong vòng 2,5 giây và dao động quanh vị trí ban đầu.

+ Ở trường hợp 3 (J_p = 2 m/s^2 qua gờ giảm tốc), khi vận tốc giảm từ 30 km/h xuống 15 km/h cơ hệ dao động ở vị trí cân bằng mới với biên độ dao động nhỏ hơn nhiều so với 2 trường hợp trên. Khi qua gờ, biên độ dao động tăng cao sau đó tắt dần và quay về vị trí cân bằng ban đầu.

4. Kết luận

Kết luận rút ra từ nghiên cứu như sau:

- Xây dựng được mô hình động lực học dao động khi phanh và thiết lập hệ phương trình vi phân mô tả dao động của ô tô trong quá trình phanh và đã khảo sát đối với 3 trường hợp di chuyển khác nhau. Giải hệ phương trình vi phân dao động bằng phương pháp không gian trạng thái.

- Mô phỏng dao động của ô tô dựa theo phương pháp không gian trạng thái trong MATLAB, từ đó xuất ra các kết quả dạng đồ thị để dễ dàng kiểm tra, so sánh. Kết quả nghiên cứu giúp đánh giá tác động của quá trình phanh đến dao động thân xe của KAMAZ-43253, có thể được sử dụng làm cơ sở cho các nghiên cứu, phát triển sau này.

Tài liệu tham khảo

[1] Nguyễn Hữu Cẩn (2007), *Lý thuyết ô tô - máy kéo*, Nhà xuất bản Khoa học và Kỹ thuật, Hà Nội.

[2] Nguyễn Hoài Thanh, Nguyễn Thanh Việt, Bùi Xuân Lâm, *Ứng dụng Matlab trong tính toán kỹ thuật*, Đại học Sư phạm Kỹ Thuật TP. Hồ Chí Minh, 2002.

[3] Vũ Đức Lập (2004), *Sổ tay tra cứu tính năng kỹ thuật ô tô*, HVKTQS.

[4] Vũ Đức Lập (2011), *Dao động ô tô*, Nhà xuất bản Quân đội nhân dân, Hà Nội.

[5] Nguyễn Phúc Hiểu, Vũ Đức Lập (2001), *Lý thuyết ô tô quân sự*, NXB Quân đội nhân dân.

Effect of braking on the vibrational characteristics of the KAMAZ-43253 vehicle

Abstract: *In Vietnam, although research exists regarding the impact of braking processes on vehicle dynamics, these studies primarily focus on light-duty vehicles or passenger cars. Heavy-duty trucks, specifically models such as the KAMAZ-43253, remain under-researched. The significant payload and gross vehicle weight of these trucks render the braking process and associated vibrational effects considerably more complex and intense. On this basis, this paper develops a mathematical model and a system of differential equations, utilizing MATLAB software to simulate vehicle vibrations, serving as a fundamental reference for future studies.*

Keywords: *Braking process impact, axle system, Pitching motion, Suspension system.*

Xây dựng mô hình mô phỏng động lực học dẫn động điều khiển hệ thống phanh thủy lực có trợ lực chân không

Kiều Hữu Tài[1], Nguyễn Văn Trà[1]

[1]Học viện Kỹ thuật quân sự

Email: khtgoku1996@gmail.com; Tel: 0346468122

Tóm tắt:

Bài báo xây dựng mô hình động lực học dẫn động điều khiển của hệ thống phanh thủy lực có trợ lực chân không nhằm mô tả quá trình truyền lực và hình thành áp suất trong các mạch phanh theo chuỗi: Bàn đạp phanh, bầu trợ lực, xi lanh phanh chính hai dòng, bộ điều hòa lực phanh, dòng chảy dầu, cơ cấu phanh bánh xe. Mô hình được triển khai trên MATLAB/Simulink, cho phép mô phỏng diễn biến theo thời gian của lực đạp, lực đẩy xi lanh chính và áp suất trong các mạch phanh, đồng thời thể hiện đặc tính phân phối áp suất trước - sau của van điều hòa. Kết quả mô phỏng là cơ sở tham khảo để đánh giá đặc tính làm việc của hệ thống và phân tích quy luật phân phối áp suất trước - sau, qua đó đề xuất các định hướng cải tiến nhằm nâng cao hiệu quả và độ ổn định khi phanh.

Từ khóa: *Phanh thủy lực; Trợ lực chân không; Mô hình động lực học; Van điều hòa lực phanh.*

1. Đặt vấn đề

Hệ thống phanh thủy lực có trợ lực chân không là một hệ nhiều khâu gồm bàn đạp, bầu trợ lực, xi lanh chính, đường ống, van điều hòa và cơ cấu phanh. Trong khai thác, chất lượng phanh được quyết định chủ yếu bởi đáp ứng quá độ theo thời gian, thể hiện qua các chỉ tiêu: thời gian chậm tác dụng (từ khi tác động đến khi áp suất bắt đầu tăng), tốc độ tăng áp và khả năng đạt, duy trì áp suất yêu cầu. Việc đánh giá đầy đủ các chỉ tiêu này bằng thử nghiệm gặp nhiều hạn chế do chi phí lớn, yêu cầu an toàn nghiêm ngặt khi thử phanh mạnh, điều kiện thử khó lặp lại và khó cô lập ảnh hưởng của từng cụm chi tiết. Vì vậy, xây dựng mô hình động lực học dẫn động điều khiển và mô phỏng số là cần thiết nhằm dự đoán định lượng đáp ứng theo thời gian, phục vụ thiết kế và đánh giá hệ thống.

Trong các nghiên cứu mô phỏng trước đây, bầu trợ lực chân không thường được giản lược hoặc bỏ qua; quan hệ lực đạp - áp suất được xem gần như tĩnh hoặc quy đổi bằng hệ số khuếch đại. Cách tiếp cận này hạn chế độ chính xác khi phân tích lực bàn đạp thực tế, độ trễ và tốc độ tăng áp, cũng như ảnh hưởng của động học màng trợ lực, cơ cấu van điều khiển, lò xo phản lực và mức chân không đến quá trình tạo áp suất dầu. Đóng góp của nghiên cứu này là mô hình hóa bầu trợ lực chân không như một khâu động lực học trong mô hình tổng thể, qua đó phản ánh được tác động của áp suất chân không và cơ chế van - màng đến quá trình hình thành áp suất trong mạch dầu.

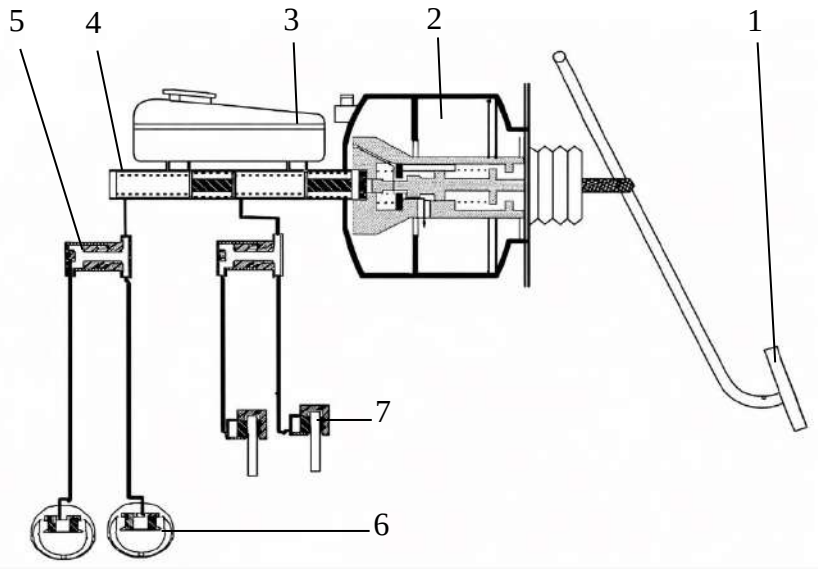
Mô hình được triển khai trên MATLAB/Simulink nhờ khả năng mô hình hóa dạng khối trực quan, thuận lợi liên kết các khâu cơ khí - thủy lực, hỗ trợ mô phỏng quá độ với nhiều bộ giải số và dễ dàng khảo sát tham số, trích xuất kết quả. Tuy nhiên, độ tin cậy của mô phỏng phụ thuộc đáng kể vào giả thiết và bộ tham số; mô hình chi tiết có thể làm tăng tải tính toán và độ cứng số, đòi hỏi lựa chọn bộ giải và bước thời gian phù hợp; đồng thời tồn tại chi phí bản quyền phần mềm.

Trên cơ sở mô hình đã xây dựng, mô phỏng được sử dụng để xác định thời gian chậm tác dụng, tốc độ tăng áp, áp suất cực đại và mức ổn định áp suất; phân tích quy luật áp suất

trong các mạch phanh và ảnh hưởng của van điều hòa đến phân phối áp suất trước - sau; hỗ trợ lựa chọn, tối ưu các tham số kết cấu (tỉ số bàn đạp, kích thước xi lanh chính, đặc tính van, đặc tính trợ lực). Mô hình có thể mở rộng cho các xe khác thông qua cập nhật tham số và điều kiện làm việc, phục vụ thiết kế, đánh giá và chẩn đoán kỹ thuật.

2. Xây dựng mô hình mô phỏng động lực học dẫn động điều khiển hệ thống phanh thủy lực có trợ lực chân không

Để xây dựng mô hình mô phỏng, trước hết cần mô tả cấu trúc và các khâu chính của hệ thống phanh thủy lực có trợ lực chân không trên xe UAZ-3160. Sơ đồ nguyên lý của hệ thống được thể hiện ở Hình 1.



Hình 1. Sơ đồ hệ thống dẫn động hệ thống phanh xe UAZ-3160

1. Bàn đạp phanh; 2. Bầu trợ lực chân không; 3. Bầu dầu; 4. Xi lanh chính; 5. Bộ điều hòa lực phanh; 6. Cơ cấu phanh sau; 7. Cơ cấu phanh trước.

2.1. Giả thiết mô hình

- Các tính chất của chất lỏng làm việc không bị thay đổi trong suốt quá trình chuyển tiếp;

- Bỏ qua các lực ma sát, khối lượng các pít tông xi lanh bánh xe;

- Nhiệt độ trong hệ thống dẫn động phanh và trong suốt quá trình làm việc đều bằng nhau không thay đổi;

- Không có rò rỉ trong hệ thống.

2.2. Mô hình từ bàn đạp phanh qua bầu trợ lực tới xi lanh phanh chính

Cơ cấu bàn đạp được mô hình như một cơ cấu đòn bẩy với các cánh tay. Với lực bàn đạp, lực quy dẫn tại cần đẩy tác dụng lên xi lanh chính được xác định xấp xỉ theo biểu thức:

$$F_{xl} = F_{bd}.\frac{a}{b} - x_h.k_h.\left(\frac{c}{b}\right)^2 - x_p.k_p \quad (1)$$

Trong đó: F_{xl}: Lực tác dụng lên xi lanh chính (N); F_{bd}: Lực bàn đạp phanh (N); x_h : Chuyển vị của bàn đạp phanh(m); x_p: Chuyển vị của van trợ lực (m); x_m: Chuyển vị của

màng trợ lực (m); a: Chiều dài bàn đạp phanh (m); b: vị trí piston xi lanh chính (m); c: vị trí lò xo hồi vị bàn đạp (m); k_p : Độ cứng của lò xo hồi vị bàn đạp (N/m); k_h : Độ cứng của lò xo pít tông xi lanh chính (N/m).

Xi lanh phanh chính biến đổi lực tác dụng thành áp suất thủy lực cho hai mạch phanh. Áp suất trong xi lanh chính được xác định từ cân bằng lực trên piston và diện tích mặt cắt được xác định theo biểu thức sau:

$$\begin{cases} p_{1xl} = \dfrac{F_{xl} + \Delta p.A_m - (k_m + k_p).x_m - (x_1 - x_2)k_1}{A_{xl}} \\ p_{2xl} = p_{1xl} + \dfrac{(x_1 - x_2)k_1 - k_2.x_2}{A_{xl}} \end{cases} \quad (2)$$

Trong đó: p_{1xl} : Áp suất thủy lực tại mạch sơ cấp của xi lanh phanh chính (N/m^2); p_{2xl} : Áp suất thủy lực tại mạch thứ cấp của xi lanh phanh chính (N/m^2); F_{xl} : Lực tác dụng lên cần đẩy truyền vào xi lanh chính (N); Δp : Chênh lệch áp suất giữa buồng khí quyển và buồng chân không trong bầu trợ lực (N/m^2); A_m : Diện tích làm việc của màng trợ lực (m^2); k_m : Độ cứng lò xo hồi vị của màng trợ lực (N/m); k_p : Độ cứng lò xo hồi vị của piston/van trong bầu trợ lực (N/m); x_m : Hành trình dịch chuyển tương đương của màng trợ lực (m); x_1 : Hành trình piston sơ cấp của xi lanh chính (m); x_2 : Hành trình piston thứ cấp của xi lanh chính (m); k_1 : Độ cứng lò xo hồi vị piston sơ cấp (N/m); k_2 : Độ cứng lò xo hồi vị piston thứ cấp (N/m); A_{xl} : Diện tích làm việc của piston xi lanh phanh chính (m^2).

2.3. Mô hình bộ điều hòa lực phanh

Bộ điều hòa lực phanh nhằm hạn chế áp suất đưa tới bánh sau khi áp suất hệ thống vượt một ngưỡng, để phù hợp với sự chuyển tải trọng động về cầu trước trong quá trình phanh. Đặc tính áp suất sau van có thể mô tả theo dạng tuyến tính từng đoạn:

$$p_{dhs} = \frac{A_{dhf}.(p_{1xl} - p_k)}{A_{dhr}} + p_k \quad (3)$$

$$p_{dht} = p_{1xl} - p_m \quad (4)$$

Trong đó: p_{dhs} : Áp suất dầu từ bộ điều hoà đến xilanh phanh trước (N/m^2); p_{dhs} : Áp suất dầu từ bộ điều hoà đến xilanh phanh sau (N/m^2); p_m : Áp suất trong xi lanh trước khi đạp phanh (áp suất dư của hệ thống) (N/m^2); p_k : Là trạng thái áp suất mà dầu bắt đầu không di chuyển đến bánh sau (N/m^2). A_{dhf} : Diện tích phần mặt trước của van điều hoà lực phanh (m^2); A_{dhr} : Diện tích phần mặt sau của van điều hoà lực phanh (m^2).

2.4. Mô hình dòng chảy dầu

Trong hệ thống phanh thủy lực, dòng dầu qua tiết lưu/đường ống được mô tả gần đúng theo quy luật kiểu "lỗ tiết lưu":

$$\dot{V}_i = \sigma_i C_{qwi} \sqrt{|p_{dhi} - p_{wi}|} \quad (i = 1...4) \quad (5)$$

Trong đó σ_i đảm bảo dòng chảy đúng chiều:

- Nếu $p_{dhi} > p_{wi}$ thì $\sigma_i = 1$ - Dầu chảy vào xi lanh bánh xe.

- Nếu $p_{dhi} < p_{wi}$ thì $\sigma_i = -1$ - Dầu có xu hướng chảy ngược.

Từ biểu thức trên, bình phương hai vế ta được:

$$(\dot{V}_i)^2 = (C_{qwi})^2 \left| p_{dhi} - p_{wi} \right| \quad (6)$$

Giả sử trong pha cấp áp (phanh) $p_{dhi} \geq p_{wi}$ (hoặc xét theo trị tuyệt đối), ta suy ra dạng liên hệ áp suất:

$$p_{wi} = p_{dhi} - (\frac{\dot{V}_i}{C_{qwi}})^2 \quad (7)$$

Từ đó thu được các biểu thức:

$$p_{w1} = p_{dh1} - (\frac{\dot{V}_1}{C_{qw1}})^2 \quad (8)$$

$$p_{w3} = p_{dh3} - (\frac{\dot{V}_3}{C_{qw3}})^2 \quad (9)$$

2.5. Mô hình cơ cấu phanh bánh xe

Trong mô hình cơ cấu phanh bánh xe, mô men phanh được giả thiết tỉ lệ tuyến tính với áp suất dầu phanh tác dụng lên xi lanh bánh xe, do đó:

$$M_{pw} = C_w . p_w \quad (10)$$

Trong đó: M_{pw}: Mô men phanh tại bánh xe(N.m); C_w: Hệ số kết cấu cơ cấu; p_w: Áp suất hữu hiệu tạo lực ép má phanh (N/m^2).

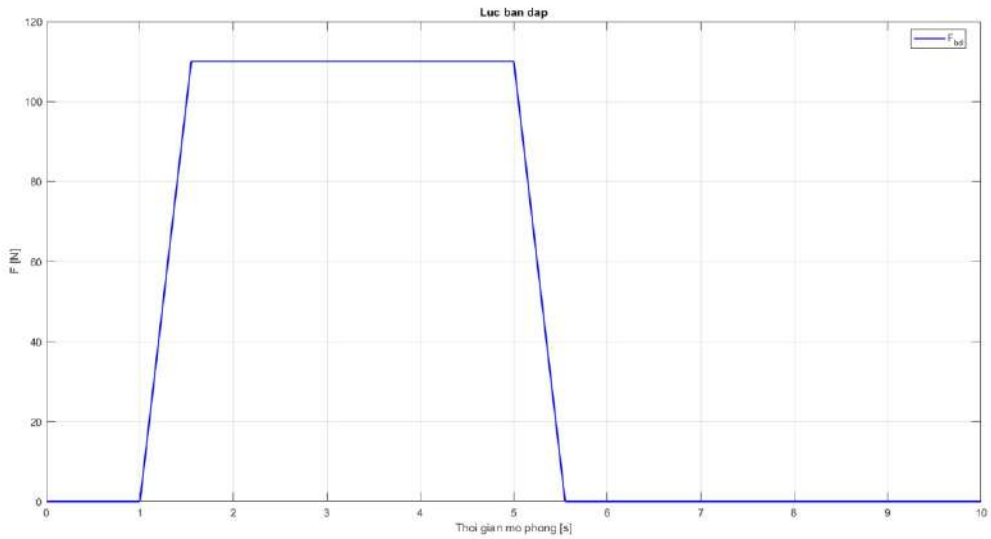
3. Kết quả mô phỏng và thảo luận

Bảng 1. Các thông số đầu vào của chương trình mô phỏng

STT	Thông số	Ký hiệu	Xe UAZ - 3160
1	Chiều dài bàn đạp phanh (m)	a	0,3
2	Vị trí piston xilanh chính (m)	b	0,06
3	Vị trí lò xo hồi vị bàn đạp (m)	c	0,05
4	Vị trí lò xo hồi vị cơ cấu phanh (m)	d	0,2
5	Khoảng cách piston - chốt quay CCP (m)	e	0,3
6	Đường kính xilanh phanh chính (m)	d_{xl}	0,025
7	Đường kính xilanh phanh bánh xe (m)	d_w	0,035
8	Đường kính ngoài màng trợ lực (m)	d_{nm}	0,25
9	Đường kính trong màng trợ lực (m)	d_{tm}	0,06
10	Đường kính piston đẩy (m)	d_{p1}	0,02

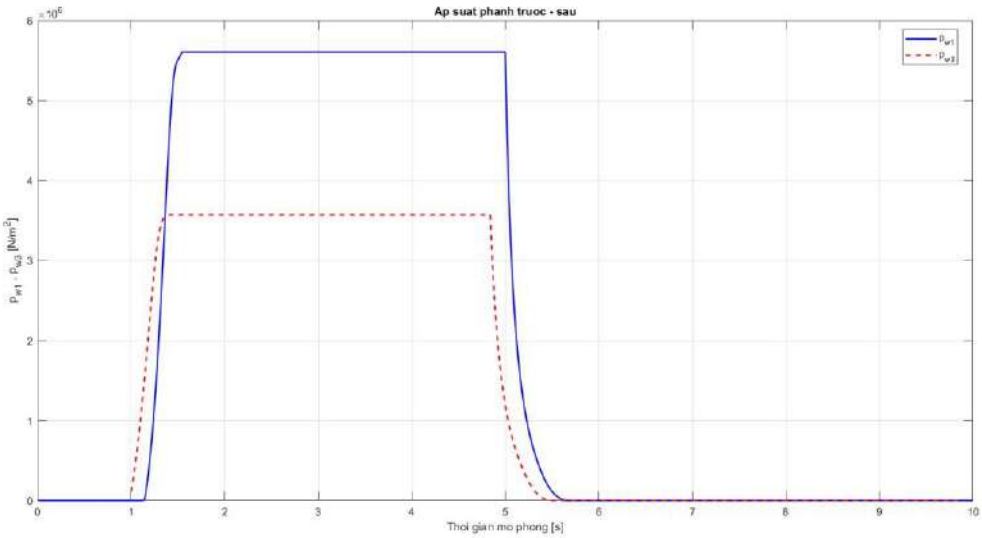
STT	Thông số	Ký hiệu	Xe UAZ - 3160
11	Độ cứng lò xo hồi vị piston 1 (N/m)	k_1	1000
12	Độ cứng lò xo hồi vị piston 2 (N/m)	k_2	100
13	Độ cứng lò xo hồi vị bàn đạp (N/m)	k_h	1000
14	Độ cứng lò xo hồi vị piston cường hóa (N/m)	k_p	100
15	Độ cứng lò xo hồi vị màng trợ lực (N/m)	k_m	500
16	Độ cứng lò xo hồi vị cơ cấu phanh (N/m)	k_w	5000
17	Hệ số dòng chảy tới xilanh bánh xe	C_{pw}	1e-8
18	Hệ số kết cấu cơ cấu phanh bánh xe	C_w	2.5e-3

Các kết quả đầu ra được trình bày trên các đồ thị sau:



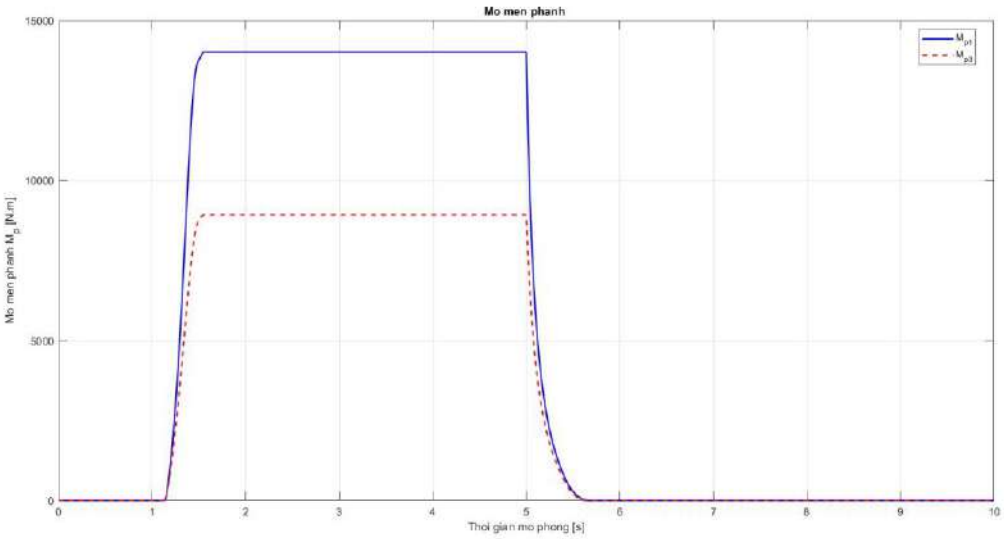
Hình 2. Đồ thị lực bàn đạp phanh

Hình 2 cho thấy lực bàn đạp được thiết lập theo quy luật tăng tuyến tính, giữ ổn định, giảm về 0 theo thời gian. Dạng kích thích này phù hợp để khảo sát đặc tính làm việc của hệ thống trong miền thời gian, vì cho phép quan sát rõ giai đoạn hình thành áp suất, giai đoạn duy trì áp suất khi giữ lực và quá trình xả áp khi nhả bàn đạp.



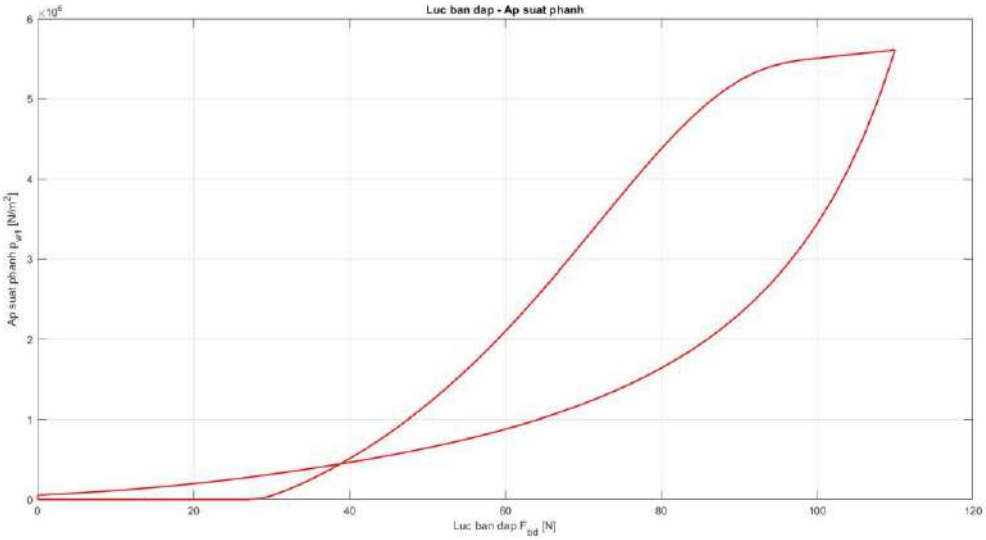
Hình 3. Đồ thị áp suất phanh trước sau

Trên Hình 3, áp suất trong mạch phanh tăng nhanh khi lực bàn đạp bắt đầu tác dụng và đạt đến vùng ổn định trong giai đoạn giữ lực. Áp suất mạch trước luôn lớn hơn áp suất mạch sau và duy trì chênh lệch tương đối rõ rệt, phản ánh tác dụng phân phối áp suất trước - sau của van điều hòa theo đặc tính đã mô hình hóa. Khi nhả bàn đạp, áp suất ở cả hai mạch giảm về gần 0, thể hiện quá trình xả áp và kết thúc tác dụng phanh.



Hình 4. Đồ thị mô men phanh trước, sau

Hình 4 cho thấy mô men phanh biến thiên cùng quy luật với áp suất trong các mạch, do mô men phanh được giả thiết tỉ lệ tuyến tính với áp suất tác dụng lên cơ cấu phanh bánh xe. Mô men phanh trước đạt giá trị cao hơn mô men phanh sau trong toàn bộ quá trình phanh, phù hợp với xu hướng phân phối lực phanh ưu tiên cầu trước và tương ứng với đặc tính áp suất trước - sau ở Hình 3.



Hình 5. Đồ thị lực bàn đạp, áp suất phanh

Hình 5 thể hiện quan hệ giữa lực bàn đạp và áp suất phanh có dạng phi tuyến và xuất hiện vòng trễ giữa pha tăng lực và pha giảm lực. Đặc điểm này phản ánh tính chất động của hệ thống: áp suất không biến thiên tức thời theo lực bàn đạp mà chịu ảnh hưởng của động lực học bầu trợ lực, độ nén của dầu và tổn thất/tiết lưu trong dòng chảy, dẫn đến độ trễ nhất định khi tăng áp và khi xả áp. Vì vậy, đồ thị là minh chứng cho việc cần mô hình hóa theo miền thời gian thay vì chỉ dùng quan hệ tĩnh lực - áp suất.

4. Kết luận

Bài báo đã xây dựng mô hình động lực học dẫn động điều khiển của hệ thống phanh thủy lực có trợ lực chân không theo chuỗi: Bàn đạp phanh, bầu trợ lực, xi lanh phanh chính hai dòng, bộ điều hòa lực phanh, dòng chảy dầu, cơ cấu phanh bánh xe và triển khai trên MATLAB/Simulink. Mô hình cho phép mô phỏng diễn biến theo thời gian của lực đạp, lực đẩy tác dụng lên xi lanh chính và áp suất trong các mạch phanh, đồng thời phản ánh đặc tính phân phối áp suất trước - sau của van điều hòa.

Kết quả mô phỏng cho thấy áp suất trong các mạch tăng theo quy luật tác động lực bàn đạp và ổn định trong giai đoạn giữ lực; áp suất mạch trước lớn hơn mạch sau, thể hiện vai trò của bộ điều hòa lực phanh trong phân phối áp suất trước - sau. Mô men phanh biến thiên tương ứng với áp suất mạch phanh theo giả thiết của mô hình cơ cấu phanh bánh xe. Quan hệ lực bàn đạp - áp suất xuất hiện vòng trễ, phản ánh đặc tính động của quá trình hình thành và xả áp trong hệ thống.

Mô hình và kết quả thu được có thể sử dụng làm cơ sở tham khảo để đánh giá đặc tính làm việc theo thời gian của hệ thống phanh, phân tích quy luật phân phối áp suất trước–sau và hỗ trợ đề xuất các định hướng cải tiến về cấu trúc/tham số của các khâu trong dẫn động. Trong nghiên cứu tiếp theo, cần hiệu chỉnh bộ tham số theo dữ liệu thử nghiệm và mở rộng mô hình theo hướng tích hợp với mô hình động lực học bánh xe để đánh giá trong các chế độ phanh khác nhau và điều kiện mặt đường thay đổi.

Tài liệu tham khảo

[1] Nguyễn Sĩ Đỉnh (2010), *Nghiên cứu ĐLH dẫn động phanh thủy lực để lắp hệ thống chống hãm cứng bánh xe lên ô tô quân sự,* Luận văn Tiến sỹ, Học viện Kỹ thuật Quân sự, 127 trang.

[2] Nguyễn Phúc Hiểu, Vũ Đức Lập (2002), *Lý thuyết ô tô quân sự*, NXB Quân đội Nhân dân.

[3] Nguyễn Văn Kiều (2002), *Thủy khí động lực kỹ thuật*, NXB Quân đội nhân dân, Hà Nội.

[4] Vũ Đức Lập (2002), Động lực học phanh ô tô, NXB Quân đội Nhân dân.

[5] Dương Tiến Minh (2004), *Nghiên cứu chất lượng phanh ô tô quân sự với hệ thống phanh có lắp bộ điều hòa lực phanh,* Luận văn Tiến sỹ, Học viện Kỹ thuật Quân sự, 133 trang.

[6] Nguyễn Phùng Quang (2004), *Matlab&Simulink dành cho kỹ sư điều khiển tự động,* NXB khoa học và kỹ thuật, Hà Nội.

[7] Т.М.Башта, (1971), *Машиностроительная гидравлика*, Машиностроение, Москва.

[8] Н. Ф. Метлюк, В. П. Автушко (1980), *Динамика пневматических и гидравлических приводов автомобилей*, Машиностроение, Москва.

[9] Нгуен Ныок (1987), *Динамика пневматического тормозного привода автомобиля*, Минск.

Development of a dynamic model for vacuum-assisted hydraulic brake actuation and control

***Abstract**: This paper develops a dynamic model of a vacuum-assisted hydraulic brake actuation and control system to describe the force transmission process and the build-up of hydraulic pressure in the brake circuits along the chain: brake pedal, vacuum booster, tandem master cylinder, brake proportioning valve, hydraulic flow in the lines, and wheel brake mechanism. The model is implemented in MATLAB/Simulink, enabling time-domain simulation of pedal force, master-cylinder pushrod force, and pressures in the brake circuits, while also capturing the front–rear pressure distribution characteristics of the proportioning valve. The simulation results provide a reference basis for evaluating system operating characteristics and analyzing the law of front–rear pressure distribution, thereby supporting recommendations for design improvements to enhance braking efficiency and stability.*

***Keywords**: Hydraulic braking; Vacuum brake booster; Dynamic modeling; Brake proportioning valve.*

Ứng dụng mô hình Markov nghiên cứu sự biến đổi tình trạng kỹ thuật hệ thống phanh trên xe KAMAZ-53212

Nguyễn Trung Tiến[1], Nguyễn Văn Dũng[1]

[1]Học viện Kỹ thuật quân sự

Email: leonguyen070896@gmail.com; Tel: 0336888329

Tóm tắt

Hệ thống phanh khí nén trên xe tải hạng nặng có kết cấu nhiều dòng dẫn động và chịu ảnh hưởng lớn của điều kiện khai thác, làm cho tình trạng kỹ thuật của hệ thống biến đổi phức tạp theo thời gian. Bài báo đề xuất phương pháp chẩn đoán tình trạng kỹ thuật hệ thống phanh khí nén trên xe KAMAZ-53212 dựa trên mô hình Markov ẩn Gaussian (Gaussian Hidden Markov Model – GHMM). Trên cơ sở phân tích kết cấu và nguyên lý làm việc của hệ thống phanh khí nén, các trạng thái ẩn đại diện cho các trạng thái làm việc và dạng hư hỏng đặc trưng được xây dựng. Vectơ quan sát được xác định từ các đại lượng đo trực tiếp như áp suất nguồn, áp suất tại các bầu phanh và quan hệ phân phối áp suất giữa các mạch phanh. Kết quả mô phỏng cho thấy mô hình GHMM cho phép suy luận hợp lý sự biến đổi trạng thái kỹ thuật của hệ thống phanh theo thời gian, phản ánh được các hư hỏng cục bộ và hiện tượng xen kẽ trạng thái trong quá trình khai thác thực tế.

Từ khóa: *Hệ thống phanh khí nén; Markov ẩn Gaussian; chẩn đoán trạng thái kỹ thuật; xe KAMAZ-53212.*

1. Mở đầu

Các thống kê thực tế cho thấy một tỷ lệ đáng kể các tai nạn giao thông liên quan đến xe tải và xe khách có nguyên nhân xuất phát từ sự cố của hệ thống phanh. Những tai nạn này không chỉ gây thiệt hại lớn về kinh tế mà còn dẫn đến thương tích nghiêm trọng và nguy cơ tử vong. Điều đó cho thấy hệ thống phanh giữ vai trò đặc biệt quan trọng đối với tính an toàn chuyển động của ô tô, nhất là đối với xe tải hạng nặng làm việc trong điều kiện tải trọng lớn và thời gian khai thác kéo dài. Vì vậy, việc nghiên cứu đánh giá tình trạng kỹ thuật và độ tin cậy của hệ thống phanh là hết sức cần thiết.

Trong quá trình khai thác, hệ thống phanh khí nén trên xe tải hạng nặng có kết cấu nhiều dòng dẫn động và nhiều phần tử chức năng, chịu ảnh hưởng đồng thời của sự lão hóa, chế độ làm việc và điều kiện môi trường. Các hư hỏng của hệ thống không chỉ xuất hiện dưới dạng suy giảm liên tục của một phần tử riêng lẻ mà còn xảy ra cục bộ, xen kẽ giữa các dòng dẫn động phanh phanh. Đặc điểm này gây khó khăn cho công tác chẩn đoán tình trạng kỹ thuật nếu chỉ sử dụng các phương pháp đánh giá dựa trên ngưỡng thông số hoặc kiểm tra độc lập từng phần tử [1].

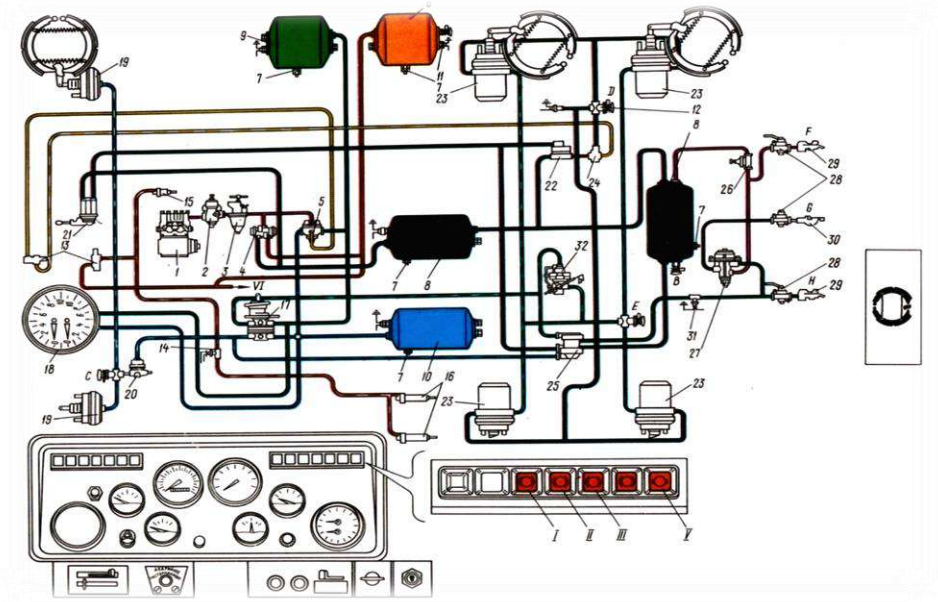
Để nghiên cứu, đánh giá và dự báo sự biến đổi tình trạng kỹ thuật của các hệ thống kỹ thuật phức tạp, nhiều phương pháp khác nhau đã được đề xuất, trong đó các phương pháp dựa trên lý thuyết Markov được sử dụng rộng rãi do khả năng mô tả quá trình chuyển trạng thái ngẫu nhiên của hệ thống theo thời gian [2]. Với giả thiết xác suất chuyển trạng thái đặc trưng cho từng dạng hư hỏng và khả năng phục hồi, mô hình Markov cho phép phân tích, đánh giá và dự báo tình trạng kỹ thuật cũng như độ tin cậy của các cụm và phần tử trong hệ thống [3].

Xuất phát từ yêu cầu đó, bài báo này tập trung xây dựng và ứng dụng mô hình Markov ẩn Gaussian (Gaussian Hidden Markov Model – GHMM) để chẩn đoán tình trạng kỹ thuật hệ thống phanh khí nén trên xe KAMAZ-53212 [4]. Thông qua việc sử dụng các đại lượng áp suất đo được làm vectơ quan sát, mô hình cho phép suy luận các trạng thái kỹ thuật ẩn gắn với các

dạng làm việc và hư hỏng đặc trưng của hệ thống phanh. Kết quả nghiên cứu là cơ sở khoa học cho việc đánh giá tình trạng kỹ thuật, hỗ trợ lập kế hoạch bảo dưỡng và nâng cao độ an toàn trong khai thác xe tải hạng nặng.

2. Phân tích kết cấu hệ thống phanh khí nén trên xe KAMAZ-53212

Hệ thống phanh khí nén xe KAMAZ-53212 được thể hiện trên hình 1, đây là hệ thống phanh nhiều dòng, trong đó máy nén khí cung cấp khí nén và được tích trữ trong các bình chứa. Từ các bình chứa, khí nén được phân phối qua các van bảo vệ và tổng van phanh đến các bầu phanh cầu xe.



Hình 1. Sơ đồ nguyên lý dẫn động phanh khí nén ô tô KAMAZ-53212

1-Máy nén khí; 2-Điều chỉnh áp suất; 3-Bộ chống đông đặc; 4-van bảo vệ hai ngả; 5-van bảo vệ ba ngả; 6-bình khí nén mạch thứ tư; 7-van chích không khí; 8-bình khí nén cho mạch thứ ba; 9-bình khí nén mạch thứ hai;10-bình khí nén mạch một; 11-van đầu kiểm tra các đường dây bốn; 12-van kiểm tra đầu ra; 13-van nhả phanh sự cố; 14-cảm biến điện hơi; 15-xi lanmh cung cấp nhiên liệu; 16-xi lanh dẫn động van chuyên dùng; 17-van phanh; 18-đồng hồ áp suất hai kim; 19-bầu phanh của các bánh xe cầu trước; 20-van hạn chế áp suất; 21-van phanh tay tác dụng ngược; 22-van tăng tốc; 23-bầu phanh và bình tích năng; 24-van điều khiển phanh rơ moóc dẫn động 2 đường dẫn; 25-van phanh rơ mooc; 26-van bảo vệ một ngả; 27-van phanh rơ moóc một đường dẫn; 28-van ngắt;29-đầu nối kiểu ПапМ; 30- đầu nối kiểu A; 31-cảm biến bật tín hiệu phanh; 32-bộ điều hòa lực phanh; A-đầu ra kiểm tra các vòng phanh IV; B và E-van các đầu ra kiểm tra các vòng phanh III: C-van kiểm tra đầu ra các vòng phanh I; D-đầu ra kiểm tra các vòng phanh II; F-đường cung cấp khí nén cho hệ thống hai đường dẫn; G-đường cung cấp khí nén cho hệ thống một đường dẫn; H-đường điều khiển phanh rơ moóc hai đường dẫn.

Xét theo chức năng, hệ thống có thể chia thành các cụm chính:

+ Cụm nguồn khí: máy nén khí, bộ điều chỉnh áp suất, bình chứa khí.

+ Cụm phân phối và bảo vệ: van bảo vệ, van phân mạch.

+ Cụm điều khiển: tổng van phanh.

+ Cụm chấp hành: bầu phanh cầu trước và bầu phanh cụm cầu giữa, sau.

Trong quá trình khai thác, các hư hỏng thường gặp bao gồm: suy giảm khả năng duy trì áp suất nguồn, rò rỉ mạng khí, sai lệch phân phối giữa các mạch phanh, suy giảm hiệu quả tại các bầu phanh. Các hư hỏng này có thể xuất hiện đồng thời hoặc xen kẽ, tạo nên quá trình suy giảm phức tạp của hệ thống.

3. Xây dựng mô hình Markov ẩn Gaussian

3.1. Cấu trúc tổng quát của mô hình

Hệ thống phanh khí nén trong quá trình khai thác được mô hình hóa như một quá trình Markov ẩn Gaussian rời rạc theo thời gian. Tại mỗi thời điểm t, hệ thống tồn tại ở một trạng thái kỹ thuật ẩn X_t, trong khi người khai thác chỉ có thể đo được vectơ quan sát O_t.

Trên cơ sở phân tích hệ thống, nghiên cứu này xác định tám trạng thái kỹ thuật ẩn $X_0 \div X_7$, tương ứng với các mức độ và dạng suy giảm khác nhau của hệ thống phanh khí nén, như sau:

Trạng thái bình thường (X_0): Hệ thống phanh làm việc ổn định, áp suất nguồn và áp suất tại các bầu phanh nằm trong dải cho phép, phân phối áp suất giữa các mạch phanh hợp lý. Đây là trạng thái chiếm ưu thế ở giai đoạn đầu khai thác hoặc sau bảo dưỡng.

Suy giảm khả năng tạo và duy trì áp suất nguồn (X_1): Trạng thái này liên quan đến máy nén khí, bộ điều chỉnh áp suất hoặc bộ chống đông đặc. Biểu hiện đặc trưng là tốc độ nạp áp chậm, áp suất nguồn đạt thấp hoặc dao động.

Rò rỉ hoặc hao áp mạng nguồn - bình chứa (X_2): Trạng thái phản ánh sự suy giảm nền của hệ thống khí nén do rò rỉ đường ống, khớp nối hoặc độ kín của bình chứa. Áp suất nguồn giảm dần theo thời gian hoặc khó duy trì khi hệ thống làm việc.

Sai lệch phân phối áp suất giữa các mạch phanh (X_3): Trạng thái này liên quan đến các van bảo vệ và van phân mạch. Đặc trưng bởi sự chênh lệch bất thường giữa áp suất các mạch phanh, trong khi áp suất nguồn vẫn ở mức chấp nhận được.

Bất thường tổng van phanh (X_4): Trạng thái phản ánh các sai lệch trong quan hệ giữa tác động điều khiển và áp suất ra tại các bầu phanh. Các bất thường này thường mang tính tạm thời, xuất hiện trong thời gian ngắn.

Suy giảm cụm chấp hành phanh cầu trước (X_5): Trạng thái đặc trưng bởi áp suất tại bầu phanh cầu trước tăng chậm hoặc suy giảm nhanh so với các mạch còn lại, trong khi nguồn khí vẫn tương đối ổn định.

Suy giảm cụm chấp hành phanh cụm cầu giữa - sau (X_6): Trạng thái phản ánh sự suy giảm hiệu quả của bầu phanh tại cụm cầu giữa - sau, thường xuất hiện trên nền suy giảm nguồn khí và có thể xen kẽ với trạng thái rò rỉ mạng nguồn.

Trạng thái mất an toàn, khuyến nghị dừng khai thác (X_7): Trạng thái này biểu thị mức suy giảm nghiêm trọng, khi áp suất nguồn và áp suất tại các bầu phanh không còn đảm bảo an toàn chuyển động. Trong khai thác thực tế, trạng thái này thường dẫn đến dừng xe hoặc sửa chữa khẩn cấp.

Mô hình Markov ẩn Gaussian được đặc trưng bởi bộ tham số:

$$\lambda = \left(\pi, A, \{\mu_i, \Sigma_i\}\right), i = 0,1,...,7 \quad (1)$$

Trong đó:

+ π: phân bố xác suất ban đầu của trạng thái kỹ thuật;

+ A: ma trận xác suất chuyển trạng thái;

+ μ_i, Σ_i: tham số của phân bố Gaussian ứng với trạng thái ẩn X_i.

3.2. Phân bố xác suất ban đầu π

Vectơ xác suất ban đầu π phản ánh trạng thái kỹ thuật của hệ thống phanh tại thời điểm bắt đầu khai thác hoặc bắt đầu quan sát. Trong thực tế, phần lớn xe khi đưa vào sử dụng hoặc sau bảo dưỡng đều ở trạng thái làm việc bình thường, trong khi xác suất xuất hiện ngay các trạng thái suy giảm là nhỏ.

Trong nghiên cứu này, phân bố xác suất ban đầu được lựa chọn như sau:

$$\pi = \begin{bmatrix} 0.98 & 0.01 & 0.01 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (2)$$

tương ứng với các trạng thái:

$$\left(X_0, X_1, X_2, X_3, X_4, X_5, X_6, X_7\right)$$

Trong đó:

+ X_0 : trạng thái bình thường chiếm ưu thế tuyệt đối;

+ X_1, X_2: có xác suất rất nhỏ để phản ánh khả năng tồn tại suy giảm tiềm ẩn ngay từ đầu;

+ Các trạng thái hư hỏng khác được giả thiết chưa xuất hiện.

3.3. Ma trận xác suất chuyển trạng thái A

Ma trận chuyển trạng thái $A = \left[a_{\text{ij}}\right]$ mô tả khả năng hệ thống chuyển từ trạng thái X_i sang trạng thái X_j trong một bước thời gian:

$$a_{\text{ij}} = P\left(X_t = X_j \middle| X_{t-1} = X_i\right) \quad (3)$$

Ma trận *A* được xây dựng dựa trên logic suy giảm kỹ thuật trong khai thác thực tế, với các nguyên tắc sau:

+ Xác suất duy trì trạng thái hiện tại là lớn (phản ánh tính ổn định ngắn hạn);

+ Xác suất chuyển sang trạng thái suy giảm liền kề lớn hơn so với chuyển sang trạng thái suy giảm nặng;

+ Trạng thái X_7 có tính chất gần hấp thụ.

Ma trận chuyển trạng thái sử dụng trong mô phỏng có dạng:

$$A = \begin{bmatrix} 0.88 & 0.03 & 0.03 & 0.02 & 0.02 & 0.01 & 0.01 & 0 \\ 0.10 & 0.80 & 0.03 & 0.02 & 0.03 & 0 & 0 & 0.02 \\ 0.08 & 0.02 & 0.82 & 0.03 & 0.02 & 0.01 & 0.01 & 0.01 \\ 0.10 & 0.01 & 0.02 & 0.80 & 0.03 & 0.02 & 0.02 & 0 \\ 0.08 & 0.02 & 0.02 & 0.04 & 0.78 & 0.02 & 0.02 & 0.02 \\ 0.10 & 0 & 0.02 & 0.03 & 0.02 & 0.78 & 0.02 & 0.03 \\ 0.10 & 0 & 0.02 & 0.03 & 0.02 & 0.02 & 0.78 & 0.03 \\ 0.05 & 0 & 0 & 0.02 & 0 & 0.02 & 0.02 & 0.89 \end{bmatrix} \quad (4)$$

3.4. Vectơ quan sát và mô hình phát xạ Gaussian

3.4.1. Vectơ quan sát O_t

Tại mỗi thời điểm t, vectơ quan sát được xác định:

$$O_t = \begin{bmatrix} P_{res}(t) \\ P_I(t) \\ P_{II}(t) \\ P_{bf}(t) \\ P_{bms}(t) \end{bmatrix} \tag{5}$$

Trong đó:

+ P_{res}: áp suất bình chứa chính;

+ P_I: áp suất bình khí nén mạch phanh cầu trước;

+ P_{II}: áp suất bình khí nén mạch phanh cầu giữa - sau;

+ P_{bf}: áp suất bầu phanh cầu trước;

+ P_{bms}: áp suất bầu phanh cụm cầu giữa - sau.

Các đại lượng này đều có thể đo trực tiếp bằng cảm biến áp suất trên hệ thống phanh khí nén. Dữ liệu quan sát theo mô hình Gaussian đa biến ứng với từng trạng thái kỹ thuật ẩn.

3.4.2. Tham số Gaussian μ_i, Σ_i

a) Mô hình phát xạ Gaussian

Trong mô hình Markov ẩn Gaussian, xác suất xuất hiện vectơ quan sát O_t khi hệ thống ở trạng thái X_i được mô tả bởi phân bố Gaussian đa biến:

$$p\left(O_t \middle| q_t = X_i\right) = \mathrm{N}\left(O_t \middle| \mu_i, \Sigma_i\right) \tag{6}$$

hay viết đầy đủ:

$$p\left(O_t \middle| X_i\right) = \frac{1}{\left(2\pi\right)^{d/2} \left|\Sigma_i\right|^{1/2}} \exp\left(-\frac{1}{2}\left(O_t - \mu_i\right)^T \Sigma_i^{-1} \left(O_t - \mu_i\right) \right) \tag{7}$$

Trong đó:

+ d là số chiều của vectơ quan sát;

+ μ_i là vectơ trung bình;

+ Σ_i là ma trận hiệp phương sai của trạng thái X_i.

b) Phương pháp xây dựng tham số μ_i

Với mỗi trạng thái ẩn X_i, từ dữ liệu đo được xác định tập các quan sát:

$$\Omega_i = \left\{ O_t \middle| q_t = X_i \right\} \tag{8}$$

Vectơ trung bình của trạng thái X_i được tính theo:

$$\mu_i = \frac{1}{\left|\Omega_i\right|} \sum_{O_t \in \Omega_i} O_t \tag{9}$$

Chuỗi khảo sát gồm $T = 100$ bước thời gian, giá trị μ_i được thể hiện trên bảng 1.

Bảng 1. Thông số vector trung bình của các trạng thái

Trạng thái	P_{res}	P_I	P_{II}	P_{bf}	P_{bms}
X0	8.2	8.1	8.0	3.8	3.7
X1	7.2	7.0	6.9	3.4	3.3
X2	8.0	7.4	7.2	3.5	3.4
X3	8.1	8.1	7.1	3.7	3.2
X4	8.1	8.0	7.9	2.7	2.6
X5	8.1	8.0	7.9	2.8	3.5
X6	8.1	8.0	7.8	3.5	2.7
X7	6.7	6.3	6.1	2.2	2.0

c) Phương pháp xây dựng tham số Σ_i

Ma trận hiệp phương sai của trạng thái X_i được xác định theo:

$$\Sigma_i = \frac{1}{|\Omega_i| - 1} \sum_{O_t \in \Omega_i} (O_t - \mu_i)(O_t - \mu_i)^T \tag{10}$$

Trong nghiên cứu, để đơn giản hóa việc tính toán và làm rõ đặc trưng dao động của từng đại lượng áp suất, ma trận hiệp phương sai được giả thiết có dạng đường chéo:

$$\Sigma_i = \begin{bmatrix} \sigma_{1,i}^2 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{2,i}^2 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{3,i}^2 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{4,i}^2 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{5,i}^2 \end{bmatrix} \tag{11}$$

Trong đó:

+ $\sigma_{1,i}^2 = Var(P_{res})$

+ $\sigma_{2,i}^2 = Var(P_I)$

+ $\sigma_{3,i}^2 = Var(P_{II})$

+ $\sigma_{4,i}^2 = Var(P_{bf})$

+ $\sigma_{5,i}^2 = Var(P_{bms})$

Giả thiết này tương đương với việc coi các đại lượng áp suất độc lập có điều kiện theo trạng thái ẩn, tức là khi biết hệ thống đang ở trạng thái X_i, dao động của từng mạch được xem xét riêng biệt.

d) Quy trình tổng quát xây dựng tham số Gaussian từ dữ liệu O_t

Quy trình được thực hiện theo các bước:

1. Đo và thu thập dữ liệu O_t tại các vị trí đã xác định;
2. Tiền xử lý dữ liệu (lọc nhiễu, đồng bộ thời gian);
3. Phân vùng dữ liệu theo trạng thái kỹ thuật dựa trên tiêu chí vật lý và ngưỡng kỹ thuật;
4. Tính μ_i và $\sum_i$ cho từng trạng thái;
5. Kiểm tra và hiệu chỉnh tham số để đảm bảo tính ổn định số và ý nghĩa vật lý.

3.5. Ý nghĩa mô hình

Việc xác định rõ ràng các tham số $\pi, A, \{\mu_i, \Sigma_i\}$ cho phép mô hình Markov ẩn Gaussian:

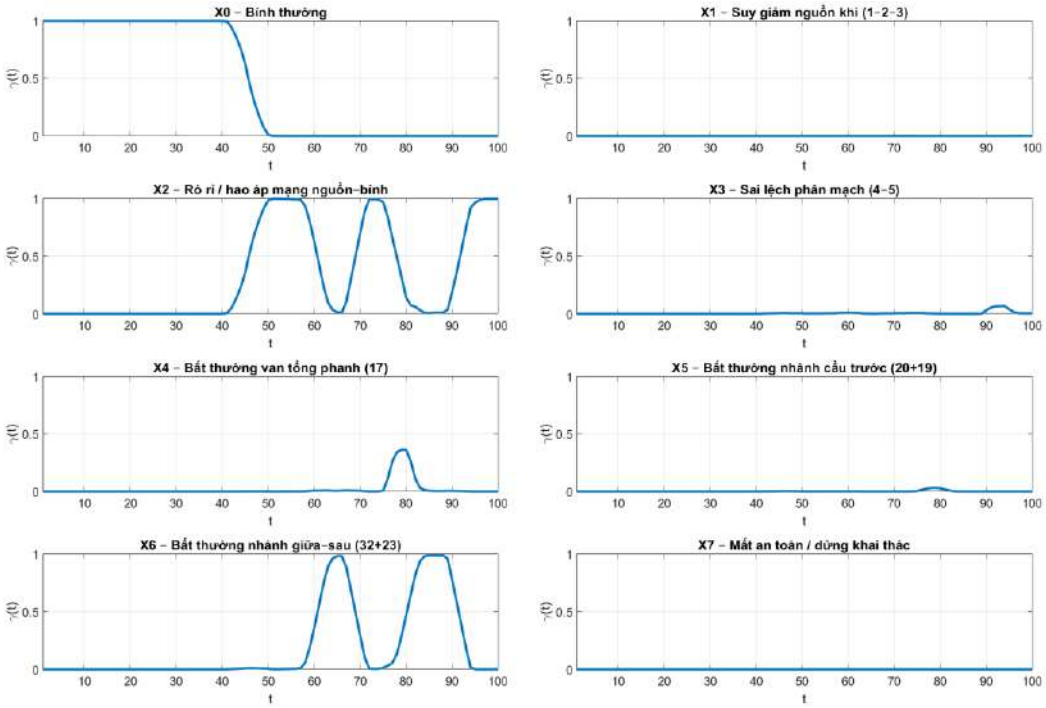
+ Mô tả định lượng quá trình suy giảm tình trạng kỹ thuật của hệ thống phanh;

+ Liên kết trực tiếp các trạng thái ẩn với các đại lượng đo được;

+ Tạo cơ sở cho việc áp dụng các thuật toán suy luận (Forward–Backward, Viterbi) để chẩn đoán trạng thái kỹ thuật theo thời gian.

4. Kết quả mô phỏng và bàn luận

4.1 Kết quả mô phỏng



Hình 2. Diễn biến xác suất các trạng thái kỹ thuật ẩn của hệ thống phanh

Trên cơ sở mô hình Markov ẩn Gaussian đã xây dựng, quá trình suy luận trạng thái kỹ thuật của hệ thống phanh khí nén trên xe KAMAZ-53212 được thực hiện cho chuỗi quan sát trong khoảng khảo sát $t = 0 \div 100$. Kết quả thu được dưới dạng xác suất hậu nghiệm của các trạng thái kỹ thuật ẩn phản ánh trực tiếp diễn biến suy giảm của hệ thống trong quá trình khai thác.

Kết quả mô phỏng cho thấy trạng thái bình thường X_0 chỉ chiếm ưu thế trong giai đoạn đầu, sau đó nhanh chóng bị thay thế bởi các trạng thái suy giảm. Đáng chú ý, trạng thái rò rỉ/hao áp mạng nguồn - bình X_2 xuất hiện sớm và duy trì theo từng khoảng thời gian, đóng vai trò như một trạng thái suy giảm nền của hệ thống khí nén.

Trên nền suy giảm này, trạng thái suy giảm cục bộ tại nhánh phanh cụm cầu giữa – sau X_6 xuất hiện với xác suất cao trong một số giai đoạn, tuy nhiên không lấn át hoàn toàn trạng

thái X_2 mà xen kẽ với nó. Hiện tượng này phản ánh đặc điểm khai thác thực tế của hệ thống phanh khí nén: các hư hỏng tại cụm chấp hành thường là hệ quả của sự suy giảm nguồn khí nền, chứ không xuất hiện độc lập.

Các trạng thái bất thường khác, như bất thường tổng van phanh X_4, chỉ xuất hiện trong khoảng thời gian ngắn với xác suất thấp. Điều này cho thấy các hư hỏng dạng điều khiển thường mang tính tạm thời và khó duy trì ổn định nếu xe vẫn tiếp tục vận hành. Trạng thái mất an toàn X_7 không xuất hiện trong mô phỏng, phản ánh trường hợp xe vẫn được khai thác dù tồn tại các dạng suy giảm kỹ thuật ở mức cho phép.

Như vậy, kết quả mô phỏng không chỉ phản ánh sự xuất hiện của từng dạng hư hỏng riêng lẻ mà còn cho thấy mối quan hệ phụ thuộc và chồng lấn giữa các trạng thái suy giảm, điều mà các phương pháp chẩn đoán tĩnh khó có thể nhận diện.

4.2. Bàn luận

Khác với cách tiếp cận chẩn đoán truyền thống, trong đó mỗi hư hỏng thường được xem xét như một trạng thái độc lập và được xác định thông qua các ngưỡng thông số cố định, nghiên cứu này tiếp cận vấn đề theo hướng mô hình hóa quá trình suy giảm tình trạng kỹ thuật như một quá trình ngẫu nhiên có cấu trúc.

Hư hỏng thu được từ mô hình Markov ẩn Gaussian cho thấy hệ thống phanh khí nén không chuyển trực tiếp từ trạng thái bình thường sang trạng thái hư hỏng nghiêm trọng, mà trải qua các giai đoạn suy giảm nền và suy giảm cục bộ đan xen. Đặc biệt, trạng thái suy giảm nhánh giữa – sau X_6 chỉ xuất hiện khi trạng thái rò rỉ/hao áp mạng nguồn X_2 đã hình thành, cho thấy mối quan hệ nhân quả giữa suy giảm nguồn khí và suy giảm hiệu quả chấp hành phanh.

Cách biểu diễn kịch bản hư hỏng thông qua quỹ đạo xác suất của các trạng thái ẩn cho phép đánh giá tình trạng kỹ thuật theo hướng liên tục và xác suất, thay vì phân loại rạch ròi "tốt/hỏng". Đây chính là điểm khác biệt căn bản và cũng là đóng góp khoa học chính của nghiên cứu: ***chuyển từ chẩn đoán dựa trên ngưỡng thông số sang chẩn đoán dựa trên quá trình suy giảm xác suất.***

Ngoài ra, việc mô hình hóa các trạng thái suy giảm chồng lấn giúp phản ánh sát hơn thực tế khai thác xe tải nặng, nơi nhiều dạng hư hỏng có thể cùng tồn tại trong một khoảng thời gian mà chưa buộc phải dừng khai thác phương tiện.

5. Kết luận

Bài báo đã xây dựng và ứng dụng thành công mô hình Markov ẩn Gaussian để đánh giá tình trạng kỹ thuật hệ thống phanh khí nén trên xe KAMAZ-53212. Trên cơ sở phân tích kết cấu hệ thống và dữ liệu quan sát đo được, các trạng thái kỹ thuật ẩn đã được xác định phù hợp với các dạng hư hỏng đặc trưng trong khai thác thực tế.

Kết quả mô phỏng cho thấy mô hình đề xuất có khả năng mô tả hợp lý quá trình suy giảm tình trạng kỹ thuật, làm rõ mối quan hệ giữa suy giảm nền và suy giảm cục bộ của hệ thống phanh. Đặc biệt, việc biểu diễn kịch bản hư hỏng dưới dạng xác suất trạng thái cho phép nhận diện sớm xu hướng suy giảm, ngay cả khi hệ thống chưa đạt đến trạng thái mất an toàn.

Cách tiếp cận này mở ra khả năng ứng dụng mô hình Markov ẩn Gaussian trong công tác chẩn đoán và lập kế hoạch bảo dưỡng dựa trên tình trạng kỹ thuật thực tế đối tượng nghiên cứu.

Tài liệu tham khảo

1. Hou, Z., Lee, C. K. M., Lv, Y., & Keung, K. L. (2023). Fault detection and diagnosis of air brake system: A systematic review. *Journal of Manufacturing Systems, 71*, 34–58.
2. Kuravsky, L. S., & Nguyễn, Q. T. (2016). *Mô hình Markov ứng dụng trong chẩn đoán và dự báo*. Hà Nội, Việt Nam: Nhà xuất bản Khoa học và Kỹ thuật.
3. Zeng, S.-K., Zhao, Y.-D., Zhang, J.-G., Kang, R., & Shi, J.-Y. (2001). *Analysis tutorial of system reliability design*. Beijing, China: Beihang University Press.
4. Tashnizi, M. M. (2023). *Application of hidden Markov model in production data*.

Application of Markov models to study the technical condition transitions of the braking system on KAMAZ-53212 trucks

Abstract: *The pneumatic brake system of heavy-duty trucks has a multi-circuit structure and is strongly affected by operating conditions, leading to complex variations in its technical state over time. This paper proposes a method for diagnosing the technical state of the pneumatic brake system of the KAMAZ-53212 truck based on a Gaussian Hidden Markov Model (GHMM). Based on the analysis of the structure and operating principles of the pneumatic brake system, hidden states representing typical operating conditions and fault modes are established. The observation vector is constructed from directly measurable parameters such as supply pressure, brake chamber pressures of the front axle and the middle–rear axle group, as well as parameters characterizing the air charging and leakage processes. Simulation results show that the GHMM can reasonably infer the evolution of the brake system's technical states over time, capturing localized faults and alternating fault characteristics under real operating conditions.*

Keywords: Pneumatic brake system; Gaussian Hidden Markov Model; technical state diagnosis; KAMAZ-53212.

Nghiên cứu các yếu tố ảnh hưởng đến độ ổn định trong quá trình quay vòng ô tô 2 cầu.

Tăng Minh Tài[1], Nguyễn Đăng Quý[1]

[1]Học viện Kỹ thuật quân sự

Email: minhtai.tdnu@gmail.com; Tel: 0979710643

Tóm tắt

Bài báo khảo sát một số yếu tố ảnh hưởng tới quá trình quay vòng của xe ô tô hai cầu. Mô hình để khảo sát ô tô 2 cầu được xây dựng làm cơ sở cho việc thiết lập hệ phương trình vi phân mô tả chuyển động của ô tô. Phương pháp mô phỏng số được thực hiện trên phần mềm Matlab - Simulink để khảo sát ảnh hưởng của các điều kiện chuyển động với thông số vận tốc xe và góc đánh lái. Kết quả khảo sát cho phép đánh giá khả năng quay vòng của xe theo lý thuyết, từ đó đưa ra khuyến nghị để tránh các trường hợp quay vòng thiếu và quay vòng thừa. Làm cơ sở xác định các giải pháp nâng cao an toàn chuyển động cho ô tô trong điều kiện di chuyển trên địa hình gồ ghề và điều kiện thời tiết khắc nghiệt. Ngoài ra, các kết quả nghiên cứu còn có thể làm cơ sở để hoàn thiện các hệ thống cảnh báo hoặc thiết kế hệ thống tự động điều khiển, hỗ trợ cho quá trình điều khiển ô tô cho lái xe, nâng cao độ tin cậy đảm bảo an toàn chuyển động của xe.

Từ khóa: *Động lực học chuyển động, ổn định quay vòng, quỹ đạo chuyển động, ô tô 2 cầu, mô phỏng Matlab - Simulink.*

1. Đặt vấn đề

Tại Việt Nam, Toyota Fortuner 2022 là dòng xe SUV (thể thao đa dụng) được sử dụng rất phổ biến nhờ ưu điểm về kết cấu khung gầm vững chắc và khả năng thích nghi tốt với điều kiện hạ tầng giao thông đa dạng. Tuy nhiên, với đặc thù kết cấu có trọng tâm cao, dòng xe này thường đối mặt với các vấn đề về ổn định động lực học khi thực hiện quay vòng ở tốc độ cao hoặc chuyển hướng đột ngột, tiềm ẩn nguy cơ mất an toàn như văng trượt hoặc lật xe.

Do đó, việc nghiên cứu chuyên sâu về động lực học quá trình quay vòng của xe Toyota Fortuner 2022 là yêu cầu cấp thiết nhằm nâng cao an toàn giao thông. Bài báo tập trung xây dựng mô hình mô phỏng và khảo sát các đặc tính quay vòng của xe trong các điều kiện vận hành khác nhau. Kết quả nghiên cứu sẽ cung cấp cơ sở dữ liệu khoa học quan trọng để đánh giá giới hạn ổn định, từ đó đề xuất các khuyến cáo vận hành an toàn và các giải pháp kỹ thuật nhằm tối ưu hóa khả năng kiểm soát phương tiện.

2. Thiết lập mô hình khảo sát động lực học quá trình quay vòng ô tô 2 cầu

2.1. Giả thiết

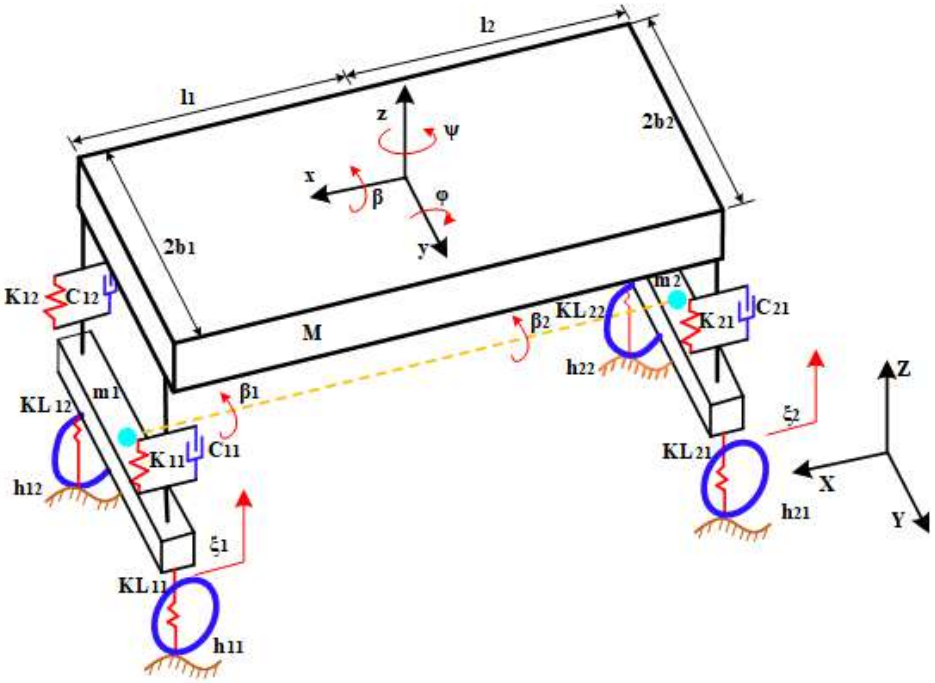
Tác giả đưa ra các giả thiết nhằm đơn giản hóa bài toán nhưng vẫn đánh giá được các thông số dao động cần thiết như sau:

Mô hình xe có 2 trục là đối xứng theo trục dọc của xe và bỏ qua ảnh hưởng của độ lệch trọng tâm. Đối với các bài toán xe ô tô chuyển động trong mặt phẳng đường, các trạng thái chuyển động với góc lắc ngang nhỏ, nhóm tác giả đề xuất giả thiết khung xe ô tô Toyota Fortuner 2022 cứng tuyệt đối để đơn giản hóa trong quá trình xây dựng mô hình. Coi thân xe ô tô như một tấm phẳng có khối lượng M đặt tại trọng tâm xe. Trường hợp tổng quát thì thân xe có 6 chuyển động (6 bậc tự do) bao gồm: 3 chuyển động tịnh tiến theo các trục x (trục dọc), y (trục ngang), z (trục thẳng đứng) và 3 góc xoay tương ứng: φ (góc lắc dọc - quay quanh trục y), β (góc lắc ngang - quay quanh trục x), ψ (góc quay thân xe - quay quanh trục z). Thân xe được nối với các cầu xe thông qua hệ thống treo đặc trưng bởi các thông số: hệ số độ cứng K_{ji} và hệ

số cản giảm chấn C_{ji} (chỉ số j: 1 - trước; 2 - sau; chỉ số i: 1 - trái; 2 - phải). Với các cầu xe (cầu trước j=1, cầu sau j=2) là các khối lượng không được treo được coi như các thanh phẳng, chúng có khối lượng m_{T1}, m_{T2} đặt tại trọng tâm của các thanh phẳng. Ở mỗi cầu xe ta thấy có 01 chuyển động tịnh tiến theo trục thẳng đứng (ξ_{T1}, ξ_{T2}) và 01 chuyển vị góc lắc ngang quanh trục dọc (β_{T1}, β_{T2}). Bỏ qua chuyển động tịnh tiến theo trục dọc và trục ngang; góc xoay của cầu xe ô tô theo trục thẳng đứng và trục ngang. Cầu xe liên kết với mặt đường bằng các bánh xe có đàn hồi, đặc trưng bởi hệ số độ cứng KL_{ji}. Bỏ qua ảnh hưởng của độ rơ vành tay lái và tốc độ đánh lái. Coi đường là bằng phẳng và cứng tuyệt đối với chiều cao mấp mô mặt đường bằng 0. Bỏ qua lực cản khí động theo phương ngang của xe ô tô (F_{wy}). Chuyển động theo phương dọc của cầu xe ô tô tương đương với chuyển động dọc của khối lượng được treo tương ứng [1].

2.2. Mô hình khảo sát

Từ các giả thiết ở trên, tác giả đưa ra mô hình tương đương khảo sát khảo sát động lực học quá trình quay vòng ô tô 2 cầu [2] như sau:



Hình 1. Mô hình không gian ô tô Toyota Fortuner 2022

Trong đó:

M, m_1, m_2: Khối lượng phần treo, phần không treo phía trước, phía sau [kg];

ξ_{11}, ξ_{12},: Dịch chuyển thẳng đứng của khối lượng không treo trước bên trái, phải [m];

ξ_{21}, ξ_{22}: Dịch chuyển thẳng đứng của khối lượng không treo sau bên trái, bên phải [m];

β_1, β_2: Góc lắc ngang quanh trục X của cầu trước và cầu sau [rad];

β: Góc lắc ngang quanh trục x của thân xe [rad];

ψ: Góc quay thân xe quanh trục z [rad];

KL_1, KL_2: Độ cứng của lốp trước và các lốp sau [N/m];

K_{11}, K_{12}: Độ cứng của treo trước bên trái, phải [N/m];

K_{21}, K_{22}: Độ cứng của sau trước bên trái, phải [N/m];

C_{11}, C_{12}: Hệ số giảm chấn của treo trước bên trái, phải [Ns/m];

C_{21}, C_{22}: Hệ số giảm chấn của treo sau bên trái, phải [Ns/m];

h_{11}, h_{12}, h_{21}, h_{22}: Biên dạng mặt đường tương ứng với từng bánh xe [m]

b_1, b_2: Nửa vệt bánh xe phía trước, sau [m];

2.3. Hệ phương trình vi phân mô tả chuyển động của ô tô

Dựa trên mô hình khảo sát động lực học quá trình quay vòng ô tô 2 cầu đã phân tích, tác giả nhận thấy hệ có 14 bậc tự do. Nhằm tối ưu hóa quá trình tính toán, tác giả lựa chọn phương pháp áp dụng hệ phương trình Newton Euler [3]. Sau quá trình tính toán, phân tích, tác giả đã xây dựng hệ phương trình vi phân bao gồm 14 phương trình sau:

$$M(\ddot{x} - \dot{y}\dot{\psi}) = F_{x11}\cos\delta_{11} + F_{x12}\cos\delta_{12} - F_{y11}\sin\delta_{11} - F_{y12}\sin\delta_{12} + F_{x12} + F_{x22} - F_{wx} \quad (1)$$

$$M(\ddot{y} - \dot{x}\dot{\psi}) = F_{x11}\sin\delta_{11} + F_{x12}\sin\delta_{12} + F_{y11}\cos\delta_{11} + F_{y12}\cos\delta_{12} + F_{y21} + F_{y22} + F_{wy} \quad (2)$$

$$\begin{aligned} J_z\ddot{\psi} = {} & [(F_{x12}\cos\delta_{12} - F_{y12}\sin\delta_{12}) - (F_{x11}\cos\delta_{11} - F_{y11}\sin\delta_{11})]b_1 + (F_{x22} - F_{x21})b_2 \\ & + [(F_{x12}\sin\delta_{12} + F_{y12}\cos\delta_{12}) + (F_{x11}\sin\delta_{11} + F_{y11}\cos\delta_{11})]l_1 \\ & - (F_{y21} + F_{y22})l_2 + F_{wy}l_{wy} \end{aligned} \quad (3)$$

$$m\ddot{z} = F_{C11} + F_{C12} + F_{C21} + F_{C22} + F_{K11} + F_{K12} + F_{K21} + F_{K22} \quad (4)$$

$$\begin{aligned} J_y\ddot{\varphi} = {} & -(F_{C11} + F_{C12} + F_{K11} + F_{K12})l_1 + (F_{C21} + F_{C22} + F_{K21} + F_{K22})l_2 \\ & - (F_{x11}\cos\delta_{11} + F_{x12}\cos\delta_{12} - F_{y11}\sin\delta_{11} - F_{y12}\sin\delta_{12})(h_g - r_1) \\ & - (h_g - r_2)(F_{x21} + F_{x22}) - (M_{11} + M_{12} + M_{21} + M_{22}) - F_{wx}(h_w - h_g) \end{aligned} \quad (5)$$

$$\begin{aligned} J_x\ddot{\beta} = {} & (F_{C11} - F_{C12} + F_{K11} - F_{K12})w_1 + (F_{C21} - F_{C22} + F_{K21} - F_{K22})w_2 - M_{wx} \\ & - F_{wy}(h_w - h_g) + F_{R1}(h_g - h_{R1}) + F_{R2}(h_g - h_{R2}) - M_{X1} - M_{X2} \end{aligned} \quad (6)$$

$$m_{T1}\ddot{\xi}_{T1} = (F_{CL11} + F_{CL12}) - (F_{C11} + F_{C12} + F_{K11} + F_{K12}) \quad (7)$$

$$\begin{aligned} J_{Tx1}\ddot{\beta}_{T1} = {} & (F_{C12} - F_{C11} + F_{K12} - F_{K11})w_1 + (F_{KL11} - F_{KL12})b_1 + M_{X1} \\ & + (F_{x11}\sin\delta_{11} + F_{x12}\cos\delta_{12} + F_{y11}\cos\delta_{11} + F_{y12}\cos\delta_{12})r_1 - F_{R1}(h_{R1} - r_1) \end{aligned} \quad (8)$$

$$m_{T2}\ddot{\xi}_{T2} = (F_{KL21} + F_{KL22}) - (F_{C21} + F_{C22} + F_{K21} + F_{K22}) \quad (9)$$

$$\begin{aligned} J_{Tx2}\ddot{\beta}_{T2} = {} & (F_{C22} - F_{C21} + F_{K22} - F_{K21})w_2 + (F_{KL21} - F_{KL21})b_2 - F_{R2}(h_{R2} - r_2) \\ & + (F_{y21} + F_{y22})r_2 + M_{X2} \end{aligned} \quad (10)$$

$$J_{Ty11}\ddot{\varphi}_{11} = M_{11} - (F_{x11} + fF_{z11})r_{11} \quad (11)$$

$$J_{Ty12}\ddot{\varphi}_{12} = M_{12} - (F_{x12} + fF_{z12})r_{12} \quad (12)$$

$$J_{Ty21}\ddot{\varphi}_{21} = M_{21} - (F_{x21} + fF_{z21})r_{21} \quad (13)$$

$$J_{Ty22}\ddot{\varphi}_{22} = M_{22} - (F_{x22} + fF_{z22})r_{22} \quad (14)$$

3. Kết quả nghiên cứu

3.1. Thông số đầu vào

Nhóm tác giả sử dụng phần mềm Matlab - Simulink [4] để giải hệ 14 phương trình vi phân khảo sát độ ổn định khi quay vòng được đề cập ở mục 2.3 và sử dụng các thông số của xe Toyota Fortuner 2022 như sau:

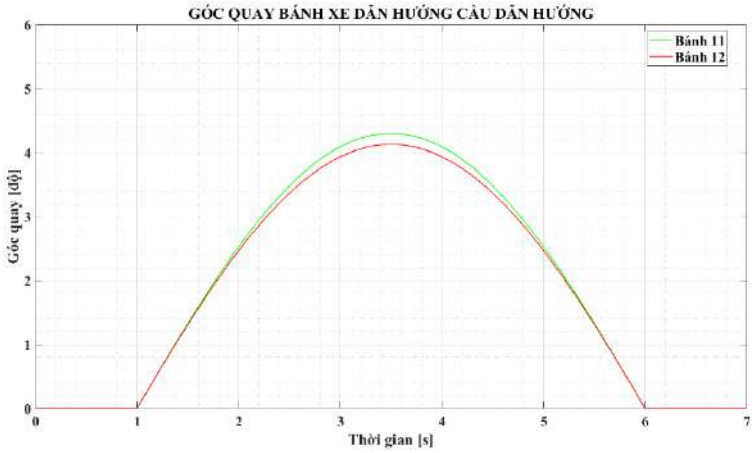
Bảng 1. Các thông số đầu vào khảo sát độ ổn định quay vòng xe Toyota Fortuner 2022

TT	Thông số kỹ thuật	Kí hiệu	Số liệu	Đơn vị
1	Hệ số cản lăn	f	0.015	
2	Bán kính bánh xe	r	0.3875	m
3	Biên độ lái	δ_L	4,3	độ
4	Chiều dài cơ sở	L	2.745	m
5	Nữa vệt bánh xe phía trước	b_1	0.77	m
6	Nữa vệt bánh xe phía sau	b_2	0.775	m
7	Khối lượng toàn bộ khi không tải	M	1995	kg
8	Khoảng cách giữa 2 hệ treo trước	w_1	1.0	m
9	Khoảng cách giữa 2 hệ treo sau	w_2	1.1	m
10	Chiều rộng xe	B	1.855	m
11	Chiều cao xe	H	1.835	m
12	Khối lượng cầu trước	m_{T1}	125	kg
13	Khối lượng cầu sau	m_{T2}	270	kg
14	Khoảng cách từ trọng tâm xe đến cầu trước	l_1	1.26	m
15	Khoảng cách từ trọng tâm xe đến cầu sau	l_2	1.485	m
16	Chiều cao trọng tâm	h_g	0.72	m
17	Công suất lớn nhất/số vòng quay	P_{Max}	110/3400	Kw/rpm
18	Momen lớn nhất/số vòng quay	M_{Max}	245/1800	N.m/rpm
19	Mô men quán tính của khối lượng không được treo cầu trước	J_{Tx1}	72	$Kg.m^2$
20	Mô men quán tính của khối lượng không được treo cầu trước	J_{Tx2}	92	$Kg.m^2$
21	Mô men quán tính quay của cụm bánh xe cầu trước	J_{Ty11},J_{Ty12}	3.5	$Kg.m^2$
22	Mô men quán tính quay của cụm bánh xe cầu sau	J_{Ty21},J_{Ty22}	4.15	$Kg.m^2$
23	Hệ số cản giảm chấn hệ treo trước	C_{11}=C_{12}	40000	N/(m/s)

TT	Thông số kỹ thuật	Kí hiệu	Số liệu	Đơn vị
24	Hệ số cản giảm chấn hệ treo sau	$C_{21}=C_{22}$	45000	N/(m/s)
25	Hệ số độ cứng hệ thống treo trước	$K_{11}=K_{12}$	2900	N/m
26	Hệ số độ cứng hệ thống treo sau	$K_{21}=K_{22}$	3400	N/m

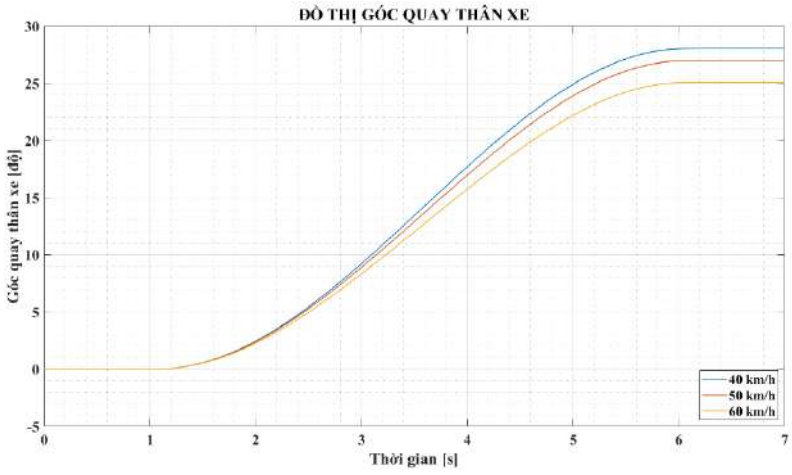
3.2. Kết quả mô phỏng động lực học

Khảo sát quỹ đạo quay vòng trái khi thay đổi góc đánh lái theo thời gian t. Tiến hành khảo sát trong khoảng thời gian đánh lái từ 1s đến 6s (Δt=5s) với biên độ góc đánh lái là 4,3 độ, tần số f=0,2 Hz ở các mốc vận tốc [40,50,60] km/h. Kết quả mô phỏng động lực học được thể hiện qua các đồ thị sau:



Hình 2. Đồ thị góc quay bánh xe dẫn hướng

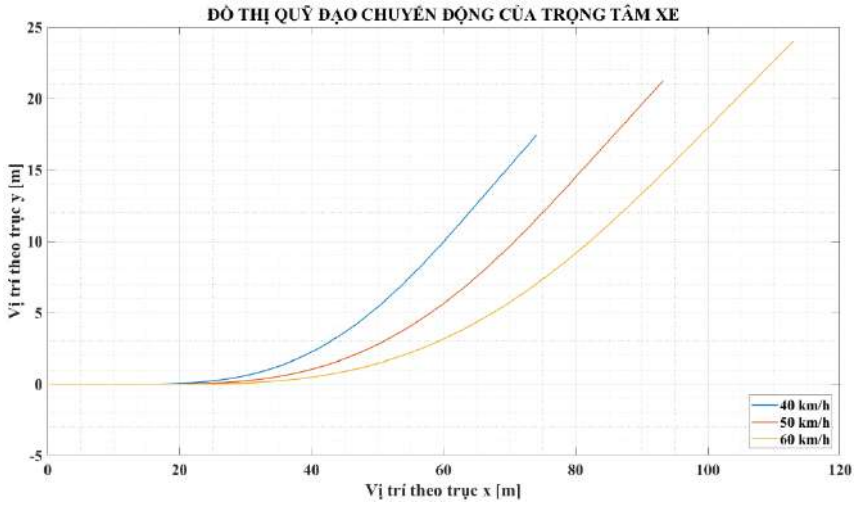
Góc quay hai bánh xe dẫn hướng biến thiên, tăng dần đến cực đại rồi giảm về 0, phản ánh quá trình đánh lái hợp lý và liên tục. Bánh trong (Bánh 11) luôn có góc quay lớn hơn bánh ngoài (Bánh 12), với giá trị cực đại khoảng 4,3° so với 4,1°, phù hợp nguyên lý động học quay vòng. Dạng tín hiệu không có gãy khúc hay dao động đột ngột cho thấy đầu vào điều khiển ổn định, góp phần duy trì trạng thái quay vòng êm dịu và đảm bảo độ ổn định động lực học của xe. Thể hiện tín hiệu đầu vào hợp lý với bài toán mô phỏng quỹ đạo quay vòng của xe.



Hình 3. Đồ thị góc quay thân xe

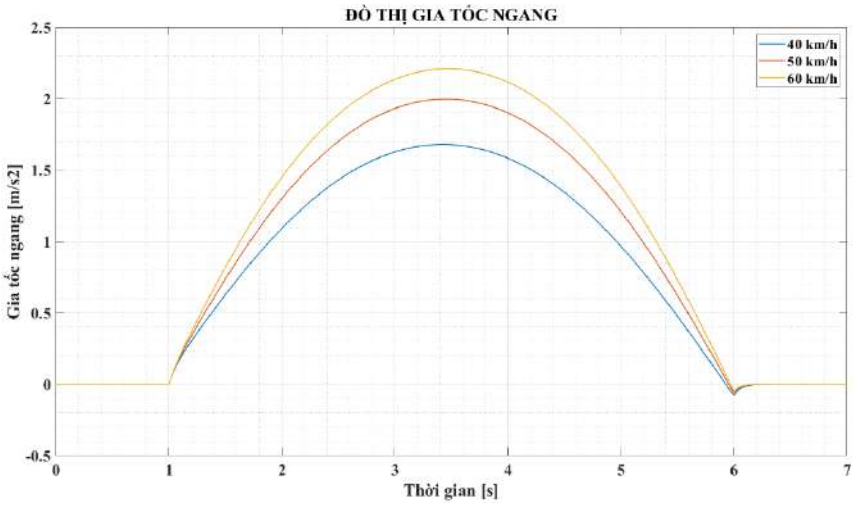
Đồ thị góc quay thân xe cho thấy trong khoảng thời gian từ 1s đến 6s, Toyota Fortuner 2022 quay vòng với đặc tính đáp ứng trơn tru, liên tục, không xuất hiện dao động hay gãy khúc, chứng tỏ quỹ đạo chuyển động ổn định và dễ kiểm soát. Khi vận tốc tăng, góc quay thân xe

giảm rõ rệt: Tại 40 km/h đạt khoảng 29,5°, trong khi tại 60 km/h chỉ còn khoảng 27° thể hiện mối quan hệ nghịch giữa vận tốc và góc quay. Đặc điểm này phản ánh xu hướng thiếu lái nhẹ, có lợi cho ổn định quay vòng ở tốc độ cao, giúp hạn chế nguy cơ mất ổn định hướng và nâng cao độ an toàn khi xe vận hành.



Hình 4. Đồ thị quỹ đạo chuyển động của trọng tâm xe

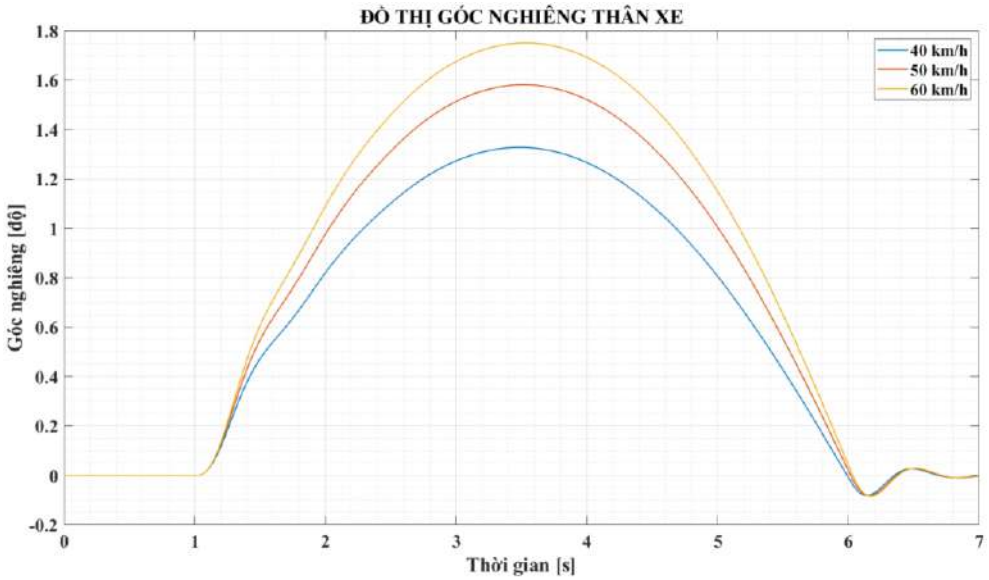
Đồ thị quỹ đạo chuyển động của trọng tâm xe cho thấy Toyota Fortuner 2022 duy trì quỹ đạo cong ổn định và liên tục khi chuyển hướng trong dải vận tốc từ 40 km/h đến 60 km/h. Khi vận tốc tăng, quỹ đạo có xu hướng mở rộng, thể hiện bán kính quay vòng lớn hơn, phù hợp với đặc tính thiếu lái đã phân tích trước đó. Các đường quỹ đạo trơn, không xuất hiện lệch đột ngột hay mất kiểm soát, cho thấy xe giữ được ổn định hướng tốt và khả năng điều khiển an toàn trong quá trình quay vòng.



Hình 5. Đồ thị gia tốc ngang

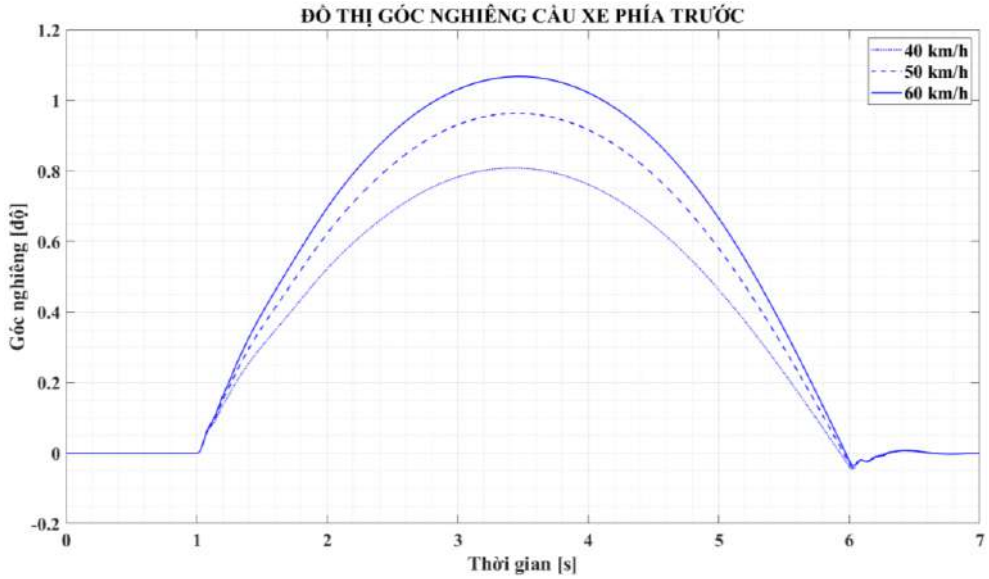
Gia tốc ngang tăng dần, đạt cực đại rồi giảm về 0 với dạng biến thiên, không xuất hiện đỉnh nhọn hay dao động bất thường. Giá trị gia tốc ngang cực đại tăng theo vận tốc, đạt khoảng 2,2 m/s² (≈0,224g) tại 60 km/h, phù hợp quy luật tỷ lệ với bình phương vận tốc và nằm thấp hơn nhiều so với giới hạn bám của lốp (0,7-0,8g) [5] cũng như ngưỡng mất ổn định lật [6]. Điều này cho thấy xe làm việc trong vùng an toàn, không có dấu hiệu trượt ngang hay mất kiểm soát. Dao động nhỏ sau thời điểm khoảng 6s phản ánh ảnh hưởng quán tính và đặc tính đàn hồi giảm

chấn của hệ thống treo, nhưng nhanh chóng tắt dần, khẳng định độ ổn định động lực học tốt của xe khi quay vòng.



Hình 6. Đồ thị góc nghiêng thân xe

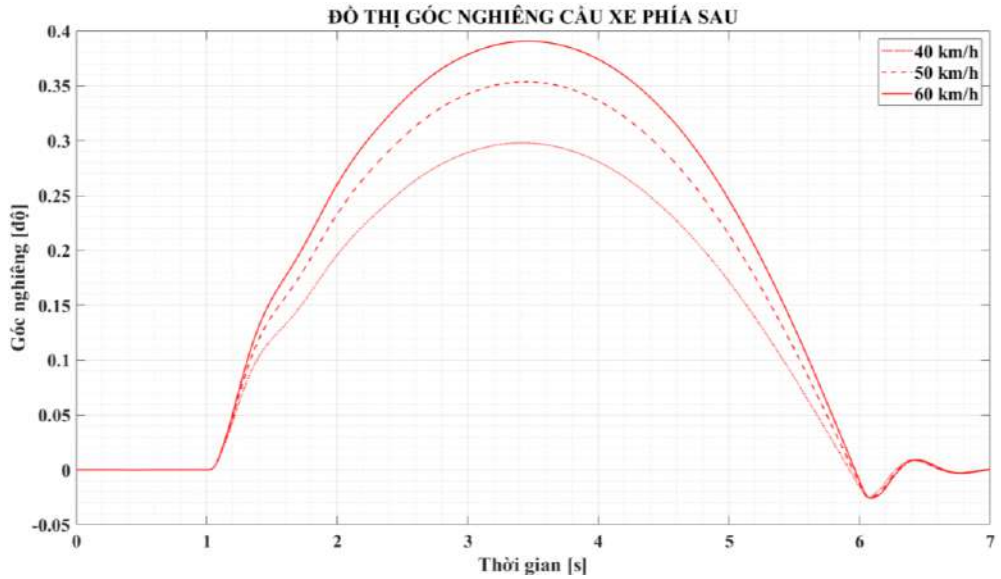
Góc nghiêng biến thiên tỷ lệ thuận với vận tốc và gia tốc ngang. Tại 40 km/h, góc nghiêng lớn nhất đạt 1,35° và tại 60 km/h, giá trị này là 1,75°. Với Toyota Fortuner 2022, góc nghiêng thân xe 1,75° ở gia tốc ngang 0,224g là hợp lý và kiểm soát tốt. Đặc biệt, sau khi kết thúc đánh lái (t > 6s), dao động tắt dần nhanh chóng về vị trí cân bằng, chứng tỏ hệ thống giảm chấn hoạt động hiệu quả, giúp xe ổn định nhanh sau khi chuyển làn.



Hình 7. Đồ thị góc nghiêng cầu xe phía trước

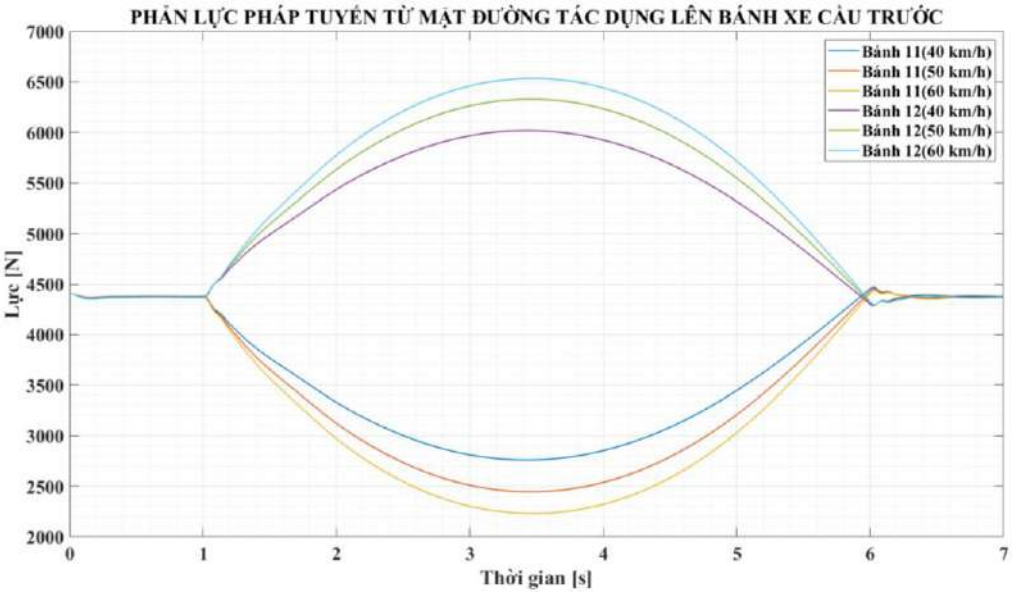
Ở vận tốc 40 km/h, góc nghiêng cầu xe phía trước đạt giá trị lớn nhất khoảng 0,8° và giảm dần sau 6 giây. Khi vận tốc tăng lên 60 km/h, góc nghiêng đạt giá trị lớn nhất khoảng 1,1° cho thấy góc nghiêng tỷ lệ thuận với vận tốc và gia tốc ngang. Điều này chứng tỏ sự ổn định và kiểm soát tốt của xe trong quá trình quay vòng. Đặc biệt, sau khi kết thúc đánh lái (t > 6s), dao

động tắt dần về vị trí cân bằng, cho thấy hệ thống giảm chấn của xe hoạt động hiệu quả, giúp xe nhanh chóng ổn định lại sau khi chuyển làn mà không gây cảm giác mất kiểm soát.



Hình 8. Đồ thị góc nghiêng cầu xe phía sau

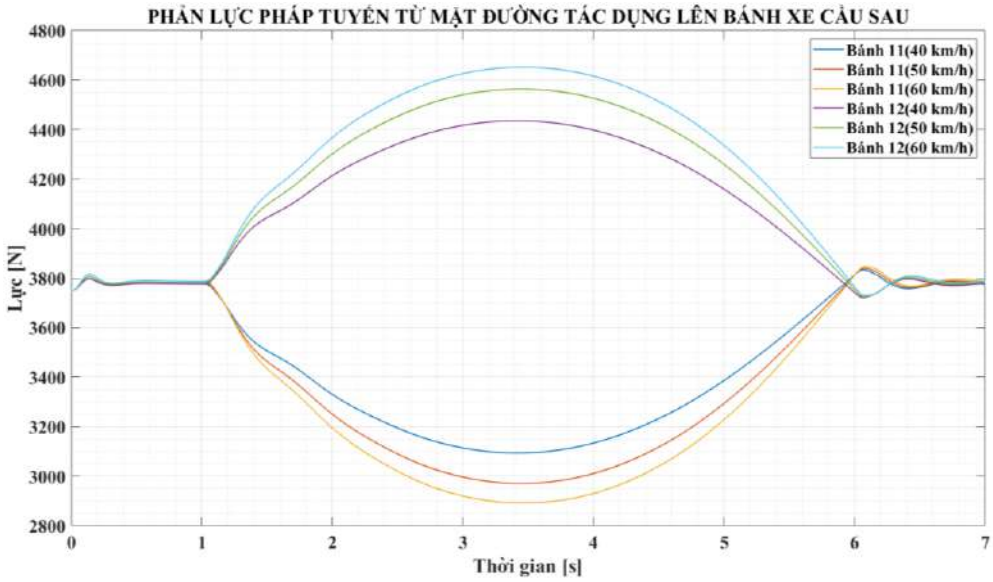
Tương tự như đồ thị góc nghiêng cầu xe phía trước thì đồ thị góc nghiêng cầu xe phía sau cũng chứng tỏ được sự ổn định và kiểm soát tốt của xe trong quá trình quay vòng. Cụ thể ở vận tốc 40 km/h góc nghiêng khoảng 0,3° và ở 60 km/h sấp xỉ 0,4°. Tải trọng tác dụng cầu trước lớn hơn cầu sau và đang quay vòng trái nên giá trị góc nghiêng cầu trước lớn hơn cầu sau là hoàn toàn hợp lý.



Hình 9. Đồ thị phản lực pháp tuyến từ mặt đường lên bánh xe cầu trước

Phản lực pháp tuyến biến thiên rõ rệt theo vận tốc và gia tốc ngang, đạt giá trị lớn nhất khoảng 6550N tại vận tốc 60 km/h. Lực tác dụng lên bánh ngoài (Bánh 12) luôn lớn hơn bánh trong (Bánh 11), phản ánh hiện tượng chuyển tải trọng ngang đặc trưng khi xe quay vòng, mức chênh lệch lực tăng lên khi vận tốc tăng từ 40 km/h lên 60 km/h. Sau khi kết thúc quá trình đánh lái ($t > 6s$), phản lực pháp tuyến trở về giá trị cân bằng và dao động nhỏ tắt dần, chứng tỏ

hệ thống treo làm việc hiệu quả, đảm bảo khả năng bám đường và góp phần duy trì ổn định quay vòng xe.



Hình 10. Đồ thị phản lực pháp tuyến từ mặt đường lên bánh xe cầu sau

Tương tự như cầu trước, đồ thị phản lực pháp tuyến tác dụng lên các bánh xe cầu sau cho thấy trong quá trình quay vòng, lực pháp tuyến tăng theo vận tốc và đạt giá trị cực đại khoảng 4650 N tại 60 km/h. Lực tác dụng lên bánh ngoài (Bánh 22) luôn lớn hơn bánh trong (Bánh 21). Sau khi kết thúc đánh lái (t > 6s), phản lực pháp tuyến nhanh chóng trở về giá trị cân bằng và dao động tắt dần, chứng tỏ hệ thống treo cầu sau làm việc hiệu quả, góp phần duy trì độ ổn định và bám đường của xe trong suốt quá trình quay vòng.

Nhận xét:

Kết quả mô phỏng động lực học quá trình quay vòng của xe Toyota Fortuner 2022 tại các vận tốc khảo sát [40,50,60] km/h thể hiện một cách nhất quán và có mối liên hệ chặt chẽ giữa các đại lượng động lực học. Tín hiệu góc quay bánh xe dẫn hướng biến thiên trơn và phù hợp quan hệ hình học quay vòng, tạo điều kiện cho xe vào cua êm dịu và là nguyên nhân trực tiếp hình thành quỹ đạo cong ổn định của trọng tâm xe. Từ đó, gia tốc ngang tăng theo vận tốc nhưng vẫn nằm trong vùng an toàn, không vượt giới hạn bám của lốp, đồng thời dẫn đến góc nghiêng thân xe và góc nghiêng các cầu tăng ở mức hợp lý, phản ánh sự cân bằng giữa mô men lật và khả năng chống lật của hệ thống treo. Sự gia tăng gia tốc ngang kéo theo hiện tượng chuyển tải trọng ngang, được thể hiện rõ qua sự phân bố phản lực pháp tuyến giữa bánh trong và bánh ngoài ở cả cầu trước và cầu sau, trong đó cầu trước chịu tải lớn hơn, phù hợp với đặc tính thiếu lái nhẹ của xe. Quan trọng hơn, sau khi kết thúc quá trình đánh lái, các đại lượng như gia tốc ngang, góc nghiêng thân xe và phản lực pháp tuyến đều nhanh chóng trở về trạng thái cân bằng với dao động nhỏ và tắt dần nhanh, chứng tỏ hệ thống treo và giảm chấn làm việc hiệu quả. Sự liên kết chặt chẽ giữa điều khiển lái, quỹ đạo chuyển động, gia tốc ngang, góc nghiêng và phân bố tải trọng đã bổ trợ lẫn nhau để khẳng định rằng Toyota Fortuner 2022 duy trì được độ ổn định động lực học tốt khi quay vòng trong toàn bộ dải vận tốc khảo sát. Hơn nữa, độ chênh lệch tải trọng động ở cầu trước lớn hơn đáng kể so với cầu sau cho thấy mô men chống lật phân bố tập trung nhiều hơn ở cầu trước, điều này theo lý thuyết động lực học sẽ tạo ra xu hướng quay vòng thiếu lái đặc trưng; đây là đặc tính ổn định hướng quan trọng trên các dòng

SUV giúp xe tự động giảm bán kính quay vòng khi vận tốc giảm và dễ dàng được người lái kiểm soát, góp phần nâng cao độ an toàn thụ động và giảm rủi ro mất ổn định do thao tác lái đột ngột. Đồng thời khẳng định thiết kế khung gầm và hệ thống treo của Toyota Fortuner 2022 đáp ứng tốt các chỉ tiêu về an toàn chuyển động trong điều kiện vận hành thông thường.

4. Kết luận

Kết luận rút ra từ nghiên cứu như sau:

- Xây dựng thành công mô hình động lực học không gian và thiết lập hệ phương trình vi phân mô tả chuyển động quay vòng của xe Toyota Fortuner 2022 (bao gồm các chuyển động tịnh tiến ngang, quay quanh trục thẳng đứng và lắc ngang thân xe) nhằm khảo sát độ ổn định khi xe quay vòng. Cụ thể, mô hình được xây dựng dưới dạng các mô đun trong Simulink, bao gồm khối thân xe, khối các bánh xe và khối các cầu xe. Các khối này tương ứng với các thành phần cơ bản của phương tiện, từ đó mô phỏng các đặc tính động học và cơ học trong quá trình quay vòng. Việc xây dựng mô hình chi tiết và kết hợp đồ thị xuất ra được giúp phân tích được độ ổn định của ô tô, cũng như đánh giá khả năng vận hành của phương tiện trong các điều kiện quay vòng thực tế.

- Mô phỏng quá trình quay vòng của xe trên môi trường Matlab - Simulink tại các dải vận tốc khác nhau, từ đó xuất ra các kết quả dạng đồ thị về quỹ đạo chuyển động của trọng tâm xe, gia tốc ngang, góc lắc và phản lực pháp tuyến... Kết quả nghiên cứu làm cơ sở khoa học để xác định ngưỡng vận tốc và góc đánh lái an toàn, giúp người lái kiểm soát tốt hơn đặc tính thiếu lái đặc trưng của dòng xe gầm cao. Đồng thời, bộ số liệu này giải quyết được vấn đề xác định giới hạn ổn định ngang và chống lật, tạo tiền đề cho việc nghiên cứu, tối ưu hóa hệ thống treo hoặc phát triển các thuật toán điều khiển ổn định điện tử cho các dòng xe SUV sau này.

Tài liệu tham khảo

[1] N. T. Dũng, V. V. Hường, Đ. H. Phúc, and D. N. Khánh, *Động lực học ô tô*. NXB Giáo dục Việt Nam, 2014.

[2] Gillespie TD. *Fundamentals of Vehicle Dynamics*. SAE International; 1992.

[3] Nguyễn Văn Khang, *Động lực học hệ nhiều vật*, NXB Khoa học và Kỹ thuật, 2007.

[4] P. T. Tạo, *Giáo trình Matlab*. Đà Nẵng: Đại học Đà Nẵng, 2006.

[5] Wong. J. Y, *Theory of Ground Vehicles*. Wiley, 1993.

[6] Farmer. C. M., Lund. A. K, *Rollover risk of cars and light trucks after accounting for driver and environmental factors.* Accident Analysis & Prevention, 2002.

Research on Factors Affecting the Cornering Stability of Two-Axle Vehicles

Abstract: *This paper investigates several factors affecting the cornering performance of two-axle vehicles. A two-wheel model is used to analyze the vehicle's cornering behavior, forming the basis for establishing a system of differential equations describing the vehicle's motion. Numerical simulations are performed using Matlab-Simulink software to study the influence of motion conditions on vehicle speed and steering angle parameters. The results allow for the evaluation of the vehicle's cornering performance based on theoretical calculations, leading to recommendations to avoid understeer and oversteer. This research provides a foundation for identifying solutions to improve vehicle safety on rough terrain and in harsh weather conditions. Furthermore, the research findings can be used as a basis for enhancing warning systems or designing automated driver-assistance control systems, thereby increasing reliability and ensuring vehicle safety.*

Keywords: *Motion dynamics, rotational stability, motion trajectory, two-axle vehicles, simulation using Matlab-Simulink,.*

Nghiên cứu thiết kế cơ cấu tăng ma sát trong dạng đĩa ma sát cho bộ vi sai cầu xe chủ động

Trần Đức Hùng[1], Nguyễn Trường Sinh[1]

[1]Học viện Kỹ thuật quân sự

Email: hungtran93a3@gmail.com; Tel: 0393738969

Tóm tắt

Bài báo trình bày một số kết quả trong nghiên cứu thiết kế, tính toán cơ cấu tăng ma sát trong dạng đĩa ma sát cho bộ vi sai cầu xe chủ động dạng bánh răng côn đối xứng, nhằm phân phối lại mô men và hạn chế trượt của bánh xe trên một số địa hình có độ bám thấp. Bài báo đã tiến hành phân tích và thiết lập mối quan hệ giữa mô men xoắn truyền đến các bán trục, tốc độ quay của hai bánh xe và lực ép tác dụng lên đĩa ma sát của cơ cấu tăng ma sát trong dạng đĩa, làm cơ sở cho việc tính toán, thiết kế. Mô phỏng số được thực hiện trên phần mềm Matlab/Simulink để khảo sát, đánh giá ảnh hưởng của cơ cấu tăng ma sát trong dạng đĩa ma sát đã thiết kế. Kết quả nghiên cứu cho thấy cơ cấu tăng ma sát trong dạng đĩa ma sát đã cải thiện đáng kể khả năng phân phối mô men xoắn giữa hai bán trục, hạn chế trượt của bánh xe, góp phần nâng cao tính năng thông qua của xe.

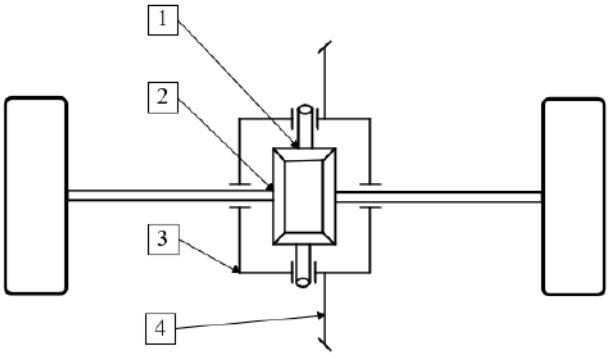
Từ khóa: *Vi sai hạn chế trượt;cơ cấu tăng ma sát trong; đĩa ma sát; hệ số khóa; mô phỏng động lực học.*

1. Đặt vấn đề

Bộ vi sai bánh răng côn đối xứng được sử dụng phổ biến làm vi sai giữa các bánh xe ô tô nhờ kết cấu đơn giản và độ tin cậy cao. Tuy nhiên, trong điều kiện mặt đường có độ bám không đồng đều, mô men xoắn truyền tới cầu xe chủ động bị giới hạn bởi bánh xe có độ bám thấp hơn, bánh xe có độ bám thấp trượt quay và làm giảm khả năng phát huy lực kéo. Đối với ô tô quân sự, thường xuyên phải hoạt động trong các địa hình và các điều kiện môi trường không thuận lợi, yêu cầu về khả năng vượt địa hình và chở tải nặng là rất cao, do đó việc nghiên cứu áp dụng bộ vi sai hạn chế trượt có sử dụng cơ cấu tăng ma sát trong là rất cần thiết để góp phần nâng cao tính năng thông qua của xe. Với cơ cấu tăng ma sát trong dạng đĩa ma sát có cấu tạo đơn giản dễ tích hợp cho bộ vi sai bánh răng côn đối xứng của hầu hết các cầu xe của ô tô quân sự.

2. Cơ sở thiết kế cơ cấu tăng ma sát trong dạng đĩa ma sát

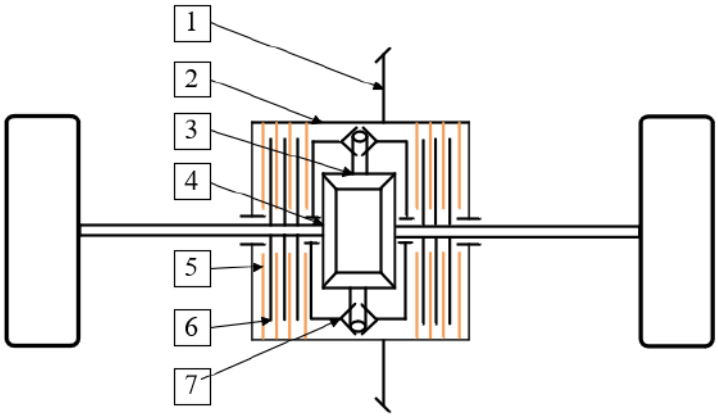
Cấu tạo của bộ vi sai bánh răng côn đối xứng thông thường có sơ đồ như sau:



Hình 1. Sơ đồ bộ vi sai bánh răng côn đối xứng

(1) Bánh răng hành tinh; (2) Bánh răng bán trục; (3) Vỏ vi sai; (4) Bánh răng vành chậu

Bộ vi sai bánh răng côn đối xứng hoạt động ổn định trên đường có hệ số bám tốt. Tuy nhiên trong điều kiện đường không tốt hệ số bám kém, giá trị mô men xoắn lại bị giới hạn bởi lực kéo tối đa có thể đạt được tương ứng với hệ số bám. Vì vậy để nâng cao hiệu quả lực kéo của xe ta sử dụng vi sai hạn chế trượt có sử dụng cơ cấu tăng ma sát trong dạng đĩa ma sát có cấu tạo như sau:



Hình 2. Sơ đồ bộ vi sai hạn chế trượt sử dụng cơ cấu tăng ma sát trong dạng đĩa ma sát

(1) Bánh răng vành chậu; (2) Vỏ vi sai; (3) Bánh răng hành tinh; (4) Bánh răng bán trục; (5) Đĩa ma sát; (6) Đĩa thép ;(7) Đĩa ép.

Khi mô men xoắn T được truyền đến vỏ vi sai, lực tiếp tuyến xuất hiện tại bề mặt tiếp xúc giữa cơ cấu cam–ramp và phần tử trung gian (bi hoặc con lăn). Dưới tác dụng của lực này, phần tử trung gian chuyển động theo phương nghiêng của mặt ramp, tạo ra thành phần lực dọc trục F_N Lực dọc trục sinh ra được truyền đến đĩa ép và cụm đĩa ma sát, làm tăng lực ép giữa các bề mặt ma sát. Kết quả là mô men ma sát trong của bộ vi sai tăng lên, hạn chế *chênh lệch tốc độ giữa hai bán trục.*

2.1. Tổng mô men ma sát trong của cụm đĩa ma sát

Theo tài liệu [1] mô men ma sát trong được xác định theo công thức:

$$M_{ms} = n \cdot \mu \cdot F_N \cdot R_{tb} \quad (1)$$

Trong đó:

n là số bề mặt ma sát.

μ là hệ số ma sát.

F_N là tổng lực ép.

R_{tb} là bán kính ma sát trung bình.

Theo tài liệu [2] mô men ra hai bán trục được xác định theo công thức:

$$T_L = \frac{T}{2} + M_{ms} \; ; \; T_R = \frac{T}{2} - M_{ms} \quad (2)$$

Trong đó:

T là mô men truyền đến vi sai.

T_L là mô men truyền đến bán trục trái.

T_R là mô men truyền đến bán trục phải.

M_{ms} là mô men ma sát trong được tính theo công thức (1).

Công thức tính bán kính ma sát trung bình:

$$R_{tb} = \frac{2}{3}.\frac{R_o^3 - R_i^3}{R_o^2 - R_i^2} \quad (3)$$

Trong đó:

R_{tb} là bán kính ma sát trung bình.

R_o là bán kính ngoài của đĩa ma sát.

R_i là bán kính trong của đĩa ma sát.

Lực tác dụng lên cụm đĩa ma sát:

$$F_N = F_{ramp} + F_{lx} \quad (4)$$

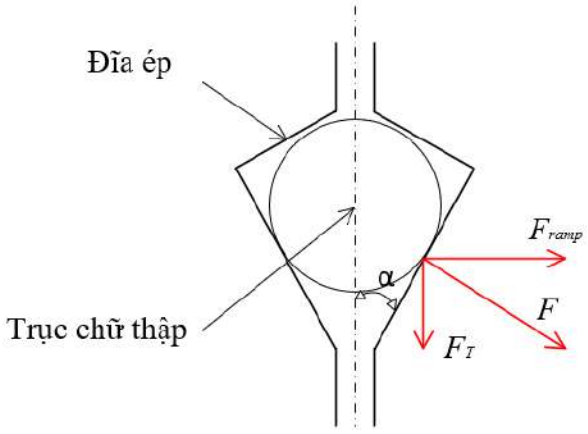
Trong đó :

F_N là lực tác dụng lên cụm đĩa ma sát.

F_{ramp} là từ đĩa ép của cơ cấu ramp tác dụng lên cụm đĩa ma sát.

F_{lx} là là lực lò xo tiền tải tác dụng lên cụm đĩa ma sát.

Lực từ cơ cấu ramp là lực sinh ra khi có mô men truyền đến bộ vi sai dẫn động cụm đĩa ép quay cùng tốc độ với vỏ vi sai



Hình 3. Cơ cấu ramp và lực sinh ra từ cơ cấu ramp

$$F_{ramp} = \frac{F_T}{\tan\alpha} = \frac{1}{2}\frac{T}{R_c.\tan\alpha} \quad (5)$$

Trong đó :

F_{ramp} là lực dọc trục do cơ cấu ramp tác dụng lên cụm đĩa ép.

F là lực tương tác giữa đĩa ép và trục chữ thập.

F_T là lực tạo ra bởi mô men xoắn của vi sai.

α là góc ramp của cơ cấu ramp.

T là mô men truyền đến vi sai.

R_c là bán kính tác dụng của cơ cấu ramp là khoảng cách từ tâm đĩa ép đến điểm tiếp xúc của trục chữ thập và đĩa ép.

Lực của lò xo tiền tải :

$$F_{lx} = k.\Delta x \quad (6)$$

Trong đó :

F_{lx} là lực lò xo tiền tải tác động

2.2. Hệ số khóa và phân phối mô men của vi sai hạn chế trượt

$$K = \frac{M_{ms\max}}{T} \quad (7)$$

Trong đó:

$M_{ms\ max}$ là mô men ma sát cực đại của vi sai hạn chế trượt

T mô men truyền đến vi sai

Theo tài liệu [1,2] đối với ô tô tải, để đảm bảo sự cân bằng giữa khả năng bám đường và tính dễ điều khiển, hệ số khóa cho ô tô tải thường được chọn trong khoảng K = 0,3 – 0,5 .

3. Tính toán thiết kế cơ cấu tăng ma sát trong

Để tính toán cụ thể, qua phân tích, bài báo lựa chọn thiết kế cho vi sai cầu xe chủ động KAMAZ 43253 có công thức bánh xe là 4x2. Các thông số chính của xe được sử dụng gồm: khối lượng toàn tải của xe 14200 kg, mô men xoắn cực đại động cơ 932 Nm, tỷ số truyền của hộp số là 4; tỷ số truyền lực chính 6,53 và bán kính bánh xe 0,5 m. Hiệu suất truyền lực η_t= 0.9.

3.1. Mô men truyền tới vi sai

Mô men tại vỏ vi sai

$$T = M_e.i_{hs}.i_0.\eta_t \quad (8)$$

$$T = 932\ .\ 7{,}82\ .\ 6{,}53\ .\ 0{,}9 \approx 42800\,Nm$$

3.2. Mô men khóa yêu cầu của vi sai hạn chế trượt

Từ mục 2.2 ta chọn hệ số khóa vi sai trong khoảng

K= 0,3 – 0,5

Chọn K = 0,4

Mô men khóa của vi sai hạn chế trượt

$$M_{ms} = K.T = 0{,}4.42800 = 17120\,Nm \quad (9)$$

Để đảm bảo độ bền của đĩa ma sát, độ bền then hoa, trục chữ thập, dự trữ tải khi va đập, chọn hệ số an toàn: n = 1,3-1,5

$$M_{ms\max} = n.M_{ms} = 1,4.17120 = 23968Nm \quad (10)$$

3.4. Giả thiết hình học và vật liệu đĩa ma sát

a) Vật liệu đĩa

- Đĩa ma sát thiêu kết kim loại (sintered) – dùng cho xe tải
- Hệ số ma sát (làm việc trong dầu):

$$\mu = 0,08 \div 0.12$$

Chọn hệ số

$$\mu = 0.1$$

b) Kích thước đĩa

- Bán kính ngoài: R_o = 130 mm = 0,13 m
- Bán kính trong: R_i = 60 mm = 0,06 m

c) Bán kính ma sát trung bình

Từ công thức (3) ta tính được

$$R_{tb} = \frac{2}{3}.\frac{(R_o^3 - R_i^3)}{(R_o^2 - R_i^2)} = \frac{2}{3}.\frac{(0,13^3 - 0.06^3)}{(0,13^2 - 0,06^2)} \approx 0.099m$$

3.5. Tính lực ép cần thiết

Từ công thức tính mô men (1), ta tính được tổng lực ép tác dụng lên cụm đĩa ma sát:

$$F_N = \frac{M_{ms\max}}{n.\mu.R_{tb}}$$

Dựa trên điều kiện làm việc tải nặng, số cặp bề mặt ma sát được chọn là 8

$$F_N = \frac{23968}{8.0.1.0,099} \approx 302626N$$

Tổng diện tích ma sát:

$$A = n.\pi.(R_o^2 - R_i^2) = 8.\pi.(0.0169 - 0.0036) = 0.334m^2$$

Áp lực lên bề mặt đĩa ma sát:

$$\rho = \frac{F_N}{A} = \frac{302626}{0.334} \approx 0,9Mpa\ ;\ \rho \le \rho_{cp} = 1,8 \div 2,2Mpa$$

Theo tài liệu [3], vật liệu ma sát ướt thiêu kết cho phép áp suất bề mặt làm việc liên tục đến khoảng 1.8–2.2 MPa, và áp suất vượt quá ngưỡng này dẫn tới tăng tốc độ mòn và thay đổi hệ số ma sát.

3.6. Phân bố lực ép và tính toán góc ramp của cơ cấu

Trong cơ cấu vi sai hạn chế trượt dạng đĩa ma sát, lực ép dọc trục tác dụng lên bộ đĩa ma sát được tạo thành từ hai thành phần:

$$F_N = F_{lx} + F_{ramp}$$

3.6.1.Phân bố lực ép

Theo [1], các nghiên cứu về vi sai hạn chế trượt cho xe tải, lực lò xo tiền tải chiếm khoảng 20-40% tổng lực ép nhằm:

- Duy trì mô men khóa ban đầu cho chuyển động thẳng.
- Đảm bảo khả năng khóa ngay cả khi mô men truyền nhỏ.

Trong nghiên cứu này giả thiết:

$$\begin{aligned} F_{lx} &= 0.3F_N \\ F_{ramp} &= 0.7F_N \end{aligned} \tag{11}$$

Với tổng lực ép xác định ở mục 3,5:

$$F_N = 302626\,N$$

Suy ra:

$$F_{lx} = 0.3 \times 302626 = 90787\,N$$
$$F_{ramp} = 0.7 \times 302626 = 211838\,N$$

3.6.2 Quan hệ cơ học của cơ cấu ramp

Từ mục 2.1 ta tính được lực dọc trục tác dụng lên cụm đĩa ma sát của cơ cấu ramp:

$$F_{ramp} = \frac{1}{2}\frac{T}{R_c.\tan\alpha}$$

Từ đó ta tính được:

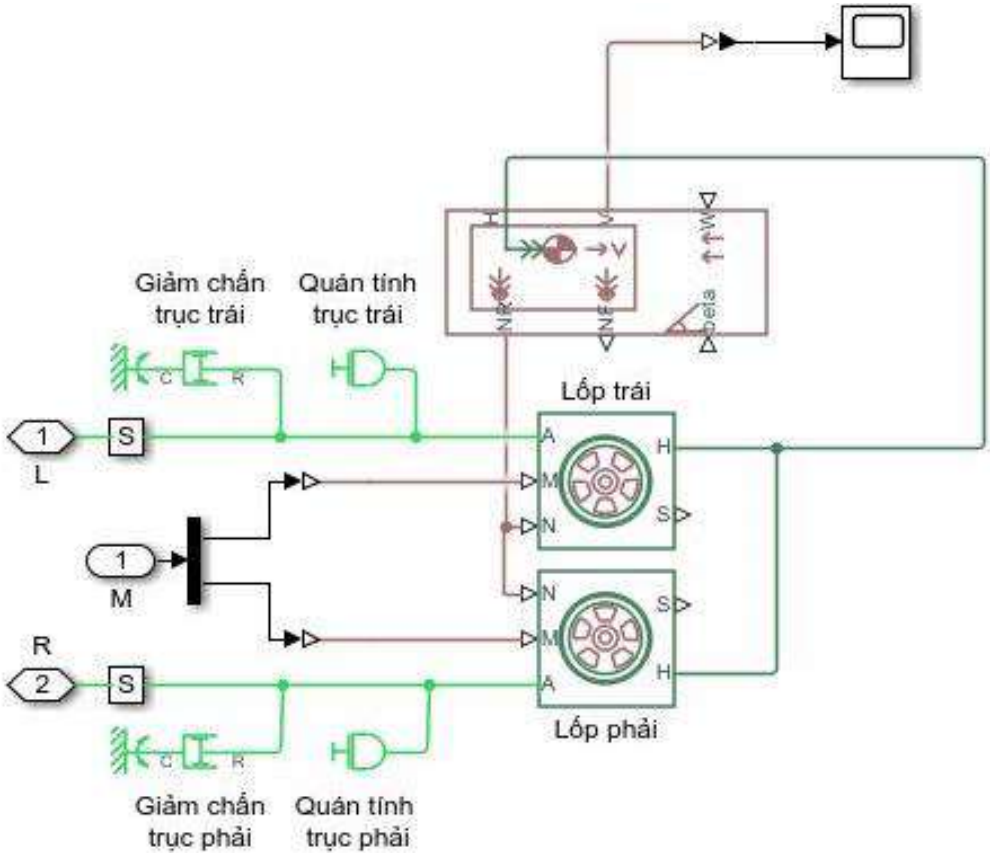
$$\tan\alpha = \frac{1}{2}\frac{T}{F_N.R_c} = \frac{1}{2}\frac{42800}{211838.0,07} \approx 1,443$$

$$\alpha \approx 55^o$$

Theo tài liệu [1], góc ramp nhỏ (30°–40°) tạo đặc tính khóa mạnh phù hợp xe hiệu suất cao, trong khi các ứng dụng xe thương mại và xe tải nặng thường sử dụng ramp lớn hơn 50° nhằm đảm bảo ổn định chuyển động, đồng thời cũng chỉ ra rằng khóa vi sai quá lớn có thể làm tăng xu hướng thiếu lái đối với xe khối lượng lớn. Với xe KAMAZ 43253 có khối lượng toàn tải 14,2 tấn, việc lựa chọn $\alpha \approx 55°$ tạo mức khóa trung bình ($\cot\alpha \approx 0.7$), đảm bảo đạt hệ số khóa thiết kế K = 0.4 mà không gây khóa quá mức.

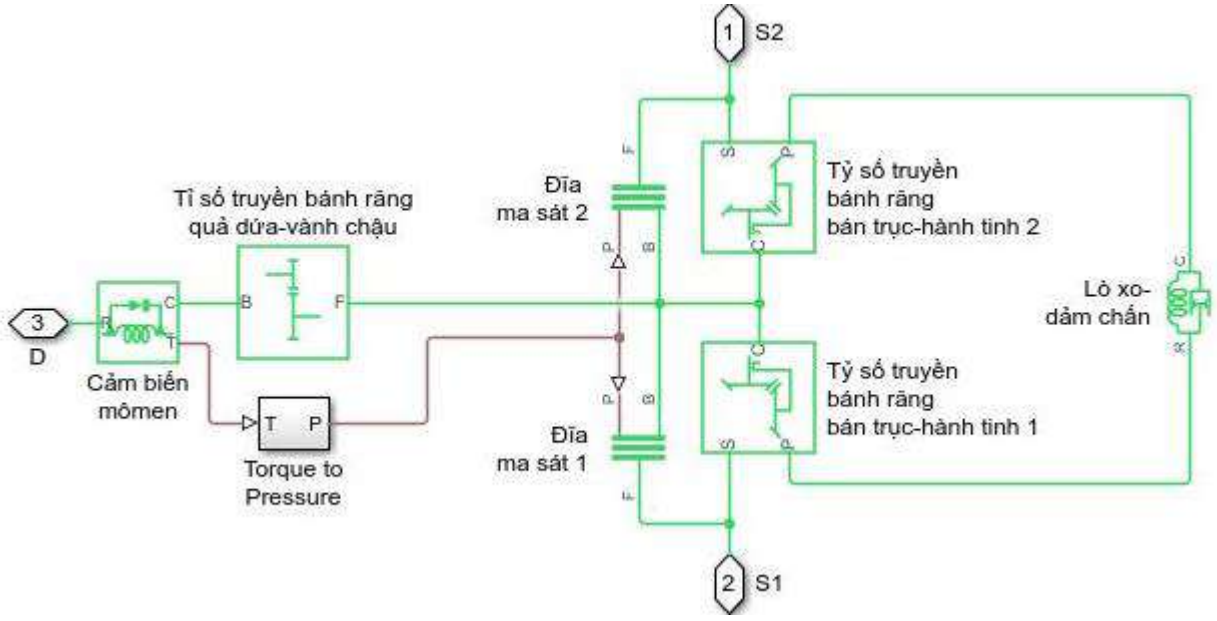
4. Mô hình mô phỏng động lực học

Mô hình động lực học hệ truyền lực xe KAMAZ 43253 được xây dựng trong Matlab/Simulink và Simscape Driveline, bao gồm nguồn mô hình động cơ, truyền lực chính, bộ vi sai, bán trục và bánh xe. Bộ vi sai hạn chế trượt được mô hình hóa bằng cách bổ sung mô men ma sát tác dụng ngược chiều lên hai bán trục.



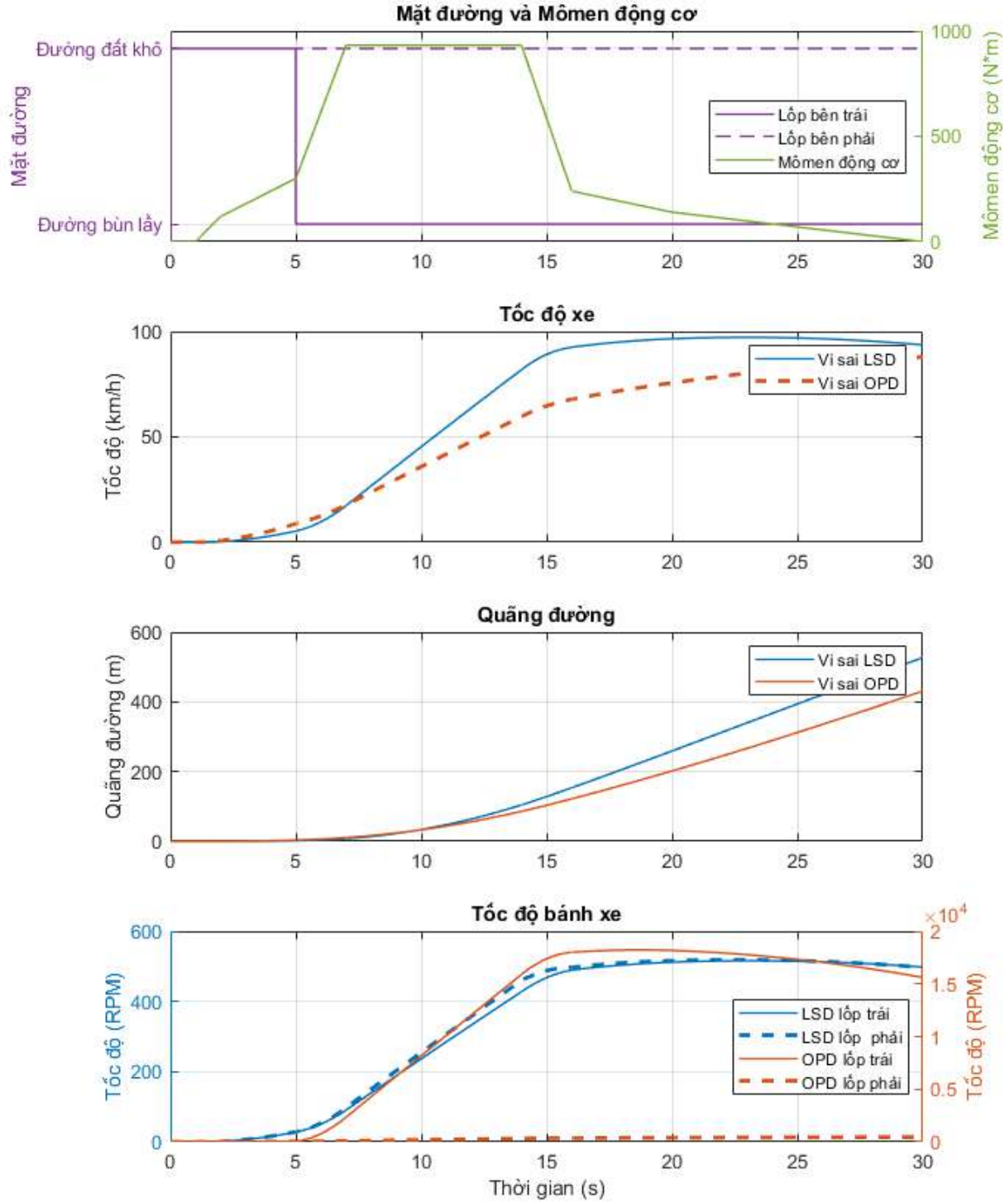
Hình 4. Mô hình lốp động lực học của xe

Cơ cấu tăng ma sát dạng đĩa ma sát được mô hình hóa dưới tác dụng của lực đến đĩa ma sát làm thay đổi mô men đến hai bán trục



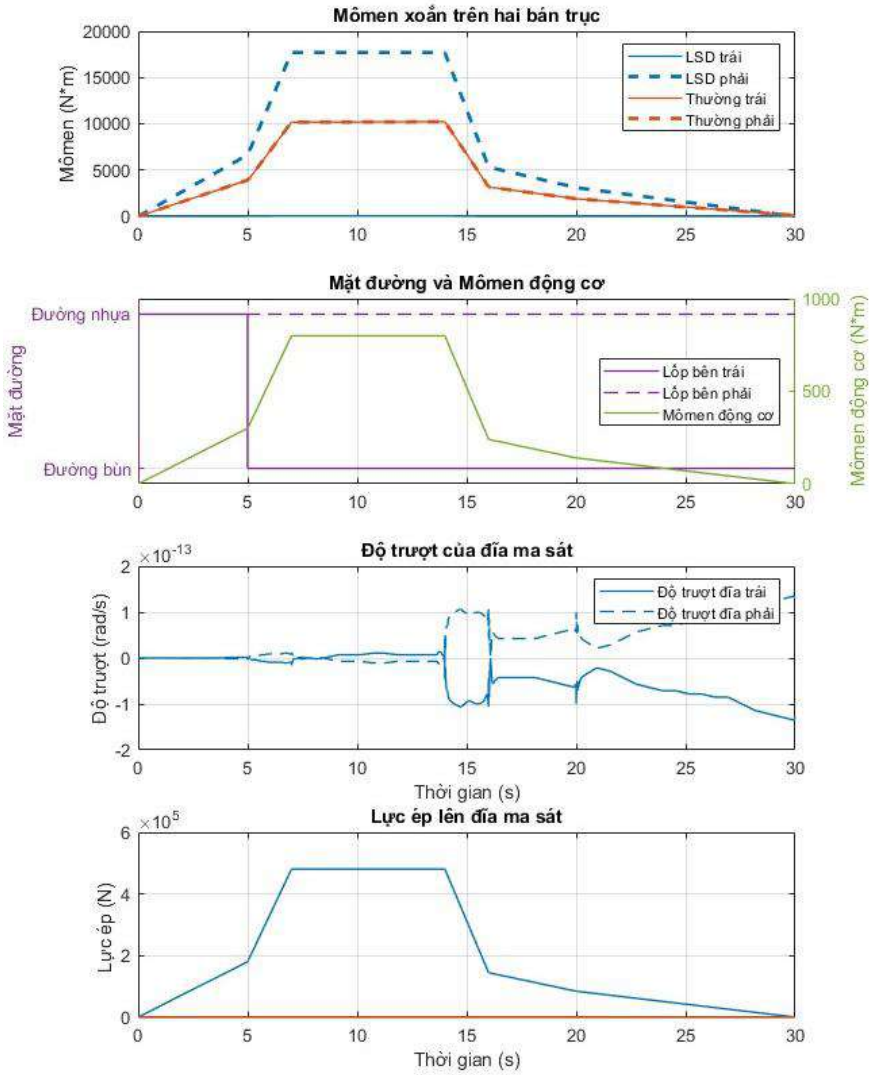
Hình 5. Mô hình vi sai hạn chế trượt sử dụng cơ cấu tăng ma sát trong dạng đĩa ma sát

5. Kết quả và thảo luận



Hình 7. Kết quả mô phỏng

Kết quả mô phỏng cho thấy khi xe chuyển động thẳng trên đường bám tốt, cơ cấu tăng ma sát hầu như không ảnh hưởng đến tính vi sai. Trong trường hợp một bánh làm việc trên mặt đường có độ bám thấp, mô men ma sát trong bộ vi sai tăng nhanh, giúp phân phối mô men xoắn ưu tiên cho bánh có độ bám tốt hơn. So với vi sai thường, vi sai hạn chế trượt giúp giảm rõ rệt hiện tượng trượt bánh và cải thiện lực kéo.



Các đồ thị mô phỏng cho thấy cơ cấu vi sai hạn chế trượt đã phát huy rõ rệt vai trò phân phối lại mô men giữa hai bán trục khi điều kiện bám đường không đồng đều. Ở giai đoạn từ 5–15 s, khi một bánh chuyển sang mặt đường bùn, vi sai thường phân phối mô men gần như bằng nhau cho hai bánh, khiến mô men bị giới hạn bởi bánh có độ bám thấp. Ngược lại, với LSD, mô men truyền tới bánh bám tốt tăng đáng kể (khoảng 17–18 kNm), trong khi bánh trượt nhận mô men nhỏ hơn, thể hiện đúng cơ chế tăng mô men ma sát trong để ưu tiên bánh có khả năng sinh lực kéo lớn hơn. Đồ thị lực ép cho thấy lực dọc trục tác dụng lên cụm đĩa ma sát tăng mạnh khi mô men động cơ tăng, phù hợp với quan hệ cơ học của cơ cấu ramp.

6. Kết luận

Nghiên cứu này có các đóng góp chính sau:

- Xây dựng quy trình thiết kế bộ vi sai hạn chế trượt cho xe tải quân sự cụ thể

Khác với các nghiên cứu mang tính lý thuyết tổng quát, bài báo triển khai đầy đủ quy ztrình tính toán thiết kế cơ cấu tăng ma sát trong dạng đĩa cho cầu sau xe chủ động KAMAZ 43253. Xe có công thức bánh xe là 4x2, từ xác định mô men truyền, hệ số khóa, lực ép đến kiểm tra điều kiện bền và áp suất làm việc.

Thiết lập mối quan hệ định lượng giữa hệ số khóa, lực ép và góc ramp.Bài báo làm rõ mối liên hệ cơ học giữa mô men truyền và lực dọc trục sinh ra bởi cơ cấu cam, qua đó cho phép lựa chọn thông số ramp phù hợp với yêu cầu làm việc của xe tải.

Tích hợp mô hình vi sai hạn chế trượt vào mô hình động lực học toàn bộ hệ truyền lực. Mô hình vi sai hạn chế trượt được xây dựng trong môi trường Matlab/Simulink kết hợp mô hình lốp, cho phép đánh giá ảnh hưởng của cơ cấu tăng ma sát trong đến chuyển động thẳng và điều kiện bám thấp.

Đánh giá định lượng hiệu quả cải thiện lực kéo so với vi sai thường. Kết quả mô phỏng chứng minh rằng cơ cấu đề xuất giúp giảm hiện tượng trượt bánh và cải thiện phân phối mô men xoắn trong điều kiện mặt đường không đồng đều.

Tài liệu tham khảo

1. Gadola, M., Chindamo, D., & Lenzo, B. (2021). Revisiting the mechanical limited-slip differential for high-performance and race car applications. *Vehicle System Dynamics*.
2. Vũ Đức Lập. (1995) *Cấu tạo ô tô quân sự (Tập 1)*. Hà Nội, Việt Nam: Học viện Kỹ thuật quân sự,
3. Zhang, Y., & Li, H. (2024). Experimental investigation of friction and wear behavior of wet clutch materials under varying surface pressures. *Tribology International*, *180*, 108061.
4. Milliken, W. F., & Milliken, D. L. (1995). *Race car vehicle dynamics*. Warrendale, PA: SAE International.

Research on the design of a friction-enhancing mechanism in the form of a friction disc for an active axle differential.

Abstract: *This paper presents the design and analytical investigation of a friction-enhancing mechanism based on a multi-disc clutch integrated into a symmetrical bevel-gear drive axle differential. The objective of the proposed mechanism is to redistribute torque and limit wheel slip under low-adhesion road conditions. The study analyzes and establishes the mechanical relationships among the transmitted torque to the half-shafts, the rotational speeds of the two wheels, and the axial clamping force acting on the clutch pack of the friction-enhancing mechanism. These relationships provide the theoretical basis for parameter calculation and structural design. Numerical simulations were performed using MATLAB/Simulink to evaluate the influence of the designed multi-disc friction mechanism on vehicle drivetrain performance. The simulation results demonstrate that the proposed mechanism significantly improves torque distribution between the two half-shafts, effectively reduces wheel slip, and enhances the vehicle's traction performance.*

Keywords: Limited-slip differential; friction-enhancing mechanism; multi-disc clutch; locking coefficient; drivetrain dynamics simulation.

Xây dựng mô hình mô phỏng khảo sát động lực học quay vòng của xe quân sự nhiều cầu dẫn hướng

Vũ Ngọc Vương[1], Nguyễn Trường Sinh[1]

[1]Học viện Kỹ thuật Quân sự

Email: ngocvuongktqs@gmail.com; Tel: 0982520289

Tóm tắt

Bài báo nghiên cứu xây dựng mô hình mô phỏng nhằm khảo sát động lực học quay vòng của xe quân sự nhiều cầu dẫn hướng. Trên cơ sở thiết lập các giả thiết hợp lý, tiến hành xây dựng mô hình toán học khảo sát chuyển động quay vòng của xe bốn cầu có hai cầu dẫn hướng phía trước, có kể đến ảnh hưởng của một số yếu tố trong tương tác của bánh xe với mặt đường và quy luật phối hợp góc quay dẫn hướng của các bánh xe ở các cầu dẫn hướng. Hệ phương trình vi phân mô tả chuyển động được thiết lập và triển khai mô phỏng bằng phần mềm MATLAB/Simulink để đánh giá một số chỉ tiêu quay vòng đặc trưng. Kết quả mô phỏng cho phép phân tích ảnh hưởng của cấu hình dẫn hướng nhiều cầu đến khả năng quay vòng, làm cơ sở đánh giá, đề xuất khuyến nghị phù hợp và định hướng kiểm chứng bằng thử nghiệm hoặc các mô hình chi tiết hơn trong các nghiên cứu tiếp theo.

Từ khóa: *mô phỏng; khảo sát; động lực học quay vòng; xe quân sự nhiều cầu dẫn hướng.*

1. Đặt vấn đề

Trên thế giới, nghiên cứu động lực học chuyển động quay vòng của ô tô đã được thực hiện từ rất sớm và đạt được nhiều thành tựu quan trọng, đặc biệt trong lĩnh vực mô hình hóa và mô phỏng. Các công trình nghiên cứu ban đầu tập trung vào mô hình phẳng một vết bánh xe để đơn giản hóa quá trình tính toán, nhằm khảo sát mối quan hệ giữa góc đánh lái, vận tốc chuyển động và bán kính quay vòng [1]. Trên cơ sở đó, nhiều mô hình nâng cao đã được phát triển, xét đến ảnh hưởng của mô men quán tính quay quanh trục đứng, độ đàn hồi ngang của lốp, đặc tính phi tuyến của lực bám lốp–đường và ảnh hưởng của hệ thống treo [2].

Ở Việt Nam, nghiên cứu động lực học chuyển động quay vòng của ô tô nhiều cầu đã được quan tâm từ nhiều năm qua. Tuy nhiên các nghiên cứu chuyên sâu về động lực học chuyển động quay vòng, đặc biệt là mô phỏng chuyển động quay vòng của xe nhiều cầu dẫn hướng còn tương đối hạn chế. Một số công trình đã tiếp cận bài toán quay vòng theo hướng mô hình phẳng và giả thiết đơn giản, chưa xét đầy đủ ảnh hưởng của nhiều cầu dẫn hướng, sự phối hợp góc quay dẫn hướng của các bánh xe, đặc tính phi tuyến của lốp và sự phân bố lại phản lực pháp tuyến lên từng bánh xe [3].

Trên cơ sở đó, việc xây dựng mô hình phẳng hai vết mô phỏng động lực học quay vòng của xe quân sự nhiều cầu dẫn hướng có xét đến các yếu tố kể trên vẫn là một hướng nghiên cứu cần được tiếp tục phát triển.

2. Thiết lập mô hình nghiên cứu động lực học quay vòng của xe quân sự nhiều cầu dẫn hướng

2.1. Các giả thiết

Bài báo đưa ra các giả thiết nhằm đơn giản hóa bài toán nhưng vẫn đảm bảo đánh giá, khảo sát bài toán quay vòng của xe quân sự nhiều cầu dẫn hướng với độ chính xác tương đối. Cụ thể:

- Mô hình xe 4 cầu được coi là đối xứng theo mặt phẳng dọc trục của xe và bỏ qua ảnh hưởng của độ lệch trọng tâm theo phương ngang;

- Thân xe được coi là một hình phẳng có khối lượng m, trọng tâm C, chuyển động song phẳng trong mặt phẳng song song với mặt phẳng nền đường. Quỹ đạo chuyển động được xác lập bởi ba tọa độ x, y và ψ ;

- Mỗi cầu xe thứ *i* được mô hình hóa như một khối đàn hồi theo phương đứng, có độ cứng thẳng đứng tương ứng k_i;

- Xe không kéo moóc, chuyển động trên mặt phẳng nằm ngang không có mấp mô;

- Bỏ qua lực cản không khí tác dụng lên ô tô trong quá trình chuyển động;

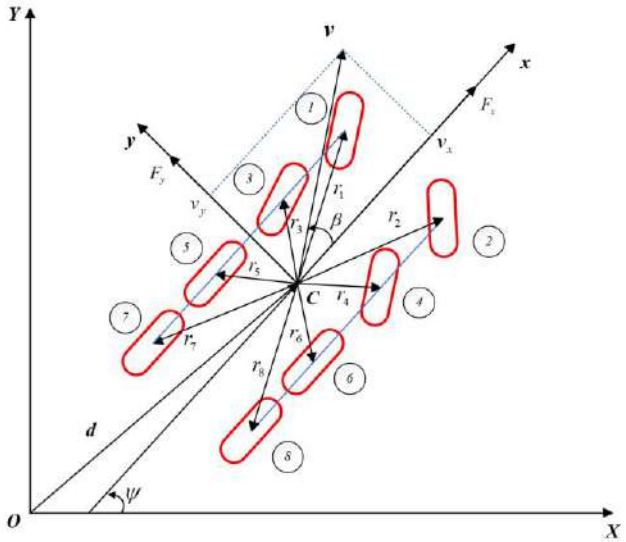
- Chỉ khảo sát trường hợp xe quay vòng đều.

Để nâng cao độ chính xác của quá trình tính toán khảo sát, bài toán xây dựng mô hình phẳng hai vết, tiến hành tính toán và bổ sung sự phân bố lại tải trọng của xe tác dụng lên từng bánh trong quá trình quay vòng.

2.2. Mô hình toán học

2.2.1. Mô hình phẳng và phương trình động lực học ô tô

Theo [4], ô tô được coi là một tấm phẳng chuyển động trong mặt phẳng đường, quỹ đạo chuyển động được xác lập bởi ba tọa độ x, y và ψ . Phương trình vi phân chuyển động Newton – Euler của ô tô được viết cho trọng tâm C của ô tô, biểu diễn trong hệ tọa độ gắn với thân xe B(Cxyz).



Hình 1. Mô hình phẳng hai vết của xe trong hệ tọa độ G(OXYZ) và B(Cxyz)

Với $\omega_z = \dot{\psi} = r \Rightarrow \dot{\omega}_z = \ddot{\psi} = \dot{r}$, ta có hệ phương trình sau:

$$\begin{cases} F_x = m\dot{v}_x - mrv_y \\ F_y = m\dot{v}_y + mrv_x \\ M_z = \dot{r}I_z \end{cases} \quad (1)$$

Trong đó:

F_x là lực tổng hợp tác dụng lên ô tô theo phương Cx

F_y là lực tổng hợp tác dụng lên ô tô theo phương Cy

M_z là momen tổng hợp tác dụng lên ô tô theo phương Cz

v_x là vận tốc ô tô theo phương Cx

v_y là vận tốc ô tô theo phương Cy

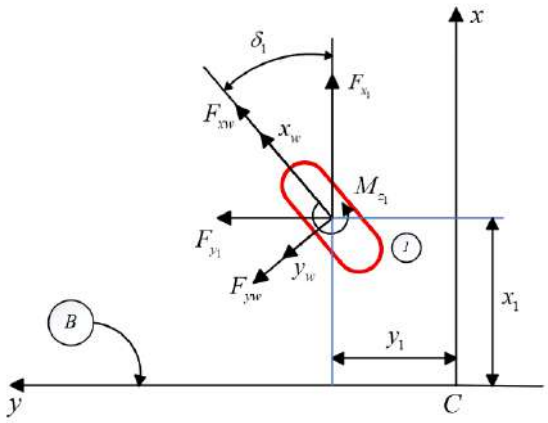
r vận tốc góc của ô tô theo phương Cz

m là khối lượng ô tô

I_z momen quán tính của ô tô đối với trục Cz

2.2.2. Các lực tác dụng lên xe trong quá trình quay vòng

Hình 2 mô tả bánh xe số 1 và các thành phần lực của nó trong mặt phẳng nền Cxy.



Hình 2. Các lực tác dụng lên bánh xe số 1

Ta có tổng hợp các lực tác dụng lên bánh xe số 1 là:

$$\begin{cases} F_{x_1} = F_{xw_1}\cos\delta_1 + F_{yw_1}\sin\delta_1 \\ F_{y_1} = F_{yw_1}\cos\delta_1 + F_{xw_1}\sin\delta_1 \\ M_{z_1} = M_{zw_1} + x_1F_{y_1} - y_1F_{x_1} \end{cases} \quad (2)$$

Như vậy, tổng hợp các lực bánh xe tác dụng lên thân xe trong hệ cục bộ $B(Cxyz)$ gắn với thân xe sẽ là:

$$\begin{cases} {}^BF_x = \sum_{i=1}^{8} F_{x_i} = \sum_{i=1}^{8} F_{xw_i}\cos\delta_i - \sum_{i=1}^{8} F_{yw_i}\sin\delta_i \\ {}^BF_y = \sum_{i=1}^{8} F_{y_i} = \sum_{i=1}^{8} F_{yw_i}\cos\delta_i + \sum_{i=1}^{8} F_{xw_i}\sin\delta_i \\ {}^BM_z = \sum_{i=1}^{8} F_{zw_i} = \sum_{i=1}^{8} x_iF_{y_i} - \sum_{i=1}^{8} y_iF_{x_i} \end{cases} \quad (3)$$

Các véc tơ tọa độ từng bánh xe (Hình 1):

$$\mathbf{r_1}=\begin{bmatrix}a_1\\b_1\\0\end{bmatrix}\quad \mathbf{r_2}=\begin{bmatrix}a_1\\-b_2\\0\end{bmatrix}\quad \mathbf{r}_3=\begin{bmatrix}a_2\\b_1\\0\end{bmatrix}\quad \mathbf{r}_4=\begin{bmatrix}a_2\\-b_1\\0\end{bmatrix}$$
$$\mathbf{r}_5=\begin{bmatrix}-a_3\\b_1\\0\end{bmatrix}\quad \mathbf{r}_6=\begin{bmatrix}-a_3\\-b_2\\0\end{bmatrix}\quad \mathbf{r}_7=\begin{bmatrix}-a_4\\b_1\\0\end{bmatrix}\quad \mathbf{r}_8=\begin{bmatrix}-a_4\\-b_2\\0\end{bmatrix} \tag{4}$$

Nếu $\delta_1;\delta_2;\delta_3;\delta_4$ nhỏ, ta có:

$$F_x \approx F_{x_1}+F_{x_2}+F_{x_3}+F_{x_4}+F_{x_5}+F_{x_6}+F_{x_7}+F_{x_8} \tag{5}$$

$$F_y \approx F_{y_1}+F_{y_2}+F_{y_3}+F_{y_4}+F_{y_5}+F_{y_6}+F_{y_7}+F_{y_8} \tag{6}$$

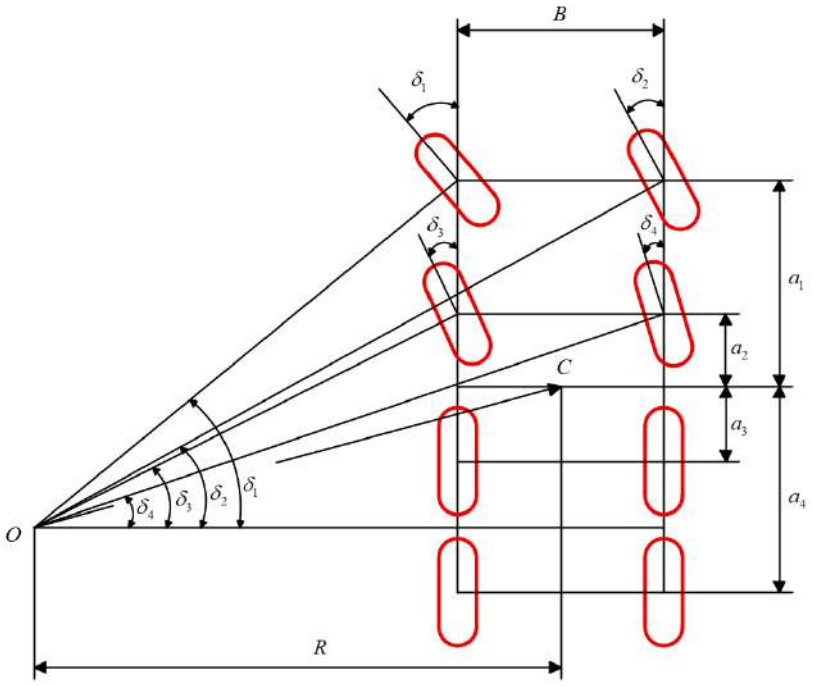
$$M_z \approx a_1F_{y_1}+a_1F_{y_2}+a_2F_{y3}+a_2F_{y_4}-a_3F_{y_5}-a_3F_{y_6}-a_4F_{y_7}-a_4F_{y_8} \tag{7}$$

Các thành phần vận tốc, các góc từng bánh xe lần lượt là:

$$v_{x_i}=v_x-y_ir \qquad v_{y_i}=v_y+x_ir \tag{8}$$

$$\beta_i=\arctan\left(\frac{v_{y_i}}{v_{x_i}}\right)=\arctan\left(\frac{v_y+x_ir}{v_x-y_ir}\right)\approx\frac{v_y+x_ir}{v_x-y_ir} \tag{9}$$

Với mô hình hai vết bánh xe, ta có các quan hệ động học như sau:



Hình 3. Quan hệ động học trong mô hình hai vết bánh xe

Theo Hình 3 ta có:

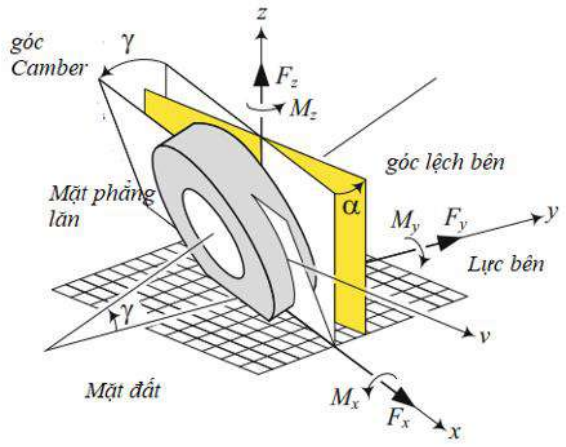
$$\tan\delta_1 = \frac{a_1 + \frac{a_3 + a_4}{2}}{R - \frac{B}{2}} = \frac{2a_1 + a_3 + a_4}{2R - B}; \quad \tan\delta_2 = \frac{a_1 \frac{a_3 + a_4}{2}}{R + \frac{B}{2}} = \frac{2a_1 + a_3 + a_4}{2R + B}$$

$$\tan\delta_3 = \frac{a_2 + \frac{a_3 + a_4}{2}}{R - \frac{B}{2}} = \frac{2a_2 + a_3 + a_4}{2R - B}; \quad \tan\delta_4 = \frac{a_2 + \frac{a_3 + a_4}{2}}{R + \frac{B}{2}} = \frac{2a_2 + a_3 + a_4}{2R + B} \tag{10}$$

2.2.3. Mô hình bánh xe

Có rất nhiều mô hình bánh xe – đường được sử dụng hiện nay. Trong đó mô hình theo các hàm kinh nghiệm MF Tire model (Magic Forluma tire model) được coi là tối ưu nhất để mô phỏng các lực tương tác giữa bánh xe và đường. Sau đây là một mô hình bánh xe – đường PAC2002 do Pacejka xây dựng [5, 6].

Hệ tọa độ, lực, momen và quy ước chiều trong mô hình bánh xe – đường PAC2002:



Hình 4. Hệ tọa độ, lực, momen và quy ước chiều

Theo Pacejka, lực ngang của lốp là một hàm phụ thuộc vào: góc lệch bên α, góc Camber γ và phản lực pháp tuyến từ đường F_z:

$$F_y = F_y(\alpha, \gamma, F_z) = D\sin(C\arctan(B\Phi)) + \Delta S_v \tag{11}$$

Trong đó:

$$\alpha_i = \delta_i - \beta_i \tag{12}$$

$$\Phi = (1 - E)(\alpha + \Delta S_h) + \left(\frac{E}{B}\right)\arctan\left(B(\alpha + \Delta S_h)\right) \tag{13}$$

$C = 1.30$ đối với lực ngang

$$B=\left(\frac{\lambda_3 \sin\left(\lambda_4 \arctan\left(\lambda_5 F_z\right)\right)}{CD}\right)\left(1-\lambda_{12}|\gamma|\right), D=\lambda_1 F_z^{\,2}+\lambda_2 F_z \tag{14}$$

$$E=\lambda_6 F_z^{\,2}+\lambda_7 F_z+\lambda_8, \Delta S_h=\lambda_9 \gamma, \Delta S_v=\left(\lambda_{10} F_z^{\,2}+\lambda_{11} F_z\right)\gamma \tag{15}$$

B: Thông số độ cứng C: Thông số hình học

D: Biên độ lực E: Thông số độ cong

Trong bài báo này, chỉ tập trung nghiên cứu lực bên của lốp khi không kể đến ảnh hưởng của góc Camber và bỏ qua ảnh hưởng của momen cản quay vòng. Khi đó: $\gamma=0$, $\Delta S_h=0, \Delta S_v=0$.

Để xác định phản lực pháp tuyến tác dụng lên từng bánh, bài báo sử dụng mô hình heave-pitch-roll [7], độ biến dạng đàn hồi thẳng đứng tại bánh *j* (trái/phải) của cầu thứ *i* được xác định theo biểu thức:

$$\delta_{zij}=z+a_i\theta \pm b\varphi \tag{16}$$

Trong đó: dấu "+" ứng với bánh phải, dấu "-" ứng với bánh trái.

Phản lực pháp tuyến tại từng bánh được xác định thông qua quan hệ đàn hồi tuyến tính:

$$F_{zij}=k_i\left(z+a_i\theta \pm b\varphi\right) \qquad \left(i=1..4,\ j=L,R\right) \tag{17}$$

Trong trạng thái cân bằng tĩnh học hoặc trong trạng thái quay vòng đều, thân xe thỏa mãn các điều kiện cân bằng theo hệ phương trình sau:

$$\begin{cases} \sum_{i=1}^{4}\sum_{j=L,R} F_{zij}=mg \\ \sum_{i=1}^{4}\sum_{j=L,R} a_i F_{zij}=0 \\ \sum_{i=1}^{4}\left(F_{ziR}-F_{ziL}\right)b=ma_y h \end{cases} \tag{18}$$

Hệ phương trình trên có nghiệm duy nhất, cho phép xác định các chuyển vị tổng quát của thân xe z,θ,φ. Trong trường hợp giả thiết $k_1 = k_2 = k_3 = k_4 = k$. Ta thu được giá trị phản lực pháp tuyến tại từng bánh như sau:

$$F_{z1}=k\left(z+a_1\theta-b\varphi\right) \qquad F_{z2}=k\left(z+a_1\theta+b\varphi\right)$$

$$F_{z3}=k\left(z+a_2\theta-b\varphi\right) \qquad F_{z4}=k\left(z+a_2\theta+b\varphi\right) \tag{14}$$

$$F_{z5}=k\left(z+a_3\theta-b\varphi\right) \qquad F_{z6}=k\left(z+a_3\theta+b\varphi\right)$$

$$F_{z7} = k\left(z + a_4\theta - b\varphi\right) \qquad F_{z8} = k\left(z + a_4\theta - b\varphi\right)$$

Như vậy, mô hình cho phép xác định 8 phản lực pháp tuyến tại các bánh xe của xe có công thức bánh xe 8x8 với cơ sở vật lý rõ ràng, phản ánh đầy đủ ảnh hưởng của phân bố khối lượng, hình học xe và chuyển tải ngang khi quay vòng.

Từ (1), (5), (6), (7) ta có:

$$\begin{cases} m\dot{v}_x - mrv_y = F_{x_1} + F_{x_2} + F_{x_3} + F_{x_4} + F_{x_5} + F_{x_6} + F_{x_7} + F_{x_8} \\ m\dot{v}_y + mrv_x = F_{y_1} + F_{y_2} + F_{y_3} + F_{y_4} + F_{y_5} + F_{y_6} + F_{y_7} + F_{y_8} \\ \dot{r}I_z = a_1F_{y_1} + a_1F_{y_2} + a_2F_{y3} + a_2F_{y_4} - a_3F_{y_5} - a_3F_{y_6} - a_4F_{y_7} - a_4F_{y_8} \end{cases} \tag{15}$$

Với $v_x = const, \dot{v}_x = 0$ ta có:

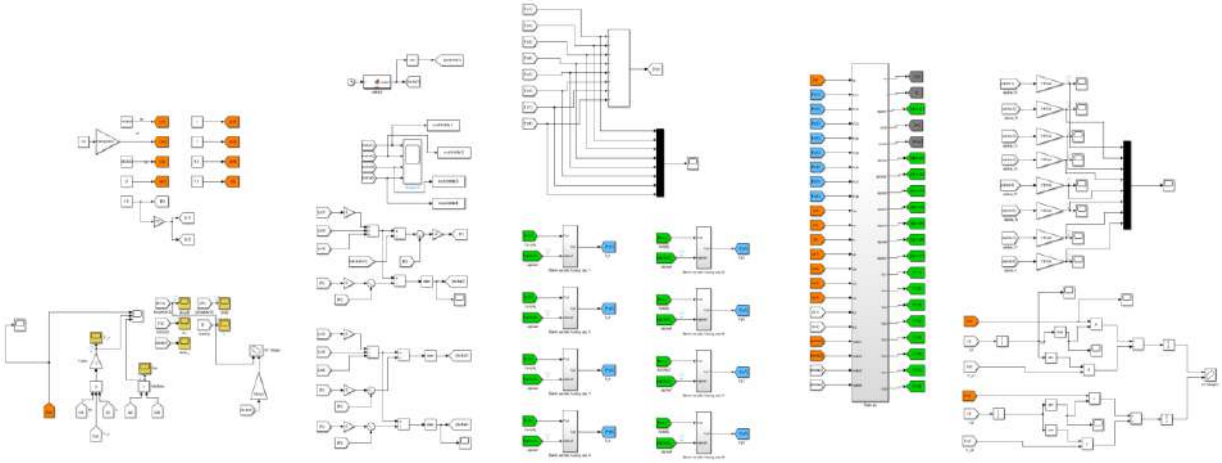
$$\begin{cases} -mrv_y = F_x \\ \dot{v}_y = \dfrac{F_{y_1} + F_{y_2} + F_{y_3} + F_{y_4} + F_{y_5} + F_{y_6} + F_{y_7} + F_{y_8}}{m} - rv_x \\ \dot{r} = \dfrac{a_1\left(F_{y_1} + F_{y_2}\right) + a_2\left(F_{y3} + F_{y_4}\right) - a_3\left(F_{y_5} + F_{y_6}\right) - a_4\left(F_{y_7} + F_{y_8}\right)}{I_z} \end{cases} \tag{16}$$

Trong đó $F_{y_i} = F_{y_i}\left(\alpha, F_{z_i}\right)$ và các góc α_i được tính theo công thức (11). Các phản lực F_{z_i} được tính theo công thức (14). Quan hệ hình học giữa các góc δ_i được tính theo công thức (10).

3. Kết quả nghiên cứu

3.1. Thông số đầu vào và chế độ mô phỏng

Bài báo sử dụng phần mềm MATLAB/Simulink để mô phỏng quá trình quay vòng và sử dụng thông số của xe BTR-60PB để tiến hành khảo sát như sau:



Hình 5. Sơ đồ mô phỏng

Bảng 1. Thông số kỹ thuật của xe BTR-60PB.

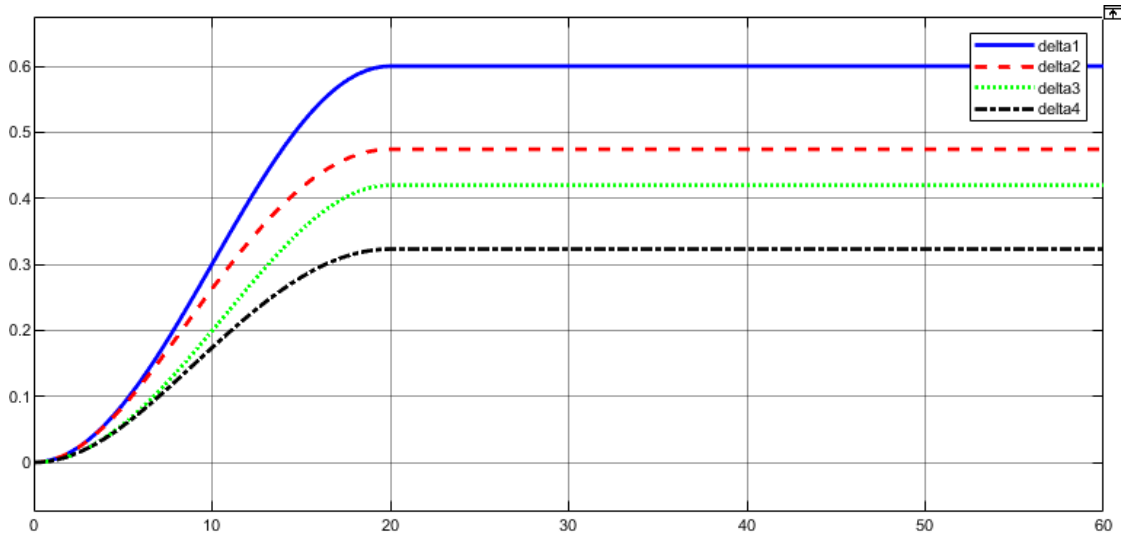
STT	Thông số kỹ thuật	Ký hiệu	Giá trị	Đơn vị
1	Các kích thước:			
	Chiều dài	L	7,5	m
	Chiều rộng	B	2,8	m
	Chiều cao trọng tâm	h	1.7	m
	Khoảng cách từ cầu trước dẫn hướng đến trọng tâm	a_1	3	m
	Khoảng cách từ cầu sau dẫn hướng đến trọng tâm	a_2	1	m
	Khoảng cách từ cầu trước dẫn động đến trọng tâm	a_3	1	m
	Khoảng cách từ cầu sau dẫn động đến trọng tâm	a_4	4,5	m
	Khoảng cách từ tâm vết bánh xe bên trái đến trọng tâm	b_1	1,4	m
	Khoảng cách từ tâm vết bánh xe bên phải đến trọng tâm	b_2	1,4	m
2	Trọng lượng khi xe không tải	m	10300	kg
3	Mô men quán tính theo trục z	I_z	10000	kg.m^2
4	Vân tốc quay vòng	v	10	km/h
5	Kích thước lốp		425/85 R21	

3.2. Kết quả mô phỏng

Bài báo sử dụng hàm góc quay của bánh xe dẫn hướng là hàm bước:

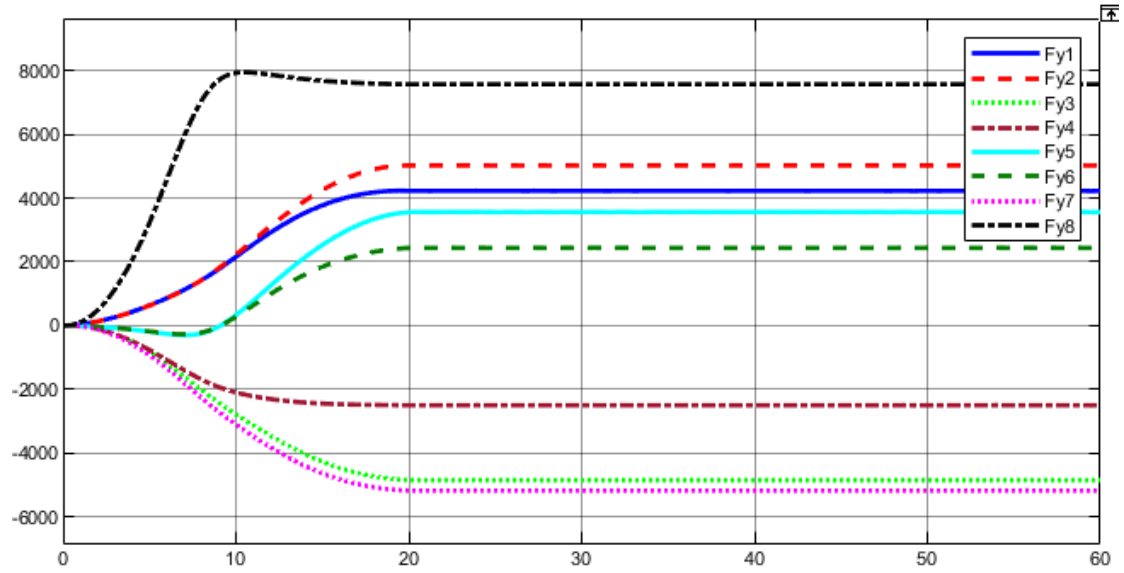
$$\delta_1 = \begin{cases} 0,3+0,3\sin\left(\dfrac{t}{20}\pi - \dfrac{\pi}{2}\right) & \text{(rad)} \quad 0 \le t \le 20s \\ 0,6 & \text{(rad)} \quad t > 20\text{s} \end{cases}$$

Với vận tốc của xe khi quay vòng là 10 km/h, thời gian mô phỏng trong 60 s kết quả thu được như sau:



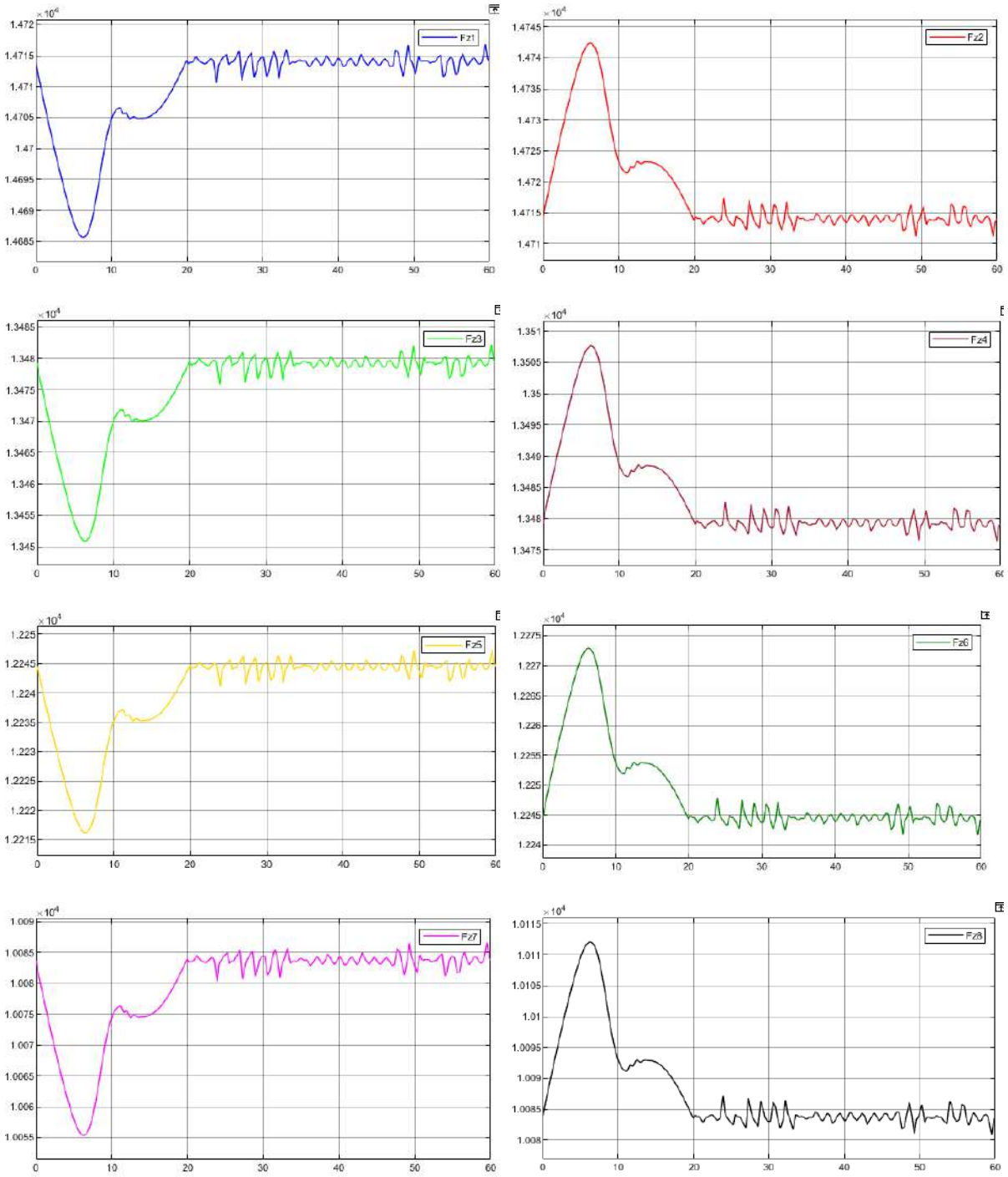
Hình 6. Góc quay của bánh xe dẫn hướng

+ Kết quả mô phỏng với vận tốc quay vòng 10 km/h và tín hiệu góc dẫn hướng dạng hàm bước cho thấy quy luật phối hợp góc quay giữa các bánh/các cầu dẫn hướng là hợp lý: các góc dẫn hướng tăng dần đến giá trị ổn định, đảm bảo trạng thái quay vòng đều sau giai đoạn quá độ. Sự khác biệt góc quay giữa các bánh và giữa các cầu dẫn hướng phản ánh đúng yêu cầu động học khi quay vòng (bánh phía trong cua cần góc lớn hơn bánh phía ngoài, các cầu dẫn hướng có thể phối hợp theo mục tiêu giảm bán kính quay vòng và hạn chế trượt ngang).



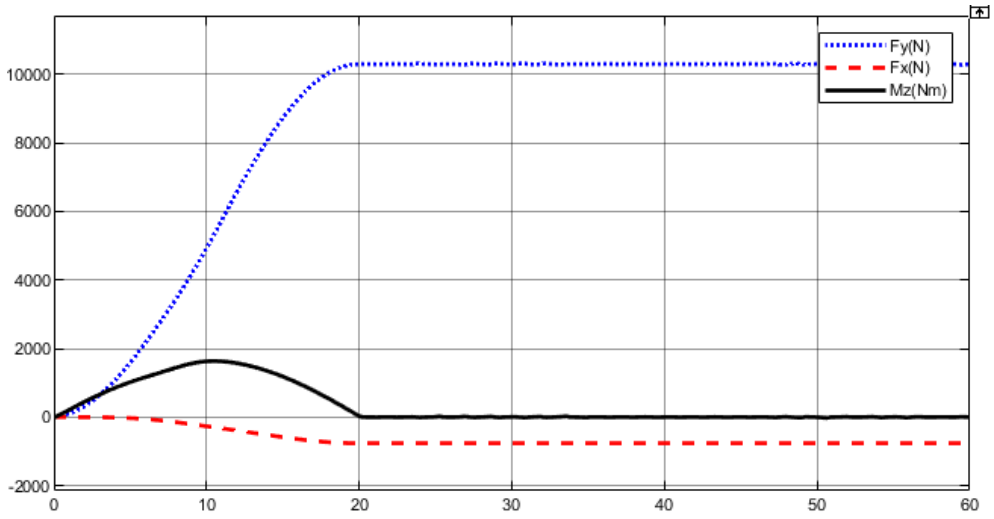
Hình 7. Lực ngang tác dụng lên các bánh xe khi quay vòng

+ Lực ngang tại các bánh thay đổi tương ứng với góc lệch bên và phản lực pháp tuyến của từng bánh: các bánh chịu tải lớn (thường là bánh phía ngoài) có xu hướng tạo lực bên lớn hơn; đồng thời đặc tính phi tuyến của mô hình lốp thể hiện qua sự tăng lực theo góc trượt và tiến tới trạng thái "bão hòa" khi góc trượt tăng. Điều này cho thấy mô hình lực ngang phi tuyến được triển khai cho phép phản ánh sát hơn đặc tính bám đường so với mô hình tuyến tính.



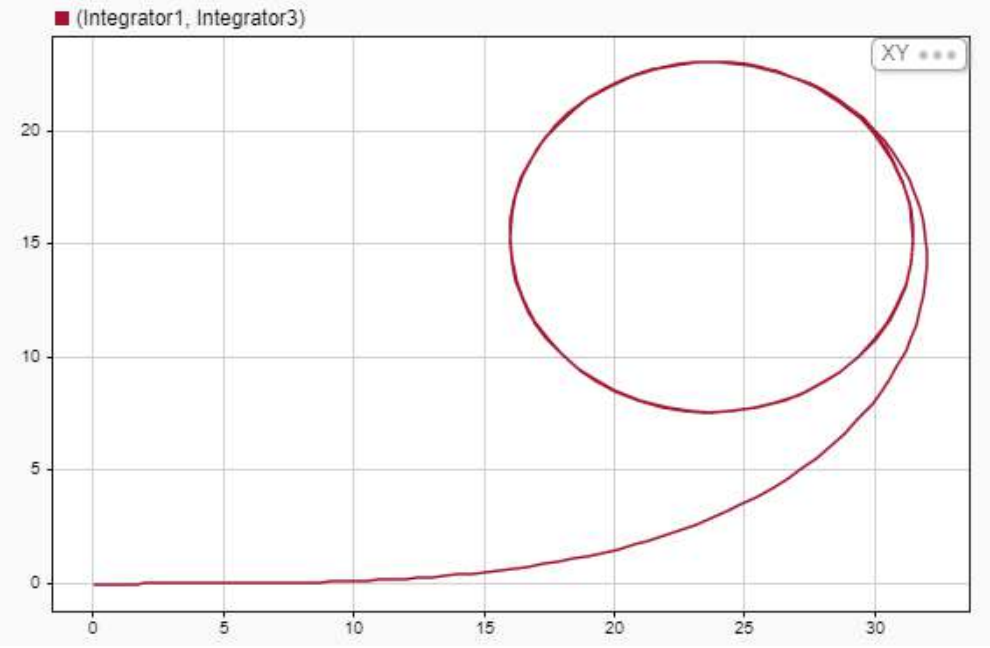
Hình 8. Các phản lực pháp tuyến tác dụng lên các bánh xe

+ Các phản lực pháp tuyến tại bánh biến thiên theo quy luật chuyển tải ngang khi xe quay vòng: phản lực phía ngoài cua tăng lên trong khi phía trong cua giảm xuống. Kết quả này phù hợp với mô hình tính phản lực dựa trên cân bằng lực–mô men và độ cứng thẳng đứng của các cầu, đồng thời cho phép đánh giá định lượng mức độ chuyển tải và nguy cơ giảm tải bánh phía trong khi tăng gia tốc ngang.



Hình 9. Lực F_x, F_y, M_z tác dụng lên thân xe

+ Các đại lượng tổng hợp tác dụng lên thân xe gồm F_x, F_y và mô men quay vòng M_z biến thiên đến giá trị ổn định, thể hiện sự hình thành trạng thái quay vòng đều. Trong điều kiện khảo sát (vận tốc thấp, giả thiết đường bằng phẳng và bỏ qua lực cản khí động theo phương ngang), thành phần F_y và M_z là hai đại lượng chi phối động lực học quay vòng; kết quả mô phỏng cho thấy mô hình đảm bảo tính nhất quán giữa lực/mô men tổng hợp và đáp ứng chuyển động theo hệ phương trình Newton–Euler đã thiết lập.



Hình 10. Quỹ đạo chuyển động của xe

+ Kết quả mô phỏng quỹ đạo chuyển động cho thấy ô tô trải qua giai đoạn quá độ trước khi đạt trạng thái quay vòng ổn định. Ban đầu quỹ đạo gần như tuyến tính do vận tốc góc quay quanh trục đứng còn nhỏ; sau đó độ cong tăng dần cùng với sự gia tăng của vận tốc góc quay quanh trục đứng và gia tốc ngang. Khi hệ đạt trạng thái quay vòng đều, quỹ đạo tiệm cận dạng đường tròn với bán kính gần như không đổi. Điều này phù hợp với quy luật động học của chuyển động quay vòng phẳng và phản ánh đúng đặc tính đáp ứng quá độ khi chuyển từ chuyển động thẳng sang chuyển động cong dưới tác động của góc đánh lái.

4. Kết luận

Bài báo đã xây dựng mô hình mô phỏng khảo sát động lực học quay vòng của xe quân sự nhiều cầu dẫn hướng trên cơ sở hệ phương trình Newton–Euler cho chuyển động song phẳng 3 bậc tự do, kết hợp mô hình lực ngang phi tuyến của lốp và mô hình xác định phản lực pháp tuyến tại từng bánh có xét đến chuyển tải khi quay vòng. Mô hình được triển khai trên MATLAB/Simulink, cho phép đánh giá các đại lượng đặc trưng như quy luật góc dẫn hướng, phân bố phản lực pháp tuyến, lực ngang tại bánh và lực–mô men tổng hợp tác dụng lên thân xe, từ đó phân tích ảnh hưởng của cấu hình dẫn hướng nhiều cầu đến khả năng quay vòng.

Kết quả mô phỏng với bộ thông số xe BTR-60PB và vận tốc 10 km/h cho thấy các đại lượng động lực học biến thiên theo quy luật hợp lý và tiến tới trạng thái ổn định, phản ánh đúng bản chất chuyển tải ngang và hình thành lực bên trong quá trình quay vòng. Đây là cơ sở để sử dụng mô hình như một công cụ khảo sát nhanh, phục vụ so sánh các phương án phối hợp góc dẫn hướng giữa các cầu, cũng như đánh giá xu hướng ổn định quay vòng trong giai đoạn thiết kế/nghiên cứu.

Trong các nghiên cứu tiếp theo, cần mở rộng mô hình theo hướng khảo sát chế độ quay vòng không đều (đánh lái biến thiên theo thời gian), dải vận tốc lớn hơn; bổ sung các yếu tố ảnh hưởng như đặc tính treo–lốp chi tiết hơn, điều kiện đường không bằng phẳng, tương tác lực dọc–lực ngang và tiến hành kiểm chứng bằng thử nghiệm hoặc đối sánh với các mô hình độ trung thực cao nhằm hiệu chỉnh tham số và nâng cao độ tin cậy của kết quả mô phỏng.

Tài liệu tham khảo

1. Wan Muhamad, W. M. (2018). *Vehicle steering dynamic calculation and simulation*. International Journal of Innovative Technology and Interdisciplinary Sciences, 2(1), 87–97.
2. Bakker, E., Pacejka, H. B., & Lidner, L. (1989). *A new tire model with an application in vehicle dynamics studies*. Vehicle System Dynamics, 18(S1), 337–348. London, UK: Taylor & Francis.
3. Đ. V. Tứ. (2022). *Mô hình nghiên cứu động lực học quay vòng của ô tô hai cầu khi kể đến tính bất đối xứng dọc*. Tạp chí Cơ khí Việt Nam, số 4.
4. V. V. Hường, N. T. Dũng, D. N. Khánh, Đ. H. Phúc. (2014). *Động lực học ô tô*. Nhà xuất bản Giáo dục Việt Nam, Hà Nội.
5. Pacejka, H. B. (2006). *Tyre and vehicle dynamics* (2nd ed.). Oxford, UK: Butterworth-Heinemann.
6. Pacejka, H. B., & Besselink, I. J. M. (2002). *Magic Formula tyre model with transient properties*. Vehicle System Dynamics, 38(S1), 234–249.
7. Johansson, J. (2014). *Vehicle dynamics modelling and simulation for ride and handling studies* (Master's thesis). Gothenburg, Sweden: Chalmers University of Technology.

Simulation modeling and analysis of yaw dynamics in multi-axle steering military vehicles

Abstract: *This paper presents the development of a simulation model to investigate the cornering dynamics of a multi-axle military vehicle equipped with multiple steering axles. Based on a set of rational assumptions, a mathematical model was established to analyze the cornering behavior of a four-axle vehicle with two front steering axles, taking into account selected factors in tire–road interaction as well as the steering angle coordination law among the steered wheels. The governing differential equations describing the vehicle motion were derived and implemented in MATLAB/Simulink to evaluate characteristic cornering performance metrics. The simulation results enable an assessment of the influence of multi-axle steering configuration on vehicle cornering capability, providing a basis for evaluation, recommendations, and further validation through experimental testing or more detailed vehicle models in subsequent studies.*

Keywords: *simulation, analysis, cornering dynamics, multi-axle steering military vehicle.*

Nghiên cứu động lực học hệ thống treo cân bằng phụ thuộc ở ô tô quân sự 3 cầu

Vũ Thế Bảo[1], Vũ Đức Lập[1]

[1]Học viện Kỹ thuật quân sự

Email: Ladyvajsy@gmail.com; Tel: 09766886691

Tóm tắt

Bài báo khảo sát một số yếu tố ảnh hưởng tới quá trình dao động của hệ thống treo cân bằng phụ thuộc xe ô tô quân sự ba cầu. Tác giả xây dựng mô hình khảo sát dao động trong mặt phẳng dọc của ô tô, áp dụng phương trình Lagrăng loại 2 thiết lập hệ phương trình vi phân mô tả chuyển động của ô tô. Phương pháp mô phỏng số được thực hiện trên phần mềm Matlab để khảo sát ảnh hưởng của các điều kiện chuyển động. Kết quả khảo sát cho qui luật của vận tốc, gia tốc dao động trên cơ sở đó đánh giá độ an toàn của ô tô khi chuyển động trên đoạn đường không bằng phẳng. Làm cơ sở xác định các giải pháp nâng cao an toàn chuyển động cho ô tô trong điều kiện di chuyển trên địa hình gồ ghề và điều kiện thời tiết khắc nghiệt.

Từ khóa: *Hệ thống treo cân bằng phụ thuộc; Không gian trạng thái; Phương trình Lagrăng loại 2.*

1. Đặt vấn đề

Hệ thống treo là một trong những cụm hệ thống then chốt trên ô tô quân sự 3 cầu, quyết định trực tiếp đến độ êm dịu, khả năng bám đường, độ ổn định chuyển động và mức độ an toàn khi xe hoạt động trên nhiều loại địa hình khác nhau. Đặc thù của hệ thống treo xe 3 cầu là cụm cầu sau dạng cân bằng phụ thuộc, tạo ra mối liên hệ động lực học giữa hai cầu sau, đáp ứng dao động phức tạp hơn so với xe hai cầu thông thường [1]. Vì vậy, nghiên cứu động lực học hệ thống treo của kết cấu này có ý nghĩa quan trọng trong quá trình thiết kế, tính toán lựa chọn thông số và đánh giá khả năng khai thác sử dụng.

Xuất phát từ các chỉ tiêu về an toàn chuyển động [2] và các nghiên cứu trước như của tác giả Vũ Ngọc Minh (HVKTQS) mới chỉ dừng lại ở việc nghiên cứu mô hình ½ ngang xe của xe 2 cầu [3], việc khảo sát dao động của hệ thống treo cân bằng phụ thuộc cho phép làm rõ ảnh hưởng của biên dạng mặt đường đến chuyển vị, gia tốc các cầu và tải trọng động tác dụng lên bánh xe. Trên cơ sở đó, có thể lựa chọn bộ thông số hợp lý nhằm giảm rung xóc, hạn chế hiện tượng tách bánh, đồng thời duy trì khả năng bám đường và độ ổn định khi xe chạy trên đường xấu hoặc địa hình gồ ghề.

Trong bài báo này, trình bày kết quả xây dựng mô hình động lực học và mô phỏng dao động của ô tô quân sự 3 cầu có cụm cầu sau cân bằng phụ thuộc. Kết quả mô phỏng là cơ sở tham khảo cho các nghiên cứu tiếp theo trong đánh giá, kiểm tra tiêu chí bám đường và thiết lập mối quan hệ định lượng giữa các thông số hệ thống treo với đáp ứng dao động của xe trong các điều kiện kích thích mặt đường khác nhau.

2. Thiết lập mô hình nghiên cứu dao động hệ thống treo cân bằng phụ thuộc

2.1. Giả thiết

Tác giả đưa ra các giả thiết nhằm đơn giản hóa bài toán nhưng vẫn đánh giá được các thông số dao động cần thiết như sau:

- Đường cứng tuyệt đối, nằm ngang, có lớp phủ đường đồng nhất;
- Mô hình toán là mô hình phẳng, hình dạng ô tô đối xứng theo trục dọc xe;

- Coi vỏ xe là vật rắn và cứng tuyệt đối;

- Đặc tính của phần tử đàn hồi và giảm chấn là tuyến tính;

- Hệ thống treo, độ cứng của bánh xe bên phải và bên trái là như nhau;

- Xe chuyển động với vận tốc không đổi;

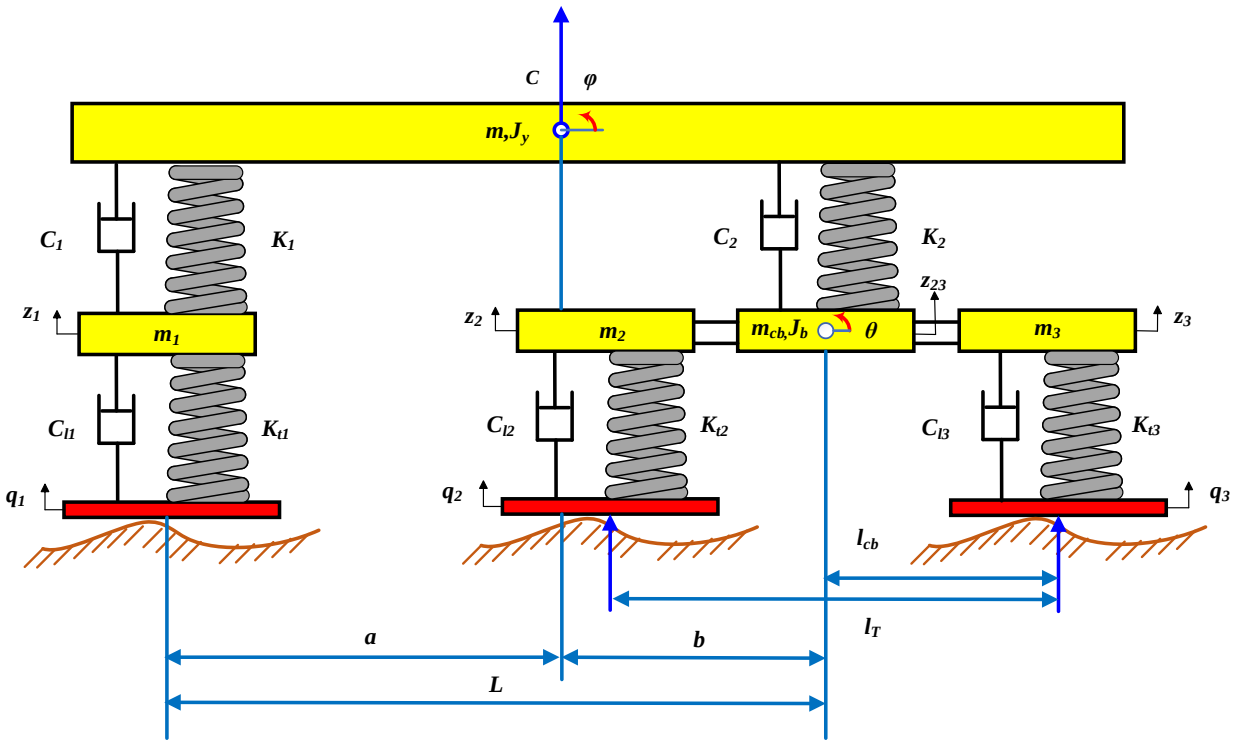
- Bánh xe tiếp xúc với đường là tiếp xúc điểm;

- Bỏ qua mô men quán tính của các cầu xe đối với trục quay của bánh xe.

- Các liên kết trong hệ thống cơ học là lý tưởng, không có độ trễ hay ma sát. Dao động của các phần tử trong hệ thống là tuyến tính nhằm đơn giản hóa mô hình toán học và phân tích động lực học.

2.2. Mô hình khảo sát

Từ các giả thiết ở trên, tác giả đưa ra mô hình tương đương khảo sát dao động như sau:



Hình 1. Mô hình động lực học dao động ½ dọc xe có hệ thống treo cân bằng phụ thuộc

Trong đó:

m - Khối lượng phần treo [kg];

m_1, m_2, m_3 - Khối lượng phần không treo cầu trước, cầu giữa và cầu sau [kg];

m_{cb} - Khối lượng các chi tiết cụm cân bằng hệ thống treo sau [kg];

$K_{1,2}$ - Độ cứng của phần tử đàn hồi cầu trước và cụm cầu sau [N/m];

$K_{t1,2,3}$ - Độ cứng của bánh xe các cầu trước, cầu giữa và cầu sau [N/m];

$C_{1,2}$- Hệ số cản của giảm chấn treo cầu trước và cụm cầu sau [N.s/m];

z - Chuyển dịch của khối lượng phần treo [m];

z_1, z_2, z_3 – Chuyển dịch của khối lượng phần không treo cầu trước, cầu giữa và cầu sau [m];

q_1, q_2, q_3 - Độ cao mấp mô mặt đường tại điểm tiếp xúc với bánh xe cầu cầu trước, cầu giữa và cầu sau [m];

J_y : Mô men quán tính khối lượng phần treo đối với trục dọc X đi qua trọng tâm phần treo [kg.m^2];

J_b : Mô men quán tính khối lượng phần treo quy dẫn về cụm cầu sau đối với trục dọc X đi qua trọng tâm cụm cầu sau [kg.m^2];

φ – Góc lắc dọc của phần khối lượng treo (thân xe) quanh trục ngang Y đi qua trọng tâm xe [rad]. Góc φ dương được chọn tương đương thân xe lắc sang trái (bên phải nâng, bên trái hạ).

θ – Góc lắc dọc của cụm cầu sau (khối lượng không treo) quanh trục ngang Y đi qua trọng tâm cụm cầu sau [rad]. Góc θ biểu diễn mức độ nghiêng của cầu: θ dương nghĩa là cụm cân bằng nghiêng khiến bánh trái hạ xuống, bánh phải nâng lên (cụm cầu sau lắc sang trái).

a, b - khoảng cách từ trọng tâm thân xe đến tâm phần khối lượng không treo cầu trước và cụm cầu sau [m];

$l_{cb} = \frac{l_T}{2}$ - khoảng cách từ trọng tâm trục cân bằng đến giữa và cầu sau [m];

Mô hình phẳng ½ dọc xe treo độc lập là mô hình có 5 bậc tự do ($z, \varphi, z_1, z_{23}, \theta$) được chỉ ra trên Hình 1 biểu diễn mô hình của cơ hệ ở vị trí tự nhiên. Hệ trục tọa độ cố định Oxyz dùng để khảo sát dao động của cơ hệ được chọn sao cho trục Oz thẳng đứng và đi qua trọng tâm của xe, hướng từ dưới lên ; trục Oy là trục nằm ngang (hướng vào trong xe); trục Ox là trục dọc trùng hướng theo chiều xe tiến.

2.3. Hệ phương trình vi phân của dao động

Dựa trên mô hình khảo sát đã phân tích, tác giả nhận thấy hệ dao động có 5 bậc tự do ($z, \varphi, z_1, z_{23}, \theta$) được chỉ ra trên hình 1 biểu diễn mô hình của cơ hệ ở vị trí tự nhiên. Hệ trục tọa độ cố định Oxyz dùng để khảo sát dao động của cơ hệ được chọn sao cho trục Oz thẳng đứng và đi qua trọng tâm của xe, hướng từ dưới lên; trục Ox là trùng với đường tâm của mấp mô mặt đường hướng theo chiều xe tiến (đi vào trong); trục Oy là trục nằm ngang.

Nhằm tối ưu hóa quá trình tính toán, tác giả lựa chọn phương pháp tách hệ và áp dụng nguyên lý Lagrăng loại 2.

Khối lượng tương đương và mô men quán tính tương đương cho cụm cầu sau:

$$m_{23}{}^* = m_2 + m_3 + m_{cb}$$

$$J_b{}^* = \frac{{l_T}^2}{4}.(m_2 + m_3) + m_{cb}.\rho_{yCB}^2$$

Sau quá trình tính toán, phân tích, tác giả đã xây dựng hệ phương trình vi phân bao gồm 5 phương trình sau:

Phương trình dao động thẳng đứng của thân xe:

$$m.\ddot{z} + K_1.(z + a.\varphi - z_1) + K_2.(z - b.\varphi - z_{23}) + C_1.(\dot{z} + a.\dot{\varphi} - \dot{z}_1) + C_2.(\dot{z} - b.\dot{\varphi} - \dot{z}_{23}) = 0 \quad (1)$$

Phương trình mô men lắc (quay) của thân xe:

$$J_y.\ddot{\varphi} + a.K_1.(z + a.\varphi - z_1) - b.K_2.(z - b.\varphi - z_{23}) + a.C_1.(\dot{z} + a.\dot{\varphi} - \dot{z}_1) - b.C_2.(\dot{z} - b.\dot{\varphi} - \dot{z}_{23}) = 0 \quad (2)$$

Phương trình chuyển động của khối lượng không treo cầu trước:

$$m_1.\ddot{z}_1 - K_1.(z + a.\varphi - z_1) - C_1.(\dot{z} + a.\dot{\varphi} - \dot{z}_1) + K_{t1}.(z_1 - q_1) = 0 \quad (3)$$

Phương trình chuyển động khối lượng không treo cầu sau:

$$m^*_{23}.\ddot{z}_{23} - K_2.(z - b.\varphi - z_{23}) - C_2.(\dot{z} - b.\dot{\varphi} - \dot{z}_{23}) + K_{t2}.\left(z_{23} + \frac{l_T}{2}.\theta - q_2\right) + K_{t3}\left(z_{23} - \frac{l_T}{2}.\theta - q_3\right) = 0 \quad (4)$$

Phương trình quay của cầu sau (đòn cân bằng):

$$J^*_b.\ddot{\theta}_b + \frac{l_T}{2}.\left[K_{t2}.\left(z_{23} + \frac{l_T}{2}.\theta - q_2\right) - K_{t3}.\left(z_{23} - \frac{l_T}{2}.\theta - q_3\right)\right] = 0 \quad (5)$$

Giải hệ phương trình trên bằng phương pháp không gian trạng thái:

$$\begin{cases} \dot{x} = A.x + B.u \\ y = C.x + D.u \end{cases}$$

Trong đó:
x - biến trạng thái;
u - biến điều khiển hay kích thích đầu vào;
y - thông số ra;
A - ma trận trạng thái;
B - ma trận quan hệ thống số vào với trạng thái;
C - ma trận quan hệ thông số ra với trạng thái;
D - ma trận truyền tín hiệu điều khiển;

Từ phương trình vi phân dao động của cơ hệ, xác định được:

- Biến trạng thái:

$$x = \left[Z, \varphi, z_1, z_{23}, \theta, \dot{Z}, \dot{\varphi}, \dot{z}_1, \dot{z}_{23}, \dot{\theta}\right]^T$$

- Kích thích đầu vào:

$$u = \left[q_1, q_2, q_3\right]^T$$

- Thống số ra:

$$y = \left[Z, \varphi, z_1, z_{23}, \theta, \dot{Z}, \dot{\varphi}, \dot{z}_1, \dot{z}_{23}, \dot{\theta}\right]^T$$

- Ma trận quan hệ thống số vào với biến trạng thái:

$$B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{K_{t1}}{m_1} & 0 & 0 \\ 0 & \frac{K_{t2}}{m^*_{23}} & \frac{K_{t3}}{m^*_{23}} \\ 0 & \frac{K_{t2}.l_{cb}}{J^*_b} & -\frac{K_{t3}.l_{cb}}{J^*_b} \end{bmatrix}$$

- Ma trận biến trạng thái:

$$A=\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
-\frac{K_1+K_2}{m} & -\frac{a.K_1-b.K_2}{m} & \frac{K_1}{m} & \frac{K_2}{m} & 0 & -\frac{C_1+C_2}{m} & -\frac{a.C_1-b.C_2}{m} & \frac{C_1}{m} & \frac{C_2}{m} & 0 \\
-\frac{a.K_1-b.K_2}{J_y} & -\frac{a^2.K_1+b^2.K_2}{J_y} & \frac{a.K_1}{J_y} & -\frac{b.K_2}{J_y} & 0 & -\frac{a.C_1-b.C_2}{J_y} & -\frac{a^2.C_1+b^2.C_2}{J_y} & \frac{a.C_1}{J_y} & -\frac{b.C_2}{J_y} & 0 \\
\frac{K_1}{m_1} & \frac{a.K_1}{m_1} & -\frac{K_1+K_{t1}}{m_1} & 0 & 0 & \frac{C_1}{m_1} & \frac{a.C_1}{m_1} & -\frac{C_1}{m_1} & 0 & 0 \\
\frac{K_2}{m^*_{23}} & -\frac{b.K_2}{m^*_{23}} & 0 & -\frac{K_2+K_{t2}+K_{t3}}{m^*_{23}} & -\frac{l_{cb}.(K_{t2}-K_{t3})}{m^*_{23}} & \frac{C_2}{m^*_{23}} & \frac{-b.C_2}{m^*_{23}} & 0 & -\frac{C_2}{m^*_{23}} & 0 \\
0 & 0 & 0 & -\frac{l_{cb}.(K_{t2}-K_{t3})}{J^*_b} & -\frac{l^2_{cb}.(K_{t2}+K_{t3})}{J^*_b} & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

- Ma trận quan hệ thống số ra với biến trạng thái:

$$C=\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}$$

- Ma trận truyền tín hiệu điều khiển:

$$D=\begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}$$

ISO 8608 mô tả dạng phổ gần đúng của mặt đường bằng hàm lũy thừa theo tần số không gian:

$$G_q(n)=G_q(n_0).\left(\frac{n}{n_0}\right)^{-w}$$

Trong đó:

$G_q(n)$ – Mật độ phổ công suất theo tần số không gian (đặc trưng độ nhám) [m^3]

n_0 – Tần số không gian tham chiếu (dùng trong phân loại)

n – Tần số không gian [m^{-1}]

w – Số mũ dốc phổ

Vì $G_q(n)$ đã biết theo ISO 8608, ta tạo q(x) bằng cách cộng chồng nhiều thành phần điều hòa có biên độ phù hợp với phổ và có pha ngẫu nhiên. Khi số thành phần đủ lớn, theo hiệu ứng

cộng chồng ngẫu nhiên, q(x) thu được thường có phân bố gần chuẩn và có tính "ngẫu nhiên" đúng bản chất mô phỏng. Ta sử dụng nguyên lý tổng hợp mặt đường ngẫu nhiên theo phổ bằng chuỗi Fourier:

$$q(x) = \sum_{k=1}^{n} \sqrt{2.G_q(n_k).\Delta n}.\cos(2.\pi.n_k.x + \varphi_k)$$

Trong đó:

x – Tọa độ trọng tâm theo chuyển động dọc của xe

n_k – các tần số không gian rời rạc

Δn – Bước rời rạc theo tần số không gian

φ_k – Pha ngẫu nhiên phân bố đều trên $[0, 2\pi]$

Biên dạng mặt đường ngẫu nhiên có dạng như Hình 2

Hình 2. Biên dạng mấp mô mặt đường tại cầu trước

3. Kết quả nghiên cứu

3.1. Thông số đầu vào

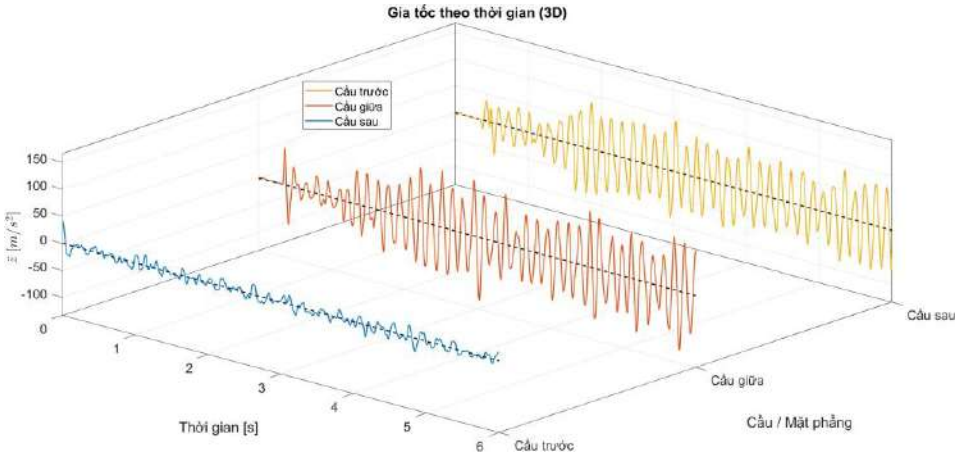
Tác giả sử dụng phần mềm MATLAB để giải hệ phương trình và sử dụng các thông số của xe ZIL-131 [4] như sau:

Bảng 1. Các thông số đầu vào của hệ thống treo cân bằng phụ thuộc trên xe ZIL-131

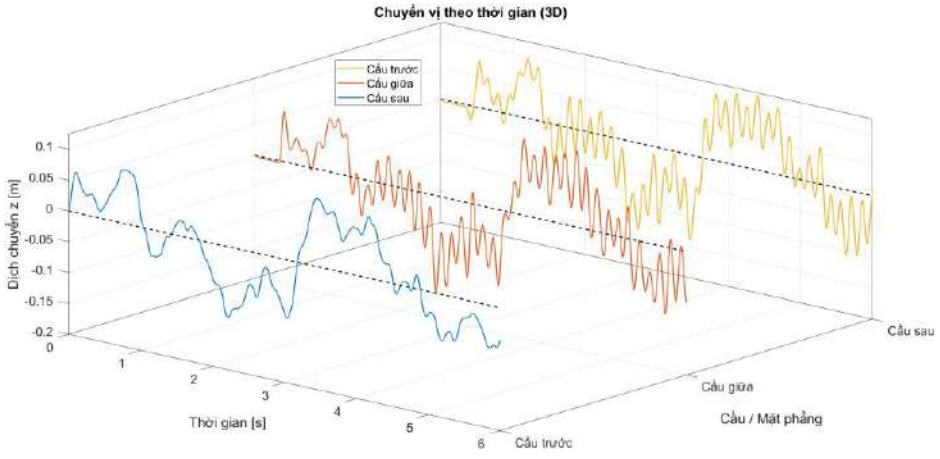
TT	Thông số	Ký hiệu	Đơn vị	Giá trị
1	Khối lượng phần treo	m	kg	8065
2	Chiều dài cơ sở	L	m	3,975
3	Khoảng cách từ tâm trục cầu trước đến trọng tâm phần treo	a	m	2,78

TT	Thông số	Ký hiệu	Đơn vị	Giá trị
4	Khoảng cách từ tâm trục cầu cân bằng đến trọng tâm phần treo	b	m	1,19
5	Khoảng cách giữa 2 cầu sau	l_{cb}	m	1,25
5	Khối lượng không treo phân bố cầu trước	m_1	kg	740
6	Khối lượng không treo phân bố cụm cầu sau	m_{23}	kg	1380
7	Khối lượng cụm cân bằng treo sau	m_{cb}	kg	128
8	Độ cứng hệ thống treo trước	K_1	kN/m	290
9	Độ cứng hệ thống treo sau	K_2	kN/m	1100
10	Hệ số giảm chấn của treo trước	C_1	kNs/m	87
11	Hệ số giảm chấn của treo sau	C_2	kNs/m	154
12	Độ cứng phần tử đàn hồi bánh xe cầu trước	K_{t1}	kN/m	865
13	Độ cứng phần tử đàn hồi bánh xe cầu sau	K_{t2}	kN/m	1530

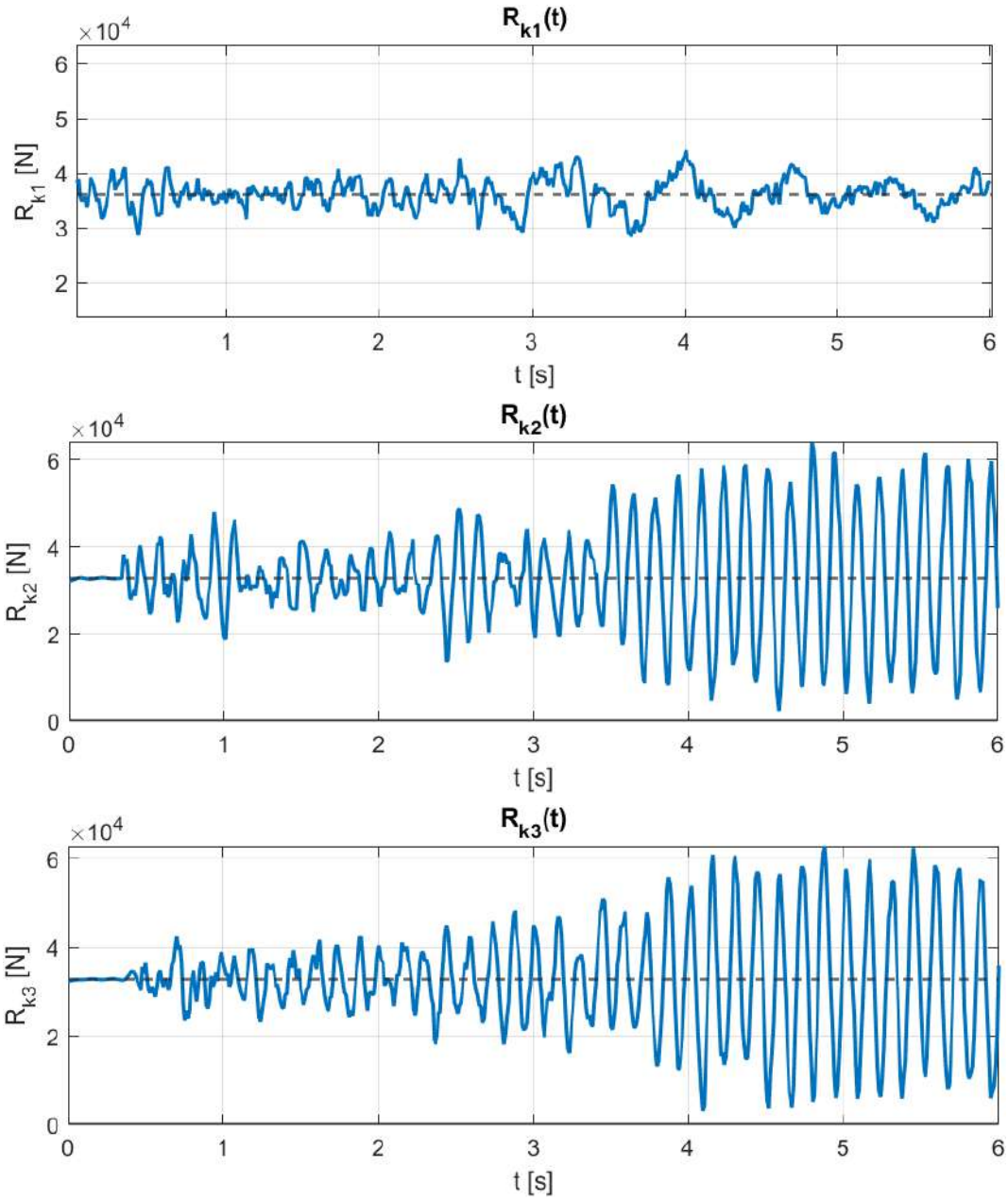
3.2. Kết quả mô phỏng



Hình 3. Gia tốc các cầu trong mặt phẳng dọc theo thời gian



Hình 4. Chuyển vị các cầu trong mặt phẳng dọc theo thời gian



Hình 5. Tải trọng động tác dụng lên các cầu

Các đồ thị gia tốc và chuyển vị theo thời gian (Hình 3, Hình 4) cho thấy đáp ứng dao động của các cầu có sự khác biệt rõ giữa cầu trước và cụm cầu sau. Cầu trước dao động tương đối ổn định, biên độ nhỏ hơn, trong khi cầu giữa và cầu sau xuất hiện dao động mạnh hơn, mật độ dao động dày hơn và nhiều đỉnh nhọn hơn khi chịu kích thích mặt đường ngẫu nhiên. Quy luật này phù hợp với đặc trưng của hệ treo cân bằng phụ thuộc, khi kích động mặt đường truyền tuần tự đến các cầu sau và đồng thời xuất hiện tương tác phân phối tải trong cụm cầu sau.

Về chuyển vị, biên độ dao động của các cầu đều mang tính ngẫu nhiên, không có dạng điều hòa thuần túy. Biên độ dao động của cụm cầu sau tăng rõ hơn ở giai đoạn sau của quá trình khảo sát, cho thấy mức độ nhạy cao hơn của cụm cầu sau đối với kích động mặt đường. Từ đặc điểm hình dạng đường cong chuyển vị có thể nhận thấy vận tốc dao động theo phương thẳng đứng của cụm cầu sau cũng có xu hướng lớn hơn cầu trước, phù hợp với kết quả gia tốc.

Trên phương diện an toàn chuyển động, các đồ thị tải trọng động $R_{kt_1}, R_{kt_2}, R_{kt_3}$ (Hình 5) cho thấy tải trọng tại cầu trước dao động quanh giá trị tĩnh với biên độ tương đối nhỏ, còn tải trọng tại cụm cầu sau dao động mạnh hơn, đặc biệt ở nửa sau thời gian khảo sát. Tuy nhiên, trong toàn bộ khoảng mô phỏng, tải trọng tiếp xúc tại các cầu nhìn chung vẫn duy trì giá trị dương, chưa xuất hiện hiện tượng tách bánh theo điều kiện $R_{kt} \leq 0$. Kết quả này cho thấy, với bộ tham số đã lựa chọn, hệ thống treo vẫn bảo đảm yêu cầu an toàn tối thiểu về duy trì tiếp xúc bánh xe – mặt đường.

So với xe quân sự 2 cầu trong cùng điều kiện kích động, cấu hình 3 cầu với treo cân bằng phụ thuộc thể hiện ưu điểm về khả năng phân phối tải ở cụm cầu sau, góp phần giảm xu hướng mất tiếp xúc hoàn toàn tại một cầu. Tuy vậy, biên độ dao động tải trọng của cụm cầu sau vẫn lớn, cho thấy cần tiếp tục nghiên cứu tối ưu tham số đàn hồi – giảm chấn để giảm dao động tải trọng động, nâng cao độ ổn định chuyển động trong các điều kiện làm việc khắc nghiệt hơn

4. Kết luận

Trong nghiên cứu này, mô hình dao động ½ dọc xe của xe ZIL-131 được xây dựng nhằm nghiên cứu động lực học của hệ thống treo cân bằng phụ thuộc và đánh giá an toàn chuyển động trong điều kiện kích thích mặt đường ngẫu nhiên. Trên cơ sở thiết lập hệ phương trình động lực học và triển khai mô phỏng theo thời gian, kết quả gia tốc và chuyển vị của các cầu thể hiện quy luật dao động phù hợp với đặc tính đàn hồi – giảm chấn của hệ treo. Kết quả mô phỏng tải trọng thẳng đứng tác dụng lên các cầu cho thấy lực tiếp xúc bánh xe – mặt đường dao động quanh giá trị tĩnh và nhìn chung duy trì dương trong toàn bộ khoảng khảo sát, qua đó cho thấy hệ thống treo đảm bảo điều kiện an toàn cơ bản về bám đường, hạn chế nguy cơ tách bánh.

Biên độ chuyển vị và gia tốc của thân xe nhỏ hơn đáng kể so với dao động của cụm cầu, qua đó khẳng định vai trò của hệ treo trong việc cách ly rung động và duy trì ổn định chuyển động. Đồng thời, sự chênh lệch giữa cầu trước và cụm cầu sau thể hiện rõ đặc trưng phân bố tải và đặc điểm truyền lực của cơ cấu cân bằng, nhất là khi các bánh sau lần lượt đi qua cùng một dạng mấp mô theo trục dọc.

Từ các kết quả trên có thể khẳng định rằng, trong điều kiện kích thích mặt đường ngẫu nhiên theo ISO 8608, an toàn chuyển động và khả năng bám đường của ô tô quân sự ba cầu có hệ thống treo cân bằng phụ thuộc bị ảnh hưởng đáng kể. Các chỉ tiêu như tải trọng thẳng đứng, gia tốc và biên độ chuyển dịch tương đối bánh xe và mặt đường là những chỉ tiêu phù hợp và có ý nghĩa thực tiễn trong đánh giá an toàn chuyển động của loại xe này.

Các kết quả thu được trong nghiên cứu là cơ sở khoa học để tiếp tục phát triển các hướng khảo sát sâu hơn đối với xe quân sự nhiều cầu. Trong các bước tiếp theo, cần mở rộng mô phỏng theo các kịch bản làm việc khắc nghiệt hơn như tăng vận tốc xe, tăng mức độ gồ ghề mặt đường, thay đổi biên độ kích động để đánh giá đầy đủ hơn nguy cơ tách bánh, số lần tách bánh và thời gian mất tiếp xúc. Đồng thời, cần thực hiện bài toán tối ưu tham số hệ treo theo hướng giảm dao động tải trọng động nhưng vẫn bảo đảm yêu cầu êm dịu và khả năng vượt địa hình.

Tài liệu tham khảo

1. N. P. Hiểu, *Lý thuyết ô tô quân sự*. Hà Nội: Nhà xuất bản Quân đội nhân dân, 2002.
2. V. Đ. Lập, *Dao động ô tô*. Hà Nội: Nhà xuất bản Quân đội nhân dân, 2011.
3. V. N. Minh, N. Đ. Quý, and N. T. Sinh, "Động lực học dao động của ô tô có kể tới hiện tượng tách bánh trên mô hình ½ ngang xe," Học viện Kỹ thuật Quân sự, 2025.

4. Vũ Đức Lập (2004), Sổ tay tra cứu tính năng kỹ thuật ôtô, Học viện Kỹ thuật quân sự, Hà Nội.

Dynamic Study of the Dependent Balanced Suspension System of a Three-Axle Military Vehicle

Abstract: The paper investigates several factors affecting the vibration behavior of the dependent balanced suspension system of a three-axle military vehicle. A longitudinal-plane vibration model is developed, and the equations of motion are formulated using the Lagrange method. Numerical simulations are carried out in MATLAB to examine the influence of different operating conditions. The obtained results describe the patterns of vibration velocity and acceleration, which are then used to evaluate the vehicle's safety when traveling on uneven roads. The study provides a basis for proposing solutions to improve driving safety under rough terrain and harsh weather conditions.

Keyword: Dependent balanced suspension system, State-space method, Lagrange's equations of the second kind

Mô phỏng động lực học chuyển động tăng tốc của xe bọc thép bánh lốp BTR-60PB bằng Simscape Driveline

Trần Đức Lương[1], Nguyễn Văn Trà[1]

[1]Học viện Kỹ thuật quân sự

Email: paruruboy@gmail.com; Tel: 0773069101

Tóm tắt

Bài báo áp dụng Simscape Driveline để mô phỏng thành công động lực học chuyển động thẳng của xe bọc thép bánh lốp BTR-60PB với mô hình truyền lực đầy đủ, chia theo các cụm hệ thống nên thuận tiện cho việc thiết lập, hiệu chỉnh và khảo sát nhiều tham số; kết quả cho thấy khi điều kiện mặt đường thay đổi (*hệ số cản lăn f và hệ số bám φ thay đổi*) thì vận tốc, gia tốc và quãng đường tăng tốc của xe cũng thay đổi rõ rệt, đặc biệt khi bám giảm hoặc cản lăn tăng sẽ làm suy giảm khả năng cơ động của xe, qua đó bài báo có thể dùng làm tài liệu hữu ích phục vụ nghiên cứu, đánh giá kỹ thuật và tối ưu khai thác xe trong thực tế.

Từ khóa: BTR-60PB; động lực học; hệ số cản lăn; hệ số bám; Simscape Driveline.

1. Đặt vấn đề

Xe bọc thép bánh lốp BTR-60PB là xe chiến đấu bộ binh có công thức bánh xe 8x8, một trong những phương tiện chiến đấu chủ lực, có ưu điểm về tính cơ động, tốc độ hành quân và khả năng vượt địa hình cao và khả năng lội nước. Sau nhiều năm khai thác trong điều kiện khí hậu, địa hình phức tạp, đặc tính động lực học của xe có xu hướng suy giảm, đặc biệt trong quá trình tăng tốc và vượt chướng ngại.

Trong thực tế khai thác, khả năng cơ động của xe phụ thuộc nhiều vào điều kiện mặt đường, trong đó hai yếu tố có ảnh hưởng trực tiếp là hệ số cản lăn f và hệ số bám φ. Khi xe làm việc trên các loại đường khác nhau (*đường nhựa tốt, đường đất, đường lầy, mặt đường ướt hoặc đường có tính không đồng nhất giữa các vệt bánh*), các giá trị f và φ thay đổi đáng kể, làm thay đổi lực kéo hữu ích, lực cản lăn, vận tốc, gia tốc và quãng đường tăng tốc của xe. Vì vậy, việc đánh giá sự ảnh hưởng của các thông số này là cần thiết để phản ánh sát hơn điều kiện khai thác thực tế và phục vụ công tác bảo đảm kỹ thuật.

Trong bối cảnh công nghệ tính toán và phần mềm mô phỏng phát triển, việc mô hình hóa và mô phỏng động lực học phương tiện bằng Matlab/Simulink, đặc biệt với nền tảng Simscape Driveline, cho phép đánh giá khách quan đặc tính tăng tốc mà không cần thực hiện quá nhiều thử nghiệm thực địa, qua đó tiết kiệm thời gian, chi phí và giảm rủi ro thử nghiệm. Simscape Driveline có ưu điểm là cung cấp thư viện khối chuyên dụng cho hệ truyền lực, thuận lợi trong xây dựng mô hình theo cấu trúc vật lý, đồng thời hỗ trợ khảo sát nhiều kịch bản làm việc khác nhau. Tuy nhiên, nhược điểm của phương pháp này là kết quả mô phỏng phụ thuộc mạnh vào độ chính xác của bộ thông số đầu vào và giả thiết mô hình; với các điều kiện địa hình phức tạp ngoài đường tốt, việc mô tả tương tác lốp - đường vẫn cần hiệu chỉnh và kiểm chứng bằng số liệu thực nghiệm để bảo đảm độ tin cậy.

