

**Kết hợp trí tuệ nhân tạo (AI) và Google Earth Engine (GEE)
để phân loại các lớp phủ từ ảnh Sentinel-2:
trường hợp nghiên cứu tại xã Quảng Sơn và xã Tà Đùng, tỉnh Lâm Đồng**

**Nguyễn Trọng Cường¹, Trần Quang Bảo², Trương Tất Đo^{2*},
Nguyễn Thị Hà³, Phạm Văn Duẩn¹, Nguyễn Hải Hoà¹**

¹Trường Đại học Lâm nghiệp

²Cục Lâm nghiệp và Kiểm lâm

³Trường Đại học Lâm nghiệp - Phân hiệu Đồng Nai

**Integrating artificial intelligence (AI) and Google Earth Engine (GEE)
for landcover classification using Sentinel-2 imagery:
a case study in Quang Son and Ta Dung communes, Lam Dong province**

**Nguyen Trong Cuong¹, Tran Quang Bao², Truong Tat Do^{2*},
Nguyen Thi Ha³, Pham Van Duan¹, Nguyen Hai Hoa¹**

¹Vietnam National University of Forestry

²Vietnam Administration of Forestry

³Vietnam National University of Forestry - Dong Nai Campus

*Corresponding author: truongtatdo@gmail.com

<https://doi.org/10.55250/jo.vnuf.14.7.2025.060-070>

Thông tin chung:

Ngày nhận bài: 07/10/2025

Ngày phản biện: 10/11/2025

Ngày quyết định đăng: 02/12/2025

Từ khóa:

EUDR, GEE, không gây mất rừng,
phân loại lớp phủ,
trí tuệ nhân tạo.

Keywords:

EUDR, GEE, deforestation,
landcover classification, artificial
Intelligence,

TÓM TẮT

Nghiên cứu này đã tích hợp giữa mô hình học sâu Vision Transformer (ViT) và Google Earth Engine (GEE) để phân loại lớp phủ đất và thành lập bản đồ các khu vực canh tác cà phê và các khu vực có rừng từ ảnh Sentinel-2 năm 2020 tại xã Quảng Sơn và xã Tà Đùng, tỉnh Lâm Đồng. Tổng cộng 10.394 điểm mẫu được sử dụng, trong đó 8.352 điểm huấn luyện và 2.042 điểm kiểm tra. Kết quả đào tạo mô hình AI cho thấy độ chính xác tổng thể đạt 91%, điểm F1 trung bình có trọng số 0,91, điểm F1 trung bình không có trọng số đạt 0,92. Kết quả đánh giá độ chính xác bằng dữ liệu kiểm tra cho độ tin cậy cao với độ chính xác tổng thể đạt 88%, độ tin cậy của kết quả kiểm tra đối với cà phê đạt 83%, rừng đạt 96%, mặt nước 99%, đất nông nghiệp và cây trồng khác đạt 71%, đất xây dựng đạt 95%. Kết quả phân loại được kiểm chứng ngoài thực địa ở 46 điểm năm 2024 đạt 91,4%. Kết quả tính toán diện tích lớp phủ khu vực nghiên cứu cho thấy có khoảng 55.264 ha rừng (chiếm hơn 56% tổng diện tích của hai xã), hơn 26.000 ha cà phê (chiếm khoảng 26,5% tổng diện tích của hai xã). Kết quả nghiên cứu đã khẳng định hiệu quả của việc kết hợp thuật toán AI và GEE trong nhận diện các vùng canh tác cà phê và rừng điển hình ở vùng Tây Nguyên, đồng thời cung cấp cơ sở khoa học cho giám sát và chứng minh vùng trồng cà phê không gây mất rừng nhằm đáp ứng Quy định chống mất rừng của Liên minh châu Âu (EUDR).

ABSTRACT

This study integrated the Vision Transformer (ViT) deep learning model with Google Earth Engine (GEE) to classify land-cover types and generate maps of coffee plantation areas and forested regions using 2020 Sentinel-2 imagery for Quang Son and Ta Dung communes, Lam Dong province. A total of 10,394 samples were used, including 8,352 training samples and 2,042 testing samples. The AI model achieved an overall accuracy of 91%, with a weighted average F1-score of 0.91 and a macro average F1-score of 0.92. Accuracy assessment using the test dataset demonstrated high reliability, with an overall accuracy of 88%; class-specific accuracies included 83% for coffee, 96% for forest, 99% for water, 71% for agricultural and other crop lands, and

95% for built-up areas. Field verification conducted at 46 sites in 2024 confirmed an accuracy of 91.4%. The land-cover area estimates indicate approximately 55,264 ha of forest (over 56% of the two communes' natural area) and more than 26,000 ha of coffee plantations (about 26.5%). These results confirm the effectiveness of integrating AI algorithms with GEE in identifying typical forest and coffee cultivation patterns in the Central Highlands, while providing a scientific basis for monitoring deforestation-free coffee production in compliance with the EUDR.

1. ĐẶT VẤN ĐỀ

Suy thoái tài nguyên đất và rừng là một trong những thách thức lớn nhất đối với phát triển bền vững trong thế kỷ 21, ảnh hưởng trực tiếp đến an ninh lương thực, bảo tồn đa dạng sinh học và khả năng thích ứng với biến đổi khí hậu [1]. Việc giám sát chính xác sự thay đổi lớp phủ là nền tảng để đưa ra các kế hoạch quản lý hiệu quả. Tại nhiều quốc gia, sự chuyển đổi đất rừng sang canh tác cây công nghiệp như dầu cọ, cao su, cà phê đã đặt ra áp lực ngày càng tăng từ phía cộng đồng quốc tế về minh bạch chuỗi cung ứng [2]. Đặc biệt là Quy định về việc sản xuất và thương mại các sản phẩm hàng hoá không mất rừng và suy thoái rừng của Liên minh châu Âu (EUDR), theo đó EUDR quy định các hàng hóa nhập khẩu và xuất khẩu vào EU (trong đó bao gồm cà phê) phải chứng minh không liên quan đến mất rừng và suy thoái rừng sau ngày 31/12/2020 [3]. Điều này đặt ra yêu cầu cấp thiết về các phương pháp phân loại và giám sát lớp phủ có độ chính xác cao, ứng dụng được trên phạm vi rộng và đủ thông tin và bằng chứng khoa học để chứng minh các sản phẩm hàng hoá không gây mất rừng, đáp ứng yêu cầu trong thương mại quốc tế trước bối cảnh mới.

Trong nhiều năm trở lại đây, Google Earth Engine (GEE) trở thành nền tảng hữu hiệu trong thác và xử lý dữ liệu viễn thám nhờ khả năng truy cập kho dữ liệu vệ tinh khổng lồ, tích hợp hạ tầng tính toán đám mây và cung cấp công cụ lập trình mạnh mẽ [4]. Các nghiên cứu trên thế giới đã khai thác tối đa GEE để xây dựng bản đồ lớp phủ quy mô toàn cầu [5], giám sát biến động rừng [6] hay phát hiện sớm cháy rừng [7, 8]. Tuy nhiên, GEE chủ yếu cung cấp các thuật toán học máy truyền thống (Random Forest, SVM) mà chưa tích hợp đầy đủ các mô hình học sâu tiên tiến.

