

NGHIÊN CỨU ỨNG DỤNG ẢNH VỆ TINH SENTINEL 2 ĐỂ XÁC ĐỊNH VỊ TRÍ TRƯỢT LỞ ĐẤT BẰNG MÔ HÌNH PHÂN LOẠI RANDOM FOREST

Đoàn Viết Long¹, Nguyễn Chí Công¹, Phạm Thành Hưng¹, Nguyễn Tiến Cường²

Tóm tắt: Trong nghiên cứu về trượt lở đất ở Việt Nam, nguồn dữ liệu về hiện trạng trượt lở chưa được thu thập đầy đủ do khó khăn trong công tác đo đạc xác định vị trí và thời gian trượt lở. Với sự phát triển của khoa học quan sát Trái đất và khoa học máy tính, công nghệ xử lý ảnh viễn thám có thể giải quyết vấn đề này. Nghiên cứu này sử dụng kỹ thuật phân tích ảnh viễn thám Sentinel 2 để nhận dạng vị trí trượt lở. Hai ảnh vệ tinh trước và sau khi xảy ra trượt lở được sử dụng để phân tích. Vị trí các điểm trượt lở được xác định dựa trên phân tích sự thay đổi của chỉ số thảm thực vật NDVI, sử dụng mô hình phân loại Random Forest (RF) và kỹ thuật chồng chấp bản đồ. Kết quả kiểm định cho thấy mô hình RF cho độ chính xác toàn cục đạt 98.2% và chỉ số Kappa đạt 0.95. Kết quả kiểm chứng tại 2 vị trí trượt lở bằng dữ liệu ảnh chụp thực tế đã cho thấy khả năng áp dụng của mô hình này.

Từ khóa: Trượt lở đất, Sentinel 2, phần mềm SNAP, Random Forest, NDVI

1. GIỚI THIỆU CHUNG

Trượt lở đất được định nghĩa là sự dịch chuyển của khối đất, đá trên sườn dốc dưới tác động của trọng lực (Cruden & Varnes, 1996). Đây là loại hình thiên tai khá phổ biến ở trên thế giới, gây hậu quả hết sức nghiêm trọng (Reichenbach et al., 2018). Trong nghiên cứu về hiện tượng này, công tác thu thập dữ liệu các vụ trượt lở (vị trí, thời gian, loại hình, quy mô ...) đã xảy ra là rất quan trọng. Việc này có thể được thực hiện bằng khảo sát đo đạc thực tế hoặc thông qua phân tích dữ liệu ảnh từ máy bay (Scaioni et al., 2014). Quá trình này tiêu tốn khá nhiều thời gian khi phải thực hiện trên một phạm vi rất rộng với địa hình phức tạp (Qin et al., 2018). Ngày nay, với sự phát triển của khoa học quan sát Trái đất, kỹ thuật viễn thám đã được ứng dụng giúp nhận diện nhanh các vụ trượt lở đất (Guzzetti et al., 2012). Ảnh viễn thám có thể thu thập từ nhiều nguồn vệ tinh khác nhau,

có thể phân chia làm 2 nhóm: ảnh quang học và ảnh radar. Ảnh quang học có thể lấy từ các vệ tinh như: Landsat 4, 5, 7 và 8, Sentinel-2A/B (Amatya et al., 2019; Budha & Bhardwaj, 2019; Li et al., 2016; Qin et al., 2018; Zhao et al., 2017). Trong khi đó, ảnh radar có thể thu thập từ vệ tinh radar khẩu độ tổng hợp (Synthetic Aperture Radar - SAR) như Sentinel-1A/B, ERS-1/2, ENVISAT, ALOS PALSAR, RADASAT-2 (Barra et al., 2016; Liu et al., 2018; Plank et al., 2016; Strozzi et al., 2018). Ưu điểm của ảnh quang học so với ảnh SAR là có độ phân giải cao hơn, các bước tiền xử lý và phân tích đơn giản hơn. Do đó, ảnh quang học được sử dụng nhiều hơn trong các trường hợp phân tích rủi ro thiên tai khi cần kết quả nhanh chóng.

Có nhiều phương pháp nhận dạng vết trượt lở đất từ ảnh không gian, có thể chia làm 5 nhóm: (i) Dựa vào trực quan (visual interpretation-based), (ii) dựa vào đặc trưng (feature-based), (iii) dựa vào phát hiện sự thay đổi (change detection-based), (iv) dựa vào mô phỏng địa hình (topographic model-based) và (v) phương

¹ Khoa Xây dựng Công trình thủy, Trường Đại học Bách khoa, Đại học Đà Nẵng

² Khoa Kỹ thuật Ô tô và Năng lượng, Trường Đại học Phenikaa

pháp học máy (machine learning-based). Trong đó, tùy thuộc vào loại dữ liệu ảnh và đặc trưng của từng khu vực nghiên cứu để chọn phương pháp phân tích phù hợp (Li et al., 2016). Zhao et al. (2017) đã ứng dụng phương pháp phát hiện sự thay đổi để phát hiện các vụ trượt lở sau động đất dựa trên dữ liệu ảnh Landsat 8, kết quả thu được khá chính xác khi so sánh với dữ liệu ảnh có độ phân giải cao từ Google Earth, tuy nhiên độ phân giải của ảnh vẫn khá thô (30m x 30m). Qin et al. (2018) đã sử dụng ảnh Sentinel 2 có độ phân giải tốt hơn (10m x 10m) và phương pháp CDRMF (Change Detection-based Markov Random Field) để phát hiện sự thay đổi dữ liệu ảnh vệ tinh ở thời điểm trước và sau các vụ trượt lở, kết quả thu được khá tốt với nhiều điểm trượt lở nhỏ đã được xác định. Phương pháp nhận dạng vết trượt lở dựa vào phân tích địa hình được ứng dụng cho các nguồn ảnh có chất lượng tốt và độ phân giải cao. Nhiều nghiên cứu đã sử dụng nguồn dữ liệu LiDAR trong phân tích và thu được kết quả tốt, tuy nhiên dữ liệu ảnh này thường không phổ biến và có chi phí cao (Bernat Gazibara et al., 2019; Bunn et al., 2019; Görüm, 2019; Pradhan et al., 2020). Với sự phát triển của khoa học máy tính, kỹ thuật học máy cũng đã được ứng dụng và cho thấy sự hiệu quả cao trong việc nhận dạng các vụ trượt lở đất. Các mô hình học máy được sử dụng trong các nghiên cứu này có thể kể đến như MLP-NN (Multilayer Perceptron Neural Network), RF (Random Forest), SVM (Support Vector Machine). (Ghorbanzadeh et al., 2019; Prakash et al., 2020; Tavakkoli Piralilou et al., 2019).

