

BỘ GIÁO DỤC VÀ ĐÀO TẠO

BỘ QUỐC PHÒNG

HỌC VIỆN KỸ THUẬT QUÂN SỰ

LÊ HÀ KHÁNH

NGHIÊN CỨU NÂNG CAO HIỆU NĂNG PHÂN LOẠI
ĐIỀU CHẾ TỰ ĐỘNG ỨNG DỤNG MẠNG NƠ-RON

LUẬN ÁN TIẾN SĨ KỸ THUẬT

HÀ NỘI - NĂM 2026

BỘ GIÁO DỤC VÀ ĐÀO TẠO

BỘ QUỐC PHÒNG

HỌC VIỆN KỸ THUẬT QUÂN SỰ

LÊ HÀ KHÁNH

NGHIÊN CỨU NÂNG CAO HIỆU NĂNG PHÂN LOẠI
ĐIỀU CHẾ TỰ ĐỘNG ỨNG DỤNG MẠNG NƠ-RON

LUẬN ÁN TIẾN SĨ KỸ THUẬT

Chuyên ngành: KỸ THUẬT ĐIỆN TỬ
Mã số: 9 52 02 03

NGƯỜI HƯỚNG DẪN KHOA HỌC:
TS ĐOÀN VĂN SÁNG
PGS. TS HOÀNG VĂN PHÚC

HÀ NỘI - NĂM 2026

LỜI CAM ĐOAN

Tôi xin cam đoan Luận án và các kết quả trình bày trong luận án là công trình nghiên cứu của tôi dưới sự hướng dẫn của cán bộ hướng dẫn. Các số liệu, kết quả trình bày trong luận án là hoàn toàn trung thực và chưa được công bố trong bất kỳ công trình nào trước đây. Các kết quả sử dụng tham khảo đều đã được trích dẫn đầy đủ và theo đúng quy định.

Hà Nội, ngày tháng năm 2026

Tác giả

Lê Hà Khánh

LỜI CẢM ƠN

Trong quá trình học tập, nghiên cứu và hoàn thành luận án, nghiên cứu sinh đã nhận được nhiều sự giúp đỡ và đóng góp quý báu.

Trước hết nghiên cứu sinh xin bày tỏ lòng cảm ơn sâu sắc đến tập thể thầy giáo hướng dẫn PGS.TS Hoàng Văn Phúc và TS Đoàn Văn Sáng đã tận tình hướng dẫn, giúp đỡ nghiên cứu sinh trong suốt quá trình nghiên cứu khoa học và hoàn thành luận án.

Nghiên cứu sinh cũng chân thành cảm ơn các thầy giáo tập thể Bộ môn Lý thuyết mạch - Đo lường, Khoa Vô tuyến điện tử, tập thể Bộ môn Công nghệ điện tử, Viện Tích Hợp Hệ Thống, Học viện Kỹ thuật Quân sự, nơi nghiên cứu sinh làm việc, đã quan tâm và tận tình giúp đỡ và tạo điều kiện mọi mặt trong suốt thời gian nghiên cứu sinh học tập, nghiên cứu tại đây.

Nghiên cứu sinh cũng chân thành cảm ơn Phòng Sau đại học; Hệ Quản lý học viên sau đại học - Học viện KTQS; Trường Sĩ quan thông tin - Binh chủng Thông tin liên lạc là các đơn vị chủ quản đã thường xuyên hỗ trợ, tạo điều kiện và giúp đỡ nghiên cứu sinh hoàn thành luận án này.

Cuối cùng, nghiên cứu sinh trân trọng cảm ơn vợ, con, người thân trong gia đình, các đồng nghiệp đã luôn quan tâm, động viên, chia sẻ những khó khăn tạo động lực rất lớn để nghiên cứu sinh hoàn thành công trình này.

Xin chân thành cảm ơn!

MỤC LỤC

MỤC LỤC.....	i
DANH MỤC CÁC TỪ VIẾT TẮT	v
DANH MỤC HÌNH VẼ.....	viii
DANH MỤC BẢNG	x
DANH MỤC CÁC KÝ HIỆU TOÁN HỌC	xi
Chương 1. TỔNG QUAN VỀ BÀI TOÁN PHÂN LOẠI ĐIỀU CHẾ TỰ ĐỘNG ỨNG DỤNG MẠNG NƠ-RON HỌC SÂU.	10
1.1. Tổng quan về phân loại điều chế tự động.....	10
1.2. Cơ sở lý thuyết bài toán phân loại điều chế tự động.....	12
1.2.1. Mô hình toán học của hệ thống truyền dẫn.....	12
1.2.2. Các thách thức từ môi trường truyền dẫn vô tuyến.....	14
1.2.3. Định nghĩa bài toán AMC dưới góc độ học máy.....	16
1.3. Tổng quan về phân loại điều chế tự động cổ điển.....	17
1.3.1. Các phương pháp AMC cổ điển	17
1.3.2. Tổng quan về phân loại điều chế tự động dựa trên machine learning và deep learning	23
1.4. Cơ sở lý thuyết mạng Nơ-ron cho xử lý tín hiệu.....	29
1.4.1. Mạng Nơ-ron tích chập (CNN) trong miền thời gian.....	30
1.4.2. Mạng Nơ-ron hồi quy (RNN) và cơ chế BiLSTM	32

1.4.3. Tầng phân loại (Classification layer)	34
1.5. Tập dữ liệu mô phỏng	35
1.5.1. Tập dữ liệu HisarMod2019.1	35
1.5.2. Tập dữ liệu RadioML 2018.01A	37
1.5.3. Các tiêu chí đánh giá hiệu năng mô hình	38
1.6. Các công trình nghiên cứu có liên quan	41
1.6.1. Các công trình quốc tế nổi bật	41
1.6.2. Các nghiên cứu trong nước	44
1.6.3. Đánh giá chung	46
1.6.4. Hướng nghiên cứu chính của luận án	46
1.7. Kết luận chương 1	48
Chương 2. NÂNG CAO HIỆU NĂNG PHÂN LOẠI ĐIỀU CHẾ	
TỰ ĐỘNG SỬ DỤNG MẠNG NƠ-RON TÍCH CHẬP	50
2.1. Phân tích ảnh hưởng của cấu trúc thời gian, đặc trưng phổ và nhiễu cộng	51
2.1.1. Cấu trúc thời gian của tín hiệu trong kênh đa đường	52
2.1.2. Đặc trưng phổ và ảnh hưởng của sai lệch tần số	53
2.1.3. Ảnh hưởng của nhiễu cộng	54
2.2. Biểu diễn tín hiệu trong miền tần số và thời gian–tần số	57
2.2.1. Biểu diễn trong miền tần số bằng FFT	57
2.2.2. Biểu diễn thời gian–tần số bằng STFT	58
2.2.3. Kết quả thực nghiệm	59
2.2.4. Kết hợp STFT và Radon Transform để tăng cường đặc trưng	62

2.2.5. So sánh tổng thể với tín hiệu IQ gốc	65
2.3. Thiết kế mô hình DFENet	68
2.3.1. Đề xuất mô hình CNN	68
2.3.2. Cấu trúc tổng quan và luồng dữ liệu	70
2.3.3. Thiết kế khối DFE (Diverse Feature Extraction Block)	72
2.3.4. Thiết kế khối FEN (Feature Extraction and Normalization Block) 75	
2.3.5. So sánh với các mô hình hiện có	77
2.3.6. Mô tả tập dữ liệu và thiết lập thử nghiệm	79
2.3.7. Kết quả thực nghiệm và thảo luận	81
2.4. Kết luận chương 2	102
Chương 3. ĐỀ XUẤT GIẢI PHÁP KẾT HỢP KIẾN TRÚC CÁC MẠNG NƠ-RON CHO NÂNG CAO HIỆU NĂNG PHÂN LOẠI ĐIỀU CHẾ TỰ ĐỘNG	104
3.1. Thiết kế mô hình lai ghép AMC-CNN-BiLSTM	106
3.1.1. Xây dựng mô hình lai ghép	106
3.1.2. Kết quả thực nghiệm và thảo luận	113
3.2. Phương pháp học kết hợp để cải thiện hiệu năng AMC	120
3.2.1. Mô hình học kết hợp CNN	123
3.2.2. Mô hình học kết hợp và phương pháp kết hợp quyết định ..	124
3.2.3. Phương pháp học kết hợp CNN cho phân loại điều chế	125
3.2.4. Kết quả thực nghiệm và phân tích hiệu năng học kết hợp ..	127
3.2.5. Nhận xét, hạn chế và liên hệ	132

3.3. Kết luận chương 3.....	133
KẾT LUẬN VÀ HƯỚNG NGHIÊN CỨU.....	135
DANH MỤC CÁC CÔNG TRÌNH ĐÃ CÔNG BỐ.....	141
TÀI LIỆU THAM KHẢO.....	143

DANH MỤC CÁC TỪ VIẾT TẮT

Từ viết tắt	Nghĩa Tiếng Anh	Nghĩa Tiếng Việt
AMC	Automatic Modulation Classification	Phân loại điều chế tự động
ANN	Artificial Neural Networks	Mạng nơ-ron nhân tạo
AI	Artificial Intelligence	Trí tuệ nhân tạo
AWGN	Additive White Gaussian Noise	Tạp âm Gauss trắng cộng
ALRT	Average Likelihood Ratio Test	Kiểm định tỷ số hợp lý trung bình
BiLSTM	Bidirectional Long Short-Term Memory	LSTM hai chiều
BN	Batch Normalization	Chuẩn hóa theo lô
CNN	Convolutional Neural Network	Mạng nơ-ron tích chập
CR	Cognitive Radio	Vô tuyến nhận thức
CWT	Continuous Wavelet Transform	Biến đổi Wavelet liên tục
CFO	Carrier Frequency Offset	Lệch tần số sóng mang
CW	Continuous Wave	Sóng liên tục
DSA	Dynamic Spectrum Access	Truy nhập phổ động
DL	Deep Learning	Học sâu
DFENet	Diverse Feature Extraction Network	Mạng trích xuất đặc trưng đa dạng

DFE Block	Diverse Feature Extraction Block	Khối trích xuất đặc trưng đa dạng
DNN	Deep Neural Network	Mạng nơ-ron sâu
DOA	Direction of Arrival	Hướng góc đến
FFT	Fast Fourier Transform	Biến đổi Fourier nhanh
FEN Block	Feature Extraction and Normalization Block	Khối trích xuất và chuẩn hóa đặc trưng
FC	Fully Connected	Kết nối đầy đủ
FNN	Feedforward Neural Network	Mạng nơ-ron nhân tạo truyền thẳng
GLRT	Generalized Likelihood Ratio Test	Kiểm định tỷ số hợp lý tổng quát
GPU	Graphical Processing Unit	Bộ xử lý đồ họa
HLRT	Hybrid Likelihood Ratio Test	Kiểm định tỷ số hợp lý lai
HOS	Higher-Order Statistics	Đặc trưng thống kê bậc cao
IoT	Internet of Things	Internet vạn vật
I/Q	In-phase / Quadrature	Thành phần đồng pha / vuông pha
KNN	K-Nearest Neighbors	Thuật toán K láng giềng gần nhất
LSTM	Long Short-Term Memory	Mạng nơ-ron hồi quy bộ nhớ dài-ngắn hạn
ML	Machine Learning	Học máy
M	Million	Triệu
m	minute	Phút
ms	milliseconds	Mili giây
NBC	Naive Bayes Classifier	Bộ phân loại Naive Bayes
NLOS	Non-Line-of-Sight	Truyền dẫn không tầm nhìn thẳng

PCA	Principal Component Analysis	Phân tích thành phần chính
PDF	Probability Density Function	Hàm mật độ xác suất
ReLU	Rectified Linear Units	Hàm kích hoạt phi tuyến
RNN	Recurrent Neural Network	Mạng nơ-ron hồi quy
SGD	Stochastic Gradient Descent	Thuật toán tối ưu với độ dốc giảm dần ngẫu nhiên
SNR	Signal-to-Noise Ratio	Tỷ lệ tín trên tạp
STFT	Short-Time Fourier Transform	Biến đổi Fourier thời gian ngắn
SVD	Singular Value Decomposition	Phương pháp phân tích suy biến
SPWVD	Smoothed Pseudo Wigner-Ville Distribution	Phân bố Wigner–Ville giả làm trơn
SVM	Support Vector Machine	Máy véc-tơ hỗ trợ
t-SNE	t-distributed Stochastic Neighbor Embedding	Trực quan hóa không gian đặc trưng bằng t-SNE
WVD	Wigner-Ville Distribution	Phân bố Wigner–Ville

DANH MỤC HÌNH VẼ

1.1	AMC trong hệ thống điều chế thích nghi.	12
1.2	Hệ thống phân loại điều chế tự động (AMC) cổ điển dựa trên trích chọn đặc trưng thủ công.	17
1.3	Phân loại các phương pháp phân loại điều chế tự động.	18
1.4	Hệ thống AMC dựa trên Deep Learning với khả năng học đặc trưng tự động.	23
2.1	So sánh độ chính xác phân loại các CNN khác nhau sử dụng STFT và FFT.	60
2.2	So sánh độ chính xác phân loại của các CNN khác nhau sử dụng STFT và STFT-Radon.	64
2.3	So sánh độ chính xác phân loại của các CNN khác nhau sử dụng STFT-Radon và IQ.	67
2.4	Cấu trúc CNN cho AMC: (a) mô hình tổng thể của DFENet, (b) cấu trúc khối DFE, (c) cấu trúc khối FEN.	71
2.5	Độ chính xác của DFENet với số lượng bộ lọc khác nhau.	81
2.6	Ảnh hưởng của số lượng bộ lọc đến độ chính xác và độ phức tạp mô hình.	82
2.7	Độ chính xác phân loại điều chế với chiều dài tín hiệu khác nhau.	83
2.8	Độ chính xác của DFENet trên 26 loại điều chế khác nhau trong các điều kiện SNR khác nhau.	85
2.9	Ma trận nhầm lẫn cho phân loại 26 dạng điều chế tại mức SNR = +0dB.	87
2.10	Ma trận nhầm lẫn cho phân loại 26 dạng điều chế tại mức SNR = +10dB.	87

2.11	Đặc trưng tại đầu ra của lớp tích chập đầu tiên ở mức SNR = +0 dB.	89
2.12	Đặc trưng tại đầu ra của lớp FC ở mức SNR = +0 dB.	90
2.13	Đặc trưng tại đầu ra của lớp FC ở mức SNR = +10 dB.	91
2.14	Độ chính xác phân loại với số lượng khối DFE khác nhau.	94
2.15	Precision, Recall và F1-score của mô hình DFE.	95
2.16	So sánh độ chính xác phân loại của các mô hình CNN đơn lẻ khác nhau đối với tập dữ liệu HisarMod2019.	97
2.17	So sánh độ chính xác phân loại của các mô hình CNN đơn lẻ khác nhau đối với tập dữ liệu RadioMl2018.	102
3.1	Mô hình đề xuất của mô hình AMC-CNN-BiLSTM.	108
3.2	Độ chính xác phân loại với số lượng bộ lọc khác nhau.	114
3.3	Độ chính xác phân loại với các độ dài tín hiệu khác nhau.	115
3.4	Độ chính xác phân loại cho 26 loại điều chế khác nhau.	116
3.5	Ma trận nhầm lẫn của 26 phân loại điều chế ở mức SNR +10 dB.	117
3.6	So sánh độ chính xác phân loại của các mô hình CNN khác nhau.	118
3.7	So sánh độ chính xác phân loại của các mô hình CNN đơn lẻ khác nhau đối với tập dữ liệu RadioMl2018.	120
3.8	So sánh độ chính xác phân loại của các CNN đơn.	128
3.9	So sánh độ chính xác giữa các mô hình CNN đơn và các mô hình học kết hợp.	129

DANH MỤC BẢNG

1.1	So sánh các nhóm phương pháp học sâu trong AMC	26
1.2	Tập dữ liệu HisarMod2019.1	36
1.3	Các kiểu điều chế trong tập dữ liệu RadioML 2018.01A	37
1.4	Ma trận nhầm lẫn với bài toán phân loại 3 lớp A, B và C.	39
2.1	Độ chính xác của mô hình ResNet18 với các hàm cửa sổ khác nhau.	61
2.2	So sánh mô hình mạng DFE với các mô hình đa nhánh và đa quy mô	77
2.3	Cấu trúc mô hình 4 khối DFENet.	79
2.4	Nghiên cứu cắt bỏ trên bộ dữ liệu HisarMod2019	91
2.5	Độ phức tạp, thời gian tính toán, FLOPs của mô hình.	99
3.1	So sánh hiệu năng và độ phức tạp giữa các mô hình CNN đơn và mô hình ensemble	131

DANH MỤC CÁC KÝ HIỆU TOÁN HỌC

Ký hiệu	Ý nghĩa
$*$	Phép tích chập
$\odot$	Nhân từng phần tử
$\oplus$	Cộng từng phần tử
$h(t)$	Đáp ứng xung của kênh truyền.
$n(t)$	Tạp âm cộng.
$A(t), \phi(t)$	Biên độ tức thời và pha tức thời của tín hiệu.
$\mathbf{R}$	Vector tín hiệu thu chứa nhiễu và fading.
$\hat{y}$	Loại điều chế được dự đoán.
$r(t), r[n]$	Tín hiệu thu
X	Tín hiệu sạch (input signal)
$\tilde{X} = X + N$	Tín hiệu quan sát có nhiễu
$N, n(t), n[n]$	Nhiều cộng
$h(t)$	Đáp ứng xung của kênh truyền
$\mathbf{R}$	Vector tín hiệu thu
$A(t), \phi(t)$	Biên độ và pha tức thời
$n[n] \sim \mathcal{CN}(0, \sigma^2)$	Nhiều Gaussian phức
σ^2	Phương sai nhiễu
$\mathbb{E}[\cdot]$	Toán tử kỳ vọng
SNR	Tỷ số tín hiệu trên nhiễu
$p(x)$	Hàm mật độ xác suất
$H(X)$	Entropy của biến ngẫu nhiên X
$H(Y X)$	Entropy có điều kiện
$I(X; Y)$	Thông tin tương hỗ

$H(X + N)$	Entropy của tín hiệu có nhiễu
$I(\tilde{X}; Y)$	Thông tin tương hỗ khi có nhiễu
$Z = f(\tilde{X})$	Đặc trưng học được từ mạng
$f(\cdot)$	Hàm ánh xạ (mô hình học sâu)
θ	Tập tham số mô hình
$\hat{y}$	Nhãn dự đoán
y	Nhãn thực
M	Số lượng lớp điều chế
$p(r M_i)$	Xác suất có điều kiện của tín hiệu thu
<i>Precision</i>	Độ chính xác dương
<i>Recall</i>	Độ thu hồi
<i>F1</i>	F1-score

MỞ ĐẦU

Trong bối cảnh chiến tranh hiện đại chuyển mạnh sang không gian điện từ, *trình sát điện tử* đóng vai trò then chốt trong việc thu thập, phân tích và khai thác thông tin từ các nguồn phát xạ vô tuyến. Khả năng phát hiện, nhận dạng và phân loại tín hiệu một cách nhanh chóng, chính xác là nền tảng để xây dựng bức tranh tác chiến điện tử, hỗ trợ ra quyết định chiến thuật và đảm bảo ưu thế trên phổ tần. Đặc biệt, trong các hệ thống *tác chiến điện tử*, việc nhận dạng dạng điều chế của tín hiệu thu không chỉ phục vụ giám sát phổ mà còn trực tiếp liên quan đến phát hiện mục tiêu, phân tích ý đồ đối phương và triển khai các biện pháp đối kháng hiệu quả.

Song song với đó, phổ tần vô tuyến ngày càng trở nên phức tạp do sự gia tăng nhanh chóng của mật độ thiết bị phát xạ, sự đa dạng của các chuẩn truyền thông và tính động cao của môi trường truyền dẫn. Trong các kịch bản quân sự, tín hiệu thường được thiết kế theo hướng *khó bị phát hiện*, sử dụng các kỹ thuật điều chế thích nghi, nhảy tần hoặc trải phổ, đồng thời chịu tác động của gây nhiễu chủ động và các dạng can nhiễu Gauss. Những đặc điểm này làm gia tăng đáng kể độ khó của bài toán nhận dạng tín hiệu, đặc biệt trong điều kiện tỷ số tín hiệu trên nhiễu thấp.

Trong bối cảnh đó, bài toán *phân loại điều chế tự động* (Automatic Modulation Classification – AMC) không chỉ là một vấn đề kỹ thuật thuần túy mà còn là thành phần cốt lõi trong các hệ thống trình sát điện tử, giám sát phổ tần và thông tin vô tuyến thông minh. Tuy nhiên, trong thực tế, tín hiệu thu luôn chịu ảnh hưởng đồng thời của nhiễu Gauss trắng, fading đa đường, sai

lệch tần số, dịch Doppler và các yếu tố phi tuyến của kênh truyền. Những tác động này làm biến dạng mạnh đặc trưng biên độ và pha của tín hiệu, khiến việc nhận dạng dựa trên đặc trưng trở nên khó khăn, đặc biệt trong các kịch bản tác chiến có cường độ nhiễu cao.

Các phương pháp AMC truyền thống, dựa trên lý thuyết quyết định hoặc trích xuất đặc trưng thủ công, thường yêu cầu các giả thiết chặt chẽ về mô hình kênh và phụ thuộc lớn vào kinh nghiệm chuyên gia. Khi môi trường thay đổi hoặc xuất hiện các dạng tín hiệu mới, các phương pháp này bộc lộ hạn chế về khả năng thích nghi và suy giảm hiệu năng đáng kể. Điều này đặt ra yêu cầu cấp thiết đối với các phương pháp có khả năng tự động học đặc trưng, tổng quát hóa tốt và hoạt động ổn định trong môi trường vô tuyến phức tạp, đặc biệt là trong các hệ thống quân sự phi hợp tác.

Sự phát triển mạnh mẽ của trí tuệ nhân tạo và học sâu trong những năm gần đây đã mở ra một hướng tiếp cận đầy triển vọng cho bài toán AMC. Các mạng nơ-ron sâu có khả năng học trực tiếp từ dữ liệu tín hiệu thô, tự động trích xuất đặc trưng và mô hình hóa các quan hệ phi tuyến phức tạp mà không cần xây dựng mô hình toán học tường minh cho từng điều kiện kênh truyền. Đây là ưu thế đặc biệt quan trọng trong các hệ thống vô tuyến quân sự, nơi môi trường hoạt động có tính biến động cao và khó mô hình hóa chính xác trước.

Xuất phát từ những yêu cầu thực tiễn nêu trên, với mục tiêu nghiên cứu chuyên sâu trong lĩnh vực xử lý tín hiệu và học máy nhằm đề xuất các giải pháp nâng cao hiệu năng nhận dạng tín hiệu vô tuyến trong các điều kiện bất lợi, nghiên cứu sinh đã lựa chọn đề tài **“Nghiên cứu nâng cao hiệu năng phân loại điều chế tự động ứng dụng mạng nơ-ron”** cho luận

án tiến sĩ. Đề tài không chỉ có ý nghĩa khoa học trong việc phát triển các mô hình học sâu tiên tiến cho bài toán AMC, mà còn mang giá trị thực tiễn cao trong các ứng dụng trinh sát điện tử, tác chiến điện tử và đảm bảo an ninh – quốc phòng, góp phần nâng cao hiệu quả giám sát và khai thác phổ tần vô tuyến trong các kịch bản hiện đại.

1. Mục tiêu nghiên cứu của luận án

Mục tiêu tổng quát của luận án là nghiên cứu, đề xuất và phát triển các phương pháp học sâu nhằm nâng cao hiệu năng nhận dạng tín hiệu vô tuyến, đặc biệt là bài toán phân loại điều chế tự động và bài toán kết hợp ước lượng tham số tín hiệu trong môi trường kênh truyền phức tạp, đáp ứng yêu cầu của các hệ thống vô tuyến thông minh hiện đại.

Trên cơ sở mục tiêu tổng quát đó, luận án hướng tới các mục tiêu sau:

- Nghiên cứu cơ sở lý thuyết và tổng quan các phương pháp phân loại điều chế tự động truyền thống và hiện đại, bao gồm các phương pháp dựa trên trích chọn đặc trưng thủ công và các phương pháp dựa trên học máy, học sâu, từ đó chỉ ra các hạn chế còn tồn tại trong môi trường nhiễu và fading đa đường.
- Đề xuất các mô hình mạng nơ-ron sâu mới hoặc cải tiến, có khả năng tự động học đặc trưng hiệu quả từ tín hiệu vô tuyến đầu vào (tín hiệu IQ hoặc các biểu diễn miền thời–tần số), nhằm nâng cao độ chính xác phân loại điều chế trong điều kiện SNR thấp và kênh truyền phức tạp.
- Nghiên cứu và đánh giá hiệu quả của các kỹ thuật tiền xử lý và biểu diễn tín hiệu (như STFT, phổ thời–tần số, đặc trưng học được), cũng như ảnh hưởng của các tham số xử lý tín hiệu đến hiệu năng của các mô

hình học sâu trong bài toán AMC.

- Đề xuất và phân tích phương pháp học kết hợp (Ensemble Learning) giữa nhiều mô hình CNN khác nhau nhằm khai thác tính bổ sung về đặc trưng và khả năng khái quát của từng mô hình, từ đó cải thiện độ chính xác và độ ổn định của hệ thống phân loại điều chế.
- Đánh giá toàn diện hiệu năng của các mô hình đề xuất thông qua mô phỏng trên các bộ dữ liệu chuẩn, xét theo nhiều tiêu chí như độ chính xác phân loại, độ ổn định theo SNR, độ phức tạp tính toán và khả năng triển khai thực tế.

Thông qua việc đạt được các mục tiêu trên, luận án hướng tới việc đóng góp các giải pháp có giá trị khoa học và thực tiễn cho lĩnh vực nhận dạng tín hiệu vô tuyến, đặc biệt trong bối cảnh các hệ thống thông tin vô tuyến thế hệ mới và các ứng dụng an ninh – quốc phòng.

2. Đối tượng nghiên cứu

Đối tượng nghiên cứu của luận án là các phương pháp và mô hình xử lý tín hiệu vô tuyến dựa trên mạng nơ-ron nhân tạo, đặc biệt là các mô hình học sâu (Deep Learning), được ứng dụng cho bài toán nhận dạng tín hiệu vô tuyến trong môi trường kênh truyền phức tạp.

Cụ thể, luận án tập trung nghiên cứu các đối tượng sau:

- Tín hiệu vô tuyến số thu được tại máy thu, bao gồm tín hiệu IQ và các dạng biểu diễn dẫn xuất trong miền thời gian, miền tần số và phổ tần.
- Bài toán phân loại điều chế tự động trong điều kiện nhiễu, suy hao và fading đa đường, với nhiều mức SNR khác nhau.

- Các mô hình mạng nơ-ron học sâu như mạng nơ-ron tích chập, các mô hình CNN cải tiến và các phương pháp học kết hợp dùng cho nhận dạng điều chế.

Các đối tượng nghiên cứu trên được xem xét trong bối cảnh các hệ thống vô tuyến thông minh, vô tuyến nhận thức và các ứng dụng trinh sát, giám sát phổ vô tuyến.

3. Phạm vi nghiên cứu

Trong khuôn khổ của luận án, phạm vi nghiên cứu được giới hạn như sau:

- Luận án tập trung vào bài toán nhận dạng điều chế tín hiệu vô tuyến số ở lớp vật lý, không đi sâu vào các lớp cao hơn của hệ thống thông tin như lớp liên kết dữ liệu hay lớp mạng.
- Các dạng điều chế được xem xét chủ yếu là các điều chế số phổ biến trong thông tin vô tuyến hiện đại, bao gồm các điều chế số pha (PSK), điều chế biên độ vuông góc (QAM) và một số dạng điều chế liên quan, phù hợp với các bộ dữ liệu và kịch bản mô phỏng được sử dụng.
- Môi trường kênh truyền được mô hình hóa thông qua các kịch bản nhiễu và fading điển hình (AWGN, fading đa đường), nhằm đánh giá tính bền vững của các mô hình đề xuất trong điều kiện truyền dẫn thực tế.
- Các phương pháp học sâu được nghiên cứu và đánh giá chủ yếu thông qua mô phỏng bằng phần mềm Matlab và các bộ dữ liệu chuẩn; luận án không đi sâu vào triển khai phần cứng hoặc thử nghiệm thực địa.

Với phạm vi nghiên cứu như trên, luận án hướng tới việc làm rõ hiệu quả và tiềm năng ứng dụng của các mô hình học sâu trong bài toán nhận dạng

tín hiệu vô tuyến, đồng thời đảm bảo tính khả thi và tính khoa học của các kết quả đạt được.

4. Phương pháp nghiên cứu

- **Phương pháp nghiên cứu lý thuyết:** Luận án tiến hành phân tích và tổng hợp cơ sở lý thuyết về tín hiệu vô tuyến số, các mô hình kênh truyền, bài toán nhận dạng điều chế tự động và các phương pháp học máy, học sâu liên quan. Trên cơ sở đó, các đặc trưng tín hiệu trong miền thời gian, miền tần số và miền thời gian – tần số được khảo sát nhằm làm rõ khả năng biểu diễn và phân tách thông tin điều chế trong điều kiện nhiễu và fading.
- **Phương pháp mô hình hóa và thiết kế thuật toán:** Dựa trên cơ sở lý thuyết, luận án đề xuất và xây dựng các mô hình mạng nơ-ron học sâu cho bài toán nhận dạng tín hiệu vô tuyến. Các mô hình mạng được thiết kế theo hướng tự động trích xuất đặc trưng từ dữ liệu thô, bao gồm các mô hình CNN cải tiến, mô hình kết hợp nhiều nhánh đặc trưng, mô hình ensemble và mô hình mạng nơ-ron đa đầu vào – đa đầu ra cho bài toán kết hợp phân loại điều chế.
- **Phương pháp mô phỏng và thực nghiệm số:** Các mô hình đề xuất được đánh giá thông qua mô phỏng bằng phần mềm Matlab. Tín hiệu vô tuyến được sinh tổng hợp hoặc sử dụng từ các bộ dữ liệu chuẩn, với nhiều mức SNR và các điều kiện kênh truyền khác nhau. Các chỉ tiêu đánh giá bao gồm độ chính xác phân loại, độ ổn định theo SNR, độ phức tạp tính toán và thời gian suy luận của mô hình.
- **Phương pháp so sánh và đối chứng:** Hiệu năng của các mô hình đề

xuất được so sánh với các phương pháp AMC cổ điển và các mô hình học sâu phổ biến trong tài liệu nghiên cứu. Thông qua các thí nghiệm đối chứng, luận án phân tích ưu điểm và hạn chế của từng phương pháp, từ đó làm rõ hiệu quả của các giải pháp đề xuất.

- **Phương pháp phân tích và thảo luận kết quả:** Các kết quả mô phỏng được tổng hợp, phân tích và thảo luận một cách có hệ thống nhằm đánh giá tính đúng đắn, tính bền vững và khả năng ứng dụng thực tiễn của các mô hình. Từ đó, luận án rút ra các kết luận khoa học và đề xuất các hướng nghiên cứu tiếp theo.

5. Các kết quả đạt được

Một số đề xuất đóng góp chính của luận án được tóm tắt như sau:

- Xây dựng và đánh giá có hệ thống các phương pháp tiền xử lý và biểu diễn tín hiệu trong các miền khác nhau, bao gồm tín hiệu IQ miền thời gian, phổ tần số (FFT), biểu diễn thời-tần (STFT) và STFT kết hợp biến đổi Radon. Thông qua phân tích lý thuyết và thực nghiệm định lượng, luận án xác định dạng biểu diễn tối ưu làm cơ sở cho thiết kế kiến trúc học sâu chuyên biệt, qua đó làm rõ vai trò quyết định của không gian biểu diễn đối với hiệu năng AMC. Kết quả nghiên cứu được công bố trong công trình [C2, C3] trong danh mục các công trình công bố.
- Đề xuất mô hình **DFENet (Diverse Feature Extraction Network)** cho bài toán phân loại điều chế tự động. Mô hình được thiết kế theo hướng trích xuất đặc trưng đa nhánh trực tiếp từ tín hiệu IQ, khai thác hiệu quả cấu trúc không gian của tín hiệu vô tuyến mà không phụ thuộc vào đặc trưng thủ công. DFENet cho phép nâng cao độ chính xác phân

loại trong dải SNR rộng và dưới các điều kiện kênh truyền fading phức tạp. Kết quả nghiên cứu được công bố trong công trình [J1] trong danh mục các công trình công bố.

- Đề xuất kiến trúc lai **AMC-CNN-BiLSTM** nhằm khai thác đồng thời đặc trưng không gian và phụ thuộc thời gian của tín hiệu điều chế, khắc phục hạn chế của mạng CNN thuần túy trong việc mô hình hóa chuỗi dài hạn. Trên cơ sở đó, luận án phát triển giải pháp học kết hợp các mô hình khác biệt tại tầng đầu ra, tận dụng tính bổ sung giữa các kiến trúc để nâng cao độ chính xác tổng thể và tăng tính bền vững của hệ thống trong điều kiện SNR thấp. Kết quả nghiên cứu được công bố trong công trình [C4,J2] trong danh mục các công trình công bố.

6. Bố cục của luận án Các kết quả nghiên cứu và những đóng góp mới của luận án được trình bày chi tiết trong các chương sau:

Chương 1: Tổng quan về bài toán phân loại điều chế tự động ứng dụng mạng nơ-ron học sâu

Chương này trình bày tổng quan cơ sở lý thuyết của bài toán phân loại điều chế tự động trong hệ thống thông tin vô tuyến, bao gồm các khái niệm cơ bản, mô hình tín hiệu và các thách thức trong môi trường nhiễu và fading đa đường. Bên cạnh đó, chương cũng phân tích các phương pháp AMC cổ điển dựa trên trích chọn đặc trưng và lý thuyết quyết định, cũng như các phương pháp AMC hiện đại dựa trên Machine Learning và Deep Learning. Trên cơ sở khảo sát các công trình nghiên cứu trong và ngoài nước, chương này chỉ ra những ưu điểm, hạn chế và các khoảng trống nghiên cứu, làm cơ sở khoa học cho việc đề xuất các mô hình và phương pháp mới trong các

chương tiếp theo của luận án.

Chương 2: Nâng cao hiệu năng phân loại điều chế tự động sử dụng mạng nơ-ron tích chập.

Chương này tập trung khảo sát có hệ thống các phương pháp biểu diễn tín hiệu vô tuyến, bao gồm miền thời gian (IQ), miền tần số (FFT) và miền thời gian–tần số (STFT và STFT kết hợp Radon), nhằm đánh giá ảnh hưởng của không gian biểu diễn đến hiệu năng phân loại. Trên cơ sở phân tích lý thuyết và thực nghiệm, luận án xác định tín hiệu IQ là biểu diễn phù hợp và đề xuất kiến trúc DFENet (Diverse Feature Extraction Network) với cơ chế trích xuất đặc trưng đa nhánh nhằm khai thác hiệu quả cấu trúc của tín hiệu IQ. Các kết quả đánh giá định lượng chứng minh ưu thế của mô hình đề xuất trong môi trường nhiễu và fading phức tạp.

Chương 3: Đề xuất giải pháp kết hợp kiến trúc các mạng nơ-ron cho nâng cao hiệu năng phân loại điều chế tự động.

Chương này phát triển mô hình AMC-CNN-BiLSTM nhằm khai thác đồng thời đặc trưng không gian và phụ thuộc theo thời gian của tín hiệu điều chế. Bên cạnh đó, luận án đề xuất phương pháp học kết hợp (ensemble learning) tại tầng đầu ra giữa các mô hình khác biệt về cấu trúc, nhằm tận dụng tính bổ sung giữa các đặc trưng đã học. Các thí nghiệm so sánh toàn diện được thực hiện để đánh giá mức độ cải thiện về độ chính xác và tính bền vững của hệ thống trong điều kiện SNR thấp và môi trường fading đa dạng.

Trong phần kết luận, luận án tổng hợp các kết quả nghiên cứu đạt được, nhấn mạnh những đóng góp khoa học chính, đồng thời đề xuất các hướng nghiên cứu tiếp theo nhằm mở rộng và hoàn thiện các phương pháp đã được trình bày.

Chương 1

TỔNG QUAN VỀ BÀI TOÁN PHÂN LOẠI ĐIỀU CHẾ TỰ ĐỘNG ỨNG DỤNG MẠNG NƠ-RON HỌC SÂU

1.1. Tổng quan về phân loại điều chế tự động

Trong những năm gần đây, sự phát triển nhanh chóng của các hệ thống thông tin vô tuyến, đặc biệt với sự chuyển dịch sang 5G/6G và sự bùng nổ của Internet vạn vật (IoT), đã làm gia tăng đáng kể mật độ thiết bị và lưu lượng truyền dẫn. Điều này tạo ra áp lực lớn lên tài nguyên phổ tần vô tuyến, vốn hữu hạn và có giá trị cao, đòi hỏi các cơ chế khai thác linh hoạt và hiệu quả hơn [1], [2].

Các phương pháp phân bổ phổ tần số truyền thống dần bộc lộ hạn chế do hiện tượng sử dụng không đồng đều giữa các băng tần. Trong bối cảnh đó, các mô hình vô tuyến thông minh như vô tuyến nhận thức (Cognitive Radio – CR) và truy nhập phổ động (Dynamic Spectrum Access – DSA) được đề xuất nhằm tối ưu hóa việc sử dụng phổ thông qua khả năng cảm nhận và thích nghi với môi trường vô tuyến [3], [4].

Trong các hệ thống này, bài toán *phân loại điều chế tự động* (Automatic Modulation Classification – AMC) đóng vai trò then chốt tại lớp vật lý. AMC là quá trình xác định dạng điều chế của tín hiệu thu từ các quan sát nhiễu mà không cần thông tin tiên nghiệm, làm cơ sở cho việc lựa chọn chiến lược giải điều chế phù hợp [5], [6]. Không chỉ giới hạn trong các hệ thống thông tin dân sự, AMC còn có ý nghĩa đặc biệt quan trọng trong các ứng dụng

quân sự và an ninh. Trong tác chiến điện tử và trinh sát vô tuyến, AMC cho phép nhận dạng các nguồn phát không hợp tác, hỗ trợ phân tích tín hiệu, gây nhiễu có chọn lọc và thu thập thông tin tình báo. Đồng thời, trong quản lý phổ tần dân sự, AMC là công cụ hiệu quả để giám sát việc sử dụng phổ, phát hiện các nguồn phát trái phép và đảm bảo tuân thủ các quy định về tần số [7], [8].

Tuy nhiên, hiệu năng của AMC trong thực tế bị suy giảm đáng kể do ảnh hưởng của nhiễu, fading đa đường, sai lệch tần số và pha. Các phương pháp truyền thống dựa trên đặc trưng thủ công thường kém ổn định và khó thích nghi với môi trường kênh phức tạp.

Gần đây, học sâu đã trở thành một hướng tiếp cận hiệu quả cho bài toán AMC nhờ khả năng học đặc trưng tự động từ dữ liệu tín hiệu. Các mô hình như CNN và RNN/LSTM giúp cải thiện đáng kể độ chính xác phân loại, tuy nhiên vẫn đặt ra thách thức trong việc thiết kế mô hình phù hợp với bản chất chuỗi thời gian và dữ liệu phức của tín hiệu vô tuyến, cũng như cân bằng giữa hiệu năng và chi phí tính toán.

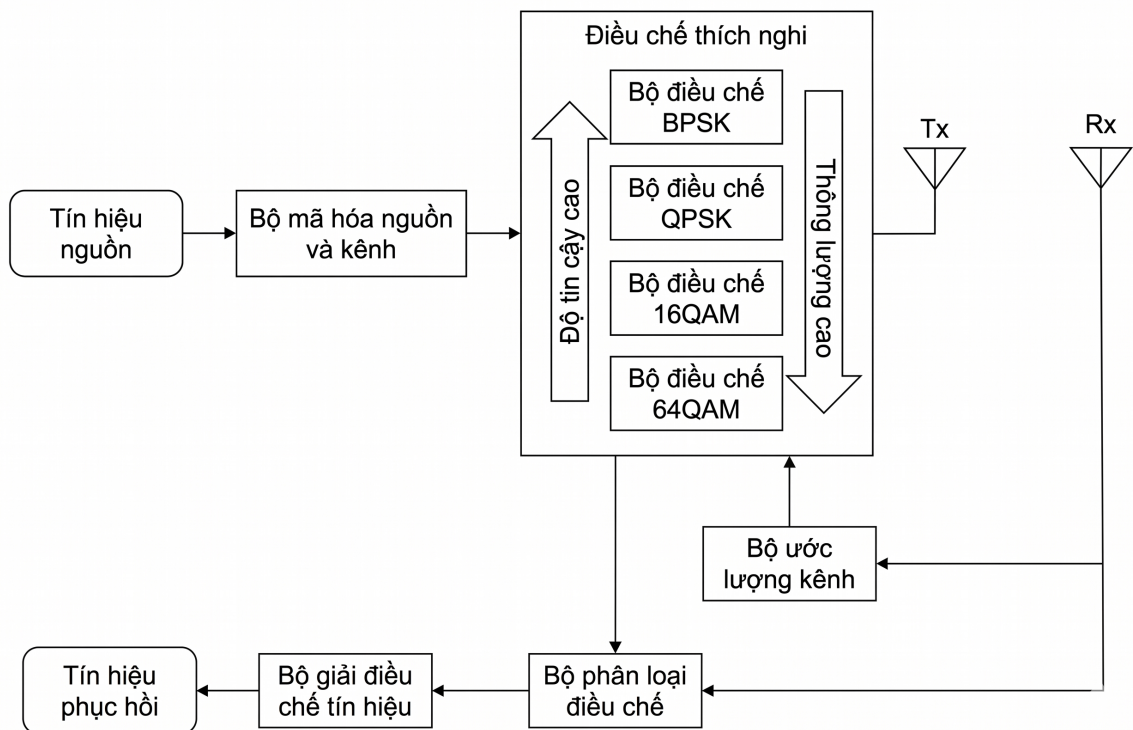
Xuất phát từ đó, luận án này tập trung nghiên cứu các phương pháp nâng cao hiệu năng AMC thông qua việc kết hợp học sâu với các kỹ thuật xử lý tín hiệu tiên tiến. Các mô hình đề xuất như DFENet và AMC-CNN-BiLSTM, cùng với các phương pháp như tiền xử lý miền thời gian–tần số, học kết hợp và mô hình đa đầu vào, được phát triển nhằm nâng cao độ chính xác và tính bền vững của hệ thống trong môi trường vô tuyến phức tạp.

1.2. Cơ sở lý thuyết bài toán phân loại điều chế tự động

Để xây dựng các giải pháp nhận dạng tín hiệu hiệu quả, trước hết cần phải mô hình hóa chính xác quá trình sinh, truyền dẫn và thu tín hiệu trong một hệ thống thông tin vô tuyến tổng quát. Bản chất của bài toán AMC là một bài toán ước lượng và phân loại thống kê, trong đó máy thu phải đưa ra quyết định về định dạng điều chế của tín hiệu dựa trên các mẫu quan sát bị nhiễu và méo dạng [9].

1.2.1. Mô hình toán học của hệ thống truyền dẫn

Xét một hệ thống thông tin vô tuyến đơn giản bao gồm ba thành phần chính: máy phát (transmitter), kênh truyền (channel) và máy thu (receiver) được minh họa ở Hình 1.1 [10].



Hình 1.1: AMC trong hệ thống điều chế thích nghi.

Tại phía phát, chuỗi tin tức nhị phân được ánh xạ vào các ký hiệu (symbols) và sau đó được điều chế lên sóng mang cao tần để tạo thành tín hiệu phát $s(t)$. Tín hiệu này là một hàm phức của thời gian, mang thông tin về biên độ, pha hoặc tần số tùy thuộc vào phương thức điều chế được sử dụng (ví dụ: QAM, PSK, FSK).