Sự bùng nổ của trí tuệ nhân tạo (AI) và học sâu (Deep learning) đã tạo ra những đột phá lớn trong phân tích ảnh viễn thám [9]. Mặc dù ở Việt Nam đã có nhiều nghiên cứu ứng dụng GEE hoặc AI trong giám sát lớp phủ, tuy nhiên sự kết hợp trực tiếp giữa GEE và các mô hình học sâu vẫn chưa được khai thác đầy đủ, đa số công trình nghiên cứu tập trung vào các mô hình học máy có sẵn trên GEE hoặc sử dụng GEE để tiền xử lý dữ liệu, sau đó xuất dữ liệu ra ngoài để huấn luyện mô hình AI, việc này gây gián đoạn quá trình và khó mở rộng quy mô do phải chuyển dữ liệu qua thiết bị thứ ba mà không thể xuất dữ liệu từ GEE sang nền tảng AI tự động và liên tục [9]. Bên cạnh các thuật toán học sâu phổ biến như CNN, FCNN, mô hình Transformer cũng đã được ứng dụng trong phân loại rừng, giám sát cây trồng, nhận dạng đô thị với độ chính xác vượt trội so với các phương pháp truyền thống [10]. Hiện nay, ứng dụng Transformer trong viễn thám còn khá mới mẻ, đặc biệt tại các quốc gia đang phát triển nhưng tiềm năng rất lớn để giải quyết các bài toán phân loại phức tạp như nhận diện cây công nghiệp trong môi trường nhiệt đới [9]. Các nghiên cứu cho thấy Transformer có khả năng học đặc trưng không gian và phổ hiệu quả, xử lý ảnh đa nguồn và cải thiện đáng kể kết quả phân loại lớp phủ đất [11]. Ngoài ra, hầu hết các nghiên cứu tập trung vào phân loại rừng và đô thị, trong khi các cây công nghiệp dài ngày như cà phê lại ít được nghiên cứu, nhất là trong bối cảnh gắn với các yêu cầu minh bạch chuỗi cung ứng sản phẩm nông lâm sản.

Để giải quyết các khoảng trống trên, nghiên cứu này triển khai một cách tiếp cận mới theo hướng kết hợp GEE để tiền xử lý và trích xuất ảnh Sentinel-2 đa thời gian và toàn bộ dữ liệu gồm các điểm mẫu nghiên cứu, giá trị phổ xạ

của các kênh ảnh Sentinel 2 và giá trị các chỉ số thực vật của toàn bộ các điểm mẫu trực tiếp vào Google Driver và đưa trực tiếp các dữ liệu vào mô hình Vision Transformer huấn luyện thông qua nền tảng Google Colab. Kết hợp Google Colab với GEE tạo ra một lợi thế mà các nền tảng Python hay Jupyter Notebook thông thường khó đạt được bởi Google Colab và GEE cho phép triển khai hạ tầng đám mây kép theo hướng GEE xử lý dữ liệu viễn thám quy mô hành tinh trực tiếp trên máy chủ của Google, trong khi Colab cung cấp GPU/TPU mạnh để huấn luyện các mô hình học sâu. Ngoài ra, việc kết nối trực tiếp giữa GEE và Colab giúp loại bỏ hoàn toàn nhu cầu tải dữ liệu thủ công, giảm rủi ro lỗi và tăng tốc đáng kể quá trình nghiên cứu. Đồng thời, toàn bộ quá trình từ tiền xử lý, mô hình hóa, đánh giá đến trực quan hóa đều chạy hoàn toàn trên đám mây, không phụ thuộc cấu hình máy tính cá nhân và không yêu cầu phần mềm bản quyền. Do đó, kết hợp GEE và Colab trở thành lựa chọn tối ưu cho các nghiên cứu viễn thám hiện đại bởi ưu điểm nhanh, ổn định, linh hoạt, chi phí thấp và đặc biệt phù hợp cho các mô hình AI quy mô lớn mà Jupyter Notebook

cực bộ khó vận hành hiệu quả [11].

Nghiên cứu tiến hành phân loại một số trạng thái lớp phủ chủ yếu tại xã Quảng Sơn và xã Tà Đùng, tỉnh Lâm Đồng gồm cà phê, rừng, mặt nước, đất xây dựng, đất nông nghiệp. Khu vực này vốn có thế mạnh về sản xuất cà phê xuất khẩu, đồng thời cũng có tốc độ chuyển đổi cây trồng khác sang cà phê lớn, thực tế này đặt ra áp lực rất lớn cho công tác giám sát, chứng minh các khu vực trồng cà phê không gây mất rừng đáp ứng yêu cầu của EUDR.

2. PHƯƠNG PHÁP NGHIÊN CỨU

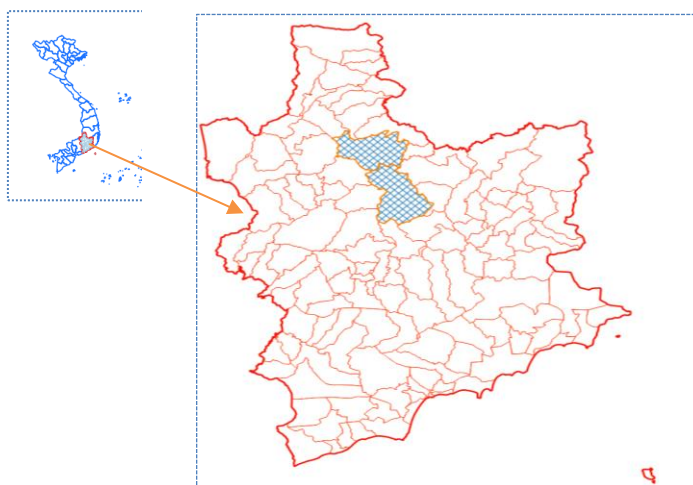
2.1. Khu vực nghiên cứu

Nghiên cứu được thực hiện tại xã Quảng Sơn và xã Tà Đùng, tỉnh Lâm Đồng một trong những địa bàn trọng điểm về phát triển cây công nghiệp, đặc biệt là cà phê (Hình 1).

2.2. Dữ liệu nghiên cứu

2.2.1. Tư liệu ảnh vệ tinh Sentinel-2 MSI

Nghiên cứu sử dụng ảnh vệ tinh Sentinel-2 (mức xử lý 2A, độ phân giải không gian 10 m) được truy cập miễn phí trên Google Earth Engine. Để giảm thiểu ảnh hưởng của mây và nhiễu khí quyển, ảnh được tiền xử lý bằng phương pháp Cloud masking.



Hình 1. Sơ đồ khu vực nghiên cứu

Để thu được ảnh có tỉ lệ mây thấp, đảm bảo kết quả phân loại tốt các trạng thái lớp phủ, nghiên cứu đã thu thập toàn bộ ảnh Sentinel-2 có tỉ lệ mây nhỏ hơn 20% từ tháng 11 - 12/2020 đây là thời điểm bắt đầu mùa khô ở Tây

Nguyên, việc sử dụng ảnh vào cuối năm 2020 cũng đáp ứng yêu cầu của EUDR về sản phẩm hàng hoá không gây mất rừng kể từ sau ngày 31/12/2020. Tổng số 06 cảnh ảnh Sentinel-2 đảm bảo điều kiện đã được thu thập (Bảng 1).

