



VIỆN KHOA HỌC THỦY LỢI VIỆT NAM

Tổng biên tập

GS.TS Nguyễn Tùng Phong

Phó Tổng biên tập

GS.TS Trần Đình Hòa

Ủy viên thường trực

PGS.TS Nguyễn Thanh Bằng

Hội đồng biên tập

TS Nguyễn Tiếp Tân

GS.TS Tăng Đức Thắng

GS.TS Lars Ribbe

GS.TS Lê Văn Nghị

PGS. TS Nguyễn Văn Thịnh

PGS.TS Đoàn Doãn Tuấn

PGS.TS Nguyễn Quốc Huy

PGS.TS Đặng Hoàng Thanh

PGS.TS Trần Bá Hoàng

PGS.TS Đỗ Hoài Nam

TS Trần Văn Đạt

Thư ký Tạp chí

CN Vũ Thị Tinh

Giấy phép xuất bản số

675/GP-BTTTT

ngày 19 tháng 10 năm 2021

In tại:

Công ty TNHH Một thành viên in Tạp chí Cộng sản

[TRONG SỐ NÀY]

KHOA HỌC CÔNG NGHỆ

- T2** Xây dựng bản đồ nguy cơ trượt lở đất theo chuỗi dữ liệu mưa thời gian thực khu vực Tả Van-Hầu Thào, Sa Pa-Lào Cai và đề xuất phương án cảnh báo sớm
Nguyễn Hà Phương, Hoàng Việt Hùng, Lê Thị Thanh Thủy, Nguyễn Tiếp Tân
- T9** Nghiên cứu thí điểm hiện đại hóa một số tiêu chí thủy lợi phục vụ chuyển đổi cơ cấu cây trồng hồ chứa Lái Bay, xã Phồng Lái, huyện Thuận Châu, tỉnh Sơn La
Nguyễn Xuân Lâm, Nguyễn Thị Kim Dung
- T18** Đánh giá mức độ và xu thế biến đổi lượng mưa lưu vực sông Cu Đê dựa trên dữ liệu mưa từ vệ tinh CHIRPS
Nguyễn Thành Luân, Lê Văn Nghị, Phạm Ngọc Tú, Âu Thị Bích Hoà, Hoàng Anh Quý
- T27** Nghiên cứu sử dụng phụ phẩm nông nghiệp và công nghiệp làm phụ gia khoáng để sản xuất bê tông tự lèn ứng dụng cho các công trình xây dựng
Ngô Cảnh Tùng, Nguyễn Quang Phú, Nguyễn Sỹ Tuấn, Nguyễn Minh Bảo, Trần Nhật Lâm
- T38** Phân loại sử dụng đất bằng mạng nơron tích chập trong nghiên cứu thủy văn cho huyện Mù Cang Chải
Lê Văn Thìn
- T45** Đa dạng sinh học cá và các đặc tính môi trường vùng biển và hải đảo tỉnh Cà Mau
Phạm Văn Tùng, Huỳnh Vũ Ngọc Quý
- T56** Nghiên cứu nguyên nhân và diễn biến xói lở mái hạ lưu đập đất hồ Sông Sắt, tỉnh Ninh Thuận
Trần Văn Vững
- T66** Đánh giá định lượng sự biến đổi lượng mưa trên lưu vực sông Hiếu, Nghệ An giai đoạn 1969-2024
Lê Thị Thu Hiền, Phạm Văn Chiến
- T74** Nghiên cứu đề xuất giải pháp phát triển hệ thống thủy lợi tỉnh Nam Định bằng phương pháp ma trận SWOT
Bùi Anh Tú
- T82** Xác định giá sản phẩm, dịch vụ thủy lợi phản ánh đầy đủ chi phí: Nghiên cứu điển hình tại hệ thống thủy lợi sông Nhuệ
Lê Văn Chính, Vương Việt Hưng, Nguyễn Minh Đức
- T90** Ứng dụng mô hình quan hệ mưa - dòng chảy - ngập lụt (RRI) trong xây dựng bản đồ ngập lụt cho lưu vực sông Đắc Bla
Trần Quốc Lập
- T98** Đánh giá chất lượng sản phẩm mưa lưới cho tỉnh Ninh Thuận và tỉnh Bình Thuận phục vụ công tác giám sát hạn
Nguyễn Hồ Phương Thảo, Hoàng Thanh Tùng
- T107** Đánh giá diễn biến chất lượng nước mặt theo xu thế phát triển kinh tế vùng, tỉnh Trà Vinh
Nguyễn Thị Phương Thảo, Lê Văn Tinh

