

THEO DÕI BIẾN ĐỘNG SỬ DỤNG ĐẤT HUYỆN GIA LÂM, THÀNH PHỐ HÀ NỘI GIAI ĐOẠN 2018 - 2023

Phạm Anh Tuấn

Trường Đại học Tài nguyên và Môi trường Hà Nội

Tóm tắt

Nghiên cứu sử dụng thuật toán học máy Random Forest (RF) thực hiện việc phân loại, giám sát các lớp phủ bề mặt sử dụng đất tại khu vực Gia Lâm, Hà Nội giai đoạn 2018 - 2023. Kết quả biến động các loại hình sử dụng đất tại khu vực nghiên cứu được thể hiện rõ qua từng năm. Trong đó, tốc độ suy giảm trung bình hàng năm của lớp đất trống là 0,51 % ngược lại, tốc độ tăng trung bình của các khu vực dân cư đạt tới 1,80 % so với tổng diện tích tự nhiên của khu vực Gia Lâm.

Từ khóa: Học máy; Rừng ngẫu nhiên; Lớp phủ sử dụng đất; Giám sát.

Abstract

Monitoring land use changes in Gia Lam district, Hanoi during the period of 2018 - 2023

The study utilized the Random Forest (RF) machine learning algorithm to classify and monitor land use in the Gia Lam district, Hanoi, during 2018 - 2023. The results of land use changes in the study area are clearly demonstrated annually. The results study showed that the average annual decline rate of bare land changes is 0,51 %. In contrast, the average increase rate of built-up areas reaches 1,80 % compared to the total natural area of the Gia Lam area.

Keywords: Machine learning; Random Forest; Land use land cover; Monitoring.

Nhận bài: 15/5/2024; Phản biện xong: 12/6/2024; Duyệt đăng: 26/9/2024

Tác giả liên hệ, Email: patuan@hunre.edu.vn

DOI: <https://doi.org/10.63064/khtnmt.2024.604>

1. Mở đầu

Hiện nay, phương pháp khai thác và xử lý ảnh vệ tinh trực tuyến dựa trên hệ thống mạng Internet toàn cầu đã phát triển đem lại hiệu quả cao cho nhiều ứng dụng khác nhau, trong đó có việc ứng dụng để phân loại ảnh vệ tinh, theo dõi biến động sử dụng đất. Đồng hành với sự phát triển của công nghệ viễn thám, các phương pháp xây dựng thông tin phục vụ phân tích, đánh giá về quản lý đất đai, về đô thị hóa ngày càng có hiệu quả cao. Các thuật toán áp dụng trong phân loại lớp phủ sử dụng đất từ ảnh vệ tinh đem lại

kết quả có độ tin cậy cao, quá trình xử lý nhanh. Google Earth Engine (GEE) là nền tảng điện toán đám mây, có khả năng xử lý dữ liệu viễn thám online đáp ứng kịp thời các yêu cầu cung cấp thông tin, dữ liệu, phục vụ quá trình chuyển đổi số trong quản lý đất đai. Trước đây, đã có những nghiên cứu sử dụng một số thuật toán như Maximum Likelihood Classifier (MLC), Minimum Distance Classifier (MDC), K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest (RF) và Classification and regression tree (CART) [2, 3, 7] dựa trên

Nghiên cứu

các phần mềm thương mại chuyên dụng để phân loại lớp phủ bề mặt đất từ ảnh vệ tinh [8, 11]. Các hệ thống phần mềm này đã chứng minh được hiệu quả của mình, tuy nhiên kinh phí lớn và khả năng xử lý trực tuyến chưa thuận lợi. Ảnh vệ tinh trực tuyến, miễn phí là nguồn dữ liệu rất hiệu quả trong phân loại, theo dõi các biến động về đất đai. Trong đó có ảnh Landsat, Sentinel, Modis,... từ năm 2015 đến nay, ảnh Sentinel-2 là ảnh vệ tinh quang học với độ phân giải không gian cao (10 m) được cung cấp liên tục và miễn phí trên toàn cầu [12]. Đây là nguồn dữ liệu phù hợp cho nhóm nghiên cứu biến động sử dụng đất tại khu vực Gia Lâm. Trong nghiên cứu này, tác giả sử dụng thuật toán học máy (Machine Learning) là thuật toán Rừng ngẫu nhiên (Random Forest - RF) và ngôn ngữ lập trình Java Script để xây dựng chương trình tự động hóa việc phân loại, giám sát biến động các lớp phủ sử dụng đất dựa trên nền tảng dữ liệu điện toán đám mây GEE từ ảnh vệ tinh Sentinel-2.

