

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



**BÁO CÁO THỰC TẬP TỐT NGHIỆP
GIÁM SÁT BIẾN ĐỘNG RỪNG CÀ MAU
SỬ DỤNG DỮ LIỆU VIỄN THÁM VÀ
MẠNG NƠI RON NHÂN TẠO**

Học phần:

Thực tập tốt nghiệp

Mã học phần:

AER4002 21

Giảng viên phụ trách học phần:

TS. Nguyễn Văn Thương

Đơn vị thực tập:

CÔNG TY TNHH TƯ VẤN VÀ PHÁT TRIỂN ĐỒNG XANH

Cán bộ hướng dẫn tại công ty:

TS. Hoàng Việt Anh

Giảng viên hướng dẫn tại trường:

TS. Hà Minh Cường

Sinh viên thực hiện:

Ninh Hải Đăng

Mã sinh viên:

21021411

Lớp:

QH-2021-I/CQ-S-AE

Hà Nội, tháng 12 năm 2025

MỤC LỤC

MỤC LỤC	i
LỜI CẢM ƠN	iii
Chương 1 NHẬN XÉT VÀ XÁC NHẬN	1
1.1. Nhận xét của cán bộ hướng dẫn nơi thực tập	1
1.2. Nhận xét của giảng viên phụ trách sinh viên	2
Chương 2 KHÁI QUÁT BÁO CÁO THỰC TẬP	3
2.1. Giới thiệu chung	3
2.2. Khái quát nội dung thực tập	4
2.3. Kết quả đạt được	4
2.4. Kinh nghiệm rút ra	5
2.5. Kết luận	5
Chương 3 BÁO CÁO CHI TIẾT KẾT QUẢ ĐẠT ĐƯỢC	6
3.1. Nghiên cứu đề bài và cơ sở lý thuyết	6
3.1.1. Tổng quan về rừng và biến động rừng	6
3.1.2. Cơ sở lý thuyết về viễn thám vệ tinh	6
3.1.3. Học sâu và Mạng nơ-ron tích chập	8
3.1.4. Tổng quan tình hình nghiên cứu	9
3.2. Báo cáo chi tiết nội dung nghiên cứu và thực hiện	10
3.2.1. Dữ liệu nghiên cứu	10
3.2.2. Phương pháp nghiên cứu	13
3.2.3. Kết quả thực nghiệm và thảo luận	17
3.3. Các sản phẩm hoàn thành	25

3.4. Các kiến nghị đề xuất	26
TÀI LIỆU THAM KHẢO	27

LỜI CẢM ƠN

Trước tiên, em xin chân thành cảm ơn Công ty TNHH Tư vấn và Phát triển Đồng Xanh đã tạo điều kiện cho em thực hiện đề tài thực tập tốt nghiệp này. Môi trường làm việc chuyên nghiệp và sự hỗ trợ nhiệt tình từ công ty đã giúp em có cơ hội áp dụng kiến thức đã học vào thực tiễn.

Em xin gửi lời cảm ơn sâu sắc đến TS. Hoàng Việt Anh – cán bộ hướng dẫn tại công ty, người đã dành thời gian hướng dẫn, giải đáp các thắc mắc và định hướng cho em trong suốt quá trình thực tập.

Em xin chân thành cảm ơn TS. Hà Minh Cường và ThS. Hoàng Tích Phúc – giảng viên Viện Công nghệ Hàng không Vũ trụ, đã tận tình chỉ bảo, theo dõi và hỗ trợ em trong suốt quá trình thực hiện đề tài. Sự hướng dẫn tận tâm của các thầy đã giúp em nâng cao kiến thức chuyên môn và hoàn thiện kỹ năng nghiên cứu khoa học.

Cuối cùng, em xin cảm ơn gia đình, bạn bè và đồng nghiệp đã luôn động viên, chia sẻ và tạo điều kiện tốt nhất để em hoàn thành kỳ thực tập này.

Do thời gian và kinh nghiệm còn hạn chế, báo cáo không tránh khỏi những thiếu sót. Em rất mong nhận được sự góp ý của các thầy để báo cáo được hoàn thiện hơn.

Em xin chân thành cảm ơn!

Hà Nội, tháng 12 năm 2025
Sinh viên

Ninh Hải Đăng

PHẦN 1. NHẬN XÉT VÀ XÁC NHẬN

1.1. Nhận xét của cán bộ hướng dẫn noi thực tập

Điểm:/10

....., ngày tháng năm

Cán bộ hướng dẫn

(Ký và ghi rõ họ tên)

TS. Hoàng Việt Anh

1.2. Nhận xét của giảng viên phụ trách sinh viên

Điểm:/10

....., ngày tháng năm

Giảng viên phụ trách

(Ký và ghi rõ họ tên)

TS. Hà Minh Cường

PHẦN 2. KHÁI QUÁT BÁO CÁO THỰC TẬP

2.1. Giới thiệu chung

Nằm ở cực Nam của Tổ quốc, tỉnh Cà Mau sở hữu hệ sinh thái rừng ngập mặn đặc thù và vô cùng quan trọng đối với sự cân bằng môi trường cũng như đời sống kinh tế xã hội của vùng Đồng bằng sông Cửu Long. Tuy nhiên, trước áp lực của biến đổi khí hậu, nước biển dâng và các hoạt động kinh tế như nuôi trồng thủy sản, diện tích rừng tại đây đang đứng trước nguy cơ suy giảm nghiêm trọng. Việc giám sát biến động tài nguyên rừng theo phương pháp truyền thống thường tốn kém nhiều thời gian, công sức và khó triển khai trên diện rộng. Trong bối cảnh đó, sự phát triển mạnh mẽ của công nghệ viễn thám và trí tuệ nhân tạo, đặc biệt là các mô hình học sâu, đã mở ra hướng đi mới đầy tiềm năng cho công tác quản lý và bảo vệ rừng. Xuất phát từ thực tiễn cấp thiết này, báo cáo thực tập tập trung vào đề tài “Giám sát biến động rừng tỉnh Cà Mau sử dụng dữ liệu viễn thám và mạng nơ-ron nhân tạo”, với mục tiêu ứng dụng công nghệ hiện đại để giải quyết bài toán môi trường tại địa phương.

Quá trình thực tập được thực hiện tại Công ty TNHH Tư vấn và Phát triển Đồng Xanh dưới sự hướng dẫn của TS. Hoàng Việt Anh. Đây là đơn vị có nhiều kinh nghiệm trong lĩnh vực tư vấn giải pháp công nghệ quản lý tài nguyên và là đối tác chiến lược của Chi cục Kiểm lâm tỉnh Cà Mau. Trong thời gian thực tập từ ngày 15/9/2025 đến 15/12/2025, em đã trực tiếp tham gia vào dự án xây dựng hệ thống giám sát rừng tự động, từ khâu thu thập dữ liệu vệ tinh đa nguồn, xử lý ảnh, đến thiết kế và huấn luyện các mô hình học sâu tiên tiến.

Công ty TNHH Tư vấn và Phát triển Đồng Xanh (GFD) là doanh nghiệp tiên phong trong lĩnh vực công nghệ và kỹ thuật, chuyên sâu về các giải pháp GIS (Hệ thống Thông tin Địa lý), Viễn thám và Tư vấn quản lý tài nguyên thiên nhiên. Với hơn 15 năm hình thành và phát triển, GFD đã khẳng định vị thế vững chắc thông qua việc triển khai thành công hơn 100 dự án GIS và 300 dự án tư vấn cho các khách hàng lớn trong nước (Bộ Nông nghiệp và Phát triển Nông thôn, Bộ Tài nguyên và Môi trường) và các tổ chức quốc tế (World Bank, ADB, JICA, USAID). Quy trình sản xuất phần mềm và quản lý an toàn thông tin của công ty tuân thủ nghiêm ngặt các tiêu chuẩn quốc tế ISO 9001:2015 và ISO 27001:2022.

Lĩnh vực hoạt động cốt lõi của GFD bao gồm phát triển các giải pháp phần mềm

GIS và Viễn thám toàn diện, cung cấp chuỗi dịch vụ khép kín từ khảo sát thực địa, số hóa dữ liệu đến thiết kế phần mềm trên nền tảng dựa trên các công nghệ hiện đại như Open Source GIS, điện toán đám mây và các giải pháp thương mại từ ESRI, Bentley; đồng thời, công ty còn cung cấp các dịch vụ tư vấn chuyên sâu về quản lý tài nguyên thiên nhiên như quản lý rừng bền vững, bảo tồn đa dạng sinh học, đo đạc carbon lâm nghiệp và đánh giá thiệt hại thiên tai, luôn đảm bảo tuân thủ hệ thống pháp luật hiện hành và đáp ứng các yêu cầu khắt khe của các tổ chức quốc tế.

2.2. Khái quát nội dung thực tập

Nội dung thực tập trung vào việc nghiên cứu và giải quyết bài toán phát hiện mất rừng và phục hồi rừng thông qua phân tích ảnh vệ tinh. Cụ thể, em đã tiến hành nghiên cứu cơ sở lý thuyết về viễn thám, đặc điểm phổ của các đối tượng rừng ngập mặn, cũng như các kiến trúc mạng nơ-ron tích chập (CNN). Dựa trên nền tảng lý thuyết đó, quá trình thực nghiệm được triển khai với việc thu thập dữ liệu ảnh vệ tinh Sentinel-1 và Sentinel-2 trên nền tảng Google Earth Engine. Việc kết hợp dữ liệu quang học và dữ liệu radar được xem là chìa khóa để khắc phục hạn chế về mây che phủ thường xuyên tại khu vực nhiệt đới, đồng thời cung cấp thêm thông tin về cấu trúc bề mặt và độ ẩm của thảm thực vật.

Sau khi xây dựng được bộ dữ liệu huấn luyện chất lượng cao với các nhãn được kiểm chứng, em đã thiết kế và tối ưu hóa một kiến trúc mạng CNN chuyên biệt, phù hợp với quy mô dữ liệu và đặc thù bài toán. Mô hình được huấn luyện để phân loại bốn trạng thái biến động chính gồm rừng ổn định, mất rừng, phi rừng và phục hồi rừng. Quá trình đánh giá mô hình được thực hiện nghiêm ngặt thông qua kỹ thuật kiểm định chéo (cross-validation) để đảm bảo độ tin cậy và khả năng tổng quát hóa của kết quả. Cuối cùng, mô hình tối ưu đã được áp dụng để phân loại toàn bộ khu vực quy hoạch lâm nghiệp của tỉnh Cà Mau, tạo ra bản đồ biến động rừng chi tiết và trực quan.

