

**ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ**



Ninh Hải Đăng

**ỨNG DỤNG VIỄN THÁM VÀ HỌC SÂU TRONG
GIÁM SÁT BIỂN ĐỘNG RỪNG TỈNH CÀ MAU**

ĐỒ ÁN TỐT NGHIỆP ĐẠI HỌC HỆ CHÍNH QUY

Ngành: Công nghệ Hàng không Vũ trụ

HÀ NỘI - 2025

**ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ**



Ninh Hải Đăng

**ỨNG DỤNG VIỄN THÁM VÀ HỌC SÂU TRONG GIÁM
SÁT BIỂN ĐỘNG RỪNG TỈNH CÀ MAU**

ĐỒ ÁN TỐT NGHIỆP ĐẠI HỌC HỆ CHÍNH QUY

Ngành: Công nghệ Hàng không Vũ trụ

Cán bộ hướng dẫn: TS. Hà Minh Cường

Cán bộ đồng hướng dẫn: ThS. Hoàng Tích Phúc

HÀ NỘI - 2025

TÓM TẮT

Tóm tắt: Đồ án này nghiên cứu ứng dụng mạng Neural Tích chập (CNN) kết hợp với dữ liệu viễn thám đa nguồn từ vệ tinh Sentinel-1 và Sentinel-2 để phát hiện và phân loại biến động rừng tại tỉnh Cà Mau — khu vực rừng ngập mặn quan trọng bậc nhất tại Việt Nam. Nghiên cứu sử dụng 27 đặc trưng bao gồm các kênh phổ, chỉ số thực vật (NDVI, NBR, NDMI) và dữ liệu SAR từ hai thời kỳ (01/2024 và 02/2025) với độ phân giải 10m.

Kiến trúc CNN nhẹ với khoảng 36,000 tham số được thiết kế, sử dụng patches 3×3 để khai thác ngữ cảnh không gian, áp dụng các kỹ thuật regularization (Batch Normalization, Dropout, Weight Decay) để tránh overfitting trên bộ dữ liệu thực địa gồm 2,630 điểm với 4 lớp phân loại: Rừng ổn định, Mất rừng, Phi rừng và Phục hồi rừng. Quy trình đánh giá triển khai với stratified random split kết hợp 5-Fold Cross Validation, đảm bảo phân bố cân bằng các lớp và giữ lại 20% dữ liệu làm tập test cố định.

Kết quả cho thấy mô hình đạt độ chính xác 98.86% trên tập test với ROC-AUC 99.98%, Cross Validation cho kết quả trung bình $98.15\% \pm 0.28\%$ chứng tỏ mô hình ổn định và có khả năng tổng quát hóa tốt. Ứng dụng thực tế phân loại toàn vùng quy hoạch lâm nghiệp Cà Mau (170,179 ha ranh giới, 162,469 ha phân loại thực tế), phát hiện 7,282 ha mất rừng (4.48%) và 4,941 ha phục hồi rừng (3.04%) trong giai đoạn nghiên cứu, đóng góp phương pháp hiệu quả cho công tác giám sát tài nguyên rừng.

Từ khóa: CNN, viễn thám, biến động rừng, Cà Mau.

LỜI CAM ĐOAN

Tôi tên là Ninh Hải Đăng, sinh viên lớp QH-2021-I/CQ-S-AE, Viện Công nghệ Hàng không Vũ trụ, Trường Đại học Công nghệ – Đại học Quốc gia Hà Nội. Tôi xin cam đoan rằng Đồ án tốt nghiệp với đề tài “Ứng dụng viễn thám và học sâu trong giám sát biến động rừng tỉnh Cà Mau” là kết quả nghiên cứu khoa học do chính tôi thực hiện dưới sự hướng dẫn của TS. Hà Minh Cường và ThS. Hoàng Tích Phúc.

Mọi sự hỗ trợ, hướng dẫn trong quá trình thực hiện đều đã được ghi nhận và cảm ơn; các thông tin, số liệu, tài liệu tham khảo trong đồ án đều được trích dẫn đầy đủ và được phép sử dụng. Tôi xin đảm bảo rằng tất cả dữ liệu nghiên cứu và kết quả trình bày trong đồ án là trung thực và chính xác. Nếu phát hiện bất kỳ sai sót nào, tôi xin hoàn toàn chịu trách nhiệm trước cơ quan nhà trường và các cơ quan liên quan.

Sinh viên

Ninh Hải Đăng

LỜI CẢM ƠN

Đối với mỗi sinh viên, đồ án tốt nghiệp là một cột mốc quan trọng, phản ánh kết quả học tập và quá trình rèn luyện suốt thời gian tại trường đại học. Trong quá trình thực hiện đồ án, bên cạnh những nỗ lực của bản thân, em đã nhận được rất nhiều sự giúp đỡ quý báu từ các thầy cô và những người xung quanh, nhờ đó công trình này mới có thể hoàn thành.

Trước hết, em xin bày tỏ lòng biết ơn sâu sắc tới các thầy, cô là cán bộ, giảng viên trong trường Đại học Công nghệ nói chung và Viện Công nghệ Hàng không Vũ trụ nói riêng, những người đã tạo điều kiện và truyền đạt kiến thức quý báu trong suốt thời gian học tập.

Đặc biệt, em xin gửi lời cảm ơn chân thành tới TS. Hà Minh Cường và ThS. Hoàng Tích Phúc – giảng viên Viện Công nghệ Hàng không Vũ trụ, trường Đại học Công nghệ - Đại học Quốc gia Hà Nội, những người đã tận tình hướng dẫn, chỉ bảo và đồng hành cùng em trong suốt quá trình thực hiện đồ án.

Em cũng xin trân trọng cảm ơn TS. Hoàng Việt Anh và ThS. Vũ Văn Thái – Công ty TNHH Tư vấn và Phát triển Đồng Xanh (GFD), đã hỗ trợ em trong việc cung cấp dữ liệu, cơ sở vật chất và tạo điều kiện thuận lợi để hoàn thành nghiên cứu.

Cuối cùng, em xin gửi lời cảm ơn sâu sắc tới gia đình, bạn bè và các đồng nghiệp tại công ty đã luôn động viên, chia sẻ và đồng hành cùng em trong suốt quá trình học tập và thực hiện đồ án.

Mặc dù đã rất cố gắng, song do những hạn chế về mặt thời gian và kinh nghiệm, đồ án tốt nghiệp này chắc chắn sẽ không tránh khỏi những thiếu sót. Em kính mong nhận được những ý kiến đóng góp quý báu từ quý thầy cô để bản thân có thể hoàn thiện hơn nữa trong tương lai.

Em xin trân trọng cảm ơn!

MỤC LỤC

DANH MỤC HÌNH ẢNH

DANH MỤC BẢNG

DANH MỤC THUẬT NGỮ VÀ TỪ VIẾT TẮT

Từ	Viết tắt	Dịch - Nghĩa
Ablation Study	-	Nghiên cứu loại trừ (phân tích ảnh hưởng từng thành phần)
Accuracy	-	Độ chính xác tổng thể (tỷ lệ dự đoán đúng trên tổng số mẫu)
Activation Function	-	Hàm kích hoạt (hàm phi tuyến trong mạng neural)
AdamW	-	Thuật toán tối ưu Adam với phân rã trọng số tách rời
Artificial Intelligence	AI	Trí tuệ nhân tạo
Backscatter	-	Tán xạ ngược (năng lượng radar phản hồi về cảm biến)
Batch Normalization	BN	Chuẩn hóa theo lô (chuẩn hóa đầu vào mỗi lớp theo mini-batch)
Bi-temporal	-	Hai thời điểm (phân tích so sánh hai thời kỳ)
C-band	-	Băng tần C (bước sóng radar 4-8 GHz, xuyên qua mây)
Classification	-	Phân loại (gán nhãn cho đối tượng)
Compute Unified Device Architecture	CUDA	Nền tảng tính toán song song của NVIDIA trên GPU
Confusion Matrix	-	Ma trận nhầm lẫn (bảng so sánh dự đoán và thực tế)
Convolutional Neural Network	CNN	Mạng nơ-ron tích chập
Cross Validation	CV	Kiểm định chéo (đánh giá mô hình trên nhiều phần dữ liệu)
CrossEntropyLoss	-	Hàm mất mát entropy chéo (độ sai khác phân phối xác suất)

Deep Learning	DL	Học sâu (học máy với nhiều lớp ẩn)
Deforestation	-	Mất rừng (chuyển đổi từ rừng sang phi rừng)
Dropout	-	Loại bỏ ngẫu nhiên (tắt ngẫu nhiên một số neuron khi huấn luyện)
Early Stopping	-	Dừng sớm (ngừng huấn luyện khi validation loss không giảm)
Enhanced Vegetation Index	EVI	Chỉ số thực vật cải tiến (nhạy hơn NDVI ở vùng rậm rạp)
European Space Agency	ESA	Cơ quan Vũ trụ Châu Âu
European Union	EU	Liên minh Châu Âu
F1-Score	-	Điểm F1 (trung bình điều hòa của độ chuẩn xác và độ phủ)
False Negative	FN	Âm tính giả (dự đoán âm nhưng thực tế dương)
False Positive	FP	Đương tính giả (dự đoán dương nhưng thực tế âm)
False Positive Rate	FPR	Tỷ lệ dương tính giả
Feature	-	Đặc trưng (thuộc tính đầu vào của mô hình)
Feature Extraction	-	Trích xuất đặc trưng (rút trích thông tin từ dữ liệu thô)
Feature Map	-	Bản đồ đặc trưng (đầu ra của lớp tích chập)
Filter/Kernel	-	Bộ lọc/Nhân tích chập (ma trận trọng số trượt qua ảnh)
Fold	-	Phần gập (một phần dữ liệu trong Cross Validation)
Food and Agriculture Organization	FAO	Tổ chức Lương thực và Nông nghiệp Liên Hợp Quốc
Forest Change Detection	-	Phát hiện biến động rừng

Fully Connected Layer	FC	Lớp kết nối đầy đủ (mỗi neuron kết nối với tất cả neuron lớp trước)
Geographic Information System	GIS	Hệ thống thông tin địa lý
Global Average Pooling	GAP	Gộp trung bình toàn cục (lấy trung bình toàn bộ bản đồ đặc trưng)
Global Forest Watch	GFW	Giám sát rừng toàn cầu
Graphics Processing Unit	GPU	Bộ xử lý đồ họa (phân cứng tính toán song song)
Ground Truth	-	Dữ liệu mẫu (nhãn thực tế dùng để huấn luyện và đánh giá mô hình, thường thu thập từ khảo sát thực địa hoặc giải đoán ảnh)
Hyperparameter	-	Siêu tham số (tham số cấu hình trước khi huấn luyện)
Interferometric Wide	IW	Chế độ giao thoa rộng (chế độ chụp chính của Sentinel-1)
Intergovernmental Panel on Climate Change	IPCC	Ủy ban Liên chính phủ về Biến đổi Khí hậu
Learning Rate	LR	Tốc độ học (bước cập nhật trọng số mỗi lần lặp)
Lightweight	-	Kiến trúc nhẹ (mô hình có ít tham số, tính toán nhanh)
Logits	-	Giá trị logit (đầu ra thô của mạng trước khi áp dụng softmax)
Machine Learning	ML	Học máy
Mangrove Forest	-	Rừng ngập mặn
Multi-Layer Perceptron	MLP	Perceptron đa lớp (mạng neural nhiều lớp kết nối đầy đủ)
Multispectral	-	Đa phổ (ảnh chụp ở nhiều dải bước sóng)

Near-Infrared	NIR	Cận hồng ngoại (bước sóng 0.7–1.4 μm)
Non-forest	-	Phi rừng (vùng không có rừng che phủ)
Normalized Burn Ratio	NBR	Chỉ số cháy chuẩn hóa (phát hiện vùng cháy rừng)
Normalized Difference Moisture Index	NDMI	Chỉ số độ ẩm chuẩn hóa (đánh giá độ ẩm thực vật)
Normalized Difference Vegetation Index	NDVI	Chỉ số thực vật chuẩn hóa (đánh giá mức độ xanh tươi)
Normalized Difference Water Index	NDWI	Chỉ số nước chuẩn hóa (phát hiện mặt nước)
Optical	-	Quang học (ảnh vệ tinh dùng ánh sáng khả kiến và hồng ngoại)
Overfitting	-	Quá khớp (mô hình học thuộc dữ liệu huấn luyện, kém tổng quát)
Padding	-	Đệm viền (thêm pixel xung quanh ảnh khi tích chập)
Parameter	-	Tham số (trọng số học được trong quá trình huấn luyện)
Patch	-	Mảnh ảnh (vùng ảnh nhỏ trích xuất từ ảnh gốc)
Pixel	-	Điểm ảnh (đơn vị nhỏ nhất của ảnh số)
Polarization	-	Phân cực (hướng dao động của sóng radar: VV, VH)
Pooling	-	Gộp (giảm kích thước không gian của feature map)
Precision	-	Độ chuẩn xác (tỷ lệ dự đoán dương đúng trên tổng dự đoán dương)
Prediction	-	Dự đoán (kết quả đầu ra của mô hình)

Raster		-	Dữ liệu raster (ảnh dạng lưới điểm ảnh)
Recall		-	Độ phủ (tỷ lệ phát hiện đúng trên tổng số thực tế dương)
Receiver Operating Characteristic	ROC		Đường cong ROC (biểu đồ đánh giá khả năng phân loại)
Rectified Linear Unit	ReLU		Hàm ReLU (hàm kích hoạt: $f(x) = \max(0, x)$)
Reforestation		-	Phục hồi rừng (tái sinh hoặc trồng lại rừng)
Regularization		-	Chính quy hóa (kỹ thuật giảm overfitting)
Remote Sensing		-	Viễn thám (thu thập thông tin từ xa qua vệ tinh/máy bay)
ROC-AUC		-	Diện tích dưới đường cong ROC (đo khả năng phân biệt lớp)
Scheduler		-	Bộ điều chỉnh tốc độ học (thay đổi learning rate theo epoch)
Short-Wave Infrared	SWIR		Hồng ngoại sóng ngắn (bước sóng 1.4–3 μm)
Softmax		-	Hàm softmax (chuyển logits thành phân phối xác suất)
Spectral Signature		-	Phổ phản xạ đặc trưng (đặc điểm phản xạ theo bước sóng)
Stratified		-	Phân tầng (chia dữ liệu giữ nguyên tỷ lệ các lớp)
Synthetic Aperture Radar	SAR		Radar khẩu độ tổng hợp (cảm biến chủ động, hoạt động mọi thời tiết)
Test Set		-	Tập kiểm tra (dữ liệu đánh giá cuối cùng, không dùng khi huấn luyện)
True Positive Rate	TPR		Tỷ lệ dương tính thật (độ nhạy/độ phủ)

Training		-	Huấn luyện (quá trình học tham số từ dữ liệu)
Universal Mercator	Transverse UTM		Hệ tọa độ UTM (phép chiếu bắn đồ chia thành 60 múi)
Validation		-	Xác thực (đánh giá mô hình trong quá trình huấn luyện)
Weight Decay		-	Phân rã trọng số (thêm phạt L2 vào hàm mất mát)
World Geodetic System	WGS		Hệ trắc địa thế giới (hệ tọa độ toàn cầu, WGS84)

MỞ ĐẦU

Đặt vấn đề

Rừng đóng vai trò quan trọng trong việc duy trì cân bằng sinh thái, điều hòa khí hậu, lưu giữ carbon và bảo vệ đa dạng sinh học. Tuy nhiên, tình trạng mất rừng đang diễn ra nghiêm trọng trên toàn cầu, đặc biệt tại các quốc gia đang phát triển. Theo báo cáo “Global Forest Resources Assessment 2020” của Tổ chức Lương thực và Nông nghiệp Liên hợp quốc [?], thế giới đã mất ròng (net loss) khoảng 178 triệu hecta rừng trong giai đoạn 1990–2020, tương đương diện tích của Libya.

Tại Việt Nam, mặc dù độ che phủ rừng đã tăng từ 37% (năm 2000) lên 42% (năm 2020) nhờ các chương trình trồng rừng, nhưng tình trạng suy thoái và mất rừng tự nhiên vẫn đáng báo động, đặc biệt tại các tỉnh ven biển và đồng bằng sông Cửu Long. Tỉnh Cà Mau, nằm ở cực Nam Tổ Quốc, sở hữu hệ sinh thái rừng ngập mặn quan trọng nhưng đang phải đối mặt với áp lực từ nuôi trồng thủy sản, xâm nhập mặn và biến đổi khí hậu.

Phương pháp giám sát rừng truyền thống dựa trên điều tra thực địa tốn kém thời gian, chi phí và khó áp dụng cho diện tích rộng. Công nghệ viễn thám vệ tinh cung cấp giải pháp hiệu quả, cho phép giám sát liên tục, diện rộng với chi phí hợp lý. Chương trình Copernicus của Liên minh Châu Âu cung cấp dữ liệu miễn phí từ các vệ tinh Sentinel-1 (SAR) và Sentinel-2 (Optical) với độ phân giải không gian 10m và chu kỳ quay trở lại ngắn (5–6 ngày), phù hợp cho giám sát rừng nhiệt đới.

Trong những năm gần đây, trí tuệ nhân tạo và học sâu (Deep Learning) đã đạt được những bước tiến vượt bậc trong xử lý ảnh và nhận dạng mẫu. Mạng Neural Tích chập (Convolutional Neural Networks – CNN) đặc biệt hiệu quả trong phân loại ảnh nhờ khả năng tự động học đặc trưng không gian từ dữ liệu thô.

Xuất phát từ nhu cầu thực tiễn về giám sát rừng hiệu quả và xu hướng ứng dụng công nghệ AI tiên tiến, đồ án này lựa chọn đề tài **“Ứng dụng viễn thám và học sâu trong giám sát biến động rừng tỉnh Cà Mau”** nhằm phát triển mô hình phát hiện mất rừng với độ chính xác cao.

Mục tiêu nghiên cứu

Mục tiêu tổng quát của đồ án là phát triển mô hình học sâu dựa trên kiến trúc CNN để phát hiện và phân loại tự động các khu vực biến động rừng từ ảnh vệ tinh đa nguồn (Sentinel-1 SAR và Sentinel-2 Optical) tại tỉnh Cà Mau.

Để đạt được mục tiêu tổng quát, đề tài tập trung vào năm mục tiêu cụ thể. Thứ nhất, xây dựng bộ dữ liệu huấn luyện thông qua thu thập và xử lý dữ liệu ảnh vệ tinh Sentinel-1/2 đa thời gian, kết hợp với ground truth points để tạo bộ dữ liệu huấn luyện chất lượng cao. Thứ hai, thiết kế kiến trúc CNN nhẹ (lightweight) phù hợp với bộ dữ liệu có quy mô vừa phải (khoảng 2.600 mẫu). Thứ ba, triển khai phương pháp phân chia dữ liệu khoa học bằng stratified random split kết hợp với 5-Fold Cross Validation. Thứ tư, huấn luyện và tối ưu hóa mô hình bằng các kỹ thuật huấn luyện tiên tiến. Thứ năm, áp dụng mô hình để phân loại biến động rừng toàn bộ khu vực nghiên cứu.

Đối tượng nghiên cứu

Đối tượng nghiên cứu của đồ án là khu vực ranh giới lâm nghiệp tỉnh Cà Mau (theo địa giới hành chính mới có hiệu lực từ ngày 01/07/2025, sau khi sáp nhập tỉnh Cà Mau cũ và tỉnh Bạc Liêu) với tổng diện tích 170.179 hecta (tương đương 1.701,79 km²). Khu vực này bao gồm các loại hình rừng tự nhiên và rừng trồng, trong đó chủ yếu là rừng ngập mặn và rừng phòng hộ ven biển. Các trạng thái biến động rừng được phân loại thành bốn nhóm: Rừng ổn định (Forest Stable), Mất rừng (Deforestation), Phi rừng (Non-forest) và Tái sinh rừng (Reforestation). Dữ liệu sử dụng bao gồm ảnh vệ tinh Sentinel-1 (radar băng C) và Sentinel-2 (quang học đa phổ) trong giai đoạn từ tháng 01/2024 đến tháng 02/2025. Diện tích thực tế được phân loại là 162.469 hecta (khoảng 95,5% diện tích ranh giới, phần còn lại bị loại do mây che hoặc dữ liệu không hợp lệ).

Phương pháp nghiên cứu

Đồ án áp dụng phương pháp nghiên cứu thực nghiệm, kết hợp giữa viễn thám và học sâu, bao gồm năm giai đoạn chính.

Giai đoạn thứ nhất là thu thập và tiền xử lý dữ liệu, trong đó ảnh vệ tinh Sentinel-1 và Sentinel-2 được thu thập từ nền tảng Google Earth Engine. Quá trình tiền xử lý bao

gồm hiệu chỉnh khí quyển, tạo ảnh composite theo mùa và loại bỏ nhiễu.

Giai đoạn thứ hai là trích xuất đặc trưng, bao gồm việc tính toán các chỉ số phổ từ dữ liệu Sentinel-2 (NDVI, NDWI, EVI, SAVI, BSI, NDBI) và các chỉ số backscatter từ Sentinel-1 (VV, VH, VV/VH ratio), tổng hợp 18 đặc trưng cho mỗi điểm mẫu.

Giai đoạn thứ ba là chuẩn bị mẫu huấn luyện, trong đó bộ dữ liệu 2.600 mẫu với 4 lớp biến động rừng được xây dựng. Kỹ thuật chuẩn hóa dữ liệu được áp dụng và dữ liệu được phân chia theo tỷ lệ 70% huấn luyện, 15% validation và 15% kiểm tra.

Giai đoạn thứ tư là huấn luyện mô hình, bao gồm thiết kế kiến trúc CNN nhẹ với 3 lớp tích chập, sử dụng các kỹ thuật regularization (Dropout, Batch Normalization) và áp dụng 5-Fold Cross Validation để đánh giá độ ổn định của mô hình.

Giai đoạn thứ năm là áp dụng mô hình để phân loại toàn vùng nghiên cứu bằng phương pháp sliding window, tổng hợp kết quả và đánh giá độ chính xác.

Ý nghĩa khoa học và thực tiễn của đề tài

Về mặt khoa học, đồ án đề xuất kiến trúc CNN nhẹ và hiệu quả cho bài toán phân loại ảnh viễn thám với bộ dữ liệu quy mô vừa phải. Nghiên cứu chứng minh hiệu quả của việc tích hợp đa nguồn dữ liệu (SAR và Optical) trong phát hiện biến động rừng, đồng thời đóng góp vào hướng nghiên cứu ứng dụng học sâu trong lĩnh vực viễn thám và giám sát môi trường.

Về ý nghĩa thực tiễn, mô hình cung cấp công cụ tự động phát hiện mất rừng với độ chính xác cao (trên 98%), giúp giảm đáng kể thời gian và chi phí so với phương pháp điều tra thực địa truyền thống. Kết quả nghiên cứu có thể hỗ trợ các cơ quan quản lý lâm nghiệp trong việc giám sát và bảo vệ rừng tại tỉnh Cà Mau. Ngoài ra, mô hình có thể mở rộng áp dụng cho các khu vực khác có điều kiện tương tự.