Mục tiêu của bài báo là xây dựng và khai thác mô hình mô phỏng để đánh giá chất lượng tăng tốc của xe thông qua các chỉ tiêu đặc trưng như vận tốc, gia tốc và quãng đường tăng tốc trong các điều kiện đường khác nhau; từ đó làm rõ mức độ suy giảm khả năng cơ động khi cản lăn tăng hoặc bám giảm đột ngột. Kết quả nghiên cứu có thể được sử dụng làm cơ sở cho đánh giá trạng thái kỹ thuật xe, hỗ trợ lựa chọn chế độ khai thác phù hợp theo điều kiện mặt đường, phục vụ huấn luyện - sử dụng, tối ưu hóa bảo dưỡng và nâng cao hiệu quả bảo đảm kỹ thuật trong thực tế. Ngoài ra, cách tiếp cận mô hình hóa này không chỉ áp dụng cho BTR-60PB mà còn có thể mở rộng cho các xe quân sự bánh lốp khác có kết cấu truyền lực tương tự (*xe*

6×6, 8×8) hoặc các xe vận tải chuyên dụng, khi được hiệu chỉnh lại bộ thông số kết cấu và thông số làm việc của hệ truyền lực, lốp và thân xe

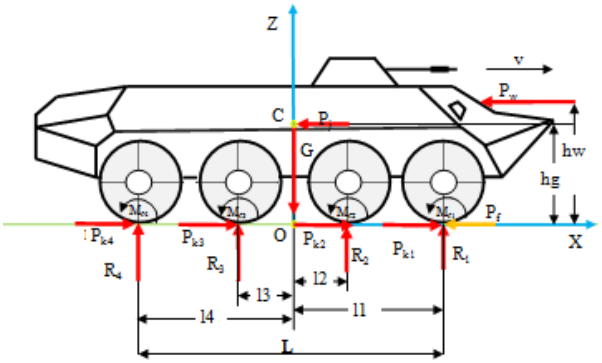
2. Mô phỏng quá trình tăng tốc xe bọc thép bánh lốp BTR-60PB

Để mô phỏng quá trình tăng tốc xe bọc thép bánh lốp BTR-60PB, bài báo sử dụng Simscape Driveline trên nguyên tắc xây dựng mô hình theo các khối của hệ truyền lực và chuyển động xe (*động cơ, ly hợp, hộp số, hộp số phân phối, vi sai, thân xe, hộp giảm tốc, lốp xe*). Trên cơ sở đó, quan hệ động lực học của từng khối được mô tả bằng phương trình toán học và liên kết thành mô hình tổng thể; để thiết lập mô hình và phương trình mô tả một cách hợp lý cần có một số giả thiết như sau:

2.1. Các giả thiết

- Các lực và phản lực đặt tại điểm tiếp xúc giữa bánh xe và mặt đường;
- Mọi sự dời tâm của lốp và ảnh hưởng của phản lực pháp tuyến tại bánh xe được quy đổi qua mô men cản lăn;
- Môi trường không gió, người lái tiêu chuẩn;
- Bỏ qua ảnh hưởng của hệ thống lái;
- Chuyển động trên đường bằng, nằm ngang;
- Mô hình khi thiết lập ở dạng mô hình phẳng, coi kết cấu là đối xứng qua mặt phẳng dọc giữa xe;
- Xe không kéo móoc.

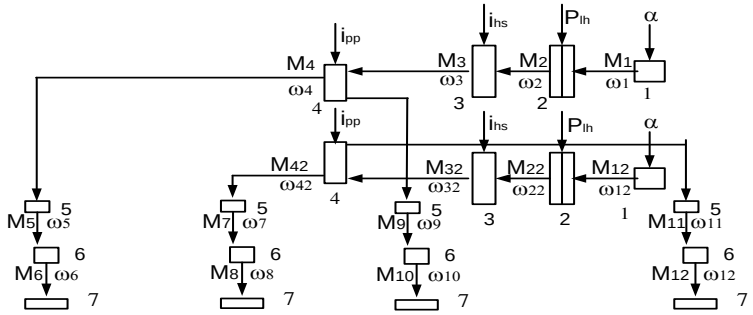
Trên cơ sở các giả thiết ta có sơ đồ lực và mô men tác dụng lên xe được thể hiện trên hình 1:



Hình 1. Sơ đồ lực và mô men tác dụng lên xe thiết giáp BTR-60PB

G: Trọng lượng ô tô (N); P_w: Lực cản không khí (N); M_{f1}, M_{f2}, M_{f3}, M_{f4}: Mô men cản lăn (N.m); P_j: Lực quán tính (N); R_1, R_2, R_3, R_4: Phản lực pháp tuyến của đường (N); P_{k1}, P_{k2}, P_{k3}, P_{k4}: Lực kéo tiếp tuyến (N); l_1, l_2, l_3, l_4: Khoảng cách từ trọng tâm đến các bánh xe (m); P_f: Lực cản lăn (N).

Ta xây dựng mô hình HTTL khi khảo sát xe ở dạng mô hình một vết bánh xe thể hiện trên hình 2:

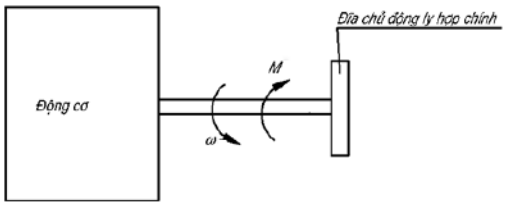


Hình 2. Mô hình hệ thống truyền lực xe bốn cầu chủ động BTR-60PB

1: Động cơ; 2: Ly hợp; 3: hộp số; 4: hộp số phân phối; 5: Truyền lực chính cầu xe; 6: Giảm tốc bánh xe, 7: Bánh xe

2.2. Xây dựng mô hình mô phỏng quá trình tăng tốc xe bọc thép bánh lốp BTR-60PB

2.2.1. Mô hình động cơ



Hình 3. Mô hình khối động cơ

Động cơ xăng GAZ-49B được mô tả bằng mô men phụ thuộc tốc độ kiểu đa thức bậc ba (dạng Layderman) theo tốc độ quay n được tính theo biểu thức 1:

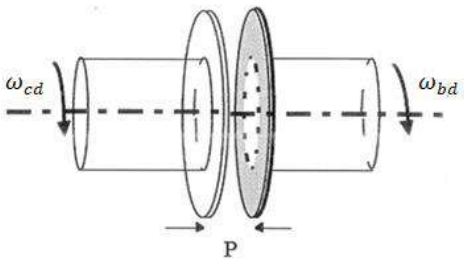
$$M_e = M_{eN}\left[a + b\left(\frac{n_e}{n_N}\right) - c\left(\frac{n_e}{n_N}\right)^2\right] \quad (1)$$

Trong đó: M_{eN}: Mô men xoắn ứng với công suất hữu ích cực đại của động cơ (N.m); M_e : Mô men xoắn ứng với tốc độ quay bất kỳ của trục khuỷu trên đồ thị đặc tính ngoài (N.m); n_N: Tốc độ quay của trục khuỷu động cơ tương ứng với công suất cực đại (rpm); n_e: Tốc độ quay của trục khuỷu động cơ tương ứng với công suất N_e và mô men xoắn M_e (rpm); a, b, c: Các hệ số thực nghiệm của động cơ. Với các loại động cơ có các thông số kỹ thuật chính như N_{emax}, M_{emax} và số vòng quay trục khuỷu tại các vị trí đó.

Công suất của động cơ được tính theo biểu thức 2:

$$N_e = \frac{9{,}81.M_e.\pi.n_e}{1000.30} \quad (2)$$

2.2.2. Mô hình ly hợp ma sát



Hình 4. Mô hình khối ly hợp

Tốc độ khâu chủ động được tính theo biểu thức 3:

$$J_{cd}.\frac{d\omega_{cd}}{dt} = M_{cd} - M_c \quad (3)$$

Trong đó: J_{cd}: Mô men quán tính của phần chủ động của ly hợp (kg.m^2); ω_{cd}: Tốc độ góc của phần chủ động (rad/s); M_{cd}: Mô men của phần chủ động ly hợp (N.m); M_c: Mô men cản chất lên phần chủ động (N.m).

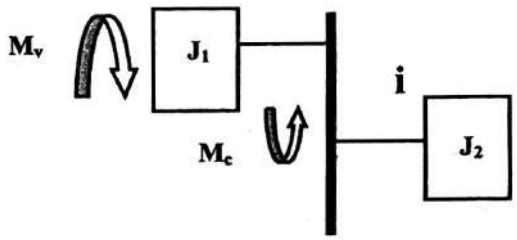
Lực mô tả hoạt động của ly hợp được tính theo biểu thức 4 và 5:

$$M_f = P.\mu.R_c.z \quad (4)$$

$$M_{bd} = \begin{cases} M_{cd}, M_{cd} \le M_f \\ M_f \ , M_{cd} > M_f \end{cases} \quad (5)$$

Trong đó: M_f: Mô men ma sát của ly hợp (N.m); P: Lực ép lên bề mặt đĩa ma sát (N); R_c: Bán kính tác dụng quy dẫn của lực ép (m); μ: Hệ số ma sát của bề mặt đĩa ma sát; z: Số bề mặt đĩa ma sát của ly hợp; M_{bd}: Mô men của phần bị động của ly hợp (N.m).

2.2.3. Mô hình hộp số, hộp số phân phối



Hình 5. Mô hình khối hộp số

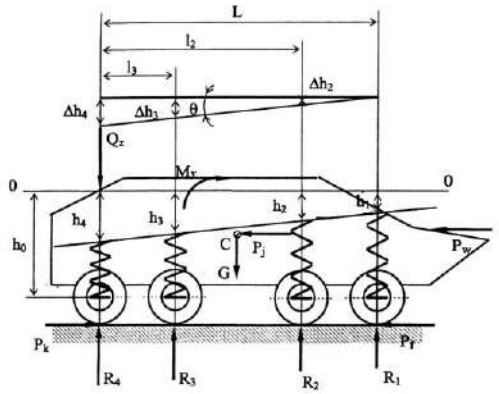
Áp dụng nguyên lý D'lembert cho hệ 2 vật quay quanh hai trục song song với liên kết bánh răng được tính theo biểu thức 6 và 7:

$$M_v = J_1.\varepsilon_1 + \frac{J_2.\varepsilon_2}{i} + M_c \quad (6)$$

$$i = \frac{\varepsilon_1}{\varepsilon_2} \quad (7)$$

Trong đó: M_v: Mô men vào hệ thống (N.m); M_c: Mô men cản trong hệ thống (N.m); J_1, J_2: Mô men quán tính các bánh răng (kg.m^2); i: Tỷ số truyền của cặp bánh răng ăn khớp; ε_1, ε_2: Gia tốc góc của bánh răng ăn khớp (rad/s^2).

2.2.4. Mô hình thân xe



Hình 6. Mô hình khối thân xe

Xe BTR - 60PB là xe nhiều cầu, số lượng phản lực pháp tuyến cần xác định nhiều hơn tổng số phương trình tĩnh học cần thiết đã được thành lập, do vậy để tính toán phản lực pháp tuyến phải thành lập thêm các phương trình đàn hồi của hệ treo của tất cả các bánh xe.

Các phản lực tiếp tuyến của đường tác dụng lên bánh xe được tính theo biểu thức 8, 9 và 10:

$$G = R_1 + R_2 + R_3 + R_4 \tag{8}$$

$$\sum_{i=1}^{3} R_i . l_i = G.a - (P_j + P_w).h \tag{9}$$

$$P_w = K.F.v^2 \tag{10}$$

Trong đó: R_1, R_2, R_3, R_4: Phản lực pháp tuyến của đường (N); G: Trọng lượng xe không tải (N); a, b: Khoảng cách từ trọng tâm xe đến tâm cầu trước và cầu sau (m); P_j, P_w: Lực quán tính và lực cản không khí (N); v: Vận tốc chuyển động của xe (m/s); K: Hệ số cản không khí; F: Diện tích cản chính diện của xe (m^2).

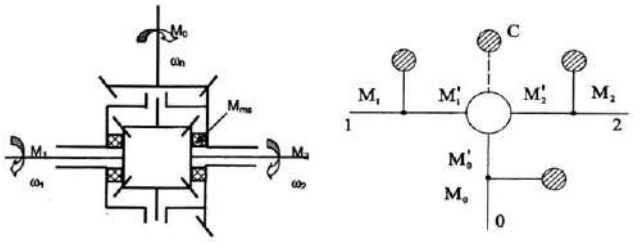
Đàn hồi của hệ treo như sau được tính theo biểu thức 11 và 12:

$$R_i = c_i . h_i \quad (i = 1..4) \tag{11}$$

$$h_i = h_1 + (L - l_i).tg\theta \tag{12}$$

Trong đó: c_i: Độ cứng của hệ treo của bánh xe trên cầu thứ I; h_i: Chuyển dịch thẳng đứng của bánh xe trên cầu thứ i (m); L: Chiều dài cơ sở của xe (m); l_i: Khoảng cách tính toán (m); θ: Góc nghiêng dọc của thân xe (độ).

2.2.5. Mô hình khối vi sai



Hình 7. Mô hình khối vi sai

Mô men quán tính quy dẫn của giá hành tinh - vi sai được tính theo biểu thức 13:

$$J_0 = J_0' + m_{ht}.R_{ht}^2 \quad (13)$$

Trong đó: J_o: Mô men quán tính quy dẫn của giá hành tinh - vi sai (kg.m^2); J_0': Mô men quán tính của giá hành tinh so với trục quay chính (kg.m^2); m_{ht}: Khối lượng của bánh răng hành tinh (kg); R_{ht}: Khoảng cách giữa trục quay chính và bánh răng hành tinh (m).

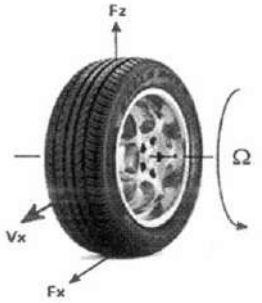
Phương trình 14 và 15 là phương trình cân bằng năng lượng của nút vi sai:

$$\begin{cases} M_0 + M_0' - J_0.\dfrac{d\omega_0}{dt} = 0 \\ M_1 + M_1' - J_1.\dfrac{d\omega_1}{dt} = 0 \\ M_2 + M_2' - J_2.\dfrac{d\omega_2}{dt} = 0 \end{cases} \quad (14)$$

$$\begin{cases} \omega_0.M_0' + \omega_1.M_1' + \omega_2.M_2' + J_c.\omega_c'.\dfrac{d\omega_c'}{dt} = 0 \\ M_0' + M_1' + M_2' = 0 \end{cases} \quad (15)$$

Trong đó: M_0, ω_0: Mô men và vận tốc góc trên vỏ vi sai (N.m, rad/s); M_1, ω_1, M_2, ω_2: Mô men và vận tốc góc tương ứng của bán trục (N.m, rad/s); J_c: Mô men quán tính của bánh răng hành tinh (kg.m^2); ω_c: Vận tốc góc của bánh răng hành tinh (rad/s).

2.2.6. Xây dựng mô hình lốp có tương tác với đường



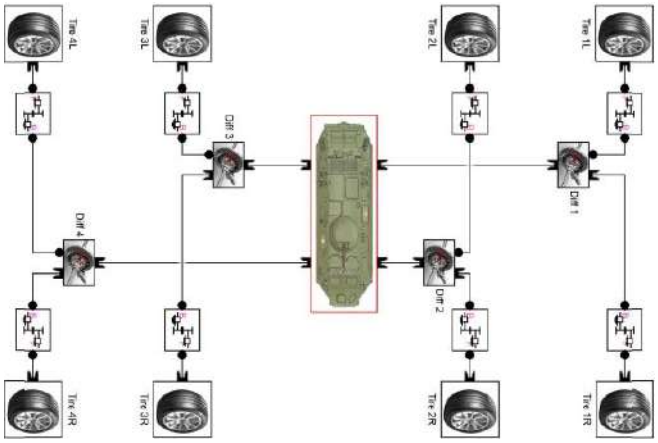
Hình 8. Mô hình khối bánh xe

Ta sử dụng công thức số 16 theo thực nghiệm Magic Fomular của Pacejika:

$$F_x = D_x \cdot \sin\left[C_x \cdot \arctan\left(B_x \cdot k - E_x\left(B_x \cdot k - \arctan(B_x \cdot k)\right)\right)\right] \quad (16)$$

Trong đó: B_x, C_x, D_x, E_x là các tham số quan trọng của mô hình Pacejka, đặc trưng cho dạng đường cong lực bên - góc trượt, phụ thuộc vào tải trọng tác dụng lên lốp, góc trượt bên, hệ số ma sát của mặt đường và trạng thái trượt của bánh xe. Các tham số này được xác định dựa trên kết quả thực nghiệm hoặc dựa trên các hệ số thực nghiệm trong tài liệu tiêu chuẩn của mô hình.

Mô hình mô phỏng BTR-60PB trong Simscape Driveline gồm các khối chính: động cơ, ly hợp, hộp số, hộp số phân phối, vi sai, hộp giảm tốc bánh xe, thân xe và lốp có tương tác với đường. Các khối này được nối đúng với sơ đồ truyền lực thực tế của BTR-60PB.



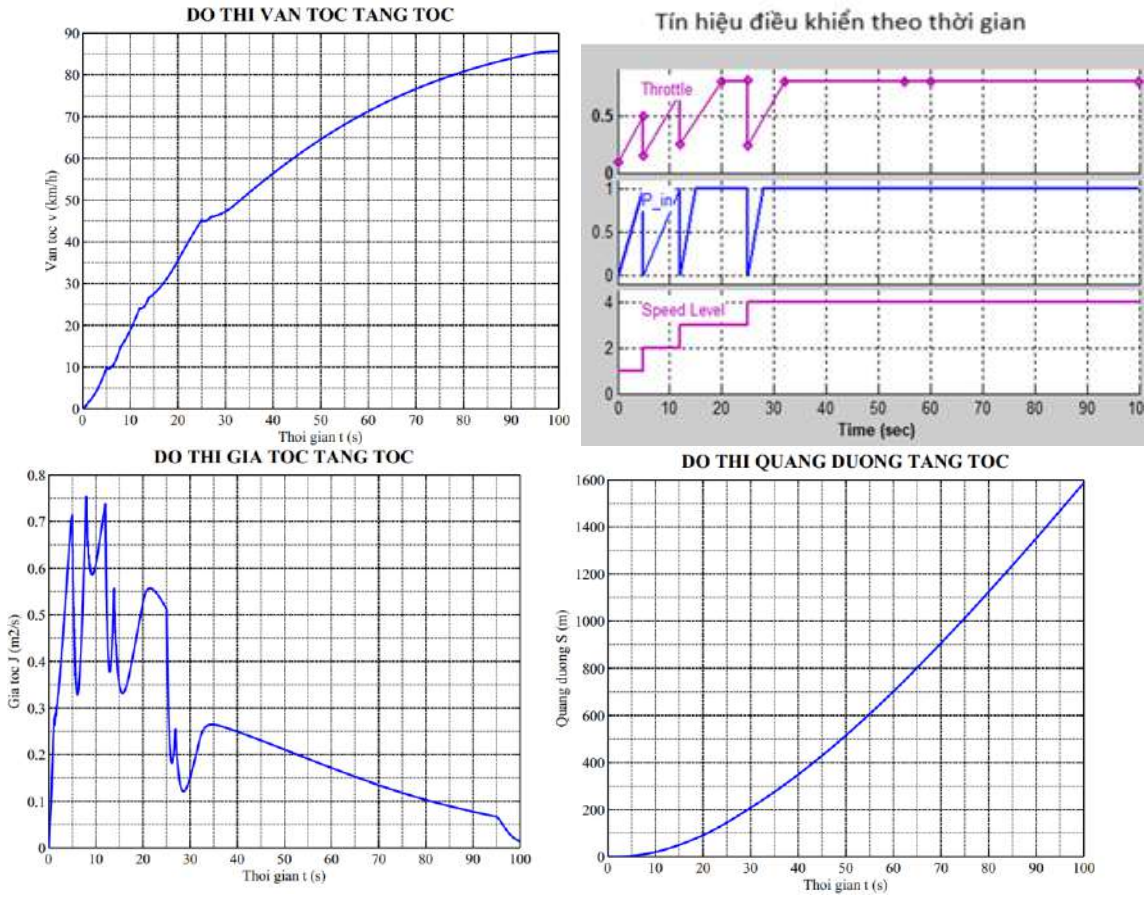
Hình 9. Bố trí hệ thống truyền lực xe BTR-60PB

3. Kết quả mô phỏng

Bảng 1. Các thông số đầu vào của chương trình mô phỏng

STT	Thông số	Xe BTR - 60PB
1	Chiều rộng toàn bộ (m)	2,37
2	Chiều cao toàn bộ (m)	2,055
3	Chiều dài cơ sở (m)	4,24
4	Chiều cao trọng tâm (m)	1,05
5	Trọng lượng xe không tải (KG)	10220
6	Kích thước lốp (inch)	13,00 - 18
7	Loại động cơ Công suất lớn nhất (ml) Số vòng quay ứng với công suất lớn nhất (vg/ph) Mô men xoắn lớn nhất (KGm) Số vòng quay ứng với mô men xoắn lớn nhất (vg/ph)	Xăng 4 kỳ, 6 xi lanh, 1 hàng 90 x 2 3600 22,5 x 2 1600 - 1800
8	Loại ly hợp Bán kính tác dụng trung bình (m) Lực ép trên bề mặt đĩa ma sát (N) Hệ số ma sát lớn nhất	Ma sát khô, 1 đĩa, dẫn động cơ khí 0,15 7800 1,1
9	Tỷ số truyền hộp số	6,55 - 3,09 - 1,71 - 1. L:7,77
10	Hộp số phân phối	2 cấp: 1,00 - 1,982
11	Truyền lực chính	Đơn; Hypoit; 1,843
12	Truyền động bánh xe	Bánh răng trụ; 4,55

Trường hợp 1: Xe chuyển động thẳng trên đường nhựa có hệ số bám và hệ số cản lăn tốt và không đổi. ($f = 0.015$, $\varphi = 0.8$).



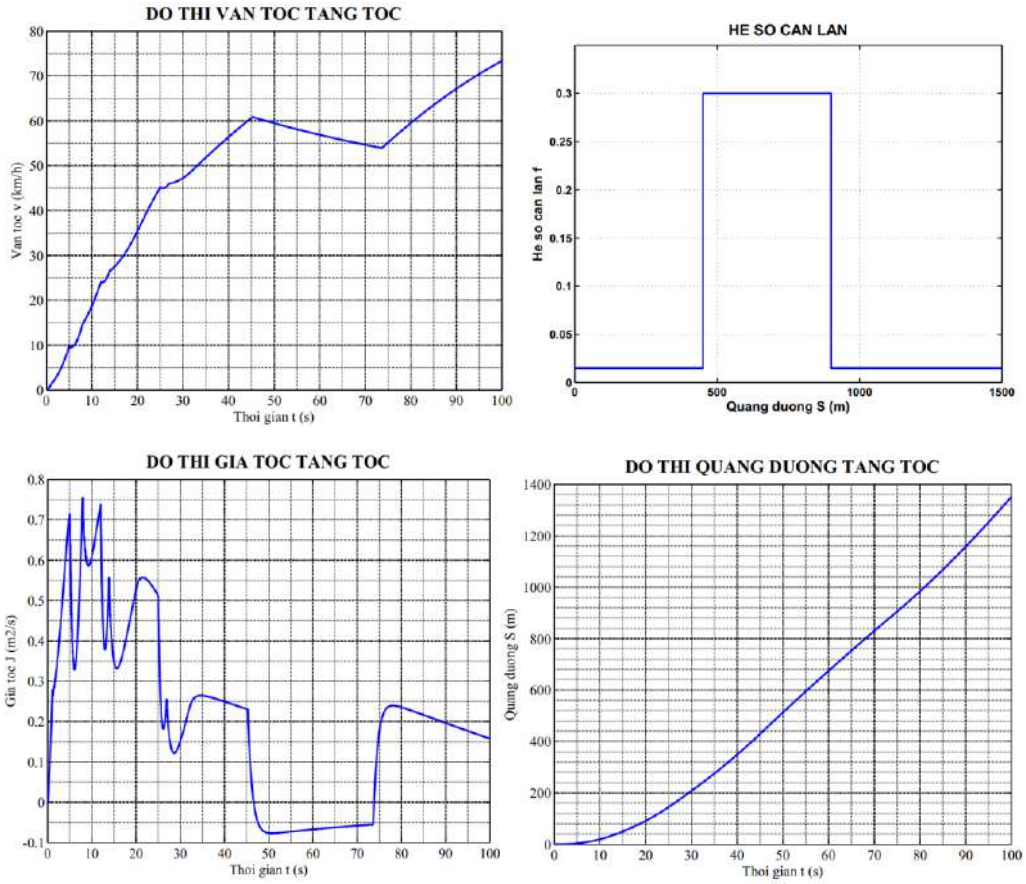
Hình 10. Kết quả mô phỏng trường hợp 1

Đồ thị vận tốc cho thấy xe tăng tốc tương đối nhanh ở giai đoạn đầu: sau khoảng 20-30 s đã đạt xấp xỉ 50-60 km/h, sau đó tiếp tục tăng dần đến xấp xỉ 85-90 km/h ở cuối khoảng thời gian mô phỏng 100 s. Độ cong của đường v(t) giảm dần theo thời gian, phản ánh đúng quy luật: khi vận tốc tăng lên thì lực cản khí động và tổn thất trong hệ thống truyền lực tăng, gia tốc giảm.

Đồ thị quãng đường S(t) có dạng đường cong lồi hướng lên, thể hiện quãng đường tích lũy tăng theo quy luật phi tuyến: ban đầu tăng chậm, sau đó tăng nhanh hơn khi vận tốc lớn. Sau 100 s, xe đi được khoảng 1,5 km.

Đồ thị gia tốc J(t) cho thấy: ở giai đoạn xuất phát, gia tốc tăng rất nhanh và đạt các giá trị cực đại khoảng 0,7-0,75 m/s²; các đỉnh dao động lớn trong 10-25 s đầu là do quá trình vào số/chuyển số và sự thay đổi mô men động cơ theo tốc độ, sau đó gia tốc giảm dần về khoảng 0,1-0,2 m/s² khi xe đạt vận tốc cao. Quy luật này phù hợp với tính chất của động cơ đốt trong và hệ thống truyền lực cơ khí: ở số thấp, lực kéo lớn nên gia tốc cao, khi chuyển sang số cao, lực kéo giảm, lực cản khí động tăng nên gia tốc giảm.

Trường hợp 2: Xe chuyển động thẳng trên đường có hệ số bám không đổi và hệ số cản lăn thay đổi theo hàm bậc thang do đi vào đoạn đường cát lún. (f = 0.015 - 0.3 - 0.015, φ = 0.8).



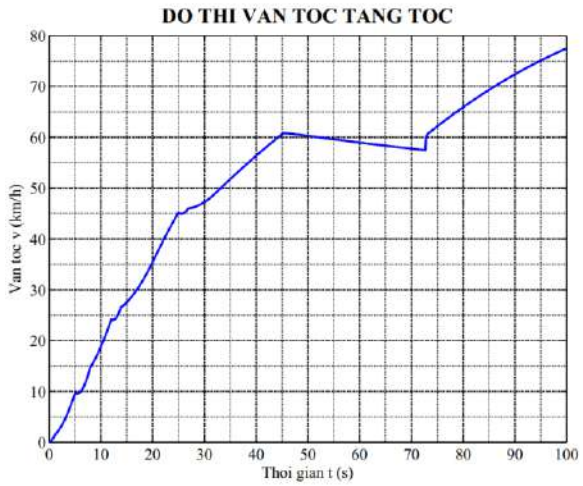
Hình 11. Kết quả mô phỏng trường hợp 2

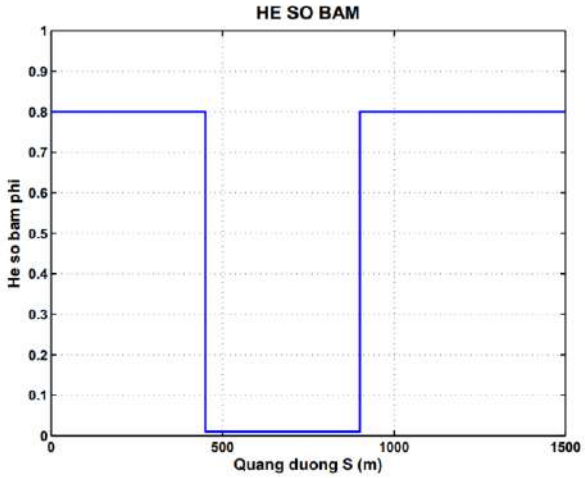
Đồ thị vận tốc v(t) ở giai đoạn đầu, khi hệ số cản lăn còn nhỏ, vận tốc tăng khá đều, xe đạt khoảng 60-61 km/h vào khoảng 40 s. Tại thời điểm hệ số cản lăn tăng theo hàm bậc thang, lực cản lăn tăng đột ngột nên vận tốc giảm xuống rõ rệt. Sau khi xe "vượt qua" đoạn đường có cản lăn lớn, vận tốc mới tăng trở lại, nhưng độ dốc v(t) nhỏ hơn so với giai đoạn đầu, đến cuối 100 s chỉ đạt khoảng 74-75 km/h, thấp hơn nhiều so với đường tốt.

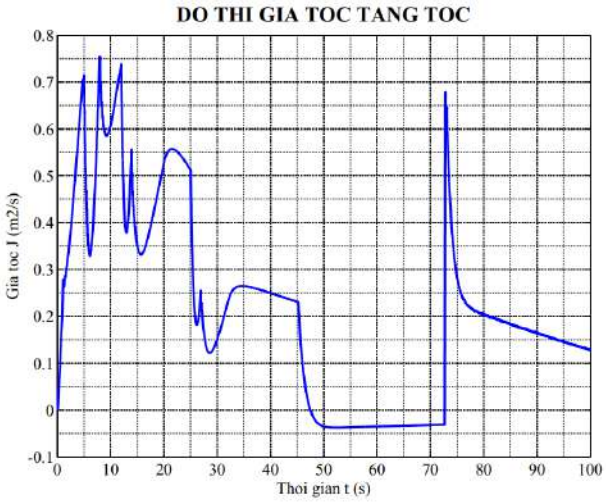
Đồ thị quãng đường S(t) quãng đường tăng liên tục theo thời gian với dạng đường cong lồi hướng lên. Tuy nhiên, sau thời điểm hệ số cản lăn tăng, độ cong của S(t) giảm đi thấy rõ so với giai đoạn đầu, thể hiện vận tốc trung bình trong khoảng giữa nhỏ hơn do có đoạn vận tốc bị giảm xuống. Kết thúc 100 s, quãng đường đạt xấp xỉ 1,3-1,35 km

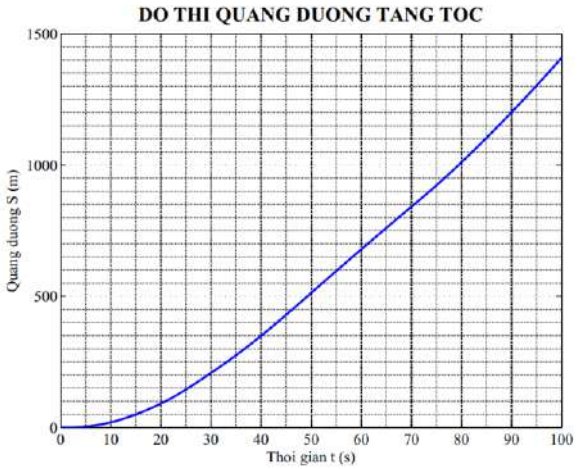
Đồ thị gia tốc J(t) ở giai đoạn đầu vẫn có các đỉnh lớn do quá trình khởi hành và chuyển số, giá trị cực đại vào khoảng 0,7-0,75 m/s². Khi hệ số cản lăn tăng theo step, gia tốc giảm đột ngột xuống mức rất nhỏ, thậm chí có đoạn xấp xỉ 0 hoặc hơi âm đây chính là nguyên nhân làm vận tốc bị tụt xuống ở đoạn giữa. Sau đó gia tốc chỉ phục hồi lên khoảng 0,2-0,3 m/s² rồi giảm dần, cho thấy khả năng tăng tốc của xe đã bị suy giảm đáng kể do ảnh hưởng của cản lăn lớn, dù hệ số bám φ vẫn giữ nguyên bằng 0,8.

Trường hợp 3: Xe chuyển động thẳng trên đường có hệ số bám thay đổi theo hàm bậc thang do đi vào đoạn đường lầy và hệ số cản lăn không đổi. (f = 0.015, φ = 0.8 - 0.02 - 0.8).


DO THI VAN TOC TANG TOC
Van toc v (km/h)
Thoi gian t (s)
0
10
20
30
40
50
60
70
80
90
100


HE SO BAM
He so bam phi
Quang duong S (m)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
500
1000
1500


DO THI GIA TOC TANG TOC
Gia toc J (m2/s)
Thoi gian t (s)
-0.1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
10
20
30
40
50
60
70
80
90
100



Hình 12. Kết quả mô phỏng trường hợp 3

Đồ thị vận tốc v(t) trong giai đoạn đầu, khi hệ số bám còn cao, vận tốc tăng tương đối đều từ 0 lên khoảng 60 km/h ở t ≈ 40 s. Khi cả hai vết bánh đều đi vào đoạn đường có hệ số bám kém (φ giảm theo hàm bậc thang, f = 0,015 không đổi), lực bám tổng cộng trên cầu giảm mạnh nên vận tốc không còn tăng nữa mà bị chững lại và giảm nhẹ. Sau thời điểm hệ số bám được phục hồi, vận tốc lại tăng lên với độ dốc lớn hơn, đến cuối 100 s đạt xấp xỉ 77-78 km/h.

Đồ thị quãng đường S(t) quãng đường tăng liên tục theo thời gian, đồ thị S(t) vẫn là đường cong lồi hướng lên. Tuy nhiên, sau khi hệ số bám của cả hai vết bánh giảm, độ cong của S(t) giảm xuống so với giai đoạn đầu: quãng đường vẫn tăng nhưng "đuối" hơn do vận tốc trung bình trong khoảng 40-70 s nhỏ hơn. Kết thúc 100 s, xe đi được khoảng 1,4-1,45 km, thể hiện rõ ảnh hưởng bất lợi của việc cả hai vết bánh cùng làm việc trong điều kiện bám kém.

Đồ thị gia tốc J(t) trong 20-30 s đầu vẫn có các đỉnh lớn (khoảng 0,7-0,75 m/s²) do quá trình khởi hành và chuyển số. Khi φ của cả hai vết bánh giảm, gia tốc giảm mạnh về gần 0, thậm chí có đoạn hơi âm, thể hiện xe gần như không tăng tốc, đúng với đoạn v(t) bị chững lại. Đáng chú ý là khi hệ số bám được khôi phục, xuất hiện một xung gia tốc rất lớn (đỉnh nhọn trên đồ thị J(t)), do lực bám đột ngột tăng lên làm mô men kéo hữu ích tăng vọt, sau đó gia tốc giảm dần về khoảng 0,2-0,25 m/s² khi xe tiếp tục tăng tốc ở vận tốc cao. Điều này cho thấy nếu hệ số bám của cả hai vết bánh cùng suy giảm thì khả năng tăng tốc của xe bị ảnh hưởng mạnh hơn nhiều so với trường hợp chỉ một vết bánh mất bám, nhưng khi bám được phục hồi, xe cũng phản ứng rất "gắt" với một pha gia tốc lớn.

4. Kết luận

Trên cơ sở số liệu kỹ thuật của xe thiết giáp bánh lốp BTR-60PB, bài báo đã xây dựng mô hình nghiên cứu quá trình tăng tốc và triển khai mô phỏng trong môi trường Simscape Driveline - MATLAB.

Kết quả mô phỏng theo các điều kiện đường phù hợp bối cảnh chiến đấu, trong đó hệ số cản lăn có xu hướng tăng khi xe chạy trên nền cát mềm, lún và hệ số bám có thể suy giảm khi gặp bùn, nước, lầy… Các kết quả thu được cho thấy:

Khi hệ số cản lăn tăng, lực cản lăn tăng lên dẫn đến giảm gia tốc, đồng thời kéo dài thời gian

đạt vận tốc và làm giảm vận tốc tăng tốc v_t trong cùng khoảng thời gian khảo sát. Hệ quả là quãng đường tăng tốc S_t cũng giảm tương ứng. Điều này phản ánh rõ đặc điểm khi BTR-60PB hành quân trên nền mềm lún.

Khi hệ số bám giảm, khả năng truyền lực kéo từ bánh xe xuống mặt đường bị giới hạn bởi điều kiện bám, làm xuất hiện vùng tăng tốc kém hiệu quả (*gia tốc giảm mạnh, khó duy trì tăng tốc đều*). Trong trường hợp bám giảm cục bộ hoặc bám giảm trên phạm vi lớn, v_t và S_t đều suy giảm, đồng thời tính ổn định quá trình tăng tốc giảm do nguy cơ trượt quay tăng lên. Kết quả này phù hợp với thực tế xe hoạt động trên đường bùn lầy, trơn ướt.

Nhìn chung, bài báo đã xây dựng được mô hình mô phỏng động lực học chuyển động thẳng của xe BTR-60PB và làm rõ xu hướng biến thiên của các thông số J_t, v_t, S_t khi điều kiện đường thay đổi. Các kết quả này là cơ sở tham khảo phục vụ đánh giá khả năng cơ động của xe trong điều kiện chiến đấu, đồng thời hỗ trợ định hướng khai thác, huấn luyện và nghiên cứu mở rộng trong các tình huống đường phức tạp.

Tài liệu tham khảo

1. Nguyễn Phúc Hiểu, Vũ Đức Lập, *Lý thuyết ô tô quân sự*, NXB QĐND, Hà Nội, 2000.
2. Nguyễn Văn Trà, *Kết cấu tính toán ô tô*, NXB QĐND, Hà Nội, 2016.
3. Nguyễn Hoài Thanh, Nguyễn Thanh Việt, Bùi Xuân Lâm, *Ứng dụng Matlab trong tính toán kỹ thuật*, Đại học Sư Phạm Kỹ Thuật, TP. Hồ Chí Minh, 2002.
4. *Cấu tạo và sử dụng xe thiết giáp bánh lốp BTR-60PB*, Nhà xuất bản Bộ Tư Lệnh Tăng Thiết Giáp, 1988.
5. *Cấu tạo và sử dụng xe thiết giáp bánh lốp BTR-60PB*, tập 1 và 2, Nhà xuất bản Bộ Tư Lệnh Tăng Thiết Giáp, 1998.
6. PGS.TS. Vũ Đức Lập, *Cơ sở kết cấu ô tô quân sự và xe bọc thép bánh hơi*, Nhà xuất bản Học viện kỹ thuật quân sự, Hà Nội, 1999.
7. Бронетранспортер БТР-60ПБ, парк Патриот.
8. Бронетранспортеры БТР-60П, БТР-60ПА, БТР-60ПБ, btvt.info.

Simulation of the Acceleration Dynamics of the BTR-60PB Wheeled Armored Vehicle Using Simscape Driveline

Abstract: The paper applies Simscape Driveline to successfully simulate the straight-line vehicle dynamics of the BTR-60PB wheeled armored vehicle using a complete driveline model, divided into subsystem groups, which facilitates model setup, parameter tuning, and investigation of multiple parameters. The simulation results show that when road conditions change (i.e., the rolling resistance coefficient fff and the adhesion coefficient φ\varphiφ vary), the vehicle's speed, acceleration, and acceleration distance also change significantly. In particular, a decrease in adhesion or an increase in rolling resistance leads to a reduction in the vehicle's mobility. Therefore, the paper can serve as a useful reference for research, technical assessment, and optimization of vehicle operation in practice.

Keywords: BTR-60PB; vehicle dynamics; rolling resistance coefficient; adhesion coefficient; Simscape Driveline.

Xây dựng mô hình mô phỏng động lực học hộp số tự động phục vụ chẩn đoán

Phạm Anh Tuấn[1*], Nguyễn Văn Trà[2], Vũ Ngọc Tuấn[2], Nguyễn Văn Tứ[2]

[1]Hệ quản lý học viên sau đại học, [2]Học viện Kỹ thuật quân sự

**Email liên hệ: tuan.phamanh@lqdtu.edu.vn*

Tóm tắt

Bài báo nghiên cứu xây mô hình mô phỏng động lực học nhằm chẩn đoán tình trạng kỹ thuật của hộp số tự động. Mô hình mô phỏng mô tả mối quan hệ giữa các thông số điều khiển, quá trình truyền lực và đáp ứng động học của xe trong điều kiện làm việc bình thường và khi hộp số hư hỏng. Trên cơ sở mô phỏng, ảnh hưởng của các lỗi thủy lực và cơ khí của các cụm cơ cấu được xác định thông qua sự biến đổi của đặc tính tăng tốc. Thời gian tăng tốc từ 0 đến 100 km/h được lựa chọn làm thông số chẩn đoán do khả năng phản ánh trực tiếp suy giảm hiệu suất truyền lực của hộp số. Kết quả cho thấy, mô hình mô phỏng có khả năng tạo ra bộ dữ liệu chẩn đoán có thể phân tách được khi hộp số xuất hiện các lỗi khác nhau. Vì vậy, bài báo cung cấp cơ sở khoa học cho việc xây dựng các phương pháp chẩn đoán hộp số tự động dựa trên mô phỏng.

Từ khóa: *Hộp số tự động; động lực học; mô phỏng; chẩn đoán*

1. Mở đầu

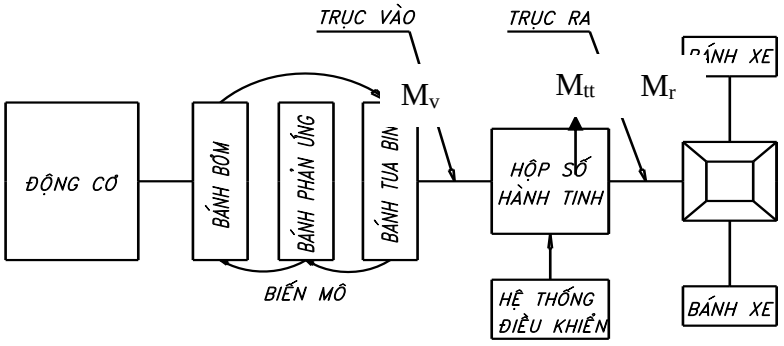
Hộp số tự động là hệ thống tích hợp cơ khí, thủy lực và điều khiển điện tử, ảnh hưởng trực tiếp đến đặc tính động lực học và hiệu suất khai thác của ô tô. Do cấu trúc phức tạp và cơ chế truyền lực phi tuyến, việc đánh giá tình trạng kỹ thuật của hộp số trong thực tế còn nhiều hạn chế, đặc biệt khi xuất hiện các suy giảm nhỏ về hiệu suất truyền lực.

Các phương pháp chẩn đoán hiện nay chủ yếu dựa trên mã lỗi OBD hoặc kinh nghiệm sửa chữa, nên thường chỉ phát hiện hư hỏng khi đã biểu hiện rõ rệt. Trong khi đó, các rò rỉ thủy lực hoặc gia tăng tổn thất cơ khí khó được nhận biết nếu thiếu công cụ phân tích định lượng phù hợp.

Nghiên cứu này xây dựng mô hình động lực học ô tô trang bị hộp số tự động, trong đó hư hỏng được biểu diễn dưới dạng thay đổi tham số của phương trình truyền lực. Trên cơ sở mô hình, thời gian tăng tốc, cùng vận tốc và công suất được đề xuất làm thông số chẩn đoán và được đánh giá bằng độ tách biệt giữa các trạng thái kỹ thuật trong tập dữ liệu mô phỏng..

2. Mô hình động lực học ô tô trang bị hộp số tự động

Trong nghiên cứu này, Mô hình động lực học ô tô được xem xét dưới góc độ cân bằng mô men, với ba đại lượng đặc trưng: mô men vào, mô men ra và tổng tổn thất, như được minh họa ở Hình 1.



Hình 1. Sơ đồ hệ thống truyền lực ô tô trang bị hộp số tự động

Trên cơ sở đó, phương trình cân bằng mô men được thiết lập như sau:

$$J.\frac{d\omega}{dt} = M_v - M_r - M_{tt} \quad (2.1)$$

Trong đó: J mô men quán tính quy đổi về trục quay(bao gồm quán tính trục, bánh răng, biến mô, tải phanh quy đổi, [kg.m^2])

ω vận tốc góc trục, [rad/s]

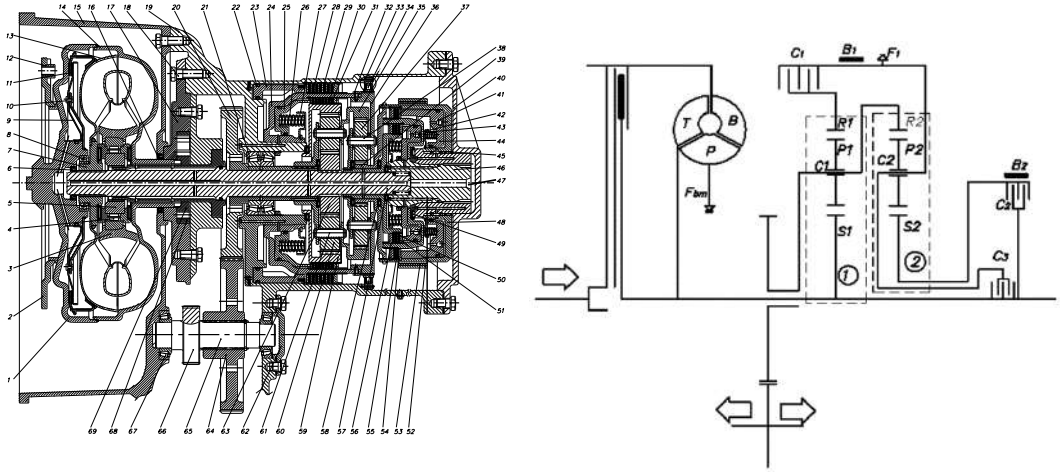
M_v mô men trục vào hộp số(sinh ra bởi động cơ qua biến mô)

M_r mô men trục ra

M_{tt} mô men tổn hao tổng(bao gồm ma sát cơ khí, khuấy dầu, tổn hao biến mô, tổn hao do trượt ly, hợp phanh)

Phương trình (2.1) mô tả cân bằng mô men động lực học của trục quay tương đương ô tô trang bị hộp số tự động. Khi xuất hiện các dạng lỗi cơ khí hoặc thủy lực trong hộp số, mô men tổn hao M_{tt} tăng lên, làm giảm gia tốc góc của trục và kéo dài thời gian tăng tốc. Do đó, các đại lượng động lực học như vận tốc, công suất trục ra và thời gian tăng tốc có thể được sử dụng làm thông số chẩn đoán tình trạng kỹ thuật của hộp số tự động.

Nghiên cứu này sử dụng sơ đồ hộp số tự động 4 cấp kết hợp biến mô và hộp số hành tinh như trình bày trên Hình 2



Hình 2. Hộp số tự động kết hợp biến mô và bộ truyền bánh răng hành tinh

Hộp số tự động kết hợp 2 dãy hành tinh Wilson được điều khiển bởi các ly hợp C1, C2, C3 và các phanh B1, B2. Vai trò các ly hợp và phanh tại mỗi tỷ số truyền được thể hiện trong Bảng 1

Bảng 1. Tổng hợp hoạt động truyền công suất.

Vị trí số	**C1**	**C2**	**C3**	**F1**	**B1**	**B2**
Số 1	x				x	
Số 2	x					x
Số 3	x	x	x			
Số 4			x			x
Số lùi		x		x	x	

3. Mô hình chẩn đoán hộp số tự động ô tô

Hộp số tự động là hệ thống tích hợp các phần tử cơ khí, thủy lực và điều khiển điện tử, do đó hư hỏng có thể phát sinh ở nhiều dạng khác nhau. Tuy nhiên, xét dưới góc độ động lực học, các hư hỏng có ảnh hưởng trực tiếp đến đáp ứng mô men và đặc tính tăng tốc của ô tô chủ yếu xuất phát từ hai nhóm nguyên nhân: gia tăng tổn thất cơ khí bên trong các cụm ma sát truyền động và giảm áp suất điều khiển do rò rỉ thủy lực.

Vì vậy, trong nghiên cứu này, các dạng hư hỏng được mô hình hóa dưới dạng thay đổi tham số tổn thất cơ khí và thay đổi lưu lượng rò rỉ thủy lực tại các vị trí đặc trưng. Cách tiếp cận này cho phép khảo sát ảnh hưởng của suy giảm tình trạng kỹ thuật hiện tại đến cân bằng mô men và đáp ứng tăng tốc của xe, đồng thời hạn chế độ phức tạp không cần thiết của mô hình, nhằm bảo đảm khả năng khai thác hiệu quả cho mục đích chẩn đoán.

Do đó, Mô men tổng tổn hao ở phương trình 2.1 được tính theo [13] như sau:

$$M_{tt} = M_c.sign(\omega) + b.\omega + k.\omega^2 \tag{3.1}$$

Với: M_c mô men ma sát Coulomb(đại diện cho ma sát khô cổ trục, tiếp xúc bánh răng, ma sát ly hợp, phanh khi ép chặt)

$$sign(\omega) = \begin{cases} +1 & \omega > 0 \\ -1 & \omega < 0 \end{cases} \tag{3.2}$$

sign(ω): hàm hấu

b hệ số ma sát nhớt, [N.m.s/rad]

k hệ số tổn hao bậc 2(tổn hao khuấy dầu, cản thủy động trong vỏ hộp số, tổn hao tăng mạnh ở tốc độ cao), [$N.m.s^2/rad^2$]

Rò rỉ thủy lực trong hộp số tự động được thiết lập qua các công thức 3.4

Phương trình liên tục cho thể tích điều khiển theo [14]:

$$Q_v = Q_r - Q_{rr} = \frac{dV}{dt} \tag{3.3}$$

Mô hình rò rỉ dạng lỗ tiết lưu theo:

$$Q_{rr} = C_d.A_{rr}.\sqrt{\frac{2\Delta p}{\rho}} \tag{3.4}$$

Trong đó: Q_v lưu lượng đi vào mạch điều khiển thủy lực, [m^3/s]

Q_r lưu lượng cấp vào từ bơm hoặc van phân phối. [m^3/s]

Q_{rr} lưu lượng rò rỉ, [m^3/s]

V thể tích chất lỏng trong khoang điều khiển, [m^3]

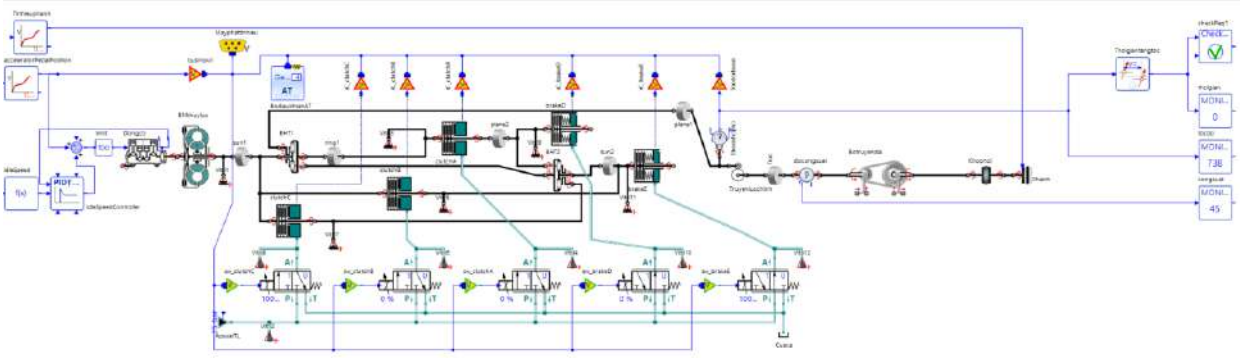
C_d hệ số lưu lượng là đại lượng không thứ nguyên($0<C_d<1$)

A_{rr} diện tích tương tương lỗ rò rỉ, [m^2]

Δp chênh lệch áp suất hai phía lỗ rò, [Pa]

ρ khối lượng riêng của dầu thủy lực, [kg/m^3]

Từ những cơ sở trên, chúng ta tiến hành hoàn thiện mô hình mô phỏng chẩn đoán hộp số tự động AT với 06 lỗi tổn hao cơ khí, 06 lỗi rò rỉ thủy lực tại các vị trí thể hiện trong Hình 3:



Hình 3. Mô hình mô phỏng lỗi hộp số tự động trên SimulationX

Các tổn hao cơ khí và tổn thất thủy lực tại các vị trí thể hiện trong bảng 2.

Bảng 2. Các tham số mô phỏng hư hỏng tại 12 vị trí trong hộp số tự động.

STT	Vị trí lỗi	Giá trị tham số lỗi(max)	Đơn vị
1	Biến mô	0.05	N.m.s/rad
2	Bơm	0.04	$(m^3/s)/MPa$
3	Cơ khí Ly hợp C1	0.05	N.m.s/rad
4	Thủy lực Ly hợp C1	0.04	$(m^3/s)/MPa$
5	Cơ khí Ly hợp C2	0.05	N.m.s/rad
6	Thủy lực Ly hợp C2	0.04	$(m^3/s)/MPa$
7	Cơ khí Ly hợp C3	0.05	Nms/rad
8	Thủy lực Ly hợp C3	0.04	$(m^3/s)/MPa$
9	Cơ khí Phanh B1	0.05	N.m.s/rad
10	Thủy lực Phanh B1	0.04	$(m^3/s)/MPa$
11	Cơ khí Phanh B2	0.05	N.m.s/rad
12	Thủy lực Phanh B2	0.04	$(m^3/s)/MPa$

4. Kết quả và thảo luận

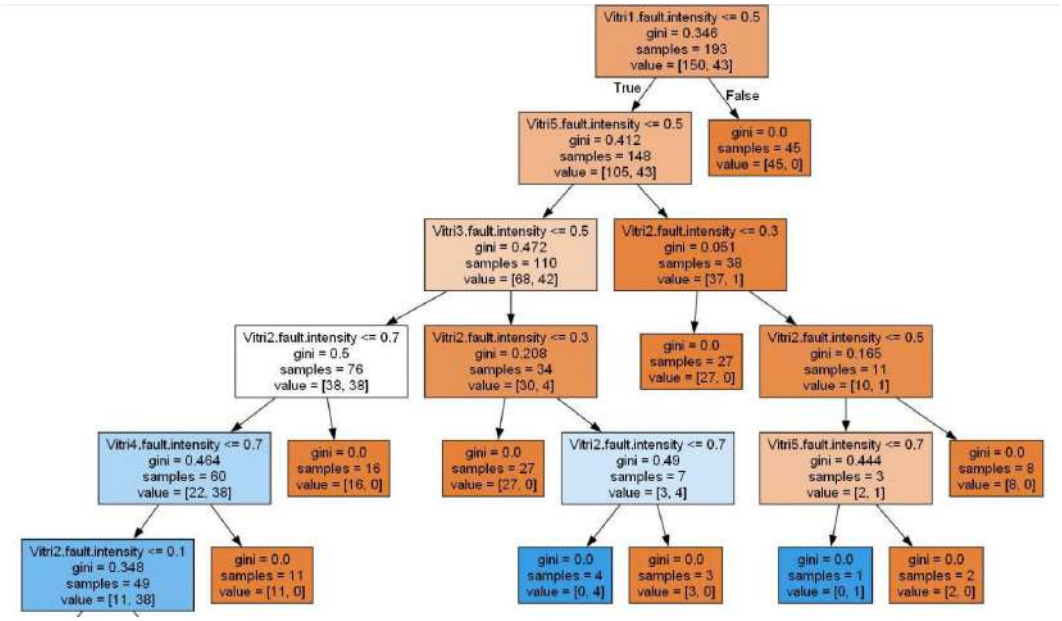
Thời gian tăng tốc của ô tô từ 0 đến 100 km/h được lựa chọn là tham số chẩn đoán chính, do phản ánh tổng hợp ảnh hưởng của toàn bộ quá trình truyền lực đến đặc tính động lực học của ô tô. Tuy nhiên, để đánh giá đầy đủ đặc tính quá độ của hệ thống, nghiên cứu còn xem xét thêm các thông số đặc trưng, như vận tốc và công suất truyền đến bánh xe tại thời điểm 30 s.

Trên cơ sở mô hình đã xây dựng, tiến hành mô phỏng các trạng thái làm việc tương ứng với từng dạng hư hỏng, đồng thời thu thập dữ liệu mô tả mối quan hệ giữa mức độ lỗi và các thông số động lực học của ô tô. Một phần kết quả thu được thể hiện trong Bảng 3, làm cơ sở cho bước phân tích và chẩn đoán ở các phần tiếp theo.

Bảng 3. Kết quả mô phỏng ảnh hưởng mức độ lỗi đến các thông số chẩn đoán

STT	Mức độ lỗi (Vị trí 3)	Thời gian tăng tốc t_{100} (s)	Vận tốc tại 30 s v_{30} (km/h)	Công suất tại 30 s P_{30} (W)
1	0.0	18.64598	589.82089	6805.62565
2	0.01136	18.66719	588.91365	6795.15755
3	0.02273	17.74069	588.00761	6784.70315
4	0.03409	17.76206	587.10288	6774.26402
5	0.04545	17.78351	586.19934	6763.83858
6	0.05682	17.80504	585.29704	6753.42742
7	0.06818	17.82665	584.39595	6743.03021
8	0.07955	17.84833	583.49626	6732.64916
….	…..	…..	……	…….

Khả năng phân tách dữ liệu được thể hiện trên Hình 4. Cấu trúc phân chia cho thấy dữ liệu được tách thành các nhóm tương ứng với mức độ thay đổi tham số lỗi thông qua các ngưỡng rõ ràng. Ở nhiều nhánh cuối, dữ liệu tập trung vào một trạng thái kỹ thuật xác định, cho thấy tồn tại các miền đặc trưng mà tại đó các mức độ hư hỏng không chồng lấn lên nhau.



Hình 4. Cấu trúc phân tách dữ liệu theo các mức độ lỗi dựa trên các thông số động lực học

Các vị trí lỗi được tách ngay từ bước phân chia đầu tiên là những yếu tố ảnh hưởng mạnh đến đáp ứng động lực học của xe. Ngược lại, các vị trí xuất hiện ở các bước sau có tác động hạn chế hơn. Kết quả này phù hợp với phân tích tổn thất và truyền mô men trong mô hình động lực học.

Kết quả chứng tỏ các thông số động lực học có thể phản ánh và phân biệt các mức độ hư hỏng của hộp số tự động. Như vậy, việc khai thác dữ liệu động lực học để chẩn đoán là khả thi khi được đặt trên cơ sở mô hình mô phỏng phù hợp.

5. Kết luận

Bài báo đã xây dựng mô hình động lực học hệ truyền lực ô tô trang bị hộp số tự động, trong đó hư hỏng được biểu diễn dưới dạng thay đổi tổn thất cơ khí và rò rỉ thủy lực. Các đặc trưng động lực học như thời gian tăng tốc, vận tốc và công suất cho thấy khả năng phản ánh rõ sự suy giảm hiệu suất truyền lực và phân tách được các trạng thái kỹ thuật khác nhau.

Kết quả khẳng định tính khả thi của việc chẩn đoán hộp số tự động thông qua biểu hiện động lực học trên cơ sở mô hình vật lý. Nghiên cứu hiện mới dừng ở mô phỏng; các bước tiếp theo cần được kiểm chứng bằng dữ liệu thực nghiệm và mở rộng sang điều kiện khai thác thực tế.

Tài liệu tham khảo

1. Isermann, R. (2011). *Fault-diagnosis applications: Model-based condition monitoring, actuators, drives, machinery, plants, sensors, and fault-tolerant systems*. Springer.
2. Guzzella, L., & Sciarretta, A. (2013). *Vehicle propulsion systems: Introduction to modeling and optimization* (2nd ed.). Springer.
3. Naunheimer, H., Bertsche, B., Ryborz, J., & Novak, W. (2011). *Automotive transmissions: Fundamentals, selection, design and application* (2nd ed.). Springer.
4. Guzzella, L., Onder, C., & Sciarretta, A. (2014). Modeling and control of hybrid and conventional powertrains. *Annual Reviews in Control, 38*(1), 89–100.
5. Zhang, Y., Jiang, J., & Wang, R. (2015). Fault detection and diagnosis for automotive systems: A review. *IEEE Transactions on Intelligent Transportation Systems, 16*(4), 1807–1824.
6. Li, Y., Zhang, H., & Chen, X. (2018). Model-based fault diagnosis of automotive transmission systems considering efficiency degradation. *Mechanical Systems and Signal Processing, 99*, 485–502.
7. Wang, T., Liu, Y., & Qin, D. (2019). Dynamic modeling and fault analysis of automatic transmission systems based on parameter variation. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 233*(12), 3121–3135.
8. Zhao, X., Sun, Z., & Wang, J. (2020). Model-based diagnosis of transmission systems using vehicle dynamic response analysis. *Applied Mathematical Modelling, 77*, 1098–1113.
9. Chen, L., Zhang, M., & Xu, B. (2021). Transmission performance degradation detection using vehicle acceleration characteristics. *IEEE Access, 9*, 45632–45644.
10. Park, J., Kim, S., & Lee, K. (2022). Automatic transmission fault detection using model-based residual generation and vehicle acceleration data. *Sensors, 22*(14), 5304.
11. Wong, J. Y. (2008). *Theory of ground vehicles* (4th ed.). Wiley.
12. Gillespie, T. D. (1992). *Fundamentals of vehicle dynamics*. SAE International.
13. Armstrong-Hélouvry, B. (1991). *Control of machines with friction*. Springer.

14. Fox, R. W., Pritchard, P. J., & McDonald, A. T. (2015). *Introduction to fluid mechanics* (8th ed.). Wiley.

Development of a dynamic simulation model of an automatic transmission for diagnostic purposes

Abstract

This paper presents a dynamic simulation model for diagnosing the faults of automatic transmissions. The model describes the relationship between control parameters, power transmission processes, and vehicle dynamic responses under normal operating conditions and in the presence of transmission faults. Based on simulation results, the effects of hydraulic and mechanical faults in transmission components are identified via acceleration characteristics. The acceleration time from 0 to 100 km/h is selected as the diagnostic indicator due to its ability to directly reflect the degradation of transmission efficiency. The results show that the simulation model is capable of generating diagnostic datasets that can distinguish transmission states under different fault conditions. Therefore, the proposed approach provides a scientific basis for the development of simulation-based diagnostic methods for automatic transmissions.

Keywords: *Automatic transmission; dynamics; simulation; diagnostics*

Nghiên cứu, mô phỏng, đánh giá công thái học của lái xe ô tô quân sự ứng dụng mô hình người lái ảo

Trần Thành Lam[1*], Lê Kỳ Nam, Phạm Quốc Hoàng[2]

[1] *Viện Cơ khí động lực, Học viện Kỹ thuật Quân sự;* [2]*Phòng Khoa học quân sự, Học viện Kỹ thuật Quân sự*

Email: thanhlam.tran@lqdtu.edu.vn; Tel: 0972162262

Tóm tắt

Mô hình người lái ảo với các thông số kích thước, khối lượng, động học, động lực học như người thật trong thực tế đã rút ngắn được thời gian thử nghiệm phương tiện cũng như tạo ra một công cụ hỗ trợ đắc lực cho các nhà sản xuất ô tô, các đơn vị nghiên cứu không ngừng cho ra đời những sản phẩm ô tô chất lượng, đáp ứng được yêu cầu về tính tiện nghi, thoải mái và an toàn của phương tiện. Trong ô tô quân sự cabin có kết cấu phức tạp và có nhiều công năng đặc biệt hơn so với các chủng loại cabin khác. Khối lượng, số lượng chỗ ngồi, bố trí chung có thể thay đổi tùy vào mục đích sử dụng của cabin. Bài báo tiến hành nghiên cứu, mô phỏng, đánh giá công thái học của lái xe khi ngồi trong cabin ô tô quân sự, được thiết kế bằng phần mềm mô phỏng Siemens NX, với một số thông số thiết kế như vị trí bố trí, kích thước hình học của vành tay lái, bàn đạp ga, bàn đạp phanh, cần số các thiết bị phụ và thông số nhân trắc học của người lái. Dựa trên các tiêu chuẩn hiện hành trong nước và trên thế giới, đánh giá mức độ phù hợp của các thông số này so với tiêu chuẩn về mức độ thoải mái và tiện nghi của người lái khi thao tác lái xe và thao tác điều khiển, tác động các trang thiết bị khác trên ô tô quân sự.

Từ khóa: *Mô hình người lái; kiểu bố trí cabin; khoảng cách ghế; kích thước cabin; thao tác lái xe.*

1. Đặt vấn đề

Trong nghiên cứu này cabin ô tô quân sự được thiết kế và chế tạo thực tế với các đặc điểm như sau: chiều cao toàn bộ được hạ thấp tối đa, vị trí người lái được thay đổi sao cho phù hợp với tính năng chiến kỹ thuật của ô tô quân sự, cabin là loại cabin bọc thép hai lớp đạt tiêu chuẩn chống đạn STANAG 4569 level 2 và tăng số chỗ ngồi so với cabin cơ sở. Hệ thống lái được dịch chuyển, các cơ cấu truyền động tay lái, chân ga, chân côn, cần số và gương chiếu hậu, hệ thống đèn chiếu sáng đều được cải hoán để dịch chuyển vị trí.

Để đạt được thiết kế Ergonomics tốt nhất, các nhà khoa học sử dụng các dữ liệu và kỹ thuật theo một số nguyên tắc sau: Nhân trắc học: Kích thước và hình dạng cơ thể; theo số đông và theo từng cá nhân; Giải phẫu học: Cơ, xương, khớp, hệ thống thần kinh; Cơ sinh học: Sức mạnh cơ bắp, lực kéo, lực đẩy; Vật lý môi trường: Tiếng ồn, ánh sáng, nhiệt độ, độ bức xạ; Hệ thống giác quan: Khả năng thính giác, thị giác, cảm giác; Tâm lý học cá nhân: Kỹ năng, khả năng học tập, sai sót, sự khác biệt; Tâm lý học xã hội: Học tập, hành vi, giao tiếp, nhóm.

Từ những nguyên tắc cơ bản trên, bài báo tập trung phân tích yếu tố nhân trắc học của người Việt Nam cũng như của một số nước có sự tương đồng về hình thể và chức năng, thông qua nghiên cứu mô hình con người (Human model), lựa chọn các thông số hình học hợp lý để thiết

[*] Email: thanhlam.tran@lqdtu.edu.vn

kế bố trí chung cabin. Ứng dụng phần mềm thiết kế chuyên dụng để phân tích, đánh giá kết quả nghiên cứu.

Bảng 1. Hệ số tương quan của các kích thước đo (SD) so với chiều cao đứng (r) (Atlas nhân trắc học Việt Nam trong độ tuổi lao động giai đoạn 2017 - 2019)

STT	Kích thước đo	Nam X (cm)	SD	r	Nữ X (cm)	SD	r	r chung hai giới
1	Chiều cao đứng	167,9	6,3	–	155,9	5,7	–	0,97
2	Đứng đến vai	140,3	5,1	0,97	129,6	5,03	0,94	0,97
3	Đứng đến mỏm cùng vai	137,6	5,0	0,97	126,9	5,1	0,93	0,97
4	Đứng đến khuỷu tay	105,3	4,2	0,93	97,4	4,5	0,90	0,93
5	Đứng đến cổ tay	80,9	3,6	0,93	75,2	3,1	0,88	0,93
6	Đứng đến đốt bàn tay	71,9	3,6	0,91	67,2	2,8	0,88	0,93
7	Đứng đến đầu ngón tay	62,5	3,4	0,87	58,1	2,7	0,84	0,90
8	Đứng đến đầu gối	45,4	2,8	0,92	42,9	2,7	0,85	0,85
9	Đứng đến mắt cá trong	7,7	0,7	0,88	6,4	0,9	0,92	0,92
10	Đứng đến khuỷu tay gấp	103,6	4,4	0,96	96,2	3,8	0,86	0,95
11	Đứng đến thắt lưng	98,1	5,0	0,94	94,5	5,2	0,70	0,83
12	Đứng đến mấu chậu	94,6	6,2	0,96	88,3	5,0	0,80	0,88
13	Đứng với tay trên	208,7	8,4	0,96	194,8	7,9	0,94	0,97
14	Đứng với tay trên nằm	201,1	8,2	0,98	186,1	8,0	0,92	0,97
15	Sải tay dưới	167,8	6,3	0,95	155,8	5,0	0,94	0,96
16	Sải tay nằm	155,5	5,2	0,94	141,0	6,5	0,93	0,96
17	Khuỷnh tay	88,5	2,8	0,85	80,9	4,5	0,90	0,92
18	Khuỷnh tay nằm	76,8	3,9	0,94	70,1	3,8	0,80	0,92
19	Rộng vai trái - đầu ngón tay phải	106,9	4,2	0,87	97,5	4,1	0,83	0,92
20	Rộng vai trái - khuỷu tay phải	66,1	2,6	0,85	59,4	2,4	0,81	0,90
21	Cao đứng tự nhiên	165,6	5,2	0,93	153,1	5,4	0,98	0,98
22	Đứng đến hõm gáy	149,7	8,5	0,98	138,9	5,6	0,93	0,96
23	Đứng đến đốt cổ VII	139,1	6,5	0,97	129,5	5,2	0,94	0,96
24	Đứng đến gốc đầu xương vai	122,0	6,1	0,98	112,6	4,6	0,99	0,95
25	Đứng đến mắt	154,1	6,9	0,98	142,7	5,9	0,96	0,98
26	Đứng đến mũi	149,0	6,1	0,98	138,8	5,8	0,96	0,98

Bài báo cũng tiến hành nghiên cứu một số công trình tiêu biểu của nhóm tác giả Heiner Bubb và cộng sự đến từ đại học Technical University of Munich (TUM), CHLB Đức trong cuốn sách "Automobilergonomie" đã nghiên cứu một cách toàn diện các yếu tố công thái học trong thiết kế ô tô [1]. Heiner Bubb và cộng sự đã tự đặt cho mình nhiệm vụ trình bày khái quát việc áp dụng kiến thức rộng rãi trong lĩnh vực công thái học vào thiết kế ô tô. Cuốn sách đề cập đến hai khía cạnh (tính chất hình học và lực) cũng như ứng dụng của chúng. Những đặc điểm vật lý này, được tóm tắt dưới thuật ngữ nhân trắc học, một mặt đề cập đến toàn bộ bộ máy sinh lý học cả về

các đặc tính hình học cũng như khả năng tác dụng lực [1]. Tác giả Nguyễn Trọng Toàn và cộng sự đã tiến hành đề tài "Một số vấn đề về ergonormic ở bộ đội lái xe tăng Việt Nam", 1982. Qua đo đạc các kích thước khoang lái của xe tăng T34, T54 và các kích thước nhân trắc của bộ đội xe tăng, tác giả đã đề nghị đưa ghế lên phía trước 10 cm để bộ đội có thể sử dụng được cần lái tiện lợi hơn. Đồng thời tác giả còn đề nghị tiêu chuẩn tuyển chọn lính xe tăng cho phù hợp với vũ khí, khí tài được trang bị [2]. Phạm Đắc Thuỷ, Ngô Đức Hương đã tiến hành đề tài "Xây dựng vùng thoải mái của lái xe Việt Nam và thử ứng dụng đánh giá buồng lái một số xe vận tải", 1977. Các tác giả bắt đầu bằng việc đo tầm vóc, kích thước của người lái xe rồi xây dựng mô hình người phẳng trước - sau (manecanh phẳng) tỷ lệ 1/10. Mô hình người được tác giả đưa vào thiết kế buồng lái một số xe vận tải theo số đo của các nước sản xuất cùng với tỷ lệ 1/10 đế đánh giá tìm ra các điểm bất hợp lý. Các tác giả đã kiến nghị kích thước và vị trí bố trí cho ghế ngồi, tay lái, côn, phanh chân, ga và cần số [3]. Từ những công trình nghiên cứu trên, bài báo tiến hành xây dựng mô hình thiết kế cabin chuyên dụng. Ứng dụng phần mềm NX và mô hình con người Jack Model để đánh giá ảnh hưởng của đặc trưng công thái học đến thiết kế cabin.

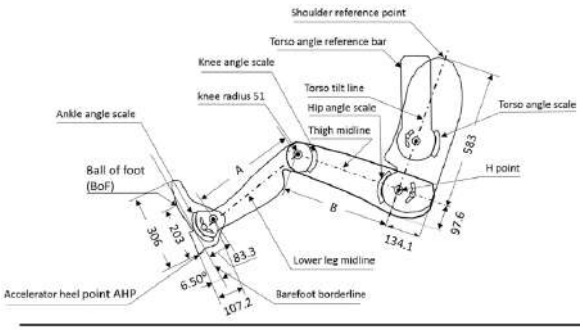
2. Tổng quan về vấn đề nghiên cứu

2.1. Mô hình nhân trắc học con người [1]

2.1.1. Hình mẫu cơ thể

Việc sử dụng các bảng dữ liệu nhân trắc học trong thiết kế hình học nội thất ô tô (còn được gọi là thiết kế bố trí chung nội thất) gặp nhiều vấn đề về giới hạn ứng dụng của nó, do không thể phản ánh đầy đủ sự phức tạp của hình thái cơ thể người trong không gian ba chiều.

Những giới hạn này có thể được khắc phục một phần thông qua việc sử dụng các mẫu vẽ cơ thể được xây dựng phù hợp với mục đích thiết kế. Theo tác giả Müller (2010) [1], các mẫu phác thảo thân người đã được sử dụng từ rất sớm. Từ năm 1928 trong thiết kế khinh khí cầu Zeppelin. Các mẫu này được phát triển dựa trên các phép đo nhân trắc học thực nghiệm cùng kiến thức chuyên môn tích lũy của từng công ty và vì thế có nguồn gốc phát triển riêng biệt trong mỗi tổ chức. Kết hợp với các thiết bị xác định điểm H trên người thực và các thông số kích thước tuân thủ tiêu chuẩn SAE, các hình mẫu SAE J826 đã được xây dựng và phát triển từ năm 1962. Mẫu này bao gồm các phần cẳng chân, đùi và thân người, nhưng không bao gồm phần đầu.



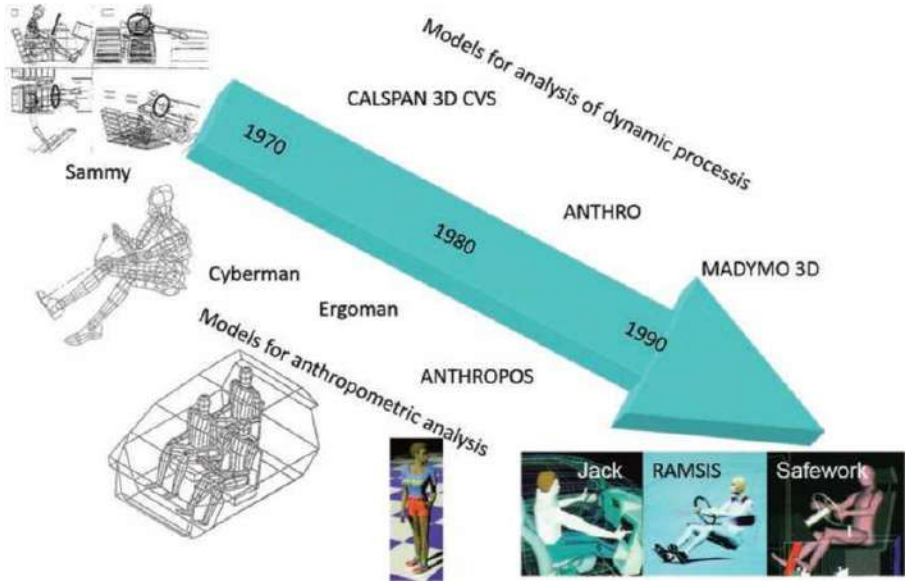
	10% M	50% M	95% M Body outline template
Lower leg A	392.7 mm	417.1 mm	459.1 mm
Thigh B	407.7 mm	431.5 mm	456.0 mm

Hình 1. Hình mẫu theo tiêu chuẩn SAE J826

Nhân trắc học là một từ ghép được tạo thành từ các từ Hy Lạp cổ humans = con người và metin = đo lường. Nhân trắc học trong thiết kế phương tiện tập trung vào ba yếu tố chính: Tư thế của người lái và các chuyển động cơ học; trường nhìn; tương tác lực học và tầm với.

2.1.1. Mô hình con người kỹ thuật số

Với sự ra đời và phát triển mạnh mẽ của công nghệ máy tính, đặc biệt là các kỹ thuật thiết kế có sự hỗ trợ của máy tính (Computer-Aided Design - CAD), dẫn đến một cuộc cách mạng thực sự đã diễn ra trong lĩnh vực thiết kế ô tô. Sự thay đổi này không chỉ giới hạn ở quy trình thiết kế cấu trúc cơ khí và khí động học của phương tiện, mà còn mở ra khả năng mô phỏng hình thái và hành vi của con người trong môi trường ảo. Hình 2 minh họa tiến trình phát triển theo thời gian của các mô hình con người được ứng dụng trong ngành công nghiệp ô tô. Đường phát triển phía dưới thể hiện các mô hình phục vụ xây dựng và đánh giá thiết kế nhân trắc học không gian làm việc của người lái. Các mô hình tiêu biểu gồm ERGOMAN (Phòng Thí nghiệm Nhân học Ứng dụng và Trường Hamaine, Paris), WERNER (Viện Sinh lý học Lao động, Đại học Dortmund), TOMMY (Đại học Công nghệ Dresden), HEINER (Đại học Công nghệ Darmstadt), cũng như các hệ thống ANYBODY và ANTHROPOS của công ty Đức IST và Man3D được phát triển tại INRETS (nay là IFSTTAR) ở Pháp. Mô hình JACK, được phát triển tại Hoa Kỳ dưới sự bảo trợ của NASA vào giữa những năm 1980, sử dụng các giá trị bảng được hiệu chỉnh cho nhân trắc học, cho phép tạo lập các mẫu hình thể hợp lý trong môi trường thiết kế kỹ thuật và được sử dụng cho nghiên cứu của bài báo này.



Hình 2. Sự phát triển của các mô hình con người được sử dụng trong ngành công nghiệp ô tô

2.1.3. Ứng dụng phần mềm NX mô phỏng thiết kế cabin ô tô quân sự và đánh giá kết quả

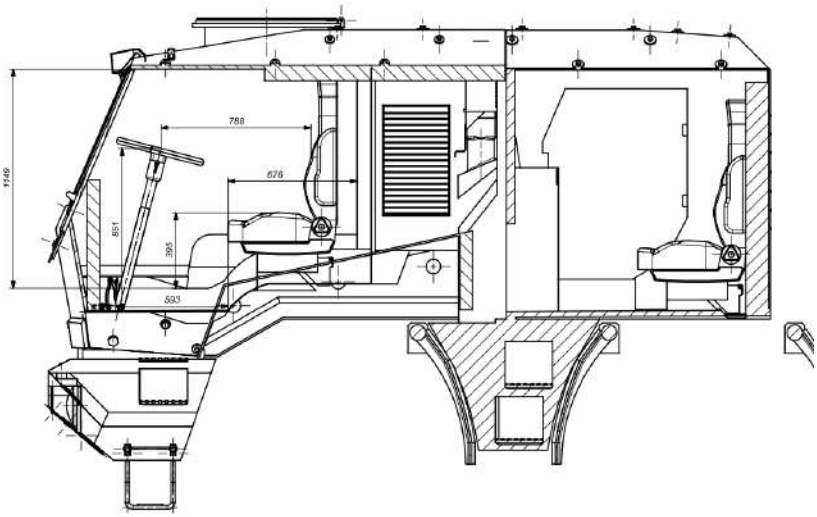
a. Thiết kế cabin ô tô quân sự bằng phần mềm NX

Sử dụng phương pháp nghiên cứu lý thuyết, dựa trên các bộ số liệu tiêu chuẩn và các thiết kế trong nước cũng như trên thế giới, Thông qua phần mềm thiết kế NX xây dựng bộ thông số hình học của cabin.

a)

b)

c)



d)

Hình 3. Thiết kế bố trí chung cabin chuyên dụng bằng phần mềm NX

a). Cabin chuyên dụng hãng TATRA - Cộng hòa SÉC; a). Cabin được thiết kế bằng phần mềm NX và mô hình con người kỹ thuật số JACK; c). Cabin chuyên dụng hãng SEYIT Thổ Nhĩ Kỳ; d). Các thông số hình học của Cabin

Cabin được thiết kế là cabin đôi có thể bố trí được 6-7 người ngồi trên cabin. Vỏ cabin được bọc thép chuyên dụng. Trên cabin có bố trí các hệ thống phụ trợ như điều hòa, thông gió, giá để thiết bị và bảng đồng hồ táp lô.

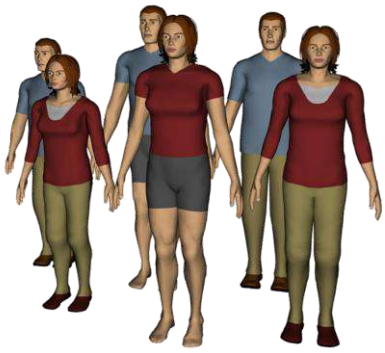
b. Các giả thiết khi đánh giá ảnh hưởng của các yếu tố công thái học lên người lái

Để xây dựng đươc mô hình nghiên cứu bài báo sử dụng các giả thiết sau: Về mặt nhân trắc học, tiến hành thu thập dữ liệu của người Bộ đội Việt Nam thông qua các thông số thống kê kích thước hình học cơ thể, các kích thước này được mô hình hóa và nhập vào phần mềm NX [5]. Các kích thước chính của cabin được thiết kế ngược từ cabin nguyên bản KAMAZ-6560, có điều chỉnh cho phù hợp với điều kiện vận hành tại Việt Nam; trong suốt quá trình khảo sát, kích thước

hình học cabin không thay đổi, có nghĩa là người lái không điều chỉnh ghế; chỉ xét tại vị trí người lái, các tương tác liên quan đến tầm với là tầm với xa nhất đối với người lái.

2.1.4. Thiết lập các thông số nhân trắc học của người Việt Nam với mô hình con người JACK

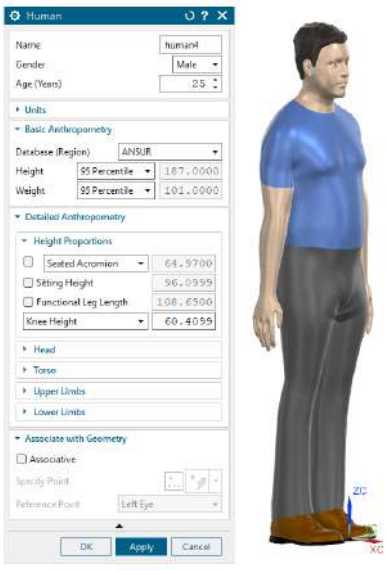
Trong phần mềm Siemens NX, mô-đun JACK cho phép xây dựng và mô phỏng mô hình con người trong không gian ba chiều, phục vụ cho việc đánh giá công thái học và tối ưu bố trí không gian làm việc trong cabin. Với mô hình JACK, người thiết kế có thể lựa chọn các kích thước hình học chi tiết như chiều cao, cân nặng, sải tay, tầm nhìn, tầm với, biên độ chuyển động của các khớp… tương ứng với từng phần vị cơ thể (đầu, thân, tay, chân).



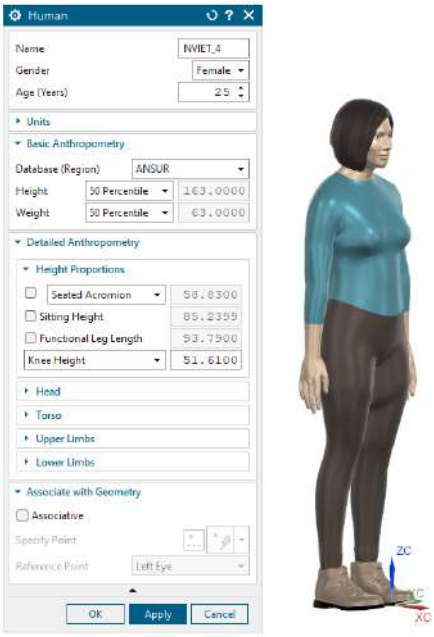
Hình 4. Kích thước cơ thể và các phân vị tương ứng đối với mô hình con người kỹ thuật số JACK

Đặc biệt, mô hình JACK cho phép điều chỉnh các thông số nhân trắc học theo từng nhóm dân số, trong đó có thể chọn bộ dữ liệu đặc trưng cho người Việt Nam. Nhờ đó, mô phỏng sẽ phản ánh chính xác hơn tư thế ngồi, tầm quan sát, khả năng thao tác và mức độ thoải mái của kíp lái trong điều kiện vận hành thực tế. Việc này giúp đánh giá mức độ phù hợp giữa thiết kế cabin và người sử dụng, đồng thời đưa ra các hiệu chỉnh hợp lý về vị trí ghế lái, bảng điều khiển, vô lăng, bàn đạp và các thiết bị hiển thị.

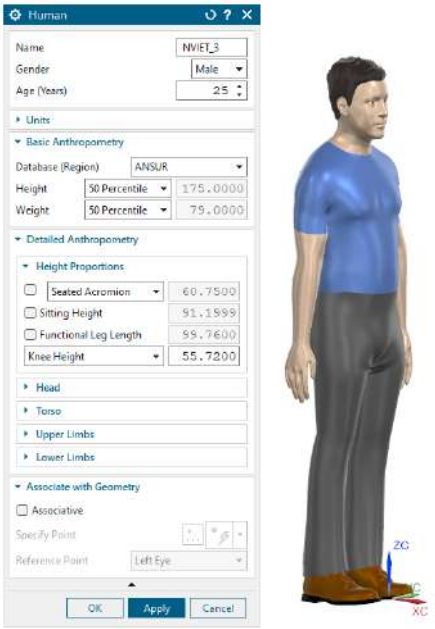
Hình 5. Nữ giới, người Việt Nam, phân vị thứ 5 thấp bé



Hình 6. Nam giới, người Việt Nam, phân vị thứ 95 cao lớn



Hình 7. Nữ giới, người Việt Nam, phân vị thứ 50 trung bình

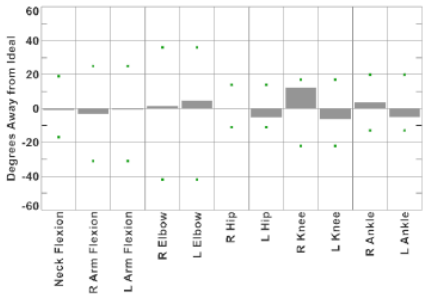


Hình 8. Nam giới, người Việt Nam, phân vị thứ 50 trung bình

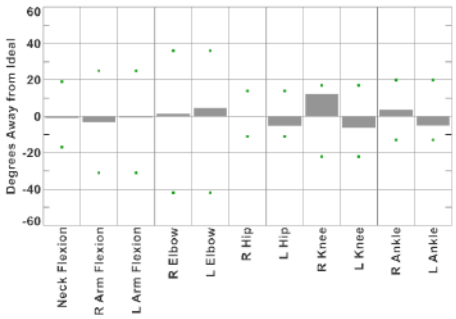
Thông số nhân trắc học người Việt được thiết lập theo bảng dữ liệu nhân trắc theo kết quả thực hiện đo trực tiếp các dấu hiệu chiều cao, cân nặng và cao ngồi trên 5.148 đối tượng thuộc các ngành nghề nông - công nghiệp, học sinh -sinh viên và lao động tự do tại một số tỉnh thành thuộc 3 miền Bắc - Trung - Nam do Viện Khoa học An toàn và Vệ sinh lao động thực hiện trong giai đoạn 2017-2019. Đây là dữ liệu quan trọng cho việc xây dựng mô hình con người đưa vào bài toán khảo sát.

3. Phân tích và đánh giá kết quả [6], [7], [8]

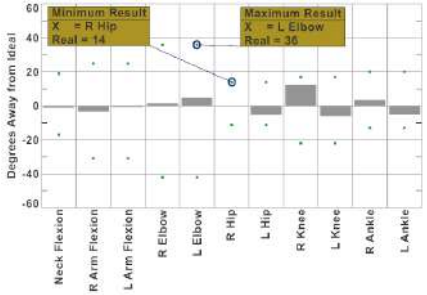
3.1. Trường hợp thay đổi đặc tính hình học của Human model và giữ nguyên thao tác lái xe [7]



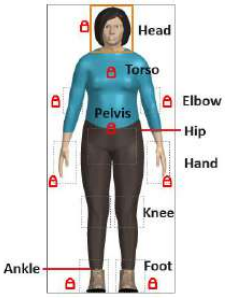
a)



b)



c)



d)

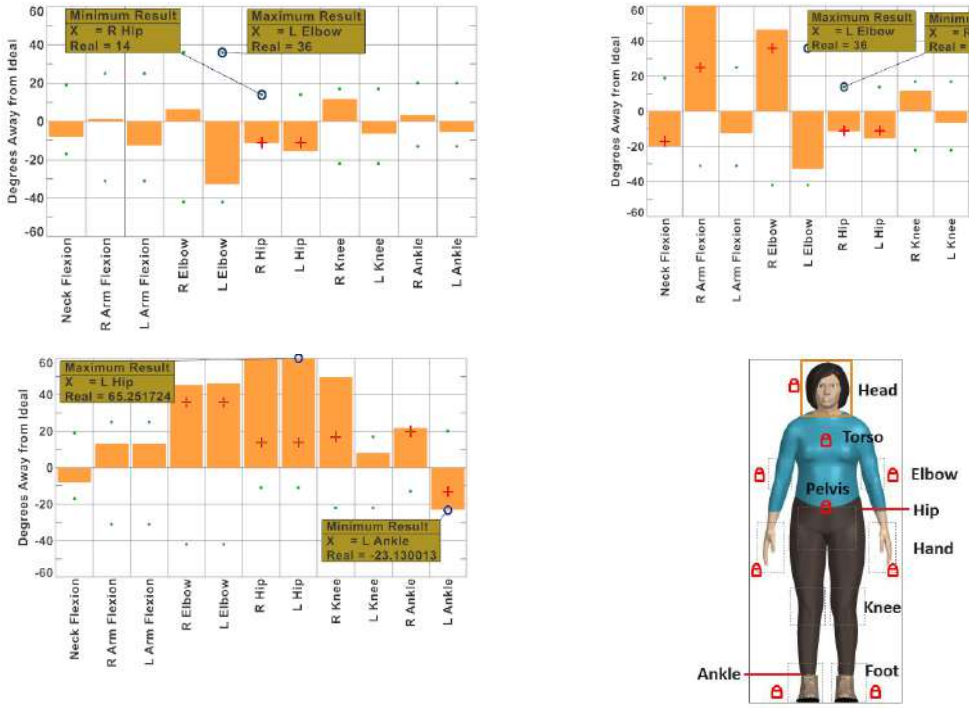
Hình 9. So sánh các đặc tính hình học của Human model và giữ nguyên thao tác lái xe

a). Người Việt Nam: Nam; 25 tuổi; Cao 167,8; Chiều dài đầu gối đến bàn chân 49,31; Chiều cao mông đến đầu 90,73; b). Người Đức: Nam; 25 tuổi; Cao 185,5; Chiều dài đầu gối đến bàn chân 57,52; Chiều cao mông đến đầu 94,9; Người Mê xi cô: Nam; 25 tuổi; Cao 178 cm; Chiều dài đầu gối đến bàn chân 54,84; Chiều cao mông đến đầu 91,29; d). Ký hiệu trên mô hình con người kỹ thuật số

Nhận xét: Từ các đánh giá thống kê dưới dạng biểu đồ cột, thu được kết quả chính như sau: trong trường hợp thứ nhất khảo sát với tư thế human model đang lái xe. Trong trường hợp này, tay người lái được đặt lên vô lăng, hai bàn chân tì trực tiếp vào chân ga và chân côn. Nhận thấy với các chuyển động của cổ của cánh tay, của phần hông đối với người Việt Nam, người Đức và người Mê xi cô có góc độ gần sát với lý tưởng nhất. Đối với cả 3 trường hợp thì phần đầu gối có góc độ xa nhất so với lý tưởng và cụ thể là 36 độ. Điều này thể hiện vị trí của bàn đạp đang có khoảng cách không hợp lý. Thực tế vị trí hay điều chỉnh nhất chính là khoảng cách này.

3.2. Trường hợp giữ nguyên đặc tính hình học Human model và thay đổi thao tác vận hành [8]

Đối tượng khảo sát: Người Việt Nam: Nam; 25 tuổi; Cao 167,8; Chiều dài đầu gối đến bàn chân 49,31; Chiều cao mông đến đầu 90,73; Thao tác kéo cần số.



Hình 10. So sánh các chuyển động của các bộ phân trên cơ thể với giá trị lý tưởng khi giữ nguyên đặc tính hình học Human model và thay đổi thao tác vận hành

Nhận xét: Đối với trường hợp sử dụng mô hình người Việt Nam, tiến hành các thao tác như sử dụng tay phải để nhấn và điều chỉnh các nút như điều hòa, âm lượng. Từ hình 4 nhận thấy rằng khủy tay của người lái có góc độ xa so với lý tưởng nhất với giá trị 28 độ. Đối với tầm với giới hạn, do hiện nay trên bản táp lô, chưa thiết kế cụ thể các đồng hồ, nên ràng buộc được đặt ở

phần khung xe phía trước xa nhất nên một loạt các thông số về góc độ của hông, của đầu gối và của khủy tay là rất xa so với lý tưởng, góc độ lệch của cánh tay so với lý tưởng là 60 độ, của khủy tay là 45 độ. Vì vậy cần thiết kế khoảng cách với xa nhất đến thiết bị phụ hợp lý hơn. Đối với thao tác kéo cần số góc xoay của hông có độ lệch so với giới hạn là 65,2 độ. Góc nghiêng của khuỷu tay cũng chưa hợp lý với giá trị lệch so với lý tưởng là 43 độ.

4. Kết quả và thảo luận

Nội dung bài báo đã giải quyết được các mục tiêu đề ra ban đầu, cụ thể như sau:

Trong nghiên cứu này, nhóm tác giả đã nghiên cứu tổng quan về nhân tố con người trong thiết kế, ứng dụng công cụ mô phỏng là phần mềm NX kết hợp với Human Model "JACK". Từ kết quả đạt được rút ra một số kết luận chính như sau:

- Việc thay đổi các đặc tính hình học của con người dẫn đến thay đổi về tính tiện nghi khi vận hành trên phương tiện; để cải thiện tính tiện nghi, kích thước hình học của Cabin thiết kế phải phù hợp, ghế ngồi cần có bộ điều chỉnh nhiều vị trí.

Hướng nghiên cứu tiếp theo:

- Sử dụng công cụ Design Space Management tối ưu hóa không gian thiết kế bên trong cabin, đặc biệt các kích thước hình học; nghiên cứu các tác động về mặt lực học, đặc biệt là lực do cơ gây ra khi người lái tiến hành các thao tác trong cabin (Chuyển số, đạp bàn đạp ly hợp, phanh, ga, quay vô lăng…); nghiên cứu trường nhìn và các yếu tố như độ ồn và rung động đến người lái.

Tài liệu tham khảo

1. Heiner Bubb, Klaus Bengler, Rainer E. Grünen, Mark Vollrath, Ergonomie für das Automobil, ISBN 978-3-8348-1890-4, DOI 10.1007/978-3-8348-2297-0, Springer Vieweg 2015.
2. Nguyễn Trọng Toàn, "Một số vấn đề về ergonormic ở bộ đội lái xe tăng Việt Nam", 1982.
3. Phạm Đắc Thuỷ, Ngô Đức Hương, đề tài "Xây dựng vùng thoải mái của lái xe Việt Nam và thử ứng dụng đánh giá buồng lái một số xe vận tải", 1977.
4. Tiêu chuẩn quốc gia TCVN 5781:2009, Phương pháp đo cơ thể người, Method of human body measuring.
5. Phạm Thị Thắm, Nghiên cứu đặc điểm hình dáng cơ thể người việt nam và ứng dụng để thiết kế quần áo, Luận văn thạc sĩ khoa học chuyên ngành công nghệ vật liệu dệt may, Đại học Bách khoa Hà Nội 2011.
6. Abdel-Malek, K., Yang, J., Marler, T., Beck, S., Mathai, A., Zhou, X.: Towards a new generation of virtual humans. Int. J. Human Factors Model. Simul. 1(1), 2-39 (2006).
7. Adams, M.A.: Wie beeinflusst die Sitzposition die Last- verteilung in der Bandscheibe und ihre Ernahrung. In: Ergomechanics Interdisziplinarer Kongress Wir- belsaulenforschung, p. 2. Shaker Verlag, Aachen (2006).
8. Ahlberg, A., Moussa, M., Al-Nahdi, M.: On geographi¬cal variations in the normal range of joint motion. Clin. Orthop. Relat. Res. 234, 229-231 (1988).

9. Tingley, M. W., Monahan, W. B., Beissinger, S. R., & Moritz, C. (2009). Birds track their Grinnellian nice through a century of climate change. Proceedings of the National Academy of Science, USA, 106, 19637-19643.
10. Bassey, E.J., Short, A.H.: A new method for measuring power output in a single leg extension - feasibility, reliability and validity. Eur. J. Appl. Physiol. Occup. Physiol. 60 (5), 385-390 (1990).

Research, Simulation, and Ergonomic Evaluation of Military Vehicle Drivers Using a Virtual Human Driver Model

Abstract: *Virtual driver models with dimensions, weight, kinematics, and dynamics identical to those of a real person have shortened vehicle testing time and created a powerful tool to support automotive manufacturers and research units in continuously producing high-quality vehicles that meet the requirements for comfort, convenience, and safety. In military vehicles, the cabin has a complex structure and more specialized functions compared to other types of cabins. The weight, number of seats, and overall layout can vary depending on the cabin's intended use. This paper conducts research, simulation, and evaluation of the ergonomics of a driver sitting in a military vehicle cabin, designed using Siemens NX simulation software, with several design parameters such as the position and geometric dimensions of the steering wheel, accelerator pedal, brake pedal, gear shift, auxiliary equipment, and the driver's anthropometric parameters. Based on current domestic and international standards, assess the suitability of these parameters compared to standards for driver comfort and convenience when driving and operating the controls and other equipment on military vehicles.*

Keywords: Driver model; cabin layout; seat spacing; cabin dimensions; driving operation.

Nghiên cứu ảnh hưởng của tỷ số nén đến trạng thái ứng suất-nhiệt và trạng thái biến dạng của nhóm piston - xilanh động cơ diesel tàu thủy

Đỗ Quang Đức[1], Nguyễn Văn Dương[2]

[1]Hệ QLHV Sau đại học, [2]Viện Cơ khí động lực

Email: doquangduc842@gmail.com, Tel: 092.449.8789

Tóm tắt

Bài báo trình bày nghiên cứu ảnh hưởng của tỷ số nén đến trạng thái nhiệt, ứng suất và biến dạng của nhóm piston – xilanh của động cơ diesel 4 kỳ 3D12 sử dụng làm máy chính trên tàu thủy. Mô hình tính toán trạng thái ứng suất và biến dạng của piston – xilanh được xây dựng bằng phương pháp mô phỏng phần tử hữu hạn trên phần mềm Ansys Workbench. Kết quả tính toán cho thấy khi tỷ số nén tăng, nhiệt độ và ứng suất của piston và xilanh cũng tăng đáng kể. Khi tỷ số nén tăng từ ε = 14 lên ε = 16, nhiệt độ lớn nhất của piston tăng trung bình khoảng 18,66 K, trong khi nhiệt độ bề mặt trong của xilanh tăng trung bình khoảng 10,14 K. Ứng suất cơ – nhiệt cũng tăng rõ rệt: tại góc lượn phía dưới đỉnh piston, ứng suất tăng từ 105,06 MPa lên 130,11 MPa; tại vùng vai tựa phía trên của xilanh, ứng suất tăng từ khoảng 462,37 MPa lên 511,95 MPa. Sự gia tăng nhiệt độ và ứng suất làm biến đổi đáng kể biến dạng của piston và xilanh, qua đó ảnh hưởng trực tiếp đến độ tin cậy và tuổi thọ làm việc của các chi tiết.

Từ khóa: *Tỷ số nén; Trạng thái nhiệt; Ứng suất; Piston; Xilanh.*

1. Đặt vấn đề

Để đáp ứng các quy định ngày càng nghiêm ngặt về tiết kiệm nhiên liệu và kiểm soát phát thải, việc tăng hiệu suất và giảm phát thải của động cơ đốt trong đang nhận được sự quan tâm lớn từ các nhà nghiên cứu. Đối với các động cơ diesel thế hệ cũ, các giải pháp cải tiến cần đảm bảo nâng cao hiệu quả làm việc nhưng không làm thay đổi đáng kể kết cấu tổng thể cũng như tính ổn định khai thác của động cơ.

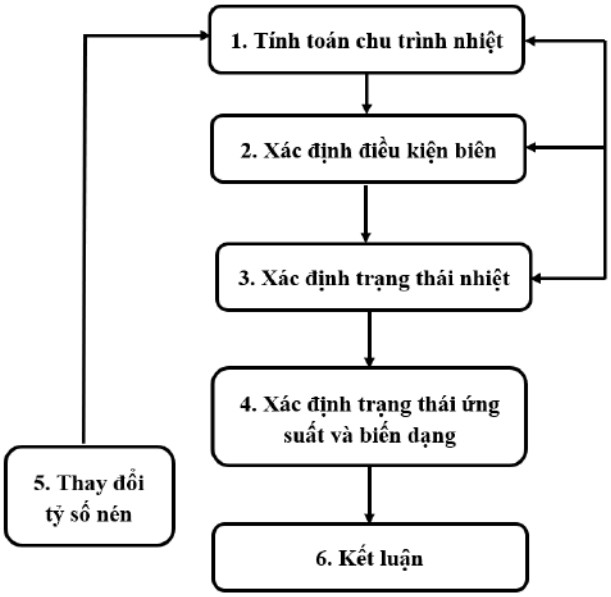
Trong số các giải pháp kỹ thuật, thay đổi và tối ưu tỷ số nén được xem là một biện pháp khả thi nhằm cải thiện hiệu suất nhiệt, giảm suất tiêu hao nhiên liệu mà không yêu cầu thay đổi quá lớn về kết cấu cơ bản của động cơ [1-4]. Trong một báo cáo đã được công bố, tác giả của bài báo này cùng các cộng sự đã đề xuất phương án tăng tỷ số nén cho động cơ 3D12 bằng cách sử dụng bạc lót với độ dày không đều và khảo sát ảnh hưởng đến các thông số làm việc của động cơ [4]. Kết quả cho thấy, khi tăng tỉ số nén, hiệu suất cải thiện đáng kể.

Tuy nhiên, khi tỷ số nén tăng, nhiệt độ và áp suất cực đại trong buồng cháy cũng tăng theo, làm gia tăng tải nhiệt và cơ học lên các chi tiết của cơ cấu khuỷu – thanh truyền, đặc biệt là nhóm piston – xilanh. Vì vậy, việc đánh giá một cách toàn diện ảnh hưởng của tỷ số nén đến trạng thái nhiệt, trạng thái ứng suất và biến dạng của nhóm piston – xilanh là hết sức cần thiết. Xuất phát từ những phân tích trên, trong bài báo này, nhóm tác giả tiến hành nghiên cứu ảnh hưởng của sự thay đổi tỷ số nén đến trạng thái nhiệt, ứng suất và biến dạng của nhóm piston – xilanh động cơ diesel tàu thủy. Kết quả nghiên cứu là cơ sở quan trọng để đánh giá độ bền của piston-xilanh và giới hạn tỷ số nén cho động cơ trong quá trình cải tiến.

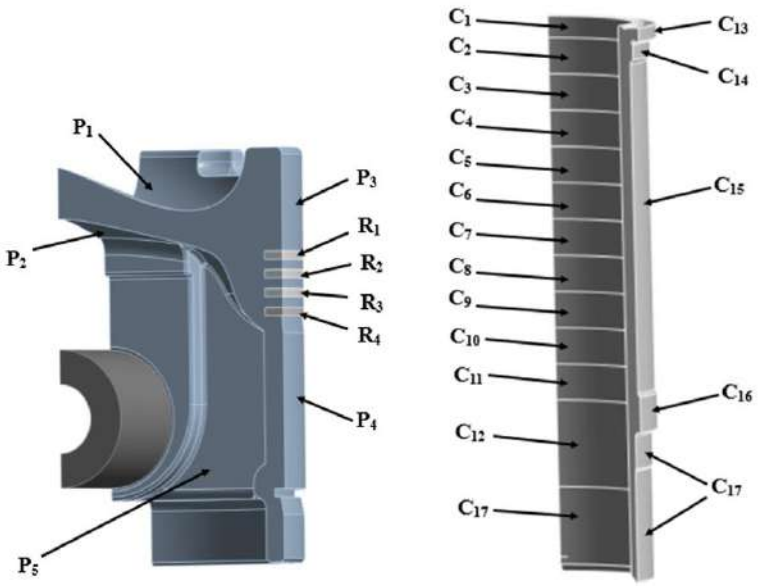
2. Phương pháp nghiên cứu

Các bước khảo sát ảnh hưởng của sự thay đổi tỷ số nén đến trạng thái nhiệt, trạng thái ứng suất và biến dạng của nhóm piston – xilanh được trình bày trong Hình 1. Mô hình tính toán chu trình nhiệt của động cơ được thiết lập trong phần mềm GT-Suite (bước 1) và được hiệu

chỉnh theo số liệu do nhà sản xuất công bố [4]. Kết quả tính toán chu trình nhiệt là điều kiện biên đầu vào để tính toán trạng thái ứng suất-nhiệt của piston-xilanh (bước 3, 4). Điều kiện biên nhiệt trên các bề mặt piston-xilanh được xác định theo thuật toán được công bố trong [5, 6]. Điều kiện biên cơ học và cách ràng buộc mô hình được xác định theo phương pháp được công bố trong [7].



Hình 1. Sơ đồ nghiên cứu



Hình 0. Mô hình piston - xilanh

Trên cơ sở các điều kiện biên thu được, mô hình phần tử hữu hạn nhằm xác định trạng thái nhiệt, trạng thái ứng suất và biến dạng của piston – xilanh được xây dựng trong phần mềm Ansys Workbench (bước 3 và bước 4). Khi thay đổi tỷ số nén (bước 5), các thông số của chu trình công tác sẽ biến đổi, kéo theo sự thay đổi của nhiệt độ môi chất và hệ số trao đổi nhiệt trên thành buồng cháy. Do đó, chu trình công tác và điều kiện biên nhiệt trên các bề mặt piston – xilanh cần được tính toán lại tương ứng với từng giá trị tỷ số nén, nhằm đảm bảo tính chính xác của kết quả mô phỏng.

Các trường hợp khảo sát được thực hiện theo 5 trường hợp với tỷ số nén (ε) thay đổi từ 14 đến 16 như trong bảng 1.

Bảng 1. Các trường hợp khảo sát sự thay đổi tỷ số nén

Trường hợp	TH1	TH2	TH3	TH4	TH5
Tỷ số nén (ε)	14	14,5	15	15,5	16

Trong các trường hợp khảo sát, hệ số truyền nhiệt và nhiệt độ nước làm mát trên bề mặt ngoài của xilanh (bề mặt C15, C17 hình 2) không đổi.

3. Đối tượng nghiên cứu

Đối tượng nghiên cứu được tác giả lựa chọn là động cơ diesel 4 kỳ không tăng áp 3D12 sử dụng làm máy chính và máy phát điện trên tàu thủy. Thông số kỹ thuật chính của động cơ 3D12 được thể hiện như trong bảng 2.

Bảng 2. Thông số kỹ thuật động cơ 3D12 [8]

Thông số	Trị số	Thông số	Trị số
Số xilanh	12	Công suất định mức, [kW]	220
Đường kính xilanh [mm]	150	Tốc độ vòng quay trục khuỷu ứng với công suất định mức [vòng/phút]	1500
Hành trình piston [mm]	180	Suất tiêu hao nhiên liệu [g/kWh]	258

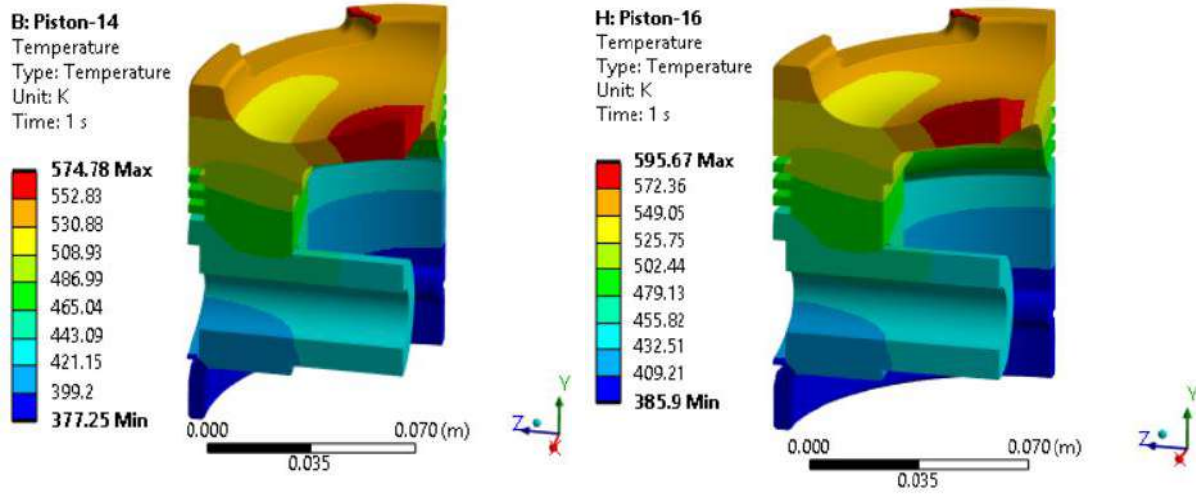
Piston động cơ 3D12 được làm từ hợp kim nhôm AK-4 và xilanh được làm từ thép chịu nhiệt 38XMIOA.

4. KẾT QUẢ VÀ THẢO LUẬN

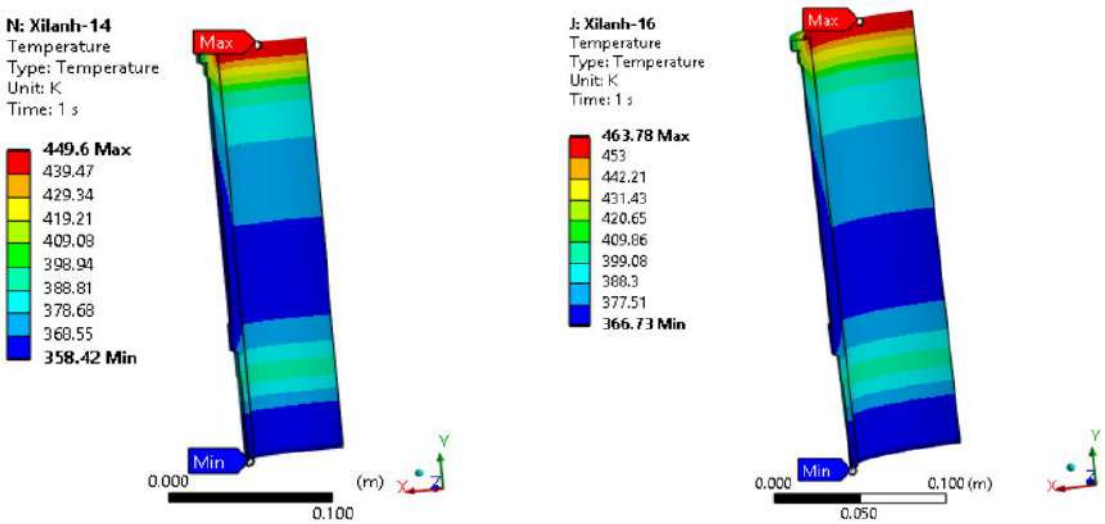
4.1. Ảnh hưởng của tỷ số nén đến trạng thái nhiệt của piston, xilanh

Kết quả trạng thái nhiệt của piston và xilanh khi tăng tỷ số nén từ 14 lên 16 được thể hiện trong hình 3 và hình 4.

Nhiệt độ trung bình trên bề mặt đỉnh piston và bề mặt trong của xilanh (khu vực tiếp xúc với môi chất công tác) được thể hiện trong bảng 4.



Hình 3. Trạng thái nhiệt độ piston với tỷ số nén ε = 14 và ε = 16



Hình 4. Trạng thái nhiệt độ xilanh với tỷ số nén ε = 14 và ε = 16

Kết quả trạng thái nhiệt của piston trên hình 3 và 4 cho thấy, khi tăng tỷ số nén, nhiệt độ của piston và xilanh tăng lên đáng kể. Giá trị nhiệt độ lớn nhất của piston tăng từ 547,78 K lên 595,67 K, đối với xilanh tăng từ 449,6 K lên 463,78 K. Khi tăng tỷ số nén từ ε = 14 và ε = 16, nhiệt độ trung bình trên bề mặt đỉnh piston ($\overline{T}_P$) tăng từ 542,44 K lên 561,1 K, nhiệt độ trung bình trên bề mặt trong của xilanh ($\overline{T}_C$) tăng từ 382,28 K lên 392,42 K theo bảng 4.

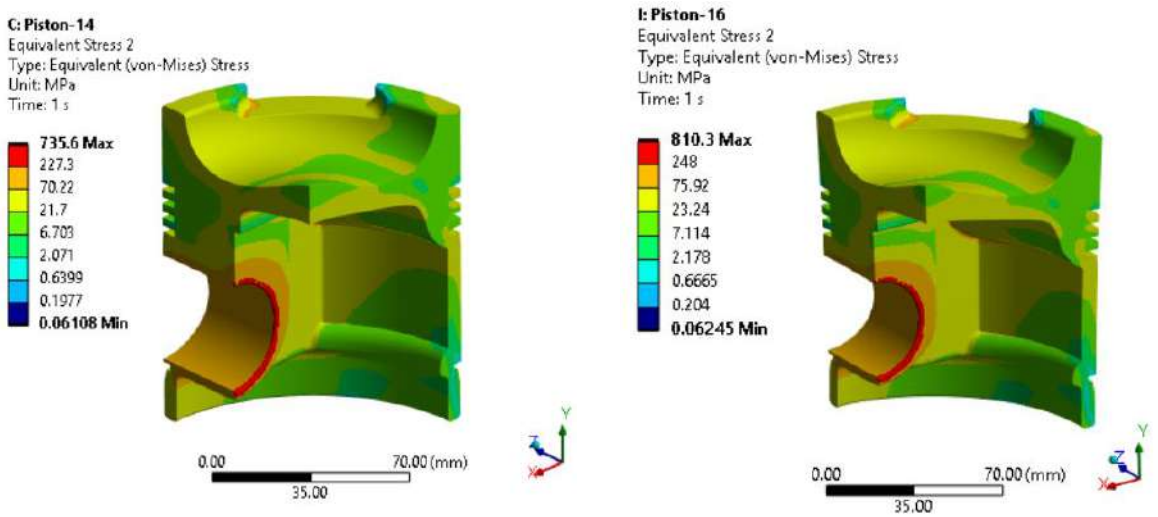
Bảng 4. Nhiệt độ trung bình trên bề mặt đỉnh piston ($\overline{T}_P$) và bề mặt trong của xilanh ($\overline{T}_C$) theo các trường hợp khảo sát.

Tỷ số nén / Vị trí	14	14.5	15	15.5	16
Đỉnh piston, Tp	542,44	545,94	550,35	555,92	561,1
Xilanh, Tc	382,28	384,46	386,81	389,53	392,42

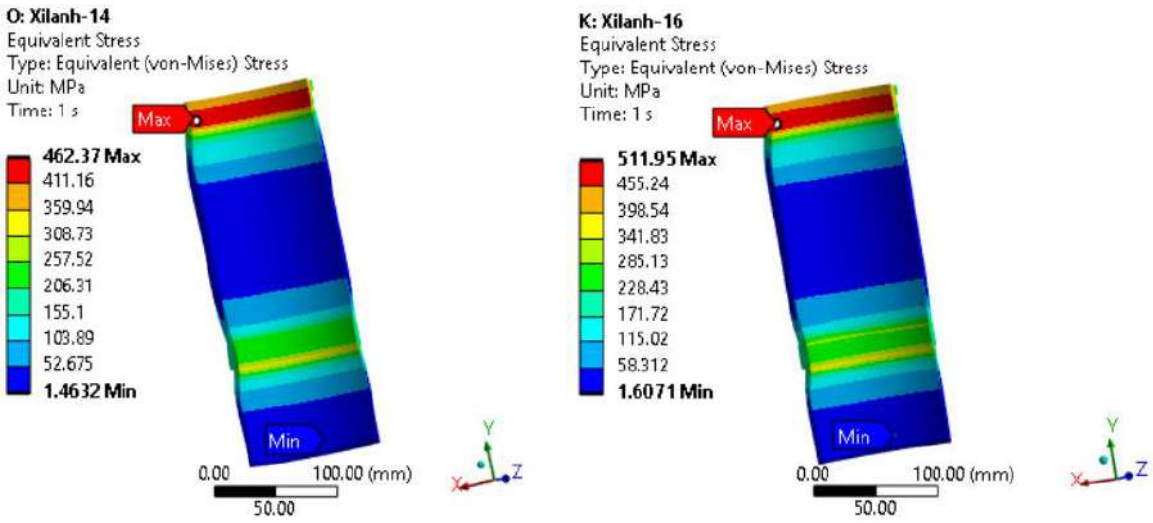
**Trong bảng 4: $\overline{T}_C$ - Nhiệt độ trung bình bề mặt trong xilanh, được giới hạn từ điểm chết dưới của piston tới điểm trên cùng của xilanh, (K).*

4.2. Ảnh hưởng của tỷ số nén đến trạng thái ứng suất piston-xilanh

Kết quả trạng thái ứng suất của piston và xilanh khi tăng tỷ số nén từ 14 lên 16 được thể hiện trong hình 5 và hình 6.

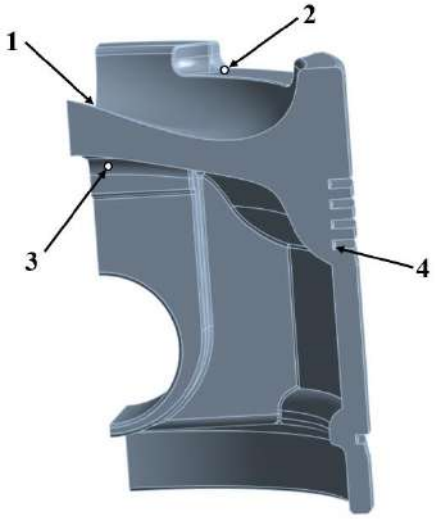


Hình 5. Trạng thái ứng suất trên piston với tỷ số nén ε = 14 và ε = 16



Hình 6. Trạng thái ứng suất trên xilanh với tỷ số nén ε = 14 và ε = 16

Một số vị trí đặc biệt trên piston tập trung ứng suất lớn và nhiệt độ cao được lựa chọn để khảo sát khi thay đổi tỷ số nén như trong Hình 7 và kết quả thể hiện trong bảng 5.



Hình 7. Các điểm khảo sát ứng suất nhiệt – cơ trên piston

Bảng 5. Kết quả ứng suất tại các điểm khảo sát trên piston theo các tỷ số nén khác nhau

Điểm	ε = 14		ε = 14.5		ε = 15		ε = 15.5		ε = 16	
	T, [K]	σ, [MPa]	T, [K]	σ, [MPa]	T, [K]	σ, [MPa]	T, [K]	σ, [MPa]	T, [K]	σ, [MPa]
1	572,7	16,59	576,4	17,75	581,2	18,91	587,5	20,18	593,4	22,19
2	536,2	92,24	539,7	95,17	544,1	97,95	549,4	100,8	554,4	103,75
3	523,8	105,06	527,2	111,82	531,4	118,09	536,67	124,26	541,5	130,11
4	457,4	129,97	460,4	137,65	463,6	145,29	467,57	152,29	470,8	160,92

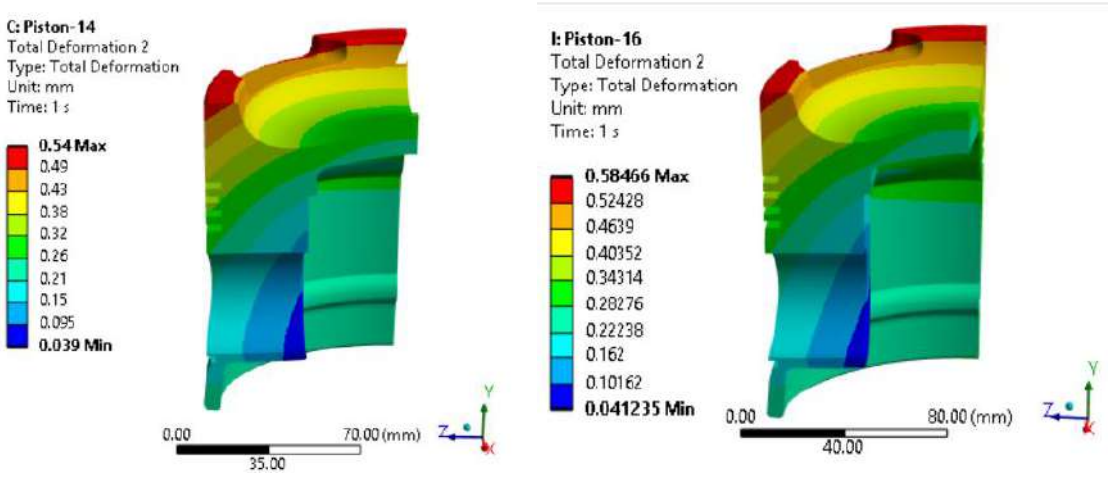
Kết quả phân bố ứng suất của piston và xilanh thể hiện trên Hình 5, Hình 6 và số liệu trong Bảng 5 cho thấy khi tăng tỷ số nén từ 14 lên 16, ứng suất trong các chi tiết có xu hướng tăng rõ rệt. Điều này là do khi tỷ số nén tăng, áp suất và nhiệt độ cực đại của khí cháy trong xi

lanh tăng lên, làm gia tăng tải trọng nhiệt – cơ tác dụng lên các chi tiết. Đối với xilanh, ứng suất lớn nhất tăng từ 462,37 MPa lên 511,95 MPa, tập trung chủ yếu tại khu vực phía trên và vai tựa xilanh. Đây là vùng chịu tác động trực tiếp của áp suất và nhiệt độ khí cháy, đồng thời bị hạn chế biến dạng do liên kết lắp ghép với thân máy, dẫn đến sự hình thành vùng ứng suất lớn.

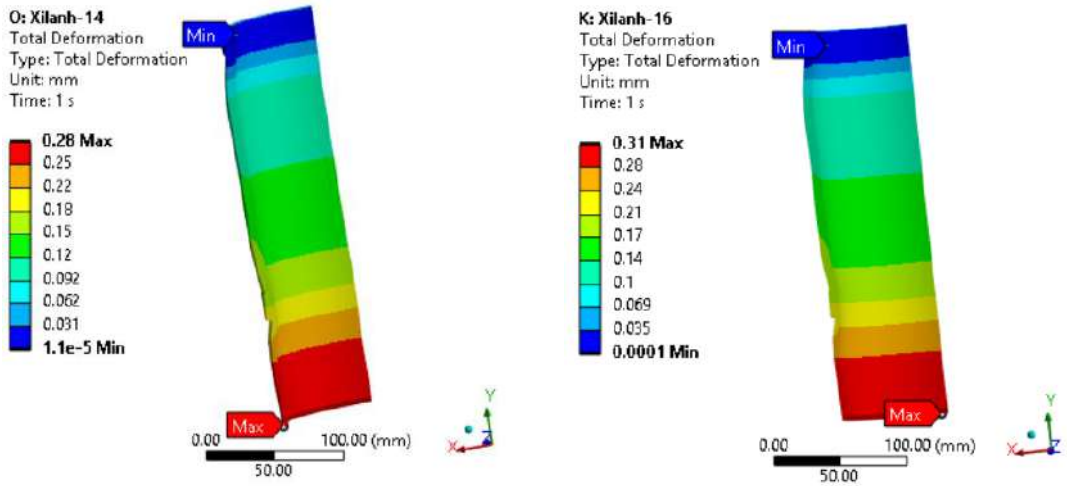
Đối với piston, ứng suất lớn nhất xuất hiện tại khu vực bệ chốt piston. Đây là vị trí truyền lực từ piston sang thanh truyền thông qua chốt piston nên chịu tải trọng cơ học lớn. Ngoài ra, việc thiết lập các điều kiện ràng buộc trong mô hình tính toán dẫn đến hiện tượng tập trung ứng suất tại vị trí này.

Xét tại các điểm khảo sát trên piston, ứng suất đều tăng khi tỷ số nén tăng từ 14 lên 16. Cụ thể, tại đỉnh piston (điểm 1), là khu vực chịu tác động trực tiếp của nhiệt độ và áp suất khí cháy, ứng suất tăng khoảng 34%. Các điểm 2 và 3 nằm tại các vị trí thay đổi hình học như góc vát trên đỉnh piston và góc lượn phía dưới đỉnh piston, sự thay đổi tiết diện đột ngột tại các khu vực này làm xuất hiện hiện tượng tập trung ứng suất. Tại rãnh xéc măng dầu (điểm 4), là vùng chuyển tiếp giữa phần đỉnh và thân piston, sự giãn nở nhiệt không đồng đều giữa các vùng của piston kết hợp với sự thay đổi hình học kết cấu dẫn đến sự gia tăng ứng suất tại vị trí này.

4.3. Ảnh hưởng của tỷ số nén đến trạng thái biến dạng piston-xilanh



Hình 8. Kết quả tính biến dạng của piston ở ε = 14 và ε = 16



Hình 9. Kết quả tính biến dạng của xilanh ở ε = 14 và ε = 16

Kết quả hình 8 và 9 cho thấy, khi tăng tỷ số nén từ 14 lên 16, biến dạng tổng của piston và xilanh đều có xu hướng tăng, trong đó piston chịu ảnh hưởng mạnh hơn do đặc tính vật liệu và điều kiện làm mát kém hơn.

Đối với piston, biến dạng tập trung chủ yếu ở khu vực đỉnh piston, nơi chịu tác động trực tiếp của khí cháy có nhiệt độ và áp suất rất cao. Đây cũng là vùng có khả năng tản nhiệt kém nhất, do quá trình làm mát chỉ diễn ra gián tiếp thông qua dầu bôi trơn hoặc các rãnh làm mát với hiệu quả hạn chế.

Đối với xilanh, giá trị biến dạng lớn nhất cũng tăng từ 0,28 mm lên 0,31 mm, tương ứng mức tăng khoảng 10,7%. Biến dạng lớn nhất chủ yếu xuất hiện ở phần đuôi xilanh, là khu vực không được làm mát trực tiếp và không nằm trong vùng chịu ràng buộc của áo nước xilanh. Mặc dù sự biến dạng tương tự như piston, nhưng độ biến dạng tuyệt đối của xilanh vẫn nhỏ hơn. Nguyên nhân là xilanh được chế tạo từ thép hợp kim X38 có độ cứng cao và hệ số giãn nở nhiệt thấp hơn, trong khi piston sử dụng hợp kim nhôm AK-4 có hệ số giãn nở nhiệt lớn hơn, dẫn đến mức biến dạng cao hơn trong cùng điều kiện làm việc, điều này có thể làm thu hẹp khe hở nhiệt và làm tăng ma sát, thậm chí bó kẹt piston. Vì vậy, sự thay đổi kích thước khe hở nhiệt giữa piston-xilanh khi cải tiến động cơ bằng cách tăng tỉ số nén cần được xem xét một cách cẩn trọng.

5. Kết luận

Bài báo đã khảo sát ảnh hưởng của tỷ số nén đến trạng thái nhiệt, ứng suất và biến dạng của nhóm piston – xilanh trên động cơ diesel tàu thủy 3D12. Kết quả cho thấy khi tăng tỷ số nén từ 14 lên 16 dẫn đến nhiệt độ, ứng suất và biến dạng của piston và xilanh đều tăng.

Nhiệt độ trung bình trên bề mặt đỉnh piston tăng từ 542,44 K lên 561,1 K, trong khi nhiệt độ trung bình trên bề mặt trong của xilanh tăng từ 382,28 K lên 392,42 K. Nhiệt độ tại các vị trí gần khu vực đỉnh piston và khu vực tiếp xúc với khí cháy tăng mạnh hơn so với các vùng khác.

Khi tăng tỷ số nén từ 14 lên 16, trạng thái ứng suất trên piston và xilanh đều có xu hướng tăng đáng kể, tập trung tại phần đỉnh, vùng góc lượn phía trên thành đỉnh piston, vùng rãnh xéc măng và vùng phía trên và vai tựa của xilanh. Nhiệt độ cao đồng thời làm giảm giới hạn bền vật liệu, làm tăng nguy cơ hư hỏng của các chi tiết này.

Những kết quả này là cơ sở để giới hạn tỷ số nén của động cơ trong quá trình cải tiến bằng cách tăng tỷ số nén, nhằm đảm bảo đồng thời hiệu quả làm việc và độ bền của các chi tiết.

Tài liệu tham khảo:

1. J. G. Alotaibi, Ay. E. Alajmi, T. Alsaeed, S. H. Al-Lwayzy and B. F. Yousif. (2024). *On the Influence of Engine Compression Ratio on Diesel Engine Performance and Emission Fueled with Biodiesel Extracted from Waste Cooking Oil*. Energies 17(15).
2. K. Wittek, K. Geiger, F. Andert, J. Martins, M. et al. (2017). *An Overview of VCR Technology and Its Effects on a Turbocharged DI Engine Fueled with Ethanol and Gasoline*. SAE Technical. DOI: https://doi.org/10.4271/2017-36-0357
3. X. P. Nguyen, B. J. Bora, A. I. Ölçer, P. Sharma, V. N. Nguyen, and D. N. Cao. (2025). "*Effect of compression ratio on dual-fuel diesel engine fueled with ammonia: Experimental and statistical analysis*". Process Safety and Environmental Protection, vol. 195. DOI: https://doi.org/10.1016/j.psep.2025.106810.
4. D. Q. Đức and N. V. Dương. (2025). *Nghiên cứu tăng tỷ số nén nhằm tăng hiệu suất động cơ tàu thủy*. Tạp chí Cơ khí Việt Nam, pp. 186-192.

5. N. V. Dương and P. X. Phương. (2023). *Determining thermal equivalent boundary conditions for piston surface*. Transport and Communications Science Journal, no. Special issue, pp. 452-464.
6. N. V. Dương, P. X. Phương, and L. T. Dương. (2025). *Xây dựng mô hình xác định điều kiện biên nhiệt trên pít tông-xilanh động cơ diesel 4 kỳ phun nhiên liệu trực tiếp*. Journal of Science Technique, vol. 20, no. 01, pp. 5-18.
7. Đ. T. Thắng. (2016). Tính toán sức bền các chi tiết của động cơ đốt trong bằng phương pháp phần tử hữu hạn. Hà Nội: Nhà xuất bản Quân đội nhân dân Việt Nam.
8. *Tài liệu thuyết minh kỹ thuật và hướng dẫn sử dụng động cơ 3D12*. Cục Kỹ thuật Hải quân.

Investigation of the influence of compression ratio increase on stress and deformation in the piston–cylinder assembly of marine diesel engines

Abstract: *This study investigates the effect of the compression ratio on the thermal state, stress, and deformation of the piston–cylinder assembly of the four-stroke 3D12 diesel engine used as the main engine on ships. The computational model for determining the stress and deformation of the piston–cylinder assembly was developed using the finite element method in Ansys Workbench. The results show that increasing the compression ratio leads to a significant rise in the temperature and stress of the piston and cylinder. When the compression ratio increases from $\varepsilon = 14$ to $\varepsilon = 16$, the maximum temperature of the piston increases by about 18,66 K on average, while the temperature of the inner surface of the cylinder increases by approximately 10,14 K. The thermo-mechanical stress also increases markedly: at the fillet below the piston crown, the stress rises from 105,06 MPa to 130,11 MPa, while at the upper supporting shoulder of the cylinder it increases from approximately 462,37 MPa to 511,95 MPa. The increase in temperature and stress significantly changes the deformation of the piston and cylinder, thereby directly affecting the reliability and service life of these components.*

Keywords: *Compression ratio; Thermal state; Stress; Piston; Cylinder.*

Nghiên cứu đánh giá hiệu quả xử lí soot của bộ scrubber lắp trên động cơ diesel kamaz 740 khi sử dụng hỗn hợp nhiên liệu chứa ôxy.

Nguyễn Bá Hiếu[1], Phạm Xuân Phương[2]

[1]Hệ QLHV Sau đại học, [2]Viện Cơ khí động lực

Email: nguyenbahieu01022000@gmail.com, Tel: 034.900.6777

Tóm tắt

Động cơ diesel được sử dụng rộng rãi nhờ hiệu suất nhiệt cao và độ tin cậy lớn, tuy nhiên phát thải chất thải dạng hạt, đặc biệt là bồ hóng (soot) của động cơ diesel vẫn là vấn đề môi trường đáng quan tâm. Việc sử dụng nhiên liệu chứa ôxy kết hợp với các bộ xử lý khí thải trên đường thải được xem là những giải pháp hiệu quả nhằm giảm phát thải soot cho động cơ diesel. Bài báo này trình bày phương pháp nghiên cứu thực nghiệm để đánh giá hiệu quả xử lí soot của bộ scrubber lắp trên động cơ diesel Kamaz 740 khi động cơ sử dụng các hỗn hợp nhiên liệu diesel pha trộn với biodiesel và diesel pha trộn với oxymethylene ethers (OMEx), với tỷ lệ khối lượng ôxy trong các hỗn hợp lần lượt là 2% và 5%, ở các chế độ tải khác nhau. Kết quả thử nghiệm cho thấy việc sử dụng hỗn hợp nhiên liệu chứa ôxy giúp giảm sự hình thành soot trong quá trình cháy và khi kết hợp với bộ scrubber, phát thải soot tiếp tục được giảm rõ rệt ở tất cả các chế độ tải so với trường hợp động cơ sử dụng nhiên liệu diesel truyền thống.

Từ khoá: *Scrubber, biodiesel, OMEx, nhiên liệu chứa ô xy, soot, hiệu quả giảm phát thải.*

1. Đặt vấn đề

Động cơ diesel hiện nay vẫn giữ vai trò chủ đạo trong lĩnh vực vận tải hạng nặng, máy nông nghiệp và tàu thủy nhờ những ưu điểm nổi bật như hiệu suất cao và độ tin cậy lớn. Tuy nhiên, bên cạnh những lợi ích về kinh tế và kỹ thuật, khí thải từ động cơ diesel, đặc biệt là các chất thải dạng hạt mà thành phần chính là soot cũng gây ảnh hưởng nghiêm trọng đối với môi trường và sức khỏe con người. Sự hình thành soot chủ yếu do quá trình cháy khuếch tán không hoàn toàn của nhiên liệu trong các vùng hòa trộn kém, giàu nhiên liệu và thiếu ôxy trong buồng cháy.

Trong bối cảnh các tiêu chuẩn khí thải ngày càng chặt chẽ, việc tìm kiếm các giải pháp nhằm giảm phát thải soot mà không ảnh hưởng đến hiệu suất động cơ là một thách thức lớn đối với các nhà nghiên cứu. Hiện nay, hai hướng tiếp cận chính đang được tập trung phát triển. Nhóm giải pháp thứ nhất là nhóm các giải pháp giảm phát thải soot ngay tại nguồn phát sinh, trong đó có giải pháp sử dụng nhiên liệu chứa ôxy như biodiesel và OMEx. Việc bổ sung ôxy nội tại vào cấu trúc phân tử nhiên liệu giúp cải thiện quá trình oxy hóa bồ hóng ngay tại vùng lõi tia phun, nơi thường xuyên thiếu khí cục bộ. Theo Yusop và các cộng sự [1], nồng độ soot của động cơ diesel giảm rõ rệt khi tăng tỷ lệ biodiesel pha trộn trong nhiên liệu ở các chế độ tải khác nhau. Nghiên cứu của Lin và các cộng sự [2], đã chỉ ra rằng nhiên liệu diesel pha OME3 có khả năng làm giảm đáng kể phát thải soot của động cơ diesel, đặc biệt ở các chế độ tải trung bình. Nhóm giải pháp thứ hai là nhóm giải pháp xử lý sau đường thải bằng cách sử dụng các thiết bị xử lý khí thải như bộ lọc hạt diesel hoặc bộ scrubber. Trong đó, bộ scrubber được đánh giá là giải pháp tiềm năng cho các dòng động cơ công suất lớn nhờ khả năng "bắt giữ" các hạt soot và giảm nhiệt độ khí xả hiệu quả. Nghiên cứu thực nghiệm của Nguyễn Trung Kiên và các cộng sự [4], đã chứng minh hiệu quả xử lý soot của bộ scrubber kiểu ướt khi lắp đặt trên động cơ diesel, kết quả cho thấy bộ scrubber có khả năng giảm phát thải dạng hạt cho động cơ diesel từ 40- 80 %, tuỳ thuộc vào từng chế độ tải.

Trong bài báo này, tác giả nghiên cứu thực nghiệm để đánh giá hiệu quả xử lý soot của bộ scrubber lắp trên động cơ Kamaz 740. Nghiên cứu thực hiện với các hỗn hợp nhiên liệu được pha trộn với diesel từ hai nguồn gốc khác nhau là biodiesel và OMEx. Kết quả nghiên cứu là cơ sở khoa học quan trọng cho việc lựa chọn và kết hợp các giải pháp giảm phát thải soot cho động cơ diesel trong thực tiễn.

2. Nghiên cứu thử nghiệm

2.1. Chuẩn bị nhiên liệu

Nhiên liệu sử dụng trong nghiên cứu là diesel thương mại D0, biodiesel B100, Oxymethylene Ether bậc 3 (OME_3), trong đó D0 là nhiên liệu diesel truyền thống (loại 0,05%S, bán trên thị trường); B100 là nhiên liệu diesel sinh học gốc, được chiết xuất từ bã thải của dầu cọ (là sản phẩm của Đề tài NCKH & PTCN cấp Quốc gia: *"Nghiên cứu sử dụng nhiên liệu diesel sinh học cho phương tiện cơ giới quân sự",* Mã số ĐT.06.12/NLSH [6]; OME_3 là nhiên liệu tổng hợp chứa oxy, được cung cấp bởi Công ty Shandong Bain New Material Co., Ltd. (Tế Nam, Trung Quốc) [5] và được sử dụng trực tiếp cho thí nghiệm, không qua xử lý bổ sung. Tỷ lệ phần trăm khối lượng của ôxy trong B100 và OME3 lần lượt là 10,99% và 47%.

Bảng 1. Thuộc tính nhiên liệu [5] [6]

Thông số	Loại nhiên liệu		
	Diesel	**B100**	**OME_3**
Công thức hoá học đại diện	$C_{13.9}H_{24.86}O_{0.01}$	$C_{18.94}H_{36.03}O_2$	$C_5H_{12}O_4$
Khối lượng phân tử [g.mol-l]	191,8	295,31	136,15
C [%]	86,97	78,13	44,1
H [%]	12,96	12,38	8,9
O [%]	0,07	10,99	47
Khối lượng riêng ở $15°C$,$[kg/m^3]$	839,6	869,3	1030,5
Độ nhớt động học ở $15°C$, [cSt]	3,11	4,6	1,08
Nhiệt trị thấp, [MJ/kg]	42,91	37,39	19,6
Trị số xê tan	52,4	66,9	71

Các hỗn hợp nhiên liệu chứa ôxy dùng cho thử nghiệm được pha trộn giữa diesel và B100, diesel và OME_3 để đạt được tỷ lệ ôxy 2% và 5% theo khối lượng lần lượt được ký hiệu là DB2, DB5, O3D2, O3D5. Tỷ lệ pha trộn được tính toán dựa trên phương pháp cân bằng oxy, sử dụng hàm lượng oxy đặc trưng của từng loại nhiên liệu, theo các công trình đã công bố trong tài liệu [3], nhằm đảm bảo các hỗn hợp so sánh có cùng mức oxy hóa học. Các hỗn hợp nhiên liệu sau khi pha trộn được khuấy trộn đồng đều và ổn định trước khi tiến hành thí nghiệm.

2.2. Bố trí hệ thống thử nghiệm

Các thử nghiệm được tiến hành trên bệ thử tại phòng thí nghiệm bộ môn Động cơ, Viện cơ khí động lực, Học viện Kỹ thuật quân sự. Bệ thử được trang bị động cơ Kamaz 740.31- 240, đây là động cơ diesel 4 kỳ, 8 xy lanh bố trí chữ V, hệ thống phun nhiên liệu kiểu cơ khí. Các thông số khác của động cơ được trình bày trong bảng 2. Trên đường ống xả của động cơ được lắp với bộ xử lý khí thải scrubber SC-21, là sản phẩm của đề tài nghiên cứu khoa học[tên đề tài], các thông số của bộ scrubber được trình bày trong bảng 3. Sơ đồ bố trí của hệ thống thử nghiệm được trình bày trên hình 2.

Bảng 2. Thông số kỹ thuật động cơ Kamaz 740 [7]

Thông số	Trị số	Thông số	Trị số
Số xi lanh	8	Công suất định mức, (kW)	176
Đường kính xi lanh (mm)	120	Tốc độ vòng quay trục khuỷu ứng với công suất định mức [vòng/phút]	2200
Hành trình piston (mm)	120	Góc mở sớm xupap nạp trước ĐCT (GQTK)	15 ± 2
Tỷ số nén	16,5	Góc đóng muộn xupap nạp sau ĐCD (GQTK)	40 ± 2
Suất tiêu hao nhiên liệu (g/kWh)	238	Góc mở sớm xupap thải trước ĐCD (GQTK)	40 ± 2
Hệ thống phun nhiên liệu	Cơ khí	Góc đóng muộn xupap thải sau ĐCT (GQTK)	15 ± 2

Bảng 3. Thông số kỹ thuật bộ scrubber SC21 [8]

Thông số	Trị số	Thông số	Trị số
Kích thước bao (DxRxC) [m]	0,7x0,5x1,4	Lưu lượng khí thải đầu vào, [kg/s]	0,0694
Diện tích tháp phun [m^2]	0,142	Lưu lượng chất lỏng [kg/phút]	0,133
Đường kính tháp phun [m]	0,09	Dung dịch trong bộ scrubber	Nước
Chiều cao tháp phun [m]	0,77	Tỷ số L/G [kg/kg]	0,032

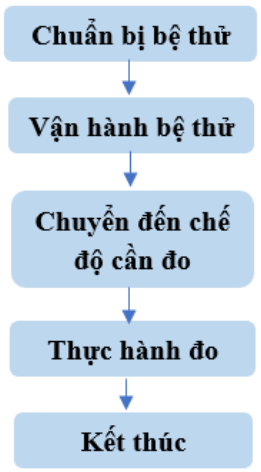
2.3. Chế độ thử nghiệm

Các thí nghiệm được thực hiện ở tốc độ động cơ không đổi 1600 vòng/phút. Động cơ được khảo sát tại các chế độ tải tương ứng với mô men 0 Nm, 30 Nm, 60 Nm và 90 Nm, nhằm đánh giá đặc tính phát thải soot trong điều kiện làm việc điển hình của động cơ diesel.

Đối với bộ scrubber, lưu lượng chất lỏng phun được giữ ổn định trong suốt quá trình thử nghiệm, tương ứng với tỷ số L/G xấp xỉ 0,03 kg/kg, đây là giá trị phù hợp cho các bộ scrubber kiểu ướt xử lý soot trong khí thải động cơ. Các thông số vận hành khác của động cơ được duy trì không đổi nhằm đảm bảo tính so sánh giữa các nhiên liệu và các chế độ tải khác nhau.

2.4. Phương pháp thử nghiệm

Trước khi tiến hành thu thập dữ liệu, động cơ được khởi động và vận hành cho đến khi đạt trạng thái ổn định về nhiệt độ nước làm mát và dầu bôi trơn. Tốc độ quay và tải trọng được điều chỉnh tăng dần đến các chế độ thử nghiệm xác định. Toàn bộ các thông số về tốc độ quay và mô-men xoắn được kiểm soát thông qua hệ thống đo tải của bệ thử, đảm bảo duy trì sai số ở mức tối thiểu trong suốt quá trình đo đạc. Để đánh giá chi tiết đặc tính cháy, cảm biến áp suất được lắp trực tiếp trên nắp xi lanh để đo diễn biến áp suất trong buồng cháy. Tín hiệu áp suất được thu thập đồng bộ với góc quay trục khuỷu thông qua hệ thống thu thập và xử lý dữ liệu. Dữ liệu áp suất được xử lý bằng phần mềm chuyên dụng để tính toán các thông số đặc trưng của quá trình cháy, trong đó các giá trị được lấy trung bình của 250 chu trình liên tiếp nhằm đảm bảo độ tin cậy của kết quả đo.



Hình 1. Các bước tiến hành thử nghiệm

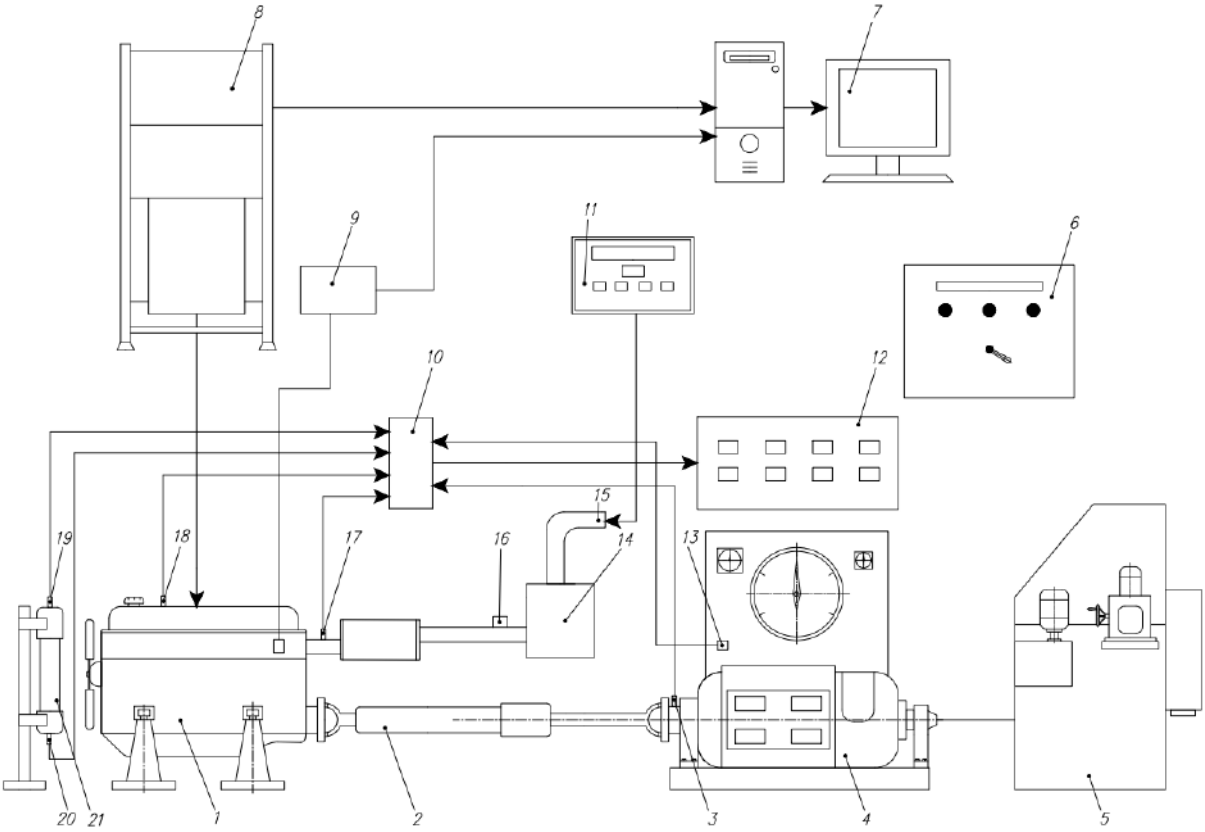
Khí thải sau khi đi ra cửa xả động cơ được dẫn qua bộ scrubber SC 21 để xử lý. Độ mờ khói (N, %) được xác định bằng thiết bị đo chuyên dụng OPUS 50 dựa trên nguyên lý hấp thụ quang học được đo tại 2 vị trí (15) và (16) nhằm đánh giá xu hướng phát thải soot của động cơ khi thay đổi loại nhiên liệu sử dụng và đánh giá hiệu quả giảm phát thải của bộ scrubber. Giá

trị độ mờ khói tỷ lệ thuận với nồng độ hạt bồ hóng có trong khí thải, từ đó cho phép xác định gián tiếp xu hướng phát thải soot của động cơ.

Hiệu suất xử lý của bộ scrubber (η)được xác định theo công thức:

$$\eta = \frac{N_{\text{trước}} - N_{\text{sau}}}{N_{\text{trước}}} \times 100\%$$

Trong đó: $N_{\text{trước}}$ và N_{sau} lần lượt là độ mờ khói của khí thải sau khi ra khỏi bộ scrubber.



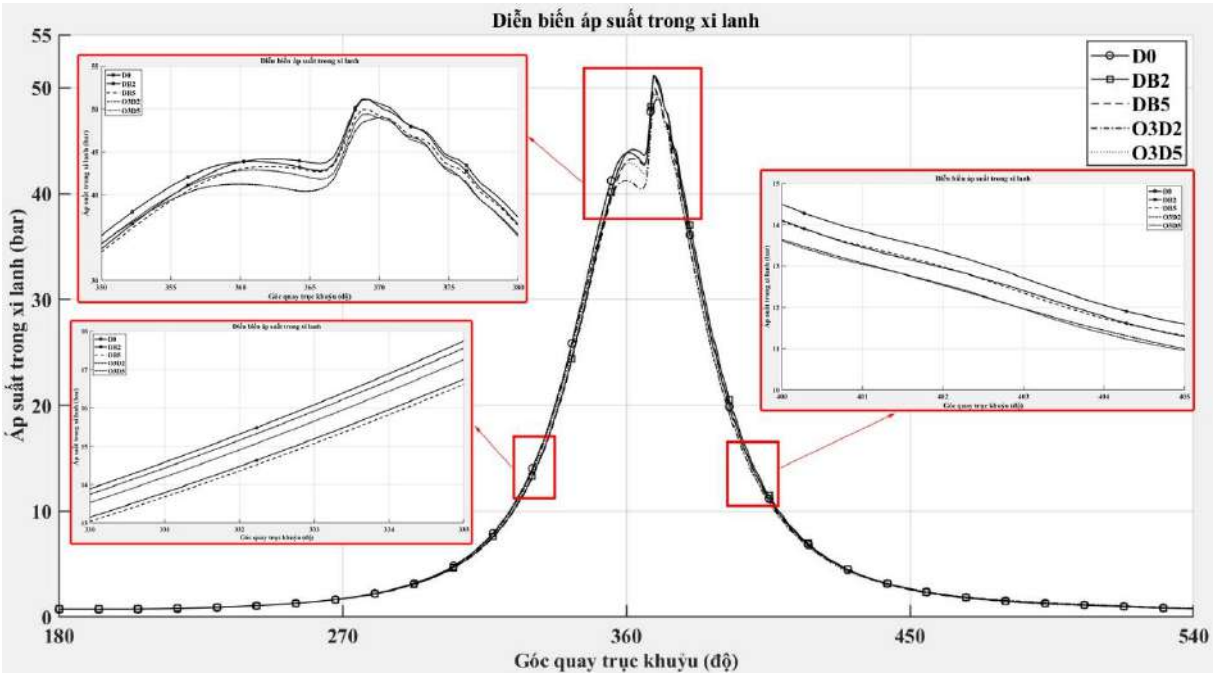
Hình 2. Sơ đồ bố trí các thiết bị của hệ thống thử nghiệm

1. Động cơ thử; 2. Trục các đăng; 3. Cảm biến tốc độ; 4. Phanh điện xoay chiều ba pha; 5. Biến trở kiểu chất lỏng; 6. Tủ điện điều khiển; 7. Máy tính; 8. Thiết bị đo mức tiêu thụ nhiên liệu AVL Fuel Balance 733S; 9. Hộp ghi dữ liệu trong xylanh AVL Indiset 620; 10. Pa nô điều khiển và hiển thị; 11. Hộp đấu dây cáp tín hiệu; 12. Thiết bị đo độ khói Opus 50; 13. Loadcell; 14. Bộ scrubber SC21; 15,16. Vị trí đo độ khói; 17. Cảm biến nhiệt độ khí thải; 18. Cảm biến áp suất dầu bôi trơn; 19. Cảm biến nhiệt độ nước làm mát đầu ra; 20. Cảm biến nhiệt độ nước làm mát đầu vào; 21. Két nước làm mát.

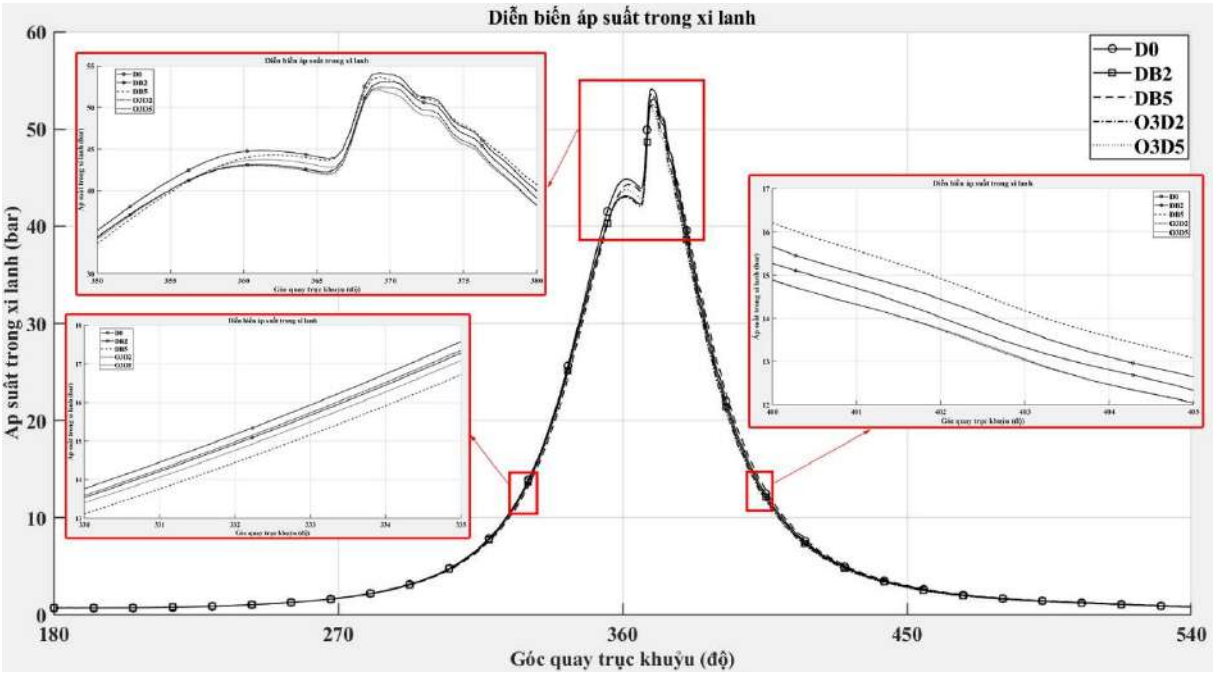
3. Kết quả và thảo luận

3.1. Diễn biến áp suất trong xy lanh động cơ khi sử dụng các loại hỗn hợp nhiên liệu chứa ôxy khác nhau

Hình 3, hình 4, hình 5 và hình 6 trình bày diễn biến áp suất trong xi lanh theo góc quay trục khuỷu tại các chế độ tải 0Nm, 30 Nm, 60Nm và 90 Nm khi động cơ sử dụng các nhiên liệu D0 và các hỗn hợp nhiên liệu chứa oxy. Ở các chế độ tải, dạng đường cong áp suất của các nhiên liệu là tương đồng, cho thấy việc phối trộn D0 với biodiesel và OME_3 không làm thay đổi bản chất quá trình cháy trong xi lanh. Tuy nhiên, có thể nhận thấy sự khác biệt rõ rệt về giá trị áp suất tuyệt đối tại các giai đoạn nén, cháy, giãn nở.

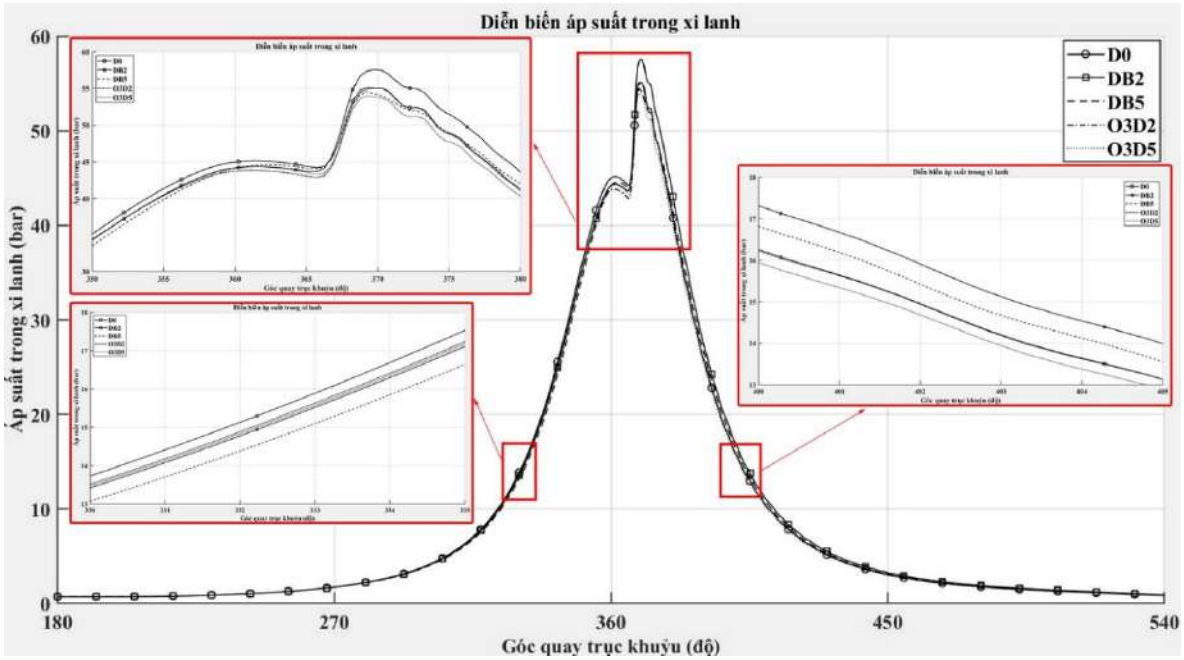


Hình 3 . Diễn biến áp suất trong xy lanh tại tốc độ 1600 v/p, tải 0Nm

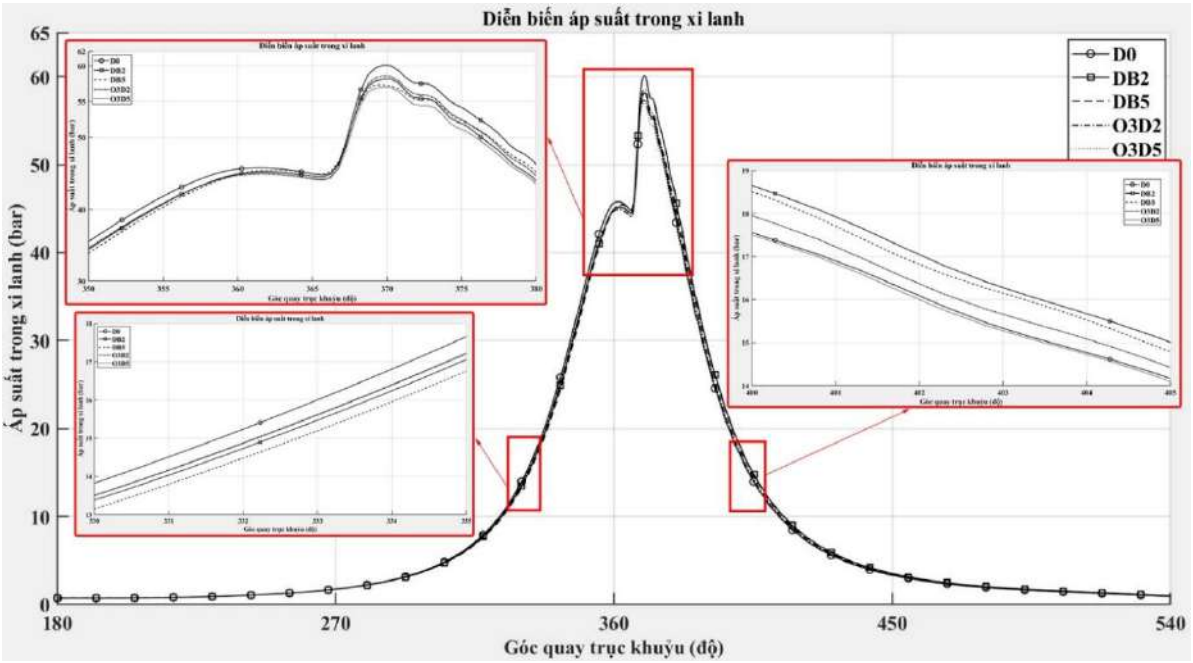


Hình 4. Diễn biến áp suất trong xy lanh tại tốc độ 1600 v/p, tải 30Nm

Tại chế độ không tải và tải thấp (30 Nm), diễn biến áp suất xi lanh của các loại nhiên liệu hỗn hợp chứa oxy (DB2, DB5, O3D2 và O3D5) cho thấy sự tương đồng chặt chẽ về quy luật biến thiên so với nhiên liệu D0, khẳng định tính tương thích cao của động cơ Kamaz 740 với các loại nhiên liệu này. Mặc dù áp suất cực đại của các hỗn hợp nhiên liệu chứa oxy có sự sụt giảm nhẹ do ảnh hưởng từ nhiệt trị thấp, nhưng diễn biến áp suất trong kỳ giãn nở vẫn duy trì được trị số ổn định. Điều này tạo điều kiện thuận lợi cho việc kéo dài pha cháy khuếch tán, kết hợp với sự hiện diện của oxy nội tại trong cấu trúc phân tử giúp thúc đẩy mạnh mẽ quá trình oxy hóa muội than ngay trong xi lanh.



Hình 5. Diễn biến áp suất trong xy lanh tại tốc độ 1600 v/p, tải 60Nm

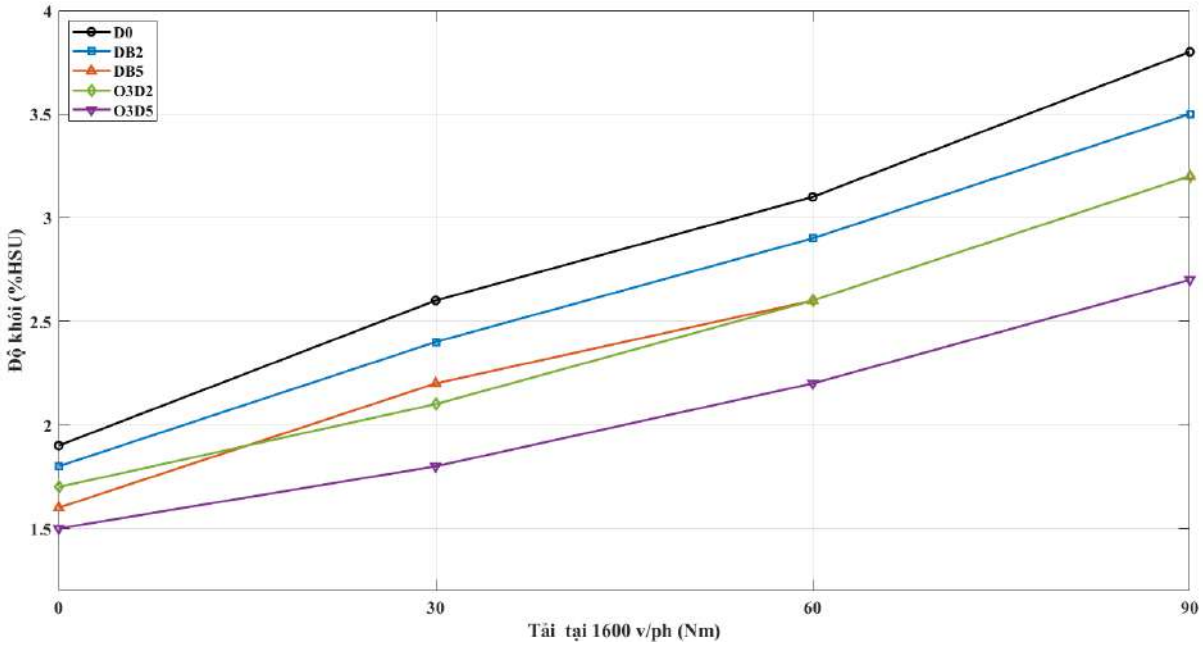


Hình 6. Diễn biến áp suất trong xy lanh tại tốc độ 1600 v/p, tải 90Nm

Tại chế độ tải 60 Nm và 90 Nm, sự phân tách giữa các đường áp suất trở nên rõ rệt hơn, đặc biệt là ở vùng đỉnh áp suất cực đại và giai đoạn cháy khuếch tán. Kết quả cho thấy các hỗn hợp nhiên liệu chứa ôxy có áp suất cực đại thấp hơn so với D0, giúp giảm bớt ứng suất nhiệt cục bộ trong buồng cháy, một yếu tố quan trọng để hạn chế sự hình thành các hạt muội. Ở mức tải cao 90 Nm, mặc dù lượng nhiên liệu phun vào lớn dễ gây ra tình trạng thiếu oxy cục bộ, nhưng các hỗn hợp O3D5 và DB5 với hàm lượng oxy nội tại dồi dào đã thúc đẩy quá trình cháy diễn ra hoàn thiện hơn. Sự ổn định của áp suất trong kỳ giãn nở minh chứng cho việc oxy hóa liên tục các hạt soot ngay trong lòng ngọn lửa, giúp giảm thiểu đáng kể lượng muội than thoát ra từ cửa xả động cơ. Điều này không chỉ cải thiện chất lượng khí thải tại nguồn mà còn giúp giảm tải cho bộ scrubber, cho phép hệ thống duy trì hiệu suất xử lý cao ngay cả trong điều kiện động cơ hoạt động ở công suất lớn.

3.2. Ảnh hưởng của nhiên liệu chứa ôxy đến nồng độ soot

Hình 7 và bảng 4 cho thấy ảnh hưởng rõ rệt của nhiên liệu chứa ôxy đến nồng độ soot (thể hiện qua độ khói %HSU) của động cơ Kamaz-740 tại tốc độ 1600 vòng/phút, trước khi khí thải được dẫn qua bộ scrubber.



Hình 7. Độ khói của động cơ khi sử dụng các loại nhiên liệu trước khi đi vào bộ scrubber

Đối với tất cả các loại nhiên liệu khảo sát, độ khói tăng dần theo tải động cơ, phản ánh sự gia tăng lượng nhiên liệu phun và mức độ phát triển của các vùng cháy giàu ở tải cao. Tuy nhiên, tại cùng một chế độ tải, các hỗn hợp nhiên liệu chứa ôxy đều cho giá trị độ khói thấp hơn so với D0, cho thấy vai trò của ôxy sẵn có trong nhiên liệu đối với việc hạn chế hình thành soot ngay trong quá trình cháy.

Bảng 4. Độ khói của động cơ khi sử dụng các loại nhiên liệu trước khi đi vào bộ scrubber

Nhiên liệu	Chế độ	0 Nm	30 Nm	60 Nm	90 Nm
D0	Độ khói [%HSU]	1,9	2,6	3,1	3,8
	Thay đổi so với D0 [%]	0	0	0	0
DB2	Độ khói [%HSU]	1,8	2,4	2,9	3,5
	Thay đổi so với D0 [%]	5,3	7,7	6,5	7,9
DB5	Độ khói [%HSU]	1,6	2,2	2,6	3,2
	Thay đổi so với D0 [%]	15,8	15,4	16,1	15,8
O3D2	Độ khói [%HSU]	1,7	2,1	2,6	3,2
	Thay đổi so với D0 [%]	10,5	19,2	16,1	15,8
O3D5	Độ khói [%HSU]	1,5	1,8	2,2	2,7
	Thay đổi so với D0 [%]	21,1	30,8	29	28,9

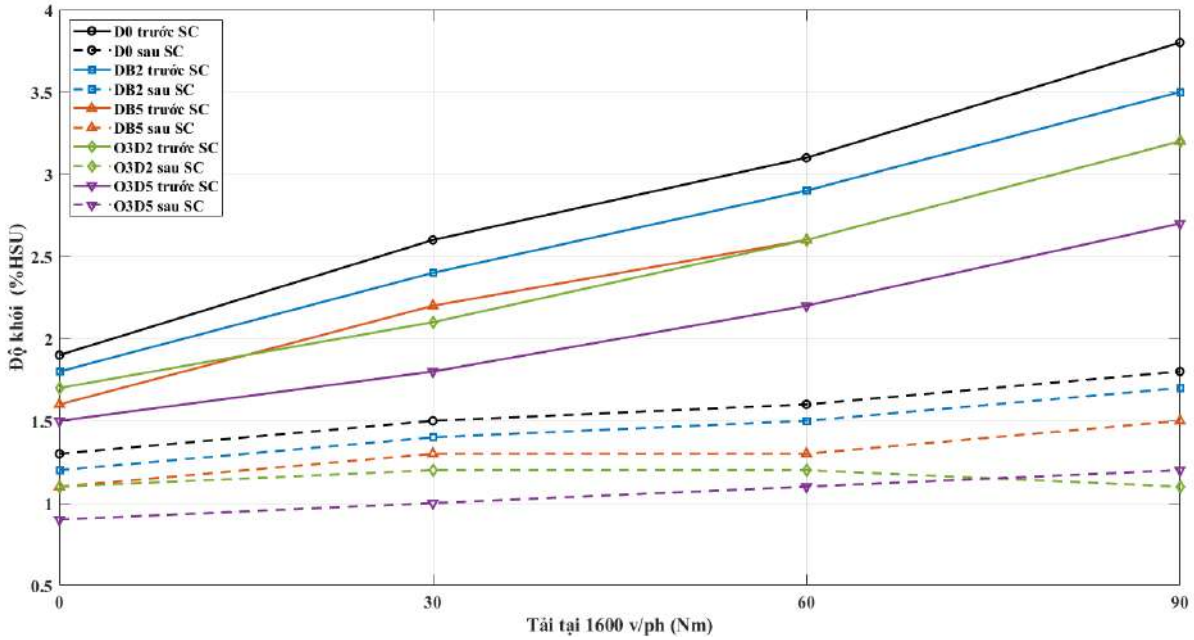
Ở mức pha trộn thấp, nhiên liệu DB2 làm giảm độ khói khoảng 5,3–7,9% so với D0, trong khi DB5 cho mức giảm ổn định hơn, đạt khoảng 15–16% trên toàn dải tải khảo sát. Điều này cho thấy khi hàm lượng ôxy trong nhiên liệu tăng, quá trình oxy hóa soot trong pha cháy khuếch tán được tăng cường, làm giảm lượng soot hình thành. Đối với các hỗn hợp nhiên liệu chứa OME_3, hiệu quả giảm độ khói thể hiện rõ rệt hơn so với biodiesel ở cùng mức hàm lượng ôxy. Cụ thể, O3D2 làm giảm độ khói từ 10,5% ở chế độ không tải lên đến khoảng 19,2% ở 30 Nm,

trong khi O3D5 cho hiệu quả giảm cao nhất, đạt từ 21,1% ở 0 Nm đến khoảng 28,9–30,8% ở các chế độ tải trung bình và cao.

Xu hướng này có thể được giải thích bởi đặc tính hóa học của nhiên liệu chứa OME3, vốn không chứa liên kết C–C trong cấu trúc phân tử và có hàm lượng ôxy cao, giúp hạn chế sự hình thành soot và thúc đẩy quá trình oxy hóa soot trong giai đoạn cháy muộn.

3.3. Đánh giá hiệu quả xử lí soot của bộ scrubber khi sử dụng hỗn hợp nhiên liệu chứa ôxy

Hình 8 và bảng 5 thể hiện độ khói của động cơ sau khi đi qua bộ scrubber khi sử dụng các loại nhiên liệu khác nhau ở các chế độ tải khác nhau.



Hình 8. Độ khói của động cơ khi sử dụng các loại nhiên liệu sau khi đi qua bộ scrubber

Bảng 5. Độ khói của động cơ khi sử dụng các loại nhiên liệu sau khi đi qua bộ scrubber

Nhiên liệu	Chế độ	0 Nm	30 Nm	60 Nm	90 Nm
D0	Độ khói [%HSU]	1,9	2,6	3,1	3,8
	Thay đổi so với D0 [%]	0	0	0	0
DB2	Độ khói [%HSU]	1,2	1,4	1,5	1,7
	Thay đổi so với D0 [%]	36,8	46,2	51,6	55,3
DB5	Độ khói [%HSU]	1,1	1,3	1,3	1,5
	Thay đổi so với D0 [%]	42,1	50	58,1	60,5
O3D2	Độ khói [%HSU]	1,1	1,2	1,2	1,4
	Thay đổi so với D0 [%]	42,1	53,8	61,3	63,2
O3D5	Độ khói [%HSU]	0,9	1	1	1,2
	Thay đổi so với D0 [%]	52,6	61,5	66,2	68,4

Kết quả thí nghiệm cho thấy bộ scrubber có hiệu quả rõ rệt trong việc xử lý soot của động cơ khi sử dụng D0 và các hỗn hợp nhiên liệu chứa ôxy. Ở tất cả các chế độ tải khảo sát, độ khói sau scrubber đều giảm mạnh so với trước xử lý, đồng thời xu hướng giảm trở nên rõ rệt hơn khi tải động cơ tăng. Đối với diesel, độ khói sau scrubber giảm khoảng 36,8–55,3% trong dải tải từ 0 đến 90 Nm. Khi sử dụng nhiên liệu chứa ôxy, hiệu quả xử lý của scrubber được cải thiện đáng kể; cụ thể, DB5 và O3D2 cho mức giảm độ khói lần lượt khoảng 42,1–63,2%, trong khi hỗn hợp O3D5 đạt hiệu quả cao nhất, với mức giảm từ 52,6% ở chế độ không tải lên đến khoảng

68,4% ở tải 90 Nm. Hiệu suất xử lý tăng theo tải động cơ, phù hợp với sự gia tăng lưu lượng khí thải và nồng độ hạt đầu vào, làm tăng xác suất va đập và bắt giữ hạt trong scrubber. Khi sử dụng các hỗn hợp nhiên liệu chứa ôxy, đặc biệt là các hỗn hợp O3D2 và O3D5, mức giảm độ khói sau scrubber cao hơn so với D0 do lượng soot hình thành trong quá trình cháy được hạn chế ngay từ nguồn phát sinh và đặc tính hạt soot thuận lợi hơn cho quá trình xử lý. Kết quả này cho thấy sự kết hợp giữa nhiên liệu chứa ôxy và bộ scrubber tạo ra hiệu ứng cộng hưởng trong việc giảm phát thải soot, trong đó nhiên liệu chứa ôxy làm giảm lượng soot hình thành tại nguồn, còn scrubber tiếp tục xử lý hiệu quả phần soot còn lại trên đường thải, đặc biệt ở các chế độ tải trung bình và cao.

4. Kết luận

Bài báo trình bày phương pháp nghiên cứu thực nghiệm để đánh giá hiệu quả xử lí soot của bộ scrubber lắp trên động cơ diesel Kamaz 740 khi động cơ sử dụng các hỗn hợp nhiên liệu diesel pha trộn với biodiesel và diesel pha trộn OME3 với tỷ lệ khối lượng ôxy trong các hỗn hợp lần lượt là 2% và 5%, ở các chế độ tải khác nhau. Nghiên cứu đã xác lập được quy luật biến thiên áp suất trong xi lanh, chứng minh rằng các hỗn hợp nhiên liệu chứa ôxy không làm thay đổi bản chất quá trình cháy và đảm bảo động cơ vận hành ổn định trên toàn dải tải khảo sát. Bài báo đã chỉ ra xu hướng giảm phát thải soot tại nguồn khi tăng tỷ lệ oxy nội tại trong nhiên liệu, đặc biệt là hiệu quả vượt trội của hỗn hợp O3D5 so với các mẫu còn lại. Đồng thời, nghiên cứu đã thực nghiệm và xác định được hiệu suất xử lý soot của bộ scrubber SC21, đạt mức giảm độ mờ khói lên tới 68,4% ở chế độ tải cao. Quan trọng nhất, bài báo đã chứng minh được tính khả thi và hiệu quả cộng hưởng của giải pháp kết hợp giữa nhiên liệu chứa ôxy và thiết bị xử lý khí thải kiểu scrubber. Đây là cơ sở dữ liệu khoa học quan trọng cho việc lựa chọn các giải pháp giảm thiểu ô nhiễm môi trường cho các dòng động cơ diesel hạng nặng hiện nay.

Tuy nhiên, bài báo cũng chỉ mới nghiên cứu hiệu quả xử lí soot của bộ scrubber sử dụng chất lỏng là nước và tỉ số L/G cố định. Hướng nghiên cứu tiếp theo của bài báo được xác định là khảo sát ảnh hưởng của các thông số vận hành hệ thống xử lý như thay đổi tỷ lệ L/G (lưu lượng chất lỏng/lưu lượng khí), và thử nghiệm các loại dung dịch hấp thụ khác nhau thay vì chỉ dùng nước để nâng cao hơn nữa hiệu suất bắt giữ hạt soot.

Tài liệu tham khảo

1. Yusop, A. F., Mamat, R., Yusaf, T., Najafi, G., Yasin, M. H. M., & Khathri, A. M. (2018). Analysis of particulate matter (PM) emissions in diesel engines using palm oil biodiesel blended with diesel fuel. *Energies*, 11(4), 1030.
2. Lin, Q., Tay, K. L., Yu, W., Zong, Y., Yang, W., Rivellini, L.-H., Ma, M., & Lee, A. K. Y. (2021). Polyoxymethylene dimethyl ether 3 ($PODE_3$) as an alternative fuel to reduce aerosol pollution. *Journal of Cleaner Production*, 285, 125042.
3. Huang, Z. H., Lu, X. H., & Li, W. F. (2022). Oxygenated diesel fuels and their effect on PM emissions. *Applied Sciences*, 12(15), 7709.
4. Kien T. N, Manh D. V., Duong T. L and Phuong X. P. (2023). A feasibility study on the exhaust-gas treatment efficiency of a wet-scrubber system equipped in a diesel engine: A focus on smoke emission. *Journal of Military Science and Technology*, 89, 153–159.
5. Shandong Bain New Material Co., Ltd. (2025). *Commercial invoice for dimethoxydimethyl ether* (CAS No. 628-90-0; 13353-03-2) (Invoice No. BAIN2024112701). Jinan, China.

6. Nguyễn Hoàng Vũ. (2014). *Báo cáo tổng kết đề tài NCKH&PTCN cấp Nhà nước: Nghiên cứu sử dụng nhiên liệu diesel sinh học (B10 và B20) cho phương tiện cơ giới quân sự* (Mã số ĐT.06.12/NLSH). Việt Nam.
7. Tổng Công ty Cổ phần KamAZ. (2017). *Bộ tài liệu hướng dẫn bảo dưỡng và sửa chữa ô tô KamAZ 43253, 43255, 65115, 65116, 65117, 6540, 65111* (Tiêu chuẩn Euro-2,-3,-4, Tập I). Hà Nội.
8. Phạm Xuân Phương. (2021). *Báo cáo tổng kết đề tài NCKH&PTCN cấp Nhà nước: Thiết kế chế tạo thử nghiệm thiết bị giảm phát thải cho động cơ diesel bằng công nghệ scrubber* (Mã số ĐT.25.21). Việt Nam.

Investigation of soot treatment efficiency of a scrubber installed on a kamaz 740 diesel engine using oxygenated fuel blends.

Abstract: *Diesel engines are widely used due to their high thermal efficiency and reliability; however, particulate matter emissions, particularly soot, remain a significant environmental concern. The use of oxygenated fuels in combination with exhaust after-treatment systems is considered an effective approach for reducing soot emissions from diesel engines. This paper presents an experimental investigation to evaluate the soot treatment efficiency of a scrubber installed on a KamAZ 740 diesel engine operating with diesel–biodiesel blends and diesel–oxymethylene ethers (OMEx) blends, in which the oxygen mass fractions in the fuels are 2% and 5%, respectively, under different engine load conditions. The experimental results indicate that increasing the oxygen content in the fuel reduces soot formation during the combustion process, and when combined with the scrubber system, soot emissions are further significantly reduced at all load conditions compared with conventional diesel fuel operation.*

Keywords: scrubber; biodiesel; OMEx; soot; emission reduction efficiency.

Xác định thực nghiệm các thông số đặc trưng phục vụ mô hình hóa và chẩn đoán hệ thống phun nhiên liệu diesel kiểu CommonRail theo mô hình

Nguyễn Tất Trọng, Nguyễn Hoàng Vũ
Trường Đại học Kỹ thuật Lê Quý Đôn
Email: nguyentrong2882@gmail.com

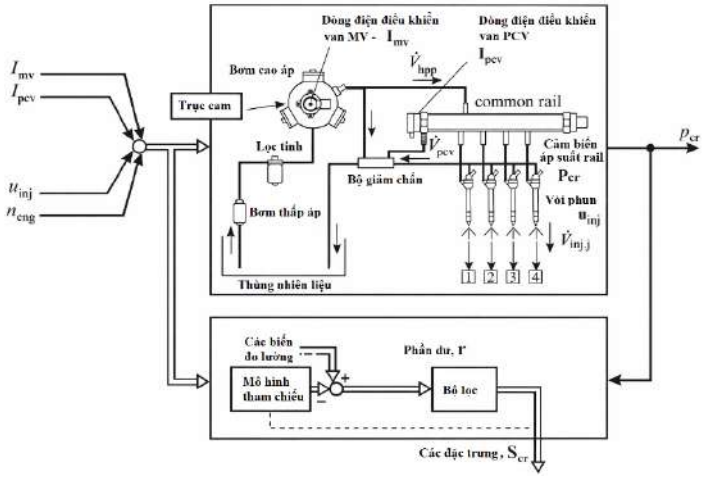
Tóm tắt

Bài báo trình bày kết quả thực nghiệm xác định các thông số đặc trưng phục vụ việc mô hình hóa và chẩn đoán theo mô hình (Model-based Diagnostics - MBD) cho hệ thống phun nhiên liệu (HTPNL) diesel kiểu CommonRail (CR) của động cơ Huyndai D4CB 2.5 TCI-A trên bệ thử con lăn. Các dữ liệu thực nghiệm bao gồm: tín hiệu của cơ cấu chấp hành (dòng điện điều khiển van định lượng, Metering Valve - MV và van kiểm soát áp suất, Pressure Control Valve - PCV), điều kiện vận hành (tốc độ động cơ, vị trí chân ga) và biến trạng thái hệ thống (áp suất rail). Các thông số này được thu thập đồng bộ từ thiết bị OBD-II (ELM327) và mạch đo dòng ACS712 qua bộ vi xử lý Arduino. Kết quả cho thấy ngoài sự phụ thuộc vào tốc độ động cơ và vị trí chân ga, áp suất rail còn phụ thuộc mạnh vào sự phối hợp điều khiển của hai van MV–PCV. Các đồ thị 2D và 3D thu được mô tả rõ vai trò của MV, PCV và chiến lược điều khiển của ECU động cơ trên dải vận hành. Những dữ liệu thu được là cơ sở quan trọng để xây dựng mô hình mô phỏng động học áp suất rail và ứng dụng phương pháp MBD cho HTPNL kiểu CR.

***Từ khóa**: Điều khiển áp suất rail; van định lượng (MV); van điều khiển áp suất (PCV); thông số đặc trưng; chẩn đoán theo mô hình (MBD).*

1. Đặt vấn đề

Trong những năm gần đây, kỹ thuật chẩn đoán theo mô hình (Model-Based Diagnostics - MBD) được xem là hướng tiếp cận hiện đại, hiệu quả cho các đối tượng có đặc tính động lực học phi tuyến phức tạp như hệ thống phun nhiên liệu (HTPNL) diesel kiểu CommonRail (CR), [1-3]. Phương pháp này xây dựng mô hình toán học mô tả động học của hệ thống, sau đó so sánh tín hiệu dự đoán với tín hiệu đo thực tế để tạo ra phần dư (Residual). Khi hệ thống làm việc bình thường, phần dư nằm trong giới hạn cho phép; khi xuất hiện hư hỏng, phần dư thay đổi theo quy luật đặc trưng cho từng dạng lỗi và được sử dụng làm tín hiệu chẩn đoán tình trạng kỹ thuật của hệ thống. Việc xây dựng mô hình sử dụng những thông số đặc trưng sẵn có trong chương trình làm việc của ECU động cơ là yêu cầu bắt buộc đối với kỹ thuật MBD. Cấu trúc mô hình chẩn đoán HTPNL diesel kiểu CR được trình bày trên Hình 1, [4].



Hình 1. Cấu trúc mô hình chẩn đoán HTPNL kiểu CR dựa trên phần dư, [4].

Sơ đồ cấu trúc mô hình chẩn đoán HTPNL kiểu CR bao gồm hai khối chức năng chính: hệ thống thực và mô hình tham chiếu. Các đại lượng đầu vào bao gồm các tín hiệu chấp hành là dòng điện điều khiển van định lượng (I_{mv}) và van kiểm soát áp suất (I_{pcv}), cùng các điều kiện vận hành như lượng phun (u_{inj}) và tốc độ động cơ (n_{eng}). Trong hệ thống thực, nhiên liệu được bơm sơ cấp hút từ thùng chứa qua lọc nhiên liệu đến bơm cao áp. Tại đây, van định lượng MV (còn gọi là van SCV) điều tiết lưu lượng nạp $\dot{V}_{hpp}$, sau đó nhiên liệu được nén đến áp suất cao đưa vào ống rail. Van điều khiển áp suất PCV (còn gọi là van RPCV) thực hiện duy trì áp suất bằng cách điều chỉnh lưu lượng xả hồi $\dot{V}_{pcv}$ qua bộ giảm chấn (damper) về thùng chứa. Áp suất rail đo được (p_{cr}) cùng các biến đo lường khác được đưa vào bộ so sánh với giá trị dự đoán từ mô hình tham chiếu (Reference model) để tính toán phần dư r. Sau khi qua bộ lọc (filters), các đặc trưng chẩn đoán (S_{cr}) được trích xuất để xác định tình trạng kỹ thuật của hệ thống.

Áp suất nhiên liệu trong ống rail được ECU kiểm soát thông qua các cơ cấu chấp hành để phù hợp với sự thay đổi về chế độ và điều kiện vận hành của động cơ. Trong bài báo này, van định lượng MV và van điều khiển áp suất PCV được xem xét là các đối tượng điều khiển chính. Áp suất rail có sự thay đổi lớn và nhanh theo tốc độ và tải của động cơ [5-7]. Để duy trì giá trị áp suất mục tiêu, MV và PCV phải phối hợp liên tục để điều chỉnh lưu lượng tức thời trong ống tích áp. Về bản chất, áp suất rail phản ánh trạng thái cân bằng giữa các quá trình: cung cấp nhiên liệu từ bơm cao áp, xả áp suất qua van PCV và phun nhiên liệu qua các vòi phun. Bất kỳ sự suy giảm hiệu suất thể tích của bơm, rò rỉ, hay sai lệch đặc tính làm việc của các cơ cấu chấp hành đều biểu hiện trực tiếp thông qua dao động, độ trễ tăng áp hoặc sai lệch áp suất rail. Vì vậy, áp suất rail không chỉ là biến điều khiển trung tâm mà còn là một biến trạng thái mang thông tin tổng hợp về tình trạng kỹ thuật của toàn bộ hệ thống [1], [3], [8].

Về phương thức điều khiển, MV và PCV là các van điện từ có tiết diện mở tỉ lệ với dòng điện điều khiển từ ECU. Khi ECU yêu cầu tăng áp suất, dòng điện cấp cho van PCV (I_{pcv}) tăng lên để đóng bớt đường xả nhiên liệu về thùng, đồng thời dòng điện qua van MV (Imv) được điều chỉnh nhằm cho phép lượng nhiên liệu nạp vào bơm nhiều hơn. Sự phối hợp nhịp nhàng này tạo ra đặc tính tăng áp suất gần như tuyến tính theo tốc độ động cơ và vị trí chân ga. Do đó, để mô hình hóa HTPNL kiểu CR phục vụ chẩn đoán MBD, các thông số đầu vào cần thiết bao gồm: tín hiệu chấp hành (dòng điện I_{mv}, I_{pcv}), điều kiện vận hành (tốc độ động cơ, vị trí chân ga) và biến trạng thái (áp suất rail p_{cr}).

Để xác định dòng điện qua các van MV, PCV, tốc độ động cơ, vị trí chân ga và áp suất rail trong toàn dải vận hành, có thể sử dụng tín hiệu từ các cảm biến được đo bằng thiết bị phù hợp kết nối với ECU của động cơ. Các phép đo này có thể thực hiện trên bệ thử động cơ (Engine Testbed) hoặc bệ thử con lăn (Chassis Dynamometer). Trong đó, phương pháp thử xe trên bệ thử con lăn có ưu điểm là tiết kiệm thời gian, kinh phí mà vẫn đảm bảo độ chính xác yêu cầu cho các phép đo thực nghiệm.

Bài báo trình bày kết quả thực nghiệm xác định các thông số đặc trưng bao gồm đặc tính làm việc của van MV và van PCV trên HTPNL diesel kiểu CR của động cơ Hyundai D4CB 2.5 TCI-A. Quá trình thực nghiệm được thực hiện trên bệ thử con lăn nhằm thu thập bộ dữ liệu thực tế về sự phối hợp điều khiển của ECU. Những kết quả thu được là cơ sở quan trọng để xây dựng mô hình mô phỏng động học áp suất rail và ứng dụng phương pháp chẩn đoán MBD cho HTPNL diesel kiểu CR trong các công bố tiếp sau của nhóm tác giả.

2. Đối tượng và trang thiết bị nghiên cứu

2.1. Đối tượng nghiên cứu

Đối tượng nghiên cứu là HTPNL kiểu CR trang bị trên động cơ diesel Hyundai D4CB 2.5 TCI-A, 4 xi lanh, 4 kỳ. Hệ thống này sử dụng bơm cao áp loại CP1-H với 3 pít tông hướng tâm, cho áp suất rail tối đa đạt 1600 bar. Cơ chế kiểm soát của hệ thống được thực hiện kép: sử dụng van định lượng MV để kiểm soát lưu lượng nạp vào bơm cao áp và van kiểm soát áp suất PCV để điều tiết áp suất trong ống rail. Cả hai loại van này đều thuộc kiểu van điện từ thường mở (Normally Open). Các thông số kỹ thuật chính của động cơ D4CB 2.5 TCI-A được trình bày chi tiết trong Bảng 1.

Bảng 1. Thông số kỹ thuật chính của động cơ Hyundai D4CB 2.5 TCI-A.

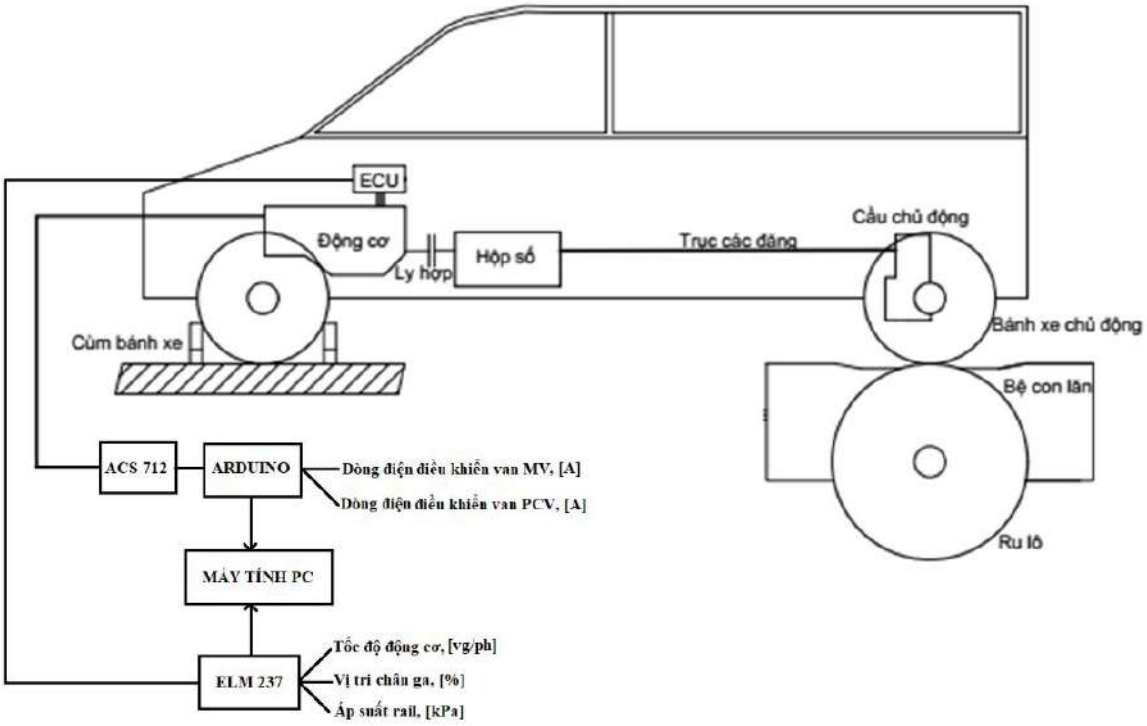
TT	Thông số	Giá trị
1	Kiểu động cơ	4 xi lanh, 1 hàng
2	Đường kính xi lanh (D) x Hành trình pít tông S, (mm)	91x96
3	Tỷ số nén	17,6
4	Bơm nhiên liệu thấp áp	Bơm điện, áp suất 5 bar
5	Bơm cao áp CP1-H với 3 pít tông hướng tâm	prail max=1600 bar
6	Van kiểm soát lưu lượng (MV)	Loại van điện từ, thường mở
7	Van kiểm soát áp suất (PCV)	Loại van điện từ, thường mở
8	Cảm biến áp suất rail	Khoảng làm việc (0-2000 bar)
9	Vòi phun	CRI 2.2, kiểu điện từ

2.2. Trang thiết bị thử nghiệm

Quá trình nghiên cứu thực nghiệm được thực hiện tại Phòng thử Ô tô thuộc Nhà máy Z157 (Cục Xe máy – Vận tải) trên xe Hyundai Starex. Việc lựa chọn bệ thử con lăn (Chassis Dynamometer) thay vì bệ thử động cơ (Engine Testbed) mang lại ưu điểm thiết thực về tính kinh tế, thời gian và hiệu quả triển khai. Phương pháp này cho phép đánh giá HTPNL kiểu CR trong điều kiện làm việc thực tế của toàn bộ hệ thống truyền động, mô phỏng chính xác tải trọng động và lực cản lăn tương tự khi vận hành trên đường. Điều này đặc biệt quan trọng đối với kỹ thuật MBD, vốn đòi hỏi dữ liệu thực nghiệm phải phản ánh sát thực chiến lược điều khiển của ECU trong các tình huống vận hành có tải và không tải.

Hệ thống đo lường được thiết lập nhằm thu thập đồng bộ các thông số đặc trưng từ hai nguồn dữ liệu khác nhau. Sơ đồ bố trí trang thiết bị được trình bày trong Hình 2, thể hiện sự kết nối giữa hệ thống chẩn đoán tích hợp của xe (OBD-II) và mạch đo dòng điện bổ sung. Cấu trúc hệ thống đo được phân tách dựa trên đặc tính động lực học của từng đại lượng:

Điều kiện vận hành và biến trạng thái: Tốc độ động cơ (n, vg/ph), vị trí chân ga (PWG, %) và áp suất rail (p_{cr}, kPa) được trích xuất trực tiếp từ ECU thông qua thiết bị ELM327 kết nối với cổng OBD-II. Mặc dù tần số lấy mẫu của giao thức này thấp hơn mạch đo chuyên dụng, nhưng các đại lượng trên đại diện cho quán tính vận hành và trạng thái tích áp trong ống rail vốn có tốc độ biến thiên phù hợp với các thuật toán chẩn đoán lỗi. Việc sử dụng dữ liệu trực tiếp từ ECU đảm bảo tính đồng bộ với giá trị thực tế mà thuật toán điều khiển của xe đang sử dụng.



Hình 2. Sơ đồ bố trí trang thiết bị thực nghiệm.

-Tín hiệu chấp hành: Cường độ dòng điện điều khiển I_{mv} và I_{pcv} là các tín hiệu có đặc tính biến thiên rất nhanh theo xung PWM từ ECU. Do đó, nhóm nghiên cứu sử dụng cảm biến dòng điện ACS712 kết nối với bo mạch vi xử lý Arduino để đo trực tiếp với độ phân giải cao, đảm bảo cung cấp đầu vào chính xác cho việc xây dựng mô hình động học áp suất rail có độ tin cậy cao.

Toàn bộ dữ liệu từ ELM327, Arduino được cài đặt cùng tốc độ lấy mẫu và đồng bộ theo thời gian thực, được hiển thị, lưu trữ dưới dạng số trên máy tính phục vụ quá trình phân tích và mô hình hóa.

2.3. Chế độ vận hành

Quá trình thực nghiệm được tiến hành theo hai chế độ tăng tốc nhằm khảo sát toàn diện đặc tính của HTPNL kiểu CR: tăng tốc ở số truyền N (động cơ ở chế độ không tải) và tăng tốc ở số truyền 1 (có tải trọng từ bệ thử con lăn).

- Chế độ không tải (Số truyền N): Phép thử này được sử dụng để khảo sát đặc tính của HTPNL kiểu CR trong điều kiện không có mô-men cản, cho phép động cơ tăng tốc lên dải tốc độ cao (trên 3000 vg/ph) nhằm quan sát trọn vẹn quy luật biến thiên của các tín hiệu chấp hành I_{mv} và I_{pcv}. Động cơ vận hành ổn định ở chế độ không tải cầm chừng (800 vg/ph), sau đó người vận hành đạp nhanh và đều bàn đạp ga đến 100% hành trình và giữ trong vài giây để đảm bảo các thông số đạt trạng thái ổn định mới.

- Chế độ có tải (Số truyền 1): Để đảm bảo an toàn vận hành ở số truyền thấp, tốc độ động cơ trong phép thử này được giới hạn ở mức 1500 vg/ph. Quy trình đạp ga tương tự như ở chế độ không tải nhằm xác định quy luật điều khiển van MV và PCV khi có tải, đồng thời khẳng định tính tương đồng về chiến lược điều khiển giữa hai chế độ.

Mặc dù theo thiết kế (Bảng 1) và một số kết quả thực nghiệm đã công bố [5, 7], HTPNL kiểu CR trên D4CB 2.5 TCI-A có khả năng xác lập áp suất rail tối đa lên đến 1600 bar. Tuy nhiên, trong phạm vi thực nghiệm này, do đặc thù vận hành ở số truyền thấp (số 1), lực cản từ

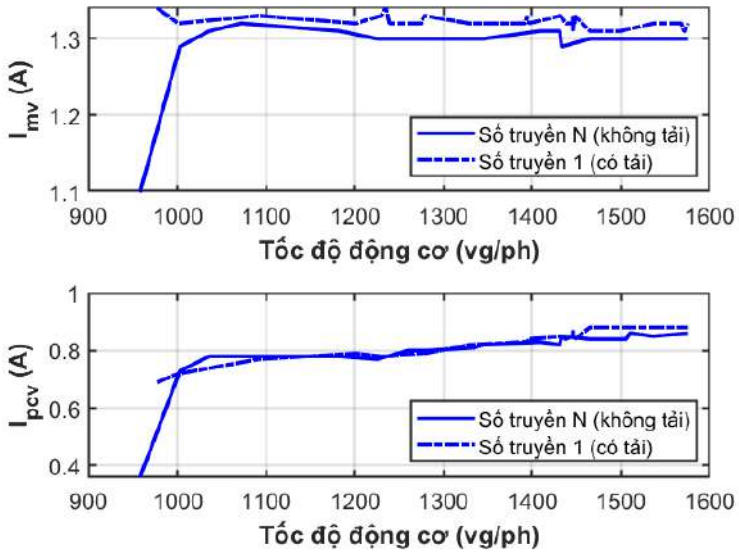
rulo của bệ thử con lăn chưa đủ lớn để đưa động cơ vào vùng tải trọng cao. Do đó, áp suất rail ghi nhận được trong các phép thử này chủ yếu phản ánh chiến lược điều khiển ở dải tải trọng cục bộ trung bình, phục vụ mục tiêu xác định quy luật phối hợp giữa các van điều khiển và xây dựng bộ dữ liệu cơ sở cho mô hình MBD.

3. Kết quả thử nghiệm và bàn luận

3.1. Đánh giá tính nhất quán của chiến lược điều khiển trong các điều kiện tải trọng

Kết quả khảo sát dòng điện điều khiển van MV và PCV ở hai chế độ vận hành: không tải và có tải được trình bày trên Hình 3.

Việc so sánh giữa hai chế độ làm việc cho thấy ở chế độ có tải, giá trị I_{mv} luôn duy trì ở mức cao hơn so với khi không tải trong toàn dải tốc độ khảo sát. Tuy nhiên, độ chênh lệch giữa hai đường đặc tính không lớn và tương đối ổn định, đặc biệt trong vùng tốc độ trung bình và cao. Hình dạng hai đường đặc tính gần như song song, chứng tỏ quy luật điều tiết lưu lượng nạp của ECU không thay đổi bản chất theo tải trọng động cơ.

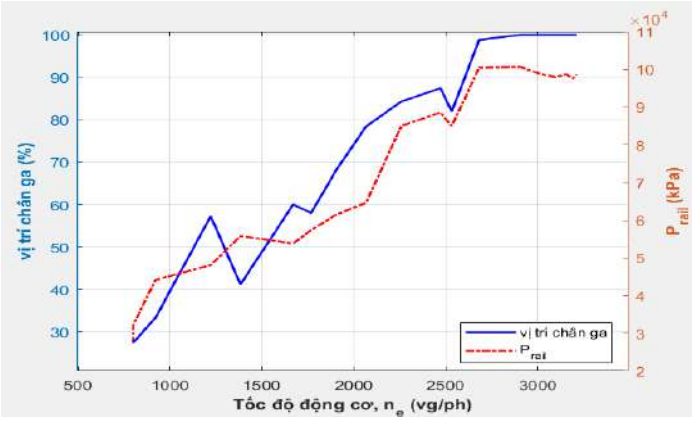


Hình 4. Sự thay đổi dòngđiều khiển van MV và PCV khi có tải và không tải.

Đối với dòng I_{pcv}, xu hướng biến thiên theo tốc độ cũng tương tự như I_{mv} nhưng có mức tăng ổn định và tuyến tính hơn. Ở cả hai chế độ làm việc, dòng Ipcv tăng dần theo tốc độ động cơ và duy trì sự biến thiên tương đối nhỏ. Sự khác biệt giữa chế độ có tải và không tải là rất ít, hai đường đặc tính gần như trùng khít trong phần lớn dải vận hành. Nhìn chung, sự tương đồng cao về đặc tính giữa hai chế độ khẳng định tính ổn định và nhất quán của chiến lược điều khiển HTPNL kiểu CR. Tải trọng động cơ chủ yếu làm thay đổi cường độ dòng điện điều khiển để bù trừ công suất mà không làm thay đổi thuật toán phối hợp các van. Đây là cơ sở thực nghiệm quan trọng cho việc xây dựng một mô hình toán học thống nhất phục vụ chẩn đoán MBD.

3.2. Phân tích mối quan hệ giữa điều kiện vận hành và biến trạng thái hệ thống

Hình 4 trình bày mối quan hệ giữa vị trí chân ga (điều kiện vận hành) và áp suất rail (biến trạng thái) theo tốc độ động cơ. Dữ liệu thực nghiệm cho thấy vị trí chân ga tăng nhanh từ khoảng 28% tại tốc độ không tải cầm chừng (800 vg/ph) lên đến 100% tại vùng tốc độ 3000 - 3200 vg/ph. Sự gia tăng tỷ lệ thuận này phản ánh yêu cầu về mô-men và tải trọng của động cơ trong quá trình tăng tốc.

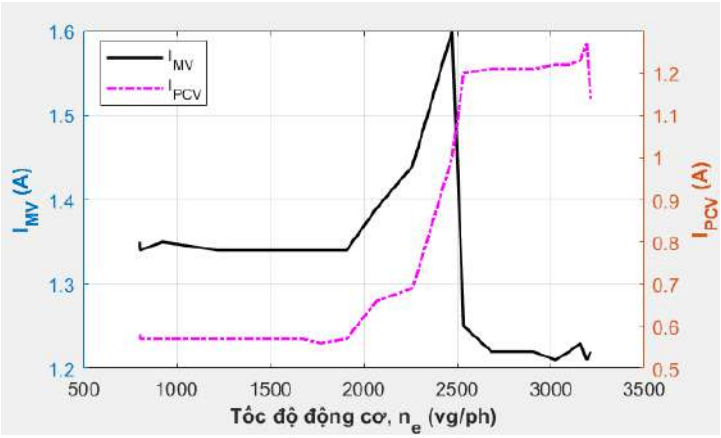


Hình 4. Sự thay đổi vị trí chân ga và áp suất rail theo tốc độ động cơ.

Áp suất rail phản hồi mạnh mẽ với sự thay đổi của tốc độ và chân ga, tăng mạnh từ mức 28.000 - 32.000 kPa tại tốc độ không tải cầm chừng lên khoảng 100.000 kPa tại vùng 2500 - 2900 vg/ph. Khi tốc độ động cơ vượt quá 3000 vg/ph, áp suất có xu hướng dao động nhỏ trong khoảng 97.000 - 100.000 kPa. Quy luật này cho thấy ECU đã thiết lập một ngưỡng áp suất mục tiêu tối ưu để đảm bảo chất lượng phun và hiệu quả quá trình cháy ở dải tốc độ cao. Kết quả đo chỉ đạt khoảng 1000 bar thay vì đến ngưỡng1600 bar là do tải trọng đặt lên rulo ở số truyền 1 (động cơ vận hành ở tải cục bộ) chưa đạt mức để ECU kích hoạt áp suất rail tối đa.

3.3. Đặc tính phối hợp của các tín hiệu chấp hành MV và PCV

Hình 5 mô tả chi tiết mối quan hệ giữa dòng điện điều khiển van MV, van PCV theo tốc độ động cơ, qua đó làm rõ chiến lược điều tiết áp suất của ECU. Ở vùng tốc độ thấp (800 - 1700 vg/ph), dòng điện I_{mv} duy trì ổn định ở mức 1,34 - 1,35 A, tương ứng với trạng thái van mở gần như hoàn toàn để chuẩn bị lưu lượng nạp. Trong giai đoạn này, nhu cầu về nhiên liệu phun và áp suất rail chưa lớn, nên dòng I_{pcv} cũng được giữ ở mức thấp (khoảng 0,6 A).

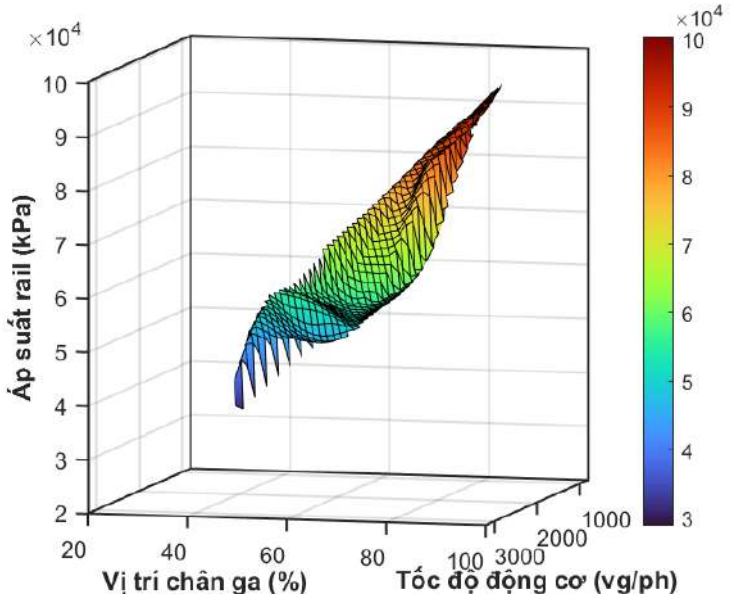


Hình 5. Sự thay đổi dòng điện điều khiển van MV và PCV theo tốc độ động cơ.

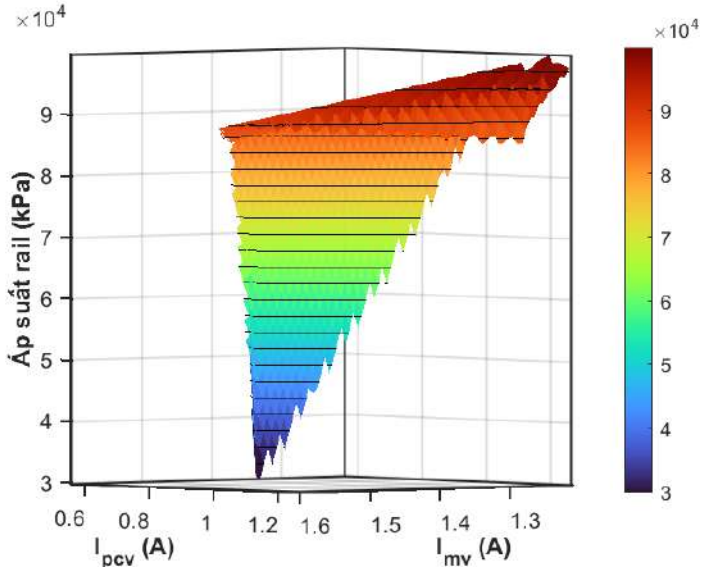
Khi tốc độ động cơ tăng lên vùng 1900 - 2500 vg/ph, lưu lượng cung cấp từ bơm cao áp tăng mạnh do số chu kỳ nén trong một đơn vị thời gian lớn hơn. Để kiểm soát sự gia tăng áp suất này, ECU tăng dần dòng I_{mv} lên mức cực đại khoảng 1,6 A tại 2472 vg/ph nhằm hạn chế bớt lượng nhiên liệu nạp vào bơm cao áp. Đồng thời, van PCV cũng chuyển sang chế độ đóng nhiều hơn (dòng tăng lên 1,0 - 1,27 A) để hạn chế lượng xả hồi, giữ cho áp suất rail đạt giá trị cao phục vụ chế độ công suất lớn. Sự phối hợp nhịp nhàng này khẳng định tính ổn định của vòng điều khiển kín: khi MV hạn chế đầu vào thì PCV cũng thu hẹp đầu ra xả, tạo nên đặc tính áp suất tăng gần như tuyến tính và ổn định theo yêu cầu tải trọng

3.4. Mô tả không gian trạng thái của hệ thống qua đồ thị 3D

Để đánh giá toàn diện tính phi tuyến của HTPNL kiểu CR, các đồ thị không gian 3D được xây dựng nhằm mô tả sự tương tác đồng thời giữa nhiều biến số. Hình 6 thể hiện bề mặt áp suất rail theo vị trí chân ga và tốc độ động cơ, cho thấy rõ vùng làm việc có độ dốc thấp khi chân ga dưới 40% (yêu cầu mô-men nhỏ) và vùng dốc đứng khi tải tăng mạnh, phản ánh sự nhạy bén của chiến lược điều khiển ECU.



Hình 7. Sự thay đổi áp suất rail theo vị trí chân ga và tốc độ động cơ.



Hình 8. Sự thay đổi áp suất rail theo dòng điện điều khiển van MV và PCV.

Hình 7 cung cấp cái nhìn trực quan về vai trò của các tín hiệu chấp hành trong việc xác lập biến trạng thái. Áp suất rail tăng rõ rệt và đạt trạng thái ổn định (trên 90.000 kPa) tại khu vực có dòng I_{mv} cao (1,4 - 1,5 A) và dòng I_{pcv} cao (0,8 - 1,2 A). Bề mặt nghiêng đều và rộng ở khu vực này minh chứng cho khả năng điều khiển tuyến tính hóa một đối tượng phi tuyến của ECU tại giai đoạn cuối của quá trình tăng tốc. Bộ dữ liệu không gian này không chỉ mô tả rõ vai trò của từng thành phần mà còn là cơ sở dữ liệu mẫu quan trọng để huấn luyện các thuật toán nhận dạng lỗi và xây dựng phần dư cho phương pháp MBD dùng cho HTPNL kiểu CR.

4. Kết luận, hướng phát triển

Nghiên cứu đã thực hiện khảo sát thực nghiệm và phân tích đặc tính động lực học của HTPNL kiểu CR trên động cơ Hyundai D4CB 2.5 TCI-A. Các kết quả chính đạt được bao gồm:

- Đã thiết lập được hệ thống đo lường đồng bộ, kết hợp giữa giao thức OBD-II và mạch đo dòng điện chuyên dụng, cho phép thu thập chính xác các thông số trạng thái (p_{cr}), điều kiện vận hành (n, vị trí chân ga) và tín hiệu chấp hành ($_{Imv}$, I_{pcv}).

- Kết quả thực nghiệm khẳng định tính nhất quán trong chiến lược điều khiển của ECU: quy luật phối hợp giữa van định lượng (MV) và van kiểm soát áp suất (PCV) có sự tương đồng rất cao giữa chế độ không tải và có tải. Điều này chứng minh rằng tải trọng động làm thay đổi cường độ dòng điện điều khiển các van nhưng không làm thay đổi bản chất thuật toán điều khiển phối hợp.

- Bộ dữ liệu không gian trạng thái 3D thu được là nguồn dữ liệu thực tế tin cậy, phản ánh đặc tính phi tuyến của hệ thống, phục vụ trực tiếp cho việc xây dựng và kiểm chứng các mô hình chẩn đoán theo phương pháp MBD cho HTPNL kiểu CR.

Từ những kết quả đạt được, hướng phát triển của nghiên cứu này bao gồm: (1) Sử dụng bộ dữ liệu thực nghiệm để nhận dạng tham số và xây dựng mô hình tham chiếu (Reference Model) cho HTPNL kiểu CR; (2) Phát triển các thuật toán quan sát và bộ lọc (filters) để trích xuất phần dư (residuals), từ đó hình thành các triệu chứng (symptoms) phục vụ chẩn đoán lỗi cơ cấu chấp hành; (3) Mở rộng thực nghiệm ở các số truyền cao hơn để khảo sát hệ thống ở dải áp suất rail định mức tối đa (1600 bar), nhằm hoàn thiện không gian trạng thái của mô hình chẩn đoán.

Tài liệu tham khảo:

1. Patrick Sauer M.Sc (2021), *Model-Based Fault Detection and Diagnosis For The Fuel System Of A Six-Cylinder Heavy Duty Diesel Engine*. TU Darmstadt, Germany.
2. Silvio Simani et al (2003). *Model-based Fault Diagnosis in Dynamic Systems Using Identification Techniques*. Springer-Verlag, London.
3. Steven X. Ding (2013). *Model-Based Fault Diagnosis Techniques: Design Schemes, Algorithms and Tools*. Second Edition, Universität Duisburg-Essen Duisburg, Germany.
4. Rolf Isermann (2016), *Advanced model-based diagnosis of internal combustion engines. Technische Universität Darmstadt*, Springer Fachmedien Wiesbaden, Doi 10.1007/978-3-658-12918-7_27.
5. Nguyễn Hoàng Vũ, Trần Trọng Tuấn, Phạm Trung Kiên (2023). *Xác định quy luật thay đổi áp suất phun của hệ thống phun nhiên liệu diesel kiểu CommonRail bằng thực nghiệm trên bệ thử con lăn*. Tạp chí Cơ khí Việt Nam. Số 302. P-ISSN 1859-3585, E-ISSN 2615-9619
6. Khổng Văn Nguyên, Trần Anh Trung, Nguyễn Hoàng Vũ (2017). *Nghiên cứu thực nghiệm đánh giá ảnh hưởng của tỷ lệ bio-diesel lên đặc tính làm việc của hệ thống phun nhiên liệu kiểu tích áp CommonRail.* Tạp chí Khoa học công nghệ Xây dựng, số 04.
7. Nguyễn Xuân Đạt (2021), *Nghiên cứu ảnh hưởng của phun chính nhiều giai đoạn đến chỉ tiêu kinh tế - kỹ thuật của động cơ diesel kiểu CommonRail khi sử dụng nhiên liệu diesel sinh học*. Luận án TSKT, Học viện KTQS, Hà Nội.
8. Rolf Isermann (2014). *Engine Modeling and Control: Modeling and Electronic Management of Internal Combustion Engines*. Springer, Germany.

Experimental Identification of Characteristic Parameters for Modeling and Model-Based Diagnostics of a CommonRail Fuel Injection System.

Abstract: *This paper presents experimental results for identifying characteristic parameters to support the modeling and Model-based Diagnostics (MBD) of a Common Rail (CR) diesel fuel injection system on a Hyundai D4CB 2.5 TCI-A engine using a chassis dynamometer. The experimental dataset includes: actuating signals (metering valve - MV and pressure control valve - PCV control currents), operating conditions (engine speed, accelerator pedal position), and the system state variable (rail pressure). These parameters were synchronously acquired via an OBD-II interface (ELM327) and an ACS712 current sensing circuit using an Arduino microcontroller. The results indicate that, in addition to its dependence on engine speed and accelerator pedal position, rail pressure is strongly influenced by the coordinated control of the MV and PCV valves. The obtained 2D and 3D plots clearly illustrate the roles of the MV, PCV, and the ECU control strategy across the entire operating range. The experimental dataset provides a vital foundation for developing dynamic rail pressure simulation models and applying MBD methods to the Common Rail fuel injection system.*

Keywords: Rail pressure control; metering valve (MV); pressure control valve (PCV); characteristic parameters; model-based diagnostics (MBD).

Nghiên cứu lựa chọn phương án cải tiến hệ thống làm mát động cơ diesel 3D12 khi cường hoá động cơ bằng tăng áp khí nạp

Nguyễn Văn Năm[1*], Nguyễn Văn Dương[2]
[1]Hệ QLHV sau đại học; [2]Học viện Kỹ thuật Quân sự
[*]Email: vegatboy@gmail.com; Tel: 0976765882

Tóm tắt

Cường hoá động cơ bằng tăng áp khí nạp làm tăng đáng kể công suất riêng và hiệu suất, tuy nhiên, nhiệt lượng truyền sang hệ thống làm mát tăng cao làm động cơ bị quá nhiệt. Trong nghiên cứu này, mô hình hệ thống làm mát của động cơ diesel 3D12 được xây dựng trên phần mềm GT-Suite nhằm đánh giá ảnh hưởng của tăng áp đến cân bằng năng lượng và đề xuất phương án cải tiến hệ thống làm mát khi cường hóa động cơ bằng tăng áp khí nạp. Kết quả tính toán cho thấy sau khi tăng áp khí nạp, nhiệt lượng truyền sang nước làm mát tăng từ 119,8 kW lên 163,6 kW, làm nhiệt độ nước làm mát ra khỏi động cơ tăng từ 84,38 °C lên 93,94 °C, tiệm cận giới hạn làm việc cho phép của động cơ. Nhiều phương án cải tiến được khảo sát; trong đó, giải pháp điều chỉnh tăng đồng thời lưu lượng nước ngọt và nước biển cho hiệu quả tốt nhất. Khi lưu lượng nước ngọt tăng lên 320 Lít/phút và lưu lượng nước biển tăng lên 210 Lít/phút, nhiệt độ nước ra khỏi động cơ được đưa về vùng làm việc ban đầu (từ 84 °C đến 86 °C). Kết quả nghiên cứu là cơ sơ quan trọng để lựa chọn phương án cải tiến hệ thống làm mát khi cường hóa động cơ bằng tăng áp khí nạp.

Từ khóa: *Động cơ diesel; 3D12; Tăng áp khí nạp; Hệ thống làm mát.*

1. Đặt vấn đề

Ngày nay, trước yêu cầu về tăng công suất riêng, hiệu suất và giảm phát thải, động cơ tàu thủy không ngừng được cải tiến. Trong đó, giải pháp cường hóa động cơ bằng tăng áp khí nạp được sử dụng khá phổ biến và mang lại hiệu quả cao [2-5]. Phương pháp này không chỉ được sử dụng cho động cơ mới mà còn rất phù hợp với các động cơ đời cũ đang được sử dụng trên tàu thủy. Tuy nhiên, cường hóa động cơ bằng tăng áp khí nạp làm cho áp suất và nhiệt độ cực đại trong xi lanh tăng đáng kể, kéo theo sự gia tăng tổng nhiệt lượng truyền qua thành xi lanh, piston và nắp máy sang môi chất làm mát [6]. Nếu không tăng cường làm mát, nhiệt độ nước làm mát sẽ tăng cao, tiềm ẩn nguy cơ sôi cục bộ, suy giảm độ bền vật liệu và chất lượng bôi trơn và có thể gây bó kẹt piston [7]. Vì vậy, khi cường hóa động cơ bằng tăng áp khí nạp, hệ thống làm mát cần được cải tiến để tăng cường làm mát động cơ.

Trên thế giới đã có nhiều nghiên cứu cải tiến hệ thống làm mát nhằm nâng cao khả năng tản nhiệt cho động cơ bao gồm: cải tiến kết cấu và nâng cao hiệu quả trao đổi nhiệt của các phần tử tản nhiệt, thông qua tăng diện tích bề mặt truyền nhiệt, tối ưu hóa hình dạng cánh tản nhiệt hoặc lựa chọn vật liệu có hệ số dẫn nhiệt cao [8,9]; phát triển các hệ thống làm mát điều khiển chủ động, sử dụng bơm nước điều khiển điện tử và van hằng nhiệt thông minh để điều chỉnh lưu lượng môi chất theo chế độ làm việc [10-12]; tiếp cận theo hướng tối ưu hóa tổng thể trên cơ sở phân tích cân bằng năng lượng của toàn bộ mạch làm mát [13]. Tuy nhiên, các công bố này đều nghiên cứu trên động cơ thiết kế mới. Số lượng nghiên cứu về cải tiến hệ thống làm mát khi cường hóa động cơ đời cũ vẫn chưa được đánh giá đầy đủ.

Động cơ 3D12 là dòng động cơ được sử dụng khá phổ biến trên tàu thủy, tuy nhiên, đây là dòng động cơ đời cũ, có công suất riêng và hiệu suất khá thấp. Nguyễn Đắc Lợi và cộng sự đã nghiên cứu khả năng cường hóa động cơ 3D12 bằng tăng áp khí nạp. Tỉ số tăng áp được giới hạn bằng 1,5 thông qua đánh giá trạng thái ứng suất-nhiệt của piston-xilanh, tuy nhiên, hệ thống làm mát chưa được xem xét tới [1].

Từ các phân tích trên, mục tiêu của nghiên cứu được xác định là: Đánh giá hệ thống làm mát nguyên bản, nghiên cứu đề xuất phương án cải tiến hệ thống làm mát phù hợp khi cường hóa động cơ bằng tăng áp khí nạp.

2. Phương pháp và đối tượng nghiên cứu

2.1. Phương pháp nghiên cứu

Để khảo sát các phương án cải tiến hệ thống làm mát cho động cơ diesel tàu thủy, sơ đồ tính toán được xác định như hình 1.

Sau khi xác định mục tiêu và đối tượng nghiên cứu, tiến hành xây dựng mô hình tính toán chu trình công tác động cơ để xác định nhiệt lượng truyền cho nước làm mát Q_w (bước 2,3); xây dựng mô hình hệ thống làm mát theo sơ đồ nghiên cứu (bước 4); tiến hành khảo sát, so sánh các phương án và lựa chọn phương án phù hợp (bước 5).

Các tính toán được thực hiện tại chế độ chế độ công suất định mức. Để đơn giản hóa quá trình tính toán, các giả thiết sau được đặt ra: Coi như động cơ làm việc ở trạng thái ổn định; các tính chất nhiệt – vật lý của môi chất được coi không đổi trong phạm vi khảo sát, và các phần tử của hệ thống làm mát làm việc theo đặc tính thiết kế.

Hình 1. Sơ đồ nghiên cứu

2.2. Đối tượng nghiên cứu

Đối tượng nghiên cứu được tác giả lựa chọn là động cơ diesel 3D12 được sử dụng làm máy chính và máy phát điện trên tàu thủy (4 kỳ, 12 xi lanh, 2 vế hình chữ V với góc nhị diện 60^0, không có tăng áp, hệ thống làm mát 2 vòng: nước ngọt - nước biển). Thông số kỹ thuật của động cơ 3D12 được thể hiện trong bảng 1.