PHÂN LOẠI SỬ DỤNG ĐẤT BẰNG MẠNG NƠ-ron TÍCH CHẬP TRONG NGHIÊN CỨU THỦY VĂN CHỖ HUYỆN MÙ CANG CHẢI

Lê Văn Thìn

Phòng Thí nghiệm trọng điểm Quốc gia về động lực học sông biển

Tóm tắt: Nghiên cứu này xây dựng mô hình học sâu sử dụng mạng nơ-ron tích chập (CNN) để phân loại sử dụng đất và che phủ đất (LULC) huyện Mù Cang Chải thành bảy lớp, nhằm hỗ trợ nghiên cứu dòng chảy thủy văn. Dữ liệu đầu vào bao gồm 16 tham số từ ảnh vệ tinh Sentinel-2 và dữ liệu radar Sentinel-1, được xử lý để phản ánh các đặc trưng thực vật, độ ẩm, và bề mặt không thấm nước. Các lớp LULC được định nghĩa dựa trên tài liệu của Cục Bảo tồn Tài nguyên Thiên nhiên Hoa Kỳ (NRCS), có liên quan chặt chẽ đến các điều kiện thủy văn như độ thấm thấu và dòng chảy bề mặt. Mô hình CNN được thiết kế với các lớp tích chập giãn nở và kernel lớn để nắm bắt mối quan hệ không gian, đạt độ chính xác xác thực 98.3% sau 50 epoch. So với các sản phẩm LULC hiện có như ESRI Land Cover và JAXA ALOS, phương pháp này cung cấp các lớp phân loại phù hợp hơn với nghiên cứu thủy văn. Kết quả có tiềm năng ứng dụng trong quản lý tài nguyên nước và dự báo lũ lụt.

Từ khóa: Mạng nơ-ron tích chập, CNN, LULC, Mù Căng Chải, thủy văn.

Summary: This study develops a deep learning model utilizing convolutional neural networks (CNNs) to classify land use and land cover (LULC) in Mu Cang Chai District into seven categories, aimed at supporting hydrological runoff studies. The input data consist of 16 parameters derived from Sentinel-2 optical imagery and Sentinel-1 radar data, processed to reflect vegetation characteristics, moisture levels, and impervious surfaces. The LULC classes are defined based on guidelines from the United States Natural Resources Conservation Service (NRCS), closely tied to hydrological conditions such as infiltration and surface runoff. The CNN model is designed with dilated convolutional layers and large kernels to capture spatial relationships, achieving a validation accuracy of 98.3% after 50 epochs. Compared to existing LULC products such as ESRI Land Cover and JAXA ALOS, this approach provides classifications more suited to hydrological research. The results hold potential for applications in water resource management and flood forecasting.

Keywords: Convolutional neural network, CNN, LULC, Mu Cang Chai, hydrology.

1. GIỚI THIỆU

Sử dụng đất và che phủ đất (LULC) đóng vai trò quan trọng trong nghiên cứu thủy văn, đặc biệt là trong việc ước lượng dòng chảy bề mặt và thấm thấu [1]. Các đặc trưng LULC ảnh hưởng trực tiếp đến các quá trình thủy văn như tốc độ dòng chảy, khả năng giữ nước, và nguy cơ lũ lụt. Theo tài liệu của Cục Bảo tồn Tài nguyên Thiên nhiên Hoa Kỳ (NRCS), các lớp

LULC được phân loại dựa trên điều kiện thủy văn, chẳng hạn như "Good", "Fair", hoặc "Poor", phản ánh khả năng thấm thấu của đất [2]. Tuy nhiên, các sản phẩm LULC hiện có như ESRI Land Cover [3] hoặc JAXA ALOS [4] thường tập trung vào các mục đích chung như giám sát môi trường hoặc đô thị hóa, thiếu sự tập trung vào các đặc trưng thủy văn cần thiết cho mô phỏng dòng chảy.

Nghiên cứu này xây dựng mô hình phân loại LULC mới sử dụng mạng nơ-ron tích chập (CNN), tận dụng dữ liệu viễn thám từ Sentinel-1 và Sentinel-2. Phương pháp này sử

Ngày nhận bài: 11/4/2025

Ngày thông qua phản biện: 04/5/2025

Ngày duyệt đăng: 02/6/2025

dụng 16 tham số đầu vào, bao gồm các chỉ số có liên quan đến thực vật và xây dựng (NDVI, EVI, GNDVI, GRVI, BSI, NDBI), chỉ số liên quan đến nước (NDWI1, NDWI2, NDWI, GSI), chỉ số màu tự nhiên (RED, GREEN, BLUE), độ dốc, và dữ liệu radar (VV, VH). Để phân loại LULC thành bảy lớp có liên quan trực tiếp đến nghiên cứu thủy văn. Các lớp này được thiết kế dựa trên tài liệu NRCS [2], đảm bảo khả năng áp dụng trong mô phỏng dòng chảy và quản lý tài nguyên nước. So với các sản phẩm LULC hiện có [3, 4], phương pháp này cung cấp độ phân giải không gian cao hơn và các lớp phân loại và phù hợp hơn với các điều kiện thủy văn.

