

PHÁT HIỆN SẠT LỞ ĐẤT DỰA TRÊN DỮ LIỆU ẢNH UAV SỬ DỤNG AI TẠO SINH VÀ THUẬT TOÁN PHÂN ĐOẠN ĐIỂM ẢNH

Lê Đức Quyền⁽¹⁾, Phạm Minh Vương⁽²⁾

⁽¹⁾Viện Khoa học Khí tượng Thủy văn và Biến đổi khí hậu

⁽²⁾Khoa Kỹ thuật Xây dựng, Trường Đại học Kiến trúc Đà Nẵng

Ngày nhận bài: 2/2/2025; ngày chuyển phản biện: 3/2/2025; ngày chấp nhận đăng: 21/2/2025

Tóm tắt: Phát hiện các điểm sạt lở đất dựa trên hình ảnh từ phương tiện máy bay không người lái (UAV) được sử dụng rộng rãi cho việc lập bản đồ dự báo và đánh giá rủi ro cho các khu vực dễ xảy ra sạt lở đất. Các nghiên cứu gần đây đã tập trung vào việc sử dụng hình ảnh UAV có độ phân giải cao và các thuật toán học sâu để phát hiện các điểm sạt lở đất thay thế cho các phương pháp thủ công truyền thống. Tuy nhiên, trong thực tế, điều kiện ánh sáng thay đổi khiến chất lượng hình ảnh UAV giảm sút, làm giảm hiệu quả phát hiện vị trí sạt lở đất của các thuật toán học sâu. Vì vậy, nghiên cứu này đề xuất một phương pháp tiếp cận mới kết hợp thuật toán phân đoạn hình ảnh với trí tuệ nhân tạo (AI) tạo sinh để tăng cường độ chính xác phát hiện sạt lở đất dựa trên hình ảnh UAV trong điều kiện ánh sáng thay đổi. Trong nghiên cứu này, AI tạo sinh được áp dụng để nâng cao chất lượng hình ảnh đầu vào, giúp cải thiện khả năng nhận diện sạt lở đất. Thuật toán phân đoạn ảnh dựa trên mạng nơ-ron Unet được ứng dụng để xác định vị trí sạt lở đất trên hình ảnh tái tạo. Kết quả nghiên cứu cho thấy AI tạo sinh giúp cải thiện đáng kể chất lượng hình ảnh trong điều kiện ánh sáng kém. Phương pháp đề xuất giúp độ chính xác phát hiện sạt lở đất tăng đáng kể, với giá trị F1 đạt 0,878-0,891 trên ảnh tái tạo, so với chỉ 0,043-0,616 trên ảnh gốc. Kết quả này góp phần vào việc phát triển phương pháp tự động phát hiện và dự báo sạt lở đất dựa trên dữ liệu UAV cho các khu vực có nguy cơ sạt lở đất cao.

Từ khóa: Phát hiện sạt lở đất, AI tạo sinh, phân đoạn ảnh, UAV.

1. Mở đầu

Sạt lở đất ngày càng diễn ra thường xuyên và nghiêm trọng hơn do tác động của biến đổi khí hậu, gây ra mối đe dọa lớn đối với con người và cơ sở hạ tầng [1-3]. Việc xác định kịp thời các khu vực sạt lở đất giúp chính quyền có thể triển khai các biện pháp ứng phó khẩn cấp một cách nhanh chóng. Thông tin chi tiết về vị trí các điểm sạt lở đất đóng vai trò quan trọng trong phân tích biến dạng địa hình, đánh giá mức độ rủi ro, lập bản đồ sạt lở và phát triển các mô hình dự báo sạt lở đất trong tương lai.

Hiện nay, công nghệ phát hiện sạt lở đất dựa trên dữ liệu hình ảnh từ máy bay không người lái (Unmanned Aerial Vehicle - UAV) được ứng dụng rộng rãi. Hình ảnh UAV, được chụp từ góc nhìn trên cao, cung cấp dữ liệu có độ phân giải

cao, thường đạt mức vài centimet trên mỗi pixel. Độ phân giải cao này cho phép phân tích chính xác các đặc điểm địa hình. Ngoài ra, hình ảnh UAV đóng vai trò quan trọng trong mô hình hóa địa hình 3D, cung cấp thông tin chi tiết về cấu tạo địa hình, hỗ trợ nghiên cứu chuyên sâu về ổn định mái dốc [4-6].

Với sự phát triển nhanh chóng của trí tuệ nhân tạo (Artificial Intelligence - AI) và các thuật toán thị giác máy tính (Computer Vision), các phương pháp như phân loại ảnh, phát hiện đối tượng trong ảnh và phân đoạn ảnh đã được ứng dụng hiệu quả trong việc giám sát nguy cơ sạt lở đất (Landslide) và lũ bùn đá (Debris Flow) [7-9]. Trong số đó, các thuật toán phân đoạn hình ảnh thu hút sự chú ý đáng kể nhờ khả năng phân loại từng pixel trong ảnh, cho phép nhận diện sạt lở đất một cách chính xác và tự động. Ji và cộng sự (2020) [10] đã áp dụng mạng nơ-ron tích chập (Convolutional Neural Network - CNN)

Liên hệ tác giả: Phạm Minh Vương

Email: vuongpm@dau.edu.vn

để phát hiện sạt lở đất từ hình ảnh vệ tinh. Mô hình CNN được huấn luyện và kiểm thử trên bộ dữ liệu gồm 770 điểm sạt lở đất, đạt độ chính xác với F1-score lên đến 96,62%. Bragagnolo và cộng sự [11] sử dụng Unet để xác định vết sạt lở trong ảnh vệ tinh, đạt độ chính xác với F1-score là 67%. Sandric và cộng sự [12] kết hợp UAV với các mô hình Unet và DeepLab để khảo sát các vết nứt sạt lở trên diện rộng tại khu vực sạt lở Livadea, Romania, đạt độ chính xác trên 93% đối với hình ảnh UAV. Pham và Kim [13] đã đề xuất một phương pháp mới dựa trên mạng CNN sâu để xác định dòng chảy bùn đá và đo vận tốc của chúng từ hình ảnh kỹ thuật số thu được bằng camera cố định, với độ chính xác trung bình dao động từ 87,1% đến 97,3%.