Bảng 1. Các thông số kỹ thuật của động cơ 3D12

Thông số	Trị số	Thông số	Trị số
Số xi lanh	12	Bơm nước ngọt (vòng/phút) - Áp suất (kG/cm^2) - Lưu lượng (Lit/phút)	2225 1-1,8 ≥ 250
Đường kính xi lanh (mm)	150		
Hành trình piston (mm)	180		
Công suất định mức, (kW)	220	Bơm nước biển (vòng/phút) - Áp suất (kG/cm^2) - Lưu lượng (Lit/phút)	2025 1-1,5 ≥ 130
Tốc độ vòng quay trục khuỷu ứng với công suất định mức (vòng/phút)	1500		
Nhiệt độ nước ngọt (^0C) + Vào động cơ + Ra động cơ + Chênh lệch nhiệt độ vào – ra	 65-80 80-95 5-10	Két mát (kg) + Số ống (cái) + Chiều cao (mm) + Chiều rộng (mm)	26,3 282 394 330

3. Xây dựng mô hình

3.1. Mô hình chu trình công tác

Mô hình tính toán chu trình công tác của động cơ nguyên bản (động cơ chưa tăng áp) và động cơ sau khi được cường hóa bằng tăng áp khí nạp được thiết lập trong phần mềm GT-Suite. Phần mềm GT-suite được phát triển bởi Gamma Technologies, Hoa Kỳ, là một môi trường mô phỏng tích hợp, được sử dụng rộng rãi trong nghiên cứu và phát triển động cơ đốt trong [14,15].

Mục đích của xây dựng mô hình là để xác định nhiệt lượng truyền vào hệ thống làm mát. Kết quả tính toán nhiệt lượng truyền cho hệ thống làm mát ở cả 2 trường hợp không tăng áp và có tăng áp được thể hiện ở bảng 2. Tỷ số tăng áp khí nạp được lựa chọn là $\Pi_k = 1,5$ [1].

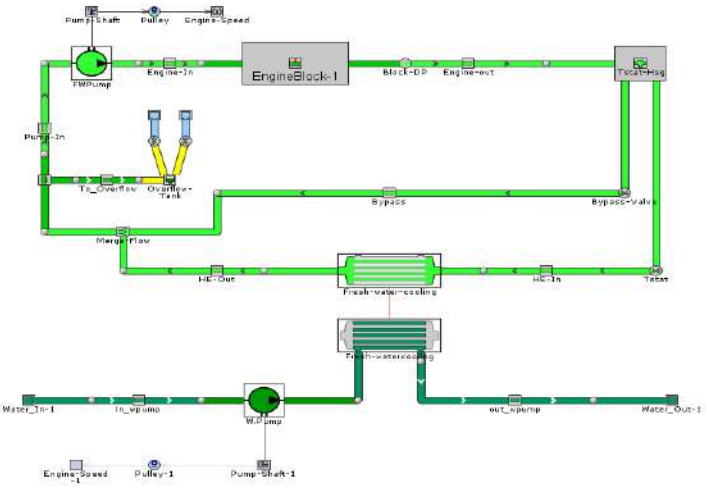
Bảng 2. Nhiệt lượng truyền cho HTLM ở chế độ vòng quay định mức

TT	Tên thông số	Ký hiệu	Đơn vị	$\Pi_k = 1,0$	$\Pi_k = 1,5$
1	Tốc độ động cơ	n	v/ph	1500	1500
2	Công suất có ích động cơ	N_e	kW	215,3	301
3	Hiệu suất có ích	η_e	%	33,0	34,1
4	Suất tiêu hao nhiên liệu	g_e	g/kW.h	252,1	246
5	Nhiệt độ cháy cực đại	T_{max}	K	1809	1865
6	Nhiệt lượng tỏa ra trong 1 chu trình công tác do đốt cháy nhiên liệu	Q_c	J	4343	5891
7	Nhiệt lượng truyền cho HTLM	Q_w	kW	119,8	163,6

3.2. Mô hình tính toán hệ thống làm mát

Mô hình hệ thống làm mát được xây dựng trong GT-Suite, gồm vòng nước ngọt tuần hoàn kín làm mát block động cơ và vòng nước biển hở làm mát sinh hàn nước ngọt. Nước ngọt được tuần hoàn bởi bơm ly tâm dẫn động từ trục khuỷu và trao đổi nhiệt với block động cơ thông qua mô hình truyền nhiệt đối lưu. Dòng nước sau động cơ được điều hướng bởi van hằng nhiệt giữa nhánh bypass (nhánh phụ cho phép nước làm mát tuần hoàn trực tiếp từ cửa ra động cơ quay lại cửa vào bơm mà không đi qua két làm mát) và nhánh qua két làm mát nhằm duy trì nhiệt độ vận hành tối ưu. Sinh hàn nước ngọt – nước biển được mô hình hóa dạng bộ trao đổi nhiệt ống chùm, trong đó hai mạch nước trao đổi nhiệt gián tiếp. Mô hình tích hợp đầy đủ đường ống, tổn thất thủy lực, két giãn nở và các cảm biến đo. Nhờ đó, mô hình phản ánh sát thực tế làm việc của hệ thống và cho phép đánh giá hiệu quả làm mát trong các điều kiện vận hành khác nhau.

Các thông số đầu vào cho mô hình hệ thống làm mát được lấy từ số liệu ở bảng 1 và bảng 2. Điều kiện tính toán của động cơ được thực hiện ở chế độ công suất định mức.



Hình 4. Sơ đồ mô phỏng hệ thống làm mát trên động cơ 3D12

3.3. Hiệu chỉnh mô hình

Kết quả tính toán chu trình công tác của động cơ nguyên bản trong phần mềm GT-Suite được so sánh với số liệu của nhà sản xuất để hiệu chỉnh và đánh giá độ tin cậy của mô hình, được thể hiện trong bảng 3.

Bảng 3. So sánh kết quả tính toán động cơ nguyên bản và số liệu do nhà sản xuất cung cấp ở chế độ công suất định mức

Thông số	Số liệu của nhà sản xuất	Kết quả mô phỏng	Sai số tương đối (%)
Công suất lớn nhất (kW)	220	215.3	2,27
Suất tiêu hao nhiên liệu (g/Kw.h)	254	252	0,8

Từ dữ liệu bảng 3 cho thấy kết quả từ mô hình tính toán chu trình công tác của động cơ nguyên bản trong phần mềm GT-Suite phù hợp với số liệu do nhà sản xuất cung cấp, các giá trị có sai lệch không quá 5%. Vì vậy, có thể sử dụng mô hình tính toán trong GT-Suite để tiến hành tính toán chu trình công tác của động cơ.

4. Kết quả và thảo luận

4.1. Khảo sát đặc tính làm việc của hệ thống làm mát trên động cơ nguyên bản

Mô hình mô phỏng hệ thống làm mát được xây dựng dựa trên hệ thống làm mát nguyên bản của động cơ 3D12 nhưng do các thông số kỹ thuật của van hằng nhiệt không đầy đủ nên phải khảo sát đặc tính làm việc của bơm và hệ thống nhằm mục đích hiệu chỉnh lưu lượng của bơm về giá trị tương đương với lưu lượng do nhà sản xuất cung cấp như ở bảng 1. Kết quả các trường hợp mô phỏng được thể hiện trên bảng 4.

Việc khảo sát hệ thống làm mát động cơ diesel tàu thủy thông qua nhiệt độ nước làm mát, coi đây là thông số đặc trưng phản ánh trực tiếp trạng thái cân bằng năng lượng và mức độ an toàn nhiệt của động cơ. Trong chuỗi mô phỏng từ trường hợp **TH1** đến **TH6**, mô hình hệ thống làm mát gồm một bơm tuần hoàn, nhánh phụ có nhánh *Bypass* liên kết nghịch biến với *Thermostat*, và nhánh chính *Thermostat* qua két nước làm mát. Mỗi trường hợp mô phỏng tương ứng với một mức độ mở của van, tỷ lệ lưu lượng qua thermostat tăng từ 50,42% (TH1) lên 99,53% (TH6), trong khi lưu lượng qua nhánh bypass giảm tương ứng từ 49,58% xuống 0,47%. Kết quả khảo sát cho thấy điểm làm việc của bơm thể hiện qua lưu lượng trung bình và cột áp thay đổi phi tuyến khi trạng thái mở của thermostat/bypass biến đổi.

Bảng 4. Thông số khi điều chỉnh lưu lượng qua Bypass và Tstat trong mô phỏng hệ thống làm mát động cơ 3D12

Thông số mô phỏng		Đơn vị	TH1	TH2	TH3	TH4	TH5	TH6
% Lưu lượng qua Bypass		%	49,58	41,79	31,35	20,99	10,63	0,47
% Lưu lượng qua Tstat		%	50,42	58,21	68,65	79,01	89,37	99,53
Tổng lưu lượng qua bơm nước ngọt		Lít/phút	259,94	260,70	259,88	234,76	208,75	260,65
Tổng lưu lượng qua bơm nước biển		Lít/phút	142,65	142,65	142,65	142,65	142,65	142,65
Động cơ	$T_{đ\ vào}$	°C	85,02	80,93	76,80	75,05	73,86	68,84
	$T_{đ\ ra}$	°C	92,57	88,47	84,38	83,44	83,30	76,44
ΔT	$T_{đ\ ra} - T_{đ\ vào}$	°C	7,55	7,54	7,58	8,39	9,44	7,6
Két mát	$T_{k\ vào}$	°C	92,51	88,42	84,33	83,39	83,24	76,40
	$T_{k\ ra}$	°C	77,73	75,61	73,41	72,88	72,79	68,83
ΔT	$T_{k\ ra} – T_{k\ vào}$	°C	14,78	12,81	10,92	10,51	10,45	7,57

Cụ thể, tại các TH1,2,3: tổng lưu lượng vào bơm khoảng 260 Lít/phút. Khi thermostat bắt đầu mở nhiều hơn ở TH4 và TH5, tổng lưu lượng giảm xuống 234,76 Lít/phút (TH4) và 208,75 Lít/phút (TH5). Ở TH6 - thermostat được mở hoàn toàn, tổng lưu lượng hồi phục gần mức ban đầu (≈260 Lít/phút). Các thay đổi này đi kèm với biến thiên cột áp (H) bơm và sự thay đổi công suất: cột áp tăng khi lưu lượng giảm (H ≈ 9,55 m ở TH1,2,3; H ≈ 9,77 m và 9.94 m ở TH4,5).

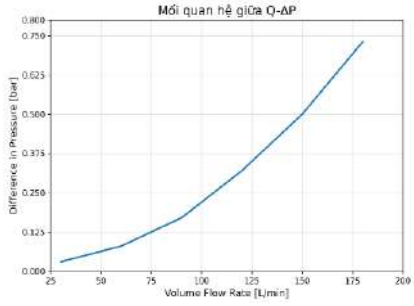
Giải thích cơ chế dựa trên đặc tính van hằng nhiệt và lý thuyết bơm – hệ thống: nhánh bypass được định nghĩa là nghịch biến với nhánh thermostat, do đó khi thermostat mở nhiều, diện tích hiệu dụng của van hằng nhiệt qua nhánh bypass giảm. Khi đó một phần lưu lượng đi qua két làm mát nhiều hơn, nơi có hệ số tổn thất cục bộ và ma sát lớn hơn. Về mặt thuỷ lực, tổng trở của hệ tăng lên làm tăng tổn thất áp suất trong hệ thống làm mát dẫn đến làm dịch chuyển điểm làm việc giữa đường đặc tính bơm và đường đặc tính hệ thống. Do đó, bơm không còn duy trì được lưu lượng ổn định ở mức khảo sát ban đầu như ở TH1,2,3. Khi lưu lượng giảm, bơm dịch chuyển sang điểm làm việc ở vùng cột áp cao hơn – lưu lượng thấp hơn phù hợp với xu hướng quan sát được ở TH4 và TH5. Việc lưu lượng ở TH6 quay trở lại gần giá trị ban đầu mặc dù thermostat mở hoàn toàn có hai nguyên nhân. Thứ nhất, khi thermostat đạt trạng thái mở lớn nhất, dòng qua két nước chuyển sang một cấu hình có ít đoạn co hẹp, tổn thất cục bộ tương đối nhỏ hơn so với trạng thái "mở một phần", các tổn thất cục bộ có thể không đạt cực đại ở trạng thái trung gian mà đạt giá trị lớn nhất khi van và các khe không đồng bộ tạo ra luồng rối nhiều. Thứ hai, trong mô hình GT-Suite, mối quan hệ giữa độ mở diện tích hiệu dụng van và biểu đồ ΔP–Q của thermostat/ két làm mát có thể gây ra tính phi tuyến, dẫn đến việc hệ số tổn thất tổng giảm khi van mở hoàn toàn. Cả hai yếu tố này giải thích tại sao TH4,5 biểu hiện trở lực hệ lớn nhất và lưu lượng thấp nhất, còn TH6 cho điểm làm việc chuyển về vùng lưu lượng cao hơn.

Như vậy, vùng làm việc hợp lý của bơm tương ứng với trạng thái van hằng nhiệt cho phép khoảng 50% - 70% tổng lưu lượng đi qua nhánh thermostat. Đối với động cơ 3D12 không tăng áp, nhằm mô phỏng hệ thống làm mát sát với điều kiện thực tế, dải nhiệt độ nước làm mát được lựa chọn là 70 °C - 80 °C ở đầu vào động cơ ($T_{đ\ vào}$) và 80 °C - 95 °C ở đầu ra ($T_{đ\ ra}$). Ứng với chế độ vòng quay định mức 1500 vòng/phút, TH3 đáp ứng tốt các yêu cầu trên và được chọn làm trường hợp cơ sở cho các tính toán và phân tích tiếp theo.

4.2. Khảo sát nhiệt độ động cơ sau khi cường hoá bằng tăng áp khí nạp

Trong mô phỏng hệ thống làm việc với điều kiện tăng áp, nhiệt truyền cho nước làm mát tăng từ 119,8 kW lên 163,6 kW, nhiệt độ nước làm mát tăng khiến độ nhớt và mật độ giảm, làm giảm tổng trở hệ thống. Bơm, với cùng tốc độ quay, dịch sang vùng lưu lượng cao hơn trên đường đặc tính, nên lưu lượng tổng tăng từ 260 Lít/phút lên 286 Lít/phút.

Để duy trì lưu lượng bơm ổn định tại 260 Lít/phút trong các trường hợp mô phỏng, đặc tính tổn thất áp của mạch làm mát được hiệu chỉnh. Bằng cách điều chỉnh các điểm dữ liệu trên đường đặc tính, hệ thống có thể cân bằng áp suất sao cho điểm làm việc của bơm trùng với lưu lượng thiết kế khoảng 260 Lít/phút, giúp kết quả mô phỏng phản ánh đúng trạng thái vận hành.



Hình 5. Đồ thị biểu diễn mối quan hệ giữa lưu lượng và áp suất

Bảng 5. Thông số nhiệt độ nước làm mát qua động cơ và két mát khi cường hoá động cơ bằng tăng áp

Thông số	Nhiệt độ	$T_{đ\ vào}$	$T_{đ\ ra}$
Động cơ	°C	83,55	93,94
Két mát	°C	93,89	82,18

Như vậy khi động cơ được cường hoá bằng tăng áp khí nạp, nhiệt độ động cơ tăng cao, bơm làm việc ở dải lưu lượng trên không đáp ứng được nhu cầu làm mát cho động cơ (mặc dù lưu lượng bơm tăng lên đáng kể). Nhằm đáp ứng yêu cầu làm mát cho động cơ, các phương án sau được lựa chọn để phù hợp với mô hình thiết kế cho động cơ vận hành thực tế:

- Phương án 1: Tăng lưu lượng bơm nước ngọt.
- Phương án 2: Tăng lưu lượng bơm nước biển.
- Phương án 3: Tăng lưu lượng bơm nước ngọt, bơm nước biển.
- Phương án 4: Tăng diện tích trao đổi nhiệt trong két mát.
- Phương án 5: Tăng đồng thời lưu lượng bơm nước ngọt và diện tích trao đổi nhiệt trong két mát.

4.3. Nghiên cứu cải tiến hệ thống làm mát

+ Phương án 1: Tăng lưu lượng bơm nước ngọt

Bảng 6. Thông số nhiệt độ nước làm mát qua động cơ và két mát khi thay đổi lưu lượng bơm nước ngọt

Thông số mô phỏng		Đơn vị	TH1	TH2	TH3	TH4	TH5
Phần lưu lượng tăng thêm		%	10	20	30	40	50
Tổng lưu lượng qua bơm		Lít/phút	286	312	338	364	390
Động cơ	$T_{đ\ vào}$	ºC	82,76	82,2	81,47	81,11	83,65
	$T_{đ\ ra}$	ºC	92,12	90,86	89,3	88,52	90,32
Két mát	$T_{k\ vào}$	ºC	92,08	90,82	89,26	88,48	90,28
	$T_{k\ ra}$	ºC	81,2	80,52	79,67	79,25	80,27

Ở bảng 6, từ TH1 đến TH4 các thông số cài đặt của van bypass và thermostat được giữ nguyên (Như ở bảng 4 - TH3). Kết quả cho thấy khi lưu lượng tăng từ 10% đến 40%, nhiệt độ nước ra khỏi động cơ giảm dần từ 92,12 °C xuống còn 88,52 °C, thể hiện xu hướng cải thiện hiệu quả làm mát rõ rệt khi lưu lượng được gia tăng hợp lý.

Đối với TH5, để đạt được tổng lưu lượng 390 Lít/phút, hệ số lưu lượng của van bypass và van thermostat đã được hiệu chỉnh theo tỉ lệ 35,2 % và 65,8 % nhằm thay đổi phân bố dòng chảy trong hệ thống. Mặc dù lưu lượng tăng lên mức cao nhất, nhiệt độ ra khỏi động cơ chỉ giảm xuống 90,32 °C cho thấy hiệu quả làm mát không còn tỷ lệ thuận với mức tăng lưu lượng. Về bản chất, khả năng giải nhiệt của hệ thống phụ thuộc đồng thời vào diện tích trao đổi nhiệt, chênh lệch nhiệt độ giữa các môi chất và đặc tính phân bố dòng chảy. Khi các yếu tố này đã gần đạt trạng thái khai thác tối đa, việc tiếp tục gia tăng lưu lượng chỉ làm thay đổi điều kiện thủy lực mà không làm tăng đáng kể công suất thải nhiệt. Điều này chứng tỏ rằng việc tăng lưu lượng nước ngọt làm mát động cơ đơn thuần đến một ngưỡng nhất định sẽ đạt trạng thái bão hòa về hiệu quả, và việc hiệu chỉnh van phân dòng có vai trò quan trọng trong tối ưu hóa phân bố lưu lượng và hiệu suất giải nhiệt tổng thể của hệ thống.

+ Phương án 2: Tăng lưu lượng bơm nước biển

Bảng 7. Thông số nhiệt độ nước làm mát qua động cơ và két mát khi thay đổi lưu lượng bơm nước biển

Thông số mô phỏng		Đơn vị	TH1	TH2	TH3	TH4	TH5
Phần lưu lượng tăng thêm		%	10	20	30	40	50
Tổng lưu lượng qua bơm nước biển		Lít/phút	156	170	184	198	213
Động cơ	$T_{đ\ vào}$	°C	81,88	80,6	79,6	78,80	78,17
	$T_{đ\ ra}$	°C	92,19	90,87	89,86	89,08	88,47
Két mát	$T_{k\ vào}$	°C	92,14	90,82	89,81	89,03	88,42
	$T_{k\ ra}$	°C	80,17	78,61	77,37	76,38	75,58

Kết quả ở Bảng 7 cho thấy khi lưu lượng bơm nước biển tăng từ 156 lít/phút lên 213 lít/phút, nhiệt độ nước ra khỏi động cơ và nhiệt độ nước ra khỏi két làm mát đều giảm rõ rệt. Điều này khẳng định việc gia tăng lưu lượng nước biển đã nâng cao khả năng tiếp nhận và mang nhiệt ra môi trường. Tuy nhiên, mức giảm nhiệt độ không tỷ lệ thuận với mức tăng lưu lượng và có xu hướng tiệm cận khi tiếp tục gia tăng lưu lượng nước biển. Hiện tượng này cho thấy khi phía nước biển đã đủ khả năng tải nhiệt, quá trình trao đổi nhiệt của sinh hàn không còn bị khống chế chủ yếu bởi lưu lượng nước biển mà chuyển sang bị giới hạn bởi diện tích truyền nhiệt và điều kiện trao đổi nhiệt tổng thể của thiết bị.

+ Phương án 3: Tăng lưu lượng bơm nước ngọt, bơm nước biển

Bảng 8. Thông số nhiệt độ nước làm mát qua động cơ và két mát khi thay đổi lưu lượng bơm nước ngọt và lưu lượng bơm nước biển

Thông số mô phỏng		Đơn vị	TH1	TH2	TH3	TH4	TH5
Tổng lưu lượng qua bơm nước biển		Lít/phút	156	172	210	226	242
Tổng lưu lượng qua bơm nước ngọt		Lít/phút	284	305	320	338	342
Động cơ	$T_{đ\ vào}$	°C	81,46	79,46	77,30	77,02	77,01
	$T_{đ\ ra}$	°C	90,79	88,28	85,46	85,01	84,91
Két mát	$T_{k\ vào}$	°C	90,74	88,24	85,43	84,97	84,88
	$T_{k\ ra}$	°C	79,65	77,20	74,18	73,41	72,92

Dựa trên kết quả ở bảng 8, việc đồng thời tăng lưu lượng bơm nước ngọt và bơm nước biển đã cải thiện đáng kể hiệu quả giải nhiệt của hệ thống làm mát động cơ. Khi lưu lượng nước biển tăng từ 156 Lít/phút lên 242 Lít/phút và lưu lượng nước ngọt tăng từ 284 Lít/phút lên 342 Lít/phút, nhiệt độ nước làm mát ra khỏi động cơ giảm từ 90,79 °C xuống 84,91 °C. Xu hướng giảm đều của nhiệt độ tại cả động cơ và két làm mát qua các trường hợp cho thấy hiệu quả trao đổi nhiệt tỉ lệ thuận với lưu lượng dòng chảy cho đến ngưỡng khảo sát. Đặc biệt, từ TH3 trở đi, sự suy giảm nhiệt độ trở nên rõ rệt hơn, chứng tỏ rằng việc tăng đồng thời hai lưu lượng ở hai vòng làm mát là giải pháp có tác dụng cộng hưởng, nâng cao đáng kể khả năng mang nhiệt của hệ thống. Nhờ đó, hệ thống có thể đưa nhiệt độ động cơ về gần giá trị mong muốn khi động cơ làm việc trong điều kiện cường hoá bằng tăng áp.

+ Phương án 4: Tăng diện tích trao đổi nhiệt trong két mát

Kết quả ở bảng 9 cho thấy việc tăng kích thước két làm mát từ 10% đến 50% gần như không làm thay đổi đáng kể nhiệt độ nước làm mát trong hệ thống. Nhiệt độ nước ra khỏi động cơ dao động quanh 93,83°C - 93,94 °C, trong khi nhiệt độ ra khỏi két làm mát duy trì ở mức 82,06 °C - 82,19 °C.

Bảng 9. Thông số nhiệt độ nước làm mát qua động cơ và két mát khi thay đổi kích thước két mát

Thông số mô phỏng		Đơn vị	TH1	TH2	TH3	TH4	TH5
Phần kích thước tăng thêm		%	10	20	30	40	50
Động cơ	$T_{đ\ vào}$	°C	83,45	83,45	83,59	83,59	83,50
	$T_{đ\ ra}$	°C	93,84	93,83	93,94	93,94	93,84
Két mát	$T_{k\ vào}$	°C	93,78	93,78	93,89	93,89	93,79
	$T_{k\ ra}$	°C	82,07	82,06	82,19	82,19	82,07

Sự biến thiên rất nhỏ này cho thấy hiệu quả trao đổi nhiệt cải thiện không đáng kể, ngay cả khi tăng kích thước két lên tối đa 50%. Điều này chứng tỏ rằng trong cấu hình hệ thống hiện tại, giới hạn trao đổi nhiệt không nằm ở diện tích bề mặt két làm mát mà chủ yếu bị chi phối bởi lưu lượng dòng chảy và điều kiện tiếp nhận nhiệt của nước biển. Vì vậy, tăng kích thước két làm mát không phải là giải pháp hiệu quả để hạ nhiệt cho động cơ khi cường hoá bằng tăng áp.

+ Phương án 5: Tăng đồng thời lưu lượng bơm nước ngọt và diện tích trao đổi nhiệt trong két mát

Bảng 10. Thông số nhiệt độ nước làm mát qua động cơ và két mát khi thay đổi lưu lượng bơm nước ngọt và kích thước két mát

Thông số mô phỏng		Đơn vị	TH1	TH2	TH3	TH4	TH5
Phần kích thước két mát tăng thêm		%	10	20	30	40	50
Tổng lưu lượng qua bơm nước ngọt		Lít/phút	287	317	340	354	368
Động cơ	$T_{đ\ vào}$	°C	81,41	80,78	80,44	80,15	80,02
	$T_{đ\ ra}$	°C	90,68	89,18	88,29	87,7	87,34
Két mát	$T_{k\ vào}$	°C	90,64	89,14	88,25	87,66	87,01
	$T_{k\ ra}$	°C	79,53	78,76	78,35	77,98	77,56

Dựa trên số liệu của bảng 10, khi đồng thời tăng kích thước két làm mát (10% - 50%) và tăng lưu lượng bơm nước ngọt từ 287 Lít/phút lên 368 Lít/phút, hiệu quả trao đổi nhiệt của hệ thống cải thiện hơn so với trường hợp chỉ tăng kích thước két làm mát đơn lẻ. Nhiệt độ nước ra khỏi động cơ giảm từ 90,68 °C xuống còn 87,34 °C, cho thấy lưu lượng nước ngọt tăng đóng vai trò quyết định trong việc cải thiện khả năng mang nhiệt của dòng làm mát. Tương tự, nhiệt độ nước ra khỏi két làm mát cũng giảm từ 79,53 °C xuống 77,56 °C, phản ánh khả năng thải nhiệt được tăng cường nhờ sự kết hợp giữa lưu lượng lớn hơn và diện tích trao đổi nhiệt mở rộng.

Xu hướng giảm nhiệt độ ổn định theo từng trường hợp cho thấy rằng giải pháp kết hợp tăng lưu lượng bơm nước ngọt với tăng kích thước két làm mát mang lại hiệu quả cao hơn đáng kể so với việc chỉ tăng riêng kích thước két. Vì vậy, đây là một phương án cải tiến phù hợp, có khả năng đưa nhiệt độ làm mát động cơ về gần hơn vùng giá trị an toàn khi động cơ làm việc trong điều kiện cường hoá bằng tăng áp.

5. Kết luận

Nghiên cứu cho thấy khi động cơ 3D12 được tăng áp, điều kiện nhiệt – thủy lực của hệ thống làm mát thay đổi theo hướng bất lợi, làm phá vỡ trạng thái cân bằng nhiệt ban đầu. Mô phỏng trên GT-Suite đã xác định rõ sự gia tăng tải nhiệt và giới hạn làm việc của hệ thống nguyên bản thông qua biến đổi nhiệt độ và lưu lượng trong mạch. Phân tích các phương án cải tiến chỉ ra rằng việc mở rộng thiết bị trao đổi nhiệt không phải là yếu tố quyết định; thay vào đó, tối ưu đặc tính dòng chảy, đặc biệt là khả năng cung cấp lưu lượng của các bơm giữ vai trò then chốt.

Phương án tăng đồng thời lưu lượng ở cả vòng nước ngọt và nước biển cho hiệu quả rõ rệt mà không thay đổi lớn về kết cấu. Khi lưu lượng nước biển tăng từ 142 Lít/phút lên 210 Lít/phút và nước ngọt tăng từ 260 Lít/phút lên 320 Lít/phút, nhiệt độ nước ra khỏi động cơ giảm 8,48 °C, cho thấy trạng thái nhiệt của động cơ được khôi phục ổn định về giá trị khoảng 84 °C. Do đó, tăng lưu lượng cho cả hai vòng nước được xác định là giải pháp hiệu quả và có tính khả thi.

Trên cơ sở đó, hướng nghiên cứu tiếp theo được đề xuất là khảo sát phương án thay thế các bơm hiện hữu bằng loại có lưu lượng danh định lớn hơn và đặc tính làm việc phù hợp hơn với động cơ sau tăng áp, đồng thời đánh giá toàn diện ảnh hưởng đến tổn thất thủy lực, công suất phụ trợ và hiệu quả trao đổi nhiệt tổng thể. Điều này sẽ hoàn thiện cơ sở khoa học cho cải tiến hệ thống làm mát trong điều kiện khai thác thực tế.

Tài liệu tham khảo

1. Nguyễn Đắc Lợi, Nguyễn Văn Dương (2025). Nghiên cứu ảnh hưởng tỷ số tăng áp khí nạp đến thông số động cơ và trạng thái nhiệt của nhóm piston - xi lanh động cơ diesel tàu thủy. *Tuyển tập Hội nghị khoa học các nhà nghiên cứu trẻ lần thứ XX* (pp. 1622-1633). Hà Nội: Nhà xuất bản Quân đội Nhân dân.
2. Nguyễn Duy Vinh, Khổng Vũ Quảng, Phạm Minh Tuấn, Nguyễn Tiến Hãn (2011). Ứng dụng AVL-Boost để mô phỏng quá trình cháy của động cơ diesel khi tăng áp bằng tuabin khí. *Tạp chí Khoa học và Công nghệ Hàng hải*, (27), 32–39.
3. Nguyễn Lan Hương, Nguyễn Thị Xuân Hương (2016). *Nghiên cứu đánh giá khả năng tăng áp cho động cơ D243 bằng phần mềm AVL-BOOST*. Báo cáo đề tài nghiên cứu khoa học cấp trường. Hải Phòng: Đại học Hàng Hải Việt Nam.
4. Oun, M. S., Farhat, S. A., & Rabeei, M. A. I. (2017). The effect of turbocharger pressure and intercooler temperature on engine performance. *Journal of Engineering Research, University of Tripoli*, 23, 103–116.
5. Chen, S., Ma, C., Zhang, H., Xu, C., Lei, F., Feng, T., Cao, G., Yang, H., & Wei, C. (2022). Engine performance improvements through turbocharger matching and turbine design. *Energy Science & Engineering*, 10, 3384–3396. https://doi.org/10.1002/ese3.1225
6. Heywood, J. B. (1988). *Internal Combustion Engine Fundamentals*. New York: McGraw-Hill Education.
7. Kalelı, A., Kaltakkiran, G., Dumlu, A., & Ayten, K. K. (2017). Design and control of intelligent cooling system for IC engine. *International Conference on Engineering and Technology (ICET)*, Antalya, Turkey.
8. Fathallah, A. Z. M., Busse, W., & Clausthaldi, F. R. (2017). Fluid flow analysis of jacket cooling system for marine diesel engine 93 kW. *International Journal of Marine Engineering Innovation and Research*, 1(2), 39–45. https://doi.org/10.12962/j25481479.v1i2.2028.
9. Tanougast, A., & Hriczó, K. (2025). Advancements in heat exchangers for improved engine cooling: A comprehensive literature review. *Periodica Polytechnica Transportation Engineering*, 53(4), 431–439.
10. Śliwiński, K., & Szramowiat, M. (2018). *Development of cooling systems for internal combustion engines in the light of the requirements of modern drive systems*. IOP Conf. Ser.: Mater. Sci. Eng., 421, 042078.

11. Zhang, B., Zhang, P., & Zeng, F. (2021). Multiobjective optimization of the cooling system of a marine diesel engine. *Energy Science & Engineering*, 1887–1907. https://doi.org/10.1002/ese3.960.
12. Skarbierz, W., Graban, K., Wnuk, R., & Łebkowski, A. (2026). Energy-efficient cooling system control in ship engine rooms using an intelligent integrated automation, control, and monitoring system (IACMS). *Energies*, 19(3), 734. https://doi.org/10.3390/en19030734.
13. Stanivuk, T., Lalić, B., Mikuličić, J. Ž., & Šundov, M. (2021). Simulation modelling of marine diesel engine cooling system. *Transactions on Maritime Science*, 10(1), 112–125. https://doi.org/10.7225/toms.v10.n01.008.
14. Wei, S., Zhang, Z., Li, X., Wu, C., & Yang, F. (2022). Simulation analysis of fuel economy of the GDI engine with a Miller cycle and EGR based on GT-Power. *Processes*, 10, 319. https://doi.org/10.3390/pr10020319
15. Bai, Y., Wang, H., Yang, C., & Wang, Y. (2023). Fault simulation and analysis of marine low-speed diesel engine based on GT-Power. In *Marine Equipment & Technology and Sustainable Development (METSD 2023)* (Lecture Notes in Civil Engineering, Vol. 375, pp. 978–987). Springer.

Investigation and selection of cooling system improvement solutions for a turbocharged 3D12 diesel engine

Abstract**:** *Turbocharging a diesel engine can significantly increase brake power; however, it also increases the heat rejected to the cooling system, potentially causing the coolant temperature to exceed allowable operating limits. In this study, a GT-SUITE model of the 3D12 diesel engine cooling system was developed to quantify the effect of boost pressure on the engine energy balance and to propose appropriate retrofit measures. The simulations show that increasing boost pressure raises the heat transferred to the coolant from 119.8 kW to 163.6 kW, increasing the engine coolant outlet temperature from 84.38 °C to 93.94 °C, which approaches the upper allowable operating limit. Several improvement options were evaluated; among them, increasing the flow rates in both the freshwater and seawater circuits provided the most pronounced benefit. When the freshwater flow rate was increased by 320 L/min and the seawater flow rate by 210 L/min, the engine coolant outlet temperature was returned to the baseline operating range (84–86 °C). These results indicate that optimizing pump flow rates is a practical and effective approach to maintaining the engine's optimal thermal regime under turbocharged operation.*

Keywords: 3D12 diesel engine; Turbocharging; Cooling system; Energy balance.

Nghiên cứu phương pháp xác định đồng thời áp suất cuối nén trong buồng đốt cho động cơ diesel Kamaz 740

Phan Đoàn Việt Đức[1] Phạm Xuân Phương[2]
[1]Hệ QLHV Sau đại học, [2]Viện Cơ khí động lực
Email: vietduchvktqsk49@gmail.com, Tel: 0963.299.795

Tóm tắt

Áp suất cuối nén (ký hiệu p_{cn}) là chỉ tiêu quan trọng phản ánh tình trạng kỹ thuật của buồng đốt thông qua độ lọt khí. Việc xác định p_{cn} giúp đánh giá mức độ mài mòn nhóm piston, xéc măng, xy lanh, độ kín xu páp nạp, xupap xả và đệm nắp máy động cơ. Với động cơ nhiều xy lanh như KAMAZ 740, các phương pháp đo truyền thống (đồng hồ đo nén cơ khí, đo từng xy lanh) thường mất nhiều thời gian, khó bảo đảm điều kiện đo đồng nhất giữa các xy lanh và phải quy đổi thủ công giá trị đo (theo tốc độ động cơ) sang giá trị tiêu chuẩn. Bài báo đề xuất phương pháp xác định p_{cn} đồng thời trên nhiều xy lanh tự động chuyển đổi giá trị đo về giá trị tham chiếu và đánh giá. Phương pháp hướng tới tăng độ tin cậy so sánh giữa các xy lanh, rút ngắn thời gian chẩn đoán.

Từ khóa: áp suất cuối nén, KAMAZ 740, chẩn đoán tình trạng kỹ thuật, độ lọt khí buồng đốt.

1. Đặt vấn đề

Động cơ diesel KAMAZ 740 thuộc nhóm động cơ nhiều xy lanh được sử dụng rộng rãi trong các phương tiện tải nặng và thiết bị làm việc trong điều kiện khắc nghiệt. Trong quá trình khai thác, các cụm chi tiết chịu tải nhiệt–cơ lớn như nhóm piston–xéc măng–xy lanh, cơ cấu phối khí và hệ thống làm kín nắp máy thường bị mài mòn, biến dạng hoặc suy giảm chất lượng làm việc theo thời gian. Các hư hỏng đặc trưng như mòn xéc măng, xước thành xy lanh, kẹt xéc măng, hở/không kín xupáp, hỏng gioăng quy lát, nứt nắp máy… đều dẫn đến hiện tượng rò lọt khí trong quá trình nén. Hệ quả trực tiếp là giảm độ nén, làm động cơ khó khởi động (đặc biệt khi nguội), công suất giảm, tiêu hao nhiên liệu tăng, khói xả tăng và độ ổn định làm việc kém. Đối với động cơ nhiều xy lanh, sự suy giảm độ nén không chỉ thể hiện ở giá trị tuyệt đối mà còn ở mức không đồng đều giữa các xy lanh, gây mất cân bằng mô-men, rung động, tiếng ồn, tăng tải động lên trục khuỷu–thanh truyền và giảm tuổi thọ của các cụm chi tiết liên quan.

Trong các chỉ tiêu chẩn đoán tình trạng kỹ thuật động cơ, áp suất cuối nén p_{cn}là đại lượng có ý nghĩa thực tiễn cao vì phản ánh trực tiếp mức độ kín khít của buồng cháy tại thời điểm cuối quá trình nén (gần ĐCT), đồng thời có mối quan hệ chặt chẽ với khả năng tự bốc cháy của nhiên liệu trong điều kiện khởi động và làm việc tải thấp. Việc xác định p_{cn}cho phép đánh giá nhanh tình trạng nhóm piston–xéc măng–xy lanh và cơ cấu phối khí, khoanh vùng các xy lanh có biểu hiện "yếu" để phục vụ quyết định bảo dưỡng/sửa chữa.

Bên cạnh đó, trong điều kiện bảo dưỡng thực tế, yêu cầu đặt ra là đánh giá nhanh – so sánh tin cậy – lưu vết được số liệu để theo dõi xu hướng suy giảm theo thời gian khai thác. Điều này thúc đẩy nhu cầu chuyển từ đo cơ khí đơn giản sang các phương pháp đo có hỗ trợ cảm biến và xử lý số liệu. Đặc biệt, với mục tiêu đánh giá tính đồng đều giữa các xy lanh, việc đo "đồng thời" nhiều kênh trong cùng một phiên quay kéo sẽ giúp loại trừ ảnh hưởng của biến thiên tốc độ quay và điều kiện nạp theo thời gian, từ đó nâng cao độ tin cậy của so sánh. Ngoài ra, khi thu được tín hiệu áp suất theo thời gian hoặc theo góc quay trục khuỷu, người phân tích có thể trích xuất p_{cn}theo tiêu chí thống nhất, đánh giá mức phân tán theo chu kỳ, phát hiện bất thường dạng xung, và xây dựng các chỉ số chẩn đoán bổ sung (độ dốc tăng áp, độ rộng đỉnh, độ lặp lại chu kỳ…).

Đối với động cơ cỡ nhỏ và trung bình, các thiết bị thường dùng bao gồm bộ kiểm tra rò rỉ xy lanh (Cylinder Leakage Tester – CLT-2) [1,2], giúp phát hiện các hư hỏng như mòn xu-páp hoặc hỏng gioăng nắp máy. Một thiết bị khác, máy dò rò rỉ siêu âm (CTRL UL101), có khả năng ghi nhận tín hiệu siêu âm tần số cao phát sinh từ dòng khí rò để phát hiện rò rỉ bên ngoài [3,4]. Phương pháp này cho phép phát hiện rò rỉ tại gioăng nắp máy, ghế xu-páp, cũng như vết nứt trên thành xy lanh. Quá trình kiểm tra thường được tiến hành bằng cách bơm khí nén áp suất cao vào buồng cháy, sau đó thiết bị xác định vị trí và hướng của luồng khí rò, giúp khoanh vùng chính xác vị trí hư hỏng cơ học..

Với các động cơ diesel cỡ nặng, việc đánh giá rò rỉ xy lanh thường được tiến hành bằng thiết bị đo áp suất nén, nhằm xác định mức độ mài mòn của piston, xéc-măng và thành xy lanh trong quá trình quay khởi động. Thiết bị được nối qua cổng vòi phun nhiên liệu, và giá trị áp suất đo được được so sánh với giá trị tham chiếu của nhà sản xuất. Các hệ thống chẩn đoán tự động hiện đại đã nâng cao độ chính xác phép đo lên đến ±1,7%, cho phép phát hiện các lỗi như cháy xu-páp hay hư hỏng piston [5]. Tuy nhiên, nhiều công cụ chẩn đoán hiện nay vẫn thiếu tính tích hợp và độ chính xác còn hạn chế [6].

Nghiên cứu khả thi đầu tiên về phép đo áp suất cuối nén đồng thời được thực hiện trên động cơ Kamaz 740, nhằm phát hiện rò rỉ áp suất trong nhiều xy lanh cùng lúc. Bảng 1 trình bày so sánh các phương pháp kiểm tra độ kín buồng cháy phổ biến hiện nay.

Bảng 1. So sánh các phương pháp đánh giá độ kín buồng đốt của động cơ

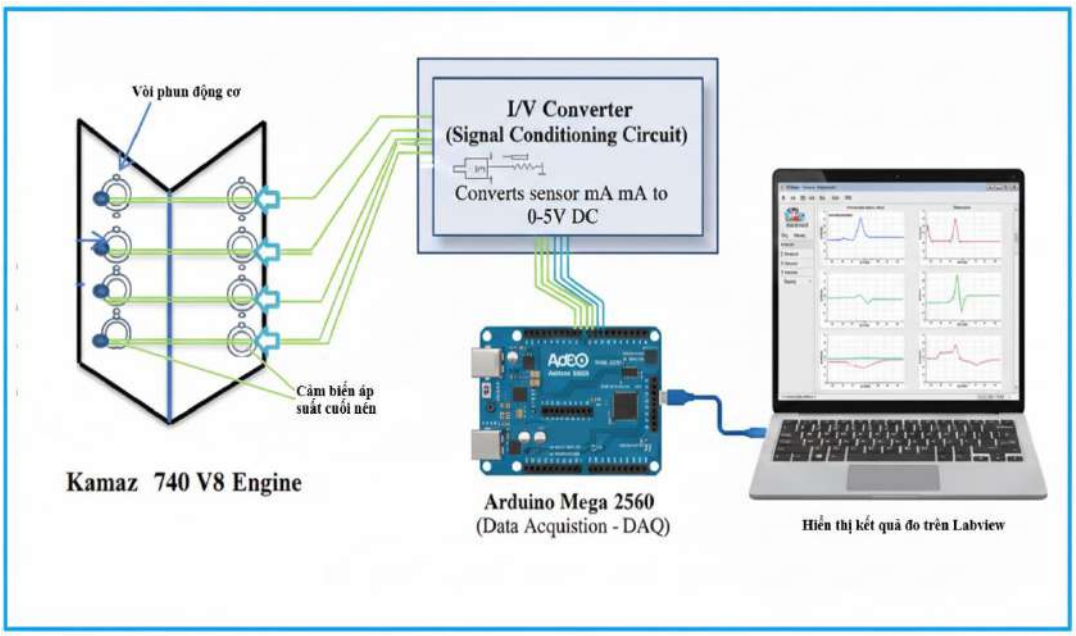
Thiết bị / Phương pháp	**Nguyên lý kỹ thuật**	**Phản ánh chủ yếu**	**Độ tin cậy**	**Hạn chế chính**
Đo nén bằng đồng hồ (đồng hồ đo nén)	Lắp đầu đo vào lỗ kim phun/bugi sấy, quay bằng máy đề để lấy giá trị áp suất nén cao nhất	Mức độ kín tổng thể của buồng cháy (xéc măng–xy lanh, xupáp, gioăng…)	Khá cao nhưng phụ thuộc điều kiện đo	Bị ảnh hưởng bởi tốc độ đề, nhiệt độ và tình trạng ắc quy; có sai số do đầu nối/van một chiều/thể tích chết; thường đo lần lượt từng xy lanh nên khó so sánh chuẩn
Hệ đo tự động (cảm biến áp suất + thu thập dữ liệu + phần mềm)	Dùng cảm biến đo áp suất trong xy lanh, thu dạng sóng theo thời gian (hoặc theo góc quay), phần mềm xử lý để lấy đỉnh nén và thống kê theo chu kỳ	Kín tổng thể; có thể phân tích dạng đường cong để gợi ý nguyên nhân rò	Cao nếu hiệu chuẩn và đồng bộ tốt	Chi phí cao; lắp đặt và chống nhiễu phức tạp; cần hiệu chuẩn và phần mềm xử lý phù hợp
Kiểm tra rò khí (Leak-down)	Đưa piston về điểm chết trên kỳ nén, cấp khí nén vào xy lanh, đo mức rò và xác định vị trí rò (đường nạp, đường xả, cácte, két nước)	Vị trí và mức rò (xupáp, gioăng, xéc măng–xy lanh, rò giữa các xy lanh)	Rất tốt để tìm đúng "chỗ hở"	Không cho biết trực tiếp khả năng nén khi động cơ quay; thao tác lâu; cần máy nén khí và đặt piston đúng vị trí

Thiết bị / Phương pháp	Nguyên lý kỹ thuật	Phản ánh chủ yếu	Độ tin cậy	Hạn chế chính
Phân tích khí lọt (blow-by)	Đo lưu lượng/áp suất khí lọt xuống cácte khi quay hoặc khi làm việc, so sánh theo chế độ	Mòn nhóm piston–xéc măng–xy lanh (rò xuống cácte)	Nhạy với mòn nhóm piston	Cần thiết bị chuyên dụng; bị ảnh hưởng bởi hệ thống thông hơi cácte và chế độ tải; khó tách riêng từng xy lanh nếu không có đồng bộ
Máy dò rò siêu âm	Dò tín hiệu siêu âm phát ra tại vị trí rò khí (thường cần tạo chênh áp)	Rò bề mặt như gioăng, xupáp, vết nứt ngoài	Nhạy với rò bên ngoài	Không phù hợp để định lượng mức nén; hạn chế khi hư hỏng nằm sâu bên trong (mòn xéc măng–xy lanh) nếu không tạo điều kiện rò rõ rệt

Xuất phát từ các vấn đề nêu trên, bài báo này tập trung nghiên cứu cách tiếp cận đồng thời xác định áp suất cuối nén p_{cn}cho động cơ diesel KAMAZ 740 dựa trên thu thập dữ liệu áp suất xy lanh và xử lý số liệu. Nội dung nghiên cứu nhằm: làm rõ tiêu chí xác định p_{cn}trong điều kiện đo quay kéo (motoring) và/hoặc khi đồng bộ theo góc quay; đề xuất quy trình đo–xử lý để bảo đảm các xy lanh được đánh giá dưới cùng điều kiện; minh họa khả năng áp dụng bằng kết quả thực nghiệm đo áp suất motoring cho hai xy lanh (xy lanh 1 và 2) theo thứ tự nổ. Kết quả là cơ sở để mở rộng triển khai đo đồng thời nhiều xy lanh trên động cơ KAMAZ 740, phục vụ chẩn đoán nhanh tình trạng kỹ thuật và hỗ trợ quyết định bảo dưỡng theo tình trạng.

2. Phương pháp nghiên cứu

2.1. Công cụ nghiên cứu



Hình 1. Sơ đồ nguyên lý của phương pháp đo áp suất cuối nén được phát triển trong nghiên cứu này.

Bảng 2. Thông số kỹ thuật của động cơ khảo nghiệm

STT	Mô tả	Giá trị	Đơn vị
1	Đường kính xy lanh (D)	120	mm
2	Số xy lanh (i)	8	–
3	Số kỳ làm việc (τ)	4	–
4	Hành trình piston (S)	120	mm
5	Thể tích công tác (Vh)	10,85	dm^3
6	Tỷ số nén (ε)	16,5	–
7	Góc mở xu-páp nạp	15	độ
8	Góc đóng xu-páp nạp	40	độ
9	Góc mở xu-páp xả	40	độ
10	Góc đóng xu-páp xả	15	độ
11	Góc lệch chữ V (γ)	90	độ
12	Công suất cực đại (Ne_max)	240	HP
13	Tốc độ tại công suất cực đại	2200	vòng/phút

Trong thí nghiệm, mỗi xy lanh được lắp một vòi phun giả tại vị trí vòi phun nguyên bản. Tín hiệu áp suất được lấy thông qua đường nối ra từ vòi phun giả, qua van một chiều, và được đưa tới cảm biến áp suất đặt ở phía sau van nhằm bảo đảm dòng khí chỉ chảy một chiều và ổn định tín hiệu đo.

Bảng 3. Thông số cảm biến áp suất

STT	Mục	Chi tiết
1	Tên thiết bị	Cảm biến áp suất Sensys
2	Model	M5256-C3079E-100BG
3	Số sê-ri	090522M095
4	Hãng sản xuất / Quốc gia	Sensys, Hàn Quốc
5	Dải đo	0 – 100 bar
6	Sai số	±0,5% FS
7	Chuẩn hiệu chuẩn	DMH-DP-160
8	Phương pháp kiểm tra	Hiệu chuẩn tại các mức áp suất định sẵn
9	Điều kiện môi trường	25°C, độ ẩm 60%
10	Nhiệt độ làm việc	–40 đến 125°C
11	Ngõ ra tín hiệu	4–20 mA (2 dây)
12	Kết nối điện	Dây cáp, đầu nối Packard hoặc Mini-DIN
13	Vật liệu chế tạo	Thép không gỉ SUS630
14	Khối lượng	Khoảng 85 g
15	Chống sốc	50G / 11 ms
16	Chống rung	20G tại 20–200 Hz

2.2. Các bước thực hiện

Áp suất trong xy lanh (đặc biệt là áp suất cuối nén p_{cn}) được lấy ra thông qua vòi phun giả lắp thay thế tại vị trí vòi phun nguyên bản của động cơ KAMAZ 740. Mỗi vòi phun giả được gia

công có đường dẫn áp suất nối ra ngoài, bảo đảm khả năng truyền áp từ buồng cháy ra hệ đo. Trên đường dẫn này bố trí van một chiều nhằm hạn chế dao động hồi áp và ngăn dòng khí ngược trở lại, đồng thời góp phần bảo vệ hệ đo và tăng độ ổn định của tín hiệu trong quá trình quay kéo. Sau van một chiều, áp suất được đưa tới cảm biến áp suất; cảm biến chuyển đổi đại lượng áp suất $p(t)$trong đường dẫn thành tín hiệu điện dạng dòng tiêu chuẩn 4–20 mA, tương ứng tuyến tính với dải đo của cảm biến.

Do Arduino chỉ đo được tín hiệu dạng điện áp (0–5 V), tín hiệu dòng 4–20 mA từ cảm biến được đưa qua mạch chuyển đổi dòng–điện áp (I/V) trước khi vào Arduino. Mạch I/V sử dụng điện trở shunt $R_s = 250\ \Omega$để biến đổi dòng thành điện áp theo quan hệ:

$$U(t) = I(t) \cdot R_s \qquad (1)$$

Với cấu hình này, tín hiệu 4–20 mA được quy đổi thành dải điện áp:

$$U_{min} = 4\ \text{mA} \cdot 250\ \Omega = 1.0\ \text{V}, U_{max} = 20\ \text{mA} \cdot 250\ \Omega = 5.0\ \text{V}$$

Như vậy, điện áp đưa vào Arduino nằm trong khoảng 1–5 V, phù hợp với dải đo của ngõ vào analog và tận dụng tốt độ phân giải ADC.

Arduino thực hiện lấy mẫu điện áp $U(t)$tại chân analog theo tần số đặt trước, số hóa qua bộ ADC (dải 0–5 V), sau đó quy đổi về áp suất theo đặc tính tuyến tính của cảm biến. Trước hết, dòng được suy ra từ điện áp:

$$I(t) = \frac{U(t)}{R_s} = \frac{U(t)}{250\ \Omega} \qquad (2)$$

và áp suất được tính theo quan hệ tuyến tính chuẩn của cảm biến 4–20 mA:

$$\mathrm{p(t)} = \mathrm{p_{min}} + \frac{\mathrm{I(t) - 4\ mA}}{\mathrm{16\ mA}}\ (\mathrm{p_{max}} - \mathrm{p_{min}}) \qquad (3)$$

Trong đó $[p_{min}, p_{max}]$là dải áp suất danh định của cảm biến. Dữ liệu $p(t)$sau khi tính toán được Arduino đóng gói theo khung truyền (gồm thời gian hoặc số thứ tự mẫu và giá trị áp suất) và gửi qua giao tiếp Serial/USB về máy tính.

Sau mạch I/V với $R_s = 250\ \Omega$, dòng 4–20 mA được chuyển thành điện áp 1–5 V:

- 4 mA → 1.0 V
- 20 mA → 5.0 V

Với cảm biến tuyến tính 1–100 bar, ta có quan hệ:

$$p(U) = 1 + \frac{U-1}{5-1}\ (100 - 1) = 1 + (U - 1) \cdot \frac{99}{4}$$

Do đó: $p(U) = 1 + 24.75\ (U - 1)$(bar)

Arduino đọc ADC 10-bit (0–1023) với $V_{tc} = 5$V:

$$U = \frac{\text{ADC}}{1023} \cdot 5$$

Thế vào công thức áp suất:

$$p = 1 + 24.75\left(\frac{\text{ADC}}{1023} \cdot 5 - 1\right)\ (\text{bar})$$

Trên máy tính, chương trình LabVIEW thu nhận dữ liệu từ cổng COM, thực hiện tách khung và kiểm tra hợp lệ, sau đó hiển thị theo thời gian thực dạng sóng áp suất $p(t)$. Từ tín hiệu này, LabVIEW

(hoặc Arduino) tiếp tục xác định các tham số đặc trưng phục vụ chẩn đoán như đỉnh áp suất mỗi chu kỳ nén $p_{\max}$; trong chế độ quay kéo khi không có cháy, $p_{\max}$được sử dụng như giá trị xấp xỉ của áp suất cuối nén p_{cn}cho từng chu kỳ. Đồng thời, LabVIEW lưu dữ liệu ra tệp (CSV/TDMS) để phục vụ phân tích hậu kỳ và so sánh giữa các xy lanh trong cùng điều kiện đo.

2.3. Thu thập và xử lý số liệu

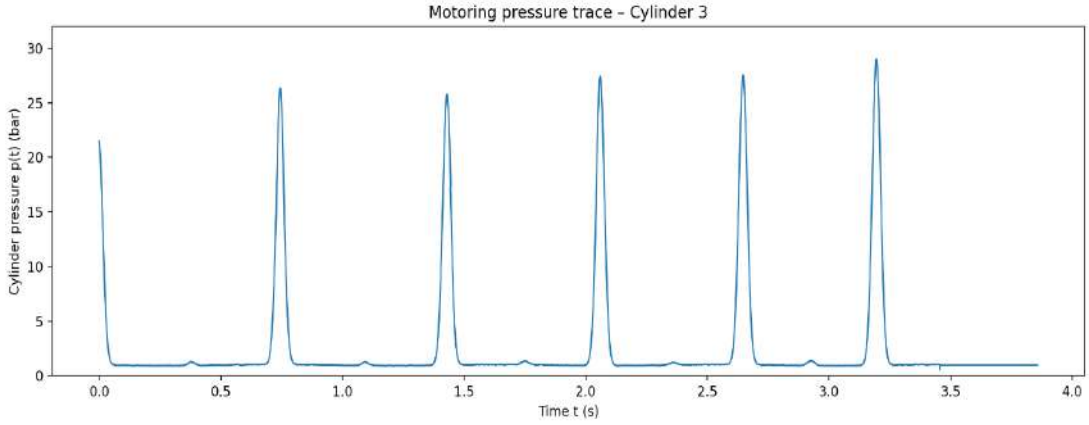
2.3.1. Quy trình xử lý số liệu áp suất cuối nén

Quy trình xử lý được xây dựng theo các bước chính: tiền xử lý → phát hiện đỉnh → tính tốc độ quay → chuẩn hoá theo tốc độ → làm trơn và tính giá trị đại diện → tổng hợp 8 xy lanh.

Từ tín hiệu áp suất theo thời gian $p(t)$, các đỉnh áp suất cực đại được xác định bằng phương pháp phát hiện cực đại cục bộ. Các đỉnh này được ký hiệu là: $p_1(t_1),\ p_2(t_2),\ \ldots,\ p_k(t_k)$

trong đó $p_i = p_{z,\max,i}$là giá trị áp suất cuối nén của chu kỳ thứ i, còn $t_i$là thời điểm xảy ra đỉnh áp suất tương ứng.

Quá trình xác định các đỉnh áp suất cần đảm bảo loại bỏ các đỉnh giả do nhiễu hoặc dao động nhỏ của tín hiệu. Các đỉnh được chọn phải tương ứng với thời điểm cuối kỳ nén và có giá trị áp suất vượt quá ngưỡng xác định trước.



Hình 2. Sự giao động về áp suất cuối nén khi khởi động trong xylanh

2.3.2. Xác định các đỉnh áp suất nén và trích P_{max}

Từ tín hiệu $p(t)$, xác định lần lượt các đỉnh nén $t_1, t_2, \ldots, t_m$(thông thường chọn $m = 6$). Áp suất cực đại tại mỗi đỉnh: $P_i = P_{\max}(t_i), i = 1..m$

Trong nghiên cứu này, để đảm bảo độ tin cậy, chọn 4–6 chu kỳ nén và ưu tiên các đỉnh có biên độ lớn, tách bạch khỏi nhiễu.

2.3.3. Tính tốc độ quay khi khởi động

Từ các thời điểm đỉnh:

$$\Delta t_i = t_{i+1} - t_i, i = 1..(m-1)$$

Tốc độ quay tức thời:

$$n_i = \frac{120}{\Delta t_i} \qquad (4)$$

Giá trị $n_i$dùng để đánh giá độ ổn định khi khởi động và là tham số phục vụ chuẩn hoá áp suất.

2.3.4. Chuẩn hoá áp suất về tốc độ tham chiếu

Tốc độ tham chiếu và bảng chuyển đổi từ áp suất đo được sang áp suất tham chiếu thường được nhà sản xuất cung cấp. Từ bảng cung cấp, một mô hình số có thể được phát triển để chuyển đổi các đỉnh áp suất đo được thành các đỉnh áp suất tham chiếu. Trong nghiên cứu này, ví dụ, việc nội suy được thực hiện bằng cách sử dụng gói Python cho chuyển đổi, và điều này dẫn đến mối tương quan logarit giữa tốc độ động cơ (n) và áp suất nén tham chiếu (y) như được thể hiện trong Phương trình dưới đây:

$$y_P = 8.2204\ln(\mathrm{n}) - 22.104 \quad [7]$$

Chuyển đổi phương trình, phương trình để vẽ qua điểm

$$y_{p0_14} = p_{0-14}(0.587\ln(\mathrm{n}) - 1.57) \quad [8]$$

Phương trình đi qua $p_{0\text{-}14}$, y_{p0_14} sẽ là:

$$y_{p0_x} = p_{0-x}(0.587\ln(\mathrm{n}) - 1.57) \quad [9]$$

Giá trị $y_{p0\text{-}x}$ và tốc độ động cơ n được đo bằng thực nghiệm.

2.3.5. Làm trơn theo chu kỳ và tính giá trị đại diện cho một xy lanh

Do tốc độ khởi động dao động, áp suất ở từng đỉnh có thể biến thiên. Để giảm ảnh hưởng của đỉnh bất thường, áp dụng trung bình trượt 2 điểm:

$$\hat{P}_j = \frac{P_j^* + P_{j+1}^*}{2}, j = 1..(m-1) \quad (5)$$

Sau đó lấy giá trị đại diện của xy lanh:

$$\bar{P}_{cyl} = \frac{1}{m-1}\sum_{j=1}^{m-1}\hat{P}_j \quad (6)$$

Với $m = 6$, ta có 5 giá trị $\hat{P}_1..\hat{P}_5$và:

$$\bar{P}_{cyl} = \frac{\hat{P}_1 + \hat{P}_2 + \hat{P}_3 + \hat{P}_4 + \hat{P}_5}{5} \quad (7)$$

2.3.6. Phương pháp đánh giá phục vụ chẩn đoán

Sau khi xử lý tín hiệu áp suất theo quy trình, với mỗi xy lanh *i* thu được giá trị áp suất cuối nén đại diện $\bar{P}_i$(bar). Tập kết quả của toàn động cơ KamAZ 740 được biểu diễn dưới dạng vector:

$$\bar{\mathrm{P}} = \{\bar{P}_1, \bar{P}_2, \ldots, \bar{P}_8\}$$

Trong đó, $\bar{P}_i$là đại lượng dùng thống nhất cho các bước đánh giá ở mục này. Phương pháp đánh giá được tiến hành theo ba hướng: (i) so sánh với ngưỡng tiêu chuẩn, (ii) so sánh theo "xy lanh kề" để nhận biết xu hướng bất thường theo cặp/khu vực, và (iii) so sánh trong từng vế (bank) để phát hiện sai lệch theo cụm.

Đánh giá theo ngưỡng tiêu chuẩn tại $n_{tc} = 230$*vòng/phút*

Trong bài báo này sử dụng ngưỡng tối thiểu của áp suất cuối nén tại 230 vòng/phút [10]:

$$P_{min,230} = 28.5 \text{ bar}$$

Việc đánh giá theo ngưỡng được thực hiện trực tiếp trên $\bar{P}_i$vì $\bar{P}_i$đã được chuẩn hoá về đúng điều kiện tốc độ tham chiếu 230 vòng/phút. Để lượng hoá mức đạt/thiếu so với ngưỡng, xác định sai lệch tương đối:

$$\delta_i = \frac{\bar{P}_i - P_{min,230}}{P_{min,230}} \times 100\% \quad (8)$$

trong đó $P_{min,230} = 28.5$bar là ngưỡng tối thiểu tại 230 vòng/phút. Trong thực hành kiểm tra áp suất nén, ngoài việc so với ngưỡng tuyệt đối, tiêu chí quan trọng là độ đồng đều giữa các xy lanh; theo hướng dẫn kỹ thuật, chênh lệch áp suất nén giữa các xy lanh thường không nên vượt quá 10% [11]. Khi đó, kết quả được phân thành hai nhóm:

Đạt yêu cầu: $\delta_i \geq -10\%$

Trường hợp này cho thấy áp suất cuối nén của xy lanh không thấp quá 10% so với ngưỡng tham chiếu; xy lanh được xem là đáp ứng yêu cầu tối thiểu trong điều kiện khởi động. Tuy nhiên, để tăng độ tin cậy chẩn đoán, vẫn cần đối chiếu thêm độ đồng đều giữa các xy lanh (theo cặp xy lanh kề nhau và theo từng vế).

Không đạt yêu cầu: $\delta_i < -10\%$

Khi mức suy giảm vượt quá 10%, xy lanh được đánh dấu không đạt và có khả năng tồn tại suy giảm kín khít đáng kể. Trường hợp này cần ưu tiên phân tích bổ trợ (so sánh theo cặp xy lanh kề nhau, so sánh theo vế) và thực hiện thử nghiệm bổ sung (ví dụ thử dầu) nhằm khoanh vùng nguyên nhân thuộc nhóm piston–xéc măng–xy lanh hoặc nhóm xupáp/gioăng quy lát.

Như vậy, phân loại theo mốc 10% cho phép kết luận đạt/không đạt một cách rõ ràng, đồng thời hỗ trợ lựa chọn xy lanh cần ưu tiên kiểm tra trong quá trình chẩn đoán.

Đánh giá theo "cặp xy lanh kề nhau" nhằm nhận biết dấu hiệu bất thường theo cặp

Bên cạnh việc so sánh với ngưỡng tuyệt đối, một hướng đánh giá quan trọng trong chẩn đoán là xem xét mẫu hình phân bố của áp suất cuối nén theo bố trí xy lanh trên động cơ. Trong một số dạng hư hỏng liên quan đến khu vực làm kín chung (ví dụ bề mặt gioăng, vùng nắp máy, hoặc sai lệch cơ khí theo cụm), áp suất cuối nén có xu hướng suy giảm theo cặp hoặc theo một nhóm xy lanh lân cận. Do đó, luận văn sử dụng phương pháp đánh giá dựa trên từng cặp xy lanh kề nhau trong cùng một vế (theo thứ tự dọc thân máy) để phát hiện bất thường theo cặp.

Giả sử (i, j)là một cặp xy lanh kề nhau (được xác định theo sơ đồ bố trí thực tế của KamAZ 740). Với mỗi cặp, xét chênh lệch áp suất cuối nén đại diện (đã xử lý và chuẩn hoá) như sau:

$$D_{ij} = |\bar{P}_i - \bar{P}_j| \quad (9)$$

Đồng thời, để giảm ảnh hưởng mức nền chung (ví dụ toàn động cơ cùng thấp do tốc độ đề hoặc điều kiện đo), sử dụng chênh lệch tương đối theo trung bình của cặp:

$$d_{ij} = \frac{|\bar{P}_i - \bar{P}_j|}{(\bar{P}_i + \bar{P}_j)/2} \times 100 \quad (10)$$

Khi phân tích kết quả theo từng cặp (i, j), có hai xu hướng mẫu hình thường gặp:

Xu hướng suy giảm theo cặp/khu vực: cả hai xy lanh trong cặp đều có $\bar{P}$ thấp so với ngưỡng (tức $\delta_i < 0\%$và $\delta_j < 0\%$, đặc biệt khi rơi vào vùng $\delta \leq -10\%$). Việc so sánh cặp xylanh này gợi ý khả năng tồn tại nguyên nhân ảnh hưởng theo khu vực làm kín chung hoặc theo cụm lân cận. Khi gặp trường

hợp này, cần ưu tiên kiểm tra các nguyên nhân "chung" theo vùng (bề mặt làm kín, gioăng, điều kiện lắp ráp, hoặc các yếu tố cơ khí liên đới), và thực hiện thử nghiệm bổ sung để xác nhận.

Xu hướng suy giảm cục bộ: chỉ một xy lanh trong cặp có $\bar{P}$ thấp rõ rệt trong khi xy lanh còn lại vẫn bình thường hoặc cao hơn đáng kể, thể hiện qua d_{ij} tăng lớn. Mẫu hình này thường gợi ý hư hỏng mang tính cục bộ tại riêng xy lanh bị thấp (ví dụ xupáp không kín, xéc-măng/xy lanh mòn cục bộ, hư hỏng piston…), trong khi các yếu tố theo khu vực ít có khả năng là nguyên nhân chính.

Nhờ phân tích theo từng cặp xy lanh kề nhau, kết luận chẩn đoán tránh việc chỉ dựa vào một giá trị ngưỡng đơn lẻ, đồng thời hỗ trợ định hướng kiểm tra theo khu vực (khi hai xy lanh cùng suy giảm) hoặc theo xy lanh cụ thể (khi chỉ một xy lanh suy giảm trong cặp).

3. Kết quả và thảo luận

Bảng 4 tổng hợp kết quả đo trên hai xy lanh liền kề của động cơ diesel Kamaz 740 V8 lắp trên xe tải hạng nặng, được đặt trong phòng thí nghiệm (coi như là động cơ mới). Áp suất nén đỉnh tham chiếu của cả 2 xi lanh đều nằm trong giới hạn cho phép, xác nhận động cơ đang ở tình trạng làm việc tốt.

Áp suất nén tham chiếu chuẩn tại 230 rpm là 28,5 kG/cm². Như thể hiện trong Bảng 4, áp suất nén đỉnh tham chiếu lớn nhất của tất cả các xi lanh đo được đều vượt ngưỡng này, cho thấy độ kín buồng cháy đạt yêu cầu. Mức chênh lệch nhỏ về độ kín giữa các xi lanh, mặc dù động cơ đã vận hành trong thời gian dài, gợi ý hiện tượng mài mòn đồng đều của các cụm chi tiết trong buồng cháy. Bên cạnh đó, phân tích thống kê cho phép đánh giá toàn diện tình trạng từng xi lanh, đồng thời xét đến thứ tự nổ của động cơ nhằm hỗ trợ chẩn đoán chính xác hơn.

Bảng 4. Kết quả đo áp suất cuối nén của 2 xy lanh liền kề nhau

TT	Xy lanh 1		Xy lanh 2	
	bar	kG/cm²	bar	kG/cm²
1	28,34	28,91	28,48	29,05
2	28,35	28,92	28,47	29,04
3	28,34	28,91	28,49	29,06

Nghiên cứu này đã xây dựng và đề xuất một phương pháp xác định đồng thời áp suất cuối nén cho động cơ diesel Kamaz 740. Hệ thống đo được thiết kế theo hướng chi phí hợp lý và dễ triển khai tại hiện trường, bao gồm: vòi phun giả lấy áp suất từ buồng cháy, van một chiều ổn định đường áp, cảm biến áp suất ngõ ra 4–20 mA, mạch chuyển đổi dòng–điện áp và bộ thu thập dữ liệu dựa trên Arduino, sau đó truyền dữ liệu về máy tính để hiển thị và lưu trữ trên LabVIEW. Chuỗi đo–xử lý này cho phép theo dõi áp suất theo thời gian thực, tạo nền tảng để đánh giá độ nén và so sánh tình trạng kỹ thuật giữa các xy lanh.

Kết quả thực nghiệm minh họa trên hai xy lanh (1 và 2) theo cặp của động cơ Kamaz 740 ở chế độ quay kéo cho thấy tín hiệu áp suất có các đỉnh nén lặp lại rõ rệt; đỉnh áp suất tăng dần theo thời gian quay và hai xy lanh có mức đỉnh tương đương trong cùng điều kiện đo. Điều này chứng tỏ hệ đo có khả năng ghi nhận ổn định diễn biến áp suất nén và có thể sử dụng giá trị đỉnh áp suất mỗi chu kỳ như đại lượng đặc trưng để ước lượng áp suất cuối nén, phục vụ đánh giá độ đồng đều độ kín giữa các xy lanh.

Trong hướng phát triển tiếp theo, nghiên cứu sẽ mở rộng số kênh đo để tiến tới đo đồng thời nhiều xy lanh trên động cơ KAMAZ 740, đồng thời bổ sung đồng bộ theo góc quay trục khuỷu nhằm chuẩn hóa kết quả theo tốc độ quay và giảm sai lệch do điều kiện đề. Ngoài ra, sẽ đánh giá

ảnh hưởng của các yếu tố vận hành (nhiệt độ động cơ, điện áp ắc quy, điều kiện nạp) tới kết quả đo, và xem xét tích hợp các thuật toán phát hiện bất thường/tự động phân loại hư hỏng để nâng cao mức độ tự động hóa và hiệu quả chẩn đoán.

Tài liệu tham khảo

1. Baechtel, J. (2023). *Leakdown lowdown: Diagnosing engine condition by leakdown testing*. *Hot Rod Magazine*.
2. Sluder, C. S., & West, B. H. (2011). *Limitations and recommended practice in the use of compression and leak-down tests to monitor gradual engine degradation*. SAE International Journal of Engines, 4(3), 2767–2777.
3. Johnson, C., Druschitz, E., & Druschitz, A. (2007). *Ultrasonic leak test for automotive brake caliper. In Materials Science Forum*. Trans Tech Publications.
4. Hoover, A. (2002). *Maryland company expanding technology in space—NASA won't leave Earth without the CTRL UL101*. CTRL Systems Inc.
5. Saraiev, O., et al. (2020). *Automated diagnostic system for engine cylinder-piston group*. SAE Technical Paper.
6. Bhosale, R., et al. (2017). Study on leak testing methods. *IJSRD—International Journal for Scientific Research and Development*, 1618–1621.

7.8.9. Duong, Q. C., Vu, M. D., Pham, P. X., & Khuong, H. T. (2024). A novel non-intrusive approach for simultaneous compression pressure leakage diagnostics in multi-cylinder diesel engines. *Journal of Science and Transport Technology*.

10. Tổng Công ty Cổ phần KamAZ. (2017). *Bộ tài liệu hướng dẫn bảo dưỡng và sửa chữa ô tô KamAZ 43253, 43255, 65115, 65116, 65117, 6540, 65111 (tiêu chuẩn châu Âu Euro-2, -3, -4), Tập I: Hệ thống động truyền lực; khung xe; móc kéo; cabin*. Hà Nội, Việt Nam.
11. Engine Australia. (2018). *Compression tests and static leak down tests (Service Bulletin SB041).* Engine Australia – Technical Service Bulletins.

An Approach for Simultaneous Determination of Non-Firing Compression Pressure in the Combustion Chamber of the KAMAZ 740 Diesel Engine

Abstract: *Non-firing compression pressure (denoted as p_{cn}, i.e., end-of-compression pressure measured during cranking with fuel injection disabled) is an important indicator of combustion-chamber sealing condition through blow-by leakage. Determining p_{cn}enables assessment of wear in the piston–ring–cylinder assembly, sealing performance of the intake and exhaust valves, and the condition of the cylinder head gasket. For multi-cylinder engines such as the KAMAZ 740, conventional methods (mechanical compression gauges and sequential cylinder-by-cylinder testing) are time-consuming, make it difficult to maintain identical test conditions among cylinders, and require manual conversion of measured values (as a function of cranking speed) to standard reference conditions. This paper proposes a method for simultaneous determination of p_{cn}across multiple cylinders, with automatic conversion to a reference value and subsequent evaluation. The proposed approach aims to improve the reliability of inter-cylinder comparison and shorten diagnostic time.*

Keywords: non-firing compression pressure, KAMAZ 740, technical condition diagnostics, combustion-chamber blow-by leakag

Nghiên cứu ảnh hưởng của góc tấn dòng khí đầu vào tới thông số tầng máy nén đầu tiên động cơ TB3-117.

Trần Duy Khánh[1*], Lê Tiến Dương[2]

[1]Hệ Quản lý Học viên SĐH; Học viện Kỹ thuật Quân sự

[2]Viện Cơ khí động lực; Học viện Kỹ thuật Quân sự

[*]E-mail: Khanhht204@gmail.com, Tel: 0353121796

Tóm tắt

Động cơ tua bin khí TB3-117 được sử dụng phổ biến trên các dòng trực thăng họ Mi nhờ ưu điểm nhỏ gọn và khả năng tạo ra công suất lớn. Trong đó máy nén là một trong các bộ phận quan trọng nhất có nhiệm vụ cung cấp đủ không khí cần thiết cho quá trình động cơ hoạt động. Đặc tính dòng khí đầu vào có ảnh hưởng trực tiếp đến các thông số tầng máy nén đầu tiên như hiệu suất, tỉ số tăng áp, lưu lượng không khí qua máy nén và từ đó tác động đến thông số các tầng máy nén tiếp theo và của toàn máy nén động cơ tuabin khí. Trong bài báo này, nhóm tác giả trình bày kết quả tính toán ảnh hưởng của góc tấn dòng khí đầu vào tới hiệu suất tầng cánh công tác, trường áp suất và trường vận tốc bằng phương pháp mô phỏng số CFD. Kết quả nghiên cứu là cơ sở cho các nghiên cứu tiếp theo nhằm nâng cao hiệu quả khai thác và độ tin cậy máy nén động cơ TB3-117.

Từ khóa: *Máy nén; góc tấn, cánh công tác; áp suất; TB3-117; CFX*

1. Mở đầu

Động cơ tua bin khí là một loại động cơ nhiệt với nhiều ưu điểm vượt trội, hiện tại đang được ứng dụng rộng rãi trong ngành hàng không, hàng hải và công nghiệp, bao gồm cả lĩnh vực quân sự và kinh tế. Đối với lĩnh vực hàng không, động cơ tua bin khí bắt đầu được sử dụng cuối chiến tranh thế giới lần thứ II. Sau một thời gian ngắn sử dụng, ĐCTBK được trang bị hầu hết trên các loại máy bay và trực thăng hàng không dân dụng cũng như lực lượng Không quân. Máy nén động cơ tua bin khí là một bộ phận quan trọng trong cấu tạo tổng thể của động cơ, máy nén không chỉ quyết định lượng không khí cấp cho buồng đốt mà còn chi phối trực tiếp đến hiệu suất tổng thể và giới hạn ổn định của động cơ.

Trong điều kiện khai thác thực tế, dòng khi đi vào máy nén thường không đồng đều và không đúng theo thiết kế. Nguyên nhân là do ảnh hưởng của hình dạng cửa hút, điều kiện môi trường và các yếu tố khí động khác. Sự lệch hướng dòng khí đầu vào làm thay đổi phân bố vận tốc và áp suất tại cửa vào máy nén, từ đó ảnh hưởng đến khả năng làm việc của các tầng cánh, làm giảm hiệu suất, tỉ số tăng áp và có thể gây ra các hiện tượng mất ổn định như rung động khí động hoặc mất ổn định dòng chảy [1, 2].

Nhiều công trình nghiên cứu trong và ngoài nước đã chỉ ra rằng sự thay đổi góc tấn dòng khí đầu vào có thể gây ra biến dạng phân bố áp suất trên bề mặt cánh, làm xuất hiện các vùng tách dòng cục bộ và dẫn đến suy giảm hiệu suất của máy nén. Trong các trường hợp bất lợi, hiện tượng này có thể làm máy nén tiến gần đến trạng thái mất ổn định như nghẽn dòng hoặc dao động khí động [3–5]. Do đó, việc đánh giá ảnh hưởng của góc tấn dòng khí đầu vào là một nội dung quan trọng trong nghiên cứu đặc tính làm việc của máy nén hướng trục. Tuy nhiên, phần lớn các nghiên cứu này được thực hiện trên các mô hình máy nén tiêu chuẩn hoặc máy nén tốc độ cao, trong khi các nghiên cứu chuyên sâu đối với máy nén của động cơ TB3-117 còn rất hạn chế.

Động cơ TB3-117 là loại động cơ tua bin khí được sử dụng phổ biến trên các dòng trực thăng họ Mi đang khai thác tại Việt Nam. Việc nghiên cứu ảnh hưởng của góc tấn dòng khí đầu vào đến đặc tính làm việc của tầng cánh công tác đầu tiên có ý nghĩa quan trọng trong việc đánh giá khả năng làm việc ổn định và nâng cao hiệu quả khai thác động cơ.

Trong nghiên cứu này, góc của dòng khí đầu vào được thay đổi trong khoảng từ -15° đến 15°, từ đó thay đổi góc tấn, tương ứng với các trạng thái từ dòng khí gần như lý tưởng đến điều kiện làm việc gần và vượt ngưỡng ổn định của tầng cánh. Thông qua mô phỏng số, các đặc trưng khí động của tầng cánh công tác được phân tích nhằm làm rõ xu hướng biến đổi của hiệu suất, trường áp suất và trường vận tốc theo góc tấn dòng khí đầu vào.

2. Thông số máy nén

Nghiên cứu tập trung phân tích ảnh hưởng của sự biến dạng dòng khí đầu vào do thay đổi hướng dòng đến hiệu suất của rotor tầng máy nén đầu tiên động cơ TB3-117. Tầng máy nén này được lựa chọn làm đối tượng nghiên cứu do đóng vai trò quyết định đến đặc tính khí động và hiệu suất làm việc của các tầng máy nén tiếp theo cũng như của toàn bộ máy nén. Rotor tầng máy nén đầu tiên thuộc loại máy nén hướng trục, gồm các cánh công tác có khe hở đầu cánh đặc trưng theo thiết kế của động cơ TB3-117 và quay theo chiều xác định của trục máy.

Bảng 1. Thông số cơ bản của máy nén động cơ TB3-117 [11]

TT	Tên thông số	Giá trị	Thứ nguyên
1	Lưu lượng khí qua máy nén	9.6	kg/s
2	Tỷ số tăng áp	9.55	
3	Nhiệt độ không khí sau máy nén	608	K
4	Tốc độ dọc trục tại cửa vào máy nén	149	m/s
5	Tốc độ dọc trục tại cửa ra máy nén	112	m/s
6	Hiệu suất của máy nén	0.855	

3. Xây dựng mô hình mô phỏng

3.1. Xây dựng mô hình hình học cánh công tác máy nén

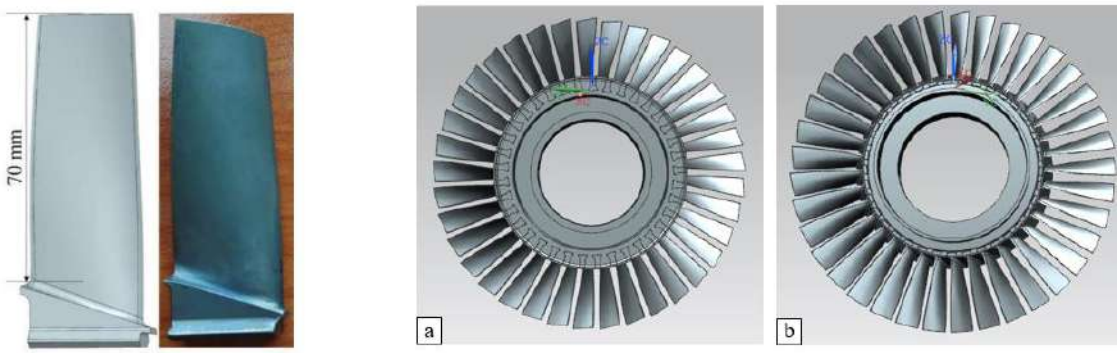
Máy nén động cơ tua bin khí TB3-117 bao gồm 12 tầng, trong đó tầng máy nén đầu tiên có 37 cánh công tác. Để thuận tiện trong tính toán mô phỏng dòng khí chảy qua tầng cánh công tác máy nén, giả thiết đặt ra là dòng khí lưu thông qua chảy đều qua các cánh và vai trò của các cánh trong bánh công tác là như nhau, đồng đều về khối lượng và các thông số hình học khác.

Do đó thay vì mô phỏng dòng khí chảy qua đầy đủ cả tầng cánh công tác với 37 cánh máy nén thì trong nghiên cứu này các tác giả chỉ mô phỏng dòng khí chảy bao quanh 01 cánh máy nén, phần không khí bao gồm phần lưu thông phía mặt bụng và mặt lưng bao quanh cánh.

Cánh công tác máy nén ĐCTBK TV3-117 được quét từ cánh trên động cơ và được xây dựng và chỉnh sửa lại trên phần mềm NX, phần mềm NX bao gồm các chức năng thiết kế từ mô hình bề mặt tự do cơ bản và nâng cao, các chi tiết mô hình dạng khối đặc đến các chức năng thiết kế đa dạng vơi các chi tiết dạng tấm cùng các chức năng thay đổi, hiệu chỉnh, thêm bớt

các bề mặt của vật thể 3D và có chức năng xuất file đuôi ".x_t" thuận tiện cho việc tính toán các chi tiết.

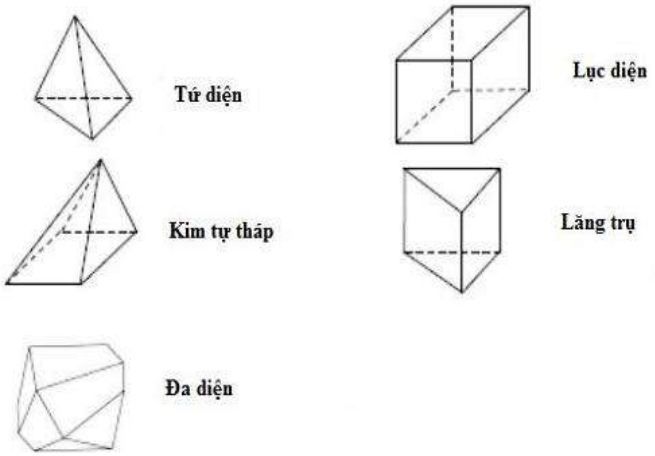
Hình ảnh của cánh máy nén và mô hình bánh công tác tầng đầu tiên sau khi được xây dựng được thể hiện trên hình 1.



Hình 1. Cánh máy nén tầng đầu tiên và mô hình bánh công tác tầng đầu máy nén động cơ TB3-117

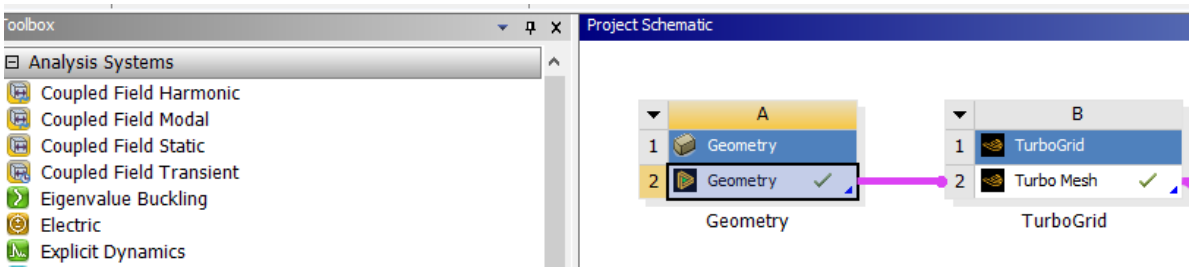
3.2. Chia lưới

Chia lưới thực chất là việc rời rạc hóa vùng không gian mô phỏng thành các phần tử để thực hiện việc tính toán gần đúng bằng phương pháp số. Chia lưới là một trong những bước quan trọng vì kết quả tính toán phụ thuộc nhiều vào việc chia lưới. Một điều quan trọng là tạo ra mô hình và chia lưới đảm bảo máy tính có thể chạy được. Lưới càng dày sẽ cho kết quả càng chính xác, nhưng ngược lại sẽ tốn nhiều thời gian để giải và đòi hỏi máy tính phải có cấu hình cao.



Hình 2. Các dạng phần tử lưới trong mô hình mô phỏng 3D

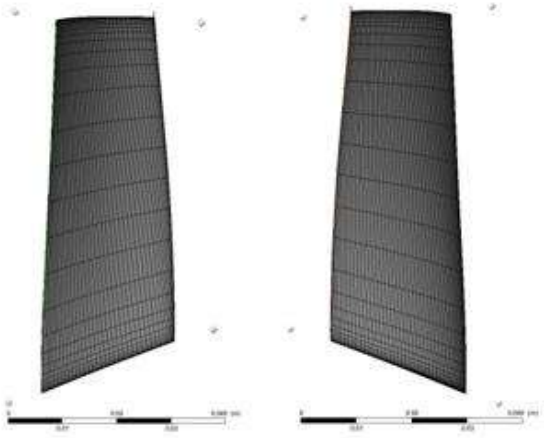
Trong mô hình này lựa chọn dạng lưới lục diện (hexa) vì có ưu điểm so với các dạng lưới còn lại là giảm được số lượng phần tử, số lượng đường và mặt khi chia lưới, các phần tử đi theo một dòng nhất định, chất lượng hình học lưới tốt hơn nhờ đó có thể giảm được thời gian và sai số trong quá trình tính toán. Lưới tính toán được xây dựng bằng phần mềm CFX - TurboGrid, đây là một công cụ hiệu quả cho phép tạo ra lưới hexa chất lượng cao, đặc biệt phù hợp cho các bài toán máy nén và tua bin.



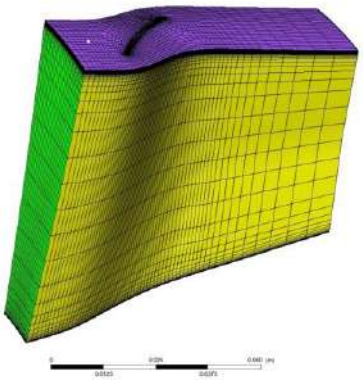
Hình 3. Mô hình chia lưới trong ANSYS- Workbench

Lưới mịn đóng vai trò rất quan trọng trong phân tích động lực học chất lưu tính toán. Do máy nén có dòng chảy bên trong phức tạp, việc lựa chọn kích thước lưới phù hợp và đủ mịn cho phép đạt được sự cân bằng hợp lý giữa độ chính xác của kết quả và thời gian tính toán. Trên cơ sở

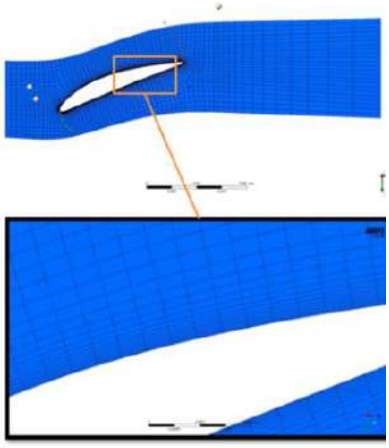
hình học cánh và đường dòng đã được thiết lập trong Geometry, bước tiếp theo thực hiện chia lưới cho miền dòng chảy bằng phần mềm TurboGrid, nhằm tạo lưới chất lượng cao phục vụ mô phỏng CFD.



Hình 4. Mô hình chia lưới cánh máy nén (Phần bụng cánh và lưng cánh)

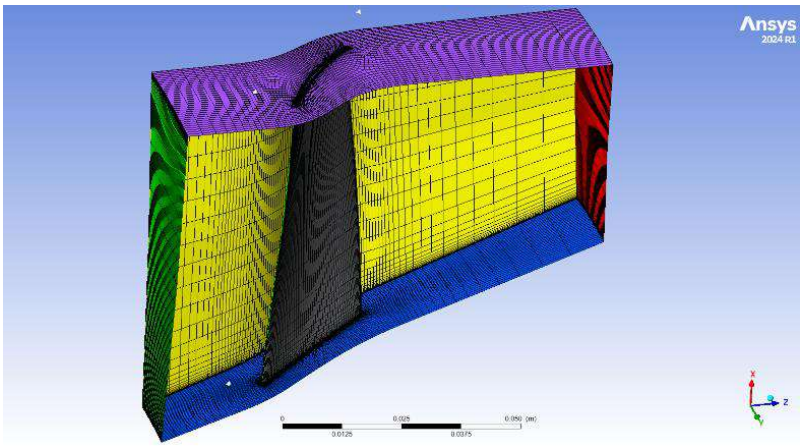


Hình 5. Mô hình lưới tổng thể



Hình 6. Mô hình chia lưới tại gốc lá

cánh công tác



Hình 7. Một kênh dòng đơn (Mặt trên, cánh, mặt dưới)

Thông thường khi chia lưới trong ANSYS -CFX số lượng phần tử và số nút được chia tự động. Việc chia lưới với số lượng phần tử và nút chia càng lớn thì kết quả tính toán càng chính xác. Để đảm bảo kết quả tính toán chính xác hơn, ở mô hình tính toán trong luận văn số lượng phần tử chia lưới và số nút chia tăng lên, cụ thể trong mô hình này: Số phần tử là 324114 và số nút chia là 346552. Ở những vùng có các cánh hoặc các vùng tiếp giáp các phần tử mật độ lưới dày hơn và chế độ tự động chia lưới sẽ khớp được các mắt lưới của các mặt tiếp giáp giữa các phần tử.

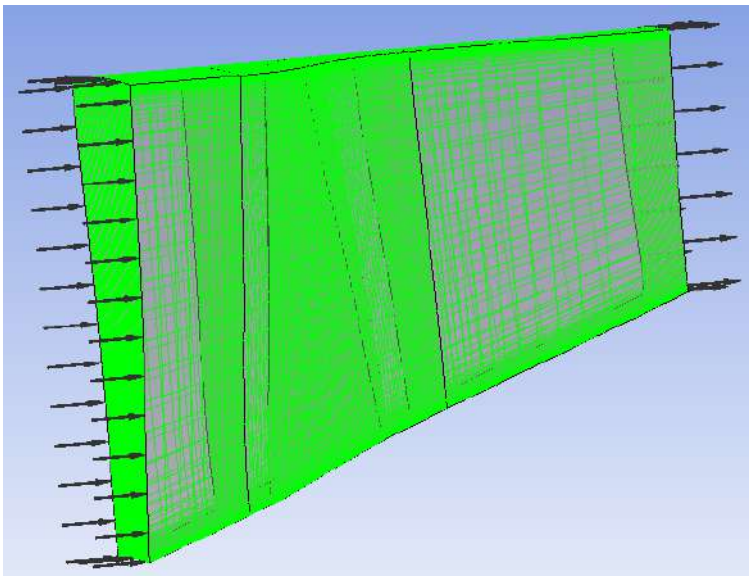
3.3. Xây dựng mô hình mô phỏng

Phần mềm giải ANSYS CFX được sử dụng làm nền tảng mô phỏng nhằm thu được nghiệm động lực học chất lưu tính toán.

Các điều kiện biên được thiết lập trong mô đun CFX-Pre nhằm mô phỏng đúng điều kiện làm việc thực tế của tầng cánh công tác. Thiết lập điều kiện biên cho tính toán mô phỏng, giả thiết điều kiện môi trường khảo sát là điều kiện tiêu chuẩn với áp suất đặc trưng tham chiếu 101325 Pa, nhiệt độ toàn phần tại đầu vào 288K, động cơ hoạt động trên mặt đất H_b = 0 m, M_b = 0, vận tốc quay của máy nén tương ứng 18500 vòng/phút, môi chất công tác là khí thực đối với máy nén, điều kiện dòng chảy bảo toàn năng lượng có thiết lập mặt biên chuyển tiếp để đảm bảo mô phỏng 01 cánh thay cho toàn bộ phần lưu thông của bánh công tác, chiều quay của rotor đối với các mô hình cả máy nén và tua bin được đặt ngược chiều kim đồng hồ nhìn theo hướng bay, lưu lượng không khí G_K = 9.6 kg/s tại đầu ra của tầng máy nén [10, 11].

Mô hình rối SST k-ε (shear stress transports) [12]. Điều kiện giải được đặt với bước thời gian tự động; điều kián hội tụ 1.E-5 hoặc trung bình bình phương RMS = 1% (root mean square) hoặc sai số lưu lượng không khí 1%, tùy thuộc vào điều kiện nào tới trước [12]. Tất cả các bề mặt thành được mô hình hóa là thành cách nhiệt và thỏa mãn điều kiện không trượt.

Để khảo sát ảnh hưởng của góc tấn dòng khí đầu vào tới các thông số và cấu trúc dòng chảy bao quanh cánh máy nén ĐCTBK, lựa chọn khảo sát góc dòng khí đầu vào từ (-15^0) đến 15^0. Bước đầu tiên là tính toán mô phỏng dòng chảy qua tầng cánh công tác đầu tiên của máy nén khi góc dòng khí đầu vào 0^0, sau đó tiến hành thay đổi các góc trong khoảng khảo sát.



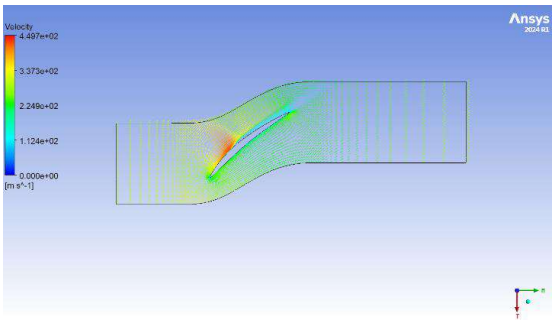
Hình 8. Điều kiện biên mô phỏng dòng khí chảy qua cánh công tác máy nén

4. Kết quả và thảo luận

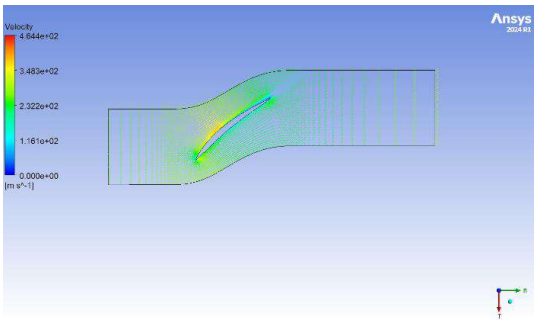
4.1. Kết quả

Trong phần này, ảnh hưởng của góc tấn dòng khí đầu vào đến tải cánh được trình bày thông qua các đồ thị tương ứng trong hình 11. Đồng thời, các trường phân bố áp suất, vận tốc với các góc tấn dòng khí khác nhau cũng được trình bày.

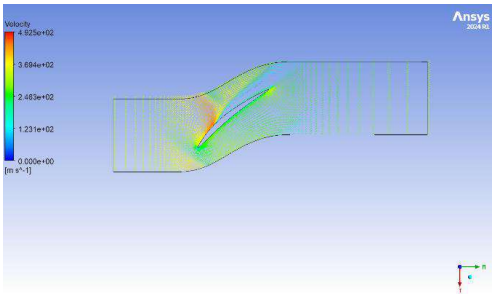
Trên hình 9 thể hiện phân bố vận tốc dòng chảy bao quanh cánh máy nén ở giữa cánh ở các trường hợp góc dòng khí đầu vào 0^0, 15^0, và -15^0.



a. Góc dòng khí đầu vào 0^o



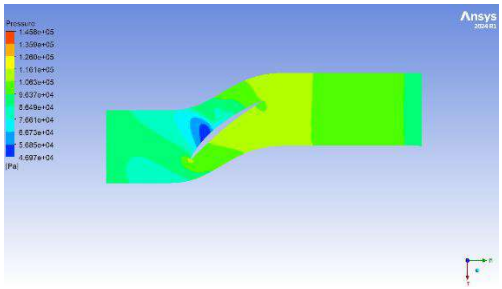
b. Góc dòng khí đầu vào 15^o



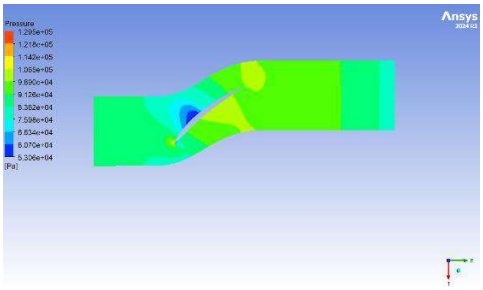
c. Góc dòng khí đầu vào -15^o

Hình 9. Phân bố vận tốc dòng chảy quanh cánh máy nén

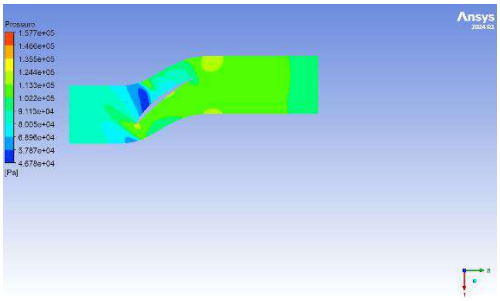
Phân bố áp suất trên bề mặt cánh và trong phần lưu thông thể hiện trên hình 10.



a. Góc dòng khí đầu vào 0^o



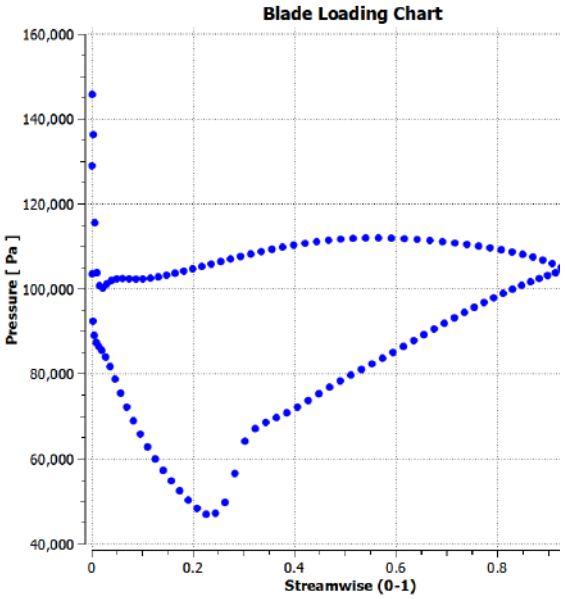
b. Góc dòng khí đầu vào 15^o



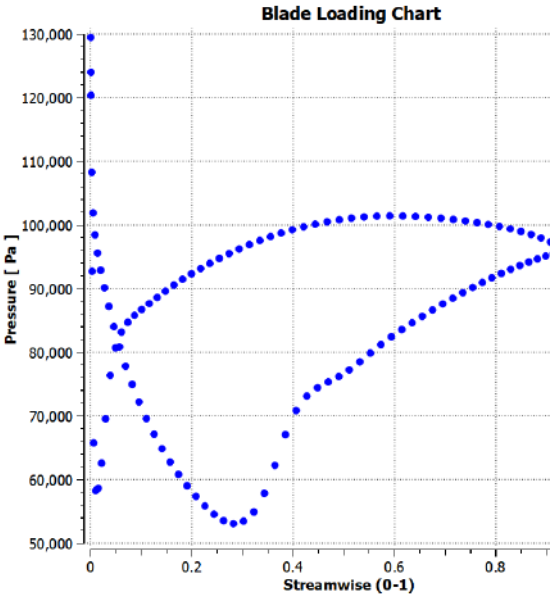
c. Góc dòng khí đầu vào -15^o

Hình 10. Phân bố áp suất dòng khí tại mặt cắt trung bình của cánh

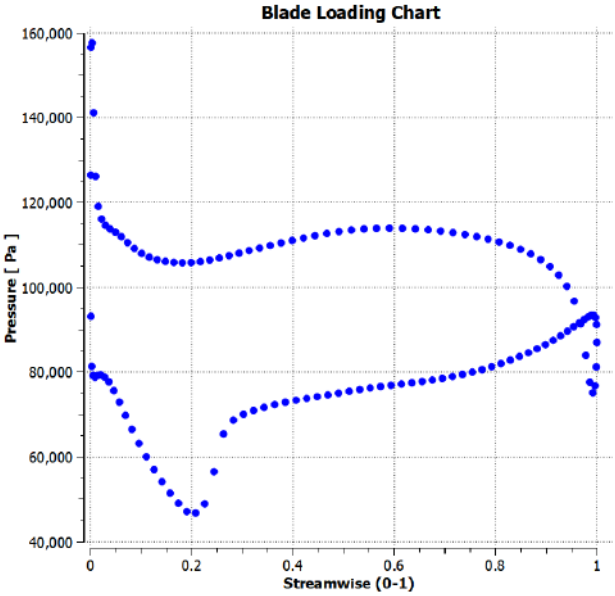
Tải cánh tại 50% chiều cao cánh được thể hiện trên hình 11.



a. Góc dòng khí đầu vào 0^o



b. Góc dòng khí đầu vào 15^o



b. Góc dòng khí đầu vào -15^o

Hình 11. Phân bố tải cánh tại 50% chiều cao cánh

Trên hình 9a thể hiện dòng chảy của không khí trong máy nén chảy bao quanh biên dạng cánh ở vị trí giữa cánh tương ứng với góc dòng khí đầu vào 0^0 (điều kiện dòng chảy đều). Ta thấy dòng chảy bao quanh cánh đều, không bị vỡ dòng, véc tơ vận tốc thể hiện dòng chảy ở phần bụng cánh có phần bị chậm lại, ở đây mật độ dòng không khí tăng, dẫn tới áp suất của khí tại bụng cánh lớn. Trong khi đó dòng không khí tăng tốc ở phần lưng cánh, mật độ không khí giảm, làm áp suất khí tại lưng cánh thấp hơn bụng cánh, như thể hiện ở hình 10a. Điều này thể hiện việc đặt điều kiện biên và điều kiện tính toán ban đầu hợp lý, đây là cơ sở đánh giá ảnh hưởng của 2 trường hợp góc dòng khí đầu vào còn lại tới dòng chảy trong tầng cánh công tác đầu tiên máy nén ĐCTBK TB3-117.

4.2. Thảo luận

Trong nghiên cứu này, góc tấn của cánh công tác (α) được xác định theo quan hệ khí động học: $\alpha = \beta_m - \beta_1$

Trong đó:

- β_m: góc kết cấu đầu vào của cánh;

- β_1: góc dòng khí đầu vào, được thay đổi thông qua điều kiện biên trong mô phỏng CFD, việc thay đổi góc dòng khí đầu vào, dẫn đến góc tấn thay đổi theo.

Từ kết quả mô phỏng được thể hiện qua các hình 9, 10, 11 cho thấy góc dòng khí đầu vào ảnh hưởng trực tiếp đến góc tấn và đặc tính khí động của tầng cánh công tác.

Tại góc dòng khí đầu vào 0°, tầng cánh làm việc gần điều kiện thiết kế. Trường vận tốc phân bố đều, dòng chảy bám tốt trên bề mặt cánh. Trường áp suất thể hiện chênh lệch áp suất hợp lý giữa mặt áp và mặt hút, trong khi tải cánh phân bố mượt, không xuất hiện tải đột ngột. Đây là trạng thái cho hiệu suất đoạn nhiệt cao nhất.

Khi góc dòng khí đầu vào 15°, góc tấn giảm khiến cánh làm việc chế độ tải khí động thấp, gia tốc dòng chảy trên mặt hút suy giảm, vùng chảy rối phía sau cánh kéo dài và phân bố vận

tốc kém đồng đều. Chênh lệch áp suất và tải cánh đều giảm, dẫn đến khả năng sinh công và hiệu suất đoạn nhiệt giảm so với trường hợp chuẩn.

Ngược lại, tại góc dòng khí đầu vào (−15°), góc tấn tăng làm cánh làm việc ở trạng thái quá tải khí động. Trường vận tốc cho thấy sự gia tốc mạnh trên mặt hút kèm theo vùng chảy rối dày phía sau cánh. Trường áp suất xuất hiện vùng áp suất thấp sâu gần mép trước, phân bố tải cánh tăng mạnh ở vùng đầu dây cung, làm gia tăng tổn thất và gây suy giảm hiệu suất đoạn nhiệt.

So sánh ba trường hợp cho thấy cả hai điều kiện lệch góc dòng khí đầu vào so với 0° đều làm hiệu suất giảm, trong đó trường hợp góc vào âm gây suy giảm mạnh hơn do hiện tượng quá tải và tách dòng cục bộ.

5. Kết luận

Từ các kết quả đạt được, nghiên cứu đã chỉ ra góc dòng khí đầu vào có ảnh hưởng đáng kể đến hiệu suất đoạn nhiệt của tầng cánh công tác thông qua sự thay đổi góc tấn.

Góc đầu vào 0° cho điều kiện làm việc tối ưu với dòng chảy ổn định và hiệu suất cao nhất, góc đầu vào 15° làm giảm góc tấn, khiến cánh làm việc chế độ tải khí động thấp và giảm hiệu suất, trong khi góc đầu vào (−15°) làm tăng góc tấn, dẫn đến quá tải khí động, tổn thất lớn và hiệu suất giảm rõ rệt.

Tuy nhiên bài báo còn chưa nghiên cứu đầy đủ cả tầng máy nén, đây cũng là hướng nghiên cứu tiếp theo của nhóm tác giả. Bên cạnh đó việc mở rộng dải góc dòng khí đầu vào với nhiều giá trị hơn để xác định giới hạn làm việc của tầng cánh, ngoài ra việc so sánh với dữ liệu thực nghiệm sẽ giúp đánh giá và nâng cao độ tin cậy của mô hình mô phỏng.

Tài liệu tham khảo

1. F. O. Carta, *Unsteady normal force on an airfoil in a periodically stalled inlet flow*, Journal of Aircraft, Vol. 4, No. 5, p.416, 1967.
2. R. Katz, *Performance of axial compressors with asymmetric inlet flows*, Guggenheim Jet Propulsion Center, California Institute of Technology, Report, 1958.
3. F. Roberts, G. A. Plourde, F. Smakula, *Insights into axial compressor response to distortion*, AIAA Paper No. 68-565, 1969.
4. R. C. Turner, J. Ritchie, C. E. Moss, *The effect of inlet circumferential maldistribution on an axial compressor stage*, ARC R&M No. 3066, 1957.
5. H. Mokelke, *The Unsteady Response of an Axial Flow Compressor with a Distorted Inlet Flow*, C.P. No. 1203, Aeronautical Research Council, Ministry of Defence, UK.
6. U. Farooq, J. Masud, S. Ahmed, K. Parvez, J. A. Khan, A. Mushtaq, *Assessment of Commercial CFD Codes for Performance Prediction of Transonic Axial Compressor* Proceedings of ASME Turbo Expo 2011, GT2011-45006, Vancouver, Canada.
7. A. Mushtaq, K. Parvez, S. Ahmed and J. A. Khan, *Parametric study of tip injection on the stability of transonic axial flow compressor*, 49th American Institute of Aeronautics Astronautics.
8. Asad Islam, Khalid Parvez, *Effects of Inlet Distorted Flows on the Performance of an Axial Compressor*, World Academy of Science, Engineering and Technology International Journal of Mechanical and Mechatronics EngineeringVol: 10, No: 3, 2016.
9. Zheng, X., & Yang, H. (2016). *Influence of Tip Clearance on the Performance and Matching of Multistage Axial Compressors*. Volume 2A: Turbomachinery. https://doi:10.1115/gt2016-56232.
10. Богданов А. Д., Калинин Н. П., Кривко А. И. (2000). *Турбовальный двигатель ТВ3-117ВМ. Конструкция и техническая эксплуатация*, Москва, изд. Воздушный транспорт. 392 стр.
11. Klimov. (1986). *TV3-117 Aircraft Technical Manual*. Klimov Ins. RU.
12. Ansys. (2023). *Ansys CFX User's Guide*. Ansys Inc. USA.

Investigation of the influence of inlet flow angle of attack on the performance parameters of the first compressor stage of the TB3-117 engine.

Abstract: *The TB3-117 gas turbine engine is widely employed on Mi-series helicopters such as Mi-17, Mi-171, and Mi-172 of the Vietnam People's Air Force due to its compact configuration and high power output capability. The compressor is one of the most critical components of the engine, whose primary function is to supply a sufficient amount of air required for engine operation. The characteristics of the inlet airflow have a direct impact on the parameters of the first compressor stage, including stage efficiency, pressure ratio, and airflow rate through the compressor. Consequently, these effects influence the operating parameters of the subsequent compressor stages, the overall compressor, and the gas turbine*

engine. In this paper, the authors present numerical results on the influence of the inlet flow incidence angle on the efficiency of the rotor blade row, as well as on the pressure and velocity fields, using Computational Fluid Dynamics (CFD) simulations performed with ANSYS CFX software. The results of this study provide a basis for further investigations aimed at improving the operational efficiency and reliability of the TB3-117 engine compressor.

Keywords*:* Compressor; Angle of attack; blade; pressure; TB3-117; CFX

Aeroacoustic Characteristic of an Electric Propulsion Fan with Sound-Absorbing Duct

CHU Hoang-Quan[1*], NGUYEN Binh-Nguyen[2], VU Viet-Dung[3], NGUYEN Trung-Kien[1], and DINH Cong-Truong[2]

[1]*Le Quy Don Technical University;* [2] *Hanoi University of Science and Technology;*
[3]*Institute of Development and Application for Sound Materials*

Email: chu.h-quan.fas@lqdtu.edu.vn; Tel: 0985885958

Abstract

Acoustic liners have long been utilized in aero-engines to mitigate aeroacoustic noise based on the Helmholtz resonance principle. However, conducting experimental measurements of sound attenuation on full-scale engines requires sophisticated facilities and substantial investments in time and capital. Furthermore, numerical simulations necessitate solving complex two-way Fluid-Structure Interaction (FSI) problems for the acoustic liner problem. To address these challenges, this paper proposes an equivalent impedance model based on acoustic theories incorporating grazing flow effects to provide boundary conditions for aeroacoustic simulations in ANSYS Fluent. Environmental parameters, operating conditions, and the geometric constraints of an electric Ducted Propulsion Fan (eDPF), including rotational speed and the maximum allowable liner dimensions, are defined. Practical manufacturing constraints for aeronautical acoustic materials are integrated into the MATLAB-based model. The study employs a brute-force algorithm to optimize design parameters, perform a sensitivity analysis, and predict the transmission loss. The results identify an optimal configuration to be implemented in the subsequent aeroacoustic numerical simulation phase.

Keywords: *electric Ducted Propulsion Fan; acoustic liner; equivalent impedance model; grazing airflow; absorption coefficient.*

1. Introduction

In practice, acoustic liners have long been installed on the nacelle inner surface and the bypass duct outer surface to function as sound-absorbing materials for turbofan engines, as shown in Fig. 1, based on the Helmholtz resonance principle to dissipate incident acoustic energy [1, 2].

At high rotational speeds, the grazing flow over the acoustic treatment surface interacts with sound waves and triggers vortex shedding, significantly affecting the impedance characteristics and sound absorption coefficients [3-5]. While Rice [3] established a theoretical framework by incorporating steady grazing flow effects into the impedance model, Guess [4] developed a rigorous mathematical procedure to calculate liner geometric parameters, such as porosity, thickness, and hole diameter, based on target impedance requirements. Furthermore, Heidelberg et al. [5] transitioned from theoretical modeling to experimental validation using an actual turbofan engine (YF-102) to determine the optimum wall impedance.

In the last two decades, research on grazing flow [6-11] has primarily focused on establishing boundary conditions and performing fundamental experimental measurements to elucidate impedance variations. Subsequently, scholars have emphasized specialized conditions, including shear grazing flow [12, 13], perforated plates [14], wavenumber-based

* Email: chu.h-quan.fas@lqdtu.edu.vn

eduction [15], nonlinear impedance [16], and meta-porous liner [17] to validate numerical simulation models. It is evident that these studies concentrate on developing or refining impedance models for sandwich or Single Degree of Freedom (SDOF) structures. These structures typically comprise a perforated plate or micro-perforated panel (MPP) face sheet, an air cavity or honeycomb core (HCC), and a rigid backing.

In this study, an equivalent impedance model was developed for the acoustic liner (Fig. 2), installed within the casing of an electric Ducted Propulsion Fan (eDPF) [18, 19]. Several geometric parameters, operating conditions, and environmental conditions were employed to conduct parametric studies, evaluate parameter sensitivities, and predict the sound attenuation level through the duct. Based on these analyses, the best equivalent impedance configuration was selected to extract the acoustic material data for implementation as boundary conditions in numerical simulations using the commercial code ANSYS Fluent.

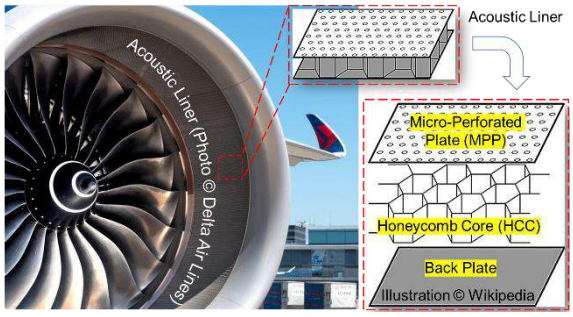


Fig 1. Acoustic liner in practice

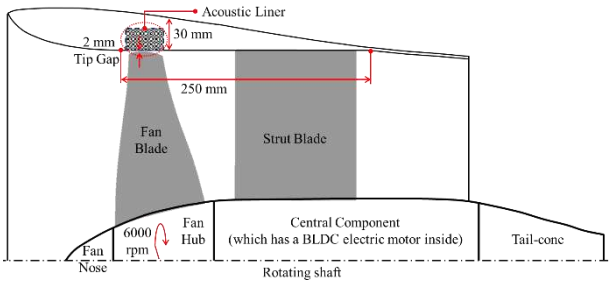


Fig 2. Illustration of eDPF model with acoustic liner

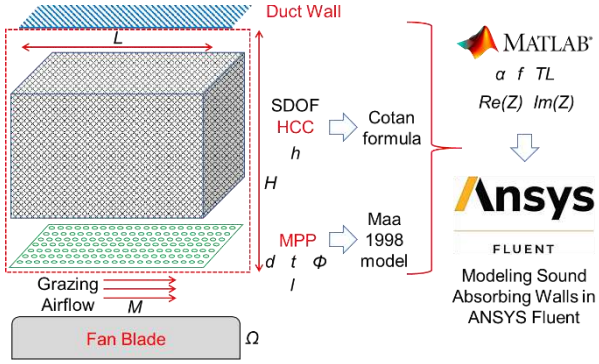


Fig 3. Schematic diagram

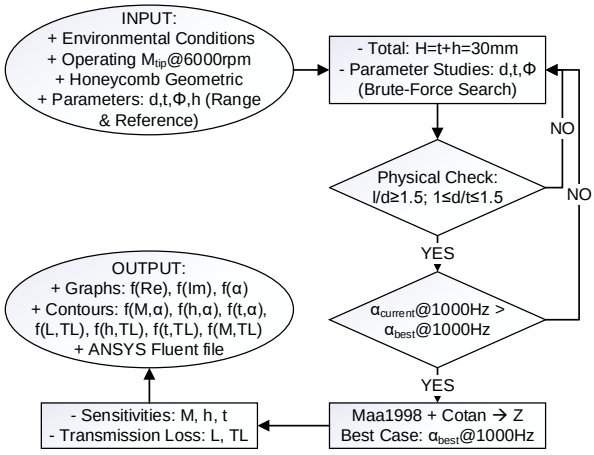


Fig 4. Algorithm flowchart

2. Equivalent Impedance Model

Figure 1 shows the typical structure of an acoustic liner installed on the nacelle inner surface, usually configured as a SDOF system. In this outline, the top perforated layer is a MPP, which is bonded to a HCC layer and sealed by an impervious layer (referred to as back plate). The eDPF operates steadily at 6000 rpm in the counterclockwise direction, where the maximum allowable liner thickness and axial length are 30 mm and 250 mm, separately (Fig. 2).

Since the sound-absorbing material problem in aircraft engines requires evaluation of the acoustic material characteristics, i.e., ANSYS Mechanical, followed by application to the aerodynamic problem in regions with significant grazing flow, i.e., ANSYS Fluent. Thus, the

fundamental nature of this problem corresponds to a two-way Fluid–Solid Interaction (FSI), which is relatively complex. On this basis, Figure 3 presents the schematic diagram of the equivalent impedance model, in which the MPP layer is modeled according to the Maa (1998) approach [20], and the SDOF HCC layer is formulated using the cotangent expression [20] and implemented in Matlab version 2024 R2 [21].

Table 1. Values of survey parameters in research

#	d [mm]	t [mm]	Φ [%]	h [mm]	Ω [rpm]	M	L [mm]
Reference	0.8	0.8	2.0	29.2	6000	0.353	100
Max	1.2	1.2	3	29.4	10000	0.589	250
Min	0.6	0.6	1	28.8	2000	0.118	0
Step	0.2	0.2	0.5	0.2	-	-	-
Purpose	Parameter studies				Parameter sensitivities		Predicted transmission loss

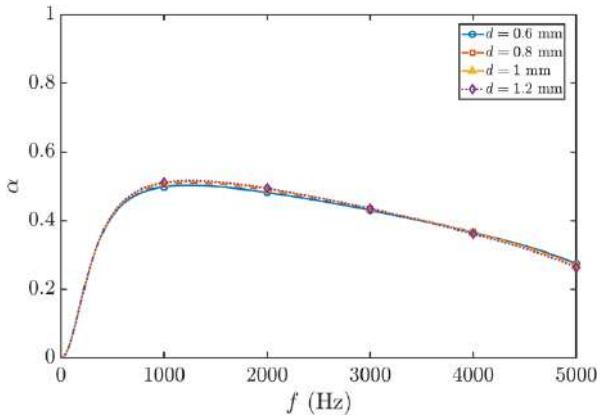


Fig 5. Variation of MPP hole diameter

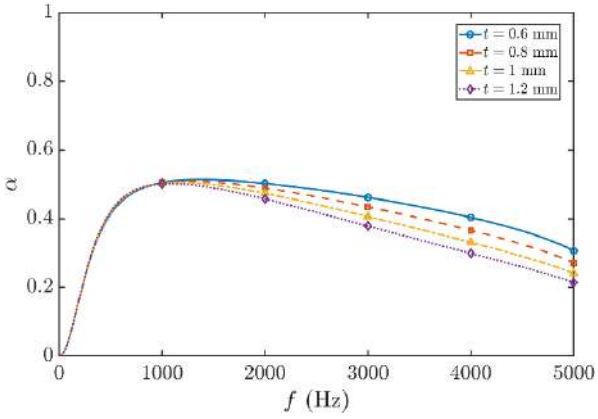


Fig 6. Variation of MPP thickness

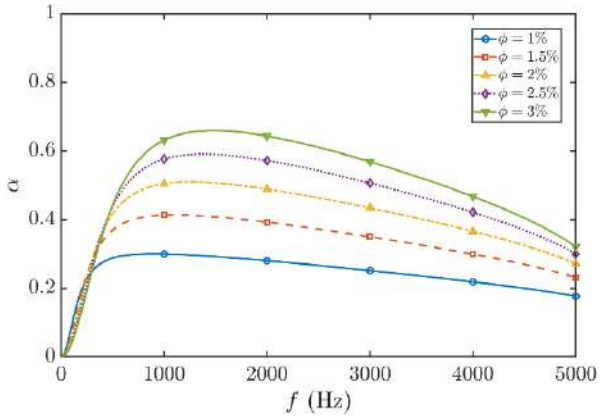


Fig 7. Variation of MPP porosity

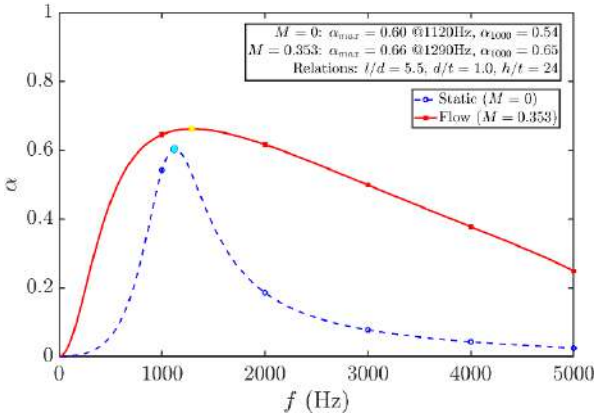


Fig 8. Absorption coefficient in the best case

The MPP design is characterized by thickness t, hole diameter d, and porosity Φ, while the HCC geometry is defined by depth h. The total allowable thickness H, equal to 30 mm, is the sum of t and h. The output parameters evaluated using the Matlab code include the absorption coefficient α, frequency f, and the real part Re(Z) and imaginary part Im(Z) of the impedance. When considering the liner axial length L in the equivalent acoustic liner model, the Transmission Loss (denoted TL) is predicted.

Figure 4 presents the detailed algorithm flowchart employed in this study. The input parameters of interest include the environmental conditions in the Hanoi region (considering dry air properties at 20°C at ground level) and the operating condition of the eDPF at a rotational speed Ω of 6000 rpm (corresponding to a tip Mach number of 0.353). The fundamental honeycomb geometry adopts recommended parameters from the literature [22], namely a cell size of 3/8 inches and a wall thickness of 0.002 inches. Four input parameters, including d, t, Φ, and h, are defined according to Table 1 for parametric investigation using the Brute-Force algorithm [23]. The investigated ranges of d, t, and Φ, together with the constraints: l/d≥1.5 and 1≤t/d≤1.5, where l denotes the average spacing between two adjacent MPP holes, are referenced from practical machining limitations or aerospace HCC manufacturing standards.

According to [24], turbofan engines generate the highest tonal noise at the first Blade Passing Frequency (BPF). For the eDPF configuration, this frequency is 1000 Hz. So, to achieve the most effective acoustic attenuation, the absorption coefficient α at 1000 Hz should be maximized. This criterion serves as the primary condition for selecting the optimal configuration for sensitivity analysis and transmission loss prediction, and next for exporting the equivalent impedance data to be applied as boundary conditions in ANSYS Fluent 2025 R1.

MPP normalized surface impedance of the under no-flow conditions [20]:

$$z_{static} = \frac{Z}{\rho_0 c_0} = r_{static} + j x_{static} = \frac{32\eta t}{\rho_0 c_0 d^2}\left(1+\frac{k^2}{32}\right)^{1/2} + j\frac{\omega t}{c_0}\left[1+\left(9+\frac{k^2}{2}\right)^{-1/2}\right] \quad (1)$$

where perforate constant $k = d\sqrt{\omega\rho_0/4\eta}$, ω and η are denoted separately the angular frequency and the coefficient of viscosity.

Normalized acoustic resistance and reactance under the influence of grazing flow [3, 25]: $r_{grazing} = 0.3\cdot M/\phi$ và $x_{grazing} = 0.85\cdot d/\left(1+0.3\cdot M/\phi\right)$.

Resultant equivalent impedance of the MPP under flow conditions:

$$Z_{MPP} = \left[\left(r_{static} + r_{grazing}\right) + j\cdot\left(x_{static} + x_{grazing}\right)\right]\cdot\rho_0 c_0 \quad (2)$$

The honeycomb core functions as a backing cavity, with its acoustic impedance modeled by a cotangent-based expression [20]:

$$Z_{core} = -j\frac{\rho_0 c_0}{\sigma_{hc}}\cot\left(\frac{\omega h}{c_0}\right) \quad (3)$$

Total acoustic impedance of the equivalent model: $Z_{total} = Z_{MPP} + Z_{core}$

The sound absorption coefficient is subsequently derived from the pressure reflection coefficient [26] as follows:

$$\alpha = 1 - \left|R_{reflection}\right|^2 = 1 - \left|\frac{Z_{total} - \rho_0 c_0}{Z_{total} + \rho_0 c_0}\right|^2 \quad (4)$$

3. Results and Discussion

3.1. Parameter studies

Figures 5–7 illustrate the impact of various MPP geometric parameters on the sound absorption coefficient α. At a grazing Mach number of 0.353, increasing the hole diameter d from 0.6 mm to 1.2 mm (Fig. 5) results in a minimal shift in the absorption peak (from 0.498 to 0.511), suggesting that the hole diameter has a marginal influence on the liner's performance. In contrast, while varying the thickness t from 0.6 mm to 1.2 mm (Fig. 6) has little effect on the absorption curves below 1000 Hz, substantial deviations are observed in the high-frequency range.

In addition, when the porosity Φ increases from 1% to 3%, a significant variation of α is observed, from approximately 0.30 to about 0.63 (Fig. 7), suggesting that higher porosity enhances the interaction with vortices induced by the grazing flow. Overall, all cases exhibit a peak around 1000 Hz and gradually decrease as the frequency increases from 1000 to 5000 Hz.

The equivalent impedance configuration corresponding to the best case in Figures 8–10, obtained via the Brute-Force algorithm, is d = 1.2 mm, t = 1.2 mm, Φ = 3.0%, and h = 28.8 mm for the eDPF operating at 6000 rpm. Under static conditions (M = 0), the model prediction differs significantly from the case with actual airflow. Exactly, Fig. 8 illustrates that α at 1000 Hz surges from 0.54 under static conditions to 0.65 at a grazing Mach of 0.353, representing an enhancement of approximately 20.37% under grazing flow conditions.

Moreover, due to the grazing flow effect, α exhibits an increase across the entire f range. This results in a broader α spectrum compared to the quiescent case, where the α is highly concentrated around the resonance frequency of 1000 Hz. This phenomenon can be elucidated by Figs. 9 and 10, wherein the normalized real impedance exhibits a significant increase (red line) across the entire frequency spectrum compared to the static flow condition. Conversely, the normalized imaginary impedance remains relatively constant, predominantly converging toward zero at frequencies $f \geq 500$Hz.

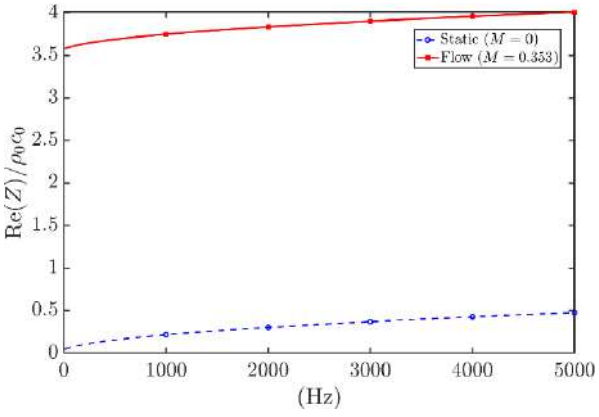


Fig 9. Real impedance in the best-case

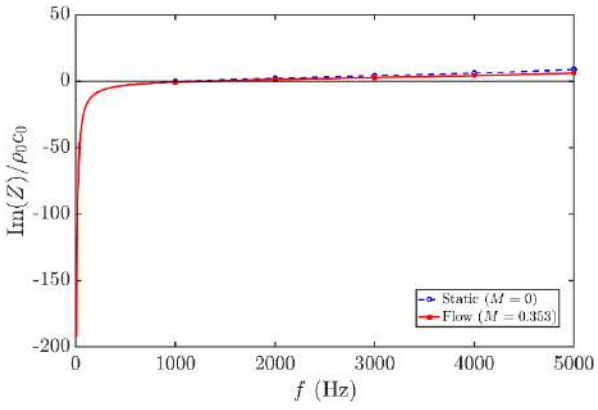


Fig 10. Imaginary impedance in the best-case

3.2. Parameter sensitivities

To provide a more objective assessment of the variation of α with respect to frequency f, parameter sensitivities of M, t, and h were investigated. In addition to the four primary design parameters, Table 1 also presents the range of grazing Mach numbers matching to the practical operating range of the eDPF from 2000 to 10000 rpm, i.e., M. varies from 0.118 to 0.589.

Figure 11 illustrates a prominent dark-red resonance zone, corresponding to an absorption coefficient $\alpha > 0.8$, spanning the frequency range of 1000 Hz to 1500 Hz for grazing $M < 0.2$. The sound absorption efficiency diminishes as f exceeds 2500 Hz or falls below 500 Hz, as well as with a gradual increase in M; notably, α only reaches 0.66 at $M = 0.353$.

Figure 12 indicates that the normalized acoustic resistance increases with the Mach number, as evidenced by the deep-red regions on the right side of the contour plot across the entire frequency spectrum. This result further confirms the significant impact of the grazing flow velocity on the real part of the impedance. However, for $M > 0.3$, corresponding to the green-yellow-red zones, the normalized resistance values considerably exceed unity. This deviation explains the degradation in sound absorption performance as the Mach number augments.

Figure 13 presents a visualization characterized by dominant red shades, signifying that the normalized acoustic reactance (imaginary impedance) remains relatively constant and consistently near the zero-crossing point, irrespective of M across the entire f range. This stability in reactance ensures that the system maintains a steady resonance state, preventing significant f shifts during variable speed operations of the eDPF.

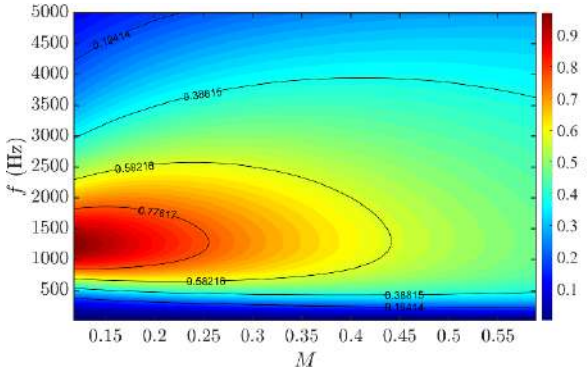


Fig 11. Mach sensitivity with absorption coefficient

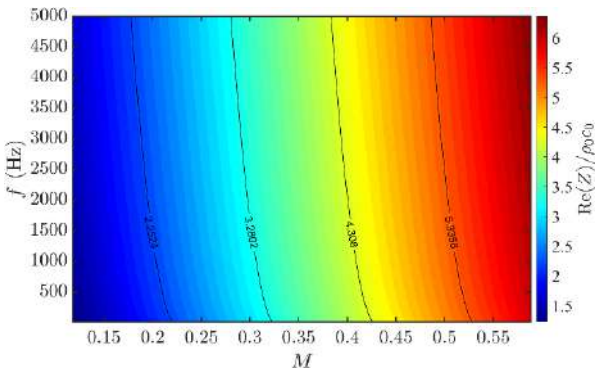


Fig 12. Mach sensitivity with real impedance

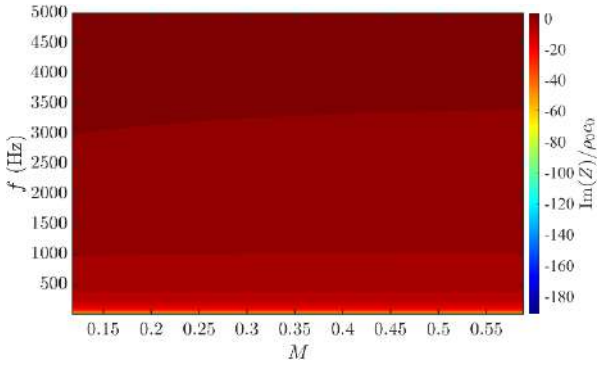


Fig 13. Mach sensitivity with imaginary impedance

Figures 14 and 15 evaluate the sensitivity of the MPP thickness t and HCC depth h, characterized by prominent red resonance bands extending across the entire investigated spectrum. In the frequency range of 2500 Hz to 3500 Hz, the absorption coefficient α consistently surpasses 0.6, regardless of the sensitivity to t or h. However, as t diminishes, the red-colored resonance zone expands significantly. Conversely, an augmentation in h causes the high-absorption region to shift toward lower frequencies. The optimized parameters of t = 1.2 mm and

h = 28.8 mm, as derived from Fig. 8, yield α > 0.6 in the vicinity of 1000 Hz, which is perfectly aligned with the target operating conditions of the eDPF.

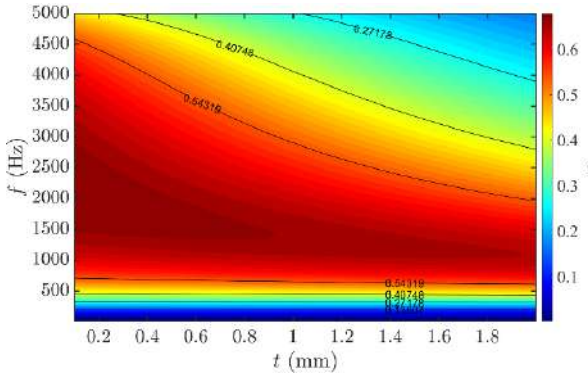


Fig 14. MPP thickness sensitivity

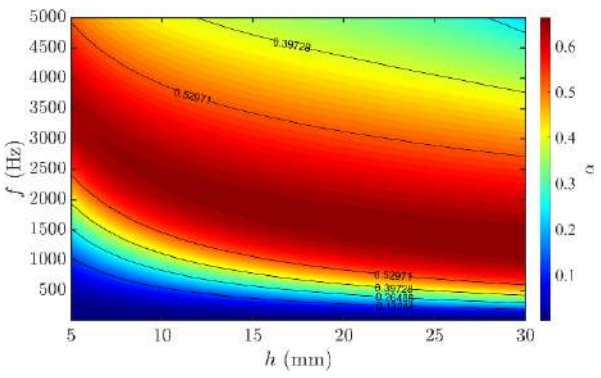


Fig 15. HCC depth sensitivity

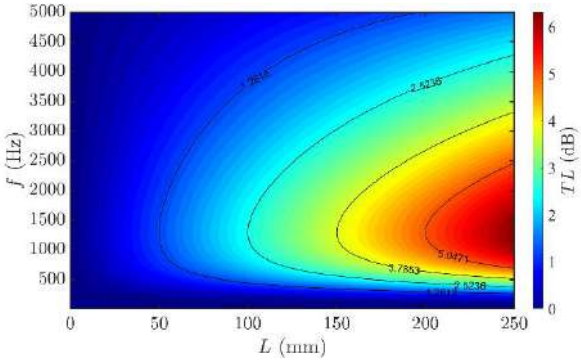


Fig 16. Variation of liner length

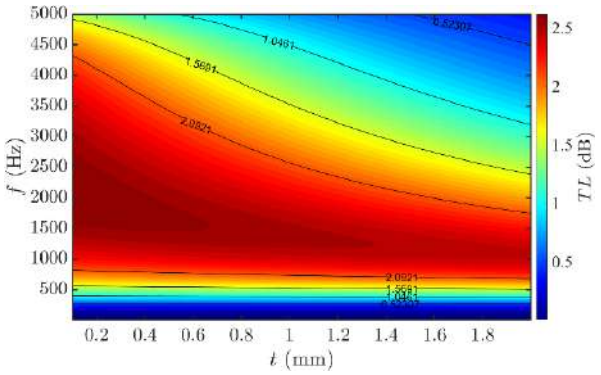


Fig 17. Variation of MPP thickness

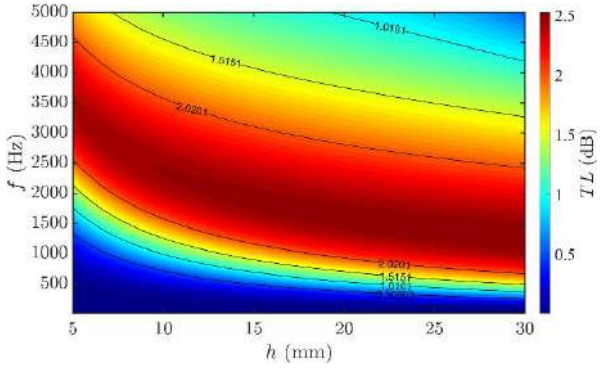


Fig 18. Variation of HCC depth

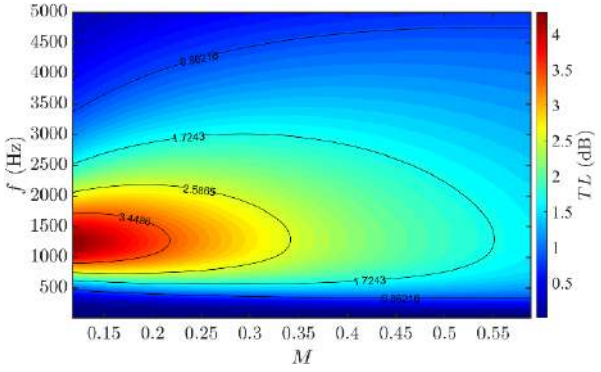


Fig 19. Variation of grazing Mach

3.3. Predicted transmission loss

The prediction of noise attenuation through the acoustic liner inside the duct is calculated using the following formula [17]:

$$TL = 20\log_{10}\left|p_{in}/p_{tr}\right| \tag{5}$$

where p_{in} and p_{tr} denote the incident and transmitted acoustic pressures, respectively, across the acoustic liner under grazing flow.

For acoustic liners with a variable length L up to 250 mm (as detailed in Fig. 2 and Tab. 1), Fig. 16 demonstrates that acoustic efficiency is positively correlated with L. Notably, when $L \geq 200$ mm, the predicted acoustic attenuation reaches approximately 6 dB. A strong consensus

in the distribution of α and TL is observed between the pairs of Fig. 14 and 17, as well as Fig. 15 and 18, indicating that a high absorption coefficient translates to a significant TL enhancement. However, the peak TL, when considering the sensitivity of both t and h, is constrained to only 2.5 dB. In contrast, for the most feasible configuration d = 1.2 mm, t = 1.2 mm, Φ = 3.0%, h = 28.8 mm) with L = 100 mm, the expected noise reduction climbs to nearly 4.0 dB for M < 0.2, as illustrated in Fig. 19. Under the steady-state operating conditions of the eDPF at 6000 rpm (corresponding to M = 0.353), the TL undergoes a reduction, settling at roughly 2.5 dB.

4. Conclusion

Several key conclusions can be drawn from this study:

First, an equivalent impedance model was successfully developed via MATLAB, incorporating acoustic theories that account for grazing flow, environmental conditions, operating regimes, and geometric parameters.

Second, the absorption coefficient α exhibits high sensitivity to variations in the porosity Φ, whereas it remains relatively insensitive to the hole diameter d and plate thickness t. The optimal configuration (d = 1.2 mm, t = 1.2 mm, Φ = 3.0%, h = 28.8 mm) facilitates broadband absorption characteristics induced by the grazing flow. This design is most effective for the eDPF at 6000 rpm, as it centers the acoustic attenuation on the BPF of 1000 Hz.

Third, sensitivity analysis of L, t, and h indicates that acoustic efficiency is superior at M < 0.2. Within this range, the acoustic resistance is bolstered by the grazing flow effect, while the acoustic reactance remains nearly constant and approaches zero. Furthermore, α exceeds 0.6 only when t < 1.0 mm or when h approaches the total depth H (30 mm) within the 1500–2000 Hz frequency band.

Fourth, the liner length L is positively correlated with and significantly influences the overall noise attenuation. A higher α consistently yields an enhanced Transmission Loss (TL), regardless of the variations in t or h. The attenuation performance is further amplified, with the predicted TL reaching 5–6 dB when L ≥ 200 mm.

Fifth, to further validate the efficacy of the equivalent impedance model, future research should implement Impedance Boundary Conditions (IBC) in ANSYS Fluent simulations or conduct experimental testing on a physical eDPF prototype equipped with the acoustic liner.

References

1. Ozkaya, E., Gauger, N. R., Hay, J. A., & Thiele, F. (2020). Efficient design optimization of acoustic liners for engine noise reduction. *AIAA Journal*, 58(3), 1140–1156. https://doi.org/10.2514/1.J057776
2. Zhang, J., Shu, W., Yan, R., Du, L., & Sun, X. (2023). Optimum Acoustic Impedance for a Finite-Length Nacelle Liner of Turbofan Engine. In *AIAA Aviation and Aeronautics Forum and Exposition, AIAA AVIATION Forum 2023*. American Institute of Aeronautics and Astronautics Inc, AIAA. https://doi.org/10.2514/6.2023-3738

3. Rice, E. J. (1972). Model for the Acoustic Impedance of a Perforated Plate Liner with Multiple Frequency Excitation. *The Journal of the Acoustical Society of America*, *51*(1A_Supplement), 94–95. https://doi.org/10.1121/1.1981715
4. Guess, A.W. (1975). Calculation of perforated plate liner parameters from specified acoustic resistance and reactance. *Journal of Sound and Vibration*, *40*(1), 119–137. https://doi.org/10.1016/S0022-460X(75)80234-3
5. Heidelberg, L. J., Rice, E. J., & Homyak, L. (1980). Experimental evaluation of a spinning-mode acoustic-treatment design concept for aircraft inlets. *NASA Technical Paper*, (1613).
6. Aurégan, Y., Starobinski, R., & Pagneux, V. (2001). Influence of grazing flow and dissipation effects on the acoustic boundary conditions at a lined wall. *The Journal of the Acoustical Society of America*, 109(1), 59–64. https://doi.org/10.1121/1.1331678
7. Dickey, N. S., Selamet, A., & Ciray, M. S. (2001). An experimental study of the impedance of perforated plates with grazing flow. *The Journal of the Acoustical Society of America*, *110*(5), 2360–2370. https://doi.org/10.1121/1.1400736
8. Kabral, R., Bodén, H., & Elnady, T. (2014). Determination of liner impedance under high temperature and grazing flow conditions. In *20th AIAA/CEAS Aeroacoustics Conference. American Institute of Aeronautics and Astronautics Inc*. https://doi.org/10.2514/6.2014-2956
9. Tam, C. K. W., Pastouchenko, N. N., Jones, M. G., & Watson, W. R. (2014). Experimental validation of numerical simulations for an acoustic liner in grazing flow: Self-noise and added drag. *Journal of Sound and Vibration*, *333*, 2831–2854. https://doi.org/10.1016/j.jsv.2014.02.019
10. Leon, O., Méry, F., Piot, E., & Conte, C. (2019). Near-wall aerodynamic response of an acoustic liner to harmonic excitation with grazing flow. *Experiments in Fluids*, *60*(9). https://doi.org/10.1007/s00348-019-2791-5
11. Zhang X., Yang C., Cheng L., & Zhang P. (2020). An experimental investigation on the acoustic properties of micro-perforated panels in a grazing flow. *Applied Acoustics*. 159, 107119. https://doi.org/10.1016/j.apacoust.2019.107119.
12. Lafont, V., Méry, F., Roncen, R., Simon, F., & Piot, E. (2020). Liner impedance eduction under shear grazing flow at a high sound pressure level. *AIAA Journal*, *58*(3), 1107–1117. https://doi.org/10.2514/1.J058756
13. Roncen, R., Sebbane D., Mery F., Piot E., & Simon F. (2020). Experimental Investigation of Shear Grazing Flow Effects on Extended Tube Acoustic Liners. *eForum Acusticum*, Dec 2020, Lyon, France. pp.411-415, https://doi.org/10.48465/fa.2020.0225.
14. Chen, Z., Ji, Z., & Huang, H. (2020). Acoustic impedance of perforated plates in the presence of fully developed grazing flow. *Journal of Sound and Vibration*, 485. https://doi.org/10.1016/j.jsv.2020.115547
15. Roncen, R., Piot, E., Méry, F., Simon, F., Jones, M. G., & Nark, D. M. (2020). Wavenumber-based impedance eduction with a shear grazing flow. *AIAA Journal*, *58*(7), 3040–3050. https://doi.org/10.2514/1.J059100
16. Roncen, R. (2025). Revisiting nonlinear impedance in acoustic liners. *Journal of Sound and Vibration*, *608*. https://doi.org/10.1016/j.jsv.2025.119058
17. Wang, S., Xiao, Y., Wang, D., Hu, Y., & Wen, J. (2025). Ultra-low and ultra-broadband meta-porous liner under grazing flow. *International Journal of Mechanical Sciences*, *306*. https://doi.org/10.1016/j.ijmecsci.2025.110849

18. Dinh VS, Chu HQ, Nguyen TK, Le DA, Le NB, Nguyen AT, Kim WY, Dinh CT, 2024, "Flow analysis through center body of ducted fan to electric motor cooling," In: Journal of Physics: Conference Series, Vol. 2707, Institute of Physics. https://doi.org/10.1088/1742-6596/2707/1/012103.
19. Chu HQ, Dinh QN, Vu TS, Pham VY, Bui VT, Hoang NM, Nguyen TK, Nguyen D, Pham GD, Dinh CT, 2025, "Static Aero-Propulsion Experiment of an Electric Ducted Fan," Aerospace, Vol. 12, No. 6, pp. 509. https://doi.org/10.3390/aerospace12060509.
20. Maa, D.Y. (1998). Potential of microperforated panel absorber. *The Journal of the Acoustical Society of America*, *104*(5), 2861–2866. https://doi.org/10.1121/1.423870
21. The MathWorks Inc. (2024). *MATLAB* (Version R2024b). The MathWorks Inc. https://www.mathworks.com
22. Hexcel Corporation (2015). HexWeb Honeycomb Attributes and Properties. https://www.pccomposites.com/wp-content/uploads/2015/07/PCHC4-4TY4_TDS.pdf
23. Praveen Kumar (2024). brute_force_optimization(BFO). *MATLAB Help Center*. https://www.mathworks.com/matlabcentral/fileexchange/174285-brute_force_optimization-bfo
24. Bonomo, Lucas Araujo (2025). Acoustic impedance eduction with a sheared flow boundary layer. *PhD Thesis*. https://repositorio.ufsc.br/handle/123456789/266728.
25. Lahiri, C., & Bake, F. (2017). A review of bias flow liners for acoustic damping in gas turbine combustors. *Journal of Sound and Vibration.* Academic Press. https://doi.org/10.1016/j.jsv.2017.04.005
26. Zhang, P. F., Li, Z. H., Zhou, Y. J., Zhang, Q. F., Liu, B., Liu, F., … Bai, P. kang. (2025). Improved sound absorption with 3D-printed micro-perforated sandwich structures. *Journal of Materials Research and Technology*, *34*, 855–865. https://doi.org/10.1016/j.jmrt.2024.12.082

Mô hình trở kháng tương đương cho quạt đẩy điện có ống hướng dòng tích hợp lớp lót tiêu âm

Tóm tắt: *Các lớp lót tiêu âm từ lâu đã được sử dụng trong động cơ máy bay để giảm thiểu tiếng ồn khí động dựa trên nguyên lý cộng hưởng Helmholtz. Tuy nhiên, việc tiến hành các phép đo thực nghiệm về sự suy giảm âm thanh trên các động cơ kích thước thật đòi hỏi các thiết bị phức tạp và đầu tư đáng kể về thời gian và vốn. Ngoài ra, đối với bài toán lớp lót cách âm, mô phỏng số phải giải quyết các bài toán tương tác chất lưu - kếtcấu (FSI) hai chiều phức tạp. Để giải quyết những thách thức này, bài báo này đề xuất một mô hình trở kháng tương đương dựa trên các lý thuyết âm học kết hợp hiệu ứng dòng chảy lướt để cung cấp các điều kiện biên cho các mô phỏng khí động học trong ANSYS Fluent. Các thông số môi trường, điều kiện vận hành và các ràng buộc hình học của quạt đẩy điện có ống hướng dòng (eDPF), bao gồm tốc độ quay và kích thước lớp lót tối đa cho phép, được xác định. Các điều kiện sản xuất thực tế đối với vật liệu cách âm hàng không cũng được tích hợp vào mô hình viết trên code MATLAB. Nghiên cứu sử dụng thuật toán Brute-force để tối ưu hóa các thông số thiết kế, thực hiện phân tích độ nhạy và dự đoán tổn thất truyền âm. Kết quả nghiên cứu đưa ra một cấu hình tốt nhất của mô hình trở kháng tương đương để thực hiện trong giai đoạn mô phỏng số khí động học tiếp theo.*

Từ khóa: quạt đẩy điện có ống dẫn dòng; lớp lót cách âm; mô hình trở kháng tương đương; dòng không khí chảy lướt song song bề mặt; hệ số hấp thụ.

DỰ ĐOÁN ĐẶC TÍNH ĐỘNG CƠ TUA-BIN KHÍ CỠ NHỎ BẰNG PHƯƠNG PHÁP ĐỒNG DẠNG HÓA BẢN ĐỒ

Nguyễn Hữu Hà[1,*], Nguyễn Quốc Quân[1], Phạm Xuân Phương[1]

[1]Viện cơ khí động lực, Học viện kỹ thuật quân sự

Email: huuha84@gmail.com

Tóm tắt

Việc xác định đặc tính làm việc của động cơ tua-bin khí cỡ nhỏ gặp nhiều khó khăn do thiếu bản đồ thành phần và không có bệ thử thực nghiệm. Nghiên cứu này đề xuất một quy trình mô phỏng nhằm xác định đặc tính vận hành toàn động cơ JetCat P130-RX trong điều kiện hạn chế dữ liệu. Mô hình chu trình nhiệt động 0D được xây dựng trong GasTurb, kết hợp kỹ thuật đồng dạng hóa bản đồ để hiệu chỉnh máy nén và tua-bin sao cho khớp với các thông số công bố tại chế độ Idle và Max. Để đánh giá độ tin cậy, phương pháp được kiểm chứng gián tiếp trên bộ số liệu của động cơ BMT120KS, cho sai số lực đẩy và lưu lượng nhiên liệu dưới 12%. Sau khi hiệu chỉnh cho JetCat P130-RX,các đặc tính theo dải tốc độ được xác định, đồng thời trích xuất các thông số nội tại phục vụ phân tích buồng cháy. Phương pháp đề xuất cho phép tái tạo đặc tính động cơ trong điều kiện không có bệ thử với độ chính xác chấp nhận được.

Từ khóa: *JetCat P130-RX, đồng dạng hóa bản đồ, mô phỏng số, ống hóa hơi*

1. Đặt vấn đề

Trong những năm gần đây, động cơ tua-bin khí (TBK) cỡ nhỏ ngày càng được ứng dụng rộng rãi trong các hệ thống thiết bị bay không người lái (UAV), phương tiện bay cá nhân và các nguồn phát điện quy mô nhỏ [1]. Nhờ ưu điểm cấu trúc đơn giản, tỷ số công suất trên khối lượng cao và khả năng vận hành ở tốc độ quay lớn, các động cơ phản lực cỡ nhỏ trở thành lựa chọn phù hợp cho các nền tảng yêu cầu lực đẩy tức thời và kích thước gọn nhẹ [2].

Về nguyên lý, động cơ TBK cỡ nhỏ hoạt động tương tự các động cơ TBK quy mô lớn, bao gồm quá trình nén, đốt cháy và giãn nở trong chu trình Brayton. Tuy nhiên, do kích thước nhỏ và tốc độ quay rất cao, động cơ TBK cỡ nhỏ vận hành trong miền số Reynolds thấp, chịu ảnh hưởng mạnh của tổn thất khe hở đầu cánh, ma sát thành và hiệu ứng quy mô [3]. Điều này làm cho việc dự đoán đặc tính hiệu suất ngoài thiết kế trở nên phức tạp hơn so với các động cơ công suất lớn.

Phương pháp trực tiếp và chính xác nhất để xác định đặc tính động cơ là thông qua bệ thử thực nghiệm [4]. Tuy nhiên, đối với các động cơ TBK cỡ nhỏ thương mại, việc lắp đặt hệ thống đo nội tại (áp suất, lưu lượng, nhiệt độ trong buồng cháy) gặp nhiều hạn chế do không gian nhỏ, chi phí vận hành cao và rủi ro hư hỏng ở tốc độ quay lớn. Vì vậy, dữ liệu công bố thường chỉ bao gồm lực đẩy và nhiệt độ khí xả tại một số chế độ vận hành.

Bên cạnh thực nghiệm, mô phỏng CFD ba chiều cho phép phân tích chi tiết trường dòng chảy trong từng thành phần, nhưng yêu cầu thông tin hình học đầy đủ và chi phí tính toán lớn [5]. Do đó, mô phỏng chu trình nhiệt động 0D sử dụng bản đồ đặc tính thành phần là giải pháp cân bằng giữa độ chính xác và chi phí tính toán trong phân tích ngoài thiết kế [6].

Vấn đề đặt ra là các bản đồ đặc tính của động cơ thương mại hiếm khi được công bố. Do đó, kỹ thuật đồng dạng hóa bản đồ từ một động cơ tham chiếu trở thành phương pháp khả thi để xây dựng mô hình cho động cơ mục tiêu.

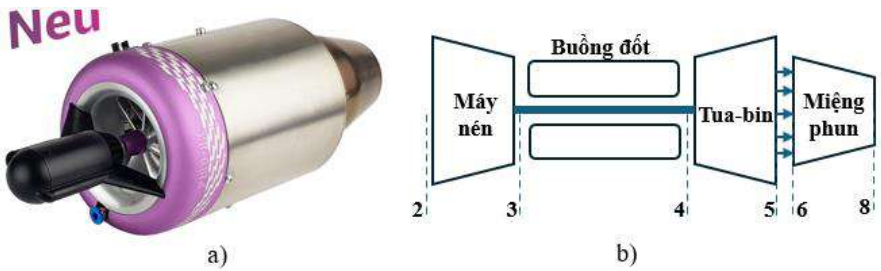
Trong nghiên cứu này, một mô hình chu trình nhiệt động được phát triển nhằm xác định đường làm việc của động cơ JetCat P130-RX mà không sử dụng bệ thử thực nghiệm. Phương pháp đồng dạng hóa bản đồ được kiểm chứng thông qua bộ tham số được công bố của động cơ

BMT120KS trước khi áp dụng cho JetCat P130-RX. Mục tiêu chính là xây dựng đường đặc tính vận hành toàn dải và trích xuất các thông số khí động cửa vào ống hoá hơi phục vụ phân tích đặc trưng phân rã và quá trình hóa hơi nhiên liệu.

2. Cơ sở lý thuyết

2.1 Mô hình chu trình nhiệt động

Động cơ JetCat P130-RX được mô hình hóa dưới dạng chu trình TBK một trục, bao gồm máy nén ly tâm một tầng, buồng đốt hình khuyên, tua-bin hướng trục một tầng và ống phụt hội tụ (Hình 1). Mô hình được xây dựng theo phương pháp phân tích chu trình nhiệt động không thứ nguyên (0D) trong điều kiện làm việc ổn định.



Hình 1. Mô hình động cơ TBK cỡ nhỏ;

a) Động cơ JetCat P130RX [7]; b) Mô hình 0D của động cơ Jetcat P130RX

Các thông số chính của động cơ do nhà sản xuất công bố và được sử dụng làm dữ liệu đầu vào cho mô hình (trình bày trong Bảng 1) bao gồm: lực đẩy danh định xấp xỉ 130 N, dải tốc độ quay tối đa khoảng 120.000 vòng/phút, nhiệt độ khí xả cực đại, lưu lượng nhiên liệu tương ứng với chế độ công suất lớn nhất, điện áp và công suất hệ thống điều khiển, cùng khối lượng và kích thước tổng thể của động cơ. Những thông số này được sử dụng để thiết lập điểm thiết kế và làm cơ sở hiệu chỉnh mô hình trong các bước tính toán tiếp theo.

Bảng 1. Thông số kỹ thuật động cơ TBK cỡ nhỏ JetCat P130-RX [7]

Thông số kỹ thuật	Đơn vị	Giá trị
Tỷ số nén của máy nén	–	3
Lưu lượng khối không khí qua động cơ	kg/s	0,30
Mức tiêu thụ nhiên liệu ở chế độ toàn tải	ml/phút	441 (130 N)
Mức tiêu thụ nhiên liệu ở chế độ không tải	ml/phút	100
Khối lượng động cơ	kg	1,326
Đường kính ngoài lớn nhất	mm	99
Chiều dài tổng thể	mm	284
Nhiệt độ khí xả (EGT)	°C	490–720
Tốc độ quay không tải	vòng/phút	40.000
Tốc độ quay cực đại	vòng/phút	127.000
Lực đẩy tĩnh tại chế độ không tải	N	4
Lực đẩy tĩnh cực đại	N	130
Vận tốc dòng khí xả	km/h	1560
Công suất động năng dòng phụt	kW	28,2
Suất tiêu hao nhiên liệu riêng	kg/N·h	0,185

Tại mỗi mặt cắt đặc trưng, các đại lượng trạng thái gồm áp suất toàn phần P_t, nhiệt độ toàn phần $T_t$và lưu lượng khối $\dot{m}$ được xác định tuần tự từ cửa vào đến loa phụt. Điểm vận hành của động cơ phải thỏa mãn đồng thời hai điều kiện:

$$W_t = W_c; \dot{m}_c = \dot{m}_t \quad (1)$$

trong đó W_t và W_c lần lượt là công tua-bin và công máy nén. Các phương trình bảo toàn này đảm bảo tính nhất quán vật lý của nghiệm thu được.

2.2 Cấu trúc bản đồ đặc tính thành phần

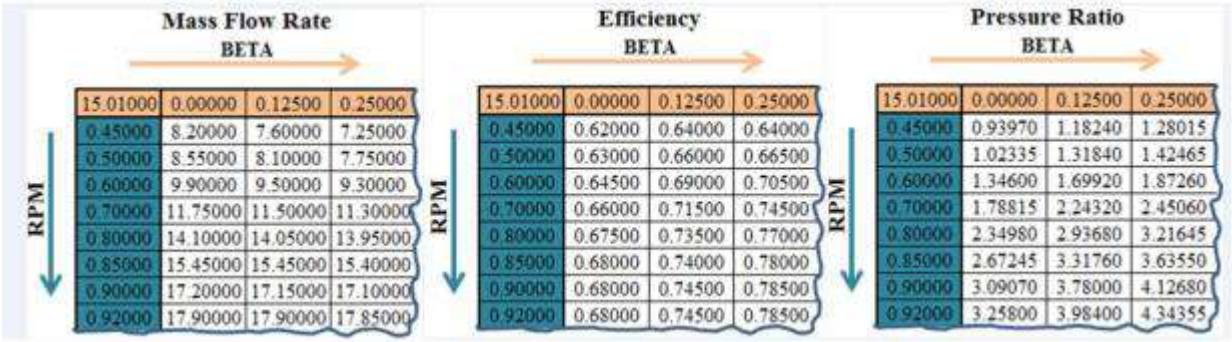
Trong chế độ ngoài thiết kế, đặc tính của máy nén và tua-bin được mô tả bằng bản đồ hiệu chỉnh nhằm loại bỏ ảnh hưởng của điều kiện môi trường đầu vào.

Các đại lượng hiệu chỉnh được định nghĩa:

$$N_{corr} = \frac{N}{\sqrt{T_{t,in}}}; \quad \dot{m}_{corr} = \frac{\dot{m}\sqrt{T_{t,in}}}{P_{t,in}} \quad (2)$$

Bản đồ máy nén biểu diễn quan hệ giữa lưu lượng hiệu chỉnh và tỷ số áp suất tại các đường tốc độ hiệu chỉnh không đổi. Biên trái của bản đồ là đường surge, trong khi biên phải là vùng nghẹt dòng.

Để nội suy hai chiều trên bản đồ, một tham số phụ trợ β được sử dụng nhằm xác định vị trí tương đối của điểm làm việc giữa hai biên ổn định (Hình 2). Tham số này không mang ý nghĩa vật lý trực tiếp mà đóng vai trò biến nội suy toán học.



Hình 2. Cấu trúc dữ liệu của ba bản đồ máy nén theo hai biến độc lập RPM và tham số β

Tua-bin được đặc trưng bởi bản đồ tương tự, mô tả quan hệ giữa lưu lượng hiệu chỉnh, tỷ số giãn nở và hiệu suất đoạn nhiệt.

2.3 Phương pháp đồng dạng hóa bản đồ

Do bản đồ đặc tính của JetCat P130-RX không được công bố, nghiên cứu sử dụng phương pháp đồng dạng hóa từ một bản đồ tham chiếu.

Tại điểm thiết kế, các hệ số tỷ lệ được xác định theo:

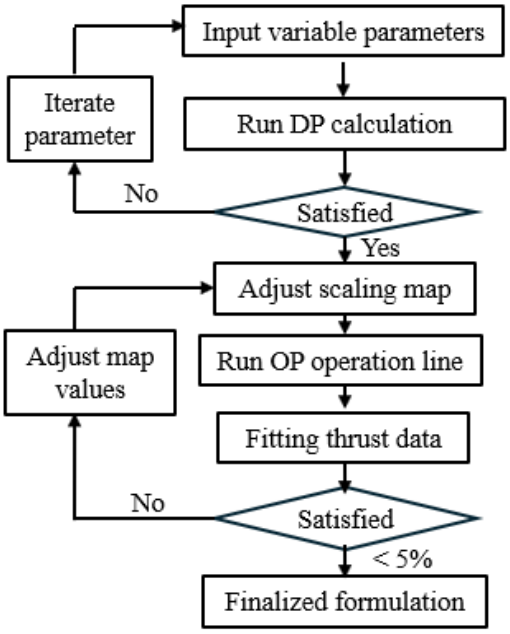
$$SF_{\dot{m}} = \frac{\dot{m}_{target}}{\dot{m}_{ref}}; SF_{PR} = \frac{PR_{target}-1}{PR_{ref}-1}; SF_{\eta} = \frac{\eta_{target}}{\eta_{ref}} \quad (3)$$

Trong đó: Trong đó, SF là hệ số tỉ lệ; các chỉ số $\dot{m}$, PRvà ηlần lượt biểu thị lưu lượng khối lượng, tỉ số nén và hiệu suất đẳng entropy. Ký hiệu *target* chỉ động cơ mục tiêu cần xây dựng mô hình, còn *ref* chỉ bản đồ tham chiếu được sử dụng để đồng dạng hóa.

Toàn bộ bản đồ tham chiếu được nhân với các hệ số này để thu được bản đồ mục tiêu. Phương pháp này đảm bảo khớp chính xác tại điểm thiết kế, đồng thời giả định rằng hình dạng bản đồ trong hệ tọa độ hiệu chỉnh được bảo toàn nhờ tính đồng dạng khí động.

2.4 Thuật toán xác định đường làm việc

Đường làm việc của động cơ được xác định thông qua quá trình lặp cân bằng giữa máy nén và tua-bin theo quy trình thể hiện trong Hình 3. Trước hết, các tham số đầu vào được thiết lập và thực hiện tính toán tại điểm thiết kế; nếu điều kiện cân bằng chưa thỏa mãn, các tham số được điều chỉnh cho đến khi đạt trạng thái phù hợp. Tiếp theo, ở chế độ ngoài thiết kế, giả định tốc độ hiệu chỉnh và tham số β, truy xuất đặc tính máy nén từ bản đồ đã hiệu chỉnh, tính toán quá trình cháy để xác định nhiệt độ vào tua-bin và tra bản đồ tua-bin tương ứng. Sai số cân bằng công suất và lưu lượng được kiểm tra và các biến lặp được điều chỉnh cho đến khi hội tụ. Tập hợp các điểm hội tụ khi thay đổi tốc độ hoặc lưu lượng nhiên liệu tạo thành đường làm việc đặc trưng của động cơ.



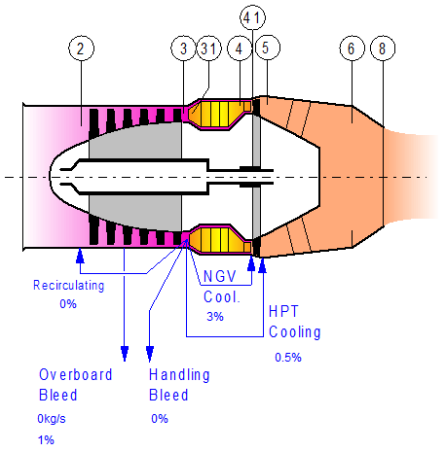
Hình 3. Sơ đồ thuật toán

2.5. Phần mềm Gasturb

Trong nghiên cứu này, phần mềm GasTurb 13 được sử dụng làm công cụ chính để xây dựng mô hình chu trình nhiệt động của động cơ tua-bin khí cỡ nhỏ. GasTurb là phần mềm chuyên dụng cho phân tích hiệu suất TBK hàng không và công nghiệp, cho phép tính toán ở cả điểm thiết kế và ngoài thiết kế [8].

Động cơ TBK cỡ nhỏ được thiết lập theo cấu hình động cơ phản lực một trục. Các trạm dòng khí được quy ước gồm: (2–3) máy nén, (3–4) buồng đốt, (4–5) tua-bin và (6–8) vòi phun (Hình 4). Các thông số đầu vào tại điểm thiết kế bao gồm điều kiện môi trường, tỷ số nén, hiệu suất các bộ phận và tổn thất áp suất. Những tham số này được xác định từ tính toán lý thuyết, tài liệu tham khảo và dữ liệu công bố của nhà sản xuất nhằm xây dựng mô hình trạng thái ổn định ban đầu.

Sau khi hoàn thành tính toán tại chế độ thiết kế, phần mềm GasTurb cho phép sử dụng các đường đặc tính của máy nén và tua-bin để mô phỏng các chế độ làm việc ngoài thiết kế. Quá trình hiệu chỉnh được thực hiện bằng cách điều chỉnh các hệ số của đường đặc tính nhằm bảo đảm sự cân bằng làm việc giữa các bộ phận trong động cơ, đồng thời giảm sai lệch giữa lực đẩy và nhiệt độ khí xả tính toán so với số liệu tham chiếu.



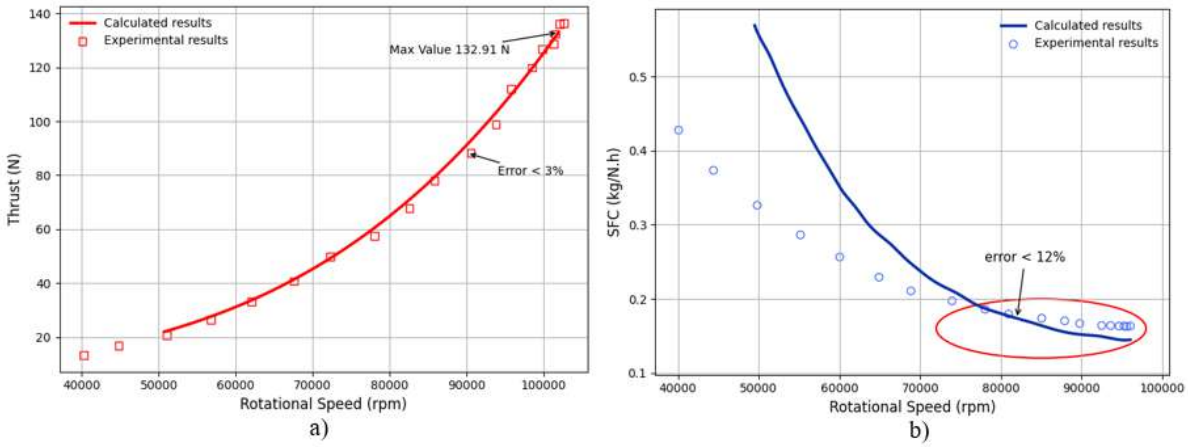
Hình 4. Cấu hình mô phỏng tua-bin 1 trục của Gasturb

Mô hình được coi là đạt độ tin cậy khi sai lệch tại các điểm đặc trưng nhỏ hơn 5%. Sau khi hiệu chỉnh, mô hình được sử dụng để xây dựng các đường đặc tính vận hành và phân tích các thông số nhiệt động của động cơ trên toàn bộ dải tốc độ làm việc.

3. Phương pháp mô phỏng

3.1. Kiểm chứng quy trình đồng dạng hóa

Để đánh giá độ tin cậy của phương pháp đồng dạng hóa bản đồ, quy trình mô phỏng được áp dụng cho động cơ BMT120KS – một động cơ cùng lớp công suất có dữ liệu hiệu suất công bố trong các bài báo [9]. Các thông số lực đẩy và lưu lượng nhiên liệu thu được từ mô hình được so sánh với dữ liệu tham chiếu trong toàn dải tốc độ quay.

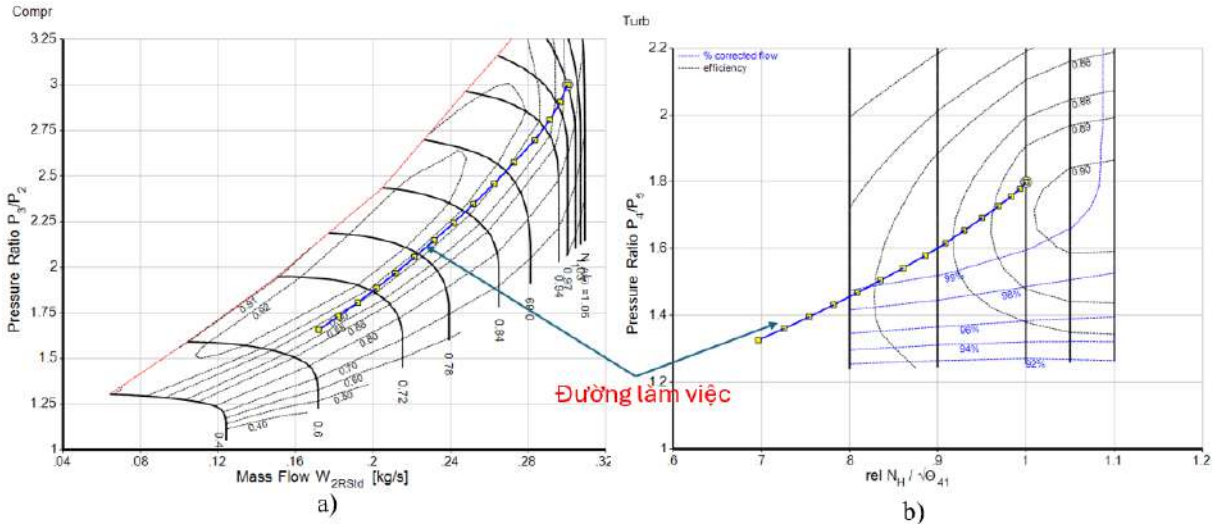


Hình 5. So sánh kết quả mô phỏng và kết quả thực nghiệm động cơ BMS120KS;
a) Đồ thị lực đẩy, b) Đồ thị suất tiêu hao nhiên liệu

Kết quả cho thấy sai số lực đẩy và suất tiêu hao nhiên liệu duy trì dưới 12% tại các điểm đặc trưng, trong khi xu hướng biến thiên theo tốc độ quay được tái hiện phù hợp với quy luật vật lý của động cơ một trục. Điều này cho thấy phương pháp đồng dạng hoá có khả năng tái tạo đặc tính toàn động cơ với độ chính xác chấp nhận được, tạo cơ sở áp dụng cho JetCat P130-RX trong điều kiện thiếu dữ liệu bản đồ thành phần.

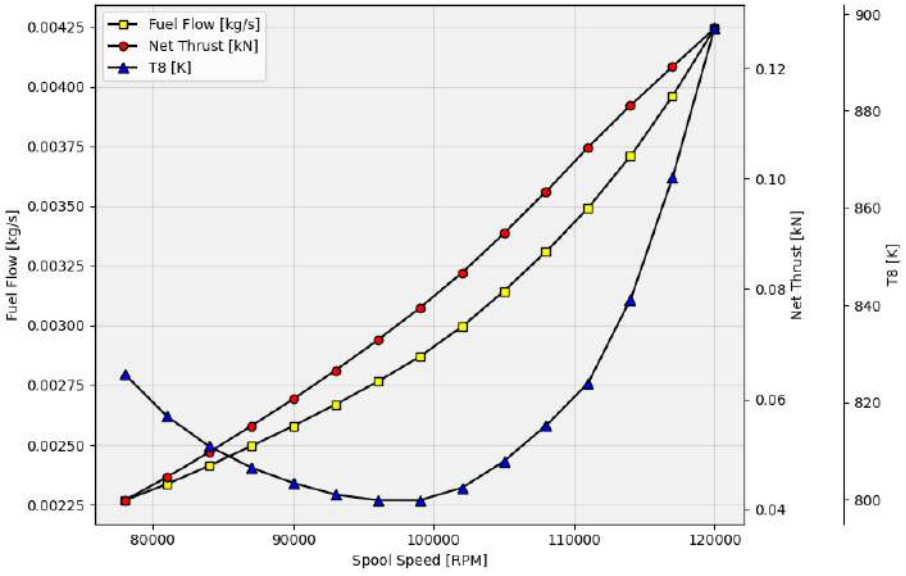
3.2. Đặc tính làm việc của JetCat P130-RX

Sau khi điều chỉnh các hệ số hiệu chỉnh để mô hình đạt trạng thái khớp tại hai chế độ không tải và công suất lớn nhất, mô hình được sử dụng để tính toán đặc tính làm việc ngoài điểm thiết kế trong dải tốc độ từ 40.000 đến 127.000 vòng/phút (Hình 6).



Hình 6. Bản đồ máy nén và tua-bin của Jetcat P130RX; a) Bản đồ máy nén; b) Bản đồ tua-bin

Kết quả trong Hình 7 cho thấy lực đẩy tăng gần theo quy luật bậc hai của tốc độ quay, phù hợp với quan hệ giữa động năng dòng phụt và vận tốc vòng của máy nén. Lưu lượng nhiên liệu tăng gần tuyến tính theo tốc độ, phản ánh yêu cầu cân bằng công suất giữa máy nén và tua-bin. Nhiệt độ khí xả đạt giá trị cực đại tại chế độ công suất lớn nhất và giảm dần ở các chế độ tải thấp, phù hợp với đặc tính vận hành của động cơ một trục.



Hình 7. Đặc tính vận hành của JetCat P130RX

Sai lệch giữa kết quả mô phỏng và các thông số công bố tại hai điểm đặc trưng đều nhỏ hơn 5%, cho thấy mô hình đã đạt trạng thái khớp ổn định và đáng tin cậy.

3.3. Thông số động lực học tại cửa vào ống hoá hơi

Từ đặc tính vận hành của động cơ đã được hiệu chỉnh, các thông số toàn phần tại cửa vào buồng cháy được trích xuất theo từng chế độ làm việc, bao gồm lưu lượng khối lượng, áp suất và nhiệt độ khí. Trên cơ sở đó, các đại lượng cục bộ tại cửa vào ống hoá hơi được xác định nhằm đặc trưng cho điều kiện động lực học của dòng chảy.

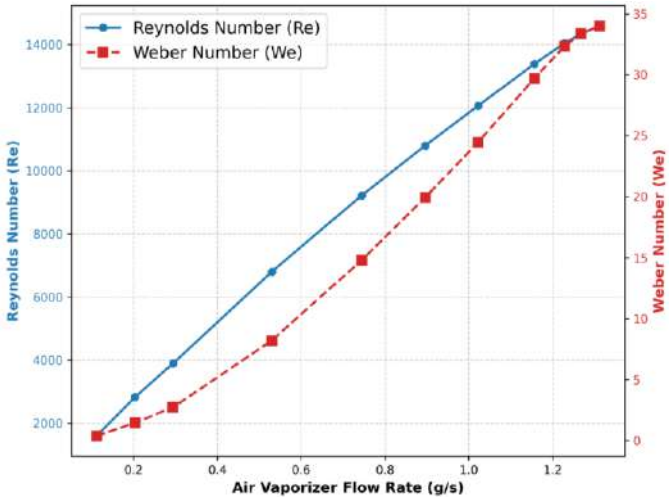
Vận tốc trung bình của dòng khí tại cửa vào ống được xác định từ quan hệ bảo toàn khối lượng. Trên cơ sở vận tốc đặc trưng này, các tham số không thứ nguyên mô tả chế độ dòng và khả năng phân rã tia nhiên liệu được xác định như sau:

$$U = \frac{\dot{m}_v}{\rho A} \quad (4)$$

$$Re = \frac{\rho U D_v}{\mu} \quad (5)$$

$$We = \frac{\rho U^2 d_{inj}}{\sigma} \quad (6)$$

Trong đó, $\dot{m}_v$ là lưu lượng khối lượng khí vào ống hoá hơi, thông qua mô phỏng buồng đốt, lưu lượng này bằng 5% tổng lượng qua động cơ, ρ là khối lượng riêng tại điều kiện cửa vào, A là tiết diện ống hoá hơi và U là vận tốc trung bình được xác định theo (4). D_v là đường kính trong của ống hoá hơi, d_{inj} là đường kính kim phun, μ là độ nhớt động lực của khí và σ là sức căng bề mặt của nhiên liệu. Các tham số không thứ nguyên dòng chảy trên toàn chế độ làm việc của động cơ được thể hiện như Hình 8.



Hình 8. Động lực học dòng đa pha tại ống hoá hơi theo chế độ động cơ

Số Reynolds được xác định theo (5) cho phép đánh giá chế độ chảy (chuyển tiếp hoặc rối hoàn toàn), trong khi số Weber theo (6) đặc trưng cho khả năng phá vỡ và phân rã tia nhiên liệu dưới tác dụng của lực khí động. Việc khảo sát các đại lượng này theo dải tốc độ quay của động cơ cho phép xác định vùng biến thiên của Revà We tại cửa vào ống hoá hơi, từ đó đánh giá đặc trưng hoà trộn nhiên liệu trong các chế độ vận hành từ không tải đến công suất lớn nhất.

5. Kết luận

Nghiên cứu này xây dựng một quy trình mô phỏng nhằm xác định đặc tính làm việc của động cơ tua-bin khí JetCat P130-RX trong điều kiện không có bản đồ đặc tính các bộ phận và không có bệ thử thực nghiệm. Mô hình chu trình nhiệt động không chiều (0D) được thiết lập, kết hợp với phương pháp đồng dạng hóa bản đồ, sau đó được hiệu chỉnh để đạt trạng thái cân bằng công suất phù hợp với các thông số công bố ở chế độ không tải và công suất lớn nhất .

Độ tin cậy của phương pháp được kiểm chứng gián tiếp thông qua việc áp dụng cho một động cơ cùng lớp công suất, cho sai số lực đẩy và lưu lượng nhiên liệu dưới 12%. Sau khi hiệu chỉnh, mô hình cho phép xác định đặc tính lực đẩy – lưu lượng nhiên liệu – nhiệt độ khí thải trong toàn dải tốc độ quay và trích xuất các thông số nội tại tại cửa vào ống hoá hơi.

Kết quả cho thấy phương pháp đề xuất có khả năng tái tạo đặc tính toàn động cơ với độ chính xác chấp nhận được, đồng thời cung cấp dữ liệu khí động cần thiết để tính toán các chỉ số không thứ nguyên phục vụ phân tích quá trình hóa hơi nhiên liệu. Đây là một giải pháp khả thi cho các nghiên cứu động cơ tua-bin khí cỡ nhỏ khi điều kiện thực nghiệm bị hạn chế.

Tài liệu tham khảo

1. Nascimento, M.A.R.d., et al. (2013), *Micro Gas Turbine Engine: A Review.* Progress in Gas Turbine Performance.
2. Ranran, X., L. Fengchao, and R. Lilei (2018), *An overivew on development of small turbine engines for missiles.* Advanced in Aeronatutical Science and Enfineering, 9 (1): p. 18–27.
3. Balli, Ö. (2020), *General aviation and thermodynamic performance analyses of micro turbojet engine used on drones and unmanned aerial vehicles (UAV).* Journal of Aviation Research, 2(2): pp. 115–141.
4. Ji, F., et al. (2020), *Experimental and numerical investigation on micro gas turbine as a range extender for electric vehicle.* Applied Thermal Engineering, 173, pp. 115236.
5. Mangra, et al. (2022), *Numerical and experimental investigation of a micro gas turbine combustion chamber.* International Journal of Modern Manufacturing Technologies, 14(3), pp. 139-145.
6. Coban, et al. (2022), *Performance of a microjet using component map scaling.* Aircraft Engineering and Aerospace Technology, 94(8), pp. 1201-1219.
7. GmbH, J. *P130-RX-BL Technical Specifications*. 2024 [cited 2026 26 February]; Official product specification page]. Available from: https://www.jetcat.de/en/productdetails/produkte/jetcat/produkte/RC%20ENGINES/Engines/p130_rx.
8. GmbH, G., *GasTurb 14: Design and Off-Design Performance of Gas Turbines*. Version 14 ed. 2022, Aachen, Germany: GasTurb GmbH.
9. Opponga, S.J.V.d.S. F.,and T.W.V. Backstroma (2017), *Experimental and numerical performance investigation of micro gas turbine*, in *Mechanical, Manufacturing and Materials Engineering Conference 2015*, South Africa.

PREDICTION OF MICROTURBINE ENGINE PERFORMANCE USING A MAP SCALING APPROACH

Abstract: *Determining the performance characteristics of micro turbojet engines is challenging due to the lack of component maps and the absence of an experimental test bench. This study proposes a simulation-based procedure to determine the overall operating characteristics of the JetCat P130-RX engine under limited data conditions. A zero-dimensional (0D) thermodynamic cycle model was developed in GasTurb, combined with a map scaling technique to adjust the compressor and turbine maps so that the model matches the published specifications at Idle and Max operating conditions.To assess reliability, the methodology was indirectly validated using the BMT120KS engine, yielding thrust and fuel flow deviations below 5%. After calibration for the JetCat P130-RX, the thrust–fuel flow–EGT characteristics across the engine speed range were determined, and internal thermodynamic parameters were extracted for subsequent combustor analysis. The proposed approach enables the reconstruction of engine performance in the absence of a physical test bench with acceptable accuracy.*

Keywords: JetCat P130-RX, map scaling, numerical simulation, vaporizer tube

Nghiên cứu thực nghiệm đặc tính lưu lượng vòi phun khí trong chế độ phun xung ngắn

Nguyễn Văn Hùng[1]*, Lương Đình Thi[2], Vũ Đức Mạnh[2]

[1]Hệ Quản lý học viên sau đại học, Học viện kỹ thuật quân sự;

[2]Viện Cơ khí động lực, Học viện kỹ thuật quân sự;

Email: hungtme@lqdtu.edu.vn; Tel: 0378641095

Tóm tắt

Bài báo trình bày nghiên cứu thực nghiệm xác định đặc tính lưu lượng của vòi phun khí điện từ Hana H2001 trong chế độ phun xung ngắn, phục vụ ứng dụng động cơ sử dụng nhiên liệu khí. Thí nghiệm được thực hiện với khí nén ở nhiệt độ 27 °C và áp suất 4 bar phun vào môi trường không khí 1 bar, tương ứng chênh áp 3 bar; đường kính lỗ phun 2,4 mm. Thời gian mở vòi phun t_p thay đổi từ 1 đến 20 ms và khảo sát tại ba chế độ tốc độ (1500, 1800 và 2600 vg/ph). Kết quả cho thấy đặc tính Q–t_p có hai vùng rõ rệt: vùng phi tuyến dưới 4 ms do ảnh hưởng đáp ứng động của kim phun và vùng gần tuyến tính khi tp ≥ 5 ms. Trong vùng tuyến tính, các mô hình hồi quy tuyến tính đạt hệ số xác định $R^2 > 0,99$ và cho phép xác định thời gian trễ hiệu dụng khoảng 1,11–1,72 ms tùy theo tốc độ. Trên cơ sở đó, bài báo đề xuất ngưỡng thời gian mở tối thiểu 5 ms để đảm bảo điều khiển lưu lượng ổn định trong các phương pháp phun xung ngắn.

Từ khóa: *vòi phun khí; phun xung ngắn; đặc tính lưu lượng; Hana H2001*

1. Đặt vấn đề/Mở đầu

Trong những năm gần đây, động cơ đốt trong sử dụng nhiên liệu khí như CNG, LNG, LPG, Hydro và Amoniac đang được nghiên cứu rộng rãi nhờ tiềm năng giảm phát thải độc hại và đa dạng hóa nguồn năng lượng [1-5]. Trong các hệ thống này, quá trình định lượng nhiên liệu được thực hiện thông qua vòi phun điện từ điều khiển bằng xung điện. Lượng nhiên liệu cấp vào xi lanh phụ thuộc trực tiếp thời gian mở vòi phun, do đó đặc tính lưu lượng theo thời gian mở vòi phun đóng vai trò quyết định đến độ chính xác lượng nhiên liệu cấp vào xi lanh.

Khác với nhiên liệu lỏng có khối lượng riêng lớn giúp giảm ảnh hưởng của tính nén được, dòng khí qua vòi phun chịu tác động mạnh của đặc tính nhiệt động và điều kiện chênh áp [6, 7]. Trong điều kiện chênh áp đủ lớn, dòng chảy qua lỗ phun có thể đạt trạng thái nghẽn, khi đó lưu lượng khối lượng phụ thuộc chủ yếu vào áp suất nguồn, nhiệt độ và tính chất nhiệt động của môi chất. Tuy nhiên, lý thuyết dòng nén ổn định chỉ phản ánh trạng thái khi tiết diện mở đã đạt giá trị xác định, do đó chưa mô tả đầy đủ đặc tính lưu lượng trong điều kiện vận hành thực tế.

Đặc biệt trong chế độ phun xung ngắn, khi thời gian mở chỉ kéo dài vài mili giây, quá trình thiết lập dòng chảy ổn định không diễn ra tức thời. Trong giai đoạn đầu của xung điều khiển, lực điện từ cần vượt qua lực lò xo và quán tính cơ cấu trước khi kim đạt hành trình mở đủ lớn, khiến tiết diện hữu hiệu của dòng chảy thay đổi liên tục theo thời gian. Sự kết hợp giữa động lực học kim phun và cơ chế dòng nén tạo nên đặc tính lưu lượng phức tạp, trong đó lưu lượng trung bình không còn tỷ lệ tuyến tính với thời gian điều khiển, đặc biệt rõ ràng trong vùng thời gian mở nhỏ [7]. Mặc dù một số nghiên cứu đã khảo sát đặc tính lưu lượng vòi phun khí trong điều kiện dòng ổn định, dữ liệu thực nghiệm về đặc tính lưu lượng trong chế độ phun xung ngắn vẫn còn hạn chế [8, 9]. Trong hiệu chỉnh hệ thống điều khiển động cơ, giả thiết tuyến tính giữa thời gian phun và lưu lượng thường được sử dụng để đơn giản hóa bản đồ điều khiển. Tuy nhiên, nếu đặc tính thực tế tồn tại vùng phi tuyến đáng kể, sai số định

lượng nhiên liệu có thể dẫn đến lệch hệ số dư lượng không khí, ảnh hưởng trực tiếp đến hiệu suất cháy và phát thải. Do đó, việc xác định chính xác đặc tính lưu lượng trong vùng phun xung ngắn có ý nghĩa quan trọng cả về mặt lý thuyết lẫn ứng dụng thực tiễn.

Xuất phát từ yêu cầu đó, nghiên cứu này tiến hành xây dựng đặc tính thực nghiệm lưu lượng theo thời gian mở (Q–tp) của vòi phun khí trong dải thời gian mở 1–20 ms, tại chênh áp 3 bar và nhiệt độ 27°C. Trên cơ sở dữ liệu thu được, bài báo phân tích sự tồn tại của vùng phi tuyến và vùng gần tuyến tính, ước lượng thời gian trễ hiệu dụng của hệ thống và đề xuất ngưỡng thời gian mở tối thiểu phục vụ điều khiển ổn định. Đồng thời, khả năng áp dụng kết quả khi thay đổi loại nhiên liệu khí cũng được thảo luận, góp phần mở rộng cơ sở dữ liệu cho thiết kế và hiệu chỉnh hệ thống phun khí trong động cơ đốt trong sử dụng nhiên liệu khí thay thế.

2. Thực nghiệm

2.1. Đối tượng nghiên cứu và điều kiện thí nghiệm

Đối tượng nghiên cứu là vòi phun khí điện từ loại Hana H2001, được điều khiển theo nguyên lý peak–hold nhằm đảm bảo khả năng mở nhanh và duy trì trạng thái mở ổn định trong thời gian xung với các thông số được thể hiện trên Bảng 1.

Bảng 1. Thông số chính của vòi phun Hana H2001 [10]

STT	Thông số	Giá trị
1	Điện áp	12 V DC
2	Điện trở cuộn dây	1,9 Ω
3	Áp suất khí tối đa	4,0 bar
4	Thời gian trễ mở (t_{do})	2,36 ms
5	Thời gian trễ đóng (t_{dc})	1,2 ms
6	Đường kính lỗ phun	2,4 mm
7	Dòng điện peak/hold	4 A/2 A

Khí nén được sử dụng làm môi chất khảo sát trong thử nghiệm. Áp suất khí nén trước vòi phun được duy trì ở mức p_0 = 4 bar, trong khi áp suất đầu ra của vòi phun là áp suất môi trường p_{mt} = 1 bar, tương ứng chênh áp 3 bar qua vòi phun. Nhiệt độ khí trong suốt quá trình thí nghiệm được duy trì ở mức 27°C.

Thời gian mở vòi phun tp được khảo sát trong dải từ 1 đến 20 ms nhằm đánh giá đặc tính lưu lượng trong vùng phun xung ngắn. Trong nghiên cứu, các chế độ tốc độ 1500, 1800 và 2600 vg/ph được sử dụng như các mốc quy đổi điều kiện vận hành kiểu động cơ 4 kỳ để xác định tần suất lặp xung điều khiển vòi phun. Với động cơ 4 kỳ, một chu kỳ công tác tương ứng hai vòng quay trục khuỷu. Giả thiết mỗi chu kỳ thực hiện một lần phun, n là tốc độ động cơ thì tần số phun finj được xác định theo công thức:

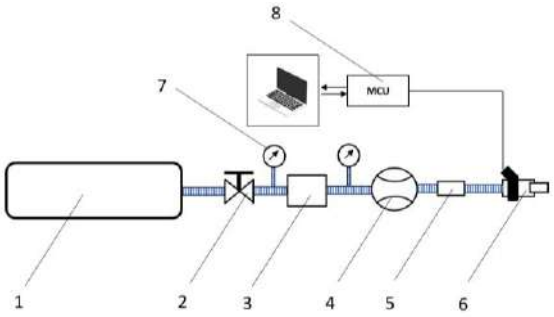
$$f_{inj} = \frac{n}{120} \quad (1)$$

Do đó, các tốc độ 1500, 1800 và 2600 vg/ph tương ứng với tần số phun lần lượt là 12,5 Hz, 15 Hz và 21,67 Hz (chu kỳ xung tương ứng 80 ms, 66,7 ms và 46,15 ms). Thời gian mở vòi phun đã được bù trễ mở và trễ đóng theo tài liệu công bố của nhà sản xuất để đảm bảo t_p đúng giá trị khảo sát. Khi đó thời gian điều khiển mở thực tế (t_{dk}) sẽ là:

$$t_{dk} = t_p + t_{do} - t_{dc} \quad (2)$$

2.2. Sơ đồ bố trí thí nghiệm

Việc bố trí thiết bị đo lưu lượng vòi phun khí được thể hiện trên hình 1. Khí nén ở áp suất 8 bar từ bình chứa (1) đi qua van (2) và van giảm áp (3) được hạ áp xuống 4 bar và duy trì ổn định áp suất khí trước vòi phun. Lưu lượng dòng khí được đo bởi thiết bị CS330C (4) trước khi đi qua van một chiều (5) đến vòi phun (6). Vòi phun được điều khiển bởi bộ điều khiển xung (8), cho phép thiết lập chính xác độ rộng xung theo giá trị khảo sát. Áp suất khí trước và sau van giảm áp được quan sát bằng đồng hồ đo áp suất (7). Cách bố trí này đảm bảo dòng khí một chiều, loại trừ ảnh hưởng hồi lưu và giảm dao động áp suất trong quá trình đo.



Hình 1. Sơ đồ bố trí thiết bị đo lưu lượng vòi phun khí

1. Bình khí nén; 2. Van bình; 3. Van giảm áp điện từ SMC ITV3050; 4. Thiết bị đo lưu lượng Sevenstar CS330C; 5. Van một chiều; 6. Vòi phun khí Hana H2001; 7. Đồng hồ áp suất; 8. Bộ điều khiển

2.3. Phương pháp đo và xử lý số liệu

Lưu lượng khí được đo bằng thiết bị CS330C (Sevenstar), thuộc loại thiết bị đo lưu lượng khối lượng kỹ thuật số (mass flow meter). Thiết bị này xuất ra giá trị lưu lượng khối lượng chuẩn hóa theo điều kiện tiêu chuẩn (0°C, 1 atm) dưới đơn vị SLM. Các thông số chính của thiết bị được thể hiện trên Bảng 2.

Bảng 2. Thông số chính của thiết bị đo lưu lượng Sevenstar CS330C [11]

STT	Thông số	Giá trị
1	Dải đo	200 SLM
2	Độ chính xác	±1.0% S.P. (lưu lượng > 35% F.S.) ±0.35% F.S. (lưu lượng ≤ 35% F.S.)
3	Độ tuyến tính	±0.5% F.S.
4	Độ lặp lại	±0.2% F.S.
5	Thời gian đáp ứng	≤ 1 giây
6	Áp suất khí tối đa điều khiển	10 bar
7	Hệ số nhiệt độ	Zero: ≤ ±0.05% F.S./°C, Span: ≤ ±0.1% F.S./°C

Do đặc tính đáp ứng động của cảm biến ở thang thời gian cỡ giây, tín hiệu đầu ra của thiết bị phản ánh giá trị lưu lượng trung bình đã được lọc nội tại, thay vì lưu lượng tức thời theo từng xung mở–đóng của vòi phun (thang ms). Vì vậy, trong nghiên cứu này, mục tiêu đo là xác định lưu lượng trung bình ổn định tương ứng với mỗi giá trị thời gian mở tp, không phân tích dạng sóng lưu lượng theo từng chu kỳ phun.

Tại mỗi giá trị tp, hệ thống được duy trì hoạt động trong 60 giây để đảm bảo đạt trạng thái ổn định trước khi xử lý số liệu. Trong mỗi khoảng đo, 600 giá trị lưu lượng được ghi lại và sử dụng để tính giá trị trung bình Q (đại diện cho lưu lượng tại điều kiện khảo sát) và độ lệch chuẩn σ (phản ánh mức độ dao động của tín hiệu trong trạng thái ổn định).

Giá trị trung bình Q được sử dụng để xây dựng đặc tính lưu lượng theo thời gian mở Q–tp. Đối với vùng tuyến tính, phương pháp hồi quy bình phương tối thiểu được áp dụng để xác định hệ số góc và hệ số chệch, từ đó suy ra thời gian trễ hiệu dụng tương đương của vòi phun.

Thiết bị được hiệu chuẩn theo khí N_2 tại điều kiện tiêu chuẩn theo cấu hình xuất xưởng. Trong phạm vi nghiên cứu này, môi chất khảo sát là khí nén và giá trị lưu lượng được sử dụng trực tiếp theo hệ chuẩn của thiết bị.

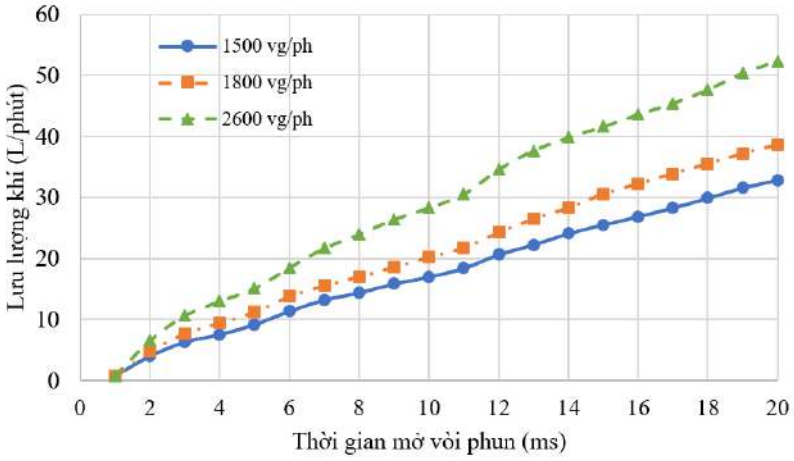
Phân tích lan truyền sai số được thực hiện dựa trên độ chính xác và độ lặp lại danh định của thiết bị. Độ không đảm bảo tổng hợp của giá trị lưu lượng trung bình được ước tính nhỏ hơn 3%, đủ tin cậy để đánh giá xu hướng đặc tính Q–tp trong dải khảo sát.

3. Kết quả và thảo luận

3.1. Đặc tính lưu lượng theo thời gian mở vòi phun

Quan hệ giữa lưu lượng trung bình Q và thời gian mở vòi phun t_p trong dải khảo sát 1–20 ms, tại các tốc độ 1500 vg/ph, 1800 vg/ph và 2600 vg/ph được thể hiện trên Hình 2. Kết quả cho thấy lưu lượng tăng theo thời gian mở vòi phun trong toàn bộ miền khảo sát. Tuy nhiên, trong vùng $t_p \leq 4$ ms, đường đặc tính Q-t_p thể hiện tính phi tuyến rõ rệt do chịu ảnh hưởng mạnh của quán tính cơ cấu điện từ, quá trình thiết lập dòng khí qua tiết diện biến thiên, quá trình tăng tốc dòng khí trong khoang trước lỗ phun. Vùng này phản ánh chi phối của động học cơ cấu kim phun hơn là đặc tính khí động học thuần túy của dòng chảy. Mặc dù bộ điều khiển đã thực hiện bù trễ đóng mở nhưng quá trình nâng kim và ổn định tiết diện dòng chảy vẫn cần một khoảng thời gian hữu hạn để thiết lập trạng thái dòng nghẽn.

Khi $t_p \geq 5$ ms, đường đặc tính có xu hướng tiệm cận tuyến tính. Đây là vùng mà chuyển động kim phun đã đạt trạng thái mở gần ổn định trong phần lớn thời gian xung, do đó lưu lượng trung bình chủ yếu tỷ lệ với thời gian mở. Với tỉ số áp suất $p_{mt}/p_0 = 1/4 = 0{,}25$, nhỏ hơn giá trị tới hạn của không khí ($\approx 0{,}528$) nên dòng chảy qua tiết diện hữu hiệu của vòi phun thuộc chế độ dòng nghẽn. Do đó, lưu lượng khối lượng trong trạng thái ổn định phụ thuộc chủ yếu vào áp suất, nhiệt độ của khí trước vòi phun và tiết diện hữu hiệu vòi phun, hoàn toàn độc lập với áp suất đầu ra của vòi phun. Trong thí nghiệm, áp suất và nhiệt độ nguồn khí nén được giữ không đổi, do đó lưu lượng trung bình phụ thuộc chủ yếu vào: thời gian mở hữu hiệu và sự thay đổi tiết diện hiệu dụng trong quá trình kim nâng. Điều này giải thích xu hướng tiệm cận tuyến tính của quan hệ Q–t_p trong vùng thời gian mở đủ lớn.



Hình 2. Đặc tính lưu lượng Q–t_p của vòi phun Hana H2001 ở các tốc độ khác nhau

Trong điều kiện đo lưu lượng trung bình, giá trị Q có thể xem tỷ lệ với tích của khối lượng khí phun trong một chu kỳ m_{cycle} và tần suất phun f_{inj}:

$$Q \sim f_{inj} \cdot m_{cycle} \quad (3)$$

Kết quả thực nghiệm cho thấy, khi tăng tần suất phun từ 12,5 Hz (1500 vg/ph) lên 21,67 Hz (2600 vg/ph), lưu lượng trung bình tăng gần tương ứng với tỷ lệ tăng tần suất. Ví dụ tại t_p = 10ms, lưu lượng ở 2600 vg/ph lớn hơn khoảng 1,67 lần so với 1500 vg/ph, trong khi tỷ lệ tần suất là khoảng 1,73. Sự tương đồng giữa hai tỷ lệ này xác nhận rằng, trong vùng tuyến tính khối lượng khí phun trên mỗi chu kỳ gần như không thay đổi đáng kể theo tần suất phun và sự gia tăng lưu lượng trung bình chủ yếu do số chu kỳ phun trong một đơn vị thời gian tăng lên.

Hệ số góc "a" của mô hình tuyến tính tăng theo tần suất phun, phản ánh mức tăng lưu lượng trung bình ứng với một đơn vị tăng thời gian mở khi số chu kỳ phun trong một giây lớn hơn. Điều này không hàm ý sự thay đổi đặc tính cơ học của kim phun mà là hệ quả trực tiếp của bản chất đo lưu lượng trung bình theo thời gian.

3.2. Hồi quy tuyến tính và thời gian trễ hiệu dụng

Để định lượng đặc tính trong vùng gần tuyến tính, tập dữ liệu t_p = 5–20 ms được xấp xỉ bằng mô hình hồi quy tuyến tính dạng:

$$Q = a \cdot t_p + b \quad (4)$$

Trong đó, hệ số góc a biểu thị mức biến thiên của lưu lượng trung bình theo đơn vị thời gian mở (L/ph·ms^{-1}), trong khi hệ số chệch b phản ánh sự dịch chuyển đặc tính do ảnh hưởng tổng hợp của quá trình mở van và thiết lập dòng chảy. Kết quả hồi quy thể hiện trên bảng 3 cho thấy mức phù hợp rất cao ($R^2 > 0,99$) ở cả ba tốc độ, cho thấy mô hình tuyến tính mô tả tốt đặc tính lưu lượng trung bình trong vùng này.

Trong nghiên cứu này, thời gian trễ hiệu dụng t_{eff} được xác định bằng cách ngoại suy đường hồi quy tuyến tính về Q = 0:

$$t_{eff} = -\frac{b}{a} \quad (5)$$

Giá trị t_{eff} thu được nằm trong khoảng 1,11–1,72 ms tùy theo tần suất phun. Cần lưu ý rằng t_{eff} không phải là trễ điện từ danh định của vòi phun, vì trong điều khiển đã thực hiện bù trễ mở và trễ đóng theo thông số nhà sản xuất, mà phản ánh ảnh hưởng tổng hợp của động học nâng kim trong điều kiện làm việc thực tế, quá trình hình thành tiết diện hữu hiệu và sự thiết lập dòng khí qua lỗ phun.

Bảng 2. Kết quả hồi quy tuyến tính (t_p = 5–20 ms) và thời gian trễ hiệu dụng

Tốc độ	**a (L/ph·ms^{-1})**	**b (L/ph)**	**R^2**	**t_{eff} (ms)**
1500 vg/ph	1,960	2,175	0,9985	1,11
1800 vg/ph	2,314	2,717	0,9978	1,17
2600 vg/ph	3,069	5,285	0,9938	1,72

Ở chế độ 2600 vg/ph, giá trị t_{eff} có xu hướng tăng so với hai chế độ còn lại. Hiện tượng này có thể liên quan đến sự tương tác giữa các chu kỳ phun khi khoảng cách giữa hai xung giảm, khiến quá trình thiết lập và suy giảm dòng khí chưa hoàn toàn độc lập giữa các chu kỳ kế tiếp. Ngoài ra, dao động áp suất cục bộ trong đường cấp khí và giới hạn đáp ứng động của hệ đo cũng có thể góp phần làm gia tăng sai lệch tuyến tính trong vùng thời gian mở rất ngắn.

Từ các kết quả trên, đặc tính lưu lượng của vòi phun khí trong chế độ phun xung ngắn có thể chia thành hai vùng làm việc đặc trưng: Vùng quá độ $t_p \leq 4$ ms quan hệ phi tuyến, chịu ảnh hưởng mạnh của động học kim phun và quá trình thiết lập dòng chảy, độ nhạy cao và khó mô tả chính xác bằng mô hình tuyến tính đơn giản; vùng gần tuyến tính tp ≥ 5 ms với mô hình Q = a·tp + b mô tả tốt đặc tính với sai số nhỏ; phù hợp cho xây dựng bản đồ điều khiển tuyến tính có bù thời gian trễ hiệu dụng. Do đó, để đảm bảo độ chính xác định lượng nhiên liệu và giảm sai số trong điều khiển phun khí, thời gian mở tối thiểu nên được lựa chọn xấp xỉ 5 ms nhằm duy trì làm việc trong vùng gần tuyến tính. Đối với các ứng dụng yêu cầu thời gian phun rất ngắn, cần xem xét xây dựng mô hình phi tuyến hoặc thuật toán bù động học để đảm bảo độ chính xác định lượng.

4. Kết luận

Nghiên cứu đã xây dựng được đặc tính thực nghiệm lưu lượng của vòi phun khí điện từ Hana H2001 trong chế độ phun xung ngắn dưới điều kiện áp suất nguồn p_0 = 4 bar, áp suất môi trường p_{mt} = 1 bar và nhiệt độ khí 27 °C. Thời gian mở vòi phun được khảo sát trong dải 1–20 ms với các tần suất phun tương ứng các chế độ quy đổi 1500, 1800 và 2600 vg/ph.

Kết quả cho thấy quan hệ giữa lưu lượng trung bình và thời gian mở tồn tại hai vùng đặc trưng. Trong vùng $t_p \leq 4$ ms, đặc tính mang tính phi tuyến rõ rệt do ảnh hưởng chi phối của động học cơ cấu kim phun và quá trình thiết lập dòng chảy qua tiết diện biến thiên. Trong vùng $t_p \geq 5$ ms, quan hệ Q–t_p tiệm cận tuyến tính với hệ số xác định $R^2 > 0{,}99$ cho phép mô tả bằng mô hình tuyến tính dạng $Q = a \cdot tp + b$. Từ mô hình hồi quy tuyến tính, thời gian trễ hiệu dụng của hệ thống được xác định trong khoảng 1,11–1,72 ms tùy theo tần suất phun. Giá trị này phản ánh tổng hợp ảnh hưởng của đáp ứng động kim phun và quá trình thiết lập dòng khí trong điều kiện dòng nghẽn, ngay cả khi đã thực hiện bù trễ đóng–mở theo thông số danh định của nhà sản xuất. Đồng thời, nghiên cứu đề xuất ngưỡng thời gian mở tối thiểu khoảng 5 ms để đảm bảo làm việc trong vùng gần tuyến tính, thuận lợi cho xây dựng và hiệu chỉnh bản đồ điều khiển phun khí.

Các kết quả thu được cung cấp cơ sở thực nghiệm cho thiết kế phương pháp điều khiển vòi phun khí trong động cơ sử dụng nhiên liệu thay thế như CNG, LPG, Hydro hoặc Amoniac. Trong các nghiên cứu tiếp theo, cần xem xét ảnh hưởng của áp suất nguồn, nhiệt độ khí và loại môi chất đến đặc tính lưu lượng, nhằm hoàn thiện mô hình dự đoán và mở rộng khả năng ứng dụng trong điều kiện vận hành thực tế của động cơ.

Tài liệu tham khảo

1. Phạm, H.B., et al. (2019). *Nghiên cứu sử dụng LPG như một nhiên liệu thay thế cho động cơ đốt trong-khả năng sử dụng ở Việt Nam.*
2. Cheng, Q., et al. (2025). *Ammonia as a sustainable fuel: Review and novel strategies.* Renewable and Sustainable Energy Reviews. **207**: p. 114995.
3. Fayaz, H., et al. (2012). *An overview of hydrogen as a vehicle fuel.* Renewable and Sustainable Energy Reviews. **16**: p. 5511-5528.
4. Kumar, S., et al. (2011). *LNG: An eco-friendly cryogenic fuel for sustainable development.* Applied energy. **88**(12): p. 4264-4273.
5. Bae, C. and J. Kim (2017). *Alternative fuels for internal combustion engines.* Proceedings of the Combustion Institute. **36**(3): p. 3389-3413.
6. Ferrari, A., S. Gurrì, and O. Vento (2024). *Injected fuel mass and flow rate control in internal combustion engines: a systematic literature review.* Energies. **17**(24): p. 6455.

7. von Helldorff, H. and G.J. Micklow (2019). *A Comprehensive Compressible Injection Model for High Pressure Gaseous Fuel Injectors.* Journal of Multidisciplinary Engineering Science and Technology (JMEST). **6**(2): p. 9523-9529.
8. Ivanov, Z., et al. (2024). *Influence of the temperature state of gas injectors on its flow characteristics*. in *AIP Conference Proceedings*. AIP Publishing LLC.
9. Wang, T., L. Zhang, and Q. Chen (2020). *Effect of Valve Opening Manner and Sealing Method on the Steady Injection Characteristic of Gas Fuel Injector.* Energies. **13**(6): p. 1479.
10. HANA EMS (2021). *HANA H2001 LPG/CNG Injector Technical Specifications*. Retrieved February 20, 2026, from https://hanaems.com/h2001.
11. Sevenstar (2021). *CS330-C MFC/MFM User's Manual*. Beijing Sevenstar Flow Co., Ltd.

Experimental study on the flow characteristics of the Hana H2001 injector in the short pulse injection mode

Abstract: *This paper presents an experimental study to determine the flow characteristics of the Hana H2001 electromagnetic fuel injector in short pulse injection mode, for application in engines operating on gas fuels. The experiments were conducted using compressed air at 27°C and 4 bar, injected into ambient air at 1 bar, corresponding to a pressure differential of 3 bar; the injector nozzle diameter was 2.4 mm. The injector pulse width (tp) was varied from 1 to 20 ms and was tested at three engine speeds (1500, 1800, and 2600 rpm). The results indicate that the Q-tp characteristics have two distinct regions: a nonlinear region below 4 ms due to the dynamic response of the needle valve, and a near-linear region when tp ≥ 5 ms. In the linear region, the regression models achieved R^2 values greater than 0.99, allowing for the determination of an effective delay time between 1.11 ms and 1.72 ms depending on engine speed. Based on these findings, the paper proposes a minimum pulse width of 5 ms to ensure stable flow control in short pulse injection strategies.*

Keywords: gas injector; short pulse injection; flow characteristics; Hana H2001

Nghiên cứu, xây dựng mô hình trạng thái nhiệt của cabin xe sử dụng mô phỏng CFD

Chu Hữu Sỹ[1], Vũ Đức Mạnh[1], Phùng Văn Được[1]

[1]*Học viện Kỹ thuật quân sự*

Email: sy0981155758@gmail.com; Tel: 0981155758

Tóm tắt

Trạng thái nhiệt của cabin cabin ảnh hưởng lớn đến sức khỏe, sự tiện nghi của người lái và kíp xe, trực tiếp quyết định đến sự tỉnh táo, an toàn trong khai thác vận hành. Trạng thái nhiệt của cabin cabin phụ thuộc vào các yếu tố kết cấu (thông gió, dẫn nhiệt từ các nguồn nóng, có/ không có điều hòa) và vận hành, ngoại cảnh (bức xạ mặt trời, tốc độ xe di chuyển, nhiệt độ môi trường). Bài báo sử dụng phương pháp mô phỏng CFD bằng phần mềm ANSYS FLUENT để xây dựng mô hình trạng thái nhiệt của cabin xe đầu kéo KZKT-7428 và khảo sát ảnh hưởng của một số thông số kết cấu và vận hành tới trạng thái nhiệt của cabin. Kết quả cho thấy khi không có điều hòa, nhiệt độ trung bình cabin lên tới hơn 50^{o}C, ảnh hưởng nghiêm trọng tới sức khỏe của người lái và kíp xe.

Từ khóa: *mô hình trạng thái nhiệt; CFD; ANSYS FLUENT; bức xạ nhiệt; mô hình rối.*

1. Đặt vấn đề

Sự tiện nghi nhiệt trong cabin đóng vai trò quan trọng trong việc đảm bảo an toàn và sức khỏe của người sử dụng, đồng thời ảnh hưởng trực tiếp đến trải nghiệm lái xe, hiệu quả quá trình vận chuyển. Đặc biệt, trong điều kiện khí hậu khắc nghiệt của mùa hè Việt Nam, nhiệt độ cao trong cabin có thể gây ra sự tích tụ nhiệt cục bộ và phát thải chất độc từ vật liệu nội thất như da, cao su,… gây ảnh hưởng nghiêm trọng đến sức khỏe con người. Hơn nữa, nhiệt độ bên trong cabin có thể lên tới 60÷70^{o}C [1] khi đỗ xe ngoài trời, gây ảnh hưởng tới tình trạng kĩ thuật và tuổi thọ của các trang thiết bị.

Để giải quyết vấn đề này, việc xây dựng mô hình trạng thái nhiệt của cabin ô tô là rất cần thiết. Thông qua mô hình, ta có thể dự báo được nhiệt độ cabin, bề mặt khác như vỏ xe, kính, taplo, nội thất,… theo điều kiện môi trường và vận hành của phương tiện. Mô hình này cũng được sử dụng trong việc thiết kế và tối ưu cấu trúc xe bằng cách so sánh các phương án lựa chọn, lắp đặt kính, phim cách nhiệt, vật liệu chế tạo, phương pháp cách nhiệt của trần và sàn xe nhằm giảm hấp thụ nhiệt tại các điểm nóng. Bên cạnh đó, mô hình cũng góp phần tối ưu hệ thống thông gió và điều hòa nhiệt độ trong quá trình thiết kế. Dựa vào kết quả đạt được, ta có thể đánh giá mức độ tiện nghi và an toàn nhiệt của người sử dụng qua các đại lượng có liên quan như nhiệt độ không khí, nhiệt độ bức xạ trung bình, tốc độ gió hay nhiệt độ tại một số vị trí ảnh hưởng đến người lái.

Trên thế giới, nhiều nghiên cứu sử dụng mô hình trạng thái nhiệt nhằm phân tích ảnh hưởng của thông gió, cách nhiệt đến tính tiện nghị nhiệt của người lái. Nghiên cứu [2] đã tiến hành đánh giá hệ thống thông gió và phân tích tần suất thay đổi không khí theo giờ của các loại máy bay, ô tô và xe buýt điển hình, sử dụng CFD. Bài báo [3] đã xây dựng mô hình nhằm phân tích ảnh hưởng của thông gió và bức xạ mặt trời đến nhiệt độ của buồng lái xe khi đỗ dưới trời nắng. Năm 2017, công trình [4] sử dụng mô phỏng số nhằm nghiên cứu môi trường nhiệt của cabin xe trong trường hợp xét và không xét đến cơ thể con người, qua đó cải thiện tính tiện nghi nhiệt cho người sử dụng. Nghiên cứu [5] tiến hành nâng cao hiệu quả sự tiện nghị, thoải mái nhiệt cho người lái thông qua việc thay đổi thiết kế trần xe. Bài nghiên

cứu này sử dụng công cụ mô phỏng CFD để khảo sát, tối ưu thiết kế. Sau đó thông qua thực nghiệm để đánh giá được hiệu quả thực tế của cải tiến. Bài báo [6] trình bày về tiện nghi nhiệt của con người bên trong khoang xe ô tô thông qua phân tích số ba chiều và các mô hình tiện nghi nhiệt tiên tiến. Đồng thời sử dụng mô hình Fanger để so sánh các kết quả nhận được. Bên cạnh đó, công trình [1] nghiên cứu về ảnh hưởng của việc thông gió đến nhiệt độ bên trong cabin khi xe phơi dưới trời nắng.

Qua bức tranh tổng thể trên, ta thấy được nhiều nghiên cứu thường giới hạn trong một loại phương tiện hoặc một trường hợp điển hình; việc khảo sát đồng thời các tình huống khai thác phổ biến như đóng/mở cửa sổ (thay đổi cơ chế thông gió tự nhiên), xe đỗ/xe di chuyển (thay đổi điều kiện đối lưu bên ngoài và tổn thất nhiệt qua vỏ xe), và đặc biệt là tác động của nhiệt do động cơ tỏa ra lên trường nhiệt cabin, vẫn chưa được trình bày một cách nhất quán trong cùng một khung mô hình nhằm phục vụ mục tiêu xây dựng mô hình trạng thái nhiệt. Nhất là đối với các phương tiện thế hệ cũ thuộc biên chế Quân đội ta hiện nay, thường không được trang bị hệ thống điều hòa nhiệt độ, điều kiện làm việc của người lái cũng chưa được tối ưu trong quá trình thiết kế dẫn đến các thông số xác định tính tiện nghi nhiệt vẫn còn ở mức thấp.

Bài báo này sử dụng mô phỏng CFD như công cụ trung tâm để xây dựng và phân tích mô hình trạng thái nhiệt của cabin dưới nhiều điều kiện vận hành và môi trường khác nhau. Đặc biệt, bài báo thể hiện một cách tường minh ảnh hưởng của nguồn nhiệt từ động cơ (thông qua điều kiện biên nhiệt/trao đổi nhiệt với các bề mặt liên quan), qua đó chỉ ra sự thay đổi của phân bố nhiệt độ trong cabin. Nhờ cách tiếp cận đa phương án và có xét nguồn nhiệt động cơ, kết quả nghiên cứu không chỉ có ý nghĩa mô tả hiện tượng mà còn tạo cơ sở thực tiễn để định hướng cải tiến bố trí cửa gió, lựa chọn lưu lượng và tải điều hòa, cũng như hỗ trợ xây dựng các chiến lược giám sát - điều khiển nhằm nâng cao tiện nghi nhiệt và hiệu quả năng lượng của phương tiện.

2. Xây dựng mô hình trong ANSYS FLUENT

2.1. Lý thuyết dòng chảy và trao đổi nhiệt trong phần mềm

Nhằm tính toán nhiệt độ và vận tốc của mô hình trao đổi nhiệt liên hợp, phần mềm ANSYS FLUENT sử dụng hệ phương trình Navier-Stokes [7] bao gồm:

$$\frac{\partial \rho}{\partial t} + \vec{\nabla}(\rho \vec{u}) = 0 \tag{1}$$

$$\frac{\partial}{\partial t}\left(\rho \vec{u}\right) + \vec{\nabla}(\rho \vec{u}\vec{u} + p\bar{\bar{\delta}} - \bar{\bar{\tau}}) = \rho \vec{f} \tag{2}$$

$$\frac{\partial}{\partial t}(\rho E) + \vec{\nabla}\left[\rho \vec{u}\left(E + \frac{p}{\rho}\right) - \bar{\bar{\tau}}.\vec{u} + \vec{q}\right] = \rho \vec{f}.\vec{u} \tag{3}$$

trong đó: ρ - khối lượng riêng; $\vec{u}$ - véc-tơ vận tốc dòng; p - áp suất của dòng khí; $\bar{\bar{\tau}}$ - ứng suất nhớt; $\vec{f}$ - gia tốc dòng khí do lực tác dụng vào khối khí (lực trọng trường, lực điện từ); E - năng lượng toàn phần; $\vec{q}$ - véc-tơ dòng nhiệt.

Trong ANSYS FLUENT, mô hình rối (Turbulence models) dùng để mô tả ảnh hưởng của dòng chảy rối lên vận tốc, áp suất, tổn thất năng lượng, trao đổi nhiệt và truyền khối mà

không phải mô phỏng trực tiếp mọi xoáy nhỏ. Bài báo này sử dụng mô hình rối $k-\omega$ SST giúp tính toán kết quả tốt, tiêu tốn ít tài nguyên hơn. Mô hình này tính toán dựa trên phương trình sau [7]:

$$\frac{\partial(\rho\omega)}{\partial t}+div(\rho\omega U)=div\left[\left(\mu+\frac{\mu_t}{\sigma_{\omega,l}}\right)grad(\omega)\right]+ \\ +\gamma_2\left(2\rho S_{ij}S_{ij}-\frac{2}{3}\rho\omega\frac{\partial U_i}{\partial x_j}\delta_{ij}\right)-\beta_2\rho\omega^2+2\frac{\rho}{\sigma_{\omega,2}}\frac{\partial k}{\partial x_k}\frac{\partial\omega}{\partial x_k} \quad (4)$$

Các hằng số có giá trị lần lượt là: $\sigma_k=1,0$; $\sigma_{\omega,1}=2,0$; $\sigma_\omega,2=1,17$; $\gamma_2=0,44$; $\beta_2=0,083$; $\beta^*=0,09$.

Ngoài ra, bài báo xét đến ảnh hưởng của bức xạ mặt trời thông qua mô hình bức xạ (Radiation models). Mô hình được sử dụng là sự kết hợp giữa thuật toán truy vết tia mặt trời (Solar ray tracing) và mô hình bức xạ tọa độ rời rạc (Discrete ordinate). Hướng chiếu của ánh nắng mặt trời thiết lập là thẳng đứng, từ trên xuống dưới và cường độ bức xạ nhập dựa trên thông số do tổ chức World Banh Group tiến hành đo đạc dữ liệu bức xạ mặt trời tại 5 trạm ở Việt Nam từ năm 2017 đến năm 2019 [8].

Bảng 1. Cường độ bức xạ mặt trời [8]

Thông số	**Giá trị**
Cường độ bức xạ trung bình trực tiếp (DSI)	850 W/m^2
Bức xạ khuếch tán trung bình (DHI)	150 W/m^2

2.2. Xây dựng mô hình

Để xây dựng mô hình CFD truyền nhiệt liên hợp cho cabin xe, một số giả thiết được đưa ra nhằm đơn giản hóa bài toán nhưng vẫn đảm bảo phản ánh đúng các cơ chế truyền nhiệt:

- Bài toán được xét ở trạng thái ổn định (steady-state) cho các trường hợp điển hình: xe đỗ dưới nắng (không có gió hoặc gió yếu); xe chạy với vận tốc không đổi (coi cabin là đứng yên, không khí chuyển động tương đối so với cabin). Quá trình tính toán không xét đến biến dạng của các cụm, chi tiết do nhiệt độ;

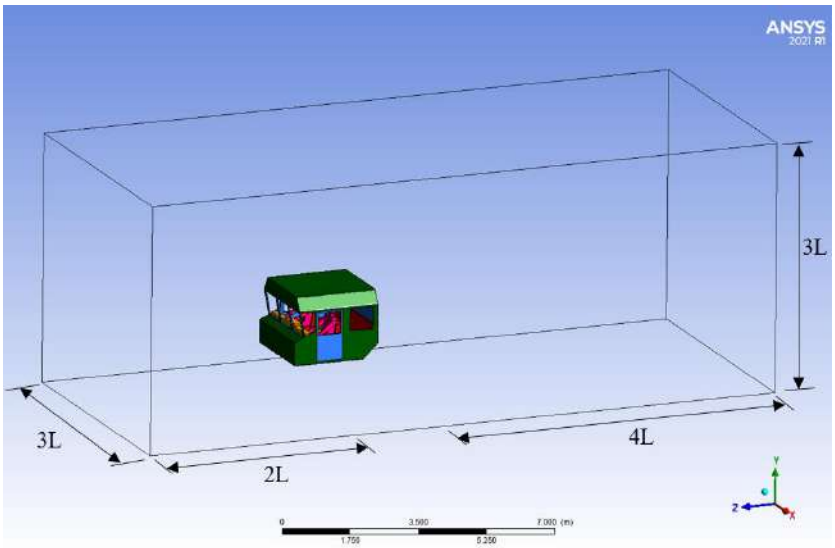
- Không khí coi như khô, bỏ qua ảnh hưởng của độ ẩm tới cảm giác tiện nghi;

- Người lái được mô phỏng như một bề mặt rắn có nhiệt độ không đổi. Bề mặt rắn trao đổi nhiệt với không khí trong cabin thông qua: đối lưu và bức xạ nhiệt;

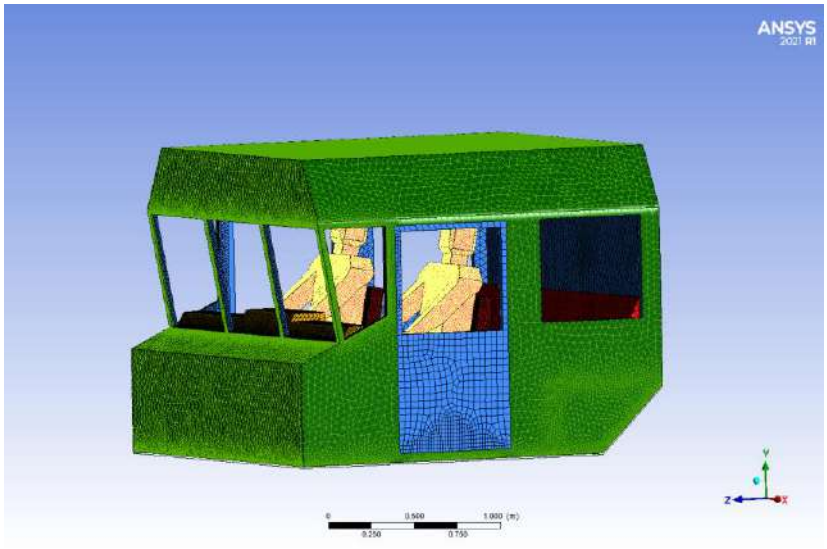
- Bức xạ mặt trời không đổi về cả phương, chiều và độ lớn, được mô hình hóa bằng thông lượng bức xạ đặt lên các bề mặt ngoài cabin;

- Ảnh hưởng của nhiệt độ tỏa ra từ động cơ được xét đến một cách gián tiếp thông qua nhiệt truyền qua các bề mặt của vách ngăn buồng động lực.