Tín hiệu khi đi qua môi trường truyền dẫn sẽ tương tác với kênh truyền $h(t)$ và chịu tác động của nhiễu cộng $n(t)$. Tín hiệu bằng cơ sở nhận được tại máy thu $r(t)$ có thể được biểu diễn dưới dạng toán học như sau [4]:

$$r(t) = s(t) * h(t) + n(t) \quad (1.1)$$

Trong đó:

- $(*)$ biểu thị phép toán tích chập (convolution).
- $h(t)$ là đáp ứng xung của kênh truyền, đặc trưng cho các hiệu ứng đa đường, suy hao và trễ truyền dẫn.
- $n(t)$ là tạp âm cộng (additive noise), thường được mô hình hóa là tạp âm Gauss trắng cộng (AWGN) sinh ra do nhiệt độ của các linh kiện điện tử tại máy thu và bức xạ nền môi trường.

Trong các hệ thống thu phát số hiện đại, tín hiệu thường được xử lý dưới dạng số phức. Do đó, tín hiệu thu $r(t)$ được phân tách thành hai thành phần trực giao là đồng pha (In-phase - I) và vuông pha (Quadrature - Q):

$$r(t) = I(t) + jQ(t) \quad (1.2)$$

Mối quan hệ giữa các thành phần này với biên độ tức thời $A(t)$ và pha tức

thời $\phi(t)$ được xác định bởi hệ phương trình:

$$I(t) = A(t) \cos(\phi(t)) \quad (1.3)$$

$$Q(t) = A(t) \sin(\phi(t)) \quad (1.4)$$

Đối với các mô hình học sâu (như CNN hay LSTM) được đề xuất trong luận án này, dữ liệu đầu vào sẽ là chuỗi rời rạc của các cặp giá trị $(I[k], Q[k])$ được lấy mẫu tại các thời điểm $t = kT_s$. Việc sử dụng biểu diễn I/Q cho phép giữ nguyên vẹn thông tin gốc của tín hiệu mà không làm mất mát các đặc trưng về pha như khi chỉ sử dụng phổ biên độ.

1.2.2. Các thách thức từ môi trường truyền dẫn vô tuyến

Khác với môi trường truyền dẫn hữu tuyến ổn định, môi trường vô tuyến chứa đựng nhiều yếu tố ngẫu nhiên làm biến dạng tín hiệu, gây khó khăn lớn cho việc nhận dạng điều chế. Luận án tập trung giải quyết ba thách thức chính [11], [12]:

Fading đa đường (Multipath Fading)

Đây là hiện tượng gây suy giảm chất lượng tín hiệu nghiêm trọng nhất. Do sự hiện diện của các vật cản (tòa nhà, cây cối, địa hình), sóng vô tuyến từ máy phát đến máy thu không chỉ đi theo một đường thẳng mà bị phản xạ, tán xạ và nhiễu xạ thành nhiều bản sao. Các bản sao này đi đến máy thu với các độ trễ và pha khác nhau, khi chồng chập lại có thể triệt tiêu lẫn nhau (fading sâu) hoặc tăng cường lẫn nhau [13].

Luận án xem xét ba mô hình kênh fading quan trọng để đánh giá độ ổn định của thuật toán:

- **Kênh Rayleigh:** Mô phỏng môi trường đô thị dày đặc, không có đường

truyền tầm nhìn thẳng (Non-Line-of-Sight - NLOS). Biên độ tín hiệu tuân theo phân bố Rayleigh, biến động mạnh và ngẫu nhiên.

- **Kênh Rice:** Mô phỏng môi trường có một thành phần truyền thẳng mạnh (LOS) trội hơn so với các thành phần tán xạ. Biên độ tuân theo phân bố Rice.
- **Kênh Nakagami-m:** Một mô hình tổng quát hóa có thể mô tả linh hoạt mức độ nghiêm trọng của fading thông qua tham số hình dáng m .

Nhiều tạp âm (Additive noise)

Ngay cả trong điều kiện kênh truyền lý tưởng, tín hiệu luôn bị ảnh hưởng bởi nhiễu nhiệt (AWGN). Tỷ số tín hiệu trên nhiễu (SNR) là thước đo quan trọng quyết định độ khó của bài toán AMC. Tại mức SNR thấp (ví dụ -20 dB, như khảo sát trong luận án), năng lượng của tín hiệu nhỏ hơn nhiều so với nhiễu, làm cho các đặc trưng điều chế bị “chìm” hoàn toàn, khiến các phương pháp truyền thống trở nên vô hiệu [13].

Các khiếm khuyết phần cứng và đồng bộ

Trong thực tế, máy phát và máy thu không bao giờ đồng bộ hoàn hảo. Các sai lệch này bao gồm:

- **Lệch tần số sóng mang (Carrier frequency offset - CFO):** Do sai lệch giữa các bộ dao động nội, gây ra hiện tượng quay chòm sao tín hiệu theo thời gian.
- **Lệch pha (Phase offset):** Làm xoay toàn bộ chòm sao đi một góc cố định.

- **Trôi đồng hồ mẫu (Sampling clock drift):** Gây sai lệch thời điểm lấy mẫu, làm méo dạng xung.

1.2.3. Định nghĩa bài toán AMC dưới góc độ học máy

Dựa trên các mô hình trên, bài toán AMC có thể được phát biểu như một bài toán phân loại mẫu M -lớp. Giả sử tập hợp các phương thức điều chế có thể có là $\mathcal{S} = \{Mod_1, Mod_2, \dots, Mod_M\}$, trong đó M là số lượng loại điều chế (trong luận án này $M = 26$).

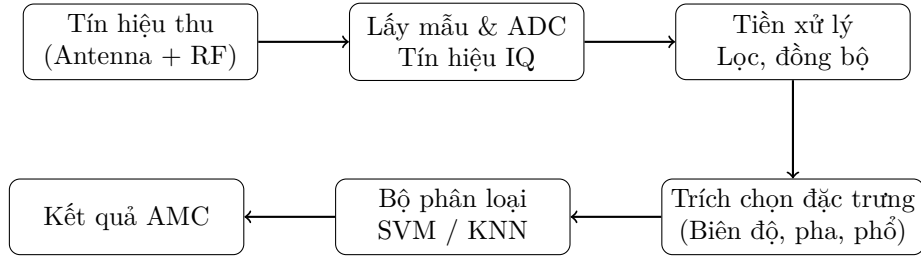
Nhiệm vụ của hệ thống AMC là tìm ra một hàm ánh xạ $f(\cdot)$ từ không gian tín hiệu quan sát $\mathbf{R}$ (chuỗi I/Q độ dài N) sang không gian nhãn lớp $\mathcal{S}$:

$$\hat{y} = f(\mathbf{R} \mid \theta) \quad (1.5)$$

Trong đó:

- $\mathbf{R}$ là vector tín hiệu thu chứa nhiễu và fading.
- θ là tập hợp các tham số của mô hình phân loại (đối với mạng nơ-ron, đây là các trọng số W và bias b).
- $\hat{y}$ là loại điều chế được dự đoán.

Mục tiêu của luận án là tối ưu hóa hàm $f(\cdot)$ thông qua việc thiết kế các kiến trúc mạng nơ-ron sâu để cực tiểu hóa xác suất sai số phân loại $P(\hat{y} \neq y)$, đặc biệt trong các điều kiện mà vector $\mathbf{R}$ bị biến dạng nặng nề bởi kênh truyền $h(t)$ và nhiễu $n(t)$ như đã mô tả ở trên.



Hình 1.2: Hệ thống phân loại điều chế tự động (AMC) cổ điển dựa trên trích chọn đặc trưng thủ công.

1.3. Tổng quan về phân loại điều chế tự động cổ điển

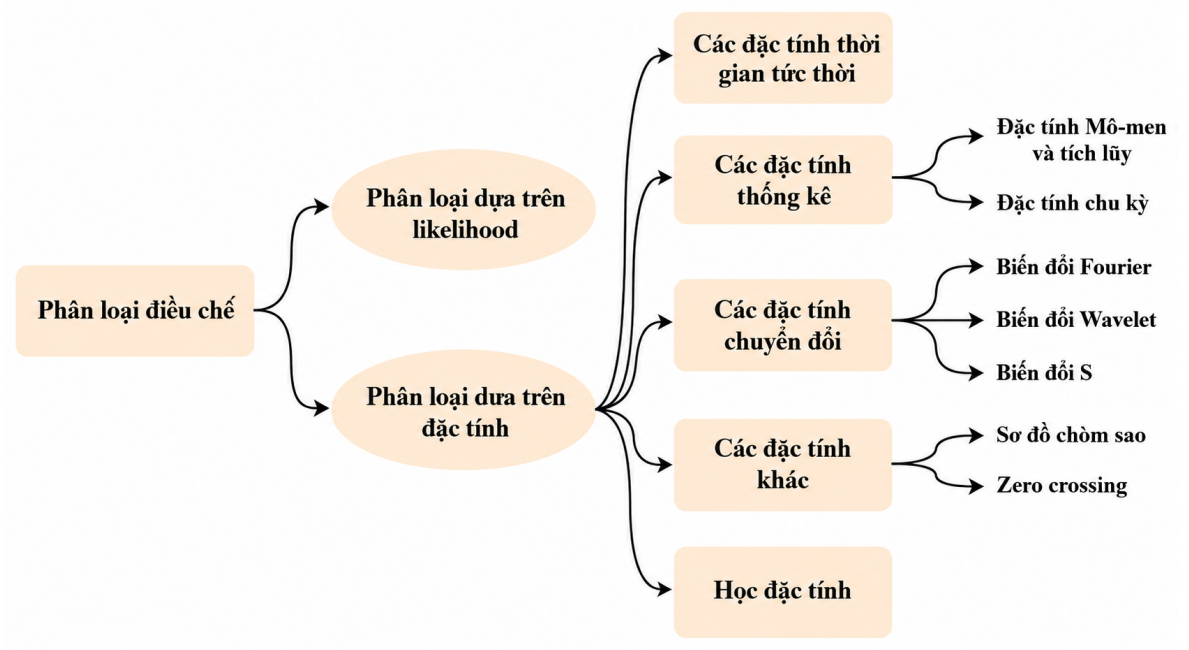
1.3.1. Các phương pháp AMC cổ điển

Trước khi các phương pháp học máy và học sâu được ứng dụng rộng rãi, bài toán phân loại điều chế tự động chủ yếu được giải quyết bằng các phương pháp cổ điển dựa trên lý thuyết thống kê và xử lý tín hiệu số được minh họa trong Hình 1.2. Các phương pháp này có thể được phân thành hai nhóm chính: (i) các phương pháp dựa trên hàm hợp lý (likelihood-based methods) và (ii) các phương pháp dựa trên trích chọn đặc trưng (feature-based methods) được minh họa trong Hình 1.3. Mỗi nhóm có những ưu điểm và hạn chế riêng, phản ánh sự đánh đổi giữa độ chính xác, độ phức tạp tính toán và khả năng thích nghi với môi trường kênh truyền [14], [15].

Phương pháp dựa trên hàm hợp lý (Likelihood-based AMC)

Các phương pháp dựa trên hàm hợp lý được xem là nhóm tiếp cận tối ưu theo nghĩa lý thuyết thống kê. Ý tưởng cốt lõi của nhóm này là xây dựng hàm xác suất có điều kiện của tín hiệu thu được ứng với từng giả thuyết điều chế, sau đó lựa chọn kiểu điều chế có giá trị hàm hợp lý lớn nhất [16].

Giả sử tập các giả thuyết điều chế là $\mathcal{M} = \{M_1, M_2, \dots, M_K\}$ và tín hiệu



Hình 1.3: Phân loại các phương pháp phân loại điều chế tự động.

thu được là $\mathbf{r}$. Bài toán AMC được phát biểu dưới dạng:

$$\hat{M} = \arg \max_{M_i \in \mathcal{M}} p(\mathbf{r}|M_i), \quad (1.6)$$

trong đó $p(\mathbf{r}|M_i)$ là hàm mật độ xác suất của tín hiệu thu khi giả sử tín hiệu được điều chế theo M_i .

Trong điều kiện lý tưởng, khi thông tin kênh truyền, nhiễu và các tham số hệ thống được biết chính xác, phương pháp này đạt hiệu năng phân loại tối ưu theo tiêu chuẩn Bayes. Tuy nhiên, trong thực tế, các tham số như SNR, pha sóng mang, tần số lệch và đặc tính fading thường không được biết trước hoặc thay đổi theo thời gian. Việc ước lượng đồng thời các tham số này làm tăng đáng kể độ phức tạp tính toán.

Các biến thể chính của phương pháp LB bao gồm [17]:

- **Kiểm định tỷ số hợp lý trung bình (Average likelihood ratio test - ALRT):** Trong thực tế, các tham số của kênh truyền như biên

độ, pha, và tần số thường không biết trước. ALRT xử lý vấn đề này bằng cách coi các tham số chưa biết là các biến ngẫu nhiên với hàm mật độ xác suất (PDF) đã biết trước. Hàm hợp lý sau đó được tính bằng cách lấy tích phân qua không gian của các tham số này. Phương pháp này yêu cầu mô hình thống kê chính xác của các tham số kênh.

- **Kiểm định tỷ số hợp lý tổng quát (Generalized likelihood ratio test - GLRT):** Thay vì lấy tích phân, GLRT thực hiện ước lượng các tham số chưa biết (thường sử dụng Ước lượng hợp lý cực đại - MLE) và thay thế các giá trị ước lượng này vào hàm hợp lý. GLRT thường được áp dụng khi không có thông tin về phân bố xác suất của các tham số kênh.
- **Kiểm định tỷ số hợp lý lai (Hybrid likelihood ratio test - HLRT):** Đây là sự kết hợp giữa ALRT và GLRT, nhằm tận dụng ưu điểm của cả hai phương pháp trong các tình huống cụ thể.

Tuy nhiên, việc áp dụng phương pháp LB vào thực tế gặp phải những rào cản to lớn, chủ yếu xuất phát từ độ phức tạp tính toán và sự nhạy cảm với sai lệch mô hình:

- **Gánh nặng tính toán:** Việc tính toán hàm hợp lý, đặc biệt là trong ALRT, đòi hỏi các phép tích phân phức tạp trên không gian tham số nhiều chiều. Đối với GLRT, quá trình tối ưu hóa để tìm ước lượng tham số cũng tiêu tốn tài nguyên tính toán lớn, làm cho việc triển khai thời gian thực trở nên bất khả thi trên các thiết bị có năng lực xử lý hạn chế.
- **Yêu cầu tri thức tiên nghiệm:** Các phương pháp LB đòi hỏi thông tin chính xác về mô hình kênh truyền và các tham số thống kê của nhiễu.

Trong các môi trường vô tuyến thực tế, nơi kênh truyền biến đổi nhanh do pha đình đa đường và các hiệu ứng phi tuyến, việc duy trì mô hình chính xác là cực kỳ khó khăn. Sự sai lệch nhỏ giữa mô hình lý thuyết và thực tế có thể dẫn đến suy giảm nghiêm trọng hiệu năng phân loại.

Do đó, mặc dù có cơ sở lý thuyết vững chắc và độ chính xác cao trong các kịch bản đơn giản, các phương pháp dựa trên hàm hợp lý AMC thường không phù hợp cho các hệ thống thời gian thực hoặc môi trường kênh truyền phức tạp.

Phương pháp dựa trên trích chọn đặc trưng (Feature-based AMC)

Nhằm khắc phục nhược điểm về độ phức tạp của các phương pháp dựa trên hàm hợp lý, các phương pháp AMC dựa trên trích chọn đặc trưng đã được nghiên cứu rộng rãi và ứng dụng phổ biến hơn trong thực tế [18].

Cách tiếp cận này bao gồm hai bước chính:

- Trích xuất một tập đặc trưng đặc trưng hóa cho tín hiệu thu.
- Sử dụng bộ phân loại (classifier) để gán nhãn điều chế dựa trên tập đặc trưng đã trích xuất.

Các đặc trưng thường được sử dụng có thể phân thành các nhóm sau [19]:

- **Đặc trưng thống kê bậc cao (Higher order statistics – HOS):**

Các mô-men và cumulant bậc cao của tín hiệu được sử dụng để phân biệt các kiểu điều chế khác nhau, đặc biệt là các điều chế tuyến tính như PSK và QAM. Ưu điểm của nhóm đặc trưng này là tính bền vững tương đối đối với nhiễu Gaussian. Tuy nhiên, hiệu năng suy giảm đáng kể trong môi trường fading mạnh.

- **Đặc trưng thống kê tức thời (Instantaneous statistical features):**
Bao gồm các thông tin về biên độ, pha và tần số tức thời của tín hiệu. Ví dụ, sự biến thiên của biên độ tức thời có thể dùng để phân biệt các điều chế biên độ (như ASK, QAM) với các điều chế pha/tần số không đổi (như PSK, FSK).
- **Đặc trưng chu kỳ dừng (Cyclostationary features):** Tín hiệu truyền thông thường mang tính chất chu kỳ dừng do các quá trình lặp lại như điều chế sóng mang, tốc độ ký hiệu, và mã hóa. Phân tích phổ tương quan chu kỳ (Spectral correlation density - SCD) cho phép phát hiện và phân loại tín hiệu ngay cả trong môi trường có tỷ số tín hiệu trên nhiễu (SNR) rất thấp hoặc có sự can nhiễu mạnh, nhờ vào khả năng tách biệt tín hiệu có chu kỳ dừng khỏi nhiễu nền dừng.
- **Đặc trưng dựa trên Biến đổi (Transform-based features):** Sử dụng các phép biến đổi như Biến đổi Wavelet (Wavelet transform) để phân tích các đặc điểm cục bộ của tín hiệu trong miền thời gian-tần số, giúp nắm bắt các sự thay đổi đột ngột hoặc các đặc tính quá độ của tín hiệu điều chế.

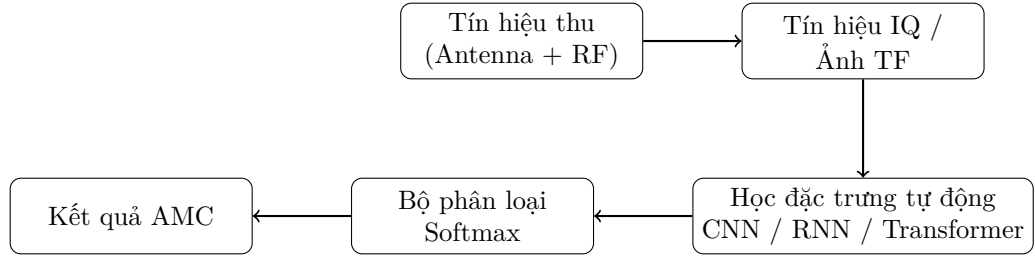
Sau khi trích xuất đặc trưng, các bộ phân loại như KNN, SVM, Decision Tree hoặc mạng nơ-ron nông (shallow neural networks) được sử dụng để thực hiện phân loại điều chế.

Phương pháp FB là độ phức tạp tính toán thấp hơn đáng kể so với LB, giúp chúng dễ dàng triển khai trên phần cứng thực tế. Ngoài ra, các phương pháp FB thường không yêu cầu thông tin tiên nghiệm chi tiết về kênh truyền, làm tăng tính linh hoạt của hệ thống.

Tuy nhiên, hạn chế cốt yếu của phương pháp FB nằm ở tính "thủ công" của quá trình thiết kế đặc trưng. Các bộ đặc trưng thường được tối ưu hóa cho một tập hợp các điều kiện kênh truyền và loại điều chế cụ thể. Khi môi trường vô tuyến thay đổi (ví dụ: xuất hiện các loại nhiễu mới, hiệu ứng pha đình phức tạp), các đặc trưng này có thể mất đi tính hiệu quả. Hơn nữa, việc tìm kiếm một tập đặc trưng tối ưu có khả năng phân biệt tất cả các loại điều chế hiện đại (đặc biệt là các điều chế bậc cao như 256QAM hay 1024QAM) là một thách thức cực lớn. Trong các điều kiện nhiễu phức tạp, hiệu năng của các phương pháp FB thường bị bão hòa và khó cải thiện thêm

Hạn chế của các phương pháp AMC cổ điển

Mặc dù các phương pháp phân loại điều chế tự động (AMC) cổ điển đã được nghiên cứu trong nhiều thập kỷ và đạt được những kết quả nhất định, chúng vẫn tồn tại những hạn chế mang tính bản chất khó khắc phục. Trước hết, các phương pháp này phụ thuộc nặng nề vào quá trình thiết kế đặc trưng thủ công dựa trên giả thiết mô hình kênh và kinh nghiệm chuyên gia, dẫn đến tính linh hoạt thấp và khó thích nghi khi môi trường truyền dẫn thay đổi. Khi xuất hiện nhiễu phức tạp, fading đa đường hoặc sai lệch tần số ngoài phạm vi giả thiết ban đầu, hiệu năng hệ thống suy giảm nhanh chóng. Đặc biệt, trong điều kiện SNR thấp hoặc đối với các điều chế bậc cao có mật độ chòm sao lớn, ranh giới phân lớp trở nên mờ nhạt, khiến các đặc trưng thủ công không còn đủ khả năng phân tách tín hiệu một cách tin cậy. Những hạn chế này cho thấy các phương pháp truyền thống khó đáp ứng yêu cầu về độ chính xác và tính bền vững trong các hệ thống vô tuyến hiện đại, đặc biệt là trong các ứng dụng quân sự và trinh sát điện tử nơi môi trường tác chiến



Hình 1.4: Hệ thống AMC dựa trên Deep Learning với khả năng học đặc trưng tự động.

luôn biến động và khó dự đoán.

Những hạn chế này chính là động lực thúc đẩy sự chuyển dịch từ các phương pháp AMC cổ điển sang các phương pháp dựa trên học máy và học sâu, nơi mà đặc trưng có thể được học tự động trực tiếp từ dữ liệu tín hiệu thô. Nội dung này sẽ được trình bày chi tiết trong Mục sau của luận án.

1.3.2. Tổng quan về phân loại điều chế tự động dựa trên machine learning và deep learning

Sự phát triển mạnh mẽ của học máy (Machine learning – ML) và đặc biệt là học sâu (Deep Learning – DL) trong thập kỷ gần đây đã tạo ra một bước ngoặt quan trọng cho bài toán phân loại điều chế tự động (AMC) [20], [21]. Không giống như các phương pháp cổ điển phụ thuộc vào trích chọn đặc trưng thủ công, các phương pháp dựa trên ML/DL cho phép mô hình tự động học các đặc trưng phân biệt trực tiếp từ dữ liệu tín hiệu, nhờ đó nâng cao đáng kể hiệu năng và khả năng thích nghi trong môi trường kênh truyền phức tạp được minh họa trong Hình 1.3.

AMC dựa trên machine learning truyền thống

Trong giai đoạn đầu, các thuật toán học máy truyền thống được kết hợp với các đặc trưng trích xuất thủ công để thực hiện AMC. Quy trình tổng quát bao gồm hai bước: (i) trích chọn đặc trưng từ tín hiệu thu và (ii) huấn

luyện bộ phân loại học máy [22], [23].

Các bộ phân loại phổ biến được sử dụng bao gồm:

- **K-Nearest neighbors (KNN):** đơn giản, dễ triển khai nhưng độ phức tạp tăng nhanh theo kích thước tập dữ liệu.
- **Support vector machine (SVM):** có khả năng phân tách tốt các lớp điều chế trong không gian đặc trưng, đặc biệt hiệu quả với tập đặc trưng có số chiều vừa phải.
- **Decision tree và random forest:** cho phép mô hình hóa các quan hệ phi tuyến, tuy nhiên dễ bị quá khớp nếu không được điều chỉnh hợp lý.

Các nghiên cứu đã cho thấy rằng AMC dựa trên ML có thể vượt trội hơn các phương pháp cổ điển trong một số kịch bản nhất định, đặc biệt khi tập đặc trưng được thiết kế tốt. Tuy nhiên, hiệu năng của các mô hình này vẫn bị giới hạn bởi chất lượng của đặc trưng đầu vào. Việc thiết kế đặc trưng tối ưu cho mọi loại điều chế và điều kiện kênh là một thách thức lớn, đòi hỏi nhiều kinh nghiệm và khó mở rộng sang các kịch bản mới.

Các phương pháp học sâu trong nhận dạng điều chế tự động

Trong những năm gần đây, học sâu (Deep Learning) đã trở thành một trong những hướng tiếp cận chủ đạo cho bài toán nhận dạng điều chế tự động (AMC) nhờ khả năng học đặc trưng trực tiếp từ dữ liệu mà không cần thiết kế đặc trưng thủ công. Các phương pháp học sâu trong AMC có thể được phân thành một số nhóm chính như mạng nơ-ron tích chập (CNN), mạng nơ-ron hồi tiếp (RNN/LSTM), các mô hình kết hợp CNN-LSTM, mạng có cơ chế chú ý (Attention) và các kiến trúc học sâu hiện đại như ResNet,

DenseNet.

Các phương pháp dựa trên CNN được sử dụng rộng rãi nhờ khả năng tự động trích xuất đặc trưng từ tín hiệu IQ hoặc các biểu diễn thời gian–tần số. O’Shea và cộng sự [8] đã chứng minh rằng CNN có thể học trực tiếp các đặc trưng phân biệt của tín hiệu điều chế mà không cần các bước xử lý tín hiệu truyền thống. Tuy nhiên, CNN chủ yếu khai thác các đặc trưng cục bộ và chưa mô hình hóa hiệu quả mối quan hệ phụ thuộc dài hạn theo thời gian.

Để khắc phục hạn chế này, các nghiên cứu dựa trên RNN và LSTM đã được đề xuất nhằm khai thác các đặc trưng tuần tự của tín hiệu vô tuyến [18]. Các mô hình này có khả năng ghi nhớ thông tin theo thời gian, tuy nhiên thường có số lượng tham số lớn và thời gian huấn luyện kéo dài.

Nhằm tận dụng ưu điểm của cả CNN và LSTM, nhiều nghiên cứu đã đề xuất các kiến trúc lai CNN-LSTM. Các mô hình này sử dụng CNN để trích xuất đặc trưng cục bộ và LSTM để học mối quan hệ theo thời gian, từ đó cải thiện hiệu năng nhận dạng trong các môi trường fading và SNR thấp [24].

Gần đây, các cơ chế chú ý (Attention) và mạng residual cũng được áp dụng cho bài toán AMC. Tiêu biểu là nghiên cứu RanNet của Huynh-The và cộng sự [9], trong đó cơ chế residual-attention được tích hợp vào CNN nhằm nâng cao khả năng tập trung vào các đặc trưng quan trọng của tín hiệu. Kết quả cho thấy mô hình đạt hiệu năng vượt trội trong các điều kiện kênh truyền phức tạp.

Ngoài ra, nhóm nghiên cứu của Huynh-The đã đề xuất kiến trúc MCNet [16], một mô hình CNN hiệu quả cho bài toán AMC với khả năng cân bằng giữa độ chính xác và độ phức tạp tính toán. Các nghiên cứu này cho thấy xu hướng phát triển hiện nay tập trung vào việc nâng cao khả năng trích xuất

đặc trưng đồng thời giảm chi phí tính toán để phục vụ triển khai thực tế.

Mặc dù đạt được nhiều kết quả tích cực, các phương pháp hiện nay vẫn tồn tại một số hạn chế như số lượng tham số lớn, độ phức tạp tính toán cao hoặc suy giảm hiệu năng trong điều kiện SNR thấp và fading nghiêm trọng. Bên cạnh đó, khả năng triển khai thời gian thực trên các nền tảng phần cứng còn chưa được nghiên cứu đầy đủ. Đây chính là khoảng trống nghiên cứu mà luận án hướng tới giải quyết thông qua việc đề xuất các kiến trúc học sâu có khả năng trích xuất đặc trưng hiệu quả, duy trì độ chính xác cao trong điều kiện bất lợi và giảm độ phức tạp tính toán.

Bảng 1.1: So sánh các nhóm phương pháp học sâu trong AMC

Nhóm phương pháp	Các mạng/công trình tiêu biểu	Đặc trưng khai thác	Ưu điểm	Hạn chế	Độ phức tạp và triển khai
Mạng nơ-ron tích chập	CNN [12], ResNet [13], MC-Net [15]	Học đặc trưng phân cấp từ tín hiệu	Khả năng học đặc trưng tự động; phép tích chập có khả năng xử lý song song	CNN thuần túy hạn chế trong mô hình hóa quan hệ thời gian dài; mạng sâu có thể làm tăng chi phí tính toán	Từ thấp đến cao tùy kiến trúc; thuận lợi cho triển khai song song
Mạng nơ-ron kết nối dày	DenseNet [14]	Tái sử dụng đặc trưng giữa các lớp	Tăng khả năng truyền và tái sử dụng thông tin	Có thể làm tăng yêu cầu bộ nhớ	Trung bình–cao
Mạng nơ-ron hồi tiếp	LSTM [17], BiLSTM	Đặc trưng tuần tự và phụ thuộc thời gian	Mô hình hóa tốt quan hệ thời gian	Tính toán tuần tự, khó song song hóa hoàn toàn	Trung bình–cao
Mạng có cơ chế chú ý	ABConv [18], Ran-Net [59]	Tập trung vào đặc trưng quan trọng	Có khả năng lựa chọn thông tin có giá trị và tăng khả năng biểu diễn	Phát sinh thêm phép tính và tham số	Trung bình–cao
Mạng kết hợp tích chập và hồi tiếp	Các kiến trúc CNN–LSTM/BiLSTM [42]	Đặc trưng cục bộ và quan hệ thời gian	Khai thác được hai loại thông tin bổ sung	Kiến trúc phức tạp và chi phí tính toán cao hơn CNN thuần	Trung bình–cao
Mô hình dựa trên cơ chế tự chú ý	Mạng Trans-former [64]	Quan hệ toàn cục giữa các thành phần tín hiệu	Mô hình hóa tốt quan hệ dài hạn	Yêu cầu tài nguyên tính toán và bộ nhớ lớn hơn trong nhiều cấu hình	Cao; cần phần cứng phù hợp
Học kết hợp nhiều mô hình	Các phương pháp học kết hợp [61, 62, 90]	Tính bổ sung giữa nhiều bộ phân loại	Có thể tăng độ ổn định và độ chính xác	Tăng số mô hình, bộ nhớ và chi phí suy luận	Cao hơn mô hình đơn

Bảng 1.1 cho thấy mỗi nhóm mô hình đều có những ưu điểm và hạn chế riêng. Các mô hình CNN có ưu thế về khả năng triển khai nhờ tính toán song song hiệu quả, trong khi RNN/LSTM và các mô hình kết hợp CNN–LSTM có khả năng khai thác thông tin chuỗi tốt hơn nhưng phải đánh đổi bằng độ phức tạp tính toán và thời gian huấn luyện lớn hơn. Các mô hình Attention

và Transformer cải thiện khả năng học phụ thuộc dài hạn nhưng yêu cầu tài nguyên tính toán và bộ nhớ cao. Trên cơ sở đó, luận án đề xuất mô hình nhằm cân bằng giữa độ chính xác nhận dạng và chi phí tính toán, hướng tới khả năng triển khai trong các hệ thống AMC thời gian thực.

Biểu diễn tín hiệu đầu vào: Hiệu quả của các mô hình học sâu trong AMC phụ thuộc đáng kể vào cách biểu diễn tín hiệu đầu vào. Do tín hiệu vô tuyến có bản chất là chuỗi thời gian số phức, các biểu diễn cần đảm bảo vừa bảo toàn thông tin, vừa làm nổi bật các đặc trưng phân biệt. Một cách tiếp cận phổ biến là sử dụng trực tiếp tín hiệu trong miền thời gian dưới dạng hai thành phần I và Q, cho phép giữ nguyên toàn bộ thông tin và phù hợp với các mô hình CNN 1D hoặc mạng lai [25], [26]. Bên cạnh đó, các biểu diễn trong miền thời–tần như STFT, WVD hoặc SPWVD được sử dụng để chuyển tín hiệu sang dạng ảnh, giúp làm nổi bật cấu trúc phổ và sự biến thiên theo thời gian, đặc biệt hữu ích đối với các tín hiệu phi dừng hoặc điều chế bậc cao [27], [28]. Ngoài ra, một số nghiên cứu còn kết hợp đồng thời nhiều dạng biểu diễn nhằm khai thác thông tin bổ sung giữa các miền khác nhau, thường đi kèm với các mô hình đa nhánh để tăng cường khả năng học đặc trưng [29], [30].

Các mô hình học sâu tiêu biểu: Dựa trên đặc điểm của dữ liệu đầu vào, nhiều kiến trúc học sâu đã được đề xuất cho bài toán AMC. Trong đó, mạng nơ-ron tích chập (CNN) là mô hình được sử dụng rộng rãi nhờ khả năng trích xuất hiệu quả các đặc trưng cục bộ và tính bất biến dịch, phù hợp với cả tín hiệu I/Q và các biểu diễn thời gian–tần số [31], [32]. Bên cạnh đó, các mô hình hồi quy như RNN, LSTM hoặc GRU được áp dụng để khai thác tính chất chuỗi và mối quan hệ theo thời gian của tín hiệu, mặc dù thường gặp

hạn chế về chi phí tính toán và khả năng song song hóa [33], [34]. Để tận dụng ưu điểm của cả hai hướng tiếp cận, các mô hình lai CNN–RNN đã được đề xuất, trong đó CNN đảm nhiệm vai trò trích xuất đặc trưng không gian, còn RNN hoặc BiLSTM mô hình hóa phụ thuộc theo thời gian, qua đó đạt hiệu năng cao trong các môi trường kênh phức tạp [35], [36]. Gần đây, các mô hình tiên tiến như ResNet, DenseNet, Attention và Transformer cũng được nghiên cứu nhằm nâng cao khả năng học đặc trưng sâu, cải thiện tốc độ hội tụ và tăng độ ổn định trước nhiễu [37], [38].

Ưu điểm và thách thức: Các phương pháp AMC dựa trên học sâu mang lại nhiều ưu điểm nổi bật, bao gồm khả năng tự động học đặc trưng, giảm phụ thuộc vào kinh nghiệm chuyên gia, và đạt hiệu năng cao trong các điều kiện kênh phức tạp như SNR thấp hoặc fading đa đường. Đồng thời, các mô hình này có khả năng mở rộng linh hoạt cho nhiều dạng điều chế và các bài toán liên quan. Tuy nhiên, vẫn tồn tại một số thách thức đáng kể, chẳng hạn như yêu cầu tập dữ liệu huấn luyện lớn và đa dạng, độ phức tạp tính toán cao gây khó khăn cho triển khai thời gian thực, và khả năng tổng quát hóa còn hạn chế khi áp dụng từ dữ liệu mô phỏng sang dữ liệu thực tế [39], [40].

Xu hướng nâng cao AMC dựa trên học sâu

Các nghiên cứu gần đây cho thấy một số xu hướng quan trọng trong AMC dựa trên học sâu [41]:

- **Tối ưu mô hình mạng:** thiết kế các khối trích xuất đặc trưng chuyên biệt cho tín hiệu vô tuyến, thay vì sử dụng trực tiếp các mô hình thị giác máy tính.

- **Kết hợp tiền xử lý tín hiệu và học sâu:** sử dụng các biểu diễn thời–tần số (STFT, WVD, Radon) để tăng cường cấu trúc đặc trưng đầu vào.
- **Học kết hợp (Ensemble Learning):** kết hợp nhiều mô hình để cải thiện độ ổn định, đặc biệt trong vùng SNR thấp.
- **Mô hình đa nhiệm và đa đầu vào/đầu ra:** đồng thời thực hiện AMC và các nhiệm vụ khác như ước lượng SNR hoặc phân loại tín hiệu.

Mặc dù đạt được nhiều thành tựu, các phương pháp DL cho AMC vẫn đối mặt với các thách thức về chi phí tính toán, khả năng tổng quát và triển khai trên các thiết bị tài nguyên hạn chế. Nội dung tổng quan trong Mục 1.3 và 1.4 là cơ sở lý thuyết quan trọng, làm nền tảng cho các nghiên cứu và mô hình được trình bày chi tiết trong các chương tiếp theo của luận án.

1.4. Cơ sở lý thuyết mạng Nơ-ron cho xử lý tín hiệu

Sự phát triển vượt bậc của Trí tuệ nhân tạo và học sâu đã mở ra một hướng tiếp cận hoàn toàn mới cho bài toán nhận dạng tín hiệu vô tuyến. Thay vì dựa vào các thuật toán trích xuất đặc trưng thủ công (hand-crafted features) vốn đòi hỏi kiến thức chuyên gia sâu rộng và kém bền vững trước các biến động của môi trường truyền sóng, các mô hình học sâu có khả năng tự động học các biểu diễn đặc trưng phức tạp từ dữ liệu thô. Để xây dựng các mô hình nhận dạng hiệu quả như DFENet và AMC-CNN-BiLSTM được đề xuất trong luận án này, cần thiết lập một nền tảng lý thuyết vững chắc về hai kiến trúc mạng nơ-ron phổ biến và hiệu quả nhất hiện nay: Mạng Nơ-ron tích chập (CNN) và Mạng bộ nhớ dài-ngắn hai chiều (BiLSTM).

1.4.1. Mạng Nơ-ron tích chập (CNN) trong miền thời gian

Mạng Nơ-ron tích chập (CNN) ban đầu được phát triển cho thị giác máy tính, nhưng đã chứng minh được hiệu quả vượt trội trong xử lý tín hiệu chuỗi thời gian, đặc biệt là tín hiệu vô tuyến định dạng I/Q. Khả năng của CNN nằm ở việc khai thác các mối tương quan cục bộ và tính bất biến dịch chuyển của tín hiệu.

Phép toán tích chập và trích xuất đặc trưng cục bộ

Thành phần cốt lõi của CNN là lớp tích chập (Convolutional layer). Đối với tín hiệu đầu vào X có kích thước $L \times 2$ (trong đó $L = 1024$ là chiều dài mẫu và 2 là hai thành phần I/Q), phép toán tích chập thực hiện việc trượt một bộ lọc (kernel/filter) W dọc theo chiều thời gian của tín hiệu. Về mặt toán học, giá trị tại một điểm trên bản đồ đặc trưng (feature map) đầu ra C được tính bằng tổng có trọng số của tích vô hướng giữa bộ lọc và một vùng cục bộ của đầu vào, cộng với một hệ số bias b [42]:

$$C[i] = (X * W)[i] + b = \sum_{k=0}^{K-1} X[i+k] \cdot W[k] + b \quad (1.7)$$

Trong đó K là kích thước của bộ lọc (kernel size). Trong bối cảnh xử lý tín hiệu, mỗi bộ lọc W đóng vai trò như một "bộ lọc phối hợp" (matched filter) học được, chuyên trách việc phát hiện một mẫu hình dạng sóng cụ thể (ví dụ: một sườn xung dựng đứng, một sự thay đổi pha đột ngột, hay một chu kỳ sóng mang). Bộ lọc kích thước nhỏ ($1 \times 1, 3 \times 3$): Có khả năng nắm bắt các chi tiết biến thiên nhanh, tần số cao của tín hiệu, giúp phân biệt các thay đổi tinh tế về biên độ hoặc pha giữa các mẫu liên kề. Bộ lọc kích thước lớn ($5 \times 5, 7 \times 7$): Cung cấp trường thụ cảm (receptive field) rộng hơn, cho phép mạng nhìn thấy ngữ cảnh dài hạn, giúp làm mịn nhiễu và nhận diện các cấu

trúc tín hiệu tổng thể.

Hàm kích hoạt phi tuyến (Non-linear activation)

Phép toán tích chập về bản chất là một phép biến đổi tuyến tính. Tuy nhiên, không gian quyết định của các loại điều chế thường phân tách phi tuyến. Do đó, sau mỗi lớp tích chập, một hàm kích hoạt phi tuyến là bắt buộc để mạng có khả năng xấp xỉ các hàm phức tạp. Luận án sử dụng hàm ReLU (Rectified Linear Unit), được định nghĩa là [43]:

$$Y = \text{ReLU}(C) = \max(0, C) \quad (1.8)$$

Về mặt lý thuyết tín hiệu, ReLU hoạt động như một bộ chỉnh lưu nửa chu kỳ. Nó loại bỏ các giá trị âm (thường đại diện cho nhiễu hoặc các đặc trưng không được kích hoạt) và giữ nguyên các giá trị dương. Ưu điểm vượt trội của ReLU so với các hàm Sigmoid hay Tanh truyền thống là tính toán đơn giản và khả năng giảm thiểu hiện tượng biến mất đạo hàm (vanishing gradient) khi huấn luyện các mạng sâu, cho phép mô hình hội tụ nhanh hơn.

Chuẩn hóa theo lô (Batch normalization - BN)

Trong quá trình huấn luyện mạng sâu, sự thay đổi liên tục của phân phối dữ liệu đầu vào tại các lớp ẩn làm chậm quá trình hội tụ. Lớp batch normalization (BN) giải quyết vấn đề này bằng cách chuẩn hóa đầu ra của lớp tích chập về dạng phân bố chuẩn tắc $N(0, 1)$ trước khi đi vào hàm kích hoạt.

Với một mini-batch $B = \{x_1, \dots, x_m\}$, quá trình BN thực hiện các bước sau:

Tính trung bình:

$$\mu_B = \frac{1}{m} \sum_{i=1}^m x_i \quad (1.9)$$

Tính phương sai:

$$\sigma_B^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2 \quad (1.10)$$

Chuẩn hóa:

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (1.11)$$

Biến đổi tỉ lệ và dịch chuyển:

$$y_i = \gamma \hat{x}_i + \beta \quad (1.12)$$

Trong đó γ và β là các tham số có thể học được (learnable parameters). Lớp BN không chỉ giúp ổn định quá trình huấn luyện mà còn đóng vai trò như một bộ điều tiết, giúp giảm thiểu việc mô hình bị quá khớp (overfitting) với nhiều của tập huấn luyện.

1.4.2. Mạng Nơ-ron hồi quy (RNN) và cơ chế BiLSTM

Mặc dù CNN rất mạnh trong việc trích xuất đặc trưng không gian, nhưng đối với các tín hiệu điều chế số, sự phụ thuộc thời gian giữa các ký hiệu là thông tin cực kỳ quan trọng. Ví dụ, trong các hệ thống có nhớ như CPFSK hay các tín hiệu có mã hóa kênh, trạng thái hiện tại phụ thuộc chặt chẽ vào quá khứ. Mạng Nơ-ron Hồi quy (RNN), và cụ thể là LSTM, được thiết kế để giải quyết vấn đề này.

Mạng bộ nhớ dài-ngắn (LSTM)

RNN truyền thống gặp khó khăn trong việc học các phụ thuộc dài hạn do vấn đề biến mất đạo hàm qua thời gian. LSTM (Long short-term memory) khắc phục điều này bằng cách đưa vào một trạng thái tế bào (cell state) C_t

chạy xuyên suốt chuỗi thời gian, được điều khiển bởi các "cổng" (gates). Tại mỗi bước thời gian t , một đơn vị LSTM thực hiện các tính toán sau:

Cổng quên (Forget gate): quyết định lượng thông tin từ quá khứ cần loại bỏ.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1.13)$$

Cổng vào (Input gate): quyết định lượng thông tin mới x_t sẽ được lưu vào trạng thái tế bào.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (1.14)$$

Cập nhật trạng thái:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (1.15)$$

Cổng ra (Output gate): tính toán trạng thái ẩn h_t để chuyển sang bước tiếp theo.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) h_t = o_t * \tanh(C_t) \quad (1.16)$$

Cơ chế này cho phép LSTM "nhớ" các đặc trưng quan trọng (như chu kỳ sóng mang, quy luật thay đổi pha) trong khoảng thời gian dài và "quên" các nhiễu tức thời.