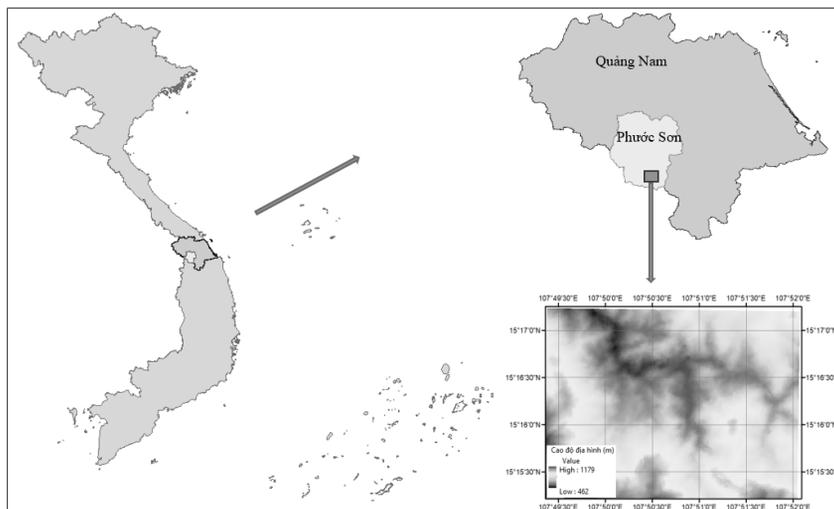
Ở Việt Nam, ứng dụng ảnh viễn thám để khôi phục dữ liệu trượt lở đất đã được một số nghiên cứu thực hiện (Le & Kawagoe, 2017; Van et al., 2017). Le & Kawagoe (2017) đã sử dụng ảnh Landsat 8 (độ phân giải 30x30m) và tính toán

chỉ số NDIs (Normalized Difference Indices) để xác định các điều kiện của thực vật, đất và nước. Qua đó nhận dạng các điểm trượt lở ở các tỉnh miền núi phía Bắc của Việt Nam. Kết quả nghiên cứu đã chỉ ra các điểm trượt lở và phân nhóm dựa vào nguyên nhân gây trượt. Tuy nhiên độ phân giải của ảnh vệ tinh vẫn còn ở mức thô (30x30m). Van et al. (2017) đã thực hiện lập bản đồ hiện trạng trượt lở đất ở 14 tỉnh miền núi phía Bắc - Việt Nam bằng phương pháp khảo sát, dữ liệu ảnh kỹ thuật số và ảnh vệ tinh. Tuy nhiên nghiên cứu này cũng nhận thấy một số hạn chế về thiếu nguồn dữ liệu ảnh, thiếu chuyên gia phân tích, khó xác minh thông tin lở đất tại các vị trí khó tiếp cận và khó khăn trong khâu cập nhật số liệu.

Trượt lở đất ở Việt Nam chủ yếu xảy ra ở các tỉnh miền núi phía Bắc và khu vực miền Trung – Tây Nguyên (Đoàn Viết Long, 2020). Đặc biệt, trượt lở đất đã xảy ra với mức độ chưa từng thấy, gây hậu quả nặng nề cho khu vực các tỉnh miền Trung trong năm 2020. Với mục tiêu cung cấp thông tin về vị trí xảy ra các vụ trượt lở đất bằng các nguồn dữ liệu tiếp cận miễn phí, bài báo này áp dụng kỹ thuật phân tích ảnh vệ tinh Sentinel 2 và thuật toán phân loại Random Forest trong việc xác định vị trí trượt lở đất.

2. KHU VỰC NGHIÊN CỨU

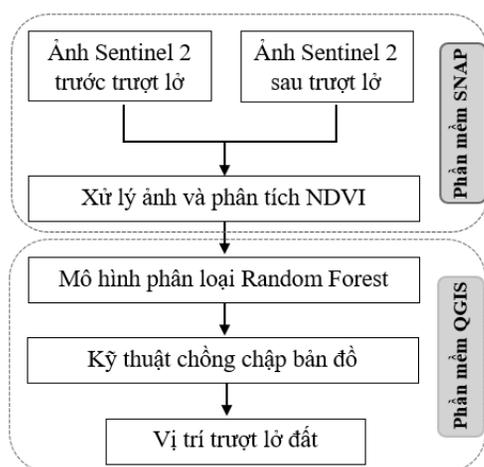
Khu vực nghiên cứu là vùng núi thuộc xã Phước Lộc, huyện Phước Sơn - tỉnh Quảng Nam, nơi xảy ra rất nhiều vụ trượt lở đất vào tháng 10 năm 2020. Đây là khu vực có địa hình hết sức phức tạp với 95% diện tích là rừng núi, do đó rất khó tiếp cận để khảo sát thực địa. Một khu vực có diện tích khoảng 17,837 km² (tọa độ địa lý từ 107°49'25"- 107°52'03" độ kinh Đông và từ 15°15'10" - 15°17'10" độ vĩ Bắc), cao độ địa hình từ 462m - 1.179m được lựa chọn để phân tích (Hình 1).