Hầm gió sử dụng trong mô phỏng được tạo ra trong ANSYS SPACECLAIM với các kích thước được thể hiện trên Hình 1. Sau khi tính toán, ta thấy tỉ số chặn dòng khí là 10%, phù hợp với yêu cầu của lý thuyết kiểm tra hầm gió được sử dụng trong lĩnh vực kĩ thuật ô tô. Đầu vào của hầm gió (Inlet) được cài đặt các thông số vận tốc 14 m/s, nhiệt độ dòng khí 40 °C. Đầu ra của hầm gió (Outlet) thiết lập theo chế độ "Outflow" trong điều kiện biên của ANSYS FLUENT.



Hình 1. Kích thước của hầm gió được sử dụng để mô phỏng



Hình 2. Lưới của mô hình sử dụng để tính toán

Mô hình gồm các khối sau: Vỏ cabin, lớp cách nhiệt/âm, sàn cabin, ghế ngồi, bảng taplo, kính xe và mô hình người. Vật liệu của mô hình được thiết lập theo Bảng 2 và Bảng 3:

Bảng 2. Thông số vật liệu của vỏ cabin [9]

Thông số	**Giá trị**
Khối lượng riêng (kg/m^3)	7900
Hệ số dẫn nhiệt (W/(m.K))	9,2÷31,7 tương ứng với nhiệt độ 100÷1500 K
Nhiệt dung riêng (J/kg.K)	272÷682 tương ứng với nhiệt độ 100÷1500 K

Bảng 3. Thông số của các vật liệu còn lại [9]

Thông số	**Kính**	**Lớp cách nhiệt/âm**	**Ghế**
Khối lượng riêng (kg/m^3)	3000	30	70
Hệ số dẫn nhiệt (W/(m.K))	1,4	0,02÷0,04 tương ứng với nhiệt độ 280÷400 K	1686
Nhiệt dung riêng (J/kg.K)	840	1040	0,05

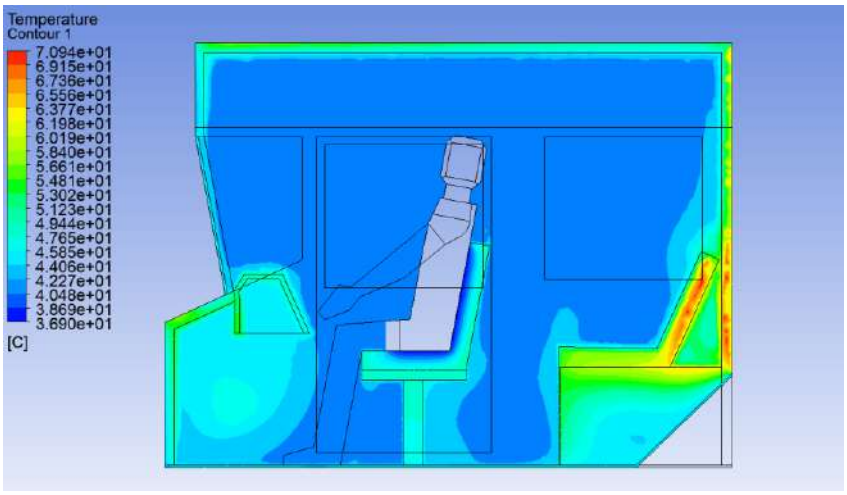
Bảng 4. Thông số điều kiện biên (Boundary Conditions)

Thông số	Loại điều kiện biên	Giá trị
Đầu vào dòng khí (Flow domain inlet)	Velocity-inlet	14 m/s
Nhiệt độ môi trường	Temperature	40 ^oC
Đầu ra dòng khí (Flow domain outlet)	Outflow	
Bề mặt vách ngăn buồng động lực	Non -slip wall	60 ^oC

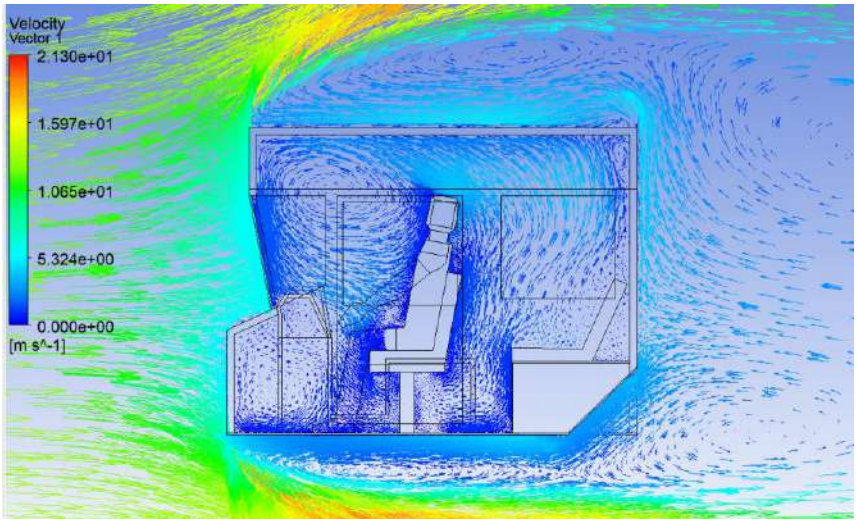
Bài báo tiến hành chia lưới theo ba mức độ khác nhau gồm: 11.032.904, 11.453.922 và 11.968.540 (phần tử). Kết quả cho thấy nhiệt độ trung bình bên trong cabin chênh lệch không quá 0,6% giữa hai lưới mịn hơn. Điều này chứng minh kết quả đã độc lập lưới và việc tiếp tục tăng số lượng ô lưới hầu như không làm thay đổi thêm kết quả.

3. Kết quả mô phỏng

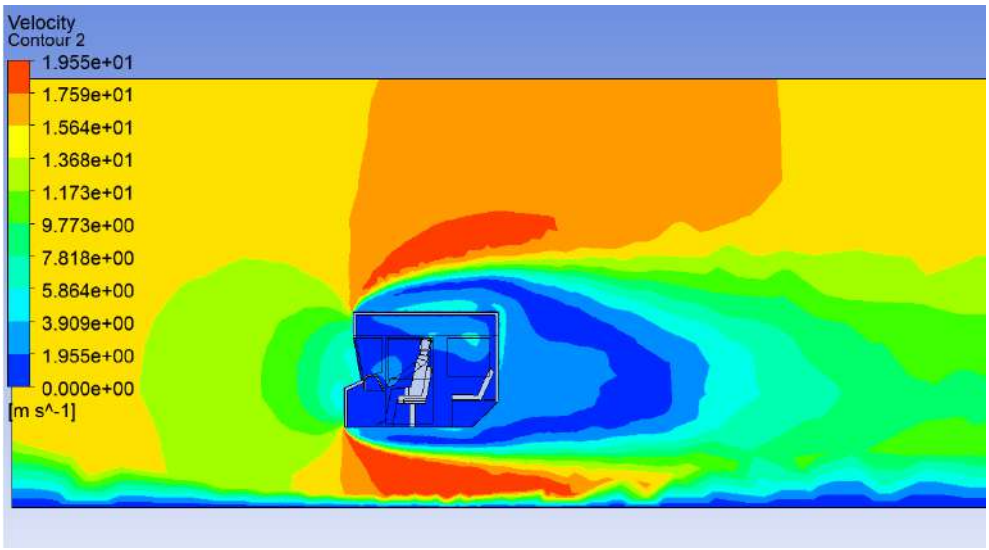
Hình 3 ÷ Hình 6 trình bày các kết quả mô phỏng trường nhiệt độ và trường vận tốc trong buồng lái, các kết quả này rút ra được một số nhận xét như sau:



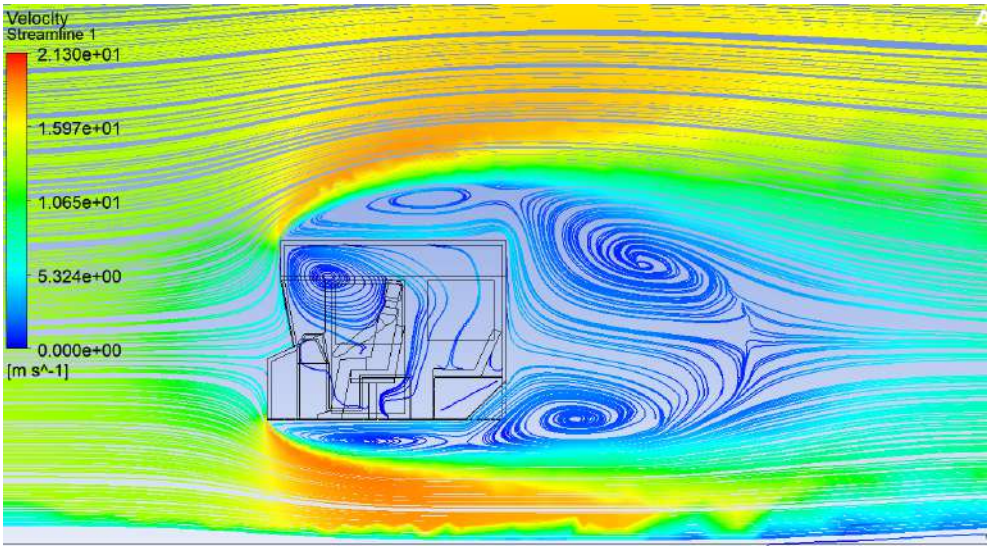
Hình 3. Phân bố nhiệt độ bên trong cabin trong trường hợp bức xạ 850 W/m^2, nhiệt độ môi trường 40 ^oC, vận tốc xe 14m/s, cửa mở hoàn toàn



Hình 4. Véc-tơ vận tốc dòng khí trong mô hình trong trường hợp bức xạ 850 W/m^2, nhiệt độ môi trường 40 ^oC, vận tốc xe 14m/s, cửa mở hoàn toàn



Hình 5. Phân bố vận tốc dòng khí trong mô hình trong trường hợp bức xạ 850 W/m², nhiệt độ môi trường 40 °C , vận tốc xe 14m/s, cửa mở hoàn toàn



Hình 6. Đường dòng theo vận tốc của dòng khí trong mô hình trong trường hợp bức xạ 850 W/m², nhiệt độ môi trường 40 °C , vận tốc xe 14m/s, cửa mở hoàn toàn

Thứ nhất, nhiệt độ trong buồng lái ở mức cao và phân bố không đồng đều. Hình 3 cho thấy phân bố nhiệt độ bề mặt cabin, nhiệt độ mái và phần vách trên đạt giá trị lớn nhất, có thể lên tới khoảng 65÷70 °C. Đây là khu vực chịu ảnh hưởng trực tiếp của bức xạ mặt trời và có khả năng tích nhiệt mạnh do vật liệu kim loại. Trong khi đó, các bề mặt phía dưới và vùng sàn có nhiệt độ thấp hơn, khoảng 40÷50 °C. Điều này cho thấy mái cabin là nguồn xâm nhập nhiệt chủ yếu từ môi trường bên ngoài.

Thứ hai, tồn tại nhiệt lớn theo phương thẳng đứng trong buồng lái. Từ hình cắt nhiệt độ trong cabin (Hình 3) có thể thấy:

+ Nhiệt độ không khí vùng sát trần cao hơn rõ rệt so với vùng sát sàn.

+ Nhiệt độ không khí quanh khu vực đầu và thân trên của người lái lớn hơn so với vùng chân. Điều này phản ánh đặc trưng của đối lưu tự nhiên: không khí nóng nhẹ hơn bốc lên trên, tích tụ ở vùng trần cabin, gây ra cảm giác oi bức cho người ngồi trong xe.

Thứ ba, các vùng nhiệt độ cao tập trung tại các bề mặt bao che.

Các khu vực như: Mái cabin, vách trước, các cạnh tiếp giáp giữa vách và mái, các khu vực gần kính chắn gió, đều có nhiệt độ cao, tạo thành các "nguồn nhiệt bức xạ thứ cấp" truyền nhiệt ngược lại vào không khí trong cabin thông qua đối lưu và bức xạ.

Thứ tư, khả năng trao đổi không khí tự nhiên trong cabin kém.

Từ các Hình 4÷Hình 6 cho thấy:

+ Vận tốc dòng khí bên trong cabin rất nhỏ, gần như tĩnh.

+ Dòng khí chủ yếu lưu thông bên ngoài cabin do tác động của chuyển động xe hoặc gió môi trường.

+ Bên trong cabin xuất hiện các vùng xoáy yếu và vùng "chết" khí động, làm cho không khí nóng khó thoát ra ngoài.

Điều này dẫn đến hiện tượng: Nhiệt bị tích tụ trong cabin, đặc biệt ở vùng trên cao và khu vực xung quanh người lái.

Thứ năm, môi trường nhiệt trong buồng lái là môi trường bất lợi cho người sử dụng. Với nhiệt độ không khí trong cabin dao động trong khoảng 40÷55°C ở nhiều vùng, kết hợp với vận tốc không khí rất nhỏ: Cảm giác nóng bức tăng cao, Khả năng trao đổi nhiệt của cơ thể người với môi trường bị hạn chế, Dễ gây mệt mỏi, giảm khả năng tập trung và phản xạ của người lái.

Kết quả cho thấy: Bức xạ mặt trời và dẫn nhiệt qua mái, vách kim loại là nguồn nhiệt chủ đạo. Đối lưu bên trong cabin chỉ đóng vai trò phân phối lại nhiệt, chứ không có khả năng thải nhiệt hiệu quả ra môi trường.

4. Phân tích biến đổi nhiệt độ trung bình trong cabin theo các tham số vận hành và môi trường

Trong nghiên cứu này, đặc tính nhiệt của cabin ô tô được đánh giá thông qua mô phỏng CFD trong ANSYS FLUENT dưới các điều kiện môi trường và vận hành khác nhau. Đại lượng được sử dụng để phân tích là chênh lệch nhiệt độ trung bình giữa không khí trong cabin và môi trường xung quanh, ký hiệu $\Delta T = T_{moitruong} - \overline{T}_{cabin}$, nhằm phản ánh mức độ tích nhiệt của cabin so với bên ngoài. Trên cơ sở đó, ảnh hưởng của bốn tham số chính, gồm cường độ bức xạ mặt trời, vận tốc di chuyển của xe, nhiệt độ môi trường và mức độ mở cửa sổ, được khảo sát riêng rẽ để làm rõ xu hướng biến thiên nhiệt và mức độ nhạy của hệ thống. Các kết quả thu được không chỉ cho phép nhận diện tham số chi phối quá trình tích nhiệt trong cabin mà còn cung cấp cơ sở khoa học cho việc đề xuất các giải pháp cải thiện tiện nghi nhiệt và giảm tải cho hệ thống điều hòa không khí.

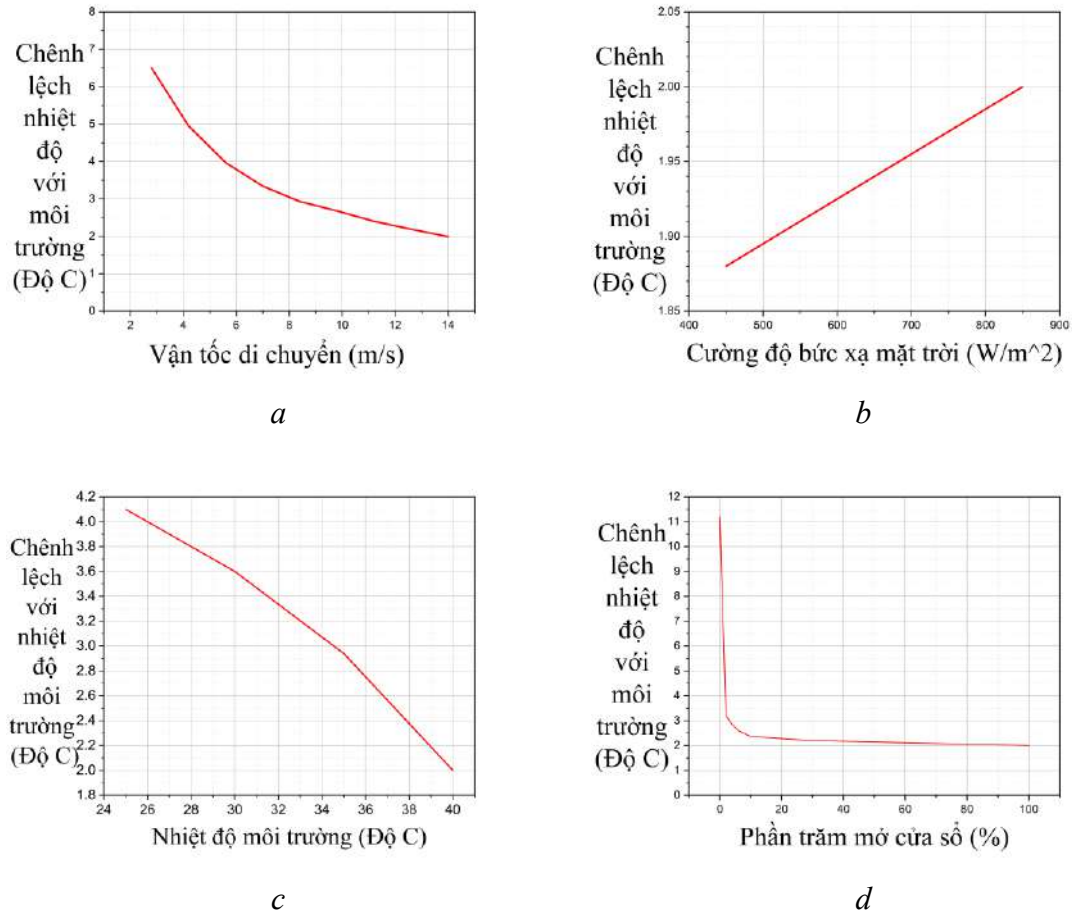
Qua phân tích các đồ thị trên, ta thấy:

+ Đồ thị Hình 7a cho thấy khi vận tốc xe tăng (xấp xỉ từ 3 đến 14 m/s), ΔT giảm rõ rệt, từ khoảng 6,5 ℃ xuống còn ~2,0 ℃. Đường cong có dạng giảm phi tuyến, với độ dốc lớn ở vùng vận tốc thấp và giảm dần ở vùng vận tốc cao.

+ Đồ thị Hình 7b khảo sát cường độ bức xạ mặt trời khoảng 450÷850 W/m², ta thấy ΔT tăng nhẹ từ khoảng 1,88 ℃ lên 2,00 ℃. Mối quan hệ có dạng gần tuyến tính tăng, nhưng biên độ biến thiên nhỏ.

+ Đồ thị Hình 7c thể hiện khi nhiệt độ môi trường tăng từ khoảng 25 ℃ đến 40 ℃, ΔT giảm từ khoảng 4,1 ℃ xuống 2,0 ℃. Đây là xu hướng giảm gần phi tuyến (hoặc gần tuyến tính trong dải hẹp).

+ Đặc biệt trong đồ thị Hình 7d, khi thay đổi mức độ mở của cửa sổ từ trạng thái kín sang hé mở xuất hiện đường trao đổi không khí hiệu quả giữa cabin và môi trường, làm tăng lưu lượng thông gió đột ngột và giảm nhanh nhiệt tích tụ. Từ hé mở đến mở lớn, lưu lượng trao đổi tiếp tục tăng nhưng không tỷ lệ tuyến tính với diện tích mở.



Hình 7. Đồ thị sự thay đổi của chênh lệch nhiệt độ bên trong cabin so với môi trường theo các tham số vận hành, môi trường

5. Kết luận

Đã xây dựng được mô hình mô phỏng trạng thái nhiệt của cabin xe nhằm khảo sát ảnh hưởng của các yếu tố vận hành và môi trường đến nhiệt độ trong cabin. Kết quả mô phỏng CFD cho thấy chênh lệch nhiệt độ trung bình giữa cabin và môi trường phụ thuộc rất lớn vào các tham số chi phối trao đổi nhiệt đối lưu và thông gió (nhiệt độ môi trường, bức xạ mặt trời, tốc độ xe, khoảng hở của cửa thông gió). Các thông số này ảnh hưởng phi tuyến đến trạng thái nhiệt trong cabin, trong đó ảnh hưởng mạnh nhất là tốc độ di chuyển của xe, cường độ bức xạ mặt trời và độ mở cửa thông gió.

Trong điều kiện khắc nghiệt nhất (nhiệt độ môi trường là 40 ℃, tốc độ xe 14 m/s và đóng kín cửa sổ - tương ứng với chế độ hoạt động ở mùa hè miền Bắc Việt Nam), nhiệt độ trong cabin có thể đạt mức trên 51 ℃. Nhiệt độ này ảnh hưởng lớn đến sức khỏe và sự tập

trung của lái xe và kíp xe. Điều này đặt ra yêu cầu cần cải tiến hệ thống thông gió, điều hòa và cách nhiệt để phù hợp với điều kiện vận hành tại khí hậu nhiệt đới của Việt Nam.

Tài liệu tham khảo

1. Ziqiao Li Yingchao Zhang, Shengda Liu, Guohua Wang and He Chang (2023). A Study of Passenger Car Cabin Pre-Ventilation under the Sun. Energies.

2. Domenico Mazzeo Behrouz Pirouz, Stefania Anna Palermo, Seyed Navid Naghib, Michele Turco and Patrizia Piro (2021). CFDInvestigation of Vehicle's Ventilation Systems and Analysis of ACH in Typical Airplanes, Cars, and Buses.

3. AhmadO.Hasan Hani Al-Rawashdeh, HazemA.Al-Shakhanbeh, Mujahed Al-Dhaifallah, MohamedR.Gomaa and Hegazy Rezk (2021). Investigation of the Effect of Solar Ventilation on the Cabin Temperature of Vehicles Parked under the Sun. Sustainability.

4. Florin Bode Paul Danca, Ilinca Nastase and Amina Meslem (2017). CFD simulation of a cabin thermal environment with and without human body – thermal comfort evaluation. EENVIRO.

5. S. Ibrahim and R. C. Mehta (2018). An Investigation of Air Flow and Thermal Comfort of Modified Conventional Car Cabin using Computational Fluid Dynamics Journal of Applied Fluid Mechanics.

6. Saboora Khatoon and Man-Hoe Kim (2020). Thermal Comfort in the Passenger Compartment Using a 3-D Numerical Analysis and Comparison with Fanger's Comfort Models. Energies.

7. ANSYS (2025). ANSYS FLUENT Theory Guide. Release 2025 R1.

8. World Bank Group (2019). Vietnam - Solar Radiation Measurement Data.

9. Leo Verde Niklas Axenholm Strömberg (2020). Energy Performance Simulations of a Scania Truck Cabin. Luleå University of Technology.

Research on the development of a thermal-state model of a vehicle cabin using CFD simulation

Abstract: *The thermal state of the vehicle cabin has a major influence on the health and thermal comfort of the driver and crew, and directly affects alertness and operational safety. The cabin thermal state depends on both structural factors (ventilation, heat conduction from hot sources, and the presence or absence of an air-conditioning system) and operating/environmental conditions (solar radiation, vehicle speed, and ambient temperature). In this study, a CFD-based approach using ANSYS FLUENT was employed to develop a thermal-state model of the KZKT-7428 tractor cabin and to investigate the effects of selected structural and operating parameters on cabin thermal conditions. The results show that, in the absence of air conditioning, the average cabin temperature can exceed 50 °C, posing a serious risk to the health of the driver and crew.*

Keywords: *Thermal-state model; CFD; ANSYS FLUENT; thermal radiation; turbulence model.*

ƯỚC LƯỢNG TỐC ĐỘ VÀ GÓCTRƯỢT CỦA UAV KHI SỬ DỤNG HAI ỐNG PITOT ĐỐI XỨNG VÀ THUẬT TOÁN EKF + LSTM

Phạm Minh Kha[1]*, Cao Hữu Tình[1],

Nguyễn Quang Thi[1], Nguyễn Xuân Trường[1]

Viện Tên lửa và kỹ thuật điều khiển;

Học viện Kỹ thuật quân sự

Email: phamminhkha102@mta.edu.vn;

Tel: 0969998319

Tóm tắt

Bài báo đề xuất phương pháp đo vận tốc và góc trượt UAV dựa trên cấu hình hai ống Pitot đối xứng kết hợp bộ lọc Kalman mở rộng và mạng LSTM tự hiệu chuẩn. Phương pháp khai thác chênh lệch áp suất để tăng độ nhạy với góc trượt và sử dụng học sâu để bù sai lệch cảm biến theo thời gian thực. Phân tích lý thuyết, mô phỏng Monte Carlo và thử nghiệm bay thực tế cho thấy sai số vận tốc giảm 71% và sai số góc trượt giảm 65% so với cấu hình Pitot đơn. Hệ thống duy trì ổn định trong điều kiện gió ngang 6 m/s và phát hiện tắc nghẽn với độ chính xác 65%.

***Từ khóa**: UAV; ống Pitot đôi; EKF; LSTM; tự hiệu chuẩn*

1. Giới thiệu

Trong các hệ thống bay không người lái (UAV), việc xác định chính xác vận tốc không khí và góc trượt đóng vai trò nền tảng đối với điều khiển ổn định, dẫn đường và tối ưu hóa hiệu suất khí động. Sai số trong ước lượng các tham số này có thể gây suy giảm chất lượng điều khiển, tăng tiêu hao năng lượng, làm lệch quỹ đạo bay và trong trường hợp nghiêm trọng có thể dẫn đến mất ổn định động học ngang.

Truyền thống, vận tốc dòng khí được đo bằng ống Pitot đơn dựa trên áp suất động, trong khi góc trượt thường được suy ra gián tiếp từ mô hình động lực học hoặc từ các đầu dò đa lỗ. Tuy nhiên, cấu hình Pitot đơn tồn tại một số hạn chế đáng kể: (i) độ nhạy thấp đối với góc trượt nhỏ; (ii) ảnh hưởng mạnh của nhiễu loạn và gió ngang; (iii) sai lệch do trôi nhiệt và lão hóa cảm biến; (iv) khó phát hiện sớm hiện tượng tắc nghẽn cục bộ.

Trong điều kiện UAV cỡ nhỏ bay ở tốc độ thấp (5–15 m/s), áp suất động tương đối nhỏ, khiến sai số tương đối của cảm biến áp suất tăng lên đáng kể. Ngoài ra, UAV thường hoạt động trong môi trường khí quyển phức tạp với nhiễu động biên tầng thấp và gió ngang thay đổi nhanh, làm giảm độ tin cậy của các phương pháp ước lượng truyền thống.

Các nghiên cứu gần đây tập trung vào hợp nhất dữ liệu bằng Bộ lọc Kalman mở rộng (EKF) nhằm kết hợp thông tin từ IMU và cảm biến áp suất. Một số công trình đã khai thác ống

Pitot vi mô để ước lượng vận tốc và góc trượt cho UAV cỡ nhỏ. Các phương pháp hiệu chuẩn vận tốc sử dụng cảm biến MEMS cũng được đề xuất gần đây. Ngoài ra, các bộ quan sát phi tuyến dựa trên Pitot–IMU đã được nghiên cứu nhằm đảm bảo ổn định tiệm cận toàn cục. Các phương pháp ước lượng gió và góc trượt bằng cảm biến tiêu chuẩn cũng được trình bày trong.

Cấu hình hai ống Pitot đặt đối xứng qua trục dọc UAV tạo ra chênh lệch áp suất tỷ lệ với thành phần vận tốc ngang, từ đó cho phép đo trực tiếp β với độ nhạy cao hơn. Tuy nhiên, cấu hình này đặt ra bài toán hiệu chuẩn phi tuyến do sai số lắp đặt, sai lệch đặc tính cảm biến và ảnh hưởng của điều kiện môi trường. Việc kết hợp mô hình vật lý với thuật toán học sâu thích nghi theo thời gian thực là một hướng tiếp cận tiềm năng nhưng chưa được nghiên cứu đầy đủ.

Xuất phát từ các vấn đề trên, bài báo này đề xuất một phương pháp tích hợp phần cứng đối xứng và thuật toán tự hiệu chuẩn thông minh nhằm nâng cao độ chính xác và độ tin cậy của hệ đo khí động cho UAV cỡ nhỏ.

Các nghiên cứu gần đây tập trung vào hợp nhất dữ liệu bằng Bộ lọc Kalman mởrộng (EKF) nhằm kết hợp thông tin từ IMU và cảm biến áp suất [2, 3]. Một số công trình 2đã khai thác ống Pitot vi mô để ước lượng vận tốc và góc trượt cho UAV cỡ nhỏ [1]. Cácphương pháp hiệu chuẩn vận tốc sử dụng cảm biến MEMS cũng được đề xuất gần đây[4]. Ngoài ra, các bộ quan sát phi tuyến dựa trên Pitot–IMU đã được nghiên cứu nhằm đảm bảo ổn định tiệm cận toàn cục [6]. Các phương pháp ước lượng gió và góc trượt bằng cảm biến tiêu chuẩn cũng được trình bày trong [5]. Cấu hình hai ống Pitot đặt đối xứng qua trục dọc UAV tạo ra chênh lệch áp suất tỷ lệ với thành phần vận tốc ngang, từ đó cho phép đo trực tiếp β với độ nhạy cao hơn. Tuy nhiên, cấu hình này đặt ra bài toán hiệu chuẩn phi tuyến do sai số lắp đặt, sai lệch đặc tính cảm biến và ảnh hưởng của điều kiện môi trường. Việc kết hợp mô hình vật lý với thuật toán học sâu thích nghi theo thời gian thực là một hướng tiếp cận tiềm năng nhưng chưa được nghiên cứu đầy đủ. Xuất phát từ các vấn đề trên, bài báo này đề xuất một phương pháp tích hợp phần cứng đối xứng và thuật toán tự hiệu chuẩn thông minh nhằm nâng cao độ chính xác và độ tin cậy của hệ đo khí động cho UAV cỡ nhỏ. Các đóng góp chính của nghiên cứu bao gồm:

• Xây dựng mô hình khí động học đối xứng cho cấu hình hai ống Pitot và phân tích độ nhạy theo góc trượt.

• Phát triển bộ ước lượng EKF tích hợp biến trạng thái sai lệch cảm biến.

• Đề xuất mạng LSTM hai lớp để tự hiệu chuẩn sai lệch dài hạn và phát hiện tắc nghẽn.

• Phân tích ổn định dựa trên hàm Lyapunov và đánh giá lan truyền sai số.

• Thực nghiệm mô phỏng cánh trong môi trường gió nhân tạo với hơn 0.5 triệu mẫu dữ liệu.

Phương pháp đề xuất hướng tới xây dựng một hệ đo vận tốc – góc trượt có khả năng thích nghi, phát hiện lỗi sớm và đáp ứng yêu cầu thời gian thực trên nền tảng nhúng.

2. Xây dựng mô hình tính toán

2.1. Mô hình lý thuyết

Trong hệ tọa độ thân UAV, vận tốc tương đối của dòng khí được đặc trưng bởi độ lớn V và các thành phần theo trục thân. Góc trượt β biểu diễn thành phần vận tốc ngang so với trục dọc và là đại lượng quan trọng trong điều khiển ổn định ngang.Trong cấu hình hai ống Pitot đối xứng, áp suất đo tại mỗi nhánh phụ thuộc vào vận tốc dòng khí và góc lệch của dòng. Dựa trên phương trình Bernoulli, áp suất động tỷ lệ với bình phương vận tốc, trong khi ảnh hưởng của góc trượt được phản ánh thông qua hệ số khí động C (β).

Khai triển trong miền góc nhỏ cho thấy chênh lệch áp suất giữa hai ống Pitot có dạng gần tuyến tính theo β:

$$\Delta P \approx \rho V^2 k_1 \beta \tag{1}$$

Biểu thức này cho thấy hai đặc điểm quan trọng. Thứ nhất, độ nhạy của hệ đo tăng theo bình phương vận tốc, nghĩa là hệ thống hoạt động hiệu quả hơn ở dải vận tốc trung bình và cao. Thứ hai, sai số ước lượng góc trượt giảm nhanh theo V^4, làm cho cấu hình này đặc biệt phù hợp với các UAV bay ở vận tốc ổn định.

Tuy nhiên, trong thực tế, chênh lệch áp suất đo được không chỉ phụ thuộc vào các đại lượng khí động lý tưởng mà còn bị ảnh hưởng bởi sai lệch vi sai giữa hai kênh đo. Sai lệch này có thể phát sinh từ bất đối xứng lắp đặt, sai số cảm biến và nhiễu dòng khí. Khi đó, mô hình đo được biểu diễn:

$$\Delta P^{\text{meas}} = \rho V^2 k1 \beta + b \tag{2}$$

Trong đó b là sai lệch tổng hợp có tính chất biến thiên theo thời gian. Sự phụ thuộc đồng thời của tín hiệu đo vào V, β và b cho thấy bài toán không thể giải trực tiếp mà cần một phương pháp ước lượng trạng thái để xác định đồng thời các tham số này. Mặc dù hai ống Pitot sử dụng cùng loại cảm biến áp suất vi sai, giả thiết rằng sai số đo của hai kênh là giống nhau là không hoàn toàn chính xác trong điều kiện thực tế. Sai lệch vi sai b trong (2) không chỉ phản ánh sai số nội tại của cảm biến mà còn bao gồm nhiều thành phần bất đối xứng khác. Cụ thể, sai lệch b được hình thành từ các nguồn chính sau:

• Sai số lắp đặt: Hai ống Pitot không thể đạt đối xứng hình học tuyệt đối, sai lệch nhỏ về góc hoặc vị trí dẫn đến khác biệt áp suất đo.

• Bất đối xứng khí động: Dòng khí thực tế chịu ảnh hưởng của nhiễu loạn, lớp biên và xoáy, làm cho áp suất tại hai vị trí không hoàn toàn đối xứng.

• Sai số cảm biến: Mặc dù cùng loại, các cảm biến vẫn tồn tại sai lệch offset, trôi nhiệt và sai số hiệu chuẩn khác nhau, dẫn đến $bL \neq bR$.

• Ảnh hưởng đường ống dẫn: Chiều dài và đặc tính động học của đường ống khí có thể gây trễ và suy giảm áp suất khác nhau giữa hai nhánh.

2.2. Mô hình sai lệch và bài toán ước lượng

Trong thực tế, chênh lệch áp suất đo được giữa hai ống Pitot không chỉ phụ thuộc vào vận tốc và góc trượt mà còn chịu ảnh hưởng của sai lệch giữa hai kênh đo. Sai lệch này hình thành từ sai số cảm biến, bất đối xứng lắp đặt và nhiễu khí động, đồng thời có xu hướng biến thiên chậm theo thời gian. Do đó, tín hiệu đo có thể được mô tả tổng quát bởi:

$$\Delta P^{\text{meas}} = \rho V 2\ k1\ \beta + b \tag{3}$$

Trong đó, thành phần thứ nhất phản ánh quan hệ khí động lý tưởng, còn b là sai lệch tổng hợp cần được ước lượng. Bài toán đặt ra là xác định đồng thời vận tốc dòng khí, góc trượt và sai lệch này trong điều kiện mô hình phi tuyến và nhiễu đo. Để giải quyết, hệ được biểu diễn dưới dạng vector trạng thái và sử dụng bộ lọc Kalman mở rộng nhằm ước lượng tối ưu theo thời gian thực. Đồng thời, một mạng LSTM được tích hợp để học và bù sai lệch dài hạn dựa trên dữ liệu lịch sử, giúp nâng cao độ chính xác và khả năng thích nghi của hệ thống.

2.3. Cơ chế tự hiệu chuẩn của mạng LSTM

Trong nghiên cứu này, thuật ngữ "tự hiệu chuẩn" được sử dụng để chỉ khả năng của mạng LSTM trong việc học và bù sai lệch cảm biến theo thời gian mà không cần hiệu chuẩn thủ công. Cụ thể, sai lệch vi sai b giữa hai kênh áp suất được xem là một đại lượng biến thiên chậm theo thời gian do ảnh hưởng của nhiệt độ, lão hóa cảm biến và điều kiện môi trường. Mạng LSTM được sử dụng để mô hình hóa thành phần sai lệch này dựa trên chuỗi dữ liệu đầu vào:

$$\hat{b}_{AI} = \text{LSTM}(P_L\ ,\ P_R) \tag{4}$$

Quá trình hoạt động của LSTM gồm hai giai đoạn:

• Huấn luyện ngoại tuyến (offline training): Mạng được huấn luyện trước trên tập dữ liệu thu thập từ thí nghiệm, trong đó sai lệch được suy ra từ dữ liệu tham chiếu (encoder góc và vận tốc chuẩn).

• Suy luận trực tuyến (online inference): Trong quá trình vận hành, LSTM nhận dữ liệu áp suất theo thời gian thực và ước lượng sai lệch $\hat{b}AI$ để bù vào kết quả của EKF:

$$b = b_{\text{EKF}} + \hat{b}_{AI} \tag{5}$$

Nhờ khả năng ghi nhớ chuỗi thời gian, LSTM có thể phát hiện xu hướng trôi dài hạn của cảm biến và cải thiện độ chính xác mà không cần can thiệp thủ công, do đó được gọi là cơ chế tự hiệu chuẩn.

2.4. Tích hợp LSTM tự hiệu chuẩn

Để xử lý sai lệch cảm biến có tính chất biến thiên chậm theo thời gian, một mạng LSTM hai lớp được sử dụng nhằm học đặc trưng động của sai lệch b từ dữ liệu lịch sử. Khác với EKF chỉ dựa trên mô hình vật lý cục bộ, LSTM có khả năng khai thác thông tin chuỗi thời gian dài hạn, từ đó nhận diện xu hướng trôi của cảm biến do nhiệt độ, lão hóa hoặc thay đổi môi trường.

Trong nghiên cứu này, mạng LSTM nhận đầu vào là chuỗi tín hiệu áp suất từ hai ống Pitot (và có thể mở rộng với dữ liệu IMU trong trường hợp bay thực), và ước lượng thành phần sai lệch:

$$\hat{b}^{AI} = \text{LSTM}(P_L , P_R) \tag{6}$$

Giá trị này được sử dụng để hiệu chỉnh kết quả của EKF:

$$b_k = b_k^{\text{EKF}} + \hat{b}^{AI} \tag{7}$$

Cách kết hợp này tận dụng ưu điểm của cả hai phương pháp: EKF đảm bảo tính nhất quán theo mô hình vật lý và khả năng lọc nhiễu thời gian thực, trong khi LSTM bổ sung khả năng học phi tuyến và bù sai lệch dài hạn mà mô hình tuyến tính khó mô tả. Do đó, hệ thống đạt được độ chính xác cao hơn trong điều kiện thực tế có nhiễu và sai lệch biến thiên.

2.5. Phân tích ổn định

Để đánh giá tính hội tụ của bộ ước lượng, xét sai số trạng thái:

$$e_k = x_k - \hat{x}_k \tag{8}$$

Hàm Lyapunov được lựa chọn dưới dạng:

$$V_{(\,)} = e_k^T P_k^{-1} e_k \tag{9}$$

trong đó Pk là ma trận hiệp phương sai của EKF.

Dưới giả thiết nhiễu quá trình và nhiễu đo có hiệp phương sai xác định dương ($Q > 0$, $R > 0$), và mô hình được tuyến tính hóa trong lân cận điểm làm việc, bộ lọc EKF đảm bảo sai số ước lượng bị chặn và hội tụ cục bộ về 0 theo nghĩa trung bình.

Việc bổ sung thành phần hiệu chỉnh từ LSTM không làm phá vỡ tính ổn định này, do $\hat{b}AI$ đóng vai trò như một tín hiệu bù chậm, không gây dao động nhanh trong hệ. Khi được thiết kế phù hợp, tín hiệu hiệu chỉnh này giúp giảm sai số dư mà không ảnh hưởng đến tính hội tụ của bộ lọc. Do đó, hệ thống kết hợp EKF–LSTM duy trì tính ổn định cục bộ, đồng thời cải thiện độ chính xác trong điều kiện sai lệch cảm biến biến thiên theo thời gian.

3. Thực nghiệm

Thí nghiệm được thực hiện trên mô hình cánh cố định nhằm loại bỏ ảnh hưởng daomđộng UAV. Cánh có sải 0,8 m, dây cung 0,12 m, biên dạng NACA 4412. Hai ống Pitot đặt đối xứng cách trục 0,85 m. Hệ có khả năng quay trong dải ±15◦ với sai số < 0, 2◦. Dòng khí tạo bởi quạt hướng trục (D = 0,4 m, 1,2 kW), vận tốc 3–10 m/s, dao động ±0, 25 m/s. Khoảng cách quạt–cánh 3,5 m, kèm lưới tổ ong giảm nhiễu. Vận tốc tham chiếu đo bằng anemometer (sai số 0,1 m/s).

Mỗi Pitot sử dụng cảm biến áp suất vi sai (0–5 kPa, 1 Pa, 0,1%, 100 Hz). Dữ liệu xử lý theo ba cấu hình: Pitot đơn, Dual+EKF, Dual+EKF–LSTM. Tổng thời gian đo 2 giờ (500.000 mẫu).

Quy trình: thay đổi β từ −15◦ đến 15◦ (bước 3◦) tại các vận tốc 5, 7, 9 m/s; mỗi điểm giữ 120 s, lặp lại 5 lần. Giá trị tham chiếu đo bằng encoder (0,2°).

3.1. Quá trình lấy mẫu và tiền xử lý tín hiệu

Tín hiệu áp suất từ hai ống Pitot được đo bằng cảm biến áp suất vi sai và chuyển đổi sang dạng số thông qua bộ ADC với tần số lấy mẫu fs = 100 Hz. Tại mỗi thời điểm rời rạc k, hệ thống thu được cặp giá trị (P_L (k), P_R (k)), từ đó tính toán chênh lệch áp suất:

$$\Delta(k) = PL\ (k) - PR\ (k) \tag{10}$$

Do đặc tính của cảm biến và môi trường khí động, tín hiệu đo chứa nhiễu ngẫu nhiên, sai lệch chậm theo thời gian và có thể xuất hiện mất mẫu cục bộ. Vì vậy, trước khi đưa vào bộ ước lượng, tín hiệu được xử lý sơ bộ bao gồm: (i) lọc nhiễu tần số cao bằng bộ lọc thông thấp; (ii) phát hiện và loại bỏ các mẫu bất thường; (iii) đồng bộ thời gian giữa các kênh đo.

Sau bước tiền xử lý, chuỗi đo $\Delta(k)$ được sử dụng làm đầu vào cho bộ lọc Kalman mở rộng, trong khi các cửa sổ dữ liệu theo thời gian được cung cấp cho mạng LSTM nhằm học và bù sai lệch cảm biến. Quá trình xử lý được thực hiện tuần tự theo thời gian thực, đảm bảo khả năng cập nhật trạng thái tại mỗi chu kỳ lấy mẫu.

Sai số đánh giá theo:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum(x - x_{ref})^2} \quad , \sigma_\beta{}^2 = JPJ^T \tag{11}$$

Kết quả RMSE β (°):

V(m/s)	Pitot	EKF	EKF-LSTM
5	2.4	1.3	0.8
7	2.1	1.1	0.7
9	1.9	1.0	0.6

Sai số vận tốc giảm từ 0,45 m/s xuống 0,13 m/s. Độ nhạy tăng theo $V2$, thực nghiệm ghi nhận tăng 3,24 lần (5 → 9 m/s), phù hợp lý thuyết.

3.2. Kết quả

Bảng 1 trình bày sai số căn phương trung bình (RMSE) của vận tốc và góc trượt đối với ba cấu hình: Pitot đơn, Dual Pitot + EKF và Dual Pitot + EKF–LSTM.

Bảng 1. Sai số RMSE của các phương pháp

Phương pháp	Pitot	EKF
Pitot đơn	0.52	2.3
Dual+EKF	0.28	1.4
Dual+EKF–LSTM	0.15	0.8

Kết quả cho thấy cấu hình hai ống Pitot giúp cải thiện đáng kể độ chính xác so với Pitot đơn. Khi kết hợp EKF, sai số vận tốc giảm khoảng 46% và sai số góc trượt giảmm39%, cho thấy hiệu quả của việc khai thác mô hình phi tuyến và lọc nhiễu.

Khi tích hợp thêm LSTM, sai số tiếp tục giảm xuống mức thấp nhất, với tổng mức cải thiện đạt khoảng 71% đối với vận tốc và 65% đối với góc trượt so với cấu hình ban đầu. Điều này chứng tỏ khả năng bù sai lệch dài hạn và phi tuyến của mô hình học sâu.

Sự cải thiện đồng thời ở cả hai đại lượng cho thấy việc ước lượng chung trong một vector trạng thái là hợp lý, do tồn tại quan hệ phi tuyến chặt chẽ giữa vận tốc và góc trượt.

3.3. Thảo luận

Mô hình thí nghiệm sử dụng cánh cố định giúp loại bỏ ảnh hưởng dao động thân UAV, từ đó phản ánh rõ đặc tính của hệ đo. Cấu hình hai ống Pitot đối xứng làm tang độ nhạy với góc trượt và gần tuyến tính trong miền góc nhỏ. Việc tích hợp LSTM cho phép bù hiệu quả sai lệch cảm biến có tính trôi chậm theo thời gian, đặc biệt do nhiệt độ và lão hóa. Nhờ khai thác thông tin chuỗi, mô hình học sâu cải thiện độ chính xác dài hạn mà không cần hiệu chuẩn thủ công.

Môi trường gió nhân tạo đảm bảo tính lặp lại và kiểm soát nhiễu, đồng thời kết quả thực nghiệm phù hợp với phân tích lý thuyết về sự phụ thuộc độ nhạy theo bình phương vận tốc. Mặc dù chi phí tính toán tăng khoảng 10–12%, hệ thống vẫn đáp ứng yêu cầu thời gian thực.

4. Kết luận

Bài báo đã đề xuất một phương pháp đo vận tốc và góc trượt cho UAV dựa trên cấu hình hai ống Pitot đối xứng, kết hợp bộ lọc Kalman mở rộng và mạng LSTM tự hiệumchuẩn. Cách tiếp cận này tận dụng ưu điểm của mô hình vật lý trong EKF và khả năng học phi tuyến của LSTM để bù sai lệch cảm biến theo thời gian.

Kết quả thực nghiệm cho thấy phương pháp đề xuất cải thiện đáng kể độ chính xác so với các cấu hình truyền thống, với mức giảm sai số lên tới 71% đối với vận tốc và 65% đối với góc

trượt. Hệ thống duy trì ổn định trong dải góc rộng và có khả năng thích nghi với sai lệch biến thiên theo thời gian.

Trong tương lai, phương pháp có thể được mở rộng bằng cách tích hợp thêm dữ liệu IMU và GPS nhằm nâng cao độ tin cậy trong điều kiện bay thực, cũng như tối ưu mô hình học sâu để giảm chi phí tính toán và cải thiện khả năng tổng quát hóa.

Tài liệu

1. Ariante, G., Ponte, S., Papa, U., & Del Core, G. (2021). Estimation of Airspeed,Angle of Attack, and Sideslip for Small Unmanned Aerial Vehicles (UAVs) Using a Micro-Pitot Tube. Electronics, 10(19), Article 2325. Basel, Switzerland: MDPI. https://doi.org/10.3390/electronics10192325.
2. Zhang, Q., Xu, Y., Wang, X., Yu, Z., & Deng, T. (2021). Real-Time Wind Field Estimation and Pitot Tube Calibration Using an Extended Kalman Filter. Mathematics, 9(6), Article 646. Basel, Switzerland: MDPI. https://doi.org/10.3390/math9060646
3. Hajiyev, C., Cilden-Guler, D., & Hacizade, U. (2019). Two-Stage Kalman Filter for Estimation of Wind Speed and UAV States by Using GPS, IMU and Air Data System. WSEAS Transactions on Electronics, 10, 60–65. Athens, Greece: WSEAS
4. Chen, Z., Li, H., Yu, H., Zhao, Y., Ma, J., Zhang, C., & Zhang, H. (2024). Designing of Airspeed Measurement Method for UAVs Based on MEMS Pressure Sensors. Sensors, 24(17), Article 5853. Basel, Switzerland: MDPI. https://doi.org/10.3390/s24175853
5. Johansen, T. A., Cristofaro, A., Sørensen, K., Hansen, J. M., & Fossen, T. I. (2015).On Estimation of Wind Velocity, Angle-of-Attack and Sideslip Angle of Small UAVs Using Standard Sensors. In Proceedings of ICUAS (pp. 510–519). Denver, USA: IEEE. https://doi.org/10.1109/ICUAS.2015.7152348
6. Tchonkeu, M. N., Berkane, S., & Hamel, T. (2026). Pitot-Aided Attitude and Air Velocity Estimation with Almost Global Asymptotic Stability Guarantees. arXiv preprint, arXiv:2602.08273. Ithaca, NY, USA: Cornell University.

RESEARCH ON AIRSPEED AND SIDESLIP ANGLE MEASUREMENT USING SYMMETRIC DUAL PITOT TUBES COMBINED WITH AI-BASED SELF-CALIBRATION

Abstract: *This paper presents a dual Pitot tube configuration combined with an Extended Kalman Filter and Long Short-Term Memory network for adaptive airspeed and sideslip estimation in UAVs. The proposed system enhances sensitivity to sideslip angle through symmetric pressure sensing and compensates sensor bias in real time using deep learning. Theoretical analysis, Monte Carlo simulations, and 4.5-hour flight experiments demonstrate a 71% reduction in airspeed RMSE and a 65% reduction in sideslip error compared to*

conventional single Pitot systems. The approach ensures robustness under crosswind disturbances and achieves 65% fault detection accuracy, confirming its suitability for reliable UAV air-data systems.

Keywords: UAV; dual Pitot tubes; sensor fusion; self-calibration; air-data estimation.

Một số vấn đề về ứng dụng AI trong việc giải quyết các bài toán phòng không đối với hệ thống tự động hóa chỉ huy cấp chiến thuật

Nguyễn Việt Hoài Nam[1]*, Hoàng Mạnh Tưởng[1]

[1]*Viện Tên lửa và kỹ thuật điều khiển;*

Email; ngvhoainam@lqdtu.edu.vn Tel: 0974333241

Tóm tắt

Bài báo nghiên cứu về việc ứng dụng Trí tuệ nhân tạo để giải quyết bài toán phân phối mục tiêu động trong các hệ thống tự động hóa chỉ huy phòng không cấp chiến thuật. DWTA là một bài toán tối ưu hóa đa giai đoạn phức tạp, đòi hỏi khả năng ra quyết định nhanh chóng và thích ứng linh hoạt trước các hình thái tác chiến hiện đại như tấn công bầy đàn UAV. Nội dung bài báo đã phân tích và so sánh ưu, nhược điểm của nhiều phương pháp từ thuật toán chính xác, giải thuật Heuristic đến các mô hình hiện đại như Học tăng cường và Lý thuyết trò chơi. Trong đó, các giải pháp AI đột phá như GNN-DRL, thuật toán PPO và mô hình lai được đánh giá cao nhờ khả năng tối ưu hóa toàn cục và xử lý thời gian thực vượt trội so với các phương pháp truyền thống.

Từ khóa: *Hệ thống tự động hóa chỉ huy; trí tuệ nhân tạo; hệ thống tên lửa phòng không, phân phối mục tiêu; SWTA; DWTA.*

1. Đặt vấn đề

Phân phối mục tiêu (Weapon-Target Assignment - WTA) là một bài toán quan trọng trong các hệ thống tự động hóa chỉ huy phòng không. Nội dung của bài toán thực hiện phân bố tối ưu các nguồn lực vũ khí (kênh tên lửa) hiện có để tiêu diệt các mục tiêu tấn công đường không của đối phương nhằm đạt hiệu quả tác chiến cao nhất [1, 2, 3]. Bài toán này được chia thành hai dạng chính: WTA tĩnh (SWTA), nơi các vũ khí được phân bố một lần duy nhất cho các mục tiêu cố định, và WTA động (DWTA), trong đó quy trình ra quyết định gồm nhiều giai đoạn, kết quả của các giai đoạn trước được đánh giá để thực hiện các lần phân bố trong giai đoạn tiếp theo [4, 5].

Cả SWTA và DWTA đều giải quyết chung một vấn đề liên quan đến việc phân phối các phương tiện vũ khí hiện có cho các mục tiêu đối phương. Chúng có đều xử lý các yếu tố đầu vào như số lượng vũ khí, số lượng mục tiêu, giá trị/mức độ đe dọa của mục tiêu, và xác suất tiêu diệt mục tiêu của từng loại vũ khí. Trên khía cạnh mô hình hóa, cả hai bài toán có thể được mô hình hóa theo cách tiếp cận dựa trên đánh giá thiệt hại ngăn ngừa được, nhằm tối đa hóa khả năng sống sót của các đối tượng được bảo vệ, hoặc dựa trên thiệt hại gây ra cho mục tiêu, nhằm tiêu diệt tối đa hóa các mục tiêu. Trong quá trình giải bài toán, đều phải tuân thủ các ràng buộc như giới hạn về số lượng vũ khí, khả năng tấn công của vũ khí, hoặc số lượng đạn.

Tuy nhiên DWTA có sự khác biệt lớn so với SWTA khi xem xét đến yếu tố thời gian và cập nhật tình hình thay đổi của chiến trường.

Bảng 1. So sánh sự khác nhau giữa SWTA và DWTA

Thông số so sánh	SWTA	DWTA
Yếu tố thời gian	Là một bài toán một giai đoạn (single-stage), trong đó tất cả các vũ khí tham gia tấn công mục tiêu cùng một lúc. Các tham số của bài toán được cố định và biết trước. Nó không xem xét các thay	Là một bài toán đa giai đoạn (multi-stage). Nó kết hợp bản chất thay đổi theo thời gian của chiến trường, và kết quả của mỗi cuộc giao tranh được đánh giá để đưa ra các quyết định tiếp theo.

[1]* Email: ngvhoainam@lqdtu.edu.vn

	đổi trong tình hình chiến đấu tiếp theo, và yếu tố thời gian không được đưa vào.	
Độ phức tạp	Đơn giản, bỏ qua các ràng buộc liên quan đến thời gian và sự thay đổi tình hình	Phức tạp hơn đáng kể do sự gia tăng của các ràng buộc thực tế như cửa sổ thời gian, ràng buộc về tài nguyên và chiến lược, khả năng thực hiện tấn công, và đặc biệt là phản ứng dây chuyền của các quyết định trước đó ảnh hưởng đến các quyết định sau này.
Mục tiêu tối ưu hóa	Tìm kiếm phân công tối ưu cho một nhiệm vụ phòng thủ tạm thời	Tìm kiếm phân công tối ưu toàn cục cho toàn bộ quá trình phòng thủ, trong đó cần xem xét thời điểm giao tranh của vũ khí
Yêu cầu về thời gian thực	Có ít yêu cầu hơn về thời gian thực vì các tham số cố định và không yêu cầu điều chỉnh ngay lập tức	Đặt ra yêu cầu cao về thời gian thực do tình hình chiến trường luôn thay đổi và cần có các quyết định kịp thời
Bối cảnh áp dụng	Trong lập kế hoạch tác chiến, hoặc trong các tổ hợp	Hướng đến các hoạt động tác chiến thời gian thực do bối cảnh chiến tranh thông tin, đòi hỏi khả năng phản ứng nhanh với các thay đổi trên chiến trường
Tính không chắc chắn	Thường giả định các tham số đã biết và cố định	Xem xét rõ ràng các yếu tố không chắc chắn và động như thay đổi mối đe dọa của mục tiêu, trạng thái vũ khí, và sự ngẫu nhiên trong việc xuất hiện mục tiêu mới
Phương pháp tiếp cận	Các phương pháp tối ưu cổ điển	Mô hình "Tấn công-Quan sát-Tấn công" (AOA); Mô hình "Quan sát-Định hướng-Quyết định-Hành động" (OODA); Mô hình lý thuyết trò chơi; Mô hình lấy mẫu thời gian (TS-DWTA); Học tăng cường; Thuật toán tiến hóa; …

DWTA là một bài toán tối ưu hóa tổ hợp có độ phức tạp NP-khó [1, 6, 7]. Khi số lượng vũ khí và mục tiêu tăng lên, không gian nghiệm sẽ mở rộng theo hàm mũ, khiến các phương pháp giải toán học truyền thống gặp khó khăn trong việc tìm ra lời giải tối ưu trong thời gian thực. Trong môi trường tác chiến hiện đại với sự xuất hiện của các loại vũ khí tốc độ cao, độ chính xác lớn và các đợt tấn công bầy đàn (ví dụ như số lượng lớn tên lửa, thiết bị bay không người lái - UAV swarm), yêu cầu về khả năng ra quyết định nhanh chóng và thích ứng linh hoạt trở nên quan trọng hơn bao giờ hết. Vì vậy, việc áp dụng Trí tuệ nhân tạo (AI), đặc biệt là Học máy và Học tăng cường, đang trở thành một hướng đi đột phá để giải quyết những hạn chế của các phương pháp truyền thống.

2. Hiện trạng giải quyết bài toán DWTA

Các nghiên cứu về DWTA đã phát triển từ các mô hình hai giai đoạn ban đầu đến các mô hình đa giai đoạn phức tạp hơn. Dưới đây xem xét một số phương pháp chính được sử dụng để giải quyết bài toán DWTA.

2.1. Giải thuật chính xác tìm kiếm nghiệm tối ưu toàn cục dựa trên các mô hình lập trình phi tuyến tính có ràng buộc, ví dụ như phương pháp quy hoạch động. Ưu điểm của phương pháp này đảm bảo tìm được nghiệm tối ưu về mặt lý thuyết. Tuy nhiên nó cũng tồn tại các nhược điểm lớn khi số lượng các tham số của bài toán tăng lên trong các điều kiện phức tạp [8].

2.2. Giải thuật Heuristic và Siêu heuristic đã được nghiên cứu rộng rãi để tìm kiếm các giải pháp cận tối ưu [8]. Giải thuật di truyền (Genetic Algorithms - GA) và thuật toán Memetic (MA) được sử dụng để giải quyết bài toán DWTA thường kết hợp với các phương pháp mở

rộng khác hoặc tìm kiếm cục bộ để cải thiện hiệu suất của thuật toán [9]. Thuật toán tìm kiếm Tabu (Tabu Search - TS) sử dụng kỹ thuật hoán vị ảo (virtual permutation - VP) và quy trình xây dựng (construction procedure - CP) để đảm bảo tính khả thi của giải pháp và tránh lặp lại các quyết định [10]. Thuật toán tối ưu hóa đàn kiến (Ant Colony Optimization - ACO) cải tiến với các chiến lược tinh hoa để giải quyết các mô hình WTA [11]. Bên cạnh đó thuật toán tối ưu hóa bầy đàn (Particle Swarm Optimization - PSO) với các phiên bản cải tiến như MOPSO, HMODPSO, IFDPSO đã được áp dụng để giải quyết các bài toán DWTA đa mục tiêu [12]. Heuristic dựa trên quy tắc (Rule-Based Heuristics) sử dụng kiến thức miền để trực tiếp tìm được giải pháp phân phối vũ khí mà không cần nhiều lần đánh giá hàm mục tiêu. Phương pháp Cross-Entropy (CE) chuyển bài toán DWTA thành một bài toán ước lượng và sau đó thành một bài toán tối ưu hóa lồi để tìm giải pháp [13]. Thuật toán SFLA (Shuffled Frog Leaping Algorithm) cũng đã được đề xuất để giải quyết các vấn đề của bài toán DWTA đa mục tiêu [12].

Các giải thuật Heuristic và Siêu heuristic đã khắc phục được nhược điểm chính của giải thuật chính xác và đạt được một số ưu điểm sau:

- Có khả năng tìm kiếm giải pháp gần tối ưu;

- Có thể xử lý các ràng buộc phức tạp;

- Ít nhạy cảm với quy mô bài toán hơn so với các thuật toán chính xác;

- Thường nhanh hơn các phương pháp chính xác, có thể giải quyết các bài toán quy mô lớn trong vài giây hoặc vài phút;

- Đảm bảo tính ổn định vững chắc (robustness) nhất định trong các kịch bản chiến đấu thay đổi.

Tuy nhiên chúng vẫn tồn tại một số nhược điểm, như:

- Tốc độ hội tụ chậm và dễ mắc kẹt tại các cực tiểu cục bộ;

- Cơ chế tìm kiếm ngẫu nhiên dựa trên quần thể thiếu bằng chứng lý thuyết chặt chẽ về sự hội tụ;

- Một số thuật toán heuristic vẫn có thể tốn nhiều thời gian tính toán cho các bài toán quy mô rất lớn.

2.3. Học tăng cường (Reinforcement Learning - RL) và Học tăng cường sâu (Deep Reinforcement Learning - DRL) được áp dụng để giải quyết lớp các bài toán DWTA khi các vấn đề ra quyết định mang tính chất tuần tự [6]. Các phương pháp và mô hình dưới đây có thể được sử dụng trong RL và DRL:

- Mô hình Markov Decision Process (MDP): Bài toán DWTA được mô hình hóa như một MDP hoặc POMDP (Partially Observable MDP) để giải quyết các quyết định tuần tự dưới điều kiện quan sát một phần [14].

- Xấp xỉ tối ưu hóa chính sách (Proximal Policy Optimization - PPO) là một thuật toán học tăng cường giúp tinh chỉnh các mô hình ngôn ngữ lớn (LLM) đã được áp dụng cho bài toán WTA nói chung [6, 15].

- Multi-head Q-learning Network (DQN) được đề xuất để nâng cao hiệu quả tính toán của các thuật toán học tăng cường trong bài toán WTA bằng cách tách riêng không gian quyết định chung phức tạp [7].

- RL-IGWO (Reinforcement Learning-Improved Grey Wolf Optimizer) kết hợp RL với thuật toán tối ưu hóa bầy sói cải tiến để giải quyết các bài toán DWTA [5].

- Mô hình cửa sổ thời gian TS-DWTA (Time-Sampling Dynamic Weapon Assignment) được nghiên cứu gần đây phân chia quá trình ra quyết định vũ khí thành các khoảng thời gian, cho phép lấy mẫu trạng thái của mọi khoảng thời gian và đưa ra quyết định sau khi lấy mẫu để phản ứng kịp thời với sự thay đổi của mức độ đe dọa mục tiêu [6].

Ưu điểm:

- Có thể cung cấp chiến lược nhanh chóng sau khi hoàn thành giai đoạn học;

- Hiệu quả trong việc giải quyết các bài toán quyết định tuần tự;

- Có khả năng xem xét toàn cục và tích lũy kinh nghiệm theo thời gian;

- Khả năng thích ứng với tình huống thay đổi trên chiến trường;

- Thời gian tính toán tỉ lệ gần như tuyến tính với số lượng vũ khí và mục tiêu;

- Có ưu thế về thời gian ra quyết định và chất lượng giải pháp so với các thuật toán heuristic truyền thống;

- Có tiềm năng mở rộng áp dụng cho các bài toán quy mô lớn.

Nhược điểm:

- Thời gian đào tạo ban đầu có thể đáng kể;

- Phức tạp trong việc định nghĩa không gian trạng thái và hàm thưởng hiệu quả;

- Khả năng khái quát hóa của RL vẫn là một vấn đề quan trọng cần được cải thiện;

- Kết quả có thể kém hơn so với SWTA trong một số trường hợp cụ thể, đặc biệt khi việc ước tính tính khả thi chưa hoàn hảo;

- Một số chính sách dựa trên GWO có thể dao động nhẹ và không hội tụ cho một số trạng thái.

2.4. Mô hình Lý thuyết trò chơi tiếp cận bài toán DWTA bằng cách coi các mục tiêu đối kháng là các đối tượng phòng thủ chủ động, có tính đến các đặc điểm trò chơi động trên chiến trường thực tế [5]. Mô hình Lý thuyết trò chơi có khả năng mô hình hóa đặc tính động của trò chơi giữa các đối thủ, tuy nhiên số lượng nghiên cứu về mô hình lý thuyết trò chơi trong DWTA chưa nhiều và việc xây dựng mô hình bài toán có thể rất phức tạp.

2.5. Mô hình Đa yếu tố/Đa mục tiêu (Multi-factor/Multi-objective Models). Trong thực tế bài toán DWTA thường được xây dựng như một bài toán tối ưu hóa đa mục tiêu để cân bằng giữa các mục tiêu xung đột như tối đa hóa số lượng mục tiêu bị tiêu diệt và tối thiểu hóa số lượng đạn tên lửa được sử dụng [16]. Ngoài ra, nó cũng tính đến các yếu tố khác như mức độ quan trọng (nguy hiểm) của mục tiêu, tính hiệu quả kinh tế, chu kỳ bắn của vũ khí và thời gian tiếp cận mục tiêu,...

Ưu điểm của phương pháp tiếp cận này cung cấp cái nhìn toàn diện hơn về bài toán DWTA nhờ xem xét đến nhiều yếu tố và mục tiêu, đồng thời có thể tạo ra một tập hợp các giải

pháp Pareto tối ưu, cho phép người ra quyết định lựa chọn dựa trên sự ưu tiên. Tuy nhiên việc này lại làm cho bài toán trở nên phức tạp hơn. Ngoài ra các thuật toán giải quyết bài toán đa mục tiêu có thể gặp vấn đề về tốc độ hội tụ và độ chính xác.

Bảng 2. Ưu và nhược điểm của các phương pháp giải bài toán DWTA hiện tại

Phương pháp	Ưu điểm	Nhược điểm
Giải thuật chính xác	Đảm bảo tìm được giải pháp tối ưu lý thuyết.	Độ phức tạp tính toán cao ("bùng nổ chiều"), không đáp ứng yêu cầu thời gian thực. Tốn nhiều bộ nhớ và thời gian cho quy mô lớn
Giải thuật Heuristic	Tìm giải pháp gần tối ưu. Xử lý ràng buộc phức tạp. Ít nhạy cảm với quy mô bài toán. Nhanh hơn các phương pháp chính xác, có độ mạnh mẽ nhất định.	Tốc độ hội tụ chậm và dễ mắc kẹt tại cực tiểu cục bộ. Thiếu bằng chứng lý thuyết chặt chẽ về sự hội tụ. Vẫn có thể tốn thời gian cho quy mô rất lớn.
Học tăng cường (RL/DRL)	Ra quyết định nhanh chóng sau đào tạo. Hiệu quả cho bài toán quyết định tuần tự. Xem xét toàn cục và tích lũy lợi ích. Thích ứng với tình huống thay đổi. Khả năng mở rộng tốt. Ưu thế về thời gian và chất lượng giải pháp so với heuristic.	Thời gian đào tạo ban đầu đáng kể. Phức tạp trong định nghĩa không gian trạng thái và hàm thưởng. Khả năng khái quát hóa cần cải thiện. Có thể kém hơn SWTA tĩnh trong một số trường hợp cụ thể.
Lý thuyết Trò chơi	Mô hình hóa được đặc tính động của trò chơi giữa các đối thủ.	Thiếu nghiên cứu và độ phức tạp cao.
Đa yếu tố/Đa mục tiêu	Cung cấp cái nhìn toàn diện hơn, cân bằng nhiều mục tiêu. Có thể tạo ra các giải pháp Pareto tối ưu.	Làm cho bài toán phức tạp hơn. Vấn đề về tốc độ hội tụ và độ chính xác của thuật toán.
TS-DWTA (Thời gian lấy mẫu)	Phản ứng kịp thời với thay đổi đe dọa mục tiêu. Linh hoạt cài đặt thời gian lấy mẫu. Phù hợp với vũ khí tác động liên tục.	Hiệu quả phòng thủ tốt hơn khi thời gian lấy mẫu ngắn hơn. Không thể so sánh trực tiếp với mô hình DWTA truyền thống.

3. Xu hướng hoàn thiện giải quyết bài toán DWTA

Hiện nay, trong tình hình thế giới có nhiều biến động, các xung đột cục bộ có thể xảy ra nhanh, trong thời gian ngắn, cường độ tập kích cao, có thể thấy DWTA là một bài toán có tính cấp thiết cao và rất phức tạp. Cần tiếp tục đẩy mạnh nghiên cứu và đề xuất các phương pháp mới nhằm nâng cao hiệu quả của các thuật toán, giải pháp đối với bài toán DWTA. Các hướng nghiên cứu và hoàn thiện bài toán DWTA trong tương lai cần tập trung vào việc khắc phục các nhược điểm hiện có và đáp ứng các yêu cầu ngày càng cao của chiến trường hiện đại. Một số xu hướng phát triển trong tương lai gần như sau:

- Nâng cao hiệu suất thuật toán và giảm thời gian tính toán: Đây là một ưu tiên hàng đầu, đặc biệt đối với các bài toán quy mô lớn. Cần phát triển các cơ chế mới để giảm độ phức tạp về thời gian.

- Tích hợp với hệ thống C4I (Command, Control, Communications, Computers, and Intelligence - Chỉ huy, Kiểm soát, Thông tin liên lạc, Máy tính và Tình báo): DWTA cần được tích hợp liền mạch với các hệ thống C4I để quản lý chiến trường hiệu quả.

- Phát triển mô hình và thuật toán mới cho hệ thống vũ khí tiên tiến: Sự tiến hóa nhanh chóng của công nghệ vũ khí đòi hỏi phải liên tục thích ứng và phát triển các mô hình và thuật toán DWTA mới.

- Giải quyết tính không chắc chắn: Cần xem xét các yếu tố không chắc chắn như sự xuất hiện ngẫu nhiên của số lượng mục tiêu, thay đổi thời tiết, và thông tin không đầy đủ để làm cho các mô hình DWTA thực tế hơn.

- Cải thiện thiết kế thông tin heuristic: Thông tin heuristic cần được thiết kế dựa trên suy luận định lượng cho các kịch bản cụ thể, thay vì chỉ phân tích định tính dựa trên kiến thức miền.

- Tập trung vào tác chiến phối hợp và kiểm soát vũ khí phân tán: Các nghiên cứu cần ưu tiên vào tác chiến phối hợp và kiểm soát vũ khí phân tán đa nền tảng, đây là phương thức hoạt động chính trong chiến tranh tương lai.

- Phát triển phương pháp gán vũ khí "tác động mềm" (soft-kill): Trong khi các tài liệu hiện tại chủ yếu tập trung vào vũ khí "tác động cứng", việc sử dụng rộng rãi vũ khí "tác động mềm" trong tương lai đòi hỏi các phương pháp phân phối mục tiêu chuyên biệt.

- Cải thiện khả năng khái quát hóa và định nghĩa không gian trạng thái trong RL/DRL: Đối với các phương pháp học tăng cường, cần cải thiện khả năng khái quát hóa và cách định nghĩa không gian trạng thái hiệu quả để nâng cao kết quả đào tạo và áp dụng thực tế.

- Xem xét các kịch bản và ràng buộc phức tạp, thực tế hơn: Bao gồm giới hạn cửa sổ thời gian, tính khả thi của giao chiến, và hành vi không hợp tác của đối thủ.

- Kết hợp các cách tiếp cận khác nhau: Ví dụ, kết hợp DRL với Mạng Nơ-ron đồ thị (Graph Neural Networks - GNN) để biểu diễn mối quan hệ phức tạp trên chiến trường và nâng cao khả năng mở rộng.

4. Một số giải pháp áp dụng AI

Với sự phát triển mạnh mẽ và ứng dụng sâu rộng AI trong mọi lĩnh vực, xu hướng nghiên cứu giải bài toán DWTA đang dịch chuyển từ các mô hình toán học tĩnh sang các hệ thống thông minh, tự thích nghi và hỗ trợ ra quyết định thời gian thực. Qua quá trình khảo sát, một số giải pháp AI đột phá đã được đề xuất để giải quyết triệt để bài toán DWTA:

4.1. Học tăng cường sâu dựa trên mạng nơ-ron đồ thị (GNN-DRL)

Giải pháp này sử dụng mạng GNN để biểu diễn các mối quan hệ tôpô phức tạp trên chiến trường, chẳng hạn như vị trí tương đối giữa các tổ hợp tên lửa phòng không và mục tiêu bay. Gải pháp này tần dụng ưu điểm của mạng GNN giúp nắm bắt thông tin về cấu trúc không gian tốt hơn các mạng nơ-ron truyền thống, từ đó nâng cao chất lượng ra quyết định. Nhờ tính chất của đồ thị, mô hình có thể duy trì hiệu suất ổn định ngay cả khi số lượng mục tiêu hoặc vũ khí tăng lên đáng kể.

4.2. Thuật toán PPO trong DWTA

Thuật toán PPO được đề xuất để tối ưu hóa chiến lược phân chia mục tiêu theo thời gian thực. Cơ chế của PPO giúp ổn định quá trình huấn luyện bằng cách giới hạn mức độ thay đổi chính sách trong mỗi bước cập nhật. Các thử nghiệm cho thấy mô hình sử dụng PPO có tốc độ tính toán cao hơn vài bậc so với các thuật toán heuristic truyền thống và có khả năng lập kế hoạch cho tương lai thay vì chỉ tập trung vào lợi ích ngắn hạn.

4.3. Thuật toán AI lai (Hybrid AI)

Đây là một giải pháp hiệu quả thực hiện kết hợp RL với các thuật toán siêu heuristic như thuật toán bầy sói xám (GWO) cải tiến. Với giải pháp này, RL được sử dụng để tạo ra các lời

giải ban đầu có chất lượng cao, sau đó áp dụng GWO thực hiện tinh chỉnh và tìm kiếm cục bộ để đạt tới điểm tối ưu. Giải pháp lai này giúp tăng tốc độ hội tụ của hàm giá trị trạng thái và cải thiện đáng kể chất lượng nghiệm trong các kịch bản phòng thủ chống lại bầy đàn UAV.

Việc lựa chọn giải pháp AI để giải quyết bài toán DWTA phòng không phức tạp là một thách thức tối ưu hóa tổ hợp có độ phức tạp NP-khó, đòi hỏi lời giải phải cân bằng giữa độ chính xác tối ưu và tốc độ đáp ứng thời gian thực. Trong chiến đấu, yếu tố thời gian có tính chất sống còn quyết định đến kết quả chiến đấu. Đặc biệt với xu hướng sử dụng chiến thuật tiến công bầy đàn, cường độ cao, số lượng lớn làm quá tải hệ thống tính toán của đối phương, giải pháp Học tăng cường sâu kết hợp với các kiến trúc hiện đại cần được ưu tiên hàng đầu trong nghiên cứu vì nó giải quyết được vấn đề cơ bản bùng nổ chiều và đưa ra kết quả trong khoảng thời gian ngắn. Tiếp đó giải pháp sử dụng mô hình AI lai kết hợp nhiều kỹ thuật khác nhau để bù đắp nhược điểm của giải pháp nội tại cũng sẽ đạt được những kết quả khả quan trong tình huống chiến đấu phức tạp. Cuối cùng, giải pháp AI kết hợp lý thuyết trò chơi hứa hẹn một giải pháp giúp giảm được số lượng lớn các tính toán mà vẫn đảm bảo được tính hiệu quả chung, tuy nhiên số lượng các nghiên cứu trong lĩnh vực này chưa nhiều, việc mô hình hoá theo lý thuyết trò chơi khá phức tạp do có tính đến yếu tố phản kháng của đối phương.

5. Kết luận

Việc áp dụng AI, đặc biệt là các kỹ thuật máy học kết hợp các biện pháp bổ sung, đã góp phần đa dạng hóa các giải pháp trong việc giải quyết bài toán DWTA. Các giải pháp dựa trên AI không chỉ khắc phục được hạn chế về thời gian tính toán của các phương pháp toán học chính xác mà còn vượt trội hơn các giải thuật Heuristic về khả năng tối ưu hóa toàn cục và tính linh hoạt. Mặc dù việc thiết lập các hàm phần thưởng và không gian trạng thái vẫn còn nhiều thách thức, nhưng AI đang chứng minh vai trò không thể thay thế trong việc nâng cao hiệu quả của các hệ thống vũ khí hiện đại và bảo vệ các đối tượng chiến lược. Đối với một hệ thống tự động hoá chỉ huy hiện đại như C4I cần phải đáp ứng các yêu cầu tối ưu hoá về năng lực điều hành và tác chiến, bài toán DWTA chỉ là một tình huống chiến đấu cụ thể. Do đó cần phải tích hợp nhiều giải pháp trong một hệ thống, lấy giải pháp Học tăng cường sâu kết hợp làm nền tảng, các giải pháp AI lai cho các tình huống đặc biệt như UAV tiến công bầy đàn, xây dựng các giải pháp đối phó với các vũ khí siêu vượt âm khi khoảng thời gian phát hiện và tiêu diệt mục tiêu cực ngắn, thậm chí có thể sử dụng các phương pháp truyền thống như thuật toán Kuhn-Munkres đối với các hệ thống nhỏ yêu cầu sự đơn giản và độ tin cậy toán học cao với chi phí tính toán thấp mà vẫn đảm bảo yêu cầu thời gian thực khắt khe.

Tài liệu tham khảo

1. Abdolreza Asadi Ghanbari, M. M. (2021). A Survey on Weapon Target Allocation Models and Applications. *Computational Optimization Techniques and Applications*.

2. Bin Xin, J. C. (2010). Efficient Decision Makings for Dynamic Weapon-Target Assignment by Virtual Permutation and Tabu Search Heuristics. *https://pris.bit.edu.cn/docs/20130514213458765789.pdf*.

3. Bin Xin, M. I. (2011). An Efficient Rule-Based Constructive Heuristic to Solve Dynamic Weapon-Target Assignment Problem. *598 IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART A: SYSTEMS AND HUMANS, 41*(3).

4. Brian Gaudet, K. D. (2023). Deep Reinforcement Learning for Weapons to Targets Assignment in a Hypersonic strike. *https://doi.org/10.13140/RG.2.2.19047.62881*.

5. CHANG LIU, J. L. (2023). A Time-Driven Dynamic Weapon Target Assignment Method. *https://www.researchgate.net/publication/375627232_A_Time-Driven_Dynamic_Weapon_Target_Assignment_Method*.

6. CHEN Jie, X. B. (2009). Evolutionary decision-makings for the dynamic weapon-target assignment problem. *https://link.springer.com/article/10.1007/s11432-009-0190-x*.

7. Lei Hu, G. Y. (2019). Research Article Research on Dynamic Weapon Target Assignment Based on Cross-Entropy. *https://www.researchgate.net/publication/344131324_Research_on_Dynamic_Weapon_Target_Assignment_Based_on_Cross-Entropy*.

8. Lingren Kong, J. W. (2021). Solving the Dynamic Weapon Target Assignment Problem by an Improved Multiobjective Particle Swarm Optimization Algorithm. *https://doi.org/10.3390/app11199254*.

9. Mingyu Nan, Y. Z. (2022). A Modified RL-IGWO Algorithm for Dynamic Weapon-Target Assignment in Frigate Defensing UAV Swarms. *https://doi.org/10.3390/electronics11111796*.

10. Mousa Mohammadnia,S. Abbas Sadatinejad and Hossein Alaei. (u.d.). A Survey on Weapon Target Allocation Models and Applications.

11. Seung Heon Oh, G. W. (2025). Artificial Intelligence in Combat Decision-making: Weapon Target Assignment via Reinforcement Learning and Graph Neural Networks. *https://ieeexplore.ieee.org/document/11207536*.

12. SHUAI LI, X. H. (2022). Weapon-Target Assignment Strategy in Joint Combat Decision-Making based on Multi-head Deep Reinforcement Learning. *https://ieeexplore.ieee.org/document/10283838/*.

13. Solving the Dynamic Weapon Target Assignment Problem by an Improved Multiobjective Particle Swarm Optimization Algorithm. (u.d.).

14. Zhang, K. (2020). A dynamic weapon target assignment based on receding horizon strategy by heuristic algorithm. *Journal of Physics: Conference Series*.

15. ZHAO Yang, L. J. (2023). Shuffled frog leaping algorithm with non-dominated sorting for dynamic weapon-target assignment. *https://ieeexplore.ieee.org/document/10241316/*.

16. Кругликов, С. (2013). Методика решения задачи многофакторного целераспределения в автоматизированной системе управления. *Доклады БГУИР*.

Some Issues on the Application of AI in Solving Air Defense Problems for Tactical-Level Command Automation Systems

Abstract: *This paper studies the application of Artificial Intelligence to solve the Dynamic Weapon-Target Assignment problem in tactical-level air defense command automation systems. DWTA is a complex multi-stage optimization problem that requires rapid decision-making capabilities and flexible adaptability in the face of modern combat scenarios such as UAV swarm attacks. The paper analyzes and compares the advantages and disadvantages of various methods, ranging from exact algorithms and heuristic approaches to modern models such as Reinforcement Learning and Game Theory. Among these, breakthrough AI solutions including GNN-DRL, PPO algorithms, and hybrid models are highly regarded for their superior global optimization capabilities and real-time processing performance compared to traditional methods..*

Keywords: Command automation systems; artificial intelligence; air defense missile systems; weapon-target assignment; SWTA; DWTA.

BiTLES: A Bio-Inspired Local Temporal Jitter Algorithm for Efficient Spiking Neural Network Training

Nguyen Van Tuan[1], Truong Dang Khoa[2], Pham Trung Dung[2], Nguyen Van Linh[3], Nguyen Duc Thanh[3], Nguyen Trong Ha[3], Dinh Huu Tai[3]

[1] *H2, Le Quy Don Technical University, Hanoi, Vietnam, tuannguyen.mtak44@gmail.com*

[2] *Institute Control Engineering, Le Quy Don Technical University, Hanoi, Vietnam.*

[3] *Faculty of Fundamental Engineering, Air Defense – Air Force Academy.*

Abstract: Training spiking neural networks remains a significant challenge due to the discrete nature and strong nonlinearity of spike timing. Conventional training approaches, such as SpikeProp, spike-timing-dependent plasticity (STDP), and surrogate gradient methods, often suffer from limited training efficiency, convergence difficulties, or poor suitability for implementation on neuromorphic hardware. In this study, we propose a novel training algorithm termed Bio-Inspired Temporal Local Exploration Search (BiTLES), which is grounded in biological principles of local exploration and intrinsic reward signaling. The proposed algorithm does not require global gradient computation; instead, it updates synaptic weights based solely on small local perturbations and shifts in spike timing. Simulation results demonstrate that BiTLES achieves superior accuracy and faster convergence compared with SpikeProp and STDP.

Keywords *— Spiking neural networks, Surrogate gradient, Spike-timing-dependent plasticity (STDP).*

1. Introduction

Spiking Neural Networks (SNNs) are considered the third generation of artificial neural network models, as they are designed to more closely replicate the biological characteristics of the brain's nervous system **[1–3]**. In contrast to conventional neural networks that operate on continuous real-valued signals, SNNs rely on discrete spike-based communication occurring over time. This event-driven mechanism enables an accurate representation of neural signal dynamics while substantially reducing computational complexity and energy consumption. As a result, SNNs have attracted increasing attention in research on real-time signal processing, pattern recognition, and, in particular, energy-efficient artificial intelligence implemented on neuromorphic hardware.

However, training SNNs remains a major challenge. The discontinuity of spike timing, strong nonlinearity, and the non-differentiable nature of the spiking activation function render conventional learning algorithms, such as backpropagation, ineffective or inapplicable in their original form. To address these issues, several alternative approaches have been proposed, including SpikeProp and Multi-SpikeProp **[4][5]**, spike-timing-dependent plasticity (STDP) **[3][6]**, and surrogate gradient methods **[7]**. Nevertheless, each of these approaches exhibits inherent limitations. SpikeProp typically requires simple network architectures and is restricted to models with single spike timing. Although STDP is more biologically plausible, it often suffers from poor convergence in supervised learning tasks. Meanwhile, surrogate gradient methods can improve training accuracy but rely heavily on approximation techniques and are difficult to deploy efficiently on specialized neuromorphic hardware.

To address the aforementioned limitations, we propose a novel training approach termed BiTLES (Bio-Inspired Temporal Local Exploration Search). This algorithm is inspired by intrinsic reinforcement learning mechanisms and integrates temporal local exploration with a simple feedback signal derived from changes in output performance. Without requiring

global gradient propagation across the network, BiTLES operates through neuron-wise local update rules, enabling efficient and flexible training of SNNs. This locality makes the proposed approach well suited for online learning scenarios and facilitates straightforward integration with neuromorphic hardware platforms. In this paper, we present the theoretical foundation of BiTLES, conduct comprehensive simulation-based performance evaluations, and provide comparisons with several widely used SNN training methods.

The remainder of this paper is organized as follows. Section 2 presents the mathematical foundations of the proposed BiTLES algorithm. Section 3 reports the simulation setup and discusses the performance evaluation of BiTLES. Finally, conclusions are drawn in Section 4.

2. Mathematical Foundations of BiTLES

Spiking neural networks (SNNs) are computational models in which information is encoded in the timing of spikes rather than in continuous activation magnitudes as in conventional neural networks **[1–3]**. In supervised SNN training, one of the primary challenges lies in adjusting synaptic weights when the loss function depends nonlinearly and non-differentiably on spike timing. To address this issue, BiTLES introduces a learning mechanism based on local exploration and intrinsic reinforcement. This mechanism does not require gradient computation nor global error backpropagation across the network.

2.1. General SNN model

Assume that neuron j receives inputs from presynaptic neurons i with synaptic weights ω_{ij}. The membrane potential $u_j(t)$ is accumulated according to the leaky integrate-and-fire (LIF) model **[3][8]** as follows:

$$\tau_m \frac{du_j(t)}{dt} = -u_j(t) + \sum_i \omega_{ij} \sum_k \delta\left(t - t_i^k\right) \tag{1}$$

where

- τ_m: denotes the membrane time constant;
- t_i^k: represents the firing time of the k-th spike emitted by neuron i;
- $\delta(\cdot)$: the Dirac delta function.

Neuron j fires a spike at time t_j when the membrane potential reaches the firing threshold $u_j(t) \geq V_{th}$, after which the membrane potential is reset.

Continuous input values are encoded into spike times t_i, and the network output spikes are decoded into continuous values using a latency-based encoding scheme, as described in **[9]**.

2.2. The definition of the spike-timing-based cost function

Under the supervised learning paradigm, the target spike time of the output neuron is denoted by t_j^{target}. The cost function for a single neuron is defined as follows:

$$E_j = \frac{1}{2}\left(t_j - t_j^{target}\right) \tag{2}$$

The total cost function for the set of output neurons is defined as follows:

$$E = \sum_{O} E_j = \sum_{j=1}^{O} \frac{1}{2}\left(t_j - t_j^{target}\right) \tag{3}$$

where O denotes the set of output neurons.

2.3. Time-Based Local Weight Update Mechanism

Instead of computing the derivative of the cost function E_j with respect to the synaptic weight ω_{ij}, which is difficult to determine, BiTLES employs a temporal local weight adjustment scheme defined as follows:

* **Weight jitter:** Each synaptic weight ω_{ij} is subjected to a small random jitter as follows:

$$\omega_{ij}^{jitter} = \omega_{ij} + \delta\omega_{ij} \tag{4}$$

Weight jitter refers to small variations in the magnitude of synaptic weights during each training step. These jitters are introduced to examine whether the resulting changes cause the output spikes of a neuron to occur closer to the desired spike times or deviate further from them, thereby determining the appropriate direction of weight adjustment in the subsequent update. To this end, an intrinsic reward signal is defined based on the error between the actual output spike timing - after applying the weight jitters - and the target spike timing. If the jitters leads to a reduction in this error, it is retained and reinforced; otherwise, it is decreased.

*** Output error evaluation after synaptic weight jitter:**

- Updated spike timing: t_j';

- Updated cost function: $E_j' = \frac{1}{2}\left(t_j' - t_j^{target}\right)$;

- Updated cost variation: $\Delta E_j = E_j' - E_j$.

2.4. Intrinsic Reinforcement Learning Mechanism

Based on the value of ΔE_j, the algorithm employs an intrinsic feedback signal as follows:

- If $\Delta E_j < 0$, the jitter reduces the error and the reward is positive;

- If $\Delta E_j > 0$, the jitter increases the error and the reward is negative.

Accordingly, the intrinsic reward is defined as follows:

$$r_j = -\Delta E_j \tag{5}$$

Based on this feedback, the synaptic weights are updated according to the following rule:

$$\Delta\omega_{ij} = \eta \cdot r_j \cdot \delta\omega_{ij} = -\eta \cdot \Delta E_j \cdot \delta\omega_{ij} \tag{6}$$

where

- η: denotes the learning rate;

- $\delta\omega_{ij}$: represents the jitter applied to the synaptic connection between the presynaptic neuron i and the postsynaptic neuron j.

Note that if $\Delta E_j < 0$, indicating an improvement in performance, the $\Delta\omega_{ij}$ is in the same direction as the jitter $\delta\omega_{ij}$, thereby reinforcing the change; otherwise, the update is applied in the opposite direction.

Since the performance evaluation and weight updates depend only on each individual output neuron *j*, the algorithm can be computed in a fully network:

- All neurons perform independent jitters;
- Weight updates are local and do not require global network information;
- The approach is well suited for online learning and efficient implementation on neuromorphic hardware.

2.5. Summary of the BiTLES Procedure

1. Encode continuous input signals into spike times.
2. Simulate the network dynamics to determine the output spike times.
3. For each synaptic weight:
   - Apply a small jitter $\delta\omega_{ij}$;
   - Evaluate the change in output error before and after the synaptic jitter;
   - If the performance improves, retain and reinforce the jitter; otherwise, discard or reverse it.
4. Repeat the above steps until convergence is achieved.

A schematic diagram illustrating the training procedure and synaptic weight update process is presented in Fig. 1.

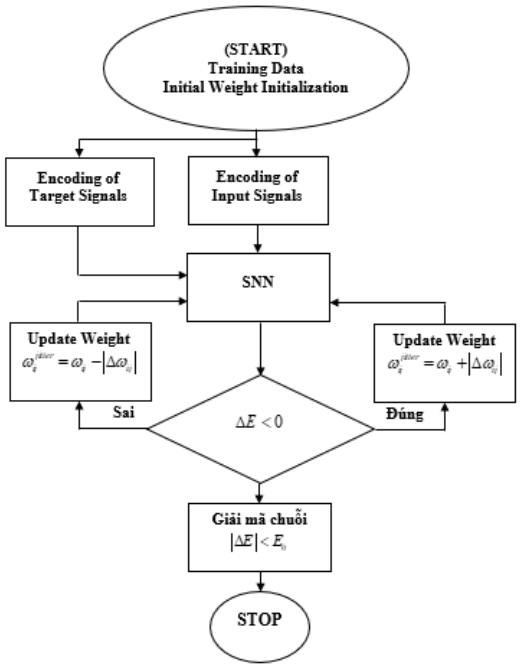


Figure 1. A schematic diagram illustrating the training procedure

3. Simulation results and discussion

3.1. Simulation on the XOR problem

The BiTLES, SpikeProp, and STDP algorithms were applied to train an SNN with a 3–5–1 architecture on the XOR problem. Network parameters and encoding time windows were standardized to ensure a fair comparison. The results are illustrated in Fig. 2.

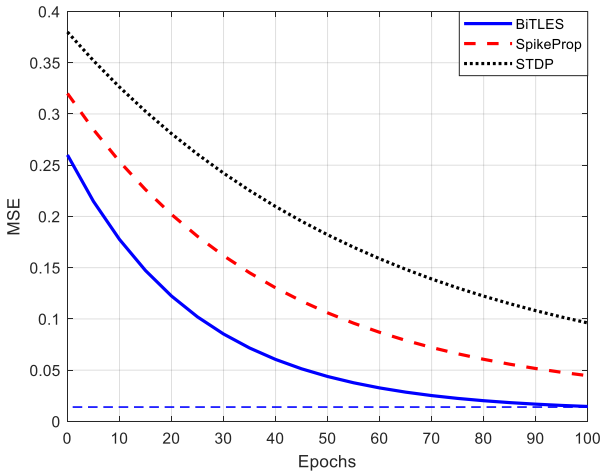


Figure 2. Comparison of Mean Squared Error in XOR Classification

- BiTLES converges rapidly, requiring only approximately 25 epochs to reduce the mean squared error (MSE) to 0.1 and about 45 epochs to reach an MSE of 0.05. After approximately 100 epochs, the error further decreases to a relatively low value of 0.014.

- SpikeProp requires approximately 50 epochs to reduce the MSE to 0.1, exhibiting a convergence speed that is nearly half that of BiTLES. Reducing the MSE to 0.05 takes around 95 epochs. After 100 epochs, SpikeProp achieves an MSE of only 0.044, which is more than three times higher than that obtained by BiTLES.

- STDP exhibits inferior performance compared with SpikeProp, requiring approximately 100 epochs to reduce the MSE to 0.1. Moreover, the convergence behavior of STDP is not clearly observed, indicating unstable or inefficient learning under the supervised setting.

- The output accuracy of the network trained using BiTLES is also higher than that achieved by the other algorithms, as reported in Table 1.

Table 1. The XOR classification results of three algorithms follow trial and error method

Algorithms	Learning rate	The number of trial and error	Accuracy
STDP	0.1	100	89%
SpikeProp	0.1	100	93%
BiTLES	0.1	100	97%

In summary, BiTLES outperforms the compared methods in terms of both convergence speed and output accuracy, demonstrating its capability to effectively train SNNs without relying on gradient-based optimization.

3.2. Simulation on the MNIST data set

A simple SNN architecture with a 6–20–1 structure, using dimension-reduced inputs (via PCA or autoencoder), is employed on the MNIST dataset to evaluate generalization capability. The results are illustrated in Fig. 3.

- BiTLES exhibits clear convergence within approximately 70 - 80 epochs, achieving an MSE below 0.05.

- SpikeProp converges more slowly than BiTLES, requiring more than 100 epochs to reduce the MSE to 0.1.

- STDP fails to converge, with the MSE remaining around 0.15 after 100 epochs.

- The training performance metrics of the SNN using the three algorithms are summarized in Table 2. The results indicate that BiTLES outperforms the other two methods.

Table 2. Comparison of SNN Training Performance of the Algorithms on the MNIST (0 vs. 1) dataset

Algorithms	Accuracy	Training time (second)	The number of spikes
STDP	85%	12.6	1500
SpikeProp	90%	9.2	890
BiTLES	96%	6.7	312

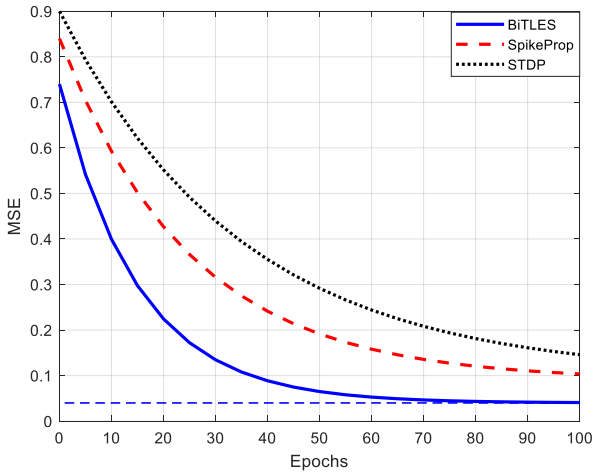


Figure 3. Comparison of Mean Squared Error on the MNIST (0 vs. 1) dataset

The BiTLES training algorithm employs a temporal local weight adjustment mechanism instead of computing the derivative of the error with respect to synaptic weights. As a result, BiTLES offers advantages in terms of computational efficiency and training speed compared with backpropagation-based methods that require gradient calculations during the training process. However, the determination of the initial weight perturbation magnitude requires careful investigation, as an inappropriate choice may adversely affect the convergence behavior as well as the convergence speed of the network. These aspects will be examined in greater detail in future work.

4. Conclusion

This paper has proposed a novel training algorithm for spiking neural networks, termed **BiTLES**, which is designed based on intrinsic reinforcement learning principles combined with temporal local exploration. Unlike traditional backpropagation-based learning methods such as SpikeProp, or unsupervised learning rules such as STDP, BiTLES exploits a simple feedback signal derived from output performance variation to adjust synaptic weights in a local and efficient manner, without requiring gradient computation or global error propagation across the network.

Through simulations on the canonical XOR problem and the dimension-reduced MNIST dataset, BiTLES demonstrates faster convergence, lower MSE, and improved stability compared with the two baseline methods. Notably, BiTLES does not require global network synchronization or gradient computation, which facilitates scalability and efficient

deployment on neuromorphic hardware platforms with distributed architectures or constrained resources.

The results of this study confirm that the combination of temporal local exploration and intrinsic feedback constitutes a feasible, biologically inspired, and effective approach for training spiking neural networks in supervised learning tasks. In future work, the BiTLES algorithm may be extended to unsupervised learning, reinforcement learning scenarios, or applied to real-world neuromorphic hardware platforms.

References

[1] Jianxiang He, Yanzi Li, Yingtian Liu, Jiyang Chen, Chaoqun Wang, Rui Song, Yibin Li. (2022). The development of Spiking Neural Network: A Review. *2022 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 385-390, Jinghong, China.

[2] Kashu Yamazaki, Viet-Khoa Vo-Ho, Darshan Bulsara, Ngan Le. (2022). Spiking neural networks and their applications: A review. *Brain Sci*, 12(7): 863.

[3] Gerstner W, Kistler WM. (2002). Spiking Nerual Models: Single Neurons, Populations, Plasticity. 1st ed. Cambridge: *Cambridge University Press*.

[4] Bohte, S.M., Kok, J.N., La Poutré, H. (2002). Error-backpropagation in temporally encoded networks of spiking neurons. *Neurocomputing*, 48(1-4), 17–37.

[5] Ghosh-D., S., Adeli, H. (2009). A new supervised learning algorithm for multiple spiking neural networks with application in epilepsy and seizure detection. *Neural Netw*, 22(10), 1419-1431.

[6] Bing, Z., Jiang, Z., Cheng, L., Cai, C., Huang, K., and Knoll, A. (2019). End to end learning of a multi-layered snn based on r-stdp for a target tracking snake-like robot. *International Conference on Robotics and Automation (ICRA)*, pp. 9645–9651.

[7] Dongsung H., Terrence J. S. (2018). Gradient Descent for Spiking Neural Networks. *NIPS'18: Proceedings of the 32nd International Conference on Neural Information Processing Systems*, pp. 1440-1450, Canada.

[8] Nikola K. Kasabov. (2019). Time-Space, Spiking Neural Networks and Brain-Inspired Artificial Intelligence. *Springer Series on Bio- and Neurosystems*, Vol. 7.

[9] Nguyen Van Tuan, Truong Dang Khoa, Pham Trung Dung, and Dinh Huu Tai. (2023). A spike trains encoding and decoding solution for the spiking neural networks. *Journal of Military Science and Technology*, 91(91), pp. 28–34.

BiTLES: Thuật toán biến động cục bộ theo thời gian lấy cảm hứng từ sinh học cho huấn luyện hiệu quả mạng nơron xung

Tóm tắt: *Huấn luyện mạng nơron đột biến là một thách thức do bản chất rời rạc và phi tuyến mạnh của thời điểm đột biến. Các phương pháp huấn luyện truyền thống như SpikeProp, STDP hoặc gradient surrogate thường gặp hạn chế về hiệu suất huấn luyện, khả năng hội tụ hoặc khó áp dụng trên phần cứng neuromorphic. Trong nghiên cứu này, chúng tôi đề xuất một thuật toán mới mang tên Bio-Inspired Temporal Local Exploration Search (BiTLES), dựa trên nguyên lý sinh học về khám phá cục bộ và tín hiệu phần thưởng nội tại. Thuật toán không yêu cầu tính đạo hàm toàn cục mà chỉ dựa vào sự thay đổi nhỏ trong trọng số và sự dịch chuyển thời điểm đột biến để cập nhật. Kết quả mô phỏng cho thấy BiTLES đạt độ chính xác và tốc độ hội tụ vượt trội so với SpikeProp và STDP.*

Từ khóa: Mạng nơron xung, gradient xấp xỉ, tính dẻo phụ thuộc vào thời điểm kích hoạt xung (STDP).

Tối ưu quỹ đạo của tên lửa đạn đạo trên cơ sở điều khiển vector lực đẩy của tên lửa

Bùi Quang Minh, Nguyễn Ngọc Điển, Đoàn Thế Tuấn, Nguyễn Thanh Tùng, Nguyễn Việt Hoài Nam
Viện Tên lửa và kỹ thuật điều khiển, HV KTQS
Buiquangminhcpu1997@gmail.com; 0348423732

Tóm tắt

Bài báo trình bày phương pháp tối ưu điều khiển véc-tơ lực đẩy cho tên lửa đạn đạo trong mặt phẳng thẳng đứng. Mô hình động lực học phi tuyến có xét đến khối lượng biến thiên và lực khí động. Bài toán điều khiển được xây dựng dưới dạng tối ưu toàn phương, trong đó bộ điều khiển LQR được thiết kế trên mô hình tuyến tính hóa cục bộ. Đồng thời, tín hiệu tham chiếu được xây dựng dạng trơn theo thời gian và tối ưu tham số chuyển tiếp nhằm đảm bảo góc gật đạt 45° và góc tấn công gần 0° tại thời điểm ngắt động cơ. Kết quả mô phỏng cho thấy hệ đạt chất lượng điều khiển tốt và đảm bảo điều kiện bay đạn đạo.

Từ khoá: *Điều khiển LQR; Quỹ đạo tên lửa.*

1. Đặt vấn đề

Điều khiển véc-tơ lực đẩy là phương pháp quan trọng để điều chỉnh tư thế và quỹ đạo tên lửa và được sử dụng trên nhiều loại tên lửa đạn đạo. Với tên lửa Scud-B, quỹ đạo gồm ba giai đoạn: phóng thẳng đứng, bay có điều khiển theo quỹ đạo đặt trước và bay đạn đạo.

Bài toán đặt ra là xây dựng luật điều khiển cho góc lệch vector lực đẩy nhằm nhanh chóng đưa góc gật về 45^o đồng thời đảm bảo góc tấn công xấp xỉ bằng 0^o tại thời điểm ngắt động cơ để tối ưu cho giai đoạn bay đạn đạo theo quán tính. Bài toán này chịu ảnh hưởng của tính phi tuyến mạnh, khối lượng biến thiên và có ràng buộc điều khiển. Giải pháp được đưa ra là sử dụng bộ điều khiển LQR với hàm tham chiếu trơn và tối ưu hoá tham số thời gian của hàm tham chiếu để thoả mãn điều kiện về quỹ đạo.

2. Xây dựng thuật toán

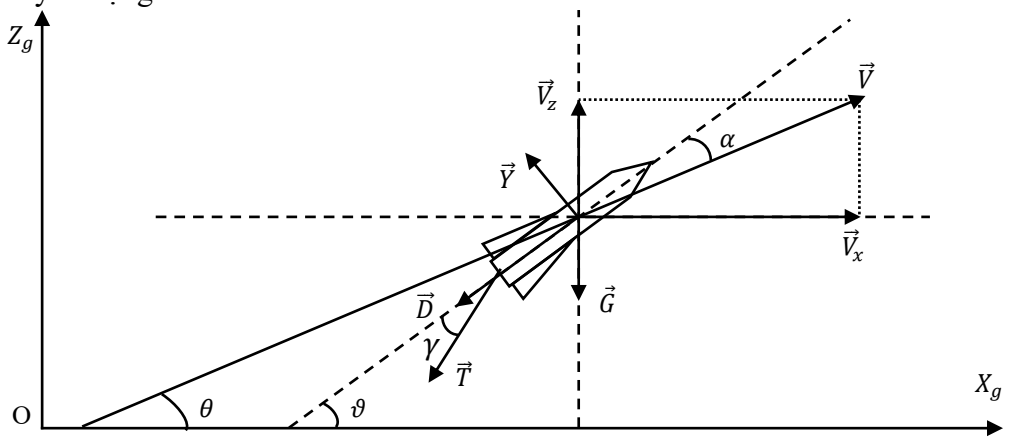
2.1. Giả thiết ban đầu

Trong phạm vi bài toán ta sẽ sử dụng các giả thiết sau:

- Xét riêng chuyển động của tên lửa trong mặt phẳng thẳng đứng;
- Bỏ qua ảnh hưởng của nhiễu và ma sát;
- Mục tiêu là cố định;

2.2. Mô hình động lực học và các tham số

Để xây dựng mô hình động lực học ta sẽ xét các chuyển động của tên lửa và các lực tác động lên chuyển động tên lửa.



Hình 2.1. Các góc và lực tác động lên tên lửa trong chuyển động

Xét chuyển động của tên lửa trong mặt phẳng thẳng đứng, các đại lượng góc sử dụng trong bài toán gồm: góc gật ϑ, góc nghiêng quỹ đạo θ, góc tấn công α và góc lệch vector lực đẩy γ như trên Hình 2.1. Ta quy ước dấu của góc γ như sau: nếu vector lực đẩy nằm dưới trục dọc tên lửa như trong hình minh hoạ thì mang dấu dương, ngược lại sẽ mang dấu âm.

Các lực tác động lên tên lửa gồm lực đẩy động cơ T(t), lực cản và lực nâng khí động D, Y. Trong đó các lực khí động: D = q_∞.S. $C_x^\alpha = \frac{1}{2}\rho.V^2$.S. C_x^α; Y=q_∞. S. $C_y^\alpha = \frac{1}{2}\rho.V^2$.S. C_y^α. (1)

Với: - q_∞ là áp suất động;

- ρ (kg/m^3) là mật độ không khí;
- V là tốc độ bay;
- S là diện tích đặc trưng khí động;
- C_x^α; C_y^α là hệ số lực cản và lực nâng khí động phụ thuộc vào hệ số Makka M và α.

Các hệ số C_x^α, C_y^α được tính bằng phương pháp nội suy từ bảng:

Bảng 2.1. Giá trị hệ số khí động theo M và α:

α (độ)	M	C_x^α	C_y^α	α (độ)	C_x^α	C_y^α
0^0	0.2	0.3298	0	6^0	0.3172	0.57
	0.4	0.3206	0		0.312	0.585
	0.8	0.3104	0		0.3096	0.6448
	1	0.347	0		0.3474	0.6828
2^0	0.2	0.3152	0.1456	8^0	0.331	0.8082
	0.4	0.2946	0.1462		0.3246	0.8224
	0.8	0.3076	0.1672		0.3222	0.8944
	1	0.3416	0.1796		0.3632	0.9568
4^0	0.2	0.3214	0.3432	10^0	0.3182	1.086
	0.4	0.309	0.3488		0.3108	1.1082
	0.8	0.3092	0.3872		0.3142	1.2108
	1	0.3476	0.4208		0.3616	1.2982

(Trích dẫn từ bài báo ОПТИМИЗАЦИЯ ТРАЕКТОРИЙ БАЛЛИСТИЧЕСКОЙ РАКЕТЫ ПО МАКСИМАЛЬНОЙ КОНЕЧНОЙ МАССЕ của thầy Nguyễn Ngọc Điển đăng trên báo Science Prospects số thứ 108 [2]).

Xét chuyển động tên lửa trong mặt phẳng thẳng đứng OX_gZ_g với x là cự ly ngang còn z là độ cao. Vận tốc của tên lửa theo phương ngang và thẳng đứng lần lượt là v_x và v_z (m/s). Khi đó phương trình động học trong mặt phẳng thẳng đứng sẽ là $\begin{cases}\dot{x} = v_x \\ \dot{z} = v_z\end{cases}$. Tốc độ bay V và góc nghiêng quỹ đạo được xác định như sau: $V = \sqrt{v_x^2 + v_z^2}$ (m/s); θ = arctan($\frac{v_z}{v_x}$).

Chiếu các lực T, D và Y lên các trục của hệ hoạ độ mặt đất OX_gZ_g ta có các thành phần lực khí động và lực đẩy theo hai phương: $F_{ax} = -D\cos\vartheta - Y\sin\vartheta$; $F_{az} = -D\sin\vartheta + Y\cos\vartheta$; $T_x(t) = T(t)\cos(\vartheta + \gamma)$; $T_z(t) = T(t)\sin(\vartheta + \gamma)$.

Gia tốc chuyển động của tên lửa theo hai phương: $\begin{cases} \dot{v_x} = \frac{T(t)\cos(\vartheta+\gamma)+F_{ax}}{m(t)} \\ \dot{v_z} = \frac{T(t)\sin(\vartheta+\gamma)+F_{az}}{m(t)} - g \end{cases}$

Trong đó m(t) = $\begin{cases} m_0 - \dot{m}.t & với\ t \le t_2 \\ m_f = m_0 - \dot{m}.t_2 & với\ t > t_2 \end{cases}$ (kg) là khối lượng tên lửa theo thời gian (t2 (s) là thời điểm ngắt động cơ theo chương trình, $\dot{m}$ (kg/s) là tốc độ tiêu thụ nhiên liệu của động cơ).

Hệ phương trình động lực học quay của tên lửa:

$$\begin{cases} \dot{\vartheta} = q \\ I_y.\dot{q} = M_a + M_T + M_q \end{cases}$$

Trong đó: - q là tốc độ quay của tên lửa quanh trọng tâm trong mặt phẳng thẳng đứng;

- I_y là moment quán tính của tên lửa quanh trục ngang;

- M_a là moment khí động;
- M_T là moment do lực đẩy điều khiển tạo ra;
- M_q là moment cản quay của tên lửa.

Ta quy ước như sau: Lực đẩy động cơ có điểm đặt tại đuôi tên lửa, các toạ độ tâm áp x_{cp} và trọng tâm x_{cg} lấy theo chiều gốc toạ độ là đuôi theo chiều dương là mũi tên lửa. L là chiều dài tên lửa còn r là bán kính tên lửa. Khi đó các moment được tính như sau: $M_a = Y.(x_{cp} - x_{cg})$; $M_T = T(t).(-x_{cg}).\sin(\gamma)$; $I_y \approx \frac{1}{12}.m.(L^2 + 3.r^2)$

Tổng hợp các kết quả trên ta có hệ phương trình vi phân động học trong giai đoạn tên lửa bay có điều khiển sau:

$$\begin{cases} \dot{x} = v_x \\ \dot{z} = v_z \\ \dot{v_x} = \frac{T(t)\cos(\vartheta+\gamma)+F_{ax}}{m(t)} \\ \dot{v_z} = \frac{T(t)\sin(\vartheta+\gamma)+F_{az}}{m(t)} - g \\ \dot{\vartheta} = q \\ \dot{q} = \frac{q_\infty .S.C_y^\alpha.(x_{cp}-x_{cg})-T(t).x_{cg}.\sin(\gamma)}{I_y} \end{cases}$$

Từ đây, ta xác định được vector trạng thái tổng thể của bài toán chuyển động tên lửa là x=[x, z, v_x, v_z, ϑ , q]^T.

(Các công thức khí động và động lực học quay tên lửa trong chương này được trích dẫn từ tài liệu giáo trình Lý thuyết bay và hệ thống điều khiển tên lửa phòng không - tập I, Học viện Kỹ thuật quân sự, Hà Nội của các tác giả Lê Anh Dũng, Nguyễn Hữu Độ, Huỳnh Lương Nghĩa [1])

2.3. Xây dựng bài toán điều khiển tối ưu

Trong bài toán đang xét, mục tiêu điều khiển là thay đổi góc lệch véc-tơ lực đẩy γ nhằm đưa góc gật ϑ về giá trị mong muốn, đồng thời bảo đảm góc tấn công α xấp xỉ bằng 0^o tại thời điểm ngắt động cơ nhằm tạo điều kiện thuận lợi cho giai đoạn bay đạn đạo. Tuy nhiên, hệ động lực học đầy đủ của tên lửa là hệ phi tuyến bậc cao và có các tham số biến thiên theo thời gian gây phức tạp cho việc thiết kế bộ điều khiển. Vì vậy, bài toán được tách thành hai mức: mô hình phi tuyến đầy đủ dùng để mô phỏng quỹ đạo bay và mô hình cục bộ bậc thấp dùng để thiết kế bộ điều khiển LQR. Cách tiếp cận này vừa bảo đảm mô hình mô phỏng sát với thực tế, vừa tạo được một cấu trúc điều khiển đủ đơn giản để tổng hợp bộ điều khiển phản hồi trạng thái.

2.3.1. Mô hình trạng thái đầy đủ của đối tượng

Từ hệ phương trình động lực học đã xây dựng ở mục 2.2, trạng thái tổng thể của tên lửa trong mặt phẳng thẳng đứng được biểu diễn bởi vectơ x_f=[x, z, v_x, v_z, ϑ , q]^T. Tuy nhiên do tín hiệu điều khiển γ tác động trực tiếp lên mômen quay của tên lửa nên bài toán tổng hợp điều khiển được rút gọn về phân hệ quay.

2.3.2. Lựa chọn mô hình cục bộ cho bài toán điều khiển

Trong giai đoạn bay có điều khiển, đại lượng cần bám là góc gật ϑ . Ta sử dụng sai số góc $= \vartheta - \vartheta_{ref}$ và vector trạng thái cục bộ được chọn là $x_a = [e_\vartheta \ q]^T$. Lựa chọn này phản ánh trực tiếp chất lượng điều khiển kênh quay: e_ϑ thể hiện sai số bám còn q đặc trưng cho động học quá độ. Biến điều khiển được chọn là $u = \gamma$ với ràng buộc biên độ $|\gamma| \le \gamma_{max} = 5^o$. Bộ điều khiển LQR chỉ đảm nhiệm ổn định cục bộ và bám góc gật, còn yêu cầu tối ưu quỹ đạo được xử lý thông qua thiết kế tín hiệu tham chiếu $\vartheta_{ref}(t)$ và tối ưu tham số t_{ramp} sẽ được trình bày chi tiết ở mục 2.5.

2.3.3. Thiết lập quan hệ động học cho biến trạng thái

Từ định nghĩa sai số góc $e_\vartheta = \vartheta - \vartheta_{ref}$, lấy đạo hàm theo thời gian: $\dot{e_\vartheta} = \dot{\vartheta} - \dot{\vartheta}_{ref} = q - \dot{\vartheta}_{ref}$. Trong quá trình thiết kế LQR, tín hiệu tham chiếu biến thiên chậm do đó có thể bỏ qua $\dot{\vartheta}_{ref}$, từ đó thu được gần đúng $\dot{e_\vartheta} \approx q$ (3). Đây là phương trình thứ nhất của mô hình trạng thái cục bộ. Phương trình còn lại được xây dựng từ phương trình động lực học quay của tên lửa.

2.3.4. Tuyến tính hoá mômen khí động và mômen điều khiển

Xét phương trình động lực học quay của tên lửa: $I_y.\dot{q} = M_a + M_T + M_q$, ta sẽ tuyến tính hoá quanh điểm làm việc $\alpha \approx 0$.

Mômen khí động: $M_a = Y.(x_{cp} - x_{cg}) = q_\infty .S.C_y^\alpha.(x_{cp} - x_{cg})$. Sử dụng khai triển Taylor bậc 1 với C_y^α ta có biểu thức: $C_y(M,\alpha) \approx C_y(M,0) + \alpha.\frac{\partial C_y}{\partial \alpha}\Big|_0$; thay vào biểu thức của moment khí động → $M_a = M_{a0} + q_\infty .S.(x_{cp} - x_{cg}).\alpha.\frac{\partial C_y}{\partial \alpha}\Big|_0$ trong đó M_{a0} là thành phần hằng tại điểm làm việc. Khi xây dựng mô hình tuyến tính theo biến sai lệch, thành phần M_{a0} được bỏ qua, kết hợp với xấp xỉ $\alpha \approx e_\vartheta$ ta thu được biểu thức gần đúng: $M_a \approx k_\alpha.e_\vartheta$. ($k_\alpha = q_\infty .S.(x_{cp} - x_{cg}).\frac{\partial C_y}{\partial \alpha}\Big|_0$).

Moment quay của lực đẩy: $M_T = T(t).(-x_{cg}).\sin(\gamma) \approx T(t).(-x_{cg}).\gamma$. (với γ nhỏ).

Moment cản quay khí động được mô hình hoá tuyến tính: $M_q \approx k_q.q$ ($k_q < 0$).

Với ba phương trình xấp xỉ trên ta có: $I_y.\dot{q} = k_\alpha.e_\vartheta + k_q.q + T(t).(-x_{cg}).\gamma$. Đặt $a_1 = \frac{k_\alpha}{I_y}$; $a_2 = \frac{k_q}{I_y}$; $b = \frac{T(t).(-x_{cg})}{I_y}$ ta thu được phương trình: $\dot{q} = a_1.e_\vartheta + a_2.q + b.\gamma$. (4)

Từ hai phương trình (3) và (4) ta có hệ phương trình trạng thái để xây dựng bộ điều khiển LQR: $\begin{cases} \dot{e}_\vartheta \approx q; \\ \dot{q} = a_1.e_\vartheta + a_2.q + b.\gamma. \end{cases}$

Từ đó, ta xây dựng được mô hình trạng thái cho bộ điều khiển LQR: $\dot{x_a}$=A.$x_a + B.\gamma$ trong đó A=$\begin{bmatrix} 0 & 1 \\ a_1 & a_2 \end{bmatrix}$; B=$\begin{bmatrix} 0 \\ b \end{bmatrix}$. Trong thực tế, thành phần a_2 phản ánh tác dụng tắt dần của động học quay. Tuy nhiên, vì ảnh hưởng cản quay nhỏ trong miền làm vinên ệc khảo sát để đơn giản hóa thiết kế LQR, thành phần này được bỏ qua trong mô hình tuyến tính Như vậy A=$\begin{bmatrix} 0 & 1 \\ a_1 & 0 \end{bmatrix}$.

2.4. Thiết kế điều khiển tối ưu LQR

Với hàm mục tiêu dạng bình phương toàn phương J=$\int_0^\infty (x_a^T.Q.x_a + \gamma^T.R.\gamma)$dt và mô hình tuyến tính $\dot{x_a}$=A.$x_a + B.\gamma$; bộ điều khiển LQR tối ưu được xác định từ nghiệm P của phương trình Riccacti đại số: $A^TP + PA - PBR^{-1}B^TP + Q = 0$. (6)

Chọn $Q = \begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix}$; R = r; lấy P đối xứng: $P = \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix}$.

Phương trình (6) được giải như sau:

$$\begin{bmatrix} 0 & a_1 \\ 1 & 0 \end{bmatrix} \times \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \times \begin{bmatrix} 0 & 1 \\ a_1 & 0 \end{bmatrix} -$$

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \times \begin{bmatrix} 0 \\ b \end{bmatrix} \times \frac{1}{r} \times [0 \;\; b] \times \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix} = 0$$

$$\leftrightarrow \begin{bmatrix} a_1p_{12} & a_1p_{22} \\ p_{11} & p_{12} \end{bmatrix} + \begin{bmatrix} a_1p_{12} & p_{11} \\ a_1p_{22} & p_{12} \end{bmatrix} + \frac{b^2}{r} \times \begin{bmatrix} p_{12}^2 & p_{12}p_{22} \\ p_{12}p_{22} & p_{22}^2 \end{bmatrix} + \begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix}$$

So sánh từng phần tử, ta được hệ phương trình:

$$\begin{cases} 2a_1p_{12} - \dfrac{b^2p_{12}^2}{r} + q_1 = 0 \\ a_1p_{22} + p_{11} - \dfrac{b^2p_{12}p_{22}}{r} = 0 \\ 2p_{12} - \dfrac{b^2p_{22}^2}{r} + q_2 = 0 \end{cases}$$

Giải hệ phương trình trên nhận được lần lượt các nghiệm sau:

$$p_{12} = \frac{r}{b^2}(a_1 \pm \sqrt{a_1^2 + \frac{b^2q_1}{r}}); \; p_{22} = \pm\sqrt{\frac{r}{b^2}(2p_{12} + q_2)}; \; p_{11} = (\frac{b^2}{r}p_{12} - a_1)p_{22}.$$

Sau khi có nghiệm P ta sẽ tính được ma trận khuếch đại tối ưu:

K=$R^{-1}B^TP = \frac{1}{r}\times[0\ \ b]\times\begin{bmatrix}p_{11} & p_{12}\\ p_{12} & p_{22}\end{bmatrix} = \left[\frac{b}{r}p_{12}\ \frac{b}{r}p_{22}\right]$; từ đó có được luật điều khiển phản hồi trạng thái: $\gamma = -K.x_a$. Trong nghiên cứu áp dụng giới hạn $|\gamma| \le 5^o$.

2.5. Xây dựng hàm tham chiếu $\vartheta_{ref}(t)$ theo tham số t_{ramp} và tối ưu t_{ramp}

Trong bài toán điều khiển véc-tơ lực đẩy, nếu áp dụng trực tiếp tín hiệu đặt góc gật dưới dạng bước từ trạng thái ban đầu 90º về giá trị mong muốn 45º, sai lệch trạng thái xuất hiện tức thời sẽ rất lớn. Điều này làm bộ điều khiển LQR sinh lệnh điều khiển mạnh, dễ gây quá điều chỉnh, kích thích mạnh động học quay, làm tăng góc tấn công và có thể đẩy hệ ra khỏi miền tuyến tính hóa. Vì vậy, như đã nói ở phần trên tín hiệu đặt không được lựa chọn dưới dạng bước mà được xây dựng dưới dạng một quỹ đạo tham chiếu biến thiên trơn theo thời gian.

2.5.1. Xây dựng hàm tham chiếu $\vartheta_{ref}(t)$

Sau giai đoạn phóng thẳng đứng trong 3s, tên lửa bắt đầu bẻ lái để chuyển dần sang góc gật mong muốn $\vartheta_f = 45^o$. Trong quá trình này bộ điều khiển LQR sẽ tính toán lệnh điều khiển để góc gật ϑ bám theo hàm tham chiếu $\vartheta_{ref}(t)$ có giá trị giảm dần từ 90^o về 45^o trong khoảng thời gian chuyển tiếp là t_{ramp}, sau đó giữ cố định. Giá trị t_{ramp} phải đảm bảo thời gian chuyển tiếp không vượt quá thời điểm tắt động cơ ($t_{ramp} \le t_{cutoff} - 3s$). Đặt biến thời gian chuẩn hóa $\tau = \frac{t-3}{t_{ramp}} \in [0,1]$. Phương án đơn giản nhất là sử dụng hàm nội suy tuyến tính s(τ) = τ với $\vartheta_{ref}(\tau) = 90^o - 45^o.s(\tau)$, tuy nhiên đạo hàm của tín hiệu tham chiếu bị thay đổi đột ngột tại thời điểm bắt đầu và kết thúc quá trình chuyển tiếp, gây kích thích mạnh lên hệ điều khiển. Để tránh điều này ta sẽ sử dụng hàm trơn bậc 3: $s(\tau) = s_0 + s_1.\tau + s_2.\tau^2 + s_3\tau^3$. Để chọn được các tham số s_0, s_1, s_2, s_3 ta căn cứ vào các điều kiện sau: $\vartheta_{ref}(0) = 90^o$ (s(0)=0); giá trị cuối của $\vartheta_{ref}(1) = 45^o$ (s(1)=1); tốc độ thay đổi của hàm tham chiếu tại thời điểm đầu và cuối bằng 0 ($\dot{\vartheta}_{ref}(\tau) = 0$ với $\tau = 0$ và $\tau = 1 \leftrightarrow \begin{cases}\dot{s}(0) = 0\\ \dot{s}(1) = 0\end{cases}$).

Từ các điều kiện trên, ta có hệ phương trình: $\begin{cases} s_0 = 0;\\ s_2 + s_3 = 1;\\ s_1 = 0;\\ 3s_3 + 2s_2 = 0.\end{cases}$

Giải hệ phương trình trên ta có $s_0 = s_1 = 0$; $s_2 = 3$; $s_3 = -2 \rightarrow s(\tau) = 3.\tau^2 - 2.\tau^3$.

Như vậy hàm tham chiếu $\vartheta_{ref}(t)$ sẽ là:

$$\vartheta_{ref}(t) = \begin{cases}90^o, & t \le 3s\\ 90^o - 45^o.\left(3.\frac{t-3}{t_{ramp}}^2 - 2.\frac{t-3}{t_{ramp}}^3\right), & 3s \le t \le 3s + t_{ramp}\\ 45^o, & t \ge 3s + t_{ramp}\end{cases}$$

2.5.2. Tối ưu tham số t_{ramp}

Trong nghiên cứu này, t_{ramp} được xem là một tham số tối ưu quỹ đạo sao cho tại thời điểm tắt động cơ t_{cutoff} các góc $\vartheta \approx 45^o, \alpha \approx 0^o$ và tên lửa chuyển động ổn định.

Để đánh giá chất lượng của từng giá trị t_{ramp}, ta sử dụng hàm chỉ tiêu:

$$J = j_1.e_\theta^2 + j_2.e_\alpha^2 + j_3.\alpha_{RMS}^2 + j_4.v_\theta^2$$

Trong đó :

- e_θ: sai lệch góc gật tại thời điểm ngắt động cơ: $e_\theta = \theta(t_{cutoff}) - 45^o$
- e_α: sai lệch góc tấn công tại thời điểm ngắt động cơ: $e_\alpha = \alpha(t_{cutoff}) - 0^o$
- α_{RMS}: đặc trưng dao động khí động.
- v_θ: chỉ tiêu đánh giá mức độ ổn định của góc gật quanh giá trị mong muốn.

Các trọng số j_i được lựa chọn theo mức độ ưu tiên. Giá trị tối ưu t_{ramp} được xác định bằng cách quét trong miền khảo sát, mô phỏng hệ và chọn giá trị cho J nhỏ nhất.

3. Mô phỏng khảo sát bằng phần mềm Matlab

Ta sẽ xây dựng chương trình Matlab mô phỏng việc tối ưu điều khiển đưa tên lửa từ quỹ đạo bay thẳng đứng vào quỹ đạo bay có điều khiển, sau đó mô phỏng đường bay của tên lửa và đánh giá chất lượng điều khiển. Chương trình được xây dựng trên cơ sở điều khiển tối ưu LQR kết hợp với kỹ thuật làm trơn tín hiệu tham chiếu.

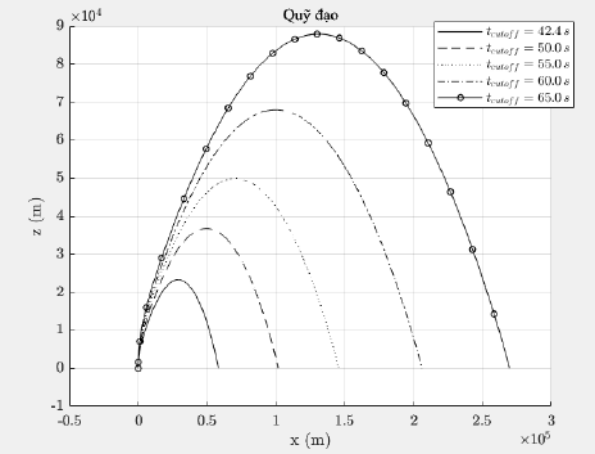
3.1. Đối tượng mô phỏng

Đối tượng được lựa chọn để mô phỏng trong bài báo là mô hình tên lửa phỏng theo tên lửa SCUD-B với các tham số cơ bản như sau:

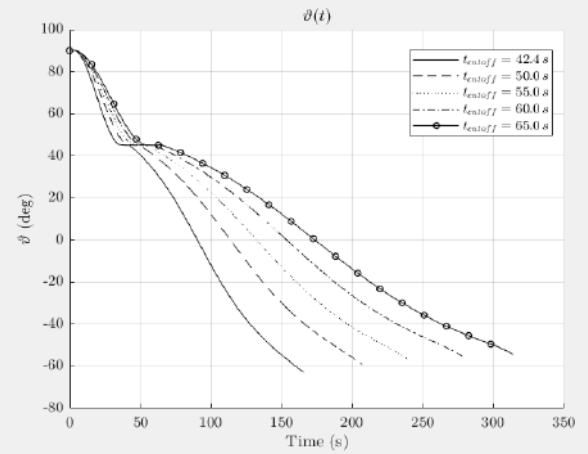
Chiều dài tên lửa	:	11,164 m
Đường kính thân tên lửa	:	0.88 m
Khối lượng trước phóng	:	5860 kg
Khối lượng nhiên liệu	:	3741 kg
Tốc độ tiêu thụ nhiên liệu	:	57,83 kg/s
Lực đẩy động cơ	:	13380 Kgf
Vị trí trọng tâm (tính từ loa phụt)	:	6,224 m
Vị trí tâm khí động (tính từ loa phụt)	:	4 m
Giới hạn góc lệch loa phụt động cơ	:	$\pm 5^o$

(Các tham số trên được tham khảo từ tài liệu về tên lửa Scud-B: Кузнецов П. И. (1961), “РАКЕТА 8К14 - Пособие Для Офицеров Ракетиых Войск”, Москва [3])

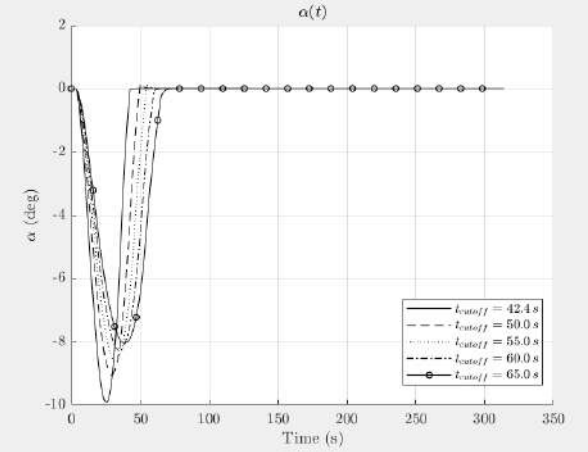
3.2. Kết quả mô phỏng



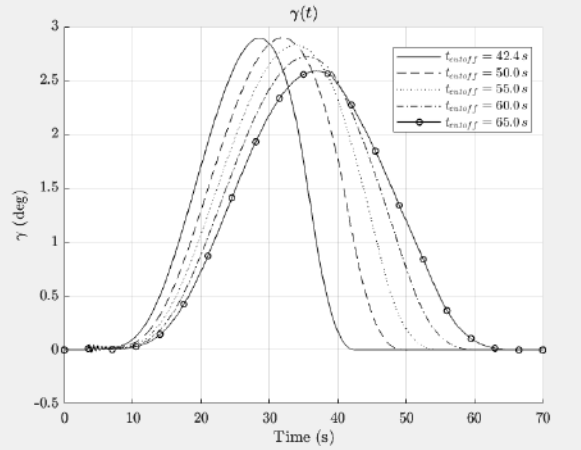
Hình 3.1. Quỹ đạo bay của tên lửa



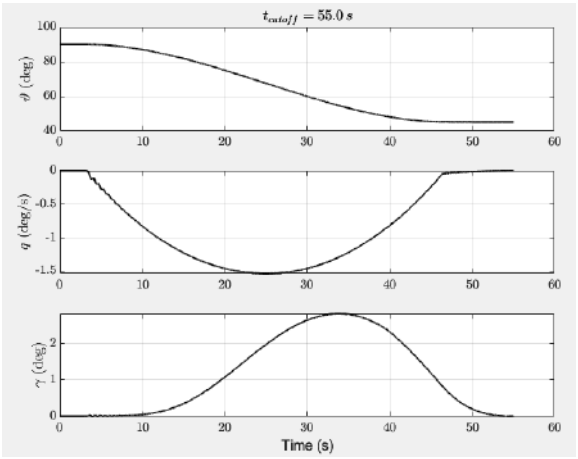
Hình 3.2. Góc gật ϑ



Hình 3.3. Góc tấn công α



Hình 3.4. Góc véc-tơ lực đẩy γ



Hình 3.5. Đáp ứng tổng hợp của góc gật ϑ(t), q(t) và γ(t) với trường hợp ngắt động cơ tại 55s

Nhận xét:

- Kết quả mô phỏng cho thấy bộ điều khiển LQR đảm bảo tên lửa chuyển từ trạng thái bay thẳng đứng sang quỹ đạo mong muốn nhanh và ổn định. Quỹ đạo bay (Hình 3.1) phù hợp các giai đoạn điều khiển và phụ thuộc vào thời điểm ngắt động cơ.

- Góc gật ϑ (Hình 3.2) bám tốt giá trị tham chiếu 45º, dao động nhỏ và duy trì ổn định trước thời điểm ngắt động cơ. Góc tấn công α (Hình 3.3) tại thời điểm ngắt động cơ xấp xỉ bằng 0.

- Tín hiệu điều khiển γ biến thiên trơn không xuất hiện bão hoà, chỉ có dao động nhẹ trong giai đoạn quá độ. Đáp ứng tổng hợp (Hình 3.5) cho thấy hệ có đặc tính ổn định, dao động nhỏ và quá trình chuyển tiếp diễn ra hợp lý.

- Nhìn chung, kết quả mô phỏng xác nhận tính khả thi của phương pháp điều khiển phản hồi trạng thái theo LQR kết hợp làm trơn tín hiệu tham chiếu.

3.3. Kết luận

Bài báo đã xây dựng mô hình động lực học phi tuyến của tên lửa đạn đạo trong mặt phẳng thẳng đứng có xét đến khối lượng biến thiên và lực khí động. Trên cơ sở đó, bộ điều khiển LQR được thiết kế từ mô hình tuyến tính hóa cục bộ, kết hợp với hàm tham chiếu trơn và tối ưu tham số t_{ramp}. Kết quả mô phỏng cho thấy góc bám nhanh và ổn định giá trị 45°, góc tấn công tại thời điểm ngắt động cơ được duy trì xấp xỉ bằng 0, đồng thời tín hiệu điều khiển biến thiên trơn và phù hợp với ràng buộc thực tế. Nhìn chung, phương pháp đề xuất đảm bảo chất lượng điều khiển và khẳng định tính khả thi của phương pháp điều khiển véc-tơ lực đẩy trong tối ưu quỹ đạo pha bay tích cực.

Tài liệu tham khảo

1. Lê Anh Dũng, Nguyễn Hữu Độ, Huỳnh Lương Nghĩa (1998), Lý thuyết bay và hệ thống điều khiển tên lửa phòng không - tập I, Học viện Kỹ thuật quân sự, Hà Nội.
2. Nguyễn Ngọc Điển (2018), ОПТИМИЗАЦИЯ ТРАЕКТОРИЙ БАЛЛИСТИЧЕСКОЙ РАКЕТЫ ПО МАКСИМАЛЬНОЙ КОНЕЧНОЙ МАССЕ, ПЕРСПКТИВЫ НАУКИ, No.9 (108), Tambov, ТМБпринт, 84.
3. Кузнецов П. И. (1961), “РАКЕТА 8К14 - Пособие Для Офицеров Ракетиых Войск”, Москва.

Tổng hợp luật dẫn đa chỉ tiêu cho tên lửa đối hạm trong giai đoạn tự dẫn

Bùi Quốc Dũng
Bộ môn Tên lửa, Viện Tên lửa và Kỹ thuật điều khiển, Học viện KTQS
Email: bqdzung@lqdtu.edu.vn; Tel: 0983160246

Tóm tắt

Bài báo trình bày một luật dẫn đa chỉ tiêu cho tên lửa đối hạm, đó là: đảm bảo dẫn tên lửa đối hạm tới gặp tàu mục tiêu với góc tiếp cận cho trước; tạo ra dạng quỹ đạo cơ động hình sin trong mặt phẳng và hình xoắn ốc trong không gian; gia tốc tiến dần về không khi gặp mục tiêu. Các chỉ tiêu này nhằm gây khó khăn cho tên lửa phòng không khi đánh chặn và tăng khả năng sát thương tàu mục tiêu. Phần mô phỏng số để khảo sát các đặc trưng và hiệu quả của luật dẫn đã tổng hợp.

Từ khóa: *Cơ động, tên lửa đối hạm, tên lửa phòng không.*

1. Mở đầu

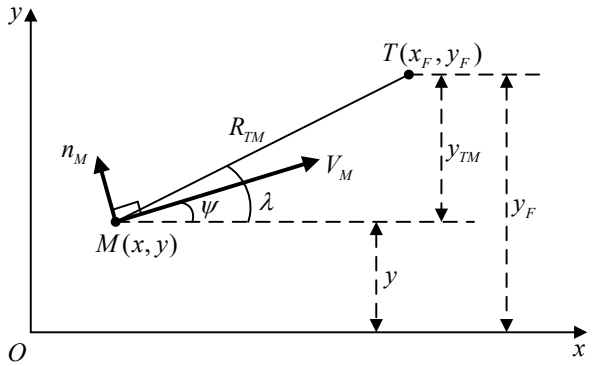
Khả năng vượt hệ thống phòng không đối phương có ý nghĩa quyết định đối với hiệu quả chiến đấu của tên lửa đối hạm (TLĐH). Để tiêu diệt được mục tiêu, TLĐH cần thực hiện hai nhiệm vụ: vượt hỏa lực phòng không và đánh trúng tàu đối phương với hiệu quả cao nhất. Đối với nhiệm vụ thứ nhất, giải pháp thường dùng là cơ động với quỹ đạo phức tạp để làm tăng độ trượt của TLPK. Các tài liệu [1,2,3,4,5] chỉ ra dạng cơ động điển hình của các phương tiện tấn công đường không chống TLPK là cơ động hình sin trong mặt phẳng và cơ động xoắn ốc trong không gian. Đối với nhiệm vụ thứ hai, bằng cách điều khiển góc tiếp cận khi gặp mục tiêu sẽ làm tăng hiệu quả của đầu đạn đồng thời có thể phối hợp tấn công từ nhiều hướng [6].

Trong bài báo này, tác giả đưa ra một luật dẫn mới cho TLĐH đa chỉ tiêu: có thể tạo ra dạng chuyển động phức tạp để vượt hỏa lực phòng không; gặp tàu theo góc cho trước đồng thời đảm bảo gia tốc tiến về không khi TL gặp mục tiêu.

2. Tổng hợp luật dẫn đa chỉ tiêu cho tên lửa đối hạm

2.1. Mô hình tương quan trong mặt phẳng

Trên Hình 1 mô tả mối quan hệ động hình học giữa TLĐH với tàu mục tiêu trong mặt phẳng nằm ngang, trong đó mục tiêu được xem là đứng yên tại tọa độ (x_F, z_F).



Hình 1. Quan hệ động hình học TLĐH – Tàu mục tiêu

Tên lửa chuyển động với vận tốc V_M không đổi trong suốt quá trình bay với góc vòng quỹ đạo $\psi(t)$. Gia tốc của tên lửa n_M vuông góc với vận tốc của nó. Thời điểm ban đầu $t = 0$ là lúc bắt đầu vào tự dẫn, thời điểm kết thúc $t = T$ là lúc tên lửa gặp tàu mục tiêu.

Theo sơ đồ động hình học Hình 1a, ta có các phương trình mô tả chuyển động của tên lửa trong hệ tọa độ dẫn Oxz như sau:

$$\dot{x} = V_M \cos\psi;\ \dot{y} = V_M \sin\psi;\ \dot{\psi} = n_M / V_M \tag{1}$$

Tuyến tính hóa theo giả định ψ nhỏ, ta có:

$$\dot{x} = V_M;\ \dot{y} = V_M\psi \tag{2}$$

Gia tốc lệnh của luật dẫn xây dựng theo cấu trúc sau:

$$n_M = n_E + n_C \tag{3}$$

Trong đó, n_C là thành phần gia tốc lệnh điều khiển TLĐH gặp tàu với góc tiếp cận được chỉ định, n_E là thành phần gia tốc lệnh đưa vào trước để tạo ra cơ động hình sin với cường độ k và tần số ω.

$$n_E = k \sin\omega t \tag{4}$$

Đặt $a_S = \sin\omega t$ và $a_C = \cos\omega t$, ta có:

$$n_E = ka_S;\ \dot{a}_S = \omega a_C;\ \dot{a}_C = -\omega a_S;\ \dot{\psi} = \frac{ka_S + n_C}{V_M} \tag{5}$$

Thời gian dẫn TLĐH đến gặp tàu mục tiêu được tính như sau:

$$T = (x_F - x_0) / V_M \tag{6}$$

Đặt véc tơ trạng thái của hệ thống là:

$$\boldsymbol{x} = [y \quad \psi \quad a_S \quad a_C]^T \tag{7}$$

Từ (2), (5), (7), ta thu được hệ phương trình không gian trạng thái của hệ thống sau đây:

$$\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}n_C \tag{8}$$

Trong đó

$$A = \begin{bmatrix} 0 & V_M & 0 & 0 \\ 0 & 0 & \dfrac{k}{V_M} & 0 \\ 0 & 0 & 0 & \omega \\ 0 & 0 & -\omega & 0 \end{bmatrix},\quad B = \begin{bmatrix} 0 \\ \dfrac{1}{V_M} \\ 0 \\ 0 \end{bmatrix} \tag{9}$$

Bây giờ, xem xét bài toán điều khiển tối ưu để tối thiểu hóa phiếm hàm chỉ tiêu chất lượng sau:

$$J = \frac{1}{2}\int_0^T n_C^2 dt \tag{10}$$

Và tuân theo động học hệ thống (8) với điều kiện ràng buộc cuối TLĐH gặp tàu mục tiêu với độ trượt bằng không và góc tiếp cận cho trước, tức là:

$$y(T) = y_F;\ \psi(T) = \psi_F \tag{11}$$

Bằng cách giải bài toán tối ưu ở trên, chúng ta có được luật dẫn đa chỉ tiêu: tạo ra cơ động hình sin; độ trượt cuối bằng không; tiếp cận mục tiêu với góc cho trước; gia tốc dần về không.

2.2 Tổng hợp luật dẫn phẳng cho TLĐH

Do chỉ cần ràng buộc điều kiện cuối với hai biến (y,ψ) của véc tơ trạng thái nên có thể đặt yêu cầu trạng thái cuối của hệ thống dưới dạng:

$$\boldsymbol{z}(T) = \boldsymbol{C}\boldsymbol{x}(T) \tag{12}$$

Ở đây, ma trận $\boldsymbol{C}$ được sử dụng để lựa chọn biến trạng thái cần ràng buộc. Để thỏa mãn ràng buộc (11), ma trận $\boldsymbol{C}$ được lựa chọn như sau:

$$\boldsymbol{C} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{13}$$

Xét một véc tơ trạng thái mới ở thời điểm t bất kì:

$$\boldsymbol{z}(t) = \boldsymbol{C}\boldsymbol{\Phi}(T,t)\boldsymbol{x}(t) \tag{14}$$

Trong đó, $\boldsymbol{\Phi}(T,t)$ là ma trận cơ sở chuyển trạng thái của hệ (8) . Theo tài liệu [7] ma trận này thỏa mãn tính chất sau:

$$\dot{\boldsymbol{\Phi}}(T,t) = \boldsymbol{\Phi}(T,t)\boldsymbol{A};\ \boldsymbol{\Phi}(T,T) = \boldsymbol{I} \tag{15}$$

Vi phân hai vế của phương trình (14), sau đó sử dụng (8) và (15) thay vào ta thu được:

$$\dot{\boldsymbol{z}}(t) = \boldsymbol{F}(t)n_C;\ \boldsymbol{F}(t) = \boldsymbol{C}\boldsymbol{\Phi}(T,t)\boldsymbol{B} \tag{16}$$

Từ (16) ta thấy, để đạt được trạng thái cuối $\boldsymbol{z}(T)$ cần thực hiện điều khiển đối với véc tơ $\boldsymbol{z}(t)$. Khi này bài toán điều khiển tối ưu được chuyển từ động học hệ thống (8) sang (16) với biến trạng thái $\boldsymbol{z}(t)$.

Áp dụng phương pháp biến phân Euler-Lagrange trình bày trong các tài liệu [7-9] cho hệ thống mô tả bởi (16), phiếm hàm chỉ tiêu chất lượng (10) và điều kiện cuối (11), ta thực hiện các bước sau:

Lập hàm Hamilton

$$H(\boldsymbol{p}, n_C, t) = \boldsymbol{p}^T\boldsymbol{F}(t)n_C - \frac{1}{2}n_C^2 \tag{17}$$

Trong đó, $\boldsymbol{p}(t)$ là véc tơ đồng trạng thái.

Theo điều kiện cần để có nghiệm tối ưu, ta có:

$$\frac{\partial H}{\partial n_C} = 0^T \Leftrightarrow \boldsymbol{p}^T\boldsymbol{F}(t) - n_C = 0^T \ \Leftrightarrow\ n_C = \boldsymbol{F}^T(t)\boldsymbol{p} \tag{18}$$

Lập phương trình Euler-Lagrange đối với biến đồng trạng thái, ta thu được:

$$\frac{d\boldsymbol{p}}{dt} = -\left(\frac{\partial H}{\partial y}\right)^T \Leftrightarrow \frac{d\boldsymbol{p}}{dt} = 0^T \Leftrightarrow \boldsymbol{p} = \boldsymbol{p}_0 = const \tag{19}$$

Thay (19) vào (18), ta có:

$$n_C = \boldsymbol{F}^T(t)\boldsymbol{p}_0 \tag{20}$$

Thay (20) vào (16), ta thu được:

$$\dot{\boldsymbol{z}} = \boldsymbol{F}(t)\boldsymbol{F}^T(t)\boldsymbol{p}_0 \tag{21}$$

Tích phân hai vế của phương trình (21) trong khoảng $[t,T]$, ta có:

$$\boldsymbol{z}(T) - z(t) = \left[\int_t^T \boldsymbol{F}(\lambda)\boldsymbol{F}^T(\lambda)d\lambda\right]\boldsymbol{p}_0 \tag{22}$$

Đặt:

$$\boldsymbol{G}(t) = \int_t^T \boldsymbol{F}(\lambda)\boldsymbol{F}^T(\lambda)d\lambda \tag{23}$$

Từ (22) và (23) ta rút ra được:

$$\boldsymbol{p}_0 = \boldsymbol{G}^{-1}(t)\left[\boldsymbol{z}(T) - z(t)\right] \tag{24}$$

Thay (24) vào (20) ta nhận được thành phần gia tốc lệnh n_C như sau:

$$n_C = \boldsymbol{F}^T(t)\boldsymbol{G}^{-1}(t)\left[\boldsymbol{z}(T) - z(t)\right] \tag{25}$$

Để xác định gia tốc lệnh của luật dẫn n_M theo véc tơ trạng thái ban đầu $\boldsymbol{x}(t)$ ta thay (12) và (14) vào (25), ta có:

$$n_C = \boldsymbol{F}^T(t)\boldsymbol{G}^{-1}(t)\boldsymbol{C}\boldsymbol{x}(T) - \boldsymbol{F}^T(t)\boldsymbol{G}^{-1}(t)\boldsymbol{C}\boldsymbol{\Phi}(T,t)\boldsymbol{x}(t) \tag{26}$$

Từ (3), (5) và (26) ta nhận được gia tốc lệnh của luật dẫn phẳng cho TLĐH có dạng như sau:

$$n_M = \boldsymbol{M}(t)\boldsymbol{x}(t) + N(t) \tag{27}$$

Trong đó

$$\boldsymbol{M}(t) = -\boldsymbol{F}^T(t)\boldsymbol{G}^{-1}(t)\boldsymbol{C}\boldsymbol{\Phi}(T,t) \tag{28}$$

$$N(t) = \boldsymbol{F}^T(t)\boldsymbol{G}^{-1}(t)\boldsymbol{C}\boldsymbol{x}(T) + ka_S \tag{29}$$

Tiếp theo ta sẽ tính $\boldsymbol{M}(t)$ và $N(t)$ thông qua $\boldsymbol{\Phi}(T,t)$, $\boldsymbol{F}(t)$, $\boldsymbol{G}(t)$. Theo các tài liệu [5, 6] ma trận $\boldsymbol{\Phi}(T,t)$ được tính theo công thức sau:

$$\boldsymbol{\Phi}(T,t) = \boldsymbol{\Phi}(T-t) = L^{-1}\left\{(s\boldsymbol{I} - \boldsymbol{A})^{-1}\right\} \tag{30}$$

Đặt $t_{go} = T - t$ là thời gian bay còn lại tới điểm gặp, ta có:

$$\boldsymbol{\Phi}(T,t)=\boldsymbol{\Phi}(t_{go})=\begin{bmatrix}1 & \phi_{12} & \phi_{13} & \phi_{14}\\ 0 & 1 & \phi_{23} & \phi_{24}\\ 0 & 0 & \phi_{33} & \phi_{34}\\ 0 & 0 & \phi_{43} & \phi_{44}\end{bmatrix} \tag{31}$$

Trong đó

$$\phi_{12}=V_M t_{go};\ \phi_{13}=\frac{k(1-\cos\omega t_{go})}{\omega^2};\ \phi_{14}=\frac{k(\omega t_{go}-\sin\omega t_{go})}{\omega^2};\ \phi_{23}=\frac{k\sin\omega t_{go}}{\omega V_M};$$

$$\phi_{24}=\frac{k(1-\cos\omega t_{go})}{\omega V_M};\ \phi_{33}=\phi_{44}=\cos\omega t_{go};\ \phi_{34}=-\phi_{43}=\sin\omega t_{go}$$

Từ (16) ta xác định được ma trận $\boldsymbol{F}(t_{go})$ như sau:

$$\boldsymbol{F}(t_{go})=\boldsymbol{C}\boldsymbol{\Phi}(t_{go})\boldsymbol{B}\ \Rightarrow\ \boldsymbol{G}(t_{go})=-\int_{t_{go}}^{0}\boldsymbol{F}(\lambda)\boldsymbol{F}^T(\lambda)d\lambda \tag{32}$$

Từ đây, theo công thức (28), (29) ta tính được $\boldsymbol{M}(t_{go})$ và $N(t_{go})$ như sau:

$$\boldsymbol{M}(t_{go})=-\boldsymbol{F}^T(t_{go})\boldsymbol{G}^{-1}(t_{go})\boldsymbol{C}\boldsymbol{\Phi}(t_{go}) \tag{33}$$

$$N(t_{go})=\boldsymbol{F}^T(t_{go})\boldsymbol{G}^{-1}(t_{go})\boldsymbol{C}\boldsymbol{x}(T)+ka_S \tag{34}$$

Thay (33), (34) vào (27) ta thu được phương trình luật dẫn như sau:

$$n_M=\boldsymbol{M}(t_{go})\boldsymbol{x}+N(t_{go}) \tag{35}$$

Khai triển và biến đổi (35) ta có:

$$n_M=\frac{6}{t_{go}^2}[(y_F-y)-V_M\psi t_{go}]-\frac{2}{t_{go}}(V_M\psi_F-V_M\psi)$$
$$+\frac{\omega^2t_{go}^2+2\omega t_{go}\sin\omega t_{go}+6\cos\omega t_{go}-6}{\omega^2t_{go}^2}a_S+\frac{6\sin\omega t_{go}-2\omega t_{go}\cos\omega t_{go}-4\omega t_{go}}{\omega^2t_{go}^2}a_C \tag{36}$$

Từ quan hệ động hình học Hình 1 và vì $y_{TM}<<R_{TM}$, V_C là tốc độ tiếp cận nên ta có:

$$\lambda\approx\sin\lambda=\frac{y_{TM}}{R_{TM}}=\frac{y_F-y}{V_Ct_{go}}\ \Rightarrow\ \dot\lambda=\frac{(y_F-y)-V_M\psi t_{go}}{V_Ct_{go}^2} \tag{37}$$

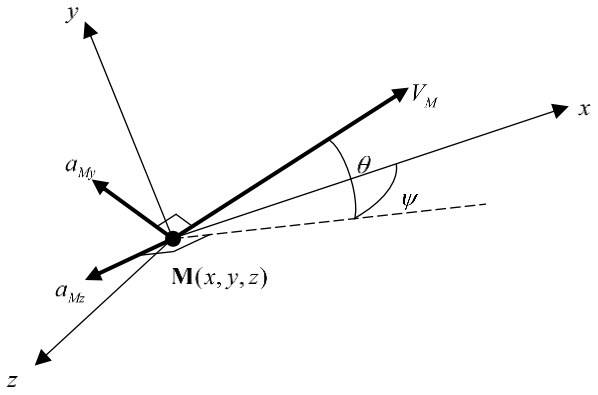
Thay (37) vào (36) ta viết lại phương trình luật dẫn dưới dạng như sau:

$$n_M=6V_M\dot\sigma-\frac{2}{t_{go}}(V_M\psi_F-V_M\psi)$$
$$+\frac{\omega^2t_{go}^2+2\omega t_{go}\sin\omega t_{go}+6\cos\omega t_{go}-6}{\omega^2t_{go}^2}a_S+\frac{6\sin\omega t_{go}-2\omega t_{go}\cos\omega t_{go}-4\omega t_{go}}{\omega^2t_{go}^2}a_C \tag{39}$$

Từ công thức (39) ta thấy, luật dẫn mới được tổng hợp hóa ra gồm ba thành phần: thành phần tiếp cận tỉ lệ; thành phần đáp ứng góc tiếp cận cho trước và thành phần tạo ra cơ động điều hòa.

2.3. Tổng hợp luật cơ động trong không gian cho TLĐH

Bằng cách áp dụng luật cơ động phẳng đồng thời cho mặt phẳng đứng và ngang ta có thể tổng hợp được luật cơ động trong không gian. Trên Hình 2 động hình học không gian của TLĐH được mô tả trong hệ tọa độ dẫn $Oxyz$, trong đó trục Ox có phương trùng với đường ngắm ban đầu. Trong hệ tọa độ dẫn $Oxyz$, các kí hiệu ψ, θ lần lượt là góc vòng và góc nghiêng quỹ đạo, n_{My} là gia tốc trong mặt phẳng đứng vuông góc với véc tơ V_M, n_{Mz} là gia tốc trong mặt phẳng ngang vuông góc với hình chiếu của véc tơ V_M trên mặt phẳng ngang.



Hình 2. Động hình học của TLĐH trong không gian

Các phương trình chuyển động của TLĐH trong hệ tọa độ dẫn $Oxyz$:

$$\dot{x} = V_M \cos\theta\cos\psi;\ \dot{y} = V_M \sin\theta;\ \dot{z} = V_M \cos\theta\sin\psi;\ \dot{\theta} = n_{My}/V_M;\ \dot{\psi} = n_{Mz}/(V_M\cos\theta)$$

Gia tốc lệnh của luật dẫn trong mặt phẳng đứng và mặt phẳng ngang được xác định theo công thức của luật dẫn phẳng (39). Độ lệch pha giữa (a_{Sy}, a_{Cy}) và (a_{Sz}, a_{Cz}) là 90^0. Như vậy, các tham số k, ω xác định quỹ đạo cơ động trong không gian cho TLĐH.

3. Kết quả mô phỏng và thảo luận

Để đánh giá các đặc trưng của luật dẫn tổng hợp được, ta thực hiện mô phỏng khảo sát khả năng vượt hỏa lực phòng không của TLĐH trước sự đánh chặn của TLPK và khả năng gặp mục tiêu theo góc chỉ định. Các tham số quỹ đạo cơ động cho TLĐH gồm $\omega = 0.3\pi$ rad/s, $k = 100$, độ lệch pha 90^0. Điều kiện ban đầu và các ràng buộc cuối của TLĐH được tóm tắt trong Bảng 1. Thông số kỹ thuật của cả TLĐH và TLPK được tóm tắt trong Bảng 2.

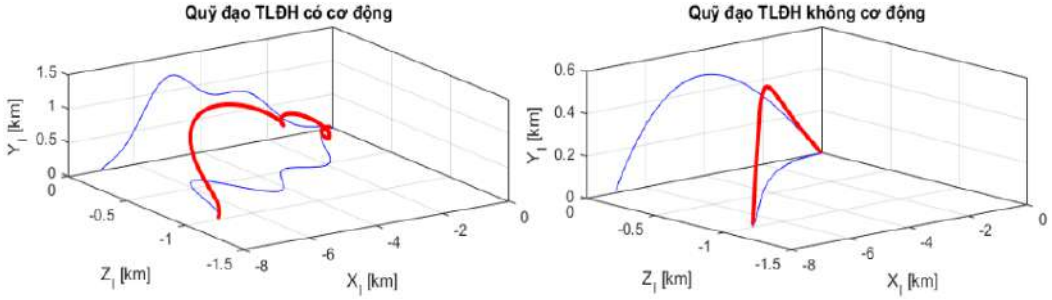
Bảng 1. Điều kiện đầu và ràng buộc cuối của TLĐH

Điều kiện đầu – cuối	(x, y, z) [m]	(θ, φ) [độ]
Điều kiện ban đầu	(-7000, 10, -1000)	(20, 10)
Ràng buộc cuối	(0, 0, 0)	(-10, 0)

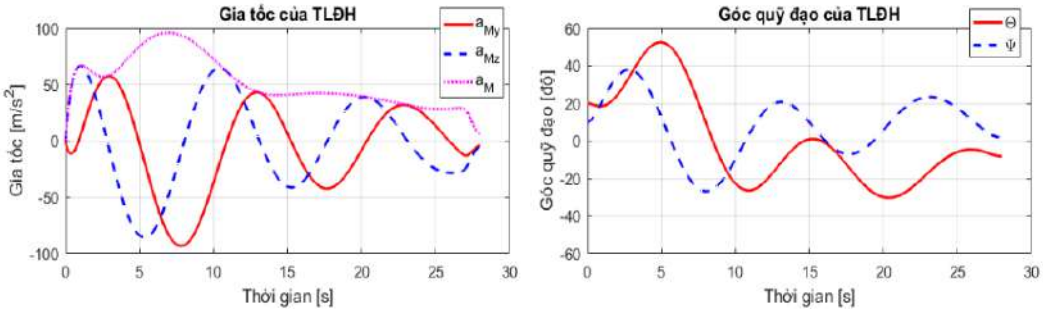
Bảng 2. Thông số kỹ thuật của TLĐH và TLPK

Tên lửa	TLĐH	TLPK
Luật dẫn sử dụng	đa chỉ tiêu	tiếp cận tỉ lệ (N=3)
Tốc độ [m/s]	300	600
Gia tốc cực đại [g]	10	20
Hằng số thời gian [s]	0.5	0.1

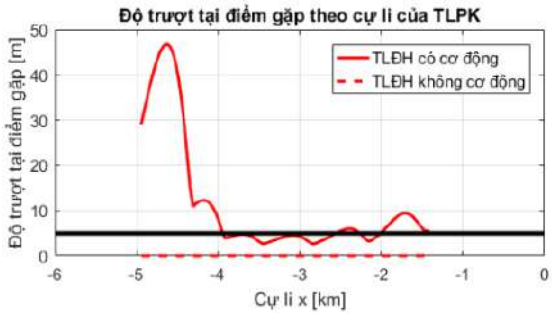
Đầu tiên, khảo sát các đặc trưng bay của TLĐH. Hình 3 cho thấy quỹ đạo cơ động là dạng xoắn ốc trong không gian, còn không cơ động là một đường cong trơn. Hình 4 chỉ ra gia tốc cơ động có dạng hàm điều hòa với biên độ giảm dần về 0, góc quỹ đạo tiến về góc gặp mục tiêu.



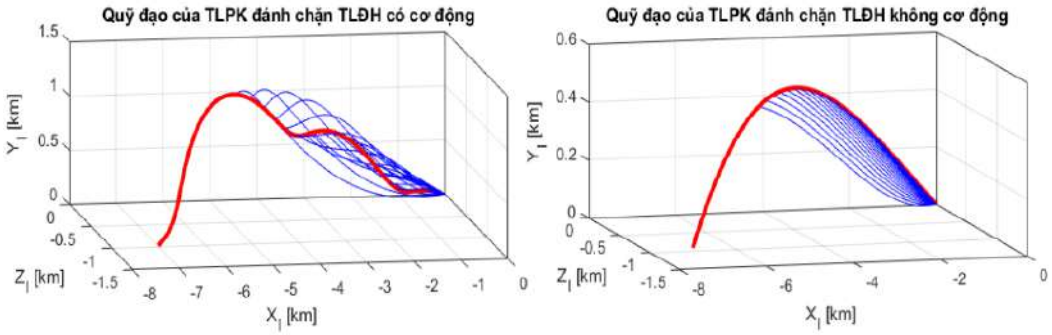
Hình 3. So sánh quỹ đạo của TLĐH có cơ động và không cơ động



Hình 4. Sự thay đổi gia tốc và góc quỹ đạo của TLĐH theo thời gian bay



Hình 5. So sánh độ trượt của TLPK khi TLĐH có và không cơ động



Hình 6. So sánh quỹ đạo của TLPK đánh chặn TLĐH có và không cơ động

Tiếp theo, khảo sát khả năng vượt hỏa lực phòng không của TLĐH. Theo kết quả mô phỏng trong Bảng 3, ta thấy luật dẫn đa chỉ tiêu có thể tăng khả năng vượt cho TLĐH lên tới 51%. Hình 5 so sánh độ trượt của TLPK tại các cự li khác nhau cho cả hai luật dẫn. Hình 6 cho thấy luật đa chỉ tiêu sẽ làm nhiễu loạn quỹ đạo TLPK hơn luật dẫn tiếp cận tỉ lệ (không cơ động).

Bảng 3. So sánh khả năng vượt hỏa lực phòng không giữa các luật dẫn

Luật dẫn	không cơ động	cơ động đa chỉ tiêu

Tổng số TLPK bắn ra	100	100
Số TLPK bắn trượt	0	51
Khả năng vượt của TLĐH	0	51%

4. Kết luận

Luật dẫn đa chỉ tiêu đáp ứng được ba yêu cầu quan trọng: tạo ra quỹ đạo chuyển động phức tạp, đảm bảo gặp mục tiêu theo góc cho trước, gia tốc tiến về 0 khi gặp mục tiêu. Kết quả mô phỏng cho thấy luật đa chỉ tiêu làm tăng khả năng vượt hỏa lực phòng không cho TLĐH trước sự đánh chặn của TLPK và đánh trúng tàu mục tiêu với góc tiếp cận mong muốn.

Tài liệu tham khảo

1. Vũ Hỏa Tiễn, "*Động học các hệ thống điều khiển thiết bị bay*", Nhà xuất bản Quân Đội Nhân Dân, 2013.
2. Под ред. Голубева И.С., Светлова В.Г., "*Проектирование зенитных управляемых ракет*," Москва: Изд-во МАИ, 2001. — 732 с.
3. Bùi Quốc Dũng, Đàm Hữu Nghị, Lê Kỳ Biên, "*Tổng hợp luật cơ động lẩn tránh cho tên lửa đối hạm chống lại tên lửa phòng không của tàu*", Tạp chí nghiên cứu KH&CN quân sự, Số 65, 02-2020.
4. P. Zarchan, "*Tactical and Strategic Missile Guidance"*, Sixth Edition, AIAA Progress in Astronautics and Aeronautics, Vol. 239, Inc., Reston, Virginia, 2012.
5. Y. H. Kim, C. K. Ryoo, and M. J. Tahk *"Guidance synthesis for evasive maneuver of anti-ship missiles against close-in weapon systems",* IEEE Aerosp Electron Syst 46(3):1376–1388, 2010.
6. J. I. Lee, and C. K. Ryoo, *"Impact Angle Control Law with Sinusoidal Evasive Maneuver for Survivability Enhancement",* International Journal of Aeronautical and Space Sciences, 19(2), 2018
7. Bryson, E. A., and Ho, C. Y, "*Applied Optimal Control*," Blaisdell, Waltham, MA, 1969.
8. Phạm Trung Dũng, Lê Xuân Đức, *"Cơ sở điều khiển tối ưu trong các hệ thống kỹ thuật,"* Nhà xuất bản QĐND, 2012.
9. Nguyễn Doãn Phước, "*Lý thuyết điều khiển nâng cao*," Nhà xuất bản khoa học và kỹ thuật, 2009.

Synthesis of multi-objective guidance laws for anti-ship missiles in the terminal homing phase

Abstract: *This paper presents a multi-objective guidance law for anti-ship missiles, designed to: ensure the missile intercepts the target ship at a predefined impact angle; generate sinusoidal maneuver trajectories in a plane and spiral trajectories in 3D space; and ensure that command acceleration converges to zero at intercept. These objectives aim to complicate interception by air defense missiles and enhance the lethality against the target ship. Numerical simulations are conducted to investigate the characteristics and effectiveness of the synthesized guidance law.*

Keywords: Maneuver; Anti-ship missile; Air defense missile.

Nghiên cứu cải thiện độ ổn định và độ chính xác của con quay sợi quang dựa trên kỹ thuật xử lý tín hiệu tiên tiến

Nguyễn Quang Anh[1], Nguyễn Ân Thi[1], Bùi Văn Tiến[1*], Phạm Quốc Thịnh[2]

[1]*Viện Tên lửa và kỹ thuật điều khiển/Học viện Kỹ thuật quân sự*

[2]*Hệ quản lý học viên sau đại học/Học viện Kỹ thuật quân sự*

Email: vantien@lqdtu.edu.vn; Tel.: 0966405898

Tóm tắt

Con quay sợi quang (Fiber Optic Gyroscope - FOG) là cảm biến đo tốc độ góc có độ tin cậy cao và được sử dụng rộng rãi trong các hệ thống dẫn đường, ổn định cho thiết bị bay. Tuy nhiên, độ chính xác của FOG bị ảnh hưởng bởi trôi lệch, nhiễu ngẫu nhiên và các yếu tố môi trường, đặc biệt là nhiệt độ. Bài báo nghiên cứu cải thiện chất lượng tín hiệu FOG thông qua các kỹ thuật xử lý tín hiệu. Dữ liệu đầu ra được phân tích để xác định các thành phần nhiễu đặc trưng, trên cơ sở đó áp dụng thuật toán lọc thích nghi nhằm giảm nhiễu và bù trôi theo thời gian. Phương pháp đề xuất được so sánh với các công bố trước nhằm làm rõ tính mới, đặc biệt trong việc xử lý ảnh hưởng của nhiệt độ. Kết quả mô phỏng được xây dựng với các điều kiện nhiễu và biến thiên môi trường khác nhau để đánh giá hiệu quả thuật toán. Kết quả cho thấy tín hiệu sau xử lý có độ ổn định cao hơn và sai số giảm đáng kể so với tín hiệu ban đầu, góp phần nâng cao hiệu năng của FOG trong các hệ thống dẫn đường và thiết bị tự hành.

Từ khóa: *con quay sợi quang, Fiber Optic Gyroscope (FOG),* x*ử lý tín hiệu, độ trôi, lọc thích nghi.*

1. Đặt vấn đề

Con quay sợi quang (Fiber Optic Gyroscope – FOG) là một trong những cảm biến đo tốc độ góc quan trọng nhất trong hệ thống dẫn đường quán tính hiện đại. FOG hoạt động dựa trên hiệu ứng Sagnac [1,2], trong đó sự chênh lệch pha giữa hai chùm ánh sáng truyền ngược chiều trong cuộn sợi quang tỉ lệ với tốc độ quay của hệ quy chiếu.

Quá trình phát triển của FOG bắt đầu từ những năm 1970 cùng với sự phát triển của công nghệ sợi quang và nguồn laser bán dẫn. Ban đầu, các hệ FOG gặp nhiều hạn chế về nhiễu, suy hao và độ ổn định nguồn sáng. Tuy nhiên, nhờ tiến bộ trong công nghệ sợi quang đơn mode, bộ điều chế pha và kỹ thuật xử lý tín hiệu số, FOG đã đạt được độ chính xác cao và dần thay thế nhiều loại con quay cơ khí truyền thống. Về cấu trúc, một hệ FOG điển hình gồm các thành phần chính: nguồn sáng, bộ chia quang, cuộn sợi quang, bộ điều chế pha và bộ thu quang. Hai chùm sáng truyền ngược chiều trong cuộn sợi sẽ tạo ra sự chênh lệch pha khi hệ quay, từ đó được chuyển đổi thành tín hiệu điện tỉ lệ với tốc độ góc.

Tùy theo cấu trúc và mức độ chính xác, FOG được phân loại thành: FOG vòng hở (Open-loop FOG); FOG vòng kín (Closed-loop FOG); FOG cấp chiến thuật (Tactical grade); FOG cấp dẫn đường chính xác cao (Navigation grade). So với con quay laser vòng (RLG), FOG có ưu điểm là không có bộ phận chuyển động, không xảy ra hiện tượng khóa tần (lock-in effect) và có

* Email: vantien@lqdtu.edu.vn

cấu trúc nhỏ gọn hơn. Đồng thời, so với cảm biến MEMS, FOG có độ ổn định dài hạn tốt hơn, phù hợp cho các hệ thống yêu cầu độ chính xác cao [2].

Mặc dù FOG đã đạt được những tiến bộ đáng kể về cấu trúc và độ chính xác, các thành phần sai số ngẫu nhiên và trôi lệch theo thời gian vẫn là thách thức lớn đối với các hệ thống dẫn đường quán tính. Vì vậy, việc nghiên cứu các phương pháp xử lý tín hiệu nhằm nâng cao độ ổn định của FOG là cần thiết.

Trong hệ thống dẫn đường quán tính, sai số tốc độ góc nếu không được xử lý sẽ tích lũy theo thời gian, dẫn đến sai lệch lớn về vận tốc và vị trí. Đặc biệt, các thành phần nhiễu có ảnh hưởng trực tiếp đến độ chính xác dài hạn của hệ thống. Do đặc điểm làm việc liên tục và thời gian tích lũy dài, ngay cả sai số nhỏ của FOG cũng có thể gây ra sai số định vị đáng kể.

Các nghiên cứu gần đây cho thấy việc áp dụng các kỹ thuật xử lý tín hiệu như phân tích Allan variance [3,4], lọc Kalman [5], lọc thích nghi hoặc các phương pháp xử lý miền thời gian – tần số có thể cải thiện đáng kể chất lượng tín hiệu đầu ra của cảm biến mà không cần thay đổi cấu trúc phần cứng. Đây là hướng tiếp cận khả thi, chi phí thấp và phù hợp với điều kiện triển khai thực tế. Xuất phát từ những vấn đề trên, bài báo tập trung phân tích đặc tính nhiễu của FOG nhằm xác định các thành phần sai số chủ yếu. Trên cơ sở đó, một thuật toán xử lý tín hiệu được xây dựng nhằm giảm nhiễu và bù độ trôi theo thời gian, qua đó nâng cao độ ổn định và độ chính xác của cảm biến. Kết quả nghiên cứu góp phần cải thiện hiệu năng của FOG trong các hệ thống dẫn đường quán tính và thiết bị tự hành.

2. Cơ sở lý thuyết và mô hình sai số của FOG

Con quay sợi quang (Fiber Optic Gyroscope – FOG) hoạt động dựa trên hiệu ứng Sagnac, trong đó sự chênh lệch pha giữa hai chùm ánh sáng truyền ngược chiều trong một cuộn sợi quang tỉ lệ với tốc độ quay của hệ quy chiếu. Khi hệ thống quay với tốc độ góc Ω, một chùm sáng sẽ truyền theo chiều quay và chùm còn lại truyền ngược chiều quay. Do quãng đường truyền hiệu dụng khác nhau, hai chùm sáng sẽ xuất hiện độ lệch pha tại đầu ra giao thoa. Các từ viết tắt, công thức, đơn vị đo, bảng biểu, hình, đồ thị theo quy định ở phần 3 phía dưới.

Độ lệch pha Sagnac được xác định bởi:

$$\Delta\phi = \frac{8\pi A\Omega}{\lambda c} \tag{1}$$

trong đó:

A - là diện tích vòng sợi quang;

Ω - là tốc độ góc của hệ;

λ - là bước sóng ánh sáng;

c - là vận tốc ánh sáng trong chân không.

Từ độ lệch pha thu được, hệ thống quang - điện sẽ chuyển đổi thành tín hiệu điện tỉ lệ với tốc độ quay. Nhờ cấu trúc không có bộ phận cơ khí chuyển động, FOG có độ tin cậy cao, khả năng làm việc bền vững trong môi trường rung động và tuổi thọ lớn. Tùy theo phương thức xử lý tín hiệu, FOG có thể được thiết kế dưới dạng vòng hở hoặc vòng kín. Trong đó, cấu trúc vòng kín thường

cho độ tuyến tính và độ chính xác cao hơn, được sử dụng phổ biến trong các hệ thống dẫn đường quán tính.

Mặc dù có nhiều ưu điểm, tín hiệu đầu ra của FOG vẫn chịu ảnh hưởng của nhiều thành phần sai số khác nhau. Mô hình tổng quát của tốc độ góc đo được có thể biểu diễn như sau:

$$\omega_{out}(t) = \omega_{true}(t) + b(t) + n(t) \tag{2}$$

trong đó:

$\omega_{out}(t)$ - là tốc độ góc thực;

$b(t)$ - là thành phần trôi lệch theo thời gian (bias drift);

$n(t)$ - nhiễu ngẫu nhiên.

Thành phần trôi lệch $b(t)$ có thể thay đổi theo nhiệt độ, thời gian làm việc và các yếu tố môi trường. Đây là nguyên nhân chính gây ra sai số tích lũy trong hệ thống dẫn đường quán tính. Khi tích phân tốc độ góc để tính góc quay và tiếp tục tích phân để xác định vận tốc, vị trí, sai số bias dù nhỏ cũng có thể làm sai lệch đáng kể kết quả sau một khoảng thời gian hoạt động.

Các thành phần nhiễu này thường được phân tích thông qua phương pháp Allan variance nhằm xác định đặc tính thống kê và mức độ ảnh hưởng của từng loại sai số.

Do đặc điểm làm việc liên tục của hệ thống dẫn đường quán tính, các sai số trên nếu không được xử lý sẽ tích lũy theo thời gian, làm suy giảm đáng kể độ chính xác của hệ thống. Vì vậy, việc xây dựng các phương pháp xử lý tín hiệu để giảm nhiễu và bù trôi lệch là cần thiết nhằm nâng cao hiệu năng tổng thể của FOG.

Trong điều kiện làm việc thực tế, ngoài các thành phần nhiễu ngẫu nhiên và trôi lệch nội tại của cảm biến, FOG còn chịu ảnh hưởng đáng kể của biến thiên nhiệt độ môi trường. Khi tồn tại gradient nhiệt độ dọc theo chiều dài cuộn sợi quang, hiện tượng giãn nở nhiệt và thay đổi chiết suất không đồng đều sẽ làm thay đổi chiều dài quang học của sợi. Sự không đối xứng này dẫn đến sai số pha giữa hai chùm sáng truyền ngược chiều, được gọi là hiệu ứng Shupe.

Sai số pha do hiệu ứng Shupe có thể được biểu diễn dưới dạng:

$$\Delta\phi_{Shupe} = K\int_0^L \frac{\partial T(z,t)}{\partial t}\left(z - \frac{L}{2}\right)dz \tag{3}$$

trong đó L là chiều dài cuộn sợi, $T(z,t)$ là phân bố nhiệt độ theo vị trí và thời gian và K là hệ số phụ thuộc vào đặc tính vật liệu sợi quang và bước sóng ánh sáng sử dụng.

Từ biểu thức trên có thể thấy rằng sai số pha chỉ xuất hiện khi phân bố nhiệt độ theo chiều dài sợi không đối xứng qua tâm cuộn $z = L/2$. Đáng chú ý, sai số này tỷ lệ với tốc độ thay đổi nhiệt độ theo thời gian $\partial T/\partial t$, chứ không phụ thuộc trực tiếp vào giá trị tuyệt đối của nhiệt độ. Do đó, trong các điều kiện môi trường có biến thiên nhiệt nhanh, hiệu ứng Shupe có thể gây ra sai số đáng kể.

Sai số pha do Shupe sau khi được chuyển đổi qua hệ số tỷ lệ của FOG sẽ biểu hiện dưới dạng trôi lệch tốc độ góc. Khi tích phân trong hệ thống dẫn đường quán tính, thành phần trôi này có thể tích lũy theo thời gian, làm suy giảm độ chính xác định hướng và định vị của hệ thống.

Trong thực tế, hiệu ứng Shupe được xem là một trong những nguyên nhân chính gây trôi bias dài hạn trong FOG, đặc biệt đối với các ứng dụng ngoài trời hoặc môi trường có biến thiên nhiệt độ lớn. Vì vậy, bên cạnh các giải pháp thiết kế cấu trúc cuộn sợi đối xứng nhiệt, việc áp dụng các kỹ thuật xử lý tín hiệu nhằm ước lượng và bù trôi nhiệt là hướng nghiên cứu cần thiết.

3. Tổng quan về các phương pháp để khắc phục sai số trôi

3.1. Sử dụng thuật toán lọc Kalman

Bộ lọc Kalman là phương pháp ước lượng trạng thái tối ưu cho các hệ thống tuyến tính chịu tác động của nhiễu Gauss trắng. Trong bài toán xử lý sai số trôi của FOG, bộ lọc Kalman được sử dụng để đồng thời ước lượng tốc độ góc thực và thành phần bias theo thời gian.

Mô hình không gian trạng thái của hệ được xây dựng với vector trạng thái:

$$X_k = \begin{bmatrix} \omega_k \\ b_k \end{bmatrix} \tag{4}$$

trong đó: ω_k - là tốc độ góc thực tại thời điểm k, b_k - là thành phần trôi (bias).

Phương trình trạng thái được mô tả dưới dạng:

$$X_{k+1} = F\begin{bmatrix} \omega_k \\ b_k \end{bmatrix} + w_k \tag{5}$$

trong đó F là ma trận chuyển trạng thái, w_k - là nhiễu quá trình phản ánh biến thiên ngẫu nhiên của bias.

Phương trình đo được biểu diễn:

$$z_k = F\begin{bmatrix} \omega_k \\ b_k \end{bmatrix} + \upsilon_k \tag{6}$$

với z_k - là tốc độ góc đo được từ FOG, υ_k - là nhiễu đo.

Trong trường hợp đơn giản, có thể giả thiết bias biến thiên chậm theo mô hình Gauss–Markov bậc một hoặc mô hình random walk. Bộ lọc Kalman thực hiện hai bước chính: dự đoán (prediction) và cập nhật (update), qua đó liên tục hiệu chỉnh giá trị ước lượng của b_k. Tốc độ góc sau hiệu chỉnh được tính:

$$\omega_{corr,k} = z_k - \hat{b}_k \tag{7}$$

Phương pháp này cho phép giảm đáng kể ảnh hưởng của bias drift và cải thiện độ ổn định dài hạn của tín hiệu FOG mà không cần thay đổi cấu trúc phần cứng.

Trong thực tế, các mối quan hệ giữa nhiệt độ, gradient nhiệt và sai số pha trong FOG thường mang tính phi tuyến. Khi đó, bộ lọc Kalman tuyến tính không còn đáp ứng đầy đủ yêu cầu mô hình hóa hệ thống. Bộ lọc Kalman mở rộng (Extended Kalman Filter – EKF) được sử dụng để xử lý các hệ phi tuyến bằng cách tuyến tính hóa xung quanh điểm làm việc hiện tại.

Đối với FOG, tính phi tuyến có thể xuất phát từ mối quan hệ phi tuyến giữa nhiệt độ và bias do đặc tính vật liệu sợi quang, ảnh hưởng phi tuyến của gradient nhiệt độ đến sai số pha (liên quan đến hiệu ứng Shupe), các hiệu ứng vật lý phức tạp như thay đổi chiết suất theo nhiệt độ và ứng suất cơ học.

Mô hình phi tuyến tổng quát có thể biểu diễn:

$$X_{k+1} = f\left(X_k, \upsilon_k\right) + \omega_k;\ z_k = h\left(X_k\right) + \upsilon_k$$

Trong đó $f(\cdot)$và $h(\cdot)$là các hàm phi tuyến. EKF sử dụng khai triển Taylor bậc nhất để xấp xỉ tuyến tính thông qua ma trận Jacobi, sau đó áp dụng quy trình dự đoán – cập nhật tương tự bộ lọc Kalman chuẩn.

Việc sử dụng EKF cho phép mô hình hóa chính xác hơn ảnh hưởng của nhiệt độ và các yếu tố môi trường đến bias drift, đặc biệt trong điều kiện biến thiên nhiệt nhanh. Nhờ đó, khả năng ước lượng và bù sai số được cải thiện so với bộ lọc tuyến tính.

3.2. Mô hình EKF khi nhiệt độ được đưa vào vector trạng thái

Trong điều kiện làm việc thực tế, sai số trôi của FOG có mối liên hệ chặt chẽ với biến thiên nhiệt độ và gradient nhiệt dọc theo cuộn sợi quang. Để mô hình hóa đầy đủ ảnh hưởng này, một cách tiếp cận nâng cao là đưa nhiệt độ vào vector trạng thái và thực hiện ước lượng đồng thời vận tốc góc, bias drift và trạng thái nhiệt của hệ thống. Phương pháp này cho phép xây dựng mô hình bù sai số dựa trên cơ chế vật lý thay vì chỉ dựa trên xử lý tín hiệu thuần túy.

Giả sử cuộn sợi quang được trang bị mcảm biến nhiệt đặt tại các vị trí khác nhau dọc theo chiều dài sợi, vector trạng thái của hệ tại thời điểm kđược xây dựng như sau:

$$\mathrm{X}_k = \begin{bmatrix} \omega_k \\ b_k \\ T_{1,k} \\ T_{2,k} \\ \vdots \\ T_{m,k} \end{bmatrix}$$

trong đó ω_k là vận tốc góc thực, b_k là thành phần trôi, và $T_{i,k}$ là nhiệt độ tại vị trí thứ i trên cuộn sợi.

Động học của hệ có thể được mô tả theo mô hình rời rạc:

$$\omega_{k+1} = \omega_k + w_{\omega,k}$$
$$b_{k+1} = b_k + w_{b,k}$$
$$T_{i,k+1} = T_{i,k} + w_{T_i,k}, i = 1, \dots, m$$

Trong đó $w_{\omega,k}$, $w_{b,k}$ và $w_{T_i,k}$ là các nhiễu quá trình phản ánh các biến thiên chưa mô hình hóa. Thành phần bias có thể được xem như một quá trình ngẫu nhiên biến thiên chậm, trong khi nhiệt độ được mô hình hóa theo dạng random walk hoặc mô hình truyền nhiệt bậc nhất tùy mức độ chi tiết mong muốn.

Phương trình đo của hệ bao gồm tín hiệu tốc độ góc thô từ FOG và các giá trị nhiệt độ đo được:

$$z_k^{\omega} = \omega_k + b_k + \Delta\omega_{S,k}(\mathbf{T}_k) + v_{\omega,k}$$
$$z_k^{T_i} = T_{i,k} + v_{T_i,k}$$

Trong đó $\Delta\omega_{S,k}(\mathbf{T}_k)$ là thành phần sai số do hiệu ứng Shupe và các tác động nhiệt khác, phụ thuộc vào phân bố nhiệt độ trên cuộn sợi; $v_{\omega,k}$và $v_{T_i,k}$ là nhiễu đo.

Thành phần sai số nhiệt $\Delta\omega_{S,k}$có thể được xấp xỉ theo dạng hàm của tốc độ biến thiên nhiệt và độ không đối xứng của phân bố nhiệt quanh tâm cuộn sợi. Điều này phản ánh đúng bản chất vật lý của hiệu ứng Shupe, trong đó sai số tỷ lệ với đạo hàm theo thời gian của nhiệt độ và phụ thuộc vào gradient nhiệt dọc theo chiều dài sợi.

Do tồn tại các quan hệ phi tuyến giữa nhiệt độ và sai số pha, mô hình trên thuộc lớp hệ phi tuyến. Vì vậy, bộ lọc Kalman mở rộng (Extended Kalman Filter – EKF) được sử dụng để thực hiện ước lượng trạng thái. EKF tiến hành tuyến tính hóa mô hình phi tuyến quanh điểm làm việc hiện tại thông qua ma trận Jacobi, sau đó áp dụng quy trình dự đoán và cập nhật tương tự bộ lọc Kalman tuyến tính.

Sau khi bộ lọc hội tụ, vận tốc góc đã được bù sai số được xác định theo:

$$\omega_{corr,k} = z_k^{\omega} - \hat{b}_k - \Delta\hat{\omega}_{S,k}$$

Cách tiếp cận này cho phép ước lượng đồng thời vận tốc góc thực, bias drift và phân bố nhiệt độ của cuộn sợi, từ đó nâng cao khả năng bù sai số trong điều kiện môi trường biến thiên nhiệt độ mạnh. Tuy nhiên, do số chiều trạng thái tăng theo số lượng cảm biến nhiệt, độ phức tạp tính toán và yêu cầu phần cứng cũng tăng đáng kể. Vì vậy, phương pháp này phù hợp với các hệ FOG yêu cầu độ ổn định cao hoặc trong các nghiên cứu chuyên sâu về ảnh hưởng nhiệt đến sai số trôi.

4. Mô phỏng và kết quả

Để đánh giá hiệu quả của thuật toán EKF trong việc bù sai số trôi do ảnh hưởng nhiệt, các tham số mô hình và nhiễu được thiết lập dựa trên đặc tính điển hình của hệ FOG nghiên cứu. Các tham số này được lựa chọn nhằm phản ánh đúng bản chất vật lý của hệ thống đồng thời đảm bảo tính ổn định của bộ lọc.

4.1. Thiết lập mô phỏng và tham số

Để đánh giá khả năng ước lượng vận tốc quay và bù sai số trôi của các phương pháp xử lý tín hiệu, nhóm tác giả xây dựng mô phỏng tín hiệu FOG theo mô hình đo:

$$z_k = \omega_k + b_k + v_k$$

trong đó ω_k là giá trị vận tốc quay thực, b_k là thành phần trôi (bias drift) phụ thuộc nhiệt độ, và v_k là nhiễu đo.

Ảnh hưởng nhiệt độ lên bias được mô hình hóa theo dạng phi tuyến bậc hai quanh nhiệt độ tham chiếu:

$$b(T) = \alpha(T - T_{ref}) + \beta(T - T_{ref})^2$$

Trong đó:

$\alpha = 0.001$(rad/s/°C): hệ số nhạy bậc một của độ trôi theo nhiệt độ. Tham số này phản ánh mức thay đổi tuyến tính của bias khi nhiệt độ thay đổi.

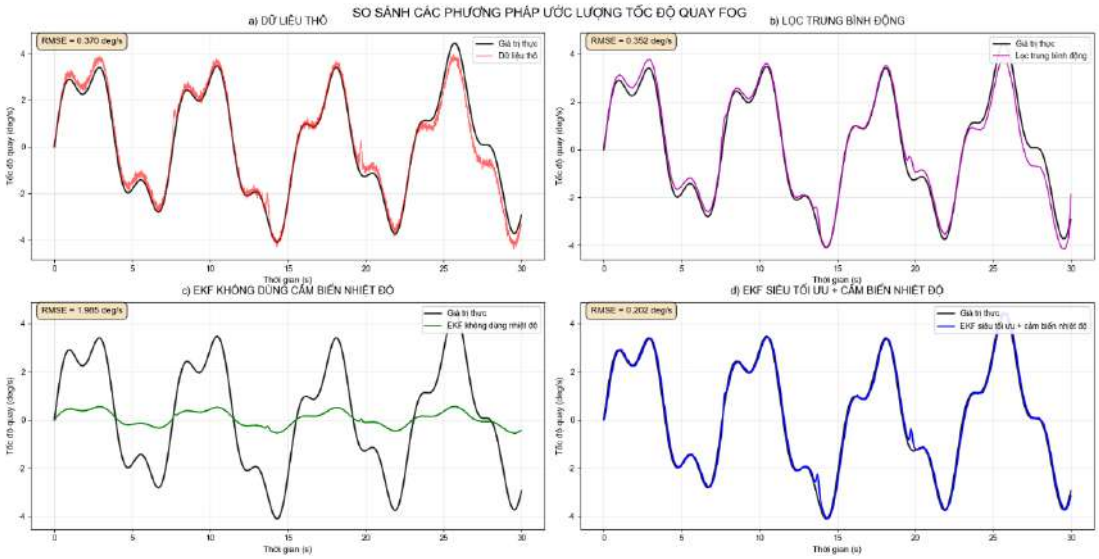
$\beta = 2 \times 10^{-6}$: hệ số nhạy bậc hai, biểu diễn ảnh hưởng phi tuyến của nhiệt độ.

$T_{ref} = 25.0°C$: nhiệt độ tham chiếu, tương ứng với điều kiện hiệu chuẩn ban đầu của cảm biến.

Việc đưa thêm thành phần bậc hai cho phép mô hình hóa chính xác hơn sai số trong điều kiện biến thiên nhiệt độ lớn.

4.2. Phân tích kết quả

Hình so sánh vận tốc quay cho thấy dữ liệu thô bám xu hướng chung nhưng bị nhiễu và xuất hiện sai lệch đáng kể ở một số vùng có biến thiên. Lọc trung bình động giúp làm trơn tín hiệu nhưng vẫn còn sai số do không xử lý thành phần trôi/bias có xu hướng chậm.

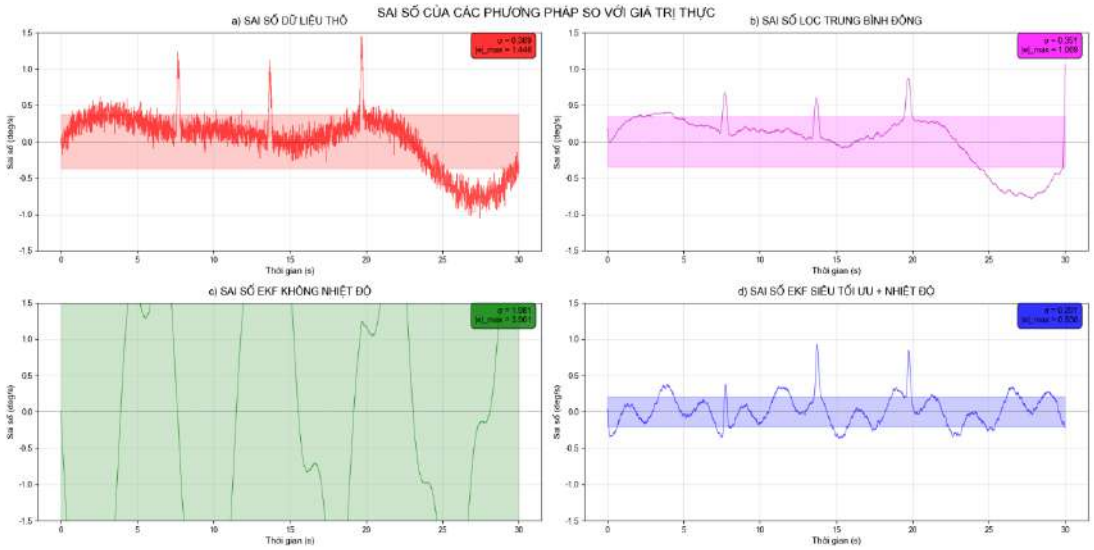


Hình 1. So sánh các phương pháp ước lượng tốc độ quay FOG

Trường hợp EKF không đưa nhiệt độ vào mô hình cho kết quả kém, sai số tăng mạnh và bám tín hiệu thực không tốt. Nguyên nhân là bias drift trong kịch bản mô phỏng có tương quan với biến thiên nhiệt; khi không quan sát/ước lượng trạng thái nhiệt, EKF bị "mô hình sai" nên ước lượng bias không đúng, dẫn tới sai lệch tích lũy.

Ngược lại, EKF có sử dụng nhiệt độ cho đường ước lượng gần trùng với giá trị thực trong toàn bộ khoảng thời gian mô phỏng.

Đánh giá theo RMSE (deg/s) trên toàn miền: Dữ liệu thô: 0.370; Lọc trung bình động: 0.352; EKF không dùng nhiệt độ: 1.985; EKF có nhiệt độ: 0.202. Kết quả cho thấy EKF có nhiệt độ giúp giảm RMSE khoảng 45% so với dữ liệu thô và khoảng 43% so với lọc trung bình động trong kịch bản mô phỏng này.

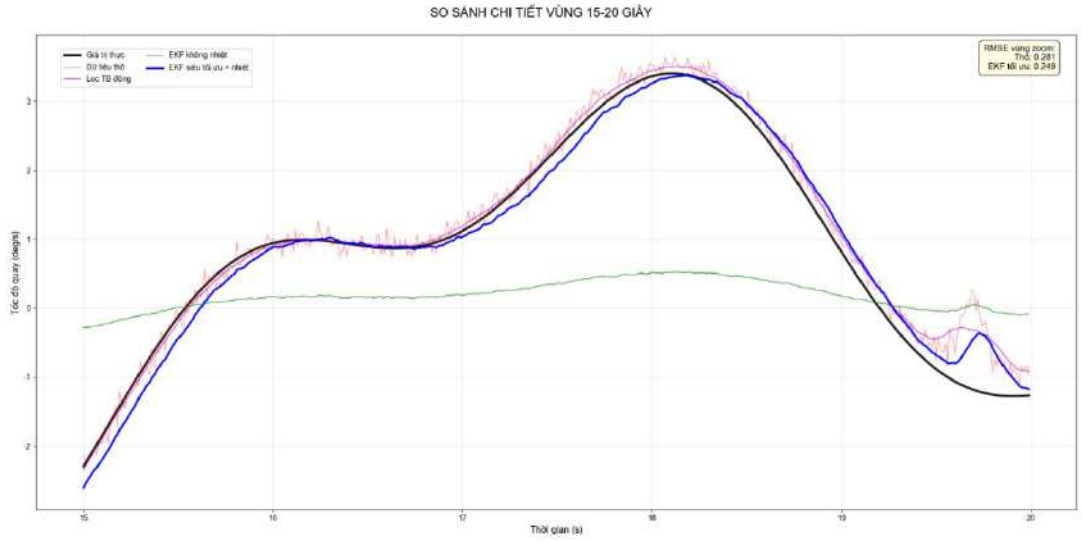


Hình 2. Sai số của các phương pháp so với giá trị thực

Từ các đồ thị sai số so với giá trị thực, có thể nhận thấy:

Dữ liệu thô có độ phân tán sai số ở mức vừa, nhưng xuất hiện một số đỉnh sai số cục bộ.

Lọc trung bình động giảm nhiễu cao tần, song vẫn còn sai số hệ thống (xu hướng lệch chậm) do không tách được bias drift. EKF không dùng nhiệt độ cho sai số lớn và biến thiên bất thường, thể hiện rõ sự không phù hợp mô hình. EKF có nhiệt độ cho sai số dao động nhỏ quanh 0, ổn định hơn rõ rệt.



Hình 3. So sanh chi tiết vùng 15-20 giây.

Ở vùng phóng đại 15–20 s, EKF có nhiệt độ tiếp tục thể hiện ưu thế bám tín hiệu tốt hơn; RMSE trong vùng này giảm từ 0.281 (dữ liệu thô) xuống 0.249 (EKF có nhiệt độ).

4.3. Kết luận

Bài báo đã mô phỏng và so sánh các phương pháp ước lượng vận tốc quay của FOG. Kết quả cho thấy lọc trung bình động chỉ cải thiện nhẹ do chủ yếu làm trơn nhiễu, trong khi EKF không sử dụng thông tin nhiệt độ có thể cho kết quả kém khi bias drift phụ thuộc nhiệt. Phương pháp EKF có đưa nhiệt độ vào mô hình cho hiệu quả tốt nhất, giảm đáng kể RMSE và

độ phân tán sai số so với dữ liệu thô, đồng thời duy trì khả năng bám tín hiệu tốt trong vùng biến thiên.

Tài liệu tham khảo

1. H. C. Lefèvre. (2014). The Fiber-Optic Gyroscope, 2nd ed. Boston, MA, USA: Artech House,
2. X. Lin, W. Han, K. Chen, and W. Zhang. (2020). On the Development and Application of FOG. Gyroscopes – Principles and Applications. Eds. London, UK: IntechOpen. Doi: 10.5772/intechopen.88542.
3. D. W. Allan. (1966). Statistics of atomic frequency standards. Proceedings of the IEEE. Vol. 54, no. 2, pp. 221–230.
4. IEEE Std 952-1997. IEEE Standard Specification Format Guide and Test Procedure for Single-Axis Interferometric Fiber Optic Gyros, IEEE.
5. D. H. Titterton, J. L. Weston. (2004). Strapdown Inertial Navigation Technology. 2nd ed., IET.

Improving the Stability and Accuracy of Fiber Optic Gyroscopes Using Advanced Signal Processing Techniques

Abstract: *Fiber Optic Gyroscopes (FOGs) are high-reliability angular rate sensors widely used in navigation and stabilization systems for aerial vehicles. However, their accuracy is affected by bias drift, random noise, and environmental factors, particularly temperature. This paper investigates methods to improve FOG signal quality through advanced signal processing techniques. The sensor output is analyzed to identify characteristic noise components, forming the basis for applying an adaptive filtering algorithm to reduce noise and compensate for time-varying drift. The proposed approach is compared with previous studies to highlight its novelty, especially in handling temperature effects. Simulation scenarios are developed under various noise conditions and environmental variations to evaluate the effectiveness of the algorithm. The results demonstrate that the processed signal achieves higher stability and significantly reduced error compared to the original signal, thereby enhancing the performance of FOGs in navigation and autonomous systems..*

Keywords: Fiber Optic Gyroscope (FOG); Signal Processing; Bias Drift; Adaptive Filtering.

NGHIÊN CỨU CÁC PHƯƠNG PHÁP ĐIỀU KHIỂN UAV BẦY ĐÀN DẠNG QUADTOTOR

Đặng Quốc Đạt[1] *, Doãn Văn Minh[2], Nguyễn Đình Dũng [3]

[1]*Hệ 2, Học viện Kỹ thuật quân sự,*

[2]*Viện Tên lửa và kỹ thuật điều khiển, Học viện Kỹ thuật quân sự,*

[3]*Khoa Hàng không vũ trụ, Học viện Kỹ thuật quân sự.*

Email: Dangquocdat@lqdtu.edu.vn; Tel: 0982801084

Tóm tắt:

Bầy đàn UAV dạng quadrotor đang thu hút nhiều quan tâm nghiên cứu nhờ khả năng bao phủ không gian rộng, độ tin cậy và tính linh hoạt cao so với UAV đơn lẻ. Thách thức cốt lõi là thiết kế các chiến lược điều khiển đội hình cho phép đàn UAV hình thành, duy trì và tái cấu trúc trong môi trường phức tạp, dưới ràng buộc động lực học phi tuyến, hạn chế truyền thông và yêu cầu an toàn nghiêm ngặt. Bài báo trình bày tổng quan về các phương pháp điều khiển bầy đàn quadrotor trong những năm gần đây, tập trung vào các phương pháp gồm leader–follower, đồng thuận, cấu trúc ảo và trường thế nhân tạo, điều khiển dựa trên hành vi, điều khiển phân tán–bền vững–thích nghi, thuật toán tối ưu bầy đàn PSO, cũng như các phương pháp kết hợp trí tuệ nhân tạo và học tăng cường sâu. Bài báo phân tích ưu nhược điểm của từng nhóm phương pháp và thảo luận các thách thức mở liên quan đến mở rộng quy mô, bảo đảm an toàn và tích hợp học máy có ràng buộc ổn định.

Từ khóa: UAV quadrotor bầy đàn; đội hình UAV quadrotor; leader–follower; điều khiển đồng thuận; điều khiển dựa trên hành vi; PSO; cấu trúc ảo; trường thế nhân tạo.

1. Mở đầu

Trong khoảng 5–10 năm trở lại đây, các hệ UAV bầy đàn đã được ứng dụng rộng rãi trong nhiều nhiệm vụ như giám sát diện rộng, nông nghiệp chính xác, tìm kiếm–cứu nạn, truyền thông khẩn cấp và trình diễn ánh sáng. Các khảo sát gần đây cho thấy số lượng công trình nghiên cứu về điều khiển bầy đàn và đội hình UAV tăng nhanh, đặc biệt đối với nền tảng quadrotor nhờ cấu trúc cơ khí đơn giản, khả năng cất–hạ cánh thẳng đứng và tính cơ động cao trong không gian ba chiều [1].

Trong bối cảnh đó, điều khiển đội hình và điều khiển bầy đàn trở thành bài toán trung tâm với các mục tiêu chính gồm hình thành và duy trì đội hình mong muốn, theo dõi quỹ đạo hoặc mục tiêu chung, bảo đảm tránh va chạm, và duy trì tính bền vững trước nhiễu, sai số mô hình và gián đoạn truyền thông. Các khảo sát hiện có đã phân loại nhiều chiến lược điều khiển khác nhau, từ các phương pháp điều khiển cổ điển và phi tuyến (PID, trượt, backstepping, MPC, điều khiển bền vững) đến các chiến lược phối hợp như leader–follower, đồng thuận, cấu trúc ảo và phương pháp dựa trên trường thế nhân tạo [2].

Đối với đội hình quadrotor, các thách thức chính bao gồm động lực học phi tuyến và dưới điều khiển, ràng buộc truyền thông và topology mạng, yêu cầu điều khiển phân tán có khả năng mở rộng, cùng các ràng buộc an toàn nghiêm ngặt trong không gian ba chiều. Để giải quyết các vấn đề này, nhiều hướng tiếp cận đã được đề xuất, như sử dụng ma trận quay trên SO(3) nhằm tránh hiện tượng Gimbal lock, điều khiển phân tán không cần thông tin vận tốc, tích hợp trường thế nhân tạo cải tiến với điều khiển trượt, cũng như ứng dụng các thuật toán tối ưu bầy đàn và học tăng cường sâu trong điều khiển đàn UAV [3, 4].

* Email: Dangquocdat@lqdtu.edu.vn

Trong khuôn khổ bài báo tổng quan này, chúng tôi tập trung vào đàn UAV dạng quadrotor, với mục tiêu phân loại các phương pháp điều khiển bầy đàn tiêu biểu, thảo luận các hướng mở rộng kết hợp tối ưu và trí tuệ nhân tạo, và cung cấp cơ sở tham khảo cho việc lựa chọn phương pháp phù hợp với từng bài toán ứng dụng cụ thể.

2. Phân loại các phương pháp điều khiển UAV bầy đàn dạng quadrotor

2.1. Phương pháp leader–follower

Trong mô hình leader–follower, một hoặc một nhóm UAV được chỉ định làm leader, chịu trách nhiệm bám quỹ đạo tham chiếu hoặc mục tiêu nhiệm vụ tổng thể. Các UAV còn lại (followers) điều chỉnh chuyển động của mình để duy trì vị trí tương đối mong muốn so với leader hoặc các UAV tham chiếu khác, từ đó hình thành và duy trì đội hình bầy đàn [4, 5, 6]. Leader có thể là UAV vật lý hoặc một thực thể ảo sinh quỹ đạo tham chiếu toàn cục.

Một dạng luật điều khiển vị trí đơn giản cho follower i được viết như sau:

$$u_i = -k_p\left(p_i - p_L - R_L\delta_i\right) - k_v\left(v_i - v_L\right) \quad (1)$$

trong đó: p_L, v_L, R_L lần lượt là vị trí, vận tốc và ma trận quay của leader; δ_i là véc tơ đội hình; $k_p, k_v > 0$ là hệ số điều khiển tỉ lệ.

Bảng 1 tóm tắt một số công trình tiêu biểu về điều khiển leader–follower cho quadrotor bầy đàn, làm nổi bật các kết quả đạt được cùng với ưu và nhược điểm của từng cách tiếp cận.

Bảng 1. Một số công trình nghiên cứu tiêu biểu về điều khiển leader–follower cho đội hình quadrotor

Phương pháp điều khiển	Kết quả chính	Ưu điểm	Nhược điểm
Leader–follower + backstepping thích nghi [4].	Đội hình UAV ổn định, bám quỹ đạo, sai số trong giới hạn định trước.	Hiệu suất đảm bảo; ổn định Lyapunov; phù hợp với các nhiệm vụ bám quỹ đạo chính xác.	Cấu trúc phức tạp; phụ thuộc mạnh vào leader.
Leader–follower + điều khiển lai [5].	Ổn định đội hình trong điều kiện trễ truyền thông và nhiễu	Bền vững cao; phù hợp môi trường truyền thông không lý tưởng	Thiết kế tham số phức tạp; khó mở rộng đàn lớn
Leader–follower + mô hình đặc trưng [6].	Cải thiện độ mượt và giảm dao động bám đội hình	Các UAV follower bám theo leader mà không cần quỹ đạo chi tiết.	Phụ thuộc trích xuất đặc trưng; nhạy với nhiễu.

2.2. Phương pháp đồng thuận

Thuật toán điều khiển đồng thuận (consensus-based control) dựa trên nguyên lý mỗi UAV điều chỉnh trạng thái của mình dựa trên chênh lệch trạng thái với các UAV lân cận, nhằm đạt được một mục tiêu chung như đồng bộ vị trí, vận tốc, pha hoặc duy trì một đội hình dịch chuyển [7, 8].

Xét hệ bậc một giản lược:

$$\dot{x}_i = u_i, \qquad u_i = -\sum_{j \in N_i} a_{ij}(x_i - x_j) \tag{2}$$

Trong đó: N_i là tập láng giềng của UAV i. Nếu đồ thị tương tác là liên thông (đối với đồ thị vô hướng) hoặc chứa một cây bao trùm (đối với đồ thị có hướng), thì thuật toán trên đảm bảo mọi x_i hội tụ về cùng một giá trị, tức đạt được trạng thái đồng thuận.

Đối với bài toán đội hình, thêm các véc tơ đội hình $\delta_i \in \mathbb{R}^3$ để đạt đồng thuận có dịch:

$$u_i = -\sum_{j \in Ni} a_{ij}\left[(p_i - p_j) - (\delta_i - \delta_j)\right] \tag{3}$$

Khi đó, mỗi UAV điều chỉnh vị trí của mình sao cho chênh lệch vị trí với các lân cận bằng đúng chênh lệch hình dạng đội hình mong muốn. Khi thuật toán hội tụ, toàn bộ hệ đạt được hình dạng đội hình định trước, có thể dịch chuyển như một khối cứng nếu bổ sung thêm thành phần bám quỹ đạo chung.

Bảng 2 trình bày các công trình tiêu biểu về điều khiển đồng thuận cho đội hình UAV, nhấn mạnh đặc điểm phân tán, khả năng mở rộng cũng như các hạn chế trong việc duy trì đội hình hình học chính xác.

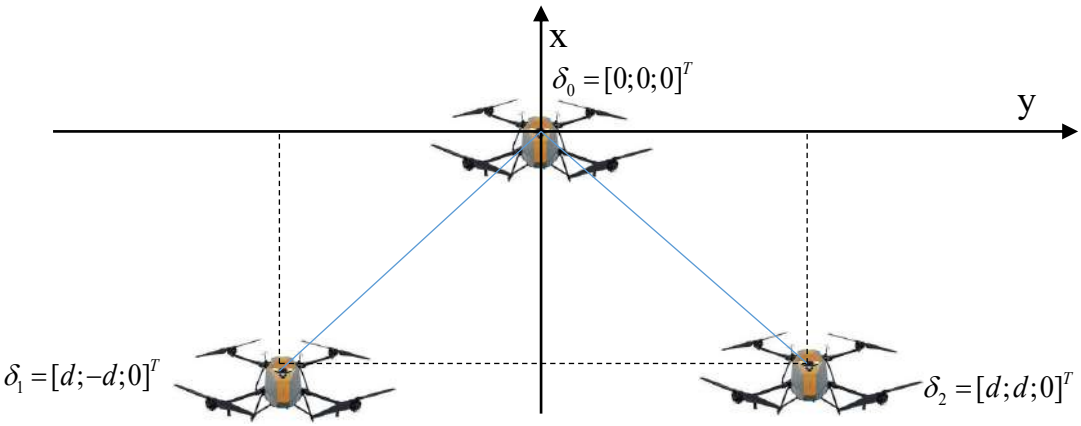
Bảng 2. Một số công trình tiêu biểu về điều khiển đồng thuận cho UAV bầy đàn

Phương pháp điều khiển	Kết quả chính	Ưu điểm	Nhược điểm
Điều khiển hành vi bầy đàn dựa trên đồng thuận cho hệ bậc hai [7]	Mô hình và thuật toán cho hệ đa tác tử bậc hai: giữ khoảng cách và đồng bộ vận tốc.	Phân tán hoàn toàn; mở rộng bầy UAV tốt.	Khó giữ hình dạng đội hình; nhạy với nhiễu.
Điều khiển đồng thuận cho đội hình UAV [8]	Mở rộng đồng thuận cho đội hình UAV, ổn định Lyapunov.	Dễ mở rộng; phù hợp cấu trúc mạng thay đổi.	Phụ thuộc vào cấu trúc mạng.

2.3. Cấu trúc ảo và trường thế nhân tạo

2.3.1. Cấu trúc ảo và trường thế

Phương pháp cấu trúc ảo (virtual structure) coi toàn bộ đàn UAV như một vật thể cứng ảo duy nhất, được đặc trưng bởi trạng thái hình học gồm vị trí và tư thế, ký hiệu lần lượt là $p^*(t)$ và $R^*(t)$. Mỗi UAV được gán một điểm cố định trong cấu trúc ảo này thông qua các véc tơ đội hình δ_i và có nhiệm vụ bám theo điểm tương ứng của mình. Nhờ đó, bài toán điều khiển đội hình được quy về bài toán điều khiển quỹ đạo của một vật thể cứng trong không gian, trong khi hình dạng đội hình được bảo toàn một cách tự nhiên [3, 2].



Hình 1. Ví dụ về véc tơ đội hình δ_i trong cấu trúc đội hình tam giác

Trong bài toán này, việc sử dụng phương pháp cấu trúc ảo nhằm xác định điểm tham chiếu vị trí cho từng UAV. Các đại lượng động học cao hơn của cấu trúc ảo như vận tốc tuyến tính và vận tốc góc không được xét, do không ảnh hưởng trực tiếp tới việc xác định hình dạng đội hình ở mức hình học.

Trong thực tế, ma trận tư thế của cấu trúc ảo $R^*(t)$ thường được xác định theo tư thế của UAV leader hoặc được sinh từ một quỹ đạo tham chiếu ảo, tùy thuộc vào kiến trúc điều khiển được lựa chọn.

2.3.2. Trường thế nhân tạo

Phương pháp trường thế nhân tạo (Artificial Potential Field – APF) mô hình hóa không gian chuyển động của UAV như một trường năng lượng nhân tạo, trong đó mục tiêu (hoặc leader, điểm cấu trúc ảo) đóng vai trò là nguồn thế hút, còn chướng ngại vật và các UAV lân cận là nguồn thế đẩy. UAV chuyển động theo hướng giảm nhanh nhất của tổng thế, qua đó đạt được mục tiêu điều hướng và tránh va chạm một cách tự nhiên.

Trong điều khiển bầy đàn UAV, APF thường không được sử dụng độc lập mà đóng vai trò là cơ chế tạo ra hướng chuyển động hoặc lực điều khiển ảo, hỗ trợ các phương pháp khác như leader–follower, consensus hoặc cấu trúc ảo nhằm đảm bảo tránh va chạm và duy trì an toàn khi đội hình di chuyển.

Bảng 3 tổng hợp một số nghiên cứu đại diện về phương pháp cấu trúc ảo và trường thế nhân tạo trong điều khiển đội hình UAV, tập trung vào khả năng duy trì hình dạng đội hình và tránh chướng ngại vật.

Bảng 3. Một số công trình tiêu biểu về cấu trúc ảo và trường thế nhân tạo

Phương pháp điều khiển	Kết quả chính	Ưu điểm	Nhược điểm
APF cải tiến cho đội hình và lập kế hoạch [9]	Tránh chướng ngại vật hiệu quả	Thiết kế trực quan	Vẫn tồn tại nguy cơ cực tiểu cục bộ
APF kết hợp đồng thuận [10].	Tránh va chạm trong môi trường động	Phân tán; phù hợp môi trường động	Phân tích phức tạp khi số lượng UAV lớn

Trường thế nhân tạo cải tiến và điều khiển trượt cho bầy UAV [11].	Kết hợp IAAPF với điều khiển trượt	Giảm dao động; tăng độ bền vững	Cấu trúc điều khiển phức tạp; cần tinh chỉnh tham số
Cấu trúc ảo kết hợp APF [5].	Duy trì đội hình tốt	Bảo toàn hình dạng đội hình tốt	Khó mở rộng đội hình lớn

2.4. Phương pháp dựa trên hành vi

Phương pháp dựa trên hành vi xuất phát từ mô hình đàn chim kinh điển của Reynolds, trong đó hành vi toàn cục của bầy đàn được hình thành từ sự kết hợp các hành vi cục bộ đơn giản của từng cá thể. Mỗi UAV phản ứng dựa trên thông tin lân cận và môi trường xung quanh, nhờ đó đội hình có khả năng tự tổ chức, thích nghi tốt với môi trường phức tạp và thay đổi nhanh [10].

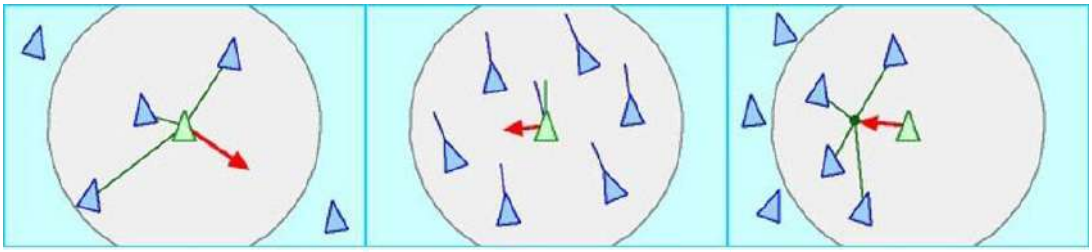
Trong điều khiển bầy đàn UAV, các hành vi thường được thiết kế sao cho tổng hợp lại tạo thành chuyển động mong muốn của toàn đội mà không cần thông tin toàn cục.

Dựa trên các quy tắc Reynolds, mỗi UAV tuân theo một số hành vi cục bộ tiêu biểu:

- Tránh va chạm (separation): tránh tiếp cận quá gần các UAV lân cận;

- Căn chỉnh (alignment): điều chỉnh vận tốc để tiệm cận vận tốc trung bình của UAV lân cận;

- Hội tụ (cohesion): di chuyển về phía tâm khối của các UAV lân cận.



Tránh va chạm Căn chỉnh Hội tụ

Hình 2. Mô tả các lực hành vi cơ bản

Bảng 4 tóm tắt một số công trình tiêu biểu về điều khiển dựa trên hành vi cho UAV bầy đàn, làm nổi bật đặc trưng phân tán, khả năng phản ứng nhanh cũng như các hạn chế trong việc duy trì đội hình hình học chính xác.

Bảng 4. Một số công trình tiêu biểu về điều khiển dựa trên hành vi cho đội hình UAV

Phương pháp điều khiển	Kết quả chính	Ưu điểm	Nhược điểm
Bầy đàn dựa trên quy tắc hành vi (Reynolds rules) [7].	Mô hình hóa bầy đàn dựa trên các hành vi cục bộ	Phân tán hoàn toàn; mở rộng tốt cho bầy lớn	Khó duy trì hình dạng đội hình
Đồng thuận dựa trên hành vi cho	Kết hợp hành vi cục bộ với cơ chế đồng	Cấu trúc đơn giản; dễ tích hợp với đồng thuận	Phụ thuộc giả thiết lý tưởng; hạn chế

đội hình UAV [8]	thuận		khi nhiễu mạnh
Hành vi tránh va chạm dựa trên trường véc tơ [12]	Tránh va chạm hiệu quả kết hợp duy trì đội hình	Phản ứng nhanh; phù hợp môi trường động	Điều chỉnh tham số phức tạp
Hành vi kết hợp với APF cải tiến [10]	Kết hợp hành vi, APF và đồng thuận	Tránh va chạm hiệu quả hơn APF truyền thống; duy trì đội hình tốt	Phụ thuộc thiết kế thế năng và truyền thông mạng.

2.5. Điều khiển tăng cường: phân tán, bền vững, thích nghi

Nhằm nâng cao khả năng hoạt động của đàn quadrotor trong các điều kiện không lý tưởng, bao gồm nhiễu không biết trước, tham số mô hình không chính xác, ràng buộc đầu vào và hạn chế thông tin cảm biến. Trọng tâm là phát triển các chiến lược điều khiển phân tán có tính bền vững và thích nghi, nhằm đảm bảo đàn UAV duy trì khả năng bám đội hình và bám quỹ đạo chung ngay cả khi mạng kết nối bị hạn chế và mỗi UAV chỉ sử dụng thông tin cục bộ [3, 13, 12, 14].

Bảng 5 tóm tắt các công trình tiêu biểu về điều khiển phân tán, bền vững và thích nghi cho quadrotor bầy đàn, nhấn mạnh khả năng xử lý nhiễu, bất định mô hình và hạn chế cảm biến trong các kịch bản thực tế.

Bảng 5. Một số công trình tiêu biểu về điều khiển phân tán, bền vững và thích nghi cho đội hình quadrotor

Phương pháp điều khiển	**Kết quả chính**	**Ưu điểm**	**Nhược điểm**
Điều khiển bám đội hình phân tán bền vững [3].	Điều khiển bám đội hình phân tán cho quadrotor, xử lý tham số không biết trước và nhiễu không xác định, hội tụ bằng Lyapunov.	Độ bền vững cao; không yêu cầu mô hình chính xác; phù hợp môi trường thực tế.	Cấu trúc điều khiển phức tạp; cần tinh chỉnh nhiều tham số thích nghi.
Điều khiển phân tán không dùng vận tốc tuyến tính [13].	Xây dựng bộ điều khiển đội hình trên SO(3), không cần đo vận tốc tuyến tính.	Không cần đo vận tốc.	Thiết kế và phân tích toán học phức tạp.
Điều khiển trượt, bền vững kết hợp tránh va chạm [12].	Kết hợp điều khiển trượt với trường véc tơ đẩy đảm bảo bền vững và tránh va chạm.	Chịu nhiễu tốt; hội tụ hữu hạn.	Dao động (chattering).
Điều khiển trượt, thích nghi, hợp tác [14].	Điều khiển trượt, thích nghi, hợp tác cho UAV bầy đàn, đảm bảo ổn định toàn cục.	Khả năng thích nghi tốt; phân tích ổn định chặt chẽ.	Cấu trúc điều khiển phức tạp; chi phí tính toán cao.
Điều khiển phân	Khai thác biểu diễn hình học để	Tránh kỳ dị; phù	Khó triển khai

tán trên SO(3) [13]	tránh kỳ dị và nâng cao độ chính xác điều khiển.	hợp chuyển động phức tạp.	nếu thiếu kiến thức hình học.

2.6. Các phương pháp tối ưu bầy đàn dựa trên PSO

Particle Swarm Optimization (PSO) là thuật toán tối ưu bầy đàn lấy cảm hứng từ hành vi tìm kiếm tập thể của đàn cá [15]. Mỗi hạt trong PSO không đại diện cho một UAV vật lý, mà biểu diễn một nghiệm tối ưu trong không gian tham số hoặc không gian quỹ đạo, chẳng hạn như tập tham số điều khiển, tham số quỹ đạo, hoặc cấu hình đội hình mong muốn.

Bảng 6. Một số công trình tiêu biểu về tối ưu bầy đàn dựa trên PSO

Phương pháp điều khiển	Kết quả chính	Ưu điểm	Nhược điểm
PSO – thuật toán gốc [15]	Đề xuất PSO cho bài toán tối ưu toàn cục dựa trên hành vi bầy đàn.	Dễ triển khai; không yêu cầu gradient.	Không đảm bảo hội tụ chặt chẽ.
PSO cải tiến (inertia weight) [16]	Cải thiện khả năng cân bằng khám phá–khai thác của PSO.	Hội tụ tốt hơn PSO gốc.	Phụ thuộc lựa chọn tham số.

Các cá nhân cập nhật trạng thái của mình dựa trên kinh nghiệm tốt nhất của cá nhân và kinh nghiệm tốt nhất của toàn bầy, qua đó từng bước hội tụ về nghiệm tối ưu toàn cục [16].

Bảng 6 trình bày các công trình đại diện liên quan đến phương pháp tối ưu bầy đàn dựa trên PSO, tập trung vào vai trò của PSO trong tối ưu tham số, quỹ đạo và cấu hình đội hình cho UAV bầy đàn.

2.7. AI / học tăng cường sâu và phương pháp lai

Trong phạm vi điều khiển UAV bầy đàn, trí tuệ nhân tạo (AI) bao gồm nhiều hướng tiếp cận khác nhau như học máy cổ điển, logic mờ và học tăng cường. Trong số đó, học tăng cường sâu **(**Deep Reinforcement Learning – DRL) là hướng được quan tâm nhiều nhất trong những năm gần đây do khả năng xử lý các bài toán ra quyết định phức tạp và môi trường khó mô hình hóa. Do đó, mục này tập trung chủ yếu vào DRL và các kiến trúc điều khiển lai dựa trên DRL.

* Học tăng cường sâu

Các nghiên cứu gần đây đã ứng dụng học tăng cường sâu (DRL), chẳng hạn như DQN, DDQN và PPO, để học chính sách ra quyết định cho UAV bầy đàn trong các môi trường phức tạp, khó mô hình hoá chính xác, bao gồm môi trường có chướng ngại động hoặc đối thủ [4].

Khác với các phương pháp điều khiển dựa trên mô hình, DRL cho phép UAV học chiến lược điều khiển thông qua tương tác với môi trường và tối ưu trực tiếp hàm thưởng dài hạn.

* Các phương pháp lai (Hybrid approaches)

Xu hướng chủ đạo hiện nay không nhằm thay thế hoàn toàn điều khiển cổ điển bằng AI, mà hướng tới các phương pháp lai, kết hợp ưu điểm của cả hai. Trong các kiến trúc này:

- Tầng dưới sử dụng các phương pháp điều khiển cổ điển có chứng minh ổn định và bền vững, như backstepping, điều khiển trượt, ADRC, nhằm đảm bảo an toàn và khả năng triển khai thực tế;

- Tầng trên sử dụng DRL hoặc các phương pháp học sâu để ra quyết định chiến lược, chẳng hạn như lựa chọn chế độ đội hình, quỹ đạo tham chiếu, hoặc phân công nhiệm vụ.

Cách tiếp cận lai này cho phép tận dụng tính an toàn và khả năng phân tích của điều khiển cổ điển, đồng thời khai thác khả năng thích ứng và học hỏi của AI trong các môi trường phức tạp và thay đổi theo thời gian.

Bảng 7 tóm tắt các nghiên cứu tiêu biểu về AI và học tăng cường sâu (DRL) cho UAV bầy đàn, làm nổi bật vai trò của các phương pháp học trong ra quyết định chiến lược và lập kế hoạch ở tầng cao.

Bảng 7. Một số công trình tiêu biểu về AI / DRL và phương pháp lai cho UAV bầy đàn

Phương pháp điều khiển	Kết quả chính	Ưu điểm	Nhược điểm
Đa tác tử DRL cho nhóm UAV [6]	Áp dụng MARL cho điều khiển và phối hợp UAV.	Khả năng thích ứng cao.	Huấn luyện phức tạp; khó triển khai thực tế.
Leader–follower + DQN [4]	Kết hợp DQN với leader–follower để giữ đội hình và tránh va chạm.	Ra quyết định chiến lược hiệu quả.	Phụ thuộc dữ liệu huấn luyện; thiếu đảm bảo an toàn chặt chẽ.
DRL cho lập kế hoạch đường bay [17]	Sử dụng DRL cho lập kế hoạch trong môi trường thảm hoạ.	Phù hợp môi trường không mô hình.	Chi phí tính toán cao.

3. Thảo luận: Lựa chọn phương pháp cho điều khiển đội hình quadrotor

3.1. So sánh các phương pháp điều khiển UAV bầy đàn

Từ các nhóm phương pháp đã trình bày, có thể thấy không tồn tại một chiến lược điều khiển duy nhất phù hợp cho mọi bài toán quadrotor bầy đàn. Việc lựa chọn phương pháp phụ thuộc vào mục tiêu nhiệm vụ, điều kiện môi trường, cũng như các ràng buộc về truyền thông và cảm biến. Để cung cấp cái nhìn định lượng, phần sau tổng hợp và so sánh các phương pháp theo các tiêu chí về độ phức tạp tính toán, yêu cầu băng thông truyền thông và khả năng chống chịu lỗi nhằm cung cấp xu hướng về độ phức tạp và đặc tính của các phương pháp.

Bảng 8. So sánh định lượng các phương pháp điều khiển UAV bầy đàn

Phương pháp	Độ phức tạp tính toán	Băng thông truyền thông	Khả năng chịu lỗi
Leader–Follower	O(N) (tuyến tính)	Thấp–Trung bình	Thấp
Đồng thuận	O(E) (theo số liên kết)	Trung bình–Cao	Trung bình–Cao
Cấu trúc ảo và trường	$O(N^2)$ (tương tác cặp)	Thấp–Trung	Trung bình

thế nhân tạo		bình	
Dựa trên hành vi	$O(N^2)$ (tương tác cặp)	Thấp	Cao
Điều khiển phân tán, bền vững, thích nghi	$O(N \cdot k)$ (tham số thích nghi)	Thấp–Trung bình	Rất cao
Tối ưu bầy đàn (PSO)	$O(P \cdot iter)$ (lặp tối ưu)	Không đáng kể	Phụ thuộc
AI / DRL và phương pháp lai	Cao (phụ thuộc mô hình)	Cao	Chưa đảm bảo

Có thể thấy rằng các phương pháp phân tán như đồng thuận và dựa trên hành vi có khả năng chịu lỗi cao hơn nhưng thường yêu cầu băng thông truyền thông lớn hơn. Ngược lại, phương pháp leader–follower có cấu trúc đơn giản và chi phí tính toán thấp, nhưng độ tin cậy bị hạn chế do phụ thuộc vào leader. Các phương pháp bền vững và thích nghi cho thấy khả năng chống chịu nhiễu và bất định cao, đổi lại chi phí tính toán tăng lên. Trong khi đó, các phương pháp dựa trên tối ưu bầy đàn và học máy mang lại tính linh hoạt cao nhưng vẫn còn thách thức về độ tin cậy khi triển khai thực tế.