Khu vực nghiên cứu phần lớn là huyện Mù Căng Chải trong nỗ lực xác định chỉ số CN từ phân loại LULC và nhóm đất thủy văn (HSG) nhằm đánh giá tốt hơn về dòng chảy, qua đó đánh giá khả năng hình thành lũ trong khu vực. Khu vực nghiên cứu được thể hiện trong Hình 1.



Hình 1: Khu vực nghiên cứu

2. PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Dữ liệu sử dụng

Dữ liệu đầu vào được thu thập từ hai nguồn chính: ảnh quang học Sentinel-2 [5] và ảnh

radar Sentinel-1 [6], cả hai đều có độ phân giải không gian 10m. Sentinel-2 cung cấp các băng quang học (Red, Green, Blue) và các chỉ số thực vật như NDVI (Normalized Difference Vegetation Index), EVI (Enhanced Vegetation Index), GNDVI (Green Normalized Difference Vegetation Index), và GRVI (Green-Red Vegetation Index). Các chỉ số nước như NDWI (Normalized Difference Water Index), GSI (Green-SWIR Index), NDWI1, và NDWI2 được tính toán để phát hiện các vùng nước và độ ẩm. Chỉ số xây dựng NDBI (Normalized Difference Built-up Index) và chỉ số đất trống BSI (Bare Soil Index) được sử dụng để xác định các khu vực đô thị và đất không có thực vật. Dữ liệu radar từ Sentinel-1 (VV và VH) cung cấp thông tin về cấu trúc bề mặt và độ ẩm, đặc biệt hữu ích trong điều kiện thời tiết xấu.

Dữ liệu độ dốc (Slope) được trích xuất từ mô hình độ cao số (DEM) của ALOS [7] ở độ phân giải 12,5m, phản ánh ảnh hưởng của địa hình đến dòng chảy. Tổng cộng, 16 tham số được xử lý thành các mảng 5x5 pixel, với mỗi pixel chứa 16 kênh tương ứng. Dữ liệu được thu thập từ huyện Mù Cang Chải, trong khoảng thời gian từ tháng 1 đến tháng 12 năm 2023, để đảm bảo bao quát các mùa khác nhau. Nhân đầu ra được gán thủ công dựa trên khảo sát thực địa và tham chiếu tài liệu NRCS [2], được trình bày chi tiết trong Bảng 1. Bên cạnh đó, để phân loại các nhân trong kết quả đầu ra, nghiên cứu sử dụng dữ liệu sử dụng đất kết hợp với dữ liệu rừng của huyện Mù Cang Chải [8]. Rừng rậm thuộc 2 nhóm “rừng giàu” và “rừng rất giàu” và rừng thưa thuộc 2 nhóm “rừng trung bình” và “rừng nghèo” theo cơ sở dữ liệu tài nguyên rừng [8].

Bảng 1: Các lớp LULC và mối liên hệ với thủy văn

| Lớp | Mô tả | Điều kiện thủy văn | Ảnh hưởng đến dòng chảy |
|-----|---|--------------------|--|
| 1 | Mặt nước (Hồ, sông, đầm lầy) | Bề mặt nước | Tạo dòng chảy trực tiếp, không thấm thấu |
| 2 | Rừng rậm (Rừng dày, điều kiện Good) | Good | Tăng thấm thấu, thấm thấu tốt |
| 3 | Đất trống (Fallow hoặc đất không có thực vật) | Poor | Tăng dòng chảy bề mặt, giảm thấm thấu |
| 4 | Khu xây dựng (Bề mặt không thấm nước) | Impervious | Tạo dòng chảy nhanh, không thấm thấu |
| 5 | Rừng thưa (Rừng thưa, điều kiện Fair/Poor) | Fair/Poor | Thấm thấu trung bình, dòng chảy trung bình |

| Lớp | Mô tả | Điều kiện thủy văn | Ảnh hưởng đến dòng chảy |
|-----|---|--------------------|---|
| 6 | Ruộng bậc thang (Đất canh tác với hệ thống terracing) | Contoured/terraced | Giảm dòng chảy nhờ lưu trữ nước trong ruộng |
| 7 | Đất trồng lúa (Đất canh tác ngập nước) | Wet meadow | Giữ nước lâu dài, dòng chảy chậm |