2.3. Kết quả đạt được

Kết quả thực tập đã đạt được những thành tựu đáng kể cả về mặt khoa học và ứng dụng thực tiễn. Nghiên cứu đã chứng minh hiệu quả vượt trội của việc tích hợp dữ liệu đa nguồn Sentinel-1 và Sentinel-2 so với việc chỉ sử dụng đơn lẻ dữ liệu quang học. Mô hình đề xuất đạt độ chính xác ẩn tượng lên tới 98,86% trên tập dữ liệu kiểm tra độc lập, với chỉ số ROC-AUC đạt 99,98%. Điều này khẳng định khả năng phân biệt chính xác

các đối tượng rừng và biến động rừng của thuật toán đã phát triển.

Về mặt sản phẩm, báo cáo đã xây dựng thành công bản đồ phân loại biến động rừng tỉnh Cà Mau giai đoạn 2024 – 2025 với độ phân giải 10m. Kết quả phân tích không gian cho thấy bức tranh tổng thể về tài nguyên rừng của tỉnh, trong đó phát hiện khoảng 7.282 ha rừng bị mất và 4.941 ha rừng được phục hồi. Bên cạnh đó, em cũng đã xây dựng một ứng dụng web trực tuyến trên nền tảng Google Earth Engine, cho phép người dùng truy cập, tương tác hiển thị dữ liệu và kết quả phân loại biến động rừng một cách dễ dàng và thuận tiện.

2.4. Kinh nghiệm rút ra

Quá trình thực tập không chỉ mang lại kiến thức chuyên môn mà còn giúp em tích lũy nhiều kinh nghiệm quý báu. Bài học lớn nhất là tầm quan trọng của việc xử lý và làm sạch dữ liệu. Trong các bài toán học sâu, chất lượng dữ liệu đầu vào quyết định phần lớn hiệu suất của mô hình. Việc hiểu rõ đặc điểm vật lý của ảnh viễn thám và thực hiện các bước tiền xử lý kỹ lưỡng là yếu tố then chốt để đạt được kết quả cao. Ngoài ra, kỹ năng tối ưu hóa mã nguồn và quản lý tài nguyên tính toán cũng được rèn luyện khi phải làm việc với lượng dữ liệu lớn trên quy mô toàn tỉnh. Cuối cùng, khả năng phân tích và biện giải kết quả từ "hộp đen" của mô hình học sâu giúp em hiểu sâu sắc hơn về cơ chế hoạt động của thuật toán và đưa ra những cải tiến phù hợp.

2.5. Kết luận

Kỳ thực tập đã hoàn thành tốt các mục tiêu đề ra, góp phần đưa ra một giải pháp công nghệ hiệu quả cho công tác giám sát rừng tại Cà Mau. Phương pháp tiếp cận kết hợp viễn thám đa nguồn và học sâu đã chứng minh được tính ưu việt và khả năng ứng dụng thực tiễn. Những kết quả đạt được không chỉ có giá trị học thuật mà còn là công cụ hỗ trợ đắc lực cho các nhà quản lý trong việc ra quyết định bảo vệ và phát triển rừng bền vững. Mặc dù vẫn còn một số hạn chế nhất định cần tiếp tục nghiên cứu, nhưng thành công bước đầu này là tiền đề vững chắc cho các hướng phát triển tiếp theo, mở rộng phạm vi ứng dụng sang các khu vực khác và nâng cao hơn nữa độ chính xác của mô hình.

PHẦN 3. BÁO CÁO CHI TIẾT KẾT QUẢ ĐẠT ĐƯỢC

3.1. Nghiên cứu đề bài và cơ sở lý thuyết

Để thực hiện đề tài giám sát biến động rừng tỉnh Cà Mau, quá trình nghiên cứu đã được triển khai một cách bài bản, bắt đầu từ việc tìm hiểu tổng quan về đối tượng nghiên cứu là tài nguyên rừng, phân tích bối cảnh mất rừng trên thế giới và tại Việt Nam, đến việc nắm vững các công nghệ lõi bao gồm viễn thám và học sâu.

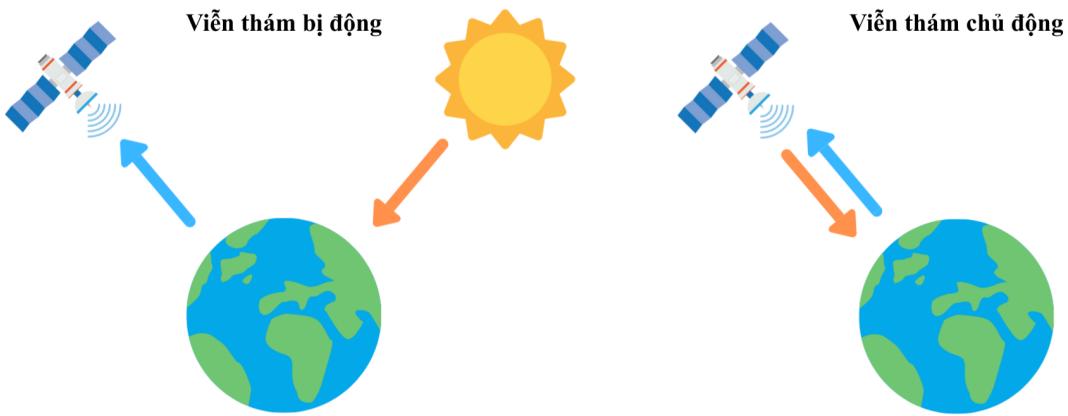
3.1.1. Tổng quan về rừng và biến động rừng

Rừng là một hệ sinh thái phức tạp đóng vai trò then chốt trong việc duy trì sự sống trên Trái Đất. Theo định nghĩa của Tổ chức Lương thực và Nông nghiệp Liên Hợp Quốc (FAO), rừng được xác định là vùng đất có diện tích tối thiểu 0,5 ha với độ che phủ tán cây trên 10% và chiều cao cây khi trưởng thành đạt ít nhất 5 mét [9]. Đối với tỉnh Cà Mau, hệ sinh thái rừng ngập mặn chiếm ưu thế và có ý nghĩa đặc biệt quan trọng. Rừng ngập mặn không chỉ là bức tường xanh chắn sóng, bảo vệ bờ biển khỏi xói mòn mà còn là nơi lưu trữ carbon hiệu quả gấp 3–5 lần so với rừng nhiệt đới trên cạn [6], góp phần giảm thiểu tác động của biến đổi khí hậu.

Tuy nhiên, tình trạng mất rừng đang diễn ra ở mức báo động trên phạm vi toàn cầu. Các báo cáo gần đây chỉ ra rằng thế giới đã mất hàng trăm triệu hécta rừng trong vài thập kỷ qua. Tại các khu vực nhiệt đới như Amazon hay Đông Nam Á, rừng đang bị suy giảm nhanh chóng do áp lực chuyển đổi mục đích sử dụng đất sang nông nghiệp, chăn nuôi và khai thác gỗ. Mặc dù tại Việt Nam, độ che phủ rừng đã có xu hướng tăng nhờ các nỗ lực trồng rừng [1], nhưng chất lượng rừng tự nhiên vẫn còn nhiều lo ngại. Tại khu vực Đồng bằng sông Cửu Long, đặc biệt là Cà Mau, rừng ngập mặn chịu áp lực lớn từ việc mở rộng diện tích nuôi trồng thủy sản [3], cũng như tác động tiêu cực của biến đổi khí hậu gây ra hiện tượng xâm nhập mặn và sạt lở bờ biển [16].

3.1.2. Cơ sở lý thuyết về viễn thám vệ tinh

Công nghệ viễn thám đóng vai trò nền tảng trong việc cung cấp dữ liệu đầu vào cho hệ thống giám sát rừng. Báo cáo tập trung nghiên cứu hai hệ thống vệ tinh tiên tiến thuộc chương trình Copernicus của Cơ quan Vũ trụ Châu Âu (ESA) là Sentinel-1 và Sentinel-2.



Hình 3.1: Nguyên lý viễn thám bị động và chủ động

Sentinel-2 là vệ tinh quang học (viễn thám bị động), hoạt động dựa trên nguyên lý ghi nhận bức xạ mặt trời phản xạ từ bề mặt Trái Đất. Hệ thống này cung cấp dữ liệu đa phô với 13 kênh bước sóng khác nhau, từ vùng ánh sáng nhìn thấy đến hồng ngoại sóng ngắn. Đặc biệt, các kênh có độ phân giải không gian 10m và 20m cung cấp thông tin chi tiết về đặc tính phổ của thảm thực vật. Các chỉ số thực vật được tính toán từ dữ liệu Sentinel-2 đóng vai trò quan trọng trong việc đánh giá sức khỏe rừng. Chỉ số thực vật chuẩn hóa (NDVI) được tính bằng công thức:

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (3.1)$$

trong đó NIR là phản xạ ở kênh hồng ngoại gần và Red là phản xạ ở kênh đỏ. NDVI phản ánh mật độ và sức khỏe của diệp lục [20].

Bên cạnh đó, chỉ số cháy chuẩn hóa (NBR) cũng được sử dụng để phát hiện các dấu hiệu suy thoái rừng, được tính bằng công thức:

$$NBR = \frac{NIR - SWIR2}{NIR + SWIR2} \quad (3.2)$$

trong đó SWIR2 là phản xạ ở kênh hồng ngoại sóng ngắn 2 (B12) [15].

Chỉ số độ ẩm chuẩn hóa (NDMI) phản ánh hàm lượng nước trong tán lá, được tính như sau:

$$NDMI = \frac{NIR - SWIR1}{NIR + SWIR1} \quad (3.3)$$

trong đó SWIR1 là phản xạ ở kênh hồng ngoại sóng ngắn 1 (B11) [10].

Tuy nhiên, hạn chế lớn nhất của viễn thám quang học là sự phụ thuộc vào điều kiện thời tiết, đặc biệt là mây che phủ thường xuyên ở các vùng nhiệt đới. Để khắc phục điều

này, dữ liệu từ vệ tinh Sentinel-1 sử dụng công nghệ radar khẩn độ tổng hợp (SAR) được tích hợp [7, 22]. Đây là hệ thống viễn thám chủ động, tự phát ra sóng vô tuyến băng C và ghi nhận tín hiệu tán xạ ngược. Do sử dụng sóng vi ba có bước sóng dài, tín hiệu radar có khả năng xuyên qua mây và hoạt động cả ngày lẫn đêm. Trong chế độ giao thoa ké (Interferometric Wide - IW), Sentinel-1 cung cấp dữ liệu ở hai kênh phân cực VV và VH. Phân cực VH đặc biệt nhạy cảm với cấu trúc thể tích của tán lá rùng, trong khi phân cực VV phản ánh độ nhám và độ ẩm bề mặt. Việc kết hợp dữ liệu quang học và radar cho phép khai thác ưu điểm của cả hai loại cảm biến, nâng cao độ tin cậy của việc phát hiện biến động rùng.