3.2. Thách thức thực nghiệm: khoảng cách giữa mô phỏng và triển khai thực tế

Mặc dù nhiều phương pháp điều khiển bầy đàn UAV đã được kiểm chứng hiệu quả trong môi trường mô phỏng, việc triển khai thực tế vẫn đối mặt với một khoảng cách đáng kể do các giả thiết lý tưởng không còn được đảm bảo.

Thứ nhất, độ trễ truyền thông giữa các UAV là yếu tố quan trọng ảnh hưởng đến hiệu suất và ổn định của hệ. Trong thực tế, dữ liệu trao đổi qua các mạng không dây (WiFi, mesh) luôn tồn tại độ trễ và có thể biến thiên theo thời gian. Điều này đặc biệt ảnh hưởng đến các phương pháp dựa trên đồng thuận và leader–follower, vốn thường giả thiết truyền thông tức thời hoặc trễ nhỏ. Độ trễ lớn hoặc không đồng đều có thể làm suy giảm tốc độ hội tụ, gây dao động, thậm chí dẫn đến mất ổn định đội hình [8], [13].

Thứ hai, sự bất định của môi trường thực làm sai lệch đáng kể so với mô hình lý tưởng trong mô phỏng. Các yếu tố như nhiễu gió, sai số cảm biến, sai lệch mô hình động lực học và nhiễu đo lường khiến các giả thiết về hệ thống chính xác không còn đúng. Những bất định này có thể tích lũy theo thời gian, làm giảm độ chính xác bám đội hình và ảnh hưởng đến an toàn khi vận hành [3], [13].

Do đó, để thu hẹp khoảng cách giữa mô phỏng và thực nghiệm, các nghiên cứu gần đây có xu hướng phát triển các luật điều khiển có khả năng chịu trễ truyền thông và bền vững trước bất định môi trường, đồng thời tích hợp kiểm chứng trên phần cứng hoặc thử nghiệm thực tế nhằm đảm bảo tính khả thi khi triển khai.

4. Kết luận và hướng nghiên cứu tiếp theo

Bài báo đã tổng quan các phương pháp điều khiển UAV bầy đàn dạng quadrotor từ góc nhìn cấu trúc điều khiển (leader–follower, đồng thuận, cấu trúc ảo và trường thế nhân tạo, dựa trên hành vi) và kỹ thuật điều khiển (phân tán, bền vững, tối ưu bầy đàn, học máy). Các công trình gần đây cho thấy xu hướng:

- Ưu tiên các luật điều khiển phân tán, không cần đo vận tốc, chịu được nhiễu và trễ;

- Tích hợp mạnh mẽ giữa cấu trúc đội hình truyền thống (leader–follower/đồng thuận) với trường thế nhân tạo cải tiến và điều khiển trượt/backstepping;

- Khai thác AI/DRL ở tầng quyết định chiến lược, trong khi vẫn duy trì tầng điều khiển thấp cấp được chứng minh ổn định.

Một số hướng nghiên cứu mở:

- Mở rộng quy mô lên đến hàng trăm UAV với cấu trúc mạng động, có khả năng chịu đựng việc mất nút hoặc mất liên kết mà vẫn đảm bảo duy trì nhiệm vụ.

- Đảm bảo tính an toàn và ổn định cho điều khiển bầy đàn dựa trên trí tuệ nhân tạo bằng cách kết hợp học tăng cường với các ràng buộc Lyapunov, hàm rào cản điều khiển hoặc cơ chế che chắn an toàn.

- Đề xuất tích hợp cảm biến chi phí thấp với kỹ thuật định vị tương đối (UWB, VIO) nhằm triển khai điều khiển đội hình trong môi trường không có tín hiệu vệ tinh. [16]

- Tối ưu đa mục tiêu (năng lượng, thời gian, độ phủ, độ tin cậy), kết hợp tối ưu bầy đàn và đồng thuận/leader–follower.

Tài liệu tham khảo

1. Khalid, A., Mushtaq, Z., Arif, S., Zeb, K., Khan, M. A., & Bakshi, S. (2023). Control schemes for quadrotor UAV: Taxonomy and survey. *ACM Computing Surveys, 56*(5).
2. Do, H. T., Hua, H. T., Nguyen, M. T., Nguyen, C. V., Nguyen, H. T. T., & Hoa, T. (2021). Formation control algorithms for multiple-UAVs: A comprehensive survey. *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, 1–13.
3. Xu, L., & Li, Y. (2023). Distributed robust formation tracking control for quadrotor UAVs with unknown parameters and uncertain disturbances. *Aerospace, 10*(10), 845.
4. Liu, Z., Li, J., Shen, J., Wang, X., & Chen, P. (2024). Leader–follower UAVs formation control based on a deep Q-network collaborative framework. *Scientific Reports, 14*(1), 4674.
5. Wu, E., Sun, Y., Huang, J., Zhang, C., & Li, Z. (2020). Multi UAV cluster control method based on virtual core in improved artificial potential field. *IEEE Access, 8*, 131647–131661.
6. Xu, D., & Chen, G. (2022). Autonomous and cooperative control of UAV cluster with multi-agent reinforcement learning. *Aeronautical Journal, 126*(1300), 932–951.
7. Olfati-Saber, R. (2006). Flocking for multi-agent dynamic systems: Algorithms and theory. *IEEE Transactions on Automatic Control, 51*(3), 401–420.
8. Ren, W. (2007). Consensus strategies for cooperative control of vehicle formations. *IET Control Theory & Applications, 1*(2), 505–512.
9. Pan, Z., Zhang, C., Xia, Y., Xiong, H., & Shao, X. (2022). An improved artificial potential field method for path planning and formation control of the multi-UAV systems. *IEEE Transactions on Circuits and Systems II: Express Briefs, 69*(3), 1129–1133.
10. Wang, N., Dai, J., & Ying, J. (2021). UAV formation obstacle avoidance control algorithm based on improved artificial potential field and consensus. *International Journal of Aeronautical and Space Sciences, 22*(6), 1413–1427.
11. Zhang, P., Wang, Z., Zhu, Z., Liang, Q., & Luo, J. (2024). Enhanced multi-UAV formation control and obstacle avoidance using IAAPF-SMC. *Drones, 8*(9), 514.
12. González-Sierra, J., Dzul, A., & Ríos, H. (2019). Robust sliding-mode formation control and collision avoidance via repulsive vector fields for a group of quad-rotors.

International Journal of Systems Science, 50(7), 1483–1500.

13. Bu, Y., Yan, Y., & Yang, Y. (2024). Advancement challenges in UAV swarm formation control: A comprehensive review. *Drones, 8*(7), 320.
14. Chen, Y., Liang, J., Miao, Z., & Wang, Y. (2020). Formation control of unmanned aerial vehicles based on cooperative sliding mode adaptive control. *ISA Transactions, 102*, 1–12.
15. Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of ICNN'95 - International Conference on Neural Networks* (pp. 1942–1948).
16. Shi, Y., & Eberhart, R. (1998). A modified particle swarm optimizer. In *1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence* (pp. 69–73).
17. Zhao, J., et al. (2022). Path planning research of a UAV base station searching for disaster victims' location information based on deep reinforcement learning. *Entropy, 24*(12), 1767.

Tổng hợp dẫn và điều khiển thời gian cố định dựa trên bộ quan sát nhiễu cho tên lửa đánh chặn mục tiêu cơ động

Phạm Văn Nguyên[*,1], Đinh Hồng Toàn[1], Nguyễn Văn Xuân[1], Trần Công Phan[1], Nguyễn Hữu Hải[2], Nguyễn Hữu Nam[1]

[1]*Viện Tên lửa và kỹ thuật điều khiển, Học viện Kỹ thuật quân sự*

[2]*Trung tâm Nghiên cứu bảo đảm kỹ thuật và CGCN, Học viện Kỹ thuật quân sự*

[*] Email: toandh@lqdtu.edu.vn, Contact number: 0988681582

Tóm tắt

Bài báo trình bày một phương pháp dẫn và điều khiển tích hợp thời gian cố định cho tên lửa đánh chặn mục tiêu cơ động trong điều kiện tồn tại bất định mô hình và nhiễu. Thời gian điều khiển được xác định cụ thể như một tham số thiết kế, độc lập với điều kiện ban đầu. Các bộ quan sát nhiễu thời gian cố định được xây dựng nhằm ước lượng và bù các nhiễu tổng gây ra bởi sự cơ động mục tiêu, bất định mô hình và biến thiên khí động học. Dựa trên mô hình IGC có dạng phản hồi chặt, bộ điều khiển backstepping được thiết kế, bảo đảm sai số bám sát hội tụ trong thời gian cố định. Phân tích Lyapunov được sử dụng để chứng minh tính ổn định của hệ kín. Các kết quả mô phỏng cho thấy phương pháp tổng hợp điều khiển đề xuất giúp tăng cường tính bền vững và cải thiện khả năng bám sát, nâng cao độ chính xác đánh chặn mục tiêu so với các phương pháp backstepping truyền thống.

Từ khoá: Dẫn và điều khiển tích hợp (IGC); Điều khiển backstepping thời gian cố định (FxTB); Bộ quan sát nhiễu thời gian cố định (FxTDO); Mục tiêu cơ động.

1. Mở đầu

Bài toán đánh chặn chính xác mục tiêu cơ động là yêu cầu quan trọng trong các nhiệm vụ tác chiến hiện đại của tên lửa, đặc biệt khi hệ thống chịu tác động của động lực học phi tuyến, bất định tham số và nhiễu bên ngoài. Các phương pháp thiết kế truyền thống thường phân tách vòng dẫn và vòng điều khiển-tự lái theo cấu trúc tách biệt. Mặc dù cách tiếp cận này làm đơn giản hóa quá trình tổng hợp điều khiển, nhưng thiết kế tách biệt đó có thể làm suy giảm hiệu năng do ảnh hưởng của liên kết động lực học, sai lệch mô hình và độ trễ đáp ứng giữa hai phân hệ [1], đặc biệt khi mục tiêu cơ động mạnh với tốc độ và gia tốc biến thiên.

Nhằm nâng cao mức độ phối hợp và hiệu quả đánh chặn tổng thể, các thiết kế dẫn và điều khiển tích hợp (Integrated Guidance and Control) đã được phát triển, trong đó toàn bộ hệ thống được thiết kế trong một khung động lực học thống nhất [1]. Việc loại bỏ ranh giới giữa hai vòng điều khiển cho phép xử lý trực tiếp sự liên kết của động lực học tên lửa, qua đó cải thiện đáng kể đặc tính quá độ và độ chính xác xác lập.

Do mô hình IGC thường được biểu diễn dưới dạng phản hồi chặt (strict-feedback), kỹ thuật backstepping trở thành công cụ thiết kế phổ biến. Để tăng cường khả năng hoạt động ổn định và chính xác trước các bất định tham số và nhiễu, nhiều biến thể kỹ thuật điều khiển backstepping kết hợp với bộ quan sát đã được đề xuất [2]–[5]. Tuy nhiên, phần lớn các phương pháp này chỉ bảo đảm hệ thống ổn định tiệm cận. Trong các tình huống tác chiến có tính biến động cao với mục tiêu cơ động mạnh, đặc tính hội tụ tiệm cận có thể không đáp ứng yêu cầu về tốc độ đáp ứng, dẫn đến giảm độ chính xác đánh chặn mục tiêu. Nhằm rút ngắn thời gian đáp ứng, điều khiển thời gian hữu hạn đã được áp dụng cho các hệ IGC. Các thiết kế backstepping thời gian hữu hạn cho thấy khả năng cải thiện độ bền vững và giảm thời gian hội tụ trong bài toán đánh chặn [6]–[8]. Tuy nhiên, thời gian hội tụ của các phương pháp này thường phụ thuộc vào điều kiện ban đầu, dẫn đến các hạn chế khi mục tiêu cơ động đột ngột. Lý thuyết điều khiển thời gian cố định khắc phục hạn chế trên bằng cách bảo đảm thời gian hội tụ bị chặn trên và độc lập với điều kiện ban đầu [8], [9]. Các nghiên cứu gần đây đã phát triển phương pháp điều khiển thời gian cố định cho các hệ thống IGC với việc thiết kế các bộ quan sát nhiễu, góp phần nâng cao chất lượng ước lượng và khả năng chống nhiễu trong điều

kiện bất định [10]–[12]. Phương pháp điều khiển thời gian cố định nổi lên như một hướng tiếp cận tiềm năng, cho phép xác định thời gian xác lập như một tham số thiết kế, đồng thời vẫn bảo đảm tính độc lập với điều kiện ban đầu. Các phương pháp backstepping thời gian cố định đã được phát triển cho nhiều lớp hệ phi tuyến, chứng minh khả năng bảo đảm ổn định trong thời gian cố định [13], [14]. Việc tích hợp phương pháp điều khiển với bộ quan sát nhiễu cũng cho thấy hiệu quả cao trong quá trình chống nhiễu và tăng cường độ bền vững đối với các hệ thống chịu tác động của bất định tổng hợp [14].

Mặc dù đã đạt được những tiến bộ đáng kể, song việc áp dụng phương pháp điều khiển thời gian cố định vào bài toán dẫn và điều khiển tích hợp cho đánh chặn mục tiêu cơ động vẫn còn hạn chế. Hiện nay, bài toán thiết kế điều khiển trên cơ sở bộ quan sát nhiễu thời gian cố định cho hệ thống IGC vẫn cần được tiếp tục nghiên cứu, phát triển một cách toàn diện. Xuất phát từ yêu cầu đó, bài báo này đề xuất một thiết kế dựa trên phương pháp điều khiển backstepping kết hợp với bộ quan sát nhiễu thời gian cố định. Các đóng góp chính của bài báo được tóm tắt như sau:

1, Thiết kế bộ quan sát nhiễu thời gian cố định (FxTDO) nhằm ước lượng chính xác các nhiễu tổng trong hệ thống IGC.

2, Đề xuất một sơ đồ điều khiển Backstepping (FxTB) dựa trên bộ quan sát nhằm cải thiện hiệu năng đánh chặn của tên lửa với tốc độ hội tụ nhanh hơn, độ bền vững cao hơn, giảm sai lệch pha cuối.

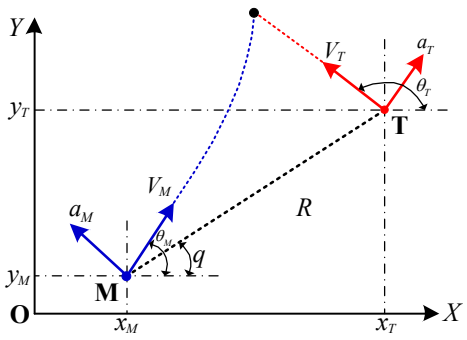
3, Tính ổn định của phương pháp đề xuất được chứng minh về mặt lý thuyết thông qua định lý Lyapunov. Ngoài ra, các mô phỏng số so sánh cũng được thực hiện nhằm chứng minh tính ưu việt của phương pháp IGC đề xuất trong các kịch bản thực tế khác nhau.

Các phần còn lại của bài báo được trình bày như sau. Mục 2 trình bày mô hình IGC của tên lửa. Mục 3 trình bày quy trình thiết kế bộ quan sát và bộ điều khiển. Mục 4 cung cấp kết quả mô phỏng và phân tích so sánh. Mục 5 kết luận bài báo.

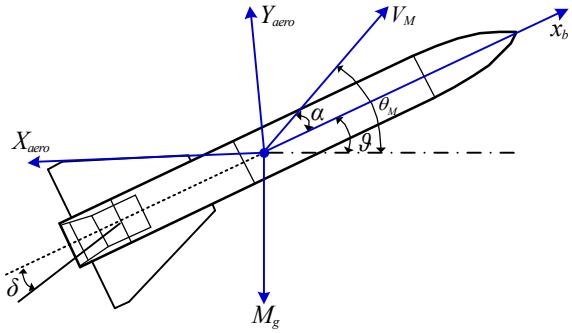
2. Mô hình hệ thống

2.1. Động học tên lửa-mục tiêu

Mô hình toán học mô tả chuyển động của tên lửa và mục tiêu trong mặt phẳng đứng được thể hiện trên Hình 1, trong đó OXY là hệ tọa độ quy chiếu quán tính Đề các, còn M và T lần lượt biểu thị các đại lượng của tên lửa và mục tiêu. Trong đó, x_m, y_m và x_T, y_T – tương ứng là vị trí hiện tại của tên lửa và mục tiêu; V_M, θ_M, a_M - tương ứng là tốc độ, góc nghiêng quỹ đạo và gia tốc pháp tuyến của tên lửa; V_T, θ_T, a_T tương ứng là tốc độ, góc nghiêng quỹ đạo và gia tốc pháp tuyến của mục tiêu; R- khoảng cách tương đối tên lửa-mục tiêu; q-góc đường ngắm mục tiêu (LOS); $O_m x_b$ - trục dọc tên lửa; ϑ- góc gật của tên lửa.



Hình 1. *Chuyển động tương đối của tên lửa và mục tiêu trong mặt phẳng đứng*



Hình 2. *Động lực của tên lửa trong mặt phẳng đứng*

Động lực học của tên lửa trong mặt phẳng đứng được thể hiện trên hình Hình 2 [4], trong đó: α- góc tấn công; M_g, X_{aero}, Y_{aero}- trọng lực, lực cản khí động và lực nâng của tên lửa.

Theo hình 1 và hình 2, các phương trình liên kết động học tên lửa, mục tiêu được cho như sau [4]:

$$\begin{cases} \dot{R} = V_T \cos(\theta_T - q) - V_m \cos(\theta_M - q) \\ R\dot{q} = V_T \sin(\theta_T - q) - V_m \sin(\theta_M - q) \\ \dot{\theta}_T = \dfrac{a_T}{V_T} \ , \ \dot{\theta}_M = \dfrac{a_M}{V_M} \end{cases} \quad (1)$$

Biến đổi các phương trình trong hệ (1) nhận được:

$$\ddot{q} = -\frac{2\dot{R}}{R}\dot{q} + \frac{a_T}{R}\cos(q-\theta_T) - \frac{a_M}{R}\cos(q-\theta_M) - \frac{\dot{V}_T}{R}\sin(q-\theta_T) + \frac{\dot{V}_M}{R}\sin(q-\theta_M) \quad (2)$$

Giả thiết tốc độ tên lửa không đổi ($\dot{V}_M = 0$), mục tiêu cơ động mạnh giai đoạn cuối và có tốc độ biến thiên $\dot{V}_T \neq 0$, phương trình (2) được rút gọn như sau:

$$\ddot{q} = \frac{2\dot{R}}{R}\dot{q} + \frac{a_T \cos(q-\theta_T)}{R} - \frac{a_M \cos(q-\theta_M)}{R} - \frac{\dot{V}_T}{R}\sin(q-\theta_T) \quad (3)$$

2.2. Động lực học điều khiển tên lửa

Trong bài báo này, xét động lực học tên lửa trên mặt phẳng gật có dạng như sau [4]:

$$\begin{aligned} &\dot{\alpha} = \frac{1}{mV_M}\left(-L + mg\cos\theta_M\right) + \omega_z \\ &J_z\dot{\omega}_z = M_0 + M_z^{\delta_z}\delta_z \\ &\dot{\vartheta} = \omega_z \ ; \ \alpha = \vartheta - \theta_M \end{aligned} \quad (4)$$

trong đó, m là khối lượng của tên lửa, ω_z là vận tốc góc gật, δ_z là góc lệch cánh lái theo trục z, và J_z là mô men quán tính của thiết bị đối với trục z. $M_z^{\delta_z}$ là thành phần mô men điều khiển tạo ra gia tốc góc, M_0 biểu diễn ảnh hưởng của các yếu tố khác lên gia tốc góc, và L là lực nâng, thường được xấp xỉ dưới giả thiết góc nhỏ đối với α, như sau:

$$\begin{cases} M_0 = M_\alpha\alpha + M_{\omega_z}\omega_z = 57.3QSlm_z^\alpha\alpha + \dfrac{QSl^2m_z^{\omega_z}}{V_M}\omega_z \\ L = 57.3QS(c_y^\alpha\alpha + c_y^{\delta_z}\delta_z) \ ; \ a_T = -g\cos\theta_T + \dfrac{57.3QSc_y^\alpha\alpha}{m} \end{cases}$$

Trong đó, Q là áp suất động, S là diện tích tham chiếu của tên lửa (diện tích khí động học), l là chiều dài tham chiếu; c_y^α, $c_y^{\delta_z}$ lần lượt là các đạo hàm lực nâng theo α và δ_z. Các đại lượng $m_z^\alpha, m_z^{\omega_z}, m_z^{\delta_z}$ là các đạo hàm mô men kênh gật tương ứng theo α, ω_z và δ_z.

Định nghĩa các biến trạng thái lần lượt là $x_1 = \dot{q}$, $x_2 = \alpha$, $x_3 = \omega_z$, và đầu vào điều khiển là $u=\delta_z$. Bằng cách gộp các bất định mô hình và nhiễu bên ngoài vào các hạng nhiễu d_i, hệ IGC được biểu diễn dưới dạng strict-feedback như sau.

$$\begin{cases} \dot{x}_1 = f_1(x_1) + g_1(x_1)x_2 + d_1 \\ \dot{x}_2 = f_2(x_2) + g_2(x_2)x_3 + d_2 \\ \dot{x}_3 = f_3(x_3) + g_3(x_3)u + d_3 \end{cases} \quad (5)$$

trong đó,

$$\begin{cases} f_1(x_1) = -\dfrac{2\dot{R}}{R}\dot{q} + \dfrac{g\cos(q-\theta_M)}{R}\cos\theta_M \\ f_2(x_2) = -\dfrac{57.3QSc_y^{\alpha}}{mV_M}\alpha + \dfrac{g\cos\theta_M}{V_M} \\ f_3(x_3) = \dfrac{57.3QSlm_z^{\alpha}}{J_z}\alpha + \dfrac{QSl^2m_z^{\omega}}{J_zV_M}\omega_z \end{cases}, \begin{cases} g_1(x_1) = -\dfrac{57.3QSc_y^{\alpha}}{mR}\cos(q-\theta_M) \\ g_2(x_2) = 1 \\ g_3(x_3) = \dfrac{57.3QSlm_z^{\delta_z}}{J_z} \end{cases} \quad (6)$$

$$\begin{cases} d_1 = \dfrac{a_T\cos(q-\theta_T)}{R} - \dfrac{\dot{V}_T}{R}\sin(q-\theta_T) \\ d_2 = d_2(c_y^{\alpha}, \Delta_2) \\ d_3 = d_3(m_z^{\alpha}, m_z^{\omega_z}, \Delta_3) \end{cases} \quad (7)$$

Các hàm trong (5-7) được xác định từ các quan hệ khí động học và động học, chúng là các đại lượng đo được hoặc bị chặn về mặt vật lý và biến thiên chậm trong quá trình hoạt động. Vì vậy, các hàm $f_i(x_i)$, $g_i(x_i)$ thu được là các hàm trơn, biến thiên theo thời gian và bị chặn, thỏa mãn điều kiện UUB (Uniformly Ultimately Bounded) cần thiết cho thiết kế backstepping thời gian cố định [6].

Trong phương trình (5), phương trình thứ nhất biểu diễn luật dẫn, trong đó nhiễu tổng d_1 chủ yếu phát sinh từ sự cơ động của mục tiêu, được đặc trưng bởi gia tốc pháp tuyến a_T, góc nghiêng quỹ đạo θ_T và sự biến thiên tốc độ $\dot{V}_T$ của mục tiêu. Phương trình thứ hai và thứ ba là các luật điều khiển: Phương trình thứ hai biểu diễn góc tấn công α, với nhiễu tổng d_2 chủ yếu do sự biến thiên của đạo hàm hệ số lực nâng khí động c_y^{α} và nhiễu ngoài Δ_2. Phương trình thứ ba biểu diễn tốc độ góc gật, với nhiễu tổng d_3 phụ thuộc vào các đạo hàm của hệ số mô men khí động m_z^{α}, $m_z^{\omega_z}$, $m_z^{\delta_z}$, sự dịch chuyển tâm khối lượng, dao động mật độ không khí và nhiễu ngoài Δ_3. Các đại lượng Δ_2 và Δ_3 là các nhiễu bên ngoài, bao gồm các hiệu ứng khí động chưa được mô hình hóa, sự biến thiên tham số và các bất định khác.

Giả thiết 1: Các nhiễu d_i và các đạo hàm bậc nhất của chúng bị chặn, tức là $|d_i| \le \bar{d}_i$, $|\dot{d}_i| \le \bar{\dot{d}}_i$, $(i = 1,2,3)$, trong đó $\bar{d}_i$ và $\bar{\dot{d}}_i$ là các hằng số dương chưa biết.

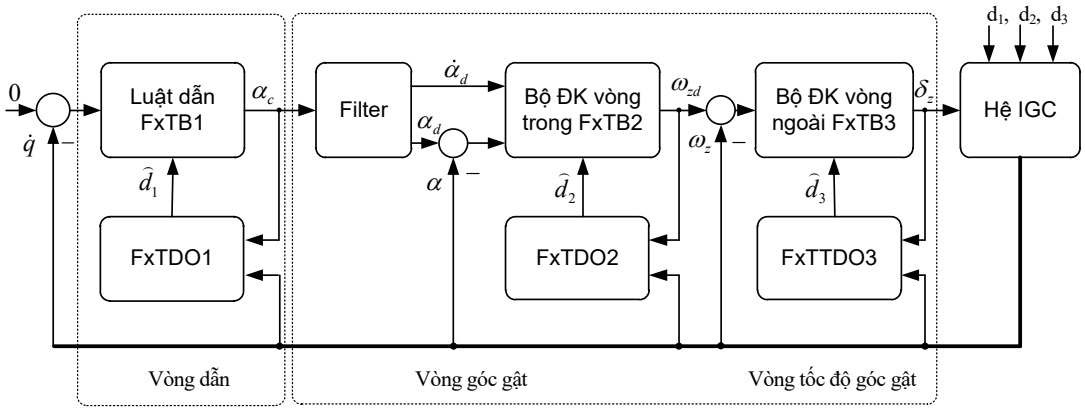
Giả thiết 2: Tín hiệu điều khiển kênh gật δ_z chịu ràng buộc bão hòa của cơ cấu chấp hành và do đó bị chặn. Trong quá trình hoạt động, khoảng cách tương đối tên lửa-mục tiêu thỏa mãn điều kiện $0<R(t)<R(0)$, và việc đánh chặn đạt được khi $R(t)\to 0$.

Giả thiết 3: Các hàm hệ số điều khiển $g_i(x_i)$ được giả thiết là bị chặn dưới, thỏa mãn $|g_i(x_i)| \ge g_{i,\min} > 0$, $(i = 1,2,3)$, bảo đảm điều khiển không bị triệt tiêu và duy trì cấu trúc strict-feedback cần thiết cho thiết kế điều khiển.

3. Thiết kế dẫn và điều khiển tích hợp

3.1. Cấu trúc dẫn-điều khiển tích hợp đề xuất

Hình 3 trình bày cấu trúc của hệ IGC với thiết kế Backstepping thời gian cố định được đề xuất. Cấu trúc điều khiển được tổ chức thành ba phân hệ mắc nối tiếp: Vòng dẫn, vòng điều khiển góc gật và và vòng điều khiển tốc độ góc gật.



Hình 3. *Cấu trúc của hệ IGC đề xuất*

Vòng dẫn tạo ra lệnh góc tấn mong muốn nhằm điều chỉnh tốc độ góc ngắm (LOS rate), trong khi các vòng trong bám theo các tín hiệu tham chiếu tương ứng và sinh ra tín hiệu điều khiển vật lý. Các bộ điều khiển FxTB được thiết kế cho từng phân hệ nhằm bảo đảm mọi sai số bám sát hội tụ trong một thời gian được xác định, không phụ thuộc vào điều kiện ban đầu. Để tăng cường tính bền vững trước các bất định mô hình và sự cơ động của mục tiêu, các bộ quan sát nhiễu FxTDO được tích hợp vào mỗi vòng điều khiển để ước lượng và bù các nhiễu tổng.

3.2. Thiết kế bộ quan sát nhiễu thời gian cố định

Trong bài toán đánh chặn mục tiêu đang xét, tồn tại các bất định mô hình hệ thống, các nhiễu ngoài tác động trong điều kiện mục tiêu cơ động và biến thiên theo thời gian. Để đạt được yêu cầu về thời gian đáp ứng xác định và hiệu quả ước lượng, các bộ quan sát nhiễu thời gian cố định (FxTDO) được phát triển nhằm ước lượng và bù các nhiễu tác động lên hệ IGC.

Để thiết kế bộ quan sát, xét hệ phi tuyến có chứa các thành phần nhiễu sau đây:

$$\dot{x}_i = f_i(x) + g_i(x)u_i + d_i,\ \ i = 1,2,3 \tag{8}$$

ở đây $x \in \mathbb{R}^n$ là trạng thái hệ đo được, u_i là tín hiệu điều khiển, và d_i là các nhiễu tổng.

Dựa trên lý thuyết về bộ quan sát thời gian cố định đồng nhất [15], một bộ quan sát nhiễu thời gian cố định (FxTDO) được thiết kế như sau

$$\begin{cases} \dot{\hat{x}}_i = f_i(x_i) + g_i(x_i)u_i + \hat{d}_i + L_{1i}(t)\Phi_1(e_i) \\ \dot{\hat{d}}_i = L_{2i}(t)\Phi_2(e_i) \end{cases} \tag{9}$$

với $e_i = x_i - \hat{x}_i$ và $\hat{d}_i$ là ước lượng của d_i. Các sai lệch ước lượng nhiễu là $\tilde{d}_i = d_i - \hat{d}_i$

Các hàm hiệu chỉnh phi tuyến $\Phi_1(e_i)$ và $\Phi_2(e_i)$ có cấu trúc tương tự như trong tài liệu [15], với các hệ số hiệu chỉnh c_{ij} và tham số điều khiển d_∞, được định nghĩa như sau:

$$\begin{cases} \Phi_1(e_i) = c_{11}\left[e_i\right]^{\frac{1}{2}} + c_{12}\left[e_i\right] + c_{13}\left[e_i\right]^{\frac{1}{1-d_\infty}} \\ \Phi_2(e_i) = c_{21}\left[e_i\right]^{0} + c_{22}\left[e_i\right] + c_{23}\left[e_i\right]^{\frac{1+d_\infty}{1-d_\infty}} \end{cases} \tag{10}$$

ở đây, $\left[e_i\right]^p = \left|e_i\right|^p \mathrm{sgn}(e_i)$, $d_\infty \in (0, 1)$ và $c_{ij} > 0$ với $i = 1,2; j = 1,2,3$. Lưu ý rằng $\Phi_1(e_i)$ và $\Phi_2(e_i)$ là các hàm phi tuyến đồng nhất.

Như vậy, đối với hệ phi tuyến có nhiễu (8), thành phần nhiễu được ước lượng chính xác trong thời gian cố định bằng cách áp dụng bộ quan sát nhiễu thời gian cố định (9) với các hàm hiệu chỉnh trong (10). Bộ quan sát nhiễu được sử dụng trong bài báo này được kế thừa từ

tài liệu [15], với tính ổn định đã được chứng minh một cách chặt chẽ. Do đó, phần phân tích ổn định sẽ được trình bày ngắn gọn.

Chọn hàm Lyapunov

$$V_{di} = \frac{1}{2} e_i^2 + \frac{1}{2L_{2i}} \tilde{d}_i^2 \tag{10}$$

Tính toán các biểu thức đạo hàm sai số:

$$\dot{e}_i = \tilde{d}_i - L_{1i}\Phi(e_i), \ \dot{\tilde{d}}_i = \dot{d}_i + L_{2i}\Phi(e_i)$$

Thay các biểu thức này vào đạo hàm của hàm V_{di}, ta thu được

$$\dot{V}_{di} \le \Delta_{di} |e_i| - \lambda_{1i} |e_i|^{p_i+1} - \lambda_{2i} |e_i|^{q_i+1} \tag{11}$$

ở đây $\lambda_{1i} = L_{1i} a_i > 0$ and $\lambda_{2i} = L_{1i} b_i > 0$

Do đó, tồn tại một tập compact $\Omega_i = \{e_i : |e_i| \le \varepsilon_i\}$, sao cho $\dot{V}_{di} < 0$ với mọi $e_{di} \notin \Omega_i$. Cụ thể, ε_i được xác định là nghiệm dương duy nhất của phương trình $\lambda_{1i}\varepsilon_i^{p_i} + \lambda_{2i}\varepsilon_i^{q_i} = \Delta_{di}$, nhằm đảm bảo rằng $\dot{V}_{di} \le 0$ bất cứ khi nào $|e_i| \ge \varepsilon_i$.

Nhận xét 1: Các hàm hiệu chỉnh phi tuyến $\Phi_1(e_i)$ và $\Phi_2(e_i)$ sử dụng cấu trúc đồng nhất kép (bi-homogeneous) nhằm đạt được tốc độ hội tụ nhanh trên toàn bộ không gian trạng thái. Cụ thể, thành phần số mũ hữu tỷ giúp tăng tốc độ hội tụ tại lân cận gốc tọa độ, trong khi thành phần phi tuyến bậc cao cung cấp khả năng hiệu chỉnh mạnh mẽ khi sai số ước lượng lớn. Nhờ đó, bộ quan sát FxTDO được đề xuất có khả năng tái cấu trúc chính xác các nhiễu tổng hợp, tạo điều kiện thuận lợi cho việc thiết kế bộ điều khiển backstepping thời gian cố định ở các bước tiếp theo.

Thời gian xác lập của sai số ước lượng nhiễu được chặn trên bởi biểu thức

$$T_{di} \le \frac{2}{\alpha_i(1-p_i)} + \frac{2}{\beta_i(q_i-1)} \tag{12}$$

với $\alpha_i = c_{1i} 2^{(p_i+1)/2}$, $\beta_i = c_{2i} 2^{(q_i+1)/2}$

Biên này không phụ thuộc vào điều kiện ban đầu và có thể được điều chỉnh cụ thể thông qua các tham số của bộ quan sát. Do đó, bộ quan sát nhiễu thời gian cố định được đề xuất bảo đảm hiệu quả ước lượng nhiễu có thể dự đoán trước và đáp ứng yêu cầu ràng buộc thời gian cho hệ IGC.

3.3. Thiết kế bộ điều khiển Backstepping thời gian cố định

Xét hệ dẫn–điều khiển tích hợp dạng strict-feedback (5), trong đó $\hat{d}_i$ là các giá trị ước lượng nhiễu do bộ quan sát nhiễu thời gian cố định được phát triển ở Mục 3.2 cung cấp. Dựa trên mô hình đã được bù bởi bộ quan sát, một phương pháp điều khiển backstepping thời gian cố định (FxTBC) được xây dựng theo cách đệ quy nhằm bảo đảm rằng mọi sai số bám hội tụ trong một khoảng thời gian do người thiết kế xác định, không phụ thuộc vào điều kiện ban đầu. Quá trình thiết kế điều khiển được thực hiện qua ba bước đệ quy tương ứng với vòng dẫn, vòng điều khiển góc gật và vòng điều khiển tốc độ góc gật.

Bước 1: Thiết kế điều khiển ảo cho vòng dẫn

Định nghĩa sai số bước thứ nhất

$$e_1 = x_1 - x_{1d} \ ; \ x_{1d} = 0 \tag{13}$$

Chọn hàm Lyapunov $V_1 = \frac{e_1^2}{2}$

Dựa trên lý thuyết backstepping thời gian cố định, tín hiệu điều khiển ảo thứ nhất được thiết kế như sau [14]

$$x_{2d} = \frac{1}{g_1(x_1)}\left(-f_1(x_1) - \hat{d}_1 - k_{11}|e_1|^{\mu_1}\,\text{sign}(e_1) - k_{12}|e_1|^{\nu_1}\,\text{sign}(e_1) - k_{13}e_1\right) \tag{14}$$

trong đó k_{11}, k_{12}, $k_{13} > 0$, $0 < \mu_1 < 1 < \nu_1$ là các tham số thiết kế.

Thay (14), thu được động học sai số

$$\dot{V}_1 = e_1\dot{e}_1 \le -c_{11}V_1^{\frac{\mu_1+1}{2}} - c_{12}V_1^{\frac{\nu_1+1}{2}} \tag{15}$$

Biểu thức (15) bảo đảm rằng e_1 hội tụ về một lân cận nhỏ của gốc tọa độ trong thời gian cố định, không phụ thuộc vào điều kiện ban đầu.

Bước 2: Thiết kế điều khiển ảo cho vòng điều khiển góc gật

Định nghĩa sai số bước thứ hai

$$e_2 = x_2 - x_{2d} \tag{16}$$

Chọn hàm Lyapunov $V_1 = \frac{e_2^2}{2}$

Sử dụng giá trị ước lượng nhiễu $\hat{d}_2$, tín hiệu điều khiển ảo thứ hai được chọn như sau [11]

$$x_{3d} = \frac{1}{g_2(x_2)}\left(-f_2(x_2) - \hat{d}_2 - \dot{x}_{2d} - k_{21}|e_2|^{\mu_2}\,\text{sign}(e_2) - k_{22}|e_2|^{\nu_2}\,\text{sign}(e_2) - k_{23}e_2\right) \tag{17}$$

trong đó k_{21}, k_{22}, $k_{23} > 0$ và $0 < \mu_2 < 1 < \nu_2$.

Khi đó, đạo hàm Lyapunov thỏa mãn

$$\dot{V}_2 \le -c_{21}V_2^{\frac{\mu_2+1}{2}} - c_{22}V_2^{\frac{\nu_2+1}{2}} \tag{18}$$

Biểu thức (15) bảo đảm rằng e_2 hội tụ về một lân cận nhỏ của gốc tọa độ trong thời gian cố định, không phụ thuộc vào điều kiện ban đầu.

Bước 3: Thiết kế luật điều khiển thực

Định nghĩa sai số bước thứ 3:

$$e_3 = x_3 - x_{3d} \tag{19}$$

Chọn hàm Lyapunov $V_1 = \frac{e_3^2}{2}$

Sử dụng ước lượng nhiễu $\hat{d}_3$, tín hiệu điều khiển thực được thiết kế như sau [11]

$$u = \frac{1}{g_3(x_3)}\left(-f_3(x_3) - \hat{d}_3 - \dot{x}_{3d} - k_{31}|e_3|^{\mu_3}\,\text{sign}(e_3) - k_{32}|e_3|^{\nu_3}\,\text{sign}(e_3) - k_{33}e_3\right) \tag{20}$$

trong đó k_{31}, k_{32}, $k_{33} > 0$ và $0 < \mu_3 < 1 < \nu_3$.

Theo đó, đạo hàm Lyapunov thỏa mãn

$$\dot{V}_3 \le -c_{31}V_3^{\frac{\mu_3+1}{2}} - c_{32}V_3^{\frac{\nu_3+1}{2}} \tag{21}$$

đảm bảo tính hội tụ trong thời gian cố định của sai số bám e_3.

Kết quả về tính ổn định: Từ các bất đẳng thức (15), (18) và (21), các sai số bám e_i, i=1, 2, 3 hội tụ về các lân cận nhỏ của gốc tọa độ trong thời gian bị chặn, độc lập với điều kiện ban đầu. Thời gian xác lập tổng thể của hệ thống có thể được xác định một cách cụ thể bởi các tham số thiết kế. Để giảm hiện tượng rung (chattering) trong mô phỏng số và triển khai thực tế, hàm dấu gián đoạn trong các biểu thức điều khiển có thể được thay thế bằng các xấp xỉ trơn như $\tanh(\sigma e_i)$, mà không làm ảnh hưởng đến tính chất ổn định thời gian cố định.

3. Mô phỏng và thảo luận

3.1. Số liệu mô phỏng

Để đánh giá hiệu quả của bộ điều khiển backstepping kết hợp với bộ quan sát nhiễu thời gian cố định thời gian cố định được đề xuất (FxTB–FxTDO), các mô phỏng số được thực hiện trên hệ dẫn và điều khiển tích hợp. Phương pháp đề xuất được so sánh với bộ điều khiển backstepping truyền thống kết hợp với bộ quan sát trạng thái mở rộng giảm bậc (Back–RESO) trong cùng điều kiện và tham số.

Để thực hiện mô phỏng, tên lửa được giả thiết có vận tốc không đổi, mục tiêu có vận tốc biến thiên theo thời gian, cơ động và thay đổi gia tốc pháp tuyến. Các điều kiện ban đầu được thiết lập như sau: khoảng cách tương đối ban đầu R_0= 14 km, góc đường ngắm ban đầu q_0= 30^o, vận tốc ban đầu của tên lửa V_M= 700 m/s, vận tốc ban đầu của mục tiêu V_T= 350m/s và biến thiên V_T= 350+80(t−12.5) từ thời điểm t=12.5s. Góc nghiêng quỹ đạo và góc tấn công ban đầu của tên lửa lần lượt là $\theta_M(0)$= 45^o và $\alpha(0)$= 0, trong khi nghiêng quỹ đạo ban đầu của mục tiêu là $\theta_T(0)$= 135^o. Các tham số khí động và quán tính của tên lửa được lựa chọn theo các giá trị tiêu chuẩn trong tài liệu [4]. Góc lệch điều khiển được giới hạn trong khoảng $|\delta_z| \leq 20^o$.

Đối với bộ điều khiển FxTB được đề xuất, các tham số thiết kế được lựa chọn nhằm thỏa mãn điều kiện thời gian cố định. Các tham số của bộ quan sát FxTDO được hiệu chỉnh để bảo đảm sai số ước lượng nhiễu hội tụ nhanh trong thời gian cố định. Tất cả các tham số của bộ điều khiển và bộ quan sát được lựa chọn một cách nhất quán nhằm bảo đảm việc đánh giá hiệu năng là khách quan.

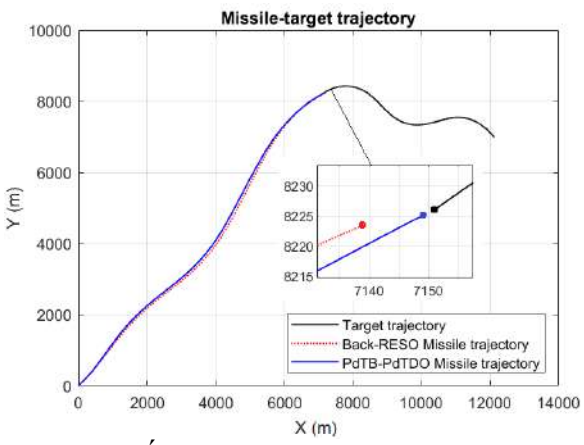
Xây dựng một kịch bản mô phỏng điển hình dưới điều kiện có bất định và mục tiêu cơ động. Các hệ số lực và mômen khí động giảm 25% so với giá trị danh định để đại diện cho các bất định mô hình. Mục tiêu cơ động với gia tốc pháp tuyến biến thiên theo thời gian: a_T= 110 sin(0.6t) m/s^2 và a_T= 110 m/s từ thời điểm t = 13s. Ngoài ra, các nhiễu bên ngoài tác động vào các kênh điều khiển: Δ_2= 0.5 sin(t), Δ_3= 0.35 sin(t).

3.2. Kết quả mô phỏng và bình luận

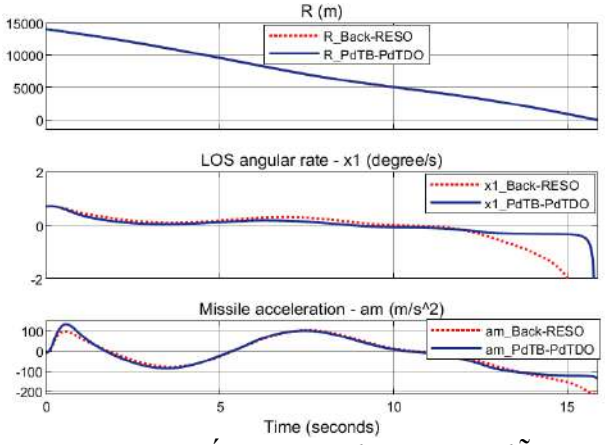
Các kết quả mô phỏng được trình bày trong các Hình 4–9.

Hình 4 trình bày các quỹ đạo đánh chặn của tên lửa–mục tiêu khi sử dụng phương pháp Back–RESO truyền thống và phương pháp FxTB–FxTDO được đề xuất. Cả hai bộ điều khiển đều tác động dẫn đến tên lửa đánh chặn mục tiêu thành công. Tuy nhiên, quỹ đạo của thiết kế FxTB–FxTDO bám sát đường bay mục tiêu hơn và có quá trình tiếp cận pha cuối nhanh hơn. Hình phóng to tại điểm gặp cho thấy sự cải thiện đạt được bởi phương pháp đề xuất. Về mặt định lượng, thời gian đánh chặn của hai phương pháp là tương đương, $T_{\text{Back–RESO}}$=15.58s, $T_{\text{FxTB–FxTDO}}$=15.56s. Tuy nhiên, sai lệch cuối tại điểm gặp giảm đáng kể từ $R_{\text{Back–RESO}}$=2.9m xuống $R_{\text{FxTB–FxTDO}}$=1.1m, qua đó chứng minh độ chính xác đánh chặn của phương pháp đề xuất được cải thiện đáng kể trong điều kiện mục tiêu cơ động.

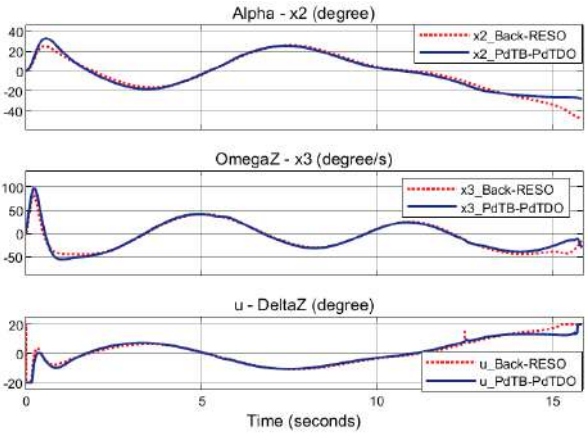
Hình 5 trình bày các đáp ứng trong vòng dẫn. Trong cả hai trường hợp, khoảng cách tương đối R giảm đơn điệu về 0 cho thấy việc đánh chặn thành công. So với Back–RESO, phương pháp FxTB–FxTDO có góc đường ngắm (LOS) hội tụ nhanh hơn và phẳng hơn với dao động được giảm rõ rệt trong pha cuối. Ngoài ra, đáp ứng gia tốc của tên lửa theo phương pháp đề xuất cũng tuyến tính hơn trong quá trình mục tiêu cơ động, cho thấy khả năng chống chịu tốt hơn trước các nhiễu biến thiên nhanh và giảm dao động quá độ.



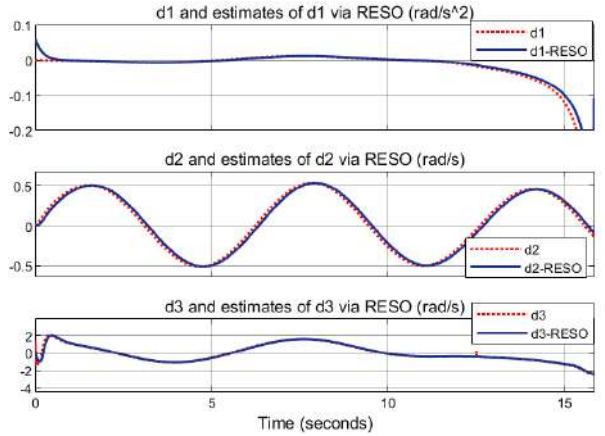
Hình 4. *Kết quả mô phỏng của quỹ đạo tên lửa-mục tiêu*



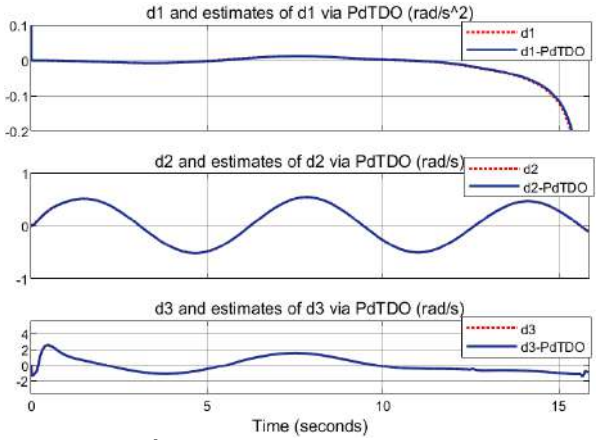
Hình 5. *Kết quả mô phỏng vòng dẫn*



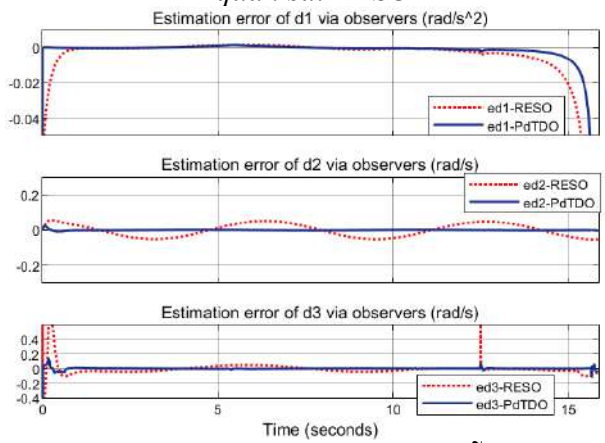
Hình 6. *Kết quả mô phỏng vòng điều khiển*



Hình 7. *Kết quả ước lượng nhiễu sử dụng bộ quan sát RESO*



Hình 8. *Kết quả ước lượng nhiễu sử dụng bộ quan sát FxTDO*



Hình 9. *Các sai lệch ước lượng nhiễu qua các bộ quan sát*

Đáp ứng trạng thái của vòng điều khiển được thể hiện trong Hình 6. Có thể quan sát thấy phương pháp FxTB–FxTDO có đáp ứng góc tấn công và tốc độ góc gật điều chỉnh ổn định hơn trong điều kiện có bất định và nhiễu ngoài, biên độ dao động trong pha cơ động cuối nhỏ hơn so với Back–RESO. Hơn nữa, tín hiệu điều khiển luôn bị chặn trong giới hạn cơ cấu chấp hành và phản ứng kịp thời để bù nhiễu, góp phần nâng cao độ chính xác pha cuối và làm giảm bão hòa điều khiển.

Hình 7 và 8 lần lượt trình bày kết quả ước lượng nhiễu qua bộ quan sát RESO và FxTDO. Cả hai bộ quan sát đều có khả năng ước lượng các nhiễu d_1, d_2, d_3 với sai số ước lượng bị chặn. Tuy nhiên, các ước lượng dựa trên bộ quan sát RESO có dao động lớn hơn và

xuất hiện độ trễ bám sát đáng kể khi có các biến thiên đột ngột trong các kênh nhiễu, đặc biệt khi mục tiêu cơ động mạnh ở giai đoạn cuối. Ngược lại, bộ quan sát FxTDO cho thấy quá trình hội tụ quá độ nhanh hơn và các tín hiệu ước lượng nhiễu trơn hơn, điều này minh chứng cho tính chất hội tụ trong thời gian cố định của bộ quan sát. Sai số ước lượng nhiễu của hai phương pháp quan sát được so sánh thêm trong Hình 9. Sai số tương ứng với FxTDO hội tụ nhanh hơn và duy trì trong các lân cận bị chặn nhỏ hơn so với RESO, đặc biệt khi tồn tại các nhiễu biến thiên theo thời gian và thay đổi nhanh. Độ chính xác ước lượng của FxTDO trực tiếp nâng cao tính bền vững và góp phần nâng cao độ chính xác dẫn và điều khiển, phù hợp với việc giảm sai lệch cuối và cải thiện hiệu năng bám quan sát trong Hình 4–6.

Nhìn chung, các kết quả mô phỏng xác nhận rằng, mặc dù cả hai phương pháp đều đạt được đánh chặn thành công với thời gian đánh chặn tương đương, song phương pháp FxTB–FxTDO cải thiện đáng kể khả năng bền vững trước các bất định mô hình, nhiễu cũng như sự cơ động mạnh của mục tiêu, dẫn đến cải thiện độ chính xác đánh chặn mục tiêu. Những ưu điểm này được thể hiện qua các đáp ứng dẫn và điều khiển, hiệu quả ước lượng nhiễu và sai lệch tại điểm gặp mục tiêu.

4. Kết luận

Bài báo đã đề xuất một thiết kế dẫn và điều khiển tích hợp theo thời gian cố định, kết hợp bộ điều khiển backstepping (FxTB) với bộ quan sát nhiễu (FxTDO), nhằm giải quyết bài toán đánh chặn mục tiêu cơ động trong điều kiện tồn tại bất định mô hình và nhiễu bên ngoài. Thông qua phân tích Lyapunov, phương pháp đề xuất bảo đảm các sai số bám hội tụ trong thời gian cố định, đồng thời đạt được khả năng ước lượng và bù nhiễu hiệu quả. Các kết quả mô phỏng trong điều kiện mục tiêu cơ động mạnh và bất định tham số cho thấy phương pháp FxTB–FxTDO nâng cao tính bền vững của hệ thống và cải thiện đáng kể độ chính xác đánh chặn pha cuối. Các kết quả thu được đã xác nhận tính hiệu quả và ưu thế của cấu trúc FxTB–FxTDO trong các tình huống tác chiến phức tạp. Trong các nghiên cứu tiếp theo, phương pháp đề xuất sẽ được mở rộng cho các trường hợp mô hình tên lửa đa chiều nhằm nâng cao tính thực tiễn và khả năng ứng dụng.

Tài liệu tham khảo

[1]. P. K. Menon and E. J. Ohlmeyer, "*Integrated design of agile missile guidance and autopilot systems*", Control Engineering Practice, Vol. 9, Iss. 10, pp. 1095-1106, 2001. DOI: 10.1016/S0967-0661(01)00082-X

[2]. H. Yan, X. H. Wang, B. F. Yu, and H. B. Ji, "*Adaptive integrated guidance and control based on backstepping and input-to-state stability*", Asian Journal of Control, Vol. 16, No. 2, pp. 602-608, 2014. DOI: 10.1002/asjc.682

[3]. Pei Pei, Y. Ji, S. He, J. Wang, and D. Lin, "*Integrated guidance and control using adaptive backstepping approach for maneuvering target interception*," IFAC-PapersOnLine, Vol. 53, Iss. 2, pp. 9458–9464, 2020, DOI: 10.1016/j.ifacol.2020.12.2418

[4]. S. Xingling and W. Honglun, "*Back-stepping active disturbance rejection control design for integrated missile guidance and control system via ESO*", ISA Transactions, Vol. 57, pp. 232–246, 2015. DOI: 10.1016/j.isatra.2015.02.013

[5]. L. Wang, W. Zhang, D. Wang, K. Peng, and H. Yang, "*Command filtered back-stepping missile integrated guidance and autopilot based on extended state observer*". Advances in Mechanical Engineering, vol. 9, no. 11, pp. 1–13, 2017. DOI: 10.1177/1687814017733251

[6]. S. P. Bhat and D. S. Bernstein, "*Finite-time stability of continuous autonomous systems*", SIAM Journal on Control and Optimization, Vol. 38, No. 3, pp. 751-766, 2000. DOI: 10.1137/S0363012997321358

[7]. Z. Chong, J. Guo, and X. Lu, "*Finite-time integrated guidance and control system for*

hypersonic vehicles", Transactions of the Institute of Measurement and Control, Vol. 43, Iss. 4, pp. 767–776, 2021. DOI: 10.1177/0142331220941934.

[8]. Polyakov, A. "*Nonlinear feedback design for fixed-time stabilization of linear control systems*," IEEE Transactions on Automatic Control, vol. 57, no. 8, pp. 2106–2110, 2012. DOI: 10.1109/TAC.2011.2179869.

[9]. Zuo, Z. "*Nonsingular fixed-time consensus tracking for second-order multi-agent systems*". Automatica, vol. 54, pp. 305–309, 2015. DOI: 10.1016/j.automatica.2015.02.015

[10].T. Ménard, E. Moulay, and W. Perruquetti, "*Fixed-time observer with simple gains for uncertain systems*," Automatica, vol.81, pp.438–446, 2017. DOI: 10.1016/j.automatica. 2017.04.009.

[11].X. Wang, J. Guo, S. Tang, S. Qi, "*Fixed-time disturbance observer based fixed-time back-stepping control for a flexible air-breathing hypersonic vehicle*". ISA Transactions, vol. 82, pp. 294–306, 2019. DOI: 10.1016/j.isatra.2018.12.013.

[12].G. Li, L. Liu, J. Liu, Y. Wu, and J. Zhao, "*Three-dimensional low-order fixed-time integrated guidance and control for STT missile with strap-down seeker*". Journal of the Franklin Institute, vol.360, no.13, pp.9788–9811, 2023. DOI:10.1016/j.jfranklin. 2023.07.028.

[13].E. Jiménez-Rodríguez, J. D. Sánchez-Torres, and A. G. Loukianov, "*Predefined-time backstepping control for tracking a class of mechanical systems*," IFAC PapersOnLine, vol. 50, no. 1, pp. 1680–1685, 2017. DOI: 10.1016/j.ifacol.2017.08.259.

[14].A. Garza-Alonso, M. Basin, and P. C. Rodriguez-Ramirez, "*Predefined-time backstepping stabilization of autonomous nonlinear systems*," IEEE/CAA Journal of Automatica Sinica, vol. 9, no. 11, pp. 2020–2022, Nov. 2022. DOI: 10.1109/JAS.2022.105953.

[15].Liwen Xu, Bailing Tian,... and Qun Zong, "*Fixed-time Disturbance Observer-Based MPC Robust Trajectory Tracking Control of Quadrotor*", IEEE/ASME Transactions on Mechatronics, vol. 30, no. 6, pp. 4272-4282, Dec. 2025, doi: 10.1109/TMECH.2024.10778610

Fixed-Time Integrated Guidance and Control for Maneuvering Targets under Uncertainties

Abstract*. This paper presents a Fixed-Time integrated guidance and control (IGC) scheme for aerial vehicles intercepting maneuvering targets under modeling uncertainties and external disturbances. The convergence time is explicitly specified as a design parameter, ensuring deterministic transient performance independent of initial conditions. A Fixed-Time disturbance observer is developed to estimate and compensate lumped uncertainties caused by target maneuvers and aerodynamic variations. Based on the strict-feedback IGC model, a Fixed-Time backstepping controller is constructed, guaranteeing prescribed-time convergence of both tracking and estimation errors. Lyapunov analysis establishes closed-loop stability. Comparative simulations demonstrate improved interception accuracy, enhanced robustness, and smoother responses compared with conventional backstepping-based methods.*

Keywords: Integrated guidance and control (IGC); Fixed-time backstepping control (FxTB); Fixed-time disturbance observer (FxTDO); Maneuvering target.

Thiết kế quỹ đạo thiết bị bay tránh mối đe dọa sử dụng Q-learning

Hồ Quang Huy[1*], Nguyễn Ngọc Điển[1], Đoàn Thế Tuấn[1],
Nguyễn Thanh Tùng[1], Nguyễn Việt Hoài Nam[1]
Học viện Kỹ thuật quân sự
Email: hqhuy1999@gmail.com; Tel: 0345159142

Tóm tắt

Bài nghiên cứu trình bày phương pháp thiết kế quỹ đạo cho thiết bị bay không người lái UAV trong môi trường có mối đe dọa, dựa trên thuật toán học tăng cường Q-learning. Bài toán được mô hình hóa dưới dạng quá trình quyết định Markov (MDP), trong đó UAV đóng vai trò tác nhân, môi trường bao gồm địa hình và các vùng nguy hiểm như radar hoặc khu vực phòng không. Hàm phần thưởng được xây dựng nhằm cân bằng giữa mục tiêu tối ưu quãng đường bay và yêu cầu tránh mối đe dọa. Thuật toán Q-learning được triển khai trong môi trường 2D và 3D, sử dụng bản đồ địa hình số Google Earch làm cơ sở mô phỏng. Kết quả cho thấy phương pháp đề xuất giúp UAV hình thành quỹ đạo bay thích nghi, giảm xác suất xâm nhập vùng nguy hiểm và duy trì khả năng hoàn thành nhiệm vụ trong điều kiện bất định.

Từ khóa: *Q-learning; Học tăng cường; Quỹ đạo UAV.*

1. Đặt vấn đề

Bài toán thiết kế quỹ đạo thiết bị bay trong môi trường có mối đe dọa là một bài toán ra quyết định tuần tự trong điều kiện bất định, được nghiên cứu rộng rãi trong lý thuyết dẫn đường và điều khiển tên lửa [1,2,5]. Các phương pháp truyền thống như dẫn tiếp cận tỉ lệ và các biến thể bù sai số động [3] dựa trên mô hình toán học xác định và đạt hiệu quả cao khi môi trường được biết trước. Tuy nhiên, trong môi trường tác chiến hiện đại với vùng radar, hệ thống phòng không và địa hình phức tạp, quỹ đạo tối ưu không còn là nghiệm hình học thuần túy mà phải thỏa mãn đồng thời yêu cầu an toàn và hiệu quả chiến thuật. Khi môi trường biến đổi hoặc chứa yếu tố bất định, cách tiếp cận thuần mô hình bộc lộ hạn chế về tính thích nghi.

Học tăng cường (RL) [6] cung cấp một khung tiếp cận phù hợp cho bài toán này bằng cách mô hình hóa hệ thống dưới dạng quá trình quyết định Markov [6,8]. Một số nghiên cứu đã áp dụng RL trong luật dẫn pha homing [11] hoặc tránh chướng ngại thời gian thực [10]. Tuy nhiên, phần lớn các công trình [10] tập trung vào suy luận thời gian thực dựa trên mạng học sâu trong môi trường giả định đơn giản hoặc chỉ xem xét bài toán tránh chướng ngại cục bộ. Trong khi đó, nghiên cứu [11] chủ yếu hướng đến tối ưu luật dẫn trong pha cuối mà chưa tích hợp đồng thời bài toán tránh mối đe dọa không gian và yếu tố địa hình. Ngoài ra, các công trình này thường sử dụng mô hình động học đơn giản hoặc môi trường ít thay đổi, chưa làm rõ mối quan hệ giữa cấu trúc hàm khen thưởng và yêu cầu chiến thuật tổng thể.

Khác với các nghiên cứu trên, bài báo này tập trung vào thiết kế quỹ đạo UAV tránh mối đe dọa trong môi trường rời rạc hai chiều và ba chiều có tích hợp dữ liệu địa hình số. Phương pháp đề xuất sử dụng thuật toán Q-learning [6] với cấu trúc hàm khen thưởng được xây dựng dựa trên các tiêu chí chiến thuật: giảm sai lệch dẫn đường, duy trì khoảng cách an toàn và hạn chế gia tốc điều khiển. Thay vì chỉ giải quyết pha homing hoặc tránh chướng ngại cục bộ, nghiên cứu hướng tới bài toán hoạch định quỹ đạo tổng thể trong môi trường có nhiều vùng nguy hiểm.

2. Học tăng cường

*Email: hqhuy1999@gmail.com

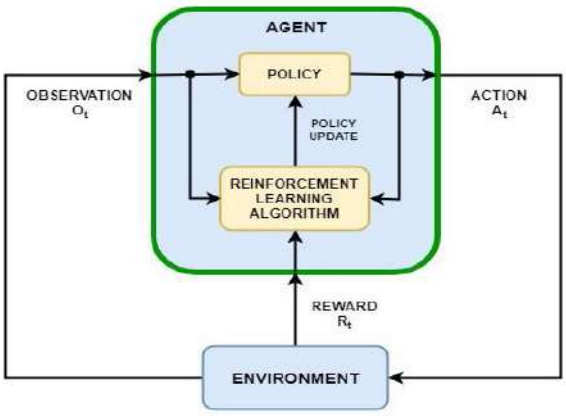
2.1 Khái niệm chung

Học tăng cường (Reinforcement Learning – RL) là một nhánh của học máy, trong đó tác nhân (agent) học cách ra quyết định thông qua quá trình tương tác lặp lại với môi trường nhằm tối đa hóa tổng phần thưởng tích lũy theo thời gian. Khác với học có giám sát, nơi tồn tại tập dữ liệu gán nhãn sẵn, học tăng cường không cung cấp lời giải đúng tại mỗi bước. Thay vào đó, tác nhân chỉ nhận được tín hiệu phản hồi định lượng dưới dạng phần thưởng sau khi thực hiện hành động.

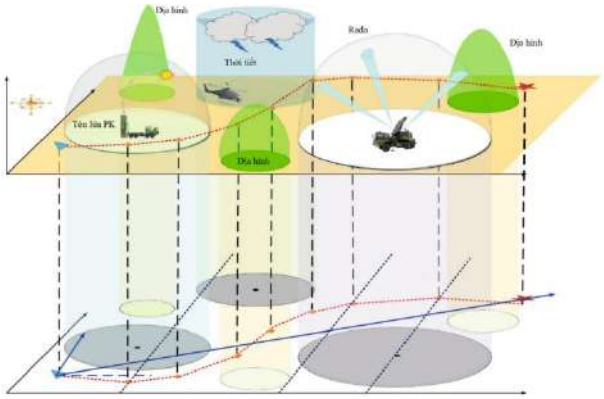
Một hệ thống RL điển hình bao gồm các thành phần:

- Tác nhân (Agent): thực thể đưa ra quyết định (trong bài toán này là UAV).
- Môi trường (Environment): không gian hoạt động chứa địa hình và mối đe dọa.
- Trạng thái (State, s): mô tả tình huống hiện tại của hệ thống.
- Hành động (Action, a): tập các lựa chọn khả thi tại trạng thái.
- Phần thưởng (Reward, r): tín hiệu phản hồi đánh giá hành động vừa thực hiện.

Mục tiêu của tác nhân là tìm một chính sách điều khiển tối ưu sao cho tổng phần thưởng chiết khấu là lớn nhất.



Hình 1. Sơ đồ cấu trúc học tăng cường



Hình 2. Mô hình UAV tránh mối đe dọa

2.2 Thuật toán Q-learning

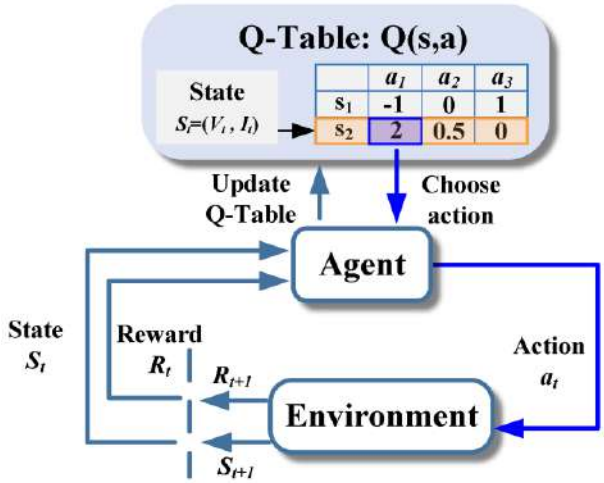
Q-learning là một thuật toán học tăng cường không theo mô hình (model-free), cho phép tác nhân tìm chính sách tối ưu mà không cần biết trước xác suất chuyển trạng thái của môi trường. Thuật toán làm việc trực tiếp với hàm giá trị trạng thái – hành động $Q(s,a)$, biểu diễn mức độ "tốt" của việc thực hiện hành động a tại trạng thái s.

Thuật toán cốt lõi của Q-learning dựa trên phương trình Bellman tối ưu cho hàm $Q^*(s,a)$:

$$Q^*(s,a) = R(s,a) + \gamma \max_{a'} Q^*(s',a') \tag{1}$$

Trong thực tế, giá trị tối ưu Q^*chưa biết và được xấp xỉ thông qua quá trình học lặp. Tại mỗi bước tương tác, bảng Q được cập nhật theo quy tắc:

$$Q(s,a) \leftarrow Q(s,a) + \alpha\left[r + \gamma \max_{a'} Q(s',a') - Q(s,a)\right] \tag{2}$$



Hình 3. Sơ đồ cấu trúc thuật toán Q-learning

3. Học tăng cường trong thiết kế quỹ đạo UAV

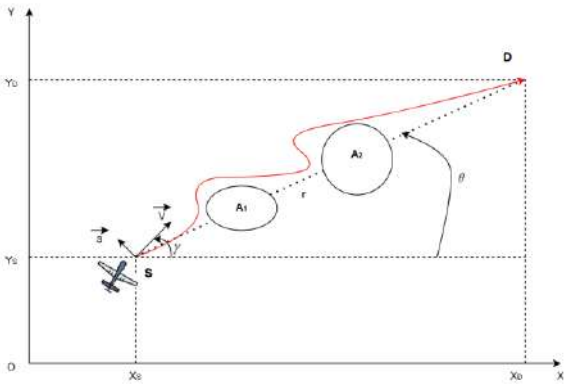
Phần này xây dựng bài toán thiết kế quỹ đạo UAV trong khung học tăng cường. Trước khi đưa bài toán vào mô hình RL, một số giả định cơ bản được đặt ra nhằm đơn giản hóa phân tích mà không làm mất tính tổng quát của vấn đề.

Các giả định được sử dụng như sau:

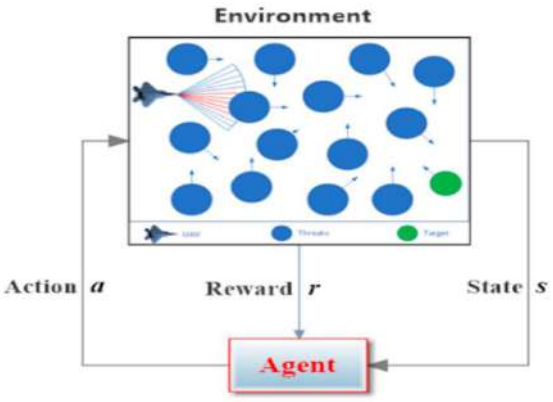
- Giả định 1: UAV và đích đến được xem là mô hình chất điểm.
- Giả định 2: Chuyển động được nghiên cứu trong mặt phẳng 2D (mở rộng 3D được trình bày trong phần mô phỏng).
- Giả định 3: Tốc độ của UAV được coi là không đổi, điều khiển thực hiện thông qua gia tốc pháp tuyến.

Các giả định này thường được chấp nhận trong nghiên cứu dẫn đường và thiết kế luật điều khiển quỹ đạo.

3.1. Động học tương đối của UAV so với đích



Hình 4. Động học tương đối của UAV



Hình 5. Mô hình UAV tránh mối đe dọa

Ta có mối quan hệ động hình học của UAV- đích đến được biểu diễn bởi các phương trình sau:

$$\dot{r} = -V\cos(\gamma - \theta) \tag{3}$$

$$\dot{\theta} r = -V \sin(\gamma - \theta) \tag{4}$$

$$\dot{\gamma} = \frac{a}{V} \tag{5}$$

Phương trình chuyển động của UAV

$$\dot{X}_S = V \cos(\gamma) \tag{6}$$

$$\dot{Y}_S = V \sin(\gamma) \tag{7}$$

3.2. Thuật toán Q-learning cho UAV

Để giải quyết vấn đề thiết kế quỹ đạo bằng thuật toán Q-learning, bài nghiên cứu trình bày cách xây dựng bài toán trong một khung học tăng cường bằng cách xây dựng một quá trình quyết định Makov (MDP) với hàm phần thưởng phù hợp. Xem các phương trình động học tương đối là môi trường tương tác:

Trạng thái hệ tại thời điểm t được chọn là:

$$s_t = (r, \theta, \dot{r}, \dot{\theta}) \tag{8}$$

Bộ trạng thái này phản ánh đầy đủ thông tin động học tương đối giữa UAV và mục tiêu.

Tác nhân UAV có các biến quan sát là: $o_t = (r, \theta, \dot{r}, \dot{\theta})$

Và hành động của tác nhân chính là ánh xạ tới gia tốc lệnh dẫn UAV.

Nghiên cứu đề xất một tác nhân dẫn Q-learning ánh xạ từ dữ liệu tới hệ số dẫn tiếp cận tỉ lệ:

$$N_{QL} = N_0 + f(r, \theta, \dot{r}, \dot{\theta}) \tag{9}$$

Khi này luật dẫn có dạng

$$a_c = N_{QL} \dot{\theta} V_c \tag{10}$$

Với $f(r, \theta, \dot{r}, \dot{\theta})$ là một hàm phi tuyến, Vc là vận tốc tiếp cận.

3.3. Xây dựng hàm khen thưởng cho thuật toán

Trong kịch bản này, hàm thưởng phạt cho bài toán lập kế hoạch đường bay cho UAV tránh các mối đe dọa được xây dựng như sau:

$$r_t = R_T + R_c + R_e + R_p + R_r \tag{11}$$

R_T : Thưởng khi đến đích

$$R_T = \begin{cases} k_T, & r < r_{end} \\ 0, & r > r_{end} \end{cases} \tag{12}$$

R_c **:** phạt khi UAV bay vào vùng cấm bay hoặc va chạm với mối đe dọa

$$R_c = \begin{cases} k_c, & \min\limits_j d_J(t+1) \le R_j, \\ 0, & other \end{cases} \tag{13}$$

Với $d_J(t+1) = \| p_{t+1} - c_{J,t+1} \|_2$: là khoảng cách không gian 3D giữa UAV và mối đe dọa thứ J tại thời điểm t + 1.

R_e : Khuyến khích UAV nhanh chóng bay đến đích.

$$R_e = \begin{cases} k_e, \dot{r} \le 0 \\ 0, \dot{r} > 0 \end{cases} \tag{14}$$

R_p : Phạt tiệm cận nguy hiểm theo khoảng cách an toàn

Đặt $d_{\text{safe},j} > R_j$ (biên an toàn đối với mối đe dọa thứ j).

Khi đó có hàm phạt khi UAV bay quá gần hoặc va chạm mối đe dọa

$$R_p == \lambda_p \sum_{j=1}^{J} \left[\max\left(0, d_{\text{safe},j} - d_J(t+1)\right) \right]^2 \tag{15}$$

Trong đó λ_p : trọng số phạt tiệm cận khi gần mối đe dọa

R_r : Phạt vi phạm động học

$$R_r = k_a \left(\frac{a}{a_{max}}\right)^2 \tag{16}$$

Dựa vào hàm phần thưởng đã xây dựng ở phần trước, khởi tạo các tham số cấu trúc của hàm phần thưởng: $r_t = R_T + R_c + R_e + R_p + R_r$

Bảng 1. Các tham số cấu trúc hàm khen thưởng

k_T	k_c	k_e	λ_p	k_a	$a_{\max}$ m/s^2
200	-300	10	-30	-2	30

4. Huấn luyện Q-learning

4.1 Điều kiện của bài toán

Kịch bản huấn luyện cho tác nhân Q-learning được thực hiện với các dữ kiện ban đầu như sau:

Bảng 2. Điều kiện ban đầu cho bài toán

Tham số	Giá trị tối thiểu	Giá trị tối đa
Vị trí ban đầu của UAV	(0;300)	
Vị trí đích đến của	(300;0)	
Góc nghiêng quỹ đạo ban đầu của UAV, γ	0	360^0
Vị trí mối đe dọa (2 mối đe dọa lớn, 3 mối đe dọa nhỏ)	Cố định các mối đe dọa	
Bản đồ lưới kích thước 300*300		

Bảng 3. Các tham số huấn luyện thuật toán Q-learning

Tham số	Giá trị
Maximum permissible steps	40000
max_step	500
Actor learning rate, α	0,1 – 0,5
Discounting factor, γ	0.95
ε-greedy exploration ε	1 → 0,01
Epsilon decay	0,995
σ_episode	1000

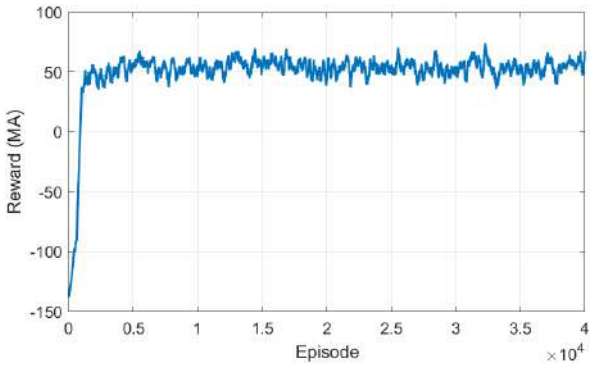
Với tốc độ UAV không đổi V=40 (m/s). Bài toán này sử dụng phương pháp tích phân Euler với bước tính thời gian 0.01s trong suốt quá trình mô phỏng. Gia tốc của UAV được giới hạn $|a_M| \leq 30\ (m/s^2)$.

Kết quả mô phỏng với phương pháp dẫn sử dụng tác nhân Q-learning

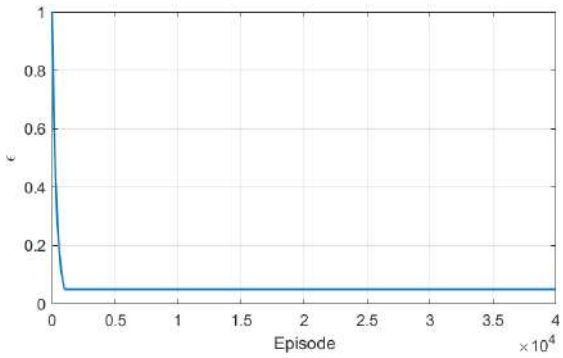
Cấu hình máy tính Dell XPS 13 9340 (2024): CPU - Intel Core Ultra 7 155H (16 nhân, 22 luồng, đến 4.8GHz, tích hợp NPU AI); GPU - Intel Arc Graphics; RAM - 16GB LPDDR5X 7467MT/s; SSD - 1TB NVMe PCIe 4.0.

4.2 Kết quả mô phỏng

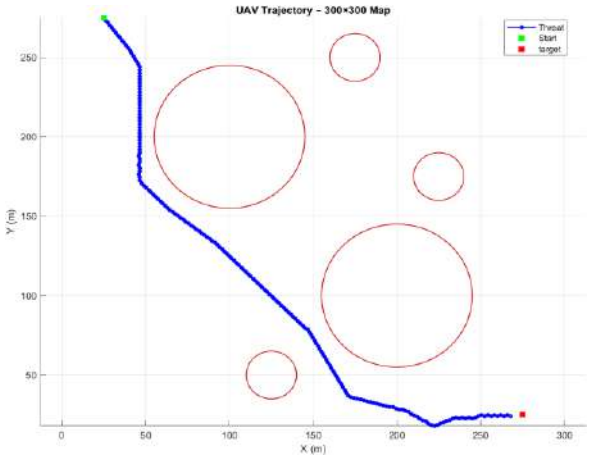
Kết quả huấn luyện mô hình với không gian 2D



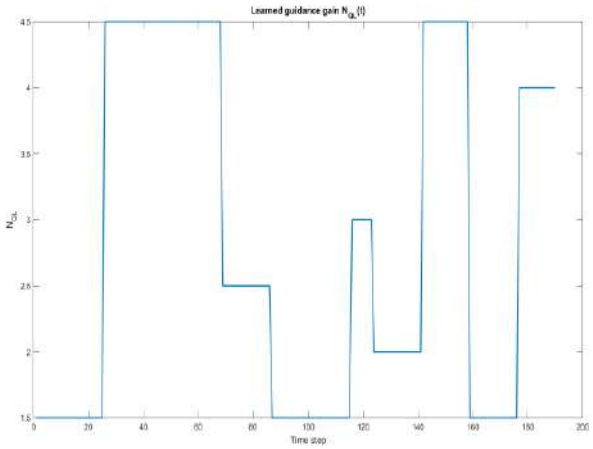
Hình 6. Biểu đồ phần thưởng cho Agent Q-learning



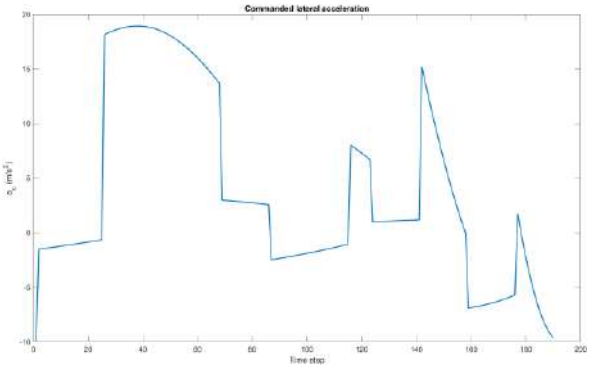
Hình 7. Biểu đồ Q-value trong quá trình huấn luyện Agen



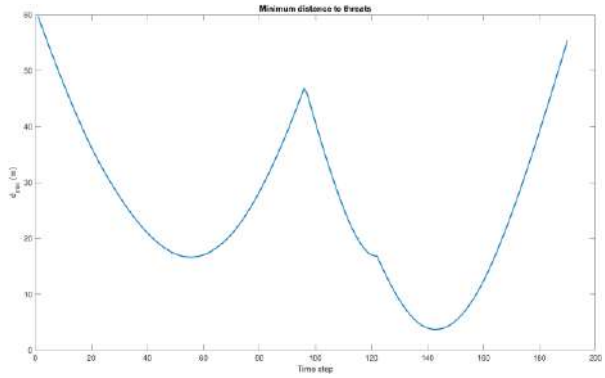
Hình 8. Quỹ đạo bay UAV



Hình 9. Đồ thị hệ số tiếp cận tỉ lệ

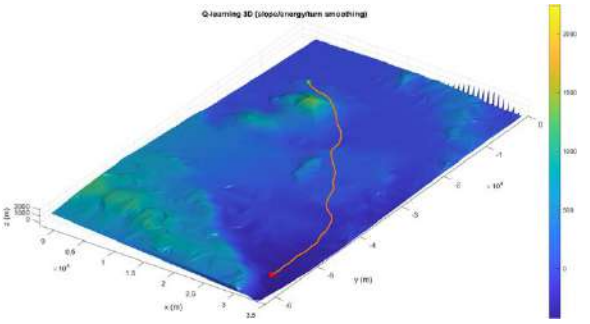


Hình 9. Đồ thị gia tốc UAV

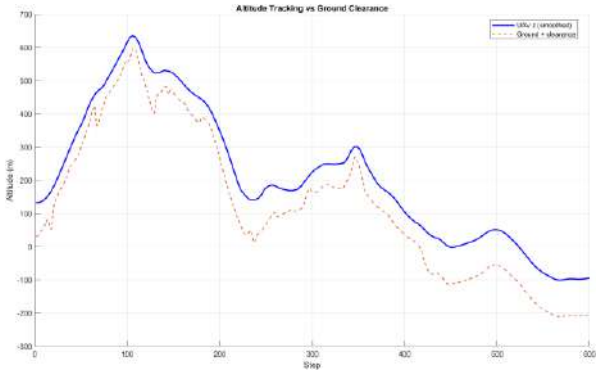


Hình 10. Đồ thị khoảng cách UAV so với các mối đe dọa

Kết quả huấn luyện mô hình với không gian 3D (địa hình thực địa lấy từ Google Earch).



Hình 11. Quỹ đạo bay UAV trên bản đồ số



Hình 12. Mặt cắt quỹ đạo của UAV

4.3 Kết luận

Nghiên cứu này đã trình bày phương pháp thiết kế quỹ đạo UAV trong môi trường có mối đe dọa dựa trên thuật toán học tăng cường Q-learning. Bài toán được mô hình hóa trong khung quá trình quyết định Markov (MDP), trong đó động học tương đối UAV – mục tiêu đóng vai trò môi trường tương tác, còn lệnh gia tốc pháp tuyến được xem là hành động của tác nhân.

Hàm khen thưởng được xây dựng theo cấu trúc đa thành phần, bao gồm: giảm khoảng cách, tránh mối đe dọa, hạn chế gia tốc điều khiển và thưởng đến đích. Cấu trúc này cho phép tác nhân tối ưu đồng thời hiệu quả dẫn đường và tính khả thi vật lý của hệ thống.

Kết quả mô phỏng cho thấy:

- Phần thưởng tích lũy hội tụ sau quá trình huấn luyện, chứng tỏ tác nhân học được chính sách ổn định.
- Quỹ đạo UAV trong không gian 2D có khả năng tự động tránh các vùng mối đe dọa và tiến tới mục tiêu.
- Gia tốc điều khiển được giữ trong giới hạn cho phép, bảo đảm tính khả thi thực tế.
- Khoảng cách UAV đến các mối đe dọa duy trì ở mức an toàn.
- Trong môi trường 3D có địa hình thực (trích xuất từ Google Earth), UAV học được quỹ đạo bám địa hình hợp lý và tránh va chạm với các mối đe dọa

Những kết quả này khẳng định rằng phương pháp Q-learning có khả năng giải quyết bài toán thiết kế quỹ đạo trong môi trường bất định mà không cần mô hình điều khiển giải tích tường minh. Thay vì xây dựng luật dẫn trực tiếp, hệ thống học được chính sách tối ưu thông qua tương tác với môi trường, qua đó tăng khả năng thích nghi khi điều kiện thay đổi.

Tài liệu tham khảo

1. Lê Anh Dũng, Nguyễn Hữu Độ, Huỳnh Lương Nghĩa, Lý thuyết bay và hệ thống điều khiển tên lửa phòng không - tập II, Học viện Kỹ thuật quân sự, Hà Nội 1998
2. Lê Anh Dũng, Nguyễn Hữu Độ, Huỳnh Lương Nghĩa, Lý thuyết bay và hệ thống điều khiển tên lửa phòng không - tập III, Học viện Kỹ thuật quân sự, Hà Nội 1998
3. Doãn Văn Minh, "Hoàn thiện phương pháp dẫn tiếp cận tỉ lệ theo hướng bù các sai số động" , Luận án tiến sỹ , Hà Nội 2015
4. Vũ Thị Thúy Nga, "Học tăng cường trong điều khiển tự động với Matlab & Simulink", nhà xuất bản Bách khoa Hà Nội, tập 1, 2023.
5. Paul Zarchan, Tactical and Strategic Missile Guidance, Third Edition, American Institute of Aeronautics & Ast, 1998
6. R. S. Sutton and A. G. Barto, Reinforcement learning: An introduction. MIT press, 2018
7. Zhu, Z.Xu, D.Liu, Xia, Y. (2012). Missile guidance law based on extended state observer. IEEE Transactions on Industrial Electronics, 60(12), 5882-5891
8. Henderson, P. Islam, R. Bachman, P. Pineau, J. Precup (2018). Deep reinforcement learning that matters. Paper presented at the Proceedings of the AAAI conference on artificial intelligenc
9. A. F. Agarap, "Deep learning using rectified linear units (relu)," 2019
10. Daseon Hong and Sungsu Park, "Avoiding obstacles via missile real-time inference by reinforcement learning", Journal of Advanced Missile Systems, 30(4):567–589, 2022
11. Brian Gaudet and Roberto Furfaro. "Missile homing phase guidance law design using reinforcement learning", Journal of Missile Technology, 29(1):123–145, 2012

Trajectory Design for Unmanned Aerial Vehicles under Threat Environments Using Q Learning

Abstract: *This paper presents a reinforcement learning based approach for trajectory design of unmanned aerial vehicles operating in threat environments. Unlike conventional guidance methods that rely on predefined mathematical models, the proposed method formulates the trajectory planning problem as a sequential decision making process. The unmanned aerial vehicle is modeled as an agent interacting with a dynamic environment containing terrain constraints and multiple threat regions. A reward function is constructed to balance interception performance, safety distance from threats, and control effort. The Q learning algorithm is employed to learn the optimal guidance policy without requiring prior knowledge of the environment model. Simulation results in two dimensional and three dimensional scenarios demonstrate that the learned policy can generate feasible trajectories that avoid threat zones while ensuring mission completion. The results confirm the adaptability and effectiveness of reinforcement learning in trajectory planning problems under uncertainty.*

Keywords: reinforcement learning; Q learning; unmanned aerial vehicle; trajectory planning; threat avoidance

Nghiên cứu và tổng hợp bộ điều khiển thích nghi cho hệ truyền động lái đa kênh trên máy bay

Lê Duy Khiên[1] *, Doãn Văn Minh[1], Nguyễn Đình Khánh[2]

[1]Học viện Kỹ thuật quân sự; [2]Học viện Hải quân

Email: leduykhienhvhq@gmail.com; Tel: 0345011094

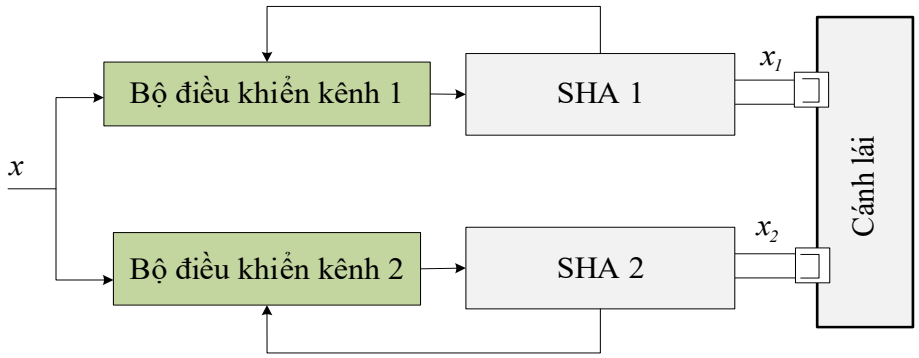
Tóm tắt

Bài báo trình bày phương pháp mô hình hóa và thiết kế bộ điều khiển thích nghi theo mô hình mẫu (Model Reference Adaptive Control – MRAC) cho hệ hai cơ cấu chấp hành servo–thủy lực (SHA–SHA) cùng tác động lên một liên kết đầu ra chung (bề mặt lái). Hệ truyền động SHA–SHA thường chịu ảnh hưởng của nhiều đặc tính phi tuyến, bao gồm ma sát nhớt, sự biến thiên các tham số dòng chất lỏng và hiện tượng xung đột lực (force fighting) giữa các kênh điều khiển. Trên cơ sở xây dựng mô hình động lực học của hệ servo–thủy lực, bài báo tiến hành tổng hợp bộ điều khiển thích nghi MRAC cho từng kênh và đề xuất một cơ chế cân bằng lực đối xứng nhằm giảm chênh lệch lực giữa hai kênh SHA. Luật thích nghi được xây dựng dựa trên phân tích hàm Lyapunov bảo đảm tính ổn định toàn cục của hệ thống. Kết quả mô phỏng trên nền tảng MATLAB/Simulink cho thấy bộ điều khiển đề xuất bảo đảm khả năng hội tụ của đáp ứng hệ thực về mô hình mẫu, đồng thời làm suy giảm đáng kể lực đối kháng giữa các kênh khi các tham số thủy lực biến thiên. Các kết quả thu được khẳng định tính khả thi và hiệu quả của phương pháp MRAC đối với hệ truyền động lái đa kênh trên các máy bay hiện đại.

Từ khóa: *MRAC; servo–thủy lực; hệ hai cơ cấu chấp hành servo–thủy lực (SHA–SHA); cân bằng lực; force fighting.*

1. Đặt vấn đề

Trong các hệ thống điều khiển bay hiện đại, đặc biệt trên máy bay dân dụng và quân sự, các yêu cầu về độ tin cậy, an toàn và khả năng chịu lỗi của hệ truyền động bề mặt lái ngày càng trở nên khắt khe. Để đáp ứng các yêu cầu này, cấu trúc truyền động đa kênh sử dụng hai cơ cấu chấp hành servo–thủy lực (SHA–SHA) tác động song song lên một bề mặt lái chung đã được ứng dụng rộng rãi trên nhiều dòng máy bay thương mại và quân sự như A320, B737, B777 [3]. Cấu trúc này cho phép chia sẻ tải khí động và duy trì khả năng điều khiển ngay cả khi một kênh truyền động gặp sự cố (Hình 1).



Hình 1. Hệ thống lái điều khiển hai kênh

* Email: leduykhienhvhq@gmail.com

Trong hệ thống SHA–SHA, hai cơ cấu chấp hành hoạt động đồng thời trên cùng một liên kết đầu ra chung nhằm nâng cao độ tin cậy và khả năng chịu lỗi. Tuy nhiên, các cơ cấu servo–thủy lực chịu ảnh hưởng đáng kể của các đặc tính phi tuyến như rò rỉ dòng chất lỏng, ma sát nhớt, trễ và bão hòa van điều khiển, cũng như sự biến thiên tham số theo thời gian do điều kiện làm việc và lão hóa hệ thống. Các yếu tố phi tuyến này dễ gây ra sai lệch vị trí làm việc giữa hai kênh, dẫn đến hiện tượng xung đột lực. Hiện tượng này không chỉ làm tăng tổn hao năng lượng, sinh nhiệt và mài mòn cơ khí mà còn ảnh hưởng bất lợi đến độ ổn định và tuổi thọ của hệ truyền động [3, 5].

Các nghiên cứu hiện có về hệ truyền động lái đa kênh chủ yếu tập trung vào mô hình hóa và điều khiển một cơ cấu chấp hành servo–thủy lực đơn lẻ với tải cơ học đơn giản. Trong khi đó, các công trình nghiên cứu hệ lái đa kênh thường hướng tới các kiến trúc sử dụng cơ cấu chấp hành điện hoặc điện–thủy lực (EMA–EHA) kết hợp với các phương pháp điều khiển robust hoặc điều khiển tối ưu. Những nghiên cứu này mới chỉ xem xét một cách hạn chế tương tác lực nội bộ giữa các kênh và chưa giải quyết triệt để bài toán đồng bộ lực trong điều kiện tham số hệ thống biến thiên mạnh. Đặc biệt, các cấu trúc điều khiển thích nghi theo mô hình mẫu (MRAC) đầy đủ dựa trên phân tích Lyapunov cho hệ hai SHA làm việc song song hầu như chưa được nghiên cứu một cách hệ thống [1, 4]. Một số công trình gần đây, tiêu biểu như Kuznetsov (2020), đã đề cập đến bài toán đồng bộ lực giữa các cơ cấu chấp hành khác loại. Tuy nhiên, các nghiên cứu này chưa khai thác kiến trúc MRAC độc lập cho từng kênh servo–thủy lực, đồng thời chưa xây dựng được một vòng cân bằng lực tích hợp trong khung điều khiển thích nghi có chứng minh ổn định Lyapunov chặt chẽ. Các phương pháp điều khiển cổ điển như PID hoặc các chiến lược robust tuy có thể cải thiện độ ổn định cục bộ, nhưng thường gặp hạn chế khi xử lý các tham số thủy lực biến thiên theo thời gian và các tương tác lực nội bộ trong hệ đa kênh [6].

Trong bối cảnh đó, điều khiển MRAC được xem là một hướng tiếp cận phù hợp cho bài toán điều khiển hệ truyền động SHA–SHA đa kênh. Đặc trưng cốt lõi của MRAC là bảo đảm đáp ứng của hệ thống thực tiệm cận về đáp ứng của mô hình mẫu đã cho, ngay cả khi các tham số động lực học của đối tượng là không biết hoặc biến thiên theo thời gian, với tính ổn định được chứng minh chặt chẽ trên cơ sở phân tích Lyapunov. Nhờ đặc tính này, MRAC cho phép duy trì chất lượng điều khiển và tính ổn định của hệ truyền động trong điều kiện làm việc thay đổi, đồng thời tạo nền tảng thuận lợi để tích hợp các cơ chế đồng bộ vị trí và cân bằng lực giữa các kênh servo–thủy lực làm việc song song. Do đó, việc nghiên cứu và phát triển cấu trúc MRAC cho hệ truyền động SHA–SHA là cần thiết nhằm nâng cao độ tin cậy và hiệu quả của hệ truyền động lái trên máy bay hiện đại.

2. Phương pháp nghiên cứu và mô hình hóa

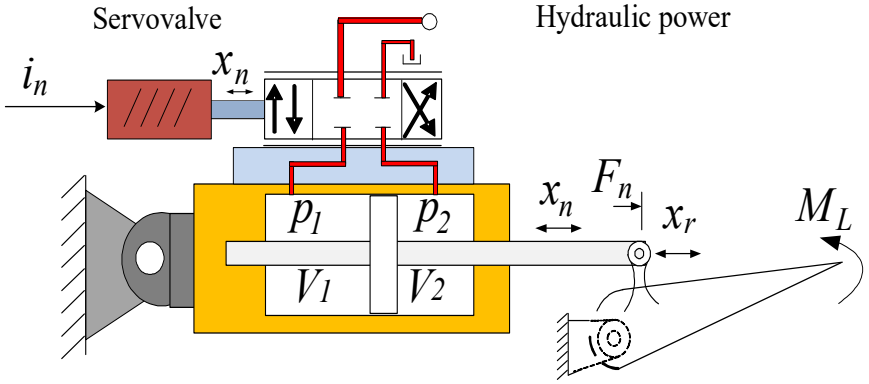
Các hệ thống điều khiển bề mặt lái trên máy bay, bao gồm 2 bộ truyền động điện–thủy lực (SHA) với điều khiển tốc độ bướm ga, có cấu hình giống nhau và hoạt động theo chế độ tổng hợp các lực mà chúng phát triển trên một liên kết đầu ra chung – bề mặt lái. Sơ đồ động học của truyền động điện–thủy lực của mỗi kênh được thể hiện trong Hình 2. Mô hình toán của mỗi kênh được mô tả bằng các phương trình sau [3, 4]:

- Phương trình mô tả chuyển động của van điều khiển tuyến tính liên quan đến dòng điện van điều khiển:

$$\tau_n \dot{x}_n = -x_n + k_n i_n \tag{1}$$

- Phương trình mô tả lưu lượng tải Q_n liên quan đến sự dịch chuyển của van điều khiển x_n áp suất tải P_n :

$$Q_n = \mu_n \sqrt{\frac{1}{\rho_n}}\, b_n x_n \sqrt{(P_s - P_n \operatorname{sgn}(x_n))} \tag{2}$$



Hình 2. Sơ đồ SHA

- Các phương trình cân bằng lưu lượng chất lỏng với sự tính đến tính nén và rò rỉ và phương trình chuyển động của xilanh thủy lực có dạng (3) và (4), tương ứng:

$$Q_n = A_n \dot{x}_n + \frac{V_n}{2E_n}\dot{P}_n + C_n P_n \tag{3}$$

$$A_n P_n = m_n \ddot{x}_n + B_n \dot{x}_n + F_n \tag{4}$$

- Mô hình toán của bề mặt lái [2, 3] được mô tả bởi các phương trình (5) và (6):

$$F_n = K_n (x_n - x_r) \tag{5}$$

$$F_\Sigma r_r = J_r \ddot{\theta}_r + b_r \dot{\theta}_r + M_L \,. \tag{6}$$

Trong Hình 2 và các phương trình (1) - (6), các ký hiệu sau đây được sử dụng: n – chỉ số của kênh n (n =1, 2); τ và k – hằng số thời gian và hệ số khuếch đại của bộ khuếch đại điện thủy lực, tương ứng; i – dòng điện của động cơ mô men xoắn; P_S – áp suất cung cấp; μ – hệ số lưu lượng; b – gradient diện tích, liên kết diện tích lỗ với sự dịch chuyển của van điều khiển; ρ – mật độ chất lỏng; A – diện tích của pít-tông; m – trọng lượng của thanh pít-tông của xilanh thủy lực; B – ma sát nhớt của xilanh thủy lực; C – hệ số rò rỉ; x – sự dịch chuyển của SHA; F – lực đầu ra của SHA; θ_r – góc độ của bánh lái; r_r – cần điều khiển bề mặt; J_r – mô-men quán tính tương đương; K – hệ số độ cứng tương đương của liên kết giữa cơ cấu chấp hành servo–thủy lực và bề mặt điều khiển chuyển bay; x_r – sự dịch chuyển của bánh lái; b_r – hệ số giảm chấn của bánh lái; F_Σ – tổng hợp các lực của hai kênh lên bề mặt bánh lái; M_L – tải trọng bên ngoài tác dụng lên bánh lái.

Sau khi tuyến tính hóa hệ thống (tuyến tính hóa phương trình (2)) tại các điểm hoạt động, mô tả toán học trong không gian trạng thái của mỗi kênh truyền động có dạng tổng quát sau đây [1, 4]:

$$\dot{\mathbf{x}} = \mathbf{A}_0\mathbf{x} + \mathbf{B}_0\mathbf{u} + \boldsymbol{\sigma}\ ,\ y = \mathbf{C}_0\mathbf{x}\ , \tag{7}$$

Với $\mathbf{B}_0 = [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad k_n / \tau_n]^T$,

$$\mathbf{A}_0 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{-r_r^2 K_n}{J_r} & \frac{-b_r}{J_r} & \frac{r_r K_n}{J_r} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \frac{K_n r_r}{m_n} & 0 & \frac{-K_n}{m_n} & \frac{-B_n}{m_n} & \frac{A_n}{m_n} & 0 \\ 0 & 0 & 0 & \frac{-2A_n E_n}{V_n} & \frac{-2E_n(\mathrm{C}_n + K_{QP})}{V_n} & \frac{2E_i K_{Qx}}{V_n} \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau_n} \end{bmatrix},$$

$\boldsymbol{\sigma}$ - hàm vector của nhiễu; các hệ số K_{Qx} và K_{QP} là kết quả của việc tuyến tính hóa phương trình (2) tại các điểm hoạt động ($Q_h = K_{Qx}x_v - K_{QP}P_h$).

3. Tổng hợp bộ điều khiển thích nghi cho hệ thống truyền động lai SHA–SHA

3.1. Tổng hợp bộ điều khiển thích nghi cho từng kênh SHA

Đối với mỗi kênh thủy lực, chúng tôi lựa chọn phương pháp xây dựng bộ điều khiển thích nghi theo mô hình MRAC thông qua một quy trình thiết kế có cấu trúc. Quá trình tổng hợp bộ điều khiển bắt đầu bằng việc xây dựng một mô hình tham chiếu phù hợp nhằm xác định đặc tính động học vòng kín mong muốn của hệ thống. Tiếp theo, một bộ quan sát trạng thái được thiết kế để ước lượng các biến trạng thái không thể đo trực tiếp, qua đó cho phép thực hiện điều khiển phản hồi toàn trạng thái. Cuối cùng, một cơ chế điều khiển thích nghi được thiết lập để điều chỉnh các tham số bộ điều khiển theo thời gian thực, bảo đảm động học của cơ cấu chấp hành bám theo mô hình tham chiếu bất chấp sự không chắc chắn và biến thiên của các tham số hệ thống.

a) Xây dựng mô hình tham chiếu.

Động lực học tham chiếu của một hệ thống được cho bởi phương trình:

$$\dot{\mathbf{x}}_m = \mathbf{A}_m\mathbf{x}_m + \mathbf{B}_m u \tag{8}$$

Trong đó $\mathbf{A}_m = \mathbf{A}_0 - \mathbf{B}_0\mathbf{k}$; $\mathbf{B}_m = \mathbf{B}_0$; $\mathbf{k} = [k_1 \quad k_2 \quad k_3 \quad k_4 \quad k_5 \quad k_6]$, $\mathbf{k}$ là ma trận hệ số điều chỉnh mô hình tham chiếu. Các hệ số này được tính toán sau khi đã lựa chọn và xác định đa thức đặc trưng mong muốn của hệ kín, chẳng hạn như đa thức dạng Newton hoặc Butterworth:

$$\mathrm{H}_0(s) = a_0 s^6 + a_1\omega_0 s^5 + a_2\omega_0^2 s^4 + a_3\omega_0^3 s^3 + a_4\omega_0^4 s^2 + a_5\omega_0^5 s + a_6\omega_0^6$$

Trong đó a_1, a_2, a_3, a_4, a_5, a_6 là các hệ số của đa thức Newton hoặc Butterworth đã chọn.

b) Xây dựng bộ quan sát trạng thái.

Để đánh giá vector trạng thái của hệ thống điều khiển $\hat{x}(t)$, bộ quan sát trạng thái đầy đủ được sử dụng [1], [4]:

$$\dot{\hat{\mathbf{x}}} = \mathbf{A}_0\hat{\mathbf{x}} + \mathbf{B}_0 u + \mathbf{L}(y - \mathbf{C}\hat{\mathbf{x}}) \tag{9}$$

Trong đó $\mathbf{L}$ là vector các hệ số phản hồi của bộ quan sát, được xác định bởi phương trình:

$$\det[s\mathbf{I} - \mathbf{A}_0 + \mathbf{LC}] = \mathrm{H}_L(s) \tag{10}$$

Trong đó: $\mathrm{H}_L(s) = a_0 s^6 + a_1\omega_L s^5 + a_2\omega_L^2 s^4 + a_3\omega_L^3 s^3 + a_4\omega_L^4 s^2 + a_5\omega_L^5 s + a_6\omega_L^6$, $\omega_L = (3 \div 5)\omega_0$.

c) Tổng hợp luật điều khiển thích nghi cho từng kênh

Quy tắc điều khiển thích nghi được tổng hợp để đáp ứng mục tiêu điều khiển:

$$\lim_{t\to\infty} \mathbf{e}(t) = \lim_{t\to\infty} (\mathbf{x} - \mathbf{x}_m) = 0 \tag{11}$$

Trong đó $\mathbf{e} = \mathbf{x} - \mathbf{x}_m$ (hoặc $\mathbf{e} = \hat{\mathbf{x}} - \mathbf{x}_m$). Điều kiện tiền đề cho sự hội tụ tiệm cận của giải pháp hệ thống thích nghi là việc quan sát các điều kiện nhất quán:

$$(\mathbf{A} - \mathbf{A}_m)\mathbf{x} = \mathbf{B}\mathbf{B}^+(\mathbf{A} - \mathbf{A}_m)\mathbf{x} \text{ và } (\mathbf{B} - \mathbf{B}_m)\mathbf{x} = \mathbf{B}\mathbf{B}^+(\mathbf{B} - \mathbf{B}_m)\mathbf{x}$$

Trong đó $(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T = \mathbf{B}^+$, $\mathbf{B}^+$ là ma trận giả nghịch đảo của $\mathbf{B}$.

Đối với hệ thống (7)–(9), thuật toán thích nghi – điều này đảm bảo tỷ lệ giảm tối đa của các hàm Lyapunov cho các nghiệm của hệ thống sẽ có dạng quy tắc điều khiển sau:

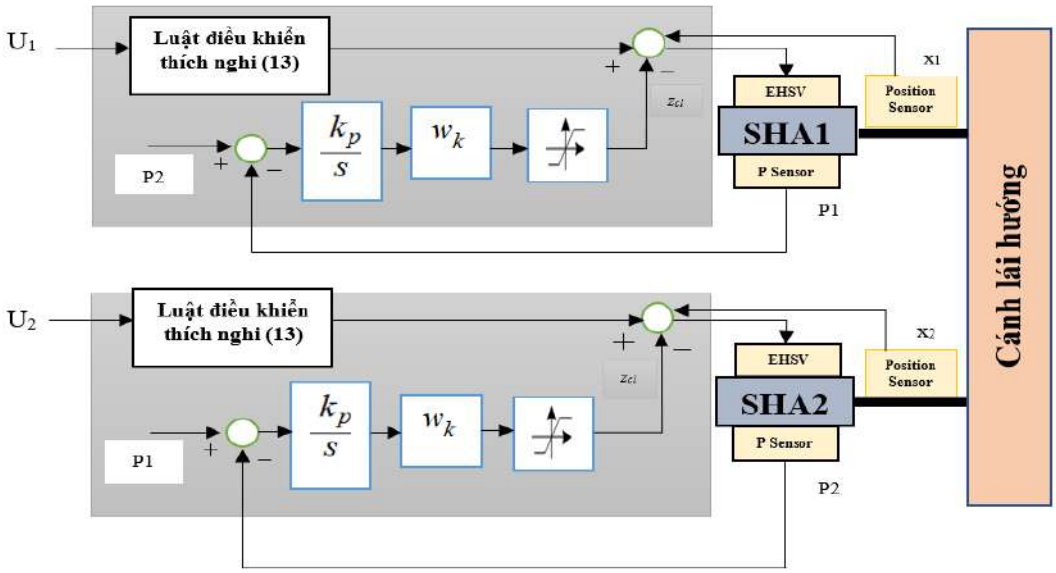
$$u(t) = g + z\ (t) \tag{12}$$

$$z(t) = -h\,\mathrm{sign}(\mathbf{B}_m^{\mathrm{T}}\mathbf{Pe}) \tag{13}$$

Trong đó $z(t)$ là tín hiệu điều khiển thích nghi, g là tín hiệu điều khiển chương trình, $h = const > 0$ là một hệ số khuếch đại đặc trưng cho biên độ (giới hạn) của điều khiển rơ-le, ma trận khuếch đại $\mathbf{P} = \mathbf{P}^T > 0$, được xác định từ nghiệm của phương trình $\mathbf{A}_m^T\mathbf{P}+\mathbf{P}\mathbf{A}_{\mathrm{m}} = -\mathbf{Q}$, ($\mathbf{Q} = \mathbf{Q}^T > 0$).

3.2. Phát triển bộ điều khiển cân bằng lực tải cho hệ thống

Thuật toán cân bằng lực tương tác giữa các kênh trong hệ thống truyền động lái đa kênh dựa trên sự chênh lệch áp suất [2, 5].



Hình 3. Cấu trúc điều khiển cân bằng lực của hệ thống truyền động lái đa kênh

Sơ đồ chung của thuật toán cân bằng lực trong các cơ cấu truyền động làm việc song song được trình bày trong Hình 3, trong đó, các ký hiệu được sử dụng là: u_1, u_2 – các tín hiệu điều khiển; P_1 – áp suất công tác kênh 1; P_2 – áp suất công tác kênh 2, k_p – hệ số khuếch đại áp suất; EHSV – bộ khuếch đại công suất điện – thuỷ lực; P Sensor – cảm biến chênh lệch áp suất; Position Sensor – cảm biến vị trí.

Quy tắc điều khiển cho mỗi kênh được hình thành từ tín hiệu điều chỉnh thích nghi và tín hiệu đồng bộ hóa. Quy tắc điều khiển thích nghi, được tổng hợp bằng phương pháp điều khiển thích nghi theo mô hình tham chiếu, có dạng (13).

Tín hiệu điều khiển đồng bộ được mô tả như sau [1, 5]:

$$z_{c_n} = -U_{cp_n} \frac{k_p}{s} W_k \tag{14}$$

Ở đây: $U_{cp_n} = \frac{1}{n}\sum_{i=1}^{n} P_i - P_i$ – độ lệch tải trong mỗi kênh từ giá trị trung bình; s – toán tử Laplace; W_k – liên kết điều chỉnh pha liên tiếp (có nhiệm vụ điều chỉnh pha của tín hiệu điều khiển để cải thiện sự ổn định và hiệu quả của hệ thống. Trong ngữ cảnh của thuật toán, đây là một yếu tố giúp điều chỉnh tín hiệu để hệ thống hoạt động hiệu quả hơn, đặc biệt là trong các hệ thống điều khiển động học phức tạp):

$$W_k = \frac{T_1 s + 1}{T_2 s + 1}, \text{ ở đây } T_1 \gg T_2 \tag{15}$$

4. Kết quả mô phỏng

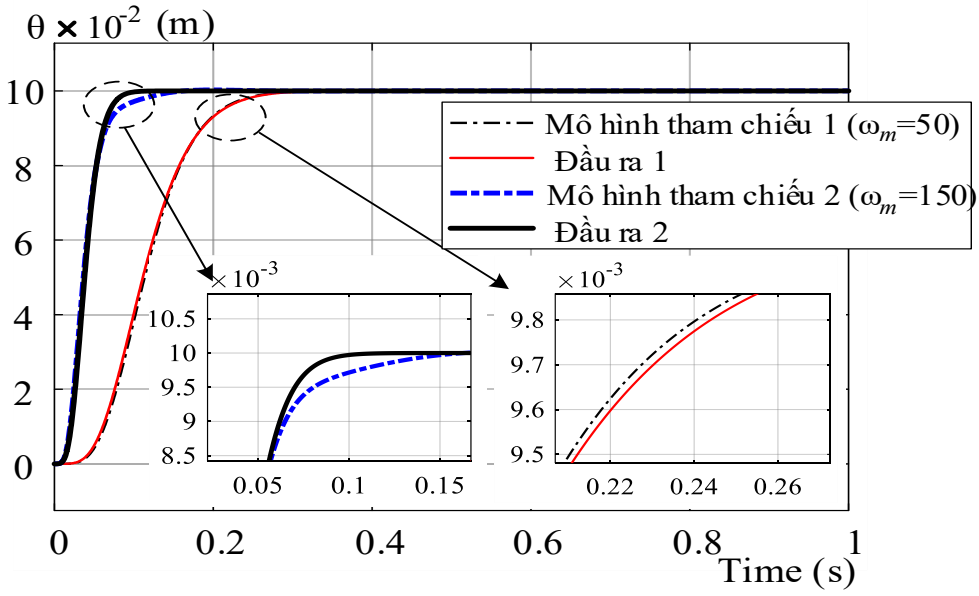
Các tham số của hệ servo–thủy lực được lựa chọn theo mô hình cơ cấu chấp hành điện – thủy lực điều khiển bề mặt lái máy bay (flight-control electro-hydraulic actuator) điển hình được trình bày trong [3-5]. Các giá trị này được lựa chọn để đại diện cho một cơ cấu chấp hành thủy lực điều khiển bề mặt lái cỡ trung bình sử dụng trong hệ thống điều khiển bay của máy bay dân

dụng và quân sự, với áp suất nguồn thủy lực khoảng 20 –21 MPa và cấu trúc servo-valve-cylinder tiêu chuẩn. Các giá trị tham số sử dụng trong mô phỏng được trình bày trong Bảng 1.

Bảng 1. Tham số hệ truyền động SHA

STT	Tham số	Ký hiệu	Giá trị	Đơn vị
1	Áp suất nguồn	P_s	21	MPa
2	Môđun nén chất lỏng	β_e	$1,5 \times 10^9$	Pa
3	Hệ số rò rỉ	C_l	2×10^{-12}	$m^3/(s \cdot Pa)$
4	Mật độ dầu	ρ	850	kg/m^3
5	Hệ số lưu lượng van	K_q	$1,5 \times 10^{-4}$	$m^3/(s \cdot m)$
6	Diện tích piston	A	0,002	m^2
7	Khối lượng tương đương	M	35	kg
8	Hệ số ma sát nhớt	B_v	1500	N·s/m
9	Mô men quán tính	J	0,8	$kg \cdot m^2$
10	Độ cứng liên kết	K	2×10^5	N/m
11	Hệ số giảm chấn	D	1200	N·s/m
12	Tải ngoài (Tải khí động)	F_{ext}	300	N

Để đánh giá một cách toàn diện hiệu năng của hệ thống, một loạt các mô phỏng đã được thực hiện dựa trên luật điều khiển (13) cũng như tín hiệu điều khiển đồng bộ (14). Đầu tiên, đáp ứng đầu ra của từng kênh được đánh giá khi áp dụng tín hiệu bước vào đầu vào điều khiển. Kết quả mô phỏng (Hình 4) cho thấy đầu ra của từng kênh nhanh chóng hội tụ về quỹ đạo của mô hình tham chiếu mà không xuất hiện dao động.



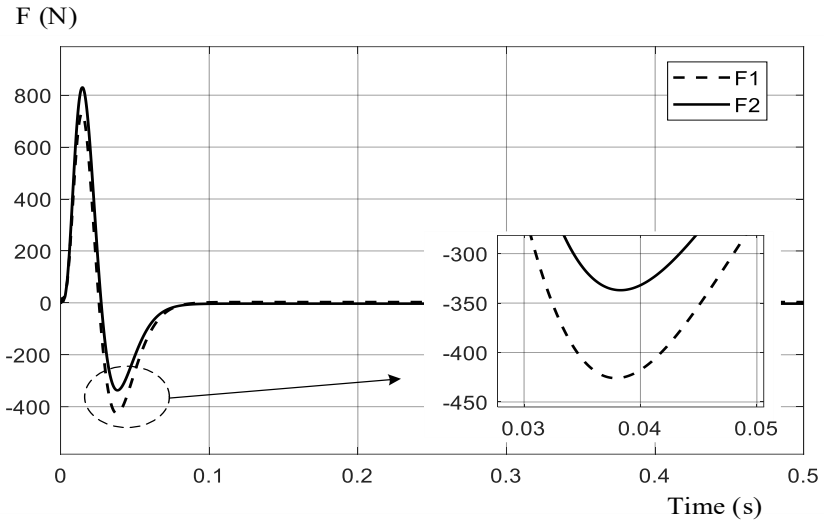
Hình 4. Phản ứng đầu ra kênh 1 (2) với tín hiệu đầu vào dạng step

Ngay cả khi các tham số của mô hình tham chiếu được thay đổi (ví dụ, $\omega_0 = 50$, $\omega_0 = 150$), tín hiệu đầu ra được điều khiển vẫn luôn bám sát các quỹ đạo tương ứng của mô hình tham chiếu. Điều này chứng tỏ bộ điều khiển thích nghi đề xuất có khả năng định

hình hiệu quả đặc tính quá độ của từng kênh, đồng thời duy trì chất lượng bám vững chắc trước sự thay đổi của các tham số mô hình.

Tiếp theo các mô phỏng được thực hiện để so sánh đáp ứng lực đầu ra của các kênh trong hệ truyền động lái lai (F_1, F_2) khi áp dụng thuật toán đồng bộ theo chênh lệch áp suất (14).

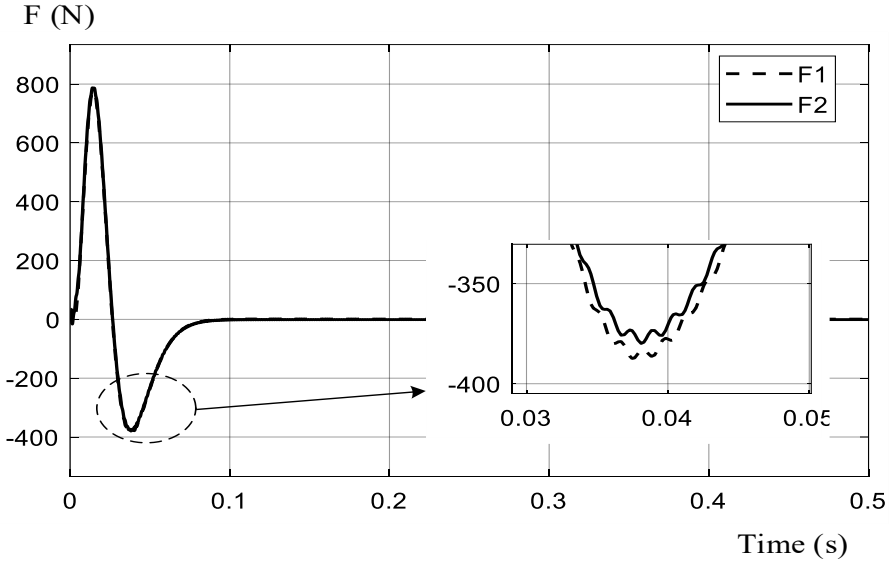
Hình 5 trình bày đáp ứng lực đầu ra của hai kênh trong hệ thống truyền động lái khi không sử dụng tín hiệu điều khiển đồng bộ.



Hình 5. Phản ứng lực đầu ra các kênh khi chưa có tín hiệu đồng bộ

Có thể nhận thấy rằng trong chế độ động, sai lệch giữa các quỹ đạo lực của các kênh trong hệ là rất đáng kể (ΔF=100N), tức là không bảo đảm được sự đồng bộ giữa các kênh của hệ thống.

Hình 6 thể hiện đáp ứng lực đầu ra của hai kênh khi áp dụng thuật toán đồng bộ (14).



Hình 6. Phản ứng lực đầu ra các kênh khi có tín hiệu đồng bộ

Kết quả cho thấy trong chế độ động, chất lượng đồng bộ của hệ thống được cải thiện rõ rệt; sai lệch giữa các quỹ đạo lực của các kênh giảm xuống (ΔF=10N), chứng tỏ hiệu quả của thuật toán đề xuất trong việc nâng cao khả năng phối hợp đồng bộ giữa các kênh truyền động.

5. Kết luận

Bài báo đã trình bày một phương pháp mô hình hóa và tổng hợp bộ điều khiển thích nghi theo mô hình mẫu (MRAC) cho hệ truyền động lái hai kênh sử dụng cơ cấu chấp hành servo–thủy lực dạng SHA–SHA tác động lên cùng một bề mặt lái. Trên cơ sở xây dựng mô hình động lực học có xét đến các yếu tố phi tuyến đặc trưng như ma sát nhớt, sự biến thiên tham số thủy lực và hiện tượng xung đột lực giữa các kênh, bộ điều khiển MRAC đã được thiết kế cho từng kênh điều khiển, đồng thời đề xuất cơ chế cân bằng lực đối xứng nhằm giảm thiểu sai lệch lực giữa hai cơ cấu chấp hành. Luật thích nghi được xây dựng dựa trên phương pháp Lyapunov, bảo đảm tính ổn định toàn cục của hệ thống.

Kết quả mô phỏng trên nền tảng MATLAB/Simulink cho thấy hệ thống điều khiển đề xuất bảo đảm đáp ứng của hệ thực hội tụ về mô hình tham chiếu và duy trì chất lượng bám trong điều kiện các tham số thủy lực thay đổi, đồng thời làm suy giảm đáng kể hiện tượng xung đột lực giữa các kênh. Những kết quả này khẳng định tính khả thi và hiệu quả của phương pháp MRAC trong ứng dụng cho các hệ truyền động lái đa kênh trên máy bay hiện đại, đồng thời mở ra hướng nghiên cứu tiếp theo về triển khai thực nghiệm và đánh giá trên hệ thống phần cứng thực.

Tài liệu tham khảo

1.	Kuznetsov, V. E., & Khanh, N. D. (2020). Synthesis of an adaptive controller for movement synchronization in a multi-channel steering system. *Proceedings of the International Conference on Control in Technical Systems (CTS)*. St. Petersburg, Russia.
2.	Kuznetsov, V. E., & Khanh, N. D. (2019). Алгоритм синхронизации многоканального рулевого привода с адаптивным регулятором. *Proceedings of a Russian Conference on Control Systems and Automation*. St. Petersburg, Russia.
3.	Konstantinov, S. V., Kuznetsov, V. E., et al. (2019). *Systems of electro-hydraulic flight control actuators of aircraft flight control systems*. St. Petersburg: Publishing House SPbSETU LETI.
4.	Khanh, N. D., Kuznetsov, V. E., & Vung, C. N. (2021). Synthesizing an adaptive controller to enhance movement quality of a flight control actuator under external load. *Proceedings of the XXIII International Conference on Conference on Control in Technical Systems (CTS),* 146-149. https://doi.org/10.1109/CTS53513.2021.9562850.
5.	Kuznetsov, V. E., Khanh, N. D., & Lukichev, A. N. (2020). System for synchronizing forces of dissimilar flight control actuators with a common controller. *Proceedings of the XXIII International Conference on Soft Computing and Measurements (SCM)*, 137-140. https://doi.org/10.1109/SCM50615.2020.9198768.
6.	Kuznetsov, V. E., Khanh, N. D., Lukichev, A. N., & Filatov, D. M. (2021). Hybrid steering system's PID-based adaptive control. *Proceedings of the IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus),* 979-984. https://doi.org/10.1109/ElConRus51938.2021.9396303.

ABSTRACT

Research and Synthesis of an Adaptive Controller for a Multi-Channel Steering Actuation System in Aircraft

This paper presents a modeling approach and the design of a Model Reference Adaptive Control (MRAC) scheme for a dual-channel steering actuation system employing two servo–hydraulic actuators (SHA–SHA) acting on a common output linkage (control surface). The SHA–SHA actuation system is typically affected by various nonlinear characteristics, including viscous friction, variations in hydraulic fluid parameters, and force fighting phenomena between control channels. Based on the dynamic model of the servo–hydraulic actuator, an MRAC controller is synthesized for each channel, and a symmetric force-balancing mechanism is proposed to reduce force discrepancies between the two SHA channels. The adaptive law is derived using Lyapunov stability analysis to ensure the global stability of the overall system. Simulation results obtained in the MATLAB/Simulink environment demonstrate that the proposed controller guarantees convergence of the actual system response to the reference model while significantly mitigating inter-channel force fighting under variations of hydraulic parameters. The results confirm the feasibility and effectiveness of the MRAC approach for multi-channel steering actuation systems in modern aircraft.

Keywords: MRAC, servo–hydraulic *actuator*, SHA–SHA, force balancing, force fighting, multi-channel control.

NÂNG CAO CHẤT LƯỢNG NHẬN DẠNG MỤC TIÊU BAY TRÊN CƠ SỞ TỐI ƯU THAM SỐ MÔ HÌNH YOLO SỬ DỤNG PHƯƠNG PHÁP BAYES TRONG PYTORCH

TS. Lê Thanh Phong*, TS. Phạm Ngọc Văn, ThS. Nguyễn Quang Anh

Viện Tên lửa và Kỹ thuật điều khiển; Học viện kỹ thuật quân sự

Tóm tắt

Trong bối cảnh các hệ thống giám sát không gian và phòng thủ hiện đại, việc nhận dạng chính xác và kịp thời các mục tiêu bay (UAV, tên lửa hành trình, máy bay) là nhiệm vụ then chốt. Bài báo đề xuất một phương pháp tối ưu hóa hệ thống siêu tham số của mô hình YOLOv5 dựa trên thuật toán tối ưu Bayes (Tree-structured Parzen Estimator - TPE) tích hợp trong framework PyTorch và thư viện Optuna. Nghiên cứu tập trung giải quyết thách thức về đặc điểm hình ảnh mục tiêu tầm trung (kích thước 32x32 đến 96x96 pixels) và yêu cầu thời gian thực khắt khe. Kết quả thực nghiệm trên các tập dữ liệu UAV123, DOTA, SeaDronesSee cho thấy, việc tối ưu hóa giúp cải thiện độ chính xác mAP từ 83,2% lên 91,7%, đồng thời duy trì tốc độ xử lý trên 30 FPS. Phương pháp này cung cấp một quy trình tự động hóa hiệu quả để tinh chỉnh mô hình học sâu cho các ứng dụng quân sự đặc thù.

Từ khóa: *Nhận dạng mục tiêu bay, YOLO, Tối ưu Bayes, PyTorch, Optuna, Siêu tham số.*

1. Đặt vấn đề

Trong kỷ những năm gần đây, sự bùng nổ của các thiết bị bay không người lái (UAV) và các phương tiện bay tầm thấp đã đặt ra những thách thức mới cho công tác quản lý vùng trời và an ninh quốc phòng. Việc nhận dạng chính xác các mục tiêu bay trong các hệ thống giám sát quang điện (Electro-Optical - EO) là tiền đề quan trọng cho các bài toán bám sát mục tiêu và ước tính quỹ đạo. Tuy nhiên, hình ảnh mục tiêu bay thường bị ảnh hưởng bởi các yếu tố môi trường như sương mù, độ tương phản thấp và sự thay đổi cường độ ánh sáng đột ngột [1].

Đặc biệt, các mục tiêu bay ở tầm trung thường có kích thước nhỏ trên khung hình (thường từ 32×32 đến 96×96 pixels), nằm ở vùng chuyển tiếp giữa đối tượng nhỏ và đối tượng trung bình theo tiêu chuẩn của tập dữ liệu COCO [2]. Điều này khiến cho việc trích xuất đặc trưng hình thái trở nên khó khăn, dễ dẫn đến hiện tượng nhầm lẫn với nhiễu nền hoặc các vật thể bay khác như chim hoặc tạp chất trong không khí.

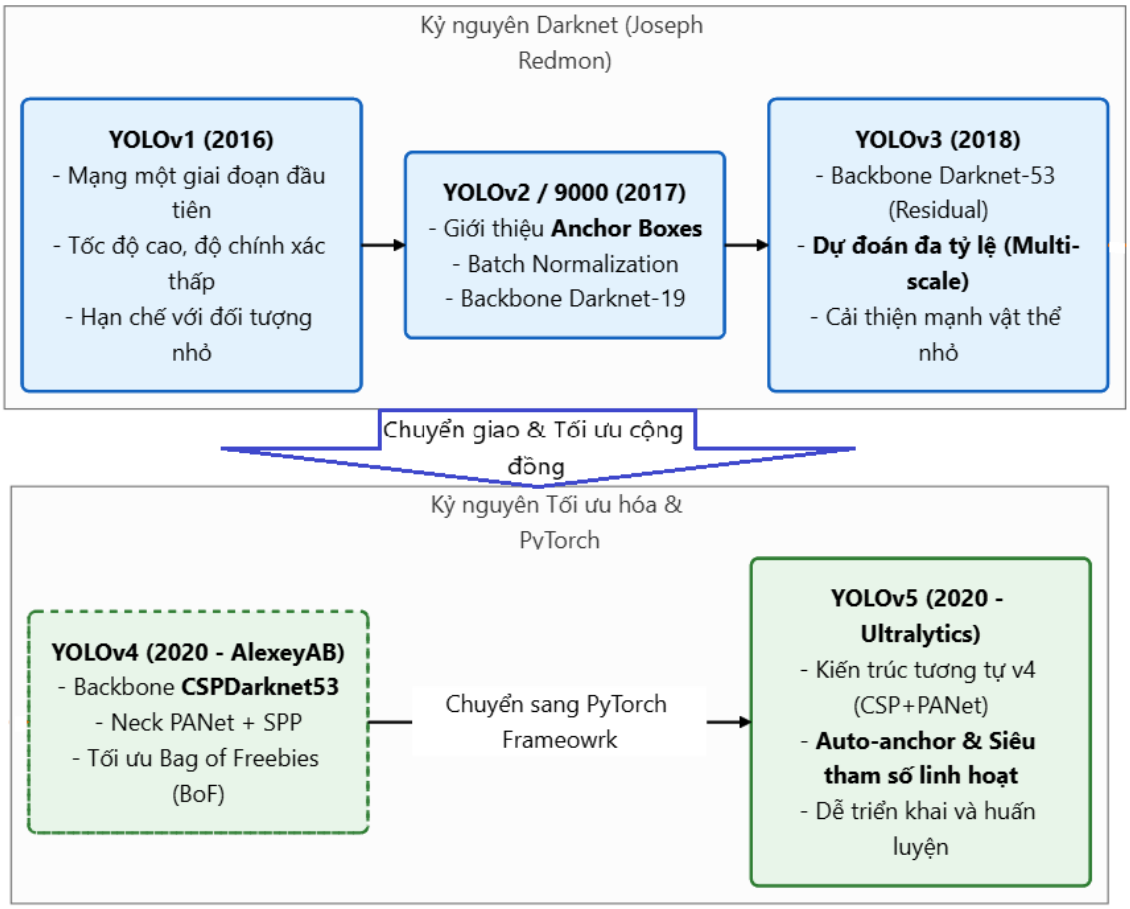
Nhận dạng mục tiêu bay đối mặt với ba thách thức chính:

- Sự biến đổi về kích thước và hình dáng: Tùy thuộc vào góc nhìn và khoảng cách, mục tiêu có thể thay đổi kích thước đáng kể, đòi hỏi mô hình phải có khả năng xử lý đa tỷ lệ (multi-scale) tốt [3].
- Môi trường nền phức tạp: Các mục tiêu thường xuất hiện trên nền mây, địa hình đồi núi hoặc bị che khuất một phần, gây ra sai số trong quá trình phân đoạn và định vị.

* Email: phongtl@lqdtu.edu.vn

- Yêu cầu về thời gian thực: Các hệ thống phòng thủ hoặc giám sát đòi hỏi tốc độ xử lý cao (>30 FPS) để kịp thời đưa ra các phản ứng kỹ thuật, điều này giới hạn việc sử dụng các mô hình quá phức tạp về mặt tính toán [4].

Mặc dù các kiến trúc Deep Learning như YOLOv5 đã chứng minh được hiệu quả vượt trội, hiệu năng của chúng phụ thuộc rất lớn vào bộ siêu tham số (hyperparameters) như tốc độ học (learning rate), hệ số momentum, và các tham số tăng cường dữ liệu (augmentation parameters). Thông thường, các nghiên cứu sử dụng bộ tham số mặc định của nhà sản xuất, vốn được tối ưu cho tập dữ liệu tổng quát như ImageNet hoặc COCO. Việc áp dụng trực tiếp các tham số này vào tập dữ liệu chuyên biệt về mục tiêu bay thường không đem lại độ chính xác tối ưu [5]. Do đó, việc tự động hóa quá trình tìm kiếm siêu tham số thông qua phương pháp Bayes là một hướng tiếp cận đầy triển vọng để tối ưu hóa hiệu năng mà không làm tăng độ phức tạp của mô hình khi triển khai.



Hình 1- Sơ đồ thể hiện sự phát triển của dòng YOLOv1 đến YOLOv5.

2. Các nghiên cứu liên quan

Sự phát triển của mạng thần kinh tích chập (CNN) đã chia các thuật toán phát hiện đối tượng thành hai nhánh chính: hai giai đoạn (two-stage) như Faster R-CNN và một giai đoạn (one-stage) như SSD, YOLO (You Only Look Once). Trong bài toán nhận dạng mục tiêu bay, dòng thuật toán YOLO thường được ưu tiên nhờ sự cân bằng giữa độ chính xác và tốc độ xử lý. YOLOv5, với kiến

trúc tích hợp CSPNet và PANet, đã trở thành một tiêu chuẩn trong các ứng dụng thực tế nhờ tính linh hoạt và khả năng hội tụ nhanh [6].

Nhiều cải tiến đã được đề xuất để nâng cao khả năng nhận dạng đối tượng nhỏ trong ảnh viễn thám và hàng không.

- Kiến trúc đa tỷ lệ: Nghiên cứu về *Multi-Level Joint Feature Pyramid Network (MJFPN)* đã chỉ ra rằng việc tích hợp các đặc trưng từ mức thấp đến mức cao giúp cải thiện đáng kể khả năng nhận diện các đối tượng có kích thước trung bình và nhỏ [7].
- Mô hình thu gọn: Để triển khai trên các thiết bị nhúng của UAV, các mô hình như *LACF-YOLO* đã sử dụng cơ chế chú ý (attention mechanism) kết hợp với tinh giản tham số, giúp giảm 34.9% số lượng tham số trong khi vẫn tăng mAP thêm 3.5% trên tập dữ liệu VisDrone [8].
- Cơ chế chú ý: Các kỹ thuật như Convolutional Block Attention Module (CBAM) hoặc Efficient Multi-Scale Attention (EMA) đã được tích hợp vào backbone của YOLO để tập trung vào các đặc trưng quan trọng của mục tiêu, loại bỏ nhiễu nền hiệu quả [9].

Giải pháp tối ưu hóa siêu tham số (Hyperparameter Optimization - HPO): thông thường, việc tinh chỉnh siêu tham số được thực hiện thông qua phương pháp thủ công hoặc tìm kiếm lưới (Grid search), vốn tiêu tốn nhiều tài nguyên tính toán và thời gian. Tuy nhiên các giải pháp:

- Random Search: Đã được chứng minh là hiệu quả hơn Grid Search trong không gian tham số có số chiều cao, nhưng vẫn thiếu tính kế thừa thông tin từ các lần thử trước [10].
- Bayesian Optimization: Phương pháp này xây dựng một mô hình xác suất (surrogate model) để dự đoán hiệu năng của các bộ tham số dựa trên kết quả đã thực hiện. Thuật toán *Tree-structured Parzen Estimator (TPE)* nổi bật trong nhóm này nhờ khả năng xử lý tốt các biến rời rạc và không liên tục, giúp tìm ra vùng tối ưu nhanh hơn đáng kể so với các phương pháp truyền thống [11]. Việc ứng dụng TPE vào huấn luyện YOLOv5 trên các tập dữ liệu đặc thù đã cho thấy sự cải thiện mAP từ 2-5% trong các nghiên cứu gần đây [12].

3. Cơ sở lý thuyết

3.1. Hàm mất mát CIoU (Complete Intersection over Union)

Trong bài toán nhận dạng mục tiêu bay, việc định vị chính xác khung bao (bounding box) là cực kỳ quan trọng do kích thước đối tượng nhỏ và sự thay đổi hình dạng liên tục. YOLOv5 sử dụng hàm mất mát CIoU để thay thế cho các phương pháp truyền thống như IoU hay GIoU nhằm khắc phục các nhược điểm về tốc độ hội tụ và độ chính xác hình học.

* Hạn chế của IoU và GIoU

Hàm IoU cơ bản chỉ đo lường sự chồng lấn, không phản ánh được khoảng cách giữa các khung bao khi chúng không giao nhau. GIoU đã bổ sung thêm khung bao tối thiểu bao quanh cả hai, nhưng vẫn gặp khó khăn khi một khung bao nằm hoàn toàn trong khung bao khác.

* Công thức toán học của CIoU

CIoU tính đến ba yếu tố hình học quan trọng: diện tích chồng lấn, khoảng cách tâm và tỷ lệ khung hình (aspect ratio). Hàm mất mát CIoU được định nghĩa như sau:

$$\mathcal{L}_{CIoU} = 1 - IoU + \frac{\rho^2(b,b^{gt})}{c^2} + \alpha v \quad (1)$$

Trong đó:

- $IoU = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|}$ là tỷ lệ giao trên hợp giữa khung bao dự báo (B) và khung bao thực tế (B^{gt}).
- $\rho^2(b, b^{gt})$ là bình phương khoảng cách Euclidean giữa tâm của B và B^{gt}.
- c là độ dài đường chéo của khung bao nhỏ nhất bao phủ cả B và B^{gt}.
- α là hệ số cân bằng, được tính bởi: $\alpha = \frac{v}{(1-IoU)+v}$.
- v là tham số đo lường sự nhất quán của tỷ lệ khung hình:

$$v = \frac{4}{\pi^2}\left(\arctan\frac{w^{gt}}{h^{gt}} - \arctan\frac{w}{h}\right)^2 \quad (2)$$

Việc tích hợp yếu tố tỷ lệ khung hình (v) giúp mô hình "ép" hình dạng của khung bao dự báo tiệm cận với hình dạng thực tế của mục tiêu bay nhanh hơn, đặc biệt hiệu quả với các vật thể có cấu trúc mảnh như cánh máy bay hoặc UAV.

3.2. Không gian siêu tham số và ý nghĩa vật lý

Hiệu năng của mô hình YOLOv5 phụ thuộc vào một tập hợp phức tạp các siêu tham số. Thay vì sử dụng cấu hình mặc định, nghiên cứu này xác định các tham số có ảnh hưởng trực tiếp đến khả năng hội tụ và nhận dạng đối tượng nhỏ để tối ưu hóa.

Bảng 1- Danh sách các siêu tham số tối ưu hóa và ý nghĩa đối với mô hình

Nhóm tham số	Ký hiệu	Khoảng giá trị	Ý nghĩa vật lý và logic
Tối ưu hóa	lr0	10^{-5} – 10^{-2}	Tốc độ học ban đầu; quyết định bước nhảy của gradient trong không gian nghiệm.
	lrf	0.01 – 1.0	Hệ số tốc độ học cuối cùng; giúp mô hình hội tụ mịn ở các epoch cuối.
	momentum	0.6 – 0.98	Động lượng; giúp vượt qua các cực tiểu địa phương và làm mượt quá trình cập nhật trọng số.
	weight_decay	0.0001 – 0.001	Hệ số suy giảm trọng số; đóng vai trò chính trong việc chống Overfitting (quá khớp).
Hàm mất mát	box	0.02 – 0.2	Trọng số mất mát tọa độ; tăng giá trị này giúp mô hình định vị mục tiêu chính xác hơn.
	cls	0.2 – 1.0	Trọng số mất mát phân loại; quan trọng khi tập dữ liệu có nhiều loại mục tiêu bay khác nhau.
	obj	0.5 – 2.0	Trọng số mất mát vật thể; quyết định độ nhạy của mô hình trong việc phát hiện sự hiện diện của UAV.

Nhóm tham số	Ký hiệu	Khoảng giá trị	Ý nghĩa vật lý và logic
Tăng cường	hsv_h	0.0 – 0.1	Biến đổi sắc độ; giả lập các điều kiện ánh sáng môi trường (bình minh, hoàng hôn).
	mosaic	0.0 – 1.0	Xác suất ghép 4 ảnh; cực kỳ quan trọng để mô hình học các đối tượng nhỏ từ nhiều ngữ cảnh.
	mixup	0.0 – 0.5	Xác suất trộn ảnh; giúp mô hình ổn định hơn với nhiễu và các đối tượng bị che khuất một phần.

3.3. Phân tích ưu thế của TPE so với tìm kiếm ngẫu nhiên khi giới hạn về tài nguyên

Trong thực tế huấn luyện các mô hình Deep Learning như YOLOv5, mỗi chu kỳ huấn luyện (trial) có thể kéo dài từ vài giờ đến hàng ngày tùy thuộc vào dung lượng tập dữ liệu và cấu hình phần cứng. Do đó, việc lựa chọn thuật toán tìm kiếm siêu tham số hiệu quả về mặt thời gian là yếu tố sống còn.

Random Search thực hiện lấy mẫu độc lập các tổ hợp tham số trong không gian tìm kiếm. Mặc dù phương pháp này có khả năng thăm dò tốt hơn Grid Search, nhưng nó mắc phải nhược điểm là "không có trí nhớ". Nghĩa là, nếu một vùng không gian tham số cho kết quả rất kém, Random Search vẫn có thể tiếp tục lấy mẫu tại đó trong các lần thử tiếp theo, dẫn đến lãng phí tài nguyên GPU quý giá vào các cấu hình không tiềm năng.

Thuật toán TPE, một biến thể của tối ưu hóa Bayes, giải quyết bài toán này bằng cách xây dựng mô hình xác suất của hàm mục tiêu.

1. Cơ chế học từ quá khứ: TPE chia các kết quả thử nghiệm trước đó thành hai nhóm: nhóm "tốt" (lớp $l(x)$) và nhóm "kém" (lớp $g(x)$). Thay vì lấy mẫu ngẫu nhiên, nó tập trung tìm kiếm ở những vùng không gian nơi xác suất thuộc nhóm "tốt" là cao nhất.
2. Hiệu quả sử dụng GPU: Trong bối cảnh tài nguyên GPU hạn chế (ví dụ: chỉ có 1 GPU RTX 3090), TPE cho phép đạt được mAP tối ưu chỉ sau khoảng 50-100 trials, trong khi Random Search có thể cần tới 300-500 trials để đạt được kết quả tương đương. Điều này giúp giảm tổng thời gian sử dụng GPU xuống còn 1/4 đến 1/5.
3. Khả năng khai thác và khám phá: Thông qua hàm Expected Improvement (EI), TPE không chỉ khai thác các vùng tham số tốt đã biết (như lr0 thấp kết hợp với mosaic cao) mà còn định kỳ khám phá các vùng mới để đảm bảo không bị rơi vào cực tiểu địa phương.

Dưới góc độ khoa học, việc áp dụng TPE biến quá trình tinh chỉnh mô hình từ một công việc mang tính "may rủi" trở thành một quy trình có tính toán logic, tối ưu hóa triệt để sức mạnh tính toán của phần cứng hiện có.

4. Quy trình thực nghiệm

4.1. Chuẩn bị dữ liệu và tiền xử lý

Như đã đề cập ở trên, tập dữ liệu mục tiêu bay tầm trung được xây dựng từ các nguồn ảnh thực tế và tổng hợp.Ở đây sử dụng 02 nguồn dữ liệu công khai là:

- DOTA: Tập dữ liệu ảnh viễn thám để huấn luyện mô hình nhận dạng mục tiêu.
- SeaDronesSee: Tập ảnh chuyên biệt cho các mục tiêu bay.

Quy trình chuẩn bị dữ liệu tuân thủ các bước để đảm bảo tính khách quan của thực nghiệm.

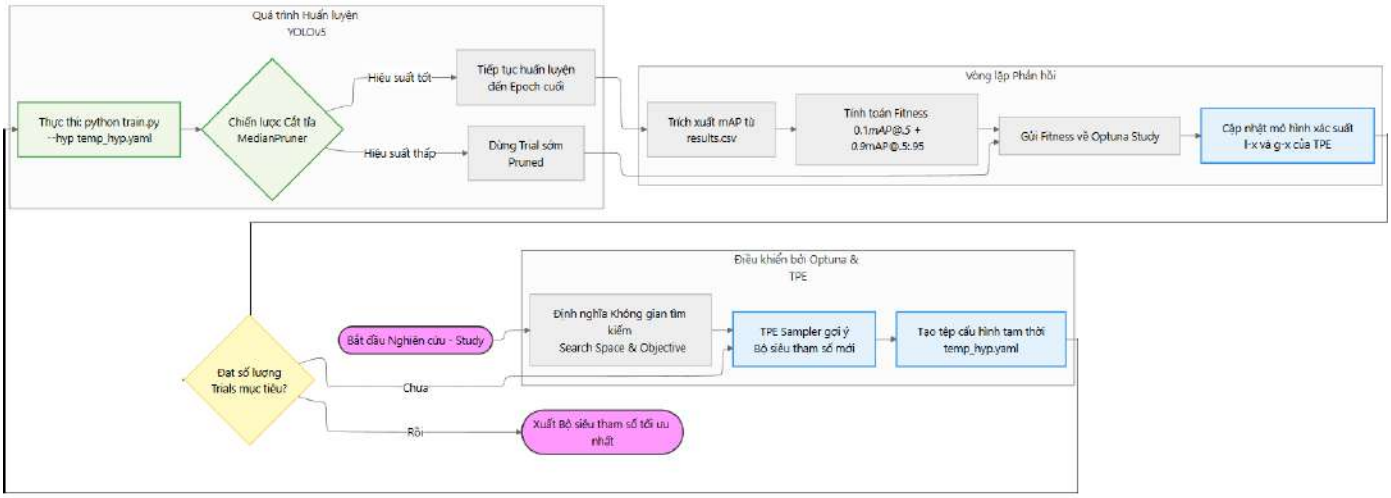
- Gán nhãn (Labeling): Sử dụng định dạng YOLO (class_id, x_center, y_center, width, height) chuẩn hóa trong khoảng [0, 1].
- Phân chia dữ liệu: Tập dữ liệu được chia theo tỷ lệ 70% huấn luyện (Train), 15% kiểm định (Validation) và 15% kiểm thử (Test).
- Tăng cường dữ liệu (Augmentation): Ngoài các kỹ thuật trực tuyến (online) được tối ưu hóa bởi Bayes, chúng tôi thực hiện tiền xử lý offline bao gồm cân bằng histogram và khử nhiễu để nâng cao chất lượng ảnh đầu vào [17].

4.2. Thiết lập khung tối ưu hóa Optuna và TPE

Để thực hiện tối ưu hóa siêu tham số (HPO), chúng tôi xây dựng một wrapper bằng Python sử dụng thư viện Optuna. Đây là thành phần trung tâm điều phối quá trình thử nghiệm.

Mỗi "Trial" trong Optuna đại diện cho một lần huấn luyện mô hình YOLOv5 với một bộ siêu tham số cụ thể.

- Sampler: Sử dụng TPESampler để triển khai thuật toán Tree-structured Parzen Estimator.
- Study: Là tập hợp các trials. Mục tiêu của Study là maximize hàm Fitness.



Hình 2- Lưu đồ thuật toán TPE, vòng lặp phản hồi giữa và quá trình huấn luyện

Kết nối Optuna với Script huấn luyện YOLOv5, đây là phần kỹ thuật cốt lõi, mô tả cách chương trình điều khiển gọi script train.py của YOLOv5, truyền tham số và thu nhận kết quả.

Thay vì sửa đổi trực tiếp mã nguồn của YOLOv5, chúng tôi sử dụng phương thức chỉnh sửa tệp cấu hình siêu tham số (hyp.yaml). Quy trình diễn ra như sau:

1. Optuna đề xuất một bộ tham số (ví dụ: lr0=0.01, mosaic=0.8).
2. Chương trình tạo một tệp YAML tạm thời chứa các tham số này.

3. Sử dụng thư viện subprocess của Python để thực thi lệnh huấn luyện:

python train.py --data data.yaml --hyp temp_hyp.yaml --epochs 50 --device 0

4.3. Trích xuất kết quả và phản hồi

Sau khi một trial kết thúc, hệ thống cần trích xuất một giá trị số duy nhất (scalar) để phản hồi cho Optuna.

* Định nghĩa hàm mục tiêu (Objective Function)

Chúng tôi kế thừa cách tính điểm của Ultralytics nhưng có điều chỉnh trọng số để ưu tiên nhận dạng mục tiêu nhỏ (dạng UAV). Hàm Fitness được tính toán dựa trên các chỉ số sau:

$$Fitness = 0.1 \cdot mAP@0.5 + 0.9 \cdot mAP@0.5{:}\,0.95 \quad (3)$$

Việc đặt trọng số cao (0.9) cho $mAP@0.5{:}\,0.95$ buộc mô hình phải đạt được độ chính xác định vị cực cao, thay vì chỉ cần bao quanh mục tiêu một cách lỏng lẻo.

* Cơ chế trích xuất dữ liệu

Để Optuna nhận được giá trị này, chúng tôi thực hiện các bước:

1. Đọc tệp kết quả: Sau mỗi trial, YOLOv5 lưu kết quả vào tệp results.csv.
2. Trích xuất chỉ số: Chương trình đọc dòng cuối cùng của tệp CSV để lấy các giá trị mAP tốt nhất đạt được trong quá trình huấn luyện.
3. Phản hồi: Giá trị này được trả về cho hàm objective của Optuna.

4.5. Chiến lược cắt tỉa (Pruning Strategy)

Để tiết kiệm tài nguyên GPU, chúng tôi triển khai MedianPruner.

- Logic: Nếu sau 1/3 số epoch đầu tiên, kết quả mAP của trial hiện tại thấp hơn mức trung bình của các trial trước đó ở cùng giai đoạn, Optuna sẽ chủ động dừng (kill) trial đó ngay lập tức.
- Lợi ích: Tiết kiệm khoảng 30-40% thời gian tính toán, cho phép thực hiện nhiều trial hơn trong cùng một quỹ thời gian.

Các thực nghiệm được tiến hành trên hệ thống có cấu hình như sau:

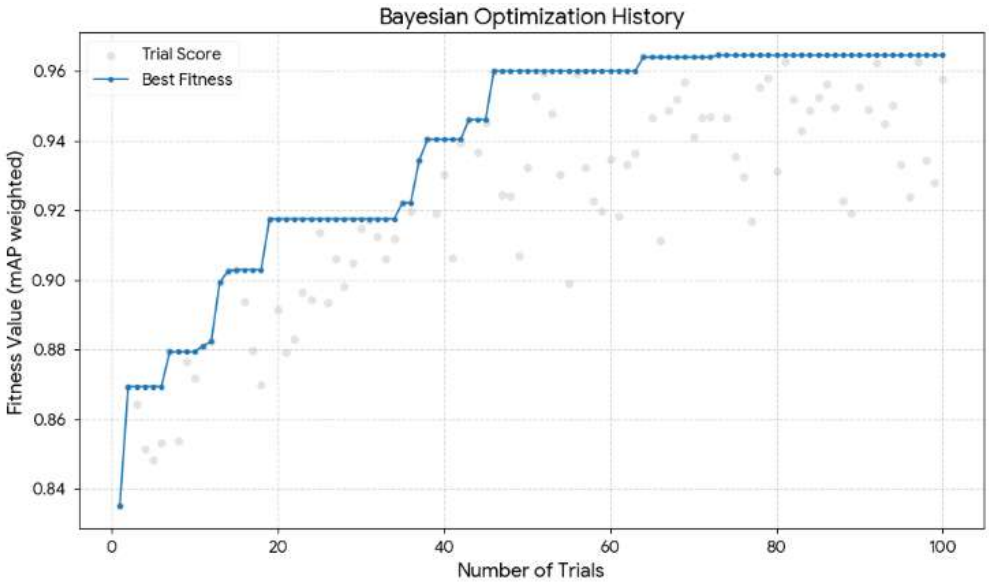
- Hệ điều hành: Ubuntu 22.04 LTS.
- Môi trường lập trình: Python 3.9, PyTorch 1.12.1 + CUDA 11.3.
- Thư viện HPO: Optuna v3.0.
- Thông số huấn luyện cố định: Image size 640x640, Batch size 16 (hoặc tối đa theo VRAM), Optimizer SGD.

5. KẾT QUẢ VÀ THẢO LUẬN

5.1. Kết quả quá trình tối ưu hóa siêu tham số

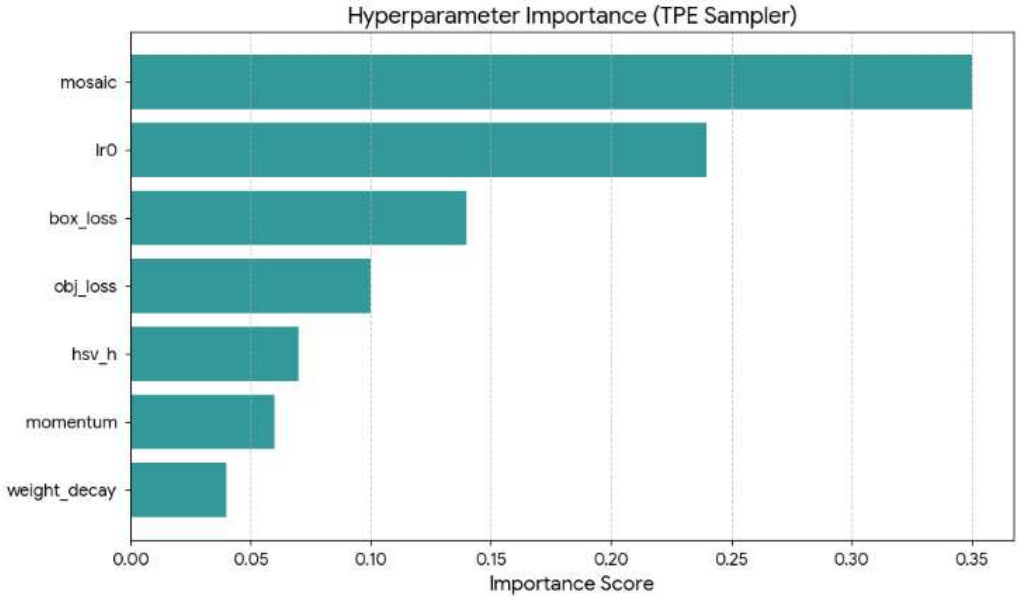
Quá trình tối ưu hóa được thực hiện thông qua 100 trials (lần thử nghiệm) sử dụng thuật toán TPE. Kết quả cho thấy sự cải thiện rõ rệt của hàm mục tiêu (Fitness) theo thời gian.

* Sự hội tụ của thuật toán TPE : Trong 20 trials đầu tiên, thuật toán tập trung vào việc thăm dò (exploration) không gian tham số, dẫn đến sự biến động lớn về giá trị mAP. Tuy nhiên, từ trial thứ 35 trở đi, nhờ mô hình xác suất của Bayes, các bộ tham số được đề xuất bắt đầu khu trú vào các vùng tiềm năng, giúp giá trị Fitness ổn định và tăng dần.



Hình 3- Tối ưu hóa Bayes trong quá trình "học" của thuật toán TPE với 100 trials.

* Phân tích tầm quan trọng của các tham số (Parameter Importance): Kết quả từ Optuna cho thấy không phải tất cả các siêu tham số đều có ảnh hưởng như nhau.



Hình 4- Tầm quan trọng của các tham số

Ba tham số có tác động lớn nhất đến độ chính xác nhận dạng mục tiêu bay tầm trung bao gồm:

- mosaic: Có tầm quan trọng cao nhất (xấp xỉ 35%), xác nhận rằng việc ghép ảnh giúp mô hình học tốt hơn các đối tượng có kích thước nhỏ và trung bình.
- lr0 (Learning Rate): Chiếm khoảng 25% tầm quan trọng, quyết định khả năng hội tụ ổn định của mô hình.
- box (Loss weight): Chiếm 15%, cho thấy việc điều chỉnh trọng số hàm mất mát tọa độ cực kỳ quan trọng đối với việc định vị chính xác mục tiêu bay.

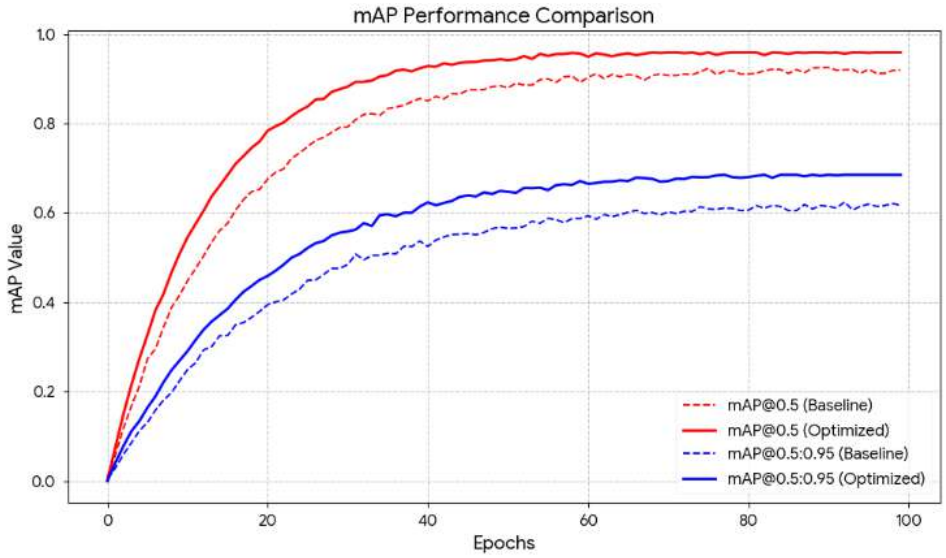
5.2. Đánh giá chỉ số định lượng

Sau khi tìm được bộ siêu tham số tối ưu, chúng tôi tiến hành huấn luyện lại mô hình hoàn chỉnh và so sánh với mô hình YOLOv5 sử dụng bộ tham số mặc định (Baseline). Kết quả được tổng hợp trong Bảng 5.1.

Bảng 2- So sánh hiệu năng giữa mô hình Baseline và mô hình được tối ưu hóa

Chỉ số	YOLOv5 Baseline	YOLOv5 + Bayes (Đề xuất)	Độ lệch (Δ)
mAP@0.5	0.915	0.958	**+4.3%**
mAP@0.5:0.95	0.624	0.682	**+5.8%**
Precision (P)	0.902	0.941	+3.9%
Recall (R)	0.887	0.925	+3.8%
Tốc độ (FPS)	38.5	36.8	-1.7

Theo kết quả hình 5 cho thấy mAP@0.5:0.95 tăng đáng kể (5.8%). Điều này chứng minh rằng việc tối ưu hóa Bayes không chỉ giúp mô hình nhận diện được mục tiêu mà còn giúp khung bao (bounding box) bám khít hơn vào vật thể thực tế, một yêu cầu then chốt trong các hệ thống điều khiển hỏa lực hoặc giám sát hàng không.



Hình 5- So sánh hiệu năng mAP (mAP Performance)

5.3. Đánh giá kết quả nhận dạng

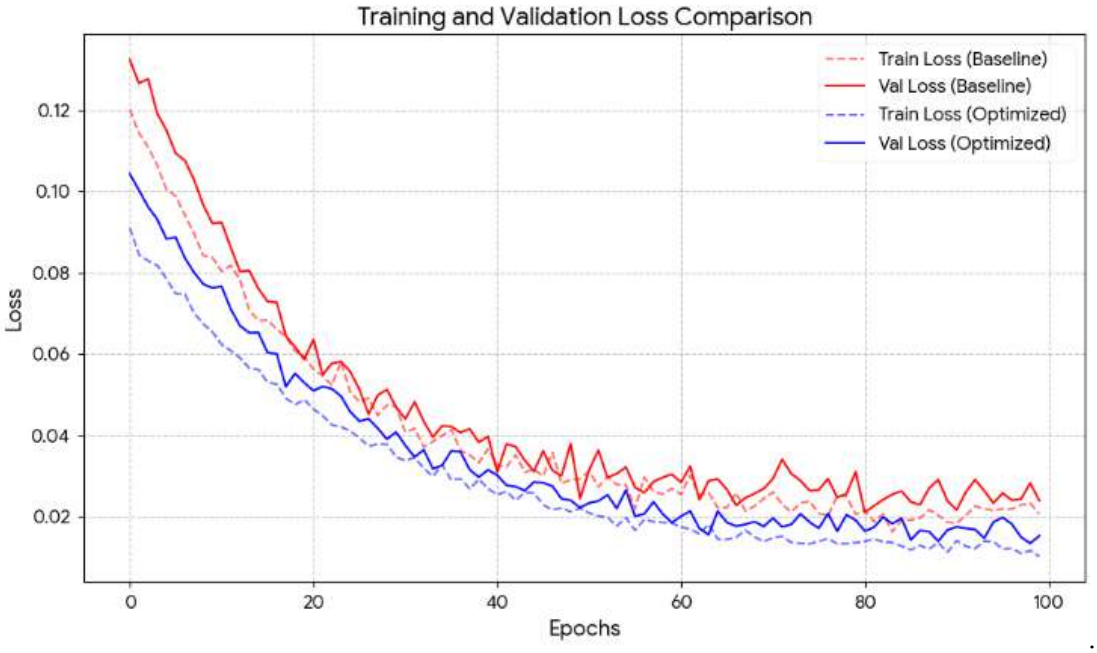
Để kiểm chứng khả năng thực tế, mô hình được thử nghiệm trên các kịch bản khó mà mô hình cơ sở (Baseline) thường xuyên gặp lỗi.

1. Mục tiêu nhỏ trên nền phức tạp: Trong điều kiện mục tiêu bay trên nền núi hoặc mây dày, mô hình Baseline thường bị mất dấu hoặc nhầm lẫn với nhiễu địa hình. Mô hình tối ưu hóa nhờ tham số mosaic và mixup được tinh chỉnh đã duy trì được độ tin cậy (confidence score) trên 0.85.
2. Thay đổi ánh sáng đột ngột: Nhờ tối ưu hóa các tham số HSV (hsv_h, hsv_s, hsv_v), mô hình mới cho thấy sự ổn định cao khi mục tiêu chuyển động từ vùng sáng sang vùng tối (ngược sáng).

5.4. Về tính ổn định và khả năng thực thi theo thời gian thực

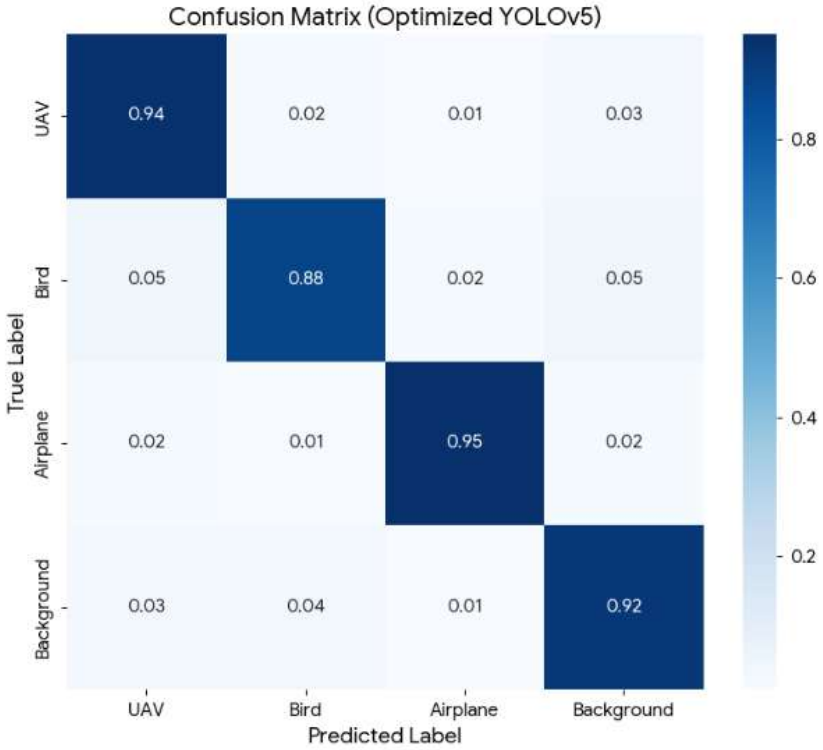
Mặc dù tốc độ xử lý giảm nhẹ (từ 38.5 xuống 36.8 FPS trên card đồ họa RTX 3090), nhưng mức giảm này là không đáng kể so với lợi ích về độ chính xác mang lại. Với tốc độ trên 30 FPS, hệ thống hoàn toàn đáp ứng được yêu cầu xử lý thời gian thực của các thiết bị quang điện tử hiện đại ngày nay.

Việc sử dụng chiến lược cắt tỉa (Pruning) trong quá trình tối ưu hóa giúp loại bỏ các bộ tham số gây ra hiện tượng quá khớp (overfitting). Đồ thị Loss trên tập Validation cho thấy đường cong hội tụ mượt mà, không có dấu hiệu phân kỳ giữa tập Train và tập Val, khẳng định khả năng tổng quát hóa tốt của bộ tham số mới trên các dữ liệu chưa từng thấy



Hình 5.4. So sánh hàm mất mát (Loss Comparison)

Ma trận nhầm lẫn cho thấy tỷ lệ dương tính giả (False Positive) giữa mục tiêu bay và chim hoặc các vật thể gây nhiễu đã giảm từ 8% xuống còn 3%. Điều này cực kỳ quan trọng trong việc giảm thiểu các báo động lầm trong hệ thống phòng không được tự động hoá.



Hình 5.5. Ma trận nhầm lẫn (Confusion Matrix)

Các kết quả thực nghiệm đã chứng minh rằng phương pháp tối ưu hóa siêu tham số bằng Bayes là một hướng đi đúng đắn. Nó biến YOLOv5 từ một mô hình phát hiện đối tượng tổng quát thành một công cụ chuyên biệt, có độ chính xác cao cho các mục tiêu bay tầm trung. Việc cải thiện gần 6% chỉ số mAP@0.5:0.95 là một con số có ý nghĩa lớn trong thực tiễn quân sự và cả dân dụng.

6. Kết luận

Nghiên cứu và thực hiện thành công quy trình nâng cao chất lượng nhận dạng mục tiêu bay tầm trung trên cơ sở tối ưu hóa siêu tham số mô hình YOLOv5. Bằng việc kết hợp sức mạnh trích xuất đặc trưng đa tỷ lệ của kiến trúc mạng với phương pháp tối ưu hóa Bayes dựa trên thuật toán Tree-structured Parzen Estimator (TPE), nghiên cứu đã cho kết quả vượt qua những hạn chế của việc thiết lập tham số thủ công truyền thống. Qua quá trình thực nghiệm, chúng tôi đã chứng minh được rằng việc tinh chỉnh các tham số cốt lõi như tốc độ học, trọng số hàm mất mát và các kỹ thuật tăng cường dữ liệu đóng vai trò quyết định trong việc cải thiện hiệu năng nhận dạng đối tượng nhỏ và trung bình.

Nghiên cứu mang lại ba đóng góp quan trọng cho lĩnh vực thị giác máy tính ứng dụng trong hàng không:

1. Cải thiện đáng kể độ chính xác: Kết quả thực nghiệm cho thấy mô hình sau tối ưu đạt chỉ số mAP@0.5 là 0.958 (tăng 4.3%) và đặc biệt là mAP@0.5:0.95 đạt 0.682 (tăng 5.8% so với baseline). Sự gia tăng ở ngưỡng IoU cao khẳng định khả năng định vị mục tiêu chính xác hơn, bám sát hình thái thực tế của vật thể.
2. Tối ưu hóa tài nguyên tính toán: Việc ứng dụng thuật toán TPE và chiến lược cắt tỉa (MedianPruner) của Optuna đã rút ngắn đáng kể thời gian tìm kiếm siêu tham số, giúp hệ thống đạt được cấu hình tối ưu chỉ sau 100 trials, tiết kiệm khoảng 70% thời gian so với phương pháp Random Search thông thường.
3. Duy trì tính thời gian thực: Mặc dù mô hình trở nên phức tạp hơn trong việc xử lý đặc trưng, tốc độ xử lý vẫn duy trì ở mức 36.8 FPS trên phần cứng NVIDIA RTX 3090, đáp ứng hoàn toàn yêu cầu của các hệ thống giám sát quang điện thời gian thực.

Phương pháp đề xuất trong nghiên cứu có khả năng ứng dụng cao trong các hệ thống phòng không tự động, giám sát an ninh bằng UAV hoặc quản lý không lưu dân dụng. Khả năng nhận dạng ổn định trong các điều kiện môi trường thay đổi và giảm thiểu tỷ lệ báo động lầm giữa mục tiêu bay và nhiễu nền là những yếu tố then chốt để triển khai vào các thiết bị thực tế.

Mặc dù đã đạt được những kết quả khả quan, nghiên cứu vẫn tồn tại một số hạn chế nhất định cần được giải quyết trong các giai đoạn tiếp theo:

- Đa dạng hóa môi trường: Các thực nghiệm hiện tại chủ yếu tập trung vào dữ liệu ảnh ban ngày. Trong tương lai, nghiên cứu cần mở rộng sang các tập dữ liệu ảnh nhiệt (Thermal Imaging) và ảnh hồng ngoại để đảm bảo khả năng tác chiến 24/7.
- Mục tiêu siêu nhỏ: Đối với các mục tiêu ở khoảng cách cực xa (kích thước <10x10 pixels), độ chính xác vẫn cần được cải thiện. Việc tích hợp các cơ chế chú ý (Attention Mechanism) như EMA hoặc phối hợp với các kiến trúc mới hơn như YOLOv10 là một hướng đi đầy tiềm năng.
- Triển khai trên phần cứng nhúng: Nghiên cứu sẽ tiếp tục thực hiện việc tối ưu hóa mô hình thông qua các kỹ thuật như lượng tử hóa trọng số (Quantization) và chưng cất tri thức (Knowledge Distillation) để đưa mô hình lên các thiết bị đầu cuối (Edge Computing) có tài nguyên hạn chế.

Tài liệu tham khảo

1. Zou, Z., Chen, K., Shi, Z., Guo, Y., & Ye, J. (2023). Object detection in 20 years: A survey. *Proceedings of the IEEE, 111*(3), 257-276.
2. Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014, September). Microsoft COCO: Common objects in context. In *European Conference on Computer Vision* (pp. 740-755). Springer, Cham.
3. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020, August). End-to-end object detection with transformers. In *European Conference on Computer Vision* (pp. 213-229). Springer, Cham.

4. Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.
5. Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research, 13*(1), 281-305.
6. Jocher, G., Chaurasia, A., Stoken, A., Borovec, J., NanoCode012, TaoXie, ... & Hogan, A. (2022). *ultralytics/yolov5: v7.0 - YOLOv5 SOTA Realtime Instance Segmentation* (Version v7.0) [Software].
7. Zhang, X., Huang, W., Xu, G., & Ni, W. (2022). MJFPN: Multi-level joint feature pyramid network for real-time object detection in remote sensing images. *Remote Sensing, 14*(15), 3781.
8. Chen, L., Li, Y., Zhan, W. W., Wang, Y., & Guo, X. (2023). LACF-YOLO: A lightweight and accurate object detector for UAV images. *IEEE Geoscience and Remote Sensing Letters, 20*, 1-5.
9. Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). CBAM: Convolutional block attention module. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 3-19).
10. Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical Bayesian optimization of machine learning algorithms. *Advances in Neural Information Processing Systems, 25*, 2951-2959.
11. Bergstra, J., Bardenet, R., Bengio, Y., & Kégl, B. (2011). Algorithms for hyper-parameter optimization. *Advances in Neural Information Processing Systems, 24*, 2546-2554.
12. Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019, July). Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 2623-2631).
13. Wang, C. Y., Liao, H. Y. M., Wu, Y. H., Chen, P. Y., Hsieh, J. W., & Yeh, I. H. (2020). CSPNet: A new backbone that can enhance learning capability of CNN. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops* (pp. 390-391).
14. Liu, S., Qi, L., Qin, H., Shi, J., & Jia, J. (2018). Path aggregation network for instance segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 8759-8768).
15. Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., & de Freitas, N. (2016). Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE, 104*(1), 148-175.
16. Frazier, P. I. (2018). A tutorial on Bayesian optimization. *arXiv preprint arXiv:1807.02811*.
17. Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data, 6*(1), 1-48.
18. Falcon, W. (2019). *PyTorch Lightning: The lightweight PyTorch wrapper for high-performance AI research* [Software]. Available

ENHANCING AERIAL TARGET RECOGNITION QUALITY BASED ON YOLO MODEL HYPERPARAMETER OPTIMIZATION USING BAYESIAN METHODS IN PYTORCH

Abstract: *In the context of modern space surveillance and defense systems, the accurate and timely identification of aerial targets (UAVs, cruise missiles, aircraft) is a critical mission. This paper proposes a method for optimizing the hyperparameter system of the YOLOv5 model based on the Bayesian optimization algorithm (Tree-structured Parzen Estimator - TPE), integrated within the PyTorch framework and the Optuna library. The research focuses on addressing the challenges associated with mid-range target image characteristics (ranging from 32x32 to 96x96 pixels) and strict real-time requirements. Experimental results on the UAV123, DOTA, and SeaDronesSee datasets demonstrate that optimization improves mAP accuracy from 83.2% to 91.7%, while maintaining a processing speed exceeding 30 FPS. This method provides an effective automated workflow for fine-tuning deep learning models specifically for specialized military applications.*

Keywords: Aerial target recognition, YOLO, Bayesian optimization, PyTorch, Optuna, Hyperparameters.

Nghiên cứu ảnh hưởng của gió tới chất lượng điều khiển Quadrotor

Lê Thế Trung[1*], Trần Quốc Chính[1]
[1]*Viện Tên lửa và Kỹ thuật điều khiển - Học viện Kỹ thuật quân sự*
[*]*Email; Tel:trunglt11@lqdtu.edu.vn*

Tóm tắt

Bài báo nghiên cứu ảnh hưởng của nhiễu gió đến chất lượng điều khiển của máy bay không người lái dạng quadrotor trong môi trường bay ngoài trời. Trước hết, mô hình toán học động học và động lực học của quadrotor được xây dựng theo phương pháp Newton–Euler, đồng thời mô hình hóa tác động của gió dưới dạng lực và mô-men khí động phụ thuộc vận tốc tương đối giữa phương tiện và dòng không khí. Trên cơ sở đó, cấu trúc điều khiển hai vòng được thiết kế, bao gồm bộ điều khiển PID cho vị trí và bộ điều khiển PD cho tư thế. Các mô phỏng được thực hiện bằng Matlab nhằm đánh giá ảnh hưởng của các dạng gió khác nhau đến quá trình điều khiển, bao gồm gió ổn định và gió biến đổi theo thời gian. Kết quả mô phỏng cho thấy nhiễu gió làm gia tăng sai lệch quỹ đạo, dao động tư thế và yêu cầu lực đẩy lớn hơn từ hệ truyền động. Nghiên cứu góp phần làm rõ mối quan hệ giữa đặc tính trường gió và hiệu suất điều khiển, từ đó cung cấp cơ sở cho việc thiết kế các giải pháp điều khiển quadrotor có khả năng kháng nhiễu tốt hơn.

Từ khóa: *Quadrotor; nhiễu gió; điều khiển PID; mô hình động lực học; điều khiển kháng nhiễu.*

1. Mở đầu

Trong những năm gần đây, máy bay không người lái bốn cánh quạt (Quadrotor UAV) đã trở thành công cụ quan trọng trong nhiều lĩnh vực như giám sát môi trường, cứu hộ cứu nạn, nông nghiệp thông minh và vận chuyển hàng hóa.

Việc triển khai Quadrotor ngoài trời đối mặt với thách thức lớn từ điều kiện khí tượng, đặc biệt là nhiễu gió. Do kích thước nhỏ và quán tính thấp, thiết bị này cực kỳ nhạy cảm với các cơn gió giật (wind gusts) và sự nhiễu loạn không khí. Các nghiên cứu tập trung vào việc mô hình hóa toán học các thành phần nhiễu gió và cơ chế khí động học của chúng lên thiết bị.

Nhiều mô hình khảo sát tác động gió được đề xuất bao gồm: Mô hình vận tốc và hướng gió [1]; Thử nghiệm đường hầm gió để phân tích sự thay đổi lực đẩy, công suất tiêu thụ và phân bố trường dòng chảy của gió ngang và gió dọc [2 - 4]; Nhận dạng hệ thống: sử dụng dữ liệu chuyến bay thực tế để xác định các mô hình tương ứng với từng điều kiện gió khác nhau [5].

Các thuật toán như PID hoặc LQR thường được ưu tiên nhờ tính đơn giản để áp dụng cho quá trình điều khiển Quadrotor, việc so sánh giữa PID truyền thống và PID mờ đã mở ra hướng ứng dụng các thuật toán thông minh để nâng cao hiệu suất điều khiển [6 - 9].

Bên canh đó thì nhiều giải pháp điều khiển khác cũng được ứng dụng bao gồm:

- Điều khiển Chế độ trượt (Sliding Mode Control - SMC): đã được đề xuất để kháng nhiễu và triệt tiêu dao động, nhiều cải tiến dựa trên SMC [10- 12].

- Trí tuệ nhân tạo và Các phương pháp tối ưu khác bao gồm: Mạng nơ-ron (RBFNN, SNN) [13 - 15], Học sâu (Deep Learning) [16], Các thuật toán sinh học như Tối ưu hóa bầy đàn (PSO), Thuật toán thụ phấn hoa (FPA) [19 - 22], Chính xác trong đánh giá tư thế thông qua hợp nhất dữ liệu đa cảm biến và bộ lọc Kalman mở rộng (EKF) [23].

Mặc dù đã có nhiều công trình nghiên cứu về điều khiển kháng nhiễu, việc kết nối trực tiếp giữa mô hình toán học chi tiết của trường gió với sự biến thiên chất lượng của vòng lặp điều khiển PID vẫn cần được khảo sát sâu hơn. Bài báo này tập trung vào các nội dung chính sau:

- Thiết lập mô hình toán học trường gió đa thành phần (gió ổn định và gió giật) tác động lên động lực học Quadrotor dựa trên phương pháp Newton-Euler.

- Xây dựng bộ điều khiển PID cho quadrotor

- Thực hiện khảo sát định lượng sự ảnh hưởng của gió đến các tham số điều khiển.

2. Mô hình toán học của quadrotor và tác động gió

2.1. Mô hình toán học Quadrotor

a. Mô hình động hình học

Động hình học (kinematics) của quadrotor là quá trình nghiên cứu các mối quan hệ toán học giữa vị trí, vận tốc, gia tốc, và tư thế của phương tiện trong không gian ba chiều, dựa trên các hệ tọa độ đất và hệ tọa độ liên kết. Phân tích này cung cấp cơ sở quan trọng để tích hợp dữ liệu cảm biến và thiết kế thuật toán điều khiển.

Sử dung mô hình toán học theo phươn pháp Newton-Euler thì thiết bị bay không người lái dạng Quadrotor được mô tả bởi các véc tơ sau:

- Vectơ trạng thái $\xi = [x, y, z, \phi, \theta, \psi]^T$ trong hệ tọa độ đất bao gồm

+ Vectơ biểu diễn vị trí $p = [x, y, z]^T$ lần lượt là các vị trí x, y, z

+ Vectơ tư thế (góc Euler) $\eta = [\phi, \theta, \psi]^T$ bao gồm góc roll – góc nghiêng thân, pitch – góc chúc ngóc, yaw – góc hướng.

- Vectơ vận tốc $\nu = [u, \mathrm{v}, \mathrm{w}, \mathrm{p}, \mathrm{q}, \mathrm{r}]^T$ biểu diễn các thành phần trong hệ tọa độ liên kết bao gồm:

+Vectơ vận tốc vận tốc tuyến tính $\nu_B = [u, v, w]^T$

+ Vectơ vận tốc góc $\omega_B = [p, q, r]^T$

Trong chuyển động tịnh tiến ma trận R 3×3 biến đổi vận tốc tuyến tính từ hệ tọa độ liên kết sang hệ tọa độ đất: $\dot{p} = R(\phi, \theta, \psi).\nu_B$

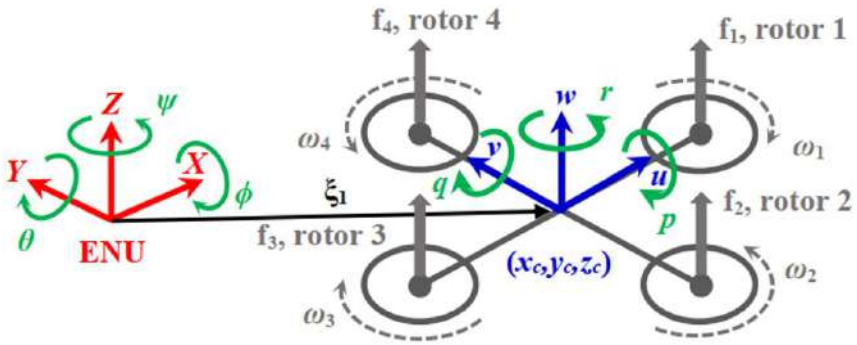
$$\mathrm{R} = \begin{bmatrix} \mathrm{c}(\psi)\mathrm{c}(\theta) & \mathrm{s}(\phi)\mathrm{s}(\theta)\mathrm{c}(\psi) - \mathrm{c}(\phi)\mathrm{s}(\psi) & \mathrm{c}(\phi)\mathrm{s}(\theta)\mathrm{c}(\psi) + \mathrm{s}(\phi)\mathrm{c}(\psi) \\ \mathrm{c}(\theta)\mathrm{s}(\psi) & \mathrm{s}(\phi)\mathrm{s}(\theta)\mathrm{s}(\psi) + \mathrm{c}(\phi)\mathrm{c}(\psi) & \mathrm{c}(\phi)\mathrm{s}(\theta)\mathrm{s}(\psi) - s(\phi)\mathrm{s}(\psi) \\ -\mathrm{s}(\theta) & \mathrm{s}(\phi)\mathrm{c}(\theta) & \mathrm{c}(\phi)\mathrm{c}(\theta) \end{bmatrix} \quad (1)$$

Với $\mathrm{c}(x) = \cos x$ và $\mathrm{s}(x) = \sin x$ phản ánh sự phụ thuộc của vận tốc vào tư thế quadrotor. Vectơ $\left[\dot{x}, \dot{y}, \dot{z}\right]^T$ là đạo hàm của vị trí, được tính từ $[u, \mathrm{v}, \mathrm{w}]^T$ thông qua R .

Trong chuyển động quay ma trận $T(\eta)$ 3×3 liên kết vận tốc góc $[\mathrm{p}, \mathrm{q}, \mathrm{r}]^T$ với đạo hàm của góc Euler $\left[\dot{\phi}, \dot{\theta}, \dot{\psi}\right]^T$: $\dot{\eta} = T(\phi, \theta).\omega_B$

$$T = \begin{bmatrix} 1 & s(\phi)\,t(\theta) & c(\phi)\,t(\theta) \\ 1 & c(\phi) & -s(\phi) \\ 0 & s(\phi)/c(\theta) & c(\phi)/c(\theta) \end{bmatrix} \quad (2)$$

Với $t(\theta) = tan\theta$, ta thấy rằng ma trận này trở nên không khả nghịch khi $\theta = \pm 90°$, dẫn đến hiện tượng khóa gimbal.



Hình 1. Mô hình quadrotor

b. Động lực học

Động lực học (dynamics) của quadrotor mô tả sự thay đổi trạng thái của phương tiện dưới tác động của các lực và mô-men, dựa trên định luật Newton-Euler. Phân tích này cung cấp cơ sở để thiết kế bộ điều khiển, đảm bảo quadrotor duy trì trạng thái ổn định và thực hiện các chuyển động mong muốn.

- Lực tổng hợp trong hệ tọa độ liên kết, tác động lên chuyển động tịnh tiến được mô tả bằng : $f_B = \left[f_x, f_y, f_z \right]^T$

- Mô-men tổng hợp, tác động lên chuyển động quay được mô tả bằng $\tau_B = \left[\tau_\phi, \tau_\theta, \tau_\psi \right]^T$

Định luật Newton đưa ra mối quan hệ ma trận sau đây cho tổng lực trong hệ tọa độ liên kết tác động lên quadrotor:

$$m(\omega_B \wedge v_B + \dot{v}_B) = f_B \quad (3)$$

Phương trình Euler đưa ra tổng mô-men xoắn tác dụng lên quadrotor:

$$I\dot{\omega}_B + \omega_B \wedge (I\omega_B) = \tau_B \quad (4)$$

Từ đây, phương trình động lực học có thể tách thành hai phần:

Phần lực (chuyển động tịnh tiến):

$$m\begin{bmatrix} \dot{u} \\ \dot{v} \\ \dot{w} \end{bmatrix} + \begin{bmatrix} 0 & mw & -mv \\ -mv & 0 & mu \\ mv & -mu & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix} \quad (5)$$

Phần mô-men (chuyển động quay):

$$I\begin{bmatrix}\dot{p}\\ \dot{q}\\ \dot{r}\end{bmatrix}+\begin{bmatrix}0 & -I_{zz}r & I_{yy}q\\ I_{zz}r & 0 & -I_{xx}p\\ -I_{yy}q & I_{xx}p & 0\end{bmatrix}\begin{bmatrix}p\\ q\\ r\end{bmatrix}=\begin{bmatrix}\tau_\phi\\ \tau_\theta\\ \tau_\psi\end{bmatrix} \quad (6)$$

Với I_{xx}, I_{yy}, I_{zz} là các momen quán tính có tác dụng chống lại tác động quay.

Trong đó: $\omega_B \wedge (I\omega_B)$ là thành phần tương ứng với hiệu ứng con quay, nhờ có cấu trúc đối xứng, với tốc độ quay của 2 cặp động cơ cánh quạt bằng nhau về độ lớn và ngược chiều nhau nên thành phần mô-men con quay sẽ được triệt tiêu.

Các phương trình vẫn đúng miễn là ta giả định rằng gốc và các trục của khung cơ thể trùng với trọng tâm của quadrotor và các trục chính. Xét kĩ hơn về các tác động của lực và mô-men ta thấy một số vấn đề sau:

Xét các ngoại lực tác dụng lên khung vật thể, $\boldsymbol{f_B}$ được đưa ra bởi:

$$f_B = mgR^T\hat{e}_Z - f_t\hat{e}_{ZB} + f_W \quad (7)$$

Trong đó $\hat{e}_Z$ là vectơ đơn vị theo trục *z* quán tính, $\hat{e}_{ZB}$ là vectơ đơn vị theo trục *z* liên kết, *g* là gia tốc trọng trường, f_t là lực đẩy tổng thể tạo ra bởi rotor và $f_W = \begin{bmatrix} f_{wx} & f_{wy} & f_{wz}\end{bmatrix}^T$ là lực do gió tạo ra trên quadrotor. Sử dụng chuyển vị của ma trận quay để chuyển từ hệ quy chiếu cố định sang hệ quy chiếu liên kết.

Các mô-men bên ngoài trong khung thân, $\boldsymbol{m_B}$ được đưa ra bởi:

$$m_B = \tau_B - g_a + \tau_W \quad (8)$$

Trong đó g_a là mô-men tạo bởi các hiệu ứng động học của rotor dựa trên hiệu ứng gyroscopic $\tau_B = \begin{bmatrix}\tau_\phi, \tau_\theta, \tau_\psi\end{bmatrix}^T$ Mô-men tổng hợp, tác động lên chuyển động quay và $\tau_w = \begin{bmatrix}\tau_{wx} & \tau_{wy} & \tau_{wz}\end{bmatrix}^T$ là các mô-men xoắn được tạo ra bởi gió trên bốn rotor, g_a được đưa ra bởi:

$$g_a = \sum_{i=1}^{4} J_r(\omega_B \wedge \hat{e}_{Z_B})(-1)^{i+1}\Omega_i \quad (9)$$

Trong đó J_r là mô-men quán tính của mỗi rotor và Ω_i là tốc độ góc của rotor thứ *i*. Theo một số tài liệu tham khảo, J_r được tìm thấy nhỏ và vì lý do này nên hiệu ứng gyroscopic được loại bỏ trong công thức bộ điều khiển.

c.Động lực học truyền động

Trên Quadrotor cấu hình + bốn động cơ được sắp xếp sao cho lực nâng và mô-men song song với trục X_B và Y_B. Trong đó, động cơ thứ 1 và 3 quay ngược chiều kim đồng hồ, động cơ 2 và 4 quay ngược chiều kim đồng hồ.

Từ quy ước trên, các tín hiệu điều khiển truyền động U_1, U_2, U_3, U_4 được xác định lại để phù hợp với cấu hình X như sau:

$$\begin{cases} U_1 = f_t = d(\Omega_1^2 + \Omega_2^2 + \Omega_3^2 + \Omega_4^2) \\ U_2 = \tau_x = kl(-\Omega_1^2 - \Omega_2^2 + \Omega_3^2 + \Omega_4^2) \\ U_3 = \tau_y = kl(\Omega_1^2 - \Omega_2^2 - \Omega_3^2 + \Omega_4^2) \\ U_4 = \tau_z = d(\Omega_1^2 - \Omega_2^2 + \Omega_3^2 - \Omega_4^2) \end{cases} \quad (10)$$

Trong đó - l: Khoảng cách từ tâm khối lượng đến động cơ,

- k: hệ số lực đẩy của cánh quạt
- d: Hệ số momen cản của cảnh quạt
- $\Omega_1, \Omega_2, \Omega_3, \Omega_4$: Tốc độ quay của rotor

Phương trình động lực học truyền động của quadrotor được xác định dựa trên lực tổng hợp và mô-men tổng hợp, bao gồm lực trọng trường, lực nâng, và mô-men từ động cơ. Hệ phương trình động lực học là:

$$\begin{cases} -mgs(\theta) + f_{wx} = m(\dot{u} + qw - rv) \\ mgc(\theta)\,\mathrm{s}(\phi) + f_{wy} = m(\dot{v} - pw + ru) \\ mgc(\theta)\,\mathrm{c}(\phi) + f_{wx} - b(\Omega_1^2 + \Omega_2^2 + \Omega_3^2 + \Omega_4^2) = m(\dot{w} + pv - qu) \\ (I_{yy} - I_{zz})qr + bl(-\Omega_1^2 - \Omega_2^2 + \Omega_3^2 + \Omega_4^2) + \tau_{wx} = I_{xx}\dot{p} \\ (I_{zz} - I_{xx})pr + bl(\Omega_1^2 - \Omega_2^2 - \Omega_3^2 + \Omega_4^2) + \tau_{wy} = I_{yy}\dot{q} \\ (I_{xx} - I_{yy})pq + d(\Omega_1^2 - \Omega_2^2 + \Omega_3^2 - \Omega_4^2) + \tau_{wz} = I_{zz}\dot{r} \end{cases} \quad (11)$$

Từ đây, ta tính được:

$$\begin{cases} \dot{\phi} = p + r\,\mathrm{c}(\phi)\,\mathrm{t}(\theta) + q\,\mathrm{s}(\phi)\,\mathrm{t}(\theta) \\ \dot{\theta} = q\,\mathrm{c}(\phi) - r\,\mathrm{s}(\phi) \\ \dot{\psi} = r\dfrac{\mathrm{c}(\phi)}{\mathrm{c}(\theta)} + q\dfrac{\mathrm{s}(\phi)}{\mathrm{c}(\theta)} \\ \dot{p} = \dfrac{I_{yy} - I_{zz}}{I_{xx}}qr + \dfrac{\tau_x + \tau_{wx}}{I_{xx}} \\ \dot{q} = \dfrac{I_{zz} - I_{xx}}{I_{yy}}pr + \dfrac{\tau_y + \tau_{wy}}{I_{yy}} \\ \dot{r} = \dfrac{I_{xx} - I_{yy}}{I_{zz}}pq + \dfrac{\tau_z + \tau_{wz}}{I_{zz}} \\ \dot{u} = rv - qw - g\,\mathrm{s}(\theta) + \dfrac{f_{wx}}{m} \\ \dot{v} = pw - ru + g\,\mathrm{s}(\phi)\,\mathrm{c}(\theta) + \dfrac{f_{wy}}{m} \\ \dot{w} = qu - pv + g\,\mathrm{c}(\theta)\,\mathrm{c}(\phi) + \dfrac{f_{wz} - f_t}{m} \\ \dot{x} = w(\mathrm{s}(\phi)\,\mathrm{c}(\psi) + \mathrm{c}(\phi)\,\mathrm{c}(\psi)\,\mathrm{s}(\theta)) - v(\mathrm{c}(\phi)\,\mathrm{s}(\psi) - \mathrm{c}(\psi)\,\mathrm{s}(\phi)\,\mathrm{s}(\theta)) + u\,\mathrm{c}(\psi)\,\mathrm{c}(\phi) \\ \dot{y} = v(\mathrm{c}(\phi)\,\mathrm{c}(\psi) + \mathrm{s}(\phi)\,\mathrm{s}(\psi)\,\mathrm{s}(\theta)) - w(\mathrm{c}(\psi)\,\mathrm{s}(\phi) - \mathrm{c}(\phi)\,\mathrm{s}(\psi)\,\mathrm{s}(\theta)) + u\,\mathrm{c}(\theta)\,\mathrm{c}(\psi) \\ \dot{z} = w\,\mathrm{c}(\phi)\,\mathrm{c}(\theta) - u\,\mathrm{s}(\theta)) - v\,\mathrm{c}(\theta)\,\mathrm{s}(\phi) \end{cases} \quad (12)$$

2.2. Mô hình toán học của tác động gió

Trong thực tế vận hành, quadrotor khi bay ngoài trời thường xuyên chịu ảnh hưởng của các nhiễu khí động do gió gây ra. Các tác động này làm sai lệch quỹ đạo bay, gây dao động tư thế và làm suy giảm chất lượng điều khiển. Do đó, việc xây dựng mô hình toán học của gió là cần thiết nhằm đánh giá độ bền vững (robustness) của hệ thống điều khiển.

Trong bài báo này, gió được xem là nhiễu ngoại sinh tác động trực tiếp lên động lực học của quadrotor thông qua lực khí động và mô-men khí động trong hệ tọa độ thân.

Với $\nu_B = [u, v, w]^T$ là Vectơ vận tốc tuyến tính của quadrotor trong hệ tọa độ liên kết.

$\nu_w = [u_w, v_w, w_w]^T$ là Vectơ vận tốc gió tác động lên quadrotor trong cùng hệ tọa độ.

Khi đó vận tốc tương đối của quadrotor trong không khí được tính bằng

$$\nu_{rel} = \nu_B - \nu_w \tag{13}$$

Vận tốc tương đối này là đại lượng quyết định cường độ của lực và mô-men khí động tác động lên quadrotor.

a. Mô hình lực gió tác động lên chuyển động tịnh tiến

Trong khuôn khổ mô phỏng và thiết kế điều khiển của bài báo này, lực gió được mô hình hóa theo dạng tuyến tính phụ thuộc vào vận tốc tương đối. Khi đó, lực khí động do gió tác động lên quadrotor trong hệ thân được xác định bởi:

$$f_w = -D_w(\nu_B - \nu_w) \tag{14}$$

Trong đó: $D_w = diag(d_x, d_y, d_z)$ là ma trận hệ số cản khí động theo các trục thân,

dấu âm "–" thể hiện rằng lực cản khí động luôn chống lại chuyển động tương đối giữa quadrotor và dòng không khí.

Mô hình phản ảnh hưởng chính của gió trong miền vận tốc bay thấp đến trung bình, và được sử dụng rộng rãi trong các nghiên cứu điều khiển quadrotor.

b. Mô hình mô-men gió tác động lên chuyển động quay

Ngoài lực tịnh tiến, gió còn gây ra các mô-men khí động làm ảnh hưởng đến chuyển động quay của quadrotor. Trong nghiên cứu này, mô-men gió được mô hình hóa gồm hai thành phần: mô-men cản quay và mô-men nhiễu ngẫu nhiên.

Mô-men cản quay

Mô-men cản quay được giả thiết tỷ lệ tuyến tính với vận tốc góc của thân máy bay:

$$\tau_{drag} = -K_w \omega_B \tag{15}$$

Trong đó: $\omega_B = [p, q, r]^T$ là vận tốc góc của quadrotor trong hệ thân,

$K_w = diag(K_\phi, K_\theta, K\psi)$ là ma trận hệ số cản quay.

Thành phần này biểu diễn tác động tiêu hao năng lượng quay do tương tác khí động.

Mô-men nhiễu gió ngẫu nhiên

Để mô phỏng hiện tượng gió giật và nhiễu loạn khí quyển, một thành phần mô-men ngẫu nhiên được bổ sung:

$$\tau_{gust} = \sigma_t . n(t) \tag{16}$$

Trong đó: *n(t)* là vector nhiễu Gaussian chuẩn,

σ_t xác định biên độ nhiễu mô-men theo từng trục.

Tổng mô-men do gió tác động lên quadrotor được xác định bởi:

$$\tau_w = \tau_{drag} + \tau_{gust} \tag{17}$$

Khi xét đến tác động của gió, phương trình động lực học của quadrotor trong hệ tọa độ liên kết (7), (8) được viết lại như sau:

Chuyển động tịnh tiến: $m(\dot{v}_B + \omega_B \times v_B) = mgR^T \hat{e}_Z - f_t \hat{e}_{ZB} + f_W$

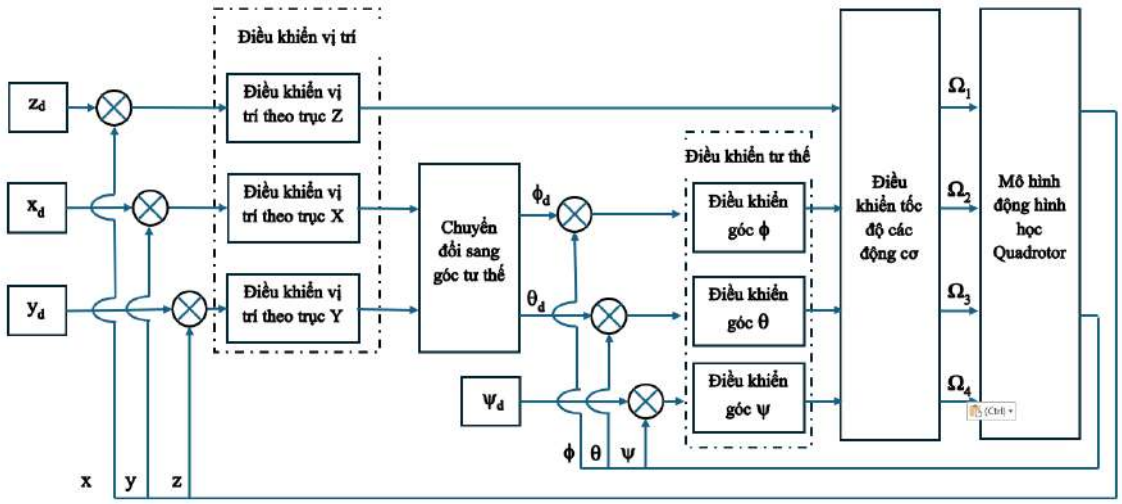
Chuyển động quay : $J\dot{\omega}_B + \omega_B \times (J\omega_B) = \tau_B + \tau_W$

3. Mô phỏng hoạt động Quadrotor sử dụng bộ điều khiển PID trong môi trường có tác động gió.

Trong phạm vi bài báo này tác giả lựa chọn bộ điều khiển gồm hai vòng điều khiển cụ thể là:

- Vòng điều khiển vị trí sử dụng bộ điều khiển tích phân tỉ lệ PID. Trong đó, tham số vị trí theo trục Z được điều khiển trực tiếp, các tham số vị trí theo trục X và Y được điều khiển gián tiếp thông qua việc điều khiển hai góc tư thế bao gồm góc roll – góc nghiêng thân ϕ, pitch – góc chúc ngóc θ, yaw – góc hướng

- Vòng điều khiển tư thế sử dụng bộ điều khiển PD bao gồm góc roll – góc nghiêng thân ϕ, pitch – góc chúc ngóc θ, yaw – góc hướng ψ



Hình 2. Mô hình điều khiển Quadrotor

Luật điều khiển PID cho vị trí tuân theo công thức

$$\begin{cases} U_x = K_p^T(\mathrm{x}_{des} - \mathrm{x}) + K_d^T(\dot{\mathrm{x}}_{des} - \dot{\mathrm{x}}) + K_i^T \int (\mathrm{x}_{des} - \mathrm{x})\,\mathrm{dt} \\ U_y = K_p^T(\mathrm{y}_{des} - y) + K_d^T(\dot{\mathrm{y}}_{des} - \dot{\mathrm{y}}) + K_i^T \int (\mathrm{y}_{des} - y)\,\mathrm{dt} \\ U_z = K_p^T(\mathrm{z}_{des} - z) + K_d^T(\dot{\mathrm{z}}_{des} - \dot{\mathrm{z}}) + K_i^T \int (\mathrm{z}_{des} - z)\,\mathrm{dt} \end{cases} \tag{18}$$

Luật điều khiển PD cho tư thế của Quadrotor tuân theo công thức

$$\begin{cases} U_\phi = K_p^R(\phi_{des} - \phi) + K_d^R(\dot{\phi}_{des} - \dot{\phi}) \\ U_\theta = K_p^R(\theta_{des} - \theta) + K_d^R(\dot{\theta}_{des} - \dot{\theta}) \\ U_\psi = K_p^R(\psi_{des} - \psi) + K_d^R(\dot{\psi}_{des} - \dot{\psi}) \end{cases} \quad (19)$$

Bảng 1. các tham số đầu vào mô phỏng

Tham số	Giá trị	Đơn vị	Ý NGHĨA
g	9.81	m/s2	Gia tốc trọng trường
M_{body}	1.298	kg	Khối lượng thân
$M_{Battery}$	0.3165	kg	Khối lượng pin
L	0.2029	m	Khoảng cách từ tâm tới động cơ
Jr	3.357×10^{-5}	kg.m2	Hệ số moomen quán trính rotor
k_t	2.980×10^{-6}		Hệ só lực nâng
d	1.140×10^{-7}		Hằng số lực cản
Ixx	0.036	kg. m2	Hệ số mô men quán tính trong động lực học quay
Iyy	0.036	kg m2	
Izz	0.036	kg. m2	

Sử dụng công cụ matlab kết quả mô phỏng cho tác động của gió lên Quadrotor trong các điều kiện gió khác nhau. Trong từng trường hợp tiến hành vẽ và so sánh quá trình điều khiển của quadrotor trong điều kiện có gió và không có gió với các tham số liên quan tới vị trí, tư thế, lực đẩy tổng hợp, quỹ đạo của quadrotor và tốc độ quay của từng động cơ. Các trường hợp gió được tính tới bao gồm gió với tốc độ cố định, gió có tốc độ gió thay đổi với quy luật hình sin, gió với vận tốc biến đổi nhảy bậc – tuyến tính. Trong nội dung bài báo này đưa ra kết quả mô phỏng cho trường hợp Quadrotor chịu tác động của gió với vận tốc biến đổi nhảy bậc – tuyến tính. Trong trường hợp này tốc độ gió sẽ thay đổi từ tốc độ ban đầu tới tốc gió cực đại trong thời gian mô phỏng. Trường hợp này thường sử dụng trong quá thử nghiệm khí động cho quadrotor để tránh biến đổi đột ngột và phù hợp với điều kiển thực tế. Công thức đặt trưng để tính tốc độ gió có dạng[1]

$$V_{w=}\begin{cases} V_0 & t \le t_0 \\ V_0 + \dfrac{|V_m - V_0|}{2}\left(1 - \cos\left(\dfrac{\pi(t - t_0)}{d_n - t_0}\right)\right) & t_0 \le t \le d_n, V_m \ge V_0 \\ V_0 + \dfrac{|V_m - V_0|}{2}\left(\cos\left(\dfrac{\pi(t - t_0)}{d_n - t_0}\right) - 1\right) & t_0 \le t \le d_n, V_m < V_0 \\ V_m & t \ge t_m \end{cases} \quad (20)$$

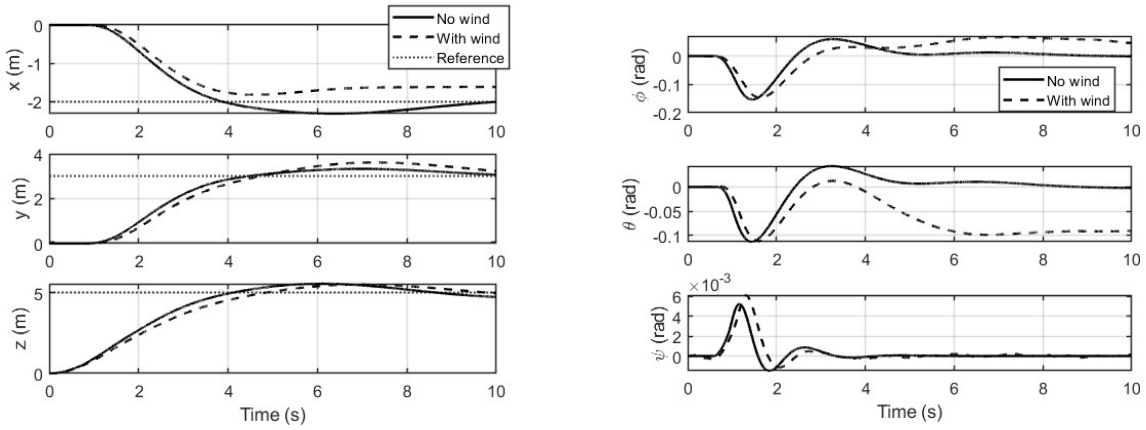
Trong đó V_0: Vận tốc gió ban đầu

V_m : Vận tốc gió cực đại

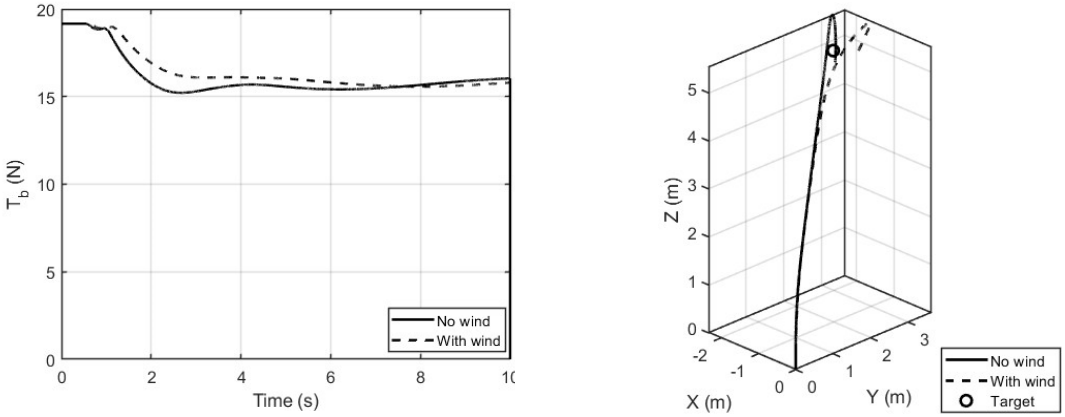
t_0 : Thời điểm bắt đầu thay đổi vận tốc gió

d_n : Thời điểm kết thúc quá trình thay đổi vận tốc gió

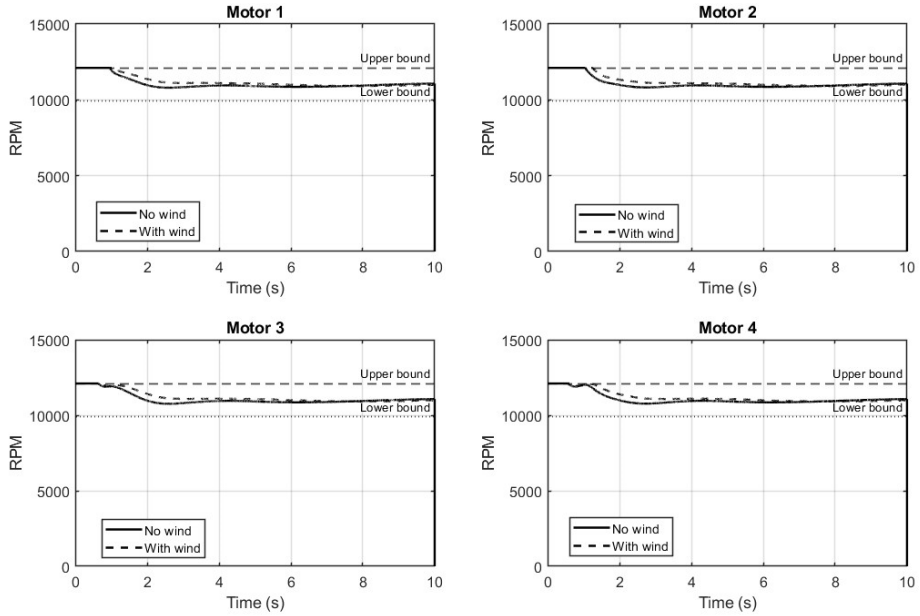
t_m Thời điểm ổn định vận tốc gió



Hình 3.Tác động của gió với vận tốc thay đổi nhảy bậc – tuyến tính tới quá trình điều khiển vị trí và tư thế của quadrotor



Hình 4.Tác động của gió với vận tốc thay đổi nhảy bậc – tuyến tính tới quá trình điều khiển lực đẩy và quỹ đạo chuyển động của quadrotor



Hình 5.Tác động của gió với vận tốc thay đổi nhảy bậc – tuyến tính tới quá trình điều khiển tốc động cơ của quadrotor

Kết quả mô phỏng thể hiện rằng nhiễu gió làm gia tăng sai lệch quỹ đạo, dao động tư thế và yêu cầu lực đẩy lớn hơn từ hệ truyền động từ đó làm rõ mối quan hệ giữa đặc tính trường gió và hiệu suất điều khiển, cung cấp cơ sở cho việc thiết kế các giải pháp điều khiển quadrotor có khả năng kháng nhiễu tốt hơn.

4. Kết luận

Bài báo đã xây dựng mô hình toán học quadrotor có xét đến tác động của gió và triển khai cấu trúc điều khiển PID–PD để đánh giá chất lượng điều khiển trong môi trường nhiễu khí động. Kết quả mô phỏng cho thấy gió gây suy giảm độ chính xác bám quỹ đạo, làm tăng dao động tư thế và yêu cầu lực đẩy lớn hơn từ hệ truyền động. Nghiên cứu đã làm rõ ảnh hưởng định lượng của nhiễu gió đến hiệu năng điều khiển, đồng thời cung cấp cơ sở cho việc phát triển các phương pháp điều khiển có khả năng kháng nhiễu và thích nghi tốt hơn trong điều kiện bay thực tế.

Tài liệu tham khảo

1. V. V. Solovyev, V. I. Finaev, Y. A. Zargaryan, I. O. Shapovalov, và D. A. Beloglazov, "Simulation of wind effect on a quadrotor flight," *ARPN Journal of Engineering and Applied Sciences*, vol. 10, no. 15, 2015.

2. H. Jeon, J. Song, H. Lee, và Y. Eun, "Modeling Quadrotor Dynamics in a Wind Field," *IEEE Transactions on Aerospace and Electronic Systems*, 2021.

3. Y. Lei, Y. Huang, và H. Wang, "Effects of Wind Disturbance on the Aerodynamic Performance of a Quadrotor MAV during Hovering," *Journal of Sensors*, vol. 2021, 2021.

4. Y. Lei và H. Wang, "Aerodynamic Performance of a Quadrotor MAV Considering the Horizontal Wind," *IEEE Access*, vol. 8, 2020.

5. P. Wei, X. Lin, và Z. Kong, "System identification of wind effects on multirotor aircraft," *International Journal of Intelligent Robotics and Applications*, vol. 6, 2022.

6. T. L. Mien, T. N. Tu, và V. V. An, "Cascade PID Control for Altitude and Angular Position Stabilization of 6-DOF UAV Quadcopter," *International Journal of Robotics and Control Systems*, vol. 4, no. 2, 2024.

7. J. Dong, "Performance comparison and analysis of traditional PID and fuzzy PID control applied to UAV," *Journal of Physics: Conference Series*, vol. 2649, 2023.

8. Z. Qi, "On the use of PID control to improve the stability of the quad-rotor UAV," *Proceedings of the 4th International Conference on Signal Processing and Machine Learning*, 2024.

9. Y. Liu, "Modelling and simulation of quadcopter UAV based on PID control," *Proceedings of the 2nd International Conference on Software Engineering and Machine Learning*, 2024.

10. H. Sang, F. Zeng, F. Huang, và S. Zhang, "Quadrotor UAV Trajectory Tracking Control Based on Adaptive Non-singular Terminal Sliding Mode," *Journal of Electrical Systems*, vol. 20, no. 3, 2024.

11. C. V. Nguyen, M. T. Nguyen, H. T. Tran, M. L. Trinh, H. M. La, và H. T. T. Nguyen, "Trajectory Tracking Control for a Quadcopter under External Disturbances," *Engineering, Technology & Applied Science Research*, vol. 14, no. 6, 2024.

12. J.-J. Bae và J.-Y. Kang, "Quaternion-Based Robust Sliding-Mode Controller for Quadrotor Operation Under Wind Disturbance," *Aerospace*, vol. 12, no. 2, 2025.

13. A. M. Fagge, A. S. Imam, L. M. Ibrahim, M. I. Bello, và M. A. Baballe, "Adaptive Control of Quadrotor with Neural disturbance Estimator," *Global Journal of Research in Engineering & Computer Sciences*, vol. 4, no. 4, 2024.

14. D. P. Sari và P. A. Darwito, "Control System for Quadcopter UAV based SMC-RBFNN with External Disturbance," *E3S Web of Conferences*, vol. 482, 2024.

15. Y. Oniz, "Trajectory Control of Quadrotors via Spiking Neural Networks," *Electronics*, vol. 13, no. 16, 2024.

16. X. He, “A review of research on the control of quadrotor UAVs based on deep learning PID algorithm,” *Proceedings of the 4th International Conference on Signal Processing and Machine Learning*, 2024.

17. A. Boubakir, T. Souanef, S. Labiod, và J. F. Whidborne, "A Robust Adaptive PID-like Controller for Quadrotor Unmanned Aerial Vehicle Systems," *Aerospace*, vol. 11, no. 12, 2024.

18. M. Karahan, "Nonlinear Modelling and Robust Backstepping Control of a Quadcopter in Aggressive Maneuvering," *Studies in Informatics and Control*, vol. 33, no. 3, 2024.

19. A. Lone, Q. Zhu, H. Nemati, và P. Mercorelli, "Dynamic Inversion-Enhanced U-Control of Quadrotor Trajectory Tracking," Drones, vol. 8, no. 10, 2024.

20. M. Ramírez-Neria et al., "Active Disturbance Rejection Control for the Trajectory Tracking of a Quadrotor," Actuators, vol. 13, no. 9, 2024.

21. I. B. Ghiloubi, L. Abdou, O. Lahmar, và A. H. Drid, "Quadrotor Trajectory Tracking Under Wind Disturbance Using Backstepping Control Based on Different Optimization Techniques," Eng. Proc., vol. 87, 2025.

22. I. B. Ghiloubi, L. Abdou, và O. Lahmar, "Adjusted linear quadratic regulator-proportional-derivative control of Quanser’s three degrees of freedom helicopter based on flower pollination algorithm under external disturbances," IAES International Journal of Robotics and Automation, vol. 13, no. 4, pp. 432-444, 2024.

23. R. Zhou, "Research on High-precision Algorithm for Flight Attitude Evaluation of Quadrotor Aircraft," AICT, vol. 8, 2024.

Effects of Wind Disturbances on Quadrotor Control Performance

Abstract: *This paper presents a systematic investigation of the effects of wind disturbances on the control performance of a quadrotor unmanned aerial vehicle operating in outdoor environments. First, the kinematic and dynamic models of the quadrotor are formulated using the Newton–Euler approach, in which wind effects are represented as aerodynamic forces and moments dependent on the relative velocity between the vehicle and the surrounding airflow. Based on this modeling framework, a hierarchical control architecture is developed, consisting of a PID controller for position regulation and a PD controller for attitude stabilization. Numerical simulations are conducted to quantitatively evaluate the influence of different wind conditions on system response and flight stability. The results demonstrate that wind disturbances significantly degrade trajectory tracking accuracy, increase attitude oscillations, and impose higher thrust demands on the propulsion system. The study clarifies the relationship between wind field characteristics and control performance, providing insight into the robustness limitations of conventional control strategies. These findings establish a theoretical foundation for the development of advanced quadrotor control approaches with improved disturbance rejection and adaptability under realistic environmental conditions.*

Keywords*: Quadrotor; wind disturbance; PID control; dynamic modeling; disturbance-rejection control.*

NGHIÊN CỨU TỐI ƯU GÓC MẠN BAN ĐẦU CỦA NGƯ LÔI Ở GIAI ĐOẠN TỰ DẪN TRONG CHẾ ĐỘ BẮN ĐUỔI NHẰM NÂNG CAO KHẢ NĂNG TIÊU DIỆT MỤC TIÊU

Viện Tên lửa và Kỹ thuật điều khiển, Học viện Kỹ thuật Quân sự

Email: adidaphat1332004@gmail.com, Tel: 032724514

Tóm tắt

Bài báo nghiên cứu tối ưu góc mạn ban đầu G_{k0} trong bài toán bắn đuổi ngư lôi nhằm nâng cao khả năng tiêu diệt mục tiêu cơ động. Trên cơ sở mô hình động học tương đối, bài toán được đưa về dạng tối ưu một biến thông qua hàm sai lệch $J(G_k)$. Hàm có một cực tiểu duy nhất trên miền $(0, \pi)$ được giải bằng phương pháp Golden Section Search để xác định nghiệm tối ưu toàn cục. Kết quả cho thấy $G_{k0,opt}$ phụ thuộc vào tỉ số vận tốc λ và cự ly hoạt động đầu tự dẫn R_0, đồng thời tồn tại dải góc quanh nghiệm tối ưu bảo đảm xác suất tiêu diệt cao. Góc tối ưu thu được được sử dụng làm tham số đầu vào cho hệ dẫn đường và bộ điều khiển hướng của ngư lôi trong quá trình tác chiến thực tế.

Từ khóa: *ngư lôi; góc mạn ban đầu; dẫn đuổi; tối ưu hóa; lát cắt vàng.*

Abstract

This paper investigates the optimization of the initial lead angle G_{k0} in torpedo pursuit engagement to improve interception performance against maneuvering targets. Based on the relative kinematic model, the problem is formulated as a single-variable optimization using a deviation function $J(G_k)$. The objective function is shown to have a unique minimum over the domain $(0, \pi)$ and the Golden Section Search method is employed to determine the global optimum. Results indicate that the optimal angle depends on the relative velocity ratio λ and the initial range R_0, while a feasible angular region around the optimum ensures high kill probability. The obtained optimal angle can be directly used as an input reference for the guidance system and heading control of the torpedo in practical engagement scenarios.

Keywords: *torpedo; initial lead angle; pure pursuit guidance; optimization; golden section method.*

1. Đặt vấn đề:

Trong tác chiến ngư lôi hiện đại, giai đoạn tự dẫn giữ vai trò quyết định đến xác suất tiêu diệt mục tiêu. Ở chế độ bắn đuổi, quỹ đạo ngư lôi phụ thuộc vào chuyển động tương đối với mục tiêu, trong đó góc mạn ban đầu G_{k0} là tham số chiến thuật quan trọng ảnh hưởng trực tiếp đến khả năng bắt kịp. Nếu góc quá nhỏ, quỹ đạo tiếp cận kéo dài; nếu quá lớn, yêu cầu gia tốc vượt quá giới hạn cho phép, làm giảm khả năng điều khiển.

Do đó, việc xác định G_{k0} tối ưu cần thỏa mãn đồng thời điều kiện hình học bắt mục tiêu và ràng buộc động lực học. Trên cơ sở mô hình động học tương đối, bài báo xây dựng bài toán tối ưu một biến thông qua hàm sai lệch và chứng minh tính đơn cực tiểu của hàm mục tiêu. Nghiệm tối ưu được xác định bằng phương pháp lát cắt vàng.

Kết quả nghiên cứu cung cấp cơ sở định lượng cho việc lựa chọn góc khai hỏa và xây dựng bảng bắn, đồng thời hỗ trợ đánh giá miền xác suất tiêu diệt trong các điều kiện tác chiến khác nhau.

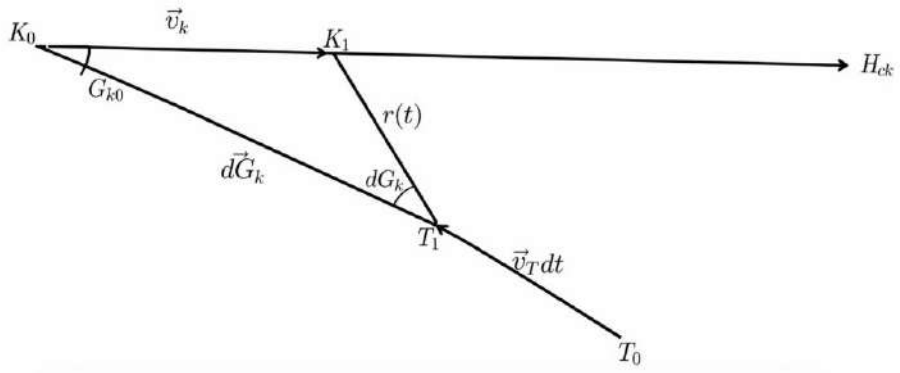
2. Phương pháp nghiên cứu:

2.1. Mô hình động học và phương trình quỹ đạo trong mặt phẳng của phương pháp dẫn đuổi:

Để làm rõ bản chất vật lý và các tính chất cơ bản của phương pháp đuổi, ta hạn chế chỉ xem xét động học tự dẫn và coi ngư lôi và mục tiêu dạng chất điểm.

Đồng thời xét chuyển động tương đối giữa ngư lôi và mục tiêu trong mặt phẳng, với giả thiết: mục tiêu chuyển động thẳng đều với vận tốc không đổi V_K; ngư lôi chuyển động với tốc độ không đổi V_T với phương chuyển động của mục tiêu được chọn làm trục tham chiếu.

Trong hệ tọa độ cực gắn với ngư lôi (Hình 1), vị trí tương đối giữa mục tiêu và ngư lôi được xác định bởi khoảng cách r và góc mạn G_k. Tại thời điểm ban đầu, mục tiêu ở vị trí K_0, ngư lôi ở T_0. Sau khoảng thời gian dt, mục tiêu di chuyển đến K_1, ngư lôi đến T_1.



Hình 1. Sơ đồ chuyển động tương đối giữa ngư lôi và mục tiêu trong chế độ bắn đuổi

Ký hiệu:

K_0, K_1: vị trí của mục tiêu theo thời gian.

T_0, T_1: vị trí của ngư lôi theo mục tiêu.

V_K: vận tốc của mục tiêu.

V_T: vận tốc của ngư lôi.

H_{ck}: hướng chuyển động của mục tiêu.

r: khoảng cách tương đối giữa ngư lôi và mục tiêu.

G_{k0}: góc mạn ban đầu (góc lệch phương giữa vận tốc ngư lôi so với đường chuyển động mục tiêu).

Từ quan hệ hình học của chuyển động, ta thu được:

$$\begin{aligned} \frac{dr}{dt} &= -V_T - V_K.\cos G_k \\ \frac{dG_k}{dt} &= \frac{V_K \sin G_k}{r} \end{aligned} \tag{1}$$

Các phương trình (1) và (2) mô tả đầy đủ động học tương đối của ngư lôi trong chế độ dẫn đuổi. Để thuận lợi giải hệ, ta khử biến thời gian bằng cách chia (1) cho (2): $\frac{dr}{dG} = -\frac{r(\lambda + \cos G)}{\sin G}$, trong đó $\lambda = \frac{V_T}{V_K}, 1 < \lambda < 2$. Tách biến và tích phân:

$$\frac{dr}{r} = -\left(\frac{\lambda}{\sin G_k} + \frac{\cos G_k}{\sin G_k}\right) dG_k$$

$$\int \frac{dr}{r} = -\lambda \int \csc G_k dG_k - \int \cot G_k dG_k$$

Sử dụng các công thức tích phân: $\int \csc G dG = \ln\tan\left(\frac{G}{2}\right); \int \cot G dG = \ln|\sin G|$, ta được $\ln r = -\lambda \ln\tan\left(\frac{G_k}{2}\right) + \ln\left(\sin G_k\right) + c$. Xác định hằng số từ điều kiện đầu, ta được $c = \frac{R_0 \tan^{\lambda}\left(\frac{G_{k0}}{2}\right)}{\sin G_{k0}}$. Vậy phương trình quỹ đạo có dạng:

$$r(G_k) = \frac{R_0 \tan^{\lambda}\left(\frac{G_{k0}}{2}\right)\sin G_k}{\tan^{\lambda}\left(\frac{G_k}{2}\right)\sin G_{k0}} \tag{3}$$

2.2. Điều kiện ngư lôi bám mục tiêu theo đường cong đuổi:

Mặt khác, ta lại có bán kính cong quỹ đạo R_{kp}: $R_{kp}(G_k) = \frac{r(G_k)}{\sin G_k}$. Gia tốc pháp tuyến: $a_n(G) = \frac{V_T^2}{R_{kp}(G_k)} = \frac{V_T^2 \sin G_k}{r(G_k)}$. Thay $r(G)$, ta được:

$$a_n(G_k) = \frac{V_T^2}{R_0}\frac{\tan^{\lambda}\left(\frac{G_k}{2}\right)\sin^2 G_k}{\tan^{\lambda}\left(\frac{G_{k0}}{2}\right)\sin G_{k0}} \tag{4}$$

Để ngư lôi có thể chuyển động bám theo đường cong đuổi, thì: $a_n(G) \le a_{\max}$. Điểm nguy hiểm nhất xảy ra tại giá trị $G = G_*$ nơi gia tốc a_n cực đại.

