



Original Article

Study Model for Information Reconstruction on Cloud Contaminated Area for Single Multispectral Remote Sensing Sentinel-2 Imagery using Generative Adversarial Network

Do Thi Nhung, Pham Vu Dong, Bui Quang Thanh, Pham Van Manh*

VNU University of Science, 334 Nguyen Trai, Thanh Xuan, Ha Noi, Viet Nam

Received 26 July 2021

Revised 15 August 2021; Accepted 24 August 2021

Abstract: Cloud and cloud shadow cause information loss in optical remote sensing analysis. South East Asia, especially Vietnam, Sentinel-2 imagery has short re-visit cycle and observations tend to be contaminated with cloud and cloud shadow. Traditional cloud removal methods require close date multi-temporal data to avoid seasonal land cover changes. In this study, a method of integrating Deep Convolutional Neural Networks (DCNN) and Generative Adversarial Network (GAN) was proposed. This machine learning model estimates the information loss over cloud contaminated areas on a single Sentinel-2 image. The results show that for images with cloud cover rate under 25%, our model can reconstruct cloudless images with PSNR (25 – 40 dB) and SSIM (0.86 – 0.93) compared to real clear images. On the other hand, with cloud cover rate up to 40%, the model performance will be affected heavily by the distribution of cloud and cloud shadow areas. By investigating DCNN and GAN, our method has proven to be an effective tool to remove cloudy images with low and medium rates, which enriches the clear optical remote sensing data sources for environment monitoring.

Keywords: Cloud and cloud shadow, Remote sensing, Sentinel-2, Generative adversarial network, Single image cloud removal.

* Corresponding author.

E-mail address: manh10101984@gmail.com

<https://doi.org/10.25073/2588-1094/vnuees.4815>

Nghiên cứu mô hình tái tạo vùng dữ liệu bị mây che phủ trên ảnh đơn viễn thám đa phổ Sentinel-2 sử dụng thuật toán mạng đối lập

Đỗ Thị Nhung, Phạm Vũ Đông, Bùi Quang Thành, Phạm Văn Mạnh*

*Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội,
334 Nguyễn Trãi, Thanh Xuân, Hà Nội, Việt Nam*

Nhận ngày 26 tháng 7 năm 2021

Chỉnh sửa ngày 15 tháng 8 năm 2021; Chấp nhận đăng ngày 24 tháng 8 năm 2021

Tóm tắt: Mây và bóng mây là những yếu tố gây ảnh hưởng đến việc mất thông tin trên ảnh viễn thám quang học. Đặc biệt, trên lãnh thổ Việt Nam và các nước Đông Nam Á nói chung, ảnh Sentinel-2 có chu kỳ bay ngắn và hầu hết các ảnh thu nhận đều bị nhiễm mây. Việc lọc mây bằng các phương pháp truyền thống yêu cầu việc sử dụng các ảnh đa thời gian. Tuy nhiên, điều này chỉ nên được áp dụng với các bộ ảnh có chu kỳ bay chụp ngắn nhằm giảm sự sai lệch về giá trị phổ giữa các mùa và sự thay đổi của lớp phủ. Trong nghiên cứu này, nhóm tác giả đề xuất lựa chọn một phương pháp tái tạo dữ liệu vùng nhiễm mây bằng cách chỉ sử dụng ảnh đơn dựa trên mô hình kết hợp mạng nơron nhân tạo phức hợp sâu (DCNN) và kỹ thuật mạng đối lập (GAN). Phương pháp sử dụng trong nghiên cứu này có khả năng lấp thông tin một cách chính xác cho các vùng thông tin bị mây che phủ trên ảnh Sentinel-2. Kết quả cho thấy, ảnh Sentinel-2 với độ phủ mây và bóng mây dưới 25% thì ảnh tái tạo có độ tương đồng cao với giá trị của PSNR (25-40 dB) và SSIM (0,86-0,93). Trong khi đó, với các ảnh có độ phủ mây và bóng mây trên 40%, mô hình vẫn có thể tái tạo được nhưng độ chính xác thông tin trên ảnh sẽ phụ thuộc vào mật độ phân bố không gian của mây và bóng mây. Phương pháp kết hợp DCNN và GAN trong nghiên cứu này, đối với các ảnh nhiễm mây và bóng mây với tỉ lệ nhỏ và trung bình có thể được tái tạo thành ảnh không mây, cung cấp nguồn dữ liệu ảnh vệ tinh sạch cho mục đích giám sát tài nguyên và môi trường.

Từ khóa: Mây và bóng mây, Viễn thám, Sentinel-2, Mạng đối lập, Loại bỏ mây hình ảnh đơn.

1. Mở đầu

Trong lĩnh vực viễn thám quang học, chất lượng của ảnh vệ tinh phụ thuộc rất nhiều vào điều kiện thời tiết trong quá trình thu nhận [1]. Một trong những vấn đề chính ảnh hưởng tới chất lượng ảnh là sự xuất hiện của mây và bóng mây. Chính vì thế, quá trình loại bỏ mây và bóng mây là một trong những giai đoạn của quá trình tiền xử lý ảnh viễn thám quang học, để phục vụ

cho công việc phân tích ảnh sau này. Việc lọc mây và bóng mây trên ảnh viễn thám quang học có thể được thực hiện bằng nhiều cách khác nhau, như sử dụng nhiều cặp ảnh chụp liên kế thời gian trên cùng một vị trí [2]. Tại vị trí ảnh chính có những khu vực bị nhiễm mây và bóng mây, vùng thông tin của các khu vực này sẽ được lấy từ các ảnh tại thời điểm khác và chồng lấp vào vị trí đó. Bằng cách sử dụng hai hoặc nhiều hơn các ảnh tại cùng một vị trí sẽ tạo ra được một

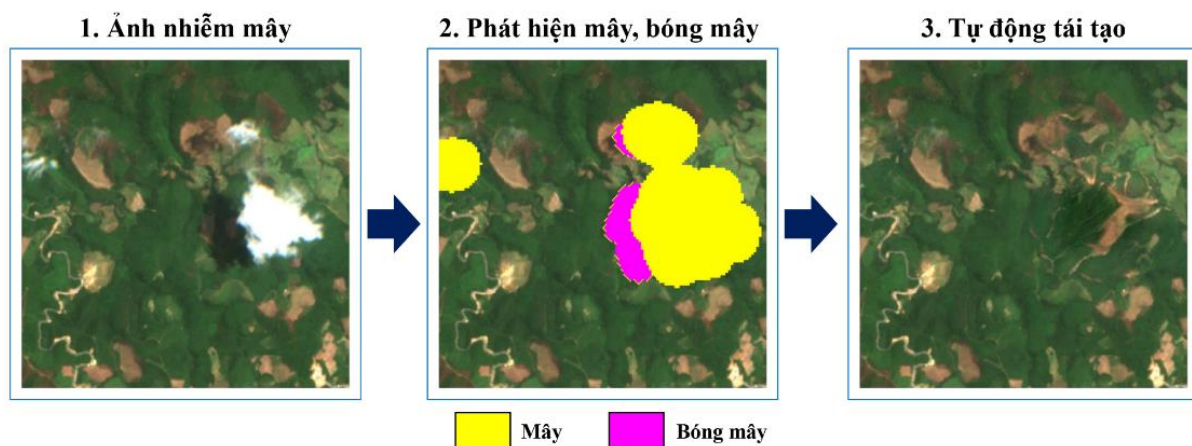
* Tác giả liên hệ.

