



Tích hợp viễn thám và học máy trên nền tảng điện toán đám mây để lập bản đồ lớp phủ đất và phân tích ngập lụt tại Quảng Ngãi

Integrating Remote Sensing and Machine Learning on a Cloud Computing Platform for Land Use/Land Cover Mapping and Inundation Analysis in Quang Ngai

Hoàng Nhật Đức^{a*, b}, Nguyễn Quốc Lâm^b
Hoang Nhat Duc^{a*, b}, Nguyen Quoc Lam^b

^a*Viện Nghiên cứu và Phát triển Công nghệ Cao, Đại học Duy Tân, Đà Nẵng, Việt Nam*

^a*Institute of Research and Development, Duy Tan University, Da Nang, 550000, Viet Nam*

^b*Khoa Xây dựng, Trường Công nghệ, Đại học Duy Tân, Đà Nẵng, Việt Nam*

^b*Faculty of Civil Engineering, School of Engineering and Technology, Duy Tan University, Da Nang, 550000, Viet Nam*

(Ngày nhận bài: 04/4/2026, ngày phản biện xong: 28/7/2026, ngày chấp nhận đăng: 18/8/2026)

Tóm tắt

Nghiên cứu của chúng tôi trình bày việc ứng dụng nền tảng điện toán đám mây Google Earth Engine (GEE) trong phân loại hiện trạng lớp phủ đất và đánh giá tác động của ngập lụt tại khu vực đô thị ven biển thuộc tỉnh Quảng Ngãi, Việt Nam. Ảnh Sentinel-2 thu thập vào năm 2025 được xử lý bằng các thuật toán học máy, bao gồm mô hình rừng phân loại và cây phân loại tăng cường. Hiệu suất của các mô hình được đánh giá thông qua phương pháp kiểm định chéo, trong đó mô hình cây phân loại tăng cường đạt độ chính xác trung bình khoảng 94%, cao hơn so với mô hình rừng phân loại (92%). Dựa trên mô hình cây phân loại tăng cường, bản đồ hiện trạng lớp phủ đất chi tiết cho năm 2025 đã được xây dựng. Dữ liệu vùng ngập từ nền tảng Sentinel Asia được tích hợp nhằm phân tích ảnh hưởng của đợt mưa lớn và ngập lụt xảy ra trong giai đoạn 16–19 tháng 11 năm 2025. Kết quả phân tích giúp xác định các khu vực đô thị bị ảnh hưởng bởi ngập lụt và cho thấy tính hữu ích của việc kết hợp dữ liệu viễn thám với học máy và xử lý dựa trên điện toán đám mây. Kết quả nghiên cứu cho thấy hiệu quả của nền tảng GEE trong việc xác lập bản đồ hiện trạng lớp phủ đất nhanh chóng, đáng tin cậy và đánh giá mức độ ảnh hưởng của mưa lũ cho các khu vực đô thị ven biển.

Từ khóa: viễn thám, phân loại lớp phủ đất, đánh giá ảnh hưởng của ngập lụt, học máy, điện toán đám mây

Abstract

This study presents an application of the Google Earth Engine (GEE) cloud computing platform for land cover classification and flood impact assessment in an urban center of the coastal region of Quang Ngai, Vietnam. Sentinel-2 imagery acquired in 2025 was processed using machine learning algorithms, including random forest and gradient boosting classification tree models. Model performance was evaluated through cross-validation, with Gradient Boosting achieving the highest average accuracy of approximately 94%, slightly outperforming Random Forest (92%). Using the gradient boosting classification tree, a detailed land cover map for 2025 was generated. Flood-inundation data from the Sentinel Asia platform were incorporated to analyze the impacts of the heavy rainfall and flooding event that occurred between November 16–19, 2025. The analysis identified key urban areas affected by inundation and demonstrated the usefulness of remote sensing data combined with machine learning and cloud-based processing. The results highlight the potential of the GEE platform for rapid and reliable land cover mapping and inundation assessment in urban coastal environments.

Keywords: remote sensing, land cover classification, inundation assessment, machine learning, cloud computing

*Tác giả liên hệ: Hoàng Nhật Đức

Email: hoangnhatduc@duytan.edu.vn

1. Giới thiệu

Việc lập bản đồ và giám sát lớp phủ đất cung cấp thông tin không gian quan trọng cho quy hoạch đô thị vì các bản đồ này hiển thị sự thay đổi của các khu vực trong đô thị theo thời gian [1]. Bằng cách trực quan hóa các thay đổi về bề mặt xây dựng (như bê tông và đường nhựa), thảm thực vật và mặt nước, dữ liệu lớp phủ giúp xác định các khu vực mà quá trình đô thị hóa nhanh đang làm gia tăng dòng chảy mặt, từ đó trực tiếp làm tăng tần suất và mức độ nghiêm trọng của ngập lụt đô thị [2]. Việc giám sát lớp phủ đất thường xuyên hỗ trợ công tác đánh giá ảnh hưởng của ngập lụt bằng cách liên hệ thiệt hại quan sát được với các loại hình sử dụng đất cụ thể (ví dụ như đất ở đô thị hay đất trồng trọt), qua đó cung cấp cơ sở cho điều chỉnh quy hoạch sử dụng đất, ban hành các quy định sử dụng đất trong khu vực có rủi ro ngập lụt cao và xây dựng các chiến lược phát triển đô thị thích ứng khí hậu trong dài hạn [3, 4].