Mặc dù đã có nhiều tiến bộ, các phương pháp phát hiện sạt lở đất hiện nay vẫn gặp hạn chế do phụ thuộc vào hình ảnh có độ phân giải cao. Nhiều thuật toán gặp khó khăn trong việc xử lý hình ảnh bị suy giảm chất lượng do điều kiện ánh sáng thay đổi, làm ảnh hưởng đến độ chính xác của mô hình AI. Giải quyết những thách thức này là điều cần thiết để nâng cao khả năng ứng dụng của các hệ thống giám sát sạt lở đất dựa trên AI trong điều kiện thực tế. Do đó, nghiên cứu này đề xuất một phương pháp mới tích hợp các thuật toán phân đoạn điểm ảnh với AI tạo sinh. Cụ thể, AI tạo sinh được sử dụng để cải thiện chất lượng hình ảnh đầu vào trong điều kiện thiếu sáng, sau đó thuật toán phân đoạn điểm ảnh dựa trên mạng nơ-ron Unet sẽ xác định vị trí sạt lở trên các hình ảnh đã được cải thiện. Cách tiếp cận này giúp tăng độ chính xác của các mô hình AI trong việc phát hiện sạt lở đất khi thu thập dữ liệu bằng UAV trong điều kiện ánh sáng kém. Bằng cách kết hợp công nghệ AI tiên tiến với hệ thống UAV, chúng ta có thể cải thiện đáng kể khả năng đánh giá rủi ro sạt lở đất và quản lý thiên tai một cách hiệu quả hơn.

2. Phương pháp nghiên cứu

Hình 1 trình bày sơ đồ nghiên cứu về phương pháp xác định vị trí sạt lở đất từ ảnh UAV bằng AI tạo sinh và các thuật toán phân đoạn ảnh. Quy trình nghiên cứu gồm hai bước chính: (1) AI tạo sinh được sử dụng để cải thiện chất lượng hình ảnh đầu vào; (2) Mô hình phân đoạn điểm

ảnh được áp dụng để phát hiện vị trí sạt lở đất. Chi tiết các bước được mô tả trong các phần dưới đây:

2.1. Cải thiện chất lượng hình ảnh đầu vào bằng AI tạo sinh

Việc cải thiện chất lượng ảnh đầu vào là một bước tiến xử lý quan trọng nhằm nâng cao độ chính xác của các mô hình AI trong việc nhận diện vị trí sạt lở, đặc biệt trong điều kiện ánh sáng thay đổi. Hình ảnh thiếu sáng thường xuất hiện khi UAV chụp vào sáng sớm, chiều tối hoặc trong điều kiện trời nhiều mây. Khi đó, tầm nhìn bị giảm và nhiễu ảnh tăng cao, gây khó khăn cho việc phân đoạn chi tiết các đối tượng trong ảnh [14]. Hình 2 minh họa sự phân bố cường độ pixel của hình ảnh thiếu sáng. Các hình ảnh này có giá trị độ sáng tập trung trong một phạm vi hẹp, với cường độ điểm ảnh dao động từ 0 đến 50, trong khi hình ảnh chụp trong điều kiện đủ sáng có cường độ điểm ảnh phân bố đều từ 50 đến 200.

Mạng đối sinh (Generative Adversarial Networks-GANs) đã trở thành một công cụ mạnh mẽ trong lĩnh vực xử lý ảnh nhờ khả năng tạo ra đầu ra chất lượng cao dựa trên nguyên lý khuếch tán pixel và phân phối dữ liệu [15]. Các mô hình dựa trên GAN có thể cải thiện đáng kể độ rõ nét, độ tương phản và mức độ chi tiết của hình ảnh. Hình 3 minh họa cấu trúc của GANs, bao gồm hai mạng nơ-ron chính: Mạng nơ-ron tạo ảnh (Generator- G): Học cách tạo ra hình ảnh chất lượng cao từ hình ảnh đầu vào chất lượng thấp. Mạng nơ-ron phân biệt ảnh (Discriminator-D): Phân biệt giữa ảnh thực và ảnh giả [16]. Hai mạng này được huấn luyện đồng thời theo thuật toán minmax, được định nghĩa theo Phương trình (1). Quá trình huấn luyện GANs sử dụng các phương pháp tối ưu hóa dựa trên hàm tối ưu gradient để cập nhật tham số của cả mạng nơ-ron tạo ảnh và mạng nơ-ron phân biệt ảnh nhằm phân loại ảnh thật và ảnh giả. Hàm mất mát của mô hình tạo ảnh LG và mô hình phân biệt ảnh LD được xác định trong Phương trình (2) và (3).

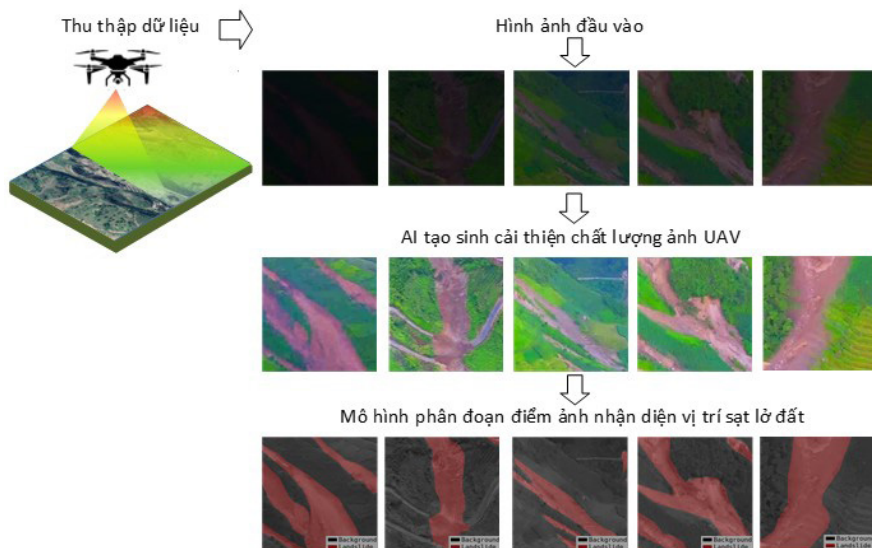
$$\min_G \max_D V(G, D) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (1)$$

$$L_D = -(\mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{x \sim p_z(z)} [\log(1 - D(G(z)))])$$

