

Sản phẩm toàn cầu ứng dụng trí tuệ nhân tạo phục vụ đào tạo, nghiên cứu trong lĩnh vực quản lý đất đai

Tạ Minh Ngọc*

*ThS. Trường Đại học Tài nguyên và Môi trường Hà Nội

Received: 12/10/2023; Accepted: 22/10/2023; Published: 30/10/2023

Abstract: This study investigates the feasibility of utilizing global products, namely Dynamic World (DW) and ESA's World Cover (WC) 2020, in the Hanoi region for the purposes of training and research in land use cover classification. The findings of this study indicate that these products can be applied effectively across various specialized fields in Vietnam, proving highly beneficial for training and research in the field of land management.

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

1. Đặt vấn đề

Ứng dụng các thuật toán học máy, trí tuệ nhân tạo (AI) hiện nay trên thế giới đã có các sản phẩm lớp phủ bề mặt và lớp phủ sử dụng đất (LULC) ở phạm vi toàn cầu. Các sản phẩm này đã được cách mạng hóa rất nhiều nhờ dữ liệu vệ tinh có độ phân giải trung bình và độ phân giải tương đối cao có sẵn như Landsat và Sentinel. Để theo dõi, đánh giá độ tin cậy của sản phẩm, người ta đã so sánh chéo các bản đồ LULC độ phân giải 10m toàn cầu và kết quả đã tìm thấy một số sai lệch về không gian và còn tồn tại sự chưa thống nhất về lớp phủ trong mỗi sản phẩm khác nhau. Cụ thể, đã có so sánh chéo và đánh giá độ chính xác của các sản phẩm Google's Dynamic World (DW), ESA's World Cover (WC) và Esri's Land Cover (Esri) để xem xét về việc khả năng áp dụng sản phẩm này trong tương lai. Đối với năm 2020, đã có ba sản phẩm bản đồ LULC toàn cầu là DW, WC, Esri được đánh giá và kết quả cho thấy sự tương ứng về không gian mạnh mẽ (tức là ước tính diện tích gần bằng nhau) cho các lớp phủ mặt nước, khu vực công trình xây dựng, thực phủ và một số loại cây trồng. Tuy nhiên, kết quả đánh giá cũng cho thấy rằng sản phẩm của WC thiên về việc ước tính các lớp phủ là lớp cỏ, trong khi đó Esri thiên về lớp phủ cây bụi còn sản phẩm DW thiên về các lớp phủ băng tuyết. [1,2].

Nghiên cứu này ứng dụng mô hình trí tuệ nhân tạo, khai thác sản phẩm DW và phục vụ đánh giá khả năng sử dụng của các lớp phủ sử dụng đất đối với các sản phẩm trên tại WC tại khu vực đô thị Hà Nội.

2. Nội dung nghiên cứu

2.1. Dữ liệu và khu vực nghiên cứu

Sản phẩm DW được xây dựng phù hợp với

Nguyên tắc AI của Google. DW được phát triển bằng cách sử dụng dữ liệu đào tạo của Hiệp hội địa lý quốc gia (National Geographic Society). Tập dữ liệu này được thành lập bởi dự án hợp tác giữa Hiệp hội Địa lý Quốc gia và Google cùng với Viện Tài nguyên Thế giới (World Resources Institute). Sản phẩm bao gồm 9 loại lớp phủ, bao gồm "Mặt nước", "Cây cối", "Cỏ", "Thảm thực vật ngập nước", "Cây trồng", "Cây bụi", "Khu vực xây dựng", "Sa mạc", "Tuyết và Băng".

Sản phẩm của WC cơ bản bao gồm 11 lớp phủ mô tả bề mặt trái đất với độ phân giải không gian là 10m trong đó bao gồm cả lớp phủ: "Rừng", "Cây bụi", "Đồng cỏ", "Đất trồng trọt", "Khu vực xây dựng", "Cây cối trụi/thưa thớt", "Băng tuyết", "Các vùng nước vĩnh viễn và theo mùa", "Đất ngập nước cây cỏ", "Rừng ngập mặn", "Rêu và địa y". Sản phẩm ESA WC đã được xác nhận độc lập bởi Đại học Wageningen (thống kê) và IIASA (không gian).

Các dữ liệu là các sản phẩm về LULC toàn cầu trên bao phủ toàn bộ khu vực đô thị Hà Nội.

2.2. Phương pháp nghiên cứu

Nghiên cứu sử dụng nền tảng GEE và ngôn ngữ lập trình JavaScript để trích xuất các dữ liệu toàn cầu DW, WC theo phạm vi hành chính khu vực đô thị Hà Nội. Cụ thể, nhóm tác giả đã thu nhận dữ liệu LULC của các sản phẩm toàn cầu là DW, WC dựa trên mô hình AI. Trong đó, 6 lớp phủ được trích xuất lần lượt bao gồm: 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, 6) lớp phủ Giao thông. Từ các kết quả trên, lựa chọn 3 lớp phủ cơ bản của khu vực để tiến hành so sánh, đánh giá các sản phẩm nghiên cứu.