Địa chỉ email: manh10101984@gmail.com

<https://doi.org/10.25073/2588-1094/vnuees.4815>

ảnh có lượng mây che phủ ít nhất. Tuy nhiên, việc lấy thông tin từ các ảnh khác thời điểm chụp sẽ nảy sinh một số vấn đề như: i) Giá trị phổ phản xạ không đồng nhất, tạo ra hiện tượng các vùng sau khi lọc mây sẽ khác biệt với các vùng xung quanh; và ii) Yêu cầu phải thu thập nhiều hơn một ảnh cho một khu vực nghiên cứu và các ảnh thu thập phải có vùng mây che phủ không trùng lặp. Trong các dữ liệu ảnh viễn thám quang học đa phổ, ảnh vệ tinh Sentinel-2 là nguồn dữ liệu mở và miễn phí có độ phân giải không gian cao nhất (10 m) so với các ảnh viễn thám quang học miễn phí khác. Tuy nhiên, với đặc điểm thời gian bay chụp lặp lại tập trung chủ yếu ở khu vực Châu Âu. Trong khi đó, các khu vực khác như khu vực Đông Nam Á, đặc biệt là tại Việt Nam, số lượng ảnh lặp lại trên một cảnh ảnh trong một năm là rất ít. Chính vì thế, ảnh Sentinel-2 chụp tại các khu vực ở Việt Nam thường bị nhiễm mây

và rất khó có thể tìm được các ảnh gần thời điểm với nhau để tạo ra một ảnh không mây. Mây và bóng mây là những yếu tố cản trở việc giám sát các hiện vật trên bề mặt Trái Đất từ ảnh vệ tinh quang học, chúng làm cho các thông tin trên ảnh bị mờ hoặc mất hoàn toàn thông tin [3, 4]. Trong những năm qua, các nghiên cứu trong lĩnh vực viễn thám đã được thực hiện để tiến hành phát hiện và loại bỏ mây và bóng mây trên ảnh viễn thám quang học [5]. Trong các nghiên cứu phát hiện về nhiễm mây, đã kết hợp sử dụng các ảnh viễn thám đa thời gian Formosat-2 và Landsat để lọc các các vùng pixel nhiễm mây trên ảnh [6]. Ngoài ra, một hàm toán F-mask đã được sử dụng, đây là một trong những hàm toán nổi tiếng dùng để phát hiện mây và bóng mây bằng cách tận dụng kênh ảnh Cirrus trên dữ liệu ảnh Landsat-8 và Sentinel-2 [4, 5].



Hình 1. Minh họa mô hình tái tạo vùng nhiễm mây trên ảnh vệ tinh Sentinel-2.

Cùng với các nghiên cứu về phát hiện mây và bóng mây, các nghiên cứu về tái tạo thông tin và loại bỏ vùng mây và bóng mây được nghiên cứu trên thế giới với nhiều các cách tiếp cận khác nhau [5]. Một nghiên cứu điển hình đã sử dụng phương pháp thay thế pixel tương đương bằng các bộ ảnh viễn thám đa thời gian [7]. Hàm MRF (Markov Random Fields) được sử dụng để tìm các pixel trên ảnh có giá trị giống nhau cao nhất. Trong một nghiên cứu khác, phương pháp phân ngưỡng được sử dụng để loại bỏ phân mây dày

trên ảnh vệ tinh và cân bằng giá trị phản xạ phổ bằng các ảnh viễn thám đa thời gian [10]. Có thể nhận thấy, hầu hết các phương pháp đều tiếp cận các dữ liệu ảnh đa thời gian với nguyên tắc lấy dữ liệu từ ảnh không mây để phục hồi cho ảnh gốc bị nhiễm mây và bóng mây. Việc tái tạo dữ liệu từ một ảnh đơn lẻ như chưa được nghiên cứu trên thế giới. Hiện nay, trong lĩnh vực trực quan máy tính (Computer Vision), tái tạo thông tin trên ảnh (Image Inpainting) là một trong những kỹ thuật mới được phát triển trong việc

khôi phục vùng thông tin đã mất trên ảnh. Mô hình Image Inpainting là một cấu trúc DCNN (Deep Convolutional Neural Network) có khả năng nhận dữ liệu đầu vào là ảnh với các vùng thông tin bị mất (invalid pixel) và tạo ra một ảnh đầu ra đã được tái tạo thông tin cho các vùng thiếu thông tin trong hình ảnh. Các mô hình Image Inpainting điển hình như trong các nghiên cứu [11-13]. Tuy nhiên, một vấn đề trong các phương pháp này đó là chúng chỉ nhận những ảnh có vùng thông tin bị mất (invalid pixel) ở dạng đa giác (rectangular hole). Trong khi đó, các vùng bị nhiễm mây trên ảnh vệ tinh có nhiều các hình dạng phức tạp khác nhau và không tuân theo một quy tắc hay đặc điểm nhận dạng nào. Để khắc phục điều này, mô hình Image Inpainting cho vùng thông tin bị mất đa hình dạng (Irregular hole) là một trong những nghiên cứu đầu tiên về kỹ thuật tái tạo ảnh với vùng thông tin bị mất không nhất quán [14]. Do đó, trong nghiên cứu này với cách tiếp cận vấn đề tái tạo dữ liệu vùng nhiễm mây (mây và bóng mây) bằng cách phân loại các pixels trên ảnh vệ tinh bị nhiễm mây là các “invalid pixels”. Từ đó xây dựng mô hình học máy sâu CNN để tự động thêm thông tin không gian của các đối tượng bị nhiễm mây vào những vị trí pixels đó trên hình ảnh viễn thám quang học như minh họa trong (Hình 1).

2. Phương pháp nghiên cứu và cách tiếp cận

Để loại bỏ mây và bóng mây, các phương pháp sử dụng hiện nay có thể được chia thành các phương pháp theo không gian - thời gian, quang phổ, đa cảm biến và phương pháp kết hợp, tùy theo loại dữ liệu đầu vào khác nhau của các mô hình để loại bỏ các vùng nhiễm mây [15-17]. Dựa trên điều tra sâu rộng và công việc trước đó, nghiên cứu này sử dụng kỹ thuật mạng đối lập (GAN) để loại bỏ mây và bóng mây trên hình ảnh viễn thám quang học và đề xuất một mô hình kết hợp DCNN và GAN cho ảnh đơn không bị nhiễm mây và bóng mây bằng cách tự tạo mây giả (mây nhân tạo). Mô hình này mô phỏng cơ chế trực quan và đồng thời chú ý không gian cục bộ ở khu vực mây và bóng mây đến dữ liệu bên

ngoài vùng nhiễm mây để tái tạo hay điền vào vùng dữ liệu còn thiếu tạo ra hình ảnh không có mây và bóng mây với chất lượng tốt hơn. Đồng thời, kỹ thuật inpainting được đề xuất có khả năng hoạt động tốt hơn cho kết nối vùng [18]. Do đó, dữ liệu bị thiếu cho các vùng mây và bóng mây đã được tổng hợp hiệu quả.