2.2. Mô hình mạng nơ-ron tích chập (CNN)

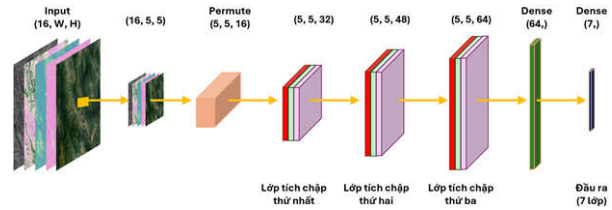
Mô hình CNN là mô hình học sâu được thiết kế để tận dụng mối quan hệ không gian giữa các pixel lân cận, một yếu tố quan trọng trong phân loại LULC do sự khác biệt về cấu trúc bề mặt giữa các lớp. Mô hình bao gồm ba lớp tích chập chính:

Lớp tích chập đầu tiên: Sử dụng kernel 5x5 để nắm bắt các đặc trưng không gian rộng hơn, chẳng hạn như ranh giới giữa rừng và đất trồng. Hàm kích hoạt LeakyReLU (negative_slope=0.1) được sử dụng để tránh vấn đề "dying ReLU", trong khi bình thường hóa hàng loạt (batch normalization) giúp ổn định quá trình huấn luyện.

Lớp tích chập giãn nở (thứ hai): Áp dụng dilation rate 2x2 để mở rộng trường tiếp nhận mà không tăng số lượng tham số, cho phép mô hình phát hiện các mẫu không gian lớn hơn như các vùng nước hoặc khu xây dựng. Lớp này sử dụng kernel 3x3 và cũng được theo sau bởi LeakyReLU và batch normalization.

Lớp tích chập thứ ba: Sử dụng kernel 3x3 để xử lý các đặc trưng kết hợp từ các lớp trước, tập trung vào các chi tiết nhỏ hơn như ranh giới giữa ruộng bậc thang và đất trồng lúa.

Sau các lớp tích chập, một lớp tổng hợp toàn cục (global average pooling) được sử dụng để giảm số chiều, tiếp theo là hai lớp dày đặc (dense) với 64 và 7 nơ-ron, tương ứng với số lớp đầu ra. Lớp dropout (0.2) được thêm vào để ngăn chặn quá khớp. Mô hình được tối ưu hóa bằng thuật toán Adam với tốc độ học ban đầu 0.00005 và clipnorm=1.0 để kiểm soát gradient lớn. Hàm mất mát sparse categorical crossentropy được sử dụng do nhãn là các giá trị nguyên. Cấu trúc mô hình được minh họa trong Hình 2.



Hình 2: Cấu trúc mạng CNN

Trong thiết kế mô hình CNN, số lượng đặc trưng (feature maps) được tăng dần qua các lớp tích chập, từ 32 ở lớp đầu tiên lên 48 ở lớp giãn nở và 64 ở lớp cuối cùng. Chiến lược này nhằm mục đích học các đặc trưng phức tạp hơn ở các tầng sâu, tận dụng mối quan hệ không gian giữa các pixel lân cận để phân biệt các lớp LULC có cấu trúc bề mặt đa dạng, chẳng hạn như ranh giới giữa rừng rậm và đất trồng hoặc các mẫu không gian lớn của khu xây dựng. So với các phương pháp thu hẹp dần (ví dụ: sử dụng các lớp max pooling hoặc stride lớn để giảm kích thước không gian), cách tiếp cận tăng dần đặc trưng giữ nguyên độ phân giải không gian qua các lớp tích chập, cho phép mô hình duy trì thông tin chi tiết về ranh giới và kết cấu bề mặt. Điều này đặc biệt quan trọng trong phân loại LULC, nơi các đặc trưng không gian như ranh giới giữa ruộng bậc thang và đất trồng lúa có thể ảnh hưởng đến kết quả thủy văn. Ngoài ra, việc sử dụng lớp tổng hợp toàn cục (global average pooling) thay vì các lớp thu hẹp dần ở giai đoạn sau giúp giảm số chiều mà không làm mất thông tin tổng quát, đồng thời giảm nguy cơ quá khớp so với các lớp dày đặc truyền thống. Tuy nhiên, nhược điểm của chiến lược này là yêu cầu tính toán cao hơn do kích thước không gian được giữ nguyên qua các lớp, có thể làm tăng thời gian huấn luyện và tiêu tốn tài nguyên trên các tập dữ liệu lớn. Hơn nữa, việc tăng số đặc trưng có thể dẫn đến nguy cơ học các đặc trưng dư thừa nếu không được kiểm soát chặt chẽ bởi các kỹ thuật như dropout và bình thường hóa hàng loạt.