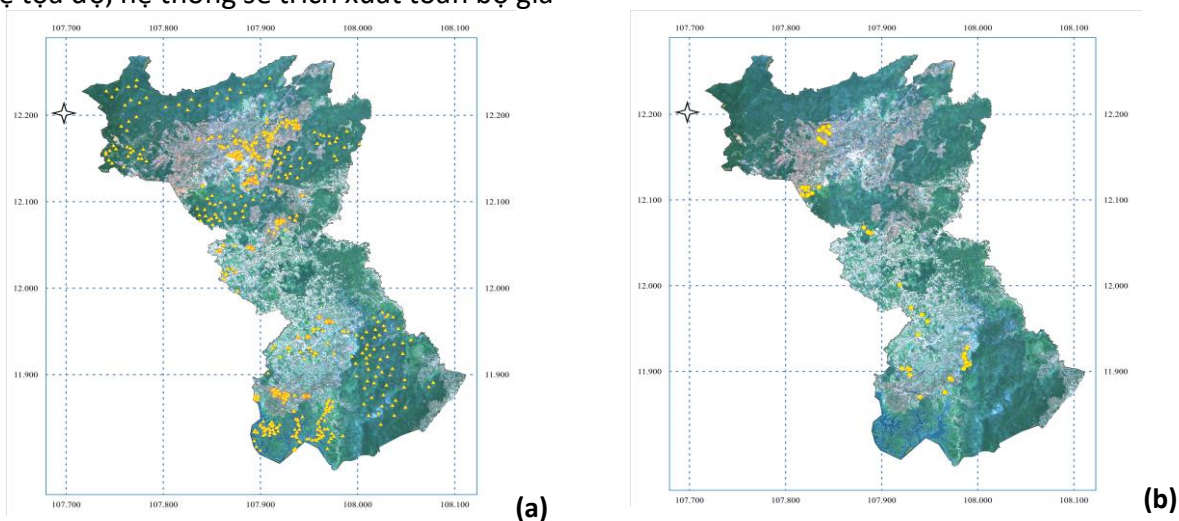
Bảng 1. Danh sách cảnh ảnh Sentinel-2 được sử dụng trong nghiên cứu

TT	Mã cảnh ảnh	Ngày chụp	Tỉ lệ mây (%)
1	20201113T032312_T48PYU	13/11/2020	7,4
2	20201118T032219_T48PYU	18/11/2020	2,9
3	20201123T031414_T48PYU	23/11/2020	22,0
4	20201203T032317_T48PYU	03/12/2020	12,2
5	20201213T031451_T48PYU	13/12/2020	18,4
6	20201228T032330_T48PYU	28/12/2020	13,2

2.2.2. Dữ liệu điều tra thực địa và tạo các vùng quan tâm

Từ các cuộc điều tra thực địa năm 2020, kết hợp với các số liệu từ các nghiên cứu trước đó, tổng số 1.140 điểm mẫu được lấy trên ảnh Google Earth độ phân giải cao vào thời điểm tháng 12/2020 đã được thu thập một cách kỹ lưỡng nằm ở trung tâm các khu vực có bán kính xung quanh tối thiểu 30 m để khi mở rộng, đảm bảo tính đại diện và nằm ở trung tâm các điểm trạng thái thực địa, gồm 315 điểm rừng, 211 điểm mặt nước, 236 điểm cà phê, 250 điểm đất nông nghiệp (gồm đất trồng cây nông nghiệp hàng năm, đất trồng cây công nghiệp khác ngoài cà phê và đất đồng cỏ), 128 điểm đất xây dựng (gồm khu vực dân cư và đất xây dựng). Trong toàn bộ các điểm mẫu đã được thu thập, sử dụng 80% số điểm mẫu làm mẫu huấn luyện và 20% làm mẫu kiểm tra. Các điểm mẫu này sau đó tiếp tục được mở rộng ra 20 m bằng hàm mở rộng (buffer) trong GEE, việc này sẽ tạo ra các điểm mẫu trong vòng bán kính 20 m xung quanh điểm gốc với cùng giá trị phản xạ và hệ tọa độ, hệ thống sẽ trích xuất toàn bộ giá

trị phổ của các pixel nằm trong vùng đệm được tạo, sau đó tính trung bình đặc trưng cho vùng mẫu thay vì chỉ lấy giá trị tại một pixel đơn lẻ, điều này đảm bảo dung lượng mẫu để mô hình AI huấn luyện và kiểm tra chính xác hơn. Sau khi mở rộng, có tổng số 10.394 điểm mẫu mở rộng đã được tạo tự động, trong đó 8.352 điểm được sử dụng làm mẫu huấn luyện cho quá trình phân loại và 2.042 điểm được sử dụng làm mẫu kiểm tra. Toàn bộ dữ liệu này được đưa sang mô hình Vision Transformer tích hợp trên Google Colab để huấn luyện và phân loại. Trong chuyến thu thập dữ liệu ngoại nghiệp vào năm 2024, có 46 điểm kiểm chứng đã được thu thập để phục vụ cho việc kiểm chứng kết quả phân loại ngoài hiện trường gồm 27 điểm cà phê, 3 điểm rừng, 16 điểm cây nông nghiệp và cây trồng khác. Các điểm này được xác định là không có sự thay đổi trạng thái từ năm 2020 đến 2024 thông qua khảo sát chủ rừng và các chủ đất tại khu vực nghiên cứu. Vị trí các điểm mẫu và điểm kiểm chứng thực địa sau phân loại thể hiện trong Hình 2.



Hình 2. Vị trí các điểm mẫu khu vực nghiên cứu trên ảnh vệ tinh Sentinel-2 Quý IV/2020 (a) và điểm kiểm chứng thực địa thu thập năm 2024 (b)

2.2.3. Tiền xử lý và trích xuất đặc trưng ảnh

Trên nền tảng GEE, tiến hành trích xuất giá trị các chỉ số đặc trưng phổ trong phân loại lớp

phủ được tính toán từ các băng tần của ảnh Sentinel 2, các chỉ số và công thức toán học chi tiết thể hiện trong Bảng 2.

Bảng 2. Các chỉ số thực vật đặc trưng sử dụng cho phân loại

TT	Chỉ số	Công thức toán học	Nguồn
1	NDVI (Normalized Difference Vegetation Index)	$NDVI = \frac{(B8 - B4)}{(B8 + B4)}$	[12]
2	NDMI (Normalized Difference Moisture Index)	$NDMI = \frac{(B8 - B11)}{(B8 + B11)}$	[13]
3	NDBI (Normalized Difference Built-up Index)	$NDBI = \frac{(B11 - B8)}{(B11 + B8)}$	[14]
4	NDBaI (Normalized Difference Bareness Index)	$NDBaI = \frac{(B11 - B4)}{(B11 + B4)}$	[15]
5	NDWI (Normalized Difference Water Index)	$NDWI = \frac{(B3 - B8)}{(B3 + B8)}$	[16]
6	SAVI (Soil-Adjusted Vegetation Index)	$SAVI = \frac{(B8 - B4)}{(B8 + B4 + L)} \times (1 + L)$	[17]
7	EVI (Enhanced Vegetation Index)	$EVI = 2.5 \times \frac{B8 - B4}{B8 + 6 \times B4 - 7.5 \times B2 + 1}$	[18]

Trong đó: B3, B4, B8, B11 lần lượt là các kênh GREEN, RED, NIR, SWIR1 trên ảnh Sentinel-2, trong nghiên cứu này giá trị L được lấy bằng 0,5.