Để hoàn thành mục tiêu và định hướng phát triển đến năm 2030 là xây dựng hệ thống quản lý đất đai hiện đại, đạt trình độ của các nước phát triển trong khu vực thì việc đánh giá công tác quản lý và sử dụng đất đai của các địa phương từ khi có Luật Đất đai 2013 đến nay và sắp tới thực hiện Luật Đất đai mới năm 2024 là rất quan trọng, giúp cơ quan quản lý Nhà nước đánh giá tổng quan tình hình quản lý, sử dụng đất trong thời gian qua. Từ đó, hoạch định các chính sách quản lý trong thời gian tới, để góp phần nâng cao hiệu lực, hiệu quả quản lý Nhà nước về đất đai nói chung và tình hình biến động sử dụng đất nói riêng. Về mặt địa lý, huyện Gia Lâm được phân ra làm ba khu vực, ngăn

cách bởi dòng Sông Đuống bao gồm cụm Bắc Đuống, cụm Nam Đuống, cụm Sông Hồng. Với các dự án đô thị, chính sách và chiến lược phát triển hạ tầng, kinh tế và xã hội của huyện Gia Lâm sẽ có tác động không nhỏ đến biến động sử dụng đất.

2. Dữ liệu và khu vực nghiên cứu

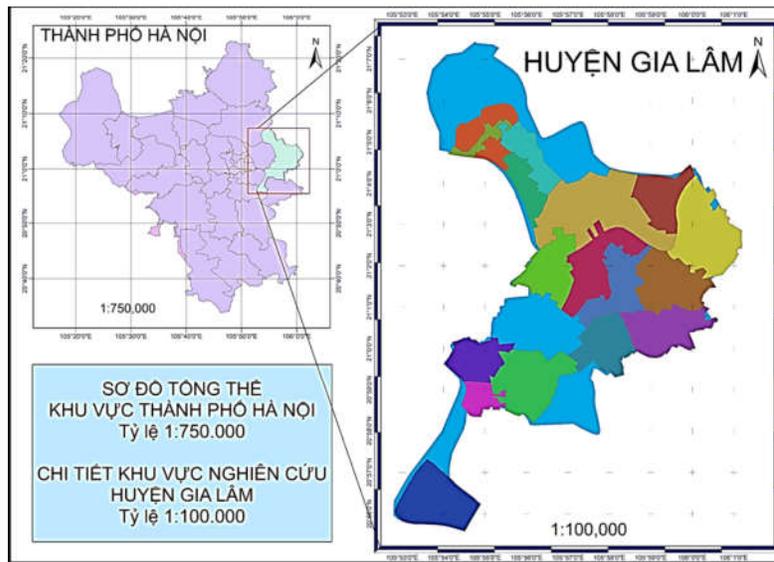
2.1. Khu vực nghiên cứu

Huyện Gia Lâm nằm ở phía Đông Hà Nội, cách trung tâm thành phố Hà Nội khoảng 12 km. Về mặt địa lý, huyện Gia Lâm được phân ra làm ba khu vực, ngăn cách bởi dòng Sông Đuống bao gồm cụm Bắc Đuống, cụm Nam Đuống, cụm Sông Hồng, với diện tích 116,64 km². Dân số năm 2022 là 309.353 người, mật độ dân số đạt 2.652 người/km² [6]. Gia Lâm có địa hình bằng phẳng, khí hậu chia làm hai mùa rõ rệt trong năm. Mùa nóng ẩm từ tháng 4 đến tháng 10, mùa khô hanh từ tháng 11 đến tháng 3 năm sau, giữa hai mùa nóng ẩm và mùa khô hanh có các thời kỳ chuyển tiếp tạo ra dạng khí hậu bốn mùa xuân, hạ, thu, đông. Nhiệt độ không khí trung bình năm 23,5 °C, mùa nóng khoảng 27,4 °C. Lượng mưa từ 1.400 mm đến 1.600 mm, tập trung nhiều vào tháng 7 và tháng 8 hàng năm, độ ẩm tương đối trung bình hàng năm là 80 %. Với điều kiện thời tiết này, có thể thấy rằng dữ liệu ảnh vệ tinh quang học thu nhận trong khoảng tháng 5, tháng 6 hàng năm thường là có chất lượng rõ ràng nhất, đảm bảo chất lượng hình ảnh, thông tin cho việc khai thác, phân loại các lớp phủ sử dụng đất tại khu vực nghiên cứu.

