

Ứng dụng thuật toán học máy theo dõi lớp phủ mặt nước phục vụ đào tạo, nghiên cứu trong lĩnh vực quản lý đất đai

Võ Ngọc Hải*

* Trường Đại học Tài nguyên và Môi trường Hà Nội

Received: 12/12/2023; Accepted: 20/12/2023; Published: 28/12/2023

Abstract: Machine learning algorithms have effectively contributed to monitoring changes in land use land cover, including water bodies. This study uses the Classification and regression tree (Cart) algorithm to classify land cover and monitor water bodies layer changes in Ba Vi district, Hanoi from 2018 to 2023. Results of the study has shown a decline in water bodies in the last 5 years, the total water surface area has decreased by about 1% compared to the total natural area.

Keywords: Dynamic World, World Cover, Land use, Land cover.

1. Đặt vấn đề

Việc xác định và theo dõi biến động diện tích mặt nước bằng ảnh vệ tinh có thể dùng chỉ số như NDWI, MNDWI, WNDWI để phát hiện các pixel là mặt nước và các pixel không phải là mặt nước [1, 2]. Ngoài ra cũng có thể sử dụng các phương pháp phân loại các lớp phủ bề mặt, trong đó có lớp phủ là mặt nước, phương pháp này cũng đã được chứng minh về độ tin cậy và tính hiệu quả trong việc theo dõi các biến động lớp phủ bề mặt [3].

Hiện nay, việc sử dụng nền tảng điện toán đám mây của Google Earth Engine (GEE) 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 theo dõi, giám sát biến động lớp phủ bề mặt, trong đó có việc theo dõi biến động lớp phủ mặt nước một cách liên tục theo thời gian gần thực. 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) [4] dựa trên các phần mềm thương mại để phân loại lớp phủ bề mặt đất từ ảnh vệ tinh. 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. Trong nghiên cứu này, nhóm nghiên cứu sử dụng mô hình trí tuệ nhân tạo (AI - Artificial Intelligence) với thuật toán học máy (Machine Learning) Classification and regression tree (Cart) để viết code chương trình tự động hóa việc phân loại, giám sát biến động các lớp phủ mặt nước. Các dữ liệu ảnh vệ tinh và các kết quả phân tích được thực hiện trên nền tảng điện toán đám mây GEE với ngôn ngữ lập trình Javascript.

2. Nội dung nghiên cứu

2.1. Dữ liệu và khu vực nghiên cứu

Từ khi Hà Nội mở rộng cho đến nay đã có nhiều cơ sở hạ tầng được bổ sung, phát triển, đã có những biến động rõ rệt về các loại hình sử dụng đất, góp phần vào sự phát triển kinh tế, xã hội của khu vực tuy nhiên qua đó cũng phần nào gây suy giảm diện tích mặt nước, ảnh hưởng đến môi trường chung của thủ đô. Khu vực Ba Vi cũng chịu ảnh hưởng nhất định theo xu hướng biến động chung của vùng thủ đô.

Dữ liệu nghiên cứu sử dụng tư liệu ảnh vệ tinh Sentinel 2 từ năm 2018 đến 2023 được khai thác trực tuyến trên nền tảng GEE. Các dữ liệu ảnh vệ tinh được lựa chọn đảm bảo có chất lượng rõ ràng, độ phủ mây thấp nhất. Kết quả đã lựa chọn ảnh Sentinel 2 là dữ liệu tổng hợp của các ảnh có độ phủ mây thấp nhất trong khoảng thời gian từ tháng 3 tới tháng 6 hàng năm. Là thời gian cuối mùa xuân và đầu mùa hè, thời tiết chuyển mùa từ mưa phùn sương mù sang thời tiết quang đãng hơn, ít có mưa là điều kiện thuận lợi để có được chất lượng ảnh viễn thám quang học tốt. Trong đó ảnh Sentinel 2 có độ phân giải không gian 10m có che phủ mây rất thấp, đảm bảo chất lượng để phân tích, tính toán kết quả..

2.2. Phương pháp nghiên cứu

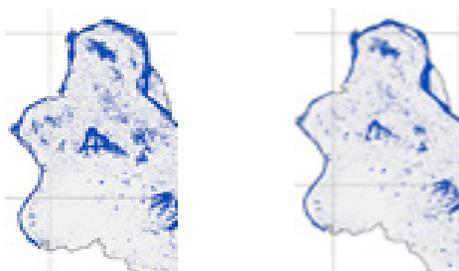
Nghiên cứu này tiến hành phân loại ảnh vệ tinh bằng ngôn ngữ lập trình JavaScript trên nền tảng dữ liệu điện toán đám mây GEE. Trong đó, thuật toán học máy Cart được dùng để tính toán, phân loại lớp phủ bề mặt cho dữ liệu đầu vào là ảnh vệ tinh Landsat 8, Sentinel 2. Trong nghiên cứu tại khu vực này, tác giả lựa chọn 4 lớp cơ bản để phân loại lần lượt bao gồm: 1) lớp phủ Đất trống, 2) lớp phủ Mặt nước, 3) lớp phủ Rừng, 4) lớp phủ cây nông nghiệp 5) lớp phủ Công trình xây dựng, 6) lớp phủ Giao thông cơ

bản. Các lớp phủ được lựa chọn để phân loại là hiện trạng tại thời điểm thu nhận ảnh, trong đó với lớp 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 Rừng vật bao gồm những khu vực có cây cối, dày đặc, rừng. Lớp phủ Cây nông nghiệp bao gồm bãi cỏ, ruộng vườn tại thời điểm thu nhận ảnh đang có cây trồng. Lớp phủ Công trình xây dựng 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. Lớp phủ Giao thông cơ bản gồm các loại đường lớn, đường chính trong khu vực nghiên cứu.