Giá trị các chỉ số này được tính trung bình cho toàn bộ các cảnh ảnh nghiên cứu. Bên cạnh đó, các giá trị phản xạ phổ trung bình của toàn bộ 11 băng phổ gốc làm đặc trưng đầu vào cho mô hình phân loại trong mô hình trí tuệ nhân tạo. Như vậy, bộ mẫu huấn luyện và kiểm tra gồm 10.394 điểm, mỗi điểm có 11 giá trị của 11 băng phổ ảnh Sentinel-2 và 7 giá trị của 7 chỉ số thực vật tại bảng 2 đã được tính toán trên nền tảng GEE phục vụ cho việc phân loại bằng AI.

2.2.4. Mô hình Vision Transformer và huấn luyện mô hình

Mô hình Vision Transformer (ViT) là một hướng tiếp cận tiên tiến trong lĩnh vực học sâu, được giới thiệu bởi Dosovitskiy vào năm 2021 [11]. Khác với các mạng tích chập truyền thống, ViT chia ảnh đầu vào thành các ô nhỏ (patch) có kích thước cố định, sau đó mã hóa (encode) từng ô thành một vector đặc trưng (feature vector) trước khi đưa qua các khối mã hóa Transformer, bao gồm hai cơ chế cốt lõi là cơ chế tự chú ý (self-attention) và mạng truyền thẳng (feed-forward network). Cơ chế tự chú ý cho phép mô hình nhận biết các mối quan hệ toàn cục giữa các vùng trong ảnh, giúp ViT vượt

trội trong việc nhận dạng những đối tượng có đặc trưng phổ tương đồng hoặc phân bố không gian phức tạp, chẳng hạn như phân biệt giữa rừng tự nhiên, cà phê, đất nông nghiệp [10, 19]. Ưu điểm nổi bật của ViT là khả năng hiểu và học ngữ cảnh toàn ảnh, từ đó cải thiện đáng kể độ chính xác phân loại lớp phủ đất trong các khu vực có phổ phản xạ tương tự nhau.

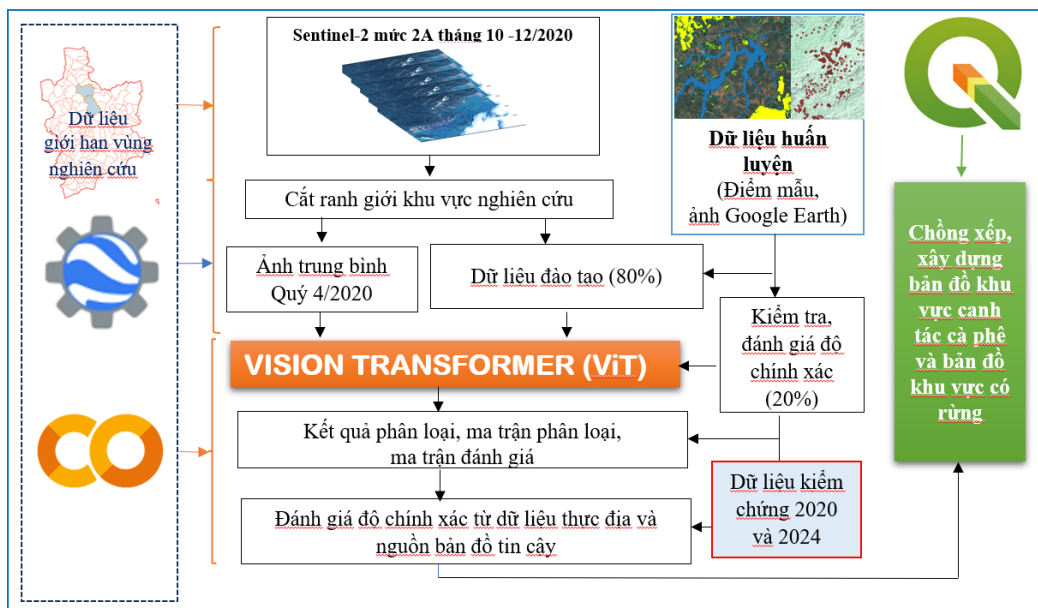
Bên cạnh đó, ViT không phụ thuộc vào tầng tích chập như CNN, giúp mô hình có thể mở rộng linh hoạt cho ảnh vệ tinh đa phổ và độ phân giải cao, điển hình là dữ liệu từ vệ tinh Sentinel-2. Trong nghiên cứu này, ViT được triển khai trực tiếp trên nền tảng Google Colab, tích hợp với Google Drive để truy cập và huấn luyện dữ liệu ảnh Sentinel-2 thông qua thư viện TensorFlow/Keras. Mô hình sử dụng lớp nhẵn ảnh 3x3 để trích xuất đặc trưng phổ cục bộ, lớp gộp cực đại để giảm chiều dữ liệu kết hợp với các lớp kết nối dày đặc (Dense, hàm kích hoạt Relu) nhằm học biểu diễn phi tuyến và lớp Softmax ở đầu ra để xác định xác suất phân loại cho từng lớp phủ. Quá trình huấn luyện được thực hiện với hàm mất mát và bộ tối ưu tiêu chuẩn, tốc độ học (learning rate) 0,001, kích

thước lô (batch size) 32 trong 200 vòng lặp (epochs = 200). Đồng thời cơ chế dừng sớm (early stopping) được tích hợp để ngăn hiện tượng quá khớp (overfitting) mô hình tự động dừng huấn luyện nếu sai số trên tập kiểm tra không được cải thiện sau 10 vòng lặp (patience = 10).

2.2.5. Đánh giá độ chính xác

Quá trình đánh giá độ chính xác phân loại

được thực hiện tự động thông qua ma trận nhầm lẫn và các chỉ số thống kê gồm: độ chính xác tổng thể (OA), độ chính xác của nhà sản xuất (PA), độ chính xác của người dùng (UA), điểm F1 (F1_score). Để kiểm định thống kê, phương pháp lấy mẫu lại Bootstrap được áp dụng nhằm đánh giá độ tin cậy của các chỉ số. Toàn bộ quy trình nghiên cứu được thể hiện tại Hình 3.