Gần đây, Gia Lâm phát triển nhanh chóng, các khu đô thị mới được hình thành như khu đô thị Đặng Xá, khu đô thị Vinhomes Ocean Park, MasteriseHome

WaterFront Park, Handhomes BlueStar, Highway5 Residence, EuroWindow Twin Park, Gia Lâm Metropolitan, Oasis Cổ Bi, khu 31 ha Trâu Quỳ, Ninh Hiệp,... Điều này giúp tăng đáng kể dân số của huyện cũng như tỉ lệ dân thành thị. Ngày 22/9/2023, tại kỳ họp thứ mười ba, Hội đồng nhân dân thành phố Hà Nội đã biểu quyết thông qua Nghị quyết về việc tán thành chủ trương thành lập quận Gia Lâm và các phường

thuộc quận Gia Lâm. Theo đó, thành lập quận Gia Lâm trên cơ sở nguyên trạng diện tích tự nhiên và quy mô dân số của huyện Gia Lâm và thành lập 16 phường trên cơ sở 20 xã, 02 thị trấn của huyện Gia Lâm [5]. Với các dự án đô thị, chính sách và chiến lược phát triển hạ tầng, kinh tế và xã hội của huyện Gia Lâm như trên sẽ có tác động không nhỏ đến biến động sử dụng đất trên địa bàn nghiên cứu.



Hình 1: Vị trí khu vực nghiên cứu

2.2. Dữ liệu ảnh vệ tinh sử dụng trong nghiên cứu

Trong khu vực có một số dữ liệu ảnh vệ tinh có thể được khai thác miễn phí như ảnh Landsat, Modis, Sentinel, ... Trong đó, dữ liệu ảnh vệ tinh Sentinel-2 được khai thác trực tuyến trên nền tảng GEE và sử dụng trong nghiên cứu này. Đây là ảnh có sự phù hợp nhất với chu kỳ chụp lặp lại là 5 ngày và có độ phân giải không gian cao, đạt tới 10 m. Kết quả đã lựa chọn ảnh Sentinel-2 trong giai đoạn 2018 đến 2023. Thời gian ảnh vệ tinh quang học Sentinel-2 được thu nhận là các ảnh tổng hợp trong tháng 6 hàng năm, riêng năm 2018 có ảnh thu nhận trong tháng 7 (do

tháng 6 ảnh có độ phủ mây lớn) các ảnh đảm bảo điều kiện có chất lượng rõ ràng, độ phủ mây thấp nhất. Các thông tin cơ bản về dữ liệu ảnh vệ tinh được trình bày tại Bảng 1.

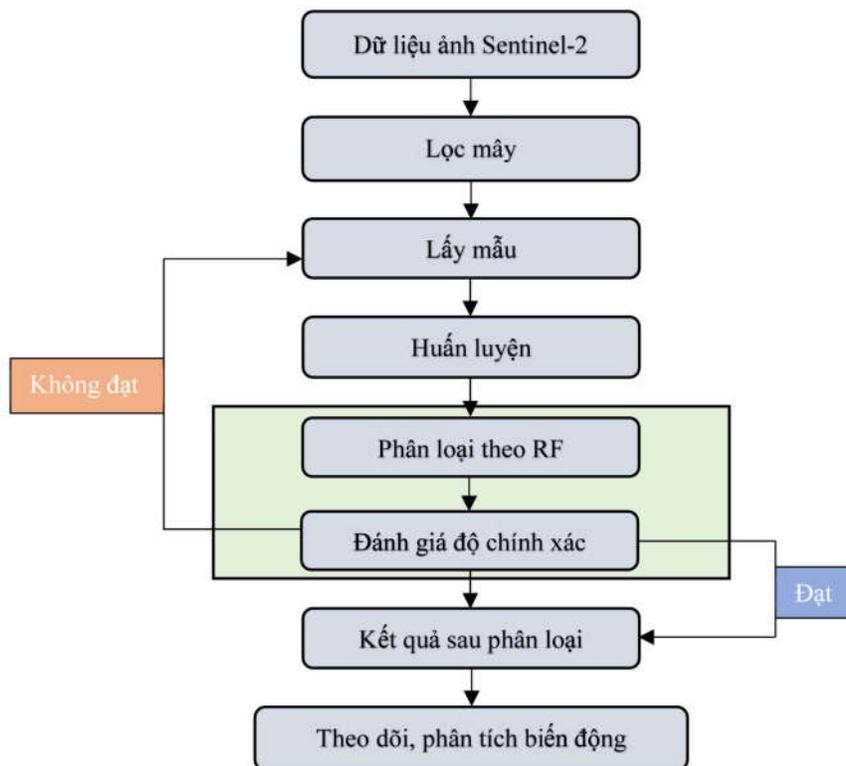
Bảng 1. Thông tin cơ bản về dữ liệu ảnh vệ tinh trong khu vực nghiên cứu

| Thời gian thu nhận ảnh | Loại ảnh | Độ phủ mây |
|------------------------|------------|------------|
| 07 - 2018 | Sentinel-2 | 1,85 % |
| 06 - 2019 | Sentinel-2 | 0,12 % |
| 06 - 2020 | Sentinel-2 | 1,53 % |
| 06 - 2021 | Sentinel-2 | 0,73 % |
| 06 - 2022 | Sentinel-2 | 1,76 % |
| 06 - 2023 | Sentinel-2 | 1,23 % |