Bảng 3.1: Thông số kỹ thuật của Sentinel-1 và Sentinel-2

Thông số	Sentinel-1	Sentinel-2
Loại cảm biến	Radar (chủ động)	Quang học (bị động)
Dải sóng	C-band (5,55 cm)	443–2.190 nm
Số kênh/phân cực	2 (VV, VH)	13 dải phổ
Độ phân giải không gian	10m (IW mode)	10/20/60m
Độ rộng dải quét	250 km	290 km
Chu kỳ quay lại	6–12 ngày	5–10 ngày
Hoạt động qua mây	Có	Không
Thông tin thu nhận	Cấu trúc, độ ẩm, độ nhám	Phản xạ phổ, chỉ số thực vật

3.1.3. Học sâu và Mạng nơ-ron tích chập

Trong kỷ nguyên số, học sâu đã tạo ra cuộc cách mạng trong lĩnh vực thị giác máy tính và xử lý ảnh [18]. Mạng nơ-ron tích chập là kiến trúc mạng nơ-ron chuyên biệt cho dữ liệu dạng lưới như hình ảnh, được xem là công cụ mạnh mẽ nhất hiện nay để phân loại ảnh viễn thám [24, 25].

Khác với các phương pháp học máy truyền thống yêu cầu trích xuất đặc trưng thủ công, CNN có khả năng tự động học các đặc trưng từ mức thấp đến mức cao thông qua các lớp tích chập. Thành phần cốt lõi của CNN là phép toán tích chập 2D, được mô tả bởi công thức:

$$(I * K)(i, j) = \sum_m \sum_n I(i + m, j + n) \times K(m, n) \quad (3.4)$$

trong đó I là ảnh đầu vào và K là kernel. Các kernel này trượt qua toàn bộ ảnh để phát

hiện các đặc trưng cục bộ như cạnh, góc, kết cấu bề mặt. Thông qua việc xếp chồng nhiều lớp tích chập, mạng có thể học được các mẫu phức tạp mang tính trừu tượng.

Bên cạnh các lớp tích chập, kiến trúc CNN còn bao gồm các hàm kích hoạt phi tuyến như ReLU (Rectified Linear Unit) giúp mô hình có khả năng học các mối quan hệ phi tuyến tính phức tạp trong dữ liệu. Các lớp gộp (Pooling), điển hình là Max Pooling hoặc Global Average Pooling, được sử dụng để giảm chiều dữ liệu, giảm số lượng tham số cần tính toán và tăng tính bất biến của mô hình đối với các biến đổi nhỏ của đối tượng.

Quá trình huấn luyện mạng CNN thực chất là bài toán tối ưu hóa nhằm tìm ra bộ trọng số sao cho sai số giữa dự đoán của mô hình và nhãn thực tế là nhỏ nhất. Hàm mất mát Cross-Entropy thường được sử dụng cho bài toán phân loại đa lớp:

$$L = - \sum_i y_i \log(\hat{y}_i) \quad (3.5)$$

trong đó y_i là nhãn thực tế và \hat{y}_i là xác suất dự đoán. Thuật toán tối ưu hóa như AdamW [19] được áp dụng để cập nhật trọng số dựa trên đạo hàm của hàm mất mát. Để ngăn ngừa hiện tượng quá khớp (overfitting) thường gặp khi huấn luyện với bộ dữ liệu hạn chế, các kỹ thuật điều chỉnh như Dropout [21] (ngắt ngẫu nhiên các kết nối) và Batch Normalization [14] (chuẩn hóa theo lô) được tích hợp vào kiến trúc mạng.

3.1.4. Tổng quan tình hình nghiên cứu

Lĩnh vực giám sát rừng bằng viễn thám đã trải qua một quá trình phát triển dài từ các phương pháp giải đoán thủ công, đến các thuật toán học máy cơ bản như Rừng ngẫu nhiên (Random Forest) [4] hay Support Vector Machine (SVM) [5]. Các nghiên cứu kinh điển như của Hansen và cộng sự [11] đã xây dựng bản đồ mất rừng toàn cầu dựa trên dữ liệu Landsat. Tuy nhiên, trong những năm gần đây, xu hướng nghiên cứu đã chuyển dịch mạnh mẽ sang ứng dụng học sâu.

Nhiều công trình nghiên cứu trên thế giới đã chứng minh tính ưu việt của CNN so với các phương pháp truyền thống trong việc phân loại sử dụng đất và phát hiện biến động [13, 17]. Ví dụ, các nghiên cứu tại khu vực Amazon hay Indonesia đã áp dụng thành công mô hình U-Net hay ResNet để phát hiện chặt phá rừng với độ chính xác rất cao. Riêng về rừng ngập mặn Cà Mau, nghiên cứu của Vo và cộng sự [23] đã thực hiện giám sát động thái rừng ngập mặn hàng năm sử dụng dữ liệu Landsat-7-8. Tại Việt Nam, các nghiên cứu ứng dụng viễn thám giám sát rừng cũng đang phát triển, tuy nhiên

phần lớn vẫn dựa trên các phương pháp học máy truyền thống và sử dụng đơn nguồn dữ liệu quang học. Việc ứng dụng các mô hình học sâu hiện đại kết hợp dữ liệu đa nguồn Sentinel-1 và Sentinel-2 [12] cho đối tượng rừng ngập mặn Cà Mau vẫn là một hướng đi mới mẻ, tiềm năng và hứa hẹn mang lại những đột phá về độ chính xác và hiệu quả giám sát.

3.2. Báo cáo chi tiết nội dung nghiên cứu và thực hiện

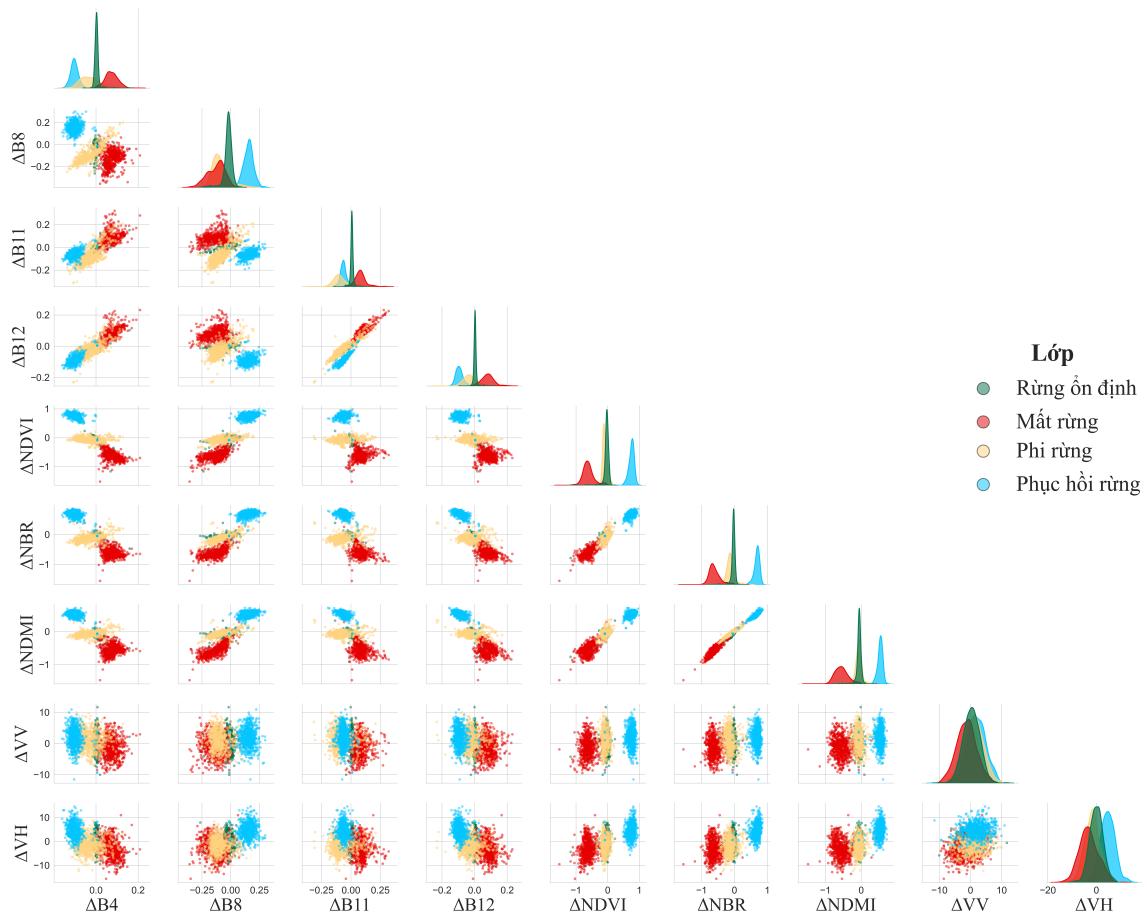
3.2.1. Dữ liệu nghiên cứu

Dữ liệu đóng vai trò quyết định trong các mô hình học sâu. Để xây dựng hệ thống giám sát biến động rừng tỉnh Cà Mau, nghiên cứu này đã thiết lập một quy trình thu thập và xử lý dữ liệu nghiêm ngặt từ các nguồn vệ tinh miễn phí nhưng chất lượng cao của Chương trình Copernicus. Dữ liệu thực địa được Công ty TNHH Tư vấn và Phát triển Đồng Xanh (GFD) cung cấp trong khuôn khổ đề tài “Ứng dụng công nghệ viễn thám, ảnh vệ tinh để quản lý, giám sát tài nguyên rừng và sạt lở ven biển trên địa bàn tỉnh Cà Mau” [2].

Nguồn dữ liệu chính được sử dụng bao gồm ảnh quang học từ vệ tinh Sentinel-2 [8] và ảnh radar khẩu độ tổng hợp (SAR) từ vệ tinh Sentinel-1 [7, 22]. Đối với Sentinel-2, nghiên cứu lựa chọn sản phẩm cấp độ L2A (Surface Reflectance) đã được hiệu chỉnh khí quyển để loại bỏ ảnh hưởng của sương mù và khí dung, đảm bảo giá trị phản xạ bề mặt chính xác. Các ảnh được thu thập trong giai đoạn mùa khô, cụ thể là tháng 01/2024 (kỳ trước) và tháng 02/2025 (kỳ sau), nhằm giảm thiểu tối đa sự che phủ của mây và đảm bảo tính tương đồng về điều kiện chiếu sáng giữa hai thời điểm. Toàn bộ quy trình truy xuất được thực hiện trên nền tảng điện toán đám mây Google Earth Engine (GEE), cho phép lọc và xử lý khối lượng lớn dữ liệu một cách hiệu quả. Các ảnh Sentinel-2 sau khi lọc mây với ngưỡng xác suất 50% đã được dùng để tính toán các chỉ số thực vật quan trọng.