Hình 1. Vị trí khu vực nghiên cứu

3. PHƯƠNG PHÁP NGHIÊN CỨU

Ảnh vệ tinh Sentinel 2 có chất lượng tốt ít bị ảnh hưởng bởi mây được thu thập ở các thời điểm trước và sau vụ trượt lở. Nghiên cứu sử dụng phần mềm SNAP để xử lý ảnh, tính toán chỉ số NDVI. Sau đó, thuật toán phân loại RF (tích hợp trong phần mềm QGIS) được sử dụng để phân loại bản đồ NDVI thành 2 nhóm: có thảm phủ và không có thảm phủ. Các vết trượt lở được nhận dạng dựa vào kỹ thuật chồng chập bản đồ, được thực hiện trên phần mềm QGIS. Sơ đồ nghiên cứu được trình bày ở Hình 2.



Hình 2. Sơ đồ phương pháp nghiên cứu

3.1. Thu thập và xử lý ảnh vệ tinh Sentinel 2

Vệ tinh Sentinel-2 là một hệ thống quan sát Trái đất đa quang phổ được thực hiện bởi tổ chức

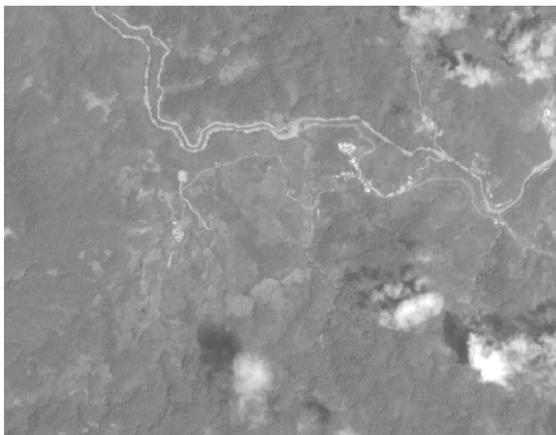
Giám sát Môi trường và An ninh toàn cầu (GMES) và được đồng tổ chức bởi Ủy ban Châu Âu (EC) và Cơ quan Vũ trụ Châu Âu (ESA) (Qin et al., 2018). Hệ thống này bao gồm hai vệ tinh quay quanh cực (tức là Sentinel-2A và Sentinel-2B) trên cùng một quỹ đạo nhưng lệch pha nhau 180°. Các vệ tinh mang thiết bị đa quang phổ (MSI) tạo mẫu gồm 12 dải phổ (dải 10 m, 20 m, 60 m) từ dải nhìn thấy và hồng ngoại gần (VNIR) đến dải hồng ngoại sóng ngắn (SWIR). Độ phân giải về thời gian là 5 ngày nếu kết hợp cả hai vệ tinh. Các cấp sản phẩm Cấp-1C và Cấp-2A được cung cấp miễn phí cho người dùng để giám sát đất đai, quản lý khẩn cấp và lập bản đồ rủi ro.

Đối với khu vực nghiên cứu này, dữ liệu ảnh Cấp-1C được thu thập từ web-site của Cơ quan Vũ trụ Châu Âu (ESA) <https://scihub.copernicus.eu/>. Dữ liệu ảnh vệ tinh Sentinel-2B với độ phân giải cao nhất 10mx10m và chu kỳ lặp lại là 10 ngày sẽ được thu thập để phân tích. Để đảm bảo chất lượng trong quá trình phân tích, nghiên cứu chỉ lựa chọn những bức ảnh có tỷ lệ mây che phủ dưới 20%. Ngày 28/10/2020, liên tiếp nhiều vụ trượt lở được ghi nhận trên địa bàn huyện Phước Sơn do ảnh hưởng của mưa lớn. Trên cơ sở đó, nghiên cứu này đã lựa chọn 2 ảnh vệ tinh Sentinel 2 có chất lượng tốt và được chụp vào thời điểm trước trượt lở (ngày 14/09/2020, cấp 1C) và sau thời

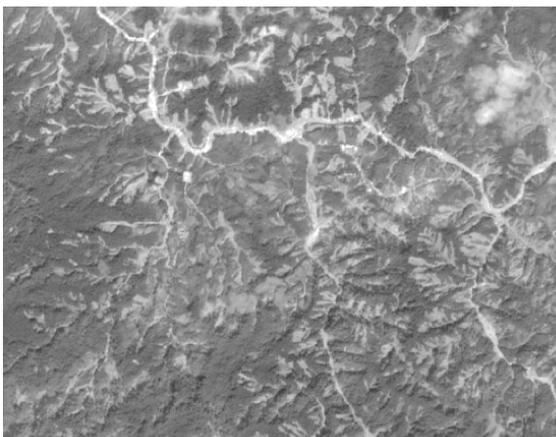
điểm trượt lở (ngày 28/12/2020, cấp 2A) để phân tích (Hình 3 và Hình 4)

Bảng 1. Thông số chính các dải phổ dùng để phân tích NDVI của ảnh vệ tinh Sentinel 2B

Dải phổ	Bước sóng (nm)	Độ phân giải (m)
Band 4 – Red	664.9	10
Band 8 – NIR	832.9	10



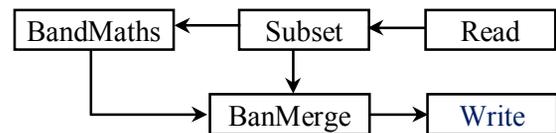
Hình 3. Ảnh Sentinel 2 trước vụ trượt lở (ngày 14/09/2020)