Hình 3. Sơ đồ quá trình nghiên cứu

3. KẾT QUẢ NGHIÊN CỨU

3.1. Kết quả phân loại tổng thể

3.1.1. Kết quả phân loại các trạng thái trên ảnh Sentinel 2

Từ các điểm mẫu huấn luyện đã được gán các giá phản xạ của 11 kênh ảnh Sentinel-2 và giá trị của 7 chỉ số thực vật được tính toán trên GEE, các điểm này được trích xuất tự động sang nền tảng Colab thông qua Google Driver để phân loại bằng Transformer. Kết quả sau 22 lần huấn luyện cho thấy độ chính xác tổng thể (OA) đạt 91% và hệ số F1 trung bình có trọng số đạt 0,91 phản ánh độ tin cậy cao của mô hình trong việc nhận dạng các lớp phủ chính tại khu vực nghiên cứu và khẳng định mô hình có khả năng học hiệu quả từ tập dữ liệu và tái tạo đặc trưng phổ của các lớp phủ trong khu vực nghiên cứu. Kết quả tính toán các chỉ số thống kê đối với kết quả phân loại thể hiện hiệu suất phân loại rất cao của mô hình khi áp dụng trên toàn bộ 8.352

mẫu huấn luyện với lớp rừng có độ tin cậy 0,96, điểm hệ số F1 đạt 0,96 trên 1.773 mẫu, lớp phủ cà phê đạt độ tin cậy 0,84, với F1 bằng 0,87 trên 2.198 mẫu, lớp mặt nước đạt mức phân loại với độ tin cậy 1,00 và hệ số F1 bằng 0,99, với 1.903 mẫu huấn luyện, đất nông nghiệp đạt độ tin cậy tương đối cao (83%), khu vực đất đô thị tuy có số mẫu nhỏ hơn (957 mẫu) nhưng cho kết quả khả quan với độ tin cậy 98% và F1 đạt 0,96.

3.1.2. Đánh giá độ chính xác sau phân loại bằng dữ liệu kiểm tra

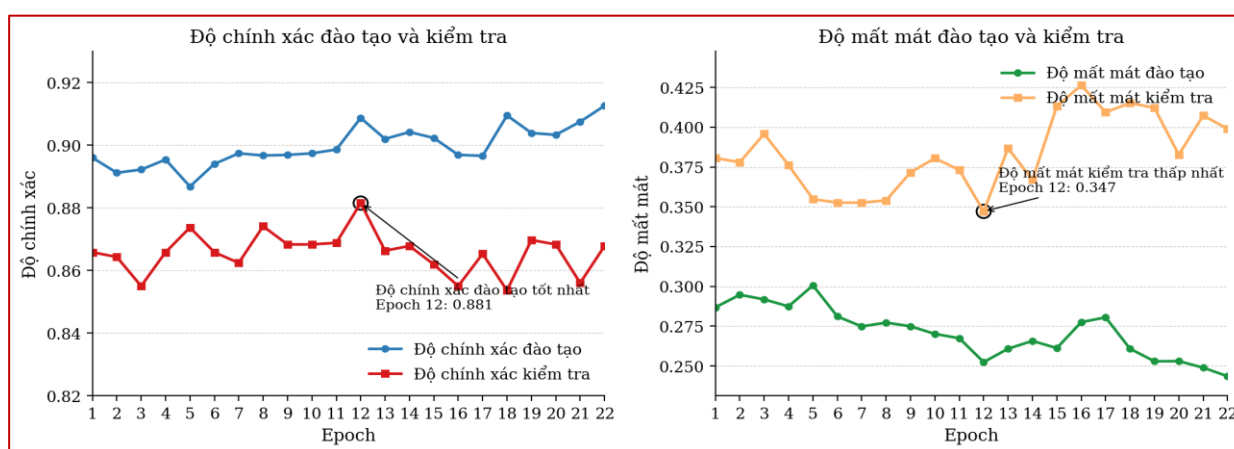
Bảng 3 thể hiện kết quả đánh giá độ chính xác sau phân loại thông qua 2.042 điểm kiểm tra của mô hình phân loại. Kết quả cho thấy hiệu suất cao trong nhận diện lớp phủ với độ chính xác tổng thể (OA) kiểm tra đạt 88% và F1 trung bình đạt 0,88. Ngoài lớp mặt nước với độ tin cậy của kết quả kiểm tra rất cao thì lớp rừng, cà phê và đất xây dựng đều có độ tin cậy và độ nhạy phân loại trên 80%.

Bảng 3. Kết quả đánh giá độ chính sau phân loại bằng dữ liệu kiểm tra

Chỉ số Lớp phân loại	Độ tin cậy	Độ nhạy phân loại	Điểm F1	Số mẫu kiểm chứng
Rừng	0,96	0,97	0,97	432
Mặt nước	0,99	0,97	0,98	396
Cà phê	0,83	0,82	0,83	561
Đất nông nghiệp	0,71	0,72	0,72	384
Đất xây dựng	0,95	0,95	0,95	269
Trung bình không trọng số	0,89	0,89	0,89	2.042
Trung bình có trọng số	0,88	0,88	0,88	
Độ chính xác tổng thể	0,88			

Toàn bộ quá trình huấn luyện và kiểm tra mô hình phân loại được thể hiện chi tiết tại biểu đồ

quá trình huấn luyện và kiểm tra (Hình 4).



Hình 4. Biểu đồ quá trình đào tạo và kiểm tra của mô hình

Kết quả qua 22 epoch cho thấy mô hình nhanh chóng đạt đến trạng thái hội tụ tương đối ổn định ngay từ những vòng lặp đầu tiên. Ở epoch 1, mô hình đạt độ chính xác huấn luyện 0,89 và độ chính xác kiểm tra 0,86 đi kèm với mất mát huấn luyện 0,28 và mất mát kiểm tra 0,38. Điều này thể hiện khả năng học ban đầu rất tốt, đồng thời cho thấy bộ đặc trưng đầu vào đã cung cấp thông tin đủ tách biệt cho mô hình. Từ epoch 2 - 10, độ chính xác dao động rất nhỏ, chỉ khoảng 0,01 cho thấy mô hình đã bước vào trạng thái ổn định, không xảy ra hiện tượng quá khớp. Mặc dù độ chính xác kiểm tra dao động giữa 0,85 – 0,87, mức sai lệch giữa đào tạo và kiểm tra luôn nhỏ hơn 3%, chứng tỏ mô hình duy trì được khả năng tổng quát hóa tốt. Tại epoch 12, mô hình đạt điểm đột phá với độ chính xác đào tạo 0,91 và độ chính xác kiểm