Bảng 2: Mô tả mô hình CNN

| TT | Lớp | Loại | Kích thước | Số filter/ units | Đặc điểm đặc biệt | Kích thước đầu ra |
|----|---------------|-------------------|------------|------------------|---|-------------------|
| 0 | Input | Input | - | - | 16 bands, patch 5×5 | (16, 5, 5) |
| 1 | Permute | Reshape | - | - | Chuyển (16,5,5) → (5,5,16) | (5, 5, 16) |
| 2 | Conv2D | Tích chập | 5×5 | 32 | padding='same', nắm bắt lân cận | (5, 5, 32) |
| 3 | LeakyReLU | Kích hoạt | - | - | negative_slope=0.1 | (5, 5, 32) |
| 4 | BatchNorm | Chuẩn hóa | - | - | Chuẩn hóa theo batch | (5, 5, 32) |
| 5 | Conv2D | Tích chập giãn nở | 3×3 | 48 | dilation_rate=(2,2), mở rộng trường tiếp nhận | (5, 5, 48) |
| 6 | LeakyReLU | Kích hoạt | - | - | negative_slope=0.1 | (5, 5, 48) |
| 7 | BatchNorm | Chuẩn hóa | - | - | Chuẩn hóa theo batch | (5, 5, 48) |
| 8 | Conv2D | Tích chập | 3×3 | 64 | padding='same', xử lý đặc trưng kết hợp | (5, 5, 64) |
| 9 | LeakyReLU | Kích hoạt | - | - | negative_slope=0.1 | (5, 5, 64) |
| 10 | BatchNorm | Chuẩn hóa | - | - | Chuẩn hóa theo batch | (5, 5, 64) |
| 11 | GlobalAvgPool | Pooling | - | - | Trung bình toàn cục | (64) |
| 12 | Dense | Kết nối đầy đủ | - | 64 | Tầng ẩn | (64) |
| 13 | LeakyReLU | Kích hoạt | - | - | negative_slope=0.1 | (64) |
| 14 | BatchNorm | Chuẩn hóa | - | - | Chuẩn hóa theo batch | (64) |
| 15 | Dropout | Regularization | - | - | rate=0.2, giảm overfitting | (64) |
| 16 | Dense | Kết nối đầy đủ | - | 7 | activation='softmax', phân loại 7 lớp | (7) |

Tập dữ liệu bao gồm 100,000 mẫu (trên tổng số 122,393 mẫu dựa trên thuật toán không chệch số lượng mẫu tối đa) được chia thành tập huấn luyện (80%) và tập xác thực (20%) bằng phương pháp phân chia ngẫu nhiên có phân tầng để đảm bảo phân bố nhãn đồng đều. Toàn bộ mẫu được lấy ở dạng vùng và chuyển về dạng điểm thông qua công cụ rasterize trong QGIS. Mô hình được huấn luyện trong 50 epoch với kích thước batch là 32. Các callback được sử dụng bao gồm:

EarlyStopping: Dừng huấn luyện nếu hàm mất mát xác thực không cải thiện sau 5 epoch, đồng thời khôi phục trọng số tốt nhất.

ReduceLROnPlateau: Giảm tốc độ học xuống 50% nếu hàm mất mát xác thực không cải thiện sau 3 epoch, với tốc độ học tối thiểu là 1e-6.

Phân bố nhãn trong tập dữ liệu được trình bày trong Bảng 3, phản ánh sự đa dạng của các lớp

LULC trong khu vực nghiên cứu. Bài báo không trình bày bản đồ tập dữ liệu huấn luyện và kiểm tra của mô hình do các mẫu được lấy ở dạng vùng, sau đó chuyển về dạng điểm với độ phân giải là 12,5m (quá nhỏ và gần nhau so với kích thước khu vực nghiên cứu) gây ra hiệu ứng chồng lấn và thể hiện không rõ ràng trong các bản đồ.

Bảng 3: Phân bố nhãn trong tập dữ liệu huấn luyện

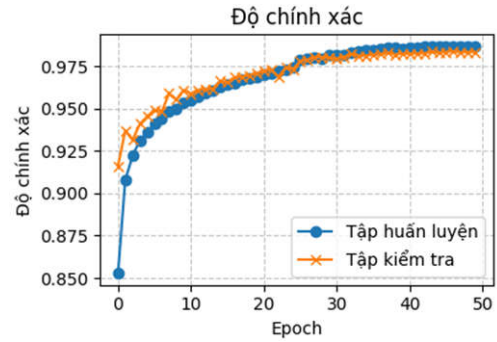
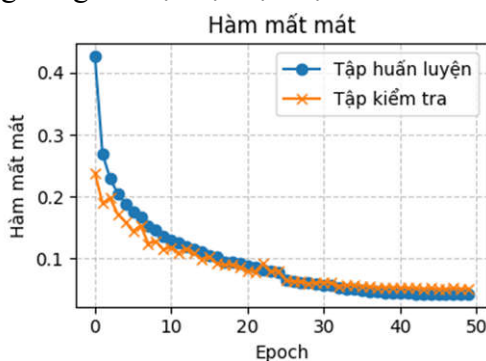
| Lớp | Tên | Số mẫu | Tỷ lệ |
|-----|-----------------|--------|--------|
| 1 | Mặt nước | 4,929 | 4.93% |
| 2 | Rừng rậm | 24,223 | 24.22% |
| 3 | Đất trống | 17,784 | 17.78% |
| 4 | Khu xây dựng | 25,505 | 25.51% |
| 5 | Rừng thưa | 4,986 | 4.99% |
| 6 | Ruộng bậc thang | 17,012 | 17.01% |
| 7 | Đất trồng lúa | 5,561 | 5.56% |