Các cảm biến trên vệ tinh thu nhận các quan sát bề mặt Trái Đất trên những khu vực rộng lớn với nhiều độ phân giải không gian và thời gian khác nhau. Lợi thế này giúp dữ liệu vệ tinh rất phù hợp cho việc giám sát các khu vực đô thị rộng lớn và có tính biến động cao. Các cảm biến vệ tinh đa phổ là nguồn dữ liệu chính cho phân loại lớp phủ đất. Trong số đó, ảnh Sentinel-2 được cung cấp miễn phí và rất hiệu quả cho việc lập bản đồ giám sát hiện trạng bề mặt đất [5]. Với lượng dữ liệu lớn thu thập từ các vệ tinh, việc xử lý và trích xuất dữ liệu là một thách thức lớn. Tuy nhiên, trong những năm gần đây, Google Earth Engine (GEE), một nền tảng xử lý địa không gian dựa trên điện toán đám mây, ngày càng được ứng dụng rộng rãi trong việc lưu trữ và phân tích các bộ dữ liệu này [6]. GEE cung cấp các kho lưu trữ đồ sộ ảnh vệ tinh cùng thư viện phong phú về các công cụ xử lý ảnh và thuật toán học máy. GEE đã đơn giản hóa đáng kể quy trình tiền xử lý và phân loại ảnh vệ tinh, dẫn đến

việc được ứng dụng rộng rãi trong các bài toán như giám sát và lập bản đồ lớp phủ đất [5].

Phương pháp có giám sát thường được sử dụng cho phân loại lớp phủ đất vì chúng thường giúp đạt được độ chính xác phân loại cao. Trong số các phương pháp này, rừng phân loại và cây phân loại tăng cường được tích hợp trong GEE và có thể được sử dụng để huấn luyện các mô hình tính toán dựa trên dữ liệu trích xuất từ ảnh vệ tinh Sentinel-2. Rừng phân loại có ưu điểm nổi bật là xử lý tốt dữ liệu có số lượng lớn biến đầu vào và ít nhạy với độ nhiễu trong dữ liệu. Trong khi đó, cây phân loại tăng cường xây dựng mô hình bằng cách kết hợp tuần tự nhiều cây quyết định yếu để dần giảm sai số, nên thường đạt độ chính xác rất cao.

Trong bối cảnh biến đổi khí hậu, tần suất và cường độ các hiện tượng thời tiết cực đoan như mưa lớn kéo dài đang gia tăng, gây áp lực ngày càng lớn lên các đô thị ven biển miền Trung Việt Nam, nơi địa hình hẹp, dốc và hệ thống sông ngòi dày đặc khiến nguy cơ lũ lụt và sạt lở đặc biệt cao. Trong bối cảnh đó, việc kết hợp lập bản đồ và giám sát lớp phủ đất với đánh giá ảnh hưởng của ngập lụt ở đô thị đóng vai trò then chốt. Bản đồ lớp phủ đất giúp nhận diện các khu vực đô thị có mật độ xây dựng cao có nguy cơ ngập lụt. Những thông tin này cung cấp cơ sở khoa học cho điều chỉnh quy hoạch sử dụng đất và ban hành các quy định kiểm soát xây dựng tại các khu vực đô thị ven biển.

Trong những năm gần đây, sự phát triển vượt bậc của học máy và dữ liệu viễn thám mở đã tạo ra bước tiến lớn trong công tác quản lý tài nguyên và giám sát thiên tai đô thị. Đối với bài toán phân loại lớp phủ đất, các thuật toán học máy phổ biến như Random Forest, Support Vector Machine hay Gradient Boosting Machine đã chứng minh khả năng xử lý xuất sắc các tập dữ liệu ảnh vệ tinh đa phổ (như Sentinel, Landsat) nhờ năng lực tự động trích xuất đặc trưng và xử lý tốt tính phi tuyến của dữ liệu [5]. Việc xác định chính xác hiện trạng và biến động