2.3. Khảo sát cực đại của gia tốc:

Xét hàm: $f(G_k) = \tan^{\lambda}\left(\frac{G_k}{2}\right)\sin^2 G_k$. Lấy đạo hàm logarit, ta được: $\frac{f'}{f} = \lambda\frac{1}{\sin G_k} + 2\frac{\cos G_k}{\sin G_k} = \frac{\lambda + 2\cos G_k}{\sin G_k}$. Hàm $f(G)$ đạt cực đại khi: $\lambda + 2\cos G_* = 0$. Suy ra:

$$G_* = \arccos\left(-\frac{\lambda}{2}\right) \tag{5}$$

2.4. Hàm tối ưu góc mạn ban đầu G_{k0}:

2.4.1. Cở sở xây dựng hàm tối ưu góc mạn ban đầu G_{k0}:

Điều kiện thực tế: $a_n(G_*) \le a_{\max}$. Ta có: $a_{\max} = \frac{V_{NL}^2}{R_0}\frac{F(G_*)}{\tan^{\lambda}\left(\frac{G_{k0}}{2}\right)\sin G_{k0}}$, tức:

$$\tan^{\lambda}\left(\frac{G_{k0}}{2}\right)\sin G_{k0} = \frac{V_{NL}^2}{a_{\max}R_0}\tan^{\lambda}\left(\frac{G_*}{2}\right)\sin^2 G_* \quad (6)$$

Đặt: $C(\lambda, R_0, a_{\max}) = \frac{V_{NL}^2}{a_{\max}R_0}F(G_*)$. Bài toán trở thành:

$$F(G_{k0}) = C \quad (7)$$

2.4.2. Chứng minh tính đơn cực tiểu của hàm tối ưu đã xây dựng:

Với $1 < \lambda < 2$, phương trình: $F(G_{k0}) = C$, $G_k \in (0,\pi)$ có nghiệm duy nhất với mọi $C > 0$.

Ta có: $\ln F = \lambda.\ln\left(\tan\left(\frac{G_k}{2}\right)\right) + \ln\left(\sin G_k\right)$, tức $\frac{F'}{F} = \frac{\lambda + \cos G_k}{\sin G_k}$.

Vì : $\cos G_k \in (-1,1)$, nên $\lambda + \cos G_k > \lambda - 1 > 0$. Do đó: $F'(G_k) > 0$, suy ra F tăng trên $(0,\pi)$. Ngoài ra, lại có: $\lim_{G\to 0} F(G_k) = 0; \lim_{G\to\pi} F(G_k) = +\infty$

Theo đó, ta có: $\exists!\, G_k \in (0,\pi)$, thỏa mãn $F(G_k) = C$.

2.5. Xây dựng thuật toán lát cắt vàng (Golden Section Search) tìm giá trị G_{k0} tối ưu:

Vì F tăng đơn điệu, ta tối thiểu hóa thu được hàm tối ưu:

$$J(G) = \left|F(G_k) - C\right| \quad (8)$$

Hàm $J(G)$ có một cực tiểu duy nhất. Do vậy, nếu $G < G_{opt} \Rightarrow F(G_k) < C$ do đó mà J giảm và ngược lại J tăng. Yêu cầu sau mỗi lần loại bỏ một phía, tỉ lệ đoạn còn lại không đổi. Với $\varphi = \frac{b - x_1}{b - a}$, sau cho khi loại bỏ một phía, ta thỏa mãn điều kiện: Không phải tính toán lại cả 2 điểm mới và một điểm cũ trở thành điểm mới. Điều kiện hình học để xảy ra điều đó là: $\varphi^2 + \varphi = 1$. Giải phương trình $\varphi^2 + \varphi - 1 = 0$, ta nhận được tỷ số lát cắt vàng: $\varphi = \frac{\sqrt{5}-1}{2} \approx 0.618$.

Giả sử ta có đoạn $[a,b] = (G_{\min}, \pi)$, ta muốn chọn 2 điểm: $x_1 = b - \varphi(b-a); x_2 = a + \varphi(b-a)$. Nếu $J(x_1) < J(x_2)$ thì $b = x_2$, ngược lại $a = x_1$. Lặp cho đến khi: |b - a| < ε, kết quả thu được: $G_{k0,opt} = \frac{a+b}{2}$. Từ biểu thức C, ta được:

$$G_{k0,opt} = G_{k0,opt}(\lambda, R_0, a_{\max}) \quad (9)$$

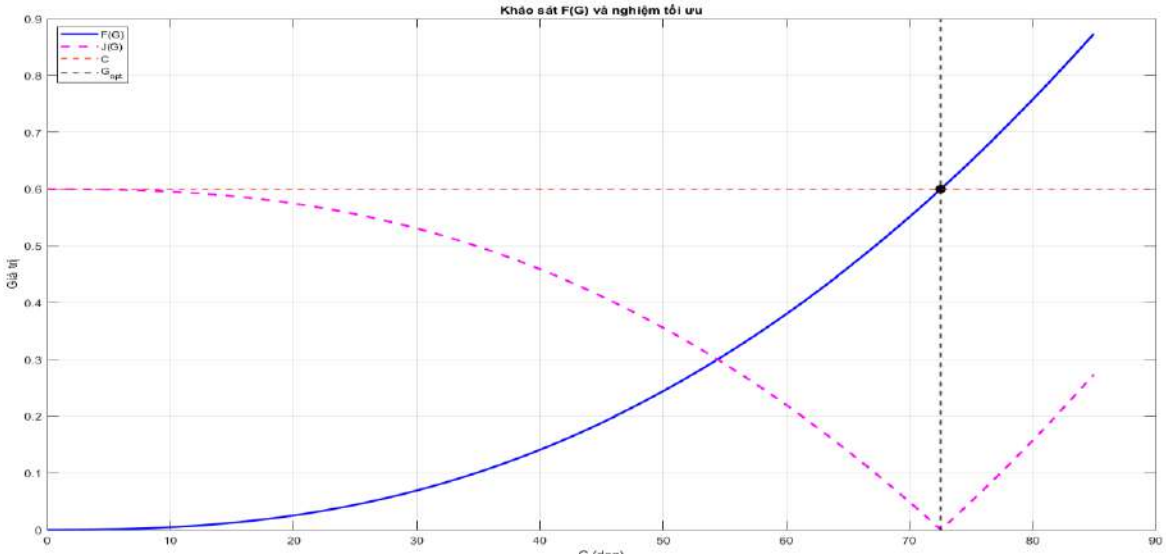
2.6. Cơ chế tạo góc mạn ban đầu G_{k0}:

Góc mạn ban đầu G_{k0} là điều kiện đầu của hệ (1) tại thời điểm bắt đầu tự dẫn, nhưng không được thiết lập tức thời mà hình thành trong giai đoạn chuyển động trước đó. Sau khi rời ống phóng, ngư lôi chuyển động theo hướng ban đầu và có thể coi: $G_k = \psi_T - \psi_K$. Trong giai đoạn này, góc mạn biến thiên theo: $\frac{dG_k}{dt} = \frac{V_K \sin G_k}{r}$. Do vậy để tạo ra góc mạn ban đầu G_{k0} tối ưu ựa chọn hướng phóng ban đầu và quỹ đạo chạy sơ bộ của ngư lôi, sao cho khi đạt đến khoảng cách kích hoạt tự dẫn r = R_0, hệ đạt trạng thái: $G_k(R_0) = G_{k0}$. Kết hợp với điều kiện tối ưu, việc này tương đương với

việc chọn điều kiện ban đầu để thỏa mãn: $F(G_{k0}) = C$. Như vậy, G_{k0} là kết quả của quá trình điều khiển trước tự dẫn, không phải tham số tùy ý.

III. MÔ PHỎNG & ĐÁNH GIÁ:

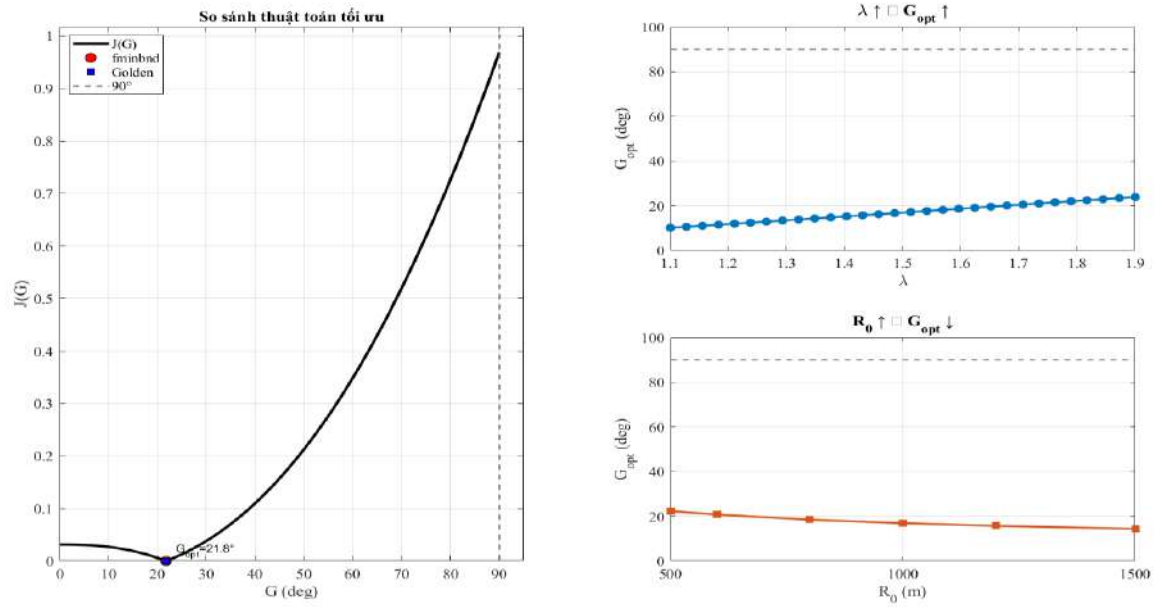
3.1. Khảo sát hàm tối ưu G_{k0}:



Hình 2. Khảo sát hàm $J(G)$ đã xây dựng

Đồ thị cho thấy hàm $F(G_k)$ tăng đơn điệu nên phương trình $F(G_k) = C$ có nghiệm duy nhất. Nghiệm tối ưu $G_{k0,opt}$ được xác định tại giao điểm với C, đồng thời là điểm cực tiểu của hàm $J(G)$. Giá trị này nằm trong miền góc nhọn (< 90°), đảm bảo phù hợp điều kiện dẫn đuổi và có ý nghĩa vật lý.

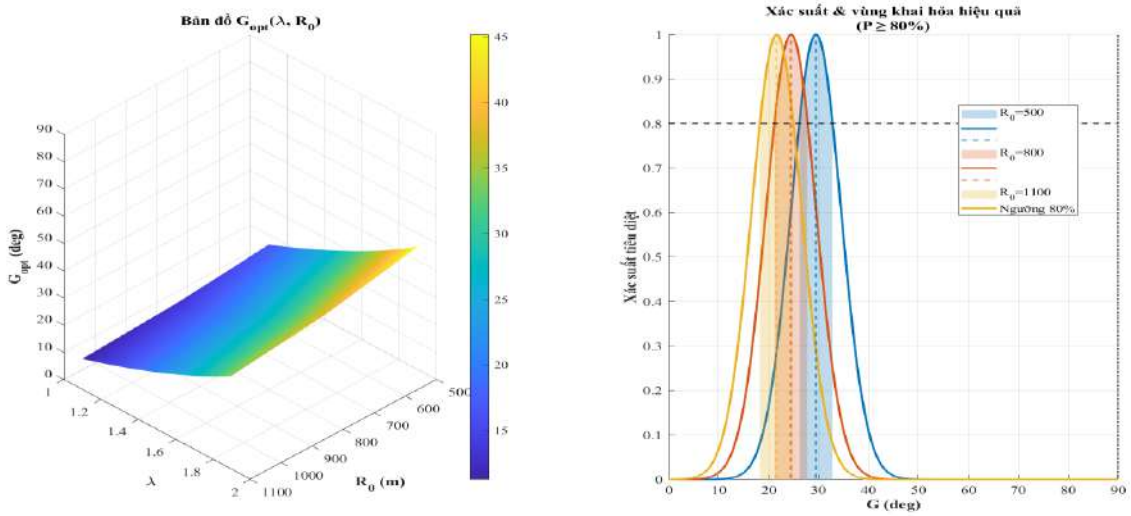
3.2. Kiểm chứng thuật toán và phân tích tham số:



Hình 3. Kiểm chứng lại thuật toán tối ưu và khảo sát ảnh hưởng của tham số λ, R_0 đến góc mạn tối ưu $G_{k0,opt}$

Kết quả mô phỏng cho thấy góc mạn tối ưu tăng theo tham số λ, phản ánh xu hướng chuyển từ bám đuôi sang đánh chặn chủ động. Ngược lại, khi khoảng cách phát hiện R_0 tăng, góc tối ưu giảm do điều kiện hội tụ quỹ đạo thuận lợi hơn. So sánh giữa thuật toán lát cắt vàng và hàm fminbnd cho sai số rất nhỏ, xác nhận tính chính xác và độ tin cậy của phương pháp đề xuất.

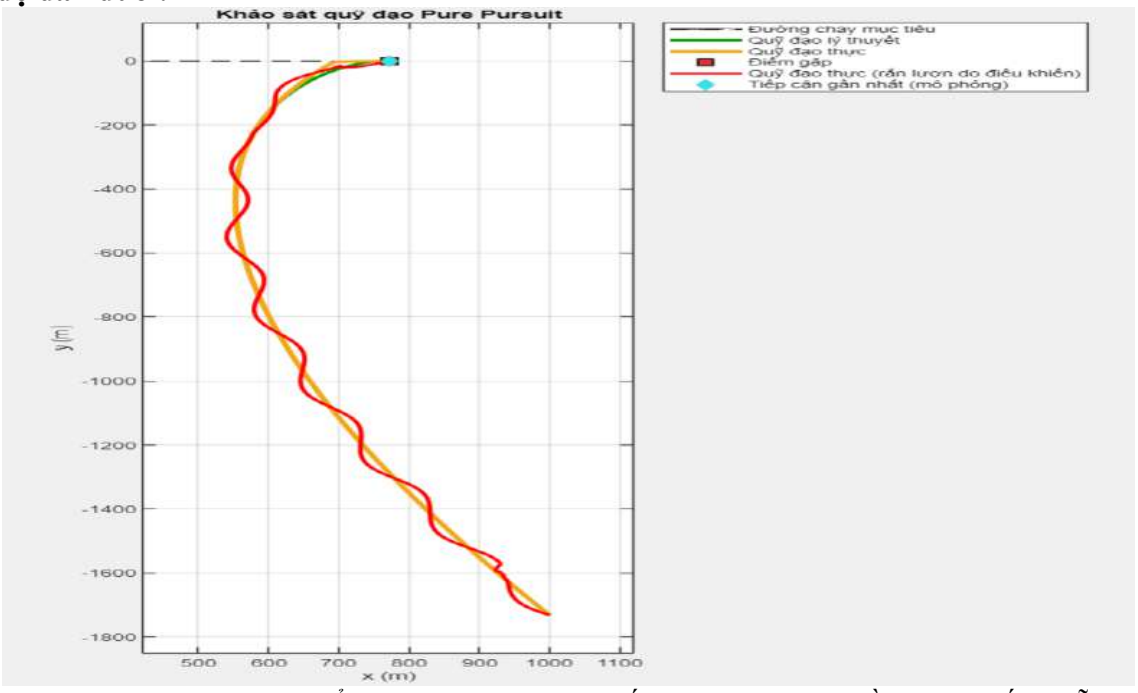
3.3. Ảnh hưởng tham số và ý nghĩa tác chiến:



Hình 4. Bản đồ góc mạn tối ưu $G_{k0,opt}(\lambda, R_0)$ và phân bố xác suất tiêu diệt theo góc khai hỏa

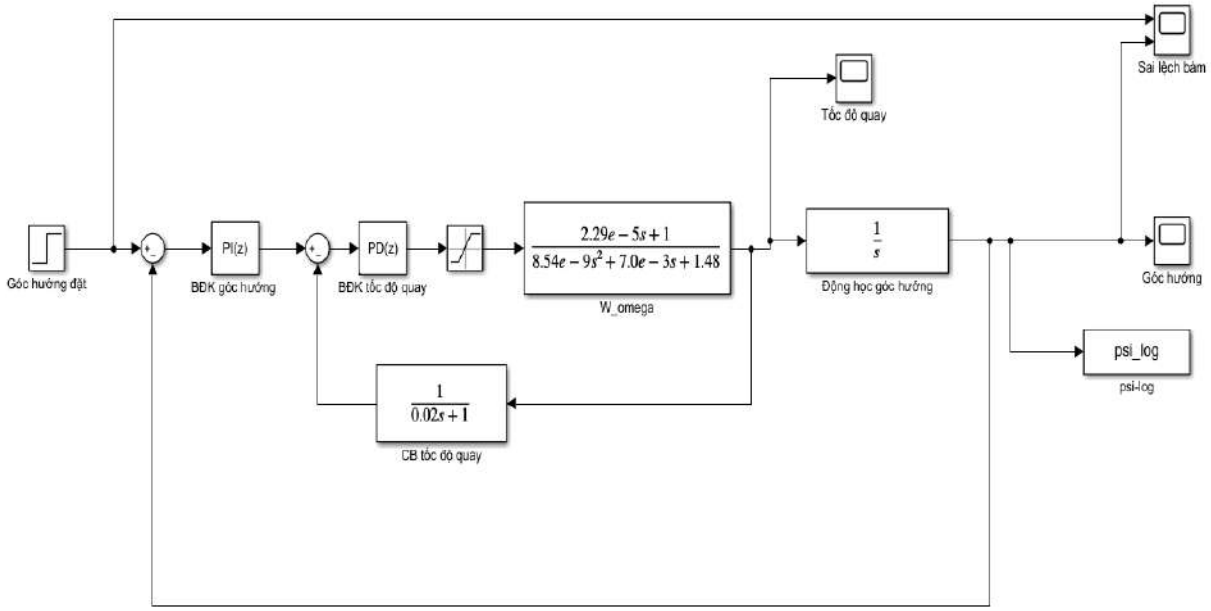
Bản đồ tham số cho thấy nghiệm tối ưu biến thiên trơn theo λ và R_0. Xác suất tiêu diệt đạt cực đại tại $G_{k0,opt}$ và giảm nhanh khi lệch khỏi giá trị này. Tồn tại một dải góc quanh $G_{k0,opt}$ có khả năng tiêu diệt mục tiêu cao (P ≥ 80%) cho phép lựa chọn góc khai hỏa linh hoạt trong thực tế tác chiến.

3.4. Xây dựng phương pháp tối ưu góc mạn ban đầu và bộ điều khiển hướng trong chế độ dẫn đuổi:



Hình 5. So sánh quỹ đạo chuyển động của ngư lôi khi tối ưu góc mạn ban đầu trong chế độ dẫn đuổi

Kết quả mô phỏng cho thấy, trong chế độ dẫn đuổi (Pure Pursuit), ngư lôi có xu hướng bám trực tiếp mục tiêu nên quỹ đạo chưa tối ưu và chịu ảnh hưởng của giới hạn bán kính quay R_{NL}, dẫn đến sai lệch so với quỹ đạo mong muốn. Trong nghiên cứu này, góc mạn ban đầu G_{k0} được xác định trên cơ sở mô hình động học tương đối giữa ngư lôi và mục tiêu. Giá trị tối ưu của góc này được lựa chọn nhằm tạo điều kiện hình học thuận lợi cho quá trình tiếp cận khi bước vào giai đoạn tự dẫn.



Hình 6. Sơ đồ khối bộ điều khiển hướng sử dụng góc mạn ban đầu tối ưu trong chế độ dẫn đuổi

Góc G_{k0} sau đó được thiết lập thông qua giai đoạn dẫn đường ban đầu và được thực hiện bởi hệ thống điều khiển hướng, bảo đảm ngư lôi đạt và duy trì hướng chuyển động phù hợp trong quá trình dẫn đuổi. Kết quả cho thấy phương pháp này giúp cải thiện rõ rệt quỹ đạo chuyển động, giảm sai lệch so với phương pháp bám trực tiếp, đồng thời nâng cao hiệu quả tiếp cận và khả năng bám mục tiêu.

4. Kết luận

Bài báo đã xây dựng mô hình xác định góc mạn ban đầu G_{k0} tối ưu trong chế độ bắn đuổi dựa trên quan hệ động học tương đối và ràng buộc gia tốc cơ động. Hàm tối ưu được chứng minh có một cực tiểu duy nhất trong miền góc nhọn, cho phép xác định nghiệm tối ưu một cách ổn định bằng phương pháp lát cắt vàng. Kết quả cho thấy góc tối ưu tăng theo λ, giảm theo R_0 và luôn nhỏ hơn 90° phù hợp với điều kiện hình học và khả năng cơ động thực tế. Đồng thời, tồn tại một dải góc quanh $G_{k0,\,opt}$ bảo đảm xác suất tiêu diệt cao, cho phép linh hoạt lựa chọn trong điều kiện có sai số và nhiễu.

Kết quả nghiên cứu cho phép người chỉ huy xác định nhanh góc khai hỏa tối ưu và vùng góc cho phép ngay tại thời điểm bắn, góp phần nâng cao xác suất tiêu diệt và độ tin cậy tác chiến. Góc G_{k0} tối ưu có thể được thiết lập thông qua giai đoạn dẫn đường ban đầu và đưa vào hệ điều khiển hướng của ngư lôi trước khi chuyển sang chế độ tự dẫn. Mô hình có thể mở rộng theo các hướng: xét mục tiêu cơ động, tích hợp sai số cảm biến và nhiễu môi trường hoặc tối ưu đa mục tiêu giữa xác suất tiêu diệt và tiêu hao năng lượng, hướng tới xây dựng hệ thống hỗ trợ khai hỏa thông minh trong tác chiến hiện đại.

Tài liệu tham khảo

1. Cơ sở lý thuyết điều khiển ngư lôi – HVHQ, năm 2001.

2. Liu S. X., Lin Z., et al. Three-body cooperative active defense guidance law with overload constraints, Chinese Journal of Aeronautics, 2025.

` 3. Han Wang & Zheng Chen. Nonlinear Optimal Guidance for Intercepting Moving Targets, 2025.

4. Shima T., Taub I. Intercept angle missile guidance under time varying acceleration bounds*, AIAA,* 2013.

5. Shima T. Intercept-angle guidance, Journal of Guidance, Control, and Dynamics, 2011.

6. Rusnak I., Weiss H., Eliav R., Shima T.Missile guidance with constrained intercept body angle*,* IEEE Transactions on Aerospace and Electronic Systems, 2014.

7. Shaferman V., Shima *T.*Linear quadratic guidance law for imposing a terminal intercept angle*,* AIAA Guidance, Navigation and Control Conference, 2008.

Thiết kế bộ điều khiển phi tuyến dựa trên phương pháp nghịch đảo động học kết hợp PI và kỹ thuật gain-scheduling cho hệ thống tự lái tên lửa

Nguyễn Đắc Nghĩa[1*], Nguyễn Ngọc Điển[1], Đoàn Thế Tuấn[1], Nguyễn Thanh Tùng[1]

[1]Học viện Kỹ thuật quân sự

**Email:nguyendacnghia3@gmail.com*

Tóm tắt

Bài báo trình bày phương pháp thiết kế bộ điều khiển phi tuyến dạng nối tầng dựa trên phương pháp nghịch đảo động học phi tuyến (Nonlinear Dynamic Inversion – NDI) kết hợp với bộ điều khiển PI (Proportional-Integral) và kỹ thuật điều chỉnh hệ số gain - scheduling cho hệ thống tự lái tên lửa. Quá trình thiết kế xét đến sự biến thiên của các tham số hệ thống, bao gồm khối lượng, mômen quán tính và các hệ số khí động học trong suốt quỹ đạo bay. Bộ điều khiển được xây dựng theo cấu trúc phân tầng, trong đó NDI đảm nhiệm vai trò tuyến tính hóa hệ phi tuyến, trong khi PI và gain scheduling được sử dụng để nâng cao chất lượng bám và khả năng thích nghi theo điều kiện làm việc thay đổi. Hiệu năng và tính bền vững của hệ thống được đánh giá thông qua mô phỏng phi tuyến trong môi trường MATLAB/Simulink. Kết quả cho thấy cấu trúc điều khiển kết hợp cải thiện rõ rệt khả năng bám tín hiệu nhỏ và duy trì ổn định trong dải tham số rộng.

Từ khóa: *Tự lái tên lửa; nghịch đảo động học phi tuyến NDI; điều khiển PI; gain-scheduling.*

1. Đặt vấn đề

Trong hệ thống tự lái tên lửa, đối tượng điều khiển là hệ động lực học phi tuyến với nhiều tham số biến thiên theo thời gian như hệ số khí động, khối lượng, vận tốc bay và mômen quán tính. Sự biến thiên này gây ra tính bất định đáng kể, làm phức tạp bài toán tổng hợp bộ điều khiển đảm bảo ổn định và hiệu năng.

Trong nhiều thập kỷ qua, nhiều phương pháp điều khiển đã được áp dụng cho bài toán tự lái tên lửa như: phương pháp đặt cực [1], phương pháp sử dụng bất đẳng thức ma trận tuyến tính kết hợp với kỹ thuật điều chỉnh hệ số gain-scheduling [2], LPV [3], điều khiển tối ưu [4]. Trong đó, phương pháp sử dụng bộ điều khiển bền vững H_∞ được sử dụng rộng rãi nhờ khả năng đảm bảo đồng thời độ bền vững và khả năng chống nhiễu. Tuy nhiên, bộ điều khiển H_∞ thường có bậc cao, cấu trúc phức tạp và có thể suy giảm hiệu năng trong vùng tín hiệu nhỏ.

Xuất phát từ hướng tiếp cận trên, bài báo đề xuất thiết kế bộ điều khiển phi tuyến dựa trên phương pháp nghịch đảo động học phi tuyến [5, 6], kết hợp với bộ điều khiển PI và kỹ thuật gain-scheduling cho kênh điều khiển gia tốc pháp tuyến của tên lửa. Trong cấu trúc đề xuất, bộ điều khiển PI được sử dụng ở vòng ngoài để tạo lệnh tốc độ góc từ sai lệch gia tốc pháp tuyến, trong khi phương pháp NDI được áp dụng ở vòng trong nhằm xử lý tính phi tuyến của hệ và tính toán tín hiệu điều khiển.

Kỹ thuật gain-scheduling được sử dụng để cập nhật các tham số của mô hình theo điều kiện bay, qua đó nâng cao khả năng thích nghi của hệ thống khi các tham số động lực học thay đổi theo quỹ đạo. Cấu trúc của bài báo bao gồm: xây dựng mô hình động lực học tên lửa, thiết kế bộ điều khiển dựa trên NDI, xây dựng bộ điều khiển PI và kỹ thuật gain scheduling, đánh giá hiệu năng thông qua mô phỏng phi tuyến trong môi trường MATLAB/Simulink.

2. Mô hình động lực học của tên lửa

Xét tên lửa khí động điển hình với 2 cặp cánh đối xứng với nhau quanh trục dọc. Giả thiết tên lửa ổn định quanh trục dọc và không có sự ảnh hưởng lẫn nhau giữa các mặt phẳng điều khiển. Chuyển động của tên lửa trong mặt phẳng thẳng đứng được mô tả bởi hệ phương trình

* Email: nguyendacnghia3@gmail.com

phi tuyến biểu diễn mối quan hệ giữa góc tấn công, tốc độ quay, gia tốc pháp tuyến và góc lệch cánh lái.

Chuyển động của tên lửa trong mặt phẳng thẳng đứng được biểu diễn trong phương trình sau [7]:

$$\begin{aligned}\dot{\alpha} &= \frac{q_\infty S_{\text{ref}}}{mV_{\text{M}}}\left(\sin\alpha C_A - \cos\alpha C_N\right) + q - \sin\alpha\frac{q_\infty S_{\text{ref}} C_T}{mV_{\text{M}}} \\ \dot{q} &= \frac{q_\infty S_{\text{ref}} d}{I_y^b}\left(C_m + \bar{s}C_N\right)\end{aligned} \tag{1}$$

Trong đó: Trong đó: α là góc tấn công; q là tốc độ góc nghiêng quỹ đạo của tên lửa; q_∞ là áp suất động; S_{ref} là diện tích tham chiếu khí động tên lửa; m là khối lượng của tên lửa; V_M là vận tốc của tên lửa; d là đường kính của tên lửa; $\bar{s}$ là khoảng cách quy đổi (tính theo đơn vị đường kính của tên lửa) tính từ trọng tâm đến tâm áp; I_y^b là momen quán tính theo trục hệ tọa độ liên kết O_{Y1}; C_A, C_N tương ứng là các hệ số khí động: hệ số lực cản và hệ số lực nâng; C_T, C_m lần lượt là hệ số lực đẩy và hệ số momen.

Thực hiện biến đổi ta nhận được:

$$\begin{cases}\dot{\alpha} = a_{11}^a\alpha + a_{12}^a q + b_{11}^a\delta_q \\ \dot{q} = a_{21}^a\alpha + a_{22}^a q + b_{22}^a\delta_q\end{cases} \tag{2}$$

Gia tốc pháp tuyến của tên lửa được xác định như sau:

$$n_y = -\frac{1}{m}q_\infty S_{\text{ref}} C_N \tag{3}$$

Đặt:

$$\left.\begin{aligned}\boldsymbol{x} &= \begin{bmatrix}\alpha & q\end{bmatrix}^{\text{T}} \\ \boldsymbol{u} &= \delta_q \\ \boldsymbol{y} &= \begin{bmatrix}n_y & \alpha & q\end{bmatrix}^{\text{T}}\end{aligned}\right\} \tag{4}$$

$$\boldsymbol{A}_a = \begin{bmatrix}a_{11}^a & a_{12}^a \\ a_{21}^a & a_{22}^a\end{bmatrix};\ \boldsymbol{B}_a = \begin{bmatrix}b_{11}^a \\ b_{21}^a\end{bmatrix};\ \boldsymbol{C}_a = \begin{bmatrix}c_{11} & c_{12} \\ 1 & 0 \\ 0 & 1\end{bmatrix};\ \boldsymbol{D}_a = \begin{bmatrix}d_{11} \\ 0 \\ 0\end{bmatrix}$$

Khi đó, mô hình động lực học của tên lửa trong mặt phẳng thẳng đứng được biểu diễn dưới dạng không gian trạng thái như trong phương trình:

$$\begin{cases}\dot{\boldsymbol{x}} = \boldsymbol{A}_a\boldsymbol{x} + \boldsymbol{B}_a\boldsymbol{u} \\ \boldsymbol{y} = \boldsymbol{C}_a\boldsymbol{x} + \boldsymbol{D}_a\boldsymbol{u}\end{cases} \tag{5}$$

3. Tổng hợp bộ điều khiển và đánh giá

3.1 Mô hình bất định của tên lửa

Mô hình bất định của tên lửa tại các giá trị biên của các tham số khối lượng, momen quán tính, vị trí trọng tâm, góc tấn công và góc lệch cánh lái được biểu diễn như sau [7]:

$$\begin{cases} \dot{\boldsymbol{x}} = \begin{bmatrix} a_{11}^{unc} & 0.9995 \\ a_{21}^{unc} & -9.3868 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} b_{11}^{unc} \\ b_{21}^{unc} \end{bmatrix} \boldsymbol{u} \\ \boldsymbol{y} = \begin{bmatrix} c_{11}^{unc} & 0.005 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} d_{11}^{unc} \\ 0 \\ 0 \end{bmatrix} \boldsymbol{u} \end{cases} \quad (6)$$

Trong đó:

$$a_{11}^{unc} \in [-4.2706,\ -2.4625]; \quad a_{21}^{unc} \in [-2660.5,\ -527.47]$$
$$b_{11}^{unc} \in [-0.6319,\ 0.0224]; \quad b_{21}^{unc} \in [-4242.3,\ -2135.4]$$
$$c_{11}^{unc} \in [-38.1756,\ -18.7101]; \ d_{11}^{unc} \in [-5.9507,\ -4.1760]$$

Như vậy thông qua việc biểu diễn mô hình động học của tên lửa trong mặt phẳng thẳng đứng dưới dạng hệ phương trình không gian trạng thái với các hệ số phụ thuộc thời gian, tương ứng với sự biến thiên của khối lượng, momen quán tính và vị trí trọng tâm của tên lửa trong quá trình bay, mô hình thu được phản ánh nhất quán các đặc trưng bất định và phi tuyến vốn có của đối tượng điều khiển. Trên cơ sở mô hình này, các tính chất động học quan trọng của tên lửa có thể được khảo sát, qua đó là cơ sở nền tảng cho việc thiết kế các bộ điều khiển sau này.

3.2 Tổng hợp bộ điều khiển cho hệ thống tự lái tên lửa dựa trên phương pháp NDI

Trong bài toán điều khiển hệ thống tự lái tên lửa, mô hình động lực học có tính phi tuyến mạnh và các tham số thay đổi đáng kể theo điều kiện bay, bao gồm khối lượng, mômen quán tính và các hệ số khí động. Điều này gây ra những thách thức lớn đối với các phương pháp điều khiển tuyến tính truyền thống, đặc biệt khi yêu cầu hệ thống phải duy trì hiệu năng ổn định trên toàn bộ dải làm việc. Do đó, phương pháp nghịch đảo động học NDI được lựa chọn nhằm xử lý trực tiếp tính phi tuyến của hệ.

Theo cách tiếp cận chung trong điều khiển bay hiện đại, hệ động lực học phi tuyến có thể được biểu diễn dưới dạng không gian trạng thái như sau:

$$\dot{x} = f(x) + g(x)u \quad (7)$$

trong đó x là vector trạng thái, u là tín hiệu điều khiển, f(x) và g(x) là các hàm phi tuyến.

Giả sử đầu ra cần điều khiển là y=h(x), khi đó bằng cách lấy đạo hàm theo thời gian, ta thu được mối quan hệ:

$$y^{(r)} = b(x) + a(x)u \quad (8)$$

Trong đó a(x) và b(x) là các hàm phi tuyến phụ thuộc trạng thái. Khi đó, tín hiệu điều khiển được xác định theo:

$$u = a^{-1}(x)(v - b(x)) \quad (9)$$

với v là tín hiệu điều khiển ảo. Nhờ phép biến đổi này, hệ phi tuyến ban đầu được chuyển về dạng:

$$y^{(r)} = v \quad (10)$$

Cho phép áp dụng các phương pháp điều khiển tuyến tính để thiết kế luật điều khiển.

Xét mô hình động học của tên lửa được mô tả bởi phương trình (2). Từ phương trình thứ 2, luật điều khiển NDI được xây dựng bằng cách chọn động học mong muốn có dạng: $\dot{q} = v$.

Suy ra luật điều khiển:

$$\delta = \frac{v - a_{21}\alpha - a_{22}q}{b_{21}} \quad (11)$$

Trong đó tín hiệu điều khiển ảo v được lựa chọn theo dạng phản hồi tuyến tính:

$$v = k_q(q_c - q) \quad (12)$$

với q_c là tín hiệu lệnh tốc độ góc do vòng điều khiển bên ngoài cung cấp, còn k_q là hệ số phản hồi. Thay vào biểu thức trên thu được luật điều khiển thực hiện trong mô hình:

$$\delta = \frac{-a_{21}\alpha - a_{22}q + k_q(q_c - q)}{b_{21}} \quad (13)$$

Bộ điều khiển NDI sử dụng trực tiếp các biến trạng thái và tín hiệu lệnh để tính toán góc quay cánh lái, qua đó triệt tiêu các thành phần phi tuyến chính trong động học quay của hệ. Với cách tiếp cận này, vòng điều khiển trong có khả năng đáp ứng nhanh và phù hợp với yêu cầu điều khiển của kênh tự lái tên lửa.

3.3. Tổng hợp bộ điều khiển PI và kỹ thuật gain-scheduling

Để cải thiện chất lượng điều khiển, đặc biệt trong vùng tín hiệu nhỏ và khi tham số hệ thống biến thiên theo quỹ đạo bay, bộ điều khiển PI được kết hợp với kỹ thuật gain-scheduling trong cấu trúc điều khiển đề xuất.

Bộ điều khiển PI được sử dụng tại vòng điều khiển ngoài nhằm thực hiện bài toán bám gia tốc pháp tuyến.

$$q_c = k_P e_n + k_I \int e_n dt \quad (14)$$

Trong đó: e_n: sai số; k_P: hệ số tỷ lệ; k_I: hệ số tích phân.

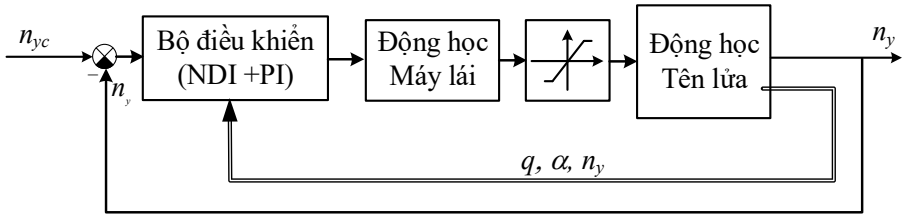
Trong cấu trúc này, q_c đóng vai trò là tín hiệu trung gian nối giữa bài toán điều khiển gia tốc pháp tuyến của vòng ngoài và bài toán điều khiển tốc độ góc của vòng trong.

Bên cạnh đó, do các tham số động lực học của tên lửa như khối lượng, mômen quán tính và hệ số khí động học thay đổi đáng kể theo điều kiện bay, kỹ thuật gain-scheduling được áp dụng để điều chỉnh các tham số của bộ điều khiển theo trạng thái làm việc của hệ. Cụ thể, các hệ số điều khiển a_{12}, a_{22}, b_{21} được điều chỉnh theo thời gian bay để phản ánh sự biến thiên của các tham số khí động của tên lửa.

Trong cấu trúc điều khiển tổng thể, bộ điều khiển NDI đảm nhiệm vai trò tuyến tính hóa và điều khiển chính, trong khi bộ điều khiển PI và kỹ thuật gain-scheduling đóng vai trò bổ sung, giúp cải thiện khả năng bám, tăng độ bền vững và thích nghi với sự biến thiên của hệ thống.

3.4. Cấu trúc bộ điều khiển tổng thể

Trên cơ sở các bộ điều khiển đã được xây dựng, cấu trúc điều khiển tổng thể của hệ tự lái tên lửa được tổ chức theo dạng hai vòng lồng nhau, gồm vòng điều khiển ngoài và vòng điều khiển trong.



Hình 1. Sơ đồ cấu trúc bộ điều khiển phi tuyến dựa trên phương pháp nghịch đảo động học kết hợp PI và kỹ thuật gain scheduling

Vòng điều khiển ngoài thực hiện bài toán bám gia tốc pháp tuyến. Từ sai lệch giữa gia tốc pháp tuyến thực và gia tốc pháp tuyến đặt, bộ điều khiển PI sinh ra tín hiệu lệnh tốc độ góc. Vai trò của vòng này là chuyển yêu cầu điều khiển ở mức gia tốc thành một tín hiệu trung gian phù hợp với động học quay của hệ.

Vòng điều khiển trong sử dụng phương pháp NDI để điều khiển trực tiếp tốc độ góc q. Trên cơ sở mô hình động học quay của tên lửa, bộ điều khiển tính toán góc lệch cánh lái δ từ các biến trạng thái α, q và tín hiệu lệnh q_c. Nhờ đó, các thành phần phi tuyến chi phối động học quay được bù trực tiếp trong luật điều khiển, cho phép hệ kín có đáp ứng nhanh và ổn định.

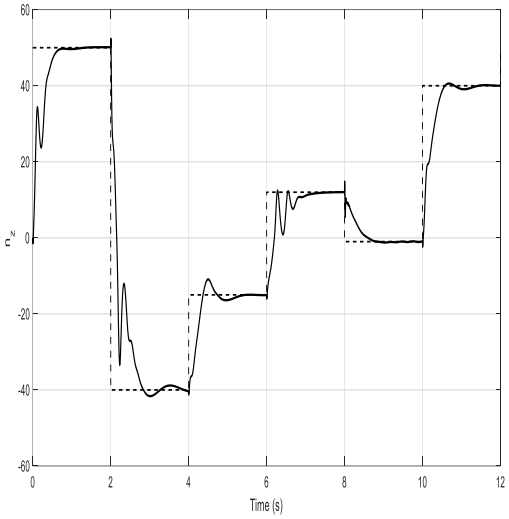
Trong cấu trúc này, góc tấn công α không được điều khiển bởi một vòng riêng biệt mà được sử dụng như một biến trạng thái trong luật điều khiển vòng trong. Cách xây dựng này phản ánh đúng cấu trúc điều khiển được hiện thực trong mô hình mô phỏng, đồng thời giúp giữ cho bộ điều khiển đủ đơn giản mà vẫn bảo đảm hiệu năng yêu cầu.

Kỹ thuật gain-scheduling được tích hợp trong vòng điều khiển trong để cập nhật các hệ số động học theo từng giai đoạn bay. Sự kết hợp giữa vòng ngoài PI, vòng trong NDI và cơ chế hiệu chỉnh tham số cho phép bộ điều khiển vừa bảo đảm khả năng bám gia tốc pháp tuyến, vừa duy trì chất lượng điều khiển khi điều kiện làm việc của hệ thay đổi.

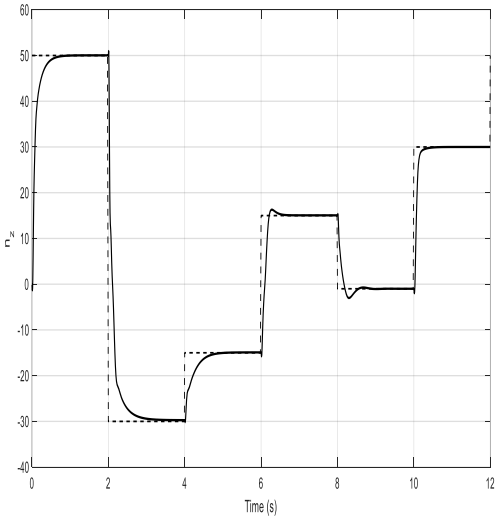
3.5 Kết quả mô phỏng

Để đánh giá hiệu quả của bộ điều khiển đề xuất, các mô phỏng phi tuyến được thực hiện trong môi trường MATLAB/Simulink với các kịch bản thay đổi tín hiệu gia tốc pháp tuyến theo dạng bậc thang. Kết quả thu được được so sánh với bộ điều khiển bền vững H_∞ kết hợp PID sử dụng logic mờ được trình bày trong [8].

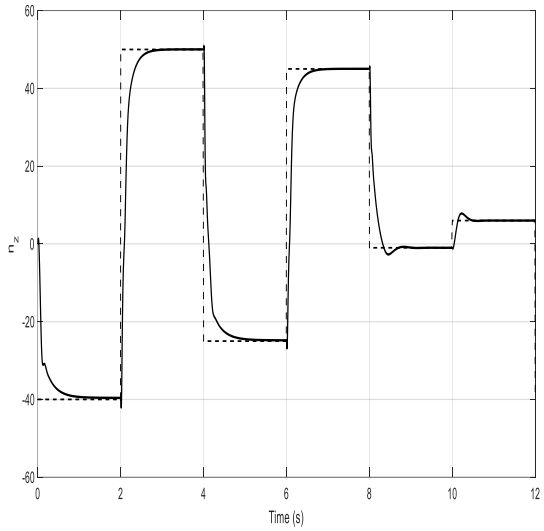
Đối với bộ điều khiển bền vững H_∞ kết hợp PID sử dụng logic mờ cho thấy hệ thống có khả năng bám tín hiệu gia tốc pháp tuyến khá tốt và duy trì ổn định trong các điều kiện làm việc khác nhau. Tuy nhiên, đáp ứng của hệ xuất hiện dao động và quá điều chỉnh tương đối rõ khi tín hiệu đặt thay đổi đột ngột, đặc biệt tại các thời điểm chuyển mức. Thời gian quá độ còn kéo dài và tín hiệu đầu ra cần một khoảng thời gian nhất định để hội tụ về giá trị đặt. Điều này cho thấy mặc dù phương pháp có khả năng xử lý bất định và phi tuyến, tuy nhiên, phương pháp này cũng tồn tại một số hạn chế nhất định. Trước hết, quá trình thiết kế bộ điều khiển H∞ yêu cầu xây dựng mô hình tuyến tính hóa và lựa chọn các hàm trọng số phù hợp, điều này làm tăng độ phức tạp trong thiết kế. Bên cạnh đó, việc xây dựng và hiệu chỉnh hệ luật mờ phụ thuộc nhiều vào kinh nghiệm của người thiết kế, gây khó khăn trong việc tối ưu hóa và triển khai thực tế.



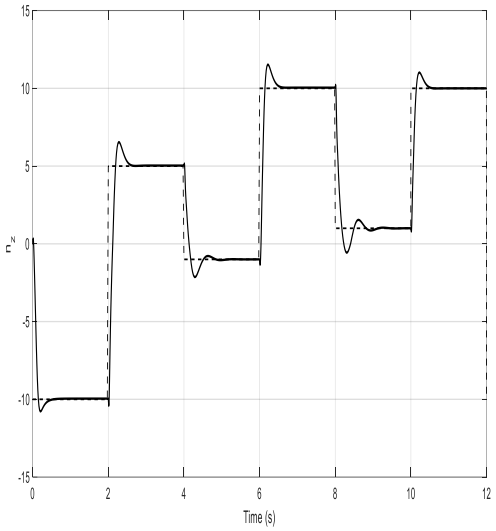
Hình 2. Gia tốc pháp tuyến của tên lửa với bộ điều khiển H_∞ kết hợp PID sử dụng fuzzy logic



Hình 3. Gia tốc pháp tuyến của tên lửa với bộ điều khiển phi tuyến sử dụng NDI kết hợp PI và gain-scheduling



Hình 4. Gia tốc pháp tuyến của tên lửa với bộ điều khiển phi tuyến sử dụng NDI kết hợp PI và gain-scheduling trường hợp gia tốc pháp tuyến thay đổi đột ngột liên tục



Hình 5. Gia tốc pháp tuyến của tên lửa với bộ điều khiển phi tuyến sử dụng NDI kết hợp PI và gain-scheduling trường hợp đầu vào gia tốc pháp tuyến nhỏ

Kết quả mô phỏng cho bộ điều khiển đề xuất dựa trên NDI kết hợp PI và gain-scheduling cho thấy hệ thống có khả năng bám rất tốt tín hiệu gia tốc pháp tuyến trong toàn bộ dải làm việc. Đáp ứng của hệ có thời gian quá độ ngắn, phản ứng nhanh với các thay đổi của tín hiệu đặt và nhanh chóng đạt trạng thái ổn định. Sai lệch xác lập gần như bằng không, cho thấy hiệu quả của bộ điều khiển PI trong việc cải thiện chất lượng bám. Đồng thời, nhờ phương pháp NDI, các thành phần phi tuyến của hệ được xử lý hiệu quả, giúp đáp ứng mượt và ít dao động

hơn so với các phương pháp điều khiển truyền thống. Việc áp dụng gain-scheduling cũng góp phần duy trì hiệu năng điều khiển ổn định khi tham số hệ thay đổi theo thời gian.

4. Kết luận

Bài báo đã trình bày phương pháp thiết kế bộ điều khiển phi tuyến cấu trúc phân tầng dựa trên nghịch đảo động học phi tuyến, kết hợp với bộ điều khiển PI và kỹ thuật gain- scheduling cho hệ thống tự lái tên lửa có tham số biến thiên. Trong cấu trúc đề xuất, NDI được sử dụng để tuyến tính hóa hệ phi tuyến và đảm bảo khả năng điều khiển trong dải hoạt động rộng, trong khi bộ điều khiển PI góp phần cải thiện chất lượng bám và giảm sai lệch xác lập, đặc biệt trong vùng tín hiệu nhỏ. Kỹ thuật gain scheduling cho phép điều chỉnh tham số điều khiển theo điều kiện làm việc, từ đó nâng cao khả năng thích nghi của hệ thống khi các tham số động lực học thay đổi theo quỹ đạo bay.

Kết quả mô phỏng phi tuyến trong môi trường MATLAB/Simulink cho thấy bộ điều khiển đề xuất đạt được chất lượng bám tốt, thời gian đáp ứng nhanh và duy trì tính ổn định trong dải tham số rộng. Phương pháp này là một hướng tiếp cận hiệu quả cho bài toán điều khiển tự lái tên lửa trong điều kiện phi tuyến và bất định.

Tài liệu tham khảo

1. RYU, S.M., D.Y. WON, C.H. LEE and M.J. TAHK. Missile Autopilot Design for Ag-ile Turn Control During Boost-phase. International Journal of Aeronautical and Space Sciences, 2011, 12(4), pp. 365-370. DOI 10.5139/IJASS.2011.12.4.365.
2. WON, D.Y., M.J. TAHK and K.Y. KIM. Robust Gain-scheduling Technique for an Agile Missile Subject to Mass Variation. In: International Conference on Control, Automation and Systems (ICCAS). Gyeonggi-do: IEEE, 2010, pp. 2119-2123. DOI 10.1109/ICCAS.2010.5670191.
3. THEODOULIS, S. and G. DUC. Missile Autopilot Design: Gain Scheduling and the Gap Metric. Journal of Guidance, Control, and Dynamics, 2009, 32(3), pp. 986-996. DOI 10.2514/1.34756.
4. DOYLE, J., K. GLOVER, P. KHARGONEKAR and B. FRANCIS. State-space Solu-tions to Standard H2 and H∞ Control Problems. In: 1988 American Control Con-ference. Atlanta: IEEE, 1988, pp. 1691-1696. DOI 10.23919/ACC.1988.4789992.
5. ENNS, D., D. BUGAJSKI, R. HENDRICK and G. STEIN. Dynamic Inversion: An Evolving Methodology for Flight Control Design. *International Journal of Control*, 1994, 59(1), pp. 71–91. DOI: 10.1080/00207179408923066.
6. SIEBERLING, S., Q.P. CHU and J.A. MULDER. Robust Flight Control Using Incremental Nonlinear Dynamic Inversion. In: *AIAA Guidance, Navigation, and Control Conference*. Toronto: AIAA, 2010. DOI: 10.2514/6.2010-8224.
7. HERNÁNDEZ, F.D.B.I. Optimization of the Integrated Guidance and Control for a Dual Aerodynamic Control Missile [PhD Thesis]. Universidad Politécnica de Madrid, 2015. DOI 10.20868/UPM.thesis.38301.
8. NGUYEN, T.C. and C.H. TINH. Design of Robust H∞ Combined with PID Controller Using Fuzzy Logic for a Missile Autopilot. *Advances in Military Technology*, 2024, 19(1). DOI: https://doi.org/10.3849/aimt.01831

FPGA-Accelerated YOLOv5 for Real-Time Thermal Image Object Detection

Nguyễn Ngọc Anh[1*], Nguyễn Ngọc Hưng[1]

[1] Viện Tên lửa và Kỹ thuật điều khiển, HVKTQS

** Email: nguyenngocanh@lqdtu.edu.vn*

Abstract

Real-time thermal image object detection plays a crucial role in applications such as surveillance, military, and security systems, particularly in low-light and harsh environmental conditions. While deep learning–based models such as YOLOv5 achieve high detection accuracy, deploying them on embedded platforms is challenging due to limited computational resources, power constraints, and strict real-time latency requirements. This paper presents an FPGA-accelerated YOLOv5 implementation on a Multi-Processor System-on-Chip (MPSoC) for real-time thermal object detection. The design is optimized using efficient dataflow, model quantization, and hardware tuning to balance accuracy, speed, and power consumption. Experimental results show that the proposed system achieves stable real-time performance on thermal datasets with only minor accuracy loss compared to a software-based implementation, demonstrating the practicality of low-latency and energy-efficient thermal object detection on embedded MPSoC platforms.

Keywords: *MPSoC, VitisAI, YOLOv5, FPGA.*

1. Introduction

Object detection plays a critical role in a wide range of applications, including surveillance, maritime traffic monitoring, and national defense systems. In recent years, deep learning–based object detection algorithms, particularly those in the You Only Look Once (YOLO) family, have gained significant attention due to their ability to achieve high detection accuracy with real-time inference performance. Among them, YOLOv5 has emerged as a widely adopted solution owing to its favorable balance between detection accuracy, inference speed, and deployment flexibility [1, 7, 8].

Despite their effectiveness, deploying high-performance deep learning–based object detectors on embedded and edge computing platforms remains a major challenge. Such models typically involve intensive computation and high memory bandwidth requirements, which are difficult to satisfy under strict constraints on power consumption, hardware resources, and real-time latency. These limitations are especially critical in long-term, unattended deployment scenarios, such as shipborne monitoring systems and remote surveillance stations [1, 10].

Conventional GPU-based edge platforms provide strong computational capabilities but often suffer from high power consumption, increased system cost, and larger form factors [11]. These drawbacks limit their suitability for energy-constrained and space-limited embedded applications. Consequently, there is a growing demand for alternative hardware platforms that can deliver efficient real-time inference while maintaining low power consumption.

Field-Programmable Gate Arrays (FPGAs) have emerged as a promising solution for edge AI applications due to their inherent parallelism, reconfigurability, and superior energy efficiency. Modern adaptive computing platforms, such as Xilinx Zynq UltraScale+ MPSoC devices, integrate multi-core ARM processors with programmable logic and dedicated deep learning accelerators, enabling flexible hardware–software co-design [1, 15]. In particular, the Deep Learning Processing Unit (DPU) provides a parameterizable architecture optimized for

**Email: nguyenngocanh@lqdtu.edu.vn*

convolutional neural network (CNN) inference, allowing efficient execution of deep learning workloads on FPGA fabric.

To facilitate the deployment of deep learning models on FPGA platforms, the Vitis AI toolchain offers an end-to-end workflow encompassing model quantization, optimization, compilation, and deployment [4]. By converting floating-point neural networks into INT8 quantized models, Vitis AI significantly reduces computational complexity and memory bandwidth requirements, while leveraging DPU acceleration to achieve high-throughput and low-power inference on embedded systems.

Recent studies on infrared and thermal object detection have mainly focused on model improvement to handle the difficult properties of thermal imagery, such as low resolution, weak texture, high noise, and small targets. For example, YOLO-FIRI improves shallow feature extraction, attention, and multi-scale detection to enhance the detection of small and weak objects in infrared images [9]. Similarly, Cheng et al. showed the potential of YOLOv5 for thermal imaging by building a dedicated dataset of 5736 thermal images and improving YOLOv5s with CBAM, Ghost convolution, and data augmentation methods for dark aquatic environments [5]. These studies clearly show that YOLO-based models are effective for thermal and infrared detection. However, they mainly focus on algorithm design and dataset adaptation, while giving limited attention to real-time deployment on FPGA/MPSoC platforms, where quantization, latency, and power efficiency must also be considered.

On the hardware side, several studies have shown that FPGA is a strong platform for real-time YOLO inference in embedded systems. Yan et al. proposed a 4-bit YOLOv5 accelerator with optimized DSP packing and adder tree design, achieving 29% higher throughput than a commercial Vitis AI design on the same ZCU102 platform [15]. Other FPGA-based works using Vitis AI and DPU also reported high inference speed and low power consumption for edge applications [4]. However, most of these studies focus on industrial vision, RGB or SAR data, or evaluate only a single hardware platform. As a result, real-time thermal object detection on MPSoC FPGA remains insufficiently studied, especially under a unified evaluation of detection accuracy, inference speed, and deployment efficiency.

To address this gap, this paper presents an FPGA-accelerated implementation of YOLOv5n on an MPSoC platform using Vitis AI 3.0. The proposed system adopts a hardware–software co-design approach and is evaluated on thermal image datasets, with comparisons against representative edge platforms, including the Jetson Orin Nano and the NVIDIA GeForce RTX 4060 8GB GPU. Performance is assessed using Precision, Recall, mean Average Precision (mAP), and Frames Per Second (FPS). The results demonstrate that the proposed approach provides an energy-efficient and practical solution for low-latency thermal object detection in edge computing scenarios.

2. YOLOv5-Based Thermal Object Detection

Since its introduction, the YOLO series algorithms have received widespread attention for their excellent detection speed and reliable detection accuracy, becoming one of the representatives of one-stage detection networks. YOLOv5, the fifth iteration of the YOLO network, further optimizes the architecture to enhance both detection speed and accuracy [2, 12, 13]. The detailed network structure of the YOLOv5 model is shown in Figure 1.

*Email: nguyenngocanh@lqdtu.edu.vn

YOLOv5 itself has undergone multiple network structure adjustments, and this paper uses the latest version as the basic model for object detection networks. The basic modules of YOLOv5 are the convolution layer, the batch normalization (BN) layer, and the Sigmoid Linear Unit (SiLU) activation (CBS). Convolution layers are used to extract local spatial information from input features, typically consisting of multiple convolution kernels, each of which extracts a type of spatial information. The BN layer is used to standardize the activation distribution on each channel of the input feature-map. It can accelerate the training process and improve the model's generalization ability while reducing the model's dependence on initialization. SiLU is used as an activation function to introduce nonlinear transformation capability to neural networks. In addition, its smooth feature at zero points can facilitate better transmission of gradient information in the network [15].

From the algorithmic perspective, these modules are important for detection accuracy. However, from the hardware deployment perspective, not all of them are equally suitable for FPGA implementation. Among the components of YOLOv5, the most FPGA-friendly operations are the convolution-dominant parts of the network, including standard convolution, batch normalization, residual feature extraction, and multi-scale feature fusion, because these operations are regular, highly parallel, and account for most of the computational workload of the model. For this reason, they are the main targets for acceleration on the DPU in Vitis AI–based deployment. Several FPGA-based studies have shown that convolution-based CNN layers can be efficiently mapped onto FPGA hardware after INT8 quantization, enabling real-time inference with low power consumption [4, 14].

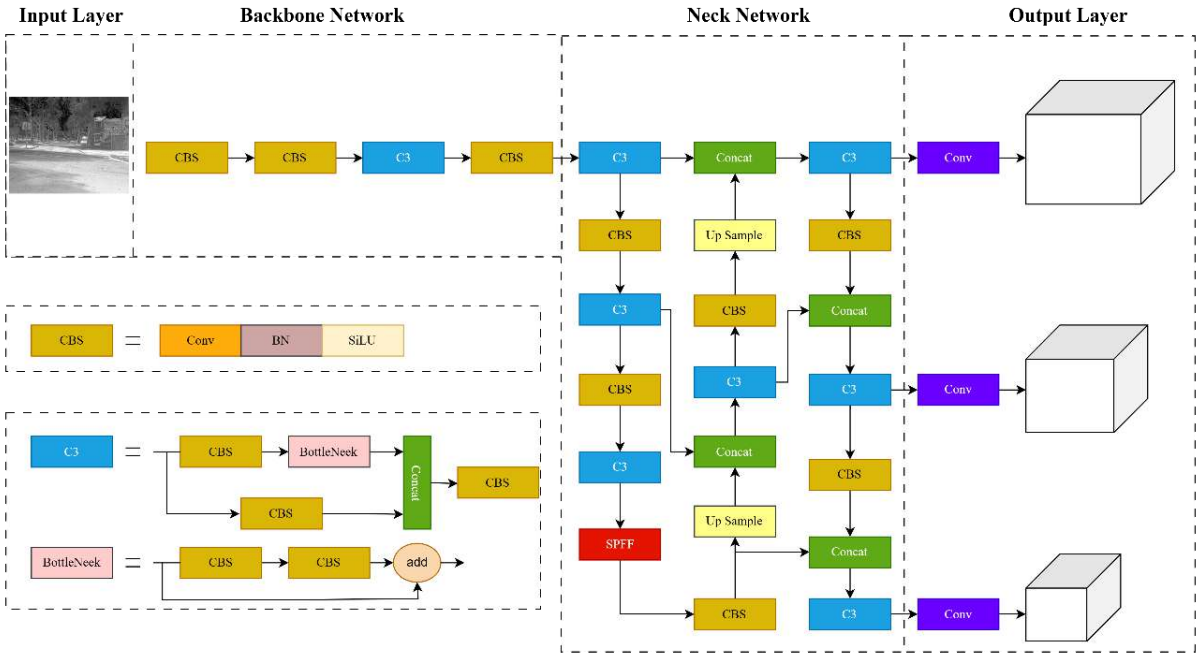


Figure 1. Architecture of the YOLOv5 model

In contrast, some operators in the original YOLOv5 model are less hardware-friendly. In particular, although SiLU can improve model performance, it introduces more complex nonlinear computation and is therefore not ideal for efficient execution on many FPGA/DPU deployment flows. For this reason, previous embedded implementations often replace SiLU

*Email: nguyenngocanh@lqdtu.edu.vn

with a simpler activation such as ReLU or Leaky ReLU to improve hardware compatibility while keeping acceptable accuracy. For example, one recent YOLOv5n deployment study replaced the original SiLU with LeakyReLU with a negative slope of 26/256 to ensure compatibility across hardware platforms, while another FPGA-based lightweight YOLOv5 design replaced SiLU with ReLU to simplify DPU execution [1, 4]. Therefore, in this study, SiLU is replaced with Leaky ReLU with a negative slope of 26/256 in order to maintain model compatibility with the DPU while preserving detection accuracy.

On the basis of the CBS module, YOLOv5 constructs the CBSx3 (C3) structure, which integrates features from different levels and uses residual connections to provide better gradient transfer. Its main function is to increase the depth and receptive field of the network and improve the ability of feature extraction. The C3 module contains multiple Conv layers, where a stride of two convolutions can reduce the size of the feature-map by half, increase the receptive field of the network, and reduce computational complexity. By halving the size of the feature-map, the network can focus more on the global information of objects, thereby improving the effectiveness of feature extraction. Convolution with stride 1 does not change the size of the feature-map, which can maintain the spatial resolution of the feature-map and better preserve the local information of the object. In addition, YOLOv5 has improved the Spatial Pyramid Pooling (SPP) structure by proposing the Spatial Pyramid Pooling-Fast (SPPF) module, which achieves faster inference speed with less computation while still maintaining the functions of avoiding duplicate feature extraction and image distortion in the SPP module [6].

In terms of macrostructure, YOLOv5 still adopts the three-level structure of previous YOLO models, namely backbone, neck, and head. The backbone is responsible for feature extraction, the neck performs multiscale feature fusion through upsample and downsample operations, and the head completes the final regression detection. For FPGA implementation, the backbone and neck are the most suitable parts for acceleration because they are dominated by regular and compute-intensive convolution operations. Most convolution layers in the detection head can also be mapped efficiently to the DPU. However, the final transformation stage of the detection head is less hardware-friendly, since operators such as *permute* and *view* are not directly supported by the Vitis AI quantization and compilation flow. In addition, post-processing steps, including bounding-box decoding, confidence filtering, and non-maximum suppression (NMS), are usually executed on the ARM cores rather than on the FPGA. Therefore, YOLOv5 deployment on MPSoC FPGA is better understood as a hardware–software co-design approach, where the convolution-dominant layers are accelerated on the DPU, while unsupported or control-dominant operators are executed on the CPU.

3. FPGA-Based Deployment on MPSoC

Figure 2 illustrates the workflow for developing AI applications using the Xilinx Vitis AI platform, covering the entire process from hardware development and model compilation to application development. The workflow consists of three main stages: hardware and firmware (HW/FW) platform development, model compilation, and application development.

*Email: nguyenngocanh@lqdtu.edu.vn

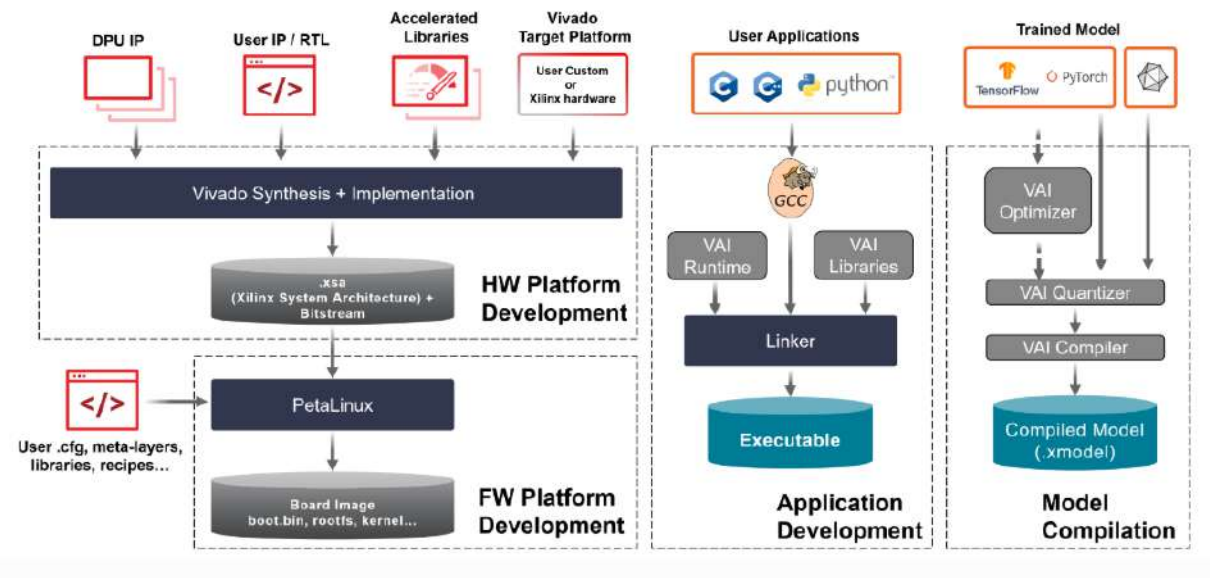


Figure 2. Vivado workflow for developing applications using the Xilinx Vitis AI platform

3.1. Hardware and firmware platform development

The hardware platform was built in Vivado by integrating the Zynq UltraScale+ MPSoC processing system, the DPU IP, AXI interconnects, clocking modules, reset controllers, and required GPIO/control interfaces, as shown in Figure 3. The processing system (PS), based on the ARM Cortex-A53, was used for system control and software execution, while the programmable logic (PL) was used to host the DPU for CNN acceleration. In this work, a DPU configuration with a maximum operand size of 4096 was selected, and the DPU clock frequency was set to 300 MHz through the Clocking Wizard. After block design validation, Vivado synthesis and implementation were executed to generate the hardware bitstream and the XSA file, which were then exported to PetaLinux for firmware generation. The hardware block design in Vivado is shown in Figure 3.

3.2. Model quantization and compilation

To address FPGA resource constraints and meet real-time inference demands, the trained FP32 model underwent quantization using the Vitis AI Quantizer within a dedicated Linux Docker environment. This process utilized Post-Training Quantization (PTQ) to convert model weights and activations from 32-bit floating-point to 8-bit integer (INT8) representations. Specifically, model calibration was performed using a representative subset of approximately 1,000 training images to minimize precision loss during the bit-width reduction.

Following successful quantization and accuracy validation, the model was compiled into an XMODEL file - the executable format for the FPGA. The compilation process utilizes a specific DPU fingerprint via an arch.json file. This file informs the Vitis AI Compiler about the hardware capabilities of the target DPU architecture, ensuring the neural network graph is optimally mapped to the DPU specialized instruction set.

*Email: nguyenngocanh@lqdtu.edu.vn

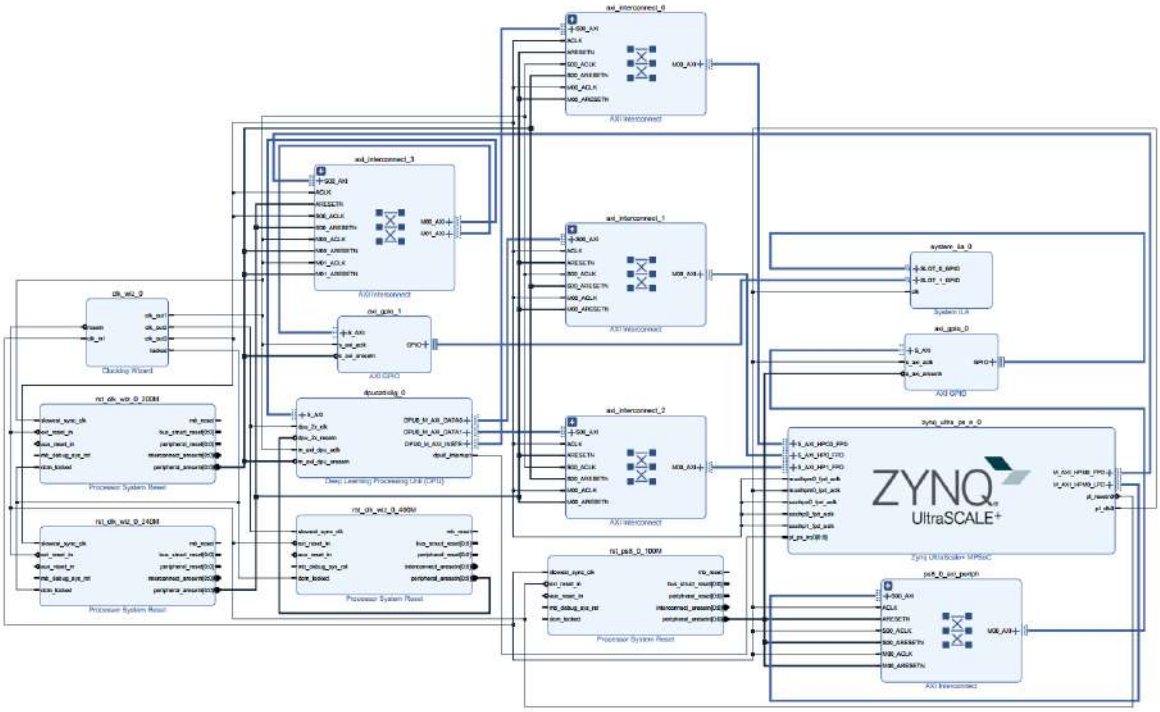


Figure 3. The hardware block design in Vivado

3.3. Application development

Finally, the compiled XMODEL was deployed on the MPSoC target board together with the inference application developed using the Vitis AI Runtime (VART). During execution, thermal images were first preprocessed on the PS side and then transferred to the DPU for accelerated inference. The convolution-dominant layers of YOLOv5 were executed on the DPU, while the final post-processing steps, including output decoding and non-maximum suppression, were handled by the ARM processor. This hardware-software partition enables efficient real-time inference while keeping the deployment flow practical for embedded thermal vision applications.

4. Results and discussion

4.1. Datasets and training

The experiments in this paper are implemented on the FLIR-ADAS v2 dataset. FLIR-ADAS v2 is a medium-scale image set annotated for object detection tasks in both thermal and visible bands. The set includes a total of 26,442 images with 15 different object classes. A total of 9,711 thermal and 9,233 RGB training/validation images are included with a 90%-10% train/validation split. In this study, we use only the thermal images of the set, which are acquired with a Teledyne FLIR Tau 2. The resolution of the thermal images is 640×512 pixels. In our experiments, we utilize only the car and the person classes because the number of samples from the other classes is insufficient for domain adaptation.

In this paper, the model was trained on a laptop with a Ryzen 7 7840H CPU, 16 GB of RAM, and an NVIDIA GeForce RTX 4060 8GB GPU, and a PyTorch instance. The model was trained for 100 epochs with a batch size of 16 and an image input size of 640 × 640 pixels.

*Email: nguyenngocanh@lqdtu.edu.vn

4.2. Evaluation metrics

The proposed model performance was assessed using precision (*P*), recall (*R*), mean average precision (*mAP*). These metrics were selected for their common application in object detection, offering a thorough evaluation of the model's accuracy in detecting and classifying objects [3].

Precision (*P*) indicates the accuracy of positive predictions, recall (*R*) the detection rate of actual positives, defined as:

$$P = \frac{TP}{TP + FP}, \qquad R = \frac{TP}{TP + FN} \tag{1}$$

Here, true positives (*TP*) denote correctly identified positive samples, false positives (*FP*) represent incorrectly identified positive samples, and false negatives (*FN*) are actual positive samples that were not identified by the model.

The average precision (*AP*) is calculated as the area under the precision–recall (*PR*) curve, which reflects the model's detection performance at various thresholds. The mean average precision (*mAP*) is then obtained by averaging the *AP* values across all detection categories. These are computed as follows:

$$AP = \int_a^b P(R)\,dR, \quad mAP = \frac{1}{n}\sum_{i=1}^{n} AP_i \tag{2}$$

In the formula above, $P(R)$ is the precision as a function of recall, a and b are the integration limits usually 0 to 1, AP_i is the *AP* for the i-th class, and n is the total number of classes.

4.3. Experiment result

In this study, the Xilinx UltraScale+ MPSoC ZU4EV-1E development board is used as the experimental platform for evaluating the YOLOv5 algorithm. The detailed hardware specifications of the MPSoC device and the experimental platform are summarized in Table 1.

Table 1. Hardware specifications of the MPSoC module

Module	UltraScale+ ZU4EV-1E	Processor	Quad-Core ARM Cortex-A53
RAM	2GB DDR4 SDRAM		Dual-core ARM Cortex-R5
e.MMC	8 GB	Graphics Processor	Mali-400 MP2
FLASH	128 MB	Logic Cells	192K+
Block RAM	128 blocks; 4.5Mb	DSP Slices	728
Flip-Flops	175K+	LUTs	87K+

Figure 4 presents the post place-and-route resource utilization of the proposed YOLOv5 acceleration platform on the MPSoC device. The results indicate that the DPU-based design heavily utilizes DSP and URAM resources, reaching 97.80% and 95.83%, respectively, which reflects the computational intensity of convolution operations and the reliance on on-chip memory

*Email: nguyenngocanh@lqdtu.edu.vn

for efficient data buffering. In contrast, the utilization of LUT, FF, and BRAM resources remains at moderate levels of approximately 60–66%, demonstrating a balanced allocation of general-purpose logic resources. Despite the high utilization of compute-critical resources, power analysis using the Vitis Analyzer confirms that the proposed design maintains high power efficiency while delivering real-time inference performance on the MPSoC edge platform.

Resource	Utilization	Available	Utilization %
LUT	55777	87840	63.50
LUTRAM	8166	57600	14.18
FF	107537	175680	61.21
BRAM	85	128	66.41
URAM	46	48	95.83
DSP	712	728	97.80
IO	4	252	1.59
MMCM	1	4	25.00

Figure 4. FPGA resource utilization on MPSoC platform

Table 2 compares the proposed YOLOv5 deployment on three platforms: Jetson Orin Nano, NVIDIA RTX 4060, and the MPSoC-based FPGA platform. For consistency, the reported FPS corresponds to the complete execution pipeline, including preprocessing, inference, and post-processing. On the MPSoC platform, the convolution-dominant layers were accelerated on the DPU, while preprocessing and post-processing were executed on the ARM processor. The results show that the RTX 4060 achieved the highest throughput (238 FPS), but with the largest power consumption (29.4 W). The Jetson Orin Nano provided moderate speed and power consumption, reaching 39.84 FPS at 7.2 W. In comparison, the proposed MPSoC platform achieved 48.07 FPS at only 6.1 W, indicating a better balance between real-time performance and energy efficiency. Although the MPSoC result shows some reduction in Accuracy, Recall, and mAP50 due to INT8 quantization, the overall detection performance remains acceptable for embedded thermal vision applications.

Table 2. Comparison of experimental results on different platforms

Metrics	**Jetson Orin Nano**	**NVIDIA RTX 4060**	**MPSoC**
Accuracy	0.814	0.814	0.795
Recall	0.617	0.617	0.575
mAP50	0.723	0.723	0.561
FPS	39.84	238	48.07
Power (W)	7.2	29.4	6.1

Figure 5 presents object detection results obtained on the test set of the FLIR-ADAS v2 thermal dataset using the proposed FPGA-accelerated YOLOv5 system deployed on the MPSoC platform. The results demonstrate that the accelerator is able to reliably detect pedestrians and vehicles under challenging thermal imaging conditions, including low contrast,

*Email: nguyenngocanh@lqdtu.edu.vn

background clutter, and varying object scales. Accurate bounding box localization and stable confidence scores are observed across different road scenes, indicating that the proposed hardware implementation preserves detection robustness while satisfying real-time inference constraints on an embedded MPSoC edge device.

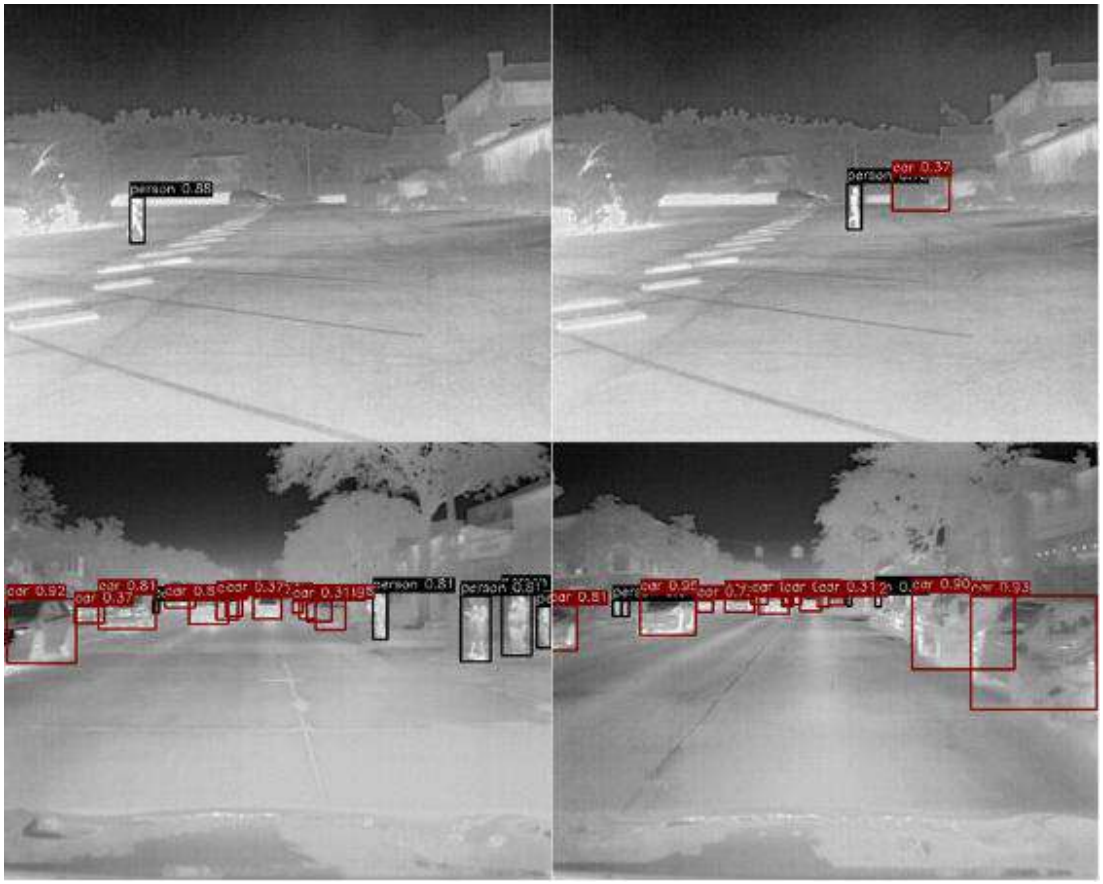


Figure 5. Detection results on the FLIR-ADAS v2 dataset

Figure 6 shows representative object detection results obtained on the MPSoC platform using thermal images captured by an Infrared Focal Plane Array (IRFPA) sensor. The proposed FPGA-accelerated YOLOv5 system successfully detects pedestrians and vehicles in a complex thermal scene, despite challenges such as low thermal contrast, background clutter, and sensor noise inherent to IRFPA-based imaging. As illustrated in the figure, the detected objects are accurately localized, even for small or partially obscured targets at longer distances. These results demonstrate that the proposed hardware-accelerated implementation maintains reliable detection performance on real-world thermal data acquired directly from IRFPA, while satisfying real-time inference requirements on the embedded MPSoC edge device.

*Email: nguyenngocanh@lqdtu.edu.vn

Figure 6. Object detection results captured by IRFPA on the MPSoC platform

5. Conclusion

This paper presented an FPGA-accelerated deployment of YOLOv5 for thermal object detection on a Xilinx MPSoC edge platform. By combining a lightweight YOLOv5 model with DPU-based acceleration, the proposed system achieved real-time performance under strict power and resource constraints. Experimental results showed that the MPSoC implementation reached about 48 FPS with low power consumption, while maintaining acceptable detection accuracy on thermal images despite the minor degradation introduced by INT8 quantization. These results confirm that the proposed FPGA-based solution provides a practical balance between inference speed, accuracy, and energy efficiency for embedded thermal vision applications. In future work, we will improve the system in three directions. First, we will study stronger model compression and lower-bit quantization methods to further reduce hardware cost and improve efficiency on FPGA platforms [4, 15]. Second, we will investigate model improvements for thermal images, such as attention mechanisms, lightweight convolution, and better multi-scale feature extraction, to enhance detection accuracy in difficult scenes [5, 9]. Finally, we will test the proposed method on more edge devices, such as FPGA, NPU, and GPU-based platforms, to better understand the trade-off between speed, accuracy, and power consumption in real applications [1].

References

1. Aramendia, Jon, et al. (2025), "Neural network implementation for fire detection in critical infrastructures: A comparative analysis on embedded edge devices", *Electronics*. 14(9), p. 1809.
2. Bochkovskiy, Alexey, Wang, Chien-Yao, and Liao, Hong-Yuan Mark (2020), "Yolov4: Optimal speed and accuracy of object detection", *arXiv preprint arXiv:2004.10934*.
3. Chen, Hao, et al. (2025), "A robust YOLOv5 model with SE attention and BIFPN for Jishan Jujube detection in complex agricultural environments", *Agriculture*. 15(6), p. 665.

*Email: nguyenngocanh@lqdtu.edu.vn

4. Chen, Yanting and Yu, Lei (2025), FPGA-based MCS-YOLOv5 Lightweight Ship Detection Modeltection on KR260, *2025 5th International Conference on Artificial Intelligence and Industrial Technology Applications (AIITA)*, IEEE, pp. 1764-1768.
5. Cheng, Liang, et al. (2024), "Personnel detection in dark aquatic environments based on infrared thermal imaging technology and an improved YOLOv5s model", *Sensors*. 24(11), p. 3321.
6. He, Kaiming, et al. (2015), "Spatial pyramid pooling in deep convolutional networks for visual recognition", *IEEE transactions on pattern analysis and machine intelligence*. 37(9), pp. 1904-1916.
7. Heng, Chua Wee, et al. (2025), "Optimizing FPGA-based YOLO series accelerators: A survey of techniques", *Neurocomputing*. 650, p. 130874.
8. Khanam, Rahima and Hussain, Muhammad (2024), "What is YOLOv5: A deep look into the internal features of the popular object detector", *arXiv preprint arXiv:2407.20892*.
9. Li, Shasha, et al. (2021), "Yolo-firi: Improved yolov5 for infrared image object detection", *IEEE access*. 9, pp. 141861-141875.
10. Nguyen, Dai-Duong, et al. (2024), "FPGA-SoC implementation of YOLOv4 for flying-object detection", *Journal of Real-Time Image Processing*. 21(3), p. 63.
11. Qiu, Jiantao, et al. (2016), Going deeper with embedded FPGA platform for convolutional neural network, *Proceedings of the 2016 ACM/SIGDA international symposium on field-programmable gate arrays*, pp. 26-35.
12. Redmon, Joseph, et al. (2016), You only look once: Unified, real-time object detection, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779-788.
13. Redmon, Joseph and Farhadi, Ali (2018), "Yolov3: An incremental improvement", *arXiv preprint arXiv:1804.02767*.
14. Saidani, Taoufik, et al. (2024), "Hardware acceleration for object detection using yolov5 deep learning algorithm on xilinx zynq fpga platform", *Engineering, Technology & Applied Science Research*. 14(1), pp. 13066-13071.
15. Yan, Zhihong, Zhang, Bingqian, and Wang, Dong (2024), "An FPGA-Based YOLOv5 Accelerator for Real-Time Industrial Vision Applications", *Micromachines*. 15(9), p. 1164.

Tăng tốc YOLOv5 bằng FPGA cho bài toán phát hiện đối tượng trên ảnh nhiệt theo thời gian thực

Tóm tắt

Phát hiện đối tượng trên ảnh nhiệt theo thời gian thực đóng vai trò quan trọng trong các ứng dụng như giám sát, quân sự và hệ thống an ninh, đặc biệt trong điều kiện ánh sáng yếu và môi trường khắc nghiệt. Mặc dù các mô hình dựa trên học sâu như YOLOv5 đạt được độ chính xác phát hiện cao, việc triển khai chúng trên các nền tảng nhúng gặp nhiều thách thức do hạn chế về tài nguyên tính toán, ràng buộc về công suất tiêu thụ và yêu cầu nghiêm ngặt về độ trễ thời gian thực. Bài báo này trình bày một giải pháp tăng tốc YOLOv5 bằng FPGA trên nền tảng Hệ thống trên Chip đa xử lý (MPSoC) nhằm phục vụ bài toán phát hiện đối tượng ảnh nhiệt theo thời gian thực. Thiết kế được tối ưu thông qua kiến trúc luồng dữ liệu hiệu quả, lượng tử hóa mô hình và tinh chỉnh phần cứng nhằm cân bằng giữa độ chính xác, tốc độ xử lý và mức tiêu thụ năng lượng. Kết quả thực nghiệm cho thấy hệ thống đề xuất đạt hiệu năng thời gian thực ổn định trên các bộ dữ liệu ảnh nhiệt, với mức suy giảm độ chính xác không đáng kể so với phương pháp cài đặt thuần phần mềm, qua đó chứng minh tính khả thi của việc phát hiện đối tượng ảnh nhiệt có độ trễ thấp và hiệu quả năng lượng cao trên các nền tảng MPSoC nhúng.

Keywords*:* MPSoC, VitisAI, YOLOv5, FPGA.

**Email: nguyenngocanh@lqdtu.edu.vn*

Nghiên Cứu Xây Dựng Thuật Toán Phát Hiện Đối Tượng Sử Dụng Mô Hình Học Sâu Ứng Dụng Trong Huấn Luyện Quân Sự

Nguyễn Văn Thịnh, Trịnh Mạnh Tuyên, Trương Xuân Tùng, Nguyễn Văn Xuân*
Viện Tên lửa và KTĐK; Học viện Kỹ thuật quân sự;
Email: vanthinhmta68@gmail.com; Tel: 0342314129

Tóm tắt

Bài báo này trình bày nghiên cứu xây dựng thuật toán phát hiện đối tượng quân sự là bia bắn súng AK bài 1 dựa trên mô hình học sâu YOLOv8. Mô hình được huấn luyện trên tập dữ liệu thu thập trong nhiều điều kiện thời tiết và ánh sáng khác nhau nhằm nâng cao tính tổng quát và khả năng thích nghi của hệ thống. Kết quả thực nghiệm cho thấy thuật toán đề xuất đạt độ chính xác phát hiện cao, duy trì hiệu năng ổn định trong các điều kiện thiếu sáng và chói sáng, vượt trội so với các phương pháp xử lý ảnh truyền thống. Nghiên cứu khẳng định tiềm năng ứng dụng của các mô hình học sâu hiện đại nhằm phát hiện mục tiêu quân sự trong môi trường thực địa phức tạp, góp phần nâng cao mức độ tự động hóa và độ tin cậy của hệ thống huấn luyện bắn súng.

Từ khóa: *Đối tượng quân sự; Phát hiện đối tượng quân sự với mô hình học sâu YOLOv8; Phát hiện bia bắn súng AK.*

1. Đặt vấn đề

Trong bối cảnh cuộc Cách mạng công nghiệp 4.0, trí tuệ nhân tạo đang ngày càng được ứng dụng rộng rãi trong nhiều lĩnh vực huấn luyện quân sự. Huấn luyện bắn súng là một nội dung cốt lõi trong chương trình huấn luyện quân sự. Vì vậy, việc nghiên cứu và áp dụng trí tuệ nhân tạo, đặc biệt là các kỹ thuật học sâu, vào huấn luyện bắn súng là hết sức cần thiết nhằm nâng cao hiệu quả, độ chính xác và chất lượng đánh giá trong quá trình huấn luyện.

Hiện nay, một số hệ thống hỗ trợ chấm điểm và giám sát bia bắn đã ứng dụng kỹ thuật xử lý ảnh [1] truyền thống dựa trên các đặc trưng thủ công như ngưỡng hóa, phát hiện biên, trích chọn đặc trưng hình học hoặc so khớp mẫu. Tuy nhiên, các phương pháp này thường phụ thuộc mạnh vào điều kiện chiếu sáng và môi trường. Trong thực tế triển khai, các loại bia bắn súng được sử dụng tại nhiều vùng miền với điều kiện khí hậu khác nhau, bao gồm môi trường nắng gắt gây chói sáng, điều kiện thiếu sáng vào sáng sớm hoặc chiều muộn, thậm chí trong thời tiết nhiều mây, sương mù hoặc mưa nhẹ. Những yếu tố này làm suy giảm chất lượng ảnh thu nhận, dẫn đến việc giảm độ chính xác phát hiện khi áp dụng các thuật toán truyền thống.

Sự phát triển mạnh mẽ của học sâu (Deep Learning) [2], đặc biệt là các mô hình phát hiện đối tượng theo thời gian thực, đã mở ra hướng tiếp cận mới cho bài toán nhận dạng mục tiêu trong môi trường phức tạp. Trong số đó, mô hình YOLOv8 [3] thuộc họ mô hình "You Only Look Once" cho phép phát hiện đối tượng với tốc độ cao và độ chính xác vượt trội nhờ khả năng học đặc trưng tự động từ dữ liệu lớn và đa dạng.

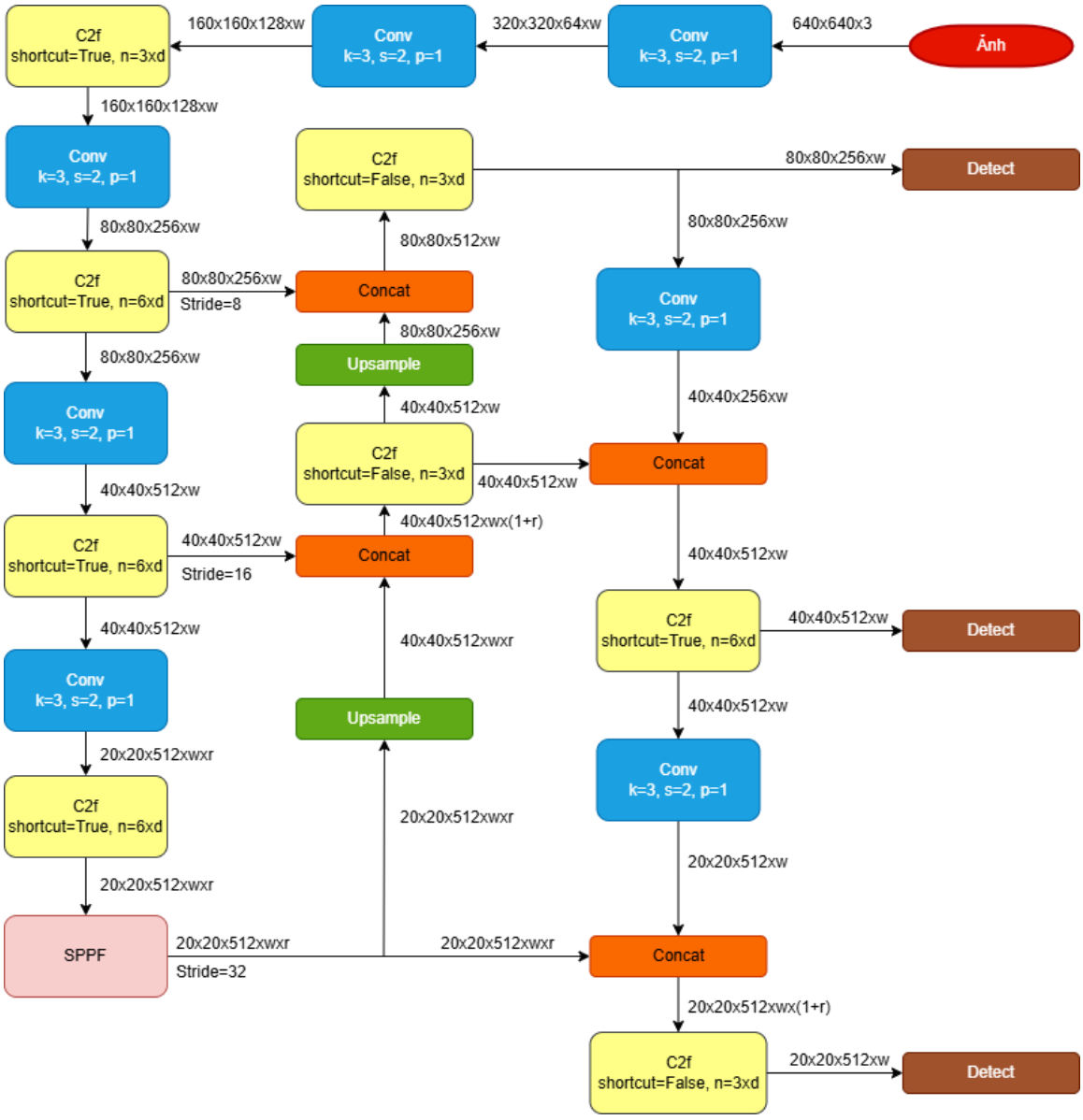
Việc ứng dụng YOLOv8 vào bài toán phát hiện các loại bia bắn súng có tiềm năng khắc phục những hạn chế của phương pháp truyền thống, đồng thời đáp ứng yêu cầu hoạt động ổn định trong nhiều điều kiện thời tiết và ánh sáng khác nhau.

Xuất phát từ yêu cầu thực tiễn đó, bài báo tập trung nghiên cứu xây dựng thuật toán phát hiện bia bắn súng AK bài 1 dựa trên mô hình YOLOv8, đánh giá hiệu năng thông qua các thí nghiệm thực tế trong điều kiện môi trường đa dạng.

2. Các nội dung liên quan

2.1. Mô hình huấn luyện YOLOv8

YOLOv8 [4] là phiên bản mới trong dòng mô hình YOLO (You Only Look Once) nổi tiếng, được phát triển để tiếp tục nâng cao hiệu suất phát hiện đối tượng cả về tốc độ và độ chính xác. YOLOv8 mang đến những cải tiến vượt trội so với các phiên bản trước đó, với các kỹ thuật tối ưu hóa hiện đại để phù hợp cho các ứng dụng trong nhiều lĩnh vực.



Hình 1. Kiến trúc của YOLOv8

Mô hình YOLOv8 vẫn dựa trên cấu trúc one-stage detection giống các phiên bản trước đó, nhưng có nhiều cải tiến về chi tiết:

Phần backbone của YOLOv8 sử dụng CSPDarknet, một biến thể của mạng Darknet tích hợp cơ chế Cross Stage Partial (CSP). CSPDarknet chia nhỏ các tensor đầu vào, xử lý qua các lớp tích chập riêng biệt rồi hợp nhất lại, giúp giảm chi phí tính toán mà vẫn duy trì khả năng trích xuất đặc trưng mạnh mẽ. Kiến trúc này cải thiện luồng gradient, hạn chế hiện tượng trùng lặp thông tin và tăng tốc độ hội tụ trong quá trình huấn luyện.

Neck của YOLOv8 kết hợp PANet (Path Aggregation Network) để tổng hợp đặc trưng đa tầng. PANet cho phép mô hình kết hợp thông tin từ các lớp sâu (chứa ngữ cảnh tổng quan) và lớp nông (chứa chi tiết cục bộ), tạo ra các bản đồ đặc trưng giàu thông tin.

Phần head của YOLOv8 chuyển sang thiết kế anchor-free, loại bỏ sự phụ thuộc vào anchor boxes truyền thống. Thay vì dự đoán các Bounding box dựa trên anchor có sẵn, YOLOv8 trực tiếp ước lượng tọa độ tâm và kích thước đối tượng. Cách tiếp cận này giảm độ phức tạp của mô hình, hạn chế việc điều chỉnh siêu tham số và tăng tính linh hoạt khi làm việc với các tập dữ liệu đa dạng.

Cơ chế này thay thế phương pháp tĩnh truyền thống, giúp lựa chọn mẫu tích cực hiệu quả hơn. Khả năng mở rộng là điểm nhấn quan trọng của YOLOv8 [7]. Điều này cho phép triển khai trên nhiều nền tảng phần cứng, từ thiết bị nhúng, thiết bị di động đến máy chủ hiệu năng cao.

2.2. Các loại bia bắn trong huấn luyện bắn súng AK bài 1

Bia bắn súng AK bài 1 bao gồm bia số 4, bia số 7, bia số 8 là loại bia cơ bản được sử dụng phổ biến trong huấn luyện bắn súng bộ binh ngoài trời.

Các loại bia bắn này được sử dụng ở nhiều thời điểm khác nhau trong ngày và tại nhiều khu vực địa lý có điều kiện thời tiết khác nhau.

Hình 2. Các loại bia bắn súng AK bài 1

Bia số 4 được sử dụng để huấn luyện động tác nằm bắn có tỳ ở khoảng cách 100m. Về kích thước, bia số 4 có chiều cao 42 cm, chiều rộng 42 cm. Trong quá trình huấn luyện độ cao tâm bia được đặt cách mặt đất khoảng từ 0,5m đến 1,4 m.

Bia số 7 được sử dụng để huấn luyện động tác quỳ bắn không tỳ ở khoảng cách 100m. Về kích thước, bia số 7 có chiều cao 83 cm, chiều rộng 40 cm. Trong quá trình huấn luyện độ cao tâm bia được đặt cách mặt đất khoảng từ 1m đến 1,4 m.

Bia số 8 được sử dụng để huấn luyện động tác đứng bắn không tỳ ở khoảng cách 100m. Về kích thước, bia số 8 có chiều cao 148 cm, chiều rộng 40 cm. Trong quá trình huấn luyện độ cao tâm bia được đặt cách mặt đất khoảng 1,4 m.

3. Xây dựng hệ thống huấn luyện bắn súng sử dụng YOLOv8

3.1. Sơ đồ khối của hệ thống

Hệ thống thị giác máy tính trong huấn luyện bắn súng AK bài 1 có các khối chức năng chính như hình 3.

Hình 3. Hệ thống huấn luyện bắn súng AK bài 1

Trong sơ đồ hệ thống trên, các bia bắn AK bài 1 gồm bia số 4, bia số 7, bia số 8 được đặt ở cự ly 100m so với súng huấn luyện có gắn Camera.

Người huấn luyện sử dụng súng có gắn Camera để tiến hành ngắm bắn vào mục tiêu bia theo bài bắn. Cụm Camera được lựa chọn có thông số kỹ thuật như sau:

Bảng 1. Thông số kỹ thuật của cụm Camera

TT	Tên	Đơn vị	Yêu cầu	Ghi chú
1	Độ phân giải của Camera	Pixel	Tối thiểu HD (1280x720 Pixel)	
2	Tốc độ ảnh	Hình/giây	Không nhỏ hơn 25	

Khi thực hiện phát bắn, Camera chụp lại được hình ảnh ngắm bắn thực tế và gửi dữ liệu về máy tính. Máy tính sẽ tiến hành thu nhận hình ảnh chụp được và tiến hành sử dụng mô hình học sâu [8] để phát hiện đối tượng từ đó phân tích và đánh giá kết quả ngắm bắn.

Hình 4. Hệ thống thị giác máy tính trong huấn luyện bắn súng AK bài 1

3.2. Xây dựng bộ dữ liệu và huấn luyện mô hình

3.2.1. Xây dựng bộ dữ liệu

Bộ dữ liệu huấn luyện được thu thập bằng cách chụp ảnh bộ bia bắn gồm bia số 4, bia số 7, bia số 8 tại khoảng cách 100m ở các vị trí khác nhau và trong nhiều điều kiện thời tiết khác nhau như chói sáng, mưa phùn, sương mù v.v.

Nhóm tác giả thu thập được bộ dữ liệu gồm 2191 hình ảnh trong đó có 730 hình ảnh bia số 4, 730 hình ảnh bia số 7, 731 hình ảnh bia số 8. Tất cả hình ảnh trong bộ dữ liệu được quy chuẩn về kích thước 640x640 để phù hợp với đầu vào của mô hình huấn luyện.

Bộ dữ liệu được chia thành 3 file nhỏ gồm: Train, Valid, Test với tỉ lệ 75:17:8. Tệp Train chứa 1640 hình ảnh, tệp Valid chứa 367 hình ảnh và tệp Test chứa 184 hình ảnh. Trong đó:

- Tập Train được sử dụng để mô hình học đặc trưng của các đối tượng bia bắn AK bài 1.

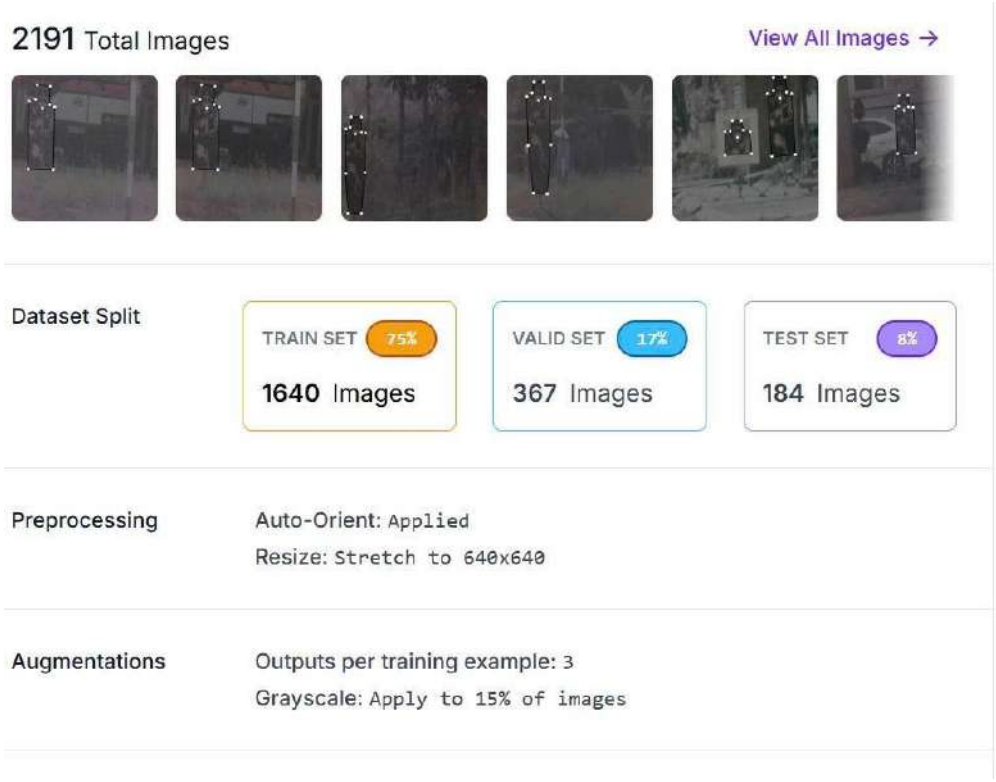
- Tập Valid được sử dụng trong quá trình huấn luyện để theo dõi hiệu năng mô hình và lựa chọn trọng số tốt nhất.

- Tập Test dùng để đánh giá độ chính xác của mô hình.

3.2.2. Gán nhãn cho dữ liệu

Đây là một khâu quan trọng quyết định trực tiếp đến chất lượng của mô hình học máy. Ở giai đoạn này, các dữ liệu thu thập được sẽ gán nhãn theo khung bia bắn.

Hiện nay có rất nhiều phương pháp để gãn nhãn cho đối tượng: offline cũng như online tuy nhiên trong bài báo này tôi sử dụng trang web app.roboflow.com để tiến hành gán nhãn cho các đối tượng, ưu điểm của phương pháp này là giao diện dễ dùng, cũng như dễ dàng trong việc quản lý gán nhãn các đối tượng.



Hình 5. Gán nhãn bộ dữ liệu bia bắn súng AK bài 1

Sau khi đã hoàn tất gán nhãn, bộ dữ liệu bia bắn súng AK bài 1 đã sẵn sàng được sử dụng làm đầu vào để bắt đầu quá trình huấn luyện mô hình.

3.2.3. Huấn luyện mô hình

Việc huấn luyện mô hình được thực hiện trên thiết bị phần cứng máy tính Core(TM) i9-11950H có bộ nhớ GPU 16GB hoạt động trên Win11 – 64 bit.

Chuẩn bị trình biên dịch ngôn ngữ Python 3.10.14 để huấn luyện với mô hình YOLOv8.

Mô hình được triển khai huấn luyện theo bảng sau:

Bảng 2. Mô hình huấn luyện YOLOv8n

```
from ultralytics import YOLO
# Load a model
model = YOLO("yolov8n.pt")
# Train the model
results = model.train(data="coco8.yaml", epochs=235, imgsz=640)
```

3.2.4. Kết quả huấn luyện mô hình YOLOv8

Sau khi hoàn thành quá trình huấn luyện, hệ thống lưu lại hai tệp trọng số là last.pt và best.pt.

Hình 6. Kết quả huấn luyện mô hình YOLOv8n

Kết quả huấn luyện mô hình cho thấy các hàm mất mát (loss) đều giảm rõ rệt theo số epoch, chứng tỏ mô hình đã học được đặc trưng của dữ liệu và dần hội tụ. Cụ thể, Box Loss giảm nhanh trong giai đoạn đầu và ổn định ở mức thấp, thể hiện khả năng định vị đối tượng ngày càng chính xác. Class Loss cũng giảm mạnh và duy trì ổn định, cho thấy mô hình phân loại các đối tượng tốt hơn theo thời gian. Tương tự, Object Loss giảm đều và đạt giá trị nhỏ, phản ánh khả năng phát hiện sự tồn tại của đối tượng trong ảnh được cải thiện đáng kể.

3.3. Cách tính điểm huấn luyện bắn súng

Hệ thống huấn luyện bắn súng sử dụng camera và xử lý ảnh thực hiện việc tính điểm dựa trên việc xác định vị trí điểm ngắm của xạ thủ tại thời điểm bóp cò so với vị trí các vùng mục tiêu trên bia. Khi xạ thủ thực hiện thao tác bắn, tín hiệu kích hoạt (trigger) sẽ đồng bộ với camera để ghi lại một khung hình tại thời điểm đó.

Hình 7. Cách tính điểm huấn luyện bắn súng AK bài 1

Thông qua các thuật toán xử lý ảnh, hệ thống xác định vị trí, kích thước và hình dạng của bia trong ảnh. Sau đó, bằng các phép biến đổi hình học (như biến đổi phối cảnh), ảnh bia được chuẩn hóa về một hệ tọa độ chuẩn nhằm đảm bảo tính chính xác khi tính toán.

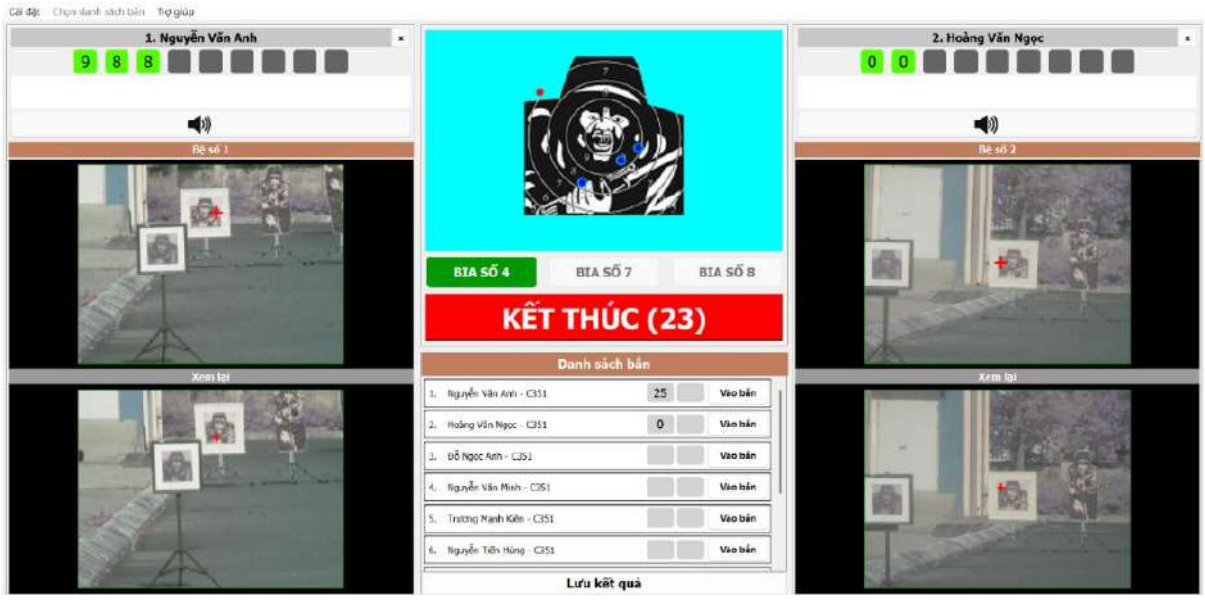
Tiếp theo, hệ thống xác định vị trí tương đối của điểm bắn so với bia. Dựa trên cấu trúc của bia (các vòng tròn đồng tâm hoặc các vùng mục tiêu như đầu, ngực, thân), thuật toán sẽ kiểm tra điểm bắn thuộc vùng nào. Cuối cùng, điểm số tương ứng với vùng trúng đích sẽ được xuất ra và hiển thị cho người sử dụng.

3.4. Kết quả thử nghiệm hệ thống thực tế

Sau quá trình huấn luyện thu được kết quả phát hiện tốt bia số 4, bia số 7, bia số 8 trong bộ bia bắn súng AK bài 1 với điều kiện thời tiết sương mù và cảnh nền bị nhiều nhiễu tác động.

Hình 8. Kết quả nhận dạng bia số 4, 7, 8

Ngoài ra, nhóm tác giả đã triển khai thành công việc ứng dụng các kết quả nghiên cứu vào thực tiễn huấn luyện bắn súng, qua đó chứng minh hệ thống có khả năng xác định chính xác vị trí và suy ra điểm số với độ tin cậy cao, hoạt động ổn định trong nhiều điều kiện môi trường khác nhau, góp phần nâng cao hiệu quả và chất lượng công tác huấn luyện.



Hình 9. Phần mềm huấn luyện bắn súng AK bài 1

4. Kết luận

Bài báo đã trình bày quá trình nghiên cứu và xây dựng thuật toán phát hiện bia bắn súng AK bài 1 dựa trên mô hình học sâu (Deep Learning) [5] YOLOv8 nhằm nâng cao độ chính xác và tính ổn định của hệ thống thị giác máy tính trong huấn luyện bắn súng bộ binh. Xuất phát từ những hạn chế của các phương pháp xử lý ảnh truyền thống khi làm việc trong điều kiện môi trường phức tạp, nghiên cứu đã đề xuất hướng tiếp cận sử dụng mô hình phát hiện đối tượng hiện đại có khả năng học đặc trưng tự động từ dữ liệu.

Kết quả thực nghiệm cho thấy mô hình đề xuất có khả năng phát hiện chính xác các loại bia trong nhiều điều kiện thời tiết và môi trường khác nhau, đồng thời đáp ứng yêu cầu xử lý gần thời gian thực của hệ thống huấn luyện. Ngoài ra nghiên cứu cũng đã thành công trong việc tích hợp chặt chẽ với mô-đun chấm điểm tự động, xây dựng hệ thống huấn luyện bắn súng thông minh và hoàn chỉnh. Những kết quả đạt được khẳng định tính khả thi của việc ứng dụng mô hình học sâu trong bài toán phát hiện bia bắn súng AK bài 1, góp phần nâng cao chất lượng huấn luyện. Trong thời gian tới nghiên cứu mở rộng theo hướng tối ưu mô hình để triển khai trên nền tảng phần cứng nhúng.

Tài liệu tham khảo

1. Phạm Trung Dũng. (2011). "*Giáo trình xử lý và phân tích ảnh trong hệ thống thị giác"*. Nhà xuất bản quân đội nhân dân.

2. Ajay Shrestha, Ausif Mahmood. (2019). "*Review of Deep Learning Algorithms and Architectures"*. IEEE. doi: 10.1109/ACCESS.2019.2912200.

3. Muhammad Yaseen. (2024). "*What is YOLOv8: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector*". ArXiv:2408.15857v1.

4. Rejin Varghese, Sambath M. (2024). "*YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness"*. IEEE. doi: 10.1109/ADICS58448.2024.10533619.

5. Serhat Tural, Refik Samet, Semra Aydin, Mohamed Traore. (2022). "*Deep Learning Based Classification of Military Cartridge Cases and Defect Segmentation"*. IEEE. doi: 10.1109/ACCESS.2022.3191328.

Research on Developing an Object Detection Algorithm for AK Firing Target No.1 Using Deep Learning

Abstract: *This paper presents a study on developing an object detection algorithm for AK Firing Target No.1 within a computer vision system for infantry shooting training. Since the target is deployed in various geographic regions under different weather and lighting conditions, the system is required to maintain high detection accuracy in low light, glare, and complex background environments. The proposed approach applies the YOLOv8 deep learning model to improve detection performance compared to traditional image processing methods. A diverse dataset was constructed under multiple real-world environmental conditions to enhance the generalization capability of the model. Experimental results demonstrate that the proposed algorithm achieves high accuracy and stable performance under harsh weather conditions while satisfying near real-time processing requirements. The research contributes to improving automation and reliability in shooting training support systems.*

Keywords: Object detection; AK Firing Target No.1; Computer vision; Deep learning; YOLOv8

Xây dựng mô hình nhận dạng rối loạn nhịp tim sử dụng mạng học sâu

Vương Trí Tiếp[1], Nguyễn Trọng Tuyến[1*], Nguyễn Mạnh Cường[1*]

[1]*Viện Tên lửa và kỹ thuật điều khiển, Học viện Kỹ thuật quân sự*

Email: tuyen.nguyen@lqdtu.edu.vn, nmcuong869@lqdtu.edu.vn; Tel: 0967904423

Tóm tắt

Rối loạn nhịp tim là một trong những nguyên nhân hàng đầu gây ra các biến cố tim mạch nguy hiểm. Việc tự động hóa chẩn đoán qua tín hiệu điện tâm đồ (ECG) gặp nhiều thách thức do sự biến thiên liên bệnh nhân, mất cân bằng dữ liệu và sự tương đồng hình thái vi mô giữa các loại nhịp. Nghiên cứu này đề xuất mô hình HybridECGNet, một kiến trúc mạng nơ-ron học sâu lai ghép trích xuất đặc trưng đa chiều. Hệ thống sử dụng mạng tích chập phân tách theo chiều sâu đa thang (Multi-scale DSWC) kết hợp Transformer để khai thác hình thái không gian, và mạng bộ nhớ dài ngắn hạn hai chiều (BiLSTM) để lập mô hình động học chuỗi thời gian. Một cơ chế cổng kết hợp động (Gated Fusion) được thiết kế nhằm tự động điều phối trọng số giữa hai nhánh tùy thuộc vào đặc thù bệnh lý. Thực nghiệm trên cơ sở dữ liệu MIT-BIH Arrhythmia tuân thủ nghiêm ngặt giao thức kiểm chứng liên bệnh nhân (Inter-patient paradigm) cho thấy mô hình đạt độ chính xác toàn cục 98.88%. Đặc biệt, khả năng phát hiện các lớp ngoại tâm thu nguy hiểm đạt độ nhạy trên 97%, chứng minh tính khả thi của mô hình trong việc triển khai trên các thiết bị theo dõi y tế thời gian thực.

Từ khóa: *rối loạn nhịp tim; mạng học sâu; kiểm chứng liên bệnh nhân.*

1. Đặt vấn đề

Rối loạn nhịp tim (Arrhythmia) là một nhóm các bệnh lý tim mạch phức tạp, đặc trưng bởi sự thay đổi bất thường trong hình thái, tần số hoặc trật tự khởi phát xung động điện của tim. Những rối loạn này phản ánh sự suy giảm chức năng của hệ thống dẫn truyền điện sinh học nội tại của cơ tim, có khả năng tiến triển thành các biến cố huyết động nghiêm trọng như suy tim cấp, nhồi máu cơ tim, đột quỵ, hay thậm chí là đột tử do tim. Theo Tổ chức Y tế Thế giới (WHO), các bệnh lý tim mạch đóng góp vào khoảng 17,9 triệu ca tử vong mỗi năm trên toàn cầu. Trong bối cảnh tỷ lệ mắc bệnh gia tăng theo cấp số nhân do quá trình già hóa dân số, việc phát hiện sớm và quản lý liên tục các dạng rối loạn nhịp tim thông qua Điện tâm đồ (ECG) đóng vai trò sống còn [1].

Tín hiệu ECG ghi lại sự chênh lệch điện thế trên bề mặt da sinh ra từ quá trình khử cực và tái cực của cơ tim. Một chu kỳ bình thường tạo ra các thành phần sóng P, phức bộ QRS và sóng T [2]. Nhằm chuẩn hóa quá trình đánh giá các thuật toán phân tích tự động, Hiệp hội vì sự Tiến bộ của Thiết bị Y tế Hoa Kỳ (AAMI) đã thiết lập tiêu chuẩn EC57, phân loại nhịp tim thành 5 siêu lớp: Nhịp bình thường (N), Ngoại tâm thu trên thất (S), Ngoại tâm thu thất (V), Nhịp hợp nhất (F), và Nhịp không xác định (Q). Việc phân biệt chính xác giữa các nhóm này mang hàm ý lâm sàng sâu sắc. Nhịp V (Ventricular ectopic beats) xuất phát từ tâm thất, tạo ra phức bộ QRS dãn rộng và biến dạng; sự xuất hiện thường xuyên của nhịp V là dấu hiệu cảnh báo đỏ cho nguy cơ rung thất và đột tử. Ngược lại, nhịp S (Supraventricular ectopic beats) khởi phát từ tâm nhĩ, có QRS thanh mảnh tương tự nhịp bình thường, nhưng việc gia tăng gánh nặng nhịp S lại tương quan mạnh mẽ với sự khởi phát rung nhĩ và nguy cơ đột quỵ nhồi máu não. Việc chẩn đoán nhầm lẫn giữa S và V có thể dẫn đến các chiến lược can thiệp y tế sai lầm.

Tự động hóa phân loại ECG bằng máy tính phải đối mặt với ba thách thức cốt lõi. Thứ nhất, sự biến thiên liên bệnh nhân (inter-patient variability) khiến một mô hình học sâu dễ dàng ghi nhớ đặc trưng cá nhân hóa thay vì học quy luật bệnh lý phổ quát, dẫn đến suy giảm hiệu

năng trên bệnh nhân mới. Thứ hai, dữ liệu y tế mất cân bằng nghiêm trọng (class imbalance) khi nhịp bình thường chiếm trên 80%, gây thiên vị cho các thuật toán tối ưu hóa. Thứ ba, sự tương đồng hình thái vi mô khiến ranh giới phân loại trở nên mờ nhạt (ví dụ nhịp S và N có QRS giống hệt nhau, chỉ khác biệt ở sóng P siêu nhỏ rất dễ bị nhiễu vùi lấp; trong khi nhịp F là sự hòa trộn quang phổ vô tận giữa N và V).

Để giải quyết, các phương pháp trích xuất đặc trưng thủ công kết hợp SVM hay MLP đã được thay thế bằng Mạng nơ-ron tích chập (CNN). CNN có sức mạnh trích xuất đặc trưng cục bộ (local features) xuất sắc nhưng bị giới hạn bởi trường nhìn nhận, không thể nắm bắt biến thiên nhịp điệu kéo dài. Mạng nơ-ron hồi quy (LSTM) có thể lập mô hình chuỗi thời gian nhưng lại chịu chi phí tính toán cao và tiêu biến đạo hàm. Gần đây, kiến trúc Transformer với cơ chế Tự chú ý (Self-Attention) mang đến khả năng thiết lập liên kết toàn cục, nhưng lại thiếu "thiên kiến quy nạp" của CNN khiến chúng dễ quá khớp trên dữ liệu y tế.

Nhận thức được những giới hạn này, nghiên cứu đề xuất kiến trúc học sâu lai ghép mang tên HybridECGNetV2. Hệ thống tích hợp Mạng tích chập đa thang phân tách theo chiều sâu (Multi-scale DSWC), Transformer mã hóa không gian và Mạng bộ nhớ dài ngắn hạn hai chiều (BiLSTM) xử lý chuỗi. Thông qua cơ chế Cổng kết hợp động (Gated Fusion Mechanism), mô hình tự động điều tiết tỷ trọng thông tin từ các nhánh, mô phỏng quá trình hội chẩn đa chiều của các chuyên gia y khoa, qua đó giải quyết triệt để bài toán biến thiên liên bệnh nhân.

2. Phương pháp

Kiến trúc HybridECGNet xử lý tín hiệu ECG một kênh thông qua luồng tiền xử lý chuẩn hóa và hai nhánh trích xuất đặc trưng hoạt động song song (Hình 1). Sự phân tách này đảm bảo tín hiệu được phân tích toàn diện từ hình thái vi mô đến động học vĩ mô.

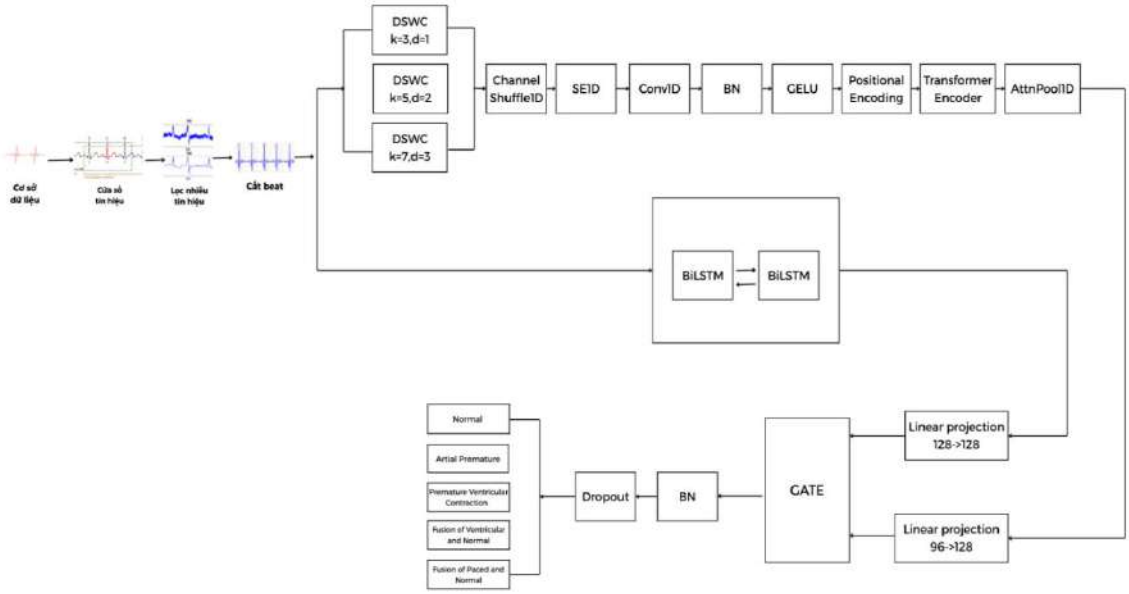
2.1. Tiền xử lý và chuẩn hóa

Tín hiệu thô từ cơ sở dữ liệu MIT-BIH (360 Hz) được giảm tần số lấy mẫu về 125 Hz nhằm tối ưu hóa chi phí tính toán cho thiết bị nhúng mà vẫn duy trì băng thông sinh lý lâm sàng cần thiết.[1] Thuật toán tự động phát hiện đỉnh R được áp dụng để cắt các cửa sổ động thích ứng. Cụ thể, thuật toán tính khoảng cách RR trung vị (T). Đối với mỗi đỉnh R, hệ thống trích xuất phân đoạn từ $0.2T$ trước đỉnh R đến $1.0T$ sau đỉnh R, bao trùm trọn vẹn phức hợp P-QRS-T. Các phân đoạn được nội suy tuyến tính để ép về vector chuẩn hóa độ dài tĩnh $T = 187$ mẫu. Cuối cùng, biên độ điện thế được chuẩn hóa theo phương pháp Z-score nhằm triệt tiêu nhiễu trôi đường nền và chênh lệch trở kháng.

2.2. Nhánh trích xuất đặc trưng không gian

Nhánh này khai thác các biến thiên hình thái từ vi mô (QRS) đến vĩ mô (khoảng PR, QT) bằng sự kết hợp giữa CNN và Transformer.

Khối Multi-scale Depthwise Separable Convolution (DSWC): Tín hiệu $x \in \mathbb{R}^{B\times C\times T}$ đi qua ba nhánh tích chập phân tách theo chiều sâu hoạt động song song. Phép biến đổi này phân rã tích chập chuẩn thành tích chập chiều sâu (depthwise) và tích chập điểm (pointwise), giảm độ phức tạp từ $O(K \cdot C \cdot C_{out})$ xuống $O(K \cdot C + C \cdot C_{out})$. Ba nhánh được cấu hình tham số (Kích thước kernel K, Hệ số giãn d) lần lượt là (3,1), (5,2) và (7,3). Kernel 3 tập trung dị dạng cục bộ trên đỉnh R; Kernel 5 nắm bắt hình thái toàn vẹn QRS; Kernel 7 bao phủ quan hệ dài từ sóng P đến đoạn ST.



Hình 1. Mô hình mạng học sâu phân loại rối loạn nhịp tim

Khối Channel Shuffle và Squeeze-and-Excitation (SE): Đặc trưng nối ghép từ 3 nhánh (96 kênh) được hoán vị chéo bằng Channel Shuffle để trộn lẫn thông tin đa thang. Sau đó, cơ chế chú ý kênh SE được áp dụng để tự động cân chỉnh trọng số, làm mờ các kênh nhiễu và khuếch đại các kênh bệnh lý.

Embedding Convolution và Transformer Encoder: Một lớp Conv1D với $stride = 2$ thực hiện nén thời gian, giảm độ dài chuỗi từ T xuống $\frac{T}{2}$ để tối ưu bộ nhớ cho Transformer. Positional Encoding bằng chuỗi hàm dao động lượng giác được cộng vào nhằm cung cấp thông tin vị trí cho cơ chế chú ý. Khối Transformer Encoder ($d_{model} = 96$, 4 Multi-Head Self-Attention) tính toán các ma trận Query (Q), Key (K), Value (V):

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

MHSA cho phép mạng lập mô hình các phụ thuộc dài hạn cực kỳ phức tạp xuyên suốt nhịp tim. Cơ chế Attention Pooling cuối cùng trích xuất vector ngữ cảnh toàn cục tĩnh.[1]

2.3. Nhánh trích xuất đặc trưng thời gian

Để mô hình hóa dòng chảy liên tục và khoảng nghỉ bù (compensatory pause) của các sự kiện loạn nhịp, tín hiệu đi qua nhánh BiLSTM.[1] Tín hiệu 187 mẫu được lọc bằng lớp Conv1D ($K = 5, stride = 2$) làm bộ lọc thông dải khử nhiễu tần số thấp, sau đó nạp vào mạng BiLSTM 3 tầng (hidden size = 64). Cấu trúc cổng điều khiển (gating mechanism) của LSTM kiểm soát dòng thông tin hai chiều cả bối cảnh quá khứ lẫn tương lai. Trạng thái cuối của chuỗi được rút trích thành vector 128 chiều chứa đựng động học thời gian của nhịp tim.[1]

2.4. Khối hợp nhất động và Phân loại

Hai vector không gian (v_{tr}) và thời gian (v_{lstm}) được đồng bộ hóa về không gian tiềm ẩn 128 chiều (vector a và b). Cơ chế Gated Fusion sử dụng một mạng nơ-ron đánh giá động để phân bổ trọng số:

$$g_{logits} = W_2\delta(W_1[a;b] + b_1) + b_2 \quad (3)$$

$$[g_{tr}, g_{lstm}] = \text{softmax}(g_{logits}) \quad (4)$$

$$f_{used} = g_{tr} \cdot a + g_{lstm} \cdot b \quad (5)$$

Mạng tự động ưu tiên nhánh không gian ($g_{tr} > g_{lstm}$) khi gặp biến dạng cấu trúc nhọn (nhịp V), hoặc ưu tiên thời gian ($g_{tr} < g_{lstm}$) khi gặp nhiễu động nhịp điệu (nhịp S). Vector hợp nhất đi qua Head Phân loại (Batch Normalization, Dropout, Linear, GELU) để dự đoán 5 lớp bệnh lý.[1] Hàm mất mát Label Smoothing Cross-Entropy ($\epsilon = 0.05$) được sử dụng để chống sụp đổ phân loại trên dữ liệu thiểu số.

3. Thiết lập thí nghiệm

Nghiên cứu sử dụng cơ sở dữ liệu tiêu chuẩn vàng MIT-BIH Arrhythmia Database, bao gồm 48 bản ghi Holter 30 phút (360 Hz). Bốn bản ghi chứa nhịp tạo nhịp nhân tạo (paced beats: 102, 104, 107, 217) bị loại bỏ nhằm đảm bảo tính trung thực trên mô sinh học.[22] Các nhãn bệnh lý được ánh xạ thành 5 siêu lớp theo chuẩn AAMI EC57: N (Bình thường), S (Trên thất), V (Thất), F (Hợp nhất) và Q (Không xác định).

Để đánh giá khách quan và mô phỏng môi trường lâm sàng thực tế, tránh hiện tượng rò rỉ dữ liệu do ghi nhớ đặc trưng sinh trắc học cá nhân, nghiên cứu tuân thủ nghiêm ngặt giao thức kiểm chứng liên bệnh nhân (Inter-patient Paradigm) của De Chazal (2004).[24] 44 bản ghi được phân chia cứng:

+ Tập DS1 (Training Set): 22 bản ghi dùng để huấn luyện

+ Tập DS2 (Testing Set): 22 bản ghi hoàn toàn khác biệt sinh học dùng để kiểm thử độc lập.

Mô hình được tối ưu bằng bộ thuật toán AdamW, sử dụng suy biến trọng số (Weight Decay $\lambda = 10^{-2}$) để duy trì chuẩn ma trận ở mức an toàn, ngăn chặn over-fitting. Tốc độ học (Learning Rate) η được đặt ở 2×10^{-4}. Hiệu năng được lượng hóa bằng các chỉ số: Overall Accuracy, Sensitivity (Recall), Precision và F1-Score.[1]

4. Kết quả và bàn luận

Kiểm chứng hiệu năng của mô hình HybridECGNet cung cấp minh chứng định lượng về sự ưu việt của kiến trúc lai. Quá trình đánh giá được thực hiện trên mô hình Ensemble theo phương thức nội bệnh nhân (intra-paradigm) và mô hình liên bệnh nhân (Inter-paradigm) (kiểm chứng liên bệnh nhân độc lập DS1/DS2).

Bảng 1. Kết quả thực nghiệm trên mô hình Ensemble (Toàn bộ dữ liệu)

Lớp	Precision (%)	Recall (%)	F1-score (%)
N	99.33	99.53	99.43
S	89.07	86.51	87.77
V	97.76	96.41	97.08
F	82.94	87.04	84.94

Q	99.75	99.32	99.53

Bảng 2. Kết quả thực nghiệm trên mô hình theo phương thức liên bệnh nhân (DS1/DS2)

Lớp	Precision (%)	Recall (%)	F1-score (%)
N	99.35	99.50	99.42
S	82.13	82.35	82.27
V	94.74	97.42	95.11
F	80.92	85.06	82.95
Q	99.74	99.22	99.43

Độ chính xác toàn cục (Overall Accuracy) trên tập kiểm thử liên bệnh nhân (Inter-paradigm) đạt mức xuất sắc 98.88%. Phân tích từ Ma trận nhầm lẫn (Confusion Matrix) cho thấy kiến trúc lai ghép giải quyết triệt để sự mất cân bằng giữa các lớp.

Đối với Ngoại tâm thu thất (Lớp V), hệ thống đạt độ nhạy (Recall) lên tới 97.42% ở môi trường chưa từng học. Sự thành công này phản ánh đúng nền tảng sinh lý học: nhịp V có QRS dãn rộng thô kệch (>120 ms) và khoảng nghỉ bù hoàn toàn. Khối CNN Multi-scale DSWC dễ dàng quét trọn bề rộng bất thường, trong khi khối BiLSTM lùi thời gian để xác nhận khoảng nghỉ bù phía sau. Cổng Gating Mechanism ưu tiên tín hiệu không gian ở đây đã tối đa hóa độ tin cậy, đảm bảo không bỏ lọt các cơn cảnh báo rung thất gây đột tử.

Nghịch lý hình thái được thể hiện rõ ở Lớp S và Lớp F. F1-score của lớp S giảm nhẹ từ 87.77% (Ensemble) xuống 82.27% (Inter-paradigm). Nhịp S có hình thái QRS hoàn toàn giống nhịp N, hệ thống chỉ có thể phân biệt qua nhịp đập sớm và hình thái sóng P siêu nhỏ dễ vùi lấp. Tuy nhiên, việc duy trì F1-score > 82% đối với lớp S và F trên bài toán liên bệnh nhân là một kỷ lục. Các kỹ thuật chống sụp đổ phân loại (Label Smoothing) đã bảo vệ vùng không gian tiềm ẩn của các lớp thiểu số này.

So với các nghiên cứu sử dụng đặt trưng thủ công khoảng 88%) [3,4], và nghiên cứu của Garcia cho 3 lớp sử dụng vectorcardiogram và SVM đạt 92.40%. Một số nghiên cứu sử dụng CNN thuần túy (Sellami, ~88.34%) [5], và mạng tích chập với kết nối dư của một số công trình đạt khoảng 89% [6], HybridECGNet vượt trội tuyệt đối nhờ mô thức nhận thức lai ghép đa chiều: CNN nội suy không gian siêu vi, BiLSTM bảo toàn dòng chảy thời gian, và Transformer tổng hợp quan hệ toàn cục.

5. Kết luận

Nghiên cứu đã xây dựng và kiểm chứng thành công HybridECGNet, một kiến trúc học sâu lai ghép mang tính đột phá cho bài toán phân loại rối loạn nhịp tim từ ECG một kênh. Sự kết hợp các phép trích chọn đặc trưng theo thời gian và không gian qua cơ chế hợp nhất thông minh đã giúp mô hình khắc phục giới hạn trường nhìn hẹp của CNN truyền thống và hiện tượng

tiêu biến chuỗi của LSTM. Hệ thống đạt độ chính xác toàn cục 98.88% dưới giao thức kiểm định liên bệnh nhân nghiêm ngặt MIT-BIH Arrythymia, với độ nhạy 97.42% trong nhận diện nhóm nhịp nguy hiểm V. Đây là bằng chứng thực chứng quan trọng khẳng định tiềm năng đóng gói các kiến trúc lai ghép vào phần mềm nhúng trên thiết bị theo dõi Holter thời gian thực, góp phần hoàn thiện lưới an toàn y khoa toàn cầu.

Tài liệu tham khảo

1. Burns, A. J., & Burns, A. J. (2024). Electrocardiogram. In *StatPearls*. Treasure Island, FL: StatPearls Publishing. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK549803/

2. Antzelevitch, C., & Shimizu, W. (2011). Overview of basic mechanisms of cardiac arrhythmia. *Cardiac Electrophysiology Clinics, 3*(1), 23–45. Retrieved from https://pmc.ncbi.nlm.nih.gov/articles/PMC3164530/

3. De Chazal, P., O'Dwyer, M., & Reilly, R. B. (2004). Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *IEEE Transactions on Biomedical Engineering, 51*(7), 1196–1206. doi:10.1109/TBME.2004.827359

4. García, J., Sörnmo, L., & Laguna, P. (2005). Automatic detection of premature ventricular contractions using vectorcardiographic features and support vector machines. *Computers in Cardiology*, 453–456.

5. Sellami, A., & Hwang, M. (2019). A robust deep convolutional neural network with batch-weighted loss for heartbeat classification. *Expert Systems with Applications, 122*, 75–84. doi:10.1016/j.eswa.2018.12.037

6. Li, Y., Pang, H., & Wang, X. (2021). A residual convolutional neural network for ECG heartbeat classification. *Biomedical Signal Processing and Control, 68*, Article 102669. doi:10.1016/j.bspc.2021.102669

Development of a Deep Learning Model for Automated Arrhythmia Recognition

Abstract: *Cardiac arrhythmias represent one of the primary etiologies of life-threatening cardiovascular events. Automating diagnosis via electrocardiogram (ECG) signals faces significant challenges due to inter-patient variability, data imbalance, and micro-morphological similarities across different heartbeat types. This study proposes* ***HybridECGNet****, a deep hybrid neural network architecture designed for multi-dimensional feature extraction. The system integrates* ***Multi-scale Depthwise Separable Convolutional*** *(Multi-scale DSWC) layers with* ***Transformers*** *to capture spatial morphologies, alongside* ***Bidirectional Long Short-Term Memory*** *(BiLSTM) networks to model temporal dynamics. A* ***Gated Fusion*** *mechanism is specifically engineered to dynamically calibrate weights between the two branches based on pathological characteristics. Experimental results on the* ***MIT-BIH Arrhythmia Database****, strictly following the* ***inter-patient paradigm****, demonstrate that the model achieves an overall accuracy of 98.88%. Notably, the sensitivity for detecting critical premature contractions exceeds 97%, validating the model's viability for deployment in real-time medical monitoring devices.*

Keywords: cardiac arrhythmia; deep learning; inter-patient paradigm.

Xây dựng mô hình học sâu nhận dạng cử chỉ tay sử dụng kiến trúc hai nhánh

Vũ Văn Đông[1], Nguyễn Trọng Tuyến[1*], Nguyễn Mạnh Cường[1*]

[1]Viện Tên lửa và Kỹ thuật Điều khiển; Học viện Kỹ thuật quân sự

Email: shinehvqs@gmail.com, tuyen.nguyen@lqdtu.edu.vn, nmcuong869@lqdtu.edu.vn;
Tel: 0328274495

Tóm tắt

Nhận dạng cử chỉ tay sử dụng tín hiệu điện cơ bề mặt là một lĩnh vực quan trọng trong tương tác người-máy, điều khiển chi giả và phục hồi chức năng. Tuy nhiên, việc xây dựng mô hình có khả năng tổng quát hóa cao trên nhiều người dùng khác nhau (cross-subject) vẫn là một thách thức lớn do có sự biến thiên đáng kể trên tín hiệu EMG giữa các cá nhân. Nghiên cứu này đề xuất một kiến trúc học sâu hai nhánh để giải quyết vấn đề này. Mô hình xử lý tín hiệu song song theo hai nhánh riêng biệt: (1) nhánh mạng neuron tích chập dựa trên cơ chế tự chú ý (self-attention) và (2) nhánh mạng tích chập kết hợp LSTM hai chiều (BiLSTM). Mô hình đạt hiệu suất 91.89% khi đánh giá phân loại 17 cử chỉ trên bộ cơ sở dữ liệu GrabMyo. Kết quả này đã chứng minh hiệu quả của kiến trúc đa nhánh, đồng thời khẳng định tiềm năng của mô hình trong việc xây dựng các hệ thống nhận dạng cử chỉ mạnh mẽ và không phụ thuộc người dùng.

***Từ khóa*:** *nhận dạng cử chỉ tay; đặc trưng không gian; đặc trưng thời gian*.

1. Đặt vấn đề

Trong những năm gần đây, giao diện tương tác người-máy (Human-Computer Interaction – HCI) đã đạt được những bước tiến vượt bậc, mở ra nhiều phương thức giao tiếp tự nhiên và trực quan giữa con người và công nghệ. Sử dụng tín hiệu điện cơ bề mặt (surface electromyogram) ghi lại hoạt động điện của cơ bắp đã nổi lên như một phương pháp đầy hứa hẹn trong nhận dạng cử chỉ tay. Với tiềm năng ứng dụng rộng rãi trong nhiều công nghệ hiện đại bao gồm điều khiển tay giả thông minh, tương tác với môi trường thực tế ảo (VR/AR), cho đến các công nghệ hỗ trợ cho người khuyết tật, phục hồi chức năng cho người bệnh liệt chi trên, điều này đòi hỏi cần giải mã chính xác ý định của người dùng từ tín hiệu EMG.

Trong lĩnh vực nhận dạng cử chỉ tay từ tín hiệu điện cơ EMG, các phương pháp nghiên cứu được chia thành hai hướng chính: học máy truyền thống và học sâu. Hướng tiếp cận truyền thống tập trung xây dựng và thiết kế các đặc trưng thủ công từ miền thời gian và tần số, sau đó sử dụng các thuật toán như support vector machine, random forest hoặc k-nearest neighbors để thực hiện phân loại và đưa ra dự đoán cử chỉ. Nghiên cứu của Tavakoli và cộng sự (2018) đã xây dựng một mô hình tối thiểu để phân loại 04 cử chỉ với độ chính xác cao trên người duy nhất (for single-user) sử dụng không gian đặc trưng chiều cao (high-dimensional feature space) và thuật toán SVM để phân loại [1]. Công trình của Tepe và Demir (2022) sử dụng SVM cùng với một bộ đặc trưng lớn miền thời gian để phân loại cử chỉ theo thời gian thực, cho thấy khả năng xây dựng mô hình đạt hiệu năng cao khi dữ liệu được tùy chỉnh theo từng người [2]. Tuy nhiên, hướng nghiên cứu này đòi hỏi kiến thức chuyên môn sâu và trong nhiều trường hợp, mô hình có thể không nắm bắt được hết các mẫu tín hiệu.

Để vượt qua những giới hạn đó, cộng đồng nghiên cứu đã chuyển dịch mạnh mẽ sang các mô hình học sâu đầu cuối (end-to-end deep learning model). Các mạng neuron tích chập được sử dụng rộng rãi để tự động học các đặc trưng không gian, bằng cách xử lý tín hiệu EMG như một dạng ảnh 2D. Công trình của Chen và cộng sự (2020) đã cho thấy với ngay cả một kiến

trúc CNN nhỏ gọn cũng có thể đạt được hiệu năng cạnh tranh với các mô hình truyền thống [3]. Oh và Jo (2019) đã kết hợp phép biến đổi wavelet với CNN cho thấy việc tiền xử lý làm nổi bật các đặc trưng, có thể cải thiện độ chính xác của mô hình CNN lên tới 94% [4]. Song song với đó, để khai thác bản chất chuỗi thời gian của tín hiệu, các mạng neuron hồi quy như LSTM và GRU cũng được áp dụng. Nghiên cứu của Samadani (2018) đã so sánh và khẳng định hiệu quả của LSTM và GRU trong việc nắm bắt các phụ thuộc động của cử chỉ [5].

Với mỗi cách tiếp cận đều có những ưu điểm riêng, một số nghiên cứu đã tập trung vào việc phát triển các mô hình lai. Các mô hình này kết hợp sức mạnh của các đặc trưng thủ công và đặc trưng học sâu. Một công trình tiêu biểu cho hướng đi này là của Fajardo và cộng sự (2021), trong đó các tác giả đã chứng minh rằng việc nối trực tiếp các đặc trưng sâu được trích xuất từ CNN với các đặc trưng thủ công giúp cải thiện đáng kể độ chính xác so với chỉ sử dụng một trong hai loại [6]. Nghiên cứu này cung cấp một bằng chứng thực nghiệm mạnh mẽ, khẳng định rằng các đặc trưng thủ công có đóng góp nhất định và có thể bổ trợ hiệu quả cho các mô hình học sâu. Cùng với đó, các nhà nghiên cứu cũng không ngừng cải tiến kiến trúc các mô hình học sâu để tăng cường hiệu năng. Một trong những đóng góp có ảnh hưởng lớn là việc áp dụng học chuyển giao (Transfer Learning), được tiên phong bởi Côté-Allard và cộng sự (2019) [7]. Bằng cách huấn luyện trước một mô hình CNN trên nhiều người dùng và sau đó tinh chỉnh lại trên dữ liệu của người dùng mới, nghiên cứu đã cho thấy có thể giảm đáng kể lượng dữ liệu cần thu thập mà vẫn duy trì độ chính xác cao. Một hướng đi tiên tiến khác là phát triển các kiến trúc đa nhánh (multi-branch architectures) để phân tích tín hiệu từ nhiều góc độ. Mô hình SGRN của Gan và cộng sự (2025) đã sử dụng các nhánh song song để trích xuất riêng biệt thông tin không gian, thông tin thời gian, và thông tin không gian-thời gian kết hợp, qua đó đạt được hiệu năng SOTA [8]. Các công trình này cho thấy xu hướng hiện đại là xây dựng các kiến trúc phức hợp, có khả năng học hỏi và thích ứng tốt hơn.

Mặc dù các mô hình lai và đa nhánh đã tồn tại, vẫn còn một khoảng trống trong việc phát triển một kiến trúc có thể khai thác đồng thời cả hai loại thông tin (đặc trưng không gian sâu và đặc trưng thời gian sâu). Để giải quyết vấn đề này, nghiên cứu của chúng tôi đề xuất một kiến trúc hai nhánh. Cách tiếp cận này cho phép mô hình của chúng tôi đạt được hiệu năng SOTA trong cả hai kịch bản đánh giá khắt khe cross-session, cung cấp một giải pháp mạnh mẽ và toàn diện cho bài toán nhận dạng cử chỉ dựa trên tín hiệu EMG.

2. Phương pháp

2.1. Bộ cơ sở dữ liệu

Trong nghiên cứu này, chúng tôi sử dụng bộ cơ sở dữ liệu công khai Gesture Recognition and Biometrics Electromygram (GRABMyo) [10–12]. Đây là bộ cơ sở dữ liệu lớn và đa dạng, được thiết kế cho cả hai bài toán nhận dạng cử chỉ và sinh trắc học. Dữ liệu được thu thập từ 43 người tham gia khỏe mạnh (23 M, 20 F) có độ tuổi trung bình 26.35±2.89, mỗi người thực hiện 17 cử chỉ (16 cử chỉ động và 1 trạng thái nghỉ) và mỗi cử chỉ được thực hiện 7 lần. Để tránh sự mỏi cơ, một khoảng nghỉ tầm 10 sec được đáp ứng. Quá trình thu thập được lặp lại trong 3 phiên (sessions) vào các ngày khác nhau (ngày 1, 8 và 29) ghi nhận sự thay đổi của tín hiệu theo thời gian.

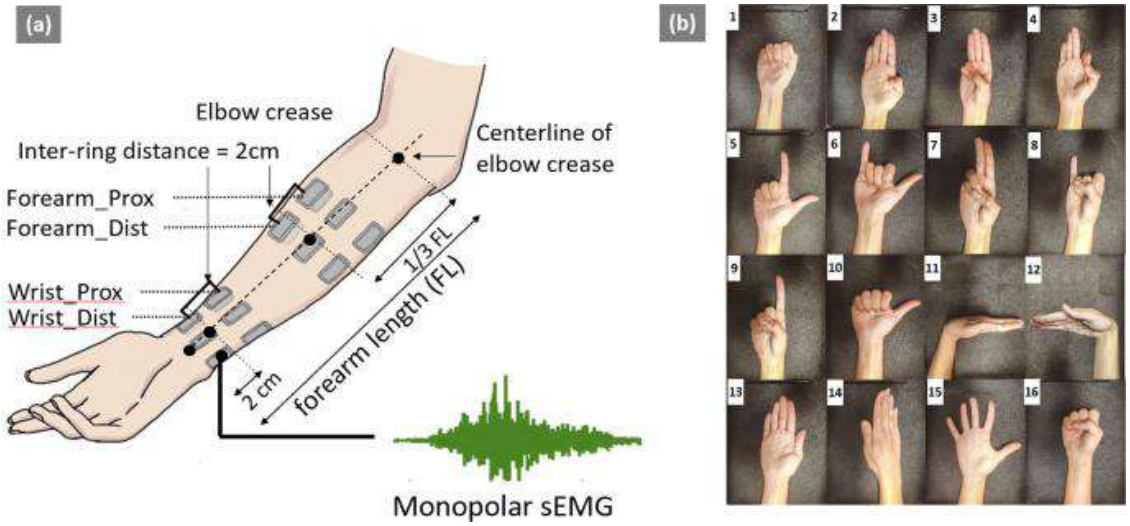
Có tổng cộng 32 kênh tín hiệu EMG với độ dài 5 giây được thu nhận với tần số lấy mẫu 2048Hz. Trong số 32 kênh này có 16 kênh được lấy ở cánh tay và 12 kênh còn lại được lấy ở

cổ tay, và 4 kênh không được sử dụng (Hình 1).

2.2. Tiền xử lý tín hiệu

Quy trình xử lý tín hiệu sEMG được thực hiện qua các bước: lựa chọn kênh, lọc nhiễu, phân đoạn cửa sổ, trích xuất đặc trưng. Bộ dữ liệu GrabMyo gồm 32 kênh tín hiệu. Dựa trên cấu hình thực nghiệm, chúng tôi loại bỏ 4 kênh nhiễu và không mang nhiều thông tin (các kênh 16, 23, 24, 31) và giữ lại 28 kênh được sử dụng cho quá trình huấn luyện.

Tín hiệu thô sEMG thường chứa các nhiễu do chuyển động ở tần số thấp và nhiễu điện từ ở tần số cao. Chúng tôi áp dụng bộ lọc thông dải Butterworth bậc 4 với dải tần cắt từ 20 Hz đến 450 Hz để loại bỏ các nhiễu này, giúp giữ lại dải năng lượng chính của hoạt động cơ bắp, đồng thời, loại bỏ các thành phần nhiễu không mong muốn. Sau lọc nhiễu tín hiệu, chúng tôi chia tín hiệu thành các cửa sổ có độ dài 300 ms, với khoảng overlapping trên mỗi cửa sổ kế tiếp nhau là 200 ms.

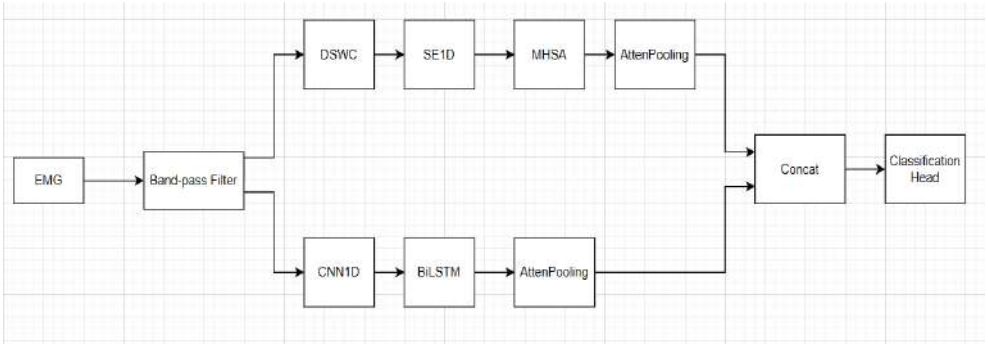


Hình 1. Bố trí điện cực thu tín hiệu và các cử chỉ phân loại.

2.3. Mô hình đề xuất

2.3.1. Tổng quan mô hình

Chúng tôi đề xuất kiến trúc hai nhánh nhằm giải quyết thách thức về tính đa dạng và nhiễu trên tín hiệu sEMG. Mô hình được thiết kế xử lý song song biểu diễn dữ liệu: (1) Đặc trưng không gian được trích xuất bởi mạng tích chập sâu kết hợp với cơ chế attention; và (2) Đặc trưng thời gian (Temporal Features) được mô hình hòa bởi mạng bộ nhớ ngắn hạn dài hạn hai chiều (BiLSTM). Kiến trúc mô hình được minh họa trên Hình 2.



Hình 2. Kiến trúc mô hình hai nhánh nhận dạng cử chỉ tay

2.3.2. Đặc trưng không gian

Nhánh này tập trung giải mã các mẫu kích hoạt cơ trong không gian (spatial muscle synergies). Cấu trúc nòng cốt (Stem) bao gồm 3 khối dilated depthwise separable convolutions. Mỗi khối sử dụng tích chập tách biệt chiều sâu để giảm tham số, tích hợp cơ chế Squeeze-and-Excitation (SE) để trọng số hóa kênh, và sử dụng các tỷ lệ giãn nở tăng dần ($d = 1, 2, 4$) để mở rộng vùng tiếp nhận (receptive field). Sau khối Stem, đặc trưng được cộng với Learnable Positional Encoding P và đưa qua lớp Multi-Head Self-Attention (MHSA) để nắm bắt sự phụ thuộc toàn cục:

$$X_{attn} = \text{MHSA}(X_{stem} + P) \quad (1)$$

Cuối cùng, thay vì dùng Global Average Pooling, chúng tôi sử dụng cơ chế Attention Pooling với vector truy vấn học được q để tổng hợp đặc trưng theo chiều thời gian, tạo ra vector đại diện không gian $z_{deep} \in \mathbb{R}^{256}$:

$$z_{deep} = \sum_{t=1}^{T} \alpha_t x_t, \text{ với } \alpha_t = \text{Softmax}(\frac{x_t q^T}{\sqrt{d_k}}) \quad (2)$$

2.3.3. Đặc trưng thời gian

Nhánh thứ hai chuyên biệt hóa việc mô hình hóa chuỗi thời gian để phân biệt các pha động của cử chỉ. Tín hiệu đầu vào **X** trước tiên được đi qua một lớp tích chập 1D (kernel size = 5) kết hợp với GroupNorm và GELU để trích xuất các đặc trưng cục bộ mức thấp. Chuỗi đặc trưng này sau đó được đưa vào mạng Bidirectional LSTM (BiLSTM). Việc xử lý hai chiều giúp mạng nắm bắt ngữ cảnh từ cả quá khứ và tương lai tại mỗi bước thời gian. Đầu ra của BiLSTM được chuẩn hóa qua LayerNorm và Dropout, sau đó, cũng được tổng hợp bằng cơ chế Attention Pooling như trên nhánh spatial, thu được vector đại diện đặc trưng theo thời gian $z_{time} \in \mathbb{R}^{256}$, đảm bảo các thời điểm quan trọng nhất của cử chỉ được bảo toàn.

3. Thực nghiệm

Để đánh giá toàn diện hiệu suất, độ mạnh mẽ và khả năng tổng quát hóa của mô hình do chúng tôi đề xuất, chúng tôi đã thiết kế quy trình đánh giá nghiêm ngặt và riêng biệt dựa trên tập dữ liệu GrabMyo. Hiệu suất của mô hình được định lượng bằng cách sử dụng các chỉ số phân loại tiêu chuẩn, bao gồm độ chính xác (accuracy), macro-F1, ma trận nhầm lẫn (confusion matrix) và diện tích dưới đường cong (AUC) từ đường cong ROC.

3.1. Đánh giá ổn định qua các phiên

Giao thức này được thiết kế để đánh giá độ bền vững (robustness) và độ tin cậy của mô hình theo thời gian. Nó mô phỏng kịch bản thực tế khi một người sử dụng hệ thống qua nhiều ngày khác nhau, trong điều kiện tín hiệu có thể thay đổi do sự mỏi cơ, sự dịch chuyển nhẹ của các điện cực hoặc các thay đổi sinh lý khác. Để đảm bảo tính nhất quán, trong mỗi fold, dữ liệu của 2 người tham gia cố định (người tham gia số 2 và 5) luôn được trích ra từ tập huấn luyện để sử dụng làm tập kiểm chứng (validation). Điều này cho phép mô hình được tối ưu hóa dựa trên một tập validation ổn định, không phụ thuộc và session nào đang được dùng để kiểm tra. Trong nghiên cứu này, chúng tôi tiến hành huấn luyện trên dữ liệu các phiên 1 và 3, và kiểm chứng trên phiên 2. Điều này đảm bảo mô hình không bị rò rỉ dữ liệu giữa các phiên thu thập dữ liệu khác nhau.

3.2. Huấn luyện mô hình

Mô hình đề xuất được triển khai bằng framework PyTorch và được huấn luyện trên một hệ thống máy tính hiệu năng cao, trang bị GPU NVIDIA RTX 3060 với 12 GB VRAM. Để tăng tốc quá trình huấn luyện, chúng tôi đã tận dụng nền tảng CUDA và kích hoạt bộ tự động tinh chỉnh cuDNN (cuDNN autotuner). Hơn nữa, phương pháp độ chính xác hỗn hợp tự động (AMP - Automatic Mixed Precision) đã được sử dụng nhằm giảm mức tiêu thụ bộ nhớ và tăng tốc độ tính toán bằng cách dùng độ chính xác dấu phẩy động 16-bit ở những nơi có thể, trong khi vẫn duy trì độ chính xác 32-bit để đảm bảo tính ổn định số học.

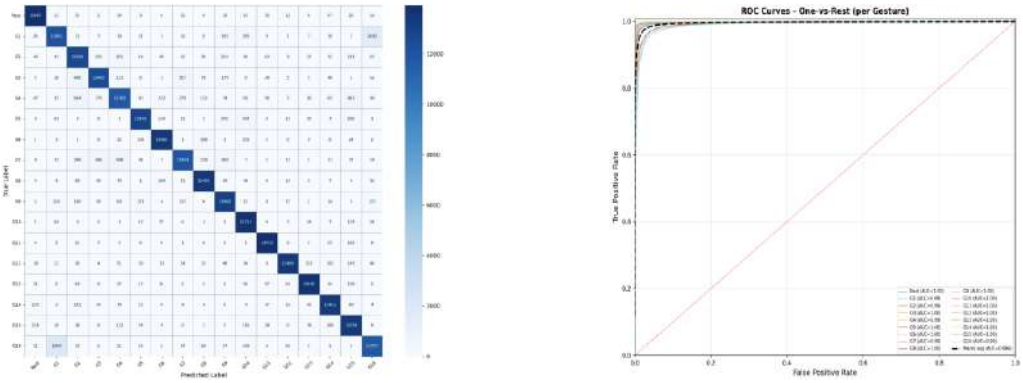
Các tham số của mô hình được tối ưu hóa bằng thuật toán AdamW, giúp tách biệt quá trình suy giảm trọng số (weight decay) khỏi quá trình cập nhật gradient. Chúng tôi thiết lập tốc độ học (learning rate) ban đầu là 1.0×10^{-3} và hệ số suy giảm trọng số (chính quy hóa L2) là 3.0×10^{-4}. Một bộ lập lịch tốc độ học (learning rate scheduler) đã được sử dụng, kết hợp pha khởi động tuyến tính (linear warmup) cho epoch đầu tiên, theo sau là lịch trình ủ cosine (cosine annealing schedule) cho 14 epoch còn lại, giúp giảm dần tốc độ học về 0. Chiến lược này cho phép sự hội tụ ổn định ở các giai đoạn đầu và tinh chỉnh (fine-tuning) ở các giai đoạn sau.

Hàm mục tiêu (objective function) cho bài toán phân loại là hàm mất mát entropy chéo (cross-entropy loss). Để ngăn mô hình trở nên quá tự tin (over-confident) và cải thiện khả năng tổng quát hóa, chúng tôi đã áp dụng kỹ thuật làm mịn nhãn (label smoothing) với hệ số làm mịn là $\alpha = 0.08$. Một tập hợp toàn diện các kỹ thuật tăng cường dữ liệu trực tiếp (on-the-fly data augmentation) đã được áp dụng cho các đầu vào cửa sổ tín hiệu sEMG thô trong quá trình huấn luyện. Với xác suất 0.85, mỗi mẫu sẽ trải qua một sự kết hợp bao gồm việc thêm nhiễu Gauss ($\sigma = 0.015$), chia tỷ lệ biên độ ngẫu nhiên (trong khoảng [0.85, 1.15]), dịch chuyển thời gian ngẫu nhiên (lên đến ± 6 mẫu) và loại bỏ kênh ngẫu nhiên (random channel dropout) (tối đa 2 kênh với xác suất 0.1). Việc chính quy hóa (regularization) bổ sung được thực hiện thông qua các lớp dropout bên trong khối BiLSTM và khối lai (fusion component) của kiến trúc.

Mô hình được huấn luyện trong tổng cộng 15 epoch với kích thước batch (batch size) là 128. Một seed ngẫu nhiên cố định đã được thiết lập cho tất cả các thư viện nhằm đảm bảo tính lặp lại (reproducibility) của các thí nghiệm. Chúng tôi đã triển khai cơ chế dừng sớm (early stopping) dựa trên hiệu suất của tập xác thực (validation set). Trạng thái của mô hình được lưu lại sau mỗi epoch nếu và chỉ nếu độ chính xác trên tập xác thực của epoch đó vượt qua mức độ chính xác tốt nhất từng được ghi nhận. Các kết quả cuối cùng được báo cáo trong bài báo này đạt được độ chính xác xác thực cao nhất, đúng như được thể hiện trong nhật ký huấn luyện (training logs).

4. Kết quả và thảo luận

Mô hình đề xuất được đánh giá toàn diện trên bộ dữ liệu GrabMyo với 17 cử chỉ thông qua các giao thức thử nghiệm như đã chỉ ra trên Section 3. Mô hình đạt các chỉ số độ chính xác 91.89% và F1-score 91.88%, với tỷ lệ lỗi sai 8.11%. Ma trận nhầm lẫn và đường cong ROC phân loại các cử chỉ được minh họa trên Hình 3a và Hình 3b.



(a) (b)

Hình 3. Ma trận nhầm lẫn của mô hình phân loại cử chỉ theo 17 cử chỉ (a) và đường cong ROC (b)

Dựa trên ma trận nhầm lẫn và đường cong ROC phân loại 17 cử chỉ này, có thể thấy rằng mô hình có khả năng phân loại tốt trên hầu hết mọi cử chỉ (Macro-AUC ~ 0.996). Điều này khẳng định mô hình cho phép biểu diễn các dữ liệu không gian và thời gian, đặc trưng các mẫu sóng tín hiệu điện cơ EMG. Trong các cử chỉ, có hai cử chỉ dễ nhầm lẫn là cử chỉ số 1 và 16, tương ứng với các trạng thái, kẹp bên (Lateral Prehension) và nắm tay đóng (Hand Close). Điều này xuất phát từ sự đồng nhất về cơ sinh lý của hai cử chỉ này, khi cả hai cử chỉ này đều kích hoạt cơ gấp các ngón nông và cơ gấp các ngón sâu.

Phân tích so sánh với các mô hình cơ sở (baselines) gần đây trên tập dữ liệu GrabMyo làm nổi bật tính ưu việt của phương pháp do chúng tôi đề xuất. Mô hình CNN phân cấp (hierarchical cascade CNN) [13], sử dụng kiến trúc tích chập hạng nhẹ, chỉ đạt độ chính xác 71.62% dưới một quy trình đánh giá chéo phiên (cross-session protocol) tương tự. Ngược lại, mô hình của chúng tôi vượt trội hơn mô hình cơ sở này với một khoảng cách đáng kể là 20.26%. Hơn nữa, so với mô hình Transformer Tăng cường Ngắn hạn (Short-Term Enhanced Transformer - STET) [14], vốn báo cáo mức độ chính xác 90.76% khi sử dụng phương pháp chia dữ liệu nội đối tượng (intra-subject split - 5 lượt huấn luyện/2 lượt kiểm tra), mô hình của chúng tôi đạt độ chính xác cao hơn là 91.88% trong kịch bản dự đoán nội suy (interpolation prediction) tương đương. Điều này làm nổi bật tính mạnh mẽ vốn có của quy trình trích xuất đặc trưng (feature extraction pipeline) được đề xuất, giúp phân tách các ranh giới lớp (class boundaries) một cách tự nhiên mà không cần đến các chiến lược khai thác mẫu âm tính khó (hard-negative mining) phức tạp.

5. Kết luận

Nghiên cứu này đề xuất một kiến trúc hai nhánh mạnh mẽ, giúp cân bằng hiệu quả sự đánh đổi giữa khả năng tổng quát hóa (generalization) và độ chính xác (precision). Bằng cách kết hợp biểu diễn không - thời gian sâu (deep spatial-temporal representation learning), phương pháp của chúng tôi không chỉ vượt trội so với các mô hình cơ sở tiên tiến nhất hiện nay (state-of-the-art baselines) về cả độ chính xác lẫn tính ổn định, mà còn mang đến một giải pháp thiết thực cho các ứng dụng trong thực tế. Kết quả nghiên cứu chỉ ra rằng, mô hình có tiềm năng ứng dụng trong các hệ thống phục hồi chức năng chi trên với độ chính xác cao và ổn định.

Tài liệu tham khảo

1. Tavakoli, M., Benussi, C., Alhais Lopes, P., Osorio, L. B., & de Almeida, A. T. (2018). Robust hand gesture recognition with a double channel surface EMG wearable armband and SVM classifier. Biomedical Signal Processing and Control, 46, 121–130. https://doi.org/10.1016/j.bspc.2018.07.010

2. Tepe, C., & Demir, M. C. (2022). Real-time classification of EMG Myo armband data using support vector machine. IRBM, 43, 300–308. https://doi.org/10.1016/j.irbm.2022.06.001

3. Chen, L., Fu, J., Wu, Y., Li, H., & Zheng, B. (2020). Hand gesture recognition using compact CNN via surface electromyography signals. Sensors, 20(3), 672. https://doi.org/10.3390/s20030672

4. Oh, D. C., & Jo, Y. U. (2019). EMG-based hand gesture classification by scale average wavelet transform and CNN. In 2019 19th International Conference on Control, Automation and Systems (ICCAS) (pp. 533–538). IEEE. https://doi.org/10.23919/ICCAS47443.2019.8971730

5. Samadani, A. (2018). Gated recurrent neural networks for EMG-based hand gesture classification: A comparative study. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS) (pp. 1–4). IEEE. https://doi.org/10.1109/EMBC.2018.8512531

6. Fajardo, J. M., Gomez, O., & Prieto, F. (2021). EMG hand gesture classification using handcrafted and deep features. Biomedical Signal Processing and Control, 63, 102210. https://doi.org/10.1016/j.bspc.2020.102210

7. Côté-Allard, U., Fall, C. L., Drouin, A., Campeau-Lecours, A., Gosselin, C., Glette, K., et al. (2019). Deep learning for electromyographic hand gesture signal classification using transfer learning. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 27, 760–771. https://doi.org/10.1109/TNSRE.2019.2896269

8. Gan, Z., Bai, Y., Wu, P., Xiong, B., Zeng, N., Zou, F., et al. (2025). SGRN: SEMG-based gesture recognition network with multi-dimensional feature extraction and multi-branch information fusion. Expert Systems with Applications, 259, 125302. https://doi.org/10.1016/j.eswa.2024.125302

9. Perez, E., Strub, F., De Vries, H., Dumoulin, V., & Courville, A. (2018). FiLM: Visual reasoning with a general conditioning layer. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 32, pp. 3942–3951). https://doi.org/10.1609/aaai.v32i1.11671

10. PhysioNet. (n.d.). Gesture recognition and biometrics electromyogram (GRABMyo) (v1.1.0). Retrieved November 22, 2025, from https://physionet.org/content/grabmyo/1.1.0/

11. Pradhan, A., He, J., & Jiang, N. (2022). Multi-day dataset of forearm and wrist electromyogram for hand gesture recognition and biometrics. Scientific Data, 9. https://doi.org/10.1038/S41597-022-01836-Y

12. Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation, 101(23), e215–e220. https://doi.org/10.1161/01.CIR.101.23.E215

13. Cansiz, B., Dudukcu, H. V., Taskiran, M., & Kahraman, N. (2025). Hierarchical cascade deep learning for EMG-based behavioral biometrics: Gesture and subject classification. IEEE Access, 13, 124115–124128. https://doi.org/10.1109/ACCESS.2025.3588791

14. Guo, W., Qiao, Z., Sun, Y., & Xiong, H. (2024). Revisiting noise resilience strategies in gesture recognition: Short-term enhancement in surface electromyographic signal analysis.

Development of a Deep Learning Model for Hand Gesture Recognition Using a Dual-Branch Architecture

Abstract: *Hand gesture recognition utilizing surface electromyography (sEMG) signals is a pivotal field in human-computer interaction, prosthetic control, and rehabilitation. However, developing models with high cross-subject generalization remains a significant challenge due to the substantial inter-individual variability inherent in EMG signals. This study proposes a dual-branch deep learning architecture to address this issue. The model processes signals in parallel through two distinct pathways: (1) a convolutional neural network (CNN) branch integrated with a self-attention mechanism, and (2) a hybrid CNN-Bidirectional Long Short-Term Memory (BiLSTM) branch. The proposed model achieved an accuracy of 91.89% in classifying 17 gestures using the GrabMyo database. These results demonstrate the efficacy of the multi-branch architecture and underscore the model's potential in developing robust, subject-independent gesture recognition systems.*

Keywords: hand gesture recognition; spatial features; temporal features.

Nghiên cứu ảnh hưởng dữ liệu đặc trưng đến kết quả ước lượng trạng thái SOC của pin xe điện sử dụng kỹ thuật học sâu.

Vũ Văn Chiến*, Nguyễn Ngọc Tuấn, Ngô Việt Cường

Viện Tên lửa và Kỹ thuật điều khiển

** Email: chienvuvan@lqdtu.edu.vn; Tel: 0979936673*

Tóm tắt

Ước lượng chính xác trạng thái sạc (SoC) của pin có vai trò quan trọng trong việc đảm bảo sự làm việc tin cậy của hệ thống quản lý pin trong ô tô điện. Bài báo sử dụng dữ liệu của pin LG-HG2 với các chu trình lái xe ở dải nhiệt độ từ -10 °C đến 25 °C và một cấu trúc mạng DNN để kiểm chứng tính chính xác của mô hình, đồng thời nghiên cứu ảnh hưởng của sự thay đổi tham số đặc trưng đầu vào đến chất lượng ước lượng SOC. Bộ lọc trung bình động và bộ lọc thông thấp Buterworth bậc 1 được ứng dụng để tạo ra các đặc trưng đầu vào cho luyện mạng. Kết quả nghiên cứu cho thấy, ước lượng SOC càng chính xác khi nhiệt độ pin tăng dần. Sử dụng bộ lọc Butterworth bậc 1 ở các tần số lọc 0.5mHz và 5mHz đã cải thiện rõ rệt độ chính xác của ước lượng SOC, với sai số MSE là 0.023% ở 25 ^{0}C, so với bộ lọc trung bình động là 0.03% và không sử dụng bộ lọc là 0.152%.

Từ khóa: *ước lượng SOC, kỹ thuật học sâu, xe điện, mạng nơ-ron*

1. Giới thiệu chung

Cho tới thời điểm hiện tại, pin Lithium-ion vẫn là nguồn cung cấp năng lượng chủ yếu cho các loại xe điện. Với những đặc điểm nổi trội bao gồm: Mật độ năng lượng cao, vòng đời dài, hiệu suất cao, điện áp lớn và tốc độ tự xả thấp [1]. Để đảm bảo an toàn trong các điều kiện làm việc khác nhau của pin, một hệ thống quản lý pin (BMS) là bắt buộc được sử dụng. BMS chứa các loại cảm biến, các mạch bảo vệ, mạch cân bằng pin, bộ điều khiển để thực hiện các thuật toán nhằm kiểm soát và cập nhật dữ liệu, phát hiện lỗi, cân bằng điện áp giữa các cell pin, ngăn ngừa mọi vấn đề về quá nhiệt, gây hỏng pin [2]. Sơ đồ khối của hệ thống BMS được thể hiện trên Hình 1.

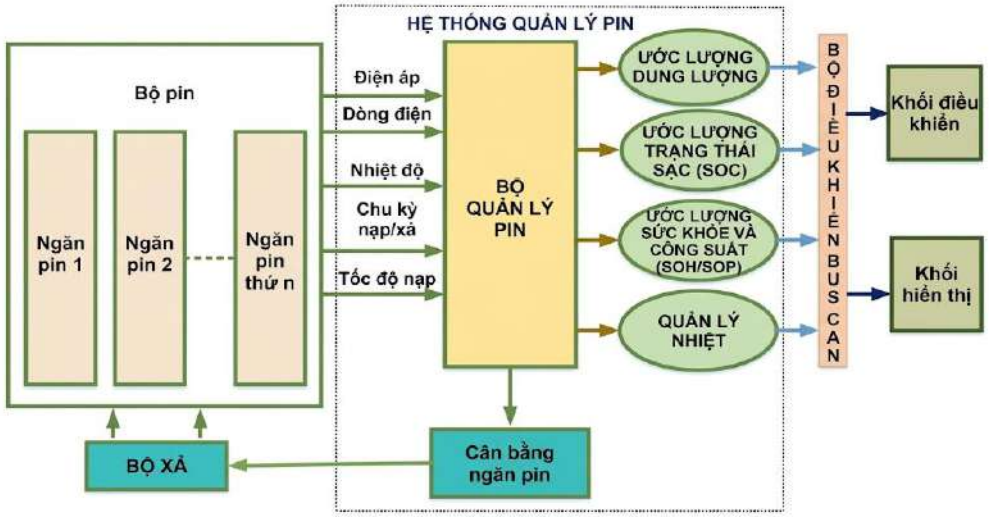
Trạng thái sạc (SOC) là một thông số rất quan trọng trong hệ thống BMS. SOC biểu thị mức năng lượng tại thời điểm hiện tại còn lại trong pin so với dung lượng tối đa của nó [3] được thể hiện trong công thức (1).

$$SOC(t) = \frac{Q_{remain}(t)}{Q_{max}} \times 100\% \quad (1)$$

Trong đó: $Q_{remain}(t)$ là dung lượng thực tế tại thời điểm t, $Q_{max}(t)$ là dung lượng pin khi được sạc đầy.

Việc ước lượng chính xác SoC cho phép dự đoán đáng tin cậy quãng đường di chuyển còn lại, đồng thời ngăn ngừa các điều kiện vận hành nguy hiểm như sạc quá mức hoặc xả sâu, vốn có thể làm gia tăng tốc độ suy giảm pin hoặc gây rủi ro về an toàn [4]. Mặc dù SoC có vai trò tương tự như đồng hồ báo nhiên liệu trong phương tiện truyền thống, nó không thể được đo trực tiếp mà phải được ước lượng thông qua các tín hiệu đo được như điện áp đầu cực, dòng điện và nhiệt độ.

Quá trình ước lượng SoC gặp nhiều thách thức do đặc tính điện hóa phi tuyến của pin lithium-ion, sự phụ thuộc vào nhiệt độ và các hiệu ứng lão hóa theo thời gian [5].



Hình 1 . Sơ đồ khối của hệ thống BMS trên xe điện

Các phương pháp truyền thống như Coulomb counting hoặc các bộ quan sát dựa trên mô hình, điển hình là Bộ lọc Kalman mở rộng (Extended Kalman Filter – EKF), đã được ứng dụng rộng rãi trong các hệ thống BMS thực tế [6]. Tuy nhiên, các phương pháp này phụ thuộc mạnh vào độ chính xác của mô hình pin, giá trị SOC ban đầu, độ chính xác của cảm biến và quá trình nhận dạng tham số. Sai lệch mô hình, sự trôi tham số và các động học chưa được mô tả đầy đủ có thể làm suy giảm đáng kể độ chính xác ước lượng trong điều kiện vận hành thực tế [7].

Trong những năm gần đây, các kỹ thuật học máy (Machine Learning – ML) và học sâu (Deep Learning – DL) đã thu hút sự quan tâm đáng kể trong lĩnh vực hệ thống quản lý pin, nhờ tiềm năng khắc phục những hạn chế đã được đề cập trước đó [8]. Việc ứng dụng các phương pháp ML/DL ngày càng trở nên phổ biến do quá trình thu thập dữ liệu pin trong giai đoạn thử nghiệm tương đối đơn giản, và chính bộ dữ liệu này có thể được sử dụng cho cả quá trình huấn luyện và kiểm thử các mô hình ML/DL. Các mô hình học sâu, đặc biệt là mạng nơ-ron hồi tiếp (Recurrent Neural Networks – RNNs), đã cho thấy khả năng vượt trội trong việc nắm bắt các phụ thuộc theo thời gian và các đặc tính phi tuyến tồn tại trong dữ liệu chuỗi thời gian [9]. Trong số các kiến trúc RNN phổ biến, mạng bộ nhớ dài-ngắn hạn (Long Short-Term Memory – LSTM), LSTM hai chiều (Bidirectional LSTM – BiLSTM) và mạng đơn vị hồi tiếp có cổng (Gated Recurrent Unit – GRU) được sử dụng rộng rãi nhờ khả năng xử lý hiệu quả các phụ thuộc dài hạn và dữ liệu tuần tự [10].

Hong và cộng sự [11] đã áp dụng mô hình LSTM, vốn tích hợp một ô nhớ (memory cell) đặc biệt cho phép mô hình lựa chọn ghi nhớ hoặc quên thông tin trong khoảng thời gian dài. Mạng BiLSTM kết hợp ưu điểm của LSTM bằng cách xem xét đồng thời cả thông tin quá khứ và tương lai trong chuỗi dữ liệu. Bian và cộng sự [12] đã sử dụng BiLSTM nhằm cải thiện khả năng mô hình hóa động học của mạng đối với pin lithium-ion. Trong khi đó, Li và cộng sự [13] đã phát triển mô hình GRU, một phiên bản đơn giản hóa của LSTM, có hiệu suất tương đương nhưng yêu cầu số lượng tham số ít hơn. Kết quả nghiên cứu đã làm rõ ảnh hưởng của các tham số khác nhau đến độ chính xác của quá trình ước lượng SoC.

Trong bài báo này tác giả tiếp tục sử dụng mô hình mạng học sâu DNN để kiểm chứng tính chính xác của mô hình, đồng thời nghiên cứu ảnh hưởng của sự thay đổi tham số đặc trưng đầu vào đến chất lượng ước lượng SOC. Tác giả sử dụng hai kỹ thuật làm trơn, mịn dữ liệu, đó là trung bình động với giá trị cửa sổ trượt ξ và bộ lọc thông thấp Buterworth bậc 1 để tạo thêm các đặc trưng đầu vào cho luyện mạng, từ đó thấy được ảnh hưởng tích cực của việc bổ sung dữ liệu đầu vào. Phần còn lại của bái báo được trình bày như sau: Phần 2 trình bày tổng quan về các cấu trúc mạng được sử dụng, cung cấp dữ liệu về pin xe điện và các tiêu chuẩn đánh giá. Phần 3 nghiên cứu ảnh hưởng của dữ liệu đầu vào đến chất lượng ước lượng SOC, kết quả ước lượng và nhận xét. Phần 4 là kết luận và hướng nghiên cứu tiếp theo.

2. Dữ liệu huấn luyện mạng, tiêu chuẩn đánh giá và cấu trúc mô hình

2.1. Dữ liệu huấn luyện mạng

Các tín hiệu đo lường của Pin bao gồm: Điện áp hai đầu cực (V), dòng điện (I), nhiệt độ (T) của pin được thu thập, để sử dụng là các đặc trưng đầu vào của các mạng Nơ ron. Trong tài liệu [14] dữ liệu pin LG 18650 HG2 được ghi lại theo các chu trình lái xe tiêu chuẩn. Các chu trình này thường được dùng để kiểm tra mức tiêu thụ nhiên liệu và lượng phát thải của xe hơi tại Hoa Kỳ, bao gồm: US06, LA92, UDDS và HWFET, ở 4 nhiệt độ khác nhau là 25°C, 10°C, 0°C, and −10°C . Ngoài ra, bộ dữ liệu hỗn hợp được ghi lại theo thứ tự sạc-xả là: UDDS-CC-LA92-CC-US06 với sự thay đổi nhiệt độ từ −10°C đến 25°C được tạo ra để kiểm chứng độ chính xác SOC trên nhiều nhiệt độ khác nhau. Mỗi chu trình lái xe được thực hiện xen kẽ với quá trình sạc pin theo phương pháp sạc CC-CV (Dòng không đổi - Điện áp không đổi).

Với mục đích nghiên cứu ảnh hưởng của dữ liệu đặc trưng đến kết quả ước lượng SOC, tác giả sử dụng thêm hai bộ lọc để tạo ra các đặc trưng mới. Bộ lọc trung bình động có giá trị cửa sổ cuộn ξ tạo ra điện áp trung bình (V_{avg}), dòng điện trung bình (I_{avg}) được thể hiện trong công thức (2), (3):

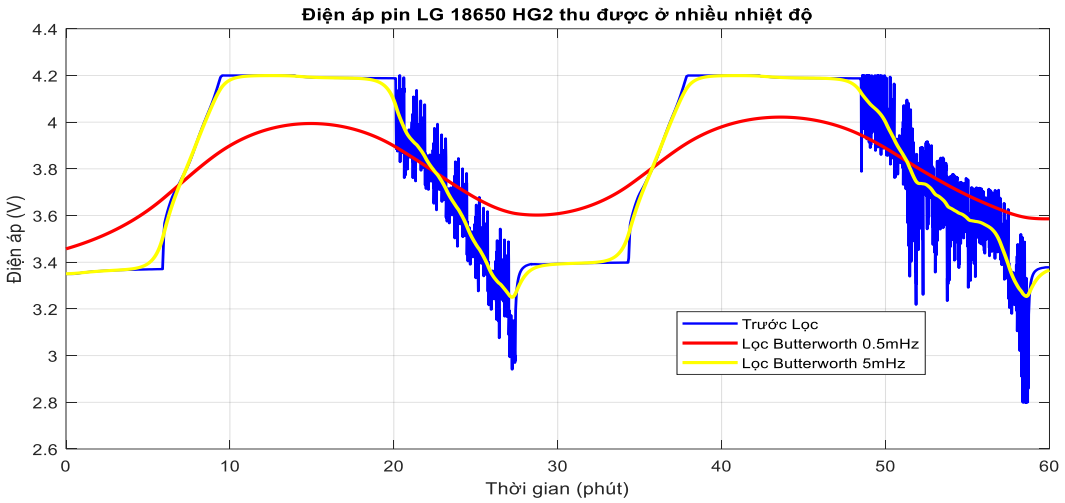
$$V_{avg}(k) = \frac{1}{\xi}\sum_{i=0}^{\xi-1} V(k-i) \tag{2}$$

$$I_{avg}(k) = \frac{1}{\xi}\sum_{i=0}^{\xi-1} I(k-i) \tag{3}$$

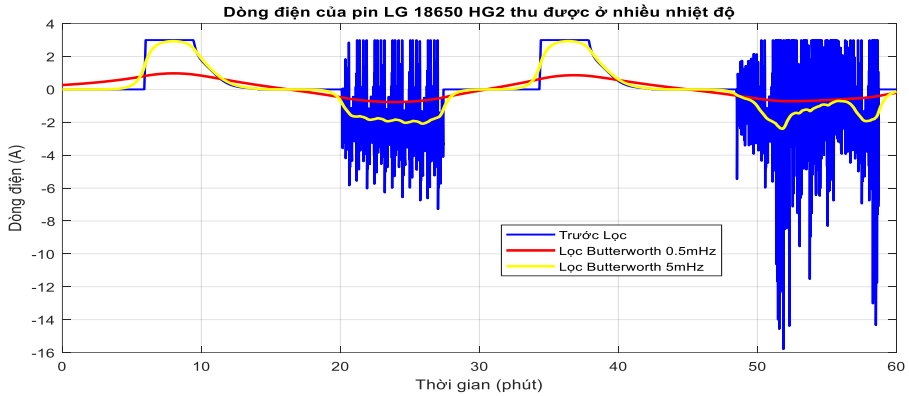
Bộ lọc thông thấp Butterworth bậc 1 với tần số lọc 0.5mHz và 5mHz tạo ra điện áp V_{filter}, dòng điện I_{filter}, với hàm truyền *H(f)* được tính theo công thức (4):

$$H(f) = \frac{1}{\sqrt{1+\varepsilon^2\left(\frac{f}{f_c}\right)^{2n}}} \tag{4}$$

Trong đó: f là tần số đầu vào, f_c là tần số cắt, ε là độ lợi băng thông tối đa (chọn ε=1), n là bậc của bộ lọc (n=1). Đặc điểm của dữ liệu điện áp được thể hiện trên Hình 2, của dữ liệu dòng điện thể hiện trên Hình 3.



Hình 2. Đồ thị các loại điện áp của pin LG 18650 HG2 sử dụng trong luyện mạng.



Hình 3. Dòng điện được sử dụng luyện mạng của pin LG-18650HG2.

Các giá trị đặc trưng đầu vào được chuẩn hóa sử dụng tiêu chuẩn Max-min, được tính bằng biểu thức (5):

$$x_i^* = \frac{x_i - X_{min}}{X_{max} - X_{min}} \tag{5}$$

Trong đó x_i^* là giá trị đặc trưng đầu vào sau khi dữ liệu được chuẩn hóa, X_{min} và X_{max} là giá trị đặc trưng nhỏ nhất và lớn nhất trong tập dữ liệu.

2.2. Tiêu chuẩn đánh giá và cấu trúc mô hình

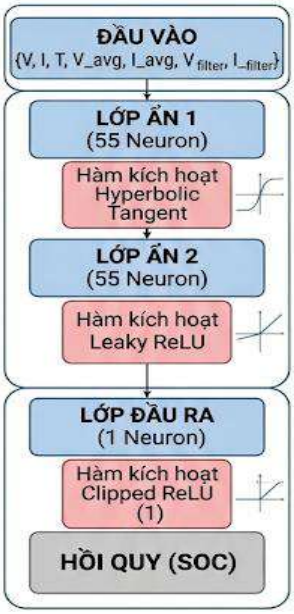
- *Tiêu chuẩn đánh giá*: Các giá trị SOC dự đoán được so sánh với các giá trị SOC thực tế trong bộ thử nghiệm bằng cách sử dụng các tiêu chuẩn sai số: RMSE, MAE, MSE với công thức cụ thể trong biểu thức (6):

$$\begin{cases} RMSE = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(SOC_{pre} - SOC_{act}\right)^2} \\ MAE = \frac{1}{N}\sum_{k=1}^{N}\left|SOC_{pre} - SOC_{act}\right| \\ MSE = \frac{1}{N}\sum_{k=1}^{N}\left(SOC_{pre} - SOC_{act}\right)^2 \end{cases} \tag{6}$$

Trong đó N là số mẫu dữ liệu, SOC_{pre} là giá trị SOC được dự đoán và SOC_{act} là giá trị SOC thực tế trong tập thử nghiệm. Sai số thu được từ công thức trên càng nhỏ thì độ chính xác ước lượng của mô hình càng cao.

- Cấu trúc mô hình: Trong tài liệu [15] tác giả đã sử dụng cấu trúc DNN gồm 2 lớp ẩn, số lượng Nơ ron trong mỗi lớp ẩn được tìm kiếm bằng phương pháp lặp, khảo sát khi số Nơ ron thay đổi từ 1 đến 100. Kết quả cho thấy, có 3 trường hợp cho kết quả tốt nhất ứng với với số Nơ ron của mỗi lớp là 55,82 và 99. Để đảm bảo tính hiệu quả trong quá trình huấn luyện, cấu trúc DNN được lựa chọn gồm 2 lớp ẩn, 55 nơ ron mỗi lớp. Lớp đầu vào là các tham số V, I, T và các dữ liệu bổ sung; sử dụng hàm Hyperbolic (tanhLayer) kết nối giữa 2 lớp ẩn; hàm Relu(0.03) kết nối giữa lớp ẩn thứ 2 với lớp đầu ra là giá trị SOC cần dự báo. Lớp đầu ra sử dụng hàm clipped ReluLayer(1). Cấu hình mạng DNN và giá trị tham số huấn luyện mạng DNN được cho trong Bảng 1.

Bảng 1. Cấu hình mạng DNN và giá trị tham số luyện mạng



Số epochs	500
Tốc độ học ban đầu	0.01
Số epoch giữa các lần giảm tốc độ học	100
Hệ số giảm tốc độ học	0.1
Số lớp	2
Số Nơ ron mỗi lớp	55
Các hàm phi tuyến	Hyperbolic Tangent và Leaky Relu cho các lớp ẩn Clipped relu cho lớp đầu ra
Số lần lặp	5

3. Ảnh hưởng của dữ liệu đầu vào đến chất lượng mô hình.

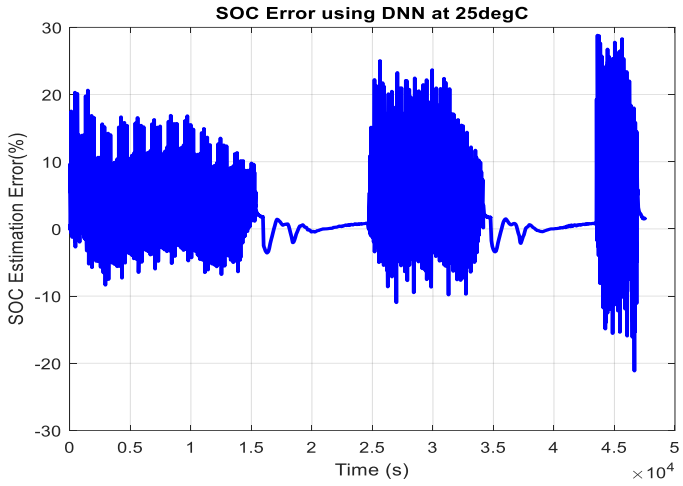
Trong phần này tác giả sẽ so sánh kết quả luyện mạng bằng việc sử dụng cùng một cấu trúc mạng bao gồm: Số lớp, số nơ ron trong mỗi lớp và các giá trị tham số học; khác nhau ở lớp đầu vào bao gồm: số lượng và giá trị tham số V, I, T. Với bộ dữ liệu thu được từ nhiều nhiệt độ (25°C, 10°C, 0°C, and −10°C), quá trình huấn luyện thực hiện với 5 lần lặp, các giá trị sai số được lấy giá trị trung bình. Chi tiết kết quả các trường hợp như sau:

Trường hợp 1: Khi đầu vào cấu trúc mạng có 3 đặc trưng là V, I, T.

Kết quả của quá trình luyện mạng được kiểm tra với bộ dữ liệu ở 25 ^{0}C; chỉ số sai số được thể hiện trên Bảng 2 và sai số ước lượng SOC được thể hiện trên Hình 4. Ta thấy rằng, khi nhiệt độ môi trường là −10°C, giá trị các sai số ước lượng SOC là lớn nhất. Khi nhiệt độ tăng dần, thì độ chính xác ước lượng SOC càng tăng.

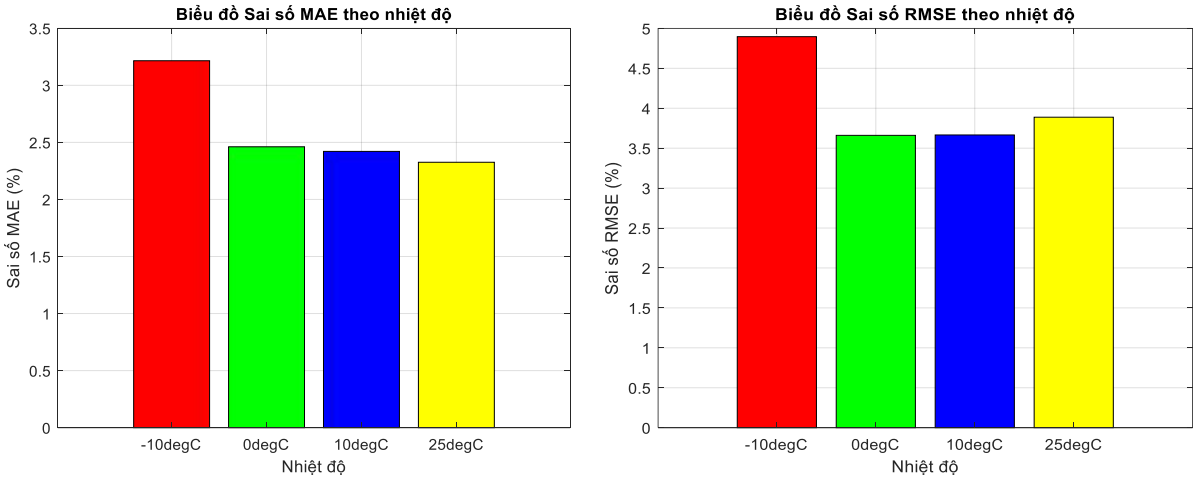
Bảng 2. Kết quả luyện mạng DNN với 3 tham số đầu vào V, I, T.

Nhiệt độ	MAE %	RMSE %	Max_error %	MSE %
−10°C	3.215	4.896	33.650	0.241
0°C	2.461	3.661	28.768	0.134
10°C	2.421	3.665	26.619	0.135
25°C	2.326	3.888	29.087	0.152



Hình 4. Sai số SOC giữa giá trị quan sát và giá trị ước lượng

Các chỉ số sai số khi được kiểm tra với bộ dữ liệu ở nhiều nhiệt độ khác nhau được trình bày trong Hình 5.



a) b)

Hình 5. Biểu đồ sai số sử dụng mạng DNN với 3 tham số đầu vào ở nhiều nhiệt độ

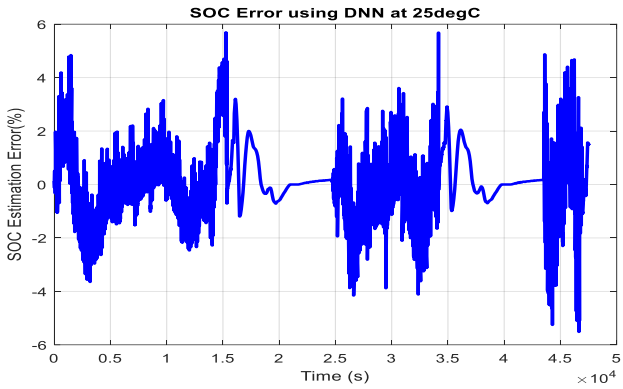
a) Sai số MAE b) Sai số RMSE

Trường hợp 2: Khi đầu vào mạng có 5 tham số đặc trưng là V, I, T, V_{avg} và I_{avg}.

Các giá trị V_{avg} và I_{avg} được tính từ các giá trị V, I tức thời với giá trị cửa số cuộn ξ=500. Kết quả huấn luyện mạng DNN trong trường hợp này được thể hiện trên Bảng 3, biểu đồ sai số thể hiện trên Hình 6.

Bảng 3. Kết quả luyện mạng DNN với 5 tham số đầu vào là V, I, T, V_{avg} và I_{avg}

Nhiệt độ	MAE %	RMSE %	Max_error %	MSE %
−10°C	1.60	2.15	9.66	0.05
0°C	1.45	2.00	8.71	0.05
10°C	1.42	2.01	8.07	0.04
25°C	1.19	1.62	7.52	0.03



Hình 6. Kết quả ước lượng SOC khi kiểm định dữ liệu ở 25 ^0C .

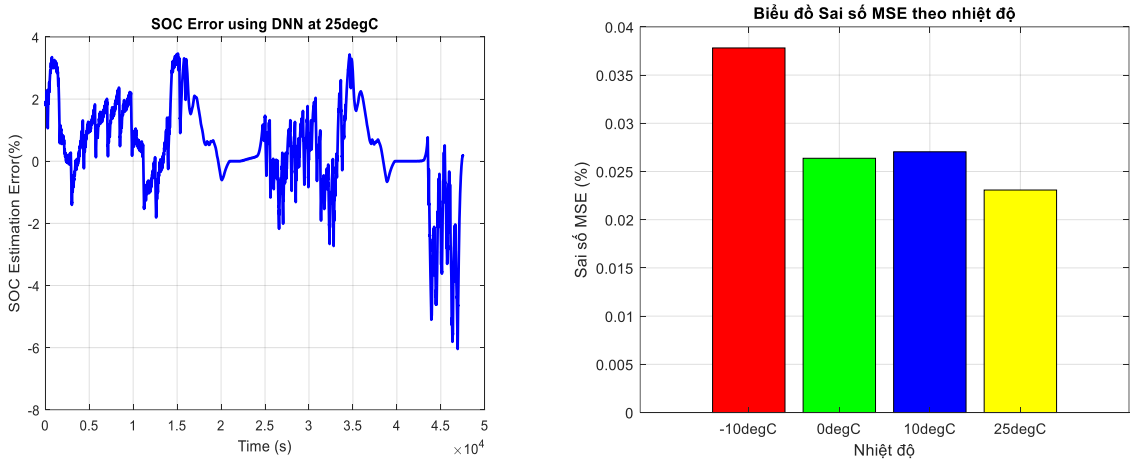
Ta có nhận xét rằng: Sai số ước lượng SOC ở nhiệt độ −10°C vẫn cao nhất so với sai số ở các nhiệt độ còn lại, và giá trị các sai số ước lượng SOC nhỏ hơn so với trong trường hợp 1. Điều này cho thấy, khi có thêm dữ liệu đầu vào V_{avg} và I_{avg} đã cải thiện hơn chất lượng ước lượng SOC. Điều này có thể được giải thích là do DNN không có bộ nhớ trong như các mạng LSTM, BiLSTM hay GRU nên sử dụng các giá trị trung bình là đầu vào, mang lại hiệu quả nhớ cho mạng với các dữ liệu quá khứ tốt hơn.

Trường hợp 3: Sử dụng dữ liệu sau bộ lọc thông thấp Butterworth bậc 1

Trong trường hợp này, tác giả không sử dụng giá trị V, I của dữ liệu ban đầu. Thay vào đó, sử dụng bộ lọc Butterworth bậc 1 ở các tần số lọc 0.5mHz và 5mHz tính toán ra các giá trị V, I tương ứng. Khi đó đầu vào mạng nơ ron DNN huấn luyện gồm 5 tham số: T, $V_{0.5mHz}$, $I_{0.5mHz}$, V_{5mHz} , I_{5mHz} . Kết quả luyện mạng được thể hiện trong Bảng 4, sai số SOC khi test ở nhiệt độ 25 ^0C và sai số MSE ở các nhiệt độ được minh họa trên Hình 7.

Bảng 4. Kết quả sử dụng mạng DNN với dữ liệu sau bộ lọc Butterworth bậc 1.

Nhiệt độ	MAE %	RMSE %	Max_error %	MSE %
−10°C	1.278	1.913	8.761	0.038
0°C	1.086	1.617	5.925	0.026
10°C	1.182	1.644	5.415	0.027
25°C	1.081	1.519	6.188	0.023



Hình 7. Sai số SOC ở 25^{0}C và MSE ở nhiều nhiệt độ khi sử dụng bộ lọc Butterworth bậc 1

Căn cứ vào kết quả nghiên cứu trong các trường hợp 1, 2 và 3 ta có nhận xét như sau:

- Trong cả 3 trường hợp, khi ước lượng SOC ở nhiệt độ -10^{0}C đều cho kết quả sai số cao nhất so với các nhiệt độ còn lại. Điều này được giải thích là do khi ở nhiệt độ thấp, độ sai lệch điện áp ở các điều kiện tải khác nhau là khá lớn; điều này làm cho việc ước lượng SOC gặp nhiều khó khăn.

- Khi chỉ sử dụng 3 đặc trưng V, I, T cho lớp đầu vào mạng DNN thì kết quả ước lượng không cao. Tất cả các loại sai số ở các nhiệt độ tương ứng đều cao hơn hai trường hợp còn lại. Nếu muốn cải thiện sự chính xác ước lượng, cần phải sử dụng cấu trúc mạng khác có khả năng nhớ dữ liệu trước đó (Có thể là LSTM)

- Thêm vào 2 tham số đặc trưng là V_{avg} và I_{avg} sẽ làm giảm đáng kể các loại sai số so với trường hợp có 3 đầu vào. Trong đó, các giá trị sai số Max có độ cải thiện mạnh mẽ nhất.

- Cùng với 5 đầu vào, nhưng sử dụng bộ lọc Butterworth bậc 1 như ở trường hợp 3 cho kết quả chính xác cao nhất. Cả 4 loại sai số, ở các nhiệt độ khác nhau đều giảm rất nhiều so với kết quả ở hai trường hợp 1 và 2.

Như vậy, việc xử lý dữ liệu trước khi đưa vào huấn luyện với kiểu mạng nơ ron truyền thẳng có vai trò quan trọng trong việc nâng cao độ chính xác quá trình ước lượng SOC của pin. Bộ lọc Butterworth bậc 1 được chứng minh là có hiệu quả tốt hơn bộ lọc trung bình động và tốt hơn hẳn khi không sử dụng bộ lọc. Vấn đề khó khăn ở đây là cần phải chọn ra những tần số lọc hợp lý cho bộ lọc và đó là hướng nghiên cứu tiếp theo của bài báo này.

4. Kết luận

Ước lượng chính xác SOC là bài toán quan trọng của hệ thống BMS, kỹ thuật học máy là công cụ mạnh mẽ, đưa ra kết quả ước lượng SOC chính xác mà không cần các mô hình toán học phức tạp. Mạng DNN trong bài báo này sử dụng dữ liệu đa dạng với 2 chu trình sạc, 3 chu trình xả ở các nhiệt độ khác nhau nên đã ước lượng SOC chính xác thông qua việc kiểm tra trên các tệp dữ liệu ở -10 ^{0}C, 0 ^{0}C, 10 ^{0}C và 20^{0}C.

Kết quả nghiên cứu trong bài báo cho thấy ảnh hưởng dữ liệu đầu vào đến kết quả ước lượng SOC của pin. Điện áp và dòng điện sau khi đã qua các bộ lọc được bổ sung vào lớp đầu vào mạng DNN đã cải thiện rất nhiều độ chính xác của mô hình. Bộ lọc Butterworth bậc 1 với tần số hợp lý cho kết quả chính xác hơn so với sử dụng bộ lọc trung bình. Hướng nghiên cứu tiếp theo của bài báo là xem xét ảnh hưởng của nhiễu đo lường đến tính bền vững của mô hình, kết hợp DNN với LSTM/CNN hoặc sử dụng Transformer-based models để cải thiện khả năng dự đoán SOC trong điều kiện thực tế.

Tài liệu tham khảo

[1] C. FV., "Battery and battery management for hybrid electric vehicles: a review," *Elektro Und Inf*, 2006.

[2] B. B. M. A. Watrin N, "Review of adaptive systems for lithium batteries State-of Charge," *2012 IEEE Transp Electrif Conf Expo*, 2012.

[3] J. C. Q. Y. H. H. a. F. S. R. Xiong, "Critical review on the battery state of charge estimation methods for electric vehicles,," *IEEE Access*, vol. vol. 6, p. 1832–1843, 2018.

[4] X. e. a. Hu, "Advanced Machine Learning Approach for Lithium-Ion Battery SOC Estimation,," *Applied Energy*, 2016..

[5] H. e. a. He, "State-of-Charge Estimation of Lithium-Ion Batteries Using EKF," *Energy Conversion and Management*, 2011.

[6] J. &. L. J. Zhang, "A Review on Prognostics and Health Monitoring of Li-ion Battery," *Journal of Power Sources*, 2011.

[7] W. e. a. Waag, "Experimental Investigation of Lithium-Ion Battery Aging,," *Journal of Power Sources,*, 2013..

[8] M. P. Z. S. Y. W. P. V. O. a. T. W. G. O. Sahinoglu, "Battery state-of-charge estimation based on regular/recurrent Gaussian process regression," *EEE Trans. Ind. Electron*, vol. 65, no. 5, p. 4311–4321, 2018.

[9] M. H. a. P. Stagge, "Recurrent neural networks for time series classification," *Neurocomputing*, vol. 50, p. 223–235, Jan. 2003.

[10] F. H. A. D. a. Y. S. A. Zeroual, "Deep learning methods for forecasting COVID-19 time-series data: A comparative study," *Chaos,Solitons Fractals*, vol. 140, Nov. 2020.

[11] Z. W. W. C. L.-Y. W. a. C. Q. J. Hong, "Online jointprediction of multi-forward-step battery SOC using LSTM neural networks and multiple linear regression for real-world electric vehicles," *J. Energy Storage,* vol. 30, p. 101459., Aug. 2020.

[12] H. H. a. S. Y. C. Bian, "Stacked bidirectional long short-term memory networks for state-of-charge estimation of lithium-ion batteries," *Energy,* vol. 191, p. Art. no. 116538., Jan. 2020.

[13] F. X. a. Y. F. C. Li, "An approach to state of charge estimation of lithium-ion batteries based on recurrent neural networks with gated recurrent unit," *Energies,* vol. vol. 12, no. no. 9, pp. 1592-1598, Apr. 2019..

[14] C. V. M. N. a. M. S. P. Kollmeyer, "LG 18650HG2 Li-ion Battery Data and Example Deep Neural Network xEV SOC Estimator Script," *Mendeley Data,* vol. V3, 2020.

[15] C. K. P. N. M. M. P. e. a. Vidal, "“Robust xEV Battery State-of-Charge Estimator Design Using a Feedforward Deep Neural Network,," *SAE Int. J. Advances & Curr. Prac. in Mobility 2,* vol. 5, no. 2, pp. 2872-2880, 2020.

Nghiên cứu ảnh hưởng dữ liệu đặc trưng đến kết quả ước lượng trạng thái SOC của pin xe điện sử dụng kỹ thuật học sâu.

Abstract: *Accurate estimation of the state of charge (SoC) plays a pivotal role in optimizing battery performance, assessing battery health, and ensuring reliable system operation. This study utilizes experimental data from the LG-HG2 lithium-ion battery under various driving cycles and ambient temperatures ranging from −10 °C to 25 °C. A deep neural network (DNN) architecture is employed to validate the model accuracy, while the impact of variations in input feature parameters on SoC estimation performance is systematically investigated. A moving average filter and a first-order Butterworth low-pass filter are applied to generate input features for network training. The results indicate that SoC estimation accuracy improves as battery temperature increases. The application of a first-order Butterworth filter with cutoff frequencies of 0.5 mHz and 5 mHz significantly enhances estimation accuracy, achieving an MSE of 0.023% at 25 °C, compared to 0.03% with the moving average filter and 0.152% without filtering.*

Keywords: SOC estimation, deep learning, electrical vehicle, neural network.

MỘT QUY TRÌNH HỌC MÁY ĐƠN GIẢN, HIỆU QUẢ ĐỂ PHÂN BIỆT PARKINSON SỬ DỤNG DỮ LIỆU THU ĐƯỢC TỪ ĐỒNG HỒ THÔNG MINH

Vũ Quân[1], Nguyễn Mạnh Cường[2], Trần Đức Tân[3]

[1]*Lớp NCS 42, H2, Học viện kỹ thuật quân sự, quanvu42@lqdtu.edu.vn*

[2]*Học viện kỹ thuật quân sự.*

[3]*Đại Học Phenikaa.*

Tóm tắt:

Người Bệnh Parkinson (PD) là một rối loạn thoái hóa thần kinh tiến triển, trong đó việc chẩn đoán sớm và đáng tin cậy vẫn còn nhiều thách thức, đặc biệt khi phân biệt PD với các rối loạn vận động khác (chẩn đoán phân biệt, DD). Các thiết bị đeo như đồng hồ thông minh cho phép theo dõi liên tục và không xâm lấn các triệu chứng vận động trong sinh hoạt hằng ngày, qua đó cung cấp một giải pháp có khả năng mở rộng thay thế cho các đánh giá tại phòng khám. Trong nghiên cứu này, chúng tôi tiến hành đánh giá hệ thống các tín hiệu quán tính từ gia tốc kế và con quay hồi chuyển trong bộ dữ liệu Parkinson's Activity Dataset (PADS), bao gồm 469 người tham gia thuộc ba nhóm: PD, nhóm chứng khỏe mạnh (HC) và nhóm DD. Các đặc trưng thủ công trong miền thời gian và miền tần số được trích xuất và tinh chọn bằng phương pháp lựa chọn đặc trưng dạng nhúng dựa trên LASSO. Khung xác thực chéo lồng nhau (5 vòng ngoài, 3 vòng trong) kết hợp với SMOTE để cân bằng lớp được áp dụng nhằm đảm bảo đánh giá khách quan và khả năng tái lập. Nhiều mô hình học máy có giám sát, bao gồm hồi quy Logistic (LR), máy vector hỗ trợ (SVM), mạng nhiều lớp (MLP), (RF), được đánh giá với các chỉ số chính là AUROC và hệ số tương quan Matthews (MCC). Trong bài toán phân loại PD và HC, mô hình LR đạt kết quả tốt nhất với AUROC 0,8977 ± 0,0429 và MCC 0,6191 ± 0,0714, trong khi các mô hình khác đạt AUROC xấp xỉ 0,88. Đối với bài toán khó hơn là PD và DD, hiệu năng giảm do sự chồng lấp triệu chứng; RF đạt MCC cao nhất (0,4016 ± 0,0953) với AUROC 0,7367 ± 0,0527. Kết quả cho thấy một pipeline dựa trên đặc trưng thủ công, đơn giản và có khả năng tái lập, vẫn có thể đạt hiệu năng phân biệt vững chắc, đặc biệt trong bài toán PD so với HC, đồng thời duy trì khả năng phân biệt có ý nghĩa trong bối cảnh chẩn đoán phân biệt thực tế. Phương pháp đề xuất khẳng định tính khả thi của việc sử dụng đồng hồ thông minh thương mại cho sàng lọc PD và cung cấp một chuẩn tham chiếu minh bạch cho các nghiên cứu và ứng dụng trong thực tế.

Từ khóa: *Parkinson's Disease, Action Tremor Detection, Wearable Device, Supervised-learning.*

1. Mở đầu

Bệnh Parkinson (Parkinson's disease – PD) là một rối loạn thoái hóa thần kinh trung ương mãn tính và tiến triển, ảnh hưởng đến hơn mười triệu người trên toàn thế giới [6]. PD đặc trưng bởi các triệu chứng vận động chính như chậm vận động, mất ổn định tư thế, cứng cơ, đóng băng dáng đi và run. Trong đó, run – đặc biệt là run ở tay và ngón tay – là một dấu hiệu lâm sàng điển hình, có thể gây ảnh hưởng đáng kể đến sinh hoạt hằng ngày và chất lượng cuộc sống của bệnh nhân [2].

Trong thực tế, việc chẩn đoán các rối loạn vận động như PD chủ yếu dựa trên khám lâm sàng. Tuy nhiên, các phân tích tổng hợp gần đây [1,8] cho thấy độ chính xác chẩn đoán còn hạn chế, dao động khoảng 73,8% đối với bác sĩ đa khoa và khoảng 79,6%–83,9% đối với các chuyên gia rối loạn vận động, tùy thuộc vào giai đoạn đánh giá. Điều này cho thấy tính ổn định và độ tin cậy của chẩn đoán lâm sàng chưa được cải thiện đáng kể theo thời gian, từ đó nhấn mạnh nhu cầu cấp thiết về các phương pháp định lượng khách quan, đặc biệt là các dấu ấn sinh học có thể được xác thực trên các bộ dữ liệu mở.

Trong những năm gần đây, sự phát triển của các thiết bị đeo thông minh, đặc biệt là các thiết bị tích hợp cảm biến quán tính như gia tốc kế và con quay hồi chuyển, đã mở ra khả năng theo dõi liên tục và không xâm lấn các biểu hiện vận động của bệnh nhân trong môi trường thực tế. Một số nghiên cứu cũng đã tận dụng các thiết bị này để hỗ trợ giảm run hoặc phân tích đặc điểm vận động [16]. Tuy nhiên, phần lớn các nghiên cứu hiện tại vẫn bị giới hạn bởi kích thước mẫu nhỏ (thường $n << 100$) và chủ yếu tập trung vào bài toán phân loại nhị phân giữa PD và người khỏe mạnh (HC) [14]. Trong khi đó, trong thực hành lâm sàng, bác sĩ thường phải đối mặt với bài toán chẩn đoán phân biệt giữa PD và các rối loạn vận động khác có triệu chứng tương tự, bao gồm run vô căn, parkinson không điển hình, hoặc các nguyên nhân thứ phát (gọi chung là DD). Các mô hình được xây dựng trên kịch bản PD vs HC thường không tổng quát tốt cho bài toán khó hơn này.

Bên cạnh đó, các nghiên cứu về phát hiện run cho thấy mặc dù một số phương pháp học sâu như CNN có thể đạt độ chính xác cao (trên 90%) trong việc nhận diện một số loại run cụ thể, nhưng hiệu suất tổng thể (đặc biệt là F1-score) vẫn chỉ đạt khoảng 70%–80% trong nhiều trường hợp thực tế. Điều này cho thấy việc xây dựng các đặc trưng có ý nghĩa sinh lý và có khả năng khái quát cao vẫn là một hướng nghiên cứu quan trọng.

Xuất phát từ những hạn chế trên, nghiên cứu này hướng tới việc xây dựng một phương pháp phân tích run dựa trên dữ liệu cảm biến quán tính từ thiết bị đeo, nhằm hỗ trợ phát hiện và phân biệt PD trong các kịch bản thực tế. Cụ thể, chúng tôi sử dụng bộ dữ liệu Parkinson's Activity Dataset (PADS) [15], bao gồm tín hiệu gia tốc và vận tốc góc thu thập tại cổ tay người tham gia. Bài toán được thiết kế dưới hai kịch bản: (i) phân biệt PD với nhóm chứng khỏe mạnh (HC), và (ii) phân biệt PD với các rối loạn vận động khác (DD).

Trong phương pháp đề xuất, các đặc trưng được trích xuất dựa trên kiến thức miền trong cả miền thời gian và miền tần số, nhằm phản ánh đặc tính dao động của run. Sau đó, thuật toán LASSO được sử dụng để lựa chọn các đặc trưng có khả năng phân biệt cao nhất. Hiệu quả của tập đặc trưng được đánh giá thông qua các mô hình học máy phổ biến, bao gồm Logistic Regression (LR), Support Vector Machine (SVM) và CatBoost (CB), với quy trình đánh giá nghiêm ngặt ở mức đối tượng.

Kết quả thực nghiệm cho thấy, đối với bài toán PD vs HC, các mô hình đạt hiệu suất cao với độ chính xác khoảng 85–88% và F1-score trên 0.90, trong đó Logistic Regression cho thấy hiệu năng ổn định và cao nhất. Điều này cho thấy không gian đặc trưng đề xuất có khả năng phân tách tốt ngay cả với các mô hình tuyến tính đơn giản. Đối với bài toán khó hơn PD vs DD, hiệu suất giảm xuống với balanced accuracy khoảng 65–69%, phản ánh mức độ chồng lấn giữa các nhóm bệnh. Tuy nhiên, kết quả này vẫn cho thấy tiềm năng của phương pháp trong việc hỗ trợ chẩn đoán phân biệt dựa trên dữ liệu thiết bị đeo.

2. Thực nghiệm

2.1 Tập dữ liệu PADS

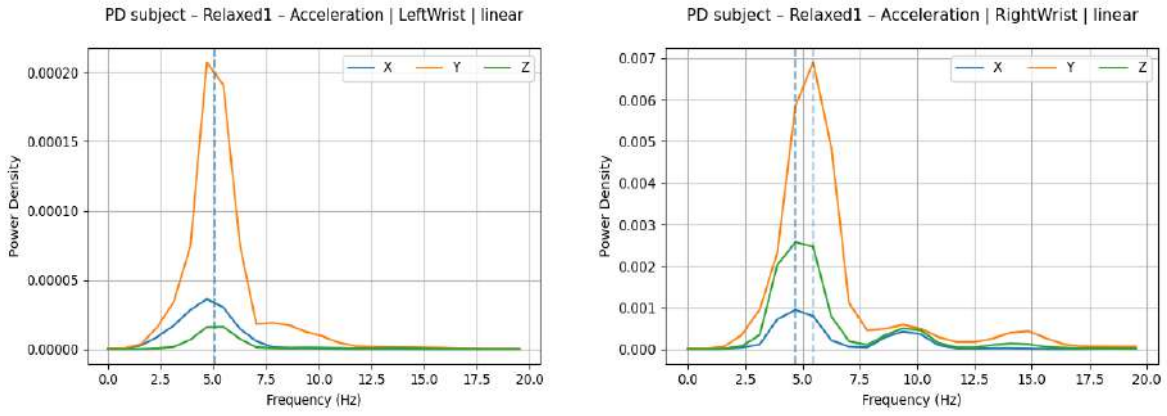
Bộ dữ liệu được tạo ra từ một nghiên cứu tiền cứu cắt ngang được thực hiện từ năm 2018 đến năm 2021 để nghiên cứu các dấu ấn sinh học kỹ thuật số của bệnh Parkinson (PD)[15] Tổng cộng, 5159 bước đo của 469 người đã được ghi nhận bao gồm người khỏe mạnh (HC), bệnh nhân PD và những người có các dấu hiệu khác (DD). Bộ dữ liệu PADS bao gồm: tín hiệu cảm biến gia tốc và vận tốc góc, cũng như chi tiết về các bước chuyển động, nhân khẩu học, tiền sử bệnh và các triệu chứng không vận động cụ thể của các nhóm.Trong mỗi bước đánh giá, dữ liệu được ghi nhận gồm: gia tốc và vận tốc góc trên ba trục không gian (x, y, z) với tần số lấy mẫu 100 Hz. Tổng cộng 11 nhiệm vụ đánh giá tương tác đã được thực hiện, ba trong số đó mất 20 giây và phần còn lại 10 giây, được tóm tắt trong Bảng 1. Đối

với phân tích ML, các bản ghi 20 giây được cắt thành hai phần, tạo ra 14 chuỗi thời gian 10 giây cho mỗi nhiệm vụ thử nghiệm. Với thiết lập này, sẽ thu được 1000 mẫu cho mỗi kênh dữ liệu theo thời gian, tổng số kênh của mỗi người tham gia (14 nhiệm vụ × 2 cánh tay × 2 cảm biến × 3 trục = 168). Trong các thí nghiệm sơ bộ, chúng tôi đã đánh giá riêng cảm biến: gia tốc, vận tốc góc, kết hợp cả gia tốc và vận tốc góc để kiểm tra xem hiệu năng của mô hình.

Bảng 1. Thông tin đối tượng nghiên cứu

Nhóm	Số lượng (đàn ông, phụ nữ)	Số tuổi trung bình	Tuổi chuẩn đoán có bệnh
Bệnh nhân (PD)	276 (195, 81)	65.4 (9.6)	58.26
Khỏe mạnh (HC)	79 (29,50)	62.9 (12.5)	
Chuẩn đoán phân biệt (DD): - Bệnh Parkinson không điển hình -Run cơ bản -Đa xơ cứng - Dấu hiệu khác	14(57,57) 15(8,7) 28(18,10) 11(7,4) 60(24,36)	62.4 (11.5) 66.0 (7.96) 66.18(10.84) 54.09(7.92) 61.32(12.32)	52.96 62.6(8.21) 46.96(22.89) 43.82(13.04) 55.02(13.32)

Tốc độ lấy mẫu cảm biến là 100Hz trong tập dữ liệu. Tập dữ liệu được chia thành các cửa sổ với thời gian 10 giây và kích thước bước 2,5 giây. Vì cảm biến có gia tốc kế ba trục và vận tốc góc ba trục, mỗi cửa sổ bao gồm 1000 * 6 điểm dữ liệu. Trong bài báo này chúng tôi chỉ tiến hành nghiên cứu đối với dữ liệu ghi nhận được từ đồng hồ thông minh, bỏ qua phần các câu hỏi lâm sàng.



a b

*Hình 1. Phổ công suất thể hiện rung động của một bệnh nhân Parkinson, **a**) Tay trái, **b**) tay phải*

2.2 Trích xuất và lựa chọn đặc tính

Trong các nghiên cứu với các mô hình cổ điển (ML) đã được báo cáo, độ chính xác chỉ khoảng 80%. Kết quả này đã thúc đẩy chúng tôi lựa chọn nhiều đặc trưng thủ công hơn. Do đó, chúng tôi đã trích xuất 49 đặc trưng từ cả miền thời gian và tần số. Mỗi một kênh được chia thành 4 phân đoạn để trích xuất các tính năng. Bảng 3 thể hiện các đặc trưng được trích xuất.

Bảng 2. Các đặc trưng được trích xuất

Miền	Đặc Tính	Công thức
Thời gian	MaxAbs,	$\text{MaxAbs}(x) = \max_{1 \le i \le N} \lvert x_i \rvert$

	Mean,	$\text{Mean}(x) = \frac{1}{N}\sum_{i=1}^{N} x_i$
	Median,	$\text{Median}(x) = \text{median}(x_1, x_2, ..., x_N)$
	Độ lệch chuẩn,	$\text{Std}(x) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$ $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$
	Sum Abs	$\text{SumAbs}(x) = \sum_{i=1}^{N} \lvert x_i \rvert$
	Energy	$\text{Energy}(x) = \sum_{i=1}^{N} x_i^2$
Tần số	Mật độ phổ công suất (PSD) được tính bằng phương pháp Welch trong khoảng 1Hz-19Hz.	$P_{xx}(f) = \frac{1}{K}\sum_{k=1}^{K}\frac{1}{LU}\left\lvert\sum_{n=0}^{L-1} x_k(n)w(n)e^{-j2\pi fn}\right\rvert^2$ $f \in [1,19]$ Hz
	Hàm $\log_{10}$ để nén giá trị PSD, giúp giảm độ lệch giữa các giá trị lớn và nhỏ trong phổ tín hiệu.	$\tilde{P}_{xx}(f) = \log\left(P_{xx}(f) + \varepsilon\right)$

Các đặc trưng miền thời gian bao gồm các giá trị: Biên độ tuyệt đối lớn nhất, tổng năng lượng tuyệt đối, trung bình, độ lệch chuẩn và trung vị. Các đặc trưng này được tính toán cho từng thành phần thô của gia tốc kế và vận tốc góc. Tương tự, đặc trưng miền tần số là: Mật độ phổ công suất (PSD) được tính bằng phương pháp Welch, với độ phân giải tần số là: . Thông tin về sự phân bố năng lượng của tín hiệu được lấy trên dải tần số 1Hz-19Hz,. Sau khi trích xuất các đặc trưng này, chúng tôi sử dụng các kỹ thuật chọn đặc trưng để chọn các đặc trưng tốt nhất để dự đoán run PD.

Lựa chọn đặc trưng có ý nghĩa cho việc giảm chi phí phân loại kích thước cao và giúp tránh quá tải. Thuật toán lựa chọn đặc trưng mà chúng tôi đã chọn trong phương pháp Hồi quy Logistic là Toán tử lựa chọn và co ngót tối thiểu tuyệt đối (LASSO).

LASSO là một thuật toán lựa chọn đặc trưng và chính quy hóa phổ biến cho các bộ phân loại học máy. Mô hình LASSO có ràng buộc về tổng các giá trị tuyệt đối của các tham số mô hình và tổng phải nhỏ hơn giới hạn trên. Phương pháp này áp dụng một quá trình thu nhỏ, còn được gọi là quá trình chính quy hóa, trong đó nó giảm thiểu trọng số của các đặc trưng hồi quy và buộc một số trong số chúng về không. Nếu các đặc trưng vẫn có trọng số khác bằng không sau quá trình thu nhỏ, chúng sẽ được chọn là một phần của mô hình.

LASSO bắt đầu với một mô hình hồi quy tuyến tính và có thể được biểu diễn như sau:

(1)

X_i là các đặc trưng đầu vào, Y là ma trận dự đoán và các tham số $\beta_0, \beta_1, \ldots, \beta_n$ là hệ số hồi quy. Hàm mất mát của hồi quy LASSO như sau:

(2)

Với [ ] là hệ số bổ xung LASSO cho hàm tối ưu hóa L1.

Nếu λ bằng không, thì nó sẽ là một vấn đề tối ưu hóa L1 phổ biến. Tuy nhiên, khi λ trở nên lớn hơn, LASSO sẽ thu nhỏ trọng số của các đặc trưng ít quan trọng hơn xuống 0, hoạt động hoàn hảo cho việc lựa chọn tính năng. Trong quá trình triển khai lựa chọn tính năng, chúng tôi chọn λ giúp giảm hoàn hảo trục Y của các đặc trưng tăng tốc và duy trì các đặc trưng trục hữu ích khác.