LSTM hai chiều (Bidirectional LSTM - BiLSTM)

Trong bài toán phân loại điều chế, chúng ta có toàn bộ chuỗi tín hiệu. Việc chỉ phân tích từ quá khứ đến hiện tại (như LSTM thường) là chưa tận dụng hết thông tin. Kiến trúc BiLSTM bao gồm hai lớp LSTM chạy song song ngược chiều nhau [44], [45]:

Lớp Forward: Xử lý chuỗi từ $t = 1$ đến L , tính toán trạng thái ẩn xuôi $\vec{h}_t$.

$$\vec{h}_t = \text{LSTM}(x_t, \vec{h}_{t-1}) \quad (1.17)$$

Lớp Backward: Xử lý chuỗi từ $t = L$ về 1, tính toán trạng thái ẩn ngược $\overleftarrow{h}_t$.

$$\overleftarrow{h}_t = \text{LSTM}(x_t, \overleftarrow{h}_{t+1}) \quad (1.18)$$

Trạng thái ẩn đầu ra cuối cùng tại mỗi thời điểm là sự kết hợp (thường là nối vector) của cả hai hướng:

$$h_t = [\vec{h}_t, \overleftarrow{h}_t] \quad (1.19)$$

Kiến trúc này cho phép mạng tại mỗi thời điểm "nhìn thấy" toàn bộ ngữ cảnh (cả quá khứ và tương lai) xung quanh mẫu tín hiệu đó. Điều này đặc biệt hữu ích khi xử lý các đoạn tín hiệu bị nhiễu nặng; thông tin từ các mẫu lân cận rõ ràng hơn (ở cả hai phía) sẽ hỗ trợ mạng suy luận và khôi phục lại đặc trưng của mẫu bị nhiễu, từ đó nâng cao đáng kể độ chính xác phân loại.

1.4.3. Tầng phân loại (Classification layer)

Sau khi các đặc trưng không gian và thời gian đã được trích xuất bởi CNN và BiLSTM, chúng được đưa vào một lớp kết nối đầy đủ (Fully connected - FC). Lớp này thực hiện việc chiếu vector đặc trưng vào không gian của các lớp điều chế (26 lớp trong bộ dữ liệu HisarMod2019.1 [33]). Cuối cùng, hàm softmax được sử dụng để chuyển đổi các giá trị đầu ra thô (logits) thành phân phối xác suất:

$$P(y = i|X) = \frac{e^{z_i}}{\sum_{j=1}^N e^{z_j}} \quad (1.20)$$

Trong đó N là số lượng lớp điều chế. Nhãn có xác suất cao nhất sẽ được chọn là kết quả dự đoán.

Để tối ưu hóa các tham số W và b của toàn bộ mạng, hàm mất mát Cross-Entropy được sử dụng để đo lường sự sai lệch giữa phân phối xác suất dự

đoán và nhân thực tế, từ đó tính toán gradient để cập nhật trọng số qua thuật toán Lan truyền ngược (Backpropagation).

1.5. Tập dữ liệu mô phỏng

Luận án sử dụng hai bộ dữ liệu RadioML2018 và HisarMod2019 nhằm đảm bảo đánh giá toàn diện mô hình. RadioML2018 là bộ dữ liệu chuẩn, được sử dụng rộng rãi trong nghiên cứu AMC, cung cấp nhiều dạng điều chế trong dải SNR rộng và được xây dựng theo quy trình mô phỏng có kiểm soát, phù hợp cho so sánh với các công trình trước [11], [46], [47].

Trong khi đó, HisarMod2019 được thiết kế với mô hình kênh phức tạp hơn, bao gồm fading đa đường và các hiệu ứng nhiễu thực tế, giúp đánh giá khả năng tổng quát hóa và độ ổn định của mô hình trong môi trường gần thực tế [33], [48], [49].

Việc kết hợp hai bộ dữ liệu này cho phép đánh giá đồng thời hiệu năng trong điều kiện lý tưởng và khả năng thích nghi trong môi trường kênh phức tạp, từ đó nâng cao độ tin cậy của kết quả nghiên cứu.

1.5.1. Tập dữ liệu HisarMod2019.1

Bộ dữ liệu HisarMod2019.1 được xây dựng nhằm phục vụ bài toán phân loại điều chế tự động trong các hệ thống thông tin vô tuyến [33], bao gồm tổng cộng 26 kiểu điều chế thuộc năm nhóm điều chế chính, bao gồm: điều chế tương tự, điều chế dịch tần (FSK), điều chế biên độ xung (PAM), điều chế dịch pha (PSK) và điều chế biên độ cầu phương (QAM). Danh sách chi tiết các kiểu điều chế được trình bày trong Bảng 1.2.

Đối với mỗi kiểu điều chế, bộ dữ liệu cung cấp 1500 tín hiệu độc lập, trong đó mỗi tín hiệu được biểu diễn dưới dạng chuỗi mẫu I/Q với độ dài 1024

Bảng 1.2: Tập dữ liệu HisarMod2019.1

Loại điều chế	Các kiểu điều chế	Loại điều chế	Các kiểu điều chế
Điều chế tương tự	AM-DSB	PSK	BPSK
	AM-SC		QPSK
	AM-USB		8-PSK
	AM-LSB		16-PSK
	FM		32-PSK
	PM		64-PSK
FSK	2-FSK	QAM	4-QAM
	4-FSK		8-QAM
	8-FSK		16-QAM
	16-FSK		32-QAM
PAM	4-PAM		64-QAM
	8-PAM		128-QAM
	16-PAM		256-QAM

mẫu. Các tín hiệu được tạo tương ứng với 20 mức tỷ số tín hiệu trên tạp âm (SNR) khác nhau, phân bố đều trong khoảng từ -20 dB đến $+18$ dB. Do đó, tổng số mẫu tín hiệu trong toàn bộ tập dữ liệu đạt 780 000 tín hiệu. Trong quá trình sinh tín hiệu, bộ lọc định dạng xung cosine nâng được sử dụng với hệ số roll-off bằng 0.35 nhằm mô phỏng đặc tính phổ của các hệ thống truyền thông thực tế.

Ngoài nhiễu Gauss trắng cộng tính (AWGN), các tín hiệu còn chịu ảnh hưởng của năm mô hình kênh truyền thông không dây khác nhau, bao gồm: kênh lý tưởng, kênh tĩnh, kênh Rayleigh, kênh pha-đỉnh Rician với hệ số $k = 3$, và kênh Nakagami- m với tham số $m = 2$. Các mô hình kênh này được phân bố đồng đều trong tập dữ liệu, tương ứng với 300 tín hiệu cho mỗi kiểu điều chế tại mỗi mức SNR.

Cụ thể, kênh lý tưởng mô phỏng trường hợp truyền dẫn không xảy ra hiện tượng pha-đỉnh nhưng có sự hiện diện của nhiễu AWGN. Trong mô hình kênh tĩnh, các hệ số kênh được khởi tạo ngẫu nhiên tại thời điểm ban đầu và giữ

Bảng 1.3: Các kiểu điều chế trong tập dữ liệu RadioML 2018.01A

STT	Kiểu điều chế	STT	Kiểu điều chế
1	BPSK	13	CPFSK
2	QPSK	14	GFSK
3	8PSK	15	PAM4
4	16QAM	16	AM-DSB
5	64QAM	17	AM-SSB
6	OQPSK	18	FM
7	$\pi/4$ -DQPSK	19	WBFM
8	DQPSK	20	CPAM
9	APSK	21	8QAM
10	32QAM	22	16APSK
11	128QAM	23	32APSK
12	256QAM	24	FSK

không đổi trong suốt quá trình truyền. Kênh Rayleigh được sử dụng để mô phỏng các điều kiện truyền không có đường truyền thẳng (non-line-of-sight – NLOS). Trong khi đó, kênh pha-đỉnh Rician với tham số $k = 3$ đại diện cho các kịch bản pha-đỉnh nhẹ, nơi tồn tại thành phần truyền thẳng chiếm ưu thế. Cuối cùng, kênh Nakagami- m với tham số $m = 2$ được sử dụng để mô phỏng các điều kiện pha-đỉnh đa dạng hơn, cho phép tập dữ liệu phản ánh toàn diện các đặc tính kênh truyền trong môi trường thông tin vô tuyến thực tế.

1.5.2. Tập dữ liệu RadioML 2018.01A

Bộ dữ liệu RadioML 2018.01A là một trong những tập dữ liệu chuẩn được sử dụng rộng rãi trong nghiên cứu và đánh giá các mô hình học máy và học sâu cho bài toán phân loại điều chế tự động được trình bày trong Bảng 1.3. Bộ dữ liệu được xây dựng nhằm phản ánh các điều kiện truyền thông vô tuyến thực tế, đặc biệt trong môi trường có nhiễu và pha-đỉnh [11].

RadioML 2018.01A bao gồm tổng cộng 24 kiểu điều chế khác nhau, được

chia thành hai nhóm chính là điều chế tương tự và điều chế số. Các kiểu điều chế tương tự bao gồm: AM-DSB, AM-SSB, và FM; trong khi nhóm điều chế số bao gồm các dạng điều chế phổ biến như BPSK, QPSK, 8PSK, 16QAM, 64QAM, GFSK, CPFSK, PAM4, cùng nhiều dạng điều chế khác. Tập hợp này cho phép đánh giá toàn diện khả năng phân biệt giữa các dạng điều chế có đặc tính phổ và cấu trúc tín hiệu gần nhau.

Mỗi mẫu tín hiệu trong bộ dữ liệu được biểu diễn dưới dạng tín hiệu I/Q với độ dài 1024 mẫu. Các tín hiệu được tạo tương ứng với nhiều mức tỷ số tín hiệu trên tạp âm (SNR) khác nhau, trải dài từ -20 dB đến 30 dB với bước tăng 2 dB. Điều này giúp mô phỏng đầy đủ các kịch bản truyền thông từ điều kiện nhiễu nghiêm trọng đến điều kiện kênh thuận lợi.

Trong quá trình sinh dữ liệu, các tín hiệu được truyền qua kênh có nhiễu Gauss trắng cộng tính (AWGN) và chịu ảnh hưởng của các hiệu ứng suy hao, lệch tần số, lệch pha, trễ thời gian và các biến dạng phi tuyến, nhằm phản ánh sát hơn các điều kiện truyền dẫn trong các hệ thống vô tuyến thực tế. Nhờ đó, RadioML 2018.01A trở thành một chuẩn đánh giá quan trọng cho khả năng tổng quát hóa và độ ổn định của các mô hình AMC trong môi trường kênh phức tạp.

1.5.3. Các tiêu chí đánh giá hiệu năng mô hình

Trong bài toán phân loại đa lớp với N lớp, hiệu năng của mô hình được đánh giá thông qua các chỉ số chính bao gồm Accuracy, Precision, Recall và F1-score. Các chỉ số này được tính toán dựa trên ma trận nhầm lẫn (confusion matrix), trong đó mỗi phần tử phản ánh mối quan hệ giữa nhãn thực tế và nhãn dự đoán, như minh họa tại Bảng 1.4.

Bảng 1.4: Ma trận nhầm lẫn với bài toán phân loại 3 lớp A, B và C.

Thực tế / Dự đoán	A	B	C
A	TP_A	FP_B	FP_C
B	FN_A	TP_B	FP_C
C	FN_A	FN_B	TP_C

Trong đó, TP_k là số mẫu được dự đoán đúng thuộc lớp k , FP_k là số mẫu bị phân loại nhầm vào lớp k , và FN_k là số mẫu thuộc lớp k nhưng bị phân loại sai sang lớp khác.

Các chỉ số đánh giá Độ chính xác (Accuracy) phản ánh tỷ lệ dự đoán đúng trên toàn bộ tập dữ liệu:

$$\text{Accuracy} = \frac{\sum_{k=1}^N TP_k}{\text{Tổng số mẫu}} \quad (1.21)$$

Precision, Recall và F1-score được định nghĩa như sau:

$$\text{Precision} = \frac{TP}{TP + FP}, \quad \text{Recall} = \frac{TP}{TP + FN} \quad (1.22)$$

$$\text{F1-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (1.23)$$

Trong bài toán AMC đa lớp, các chỉ số trên được tính riêng cho từng lớp và sau đó lấy trung bình. Để đảm bảo tính công bằng giữa các lớp, đặc biệt khi dữ liệu không cân bằng, luận án sử dụng phương pháp macro-averaging:

$$\text{Macro-F1} = \frac{1}{C} \sum_{i=1}^C \text{F1}_i \quad (1.24)$$

Các chỉ số này phản ánh các khía cạnh khác nhau của hiệu năng: Precision đánh giá độ tin cậy của dự đoán, Recall thể hiện khả năng phát hiện đúng mẫu, và F1-score cung cấp thước đo cân bằng giữa hai yếu tố này.

Đánh giá độ phức tạp và chi phí tính toán

Bên cạnh hiệu năng phân loại, độ phức tạp của mô hình là yếu tố quan trọng ảnh hưởng đến khả năng triển khai thực tế. Độ phức tạp được đánh giá thông qua số lượng tham số, số phép toán (FLOPs) và thời gian xử lý.

Số lượng tham số phản ánh khả năng biểu diễn của mô hình, nhưng nếu quá lớn có thể dẫn đến hiện tượng quá khớp và tăng yêu cầu bộ nhớ. FLOPs biểu thị chi phí tính toán lý thuyết, đặc biệt đối với các lớp tích chập:

$$FLOPs_{conv} = 2 \times H \times W \times C_{in} \times K^2 \times C_{out} \quad (1.25)$$

Thời gian huấn luyện và thời gian suy luận (inference latency) lần lượt phản ánh chi phí trong giai đoạn học và khi triển khai. Thời gian suy luận đặc biệt quan trọng đối với các hệ thống thời gian thực.

Ước lượng thời gian suy luận

Thời gian suy luận có thể được xấp xỉ dựa trên FLOPs và khả năng tính toán của phần cứng:

$$T_{infer} \approx \frac{FLOPs}{Throughput_{hardware}} \quad (1.26)$$

Trong các ứng dụng như hệ thống nháy tần, mô hình cần đáp ứng ràng buộc thời gian:

$$T_{infer} < T_{hop} \quad (1.27)$$

trong đó T_{hop} là thời gian tín hiệu tồn tại trên một tần số. Điều kiện này đảm bảo mô hình có thể đưa ra quyết định kịp thời trong môi trường hoạt động thực tế.

1.6. Các công trình nghiên cứu có liên quan

1.6.1. Các công trình quốc tế nổi bật

Nghiên cứu tự động phân loại điều chế tín hiệu đã trở thành một chủ đề nghiên cứu tích cực trong cộng đồng quốc tế, đặc biệt trong bối cảnh các hệ thống truyền thông thông minh và mạng vô tuyến thế hệ mới. Các công trình nghiên cứu này có thể được phân thành hai hướng chính: các nghiên cứu dựa trên mô hình học sâu và các giải pháp kết hợp nhiều mô hình để cải thiện độ chính xác AMC dưới các điều kiện kênh phức tạp.

Một số công trình quốc tế gần đây đã tập trung vào việc cải thiện hiệu năng AMC bằng học sâu và các kiến trúc lai:

- **Khảo sát tổng quan kiến trúc Deep Learning cho AMC:** Huynh-The và cộng sự cung cấp một bài khảo sát toàn diện về các kiến trúc DL ứng dụng trong AMC, bao gồm CNN, RNN, LSTM và các biến thể phức tạp hơn. Nghiên cứu này phân tích sâu cách các mô hình học sâu trích xuất đặc trưng từ các biểu diễn tín hiệu khác nhau như chuỗi I/Q, ảnh phổ và ảnh chòm sao, nhằm đối phó với các suy giảm do kênh truyền gây ra. Đồng thời, khảo sát chỉ ra những thách thức còn tồn tại và xu hướng phát triển trong tương lai của AMC dựa trên DL [50].
- **Mô hình CNN kết hợp đặc trưng hỗn hợp cho AMC:** Shah và cộng sự đề xuất một kiến trúc CNN kết hợp biểu diễn I/Q (thành phần thực và ảo) với biểu diễn polar (biên độ và pha), nhằm mở rộng không gian đặc trưng và cải thiện khả năng phân biệt đối với các điều chế bậc cao. Mô hình này đạt độ chính xác phân loại lên tới 96.06% trên bộ dữ liệu RadioML2018 tại SNR cao, cho thấy hiệu quả của việc sử dụng phối

hợp nhiều dạng đầu vào [51].

- **Ứng dụng kỹ thuật kết hợp CNN và LSTM trong AMC:** Trong các nghiên cứu ứng dụng vào các môi trường thực tế như UAV hoặc tín hiệu trong fading, phương pháp lai CNN–LSTM (hoặc CNN–BiLSTM) cho phép mạng vừa trích được các đặc trưng không gian cục bộ, vừa nắm bắt được phụ thuộc theo thời gian của chuỗi tín hiệu. Ví dụ, các mô hình kết hợp CNN và LSTM đã chứng minh độ chính xác sử dụng trên dữ liệu thu thực, vượt trội hơn các mô hình đơn lẻ trong dải SNR từ 0 đến 20 dB.
- **Cấu trúc CNN mạnh mẽ cho môi trường kênh thực:** Một số nghiên cứu gần đây đã đề xuất kiến trúc CNN có khối tích chập song song với kernel bất đối xứng và kết nối bỏ qua (skip connection) để cải thiện việc trích xuất đặc trưng dưới kênh có fading và lệch pha/tần. Các mô hình này đã đạt hiệu năng phân loại cao ở các mức SNR thấp và trung bình trong các điều kiện mô phỏng suy hao thực tế.
- **Học đa quy mô và cơ chế Attention:** Wu và cộng sự đề xuất sử dụng học đại diện đa quy mô với các kernel kích thước lớn kết hợp cơ chế Squeeze-and-Excitation (SE) nhằm tập trung vào các đặc trưng có phân biệt cao. Phương pháp này giúp mạng CNN tập trung tốt hơn vào các đặc trưng quan trọng của tín hiệu, đạt hiệu năng vượt các kiến trúc truyền thống trong nhiều điều kiện kênh và dải SNR khác nhau [52], [53].
- **Mixture-of-Experts (MoE) trong AMC:** Một hướng tiếp cận mới trong phân loại điều chế tự động là áp dụng mô hình Mixture-of-Experts, trong đó hệ thống bao gồm nhiều mạng con được thiết kế để chuyên trách

các miền điều kiện khác nhau của tín hiệu, đặc biệt là các mức SNR khác nhau. Thay vì sử dụng một mô hình duy nhất cho toàn bộ dải SNR, MoE sẽ tự động phân bổ trọng số cho từng expert tùy theo đặc trưng của tín hiệu đầu vào. Cách tiếp cận này cho phép mô hình xử lý hiệu quả sự khác biệt về độ khó giữa tín hiệu ở SNR thấp (nhiều chiếm ưu thế) và SNR cao (đặc trưng điều chế rõ ràng hơn).

- **Học sâu đối kháng theo miền cho AMC:** Các nghiên cứu gần đây cũng ứng dụng học đối kháng theo miền (domain adversarial) nhằm giảm thiểu sự sai lệch phân phối dữ liệu do kênh khác nhau gây ra — một vấn đề quan trọng trong việc triển khai AMC thực tế. Khung học này cải thiện đáng kể độ chính xác so với các mô hình huấn luyện đơn miền trong điều kiện kênh biến động [54].
- **Định lượng độ bất định trong mô hình AMC:** Bên cạnh việc dự đoán nhãn điều chế, xu hướng gần đây còn tập trung vào việc đánh giá độ tin cậy của dự đoán. Một số công trình đề xuất sử dụng kỹ thuật ensemble nhiều mạng CNN hoặc các phương pháp xấp xỉ Bayesian nhằm ước lượng mức độ bất định của mô hình. Thông qua phân tích phân bố xác suất đầu ra hoặc độ phân tán giữa các mô hình thành phần, hệ thống có thể cung cấp thêm thông tin về mức độ tin cậy của từng quyết định phân loại. Cách tiếp cận này đặc biệt quan trọng trong môi trường SNR thấp, nhiễu mạnh hoặc khi xuất hiện dữ liệu ngoài phân phối huấn luyện, góp phần tăng tính an toàn và độ tin cậy khi triển khai AMC trong các hệ thống thực tế. [55], [24].

Ngoài ra, nhiều nghiên cứu còn tập trung vào các chiến lược nâng cao như

tăng cường dữ liệu, thiết kế biểu diễn đầu vào (Biểu đồ mắt, ảnh phổ tần, chòm sao tín hiệu) hay biểu diễn đa kênh để cải thiện tính phân biệt của tín hiệu. Những nghiên cứu này làm nổi bật xu hướng áp dụng các kiến trúc phức tạp hơn nhằm xử lý tốt các điều chế bậc cao và môi trường kênh thực tế.

1.6.2. Các nghiên cứu trong nước

Hiện nay, ở Việt Nam các đơn vị trực của Cục Tần số vô tuyến điện thuộc Bộ Thông tin và Truyền thông được trang bị các hệ thống chuyên dụng phục vụ cho công tác quản lý và giám sát tín hiệu vô tuyến. Những thiết bị này phần lớn được đầu tư mua sắm từ các nhà cung cấp nước ngoài. Cùng với đó, các đơn vị thuộc Cục Tác chiến điện tử thuộc Bộ Quốc phòng cũng được đầu tư trang thiết bị đặc chủng, hiện đại phục vụ cho nhiệm vụ trinh sát điện tử và tác chiến điện tử. Ngoài ra, một số đề tài nghiên cứu tại các đơn vị nhà trường quân đội, viện nghiên cứu và trung tâm nghiên cứu có lĩnh vực liên quan cũng tiến hành các nghiên cứu nhằm nâng cao hiệu quả giám sát tín hiệu vô tuyến và phục vụ cho nhiệm vụ trinh sát điện tử, bảo đảm an toàn, an ninh quốc phòng. Một số đơn vị thuộc Bộ Công an cũng được trang bị các thiết bị thu, giám sát tín hiệu vô tuyến để đảm bảo an ninh quốc gia. Bằng cách sử dụng công cụ tìm kiếm của cơ sở dữ liệu cổng trích dẫn Việt Nam (Vietnam citation gateway) của Đại học Quốc gia Hà Nội để tìm các công bố khoa học liên quan với các từ khóa “tự động phân loại điều chế”, “giám sát tín hiệu vô tuyến”, và định hướng nguồn tín hiệu”, kết quả tìm kiếm cho hai cụm từ khóa đầu không cho ra các bài báo có nội dung về tự động phân loại điều chế tín hiệu vô tuyến. Từ khóa thứ ba cho một số nghiên cứu liên

quan như sau:

- Năm 2016, PGS. TS Vũ Văn Yêm đã tận dụng tính linh động và khả năng tiết kiệm năng lượng tiêu thụ của hệ thống định hướng đơn kênh để đề xuất một kỹ thuật mới áp dụng cho các hệ thống định hướng vô tuyến được thiết kế dựa trên công nghệ vô tuyến điều khiển bằng phần mềm nhằm ước lượng hướng sóng tới của các tín hiệu được thu bởi dàn ăng ten đồng dạng tròn gồm M phần tử. Theo đó trên thuật toán di truyền đã được áp dụng để xác định ra hướng tới của tín hiệu dựa trên các giá trị đo pha của tín hiệu tới từng phần tử ăng ten.

Ngoài ra, Viện Nghiên cứu và Ứng dụng Công nghệ cao Hòa Lạc giới thiệu một sản phẩm “Thiết bị giám sát vô tuyến”. Đây là thiết bị được sử dụng cho mục đích giám sát, phân tích các nguồn phát tín hiệu vô tuyến trong không gian cụ thể. Thiết bị này có cấu trúc nhỏ gọn, cho phép linh động và thuận tiện trong quá trình sử dụng, kết quả được trực quan hóa dưới dạng đồ thị, số liệu giúp dễ dàng trong việc phân tích, giám sát. Một số thông số kỹ thuật cơ bản của thiết bị như sau:

- Dải tần giám sát: 10kHz-6000MHz.
- Tốc độ giám sát: 2GHz/s khi xử lý 2 kênh hoặc 4 kênh đồng thời;
- Độ phân giải: đến 12.5kHz.
- Hỗ trợ phát hiện và định vị các nguồn bức xạ vô tuyến.
- Giải điều chế các tín hiệu: AM, FM, CW,...
- Xử lý thông minh qua các thuật toán phân tích, so sánh và giải mã các nguồn dữ liệu, từ đó đưa ra báo cáo về kết quả xác định các nguồn tín hiệu vô tuyến thực tế.

1.6.3. Đánh giá chung

Nhìn chung, các công trình quốc tế và trong nước đều đã ghi nhận sự vượt trội của các phương pháp học sâu trong bài toán AMC so với các phương pháp truyền thống, cả về độ chính xác phân loại và khả năng làm việc trong dải SNR rộng. Các nghiên cứu này tạo nền tảng cho việc phát triển các mô hình học sâu tiên tiến hơn như phần nghiên cứu của luận án, bao gồm các kiến trúc CNN tinh chỉnh, các phương pháp học kết hợp và mô hình đa nhiệm cho AMC.

Tuy nhiên, nhiều nghiên cứu hiện nay vẫn chủ yếu tập trung vào việc cải thiện độ chính xác phân loại mà chưa khai thác đầy đủ mối quan hệ giữa các nhiệm vụ liên quan như ước lượng SNR, DOA hay thông tin ngữ cảnh kênh, cũng như chưa tối ưu hóa đồng thời cả hiệu năng và độ phức tạp tính toán. Đây cũng chính là một trong những động lực chính để luận án này đề xuất các hướng nghiên cứu mới nhằm khắc phục những hạn chế nêu trên.

1.6.4. Hướng nghiên cứu chính của luận án

Từ các phân tích trên, có thể nhận diện một số hạn chế chính trong các nghiên cứu hiện tại và định hướng cải tiến như sau:

Thứ nhất, việc khai thác đặc trưng từ tín hiệu IQ vẫn chưa hiệu quả, khi các mô hình chưa tận dụng đầy đủ mối quan hệ giữa biên độ, pha và cấu trúc thời gian. Do đó, cần thiết kế các kiến trúc học sâu chuyên biệt để nâng cao khả năng trích xuất đặc trưng.

Thứ hai, các mô hình còn thiếu khả năng tích hợp thông tin ngữ cảnh của kênh (như SNR, fading), làm hạn chế khả năng thích nghi với môi trường biến đổi.

Thứ ba, độ phức tạp tính toán tăng nhanh khi mở rộng mô hình, đòi hỏi các kiến trúc vừa đảm bảo hiệu năng vừa tối ưu chi phí tính toán.

Thứ tư, khoảng cách giữa dữ liệu mô phỏng và thực tế vẫn là thách thức lớn, yêu cầu các mô hình có khả năng tổng quát hóa và thích nghi tốt hơn.

Xuất phát từ những vấn đề trên, luận án này tập trung theo đuổi các hướng nghiên cứu chính sau:

(i) *Thiết kế mô hình mạng học sâu chuyên biệt cho tín hiệu vô tuyến*

Hướng nghiên cứu thứ nhất của luận án là xây dựng mô hình học sâu chuyên biệt cho bài toán AMC dựa trên tín hiệu IQ thô. Trên cơ sở phân tích đặc tính của tín hiệu vô tuyến, luận án đề xuất mô hình DFENet với khả năng trích xuất đặc trưng đa tỷ lệ và đa nhánh, nhằm khai thác hiệu quả thông tin chứa trong tín hiệu dưới các điều kiện SNR và fading khác nhau.

(ii) *Khai thác đặc trưng không gian và quan hệ phụ thuộc thời gian dài hạn của tín hiệu vô tuyến*

Hướng nghiên cứu thứ hai mở rộng từ nền tảng CNN thuần túy sang mô hình lai CNN-BiLSTM nhằm khai thác đồng thời đặc trưng không gian và quan hệ phụ thuộc thời gian dài hạn của tín hiệu vô tuyến. Trong khi CNN có thể mạnh trong việc trích xuất đặc trưng cục bộ, thì bản chất chuỗi thời gian của tín hiệu điều chế đòi hỏi một cơ chế mô hình hóa động học để nắm bắt các phụ thuộc theo thời gian, đặc biệt trong môi trường kênh biến động. Việc tích hợp BiLSTM vào sau các tầng tích chập cho phép mô hình học được cấu trúc tuần tự hai chiều của tín hiệu, từ đó cải thiện khả năng nhận dạng trong điều kiện SNR thấp và khi chịu ảnh hưởng của fading. Hướng tiếp cận này thể hiện bước tiến từ tối ưu hóa biểu diễn đặc trưng sang mô hình hóa toàn diện động học tín hiệu.

(iii) *Nghiên cứu phương pháp học kết hợp cho AMC*

Hướng nghiên cứu thứ ba tập trung vào xây dựng phương pháp học kết hợp nhằm nâng cao hiệu năng và độ tin cậy của hệ thống AMC ở mức tổng thể. Trên cơ sở các mô hình thành phần đã được phát triển và đánh giá độc lập, luận án đề xuất khung kết hợp nhiều mô hình CNN tại mức đầu ra, khai thác tính bổ sung giữa các không gian đặc trưng khác nhau. Cách tiếp cận này không chỉ hướng tới cải thiện độ chính xác trung bình trên toàn dải SNR mà còn giảm độ biến thiên hiệu năng giữa các điều kiện kênh khác nhau. Đồng thời, luận án phân tích mối quan hệ giữa mức cải thiện hiệu năng và chi phí tính toán, từ đó đưa ra cơ sở khoa học cho việc lựa chọn cấu hình mô hình phù hợp với yêu cầu triển khai thực tế.

(v) *Định hướng ứng dụng và đóng góp khoa học*

Tổng thể, các hướng nghiên cứu của luận án được triển khai theo một lộ trình logic từ nghiên cứu và tối ưu mô hình đơn đến khai thác phương pháp học kết hợp, nhằm từng bước nâng cao hiệu năng phân loại điều chế tự động. Các đóng góp của luận án không chỉ mang ý nghĩa học thuật mà còn có tiềm năng ứng dụng cao trong các hệ thống vô tuyến nhận thức, giám sát phổ tần và tác chiến điện tử.

1.7. Kết luận chương 1

Chương 1 đã trình bày một cách có hệ thống cơ sở lý luận và bối cảnh khoa học của bài toán phân loại điều chế tự động trong các hệ thống thông tin vô tuyến hiện đại. Trước hết, chương này đã làm rõ vai trò trung tâm của bài toán AMC trong kiến trúc lớp vật lý của các hệ thống vô tuyến thông minh, đặc biệt trong các kịch bản như vô tuyến nhận thức, quản lý phổ tần

động, tác chiến điện tử và các hệ thống thông tin thích ứng thể hệ mới.

Tiếp theo, chương 1 đã tổng quan các phương pháp AMC cổ điển dựa trên lý thuyết thống kê và trích chọn đặc trưng thủ công, qua đó chỉ ra những hạn chế cố hữu của các phương pháp này về khả năng khái quát, độ ổn định trước biến đổi kênh truyền cũng như chi phí tính toán trong các kịch bản phức tạp. Trên cơ sở đó, các phương pháp AMC dựa trên Machine Learning và đặc biệt là Deep Learning đã được trình bày và phân tích một cách có hệ thống, cho thấy tiềm năng vượt trội của mô hình học sâu trong việc tự động học đặc trưng và nâng cao hiệu năng phân loại trong môi trường nhiễu động.

Từ những cơ sở lý luận và tổng quan khoa học đã trình bày, chương 2 sẽ tập trung nghiên cứu và đề xuất mô hình mạng nơ-ron sâu mới cho bài toán phân loại điều chế tự động dựa trên tín hiệu IQ và các biểu diễn đặc trưng liên quan. Tiếp theo, chương 3 sẽ trình bày chi tiết các phương pháp nâng cao hiệu năng thông qua học kết hợp ensemble và mô hình AMC-CNN-BiLSTM, cùng với các kết quả thực nghiệm và phân tích chuyên sâu.

Chương 2

NÂNG CAO HIỆU NĂNG PHÂN LOẠI ĐIỀU CHẾ TỰ ĐỘNG SỬ DỤNG MẠNG NƠ-RON TÍCH CHẬP

Như đã trình bày trong chương 1, hiệu năng của hệ thống phân loại điều chế tự động chịu ảnh hưởng trực tiếp bởi hai yếu tố cốt lõi: (i) phương thức biểu diễn tín hiệu đầu vào và (ii) khả năng trích xuất đặc trưng của mô hình học sâu. Trong khi phần lớn các nghiên cứu gần đây tập trung vào việc cải tiến kiến trúc mạng nơ-ron nhằm nâng cao năng lực học biểu diễn, vấn đề lựa chọn và tối ưu hóa dạng biểu diễn tín hiệu đầu vào vẫn chưa được khảo sát một cách hệ thống và toàn diện [56]. Điều này dẫn đến khoảng trống nghiên cứu trong việc đánh giá vai trò của biểu diễn tín hiệu đối với hiệu năng tổng thể của hệ thống AMC.

Trong môi trường truyền dẫn vô tuyến thực tế, bài toán AMC phải đối mặt với nhiều thách thức nghiêm trọng. Tín hiệu thu không chỉ bị suy hao công suất mà còn chịu ảnh hưởng đồng thời của nhiễu Gauss trắng cộng (AWGN), fading đa đường, dịch tần số, hiệu ứng Doppler và sai lệch pha. Các phương pháp nhận dạng dựa trên lý thuyết quyết định (Likelihood-based) có thể đạt được tính tối ưu về mặt lý thuyết trong điều kiện giả thiết kênh truyền lý tưởng, tuy nhiên chúng đòi hỏi chi phí tính toán lớn và yêu cầu biết chính xác các tham số thống kê của kênh truyền. Ngược lại, các phương pháp dựa trên trích chọn đặc trưng thủ công (Feature-based), chẳng hạn như Cumulants hoặc biến đổi Wavelet, phụ thuộc nhiều vào kinh nghiệm chuyên

gia và thường thiếu tính bền vững khi điều kiện kênh thay đổi hoặc mức nhiễu tăng cao [57], [58].

Tín hiệu vô tuyến có thể được biểu diễn dưới nhiều dạng khác nhau, bao gồm chuỗi IQ trong miền thời gian, phổ tần số thông qua biến đổi Fourier nhanh (FFT), hoặc biểu diễn thời gian–tần số như biến đổi Fourier ngắn hạn (STFT). Mỗi dạng biểu diễn cung cấp một góc nhìn khác nhau về cấu trúc năng lượng và sự biến thiên của tín hiệu, ảnh hưởng trực tiếp đến khả năng học đặc trưng của mô hình deep learning. Việc lựa chọn biểu diễn phù hợp không chỉ quyết định mức độ phân tách giữa các lớp điều chế mà còn tác động đến độ phức tạp tính toán và tính ổn định của hệ thống trong môi trường nhiễu và fading.

Vì vậy, chương 2 tập trung khảo sát có hệ thống các phương pháp biểu diễn tín hiệu gồm FFT, STFT và STFT kết hợp biến đổi Radon trên cùng một nền tảng thực nghiệm. Mục tiêu là xác định cấu hình biểu diễn tối ưu, bảo đảm khả năng phân biệt cao giữa các dạng điều chế, đồng thời duy trì độ ổn định trong môi trường nhiễu và fading. Trên cơ sở kết quả định lượng thu được, luận án đề xuất và xây dựng mô hình mạng nơ-ron chuyên biệt DFENet nhằm khai thác hiệu quả dạng biểu diễn đã được tối ưu, bảo đảm tính khách quan, logic và phù hợp với yêu cầu triển khai trong các hệ thống vô tuyến kỹ thuật cao.

2.1. Phân tích ảnh hưởng của cấu trúc thời gian, đặc trưng phổ và nhiễu cộng

Trong bài toán nhận dạng tín hiệu vô tuyến, tín hiệu thu chịu ảnh hưởng của nhiều yếu tố vật lý trong quá trình truyền dẫn. Ba yếu tố quan trọng

bao gồm cấu trúc thời gian do kênh đa đường, đặc trưng phổ do các biến đổi tần số và nhiễu cộng trong môi trường truyền. Khi áp dụng các mô hình học sâu, các tham số vật lý này không còn được biểu diễn tường minh mà được ánh xạ vào không gian đặc trưng phi tuyến. Do đó, việc phân tích rõ các yếu tố này là cơ sở quan trọng để thiết kế mô hình phù hợp.

2.1.1. Cấu trúc thời gian của tín hiệu trong kênh đa đường

Trong môi trường vô tuyến, tín hiệu thu chịu ảnh hưởng của hiện tượng truyền đa đường. Khi đó, tín hiệu rời rạc tại đầu thu được biểu diễn:

$$r[n] = \sum_{i=1}^P a_i s[n - \tau_i] + n[n] \quad (2.1)$$

Trong đó:

- $r[n]$: tín hiệu thu tại thời điểm mẫu n
- $s[n]$: tín hiệu phát
- P : số lượng đường truyền
- $a_i \in \mathbb{C}$: hệ số kênh của đường truyền thứ i (bao gồm suy hao biên độ và dịch pha)
- τ_i : độ trễ của đường truyền thứ i
- $n[n]$: nhiễu cộng

Biểu thức trên cho thấy tín hiệu thu là tổng của nhiều bản sao trễ của tín hiệu phát. Điều này tạo ra cấu trúc phụ thuộc theo thời gian, trong đó giá trị tín hiệu tại một thời điểm phụ thuộc vào các mẫu trước đó.

Trong mạng CNN, khả năng khai thác cấu trúc này phụ thuộc vào vùng cảm nhận, được xác định bởi:

$$RF = k + (k - 1)(L - 1) \quad (2.2)$$

Trong đó:

- RF : vùng cảm nhận
- k : kích thước kernel
- L : số lớp tích chập

Nếu $RF < \max(\tau_i)$ thì mô hình không thể khai thác đầy đủ thông tin đa đường. Ngược lại, nếu RF đủ lớn, mô hình có khả năng học được cấu trúc thời gian của tín hiệu.

2.1.2. Đặc trưng phổ và ảnh hưởng của sai lệch tần số

Tín hiệu trong miền thời gian có thể được biểu diễn trong miền tần số thông qua biến đổi Fourier:

$$S(f) = \sum_{n=-\infty}^{\infty} s[n]e^{-j2\pi fn} \quad (2.3)$$

Trong đó:

- $S(f)$: phổ tín hiệu
- $s[n]$: tín hiệu trong miền thời gian
- f : tần số chuẩn hóa

Trong thực tế, sai lệch tần số giữa bộ phát và bộ thu gây ra biến đổi:

$$r[n] = s[n] \cdot e^{j2\pi\Delta fn} \quad (2.4)$$

Trong đó:

- Δf : sai lệch tần số (frequency offset)

Áp dụng tính chất của biến đổi Fourier, ta có:

$$R(f) = S(f - \Delta f) \quad (2.5)$$

Biểu thức trên cho thấy phổ tín hiệu bị dịch chuyển theo trục tần số. Điều này làm thay đổi vị trí các đặc trưng phổ, ảnh hưởng đến khả năng nhận dạng.

Trong CNN, đặc trưng phổ không được biểu diễn trực tiếp mà được học gián tiếp thông qua các kernel tích chập. Các kernel nhỏ nhạy với thành phần tần số cao, trong khi kernel lớn nắm bắt thành phần tần số thấp.

2.1.3. Ảnh hưởng của nhiễu cộng

Trong môi trường truyền dẫn vô tuyến, tín hiệu thu luôn chịu ảnh hưởng của nhiễu cộng, thường được mô hình hóa dưới dạng nhiễu Gauss phức:

$$r[n] = s[n] + n[n], \quad n[n] \sim \mathcal{CN}(0, \sigma^2) \quad (2.6)$$

Trong đó:

- $s[n]$: tín hiệu phát
- $r[n]$: tín hiệu thu
- $n[n]$: nhiễu Gauss phức có phân phối chuẩn với trung bình bằng 0
- σ^2 : phương sai nhiễu

Mức độ ảnh hưởng của nhiễu được đặc trưng bởi tỷ số tín hiệu trên nhiễu:

$$SNR = \frac{\mathbb{E}[|s[n]|^2]}{\mathbb{E}[|n[n]|^2]} \quad (2.7)$$

Để phân tích sâu hơn ảnh hưởng của nhiễu, ta xét dưới góc độ lý thuyết thông tin. Gọi X là tín hiệu đầu vào (tín hiệu sạch) và N là nhiễu, khi đó tín hiệu quan sát được là:

$$\tilde{X} = X + N \quad (2.8)$$

Entropy của một biến ngẫu nhiên liên tục được định nghĩa:

$$H(X) = - \int p(x) \log p(x) dx \quad (2.9)$$

Trong trường hợp X và N độc lập, entropy của tổng hai biến thỏa mãn:

$$H(X + N) \geq \max\{H(X), H(N)\} \quad (2.10)$$

Đặc biệt, với nhiễu Gauss, entropy đạt giá trị lớn nhất trong các phân phối có cùng phương sai, do đó:

$$H(\tilde{X}) = H(X + N) > H(X) \quad (2.11)$$

Điều này cho thấy rằng khi có nhiễu, độ không chắc chắn của tín hiệu tăng lên. Nói cách khác, phân bố xác suất của tín hiệu trở nên “trải rộng” hơn, dẫn đến khó phân biệt giữa các lớp tín hiệu.

Xét bài toán phân loại với nhãn Y , thông tin tương hỗ giữa tín hiệu và nhãn được định nghĩa:

$$I(X; Y) = H(Y) - H(Y|X) \quad (2.12)$$

Trong đó:

- $H(Y)$: entropy của nhãn (không phụ thuộc vào mô hình)
- $H(Y|X)$: entropy có điều kiện, biểu diễn mức độ không chắc chắn của nhãn khi biết tín hiệu X

Khi tín hiệu bị nhiễu, ta xét:

$$I(\tilde{X}; Y) = H(Y) - H(Y|\tilde{X}) \quad (2.13)$$

Do nhiễu làm mất thông tin, nên:

$$H(Y|\tilde{X}) > H(Y|X) \quad (2.14)$$

suy ra:

$$I(\tilde{X}; Y) < I(X; Y) \quad (2.15)$$

Điều này cho thấy nhiễu làm giảm lượng thông tin hữu ích liên quan đến nhãn, dẫn đến suy giảm khả năng phân loại.

Trong bối cảnh này, mô hình học sâu thực hiện một ánh xạ phi tuyến:

$$Z = f(\tilde{X}) \quad (2.16)$$

với mục tiêu học được đặc trưng Z sao cho:

$$I(Z; Y) \approx \max \quad (2.17)$$

tức là giảm thiểu entropy có điều kiện $H(Y|Z)$, đồng thời loại bỏ các thành phần nhiễu không liên quan.

Có thể hiểu rằng mạng nơ-ron đóng vai trò như một bộ lọc phi tuyến, thực hiện hai nhiệm vụ chính:

- Giảm entropy của tín hiệu bằng cách loại bỏ nhiễu
- Tăng thông tin tương hỗ giữa đặc trưng và nhãn

Như vậy, dưới góc độ lý thuyết thông tin, nhiễu cộng làm suy giảm khả năng phân biệt của tín hiệu thông qua việc làm tăng entropy và giảm thông tin tương hỗ. Do đó, việc thiết kế mô hình cần hướng đến việc học các đặc

trung bền vững, có khả năng giữ lại thông tin quan trọng và loại bỏ các thành phần nhiễu.

Từ các phân tích trên, có thể thấy rằng cấu trúc thời gian, đặc trưng phổ và nhiễu cộng là các yếu tố vật lý quan trọng ảnh hưởng trực tiếp đến đặc trưng của tín hiệu vô tuyến. Khi áp dụng các mô hình học sâu, các tham số này không còn được biểu diễn tường minh mà phụ thuộc vào kiến trúc mạng.