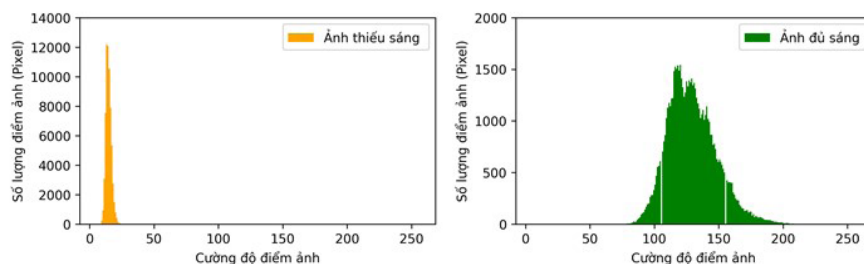
Trong đó: $\mathbb{E}_{x \sim p_{data}(x)} [\log D(x)]$ đại diện cho giá trị kỳ vọng của logarit xác suất mà mạng phân biệt ảnh gán cho các ảnh thực;

$$L_G = \mathbb{E}_{x \sim p_z(z)} [\log(1 - D(G(z)))]$$

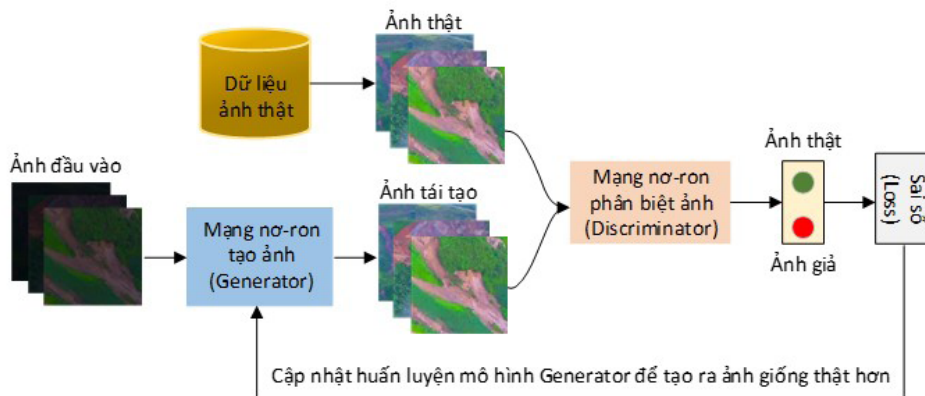
$\mathbb{E}_{x \sim p_z(z)} [\log(1 - D(G(z)))]$ đại diện cho giá trị kỳ vọng của logarit xác suất mà mạng phân biệt ảnh gán cho các ảnh giả được tạo ra bởi G.



Hình 1. Sơ đồ nghiên cứu về phát hiện sạt lở đất từ ảnh UAV bằng AI tạo sinh và thuật toán phân đoạn điểm ảnh



Hình 2. Sự phân bố cường độ điểm ảnh của hình ảnh thiếu sáng và hình ảnh bình thường



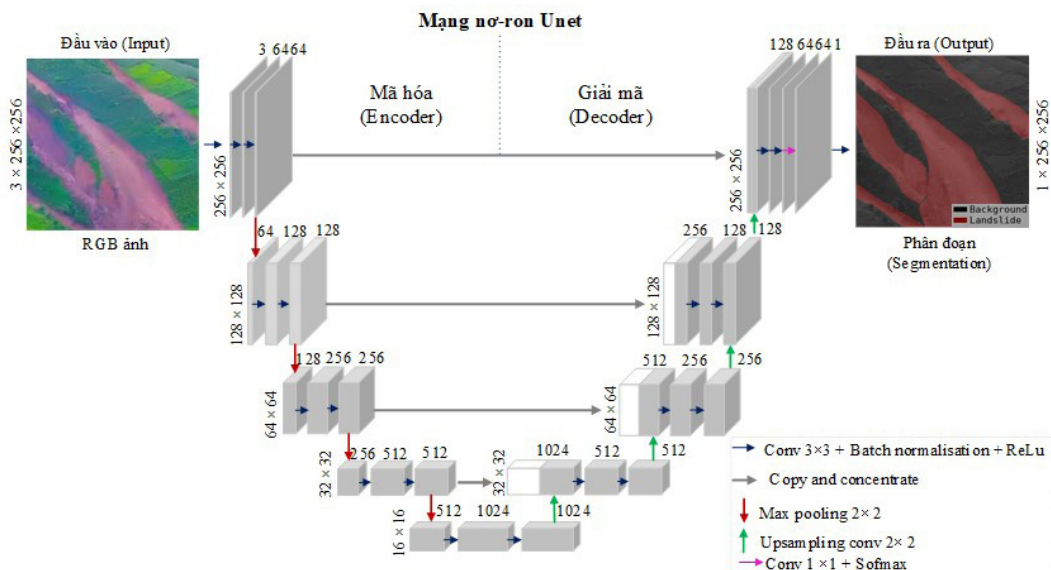
Hình 3. Kiến trúc của mạng nơ-ron nhân tạo GAN