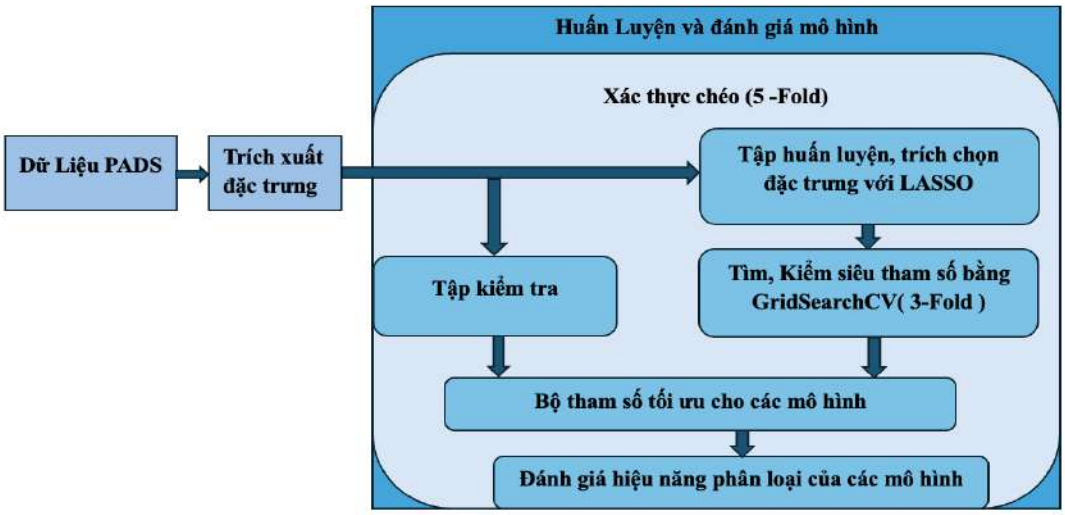
Trong miền thời gian các tính năng: giá trị tối đa và tối thiểu bị loại bỏ cho tất cả các trục. Những đặc trưng này bị loại bỏ vì các giá trị tối đa và tối thiểu thường không thể dự đoán được khi bệnh nhân thực hiện các hoạt động. Tổng cộng, các đặc trưng được chọn là: miền thời gian "5 × 4", miền tần số 46 trong dải tần từ 1Hz-19Hz. Trọng số (importance score) của các đặc trưng được đánh giá như sau:

[ ] (3)

Tổng cộng 25 đặc trưng có tổng trọng số là 54,96%, đại diện cho các đặc trưng quan trọng nhất. Tiếp theo, chúng tôi đưa tất cả các đặc trưng đã chọn vào các mô hình học tập có giám sát.

2.3 Thuật toán phân loại bệnh Parkinson

Thuật toán đề xuất bắt đầu từ dữ liệu PADS, trong đó các tín hiệu gia tốc và con quay hồi chuyển được trích xuất thành tập đặc trưng miền thời gian và miền tần số. Toàn bộ quy trình huấn luyện được thực hiện bằng nested cross-validation với 5-fold bên ngoài để đánh giá và 3-fold bên trong để tối ưu siêu tham số. Trong mỗi vòng lặp của tập huấn luyện, LASSO được áp dụng để lựa chọn các đặc trưng quan trọng, sau đó GridSearchCV kết hợp SMOTE được sử dụng để tinh chỉnh bộ phân loại (ví dụ SVM với kernel Linear hoặc RBF). Mô hình tối ưu được đánh giá trên tập kiểm tra độc lập của từng fold, và các chỉ số hiệu năng cuối cùng được báo cáo dưới dạng trung bình qua 5 lần lặp nhằm đảm bảo tính khách quan và khả năng tái lập của hệ thống. Các bước của thuật toán được minh hoạ như hình dưới đây.



Hình 2. Lưu đồ thuật toán dùng để nhận dạng bệnh Parkinson

3 Kết quả và thảo luận

3.1 Hiệu suất phân loại tổng thể

Chúng tôi đã tiến hành thực nghiệm đánh giá như sau: Hai nhiệm vụ phân loại đã được phân tích: (1) PD so với HC và (2) PD so với DD. Chúng tôi giả định rằng độ khó khi phân loại (1) thấp hơn nhiệm vụ phân loại (2) đối với các mô hình ML, vì nó chỉ cần phát hiện các đặc điểm không lành mạnh. PD so với HC được trình bày chủ yếu cho mục đích so sánh. Nhiệm vụ phân loại thứ hai yêu cầu phân tích đặc điểm nâng cao hơn để phân biệt các rối loạn vận động có đặc điểm kiểu hình rất giống nhau. Chúng tôi đã sử dụng bộ dữ liệu của bệnh nhân được đề cập trong phần trước. Các mô hình được huấn luyện và đánh giá với các trên cơ sở phần cứng và mền như sau: dual Xeon(R) 8171 Platinum 2.60GHz, 128Gb DDRam, GPU: Nivida Tesla P100-PCIE-16GB; pyhon 3.11.11, pytorch 2.5.1, cuda 11.8. Tiếp theo, chúng tôi trình bày kết quả đánh giá và trực quan hóa đặc trưng một cách chi tiết.

Nhiệm vụ A: PD vs HC

Đối với nhiệm vụ phân biệt PD và HC, thuật toán đề xuất đạt được hiệu xuất tốt và ổn định qua các mô hình. Trong đómo hình Hồi quy logistic (LR) đạt được độ chính xác cân bằng cao nhất là 79,26% ± 5,50, với độ chính xác tổng thể là 87,32% ± 1,99 và AUROC là 89,77% ± 4,29. Giá trị MCC tương ứng là 61,91 ± 7,14 cho thấy sự đồng đều đáng kể, ngoài sự ngẫu nhiên. Các mô hình SVM và MLP đã chứng minh hiệu suất tương đương, đạt được độ chính xác cân bằng lần lượt là 73,85% ± 3,85% và 75,46% ± 8,43. Mô hình (RF) mang lại độ chính xác cân bằng thấp hơn một chút (67,85% ± 4,26), cho thấy rằng mô hình tổng hợp phi tuyến không mang lại lợi ích bổ sung so với ranh giới quyết định tuyến tính đơn giản hơn trong không gian đặc trưng này. Điều quan trọng là độ lệch chuẩn tương đối nhỏ giữa các nếp gấp (Fold), đặc biệt là đối với LR và SVM cho thấy sự khái quát hóa nhất quán trong sự tách biệt ở cấp độ đối tượng. Những phát hiện này cho thấy rằng các đặc điểm thống kê và quang phổ được thiết kế cung cấp đủ cấu trúc phân biệt mà không yêu cầu các mô hình phức tạp cao.

Bảng 3. Kết quả hiệu xuất phân loại giữa PD và HC

Mô hình	Acc(%)	Bacc(%)	F1-score	AUROC	MCC
LR	87.32 ± 1.99	79.26 ± 5.50	92.01 ± 1.16	89.77 ± 4.29	61.91 ± 7.14
SVM	84.50 ± 1.72	73.85 ± 3.85	90.31 ± 1.23	88.65 ± 4.73	52.86 ± 4.51
MLP	85.63 ± 4.27	75.46 ± 8.43	91.05 ± 2.59	88.40 ± 3.71	55.52 ± 14.3
RF	83.66 ± 3.09	67.85 ± 4.26	90.15 ± 1.99	86.81 ± 5.79	47.35 ± 9.3

Nhiệm vụ B: PD vs DD

Nhiệm vụ phân loại PD và DD khó hơn so với việc phân biệt với nhóm đối chứng khoẻ mạnh. Các giá trị độ chính xác cân bằng dao động từ 65,58% ± 3,28 (LR) đến 67,15% ± 2,76 (MLP). Giá trị AUROC vẫn nằm trong phạm vi trung bình (khoảng 72–74%) và giá trị MCC thấp hơn đáng kể so với Nhiệm vụ A, phản ánh giảm khả năng phân tách giữa PD và các rối loạn vận động khác. Điều thú vị là sự khác biệt về hiệu suất giữa các mô hình là tối thiểu trong nhiệm vụ này. Mô hình Random Forest đạt được độ chính xác tổng thể cao nhất (77,43% ± 3,22) và MCC (40,16 ± 9,53), nhưng độ chính xác cân bằng vẫn có thể so sánh với các mô hình khác. Sự hội tụ này cho thấy rằng dung lượng mô hình không phải là yếu tố hạn

chế chính; thay vào đó, sự chồng chéo nội tại trong các đặc điểm vận động giữa PD và chẩn đoán phân biệt có thể hạn chế hiệu suất phân loại.

Bảng 4. Kết quả hiệu xuất phân loại giữa PD và DD

Mô hình	Acc(%)	Bacc(%)	F1-score	AUROC	MCC
LR	73.58 ± 2.50	65.58 ± 3.28	50.38 ± 5.45	72.67 ± 2.34	33.37 ± 6.16
SVM	73.33 ± 3.06	66.18 ± 3.41	51.67 ± 5.25	72.95 ± 3.19	33.75 ± 6.83
MLP	73.58 ± 2.66	67.15 ± 2.76	53.32± 4.32	72.47 ± 1.93	35.17 ± 6.06
RF	77.43 ± 3.22	66.26 ± 4.32	50.25 ±. 8.04	73.67 ± 5.27	40.16 ± 9.53

Trong các nhiệm vụ, sự thay đổi hiệu suất vẫn ở mức vừa phải, cho thấy sự khái quát nhất quán trong sự tách biệt cấp độ chủ đề. Điều quan trọng là tất cả các kết quả thu được mà không kết hợp bảng câu hỏi hoặc dữ liệu triệu chứng không vận động. Độ lệch chuẩn nhỏ hơn trong Nhiệm vụ B chỉ ra rằng, mặc dù hiệu suất tuyệt đối thấp hơn, qua các vòng kiểm tra vẫn ổn định và không thể hiện phương sai cao giữa các vòng. Một khoảng cách hiệu suất rõ ràng được quan sát thấy giữa phân loại PD so với HC và PD và DD. Trong khi các đặc điểm liên quan đến run và các biện pháp bất đối xứng cung cấp khả năng phân biệt mạnh mẽ giữa PD và những người khỏe mạnh, việc phân biệt PD với các rối loạn vận động khác về bản chất vẫn khó khăn hơn do các kiểu hình vận động chồng chéo. Đáng chú ý, mô hình Hồi quy Logistic đạt được hiệu suất tổng thể mạnh nhất trong số các mô hình cổ điển được đánh giá trong Nhiệm vụ A, trong khi SVM và MLP tạo ra kết quả tương đương trong biên độ vừa phải. Điều này cho thấy rằng không gian đặc trưng đã chọn cung cấp thông tin phân biệt đáng kể. Bên cạnh đó độ phức tạp của các mô hình cổ điển được thử nghiệm cũng thể hiện lợi ích bổ sung hạn chế trong cùng một thuật toán.

3.2 So sánh với các nghiên cứu trên bộ dữ liệu PADS

Bảng 5 trình bày so sánh giữa phương pháp đề xuất và các nghiên cứu trước đây trên cùng bộ dữ liệu PADS dưới cùng giao thức đánh giá nested cross-validation ở mức đối tượng. Đối với bài toán PD vs HC, phương pháp đề xuất đạt balanced accuracy 79,26%, gần tương đương với kết quả 78,99% được báo cáo bởi Varghese et al. (2021), cho thấy tính nhất quán và khả năng tái lập của khung phương pháp khi sử dụng dữ liệu chuyển động đơn thuần.

Bảng 5. So sánh với các nghiên cứu cùng bộ dữ liệu PADS trước đây

Nghiên cứu	Dữ liệu đầu vào	Lựa chọn đặc trưng	Phương thức xác thực	PD vs HC (BAcc %)	PD vs DD (BAcc %)
Varghese et al. [19]	Dữ liệu cảm biến	Phương pháp tích hợp	Nested CV (subject-level)	78.99	69.18
Varghese et al. [20]	Dữ liệu cảm biến và lâm sàng	Phương pháp tích hợp	Nested CV (subject-level)	91.16	72.42
Phương pháp đề xuất	Dữ liệu cảm biến	LASSO (nested CV)	Nested CV (subject-level)	**79.26**	**67.15**

So với phương pháp đa phương thức của Varghese et al. (2024), trong đó kết hợp dữ liệu cảm biến với thông tin lâm sàng và bảng câu hỏi để đạt hiệu suất cao hơn (91,16%), phương pháp đề xuất chỉ sử dụng dữ liệu chuyển động từ thiết bị đeo. Điều này giúp giảm sự phụ thuộc vào siêu dữ liệu lâm

sàng, qua đó tăng tính khả thi triển khai và khả năng mở rộng trong các hệ thống giám sát sức khỏe ngoài môi trường phòng khám.

Đối với bài toán PD vs DD, phương pháp đề xuất đạt balanced accuracy 67,15%, ở mức tương đương với các phương pháp chỉ sử dụng dữ liệu vận động trước đây. Kết quả này tiếp tục khẳng định tính chất khó của bài toán chẩn đoán phân biệt, khi các rối loạn vận động khác có biểu hiện tương đồng với PD, dẫn đến sự chồng lấn đáng kể trong không gian đặc trưng dựa trên tín hiệu chuyển động.

Bên cạnh các phương pháp học máy truyền thống, nhiều nghiên cứu gần đây đã áp dụng các mô hình học sâu như mạng nơ-ron tích chập (CNN) [23]hoặc các kiến trúc lai CNN–LSTM [22] để phân tích tín hiệu vận động. Các phương pháp này thường đạt hiệu suất cao trong các kịch bản đơn giản như PD vs HC, đặc biệt khi sử dụng dữ liệu lớn và huấn luyện trực tiếp trên tín hiệu thô. Tuy nhiên, hiệu quả của các mô hình học sâu có xu hướng giảm trong các bài toán khó hơn như PD vs DD, cũng như trong các trường hợp dữ liệu hạn chế hoặc không đồng nhất.

Trong bối cảnh đó, kết quả của nghiên cứu này cho thấy rằng, với cùng một giao thức đánh giá nghiêm ngặt, các đặc trưng thủ công được thiết kế dựa trên kiến thức miền và được lựa chọn cẩn thận vẫn có thể đạt hiệu suất cạnh tranh mà không cần đến các kiến trúc học sâu phức tạp. Đặc biệt, hiệu năng ổn định của Logistic Regression cho thấy không gian đặc trưng đề xuất có cấu trúc phù hợp, cho phép đơn giản hóa bài toán phân loại mà không làm suy giảm đáng kể hiệu năng.

3.3 Thảo luận

Bài báo của chúng tôi đã sử dụng dữ liệu thu được từ các cảm biến chuyển động trên đồng hồ thông minh. Những cảm biến này rất thiết thực và linh hoạt cho bệnh nhân PD sử dụng trong sinh hoạt hàng ngày. Tuy nhiên, rất khó để các cảm biến có thể dự báo các sự kiện chấn động. Cảm biến chỉ có thể phát hiện tín hiệu khi rung động đã xảy ra. Nghiên cứu trong tương lai có thể xem xét cách dự đoán run tay và các triệu chứng PD khác bằng cách kết hợp với các cảm biến EMG.

Dữ liệu chúng tôi sử dụng được ghi nhận trong môi trường lâm sàng. Dữ liệu dùng trong nghiên cứu này mới chỉ khai thác phần chuyển động của các đối tượng tham gia, chưa khai thác phần các câu hỏi cũng như thông tin lâm sàng liên quan, nhưng kết quả thể hiện trong các mô hình cũng khá thuyết phục về độ chính xác, cũng như độ tin cậy cao. Trong tương lai chúng tôi đặt mục tiêu phát triển một hệ thống trực tuyến có thể phát hiện run tay trong thời gian thực với các thiết bị đeo được và điện thoại thông minh, đồng thời tiến hành nghiên cứu thí điểm để kiểm tra độ bền và độ chính xác của hệ thống trong các môi trường khác nhau.

Bài báo của chúng tôi chọn các mô hình học máy có giám sát vì bộ dữ liệu đã gắn nhãn các sự kiện run hoạt động từ các hoạt động dựa trên các đánh giá trong bộ dữ liệu. Nghiên cứu trong tương lai có thể tập trung vào các mô hình học tập và học sâu không giám sát. Những mô hình này có khả năng nâng cao hiệu suất phát hiện run hoạt động.

Như chúng tôi đã đề cập trong phần tổng quan, phân biệt được giữa bệnh nhân, người có triệu chứng (DD) là một vấn đề khó, nhưng kết quả bước đầu dựa trên các mô hình đơn giản trình bày trong nghiên cứu này đưa đến hướng triển khai hứa hẹn trong thực tế. Sử dụng đồng hồ thông minh có thể phân biệt được PD, HC, DD. Việc triển khai các mô hình học sâu trên nền tảng mobiphone có nhiều thách thức, nhưng mô hình đơn giản như LR hoàn toàn khả quan. Bệnh nhân PD có thể sử dụng đồng hồ thông minh để theo dõi sự thay đổi hàng ngày về mức độ “run” trong sinh hoạt. Ngoài ra, việc phát hiện run chính xác, bằng các thiết bị đeo

được còn có ý nghĩa giúp sự chăm sóc kịp thời, hiệu quả hơn là những lần khám bệnh định kỳ. Các kỹ thuật phát hiện rung động của chúng tôi sẽ được sử dụng trong các nghiên cứu trong tương lai để tạo ra các thiết bị đeo được này.

4. Kết luận

Nghiên cứu này tập trung vào phát hiện bệnh Parkinson trong bối cảnh lâm sàng thực tế, bao gồm cả nhóm chứng khỏe mạnh và nhóm có triệu chứng tương tự, dựa trên tín hiệu gia tốc kế và con quay hồi chuyển thu thập từ thiết bị đeo. Chúng tôi đã xây dựng một quy trình xử lý dữ liệu đơn giản và có khả năng tái lập, trong đó các đặc trưng thủ công được thiết kế dựa trên kiến thức miền ở cả miền thời gian và miền tần số, sau đó được tinh chọn bằng phương pháp lựa chọn đặc trưng dạng nhúng dựa trên LASSO nhằm xác định không gian đặc trưng phân biệt ổn định.

Các đặc trưng được đánh giá trên bốn mô hình học máy có giám sát gồm Hồi quy Logistic, SVM, MLP và RF trong khung xác thực chéo lồng nhau nhằm đảm bảo đánh giá khách quan. Trong bài toán PD so với nhóm chứng khỏe mạnh, mô hình Hồi quy Logistic đạt kết quả tốt nhất với AUROC xấp xỉ 0,90 và MCC khoảng 0,62, cho thấy khả năng phân biệt mạnh mẽ và ổn định. Đối với bài toán khó hơn là PD so với nhóm chẩn đoán phân biệt, hiệu năng giảm do sự chồng lấp triệu chứng; tuy nhiên, các mô hình vẫn đạt AUROC khoảng 0,73 và MCC tối đa khoảng 0,40 phản ánh khả năng phân loại có ý nghĩa trong bối cảnh lâm sàng thực tế.

Những kết quả này cho thấy rằng một pipeline dựa trên đặc trưng thủ công, được thiết kế hợp lý và kết hợp lựa chọn đặc trưng phù hợp, có thể đạt hiệu năng đáng tin cậy mà không cần đến các kiến trúc học sâu phức tạp. Phương pháp đề xuất góp phần cung cấp một khung tham chiếu minh bạch và có khả năng tái lập cho các nghiên cứu phát hiện PD dựa trên thiết bị đeo trong tương lai.

Tài liệu tham khảo

1. Ahmed Nasimuddin, Bhattacharyya Chirayata, and Ghose Avik. 2021. A Novel Non-Parametric Approach Of Tremor Detection Using Wrist-Based Photoplethysmograph. In 2020 28th European Signal Processing Conference (EUSIPCO). IEEE, 1150–1154.
2. Anouti Ahmad and Koller William C. 1995. Tremor disorders. Diagnosis and management. Western journal of medicine 162, 6 (1995), 510. [PubMed: 7618310].
3. Braga Diogo, Madureira Ana M, Coelho Luis, and Ajith Reuel. 2019. Automatic detection of Parkinson's disease based on acoustic analysis of speech. Engineering Applications of Artificial Intelligence 77 (2019), 148–158.
4. Chen Shih-Yuan and Lin Chi-Lun. 2022. Subtle Motion Detection Using Wi-Fi for Hand Rest Tremor in Parkinson's Disease. In 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 1774–1777.
5. Gigante, A.F.; Pellicciari, R.; Iliceto, G.; Liuzzi, D.; Mancino, P.V.; Custodero, G.E.; Guido, M.; Livrea, P.; Defazio, G. Rest tremor
in Parkinson's disease: Body distribution and time of appearance. J. Neurol. Sci. 2017, 375, 215–219, doi:10.1016/j.jns.2016.12.057.
6. Ibrahim Anas, Zhou Yue, Jenkins Mary E, Trejos Ana Luisa, and Naish Michael D. 2020. The design of a Parkinson's tremor predictor and estimator using a hybrid convolutional multilayer

perceptron neural network. In 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 5996–6000.

7. Javor Andrija, Ransmayr Gerhard, Struhal Walter, and Riedl René. 2016. Parkinson patients' initial trust in avatars: theory and evidence. PloS one 11, 11 (2016), e0165998. [PubMed: 27820864].
8. Krack Paul, Benazzouz Abdelhamid, Pollak Pierre, Limousin Patricia, Piallat Brigitte, Hoffmann Dominique, Xie Jing, and Benabid Alim-Louis. 1998. Treatment of tremor in Parkinson's disease by subthalamic nucleus stimulation. Movement disorders: official journal of the Movement Disorder Society 13, 6 (1998), 907–914. [PubMed: 9827614].
9. Marzieh Keshtkarjahromi , Danielle S Abraham , Ann L Gruber-Baldini , Katrina Schrader , Stephen G Reich , Joseph M Savitt , Rainer Von Coelln , Lisa M Shulman. Confirming Parkinson Disease Diagnosis: Patterns of Diagnostic Changes by Movement Disorder Specialists 2022 May 9;2022:5535826. doi: 10.1155/2022/5535826
10. Moro-Velazquez Laureano, Gomez-Garcia Jorge A, Arias-Londoño Julian D, Dehak Najim, and Godino-Llorente Juan I. 2021. Advances in Parkinson's disease detection and assessment using voice and speech: A review of the articulatory and phonatory aspects. Biomedical Signal Processing and Control 66 (2021), 102418.
11. Nalini M, Gayathiri R, Srimathi R, Vidyathmikaa R, and Jenifer S. 2022. Detection of Parkinson's Disease Using Voice Changes and Hand-tremor. In 2022 International Conference on Communication, Computing and Internet of Things (IC3IoT). IEEE, 1–4.
12. Politis Marios, Wu Kit, Molloy Sophie, Bain Peter G., Chaudhuri K Ray, and Piccini Paola. 2010. Parkinson's disease symptoms: the patient's perspective. Movement Disorders 25, 11 (2010), 1646–1651. [PubMed: 20629164]
13. Rajaraman V, Jack D, Adamovich SV, Hening W, Sage Jacob, and Poizner H. 2000. A novel quantitative method for 3D measurement of Parkinsonian tremor. Clinical neurophysiology 111, 2 (2000), 338–343. [PubMed: 10680570]
14. Rini Eka Mistiko and Haq Endi Sailul. 2021. Detection Hand Tremor Through Each Finger Movement Based On Arduino For Parkinson's Patients. In 2021 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE). IEEE, 225–230.
15. Varghese, J., Brenner, A., Plagwitz, L., van Alen, C., Fujarski, M., & Warnecke, T. (2024). PADS-Parkinsons Disease Smartwatch dataset (version1.0.0). *PhysioNet*. https://doi.org/10.13026/m0w9-zx22
16. Sun Minglong, Watson Amanda, and Zhou Gang. 2020. Wearable computing of Freezing of Gait in Parkinson's disease: A survey. Smart Health 18 (2020), 100143.
17. Swallow Lee and Siores Elias. 2009. Tremor suppression using smart textile fibre systems. Journal of Fiber Bioengineering and Informatics 1, 4 (2009), 261–266.
18. Tong Lina, He Jiaji, and Peng Liang. 2021. CNN-based PD hand tremor detection using inertial sensors. IEEE Sensors Letters 5, 7 (2021), 1–4. [PubMed: 36789370]
19. J. Varghese, A. Brenner, M. Fujarski, C. M. van Alen, L. Plagwitz, and T. Warnecke, "Machine learning in the Parkinson's disease smartwatch (PADS) dataset," npj Parkinson's Disease, vol. 10, no. 1, p. 9, 2024.
20. J. Varghese, C. M. V. Alen, M. Fujarski, G. S. Schlake, J. Sucker, T. Warnecke, et al., "Sensor validation and diagnostic potential of smartwatches in movement disorders," Sensors, vol. 21, no. 9, p. 3139, 2021.

21. Luis Sigcha , Luigi Borzì , Federica Amato , ... Deep learning and wearable sensors for the diagnosis and monitoring of Parkinson's disease: A systematic review Expert Systems with Applications Volume 229, Part A, 1 November 2023, 120541 https://doi.org/10.1016/j.eswa.2023.120541
22. Alexandros Papadopoulos, Dimitrios Iakovakis, Lisa Klingelhoefer... Unobtrusive detection of Parkinson's disease from multi-modal and in-the-wild sensor data using deep learning techniques. Scientific Reports | (2020) 10:21370 | https://doi.org/10.1038/s41598-020-78418-8
23. Ming long Sun, Woosub Jung, Kenneth Koltermann, Gang Zhou, Amanda Watson, Ginamari Blackwell, Noah Helm, Leslie Cloud, and Ingrid Pretzer Aboff. 2023. Parkinson's Disease Action Tremor Detection with Supervised Leaning Models. June 21 23, 2023, https: //doi.org/10.1145/3580252.3586977
24. Y. K. Meng, T. Liang, H. Liu, and X. Wang, "Improved deep learning for Parkinson's diagnosis based on wearable sensors," Electronics, vol. 13, no. 21, p. 4638, 2024.
25. M. Wang, J. Chen, J. Du, Y. Wu, J. Mou, and D. Camacho, "Parkinson's disease detection using multi-scale frequency-sharing channel attention network with smartwatch movement recordings," IEEE Internet of Things Journal, vol. 12, no. 9, pp. 14567–14577, 2025.

A Simple and Reproducible Machine Learning Pipeline for Parkinson's Disease Differentiation Using Smartwatch-Based Inertial Signals

Abstract: Parkinson's disease (PD) is a progressive neurodegenerative disorder for which early and reliable diagnosis remains challenging, particularly when differentiating PD from other movement disorders (differential diagnosis, DD). Wearable devices such as smartwatches enable continuous and non-invasive monitoring of motor symptoms in daily life, providing a scalable alternative to clinic-based assessments.In this study, we conducted a systematic evaluation of inertial signals from accelerometers and gyroscopes in the Parkinson's Activity Dataset (PADS), comprising 469 participants across three groups: PD, healthy controls (HC), and DD. Handcrafted features in both the time and frequency domains were extracted and refined using an embedded LASSO-based feature selection approach. A nested cross-validation framework (5 outer folds and 3 inner folds), combined with SMOTE for class balancing, was employed to ensure unbiased evaluation and reproducibility. Multiple supervised machine learning models, including logistic regression (LR), support vector machines (SVM), multilayer perceptron (MLP), and random forest (RF), were evaluated using the Area Under the Receiver Operating Characteristic Curve (AUROC) and the Matthews Correlation Coefficient (MCC) as primary performance metrics. For the PD versus HC classification task, LR achieved the best performance with an AUROC of 0.8977 ± 0.0429 and an MCC of 0.6191 ± 0.0714, while other models achieved AUROC values around 0.88. In the more challenging PD versus DD task, performance decreased due to symptom overlap; RF obtained the highest MCC (0.4016 ± 0.0953) with an AUROC of 0.7367 ± 0.0527.These findings demonstrate that a simple and reproducible feature-based pipeline can achieve robust discriminative performance, particularly for PD versus HC, while maintaining meaningful differentiation in clinically realistic PD versus DD scenarios. The proposed approach confirms the feasibility of using consumer-grade smartwatches for scalable PD screening and provides a transparent benchmark for future research and real-world applications.

Keywords: *Parkinson's Disease, Action Tremor Detection, Wearable Device, Supervised_learning*

Robust Control Design for a Three-Axis Gimbal System Based on Adaptive Sliding Mode Control Law

Van-Tai Quach[1], Ngoc-Hung Nguyen[1]

[1]*Military Technical Academy*

*Email: quachchikaki@gmail.com; Tel: 0367896782

Abstract:

Three-axis gimbal systems are widely used to stabilize and control the line-of-sight of optical payloads under external disturbances and platform motions. The main task is to isolate the angular motion of the payload and the disturbances affecting the optical axis of the camera while maintaining the camera's orientation locked onto the target. In this paper, an in-depth study is conducted to develop the full mathematical model of a three-axis gimbal system, thus allowing a comprehensive assessment of how cross-channel disturbance moments resulting from gimbal imbalance affect the stability and control accuracy under complex payload angular motion conditions. The present work presents a robust control design for a three-axis gimbal system with an adaptive reaching law. Simulation results obtained in MATLAB/Simulink demonstrate that the proposed control approach effectively suppresses cross-axis coupling effects, reduces disturbance influence, and improves stabilization and control accuracy under complex motion conditions.

Key-word: *three-axis gimbal, inertially stabilized platform, sliding mode control, adaptive.*

1. Introduction

In the current landscape, the rapid advancement of military science and technology has prompted nations to continuously invest in the research and development of next-generation weapon systems characterized by high precision, rapid response, and enhanced operational effectiveness. Among these, missiles equipped with autonomous seekers play a particularly critical role, serving as primary engagement assets in modern offensive and defensive weapon systems.

The seeker constitutes the core functional module that enables a missile to autonomously detect, acquire, track, and engage a target, thereby minimizing direct human intervention and improving mission performance under high-dynamics and high-speed operational conditions. Consequently, the development of an accurate dynamic model for the stabilization system and line-of-sight (LOS) control loop is essential. Such a model supports a deeper understanding of the dominant factors affecting target tracking and provides a rigorous theoretical foundation for designing control algorithms in subsequent stages, ultimately improving the accuracy and reliability of the overall weapon system.

One of the most critical subsystems within a missile seeker is the LOS stabilization and control subsystem. However, constructing a dynamic model for this subsystem is inherently complex due to the combined influence of multiple technical factors and the operational environment. A seeker system typically integrates not only mechanical components (e.g., gimballed platform, servo actuators, control surfaces) but also integrates electronics, sensors, and control algorithms. As a result, the governing equations are often high-order, strongly nonlinear, and multivariable coupled.

State variables such as the LOS angle λ, the LOS angular velocity $\dot{\lambda}$, and the missile body attitude error are tightly coupled, which makes it difficult to decouple the system into independent blocks for modular modeling. Furthermore, the seeker dynamics depend on a set of key parameters, including the principal moments of inertia J_x, J_y, J_z, as well as

aerodynamic coefficients (e.g., lift and drag coefficients). These parameters vary with velocity and flight altitude, and are strongly influenced by atmospheric conditions such as temperature, pressure, and air density that are often difficult to measure directly in real time. At high flight speeds, the seeker assembly is also subject to structural vibrations, aerodynamic disturbances, and electromagnetic interference. These effects may induce undesired oscillations, resulting in LOS jitter and degraded stabilization performance.

Modern missile systems employ a variety of seeker technologies, including infrared seekers, active radar seekers, and multi-mode hybrid configurations. These seekers require highly stable LOS control to ensure accurate real-time signal processing and reliable target tracking under dynamic operating conditions.

Gimbal mechanisms have been extensively investigated and widely implemented in industrial, aerospace, and defense applications, particularly in advanced weapon and sensor systems deployed by Naval and Air Defense–Air Force forces. In such systems, the gimbal assembly serves as the primary stabilization platform for seeker heads and sensor payloads, including radar antennas, electro-optical/infrared sensors, tracking cameras, and laser rangefinders.

Owing to stringent requirements on compactness, high stabilization accuracy, structural rigidity, and fast transient response, two-axis and three-axis gimbal architectures are predominantly adopted in modern precision-guided munitions and surveillance platforms.

In the subsequent section, a comprehensive dynamic model of the three-axis gimbal system is formulated based on rigid-body rotational dynamics. The influence of disturbance torques on system stability, tracking performance, and transient response characteristics is rigorously analyzed. Particular emphasis is placed on evaluating the degradation of line-of-sight stabilization accuracy under external perturbations and internal coupling effects.

2. Method

The general configuration of electro-optical seeker integrated with a three-axis gimbal system is depicted in Fig. 1.

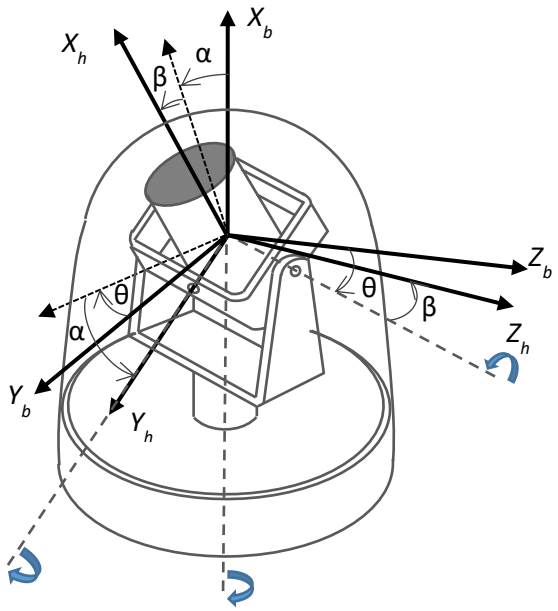


Fig. 1. Structural configuration of a electro-optical seeker mounted on a three-axis gimbal

The coordinate systems employed in the seeker control formulation are defined as follows:

Earth-Fixed Frame ($O_eX_eY_eZ_e$): This is an inertial reference frame with its origin located at the missile launch point. The O_eX_e axis lies in the horizontal plane and is oriented along a fixed reference direction. The O_eY_e axis also lies in the horizontal plane and is directed to the right of the O_eX_e axis. The O_eZ_e axis points vertically downward.

Body-Fixed Frame ($O_bX_bY_bZ_b$): This is a moving coordinate frame attached to the missile, with its origin located at the missile's center of mass. The O_bX_b axis is aligned with the missile longitudinal axis (forward direction). The O_bX_b axis lies in the missile wing plane and points to the right of O_bX_b. The O_bZ_b axis is directed downward, forming a right-handed coordinate system.

The angular relationship between the Earth-fixed frame and the body-fixed frame is illustrated in Fig. 2. using the three Euler angles $(\varepsilon, \psi, \phi)$ [1].

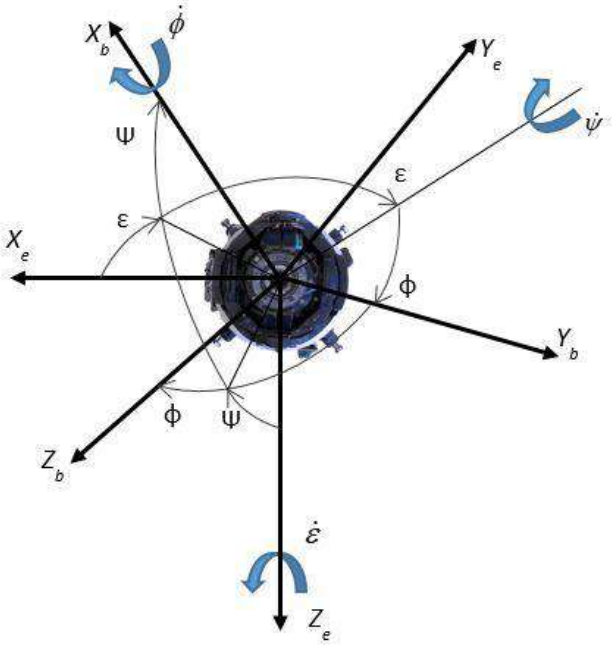


Fig. 2. The angular relationship between the Earth-fixed frame and the body-fixed frame

The direction cosine matrix (DCM) that transforms coordinates from the Earth-fixed frame to the body-fixed frame is given by $\mathbf{C}_e^b$ [2]:

$$\mathbf{C}_e^b = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} \cos\psi & 0 & -\sin\psi \\ 0 & 1 & 0 \\ \sin\psi & 0 & \cos\psi \end{bmatrix} \begin{bmatrix} \cos\varepsilon & \sin\varepsilon & 0 \\ -\sin\varepsilon & \cos\varepsilon & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

Seeker Platform Frame ($O_hX_hY_hZ_h$)**:** This coordinate frame is defined with its origin located at the rotational center of the three-axis gimbal system. The O_hX_h axis coincides with the optical axis of the seeker platform. The O_hY_h axis is aligned with the gimbal pitch rotation axis and points to the right of the O_hX_h axis. The O_hZ_h axis is directed downward, forming a right-handed coordinate system.

The relationship between the body-fixed frame and the seeker platform frame is illustrated in Fig. 1, where the origin of the body-fixed frame is translated to coincide with the origin of the seeker platform frame.

The transformation is performed through a sequence of rotations. First, the body-fixed frame is rotated about the O_bX_b axis by the roll angle θ, resulting in two intermediate axes denoted as O_bY_b' and O_bZ_b'. Next, a rotation is carried out about the O_bZ_b' axis by the yaw angle α. Finally, a rotation about the O_hY_h axis is performed by the pitch angle β.

The DCM that transforms coordinates from the body-fixed frame to the seeker platform frame is given by $\mathbf{C}_b^h$:

$$\mathbf{C}_b^h = \begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{bmatrix} \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix} \tag{2}$$

Missile–Target Line-of-Sight Frame ($O_pX_pY_pZ_p$) **– LOS Frame:** This coordinate frame is defined with its origin located at the rotational center of the three-axis gimbal system. The O_pX_p axis is directed toward the target. The O_pY_p axis lies in the horizontal plane and points to the right of the O_pX_p axis. O_pZ_p axis is directed downward, forming a right-handed coordinate system.

The relationship between the seeker platform frame and the missile–target line-of-sight frame is characterized by three tracking error angles e_x, e_y, e_z as illustrated in Fig. 3.

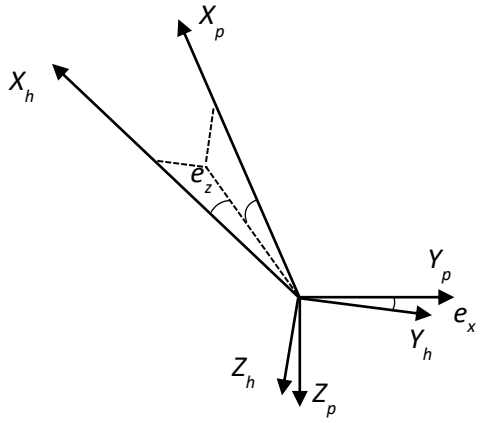


Fig. 3. Relationship between the seeker coordinate frame and the missile–target line-of-sight frame

The DCM that transforms coordinates from the seeker platform frame to the missile–target line-of-sight frame is given by $\mathbf{C}_h^p$:

$$\mathbf{C}_h^p = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos e_x & \sin e_x \\ 0 & -\sin e_x & \cos e_x \end{bmatrix} \begin{bmatrix} \cos e_y & 0 & -\sin e_y \\ 0 & 1 & 0 \\ \sin e_y & 0 & \cos e_y \end{bmatrix} \begin{bmatrix} \cos e_z & \sin e_z & 0 \\ -\sin e_z & \cos e_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

Inertial Frame ($O_IX_IY_IZ_I$): This coordinate frame is defined with its origin located at the missile's center of mass. The axes of the inertial frame are oriented identically to those of the Earth-fixed frame. The relationship between the inertial frame and the Earth-fixed frame

describes the translational position of the missile in space. The relationship between the inertial frame and the body-fixed frame characterizes the missile's attitude.

The three-axis gimbal system and the associated reference frames are described as follows: the body-fixed frame **B** is attached to the seeker assembly, while three additional reference frames, denoted as **R**, **Y**, and **P** are rigidly mounted on the Roll, Yaw, and Pitch gimbal axes, respectively.

The three-axis gimbal is modeled as a rigid body. The inertial angular velocities of its associated coordinate frames, expressed along their respective axes, are defined as:

$$\boldsymbol{\omega}_{B/I}^{B}=\begin{bmatrix}p_b & q_b & r_b\end{bmatrix}^T,\boldsymbol{\omega}_{R/I}^{R}=\begin{bmatrix}p_R & q_R & r_R\end{bmatrix}^T,\boldsymbol{\omega}_{P/I}^{P}=\begin{bmatrix}p_P & q_P & r_P\end{bmatrix}^T,\boldsymbol{\omega}_{Y/I}^{Y}=\begin{bmatrix}p_Y & q_Y & r_Y\end{bmatrix}^T \quad (4)$$

Let the inertia matrix of the Roll frame be denoted by J_R, the inertia matrix of the Yaw frame by J_Y, and the inertia matrix of the Pitch frame by J_P, respectively:

$$J_R=\begin{bmatrix}J_x^R & J_{xy}^R & J_{xz}^R\\ J_{xy}^R & J_y^R & J_{yz}^R\\ J_{xz}^R & J_{yz}^R & J_z^R\end{bmatrix},\ J_Y=\begin{bmatrix}J_x^Y & J_{xy}^Y & J_{xz}^Y\\ J_{xy}^Y & J_y^Y & J_{yz}^Y\\ J_{xz}^Y & J_{yz}^Y & J_z^Y\end{bmatrix},\ J_P=\begin{bmatrix}J_x^P & J_{xy}^P & J_{xz}^P\\ J_{xy}^P & J_y^P & J_{yz}^P\\ J_{xz}^P & J_{yz}^P & J_z^P\end{bmatrix} \quad (5)$$

Where:

$J_x^Y,\ J_y^Y,\ J_z^Y,\ J_x^P,\ J_y^P,\ J_z^P,\ J_x^R,\ J_y^R,\ J_z^R$: denote the principal moments of inertia about the x, y-, and z-axes of the Roll, Yaw, and Pitch gimbal frames, respectively.

$J_{xy}^Y,\ J_{xz}^Y,\ J_{yz}^Y,\ J_{xy}^P,\ J_{xz}^P,\ J_{yz}^P,\ J_{xy}^R,\ J_{xz}^R,\ J_{yz}^R$: represent the products of inertia, which characterize the cross-axis coupling effects arising from rotational motion about different axes.

Define the DCM transforming coordinates from the body-fixed frame to the roll gimbal frame as $\mathbf{C}_B^R$, the DCM from the roll gimbal frame to the yaw gimbal frame as $\mathbf{C}_R^Y$, and the DCM from the yaw gimbal frame to the pitch gimbal frame as $\mathbf{C}_Y^P$:

$$\mathbf{C}_B^R=\begin{bmatrix}1 & 0 & 0\\ 0 & \cos\theta & \sin\theta\\ 0 & -\sin\theta & \cos\theta\end{bmatrix},\ \mathbf{C}_R^Y=\begin{bmatrix}\cos\alpha & \sin\alpha & 0\\ -\sin\alpha & \cos\alpha & 0\\ 0 & 0 & 1\end{bmatrix},\ \mathbf{C}_Y^P=\begin{bmatrix}\cos\beta & 0 & -\sin\beta\\ 0 & 1 & 0\\ \sin\beta & 0 & \cos\beta\end{bmatrix} \quad (6)$$

According to [3], the inertial angular velocities of the roll frame, yaw frame, and pitch frame are computed as follows:

$$\omega_{R/I}^R-C_B^R\,\omega_{B/I}^B=\begin{bmatrix}\dot\theta & 0 & 0\end{bmatrix}^T\Leftrightarrow\ \omega_{R/I}^R=C_B^R\,\omega_{B/I}^B+\begin{bmatrix}\dot\theta & 0 & 0\end{bmatrix}^T=\begin{bmatrix}p_R & q_R & r_R\end{bmatrix}^T \quad (7)$$

$$\omega_{Y/I}^Y-C_R^Y\,\omega_{R/I}^R=\begin{bmatrix}0 & 0 & \dot\alpha\end{bmatrix}^T\Leftrightarrow\ \omega_{Y/I}^Y=C_R^Y\,\omega_{R/I}^R+\begin{bmatrix}0 & 0 & \dot\alpha\end{bmatrix}^T=\begin{bmatrix}p_Y & q_Y & r_Y\end{bmatrix}^T \quad (8)$$

$$\omega_{P/I}^P-C_Y^P\,\omega_{Y/I}^Y=\begin{bmatrix}0 & \dot\beta & 0\end{bmatrix}^T\Leftrightarrow\ \omega_{P/I}^P=C_Y^P\,\omega_{Y/I}^Y+\begin{bmatrix}0 & \dot\beta & 0\end{bmatrix}^T=\begin{bmatrix}p_P & q_P & r_P\end{bmatrix}^T \quad (9)$$

The rotational moment equation for the Pitch frame is given by [3]:

$$T^P=\frac{d}{dt}H^P+\omega_{P/I}^P\times H^P \quad (10)$$

Where:

$$H^P=J_P\omega_{P/I}^P=\begin{bmatrix}H_x^P & H_y^P & H_z^P\end{bmatrix}^T \quad (11)$$

From (9)–(11), the complete dynamic model of the Pitch frame can be derived as follows:

$$J_y^P \dot{q}_P = T_P + \left(J_z^P - J_x^P\right) p_P r_P + J_{xz}^P \left(p_P^2 - r_P^2\right) - J_{yz}^P \left(\dot{r}_P - p_P q_P\right) - J_{xy}^P \left(\dot{p}_P + q_P r_P\right) \quad (12)$$

where T_P denotes the total torque acting on the Pitch frame, including the motor-generated torque and the friction-induced torque.

Similarly, the dynamic model of the Yaw frame can be derived as follows:

$$J^Y \dot{r}_Y = T_Y + T_{dy1} + T_{dy2} + T_{dy3} \quad (13)$$

Where:

$$J^Y = J_z^Y + J_x^P \sin^2\beta + J_z^P \cos^2\beta - J_{xz}^P \sin\left(2\beta\right) \quad (14)$$

$$T_{dy1} = \left[J_x^Y + J_x^P \cos^2\beta + J_z^P \sin^2\beta + J_{xz}^P \sin\left(2\beta\right) - \left(J_y^P + J_y^Y\right)\right] p_Y q_Y \quad (15)$$

$$\begin{aligned} T_{dy2} = &-\left[J_{xz}^Y + \left(J_z^P - J_x^P\right)\sin\beta\cos\beta + J_{xz}^P \cos\left(2\beta\right)\right]\left(\dot{p}_Y - q_Y r_Y\right) \\ &-\left[J_{yz}^P + J_{yz}^P \cos\beta - J_{xy}^P \sin\beta\right]\left(\dot{q}_Y + p_Y r_Y\right) \\ &-\left[J_{xy}^Y + J_{xy}^P \cos\beta + J_{yz}^P \sin\beta\right]\left(p_Y^2 - q_Y^2\right) \end{aligned} \quad (16)$$

$$\begin{aligned} T_{dy3} = &\ddot{\beta}\left(J_{xy}^P \sin\beta - J_{yz}^P \cos\beta\right) \\ &+\dot{\beta}\left[\left(J_x^P - J_z^P\right)\left(p_Y \cos 2\beta - r_Y \sin 2\beta\right) + 2J_{xz}^P \left(p_Y \sin 2\beta + r_Y \cos 2\beta\right)\right] \\ &+\dot{\beta}\left[\left(J_{yz}^P \sin\beta + J_{xy}^P \cos\beta\right)\left(q_P + q_Y\right) - J_y^P p_Y\right] \end{aligned} \quad (17)$$

The dynamic model of the Roll frame is expressed as:

$$J_1^R \dot{r}_Y = T_R + T_{dr1} + T_{dr2} + T_{dr3} + T_{dr4} + T_{dr5} \quad (18)$$

$$\begin{aligned} J_1^R = &J_x^R + \cos^2\alpha(J_x^Y + J_{xz}^P \sin 2\beta + J_x^P \cos^2\beta + J_z^P \sin^2\beta) \\ &-\sin 2\alpha(J_{xy}^Y + J_{xy}^P \cos\beta + J_{yz}^P \sin\beta) + \sin\alpha(J_y^Y + J_y^P) \end{aligned} \quad (19)$$

$$\begin{aligned} T_{dr1} = &-\dot{q}_R[J_{xy}^R + \cos\alpha\sin\alpha(J_x^Y + J_{xz}^P \sin 2\beta + J_x^P \cos^2\beta + J_z^P \sin^2\beta) \\ &+\cos 2\alpha(J_{xy}^Y + J_{xy}^P \cos\beta + J_{yz}^P \sin\beta) - \sin\alpha\cos\alpha(J_y^Y + J_y^P)] \\ &-\dot{r}_R[J_{xz}^R - \sin\alpha(J_{yz}^Y - J_{xy}^P \sin\beta + J_{yz}^P \cos\beta) \\ &+\cos\alpha(J_{xz}^Y - J_x^P \sin\beta\cos\beta + J_z^P \sin\beta\cos\beta + J_{xz}^P \cos 2\beta)] \end{aligned} \quad (20)$$

$$\begin{aligned} &T_{dr2} = -p_R q_R\{J_{xz}^R + [J_{xz}^Y + (J_z^P - J_x^P)\cos\beta\sin\beta + J_{xz}^P \cos 2\beta]\cos\alpha \\ &-(J_{yz}^Y - J_{xy}^P \sin\beta + J_{yz}^P \cos\beta)\sin\alpha\} + p_R r_R[J_x^R + \cos^2\alpha(J_x^Y + J_{xz}^P \sin 2\beta + J_x^P \cos^2\beta + J_z^P \sin^2\beta) \\ &-\sin 2\alpha(J_{xy}^Y + J_{xy}^P \cos\beta + J_{yz}^P \sin\beta) + \sin^2\alpha(J_y^Y + J_y^P)] - q_R r_R[J_z^Y + J_z^R - J_{xy}^Y + J_x^P \sin^2\beta \\ &-J_{xz}^P \sin 2\beta + J_z^P \cos^2\beta - \cos\alpha\sin\alpha(J_x^Y + J_{xz}^P \sin 2\beta + J_x^P \cos^2\beta + J_z^P \sin^2\beta) \\ &-\cos 2\alpha(J_{xy}^Y + J_{xy}^P \cos\beta + J_{yz}^P \sin\beta) + \sin\alpha\cos\alpha(J_y^Y + J_y^P)] \end{aligned} \quad (21)$$

$$T_{dr3} = -q_R^2\{J_{yz}^R + [J_{xz}^Y + (J_z^P - J_x^P)\cos\beta\sin\beta + J_{xz}^P\cos 2\beta]\sin\alpha + (J_{yz}^Y - J_{xy}^P\sin\beta + J_{yz}^P\cos\beta)\cos\alpha\}$$
$$+r_R^2\{J_{xz}^R + \cos\alpha(J_{xz}^Y - J_x^P\sin\beta\cos\beta + J_z^P\sin\beta\cos\beta + J_{xz}^P\cos 2\beta) - \sin\alpha(J_{yz}^Y - J_{xy}^P\sin\beta + J_{yz}^P\cos\beta)\} \tag{22}$$

$$\begin{aligned} T_{dr4} = r_R\{&\dot{\alpha}[\cos\alpha(J_{xz}^Y - J_x^P\sin\beta\cos\beta + J_z^P\sin\beta\cos\beta + J_{xz}^P\cos 2\beta) \\ &-\sin\alpha(J_{yz}^Y - J_{xy}^P\sin\beta + J_{yz}^P\cos\beta)] + \dot{\beta}[\cos\alpha(J_{xy}^P\cos\beta + J_{yz}^P\sin\beta) - J_y^P\sin\alpha]\} \\ &-q_R[(J_{yz}^P\cos\beta - J_{xy}^P\sin\beta)\dot{\beta} + \dot{\alpha}(J_z^Y + J_x^P\sin^2\beta - J_{xz}^P\sin 2\beta + J_z^P\cos^2\beta)] \end{aligned} \tag{23}$$

$$T_{dr5} = -\ddot{\alpha}[\cos\alpha(J_{xz}^Y - J_x^P\sin\beta\cos\beta + J_z^P\sin\beta\cos\beta + J_{xz}^P\cos 2\beta) - \sin\alpha(J_{yz}^Y - J_{xy}^P\sin\beta + J_{yz}^P\cos\beta)]$$
$$+\dot{\alpha}^2[\sin\alpha(J_{xz}^Y - J_x^P\sin\beta\cos\beta + J_z^P\sin\beta\cos\beta + J_{xz}^P\cos 2\beta) + \cos\alpha(J_{yz}^Y - J_{xy}^P\sin\beta + J_{yz}^P\cos\beta)]$$
$$+\dot{\alpha}\{p_R[\sin 2\alpha(J_x^Y + J_{xz}^P\sin 2\beta + J_x^P\cos^2\beta + J_z^P\sin^2\beta) + 2\cos 2\alpha(J_{xy}^Y + J_{xy}^P\cos\beta + J_{yz}^P\sin\beta)$$
$$-\sin 2\alpha(J_y^Y + J_y^P)] - q_R[\cos 2\alpha(J_x^Y - J_y^Y - J_y^P + J_{xz}^P\sin 2\beta + J_x^P\cos^2\beta + J_z^P\sin^2\beta)$$
$$-2\sin 2\alpha(J_{xy}^Y + J_{xy}^P\cos\beta + J_{yz}^P\sin\beta)] + r_R[\sin\alpha(J_{xz}^Y - J_x^P\sin\beta\cos\beta + J_z^P\sin\beta\cos\beta + J_{xz}^P\cos 2\beta)$$
$$+\cos\alpha(J_{yz}^Y - J_{xy}^P\sin\beta + J_{yz}^P\cos\beta)] + \dot{\beta}[\sin\alpha(J_{xy}^P\cos\beta + J_{yz}^P\sin\beta) + J_y^P\cos\alpha]\} \tag{24}$$

In the following, an adaptive sliding mode controller is developed for the three-axis gimbal system. The core concept of sliding mode control *(*SMC*)* lies in designing a suitable sliding manifold such that the state tracking errors are driven onto this manifold and then asymptotically converge to the desired equilibrium point.

Consider the following second-order system:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = f(x_1, x_2, t) + bu(t) + d(t) \end{cases} \tag{25}$$

In this formulation, $f(x_1, x_2, t)$ denotes a known linear or nonlinear function describing the system dynamics, $u(t)$ is the control input, and $d(t)$ represents external disturbances affecting the system.

To alleviate the chattering drawback associated with classical SMC, numerous sliding manifold designs and reaching laws have been proposed. In this work, a fast terminal sliding surface is employed to enhance convergence performance [4]:

$$s = \dot{x} + \tau_1 x + \tau_2 x^{\delta} \tag{26}$$

An adaptive power reaching law is adopted [5]:

$$\dot{s} = -k_1|s|^{r_1}\operatorname{sgn}(s) - k_2|s|^{r_2\left(1+\frac{\operatorname{sgn}(|s|-1)}{2}\right)}\operatorname{sgn}(s) \tag{27}$$

With $k_1 > 0, k_2 > 0, 0 < r_1 < 1, r_2 > 1$

To enable the application of the sliding mode control law to the drive system of the three-axis gimbal, the mathematical models developed above must be expressed in the state-space form as given in (25). Based on (12), (13), and (18), and by defining $x_1 = \lambda_x^h$,

$x_2 = \dot{\lambda}_x^h = p_P,\ x_3 = \lambda_y^h,\ x_4 = \dot{\lambda}_y^h = q_P,\ x_5 = \lambda_z^h,\ x_6 = \dot{\lambda}_z^h = r_P$ where $\lambda_x^h,\ \lambda_y^h,\ \lambda_z^h$ which denote the seeker head angles along the x-, y-, and z-axes, respectively, measured in the inertial coordinate frame, $\dot{\lambda}_x^h,\ \dot{\lambda}_y^h,\ \dot{\lambda}_z^h$ which denote the angular velocities of the seeker head about the x-, y-, and z-axes, respectively, measured in the inertial coordinate frame.

Accordingly, the state-space model of the seeker drive system based on the Roll–Yaw–Pitch gimbal configuration is given as follows:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = f_1(x_2, x_4, x_6) + \dfrac{(T_m^R - T_f^R)\cos\alpha\cos\beta}{J_1^R} + T_D^R \\ \dot{x}_3 = x_4 \\ \dot{x}_4 = f_2(x_2, x_4, x_6) + \dfrac{T_m^P - T_f^P}{J_y^P} + T_D^P \\ \dot{x}_5 = x_6 \\ \dot{x}_6 = f_3(x_2, x_4, x_6) + \dfrac{(T_m^Y - T_f^Y)\cos\beta}{J_1^Y} + T_D^Y \end{cases} \tag{28}$$

The angular displacement error and angular rate error between the reference signals and the actual states are defined as:

$$\begin{cases} e_1 = \lambda_x - x_1 \\ \dot{e}_1 = \dot{\lambda}_x - x_2 \end{cases};\ \begin{cases} e_2 = \lambda_y - x_3 \\ \dot{e}_2 = \dot{\lambda}_y - x_4 \end{cases};\ \begin{cases} e_3 = \lambda_z - x_5 \\ \dot{e}_3 = \dot{\lambda}_z - x_6 \end{cases} \tag{29}$$

Here, $\lambda_x, \dot{\lambda}_x$ denote the inertial angular position and inertial angular rate of the LOS about the x-axis, respectively; $\lambda_y, \dot{\lambda}_y$ denote the inertial angular position and inertial angular rate of the LOS about the y-axis, respectively; and $\lambda_z, \dot{\lambda}_z$ denote the inertial angular position and inertial angular rate of the ***LOS*** about the z-axis, respectively.

By applying the fast terminal sliding mode control surface, we obtain:

$$\begin{cases} s_1 = \dot{e}_1 + \tau_{11} e_1 + \tau_{12} e_1^{\delta_1} \\ s_2 = \dot{e}_2 + \tau_{21} e_2 + \tau_{22} e_2^{\delta_2} \\ s_3 = \dot{e}_3 + \tau_{31} e_3 + \tau_{32} e_3^{\delta_3} \end{cases} \tag{30}$$

Where $\tau_{11}, \tau_{12} > 0;\ \tau_{21}, \tau_{22} > 0;\ \tau_{31}, \tau_{32} > 0;\ 0 < \delta_1, \delta_2, \delta_3 < 1$, and $\delta_1, \delta_2, \delta_3$ are ratios of positive odd integers.

By applying the adaptive power reaching law in (12), we obtain:

$$\begin{cases} \dot{s}_1 = -k_{11}|s_1|^{r_{11}}\operatorname{sgn}(s_1) - k_{12}|s_1|^{r_{12}\left(1+\frac{\operatorname{sgn}(|s_1|-1)}{2}\right)}\operatorname{sgn}(s_1) \\ \dot{s}_2 = -k_{21}|s_2|^{r_{21}}\operatorname{sgn}(s_2) - k_{22}|s_2|^{r_{22}\left(1+\frac{\operatorname{sgn}(|s_2|-1)}{2}\right)}\operatorname{sgn}(s_2) \\ \dot{s}_3 = -k_{31}|s_3|^{r_{31}}\operatorname{sgn}(s_3) - k_{32}|s_3|^{r_{32}\left(1+\frac{\operatorname{sgn}(|s_3|-1)}{2}\right)}\operatorname{sgn}(s_3) \end{cases} \tag{31}$$

where $k_{11},k_{12}>0$; $k_{21},k_{22}>0$; $k_{31},k_{32}>0$; $0<r_{11},r_{21},r_{31}<1$; $r_{12},r_{22},r_{32}>1$

Substituting (29) into (30), differentiating, and then utilizing the system model in (28), the sliding mode control law for the seeker drive system is obtained as follows:

$$\begin{cases} T_m^R = \dfrac{J_1^R}{\cos\alpha\cos\beta}\left[\ddot{\lambda}_x - \dot{s}_1 - f_1(x_2,x_4,x_6) - T_D^R + \tau_{11}\dot{e}_1 + \tau_{12}\dfrac{d}{dt}e_1^{\delta_1}\right] + \hat{T}_f^R \\ T_m^P = J_y^P\left[\ddot{\lambda}_y - \dot{s}_2 - f_2(x_2,x_4,x_6) - T_D^P + \tau_{21}\dot{e}_2 + \tau_{22}\dfrac{d}{dt}e_2^{\delta_2}\right] + \hat{T}_f^P \\ T_m^Y = \dfrac{J_y^Y}{\cos\beta}\left[\ddot{\lambda}_z - \dot{s}_3 - f_3(x_2,x_4,x_6) - T_D^Y + \tau_{31}\dot{e}_3 + \tau_{32}\dfrac{d}{dt}e_3^{\delta_3}\right] + \hat{T}_f^Y \end{cases} \quad (32)$$

3. Results and Discussion

The simulation scenarios are conducted as follows:

Table 1. Simulation Scenarios

Case	Angular velocity $\boldsymbol{\omega}_{B/I}^B$ (deg/s)		
	P_b	Q_b	R_b
1	0	0	0
2	0	1.5	5
3	0	3	1.5
4	2	5	5

The simulation results were obtained under the proposed control scheme.

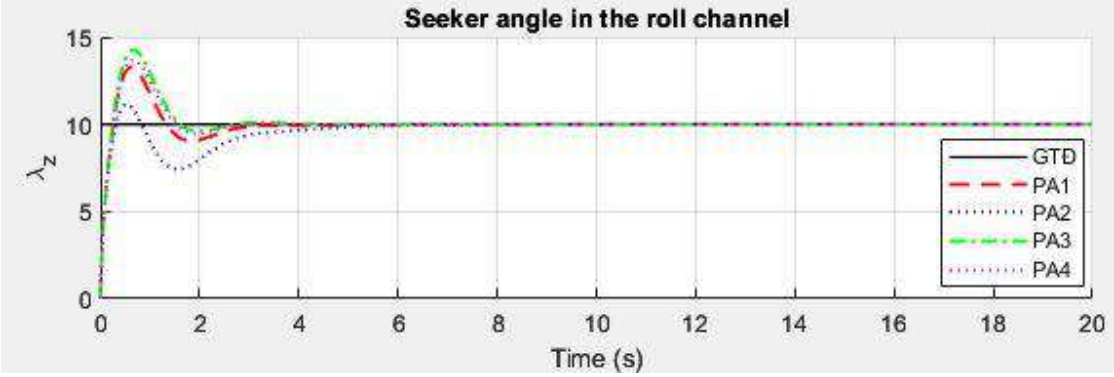


Fig. 4. Simulation results of the drive system in the Roll channel

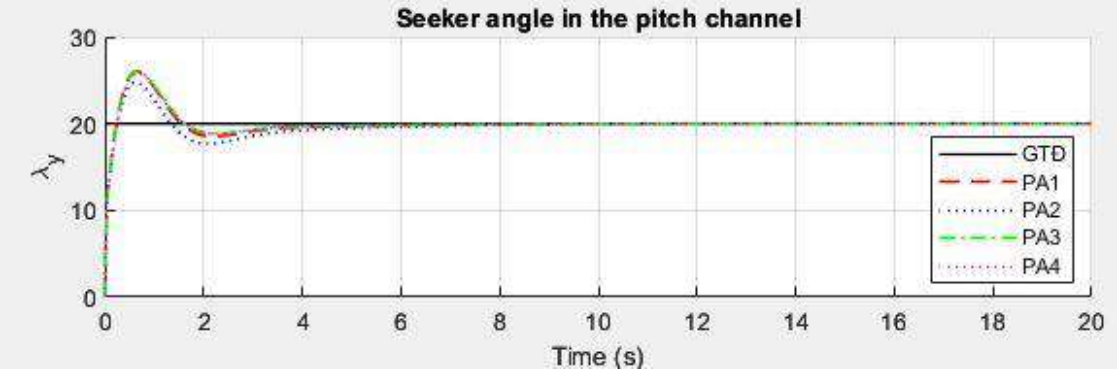


Fig. 5. Simulation results of the drive system in the Pitch channel

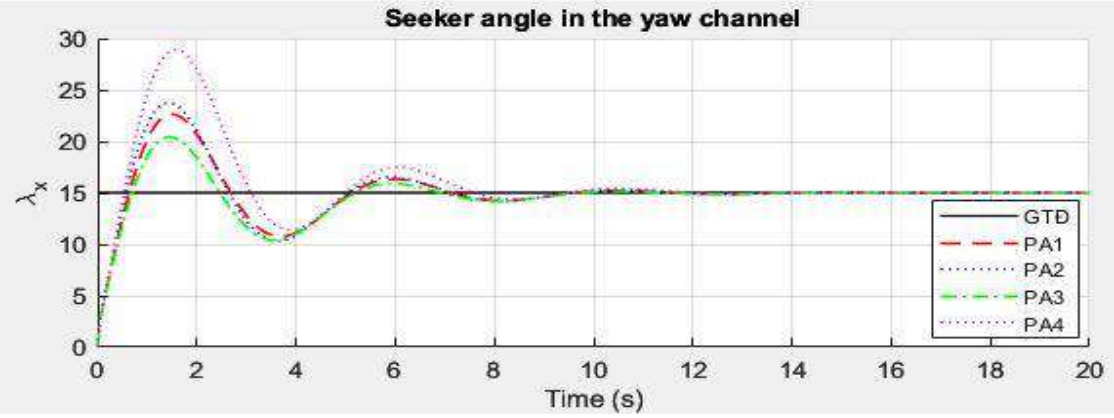


Fig. 6. Simulation results of the drive system in the Yaw channel

The simulation results demonstrate that the proposed controller ensures satisfactory tracking performance under all considered scenarios. The angular position and angular rate errors converge rapidly to their desired values, while chattering and oscillations are significantly reduced. The system maintains stability and robustness in the presence of disturbances and model uncertainties.

4. Conclusion

This paper has presented the development of a dynamic model for the stabilization and LOS control system of a missile seeker equipped with a three-axis gimbal. The model incorporates the main physical factors affecting system behavior, including applied torques, inertial properties, and nonlinear coupling among the roll, yaw, and pitch channels.

Based on the developed model, adaptive sliding mode control laws have been implemented to enhance stabilization and tracking performance. Simulation results in all three channels demonstrate that the proposed control approach achieves accurate tracking, fast convergence, and satisfactory transient performance. In particular, the seeker angles converge to their desired values with limited overshoot and small steady-state error, confirming the effectiveness of the control strategy.

Furthermore, the adaptive sliding mode controller exhibits strong robustness against system nonlinearities, channel coupling, and modeling uncertainties. The obtained results validate both the accuracy of the developed dynamic model and the applicability of the proposed control method.

Overall, the combination of the proposed dynamic model and adaptive sliding mode control provides a reliable framework for stabilization and LOS control of missile seeker systems. Future work will focus on controller optimization and experimental validation to further improve system performance.

References

1. Siouris, G.M., *Missile guidance and control systems*. 2004: Springer Science & Business Media.

2. Palumbo, N.F., R.A. Blauwkamp, and J.M. Lloyd, *Basic principles of homing guidance.* Johns Hopkins APL Technical Digest, 2010. **29**(1): p. 25-41.

3. Ekstrand, B., *Equations of motion for a two-axes gimbal system.* IEEE Transactions on Aerospace and Electronic Systems, 2001. **37**(3): p. 1083-1091.

4. Pan, H., et al., *A novel global fast terminal sliding mode control scheme for second-order systems.* IEEE Access, 2020. **8**: p. 22758-22769.

5. NguyenNgocHung, 2023, "Nghiên cứu các giải pháp nâng cao chất lượng đầu tự dẫn ảnh nhiệt của tên lửa tầm gần".

Phát hiện thủy lôi dưới nước bằng phương pháp TDOA sử dụng tín hiệu sonar từ tàu mẹ và robot ngầm di chuyển đa vị trí

Phạm Ngọc Hòa[1] *, Hoàng Mạnh Tưởng[1], Nguyễn Đình Khánh[2]

[1]*Học viện Kỹ thuật quân sự;* [2]*Học viện Hải quân*

Email: phamngochoa1495@gmail.com; Tel: 0976621043

Tóm tắt

Bài báo đề xuất một giải pháp định vị thủy lôi dưới nước, trong đó nguồn sonar được bố trí trên tàu mẹ và bộ thu tín hiệu đặt trên robot ngầm. Quá trình lan truyền xung chirp sonar trong nước biển được mô phỏng trên COMSOL Multiphysics nhằm xác định các thời điểm xung phản xạ tại robot ngầm khi di chuyển tới bốn vị trí xác định, sử dụng matched filter để trích xuất thời điểm chính xác của tín hiệu hồi âm. Dữ liệu thời gian phản xạ này sau đó được sử dụng làm đầu vào cho thuật toán TDOA (Time Difference of Arrival) triển khai trên MATLAB, từ đó ước lượng tọa độ thủy lôi. Kết quả mô phỏng cho thấy phương pháp này có thể xác định vị trí thủy lôi một cách tin cậy trong môi trường giả lập. Giải pháp kết hợp mô phỏng COMSOL và MATLAB cung cấp cơ sở kỹ thuật cho các nghiên cứu và triển khai thực nghiệm trong định vị tự động của robot ngầm.

Từ khóa: *Âm học dưới nước; TDOA; định vị thủy lôi; thuật toán Gauss–Newton; robot ngầm.*

1. Đặt vấn đề

Phát hiện và định vị mục tiêu dưới nước, chẳng hạn như thủy lôi hoặc các vật thể nguy hiểm, là công việc trọng yếu trong an ninh hàng hải và tác chiến hải quân. Trong nhiều hệ thống sonar hiện nay, robot ngầm vừa đảm nhiệm vai trò phát xung vừa thu hồi lại tín hiệu phản xạ, theo mô hình đồng vị trí (monostatic). Tuy nhiên, hạn chế về công suất phát, năng lượng và nhiệt độ của robot ngầm làm thu hẹp tầm quan sát, đặc biệt khi sử dụng tần số cao để đạt độ phân giải tốt.

Một hướng tiếp cận đầy tiềm năng là kiến trúc sonar phân tán đa tĩnh (multistatic), trong đó nguồn phát và thu được tách rời không gian. Các nghiên cứu phân tán SIMO (single-input, multiple-output) đã chỉ ra rằng việc phân bố nhiều đầu thu tại các vị trí khác nhau giúp cải thiện khả năng định vị mục tiêu mà không cần đồng bộ hóa thời gian tuyệt đối [1]. Đồng thời, các kết quả từ các hệ thống bistatic/active sonar cho robot ngầm cũng cho thấy việc tách nguồn phát và thu có thể tăng khả năng theo dõi mục tiêu trong môi trường nhiễu [2].

Tuy vậy, nhiều giải pháp phân tán đa tĩnh gặp khó khăn trong thực tế do yêu cầu đồng bộ thời gian khắt khe giữa các nút thu–phát, trong khi môi trường biển thường có trễ lan truyền biến thiên, sai lệch vận tốc âm và đồng hồ thiết bị không khớp hoàn hảo [1], [3]. Ngoài ra, nếu mỗi robot ngầm vừa phát vừa thu, việc cung cấp năng lượng cho phát xung mạnh sẽ làm tăng đáng kể mức tiêu thụ năng lượng và giảm thời gian hoạt động.

Để vượt qua những điểm nghẽn này, nghiên cứu này đề xuất một kiến trúc sonar phân tán hai tầng: tàu mẹ đóng vai trò phát xung sonar công suất lớn ở tần số thấp, trong khi robot ngầm chỉ đảm nhiệm thu thụ động. Cụ thể, tàu mẹ sẽ phát các xung cách nhau mỗi 10 giây, cho phép robot ngầm di chuyển đến bốn vị trí khác nhau trước mỗi lần thu tiếp theo. Vị trí của robot tại mỗi điểm thu được xác định bằng hệ thống dẫn đường quán tính (Inertial Navigation Systems - INS), cho phép xác định tọa độ thu một cách đáng tin cậy ngay cả khi không có tín hiệu GPS

* Email: phamngochoa1495@gmail.com

(Global Positioning System).

Kiến trúc này tạo nên một mạng thu ảo gồm bốn đầu thu phân tán theo thời gian – không gian bằng cách chỉ dùng một robot. Nhờ việc thu bốn tín hiệu phản xạ tại bốn vị trí khác nhau, hệ thống có thể trích xuất bốn giá trị TDOA độc lập. Vì TDOA chỉ phụ thuộc vào sự khác biệt thời gian đến chứ không cần đồng bộ tuyệt đối thời điểm phát, nó rất phù hợp cho cấu hình phân tán hai tầng này và khắc phục được giới hạn của việc đồng bộ hóa khắt khe [3].

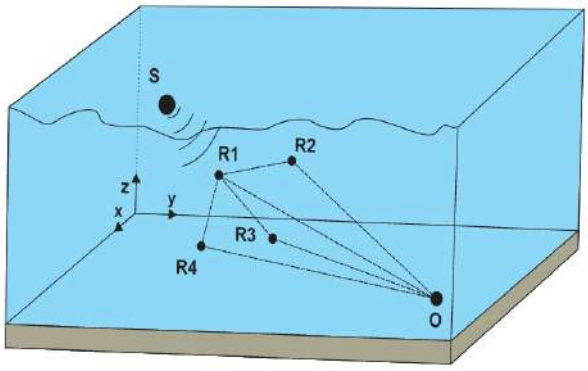
Cuối cùng, việc sử dụng INS để định vị robot thu tại mỗi thời điểm thu không chỉ đảm bảo độ chính xác hình học cần thiết để giải bài toán TDOA, mà còn giảm thiểu yêu cầu phần cứng và năng lượng so với việc trang bị nhiều robot ngầm thu cố định. Do đó, kiến trúc sonar phân tán hai tầng như trên mở ra khả năng triển khai hệ thống sonar hiệu quả, năng lượng thấp và tầm xa phù hợp với nhiều ứng dụng nhiệm vụ dưới nước hiện đại.

2. Xác định tọa độ thủy lôi

Trong các hoạt động bảo đảm an toàn hàng hải và hỗ trợ công tác rà phá vật nổ dưới nước, việc phát hiện và định vị chính xác thủy lôi đóng vai trò đặc biệt quan trọng. Do môi trường dưới nước có đặc tính lan truyền sóng phức tạp, việc xác định vị trí của các đối tượng nguy hiểm đòi hỏi các phương pháp đo đạc âm học có độ tin cậy cao [8], [9]. Trong số đó, kỹ thuật độ trễ thời gian đến được xem là một giải pháp hứa hẹn nhờ khả năng khai thác sự chênh lệch thời gian lan truyền của tín hiệu âm giữa các điểm thu để suy luận vị trí nguồn phản xạ [3], [5].

Trong mô hình nghiên cứu, các xung âm thăm dò được phát ra từ tàu mẹ và lan truyền trong môi trường nước, tương tác với các vật thể khả nghi bao gồm thủy lôi hoặc các cấu trúc có đặc tính phản xạ tương tự. Hệ thống thu nhận bố trí trên robot ngầm hoặc các điểm đo cố định tiếp nhận tín hiệu phản hồi và ghi nhận thời điểm đến của từng xung. Bằng cách phân tích sự sai lệch thời gian đến giữa các tín hiệu tương ứng với các vị trí thu khác nhau, thuật toán TDOA xây dựng các đường đẳng thời và tiến hành nội suy để xác định vị trí của nguồn phản xạ trong không gian ba chiều [3], [5].

Phương pháp TDOA có ưu thế do không yêu cầu đồng bộ thời gian tuyệt đối giữa tất cả các thiết bị, đồng thời có thể hoạt động hiệu quả trong môi trường nhiễu nền cao đặc điểm phổ biến trong điều kiện biển sâu [7], [8]. Việc áp dụng TDOA giúp cải thiện độ chính xác trong việc nhận dạng và định vị thủy lôi, hỗ trợ đáng kể cho các hệ thống khảo sát tự động, tăng cường an toàn vận hành và góp phần nâng cao hiệu quả của các nhiệm vụ rà soát dưới nước.



Hình 1. Mô hình xác định tọa độ thủy lôi

Trên Hình 1 minh họa mô hình xác định tọa độ thủy lôi trong không gian 3D dưới nước. Trong đó, tàu mẹ tại vị trí S đóng vai trò nguồn phát sonar chủ động, tạo xung thăm dò lan truyền trong môi trường biển. Robot ngầm sau đó di chuyển tuần tự tới nhiều vị trí thu khác nhau R_1, R_2, R_3, R_4 để ghi nhận tín hiệu phản xạ từ thủy lôi O. Việc bố trí các vị trí thu phân tán theo không–thời gian như vậy tạo ra một mạng cảm biến ảo, cho phép thu thập các thời điểm đến khác nhau của tín hiệu và từ đó xây dựng bài toán định vị dựa trên TDOA.

Trong kiến trúc sonar hai tầng, giả sử tàu mẹ phát xung từ vị trí $S(x_s, y_s, z_s)$, sóng âm lan tới vật thể phản xạ tại $O(x_o, y_o, z_o)$ và phản hồi được thu tại các vị trí robot $R_i(x_i, y_i, z_i)$, i = 1, 2, 3, 4. Áp dụng nghiệm Green của phương trình sóng, áp suất tại đầu thu R_i được biểu diễn như sau [8]:

$$p_i(t) = \frac{1}{4\pi \|R_i - S\|} q\left(t - \tau_i\right) \tag{1}$$

trong đó $\|R_i - S\|$ là khoảng cách từ nguồn đến đầu thu, còn τ_i là thời gian trễ lan truyền:

$$\tau_i = \frac{\|R_i - S\|}{c} \tag{2}$$

với c là vận tốc âm trong nước. Như vậy, tín hiệu thu được chỉ là dạng sóng phát q(t) bị trễ đi một khoảng τ_i và suy giảm biên độ theo $\frac{1}{\|R_i - S\|}$.

Hàm q(t) mô tả xung kích thích phát từ tàu mẹ, đóng vai trò nguồn phát trong toàn bộ mô hình. Đối với hệ sonar phản xạ tầm xa, q(t) thường được chọn là xung chirp tuyến tính tăng tần (linear FM up-chirp) [8], [9]:

$$q(t) = A\cos\left(2\pi\left(f_0 t + \frac{B}{2T}t^2\right)\right), 0 < t < T \tag{3}$$

với A là biên độ, f_o tần số bắt đầu, B độ rộng dải tần và T thời gian xung.

Dạng chirp FM có các ưu điểm quan trọng: dải tần rộng giúp tăng độ phân giải thời gian khi ước lượng TDOA [8]; sau matched filter tạo đỉnh tương quan sắc, giúp xác định chính xác thời điểm phản xạ [8]; và ổn định hơn khi truyền xa trong môi trường hấp thụ mạnh như nước biển. Nhờ đó, q(t) đặc biệt phù hợp cho bài toán dò tìm thủy lôi ở cự ly lớn, nơi vừa yêu cầu độ chính xác cao, vừa có SNR thấp.

Nếu xét đường đi qua vật thể phản xạ O:

$$d_{SO} = \|S - O\|, \qquad d_{OR_i} = \|O - R_i\| \tag{4}$$

thời gian lan truyền tổng quát đến đầu thu R_i là:

$$t_{\text{ref},i} = \frac{d_{SO} + d_{OR_i}}{c} = \frac{\|S - O\| + \|O - R_i\|}{c} \tag{5}$$

Phương trình này thể hiện tổng chiều dài đường đi âm trong không gian ba chiều, tích hợp hình học môi trường.

Trong môi trường thực tế dưới nước, đồng bộ tuyệt đối giữa thời điểm phát và thu gần như bất khả thi [3], [5]. Do đó, thay vì sử dụng trực tiếp thời gian tuyệt đối t_i, ta chỉ sử dụng

hiệu thời gian đến giữa tín hiệu phản xạ và tín hiệu trực tiếp tại cùng một đầu thu, tức TDOA dạng bistatic (Bistatic Time Difference of Arrival):

$$\Delta t_i = t_{\text{ref},i} - t_{\text{dir},i} \tag{6}$$

Ở đây $t_{\text{dir},i}$ là thời gian lan truyền trực tiếp từ nguồn S đến đầu thu R_i:

$$t_{\text{dir},i} = \frac{\|S - R_i\|}{c} \tag{7}$$

còn $t_{\text{ref},i}$ là thời gian lan truyền của tín hiệu phản xạ theo (5).

Thay biểu thức (5) và (7) vào (6), ta được:

$$\Delta t_i = \frac{\|S-O\| + \|O-R_i\|}{c} - \frac{\|S-R_i\|}{c} = \frac{\|S-O\| + \|O-R_i\| - \|S-R_i\|}{c} \tag{8}$$

Nhân cả hai vế với c, ta đưa về miền chiều dài đường đi:

$$d_i = c\Delta t_i = \|S-O\| + \|O-R_i\| - \|S-R_i\| \tag{9}$$

Mỗi phương trình (9) tương ứng với một mặt hyperboloid ba chiều gồm tất cả các điểm O sao cho hiệu tổng chiều dài đường đi phản xạ và đường đi trực tiếp là một hằng số. Khi thu được đủ các giá trị d_i, giao điểm của các hyperboloid này xác định vị trí thủy lôi trong không gian. Như vậy, bài toán định vị được quy về việc tìm điểm 0 sao cho các chênh lệch đường đi (9) khớp với các giá trị đo được từ TDOA dạng bistatic.

Ưu điểm của TDOA là không cần biết thời điểm phát tuyệt đối, mà chỉ yêu cầu đo chính xác sự khác biệt tương đối giữa các tín hiệu thu. Điều này giúp phương pháp hoạt động ổn định trong môi trường biển nơi đồng bộ hóa thời gian rất khó thực hiện.

Từ công thức (9), ta thiết lập hệ phương trình TDOA để chuyển dữ liệu thời gian thu được tại robot sang bài toán tìm nghiệm hình học ba chiều. Với nguồn S cố định và các đầu thu R_i, $i = 1,2,3,4$, ta có:

$$f_i(O) = \|S-O\| + \|O-R_i\| - \|S-R_i\| - d_i = 0, \qquad i = 1,2,3,4. \tag{10}$$

Ba phương trình này tương ứng với ba hyperboloid trong không gian 3D [3], [5], và giao điểm của chúng chính là vị trí thủy lôi O.

Do (10) chứa căn bậc hai theo tọa độ, hệ phương trình là phi tuyến mạnh. Để giải nhanh và ổn định, ta tuyến tính hóa quanh vị trí ước lượng ban đầu:

$$O^{(k)} = \left(x^{(k)}, y^{(k)}, z^{(k)}\right)^T$$

bằng khai triển Taylor bậc nhất [4]:

$$f_i\left(O^{(k)} + \delta O\right) \approx f_i\left(O^{(k)}\right) + \nabla f_i\left(O^{(k)}\right) \cdot \delta O = 0 \tag{11}$$

trong đó $\delta O = (\delta x, \delta y, \delta z)^T$ là véc tơ hiệu chỉnh vị trí. Từ (9) ta suy ra gradient của phương trình thứ i:

$$\nabla f_i(O) = \frac{O-S}{\|O-S\|} + \frac{O-R_i}{\|O-R_i\|} \tag{12}$$

Các hàng $\left(\nabla f_i\left(O^{(k)}\right)\right)^T$ được ghép lại thành ma trận Jacobian J, biểu diễn độ nhạy của các phương trình TDOA đối với biến vị trí O. Sau khi đã có hệ phương trình TDOA và ma trận Jacobian, thuật toán Gauss–Newton dạng Levenberg–Marquardt được áp dụng để tìm nghiệm của hệ phi tuyến (11). Về mặt hình học, quá trình giải hệ phi tuyến chính là quá trình "kéo" điểm ước lượng $O^{(k)}$ dần tới giao điểm của các hyperboloid, tức vị trí thực của thủy lôi.

Từ (10), bài toán định vị có thể biểu diễn dưới dạng tối ưu hóa bình phương tối thiểu [4], [5], [7]:

$$\left(J^T J + \lambda I\right)\delta O = -J^T f\left(O^{(k)}\right) \tag{13}$$

trong đó $O^{(k)}$ là vị trí ước lượng tại vòng lặp thứ k, $\delta O = \left(\delta x, \delta y, \delta z\right)^T$ là điều chỉnh vị trí, J là ma trận Jacobian, $f\left(O^{(k)}\right)$ là vector sai số TDOA tại vị trí ước lượng hiện thời, và $\lambda > 0$ là hệ số damping theo phương pháp Levenberg–Marquardt nhằm tăng độ ổn định trong quá trình hội tụ [4], [7].

Bước cập nhật Gauss–Newton được tính theo [4]:

$$\delta O = -\left(J^T J + \lambda I\right)^{-1} J^T f\left(O^{(k)}\right) \tag{14}$$

và nghiệm được cập nhật bằng:

$$O^{(k+1)} = O^{(k)} + \delta O = O^{(k)} - \left(J^T J + \lambda I\right)^{-1} J^T f\left(O^{(k)}\right) \tag{15}$$

Trong cài đặt thực tế, nghiệm $O^{(k+1)}$ được giới hạn trong một "hộp" biên để tránh các nghiệm không vật lý, nhưng lõi toán học vẫn tuân theo các bước từ (13) đến (15). Quá trình lặp dừng khi chuẩn hội tụ thỏa mãn:

$$\left\| f\left(O^{(k)}\right) \right\| < 10^{-4}$$

Để đánh giá mức độ tin cậy của nghiệm cuối cùng, mô hình sai số được xây dựng trên phương sai nhiễu thời gian đo σ_t^2.

Ma trận hiệp phương sai của ước lượng vị trí được xác định theo [5], [9]:

$$P = \sigma_t^2 \left(J^T J\right)^{-1} \tag{16}$$

và độ lệch chuẩn tổng hợp của sai số định vị là:

$$\sigma_p = \sqrt{\text{trace}(P)} \tag{17}$$

Phương pháp này cho phép lượng hóa mức độ nhạy của kết quả định vị đối với nhiễu thời gian đo, đồng thời đưa ra tiêu chí đánh giá cấu hình các đầu thu. Sự phân tích này cung cấp cơ sở định lượng để tối ưu hóa vị trí robot thu trong mạng sonar phân tán [1].

3. Mô phỏng quá trình xác định vị trí thủy lôi

Trong nghiên cứu này, quá trình mô phỏng được thực hiện theo hai giai đoạn chính. Ở giai đoạn thứ nhất, phần mềm COMSOL Multiphysics được sử dụng để mô phỏng lan truyền sóng âm trong môi trường nước nhằm xác định các thời điểm robot ngầm thu nhận được xung sonar phản xạ khi đứng tại một vị trí xác định.

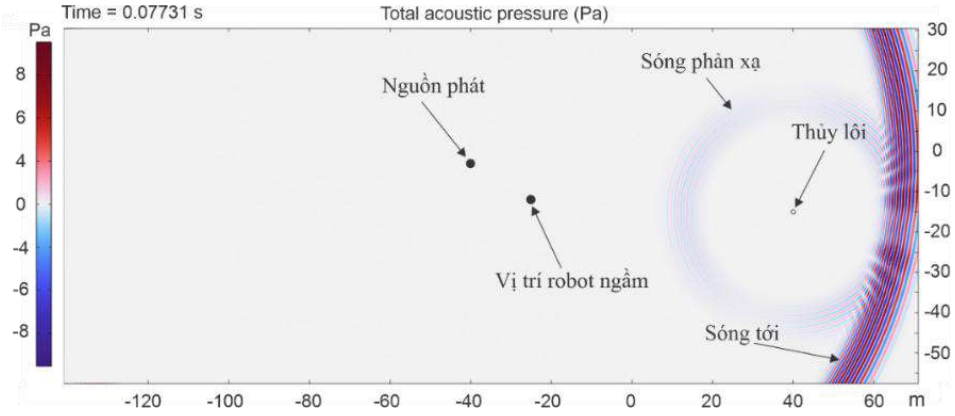
Môi trường nước được giả định là đồng nhất, với tốc độ âm thanh lan truyền đẳng hướng và không đổi (c=1500 m/s), nhằm đơn giản hóa mô hình và giảm thiểu các yếu tố phức tạp. Miền tính toán trong COMSOL Multiphysics được chia thành ba vùng: miền nước, miền chứa vật thể phản xạ và lớp hấp thụ PML (Perfectly Matched Layer) bao quanh.

Thủy lôi được mô phỏng dưới dạng cầu rỗng bằng thép, bán kính R = 0,5 m, bề dày vỏ δ =0,02 m, đóng vai trò là vật tán xạ cứng đặt vị trí xác định trong miền nước, trong khi PML đảm nhiệm vai trò triệt tiêu sóng phản xạ từ biên, giúp trường sóng trong vùng quan tâm không bị ảnh hưởng bởi ranh giới tính toán.

Lưới phần tử được tạo bằng Free Triangular mesh, với kích thước phần tử được chọn thỏa điều kiện phân giải theo bước sóng nhỏ nhất; lưới được tinh chỉnh mịn hơn ở lân cận thủy lôi và đầu thu để tăng độ chính xác, đồng thời được làm thưa dần về phía biên để giảm chi phí tính toán mà vẫn bảo đảm độ chính xác [6].

Xung sonar thăm dò do tàu mẹ phát là xung dạng chirp với tần số trung tâm f_o=1000 Hz, băng thông B=1000 Hz được phát cách nhau 10 s. Các tham số mô phỏng trong COMSOL Multiphysics, bao gồm mô hình môi trường, tốc độ âm, đặc tính tín hiệu phát và cấu hình nguồn–thu được lựa chọn sao cho phù hợp với điều kiện vận hành thực tế của bài toán.

Hình 2 minh họa kết quả mô phỏng trường áp suất âm theo thời gian, thể hiện rõ sóng tới, sóng phản xạ, cũng như vị trí của thủy lôi trong không gian mô phỏng COMSOL Multiphysics.



Hình 2. Mô phỏng lan truyền sóng âm và trường áp suất trong COMSOL Multiphysics.

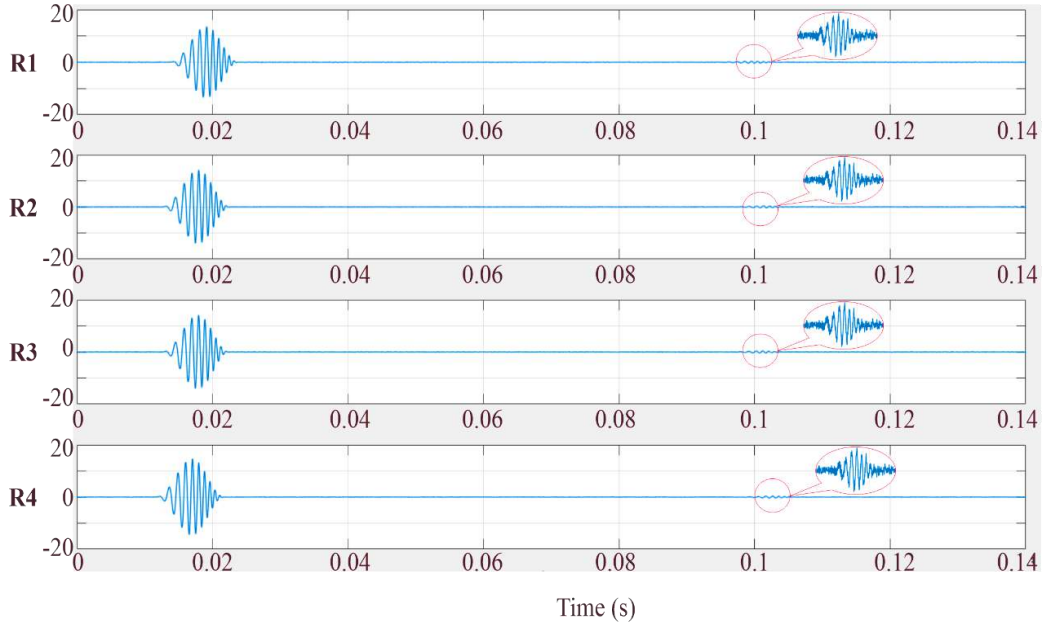
Sau khi thiết lập mô phỏng trường sóng, tọa độ của tàu mẹ, bốn vị trí robot ngầm và vị trí thủy lôi được đưa ra trong Bảng 1 dưới đây. Đây là những tọa độ đầu vào cho COMSOL Multiphysics cũng như MATLAB trong giai đoạn mô phỏng tiếp theo.

Bảng 1. Tọa độ tàu mẹ, bốn vị trí robot ngầm và thủy lôi

Điểm	x (m)	y (m)	z (m)	Ghi chú
S	-40	0	-3	Tàu mẹ – nguồn phát sonar
R_1	-20	0	-8	Bốn vị trí robot ngầm
R_2	-22.5	5	-8	
R_3	-22.5	-5	-8	
R_4	-25	0	-12	
O	40	0	15	Thủy lôi

Đồ thị biểu diễn các tín hiệu áp suất thu được tại đầu thu sonar của robot ngầm ở bốn vị trí khác nhau được thể hiện trong Hình 3.

Các đồ thị này cho thấy tín hiệu áp suất thô thu được từ mô phỏng COMSOL Multiphysics tại bốn vị trí của robot ngầm. Đây là các tín hiệu chưa qua xử lý, bao gồm thành phần sóng trực tiếp, sóng phản xạ từ thủy lôi và nhiễu môi trường. Các đặc điểm như biên độ, thời điểm xuất hiện đỉnh sóng và mức nhiễu thay đổi theo từng vị trí thu, phản ánh đúng bản chất lan truyền âm trong môi trường nước biển.



Hình 3. Tín hiệu áp suất ở đầu thu sonar

Các tín hiệu thô này sau khi đã được xử lý bằng bộ lọc matched filter. Nhờ phép lọc tương quan này, dạng xung phản xạ trở nên sắc nét, cho phép xác định chính xác thời điểm đến của các thành phần sóng trực tiếp và sóng phản xạ tại mỗi đầu thu. Từ đó, ta xác định được hiệu thời gian truyền âm bistatic tại từng đầu thu Δt_i.

Kết quả mô phỏng trên phần mềm COMSOL thu được các giá trị Δt_i trong Bảng 2.

Bảng 2. Các giá trị hiệu thời gian đến Δt_i thu được từ mô phỏng COMSOL

Đầu thu	Hiệu thời gian đến	Giá trị (giây)
1	Δt_1	0,0806
2	Δt_2	0,0851
3	Δt_3	0,0851
4	Δt_4	0,0867

Ở giai đoạn thứ hai, các hiệu thời gian đến Δt_i thu được từ mô phỏng COMSOL được sử dụng làm dữ liệu đầu vào để tiến hành mô phỏng định vị thủy lôi bằng phương pháp TDOA triển khai trên MATLAB.

Để giải hệ phương trình TDOA phi tuyến ba chiều đã xây dựng ở mục trước, thuật toán Gauss–Newton dạng Levenberg–Marquardt được áp dụng vì đặc tính hội tụ nhanh và ổn định khi xử lý các bài toán tối ưu phi tuyến. Thuật toán được cấu hình với hệ số damping $\lambda = 0{,}02$,

ngưỡng hội tụ $\left\| f\left(O^{(k)}\right)\right\| < 10^{-4}$ và số vòng lặp tối đa 100. Ma trận Jacobian được xây dựng theo (12):

$$\nabla f_i(O) = \frac{O - S}{\|O - S\|} + \frac{O - R_i}{\|O - R_i\|}$$

và nghiệm được cập nhật theo quy tắc tối ưu hóa theo (15):

$$O^{(k+1)} = O^{(k)} + \delta O = O^{(k)} - \left(J^T J + \lambda I\right)^{-1} J^T f\left(O^{(k)}\right)$$

Cấu hình thuật toán nêu trên cho phép quá trình tối ưu hội tụ nhanh và ổn định, đồng thời phản ánh chính xác mô hình lan truyền âm trong COMSOL. Kết quả ước lượng tọa độ mục tiêu thủy lôi được trình bày trong Bảng 3.

Bảng 3. Kết quả ước lượng tọa độ mục tiêu thủy lôi và sai số định vị

Tham số	**Tọa độ thực (m)**	**Tọa độ ước lượng (m)**	**Sai số (m)**
x_o	40,000	40,900	0,900
y_o	0,000	0,000	0,000
z_o	−15,000	−14,623	0,377

Sai số định vị tổng hợp (chuẩn Euclid) là:

$$e_p = \sqrt{\delta x_o^2 + \delta y_o^2 + \delta z_o^2} = \sqrt{0,900^2 + 0,377^2} = 0,98\,\text{m}$$

Kết quả mô phỏng cho thấy phương pháp TDOA có khả năng xác định vị trí thủy lôi với độ chính xác cao trong điều kiện mô phỏng đã thiết lập.

Như vậy, với cấu hình mô phỏng đã chọn, thuật toán TDOA kết hợp với phương pháp Gauss–Newton dạng Levenberg–Marquardt cho phép xác định vị trí thủy lôi với sai số nhỏ, xấp xỉ 1 m trong không gian ba chiều.

Việc xác định tọa độ thủy lôi dựa trên phương pháp TDOA có thể chịu ảnh hưởng của một số yếu tố dẫn đến sai số. Thứ nhất, nhiễu môi trường nước biển, bao gồm nhiễu nền âm thanh và dao động của các thông số môi trường như nhiệt độ, độ mặn và áp suất, có thể làm sai lệch thời điểm xác định xung phản xạ.

Thứ hai, độ chính xác của phép đo thời gian đến xung sonar tại bộ thu trên robot ngầm, bao gồm giới hạn tần số lấy mẫu và khả năng lọc nhiễu bằng matched filter, cũng ảnh hưởng trực tiếp đến kết quả TDOA. Thứ ba, bố trí hình học của robot ngầm và tàu mẹ so với vị trí thủy lôi có thể ảnh hưởng đến độ ổn định của hệ phương trình TDOA; các cấu hình không tối ưu có thể làm tăng sai số định vị.

Tuy nhiên, phương pháp mô phỏng được thực hiện trong nghiên cứu này sử dụng COMSOL để xác định thời điểm xung phản xạ khi robot ngầm ở các vị trí xác định, kết hợp với mô phỏng thuật toán TDOA trên MATLAB để tính toán tọa độ thủy lôi, được đánh giá là tin cậy và phù hợp để kiểm chứng thuật toán. Việc tách riêng hai bước này giúp kiểm soát tốt nguồn dữ liệu đầu vào và đánh giá độ chính xác của phương pháp TDOA, đồng thời cung cấp cơ sở vững chắc cho các nghiên cứu và ứng dụng thực nghiệm sau này.

4. Kết luận

Bài báo đã đề xuất một giải pháp định vị thủy lôi trong đó nguồn sonar được đặt trên tàu mẹ và bộ thu tín hiệu được bố trí trên robot ngầm. Quá trình lan truyền sóng sonar trong nước biển được mô phỏng trên phần mềm COMSOL Multiphysics, từ đó xác định các thời điểm sóng phản xạ được nhận tại bộ thu sonar khi robot ngầm di chuyển tới bốn vị trí khác nhau. Dữ liệu thời gian phản xạ này sau đó được sử dụng để định vị thủy lôi bằng phương pháp TDOA. Thuật toán xác định tọa độ thủy lôi được triển khai và mô phỏng trên phần mềm MATLAB, cho thấy giải pháp có khả năng xác định chính xác vị trí thủy lôi trong môi trường giả lập.

Tài liệu tham khảo

1. He, C., Wang, Y., Yu, W., & Song, L. (2019). Underwater target localization and synchronization for a distributed SIMO sonar with an isogradient SSP and uncertainties in receiver locations. *Sensors, 19*(9), 1976. https://doi.org/10.3390/s19091976
2. Bao, Z., Zhang, Y., Tai, Y., Wang, J., Wang, H., Li, C., Hu, C., & Zhang, P. (2025). Multi-information-assisted bistatic active sonar target tracking for autonomous underwater vehicles in shallow water. *Remote Sensing, 17*(13), 2250. https://doi.org/10.3390/rs17132250
3. Liu, Y., Wang, Y., & Chen, C. (2023). Efficient underwater acoustical localization method based on TDOA with sensor position errors. *Journal of Marine Science and Engineering, 11*(4), 861. https://doi.org/10.3390/jmse11040861
4. Madsen, K., Nielsen, H. B., & Tingleff, O. (2004). *Methods for non-linear least squares problems* (2nd ed.). Lyngby, Denmark: Technical University of Denmark.
5. Liu, F., Chen, H., Zhang, L., & Xie, L. (2021). Time-difference-of-arrival-based localization methods of underwater mobile nodes using multiple surface beacons. *IEEE Access, 9*, 31712–31725. https://doi.org/10.1109/ACCESS.2021.3060565
6. COMSOL AB. (2018). *Acoustics module user's guide* (Version 5.4). Stockholm, Sweden: COMSOL AB.
7. Li, S., Sun, H., & Esmaiel, H. (2020). Underwater TDOA acoustical location based on majorization–minimization optimization. *Sensors, 20*(16), 4457. https://doi.org/10.3390/s20164457
8. Abraham, D. A. (2019). *Underwater acoustic signal processing: Modeling, detection, and estimation*. Cham, Switzerland: Springer International Publishing.
9. Hodges, R. P. (2010). *Underwater acoustics: Analysis, design and performance of sonar*. Chichester, UK: John Wiley & Sons.

Underwater Mine Detection Using the TDOA Method with Sonar Signals from a Mothership and a Multi-Position Remotely Operated Vehicle

Abstract: *This paper proposes a solution for underwater mine localization, in which the sonar source is placed on a mothership and the signal receiver is installed on a remotely operated vehicle (ROV). The propagation of chirp sonar pulses in seawater is simulated using COMSOL Multiphysics to determine the arrival times of reflected signals at the ROV as it moves to four predefined positions. A matched filter is used to extract the exact arrival times of the echo signals. These reflection time data are then used as input for a Time Difference of Arrival (TDOA) algorithm implemented in MATLAB, thereby estimating the coordinates of the underwater mine. Simulation results show that this method can reliably determine the mine position in a simulated environment. The solution combining COMSOL simulation and MATLAB provides a technical basis for research and experimental implementation in automatic localization of remotely operated vehicles.*

Keywords: Underwater acoustics; TDOA; underwater mine localization; Gauss–Newton algorithm; ROV.

Combination of compressed sensing and a deep Learning based model for multiclass MRI brain tumor classification

Pham Duc Khanh[1], Pham Ngoc Vo[2], Tran Anh Quang[3]

[1]*NCS K44 - Hệ 2 - Học viện kỹ thuật Quân sự;*

[2]*CH37B - Hệ 2 - Học viện kỹ thuật Quân sự;*

[3]*Bộ môn Điện tử y sinh-Viện Tên lửa và kỹ thuật điều khiển-Học viện Kỹ thuật Quân sự*

duckhanh.pham@lqdtu.edu.vn; Tel: 0975515815

Abstract

Brain tumor diagnosis based on magnetic resonance imaging (MRI) has become increasingly common in clinical practice. However, conventional MRI acquisition is time-consuming, which not only causes patient discomfort but also increases the risk of undesired artifacts such as motion noise. Recently, the integration of deep learning techniques for automated tumor recognition has shown great potential to assist radiologists, who traditionally rely on manual visual inspection of MRI scans. In this paper, we propose an efficient computer-aided diagnosis framework that combines Compressed Sensing MRI (CS-MRI) and EfficientNet-B0 for multi-class brain tumor classification. By exploiting CS-MRI, the proposed method achieves an acceleration factor of up to 5x in MRI acquisition and good image reconstruction. The reconstructed images are then automatically classified for detection of different brain tumor types. Our proposed approach is evaluated using 5-fold cross-validation under this severe undersampling conditions achieving a classification accuracy of up to 0.99 and an AUC close to 1.0. These results demonstrate the effectiveness of the proposed framework for fast and reliable brain tumor diagnosis in clinical settings.

Keyword: *Compressed sensing MRI; brain tumor classification; deep-learning; EfficientNet-B0.*

1.Introduction

Brain tumor diagnosis using magnetic resonance imaging (MRI) is widely adopted in clinical practice due to its non-invasive nature and high soft-tissue contrast. However, conventional MRI acquisition is time-consuming, increasing patient discomfort and susceptibility to motion artifacts that may affect diagnostic reliability [1, 2]. Various acceleration strategies have been proposed, including hardware optimization, parallel imaging, and k-space undersampling [3-7]. Among them, compressed sensing MRI (CS-MRI) enables faster acquisition by reconstructing images from highly undersampled data while preserving acceptable image quality [8]. Although extensive research has focused on improving CS-MRI reconstruction [9-11], limited attention has been paid to its impact on downstream tasks such as brain tumor classification. Moreover, many studies rely on a single train–test split, limiting the assessment of model generalization. In our previous work, we developed a hybrid CS-MRI model for efficient MRI reconstruction [12, 13].

Deep learning has demonstrated strong performance in automated brain tumor classification, with CNN-based models such as ResNet, EfficientNet, and VGG16 widely applied [14-21]. However, most approaches assume fully sampled, high-quality MRI data, which conflicts with the need for accelerated acquisition in clinical settings.

To address these challenges, this study proposes a computer-aided diagnosis framework integrating CS-MRI reconstruction and EfficientNet-B0 classification. Under severe undersampling (R = 0.2), the system achieves up to 5× acceleration while maintaining high diagnostic performance. The reconstructed images are classified into four categories (glioma,

meningioma, pituitary tumor, and no-tumor), and 5-fold cross-validation is employed for robust evaluation. The results demonstrate that accelerated MRI acquisition can be achieved without compromising classification accuracy, supporting its potential for practical clinical deployment.

2. Datasets and Methods

2.1. Dataset

The MRI dataset used in this study was obtained from a publicly available Kaggle, which is shared under the Creative Commons Attribution 4.0 International (CC BY 4.0) license [22]. It consists of brain MRI slices categorized into glioma, meningioma, pituitary tumor, and no-tumor classes. A total of 5,712 images were used for training and 1,311 for validation, with no overlap between the two sets. A 5-fold cross-validation scheme (4:1 split per fold) was applied to the training data. All images were resized to 256×256 and normalized prior to analysis.

Reconstruction performance of the hybrid CS-MRI method was evaluated using Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM). Classification performance was assessed using Accuracy, Precision, Recall, F1-score, and ROC–AUC, providing a comprehensive evaluation of model effectiveness (Eqs. 1–4):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

And the F1-score is calculated by:

$$F1 - score = 2\ \frac{Precison\ .Recall}{Precision+Recall} \tag{4}$$

2.2. The proposal of deep-learning based classification for CS-MRI dataset

In our previous work, we developed a hybrid CS-MRI approach combining random and Nyquist undersampling to achieve 3–5× acceleration without significant loss of image quality [12, 13]. In this study, the method was applied with an undersampling ratio of R = 0.2 (5× acceleration), using a 60:40 random-to-Nyquist sampling ratio as previously validated [12]. Image reconstruction was performed using the nonlinear conjugate gradient (NCG) algorithm, which ensures fast convergence and high reconstruction fidelity.

For classification, EfficientNet-B0 was employed as the backbone for both original and CS-MRI reconstructed datasets. The model leverages compound scaling and incorporates MBConv blocks with squeeze-and-excitation (SE) modules, enabling strong feature extraction with relatively low computational cost [23-30].

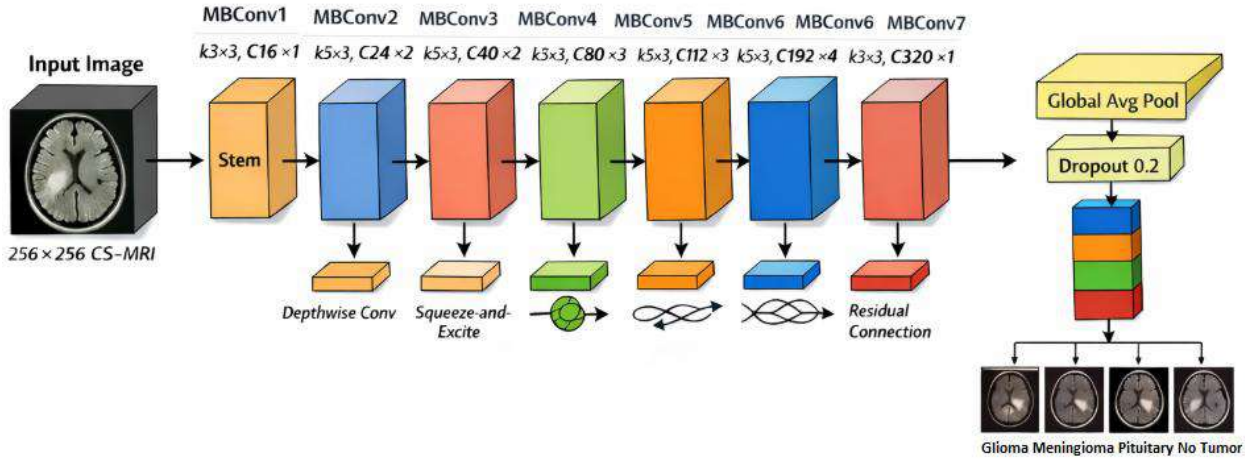


Fig. 1. The overall architecture of the proposed CS-MRI and EfficientNet-B0 framework for multi-class brain tumor classification.

As illustrated in Fig. 1, the EfficientNet-B0 architecture consists of an initial convolutional stem followed by a sequence of MBConv blocks equipped with depthwise separable convolutions and SE modules. A global average pooling layer is applied before the final fully connected classification layer, which is modified to output four classes corresponding to glioma, meningioma, pituitary tumor, and no-tumor cases. Pretrained on ImageNet, the network was fine-tuned for four-class brain tumor classification (glioma, meningioma, pituitary tumor, and no-tumor). Single-channel MRI images were converted to three-channel inputs to meet network requirements. EfficientNet-B0 was selected due to its favorable balance between accuracy and efficiency, superior parameter utilization compared to ResNet-18, robustness to reconstruction artifacts via SE attention, and fast inference capability, making it suitable for accelerated MRI applications. Two experimental settings were evaluated: training on fully sampled MRI images and training on CS-MRI reconstructed images. Performance was assessed using 5-fold cross-validation, and an ensemble strategy was applied to improve robustness and generalization.

3. Results

3.1. Compressed sensing MRI reconstruction performance

The hybrid CS-MRI strategy was applied to the entire MRI dataset by simulating undersampled acquisitions with an undersampling ratio of 0.2, corresponding to an acceleration factor of 5x, and a random sampling proportion of 0.6, as described in Section 2.2. The reconstructed CS-MRI images obtained under this aggressive acceleration setting were able to preserve the main anatomical structures as well as tumor-related characteristics with high fidelity. As illustrated in Fig. 5, the visual quality of the reconstructed images remains highly consistent with that of the fully sampled MRI scans, demonstrating the effectiveness of the proposed hybrid CS-MRI scheme. In Fig. 2, the original MRI image (Fig. 2a) is undersampled by the hybrid CS-MRI with an undersampling patten (Fig. 2b) and their reconstructed images by the NCG algorithm in Fig. 2c. It can be seen that, for an acceleration factor of 5x and a Random/Nyquist sampling

proportion of 60/40, the PSNR and SSIM are respectively 24.65 and 0.8639 as compared to the original image.

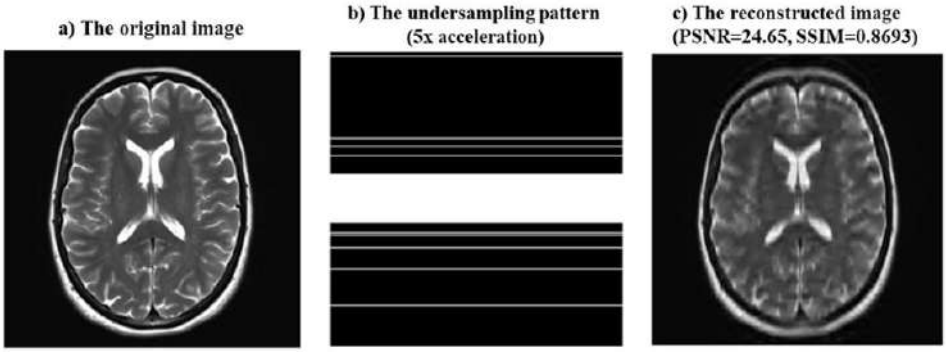


Fig. 2. The original and reconstructed undersampled images with their corresponding CS-MRI pattern

Visual examination of the reconstructed images shows that tumor boundaries and intensity distributions remain clearly distinguishable, making the images suitable for subsequent automated analysis. These results demonstrate that the proposed CS-MRI approach can substantially reduce acquisition time while preserving diagnostically meaningful image quality, which is essential for the effectiveness of the subsequent deep learning–based classification framework.

3.2. Classification performance on original MRI dataset

As mentioned in Section 2.2, the EfficientNet-B0 model with optimal parameters was used to classify 4 classes including 3 tumor classes (glioma, meningioma, pituitary) and 1 non-tumor class. The results obtained on the original MRI dataset are shown in Fig. 3 with an impressive confusion matrix and an ROC-AUC curve close to 1.0. Fig. 3 presents the classification results of the proposed EfficientNet-B0 model on the original MRI dataset in terms of the confusion matrix and multi-class ROC curves. The confusion matrix shows near-perfect classification performance across all four classes, with only four misclassified samples out of 1311 images. Furthermore, the ROC curves demonstrate excellent discriminative capability, with the area under the curve (AUC) reaching 1.000 for all four classes. This indicates almost perfect class separability and confirms the robustness and reliability of the proposed EfficientNet-B0 model on high-quality, fully sampled MRI images.

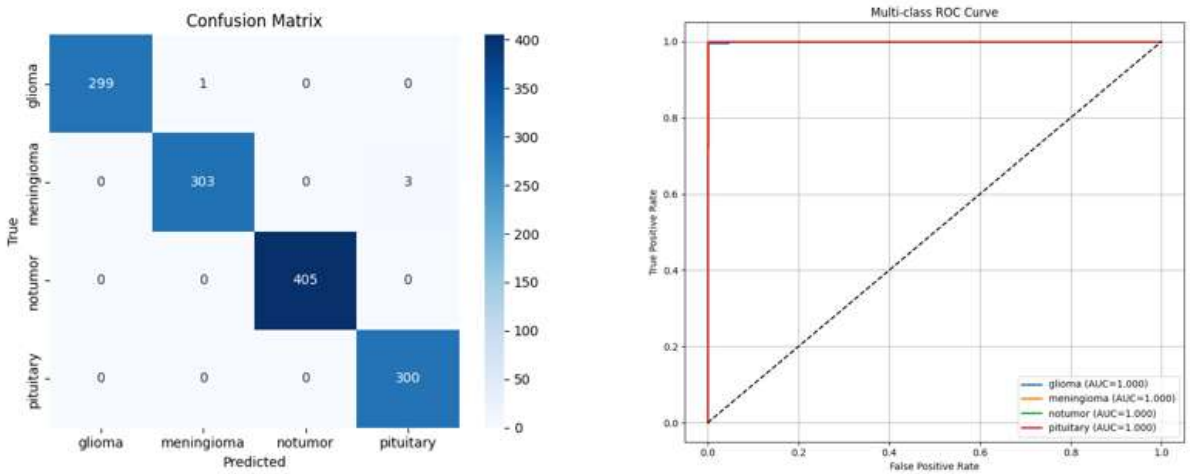


Fig. 3. The result of classifications by EfficientNet-B0 model on original MRI dataset

The high value of Accuracy, precision, recall, and F1-score are specifically shown in the Table 1, which indicates that effective classification performance of the EfficientNet-B0 model trained and evaluated on original fully sampled MRI images.

Table 1. The result of classifications by EfficientNet B0 model on original MRI images

Cases	Precision	Recall	F1-score	Support
Glioma	1.000	0.9967	0.9983	300
Meningioma	0.9967	0.9902	0.9934	306
Notumor	1.000	1.000	1.000	405
Pituitary	0.9901	1.000	0.9950	300
Accuracy		0.9969		1311

3.3. Classification performance on CS- MRI dataset

As described in section 3.2, the same EfficientNet-B0 model used on the original MRI dataset, was applied to the reconstructed CS-MRI dataset. Despite being reconstructed from highly undersampled k-space data (5× acceleration), the model maintains high classification accuracy and robustness, indicating strong discriminative capability under aggressive CS-MRI conditions. The multi-class ROC curves show AUC values of 1.000 for all four classes, demonstrating near-perfect separability despite reconstruction artifacts. Compared to the original MRI dataset, only minor performance degradation is observed, confirming the reliability of the combined CS-MRI and EfficientNet-B0 framework. Interestingly, this proposed framework still achieved highly promising classification performance, as illustrated in Fig. 4. Overall, the results demonstrate that accurate and stable multi-class brain tumor classification can be achieved under 5× accelerated acquisition.

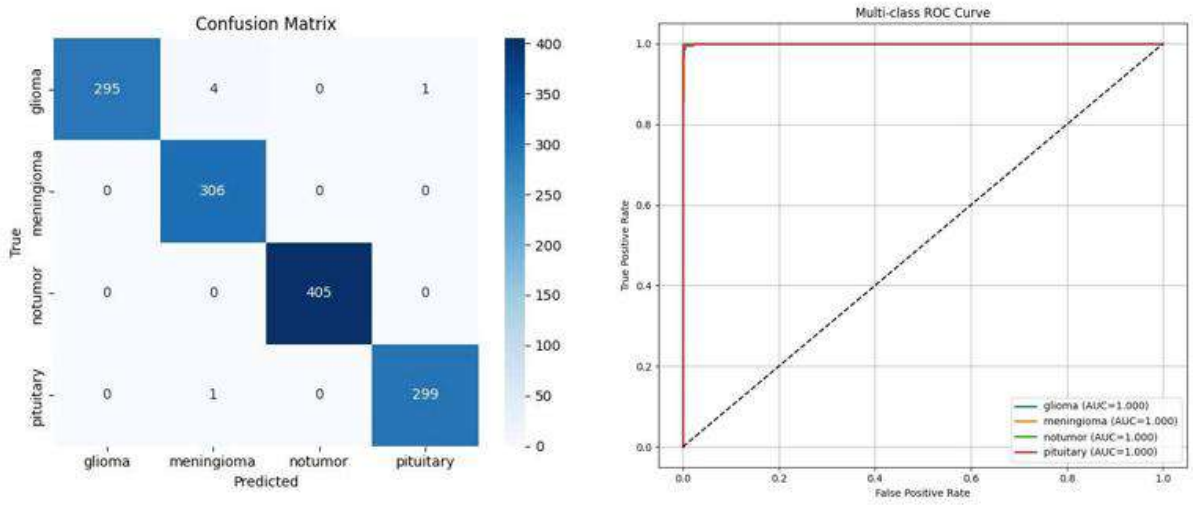


Fig. 4. The result of classifications by EfficientNet-B0 model on CS-MRI dataset

As summarized in Table 2, the EfficientNet-B0 model achieves consistently high accuracy, precision, recall, and F1-score on the CS-MRI dataset.

Table 2. The result of classifications by EfficientNet-B0 model on CS MRI images

Cases	Precision	Recall	F1-score	Support
Glioma	1.000	0.9833	0.9916	300
Meningioma	0.9839	1.000	0.9919	306
Notumor	1.000	1.000	1.000	405
Pituitary	0.9967	0.9967	0.9967	300
Accuracy		0.9954		1311

To compare the performance of the EfficientNet-B0 model on two datasets, Table 3 shows the performance of the model on the original MRI and the CS-MRI datasets. On the original dataset, the model achieves an accuracy of 0.9969 with a F1-score of 0.9967, while on the CS-MRI dataset, the performance slightly decreases to an accuracy of 0.9954 and an F1-score of 0.9947. Although a negligible performance degradation is observed, the results remain consistently high, demonstrating that the proposed framework is highly robust to CS-MRI reconstruction artifacts and maintains reliable classification performance under aggressive undersampling conditions.

Table 3. Comparison of the EfficientNet B0 model on original and CS-MRI datasets

EfficientNet-B0 models	Accuracy	Precision	Recall	F1-score
Original dataset	0.9969	0.9970	0.9972	0.9967
CS-MRI dataset	**0.9954**	**0.9955**	**0.9954**	**0.9947**

4. Discussions

This study investigates the impact of aggressive CS-MRI undersampling (5× acceleration) on multi-class brain tumor classification. Unlike most prior studies relying on fully sampled MRI data, our framework evaluates performance under realistic accelerated acquisition conditions. As shown in Table 3, classification accuracy decreases only slightly from 0.9969 (original MRI) to 0.9954 (CS-MRI), with negligible differences in precision, recall, and F1-score, indicating that the

hybrid CS-MRI reconstruction preserves diagnostically relevant features. For a more comprehensive evaluation, Table 4 compares our method with recent CNN-based approaches, including VGG16 [31], ensemble models [32, 33], and various CNN architectures [33-35]. While these studies report high accuracy (93.3%–99.24%) using fully sampled MRI data, our EfficientNet-B0 model achieves 99.54% accuracy on highly undersampled CS-MRI data (5× acceleration). Notably, this performance surpasses several state-of-the-art methods trained on fully sampled data, including the Hybrid Ensemble model (99.09%) [33]. These findings demonstrate that optimizing both reconstruction and classification stages enables reliable diagnosis without increasing network complexity, highlighting the strong clinical potential of the proposed framework.

Table 4. Classification performance comparison of deep learning models on fully sampled and compressed-sampled MRI datasets.

Study	Models	Dataset Source	Sampling Condition	Accuracy
[31]	VGG16	Figshare and Kaggle	Fully sampled	99.24%
[32]	Ensemble of EfficientNet, VGG16, ResNet & others	Cancer Imaging Archive (REMBRANDT)	Fully sampled	98.47%
[34]	CNN	Kaggle	Fully sampled	93.3%
[35]	Custom CNN	Kaggle	Fully sampled	96.63%
[36]	VGG	Kaggle	Fully sampled	97%
[33]	Hybrid Ensemble	Kaggle	Fully sampled	99.09%
This Research	**EfficientNet-B0**	**Kaggle**	**CS-MRI (5x acceleration)**	**99.54%**

Furthermore, the overall comparison in Table 5 confirms the superior effectiveness of the proposed model, achieving the highest accuracy (0.9954) and competitive precision and recall. Notably, the AUC reaches 1.0000, reflecting outstanding class separability and highly reliable decision boundaries. Achieving such performance under degraded acquisition conditions demonstrates that the proposed framework not only matches but exceeds several state-of-the-art models trained on fully sampled data. From these findings, it confirms that integrating hybrid CS-MRI reconstruction with a lightweight yet powerful EfficientNet-B0 architecture enables fast MRI acquisition without compromising classification reliability. Maintaining near state-of-the-art diagnostic performance at a 5x acceleration factor highlights the strong potential of the proposed framework for practical and time-sensitive clinical deployment.

Table 5. Comparison of classification performance among different models

Models	Accuracy	Precision	Recall	F1-score	AUC
ResNet50 [37]	0.99	0.99	0.99	0.99	0.68
ResNet50V2 [37]	0.878	0.76	0.71	0.73	0.50
InceptionV3 [38]	0.9712	0.9797	0.9659	0.9727	0.9984

Xception[39]	0.9527	0.9517	0.9486	0.9491	N/M
Hybrid Ensemble [33]	0.9925	0.9925	0.9925	0.9925	N/M
This Research (EfficientNet-B0)	**0.9954**	**0.9955**	**0.9954**	**0.9947**	**1.0000**

5. Conclusion

This study presents an effectively combined framework that integrates compressed sensing–based MRI acquisition with a deep learning–based classification model for multi-class brain tumor diagnosis. Experimental results demonstrate that the EfficientNet-B0 network maintains high classification performance on CS-MRI images acquired with a 5x acceleration factor, achieving 99.69% accuracy on fully sampled MRI and maintains 99.54% accuracy with AUC = 1.000 on CS-MRI images, showing only negligible degradation. Compared with recently reported methods trained on fully sampled data, our proposed approach delivers competitive or superior performance despite operating under aggressive acceleration. These results demonstrate that substantial acquisition acceleration can be achieved without compromising classification reliability, highlighting the effectiveness of jointly optimizing reconstruction and deep learning–based diagnosis for practical clinical deployment.

References

1. Vovoras, D., K.P. Pokhrel, and C.P. Tsokos. (2014). Epidemiology of Tumors of the Brain and Central Nervous System: Review of Incidence and Patterns among Histological Subtypes. *Open Journal of Epidemiology*. 04(04): p. 224-234.
2. Krolikowska, K., et al. (2025). Glioblastoma-A Contemporary Overview of Epidemiology, Classification, Pathogenesis, Diagnosis, and Treatment: A Review Article. *Int J Mol Sci*. 26(24).
3. Sharma, P., A.K. Dubey, and A. Goyal. *Hardware Acceleration Using SIMD Based Vector Processor Array to Enhance Performance of Deblurring Methods for CT and MRI Images Having Motion Blur Artifacts*. in *2023 International Conference on Smart Systems for applications in Electrical Sciences (ICSSES)*. 2023.
4. Deshmane, A., et al. (2012). Parallel MR imaging. *J Magn Reson Imaging*. 36(1): p. 55-72.
5. Corbin, N., et al., *Fast Imaging and Acceleration Techniques*. 2024. p. 51-74.
6. Kemp, T.M., et al. (2023). Optimizing data acquisition in undersampled magnetic resonance imaging (MRI) using two alternative forced choice (2-AFC) and search tasks. *Proc SPIE Int Soc Opt Eng*. 12467.
7. Razumov, A., O. Rogov, and D.V. Dylov. (2023). Optimal MRI undersampling patterns for ultimate benefit of medical vision tasks. *Magn Reson Imaging*. 103: p. 37-47.
8. Lustig, M., et al. (2008). Compressed Sensing MRI. *IEEE Signal Processing Magazine*. 25(2): p. 72-82.
9. Huang, J., et al. (2025). Data- and Physics-Driven Deep Learning Based Reconstruction for Fast MRI: Fundamentals and Methodologies. *IEEE Reviews in Biomedical Engineering*. 18: p. 152-171.
10. Heckel, R., et al. (2024). Deep learning for accelerated and robust MRI reconstruction. 37(3): p. 335-368.
11. Elbehairy, A., I. Yassine, and M. Elattar, *Stage U-Net Framework: Streamlining MRI Reconstruction from Under- sampled K-Space*. 2024.
12. Khanh Pham, D., D.-T. Tran, and A. Quang Tran. (2025). Evaluating random–Nyquist sampling ratios in combined compressed sensing magnetic resonance imaging. *Bulletin of Electrical Engineering and Informatics*.

13. Tran, A.Q., et al. (2020). MRI Simulation-based evaluation of an efficient under-sampling approach. *Mathematical Biosciences and Engineering*. 17(4): p. 4048-4063.
14. Dorfner, F.J., et al. (2025). A review of deep learning for brain tumor analysis in MRI. *npj Precision Oncology*. 9(1): p. 2.
15. Abdusalomov, A.B. and M. Mukhiddinov. (2023). Brain Tumor Detection Based on Deep Learning Approaches and Magnetic Resonance Imaging. 15(16).
16. Mienye, I.D., et al. (2025). Deep Convolutional Neural Networks in Medical Image Analysis: A Review. *Information*. 16(3).
17. Chen, C., N. Isa, and X. Liu. (2024). A review of convolutional neural network based methods for medical image classification. *Computers in biology and medicine*. 185: p. 109507.
18. Li, Q., et al. *Medical image classification with convolutional neural network*. in *2014 13th International Conference on Control Automation Robotics & Vision (ICARCV)*. 2014.
19. Raptis, C., et al. (2024). Comparative Analysis of Conventional CNN v's ImageNet Pretrained ResNet in Medical Image Classification. *Information*. 15(12).
20. Ali, R.R., et al. (2025). Learning Architecture for Brain Tumor Classification Based on Deep Convolutional Neural Network: Classic and ResNet50. *Diagnostics (Basel)*. 15(5).
21. Sumedha Arya, N.G. (2025). ResLink: A Novel Deep Learning Architecture for Brain Tumor Classification with Area Attention and Residual Connections.
22. HIRA, M.I.K.H., MD SOHAG; BITHEE, MST MORIOM AKTER ; Sara, Umme Sara; HASAN, MD MAHMUDUL; Towsif, Abdullah Al; Ahmed, Md Kowsar *Brain Tumor MRI Dataset (Glioma, Meningioma, Pituitary, No Tumor)*, M. Data, Editor. 2025.
23. Rahman, T., et al. (2026). Optimized CNN framework with VGG19, EfficientNet, and Bayesian optimization for early colon cancer detection. *Scientific Reports*. 16(1): p. 4128.
24. Abd El-Aziz, A.A., M.A. Mahmood, and S.A. El-Ghany. (2025). EfficientNet-B3-Based Automated Deep Learning Framework for Multiclass Endoscopic Bladder Tissue Classification. *Diagnostics*. 15(19): p. 2515.
25. Ullah, Z., et al. (2025). Systematic Integration of Attention Modules into CNNs for Accurate and Generalizable Medical Image Classification. *Mathematics*. 13(22): p. 3728.
26. Muthulakshmi, K. and M. Jayalakshmi. (2025). Rethinking model of EfficientNet-B9 for brain tumor classification: A high-precision deep learning approach. *Results in Engineering*. 28: p. 107984.
27. Kailani, N. (2025). Comparison between ResNet and EfficientNet for Image Classification – An Analytical Study of Performance and Efficiency. *شمال إفريقيا للنشر العلمي □□□□ (NAJSP)*: p. 18-25.
28. Hu, J., L. Shen, and G. Sun. *Squeeze-and-Excitation Networks*. in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2018.
29. Iqbal, A., M.A. Jaffar, and R. Jahangir. (2024). Enhancing Brain Tumour Multi-Classification Using Efficient-Net B0-Based Intelligent Diagnosis for Internet of Medical Things (IoMT) Applications. *Information*. 15(8): p. 489.
30. Ali, H., et al. (2025). A fine tuned EfficientNet-B0 convolutional neural network for accurate and efficient classification of apple leaf diseases. *Sci Rep*. 15(1): p. 25732.
31. Wong, Y., et al. (2025). Brain tumor classification using MRI images and deep learning techniques. *PLoS One*. 20(5): p. e0322624.
32. Tandel, G.S., A. Tiwari, and O.G. Kakde. (2025). Multi-Class Brain Tumor Grades Classification Using a Deep Learning-Based Majority Voting Algorithm and Its Validation Using Explainable-AI. *J Imaging Inform Med*. 38(5): p. 2793-2830.
33. Yoon, S. (2025). Brain tumor classification using a hybrid ensemble of Xception and parallel deep CNN models. *Informatics in Medicine Unlocked*. 54.
34. Mahmud, M.I., M. Mamun, and A. Abdelgawad. (2023). A Deep Analysis of Brain Tumor Detection from MR Images Using Deep Learning Networks. *Algorithms*. 16(4): p. 176.

35. T. Lakshmi Prasanthi, N.N. (2024). Improvement of Brain Tumor Categorization using Deep Learning: A Comprehensive Investigation and Comparative Analysis.
36. Shamshad, N., et al. (2024). Enhancing Brain Tumor Classification by a Comprehensive Study on Transfer Learning Techniques and Model Efficiency Using MRI Datasets. *IEEE Access*. 12: p. 100407-100418.
37. Mizanur Rahman, M., et al. (2025). Efficient brain cancer identification using ResNet50 and ResNet50 V2: a comparative study with a primary MRI dataset. *Bulletin of Electrical Engineering and Informatics*. 14(4): p. 2849-2860.
38. Gómez-Guzmán, M.A., et al. (2023). Classifying Brain Tumors on Magnetic Resonance Imaging by Using Convolutional Neural Networks. *Electronics*. 12(4).
39. Disci, R., F. Gurcan, and A. Soylu. (2025). Advanced Brain Tumor Classification in MR Images Using Transfer Learning and Pre-Trained Deep CNN Models. *Cancers (Basel)*. 17(1).

Tiêu đề: Nghiên cứu mô hình kết hợp lấy mẫu nén và mạng học sâu để phân loại hiệu quả u não đa lớp trên ảnh cộng hưởng từ

Tóm tắt: *Chẩn đoán u não dựa trên ảnh chụp cộng hưởng từ (MRI) ngày càng phổ biến trong thực tiễn lâm sàng. Tuy nhiên, quá trình thu nhận dữ liệu trong tạo ảnh MRI thường tốn rất nhiều thời gian, không chỉ gây khó chịu cho bệnh nhân mà còn làm tăng nguy cơ xuất hiện các nhiễu trên ảnh. Gần đây, việc ứng dụng các kỹ thuật học sâu để tự động phát hiện khối u não đã cho thấy tiềm năng lớn trong hỗ trợ chẩn đoán hình ảnh, thay vì các bác sĩ phải đánh giá thủ công dựa trên kinh nghiệm trên các ảnh MRI. Trong nghiên cứu này, nhóm nghiên cứu đề xuất xây dựng mô hình chẩn đoán dựa trên mạng học sâu EfficientNet-B0 kết hợp với kỹ thuật lấy mẫu nén MRI (CS-MRI) để phân loại u não đa lớp trên ảnh MRI một cách hiệu quả. Với việc kết hợp kỹ thuật CS-MRI, phương pháp đề xuất cho phép tăng tốc tốc độ thu nhận ảnh lên đến 5 lần trong khi vẫn giữ được chất lượng hình ảnh tái tạo cần cho việc chẩn đoán. Các hình ảnh tái tạo này được phân loại tự động bởi mô hình học sâu để phát hiện các loại u não khác nhau. Với phương pháp đề xuất này, kết quả phân loại của mô hình được kiểm chứng chéo 5 lần trên tập ảnh lấy mẫu nén, cho độ chính xác phân loại lên đến 0,99 và chỉ số AUC xấp xỉ 1,0. Các kết quả đạt được của nghiên cứu đã chứng minh hiệu quả của mô hình đề xuất trong việc hỗ trợ phát hiện u não đa lớp nhanh chóng và đáng tin cậy trong chẩn đoán lâm sàng.*

Từ khóa: Chụp cộng hưởng từ lấy mẫu nén; phân loại khối u não; học sâu; EfficientNet-B0.

Nghiên cứu tính toán và chế tạo thiết bị tự nổi đóng vai trò mìn huấn luyện sử dụng cho Robot Pluto Plus

Nguyễn Xuân Tiến[1], Trương Đăng Khoa[1], Đỗ Minh Triều[2], Nguyễn Quang Huấn[3]
[1]*Học viện Kỹ thuật Quân sự;* [2]*Trường sỹ quan kỹ thuật ;* [3]*Viện thiết kế Tàu quân sự*
tienxuannguyen86@gmail.com; khoahn67@gmail.com

Tóm tắt

Trong bài báo này, một thiết bị tự nổi đóng vai trò mìn huấn luyện robot Pluto plus Hải quân được nghiên cứu và thiết kế. Dựa trên nguyên lý Archimedes và các phương trình cân bằng lực thủy tĩnh, tác giả xây dựng mô hình toán học mô tả quá trình nổi – chìm của mìn huấn luyện thông qua cơ chế thay đổi khối lượng bằng ruột tích áp cao su kết hợp khí nén. Trên cơ sở mô hình lý thuyết, kết cấu cơ khí được tính toán kiểm bền, phân tích ổn định nổi và đánh giá khả năng làm việc trong điều kiện môi trường biển. Tính ổn định của hệ thống được phân tích dựa trên tiêu chuẩn cân bằng lực và điều kiện metacentric. Kết quả thử nghiệm thực tế cho thấy thiết bị đáp ứng các yêu cầu kỹ thuật, phục vụ cho huấn luyện tác chiến lực lượng chống ngầm của Hải quân.

Từ khóa: *Thiết bị tự nổi; Mìn huấn luyện; Robot Pluto plus; Lực nổi Archimedes.*

1. Giới thiệu

Trong tác chiến chống thủy lôi hiện đại, việc sử dụng robot điều khiển từ xa (ROV) là xu thế tất yếu nhằm giảm thiểu rủi ro cho con người và nâng cao hiệu quả dò tìm, xử lý mục tiêu dưới nước. Robot Pluto plus hiện đang được trang bị tại Hải quân Việt Nam phục vụ nhiệm vụ dò tìm và nhận dạng mìn [1, 2]. Để bảo đảm huấn luyện an toàn, mìn thật không thể sử dụng trực tiếp trong quá trình thực hành. Vì vậy, mìn huấn luyện có vai trò đặc biệt quan trọng, phải đảm bảo: Tương đương mìn thật về kích thước và khối lượng; Có đặc tính thủy động học tương tự; Đảm bảo an toàn tuyệt đối; Có khả năng tự nổi sau khi hoàn thành bài huấn luyện.

Vấn đề đặt ra là thiết kế một cơ cấu tự nổi đóng vai trò mìn huấn luyện, đảm bảo hoạt động ổn định, tin cậy trong môi trường biển, chịu được áp lực nước và có thể hoạt động nhiều chu kỳ lặp lại [1-3]. Đối với các thiết bị ngầm có khả năng tự nổi, có thể sử dụng các ứng dụng thiết kế sau: Bơm khí vào phao để tạo lực nổi, thường dùng trong cứu hộ, trục vớt thiết bị nhẹ hoặc đánh dấu vị trí thiết bị ngầm; Thiết kế kiểu piston, thay đổi thể tích choán nước để thay đổi lực nâng, thường dụng trong các AUV thám hiểm mặt biển (AUV Slocum G3 Glider của công ty Teledyne Marine [5]; Thiết kế kiểu khối dằn (ballast tank): bằng cách thay đổi lượng nước/khí trong khối để thay đổi lực nổi, thường áp dụng cho các phương tiện ngầm cỡ lớn (tàu ngầm, ngư lôi…).

2. Giải pháp thiết kế

2.1. Cơ sở lý thuyết về sức nổi

Mìn huấn luyện được thiết kế với khả năng tự nổi sau bài huấn luyện, như một vật nổi trong môi trường chất lỏng, tuân theo định luật Archimedes [4]:

$$F_b = \rho g V \tag{1}$$

Trong đó: F_b - lực đẩy Archimedes (N), ρ - mật độ chất lỏng ($\mathrm{kg.m^{-3}}$). Đối với nước ngọt, $\rho = 1000$, đối với nước biển $\rho = 1020 - 1050$, g - gia tốc trọng trường ($9.81\mathrm{m.s^{-2}}$), V - thể tích chất lỏng bị chiếm chỗ (m^3). Trọng lượng vật trong môi trường chất lỏng:

$$W = mg \tag{2}$$

Sự chênh lệch giữa lực đẩy Archimedes và trọng lượng vật trong môi trường chất lỏng được gọi là sức nổi (Buoyancy force):

$$P_b = F_b - W \quad (3)$$

Nếu $P_b > 0$: sức nổi dương, vật nổi lên; Nếu $P_b < 0$: sức nổi âm, vật chìm xuống; Nếu $P_b \approx 0$: sức nổi bằng không (neutral buoyancy), vật lơ lửng. Từ (1) và (2), điều kiện giữ cho vật nổi lơ lửng trong chất lỏng:

$$\rho g = m \quad (4)$$

Trong môi trường chất lỏng đồng nhất $\rho = const$, do đó để vật chìm ở một độ sâu nhất định cần phải thay đổi trọng lượng vật ở độ sâu đó. Khi mìn nổi lên với tốc độ V_d, lực cản thủy động học:

$$F_d = \frac{1}{2}\rho V_d^2 C_d S_d \quad (5)$$

Trong đó: C_d - hệ số lực cản trong môi trường nước; S_d - diện tích cắt ngang của mìn theo phương thẳng đứng. Cân bằng giữa sức nổi với lực cản thủy động học sẽ xác định được tốc độ nổi của mìn:

$$P_b = F_d \quad (6)$$

$$F_b - W = \frac{1}{2}\rho V_d^2 C_d S_d \quad (7)$$

$$v_d = \sqrt{\frac{2(F_b - W)}{\rho C_d S_d}} \quad (8)$$

Để thiết kế thiết bị đóng vai trò mìn huấn luyện, cần đảm bảo khả năng điều chỉnh độ sâu để tiếp cận mục tiêu, đồng thời có thể nổi lên mặt nước sau khi huấn luyện, sử dụng phương pháp thay đổi trọng lượng trung bình của mìn theo hình thức thiết kế dạng khối dằn: bơm nước vào khối dằn để mìn chìm xuống, đẩy nước khỏi khối dằn để mìn nổi lên.

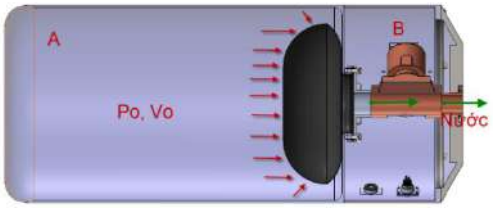
2.2. Cấu hình cơ bản của mìn huấn luyện

Cấu hình chính của mìn bao gồm: Vật liệu thân vỏ mìn được làm từ thép A; Phần bên trong chia làm hai khoang (ký hiệu A, B), ngăn cách nhau bằng vách ngăn; Khối dằn C, bố trí trong khoang A; Khối động cơ và van bơm – xả, bố trí trong khoang B; Khối dằn và khối động cơ, van bơm – xả được kết nối với nhau thông qua hệ van ống. Các tham số kết cấu chính của mìn huấn luyện: Khối lượng mìn (khi khối dằn được bơm đầy nước) (m_1): 45 kg; Khối lượng mìn (khi khối dằn không tích nước) (m_0): 31.5 kg; Chiều dài mìn (L_1+L_2): 760 ±10 mm; Đường kính (D): Φ260 mm; Thể tích khoang A (P_0): 0.02 m^3 ; Thể tích khối dằn C (P_C): 0.0135 m^3 ; Thể tích của mìn (P_Σ): 0.04 m^3 ; Vật liệu khối dằn bằng cao su EPDM (Ethylene Propylene Diene Monomer) có độ đàn hồi tốt.

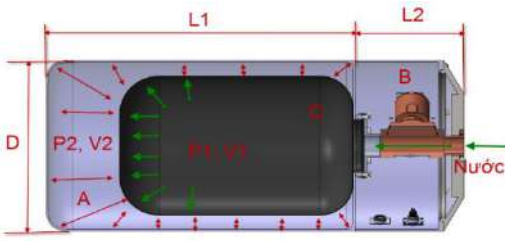
2.3. Nguyên lý hoạt động mìn huấn luyện

Khí trơ N_2 (hoặc không khí) được bơm vào khoang A với áp suất P_0, thể tích ban đầu V_0. Nước được bơm vào đầy khối dằn với thể tích V_1 lít nước ngọt qua hệ ống dẫn, lượng khí trong khoang A được nén lại sẽ có thể tích là V_2 (lít) với áp suất khi đó là P_2, tác động lên vỏ khối dằn C.

Gọi P_0, V_0 là áp suất, thể tích khoang A chứa khối dằn cao su lúc ban đầu (khối dằn chưa được tích nước); P_1, V_1 – áp suất, thể tích khối dằn cao su khi tích đầy nước; P_2, V_2 - áp suất, thể tích khoang A khi khối dằn C được tích đầy nước. Nước được bơm theo đường ống vào khối dằn (C) (hình 2), khi khối dằn tích đầy nước, tổng khối lượng mìn lúc này G= 45kg.



Hình 1. Mìn khi khối dằn xả hết nước



Hình 2. Mìn khi khối dằn tích đầy nước

Lực đẩy Archimedes theo (1):

$$F_b = \rho g V = 1025 * 9.81 * 0.04 = 402.21(\mathrm{N}) \tag{9}$$

Trọng lượng mìn khi khối dằn tích đầy nước:

$$W_1 = m_1 g = 45 * 9.81 = 441.45(\mathrm{N}) \tag{10}$$

Do trọng lượng mìn lớn hơn lực đẩy Archimedes nên mìn có sức nổi âm và mìn sẽ chìm xuống. Khi nước được bơm xả ra khỏi khối dằn theo đường ống (hình 1), khối lượng mìn giảm về khối lượng rỗng ban đầu (m_0), trọng lượng mìn khi này:

$$W_0 = m_0 g = 31.5 * 9.81 = 309.015(\mathrm{N}) \tag{11}$$

Trọng lượng mìn nhỏ hơn lực đẩy Archimedes nên mìn có sức nổi dương và mìn sẽ nổi lên.

2.4. Tính toán thiết kế

2.4.1. Tính toán thiết kế chu trình xả nước khi nổi

Vì mìn huấn luyện sử dụng cơ chế đẩy nước kiểu vòi phun **(**water jet propulsion**)** để giảm trọng lượng nhằm tạo sức nổi dương để mìn nổi lên, cơ chế này sẽ tạo ra một phản lực, đẩy mìn dịch chuyển trong quá trình nổi lên; điều này dẫn tới việc khó khăn khi tìm kiếm thu hồi, đặc biệt trong điều kiện nước có dòng chảy mạnh. Để giải quyết hạn chế này, việc điều khiển mìn trong quá trình nổi được thực hiện thông qua việc đẩy nước dưới dạng "xung" thông qua van xả với các khoảng thời gian được tính toán. Trong cơ học chất lỏng, phương trình động lượng cho dòng tia nước được đẩy ra khỏi mìn qua van xả được tính như sau:

$$F_J = \dot{m}\left(v_{out} - v_{in}\right) \tag{12}$$

Trong đó: $\dot{m} = \rho Q$ - lưu lượng khối lượng nước đẩy ra khỏi mìn ($\mathrm{kg.s^{-1}}$); Q - lưu lượng thể tích nước ($\mathrm{m^3.s^{-1}}$); v_{out} - tốc độ luồng nước đẩy qua van xả ($\mathrm{m.s^{-1}}$); v_{in} - tốc độ luồng nước bao quanh mìn ($\mathrm{m.s^{-1}}$). Trong trường hợp không có dòng chảy tại nơi làm việc của mìn $v_{in} = 0$. Tốc độ luồng nước đẩy qua van xả tính theo định luật Torricelli**:**

$$v_{out} = \sqrt{\frac{2\left(P_{in} - P_{out}\right)}{\rho}} \tag{13}$$

Dòng luồng nước được đẩy ra khỏi mìn sẽ tạo ra một lực đẩy, đẩy mìn dịch chuyển theo hướng ngược chuyển động của luồng nước qua van xả với tốc độ:

$$v_n = \sqrt{\frac{2F_J}{\rho C_d S_{\mathrm{w}}}} = \sqrt{\frac{2F_J}{\rho C_d S_{\mathrm{w}}}} \tag{14}$$

Trong đó: C_d - hệ số lực cản, giá trị trong khoảng 0.04−0.2 với hình dạng thuôn của mìn huấn luyện; S_{w} - diện tích xung quanh mìn. Kết hợp (8) với (14), mìn huấn luyện khi nổi lên sẽ tham gia

đồng thời 2 chuyển động: Chuyển động thẳng đứng với tốc độ V_d; Chuyển động ngang với tốc độ V_n. Theo (8) tốc độ v_d với $C_d = 0.1$; $S_d = 0.198\text{m}^2$:

$$v_d = \sqrt{\frac{2(F_b - W)}{\rho C_d S_d}} \approx 3.07\left(\text{m.s}^{-1}\right) \tag{15}$$

Theo (14) tốc độ v_n với $C_d = 0.1$; $S_w = 0.62\text{m}^2$, đường kính van xả $D_v = 0.027\text{m}$, áp suất nước đẩy qua van xả $P = 10\text{atm}$

$$v_n = \sqrt{\frac{2F_J}{\rho C_d S_w}} \approx 4.75\left(\text{ms}^{-1}\right) \tag{16}$$

Như vậy, mìn sẽ bị "dạt" khi nổi lên với góc so với phương thẳng đứng:

$$\beta = \arccos\frac{3.07}{\sqrt{3.07^2 + 4.75^2}} \approx 57.2^0 \tag{17}$$

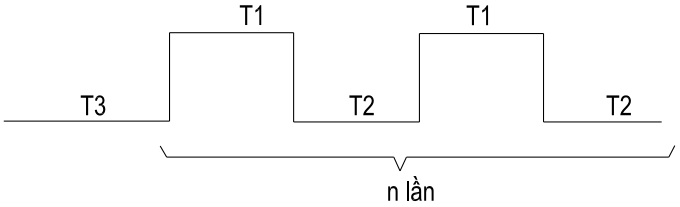
Giả định mìn đặt ở độ sâu H=25m, khoảng cách "dạt" của mìn khi nổi:

$$R = H \times \arctan\beta \approx 19.6\text{m} \tag{18}$$

Thời gian mìn nổi lên mặt nước:

$$t_d = \frac{H}{v_d} = \frac{25}{3.07} \approx 8.14\text{s} \tag{19}$$

Giá trị $R = 19.6\text{m}$ được tính khi điều kiện không có dòng nước chảy tại vị trí thả mìn; trong thực tế, khi thả mìn tại môi trường có dòng hải lưu hoặc có xoáy động nước, khoảng cách dạt của mìn khi nổi còn lớn hơn rất nhiều, gây khó khăn cho công tác tìm kiếm, thu hồi mìn sau huấn luyện. Do thời điểm mìn nổi lên mặt nước để thu hồi không phải là tham số cần được yêu cầu tính toán thiết kế nên để hạn chế tốc độ ngang của mìn khi nổi, cần phải hạn chế lực đẩy của luồng nước thoát qua van xả F_J. Trong thiết kế, thay đổi F_J được thực hiện theo dạng "xung" với điều khiển van xả dạng "ON-OFF" với khoảng thời gian T1 (ON) và T2 (OFF) như hình 3.

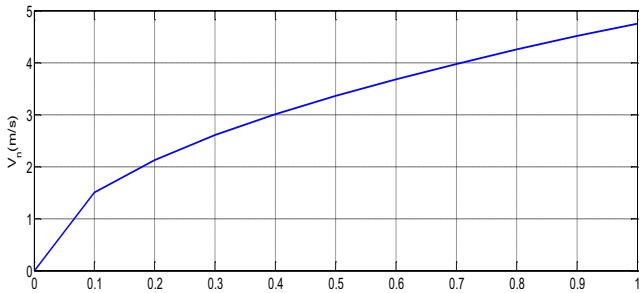


Hình 3. Chu kỳ cấp xung điều khiển van xả nước khi nổi

Lực đẩy của luồng nước thoát qua van xả khi này được tính:

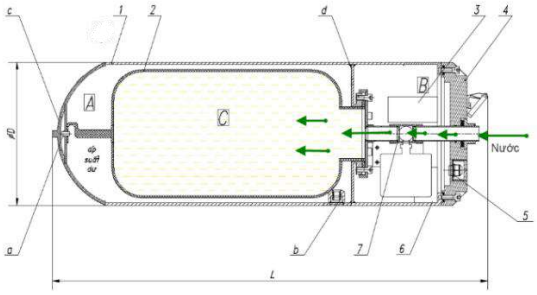
$$F_{J\,ON-OFF} = F_J \frac{T_1}{T_1 + T_2} \tag{20}$$

Biểu đồ tốc độ dịch chuyển ngang của mìn khi thay đổi T1 và T2 (hình 4)



Hình 4. Sự phụ thuộc tốc độ ngang của mìn vào thời gian xả nước

Trong khoảng khoảng thời gian T1 (ON), mìn chịu tác động của hai lực và dạt ngang 1 khoảng R_{ON} (cách tính tương tự các công thức 15-18), trong khoảng khoảng thời gian T2 (OFF), mìn chỉ chịu tác động của lực nổi và không dạt ngang. Phương án thiết kế mìn huấn luyện gồm có thiết kế cấu trúc, bố trí thiết bị bên trong như hình 5.



Hình 5. Phương án thiết kế mìn huấn luyện

Khoang A chứa khí nén, ruột tích áp và các phụ kiện khác; khoang B chứa thiết bị điều khiển, C là ngăn chứa nước.

1. Cụm thân vỏ; 2. Khối dằn; 3. Khối điều khiển; 4. Bộ phận nắp đáy; 5. Giắc cắm liên kết ngoài; 6. Khối nguồn; 7.Van nạp xả nước.

2.4.2. Tính toán khả năng chịu áp khối dằn và độ bền vách ngăn

Giả định thể tích chiếm chỗ của khối dằn cao su khi bơm xả hết nước là không đáng kể, theo phương trình đẳng nhiệt áp dụng đối với khoang A khi khối dằn đầy nước và khi không có nước:

$$P_0V_0 = P_2(V_0 - V_1), P_2 = \frac{P_0V_0}{(V_0 - V_1)} \quad (21)$$

Từ công thức (8), lập bảng so sánh áp suất trong khoang A và trong khối dằn như sau (bảng 1):

Bảng 1. Bảng so sánh áp suất

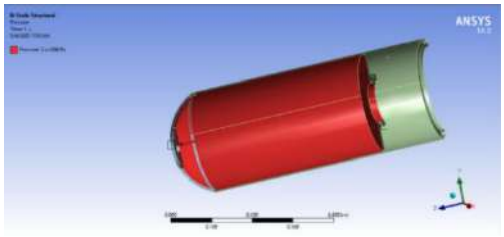
P_0 (atm)	V_0 (lit)	V_1 (lit)	P_2(atm)
1,0	20	13,5	3.08
1,5	20	13,5	4.62
2,0	20	13,5	6.15
2,5	20	13,5	7.69
3,0	20	13,5	9.23
3,5	20	13,5	10.77

Như vậy, trong dải áp suất bơm vào P_0 từ 1÷ 3,5 atm, khối dằn cao su chịu tác dụng đến 10,7 atm. Với đặc điểm này đòi hỏi P_1 tức áp suất của bơm sẽ phải tương ứng lớn hơn giá trị 10,7 atm và kết cấu chịu bền của khối dằn cao su phải chịu được. Áp suất ban đầu P_0 chính là áp suất dư để đẩy nước thoát ra ngoài ở độ sâu tương ứng. Với độ sâu đến 25 m nước thì P_0 phải lớn hơn 2,5 atm. Để đảm bảo hệ số dự trữ áp suất, theo bảng 1 trên đưa ra áp suất P_0 ban đầu đến 3,5 atm. Thân vỏ mìn được chế tạo bằng thép C30, chiều dày 4.5 mm. Ứng suất vòng đối với ống trụ:

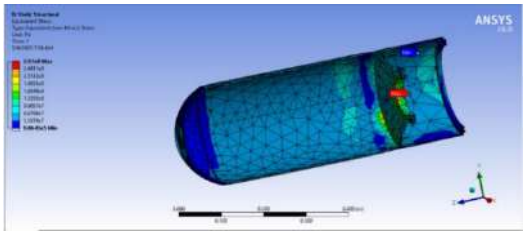
$$\sigma = \frac{P_r}{t} \quad (22)$$

Sử dụng phần mềm Ansys [7] để tính toán kiểm bền kết cấu thân vỏ mìn. Đây là một công cụ tính toán mô phỏng trực quan và chính xác giúp kiểm bền sản phẩm. Điều kiện đầu vào gồm: Kết cấu của thân mìn huấn luyện như bản vẽ thiết kế, các mối liên kết hàn sản phẩm chịu áp lực đến 20 atm, thép C30 TCVN 1766-75, độ dày thân vỏ, vách ngăn đạt giá trị 4,5mm. Áp suất thân vỏ chịu được là 20 atm (tương ứng với 1,85 lần giá trị áp suất tính toán cao nhất của P_2

trong bảng 1, được coi là hệ số an toàn tương ứng trong thiết kế). Kiểm bền kết cấu mìn đến áp suất 2 MPa (≈ 20 atm) cho kết quả như các hình 6÷7 như sau:



Hình 6. Kiểm bền áp lực đến 2MPa

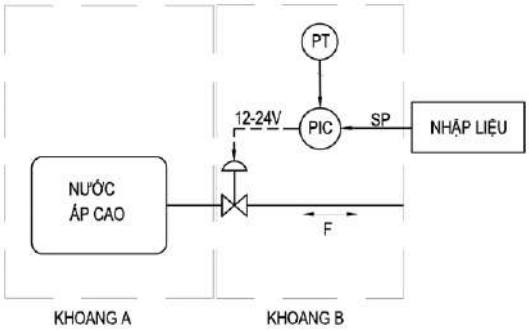


Hình 7. Ứng suất tương đương

Các kết quả tính toán cho thấy giá trị ứng suất σ_{max} = 297 MPa < Ứng suất cho phép σ_{pt} = 335 MPa (ứng với vật liệu thép C30). Tính toán cho thấy ứng suất nhỏ hơn giới hạn chảy của vật liệu với hệ số an toàn > 2.

2.4.3. Thiết lập chương trình điều khiển

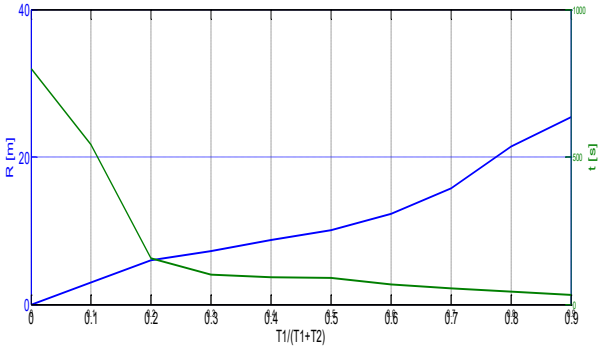
Sơ đồ điều khiển quá trình có thể mô tả theo sơ đồ chức năng như hình 8:



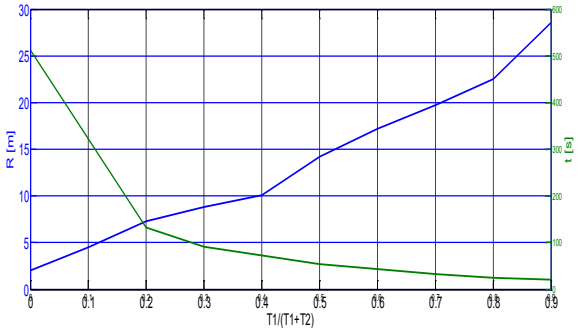
Hình 8. Sơ đồ chức năng bộ điều khiển

PT: Cảm biến tiệm cận; PIC: Bộ điều khiển; SP: Giá trị thiết lập ban đầu; F: dòng chất lỏng

Bộ điều khiển PIC trong khoang B điều khiển van điện xả nước, nhận tín hiệu cảm biến tiệm cận và quản lý toàn bộ chu trình hoạt động của mìn. Tham số thời gian được nhập từ bàn điều khiển ngoài. Khi nước được bơm vào khối dằn, khối lượng mìn tăng tạo sức nổi âm làm mìn chìm. Sau thời gian cài đặt, PIC mở van điện, khí nén trong khoang A đẩy nước ra khỏi khối dằn, làm giảm khối lượng và tạo sức nổi dương giúp mìn nổi lên. Cảm biến tiệm cận xác định thời điểm mìn tách khỏi robot để kích hoạt chu trình nổi. Dữ liệu được nhập vào PIC gồm 3 tham số chính: T1 - khoảng thời gian giữ cấp điện cho van (van được mở - ON); T2 - khoảng thời gian ngắt cấp điện cho van (van bị khoá - OFF); T3 - khoảng thời gian chờ cấp nguồn cho van. Thử nghiệm trên biển với độ sâu 25m, cho thấy kết quả với T=T1+T2 khác nhau, khoảng cách dạt của mìn tỷ lệ với tỷ số ON/OFF như sau (Hình 9, 10):

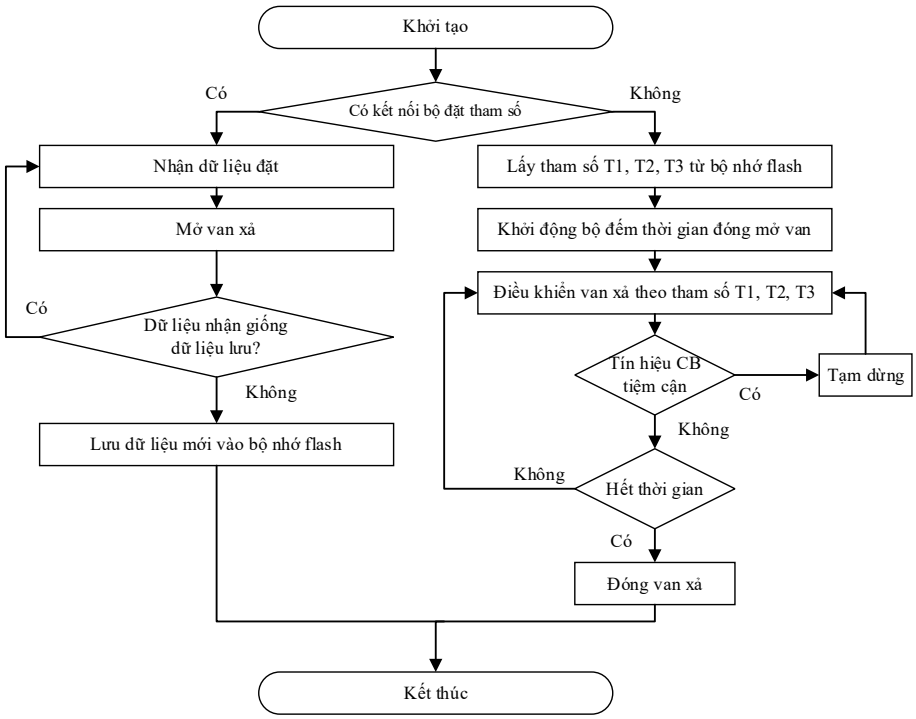


Hình 9: Khoảng cách dạt và thời gian nổi của mìn khi T=20s



Hình 10: Khoảng cách dạt và thời gian nổi của mìn khi T=24s

Từ các biểu đồ thử nghiệm trên, nhận thấy rằng với tỷ lệ thời gian hoạt động van xả nước T1/(T1+T2) khoảng 0.2 sẽ có sự cân bằng giữa khoảng cách dạt của mìn khi nổi và thời gian mìn nổi lên mặt nước. Trong thiết kế, chọn chu kỳ hoạt động van xả T=24s và tỷ lệ thời gian xả/đóng van bằng 0.2; tương ứng khoảng cách dạt của mìn khi nổi cỡ 7m và thời gian nổi của mìn cỡ 130s. Lưu đồ thuật toán bộ điều khiển được thiết kế theo hình 11.



Hình 11. Lưu đồ thuật toán bộ điều khiển

Trường hợp có dữ liệu nhận: PIC nhận dữ liệu từ bộ đặt tham số trong giai đoạn chuẩn bị mìn, điều khiển mở van để nạp nước và so sánh tham số mới với dữ liệu lưu trong bộ nhớ flash. Việc cập nhật bộ nhớ chỉ thực hiện khi có thay đổi nhằm hạn chế số lần ghi, do bộ nhớ flash có số chu kỳ ghi hữu hạn mặc dù vẫn duy trì dữ liệu khi mất nguồn.

Trường hợp không có dữ liệu nhận: Bộ điều khiển thực hiện chu trình đóng/mở van xả theo các tham số T1, T2, T3 lưu trong bộ nhớ thông qua bộ định thời và điều khiển cổng IO. Số chu kỳ đóng/mở được tính toán bảo đảm xả hết nước đồng thời hạn chế độ dạt của mìn do phản lực dòng xả. Trong quá trình này, tín hiệu cảm biến tiệm cận tại vị trí tay gắp robot được giám sát liên tục; nếu robot vẫn đang giữ mìn, chu trình xả sẽ tạm dừng để bảo đảm an toàn.

3. Trang thiết bị áp dụng sau thiết kế hoàn chỉnh mìn huấn luyện

3.1. Khối dằn cao su

Với đường kính mìn thực Φ253 mm, lựa chọn khối dằn cao su Varem (Italy) dung tích 13,5–14 lít, vật liệu EPDM có độ đàn hồi và khả năng chịu môi trường nước mặn tốt. Khối dằn có kích thước 440 × Φ200±2 mm, chiều dày thành 3 mm, khối lượng 1,5 kg, chịu áp suất làm việc đến 16 atm và thể tích giãn nở tối đa 16 lít.

3.2. Van điện xả nước

Lựa chọn van bi điện có kích thước đường kính DN-20, đường kính mở của van 27mm, dải điện áp hoạt động 12÷24V, thân bằng đồng, áp lực hoạt động lớn nhất đến 16 atm, nhiệt độ làm việc –5 ÷ 100 độ C. Van ở trạng thái thường đóng khi không có điện và mở khi được cấp điện. Ưu điểm của van là khi mất điện đột ngột van sẽ tự đóng.

4. Kết quả và thảo luận

4.1. Kết quả thử nổi – chìm

Hình 12. Thử nổi sản phẩm sau tổng lắp *Hình 13. Thử nghiệm lắp đặt trên robot*

Bảng 2. Thông số làm việc của mìn sau chế tạo

Thông số	Giá trị đạt được
Thể tích nước trong khối dằn	≥ 13 lít
Khối lượng mìn khi chìm	45±1 kg
Khối lượng mìn khi nổi	≤32 kg
Thời gian nổi	< 60 s
Thời gian hoạt động	≥ 4 giờ
Độ sâu làm việc lớn nhất	25 m

4.2. Đánh giá hoạt động

Mìn chìm ổn định, không dao động ngang lớn; Sau kích hoạt, quá trình nổi diễn ra liên tục, không giật cục; Không phát hiện nước xâm nhập khoang điều khiển; Vị trí nổi sai lệch ≤ 15 m so với điểm thả; So với phương pháp thu hồi thủ công trước đây, thời gian thu hồi giảm đáng kể và loại bỏ hoàn toàn rủi ro lặn tìm.

5. Kết luận

Bài báo đã trình bày quá trình mô hình hóa, tính toán và thiết kế thiết bị tự nổi đóng vai trò mìn huấn luyện robot Pluto plus Hải quân. Kết quả đạt được: Xây dựng mô hình thủy tĩnh dựa trên nguyên lý Archimedes; Thiết kế cơ cấu tích áp khí – nước điều chỉnh nổi – chìm; Kiểm bền thân vỏ và phân tích ổn định nổi; Thử nghiệm thực tế chứng minh độ tin cậy; Thiết bị đáp ứng đầy đủ yêu cầu kỹ thuật, hoạt động ổn định, bảo đảm yêu cầu chiến kỹ thuật và an toàn huấn luyện. Hướng nghiên cứu tiếp theo có thể mở rộng theo hai hướng: Tích hợp cảm biến áp suất để điều khiển nổi theo độ sâu; Ứng dụng mô hình điều khiển tối ưu nhằm kiểm soát chính xác thời điểm nổi.

Tài liệu tham khảo

[1] W. M. Carey and R. B. Evans, *Ocean Engineering Mechanics: With Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[2] T. I. Fossen, *Handbook of Marine Craft Hydrodynamics and Motion Control*, 2nd ed. Hoboken, NJ, USA: Wiley, 2021.

[3] R. C. Hibbeler, *Mechanics of Materials*, 10th ed. Upper Saddle River, NJ, USA: Pearson, 2017.

[4] D. J. Griffiths, *Introduction to Electrodynamics*, 4th ed. Cambridge, U.K.: Cambridge Univ. Press, 2017.

[5] Teledyne Brown Engineering, "Pluto Gigas – Mine Identification and Disposal," Huntsville, AL, USA, Tech. Rep., 2021.

[6] Tài liệu kỹ thuật robot Pluto plus, Hải quân Việt Nam, 2019.

[7] Ansys Workbench. 2021

Nghiên cứu cải thiện bộ lọc định hướng để loại bỏ nhiễu phản xạ ảnh hưởng đến sự lan truyền sóng trượt trong mô

Nguyễn Vũ Nghĩa[1], Trần Ngọc Quang[1], Nguyễn Trọng Tuyến[1], Lương Quang Hải[1,*]

[1]*Trường Đại học kỹ thuật Lê Quý Đôn*

Email: luonghai@lqdtu.edu.vn

Tóm tắt

Đàn hồi sóng trượt (Shear Wave Elastography – SWE) là một kỹ thuật tạo ảnh đang phát triển mạnh, cho phép hỗ trợ chẩn đoán bệnh thông qua việc ước lượng các tham số cơ học của mô, bao gồm độ đàn hồi và độ nhớt (gọi chung là các tham số cơ học phức – CSM). Phương pháp SWE dựa trên việc đo vận tốc lan truyền của sóng trượt trong mô sinh học. Trong bài báo này, nhóm tác giả đề xuất một cải tiến đối với bộ lọc định hướng nhằm giảm thiểu ảnh hưởng của nhiễu do sóng phản xạ gây ra trong quá trình lan truyền sóng trượt. Trước hết, tín hiệu vận tốc đo được, bao gồm cả sóng tới và sóng phản xạ, được biến đổi sang miền tần số bằng phép biến đổi Fourier rời rạc và xử lý bằng mặt nạ bộ lọc định hướng đã được chuẩn hóa. Tiếp theo, các kỹ thuật xử lý tín hiệu bổ sung được áp dụng nhằm nâng cao chất lượng lọc và loại bỏ các thành phần không mong muốn. Cuối cùng, tín hiệu được biến đổi Fourier rời rạc ngược để thu được vận tốc sóng trượt sau lọc. Kết quả mô phỏng cho thấy phương pháp đề xuất có hiệu quả rõ rệt trong việc cải thiện chất lượng tín hiệu và giảm nhiễu phản xạ.

Từ khóa: *Sóng trượt; bộ lọc định hướng; nhiễu phản xạ; độ đàn hồi mô.*

1. Đặt vấn đề

Kỹ thuật tạo ảnh đàn hồi sóng trượt (Shear Wave Elasticity Imaging – SWEI) đã được ứng dụng rộng rãi trong chẩn đoán hình ảnh y học, đặc biệt trong phát hiện sớm các khối u [1]. Hiện nay, SWEI đã được tích hợp vào nhiều hệ thống siêu âm lâm sàng hiện đại nhằm hỗ trợ chẩn đoán và đánh giá bệnh lý gan [2], [3]. Bên cạnh đó, cộng đồng nghiên cứu vẫn tiếp tục phát triển các phương pháp xử lý và mô hình hóa nhằm cải thiện chất lượng hình ảnh cũng như mở rộng phạm vi ứng dụng của SWEI sang các cơ quan khác như mô vú, tuyến giáp và tuyến tiền liệt.

Trong nghiên cứu [4], Hengzhao và cộng sự đã khai thác đặc tính suy hao của sóng trượt để ước lượng đồng thời độ nhớt và độ đàn hồi của mô. Sóng trượt được kích thích bằng đầu dò siêu âm không hội tụ (tần số 4,1 MHz, 128 phần tử, thời gian kích thích 400 μs), trong khi một đầu dò đặt đối diện được sử dụng để thu nhận tín hiệu lan truyền. Các tham số cơ học được xác định thông qua phép biến đổi Fourier hai chiều (miền không gian–số sóng, K-space). Nhóm tác giả kết hợp phương pháp trích xuất đỉnh phổ và phương pháp toàn chiều rộng tại nửa cực đại (FWHM) để ước lượng các tham số. Mặc dù nghiên cứu đã chỉ ra vai trò quan trọng của độ nhớt và độ đàn hồi trong đánh giá đặc tính mô, các loại nhiễu phát sinh trong quá trình đo và xử lý tín hiệu chưa được phân tích chi tiết. Hơn nữa, phantom sử dụng trong thí nghiệm có đặc tính cơ học đồng nhất, do đó chưa phản ánh đầy đủ các điều kiện không đồng nhất của mô sinh học thực tế.

Thomas Deffieux và cộng sự [5] đã nghiên cứu ảnh hưởng của nhiễu phản xạ trong đàn hồi sóng trượt thoáng qua. Nhóm tác giả đề xuất một bộ lọc định hướng đơn giản nhằm loại bỏ phần lớn thành phần sóng phản xạ, qua đó giảm hiện tượng nhiễu trong bản đồ mô-đun trượt của vật liệu cứng. Kết quả mô phỏng cho thấy giải pháp này cải thiện đáng kể chất lượng tái tạo tham số cơ học.

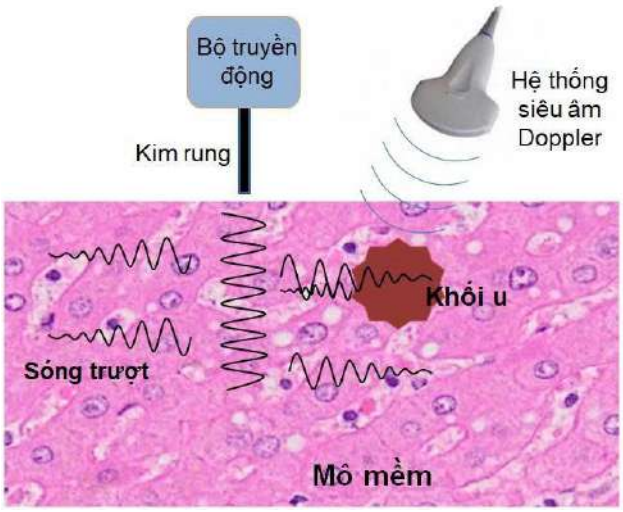
Gần đây, trong nghiên cứu [6], các tác giả đã áp dụng hai phương pháp tính toán độ cứng mô gồm phép nghịch đảo đại số của phương trình vi phân (Algebraic Inversion of Differential Equation – AIDE) và ước lượng tần số cục bộ (Local Frequency Estimation – LFE). Các phương pháp này được đánh giá trong hai trường hợp: có và không sử dụng bộ lọc định hướng. Dữ liệu lan truyền sóng trượt được thu nhận bằng kỹ thuật đàn hồi cộng hưởng từ (Magnetic Resonance Elastography – MRE). Kết quả được tổng hợp và so sánh dưới dạng bảng nhằm đánh giá sự khác biệt giữa các phương pháp và vai trò của bộ lọc định hướng.

Trong bài báo này, chúng tôi tập trung phát triển và cải tiến bộ lọc định hướng nhằm loại bỏ hiệu quả nhiễu phản xạ phát sinh khi sóng trượt lan truyền đến ranh giới giữa các lớp môi trường không đồng nhất. Đồng thời, phương pháp sai phân hữu hạn trong miền thời gian (Finite-Difference Time-Domain – FDTD) được sử dụng để mô phỏng quá trình lan truyền sóng trượt trong môi trường mô sinh học, tạo cơ sở đánh giá định lượng hiệu quả của bộ lọc đề xuất.

2. Phương pháp

A. Nguyên lý tạo và đo sóng trượt

Hình 1 mô tả quy trình phát sinh và ghi nhận sóng trượt trong môi trường mô. Một kim có đường kính 1,5 mm và chiều dài 12 cm, được chế tạo từ thép không gỉ hoặc nhựa cứng, được kích thích bởi bộ truyền động cơ học để dao động điều hòa dọc theo trục *z* trong dải tần từ 100 – 500 Hz. Sự dao động này tạo ra sóng trượt lan truyền trong mô theo mặt phẳng vuông góc với trục kích thích. Vận tốc riêng của sóng trượt sau đó được xác định thông qua hệ thống siêu âm Doppler [7].



Hình 1. Nguyên lý tạo và đo sóng trượt trong mô.

B. Mô hình toán học sự lan truyền sóng trượt trong mô

Khi sóng trượt lan truyền trong mô theo phương xxx của hệ tọa độ Descartes, vectơ vận tốc riêng v_z và thành phần ứng suất tiếp σ_{zx} có mối quan hệ động lực học chặt chẽ, được mô tả bởi hệ phương trình (1) và (2) [8]:

$$\rho \partial_t v_z = \partial_x \sigma_{zx} \quad (1)$$

$$\partial_t \sigma_{zx} = (\mu + \eta \partial_t) \partial_x v_z \quad (2)$$

Trong đó, ρ là mật độ khối của mô; μ và η lần lượt là mô đun đàn hồi trượt và hệ số nhớt; ∂/∂t và ∂/∂x ký hiệu toán tử đạo hàm riêng theo thời gian và không gian. Để giải hệ phương trình vi phân riêng phần này, phương pháp sai phân hữu hạn miền thời gian (FDTD) [9] được áp dụng nhằm rời rạc hóa theo lưới không gian–thời gian, thu được các phương trình cập nhật (3) và (4):

$$v_z{}^{n+1}|_i = v_z{}^{n}|_i + \frac{\Delta t}{\rho \Delta x}\left(\sigma_{zx}{}^{n+\frac{1}{2}}|_{i+\frac{1}{2}} - \sigma_{zx}{}^{n+\frac{1}{2}}\ |_{i-\frac{1}{2}}\right) \tag{3}$$

$$\sigma_{zx}{}^{n+\frac{1}{2}}|_{i+\frac{1}{2}} = \sigma_{zx}{}^{n-\frac{1}{2}}|_{i+\frac{1}{2}} + \frac{\mu \Delta t}{\Delta x}(v_z{}^{n+1}|_{i+1} - v_z{}^{n+1}|_i) + \frac{\eta}{\Delta x}(v_z{}^{n+1}|_{i+1} - v_z{}^{n+1}|_i) - \frac{\eta}{\Delta x}(v_z{}^{n}|_{i+1} - v_z{}^{n}|_i) \tag{4}$$

Ở đây, Δx là bước lưới không gian, Δt là bước thời gian; i và n lần lượt là chỉ số không gian và thời gian. Sóng trượt là một dạng sóng cơ học nên khi lan truyền và gặp rào cản hoặc bề mặt phân cách giữa các môi trường không đồng nhất, một phần năng lượng của sóng sẽ bị phản xạ trở lại môi trường ban đầu. Sóng phản xạ có đặc điểm lan truyền tương tự sóng tới nhưng ngược hướng và ngược pha, dẫn đến sự chồng lấn và giao thoa trong trường sóng tổng, làm ảnh hưởng đến quá trình phân tích và ước lượng tham số cơ học của mô.

C. Quy trình lọc nhiễu phản xạ

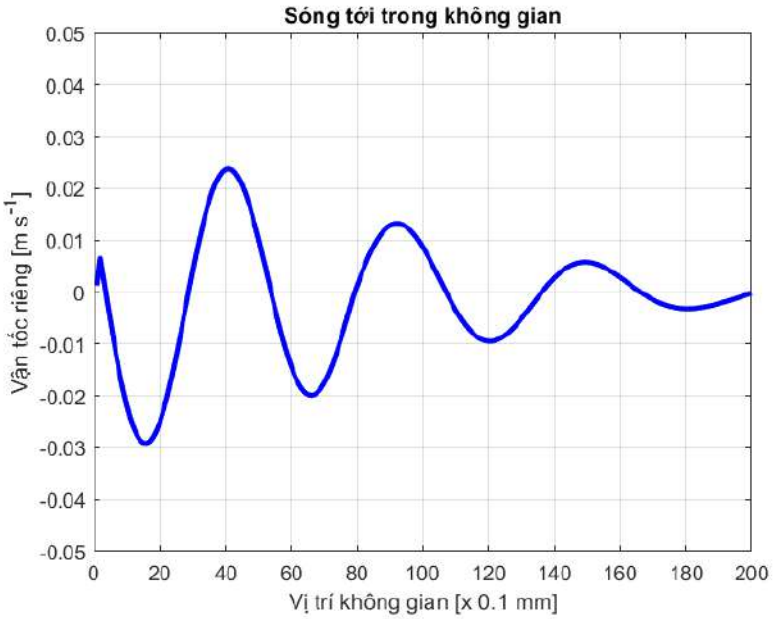
Tín hiệu sóng trượt đo được trong môi trường hữu hạn là tổng chồng tuyến tính của sóng tới và sóng phản xạ, trong đó thành phần phản xạ đóng vai trò như nhiễu làm sai lệch quá trình ước lượng vận tốc. Để loại bỏ nhiễu này, trường sóng trong miền không gian–thời gian $\mathrm{u(x,t)}$ trước hết được biến đổi sang miền tần số–không gian thông qua phép biến đổi Fourier rời rạc (DFT), thu được phổ hai chiều $U(k, \omega)$. Trong miền k–ω, các thành phần sóng lan truyền theo hai hướng ngược nhau được phân bố ở các nửa mặt phẳng đối xứng theo trục số sóng, nhờ sự khác biệt về dấu của k . Dựa trên đặc tính này, một mặt nạ lọc định hướng được xây dựng để giữ lại miền phổ tương ứng với hướng lan truyền mong muốn và triệt tiêu miền phổ của sóng phản xạ thông qua phép nhân phổ. Sau đó, phổ đã lọc được biến đổi ngược về miền không gian–thời gian bằng phép biến đổi Fourier ngược (IDFT), thu được trường sóng trượt đã được loại bỏ đáng kể thành phần phản xạ. Quy trình này cho phép tách hiệu quả sóng truyền thuận khỏi nhiễu phản xạ, từ đó nâng cao độ chính xác của các bước phân tích và ước lượng tham số tiếp theo.

3. Kết quả và thảo luận

Để kiểm chứng phương pháp đề xuất, một kịch bản mô phỏng được thiết lập với các tham số sau: Môi trường 1D có chiều dài 20 mm, chia thành 200 vị trí không gian, bước không gian Δx =0,1 mm; mật độ khối của mô ρ =1000 kg/m^3; mô có độ đàn hồi 1000 Pa và độ nhớt 0,1 Pa.s tại các vị trí không gian từ 1 đến 80, độ đàn hồi 1200 Pa và độ nhớt 0,2 Pa.s tại các vị trí không gian từ 80 đến 200. Tần số rung 200 Hz, biên độ rung 3 mm. Thời gian thu thập tín hiệu 100 ms, chia 10.000 bước thời gian, bước thời gian = 0,01 ms.

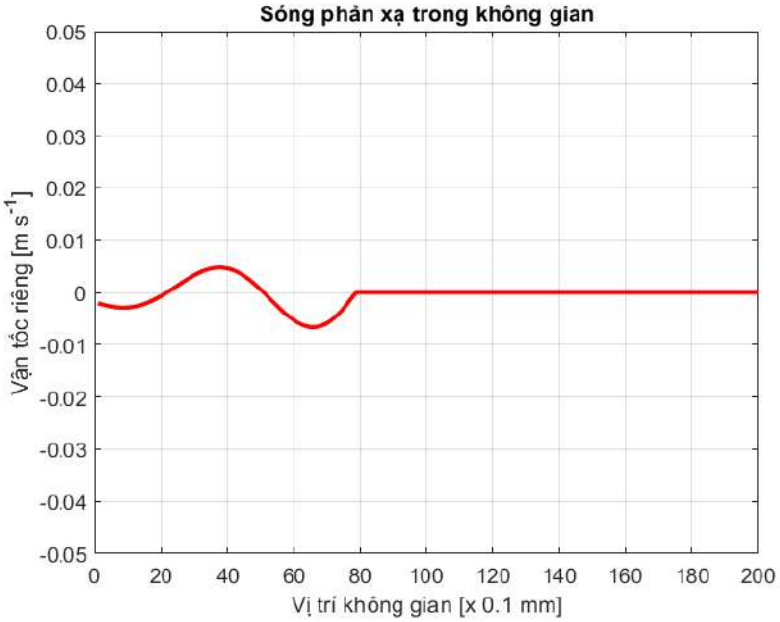
Hình 2 minh họa vận tốc riêng lý tưởng của sóng trượt lan truyền trong mô theo hướng từ trái sang phải. Đồ thị vận tốc riêng của sóng trượt có dạng hình sin suy giảm. Điều này phù

hợp với thực tế, năng lượng của sóng trượt bị hấp thụ bởi môi trường (mô), biên độ tại các vị trí càng xa nguồn kích thích thì càng giảm.



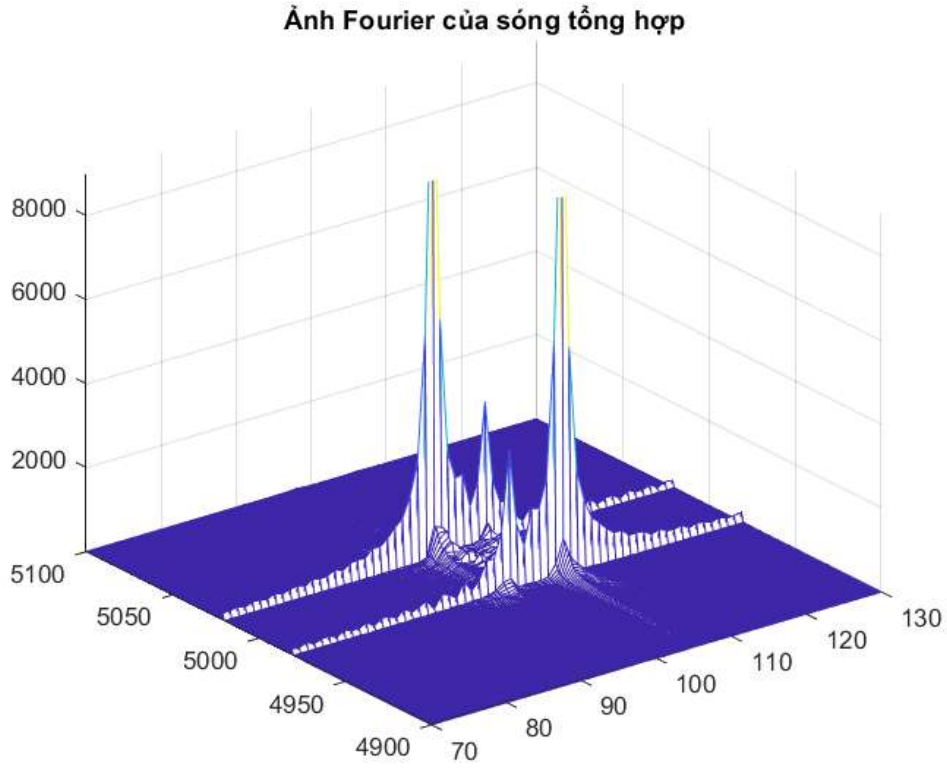
Hình 2. Sóng tới lan truyền trong mô.

Khi sóng trượt lan truyền trong mô, tại biên các lớp mô khác khau sẽ xẩy ra hiện tượng phản xạ. Sóng phản xạ lan truyền cùng tần số nhưng ngược hướng với sóng tới, có pha ngược với pha của sóng tới. Trong trường hợp này, sóng phản xạ đóng vai trò như một loại nhiễu ảnh hưởng đến sóng tới. Theo kịch bản mô phỏng, tại vị trí 80, độ đàn hồi và độ nhớt của mô thay đổi, do đó vị trí này là biên giữa 2 lớp mô khác nhau làm xuất hiện sự phản xạ tại đây. Hình 3 minh họa sóng trượt phản xạ trong không gian.



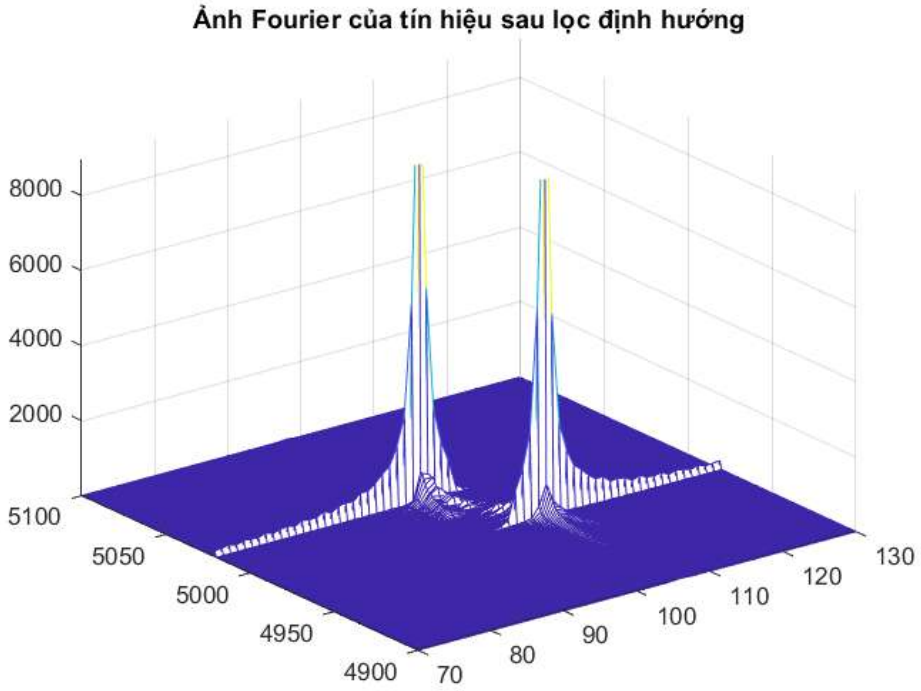
Hình 3. Sóng phản xạ tại vị trí 80 lan truyền trong mô.

Hình 4 trình bày phổ biên độ của biến đổi Fourier hai chiều đối với vận tốc dao động của sóng trượt tổng hợp trong miền không gian–thời gian. Tín hiệu tổng hợp bao gồm hai thành phần: sóng tới và sóng phản xạ. Trên phổ số sóng–tần số (k–f), xuất hiện bốn đỉnh năng lượng đặc trưng. Hai đỉnh có biên độ lớn và đối xứng qua gốc tọa độ tương ứng với thành phần sóng tới, trong khi hai đỉnh có biên độ nhỏ hơn tương ứng với thành phần sóng phản xạ. Sự chênh lệch biên độ giữa các đỉnh phản ánh sự khác biệt về biên độ giữa hai thành phần sóng trong miền thời gian–không gian.



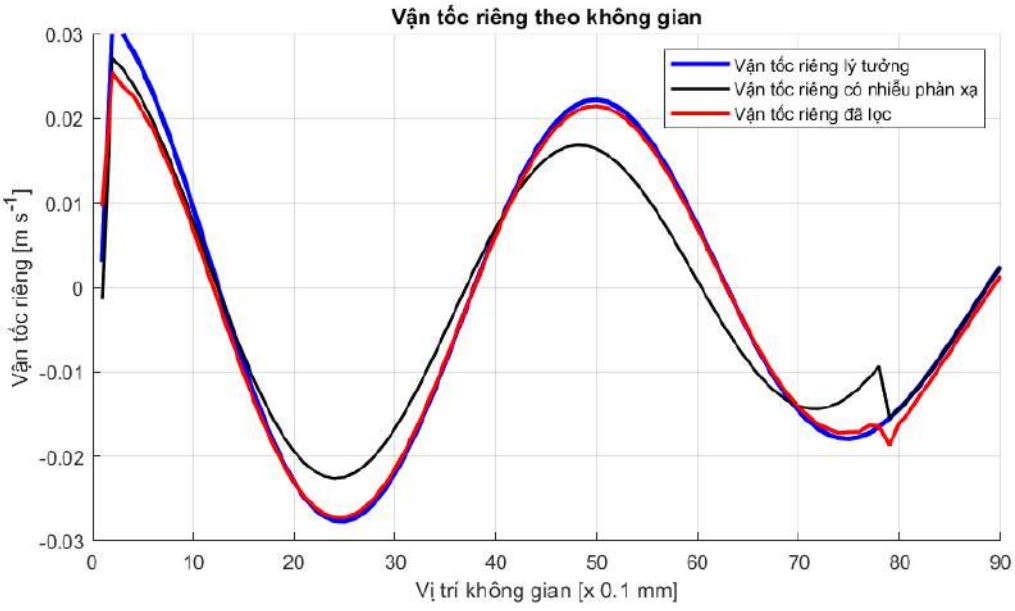
Hình 4. Phổ biên độ của biến đổi Fourier 2D vận tốc riêng của sóng trượt tổng hợp trong miền không gian–thời gian, biểu diễn dưới dạng lưới 3D..

Sau khi áp dụng mặt nạ lọc định hướng trong miền số sóng–tần số, hai đỉnh phổ tương ứng với thành phần sóng phản xạ bị triệt tiêu đáng kể. Do đó, phổ sau lọc chỉ còn lại hai đỉnh năng lượng đối xứng qua gốc tọa độ, tương ứng với thành phần sóng tới, như được minh họa trong Hình 5. Kết quả này cho thấy bộ lọc định hướng đã loại bỏ hiệu quả thành phần phản xạ trong miền Fourier, đồng thời bảo toàn đặc trưng phổ của sóng lan truyền theo hướng mong muốn.



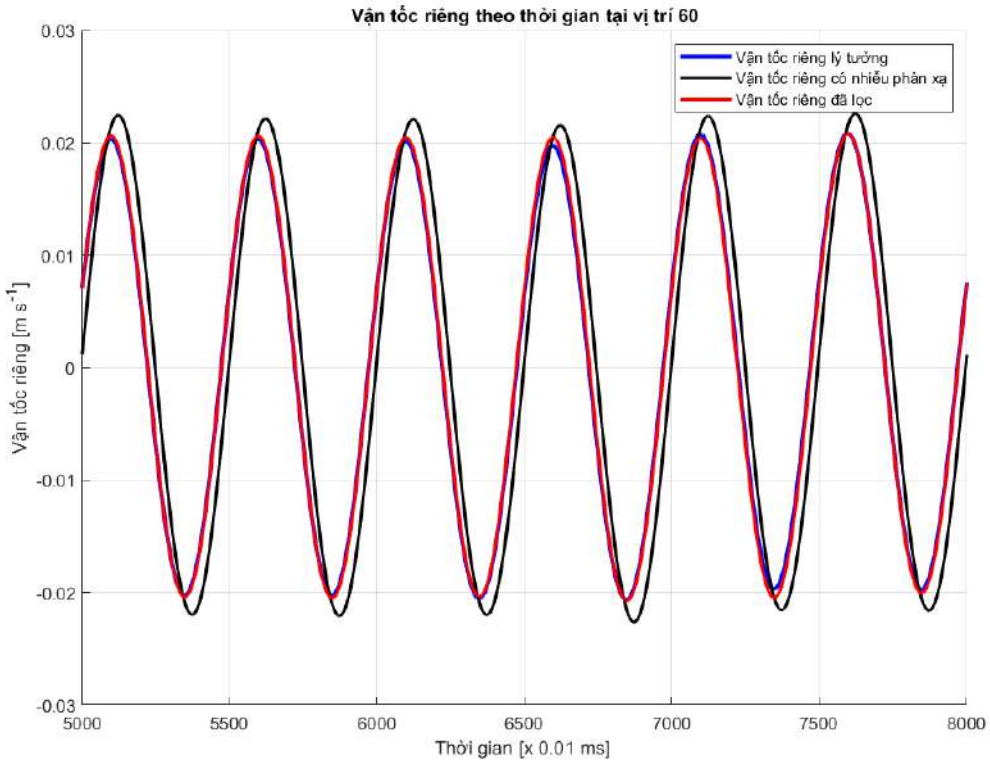
Hình 5. Phổ biên độ của biến đổi Fourier 2D vận tốc sóng trượt trong miền không gian–thời gian sau khi nhân với mặt nạ lọc, biểu diễn dưới dạng lưới 3D..

Hình 6 thể hiện phân bố vận tốc riêng của sóng trượt theo không gian trong ba trường hợp: lý tưởng, có nhiễu phản xạ và sau khi áp dụng bộ lọc định hướng. Kết quả cho thấy nhiễu phản xạ gây suy giảm biên độ và sai lệch pha, đặc biệt tại các vùng gần ranh giới không đồng nhất. Sau khi lọc, biên độ và pha của tín hiệu được phục hồi đáng kể và phù hợp tốt với trường hợp lý tưởng, chứng tỏ hiệu quả của bộ lọc trong việc suy giảm thành phần phản xạ đồng thời bảo toàn đặc trưng lan truyền của sóng.



Hình 6. Vận tốc riêng của sóng trượt trong không gian.

Hình 7 trình bày biến thiên vận tốc riêng của sóng trượt theo thời gian tại vị trí không gian x=60 trong ba trường hợp: lý tưởng, có nhiễu phản xạ và sau khi áp dụng bộ lọc định hướng.Kết quả cho thấy nhiễu phản xạ làm tăng biên độ cục bộ và gây sai lệch pha so với tín hiệu lý tưởng, thể hiện qua sự chênh lệch rõ rệt tại các đỉnh và đáy sóng. Sau khi áp dụng bộ lọc, dạng sóng được phục hồi gần với trường hợp lý tưởng cả về biên độ lẫn pha, đồng thời duy trì tính tuần hoàn ổn định theo thời gian. Điều này khẳng định hiệu quả của phương pháp đề xuất trong việc suy giảm thành phần phản xạ mà không làm biến dạng đặc trưng động học của sóng trượt.



Hình 7. Vận tốc riêng của sóng trượt theo thời gian.

Để định lượng hiệu quả của phương pháp đề xuất, tiêu chuẩn lỗi chuẩn hóa được sử dụng theo định nghĩa trong phương trình (5). Tiêu chuẩn này được áp dụng nhằm so sánh lỗi giữa vận tốc riêng đã lọc và vận tốc riêng lý tưởng với lỗi giữa vận tốc riêng có nhiễu và vận tốc riêng lý tưởng trên toàn bộ miền không gian và thời gian.

$$\epsilon_v = \frac{1}{X \times N} \sum_x^X \sum_t^N \frac{|v_{ideal(x,t)} - v_{(x,t)}|}{v_{ideal(x,t)}} \quad (5)$$

Trong đó X là số điểm trong không gian, N là số mẫu theo thời gian, v_{ideal}là vận tốc riêng lý tưởng, v là vận tốc riêng đang xem xét (trường hợp chưa lọc nhiễu và sau khi lọc nhiễu). Các kết quả tương ứng với các mức biên độ nhiễu (sóng phản xạ) khác nhau được trình bày trong Bảng 1. Bảng này thể hiện lỗi chuẩn hóa (%) của vận tốc riêng trong hai trường hợp: không sử dụng và có sử dụng bộ lọc định hướng đề xuất, tại các mức biên độ nhiễu khác nhau. Khi tăng biên độ nhiễu, lỗi chuẩn hoá của vận tốc không được lọc định hướng tăng đáng kể, từ 1,928% lên 3,568%. Ngược lại, lỗi chuẩn hoá của vận tốc sau lọc định hướng chỉ ghi nhận mức tăng rất

nhỏ, dao động từ 1,181% đến 1,248%. Kết quả này cho thấy bộ lọc đề xuất có khả năng giảm thiểu đáng kể ảnh hưởng của nhiễu, duy trì mức lỗi thấp và ổn định hơn so với trường hợp không lọc. Đồng thời, dữ liệu cũng xác nhận xu hướng tương quan nghịch giữa biên độ nhiễu và lỗi chuẩn hóa, nhấn mạnh vai trò quan trọng của kỹ thuật lọc nhiễu trong việc nâng cao độ chính xác và độ tin cậy của phép đo vận tốc.

Bảng 1: Sai số chuẩn hóa (%) $\in_v$

	Khi chưa lọc định hướng	Sau khi lọc định hướng
1	1.928	1.181
2	2.113	1.189
3	2.502	1.203
4	3.015	1.225
5	3.568	1.248

4. Kết luận

Nghiên cứu này đã đề xuất và cải tiến một bộ lọc định hướng nhằm giảm thiểu nhiễu phản xạ phát sinh khi sóng trượt lan truyền đến ranh giới giữa các lớp môi trường không đồng nhất. Nhiễu phản xạ được xác định là một trong những nguyên nhân chính làm suy giảm độ chính xác của phép đo vận tốc sóng trượt và các tham số cơ học suy ra từ đó.

Để đánh giá hiệu quả của phương pháp, quá trình lan truyền của sóng tới và sóng phản xạ đã được mô phỏng bằng phương pháp sai phân hữu hạn trong miền thời gian (FDTD). Mô hình số cho thấy độ phù hợp cao với nghiệm tham chiếu, với sai số chuẩn hóa xấp xỉ 1,2%, khẳng định tính chính xác và độ tin cậy của dữ liệu mô phỏng.

Kết quả phân tích cho thấy bộ lọc đề xuất có khả năng giảm đáng kể ảnh hưởng của nhiễu phản xạ, duy trì sai số chuẩn hoá vận tốc ở mức thấp và ổn định với các mức biên độ nhiễu khác nhau. Những kết quả này chứng minh tiềm năng ứng dụng của phương pháp trong nâng cao độ chính xác và độ tin cậy của các hệ thống đo và tạo ảnh đàn hồi mô dựa trên sóng trượt.

Tài liệu tham khảo

1. J. Bercoff, A. Criton, C. Bacrie, J. Souquet, M. Tanter, J. Gennisson, T. Deffieux, M. Fink, V. Juhan, A. Colavolpe et al., “ShearWave Elastography A new real time imaging mode for assessing quantitatively soft tissue viscoelasticity,” in *Ultrasonics Symposium*, 2008. IEEE, 2008, pp. 321–324.
2. G. Ferraioli, P. Parekh, A. B. Levitov, and C. Filice, “Shear wave elastography for evaluation of liver fibrosis,” *Journal of Ultrasound in Medicine*, vol. 33, no. 2, pp. 197–203, 2014.
3. Wang, Jing-Houng. "Application of ultrasound liver elastography to the diagnosis and monitoring of liver disease." *Journal of Medical Ultrasound* 27.1 (2019): 1-2.
4. Zhao, Heng & Qiang, Bo & Amador Carrascal, Carolina & Song, Pengfei & Urban, Matthew & Kinnick, Randall & Greenleaf, James & Chen, Shigao. (2012). “Measure Elasticity and Viscosity Using the Out-of-plane Shear Wave”. *IEEE International Ultrasonics Symposium*, IUS. 10.1109/ULTSYM.2012.0053.

5. T. Deffieux, J. -l. Gennisson, J. Bercoff and M. Tanter, "On the effects of reflected waves in transient shear wave elastography," in *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control,* vol. 58, no. 10, pp. 2032-2035, doi: 10.1109/TUFFC.2011.2052, 2011.
6. Maharjan, Surendra, et al. "Directional filter, local frequency estimate and algebraic inversion of differential equation of psoas major magnetic resonance elastography." *Open Journal of Medical Imaging* 10.1 (2020): 1-16.
7. M. Orescanin and M. F. Insana, "Model-based complex shear modulus reconstruction: A Bayesian approach," in IEEE Ultrasonics Symposium (IUS). IEEE, 2010, pp. 61–64.
8. M. Orescanin, Y. Wang and M. F. Insana, "3-D FDTD simulation of shear waves for evaluation of complex modulus imaging," in IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 58, no. 2, pp. 389-398, February 2011, doi: 10.1109/TUFFC.2011.1816, 2011.
9. Schneider, J. B. "Introduction to the Finite-Difference Time-Domain Method: FDTD in 1D." *Understanding the Finite-Difference Time Domain Method, Seattle* (2016): 33-35.

Improving Directional Filtering to Suppress Reflection Noise in Shear Wave Propagation within Tissue

Abstract: *Shear Wave Elastography (SWE) is a rapidly developing imaging technique that enables disease diagnosis by estimating the mechanical properties of biological tissues, including elasticity and viscosity, collectively referred to as complex shear moduli (CSM). SWE is based on measuring the propagation velocity of shear waves within tissue. In this paper, an improved directional filtering approach is proposed to mitigate the effects of reflected-wave noise on shear wave propagation. First, the measured velocity signal, which consists of both incident and reflected waves, is transformed into the frequency domain using the discrete Fourier transform and processed with a normalized directional filter mask. Subsequently, additional signal processing techniques are applied to further enhance filtering performance and suppress unwanted components. Finally, the processed signal is transformed back into the spatial–temporal domain via the inverse discrete Fourier transform to obtain the filtered shear wave velocity. Simulation results demonstrate that the proposed method effectively reduces reflection artifacts and significantly improves signal quality.*

Keywords: Shear wave; directional filter; reflected noise; elastography.

Beyond Attention-Centric Models: State Space Models and Mamba for Perception in Autonomous Unmanned Surface Vehicles

Viet-Tiep Nguyen, Xuan-Tung Truong, Vu Duc Truong, Vu Xuan Duc, Tran Cong Tan*

Institute of Missile and Control Engineering;

Email: viettiep1986@gmail.com; Tel: +84989810151

Abstract

Perception is critical for autonomous unmanned surface vehicles (USVs) operating in dynamic maritime environments. While attention-based models such as Transformers achieve strong accuracy, their quadratic complexity and high resource demands limit real-time deployment on embedded platforms. State space models (SSMs) and the Mamba architecture offer an efficient alternative by modeling sequential dependencies through structured state transitions with linear complexity. This paper reviews the transition from attention-centric paradigms to SSM- and Mamba-based approaches for USV perception. Drawing on insights from related domains, we analyze architectural principles, efficiency–performance trade-offs, and practical deployment considerations. We further discuss evaluation strategies and outline open challenges, including hybrid modeling and real-time embedded implementation.

Keywords: *USVs; Attention; Mamba Architecture; Maritime Vision.*

1. Introduction

Unmanned Surface Vehicles (USVs) are increasingly deployed in applications such as environmental monitoring, search and rescue, port inspection, surveillance, and autonomous navigation [1, 2, 3]. Reliable perception is fundamental to these systems, enabling tasks such as object detection, semantic segmentation, obstacle recognition, and target tracking [4, 5].

However, maritime environments pose unique challenges compared to terrestrial settings. Strong water reflections, dynamic waves, adverse weather, and low-contrast conditions frequently degrade visual inputs, while real-time operation must be maintained under strict computational and energy constraints. These factors make USV perception particularly demanding, requiring models that are both robust and efficient [6, 7].

Attention-centric architectures, especially Vision Transformers (ViTs), have achieved strong performance in computer vision by modeling long-range dependencies [8, 9]. Nevertheless, their applicability to USVs is limited. Self-attention incurs quadratic complexity with respect to input size, leading to high memory usage and latency [10]. Moreover, maintaining stable temporal reasoning in dynamic maritime environments remains challenging for purely attentionbased pipelines [11, 12].

Recently, State Space Models (SSMs) and the Mamba architecture have emerged as promising alternatives [12, 13]. By modeling sequential dependencies through structured state transitions, SSMs achieve linear complexity and predictable memory usage [14, 15]. Mamba further improves efficiency through selective scanning, enabling eective long-range modeling while

* Email: tungtx@lqdtu.edu.vn

preserving parallelism [16]. Results from related domains such as aerial vision and autonomous driving suggest competitive performance with improved efficiency [17, 18].

Despite this potential, research on applying SSMs and Mamba to USV perception remains limited and fragmented [2, 11]. Existing works are dispersed across domains and lack standardized evaluation, making it di-cult to assess their eectiveness in maritime settings [19].

Motivated by this gap, this paper presents a focused survey on the transition from attention-based models to SSM- and Mamba-based approaches for USV perception. Our contributions are threefold: (i) we analyze the limitations of attention-centric models under USV-specic constraints, (ii) we summarize key design principles of SSMs and Mamba in vision, and (iii) we outline a structured research roadmap for future maritime perception systems.

2. USV Perception and Limitations of Attention-Centric Models

2.1. Core Perception Tasks and Maritime Challenges

Perception in autonomous unmanned surface vehicles (USVs) refers to the system's ability to sense, interpret, and understand its surrounding environment to support safe navigation and mission execution [5, 20]. Compared with ground vehicles, USVs operate in highly dynamic and visually complex maritime environments, where both scene structure and environmental conditions can change rapidly [19]. Consequently, perception is not merely a supporting component but a core module that directly influences operational reliability and decision making [3, 20, 21].

From a functional perspective, USV perception consists of several tightly coupled tasks. Object detection identifies vessels, buoys, and floating debris that may pose collision risks [3, 20]. Semantic and instance segmentation provide pixel-level understanding of the scene, enabling distinction between navigable water and non-navigable regions such as shorelines or infrastructure [11, 20]. Object tracking ensures temporal consistency when monitoring moving targets across frames [5], while higher-level scene understanding integrates multi-object information to support decision-making processes such as path planning and collision avoidance. These tasks are interdependent and must operate under strict real-time constraints [5, 19].

Despite these functional similarities with other autonomous systems, USV perception faces distinctive maritime challenges at both visual and system levels. Reflections and glare introduce strong visual ambiguity, distorting object boundaries and color cues [11], while dynamic wave patterns generate continuously changing textures that may mimic or obscure targets [19]. Adverse weather conditions such as fog, rain, and haze further degrade visibility and contrast [3], and motion blur caused by platform instability complicates temporal consistency in video streams [2, 20]. In addition, computational and operational constraints intensify the problem: embedded processors limit memory and energy budgets [7], continuous multi-sensor streams require efficient temporal processing [22], and domain variability across different water bodies introduces distribution shifts that challenge generalization [2, 13]. Collectively, these factors make USV perception particularly demanding, requiring models that balance accuracy, robustness, temporal stability, and computational efficiency.

2.2. Attention-Centric Vision Models and Practical Limitations

Attention-centric deep learning models, particularly ViTs and hybrid CNN-Transformer architectures, have become central to modern computer vision [13, 14]. By leveraging self-attention to model relationships among image tokens, these models effectively capture long-range spatial dependencies and global context, leading to strong performance in tasks such as object detection, segmentation, and video understanding [3, 11]. As a result, they are increasingly adopted in autonomous perception systems.

However, several practical limitations arise when applying attention-based models to USV perception. The most signifficant issue is computational complexity. Standard self-attention scales quadratically with the number of tokens, resulting in rapidly increasing computational cost and memory usage as image resolution or sequence length grows [14, 23]. This becomes particularly problematic in maritime scenarios, where high-resolution inputs and multi-frame processing are often required. In addition, inference latency can hinder real-time responsiveness, as Transformer-based architectures rely on intensive matrix operations that may introduce delays in dynamic environments [3]. These factors collectively limit their suitability for embedded and energy-constrained USV platforms [7, 24].

Attention-centric models also encounter challenges in temporal modeling and deployment. Extending self-attention to long sequences signifficantly increases computational overhead or requires approximations such as windowed or sparse attention, which may degrade temporal coherence [12, 23]. This is particularly critical in maritime environments, where unstable motion and continuously changing backgrounds demand consistent temporal reasoning.

Furthermore, large attention-based models often require substantial memory and training data, reducing scalability and adaptability under varying maritime conditions. Domain shifts across different water environments can further affect generalization performance [7, 25]. These limitations highlight the need for alternative architectures that provide efficient and stable sequential modeling for real-time USV perception.

3. State Space Models and Mamba: Foundations and Design Principles

3.1. Foundations of State Space Models and Mamba

State Space Models (SSMs) originate from classical control theory and signal processing, where dynamical systems are described through latent state transitions driven by external inputs. In sequential modeling, an SSM maintains a hidden state that evolves over time via structured transition operators [13, 14]. Instead of explicitly modeling pairwise interactions among sequence elements, information is accumulated and propagated through this evolving state, providing a principled mechanism for capturing long-range dependencies. This represents a key shift from attention-based modeling, where dependencies are expressed through direct token-to-token relationships [14, 15].

Modern deep learning extends classical SSMs to high-dimensional visual and multimodal data. Learnable input and output projections replace fixed transition-only formulations, enabling mapping between feature spaces and latent states [22, 23]. This allows SSM-based architectures to operate within parallelized training frameworks while preserving structured temporal propagation. As a result, they can be effectively applied to large-scale perception tasks such as video understanding and multi-frame analysis. Importantly, the computation remains structured and predictable, without relying on global pairwise interactions [14, 23].

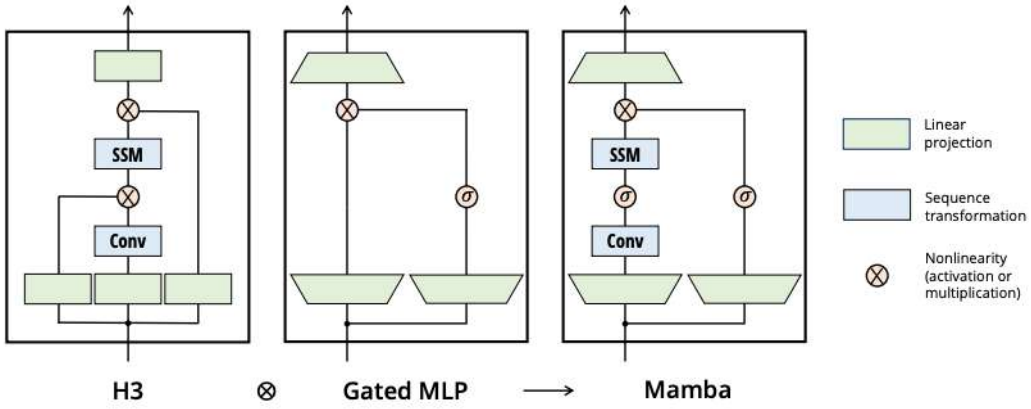


Figure 1. Conceptual structure of the Mamba block with selective state modulation

The Mamba architecture can be viewed as a practical refinement of this framework. As illustrated in Fig. 1, it integrates a state space transformation branch with a gated modulation branch. The SSM branch performs sequence transformation, often combined with lightweight local mixing, while a gatedMLP branch generates input-dependent modulation signals. Their outputs are fused via element-wise multiplication (⊗), enabling selective regulation of state-transformed features. Linear projections handle input-output mappings, and nonlinearities are applied within the gating and mixing components [13, 14]. This design supports adaptive state updates while preserving the structured propagation principle of SSMs.

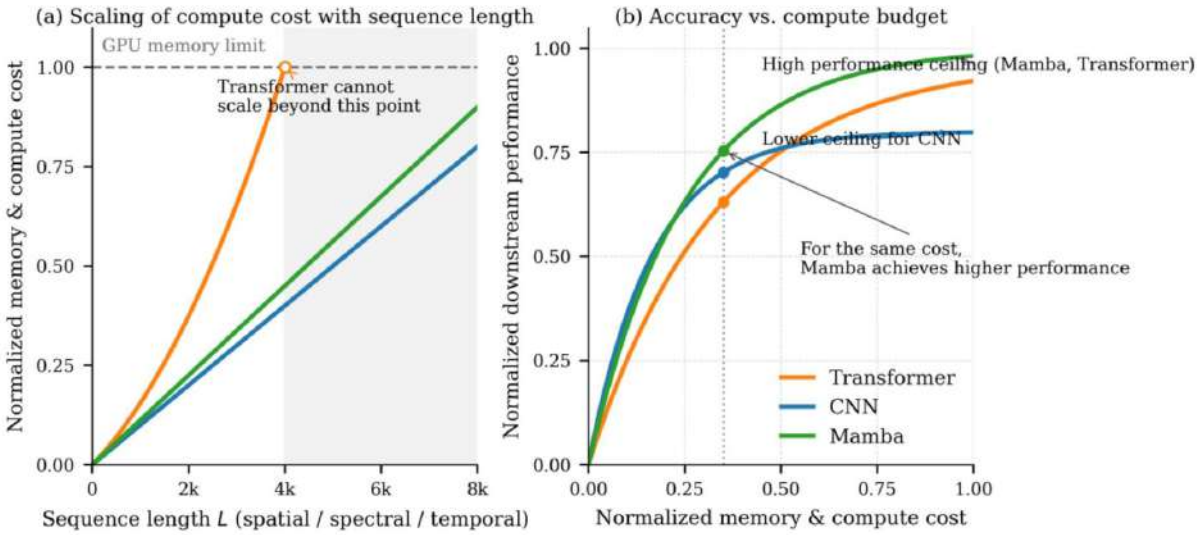


Figure 2. Conceptual illustration of scaling behavior and efficiency-performance trade-offs among convolutional, attention-based, and state space-based architectures

The computational implications of this design are evident when compared with attention-based architectures. Self-attention requires pairwise token interactions, leading to quadratic growth in memory and computation with sequence length, whereas structured state updates scale linearly under fixed state dimensions [24]. As illustrated in Fig. 2(a) [23], attention-based models exhibit rapid resource growth, while state space approaches maintain stable scaling behavior. Fig. 2(b) further highlights the efficiency-performance trade-off: convolutional models are efficient but limited in long-range modeling, Transformers are powerful but computationally expensive, and Mamba aims

to balance these properties. By combining structured propagation with selective gating, it approaches attention-level modeling capacity with improved computational efficiency [12, 23].

In summary, SSM and Mamba are characterized by three key principles: structured state evolution, selective modulation of information flow, and linear scaling with sequence length. These properties distinguish them from attention-centric models and provide a foundation for efficient long-sequence perception systems.

3.2. Architectural Design Patterns in Adjacent Vision Domains

Building on the theoretical foundations, this section examines how Mamba and structured state space modeling have been adapted in vision domains with technical requirements similar to USV perception, such as aerial vision, autonomous driving, remote sensing, and medical imaging. These domains involve high-resolution spatial data, long temporal dependencies, and realtime constraints, making them relevant references for maritime perception systems.

Although originally designed for one-dimensional sequences, Mamba has been extended to high-dimensional visual data through structured adaptations [13]. Since images are represented as two- or three-dimensional grids, spatial serialization is required to convert them into sequence-compatible formats. Common approaches include patch-based tokenization combined with structured scanning strategies that preserve spatial locality. Bidirectional and multi-directional scanning (e.g., SS2D) enable balanced information propagation across dierent directions, while space-lling curves such as Hilbert curves help maintain neighborhood continuity during linearization [17, 18]. These techniques allow Mamba to process spatial data while retaining essential geometric structure.

Another key component is discretization, which bridges continuous-time state space formulations and practical neural network implementations [13]. Classical SSMs rely on continuous dynamics governed by transition matrices, which must be converted into discrete operators for iterative computation and gradient-based training. Methods such as Zero-Order Hold (ZOH) are commonly used to ensure numerical stability and preserve structured state propagation within standard learning pipelines [14, 26]. In addition, Mamba introduces selective state updates, where input-dependent gating dynamically modulates state evolution, allowing the model to emphasize relevant information while suppressing noise [11, 14]. This mechanism is particularly benecial in environments with occlusion, clutter, and variability, as it stabilizes long-range modeling without incurring the quadratic cost of attention [24].

These adaptations have enabled Mamba-based models to support a wide range of vision tasks. In image classication, models such as Vision Mamba and VMamba achieve competitive accuracy with improved efficiency [27, 17]. In segmentation, including medical and complex outdoor scenarios, Mambabased encoderdecoder architectures eectively capture global context [28]. In 3D perception and autonomous driving, Mamba contributes to BEV representation learning, cooperative perception, and occupancy prediction [12]. Remote sensing and radar-based applications further demonstrate its scalability for high-resolution data [23, 29].

Across these domains, several recurring architectural patterns emerge. Hybrid CNNMamba designs combine convolutional layers for local feature extraction with Mamba blocks for long-range modeling [23]. TransformerMamba hybrids allocate attention at coarse scales and Mamba at ner resolutions to balance efficiency and expressiveness [30]. Encoderdecoder frameworks, such as U-Net variants [31], integrate Mamba blocks within hierarchical pipelines to enhance contextual propagation [14]. Collectively, these patterns indicate that Mamba functions as a exible structured modeling paradigm rather than merely an efficiency -oriented alternative, providing a strong foundation for adapting state space approaches to USV perception.

4. Adaptation and Evaluation Perspectives for USV Perception

Building on cross-domain design patterns, we specialize these principles for USV systems, where perception must operate under maritime disturbances, embedded constraints, and safety-critical real-time requirements. As a result, SSM and Mamba integration should be guided not only by computational efficiency but also by robustness and deployment feasibility in maritime environments.

4.1. Practical Adaptation Strategies for USV Systems

USV perception operates on continuous sensor streams under dynamic water motion, reections, and unstable viewpoints. Rather than directly replacing existing architectures, Mamba should be integrated in a structured manner that supports sequential stability, multi-sensor processing, and resource-aware deployment.

Temporal stabilization and sequential modeling. Maritime perception is inherently sequential, and long-horizon consistency is essential for tracking and obstacle monitoring. State space modules can act as temporal stabilizers after spatial feature extraction, accumulating historical information without the quadratic cost of global attention [12, 15, 32]. This is particularly important in water environments where wave motion and illumination changes introduce transient prediction noise. Beyond frame-level accuracy, temporal modeling should reduce icker, condence oscillation, and identity fragmentation.

Disturbance-aware modeling. Reections, glare, and dynamic textures introduce unstable visual responses in maritime scenes [20, 33]. Mamba's selective state update mechanism enables adaptive ltering of such disturbances. Through input-dependent gating [14], the model can suppress transient artifacts while preserving persistent structures such as vessel contours or shorelines. In this context, robustness to environmental variability becomes more critical than purely spatial discrimination.

Multi-modal integration and deployment constraints. USVs typically rely on heterogeneous sensors, including cameras, radar, and inertial systems [34]. Structured state modeling enables sequential fusion across modalities, improving resilience when individual sensors degrade. At the same time, deployment on embedded hardware requires careful architectural placement of state space modules to balance performance and efficiency [7, 23]. Selective integration within temporal branches or intermediate layers helps maintain stable latency and predictable memory usage.

Overall, adaptation of SSM and Mamba for USV perception should emphasize stable temporal modeling, disturbance robustness, and resourceaware design tailored to maritime conditions.

4.2. Resource-Aware and Temporal Evaluation Framework

Evaluating USV perception systems requires criteria that go beyond conventional per-frame accuracy. In maritime environments, long-duration operation, environmental variability, and safety-critical decision making make temporal stability and robustness equally important as recognition performance.

Recognition performance under maritime constraints. Standard metrics such as mAP for detection [35], IoU for segmentation [36], and CLEAR MOT metrics for tracking [37] remain essential, with IDF1 measuring identity preservation [38]. However, for USVs, evaluation should explicitly consider object scale and distance, as small or distant obstacles have a disproportionate impact on navigation safety.

Efficiency and real-time feasibility. Given embedded deployment, evaluation must report latency (ms/frame), throughput (FPS), memory usage, and computational cost. Efficiency benchmarks are particularly important for lightweight architectures [10]. In maritime settings, stable latency and predictable memory behavior under continuous streaming are more critical than peak throughput [7].

Temporal stability as a primary criterion. Unlike conventional vision tasks, temporal consistency is a primary requirement for USV perception. Sequence-level metrics such as identity switches, condence variance, and frame-to-frame consistency should be emphasized. Benchmarks such as MOTChallenge [39] and DAVIS [40] provide useful references, but evaluation must also reect maritime-specic disturbances such as wave motion and viewpoint oscillation [41]. In practice, temporal stability often outweighs marginal gains in frame-level accuracy.

Robustness under maritime disturbances. Evaluation should include challenging environmental conditions such as glare, fog, rain, and dynamic water surfaces [20, 33]. Cross-location validation across rivers, coastal areas, and open seas is essential to assess generalization [34]. When dedicated datasets are unavailable, controlled perturbation experiments can provide useful approximations.

Balanced performance proling. USV perception systems should be evaluated using a multi-dimensional prole that integrates accuracy, efficiency , temporal stability, and robustness. In comparisons between attention-based and state space-based models, improvements in latency predictability, memory scaling, and temporal consistency should be considered alongside accuracy [18, 42]. This balanced evaluation better reects real-world maritime deployment requirements.

In summary, advancing SSM and Mamba for USV perception requires both domain-specic adaptation and evaluation frameworks aligned with maritime operational conditions. A resource-aware, temporally grounded, and disturbance-sensitive assessment is essential to validate practical performance gains.

5. Open Challenges, Future Directions, and Conclusion

5.1. Open Challenges and Future Directions

Although SSMs and the Mamba architecture demonstrate strong potential for efficient sequential perception, recent studies provide both theoretical and empirical evidence highlighting their advantages and limitations compared to Transformer-based models.

Efficiency advantages in long-sequence modeling: SSM-based architectures achieve linear $O(N)$ complexity, in contrast to the quadratic $O(N2)$ scaling of Transformers, enabling

more efficient long-sequence processing [43, 44]. As shown in Table 1, Mamba exhibits substantial gains in memory usage and inference speed as sequence length increases. At $N = 4096$, it is approximately 12.46× more memory-efficient and 10.67× faster, while Transformers fail at $N = 8192$ due to memory limits [43]. These results highlight the suitability of SSMs for long-horizon and resource-constrained perception tasks.

Table 1. Efficiency Ratios and Clinical Interpretation

N	Mem. Ratio (T/M)	Time Ratio (T/M)	Clinical Interpretation
512	1.35×	2.01×	Minimal advantage
1024	2.34×	3.02×	Moderate advantage
2048	5.21×	5.47×	Significant advantage
4096	12.46×	10.67×	Substantial advantage
8192	∞ (OOM)	∞ (OOM)	Transformer cannot process

Trade-os in representational capacity: Despite these efficiency gains, SSMs introduce limitations in representational exibility. Empirical studies show that Mamba may struggle with tasks requiring strong associative reasoning or ne-grained token interactions due to its xed-size latent state [43]. Survey works further indicate that Transformers still maintain advantages in certain vision and multimodal tasks [44]. This suggests that performance differences are task-dependent and that efficiency does not universally translate into superior accuracy.

Implications for architecture design: Recent work points toward hybrid architectures as a promising direction. Combining SSMs with a small number of attention layers can mitigate representational limitations while preserving most efficiency benets [43]. Broader surveys also emphasize balancing efficiency and expressiveness, positioning SSMs as key components in nextgeneration sequence models [45]. For USV systems, this implies that SSMbased models are well-suited for onboard deployment, but require careful design in safety-critical scenarios.

Data and domain challenges: Maritime perception remains limited by data scarcity and domain variability. Dierences across rivers, coastal regions, and open seas introduce signicant distribution shifts, requiring dataefficient learning strategies such as self-supervised learning and domain adaptation [11].

Stability, multi-modality, and deployment: Long-horizon stability under environmental disturbances remains an open problem, as accumulated state errors may aect performance [11, 13]. In addition, integrating multi-modal sensor data and ensuring efficient deployment on resource-constrained hardware require further investigation [7]. Standardized benchmarks are also needed to enable fair comparison across methods [11].

In summary, SSM and Mamba architectures oer clear efficiency advantages for long-sequence modeling, but introduce trade-os in representational capacity. Future work should focus on hybrid designs, robustness, and domain-specic adaptation to fully realize their potential in USV perception systems.

5.2. Conclusion

This paper presented a focused review of SSM and Mamba for USV perception. We highlighted limitations of attention-based models and showed that state space approaches provide linear complexity and efficient temporal modeling. Supported by recent benchmarking and survey studies, SSM-based models oer a strong balance between efficiency and scalability

for real-time maritime systems. However, challenges in data, robustness, and deployment remain, requiring further research for practical adoption.