Song song đó, dữ liệu Sentinel-1 được thu thập ở chế độ Interferometric Wide (IW) với hai kênh phân cực VV và VH. Dữ liệu này được tiền xử lý các bước tiêu chuẩn gồm hiệu chỉnh quỹ đạo, khử nhiễu nhiệt, hiệu chỉnh bức xạ và hiệu chỉnh địa hình để tạo ra ảnh tán xạ ngược (backscatter) có đơn vị dB. Ưu điểm vượt trội của dữ liệu radar là khả năng xuyên thấu qua mây, cung cấp thông tin liên tục về cấu trúc bề mặt ngay cả khi dữ liệu quang học bị gián đoạn.

Thay vì đưa trực tiếp ảnh thô vào mô hình, nghiên cứu đã tiến hành trích xuất một tập hợp các đặc trưng giàu thông tin nhằm nâng cao khả năng phân loại của mô hình. Tổng cộng 27 đặc trưng đã được xây dựng cho mỗi điểm ảnh. Từ dữ liệu Sentinel-2, bốn kênh phổ cơ bản gồm Blue (B2), Green (B3), Red (B4) và Near-Infrared (B8) cùng hai kênh hồng ngoại sóng ngắn (SWIR) được sử dụng. Bên cạnh đó, ba chỉ số thực vật chuyên biệt được tính toán gồm NDVI (chỉ số thực vật chuẩn hóa) phản ánh mật độ xanh, NBR (tỷ số cháy chuẩn hóa) nhạy cảm với sự mất mát sinh khối, và NDMI (chỉ số độ ẩm) phản ánh hàm lượng nước trong lá. Từ dữ liệu Sentinel-1, hai kênh phân cực VV và VH cung cấp thông tin về độ nhám và cấu trúc thể tích của tán rừng. Đặc biệt, để nắm bắt thông tin về sự thay đổi, nghiên cứu đã tính toán giá trị chênh lệch (delta) của tất cả các đặc trưng này giữa hai thời điểm nghiên cứu. Như vậy, vector đặc trưng đầu vào cho mô hình bao gồm giá trị tại thời điểm trước, thời điểm sau và giá trị thay đổi của cả dữ liệu quang học và radar.

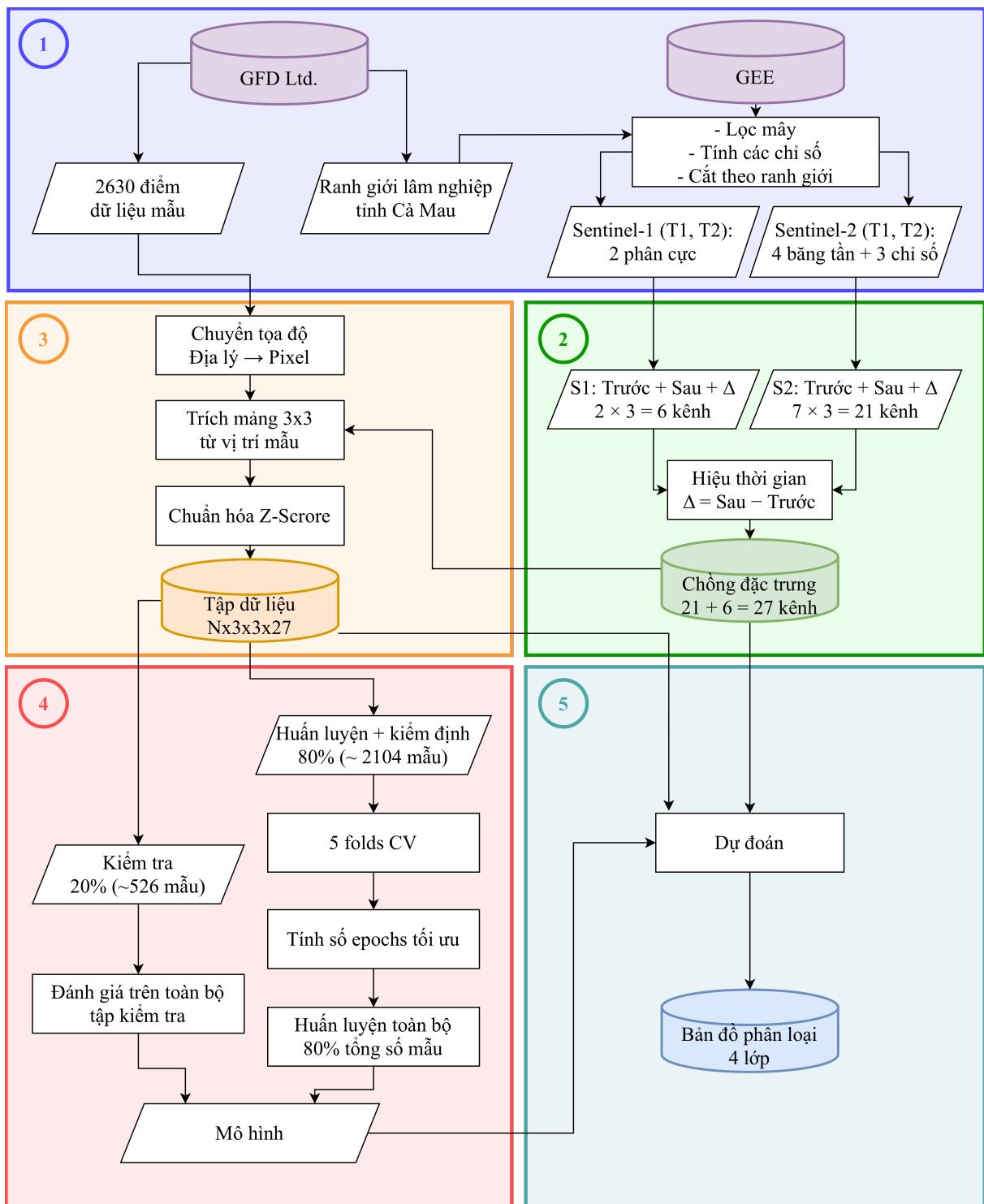


Hình 3.2: Phân bố các đặc trưng delta theo lớp biến động rừng

Kết quả phân tích phân bố cho thấy các chỉ số thực vật delta (ΔNDVI , ΔNBR ,

ΔNDMI) thể hiện khả năng phân biệt vượt trội giữa các lớp biến động, trong khi các kênh ra-đa (ΔVV , ΔVH) có sự chồng chéo đáng kể nhưng vẫn mang lại thông tin bổ sung giá trị trong điều kiện mây che phủ.

3.2.2. Phương pháp nghiên cứu



Hình 3.3: Sơ đồ quy trình tổng quát của nghiên cứu

Quy trình nghiên cứu tổng thể được thiết kế gồm năm giai đoạn chính: (1) Thu thập và tiền xử lý dữ liệu vệ tinh đa nguồn, (2) Trích xuất hệ đặc trưng tối ưu, (3) Xây dựng

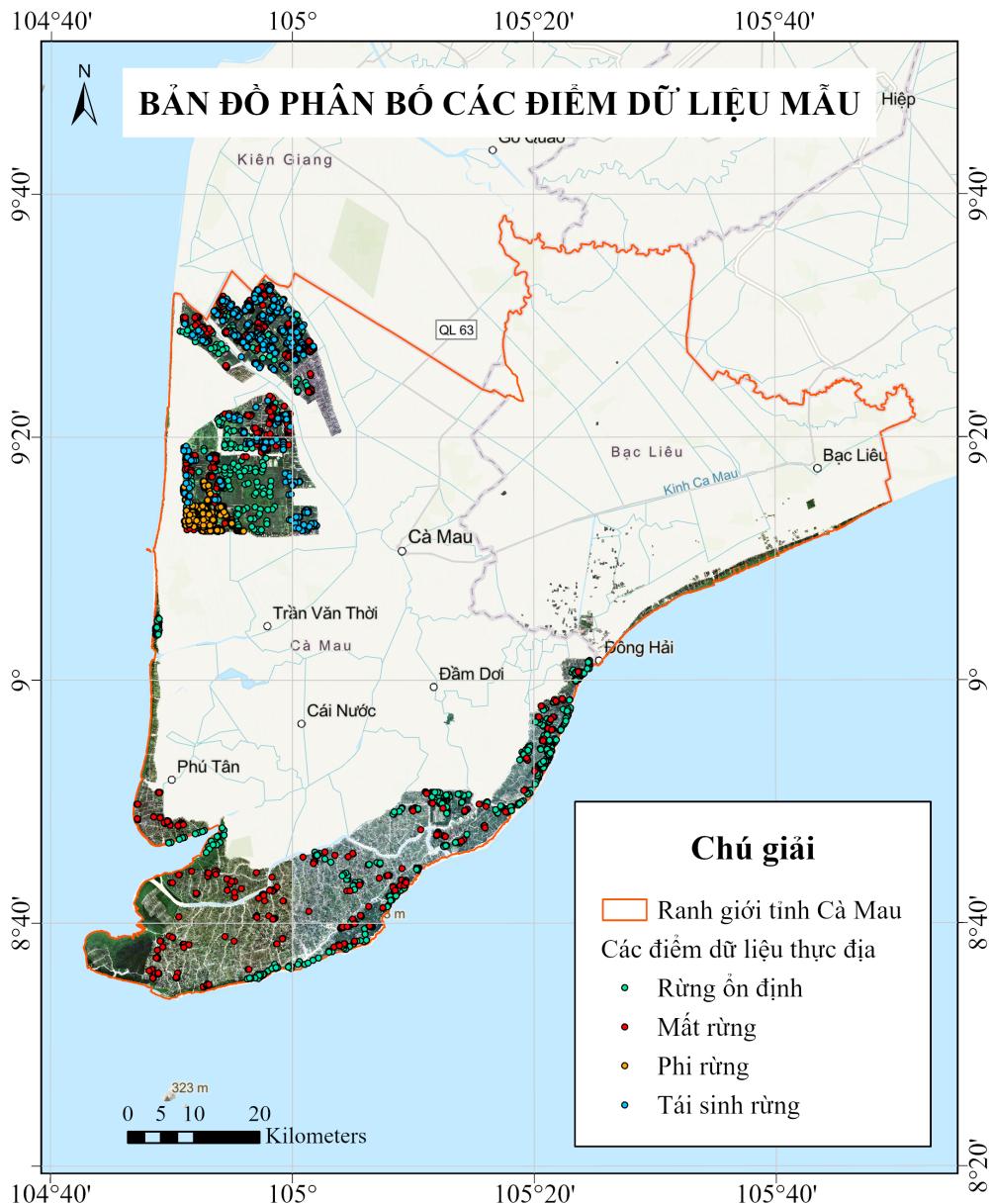
bộ dữ liệu mẫu huấn luyện, (4) Huấn luyện và tinh chỉnh mô hình học sâu, và (5) Áp dụng mô hình để thành lập bản đồ biến động rừng.