Cấu trúc đồ án

Đồ án được tổ chức thành bốn chương chính.

Chương 1: Tổng quan về vấn đề nghiên cứu trình bày tổng quan về bối cảnh mất rừng, công nghệ viễn thám, tổng quan các nghiên cứu liên quan và các khoảng trống nghiên cứu.

Chương 2: Cơ sở lý thuyết giới thiệu chi tiết về công nghệ viễn thám (Sentinel-1/2), lý thuyết về mạng Neural Tích chập (CNN), các phương pháp phân loại ảnh và các tiêu chí đánh giá mô hình.

Chương 3: Phương pháp nghiên cứu mô tả khu vực nghiên cứu, dữ liệu sử dụng, quy trình xử lý, kiến trúc mô hình CNN đề xuất, phương pháp huấn luyện và đánh giá.

Chương 4: Thực nghiệm và kết quả trình bày các kết quả huấn luyện, đánh giá mô hình, phân loại toàn vùng, phân tích lỗi và trực quan hóa. Đồng thời đưa ra kết luận và kiến nghị cho các nghiên cứu tiếp theo.

CHƯƠNG 1. TỔNG QUAN VỀ BIẾN ĐỘNG RỪNG, CÔNG NGHỆ VIỄN THÁM VÀ CÁC PHƯƠNG PHÁP GIÁM SÁT

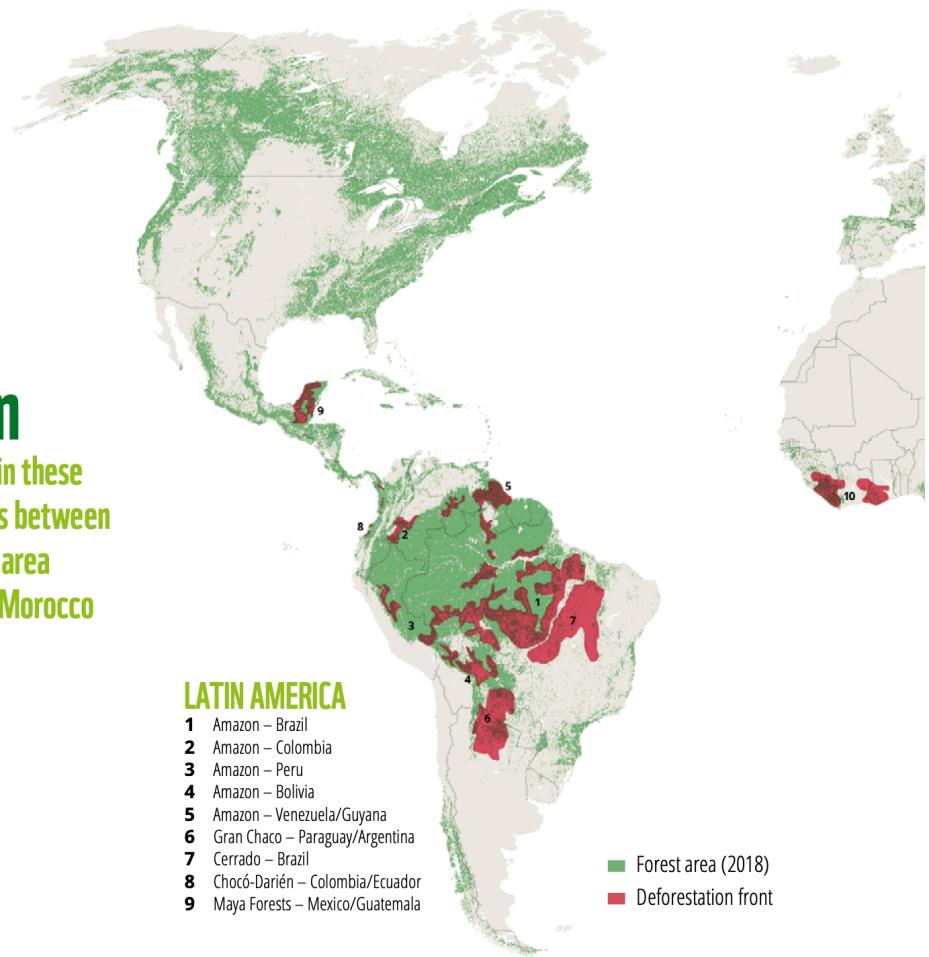
1.1. Bối cảnh và tình hình mất rừng

1.1.1. Tình hình mất rừng trên thế giới

Rừng bao phủ khoảng 31% diện tích đất liền toàn cầu [?], đóng vai trò thiết yếu trong việc điều hòa khí hậu và bảo tồn đa dạng sinh học. Tuy nhiên, tốc độ mất rừng toàn cầu vẫn đang ở mức báo động. Theo báo cáo “Global Forest Resources Assessment 2020” của FAO [?], tổng diện tích rừng bị phá (gross deforestation) từ năm 1990 đến 2020 ước tính khoảng 420 triệu hecta. Mặc dù diện tích mất rừng ròng đã giảm nhờ nỗ lực trồng rừng, nhưng việc chuyển đổi đất rừng sang nông nghiệp và chăn nuôi vẫn diễn ra phức tạp.

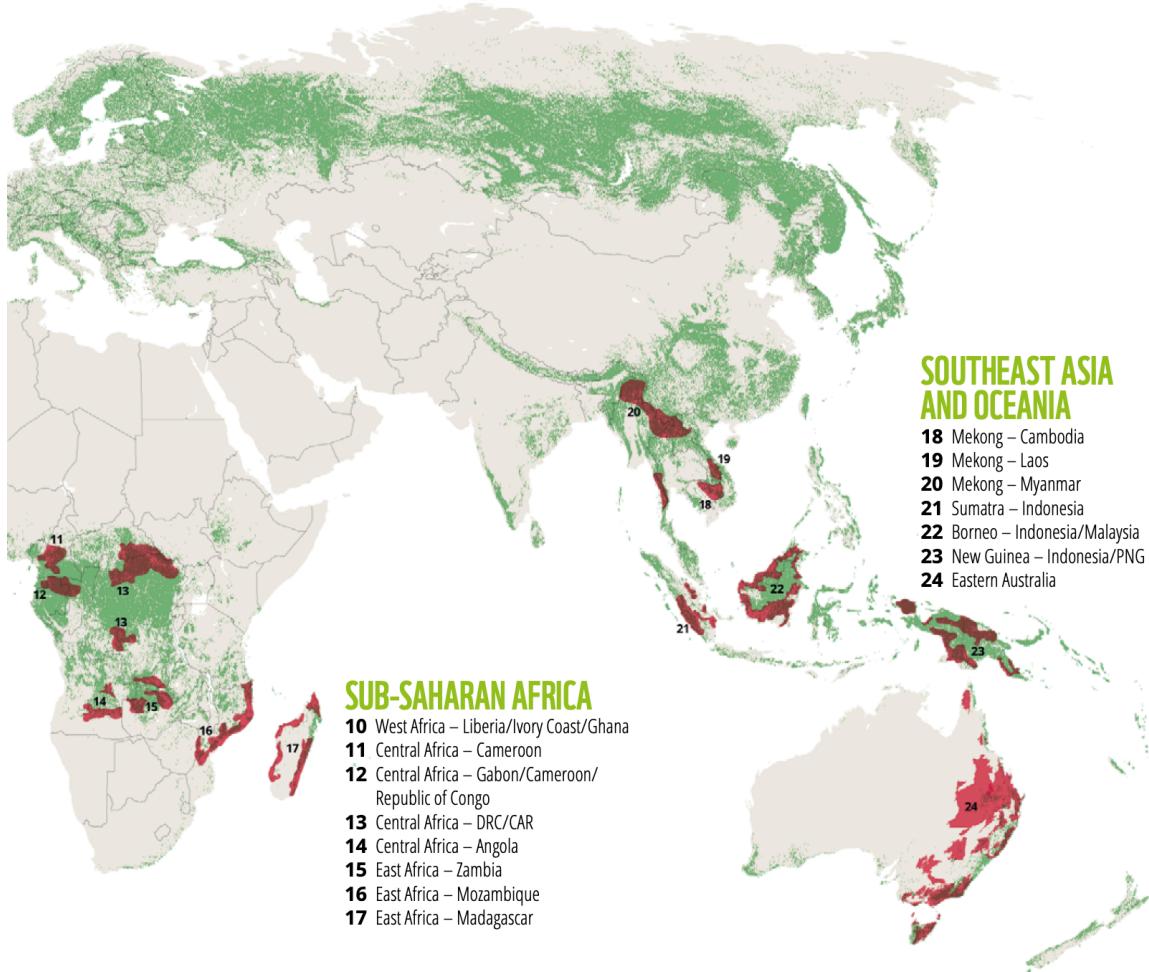
Sự suy giảm này tập trung nghiêm trọng nhất tại khu vực nhiệt đới. Theo báo cáo của WWF, chỉ tính riêng giai đoạn 2004–2017, hơn 43 triệu hecta rừng đã bị xóa sổ tại các “mặt trận” nóng bỏng nhất, diện tích tương đương quy mô nước Maroc. Trong đó, **Lưu vực Amazon (Nam Mỹ)** là nơi chịu tổn thất nặng nề nhất do áp lực từ chăn nuôi và nông nghiệp quy mô lớn.

**Over
43 million**
hectares were lost in these
deforestation fronts between
2004 and 2017, an area
roughly the size of Morocco



Hình 1.1: Các mặt trận mất rừng trọng điểm tại khu vực Mỹ Latinh. (Nguồn: WWF, 2021)

Không chỉ giới hạn ở Châu Mỹ, tình trạng phá rừng cũng đang diễn biến phức tạp tại bờ biển kia đại dương. **Lưu vực Congo (Trung Phi) và Đông Nam Á** là những điểm nóng tiếp theo. Tại Đông Nam Á, rừng nguyên sinh đang bị thu hẹp nhanh chóng tại các khu vực như sông Mekong, đảo Sumatra và Borneo để nhường chỗ cho các đồn điền cây công nghiệp.



Hình 1.2: Các mặt trận mất rừng trọng điểm tại Châu Phi và Đông Nam Á. (Nguồn: WWF, 2021)

Xu hướng này vẫn tiếp diễn trong những năm gần đây. Theo Global Forest Watch [?], thế giới mất khoảng 10 triệu hecta rừng nhiệt đới mỗi năm trong giai đoạn 2015–2020. Việc này không chỉ làm giảm khả năng hấp thụ CO₂ mà còn trực tiếp phát thải khí nhà kính từ việc đốt rừng và phân hủy sinh khối. Theo IPCC [?], phá rừng và thay đổi sử dụng đất đóng góp khoảng 23% tổng lượng phát thải khí nhà kính do con người gây ra, góp phần làm tăng hiện tượng biến đổi khí hậu toàn cầu.

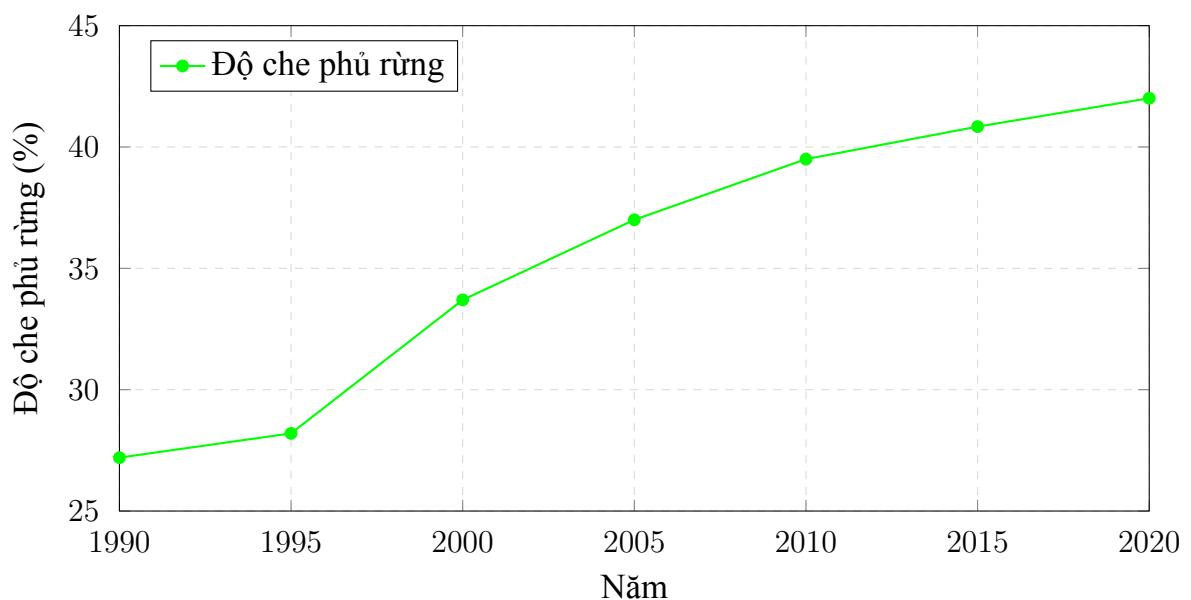
1.1.2. Tình hình mất rừng tại Việt Nam

Việt Nam đã trải qua những biến đổi lớn về độ che phủ rừng trong 30 năm qua. Sau thời kỳ suy giảm nghiêm trọng (độ che phủ chỉ còn 28% vào năm 1990 do chiến tranh và khai thác bừa bãi), Việt Nam đã thực hiện nhiều chương trình phục hồi và phát triển rừng.

Nhờ các chương trình như “Trồng 5 triệu hecta rừng” (1998-2010), độ che phủ rừng đã tăng lên 42% vào năm 2020 [?].

Tuy nhiên, chất lượng rừng là một vấn đề đáng lo ngại. Mặc dù tổng diện tích rừng tăng từ 9.4 triệu hecta (1990) lên 14.6 triệu hecta (2020) chủ yếu nhờ rừng trồng (cao su, keo, thông), chất lượng rừng tự nhiên lại suy giảm đáng kể. Theo số liệu của Bộ NN&PTNT (2020), rừng tự nhiên hiện có khoảng 10.29 triệu hecta, nhưng rừng nguyên sinh (primary forest) chỉ còn chiếm khoảng 0.25% tổng diện tích rừng [?].

Nguyên nhân chính gây mất rừng tại Việt Nam bao gồm việc chuyển đổi sang đất nông nghiệp như trồng cà phê, cao su và điều; khai thác gỗ trái phép; phát triển cơ sở hạ tầng và đô thị hóa; cháy rừng; và hoạt động nuôi trồng thủy sản, đặc biệt tại khu vực ven biển và đồng bằng sông Cửu Long.



Hình 1.3: Biến động độ che phủ rừng Việt Nam giai đoạn 1990-2020

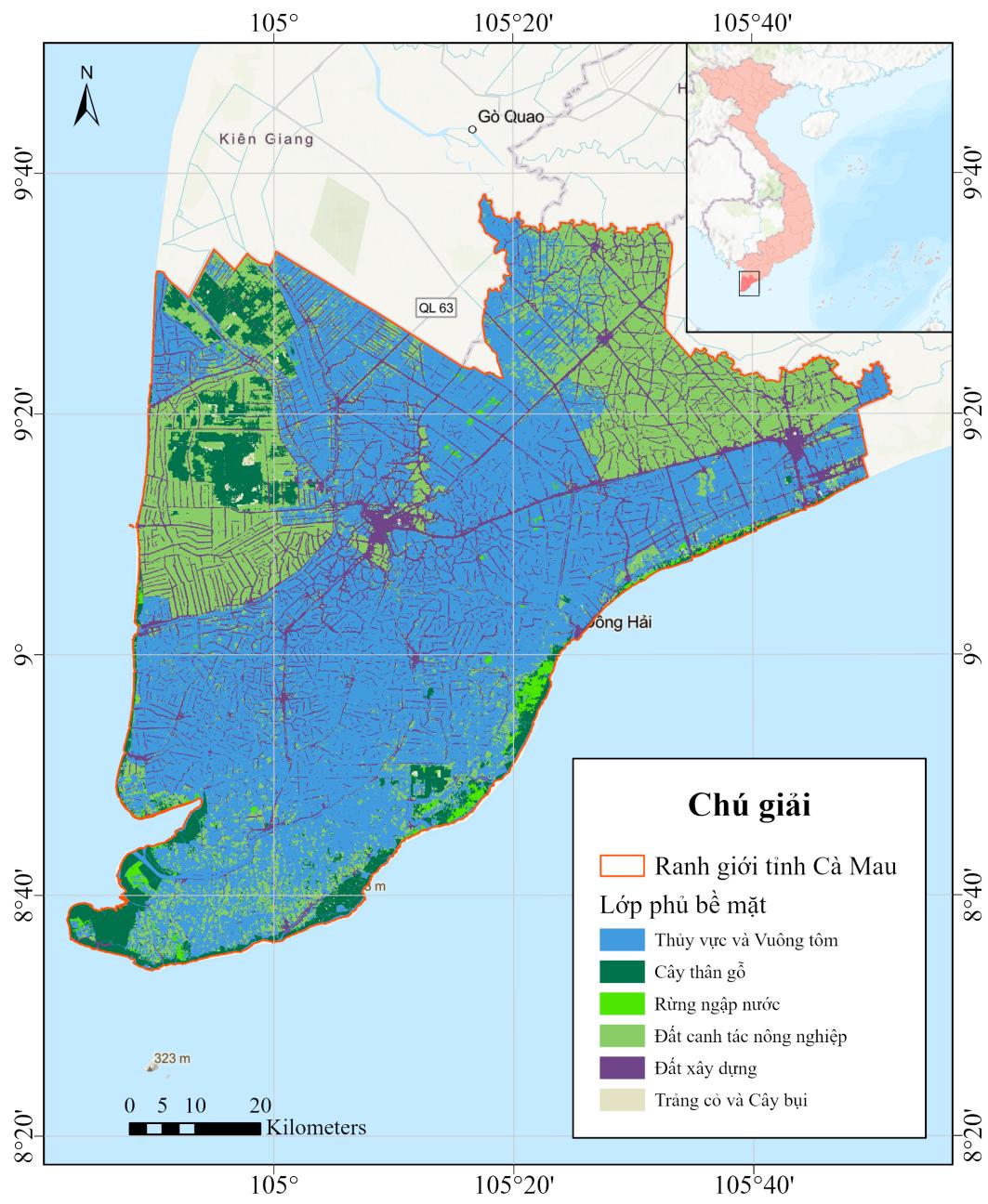
1.1.3. Tình hình rừng tại tỉnh Cà Mau

Lưu ý về địa giới hành chính: Theo Nghị quyết số 1278/NQ-UBTVQH15 ngày 24/10/2024 của Ủy ban Thường vụ Quốc hội, kể từ ngày 01/07/2025, tỉnh Cà Mau và tỉnh Bạc Liêu được sáp nhập thành tỉnh Cà Mau mới với tổng diện tích tự nhiên 7,942.38 km² và dân số khoảng 2.6 triệu người. Đề án này nghiên cứu trên phạm vi rừng của tỉnh Cà Mau mới, bao gồm cả vùng rừng thuộc địa bàn Bạc Liêu cũ.

Tỉnh Cà Mau mới nằm ở cực Nam Tổ Quốc, sở hữu hệ sinh thái rừng đa dạng bao

gồm cả rừng ngập mặn ven biển và rừng tràm nội địa. Theo số liệu trước khi sáp nhập, tỉnh Cà Mau cũ có diện tích rừng khoảng 94,319 hecta và tỉnh Bạc Liêu có khoảng 5,730 hecta rừng, tổng cộng khoảng 100,000 hecta rừng trên toàn tỉnh Cà Mau mới [?]. Trong đó, rừng ngập mặn Cà Mau chiếm khoảng 20% diện tích rừng ngập mặn của Việt Nam. Hệ thống rừng tại Cà Mau đóng vai trò then chốt trong việc phòng hộ ven biển (chắn sóng, chống xâm thực và bảo vệ bờ biển), bảo tồn đa dạng sinh học vì là môi trường sống cho nhiều loài động thực vật quý hiếm, cung cấp nguồn sinh kế thông qua các hoạt động thủy sản và du lịch sinh thái, và góp phần giảm nhẹ biến đổi khí hậu nhờ khả năng lưu giữ carbon cao, gấp khoảng 3–5 lần so với rừng nhiệt đới trên cạn [? ?].

Tuy nhiên, rừng Cà Mau đang phải đối mặt với nhiều thách thức. Trước hết là áp lực chuyển đổi sang nuôi tôm do kinh tế, khiến nhiều khu vực rừng bị chuyển đổi thành ao nuôi. Ngoài ra, hiện tượng xâm nhập mặn gia tăng do biến đổi khí hậu làm giảm sức khỏe rừng; đồng thời xói mòn bờ biển cũng làm suy giảm diện tích rừng ven biển; và tình trạng thiếu nước ngọt ảnh hưởng tới khả năng tái sinh tự nhiên của rừng. Giai đoạn 2011-2023, sát lở vùng ven biển đã làm mất hơn 6,200 hecta đất và rừng phòng hộ [?]. Việc giám sát và bảo vệ rừng tại Cà Mau là ưu tiên hàng đầu nhằm duy trì hệ sinh thái quan trọng này.



Hình 1.4: Bản đồ lớp phủ bè mặt khu vực tỉnh Cà Mau.

1.2. Công nghệ viễn thám trong giám sát rừng

1.2.1. Ưu điểm của công nghệ viễn thám

Công nghệ viễn thám vệ tinh đem lại nhiều ưu thế rõ rệt so với các phương pháp điều tra thực địa truyền thống. Trước hết, ảnh vệ tinh có khả năng bao phủ diện tích rất lớn, giúp quan sát đồng thời hàng nghìn km² rừng. Bên cạnh đó, các vệ tinh hiện đại có chu kỳ lặp ngắn, thường chỉ 3–5 ngày, tạo điều kiện phát hiện kịp thời những biến động xảy ra trong rừng. Nguồn dữ liệu từ nhiều chương trình vệ tinh hiện nay còn được cung cấp miễn phí, góp phần giảm đáng kể chi phí so với khảo sát ngoài thực địa. Hệ thống lưu trữ ảnh vệ tinh phong phú theo thời gian cũng cho phép phân tích chuỗi biến động dài hạn. Đối với những khu vực khó tiếp cận như rừng sâu, vùng núi cao hay khu vực biên giới, viễn thám vẫn có thể giám sát hiệu quả. Ngoài ra, tính khách quan và khả năng lặp lại của dữ liệu viễn thám giúp hạn chế các sai lệch do yếu tố chủ quan trong quá trình điều tra trực tiếp.