Hình 4. Ảnh Sentinel 2 trước vụ trượt lở (ngày 28/12/2020)

Ảnh Sentinel 2 thu thập được từ dữ liệu vệ tinh mới ở dạng thô. Để đáp ứng mục đích sử dụng trong nghiên cứu này, ảnh viễn thám phải trải qua

các công đoạn xử lý sau: (i) hiệu chỉnh khí quyển (Atmospheric correction) đối với ảnh cấp 1C, (ii) lựa chọn khu vực cần phân tích, (iii) phân tích và trích xuất ảnh NDVI có độ phân giải 10m x 10m. Tùy thuộc vào đặc điểm bề mặt, chỉ số NDVI của ảnh sau khi được phân tích sẽ nhận giá trị trong khoảng (-1, 1). Sơ đồ các bước thực hiện trên phần mềm SNAP được thể hiện trong Hình 5.



Hình 5. Các bước thực hiện phân tích NDVI

Chú thích các bước:

Read: đọc dữ liệu từ ảnh Sentinel-2 Level 2A (sau khi chuyển từ Level-1C).

Subset: chọn vùng nghiên cứu dựa vào region of interest (ROI).

Band Maths: Tính toán chỉ số NDVI, theo công thức (1):

$$NDVI = \frac{NIR - Red}{NIR + Red} = \frac{B8 - B4}{B8 + B4} \quad (1)$$

Với: NIR và Red lần lượt là phổ phản xạ của bước sóng cận hồng ngoại và bước sóng màu đỏ, thông số các dải này được cho ở Bảng 1.

BandMerge: ghép các bands vào thành một tệp dữ liệu.

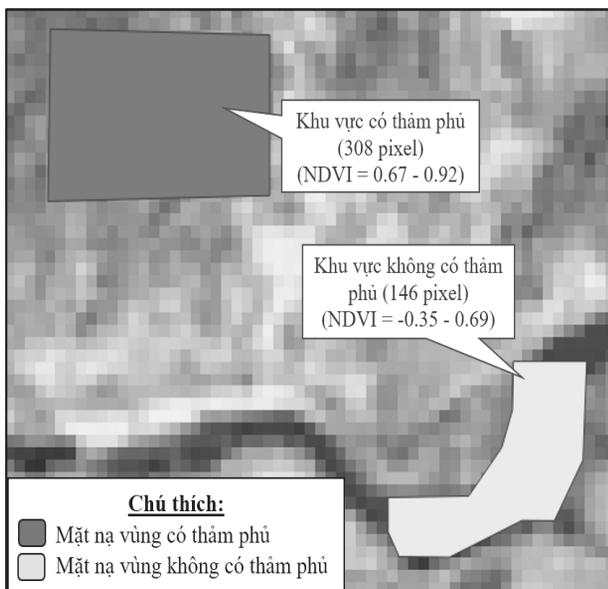
Write: xuất kết quả của quá trình xử lý.

3.2. Phân loại ảnh bằng thuật toán RF

Thuật toán RF được đề xuất bởi Breiman (2001). Đây là một thuật toán học có giám sát, được sử dụng cho cả phân lớp và hồi quy dựa trên cây quyết định. Chi tiết thông tin về thuật toán này có thể tham khảo trong nghiên cứu của Breiman, (2001).

Trong nghiên cứu này, vết trượt lở sẽ được nhận dạng dựa vào sự thay đổi chỉ số NDVI của ảnh trước và sau sự kiện. Do địa hình khu vực nghiên cứu có độ dốc lớn, các suối nhỏ và dốc, bài báo này chỉ thực hiện phân tích sự biến động thông qua 2 đặc trưng: (i) có thảm phủ thực vật, (ii) không có thảm phủ thực vật. Dữ liệu dùng

để xây dựng mô hình được lấy từ 2 vùng được lựa chọn và đã được kiểm chứng với dữ liệu ảnh từ Google Earth, bao gồm: (i) vùng có thảm phủ ở tọa độ (15°16'45'' độ vĩ Bắc, 107°50'36'' độ kinh Đông) và (ii) vùng không có thảm phủ ở tọa độ (15°16'45'' độ vĩ Bắc, 107°50'36'' độ kinh Đông) (Hình 6). Trong đó, 50% số điểm (154 điểm có thảm phủ và 73 điểm không có thảm phủ) được chọn ngẫu nhiên để huấn luyện và tạo ra bộ thông số cho mô hình dự báo. Việc kiểm định mô hình được thực hiện dựa trên 50% số điểm còn lại. Quá trình xây dựng mô hình bằng thuật toán RF được thực hiện trên công cụ Train algorithm trong phần mềm QGIS (<https://github.com/lennekade/dzetsaka>). Chỉ số độ chính xác toàn cục và chỉ số Kappa được sử dụng để đánh giá độ chính xác của mô hình dự báo. Mô hình có độ chính xác cao sẽ được áp dụng để phân loại cho toàn bộ ảnh trước và sau sự kiện. Quá trình này được thực hiện bằng công cụ Predict model trong QGIS.



Hình 6. Lựa chọn dữ liệu xây dựng mô hình:
 (i) Có thảm phủ (màu tím, ID=1);
 (ii) Không có thảm phủ (màu xanh lá, ID=0).