Quy trình chuẩn bị dữ liệu huấn luyện được thực hiện cẩn trọng để đảm bảo tính đại diện và khách quan. Dữ liệu thực địa gồm 2.630 điểm mẫu được thu thập và dán nhãn thủ công cho bốn lớp đối tượng: Rừng ổn định, Mất rừng, Phi rừng và Phục hồi rừng. Các điểm mẫu này được phân bố đều trên toàn vùng nghiên cứu để tránh thiên lệch không gian.

Bảng 3.2: Thống 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

Bộ dữ liệu thực địa bao gồm 2.630 điểm mẫu được biên tập kỹ lưỡng, với sự phân bổ tương đối cân bằng giữa bốn lớp đối tượng nghiên cứu. Việc đảm bảo cân bằng dữ liệu đóng vai trò quan trọng trong việc hạn chế hiện tượng thiên lệch (bias) của mô hình về phía các lớp chiếm đa số.



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

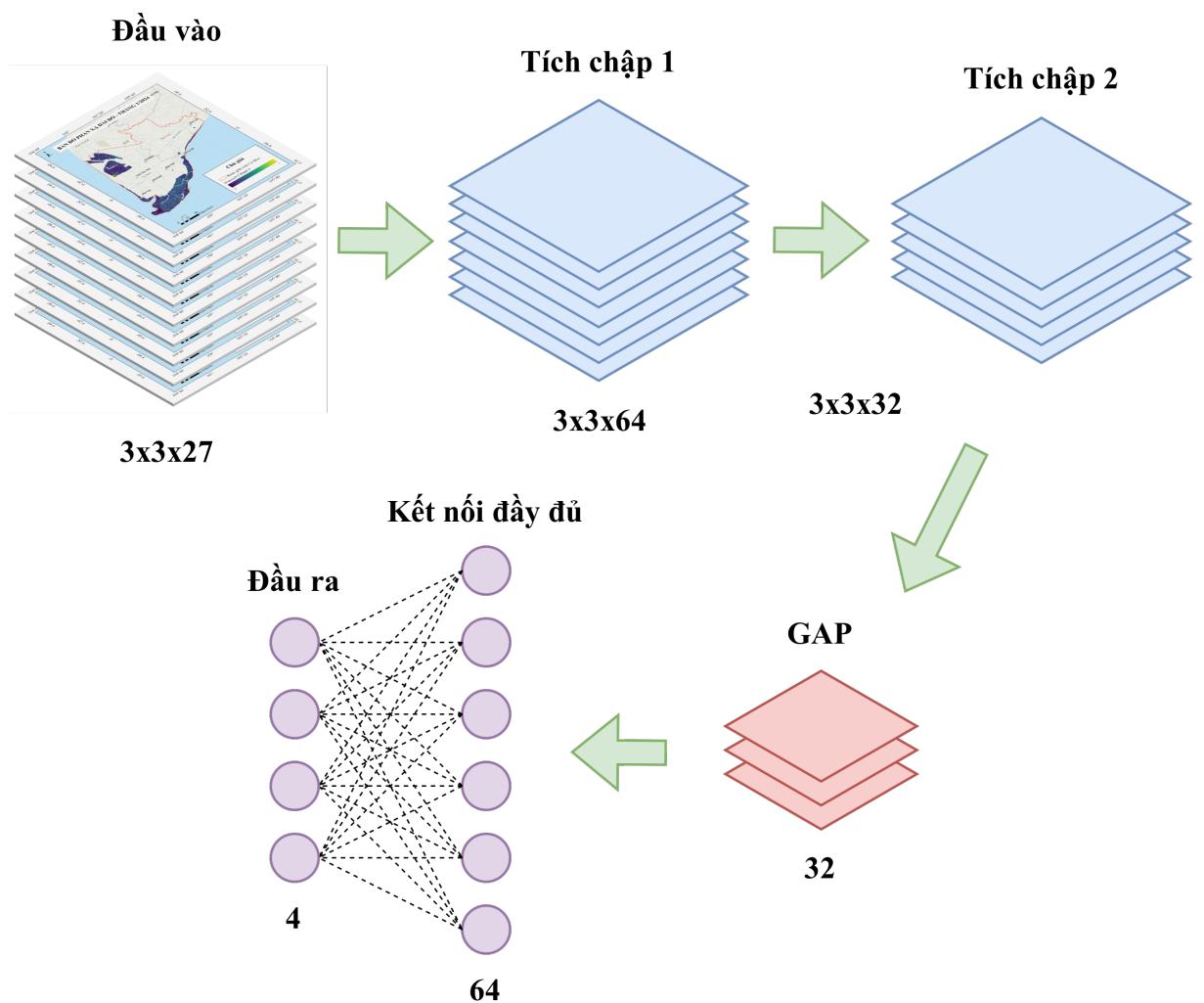
Mạng lưới điểm mẫu được thiết lập phân tán rộng khắp trên toàn bộ phạm vi địa lý của khu vực nghiên cứu, bao gồm các vùng sinh thái đặc trưng khác nhau. Chiến lược lấy mẫu này đảm bảo tính đại diện về mặt không gian cho các loại hình biến động rừng, tăng cường khả năng tổng quát hóa của mô hình.

Để khai thác thông tin ngữ cảnh không gian, thay vì phân loại từng điểm ảnh độc lập, nghiên cứu sử dụng phương pháp trích xuất các ô kích thước 3×3 điểm ảnh xung quanh mỗi điểm mẫu. Như vậy, mỗi mẫu dữ liệu đầu vào là một khối tensor có kích thước

(27, 3, 3). Trước khi đưa vào mô hình, toàn bộ dữ liệu được chuẩn hóa theo phương pháp Z-score để đưa các đặc trưng về cùng một thang đo, giúp quá trình tối ưu hóa hội tụ nhanh hơn.

Dữ liệu được chia thành hai tập độc lập: tập huấn luyện/kiểm định (80%) và tập kiểm tra (20%). Việc phân chia được thực hiện theo phương pháp phân tầng để đảm bảo tỷ lệ các lớp trong mỗi tập con tương đương với tập dữ liệu gốc. Tập kiểm tra được giữ hoàn toàn riêng biệt và chỉ được sử dụng một lần duy nhất để đánh giá hiệu năng cuối cùng của mô hình.

Nghiên cứu đề xuất một kiến trúc mạng nơ-ron tích chập được thiết kế riêng cho bài toán này, tối ưu hóa sự cân bằng giữa độ chính xác và chi phí tính toán. Kiến trúc mạng bao gồm các thành phần chính sau:



Hình 3.5: Sơ đồ kiến trúc mạng CNN để xuất

Kiến trúc mạng nơ-ron tích chập (CNN) đề xuất được thiết kế tối ưu cho dữ liệu viễn thám độ phân giải trung bình, kết hợp các kỹ thuật hiện đại như Global Average Pooling và Dropout để kiểm soát hiện tượng quá khớp (overfitting) trên tập dữ liệu mẫu hạn chế, đồng thời đảm bảo hiệu suất tính toán.

Khối tích chập đầu tiên bao gồm lớp tích chập (Convolution) với 64 bộ lọc kích thước 3×3 , tiếp theo là lớp chuẩn hóa theo lô (Batch Normalization) và hàm kích hoạt ReLU. Lớp tích chập này có nhiệm vụ trích xuất các đặc trưng không gian và quang phổ sơ cấp từ dữ liệu đầu vào. Ngay sau đó là lớp Dropout 2D với tỷ lệ 0.7, giúp ngăn ngừa nhiên các kết nối để ngăn ngừa hiện tượng quá khớp (overfitting) - một vấn đề thường gặp khi huấn luyện mạng sâu trên tập dữ liệu nhỏ.

Khối tích chập thứ hai sử dụng 32 bộ lọc để tiếp tục tinh chế các đặc trưng ở mức độ trừu tượng cao hơn, đồng thời giảm chiều dữ liệu để nén thông tin. Cấu trúc của khối này tương tự khối đầu tiên với đầy đủ các thành phần chuẩn hóa và kích hoạt.

Điểm đặc biệt của kiến trúc này là việc sử dụng lớp Global Average Pooling (GAP) ở giai đoạn cuối thay vì làm phẳng (flatten) toàn bộ tensor. GAP tính trung bình giá trị của mỗi bản đồ đặc trưng, giúp giảm mạnh số lượng tham số của mô hình và tăng khả năng bắt biến đổi với các dịch chuyển nhỏ trong ảnh. Cuối cùng, các đặc trưng được đưa qua các lớp kết nối đầy đủ (Fully Connected) để thực hiện phân loại vào 4 lớp mục tiêu. Tổng số tham số của mô hình chỉ khoảng 36.676, đảm bảo mô hình gọn nhẹ (lightweight) và có thể huấn luyện nhanh chóng trên phần cứng thông dụng.

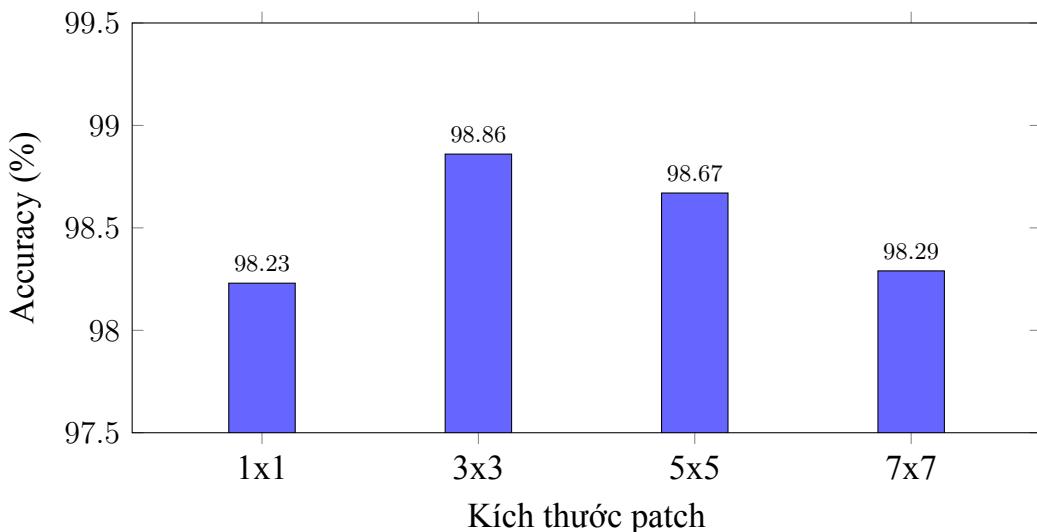
3.2.3. Kết quả thực nghiệm và thảo luận

Trước khi huấn luyện mô hình cuối cùng, một loạt các thí nghiệm đã được thực hiện để lựa chọn tham số tối ưu. Kết quả so sánh các kích thước patch khác nhau (1×1 , 3×3 , 5×5 , 7×7) cho thấy kích thước 3×3 mang lại hiệu suất cao nhất với độ chính xác 98,86%. Kích thước này đủ lớn để nắm bắt ngữ cảnh không gian cục bộ nhưng cũng đủ nhỏ để tránh nhiễu từ các đối tượng lân cận không liên quan.