Tập dữ liệu huấn luyện thể hiện sự mất cân bằng đáng kể giữa các lớp, với các lớp như "Khu xây dựng" (25.51%) và "Rừng rậm" (24.22%) chiếm tỷ lệ lớn, trong khi các lớp như "Mặt nước" (4.93%) và "Rừng thưa" (4.99%) có số mẫu ít hơn (Bảng 3). Tuy nhiên, mô hình CNN vẫn đạt độ chính xác xác thực cao (98.32%) nhờ vào các chiến lược thiết kế hiệu quả. Thứ nhất, việc sử dụng phân chia ngẫu nhiên có phân tầng (stratified random splitting) đảm bảo rằng tỷ lệ các lớp được duy trì trong cả tập huấn luyện và xác thực, giúp mô hình học được các đặc trưng của các lớp thiểu số. Thứ hai, các lớp tích chập giãn nở và kernel lớn cho phép mô hình nắm bắt các mẫu không gian phức tạp, đặc biệt là ở các lớp hiếm như "Mặt nước", vốn có đặc trưng quang học rõ rệt (NDWI, NDWI1). Ngoài ra, hàm mất mát sparse categorical crossentropy và kỹ thuật dropout (0.2) giúp giảm thiểu tình trạng quá khớp đối với các lớp chiếm ưu thế, đảm bảo mô hình tổng quát hóa tốt trên tất cả các lớp. Những yếu tố này đã giảm thiểu tác động của dữ liệu mất cân bằng, dẫn đến hiệu suất phân loại đồng đều, như được thể hiện trong ma trận nhầm lẫn (Hình 4).

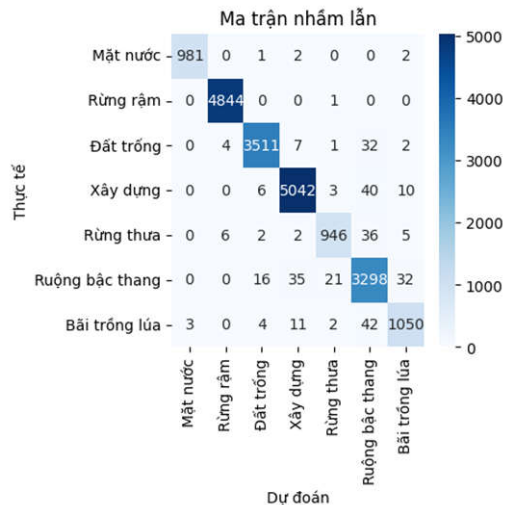
3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Hiệu suất mô hình

Mô hình CNN đạt độ chính xác huấn luyện là 98.68% và độ chính xác xác thực là 98.32% sau 50 epoch. Quá trình hội tụ được thể hiện trong Hình 3, cho thấy hàm mất mát giảm đều và ổn định, đặc biệt sau khi tốc độ học được điều chỉnh tự động. Ma trận nhầm lẫn (Hình 4) cho thấy mô hình phân loại chính xác hầu hết các lớp, với một số nhầm lẫn nhỏ giữa ruộng bậc thang (lớp 6) và bãi trồng lúa (lớp 7) do sự tương đồng về mật độ thực vật.



Hình 3: Đồ thị hàm mất mát và độ chính xác trong quá trình huấn luyện



Hình 4: Ma trận nhầm lẫn của mô hình CNN trên tập kiểm tra

Bảng 4: Độ chính xác và độ nhạy của từng lớp

| Lớp | Tên | Độ chính xác (%) | Độ nhạy (%) |
|-----|-----------------|------------------|-------------|
| 1 | Mặt nước | 99.7 | 99.49 |
| 2 | Rừng rậm | 99.79 | 99.98 |
| 3 | Đất trống | 99.18 | 98.71 |
| 4 | Khu xây dựng | 98.88 | 98.84 |
| 5 | Rừng thưa | 97.13 | 94.88 |
| 6 | Ruộng bậc thang | 95.65 | 96.94 |
| 7 | Đất trồng lúa | 95.37 | 94.42 |