2.2. Phát hiện vị trí sạt lở đất sử dụng thuật toán phân đoạn pixel

Trong nghiên cứu này, các mô hình phân đoạn ảnh dựa trên mạng nơ-ron Unet, bao gồm Res-Unet, VGG16-Unet, Mobile-Unet, Inception-Unet và Dense-Unet được lựa chọn để nhận diện sạt lở đất từ hình ảnh UAV. Những mô hình này sử dụng các kiến trúc nền tảng (backbone) khác nhau như ResNet, VGG16, MobileNet, Inception và DenseNet. ResNet [17] cải thiện quá trình lan truyền gradient nhờ các kết nối tắt (skip connections), giúp nâng cao hiệu suất thuật toán học sâu. Mạng VGG6 [18] có ưu điểm là kiến trúc đơn giản, dễ triển khai và phù hợp cho các tác vụ thị giác máy tính. Ngoài ra, nó có khả năng trích xuất đặc trưng hiệu quả nhờ các lớp tích chập sâu, giúp cải thiện độ chính xác trong nhận dạng hình ảnh. MobileNet [19] có kiến trúc gọn nhẹ, cho phép xử lý nhanh, phù hợp với các ứng dụng thời gian thực. Inception [20] có khả năng trích xuất đặc trưng ở nhiều mức tỷ lệ khác nhau, giúp cải thiện tính tổng quát của mô hình. DenseNet [21] tối ưu hóa việc tái sử dụng đặc trưng, giảm số lượng tham số nhưng vẫn tăng độ chính xác. Mạng nơ-ron Unet là một phương pháp học sâu tiên tiến, ban đầu được phát triển cho phân đoạn ảnh y sinh nhưng đã chứng minh hiệu quả ngay cả khi dữ

liệu huấn luyện bị giới hạn. So với các mạng CNN truyền thống, Unet cung cấp kết quả phân đoạn chính xác hơn [22].

Hình 4 minh họa chi tiết kiến trúc của mạng Unet. Unet tuân theo kiến trúc bộ mã hóa (encoder) - bộ giải mã (decoder), trong đó bộ mã hóa trích xuất các đặc trưng quan trọng từ ảnh đầu vào, còn bộ giải mã tái tạo ảnh đầu ra đã được phân đoạn. Mô hình này có cấu trúc hình chữ U, giúp duy trì thông tin không gian trong khi học dần các đặc trưng phân cấp. Các kết nối tắt (skip connections) giữa bộ mã hóa và bộ giải mã giúp bảo toàn chi tiết, làm cho Unet đặc biệt hiệu quả trong các bài toán phân loại từng pixel. Nhờ thiết kế này, Unet có thể hoạt động tốt ngay cả khi dữ liệu huấn luyện bị giới hạn. Bằng cách tận dụng cả đặc trưng cục bộ và toàn cục, Unet cải thiện đáng kể độ chính xác của phân đoạn, khiến nó trở thành một mô hình phổ biến trong lĩnh vực xử lý ảnh y sinh và viễn thám. Quá trình phân đoạn bao gồm việc phân loại từng pixel thuộc về khu vực sạt lở (landslide) hoặc nền ảnh (background). Mạng này nhận đầu vào là một ảnh RGB (Red-Green-Blue) có kích thước $3 \times 256 \times 256$ pixel và trích xuất đặc trưng thông qua các khối mã hóa - giải mã. Đầu ra là một ảnh nhị phân có kích thước $1 \times 256 \times 256$ pixel, trong đó mỗi pixel biểu thị xác suất thuộc về khu vực sạt lở hoặc nền ảnh.



Hình 4. Cấu trúc mạng nơ-ron nhân tạo Unet

2.3. Các tiêu chí đánh giá hiệu suất của mô hình AI tạo sinh và mô hình Unet

Để đánh giá độ chính xác của GANs trong nhiệm vụ tạo ảnh, chỉ số tương đồng cấu trúc (Structural Similarity Index- SSIM) được sử dụng. SSIM phản ánh mức độ tương đồng giữa ảnh được tạo và ảnh thực. Chỉ số SSIM giữa hai ảnh x và y được định nghĩa trong Phương trình (4).

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (4)$$

Trong đó: μ_x, μ_y lần lượt là giá trị trung bình của các điểm ảnh trên toàn bộ ảnh x và y ; μ_x^2, μ_y^2 lần lượt là phương sai của ảnh x và y ; σ_{xy} là hiệp phương sai giữa x và y ; C_1 và C_2 là các hằng số ổn định để tránh chia cho 0 ($C_1 = 0,012$ và $C_2 = 0,032$).

Để đánh giá hiệu quả của mô hình Unet, độ chính xác (Precision [P]), độ nhạy (Recall [R]) và giá trị điều hòa (F1-score) đã được sử dụng. Phương trình (5), (6) và (7) cung cấp công thức cho các tiêu chí này [23].

$$P = \frac{TP}{TP + FP} \quad (5)$$

$$R = \frac{TP}{TP + FN} \quad (6)$$

$$F1 = \frac{2 \times P \times R}{P + R} \quad (7)$$

Các điểm ảnh trong ảnh đầu ra của quá trình phân đoạn được phân loại như sau: Dương tính thật (True Positive [TP]): Các điểm ảnh được xác định chính xác là một phần của khu vực vực sạt lở. Dương tính giả (False Positive [FP]): Các điểm ảnh bị xác định nhầm là một phần của khu vực sạt lở (trên thực tế là nền). Âm tính thật (True Negative [TN]): Các điểm ảnh được xác định chính xác là nền. Âm tính giả (False Negative [FN]): Các điểm ảnh bị xác định nhầm

là nền (trên thực tế là một phần của khu vực sạt lở).

3. Dữ liệu hình ảnh

Bảng 1 và Bảng 2 trình bày tập dữ liệu hình ảnh sạt lở đất được sử dụng trong nghiên cứu này để huấn luyện và thử nghiệm các mô hình phân đoạn ảnh Unet (Res-Unet, VGG16-Unet, Mobile-Unet, Inception-Unet, Dense-Unet) và mô hình AI tạo sinh (GAN). Tập dữ liệu bao gồm hình ảnh sạt lở đất do UAV chụp tại huyện Bắc Hà, tỉnh Lào Cai, Việt Nam. Khu vực này đã trải qua nhiều vụ sạt lở đất do mưa lớn sau khi siêu bão Yagi đổ bộ vào miền Bắc Việt Nam vào tháng 9 năm 2024. Yagi là cơn bão mạnh nhất ảnh hưởng đến Việt Nam trong 70 năm qua, gây ra lũ lụt nghiêm trọng và sạt lở đất.