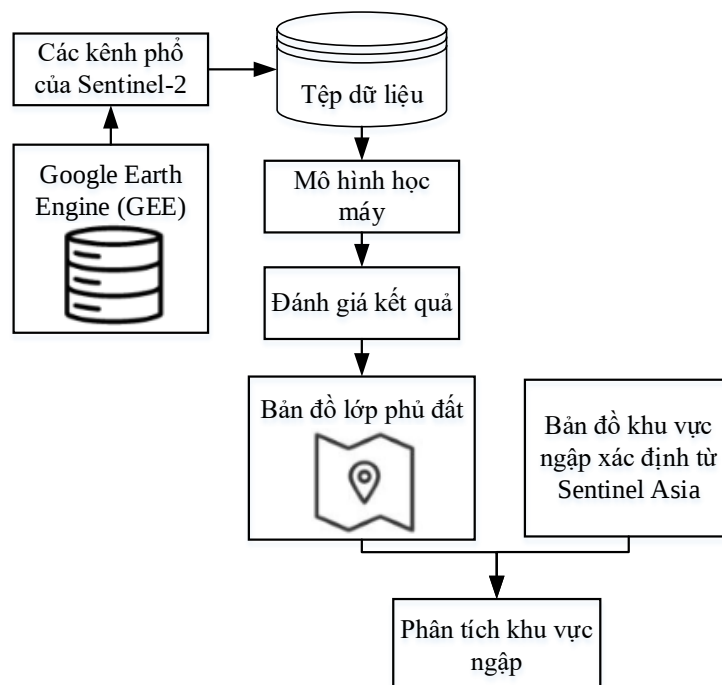
lớp phủ đất đóng vai trò là cơ sở đầu vào then chốt cho các phân tích không gian. Đặc biệt, trong bài toán phân tích ngập lụt đô thị, kết quả phân loại lớp phủ đất tích hợp cùng dữ liệu địa hình, chỉ số bề mặt ngập và các mô hình học máy giúp nhanh chóng khoanh vùng bề mặt ngập, đánh giá mức độ ngập lụt và xác định các khu vực đô thị hóa có rủi ro cao [7, 8]. Nhờ khả năng xử lý dữ liệu lớn với độ chính xác cao và thời gian phân tích nhanh, học máy đang trở thành công cụ hỗ trợ đắc lực giúp các nhà quy hoạch đưa ra chính xác các quyết định ứng phó với biến đổi khí hậu và phát triển đô thị bền vững [9].

Bài báo này đóng góp một quy trình tích hợp toàn diện từ trích xuất dữ liệu viễn thám đến phân tích rủi ro ngập lụt đô thị trên nền tảng điện toán đám mây GEE. Tính mới nổi bật của nghiên cứu nằm ở việc giải quyết khoảng trống dữ liệu và phương pháp đánh giá ngập lụt tại khu vực đô thị ven biển thuộc tỉnh Quảng Ngãi - nơi các ứng dụng tích hợp học máy trong giám sát thiên tai vẫn còn nhiều hạn chế. Nghiên cứu chứng minh sự ưu việt của thuật toán Cây phân loại tăng cường độ dốc trong việc thành lập bản đồ lớp phủ đất với độ chính xác cao, mà còn là nghiên

cứu kết hợp dữ liệu lớp phủ đất cập nhật năm 2025 với dữ liệu không gian ngập lụt (Sentinel Asia) từ đợt mưa lớn 16–19/11/2025. Kết quả đạt được cung cấp cơ sở khoa học và công cụ giám sát nhanh, chi phí thấp, hỗ trợ trực tiếp các nhà quy hoạch trong việc quản lý rủi ro thiên tai và phát triển đô thị ven biển bền vững.

2. Phương pháp nghiên cứu

Quy trình thực hiện nghiên cứu được triển khai qua các bước chính như mô tả trong Hình 1. Dữ liệu ảnh vệ tinh Sentinel-2 được trích xuất trực tiếp trên nền tảng điện toán đám mây GEE, từ đó chúng tôi chọn lọc các kênh phổ làm đặc trưng để xây dựng tập dữ liệu viễn thám đầu vào. Tiếp theo, tập dữ liệu này được đưa vào các mô hình học máy để tiến hành phân loại lớp phủ đất. Sau bước đánh giá độ chính xác của kết quả phân loại, bản đồ lớp phủ đất hoàn thiện sẽ được chồng xếp và kết hợp với bản đồ khu vực ngập lụt được xác định từ Sentinel Asia. Cuối cùng, sự tích hợp này cho phép thực hiện phân tích về diện tích ngập lụt tương ứng với từng loại lớp phủ đất. Từ các phân tích trên, chúng tôi khoanh vùng và nhận diện các khu vực trong đô thị có nguy cơ ngập lụt cao.