3.2. So sánh với các sản phẩm LULC khác

Các sản phẩm LULC hiện có như ESRI Land Cover (độ phân giải 10m) và JAXA ALOS (độ phân giải 10m) cung cấp các lớp phân loại chung như rừng, đất nông nghiệp, và khu đô thị. Tuy nhiên, chúng thiếu các lớp chuyên biệt như ruộng bậc thang hoặc đất trồng lúa, vốn có

ảnh hưởng đáng kể đến dòng chảy thủy văn. Ví dụ, ESRI Land Cover phân loại đất nông nghiệp thành một lớp duy nhất, không phân biệt giữa đất canh tác ngập nước và đất canh tác khô, dẫn đến khó phân loại trong mô phỏng dòng chảy. Tương tự, JAXA ALOS không cung cấp thông tin về điều kiện thủy văn như "Good" hoặc "Poor", làm hạn chế khả năng áp dụng trong nghiên cứu thủy văn.

Phương pháp đề xuất trong nghiên cứu này khắc phục các hạn chế trên bằng cách sử dụng các lớp LULC được thiết kế dựa trên tài liệu NRCS, tập trung vào các đặc trưng thủy văn như khả năng thấm thấu và lưu trữ nước. Bảng 5 so sánh các lớp LULC của phương pháp này với ESRI Land Cover và JAXA ALOS.

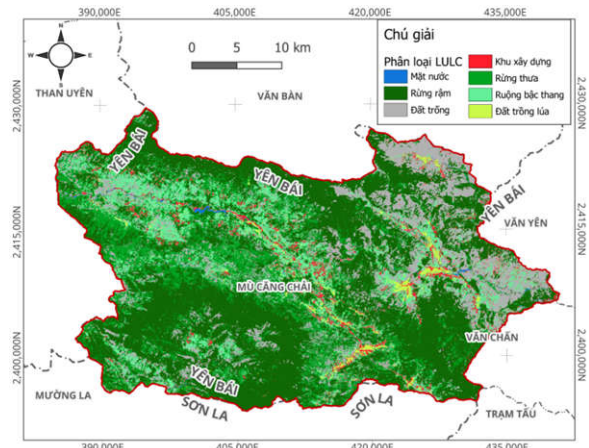
Bảng 5: So sánh các lớp LULC với các sản phẩm khác

| Lớp | Phương pháp này | ESRI Land Cover | JAXA ALOS |
|-----------------|-----------------|------------------|---------------------|
| Mặt nước | Có | Có | Có |
| Rừng rậm | Có (Good) | Có (Trees) | Có (chưa phân loại) |
| Đất trống | Có (Poor) | Có (Bare ground) | Có (Barren) |
| Khu xây dựng | Có (Impervious) | Có (Built-up) | Có (Urban) |
| Rừng thưa | Có (Fair/Poor) | Không | Có (chưa phân loại) |
| Ruộng bậc thang | Có (Contoured) | Không | Không |
| Đất trồng lúa | Có (Wet meadow) | Không | Có (rice) |

3.3. Ý nghĩa thủy văn

Các lớp phân loại sử dụng đất và che phủ đất (LULC) được đề xuất trong nghiên cứu này có mối liên hệ chặt chẽ với các quá trình thủy văn, đóng vai trò quan trọng trong việc mô phỏng dòng chảy bề mặt, thấm thấu, và quản lý tài nguyên nước tại huyện Mù Cang Chải. Ngoài các lớp phổ biến được phân loại ở nhiều nghiên cứu như mặt nước, đất trống, khu vực xây dựng, lớp rừng được phân loại chi tiết hơn theo thủy văn "Rừng rậm" với điều kiện thủy văn "Good" thể hiện khả năng thấm thấu cao

nhờ mật độ thực vật dày và tầng đất hữu cơ phong phú và lớp "Rừng thưa" với điều kiện thủy văn "Fair/Poor" thể hiện khả năng thấm thấu thấp hơn, phản ánh sự kết hợp giữa các khu vực có thực vật thưa thớt và đất bị suy thoái. Lớp này có vai trò chuyển tiếp giữa rừng rậm và đất trống, ảnh hưởng đến dòng chảy bề mặt ở mức độ vừa phải nhưng vẫn đóng góp vào việc giảm xói mòn so với đất trống. Đặc biệt, hai lớp "Ruộng bậc thang" và "Đất trồng lúa" mang ý nghĩa thủy văn quan trọng trong bối cảnh miền núi phía Bắc, đặc biệt là tại Mù Cang Chải, nơi nông nghiệp bậc thang là đặc trưng văn hóa và kinh tế. Lớp "Ruộng bậc thang" hoạt động như một hệ thống lưu trữ nước tự nhiên, giữ nước mưa trong các bậc thang và làm chậm dòng chảy, từ đó giảm thiểu nguy cơ lũ quét và xói mòn đất dốc. Hệ thống này cũng hỗ trợ việc tưới tiêu bền vững, duy trì độ ẩm đất trong mùa khô. Tương tự, lớp "Đất trồng lúa", với đặc điểm ngập nước thường xuyên, đóng vai trò như một vùng đệm thủy văn, lưu trữ nước và làm chậm dòng chảy, đồng thời giảm áp lực lên các hệ thống thoát nước tự nhiên. Những đặc điểm này không chỉ giúp kiểm soát lũ mà còn cải thiện khả năng thích ứng với biến đổi khí hậu, đặc biệt trong các khu vực dễ bị ảnh hưởng bởi mưa lớn và hạn hán xen kẽ. Bản đồ LULC được tạo ra từ mô hình CNN, như được minh họa trong Hình 5, cung cấp một cái nhìn chi tiết về sự phân bố không gian của các lớp này tại Mù Cang Chải.