Do đó, việc thiết kế mô hình cần đảm bảo khả năng khai thác cấu trúc thời gian đủ dài, biểu diễn đặc trưng phổ đa tỉ lệ và tăng cường khả năng chống nhiễu. Đây chính là cơ sở để đề xuất các kiến trúc như DFE-Net nhằm nâng cao hiệu năng nhận dạng trong các điều kiện truyền dẫn phức tạp.

2.2. Biểu diễn tín hiệu trong miền tần số và thời gian–tần số

Trong bài toán phân loại điều chế tự động, mục tiêu của quá trình biến đổi tín hiệu là làm nổi bật các đặc trưng phân biệt giữa các dạng điều chế trong khi vẫn bảo toàn tối đa thông tin hữu ích của tín hiệu gốc. Hai hướng biểu diễn phổ biến và có ý nghĩa thực tiễn cao là biểu diễn trong miền tần số thông qua biến đổi Fourier nhanh (FFT) và biểu diễn trong miền thời gian–tần số thông qua biến đổi Fourier ngắn hạn (STFT).

2.2.1. Biểu diễn trong miền tần số bằng FFT

Với tín hiệu rời rạc $x[n]$, biến đổi Fourier rời rạc được xác định bởi:

$$X[k] = \sum_{n=0}^{N-1} x[n]e^{-j2\pi kn/N}, \quad k = 0, 1, \dots, N-1. \quad (2.18)$$

Trong thực tế, FFT được sử dụng để tính nhanh phổ tần số của tín hiệu. Biểu diễn này cho phép quan sát phân bố năng lượng theo tần số, qua đó phản ánh một số đặc điểm của dạng điều chế như độ rộng phổ, sự tập trung

năng lượng hoặc cấu trúc đối xứng phổ.

Ưu điểm của FFT là:

- Chi phí tính toán thấp;
- Dễ triển khai trong hệ thống thời gian thực;
- Phù hợp với các dạng điều chế có đặc trưng phổ rõ rệt.

Tuy nhiên, FFT chỉ cung cấp thông tin trong miền tần số tổng quát và không phản ánh được sự biến thiên của phổ theo thời gian. Đối với tín hiệu vô tuyến trong môi trường fading hoặc có dịch tần số theo thời gian, việc mất thông tin thời gian có thể làm giảm khả năng phân biệt giữa các lớp điều chế.

2.2.2. Biểu diễn thời gian–tần số bằng STFT

Để khắc phục hạn chế của FFT, biến đổi Fourier ngắn hạn (STFT) được sử dụng nhằm mô tả đồng thời đặc trưng thời gian và tần số của tín hiệu. STFT của tín hiệu $x[n]$ được định nghĩa như sau:

$$\text{STFT}\{x[n]\}(m, \omega) = \sum_{n=-\infty}^{+\infty} x[n]w[n-m]e^{-j\omega n}, \quad (2.19)$$

trong đó $w[n]$ là hàm cửa sổ và m là chỉ số dịch cửa sổ theo thời gian.

STFT tạo ra biểu diễn hai chiều (thời gian–tần số), cho phép quan sát sự thay đổi phổ năng lượng theo từng khoảng thời gian ngắn. Đặc điểm này đặc biệt quan trọng đối với các tín hiệu điều chế chịu ảnh hưởng của nhiễu, fading đa đường hoặc hiệu ứng Doppler.

Ưu điểm của STFT gồm:

- Phản ánh được động thái phổ theo thời gian;

- Làm nổi bật cấu trúc năng lượng đặc trưng của từng dạng điều chế;
- Phù hợp với mô hình CNN khi biểu diễn dưới dạng ảnh phổ.

Tuy nhiên, STFT tồn tại đánh đổi giữa độ phân giải thời gian và độ phân giải tần số, phụ thuộc vào lựa chọn độ dài và dạng hàm cửa sổ. Nếu cửa sổ quá ngắn, độ phân giải tần số giảm; ngược lại, nếu cửa sổ quá dài, khả năng phát hiện biến thiên theo thời gian bị hạn chế. Do đó, việc lựa chọn tham số STFT cần được khảo sát thực nghiệm một cách có hệ thống.

2.2.3. Kết quả thực nghiệm

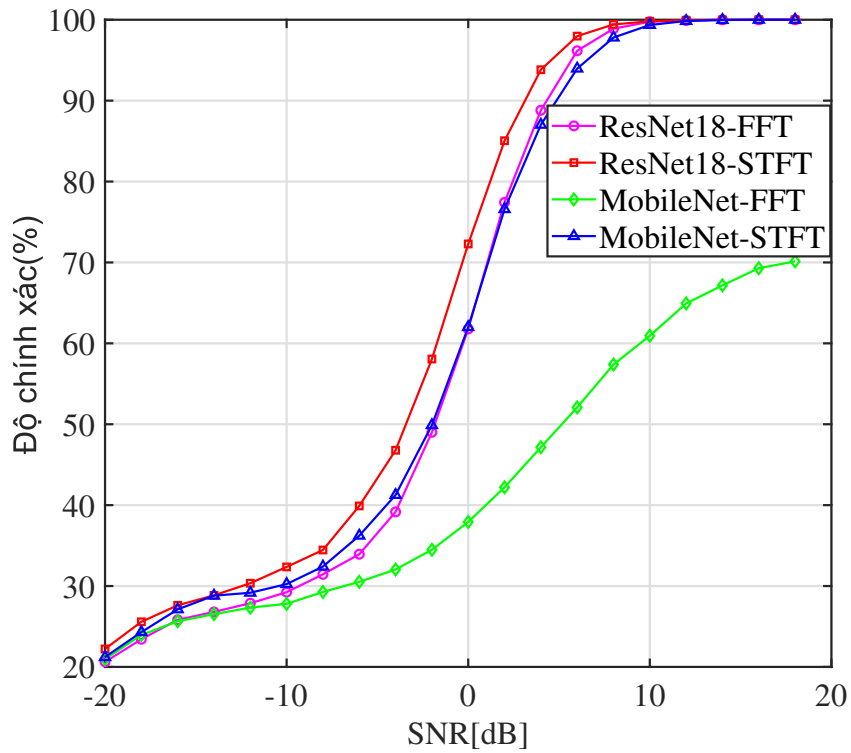
Từ góc độ triển khai hệ thống AMC, FFT có ưu thế về độ đơn giản và chi phí tính toán, trong khi STFT có khả năng biểu diễn đặc trưng phong phú hơn nhưng phức tạp hơn về mặt tham số. Vì vậy, luận án tiến hành đánh giá định lượng hiệu năng phân loại tương ứng với từng dạng biểu diễn trên cùng một mô hình học sâu và tập dữ liệu, nhằm xác định cấu hình phù hợp cho các điều kiện SNR và kênh truyền khác nhau.

So sánh FFT và STFT trên các kiến trúc CNN khác nhau.

Trong phép so sánh đầu tiên, hai kiến trúc CNN phổ biến là ResNet18 và MobileNet được kết hợp chéo với hai phương pháp tiền xử lý tín hiệu, bao gồm FFT và STFT, tạo thành bốn cấu hình mô hình: ResNet18-FFT, ResNet18-STFT, MobileNet-FFT và MobileNet-STFT. Cách tiếp cận này cho phép đánh giá đồng thời ảnh hưởng của dạng đặc trưng đầu vào cũng như năng lực trích xuất đặc trưng của từng kiến trúc mạng trong bài toán phân loại điều chế tự động.

Kết quả thực nghiệm trong Hình 2.1 cho thấy hai xu hướng rõ rệt. Thứ nhất, các mô hình sử dụng biểu diễn STFT luôn đạt độ chính xác cao hơn

so với các mô hình tương ứng sử dụng FFT trên toàn bộ dải SNR. Điều này khẳng định ưu thế của biểu diễn thời-tần số trong việc mô tả đặc tính phi dừng của tín hiệu vô tuyến, đặc biệt trong môi trường nhiễu. Thứ hai, ResNet18 cho hiệu năng vượt trội hơn MobileNet ở mọi cấu hình, phản ánh khả năng trích xuất đặc trưng mạnh mẽ hơn của kiến trúc mạng sâu có liên kết dư.



Hình 2.1: So sánh độ chính xác phân loại các CNN khác nhau sử dụng STFT và FFT.

Đáng chú ý, cấu hình ResNet18-STFT đạt độ chính xác cao nhất, lên tới 95.5% tại $\text{SNR} = +8 \text{ dB}$. Kết quả này cho thấy sự kết hợp giữa đặc trưng STFT giàu cấu trúc và các residual blocks trong ResNet18 giúp mạng học được các đặc trưng phân biệt hiệu quả hơn giữa các dạng điều chế. Ngay cả khi sử dụng đặc trưng FFT, ResNet18 vẫn duy trì hiệu năng tốt hơn MobileNet, cho thấy vai trò quan trọng của kiến trúc CNN trong việc khai

Bảng 2.1: Độ chính xác của mô hình ResNet18 với các hàm cửa sổ khác nhau.

Hàm cửa sổ	−20 dB	−4 dB	6 dB	18 dB	Độ chính xác trung bình
Rectangular	20.44	46.03	98.13	100.00	62.21
Hamming	22.22	46.78	97.97	100.00	64.72
Hanning	21.07	46.32	97.02	100.00	63.80
Kaiser	21.16	45.26	98.09	99.99	64.04
Blackman	21.08	48.06	98.22	99.98	64.86
Chebyshev	22.23	48.13	98.34	100.00	65.18

thác thông tin từ tín hiệu đầu vào.

Ảnh hưởng của các hàm cửa sổ trong STFT đến hiệu năng AMC.

Tiếp theo, nhằm đánh giá tác động của hàm cửa sổ trong biến đổi STFT đến hiệu quả phân loại điều chế, sáu loại cửa sổ khác nhau được xem xét, bao gồm: cửa sổ chữ nhật (Rectangular), Hamming, Hanning, Kaiser, Blackman và Chebyshev. Trong thí nghiệm này, mô hình ResNet18-STFT được lựa chọn làm bộ phân loại cố định để đảm bảo rằng sự thay đổi hiệu năng chỉ đến từ đặc tính của hàm cửa sổ, thay vì kiến trúc mạng.

Các kết quả định lượng được tổng hợp trong Bảng 2.1 cho thấy việc thay đổi hàm cửa sổ chỉ gây ra sự khác biệt tương đối nhỏ về độ chính xác phân loại AMC. Cụ thể, cửa sổ Chebyshev đạt độ chính xác trung bình cao nhất, đạt 65.18%, trong khi cửa sổ chữ nhật cho kết quả thấp nhất với 62.2%. Các cửa sổ còn lại, bao gồm Hamming, Hanning, Kaiser và Blackman, cho hiệu năng gần tương đương nhau, với độ chính xác trung bình xấp xỉ 64%.

Kết quả này cho thấy rằng, mặc dù hàm cửa sổ có ảnh hưởng nhất định đến chất lượng biểu diễn thời–tần số của STFT, tác động của chúng đến hiệu năng phân loại AMC là không quá đáng kể khi sử dụng một kiến trúc CNN đủ mạnh. Điều này hàm ý rằng việc lựa chọn hàm cửa sổ trong thực tế có thể ưu tiên các tiêu chí đơn giản hóa tính toán hoặc tính ổn định số, thay vì

tối ưu hóa quá mức độ chính xác phân loại.

2.2.4. Kết hợp STFT và Radon Transform để tăng cường đặc trưng

Mặc dù biểu diễn thời–tần số dựa trên STFT đã chứng minh được hiệu quả trong việc cung cấp thông tin cục bộ về sự phân bố năng lượng của tín hiệu điều chế, nhưng phổ thời–tần số truyền thống vẫn tồn tại những hạn chế nhất định khi được sử dụng trực tiếp làm đầu vào cho các mô hình CNN. Cụ thể, trong điều kiện nhiễu cao hoặc đối với các dạng điều chế bậc cao, các đặc trưng quan trọng trong phổ thường bị làm mờ, phân tán hoặc chồng lấn, khiến ranh giới giữa các lớp điều chế trở nên kém rõ ràng. Điều này làm giảm khả năng học đặc trưng phân biệt của CNN, đặc biệt ở vùng SNR thấp [59], [60].

Một quan sát quan trọng là nhiều dạng điều chế thể hiện các cấu trúc hình học đặc trưng trong miền thời–tần số. Ví dụ, các tín hiệu FSK thường tạo ra các vệt năng lượng nghiêng theo thời gian, các tín hiệu PSK và QAM thể hiện các mẫu phân bố đối xứng, trong khi các dạng AM và FM tạo ra các cấu trúc liên tục và trơn hơn. Tuy nhiên, các đặc trưng hình học này không được khai thác trực tiếp khi chỉ sử dụng spectrogram thông thường.

Để khắc phục hạn chế trên, luận án đề xuất kết hợp STFT với phép biến đổi Radon (Radon Transform – RT) nhằm tăng cường khả năng biểu diễn các đặc trưng hình học của tín hiệu trong miền thời–tần số. Radon Transform là một công cụ mạnh trong xử lý ảnh, cho phép ánh xạ các cấu trúc đường thẳng hoặc gần đường thẳng trong ảnh hai chiều thành các điểm nổi bật trong không gian tham số. Về mặt toán học, Radon Transform của một ảnh

hai chiều $f(x, y)$ được định nghĩa như sau:

$$\mathcal{R}(\rho, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - \rho) dx dy \quad (2.20)$$

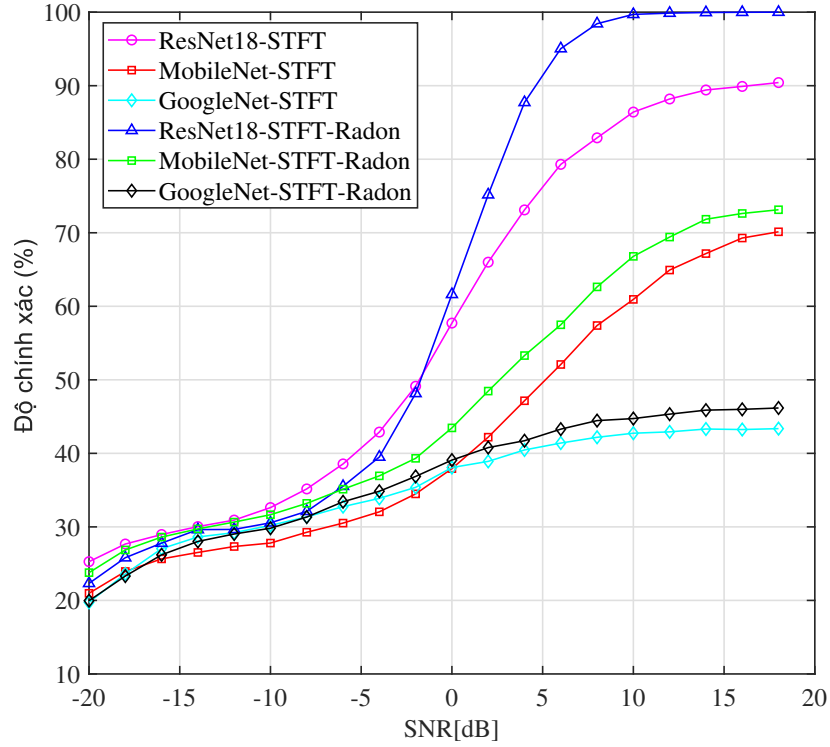
trong đó ρ là khoảng cách từ gốc tọa độ đến đường thẳng, θ là góc nghiêng của đường thẳng, và $\delta(\cdot)$ là hàm Dirac delta.

Khi áp dụng Radon Transform lên phổ tần thu được từ STFT, các vết năng lượng và cấu trúc tuyến tính trong miền thời–tần số sẽ được tích lũy và làm nổi bật trong miền Radon. Quá trình này giúp tăng cường các đặc trưng hình học bền vững trước nhiễu, đồng thời làm giảm ảnh hưởng của các thành phần nhiễu ngẫu nhiên phân bố rời rạc trong ảnh phổ. Do đó, biểu diễn kết hợp STFT–RT không chỉ bảo toàn thông tin thời–tần số, mà còn nhấn mạnh các mẫu cấu trúc mang tính phân biệt cao giữa các dạng điều chế.

Từ góc độ học sâu, ảnh đặc trưng sau khi áp dụng Radon Transform cung cấp một không gian đặc trưng thuận lợi hơn cho CNN. Các đặc trưng sau biến đổi có độ tương phản cao hơn, cấu trúc rõ ràng hơn và ít chồng lấn giữa các lớp điều chế, giúp CNN học được các mẫu phân biệt với số lượng tham số và độ sâu mạng nhỏ hơn. Đặc biệt, trong các kịch bản SNR thấp, nơi spectrogram truyền thống thường bị nhiễu mạnh, biểu diễn STFT–RT vẫn duy trì được các cấu trúc chủ đạo của tín hiệu, từ đó cải thiện đáng kể độ chính xác phân loại.

So sánh STFT và STFT–Radon trên các kiến trúc CNN khác nhau.

Hình 2.2 trình bày kết quả so sánh hiệu năng phân loại AMC của ba kiến trúc CNN phổ biến gồm ResNet18, GoogleNet và MobileNet khi sử dụng hai phương pháp tiền xử lý là STFT và STFT kết hợp Radon Transform (STFT–Radon).



Hình 2.2: So sánh độ chính xác phân loại của các CNN khác nhau sử dụng STFT và STFT-Radon.

Kết quả cho thấy ResNet18 vượt trội hơn rõ rệt so với hai mô hình còn lại trong toàn bộ miền SNR được khảo sát. Đặc biệt, mô hình ResNet18-STFT-Radon đạt độ chính xác cao nhất, lên tới 97.6% tại $\text{SNR} \geq 8$ dB. Hiệu năng này được lý giải bởi khả năng của các khối residual trong ResNet18 giúp khai thác hiệu quả các đặc trưng cấu trúc tuyến tính và hình học do Radon Transform làm nổi bật trong biểu diễn thời–tần số.

Đối với các mô hình nhẹ như GoogleNet và MobileNet, mặc dù độ chính xác tuyệt đối vẫn thấp hơn ResNet18, việc kết hợp STFT với Radon Transform vẫn mang lại cải thiện đáng kể. Cụ thể, độ chính xác trung bình của GoogleNet tăng khoảng 2%, trong khi MobileNet ghi nhận mức cải thiện lên tới 3.5% so với trường hợp chỉ sử dụng STFT. Điều này cho thấy biểu diễn

STFT-Radon giúp bù đắp phần nào hạn chế về năng lực trích xuất đặc trưng của các kiến trúc CNN gọn nhẹ.

Nhìn chung, các kết quả này khẳng định rằng việc tăng cường tiền xử lý bằng Radon Transform không chỉ nâng cao độ chính xác phân loại, mà còn giúp các mô hình CNN có kiến trúc khác nhau khai thác đặc trưng điều chế hiệu quả hơn, đặc biệt trong các điều kiện nhiễu vừa và thấp.

Như vậy, việc kết hợp STFT và Radon Transform không chỉ nâng cao chất lượng biểu diễn đầu vào cho CNN, mà còn đóng vai trò như một bước cầu nối quan trọng giữa xử lý tín hiệu truyền thống và học sâu. Hướng tiếp cận này đặt nền móng cho các nghiên cứu tiếp theo trong luận án, bao gồm việc khai thác nhiều biểu diễn đầu vào đồng thời và xây dựng các kiến trúc học sâu đa nhánh, đa đầu vào nhằm tối ưu hóa hiệu năng phân loại điều chế trong các môi trường truyền thông vô tuyến phức tạp.

2.2.5. So sánh tổng thể với tín hiệu IQ gốc

Hình 2.3 trình bày độ chính xác phân loại theo SNR của sáu mô hình, bao gồm các kiến trúc học trực tiếp từ tín hiệu IQ (ResNet18-IQ, MobileNet-IQ, GoogleNet-IQ) và các mô hình sử dụng đặc trưng thời gian-tần số kết hợp phép biến đổi Radon (ResNet18-STFT-Radon, MobileNet-STFT-Radon, GoogleNet-STFT-Radon).

Kết quả thực nghiệm cho thấy các mô hình học trực tiếp từ tín hiệu IQ đạt hiệu năng vượt trội trên toàn dải SNR. Cụ thể, tại $\text{SNR} = 0 \text{ dB}$, ResNet18-IQ đạt gần 77%, cao hơn khoảng 62% so với ResNet18-STFT-Radon. Với MobileNet và GoogleNet, mức cải thiện tương ứng vào khoảng 61% và 52%. Trung bình trong vùng $\text{SNR} \geq 0 \text{ dB}$, nhóm IQ cải thiện khoảng 10–15% so

với nhóm sử dụng biểu diễn STFT-Radon.

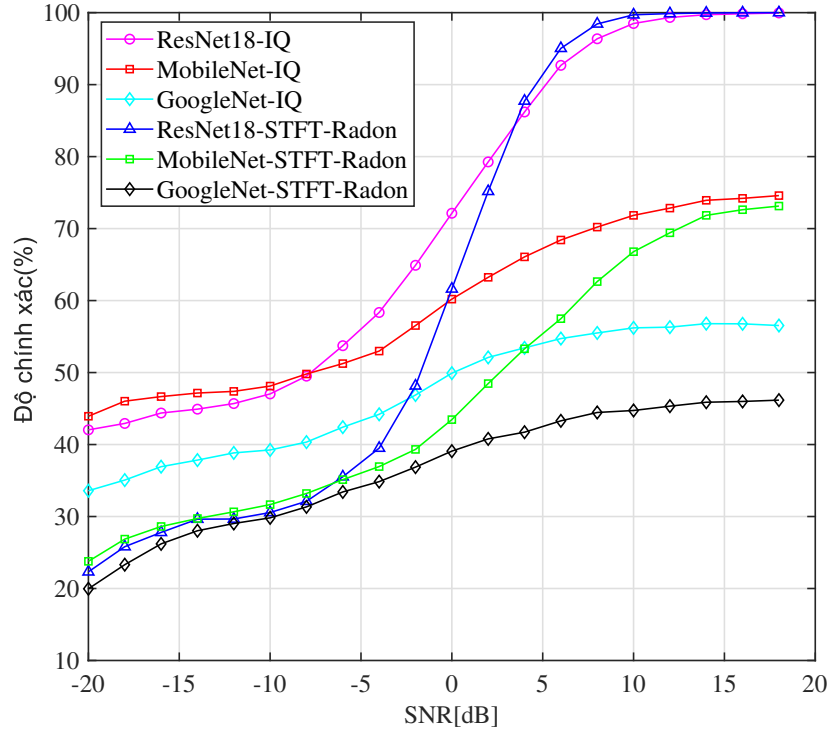
Về mặt lý thuyết, sự suy giảm hiệu năng của nhóm STFT-Radon có thể được lý giải bởi hiện tượng mất thông tin tín hiệu. Phép biến đổi STFT thực hiện chiếu cục bộ tín hiệu lên không gian thời gian-tần số, trong khi phép chiếu Radon tiếp tục tích phân dọc theo các hướng xác định. Quá trình chiếu hai bước này có thể làm mất các thành phần pha vi mô và cấu trúc tương quan phức giữa phần thực và ảo của tín hiệu IQ. Do đó, không gian đặc trưng đầu vào bị rút gọn, làm giảm thông tin phân biệt giữa các sơ đồ điều chế bậc cao.

Biểu diễn STFT-Radon làm giảm độ phức tạp đầu vào và có thể giúp giảm phương sai trong vùng SNR thấp. Tuy nhiên, sự đơn giản hóa quá mức này làm tăng sai lệch mô hình, đặc biệt trong vùng SNR trung bình và cao, nơi thông tin điều chế đã đủ rõ ràng để khai thác trực tiếp từ tín hiệu IQ. Ngược lại, học trực tiếp trên IQ bảo toàn toàn bộ cấu trúc pha-biên độ, cho phép mạng sâu học các đặc trưng bậc cao một cách linh hoạt hơn, từ đó đạt độ chính xác cao hơn và bão hòa sớm hơn theo SNR.

Kết quả tổng hợp cho thấy:

- STFT vượt trội hơn FFT trong môi trường SNR thấp;
- STFT + Radon cải thiện thêm khả năng phân tách lớp so với STFT;
- Tuy nhiên, tín hiệu IQ đạt hiệu năng cao nhất và ổn định nhất trên toàn dải SNR.

Mặc dù STFT kết hợp Radon cho kết quả cao hơn so với các biểu diễn phổ thuần túy, kết quả thực nghiệm cho thấy tín hiệu IQ gốc vẫn đạt độ chính xác cao nhất trên toàn dải SNR. Kết quả này có ý nghĩa quan trọng. Việc áp



Hình 2.3: So sánh độ chính xác phân loại của các CNN khác nhau sử dụng STFT-Radon và IQ.

dụng các biến đổi miền tần số hoặc thời gian-tần số về bản chất là quá trình ánh xạ và tái cấu trúc lại tín hiệu, trong đó có thể xảy ra mất mát thông tin pha hoặc chi tiết vi mô của tín hiệu gốc. Ngoài ra, STFT còn chịu ràng buộc bởi nguyên lý bất định Heisenberg, dẫn đến sự đánh đổi giữa độ phân giải thời gian và tần số.

Trong khi đó, tín hiệu IQ bảo toàn đầy đủ thông tin biên độ và pha, cho phép mô hình học sâu trực tiếp khai thác toàn bộ cấu trúc nguyên thủy của tín hiệu mà không bị giới hạn bởi tham số cửa sổ hoặc biến đổi trung gian. Khi sử dụng kiến trúc đủ mạnh, mạng nơ-ron có khả năng tự học các biểu diễn nội tại tối ưu từ dữ liệu IQ mà không cần phụ thuộc hoàn toàn vào tiền xử lý phức tạp. Điều này cho thấy rằng, mặc dù các kỹ thuật tiền xử lý có

thể hỗ trợ làm nổi bật đặc trưng trực quan, việc bảo toàn thông tin gốc của tín hiệu đóng vai trò quyết định đối với bài toán AMC khi sử dụng học sâu.

2.3. Thiết kế mô hình DFENet

Các kết quả thực nghiệm và phân tích lý thuyết cho thấy rằng việc bảo toàn thông tin gốc của tín hiệu đóng vai trò quyết định trong bài toán AMC.

Thay vì phụ thuộc vào tiền xử lý phức tạp, một hướng tiếp cận hợp lý hơn là thiết kế mô hình mạng có khả năng trực tiếp khai thác cấu trúc đa mức từ tín hiệu IQ nguyên thủy.

Trên cơ sở đó, luận án đề xuất mô hình DFENet (Diverse Feature Extraction Network), được thiết kế nhằm tăng cường khả năng trích xuất đặc trưng đa nhánh và đa thang đo từ dữ liệu IQ. Từ các phân tích trên, luận án lựa chọn tín hiệu IQ làm đầu vào chính cho hệ thống phân loại điều chế. Tuy nhiên, kết quả thực nghiệm cũng chỉ ra rằng việc khai thác hiệu quả dữ liệu IQ đòi hỏi một mô hình mạng có khả năng trích xuất đặc trưng đa mức, đa thang đo và duy trì tính đa dạng trong không gian biểu diễn.

Do đó, thay vì phụ thuộc vào các biến đổi tiền xử lý phức tạp, luận án đề xuất thiết kế một mô hình mạng nơ-ron chuyên biệt có khả năng trực tiếp học các đặc trưng giàu thông tin từ tín hiệu IQ nguyên thủy. Trên cơ sở đó, mô hình DFENet được xây dựng nhằm tăng cường khả năng trích xuất đặc trưng đa nhánh và nâng cao hiệu năng phân loại trong môi trường nhiễu và fading.

2.3.1. Đề xuất mô hình CNN

Dựa trên cơ sở lý thuyết đã trình bày, phần này đi sâu vào chi tiết thiết kế của mô hình đầu tiên: DFENet. Đây là một mô hình CNN chuyên sâu được

tối ưu hóa cho bài toán AMC với triết lý cốt lõi là "Trích xuất đặc trưng đa dạng".

Mục tiêu thiết kế trọng tâm của DFENet là khai thác đa tỉ lệ trong quá trình trích xuất đặc trưng. Bản chất của tín hiệu điều chế cho thấy rằng các mẫu dao động biên độ–pha nằm trên nhiều thang thời gian khác nhau: có những đặc trưng cục bộ thể hiện ở số mẫu rất ngắn (ví dụ biến thiên biên độ của PAM hoặc điểm pha sát nhau của PSK), trong khi một số đặc trưng khác mang tính toàn cục, phụ thuộc vào chu kỳ điều chế hoặc cấu trúc dạng sóng dài hơn (hoặc như dạng chòm sao của QAM bậc cao). Do đó, việc sử dụng một kích thước bộ lọc duy nhất trong mạng CNN truyền thống thường không đủ khả năng mô tả toàn diện đặc trưng của tín hiệu IQ. DFENet giải quyết vấn đề này bằng cách tích hợp cơ chế đa nhánh, trong đó mỗi nhánh đảm nhiệm một vai trò trích xuất đặc trưng trên một thang thời gian–không gian riêng biệt.

Bên cạnh đó, mô hình DFENet chú trọng đến tính ổn định của biểu diễn, vốn là yếu tố cực kỳ quan trọng trong các điều kiện kênh thực tế. Trong nhiều trường hợp, tín hiệu IQ bị nhiễu đến mức không thể phân biệt trực quan bằng các đặc trưng đơn giản, đặc biệt tại SNR thấp. Điều này làm nổi bật nhu cầu về một cơ chế tinh chỉnh – chuẩn hóa đặc trưng sau mỗi tầng học biểu diễn nhằm giảm tác động của nhiễu và đảm bảo sự ổn định của gradient trong quá trình lan truyền ngược. Khối FEN (Feature Enhancement and Normalization) trong DFENet được thiết kế nhằm thực hiện nhiệm vụ này, đóng vai trò loại bỏ nhiễu mang tính ngẫu nhiên, đồng thời tăng cường các đặc trưng có giá trị phân biệt cao. Nhờ cơ chế này, DFENet có khả năng duy trì hiệu năng tốt hơn các mô hình CNN thông thường trong điều kiện

SNR thấp và môi trường fading mạnh.

Một mục tiêu quan trọng khác trong thiết kế DFENet là giảm thiểu độ phức tạp tính toán, nhằm đảm bảo khả năng triển khai thực tế trên các hệ thống không quá mạnh về năng lực xử lý, chẳng hạn như thiết bị IoT, thiết bị SDR nhúng hoặc môi trường triển khai thời gian thực. Trái ngược với các kiến trúc CNN nghiên cứu trong thị giác máy tính vốn rất sâu và nặng về tham số, DFENet hướng tới cấu trúc gọn nhẹ nhưng hiệu quả thông qua hai nguyên tắc:

- (1) sử dụng bộ lọc kích thước nhỏ kết hợp hợp lý với bộ lọc lớn;
- (2) tránh việc ghép nối chiều kênh gây tăng mạnh số chiều và chi phí tính toán.

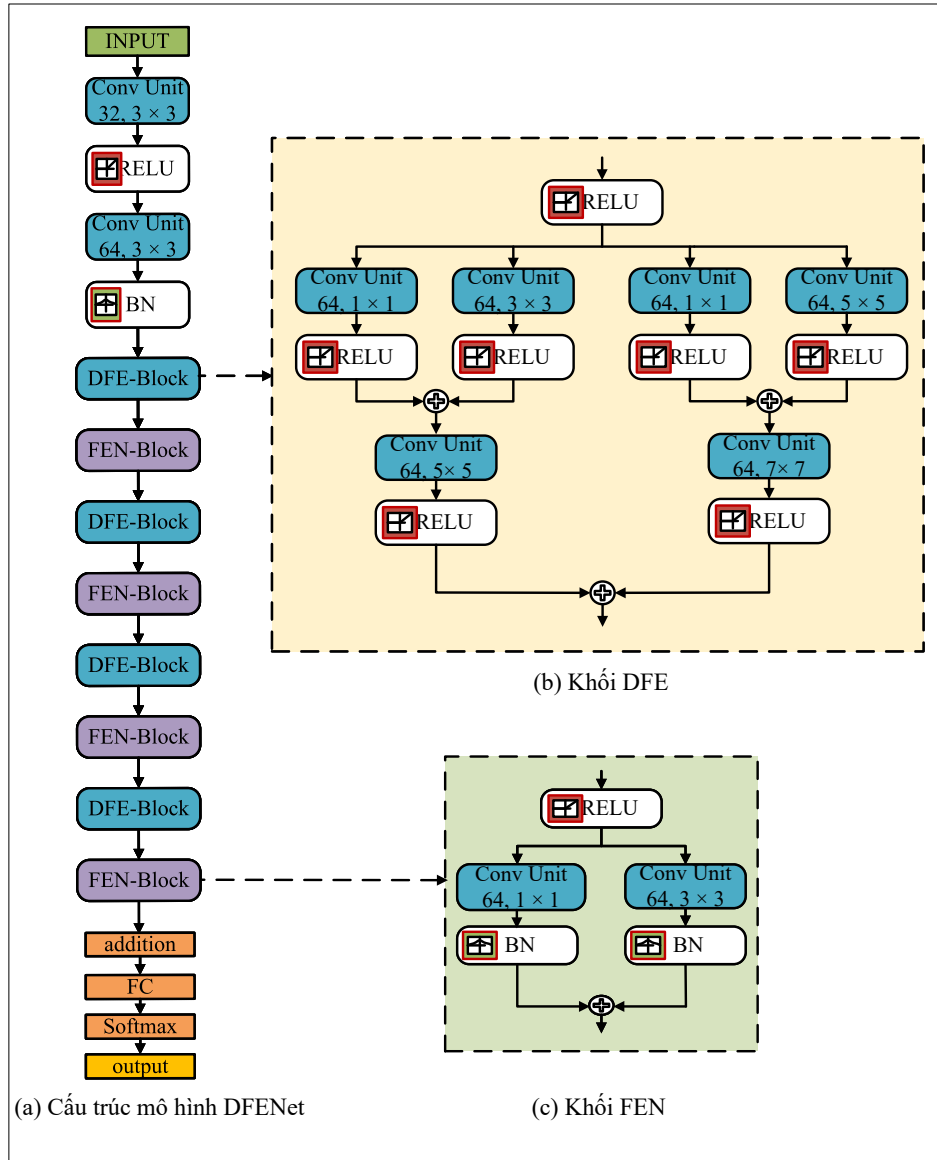
Nhờ vậy, DFENet chỉ sử dụng khoảng 2 triệu tham số – thấp hơn đáng kể so với ResNet, VGG hoặc các mô hình Inception, nhưng vẫn đạt độ chính xác vượt trội trên hai tập dữ liệu thực nghiệm lớn.

DFENet được thiết kế dựa trên ba yếu tố: đa tỉ lệ – ổn định đặc trưng – chi phí tính toán thấp. Ba yếu tố này giúp mô hình không chỉ hoạt động hiệu quả trên toàn phổ SNR mà còn phù hợp để triển khai trong các hệ thống thực tế, đáp ứng nhu cầu ngày càng cao về AMC trong mạng vô tuyến thông minh và các hệ thống thông tin 5G/6G thế hệ mới.

2.3.2. Cấu trúc tổng quan và luồng dữ liệu

Cấu trúc tổng thể của DFENet được xây dựng dựa trên sự kết nối tuần tự của các khối xử lý chuyên biệt, tạo thành một cấu trúc sâu có khả năng học các biểu diễn phân cấp của tín hiệu. Đầu vào của mạng là chuỗi tín hiệu I/Q thô với kích thước cố định 1024×2 . Luồng xử lý dữ liệu đi qua các giai

đoạn chính sau (được minh họa trong Hình 2.4a):



Hình 2.4: Cấu trúc CNN cho AMC: (a) mô hình tổng thể của DFENet, (b) cấu trúc khối DFE, (c) cấu trúc khối FEN.

1. Tầng Khởi tạo: Tín hiệu đầu vào đi qua hai lớp tích chập khởi đầu. Lớp đầu tiên sử dụng 32 bộ lọc kích thước 3×2 , theo sau là hàm kích hoạt ReLU. Lớp thứ hai tăng số lượng bộ lọc lên 64, kết hợp với lớp chuẩn hóa theo lô (Batch Normalization - BN) và ReLU để chuẩn bị các đặc trưng sơ cấp cho các khối xử lý sâu hơn.

2. Khối Trích xuất Đặc trưng (DFE - FEN Sequences): Đây là phần thân chính của mạng, bao gồm sự lặp lại của các cặp khối DFE (Diverse Feature Extraction) và FEN (Feature Extraction and Normalization). Trong cấu hình tối ưu nhất, mạng sử dụng 4 cặp khối này nối tiếp nhau.

- Các khối DFE chịu trách nhiệm mở rộng không gian đặc trưng và nắm bắt các mẫu tín hiệu ở nhiều tỷ lệ thời gian khác nhau.
- Các khối FEN thực hiện nhiệm vụ tinh chỉnh, giảm chiều dữ liệu (downsampling) và chuẩn hóa để ổn định quá trình huấn luyện.

3. Tầng Phân loại (Classification): Đầu ra cuối cùng từ chuỗi các khối đặc trưng được làm phẳng (Flatten) và đưa vào một lớp kết nối đầy đủ (Fully Connected - FC). Lớp này chiếu vector đặc trưng vào không gian 26 chiều tương ứng với 26 loại điều chế, và hàm Softmax cuối cùng sẽ tính toán xác suất dự đoán cho từng lớp.

2.3.3. Thiết kế khối DFE (*Diverse Feature Extraction Block*)

Khối DFE là đóng góp quan trọng nhất tạo nên sự khác biệt của mô hình. Thay vì sử dụng một kích thước bộ lọc cố định như các mạng CNN truyền thống (ví dụ: VGG chỉ dùng 3×3), khối DFE triển khai một kiến trúc đa nhánh song song lấy cảm hứng từ mạng Inception, nhưng được tối ưu hóa cho tín hiệu 2D.

Cấu trúc chi tiết: Khối DFE bao gồm 6 đơn vị tích chập được sắp xếp thành hai nhánh lớn hoạt động song song trên cùng một đầu vào Y như được minh họa trong Hình 2.4b):

Nhánh 1 (Xử lý chi tiết cục bộ): Đầu vào được đưa qua hai luồng con song song sử dụng bộ lọc kích thước 1×1 và 3×3 . Các lớp tích chập ở hai nhánh

con song song có kích thước bộ lọc lần lượt là 1×1 và 3×3 . Cả hai lớp đều được theo sau bởi các lớp kích hoạt ReLU.

Bộ lọc 1×1 : Đóng vai trò như một phép biến đổi tuyến tính tại mỗi điểm mẫu, giúp trộn thông tin giữa các kênh mà không thay đổi trường quan sát không gian. Trong ngữ cảnh tín hiệu IQ, nó có thể học các tổ hợp phi tuyến của biên độ I và Q.

Bộ lọc 3×3 : Nắm bắt các biến đổi cục bộ rất ngắn hạn, tương ứng với các thay đổi pha hoặc biên độ tức thời giữa các mẫu liên kề, đây là các đặc trưng tần số cao. Hai nhánh con này sau đó được kết hợp thông qua một lớp cộng (addition layer):

$$Y_1 = \text{ReLU}(\text{Conv}_{3 \times 3}(Y)) + \text{ReLU}(\text{Conv}_{1 \times 1}(Y)) \quad (2.21)$$

Sau đó, một lớp tích chập với bộ lọc kích thước 5×5 được sử dụng để tinh chỉnh đầu ra, mở rộng nhẹ trường quan sát để tổng hợp các đặc trưng cục bộ đã trích xuất:

$$Y_{1,1} = \text{ReLU}(\text{Conv}_{5 \times 5}(Y_1)) \quad (2.22)$$

Nhánh 2 (Xử lý ngữ cảnh rộng): Nhánh này được thiết kế sâu hơn để nắm bắt các phụ thuộc dài hạn. Đầu vào đi qua các luồng 1×1 và 5×5 , sau đó hợp nhất và đi qua bộ lọc lớn nhất kích thước 7×7 . Bộ lọc 7×7 cung cấp trường thụ cảm rộng, cho phép mạng nhận diện các mẫu tín hiệu kéo dài hoặc các cấu trúc lặp lại của sóng mang, đóng vai trò như một bộ lọc thông thấp (low-pass filter) giúp làm mịn nhiều tần số cao. Việc sử dụng các bộ lọc lớn hơn (5×5 và 7×7) cho phép mạng phân tích các phụ thuộc thời gian rộng hơn. Trong miền tần số, các hạt nhân lớn này hoạt động tương tự như các bộ lọc thông thấp (low-pass filters), giúp trích xuất các đặc trưng biến

đổi chậm và làm mịn tín hiệu. Khả năng này cực kỳ giá trị để nhận dạng các mẫu cấu trúc trong các phương thức điều chế bậc cao (nơi một symbol kéo dài qua nhiều mẫu) và nắm bắt động lực học tín hiệu dài hạn, đặc biệt trong các kịch bản khó khăn với kênh fading hoặc nhiễu mạnh làm biến dạng các chi tiết cục bộ.

Biểu thức toán học cho nhánh 2 :

$$Y_{2,inter} = \text{ReLU}(\text{Conv}_{5 \times 5}(Y)) + \text{ReLU}(\text{Conv}_{1 \times 1}(Y)) \quad (2.23)$$

$$Y_2 = \text{ReLU}(\text{Conv}_{7 \times 7}(Y_{2,inter})) \quad (2.24)$$

Cơ chế hợp nhất: Cơ chế cộng phần tử này cho phép tích hợp liền mạch các đặc trưng đa quy mô mà không làm tăng độ sâu của mạng (theo nghĩa tuần tự) hoặc ảnh hưởng đáng kể đến tốc độ suy luận. Đặc trưng từ hai nhánh lớn được cộng gộp để tạo ra đầu ra tổng hợp Y_{DFE} :

$$Y_{DFE} = Y_1 + Y_2 \quad (2.25)$$

Thiết kế đa tỷ lệ này cho phép DFENet đồng thời đóng vai trò như các bộ lọc thông cao (qua kernel nhỏ) để bắt các biến đổi pha nhanh, và bộ lọc thông thấp (qua kernel lớn) để bắt bao hình biên độ, từ đó nâng cao độ chính xác phân loại tổng thể.

Về bản chất vật lý, phép tích chập tương đương với một bộ lọc tuyến tính, trong đó các kernel với kích thước khác nhau đóng vai trò như các bộ lọc tại các dải tần khác nhau. Trong khi nhiễu có xu hướng phân bố rộng và biến thiên nhanh, tín hiệu lại có tính cấu trúc và tương quan theo thời gian. Do đó, các kernel lớn có xu hướng làm mượt tín hiệu và giảm nhiễu, trong khi các kernel nhỏ khai thác các biến thiên chi tiết. Cơ chế đa nhánh cho phép

mô hình đồng thời khai thác các đặc trưng ở nhiều thang, tăng khả năng tách tín hiệu khỏi nhiễu.

2.3.4. Thiết kế khối FEN (*Feature Extraction and Normalization Block*)

Mỗi khối DFE được theo sau bởi một khối FEN (được minh họa trong Hình 2.4c). Nếu DFE tập trung vào việc "mở rộng" không gian đặc trưng, khối FEN đóng vai trò bổ trợ quan trọng, tập trung vào việc "làm sạch" và nén gọn thông tin sau khi đã được trích xuất đa dạng bởi DFE.

Cấu trúc: Khối FEN tiếp nhận đầu ra từ DFE và xử lý qua hai nhánh song song với các bộ lọc kích thước nhỏ 1×1 và 3×3 . Điểm đặc biệt trong thiết kế của FEN là việc sử dụng lớp Batch Normalization (BN) ngay sau tích chập thay vì hàm kích hoạt, nhằm ổn định phân phối đặc trưng. Quá trình tính toán trong khối FEN diễn ra như sau:

Nhánh 1:

$$F_1 = \text{BN}(Y_{DFE} * W_{1 \times 1}) \quad (2.26)$$

Nhánh 2:

$$F_2 = \text{BN}(Y_{DFE} * W_{3 \times 3}) \quad (2.27)$$

Đầu ra khối:

$$F_{FEN} = F_1 + F_2 \quad (2.28)$$

Trong đó, lớp BN thực hiện chuẩn hóa dữ liệu về phân phối chuẩn tắc:

$$Y_{BN} = \gamma \frac{Y - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} + \beta \quad (2.29)$$

Trong đó: μ_B và σ_B^2 lần lượt là giá trị trung bình và phương sai của dữ liệu hiện tại.