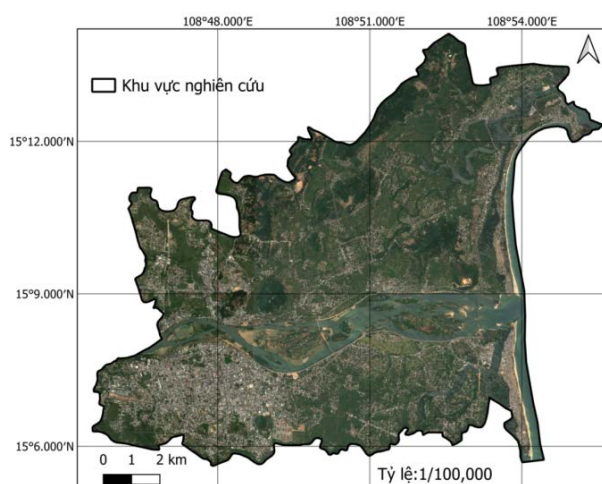
Hình 1. Quy trình thực hiện nghiên cứu

2.1. Khu vực nghiên cứu và dữ liệu viễn thám

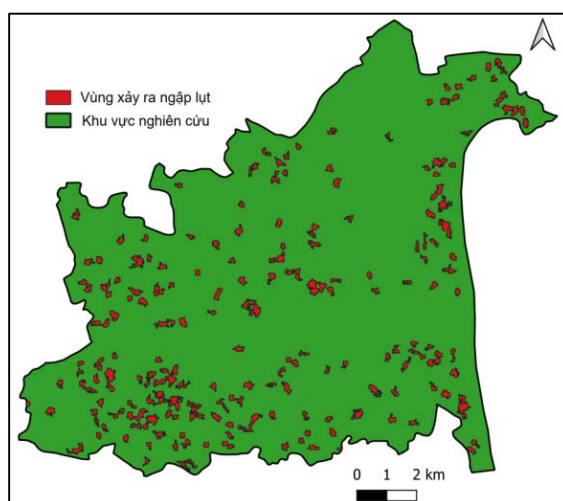
Khu vực ven biển miền Trung có địa hình hẹp ngang, dải đồng bằng ven biển thấp trũng bị kẹp giữa núi và biển. Thêm vào đó, mạng lưới sông suối ngắn, dốc thường gây tập trung dòng chảy nhanh và làm gia tăng nguy cơ ngập lụt khi xảy ra mưa lớn. Trong bài báo này, một khu vực trung tâm đô thị ven biển tại tỉnh Quảng Ngãi (xem Hình 2) được lựa chọn làm khu vực nghiên cứu. Để xây dựng bản đồ lớp phủ đất, nghiên cứu sử dụng dữ liệu ảnh Sentinel-2 trong giai đoạn từ ngày 1/1/2025 đến 31/12/2025. Dữ liệu đầu vào bao gồm các kênh B2, B3, B4, B8 (độ phân giải 10 m) và B5, B6, B7, B8A, B11, B12 (độ phân giải 20 m), trong đó các kênh 20 m được

đồng nhất về độ phân giải 10 m bằng thuật toán nội suy láng giềng gần nhất (Nearest Neighbor interpolation) trên nền tảng GEE.

Quy trình tiền xử lý bao gồm lọc các ảnh theo không gian và thời gian, đồng thời loại bỏ những cảnh có tỷ lệ mây cao bằng cách chỉ giữ lại các ảnh có mây bao phủ dưới 10%, nhằm giảm thiểu ảnh hưởng của mây và bóng mây đến chất lượng phân loại. Sau đó, các cảnh ảnh đáp ứng tiêu chí được tổng hợp bằng phép lọc trung vị để tạo thành một ảnh tổng hợp đại diện cho điều kiện lớp phủ đất trong cả năm 2025, và cuối cùng được cắt theo ranh giới khu vực nghiên cứu để phục vụ cho bước phân loại lớp phủ đất.



Hình 2. Vị trí khu vực nghiên cứu (ảnh tổ hợp màu thật từ Sentinel-2, ESA/Copernicus)



Hình 3. Khu vực ngập lụt

(Nguồn: <https://sentinel-asia.org/EO/EmergencyObservation.html>)

Trong nghiên cứu này, chúng tôi thu thập dữ liệu về các vị trí ngập lụt tại khu vực nghiên cứu trong đợt mưa rất lớn xảy ra từ ngày 16–19/11/2025. Thông tin về khu vực bị ngập được trích xuất từ kết quả phân tích ngập lụt do Cơ quan nghiên cứu phát triển Hàng không vũ trụ Nhật Bản (JAXA) cung cấp thông qua nền tảng Sentinel Asia (xem Hình 3), trong đó các khu vực bị ngập được nhận diện dựa trên dữ liệu quan sát Trái Đất và các kỹ thuật xử lý ảnh vệ tinh. Sự kiện mưa cực đoan này đã gây ngập lụt lớn, làm mực nước nhiều sông vượt đỉnh lũ lịch sử và gây ngập diện rộng tại nhiều tỉnh ven biển miền Trung. Sử dụng dữ liệu này, chúng tôi minh họa khả năng ứng dụng bản đồ lớp phủ đất kết hợp với dữ liệu ngập lụt trong phân tích rủi ro ngập đô thị.