3. Phương pháp nghiên cứu

Dựa trên dữ liệu ảnh vệ tinh được cung cấp miễn phí thông qua nền tảng GEE, nghiên cứu này sử dụng các thuật toán RF để tính toán, phân loại lớp phủ bề mặt cho dữ liệu đầu vào là ảnh vệ tinh Sentinel-2. Dữ liệu ảnh vệ tinh Sentinel-2 sử dụng trong nghiên cứu là các ảnh đã được hiệu chỉnh khí quyển, các ảnh có lượng mây che phủ thấp (Bảng 1), được hiệu chỉnh hình học trên hệ quy chiếu WGS84. Việc phân loại ảnh theo được thực hiện trực tuyến trên nền tảng dữ liệu điện toán đám mây GEE. Nghiên cứu lựa chọn phân loại ảnh theo 5 lớp phủ bề mặt tương đương với các lớp: 1) Lớp phủ Đất trống, 2) Lớp phủ Mặt nước, 3) Lớp phủ Cây lâu năm, 4) Lớp phủ Cây hàng năm,

5) Lớp phủ Dân cư. Việc lấy mẫu để huấn luyện và thực hiện phân loại được lựa chọn cho từng lớp. Trong đó, đối với lớp phủ Đất trống, các điểm lấy mẫu phân loại được chọn là các khu vực như bãi đất, cát, bãi bồi, khu vực đang san lấp, các khu vực đất trống mới san lấp chuẩn bị xây dựng. Lớp phủ Mặt nước gồm các khu vực ao hồ, sông suối, kênh mương. Lớp phủ Cây lâu năm bao gồm những khu vực có cây cối lớn, dày đặc, lâu năm. Lớp phủ Cây hàng năm gồm khu vực như hoa màu, bãi cỏ, ruộng vườn tại thời điểm thu nhận ảnh đang có cây trồng hàng năm. Lớp phủ Dân cư gồm nhà cửa, công trình xây dựng độc lập, và các khu vực nhà cửa, công trình xây dựng tại làng mạc, khu chung cư, các tòa nhà cao tầng.



Hình 2: Sơ đồ quy trình phân loại ảnh

Các mẫu được lấy cụ thể cho từng loại lớp phủ phục vụ việc ghi nhận các thông tin về các mẫu cho từng loại lớp

phủ và huấn luyện cho máy học. Tổng số lượng mẫu trên ảnh cho 5 lớp phủ bao gồm 419 mẫu, chia tương đối đều về số

lượng cho từng lớp phủ. Vị trí các mẫu huấn luyện được bố trí phân bố đều theo đặc điểm của từng lớp phủ. Kết thúc quá trình huấn luyện, tiến hành phân loại các lớp phủ trong khu vực nghiên cứu theo thuật toán RF. Các bước phân loại được tiến hành theo như hình minh họa (Hình 2). Trong đó bao gồm: Thu thập dữ liệu ảnh Sentinel-2 dựa trên nền tảng dữ liệu điện toán đám mây GEE; Lọc và lựa chọn ảnh có độ phủ mây là thấp nhất; Lấy mẫu theo các vị trí để phục vụ phân loại; Tạo mẫu và huấn luyện máy học. Phân loại các lớp phủ trên ảnh vệ tinh theo các thuật toán RF; Thu nhận kết quả ảnh phân loại theo thuật toán RF; Đánh giá độ chính xác của kết quả phân loại ảnh. Sau khi có kết quả phân loại ảnh đảm bảo độ chính xác theo yêu cầu, tiến hành các bước phân tích, theo dõi biến động các lớp phủ sử dụng đất trong giai đoạn 2018 - 2023 tại khu vực nghiên cứu.