Theo các nghiên cứu đã công bố trên thế giới,

các sản phẩm LULC toàn cầu trên được tạo ra từ các kết quả tổng hợp và dựa trên việc sử dụng mô hình trí tuệ nhân tạo với các thuật toán học máy như là: Thuật toán CART 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.

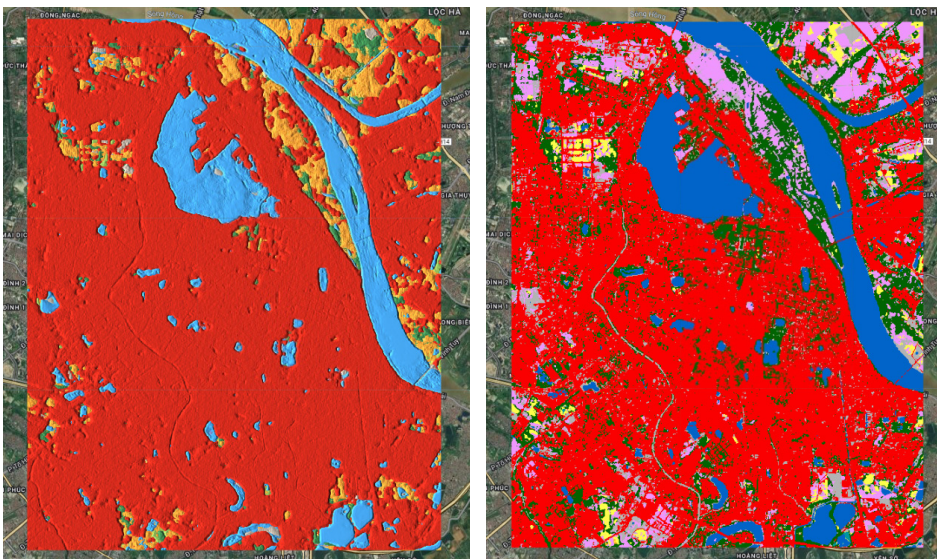
2.3. Kết quả nghiên cứu và thảo luận

Các kết quả thu nhận được có sự khác biệt rõ ràng giữa các sản phẩm DW, WC. Trong đó, có sự chênh lệch lớn về lớp Mặt nước của sản phẩm WC năm 2020 so với kết quả phân loại do nhóm tác giả thực hiện, giá trị khác biệt của lớp Đất trồng là lớn nhất, đạt tới giá trị khác biệt 1.52% so với tổng diện tích tự nhiên của khu vực nghiên cứu.

Bảng 1. Kết quả khác biệt giữa sản phẩm toàn cầu DW, WC

Loại sản phẩm	Mặt nước (%)	Đất trồng (%)	Nhà cửa (%)
DW-WC	0.88	1.52	-1.24

Sự khác biệt trên cũng có thể nhìn nhận một cách tương đối rõ ràng trên các hình ảnh của các sản phẩm phân loại lớp phủ sử dụng đất tại hình 1.



a. Sản phẩm của DW

b. Sản phẩm của WC

Hình 1. Các sản phẩm LULC toàn cầu theo mô hình AI trên nền tảng dữ liệu điện toán đám mây GEE

Các khác biệt về kết quả phân loại giữa các sản phẩm có thể do nhiều nguyên nhân bao gồm cả khách quan và chủ quan. Từ bảng kết quả 1 cho thấy kết quả tương đối giữa sản phẩm DW và WC. Để so sánh các kết quả DW và WC và đánh giá độ tin cậy của sản phẩm nào tốt hơn với các điều kiện tại khu vực nghiên cứu tại thời điểm này là thực sự còn nhiều khó khăn. Có nhiều nguyên nhân tồn tại trong việc đánh giá độ tin cậy của các sản phẩm toàn cầu DW và WC để phục vụ các công tác nghiên cứu cũng như quản lý sử dụng. Một trong những nguyên nhân ban đầu là mỗi sản phẩm hướng đến một trong những mục tiêu sử dụng khác nhau, có các lớp phủ sử dụng đất khác nhau nên gây khó khăn trong việc đánh giá độ chính xác của sản phẩm. Với 9 lớp phủ của sản phẩm DW, 11 lớp phủ của sản phẩm WC là khó khăn trong đánh giá độ tin cậy phân loại các lớp phủ của sản phẩm DW và WC. Tuy nhiên nghiên cứu này cũng đã sử dụng 3 lớp phủ cơ bản để so sánh, đánh giá các kết quả như đã phân tích ở trên. Kết quả đánh giá cũng đã cho thấy được một số sự tương đồng của các sản phẩm. Các khác biệt lớn giữa DW và WC có thể do dữ liệu sử dụng của sản phẩm WC năm 2020 được phân loại và tổng hợp từ nhiều nguồn nghiên cứu khác nhau như công bố của các nhà nghiên cứu nên khó đạt được sự thống nhất và nhất là khó có khả năng là sản phẩm được phân loại trong 1 thời điểm nhất định. Trong khi đó sản phẩm DW là sản phẩm được phân loại từ dữ liệu ảnh vệ tinh Sentinel-2, đây cũng là các sản phẩm được coi như là sản phẩm gần

thời gian thực (Near Real Time – NRT) nên có khả năng gần với thời điểm thực tế hơn tại khu vực nghiên cứu.