2.2. Các phương pháp học máy

Nhờ vào khả năng phân tích dữ liệu đa phổ phi tuyến và hạn chế sự quá khớp, các thuật toán học máy dựa trên cây quyết định như rừng phân loại (Random Forest) và cây phân loại tăng cường độ dốc (Gradient Boosting Machine) rất phù hợp cho nhiệm vụ phân loại lớp phủ đất từ dữ liệu viễn thám. Sự khác biệt cốt lõi giữa hai phương pháp nằm ở cơ chế kết hợp mô hình. Mô hình rừng phân loại áp dụng kỹ thuật Bagging để xây dựng các cây quyết định độc lập và song song, sau đó lấy kết quả dựa trên sự bầu chọn đa số nhằm giảm độ biến động. Thuật toán cây phân loại tăng cường độ dốc vận hành theo cơ chế Boosting, xây dựng các cây phân loại nối tiếp nhau sao cho cây sau tập trung sửa chữa lỗi sai của cây trước.

2.2.1. Mô hình rừng phân loại

Mô hình rừng phân loại (Random Forest) [10] là một thuật toán học máy có giám sát dựa trên tập hợp nhiều cây quyết định, trong đó mỗi cây được huấn luyện trên một mẫu con và tập biến con được chọn ngẫu nhiên, sau đó kết quả phân loại cuối cùng được xác định theo nguyên tắc bỏ phiếu đa số. Nhờ cấu trúc tổ hợp này, mô hình

có ưu điểm nổi bật là khả năng mô tả tốt các quan hệ phi tuyến phức tạp giữa tín hiệu phổ thu thập từ vệ tinh và lớp phủ đất. Mô hình này ít nhạy với nhiễu và hiện tượng quá khớp, đồng thời cho phép làm việc hiệu quả với tập dữ liệu có số lượng lớn kênh ảnh và chỉ số phổ.

Rừng phân loại phù hợp cho bài toán phân loại lớp phủ đất từ dữ liệu viễn thám đa phổ như Sentinel-2, nơi các lớp phủ đất thường có phổ chồng lấn, điều kiện quan sát thay đổi theo không gian và thời gian, và cần một mô hình vừa linh hoạt vừa ổn định để tạo ra bản đồ lớp phủ đất. Trong GEE, mô hình này có thể được sử dụng thông qua hàm `ee.Classifier.smileRandomForest()` để huấn luyện bộ phân loại Random Forest từ tập mẫu huấn luyện.

2.2.2. Mô hình cây phân loại tăng cường độ dốc

Mô hình cây phân loại tăng cường độ dốc (Gradient Boosting Machine) [11] là một thuật toán học máy có giám sát dựa trên việc kết hợp tuần tự nhiều cây quyết định yếu để tạo thành một bộ phân loại mạnh, trong đó mỗi cây mới được huấn luyện nhằm giảm dần sai số còn lại của toàn bộ mô hình. Nhờ cơ chế tăng cường theo từng bước như vậy, mô hình có khả năng mô tả rất tốt các quan hệ phi tuyến phức tạp giữa đặc trưng đầu vào và nhãn lớp, đồng thời linh hoạt trong việc điều chỉnh các siêu tham số như số lượng cây và tốc độ học để kiểm soát hiện tượng quá khớp. Trong bài toán phân loại lớp phủ đất từ dữ liệu viễn thám, cây phân loại tăng cường đặc biệt hữu ích nhờ khả năng tinh chỉnh ranh giới phân tách lớp một cách dần dần và chính xác hơn. Trong GEE, mô hình cây phân loại tăng cường độ dốc này có thể được triển khai thông qua hàm `ee.Classifier.smileGradientTreeBoost()`. Hàm này cho phép huấn luyện và áp dụng cây phân loại tăng cường độ dốc trực tiếp trên nền tảng điện toán đám mây đối với các tập dữ liệu ảnh vệ tinh có kích thước lớn.