Trong nghiên cứu này, tập dữ liệu mẫu bao gồm hình ảnh vệ tinh Sentinel-2 có độ phân giải không gian cao, được thu thập trong khoảng từ năm 2017 đến năm 2020. Ảnh vệ tinh Sentinel-2 là tập dữ liệu đa phổ, với mỗi dải phổ được lưu trữ dưới dạng hình ảnh riêng biệt. Đối với các thử nghiệm trong nghiên cứu, hình ảnh vệ tinh Sentinel-2 từ các dải phổ nhìn thấy bao gồm: B2-Blue (Xanh lam), B3-Green (Xanh lục) và B4-Red (Đỏ) tất cả đều có độ phân giải không gian 10 m. Các hình ảnh không có mây trong nghiên cứu được chọn với độ che phủ mây là 0%. Trong khi đó, đối với hình ảnh có mây và bóng mây trên ảnh Sentinel-2, nghiên cứu chọn phạm vi vùng bị nhiễm mây và bóng mây bất kỳ trong khoảng từ 10-40%. Tất cả các hình ảnh được tải xuống ở khu vực miền Bắc Việt Nam, nơi rất dễ bị ảnh hưởng bởi mây và bóng mây trong quá trình bay chụp ảnh vệ tinh. Nghiên cứu sử dụng 12 hình ảnh không mây và 5 hình ảnh có mây và bóng mây để đào tạo. Sau đó, các hình ảnh được trích xuất với các mảnh với kích thước 256×256. Các quá trình xử lý chính của mô hình kết hợp DCNN và GAN bao gồm:

2.1. Kỹ thuật tạo mây nhân tạo (mây giả) cho mô hình học máy sâu

Quá trình học máy có kiểm định của hàm phức hợp CNN yêu cầu dữ liệu đầu ảnh nhiễm mây X_i , vùng X_c và ảnh không nhiễm mây Y . Mô hình học máy sẽ được học (training) để dự đoán các thông tin bị mất dưới vùng X_c qua các thông tin từ ảnh Y . Ảnh tái tạo Y' được thể hiện dưới hàm (1):

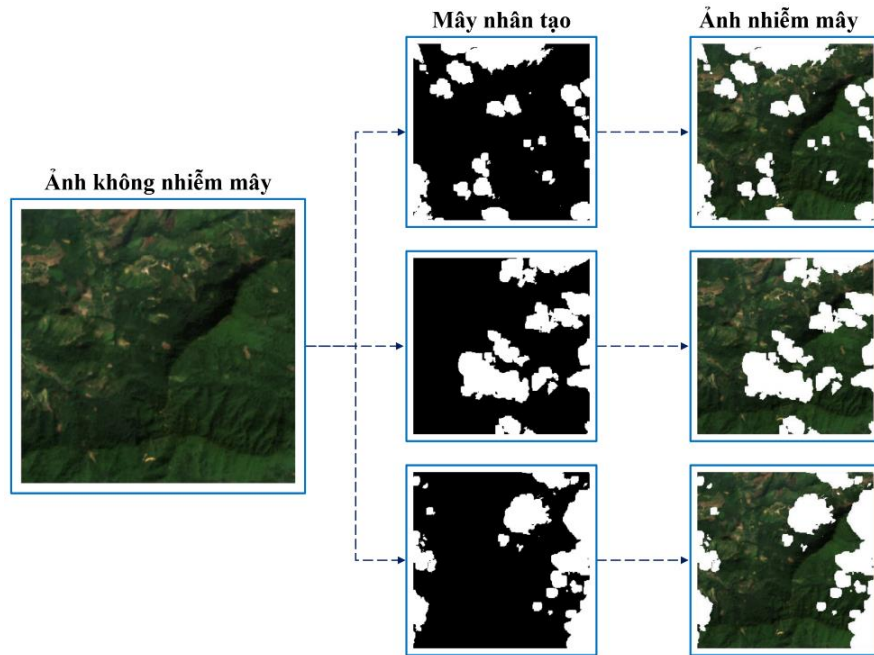
$$Y' = f(X_i, X_c) \quad (1)$$

Mô hình học máy sâu sau đó sẽ tối ưu hàm sai số Loss sao cho nhỏ nhất theo hàm (2):

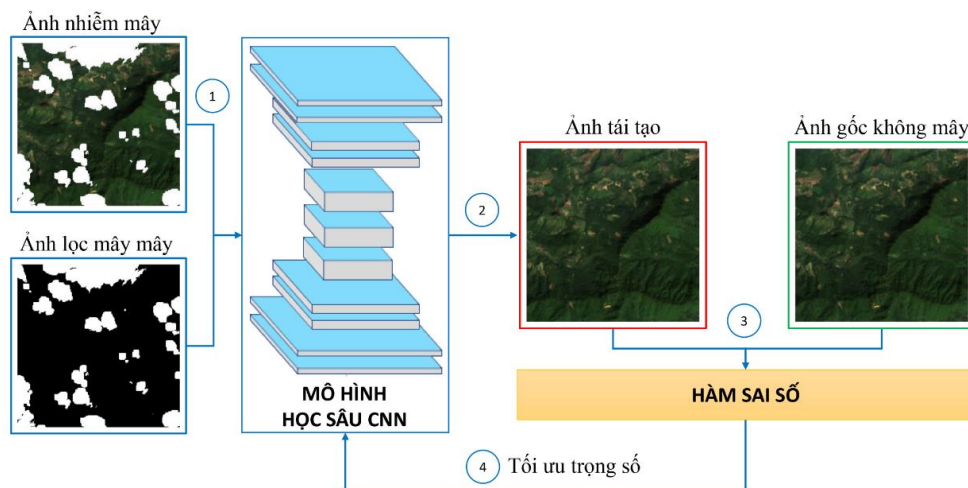
$$Loss = \minimize(|Y - Y'|) \quad (2)$$

Tuy nhiên thu thập dữ liệu ảnh bị nhiễm mây và ảnh không mây của một vị trí sẽ nảy sinh các vấn đề sau: i) thời gian chụp giữa hai ảnh Sentinel-2 tại một vị trí có thể rất xa (đặc biệt tại khu vực bay chụp châu Á). Điều này có thể dẫn đến hiện tượng thay đổi lớp phủ trên các vùng mây và bóng mây của hai ảnh X_i và Y không

đồng nhất; và ii) giá trị phổ của ảnh X_i và Y không đồng nhất tại hai thời điểm chụp, gây khó khăn cho việc tối ưu thuật toán học sâu. Chính vì thế, trong nghiên cứu này, nhóm tác giả đề xuất phương pháp tạo ra ảnh mây giả (nhân tạo) từ ảnh không mây, từ đó dùng ảnh mây giả để tối ưu hóa mô hình học sâu (Hình 2).



Hình 2. Tạo ảnh mây nhân tạo.