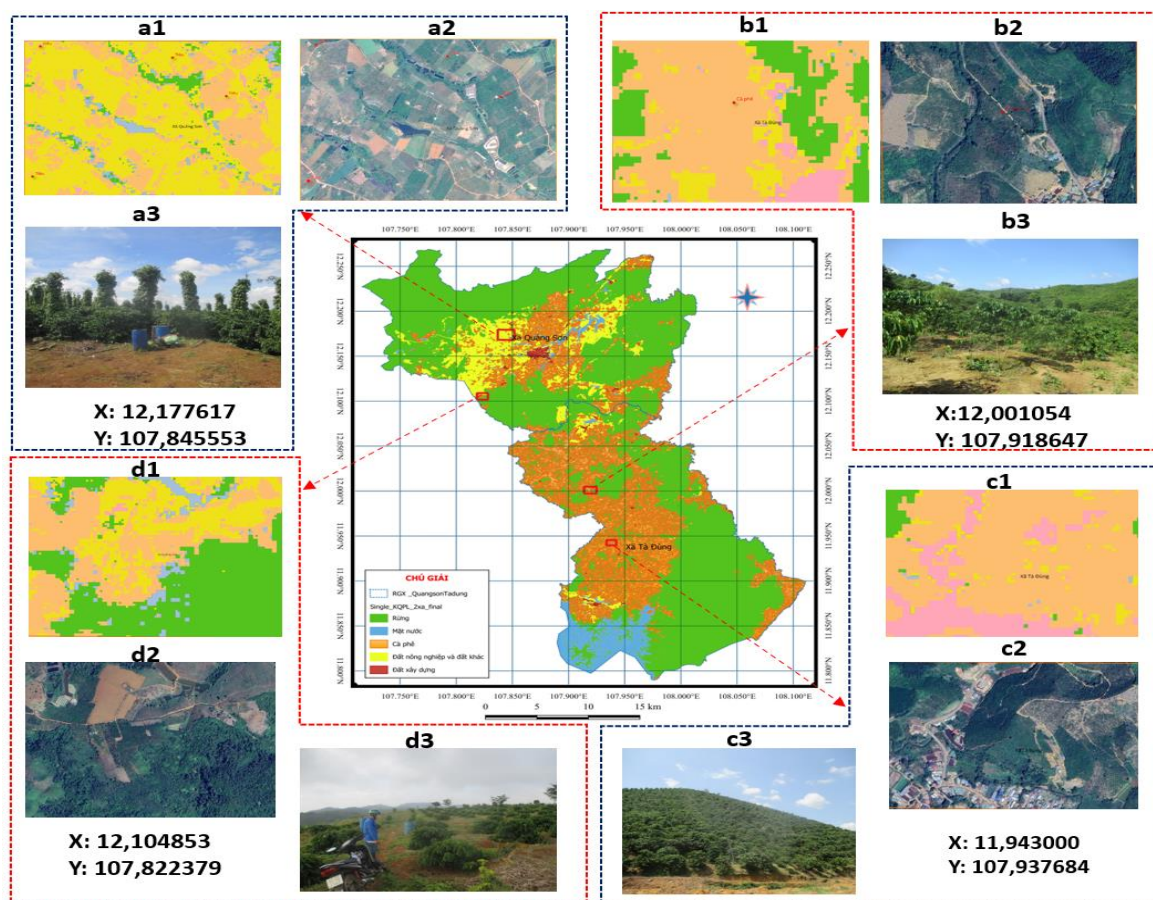
tra đạt 0,88, trong khi độ mất mát kiểm tra giảm xuống 0,347, thấp nhất trong toàn bộ quá trình huấn luyện và kiểm tra. Đây là thời điểm mô hình học được cấu trúc phổ không gian ổn định nhất từ dữ liệu. Từ epoch 13 trở đi, độ chính xác và độ mất mát đều tăng lên ở dữ liệu kiểm tra. Validation loss tăng đến khoảng 0,41. Trong các epoch từ 15 - 19, có sự giao động ở độ mất mát kiểm tra, tuy nhiên, độ chênh lệch với độ chính xác kiểm tra không lớn, đồng thời độ chính xác kiểm tra luôn duy trì trên 0,85 cho thấy mô hình không bị học nhanh. Đến epoch 22, mô hình đạt độ chính xác đào tạo tối đa bằng 0,91 và độ chính xác kiểm tra 0,87 khẳng định được kết quả đào tạo và kiểm tra đạt độ tin cậy cao.

3.2. Kết quả phân loại các trạng thái

Hình 5 trình bày kết quả kiểm chứng năm 2024 và minh họa 4 khu vực cụ thể đã được

phân loại, trong đó toàn bộ các trạng thái tại khu vực nghiên cứu đã được xác định và các

khu vực cụ thể được kiểm chứng trực quan giữa các trạng thái với nhau.

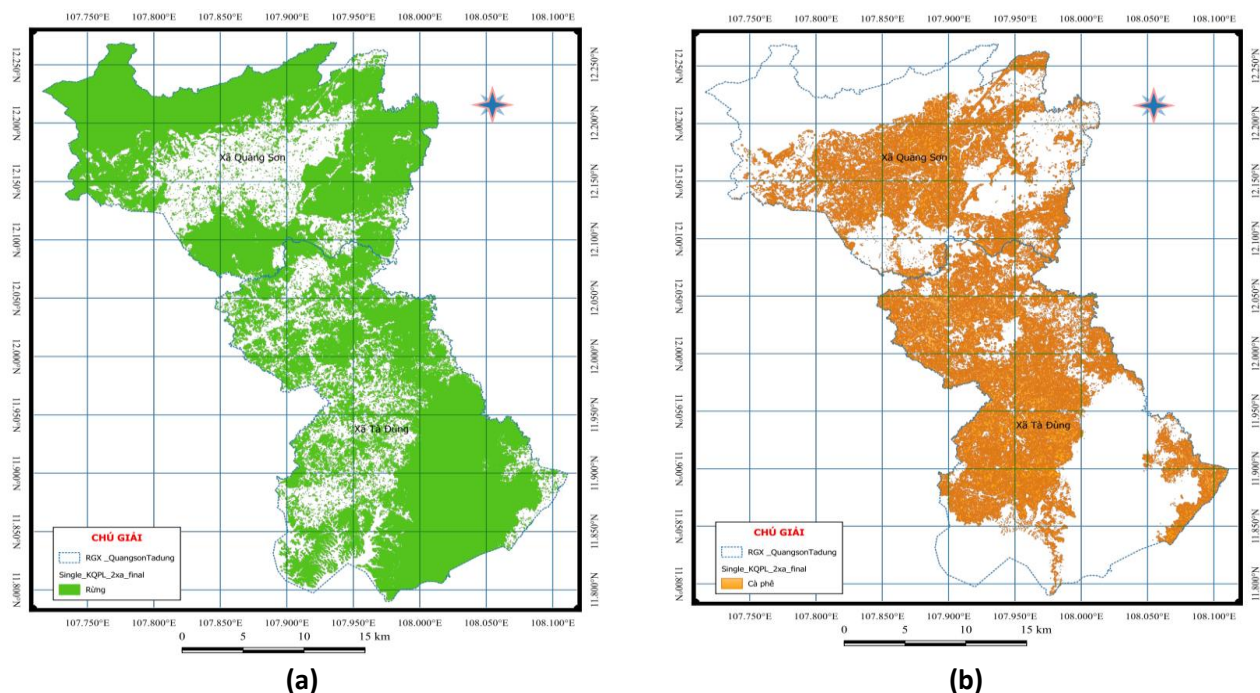


Hình 5. Kết quả phân loại các trạng thái lớp phủ năm 2020 (a1, bc, c1, d1) và kết quả phân loại được kiểm chứng thông qua ảnh Google Earth (a2, b2, c2, d2) và ảnh chụp ngoài thực địa (a3, b3, c3, d3)

Có 4 khu vực cụ thể đã được kiểm tra và đối chiếu với ảnh Google Earth vào tháng 12/2020 và ảnh chụp ngoài thực địa tại 2 xã nghiên cứu vào thời điểm tháng 12 năm 2024 ở cùng một vị trí. Trong đó, đất canh tác nông nghiệp và cà phê được phân loại chi tiết (Hình 5a) gồm: kết quả phân loại từ ViT (Hình 5a1), hình ảnh trực quan trên ảnh Google Earth (Hình 5a2) và ảnh chụp ngoài thực địa (Hình 5a3). Hình 5b1 thể hiện kết quả phân loại giữa cà phê, rừng và đất xây dựng gồm: Hình 5b2 thể hiện kết quả kiểm chứng trực quan bằng ảnh Google Earth và Hình 5b3 cho thấy hình ảnh ngoài thực địa kèm theo tọa độ địa lý. Tương tự, Hình 5c1 - 5c3 thể hiện kết quả phân loại giữa cà phê và đất xây dựng. Cuối cùng, Hình 5d1, 5d2 và 5d3 thể hiện kết quả phân loại giữa cà phê, rừng và

đất nông nghiệp.