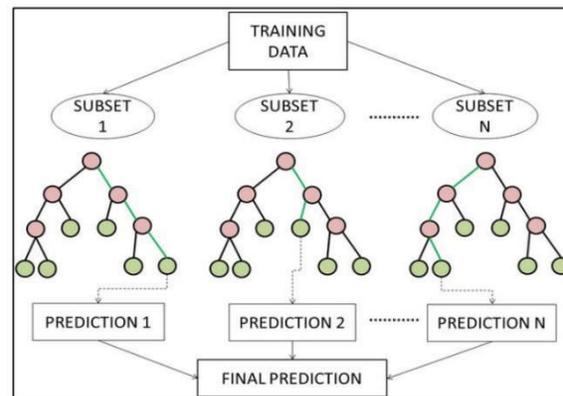
3.1. Thuật toán RF

RF là một thuật toán máy học có thể tích hợp nhiều cây quyết định và sau đó tạo thành một khu rừng. Thuật toán này kết hợp các tính năng ngẫu nhiên để tạo ra một cây. Phương pháp đóng bao được sử dụng để tạo các mẫu huấn luyện và mỗi tính năng đã chọn được rút ngẫu nhiên bằng cách thay thế N (kích thước của tập huấn luyện ban đầu). Sau đó, kết quả dự đoán cuối cùng thu được bằng cách kết hợp nhiều cây quyết định [1]. Công thức dưới đây thực hiện quyết định phân loại cuối cùng như sau:

$$H(x) = \operatorname{argmax}_Y \sum_{i=1}^k I(h_i(x) = Y) \quad (1)$$

Trong đó: $H(x)$ là mô hình kết hợp; h_i là mô hình phân loại của cây quyết định

đơn; k hệ số đàn hồi; Y là biến đầu ra (hoặc biến mục tiêu); $I(\cdot)$ là hàm chỉ báo. Công thức cho thấy rằng thuật toán RF sử dụng đa số các quyết định biểu quyết để xác định phân loại cuối cùng. Tham số điều chỉnh của thuật toán RF là số lượng cây và số lượng cây có thể được chọn theo kinh nghiệm của người xử lý tính toán. Trong các bài toán phân lớp dữ liệu thì thuật toán RF được sử dụng tương đối phổ biến. Thuật toán RF được đánh giá cao bởi tính chính xác của mô hình [10]. Nhược điểm chính của thuật toán RF là khối lượng tính toán lớn, tuy nhiên thời gian xử lý dữ liệu được đánh giá là nhanh chóng [4, 9]. Hình 3 minh họa mô hình phân loại theo thuật toán RF.



Hình 3: Mô hình phân loại theo thuật toán RF

3.2. Phương pháp đánh giá độ chính xác

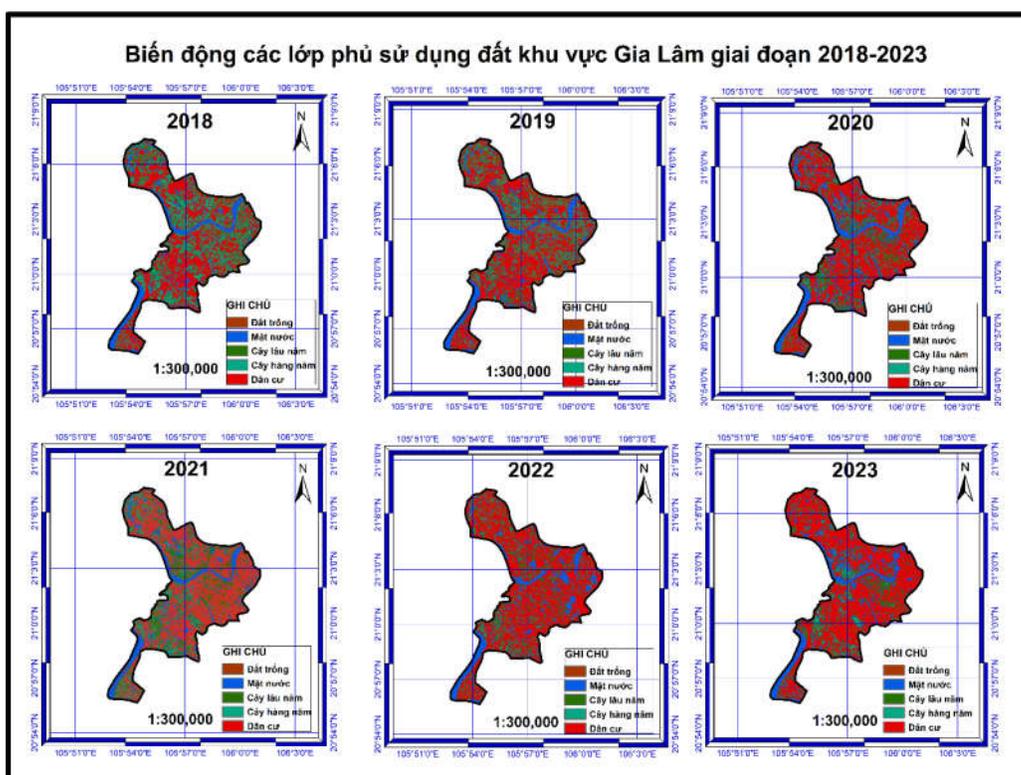
Ma trận nhầm lẫn (Confusion Matrix) được sử dụng để đánh giá độ chính xác phân loại các lớp phủ bề mặt trên ảnh. Ma trận nhầm lẫn là phương pháp quan trọng và phổ biến được sử dụng để đánh giá độ chính xác, có thể mô tả độ chính xác của phân loại và chỉ ra sự nhầm lẫn giữa các lớp đối tượng. Các thống kê cơ bản cho ma trận nhầm lẫn có đánh giá về độ chính xác tổng thể (Overall Accuracy - OA) và hệ số Kappa. Trong đó, hệ số Kappa có