3. Kết luận và đề xuất

Việc chuyển đổi phương pháp khai thác, phân tích dữ liệu ảnh vệ tinh trực tuyến và các sản phẩm toàn cầu phục vụ công tác quản lý đất đai đang ngày càng tỏ ra hiệu quả. Điều này cũng góp phần tích cực vào công cuộc chuyển đổi số nói chung

hiện nay và của ngành Quản lý đất đai nói riêng.

Sản phẩm WC năm là sản phẩm tổng hợp của phân loại từ ảnh vệ tinh Sentinel-2 và một số kết quả nghiên cứu khác, do đó cần có những nghiên cứu đánh giá chi tiết hơn về tính thống nhất và độ tin cậy của sản phẩm khi sử dụng cho các lĩnh vực cụ thể.

Sản phẩm DW là một trong những sản phẩm về các lớp phủ sử dụng đất toàn cầu có nhiều ưu điểm như là gần thời gian thực, đây là một trong những yếu tố rất quan trọng trong các phân tích, đánh giá phục vụ công tác quản lý nhà nước về đất đai, môi trường.

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.

Tài liệu tham khảo

[1] Mirmazloumi, S.M.; Kakooei, M.; Mohseni, F.; Ghorbanian, A.; Amani, M.; Crosetto, M.; Monserrat, O. (2022). *ELULC-10, a 10 m European Land Use and Land Cover Map Using Sentinel and Landsat Data in Google Earth Engine. Remote Sens.* 14, 3041. <https://doi.org/10.3390/rs14133041>

[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] Lewis, R. J. (2000). *An introduction to classification and regression tree (CART) analysis*. In Annual meeting of the society for academic emergency medicine in San Francisco, California (Vol. 14). Citeseer.

[4] Breiman L. J. M. l.l (2001). *Random forests*. Kluwer Academic Publishers. Manufactured in The Netherlands. 45, 5.

[5] Breiman L. J. M. l.l (2001). *Random forests*. Kluwer Academic Publishers. Manufactured in The Netherlands. 45, 5.

[6] Vapnik V. N. and A. Y. Chervonenkis (1971). On the uniform convergence of relative frequencies of events to their probabilities. Measures of complexity. Springer, 11.

Nâng cao chất lượng tự học.... (tiếp theo trang 299)

- Thường xuyên tự kiểm tra, đánh giá kết quả tự học của sinh viên. Thực chất, đây là biện pháp xác định “cơ chế”, tạo động lực thúc đẩy hoạt động tự học của sinh viên và phong trào tự học tự giác thường xuyên trong nhà trường. Mục đích là nhằm giúp mỗi sinh viên nhận thức được những điểm mạnh, mặt hạn chế của bản thân để điều chỉnh quá trình tự nghiên cứu, học tập. Mặt khác, thông qua kiểm tra, đánh giá kết quả tự học của sinh viên để xây dựng và rèn luyện tính tích cực, chủ động, động viên tinh thần khắc phục khó khăn của sinh viên, góp phần nâng cao chất lượng tự học của sinh viên.

3. Kết luận

Tư tưởng tự học của Chủ tịch Hồ Chí Minh có một giá trị hết sức to lớn đối với sự nghiệp giáo dục của nước nhà. Vì vậy, nâng cao chất lượng tự học của sinh viên Trường Đại học Công nghiệp Việt Hung theo tư tưởng tự học của Chủ tịch Hồ Chí Minh là rất cần thiết. Nó góp phần quyết định kết quả học tập của sinh viên và chất lượng đào tạo của nhà trường. Để đảm bảo chất lượng, hiệu quả hoạt động tự học cũng như nâng cao chất lượng tự học của sinh viên Trường

Đại học Công nghiệp Việt Hung theo tư tưởng tự học của Chủ tịch Hồ Chí Minh, một trong những vấn đề cần quan tâm là tiến hành phối hợp các giải pháp giáo dục động cơ tự học và rèn luyện, tạo điều kiện để sinh viên hình thành và phát triển năng lực tự học. Từ đó, sinh viên vận dụng liên tục vào giải quyết các bài tập nhận thức, bài tập nghiên cứu đa dạng, phù hợp với thực tiễn xã hội nảy sinh. Tư tưởng và tinh thần tự học của Người mãi mãi tỏa sáng, soi rọi cho thế hệ mai sau, đặc biệt là sinh viên Trường Đại học Công nghiệp Việt Hung phấn đấu, rèn luyện, tự vươn lên trong học tập, tư dưỡng, rèn luyện bản thân.

Tài liệu tham khảo:

1. Hồ Chí Minh (2011), toàn tập, tập 2 (2011), Nxb CTQG, Hà Nội

2. Hồ Chí Minh (2011), toàn tập, tập 4 (2011), Nxb CTQG, Hà Nội.

3. Hồ Chí Minh (2011), toàn tập, tập 5 (2011), Nxb CTQG, Hà Nội.

4. Hồ Chí Minh (2011), toàn tập, tập 6 (2011), Nxb CTQG, Hà Nội.