3.3. Xác định vị trí trượt lở

Các ảnh trước và sau sự kiện trượt lở khi

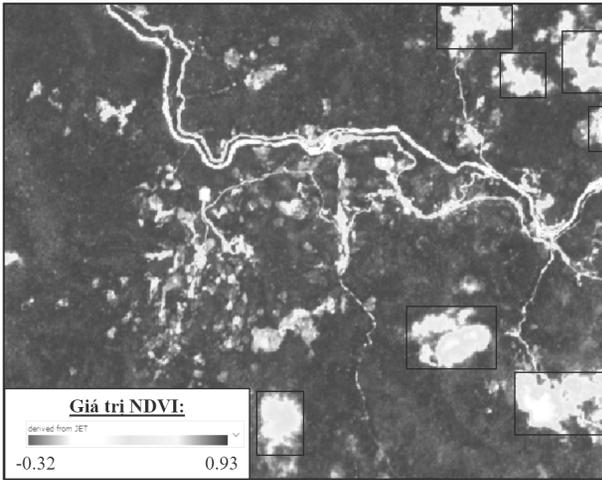
được phân loại sẽ được chồng chập với nhau để nhận diện sự thay đổi. Vùng mất đi thảm phủ xuất hiện ở ảnh sau sự kiện so với ảnh trước sự kiện được nhận diện là 1 vết trượt lở. Kỹ thuật phân lớp và chồng chập ảnh được thực hiện bằng phần mềm QGIS.

4. KẾT QUẢ VÀ THẢO LUẬN

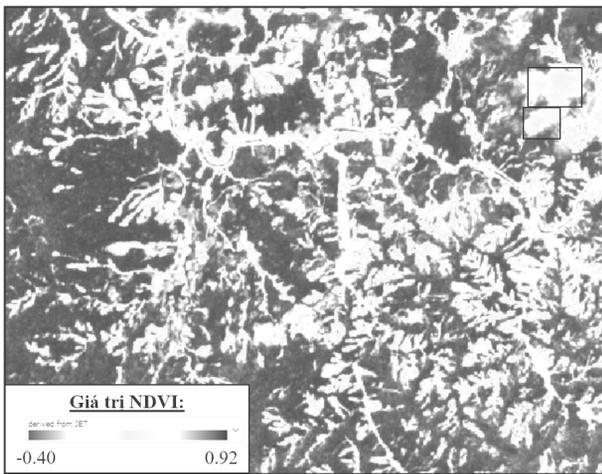
Kết quả phân tích NDVI với độ phân giải 10mx10m trước và sau sự kiện trượt lở được thể hiện lần lượt trong Hình 7 và Hình 8. Giá trị NDVI thay đổi trong khoảng từ -0.40 (biểu thị mặt nước) đến 0.92 (biểu thị đất rừng). Sự khác biệt giữa kết quả phân tích NDVI trước và sau sự kiện cho thấy sự thay đổi về thảm phủ thực vật trước và sau sự kiện trượt lở. Với độ phân giải 10m x 10m, các vết trượt lở nhỏ từ 100m² có thể được xác định. Đây có thể xem là một ưu điểm lớn của ảnh Sentinel 2 so với ảnh Landsat 8 với độ phân giải thô hơn (30m x 30m) đã áp dụng trong một số nghiên cứu trước đây (Le & Kawagoe, 2017, 2018). Ngoài ra, ảnh Sentinel 2 A/B có chu kỳ lặp lại là 5 ngày sẽ cũng cấp nhiều dữ liệu ảnh hơn so với ảnh Landsat 8 với chu kỳ lặp lại là 16 ngày.

Một hạn chế của ảnh quang học là bị che khuất bởi mây. Vệt mây và bóng mây có thể quan sát được trong ảnh đã xử lý NDVI (ô vuông đen ở Hình 7 và Hình 8). Vùng che khuất bởi mây sẽ được mô hình phân loại thành vùng không có thảm phủ. Do đó, người sử dụng mô hình cần đối chiếu ảnh NDVI để loại bỏ những vùng này ra khỏi phân tích trượt lở.

Kết quả kiểm định mô hình bằng ma trận tương quan ở Bảng 2 cho thấy rằng phương pháp RF được sử dụng cho độ chính xác cao với độ chính xác toàn cục đạt 98.2% và chỉ số Kappa đạt 0.95. Do đó, bộ tham số của mô hình này được ứng dụng để phân loại vùng có thảm thực vật và vùng không có thảm thực vật cho toàn bộ ảnh trước và sau sự kiện trượt lở.



Hình 7. Ảnh đã xử lý NDVI (trước trượt lở)



Hình 8. Ảnh đã xử lý NDVI (sau trượt lở)

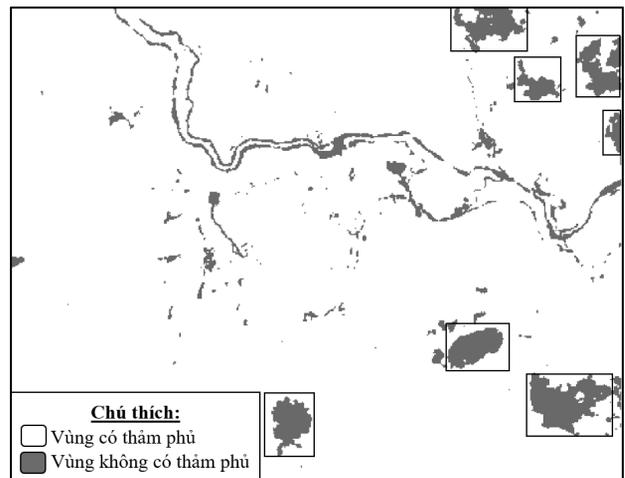
Bảng 2. Ma trận tương quan của mô hình RF

Giá trị thực tế \ Giá trị dự báo	Có thảm phủ	Không có thảm phủ
	Có thảm phủ	153
Không có thảm phủ	1	70