Kết quả phân loại đã được kiểm tra tại 46 điểm ngoài thực địa từ dữ liệu thu thập năm 2024 của nhóm nghiên cứu có kết quả trùng khớp đạt tỉ lệ 91,3% (3/3 điểm rừng, 14/16 điểm đất nông nghiệp và cây trồng khác, 25/27 điểm cà phê). Các kết quả này đã được kiểm chứng thông qua khảo sát, phỏng vấn chủ rừng và chủ lô đất để xác nhận về trạng thái thực vật năm 2020. Những kiểm chứng này củng cố tính tin cậy của bản đồ phân loại ở các vị trí không có sự thay đổi trạng thái lớp phủ từ cuối năm 2020 đến 2024. Từ kết quả phân loại các trạng thái cho phép thành lập được bản đồ các khu vực có rừng và các khu vực canh tác cà phê năm 2020 của khu vực nghiên cứu (Hình 6).



Hình 6. Hình ảnh bản đồ các khu vực có rừng (a) và khu vực canh tác cà phê (b) năm 2020 của hai xã

Số liệu thống kê diện tích các lô rừng sau khi đã loại bỏ các vùng nhỏ hơn 0,01 ha (nhỏ hơn diện tích 01 pixel ảnh Sentinel-2) tại khu vực

được nghiên cứu được phân loại các trạng thái chi tiết tại Bảng 4.

Bảng 4. Diện tích phân loại các trạng thái lớp phủ tại khu vực nghiên cứu

TT	Trạng thái	Diện tích (ha)		Tổng (ha)
		Xã Quảng Sơn	Xã Tà Đùng	
1	Khu vực có rừng	27.259,5	28.004,5	55.264,0
2	Mặt nước	1.403,6	3.724,2	5.127,8
3	Cà phê	8.831,1	17.253,3	26.084,4
4	Đất nông nghiệp và trồng cây khác	7.117,3	3.496,6	10.613,9
5	Đất xây dựng	809,8	645,3	1.455,1
Tổng:		45.421,3	53.123,9	98.545,2

Kết quả phân tích diện tích lớp phủ cho thấy diện tích rừng chiếm ưu thế với hơn 55.000 ha (chiếm hơn 56% tổng diện tích 2 xã), tiếp theo là cà phê hơn 26.000 ha (chiếm 26,5%), đất nông nghiệp 10.613,4 ha (chiếm 10,7%), mặt nước 5.127,8 ha và đất xây dựng 1.455,1 ha. Diện tích rừng và cà phê tập trung chủ yếu ở xã Tà Đùng, trong khi xã Quảng Sơn có tỉ lệ đất canh tác nông nghiệp cao hơn.

4. THẢO LUẬN

Việc sử dụng ảnh Sentinel-2 đa thời gian vào tháng 11 và tháng 12 năm 2020 giúp khắc phục hạn chế tỉ lệ mây lớn của dữ liệu quang học và phản ánh rõ đặc tính phổ ổn định của lớp phủ

đồng thời cho phép giảm được sai số trong quá trình phân loại [10, 20], đặc biệt đối với đất sản xuất nông nghiệp có thể thay đổi trong thời gian ngắn. Ngoài ra, vào thời điểm này là mùa khô ở khu vực Tây nguyên các hoạt động sản xuất nông nghiệp diễn ra rất ít, các khu vực đất trồng cà phê và nông nghiệp gần như ổn định, do vậy, có thể thấy đây là thời điểm phù hợp để phân loại cà phê với đất canh tác nông nghiệp và cây trồng khác.

Trong nghiên cứu này, trí tuệ nhân tạo và nền tảng GEE được sử dụng kết hợp một cách hiệu quả cùng với 7 chỉ số thực vật phổ biến để làm dữ liệu đầu vào cho quá trình phân loại rõ

ràng các lớp phủ. Kết quả phân loại diện tích lớp phủ bằng mô hình ViT phản ánh sát thực trạng sử dụng đất tại xã Quảng Sơn và xã Tà Đùng với cấu trúc không gian lớp phủ được thể hiện khá rõ ràng và độ chính xác đảm bảo, đặc biệt tại các khu vực tiếp giáp giữa rừng và cà phê vốn nơi thường xảy ra nhầm lẫn trong các phương pháp phân loại truyền thống. Mô hình không ghi nhận hiện tượng phân loại sai lớn tại các vùng giáp ranh cho thấy khả năng phân biệt ngữ cảnh không gian toàn cục của mô hình hoạt động hiệu quả.

Khi so sánh diện tích của các khu vực được xác định có rừng từ bản đồ rừng toàn cầu năm 2020 phiên bản số 3 (GFC 2020v3) do Trung tâm nghiên cứu chung châu Âu (JRC) cung cấp miễn phí. Nhóm tác giả thấy diện tích phát hiện rừng của JRC ở khu vực nghiên cứu khoảng 50.950 ha, thấp hơn kết quả của nghiên cứu này khoảng 4.300 ha. Tuy nhiên JRC hiện chưa có đánh giá độ chính xác cấp địa phương hoặc cấp quốc gia cho Việt Nam để có sự đối sánh một cách khách quan. Đối với dữ liệu của địa phương, diện tích các trạng thái lớp phủ ở nghiên cứu này cũng tương đồng với các thống kê năm 2020, trong đó rừng chiếm hơn 50% và cà phê khoảng 25% tổng diện tích tự nhiên. Độ chính xác của kết quả giải đoán cũng như kết quả kiểm chứng thực địa của mô hình phân loại chứng tỏ phương pháp nghiên cứu này phù hợp với xu hướng chung về sử dụng ảnh viễn thám và các thuật toán học sâu và cũng cho thấy ViT rất hiệu quả để phân biệt lớp phủ ở khu vực nghiên cứu [19, 21]. Kết quả trong nghiên cứu này cũng cho thấy ViT có thể cải thiện thêm 2 - 5% độ chính xác tổng thể so với các mô hình phân loại truyền thống khi sử dụng đặc biệt ở các lớp phủ khó phân biệt. Mặc dù vậy, mô hình vẫn gặp một số thách thức khi phân biệt các trạng thái có đặc điểm quang phổ tương đồng, điển hình là cà phê với các loài cây công nghiệp khác như cao su, hồ tiêu hoặc rừng trồng mới, các hạn chế này có thể khắc phục khi các dữ liệu thực địa đảm bảo cả về số lượng và

chất lượng. Nhìn chung, các kết quả nghiên cứu cho thấy tiềm năng ứng dụng công nghệ tiên tiến trong quản lý, giám sát tài nguyên rừng và minh bạch chuỗi sản xuất nông lâm nghiệp để đáp ứng các quy định mới của quốc tế, trong đó có EUDR, đồng thời cũng khẳng định sự phù hợp của quy trình và phương pháp phân loại, tạo tiền đề cho các nghiên cứu tương tự tiếp theo nhằm nâng cao khả năng nhận diện các lớp phủ và đất trồng cây công nghiệp lâu năm như cà phê trong trường hợp quá trình khảo sát thực địa gặp nhiều khó khăn.