γ và β là các tham số có thể học được, cho phép mạng khôi phục lại khả năng biểu diễn của dữ liệu nếu việc chuẩn hóa quá cứng nhắc.

ϵ là một giá trị nhỏ để tránh lỗi chia cho số không.

Vai trò của khối FEN gồm:

Giảm chiều dữ liệu : Các lớp tích chập trong FEN được cấu hình với stride (bước nhảy) là 2×2 . Điều này giúp giảm một nửa kích thước chiều dài của bản đồ đặc trưng sau mỗi khối (từ $1024 \rightarrow 512 \rightarrow 256 \rightarrow 128 \rightarrow 64$), qua đó giảm đáng kể khối lượng tính toán cho các tầng sau mà không cần dùng đến lớp Pooling (vốn có thể làm mất mát thông tin thời gian quan trọng).

Tinh chỉnh Đặc trưng: Các bộ lọc 1×1 và 3×3 giúp loại bỏ các nhiễu tần số cao không cần thiết và tập trung vào các đặc trưng quan trọng nhất đã được phát hiện bởi khối DFE.

Chống nhiễu và ổn định: Bằng cách áp dụng BN, khối FEN đưa phân phối của các đặc trưng về dạng chuẩn (trung bình 0, phương sai 1) trước khi truyền sang khối tiếp theo. Điều này cực kỳ quan trọng trong AMC vì nó giúp mạng trở nên bất biến hoặc ít nhạy cảm hơn với các thay đổi về biên độ tín hiệu do fading. Kết quả thực nghiệm trong cho thấy việc loại bỏ khối FEN làm giảm độ chính xác tới 3.5% ở mức SNR thấp (≤ 0 dB). Điều này chứng minh cơ chế chuẩn hóa BN trong FEN giúp mạng chống lại sự biến động mạnh của nhiễu nền và khắc phục hiện tượng biến mất đạo hàm (vanishing gradient).

Về bản chất vật lý khi xét tín hiệu đầu vào:

$$X = S + N \quad (2.30)$$

trong đó S là tín hiệu hữu ích và N là nhiễu ngẫu nhiên. FEN thực hiện ánh xạ phi tuyến $Z = f(X)$ thông qua các lớp tích chập sâu.

Quá trình này tương đương với một hệ lọc thích nghi, trong đó các thành phần có cấu trúc (thuộc tín hiệu S) được tăng cường nhờ tính tương quan theo thời gian – tần số, trong khi các thành phần nhiễu (không có cấu trúc) bị suy giảm. Kết quả là đặc trưng Z trở thành một biểu diễn “sạch hơn” của tín hiệu, làm giảm độ bất định trong phân loại và tăng khả năng phân biệt giữa các dạng điều chế, đặc biệt trong điều kiện SNR thấp.

2.3.5. So sánh với các mô hình hiện có

Để đánh giá vị thế của DFENet trong bức tranh tổng thể của lĩnh vực AMC, luận án cung cấp một sự so sánh chi tiết với các mô hình Deep Learning kinh điển khác như Inception, ResNet và MCNet thể hiện trong Bảng 2.2. Mặc dù DFENet chia sẻ nguyên lý chung về tích chập đa nhánh, thiết kế của nó chứa đựng những tùy chỉnh đặc thù nhằm giúp tối ưu hóa cho tín hiệu vô tuyến thay vì hình ảnh.

Bảng 2.2: So sánh mô hình mạng DFE với các mô hình đa nhánh và đa quy mô

Mô hình	Tích chập đa tỷ lệ	kết nối dư	Chuẩn hóa đặc trưng
Inception	Có (1×1 , 3×3 , 5×5)	Không	BN
ResNet	Không (Chỉ 3×3)	Có	BN+ReLU
MCNet	Có (đa quy mô)	Có	Không FEN
DFENet	Có (1×1 , 3×3 , 5×5 , 7×7)	Không	Khởi FEN (Chỉ BN)

Trái với Inception, vốn được phát triển ban đầu cho các bài toán thị giác máy tính và phụ thuộc nhiều vào các phép pooling có thể làm mất các quan hệ phụ thuộc theo thời gian quan trọng, DFENet loại bỏ hoàn toàn pooling nhằm bảo toàn cấu trúc thời gian chi tiết của tín hiệu IQ. Khác với ResNet, DFENet không dựa trên việc xếp chồng các khối dư sâu, từ đó tạo ra một kiến trúc gọn nhẹ hơn nhưng vẫn đạt hiệu quả cao, đồng thời khai thác các kernel đa tỷ lệ để nắm bắt cả sự chuyển tiếp ký hiệu ngắn hạn lẫn các phụ

thuộc dài hạn của tín hiệu.

So với MCNet, mặc dù cũng sử dụng tích chập đa tỷ lệ nhưng thiếu cơ chế chuẩn hóa đặc trưng rõ ràng, DFENet tích hợp đồng thời khối trích xuất đặc trưng đa dạng (khối DFE) và khối chuẩn hóa đặc trưng chuyên biệt (khối FEN). Thiết kế này giúp ổn định quá trình huấn luyện và tăng cường khả năng phân tách đặc trưng trong các điều kiện nhiễu mạnh và SNR thấp. Những cải tiến về kiến trúc này định vị DFENet như một mạng nhẹ, tối ưu theo miền tín hiệu vô tuyến, và hiệu quả cao, phù hợp cho bài toán phân loại điều chế tự động trong các môi trường truyền dẫn vô tuyến đầy thách thức.

Trong Bảng 2.3, mô hình 4 khối DFENet có khoảng 2 triệu tham số, được thiết kế nhằm đạt được sự cân bằng hợp lý giữa năng lực biểu diễn và chi phí tính toán. Việc lựa chọn bốn khối DFE không phải là ngẫu nhiên mà dựa trên sự đánh đổi có chủ đích giữa độ sâu mạng và hiệu quả suy luận.

Cụ thể, khi giảm số khối DFE ($\leq 2-3$ khối), mạng trở nên quá nông, làm hạn chế khả năng học các đặc trưng phân cấp và đa tỷ lệ cần thiết cho bài toán phân loại điều chế, đặc biệt trong điều kiện SNR thấp, nơi nhiễu làm suy giảm mạnh cấu trúc tín hiệu. Các đặc trưng trích xuất không đủ phong phú để phân biệt các điều chế bậc cao có chòm sao dày.

Ngược lại, việc tăng số khối DFE (> 4 khối) dẫn đến số tham số và độ phức tạp tính toán tăng nhanh, trong khi mức cải thiện độ chính xác không tương xứng, nhất là ở vùng SNR trung bình và cao, nơi hiệu năng của hầu hết các mô hình đã có xu hướng bão hòa. Điều này không chỉ làm giảm hiệu quả triển khai thực tế mà còn làm tăng nguy cơ overfitting trên các tập dữ liệu mô phỏng.

Nhờ sử dụng các khối DFE với tích chập song song đa kích thước (1×1 ,

3×3 , 5×5 , 7×7), 4 khối DFENet có thể khai thác hiệu quả cả đặc trưng cục bộ lẫn đặc trưng toàn cục mà không cần tăng độ sâu mạng quá mức. Do đó, cấu hình hiện tại đại diện cho một điểm vận hành tối ưu, nơi mô hình vẫn duy trì khả năng trích xuất đặc trưng đa dạng và phân cấp sâu, đồng thời giữ được kích thước gọn nhẹ và khả năng triển khai thực tế cho bài toán phân loại điều chế tự động.

Bảng 2.3: Cấu trúc mô hình 4 khối DFENet.

Lớp	Kích thước bộ lọc (K)	Bước trượt (S)	Đầu ra
Đầu vào	-	-	$1024 \times 2 \times 1$
Conv2D / ReLu	(3×2)	1×1	$1024 \times 2 \times 32$
Conv2D / BN/ ReLu	(3×2)	1×1	$1024 \times 2 \times 64$
Khối DFE	$\{1, 3, 5, 7\}$	1×1	$1024 \times 2 \times 64$
Khối FEN	$\{1, 3, \}$	2×2	$512 \times 1 \times 64$
Khối DFE	$\{1, 3, 5, 7\}$	1×1	$512 \times 2 \times 64$
Khối FEN	$\{1, 3, \}$	2×2	$256 \times 1 \times 64$
Khối DFE	$\{1, 3, 5, 7\}$	1×1	$256 \times 2 \times 64$
Khối FEN	$\{1, 3, \}$	2×2	$128 \times 1 \times 64$
Khối DFE	$\{1, 3, 5, 7\}$	1×1	$128 \times 2 \times 64$
Khối FEN	$\{1, 3, \}$	2×2	$64 \times 1 \times 64$
Lớp FC	-	-	$1 \times 1 \times 26$
Lớp Softmax	-	-	1×26

2.3.6. Mô tả tập dữ liệu và thiết lập thử nghiệm

Trong nghiên cứu này, hai bộ dữ liệu HisarMod2019.1 [33] và RadioML 2018.01A [11] được sử dụng để đánh giá hiệu năng mô hình. Hai bộ dữ liệu HisarMod2019.1 và RadioML 2018.01A được lựa chọn do đây là các bộ dữ liệu chuẩn, được sử dụng rộng rãi trong các nghiên cứu về nhận dạng điều chế, giúp đảm bảo tính khách quan và khả năng so sánh trực tiếp với các phương pháp tham chiếu. Bên cạnh đó, hai bộ dữ liệu này mô phỏng đa dạng các điều kiện kênh vô tuyến như nhiễu Gauss, fading đa đường (Rayleigh, Rician, Nakagami-m) và các hiệu ứng thực tế như dịch Doppler, sai lệch tần

số, từ đó cho phép đánh giá toàn diện khả năng hoạt động của mô hình trong các môi trường truyền dẫn khác nhau.

Bộ dữ liệu RadioML 2018.01A [11] được biết đến rộng rãi như một trong những bộ dữ liệu chuẩn phổ biến nhất trong lĩnh vực phân loại điều chế. Các tín hiệu trong bộ dữ liệu này chịu tác động của kênh đa đường Rician độc lập, với các đặc trưng khác nhau về độ trễ đường truyền, dịch Doppler, sai lệch tần số, trôi thời gian lấy mẫu do sai lệch xung nhịp, cùng với nhiễu Gauss trắng cộng tính, như đã mô tả chi tiết trong [11]. Bộ dữ liệu được chia theo tỷ lệ 70% cho huấn luyện, 15% cho xác thực, và 15% cho kiểm thử.

Đối với bộ dữ liệu HisarMod2019.1, cả mô hình đề xuất và các mô hình so sánh đều được huấn luyện trong 20 epoch, với các tham số huấn luyện gồm kích thước mini-batch là 64, tốc độ học ban đầu 0.001, và được giảm 10 lần sau mỗi 5 epoch. Đối với bộ dữ liệu RadioML2018.01A, bộ tối ưu ADAM được sử dụng, số epoch tối đa được thiết lập là 90, tốc độ học ban đầu là 0.001, và kích thước mini-batch được đặt ở mức 256.

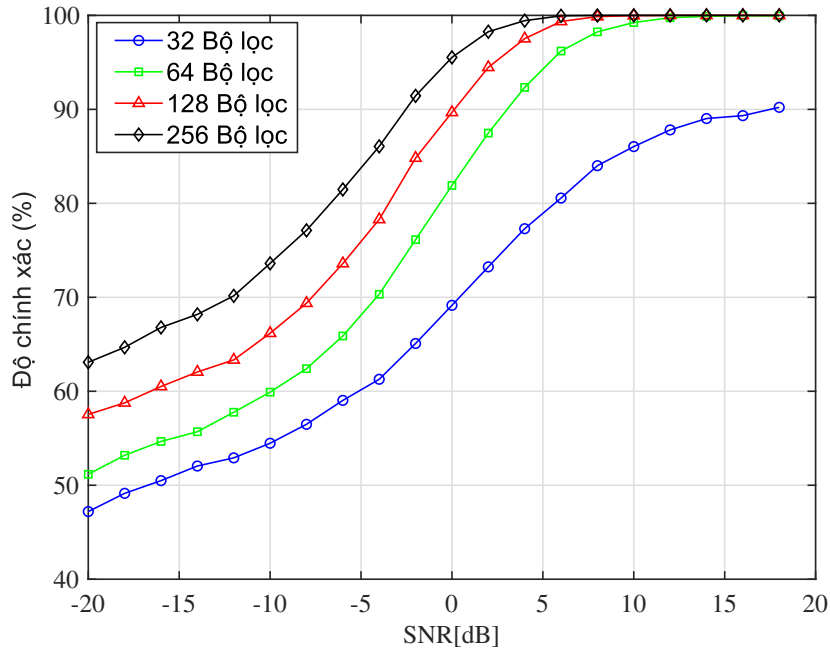
Thí nghiệm được thực hiện trên một hệ thống máy tính trang bị CPU 3.70 GHz, 32 GB RAM (2×16 GB) và GPU NVIDIA GeForce RTX 3060 Ti.

Trong các phần tiếp theo, hiệu năng của mô hình DFE được đánh giá trên hai bộ dữ liệu HisarMod2019 và RadioML2018, đồng thời được so sánh với các kiến trúc mạng hiện đại, bao gồm ResNet, SqueezeNet, MobileNet, GoogleNet, RepVGG, MCNet, CLDNN và LSTM.

2.3.7. Kết quả thực nghiệm và thảo luận

Đánh giá hiệu năng của DFENet khi thay đổi số lượng bộ lọc

Trong thí nghiệm đầu tiên, ảnh hưởng của số lượng bộ lọc trong các lớp tích chập đến hiệu năng phân loại điều chế của mô hình DFENet được khảo sát trên bộ dữ liệu HisarMod2019. Việc đánh giá được thực hiện trên nhiều mức SNR khác nhau, trong đó số lượng bộ lọc tại mỗi lớp tích chập lần lượt được thiết lập là 32, 64, 128 và 256, như thể hiện trong Hình 2.5.

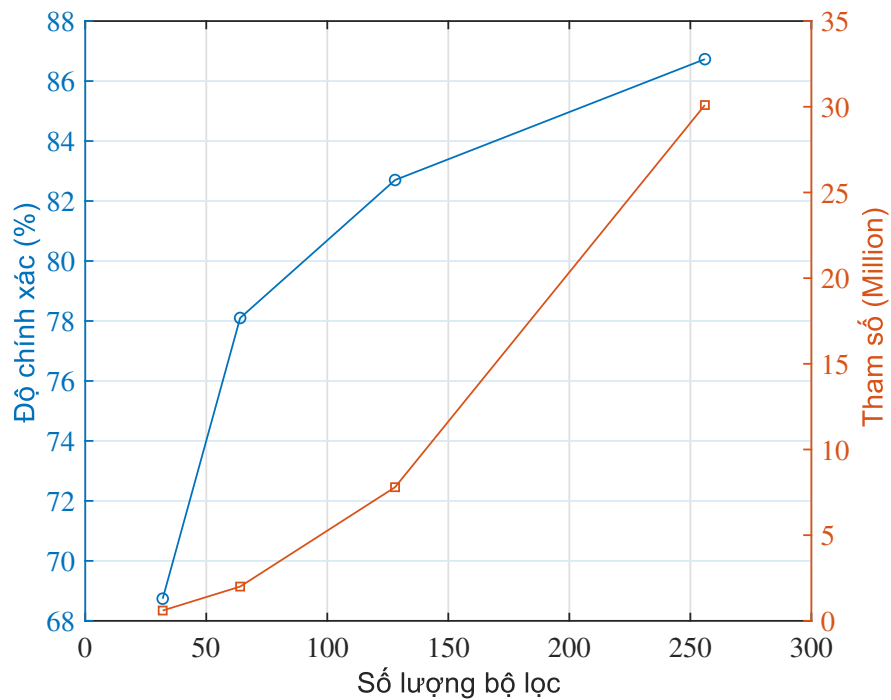


Hình 2.5: Độ chính xác của DFENet với số lượng bộ lọc khác nhau.

Kết quả thực nghiệm cho thấy độ chính xác phân loại tăng một cách nhất quán khi số lượng bộ lọc tăng, phản ánh khả năng trích xuất đặc trưng phân biệt tốt hơn của mạng khi dung lượng biểu diễn được mở rộng. Cụ thể, việc tăng số lượng bộ lọc mang lại mức cải thiện độ chính xác trung bình khoảng 6%, với lợi ích rõ rệt hơn trong các điều kiện SNR thấp.

Tuy nhiên, sự cải thiện này đi kèm với sự gia tăng đáng kể về độ phức tạp

của mô hình. Khi số lượng bộ lọc được tăng lên 256, mức cải thiện độ chính xác bắt đầu bão hòa, với mức tăng chỉ khoảng 4% so với cấu hình 128 bộ lọc. Hiện tượng này cho thấy việc bổ sung thêm bộ lọc chỉ đóng góp hạn chế vào khả năng học đặc trưng khi mạng đã tiến gần đến giới hạn hiệu quả trích xuất đặc trưng.

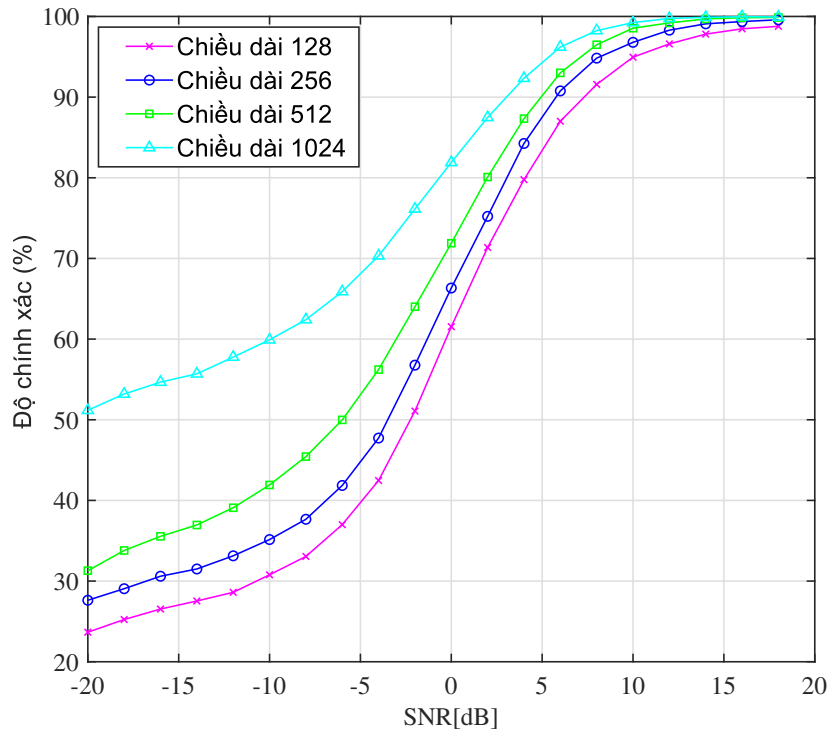


Hình 2.6: Ảnh hưởng của số lượng bộ lọc đến độ chính xác và độ phức tạp mô hình.

Bên cạnh đó, Hình 2.6 xem xét sự ảnh hưởng của số lượng bộ lọc đến độ chính xác và độ phức tạp mô hình. Kết quả cho thấy độ chính xác tăng theo số lượng bộ lọc nhưng với hiệu quả biên giảm dần, trong khi số lượng tham số tăng rất nhanh. Cụ thể, giai đoạn từ 32 lên 64 mang lại cải thiện đáng kể, nhưng từ 128 lên 256 mức tăng độ chính xác không còn tương xứng với chi phí tính toán gia tăng. Do đó, cấu hình số lượng bộ lọc $k = 64$ được lựa chọn như một điểm cân bằng hợp lý, khi đã khai thác phần lớn hiệu năng trong khi vẫn kiểm soát hiệu quả độ phức tạp của mô hình.

Đánh giá ảnh hưởng của độ dài tín hiệu với hiệu năng mô hình

Độ dài tín hiệu đầu vào có ảnh hưởng trực tiếp đến độ chính xác phân loại điều chế (AMC) cũng như độ phức tạp tính toán của mô hình. Do đó, trong thí nghiệm thứ hai, ảnh hưởng của độ dài tín hiệu khác nhau đến hiệu năng phân loại của mô hình DFENet được khảo sát. Cụ thể, mô hình được đánh giá tại nhiều mức SNR, với độ dài tín hiệu lần lượt được thiết lập là 128, 256, 512 và 1024 mẫu, như minh họa trong Hình 2.7.



Hình 2.7: Độ chính xác phân loại điều chế với chiều dài tín hiệu khác nhau.

Kết quả thực nghiệm cho thấy độ dài tín hiệu đóng vai trò then chốt trong hiệu năng phân loại, trong đó các tín hiệu có độ dài lớn hơn luôn mang lại độ chính xác AMC cao hơn. Các tín hiệu dài cung cấp thông tin phong phú và mang tính đại diện cao hơn về đặc trưng điều chế, giúp mô hình phân biệt các sơ đồ điều chế một cách hiệu quả hơn. Ngược lại, các tín hiệu ngắn,

chẳng hạn với độ dài 128 mẫu, hạn chế khả năng trích xuất đặc trưng, dẫn đến suy giảm đáng kể độ chính xác phân loại.

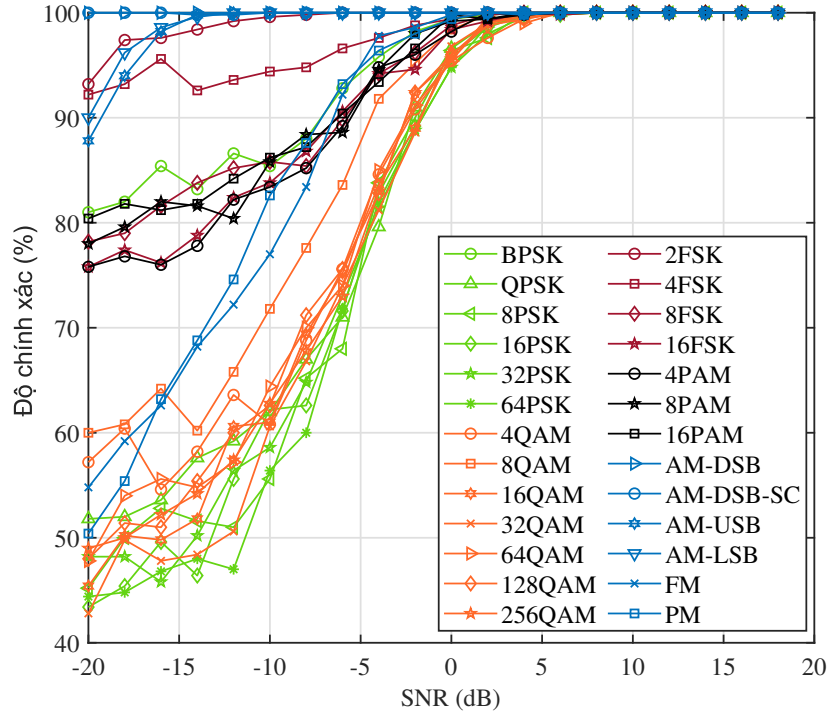
Đặc biệt, tín hiệu có độ dài 1024 mẫu cho phép mô hình khai thác các mẫu hình theo thời gian chi tiết hơn, từ đó cải thiện khả năng nhận dạng các dạng điều chế phức tạp, ngay cả trong điều kiện nhiễu cao. Tuy nhiên, lợi ích về hiệu năng này đi kèm với chi phí tính toán và thời gian xử lý tăng lên, do khối lượng dữ liệu đầu vào lớn hơn.

Kết quả trong Hình 2.5 và Hình 2.7 cho thấy mối quan hệ đánh đổi rõ ràng giữa độ chính xác, độ phức tạp tính toán và số lượng bộ lọc, chiều dài tín hiệu đầu vào. Việc lựa chọn cấu hình trong luận án (chiều dài tín hiệu 1024 và số bộ lọc 64) được xem là điểm cân bằng hợp lý giữa độ chính xác và chi phí tính toán. Do đó, cấu hình này được sử dụng cố định trong toàn bộ các thí nghiệm tiếp theo và không được điều chỉnh thêm.

Đánh giá hiệu năng mô hình đối với các dạng điều chế khác nhau

Trong thí nghiệm thứ ba, độ chính xác phân loại điều chế (AMC) của mô hình DFENet được đánh giá trên 26 sơ đồ điều chế khác nhau. Kết quả thể hiện trong Hình 2.8 cho thấy các sơ đồ điều chế có tốc độ thông tin thấp, chẳng hạn như AM và FM, đạt độ chính xác phân loại cao ngay cả trong điều kiện SNR thấp. Cụ thể, tại mức -10 dB, AM đạt độ chính xác 100%, trong khi FM và PM đạt xấp xỉ 76%. Hiệu năng cao này chủ yếu xuất phát từ đặc tính cấu trúc đơn giản của các dạng điều chế này, giúp chúng duy trì tính phân biệt tốt hơn trước tác động của nhiễu.

Đối với các tín hiệu FSK và PAM, mô hình DFENet cũng thể hiện khả năng nhận dạng hiệu quả, đạt độ chính xác lần lượt là 87% và 83% tại mức



Hình 2.8: Độ chính xác của DFENet trên 26 loại điều chế khác nhau trong các điều kiện SNR khác nhau.

-10 dB. Ngược lại, độ chính xác AMC đối với các nhóm điều chế PSK và QAM có xu hướng giảm khi bậc điều chế tăng lên. Nguyên nhân là do các sơ đồ điều chế bậc cao nhạy cảm hơn với suy hao kênh và nhiễu, dẫn đến sự chồng lấn đáng kể giữa các đặc trưng thống kê. Cụ thể, tại -10 dB, độ chính xác của PSK và QAM giảm xuống còn 54% và 61%, tương ứng. Khi mức nhiễu giảm xuống 0 dB, các giá trị này được cải thiện rõ rệt, đạt 94% đối với PSK và 96% đối với QAM.

Những kết quả này cho thấy mô hình DFENet đặc biệt hiệu quả trong việc phân loại các dạng điều chế bậc thấp trong điều kiện kênh khắc nghiệt, đồng thời cũng phản ánh thách thức cố hữu trong việc nhận dạng chính xác các dạng điều chế bậc cao, vốn dễ bị ảnh hưởng bởi nhiễu và méo kênh.

Phân tích chi tiết ma trận nhầm lẫn cho bài toán phân loại 26 dạng điều

chế tại các mức $\text{SNR} = +0 \text{ dB}$ và $\text{SNR} = +10 \text{ dB}$, như minh họa trong Hình 2.9 và Hình 2.10, cung cấp những hiểu biết quan trọng về mẫu hình nhầm lẫn giữa các loại điều chế khác nhau.

Tại mức $+0 \text{ dB}$, các dạng điều chế như PAM, FSK và AM thể hiện tỷ lệ nhầm lẫn thấp và đạt độ chính xác cao. Như thể hiện trong ma trận nhầm lẫn (Hình 2.9), các lỗi phân loại đáng chú ý nhất chủ yếu xảy ra giữa các sơ đồ QAM bậc cao, đặc biệt là giữa 64QAM và 16QAM. Nguyên nhân của hiện tượng này là do trong điều kiện SNR thấp, các điểm chòm sao của 64QAM bị co cụm và chồng lấn một phần với vùng chòm sao của QAM bậc thấp, khiến các đặc trưng thống kê trở nên khó phân biệt. Sự nhầm lẫn này có thể dẫn đến suy giảm hiệu năng đáng kể trong các hệ thống thông tin thực tế, việc nhận dạng sai điều chế bậc cao thành điều chế bậc thấp sẽ ảnh hưởng trực tiếp đến quá trình giải điều chế, giải mã ký hiệu và ước lượng thông lượng.

Ngược lại, các nhóm điều chế QAM và PSK đạt độ chính xác thấp hơn so với các sơ đồ đơn giản, với các giá trị cụ thể lần lượt là 96% đối với 128QAM, 95.4% đối với QPSK và 96.6% đối với 64PSK tại mức SNR tương ứng.

Đáng chú ý, tại mức $+10 \text{ dB}$, ma trận nhầm lẫn cho thấy mô hình đạt độ chính xác xấp xỉ 100% đối với phần lớn các dạng điều chế trong điều kiện nhiễu thấp. Điều này cho thấy DFENet hoạt động rất hiệu quả khi chất lượng tín hiệu được cải thiện. Tuy nhiên, vẫn có thể quan sát thấy sự suy giảm nhẹ về độ chính xác đối với các dạng điều chế có độ phức tạp cao, chẳng hạn như QAM và PSK bậc cao, khi bậc điều chế tăng lên.

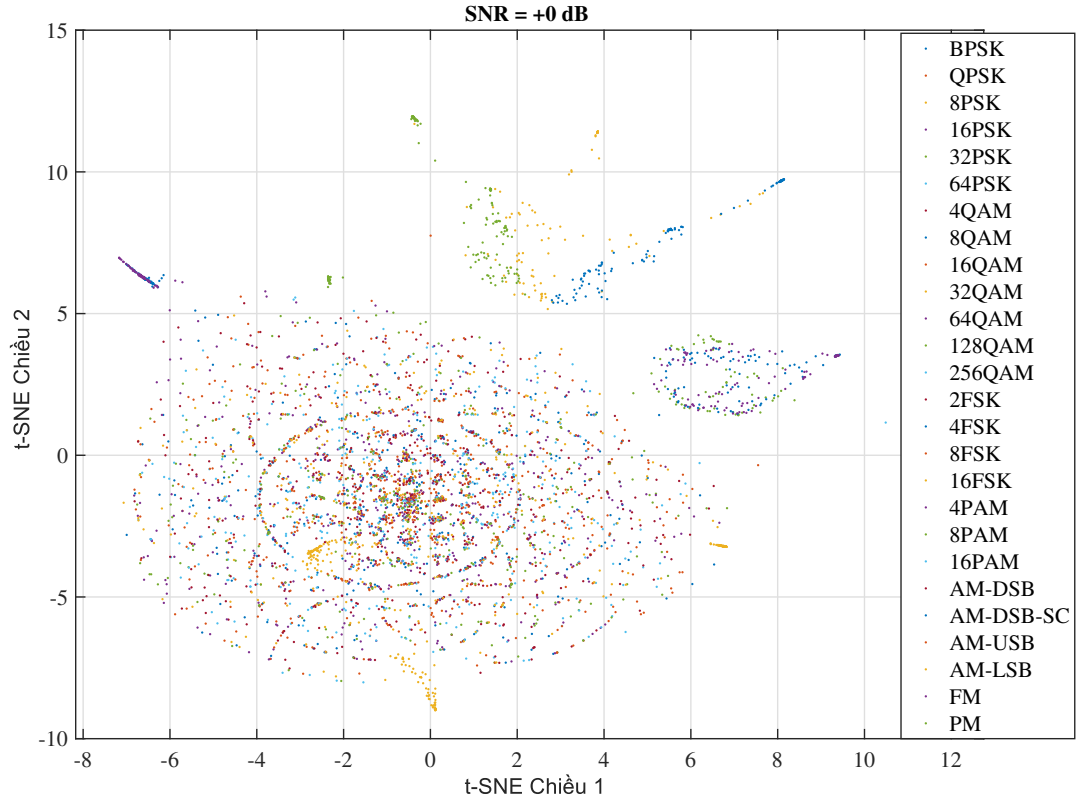
Trực quan hóa không gian đặc trưng bằng t-SNE

Để làm rõ khả năng phân tách đặc trưng cũng như năng lực biểu diễn của mô hình DFENet, nghiên cứu này áp dụng phương pháp t-distributed Stochastic Neighbor Embedding (t-SNE) lên các vector đặc trưng được trích xuất tại hai giai đoạn tiêu biểu của mạng: (i) ngay sau lớp tích chập đầu tiên, và (ii) tại đầu ra của lớp fully connected (FC) cuối cùng.

Hình 2.11 minh họa kết quả chiếu t-SNE của các đặc trưng thu được từ lớp tích chập đầu tiên tại mức SNR = +0 dB. Mặc dù đầy đủ 26 dạng điều chế đều xuất hiện trong không gian đặc trưng, các cụm điểm bị chồng lấn mạnh và không hình thành ranh giới phân tách rõ ràng. Điều này cho thấy các đặc trưng mức thấp chủ yếu phản ánh biến thiên tín hiệu tổng quát, thay vì mang tính phân biệt theo lớp điều chế.

Ngược lại, Hình 2.12 trình bày kết quả t-SNE tại lớp FC cuối cùng ở mức SNR = +0 dB. Mặc dù vẫn chịu ảnh hưởng đáng kể của nhiễu, mô hình đã hình thành được các cụm đặc trưng tương đối gọn đối với nhiều dạng điều chế như BPSK, QPSK, 2FSK và FM, cho thấy DFENet có khả năng học hiệu quả các cấu trúc tín hiệu mức thấp. Tuy nhiên, hiện tượng chồng lấn đặc trưng vẫn tồn tại rõ rệt giữa một số nhóm điều chế, cụ thể:

- Các dạng QAM bậc cao (64QAM, 128QAM, 256QAM) thể hiện sự trộn lẫn mạnh trong không gian đặc trưng. Ở mức SNR thấp và trung bình, cụm đặc trưng của 64QAM kém cô đặc và chồng lấn đáng kể với các cụm QAM lân cận. Nguyên nhân chính là do khoảng cách Euclid rất nhỏ giữa các điểm chòm sao trong QAM bậc cao, khiến nhiễu làm biến dạng biên độ và pha theo cách tương tự giữa các sơ đồ QAM khác nhau. Điều này



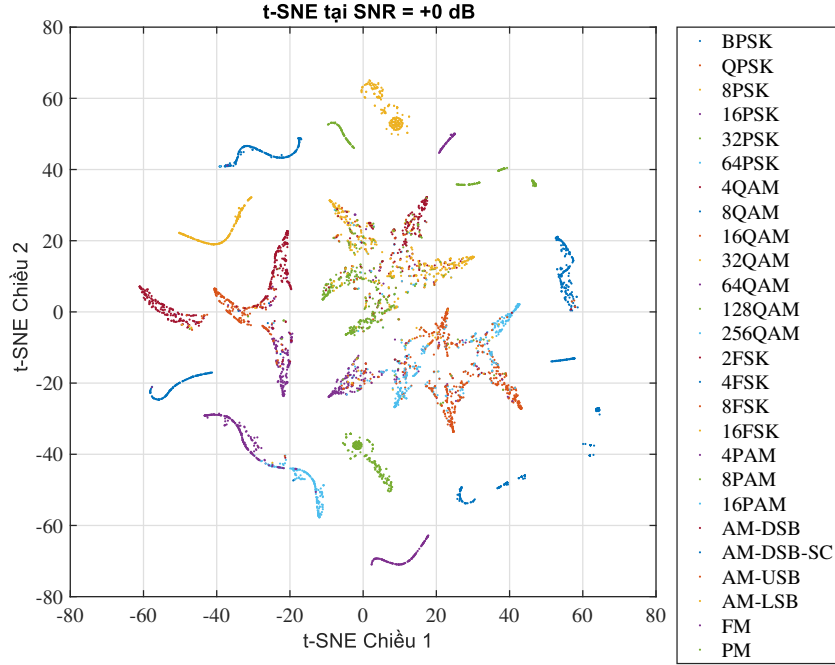
Hình 2.11: Đặc trưng tại đầu ra của lớp tích chập đầu tiên ở mức SNR = +0 dB.

cản trở việc hình thành biểu diễn tiềm ẩn tách biệt và giải thích cho tỷ lệ nhầm lẫn cao quan sát được trong ma trận nhầm lẫn, cũng như sự suy giảm thông lượng hiệu dụng đối với 64QAM.

- Các dạng PAM (8PAM, 16PAM) khó phân biệt do đặc trưng tín hiệu chủ yếu dựa trên mức biên độ tuyến tính.
- Các sơ đồ PSK bậc cao (16PSK, 32PSK) có độ tách góc giảm, dẫn đến sự giao thoa đặc trưng giữa các lớp.
- Các dạng điều chế AM (AM-DSB, AM-USB, AM-LSB) có phổ tín hiệu tương đồng, làm giảm khả năng phân tách.

Những kết quả này cho thấy, mặc dù DFENet đã bắt đầu trích xuất các

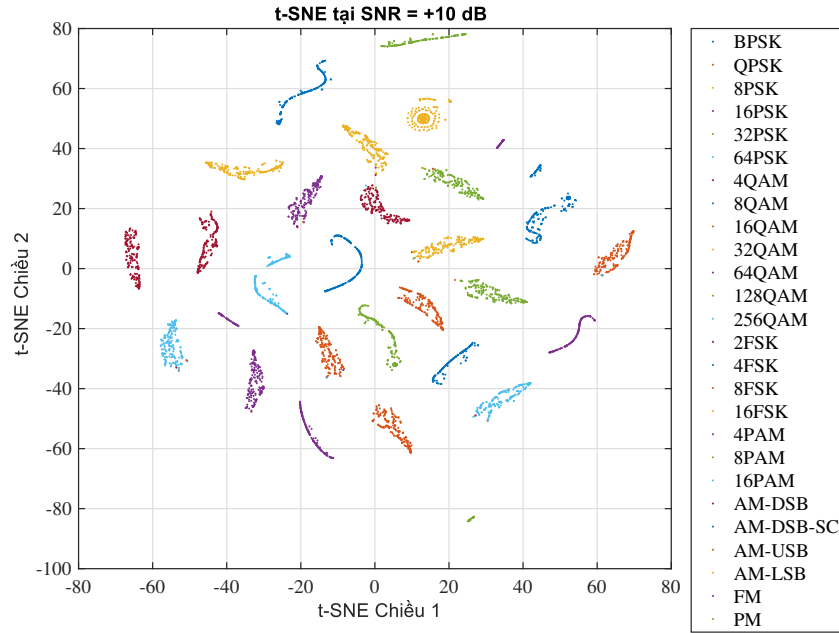
đặc trưng mang tính điều chế, mức độ phân tách hoàn toàn vẫn bị giới hạn trong điều kiện nhiễu cao.



Hình 2.12: Đặc trưng tại đầu ra của lớp FC ở mức SNR = +0 dB.

Hơn thế nữa, hình 2.13 thể hiện trực quan t-SNE tại lớp FC cuối cùng khi SNR = +10 dB. Trong trường hợp này, có thể quan sát thấy sự cải thiện rõ rệt về khả năng phân tách đặc trưng, với các cụm trở nên cô đặc và tách biệt hơn. Ngay cả các dạng điều chế từng gây nhầm lẫn ở mức SNR thấp, chẳng hạn như 64QAM và 16PAM, cũng được phân tách tương đối rõ ràng. Kết quả này khẳng định rằng kiến trúc học sâu phân cấp của DFENet có khả năng chuyển hóa các đặc trưng mức thấp còn mơ hồ thành các biểu diễn có tính phân biệt cao khi chất lượng tín hiệu được cải thiện.

Tổng thể, các biểu đồ t-SNE cung cấp bằng chứng định tính mạnh mẽ cho thấy DFENet học được các biểu diễn có cấu trúc và ý nghĩa, trong đó các đặc trưng tiến hóa từ chồng lấn và nhiễu sang tách biệt và rõ ràng khi lan



Hình 2.13: Đặc trưng tại đầu ra của lớp FC ở mức SNR = +10 dB.

truyền qua các tầng mạng và khi điều kiện kênh trở nên thuận lợi hơn.

Ảnh hưởng của cấu trúc mô hình

Để đánh giá mức độ đóng góp của từng thành phần kiến trúc trong DFENet, nghiên cứu sinh tiến hành thí nghiệm loại bỏ từng khối (ablation study) trên tập dữ liệu HisarMod2019. Mô hình DFENet được đánh giá dưới các cấu hình khác nhau, như tóm tắt trong Bảng 2.4.

Bảng 2.4: Nghiên cứu cắt bỏ trên bộ dữ liệu HisarMod2019

Mô hình	Khối		Bộ lọc đa quy mô	Độ chính xác trung bình (%)
	DFE	FEN		
DFENet- đầy đủ	✓	✓	✓	80.8
Không-FEN	✓	□	✓	70.31
Không-DFE	□	✓	✓	70.59
Không bộ lọc đa quy mô	✓	✓	3×3	71.25
ShallowNet	□	□	□	14.95

Bảng 2.4 trình bày kết quả ablation của DFENet, qua đó làm rõ tác động của từng thành phần kiến trúc đến độ chính xác phân loại điều chế (AMC).

Mô hình DFENet đầy đủ, kết hợp cả khối trích xuất đặc trưng đa dạng

(DFE) và khối tinh chỉnh đặc trưng (FEN) cùng với bộ lọc đa tỉ lệ, đạt độ chính xác trung bình cao nhất là 80.8%. Kết quả này cho thấy hiệu quả vượt trội của việc kết hợp trích xuất đặc trưng đa tỉ lệ với tăng cường và chuẩn hóa đặc trưng sâu trong bài toán phân loại điều chế vô tuyến.

Khi loại bỏ khối FEN (Không-FEN), độ chính xác giảm mạnh xuống 70.31%, khẳng định vai trò quan trọng của cơ chế tinh chỉnh dựa trên batch normalization trong việc ổn định quá trình huấn luyện và giảm nhiễu trong không gian đặc trưng. Tương tự, khi loại bỏ khối DFE (Không-DFE), độ chính xác chỉ còn 70.59%, cho thấy trích xuất đặc trưng đa nhánh là yếu tố then chốt để nắm bắt các quan hệ phụ thuộc theo thời gian trong tín hiệu IQ.

Ngoài ra, việc thay thế bộ lọc đa tỉ lệ bằng một kernel duy nhất 3×3 (Không bộ lọc đa quy mô) làm hiệu năng giảm xuống 71.25%, chứng tỏ rằng việc sử dụng đa kích thước kernel (từ 1×1 đến 7×7) giúp cải thiện đáng kể khả năng mô hình hóa ngữ cảnh thời gian ở nhiều thang đo khác nhau. Điều này xác nhận lợi ích của việc khai thác đặc trưng tín hiệu ở các độ phân giải khác nhau thay vì chỉ dựa vào một trường tiếp nhận cố định.

Cuối cùng, cấu hình ShallowNet, không sử dụng DFE, FEN cũng như bộ lọc đa tỉ lệ, cho hiệu năng thấp nhất với độ chính xác chỉ 14.95%. Kết quả này khẳng định rằng các mạng nông, thiếu cấu trúc xử lý chuyên biệt, là không đủ khả năng giải quyết bài toán phân loại điều chế trong môi trường kênh phức tạp. Đáng chú ý, DFENet cải thiện hiệu năng lên gần 66% so với ShallowNet, làm nổi bật giá trị vượt trội và khả năng trích xuất đặc trưng mạnh mẽ của kiến trúc đa thành phần được đề xuất.

