

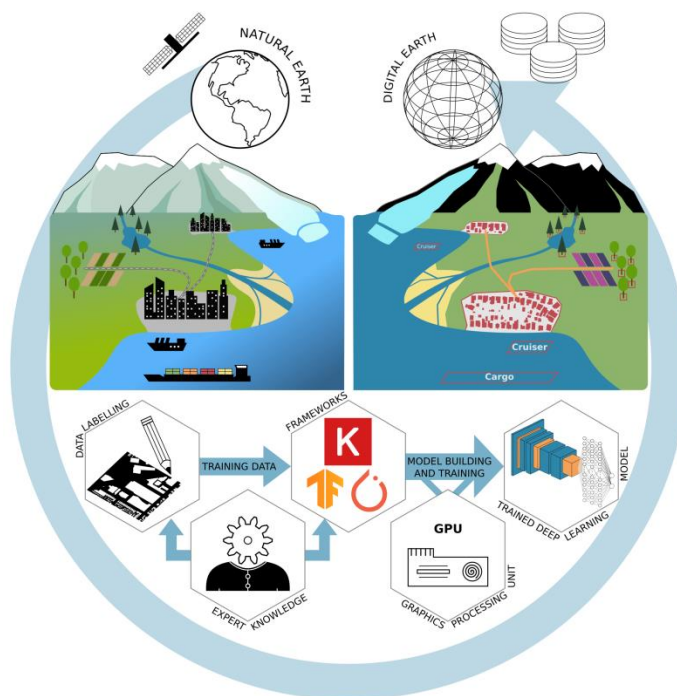
ĐẠI HỌC MỎ - ĐỊA CHẤT
KHOA CÔNG NGHỆ THÔNG TIN

BÁO CÁO HỌC THUẬT

**VIỄN THÁM VÀ DỮ LIỆU LỚN, NGHIÊN CỨU VÀ VẬN DỤNG DỮ
LIỆU LỚN TỪ ẢNH VIỄN THÁM**

Người thực hiện báo cáo

ThS. Trương Xuân Bình



Hà Nội – 2021

MỤC LỤC

DANH MỤC HÌNH VẼ	3
DANH MỤC CÁC TỪ VIẾT TẮT	4
MỞ ĐẦU	5
CHƯƠNG 1: 7DỮ LIỆU LỚN VIỄN THÁM	7
1.1 Xử lý dữ liệu lớn viễn thám.....	9
1.1.1 Xử lý hình học (xử lý không gian)	9
1.1.2. Xử lý bức xạ (quang phổ)	10
1.1.3. Mặt nạ đám mây	11
1.1.4. Hợp nhất dữ liệu.....	12
1.2. Phân tích dữ liệu lớn viễn thám.....	16
1.2.1. Phát hiện đối tượng và trích xuất đối tượng	16
1.2.2. Phân loại hình ảnh sử dụng đất và lớp phủ	18
1.2.3. Phát hiện các yếu tố thay đổi và biến động.....	19
1.2.4. Phân tích viễn thám đa thời gian	20
1.2.5. Các nhiệm vụ phân tích viễn thám khác	21
CHƯƠNG 2:CÁC MÔ HÌNH HỌC SÂU PHỔ BIẾN	23
2.1. Mạng nơ-ron tích chập (Convolutional neural networks - CNN)	24
2.2. Mạng nơron hồi quy (Recurrent Neural Networks RNN).....	27
2.3. Mạng Nơ ron tự lập mã (Autoencoder neural networks – AE).....	28
2.4. Mạng Nơ ron đối luyện (Generative adversarial networks – GAN)	29
2.5. Mạng nơ-ron đồ thị (Graph neural networks – GNN).....	30
CHƯƠNG 3: TÌM HIỂU CÁC MÔ HÌNH HỌC MÁY TRONG NHẬN DIỆN TRƯỢT LỖ TỪ DỮ LIỆU ẢNH VỆ TINH	32
3.1 Mô hình mạng nơ ron đa lớp (MLP -NN)	32
3.2. Mô hình Hồi quy logistic (LR).....	33
3.3 Mô hình Rừng Ngẫu nhiên (RF)	34
3.4 Phương pháp mô hình xếp chồng	34
3.5 Mô hình lai tích hợp giữa mạng nơ ron đa lớp và OBIA	35
KẾT LUẬN	36