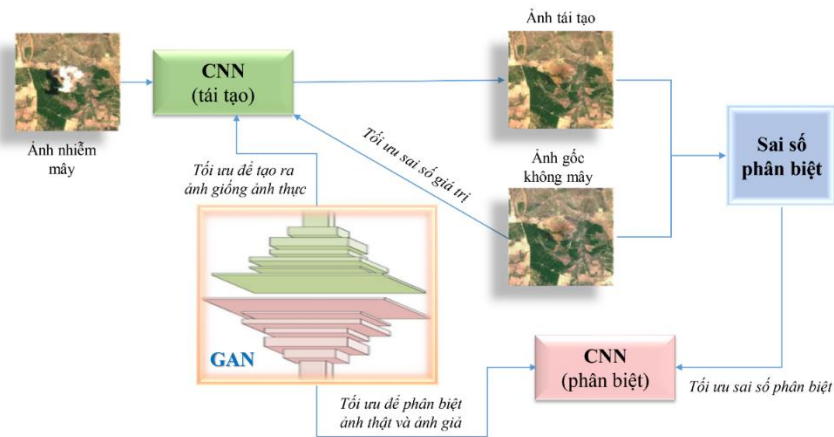


Hình 3. Quy trình học máy gồm: (1) Ảnh nhiễm mây và ảnh lọc mây được đưa vào mô hình CNN, (2) Mô hình tái tạo lại dữ liệu vùng nhiễm mây, (3) Sai số giữa ảnh tái tạo và ảnh gốc không mây được tính toán, (4) Giá trị sai số sau đó được sử dụng để tối ưu mô hình CNN.

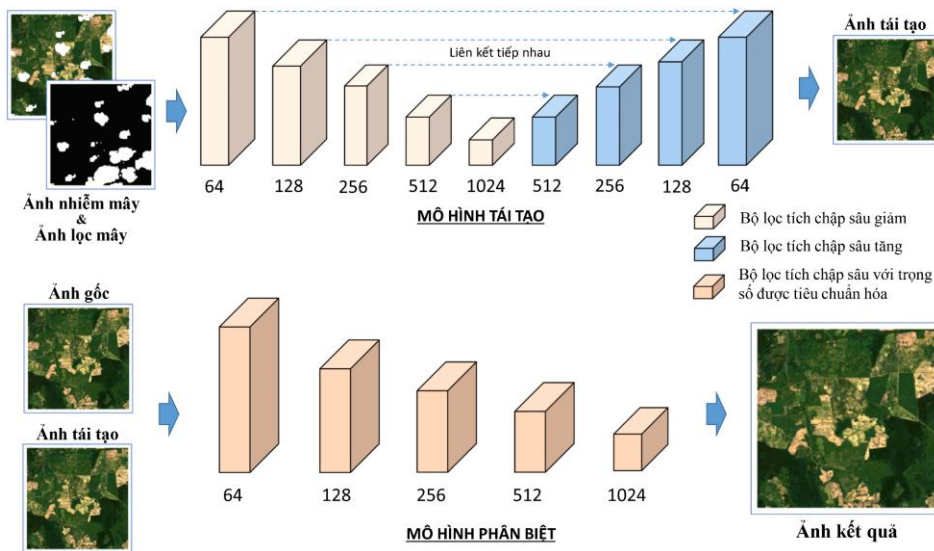
Dữ liệu ảnh không mây và mây nhân tạo được lấy ngẫu nhiên trên các ảnh Sentinel-2 được chụp tại nhiều khu vực che phủ trên lãnh thổ miền Bắc Việt Nam. Với mỗi ảnh đơn có thể kết hợp nhiều ảnh mây nhân tạo để tạo ra nhiều ảnh nhiễu mây giả ở các vị trí khác nhau. Kỹ thuật này giúp cho bộ dữ liệu ảnh nhiễu mây và ảnh mây X_i và ảnh không mây Y hoàn toàn đồng nhất về mặt thông tin dưới lớp mây và giá trị phổ. Trong nghiên cứu này, độ che phủ của mây nhân tạo dao động trong khoảng 10% - 40% của ảnh. Với kỹ thuật tạo ảnh mây nhân tạo, mô hình học máy được minh họa như trong (Hình 3).

2.2. Mô hình học máy sâu mạng đối lập GAN

Quy trình mạng đối lập GAN được cấu tạo từ hai mô hình mạng nơron nhân tạo phức hợp Convolutional Neural Network (CNN) (Hình 4). Hai mô hình được tối ưu hóa trọng số theo hai hàm toán đối lập nhau: mô hình CNN tái tạo được tối ưu để tạo ra ảnh không mây giống với ảnh thực. Trong khi đó, mô hình CNN phân biệt được tối ưu để xác định ảnh không mây là ảnh thực hay ảnh tái tạo (ảnh giả).



Hình 4. Quy trình kết hợp giữa hai mô hình học sâu CNN với hai hàm toán tối ưu khác nhau để tạo ra mô hình học máy sâu mạng đối lập GAN.



Hình 5. Chi tiết cấu trúc 2 mô hình CNN tái tạo và CNN phân biệt trong mạng đối lập GAN.

Mạng đối lập Generative Adversarial Network (GAN) là mô hình học sâu có khả năng tạo ra các hình ảnh nhân tạo giống với ảnh tự nhiên [19]. Kỹ thuật mạng đối lập dựa trên nguyên tắc như: i) Mô hình tạo Generator (G), và (ii) Mô hình phân biệt Discriminator (D). Trong quá trình học máy, mô hình G sẽ tạo ra các ảnh sao cho giống ảnh tự nhiên nhất có thể, trong khi đó mô hình D sẽ đánh giá ảnh do G tạo ra là ảnh tự nhiên thật (real) hay ảnh tự nhiên giả (reconstructed). Trong nghiên cứu này, nhóm tác giả xây dựng mô hình G và D dựa trên hàm lọc Gated Convolution với kiến trúc U-Net [16]. Hai mô hình Generator và Discriminator được minh họa như (Hình 5).

2.3. Hàm sai số

Sai số mạng đối lập

Theo kỹ thuật GAN [14], nghiên cứu xác định hai mô hình, Generator (G) và Discriminator (D), với các giá trị trọng số tương ứng θ_G và θ_D . Cho ảnh nhiễu mây X_i và ảnh lọc mây X_c , mô hình G tạo ra ảnh tái tạo $Y' = G(X_i, X_c)$, trong khi mô hình D sẽ dự đoán ảnh không mây thật Y và ảnh không mây tái tạo Y' là ảnh thật hay ảnh giả. Chính vì thế, mô hình D được học để tối ưu hóa khả năng dự đoán đúng nhãn từ hàm $\log(D(Y))$ và dữ liệu tạo ra từ mô hình G $\log(1 - D(G(X_i, X_c)))$. Ở chiều còn lại, mô hình G được học để giảm thiểu sai số của hàm $\log(1 - D(G(X_i, X_c)))$. Sai số mạng đối lập được thể hiện theo các hàm (3) và (4):

$$L_D = \text{maximize} \left(\left[\log(D_{\theta_D}(Y)) \right] + \left[\log(1 - D_{\theta_D}(G_{\theta_G}(X_i, X_c))) \right] \right) \quad (3)$$

$$L_{G_{\text{adversarial}}} = \text{minimize} \left(\left[\log(1 - D_{\theta_D}(G_{\theta_G}(X_i, X_c))) \right] \right) \quad (4)$$

Sai số thông tin

Sai số thông tin của ảnh tái tạo Y' và ảnh không mây Y được tính theo giá trị sai số tuyệt đối giữa hai ảnh được tính toán theo hàm (5):

$$L_1 = \frac{1}{C \times W \times H} \sum_{z=1}^C \sum_{x=1}^W \sum_{y=1}^H |Y - Y'| \quad (5)$$

Trong đó: H, W, C là chiều dài, rộng và số kênh của ảnh.