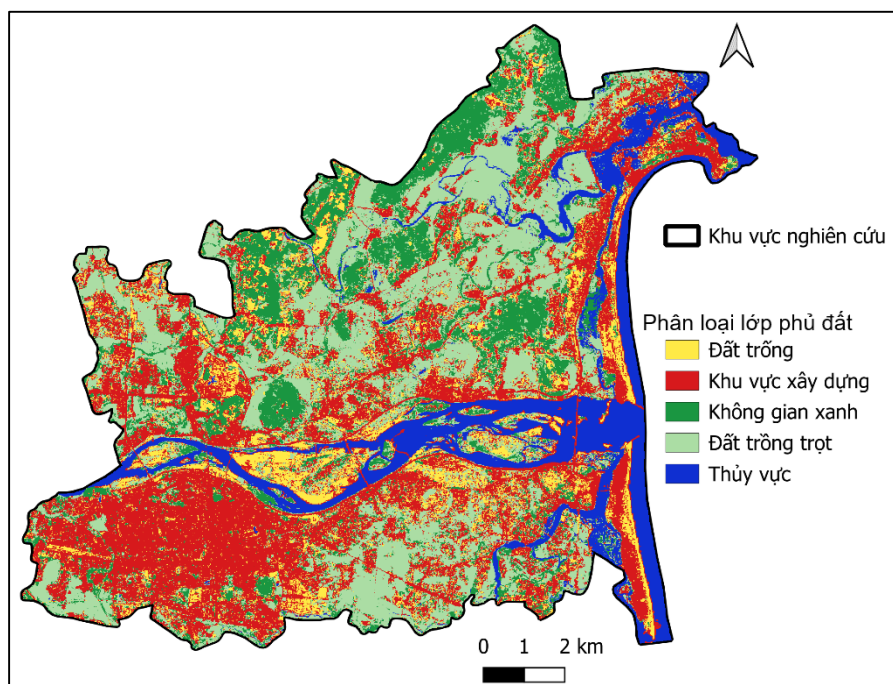
3. Kết quả và thảo luận

Trong nghiên cứu này, hai mô hình rừng phân loại (RF) và cây phân loại tăng cường độ dốc (GBM) được đánh giá dựa trên cùng một tập dữ liệu mẫu huấn luyện và kiểm định. Năm lớp phủ đất bao gồm đất trống, khu vực xây dựng, không gian xanh, đất trồng trọt và thủy vực. Các lớp phủ đất này được lựa chọn dựa trên đặc tính thảm nước phục vụ phân tích ngập lụt đô thị, đồng thời phù hợp với khả năng phân giải không gian 10 m của ảnh Sentinel-2 nhằm tối ưu hóa quá trình

gán nhãn dữ liệu và đảm bảo độ chính xác phân loại cho mô hình. Mỗi lớp được thu thập 200 mẫu, tổng cộng 1000 mẫu điểm. Các mẫu này được gán nhãn thông qua quan sát từ ảnh có độ phân giải cao trên Google Earth Pro. Tập mẫu này được chia ngẫu nhiên thành 5 tập con và áp dụng quy trình đánh giá chéo: ở mỗi lượt, bốn phần dữ liệu được dùng để huấn luyện mô hình, phần còn lại dùng để kiểm định, sau đó chỉ số phân loại chính xác được tính trung bình qua 5 lượt.

Bảng 1. Tỷ lệ phân loại chính xác theo lớp phủ đất (%) của hai mô hình RF và GBM

Lớp phủ đất	RF	GBM
Đất trống	88,83	91,88
Khu vực xây dựng	89,55	90,05
Không gian xanh	95,16	97,31
Đất trồng trọt	94,55	96,04
Thủy vực	93,93	93,46



Hình 4. Bản đồ phân loại lớp phủ đất

Mô hình rừng phân loại được xây dựng với 200 cây quyết định; các tham số còn lại được giữ ở giá trị mặc định của GEE. Với mô hình cây phân loại tăng cường độ dốc, số lượng cây quyết định được sử dụng là 150 cây và tốc độ học là 0,1; các tham số khác của mô hình được lấy theo

giá trị mặc định của GEE. Kết quả đánh giá (xem Bảng 1) cho thấy cả hai mô hình đều đạt độ chính xác phân loại theo lớp phủ tương đối cao, trong đó các lớp không gian xanh, đất trồng trọt và thủy vực được nhận diện tốt nhất với tỷ lệ chính xác đều trên khoảng 93–97%. Độ chính xác

trung bình cho tất cả các lớp phủ của mô hình rừng phân loại (RF) đạt khoảng 92%, trong khi mô hình cây phân loại tăng cường độ dốc (GBM) cho giá trị cao hơn, khoảng 94%. Kết quả này cho thấy GBM có hiệu năng nhìn hơn RF trong bài toán phân loại lớp phủ tại khu vực nghiên cứu. Bản đồ lớp phủ xây dựng bởi GBM được cho trong Hình 4.

Kết quả phân loại được thể hiện trong Bảng 2 cho thấy khu vực xây dựng chiếm diện tích lớn

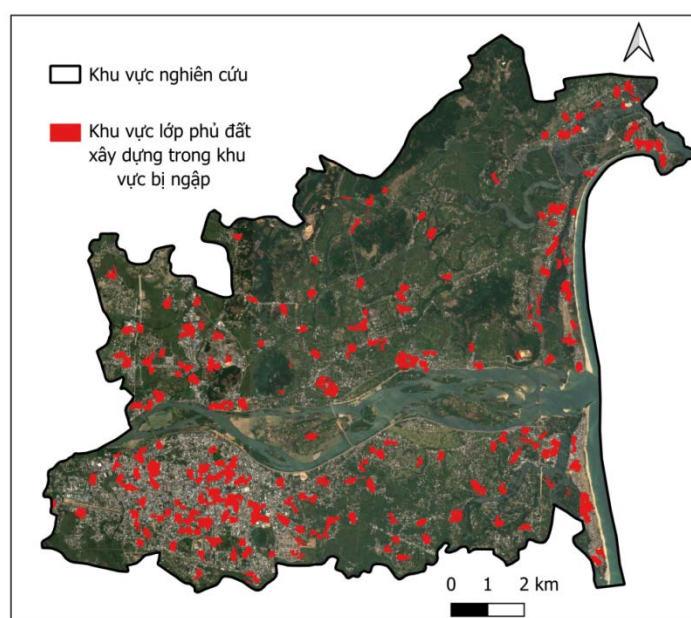
nhất, khoảng 53,98 km², tương ứng 33,71% tổng diện tích khu vực nghiên cứu, phản ánh mức độ đô thị hóa cao tại khu đô thị ven biển này. Đất trồng trọt đứng thứ hai với 46,19 km² (28,84%), cho thấy nông nghiệp vẫn là một loại hình sử dụng đất quan trọng xen kẽ với không gian đô thị. Không gian xanh chiếm 24,88 km², tương đương 15,53%, trong khi đất trống và thủy vực lần lượt chiếm 16,56 km² (10,34%) và 18,54 km² (11,58%).