Tổng thể, các kết quả ablation cho thấy Bảng 2.4 phản ánh rõ ràng đóng

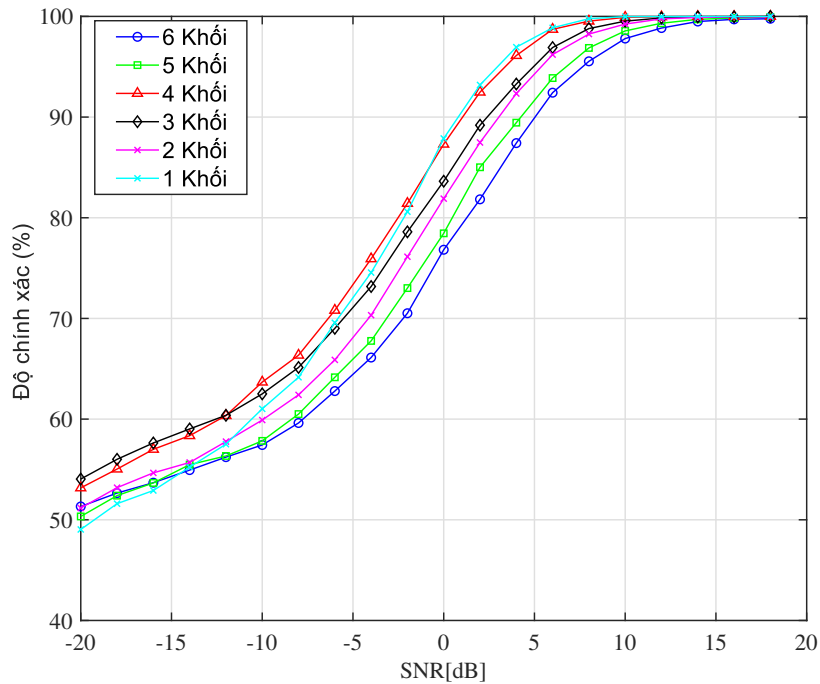
góp riêng lẻ cũng như tổng hợp của từng thành phần kiến trúc trong DFENet. Đặc biệt, sự kết hợp đồng thời giữa các khối DFE và FEN đóng vai trò then chốt trong việc đạt được cải thiện đáng kể về độ chính xác phân loại, nhất là trong điều kiện SNR thấp.

Hiệu năng mô hình với số lượng khối DFE khác nhau

Thí nghiệm này cung cấp những nhận định quan trọng về mặt thiết kế kiến trúc đối với vai trò của các khối DFE trong mô hình DFENet. Bằng cách giữ cố định số lượng bộ lọc tích chập ở mức 64 và chỉ thay đổi số lượng khối DFE, ảnh hưởng của độ sâu mạng đến hiệu năng phân loại điều chế được đánh giá một cách độc lập.

Như thể hiện trong Hình 2.14, việc tăng số lượng khối DFE từ một lên ba mang lại sự cải thiện độ chính xác rõ rệt trong miền SNR thấp (≤ -5 dB), với mức tăng trung bình khoảng 2% cho mỗi khối bổ sung. Kết quả này cho thấy quá trình trích xuất đặc trưng theo tầng sâu cho phép mạng từng bước tinh lọc các biểu diễn tín hiệu bị suy giảm bởi nhiễu và fading. Trong điều kiện SNR thấp, nơi các cấu trúc đặc trưng của điều chế bị che khuất nghiêm trọng, việc xếp chồng nhiều khối DFE giúp tích lũy các đặc trưng mang tính phân biệt thông qua các trường tiếp nhận đa tỉ lệ.

Ngược lại, tại các mức SNR cao, sự khác biệt về hiệu năng giữa các cấu hình có số lượng khối DFE khác nhau trở nên không đáng kể. Điều này cho thấy khi chất lượng tín hiệu đủ tốt, các đặc trưng điều chế đã được tách biệt hiệu quả ngay từ các tầng đầu, và việc gia tăng thêm chiều sâu kiến trúc chỉ mang lại lợi ích biên rất hạn chế. Từ góc độ thiết kế, kết quả này khẳng định rằng kiến trúc sâu chủ yếu phát huy tác dụng trong các kịch bản kênh khắc



Hình 2.14: Độ chính xác phân loại với số lượng khối DFE khác nhau.

nghiệt – vốn là bài toán trọng tâm của các hệ thống AMC thực tế.

Đáng chú ý, kết quả thí nghiệm cũng chỉ ra một ngưỡng bão hòa rõ ràng. Khi số lượng khối DFE vượt quá ba, độ chính xác phân loại bắt đầu suy giảm, với mức giảm khoảng 1% khi tăng từ bốn lên sáu khối. Hiện tượng này phản ánh sự đánh đổi giữa năng lực biểu diễn và hiệu quả mô hình. Các kiến trúc quá sâu có thể tạo ra các biến đổi đặc trưng dư thừa, làm khuếch đại nhiễu tích lũy hoặc dẫn đến hiện tượng quá khớp, đặc biệt khi số lượng tham số tăng nhanh nhưng dữ liệu huấn luyện không đủ để hỗ trợ.

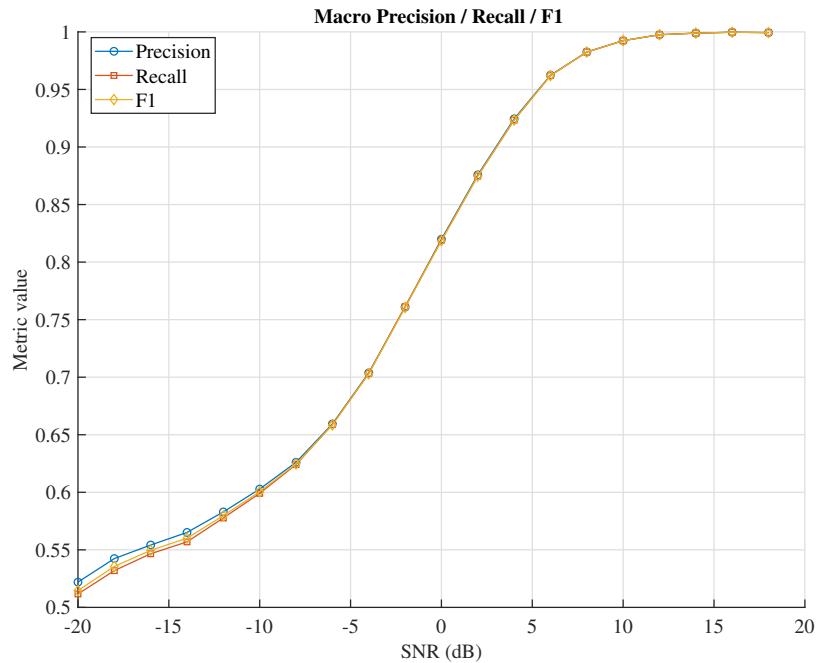
Các mô hình quá phức tạp với số lượng khối DFE dư thừa có nguy cơ quá khớp (overfitting) hoặc phát sinh chi phí tính toán lớn, từ đó làm suy giảm hiệu quả tổng thể. Do đó, kết quả thí nghiệm này cho thấy rằng việc lựa chọn số lượng khối DFE tối ưu là yếu tố then chốt để đạt được sự cân bằng giữa độ chính xác phân loại và hiệu quả tính toán, đặc biệt trong các kịch bản

AMC thực tế dưới điều kiện SNR thấp.

Những kết quả này cung cấp một nguyên tắc thiết kế rõ ràng cho AMC: việc gia tăng độ sâu mạng chỉ mang lại lợi ích trong miền SNR thấp và phải được kiểm soát chặt chẽ để tránh dư thừa kiến trúc và suy giảm hiệu năng. DFENet hiện thực hóa nguyên tắc này bằng một cấu hình cân bằng giữa khả năng trích xuất đặc trưng và hiệu quả mô hình.

Đánh giá hiệu năng bằng Precision, Recall và F1-score

Để bổ sung cho chỉ số độ chính xác tổng thể (Accuracy) và cung cấp một đánh giá toàn diện hơn về hiệu năng phân loại của DFENet, luận án sử dụng ba thước đo phổ biến gồm Precision, Recall và F1-score, như minh họa trong Hình 2.15.



Hình 2.15: Precision, Recall và F1-score của mô hình DFE.

Kết quả cho thấy rằng tại các mức SNR thấp, cả ba chỉ số đánh giá đều tăng dần khi SNR được cải thiện. Trong miền SNR rất thấp (từ -20 dB đến

0 dB), DFENet vẫn duy trì hiệu năng tương đối ổn định, với giá trị F1-score tăng từ khoảng 0.52 lên gần 0.80. Điều này chứng tỏ rằng ngay cả trong điều kiện nhiễu nghiêm trọng, mô hình vẫn có khả năng trích xuất các đặc trưng mang ý nghĩa từ tín hiệu đầu vào.

Trong khoảng SNR từ 0 dB đến +10 dB, các chỉ số đánh giá tăng nhanh và sau +10 dB, F1-score tiến gần 0.99, phản ánh khả năng phân loại gần như hoàn hảo của DFENet khi chất lượng tín hiệu được cải thiện.

Đáng chú ý, sự chênh lệch giữa Precision và Recall luôn ở mức rất nhỏ trên toàn bộ dải SNR, cho thấy mô hình duy trì hành vi dự đoán cân bằng, không thiên lệch về phía tăng sai dương tính hay sai âm tính. Điều này khẳng định rằng DFENet không chỉ đạt độ chính xác cao mà còn đảm bảo sự ổn định và công bằng giữa các lớp điều chế.

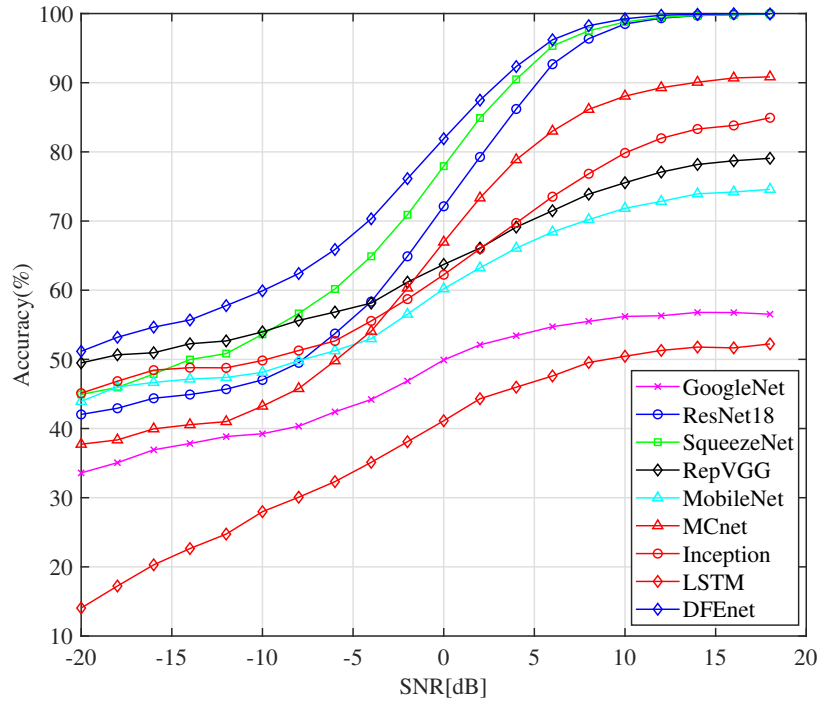
Tổng hợp lại, các kết quả trên cho thấy DFENet có khả năng tổng quát hóa mạnh mẽ trong nhiều điều kiện kênh khác nhau, đồng thời hiệu năng F1 cao xác nhận rằng mô hình đạt được cả độ nhạy và hiệu quả, đáp ứng tốt các yêu cầu khắt khe của bài toán phân loại điều chế trong môi trường vô tuyến phức tạp.

So sánh DFENet với các mô hình phổ biến hiện nay

Trong thí nghiệm cuối cùng, luận án đánh giá hiệu năng của DFENet trên hai bộ dữ liệu HisarMod2019 và RadioML2018 nhằm kiểm chứng khả năng hoạt động trong các điều kiện kênh đa dạng cũng như đánh giá khả năng tổng quát hóa của mô hình so với các kiến trúc học sâu tiêu biểu hiện nay.

Trước hết, độ chính xác phân loại AMC của DFENet được so sánh với các mô hình hiện có trên bộ dữ liệu HisarMod2019, như minh họa trong

Hình 2.16. Kết quả cho thấy mô hình LSTM đạt hiệu năng thấp nhất, với độ chính xác tối đa chỉ khoảng 52% tại mức SNR +18 dB. Nguyên nhân chủ yếu là do LSTM có khả năng trích xuất đặc trưng hạn chế khi xử lý trực tiếp tín hiệu IQ, đặc biệt trong môi trường nhiễu và fading phức tạp.



Hình 2.16: So sánh độ chính xác phân loại của các mô hình CNN đơn lẻ khác nhau đối với tập dữ liệu HisarMod2019.

Ngược lại, GoogLeNet thể hiện hiệu năng vượt trội hơn nhờ cấu trúc tích chập nhiều tầng, qua đó tăng cường khả năng học các đặc trưng phân cấp. Các mô hình như MobileNet, RepVGG và Inception cũng cho thấy sự cải thiện đáng kể, với mức tăng độ chính xác lần lượt khoảng 20%, 27% và 32% so với LSTM trong dải SNR từ -20 dB đến $+18$ dB. Điều này cho thấy vai trò quan trọng của các kiến trúc CNN sâu trong bài toán phân loại điều chế.

Bên cạnh đó, các mô hình SqueezeNet (sử dụng nhiều lớp tích chập với số lượng bộ lọc tăng dần), ResNet (dựa trên các khối liên kết dư) và MCNet

(kết hợp các khối tích chập song song đa tỉ lệ) đạt được độ chính xác cao hơn, đặc biệt trong miền SNR từ +2 dB đến +18 dB. Tuy nhiên, hiệu năng của các mô hình này vẫn suy giảm đáng kể trong điều kiện SNR thấp.

Đáng chú ý, với cấu hình sử dụng 64 bộ lọc, DFENet đạt độ chính xác trung bình 78,2% trên tập kiểm thử trong dải SNR được sử dụng để đánh giá. DFENet đạt hiệu năng cao nhất trong tất cả các mô hình được so sánh. Mô hình đặc biệt nổi bật trong miền SNR thấp, khi vẫn đạt độ chính xác trên 51.2% tại -20 dB và vượt 90% tại +2 dB. Kết quả này chứng minh khả năng thích nghi mạnh mẽ của DFENet trong các điều kiện kênh khắc nghiệt.

Hiệu năng vượt trội của DFENet xuất phát từ cơ chế trích xuất đặc trưng đa dạng và đa tỉ lệ thông qua các khối DFE, kết hợp với cấu trúc kiến trúc được thiết kế có kiểm soát. Việc sử dụng nhiều kích thước bộ lọc và các khối xử lý mở rộng giúp mô hình nắm bắt đồng thời các đặc trưng cục bộ và dài hạn của tín hiệu, từ đó nâng cao khả năng phân biệt giữa các dạng điều chế phức tạp. Điều này cho phép DFENet vượt trội hơn so với các mô hình CNN hiện có trong cả điều kiện SNR thấp và cao.

DFENet thể hiện hiệu năng đáng chú ý so với các mô hình khác, như được trình bày trong Bảng 2.5. DFENet cho thấy sự cân bằng nổi bật giữa hiệu năng phân loại và hiệu quả tính toán so với các mô hình hiện có. Nhờ kiến trúc gọn nhẹ và được thiết kế có kiểm soát, các biến thể của DFENet duy trì số lượng tham số ở mức thấp. Cụ thể, 1 khối DFENet chỉ sử dụng khoảng 1.3 triệu tham số, trong khi 4 khối DFENet có khoảng 2 triệu tham số, nhỏ hơn đáng kể so với các mô hình như RepVGG (24.8 triệu tham số) và Inception (21.8 triệu tham số). Mức độ gọn nhẹ này trực tiếp góp phần nâng cao hiệu quả tính toán và khả năng triển khai thực tế.

Về thời gian xử lý, DFENet tiếp tục thể hiện tính cạnh tranh cao. Chẳng hạn, 1 khối DFENet đạt thời gian suy luận 1.338 ms, tương đương hoặc nhanh hơn so với SqueezeNet (1.186 ms) và ResNet18 (1.218 ms). Ngay cả khi tăng số lượng khối DFE, như trong trường hợp 4 khối DFENet với thời gian 2.132 ms, mô hình vẫn duy trì hiệu năng tốt hơn hoặc tương đương các kiến trúc phổ biến khác như MobileNet (2.290 ms) và RepVGG (3.189 ms).

Một ưu điểm quan trọng khác của DFENet là tính mở rộng có kiểm soát. Kiến trúc dựa trên các khối DFE cho phép gia tăng số lượng tham số và chi phí tính toán một cách gần tuyến tính khi bổ sung thêm khối mới. Tuy nhiên, mức gia tăng này được kiểm soát hiệu quả, giúp mô hình mở rộng năng lực biểu diễn mà không gây ra sự bùng nổ về độ phức tạp.

Nhìn chung, việc sử dụng tham số hiệu quả cùng với tốc độ suy luận cạnh tranh cho thấy DFENet là một mô hình được thiết kế theo hướng hiệu năng cao. Mô hình không chỉ đạt hiệu năng phân loại cao mà còn đáp ứng tốt các ràng buộc về tài nguyên tính toán, qua đó trở thành một lựa chọn phù hợp cho các ứng dụng AMC yêu cầu độ chính xác cao, độ trễ thấp và khả năng triển khai trong môi trường thực tế.

Bảng 2.5: Độ phức tạp, thời gian tính toán, FLOPs của mô hình.

Mô hình	Tham số(m)	Thời gian (ms)	FLOPs (MFLOPs)
GoogleNet	5.9	1.361	103.89
ResNet18	11.1	1.218	41.14
SqueezeNet	2.4	1.186	13.10
RepVGG	24.8	3.189	154.29
MobileNet	2	2.290	24.86
MCnet	10.3	1.346	28.34
Inception	21.8	1.416	163.72
LSTM	1.1	0.9725	137.36
DFEnet-1 khối	1.3	1.338	123.78
DFEnet-2 khối	1.4	1.422	195.36
DFEnet-3 khối	1.7	1.485	266.94
DFEnet-4 khối	2	2.13	338.52

Như trình bày trong Bảng 2.5, độ phức tạp mô hình là một trong những ưu điểm nổi bật của DFENet. Cụ thể, biến thể 4 khối DFENet chỉ bao gồm khoảng 2.0 triệu tham số, giảm đáng kể so với các mô hình hiệu năng cao khác như RepVGG (24.8 triệu tham số) và Inception (21.8 triệu tham số). Kích thước mô hình nhỏ gọn này trực tiếp dẫn đến yêu cầu bộ nhớ thấp hơn, một lợi thế quan trọng đối với các thiết bị hạn chế tài nguyên, chẳng hạn như các nút IoT và hệ thống điện toán biên.

Bên cạnh yêu cầu bộ nhớ, tốc độ suy luận là một chỉ số then chốt đối với các ứng dụng AMC thời gian thực. Để bảo đảm tính khách quan trong so sánh, tất cả các mô hình đều được huấn luyện và đánh giá trên cùng bộ dữ liệu, cùng môi trường phần cứng và phần mềm. Thời gian suy luận của các mô hình được đo theo cùng quy trình mô tả tại Mục 2.3.6, bảo đảm các kết quả có thể được so sánh trực tiếp. Kết quả trong Bảng 2.5 cho thấy 4 khối DFENet đạt thời gian suy luận nhanh, chỉ 2.132 ms cho mỗi mẫu, thể hiện tính cạnh tranh cao so với các kiến trúc hiện có. Sự kết hợp giữa số lượng tham số thấp và tốc độ suy luận cao cho thấy DFENet đạt được sự cân bằng hiệu quả giữa độ chính xác phân loại và chi phí tính toán.

Từ Bảng 2.5 có thể thấy sự khác biệt đáng kể về số lượng FLOPs giữa các mô hình. Các kiến trúc nhẹ như SqueezeNet và MobileNet có FLOPs thấp, phù hợp cho các hệ thống hạn chế tài nguyên, trong khi các mô hình sâu như Inception và RepVGG có FLOPs cao hơn do cấu trúc phức tạp.

Đáng chú ý, các biến thể DFENet có số lượng tham số nhỏ nhưng FLOPs tăng nhanh theo số khối, cho thấy chi phí tính toán chủ yếu đến từ độ sâu và số phép tích chập thay vì số tham số. Điều này phản ánh sự đánh đổi giữa khả năng trích xuất đặc trưng và chi phí tính toán. Ngoài ra, FLOPs không

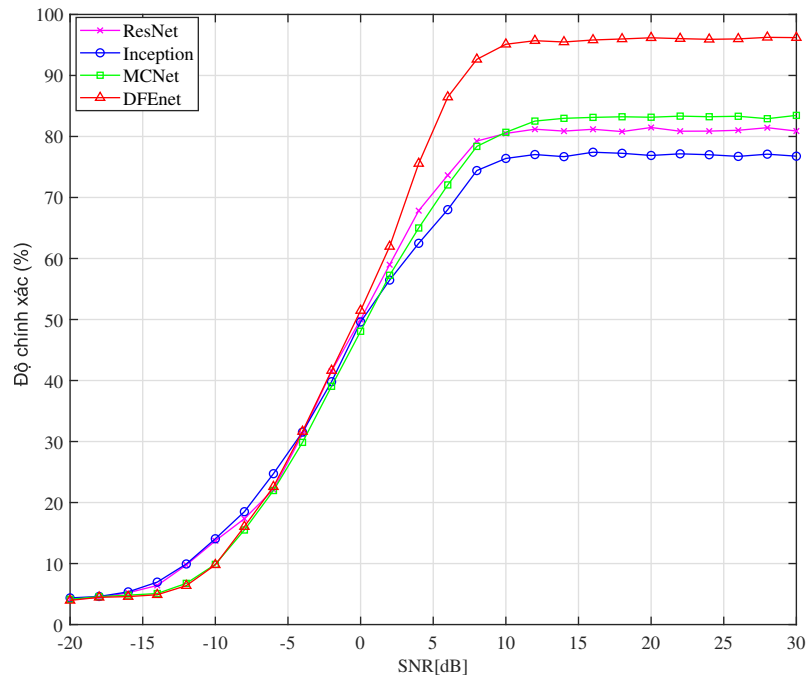
hoàn toàn tương quan với thời gian suy luận, cho thấy hiệu năng thực tế còn phụ thuộc vào cấu trúc mạng và khả năng tối ưu phần cứng.

Tổng thể, các phân tích trên xác nhận rằng DFENet không chỉ là một mô hình có hiệu năng cao trong phòng thí nghiệm mà còn là một giải pháp khả thi và mạnh mẽ cho các hệ thống AMC thực tế, đáp ứng tốt các yêu cầu về độ trễ thấp, tiết kiệm tài nguyên và khả năng triển khai linh hoạt trong môi trường vô tuyến hiện đại.

Luận án tiếp tục đánh giá mô hình DFENet với ba mô hình tiêu biểu khác, bao gồm MCNet, Inception và ResNet, trên bộ dữ liệu RadioML2018, như minh họa trong Hình 2.17. Kết quả cho thấy mô hình Inception đạt độ chính xác phân loại thấp nhất, với giá trị tối đa chỉ khoảng 73% tại $\text{SNR} = +18$ dB. Hiệu năng hạn chế này có thể xuất phát từ việc mô hình Inception, vốn được thiết kế cho bài toán thị giác máy tính, chưa tối ưu cho việc trích xuất các đặc trưng thời-tần của tín hiệu IQ.

Ngược lại, ResNet và MCNet đạt được độ chính xác cao hơn nhờ khai thác hiệu quả các liên kết dư, giúp ổn định quá trình huấn luyện và cải thiện khả năng học đặc trưng. Cụ thể, tại mức $\text{SNR} = +18$ dB, ResNet và MCNet lần lượt đạt độ chính xác khoảng 81% và 83%.

DFENet vượt trội hơn tất cả các mô hình được so sánh khi đạt độ chính xác trên 93% tại các mức $\text{SNR} \geq +6$ dB. Không chỉ vậy, DFENet còn thể hiện độ ổn định cao trong điều kiện SNR thấp, với độ chính xác AMC vượt 60% tại -20 dB và trên 90% tại -2 dB. Sự cải thiện đáng kể này cho thấy khả năng thích nghi vượt trội của DFENet trong môi trường nhiễu mạnh.



Hình 2.17: So sánh độ chính xác phân loại của các mô hình CNN đơn lẻ khác nhau đối với tập dữ liệu RadioM12018.

2.4. Kết luận chương 2

Chương 2 đã thực hiện một khảo sát toàn diện về vai trò của các phương pháp tiền xử lý và không gian biểu diễn tín hiệu trong bài toán phân loại điều chế dựa trên học sâu. Kết quả cho thấy biểu diễn miền thời gian-tần số (STFT) mang lại hiệu năng vượt trội so với biểu diễn phổ thuần túy (FFT), đặc biệt trong vùng SNR thấp. Việc kết hợp biến đổi Radon giúp tăng cường cấu trúc hình học của ảnh phổ và cải thiện thêm khả năng phân tách giữa các lớp điều chế. Tuy nhiên, so sánh tổng thể cho thấy tín hiệu IQ nguyên thủy vẫn đạt độ chính xác cao nhất và ổn định nhất trên toàn dải SNR.

Trên cơ sở các kết luận đó, luận án đã đề xuất mô hình DFENet, được thiết kế nhằm khai thác trực tiếp và hiệu quả các đặc trưng đa mức từ tín hiệu IQ nguyên thủy. DFENet sử dụng cấu trúc trích xuất đặc trưng đa nhánh nhằm

tăng cường tính đa dạng trong không gian biểu diễn, đồng thời duy trì khả năng học đặc trưng cục bộ và toàn cục của tín hiệu. Kết quả thực nghiệm cho thấy DFENet cải thiện đáng kể độ chính xác phân loại so với các kiến trúc CNN truyền thống trong cùng điều kiện huấn luyện.

Mặc dù mô hình DFENet đạt được kết quả phân loại cao trên các bộ dữ liệu chuẩn, mô hình vẫn chủ yếu tập trung vào việc khai thác đặc trưng không gian từ tín hiệu IQ thông qua các phép tích chập. Do đó, khả năng mô hình hóa các mối quan hệ phụ thuộc theo thời gian của tín hiệu còn hạn chế, đặc biệt đối với các tín hiệu có đặc tính động hoặc biến thiên phức tạp. Bên cạnh đó, hiệu năng phân loại vẫn suy giảm ở các mức SNR thấp và đối với một số dạng điều chế có đặc trưng tương đồng, cho thấy việc sử dụng một kiến trúc đơn lẻ chưa khai thác đầy đủ thông tin chứa trong tín hiệu vô tuyến.

Từ những hạn chế trên, hướng nghiên cứu tiếp theo tập trung phát triển các mô hình có khả năng khai thác đồng thời đặc trưng không gian và quan hệ phụ thuộc theo thời gian của tín hiệu, điển hình là các mô hình kết hợp CNN và BiLSTM. Đồng thời, các phương pháp học kết hợp sẽ được nghiên cứu nhằm tận dụng tính bổ sung giữa các mô hình khác nhau, qua đó nâng cao độ chính xác, tính ổn định theo SNR và khả năng tổng quát hóa của hệ thống phân loại điều chế tự động.

Chương 3

GIẢI PHÁP KẾT HỢP KIẾN TRÚC CÁC MẠNG NƠ-RON CHO NÂNG CAO HIỆU NĂNG PHÂN LOẠI ĐIỀU CHẾ TỰ ĐỘNG

Chương 2 đã chứng minh rằng việc bảo toàn tối đa thông tin của tín hiệu IQ là yếu tố then chốt quyết định hiệu năng của hệ thống phân loại điều chế dựa trên học sâu. Trên cơ sở phân tích lý thuyết và đánh giá thực nghiệm, luận án đã lựa chọn tín hiệu IQ làm đầu vào chính và đề xuất kiến trúc DFENet nhằm tăng cường khả năng trích xuất đặc trưng đa nhánh, đa thang đo trong miền không gian đặc trưng. Kết quả cho thấy DFENet vượt trội so với các kiến trúc CNN truyền thống trong cùng điều kiện huấn luyện, đặc biệt ở vùng SNR thấp.

Tuy nhiên, mặc dù các mô hình CNN có khả năng học hiệu quả các đặc trưng cục bộ và cấu trúc không gian của tín hiệu, chúng vẫn tồn tại hạn chế trong việc mô hình hóa phụ thuộc theo chuỗi thời gian dài hạn. Trong thực tế, tín hiệu điều chế vô tuyến mang tính chất động học theo thời gian, cấu trúc biên độ và pha không chỉ chứa thông tin cục bộ mà còn phụ thuộc vào mối quan hệ giữa các mẫu tín hiệu theo chuỗi.

Đối với các điều chế bậc cao hoặc trong môi trường fading mạnh, sự phụ thuộc theo thời gian dài hạn có thể mang thông tin phân lớp quan trọng mà các lớp tích chập thuần túy khó khai thác triệt để. Điều này đặt ra yêu cầu phát triển các kiến trúc có khả năng kết hợp đồng thời trích xuất đặc trưng

không gian và mô hình hóa động học theo thời gian.

Xuất phát từ yêu cầu đó, chương 3 đề xuất mô hình AMC-CNN-BiLSTM, trong đó mạng tích chập CNN được sử dụng để trích xuất đặc trưng không gian từ tín hiệu IQ, trong khi mạng bộ nhớ dài hạn hai chiều (BiLSTM) đảm nhiệm việc học phụ thuộc chuỗi theo cả hướng tiến và lùi. Sự kết hợp này cho phép mô hình khai thác toàn diện cấu trúc không gian-thời gian của tín hiệu điều chế. Mô hình đề xuất sẽ được đánh giá toàn diện thông qua phân tích định lượng và so sánh với các phương pháp hiện có, qua đó chứng minh hiệu quả của cách tiếp cận kết hợp không gian-thời gian trong bài toán AMC.

Bên cạnh mô hình kết hợp không gian-thời gian, chương cũng nghiên cứu phương pháp học kết hợp giữa các mô hình có tính bổ sung về đặc trưng, nhằm tăng tính ổn định và khả năng tổng quát hóa trong điều kiện nhiễu mạnh và fading biến động. Thay vì phụ thuộc vào một kiến trúc đơn lẻ, phương pháp học kết hợp hướng tới khai thác sự đa dạng trong không gian mô hình để cải thiện độ tin cậy của quyết định phân loại.

Như vậy, nếu chương 2 tập trung vào tối ưu hóa biểu diễn và thiết kế kiến trúc trích xuất đặc trưng từ tín hiệu IQ, thì chương 3 mở rộng phạm vi nghiên cứu sang khai thác phụ thuộc theo thời gian và tăng cường độ ổn định thông qua chiến lược kết hợp mô hình. Các phương pháp đề xuất sẽ được đánh giá toàn diện và so sánh với các kiến trúc đã trình bày trước đó nhằm làm rõ mức độ cải thiện và giá trị bổ sung của cách tiếp cận không gian-thời gian trong bài toán phân loại điều chế.

3.1. Thiết kế mô hình lai ghép AMC-CNN-BiLSTM

3.1.1. Xây dựng mô hình lai ghép

Trong phần trước của chương 2, luận án đã trình bày mô hình DFENet với trọng tâm là khai thác triệt để các đặc trưng "không gian" của tín hiệu thông qua cơ chế đa tỷ lệ. Các mô hình dựa trên CNN như DFENet đã chứng minh được sự ưu việt trong việc nhận diện các đặc điểm biên độ - pha cục bộ trong không gian I/Q. Tuy nhiên, khi xét đến bản chất vật lý của tín hiệu vô tuyến trong môi trường thực tế, cách tiếp cận thuần túy không gian này bộc lộ những hạn chế nhất định cần được khắc phục để đạt được độ chính xác và độ tin cậy cao hơn.

Thứ nhất, hạn chế về khả năng nắm bắt sự phụ thuộc thời gian: Tín hiệu vô tuyến về bản chất là dữ liệu chuỗi thời gian, nơi mà thông tin điều chế không chỉ được mã hóa tại một thời điểm rời rạc mà còn ẩn chứa trong quy luật diễn biến của chuỗi ký hiệu theo thời gian. Đặc biệt đối với các phương thức điều chế có nhớ (như CPFSK) hoặc trong các môi trường kênh truyền biến đổi nhanh, trạng thái của tín hiệu tại thời điểm hiện tại t có mối liên hệ mật thiết với các trạng thái trong quá khứ $t - 1, t - 2 \dots$ và tương lai. Các bộ lọc tích chập của CNN, do giới hạn bởi kích thước cửa sổ trượt, chỉ có khả năng quan sát cục bộ và thường gặp khó khăn trong việc liên kết các thông tin ngữ cảnh dài hạn nằm cách xa nhau trong chuỗi dữ liệu.

Thứ hai, thách thức từ môi trường Fading đa đường: Trong các kênh truyền như Rayleigh hay Rician (được mô phỏng trong tập dữ liệu HisarMod2019.1), tín hiệu chịu tác động của hiện tượng fading, khiến biên độ và pha bị suy lạc nghiêm trọng tại một số thời điểm ngẫu nhiên. Nếu chỉ dựa vào đặc trưng

hình ảnh cục bộ tại các điểm fading này, mạng nơ-ron rất dễ đưa ra quyết định sai lầm do thông tin bị mờ hoặc mất hẳn. Để vượt qua điều này, hệ thống cần một cơ chế có khả năng "ghi nhớ" và "suy luận" ngữ cảnh toàn cục: sử dụng thông tin từ các đoạn tín hiệu rõ ràng ở lân cận để khôi phục và dự đoán nhãn cho các đoạn tín hiệu bị nhiễu.

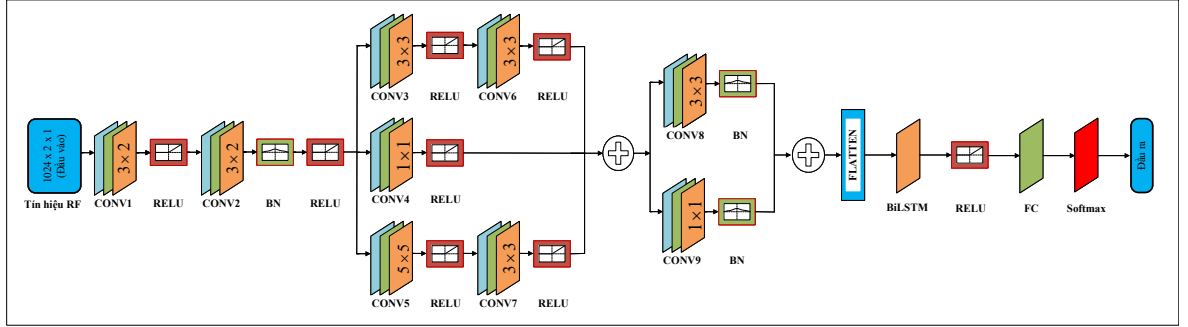
Giải pháp tiếp cận Lai ghép: Xuất phát từ những phân tích trên, nhu cầu kết hợp sức mạnh trích xuất đặc trưng hình thái học của CNN với khả năng xử lý chuỗi của Mạng nơ-ron hồi quy (RNN) trở nên cấp thiết. Tuy nhiên, việc áp dụng trực tiếp RNN (như LSTM) lên dữ liệu thô I/Q có chiều dài lớn (1024 mẫu) thường dẫn đến chi phí tính toán khổng lồ và khó hội tụ do vấn đề bùng nổ đạo hàm. Do đó, luận án đề xuất mô hình AMC-CNN-BiLSTM, một kiến trúc lai ghép được thiết kế theo chiến lược phân tầng tối ưu.

Tầng CNN đa nhánh: Đóng vai trò là bộ trích xuất đặc trưng ở mức thấp. Nhiệm vụ của nó là lọc nhiễu, nén chiều dữ liệu và chuyển đổi các mẫu tín hiệu thô thành một chuỗi các vector đặc trưng trừu tượng, giàu thông tin nhưng gọn nhẹ hơn.

Tầng BiLSTM (Bidirectional LSTM): Đóng vai trò là bộ phân tích chuỗi ở mức cao. Điểm đột phá ở đây là việc sử dụng cơ chế hai chiều thay vì một chiều. Trong bài toán phân loại điều chế, việc mạng nơ-ron có thể "nhìn" thấy cả quá khứ và tương lai cho phép xây dựng một vector ngữ cảnh toàn diện nhất, giúp mô hình đưa ra quyết định chính xác ngay cả khi tín hiệu bị phân mảnh do nhiễu. Mô hình này không chỉ là sự kế thừa mà là một bước phát triển logic từ DFENet, hướng tới việc giải quyết bài toán AMC một cách toàn diện trên cả hai miền: Không gian và Thời gian.

Mô hình tổng quan

Mô hình được xây dựng với dữ liệu đầu vào là ma trận I/Q kích thước 1024×2 . Mô hình tổng thể của mô hình AMC-CNN-BiLSTM được minh họa trong Hình 3.1.



Hình 3.1: Mô hình đề xuất của mô hình AMC-CNN-BiLSTM.

Luồng xử lý bao gồm 3 giai đoạn chính:

Giai đoạn Không gian (Spatial Feature Extraction): Sử dụng mạng CNN đa nhánh để chuyển đổi tín hiệu thô thành bản đồ đặc trưng trừu tượng.

Giai đoạn Thời gian (Temporal Feature Learning): Sử dụng lớp BiLSTM để học các mối quan hệ ngữ cảnh hai chiều.

Giai đoạn Phân loại (Classification): Sử dụng các lớp kết nối đầy đủ (Fully Connected) để đưa ra dự đoán.

Thiết kế tầng CNN đa nhánh (Multi-branch CNN)

Tầng CNN được thiết kế đặc biệt để tối đa hóa khả năng biểu diễn đặc trưng trước khi đưa vào lớp hồi quy.

a. Khởi tạo và Ổn định: Tín hiệu đầu vào X trước tiên đi qua lớp tích chập khởi tạo (Conv1) để trích xuất các đặc trưng biên độ và tần số sơ cấp. Ngay sau đó, lớp tích chập thứ hai (Conv2) kết hợp với Chuẩn hóa theo BN (Batch Normalization) được sử dụng.

Công thức lớp khởi tạo:

$$Y_1 = \text{ReLU}(\text{Conv}(X, W_1) + b_1), \quad (3.1)$$

trong đó X biểu diễn tín hiệu đầu vào, W_1 là ma trận trọng số của lớp tích chập, và b_1 là tham số sai lệch.

Lớp này có nhiệm vụ trích xuất các đặc trưng cục bộ ban đầu của tín hiệu, chẳng hạn như sự biến thiên biên độ, pha và các thành phần tần số ngắn hạn. Hàm kích hoạt ReLU được sử dụng nhằm đưa tính phi tuyến vào mô hình, cho phép mạng nơ-ron học được các quan hệ phức tạp vốn tồn tại trong tín hiệu điều chế.

Công thức lớp ổn định:

$$Y_2 = \text{ReLU}(\text{Conv}(Y_1, W_2) + b_2), \quad (3.2)$$

Cấu trúc này giúp tinh chỉnh quá trình trích xuất đặc trưng bằng cách học các biểu diễn mức cao hơn. Việc tích hợp lớp Batch Normalization giúp ổn định quá trình huấn luyện bằng cách giảm hiện tượng dịch chuyển phân phối nội bộ, từ đó tăng tốc độ hội tụ và cải thiện khả năng tổng quát hóa của mô hình. Đồng thời, hàm kích hoạt ReLU tiếp tục duy trì khả năng xử lý phi tuyến, đảm bảo rằng các đặc trưng trích xuất được có khả năng biểu diễn hiệu quả các dạng điều chế phức tạp trong điều kiện kênh truyền có nhiễu.

b. Cấu trúc Rẽ nhánh song song:

Sau các lớp trích xuất đặc trưng ban đầu, mô hình được phân tách thành ba nhánh xử lý song song nhằm tăng cường khả năng học đặc trưng đa tỷ lệ. Cấu trúc đa nhánh này được thiết kế với mục tiêu khai thác đồng thời các đặc trưng của tín hiệu ở nhiều mức độ trừu tượng khác nhau, từ các biến thiên cục bộ đơn giản đến các mẫu cấu trúc phức tạp hơn.

Nhánh 1 bao gồm một lớp tích chập duy nhất, theo sau là hàm kích hoạt ReLU, được mô tả như sau:

$$Y_{3a} = \text{ReLU}(\text{Conv}(Y_2, W_{3a}) + b_{3a}). \quad (3.3)$$

Nhánh này tập trung vào việc trích xuất các đặc trưng cục bộ cơ bản của tín hiệu, chẳng hạn như các thay đổi ngắn hạn về biên độ và pha, đóng vai trò như một bộ trích xuất đặc trưng mức thấp.

Nhánh 2 được cấu thành bởi hai lớp tích chập liên tiếp, mỗi lớp đều được kết hợp với hàm kích hoạt ReLU:

$$Y_{3b1} = \text{ReLU}(\text{Conv}(Y_2, W_{3b1}) + b_{3b1}), \quad (3.4)$$

$$Y_{3b2} = \text{ReLU}(\text{Conv}(Y_{3b1}, W_{3b2}) + b_{3b2}). \quad (3.5)$$

Cấu trúc sâu hơn của nhánh này cho phép mô hình học được các đặc trưng có mức độ trừu tượng cao hơn, phản ánh các mối quan hệ phi tuyến và các mẫu điều chế phức tạp trong tín hiệu.

Nhánh 3 có kiến trúc tương tự nhánh 2, nhưng sử dụng các cấu hình bộ lọc khác nhau nhằm khai thác các tập đặc trưng bổ sung:

$$Y_{3c1} = \text{ReLU}(\text{Conv}(Y_2, W_{3c1}) + b_{3c1}), \quad (3.6)$$

$$Y_{3c2} = \text{ReLU}(\text{Conv}(Y_{3c1}, W_{3c2}) + b_{3c2}). \quad (3.7)$$

Nhờ sự đa dạng trong cấu hình bộ lọc, nhánh này giúp mô hình nắm bắt các đặc trưng khác biệt về mặt phổ và cấu trúc thời gian, vốn khó có thể được biểu diễn đầy đủ bởi một kiến trúc đơn nhánh.

Cách tiếp cận đa nhánh cho phép mô hình đồng thời học các đặc trưng ở nhiều mức độ trừu tượng khác nhau. Trong đó, nhánh nông tập trung vào

các đặc trưng mức thấp, trong khi các nhánh sâu hơn đóng vai trò trích xuất các đặc trưng mức cao và mang tính ngữ nghĩa hơn. Sự kết hợp này đảm bảo rằng mô hình có khả năng nhận diện một phổ rộng các đặc trưng tín hiệu tồn tại ở nhiều dạng và nhiều thang đo khác nhau.

c. Hợp nhất đặc trưng:

Đầu ra của ba nhánh được cộng gộp để tạo ra một bản đồ đặc trưng thống nhất, giàu thông tin:

$$Y_4 = \text{ReLU}(Y_{3a} + Y_{3b} + Y_{3c}). \quad (3.8)$$

Phép tổng hợp này cho phép tích lũy thông tin từ các đường xử lý song song, đồng thời vẫn bảo toàn đóng góp riêng của từng nhánh. Hàm kích hoạt ReLU được áp dụng nhằm tăng cường khả năng biểu diễn phi tuyến, qua đó nâng cao năng lực phân biệt giữa các dạng điều chế có cấu trúc tín hiệu tương đồng nhưng khác biệt tinh vi.

Tiếp theo, mô hình được tách thành hai nhánh song song, mỗi nhánh bao gồm một lớp tích chập theo sau bởi chuẩn hóa theo BN, nhằm tiếp tục tinh chỉnh và ổn định biểu diễn đặc trưng. Thiết kế này đảm bảo tính ổn định của đặc trưng trước các biến động của nhiễu.

$$Y_{5a} = \text{BatchNorm}(\text{Conv}(Y_4, W_{5a}) + b_{5a}) \quad (3.9)$$

$$Y_{5b} = \text{BatchNorm}(\text{Conv}(Y_4, W_{5b}) + b_{5b}) \quad (3.10)$$

$$Y_6 = Y_{5a} + Y_{5b} \quad (3.11)$$

Việc sử dụng hai nhánh tích chập song song ở giai đoạn này cho phép mô hình học các biến thể đặc trưng khác nhau từ cùng một không gian đầu vào, trong khi Batch Normalization giúp chuẩn hóa các bản đồ đặc trưng, giảm

độ nhạy đối với điều kiện khởi tạo và cải thiện tính ổn định trong quá trình huấn luyện. Phép cộng phần tử được sử dụng để hợp nhất đầu ra của hai nhánh, nhằm bảo toàn thông tin quan trọng và tăng cường độ bền vững của các đặc trưng đã học, đặc biệt trong điều kiện nhiễu cao.