2.4. Đánh giá hiệu suất mô hình

Dữ liệu đầu vào của mô hình là ảnh vệ tinh Sentinel-2 bị nhiễu mây và bóng mây và đầu ra là ảnh không có mây và bóng mây sau khi loại bỏ các vùng bị nhiễu mây. Để đo chất lượng của hình ảnh không có mây và bóng mây được tạo ra và khả năng loại bỏ mây và bóng mây từ mô hình kết hợp DCNN và GAN. Trong nghiên cứu này, hai tỷ số đánh giá chất lượng ảnh sau khi loại bỏ vùng nhiễu mây, (i) tỷ lệ tín hiệu trên độ nhiễu (Peak Signal to Noise Ratio - PSNR); và (ii) chỉ số cấu trúc tương đồng (Structural Similarity Index - SSIM). Đây là hai chỉ số được sử dụng phổ biến để đánh giá chất lượng về sự tương đồng giữa hai ảnh [21, 22].

PSNR là phép đo khách quan để đánh giá chất lượng hình ảnh tái tạo, công thức tính toán được thể hiện theo hàm (6).

$$PSNR = 10 \log_{10} \left(\frac{(2^n - 1)^2}{MSE} \right) \quad (6)$$

Trong đó, n là các bits của giá trị độ xám của pixel (8 bits); MSE là lỗi bình phương trung bình giữa hình ảnh gốc và ảnh tái tạo, được tính toán theo hàm (7).

$$MSE = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W (X(i, j) - Y(i, j))^2 \quad (7)$$

Giá trị của PSNR nằm trong khoảng 20 đến 40, giá trị càng lớn thể hiện khoảng cách gần hơn giữa hình ảnh tái tạo và hình ảnh gốc, và chất lượng ảnh tái tạo tốt hơn.

SSIM là chỉ số đánh giá mức độ giống nhau của hai hình ảnh (gốc và tái tạo) thông qua các khía cạnh về độ sáng, độ tương phản và cấu trúc, có các hàm tương ứng lần lượt là (8), (9), và (10).

$$l(X, Y) = \frac{2\mu_X \mu_Y + C_1}{\mu_X^2 + \mu_Y^2 + C_1} \quad (8)$$

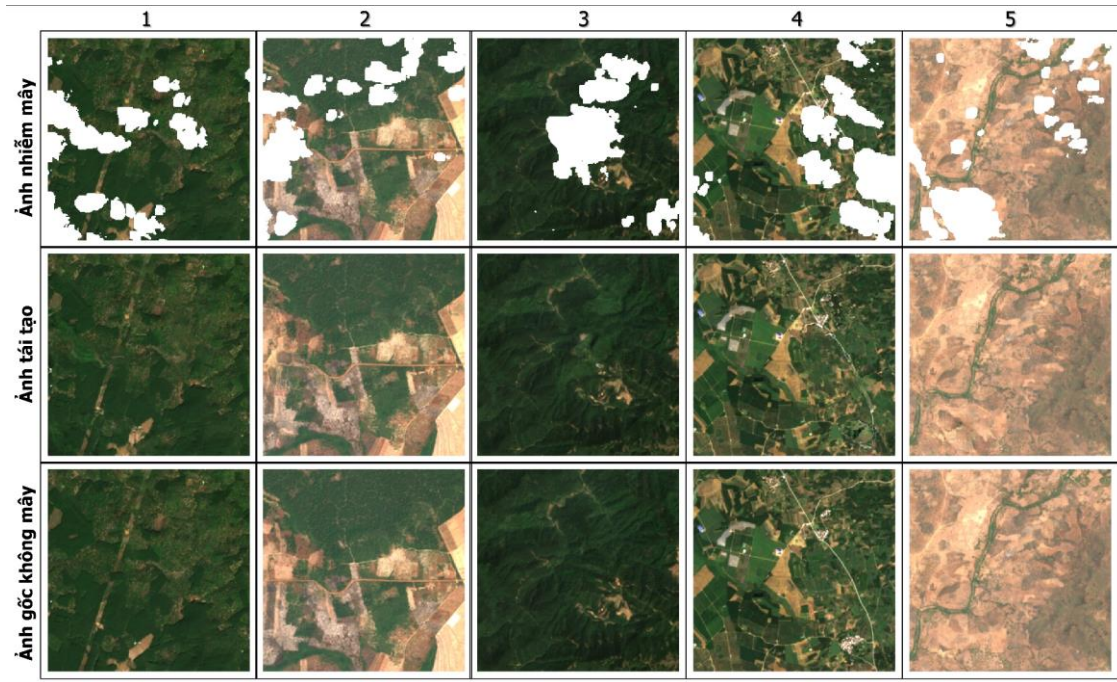
$$c(X, Y) = \frac{2\sigma_X \sigma_Y + C_2}{\sigma_X^2 + \sigma_Y^2 + C_2} \quad (9)$$

$$s(X, Y) = \frac{\sigma_{XY} + C_3}{\sigma_X \sigma_Y + C_3} \quad (10)$$

Trong đó, C_1, C_2, C_3 là các hằng số để tránh lỗi phân chia số không. Lần lượt μ, σ là giá trị trung bình và phương sai của hình ảnh, σ_{XY} là phương sai của hai ảnh gốc và tái tạo. Do đó, công thức tính toán tổng thể của SSIM được thể hiện theo phương trình (11).

$$SSIM = I(X, Y) \cdot c(X, Y) \cdot s(X, Y) \quad (11)$$

Phạm vi giá trị của SSIM có giá trị từ 0 đến 1, giá trị lớn hơn có nghĩa là giữa hai hình ảnh (gốc và tái tạo) giống nhau hơn. Nếu giá trị là 1 thì hai ảnh gốc và tái tạo hoàn toàn giống nhau.



Hình 6. Kết quả tái tạo dữ liệu trên ảnh mây nhân tạo.

3. Kết quả và thảo luận

3.1. Hiệu suất mô hình DCNN và GAN

Kết quả của mô hình tái tạo dữ liệu trên ảnh mây nhân tạo được thể hiện trong Hình 6. Có thể nhận thấy, các giá trị của chỉ số PSNR và SSIM trên các khu vực đồi núi thường cho kết quả cao, bởi vì các đối tượng ở khu vực này phân bố không gian một cách liên tục, khó có biến động dưới vùng mây. Trong khi ở các vùng đồng bằng có nhiều đối tượng không liên tục, làm cho mô hình dự đoán thiếu độ chính xác, làm giảm giá trị các chỉ số.

Kết quả phân tích định lượng mô hình kết hợp DCNN-GAN trên bộ dữ liệu thử nghiệm được thể hiện trong Bảng 1. Nhận thấy, các chỉ số đánh giá PSNR và SSIM của mô hình kết hợp

DCNN-GAN dao động trong khoảng từ 24,11 dB – 32,60 dB (PSNR) và SSIM (0,87 – 0,93). Đây là các mức chấp nhận được (tốt) của giá trị của PSNP và giá trị của SSIM.