1.2.2. Chương trình Copernicus và vệ tinh Sentinel

Chương trình Copernicus của Liên minh Châu Âu (EU) là một trong những chương trình quan sát Trái Đất lớn nhất thế giới, cung cấp dữ liệu miễn phí và mở. Hai vệ tinh quan trọng cho giám sát rừng là:

Sentinel-1 (SAR - Synthetic Aperture Radar):

Vệ tinh Sentinel-1 hoạt động ở dải sóng C-band (xấp xỉ 5.5 cm) với hai chế độ phân cực chính là VV (Vertical-Vertical) và VH (Vertical-Horizontal); độ phân giải không gian trong chế độ Interferometric Wide (IW) là 10m [?]. Về chu kỳ quay trở lại, Sentinel-1A có chu kỳ khoảng 12 ngày; khi kết hợp với Sentinel-1B (đã ngừng hoạt động từ tháng 12/2021 do sự cố nguồn điện), chu kỳ giảm còn 6 ngày. Sentinel-1C được phóng vào tháng 12/2024 để thay thế Sentinel-1B. Do là hệ thống chủ động, Sentinel-1 có ưu điểm xuyên qua mây và khói, hoạt động được cả ngày lẫn đêm, và nhạy cảm đối với cấu trúc thực vật cũng như độ ẩm.

Sentinel-2 (Optical - Multispectral Imaging):

Vệ tinh Sentinel-2 cung cấp 13 dải phổ từ vùng nhìn thấy đến hồng ngoại ngắn (từ 443 nm đến 2190 nm) với nhiều cấp độ độ phân giải không gian: 10m cho các dải B2, B3,

B4 và B8; 20m cho các dải B5, B6, B7, B8a, B11 và B12; và 60m cho B1, B9 và B10 [?]. Chu kỳ quay trở lại của tổ hợp hai vệ tinh Sentinel-2A và Sentinel-2B vào khoảng 5 ngày, và vì có thông tin quang phổ phong phú nên Sentinel-2 rất phù hợp để tính toán chỉ số thực vật.



Hình 1.5: Vệ tinh Sentinel-1. (Nguồn: European Space Agency)



Hình 1.6: Vệ tinh Sentinel-2. (Nguồn: European Space Agency)

1.2.3. Chỉ số thực vật từ dữ liệu quang học

Các chỉ số thực vật (vegetation indices) là công cụ quan trọng trong giám sát rừng, được tính toán từ các dải phổ khác nhau:

NDVI (Normalized Difference Vegetation Index):

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (1.1)$$

NDVI có dải giá trị từ -1 đến 1; giá trị NDVI lớn hơn 0.6 thường biểu thị thực vật xanh tốt, trong khi giá trị NDVI nhỏ hơn 0.2 thường liên quan đến đất trống, nước hoặc khu vực đô thị [?].

NBR (Normalized Burn Ratio):

$$NBR = \frac{NIR - SWIR_2}{NIR + SWIR_2} \quad (1.2)$$

NBR nhạy cảm với vùng cháy; biến đổi Delta NBR (dNBR) được sử dụng để đánh giá mức độ tổn thất do cháy rừng.

NDMI (Normalized Difference Moisture Index):

$$NDMI = \frac{NIR - SWIR_1}{NIR + SWIR_1} \quad (1.3)$$

NDMI được dùng để đánh giá hàm lượng nước trong thực vật; giá trị NDMI thấp có thể chỉ ra trạng thái stress do hạn hán.

1.2.4. Tích hợp dữ liệu SAR và Optical

Việc kết hợp dữ liệu SAR (Sentinel-1) và Optical (Sentinel-2) mang lại nhiều lợi ích thực tế. Về khía cạnh bổ sung thông tin, SAR cung cấp dữ liệu về cấu trúc, độ nhám bề mặt và độ ẩm, trong khi Optical cung cấp thông tin quang phổ và các chỉ số thực vật. Về khía cạnh khắc phục hạn chế, SAR hoạt động hiệu quả trong điều kiện mây mù — điều quan trọng trong môi trường rừng nhiệt đới — còn Optical lại cung cấp dữ liệu trực quan dễ phân giải. Về khía cạnh nâng cao độ chính xác, nhiều nghiên cứu cho thấy việc kết hợp SAR và Optical giúp tăng accuracy từ khoảng 5 đến 10% so với việc sử dụng mỗi nguồn dữ liệu riêng lẻ [? ?]. Về khía cạnh phát hiện biến động đa chiều, SAR nhạy với biến đổi cấu trúc như chặt cây, trong khi Optical nhạy với biến đổi quang phổ thể hiện sức khỏe thực vật.

1.3. Sự phát triển của các phương pháp giám sát biến động rừng

Giám sát biến động rừng đã trải qua nhiều giai đoạn phát triển, từ các phương pháp thủ công truyền thống đến các kỹ thuật học sâu hiện đại. Phần này trình bày quá trình phát triển của các phương pháp theo thời gian, làm cơ sở cho việc lựa chọn phương pháp trong nghiên cứu.

1.3.1. Phương pháp truyền thống

Khảo sát thực địa (Field Survey):

Trước khi có công nghệ viễn thám, giám sát rừng chủ yếu dựa vào khảo sát thực địa trực tiếp. Nhân viên kiểm lâm đi thực địa để đo đạc, ghi nhận trạng thái rừng và lập bản đồ thủ công. Phương pháp này có độ chính xác cao tại từng điểm khảo sát nhưng tốn kém về thời gian và nhân lực, không khả thi cho giám sát diện rộng và khó cập nhật thường xuyên.

Ảnh hàng không (Aerial Photography):

Từ giữa thế kỷ 20, ảnh hàng không bắt đầu được sử dụng để lập bản đồ rừng. Các chuyên gia giải đoán ảnh hàng không bằng mắt thường để xác định ranh giới rừng và phát hiện thay đổi. Phương pháp này cho phép quan sát diện tích lớn hơn khảo sát thực địa nhưng vẫn phụ thuộc nhiều vào kinh nghiệm của người giải đoán và chi phí bay chụp cao.

1.3.2. Phương pháp viễn thám thế hệ đầu

Giải đoán ảnh trực quan (Visual Interpretation):

Với sự ra đời của vệ tinh Landsat năm 1972, lần đầu tiên có thể quan sát bề mặt Trái Đất một cách hệ thống từ không gian. Giai đoạn đầu, việc phân tích ảnh vệ tinh chủ yếu dựa vào giải đoán trực quan - chuyên gia nhìn ảnh và vẽ ranh giới các vùng đất khác nhau. Phương pháp này tận dụng được kiến thức chuyên môn của người giải đoán nhưng mang tính chủ quan, không nhất quán giữa các chuyên gia và khó tái lập.

Phân loại dựa trên ngưỡng (Threshold-based Classification):

Các phương pháp đơn giản như phân ngưỡng chỉ số thực vật NDVI được sử dụng

rộng rãi trong thập niên 1980-1990. Ví dụ, nếu NDVI < 0.2 thì phân loại là đất trống, nếu NDVI > 0.5 thì phân loại là rừng. Phương pháp này đơn giản, dễ triển khai nhưng thiếu linh hoạt, không xử lý tốt các trường hợp phức tạp và nhạy cảm với nhiều khía cạnh.

1.3.3. Phương pháp học máy truyền thống

Phân loại không giám sát (Unsupervised Classification):

Phân loại không giám sát không yêu cầu dữ liệu huấn luyện có nhãn, thay vào đó tự động nhóm các pixel có đặc trưng phổ biến tự thành các cụm. Hai thuật toán phổ biến nhất là K-means và ISODATA (Iterative Self-Organizing Data Analysis Technique) [?]. K-means phân chia dữ liệu thành k cụm dựa trên khoảng cách Euclidean đến tâm cụm. ISODATA mở rộng K-means bằng cách tự động tách, gộp hoặc loại bỏ cụm dựa trên các ngưỡng thống kê. Phương pháp không giám sát phù hợp khi không có dữ liệu thực địa hoặc khi cần khám phá cấu trúc dữ liệu ban đầu, nhưng có hạn chế là các cụm tạo ra không có ý nghĩa ngữ nghĩa rõ ràng và cần chuyên gia gán nhãn sau đó.

Phân loại có giám sát (Supervised Classification):

Từ thập niên 1990, các thuật toán phân loại thống kê bắt đầu được áp dụng rộng rãi. Maximum Likelihood Classification (MLC) giả định dữ liệu tuân theo phân phối Gaussian và phân loại pixel dựa trên xác suất. Phương pháp này đặt nền tảng cho phân loại ảnh viễn thám định lượng nhưng có hạn chế khi dữ liệu không tuân theo phân phối chuẩn [?].

Decision Trees và Random Forest:

Hansen và cộng sự [?] đã sử dụng thuật toán Decision Trees để phát triển Global Forest Change dataset — bộ dữ liệu mất rừng toàn cầu đầu tiên ở độ phân giải 30m từ chuỗi thời gian Landsat (2000-2012). Công trình này đánh dấu bước tiến quan trọng trong giám sát rừng quy mô lớn, đạt accuracy khoảng 85%. Random Forest, được giới thiệu bởi Breiman (2001), cải thiện Decision Trees bằng cách kết hợp nhiều cây quyết định và trở thành thuật toán phổ biến nhất trong phân loại ảnh viễn thám. Nguyen và cộng sự [?] áp dụng Random Forest với Sentinel-2 đa thời gian để lập bản đồ sử dụng đất tại Đăk Nông, Việt Nam và đạt overall accuracy 91.2%.

Support Vector Machine (SVM):

SVM được đề xuất cho phân loại ảnh viễn thám từ đầu những năm 2000 và nhanh

chóng chứng minh hiệu quả vượt trội so với MLC, đặc biệt với dữ liệu đa chiều và bộ mẫu nhỏ. SVM tìm siêu phẳng tối ưu để phân tách các lớp trong không gian đặc trưng, phù hợp với dữ liệu viễn thám có số chiều cao (nhiều bands).

1.3.4. Phương pháp học sâu

Convolutional Neural Networks (CNN):

Từ năm 2012, với sự thành công của AlexNet trong ImageNet Competition, học sâu bắt đầu cách mạng hóa computer vision và nhanh chóng được áp dụng vào viễn thám. Zhu và cộng sự [?] tổng hợp các ứng dụng của deep learning trong viễn thám và chỉ ra tiềm năng to lớn của CNN trong phân loại ảnh vệ tinh. Zhang và cộng sự [?] giới thiệu các kiến trúc CNN phổ biến (AlexNet, VGGNet, ResNet) và ứng dụng của chúng trong viễn thám. Kussul và cộng sự [?] áp dụng CNN cho phân loại cây trồng từ Sentinel-2 và đạt accuracy 94.5%, cao hơn đáng kể so với Random Forest.

Kiến trúc U-Net và các biến thể:

Ronneberger và cộng sự [?] đề xuất kiến trúc U-Net với cấu trúc encoder-decoder, ban đầu cho phân đoạn ảnh y sinh nhưng sau đó được áp dụng rộng rãi trong viễn thám nhờ khả năng phân đoạn ngữ nghĩa pixel-wise hiệu quả. Các biến thể như ResU-Net, Attention U-Net tiếp tục cải tiến hiệu suất cho các bài toán phân đoạn ảnh vệ tinh.

Ứng dụng CNN trong giám sát rừng:

Hethcoat và cộng sự [?] áp dụng CNN (kiến trúc ResNet) để phát hiện khai thác gỗ chọn lọc tại Amazon từ dữ liệu Sentinel-1 và Sentinel-2, đạt accuracy 94.3%. Karra và cộng sự [?] ứng dụng deep learning kết hợp Sentinel-2 để tạo bản đồ sử dụng đất toàn cầu ở độ phân giải 10m.

1.3.5. Xu hướng tích hợp đa nguồn dữ liệu

Kết hợp SAR và Optical:

Một xu hướng quan trọng trong giám sát rừng hiện đại là tích hợp dữ liệu radar (SAR) và quang học. Reiche và cộng sự [?] kết hợp Sentinel-1 (SAR) và Landsat (Optical) để phát hiện mát rừng near-real-time tại rừng nhiệt đới khô và đạt accuracy 93.8%. Hu và cộng sự [?] kết hợp Sentinel-1 và Sentinel-2 để phân loại rừng ở Madagascar

và ghi nhận accuracy tăng từ 87% (chỉ dùng optical) lên 92% (kết hợp cả hai).

Ưu điểm của tích hợp đa nguồn:

Việc kết hợp SAR và Optical mang lại nhiều lợi ích. Dữ liệu SAR có khả năng quan sát xuyên mây, đặc biệt quan trọng ở vùng nhiệt đới thường xuyên có mây. Dữ liệu quang học cung cấp thông tin phong phú về phổ phản xạ của thực vật. Hai nguồn dữ liệu bổ sung cho nhau, giúp tăng độ chính xác và độ tin cậy của kết quả phân loại.

Nghiên cứu tại Việt Nam:

Pham và cộng sự [?] đã sử dụng kết hợp ảnh QuickBird, LiDAR và chỉ số địa hình GIS để nhận dạng loài cây bản địa trong cảnh quan phức tạp. Bùi và cộng sự [?] nghiên cứu biến động rừng ngập mặn ven biển Đồng bằng sông Cửu Long bằng chuỗi thời gian Landsat giai đoạn 1990-2020, ghi nhận xu hướng suy giảm đáng kể.

1.3.6. Tổng hợp và so sánh các phương pháp

Bảng 1.1: So sánh các phương pháp giám sát biến động rừng qua các giai đoạn

Giai đoạn	Thời kỳ	Phương pháp tiêu biểu	Đặc điểm
Truyền thống	Trước 1970	Khảo sát thực địa	Chính xác nhưng tốn kém
Viễn thám đầu	1970-1990	Giải đoán trực quan	Chủ quan, khó tái lập
ML truyền thống	1990-2012	MLC, Decision Tree, RF	Khách quan, tự động hóa
Học sâu	2012-nay	CNN, U-Net	Học đặc trưng tự động
Tích hợp	2015-nay	CNN + SAR + Optical	Bổ sung, tăng độ tin cậy

Bảng 1.2: Tổng hợp các nghiên cứu tiêu biểu theo các phương pháp

Tác giả	Năm	Phương pháp	Dữ liệu	Accuracy
Hansen và cs.	2013	Decision Tree	Landsat	~85%
Kussul và cs.	2017	CNN	Sentinel-2	94.5%
Reiche và cs.	2018	Bayesian fusion	S1+Landsat	93.8%
Hethcoat và cs.	2019	CNN (ResNet)	S1+S2	94.3%
Nguyen và cs.	2020	Random Forest	Sentinel-2	91.2%
Hu và cs.	2020	ML + fusion	S1+S2	92.0%

Nhận xét: Qua quá trình phát triển, các phương pháp giám sát rừng ngày càng tự động hóa, khách quan và chính xác hơn. Xu hướng hiện tại là kết hợp học sâu với dữ liệu đa nguồn (SAR + Optical) để tận dụng ưu điểm của cả hai. Nghiên cứu này kế thừa xu hướng đó, áp dụng CNN kết hợp dữ liệu Sentinel-1 (SAR) và Sentinel-2 (Optical) cho giám sát biến động rừng tại Cà Mau.

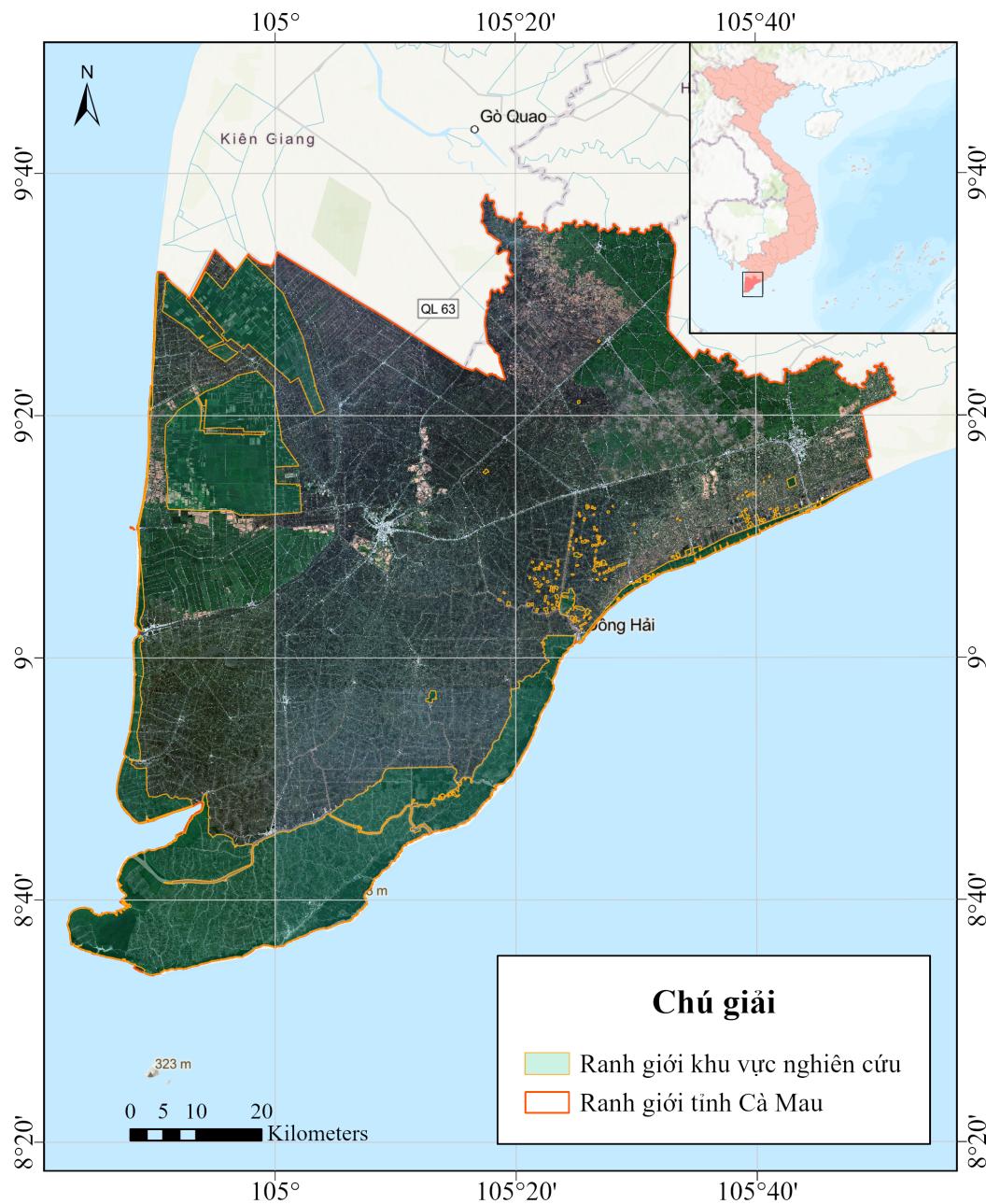
1.4. Khoảng trống nghiên cứu và định hướng đồ án

Qua tổng quan tài liệu, một số khoảng trống nghiên cứu nổi bật được xác định. Thứ nhất, thiếu nghiên cứu Deep Learning cho rừng nhiệt đới Việt Nam: phần lớn công trình tập trung ở Amazon, Congo hay Indonesia, còn ít nghiên cứu áp dụng CNN cho rừng Việt Nam, đặc biệt là rừng ngập mặn Cà Mau. Thứ hai, kiến trúc CNN cho bộ dữ liệu nhỏ: CNN thường yêu cầu tập dữ liệu lớn (hàng trăm nghìn mẫu), có ít công trình nghiên cứu về kiến trúc CNN tối ưu cho các bộ dữ liệu nhỏ trong viễn thám (khoảng 2,000–5,000 mẫu). Thứ ba, tích hợp SAR và Optical trong Deep Learning: việc tích hợp SAR và Optical trong bối cảnh Deep Learning vẫn còn nhiều thách thức và còn thiếu các khảo sát tối ưu hóa fusion trong kiến trúc CNN.

Xuất phát từ những khoảng trống nghiên cứu đã nêu, đồ án này hướng tới bốn mục tiêu chính. Thứ nhất, phát triển một kiến trúc CNN phù hợp cho các bộ dữ liệu nhỏ bằng cách thiết kế mô hình lightweight (xấp xỉ 36K tham số), áp dụng các kỹ thuật regularization mạnh như Batch Normalization, Dropout và Weight Decay. Thứ hai, triển khai một quy trình đánh giá khoa học chặt chẽ bao gồm việc sử dụng stratified random split, thực hiện 5-Fold Stratified Cross Validation và giữ lại một fixed test set (20%). Thứ ba, tối ưu hóa phương pháp fusion giữa Sentinel-1 và Sentinel-2 ở cấp độ feature, trích xuất 27 features tổng cộng (21 features từ S2 và 6 features từ S1). Thứ tư, ứng dụng thực tế tại Cà Mau, bao gồm phân loại toàn vùng quy hoạch lâm nghiệp (170,179 ha ranh giới, 162,469 ha diện tích phân loại thực tế), ước tính diện tích mất rừng và tạo bản đồ phân loại ở độ phân giải 10m.

1.5. Khu vực nghiên cứu

Đồ án tập trung vào toàn bộ vùng quy hoạch lâm nghiệp của tỉnh Cà Mau mới (đã giới thiệu tại Mục 1.1.3). Dữ liệu ranh giới quy hoạch lâm nghiệp được cung cấp bởi Công ty



Hình 1.7: Ảnh vệ tinh tổ hợp màu tự nhiên (RGB) khu vực nghiên cứu tỉnh Cà Mau.

Tổng diện tích ranh giới quy hoạch là 170,178.82 hecta (tương đương 1,701.79 km²), bao gồm 666 polygon trong file shapefile ranh giới. Diện tích thực tế được phân loại là 162,468.50 hecta (khoảng 95.5% diện tích ranh giới); phần còn lại (7,710 ha, chiếm 4.5%) bị loại do mây che phủ hoặc dữ liệu không hợp lệ (nodata) trong quá trình

xử lý ảnh vệ tinh. Kích thước raster là $12,547 \times 10,917$ pixels (ở độ phân giải 10m), sử dụng hệ quy chiếu EPSG:32648 (WGS 84 / UTM Zone 48N).

CHƯƠNG 2. CƠ SỞ LÝ THUYẾT VỀ VIỄN THÁM VÀ HỌC SÂU TRONG GIÁM SÁT BIỂN ĐỘNG RỪNG

2.1. Công nghệ viễn thám và ảnh vệ tinh

2.1.1. Nguyên lý viễn thám và các loại dữ liệu vệ tinh

Viễn thám (Remote Sensing) là khoa học và kỹ thuật thu thập thông tin về một đối tượng hoặc khu vực từ xa, thường thông qua việc ghi nhận bức xạ điện từ phản xạ hoặc phát ra từ bề mặt Trái Đất [?]. Nguyên lý cơ bản của viễn thám dựa trên tương tác giữa bức xạ điện từ và các đối tượng trên bề mặt.