DANH MỤC HÌNH VẼ

Hình 2.1: Mô hình mạng Nơ ron nhân tạo	23
Hình 2.2: Mô hình kiến trúc của mạng nơ ron tích chập - CNN.....	25
Hình 2.3: Mô phỏng lớp tích chập sử dụng cửa sổ 3x3.....	26
Hình 2.4: Kiến trúc mô hình lớp kết nối đầy đủ.....	26
Hình 2.5 Mô hình kiến trúc mạng nơ ron hồi quy - RNN	27
Hình 2.6: Sơ đồ kiến trúc mạng Nơ ron tự lập mã - AE	29
Hình 2.7: Mô hình kiến trúc mạng nơ ron đối luyện - GAN.....	30
Hình 2.8: Mô hình kiến trúc mạng nơ ron đồ thị (GNN)	31

DANH MỤC CÁC TỪ VIẾT TẮT

NN	Neural Networks	Mạng nơ ron
CNN	Convolutional neural networks	Mạng nơ ron tích chập
RNN	Recurrent neural networks	Mạng nơ ron hồi quy
LSTM	long short-term memory	Chuỗi bộ nhớ phụ thuộc ngắn
GRU	gated recurrent unit	Đơn vị nút hồi quy
AE	Autoencoder neural networks	Mạng nơ ron tự lập mã
GAN	Generative adversarial networks	Mạng nơ ron đối luyện
GNN	Graph neural networks	Mạng nơ ron đồ thị
EO	Earth observation	Vệ tinh quan sát trái đất
OBIA	object-based image analysis	Phân tích ảnh dựa trên đối tượng
GIS	Geographic Information System	Hệ thông tin địa lý
ML	Machine Learning	Học máy
RS	Remote Sensing	Viễn thám
ANN	Artificial Neural Network	Mạng nơ ron nhân tạo
MLP-NN	Multilayer perceptron	Mạng nơ ron đa lớp
PAN	Panchromatic	Hình ảnh toàn sắc
TCN	Temporal CNN	

MỞ ĐẦU

Sự phát triển nhanh chóng của công nghệ cảm biến đã thúc đẩy sự phát triển của viễn thám. Nhiều cảm biến khác nhau được lắp đặt trên vệ tinh, tàu con thoi, máy bay không người lái (UAV) và các trạm quan sát mặt đất đã tạo ra các hệ thống quan sát Trái đất theo nhiều góc độ không gian – không - đất.

. Thứ nhất, do nhu cầu to lớn đối với các ứng dụng tài nguyên và môi trường chính xác và khu vực rộng lớn, nhiều vệ tinh và chòm sao đã được phóng lên, chẳng hạn như loạt Landsat của Hoa Kỳ, dự án Copernicus của Liên minh Châu Âu và dự án lớn của Hệ thống quan sát Trái đất độ phân giải cao quốc gia của Trung Quốc (Gaofen). Hơn nữa, vốn tư nhân và những người đầu tư thương mại đang tiến tới các ứng dụng viễn thám. Ví dụ, Planet đã phóng khoảng 200 vệ tinh nhỏ, 1 vệ tinh tạo ra một lượng lớn dữ liệu viễn thám mỗi ngày. Chỉ riêng các vệ tinh Sentinel-2 A và B quang học đã tạo ra trung bình khoảng 9,54 TB dữ liệu mỗi ngày theo kế hoạch thu nhận.

Thứ hai, do những ưu điểm độc đáo của nó, ví dụ: tính linh hoạt, độ phân giải không gian cao và thu thập dữ liệu theo yêu cầu, viễn thám trên không đã được công nhận là sự bổ sung hiệu quả cho các nền tảng viễn thám dựa trên không gian truyền thống. Sự tiến bộ gần đây của công nghệ UAV với các cảm biến kích thước nhỏ và độ chính xác phát hiện cao khiến viễn thám dựa trên UAV trở thành một công nghệ rất phổ biến và ngày càng được sử dụng nhiều hơn, nó cũng tạo ra hình ảnh có độ phân giải cao lớn.

Thứ ba, do sự phát triển của công nghệ mới trong điện thoại thông minh và mạng không dây, ngày càng có nhiều trạm quan sát mặt đất được thành lập. Chúng tạo ra các quan sát tại chỗ với tần số cao, giúp làm phong phú thêm nguồn dữ liệu viễn thám. Nói chung, các hệ thống quan sát không gian-trên không-mặt đất cung cấp các RSD không lờ, đa nguồn, đa phương thức, đa tỷ lệ, chiều cao, trạng thái động và không đồng nhất.