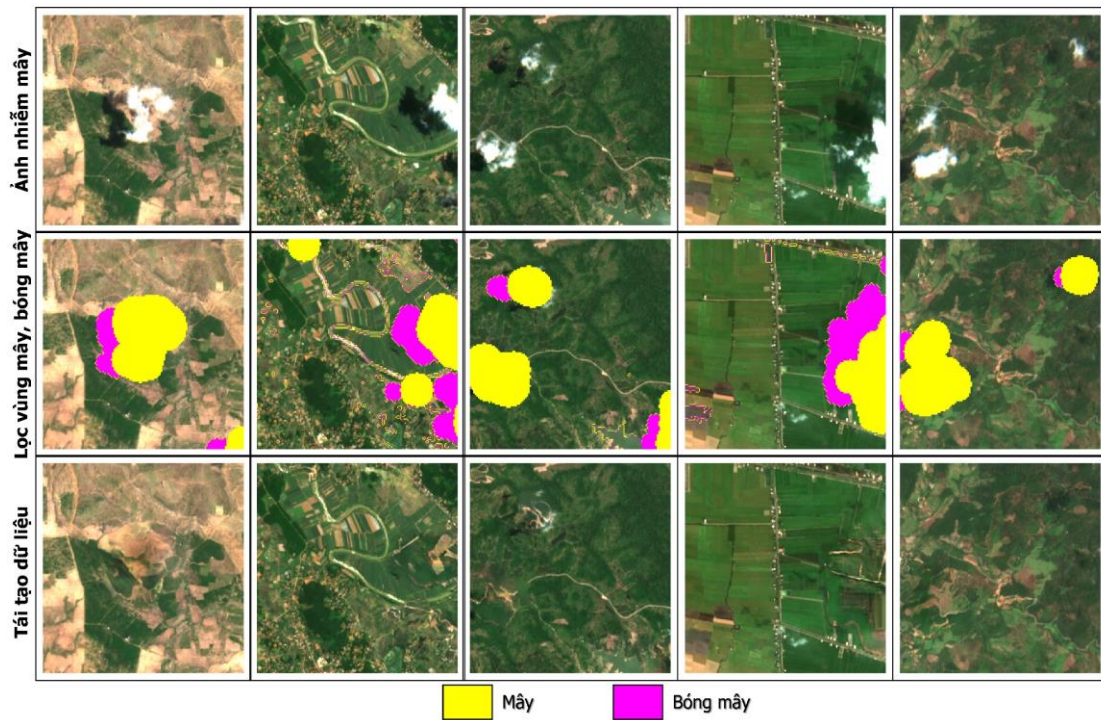
Bảng 1. Phân tích định lượng trên tập mẫu.

Tập mẫu	Chỉ số định lượng	
	PNSR	SSIM
1	31,99	0,92
2	26,74	0,91
3	32,60	0,93
4	24,11	0,87
5	26,17	0,87

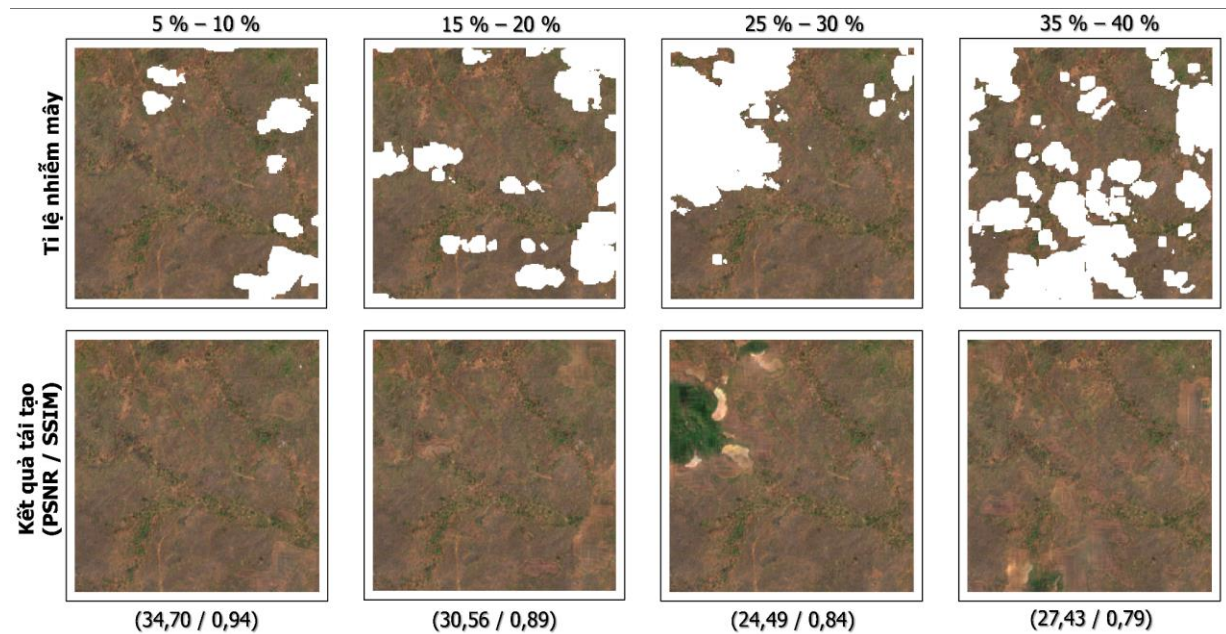
Với ảnh mây thật, vùng mây và bóng mây được xác định bằng hàm F-mask [6], sau đó dữ liệu

được mô hình tự động tính toán, loại bỏ vùng mây và tái tạo thông tin trên vùng mây đã được xác định

như minh họa trong Hình 7. Kết quả trên ảnh mây thật được đánh giá qua giá trị trực quan.



Hình 7. Kết quả tái tạo dữ liệu trên ảnh mây nhân tạo.



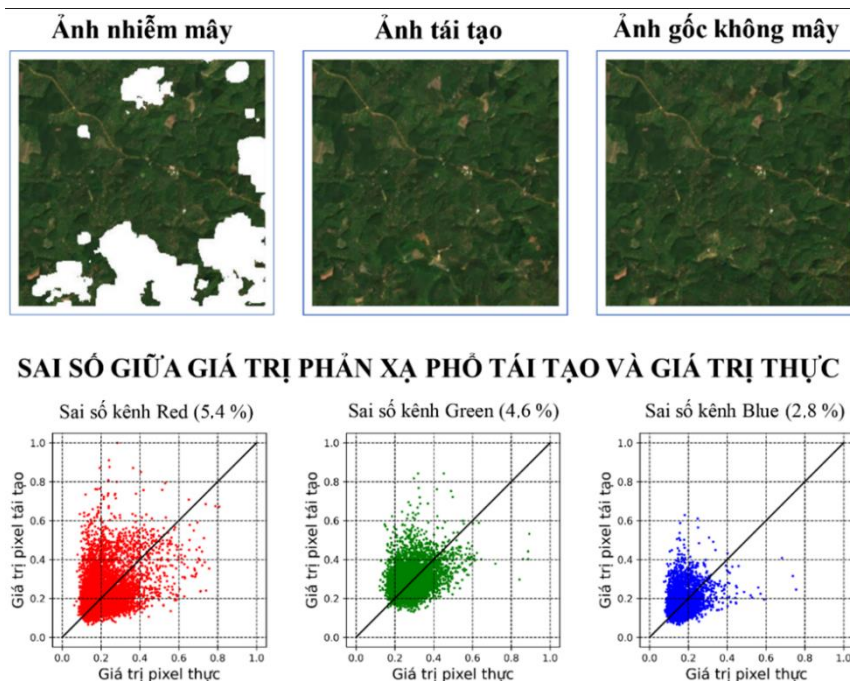
Hình 8. Chất lượng tái tạo suy giảm theo tỉ lệ mây che phủ và mức độ tập trung.