Nghiên cứu

giá trị từ 0,4 đến 0,6 được đánh giá là đạt kết quả trung bình, giá trị từ lớn hơn 0,6 đến 0,8 là tốt và hơn 0,8 đến 1,0 là rất tốt [10]. Việc thực hiện đánh giá độ chính xác của ảnh phân loại được tiến hành trên GEE. Trong đó, sử dụng 70 % số lượng mẫu dùng để phân loại ảnh và 30 % số lượng mẫu dùng để kiểm tra đánh giá. Cụ thể, có 294 mẫu dùng để phân loại các lớp phủ và 125 mẫu phục vụ đánh giá độ chính xác phân loại ảnh.

4. Kết quả nghiên cứu và thảo luận

Sau khi xử lý dữ liệu, phân loại ảnh, kết quả nghiên cứu có được là sản phẩm ảnh sau khi phân loại theo thuật toán RF. Mỗi một ảnh sau phân loại bao gồm 5 loại lớp phủ sử dụng đất bao gồm lớp phủ Đất trống, lớp phủ Mặt nước, lớp phủ Cây lâu năm, lớp phủ Cây hàng năm, lớp phủ Dân cư. Hình 4 minh họa sản phẩm ảnh sau phân loại của khu vực nghiên cứu theo thuật toán RF.



Hình 4: Các lớp phủ sử dụng đất huyện Gia Lâm, Hà Nội giai đoạn 2018 - 2023

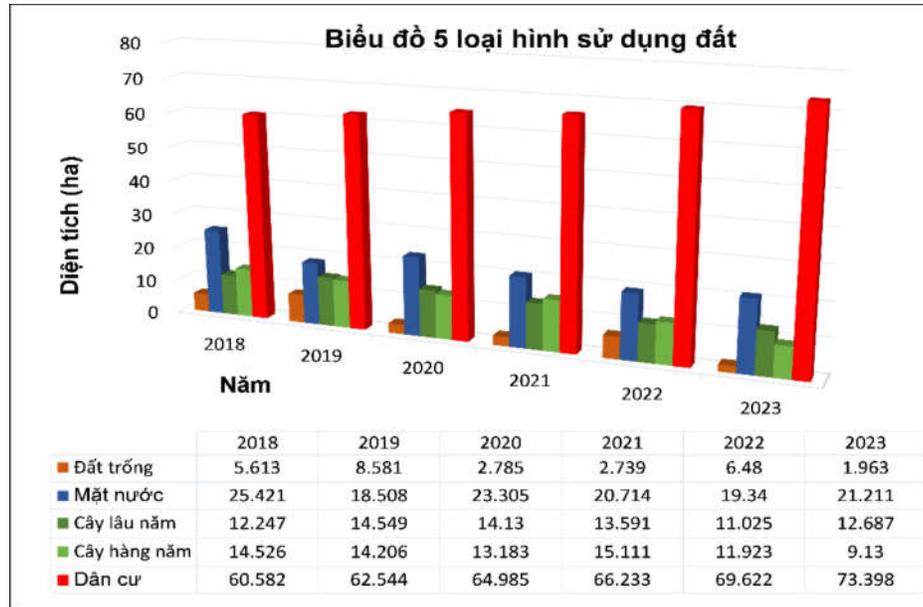
Ảnh sau phân loại được thể hiện rõ ràng, trong đó các lớp phủ chiếm đa số là lớp phủ Dân cư. Lớp phủ này có độ che phủ từ khoảng 51,17 - 62 % của diện tích tự nhiên toàn bộ khu vực nghiên cứu. Lớp phủ có tỷ lệ ít nhất là lớp Đất trống, có độ bao phủ từ khoảng 1,66 - 4,74 %. Kết quả của phân bố các lớp phủ như trên thể hiện sự phù hợp với đặc điểm tự nhiên, phân bố dân cư và tình hình phát

triển kinh tế xã hội tại huyện Gia Lâm, Hà Nội.