Bảng 3.3: So sánh các kích thước patch

Kích thước patch	Accuracy	ROC-AUC	Thời gian huấn luyện	Số tham số
1x1	98,23%	99,78%	12,5s	25.348
3x3 (nghiên cứu này)	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

Các thực nghiệm đánh giá kích thước cửa sổ trượt (patch size) cho thấy kích thước 3x3 điểm ảnh đạt hiệu suất phân loại tối ưu (Accuracy 98,86%). Kết quả này chỉ ra rằng ngũ cành không gian cục bộ trong phạm vi 30m x 30m là phù hợp nhất để đặc tả các đối tượng rừng ngập mặn vốn có tính phân mảnh cao tại khu vực nghiên cứu, trong khi các kích thước lớn hơn có xu hướng bị nhiễu từ các đối tượng lân cận.



Hình 3.6: So sánh Accuracy theo các kích thước patch

Biểu đồ so sánh độ chính xác xác nhận xu hướng giảm dần của hiệu suất khi kích thước patch tăng lên hoặc giảm xuống so với kích thước tối ưu 3x3. Điều này cũng có lựa chọn kích thước đầu vào cho mô hình, cân bằng giữa việc nắm bắt thông tin ngũ cành và duy trì độ chi tiết của đối tượng.

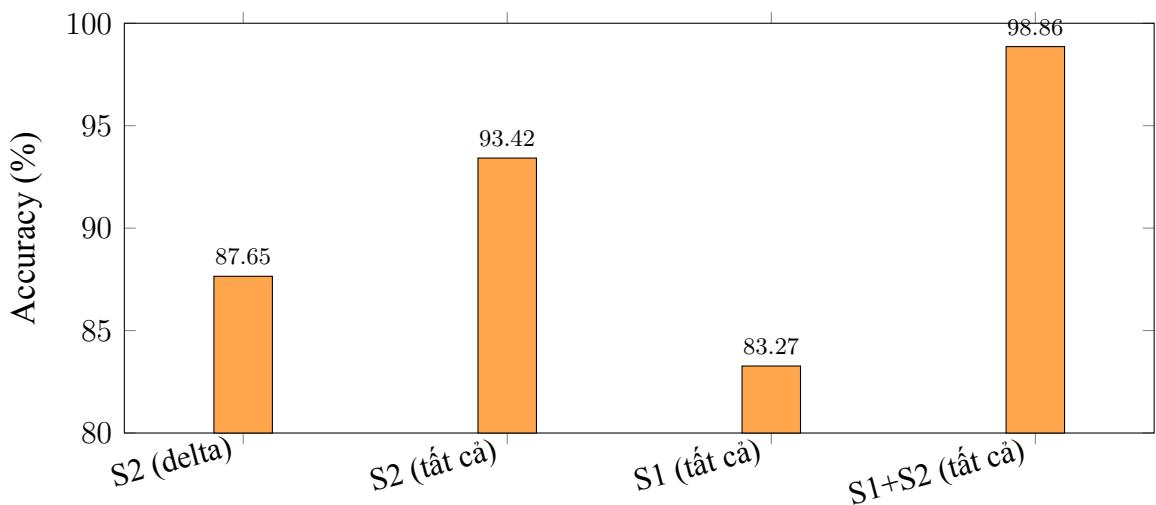
Nghiên cứu cũng tiến hành đánh giá vai trò của từng nguồn dữ liệu thông qua phương pháp loại trừ (ablation study). Kết quả chỉ ra rằng nếu chỉ sử dụng đơn lẻ dữ liệu quang học Sentinel-2, độ chính xác chỉ đạt 93,42%. Khi kết hợp thêm dữ liệu radar

Sentinel-1, độ chính xác tăng vọt lên 98,86%. Điều này khẳng định giả thuyết ban đầu rằng sự kết hợp dữ liệu đa nguồn mang lại thông tin bổ sung quan trọng: tín hiệu quang học phản ánh đặc tính sinh học của lá cây, trong khi tín hiệu radar phản ánh cấu trúc vật lý của thân và tán rừng.

Bảng 3.4: Nghiên cứu loại trừ các nguồn dữ liệu

Cấu hình	Số đặc trưng	Accuracy	ROC-AUC
Chỉ Sentinel-2 (delta)	7	87,65%	94,12%
Sentinel-2 (trước + sau + delta)	21	93,42%	97,58%
Chỉ Sentinel-1 (trước + sau + delta)	6	83,27%	91,45%
S1 + S2 (tất cả)	27	98,86%	99,98%

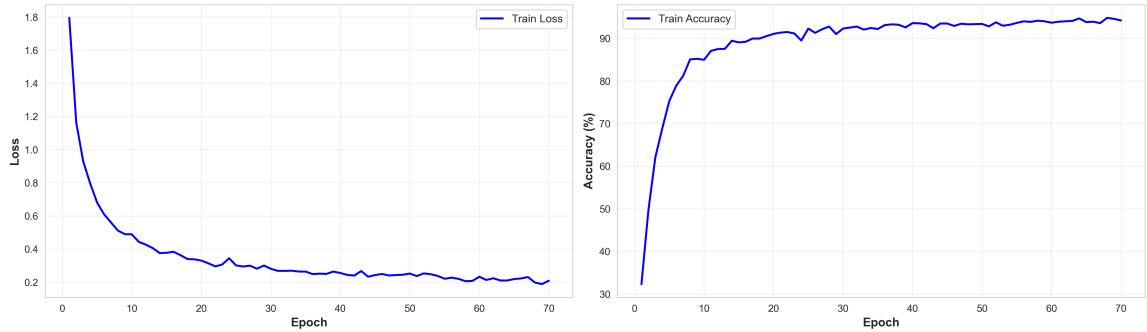
Phân tích định lượng vai trò của từng nguồn dữ liệu khẳng định tính ưu việt của phương pháp đa nguồn. Sự kết hợp giữa thông tin quang phổ từ Sentinel-2 và thông tin cấu trúc từ Sentinel-1 giúp cải thiện đáng kể độ chính xác phân loại, đặc biệt trong việc giảm thiểu sự nhầm lẫn giữa các đối tượng có đặc điểm phổ tương tự.



Hình 3.7: So sánh Accuracy theo các nguồn dữ liệu

Kết quả so sánh trực quan minh họa rõ lợi ích của việc tích hợp dữ liệu. Nhóm đặc trưng kết hợp (S1+S2) đạt được sự ổn định cao nhất và vượt trội hơn hẳn so với các trường hợp chỉ sử dụng một nguồn dữ liệu duy nhất, khẳng định giả thuyết về tính bổ sung giữa hai loại cảm biến.

Quá trình huấn luyện sử dụng chiến lược kiểm định chéo 5 phần (5-fold cross-validation) để đảm bảo độ tin cậy. Kết quả trung bình cho thấy độ chính xác ổn định ở mức 98,86% với độ lệch chuẩn rất thấp, chứng tỏ mô hình không phụ thuộc vào cách chia dữ liệu.



Hình 3.8: Diễn biến loss và accuracy trong quá trình huấn luyện mô hình cuối cùng

Diễn biến của hàm mất mát (loss) và độ chính xác (accuracy) trong quá trình huấn luyện cho thấy mô hình hội tụ ổn định sau khoảng 70 epoch. Không quan sát thấy dấu hiệu của hiện tượng quá khớp nghiêm trọng, chứng tỏ các kỹ thuật điều chỉnh (regularization) đã được áp dụng hiệu quả.

Trên tập kiểm tra độc lập, mô hình đạt độ chính xác toàn cục (Overall Accuracy) là 98,86% và chỉ số ROC-AUC đạt 99,98%. Phân tích chi tiết qua ma trận nhầm lẫn (Confusion Matrix) cho thấy mô hình phân loại gần như tuyệt đối chính xác các lớp Phi rừng và Phục hồi rừng (độ chính xác 100%). Một số ít sự nhầm lẫn (tỷ lệ lỗi khoảng 1.1%) xảy ra giữa lớp Rừng ổn định và Mất rừng.