Quá trình viễn thám bị động (Passive Remote Sensing):

Trong hệ thống viễn thám bị động, nguồn năng lượng chính là bức xạ từ Mặt Trời. Khi các sóng này truyền qua khí quyển, một phần năng lượng bị hấp thụ hoặc tán xạ. Sau đó bức xạ tương tác với bề mặt, chịu các quá trình phản xạ, hấp thụ hoặc truyền qua tùy theo đặc tính vật liệu. Tín hiệu phản xạ được vệ tinh ghi nhận bởi cảm biến và được xử lý, truyền về trạm mặt đất để phục vụ phân tích. Nguyên lý cân bằng năng lượng cho thấy năng lượng tới bằng tổng năng lượng phản xạ, hấp thụ và truyền qua.

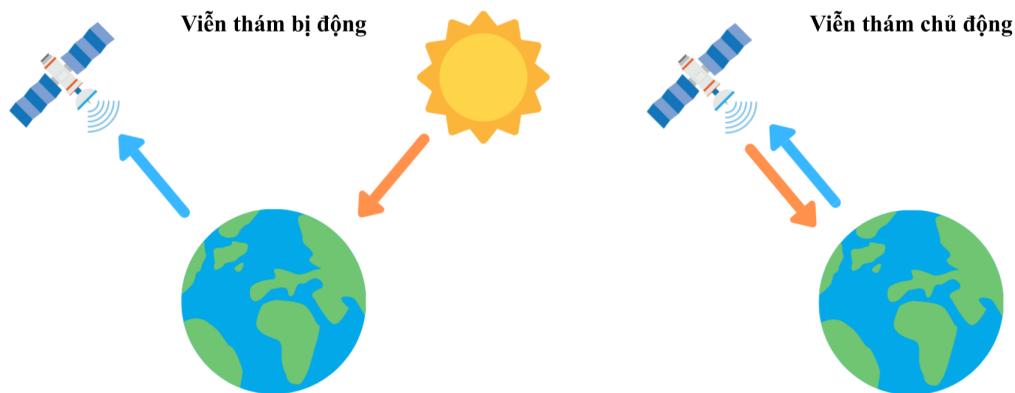
Quá trình viễn thám chủ động (Active Remote Sensing):

Khác với viễn thám bị động, hệ thống viễn thám chủ động tự phát ra nguồn năng lượng điện từ hướng về phía mục tiêu và ghi nhận tín hiệu phản xạ ngược (backscatter) từ bề mặt. Ưu điểm chính của viễn thám chủ động là khả năng hoạt động độc lập với ánh sáng Mặt Trời, cho phép thu thập dữ liệu cả ngày lẫn đêm và trong mọi điều kiện thời tiết (kể cả khi có mây che phủ). Radar khâu độ tổng hợp (SAR) là ví dụ điển hình của công nghệ viễn thám chủ động, sử dụng sóng vi ba (microwave) có khả năng xuyên qua mây và mưa.

So sánh viễn thám bị động và chủ động:

Viễn thám bị động (như Sentinel-2) cung cấp ảnh quang học đa phổ với độ phân giải không gian và phổ cao, phù hợp cho phân loại lớp phủ đất chi tiết, nhưng bị hạn chế bởi mây và điều kiện chiếu sáng. Viễn thám chủ động (như Sentinel-1 SAR) hoạt động trong mọi điều kiện thời tiết, cung cấp thông tin về cấu trúc và độ ẩm bề mặt, nhưng khó diễn giải hơn. Sự kết hợp cả hai loại dữ liệu (data fusion) cho phép tận dụng ưu điểm của

từng nguồn, đặc biệt quan trọng trong giám sát rừng nhiệt đới nơi mây che phủ thường xuyên.



Hình 2.1: Nguyên lý viễn thám bị động và chủ động

Radar khẩu độ tổng hợp (SAR):

SAR là công nghệ viễn thám chủ động đóng vai trò quan trọng trong giám sát rừng nhiệt đới nhờ khả năng quan sát trong mọi điều kiện thời tiết [? ?]. SAR là hệ thống chủ động: anten phát xung sóng điện từ về phía Trái Đất, các sóng này tương tác với bề mặt và tạo hiện tượng phản xạ ngược (backscatter) với cường độ phụ thuộc vào nhiều yếu tố như độ nhám bề mặt, hàm lượng nước (độ ẩm), hằng số điện môi và góc tới. Hệ số Backscatter (σ^0) thường được biểu diễn theo đơn vị decibel (dB), với giá trị phụ thuộc vào nhiều yếu tố. Về độ nhám bề mặt, bề mặt nhẵn như nước cho σ^0 thấp, trong khi bề mặt nhám như rùng cho σ^0 cao. Về hàm lượng nước, độ ẩm làm tăng σ^0 do hằng số điện môi lớn của nước. Về cấu trúc thực vật, khu vực rừng có cấu trúc phức tạp thường cho backscatter mạnh. SAR có thể phát và thu theo các chế độ phân cực khác nhau: VV (phát V, thu V) nhạy với độ ẩm bề mặt, VH (phát V, thu H) thường nhạy với cấu trúc thực vật (volume scattering).

Ảnh quang học đa phổ (Optical Multispectral):

Ảnh quang học từ các vệ tinh như Sentinel-2 là nguồn dữ liệu quan trọng cho phân loại lớp phủ đất và giám sát thực vật [? ?].

Dải phổ điện từ:

Ảnh quang học ghi nhận bức xạ phản xạ từ bề mặt Trái Đất ở các dải phổ khác nhau. Dải nhìn thấy (VIS) có bước sóng 400–700 nm, bao gồm Blue (450–520 nm),

Green (520–600 nm) và Red (630–690 nm). Dải cận hồng ngoại (NIR) có bước sóng 700–1400 nm, với đặc trưng phản xạ cao ở thực vật xanh do chlorophyll. Dải hồng ngoại sóng ngắn (SWIR) có bước sóng 1400–3000 nm, nhạy với độ ẩm của thực vật và đất.

Chữ ký phổ (Spectral Signature):

Mỗi loại đối tượng có chữ ký phổ đặc trưng - mẫu phản xạ qua các dải phổ khác nhau. Thực vật xanh có phản xạ thấp ở dải Red (hấp thụ bởi chlorophyll) và phản xạ cao ở dải NIR. Đất trống có phản xạ trung bình và tăng dần theo bước sóng. Nước có phản xạ thấp ở hầu hết các dải, đặc biệt là NIR và SWIR.

Bảng 2.1: Các dải phổ Sentinel-2 sử dụng trong nghiên cứu

Band	Tên	Bước sóng (nm)	Độ phân giải (m)	Ứng dụng
B4	Red	665	10	Chlorophyll absorption
B8	NIR	842	10	Biomass, NDVI
B11	SWIR1	1610	20	Độ ẩm, NDMI
B12	SWIR2	2190	20	NBR

2.1.2. Các chỉ số thực vật viễn thám

Các chỉ số thực vật (Vegetation Indices) là các công thức toán học kết hợp giá trị phản xạ từ các kênh phổ khác nhau để tăng cường thông tin về thực vật và giảm nhiễu từ các yếu tố khác như đất, khí quyển. Nghiên cứu này sử dụng ba chỉ số chính:

NDVI (Normalized Difference Vegetation Index):

NDVI là chỉ số được sử dụng phổ biến nhất để đánh giá sức khỏe và mật độ thực vật [?]:

$$NDVI = \frac{NIR - Red}{NIR + Red} = \frac{B8 - B4}{B8 + B4} \quad (2.1)$$

Thành phần: NIR (B8) là giá trị phản xạ ở dải cận hồng ngoại, Red (B4) là giá trị phản xạ ở dải đỏ. **Ý nghĩa:** Công thức khai thác sự khác biệt giữa phản xạ cao ở NIR và phản xạ thấp ở Red của thực vật xanh — chlorophyll hấp thụ mạnh ánh sáng đỏ để quang hợp nhưng phản xạ mạnh hồng ngoại gần. **Mục đích:** Định lượng mật độ và sức khỏe thực vật; giá trị NDVI cao (0.3–0.8) chỉ ra thực vật xanh khỏe mạnh, giá trị thấp (0.1–0.2) là đất trống, giá trị âm là nước.

NBR (Normalized Burn Ratio):

NBR được thiết kế ban đầu để phát hiện khu vực cháy rừng, nhưng cũng hiệu quả trong phát hiện mất rừng [?]:

$$NBR = \frac{NIR - SWIR2}{NIR + SWIR2} = \frac{B8 - B12}{B8 + B12} \quad (2.2)$$

Thành phần: *NIR* (B8) là giá trị phản xạ ở dải cận hồng ngoại, *SWIR2* (B12) là giá trị phản xạ ở dải hồng ngoại sóng ngắn 2 (2190 nm). **Ý nghĩa:** Kênh SWIR nhạy với độ ẩm và cấu trúc thực vật — khi rừng bị phá hoặc cháy, NIR giảm (do mất lá xanh) trong khi SWIR tăng (do bề mặt khô hơn). **Mục đích:** Phát hiện khu vực cháy rừng và mất rừng; NBR cao (0.3–0.8) chỉ ra rừng khỏe mạnh, NBR thấp hoặc âm chỉ ra khu vực bị tác động.

NDMI (Normalized Difference Moisture Index):

NDMI đo lường hàm lượng nước trong tán lá thực vật, là chỉ số quan trọng để phát hiện stress thực vật [?]:

$$NDMI = \frac{NIR - SWIR1}{NIR + SWIR1} = \frac{B8 - B11}{B8 + B11} \quad (2.3)$$

Thành phần: *NIR* (B8) là giá trị phản xạ ở dải cận hồng ngoại, *SWIR1* (B11) là giá trị phản xạ ở dải hồng ngoại sóng ngắn 1 (1610 nm). **Ý nghĩa:** Kênh SWIR1 nhạy với hàm lượng nước trong lá — nước hấp thụ mạnh ở dải này, nên thực vật có nhiều nước sẽ cho phản xạ SWIR thấp. **Mục đích:** Đánh giá độ ẩm tán lá và phát hiện sớm stress thực vật; NDMI dương (0.2–0.6) chỉ ra thực vật có hàm lượng nước cao, NDMI thấp hoặc âm chỉ ra thực vật bị stress hoặc khô. Đặc biệt, NDMI có thể phát hiện sự thay đổi độ ẩm trước khi NDVI phản ánh sự suy giảm sức khỏe thực vật.

Bảng 2.2: Tổng hợp các chỉ số thực vật sử dụng trong nghiên cứu

Chỉ số	Công thức	Phạm vi	Ý nghĩa
NDVI	(B8-B4)/(B8+B4)	[-1, 1]	Mật độ, sức khỏe thực vật
NBR	(B8-B12)/(B8+B12)	[-1, 1]	Phát hiện cháy/mất rừng
NDMI	(B8-B11)/(B8+B11)	[-1, 1]	Độ ẩm tán lá

2.1.3. Phát hiện biến động rừng

Phát hiện biến động rừng (Forest Change Detection) là quá trình xác định sự thay đổi về diện tích, cấu trúc hoặc trạng thái của rừng giữa hai hoặc nhiều thời điểm khác nhau [?]

]. Phương pháp này dựa trên việc so sánh các đặc trưng viễn thám thu được từ các thời điểm khác nhau.

Phương pháp phát hiện biến động dựa trên việc so sánh các đặc trưng giữa hai thời điểm, tính toán sự chênh lệch (delta) giữa “after features” và “before features”. Temporal features bao gồm các “before features” thể hiện trạng thái rừng tại thời điểm t_1 , các “after features” thể hiện trạng thái rừng tại thời điểm t_2 , và các “delta features” biểu diễn biến đổi giữa hai thời điểm.

Ví dụ với NDVI: khi $\Delta NDVI$ (hiệu NDVI sau và trước) giảm mạnh thì đó là dấu hiệu mất rừng; khi $\Delta NDVI$ xấp xỉ 0 thì vùng được xem là rừng ổn định; và khi $\Delta NDVI$ tăng mạnh thì biểu hiện tái trồng rừng.

2.2. Mạng Neural Tích chập (Convolutional Neural Networks)

Mạng neural tích chập (CNN) đã trở thành công cụ chủ đạo trong xử lý ảnh viễn thám và phân loại lớp phủ đất [? ?]. Phần này trình bày các khái niệm nền tảng về CNN làm cơ sở cho kiến trúc đề xuất trong Chương 3.

2.2.1. Kiến trúc cơ bản của CNN

Perceptron - Đơn vị cơ bản:

Một neuron nhân tạo thực hiện phép biến đổi tuyến tính và hàm kích hoạt:

$$y = f(\mathbf{w}^T \mathbf{x} + b) \quad (2.4)$$

Thành phần: $\mathbf{x} \in \mathbb{R}^n$ là input vector chứa n feature đầu vào; $\mathbf{w} \in \mathbb{R}^n$ là weight vector (trọng số) — mỗi w_i thể hiện mức độ quan trọng của feature x_i ; $b \in \mathbb{R}$ là bias (độ lệch) — cho phép dịch chuyển đường quyết định; $f(\cdot)$ là hàm kích hoạt phi tuyến; y là output (đầu ra). **Ý nghĩa:** Perceptron mô phỏng cách neuron sinh học xử lý tín hiệu — nhận nhiều đầu vào, tính tổng có trọng số, và kích hoạt khi vượt ngưỡng. **Mục đích:** Là đơn vị tính toán cơ bản của mạng neural, có khả năng học các mối quan hệ tuyến tính và phi tuyến trong dữ liệu thông qua việc điều chỉnh trọng số \mathbf{w} và bias b .

Multi-Layer Perceptron (MLP):

Một mạng neural gồm nhiều layers, trong đó mỗi layer nhận đầu ra của layer trước

làm đầu vào. Mỗi layer thực hiện phép biến đổi tuyến tính (nhân ma trận trọng số \mathbf{W} và cộng bias \mathbf{b}) sau đó áp dụng hàm kích hoạt. Các layer ẩn giúp mạng học được các biểu diễn trừu tượng của dữ liệu, với mỗi layer học các đặc trưng ở mức độ phức tạp tăng dần.

Phép tích chập 2D (2D Convolution):

Đây là thành phần cốt lõi của CNN, thực hiện phép tích chập giữa input và kernel:

$$(I * K)(i, j) = \sum_m \sum_n I(i + m, j + n) \times K(m, n) \quad (2.5)$$

Thành phần: I là input feature map (ảnh hoặc feature map từ layer trước) với kích thước $height \times width \times channels$; K là kernel (filter) có kích thước $k_h \times k_w$ — chứa các trọng số được học; (i, j) là vị trí trong output feature map; (m, n) là vị trí trong kernel. **Ý nghĩa:** Phép tích chập trượt kernel qua input, tại mỗi vị trí tính tổng các tích element-wise giữa kernel và vùng tương ứng của input — tạo ra một giá trị trong output thể hiện mức độ “khớp” giữa kernel và vùng đó. **Mục đích:** Trích xuất các đặc trưng cục bộ (local features) như cạnh, góc, texture từ input; nhờ chia sẻ trọng số (parameter sharing), CNN tiết kiệm tham số và có khả năng nhận diện đặc trưng bất kể vị trí (translation invariance).

Phép tích chập mang lại nhiều ưu điểm quan trọng. Về parameter sharing, cùng một kernel áp dụng toàn bộ input giúp tiết kiệm tham số. Về translation invariance, mạng có khả năng nhận diện các đặc trưng bất kể vị trí xuất hiện. Về local connectivity, mỗi neuron chỉ kết nối với vùng cục bộ của input, giảm số lượng kết nối cần thiết.

Quá trình tích chập diễn ra như sau: kernel có kích thước $k \times k$ trượt qua toàn bộ input với bước nhảy (stride) xác định. Tại mỗi vị trí, kernel thực hiện phép nhân element-wise với vùng tương ứng của input, sau đó cộng tổng tất cả các tích để tạo ra một giá trị trong output feature map. Kích thước output phụ thuộc vào kích thước input, kernel, stride và padding. Với input nhiều kênh (ví dụ: feature stack có 27 kênh như trong nghiên cứu này), kernel cũng có cùng số kênh và phép tích chập được thực hiện đồng thời trên tất cả các kênh, sau đó cộng tổng để tạo ra một feature map đầu ra.

Hàm kích hoạt (Activation Functions):

Hàm kích hoạt đóng vai trò quan trọng trong mạng neural, giúp mô hình học được các mối quan hệ phi tuyến trong dữ liệu [?]. ReLU (Rectified Linear Unit) là hàm kích hoạt phổ biến nhất:

$$f(x) = \max(0, x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \quad (2.6)$$

Thành phần: x là giá trị đầu vào (output của phép biến đổi tuyến tính). **Ý nghĩa:** ReLU giữ nguyên giá trị dương và chuyển giá trị âm về 0, tạo ra tính phi tuyến đơn giản nhưng hiệu quả. **Mục đích:** Cho phép mạng học các mối quan hệ phi tuyến phức tạp; ReLU có ưu điểm tính toán nhanh, giảm vấn đề vanishing gradient (gradient không bị triệt tiêu với giá trị dương), và tạo sparse activation (nhiều neuron có output = 0).

Softmax được sử dụng cho lớp đầu ra trong bài toán phân loại đa lớp:

$$\text{softmax}(\mathbf{x})_i = \frac{\exp(x_i)}{\sum_j \exp(x_j)} \quad (2.7)$$

Thành phần: \mathbf{x} là vector logits (đầu ra thô của layer cuối cùng); x_i là logit của lớp i ; $\exp(\cdot)$ là hàm mũ. **Ý nghĩa:** Softmax chuyển đổi các logits thành phân phối xác suất — giá trị lớn hơn được ánh xạ thành xác suất cao hơn, và tổng tất cả xác suất bằng 1. **Mục đích:** Cho phép mô hình đưa ra dự đoán xác suất cho bài toán phân loại đa lớp; lớp có xác suất cao nhất được chọn làm kết quả dự đoán.

2.2.2. Pooling và các kỹ thuật Regularization

Pooling là phép toán giảm chiều không gian (spatial dimensions) của feature maps, giúp giảm số lượng tham số, giảm chi phí tính toán và tăng khả năng translation invariance của mạng [?]. Có hai loại pooling phổ biến: **Max Pooling** chọn giá trị lớn nhất trong mỗi vùng cục bộ, giữ lại các đặc trưng nổi bật nhất; **Average Pooling** tính trung bình các giá trị, giữ lại thông tin tổng quát hơn.

Global Average Pooling (GAP) là trường hợp đặc biệt trong đó vùng pooling bao trùm toàn bộ feature map, biến đổi mỗi feature map thành một giá trị scalar [?]. GAP mang lại nhiều lợi ích: giảm mạnh số lượng tham số, có tính chất regularization tự nhiên, và tạo ra spatial invariance hoàn toàn. Trong nghiên cứu này, GAP được sử dụng ở cuối mạng CNN để tổng hợp thông tin từ tất cả các spatial locations trước khi đưa vào lớp classification.

Batch Normalization:

Batch Normalization là kỹ thuật chuẩn hóa được đề xuất bởi Ioffe và Szegedy nhằm tăng tốc quá trình huấn luyện mạng neural sâu [?]. Kỹ thuật này chuẩn hóa các activation trong mỗi mini-batch về phân phối có trung bình 0 và phương sai 1, sau đó áp dụng phép biến đổi tuyến tính với các tham số học được. Batch Normalization giúp tăng tốc độ

huấn luyện, giảm độ nhạy với khởi tạo trọng số, và đóng vai trò như một phương pháp regularization.

Dropout:

Dropout là kỹ thuật regularization hiệu quả được đề xuất bởi Srivastava và cộng sự để ngăn chặn overfitting trong mạng neural [?]. Trong quá trình huấn luyện, mỗi neuron bị tắt ngẫu nhiên với một xác suất nhất định, buộc mạng học các biểu diễn robust hơn. Dropout2d (Spatial Dropout) là biến thể dropout toàn bộ feature maps thay vì từng neuron, phù hợp cho CNN vì features trong cùng channel có correlation không gian cao.

2.2.3. Hàm mất mát và thuật toán tối ưu

Hàm mất mát (Loss Function) đo lường sự khác biệt giữa dự đoán của mô hình và nhãn thực tế, là cơ sở để tối ưu hóa tham số mạng [?]. Cross-Entropy Loss được sử dụng cho bài toán phân loại đa lớp:

$$L = - \sum_i y_i \log(\hat{y}_i) \quad (2.8)$$

Thành phần: y_i là nhãn thực tế (true label) được mã hóa one-hot — có giá trị 1 tại lớp đúng và 0 ở các lớp khác; \hat{y}_i là xác suất dự đoán của lớp i xuất ra từ hàm softmax; $\log(\cdot)$ là logarit tự nhiên. **Ý nghĩa:** Cross-Entropy đo lường “khoảng cách” giữa phân phối dự đoán và phân phối thực tế — giá trị loss cao khi mô hình gán xác suất thấp cho lớp đúng, và ngược lại. **Mục đích:** Làm hàm mục tiêu để huấn luyện mô hình; thuật toán tối ưu sẽ điều chỉnh trọng số để giảm thiểu loss, từ đó tăng xác suất dự đoán đúng.

Thuật toán tối ưu:

Các thuật toán tối ưu hóa đóng vai trò quan trọng trong việc huấn luyện mạng neural. **Adam** (Adaptive Moment Estimation) được đề xuất bởi Kingma và Ba là một trong những optimizer phổ biến nhất hiện nay [?]. Adam kết hợp ưu điểm của hai phương pháp: momentum (sử dụng trung bình động của gradient) và RMSprop (điều chỉnh learning rate theo từng tham số). **AdamW** là biến thể cải tiến với weight decay được tách riêng khỏi gradient update, giúp regularization hiệu quả hơn. Các hyperparameters thường sử dụng: $\beta_1 = 0.9$, $\beta_2 = 0.999$, learning rate $\eta = 0.001$.

2.3. Phương pháp phân loại ảnh viễn thám

2.3.1. Các phương pháp phân loại ảnh

Trong phân loại ảnh viễn thám, có hai phương pháp tiếp cận chính: phân loại dựa trên pixel và phân loại dựa trên patch [? ?].

Pixel-based Classification: Mỗi pixel được phân loại độc lập dựa trên vector đặc trưng của riêng nó. Ưu điểm là đơn giản, dễ triển khai, tốc độ xử lý nhanh. Nhược điểm là không tận dụng ngữ cảnh không gian, dễ tạo ra nhiều dạng salt-and-pepper.

Patch-based Classification: Trích xuất các patches (cửa sổ) xung quanh mỗi pixel và phân loại dựa trên toàn bộ patch. Ưu điểm là sử dụng được ngữ cảnh không gian, kết quả mượt hơn và phù hợp với CNN.

2.3.2. Ma trận nhầm lẫn và các độ đo đánh giá

Confusion Matrix (Ma trận nhầm lẫn) là công cụ cơ bản để đánh giá hiệu suất của mô hình phân loại, đặc biệt quan trọng trong các bài toán phân loại ảnh viễn thám [?]. Ma trận này tổng hợp kết quả dự đoán của mô hình so với nhãn thực tế.