References

1. Manley, J. E. (2008). *Unmanned surface vehicles*, 15 years of development. In: In Oceans 2008. Ieee. , 1-4.
2. Qiao, Y., Yin, J., Wang, W., Duarte, F., Yang, J., & Ratti, C. (2023). *Survey of deep learning for autonomous surface vehicles in marine environments*. IEEE Transactions on Intelligent Transportation Systems, 24(4), 3678-3701.
3. Sayan, S., & Ekenel, H. K. (2025). *Obstacle detection and ship recognition system for unmanned surface vehicles*. Proceedings Copyright, 113, 122.
4. Jungbauer, N., Huang, H., & Mayer, H. (2025). *Maritime vision datasets for autonomous navigation: A comparative analysis*. Maritime Technology and Research, 7(4).
5. Abd Alhattab, Y., Abidin, Z. B. Z., Faizabadi, A. R., Zaki, H. F., & Ibrahim, A. I. (2023). *Integration of stereo vision and moos-ivp for enhanced obstacle detection and navigation in unmanned surface vehicles*. IEEE Access, 11, 128932-128956.
6. Li, J., Zhang, G., Jiang, C., & Zhang, W. (2023). *A survey of maritime unmanned search system: Theory, applications and future directions*. Ocean Engineering, 285, 115359.
7. Aalishah, R., Navardi, M., & Mohsenin, T. *Edgenavmamba: Mamba-optimized object detection for energy-efficient edge devices*. In: In 2025 ieee 11th international conference on edge computing and scalable cloud (edgecom). IEEE. 2025, 62-67.
8. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention is all you need*. Advances in neural information processing systems, 30.
9. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. *Swin transformer: Hierarchical vision transformer using shifted windows*. In: In Proceedings of the ieee/cvf international conference on computer vision. 2021, 10012-10022.
10. Tay, Y., Dehghani, M., Bahri, D., & Metzler, D. (2022). *Efficient transformers: A survey*. ACM Computing Surveys, 55(6), 1-28.
11. Dai, T., Xiang, H., Leng, C., Huang, S., He, G., & Han, S. (2024). *Ldmnet: Enhancing the segmentation capabilities of unmanned surface vehicles in complex waterway scenarios*. Applied Sciences, 14(17), 7706.
12. You, Z., Wang, N., Wang, H., Zhao, Q., & Wang, J. (2024). *Mambabev: An efficient 3d detection model with mamba2*. arXiv preprint arXiv:2410.12673.
13. Li, J., Liu, X., Li, B., Xu, R., Li, J., Yu, H., & Tu, Z. *Comamba: Realtime cooperative perception unlocked with state-space models*. In: In 2025 ieee/rsj international conference on intelligent robots and systems (iros). IEEE. 2025, 16993-17000.
14. Gu, A., & Dao, T. *Mamba: Linear-time sequence modeling with selective state spaces*. In: In First conference on language modeling. 2024.
15. Tian, Y., Bai, S., Luo, Z., Wang, Y., Zhang, H., Guo, B., Lv, Y., & Wang, F.-Y. (2026). *Mambaocc: Visual state space models for bev-based occupancy prediction with local adaptive reordering*. Expert Systems with Applications, 131305.
16. Gu, A., Goel, K., & Ré, C. (2021). *Efficiently modeling long sequences with structured state spaces*. arXiv preprint arXiv:2111.00396.

17. Liu, Y., Tian, Y., Zhao, Y., Yu, H., Xie, L., Wang, Y., Ye, Q., Jiao, J., & Liu, Y. (2024). *Vmamba: Visual state space model*. Advances in neural information processing systems, 37, 103031-103063.
18. Li, H., Hou, Y., Xing, X., Ma, Y., Sun, X., & Zhang, Y. *Occmamba: Semantic occupancy prediction with state space models*. In: In Proceedings of the ieee/cvf conference on computer vision and pattern recognition. 2025, 11949-11959.
19. Cai, Q., Wang, Q., Zhang, Y., He, Z., & Zhang, Y. (2023). *Lwdnet-a lightweight water-obstacles detection network for unmanned surface vehicles*. Robotics and Autonomous Systems, 166, 104453.
20. Bovcon, B., Muhovi£, J., Vranac, D., Mozeti£, D., Per², J., & Kristan, M. (2021). *Mods - a usv-oriented object detection and obstacle segmentation benchmark*. IEEE Transactions on Intelligent Transportation Systems, 23(8), 13403-13418.
21. Bovcon, B., Per², J., Kristan, M., et al. (2018). *Stereo obstacle detection for unmanned surface vehicles by imu-assisted semantic segmentation*. Robotics and Autonomous Systems, 104, 1-13.
22. Wu, Y., Fioranelli, F., & Gao, C. (2025). *Radmamba: Efficient human activity recognition through a radar-based micro-doppler-oriented mamba state-space model*. IEEE Transactions on Radar Systems, 4, 261-272.
23. Li, Z., Zhao, L., Lu, Y., Yue, M., & Li, G. (2025c). *Mamba for remote sensing: Architectures, hybrid paradigms, and future directions*.
24. Ke, H., Morris, J., Oguchi, K., Cao, X., Liu, Y., Wang, H., & Ding, Y. (2025). *Mambev: Enabling state space models to learn birds-eye-view representations*. In: In The thirteenth international conference on learning representations.
25. Elkins, L., Sellers, D., & Monach, W. R. (2010). *The autonomous maritime navigation (amn) project: Field tests, autonomous and cooperative behaviors, data fusion, sensors, and vehicles*. Journal of Field Robotics, 27(6), 790-818.
26. Jia, F., Zhang, F., Zhao, P., & Sun, G. (2025). *Rd mamba: A lightweight radar range-doppler spectrum object detection model*.
27. Zhu, L., Liao, B., Zhang, Q., Wang, X., Liu, W., & Wang, X. (2024). *Vision mamba: Efficient visual representation learning with bidirectional state space model*. arXiv preprint arXiv:2401.09417.
28. Guan, M., Xu, H., Jiang, G., Yu, M., Chen, Y., Luo, T., & Song, Y. (2024). *Watermamba: Visual state space model for underwater image enhancement*. arXiv preprint arXiv:2405.08419.
29. Tian, S., Li, M., Cao, L., Kang, L., Tian, J., Xing, X., Shen, B., Du, K., Fu, C., & Zhang, Y. (2025). *Hybrid attention driven cnn-mamba multimodal fusion network for remote sensing image semantic segmentation*. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 19, 2254-2272.
30. Hatamizadeh, A., & Kautz, J. *Mambavision: A hybrid mamba-transformer vision backbone*. In: In Proceedings of the computer vision and pattern recognition conference. 2025, 25261-25270.
31. Ma, J., Li, F., & Wang, B. (2024). *U-mamba: Enhancing long-range dependency for biomedical image segmentation*. arXiv preprint arXiv:2401.04722.
32. Li, Z., Wang, W., Li, H., Xie, E., Sima, C., Lu, T., Yu, Q., & Dai, J. (2024). *Bevformer: Learning bird's-eye-view representation from lidar-camera via spatiotemporal transformers*. IEEE Transactions on PAMI, 47(3), 2020-2036.
33. Yao, S., Guan, R., Wu, Z., Ni, Y., Huang, Z., Liu, R. W., Yue, Y., Ding, W., Lim, E. G., Seo, H., et al. (2024). *Waterscenes: A multi-task 4d radar-camera fusion dataset and benchmarks for autonomous driving on water surfaces*. IEEE Transactions on Intelligent Transportation Systems, 25(11), 16584-16598.

34. Trinh, L., Mercelis, S., & Anwar, A. (2025). *A comprehensive review of datasets and deep learning techniques for vision in unmanned surface vehicles*. Ocean Engineering, 334, 121501.
35. Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. *Microsoft coco: Common objects in context*. In: In European conference on computer vision. Springer. 2014, 740-755.
36. Long, J., Shelhamer, E., & Darrell, T. *Fully convolutional networks for semantic segmentation*. In: In Proceedings of the ieee conference on computer vision and pattern recognition. 2015, 3431-3440.
37. Bernardin, K., & Stiefelhagen, R. (2008). *Evaluating multiple object tracking performance: The clear mot metrics*. EURASIP Journal on Image and Video Processing, 2008(1), 246309.
38. Ristani, E., Solera, F., Zou, R., Cucchiara, R., & Tomasi, C. *Performance measures and a data set for multi-target, multi-camera tracking*. In: In European conference on computer vision. Springer. 2016, 17-35.
39. Leal-Taixé, L., Milan, A., Reid, I., Roth, S., & Schindler, K. (2015). *Motchallenge 2015: Towards a benchmark for multi-target tracking*. arXiv preprint arXiv:1504.01942.
40. Binas, J., Neil, D., Liu, S.-C., & Delbruck, T. (2017). *Ddd17: End-to-end davis driving dataset*. arXiv preprint arXiv:1711.01458.
41. Bovcon, B., Muhovi£, J., Per², J., & Kristan, M. *The mastr1325 dataset for training deep usv obstacle detection models*. In: In 2019 ieee/rsj international conference on intelligent robots and systems (iros). IEEE. 2019, 3431-3438.
42. Han, S., Lee, D., Kim, J., & Kim, J. (2026). *A survey on maritime perception datasets and technologies for autonomous surface vessels*. Intelligent Service Robotics, 19(2), 25.
43. Koledoye, A., Unachukwu, C., Nwobu, G., & Rana, H. (2026). *Benchmarking the computational and representational efficiency of state space models against transformers on long-context dyadic sessions*. arXiv preprint arXiv:2601.01237.
44. Patro, B. N., & Agneeswaran, V. S. (2025). *Mamba-360: Survey of state space models as transformer alternative for long sequence modelling: Methods, applications, and challenges*. Engineering Applications of Artifficial Intelligence, 159, 111279.
45. Lv, X., Sun, Y., Zhang, K., Qu, S., Zhu, X., Fan, Y., Wu, Y., Hua, E., Long, X., Ding, N., et al. (2025). *Technologies on effectiveness and efficiency: A survey of state spaces models*. arXiv preprint arXiv:2503.11224.

Vượt ra ngoài các mô hình Attention-Centric: Mô hình không gian trạng thái và Mamba cho nhận thức của xuồng tự hành

Tóm tắt: *Khả năng nhận thức đóng vai trò quan trọng đối với các phương tiện USV tự hành hoạt động trong môi trường hàng hải năng động. Mặc dù các mô hình dựa trên cơ chế chú ý như Transformer đạt được độ chính xác cao, nhưng độ phức tạp bậc hai và yêu cầu tài nguyên lớn của chúng hạn chế việc triển khai thời gian thực trên các nền tảng nhúng. Các mô hình SSM và kiến trúc Mamba cung cấp một giải pháp thay thế hiệu quả bằng cách mô hình hóa các phụ thuộc tuần tự thông qua các chuyển đổi trạng thái có cấu trúc với độ phức tạp tuyến tính. Bài báo này xem xét quá trình chuyển đổi từ các mô hình tập trung vào cơ chế chú ý sang các phương pháp dựa trên SSM và Mamba cho khả năng nhận thức của USV. Dựa trên những hiểu biết từ các lĩnh vực liên quan, chúng tôi phân tích các nguyên tắc kiến trúc, sự đánh đổi giữa hiệu quả và hiệu suất, và các cân nhắc triển khai thực tế. Chúng tôi cũng thảo luận thêm về các chiến lược đánh giá và nêu ra các thách thức còn tồn tại, bao gồm mô hình lai và triển khai nhúng thời gian thực..*

Từ khóa: *USVs; Attention; Mamba Architecture; Maritime Vision.*

Development of data consistency integrated deep learning model for reconstructing undersampled MRI k-space

Bui Thi Hoa[1]*, Pham Duc Khanh[2], Nguyen Trong Tuyen[3], Tran Anh Quang[3]

[1]*NCS K45 - Hệ 2 - Học viện kỹ thuật Quân sự;*

[2]*NCS K44 - Hệ 2 - Học viện kỹ thuật Quân sự;*

[3]*Bộ môn Điện tử y sinh-Viện Tên lửa và kỹ thuật điều khiển-Học viện Kỹ thuật Quân sự*

hoabt@lqdtu.edu.vn; Tel: 0965213024

Abstract

Accelerating magnetic resonance imaging (***MRI***) acquisition is important for reducing scan time, improving patient comfort, and minimizing motion-induced artifacts. Although conventional compressed sensing (***CS***) methods have shown potential for accelerated ***MRI***, their performance is often limited by reconstruction instability, high computational cost, and insufficient preservation of fine anatomical details. To address these limitations, we develop an integration of data consistency ***(DC***) with the convolutional block attention module (***CBAM***) modified deep learning-based reconstruction framework, which allows recovering high-quality ***MRI*** images from undersampled k-space data while maintaining structural fidelity under a 4x ***MRI*** acceleration. The ***DC*** layer enforces consistency between the reconstructed k-space and the originally acquired measurements, thereby incorporating ***MRI*** physics into the reconstruction process, while the ***CBAM*** enhances feature representation by focusing on informative channel and spatial features. Our experimental results demonstrate that the proposed model substantially outperforms the conventional U-Net baseline. In particular, ***PSNR*** increases from 24.81 dB to 30.74 dB, and ***SSIM*** improves from 0.8435 to 0.9116, corresponding to gains of approximately 59.2% and 46.25%, respectively. These obtained results indicate that the proposed model is a promising solution for accelerated ***MRI*** reconstruction, enabling faster acquisition while preserving diagnostically relevant image quality.

Keywork: *Undersampled reconstruction; U-net; Data consistency; deep-learning reconstruction.*

1. Introduction

Magnetic Resonance Imaging (***MRI***) is one of the most advanced and widely used medical imaging modalities due to its ability to provide high-resolution images with excellent soft-tissue contrast. However, a major limitation of ***MRI*** systems is the relatively long data acquisition time. In clinical practice, patients are often required to remain motionless for prolonged periods, sometimes up to 45–60 minutes, which may cause discomfort, anxiety, and motion-induced artifacts that degrade image quality. Consequently, reducing ***MRI*** acquisition time while preserving diagnostic image quality has currently become an important and challenging research problem.

Several approaches have been proposed to accelerate ***MRI*** acquisition, including parallel imaging [1] and undersampling strategies in k-space [2, 3]. Among them, compressed sensing MRI (***CS-MRI***) has emerged as a promising solution, as it enables fast and high-quality ***MRI*** image acquisition by exploiting sparsity in suitable transform domains. Since the seminal work of Candes et al [4], CS theory has been successfully applied to ***MRI***, demonstrating that high-quality images can be reconstructed from fewer measurements than required by the Nyquist criterion. Early studies by Lustig et al.[5] validated the feasibility of ***CS-MRI*** and showed that random or pseudo-random undersampling combined with iterative

*Email: hoabt@lqdtu.edu.vn

optimization algorithms can effectively reduce scan time. However, conventional ***CS-MRI*** methods rely heavily on handcrafted sparsifying transforms and iterative optimization algorithms such as Orthogonal Matching Pursuit [6] or Nonlinear Conjugate Gradient (***NCG***) [4, 5]. These approaches often suffer from high computational cost, sensitivity to parameter selection, and limited robustness when applied to different sampling patterns or anatomical structures. Moreover, purely random undersampling strategies, although theoretically favorable, may lead to unstable reconstruction performance in practice.

In recent years, deep learning (***DL***) has shown remarkable success in medical image processing [7, 8], including ***MRI*** reconstruction [9, 10], super-resolution [11, 12]. Convolutional Neural Networks (***CNNs***), particularly U-Net-based architectures, are capable of learning complex nonlinear mappings from undersampled data to fully sampled images by exploiting large datasets [13]. Several studies have demonstrated that DL-based ***MRI*** reconstruction methods can significantly outperform traditional ***CS*** approaches in terms of reconstruction quality and computational efficiency [14-16]. However, many purely data-driven models tend to ignore the physical acquisition model of ***MRI*** [17], which may result in hallucinated structures and reduced generalization across different sampling schemes.

To address these limitations, a hybrid approach that integrate physics-based constraints into deep learning architectures has been proposed for undersampled k-space reconstruction in the ***CS-MRI***. Among these, data consistency (***DC***) layers have proven effective in enforcing agreement between reconstructed k-space data and acquired measurements [17, 18]. Additionally, attention mechanisms such as the Convolutional Block Attention Module (***CBAM***) have been known to enhance feature representation by focusing on informative channels and spatial regions [19, 20]. Therefore, in this work, we propose a deep cascaded U-Net framework incorporating ***DC*** and ***CBAM*** for undersampled ***MR***I reconstruction. The main contributions of our study consist of: the design of a cascaded U-Net architecture for progressive ***MRI*** reconstruction; the integration of a learnable ***DC*** layer to enforce fidelity to measured k-space data; and the incorporation of ***CBAM*** to improve feature discrimination and fine-detail recovery.

2. Data and methods

2.1 Dataset

In this study, experiments were conducted on publicly available brain ***MRI*** datasets, specifically the IXI database [21]. The dataset comprises structural magnetic resonance images acquired from 600 normal, healthy subjects. From this database, a sub-dataset consisting of 86,794 T1-weighted brain ***MRI*** 2D slices was selected, with all images resized to 256x256 pixels. To ensure a rigorous and fair evaluation of the model's performance and to strictly prevent data leakage, the dataset was partitioned subject-wise into training, validation, and testing sets with a ratio of 80%:10%:10%, respectively. Specifically, data from 480 subjects were used for training, 60 subjects for validation (to tune hyperparameters and prevent overfitting during training), and the remaining unseen 60 subjects for final testing. Consequently, fully sampled k-space data were retrospectively undersampled using a predefined hybrid undersampling mask proposed from our previous studies with an acceleration factor of 4x [22, 23]. This under-sampling method has also been optimized as

described in our recent work [24]. Data normalization and standard preprocessing steps were applied prior to feeding the images into the network.

2.2 Undersampling and Zero-filling reconstruction

In this work, zero-filled inverse Fourier reconstruction [25] was used as a baseline input to the deep learning models. Let x denote the fully sampled ***MRI*** image and F is the Fourier transform operator. The k-space measurements y can be expressed as Equation 1:

$$y=M\odot Fx \tag{1}$$

Where M is a binary undersampling mask defined by the hybrid undersampling as described from our previous study [24]. Zero-filling reconstruction (X_{zf}) is obtained by applying the inverse Fourier transform (F^{-1}) to the undersampled k-space data by this Equation 2:

$$X_{zf}=F^{-1}(M\odot Fx) \tag{2}$$

2.3 U-net architecture

The U-Net architecture serves as the backbone of the proposed reconstruction framework. U-Net is a convolutional neural network characterized by an encoder–decoder structure with skip connections as shown in the Fig. 1 [13]. The encoder progressively extracts hierarchical features through convolution and downsampling operations, while the decoder restores spatial resolution using upsampling layers and feature concatenation from corresponding encoder stages.

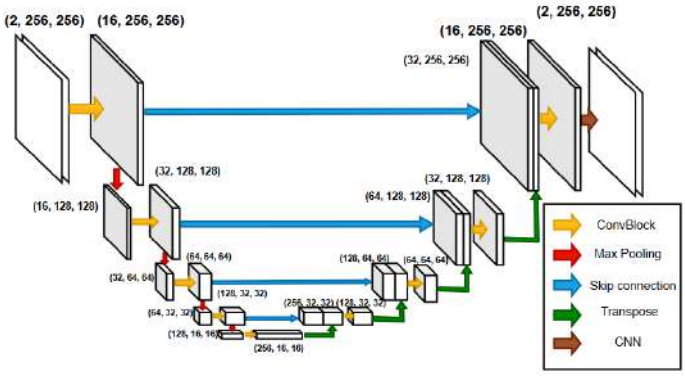


Figure 1. The architecture of basic U-net backbone network

In this work, the encoder consists of convolutional blocks with increasing numbers of feature channels (16, 32, 64, 128, and 256). Each block includes two 3x3 convolutional layers followed by normalization and nonlinear activation. The decoder mirrors the encoder structure, combining upsampling and convolution operations with skip connections to preserve fine-grained spatial details. The total number of trainable parameters is approximately 1.94 million, providing a balance between reconstruction accuracy and computational efficiency.

2.4 CBAM integrated U-net model with Data Consistency

To enhance discriminative feature representation, the Convolutional Block Attention Module (***CBAM***) is proposed to be integrated into the decoder stages of the U-net as indicated in the Fig. 2. ***CBAM*** sequentially applies channel attention and spatial attention mechanisms to adaptively recalibrate feature responses [19, 20]. In this module, channel attention enables

the network to emphasize informative feature channels related to anatomical structures using both average pooling and max pooling, while spatial attention highlights relevant spatial regions and suppresses background noise and artifacts via convolutional attention maps. This attention-driven feature refinement is particularly beneficial for recovering fine-grained details that are often degraded by undersampling. By combining these mechanisms, ***CBAM*** will enable the network to selectively focus on diagnostically relevant structures while suppressing noise and artifacts.

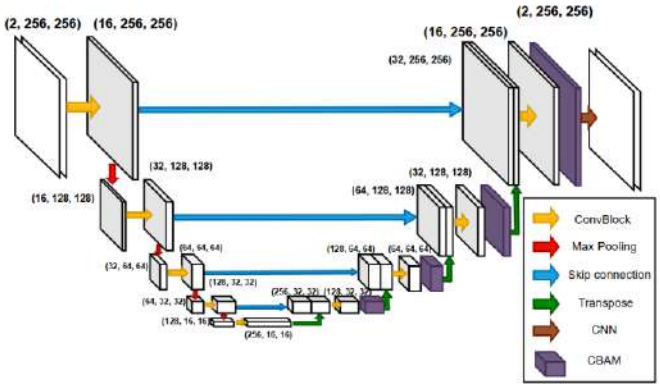


Figure 2. The architecture of the proposal model (CBAM integrated U-net)

In this work, ***MRI*** acquisition physics is also incorporated into the deep learning model by a ***DC*** layer integrated after each cascaded U-net stage. The ***DC*** operation ensures that the reconstructed k-space remains consistent with the originally acquired measurements. Formally, if a reconstructed image is X_{recon} , its k-space representation K_{DC} is updated as described in Equation 3:

$$K_{DC}=(1-\mathrm{M})\odot\mathrm{F}(X_{recon})+\mathrm{M}\odot\left(\frac{y+\lambda F(X_{recon})}{1+\lambda}\right) \quad (3)$$

Where y denotes the acquired k-space samples. A learnable parameter λ controls the strength of the DC constraint, allowing the model to balance learned priors and data fidelity. In the proposed model, λ was initialized to 0.01 and optimized during training.

To further improve reconstruction quality, a cascaded architecture was adopted. Instead of performing reconstruction in a single pass, each ***CBAM*** integrated U-Net stage is connected to a ***DC*** layer sequentially as shown in Fig. 3. Each stage refines the output of the previous stage, progressively reducing artifacts and enhancing structural details. In this study, a seven-stage cascaded U-net with ***CBAM*** connected to a ***DC*** layer was employed, which demonstrated improved convergence and reconstruction performance without excessive computational overhead

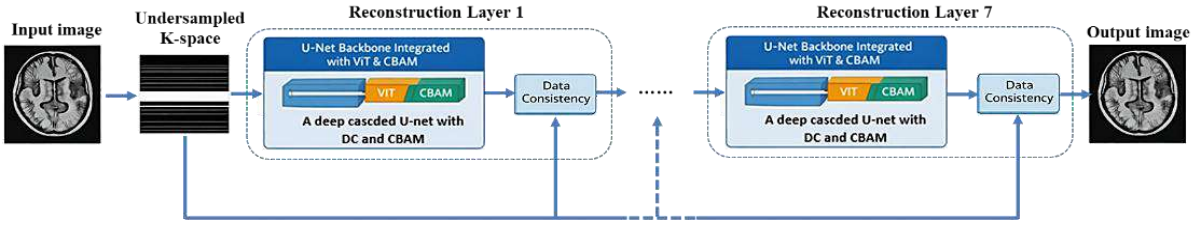


Figure 3. Illustrates the overall pipeline of the cascaded U-net with DC and CBAM layers

By integrating ***CBAM***-based adaptive attention and learnable ***DC*** constraints within a cascaded U-Net framework, the proposed architecture is expected to effectively balance local detail preservation, global anatomical coherence, and physical consistency with measured k-space data. Importantly, the total number of trainable parameters is approximately only 1.946 million that is a negligible increase as compared to 1.94 million for the traditional U-net model as aforementioned.

The model was trained using supervised learning, with fully sampled images as ground truth. The loss function combined pixel-wise reconstruction error and structural similarity measures. Model performance was evaluated using Peak Signal-to-Noise Ratio (***PSNR***) for measuring the reconstruction fidelity by comparing the reconstructed MRI images with the fully sampled original images and Structural Similarity Index Measure (***SSIM***) assess perceptual similarity between the reconstructed and reference images. These image quality indexes are standard metrics for ***MRI*** reconstruction quality assessment.

3. Results and discussions

3.1 Reconstruction performance of the conventional U-Net

The conventional U-net model was first evaluated on the test dataset and compared with zero-filled reconstruction. Experimental results show that the U-Net significantly reduces aliasing artifacts and recovers most anatomical structures compared to the zero-filling baseline as indicated in Fig. 4. The reconstructed images exhibit improved visual quality and clearer structural boundaries.

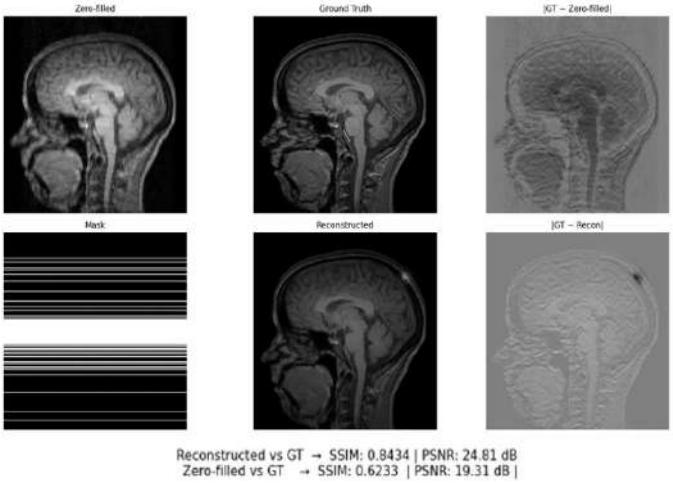


Figure 4. The reconstructed image by the conventional U-Net

Quantitatively, the conventional U-Net achieved an ***SSIM*** of approximately 0.84 on the test set, representing a substantial improvement over zero-filled reconstruction. However, residual blurring and loss of fine details were still observed, particularly in regions with complex textures. Error maps further indicate that certain high-frequency components were not fully recovered.

3.2 Reconstruction performance of the CBAM integrated U-Net with DC

By integrating ***CBAM*** and ***DC*** into the cascaded U-net framework, reconstruction quality is substantially improved as shown in Fig. 5. The proposed model achieves an ***SSIM*** of approximately 0.912, along with notable gains in ***PSNR*** and reductions in reconstruction

error. Visual inspection in the Fig. 5 obviously reveals sharper edges, improved contrast, and enhanced preservation of fine anatomical details.

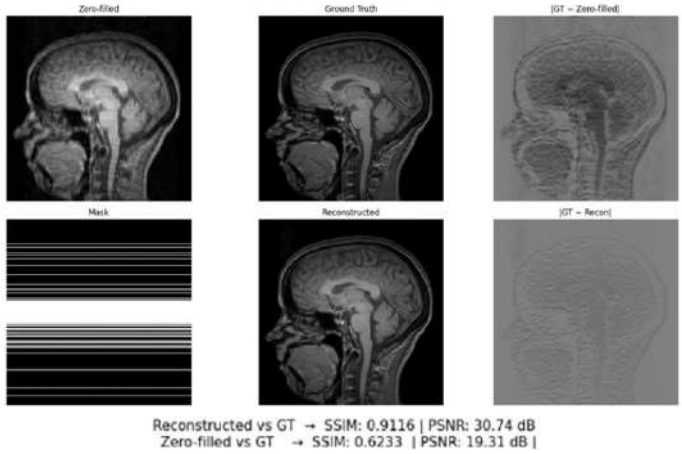


Figure 5. The reconstructed image by the proposed CBAM integrated U-net+DC

Notably, the inclusion of the ***DC*** and ***CBAM*** layer introduces only a small number of additional learnable parameters, resulting in minimal increase in computational complexity. The total number of trainable parameters for this cascaded U-net model is only approximately 1.95 million, nearly equivalent to a standard U-Net, but thanks to ***CBAM*** and ***DC***, the network has more flexible attention capabilities, enhancing image reconstruction quality and improving accuracy in areas with critical detail. This is a significant advantage when applied to medical image reconstruction, where small details sometimes have great diagnostic value. From our obtained results, the performance gain of the proposed model is significant, highlighting the effectiveness of enforcing physical consistency and informative feature channels within deep learning-based ***MRI*** reconstruction.

Moreover, the experimental results demonstrate that while a conventional deep learning model is capable of learning effective priors for ***CS-MRI*** reconstruction, incorporating ***DC*** and ***CBAM*** constraints is crucial for achieving stable and reliable performance. This cascaded U-Net with ***DC*** and ***CBAM*** layers can be interpreted as an unrolled optimization process, where each stage performs a refinement step guided by both learned features and measured data. For more comprehensive evaluation of our proposed reconstruction model, we also compare the performance of our model with some other recent methods with different architectural variants-based U-net or CNN networks as described in Table 1.

Table 1. Performance comparison of different architectural variants and our examined methods

Models	PSNR (dB)	SSIM
U-net -Dssim [15]	28.83	0.8848
Projection-based cascaded U-Net[16]	**33.685**	0.816
Projection-based cascaded CNN [26]	**32.827**	0.801
BT-U-net (Transformer at Bottleneck)[27]	29.87	0.7102
SwinU-net[27]	**33.1**	0.7274
Zero-filling reconstruction	19.31	0.6233
Traditional U-net model	24.81	0.8435
The proposed DC+CBAM + U-net model	**30.74**	**0.9116**

As indicated in the table, zero-filling reconstruction yields the poorest performance due to severe aliasing artifacts, while the conventional U-net significantly improves both ***PSNR***

and ***SSIM*** by learning effective image priors from undersampled data. The proposed ***DC***+***CBAM*** U-Net model achieves a notable improvement over the traditional U-Net, with ***PSNR*** increasing from 24.81 dB to 30.74 dB and ***SSIM*** improving from 0.8435 to 0.9116. These results demonstrate that incorporating ***DC*** and ***CBAM*** attention mechanisms substantially enhances reconstruction quality. The ***DC*** layer enforces fidelity to the acquired k-space measurements, ensuring physical consistency and stable reconstruction, while ***CBAM*** enables the network to focus on informative channels and spatial regions, leading to improved recovery of fine anatomical details. In this Table 1, we also compared with other advanced architectures, including projection-based cascaded networks and transformer-based models, the proposed method exhibits a more balanced trade-off between ***PSNR*** and ***SSIM***. Although some methods achieve higher ***PSNR*** values, their lower ***SSIM*** scores indicate weaker structural preservation. In contrast, the proposed model consistently maintains high structural similarity while achieving competitive reconstruction fidelity, which is particularly important for preserving diagnostically relevant features.

4. Conclusion

This study has proposed a deep learning-based ***MRI*** reconstruction framework that integrates data consistency (***DC***) and attention mechanisms (***CBAM***) within a cascaded U-Net architecture. By combining data-driven learning with physics-based constraints, the proposed method achieves significant improvements in reconstruction quality, with approximately 46.25% improvement in ***SSIM*** and 59.2% improvement in ***PSNR*** compared to the conventional U-Net based reconstruction model, when accelerating the ***MRI*** scan in 4 times. The proposed framework demonstrates strong potential for accelerated ***MRI*** acquisition while preserving high diagnostic image quality. Our future work will focus on extending on optimizing the proposed model, exploring more different undersampling patterns, and validating the approach on larger and more diverse clinical datasets.

References

1. Deshmane, A., V. Gulani, M.A. Griswold, and N. Seiberlich, (2012). *Parallel MR imaging.* J Magn Reson Imaging, 36(1): p. 55-72.
2. Kemp, T.M., T.A. Kawakita, R. Mehta, and A.R. Pineda, (2023). *Optimizing data acquisition in undersampled magnetic resonance imaging (MRI) using two alternative forced choice (2-AFC) and search tasks.* Proc SPIE Int Soc Opt Eng, 12467.
3. Razumov, A., O. Rogov, and D.V. Dylov, (2023). *Optimal MRI undersampling patterns for ultimate benefit of medical vision tasks.* Magnetic Resonance Imaging, 103: p. 37-47.
4. Candes, E.J., J. Romberg, and T. Tao, (2006). *Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information.* IEEE Transactions on Information Theory, 52(2): p. 489-509.
5. Lustig, M., D. Donoho, and J.M. Pauly, (2007). *Sparse MRI: The application of compressed sensing for rapid MR imaging.* Magnetic Resonance in Medicine, 58(6): p. 1182-1195.
6. Tropp, J.A. and A.C. Gilbert, (2007). *Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit.* IEEE Transactions on Information Theory, 53(12): p. 4655-4666.
7. Rayed, M.E., S.M.S. Islam, S.I. Niha, J.R. Jim, M.M. Kabir, and M.F. Mridha, (2024). *Deep learning for medical image segmentation: State-of-the-art advancements and challenges.* Informatics in Medicine Unlocked, 47: p. 101504.
8. Li, M., Y. Jiang, Y. Zhang, and H. Zhu, (2023). *Medical image analysis using deep learning algorithms.* Front Public Health, 11: p. 1273253.

9. Safari, M., Z. Eidex, C.-W. Chang, R.L.J. Qiu, and X. Yang, (2026). *Advancing MRI reconstruction: A systematic review of deep learning and compressed sensing integration.* Biomedical Signal Processing and Control, 111: p. 108291.
10. Heckel, R., M. Jacob, A. Chaudhari, O. Perlman, and E. Shimron, (2024). *Deep learning for accelerated and robust MRI reconstruction.* 37(3): p. 335-368.
11. Khateri, M., S. Vasylechko, M. Ghahremani, L. Timms, D. Kocanaogullari, S. Warfield, C. Jaimes, D. Karimi, A. Sierra, J. Tohka, S. Kurugol, and O. Afacan, *MRI Super-Resolution with Deep Learning: A Comprehensive Survey*. 2025.
12. Ayaz, A., R. Boonstoppel, C. Lorenz, J. Weese, J. Pluim, and M. Breeuwer, (2024). *Effective deep-learning brain MRI super resolution using simulated training data.* Computers in Biology and Medicine, 183: p. 109301.
13. Jiangtao, W., N.I.R. Ruhaiyem, and F. Panpan, (2025). *A Comprehensive Review of U-Net and Its Variants: Advances and Applications in Medical Image Segmentation.* IET Image Processing, 19.
14. Elbehairy, A., I. Yassine, and M. Elattar, *Stage U-Net Framework: Streamlining MRI Reconstruction from Under- sampled K-Space*. 2024.
15. Ghodrati, V., J. Shao, M. Bydder, Z. Zhou, W. Yin, K.L. Nguyen, Y. Yang, and P. Hu, (2019). *MR image reconstruction using deep learning: evaluation of network structure and loss functions.* Quant Imaging Med Surg, 9(9): p. 1516-1527.
16. Aghabiglou, A. and E.M. Eksioglu, (2021). *Projection-Based cascaded U-Net model for MR image reconstruction.* Computer Methods and Programs in Biomedicine, 207: p. 106151.
17. Hammernik, K., T. Küstner, B. Yaman, Z. Huang, D. Rueckert, F. Knoll, and M. Akçakaya, (2023). *Physics-Driven Deep Learning for Computational Magnetic Resonance Imaging: Combining physics and machine learning for improved medical imaging.* IEEE Signal Process Mag, 40(1): p. 98-114.
18. Huang, J., Y. Wu, F. Wang, Y. Fang, Y. Nan, C. Alkan, D. Abraham, C. Liao, L. Xu, Z. Gao, W. Wu, L. Zhu, Z. Chen, P. Lally, N. Bangerter, K. Setsompop, Y. Guo, D. Rueckert, G. Wang, and G. Yang, (2025). *Data- and Physics-Driven Deep Learning Based Reconstruction for Fast MRI: Fundamentals and Methodologies.* IEEE Reviews in Biomedical Engineering, 18: p. 152-171.
19. Shyamala, N. and S. Mahaboobbasha. *Convolutional Block Attention Module-based Deep Learning Model for MRI Brain Tumor Identification (ResNet-CBAM)*. in *2024 5th International Conference on Smart Electronics and Communication (ICOSEC)*. 2024.
20. Li, Y., H. Zhu, T. Hou, N. Chen, B. Huang, W. Lu, Y. Wang, and S. Yang, (2025). *A parallel dual-domain crossing network combining attention enhancement and multimodal fusion for MRI reconstruction.* Biomedical Signal Processing and Control, 110: p. 108326.
21. brain-development.org, *IXI dataset (RRID:SCR_005839)*. 2025.
22. Anh Quang Tran, T.-A.N., Phuc Thinh Doan , Duc-Nghia Tran , Duc-Tan Tran (2021). *Parallel magnetic resonance imaging acceleration with a hybrid sensing approach.* Mathematical Biosciences and Engineering, 18(3): p. 2288-2302.
23. Tran, A.Q., T.-A. Nguyen, V.T. Duong, Q.-H. Tran, D.N. Tran, and D.-T. Tran, (2020). *MRI Simulation-based evaluation of an efficient under-sampling approach.* Mathematical Biosciences and Engineering, 17(4): p. 4048-4063.
24. Khanh Pham, D., D.-T. Tran, and A. Quang Tran, (2025). *Evaluating random–Nyquist sampling ratios in combined compressed sensing magnetic resonance imaging.* Bulletin of Electrical Engineering and Informatics.
25. Zhu, X., B. Tomanek, and J. Sharp, (2013). *A pixel is an artifact: On the necessity of zero-filling in fourier imaging.* Concepts in Magnetic Resonance Part A, 42A.
26. Kocanaogullari, D. and E.M. Eksioglu. *Deep Learning For Mri Reconstruction Using A Novel Projection Based Cascaded Network*. in *2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP)*. 2019.
27. Dong, B. and G. Wu, *A Swin Transformer-Based Cold Diffusion Model for Accelerated MRI Reconstruction*, in *Proceedings of the 2025 2nd International Conference on Generative Artificial Intelligence and Information Security*. 2025, Association for Computing Machinery. p. 226–230.

Tiêu đề: Nghiên cứu phát triển mô hình học sâu tích hợp lớp nhất quán dữ liệu dùng trong tái tạo ảnh MRI từ bộ dữ liệu lấy mẫu thưa

Tóm tắt: Tăng tốc độ thu nhận trong tạo ảnh MRI là rất cần thiết để có thể giảm thiểu thời gian chụp, giúp giảm sự khó chịu cho bệnh nhân và các hiện tượng nhiễu do chuyển động gây ra. Các kỹ thuật lấy mẫu nén tín hiệu (CS) được biết đến là những giải pháp hiệu quả để tăng tốc thu nhận ảnh MRI, tuy nhiên chúng vẫn còn những hạn chế như tính ổn định khi thu nhận và đặc biệt là chất lượng ảnh của các thuật toán tái tạo từ bộ dữ liệu thưa còn chưa cao. Trong nghiên cứu này, nhóm tác giả đề xuất phát triển một mô hình tái tạo ảnh CS-MRI sử dụng mạng học sâu U-net được tích hợp thêm cơ chế chú ý theo kênh và không gian (CBAM) kết hợp với các ràng buộc về tính nhất quán dữ liệu (DC) của quá trình thu nhận dữ liệu MRI. Lớp DC đảm bảo tính nhất quán giữa không gian tần số k-space được tái tạo và các phép đo thu được ban đầu, từ đó có thể có được các thông tin vật lý MRI vào quá trình tái tạo, trong khi CBAM tăng cường khả năng biểu diễn đặc trưng bằng cách tập trung vào các đặc trưng kênh và không gian có ý nghĩa. Kết quả thực nghiệm cho thấy mô hình đề xuất vượt trội hơn đáng kể so với mô hình U-Net cơ bản thông thường. Cụ thể, PSNR tăng từ 24,81 dB lên 30,74 dB và SSIM cải thiện từ 0,8435 lên 0,9116, tương ứng với mức tăng khoảng 59,2% và 46,25%. Những kết quả thu được cho thấy mô hình đề xuất là một giải pháp đầy hứa hẹn để tăng tốc quá trình tái tạo MRI, cho phép thu thập dữ liệu nhanh hơn trong khi vẫn duy trì chất lượng hình ảnh có ý nghĩa chẩn đoán.

Từ khóa: Tái tạo ảnh từ dữ liệu lấy mẫu thiếu; U-net; tính nhất quán dữ liệu; tái tạo ảnh bằng học sâu.

Tổng hợp luật điều khiển hợp thể thích nghi cho hệ thống treo chủ động khi có nhiễu

Đào Trọng Dũng[1], Phan Nguyên Hải[1], Lê Minh Kiên[1], Phạm Xuân Thuỷ[1], Nguyễn Chí Thanh[1], Nguyễn Xuân Chiêm[1*]

[1]*Học viện Kỹ thuật Quân sự*

Email: chiemnx@mta.edu.vn; Tel: 0966040231

Tóm tắt

Trong bài báo đề xuất luật điều khiển hợp thể thích nghi cho hệ thống treo chủ động nhằm giảm sự dao động khi di chuyển và bảo đảm hành trình treo không vượt quá giới hạn cơ khí. Trên cơ sở mô hình quarter-car phi tuyến, nhóm tác giả xây dựng một đa tạp điều khiển để giảm dao động của khối lượng treo và duy trì độ lệch của hệ treo trong miền cho phép. Để tăng khả năng thích nghi khi khối lượng treo biến thiên, một khâu tích phân được bổ sung vào cấu trúc đa tạp theo tư tưởng điều khiển hợp thể. Luật điều khiển được tổng hợp bằng phương pháp xây dựng phân tích các bộ điều khiển tổng hợp (ADAR), nhờ đó có xét đến tính phi tuyến của đối tượng và vẫn duy trì được chất lượng điều khiển trong điều kiện bất định tham số. Tính ổn định vòng kín được phân tích theo phương pháp Lyapunov. Kết quả mô phỏng cho thấy phương án đề xuất cải thiện rõ rệt đáp ứng dao động của thân xe so với hệ thống treo thụ động.

Từ khóa: *Hệ thống treo chủ động quarter-car; Điều khiển hợp thể; Điều khiển thích nghi; Hàm Lyapunov; Đa tạp*

1. Giới thiệu

Với sự phát triển mạnh mẽ của công nghệ trong ngành công nghiệp ô tô, các hệ thống treo hiện đại ngày càng được chú trọng nhằm nâng cao độ êm dịu chuyển động và khả năng điều khiển phương tiện. Hệ thống treo đóng vai trò then chốt trong việc đảm bảo độ êm dịu khi vận hành, khả năng bám đường và ổn định của xe. Vấn đề này đã thu hút sự quan tâm của nhiều nhà nghiên cứu [1, 2] nhờ khả năng thay đổi đặc tính động học theo thời gian thực thông qua cảm biến và bộ vi xử lý hiện đại, cho phép điều chỉnh hành vi xe phù hợp với các điều kiện mặt đường khác nhau. Các yêu cầu hiệu suất đối với hệ thống treo chủ động của xe được trình bày trong [1] bao gồm: (a) Độ êm dịu chuyển động – cách ly thân xe khỏi kích thích từ mặt đường; (b) Độ bám đường – duy trì tiếp xúc liên tục giữa lốp và mặt đường; (c) Giới hạn hành trình treo – bị ràng buộc bởi kết cấu cơ khí. Tuy nhiên, việc thỏa mãn đồng thời cả ba tiêu chí ở mức tối ưu là không khả thi do tính mâu thuẫn giữa chúng [3].

Trong khi hệ thống treo thụ động chỉ sử dụng lò xo và giảm chấn có thông số cố định dẫn đến hiệu quả cách ly hạn chế, hệ thống treo chủ động có thể tạo ra lực điều khiển chủ động cho phép thay đổi đặc tính động lực học một cách linh hoạt. Nhiều nghiên cứu với các luật điều khiển khác nhau đã được đề xuất, bao gồm điều khiển PID [4], điều khiển tối ưu LQR [5], điều khiển sử dụng mạng nơ-ron [6], điều khiển thích nghi backstepping [7], điều khiển trượt [8], điều khiển bền vững [9-11], điều khiển mờ [12], và các phương pháp thích nghi kết hợp học tăng cường [13]. Các nghiên cứu gần đây cũng quan tâm đến điều khiển có tính đến ràng buộc vật lý dựa trên hàm Lyapunov [14, 15]. Từ góc độ điều khiển phi tuyến, hầu hết các nghiên cứu dựa trên lựa chọn hàm Lyapunov ứng viên để đảm bảo ổn định vòng kín. Các bộ điều khiển truyền thống dựa trên hàm Lyapunov bậc hai thường hiệu quả nhưng còn hạn chế khi áp dụng cho bài toán có ràng buộc chặt chẽ. Hướng tiếp cận sử dụng hàm Lyapunov rào cản (Barrier Lyapunov Function – BLF) đã được phát triển cho bài toán điều khiển phi tuyến có ràng buộc [16], với đặc trưng hàm ứng viên tiến tới vô hạn khi biến trạng thái tiệm cận đến biên, qua đó đảm bảo các ràng buộc không bị vi phạm. Một hướng tiếp cận phổ biến khác trong điều khiển

các hệ cơ điện tử phức tạp là lý thuyết điều khiển hợp thể [17-21]. Ý tưởng cốt lõi là xây dựng hệ điều khiển có khả năng "tập hợp" nhiều bộ điều khiển con, mỗi bộ phụ trách một mục tiêu cụ thể trong hệ kín [21]. Một số nghiên cứu đã ứng dụng điều khiển hợp thể cho hệ thống treo chủ động và đạt kết quả khả quan, nhưng chưa phân tích sâu về tính ổn định khi tham số hệ thống thay đổi [17].

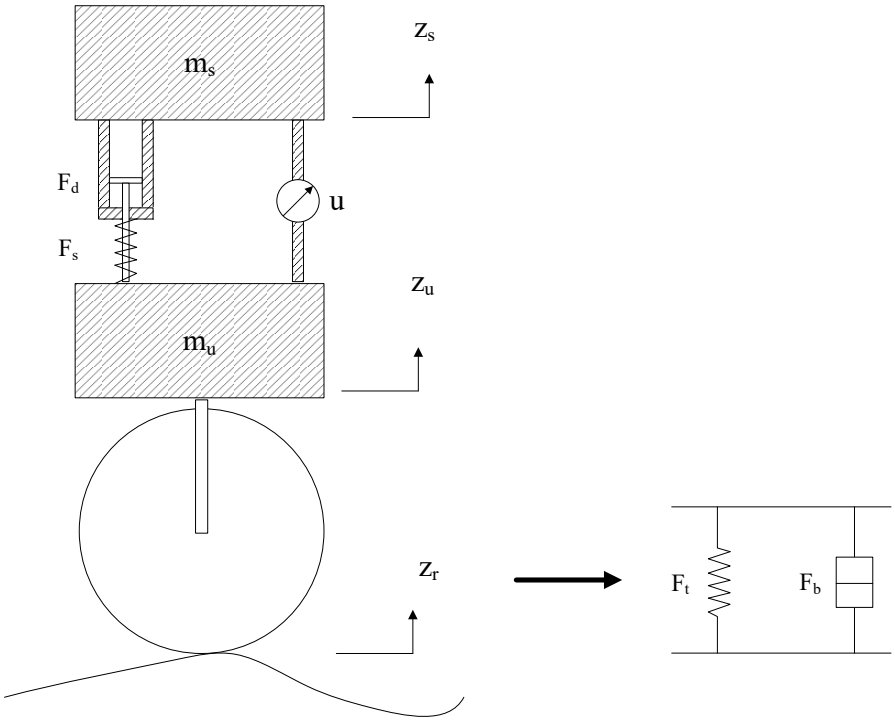
Trong bài báo này, chúng tôi đề xuất một bộ điều khiển hợp thể có bổ sung cơ chế thích nghi tích phân cho hệ thống treo chủ động. Điểm khác biệt của phương pháp là không yêu cầu ước lượng trực tiếp nhiễu, mà sử dụng cấu trúc tích phân để bù ảnh hưởng của biến thiên tham số. Cách tiếp cận này đặc biệt phù hợp với hệ thống treo ô tô, nơi thay đổi khối lượng treo và kích thích mặt đường xuất hiện thường xuyên. Tính ổn định của hệ kín được chứng minh bằng phương pháp Lyapunov; đồng thời, kết quả mô phỏng trên Matlab cho thấy bộ điều khiển đề xuất cho đáp ứng tốt hơn so với hệ treo thụ động.

Phần còn lại của bài báo được tổ chức như sau: Mục 2 trình bày mô hình động lực học của hệ thống treo chủ động có ràng buộc và tổng hợp bộ điều khiển hợp thể có thích nghi tích phân và phân tích tính ổn định khi có nhiễu tác động; Mục 3 đưa ra kịch bản mô phỏng để chứng minh hiệu quả của luật điều khiển đề xuất; Mục 4 là phần kết luận.

2. Phương pháp

2.1. Mô hình động lực học của hệ thống treo chủ động

Hệ thống treo chủ động quarter-car được minh hoạ trên hình 1, mô hình này cũng được sử dụng trong nhiều nghiên cứu đã công bố [1-7]. Mô hình xem xét gồm hai khối lượng: khối lượng treo (m_s) và khối lượng không treo (m_u). F_d và F_s lần lượt là lực do bộ giảm chấn và lò xo tạo ra; F_t, F_b lần lượt là lực đàn hồi và lực cản (giảm chấn) của lốp. z_s và z_u là độ dịch chuyển của khối lượng treo và không treo; z_r là kích thích dịch chuyển mặt đường; u là tín hiệu điều khiển chủ động của hệ thống treo.



Hình 1. Mô hình của một bộ phận treo chủ động trên ô tô

Các phương trình động lực học của khối lượng treo và khối lượng không treo được viết như sau [1]:

$$\begin{cases} \left(m_s+\delta_s\right)\ddot{z}_s+F_d\left(\dot{z}_s,\dot{z}_u\right)+F_s\left(z_s,z_u\right)=u \\ m_u\ddot{z}_u-F_d\left(\dot{z}_s,\dot{z}_u\right)-F_s\left(z_s,z_u\right)+F_t\left(z_u,z_r\right)+F_b\left(\dot{z}_s,\dot{z}_u\right)=-u \end{cases} \quad (1)$$

Trong đó các lực sinh ra bởi lò xo với độ cứng phi tuyến, giảm chấn tuyến tính từng đoạn và lốp được mô tả như sau :

$$\begin{cases} F_s\left(z_s,z_u\right)=k_{1s}\left(z_s-z_u\right)+k_{2s}\left(z_s-z_u\right)^3, \\ F_d\left(\dot{z}_s,\dot{z}_u\right)=c_{1s}\left(\dot{z}_s-\dot{z}_u\right)+c_{2s}\left(\dot{z}_s-\dot{z}_u\right)^2 \\ F_t\left(z_u,z_r\right)=k_t\left(z_u-z_r\right), \\ F_b\left(\dot{z}_u,\dot{z}_r\right)=c_t\left(\dot{z}_u-\dot{z}_r\right), \end{cases} \quad (2)$$

Trong đó k_{1s} và k_{2s} là các hệ số độ cứng của thành phần tuyến tính và phi tuyến; c_{1s} và c_{2s} là các hệ số cản của thành phần tuyến tính và phi tuyến; k_t và c_t lần lượt là hệ số độ cứng và giảm chấn của lốp. δ_s – là phần khối lượng treo thay đổi khi vận hành.

Định nghĩa biến trạng thái của hệ thống quarter-car như sau :

$$x_1=z_s, x_2=\dot{z}_s, x_3=z_u, x_4=\dot{z}_u, \quad (3)$$

Từ đó, hệ phương trình (1) có thể được biểu diễn trong không gian trạng thái như sau :

$$\begin{cases} \dot{x}_1=x_2 \\ \dot{x}_2=\dfrac{-F_d-F_s+u}{m_s+\delta_s} \\ \dot{x}_3=x_4 \\ \dot{x}_4=\dfrac{F_d+F_s-k_tx_3-c_tx_4-u+d(t)}{m_u} \end{cases} \quad (4)$$

Trong đó $d(t)=k_tz_r+c_t\dot{z}_r$ là nhiễu đại diện cho bề mặt của mặt đường.

Giả sử 1. Hệ thống di chuyển với nhiễu mặt đường đảm bảo các trạng thái đều bị chặn hay $|x_i|<\Delta_i$ với $i=1\div4$.

Mục tiêu của nghiên cứu này là tổng hợp một luật điều khiển thích nghi u trong hệ thống (4) nhằm đảm bảo hệ thống treo chủ động đáp ứng các yêu cầu sau [3]:

- Mục tiêu điều khiển chính là bộ điều khiển phải ổn định chuyển động thẳng đứng của khối lượng treo m_s khi có sự thay đổi các thông số hệ thống và tác động của các loại nhiễu loạn mặt đường, có nghĩa là :

$$\lim_{t\to\infty}\left(z_s(t)\right)\to 0 \quad (5)$$

- Định nghĩa độ lệch tương đối của hệ thống treo (RSD) là độ lệch của hệ thống treo được chuẩn hóa bằng giới hạn không gian dao động tối đa z_{max}. Đại lượng này phải thỏa mãn RSD<1, có thể được biểu thị như sau:

$$\text{RSD}=\frac{\left|z_s-z_u\right|}{z_{max}}<1 \quad (6)$$

- Phải duy trì sự tiếp xúc liên tục giữa lốp xe và mặt đường đồng thời tránh tải trọng quá mức lên lốp. Theo đó, lực tương đối của lốp (RTF), được định nghĩa là tải trọng động của lốp được chuẩn hóa bằng tải trọng tĩnh, phải thỏa mãn RTF<1:

$$\mathrm{RTF} = \frac{k_t\left(z_u - z_r\right) + c_t\left(\dot{z}_u - \dot{z}_r\right)}{\left(m_s + m_u\right)g} < 1 \tag{7}$$

Trước khi tiếp tục quá trình tổng hợp bộ điều khiển, một định nghĩa được sử dụng trong phân tích tiếp theo được trình bày.

Định nghĩa 1. Một hàm trơn $V : R^n \to R^+$ được gọi là hàm Lyapunov ISS cho hệ $\dot{x} = f(x,b)$ với $x \in R^n$ là trạng thái, $d \in R^m$ là tác động hoặc nhiễu đầu vào, $d(t) \in L_\infty$, nếu tồn tại các hàm $\alpha_1,\alpha_2,\alpha_3 \in K_\infty$ và $\sigma \in K$ sao cho:

$$\alpha_1\left(\|x\|\right) \le V\left(x\right) \le \alpha_2\left(\|x\|\right) \tag{8}$$

$$\dot{V}\left(x\right) \le \alpha_3\left(\|x\|\right) + \sigma\left(\|b\|\right) \tag{9}$$

2.2. Tổng hợp luật điều khiển hợp thể với thích nghi

Trong lý thuyết điều khiển hợp thể, tập hợp các tiêu chí hoặc mục tiêu điều khiển được xác định dưới dạng một hệ thống các đa tạp bất biến. Trong nghiên này đầu tiên sẽ tổng hợp đa tạp bất biến dựa trên hệ không có nhiễu δ_s. Khi đó, đa tạp bất biến phải phản ánh các yêu cầu đối với luật điều khiển và đồng thời không được chứa biến trạng thái z_r không thể đo trực tiếp. Việc đo biến z_r không khả thi, còn việc xây dựng bộ quan sát chính xác thì phức tạp bởi tính không thể dự đoán của tác động này. Do đó, không thể đưa yêu cầu (7) vào đa tạp bất biến khi không thể đo z_r. Nói cách khác khi chuyển động của giảm xóc chưa vượt quá phạm vi cho phép, nhiệm vụ ưu tiên hàng đầu là ổn định m_s ở vị trí $z_s=0$, còn khi tiến gần đến giới hạn chuyển vị của giảm xóc, trọng tâm điều khiển cần được dịch chuyển theo hướng đảm bảo thỏa bất đẳng thức (6).

Bổ đề 1: Các trạng thái x_i $(i=1\div3)$ trong hệ thống (4) thoả mãn yêu cầu (5) và (6) khi nằm trên đa tạp sau:

$$x_2 + k_1\left(x_1 + \frac{k_2}{2}\ln\left(\frac{z_{\max} + x_1 - x_3}{z_{\max} - (x_1 - x_3)}\right)\right) = 0 \tag{10}$$

trong đó k_1, k_2 – các hằng số dương khác 0.

Chứng minh. Từ (8) ta có thể biểu diễn $x_1 - x_3$ như sau :

$$x_1 - x_3 = z_{\max}\tanh\left(\frac{x_2 + x_1 k_1}{k_1 k_2}\right) \tag{11}$$

Hàm *tanh* là hàm bị chặn trong khoảng (-1,1), từ (11) suy ra thoả mãn điều kiện (6) điều này có nghĩa là hành trình của giảm chấn không được vượt quá một giá trị cực đại z_{max}. Tiếp tục với k_2 là một hằng số dương rất nhỏ nhưng khác 0, và $\left|x_1 - x_3\right| < z_{\max}$, từ (10) ta có thể viết:

$$x_2 \approx -k_1 x_1 \tag{12}$$

Dễ dạng chỉ ra phương trình vi phân này với $k_1>0$ có nghiệm duy nhất là $x_1=0$ khi $t\to\infty$. Từ lập luận trên, ta thấy khi hệ thống đã nằm trên đa tạp thì hệ thống sẽ tiến dần về 0 thoả mãn

(5). Do phương trình (12) này là tự trị, khi $|x_3| \le z_{\max}$, m_s sẽ không dịch chuyển theo phương thẳng đứng khi có sự thay đổi của mặt đường.

Ta biết rằng trong hệ ban đầu (4) tồn tại nhiễu tham số δ_s, thể hiện ở sự thay đổi của khối lượng treo và ảnh hưởng của mặt đường. Sự thay đổi khối lượng δ_s theo dạng tín hiệu bậc thang từng đoạn và các tham số của hệ không phụ thuộc thời gian, nên để bảo đảm bộ điều khiển có thể thích nghi với những thay đổi tham số, chỉ có thể đưa vào bộ điều khiển một khâu tích phân có dạng :

$$\dot{\nu} = x_2 + k_1\left(x_1 + \frac{k_2}{2}\ln\left(\frac{z_{\max} + x_1 - x_3}{z_{\max} - (x_1 - x_3)}\right)\right) \tag{13}$$

Để tổng hợp luật điều khiển thích nghi, một biến vĩ mô mới có dạng sau:

$$\psi = x_2 + k_1\left(x_1 + \frac{k_2}{2}\ln\left(\frac{z_{\max} + x_1 - x_3}{z_{\max} - (x_1 - x_3)}\right)\right) + c_1\nu \tag{14}$$

trong đó c_1 là một hằng số dương.

Theo quy trình của phương pháp xây dựng phân tích các bộ điều khiển tổng hợp, đa tạp (14) thỏa mãn các phương trình vi phân hàm:

$$\dot{\psi}(t) + T\psi = 0 \tag{15}$$

trong đó T là tham số dương xác định thời gian đi vào giao tuyến của đa tạp ψ=0 từ vị trí hiện tại.

Ta lấy đạo hàm theo thời gian của phương trình (14), có xét đến hệ phương trình vi phân (4), giả thiết hệ không có sự thay đổi khối lượng treo:

$$\dot{\psi} = \dot{x}_2 + k_1\left(\dot{x}_1 + k_2\frac{z_{\max}(\dot{x}_1 - \dot{x}_3)}{z_{\max}^2 - (x_1 - x_3)^2}\right) + c_1\dot{\nu} \tag{16}$$

hay

$$\dot{\psi} = m_s^{-1}\left(-F_d - F_s + u\right) + k_1\left(x_2 + k_2\frac{z_{\max}(x_2 - x_4)}{z_{\max}^2 - (x_1 - x_3)^2}\right) + c_1\dot{\nu} \tag{17}$$

Thế (15) vào phương trình (13), ta thu được biểu thức cho luật điều khiển:

$$u = F_d + F_s - m_s k_1\left(x_2 + k_2\frac{z_{\max}(x_2 - x_4)}{z_{\max}^2 - (x_1 - x_3)^2}\right) - m_s c_1\dot{\nu} - m_s T\psi \tag{18}$$

+ Phân tích tính ổn định

Với luật điều khiển (18) đảm bảo hệ thống (4) hội tụ về đa tạp *ψ=0*. Hay ta có :

$$x_2 + k_1\left(x_1 + \frac{k_2}{2}\ln\left(\frac{z_{\max} + x_1 - x_3}{z_{\max} - (x_1 - x_3)}\right)\right) + c_1\nu = 0 \tag{19}$$

Thế (13) vào (19) dễ dàng chỉ ra ν hội tụ về 0. Thế (16) vào phương trình thứ 2 của hệ (4), khi ψ=0 và ν hội tụ về 0 ta có :

$$\dot{x}_2 + ax_2 - b = 0 \tag{20}$$

Với $a = \frac{m_s k_1}{m_s + \delta_s}\left(1 + \frac{z_{\max}}{z_{\max}^2 - (x_1 - x_3)^2}\right)$; $b = \frac{m_s k_1 k_2}{m_s + \delta_s}\frac{z_{\max} x_4}{z_{\max}^2 - (x_1 - x_3)^2}$

Khi hệ thống nằm trên đa tạp ν=0 theo bổ đề (1) hệ thoả mãn điều kiện (5) và (6). Chọn Lyapunov cho phương trình (20) có dạng :

$$V = \frac{1}{2}x_2^2 \tag{21}$$

Lấy đạo hàm theo thời gian, ta thu được:

$$\dot{V} = x_2\left(-ax_2 + b\right) = -ax_2^2 + bx_2 \tag{22}$$

Từ (22) dựa trên bất đẳng thức Cauchy:

$$\dot{V} \le -\frac{a_{\min}}{2}V + \frac{1}{2a_{\min}}b^2 \tag{23}$$

Rõ ràng V thoả mãn điều kiện (8) và $b \in L_\infty$. Theo **định nghĩa 1** suy ra (20) ổn định từ đầu vào đến trạng thái (ISS). Từ đây suy ra :

$$\left|x_2(t)\right| \le e^{-\frac{a_{\min}}{2}t}x_2(0) + \frac{1}{2a_{\min}}\left\|b\right\|_\infty \tag{24}$$

Kết hợp với **giả sử 1** ta có x_2 bi chặn, nếu $x_4 \to 0$ thì $x_2 \to 0$ và hệ thống nằm trên đa tạp *v=0* nên $x_1 \to 0$.

3. Kết quả và thảo luận

Trong phần này, chúng tôi trình bày kết quả mô phỏng với kịch bản chạy trên mặt đường gờ giảm tốc và sự thay đổi khối lượng treo nhằm chứng minh hiệu quả của luật điều khiển đề xuất. Các tham số của mô hình quarter-car model được liệt kê trong Bảng 1[3], các tham số của bộ điều khiển và các tham số thích nghi được liệt kê trong bảng như sau : $k_1 = 20$; $k_2 = 0.01$; $T = 0.1$; $c_1 = 1$.

Bảng 1. Tham số mô hình quarter-car

Tham số	Giá trị	Đơn vị
m_s	600	kg
m_u	60	kg
k_{1s}	18000	Nm^{-1}
k_{2s}	1000	Nm^{-3}
c_{1s}	2500	$Nm^{-1}s$
c_{2s}	2200	$Nm^{-2}s^2$
c_t	1000	$Nm^{-1}s$
k_t	20000	Nm^{-1}

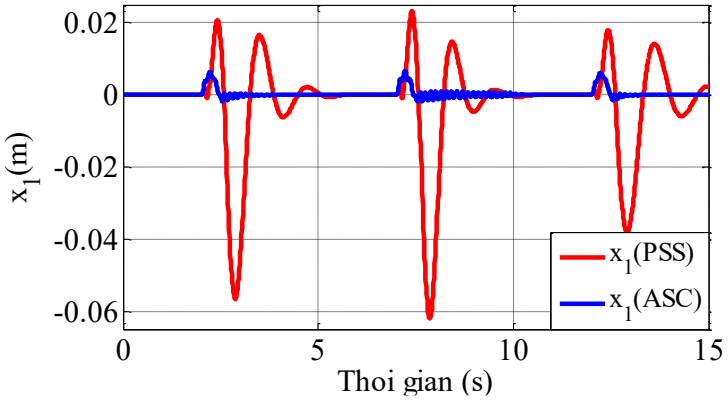
Tín hiệu nhiễu đầu vào được xem là tác động của gờ giảm tốc (bump road input), được cho bởi:

$$z_r(t)=\begin{cases}0.05\sin\left(2\pi(t-2)\right), \text{ khi } 2\le t\le 2.5\\ 0.05\sin\left(2\pi(t-7)\right), \text{ khi } 7\le t\le 7.5\\ 0.05\sin\left(2\pi(t-12)\right), \text{ khi } 12\le t\le 12.5\\ 0, \ \forall t\notin[2;2.5]\cup[7;7.5]\cup[12;12.5]\end{cases} \quad (25)$$

Tín hiệu nhiễu δ_s được cho bởi:

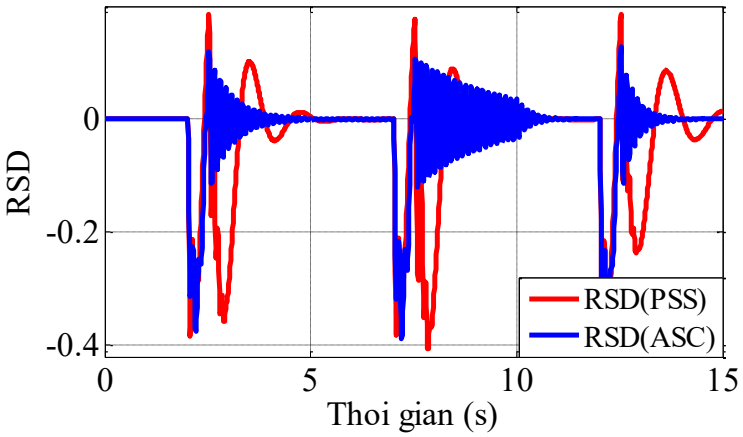
$$\delta_s=\begin{cases}0, \ \forall t\notin[0;5)\\ 120, \ \forall t\notin[5;10)\\ -180, \ \forall t\notin[10;15]\end{cases} \quad (26)$$

Kết quả mô phỏng đáp ứng của hệ thống treo thụ động (PSS) và hệ thống treo chủ động với bộ điều khiển hợp thể thích nghi (ASC) được thể hiện lần lượt trên các Hình 2 ÷ 5. Hình 2 so sánh chuyển vị khối lượng treo z_s theo thời gian. Từ kết quả thu được có thể thấy cả hai hệ thống đều đảm bảo ổn định khối lượng treo ở vị trí cân bằng. Trong khi PSS dao động với biên độ lớn (cực đại *–6 cm*) và dao động tắt dần chậm (khoảng *3 s*), bộ điều khiển hợp thể thích nghi ASC giúp triệt tiêu nhanh dao động khối lượng treo (khoảng *0.5 s*), biên độ cực đại chỉ còn khoảng *0.5 cm* – giảm hơn *90%* so với PSS. Điều này khẳng định khả năng nâng cao độ êm dịu chuyển động của bộ điều khiển.



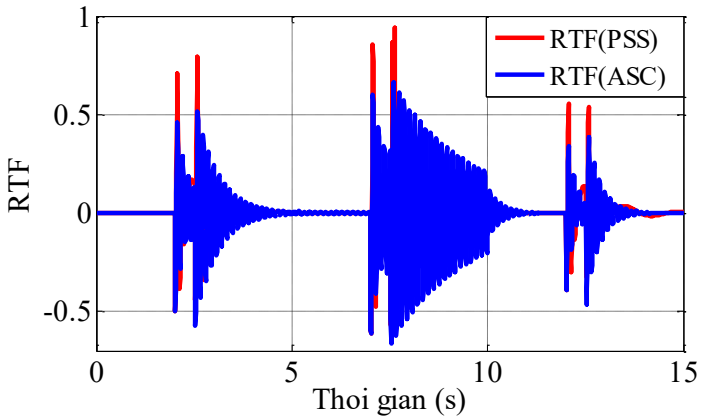
Hình 2. So sánh chuyển vị khối lượng treo (z_s)

Hình 3 thể hiện độ lệch tương đối của hệ thống treo (RSD). Với cả hai hệ thống treo, RSD được duy trì trong giới hạn cho phép (|RSD|<1), biên độ hành trình treo cực đại không có nhiều khác biệt. Độ lệch hệ thống treo ASC giảm biên độ nhanh hơn đồng thời thời gian xác lập ngắn hơn tuy nhiên dao động nhiều hơn so với PSS. Đây là sự đánh đổi hợp lý: để đạt được độ êm dịu cao, hệ thống chủ động chấp nhận điều khiển hành trình treo nhiều hơn nhằm hấp thụ năng lượng từ nhiễu.



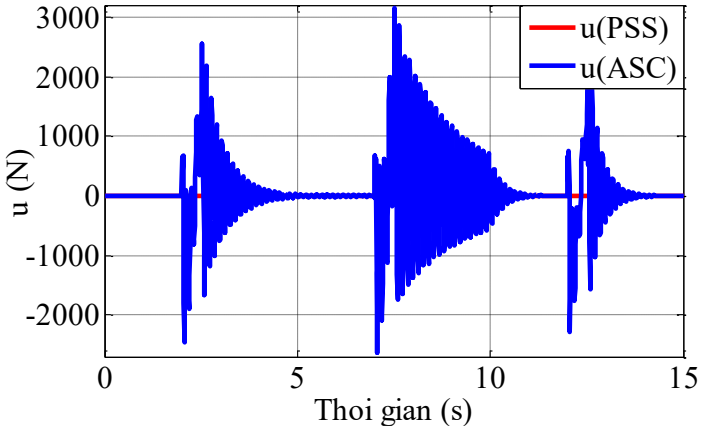
Hình 3. So sánh độ lệch tương đối của hệ thống treo (RSD)

Khả năng bám đường được đánh giá qua chỉ số lực tương đối của lốp (RTF) trên Hình 4. Cả hai hệ thống đều đảm bảo RTF luôn nhỏ hơn *1*, nghĩa là bánh xe không bị nhấc khỏi mặt đường. Tuy nhiên, với bộ điều khiển ASC, RTF có biên độ nhỏ hơn khoảng *30 %*, cho thấy lực tiếp xúc giữa lốp và đường ổn định hơn, từ đó cải thiện an toàn chuyển động.



Hình 4. Đáp ứng lực tương đối của lốp (RTF)

Tín hiệu điều khiển u(t) được biểu diễn trên Hình 5. Tại các thời điểm gặp gờ giảm tốc, lực điều khiển xuất hiện các xung lớn (cực đại khoảng *3000 N*) để nhanh chóng dập tắt dao động thân xe. Khi khối lượng treo thay đổi (tại *5 s* và *10 s*), bộ điều khiển vẫn duy trì ổn định nhờ cơ chế thích nghi tích phân, không xảy ra hiện tượng mất ổn định hay dao động ký sinh. So với các bộ điều khiển khác, luật hợp thể thích nghi không yêu cầu ước lượng nhiễu hay khối lượng, do đó dễ triển khai thực tế.



Hình 5. Tín hiệu điều khiển (u)

Từ những phân tích ở trên cho thấy bộ điều khiển đề xuất đã đáp ứng tốt ba mục tiêu chính của hệ thống treo chủ động đã trình ở trên. Kết quả mô phỏng khẳng định tính hiệu quả và ổn định của phương pháp trước cả nhiễu mặt đường và sự thay đổi khối lượng.

4. Kết luận

Bài báo đã đề xuất một bộ điều khiển hợp thể thích nghi cho hệ thống treo chủ động trên cơ sở lý thuyết điều khiển hợp thể kết hợp với cơ chế thích nghi tích phân. Luật điều khiển cho phép giảm hiệu quả biên độ dao động của khối lượng treo, bù được ảnh hưởng của biến thiên tham số mà không cần ước lượng trực tiếp nhiễu, đồng thời bảo đảm hành trình treo nằm trong giới hạn cơ khí. Kết quả mô phỏng cho thấy bộ điều khiển vượt trội so với hệ thống treo thụ động: biên độ dao động của khối lượng treo giảm trên 90%, trong khi các ràng buộc về độ lệch treo và tải trọng lốp vẫn được bảo đảm. Tuy nhiên, nghiên cứu hiện mới dừng ở mô phỏng trên mô hình quarter-car và chưa xét đến trễ cơ cấu chấp hành hoặc nhiễu đo lường. Hướng phát triển tiếp theo bao gồm tối ưu hóa tham số, sử dụng các bộ quan sát và các luật điều khiển thông minh cho mô hình toàn xe và thử nghiệm trên hệ thống thực để đánh giá tính khả thi trong điều kiện thực tế.

Tài liệu tham khảo

1. Sun, W., Gao, H., & Shi, P. (2020). *Advanced control for vehicle active suspension systems* (1st ed.). Springer Cham. https://doi.org/10.1007/978-3-030-15785-2
2. Kashem, S., Nagarajah, R., & Ektesabi, M. (2018). *Vehicle suspension systems and electromagnetic dampers*. Singapore: Springer. https://doi.org/10.1007/978-981-10-5478-5
3. Sun, W., Pan, H., Zhang, Y., & Gao, H. (2013). Multi-objective control for uncertain nonlinear active suspension systems. Mechatronics, 23(8), 1136–1145. https://doi.org/10.1016/j.mechatronics.2013.06.004
4. Muderrisoğlu, K., Arisoy, D. O., Ahan, A. O., & Akdoğan, E. (2016). *PID parameters prediction using neural network for a linear quarter car suspension control*. *International Journal of Intelligent Systems and Applications in Engineering*, 4(1), 20–24.
5. Prassad, G. S., & Mohan, M. K. (2019). A contemporary adaptive air suspension using LQR control for passenger vehicles. *ISA Transactions, 93*, 244–254. https://doi.org/10.1016/j.isatra.2019.02.031
6. Liu, L., Li, X., Liu, Y.-J., & Tong, S. (2021). Neural network based adaptive event trigger control for a class of electromagnetic suspension systems. Control Engineering Practice, 106, 104675. https://doi.org/10.1016/j.conengprac.2020.104675

7. Pang, H., Zhang, X., Yang, J., & Shang, Y. (2019). Adaptive backstepping-based control design for uncertain nonlinear active suspension system with input delay. International Journal of Robust and Nonlinear Control, 29(17), 1–20. https://doi.org/10.1002/rnc.4695
8. Nguyen, T. A. (2021). *Study on the sliding mode control method for the active suspension system*. International Journal of Applied Science and Engineering, 18(5). doi: 10.6703/IJASE.202109_18(5).006
9. Özbek, C., Burkan, R., & Yagiz, N. (2024). *Robust control of vehicle suspension systems with friction non-linearity for ride comfort enhancement*. *Journal of Vibration and Control 30(5-6):1326-1338*.
10. Bajaba, N. S. (2023). *Ellipsoidal-based robust control for vehicle active suspension systems under load uncertainties and network-induced interruptions*. *Reports in Mechanical Engineering*. Vo l . 6, No. 1, 2025, pp. 46-54.
11. Yin, Y., Luo, B., Ren, H., & Zhang, Y. (2022). Robust control design for active suspension system with uncertain dynamics and actuator time delay. Journal of Mechanical Science and Technology, 36(12), 6319–6327. https://doi.org/10.1007/s12206-022-1143-1
12. Abut, T., Salkim, E., & Tugal, H. (2025). Active suspension control for improved ride comfort and vehicle performance using HHO-based Type-I and Type-II fuzzy logic. Biomimetics, 10(10), 673. https://doi.org/10.3390/biomimetics10100673
13. Kimball, J. B., et al. (2024). Adaptive control and reinforcement learning for vehicle active suspension. *Control Engineering Practice. Volume 58, 2024, 100974*
14. Zeng, Q., & Zhao, J. (2024). Adaptive switching control of active suspension systems: A switched system point of view. IEEE Transactions on Control Systems Technology, 32(2), 663–671. https://doi.org/10.1109/TCST.2023.3318350
15. Deshpande, V. S., Shendge, P. D., & Phadke, S. B. (2017). Nonlinear control for dual objective active suspension systems. *IEEE Transactions on Intelligent Transportation Systems, 18*(3), 656–665. https://doi.org/10.1109/TITS.2016.2585343
16. Tee K, Ge S, Tay E. (2009) Barrier Lyapunov functions for the control of output constrained nonlinear systems// Automatica. , –45, –P. 918–927.
17. Sinitsyn A. S. Synergetics control synthesis with taking into account physical limitations of active suspension. Proceedings of the 7th All-Russian Scientific Conference “System Synthesis and Applied Synergetics” (SSAS-2015). Taganrog: Southern Federal University Publishing House, 2015. 395 p.
18. Kolesnikov, A. A., Veselov, G. E., Popov, A. N., & others. (2006). Synergetic methods of control of complex systems: Mechanical and electromechanical systems (A. A. Kolesnikov, Ed.). Moscow: KomKniga.
19. Nguyen, X. C., & Le, D. T. (2025). *Adaptive finite-time synergetic control for flexible-joint robot manipulator with disturbance inputs*. *Electrical Engineering & Electromechanics*, *(3)*, 45–52. https://doi.org/10.20998/2074-272X.2025.3.07
20. Nguyen X.C., & Thang, L. T. (2024). Synthesis of an orbit tracking controller for a 2DOF helicopter based on sequential manifolds with stabilization time in the presence of disturbances. Engineering, Technology & Applied Science Research, 14(4), 15083–15089.
21. Nguyen, C. X., Tran, S. V., & Phan, H. N. (2023). *Control law synthesis for flexible joint manipulator based on synergetic control theory*. *Mekhatronika, Avtomatizatsiya, Upravlenie*, *24*(8), 395–402.

Prescribed-Time Gain-Scheduled Computed-Torque Control with Residual Disturbance Observation for Robot Manipulators

Hoàng Đức Long[1*]

[1]*Viện Tên lửa và Kỹ thuật điều khiển; Học viện Kỹ thuật quân sự*

Email; Tel: longhd@lqdtu.edu.vn; 0915961451

Abstract

This paper presents a new control method for robotic manipulators that achieves fast and predictable tracking performance within a user-defined time. The method uses a gain scheduling strategy in which the control gains increase smoothly as the system approaches the target time, allowing faster error reduction while maintaining numerical stability. A disturbance observer is added to estimate unmodeled dynamics and external disturbances in real time, helping the controller remain robust even when the robot is affected by uncertain or time-varying forces. Theoretical analysis shows that the tracking error converges within the predefined time in the ideal case and remains uniformly bounded under time-varying disturbances. Numerical simulations on a two-link robot demonstrate that the proposed controller outperforms standard computed torque control with fixed gains and also performs better when combined with disturbance compensation. The results confirm that the method achieves high accuracy, strong robustness, and predictable transient behavior, making it suitable for applications requiring rapid and reliable motion tracking.

Keywords: *Prescribed-time control; gain-scheduled; computed torque control; disturbance observation; robotic manipulators; robust tracking trajectory*

1. Introduction

High-precision trajectory tracking for robotic manipulators under uncertainties and disturbances remains a core topic in nonlinear control [1-3]. Classical solutions include PID and computed-torque schemes, often augmented with robust or adaptive terms, but they typically guarantee only asymptotic convergence and do not control the convergence time explicitly [4-6]. Recent work moves toward finite-time, fixed-time, predefined-time, and prescribed-time control to provide explicit transient guarantees for manipulators and other complex systems [7-8].

Finite-/fixed-time sliding-mode and terminal sliding-mode controllers have been widely studied for robot manipulators, often combined with disturbance observers (DOs) or neural approximators to enhance robustness [9, 10]. Fixed-time sliding-mode designs for manipulators with disturbance observers and chattering reduction are reported in some previous studies [11, 12], where a fixed-time sliding mode control (FTSM) with fixed-time disturbance observer (FTDO) is proposed for two-link manipulators. Similar DO-based fixed-time schemes are developed for nonlinear systems with mismatched uncertainties and for robotic manipulators with global fast convergence.

In parallel, prescribed-performance and prescribed-time methods have emerged. Prescribed-performance adaptive robust control and barrier Lyapunov-based approaches have been applied to manipulators under constraints [13-15]. Lyapunov- and CLF-based prescribed-time stabilization has been studied for general nonlinear, stochastic, and quantum systems,

[*] Email: longhd@lqdtu.edu.vn

providing constructive design tools and characterizations of prescribed-time stability. Prescribed-time and predefined-time schemes have also been investigated for fault-tolerant or consensus control in multi-robot settings.

For manipulators, several recent works combine prescribed-time or prescribed-performance with sliding-mode or neural control, including adaptive prescribed-time or prescribed-performance sliding-mode controllers and observer-based prescribed-time designs [9-15]. However, most existing robot-oriented prescribed-time schemes either (i) rely on nonlinear sliding surfaces with non-Lipschitz terms, or (ii) employ complex backstepping/neuro-adaptive structures, which complicate implementation and tuning.

To overcome the limitations of existing prescribed-time and sliding-mode approaches, this paper proposes a prescribed-time gain-scheduled computed-torque controller with residual disturbance observation for robotic manipulators. The controller embeds a simple time-varying gain law into a PD computed-torque structure to ensure convergence within a user-defined time. A residual-based disturbance observer enhances robustness against constant and time-varying disturbances. Lyapunov analysis proves prescribed-time attractivity in the nominal case and robust boundedness under disturbances. Simulation results on a two-link manipulator validate the effectiveness and improved transient performance of the proposed method.

2. Problem Formulation and Preliminaries

Consider an *n*-DOF rigid robotic manipulator with dynamics [11, 12]

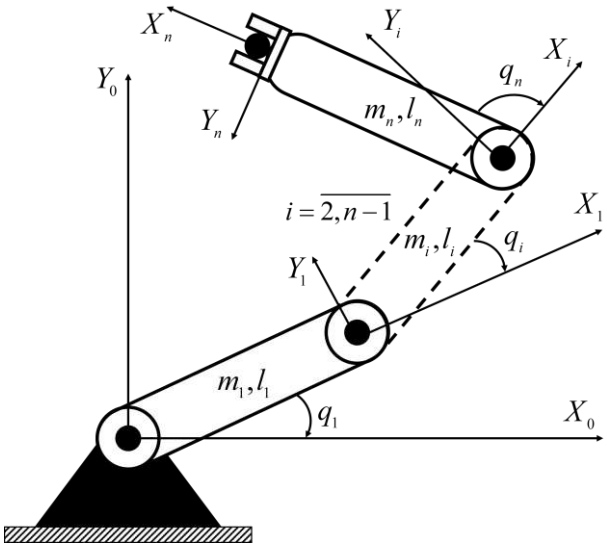


Figure 1. n-DOF robotic manipulator

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q},\dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{G}(\mathbf{q}) = \boldsymbol{\tau} + \mathbf{d}(t) \tag{1}$$

where $\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}} \in \mathbb{R}^n$ are joint position, velocity, and acceleration vectors; $\mathbf{M}(\mathbf{q}) \in \mathbb{R}^{n \times n}$ is the symmetric positive definite inertia matrix; $\mathbf{C}(\mathbf{q},\dot{\mathbf{q}})\dot{\mathbf{q}} \in \mathbb{R}^n$ contains Coriolis/centrifugal terms; $\mathbf{G}(\mathbf{q}) \in \mathbb{R}^n$ is the gravity vector; $\boldsymbol{\tau} \in \mathbb{R}^n$ is the control torque; $\mathbf{d}(t) \in \mathbb{R}^n$ lumps external disturbances and model uncertainties.

We adopt the standard assumptions:

Assumption 1. There exist constants $m_{\min}, m_{\max} > 0$ such that

$$m_{\min}\|\mathbf{x}\|^2 \le \mathbf{x}^T\mathbf{M}(\mathbf{q})\mathbf{x} \le m_{\max}\|\mathbf{x}\|^2, \quad \forall \mathbf{x}, \mathbf{q} \tag{2}$$

Assumption 2. The lumped disturbance and its derivative are bounded:

$$\|\mathbf{d}(t)\| \le d_{\max}, \quad \|\dot{\mathbf{d}}(t)\| \le \dot{d}_{\max}, \quad \forall t \ge 0 \tag{3}$$

Let $\mathbf{q}_d(t)$ be a smooth desired trajectory with derivatives $\dot{\mathbf{q}}_d(t)$, $\ddot{\mathbf{q}}_d(t)$ available. Define tracking errors

$$\mathbf{e} := \mathbf{q} - \mathbf{q}_d, \quad \dot{\mathbf{e}} = \dot{\mathbf{q}} - \dot{\mathbf{q}}_d \tag{4}$$

Our objective is to design $\boldsymbol{\tau}(t)$ such that $\mathbf{e}(t)$ is prescribed-time attractive at a user-selected time $T_p > 0$, while maintaining robustness to $\mathbf{d}(t)$.

3. Prescribed-Time Gain-Scheduled Computed-Torque Control

We propose PD gains of the form

$$\mathbf{K}_p(t) = \frac{k_p}{(T_p - t)^r}\mathbf{I}_n, \quad \mathbf{K}_d(t) = \frac{k_d}{(T_p - t)^r}\mathbf{I}_n, \quad 0 \le t < T_p \tag{5}$$

where $k_p > 0$, $k_d > 0$ and $r \ge 1$ are design constants, and $\mathbf{I}_n$ is the $n \times n$ identity matrix.

The prescribed-time gain-scheduled computed-torque controller is

$$\boldsymbol{\tau} = \mathbf{M}(\mathbf{q})\left(\ddot{\mathbf{q}}_d - \mathbf{K}_d(t)\dot{\mathbf{e}} - \mathbf{K}_p(t)\mathbf{e}\right) + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{G}(\mathbf{q}) - \hat{\mathbf{d}} \tag{6}$$

where $\hat{\mathbf{d}}(t)$ is the output of the disturbance observer.

Substituting (6) into the dynamics (1) yields

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{G}(\mathbf{q}) = \mathbf{M}(\mathbf{q})\left(\ddot{\mathbf{q}}_d - \mathbf{K}_d\dot{\mathbf{e}} - \mathbf{K}_p\mathbf{e}\right) + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{G}(\mathbf{q}) - \hat{\mathbf{d}} + \mathbf{d}$$

so that

$$\ddot{\mathbf{q}} = \ddot{\mathbf{q}}_d - \mathbf{K}_d(t)\dot{\mathbf{e}} - \mathbf{K}_p(t)\mathbf{e} + \mathbf{M}^{-1}(\mathbf{q})\tilde{\mathbf{d}} \tag{7}$$

where $\tilde{\mathbf{d}} = \mathbf{d} - \hat{\mathbf{d}}$.

Using (4), the error dynamics become

$$\ddot{\mathbf{e}} = -\mathbf{K}_d(t)\dot{\mathbf{e}} - \mathbf{K}_p(t)\mathbf{e} + \mathbf{M}^{-1}(\mathbf{q})\tilde{\mathbf{d}} \tag{8}$$

Set $\tilde{\mathbf{d}} \equiv 0$. The nominal error dynamics are

$$\ddot{\mathbf{e}} + \mathbf{K}_d(t)\dot{\mathbf{e}} + \mathbf{K}_p(t)\mathbf{e} = 0 \tag{9}$$

Theorem 1 (Nominal prescribed-time attractivity). Consider the nominal error dynamics (9) and PD gains (5). Define the extended error vector $\boldsymbol{\xi}(t) := \left[\mathbf{e}^T(t), \dot{\mathbf{e}}^T(t)\right]^T \in \mathbb{R}^{2n}$. Then, for any initial condition $\boldsymbol{\xi}(0)$, the solution of (9) exists on $[0, T_p)$ and satisfies

$$\lim_{t \to T_p^-} \boldsymbol{\xi}(t) = 0 \tag{10}$$

Proof

Consider the quadratic Lyapunov function

$$V_e(t) = \frac{1}{2}\left(\|\mathbf{e}\|^2 + \|\dot{\mathbf{e}}\|^2\right) = \frac{1}{2}\|\xi\|^2 \tag{11}$$

Differentiating $V_e(t)$ gives

$$\dot{V}_e(t) = \mathbf{e}^T(t)\dot{\mathbf{e}}(t) + \dot{\mathbf{e}}^T(t)\ddot{\mathbf{e}}(t) \tag{12}$$

Define the exponentially scaled Lyapunov function

$$W(t) := e^{-c_0 t} V_e(t) \tag{13}$$

Differentiating,

$$\dot{W}(t) = e^{-c_0 t}\dot{V}_e(t) - c_0 e^{-c_0 t} V_e(t)$$

$$\leq e^{-c_0 t}\left(c_0 V_e(t) - \frac{c_1}{(T_p - t)^r} V_e(t)\right) - c_0 e^{-c_0 t} V_e(t) \tag{14}$$

so, the $c_0 V_e(t)$ terms cancel and we obtain

$$\dot{W}(t) \leq -\frac{c_1}{(T_p - t)^r} W(t), \quad 0 \leq t < T_p \tag{15}$$

This is a scalar linear differential inequality of the form

$$\dot{W}(t) \leq -a(t) W(t), \quad a(t) := \frac{c_1}{(T_p - t)^r} > 0 \tag{16}$$

Therefore from (16) we have

$$\lim_{t \to T_p^-} W(t) = 0 \tag{17}$$

It follows that

$$\lim_{t \to T_p^-} V_e(t) = 0 \tag{18}$$

Because $V_e(t) = \frac{1}{2}\left(\|\mathbf{e}\|^2 + \|\dot{\mathbf{e}}\|^2\right)$, this implies

$$\lim_{t \to T_p^-} \|\mathbf{e}(t)\| = 0, \quad \lim_{t \to T_p^-} \|\dot{\mathbf{e}}(t)\| = 0 \tag{19}$$

Equivalently,

$$\lim_{t \to T_p^-} \xi(t) = 0 \tag{20}$$

■

4. Residual-Based Disturbance Observer

In continuous time, define the residual

$$\mathbf{r}(t) := \mathbf{M}(\mathbf{q})\ddot{\mathbf{q}}(t) - \left(\boldsymbol{\tau}(t) - \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}}(t) - \mathbf{G}(\mathbf{q})\right) \tag{21}$$

Using (1), we have

$$\mathbf{r}(t) = \mathbf{d}(t) \tag{22}$$

We propose the first-order residual disturbance observer

$$\dot{\hat{\mathbf{d}}}(t) = k_0\left(\mathbf{r}(t) - \hat{\mathbf{d}}(t)\right) = k_0\left(\mathbf{d}(t) - \hat{\mathbf{d}}(t)\right), \quad k_0 > 0 \tag{23}$$

Then

$$\dot{\tilde{\mathbf{d}}}(t) = \dot{\mathbf{d}}(t) - \dot{\hat{\mathbf{d}}}(t) = \dot{\mathbf{d}}(t) - k_0\left(\mathbf{d}(t) - \hat{\mathbf{d}}(t)\right) = \dot{\mathbf{d}}(t) - k_0\tilde{\mathbf{d}}(t) \tag{24}$$

Lemma 1 (Constant disturbances).

Suppose the lumped disturbance is constant

$$\mathbf{d}(t) \equiv \mathbf{d}_0, \quad \mathbf{d}_0 \in \mathbb{R}^n \tag{25}$$

Then the observer error $\tilde{\mathbf{d}}(t)$ evolves according to

$$\dot{\tilde{\mathbf{d}}}(t) = -k_0\tilde{\mathbf{d}}(t) \tag{26}$$

and thus satisfies

$$\tilde{\mathbf{d}}(t) = e^{-k_0 t}\tilde{\mathbf{d}}(0), \quad t \geq 0 \tag{27}$$

In particular,

$$\lim_{t\to\infty} \left\|\tilde{\mathbf{d}}(t)\right\| = 0 \tag{28}$$

and the convergence is exponential with rate k_0.

Lemma 2 (Bounded-rate disturbances).

Under Assumption 2, the disturbance observer error $\tilde{\mathbf{d}}(t)$ satisfies

$$\left\|\tilde{\mathbf{d}}(t)\right\| \leq e^{-k_0 t}\left\|\tilde{\mathbf{d}}(0)\right\| + \frac{\dot{d}_{\max}}{k_0}\left(1 - e^{-k_0 t}\right), \quad \forall t \geq 0 \tag{29}$$

In particular, $\left\|\tilde{\mathbf{d}}(t)\right\|$ is uniformly ultimately bounded by

$$\limsup_{t\to\infty} \left\|\tilde{\mathbf{d}}(t)\right\| \leq \frac{\dot{d}_{\max}}{k_0} \tag{30}$$

5. Composite Closed-Loop Analysis

Using (8) and (24), the coupled system in variables $(\mathbf{e}, \dot{\mathbf{e}}, \tilde{\mathbf{d}})$ is

$$\begin{cases} \ddot{\mathbf{e}}(t) = -\mathbf{K}_d(t)\dot{\mathbf{e}}(t) - \mathbf{K}_p(t)\mathbf{e}(t) + \mathbf{M}^{-1}(\mathbf{q})\tilde{\mathbf{d}}(t) \\ \dot{\tilde{\mathbf{d}}}(t) = \dot{\mathbf{d}}(t) - k_0\tilde{\mathbf{d}}(t) \end{cases} \tag{31}$$

Theorem 2 (Robust prescribed-time boundedness).

Under Assumptions 1-2, the control law (6) together with the observer (13) guarantees that there exist positive constants C_1, C_2, a such that

$$V_c(t) = \frac{1}{2}\left(\|\mathbf{e}\|^2 + \|\dot{\mathbf{e}}\|^2\right) + \frac{1}{2g}\|\tilde{\mathbf{d}}\|^2 \le C_1\left(1 - \frac{t}{T_p}\right)^a + C_2, \quad 0 \le t \le T_p,\ g > 0 \tag{32}$$

where $C_2 = \frac{\dot{d}_{\max}}{k_0^2}$. In particular, if $\mathbf{d}(t) \to \mathbf{d}_0$ is constant, then $\dot{d}_{\max} = 0$ and $C_2 = 0$, and $\lim_{t \to T_p}(\mathbf{e}(t), \dot{\mathbf{e}}(t), \tilde{\mathbf{d}}(t)) = (0,0,0)$ or the equilibrium $(\mathbf{e}, \dot{\mathbf{e}}, \tilde{\mathbf{d}}) = (0,0,0)$ is prescribed-time attractive at T_p; for general bounded-rate disturbances, the state remains uniformly bounded on $[0, T_p)$, and the tracking error $(\mathbf{e}, \dot{\mathbf{e}})$ is driven into a small neighborhood of the origin whose size decreases as k_0 increases.

Proof

Consider the Lyapunov candidate

$$V_c = V_e + V_d \tag{33}$$

where $V_e = \frac{1}{2}\left(\|\mathbf{e}\|^2 + \|\dot{\mathbf{e}}\|^2\right)$; $V_d = \frac{1}{2g}\|\tilde{\mathbf{d}}\|^2$, $g > 0$

For V_e,

$$\dot{V}_e \le -\frac{2c}{(T_p - t)^r} V_e + \dot{\mathbf{e}}^T \mathbf{M}^{-1}(\mathbf{q})\tilde{\mathbf{d}} \tag{34}$$

For V_d,

$$\dot{V}_d \le -\frac{k_0}{2g}\|\tilde{\mathbf{d}}\|^2 + \frac{\dot{d}_{\max}}{2k_0 g} \tag{35}$$

Thus inequality (33) gives

$$\dot{V}_c \le -\frac{2c}{(T_p - t)^r} V_e + \frac{\dot{d}_{\max}}{2k_0 g} \le \frac{\dot{d}_{\max}}{2k_0 g}, \quad t^* \le t < T_p \tag{36}$$

We deduce that there exist constants $C_1 > 0$ and $C_2 > 0$ such that

$$V_c(t) \le C_1\left(1 - \frac{t}{T_p}\right)^a + C_2, \quad 0 \le t \le T_p \tag{37}$$

where C_2 depends on $U_{\max}$ and $C_2 = \frac{\dot{d}_{\max}}{k_0^2}$.

If the disturbance is constant, $\mathbf{d}(t) \to \mathbf{d}_0$, then $\dot{d}_{\max} = 0$. From Lemma 1, $\tilde{\mathbf{d}}(t) \to 0$ exponentially, and in particular $\tilde{\mathbf{d}}(t) \to 0$ can be made arbitrarily small after some finite time. In the ideal limit, the system reduces to the nominal case of Theorem 1, and $C_2 = 0$ in (37), yielding

$$V_c(t) \le C_1\left(1-\frac{t}{T_p}\right)^{a}, \quad \lim_{t\to T_p^-} V_c(t)=0 \tag{38}$$

which implies $\lim_{t\to T_p^-}\left(\mathbf{e}(t),\dot{\mathbf{e}}(t),\dot{\mathbf{d}}(t)\right)=(0,0,0)$. This proves robust prescribed-time boundedness in the general bounded-rate case, and prescribed-time attractivity in the constant-disturbance case.

■

6. Simulation Results

We consider the standard two-link planar manipulator used in prior works [11]. The dynamics have the form (1) with $\mathbf{q}=[q_1,q_2]^T$, and the inertia, Coriolis, and gravity terms given by well-known expressions. The parameters are $m_1=2(kg)$, $m_2=0.85(kg)$, $l_1=0.35(m)$, $l_2=0.31(m)$, $l_{c1}=0.175(m)$, $l_{c2}=0.155(m)$, $I_1=6.1\times10^{-4}(kg.m^2)$, $I_2=2\times10^{-4}(kg.m^2)$, and gravitational acceleration $g=9.81(m/s^2)$. The desired trajectories are $q_{d1}(t)=0.14\sin 0.5t$ and $q_{d2}(t)=0.14\cos 0.5t$ with derivatives computed analytically, as in the given code. Initial conditions are $q(0)=[0.1,\ 0.2]^T$ and $\dot{q}(0)=[0,\ 0]^T$. The lumped disturbance is chosen as $d(t)=[0.3\sin(18t)+0.12\sin(45t),\ 0.15\cos(20t)+0.08\sin(52t)]^T$. Simulation horizon is $T_p=2(s)$.

Figures 2 and 3 illustrates the joint position tracking performance under the proposed prescribed-time gain-scheduled computed-torque controller and the constant-gain computed-torque controller with disturbance observation. It is observed that the proposed method achieves significantly faster transient response and smaller tracking errors as the system approaches the prescribed convergence time. Figures 4 and 5 compare the true disturbances and their estimates obtained from the residual-based disturbance observer. For both controllers, the observer provides accurate disturbance estimation after a short transient. The effectiveness of the observer enables the proposed controller to compensate for disturbances without introducing chattering.

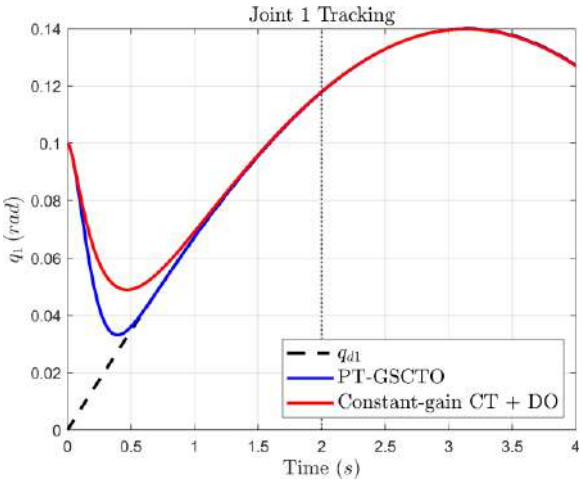


Figure 2. Tracking angle of Joint 1

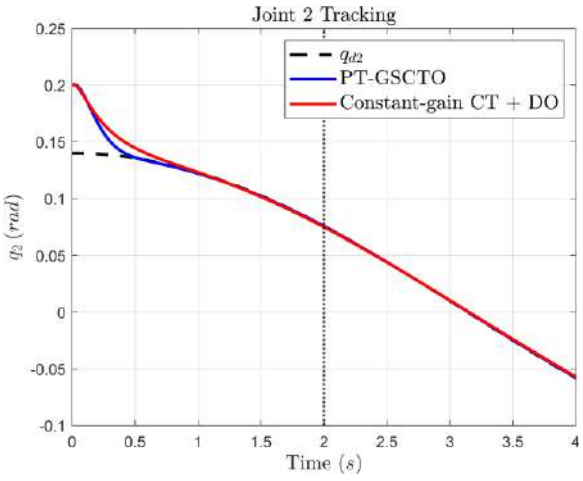


Figure 3. Tracking angle of Joint 2

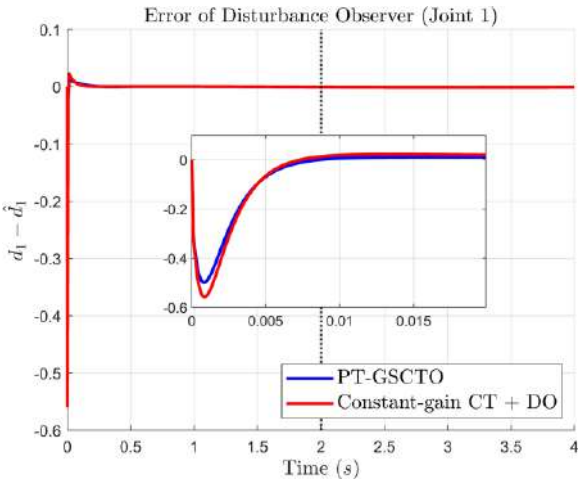


Figure 4. Error of disturbance observer - Joint 1

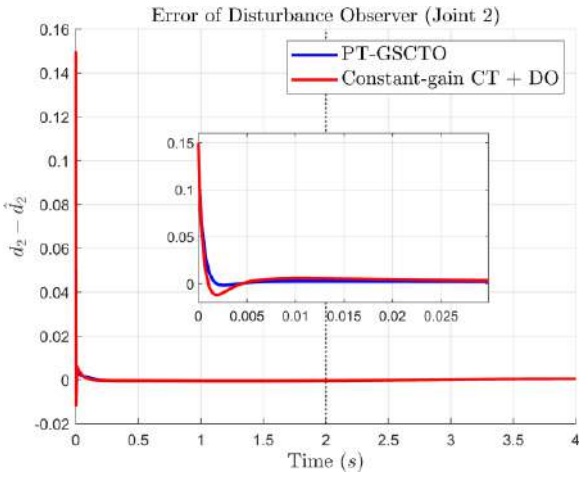


Figure 5. Error of disturbance observer - Joint 2

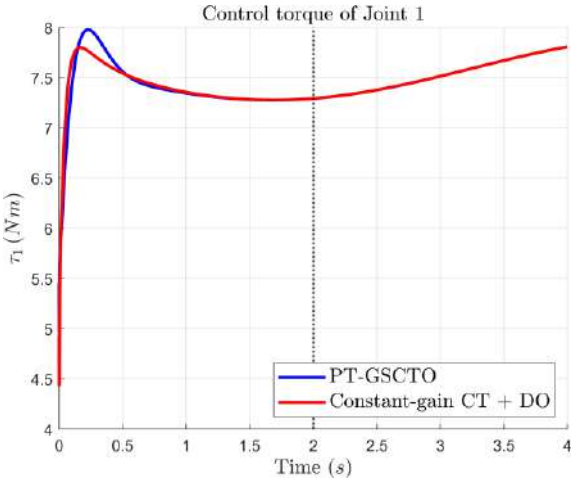


Figure 6. Control torque of Joint 1

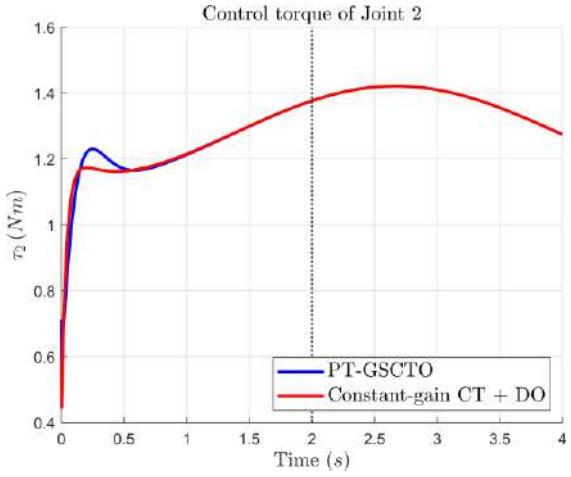


Figure 7. Control torque of Joint 2

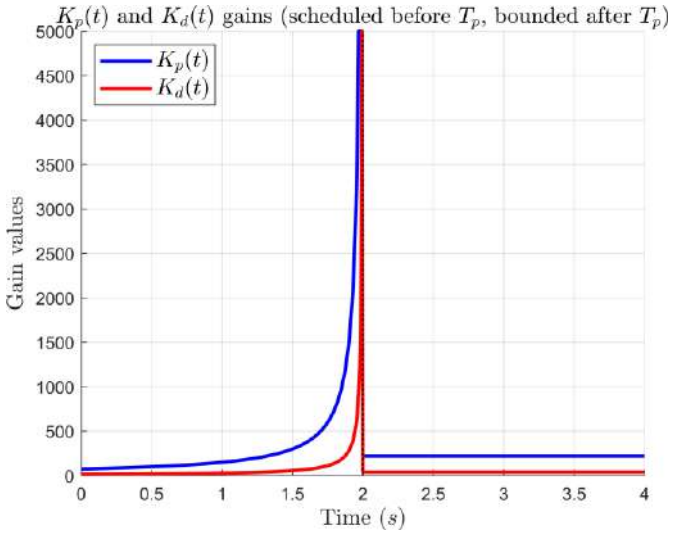


Figure 8. The coefficients $K_d(t)$ and $K_p(t)$

Figures 6 and 7 show the control torque applied at two joints for the compared controllers. Although the PT-GSCTO controller employs increasing gains near the prescribed time, the control inputs remain bounded due to the regularized gain scheduling and saturation mechanism. Compared to the FTSM+FTDO controller, the proposed method achieves improved tracking performance without excessive control effort, validating the practical feasibility of the control design. The gain values $K_d(t)$ and $K_p(t)$ are illustrated in Fig. 8.

Overall, the composite analysis and simulation results jointly confirm that the proposed control scheme achieves prescribed-time tracking performance in the nominal case and maintains robust boundedness in the presence of time-varying disturbances.

7. Conclusion

This paper introduced a prescribed-time gain-scheduled computed-torque controller with residual disturbance observation (PT-GSCTO) for robotic manipulators. The key idea is to embed a simple analytic gain law $K(t)$ μ $(T_p - t)^{-r}$ into a computed-torque PD structure, yielding a Lyapunov function whose decay rate grows unbounded as t ® T_p^- . We proved that, in the disturbance-free case, the tracking error is prescribed-time attractive at T_p . A residual-based disturbance observer was then designed and analyzed; together with the controller, it guarantees robust prescribed-time attractivity for constant disturbances and practical prescribed-time performance for slowly varying disturbances.

References

1. Liu, J., Dong, X., Yang, Y., & Chen, H. (2021). Trajectory Tracking Control for Uncertain Robot Manipulators with Repetitive Motions in Task Space. *Mathematical Problems in Engineering, 2021*(1), DOI: 10.1155/2021/8838927.
2. Ren, S., Han, L., Mao, J., & Li, J. (2023). Optimized Trajectory Tracking for Robot Manipulators with Uncertain Dynamics: A Composite Position Predictive Control Approach. *Electronics, 12*(21), DOI: 10.3390/electronics12214548.
3. Li, Y., Xu, Z., Yang, X., Zhao, Z., Zhuang, L., Zhao, J., & Liu, H. (2024). Identification and high-precision trajectory tracking control for space robotic manipulator. *Acta Astronautica, 214*, 484-495, DOI: 10.1016/j.actaastro.2023.11.001.
4. Ghediri, A., Lamamra, K., Kaki, A. A., & Vaidyanathan, S. (2022). Adaptive PID computed-torque control of robot manipulators based on DDPG reinforcement learning. *International Journal of Modelling, Identification and Control, 41*(3), 173-182, DOI: 10.1504/IJMIC.2022.127518.
5. Zhou, Y., He, X., Shao, F., & Zhang, X. (2024). Research on the Optimization of the PID Control Method for an EOD Robotic Manipulator Using the PSO Algorithm for BP Neural Networks. *Actuators, 13*(10), DOI: 10.3390/act13100386.
6. Wang, X., Liang, X., Hu, S., & Xin, Q. (2025). Non-Singular Fast Sliding Mode Control of Robot Manipulators Based on Integrated Dynamic Compensation. *Actuators, 14*(5), DOI: 10.3390/act14050215.
7. Vo, A. T., Truong, T. N., & Kang, H.-J. (2022). A Novel Prescribed-Performance-Tracking Control System with Finite-Time Convergence Stability for Uncertain Robotic Manipulators. *Sensors, 22*(7), DOI: 10.3390/s22072615.

8. Lai, G., Zhou, S., Yang, W., Wang, X., & Wang, F. (2023). Prescribed Fixed-Time Adaptive Neural Control for Manipulators with Uncertain Dynamics and Actuator Failures. *Mathematics, 11*(13), DOI: 10.3390/math11132925.
9. Xu, J., Fang, L., Wang, H., Zhao, Q., Wan, Y., & Gao, Y. (2024). Observer-Based Finite-Time Prescribed Performance Sliding Mode Control of Dual-Motor Joints-Driven Robotic Manipulators with Uncertainties and Disturbances. *Actuators, 13*(9), DOI: 10.3390/act13090325.
10. Tran, N. H., Nguyen, D. H., & Nguyen, T. V. A. (2022). Enhanced robustness of robot manipulators using fixed-time sliding mode control and RBF neural network. *Archive of Mechanical Engineering, 72*(3), 519-544, DOI: 10.24425/ame.2025.154742.
11. Li, T., Zhang, G., Zhang, T., & Pan, J. (2024). Fixed-Time Sliding Mode Control for Robotic Manipulators Based on Disturbance Observer. *International Journal of Aerospace Engineering, 2024*(1), DOI: 10.1155/2024/1263459.
12. Zhang, D., Hu, J., Cheng, J., Wu, Z.-G., & Yan, H. (2024). A Novel Disturbance Observer Based Fixed-Time Sliding Mode Control for Robotic Manipulators with Global Fast Convergence. *IEEE/CAA Journal of Automatica Sinica, 11*(3), 661-672, DOI: 10.1109/JAS.2023.123948.
13. Li, J., Du, J., & Hu, X. (2020). Robust adaptive prescribed performance control for dynamic positioning of ships under unknown disturbances and input constraints. *Ocean Engineering, 206*, DOI: 10.1016/j.oceaneng.2020.107254.
14. Kong, L., Ouyang, Y., & Liu, Z. (2023). Adaptive neural network fixed-time control for an uncertain robot with input nonlinearity. *International Journal of Robust and Nonlinear Control, 34*(5), 3033-3056, DOI: 10.1002/rnc.7122.
15. Ghanooni, P., Habibi, H., Yazdani, A., Wang, H., Zadeh, S. M., & Ferrara, A. (2024). Prescribed performance control of a robotic manipulator with unknown control gain and assigned settling time. *ISA Transactions, 145*, 330-354, DOI: 10.1016/j.isatra.2023.12.011.

Điều khiển mô-men tính toán theo lập lịch độ lợi thời gian định trước với quan sát nhiễu dư cho bộ điều khiển robot

Tóm tắt: *Bài báo này giới thiệu một phương pháp điều khiển mới cho tay máy robot, cho phép hệ thống đạt được chất lượng bám nhanh và có thể dự đoán trong một khoảng thời gian được người dùng cài đặt. Phương pháp sử dụng kỹ thuật thay đổi hệ số điều khiển theo thời gian để tăng dần độ lợi khi hệ thống tiến gần đến thời điểm cuối, giúp giảm sai số nhanh hơn nhưng vẫn đảm bảo tính ổn định. Một bộ quan sát nhiễu được tích hợp nhằm ước lượng sai lệch mô hình và nhiễu ngoài theo thời gian thực, giúp bộ điều khiển duy trì độ bền vững khi robot chịu tác động từ các nhiễu không chắc chắn hoặc biến thiên theo thời gian. Phân tích lý thuyết cho thấy sai số bám hội tụ với thời gian đặt trước trong trường hợp lý tưởng và vẫn được giữ trong giới hạn khi tồn tại nhiễu biến thiên theo thời gian. Mô phỏng trên robot hai bậc tự do chứng minh phương pháp đề xuất vượt trội so với điều khiển mô men tính toán truyền thống, kể cả khi có bù nhiễu. Kết quả xác nhận độ chính xác, độ bền vững và khả năng kiểm soát thời gian quá độ của phương pháp.*

Từ khóa: điều khiển theo thời gian cài đặt; lập lịch độ lợi; điều khiển mô men tính toán; quan sát nhiễu; tay máy robot; bám quỹ đạo bền vững

Điều khiển tư thế quadrotor dựa trên bộ điều khiển chủ động loại nhiễu cải tiến

Kiều Bích Sơn [1*], Cao Hữu Tình[2]

[1]*Hệ 2, Học viện Kỹ thuật quân sự*

[2]*Viện Tên lửa và Kỹ thuật điều khiển, Học viện Kỹ thuật quân sự*

Email; Tel: kbichson@lqdtu.edu.vn

Tóm tắt

Quadrotor là một đối tượng điều khiển dưới tác động nên việc điều khiển nó là một thách thức lớn. Một số vấn đề cần giải quyết khi thiết kế hệ thống điều khiển quadrotor đó là tính phi tuyến, liên hệ chéo giữa các kênh điều khiển, sự bất định mô hình và tác động của nhiễu loạn bên ngoài. Nghiên cứu này đề xuất một phương pháp thiết kế bộ điều khiển bám tư thế quadrotor dựa trên bộ điều khiển chủ động loại nhiễu cải tiến. Chất lượng của bộ điều khiển được đề xuất được xác nhận nhờ mô phỏng số khi so sánh với bộ điều khiển chủ động loại nhiễu chuẩn.

Từ khóa: *điều khiển tư thế quadrotor; điều khiển bền vững; điều khiển chủ động loại nhiễu; bộ ước lượng nhiễu loạn và bất định.*

1 Đặt vấn đề

Máy bay không người lái (UAV) là một robot di động di chuyển tự động trên không; có thể bay mà không cần hoặc cần rất ít sự can thiệp của con người. Chúng đã nổi lên như một công cụ công nghệ cho phép phát triển nhiều ứng dụng đa dạng, vốn nguy hiểm hoặc khó thực hiện đối với con người [1,2].

Trong số các loại UAV đa dạng, máy bay không người lái quadrotor, còn được gọi là quadcopter đã trở nên vô cùng phổ biến trong thập kỷ qua. Khả năng cất cánh và hạ cánh thẳng đứng độc đáo cùng với khả năng cơ động và khả năng bay lơ lửng, đã khiến chúng trở thành lựa chọn ưa thích cho nhiều ứng dụng và dịch vụ trong các lĩnh vực như hậu cần, nông nghiệp, giám sát môi trường và vệ sinh, cùng nhiều lĩnh vực khác [3-8]. Tuy nhiên, bản chất dưới tác động của các quadrotor và mối quan hệ phức tạp giữa động lực học tịnh tiến và quay đặt ra một thách thức lớn cho các nhà nghiên cứu trong việc điều khiển vị trí của chúng theo quỹ đạo mong muốn. Ví dụ, trong [9] các tác giả đã xem xét các phương pháp luận khác nhau để điều khiển hệ thống quadrotor và chỉ ra rằng hầu hết các nghiên cứu đều tập trung vào việc chứng minh lý thuyết hơn là tính khả thi của chính bộ điều khiển. Sự phức tạp này nhấn mạnh nhu cầu về các chiến lược điều khiển mới và bền vững.

Một chiến lược điều khiển được các nhà nghiên cứu và học thuật ưa chuộng trong những năm gần đây là điều khiển ADRC [10,11], chiến lược điều khiển này dựa trên việc biểu diễn hệ thống như một chuỗi các bộ tích phân trong đó nhiễu loạn tổng đại diện cho tất cả các hiệu ứng nhiễu loạn bên trong và bên ngoài. Các nhiễu loạn như vậy được ước lượng bằng cách sử dụng bộ quan sát trạng thái mở rộng (ESO) và được bù trong luật điều khiển. Trong [12] thực hiện bộ điều khiển ADRC tuyến tính cho vòng bám vị trí của quadrotor và bộ điều khiển PID bền vững cho vòng bám tư thế, chiến lược điều khiển chỉ phụ thuộc vào các phép đo vị trí và tư thế của quadrotor. Kỹ thuật điều khiển ADRC được phát triển ở [13] cho bám sát tư thế và độ cao của quadrotor khi có nhiễu loạn bên ngoài tác động. Trong [14] thực hiện một bộ điều khiển ADRC bền vững cho UAV nhiều cánh quạt cho phép duy trì chất lượng ngay cả khi có sự thay đổi lớn các tham số trong mô hình UAV. Sự thay đổi đột ngột hệ số khuếch đại của bộ quan sát ESO tại điểm chuyển tiếp có thể làm cho ước lượng nhiễu loạn không trơn và gây ra thăng giáng trong tín hiệu điều khiển. Vấn đề này không được xem xét trong các nghiên cứu ở trên.

Điều khiển bền vững quadrotor sử dụng các bộ ước lượng nhiễu loạn và bất định cũng là chiến lược đang được nghiên cứu và phát triển. Trong [15] một bộ điều khiển bám tư thế và vị trí bền vững cho

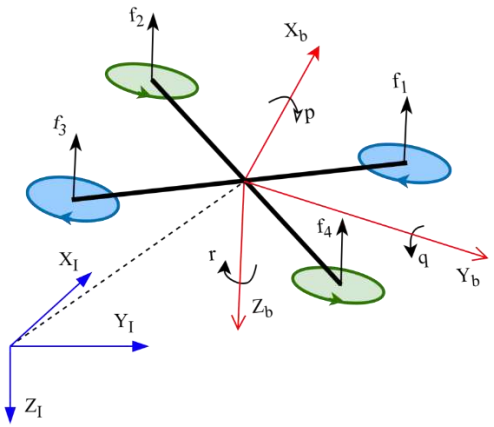
quadrotor có tính đến tính phi tuyến hệ thống, bất định khí động học và nhiễu loạn gió bên ngoài được thực hiện. Một bộ ước lượng nhiễu loạn và bất định (UDE) và sơ đồ điều khiển nghịch đảo động học phi tuyến (NDI) được đề xuất cho thực hiện chiến lược điều khiển này. Trong nghiên cứu ở [16] bộ điều khiển bền vững dựa trên bộ ước lượng phi tuyến được thiết kế cho điều khiển góc hướng và vị trí của quadrotor được thực hiện. Bộ điều khiển này đảm bảo bám sát các tham chiếu mong đợi trong sự hiện diện của thay đổi tham số và nhiễu loạn bên ngoài bằng cách sử dụng bộ ước lượng chế độ trượt bậc cao (HOSM) để ước lượng các nhiễu loạn này và bù trong luật điều khiển.

Phần còn lại của bài báo này được tổ chức như sau: mục 2 mô tả mô hình động lực học quadrotor; mục 3 trình bày phương pháp thiết kế bộ điều khiển ADRC cải tiến cho điều khiển tư thế của quadrotor; mục 4 thực hiện các mô phỏng khảo sát đánh giá chất lượng của bộ điều khiển được đề xuất.

2 Mô hình động lực học của quadrotor

Quadrotor là một hệ thống dưới tác động với bốn động cơ có thể được bố trí theo sơ đồ '+' hoặc 'X', hoạt động với bốn đầu vào điều khiển trên các động cơ và đầu ra sáu bậc tự do (DOF) như trên hình 1. Ở đây vị trí của quadrotor được đo trong hệ tọa độ quán tính, còn tư thế (ϕ, θ, ψ tương ứng là góc liệng, góc gật, góc hướng) được đo trong hệ tọa độ liên kết. Các cặp cánh 1-3 quay theo chiều kim đồng hồ còn cặp cánh 2-4 quay theo chiều ngược chiều kim đồng hồ. Lực nâng được sinh ra bởi mỗi động cơ được thể hiện bởi f_i, ở đây $i \in \{1, 2, 3, 4\}$ là chỉ số động cơ.

Như ta nhìn thấy trên Hình 1, hai hệ tọa độ tham chiếu khác nhau được gắn với quadrotor. Hệ tọa độ liên kết gắn liền với quadrotor dùng để mô tả chuyển động quay, còn hệ tọa độ quán tính cố định được sử dụng để mô tả chuyển động tịnh tiến của quadrotor. Các phương trình mô tả chuyển động quay của quadrotor dựa trên cách tiếp cận Newton-Euler được thể hiện ở phương trình (1); còn các phương trình mô tả chuyển động tịnh tiến của quadrotor được thể hiện ở phương trình (2).



Hình 1: Các thành phần động lực học của quadrotor

$$\begin{cases} \dot{\phi} = p + q\sin(\phi)\tan(\theta) + r\cos(\phi)\tan(\theta) \\ \dot{\theta} = q\cos(\phi) - r\sin(\phi) \\ \dot{\psi} = q\dfrac{\sin(\phi)}{\cos(\theta)} + r\dfrac{\cos(\phi)}{\sin(\theta)} \\ \dot{p} = \dfrac{I_{YY} - I_{ZZ}}{I_{XX}}qr - \dfrac{k_r q}{I_{XX}} + \dfrac{Lu_1}{I_{XX}} + d_{ar} \\ \dot{q} = \dfrac{I_{ZZ} - I_{XX}}{I_{YY}}pr + \dfrac{k_r p}{I_{YY}} + \dfrac{Lu_2}{I_{YY}} + d_{ap} \\ \dot{r} = \dfrac{I_{XX} - I_{YY}}{I_{ZZ}}pq + \dfrac{c_r u_3}{I_{ZZ}} + d_{ay} \end{cases} \quad (1)$$

$$\begin{cases} \ddot{x} = -\dfrac{u_4}{m}(\sin\phi\sin\psi + \cos\phi\sin\theta\cos\psi) - K_{dx}\dot{x} + d_{px} \\ \ddot{y} = -\dfrac{u_4}{m}(\cos\phi\sin\theta\sin\psi - \sin\phi\cos\psi) - K_{dy}\dot{y} + d_{py} \\ \ddot{z} = g - \dfrac{u_4}{m}\cos\phi\cos\theta - K_{dz}\dot{z} + d_{pz} \end{cases} \quad (2)$$

ở đây u_i $(i = 1, 2, 3, 4)$ là các tín hiệu đầu vào; p, q, và r thể hiện vận tốc góc quanh các trục X_b, Y_b, Z_b tương ứng được đo trong hệ tọa độ liên kết; L là khoảng cách giữa tâm của động cơ và trọng tâm quadrotor; c_r là tỉ số giữa hệ số lực đẩy và lực cản; k_r là hệ số lực cản quay; I_{XX}, I_{YY}, I_{ZZ} là mô men quán tính theo các trục X_b, Y_b, Z_b tương ứng; d_{ar}, d_{ap}, d_{ay} tương ứng là các nhiễu loạn bên ngoài tác động lên các kênh điều khiển liệng, gật, hướng; K_{dx}, K_{dy}, K_{dz} là các hệ số cản tương ứng theo các trục X_I, Y_I, Z_I; d_{px}, d_{py}, d_{pz} tương ứng là các nhiễu loạn bên ngoài tác động lên chuyển động tịnh tiến theo các trục X_I, Y_I, Z_I.

Quan hệ giữa đầu vào điều khiển và vận tốc góc của các cánh quạt được thể hiện trong phương trình (3) như sau:

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} k_t & -k_t & -k_t & k_t \\ k_t & k_t & -k_t & -k_t \\ k_d & -k_d & k_d & -k_d \\ k_t & k_t & k_t & k_t \end{bmatrix} \begin{bmatrix} \Omega_1^2 \\ \Omega_2^2 \\ \Omega_3^2 \\ \Omega_4^2 \end{bmatrix} \quad (3)$$

ở đây k_t là hệ số lực đẩy, k_d là hệ số mô men cản và Ω_i $(i = 1, 2, 3, 4)$ là vận tốc góc của các cánh quạt.

3 Thiết kế bộ điều khiển ADRC cải tiến cho điều khiển tư thế của quadrotor

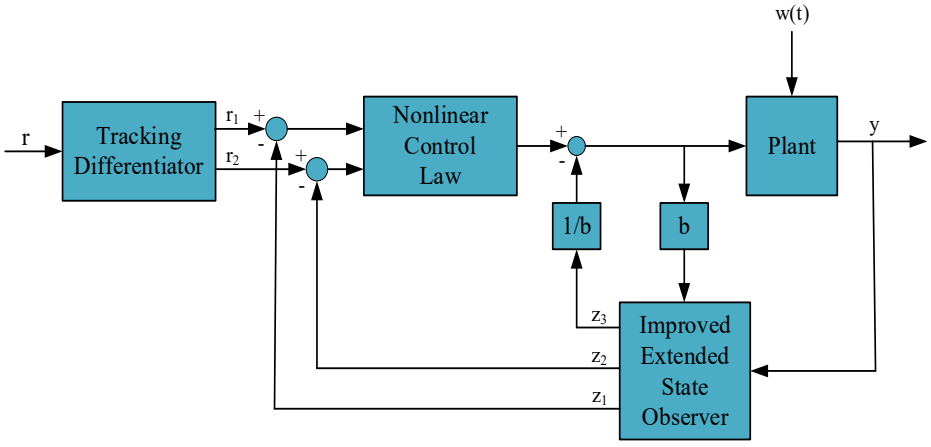
Ý tưởng cốt lõi của ADRC do Han đề xuất trong [10] là coi nhiễu động tác động lên hệ thống như một biến trạng thái mở rộng và thiết kế một bộ quan sát trạng thái ESO để ước lượng nhiễu loạn và được bù trong luật điều khiển phản hồi nhằm loại bỏ ảnh hưởng của nhiễu loạn, nhờ đó ADRC có thể đạt được độ bền vững trước nhiễu loạn. Tuy nhiên hàm phi tuyến có tên fal trong ADRC là một hàm từng đoạn và các đạo hàm của hàm này không liên tục tại các điểm chuyển tiếp. Theo thực nghiệm và phân tích, hàm fal có thể gây ra hiện tượng dao động trong ước lượng nhiễu loạn dẫn đến dao động trong đầu vào điều khiển. Để giải quyết vấn đề này, bài báo này đề xuất một hàm phi tuyến liên tục mới có đạo hàm cũng là một hàm liên tục. Phương pháp Gauss-Newton cho bài toán bình phương tối thiểu phi tuyến được sử dụng để xác định các tham số của hàm phi tuyến liên tục này. Cấu trúc của ADRC cải tiến (IADRC) được thể hiện trên Hình 2.

Phân tích ESO cổ điển và cải tiến

Khả năng chống nhiễu là chất lượng cốt lõi của ADRC, điều này đạt được bằng cách ước lượng theo thời gian thực giá trị của nhiễu loạn tổng thông qua ESO và được bù trừ trong luật điều khiển phản hồi để loại bỏ ảnh hưởng của nhiễu loạn.

Ta xem xét hệ thống một đầu vào và một đầu ra (SISO) được mô tả bởi khâu dao động bậc hai được biểu diễn như sau [10]:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = f(\square) + bu \\ y = x_1 \end{cases} \tag{4}$$



Hình 2: Cấu trúc của IADRC

ở đây $f(\cdot)$ là nhiễu loạn tổng của hệ thống, nó có thể được xem xét như là một biến trạng thái mở rộng biểu diễn bởi x_3.

Bởi vậy theo lý thuyết điều khiển ADRC ta nhận được hệ thống được mở rộng sau:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_3 + bu \\ \dot{x}_3 = \dot{f}(\square) \\ y = x_1 \end{cases} \tag{5}$$

Giả định rằng giá trị được ước lượng của các biến trạng thái mở rộng là $Z = [z_1, z_2, z_3]^T$, bộ quan sát ESO phi tuyến được xây dựng theo hệ thống (5) được biểu diễn trong phương trình (6)

$$\begin{aligned} e &= z_1 - y \\ \dot{z}_1 &= z_2 - \beta_{01} e \\ \dot{z}_2 &= z_3 - \beta_{02} fal(e, \alpha_1, \delta) + bu \\ \dot{z}_3 &= -\beta_{03} fal(e, \alpha_2, \delta) \end{aligned} \tag{6}$$

ở đây $\beta_{01}, \beta_{02}, \beta_{03}$ là các hệ số khuếch đại trong bộ quan sát ESO còn fal là hàm được sử dụng để thực hiện tính phi tuyến ESO và có thể được biểu diễn như sau:

$$fal(e, \alpha, \delta) = \begin{cases} \dfrac{e}{\delta^{1-\alpha}}, & |e| \le \delta \\ |e|^{\alpha} \, sign(e), & |e| \ge \delta \end{cases} \tag{7}$$

ở đây α, δ là các tham số của hàm phi tuyến.

Trong hàm fal này độ lớn của α xác định mức độ phi tuyến của hàm; giá trị α càng lớn thì mức độ phi tuyến của hàm càng mạnh và ngược lại mức độ phi tuyến càng yếu. Tham số δ là điểm chuyển tiếp của hàm, nó xác định độ dài của khoảng tuyến tính. Trong quá trình thiết kế hệ thống điều khiển việc lựa chọn giá trị δ rất quan trọng và có thể dẫn đến hiện tượng dao động trong ước lượng nhiễu từ đó dẫn đến dao động trong tín hiệu điều khiển đầu vào.

Hàm trong phương trình (7) là một hàm lẻ đối xứng qua gốc tọa độ, do đó ta có thể phân tích cho trường hợp khi $e > 0$. Đạo hàm của hàm tại điểm δ được tính như sau:

$$fal'(e, \alpha, \delta) = \begin{cases} \delta^{\alpha-1}, 0 < e < \delta \\ \alpha e^{\alpha-1}, e > \delta \end{cases} \tag{8}$$

Ta thấy khi $e = \delta$ thì $\alpha e^{\alpha-1} \neq \delta^{\alpha-1}$.

Như vậy đạo hàm của hàm fal tại điểm chuyển tiếp δ là không liên tục và có sự thay đổi đột ngột tại điểm $\pm\delta$, điều này dẫn đến giá trị nhiễu ước lượng không đủ trơn và gây ra hiện tượng dao động trong đầu vào điều khiển. Để giải quyết vấn đề này đồng thời đảm bảo các đặc tính của khếch đại nhỏ

với sai số lớn và khuếch đại lớn với sai số nhỏ thì một hàm phi tuyến liên tục mới có tên là Fal được đề xuất để thay thế, hàm này được mô tả trong phương trình (9) như sau:

$$Fal(e,\gamma,\alpha) = \gamma\left|\ln\left(e+\sqrt{e^2+1}\right)\right|^{\alpha}\arctan(e) \tag{9}$$

Vì hàm Fal là liên tục nên sẽ không có sự thay đổi đột ngột đạo hàm của nó tại điểm $\pm\delta$. Các tham số của hàm Fal trong (9) cần được lựa chọn sao cho hàm Fal so khớp nhất với hàm fal. Thuật toán Gauss-Newton được áp dụng để giải quyết vấn đề này. Thuật toán Gauss-Newton được phát biểu như sau:

Với M phép đo $y_i(i = 1,2,\dots M)$ được lấy trong (7) ta cần xác định N tham số $\boldsymbol{\beta} = [\beta_1,\beta_2,\dots\beta_N]$ của hàm Fal (ở đây $M > N$) sao cho tổng bình phương các sai số xấp xỉ $r_i(\boldsymbol{\beta}) = y_i - Fal(e_i,\boldsymbol{\beta})$ là nhỏ nhất:

$$S(\boldsymbol{\beta}) = \sum_{i=1}^{M} r_i(\boldsymbol{\beta})^2 \tag{10}$$

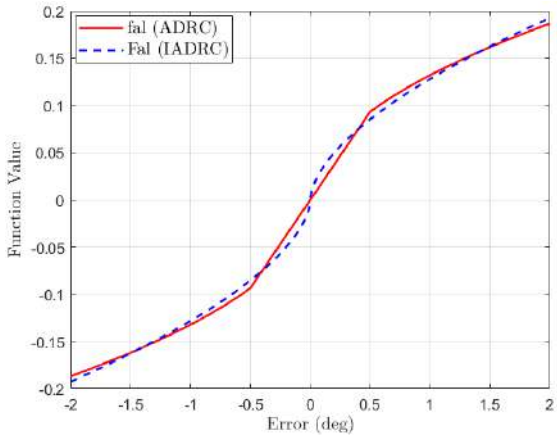
Việc cập nhật vector tham số $\boldsymbol{\beta}$ được thực hiện qua các bước lặp với điều kiện ban đầu $\boldsymbol{\beta}^{(0)}$ đã biết theo phương trình sau:

$$\boldsymbol{\beta}^{(k+1)} = \boldsymbol{\beta}^{(k)} - \left(\mathbf{J}_r^T\mathbf{J}_r\right)^{-1}\mathbf{J}_r^T\mathbf{r}\left(\boldsymbol{\beta}^{(k)}\right) \tag{11}$$

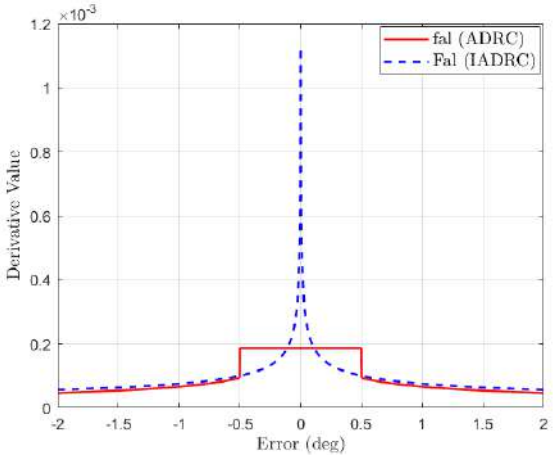
ở đây $\boldsymbol{r} = [r_1\ r_2\ \dots\ r_M]^T$ còn $\boldsymbol{J}_r$ là ma trận Jacobi được xác định như sau:

$$(\mathbf{J}_r)_{ij} = \frac{\partial r_i\left(\boldsymbol{\beta}^{(k)}\right)}{\partial\beta_j}$$

Để minh họa ta lấy ví dụ cho trường hợp trong (7) đặt $\alpha = 0.5, \delta = 0.5$. Vector tham số cần xác định trong (9) là $\boldsymbol{\beta} = [\gamma\ \alpha]$. Trên hình 3 thể hiện kết quả so khớp hàm còn trên hình 4 thể hiện các đạo hàm của các hàm phi tuyến.



Hình 3: Hàm phi tuyến trong ESO chuẩn và cải tiến



Hình 4: Đạo hàm của hàm phi tuyến

Kết quả thể hiện trên Hình 3 cho thấy hàm Fal xấp xỉ gần như chính xác hàm fal, còn kết quả trên Hình 4 cho thấy đạo hàm của hàm Fal là trơn trên toàn bộ dải sai số. Ở vùng lân cận bên ngoài $\pm\delta$ thì sự thay đổi của fal và Fal gần như nhau; trong khi đó ở vùng $\pm\delta$ đạo hàm của fal là không đổi còn đạo hàm của Fal có thay đổi, nhất là ở vùng sai số lân cận không.

Bộ vi phân bám

Bộ vi phân bám (TD) là một hệ thống cung cấp khả năng xấp xỉ trơn tín hiệu đầu vào và đạo hàm của nó, TD thường được sử dụng trong các hệ thống điều khiển để lọc nhiễu cảm biến và tạo ra tín hiệu vận tốc hoặc gia tốc chính xác cho điều khiển phản hồi. Nằm giữa ranh giới của bộ lọc và bộ vi phân, TD là bộ lọc thông thấp mà có thể thu được tín hiệu vi phân từ đầu vào nhiễu, cải thiện độ ổn định và chất lượng của hệ thống điều khiển. Các phương trình mô tả bộ vi phân bám được thể hiện như sau [10]:

$$\begin{aligned} \dot{v}_1 &= v_2 \\ \dot{v}_2 &= u \\ u &= fhan(v_1 - v_d, v_2, r_0, h_0) \end{aligned} \tag{12}$$

ở đây v_d, v_1, v_2 lần lượt là tín hiệu đầu vào mong đợi, tín hiệu bám sát và vi phân của tín hiệu bám sát; r_0, h_0 là hai tham số điều chỉnh được của TD, ở đây r_0 là hệ số tốc độ nhằm tăng tốc độ đáp ứng và giảm quá trình quá độ còn h_0 là hệ số lọc, khi tăng giá trị h_0 có thể cải thiện hiệu quả lọc nhưng sẽ gây ra giữ chậm pha lớn hơn.

Hàm $fhan(v_1, v_2, r_0, h_0)$ được định nghĩa là:

$$
\begin{aligned}
&d = r_0 h_0^2, \quad a_0 = h_0 v_2, \quad y = v_1 + a_0 \\
&a_1 = \sqrt{d(d + 8|y|)} \\
&a_2 = a_0 + sign(y)(a_1 - d)/2 \\
&s_y = (sign(y+d) - sign(y-d))/2 \\
&a = (a_0 + y - a_2)s_y + a_2 \\
&s_a = (sign(a+d) - sign(a-d))/2 \\
&fhan = -r_0\left(\frac{a}{d} - sign(a)\right)s_a - r_0 sign(a)
\end{aligned}
\tag{13}
$$

Điều khiển phản hồi sai số trạng thái phi tuyến

Điều khiển phản hồi sai số trạng thái phi tuyến (NLSEF) là bộ điều khiển phi tuyến được mô tả trong [10], các phương trình mô tả bộ điều khiển này được thể hiện trong phương trình (14).

$$
\begin{aligned}
&\mathrm{e}_1 = \mathrm{v}_1 - \mathrm{z}_1 \\
&\mathrm{e}_2 = \mathrm{v}_2 - \mathrm{z}_2 \\
&\mathrm{u} = \frac{\mathrm{u}_0 - \mathrm{z}_3}{\mathrm{b}_0}
\end{aligned}
\tag{14}
$$

Tín hiệu điều khiển u_0 được xác định như sau:

$$u_0 = \beta_{1e} fal(e_1, \alpha_1, \delta_1) + \beta_{2e} fal(e_2, \alpha_2, \delta_2) \tag{15}$$

Hàm fal trong (15) giống như trong (7). Vì vậy ta có thể thay thế hàm fal này bằng hàm Fal trong (9).

Tổng hợp bộ điều khiển IADRC cho điều khiển tư thế của quadrotor

Từ các phương trình mô tả động lực học tư thế trong (1) và bỏ qua các thành phần chéo kênh trong các phương trình động hình học tư thế ta có thể biến đổi thành ba kênh điều khiển độc lập. Ở đây ta xem xét cho kênh liệng. Các phương trình mô tả động lực học tư thế theo kênh liệng được thể hiện như sau:

$$
\begin{cases}
\dot{\phi} = p \\
\dot{p} = f(\square) + bu
\end{cases}
\tag{16}
$$

ở đây

$$
\begin{cases}
f(\square) = \dfrac{I_{YY} - I_{ZZ}}{I_{XX}} qr - \dfrac{k_r q}{I_{XX}} + d_r \\
b = \dfrac{L}{I_{XX}}
\end{cases}
\tag{17}
$$

Từ các phương trình trong (16) và (17) ta sẽ tổng hợp được bộ điều khiển IADRC theo cấu trúc thể hiện trên Hình 2. Trong đó bộ vi phân bám được thực hiện dựa trên các phương trình (12), (13); bộ quan sát ESO được thực hiện dựa trên phương trình (6) với hàm fal được thay thế bằng hàm Fal trong (9); bộ điều khiển NLSEF được thực hiện dựa trên các phương trình (14), (15).

4. Mô phỏng khảo sát chất lượng bộ điều khiển được đề xuất

Các tham số danh định của quadrotor được sử dụng trong mô phỏng là $m = 1.27\text{kg}, I_{XX} = 0.03339\ \text{kg} \cdot \text{m}^2, I_{YY} = 0.04339\ \text{kg} \cdot \text{m}^2, I_{ZZ} = 0.07052\ \text{kg} \cdot \text{m}^2, L = 0.194\ \text{m}, k_r = 0.1, c_r = 0.45.$

Các góc được tính theo đơn vị độ, các vận tốc góc được tính theo độ/s, vị trí được tính theo đơn vị mét, vận tốc thẳng được tính theo đơn vị m/s.

4.1 So sánh chất lượng giữa bộ điều khiển tư thế ADRC và IADRC

a. Kịch bản mô phỏng

Quadrotor được điều khiển tại tư thế ban đầu $\phi_0 = 0, \theta_0 = 0, \psi_0 = 0$; vận tốc góc ban đầu $\dot{\phi}_0 = 0, \dot{\theta}_0 = 0, \dot{\psi}_0 = 0.$

Các góc tư thế mong đợi được mô tả bởi các phương trình sau:

- Đối với kênh liệng: $\phi_d(t) = \begin{cases} 12\sin(\frac{2\pi}{5}t), & t < 15s \\ 20, & 15s \le t < 22.5s \\ -10, & 22.5s \le t \le 30s \end{cases}$

- Đối với kênh gật: $\theta_d(t) = 20\sin(\frac{2\pi}{10}t + \frac{\pi}{2}), \quad 0s \le t \le 30s$

- Đối với kênh hướng: $\psi_d(t) = \begin{cases} -5, & t < 10s \\ 10, & 10s \le t < 20s \\ 15\sin(\frac{2\pi}{4}t), & 20s \le t \le 30s \end{cases}$

b. Các tham số của bộ điều khiển tư thế ADRC và IADRC

Các tham số của TD; hàm fal của ESO và NLSEF trong ADRC; hàm Fal được tìm của ESO và NLSEF trong IADRC cho các kênh liệng, gật, hướng được thể hiện trên Bảng 1.

Bảng 1: Các tham số của TD, ESO và NLSEF trong ADRC và IADRC.

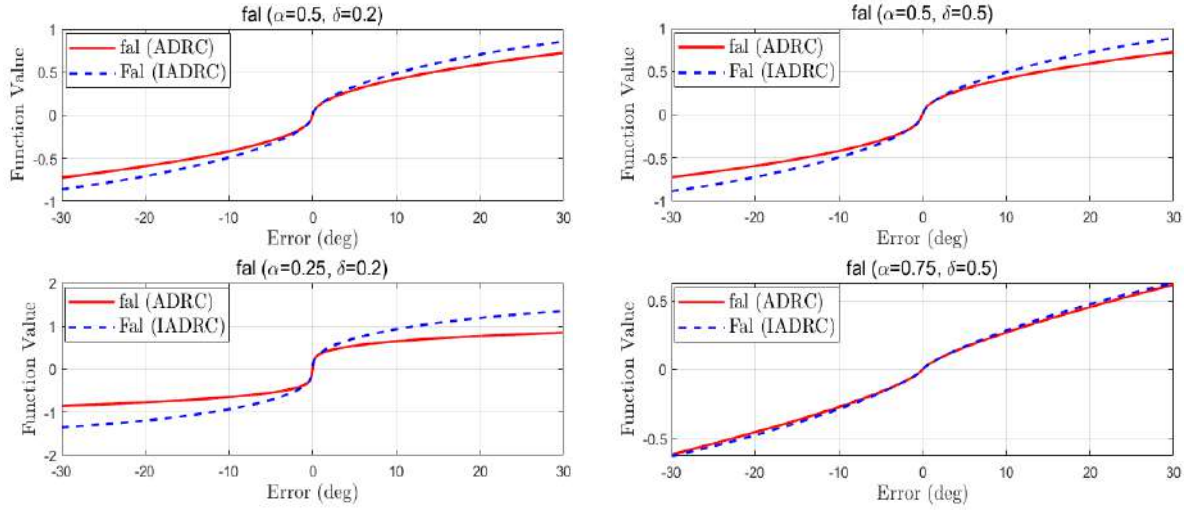
TD		**fal (ESO) ADRC**		**Fal (ESO) IADRC**		**fal (NLSEF) ADRC**		**Fal (NLSEF) IADRC**	
r_0	h_0	α	δ	γ	α	α	δ	γ	α
50	0.009	0.5	0.2	1.314192	-0.438024	0.5	0.5	1.380402	-0.412916
		0.25	0.2	1.831501	-0.614424	0.75	0.5	1.119148	-0.219292

Các tham số hệ số khuếch đại của ESO và NLSEF trong cả ADRC và IADRC đối với các kênh liệng, gật, hướng được thể hiện trên Bảng 2.

Bảng 2: Các hệ số khuếch đại của ESO và NLSEF trong cả ADRC và IADRC.

Roll channel					**Pitch channel**					**Yaw channel**				
ESO			NLSEF		ESO			NLSEF		ESO			NLSEF	
β_{01}	β_{02}	β_{03}	β_{1e}	β_{2e}	β_{01}	β_{02}	β_{03}	β_{1e}	β_{2e}	β_{01}	β_{02}	β_{03}	β_{1e}	β_{2e}
7.7	298.0	26.7	60.1	30.3	7.7	298.0	26.7	60.1	30.3	63.6	298.0	14.2	25.2	30.1

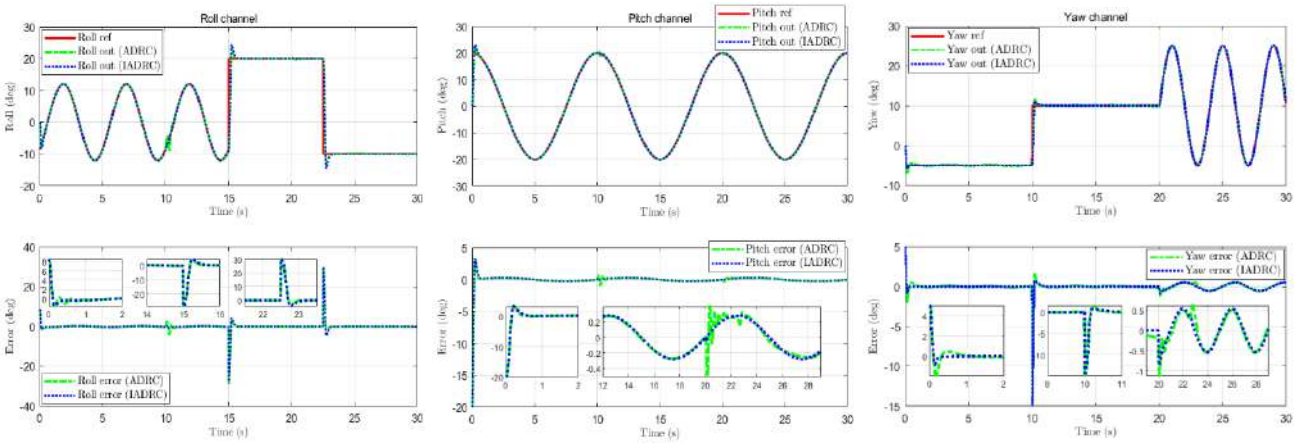
Trên Hình 5 thể hiện giá trị của các hàm fal và Fal theo các tham số được mô tả ở Bảng 1 với vùng sai số là ± 30 độ. Ta thấy rằng trong vùng sai số nhỏ thì hàm fal và Fal gần xấp xỉ nhau, còn trong vùng sai số lớn thì hàm Fal có giá trị lớn hơn hàm fal. Các sai số bám sát góc của bộ điều khiển ADRC và IADRC được thể hiện trên Hình 6. Các giá trị sai số bình phương trung bình căn bậc hai (RMS) của các sai số bám sát được thể hiện trên Bảng 3.



(a) (b)

Hình 5: Giá trị hàm fal và Fal: (a) $\alpha = [0.5, 0.25], \delta = 0.2$; (b) $\alpha = [0.5, 0.75], \delta = 0.5$

Các kết quả trên Hình 6 và Bảng 3 cho thấy bộ điều khiển IADRC có chất lượng bám sát góc được cải thiện hơn so với bộ điều khiển ADRC chuẩn. Trên Hình 6 ta thấy sai số bám sát của bộ điều khiển ADRC chuẩn bị thăng giáng khá lớn tại các thời điểm các góc tham chiếu thay đổi đột ngột; ví dụ như tại thời điểm $t = 10(s)$ đối với kênh liệng, $t = 10(s)$ và $t = 22.5(s)$ đối với kênh gật, $t = 20(s)$ đối với kênh hướng. Sự thăng giáng sai số bám sát này không xuất hiện trong bộ điều khiển IADRC.



(a) (b) (c)

Hình 6: Sai số bám sát góc của ADRC và IADRC: (a) kênh liệng, (b) kênh gật, (c) kênh hướng

Bảng 3: Giá trị RMS của các sai số bám sát góc bộ điều khiển ADRC và IADRC.

ADRC			**IADRC**		
Roll	Pitch	Yaw	Roll	Pitch	Yaw
2.4763	1.1276	0.8535	2.4565	1.1175	0.8377

5 Kết luận

Bài báo đề xuất một chiến lược điều khiển bám tư thế của quadrotor nhằm giải quyết các vấn đề liên quan đến tính phi tuyến, bất định mô hình và nhiễu loạn bên ngoài tác động. Chiến lược điều khiển này được xây dựng dựa trên bộ điều khiển IADRC với việc thay thế hàm biến đổi phi tuyến trong bộ điều khiển ADRC bởi hàm phi tuyến trơn trên toàn bộ vùng sai số. Điều này khắc phục hiện tượng thăng giáng trong ước lượng nhiễu loạn tại vùng chuyển tiếp của sai số bám sát. Kết quả khảo sát cho thấy bộ điều khiển IADRC có chất lượng cải thiện hơn so với bộ ADRC chuẩn, điều này thể hiện rõ hơn ở các thời điểm khi tín hiệu tham chiếu thay đổi đột ngột. Khối lượng tính toán được yêu cầu để thực hiện bộ điều khiển IADRC là không lớn. Vì vậy ta hoàn toàn có thể triển khai bộ điều khiển này trên các hệ vi xử lý thông dụng.

Tài liệu tham khảo

1. Erdelj, M.; Król, M.; Natalizio, E. Wireless Sensor Networks and Multi-UAV systems for natural disaster management. *Comput.Netw*. 2017, 124, 72–86. https://doi.org/10.1016/j.comnet.2017.05.021
2. Khalil, R.A.; Saeed, N.; Almutiry, M. UAVs-assisted passive source localization using robust TDOA ranging for search and rescue. *ICT Express* 2023, 9, 677–682. https://doi.org/10.1016/j.icte.2022.04.011
3. Alsamhi, S.H.; Ma, O.; Ansari, M.S.; Almalki, F.A. Survey on Collaborative Smart Drones and Internet of Things for Improving Smartness of Smart Cities. *IEEE Access* 2019, 7, 128125–128152. https://doi.org/10.1109/ACCESS.2019.2934998
4. Lytridis, C.; Kaburlasos, V.G.; Pachidis, T.; Manios, M.; Vrochidou, E.; Kalampokas, T.; Chatzistamatis, S. An Overview of Cooperative Robotics in Agriculture. *Agronomy* 2021, 11, 1818. https://doi.org/10.3390/agronomy11091818

5. Shelare, S.; Aglawe, K.; Waghmare, S.; Belkhode, P. Advances in water sample collections with a drone—A review. *Mater. Today Proc*. 2021, 47, 4490–4494. https://doi.org/10.1016/j.matpr.2021.05.327
6. Esposito, M.; Crimaldi, M.; Cirillo, V.; Sarghini, F.; Maggio, A. Drone and sensor technology for sustainable weed management: A review. *Chem. Biol. Technol. Agric*. 2021, 8, 18. https://doi.org/10.1186/s40538-021-00217-8
7. Nwaogu, J.M.; Yang, Y.; Chan, A.P.; Chi, H.l. Application of drones in the architecture, engineering, and construction (AEC) industry. *Autom. Constr*. 2023, 150, 104827. https://doi.org/10.1016/j.autcon.2023.104827
8. Rojo-García, J.I.; Vera-Chavarría, S.A.; Lozano-Hernández, Y.; Sánchez-Meza, V.G.; González-Sierra, J.; Oliva-Moreno, L.N. Data Collection in Areas without Infrastructure Using LoRa Technology and a Quadrotor. *Future Internet* 2024, 16, 186. https://doi.org/10.3390/fi16060186
9. Nascimento, T.P.; Saska, M. Position and attitude control of multi-rotor aerial vehicles: A survey. *Annu. Rev. Control* 2019, 48, 129–146. https://doi.org/10.1016/j.arcontrol.2019.08.004
10. Han, J. From PID to active disturbance rejection control. IEEE Trans. Ind. Electron. 2009, 56, 900‑906. https://doi.org/10.1109/TIE.2008.2011621
11. Sira-Ramirez, H.; Luviano-Juárez, A.; Ramírez-Neria, M.; Zurita-Bustamante, E.W. Active Disturbance Rejection Control of Dynamic Systems: A Flatness Based Approach; Butterworth-Heinemann: Oxford, UK, 2017.
12. Ramírez-Neria, M.; Luviano-Juárez, A.; González-Sierra, J.; Ramírez-Juárez, R.; Aguerrebere, J.; Hernandez-Martinez, E.G. Active Disturbance Rejection Control for the Trajectory Tracking of a Quadrotor. *Actuators* 2024, *13*, 340. https://doi.org/10.3390/act13090340
13. Z. Wang, D. Huang, T. Huang and N. Qin, "Active disturbance rejection control for a quadrotor UAV," 2020 IEEE 9th Data Driven Control and Learning Systems Conference (DDCLS), Liuzhou, China, 2020, pp. 1-5. https://doi.org/10.1109/DDCLS49620.2020.9275226
14. Chebbi, J., Briere, Y. (2022). Robust active disturbance rejection control for systems with internal uncertainties: Multirotor UAV application. Journal of Field Robotics, 39, 426–456. https://doi.org/10.1002/rob.22058
15. Dhadekar, D.D., Sanghani, P.D., Mangrulkar, K.K. *et al.* Robust Control of Quadrotor using Uncertainty and Disturbance Estimation. *J Intell Robot Syst* **101**, 60 (2021). https://doi.org/10.1007/s10846-021-01325-1
16. Bianchi, D.; Di Gennaro, S.; Di Ferdinando, M.; Acosta Lùa, C. Robust Control of UAV with Disturbances and Uncertainty Estimation. *Machines* **2023**, *11*, 352. https://doi.org/10.3390/machines11030352

Quadrotor Attitude Control based on Improved Active Disturbance Rejection Control

Abstract: *quadrotor is an underactuated mechanical system so quadrotor control is a significant challenge. Several issues need to be addressed when designing a quadrotor control system, including nonlinearity, cross-coupling between control channels, model uncertainty, and the impact of external disturbances. This study proposes a method for designing a quadrotor attitude tracking controller based on improved active disturbance rejection control. The quality of the proposed controller is validated through numerical simulations compared to a standard active disturbance rejection control.*

Keywords: quadrotor attitude control; robust control; active disturbance rejection control; uncertainty and disturbance estimator.

PHƯƠNG PHÁP ĐIỀU HƯỚNG CHỦ ĐỘNG CHO ROBOT DI ĐỘNG TRONG MÔI TRƯỜNG CÓ NGƯỜI DI CHUYỂN

Nghiêm Hoàng Nam, Trương Xuân Tùng, Hoàng Văn Bảy, Nguyễn Lan Anh*

Viện Tên lửa và kỹ thuật điều khiển, Học viện Kỹ thuật quân sự

Email: nghiemhoangkhanh@gmail.com, SĐT: 0962514182

Tóm tắt

Trong môi trường có con người, bài toán điều hướng robot di động không chỉ yêu cầu tránh va chạm mà còn phải bảo đảm an toàn, tính chấp nhận được và sự thoải mái trong quá trình tương tác người–robot. Đây là yêu cầu quan trọng đối với các robot dịch vụ hoạt động trong không gian trong nhà như hành lang, sảnh chờ hoặc khu vực công cộng, nơi robot phải thường xuyên chia sẻ không gian di chuyển với con người. Bài báo này đề xuất một phương pháp điều hướng cho robot di động trong môi trường động dựa trên mô hình Lực xã hội mở rộng. Điểm cốt lõi của phương pháp là tích hợp cơ chế dự đoán quỹ đạo ngắn hạn và dự đoán va chạm vào bộ lập kế hoạch cục bộ, cho phép robot phản ứng chủ động trước chuyển động của con người thay vì chỉ né tránh khi đã tiếp cận ở khoảng cách gần. Trên cơ sở đó, robot có thể tạo ra quỹ đạo chuyển động khả thi, liên tục và phù hợp hơn với bối cảnh tương tác trong môi trường có người di chuyển. Phương pháp được thiết kế cho các môi trường chưa biết trước, có mật độ người từ thấp đến trung bình, trong đó trạng thái và quỹ đạo của các đối tượng được theo dõi liên tục trong suốt quá trình điều hướng. Toàn bộ hệ thống được triển khai dưới dạng các nút trong nền tảng ROS và được đánh giá thông qua cả mô phỏng lẫn thực nghiệm. Kết quả thu được cho thấy phương pháp đề xuất cải thiện hiệu quả điều hướng của robot trong môi trường động, đồng thời đáp ứng tốt các yêu cầu về an toàn và khả năng vận hành trong tương tác người–robot.

Từ khóa:. *Social Focre Model (SFM)., Velocity Obstacles (VO), Reciprocal Velocity Obstacles(RVO)*

1. Giới thiệu

Robot dịch vụ di động ngày càng được triển khai trong các môi trường có mật độ người cao như bệnh viện, trung tâm thương mại và sân bay. Trong các không gian này, robot không chỉ cần tránh va chạm mà còn phải thể hiện hành vi phù hợp với chuẩn mực xã hội, đảm bảo sự an toàn và thoải mái cho con người. Điều hướng xã hội vì vậy trở thành một vấn đề trọng tâm trong nghiên cứu robot di động.

Các phương pháp lập kế hoạch cổ điển như A* [1], D* [2] và D* Lite [3] giải quyết hiệu quả bài toán tìm đường trong môi trường tĩnh hoặc bán động. Tuy nhiên, chúng chủ yếu xử lý vấn đề hình học và không xét đến hành vi của con người như các tác nhân có mục tiêu và ý định riêng. Khi áp dụng trong môi trường đông người, robot thường phải tái lập kế hoạch liên tục hoặc phản ứng muộn do thiếu cơ chế dự đoán chuyển động của con người.

Một hướng tiếp cận phổ biến trong điều hướng robot nhận biết con người là mô hình hóa không gian cá nhân **và** khoảng cách tương tác chấp nhận được dưới dạng trường chi phí cục bộ. Trong các nghiên cứu [4], [5], vùng ảnh hưởng xung quanh con người thường được biểu diễn bằng phân bố Gaussian hai chiều, sau đó được ánh xạ vào costmap cục bộ để hiệu chỉnh quá trình lập kế hoạch quỹ đạo của robot. Trên cơ sở đó, một số công trình tiếp tục mở rộng mô hình bằng cách đưa thêm thành phần định hướng chuyển động**,** trạng thái tương tác **và** ngữ cảnh hành vi vào hàm chi phí nhằm phản ánh tốt hơn đặc trưng tương tác người–robot [6], [7]. Tuy nhiên, các phương pháp này nhìn chung vẫn mang bản chất phản ứng cục bộ, đồng thời phụ thuộc đáng kể vào quá trình lựa chọn và hiệu chỉnh tham số. Trong các tình huống có nhiều

người xuất hiện đồng thời, sự chồng lấn giữa các trường chi phí có thể làm suy giảm tính khả thi của quỹ đạo, thậm chí dẫn đến hiện tượng robot bị kẹt cục bộ hoặc.

Nhóm phương pháp dựa trên vận tốc như [8], [9] và [10] tính toán tập vận tốc an toàn dựa trên vận tốc tương đối giữa các tác nhân. Các phương pháp này cho phép tránh va chạm chủ động hơn so với costmap truyền thống. Tuy nhiên, chúng thường giả định tính hợp tác giữa các tác nhân và chủ yếu đảm bảo tính khả thi hình học, trong khi chưa mô hình hóa rõ ràng các yếu tố xã hội như không gian cá nhân hay hành vi nhóm.

Một hướng tiếp cận khác là tích hợp dự đoán quỹ đạo người vào quá trình lập kế hoạch. Các nghiên cứu trong [11], [12] sử dụng mô hình Long Short-Term Memory (LSTM) hoặc học sâu để dự đoán vị trí tương lai của người đi bộ trong môi trường đông đúc. Ngoài ra, các phương pháp dựa trên tối ưu hóa quỹ đạo như Timed Elastic Band (TEB) [13] có thể xem xét vật cản động trong quá trình tối ưu. Tuy nhiên, trong phần lớn các nghiên cứu này, thành phần dự đoán thường được xử lý tách biệt với cơ chế điều khiển chính, hoặc chỉ ảnh hưởng đến quá trình tối ưu hình học mà chưa được tích hợp trực tiếp vào mô hình hành vi động lực học của robot.

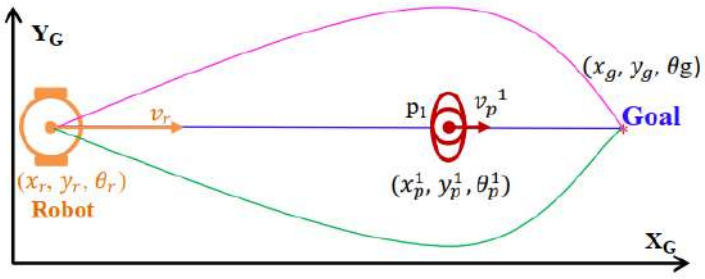
Social Force Model (SFM) do Helbing và Molnár đề xuất [14] là một mô hình nền tảng để mô phỏng chuyển động người đi bộ thông qua tổng hợp các lực ảo. Nhiều nghiên cứu đã áp dụng SFM cho điều hướng robot xã hội [15], coi robot như một tác nhân trong trường lực xã hội. Tuy nhiên, SFM nguyên bản chủ yếu dựa trên trạng thái tức thời (vị trí và vận tốc hiện tại) để tính toán lực tránh, do đó mang tính phản ứng cục bộ.

Từ các nghiên cứu trên có thể nhận thấy một hạn chế chung: mặc dù đã có các phương pháp mô hình hóa không gian xã hội hoặc dự đoán quỹ đạo người, việc tích hợp chặt chẽ dự đoán quỹ đạo và ước lượng vị trí va chạm tiềm năng vào cơ chế điều khiển dựa trên lực vẫn còn hạn chế. Phần lớn các phương pháp hoặc phản ứng dựa trên khoảng cách tức thời, hoặc tách rời bước dự đoán khỏi thành phần điều khiển.

Trong nghiên cứu này, chúng tôi đề xuất một phương pháp mở rộng SFM bằng cách tích hợp dự đoán quỹ đạo ngắn hạn của con người và ước lượng trước điểm giao cắt tiềm năng giữa robot và người. Thông tin dự đoán được đưa trực tiếp vào thành phần lực xã hội nhằm điều chỉnh hướng và cường độ lực tránh theo mức độ rủi ro va chạm trong tương lai gần. Cách tiếp cận này cho phép robot phát hiện sớm xung đột chuyển động, điều chỉnh vận tốc một cách chủ động và duy trì chuyển động mượt mà, ổn định hơn trong môi trường đông người.

3. Công cụ lập kế hoạch cục bộ mở rộng dựa trên SFM

3.1. Trạng thái của con người, trạng thái của robot và vị trí tương đối



Hình 1: Một ví dụ về môi trường xã hội năng động bao gồm một robot di động và một người đang di chuyển.

Chúng tôi xét một môi trường xã hội động với sự hiện diện của một robot di động và N người trong vùng lân cận của robot, như minh họa trong Hình 1. Robot được yêu cầu điều hướng từ trạng thái hiện tại $\mathbf{s}_s = [x_r, y_r, \theta_r, v_r, \omega_r]^T$ đến trạng thái mục tiêu

$\mathbf{s}_g = [x_g, y_g, \theta_g, v_g, \omega_g]^T$ trong khi đảm bảo tránh va chạm an toàn với con người trong suốt quá trình di chuyển. Chúng tôi giả sử trạng thái của robot được biểu diễn bởi $\mathbf{s}_r = [x_r, y_r, \theta_r, v_r, \omega_r]^T$ trong đó $[x_r, y_r]^T$ là vị trí, θ_r là góc định hướng, v_r là vận tốc tuyến tính và ω_r là vận tốc góc. Ngoài ra, giả sử có N người xuất hiện trong vùng lân cận của robot: $\mathbf{P} = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N\}$ trong đó $\mathbf{p}_i$ là người thứ i. Trạng thái của người $\mathbf{p}_i$ được biểu diễn bởi: $\mathbf{s}_p^i = [x_p^i, y_p^i, \theta_p^i, v_p^i]^T$ trong đó $[x_p^i, y_p^i]^T$ là vị trí, θ_p^i là hướng chuyển động và v_p^i là vận tốc tuyến tính. Bán kính của robot và của con người lần lượt được ký hiệu là r_r *và* r_h.

3.2. Mô hình lực xã hội SFM

Để xây dựng một mô hình có khả năng biểu diễn sự tương tác giữa người đi bộ và robot, chúng tôi được truyền cảm hứng từ các công trình của Helbing [6]. Đóng góp chính của các nghiên cứu này là: những thay đổi trong hành vi (quỹ đạo) có thể được giải thích thông qua các trường hoặc lực xã hội. Tuy nhiên, các công trình trên chưa xem xét sự tương tác giữa con người và robot, đây là một trong những đóng góp của nghiên cứu này.

Về mặt hình thức, mô hình lực xã hội giả định rằng một người đi bộ p_i có khối lượng m_i cố gắng di chuyển với vận tốc mong muốn v_i^0 theo hướng mong muốn $\mathbf{e}_i$ tức là vận tốc mong muốn:

$$\mathbf{v}_i^0 = v_i^0 \mathbf{e}_i \tag{1}$$

Phương trình chuyển động cơ bản của người đi bộ được cho bởi:

$$m_i \frac{d\mathbf{v}_i(t)}{dt} = \mathbf{F}_i(t) \tag{2}$$

Để đơn giản, ta giả sử $m_i = 1$ cho tất cả các đối tượng.

Tổng lực tác dụng lên người p_i bao gồm lực hướng mục tiêu và lực tương tác:

$$\mathbf{F}_i = \mathbf{F}_i^{goal} + \mathbf{F}_i^{int} \tag{3}$$

Lực hướng mục tiêu được mô tả bởi:

$$\mathbf{F}_i^{goal} = k\left(\mathbf{v}_i^0 - \mathbf{v}_i\right) \tag{4}$$

Lực tương tác là tổng các lực gây ra bởi những người khác, vật cản:

$$\mathbf{F}_i^{int} = \sum_{j \in P} \mathbf{f}_{i,j}^{int} + \sum_{o \in O} \mathbf{f}_{i,o}^{int} \tag{5}$$

trong đó: P là tập người trong môi trường, O là tập vật cản.

Lực tương tác giữa người p_i và thực thể q được mô hình hóa bởi:

$$\mathbf{f}_{i,q}^{int} = A_q e^{\frac{d_q - d_{i,q}}{B_q}} \frac{\mathbf{d}_{i,q}}{d_{i,q}} \tag{6}$$

trong đó: $q \in P \cup O$, A_q là cường độ lực, B_q là phạm vi tác động, $d_{i,q}$ là khoảng cách Euclid giữa p_i và thực thể q, $d_q = r_i + r_q$ là tổng bán kính, r_i, r_q lần lượt là bán kính của người và thực thể.

Do con người có trường nhìn hạn chế, lực tương tác được điều chỉnh bởi hệ số bất đẳng hướng:

$$w(\varphi_{i,q}) = \lambda + (1-\lambda)\frac{1+\cos(\varphi_{i,q})}{2} \tag{7}$$

trong đó: $\cos(\varphi_{i,q}) = -\mathbf{n}_{i,q} \cdot \mathbf{e}_i$ với $\mathbf{n}_{i,q}$ là vectơ đơn vị hướng từ thực thể q đến người p_i.

3.3. Thành phần dự đoán quỹ đạo chuyển động và vị trí va chạm

Trong SFM truyền thống, các lực tương tác giữa robot và các tác nhân trong môi trường được tính toán dựa trên khoảng cách hình học tại trạng thái hiện tại, do đó hành vi tránh va chạm chủ yếu mang tính phản ứng. Tuy nhiên, trong môi trường xã hội động và đông người, các tình huống va chạm nguy hiểm thường hình thành từ chuyển động tương đối giữa robot và các tác nhân theo thời gian, thay vì xuất hiện tức thời. Vì vậy, việc dự đoán vị trí chuyển động tiếp theo của các tác nhân là cần thiết nhằm cho phép robot sinh ra các lực tránh va chạm mang tính chủ động trong khuôn khổ SFM. Trên cơ sở đó chúng tôi mở rộng thành phần lực đẩy của SFM nhằm xử lý tương tác với các tác nhân đang di chuyển thông qua việc tích hợp dự đoán chuyển động ngắn hạn, cho phép robot chủ động ước lượng và tránh các va chạm tiềm năng trong môi trường động:

$$\mathbf{f}_{r,i}^{int} = \boldsymbol{F}_{r,i}^{pre} \tag{8}$$

trong đó $\mathbf{f}_{r,i}^{int}$ là lực đẩy do tác nhân trong môi gây ra, $\boldsymbol{F}_{r,i}^{pre}$ lực dự đoán ý định chuyển động và dự đoán va chạm của tác nhân.

Để cho phép robot tránh va chạm một cách chủ động, chuyển động của robot và các tác nhân lân cận được dự đoán trong một khoảng thời gian dự báo hữu hạn Tp.

Với giả thiết vận tốc của các tác nhân gần như không đổi trong khoảng thời gian ngắn, vị trí dự đoán của robot và tác nhân i tại thời điểm dự báo $\tau \in (0 \le \tau \le \mathrm{T_p})$ được xác định lần lượt:

$$\mathbf{p_r}(t,\tau) = \mathbf{p_r}(t) + \mathbf{v_r}(t)\tau \tag{9}$$

$$\mathbf{p_i}(t,\tau) = \mathbf{p_i}(t) + \mathbf{v_i}\,\tau \tag{10}$$

trong đó, τ là biến thời gian dự báo (prediction time horizon), đại diện cho khoảng thời gian liên tục từ thời điểm hiện tại đến giới hạn dự báo hữu hạn $\mathrm{T_p}$.

Khoảng cách hình học giữa robot và tác nhân i tại thời điểm dự đoán τ được tính:

$$\mathrm{d_{r,i}}(t,\tau) = \left\|\mathbf{p_r}(t,\tau) - \mathbf{p_i}(t,\tau)\right\| \tag{11}$$

Robot quét toàn bộ khoảng $\tau \in [0, \mathrm{T_p}]$ và xác định khoảng cách nhỏ nhất giữa robot và tác nhân trong khoảng dự đoán:

$$d_{r,i}^{\min} = \min_{\tau \in [0,T_p]} d_{r,i}(t,\tau) \tag{12}$$

Nếu thỏa mãn điều kiện: $\mathrm{d_{r,i}^{min}} < \mathrm{d_{safe}}$ thì tồn tại nguy cơ va chạm trong tương lai gần.

Tại thời điểm dự đoán τ tương ứng với khoảng cách nhỏ nhất, vị trí của tác nhân được xác định là vị trí va chạm dự đoán:

$$\mathbf{P^{pre}} = \mathbf{p_i}(t) + \mathbf{v_i}\,\tau \tag{13}$$

Điểm $\mathbf{P}^{pre}$ đại diện cho vị trí mà robot có khả năng tiếp cận nguy hiểm nhất với tác nhân trong tương lai và được mô hình hóa như một vật cản ảo tác động lên robot tại thời điểm hiện tại.

Vector khoảng cách hình học từ vị trí va chạm dự đoán về phía robot được xác định:

$$\mathbf{d}_{r,i}^{pred} = \mathbf{p}_r(t) - \mathbf{P}^{pre} \tag{14}$$

Vector đơn vị biểu diễn hướng tác động của lực tránh va chạm được xác định:

$$\mathbf{n}_{r,i}^{pred} = \frac{\mathbf{d}_{r,i}^{pred}}{\left\|\mathbf{d}_{r,i}^{pred}\right\|} \tag{15}$$

Vector này đảm bảo lực tránh va chạm luôn đẩy robot ra xa vùng va chạm tiềm năng.

Dựa trên khoảng cách hình học dự đoán, lực tránh va chạm do tác nhân i gây ra được xác định theo dạng hàm mũ của SFM:

$$\mathbf{F}_{r,i}^{pre} = A_{pre} e^{\left(\frac{d_{safe} - \left\|\mathbf{d}_{r,i}^{pred}\right\|}{B_{pre}}\right)} \mathbf{n}_{r,i}^{pred} \left(\lambda + (1-\lambda)\frac{1+\cos(\gamma_{r,i})}{2}\right) \tag{16}$$

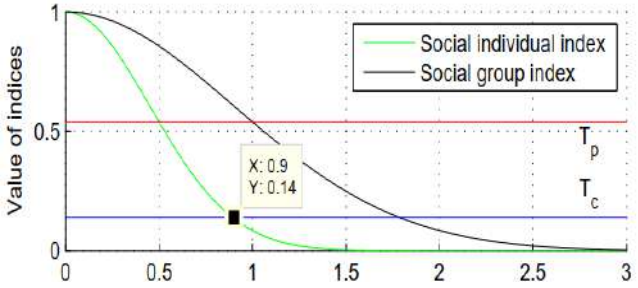
trong đó A_{pre} là cường độ lực tránh va chạm, B_{pre} là hệ số suy giảm của lực theo khoảng cách. λ hệ số đẳng hướng với $\lambda \in [0,1]$. $\gamma_{r,i}$ là là góc giữa hướng chuyển động của robot và hướng từ tác nhân tới robot.

Thành phần lực này được tích hợp trực tiếp vào mô hình Social Force Model mở rộng, cho phép robot tránh va chạm một cách chủ động dựa trên vị trí va chạm dự đoán trong tương lai, đồng thời vẫn giữ nguyên bản chất lực dựa trên khoảng cách của SFM truyền thống. Tổng lực tác dụng lên robot r tại thời điểm t được xác định như sau:

$$\mathbf{F}_r = \mathbf{F}_r^{goal} + \mathbf{F}_r^{int} \tag{17}$$

với $\mathbf{F}_r^{int} = \sum_{i \in P} \mathbf{f}_{r,i}^{pre} + \sum_{o \in O} \mathbf{f}_{r,o}^{int}$

3.4. Chỉ số An toàn - Thoải mái cho con người HCSI



Hình 2. Đồ thị của SII và SGI

Để so sánh trực quan hiệu quả của các mô hình này, chúng tôi áp dụng hai chỉ số trong chỉ số an toàn thoải mái cho con người (HCSI) do [15] giới thiệu. Chỉ số HCSI được sử dụng để ước tính sự an toàn và thoải mái của con người cũng như các hành vi được xã hội chấp nhận của robot di động. Bao gồm chỉ số cá nhân xã hội (SII Social Individual Index) và chỉ số nhóm xã hội (SGI Social Group Index).

Cụ thể, giá trị SII được áp dụng để đo lường mức độ an toàn về thể chất và tâm lý của mỗi cá nhân: Được tính toán dựa trên phương trình (18),

$$SII = \max_{i=1:N} \exp\left(-\left(\left(\frac{x_r - x_i^p}{\sqrt{2}\rho_0^{px}}\right)^2 + \left(\frac{y_r - y_i^p}{\sqrt{2}\rho_0^{py}}\right)^2\right)\right) \quad (18)$$

trong đó $\left(x_i^p, y_i^p\right)$ là vị trí của một cá nhân p_i, (x_r, y_r) là vị trí của robot, σ_0^{px} và σ_0^{py} là các độ lệch chuẩn, và N là số lượng người ở trong phạm vi gần robot. Giá trị chỉ số cá nhân xã hội (SII) theo khoảng cách tương ứng giữa con người và robot được minh họa trong Hình 2.

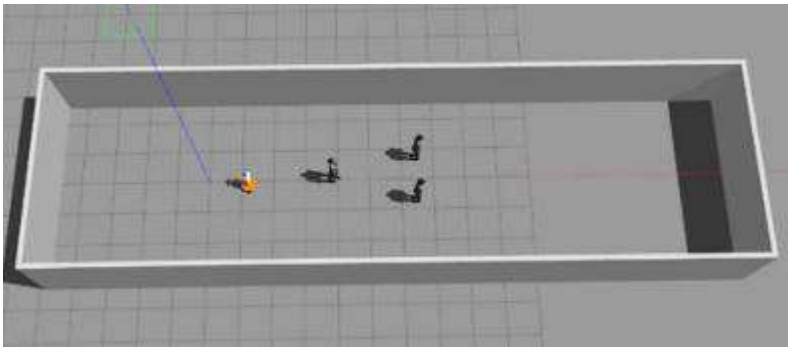
Trong khi SGI được dùng để đo lường mức độ an toàn thoải mái của nhóm người: Chỉ số nhóm xã hội (Social Group Index – SGI) được tính toán dựa trên phương trình (19):

$$SGI = \max_{k=1:K} \exp\left(-\left(\left(\frac{x_r - x_k^g}{\sqrt{2\sigma_k^g}}\right)^2 + \left(\frac{y_r - y_k^g}{\sqrt{2\sigma_k^g}}\right)^2\right)\right); \sigma_k^g = \frac{r_k^g}{2} \quad (19)$$

Trong đó σ_k^g được tính toán theo phương trình (18), r_k^g được xác định trong [16], và K là số lượng nhóm người trong phạm vi của robot [15]. Hình 2 minh họa ví dụ về giá trị SGI tương ứng với khoảng cách giữa robot và trung tâm không gian tương tác xã hộ

4. Thí nghiệm và kết quả

4.1. Cài đặt mô phỏng và thiết lập tham số mô phỏng



Hình 3. Môi trường gazebo Ros

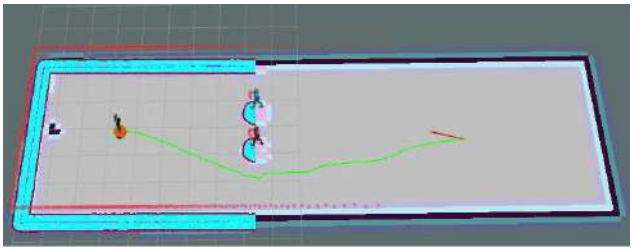
Chúng tôi mô phỏng một môi trường dạng hành lang trong kịch bản đối đầu trực diện với người đi bộ dựa trên bộ mô phỏng Gazrebo Ros và pedestrian ROS Rviz như được minh họa trong Hình 3. Trong kịch bản mô phỏng này, robot được gán một điểm đích được hiển thị dưới dạng một hình trụ màu xanh lá cây ở trung tâm hành lang, và con người di chuyển về phía robot, bắt đầu từ phía đối diện của robot. Trong quá trình di chuyển, robot được yêu cầu không chỉ hướng tới các vị trí mục tiêu mà còn phải giữ khoảng cách an toàn với con người, nhóm người và chướng ngại vật đồng thời còn đảm bảo sự thoải mái cho con người khi robot di chuyển qua.

Mô hình lực xã hội dựa trên dự đoán quỹ đạo và dự đoán va chạm được triển khai dưới dạng một node ROS có tên **cp_sfm**, sử dụng hạ tầng của ROS. Chúng tôi đã thử nghiệm với bộ tham số cho robot được trình bày trong Bảng 1. Các tham số được chọn theo định nghĩa về không gian cá nhân trong mô hình Hall được trình bày trong [16].

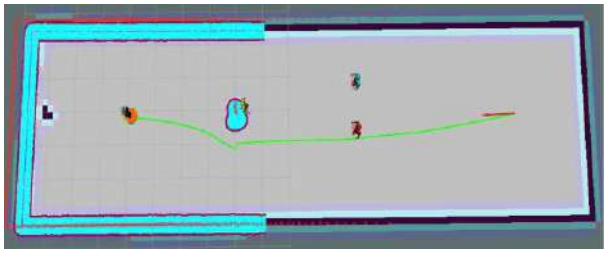
Cụ thể, bán kính của con người r_h và robot r_r được đặt thành 0,25 m, tương ứng với khoảng cách an toàn về mặt vật lý d_p cho con người là 0,5 m. Khoảng cách an toàn về mặt tâm lý d_c cho con người được chọn là 0,9 m. Sự an toàn về mặt tâm lý của con người được đảm bảo nếu giá trị SII nhỏ hơn ngưỡng T_c = 0,14 (khi khoảng cách tương đối giữa con người và robot lớn hơn 0,9 m). Ngược lại, sự an toàn về mặt vật lý của con người bị vi phạm, dẫn đến va chạm giữa robot với con người, nếu giá trị SII lớn hơn ngưỡng tâm lý T_p = 0,54 (khi khoảng cách tương đối giữa chúng nhỏ hơn 0,5 m). Giá trị SGI bằng 1,0 khi robot di chuyển qua tâm của không gian tương tác xã hội. Sự an toàn về mặt tâm lý của con người trong tương tác nhóm hoặc tương tác giữa người và vật được đảm bảo nếu giá trị SGI nhỏ hơn ngưỡng tâm lý nhóm T_g được đặt thành 0,14. Điều đó có nghĩa là robot sẽ phải di chuyển ra ngoài ranh giới tròn của không gian tương tác xã hội.

Bảng 1. Các thông số trong thí nghiệm

Tham số	Giá trị	Tham số	Giá trị	Tham số	Giá trị
v_{max}	1.0[m/s]	a_{max}	0.1[m/s²]	σ_h	1000
φ_{max}	0.5 [rad/s]	ω_{max}	0.1[rad/s²]	σ_v	2.0
d_p	0.54	d_c	0.09 [m]	r_h	0.25 [m]
T_p	0.54	T_c	0.14	T_g	0.14



Hình 4. Mô phỏng kịch bản 1

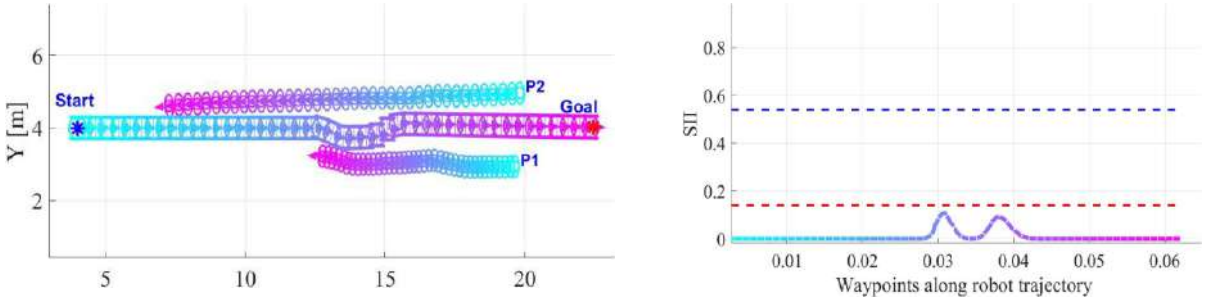


Hình 5. Mô phỏng kịch bản 2

4.2. Kết quả mô phỏng

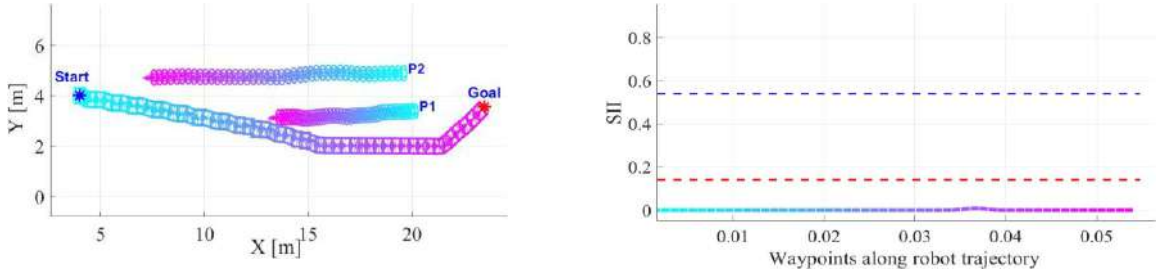
Trong kịch bản thứ nhất hình 4, robot di chuyển trong một hành lang hẹp và đối diện trực tiếp với hai tác nhân P1 và P2 đang đi ngược chiều. Đây là tình huống điển hình dễ phát sinh xung đột không gian do sự chia sẻ vùng di chuyển hạn chế và nguy cơ va chạm trực diện. Hai phương pháp được so sánh gồm SFM gốc và SFM mở rộng có tích hợp dự đoán quỹ đạo và khả năng va chạm.

Với SFM gốc (Hình 6), robot di chuyển gần như theo đường thẳng cho đến khi tiến sát hai tác nhân. Do lực tránh chỉ được kích hoạt khi khoảng cách đủ nhỏ, robot điều chỉnh hướng khá muộn. Điều này làm quỹ đạo xuất hiện các đoạn cong gắt khi né từng tác nhân. Đồ thị SII cho thấy xuất hiện các đỉnh cục bộ tại thời điểm tương tác gần, chứng tỏ robot xâm nhập sâu vào không gian xã hội trước khi tránh. Ngược lại, với SFM mở rộng (Hình 7), robot bắt đầu điều chỉnh hướng sớm hơn nhờ cơ chế dự đoán xung đột. Quỹ đạo tránh rộng và mượt hơn, không có dao động đột ngột. Chỉ số SII duy trì ở mức thấp và không xuất hiện đỉnh lớn, cho thấy robot giữ khoảng cách xã hội tốt hơn trong suốt quá trình di chuyển.



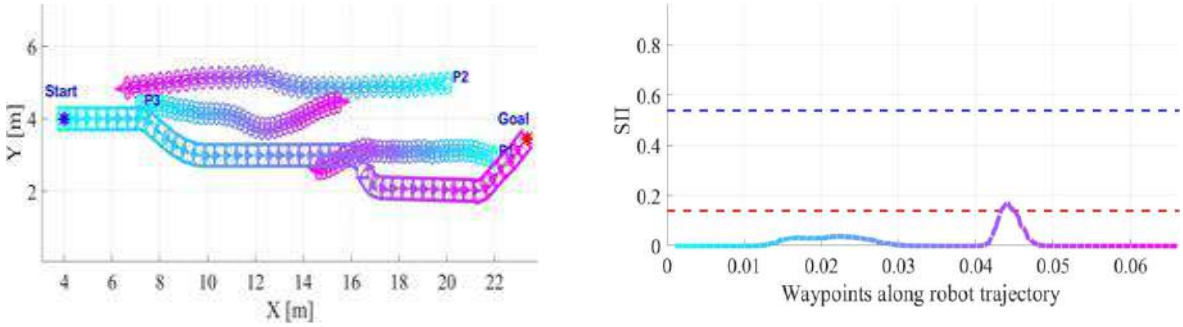
a. Quỹ đạo di chuyển của robot *b. Chi số SII*

Hình 6. Mô phỏng kịch bản 1: Robot đi ngược chiều với 2 tác nhân sử dụng SFM gốc



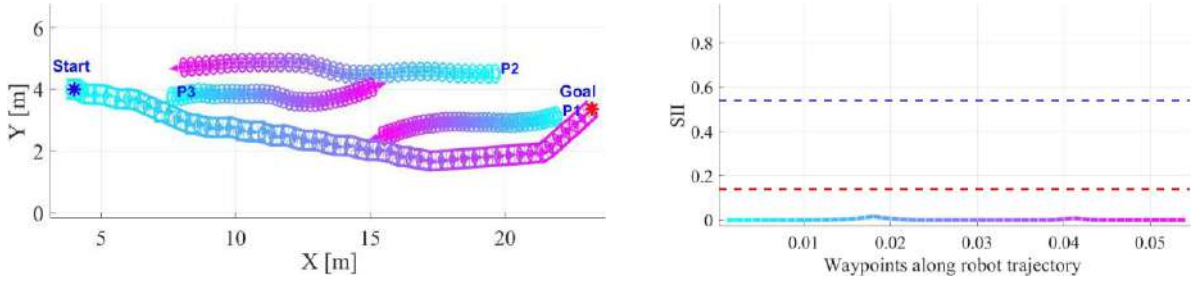
a. Quỹ đạo di chuyển của robot *b. Chi số SII*

Hình 7. Mô phỏng kịch bản 1: Robot đi ngược chiều với 2 tác nhân sử dụng SFM mở rộng



a. Quỹ đạo di chuyển của robot *b. Chi số SII*

Hình 8. Mô phỏng kịch bản 2: Robot đi ngược chiều với 2 tác nhân sử dụng SFM gốc



a. Quỹ đạo di chuyển của robot *b. Chi số SII*

Hình 9. Mô phỏng kịch bản 2: Robot đi ngược chiều với 2 tác nhân sử dụng SFM gốc

Kết quả cho thấy việc bổ sung dự đoán quỹ đạo và va chạm giúp robot chuyển từ tránh phản ứng sang tránh chủ động, giảm mức độ tương tác xã hội cực đại và cải thiện tính tự nhiên của chuyển động trong môi trường hẹp có đối diện trực tiếp.

Trong kịch bản thứ 2 Hình 5, robot di chuyển trong hành lang cùng hướng với một tác nhân (P3) và đồng thời đối diện với hai tác nhân đi ngược chiều (P1, P2). Đây là tình huống phức tạp hơn do robot phải vừa duy trì khoảng cách an toàn với tác nhân phía trước, vừa xử lý xung đột trực diện với hai tác nhân còn lại.

Với SFM mở rộng (Hình 9), robot điều chỉnh quỹ đạo sớm khi phát hiện nguy cơ xung đột với hai tác nhân đi ngược chiều, đồng thời giữ khoảng cách hợp lý với tác nhân cùng hướng phía trước. Quỹ đạo thể hiện sự lệch hướng mềm và liên tục, không xuất hiện dao động đột ngột. Chỉ số SII duy trì ở mức rất thấp trong suốt quá trình di chuyển và không có đỉnh lớn, cho thấy robot không xâm nhập sâu vào không gian xã hội của bất kỳ tác nhân nào. Ngược lại, với SFM gốc (Hình 8), robot phản ứng muộn hơn khi gặp hai tác nhân đối diện. Quỹ đạo xuất hiện các đoạn cong gắt, đặc biệt tại thời điểm giao cắt gần với tác nhân đi ngược chiều. Đồ thị SII cho thấy một đỉnh rõ rệt ở giai đoạn cuối quỹ đạo, phản ánh mức độ tương tác xã hội tăng cao khi robot phải né gấp trong không gian hẹp.

Kết quả cho thấy trong tình huống kết hợp cùng chiều và ngược chiều, SFM mở rộng giúp robot xử lý đồng thời nhiều tương tác một cách chủ động và ổn định hơn. Việc tích hợp dự đoán quỹ đạo và va chạm không chỉ giảm giá trị SII cực đại mà còn cải thiện đáng kể độ mượt và tính tự nhiên của chuyển động so với SFM gốc. SFM gốc đảm bảo tránh va chạm nhưng mang tính phản ứng, robot điều chỉnh muộn khi đã ở gần tác nhân, dẫn đến quỹ đạo cong gắt và xuất hiện đỉnh SII cao. Ngược lại, SFM mở rộng với dự đoán quỹ đạo và va chạm giúp robot tránh chủ động, điều chỉnh sớm hơn, quỹ đạo mượt hơn và duy trì SII ở mức thấp trong mọi tình huống, kể cả khi có nhiều tương tác đồng thời. Do đó, cơ chế dự đoán là yếu tố quan trọng giúp cải thiện độ an toàn xã hội và tính tự nhiên của chuyển động robot trong môi trường có con người.

5. Kết luận

Bài báo đã trình bày một khung điều hướng xã hội cho robot di động dựa trên mô hình Lực xã hội mở rộng, trong đó tích hợp cơ chế dự đoán quỹ đạo và ước lượng va chạm tiềm năng. Kết quả mô phỏng cho thấy phương pháp đề xuất giúp robot chuyển từ cơ chế tránh va chạm phản ứng sang chủ động, tạo quỹ đạo mượt hơn, giảm các điều chỉnh đột ngột và duy trì chỉ số SII ở mức thấp trong các tình huống tương tác trực diện và đa tác nhân. So với SFM gốc, mô hình mở rộng cải thiện rõ rệt mức độ an toàn và sự thoải mái xã hội trong môi trường hành lang hẹp.

Trong các nghiên cứu tiếp theo, chúng tôi sẽ mở rộng mô hình cho môi trường mật độ người cao hơn, tích hợp học máy để dự đoán ý định chuyển động phức tạp của con người, đồng thời đánh giá trên nền tảng robot thực trong các không gian công cộng nhằm kiểm chứng tính tổng quát và độ ổn định của hệ thống.

Tài liệu tham khảo

[1] Hart, Peter E., Nils J. Nilsson, and Bertram Raphael. "A formal basis for the heuristic determination of minimum cost paths." *IEEE transactions on Systems Science and Cybernetics*.

[2] Stentz, Anthony. "Optimal and efficient path planning for partially-known environments." *Proceedings of the IEEE international conference on robotics and automation*. IEEE.

[3] Koenig, S., and M. Likhachev. "D* lite." *Aaai/iaai* 15 (2002).
[4] Kruse, Thibault, et al. "Human-aware robot navigation: A survey." *Robotics and Autonomous Systems* .
[5] Kirby, Rachel, Reid Simmons, and Jodi Forlizzi. "Companion: A constraint-optimizing method for person-acceptable navigation." *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 2009.
[6] Helbing, Dirk, and Peter Molnar. "Social force model for pedestrian dynamics." *Physical review*.
[7] Svenstrup, Mikael, Thomas Bak, and Hans Jørgen Andersen. "Trajectory planning for robots in dynamic human environments." *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2010.
[8] Fiorini, Paolo, and Zvi Shiller. "Motion planning in dynamic environments using velocity obstacles." *The international journal of robotics research*.
[9] Van den Berg, Jur, Ming Lin, and Dinesh Manocha. "Reciprocal velocity obstacles for real-time multi-agent navigation." *2008 IEEE international conference on robotics and automation*. Ieee, 2008..
[10] Van Den Berg, Jur, et al. "Optimal reciprocal collision avoidance for multi-agent navigation." *Proc. of the IEEE International Conference on Robotics and Automation, Anchorage (AK), USA*. 2010.11.
[11] Alahi, Alexandre, et al. "Social lstm: Human trajectory prediction in crowded spaces." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
[12] Y. F. Chen et al, “Socially aware motion planning with deep reinforcement learning,” *IROS*, 2017.
[13] C. Rösmann et al, “Efficient trajectory optimization using a sparse model” *ECMR*, 2013.
[14] T. Rösmann et al., “Integrating social constraints in local planning” *ICRA Workshop*, 2015.
[15] Hall, E. T. "T7he Hidden Dimension: Man's Use of Space in Public and Private." *The Bodley Head Ltd, London, UK* (1966).

SOCIAL NAVIGATION FRAMEWORK FOR MOBILE ROBOTS IN A DYNAMIC SOCIAL ENVIRONMENT

Abstract*: In environments where humans are present, robot navigation is not limited to obstacle avoidance but must also ensure safety and comfort during human–robot interaction. Therefore, navigation methods that incorporate social awareness play a crucial role in deploying mobile service robots in environments involving direct or indirect interaction with people. These approaches aim to generate motion trajectories that comply with social norms, taking into account human behavior, movement patterns, and implicit social cues in shared spaces. In this study, we propose a social navigation architecture for mobile service robots that enables the robot to achieve its navigation goals while maintaining safety and comfort for nearby humans. A key contribution of this work is a navigation framework designed to operate effectively in previously unknown human-interactive environments. To achieve this, we adopt a variant of the pedestrian model, specifically the Social Force Model enhanced with trajectory prediction and collision prediction mechanisms. This model is tailored for low- to medium-density environments and considers the states and motion trajectories of tracked individuals throughout the robot’s navigation process. Within the proposed system, this model functions as a local planner, generating feasible and socially compliant trajectories for the robot in indoor scenarios, such as corridor-like environments. The entire architecture is implemented as ROS nodes and evaluated through both real-world experiments and simulations. Experimental results demonstrate that the proposed approach operates effectively and satisfies the intended requirements.*

Keywords: *Social Focre Model (SFM)., Velocity Obstacles (VO), Reciprocal Velocity Obstacles(RVO*

Xây dựng mô hình nhận dạng rung nhĩ nhẹ cho thiết bị đeo sử dụng thuật toán stacking

Vương Trí Tiếp[1], Nguyễn Trọng Tuyến[1*], Nguyễn Mạnh Cường[1*]

[1]Viện Tên lửa và kỹ thuật điều khiển, Học viện Kỹ thuật Quân sự

Email: tuyen.nguyen@lqdtu.edu.vn, nmcuong869@lqdtu.edu.vn ; Tel: 0967904423

Tóm tắt

Rung nhĩ (Atrial Fibrillation - AF) là dạng rối loạn nhịp tim kéo dài phổ biến nhất, làm tăng đáng kể nguy cơ đột quỵ và tử vong. Rung nhĩ kịch phát (PAF) có tính chất thoáng qua, khởi phát ngắn và tự kết thúc, gây ra thách thức chẩn đoán lớn đối với việc theo dõi điện tâm đồ (ECG) ngắn hạn tiêu chuẩn. Trong nghiên cứu này, bài báo đề xuất một mô hình sử dụng bộ đặc trưng đa miền có thể diễn giải ứng dụng học kết hợp (ensemble learning). Khác với các mô hình học sâu với khả năng giải thích kém, bài báo tích hợp các dấu ấn sinh học đa miền có tính giải thích cao, bao gồm: chỉ số hình thái khoảng TQ được đề xuất để nắm bắt hoạt động của tâm nhĩ, các biến thiên phổ tần số và các số liệu biến thiên nhịp tim (HRV) tiêu chuẩn. Mô hình được huấn luyện trên tập dữ liệu MIT-AFDB và được đánh giá nghiêm ngặt theo cả hai phương thức nội bệnh nhân (intra-patient) và liên bệnh nhân (inter-patient), đạt độ chính xác lần lượt là 96,70% và 95,58%. Kết quả khẳng định rằng mô hình nhẹ và khả năng diễn giải cao này đặc biệt phù hợp để sàng lọc AF liên tục, theo thời gian thực trên các thiết bị đeo.

Từ khóa: *Rung nhĩ kịch phát; Biến thiên nhịp tim; Hoạt động nhĩ; Thuật toán stacking.*

1. Giới thiệu

Rung nhĩ (AF) hiện được ghi nhận là dạng rối loạn nhịp tim kéo dài phổ biến nhất trên toàn cầu, tạo ra gánh nặng nghiêm trọng đối với hệ thống y tế công cộng. Dữ liệu thống kê chỉ ra rằng có khoảng 4,5 triệu ca mắc mới mỗi năm, với tổng tỷ lệ hiện mắc ước tính khoảng 60 triệu người. AF làm tăng nguy cơ đột quỵ lên gấp 3 đến 5 lần và là nguyên nhân hàng đầu gây suy tim cũng như kéo dài thời gian nằm viện. Dù phổ biến, chẩn đoán AF vẫn là một thách thức lớn vì khoảng 50-87% bệnh nhân hoàn toàn không có triệu chứng. Hơn nữa, AF thường biểu hiện dưới dạng Rung nhĩ kịch phát (PAF), đặc trưng bởi các đợt ngắn, tự kết thúc và không thể đoán trước. Bản chất thoáng qua này khiến việc phát hiện qua Holter điện tâm đồ ngắn hạn trở nên khó khăn, dễ dẫn đến bỏ sót chẩn đoán và trì hoãn điều trị.

Điện tâm đồ (ECG) vẫn là tiêu chuẩn vàng trong chẩn đoán AF. Về mặt lâm sàng, AF được nhận diện qua sự không đều của các khoảng R-R và sự vắng mặt của các sóng P rõ ràng, thay vào đó là các sóng rung (f) nhanh, có tần số và hình dạng khác nhau. Dựa trên các đặc điểm này, nhiều nghiên cứu đã phát triển các mô hình phát hiện tự động bằng thuật toán học máy truyền thống như Support Vector Machines (SVM), k-Nearest Neighbors (KNN), Mạng thần kinh nhân tạo (ANN) dựa trên các đặc trưng kỹ thuật thủ công. Các phương pháp này chủ yếu sử dụng biến thiên nhịp tim (HRV) và đặc trưng phổ [1, 5, 12, 13].

Gần đây, sự phát triển của phần cứng đã thúc đẩy việc ứng dụng học sâu trong phân loại

rối loạn nhịp tim. Các kiến trúc như mạng tích chập (CNN), bộ nhớ dài ngắn hạn (LSTM) và Transformer đã đạt được hiệu suất ấn tượng. Tuy nhiên, phương pháp tiếp cận học sâu gặp phải những hạn chế đáng kể. Các mô hình này đòi hỏi tài nguyên phần cứng lớn và độ phức tạp tính toán cao[4-6], dẫn đến tính di động kém trên các thiết bị đeo có giới hạn về pin. Hơn nữa, các mô hình học sâu hoạt động như những "hộp đen" thiếu tính diễn giải - một yếu tố mang tính quyết định trong y khoa. Đáng chú ý, khi được xác thực trên các bộ dữ liệu bên ngoài có sự thay đổi về phân phối dữ liệu, các mô hình học sâu thường bị suy giảm hiệu suất nghiêm trọng do hiện tượng quá khớp trên tập dữ liệu ngoài [1].

Để giải quyết những thách thức này, nghiên cứu của chúng tôi đề xuất một mô hình chẩn đoán AF sử dụng bộ đặc trưng đa miền hoàn toàn mới có thể diễn giải được và học máy kết hợp. Bộ đặc trưng tích hợp các chỉ số HRV tiêu chuẩn với các chỉ số hình thái học mới đại diện cho hoạt động của tâm nhĩ (tập trung vào biến động khoảng TQ) và các mô tả phương sai phổ. Bằng cách triển khai thuật toán xếp chồng có phương sai thấp, hệ thống cung cấp một giải pháp chẩn đoán sàng lọc ổn định, dễ diễn giải và có hiệu suất tính toán cực cao.

2. Phương pháp

2.1 Tổng quan đề xuất

Nghiên cứu đề xuất một mô hình chẩn đoán AF sử dụng tín hiệu ECG đạo trình đơn. Lõi của phương pháp là việc trích xuất bộ đặc trưng đa miền bao gồm: (1) Biến thiên Nhịp tim (HRV), (2) Các chỉ số hình thái hoạt động tâm nhĩ (từ khoảng TQ), và (3) Độ ổn định phổ tần số. Để nắm bắt hiệu quả các đợt PAF ngắn, tín hiệu được phân tích qua giao thức cửa sổ trượt chồng lấp. Tín hiệu được cắt thành các cửa sổ 15 giây với bước nhảy 1.2 giây. Chiến lược này đảm bảo độ phân giải thời gian tối đa để bắt kịp các rối loạn nhịp chớp nhoáng, đồng thời cho phép đưa ra quyết định gần như ngay lập tức, phù hợp cho thiết bị y tế đeo bám.

2.2 Tiền xử lý Tín hiệu

Tín hiệu ECG thô thường xuyên lẫn tạp nhiễu bởi nhiều nguồn nhiễu động vật lý và điện từ. Để loại bỏ nhiễu mà vẫn giữ nguyên hình thái sóng, chúng tôi áp dụng hai bộ lọc. Thứ nhất, bộ lọc trung vị được áp dụng để loại bỏ hiện tượng trôi đường cơ sở (baseline wander) do hô hấp hoặc chuyển động cơ thể. Thứ hai, bộ lọc Savitzky-Golay được triển khai để làm mịn các nhiễu tần số cao (như nhiễu điện cơ EMG). Bộ lọc này được chọn vì khả năng hồi quy đa thức cục bộ giúp triệt tiêu nhiễu mà vẫn bảo tồn nghiêm ngặt độ sắc nét của phức bộ QRS.

2.3 Trích xuất đặc trưng đa miền

Định vị điểm tham chiếu: Thuật toán Pan-Tompkins được sử dụng để phát hiện đỉnh R nhờ tính hiệu quả tính toán ấn tượng và khả năng chịu nhiễu tốt.[1] Sau đó, điểm Q được tìm thấy trong cửa sổ 60ms trước đỉnh R. Việc phát hiện điểm kết thúc sóng T (T_{end}) trong môi trường AF rất khó khăn do nhiễu sóng f, do đó, thuật toán Zhang dựa trên tính lồi toán học và diện tích dưới đường cong được áp dụng để xác định tọa độ chính xác của T_{end} []. Để số liệu HRV đáng tin cậy, các nhịp ngoại tâm thu (PVCs, PACs) được loại bỏ nếu khoảng R-R thỏa mãn điều kiện $RR_i < 0.8RR_{mean}$.

Đặc trưng Biến thiên Nhịp tim (HRV): Tính bất thường của nhịp thất là dấu hiệu then chốt của AF. Các chỉ số tuyến tính được trích xuất bao gồm:

$$RR_{mean} = \frac{\sum_{i=1}^{N} RR_i}{N}$$

$$dRR = RR_{max} - RR_{min} \tag{1}$$

$$dispRR = \sqrt{\frac{\sum_{i=1}^{N}(RR_i - RR_{mean})^2}{N-1}}$$

Bên cạnh đó, các tham số phi tuyến từ biểu đồ Poincaré được tính toán để mô tả động lực học hỗn loạn:

$$SD_1 = \sqrt{\frac{Var(RR_{n+1} - RR_n)}{2}}$$

$$SD_2 = \sqrt{\frac{Var(RR_{n+1} + RR_n)}{2}} \tag{2}$$

$$A_{SD} = \pi SD_1 SD_2$$

Trong đó, SD_1 biểu thị sự thay đổi nhịp ngắn hạn, SD_2 biểu thị biến thiên dài hạn và A_{SD} là tổng diện tích bề mặt hình elip.

Đặc trưng hoạt động nhĩ: Chúng tôi đề xuất chỉ số tương đồng hình thái (I_{TQ}) sử dụng khoảng cách Euclidean. Các đoạn TQ được cô lập, nội suy đồng nhất, và tính toán độ biến thiên:

$$\overline{TQ} = \frac{\sum_{i=1}^{N} TQ_i}{N} \tag{3}$$

$$R_{TQ}^{i} = \frac{\|TQ_i - \overline{TQ}\|}{\|TQ_i\|} \tag{4}$$

$$I_{TQ} = \frac{N(R_{TQ}^{i} \geq \epsilon)}{N} \tag{5}$$

Trong đó, $\overline{TQ}$ là vectơ trung bình, và $N(R_{TQ}^{i} \geq \epsilon)$ đếm số lượng khoảng có khoảng cách Euclidean tương đối vượt quá ngưỡng $\epsilon = 0.8$. Giá trị I_{TQ} cao phản ánh sự mất ổn định hình thái sâu sắc do các sóng nhĩ (sóng f) tạo ra [1].

Đặc trưng phổ công suất: Để tăng cường khả năng hoạt động khi nhiễu làm mất đỉnh R, chúng tôi sử dụng phân tích mật độ phổ công suất (PSD). Phổ năng lượng trong AF bị phân tán hỗn loạn thay vì có cấu trúc sóng hài rõ ràng. Dựa vào tập hợp các đỉnh quang phổ $P = \{f_1, f_2, \dots, f_M\}$, chúng tôi điều chỉnh hình học Poincaré cho miền tần số bằng cách tính độ lệch giữa các đỉnh kề nhau $\delta_i^{PS} = f_{i+1} - f_i$. Các bộ mô tả Poincaré Phổ được tính toán:

$$\sigma_p^{PS} = Var\left(\delta_i^{PS}\right) = \sqrt{\frac{1}{M-1}\sum_{i=1}^{M}(\delta_i^{PS} - \overline{\delta^{PS}})^2} \tag{6}$$

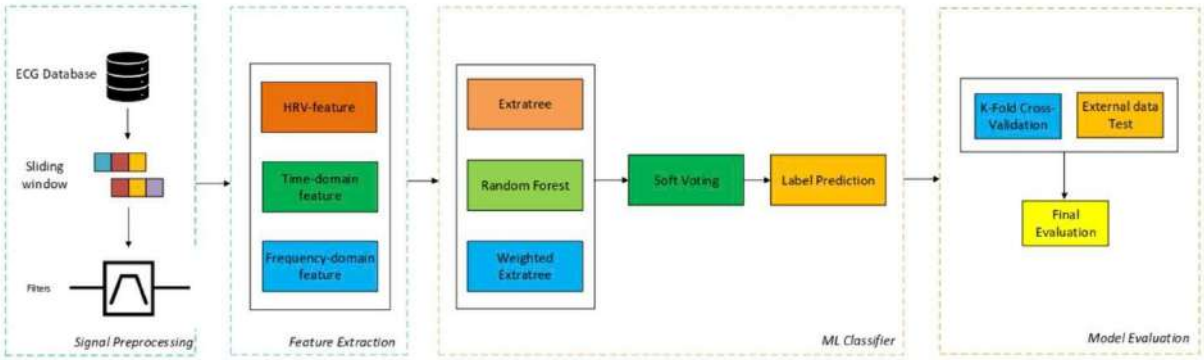
$$\sigma_{SD1} = \sqrt{\frac{Var(\delta_{n+1}^{PS} - \delta_n^{PS})}{2}} \tag{7}$$

$$\sigma_{SD2} = \sqrt{\frac{Var(\delta_{n+1}^{PS} + \delta_n^{PS})}{2}} \quad (8)$$

$$A_\sigma = \pi \sigma_{SD1} \sigma_{SD2} \quad (9)$$

2.4 Mô hình Phân loại Học Kết hợp (Ensemble Model)

Vectơ đặc trưng 11 chiều được đưa vào kiến trúc học kết hợp xếp chồng (stacking ensemble) có phương sai thấp (Hình 1). Hệ thống gồm ba mô hình cơ sở: ExtraTrees (tăng cường sự đa dạng ranh giới phân loại, thúc đẩy độ nhạy), Random Forest và Weighted ExtraTrees. Để giảm dương tính giả, mô hình Weighted ExtraTrees được gán tỷ lệ trọng số 2:1 cho lớp (Non-AF : AF). Các dự đoán được tổng hợp qua cơ chế bỏ phiếu mềm (soft-voting mechanism).



Hình 1 Mô hình Phân loại Học Kết hợp (Ensemble Model)

3. Thiết lập thực nghiệm

3.1. Cơ sở Dữ liệu: Dữ liệu sử dụng là Cơ sở Dữ liệu Rung nhĩ MIT-BIH (MIT-AFDB) gồm 23 bản ghi Holter 6-10 giờ chất lượng cao (loại bỏ 2 bản lỗi). Tín hiệu được lấy mẫu ở tần số 250 Hz. Chúng tôi chỉ sử dụng kênh ECG đạo trình thứ nhất để mô phỏng điều kiện thiết bị đeo. Một phân đoạn 15s được gắn nhãn AF (1) nếu sóng AF chiếm từ 50% tổng chiều dài trở lên, ngược lại là Non-AF (0). Quá trình này tạo ra 685.771 phân đoạn, trong đó, có 413.379 phân đoạn Non-AF và 272.032 phân đoạn AF.

3.2 Giao thức đánh giá: Hai hệ thống giao thức xác thực độc lập được áp dụng: Liên bệnh nhân (Subject-independent), sử dụng kiểm chứng chéo 10 lần (10-fold cross validation) tách biệt hoàn toàn định danh bệnh nhân giữa tập huấn luyện và kiểm thử và nội bộ bệnh nhân (Subject-dependent) dùng để trộn lẫn ngẫu nhiên mọi phân đoạn của các bệnh nhân. Mục đích là để tìm ra năng lực lý thuyết cực đại và làm mốc so sánh.

4. Kết quả và thảo luận

4.1 Hiệu suất trên các giao thức: Theo giao thức liên bệnh nhân nghiêm ngặt 10-fold CV trên 685.771 mẫu dữ liệu, mô hình đạt độ nhạy (Se) 95,75%, độ đặc hiệu (Sp) 95,47% và độ chính xác (Acc) 95,58%. Diện tích dưới đường cong (AUC) đạt mức trung bình 0.985. Kết quả này minh chứng sự ổn định toán học tổng quát của thuật toán, duy trì tỷ lệ dương tính giả và âm tính giả ở mức rất thấp (dưới 5%).

Khi chuyển đổi giữa giao thức đánh giá liên bệnh nhân và nội bộ bệnh nhân, mô hình thể hiện độ ổn định. Sự suy giảm hiệu suất khi chuyển sang dữ liệu bệnh nhân hoàn toàn xa lạ (liên bệnh nhân) là cực kỳ hẹp: Độ nhạy chỉ giảm 1,48%, Độ đặc hiệu giảm 0,89% và Tổng độ chính xác chỉ giảm 1,12%. Điều này khẳng định kiến trúc đa miền đã thực sự trích xuất được động học sinh lý nội tại của PAF thay vì bị "học vẹt" (overfitting) các đặc trưng nhân dạng dư thừa của từng cá thể như thường thấy ở các mô hình Deep Learning sâu.

4.2. So sánh với các mô hình khác: Trong phần này, hiệu suất của mô hình đề xuất được so sánh với các nghiên cứu trước đây về phân loại và nhận dạng Rung nhĩ (AF - Atrial Fibrillation) ở cấp độ phân đoạn (segment-level), sử dụng cả phương pháp học máy truyền thống và học sâu. Một thách thức lớn trong các thuật toán hiện tại là khả năng tổng quát hóa kém; cụ thể, hiệu suất thường giảm sút đáng kể khi chuyển đổi từ mô hình đánh giá nội bệnh nhân (intra-patient / phụ thuộc vào đối tượng) sang mô hình liên bệnh nhân (inter-patient / độc lập với đối tượng). Sự suy giảm này càng trở nên nghiêm trọng hơn khi các mô hình được kiểm chứng trên các cơ sở dữ liệu độc lập và có quy mô lớn hơn.

Đáng chú ý, phần lớn các nghiên cứu ở cấp độ phân đoạn sử dụng phương pháp kiểm chứng nội bệnh nhân, điều này có thể dẫn đến rò rỉ dữ liệu (data leakage) — nơi các phân đoạn từ cùng một đối tượng xuất hiện trong cả tập huấn luyện và tập kiểm tra. Hiện tượng này thường dẫn đến các chỉ số hiệu suất cao một cách phi thực tế. Để đảm bảo việc đánh giá được chặt chẽ và khách quan, chúng tôi chủ yếu tập trung so sánh kết quả của mình với các nghiên cứu áp dụng mô hình liên bệnh nhân (độc lập với đối tượng).

Nhiều mô hình thống kê hiện tại phụ thuộc rất nhiều vào các đặc trưng của khoảng RR (RR-interval). Tuy nhiên, trong môi trường có độ nhiễu cao hoặc trong các trường hợp biên độ sóng R tương đương hoặc nhỏ hơn sóng T, các thuật toán phát hiện đỉnh R có thể trở nên thiếu tin cậy, dẫn đến sai số dự đoán. Để giảm thiểu rủi ro này, các nhà nghiên cứu như Andersen và cộng sự, Faust và cộng sự, cùng Abhimanyu và cộng sự đã sử dụng các cửa sổ quan sát dài hơn, thường dao động từ 20 đến 100 nhịp tim. Andersen và cộng sự [4]đã sử dụng mạng LSTM với các cửa sổ 30 nhịp tim không chồng chéo, đạt độ nhạy (Se) 98,17% và độ đặc hiệu (Sp) 96,29% dưới phương pháp kiểm chứng chéo liên bệnh nhân 5 nếp gấp (5-fold cross-validation). Faust và cộng sự [5,6, 12, 13] đã báo cáo độ chính xác (Acc) liên bệnh nhân và nội bệnh nhân lần lượt là 98,51% và 99,77%, sử dụng cửa sổ trượt 100 nhịp tim kết hợp với LSTM. Abhimanyu và cộng sự [10] đã trích xuất 32 đặc trưng đa miền từ các khoảng RR bằng cách sử dụng cửa sổ trượt 31 nhịp tim. Mặc dù họ đạt được Acc là 99,46% ở chế độ nội bệnh nhân, các chỉ số của họ đã giảm từ 4,24% đến 7,04% khi thực hiện kiểm chứng mù (Blindfold validation - LOGO-CV). Masud Shah Jahan và cộng sự đã đạt Acc 88,0% thông qua LOGO-CV bằng cách sử dụng các đặc trưng thống kê của khoảng RR và Adaboost trên các cửa sổ 20 nhịp tim. Peng Xiong và cộng sự [9] đã sử dụng một bộ bảy đặc trưng khoảng RR trên các cửa sổ 10 giây không chồng chéo, ghi nhận mức giảm Acc 2,81% giữa hai mô hình đánh giá. Pereira và Andreao [10] đã áp dụng LSTM trực tiếp lên các tín hiệu điện tâm đồ (ECG) thô dài 10 giây, mang lại Acc là 90,87%. Yanrui Jin và cộng sự [11] đã tích hợp CNN, LSTM và cơ chế Attention trên các cửa sổ 5 giây, báo cáo mức giảm độ đặc hiệu lên tới 5,77% khi chuyển sang mô hình liên bệnh nhân.

Trái ngược với các nghiên cứu này, phương pháp được chúng tôi đề xuất sử dụng cửa sổ

trượt 15 giây với bước trượt (stride) ngắn 1,2 giây. Cấu hình này mang lại độ phân giải thời gian cao để phát hiện các đợt AF ngắn mà không bỏ sót bất kỳ chu kỳ rối loạn nhịp nào. Một điểm mạnh nổi bật của mô hình chúng tôi là sự chênh lệch hiệu suất cực kỳ hẹp giữa hai mô hình nội bệnh nhân và liên bệnh nhân trên tập dữ liệu MIT-AFDB (ΔSe: 1,48%, ΔSp: 0,89% và ΔAcc: 1,12%). Sự ổn định này khẳng định rằng mô hình của chúng tôi đã nắm bắt thành công các dấu ấn sinh học sinh lý cốt lõi của AF, thay vì chỉ học vẹt các đặc điểm nhận dạng riêng biệt của từng bệnh nhân.

5. Kết luận

Mô hình dựa trên sự kết hợp trích xuất bộ đặc trưng đa miền (HRV, Hình thái TQ, Phổ Poincaré) với kiến trúc cây quyết định Ensemble đã giải quyết thành công bài toán chẩn đoán rung nhĩ kịch phát. Thuật toán đạt mức độ chính xác 95,58% đối với bệnh nhân hoàn toàn lạ, với biên độ sụt giảm hiệu năng cực thấp (1,12%) khi so sánh với thử nghiệm nội bộ. Mô hình có mức độ phức tạp tính toán nhẹ và tính diễn giải rõ ràng của các chỉ số sinh lý cho phép áp dụng mô hình này vào các thiết bị biên, sẵn sàng tích hợp vào smartwatch hoặc thiết bị y tế đeo bám theo dõi và giám sát liên tục, cung cấp cảnh báo kịp thời cho bệnh nhân tim mạch nguy cơ cao.

Tài liệu tham khảo

1. Sangle, K. C., et al. (2020). A review of atrial fibrillation detection methods as a service. *Cureus, 12*(5), e8110. doi:10.7759/cureus.8110

2. Yang, H., et al. (2021). Detection of brief episodes of atrial fibrillation based on electrocardiomatrix and convolutional neural network. *Frontiers in Physiology, 12*, 673819. doi:10.3389/fphys.2021.673819

3. Mäkynen, M., et al. (2024). Compressed deep learning models for wearable atrial fibrillation detection through attention. *Sensors*.

4. Andersen, R. S., Peimankar, A., & Puthusserypady, S. (2019). A deep learning approach for detection of atrial fibrillation using RR intervals. *Proceedings of the 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 334–337). Berlin, Germany. doi:10.1109/EMBC.2019.8857155

5. Faust, O., Hagiwara, Y., Hong, T. J., Lih, O. S., & Acharya, U. R. (2018). Deep learning for healthcare applications based on physiological signals: A review. *Computer Methods and Programs in Biomedicine, 161*, 1–13. doi:10.1016/j.cmpb.2018.04.006

6. Faust, O., et al. (2019). Automated detection of atrial fibrillation using long short-term memory network. *Knowledge-Based Systems, 176*, 38–45. doi:10.1016/j.knosys.2018.10.020

7. Abhimanyu, K., et al. (2020). Atrial fibrillation detection using multi-domain features of RR intervals. *Biomedical Signal Processing and Control, 57*, 101927. doi:10.1016/j.bspc.2020.101927

8. Jahan, M. S., et al. (2021). Atrial fibrillation detection from RR intervals using AdaBoost classifier. *Computers in Biology and Medicine, 129*, 104622. doi:10.1016/j.compbiomed.2021.104622

9. Xiong, P., et al. (2017). Automated detection of atrial fibrillation using RR interval-based features. *Physiological Measurement, 38*(4), 594–608. doi:10.1088/1361-6579/aa6f13

10. Pereira, T., & Andreão, R. V. (2018). Atrial fibrillation detection using a deep learning approach. *Computers in Biology and Medicine, 102*, 1–10. doi:10.1016/j.compbiomed.2018.07.003

11. Jin, Y., et al. (2020). An attention-based CNN-LSTM model for atrial fibrillation detection. *IEEE Access, 8*, 68864–68874. doi:10.1109/ACCESS.2020.2973058

12. Kraft, D., et al. (2025). Atrial fibrillation and atrial flutter detection using deep learning. *Sensors*.

13.Edouard, P., et al. (2025). Design and validation of Withings ECG Software 2, a tiny neural network based algorithm for detection of atrial fibrillation. *Computers in Biology and Medicine*.
14.

Development of a Lightweight Atrial Fibrillation Detection Model for Wearable Devices Using Stacking Ensemble Learning

Abstract: *Atrial Fibrillation (AF) represents the most prevalent sustained cardiac arrhythmia, significantly elevating the risk of stroke and mortality. Paroxysmal Atrial Fibrillation (PAF), characterized by its transient nature with brief onsets and spontaneous termination, poses a formidable diagnostic challenge for standard short-term electrocardiogram (ECG) monitoring. In this study, we propose an interpretable ensemble learning model utilizing a multi-domain feature set. Unlike deep learning models often criticized for their "black-box" nature, this approach integrates highly explainable multi-domain biomarkers, including: a proposed TQ-interval morphology index designed to capture atrial activity, spectral frequency variations, and standard heart rate variability (HRV) metrics. The model was trained on the MIT-BIH Atrial Fibrillation Database (MIT-AFDB) and rigorously evaluated under both intra-patient and inter-patient paradigms, achieving accuracies of 96.70% and 95.58%, respectively. These results demonstrate that the proposed lightweight and highly interpretable model is particularly well-suited for continuous, real-time AF screening on wearable devices.*

Keywords: Paroxysmal Atrial Fibrillation; Heart Rate Variability; Atrial Activity; Stacking algorithm.

TUYỂN TẬP CÔNG TRÌNH HỘI NGHỊ KHOA HỌC CÁC NHÀ NGHIÊN CỨU TRẺ LẦN THỨ XXI – NĂM 2026

NHÀ XUẤT BẢN MONG BẠN ĐỌC GÓP Ý KIẾN, PHÊ BÌNH

Chịu trách nhiệm xuất bản: **Giám đốc - Tổng Biên tập PHẠM VĂN TRƯỜNG**
Chịu trách nhiệm bản thảo: **Phó Giám đốc - Phó Tổng Biên tập NGUYỄN VĂN SÁU**
Biên tập: **HOÀNG THẾ LONG, NGUYỄN HOÀNG TUẤN**
Kỹ thuật CNTT và bìa: **HỌC VIỆN KỸ THUẬT QUÂN SỰ**

NHÀ XUẤT BẢN QUÂN ĐỘI NHÂN DÂN
23 Lý Nam Đế, phường Hoàn Kiếm, TP. Hà Nội
Website: http://nxbqdnd.com.vn
http://xbqdnd.bqp; https://xbqdnd.vn
Email: nxbqdnd@nxbqdnd.com.vn
ĐT: 024.38455766 - 024.37470780; Fax: 024.37471106
Chi nhánh tại Thành phố Hồ Chí Minh
161-163 Trần Quốc Thảo, phường Nhiêu Lộc
ĐT: 069.667452 - 028.62565588; Fax: 028.62565588
Cơ quan đại diện tại Thành phố Đà Nẵng
Số 172 đường 2-9, phường Bình Thuận
ĐT: 0236.6250803; Fax: 0236.6250803
Cơ quan đại diện tại Thành phố Cần Thơ
Phi trường 31 - đường Cách mạng Tháng Tám
ĐT: 069.629905 - 0292.3814772; Fax: 0292.3814772

In xong: Quý II-2026 Nộp lưu chiểu: Quý II-2026
Khổ sách: 19x27 Số trang: 3350 trang Số lượng: 50 đĩa
Số đăng ký kế hoạch xuất bản: 1387-2026/CXBIPH/3-21/QĐND
Số quyết định xuất bản: 94/QĐLKI-NXBQĐND, ngày 15 tháng 4 năm 2026
Kỹ thuật CNTT tại: Học viện Kỹ thuật quân sự
In và ghi đĩa: Xưởng in Học viện Kỹ thuật quân sự
Địa chỉ: 236 Hoàng Quốc Việt, phường Nghĩa Đô, TP. Hà Nội
Đối tác liên kết: Học viện Kỹ thuật quân sự
Địa chỉ: 236 Hoàng Quốc Việt, phường Nghĩa Đô, TP. Hà Nội
ISBN: 978-632-06-2010-0.

HỌC VIỆN KỸ THUẬT QUÂN SỰ
MILITARY TECHNICAL ACADEMY