Hình 5: Bản đồ LULC được tạo ra từ mô hình CNN

4. KẾT LUẬN

Nghiên cứu này đã xây dựng mô hình phân loại LULC mới sử dụng mạng nơ-ron tích chập, đạt độ chính xác cao và phù hợp với nghiên cứu thủy văn. Bằng cách sử dụng 16 tham số đầu vào từ Sentinel-1, Sentinel-2 và ALOS, mô hình CNN đã phân loại thành công bảy lớp LULC, mỗi lớp có mối liên hệ rõ ràng với các quá trình thủy văn như thấm thấu và dòng chảy. So với các sản phẩm LULC hiện có như ESRI Land Cover và JAXA ALOS, phương pháp này cung cấp các lớp phân loại chuyên biệt hơn, đặc biệt là ruộng bậc thang, đất trồng lúa và chất lượng rừng (thông qua CSDL rừng), vốn quan

trọng trong mô phỏng dòng chảy ở các khu vực nông nghiệp miền núi phía Bắc Việt Nam. Kết quả có tiềm năng ứng dụng trong quản lý tài nguyên nước, dự báo lũ lụt, và quy hoạch nông nghiệp bền vững.

LỜI CẢM ƠN: Bài báo này là một phần của đề tài: “Nghiên cứu ứng dụng trí tuệ nhân tạo và dữ liệu địa không gian để phân vùng lũ quét quy mô cấp huyện” thuộc chương trình khoa học công nghệ tiềm năng cấp bộ. Chúng tôi xin chân thành cảm ơn Bộ Nông nghiệp và Phát triển nông thôn (nay là Bộ Nông nghiệp và Môi trường) đã cấp kinh phí tài trợ cho nghiên cứu này.

TÀI LIỆU THAM KHẢO

- [1] S. A. Alawi and S. Özkul, "Evaluation of land use/land cover datasets in hydrological modelling using the SWAT model," *H2Open Journal*, vol. 6, no. 1, pp. 63-74, mar/2023. DOI:10.2166/h2oj.2023.062.
- [2] NRCS, "Part 630 - Hydrology," 3/2020. [Trực tuyến]. Địa chỉ: <https://directives.nrcs.usda.gov/sites/default/files2/1712930634/Part%20630%20-%20Hydrology.pdf>. [Truy cập 11/12/2023].
- [3] ESRI, "Esri | Sentinel-2 Land Cover Explorer," Esri, [Trực tuyến]. Địa chỉ: <https://livingatlas.arcgis.com/landcoverexplorer/>.
- [4] JAXA, "JAXA High-Resolution Land-Use and Land-Cover Map," JAXA, [Trực tuyến]. Địa chỉ: https://www.eorc.jaxa.jp/ALOS/en/dataset/lulc/lulc_vnm_v2309_e.htm. [Truy cập 15/4/2024].
- [5] GEE, "Harmonized Sentinel-2 MSI: MultiSpectral Instrument, Level-2A (SR)," Google Earth Engine, [Trực tuyến]. Địa chỉ: https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR_HARMONIZED.
- [6] GEE, "Sentinel-1 SAR GRD: C-band Synthetic Aperture Radar Ground Range Detected, log scaling," Google Earth Engine, [Trực tuyến]. Địa chỉ: https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S1_GRD.
- [7] JAXA, "ALOS PALSAR," Japan Aerospace Exploration Agency, [Trực tuyến]. Địa chỉ: <https://www.earthdata.nasa.gov/data/catalog/alaska-satellite-facility-distributed-active-archive-center-alos-psr-version-1>.
- [8] Cục kiểm lâm, "Hệ thống thông tin tài nguyên rừng," Cục kiểm lâm, [Trực tuyến]. Địa chỉ: <http://maps.vnforest.gov.vn/vn>.