Hình ảnh UAV được thu thập bằng máy bay không người lái DJI Phantom 4 RTK, được trang bị camera RGB 20 MP. Tổng cộng, 165 điểm sạt lở đất đã được ghi nhận trong khu vực nghiên cứu. Để chuẩn bị dữ liệu cho mô hình học sâu, các hình ảnh được tiền xử lý và chia thành các ô vuông có kích thước 256×256 pixel. Kết quả tạo ra một tập dữ liệu gồm 3.538 hình ảnh sạt lở đất có gán nhãn, mỗi ảnh có độ phân giải 256×256 pixel.

Để xây dựng tập dữ liệu huấn luyện cho các mô hình AI tạo sinh, các biến thể hình ảnh chất lượng thấp đã được tạo ra bằng cách áp dụng thuật toán thay đổi độ sáng nhằm mô phỏng điều kiện ánh sáng yếu. Kết quả thu được 3.538 hình ảnh thiếu sáng với mức độ khác nhau. Hệ số điều chỉnh độ sáng được chọn ngẫu nhiên trong khoảng từ 0,1 đến 0,4 để tạo ra các hình ảnh thiếu sáng một cách đa dạng [24]. Các khu vực sạt lở thực tế được gán nhãn thủ công bằng công cụ Labelme trên toàn bộ 3.538 hình ảnh. Để đảm bảo sự cân bằng trong tập dữ liệu khi huấn luyện mô hình phân đoạn ảnh Unet và mô hình AI tạo sinh, các hình ảnh được chia ngẫu nhiên thành ba tập con: 80% (2.830 hình ảnh) để huấn luyện (training), 10% (354 hình ảnh) để xác thực (validation) và 10% (354 hình ảnh) để kiểm tra (testing).

Bảng 1. Dữ liệu được sử dụng để huấn luyện các mô hình phân đoạn ảnh (Res-Unet, VGG16-Unet, Mobile-Unet, Inception-Unet, DenseNet-Unet)

Dữ liệu	Số lượng ảnh	Số lượng nhãn sạt lở
Tổng số hình ảnh	3.538	3.538
Huấn luyện (80%)	2.830	2.830
Xác thực (10%)	354	354
Thử nghiệm (10%)	354	354

Bảng 2. Dữ liệu được sử dụng để huấn luyện mô hình AI tạo sinh (GAN)

Dữ liệu	Số lượng ảnh thiếu sáng	Số lượng ảnh trong điều kiện ánh sáng tốt
Tổng số hình ảnh	3.538	3.538
Huấn luyện (80%)	2.830	2.830
Xác thực (10%)	354	354
Thử nghiệm (10%)	354	354

4. Kết quả và thảo luận

4.1. Hiệu quả của mô hình AI tạo sinh

Các mô hình học sâu được huấn luyện trên hệ điều hành Windows 10, sử dụng CPU Intel Core i5-9600K và GPU NVIDIA GeForce RTX 4070, với phiên bản CUDA 11.1 và Python 3.9. Các tham số huấn luyện của mô hình AI tạo sinh (GAN) được thiết lập như sau: Tốc độ học (learning rate) là 0.0001, kích thước lô dữ liệu (batch size) là 5 và số vòng lặp huấn luyện (epoch) là 50.

Hình 5 trình bày hiệu quả của phương pháp AI tạo sinh trong việc nâng cao chất lượng hình ảnh từ các ảnh có điều kiện ánh sáng yếu. Khi thử nghiệm trên 354 ảnh thiếu sáng, giá trị SSIM dao động trong khoảng 0,734-0,941, cho thấy mức độ tương đồng cao giữa hình ảnh được tái tạo và hình ảnh thực tế. Kết quả này chứng minh rằng việc cải thiện hình ảnh bằng AI tạo sinh có thể nâng cao độ chính xác của các mô hình phát hiện sạt lở đất. Bằng cách giảm thiểu các biến dạng do điều kiện ánh sáng yếu, phương pháp này giúp các mô hình trích xuất tốt hơn các đặc trưng quan trọng liên quan đến sạt lở. Phương pháp đề xuất đặc biệt hữu ích trong các ứng dụng viễn thám và giám sát thảm họa, nơi mà chất lượng hình ảnh đóng vai trò quan trọng trong độ chính xác dự đoán của mô hình AI. Trong các nghiên cứu tương lai, kỹ thuật cải thiện hình ảnh này có thể được tích hợp với các mô hình học sâu để tối ưu hóa khả năng phát hiện sạt lở đất trong các điều kiện

thực tế phức tạp.