Trong báo cáo nghiên cứu này ngoài nội dung của phần Mở đầu và Kết luận, báo cáo gồm các chương sau:

CHƯƠNG 1: DỮ LIỆU LỚN VIỄN THÁM

CHƯƠNG 2: CÁC MÔ HÌNH HỌC SÂU PHỔ BIẾN

CHƯƠNG 3: TÌM HIỂU CÁC MÔ HÌNH HỌC MÁY TRONG NHẬN DIỆN
TRƯỢT LỎ TỪ DỮ LIỆU VIỄN THÁM

CHƯƠNG 1:

DỮ LIỆU LỚN VIỄN THÁM

Sự phát triển nhanh chóng của công nghệ cảm biến đã thúc đẩy sự phát triển của viễn thám. Nhiều cảm biến khác nhau được lắp đặt trên vệ tinh, tàu con thoi, máy bay không người lái (UAV) và các trạm quan sát mặt đất đã tạo ra các hệ thống quan sát Trái đất theo nhiều góc độ không gian-không-đất.

Thứ nhất, do nhu cầu to lớn đối với các ứng dụng tài nguyên và môi trường chính xác và khu vực rộng lớn, nhiều vệ tinh và chòm sao đã được phóng lên, chẳng hạn như loạt Landsat của Hoa Kỳ, dự án Copernicus của Liên minh Châu Âu và dự án lớn của Hệ thống quan sát Trái đất độ phân giải cao quốc gia của Trung Quốc (Gaofen). Hơn nữa, vốn tư nhân và những người đầu tư thương mại đang tiến tới các ứng dụng viễn thám. Ví dụ, Planet đã phóng khoảng 200 vệ tinh nhỏ, 1 vệ tinh tạo ra một lượng lớn dữ liệu viễn thám mỗi ngày. Chỉ riêng các vệ tinh Sentinel-2 A và B quang học đã tạo ra trung bình khoảng 9,54 TB dữ liệu mỗi ngày theo kế hoạch thu nhận.

Thứ hai, do những ưu điểm độc đáo của nó, ví dụ: tính linh hoạt, độ phân giải không gian cao và thu thập dữ liệu theo yêu cầu, viễn thám trên không đã được công nhận là sự bổ sung hiệu quả cho các nền tảng viễn thám dựa trên không gian truyền thống. Sự tiến bộ gần đây của công nghệ UAV với các cảm biến kích thước nhỏ và độ chính xác phát hiện cao khiến viễn thám dựa trên UAV trở thành một công nghệ rất phổ biến và ngày càng được sử dụng nhiều hơn, nó cũng tạo ra hình ảnh có độ phân giải cao lớn.

Thứ ba, do sự phát triển của công nghệ mới trong điện thoại thông minh và mạng không dây, ngày càng có nhiều trạm quan sát mặt đất được thành lập. Chúng tạo ra các quan sát tại chỗ với tần số cao, giúp làm phong phú thêm nguồn dữ liệu viễn thám. Nói chung, các hệ thống quan sát không gian-trên không-mặt đất cung cấp các RSBD khổng lồ, đa nguồn, đa phương thức, đa tỷ lệ, chiều cao, trạng thái động và không đồng nhất.

Từ dữ liệu lớn, RSBD kế thừa bốn đặc điểm “V” (sau đây gọi là 4V) bao gồm : khối lượng, đa dạng, vận tốc và giá trị. Ngoài ra, RSBD có thể được mô tả bằng các kích thước của nó.

- Khối lượng: Ngày càng nhiều cảm biến trên vệ tinh, tàu con thoi và mặt đất liên tục quan sát bề mặt Trái đất, tích lũy dữ liệu VIỄN THÁM khổng lồ.

- Đa dạng: RSBD bao gồm dữ liệu từ nhiều nguồn (vệ tinh, UAV, mặt đất, v.v.) và đa phương thức (LiDAR, radar, quang học, v.v.), đa độ phân giải (từ 1 cm đến 100 km) và đa mục (được thu thập vào các ngày không thường xuyên và thời gian) dữ liệu, cũng như dữ liệu từ các lĩnh vực khác nhau tùy thuộc vào lĩnh vực ứng dụng.