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

Bảng 2.3: Cấu trúc Confusion Matrix nhị phân

Ma trận nhầm lẫn bao gồm bốn thành phần chính. **True Positive (TP)** là số mẫu dương được dự đoán đúng là dương — trong bài toán phát hiện mắt rừng, đây là các điểm thực sự mắt rừng và mô hình dự đoán đúng. **True Negative (TN)** là số mẫu âm được dự đoán đúng là âm. **False Positive (FP)** là số mẫu âm bị dự đoán nhầm là dương (lỗi loại I). **False Negative (FN)** là số mẫu dương bị dự đoán nhầm là âm (lỗi loại II). Với bài toán phân loại K lớp, Confusion Matrix có kích thước $K \times K$, trong đó phần tử C_{ij} thể hiện số mẫu thuộc lớp i được dự đoán là lớp j .

Các độ đo đánh giá:

Các chỉ số đánh giá được tính toán dựa trên Confusion Matrix để đo lường hiệu

suất phân loại từ nhiều góc độ khác nhau [?].

Accuracy (Độ chính xác):

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.9)$$

Thành phần: TP (True Positive), TN (True Negative), FP (False Positive), FN (False Negative) như định nghĩa ở trên. **Ý nghĩa:** Tỷ lệ các mẫu được phân loại đúng trên tổng số mẫu. **Mục đích:** Đánh giá hiệu suất tổng thể của mô hình; tuy nhiên, accuracy có thể gây hiểu lầm với dữ liệu mất cân bằng.

Precision (Độ chính xác dương):

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2.10)$$

Thành phần: TP là số mẫu dương được dự đoán đúng, FP là số mẫu âm bị dự đoán nhầm là dương. **Ý nghĩa:** Trong số các mẫu được dự đoán là dương, tỷ lệ thực sự là dương. **Mục đích:** Đánh giá độ tin cậy của dự đoán dương; precision cao quan trọng khi chi phí của false positive cao (ví dụ: cảnh báo sai về mất rừng gây lãng phí nguồn lực kiểm tra).

Recall (Độ nhạy):

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2.11)$$

Thành phần: TP là số mẫu dương được dự đoán đúng, FN là số mẫu dương bị bỏ sót. **Ý nghĩa:** Trong số các mẫu thực sự dương, tỷ lệ được phát hiện đúng. **Mục đích:** Đánh giá khả năng phát hiện các mẫu dương; recall cao quan trọng khi chi phí của false negative cao (ví dụ: bỏ sót vùng mất rừng thực sự gây hậu quả nghiêm trọng về môi trường).

F1-Score (Trung bình điều hòa):

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (2.12)$$

Thành phần: Precision và Recall như định nghĩa ở trên. **Ý nghĩa:** F1-Score là trung bình điều hòa của Precision và Recall — cân bằng giữa hai độ đo này, và bị ảnh hưởng mạnh bởi giá trị thấp hơn. **Mục đích:** Đánh giá tổng hợp khi cần cân bằng giữa precision và recall; F1-Score đặc biệt hữu ích với dữ liệu mất cân bằng.

ROC-AUC (Area Under ROC Curve):

Các tiêu chuẩn diễn giải ROC-AUC theo Hosmer và Lemeshow [?]: $AUC = 0.5$ tương ứng với classifier ngẫu nhiên; $0.5 < AUC < 0.7$ là phân biệt kém; $0.7 \leq AUC < 0.8$ là chấp nhận được; $0.8 \leq AUC < 0.9$ là xuất sắc; và $AUC \geq 0.9$ là vượt trội.

2.3.3. Cross Validation, chuẩn hóa dữ liệu và rò rỉ dữ liệu

Cross Validation (Kiểm định chéo):

Cross Validation là phương pháp đánh giá mô hình giúp ước lượng khả năng generalization của mô hình trên dữ liệu chưa thấy, đồng thời giảm thiểu bias do cách chia dữ liệu [?]. Trong K-Fold Cross Validation, dữ liệu được chia thành K phần bằng nhau, mỗi vòng lặp sử dụng một fold làm tập kiểm tra và $K - 1$ folds còn lại làm tập huấn luyện; kết quả cuối cùng là trung bình của K lần đánh giá. Stratified K-Fold là biến thể đảm bảo tỷ lệ các lớp trong mỗi fold tương đương với tỷ lệ trong toàn bộ tập dữ liệu, đặc biệt quan trọng khi dữ liệu mất cân bằng giữa các lớp.

Chuẩn hóa dữ liệu (Data Normalization):

Chuẩn hóa dữ liệu là bước tiền xử lý quan trọng trong học máy, giúp các features có cùng scale và cải thiện hiệu suất huấn luyện [?]. Z-score Normalization chuyển đổi dữ liệu về phân phối với trung bình bằng 0 và độ lệch chuẩn bằng 1 (chi tiết công thức được trình bày trong Chương 3). Việc chuẩn hóa dữ liệu cho CNN là cần thiết vì các features có scale khác nhau (ví dụ: NDVI trong $[-1, 1]$, backscatter trong $[-25, 0]$ dB) sẽ ảnh hưởng không đều đến gradient. **Lưu ý quan trọng:** Các tham số chuẩn hóa phải được tính trên tập huấn luyện và áp dụng cho cả tập kiểm tra để tránh rò rỉ dữ liệu.

Dữ liệu thực địa (Ground Truth):

Dữ liệu thực địa là tập dữ liệu tham chiếu với nhãn chính xác, được sử dụng để huấn luyện và đánh giá mô hình phân loại [?]. Trong viễn thám, dữ liệu thực địa có thể thu thập từ khảo sát thực địa, diễn giải ảnh độ phân giải cao, hoặc dữ liệu lịch sử. Chất lượng dữ liệu thực địa ảnh hưởng trực tiếp đến độ tin cậy của kết quả phân loại.

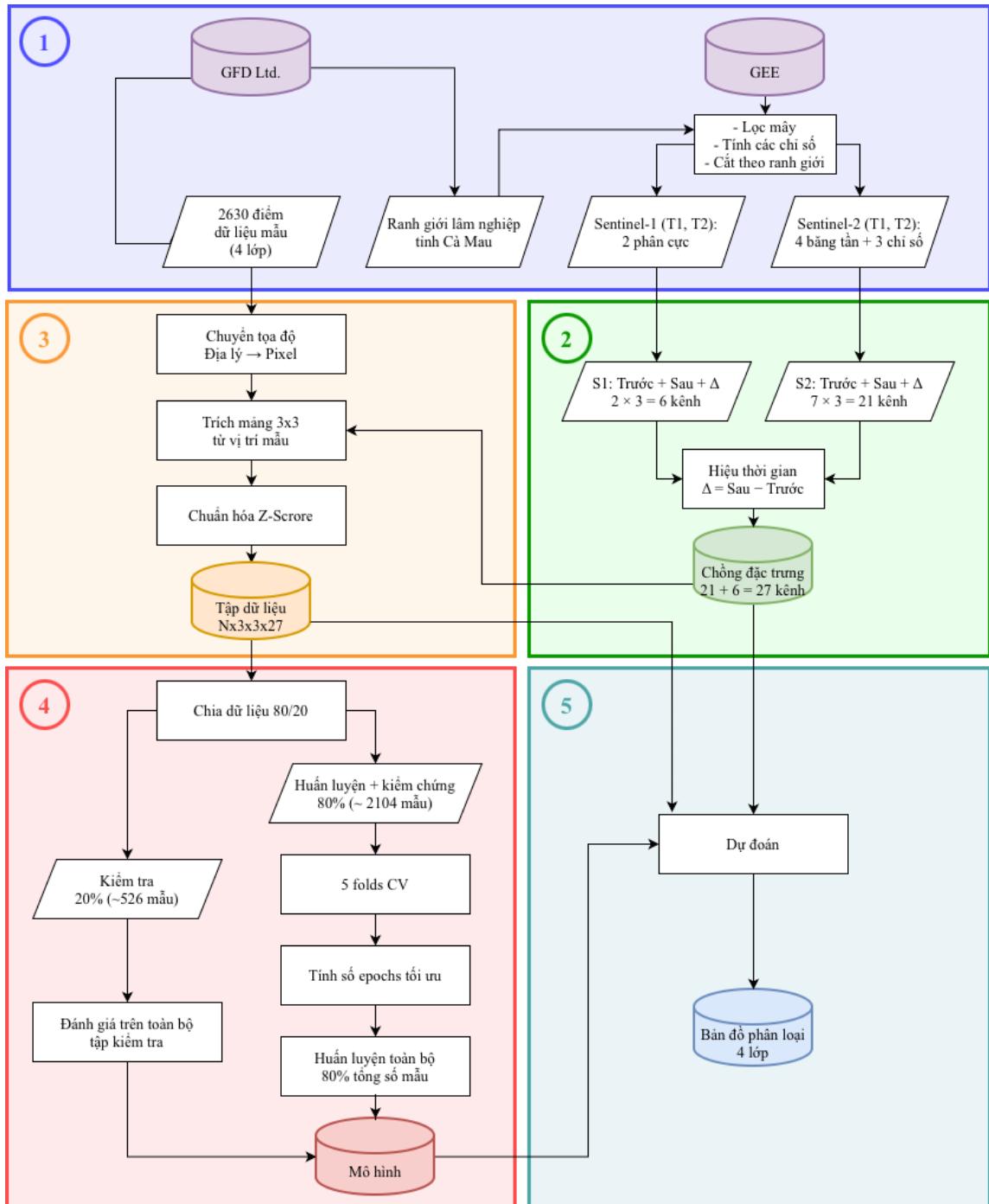
Rò rỉ dữ liệu (Data Leakage):

Rò rỉ dữ liệu xảy ra khi thông tin từ tập kiểm tra “rò rỉ” vào quá trình huấn luyện, dẫn đến kết quả đánh giá quá lạc quan [?]. Có ba dạng rò rỉ phổ biến trong viễn thám: (1) **Rò rỉ không gian** khi các điểm train/test nằm gần nhau về địa lý; (2) **Rò rỉ thời gian** khi sử dụng thông tin từ thời điểm sau để dự đoán thời điểm trước; (3) **Rò rỉ đặc trưng** khi

tính statistics trên toàn bộ dữ liệu thay vì chỉ trên tập huấn luyện. Trong nghiên cứu này, rò rỉ dữ liệu được phòng tránh bằng cách sử dụng Stratified K-Fold Cross Validation và tính toán các tham số chuẩn hóa riêng cho từng fold.

CHƯƠNG 3. DỮ LIỆU, TRÍCH XUẤT ĐẶC TRƯNG VÀ XÂY DỰNG MÔ HÌNH PHÂN LOẠI

Chương này trình bày phương pháp nghiên cứu và quy trình được minh họa trong Hình 3.1. Bao gồm 5 giai đoạn chính: (1) Thu thập và tiền xử lý dữ liệu, (2) Trích xuất đặc trưng, (3) Chuẩn bị mẫu huấn luyện, (4) Huấn luyện mô hình, và (5) Áp dụng mô hình.



Hình 3.1: Sơ đồ quy trình phương pháp nghiên cứu phát hiện biến động rừng

3.1. Thu thập và tiền xử lý dữ liệu

3.1.1. Dữ liệu viễn thám Sentinel

Việc lựa chọn ảnh vệ tinh tuân theo các tiêu chí nhằm đảm bảo chất lượng dữ liệu đầu vào. Đối với ảnh Sentinel-2, nghiên cứu sử dụng sản phẩm S2_SR_HARMONIZED (Surface Reflectance Level-2A đã đồng nhất hóa) từ Google Earth Engine, ưu tiên các ảnh trong mùa khô (tháng 1-3) để giảm thiểu ảnh hưởng của mây và đảm bảo tính so sánh giữa hai thời kỳ. Mặt nạ mây được tạo từ bộ sưu tập S2_CLOUD_PROBABILITY với ngưỡng xác suất 50% để loại bỏ các pixel bị mây che phủ. Đối với ảnh Sentinel-1, sử dụng sản phẩm S1_GRD đã được tiền xử lý bởi ESA, các cảnh được chọn có thời gian thu nhận gần nhất với ảnh Sentinel-2 tương ứng (trong phạm vi ±7 ngày) để đảm bảo tính đồng bộ về thời gian.

Bảng 3.1: Tổng quan dữ liệu viễn thám sử dụng

Nguồn dữ liệu	Độ phân giải	Kỳ ảnh	Số bands	Ghi chú
Sentinel-2 Before	10m	30/01/2024	7	Level-2A (SR), cloud <10%
Sentinel-2 After	10m	28/02/2025	7	Level-2A (SR), cloud <10%
Sentinel-1 Before	10m	04/02/2024	2	GRD, IW mode
Sentinel-1 After	10m	22/02/2025	2	GRD, IW mode
Dữ liệu thực địa	-	-	-	2,630 points
Forest Boundary	Vector	-	-	Shapefile

3.1.2. Thu thập dữ liệu trên Google Earth Engine

Toàn bộ dữ liệu ảnh vệ tinh được thu thập và xử lý trên nền tảng Google Earth Engine (GEE) — một hệ thống điện toán đám mây cho phép truy cập và xử lý khối lượng lớn dữ liệu viễn thám. Việc sử dụng GEE mang lại nhiều ưu điểm: truy cập trực tiếp kho dữ liệu Sentinel đã được tiền xử lý, khả năng xử lý song song trên hạ tầng đám mây, và đảm bảo tính nhất quán trong quy trình xử lý.

Xử lý dữ liệu Sentinel-2:

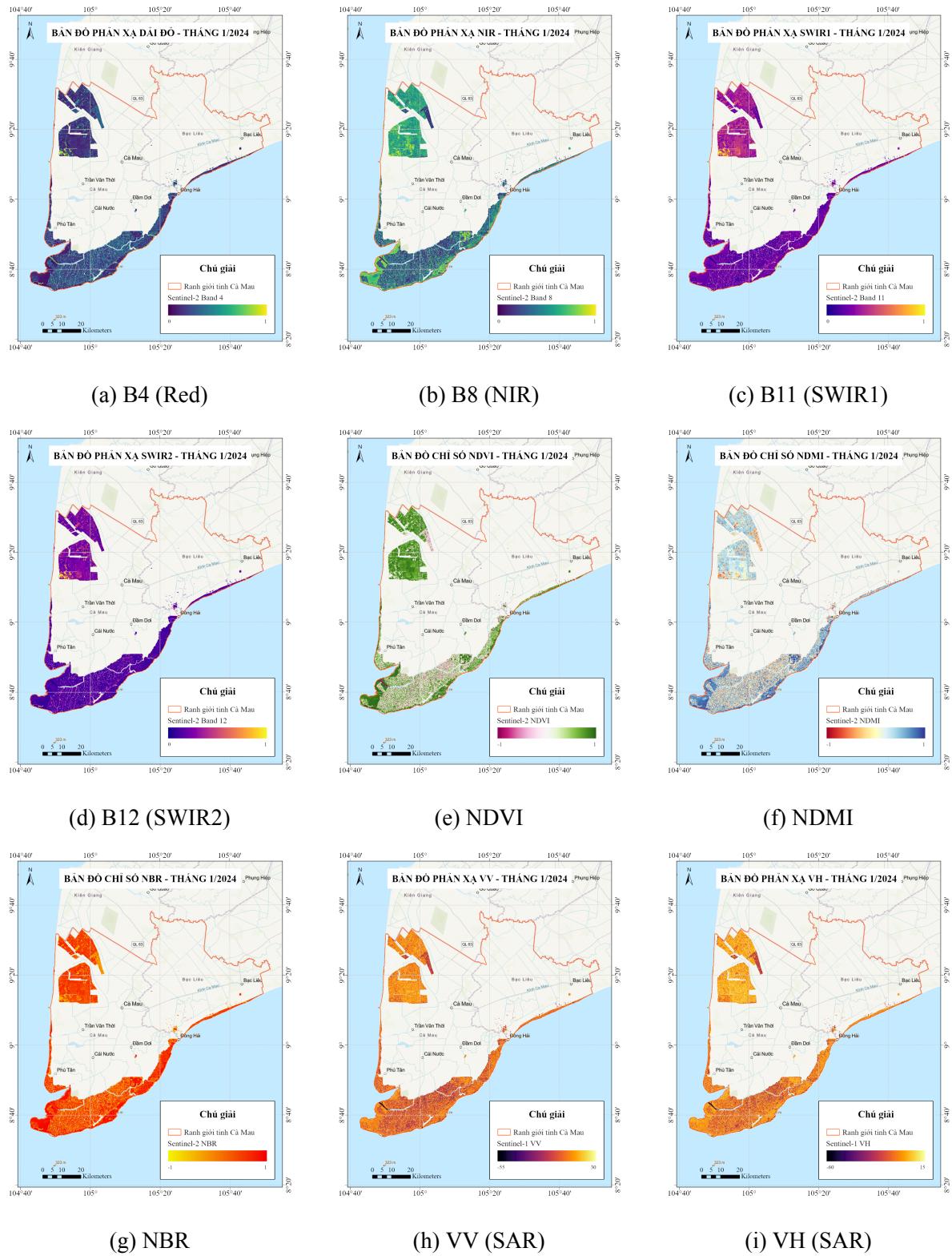
Dữ liệu Sentinel-2 được truy xuất từ bộ sưu tập COPERNICUS/S2_SR_HARMONIZED — sản phẩm Surface Reflectance Level-2A đã được hiệu chỉnh khí quyển và đồng nhất

hóa giữa các cảm biến Sentinel-2A và 2B. Quy trình xử lý bao gồm: (1) Lọc theo không gian và thời gian để chọn các cảnh phủ khu vực nghiên cứu trong ngày chỉ định; (2) Loại bỏ mây sử dụng bộ sưu tập S2_CLOUD_PROBABILITY với ngưỡng xác suất 50%; (3) Trích xuất 4 bands cần thiết (B4-Red, B8-NIR, B11-SWIR1, B12-SWIR2), chuyển đổi sang giá trị phản xạ và tính toán 3 chỉ số thực vật: NDVI, NBR, NDMI; (4) Mosaic các tiles để tạo ảnh liền mạch phủ toàn bộ khu vực nghiên cứu.

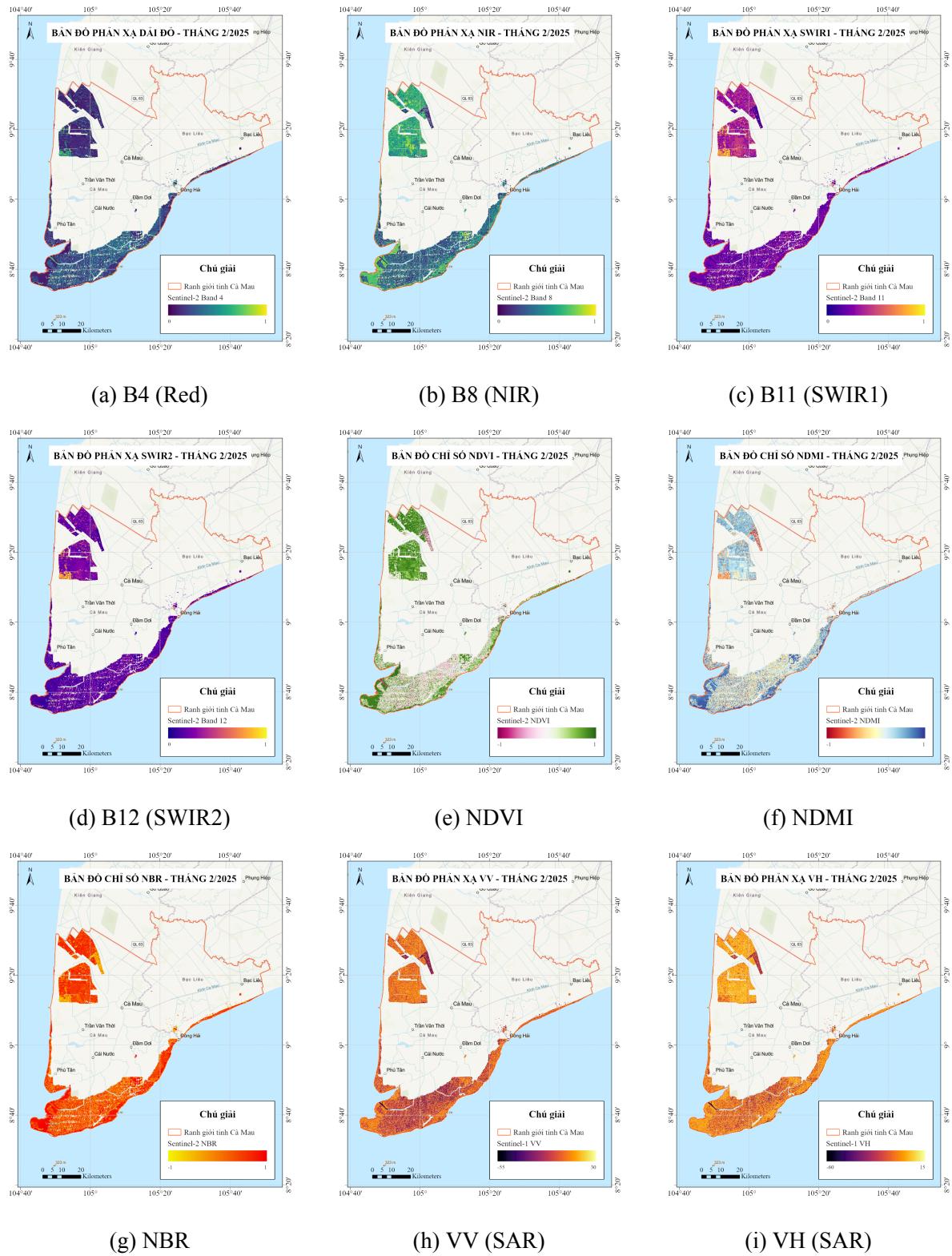
Xử lý dữ liệu Sentinel-1:

Dữ liệu Sentinel-1 được truy xuất từ bộ sưu tập COPERNICUS/S1_GRD — sản phẩm Ground Range Detected đã được tiền xử lý bởi ESA bao gồm: hiệu chỉnh quỹ đạo, loại bỏ nhiễu biên và nhiễu nhiệt, hiệu chỉnh bức xạ và hiệu chỉnh địa hình sử dụng DEM SRTM. Quy trình xử lý bổ sung bao gồm: lọc theo thời gian (± 7 ngày so với Sentinel-2), chọn chế độ Interferometric Wide (IW) với quỹ đạo đi xuống, và trích xuất hai bands VV và VH (đơn vị dB).

Sau khi xử lý, dữ liệu được xuất ra định dạng GeoTIFF với độ phân giải 10m và hệ quy chiếu EPSG:32648 (WGS 84 / UTM Zone 48N). Hình 3.2 và 3.3 minh họa trực quan các bands quang học (B4, B8, B11, B12), các chỉ số thực vật (NDVI, NDMI, NBR) và dữ liệu SAR (VV, VH) cho hai thời điểm trước và sau biến động.