3.2. Thảo luận

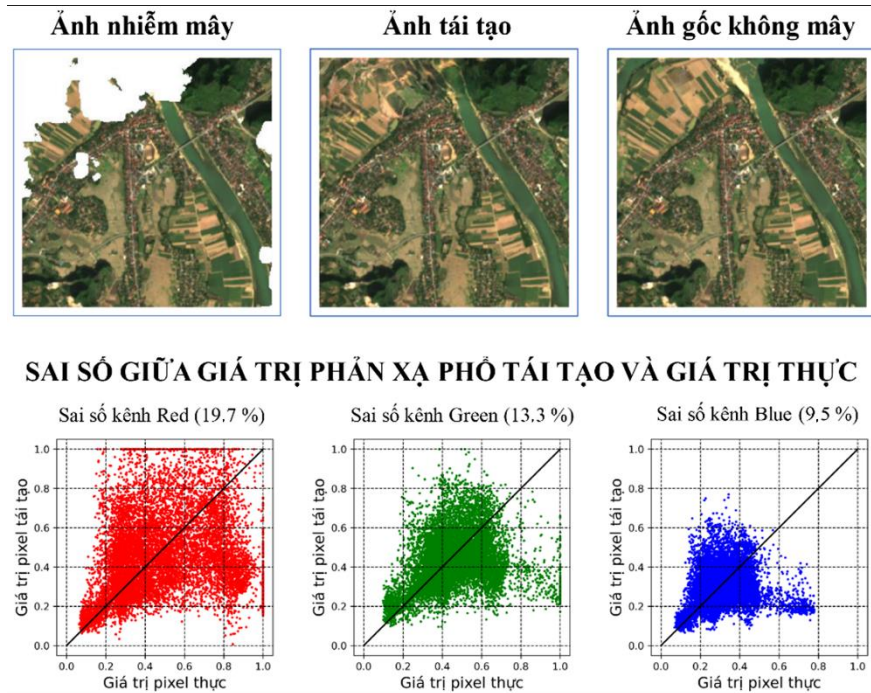
Trong nghiên cứu này, mô hình kết hợp DCNN và GAN với thuật toán mạng đối lập được đánh giá trên những ảnh Sentinel-2 có tỉ lệ nhiễm mây (mây và bóng mây) từ 15% - 25%. Đối với những ảnh có tỉ lệ nhiễm mây cao hơn, nhóm tác giả nhận thấy khả năng tái tạo dữ liệu bị giảm đi (theo giá trị của chỉ số PSNR và SSIM) và mô hình lúc này sẽ dự đoán vùng thông tin một cách ngẫu nhiên (Hình 8). Điều này có thể giải thích, bởi vì mô hình sẽ học thông tin từ các vùng xung quanh vùng nhiễm mây, nếu vùng thông tin này quá ít, mô hình sẽ không có khả năng dự đoán một cách chính xác. Đồng thời, với những ảnh có tỉ lệ nhiễm mây trung bình nhưng mật độ mây tập trung vào một vị trí, điều này có thể gây ra hiện tượng dự đoán lỗi như trên Hình 8 với lượng mây từ 25% - 30%. Nếu vùng mây phân bố rải rác thì mô hình vẫn có khả năng tái tạo với mức độ mây phủ từ 35% - 40%.

Tuy nhiên, đối với những ảnh Sentinel-2 bị nhiễm mây với tỉ lệ bao phủ lớn (> 50%), nhóm tác giả đề xuất sử dụng các cặp ảnh đa thời gian với mô hình sử dụng trong nghiên cứu này để đạt

được kết quả tốt nhất có thể. Bên cạnh đó, mô hình sử dụng trong nghiên cứu này gặp hạn chế khi tái tạo thông tin trên các khu vực có mức độ chi tiết cao và tính không đồng nhất (khu vực đô thị, dân cư). Với độ phân giải không gian 10 m, các đối tượng nhà, công trình xây dựng thường có kích thước nhỏ hơn 1 pixel, chính vì thế mô hình sẽ không có khả năng tái tạo các thông tin của những đối tượng này. Cùng với đó, nghiên cứu này cũng đánh giá thống kê giá trị phổ tái tạo cho ba kênh ảnh chính của ảnh Sentinel-2 (Red, Green, Blue) so với ảnh gốc ở các khu vực khác nhau. Trong các thí nghiệm tại khu vực đồi núi (Hình 9), đây là kiểu khu vực có các đối tượng lớp phủ liên tục đồng nhất, chính vì thế mô hình có khả năng tái tạo các kênh phổ với sai số tương đối thấp (0% - 5%) so với giá trị phản xạ phổ tại ảnh gốc. Ngược lại, ở khu vực đồng bằng, khu dân cư (Hình 10), nơi có các đối tượng phân bố không đồng nhất và liên tục, mô hình sẽ tái tạo dự đoán các giá trị phổ có xu hướng tạo ra sai số cao hơn (10% - 20%). Kênh ảnh Red của ảnh Sentinel-2 thường có giá trị sai số cao hơn kênh Green và kênh Blue.



Hình 9. Sai số phổ tái tạo tại khu vực đồi núi.



Hình 10. Sai số phổ tái tạo tại khu vực đồng bằng và khu dân cư.

4. Kết luận và kiến nghị

Trong nghiên cứu này, nhóm tác giả đã đề xuất lựa chọn phương pháp tái tạo vùng dữ liệu bị mây che phủ trên ảnh đơn Sentinel-2 bằng phương pháp học máy sâu với các mô hình DCNN và mạng đối lập GAN. Kỹ thuật inpainting kết hợp trong mô hình học máy sâu DCNN và mạng đối lập GAN cho ảnh đơn (duy nhất) bị nhiễm mây và bóng mây cho phép tái tạo lại dữ liệu bị thiếu bằng cách truyền thông tin hình học từ bên ngoài vùng mây. Phương pháp này loại bỏ các đám mây bằng cách tái tạo pixel bị nhiễm mây và bóng mây bằng các pixel tương ứng về mặt không gian gần nhất trong hình ảnh cục bộ không có mây. So với các phương pháp lọc mây truyền thống cần yêu cầu phải sử dụng nhiều ảnh đa thời gian để tái tạo dữ liệu, phương pháp được đề xuất có ưu điểm tự xác định các thông tin xung quanh đối tượng mây để dự đoán phần thông tin bị che phủ. Đây là một trong những điểm mạnh của thuật toán học máy

sâu, qua quá trình đào tạo, mô hình trong nghiên cứu này có khả năng đưa ra dự đoán thông tin vùng nhiễm mây với độ chính xác cao dựa trên kết quả PSNR trong một số cảnh ảnh thử nghiệm. Trong đó chúng tôi thấy sự cải thiện đáng kể về hiệu suất, do đó xác nhận hiệu quả của kết quả của mô hình đề xuất trong nghiên cứu. Với việc áp dụng kỹ thuật này, chất lượng dữ liệu ảnh Sentinel-2 ở Việt Nam vốn đã chịu ảnh hưởng bởi mây và chu kỳ bay chụp ngắn, sẽ được cải thiện đáng kể và giúp cho các nhà nghiên cứu viễn thám có thêm công cụ để làm dày bộ dữ liệu nghiên cứu và phục vụ cho mục đích giám sát tài nguyên và môi trường. Bên cạnh đó, việc ứng dụng phương pháp này cũng có hữu ích cho các khu vực bị thiếu hụt thông tin bằng hệ thống hình ảnh. Một trong những trường hợp được cân loại bỏ mây và bóng mây trên một hình ảnh như là phục vụ tiện ích cho ứng dụng Maps và dự án của Liên Hợp Quốc nhằm ứng phó với thiên tai bằng phân tích hình ảnh vệ tinh.