Thuật toán CART được sử dụng trong nghiên cứu là một thuật toán học máy có giám sát trong hệ thống phân loại dựa trên cây quyết định (Decision tree) và sử dụng các mẫu huấn luyện để xác định, nhận dạng, phân loại đối tượng trên ảnh viễn thám CART được sử dụng rộng rãi để phân loại viễn thám, nó còn được gọi là cây phân loại và hồi quy [3,4]; Thuật toán RF là một thuật toán học tích hợp có thể tích hợp nhiều cây quyết định và sau đó tạo thành một khu rừng. Thuật toán kết hợp các tính năng ngẫu nhiên để tạo ra một cây.

2.3. Kết quả nghiên cứu và thảo luận

Kết quả sau phân loại bao gồm 6 lớp phủ cơ bản là lớp phủ Đất trống, lớp phủ Mặt nước, lớp phủ Rừng, lớp phủ Nông nghiệp, lớp phủ Công trình xây dựng, lớp phủ Giao thông. Trong nghiên cứu này, lớp phủ mặt nước được sử dụng để theo dõi biến động trong giai đoạn 2018-2023.

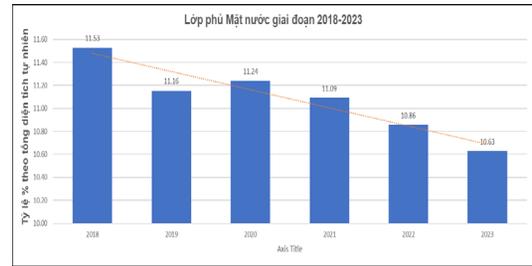


a. Mặt nước 2018

b. Mặt nước 2023

Hình 2.1. Các sản phẩm lớp phủ mặt nước được phân loại với thuật toán Cart

Kết quả giám sát biến động lớp phủ mặt nước giai đoạn 2018-2023 tại khu vực Ba Vì cho thấy rằng có sự suy giảm diện tích mặt nước qua giai đoạn 05 năm nghiên cứu. Lớp phủ Mặt nước có diện tích giảm khoảng 1% so với tổng diện tích tự nhiên của khu vực.



Hình 2.2. Kết quả biến động diện tích mặt nước tại khu vực nghiên cứu

Tổng thể, từ năm 2018 đến 2023, xu hướng giảm diện tích của lớp phủ là Mặt nước là rõ ràng. Theo kết quả phân tích, trung bình hàng năm trong giai đoạn này, tốc độ suy giảm diện tích của lớp Mặt nước trong giai đoạn từ 2018 đến 2018 giảm trung bình khoảng gần 0.2%/năm.

3. Kết luận và đề xuất

Phương pháp nghiên cứu đã đưa ra các kết quả một cách nhanh chóng, đảm bảo độ tin cậy và các dữ liệu phân tích thường là có chất lượng cao, đảm bảo yêu cầu phân tích dữ liệu, thời gian thu nhận dữ liệu ảnh vệ tinh được lựa chọn tùy ý để phù hợp yêu cầu nghiên cứu. Kết quả cũng cho thấy rõ ràng về tốc độ suy giảm lớp phủ mặt nước là gần 0.2%/năm, sau 10 năm theo dõi từ 2018 đến 2023, diện tích của lớp phủ mặt nước giảm khoảng 1%.

Các kết quả phân loại ảnh đã được đánh giá và thể hiện được độ tin cậy, tuy nhiên để có những căn cứ chắc chắn hơn nữa, các nghiên cứu sau này nên có sự so sánh với các loại bản đồ như là bản đồ hiện trạng sử dụng đất qua các thời kỳ và các số liệu đo đạc phân tích về mặt nước tại thực địa.

Tài liệu tham khảo

[1] Phạm Văn Chiến, “Nghiên cứu các chỉ số nhận biết nước từ ảnh Sentinel-2 trên nền Google Earth Engine: Áp dụng cho thành phố Sa Đéc, tỉnh Đồng Tháp,” *Tạp chí Khoa học và Công nghệ Thủy Lợi*, tập 60, 2020

[2] Đặng Thanh Tùng (2021), *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*. Kỷ yếu Hội thảo Quốc gia Giải pháp kết nối và chia sẻ hệ thống cơ sở dữ liệu phục vụ công tác đào tạo, quản lý lĩnh vực Tài nguyên và Môi trường, Nhà xuất bản Khoa học và Công nghệ – Hà Nội, trang-192.

[3] Ngô Hoàng Thanh, “Applying Random Forest approach in forecasting flash flood susceptibility area in Lao Cai region.,” *Journal of Mining and Earth Sciences*, vol. 61, no. 5, pp. 30-42, 2020.