Hình 3.2: Các bands quang học, chỉ số thực vật và dữ liệu SAR kỳ trước (T1 - 01/2024)



Hình 3.3: Các bands quang học, chỉ số thực vật và dữ liệu SAR kỳ sau (T2 - 02/2025)

3.1.3. Dữ liệu thực địa

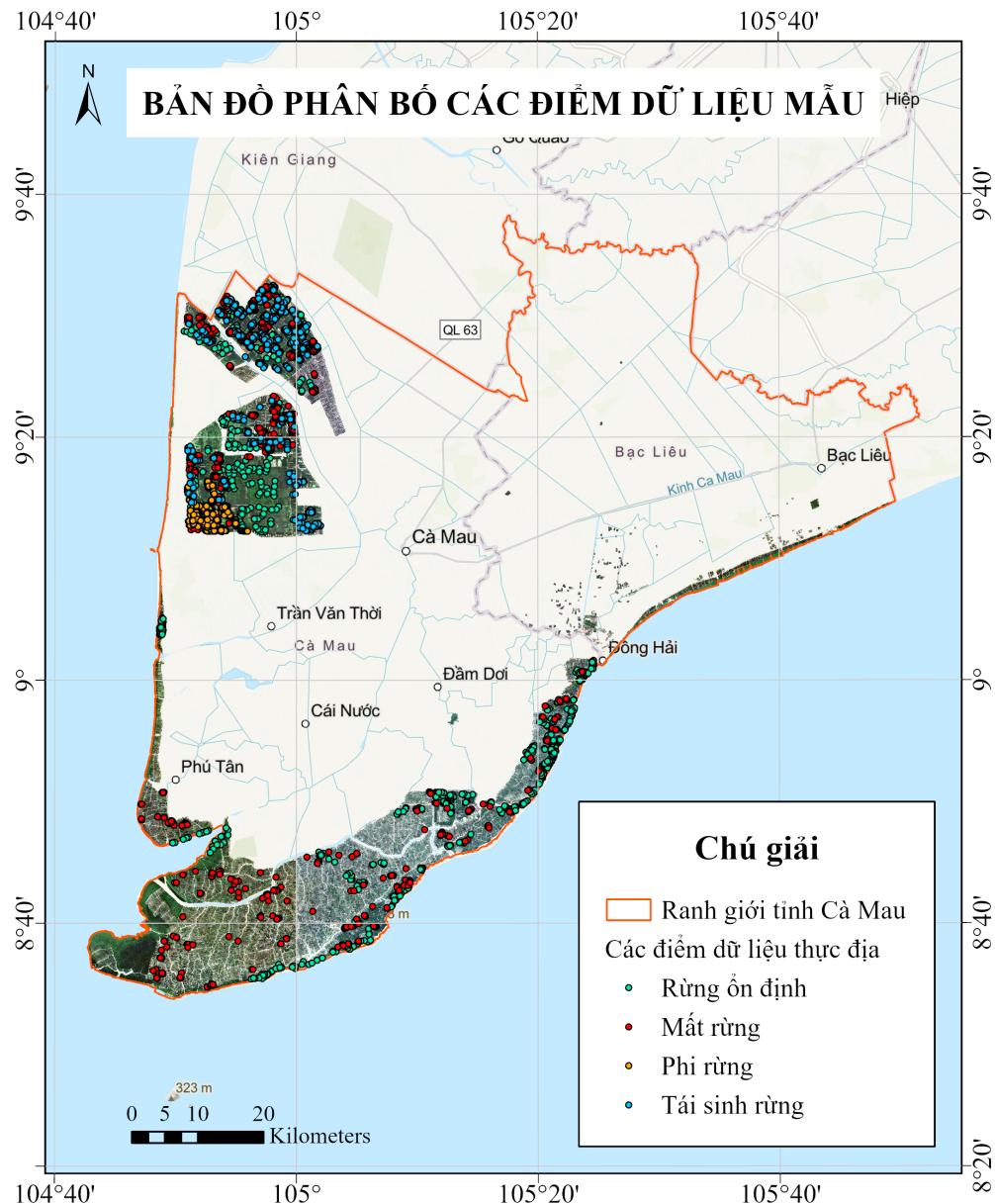
Dữ liệu thực địa được thu thập thông qua quy trình khảo sát chuyên nghiệp với sự phối hợp giữa Chi cục Kiểm lâm tỉnh Cà Mau và Công ty TNHH Tư vấn và Phát triển Đồng Xanh (GFD). Quy trình thu thập bao gồm ba giai đoạn: (1) Khảo sát thực địa bằng thiết bị bay không người lái (drone) để ghi nhận hình ảnh và xác định trạng thái rừng; (2) Số hóa các điểm dữ liệu thực địa trên phần mềm QGIS, đối chiếu với ảnh vệ tinh Sentinel-2; (3) Kiểm tra chéo và loại bỏ các điểm không rõ ràng hoặc nằm trong vùng mây che phủ.

Dữ liệu cuối cùng được xuất dưới dạng file CSV với cấu trúc gồm 4 trường: id, label, x và y (tọa độ theo hệ quy chiếu EPSG:32648).

Bảng 3.2: Thông kê dữ liệu thực địa theo lớp biến động

Lớp	Tên	Số điểm	Tỷ lệ	Mô tả
0	Rừng ổn định	656	24.9%	Có rừng ở cả 2 kỳ
1	Mất rừng	650	24.7%	Có rừng → không có rừng
2	Phi rừng	664	25.3%	Không có rừng ở cả 2 kỳ
3	Phục hồi rừng	660	25.1%	Không có → có rừng
Tổng		2,630	100%	Phân bố cân bằng

Phân bố không gian của các điểm dữ liệu thực địa được minh họa trong Hình 3.4. Các điểm mẫu được thu thập phân tán trên toàn bộ khu vực nghiên cứu, đảm bảo tính đại diện về mặt không gian cho các loại hình biến động rừng khác nhau.



Hình 3.4: Bản đồ phân bố không gian các điểm dữ liệu thực địa trên khu vực nghiên cứu tỉnh Cà Mau

3.2. Trích xuất đặc trưng

3.2.1. Xây dựng feature stack

Việc kết hợp dữ liệu SAR và quang học (data fusion) đã được chứng minh là hiệu quả trong nhiều nghiên cứu phân loại lớp phủ đất [? ?]. Cách tiếp cận này tận dụng ưu điểm

bổ sung của hai nguồn dữ liệu: SAR cung cấp thông tin về cấu trúc và độ ẩm bề mặt, trong khi quang học cung cấp thông tin về đặc tính quang phổ của thực vật.

Tổng cộng 27 features được xây dựng từ hai nguồn dữ liệu. **Sentinel-2** đóng góp 21 features, bao gồm 7 bands/chỉ số (B4, B8, B11, B12, NDVI, NBR, NDMI) cho kỳ trước, 7 bands/chỉ số tương ứng cho kỳ sau, và 7 giá trị delta (hiệu số giữa kỳ sau và kỳ trước). **Sentinel-1** đóng góp 6 features, bao gồm 2 bands (VV, VH) cho kỳ trước, 2 bands tương ứng cho kỳ sau, và 2 giá trị delta.

Bảng 3.3: Chi tiết 27 đặc trưng sử dụng trong mô hình

Chỉ số	Nguồn	Thời kỳ	Đặc trưng	Mô tả
0-6	S2	Kỳ trước	B4, B8, B11, B12, NDVI, NBR, NDMI	Quang phổ kỳ trước
7-13	S2	Kỳ sau	B4, B8, B11, B12, NDVI, NBR, NDMI	Quang phổ kỳ sau
14-20	S2	Biến đổi	$\Delta B4, \Delta B8, \dots$	Biến đổi quang phổ
21-22	S1	Kỳ trước	VV, VH	SAR kỳ trước
23-24	S1	Kỳ sau	VV, VH	SAR kỳ sau
25-26	S1	Biến đổi	$\Delta VV, \Delta VH$	Biến đổi SAR

3.2.2. Trích xuất patch 3×3

Với mỗi điểm thực địa, một patch kích thước 3×3 pixels được trích xuất từ feature stack. Kích thước 3×3 được lựa chọn vì cho phép mô hình học được thông tin ngữ cảnh không gian xung quanh pixel trung tâm, phù hợp với độ phân giải 10m của Sentinel (mỗi patch tương đương vùng $30m \times 30m$), và giảm thiểu nhiễu từ các pixel lân cận không đồng nhất.

Kết quả là mỗi mẫu có kích thước $(3, 3, 27)$ — tương ứng với chiều cao, chiều rộng và số kênh đặc trưng.

3.3. Chuẩn bị mẫu huấn luyện

3.3.1. Chuẩn hóa dữ liệu

Việc chuẩn hóa dữ liệu là bước quan trọng để đảm bảo các features có cùng phạm vi giá trị, giúp quá trình huấn luyện mô hình hội tụ nhanh và ổn định hơn. Nghiên cứu này áp

dụng phương pháp chuẩn hóa Z-score:

$$x_{normalized} = \frac{x - \mu}{\sigma} \quad (3.1)$$

trong đó x là giá trị gốc, μ là giá trị trung bình và σ là độ lệch chuẩn.

Để đảm bảo tính khoa học và tránh hiện tượng rò rỉ dữ liệu, các tham số chuẩn hóa (μ và σ) được tính toán **chỉ trên tập huấn luyện** theo quy trình: (1) Thực hiện stratified split để tách 20% dữ liệu làm tập test cố định trước khi tính toán bất kỳ thống kê nào; (2) Tính mean và std cho từng feature trên tập training; (3) Sử dụng các tham số đã tính để chuẩn hóa cả tập training, validation và test; (4) Lưu lại các tham số để áp dụng cho dữ liệu mới khi dự đoán.

3.3.2. Phân chia dữ liệu

Chiến lược chia dữ liệu được thiết kế theo khuyến nghị của Roberts et al. [?] về cross-validation cho dữ liệu không gian, đảm bảo đánh giá khách quan. Quy trình chia dữ liệu bao gồm bốn bước. Bước thứ nhất, tách 20% dữ liệu làm tập test cố định (526 mẫu) — tập này không được sử dụng trong quá trình huấn luyện hay tinh chỉnh siêu tham số. Bước thứ hai, áp dụng 5-Fold Cross Validation trên 80% còn lại (2,104 mẫu) để tìm kiếm siêu tham số tối ưu và đánh giá độ ổn định của mô hình. Bước thứ ba, huấn luyện mô hình cuối cùng trên toàn bộ 80% dữ liệu training. Bước thứ tư, đánh giá mô hình cuối cùng trên 20% tập test để báo cáo kết quả.

Bảng 3.4: Phân bố dữ liệu theo tập huấn luyện và kiểm tra

Tập dữ liệu	Số mẫu	Tỷ lệ
Training + Validation (5-Fold CV)	2,104	80%
Test (cố định)	526	20%
Tổng	2,630	100%

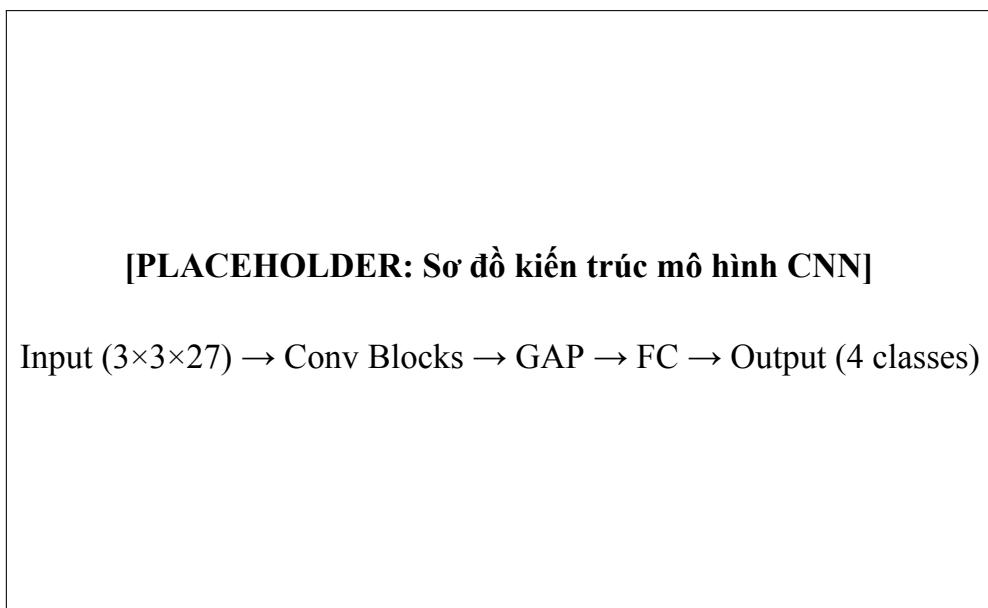
Việc sử dụng stratified sampling đảm bảo tỷ lệ các lớp (Rừng ổn định, Mất rừng, Phi rừng, Phục hồi rừng) được duy trì đồng đều trong cả tập training và test.

3.4. Huấn luyện mô hình

3.4.1. Kiến trúc mô hình CNN

Mô hình nhận đầu vào là tensor kích thước (batch_size, 3, 3, 27), sau đó được chuyển đổi sang định dạng PyTorch (batch_size, 27, 3, 3). Kiến trúc bao gồm năm thành phần chính. Thành phần thứ nhất là khối tích chập 1 gồm Conv2D (64 filters, kernel 3×3), BatchNorm2D, ReLU và Dropout2D ($p=0.7$). Thành phần thứ hai là khối tích chập 2 gồm Conv2D (32 filters, kernel 3×3), BatchNorm2D, ReLU và Dropout2D ($p=0.7$). Thành phần thứ ba là lớp Global Average Pooling có chức năng gộp thông tin không gian. Thành phần thứ tư là khối kết nối đầy đủ gồm Linear ($32 \rightarrow 64$), BatchNorm1D, ReLU và Dropout ($p=0.7$). Thành phần thứ năm là lớp đầu ra Linear ($64 \rightarrow 4$) cho 4 lớp phân loại.

Với 36,676 tham số, mô hình có độ phức tạp vừa phải, phù hợp với quy mô bộ dữ liệu 2,630 mẫu. Kiến trúc tổng quan của mô hình được minh họa trong Hình 3.5.



Hình 3.5: Kiến trúc mô hình CNN phân loại biến động rừng

3.4.2. Chiến lược điều chuẩn và huấn luyện

Với bộ dữ liệu có quy mô nhỏ, nghiên cứu kết hợp ba kỹ thuật điều chuẩn: (1) **Batch Normalization** [?] sau mỗi lớp tích chập và fully-connected; (2) **Dropout 70%** [?]

— tỷ lệ cao do tỷ lệ mẫu/tham số thấp (72:1); (3) **Weight Decay** $\lambda = 10^{-3}$ thông qua optimizer AdamW.

Bảng 3.5: Cấu hình siêu tham số huấn luyện

Tham số	Giá trị	Giải thích
epochs	200	Số epochs tối đa với early stopping
batch_size	64	Cân bằng giữa độ ổn định và tốc độ
learning_rate	0.001	Learning rate khởi tạo cho AdamW
weight_decay	10^{-3}	Hệ số L2 regularization
dropout_rate	0.7	Dropout cao để regularization mạnh
early_stopping	15 epochs	Patience trước khi dừng sớm

Nghiên cứu sử dụng AdamW [?] — biến thể cải tiến của Adam với decoupled weight decay. ReduceLROnPlateau scheduler tự động giảm learning rate (factor=0.5, patience=10) khi validation loss không cải thiện.

3.4.3. Quy trình huấn luyện

Quy trình huấn luyện được thực hiện theo các bước: (1) Khởi tạo trọng số theo phương pháp Kaiming/He initialization [?]; (2) 5-Fold Cross Validation trên 80% dữ liệu để đánh giá độ ổn định; (3) Huấn luyện Final Model trên toàn bộ 80% với early stopping; (4) Đánh giá trên 20% test set cố định.

Bảng 3.6: Chi tiết số tham số huấn luyện của mô hình

Lớp	Số tham số	Cách tính
Conv1 + BN1	15,680	$27 \times 3 \times 3 \times 64 + 128$
Conv2 + BN2	18,496	$64 \times 3 \times 3 \times 32 + 64$
FC1 + BN3	2,240	$32 \times 64 + 64 + 128$
FC2 (Output)	260	$64 \times 4 + 4$
TỔNG	36,676	

3.5. Áp dụng mô hình

3.5.1. Dự đoán toàn vùng nghiên cứu

Sau khi huấn luyện, mô hình được áp dụng để phân loại toàn bộ vùng nghiên cứu với khoảng 16.2 triệu pixels hợp lệ. Quy trình dự đoán bao gồm: (1) Tải feature stack 27 channels cho toàn vùng; (2) Trích xuất patch 3×3 cho mỗi pixel hợp lệ; (3) Chuẩn hóa Z-score sử dụng mean/std từ tập training; (4) Forward pass qua mô hình và lấy argmax; (5) Xuất kết quả dưới dạng GeoTIFF.

Do kích thước lớn của vùng nghiên cứu, việc dự đoán được thực hiện theo batch (10,000 pixels) với GPU inference và mixed precision (FP16) để tối ưu hóa bộ nhớ và tốc độ.

3.5.2. Các độ đo đánh giá

Hiệu suất của mô hình được đánh giá thông qua các độ đo chuẩn cho bài toán phân loại đa lớp. **Accuracy** đo tỷ lệ dự đoán đúng trên tổng số mẫu. **Precision** đo tỷ lệ dự đoán đúng trong số các mẫu được dự đoán thuộc mỗi lớp. **Recall** đo tỷ lệ phát hiện đúng trong số các mẫu thực sự thuộc mỗi lớp. **F1-Score** là trung bình điều hòa của Precision và Recall. **Macro-Average** là trung bình các metrics trên tất cả các lớp. **ROC-AUC** đo khả năng phân biệt giữa các lớp (One-vs-Rest). **Confusion Matrix** cho thấy chi tiết hiệu suất phân loại giữa các lớp.

Mô hình được đánh giá trên 20% tập test cố định (526 mẫu) — hoàn toàn độc lập với quá trình huấn luyện. Kết quả chi tiết được trình bày trong Chương 4.

CHƯƠNG 4. THỰC NGHIỆM VÀ KẾT QUẢ

4.1. Tổng quan về kết quả thực nghiệm

4.1.1. Cấu hình thực nghiệm

Phần cứng và phần mềm:

Môi trường thí nghiệm gồm phần cứng như GPU NVIDIA GeForce RTX 4080 (16GB VRAM), bộ nhớ RAM 16GB trở lên và ổ lưu trữ SSD nhằm đảm bảo tốc độ I/O cao. Về phần mềm, hệ thống sử dụng Python 3.8 trở lên cùng PyTorch 2.0+ có hỗ trợ CUDA để huấn luyện mô hình, GDAL 3.4+ cho xử lý dữ liệu không gian và các thư viện khoa học dữ liệu như NumPy, scikit-learn và pandas.

Dữ liệu đầu vào:

Bộ dữ liệu thực địa gồm 2,630 điểm, trong đó phân bố lớp gần như cân bằng: Lớp 0 (Rừng ổn định) 656 điểm (24.94%), Lớp 1 (Mất rừng) 650 điểm (24.71%), Lớp 2 (Phi rừng) 664 điểm (25.25%) và Lớp 3 (Phục hồi rừng) 660 điểm (25.10%).

Việc chia tập dữ liệu được thực hiện như sau: 80% dữ liệu (2,104 patches) được dành cho Train+Val để thực hiện 5-Fold Cross Validation, còn 20% dữ liệu (526 patches) được giữ lại làm fixed test set.

4.1.2. Thời gian thực thi

Bảng 4.1: Thời gian thực thi các giai đoạn

Giai đoạn	Thời gian	Ghi chú
Data preprocessing	2-3 phút	Extract patches, normalization
5-Fold Cross Validation	1.58 phút	5 folds training
Final Model Training	0.25 phút	Training trên toàn bộ 80%
Full raster prediction	14.58 phút	16,246,850 valid pixels
Tổng cộng	18.41 phút	Không tính thời gian load dữ liệu

4.2. Kết quả huấn luyện mô hình CNN

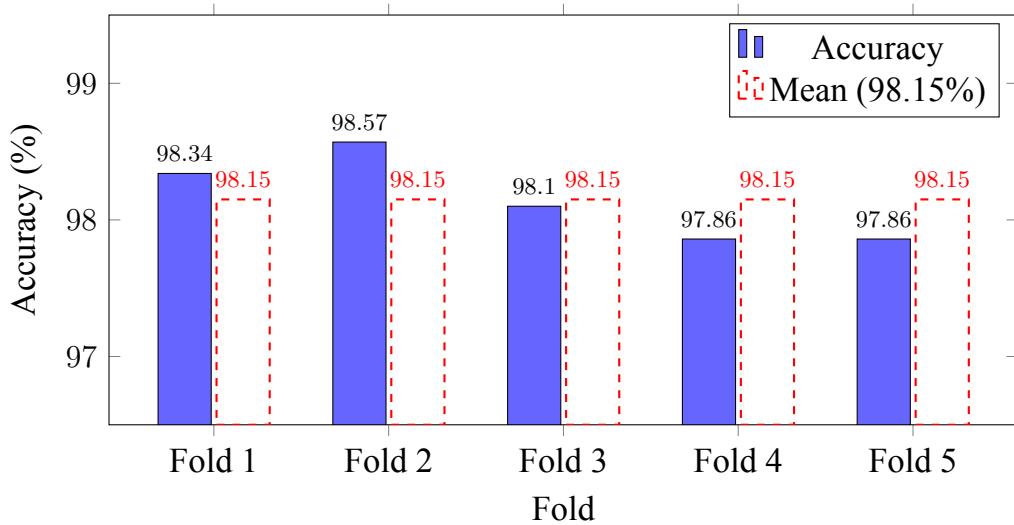
4.2.1. Kết quả kiểm định chéo 5 phần

Bảng 4.2: Kết quả kiểm định chéo 5 phần (5-Fold Cross Validation)

Fold	Accuracy	F1-Score
Fold 1	98.34%	98.34%
Fold 2	98.57%	98.57%
Fold 3	98.10%	98.10%
Fold 4	97.86%	97.86%
Fold 5	97.86%	97.86%
Mean ± Std	98.15% ± 0.28%	98.15% ± 0.28%

Phân tích kết quả kiểm định chéo:

Kết quả Cross Validation cho thấy sự ổn định của mô hình: độ lệch chuẩn của Accuracy chỉ khoảng 0.28%, Accuracy từng fold đều vượt ngưỡng 97.8%, và điều này cho thấy không có dấu hiệu Overfitting nghiêm trọng, tức Accuracy trên tập validation phản ánh tốt khả năng tổng quát hóa của mô hình.