Kết quả giám sát biến động lớp phủ sử dụng đất giai đoạn 2018 - 2023 tại khu vực Gia Lâm, Hà Nội cho thấy rằng, các lớp phủ Mặt nước và Cây lâu năm có ít biến động hơn so với các lớp phủ còn lại. Trong đó, lớp phủ Mặt nước có xu hướng giảm khoảng 0,59 % năm, lớp phủ Cây lâu năm có xu hướng tăng nhẹ khoảng

0,06 % một năm. Riêng đối với lớp phủ Đất trống tại khu vực nghiên cứu có sự biến động không có quy luật rõ ràng giữa các năm, năm cao nhất có độ che phủ 7,25 % (2019) và năm thấp nhất là 1,66 % (2023), tính trung bình là có độ suy giảm diện tích trong giai đoạn 2018 đến 2023 là 0,51 % một năm. Lớp phủ Cây hàng năm có độ suy giảm trung bình năm trong giai

đoạn nghiên cứu là 0,76 %. Nguyên nhân sự tăng giảm trên ngoài việc do tình trạng sử dụng đất thực tế mỗi năm một khác, còn có thể do tại các thời điểm thu nhận ảnh, chất lượng ảnh, cũng tạo nên những tính toán có thể tồn tại sai lệch trong phân loại. Hình 5 thể hiện biểu đồ và số liệu về biến động các lớp phủ sử dụng đất tại huyện Gia Lâm, Hà Nội.



Hình 5: Các lớp phủ sử dụng đất huyện Gia Lâm, Hà Nội giai đoạn 2018 - 2023

Đối với lớp phủ Dân cư, kết quả cho thấy có sự gia tăng liên tục về diện tích trong suốt giai đoạn nghiên cứu. Cụ thể, diện tích của lớp phủ Dân cư tăng từ 60,58 ha (2018) đến 73,40 ha (2023), trung bình đã tăng 1,80 % mỗi năm so với tổng diện tích tự nhiên của khu vực Gia Lâm, Hà Nội. Như vậy, có thể thấy rằng trong thời gian 5 năm phát triển, tốc độ phát triển của Gia Lâm là tương đối cao rất rõ ràng bởi sự phát triển, mở rộng diện tích nhanh chóng của các khu vực khu dân cư, các khu chung cư, các khu vực nhà cửa thương mại,... Ngược lại, các khu vực sử dụng đất khác hầu hết có diện tích suy giảm hàng năm.

Kết quả các độ chính xác sau phân loại được tính toán từ ma trận nhầm lẫn bao gồm các thông tin bản là độ chính xác tổng thể (OA) và hệ số Kappa được thể hiện trong bảng 2. Trong đó, độ chính xác tổng thể OA thấp nhất là 0,92 và cao nhất là 0,94. Đồng thời, hệ số Kappa cũng đạt từ 0,83 đến 0,89. Kết quả nghiên cứu cho thấy độ chính xác ảnh sau phân loại là rất cao nên kết quả này đảm bảo độ tin cậy để đưa ra các thống kê, phân tích về các lớp phủ sử dụng đất trong giai đoạn 2018 - 2023 tại khu vực Gia Lâm, Hà Nội.

Bảng 2. Độ chính xác phân loại ảnh

| | OA | Kappa |
|-------------|-----------|--------------|
| 2018 | 0,94 | 0,88 |
| 2019 | 0,92 | 0,85 |
| 2020 | 0,92 | 0,83 |
| 2021 | 0,92 | 0,84 |
| 2022 | 0,93 | 0,87 |
| 2023 | 0,94 | 0,89 |

Ngoài chất lượng của các mẫu huấn luyện, độ chính xác phân loại ảnh còn phụ thuộc vào chất lượng ảnh thu nhận. Trong nghiên cứu này, chất lượng mẫu huấn luyện và chất lượng ảnh thu nhận qua các năm là tương đối đồng đều, độ phủ mây rất thấp, ảnh có hình ảnh và các thông tin vật lý rõ ràng nên theo kết quả đánh giá tất cả các ảnh sau phân loại từ 2018 đến 2023 đều có độ tin cậy cao.

5. Kết luận

Phương pháp nghiên cứu đã đưa ra các kết quả đảm bảo độ tin cậy và nhanh chóng hơn so với các phương pháp truyền thống. Kết quả cũng cho thấy rõ về sự phát triển mở rộng nhanh chóng về diện tích của các khu vực dân cư, đô thị trong huyện Gia Lâm, Hà Nội giai đoạn 2018 - 2023. Cụ thể, kết quả phân loại ảnh vệ tinh Sentinel-2 đã cho thấy tốc độ mở rộng diện tích của các khu vực dân cư, đô thị đạt 1,80 % (tương ứng 2,13 ha) hàng năm. Trong khi đó lớp phủ Đất trống và lớp phủ Cây hàng năm có độ suy giảm trung bình năm là 0,51 % và 0,76 % tương ứng.