Bảng 2. Diện tích và phân bố các loại lớp phủ đất

Loại lớp phủ đất	Diện tích (km ²)	Tỷ lệ (%)
Đất trống	16,56	10,34
Khu vực xây dựng	53,98	33,71
Không gian xanh	24,88	15,53
Đất trồng trọt	46,19	28,84
Thủy vực	18,54	11,58

Bảng 3. Diện tích lớp phủ đất trong khu vực bị ngập

Loại lớp phủ đất	Diện tích (km ²)	Tỷ lệ (%)
Không gian xanh	0,54	6,52
Đất trống	0,72	8,71
Đất trồng trọt	0,98	11,87
Khu vực xây dựng	6,02	72,90



Hình 5. Khu vực xây dựng bị ngập

Kết quả phân tích bản đồ lớp phủ đất với các khu vực bị ngập (xem Bảng 3) cho thấy khu vực xây dựng chiếm tới khoảng 72,9% diện tích trong vùng bị ngập, trong khi đất trồng trọt, đất trống và không gian xanh chỉ lần lượt chiếm khoảng 11,9%, 8,7% và 6,5%. Tỷ lệ này cho thấy phần lớn diện tích chịu ảnh hưởng trực tiếp của ngập lụt trong vùng được khảo sát tập trung ở các khu vực đô thị hóa với mật độ xây dựng cao, nơi bề mặt không thấm nước chiếm ưu thế và hệ thống thoát nước tự nhiên bị thu hẹp. Bản đồ trong Hình 5 minh họa sự phân bố không gian của các dải, cụm khu vực xây dựng bị ngập dọc theo bờ sông và trong khu vực trung tâm đô thị. Kết quả phân tích có thể giúp nhận diện các điểm nóng ngập lụt đô thị và mối liên hệ giữa mô hình sử dụng đất với rủi ro ngập. Những thông tin này có ý nghĩa quan trọng cho quy hoạch và quản lý đô thị, vì chúng cung cấp căn cứ không gian để đề xuất các giải pháp ưu tiên như kiểm soát mở rộng xây dựng vào vùng trũng, tăng cường không gian xanh và mặt nước trữ nước, nâng cấp hệ thống thoát nước, cũng như lồng ghép rủi ro ngập lụt vào quy hoạch sử dụng đất và phát triển đô thị thích ứng khí hậu trong tương lai.

Tuy vậy, trong nghiên cứu này, khu vực ngập lụt được trích xuất từ dữ liệu của Sentinel Asia (dựa trên ảnh vệ tinh Sentinel-1), trong đó các vùng có khả năng cao bị ngập được xác định gián tiếp từ sự thay đổi tín hiệu phản xạ của bề mặt nước sau trận mưa lớn. Bản đồ này mới chỉ chỉ ra các khu vực có khả năng cao bị ngập mà chưa được kiểm chứng một cách hệ thống bằng khảo sát thực địa. Vì vậy, các khu vực bị ngập thể hiện trên bản đồ vẫn tiềm ẩn những sai số nhất định. Sai số có thể phát sinh khi tín hiệu ra-đa bị nhiễu ở khu vực đô thị có nhiều công trình cao tầng hoặc bề mặt kim loại, hay trong các vùng có thảm thực vật rậm rạp che phủ mặt nước. Do đó, các kết quả phân tích dựa trên bản đồ khu vực bị ngập này cần được diễn giải với mức độ thận trọng phù hợp và cần được kiểm chứng bổ sung bằng các khảo sát thực địa. Thêm vào đó, các

phân tích trong bài báo hiện chỉ tập trung vào việc phân loại lớp phủ đất và đánh giá diện tích lớp phủ đất trong khu vực bị ngập. Nghiên cứu hiện tại chưa mở rộng mô hình hoặc kết hợp các kịch bản biến đổi khí hậu/thủy văn để phục vụ cho vấn đề dự báo ngập lụt cho tương lai.