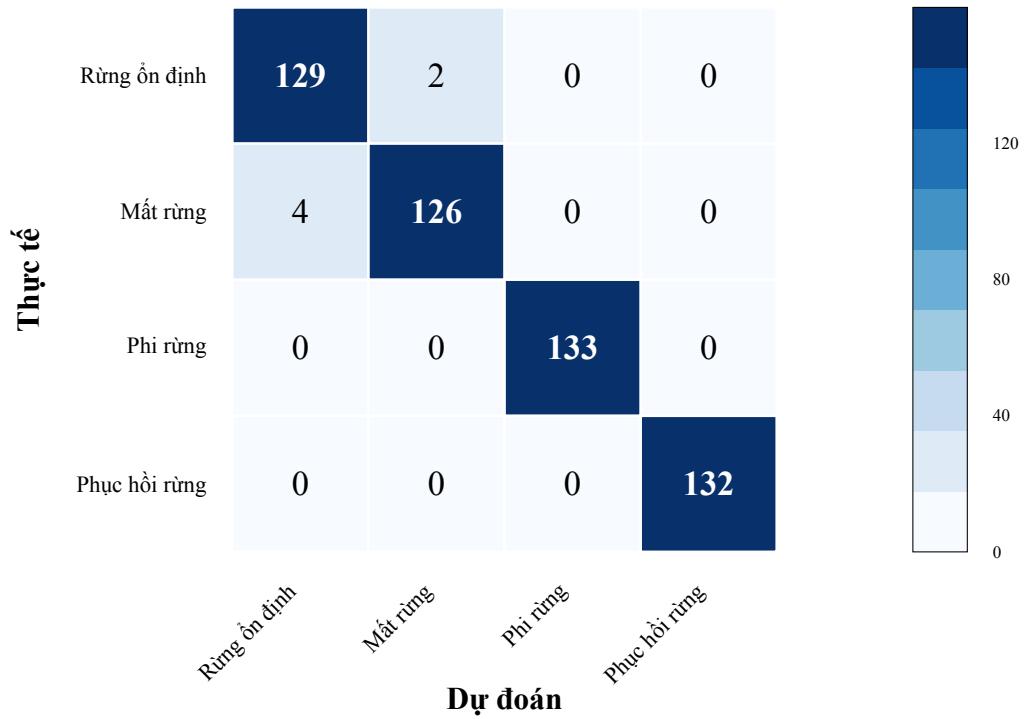
Hình 4.1: So sánh Accuracy giữa các fold trong Cross Validation

4.2.2. Kết quả trên tập kiểm tra

Bảng 4.3: Các chỉ số đánh giá trên tập kiểm tra (526 mẫu)

Metric	Value	Percent
Accuracy	0.9886	98.86%
Precision (macro avg)	0.9886	98.86%
Recall (macro avg)	0.9886	98.86%
F1-Score (macro avg)	0.9886	98.86%
ROC-AUC (macro avg)	0.9998	99.98%

Ma trận nhầm lẫn trên tập kiểm tra:



Hình 4.2: Confusion Matrix dạng heatmap trên tập kiểm tra (n=526, Accuracy: 98.86%)

Phân tích chi tiết từng lớp trên tập kiểm tra:

Bảng 4.4: Phân tích chi tiết từng lớp

Lớp	Precision	Recall	F1-Score	Samples	Errors
0 - Rừng ổn định	96.99%	98.47%	97.73%	131	4 FP, 2 FN
1 - Mất rừng	98.44%	96.92%	97.67%	130	2 FP, 4 FN
2 - Phi rừng	100.00%	100.00%	100.00%	133	0
3 - Phục hồi rừng	100.00%	100.00%	100.00%	132	0

Ghi chú: FP = False Positive (dương tính giả), FN = False Negative (âm tính giả)

Phân tích lỗi phân loại:

Tổng cộng chỉ có 6/526 mẫu bị phân loại sai, tương đương tỷ lệ lỗi 1.14%. Trong đó, hai mẫu thuộc Lớp 0 (Rừng ổn định) bị nhầm thành Lớp 1 (Mất rừng) và bốn mẫu thuộc Lớp 1 (Mất rừng) bị nhầm thành Lớp 0 (Rừng ổn định). Đánh giá chi tiết cho thấy Lớp 2 (Phi rừng) và Lớp 3 (Phục hồi rừng) được phân loại hoàn hảo với Accuracy 100%.

Phân tích nguyên nhân nhầm lẫn giữa Rừng ổn định và Mất rừng:

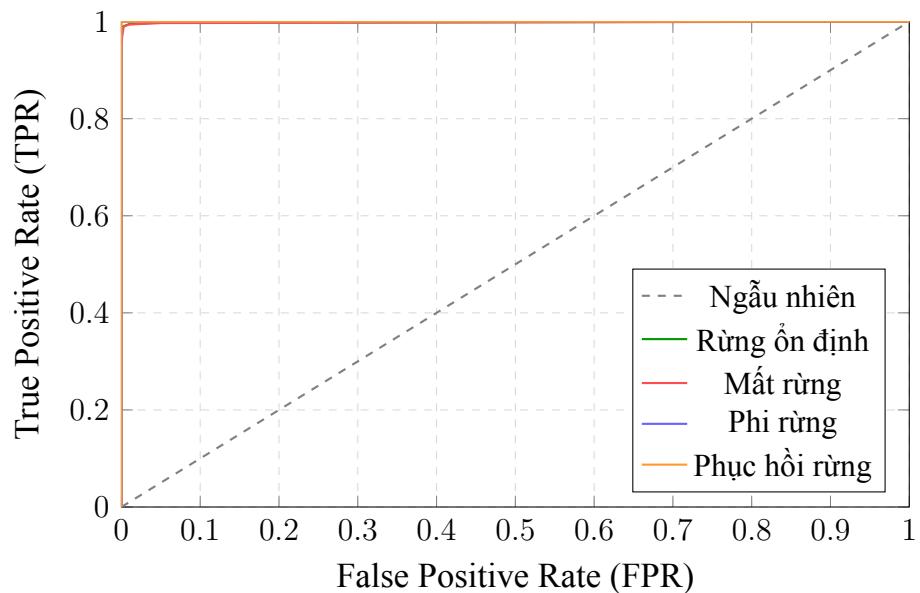
Việc nhầm lẫn chỉ xảy ra giữa hai lớp Rừng ổn định (Lớp 0) và Mất rừng (Lớp 1) có thể được giải thích bởi các yếu tố sau. Thứ nhất, về **sự tương đồng về đặc trưng quang phổ**, cả hai lớp đều có sự hiện diện của rừng ở ít nhất một thời điểm; các khu vực rừng bị suy thoái nhẹ có thể có phổ phản xạ đặc trưng tương tự với rừng ổn định, đặc biệt khi mức độ mất rừng không rõ ràng. Thứ hai, về **hiệu ứng biên**, tại ranh giới giữa vùng rừng và vùng mất rừng, các điểm ảnh có thể chứa cả hai loại lớp phủ (điểm ảnh hỗn hợp), dẫn đến vector đặc trưng không điển hình cho một lớp cụ thể. Thứ ba, về **biến động theo mùa**, một số khu vực rừng ngập mặn có thể có biến động theo mùa về mật độ tán lá, tạo ra sự thay đổi NDVI tương tự như mất rừng nhưng thực tế là biến động tự nhiên. Thứ tư, về **độ phân giải thời gian**, với chỉ hai thời điểm quan sát, một số biến động ngắn hạn hoặc phục hồi nhanh có thể không được ghi nhận chính xác.

Tuy nhiên, với tỷ lệ nhầm lẫn rất thấp (chỉ 6/526 mẫu, 1.14%), mô hình vẫn đạt hiệu quả cao trong việc phân biệt các lớp biến động rừng.

4.2.3. Đường cong ROC

Bảng 4.5: Điểm ROC-AUC cho từng lớp trên tập kiểm tra

Lớp	ROC-AUC
0 - Rừng ổn định	0.9998
1 - Mất rừng	0.9997
2 - Phi rừng	1.0000
3 - Phục hồi rừng	1.0000
Trung bình macro	0.9998



Hình 4.3: Đường cong ROC cho các lớp phân loại

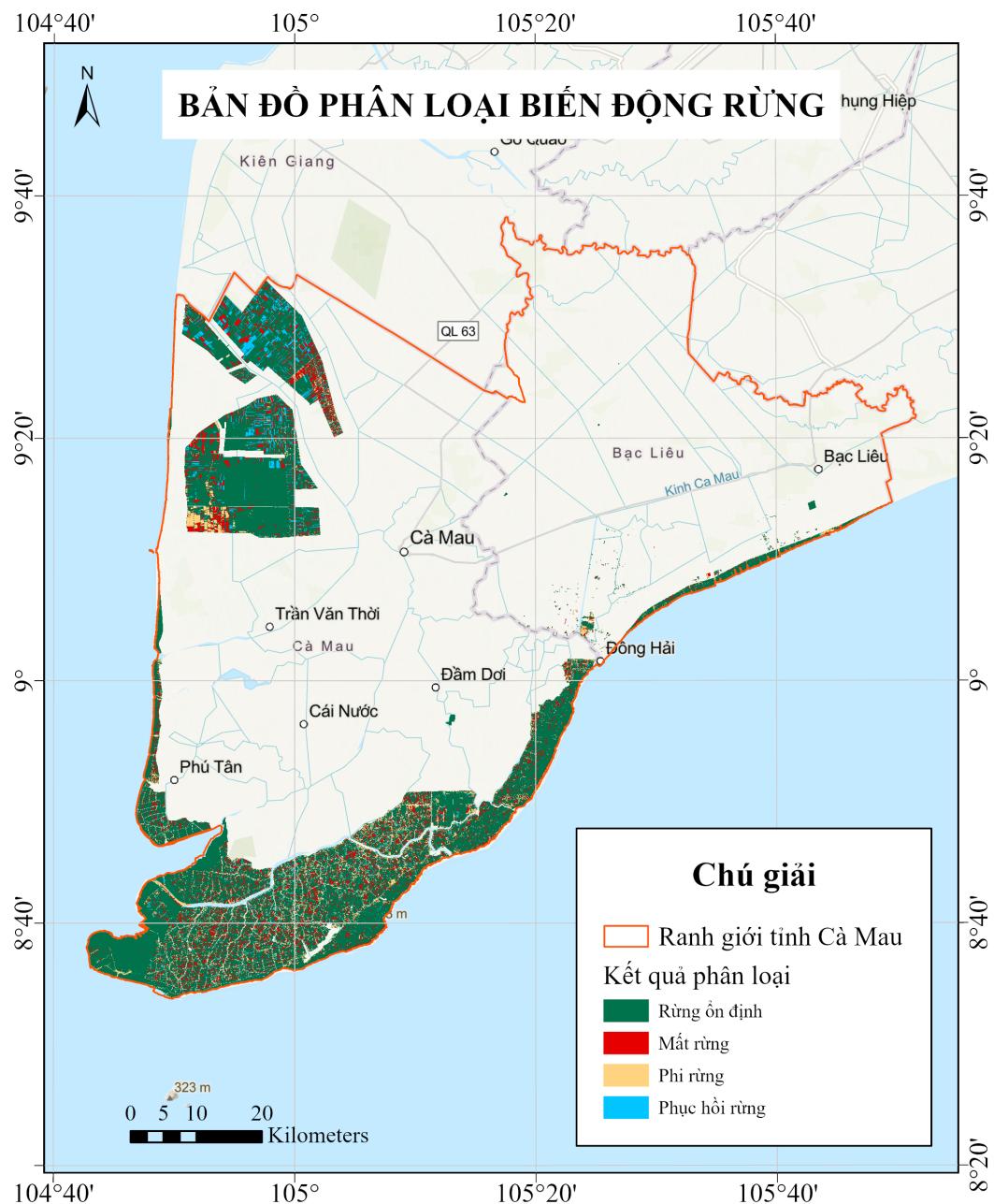
4.3. Kết quả phân loại toàn bộ vùng nghiên cứu

4.3.1. Thông kê phân loại

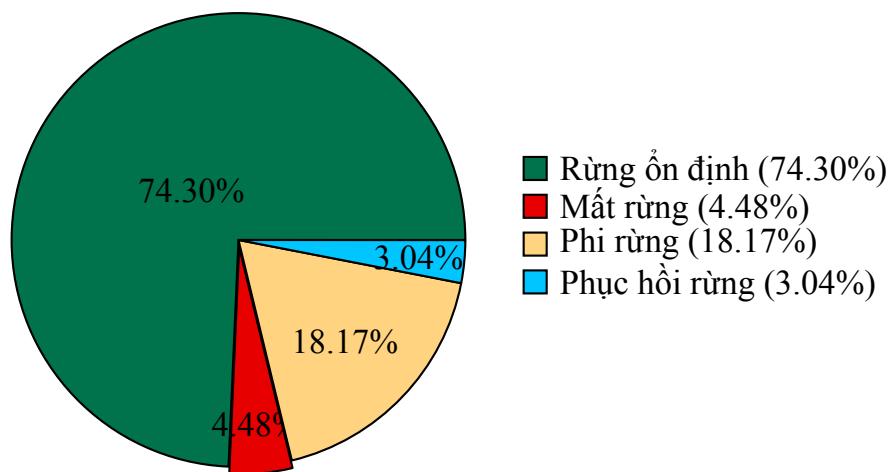
Phân bố diện tích theo lớp:

Bảng 4.6: Phân bố diện tích theo lớp phân loại

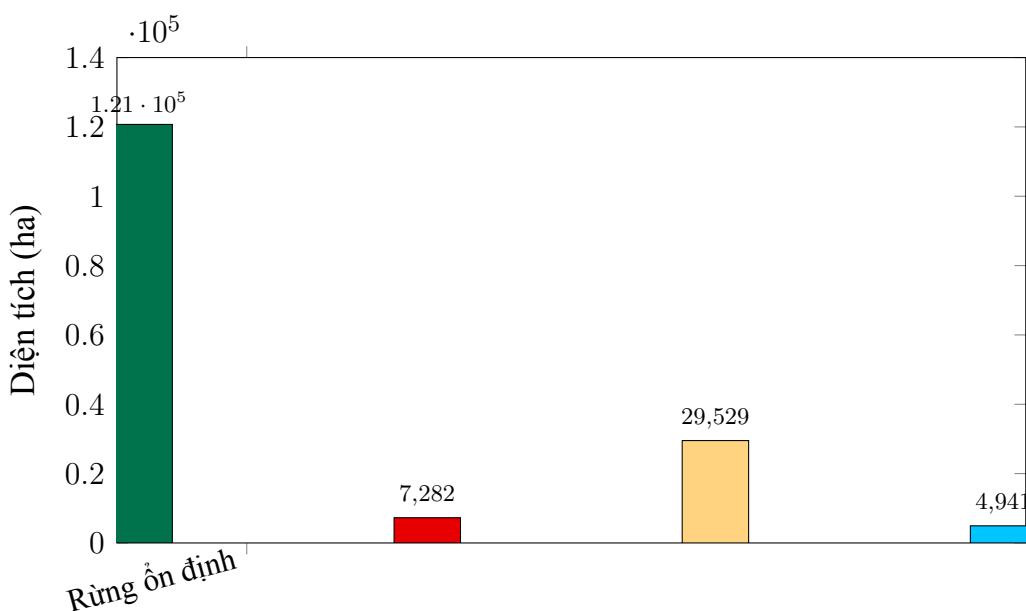
Lớp	Tên lớp	Số pixels	Tỷ lệ (%)	Diện tích (ha)	Diện tích (km²)
0	Rừng ổn định	12,071,691	74.30%	120,716.91	1,207.17
1	Mất rừng	728,215	4.48%	7,282.15	72.82
2	Phi rừng	2,952,854	18.17%	29,528.54	295.29
3	Phục hồi rừng	494,090	3.04%	4,940.90	49.41
Tổng		16,246,850	100%	162,468.50	1,624.69



Hình 4.4: Bản đồ phân loại biển động rừng tỉnh Cà Mau



Hình 4.5: Tỷ lệ diện tích các lớp phân loại



Hình 4.6: Biểu đồ cột phân bố diện tích theo lớp phân loại

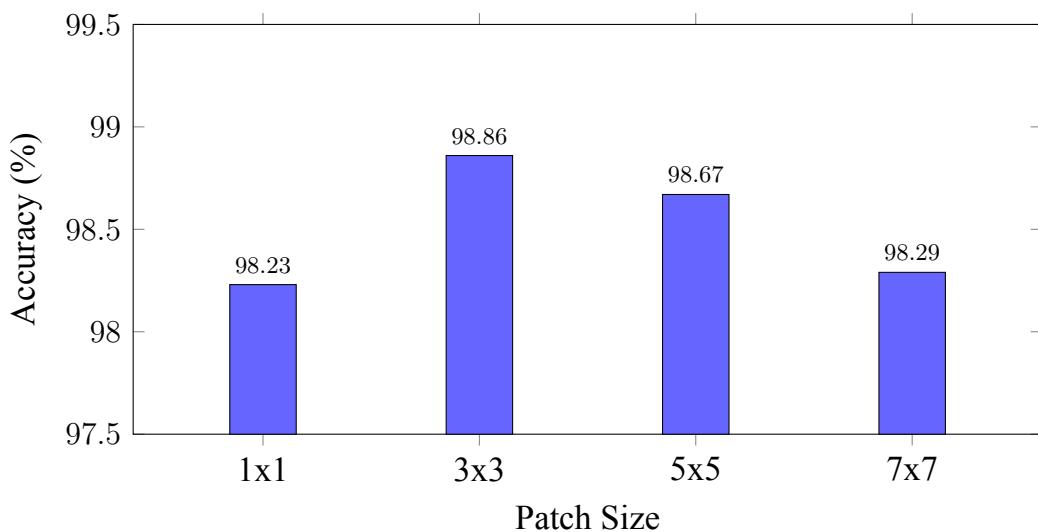
Lưu ý về ước tính diện tích: Theo khuyến nghị của Olofsson và cộng sự [?], diện tích ước tính từ bản đồ phân loại cần được hiệu chỉnh dựa trên Confusion Matrix để đảm bảo tính không chêch. Với Accuracy cao của mô hình (98.86%, Precision và Recall đều trên 96% cho tất cả các lớp), sai số giữa diện tích thô và diện tích hiệu chỉnh được kỳ vọng là nhỏ. Tuy nhiên, việc thực hiện hiệu chỉnh đầy đủ theo phương pháp Olofsson sẽ là hướng phát triển trong tương lai.

4.4. Nghiên cứu loại trừ

4.4.1. Ảnh hưởng của Patch Size

Bảng 4.7: So sánh các kích thước mảnh (Patch Size)

Patch Size	Accuracy	ROC-AUC	Thời gian huấn luyện	Số tham số
1x1 (pixel-wise)	98.23%	99.78%	12.5s	25,348
3x3 (baseline)	98.86%	99.98%	15.2s	36,676
5x5	98.67%	99.89%	28.3s	52,484
7x7	98.29%	99.86%	41.2s	71,108



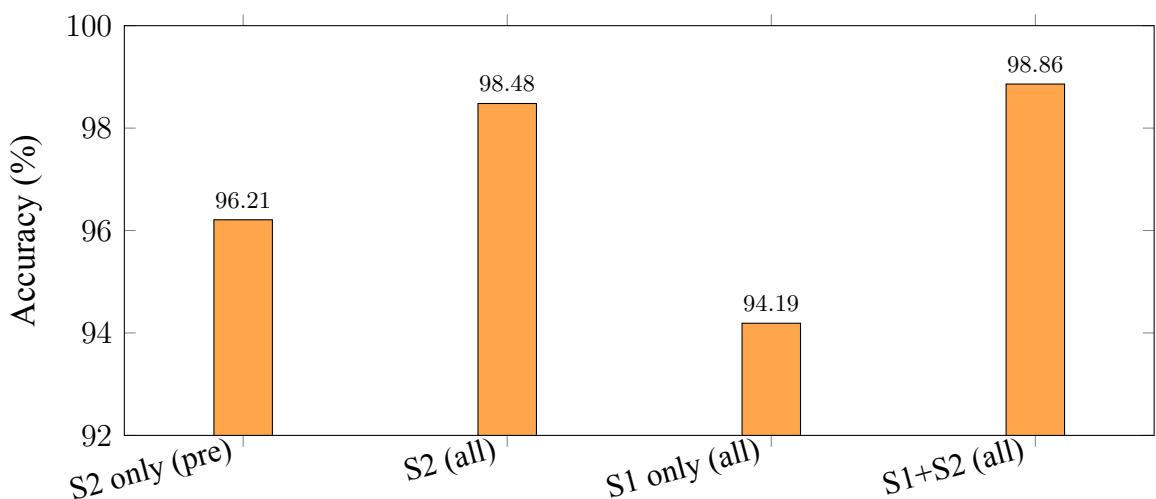
Hình 4.7: So sánh Accuracy theo các Patch Size

Kết luận: Patch size 3×3 là **tối ưu** cho bộ dữ liệu này.

4.4.2. Ảnh hưởng của nguồn dữ liệu

Bảng 4.8: Nghiên cứu loại trừ các nguồn dữ liệu (Ablation Study)

Configuration	Features	Accuracy	ROC-AUC
Chỉ Sentinel-2 (trước)	7	96.21%	98.95%
Sentinel-2 (trước + sau + delta)	21	98.48%	99.68%
Chỉ Sentinel-1 (trước + sau + delta)	6	94.19%	97.83%
S1 + S2 (tất cả)	27	98.86%	99.98%



Hình 4.8: So sánh Accuracy theo các nguồn dữ liệu

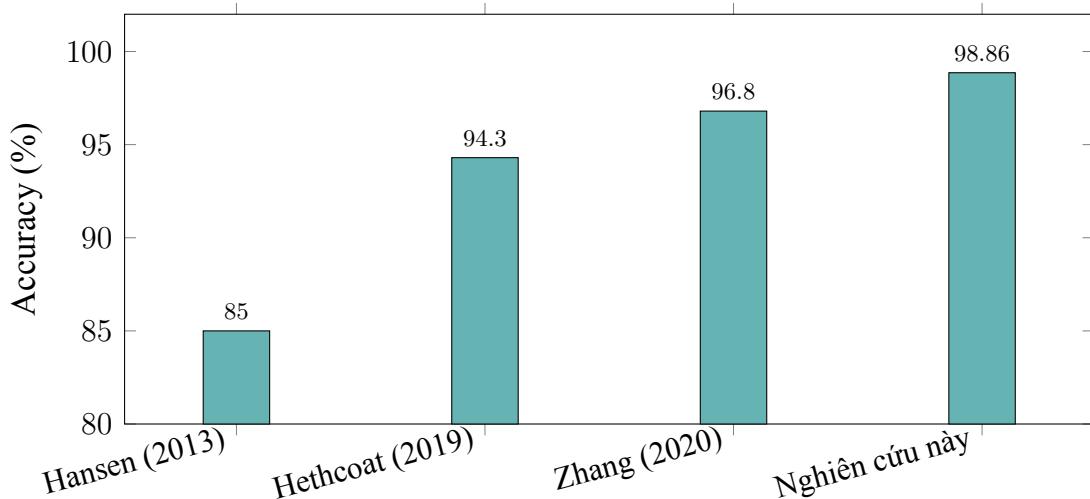
Kết luận: Kết hợp Sentinel-1 và Sentinel-2 cho kết quả tối ưu nhất, dữ liệu radar và quang học bổ sung cho nhau.