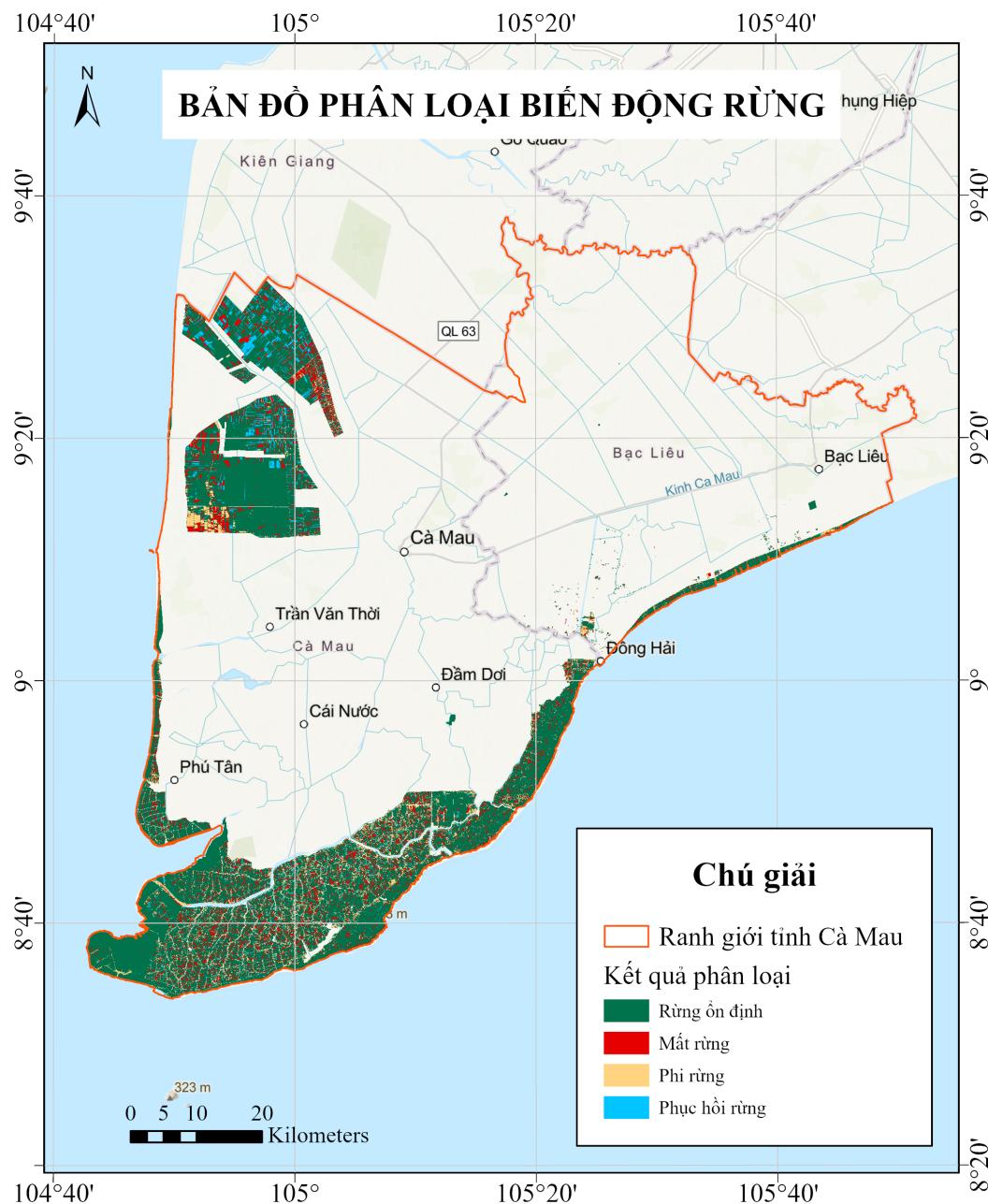
	Rừng ổn định	129	2	0	0
Thực tế	Mất rừng	4	126	0	0
	Phi rừng	0	0	133	0
	Phục hồi rừng	0	0	0	132
	Rừng ổn định	Mất rừng	Phi rừng	Phục hồi rừng	Dự đoán

Hình 3.9: Ma trận nhầm lẫn trên tập kiểm tra ($n=526$, Accuracy: 98,86%)

Kết quả đánh giá trên tập kiểm tra độc lập cho thấy mô hình đạt độ tin cậy cao với tỷ lệ phân loại sai thấp (1,14%). Các sai số chủ yếu tập trung vào việc phân biệt giữa lớp Rừng ổn định và Mất rừng, nguyên nhân có thể do sự tương đồng về đặc trưng phổ tại các khu vực rừng suy thoái nhẹ hoặc các điểm ảnh hỗn hợp tại vùng biên.

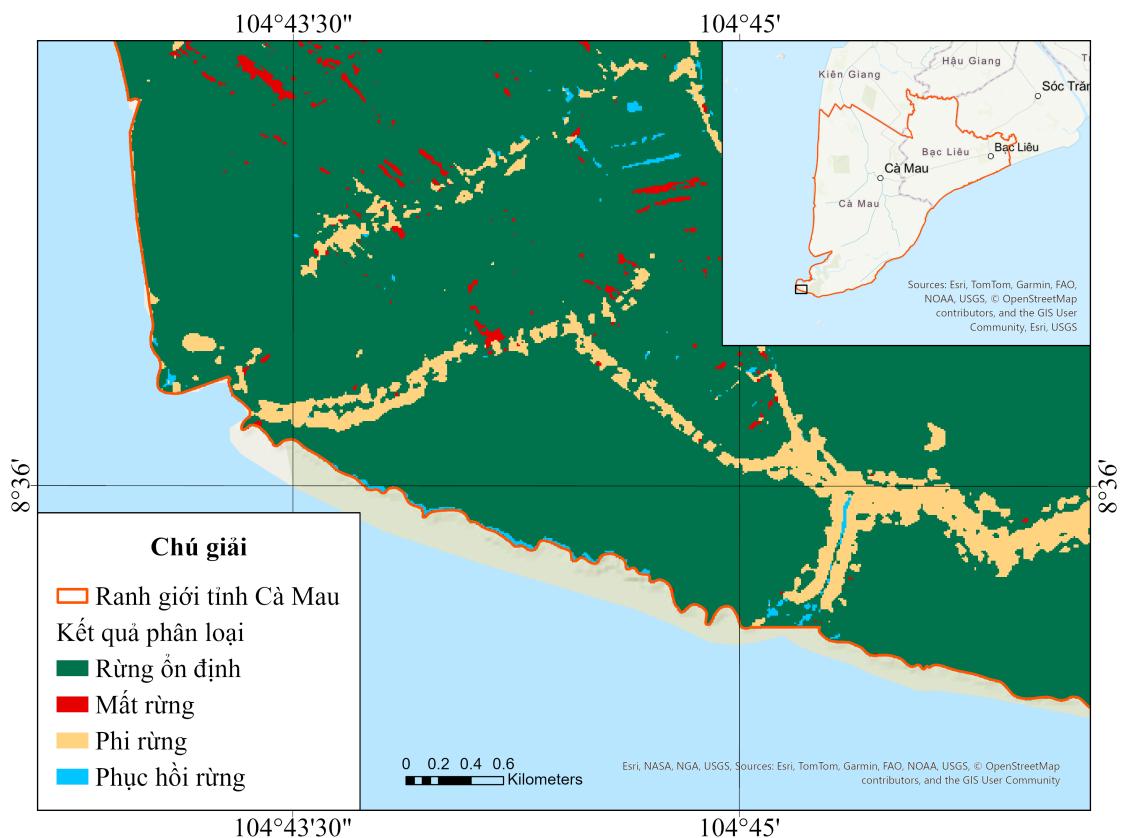
Nguyên nhân chủ yếu do các khu vực rừng bị suy thoái nhẹ hoặc mới trồng lại có đặc điểm phổ và cấu trúc chưa rõ ràng, nằm ở ranh giới giữa hai trạng thái này. Tuy nhiên, tỷ lệ lỗi này là hoàn toàn chấp nhận được trong bài toán giám sát diện rộng.

Mô hình đã huấn luyện được áp dụng để phân loại toàn bộ khu vực quy hoạch lâm nghiệp tỉnh Cà Mau với diện tích hơn 162.000 ha. Bản đồ kết quả cho thấy diện tích rừng ổn định chiếm phần lớn (khoảng 74,3%), tập trung chủ yếu ở các khu vực rừng phòng hộ và Vườn Quốc gia Mũi Cà Mau. Diện tích mất rừng được phát hiện khoảng 7.282 ha (tương đương 4,48%), phân bố rải rác và thường gắn liền với các khu vực ven biển hoặc xen kẽ trong các vùng nuôi trồng thủy sản. Ngược lại, diện tích rừng phục hồi đạt khoảng 4.941 ha (3,04%).



Hình 3.10: Bản đồ phân loại biến động rừng tỉnh Cà Mau

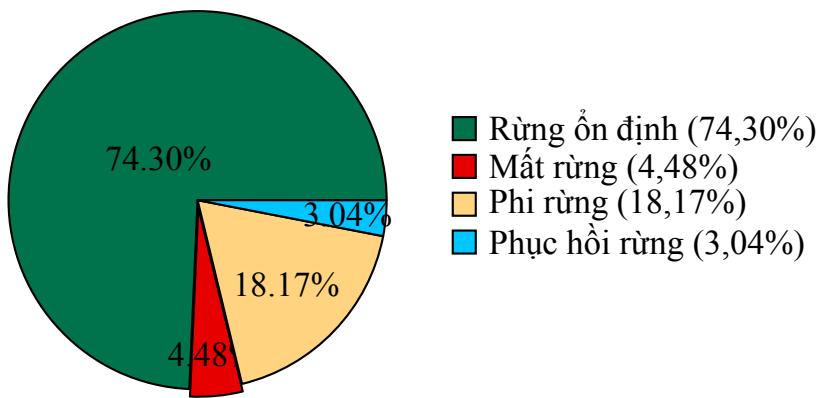
Bản đồ biến động rừng phản ánh cấu trúc không gian của lớp phủ, với diện tích rừng ổn định tập trung chủ yếu tại các khu vực Vườn Quốc gia và rừng phòng hộ. Ngược lại, các điểm nóng mất rừng phân bố phân tán, thường gắn liền với các khu vực tiếp giáp với hoạt động nuôi trồng thủy sản và ven biển.



Hình 3.11: Bản đồ phân loại biến động rừng khu vực Vườn Quốc gia Mũi Cà Mau

Tại khu vực trọng điểm Vườn Quốc gia Mũi Cà Mau, kết quả phân loại thể hiện chi tiết các biến động quy mô nhỏ. Mô hình nhận diện chính xác các khu vực mất rừng dọc theo hệ thống kênh mương và bờ bao, cũng như ghi nhận sự phục hồi tự nhiên của thảm thực vật tại các bãi bồi ven biển.

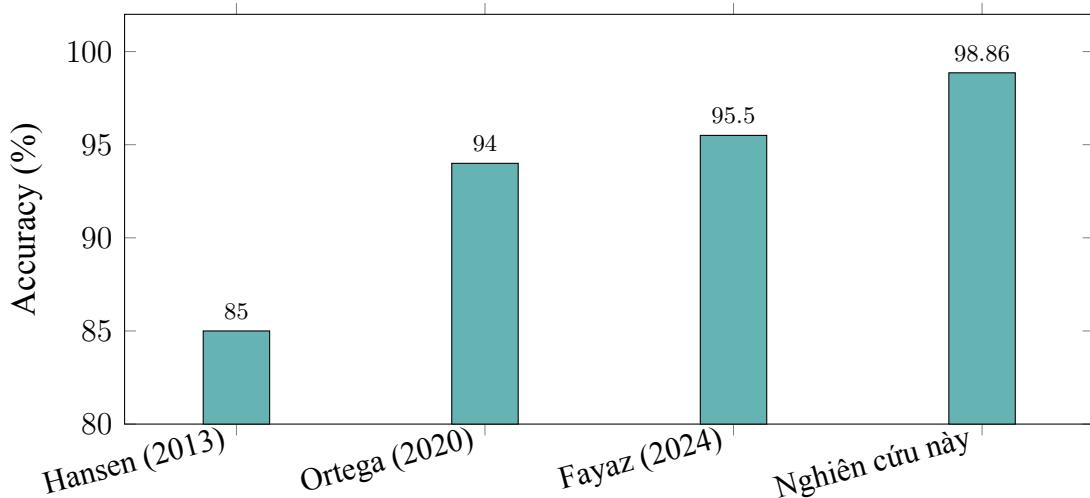
Kết quả này cho thấy một xu hướng biến động phức tạp: mặc dù có sự mất mát rừng do các nguyên nhân tự nhiên và nhân sinh (chuyển đổi mục đích sử dụng), nhưng cũng có những tín hiệu tích cực từ việc phục hồi rừng.