5. Lời cảm ơn

Nghiên cứu này được tài trợ bởi Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội trong đề tài mã số TN.21.14.

Tài liệu tham khảo

- [1] N. Q. Long, T. V. Anh, B. K. Luyen, Determination of Ground Subsidence By Sentinel-1 Sar Data (2018-2020) over Binh Duong Quarries, Vietnam, *Vnu Journal of Science: Earth and Environmental Sciences*, Vol. 37, No. 2, 2021, pp. 69-83, <https://doi.org/10.25073/2588-1094/vnuees.4605> (in Vietnamese).
- [2] C. H. Lin, P. H. Tsai, K. H. Lai, J. Y. Chen, Cloud Removal from Multitemporal Satellite Images Using Information Cloning, in *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 51, No. 1, 2013, pp. 232-241, <https://doi.org/10.1109/tgrs.2012.2197682>.
- [3] S. Huanfeng, L. H. Fang, Q. Yan., Z. L. Pei, Y. Q. Qiang, An Effective Thin Cloud Removal Procedure for Visible Remote Sensing Images, *Isprs Journal of Photogrammetry and Remote Sensing*, Vol. 96, 2012, pp. 224-235, <https://doi.org/10.1016/j.isprsjprs.2014.06.011>.
- [4] N. Q. Huy, P. V. Manh, N. X. Linh et al., Building A Travel Advisory System Using Webgis Technology for Tourism Development in Quang Binh, Vietnam, *Vnu Journal of Science: Earth and Environmental Sciences*, Vol 35, No. 3, 2019, pp. 108-114, <https://doi.org/10.25073/2588-1094/vnuees.4399> (in Vietnamese).
- [5] P. V. Dong, B. Q. Thanh, N. Q. Huy et al., Application of Deep Learning Algorithm to Build an Automated Cloud Segmentation Model Based on Open Data Cube Framework, *VNU Journal of Science: Earth and Environmental Sciences*, Vol. 36, No. 4, 2020, pp. 8-16, <https://doi.org/10.25073/2588-1094/vnuees.4441> (in Vietnamese).
- [6] O. Hagolle, M. Huc, D. V. Pascual, G. Dedieu, A Multi-temporal Method for Cloud Detection, Applied to FORMOSAT-2, VENUS, LANDSAT and SENTINEL-2 Images, *Remote Sensing of Environment*, Vol. 114, No. 8, 2010, pp. 1747-1755, <https://doi.org/10.1016/j.rse.2010.03.002>.
- [7] Z. Zhe, W. Shixiong, W. E. Curtis, Improvement and Expansion of the Fmask Algorithm: Cloud, Cloud Shadow, and Snow Detection for Landsats 4-7, 8, and Sentinel 2 Images, *Remote Sensing of Environment*, Vol. 159, 2015, pp. 269-277, <https://doi.org/10.1016/j.rse.2014.12.014>.
- [8] T. A. Tuan, N. D. Duong, Development of Landsat Cloud Free Image Data for Classification of Land Cover-case Study in Dak Lak Province, *VNU Journal of Science: Earth and Environmental Sciences*, Vol. 35, No. 4, 2019, pp. 80-87, <https://doi.org/10.25073/2588-1094/vnuees.4447>.
- [9] Q. Cheng, H. Shen, L. Zhang et al., Cloud Removal for Remotely Sensed Images by Similar Pixel Replacement Guided with A Spatio-Temporal MRF Model, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 92, 2014, pp. 54-68, <https://doi.org/10.1016/j.isprsjprs.2014.02.015>.
- [10] X. Zhang, F. Qin, Y. Qin, Study on the Thick Cloud Removal Method Based on Multi-temporal Remote Sensing Images, *International Conference on Multimedia Technology*, 2010, pp. 1-3, <https://doi.org/10.1109/ICMULT.2010.5631017>.
- [11] C. Yang, X. Lu, Z. Lin et al., High-Resolution Image Inpainting Using Multi-Scale Neural Patch Synthesis, 2017, pp. 6721-6729, <https://doi.org/10.1109/CVPR.2017.434>.
- [12] X. Zhu, Y. Qian, X. Zhao et al., A deep Learning Approach to Patch-based Image Inpainting Forensics, *Signal Processing: Image Communication*, Vol. 67, 2018, pp. 90-99, <https://doi.org/10.1016/j.image.2018.05.015>.
- [13] J. Yu, Z. Lin, J. Yang et al., Generative Image Inpainting with Contextual Attention, 2018, pp. 5505-5514, <https://doi.org/10.48550/arXiv.1801.07892>.
- [14] G. Liu, F. A. Reda, K. J. Shih et al., Image Inpainting for Irregular Holes Using Partial Convolutions, 2018, pp. 1-15, <https://doi.org/10.48550/arXiv.1804.07723>.
- [15] H. Shen, X. Li, Q. Cheng et al, Missing Information Reconstruction of Remote Sensing Data: A Technical Review, *IEEE Geosci Remote Sensing Mag*, Vol 3, No. 3, 2015, pp. 61-85, <https://doi.org/10.1109/MGRS.2015.2441912>.
- [16] H. Shen, J. Wu, Q. Cheng et al., A Spatiotemporal Fusion Based Cloud Removal Method for Remote Sensing Images with Land Cover Changes, *IEEE J Sel Top Appl Earth Observations Remote Sensing*, Vol. 12, No. 3, 2019, pp. 862-874, <https://doi.org/10.1109/JSTARS.2019.2898348>.
- [17] A. M. Martínez, E. I. Verdiguier, M. P. Maneta et al., Multispectral High Resolution Sensor Fusion for Smoothing and Gap-filling in the Cloud, *Remote Sensing of Environment*, Vol 247, 2020,

- pp. 1-19,
<https://doi.org/10.1016/j.rse.2020.111901>.
- [18] X. Zhu, Y. Qian, X. Zhao et al., A Deep Learning Approach to Patch-Based Image Inpainting Forensics, *Signal Processing: Image Communication*, Vol. 67, 2018, pp. 90-99, <https://doi.org/10.1016/j.image.2018.05.015>.
- [19] I. J. Goodfellow, J. P. Abadie, M. Mirza et al, Generative Adversarial Networks, arXiv: 1406.2661, 2014, pp. 1-9, <https://doi.org/10.48550/arXiv.1406.2661>.
- [20] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, arXiv:1505.04597, 2015, pp. 1-8, <https://doi.org/10.48550/arXiv.1505.04597>.
- [21] J. Korhonen, J. You, Peak Signal-to-noise Ratio Revisited: is Simple Beautiful?, 2012 Fourth International Workshop on Quality of Multimedia Experience, 2012, pp. 37-38, <https://doi.org/10.1109/QoMEX.2012.6263880>.
- [22] L. Zhang, L. Zhang, X. Mou, D. Zhang, FSIM: A Feature Similarity Index for Image Quality Assessment, in *IEEE Transactions on Image Processing*, Vol. 20, No. 8, 2011, pp. 2378-2386, <https://doi.org/10.1109/TIP.2011.2109730>.