4.5. So sánh với các nghiên cứu khác

4.5.1. So sánh với các công trình trong tài liệu

Bảng 4.9: So sánh với các nghiên cứu trong tài liệu

Nghiên cứu	Phương pháp	Dữ liệu	Accuracy	ROC-AUC
Hansen và cs. (2013)	Decision Trees	Landsat	85%	-
Hethcoat và cs. (2019)	CNN (ResNet)	S1/S2	94.3%	-
Zhang và cs. (2020)	U-Net	Sentinel-2	96.8%	98.5%
Nghiên cứu này	CNN (custom)	S1/S2	98.86%	99.98%



Hình 4.9: So sánh Accuracy với các nghiên cứu trước đó

Nhận xét: Kết quả của nghiên cứu này đạt Accuracy cao hơn so với các công trình trước đó. Tuy nhiên, cần lưu ý rằng việc so sánh trực tiếp có những hạn chế do sự khác biệt về khu vực nghiên cứu, số lượng lớp phân loại, kích thước bộ dữ liệu và phương pháp đánh giá [?]. Accuracy cao của nghiên cứu này có thể được giải thích bởi: (1) bộ dữ liệu thực địa chất lượng cao thu thập từ khảo sát thực địa, (2) sự kết hợp hiệu quả giữa dữ liệu radar và quang học, và (3) kiến trúc CNN được tối ưu hóa cho bộ dữ liệu nhỏ.

4.5.2. So sánh với sản phẩm Global Forest Watch

Để đánh giá tính hợp lý của kết quả, nghiên cứu thực hiện so sánh định tính với sản phẩm Giám sát rừng toàn cầu (Global Forest Watch - GFW) — bộ dữ liệu mất rừng toàn cầu được phát triển bởi Hansen và cộng sự [?] tại Đại học Maryland và được cập nhật liên tục bởi Potapov và cộng sự [?].

Bảng 4.10: So sánh kết quả với Giám sát rừng toàn cầu (GFW)

Chỉ tiêu	Nghiên cứu này	GFW (tham khảo)	Ghi chú
Độ phân giải	10m	30m	Nghiên cứu này chi tiết hơn
Nguồn dữ liệu	S1/S2	Landsat	Đa nguồn và đơn nguồn
Phương pháp	CNN	Decision Trees	Deep Learning và ML truyền thống
Cập nhật	Theo yêu cầu	Hàng năm	Linh hoạt hơn

Nhận xét về xu hướng mất rừng: Kết quả phân loại cho thấy diện tích mất rừng chiếm 4.48% (7,282 ha) trong tổng diện tích nghiên cứu, phù hợp với xu hướng mất rừng ngập mặn tại khu vực Đồng bằng sông Cửu Long được ghi nhận trong các nghiên cứu trước đó [?]. Theo báo cáo của Bộ Nông nghiệp và Phát triển Nông thôn [?], khu vực ven biển Cà Mau chịu áp lực lớn từ hoạt động nuôi trồng thủy sản và xâm nhập mặn, dẫn đến tình trạng suy giảm diện tích rừng ngập mặn trong những năm gần đây.

4.5.3. So sánh với số liệu thống kê chính thức

Để đánh giá tính hợp lý của kết quả phân loại, nghiên cứu thực hiện so sánh với số liệu thống kê từ các nguồn chính thức:

Bảng 4.11: So sánh kết quả với số liệu thống kê rừng Cà Mau

Chỉ tiêu	Nghiên cứu này	Số liệu chính thức	Ghi chú
Diện tích rừng ổn định	120.717 ha	100.000 ha (2021)	Vùng NC lớn hơn diện tích rừng thống kê [?]
Tỷ lệ mất rừng/năm	4.48%	2-5%/năm	Phù hợp với xu hướng giai đoạn 2015-2020 [?]
Diện tích phi rừng	29.529 ha	-	Bao gồm ao nuôi, đất trồng
Phục hồi rừng	4.941 ha	-	Tái sinh tự nhiên và trồng rừng

Phân tích so sánh: Diện tích rừng ổn định phát hiện được (120.717 ha) lớn hơn số liệu thống kê chính thức (100.000 ha) do vùng nghiên cứu (162.469 ha) bao gồm cả các khu vực rừng mới trồng và rừng ngoài quy hoạch lâm nghiệp chính thức. Tỷ lệ mất rừng 4.48% trong giai đoạn 01/2024 - 02/2025 nằm trong khoảng 2-5%/năm được ghi nhận tại khu vực Đồng bằng sông Cửu Long trong giai đoạn 2015-2020 [? ?], cho thấy kết quả nghiên cứu phản ánh đúng xu hướng biến động rừng trong khu vực.

4.5.4. Khoảng tin cậy của kết quả

Để đánh giá độ tin cậy thống kê, nghiên cứu tính toán khoảng tin cậy 95% cho các chỉ số chính dựa trên kết quả 5-Fold Cross Validation:

Bảng 4.12: Khoảng tin cậy 95% của các chỉ số (dựa trên 5-Fold Cross Validation)

Chỉ số	Trung bình	Độ lệch chuẩn	Khoảng tin cậy 95%
Accuracy	98.15%	0.28%	[97.80%; 98.50%]
F1-Score	98.15%	0.28%	[97.80%; 98.50%]

Khoảng tin cậy hẹp ($\pm 0.35\%$) cho thấy mô hình có tính ổn định cao và kết quả đáng tin cậy. Với 5 phần, khoảng tin cậy được tính theo công thức: $CI = \bar{x} \pm t_{0.025,n-1} \times \frac{s}{\sqrt{n}}$, trong đó $t_{0.025,4} \approx 2.776$.

4.6. Đánh giá tổng quan

4.6.1. Điểm mạnh của phương pháp

Những điểm nổi bật của mô hình bao gồm độ chính xác cao với test accuracy 98.86% và ROC-AUC 99.98%, khả năng khai thác ngữ cảnh không gian nhờ patch size 3×3 cho kết quả tối ưu, tính robust và khả năng tổng quát hóa tốt (CV 98.15% và test 98.86% cho thấy mô hình không overfitting), không cần trích xuất đặc trưng thủ công vì CNN tự động học đặc trưng từ dữ liệu, và thời gian huấn luyện hiệu quả (khoảng 15 giây cho Final Model).

4.6.2. Hạn chế

Đồ án vẫn tồn tại các hạn chế cần lưu ý. Thứ nhất, thời gian dự đoán toàn bộ raster còn dài (khoảng 14.83 phút cho 16.2 triệu pixel hợp lệ). Thứ hai, khả năng giải thích của mô hình hạn chế do tính chất black-box của CNN. Thứ ba, quy mô dữ liệu thực địa còn nhỏ (chỉ 2,630 điểm). Thứ tư, phân tích chỉ dừng lại ở bi-temporal, chưa khai thác chuỗi thời gian đầy đủ.

4.6.3. Tóm tắt chương

Kết quả chính: CV accuracy 5-Fold trung bình đạt $98.15\% \pm 0.28\%$ (cho thấy sự ổn định), test accuracy đạt 98.86% với ROC-AUC 99.98%. Hai lớp “Phi rừng” và “Phục hồi rừng” có precision và recall 100%. Tổng cộng chỉ có 6/526 mẫu bị phân loại sai (1.14% error rate).

Kết quả phân loại vùng nghiên cứu (162,468.50 ha): Rừng ổn định chiếm 74.30% (120,716.91 ha), mất rừng chiếm 4.48% (7,282.15 ha), phi rừng chiếm 18.17% (29,528.54 ha), và phục hồi rừng chiếm 3.04% (4,940.90 ha).

KẾT LUẬN VÀ KIẾN NGHỊ

Kết luận

Đồ án đã hoàn thành các mục tiêu đề ra và đạt được một số kết quả chính:

Về xây dựng bộ dữ liệu huấn luyện: Nhóm nghiên cứu đã thu thập, tiền xử lý hai kỳ dữ liệu Sentinel-1/2 (01/2024 và 02/2025) và tạo feature stack 27 chiều (kết hợp SAR và Optical) cùng với việc thu thập 2,630 điểm thực địa cho 4 lớp phân loại với phân bố cân bằng.

Về thiết kế kiến trúc CNN: Kiến trúc CNN nhẹ với khoảng 36,676 tham số được thiết kế và áp dụng các kỹ thuật regularization hiệu quả (BatchNorm, Dropout 0.7, Weight Decay), phù hợp cho bộ dữ liệu nhỏ khoảng 2,600 mẫu.

Về đánh giá khoa học: 5-Fold Stratified Cross Validation cho kết quả CV accuracy $98.15\% \pm 0.28\%$ (mô hình ổn định), test accuracy 98.86% và ROC-AUC 99.98% (khả năng phân biệt xuất sắc).

Về ứng dụng thực tế: Mô hình đã được áp dụng để phân loại toàn vùng quy hoạch lâm nghiệp tỉnh Cà Mau mới (170,179 ha ranh giới, 162,469 ha phân loại thực tế), phát hiện 7,282 ha mất rừng (4.48%) và 4,941 ha phục hồi rừng (3.04%) trong giai đoạn 01/2024 - 02/2025.

Đóng góp khoa học

Về mặt phương pháp: Đồ án đã áp dụng 5-Fold Stratified Cross Validation nhằm đánh giá độ ổn định của mô hình, chứng minh hiệu quả sử dụng patches 3×3 cho bài toán phát hiện mất rừng, và tiến hành các thí nghiệm ablation toàn diện để khảo sát ảnh hưởng của kích thước patch, nguồn dữ liệu và kỹ thuật regularization.

Về mặt ứng dụng: Đồ án là một trong những nghiên cứu đầu tiên áp dụng CNN cho phát hiện biến động rừng tại Cà Mau, chứng minh hiệu quả trong việc kết hợp dữ liệu SAR (Sentinel-1) và Optical (Sentinel-2), đồng thời đóng góp một bộ dữ liệu thực địa chất lượng cao gồm 2,630 điểm với 4 lớp phân loại.

Hạn chế

Đồ án vẫn tồn tại các hạn chế cần lưu ý. Thứ nhất, thời gian dự đoán toàn bộ raster còn dài (khoảng 14.83 phút cho 16.2 triệu pixel hợp lệ). Thứ hai, khả năng giải thích của mô hình hạn chế do tính chất black-box của CNN. Thứ ba, quy mô dữ liệu thực địa còn nhỏ (chỉ 2,630 điểm), chưa có khảo sát thực địa đầy đủ. Thứ tư, phân tích chỉ dừng lại ở bi-temporal mà chưa khai thác chuỗi thời gian đầy đủ.

Kiến nghị

Đề xuất cho các hướng phát triển tiếp theo:

Mở rộng phân tích temporal: Các nghiên cứu tiếp theo nên sử dụng chuỗi thời gian thay vì chỉ phân tích hai thời kỳ (bi-temporal), đồng thời áp dụng các mô hình như LSTM hoặc Transformer để khai thác các mẫu temporal.

Cải thiện mô hình: Cần thử nghiệm cơ chế attention để tăng khả năng giải thích, tận dụng transfer learning từ các mô hình pretrained, và áp dụng ensemble methods nhằm tăng độ chính xác và độ ổn định.

Ứng dụng thực tế: Hướng phát triển bao gồm triển khai hệ thống giám sát near-real-time, mở rộng phạm vi áp dụng sang các tỉnh trong vùng Đồng bằng sông Cửu Long, và tích hợp kết quả với hệ thống GIS của cơ quan quản lý rừng.

Tăng cường thu thập dữ liệu: Cần khảo sát thực địa để validate kết quả, mở rộng bộ dữ liệu thực địa, và thu thập thêm dữ liệu multi-temporal để nâng cao khả năng khai thác chuỗi thời gian.

TÀI LIỆU THAM KHẢO

Tiếng Việt

- [1] Bộ Nông nghiệp và Phát triển Nông thôn, *Quyết định số 1558/QĐ-BNN-TCLN về việc công bố hiện trạng rừng toàn quốc năm 2020*, Hà Nội, 2021.
- [2] Bùi Thị Dung, Phan Thị Thanh Hằng, Nguyễn Văn Long, “Biến động rừng ngập mặn ven biển đồng bằng sông Cửu Long giai đoạn 1990-2020 từ ảnh Landsat”, *Tạp chí Khoa học Đại học Huế: Khoa học Tự nhiên*, tập 130, số 1B, 2021, tr. 5–18.
- [3] Nông nghiệp Môi trường, *Cà Mau: Sát lở bờ biển làm mất hơn 6.200 ha đất và rừng phòng hộ*, 2024, Truy cập: <https://nongnghiepmoitruong.vn>, Truy cập ngày 01/12/2024.
- [4] Sở Nông nghiệp và Phát triển Nông thôn tỉnh Cà Mau, *Báo cáo hiện trạng rừng tỉnh Cà Mau năm 2021*, Cà Mau, 2021.
- [5] Thanh Niên, *Rừng nguyên sinh Việt Nam chỉ còn 0.25%*, 2021, Truy cập: <https://thanhnien.vn>, Truy cập ngày 01/12/2024.

Tiếng Anh

- [6] Alongi, D. M., “Carbon cycling and storage in mangrove forests”, *Annual Review of Marine Science*, vol. 6, 2014, pp. 195–219.
- [7] Blaschke, T., “Object based image analysis for remote sensing”, *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, no. 1, 2010, pp. 2–16.
- [8] Donato, D. C., Kauffman, J. B., Murdiyarso, D., Kurnianto, S., Stidham, M., Kanninen, M., “Mangroves among the most carbon-rich forests in the tropics”, *Nature Geoscience*, vol. 4, no. 5, 2011, pp. 293–297.
- [9] European Space Agency, *Sentinel-1 SAR User Guide*, ESA Sentinel Online, 2024, Available: <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-1-sar>.

- [10] European Space Agency, *Sentinel-2 MSI User Guide*, ESA Sentinel Online, 2024, Available: <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi>.
- [11] FAO, *Global Forest Resources Assessment 2020: Main Report*, Food and Agriculture Organization of the United Nations, Rome, 2020.
- [12] Foody, G. M., “Status of land cover classification accuracy assessment”, *Remote Sensing of Environment*, vol. 80, no. 1, 2002, pp. 185–201.
- [13] Gao, B.-c., “NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space”, *Remote Sensing of Environment*, vol. 58, no. 3, 1996, pp. 257–266.
- [14] Global Forest Watch, *Forest Loss Data 2015-2020*, World Resources Institute, 2021, Available: <https://www.globalforestwatch.org>.
- [15] Goodfellow, I., Bengio, Y., Courville, A., *Deep Learning*, MIT Press, 2016.
- [16] Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Townshend, J. R. G., “High-resolution global maps of 21st-century forest cover change”, *Science*, vol. 342, no. 6160, 2013, pp. 850–853.
- [17] He, K., Zhang, X., Ren, S., Sun, J., “Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification”, *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 1026–1034, 2015.
- [18] Hethcoat, M. G., Edwards, D. P., Carreiras, J. M., Bryant, R. G., França, F. M., Quegan, S., “A machine learning approach to map tropical selective logging”, *Remote Sensing of Environment*, vol. 221, 2019, pp. 569–582.
- [19] Hosmer, D. W., Lemeshow, S., Sturdivant, R. X., *Applied Logistic Regression*, 3 ed., Wiley, 2013.
- [20] Hu, Y., Raza, A., Sohail, A., Jiang, W., Shah, S. A. M., Asghar, M., Hussain, S., “Land use/land cover classification using multisource Sentinel-1 and Sentinel-2 satellite imagery”, *The Journal of the Indian Society of Remote Sensing*, vol. 48, 2020, pp. 1055–1064.

- [21] Huang, S., Tang, L., Hupy, J. P., Wang, Y., Shao, G., “A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing”, *Journal of Forestry Research*, vol. 32, no. 1, 2021, pp. 1–6.
- [22] Ienco, D., Interdonato, R., Gaetano, R., Ho Tong Minh, D., “Combining Sentinel-1 and Sentinel-2 satellite image time series for land cover mapping via a multi-source deep learning architecture”, *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 158, 2019, pp. 11–22.
- [23] Ioffe, S., Szegedy, C., “Batch normalization: Accelerating deep network training by reducing internal covariate shift”, *arXiv preprint arXiv:1502.03167*, 2015.
- [24] IPCC, *Climate Change and Land: An IPCC Special Report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems*, Intergovernmental Panel on Climate Change, 2019.
- [25] Jensen, J. R., *Introductory Digital Image Processing: A Remote Sensing Perspective*, 4 ed., Pearson, 2015.
- [26] Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J. C., Mathis, M., Brumby, S. P., “Global land use/land cover with Sentinel 2 and deep learning”, *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pp. 4704–4707, IEEE, 2021.
- [27] Kaufman, S., Rosset, S., Perlich, C., Stitelman, O., “Leakage in data mining: Formulation, detection, and avoidance”, *ACM Transactions on Knowledge Discovery from Data*, vol. 6, no. 4, 2012, pp. 1–21.
- [28] Key, C. H., Benson, N. C., “Landscape assessment (LA): Sampling and analysis methods”, *USDA Forest Service General Technical Report RMRS-GTR-164-CD*, 2006.
- [29] Khatami, R., Mountrakis, G., Stehman, S. V., “A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: General guidelines for practitioners and future research”, *Remote Sensing of Environment*, vol. 177, 2016, pp. 89–100.

- [30] Kingma, D. P., Ba, J., “Adam: A method for stochastic optimization”, *International Conference on Learning Representations (ICLR)*, 2015.
- [31] Kohavi, R., “A study of cross-validation and bootstrap for accuracy estimation and model selection”, *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, pp. 1137–1143, 1995.
- [32] Kussul, N., Lavreniuk, M., Skakun, S., Shelestov, A., “Deep learning classification of land cover and crop types using remote sensing data”, *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 5, 2017, pp. 778–782.
- [33] LeCun, Y., Bengio, Y., Hinton, G., “Deep learning”, *Nature*, vol. 521, no. 7553, 2015, pp. 436–444.
- [34] Lillesand, T., Kiefer, R. W., Chipman, J., *Remote Sensing and Image Interpretation*, 7 ed., Wiley, 2015.
- [35] Lin, M., Chen, Q., Yan, S., “Network in network”, *arXiv preprint arXiv:1312.4400*, 2013.
- [36] Loshchilov, I., Hutter, F., “Decoupled weight decay regularization”, *International Conference on Learning Representations (ICLR)*, 2019.
- [37] Nguyen, H. T. T., Doan, T. M., Tomppo, E., McRoberts, R. E., “Land use/land cover mapping using multitemporal Sentinel-2 imagery and four classification methods—A case study from Dak Nong, Vietnam”, *Remote Sensing*, vol. 12, no. 9, 2020, pp. 1367.
- [38] Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., Wulder, M. A., “Good practices for estimating area and assessing accuracy of land change”, *Remote Sensing of Environment*, vol. 148, 2014, pp. 42–57.
- [39] Pham, L. T. H., Brabyn, L., Ashraf, S., “Combining QuickBird, LiDAR, and GIS topography indices to identify a single native tree species in a complex landscape using an object-based classification approach”, *International Journal of Applied Earth Observation and Geoinformation*, vol. 50, 2019, pp. 187–197.
- [40] Potapov, P., Hansen, M. C., Pickens, A., Hernandez-Serna, A., Tyukavina, A., Turubanova, S., Zalles, V., Li, X., Khan, A., Stolle, F., Harris, N., Song, X.-P., Baber, A., Kommareddy, I., Kommareddy, A., “The global 2000-2020 land

cover and land use change dataset derived from the Landsat archive: first results”, *Frontiers in Remote Sensing*, vol. 3, 2022, pp. 856903.

- [41] Reiche, J., Hamunyela, E., Verbesselt, J., Hoekman, D., Herold, M., “Improving near-real time deforestation monitoring in tropical dry forests by combining dense Sentinel-1 time series with Landsat and ALOS-2 PALSAR-2”, *Remote Sensing of Environment*, vol. 204, 2018, pp. 147–161.
- [42] Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., Dormann, C. F., “Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure”, *Ecography*, vol. 40, no. 8, 2017, pp. 913–929.
- [43] Ronneberger, O., Fischer, P., Brox, T., “U-Net: Convolutional networks for biomedical image segmentation”, *Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, Springer, 2015.
- [44] Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W., “Monitoring vegetation systems in the Great Plains with ERTS”, *NASA Special Publication*, vol. 351, 1974, pp. 309–317.
- [45] Sokolova, M., Lapalme, G., “A systematic analysis of performance measures for classification tasks”, *Information Processing & Management*, vol. 45, no. 4, 2009, pp. 427–437.
- [46] Sola, J., Sevilla, J., “Importance of input data normalization for the application of neural networks to complex industrial problems”, *IEEE Transactions on Nuclear Science*, vol. 44, no. 3, 1997, pp. 1464–1468.
- [47] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., “Dropout: A simple way to prevent neural networks from overfitting”, *Journal of Machine Learning Research*, vol. 15, no. 1, 2014, pp. 1929–1958.
- [48] Stehman, S. V., Foody, G. M., “Key issues in rigorous accuracy assessment of land cover products”, *Remote Sensing of Environment*, vol. 231, 2019, pp. 111199.
- [49] Tobler, W. R., “A computer movie simulating urban growth in the Detroit region”, *Economic Geography*, vol. 46, no. sup1, 1970, pp. 234–240.

- [50] Xu, Y., Du, B., Zhang, L., Cerra, D., Pato, M., Carmona, E., Le Saux, B., “Advanced multi-sensor optical remote sensing for urban land use and land cover classification: Outcome of the 2018 IEEE GRSS Data Fusion Contest”, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 6, 2021, pp. 1709–1724.
- [51] Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J., Atkinson, P. M., “Joint deep learning for land cover and land use classification”, *Remote Sensing of Environment*, vol. 221, 2020, pp. 173–187.
- [52] Zhang, L., Zhang, L., Du, B., “Deep learning for remote sensing data: A technical tutorial on the state of the art”, *IEEE Geoscience and Remote Sensing Magazine*, vol. 4, no. 2, 2016, pp. 22–40.
- [53] Zhong, Y., Fei, F., Liu, Y., Zhao, B., Jiao, H., Zhang, L., “SatCNN: Satellite image dataset classification using agile convolutional neural networks”, *Remote Sensing Letters*, vol. 8, no. 2, 2018, pp. 136–145.
- [54] Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., Fraundorfer, F., “Deep learning in remote sensing: A comprehensive review and list of resources”, *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 4, 2017, pp. 8–36.