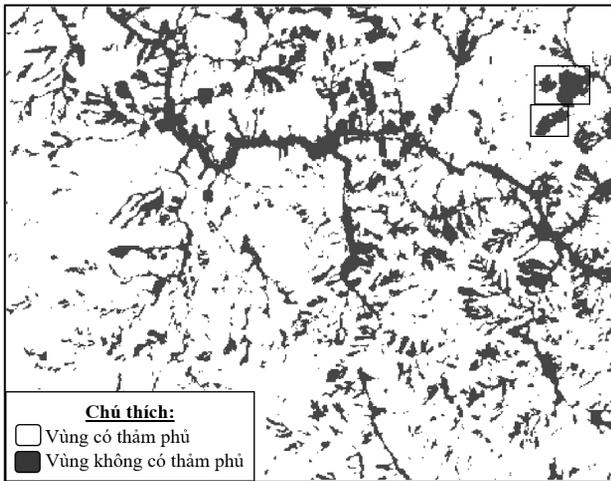
Ảnh sau khi được phân loại thể hiện bởi hai đặc trưng: (i) có thảm phủ và (ii) không có thảm phủ hoặc bề mặt nước hoặc bị ảnh hưởng bởi mây (Hình 9 và Hình 10). Vùng không có thảm phủ ở ảnh trước sự kiện biểu thị bằng màu xanh

(Hình 9) và ở ảnh sau sự kiện trượt lở biểu thị bằng màu đỏ (Hình 10). Những vùng bị che khuất bởi mây (Hình 7 và Hình 8) được mô hình dự báo là vùng không có thảm phủ sẽ không sử dụng để phân tích trượt lở (các ô vuông ở Hình 9 và Hình 10). Kết quả phân loại vùng không có thảm phủ (Hình 9 và Hình 10) khá phù hợp với những vùng không có cây xanh, đường bộ, khu dân cư và sông suối ở ảnh vệ tinh Sentinel 2 (Hình 3 và Hình 4).

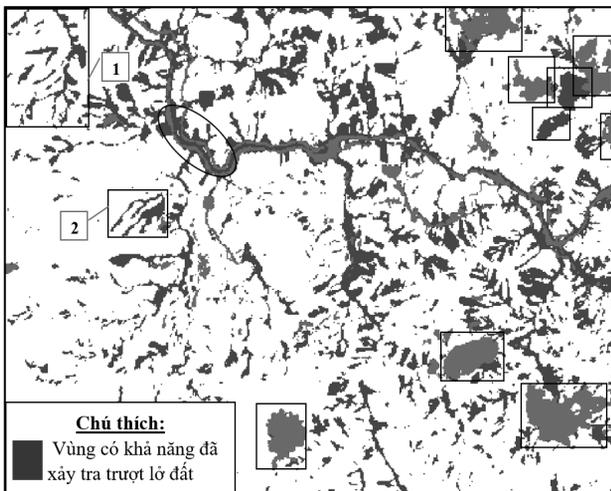
Kết quả phân loại vùng có thảm phủ và không có thảm phủ của ảnh trước và sau sự kiện được chồng chập lên nhau để xác định vùng bị trượt lở (Hình 11). Kết quả ở Hình 11 cho thấy có rất nhiều vùng không có thảm phủ (màu đỏ) xuất hiện sau sự kiện so với trước sự kiện trượt lở, đây là những vùng có khả năng cao đã xảy ra trượt lở đất, phản ánh sát thực trạng diễn ra ở khu vực nghiên cứu. Vị trí các điểm trượt lở tập trung chủ yếu trên tuyến đường bộ chạy dọc theo sông và tại vị trí các đường tụ thủy từ sườn dốc chảy ra lòng sông.



Hình 9. Ảnh trước trượt lở đã phân loại:
(i) có thảm phủ (màu trắng);
(ii) không có thảm phủ (màu xanh).
Vùng bị che khuất bởi mây (ô vuông đen)



Hình 10. Ảnh sau trượt lở đã phân loại: (i) có thảm phủ (màu trắng); (ii) không có thảm phủ (màu đen). Vùng mây (ô vuông đen)

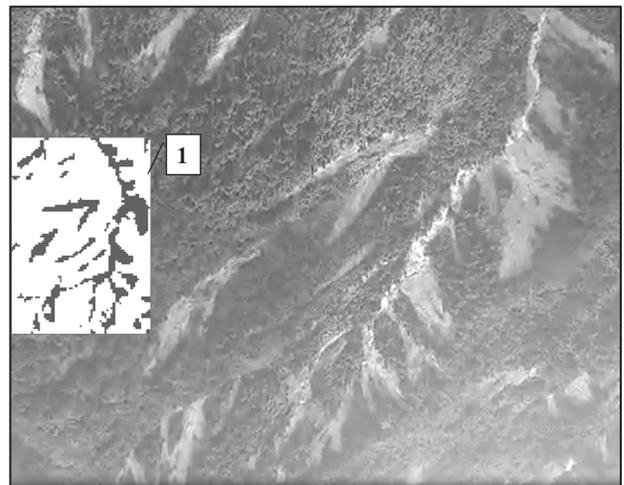


Hình 11. Chồng chụp ảnh trước trượt lở (Hình 9) và sau trượt lở (Hình 10)