Thiết kế tầng BiLSTM

Đây là thành phần tạo nên sự khác biệt so với mô hình DFENet. Sau khi được làm phẳng (Flatten), vector đặc trưng Y_7 được đưa vào lớp BiLSTM.

$$Y_7 = \text{Flatten}(Y_6) \quad (3.12)$$

$$Y_{\text{out}} = \text{BiLSTM}(Y_7) \quad (3.13)$$

Khác với các lớp tích chập vốn tập trung vào trích xuất đặc trưng cục bộ, lớp BiLSTM được thiết kế để mô hình hóa các phụ thuộc theo chuỗi và mối quan hệ dài hạn trong tín hiệu. Cụ thể, BiLSTM xử lý chuỗi đặc trưng theo cả hai hướng tiến và lùi trong thời gian, được biểu diễn như sau:

$$\vec{h}_t = \text{LSTM}(Y_7, \vec{h}_{t-1}) \quad (3.14)$$

$$\overleftarrow{h}_t = \text{LSTM}(Y_7, \overleftarrow{h}_{t+1}) \quad (3.15)$$

$$h_t = [\vec{h}_t, \overleftarrow{h}_t] \quad (3.16)$$

Trong đó, $\vec{h}_t$ và $\overleftarrow{h}_t$ lần lượt là trạng thái ẩn theo hướng tiến và hướng lùi.

Việc kết hợp hai trạng thái này cho phép mô hình đồng thời khai thác thông tin từ quá khứ và tương lai của chuỗi đặc trưng, qua đó nắm bắt đầy đủ bối cảnh thời gian của tín hiệu.

Sự tích hợp lớp BiLSTM ở giai đoạn cuối đóng vai trò then chốt trong việc nâng cao khả năng mô hình hóa các phụ thuộc thời gian phức tạp của tín hiệu vô tuyến. Điều này đặc biệt quan trọng đối với bài toán nhận dạng điều

chế tự động, nơi các dạng điều chế khác nhau có thể chia sẻ các đặc trưng cục bộ tương tự nhưng lại khác biệt ở cấu trúc chuỗi và động học theo thời gian. Nhờ đó, mô hình không chỉ cải thiện độ chính xác phân loại mà còn tăng cường khả năng khái quát hóa trước các biến thiên về nhiễu, kênh truyền và điều kiện hoạt động của hệ thống. Việc kết hợp trích xuất đặc trưng đa nhánh bằng CNN và mô hình hóa phụ thuộc thời gian hai chiều bằng BiLSTM tạo nên một kiến trúc lai hiệu quả, phù hợp với bản chất chuỗi của tín hiệu IQ.

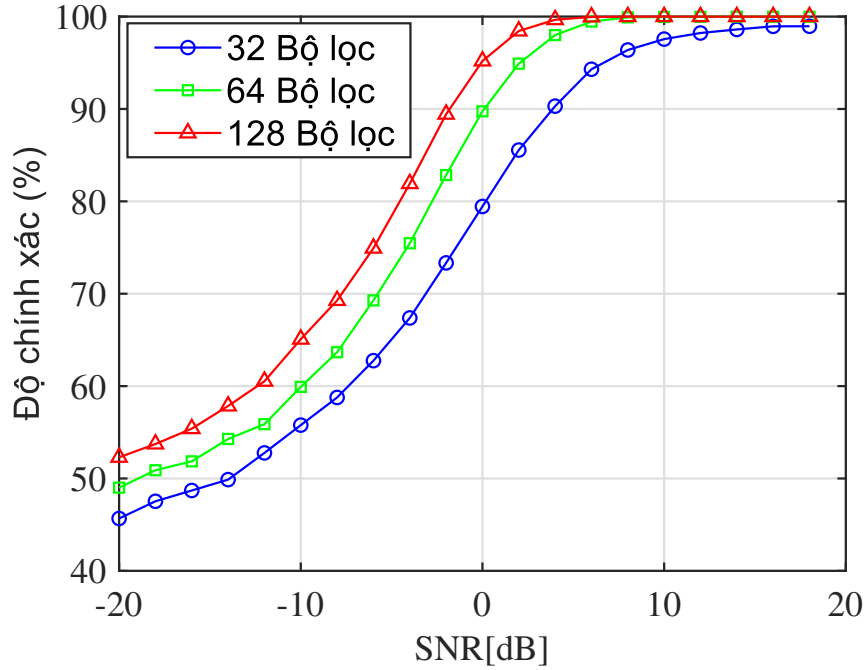
3.1.2. Kết quả thực nghiệm và thảo luận

Đánh giá hiệu năng mô hình khi thay đổi số lượng bộ lọc

Kết quả thí nghiệm về số lượng bộ lọc tích chập và độ dài tín hiệu đầu vào cung cấp những hàm ý thiết kế quan trọng, trực tiếp định hướng cho việc xây dựng kiến trúc của hai mô hình đề xuất là DFENet và AMC-CNN-BiLSTM.

Trước hết, kết quả trong Hình 3.2 cho thấy việc tăng số lượng bộ lọc tích chập giúp cải thiện đáng kể độ chính xác phân loại. Điều này phản ánh rằng bài toán phân loại điều chế tự động đòi hỏi không gian đặc trưng đủ phong phú để nắm bắt các khác biệt tinh vi giữa các dạng điều chế, đặc biệt trong điều kiện SNR thấp. Tuy nhiên, việc gia tăng số bộ lọc một cách trực tiếp cũng làm tăng đáng kể số lượng tham số và độ phức tạp tính toán, từ đó làm giảm khả năng triển khai trong các hệ thống thực tế.

Để giải quyết mâu thuẫn giữa độ chính xác và độ phức tạp, các mô hình DFENet và AMC-CNN-BiLSTM không lựa chọn cách mở rộng mạng theo chiều rộng một cách đơn thuần, mà áp dụng kiến trúc tích chập đa nhánh. Thông qua việc chia quá trình trích xuất đặc trưng thành nhiều nhánh song song với các kích thước bộ lọc khác nhau, mô hình có thể học được các đặc



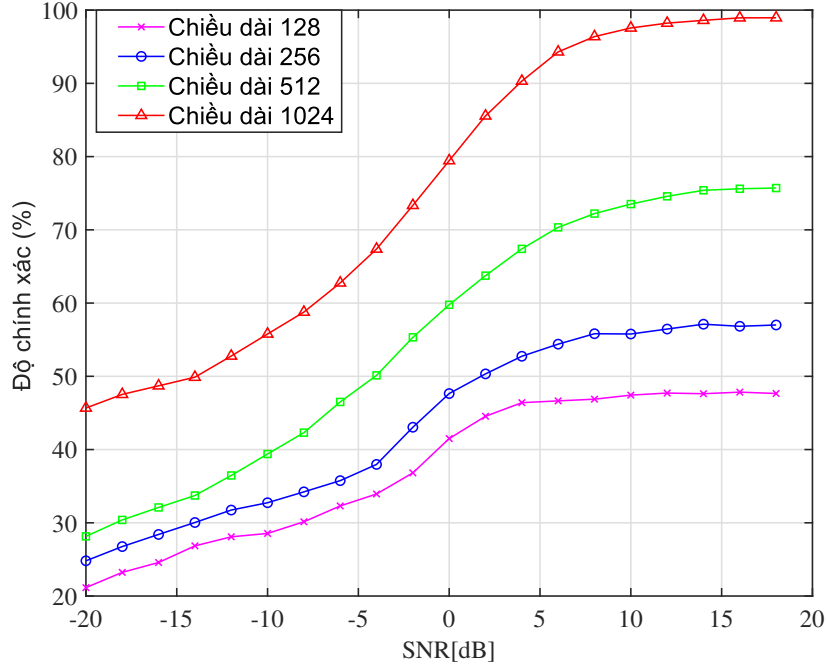
Hình 3.2: Độ chính xác phân loại với số lượng bộ lọc khác nhau.

trung đa tỉ lệ trong khi vẫn duy trì số lượng tham số ở mức hợp lý. Cách tiếp cận này cho phép khai thác tính đa dạng của đặc trưng tín hiệu mà không phải đánh đổi bằng chi phí tính toán quá lớn như các CNN một nhánh có số bộ lọc lớn.

Đánh giá hiệu năng mô hình khi thay đổi số độ dài tín hiệu

Bên cạnh đó, thí nghiệm về độ dài tín hiệu trong Hình 3.3 chỉ ra rằng tín hiệu có độ dài lớn hơn mang lại độ chính xác phân loại cao hơn trên toàn bộ dải SNR. Kết quả này khẳng định vai trò then chốt của thông tin theo miền thời gian trong bài toán AMC, bởi các tín hiệu dài cung cấp biểu diễn đầy đủ hơn về cấu trúc điều chế và động học tín hiệu. Trên cơ sở đó, mô hình AMC-CNN-BiLSTM tích hợp lớp BiLSTM nhằm mô hình hóa tường minh các phụ thuộc thời gian dài hạn, cho phép khai thác đồng thời ngữ cảnh quá khứ và tương lai của chuỗi tín hiệu. Đối với DFENet, việc xếp chồng các khối

DFE đóng vai trò tương tự, giúp tinh lọc đặc trưng theo chiều sâu và tăng cường khả năng học các quan hệ thời gian tiềm ẩn.



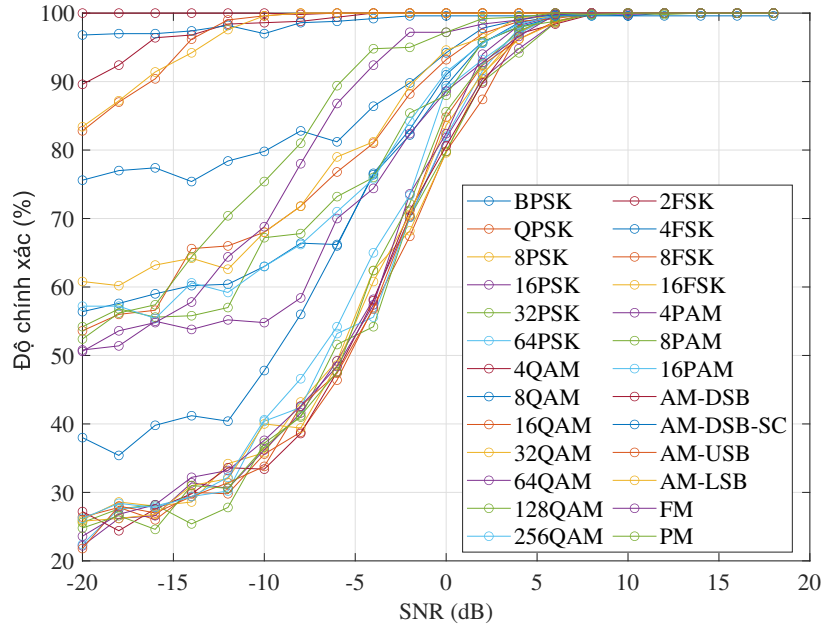
Hình 3.3: Độ chính xác phân loại với các độ dài tín hiệu khác nhau.

Các kết quả thực nghiệm đã xác nhận rằng thiết kế của DFENet và AMC-CNN-BiLSTM không mang tính thử nghiệm ngẫu nhiên mà được xây dựng dựa trên các quan sát thực nghiệm có cơ sở. Hai mô hình đạt được sự cân bằng có nguyên tắc giữa khả năng biểu diễn, mô hình hóa thời gian và hiệu quả tính toán, qua đó chứng minh tính phù hợp cho các bài toán phân loại điều chế tự động trong môi trường vô tuyến thực tế.

Đánh giá hiệu năng mô hình với các dạng điều chế khác nhau

Trong thí nghiệm đánh giá tiếp theo, nghiên cứu tập trung phân tích hiệu quả phân loại của mô hình đối với các dạng điều chế khác nhau. Cụ thể, các kết quả thực nghiệm được minh họa trong Hình 3.4 cho thấy hiệu năng phân loại được cải thiện rõ rệt đối với tất cả các dạng điều chế khi tỷ số tín hiệu

trên nhiều (SNR) đạt từ 10 dB trở lên. Đáng chú ý, các dạng điều chế như AM-DSB, AM-DSB-SC, AM-USB, AM-LSB, FM, PM, 2FSK, 4FSK, 8QAM và BPSK đạt độ chính xác phân loại rất cao, liên tục vượt mức 94% trong dải SNR từ 0 dB đến +18 dB.



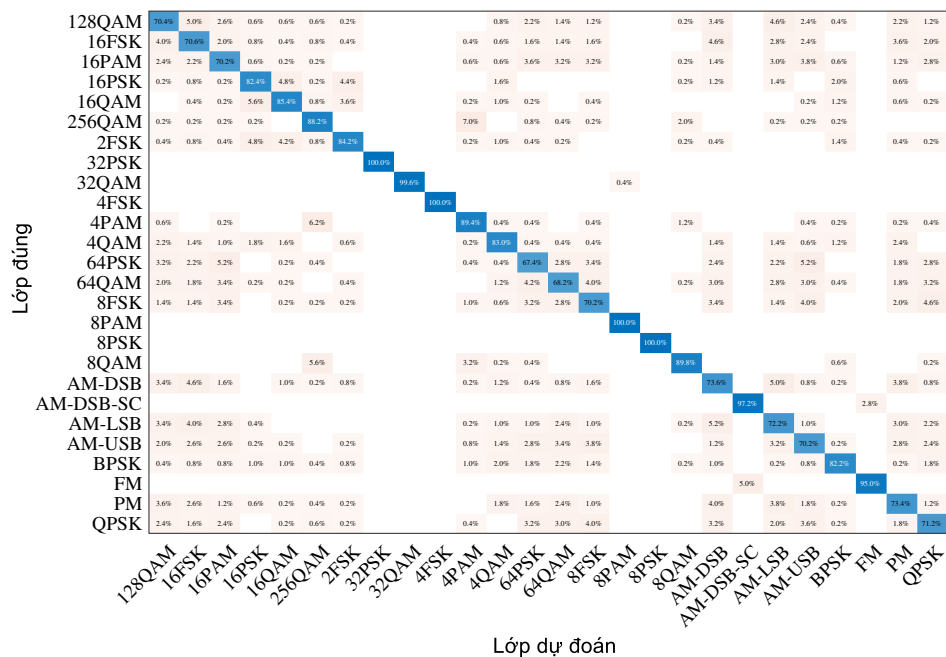
Hình 3.4: Độ chính xác phân loại cho 26 loại điều chế khác nhau.

Bên cạnh đó, các dạng điều chế như 4QAM, 8FSK, 16FSK, QPSK và 4PAM cũng thể hiện hiệu năng phân loại mạnh mẽ, với độ chính xác dao động trong khoảng từ 90.6% đến 99% tại các mức SNR từ +5 dB đến +18 dB. Những kết quả này khẳng định ưu thế của mô hình CNN đề xuất trong việc duy trì độ chính xác cao trên nhiều loại điều chế khác nhau và trong các điều kiện nhiễu đa dạng.

Khả năng đạt độ chính xác gần như tuyệt đối đối với phần lớn các dạng điều chế tại các mức SNR cao cho thấy mô hình có năng lực phân biệt hiệu quả các mẫu tín hiệu phức tạp. Tuy nhiên, đối với một số dạng điều chế bậc cao như PSK và QAM, mặc dù hiệu năng vẫn ở mức cao, độ chính xác chưa

đạt mức tối đa trong điều kiện SNR thấp. Điều này phản ánh thách thức cố hữu của bài toán AMC khi xử lý các dạng điều chế có chòm sao tín hiệu dày đặc, và đồng thời cho thấy dư địa cải tiến của mô hình trong các kịch bản nhiễu nghiêm trọng.

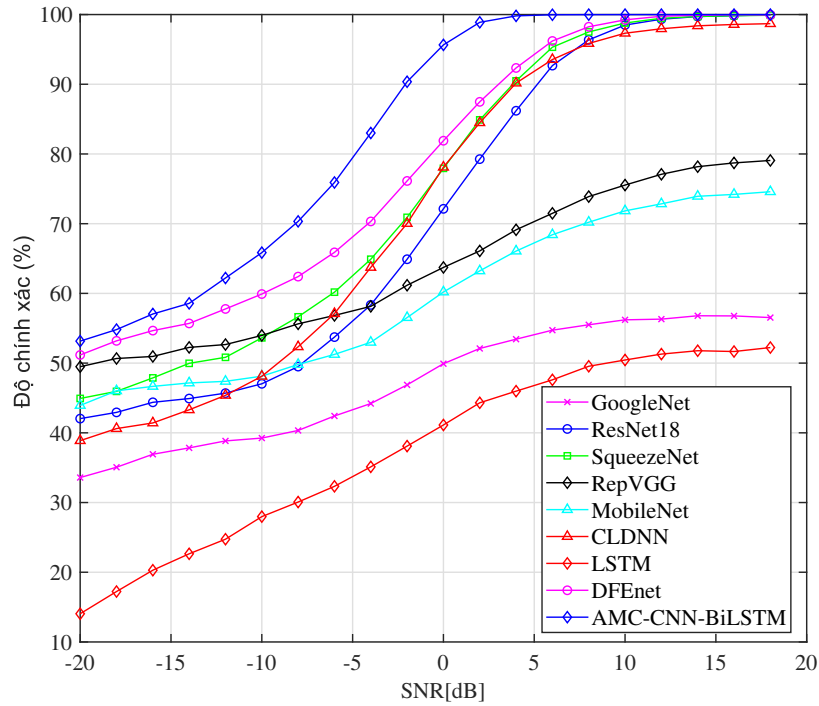
Ngoài ra, Hình 3.5 trình bày các ma trận nhầm lẫn tương ứng với 26 dạng điều chế tại mức SNR = +10 dB. Kết quả cho thấy các dạng điều chế bậc thấp như AM và FM đạt độ chính xác trên 98%, trong khi các dạng điều chế bậc cao như PSK và QAM đạt độ chính xác xấp xỉ 94%. Hiện tượng này làm nổi bật sự đánh đổi giữa tốc độ truyền dữ liệu và độ chính xác phân loại: các dạng điều chế bậc cao cho phép truyền dữ liệu với tốc độ lớn hơn, nhưng đồng thời dễ bị nhiễu hơn do khoảng cách giữa các điểm chòm sao nhỏ, dẫn đến khó khăn trong phân loại tự động điều chế trong môi trường nhiễu.



Hình 3.5: Ma trận nhầm lẫn của 26 phân loại điều chế ở mức SNR +10 dB.

So sánh mô hình với các mô hình hiện nay

Cuối cùng, nghiên cứu tiến hành đánh giá và so sánh hiệu năng của mô hình đề xuất với các mô hình học sâu tiêu biểu, bao gồm ResNet, GoogleNet, MobileNet, SqueezeNet, RepVGG, LSTM, CLDNN và AMC-CNN-BiLSTM. Kết quả so sánh được trình bày trong Hình 3.6 cho thấy mô hình CNN đề xuất mang lại sự cải thiện đáng kể về độ chính xác trong bài toán phân loại điều chế tự động (AMC).



Hình 3.6: So sánh độ chính xác phân loại của các mô hình CNN khác nhau.

Cụ thể, mô hình AMC-CNN-BiLSTM đạt độ chính xác cao nhất trên toàn bộ dải SNR từ -20 dB đến $+18$ dB, với độ chính xác trung bình đạt khoảng 79%. Trong điều kiện SNR thấp (-20 dB đến $+2$ dB), mô hình đề xuất vượt trội so với ResNet18 khoảng 10% và vượt LSTM tới 30%. Điều này cho thấy khả năng thích nghi mạnh mẽ của AMC-CNN-BiLSTM trong môi trường

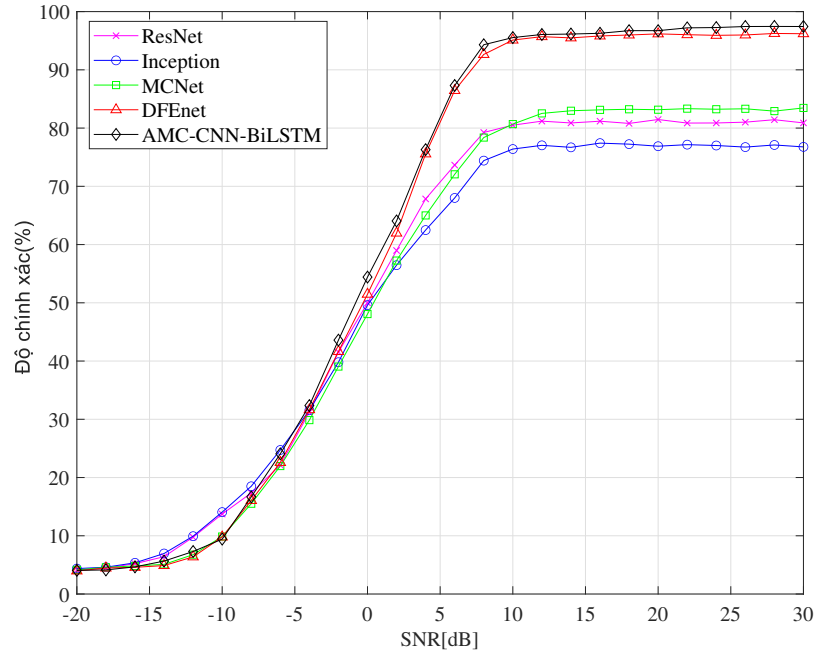
nhiều cao, nơi các đặc trưng tín hiệu trở nên khó phân biệt.

Ở dải SNR trung bình và cao (+2 dB đến +18 dB), AMC-CNN-BiLSTM tiếp tục duy trì ưu thế rõ rệt, với mức cải thiện độ chính xác khoảng 10% so với RepVGG và khoảng 6% so với SqueezeNet. Các mô hình CNN truyền thống như GoogleNet và SqueezeNet cho hiệu năng khá ổn định, tuy nhiên vẫn kém hơn so với kiến trúc lai CNN-BiLSTM được đề xuất. Trong khi đó, mô hình LSTM thuần túy cho kết quả thấp nhất trên hầu hết các mức SNR, phản ánh hạn chế của việc thiếu cơ chế trích xuất đặc trưng không gian hiệu quả.

Những kết quả này khẳng định rằng việc kết hợp trích xuất đặc trưng không gian đa nhánh bằng CNN với khả năng mô hình hóa phụ thuộc thời gian hai chiều của BiLSTM đã mang lại lợi thế rõ rệt cho AMC-CNN-BiLSTM. Mô hình này cho phép mô hình khai thác đồng thời thông tin không gian và thời gian của tín hiệu vô tuyến, từ đó nâng cao đáng kể độ chính xác phân loại so với các mô hình hiện có.

Luận án tiếp tục đánh giá mô hình AMC-CNN-BiLSTM với bốn mô hình bao gồm DFENet MCNet, Inception và ResNet, trên bộ dữ liệu RadioML2018, như minh họa trong Hình 3.7.

Kết quả trên Hình 3.7 cho thấy AMC-CNN-BiLSTM đạt hiệu năng phân loại vượt trội so với DFENet, đặc biệt trong vùng SNR trung bình và cao. Khi SNR tăng, độ chính xác của AMC-CNN-BiLSTM tăng nhanh và đạt khoảng 95–97% ở vùng SNR từ 10 dB trở lên, trong khi DFENet đạt khoảng 94–96%. Khoảng cách về độ chính xác giữa hai mô hình tuy không lớn ở vùng SNR cao nhưng AMC-CNN-BiLSTM duy trì kết quả cao và ổn định hơn trên phần lớn dải SNR. Kết quả này cho thấy việc kết hợp CNN với BiLSTM giúp



Hình 3.7: So sánh độ chính xác phân loại của các mô hình CNN đơn lẻ khác nhau đối với tập dữ liệu Radioml2018.

mô hình khai thác đồng thời đặc trưng cục bộ và quan hệ phụ thuộc theo thời gian của tín hiệu, từ đó nâng cao khả năng nhận dạng so với DFENet.

3.2. Phương pháp học kết hợp để cải thiện hiệu năng AMC

Mặc dù mô hình AMC-CNN-BiLSTM đã chứng minh được khả năng khai thác hiệu quả mô hình không gian-thời gian của tín hiệu IQ, các kết quả thực nghiệm cho thấy hiệu năng của mô hình vẫn phụ thuộc đáng kể vào điều kiện kênh truyền và mức SNR. Trong môi trường nhiễu mạnh hoặc fading phức tạp, độ tin cậy của dự đoán có thể biến thiên giữa các lần huấn luyện, phản ánh sự nhạy cảm của mô hình đơn lẻ đối với nhiễu và biến động phân bố dữ liệu.

Về bản chất, mỗi mô hình mạng học sâu chỉ là một xấp xỉ cho ánh xạ phân lớp tối ưu trong không gian giả thiết hữu hạn. Các mô hình khác nhau

có xu hướng nhấn mạnh những khía cạnh đặc trưng khác nhau của dữ liệu. Chẳng hạn, DFENet được thiết kế nhằm đa dạng hóa đặc trưng không gian thông qua cấu trúc đa nhánh, trong khi AMC-CNN-BiLSTM tập trung vào mô hình hóa phụ thuộc theo chuỗi thời gian. Do đó, các mô hình này không hoàn toàn thay thế nhau mà có thể mang tính bổ sung về năng lực biểu diễn.

Từ góc nhìn học máy thống kê, sai số của một mô hình đơn lẻ có thể được phân rã thành các thành phần kỳ vọng và phương sai. Trong các bài toán chịu ảnh hưởng mạnh của nhiễu và thay đổi phân bố theo SNR như AMC, thành phần phương sai thường đóng vai trò đáng kể. Điều này dẫn đến hiện tượng dự đoán thiếu ổn định khi điều kiện kênh thay đổi. Phương pháp học kết hợp được xem là một cơ chế hiệu quả nhằm giảm phương sai thông qua việc tổng hợp dự đoán của nhiều mô hình có sai số không hoàn toàn tương quan. Khi các mô hình thành phần học các biểu diễn khác nhau của dữ liệu, việc kết hợp chúng có thể làm giảm sai số tổng thể và tăng tính ổn định của hệ thống.

Trong bối cảnh phân loại điều chế, nơi tín hiệu chịu tác động đồng thời của nhiễu, fading, sai lệch tần số và sự chồng lấn đặc trưng giữa các điều chế bậc cao, khó có một kiến trúc đơn lẻ nào duy trì hiệu năng tối ưu trên toàn bộ không gian điều kiện hoạt động. Thực nghiệm cũng cho thấy các mô hình có cấu trúc và độ phức tạp khác nhau thường đạt hiệu năng khác nhau tùy theo miền SNR và loại điều chế. Điều này củng cố nhận định rằng tính đa dạng mô hình là yếu tố quan trọng cần được khai thác.

Do đó, thay vì lựa chọn một mô hình duy nhất, việc kết hợp nhiều mô hình có đặc tính bổ sung là một hướng tiếp cận hợp lý nhằm nâng cao độ chính xác trung bình, giảm phương sai dự đoán và cải thiện tính bền vững

của hệ thống.

Xuất phát từ nhận định này, mục này tập trung nghiên cứu hướng tiếp cận phương pháp học kết hợp cho bài toán AMC, trong đó nhiều mạng CNN được huấn luyện song song trên các dạng biểu diễn tín hiệu khác nhau và được kết hợp ở mức quyết định. Mục tiêu là nâng cao độ chính xác trung bình, giảm phương sai của mô hình, và cải thiện độ ổn định của hệ thống phân loại trong các môi trường kênh phức tạp và biến thiên mạnh.

Bên cạnh đó, cần nhấn mạnh rằng mục tiêu của phương pháp học kết hợp được trình bày trong mục này không nhằm thay thế hay so sánh trực tiếp với các mô hình DFENet và AMC-CNN-BiLSTM đã được đề xuất và phân tích chi tiết ở chương 2 và chương 3. Trong phạm vi mục 3.2 của luận án, học kết hợp được sử dụng như một công cụ phân tích nhằm đánh giá tác động của tính đa dạng kiến trúc và biểu diễn đầu vào đến hiệu năng phân loại điều chế tự động.

Cụ thể, các mô hình cơ sở được lựa chọn trong học kết hợp (ResNet, MobileNet, SqueezeNet, RepVGG, v.v.) đều là các mô hình CNN phổ biến, đại diện cho các chiến lược trích xuất đặc trưng khác nhau, từ mạng nhẹ đến mạng sâu với kết nối dư. Việc sử dụng các mô hình chuẩn này giúp làm nổi bật vai trò của phương pháp học kết hợp mô hình và tránh sự chồng lấn về mặt thiết kế với các mô hình chuyên biệt đã được đề xuất trước đó.

Ngược lại, các mô hình DFENet và AMC-CNN-BiLSTM được thiết kế với mục tiêu tối ưu hóa trích xuất đặc trưng và mô hình hóa quan hệ thời gian–tần số một cách chuyên sâu. Do đó, việc đưa trực tiếp các mô hình này vào học kết hợp trong giai đoạn này có thể làm mờ đi vai trò phân tích của học kết hợp, vốn tập trung vào khai thác tính bổ sung giữa các mô hình cơ

bản.

3.2.1. Mô hình học kết hợp CNN

Phương pháp học kết hợp

Trong bài toán AMC, tín hiệu IQ thô chứa đầy đủ thông tin về biên độ, pha và cấu trúc thời gian của tín hiệu, nhưng đồng thời cũng chịu ảnh hưởng mạnh của nhiễu và các biến dạng kênh. Khi chỉ sử dụng một mô hình CNN đơn lẻ, quá trình học đặc trưng thường bị thiên lệch theo kiến trúc cụ thể của mạng, dẫn đến hiện tượng quá khớp cục bộ hoặc khả năng tổng quát hóa hạn chế trong các điều kiện SNR thấp.

Phương pháp học kết hợp được đề xuất nhằm khắc phục hạn chế này bằng cách kết hợp nhiều bộ phân loại độc lập, mỗi bộ học một cách biểu diễn khác nhau từ cùng một dữ liệu đầu vào. Về mặt lý thuyết, học kết hợp giúp [61], [62]:

- Giảm phương sai của mô hình tổng hợp so với từng mô hình đơn lẻ.
- Khai thác tính bổ sung giữa các kiến trúc CNN khác nhau.
- Tăng độ ổn định của kết quả phân loại trong môi trường nhiễu cao.

Mô hình và quy trình học kết hợp

Trong mô hình đề xuất, nhiều mạng CNN cơ sở với mô hình khác nhau (ví dụ: CNN nông, CNN sâu, CNN với kernel kích thước khác nhau) được huấn luyện song song trên cùng tập tín hiệu IQ thô. Mỗi mạng đưa ra một phân bố xác suất dự đoán trên tập các lớp điều chế.

Giả sử K mô hình CNN cơ sở được sử dụng, đầu ra của mô hình ensemble

được xác định theo quy tắc tổng hợp sau [63]:

$$\hat{y} = \arg \max_c \left(\frac{1}{K} \sum_{k=1}^K p_k(c | \mathbf{x}) \right), \quad (3.17)$$

trong đó $p_k(c | \mathbf{x})$ là xác suất dự đoán lớp c của mô hình thứ k đối với tín hiệu đầu vào $\mathbf{x}$.

Phương pháp này cho phép mô hình học kết hợp tận dụng các quyết định đồng thuận giữa các mạng, đồng thời làm giảm ảnh hưởng của các dự đoán sai lẻ tẻ từ từng mô hình riêng biệt.

3.2.2. Mô hình học kết hợp và phương pháp kết hợp quyết định

Trong nghiên cứu này, học kết hợp được xây dựng dựa trên nhiều mô hình CNN tiêu biểu, bao gồm ResNet18, MobileNet và GoogleNet, được huấn luyện độc lập với cùng tập dữ liệu đầu vào sau tiền xử lý (STFT hoặc STFT–Radon). Mỗi mô hình đóng vai trò như một bộ phân loại cơ sở, học các đặc trưng khác nhau từ cùng một tín hiệu vô tuyến [64], [65].

Tại giai đoạn suy luận, đầu ra của mỗi CNN được biểu diễn dưới dạng vector xác suất hậu nghiệm cho các lớp điều chế. Các vector này sau đó được kết hợp bằng phương pháp hợp nhất ở mức quyết định, trong đó các xác suất dự đoán được trung bình hoặc gán trọng số để tạo ra quyết định cuối cùng của học kết hợp.

Cách tiếp cận này có hai ưu điểm chính [66], [67]. Thứ nhất, nó không yêu cầu thay đổi cấu trúc bên trong của từng mạng CNN, do đó dễ dàng mở rộng và triển khai trong thực tế. Thứ hai, việc hợp nhất ở mức quyết định cho phép hệ thống tận dụng các điểm mạnh riêng của từng mô hình, đồng thời giảm tác động của các dự đoán sai cục bộ do nhiễu hoặc đặc trưng không rõ ràng trong tín hiệu.

3.2.3. Phương pháp học kết hợp CNN cho phân loại điều chế

Trong mục này, luận án đề xuất một phương pháp học kết hợp nhằm nâng cao độ chính xác của bài toán phân loại điều chế tự động bằng cách kết hợp đầu ra của nhiều mô hình CNN độc lập. Ý tưởng cốt lõi của phương pháp học kết hợp là khai thác tính đa dạng trong quá trình học đặc trưng của các mô hình khác nhau, từ đó giảm sai lệch cục bộ của từng bộ phân loại đơn lẻ và cải thiện khả năng tổng quát hóa của hệ thống [62].

Thông qua việc xây dựng bộ phân loại tổ hợp, nghiên cứu này đánh giá mức độ cải thiện hiệu năng của mô hình học kết hợp so với từng mô hình CNN riêng lẻ. Ba phương pháp học kết hợp phổ biến được xem xét, bao gồm: (i) trung bình xác suất không trọng số, (ii) bỏ phiếu đa số không trọng số, và (iii) trung bình xác suất có trọng số.

Giả sử sử dụng n mô hình CNN độc lập, mỗi mô hình cho ra một vector xác suất $\mathbf{y}_i = [y_{i1}, y_{i2}, \dots, y_{iC}]$ sau lớp softmax, trong đó C là số lớp điều chế.

Trung bình xác suất không trọng số (Unweighted Average)

Trong phương pháp này, đầu ra softmax của mỗi mô hình CNN được xem như một phân bố xác suất độc lập. Xác suất cuối cùng của lớp điều chế thứ j được xác định bằng cách lấy trung bình cộng các xác suất tương ứng từ tất cả các mô hình:

$$p_j = \frac{1}{n} \sum_{i=1}^n y_{ij}, \quad (3.18)$$

trong đó y_{ij} là xác suất dự đoán của lớp j từ mô hình CNN thứ i , và n là số lượng mô hình trong ensemble.

Nhãn phân loại cuối cùng được quyết định bằng cách chọn lớp có xác suất p_j lớn nhất. Phương pháp này đơn giản, ổn định và đặc biệt hiệu quả khi các

mô hình thành phần có hiệu năng tương đối đồng đều [64].

Bỏ phiếu đa số không trọng số (Unweighted Majority Vote)

Khác với phương pháp trung bình xác suất, chiến lược bỏ phiếu đa số chỉ quan tâm đến nhãn dự đoán có xác suất cao nhất của mỗi mô hình CNN. Cụ thể, với mỗi mô hình i , một vector nhị phân $\hat{y}_i$ được tạo ra như sau:

$$\hat{y}_{ij} = \begin{cases} 1, & \text{nếu } y_{ij} = \max(\mathbf{y}_i), \\ 0, & \text{ngược lại.} \end{cases} \quad (3.19)$$

Sau đó, xác suất tổng hợp cho lớp j được xác định bằng cách lấy trung bình các giá trị bỏ phiếu:

$$p_j = \frac{1}{n} \sum_{i=1}^n \hat{y}_{ij}. \quad (3.20)$$

Lớp điều chế có số phiếu cao nhất sẽ được chọn làm kết quả cuối cùng. Phương pháp này có ưu điểm là ít nhạy cảm với các dự đoán xác suất không chắc chắn, nhưng có thể bỏ qua thông tin mức độ tin cậy của từng mô hình.

Trung bình xác suất có trọng số (Weighted Average)

Để tận dụng sự khác biệt về hiệu năng giữa các mô hình CNN thành phần, phương pháp trung bình xác suất có trọng số được áp dụng. Trong đó, mỗi mô hình được gán một hệ số trọng số α_i , phản ánh mức độ đóng góp của mô hình đó vào quyết định cuối cùng, với điều kiện:

$$\sum_{i=1}^n \alpha_i = 1. \quad (3.21)$$

Xác suất tổng hợp cho lớp điều chế thứ j được tính như sau:

$$p_j = \sum_{i=1}^n \alpha_i y_{ij}. \quad (3.22)$$

Cách tiếp cận này cho phép các mô hình có hiệu năng tốt hơn (ví dụ, độ chính xác cao hơn trên tập validation) đóng vai trò chi phối trong quyết định

học kết hợp, từ đó thường đạt được hiệu năng phân loại cao nhất trong thực nghiệm.

3.2.4. Kết quả thực nghiệm và phân tích hiệu năng học kết hợp

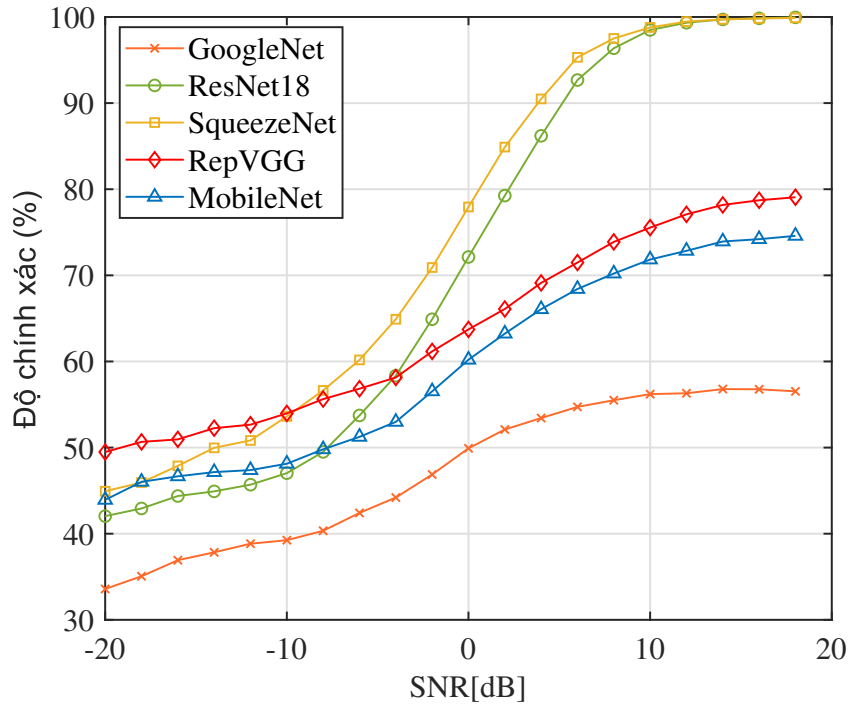
Kết quả đánh giá mô hình học kết hợp và so sánh với các mô hình CNN đơn

Trong phần này, hiệu quả của phương pháp học kết hợp được đánh giá thông qua việc so sánh với các mô hình CNN đơn lẻ phổ biến, bao gồm ResNet18, SqueezeNet, MobileNet, GoogleNet và RepVGG. Các mô hình này được lựa chọn nhằm đại diện cho các kiến trúc CNN có độ sâu, độ phức tạp và chiến lược trích xuất đặc trưng khác nhau, thường được sử dụng trong bài toán phân loại điều chế tự động.

Kết quả đánh giá các mô hình CNN đơn, như minh họa trong Hình 3.8, cho thấy rằng không tồn tại một kiến trúc đơn lẻ nào duy trì hiệu năng vượt trội trên toàn bộ dải SNR từ -20 dB đến $+18$ dB.

Cụ thể, GoogleNet cho kết quả thấp nhất với độ chính xác cực đại chỉ đạt 56.76% tại $\text{SNR} = +18$ dB. MobileNet và RepVGG cải thiện đáng kể hiệu năng so với GoogleNet, với mức tăng tương ứng khoảng 10% và 15% trên toàn dải SNR. Trong khi đó, ResNet18 và SqueezeNet đạt độ chính xác cao ở vùng SNR trung bình và cao, đặc biệt SqueezeNet đạt tới 97.5% với $\text{SNR} \geq +8$ dB. Tuy nhiên, hiệu năng của các mô hình này vẫn suy giảm rõ rệt trong điều kiện SNR thấp.

Xuất phát từ tính bổ sung về khả năng trích xuất đặc trưng của các mô hình CNN đơn, luận án đề xuất áp dụng phương pháp học kết hợp nhằm nâng cao độ chính xác phân loại điều chế, đặc biệt trong các điều kiện nhiễu mạnh. Theo đó, ba mô hình học kết hợp được xây dựng gồm EnCNN5 (kết hợp

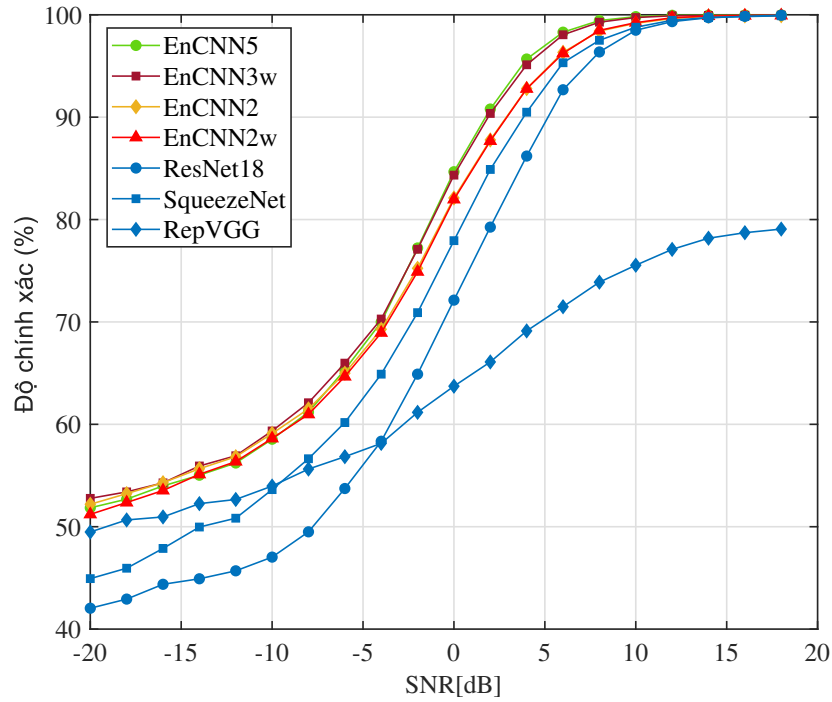


Hình 3.8: So sánh độ chính xác phân loại của các CNN đơn.

bởi ResNet18, SqueezeNet, GoogleNet, MobileNetV2, và RepVGG), EnCNN3 (kết hợp ResNet18, SqueezeNet, và RepVGG) và EnCNN2 (kết hợp RepVGG và SqueezeNet), tương ứng với các tổ hợp CNN có mức độ đa dạng kiến trúc giảm dần. Các mô hình học kết hợp này được triển khai với hai phương pháp học kết hợp: trung bình không trọng số và trung bình có trọng số.

Kết quả so sánh giữa các mô hình học kết hợp và các mô hình CNN đơn được trình bày trong Hình 3.9.