4.2. Hiệu quả của mô hình Unet trong việc nhận diện vị trí sạt lở đất

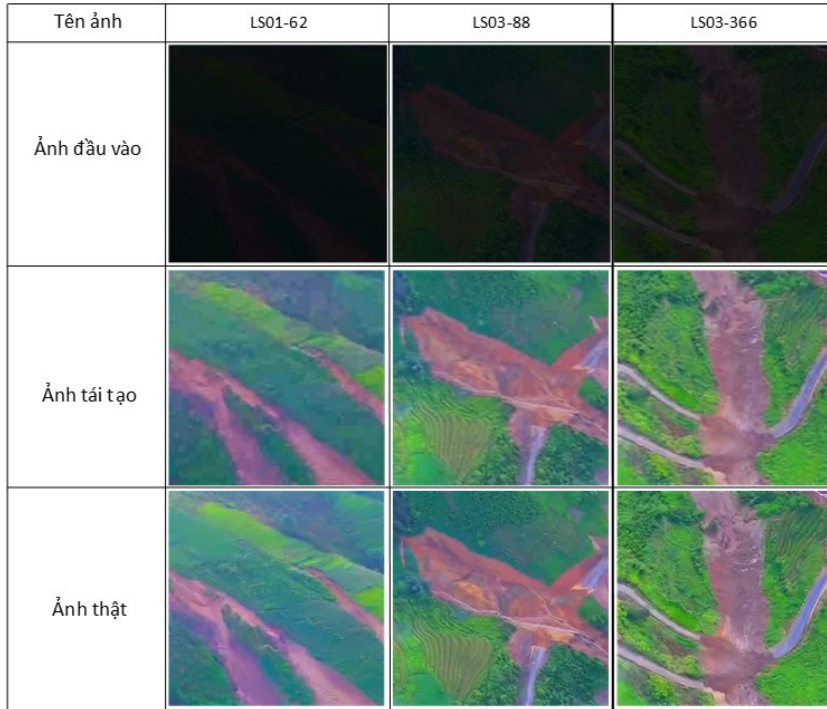
Hình 6 thể hiện hiệu suất của các mô hình Unet khác nhau (Res-Unet, VGG16-Unet, Mobile-Unet, Inception-Unet, Dense-Unet) trên tập dữ liệu ảnh thiếu sáng và ảnh đã được tái tạo. Độ chính xác (Precision) dao động từ 0,31 đến 0,679 đối với ảnh thiếu sáng và từ 0,842 đến 0,850 đối với ảnh tái tạo (Hình 6a). Độ nhạy (Recall) lần lượt đạt 0,026-0,799 và 0,910-0,948 (Hình 6b), trong khi giá trị F1 dao động từ 0,043-0,616 và 0,878-0,891 (Hình 6c). Kết quả cho thấy độ chính xác trong việc phát hiện sạt lở đất được cải thiện đáng kể khi sử dụng ảnh tái tạo so với ảnh đầu vào chất lượng thấp.

Trong số các mô hình, VGG16-Unet đạt hiệu suất cao nhất trên ảnh tạo sinh với điểm F1 là 0,891. Các mô hình Mobile-Unet, Inception-Unet, Res-Unet và Dense-Unet lần lượt đạt F1 là 0,887, 0,885, 0,878 và 0,878. Về tốc độ xử lý, VGG16-Unet, Inception-Unet, Res-Unet và Dense-Unet đạt lần lượt 12,6, 12,5, 10,5, 11,4 và 9,6 hình ảnh mỗi giây.

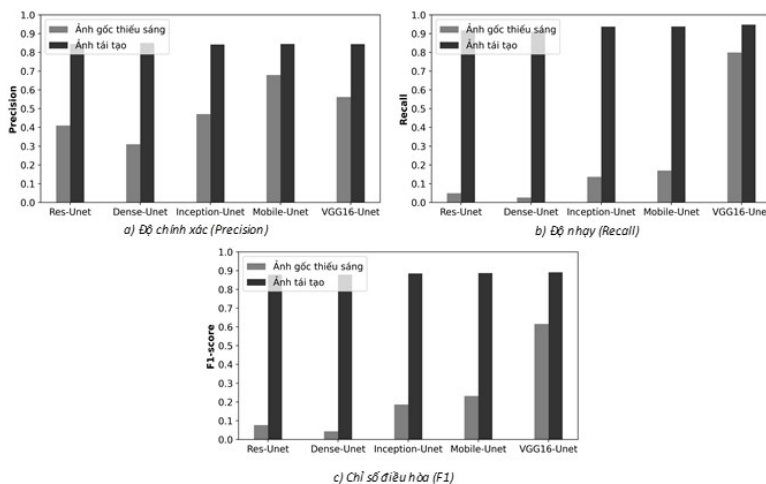
Hình 7 minh họa quá trình nhận diện sạt lở trên ảnh chất lượng thấp và ảnh đã được cải thiện bằng AI tạo sinh. Kết quả cho thấy độ chính xác trong phát hiện sạt lở tăng đáng kể trên ảnh tạo sinh, chứng minh rằng các kỹ thuật tăng cường ảnh dựa trên AI có thể khắc phục hiệu quả những hạn chế của dữ liệu đầu vào.

Giá trị F1 cao trên tất cả các biến thể của Unet khẳng định độ tin cậy của các mô hình khi áp dụng trên ảnh UAV đã được cải thiện. Hơn nữa, tốc độ xử lý cao của các mô hình này cho thấy tiềm năng ứng dụng trong thời gian thực. Những phát hiện này nhấn mạnh tính hiệu quả của việc kết hợp các mô hình phân đoạn với các kỹ thuật tăng cường hình ảnh nhằm nâng cao khả năng phát hiện sạt lở đất trong thực tế. Các nghiên cứu trong tương lai có thể tập trung vào