4. Kết luận

Nghiên cứu này đã chứng minh tính hiệu quả của việc tích hợp dữ liệu viễn thám, các thuật toán học máy và nền tảng điện toán đám mây trong lập bản đồ hiện trạng lớp phủ đất và phân tích tác động ngập lụt tại một đô thị ven biển miền Trung. Bằng cách khai thác ảnh Sentinel-2 và áp dụng hai mô hình rừng phân loại và cây phân loại tăng cường, nghiên cứu đã xây dựng được bản đồ lớp phủ đất chi tiết cho khu vực đô thị được khảo sát. Kết quả cho thấy mô hình cây phân loại tăng cường cho kết quả tốt hơn và được sử dụng để mô tả phân bố của các loại hình sử dụng đất năm 2025. Việc kết hợp bản đồ lớp phủ đất với dữ liệu vùng ngập từ Sentinel Asia cho thấy phần lớn diện tích bị ngập nằm trong khu vực xây dựng, qua đó nhấn mạnh vai trò của quá trình đô thị hóa và gia tăng bề mặt không thấm nước trong việc phân tích rủi ro ngập lụt đô thị.

Về mặt ứng dụng, cách tiếp cận được trình bày trong bài báo có thể được mở rộng cho các đô thị ven biển khác nhằm cung cấp thông tin cập nhật, đáng tin cậy phục vụ quy hoạch sử dụng đất, xác định các điểm nóng ngập lụt và ưu tiên các giải pháp giảm thiểu rủi ro như kiểm soát mở rộng xây dựng, tăng cường không gian xanh và hạ tầng thoát nước. Trong tương lai, các nghiên cứu tiếp theo có thể kết hợp đồng thời dữ liệu Sentinel-1 (SAR) và Sentinel-2 để khai thác tốt hơn các dữ liệu viễn thám để phục vụ phân loại lớp phủ đất và nhận diện vùng ngập. Bên cạnh đó, việc tích hợp thêm dữ liệu quan trắc mưa, các mô hình thủy văn – thủy lực và bộ dữ liệu ngập lụt được kiểm chứng thực địa, cũng như thử nghiệm các mô hình học máy tiên tiến ngoài nền tảng GEE, sẽ góp phần hoàn thiện hơn

nửa độ chính xác và tin cậy của phương pháp hiện tại.

Tài liệu tham khảo

- [1] Darem, A.A.; Alhashmi, A.A.; Almadani, A.M.; Alanazi, A.K.; Sutantra, G.A. (2023). "Development of a map for land use and land cover classification of the Northern Border Region using remote sensing and GIS". *The Egyptian Journal of Remote Sensing and Space Science* (26), 341-350. DOI: 10.1016/j.ejrs.2023.04.005
- [2] Prokešová, R.; Horácková, Š.; Snopková, Z. (2022). "Surface runoff response to long-term land use changes: Spatial rearrangement of runoff-generating areas reveals a shift in flash flood drivers". *Science of The Total Environment* (815):151591. DOI:10.1016/j.scitotenv.2021.151591
- [3] Aniramu, O.; Iyanda, O.; Orimoogunje, O. (2026). "Remote sensing and GIS-based modelling of land use dynamics and urban flood risk in Lagos megacity for future flood mitigation". *Scientific Reports* (16), 1-21. DOI:10.1038/s41598-026-38544-1
- [4] Tu, V.T.; Ranzi, R. (2017). "Flood risk assessment and coping capacity of floods in central Vietnam". *Journal of Hydro-environment Research* (14), 44-60. DOI: 10.1016/j.jher.2016.06.001
- [5] Arpitha, M.; Ahmed, S.A.; Harishnaika, N. (2023). "Land use and land cover classification using machine learning algorithms in google earth engine". *Earth Science Informatics* (16), 3057-3073. DOI:10.1007/s12145-023-01073-w
- [6] Yang, L.; Driscoll, J.; Sarigai, S.; Wu, Q.; Chen, H.; Lippitt, C.D. (2022). "Google Earth Engine and Artificial Intelligence (AI): A Comprehensive Review". *Remote Sensing* (14), 3253, DOI: 10.3390/rs14143253
- [7] Avand, M; Moradi, H; lasboyee, MR. (2021). "Using machine learning models, remote sensing, and GIS to investigate the effects of changing climates and land uses on flood probability". *Journal of Hydrology*, (595), 125663. DOI: 10.1016/j.jhydrol.2020.125663
- [8] Islam, T.; Zeleke, E.B.; Afroz, M.; Melesse A.M. (2025). "A Systematic Review of Urban Flood Susceptibility Mapping: Remote Sensing, Machine Learning, and Other Modeling Approaches". *Remote Sensing* (17), 524. DOI: 10.3390/rs17030524
- [9] Srivastava, A.; Maity, R. (2023). "Assessing the Potential of AI-ML in Urban Climate Change Adaptation and Sustainable Development". *Sustainability* (15), 16461. DOI: 10.3390/su152316461
- [10] Breiman, L. (2001). "Random Forests". *Machine Learning* (45), 5-32. DOI: 10.1023/A:1010933404324
- [11] Friedman, J.H. (2001). "Greedy function approximation: A gradient boosting machine". *Annals of statistics* (29), 1189-1232. DOI: 10.1214/aos/1013203451.