Cụ thể, các mô hình tổ hợp EnCNN5, EnCNN2 (sử dụng trung bình không trọng số), EnCNN3w và EnCNN2w (sử dụng trung bình có trọng số) đều đạt độ chính xác cao hơn so với các mô hình đơn lẻ trong toàn bộ dải SNR từ -20 dB đến $+18$ dB. trên toàn bộ dải SNR khảo sát. Trong vùng SNR thấp (-20 dB đến $+2$ dB), các mô hình học kết hợp cải thiện độ chính xác khoảng 10% so với ResNet18 và khoảng 4% so với RepVGG, cho thấy khả năng chống



Hình 3.9: So sánh độ chính xác giữa các mô hình CNN đơn và các mô hình học kết hợp.

hiệu vượt trội của phương pháp học tổ hợp. Ở vùng SNR trung bình và cao, mức cải thiện tiếp tục được duy trì, với độ chính xác tăng khoảng 20% so với RepVGG và khoảng 8% so với SqueezeNet trong khoảng SNR từ +2 dB đến +8 dB.

Trong số các cấu hình được khảo sát, mô hình EnCNN3 sử dụng chiến lược trung bình có trọng số (EnCNN3w) cho kết quả tốt nhất. Cụ thể, mô hình này đạt độ chính xác 52.2% tại $\text{SNR} = -20$ dB và gần đạt mức phân loại hoàn hảo với độ chính xác 99% tại $\text{SNR} = +18$ dB. Kết quả này cho thấy rằng việc kết hợp các mô hình có thể mạnh bổ sung, cùng với cơ chế gán trọng số phù hợp, giúp tăng đáng kể khả năng tổng quát hóa và độ ổn định của hệ thống AMC.

Nhìn chung, các kết quả thực nghiệm khẳng định rằng phương pháp học

kết hợp là một chiến lược hiệu quả để nâng cao độ chính xác phân loại điều chế, đặc biệt trong điều kiện SNR thấp. Những quan sát này không chỉ làm rõ vai trò của học tổ hợp trong bài toán AMC mà còn tạo tiền đề quan trọng cho việc phát triển các kiến trúc nâng cao hơn trong các nghiên cứu tiếp theo, nơi nhiều nguồn thông tin và nhiều nhiệm vụ được khai thác đồng thời.

Bảng 3.1 trình bày so sánh toàn diện giữa các mô hình CNN đơn và các mô hình học kết hợp xét trên ba khía cạnh: độ chính xác phân loại theo SNR, số lượng tham số và thời gian suy luận.

Đối với các mô hình CNN đơn, có thể nhận thấy rằng SqueezeNet và ResNet18 đạt hiệu năng rất cao trong điều kiện SNR trung bình và cao. Cụ thể, SqueezeNet đạt độ chính xác 99.2% tại $\text{SNR} = +10$ dB trong khi chỉ sử dụng 2.4 triệu tham số, cho thấy khả năng khai thác đặc trưng hiệu quả với độ phức tạp thấp. Ngược lại, RepVGG mặc dù có số tham số lớn nhất (24.8M) nhưng lại cho hiệu năng thấp hơn rõ rệt ở các mức SNR trung bình và cao, phản ánh sự đánh đổi không hiệu quả giữa độ phức tạp và độ chính xác.

Đối với các mô hình học kết hợp, tất cả các cấu hình đều cho thấy sự cải thiện rõ rệt về độ chính xác phân loại, đặc biệt trong điều kiện SNR thấp. Tại $\text{SNR} = -20$ dB, các mô hình EnCNN2, EnCNN3 và EnCNN3w đều đạt độ chính xác trên 52%, cao hơn đáng kể so với mô hình CNN đơn tốt nhất là RepVGG (49.48%).

Trong số các mô hình học kết hợp, EnCNN3w đạt hiệu năng cân bằng nhất giữa độ chính xác và độ phức tạp. Mô hình này đạt độ chính xác 52.76% tại $\text{SNR} = -20$ dB và 99.73% tại $\text{SNR} = +10$ dB, trong khi số tham số và thời gian suy luận không tăng quá nhiều so với EnCNN3. Kết quả này cho thấy

Bảng 3.1: So sánh hiệu năng và độ phức tạp giữa các mô hình CNN đơn và mô hình ensemble

Mô hình	Tham số (M)	Thời gian(ms)	Acc -20 dB (%)	Acc 0 dB (%)	Acc +10 dB (%)
GoogleNet	5.9	1.36	33.58	49.91	56.2
MobileNet	2	2.29	43.93	60.2	71.83
ResNet18	11.1	1.21	42.03	72.13	98.49
SqueezeNet	2.4	1.18	43.49	79.85	99.2
RepVGG	24.8	3.19	49.48	63.72	75.53
EnCNN2	27.2	0.62	52.19	82.1	99.16
EnCNN3	38.3	2.15	52.09	84.78	99.9
EnCNN3w	38.3	2.15	52.76	84.33	99.73
EnCNN5	46.2	3.22	51.84	84.66	99.81

việc kết hợp các mô hình CNN có đặc tính kiến trúc bổ sung bằng cơ chế trung bình có trọng số giúp nâng cao khả năng phân loại, đặc biệt trong điều kiện nhiễu nặng.

3.2.5. Nhận xét, hạn chế và liên hệ

Các kết quả thực nghiệm cho thấy phương pháp học kết hợp CNN là một bước phát triển logic và cần thiết sau các nghiên cứu về tiền xử lý và biểu diễn tín hiệu. Học kết hợp cho phép khai thác hiệu quả tính bổ sung giữa các mô hình và các không gian đặc trưng khác nhau, từ đó nâng cao độ chính xác, độ ổn định và khả năng tổng quát hóa của hệ thống AMC.

Những kết quả này cũng gợi mở hướng nghiên cứu tiếp theo: thay vì chỉ kết hợp ở mức đầu ra, có thể tích hợp thông tin từ nhiều nguồn đặc trưng và nhiều nhiệm vụ ngay trong kiến trúc mạng. Đây chính là động lực trực tiếp cho việc phát triển các mô hình multi-input và multi-output được trình bày trong chương tiếp theo.

Mặc dù phương pháp học kết hợp mang lại cải thiện rõ rệt về hiệu năng và độ ổn định, cách tiếp cận này vẫn tồn tại một số hạn chế. Thứ nhất, chi phí tính toán trong giai đoạn suy luận tăng lên do cần chạy đồng thời nhiều mô hình CNN. Thứ hai, việc kết hợp ở mức quyết định chưa khai thác triệt để mối quan hệ nội tại giữa các đặc trưng học được từ các dạng biểu diễn tín hiệu khác nhau.

Những hạn chế này đặt ra nhu cầu về một kiến trúc học sâu tích hợp hơn, trong đó các nguồn thông tin đa dạng được kết hợp trực tiếp trong quá trình học đặc trưng, thay vì chỉ ở giai đoạn ra quyết định. Đây chính là động lực dẫn tới hướng nghiên cứu tiếp theo của luận án: phát triển các mô hình học

sâu đa đầu vào – đa đầu ra, cho phép khai thác đồng thời nhiều dạng đặc trưng tín hiệu và giải quyết nhiều bài toán liên quan trong cùng một khuôn khổ thống nhất.

3.3. Kết luận chương 3

Chương 3 đã mở rộng hướng nghiên cứu từ khai thác đặc trưng không gian sang mô hình hóa cấu trúc không gian-thời gian của tín hiệu điều chế. Xuất phát từ hạn chế của các mô hình CNN thuần túy trong việc nắm bắt phụ thuộc chuỗi dài hạn, chương 3 đã đề xuất mô hình AMC-CNN-BiLSTM nhằm kết hợp khả năng trích xuất đặc trưng cục bộ của mạng tích chập với năng lực mô hình hóa động học theo thời gian của mạng bộ nhớ dài hạn hai chiều.

Kết quả thực nghiệm cho thấy việc bổ sung thành phần BiLSTM giúp cải thiện đáng kể hiệu năng phân loại, đặc biệt đối với các điều chế bậc cao và trong điều kiện SNR thấp hoặc fading mạnh. Điều này xác nhận rằng thông tin phân lớp không chỉ nằm ở cấu trúc không gian tức thời của tín hiệu IQ mà còn tồn tại trong quan hệ phụ thuộc theo chuỗi giữa các mẫu tín hiệu.

Bên cạnh đó, chương đã đề xuất phương pháp học kết hợp nhằm khai thác tính bổ sung giữa các mô hình có cấu trúc đặc trưng khác nhau. Thay vì phụ thuộc vào một kiến trúc đơn lẻ, phương pháp học kết hợp tận dụng sự đa dạng trong không gian biểu diễn để giảm phương sai dự đoán và tăng độ tin cậy của hệ thống. Kết quả cho thấy mô hình kết hợp đạt độ chính xác cao hơn và ổn định hơn so với từng mô hình thành phần riêng lẻ, đặc biệt trong vùng SNR thấp.

Nếu chương 2 tập trung vào nguyên lý bảo toàn thông tin và thiết kế kiến

trúc trích xuất đặc trưng đa nhánh từ tín hiệu IQ, thì chương 3 đã chứng minh rằng việc mô hình hóa phụ thuộc theo thời gian và khai thác tính bổ sung giữa các mô hình là bước phát triển tiếp theo cần thiết để nâng cao hiệu năng toàn diện của hệ thống AMC. Sự kết hợp giữa đặc trưng không gian, đặc trưng chuỗi thời gian và phương pháp học kết hợp đã hình thành một khuôn khổ học sâu hoàn chỉnh hơn cho bài toán phân loại điều chế trong môi trường vô tuyến thực tế.

Những kết quả đạt được trong chương này góp phần củng cố luận điểm trung tâm của luận án: hiệu năng phân loại điều chế có thể được nâng cao đáng kể khi hệ thống được thiết kế theo hướng khai thác toàn diện cấu trúc thông tin của tín hiệu, bao gồm cả miền không gian, miền thời gian và sự đa dạng mô hình.

Trên nền tảng các kết quả đã đạt được, luận án sẽ tiếp tục phát triển theo hướng mô hình hóa và giải quyết bài toán *Multi-In-Multi-Out*, trong đó mạng nơ-ron được thiết kế để đồng thời thực hiện nhận dạng điều chế và ước lượng tham số không gian (DOA), nhằm tiến gần hơn tới các hệ thống vô tuyến thông minh và tự thích nghi trong thực tế.

Những kết quả nghiên cứu này đã được công bố trong các công trình [C4, J2, C5, J3] thuộc danh mục công bố của luận án.

KẾT LUẬN VÀ HƯỚNG NGHIÊN CỨU

Luận án tập trung nghiên cứu bài toán *phân loại điều chế tự động*, một bài toán then chốt ở lớp vật lý trong các hệ thống vô tuyến thông minh, có vai trò quan trọng trong vô tuyến nhận thức, quản lý phổ tần, thông tin thích ứng và các hệ thống trinh sát vô tuyến. Trong bối cảnh môi trường truyền dẫn ngày càng phức tạp với nhiễu mạnh, fading đa đường và yêu cầu triển khai thời gian thực, các phương pháp AMC truyền thống dựa trên trích chọn đặc trưng thủ công hoặc mô hình xác suất bộc lộ nhiều hạn chế về độ chính xác và khả năng tổng quát hóa.

Xuất phát từ thực tiễn đó, luận án đã tiếp cận bài toán AMC theo hướng học sâu (Deep Learning) và đạt được các kết quả chính sau:

A. NHỮNG ĐÓNG GÓP CỦA LUẬN ÁN

Luận án đưa ra ba đóng góp kỹ thuật chính cho bài toán phân loại điều chế tự động dựa trên học sâu:

1) Đề xuất kiến trúc DFENet cho phân loại điều chế từ tín hiệu IQ nguyên thủy.

Khác với các mô hình CNN tuần tự truyền thống, DFENet sử dụng cơ chế trích xuất đặc trưng đa nhánh nhằm tăng tính đa dạng trong không gian biểu diễn và khai thác đồng thời đặc trưng cục bộ và toàn cục của tín hiệu. Kiến trúc này được thiết kế dựa trên nguyên lý bảo toàn thông tin và giảm thiểu bias tiền xử lý, qua đó cải thiện độ chính xác và tính bền vững trong môi trường SNR thấp.

2) Đề xuất kiến trúc lai CNN–BiLSTM nhằm mô hình hóa đồng thời đặc trưng không gian và phụ thuộc chuỗi theo thời gian của

tín hiệu điều chế.

Mô hình kết hợp tầng tích chập để trích xuất đặc trưng không gian và tầng BiLSTM hai chiều để học phụ thuộc dài hạn theo cả hai hướng thời gian. Cách tiếp cận này cho phép khai thác cấu trúc động của tín hiệu vô tuyến mà các kiến trúc CNN thuần túy chưa xử lý hiệu quả.

3) Đề xuất phương pháp học kết hợp ở tầng đầu ra để kết hợp các mô hình khác biệt về cấu trúc.

Luận án xây dựng cơ chế tổng hợp xác suất dự đoán theo hướng bỏ phiếu mềm có trọng số, nhằm tận dụng tính bổ sung giữa các mô hình chuyên biệt. Phương pháp này cải thiện độ chính xác tổng thể và tăng tính ổn định của hệ thống trong điều kiện fading và nhiễu mạnh.

Các kết quả mô phỏng trên các bộ dữ liệu chuẩn, kết hợp với phân tích định lượng về độ chính xác, độ phức tạp và thời gian suy luận, cho thấy các phương pháp được đề xuất trong luận án đều đạt hiệu năng cạnh tranh hoặc vượt trội so với các mô hình hiện có. Điều này khẳng định tính đúng đắn của hướng tiếp cận và giá trị khoa học cũng như thực tiễn của luận án.

B. GIỚI HẠN NGHIÊN CỨU CỦA LUẬN ÁN

Mặc dù luận án đã đạt được các mục tiêu nghiên cứu đề ra và đưa ra nhiều đóng góp khoa học có giá trị trong lĩnh vực phân loại điều chế tự động sử dụng mạng nơ-ron, vẫn tồn tại một số giới hạn nhất định cần được thừa nhận một cách khách quan như sau.

Thứ nhất, các mô hình đề xuất trong luận án được đánh giá chủ yếu trên hai bộ dữ liệu chuẩn HisarMod2019 và RadioML2018. Mặc dù các bộ dữ liệu này được xây dựng với nhiều điều kiện truyền và fading khác nhau, chúng chưa phản ánh đầy đủ sự đa dạng và phức tạp của môi trường vô tuyến thực

tế. Đặc biệt, các bộ dữ liệu sử dụng trong luận án không cung cấp thông tin hoặc nhãn riêng cho từng mẫu tín hiệu về loại kênh truyền, do đó chưa đủ cơ sở để phân tách dữ liệu và đánh giá định lượng ảnh hưởng của từng loại kênh fading đến hiệu năng phân loại. Vì vậy, luận án chủ yếu đánh giá hiệu năng của các mô hình trên các điều kiện dữ liệu và các mức SNR được cung cấp bởi bộ dữ liệu. Việc sử dụng hoặc xây dựng các bộ dữ liệu có thông tin chi tiết về loại kênh truyền để đánh giá riêng độ ổn định và khả năng thích nghi của mô hình đối với từng điều kiện kênh là một hướng nghiên cứu tiếp theo.

Thứ hai, luận án tập trung nghiên cứu bài toán phân loại điều chế trong bối cảnh tín hiệu đơn người dùng (single-user) và đơn nguồn phát. Các kịch bản phức tạp hơn như môi trường đa người dùng, đa nguồn tín hiệu chồng lấn phổ, hoặc các tình huống tín hiệu bị che khuất mạnh chưa được xem xét một cách toàn diện. Điều này có thể ảnh hưởng đến khả năng mở rộng trực tiếp của các mô hình đề xuất trong các hệ thống vô tuyến mật độ cao.

Thứ ba, mặc dù luận án đã phân tích và so sánh độ phức tạp mô hình thông qua số lượng tham số và thời gian suy luận, việc triển khai và đánh giá trực tiếp các mô hình trên phần cứng nhúng hoặc nền tảng vô tuyến thực (SDR, FPGA, hoặc SoC) chưa được thực hiện. Do đó, các kết luận về tính khả thi thời gian thực mới chỉ dừng ở mức đánh giá gián tiếp thông qua mô phỏng.

Thứ tư, phạm vi nghiên cứu của luận án chủ yếu tập trung vào các dạng tín hiệu điều chế có đặc tính tương đối ổn định trong khoảng thời gian quan sát. Các bộ dữ liệu được sử dụng chưa bao gồm các tín hiệu nhảy tần nhanh (Frequency Hopping) thường xuất hiện trong các hệ thống thông tin quân sự

hiện đại. Đối với loại tín hiệu này, thời gian lưu trú trên mỗi tần số thường rất ngắn, làm giảm thời gian quan sát và gây khó khăn cho quá trình trích xuất đặc trưng. Do chưa có bộ dữ liệu công khai phù hợp cũng như điều kiện thu thập dữ liệu thực nghiệm, khả năng áp dụng các mô hình đề xuất đối với tín hiệu nhảy tần nhanh chưa được đánh giá trong phạm vi của luận án.

Cuối cùng, các kỹ thuật học nâng cao như học liên kết, học tự giám sát, hoặc học thích nghi trực tuyến mới chỉ được đề cập ở mức định hướng, chưa được triển khai và đánh giá chi tiết trong phạm vi luận án này.

Những giới hạn nêu trên không làm giảm giá trị khoa học của các kết quả đạt được, mà đồng thời mở ra các hướng nghiên cứu tiếp theo đầy tiềm năng nhằm mở rộng và hoàn thiện hơn nữa các mô hình nhận dạng tín hiệu vô tuyến dựa trên trí tuệ nhân tạo.

C. HƯỚNG NGHIÊN CỨU TIẾP THEO

Mặc dù đã đạt được các kết quả quan trọng, luận án vẫn mở ra nhiều hướng nghiên cứu tiềm năng có thể tiếp tục phát triển trong tương lai:

- **Mở rộng sang môi trường thực tế:** Nghiên cứu và đánh giá các mô hình AMC trên dữ liệu thu thực (over-the-air) hoặc thông qua các bộ giả lập kênh chuẩn hóa như kênh 3GPP Urban Macro/Urban Micro nhằm thu hẹp khoảng cách giữa mô phỏng và triển khai thực tế. Các chỉ tiêu liên quan đến quá trình tối ưu hóa như thời gian huấn luyện, tốc độ hội tụ và độ nhạy đối với các siêu tham số chưa được khảo sát một cách hệ thống. Đây sẽ là hướng nghiên cứu tiếp theo nhằm hoàn thiện việc đánh giá đặc tính học và tối ưu hóa của các mô hình đề xuất.
- **AMC thích ứng theo SNR và kênh truyền:** Một hạn chế của các

bộ dữ liệu chuẩn được sử dụng trong luận án là các tham số vật lý của tín hiệu, như tốc độ ký hiệu, băng thông và hệ số roll-off, được thiết lập cố định trong quá trình mô phỏng và không được cung cấp dưới dạng nhãn cho từng mẫu tín hiệu. Do đó, luận án chưa thể khảo sát riêng ảnh hưởng của từng tham số đến hiệu năng phân loại. Trong thời gian tới, hướng nghiên cứu xây dựng hoặc sử dụng các bộ dữ liệu có khả năng thay đổi và gán nhãn các tham số vật lý này để đánh giá sâu hơn khả năng thích nghi và tính tổng quát của các mô hình đề xuất. Phát triển các mô hình AMC có khả năng điều kiện hóa theo SNR hoặc trạng thái kênh (SNR-conditioned hoặc channel-aware AMC) nhằm nâng cao khả năng thích ứng trong môi trường vô tuyến biến thiên.

- **Học bán giám sát và học tự giám sát:** Khai thác các phương pháp học bán giám sát (semi-supervised learning) hoặc tự giám sát (self-supervised learning) để giảm phụ thuộc vào dữ liệu gán nhãn, vốn tốn kém và khó thu thập trong thực tế.
- **Tối ưu hóa cho thiết bị biên:** Nghiên cứu các kỹ thuật nén mô hình, lượng tử hóa và triển khai trên phần cứng nhúng hoặc FPGA/SoC nhằm đáp ứng yêu cầu thời gian thực và tiêu thụ năng lượng thấp.
- **Mở rộng bài toán đa nhiệm:** Phát triển các mô hình đa nhiệm mở rộng hơn, kết hợp AMC với các bài toán khác ở lớp vật lý như phát hiện tín hiệu, nhận dạng giao thức, hoặc ước lượng tham số kênh. Một hạn chế của các bộ dữ liệu chuẩn được sử dụng trong luận án là các tham số vật lý của tín hiệu, như tốc độ ký hiệu, băng thông và hệ số roll-off, được thiết lập cố định trong quá trình mô phỏng và không được cung cấp

dưới dạng nhãn cho từng mẫu tín hiệu. Do đó, luận án chưa thể khảo sát riêng ảnh hưởng của từng tham số đến hiệu năng phân loại. Trong thời gian tới, hướng nghiên cứu xây dựng hoặc sử dụng các bộ dữ liệu có khả năng thay đổi và gán nhãn các tham số vật lý này để đánh giá sâu hơn khả năng thích nghi và tính tổng quát của các mô hình đề xuất.

Với các kết quả đạt được và các hướng phát triển nêu trên, luận án không chỉ đóng góp về mặt học thuật cho lĩnh vực xử lý tín hiệu và học sâu trong thông tin vô tuyến, mà còn đặt nền tảng cho các nghiên cứu ứng dụng và triển khai thực tế trong các hệ thống vô tuyến thông minh thế hệ mới.

DANH MỤC CÁC CÔNG TRÌNH ĐÃ CÔNG BỐ

A. Các công trình sử dụng trong luận án

- C1.** “Tự động nhận dạng điều chế dưới tác động của nhiễu pha-đỉnh đa đường sử dụng kỹ thuật học sâu”, Hội nghị Khoa học Quốc Gia REV-ECIT 2021 (10/2021).
- C2.** "Performance Analysis of Convolutional Neural Networks with Different Window Functions for Automatic Modulation Classification," 2022 13th International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea, Republic of, 2022, pp. 153-157, doi: 10.1109/ICTC55196.2022.9952750.
- C3.** "Improving Automatic Modulation Classification Accuracy Using Combination of Short-Time Fourier Transform and Radon Transform," 2024 International Conference on Advanced Technologies for Communications (ATC), Ho Chi Minh City, pp. 493-498, doi: 10.1109/ATC63255.2024.10908305.
- J1.** " DFENet: A diverse feature extraction neural network for improving automatic modulation classification accuracy in wireless communication systems." PLoS One. 2026 Jan 21;doi: 10.1371/journal.pone.0341020.
- C4.** "AMC-CNN-BiLSTM: An Advanced Hybrid Model for Automatic Modulation Classification of Radio Signals," 2024 RIVF International Conference on Computing and Communication Technologies (RIVF), Danang, Vietnam, 2024, pp. 271-275, doi: 10.1109/RIVF64335.2024.11009112.

- J2.** “Ensemble of Convolution Neural Networks for Improving Automatic Modulation Classification Performance”. Tạp Chí Khoa học Và Công nghệ - Đại học Đà Nẵng, vol 20, số p.h 6.2, Tháng Sáu 2022, tr 25-32, doi:10.31130/ud-jst.2022.293E.

B. Các công trình liên quan đến luận án

- C5.** “Nâng cao hiệu quả phân loại điều chế tự động sử dụng mạng CNN đa đầu vào”, Hội nghị Khoa học Quốc Gia REV-ECIT 2022 (12/2022).
- J3.** "Multi-In-Multi-Out Neural Network for Joint DOA Estimation and Automatic Modulation Classification," in IEEE Communications Letters, vol. 29, no. 8, pp. 1993-1997, Aug. 2025, doi: 10.1109/LCOMM.2025.3583717.

TÀI LIỆU THAM KHẢO

- [1] D. H. Al-Nuaimi, I. A. Hashim, I. S. Zainal Abidin, L. B. Salman, and N. A. Mat Isa, “Performance of feature-based techniques for automatic digital modulation recognition and classification—a review,” *Electronics*, vol. 8, p. 1407, 2019.
- [2] W. Chen, Z. Xie, L. Ma, J. Liu, and X. Liang, “A faster maximum-likelihood modulation classification in flat fading non-gaussian channels,” *IEEE Communications Letters*, vol. 23, no. 3, pp. 454–457, 2019.
- [3] J. Ma, S.-C. Lin, H. Gao, and T. Qiu, “Likelihood-ratio approaches to automatic modulation classification,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, vol. 41, no. 4, pp. 455–469, 2011.
- [4] F. A. Bhatti, M. J. Khan, A. Selim, and F. Paisana, “Shared spectrum monitoring using deep learning,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, pp. 1171–1185, 2021.
- [5] Z. Elkhatab, F. Kamalov, S. Moussa, A. B. Mnaouer, M. C. E. Yagoub, and H. Yanikomeroglu, “Radio modulation classification optimization using combinatorial deep learning technique,” *IEEE Access*, vol. 12, pp. 17 552–17 570, 2024.
- [6] R. Zhou, F. Liu, and C. W. Gravelle, “Deep learning for modulation recognition: A survey with a demonstration,” *IEEE Access*, vol. 8, pp. 67 366–67 376, 2020.
- [7] O. Kaya, M. A. Karabulut, A. F. M. Shahen Shah, and H. Ilhan, “Modulation classifier based on deep learning for beyond 5g communications,” in *47th International Conference on Telecommunications and Signal Processing (TSP)*, 2024, pp. 336–339.
- [8] T. J. O’Shea and N. West, “Radio machine learning dataset generation with gnu radio,” in *Proceedings of the GNU Radio Conference*, vol. 1, 2016.
- [9] T. Huynh-The *et al.*, “Automatic modulation classification: A deep architecture survey,” *IEEE Access*, vol. 9, pp. 142 950–142 971, 2021.

- [10] Z. Zhu and A. K. Nandi, *Automatic Modulation Classification: Principles, Algorithms and Applications*. John Wiley & Sons, 2015.
- [11] T. J. O'Shea, J. Corgan, and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 168–179, 2018.
- [12] —, "Convolutional radio modulation recognition networks," in *Communications in Computer and Information Science*, vol. 629, 2016, pp. 213–226.
- [13] J. Ma, S.-C. Lin, H. Gao, and T. Qiu, "Automatic modulation classification under non-gaussian noise: A deep residual learning approach," in *IEEE International Conference on Communications (ICC)*, 2019, pp. 1–6.
- [14] G. Huang, Z. Liu, and K. Q. Weinberger, "Densely connected convolutional networks," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2261–2269.
- [15] T. Huynh-The, C.-H. Hua, Q.-V. Pham, and D.-S. Kim, "Mcnet: An efficient cnn architecture for robust automatic modulation classification," *IEEE Communications Letters*, vol. 24, no. 4, pp. 811–815, 2020.
- [16] T. Huynh-The, C. H. Hua, V. S. Doan, and D. S. Kim, "Accurate modulation classification with reusable-feature convolutional neural network," in *IEEE Eighth International Conference on Communications and Electronics (ICCE)*, 2021, pp. 12–17.
- [17] V. Shatov, N. Shanin, B. Perner, Y. Su, N. Franchi, and M. Lübke, "Lstm-dnn-based joint wireless signal waveform classification and blind transmitter localization," in *International Symposium on Wireless Communication Systems (ISWCS)*, 2024, pp. 1–6.
- [18] C. Guo, S. Han, W. Meng, and C. Li, "Abconv: Attention based convolution for automatic modulation recognition," in *IEEE International Conference on Communications Workshops (ICC Workshops)*, 2024, pp. 1586–1591.
- [19] V.-S. Doan, T. Huynh-The, V.-P. Hoang, and D.-T. Nguyen, "Modanet: Multi-task deep network for joint automatic modulation classification and direction of arrival estimation," *IEEE Communications Letters*, vol. 26, no. 2, pp. 335–339, 2022.
- [20] J. H. Lee, K.-Y. Kim, and Y. Shin, "Feature image-based automatic modulation classification method using cnn algorithm," in *International Conference*

- on Artificial Intelligence in Information and Communication (ICAIIC)*, 2019, pp. 1–4.
- [21] J. S. Colburn and Y. Rahmat-Samii, “Patch antennas on externally perforated high dielectric constant substrates,” *IEEE Transactions on Antennas and Propagation*, vol. 47, no. 12, pp. 1785–1794, 1999.
 - [22] H. Xing, X. Zhang, S. Chang, J. Ren, Z. Zhang, J. Xu, and S. Cui, “Joint signal detection and automatic modulation classification via deep learning,” *IEEE Transactions on Wireless Communications*, 2024.
 - [23] A. Jagannath and J. Jagannath, “Multi-task learning approach for automatic modulation and wireless signal classification,” in *IEEE International Conference on Communications (ICC)*, 2021, pp. 1–7.
 - [24] A. P. Hermawan, R. R. Ginanjar, D.-S. Kim, and J.-M. Lee, “Cnn-based automatic modulation classification for beyond 5g communications,” *IEEE Commun. Lett.*, vol. 24, no. 5, pp. 1038–1041, 2020.
 - [25] S. Haykin and P. Setoodeh, “Cognitive radio networks: The spectrum supply chain paradigm,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 1, no. 1, pp. 3–28, 2015.
 - [26] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, *Discrete-Time Signal Processing*. Prentice Hall, 1999.
 - [27] J. Kaiser and R. Schaffer, “On the use of the $i0$ -sinh window for spectrum analysis,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 28, no. 1, pp. 105–107, 1980.
 - [28] Sulistyaningsih *et al.*, “Performance comparison of blackman, bartlett, hanning, and kaiser window for radar digital signal processing,” in *International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, 2019, pp. 391–394.
 - [29] B. Xiong, Y. Zhong, W. Pu, and F. Zhang, “Automatic modulation classification based on wavelet image and convolutional neural network,” in *International Conference on Computer and Communication Systems (ICCCS)*, 2023, pp. 967–971.
 - [30] C. Liu, Z. Cai, and B. Zhang, “Automatic modulation classification based on complex-valued convolutional neural network and semi-supervised learn-

- ing,” in *IEEE International Conference on Communication Technology (ICCT)*, 2023, pp. 241–246.
- [31] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
 - [32] S. Peng, H. Jiang, H. Wang, H. Alwageed, Y. Zhou *et al.*, “Modulation classification based on signal constellation diagrams and deep learning,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, pp. 1–10, 2019.
 - [33] K. Tekbıyık, A. R. Ekti, A. Görçin, G. K. Kurt, and C. Keçeci, “Robust and fast automatic modulation classification with cnn under multipath fading channels,” in *IEEE 91st Vehicular Technology Conference (VTC Spring)*, 2020, pp. 1–6.
 - [34] H. Wang and Y.-D. Yao, “Primary user boundary detection in cognitive radio networks: Estimated secondary user locations and impact of malicious secondary users,” *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4577–4588, 2018.
 - [35] M. Alymani, M. H. Alhazmi, A. Almarhabi, H. Alhazmi, A. Samarkandi, and Y.-D. Yao, “Rician k-factor estimation using deep learning,” in *Proc. 29th Wireless Opt. Commun. Conf. (WOCC)*, 2020, pp. 1–4.
 - [36] X. Qin, S. Peng, X. Yang, and Y.-D. Yao, “Deep learning based channel code recognition using textcnn,” in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN)*, 2019, pp. 1–5.
 - [37] Y. Zhou, S. Peng, and Y. Yao, “Mac protocol identification using convolutional neural networks,” in *Proc. 29th Wireless Opt. Commun. Conf. (WOCC)*, 2020, pp. 1–4.
 - [38] K. Alshathri, H. Xia, V. Lawrence, and Y.-D. Yao, “Cellular system identification using deep learning: Gsm, umts and lte,” in *Proc. 28th Wireless Opt. Commun. Conf. (WOCC)*, 2019, pp. 1–4.
 - [39] M. H. Alhazmi, M. Alymani, H. Alhazmi, A. Almarhabi, A. Samarkandi, and Y.-D. Yao, “5g signal identification using deep learning,” in *Proc. 29th Wireless Opt. Commun. Conf. (WOCC)*, 2020, pp. 1–5.

- [40] H.-C. Wu, M. Saquib, and Z. Yun, “Novel automatic modulation classification using cumulant features for communications via multipath channels,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3098–3105, 2008.
- [41] L. Xie and Q. Wan, “Automatic modulation recognition for phase shift keying signals with compressive measurements,” *IEEE Wireless Commun. Lett.*, vol. 7, no. 2, pp. 194–197, 2018.
- [42] Z. Wang, J. Peng, W. Song, X. Gao, and L. Ma, “A convolutional neural network-based classification and decision-making model for visible defect identification of high-speed train images,” *Journal of Sensors*, vol. 2021, p. 17, 2021.
- [43] L. Liu and X. Li, “Radar signal recognition based on triplet convolutional neural network,” *EURASIP Journal on Advances in Signal Processing*, vol. 2021, no. 1, p. 16, 2021.
- [44] “Signal automatic modulation based on amc neural network fusion,” *PMC PubMed Central*, 2024, pMCID: PMC11156377.
- [45] H. Huang, J. Yang, H. Huang, and Y. Liu, “Automatic modulation classification using hybrid data augmentation and lightweight neural network,” *Sensors*, vol. 23, no. 9, p. 4187, 2023.
- [46] T. J. O’Shea, T. Roy, and T. C. Clancy, “Over-the-air deep learning based radio signal classification,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 168–179, 2018.
- [47] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, “Deep learning models for wireless signal classification with distributed low-cost spectrum sensors,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 3, pp. 433–445, 2018.
- [48] Y. Zhang, Y. Wang, and F. Wang, “Automatic modulation classification based on deep learning: A survey,” *IEEE Access*, vol. 8, pp. 117 799–117 819, 2020.
- [49] S. Huang, Y. Yang, and P. Zhang, “Deep learning for automatic modulation classification: A survey,” *IEEE Access*, vol. 9, pp. 123 544–123 567, 2021.
- [50] S. Kharbech, I. Dayoub, M. Zwingelstein-Colin, and E. P. Simon, “On classifiers for blind feature-based automatic modulation classification over

- multiple-input-multiple-output channels,” *IET Commun.*, vol. 10, no. 7, pp. 790–795, 2016.
- [51] W. Xie, S. Hu, C. Yu, P. Zhu, X. Peng, and J. Ouyang, “Deep learning in digital modulation recognition using high order cumulants,” *IEEE Access*, vol. 7, pp. 63 760–63 766, 2019.
 - [52] S. H. Lee, K.-Y. Kim, J. H. Kim, and Y. Shin, “Effective feature-based automatic modulation classification method using dnn algorithm,” in *Proc. Int. Conf. Artif. Intell. Inf. Commun. (ICAIIC)*, 2019, pp. 557–559.
 - [53] Y. Tu and Y. Lin, “Deep neural network compression technique towards efficient digital signal modulation recognition in edge device,” *IEEE Access*, vol. 7, pp. 58 113–58 119, 2019.
 - [54] G. J. Mendis, J. Wei-Kocsis, and A. Madanayake, “Deep learning based radio-signal identification with hardware design,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 5, pp. 2516–2531, 2019.
 - [55] Y. Li, G. Shao, and B. Wang, “Automatic modulation classification based on bispectrum and cnn,” in *Proc. IEEE 8th Joint Int. Inf. Technol. Artif. Intell. Conf. (ITAIC)*, 2019, pp. 311–316.
 - [56] J. Shi, S. Hong, C. Cai, Y. Wang, H. Huang, and G. Gui, “Deep learning-based automatic modulation recognition method in the presence of phase offset,” *IEEE Access*, vol. 8, pp. 42 841–42 847, 2020.
 - [57] S. Zheng, P. Qi, S. Chen, and X. Yang, “Fusion methods for cnn-based automatic modulation classification,” *IEEE Access*, vol. 7, pp. 66 496–66 504, 2019.
 - [58] S. Hu, Y. Pei, P. P. Liang, and Y.-C. Liang, “Deep neural network for robust modulation classification under uncertain noise conditions,” *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 564–577, 2020.
 - [59] H. Huang, J. Yang, H. Huang, and Y. Liu, “Automatic modulation classification using hybrid data augmentation and lightweight neural network,” *Sensors*, vol. 23, no. 9, p. 4187, 2023.
 - [60] B. Tang, Y. Tu, Z. Zhang, and Y. Lin, “Digital signal modulation classification with data augmentation using generative adversarial nets in cognitive radio networks,” *IEEE Access*, vol. 6, pp. 15 713–15 722, 2018.

- [61] J. Xu, C. Luo, G. Parr, and Y. Luo, “A heterogeneous ensemble learning framework for automatic modulation classification,” *IEEE Access*, vol. 9, pp. 12 345–12 355, 2021.
- [62] O. A. Topal, S. K. Efe, and G. K. Kurt, “Ensemble learning-based automatic modulation classification for under-water acoustic communication,” *IEEE Access*, vol. 11, pp. 2534–2545, 2023.
- [63] J. N. Njoku, M. F. Hossain, and S. Nahavandi, “Cgdnet: Efficient hybrid deep learning model for robust automatic modulation recognition,” *IEEE Networking Letters*, vol. 3, no. 1, pp. 41–45, 2021.
- [64] Y. Wang, J. Liu, and X. Zhang, “A hybrid cnn-transformer framework for automatic modulation classification,” *IEEE Transactions on Vehicular Technology*, vol. 72, no. 4, pp. 5678–5689, 2023.
- [65] L. Zhang, H. Zhao, and X. Hou, “Multi-scale feature fusion network for modulation classification in complex electromagnetic environments,” *Remote Sensing*, vol. 14, no. 15, p. 3678, 2022.
- [66] K. Liu and Y. Chen, “Automatic modulation classification based on deep learning and xgboost,” in *2022 14th International Conference on Wireless Communications and Signal Processing (WCSP)*. IEEE, 2022.
- [67] T. Hu and Y. Wang, “A two-stage automatic modulation classification method based on lightgbm and deep features,” *IEEE Access*, vol. 9, pp. 89 012–89 020, 2021.
- [68] J. Norolahi, M. Mehrnia, and P. Azmi, “Blind modulation classification via combined machine learning and signal feature extraction,” *Proceedings of ISMODE*, pp. 266–271, 2022.
- [69] S. Peng, S. Sun, and Y.-D. Yao, “A survey of modulation classification using deep learning: Signal representation and data preprocessing,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 12, pp. 7020–7038, 2022.
- [70] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, “Deep learning models for wireless signal classification with distributed low-cost spectrum sensors,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 3, pp. 433–445, 2018.

- [71] —, “Deep learning models for wireless signal classification with distributed low-cost spectrum sensors,” *IEEE Trans. Cognit. Commun. Netw.*, vol. 4, no. 3, pp. 433–445, 2018.
- [72] O. S. Mossad, M. ElNainay, and M. Torki, “Deep convolutional neural network with multi-task learning scheme for modulations recognition,” in *Proc. 15th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, 2019, pp. 1644–1649.
- [73] H. Wu, Y. Li, L. Zhou, and J. Meng, “Convolutional neural network and multi-feature fusion for automatic modulation classification,” *Electron. Lett.*, vol. 55, no. 16, pp. 895–897, 2019.
- [74] D. Li, Z. Xu, L. Gu, and L. Zhu, “Recognition method of radar intra-pulse modulation type based on signal square spectrum bandwidth ratio,” *EURASIP Journal on Advances in Signal Processing*, vol. 2021, no. 1, p. 16, 2021.
- [75] J. A. Snoap, D. C. Popescu, J. A. Latshaw, and C. M. Spooner, “Deep-learning-based classification of digitally modulated signals using capsule networks and cyclic cumulants,” *Sensors*, vol. 23, no. 12, p. 5735, 2023.
- [76] Y. Wang *et al.*, “Moe-amc: Enhancing automatic modulation classification performance using mixture-of-experts,” 2023.
- [77] E. L. Ferguson, “Multitask convolutional neural network for acoustic localization of a transiting broadband source using a hydrophone array,” *The Journal of the Acoustical Society of America*, vol. 150, no. 1, pp. 248–256, 2021.
- [78] D. K. Jain, Z. Zhang, and K. Huang, “Multi angle optimal pattern-based deep learning for automatic facial expression recognition sciencedirect,” *Pattern Recognition Letters*, vol. 139, pp. 157–165, 2020.
- [79] L. Shen, Z. Shao, Y. Yu, and X. M. Chen, “Hybrid approach combining modified gravity model and deep learning for short-term forecasting of metro transit passenger flows,” *Transportation Research Record*, vol. 2675, no. 1, pp. 25–38, 2021.
- [80] P. Ajay, B. Nagaraj, B. M. Pillai, J. Suthakorn, and M. Bradha, “Intelligent ecofriendly transport management system based on iot in urban areas,” *Environment Development and Sustainability*, vol. 3, pp. 1–8, 2022.

- [81] P. Ajay, B. Nagaraj, R. A. Kumar, R. Huang, and P. Ananthi, “Unsupervised hyperspectral microscopic image segmentation using deep embedded clustering algorithm,” *Scanning*, vol. 2022, p. 9, 2022.
- [82] Q. Zeng, F. Xiao, Q. Ren, and S. Ai, “Modulation signal analysis and parameter selection for modular multi-level converters,” *Journal of Power Electronics*, vol. 20, no. 2, pp. 455–465, 2020.
- [83] S. Peng, S. Sun, and Y.-D. Yao, “A survey of modulation classification using deep learning: Signal representation and data preprocessing,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 12, pp. 7020–7038, 2022.
- [84] D. Li, Z. Xu, L. Gu, and L. Zhu, “Recognition method of radar intra-pulse modulation type based on signal square spectrum bandwidth ratio,” *EURASIP Journal on Advances in Signal Processing*, vol. 2021, no. 1, p. 16, 2021.
- [85] J. A. Snoap, D. C. Popescu, J. A. Latshaw, and C. M. Spooner, “Deep-learning-based classification of digitally modulated signals using capsule networks and cyclic cumulants,” *Sensors*, vol. 23, no. 12, p. 5735, 2023.
- [86] A. M. Abdulkarem *et al.*, “Robust automatic modulation classification using convolutional deep neural network based on scalogram information,” *Computers*, vol. 11, no. 11, p. 162, 2022.
- [87] J. Shi, S. Hong, C. Cai, and Y. Wang, “Deep learning-based automatic modulation recognition method in the presence of phase offset,” *IEEE Access*, vol. 8, pp. 42 841–42 847, 2020.
- [88] K. Yashashwi, A. Sethi, and P. Chaporkar, “A learnable distortion correction module for modulation recognition,” *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 77–80, 2019.
- [89] Z. Zhang, C. Wang, C. Gan, S. Sun, and M. Wang, “Automatic modulation classification using convolutional neural network with features fusion of spwvd and bjd,” *IEEE Transactions on Signal and Information Processing over Networks*, vol. 5, no. 3, pp. 469–478, 2019.
- [90] Y. W. H. G. X. Fu, G. Gui and F. Adachi, “Automatic modulation classification based on decentralized learning and ensemble learning,” *IEEE Transactions on Vehicular Technology*, pp. 7942–7946, 2022.

- [91] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” pp. 770–778, 2015.
- [92] T. Huynh-The, Q.-V. Pham, T.-V. Nguyen, T. T. Nguyen, D. B. d. Costa, and D.-S. Kim, “Rannet: Learning residual-attention structure in cnns for automatic modulation classification,” *IEEE Wireless Communications Letters*, vol. 11, no. 6, pp. 1243–1247, Jun. 2022.
- [93] V.-S. Doan, T. Huynh-The, and D.-S. Kim, “Underwater acoustic target classification based on dense convolutional neural network,” *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.
- [94] H. Le, V.-S. Doan, D. P. Le, H.-H. Nguyen, T. Huynh-The, K. Le-Ha, and V.-P. Hoang, “Micro-doppler-radar-based uav detection using inception-residual neural network,” in *2020 International Conference on Advanced Technologies for Communications (ATC)*. IEEE, Oct. 2020, pp. 177–181.