Ngoài ra, kết quả nghiên cứu này đã phân tích và giám sát biến động một số lớp phủ sử dụng đất, đưa ra các kết quả và xu hướng tăng, giảm diện tích của các lớp phủ sử dụng đất tại khu vực nghiên cứu. Các kết quả cũng có thể sử dụng làm số liệu tham khảo, tính toán, dự báo kết quả

tăng giảm các lớp phủ sử dụng đất cho các giai đoạn tiếp theo tại khu vực Gia Lâm, Hà Nội.

Lời cảm ơn: Bài báo được thực hiện dưới sự hỗ trợ nghiên cứu của đề tài “Nghiên cứu ứng dụng trí tuệ nhân tạo phân loại và dự báo biến động lớp phủ đất xây dựng khu vực đô thị trên địa bàn huyện Gia Lâm, thành phố Hà Nội”. Mã số: 13.01.24.M.06 do Trường Đại học Tài nguyên và Môi trường Hà Nội chủ trì.

TÀI LIỆU THAM KHẢO

[1]. David Richard Cutler, Thomas C Edwards, Karen H. Beard, Adele Cutler (2007). *Random forests for classification in ecology*. Ecology. 88, 11, pp. 2783 - 2792.

[2]. Thanh Tung Dang, Dinh Thi Thanh Huyen, Hoang Thi Thuy, Ta Minh Ngoc (2023). *Application of satellite images and artificial intelligence to monitor land cover changes in Hanoi area during 2013 - 2023 period*. VNU Journal of Science: Earth Environmental Sciences. 39, 4.

[3]. Ujjwal Maulik, Debasis Chakraborty (2017). *Remote Sensing Image Classification: A survey of support-vector-machine-based advanced techniques*. IEEE Geoscience Remote Sensing Magazine. 5, 1, pp. 33 - 52.

[4]. Nguyễn Hà Nam (2009). *Tối ưu hóa KPCA bằng GA để chọn các thuộc tính đặc trưng nhằm tăng hiệu quả phân lớp của thuật toán Random Forest*. Tạp chí Khoa học Tự nhiên và Công nghệ, Đại học Quốc gia Hà Nội. 25, pp. 84 - 93

[5]. Hội đồng nhân dân thành phố Hà Nội (2023). *Nghị quyết về việc tán thành chủ trương thành lập quận Gia Lâm và các phường thuộc quận Gia Lâm*. Ngày 22/9/2023, tại kỳ họp thứ mười ba, Hội đồng nhân dân thành phố Hà Nội.

[6]. Cổng Thông tin điện tử Chính phủ (2024). *Huyện Gia Lâm*. Available: <https://gis.chinhphu.vn/?r=PzXhopJL0mVbpY2jes40A>

- [7]. Đỗ Lan Phương, Hoàng Thị Thùy, Đỗ Như Hiệp (2024). *Nghiên cứu ứng dụng trí tuệ nhân tạo theo dõi biến động lớp phủ sử dụng đất khu vực quận Bắc Từ Liêm Hà Nội giai đoạn 2019 - 2023*. Tạp chí Trắc địa - Bản đồ. Tập 10, Số 1, trang 22 - 29.
- [8]. Sophia Rwanga, Julius Ndambukil (2017). *Accuracy assessment of land use/land cover classification using remote sensing and GIS*. International Journal of Geosciences. 8, 04, 611.
- [9]. Đặng Thanh Tùng (2020). *Khai thác trực tuyến cơ sở dữ liệu ảnh vệ tinh, so sánh thuật toán học máy về phân loại lớp phủ trên nền Google Earth Engine*. Hội thảo Quốc gia: Giải pháp kết nối chia sẻ dữ liệu Tài nguyên và Môi trường, Hà Nội, 2020.
- [10]. Sofia Visa, Brian Ramsay, Anca Ralescu, Esther van der Knaap (2011). *Confusion matrix-based feature selection*. Maics. 710, 1, pp. 120 - 127.
- [11]. Nigatu Wondrade, Øystein B. Dick, Håvard Tveite (2014). *GIS based mapping of land cover changes utilizing multi-temporal remotely sensed image data in Lake Hawassa Watershed, Ethiopia*. Environmental monitoring assessment. 186, pp. 1765 - 1780.
- [12]. Yan Zhao, Andries B Potgieter, Miao Zhang, Bingfang Wu, Graeme L Hammer (2020). *Predicting wheat yield at the field scale by combining high-resolution Sentinel-2 satellite imagery and crop modelling*. Remote Sensing. 12, 6, 1024.