5. KẾT LUẬN

Kết quả nghiên cứu đã xác lập được cơ sở khoa học và tính khả thi của việc kết hợp AI và GEE trong giám sát lớp phủ phục vụ EUDR, đồng thời khẳng định ưu việt của mô hình Vision Transformer trong phân loại lớp phủ ở khu vực Tây Nguyên, đặc biệt đối với nhận diện các khu vực canh tác cà phê tại xã Quảng Sơn và xã Tà Đùng, tỉnh Lâm Đồng. Kết quả huấn luyện đạt độ chính xác tổng thể (OA) đạt 91%, điểm F1 trung bình có trọng số đạt 0,91 trong khi kết quả kiểm tra sau phân loại vẫn duy trì OA đạt 88%, thể hiện tính ổn định và khả năng tổng quát tốt của mô hình. Kết quả nghiên cứu đã phân loại và xây dựng được bản đồ các khu vực canh tác cà phê và bản đồ ranh giới rừng bằng trí tuệ nhân tạo kết hợp với GEE. Kết quả tính toán diện tích các trạng thái cho thấy, diện tích rừng chiếm ưu thế với hơn 55.000 ha (hơn 56% tổng diện tích 2 xã), cà phê hơn 26.000 ha (chiếm 26,5%), đất nông nghiệp và cây trồng khác 10.613,4 ha (chiếm 10,7%), còn lại là mặt nước và đất xây dựng. Kết quả này không chỉ khẳng định khả năng ứng dụng trí tuệ nhân tạo trong phân loại lớp phủ có cấu trúc phức tạp ở Tây Nguyên, mà còn cung cấp cơ sở dữ liệu quan trọng phục vụ giám sát và chứng minh vùng trồng cà phê không gây mất rừng, góp phần đáp ứng yêu cầu truy xuất nguồn gốc theo EUDR, đồng thời mở ra tiềm năng ứng dụng công nghệ tiên tiến trong việc giám sát rừng ở quy mô lớn trong tương lai.

Lời cảm ơn

Nghiên cứu này là sản phẩm của Đề tài Khoa học và Công nghệ tiềm năng cấp Bộ: “*Nghiên cứu thí điểm ứng dụng GIS và công nghệ viễn thám để hỗ trợ truy xuất nguồn gốc sản xuất Cà phê không gây mất rừng tại Đắk Nông đáp ứng quy định mới của Ủy ban châu Âu (EU)*”. Trân trọng cảm ơn các tổ chức, cá nhân có liên quan và các nhà khoa học đã hỗ trợ nhóm tác giả hoàn thiện nghiên cứu này.

TÀI LIỆU THAM KHẢO

[1]. FAO (2020). Global Forest Resources Assessment 2020.

[2]. Philip G Curtis, Christy M Slay, Nancy L Harris, Alexandra Tyukavina & Matthew C Hansen (2018). Classifying drivers of global forest loss. *Science*. 361(6407): 1108-1111.

[3]. Christopher L Gilbert (2024). The EU deforestation regulation. *EuroChoices*. 23(3): 64-70.

[4]. Noel Gorelick, Matt Hancher, Mike Dixon, Simon Ilyushchenko, David Thau & Rebecca Moore (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment*. 202: 18-27.

[5]. Pardhasaradhi Teluguntla, Prasad S Thenkabail, Adam Oliphant, Jun Xiong, Murali Krishna Gumma, Russell G Congalton, Kamini Yadav & Alfredo Huete (2018). A 30-m landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform. *ISPRS journal of photogrammetry and remote sensing*. 144: 325-340.

[6]. Trần Quang Bảo, Nguyễn Trọng Cường, Lê Sỹ Doanh, Lê Nguyên Khang, Đinh Văn Tuyển & Lê Đức Nhật Minh (2023). Kết hợp ảnh vệ tinh sentinel-2 và google earth engine để xác định diện tích rừng chuyển đổi sang đất nông nghiệp giai đoạn 2016-2021: trường hợp nghiên cứu tại huyện Đắk Glong, tỉnh Đắk Nông. *Tạp chí Khoa học và Công nghệ Lâm nghiệp*. 2: 54-64. DOI: 10.55250/jo.vnuf.2023.2.054-064

[7]. Trần Quang Bảo, Nguyễn Trọng Cường & Mai Hà An (2016). Nghiên cứu xây dựng phần mềm tự động phát hiện sớm cháy rừng từ Trạm quan trắc mặt đất. *Tạp chí Khoa học Lâm nghiệp*. 3(3): 4538 - 4546.

[8]. Panagiotis Barmountis, Periklis Papaioannou, Kosmas Dimitropoulos & Nikos Grammalidis (2020). A review on early forest fire detection systems using optical remote sensing. *Sensors*. 20(22): 6442.

[9]. Xiao Xiang Zhu, Devis Tuia, Lichao Mou, Gui-Song Xia, Liangpei Zhang, Feng Xu & Friedrich Fraundorfer (2017). Deep learning in remote sensing: A

comprehensive review and list of resources. *IEEE geoscience and remote sensing magazine*. 5(4): 8-36.

[10]. Lei Ma, Yu Liu, Xueliang Zhang, Yuanxin Ye, Gaofer Yin & Brian Alan Johnson (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS journal of photogrammetry and remote sensing*. 152: 166-177.

[11]. Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold & Sylvain Gelly (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.

[12]. Nathalie Pettorelli (2013). The normalized difference vegetation index. Oxford University Press, USA.

[13]. Suming Jin & Steven A Sader (2005). Comparison of time series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances. *Remote sensing of Environment*. 94(3): 364-372.

[14]. Yong Zha, Jay Gao & Shaoxiang Ni (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International journal of remote sensing*. 24(3): 583-594.

[15]. Hongmei Zhao & Xiaoling Chen (2005). Use of normalized difference bareness index in quickly mapping bare areas from TM/ETM+. *International geoscience and remote sensing symposium*. 1666.

[16]. Bo-Cai Gao (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote sensing of environment*. 58(3): 257-266.

[17]. Alfredo R Huete (1988). A soil-adjusted vegetation index (SAVI). *Remote sensing of environment*. 25(3): 295-309.

[18]. Xiang Gao, Alfredo R Huete, Wenge Ni & Tomoaki Miura (2000). Optical–biophysical relationships of vegetation spectra without background contamination. *Remote sensing of environment*. 74(3): 609-620.

[19]. Shunzhou Wang, Tianfei Zhou, Yao Lu & Huijun Di (2022). Detail-preserving transformer for light field image super-resolution. *Proceedings of the AAAI conference on artificial intelligence*. 2522-2530.

[20]. Ewa Grabska, Patrick Hostert, Dirk Pflugmacher & Katarzyna Ostapowicz (2019). Forest stand species mapping using the Sentinel-2 time series. *Remote Sensing*. 11(10): 1197.

[21]. Ava Vali, Sara Comai & Matteo Matteucci (2020). Deep learning for land use and land cover classification based on hyperspectral and multispectral earth observation data: A review. *Remote Sensing*. 12(15): 2495.