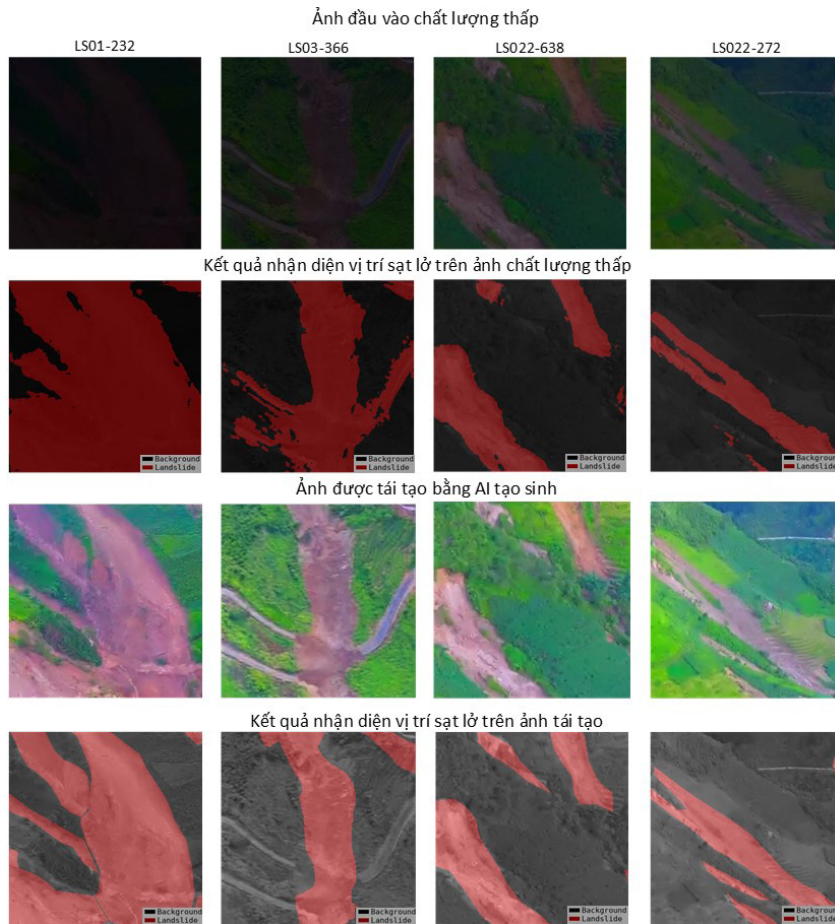
việc tối ưu hóa hình ảnh do GAN tạo ra để đạt độ chính xác cao hơn trong khi vẫn duy trì hiệu suất tính toán. Điều này có thể bao gồm việc tinh chỉnh kiến trúc GAN để tạo ra kết cấu và chi tiết thực tế hơn, đồng thời giảm thiểu các hiện tượng giả có thể ảnh hưởng đến hiệu suất mô hình. Ngoài ra, phương pháp này có thể được tích hợp vào quy trình xử lý dữ liệu UAV nhằm hỗ trợ các ứng dụng quản lý thảm họa tại những khu vực có nguy cơ sạt lở đất cao.



Hình 5. Kết quả sử dụng AI tạo sinh trong việc tạo ra hình ảnh chất lượng cao từ ảnh thiếu sáng



Hình 6. Hiệu quả của các mô hình Unet (Res-Unet, VGG16-Unet, Mobile-Unet, Inception-Unet, Dense-Unet) trên tập dữ liệu ảnh thiếu sáng và ảnh được tái tạo



Hình 7. Kết quả về nhận diện sạt lở trên ảnh UAV thiếu sáng và ảnh tạo sinh bằng mô hình VGG16-Unet

5. Kết luận

Nghiên cứu này đề xuất một phương pháp phát hiện sạt lở kết hợp giữa mạng nơ-ron học sâu và thuật toán AI tạo sinh nhằm nâng cao độ chính xác của các mô hình AI trong phát hiện sạt lở đất. Các kết quả chính thu được như sau:

Cách tiếp cận AI tạo sinh giúp cải thiện chất lượng hình ảnh trong điều kiện ánh sáng yếu, tạo ra ảnh có độ sắc nét cao và mức độ tương đồng cấu trúc mạnh mẽ so với ảnh thực tế. Chỉ số SSIM dao động từ 0,577 đến 0,921 trên tập dữ liệu thử nghiệm.

Sử dụng ảnh tạo sinh bằng GAN giúp tăng đáng kể độ chính xác trong phát hiện sạt lở so với ảnh gốc chất lượng thấp. Giá trị F1 đối với ảnh tái tạo dao động từ 0,878 đến 0,891, trong khi ảnh gốc chỉ đạt từ 0,043 đến 0,616.

Các nghiên cứu trong tương lai có thể tập trung vào tối ưu hóa hơn nữa hình ảnh do GAN tạo ra nhằm cải thiện độ chính xác trong việc nhận diện vị trí sạt lở. Ngoài ra, việc tích hợp AI tạo sinh với quy trình xử lý dữ liệu UAV có thể nâng cao khả năng ứng dụng của các hệ thống giám sát sạt lở đất dựa trên AI trong thực tế.

Đóng góp của từng tác giả trong bài báo: Xây dựng ý tưởng: Lê Đức Quyền, Phạm Minh Vương; Xử lý số liệu: Lê Đức Quyền, Phạm Minh Vương; Phương pháp: Lê Đức Quyền, Phạm Minh Vương; Viết bài: Lê Đức Quyền; Sửa bài: Phạm Minh Vương.

Lời cam đoan: Tập thể tác giả cam đoan bài báo này là công trình nghiên cứu của mình, chưa từng công bố trước đó, không sao chép, đạo văn; không có sự tranh chấp lợi ích trong nhóm tác giả.