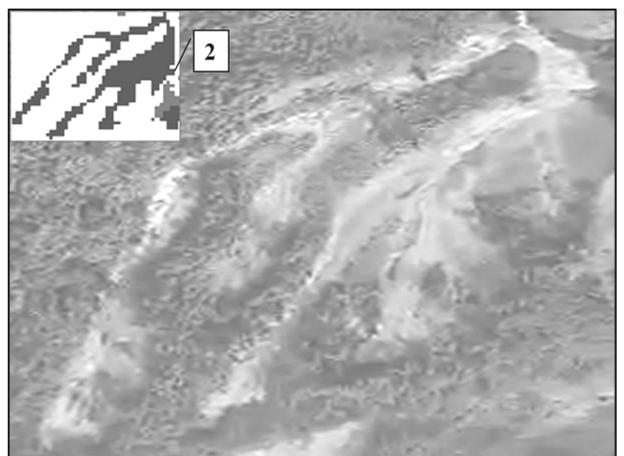
Đối với một số khu vực, kết quả chồng chụp 2 ảnh cần được phân tích thêm để loại trừ những điểm không chắc chắn. Vùng có thảm phủ thực vật trước sự kiện đã bị chuyển thành vùng không có thảm phủ thực vật sau sự kiện (vòng ellipse màu đen ở Hình 11) có thể không phải là vùng trượt lở. So sánh đối chiếu kết quả Hình 11 và ảnh gốc vệ tinh trước và sau sự kiện ở Hình 3 và Hình 4 cho thấy thảm phủ nằm giữa tuyến đường và lòng sông đã bị phá hoại và vùi lấp bởi đất đá sau trận lũ lụt. Ngoài ra, việc phân tích 2 lớp sẽ không

thấy được các biến động như ngập lụt (rừng -> mặt nước) hay biến động trượt lở (đất trồng -> vết trượt) vì NDVI của đất trồng và mặt nước trong phân loại 2 lớp là như nhau. Từ kết quả phân tích trên, nghiên cứu đề xuất cần kết hợp so sánh kết quả phân tích vùng trượt lở với bản đồ địa hình và ảnh gốc vệ tinh để đánh giá và nhận định đúng vết trượt lở.

Kết quả phân tích từ mô hình được so sánh với ảnh chụp thực tế vào ngày 04/11/2020 tại 2 vị trí: số 1 (có tọa độ $15^{\circ}17'04''$ độ vĩ Bắc, $107^{\circ}49'42''$ độ kinh Đông) và số 2 (có tọa độ $15^{\circ}16'23''$ độ vĩ Bắc, $107^{\circ}49'59''$ độ kinh Đông) (Hình 11). Qua so sánh về hình dạng, số lượng cũng như vị trí các vết trượt lở ở Hình 12 và Hình 13, có thể thấy rằng mô hình cho kết quả phù hợp thực tế.



Hình 12. Vị trí sạt lở thực tế số 1



Hình 13. Vị trí sạt lở thực tế số 2

5. KẾT LUẬN

Bài báo này áp dụng kỹ thuật phân tích ảnh vệ tinh Sentinel 2 và thuật toán phân loại Random Forest trong việc xác định vị trí trượt lở đất. Đầu tiên dữ liệu ảnh vệ tinh trước và sau sự kiện trượt lở được thu thập, xử lý và phân tích chỉ số NDVI. Sau đó, ảnh NDVI được phân loại thành vùng có thảm thực vật và không có thảm thực vật bằng thuật toán RF. Cuối cùng, vết trượt lở được xác định và đánh giá bằng cách chồng chập vùng không có thảm phủ trước và sau sự kiện trượt lở.

Nghiên cứu đã cho thấy ưu điểm của việc sử dụng ảnh vệ tinh Sentinel 2 trong việc xác định vị trí trượt lở đất tại những vùng rộng lớn có địa hình phức tạp và khó tiếp cận. Kết quả kiểm chứng tại 2 vị trí trượt lở bằng dữ liệu ảnh chụp thực tế đã cho thấy được khả năng áp dụng của phương pháp này. Bên cạnh đó, kết quả về phân lớp thảm thực vật và không có thảm phủ thực vật cũng chỉ ra rằng có

nhiều yếu tố không chắc chắn dẫn đến sai lệch trong kết quả phân loại như ảnh hưởng của mây, vùng đất trống do hoạt động của con người hay vùng thảm thực vật bị tàn phá bởi lũ lụt. Do đó, phương pháp này cần kết hợp với phân tích bản đồ địa hình để xác định tốt hơn vùng trượt lở. Việc sử dụng ảnh SAR với ưu điểm là không phụ thuộc vào các yếu tố thời tiết trong phân tích xác định vùng trượt lở có thể là một hướng nghiên cứu tiếp theo.

LỜI CẢM ƠN:

- **Đoàn Viết Long** được tài trợ bởi Tập đoàn Vingroup – Công ty CP và hỗ trợ bởi chương trình học bổng đào tạo thạc sĩ, tiến sĩ trong nước của Quỹ Đổi mới sáng tạo Vingroup (VINIF), Viện Nghiên cứu Dữ liệu lớn (VinBigdata), mã số **VINIF.2020.TS.135**.

- Bài báo này được tài trợ bởi Trường Đại học Bách khoa - Đại học Đà Nẵng với đề tài có mã số: T2020-02-19.

TÀI LIỆU THAM KHẢO

- Đoàn Viết Long, Nguyễn Chí Công, Nguyễn Quang Bình, N. T. C. (2020). *Đánh giá thực trạng và giải pháp nghiên cứu về trượt lở đất ở Việt Nam giai đoạn 2010 – 2020*. Tạp chí Khoa học và Công nghệ Thủy lợi, 61.
- Amatya, P., Kirschbaum, D., & Stanley, T. (2019). *Use of very high-resolution optical data for landslide mapping and susceptibility analysis along the Karnali highway, Nepal*. Remote Sensing, 11(19), 2284.
- Barra, A., Monserrat, O., Mazzanti, P., Esposito, C., Crosetto, M., & Scarascia Mugnozza, G. (2016). *First insights on the potential of Sentinel-1 for landslides detection*. Geomatics, Natural Hazards and Risk, 7(6), 1874–1883.
- Basten, K. (2016). *Classifying Landsat Terrain Images via Random Forests*. Bachelor Thesis Computer Science.
- Bernat Gazibara, S., Krkač, M., & Mihalić Arbanas, S. (2019). *Landslide inventory mapping using LiDAR data in the City of Zagreb (Croatia)*. Journal of Maps, 15(2), 773–779.
- Breiman, L. (2001). *Random forests*. Machine Learning, 45(1), 5–32.
- Budha, P. B., & Bhardwaj, A. (2019). *Landslide Extraction From Sentinel-2 Image In Siwalik Of Surkhet District, Nepal*. ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences, 4.
- Bunn, M. D., Leshchinsky, B. A., Olsen, M. J., & Booth, A. (2019). *A simplified, object-based framework for efficient landslide inventorying using LIDAR digital elevation model derivatives*. Remote Sensing, 11(3), 303.
- Cruden, D. M., & Varnes, D. J. (1996). *Landslides: investigation and mitigation*. Chapter 3- Landslide types and processes. Transportation Research Board Special Report, 247.

- Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., Meena, S. R., Tiede, D., & Aryal, J. (2019). *Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection*. *Remote Sensing*, 11(2), 196.
- Görüm, T. (2019). *Landslide recognition and mapping in a mixed forest environment from airborne LiDAR data*. *Engineering Geology*, 258, 105155.
- Guzzetti, F., Mondini, A. C., Cardinali, M., Fiorucci, F., Santangelo, M., & Chang, K.-T. (2012). *Landslide inventory maps: New tools for an old problem*. *Earth-Science Reviews*, 112(1–2).
- Le, T. T. T., & Kawagoe, S. (2017). *Landslide detection analysis in north vietnam base on satellite images and digital geographical information-landsat 8 satellite and historical data approaches*. *土木学会論文集 G (環境)*, 73(5), I_239-I_249.
- Le, T. T. T., & Kawagoe, S. (2018). *Study on landslide category base on temporal-spatial characteristic distribution in Northern Vietnam using satellite images*. *International Journal*, 14(43), 118–124.
- Li, Z., Shi, W., Lu, P., Yan, L., Wang, Q., & Miao, Z. (2016). *Landslide mapping from aerial photographs using change detection-based Markov random field*. *Remote Sensing of Environment*, 187, 76–90.
- Liu, X., Zhao, C., Zhang, Q., Peng, J., Zhu, W., & Lu, Z. (2018). *Multi-temporal loess landslide inventory mapping with C-, X-and L-band SAR datasets—A case study of Heifangtai Loess Landslides, China*. *Remote Sensing*, 10(11), 1756.
- Plank, S., Twele, A., & Martinis, S. (2016). *Landslide mapping in vegetated areas using change detection based on optical and polarimetric SAR data*. *Remote Sensing*, 8(4), 307.
- Pradhan, B., Al-Najjar, H. A. H., Sameen, M. I., Mezaal, M. R., & Alamri, A. M. (2020). *Landslide Detection Using a Saliency Feature Enhancement Technique From LiDAR-Derived DEM and Orthophotos*. *IEEE Access*, 8, 121942–121954.
- Prakash, N., Manconi, A., & Loew, S. (2020). *Mapping landslides on EO data: Performance of deep learning models vs. traditional machine learning models*. *Remote Sensing*, 12(3), 346.
- Qin, Y., Lu, P., & Li, Z. (2018). *Landslide inventory mapping from bitemporal 10 m sentinel-2 images using change detection based markov random field*. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 42(3).
- Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., & Guzzetti, F. (2018). *A review of statistically-based landslide susceptibility models*. *Earth-Science Reviews*, 180, 60–91.
- Scaioni, M., Longoni, L., Melillo, V., & Papini, M. (2014). *Remote sensing for landslide investigations: an overview of recent achievements and perspectives*. *Remote Sensing*, 6(10), 9600–9652.
- Strozzi, T., Klimeš, J., Frey, H., Caduff, R., Huggel, C., Wegmüller, U., & Rapre, A. C. (2018). *Satellite SAR interferometry for the improved assessment of the state of activity of landslides: A case study from the Cordilleras of Peru*. *Remote Sensing of Environment*, 217, 111–125.
- Tavakkoli Piralilou, S., Shahabi, H., Jarihani, B., Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., Meena, S. R., & Aryal, J. (2019). *Landslide detection using multi-scale image segmentation and different machine learning models in the higher himalayas*. *Remote Sensing*, 11(21), 2575.
- Van, N. T. H., Van Son, P., Ninh, N. H., Tam, N., & Huyen, N. T. (2017). *Landslide inventory mapping in the fourteen Northern provinces of Vietnam: achievements and difficulties*. *Workshop on World Landslide Forum*, 501–510.
- Zhao, W., Li, A., Nan, X., Zhang, Z., & Lei, G. (2017). *Postearthquake landslides mapping from Landsat-8 data for the 2015 Nepal earthquake using a pixel-based change detection method*. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(5), 1758–1768.

Abstract:

**STUDY ON THE APPLICATION OF SENTINEL-2 OPTICAL IMAGERY TO INVENTORY
LANDSLIDES USING RANDOM FOREST CLASSIFICATION MODEL**

In the study of landslides in Vietnam, the inventories of landslide has still been insufficient due to the difficulty in measuring and detecting location and time of landslide sites. With the development of the Earth - Observing Science and Computer Science, remote sensing technology is considered a solution to this problem. This study utilised optical imagery Sentinel 2 for landslide detection, analysed by SNAP and QGIS software. The pre-event and post-event Sentinel 2 images acquired at the same study area were selected for the analysis. Location of landslide points is determined based on the change of NDVI index, using Random Forest (RF) classification model and overlay mapping technique. The validation results showed that this model has performed well with the accuracy and kappa values are 98.2% and 0.95 respectively. In addition, the test results at 2 actual landslide locations have shown the applicability of this method.

Keywords: Landslide; Sentinel 2; SNAP software; Random Forest, NDVI.

Ngày nhận bài: 18/4/2021

Ngày chấp nhận đăng: 11/6/2021