Hình 3.12: Tỷ lệ diện tích các lớp phân loại

Thống kê diện tích cho thấy xu hướng suy giảm rõ ràng của tài nguyên rừng trong giai đoạn nghiên cứu, với tỷ lệ mất rừng (4,48%) cao hơn so với tỷ lệ phục hồi (3,04%). Sự chênh lệch này đặt ra yêu cầu cấp thiết về các biện pháp quản lý và bảo vệ rừng chặt chẽ hơn.

Bản đồ biến động độ phân giải 10m cung cấp cái nhìn chi tiết chưa từng có, cho phép các nhà quản lý xác định chính xác vị trí các điểm nóng về phá rừng cũng như đánh giá hiệu quả của các dự án trồng rừng. So sánh với các nghiên cứu khác cho thấy phương pháp đề xuất có độ chính xác cạnh tranh cao.



Hình 3.13: So sánh Accuracy với các nghiên cứu trước đó

So sánh tương quan với các nghiên cứu trước đây cho thấy phương pháp đề xuất đạt độ chính xác vượt trội (98,86%). Sự cải thiện này chủ yếu đến từ việc khai thác dữ liệu ảnh vệ tinh độ phân giải cao (10m) kết hợp với kiến trúc học sâu được tối ưu hóa

cho điều kiện địa phương, khắc phục được các hạn chế của các sản phẩm toàn cầu có độ phân giải thấp.

3.3. Các sản phẩm hoàn thành

Trong khuôn khổ đợt thực tập, dựa trên các nỗ lực nghiên cứu và thực nghiệm đã trình bày, ba sản phẩm chính đã được hoàn thành với chất lượng cao, đáp ứng trực tiếp nhu cầu giám sát rừng của đơn vị và địa phương.

Sản phẩm thứ nhất là **Bản đồ phân loại biến động rừng tỉnh Cà Mau giai đoạn 2024–2025**. Đây là kết quả của việc áp dụng mô hình CNN đã huấn luyện trên diện rộng, bao phủ toàn bộ vùng quy hoạch lâm nghiệp tỉnh Cà Mau với diện tích 162.469 ha (chiếm 95,5% diện tích ranh giới). Bản đồ có độ phân giải không gian 10m, chi tiết hơn nhiều so với các sản phẩm toàn cầu như Hansen (30m). Kết quả định lượng cho thấy bản đồ đã phát hiện 7.282 ha mất rừng (4,48%) và 4.941 ha phục hồi rừng (3,04%) trong giai đoạn nghiên cứu. Độ chính xác toàn cục của bản đồ đạt 98,86%, khẳng định độ tin cậy cao của sản phẩm. Các tập tin bản đồ được xuất dưới định dạng GeoTIFF chuẩn hóa, tương thích hoàn toàn với hệ thống GIS của Chi cục Kiểm lâm.

Sản phẩm thứ hai là **Ứng dụng Web GIS giám sát biến động rừng trực tuyến**. Em đã xây dựng thành công ứng dụng trên nền tảng Google Earth Engine Apps, cho phép truy cập công khai tại địa chỉ: <https://ee-bonglantrungmuoi.projects.earthengine.app/view/giam-sat-bien-dong-rung-ca-mau>. Ứng dụng cung cấp các công cụ trực quan hóa mạnh mẽ, cho phép người dùng xem lớp phủ rừng trên nền ảnh vệ tinh mới nhất, tra cứu thông tin chi tiết tại từng vị trí, và theo dõi diễn biến rừng mà không cần cài đặt phần mềm chuyên dụng. Đây là công cụ hỗ trợ đặc lực cho công tác báo cáo và ra quyết định nhanh chóng.

Sản phẩm thứ ba là **Bộ quy trình công nghệ và mã nguồn mở**. Toàn bộ quy trình từ thu thập, tiền xử lý dữ liệu đa nguồn (Sentinel-1/2), trích xuất đặc trưng, đến huấn luyện mô hình và dự đoán đã được đóng gói thành các module mã nguồn Python/Jupyter Notebook khoa học. Bộ mã nguồn này đã được công bố trên GitHub (<https://github.com/ninhhaidang>), đảm bảo tính minh bạch và khả năng tái lập của nghiên cứu. Đi kèm với đó là bộ dữ liệu mẫu gồm 2.630 điểm thực địa đã dán nhãn, phục vụ cho việc nghiên cứu và phát triển tiếp theo. Quy trình này có tính chuyên giao cao, sẵn sàng để đơn vị thực tập tiếp nhận và vận hành.

3.4. Các kiến nghị đề xuất

Dựa trên kết quả thực nghiệm và những hạn chế đã phân tích, em đề xuất bốn hướng phát triển chính nhằm hoàn thiện và nâng cao hiệu quả của hệ thống trong tương lai:

Thứ nhất, về phương pháp luận, cần mở rộng nghiên cứu sang phân tích chuỗi thời gian (time-series analysis) thay vì chỉ so sánh hai thời điểm. Việc áp dụng các mô hình tiên tiến như Transformer hoặc LSTM trên chuỗi ảnh vệ tinh liên tục sẽ giúp khai thác hiệu quả các đặc trưng thời gian, từ đó nắm bắt tốt hơn quy luật diễn thê của rừng và phân biệt chính xác giữa đất rừng thực sự với các thay đổi theo mùa.

Thứ hai, để cải thiện hiệu suất mô hình, cần thử nghiệm cơ chế học chuyển giao (Transfer Learning) từ các mô hình đã được huấn luyện trên bộ dữ liệu lớn, đồng thời áp dụng các phương pháp kết hợp (Ensemble Learning). Việc kết hợp kết quả dự đoán từ nhiều mô hình khác nhau có thể giúp giảm thiểu phuơng sai, tăng độ ổn định và độ chính xác tổng thể của hệ thống.

Thứ ba, về phạm vi ứng dụng thực tiễn, kiến nghị mở rộng việc áp dụng mô hình sang các tỉnh khác trong vùng Đồng bằng sông Cửu Long có điều kiện sinh thái tương tự. Đồng thời, cần đẩy mạnh việc tích hợp kết quả phân loại vào các hệ thống thông tin địa lý (GIS) hiện có của cơ quan quản lý nhà nước, tạo nên quy trình giám sát liền mạch và tự động hóa.

Cuối cùng, cần tăng cường công tác khảo sát thực địa để kiểm chứng kết quả và mở rộng bộ dữ liệu mẫu (Ground Truth). Việc thu thập thêm dữ liệu đa thời gian và thiết lập mạng lưới quan trắc cố định sẽ giúp nâng cao độ tin cậy của mô hình, đồng thời tạo cơ sở dữ liệu huấn luyện phong phú hơn cho các nghiên cứu tiếp theo.

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, trang 8.
- [2] Công ty TNHH Tư vấn và Phát triển Đồng Xanh, *Ứng dụng công nghệ viễn thám, ảnh vệ tinh để quản lý, giám sát tài nguyên rừng và sạt lở ven biển trên địa bàn tỉnh Cà Mau*, Cà Mau, 2025, Tài liệu nội bộ, Sở Khoa học và Công nghệ tỉnh Cà Mau.
- [3] Sở Nông nghiệp và Phát triển Nông thôn tỉnh Cà Mau, *Báo cáo hiện trạng rừng tỉnh Cà Mau năm 2021*, Cà Mau, 2021, Báo cáo nội bộ.

Tiếng Anh

- [4] Breiman, L., “Random forests”, *Machine Learning*, vol. 45, no. 1, 2001, pp. 5–32.
- [5] Cortes, C., Vapnik, V., “Support-vector networks”, *Machine Learning*, vol. 20, no. 3, 1995, pp. 273–297.
- [6] 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.
- [7] European Space Agency, *Sentinel-1 SAR User Guide*, 2024, Available: https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-1.
- [8] European Space Agency, *Sentinel-2 MSI User Guide*, 2024, Available: https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2.
- [9] FAO, *Global Forest Resources Assessment 2020: Main Report*, Food and Agriculture Organization of the United Nations, Rome, 2020.

- [10] 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.
- [11] Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., Townshend, J. R. G., “High-resolution global maps of 21st-century forest cover change”, *Science*, vol. 342, no. 6160, 2013, pp. 850–853.
- [12] Hu, Y., Raza, A., Syed, 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”, *Journal of the Indian Society of Remote Sensing*, vol. 48, 2020, pp. 1055–1064.
- [13] 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.
- [14] Ioffe, S., Szegedy, C., “Batch normalization: Accelerating deep network training by reducing internal covariate shift”, *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, pp. 448–456, PMLR, 2015.
- [15] Key, C. H., Benson, N. C., “Landscape assessment (LA): Sampling and analysis methods”, *USDA Forest Service General Technical Report RMRS-GTR-164-CD*, 2006, pp. 1–51.
- [16] Kuenzer, C., Bluemel, A., Gebhardt, S., Quoc, T. V., Dech, S., “Remote sensing of mangrove ecosystems: A review”, *Remote Sensing*, vol. 3, no. 5, 2011, pp. 878–928.
- [17] 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.
- [18] LeCun, Y., Bengio, Y., Hinton, G., “Deep learning”, *Nature*, vol. 521, no. 7553, 2015, pp. 436–444.

- [19] Loshchilov, I., Hutter, F., “Decoupled weight decay regularization”, *Proceedings of the 7th International Conference on Learning Representations (ICLR)*, 2019.
- [20] 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.
- [21] 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.
- [22] Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Flouri, N., Brown, M., Traver, I. N., Deghaye, P., Duesmann, B., Rosich, B., Miranda, N., Bruno, C., L’Abbate, M., Croci, R., Pietropaolo, A., Huchler, M., Rostan, F., “GMES Sentinel-1 mission”, *Remote Sensing of Environment*, vol. 120, 2012, pp. 9–24.
- [23] Vo, Q. M., Oppelt, N., Kuenzer, C., “Gap-free monitoring of annual mangrove forest dynamics in Ca Mau Province, Vietnamese Mekong Delta, using the Landsat-7-8 archives and post-classification temporal optimization”, *Remote Sensing*, vol. 12, no. 22, 2020, pp. 3729, Article 3729.
- [24] 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.
- [25] 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.