Tài liệu tham khảo

1. Hungr, O (2007), "Dynamics of rapid landslides", *Progress in landslide science*, 47-57, doi 10.1007/978-3-540-70965-7_4/COVER.
2. Regmi, N.R. et al. (2015), "A Review of Mass Movement Processes and Risk in the Critical Zone of Earth", *Developments in Earth Surface Processes*, vol 19, 319-362, doi 10.1016/B978-0-444-63369-9.00011-2.
3. Nguyen, H.H.D. et al. (2024), "A hybrid approach combining physics-based model with extreme value analysis for temporal probability of rainfall-triggered landslide", *Landslides*, 22, 149-168, doi 10.1007/S10346-024-02366-X.
4. Cheng, Z. et al. (2021), "UAV photogrammetry-based remote sensing and preliminary assessment of the behavior of a landslide in Guizhou, China", *Engineering Geology*, vol 289, 106172, doi 10.1016/j.enggeo.2021.106172.
5. Nguyen, T.N. et al. (2024), "Measurement and prediction of scour volume around a cubic artificial reef under steady flow conditions using stereo vision", *Ocean Engineering*, vol 313, 119635, doi 10.1016/j.oceaneng.2024.119635.
6. Pham, M.V. et al. (2024), "Displacement Measurement and 3D Reconstruction of Segmental Retaining Wall Using Deep Convolutional Neural Networks and Binocular Stereovision", *Structural Control and Health Monitoring*, vol 2024, Article ID 9912238, 28 pages, doi 10.1155/2024/9912238.
7. Catani, F (2021), "Landslide detection by deep learning of non-nadir and crowdsourced optical images", *Landslides*, vol 18, no 3, 1025-1044, doi 10.1007/s10346-020-01513-4.
8. Ha, Y.S. et al. (2024), "Enhancements in image quality and block detection performance for Reinforced Soil-Retaining Walls under various illuminance conditions", *Advances in Engineering Software*, vol 195, 103713, doi 10.1016/j.advengsoft.2024.103713.
9. Nguyen, H.H.D. et al. (2025), "Uncertainty effect of saturated depth on large scale assessment of landslide: Case study at Saka town, Hiroshima Prefecture, Japan", *Earth Science Informatics*, vol 18, no 2, 1-14, doi 10.1007/S12145-025-01714-2/METRICS.
10. Ji, S. et al. (2020), "Landslide detection from an open satellite imagery and digital elevation model dataset using attention boosted convolutional neural networks", *Landslides*, vol 17, no 6, 1337-1352, doi 10.1007/s10346-020-01353-2.
11. Bragagnolo, L. et al. (2021), "Convolutional neural networks applied to semantic segmentation of landslide scars", *CATENA*, vol 201, 105189, doi 10.1016/J.CATENA.2021.105189.
12. Sandric, I. et al (2024), "Using high-resolution UAV imagery and artificial intelligence to detect and map landslide cracks automatically", *Landslides*, vol 21, no 10, 2535-2543, doi 10.1007/S10346-024-02295-9/FIGURES/5.
13. Pham M.V. and Kim Y.T. (2022), "Debris flow detection and velocity estimation using deep convolutional neural network and image processing", *Landslides*, vol 19, no.10, 2473-2488, doi 10.1007/S10346-022-01931-6/METRICS.
14. Jingchun, Z. et al. (2024), "Low-light image enhancement: A comprehensive review on methods, datasets and evaluation metrics", *Journal of King Saud University-Computer and Information Sciences*, vol 36, no 10, 102234, doi 10.1016/J.JKSUCI.2024.102234.
15. Frid-Adar, M. et al. (2018), "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification", *Neurocomputing*, vol 321, 321-331, doi 10.1016/J.NEUCOM.2018.09.013.
16. Creswell, A. et al. (2018), "Generative Adversarial Networks: An Overview", *IEEE signal processing magazine*, vol 35, no 1, 53-65, doi 10.1109/MSP.2017.2765202.
17. He, K. Et al. (2016), "Deep residual learning for image recognition", *In Proceedings of the IEEE conference on computer vision and pattern recognition*, vol 2016, 770-778, doi 10.1109/CVPR.2016.90.

18. Simonyan, K. and Zisserman, A. (2015), "Very deep convolutional networks for large-scale image recognition", *arXiv preprint arXiv*, 1-14, doi 10.48550/arXiv.1409.1556.
19. Howard, A.G. et al. (2017), "Mobilenets: Efficient convolutional neural networks for mobile vision applications", *arXiv preprint arXiv*, 1704.04861, doi 10.48550/arXiv.1704.04861.
20. Szegedy, C. et al. (2014), "Rethinking the Inception Architecture for Computer Vision", *In Proceedings of the IEEE conference on computer vision and pattern recognition*, 2818-2826, doi 10.48550/arXiv.1512.00567.
21. G. Huang and K. Q. Weinberger (2016), "Densely Connected Convolutional Networks", *Computer Vision and Pattern Recognition*, v1, <https://doi.org/10.48550/arXiv.1608.06993>
22. Ronneberger, O. (2015), "U-net: Convolutional networks for biomedical image segmentation", *In Medical image computing and computer-assisted intervention-MICCAI 2015*, 234-241, doi 10.1007/978-3-319-24574-4_28.
23. Pham, M.V. et al. (2023), "Automatic detection and measurement of ground crack propagation using deep learning networks and an image processing technique", *Measurement*, vol 215, 112832, doi 10.1016/j.measurement.2023.112832.
24. Chen, C. et al. (2018), "Learning to See in the Dark", *In Proceedings of the IEEE conference on computer vision and pattern recognition*, 3291-3300, doi 10.1109/CVPR.2018.00347.

LANDSLIDE DETECTION FROM UAV IMAGERY USING GENERATIVE AI AND PIXEL SEGMENTATION ALGORITHMS

Le Duc Quyen⁽¹⁾, Pham Minh Vuong⁽²⁾

⁽¹⁾The Viet Nam Institute of Meteorology, Hydrology, and Climate Change

⁽²⁾School of Civil Engineering, Danang Architecture University

Received: 2/2/2025; Accepted: 21/2/2025

Abstract: *Landslide detection based on images from unmanned aerial vehicles (UAVs) has been widely used for landslide mapping and risk assessment. Recent studies have focused on utilizing high-resolution UAV images and deep learning algorithms to detect landslides as an alternative to traditional manual methods. However, in practice, varying lighting conditions degrade UAV image quality, reducing the performance of deep learning-based landslide detection. Therefore, this study proposes a novel approach that integrates image segmentation algorithm with generative artificial intelligence (AI) to enhance landslide detection accuracy in UAV images under changing lighting conditions. In this research, generative AI was applied to improve the quality of input images, enhancing landslide recognition capabilities. A Unet-based image segmentation algorithms were employed to identify landslide locations on reconstructed images. The results indicate that generative AI significantly improves image quality under low-light conditions. The proposed method significantly enhances detection accuracy, achieving an F1-score between 0.878 and 0.891 on reconstructed images, compared to only 0.043 to 0.616 on original images. These findings contribute to the advancement of automated landslide detection and forecasting methods using UAV data for high-risk landslide areas.*

Keywords: *Landslide detection, Generative AI, Image segmentation, UAV.*