- Tốc độ: Bề mặt Trái đất, mục tiêu quan sát Trái đất đang thay đổi và dữ liệu VIỄN THÁM được tạo với tốc độ phát triển nhanh chóng. Hơn nữa, ngày càng có nhiều nhiệm vụ yêu cầu xử lý và phân tích (gần) theo thời gian thực để tạo ra thông tin mới nhất cho việc hỗ trợ quyết định, ví dụ: hàng giờ có thể cứu hàng trăm ha rừng trong vụ cháy rừng.

Giá trị: “Giá trị” là chất lượng vốn có của dữ liệu lớn. Thông tin thu được từ RSBD có thể dự đoán những thay đổi của khí hậu toàn cầu, hướng dẫn trồng trọt nông nghiệp, hỗ trợ quy hoạch đô thị, v.v.

Ngoài các tính năng 4V phổ biến của dữ liệu lớn nói chung, RSBD trình bày các đặc điểm sau:

Dữ liệu viễn thám là đa phân giải và đa cấp độ. Do các khả năng cảm biến và khoảng cách cảm biến khác nhau, độ phân giải của dữ liệu VIỄN THÁM nằm trong khoảng từ hàng trăm km đến vài cm, mang lại hiệu quả quy mô đáng kể cho quá trình xử lý và phân tích.

Dữ liệu viễn thám thường đa phương thức, ví dụ: từ cảm biến quang học (đa và siêu kính), LiDAR và radar khẩu độ tổng hợp (SAR), trong đó hình dạng và nội dung hình ảnh hoàn toàn khác nhau.

Yếu tố trì hoãn thời gian ngày càng trở nên quan trọng. Ngoài việc theo đuổi độ phân giải không gian và quang phổ cao trong thế kỷ trước, nhu cầu về độ phân giải thời gian cao gần đây đã tăng lên. Ví dụ: các vệ tinh (ví dụ: vệ tinh A và B của Sentinel-1) hoặc vệ tinh ảo (ví dụ: Landsat 8 và Sentinel-2) bao gồm một số vệ tinh tương tự có thể đạt được chu kỳ quay lại ngắn hơn. Mặt khác, ngày càng có nhiều

nhiệm vụ đặt ra yêu cầu cao hơn về tính kịp thời của hình ảnh, chẳng hạn như giám sát mùa màng và đánh giá thảm họa thiên tai. Từ việc xác định chất lượng (chẳng hạn như phân loại hình ảnh và trích xuất đối tượng), các ứng dụng viễn thám tiến tới định lượng (chẳng hạn như đảo ngược tham số và ước tính năng suất cây trồng). Những đặc điểm này của RSBD làm cho việc xử lý và phân tích RDBD chính xác, nhanh chóng và tự động là thách thức đối với các ứng dụng viễn thám.

Dựa trên tổng quan kỹ thuật của các thuật toán học sâu chương 1 và nhu cầu thực tế trong các ứng dụng của dữ liệu lớn viễn thám, chúng tôi nhận thấy rằng học sâu có thể được áp dụng cho mọi khía cạnh của xử lý và phân tích dữ liệu viễn thám: từ các chủ đề truyền thống về phân tích bản đồ ảnh và phát hiện các đối tượng cho các nhiệm vụ thách thức gần đây của phân đoạn dựa trên ngữ nghĩa cấp cao và phân tích đa mục tiêu. Trong phần này, chúng tôi theo dõi quá trình xử lý và phân tích viễn thám để xem xét tiến sự tiến bộ của viễn thám dựa trên học sâu.

1.1 Xử lý dữ liệu lớn viễn thám

1.1.1 Xử lý hình học (xử lý không gian)

Xử lý hình học trong viễn thám nhằm xác định vị trí mọi pixel của đối tượng trong không gian địa lý, bao gồm chỉnh sửa hình học, chỉnh hình, chỉnh sửa địa hình và tái tạo mô hình 3D, v.v. Bước sơ bộ cơ bản của xử lý hình học là bản chỉnh hình ảnh, căn chỉnh nhiều hình ảnh được chụp bởi các cảm biến khác nhau (căn chỉnh đa phương thức) tại các thời điểm khác nhau (Căn chỉnh đa vùng) từ các góc nhìn khác nhau (căn chỉnh đa giác) vào cùng một hệ tọa độ. Quy trình căn chỉnh hình ảnh điển hình bao gồm ba bước sau: trích xuất đối tượng địa lý (bộ mô tả), đối sánh đối tượng địa lý và biến đổi hình ảnh.

Trích xuất đối tượng đóng một vai trò quan trọng trong việc căn chỉnh hình ảnh vì nó quyết định loại đối tượng sẽ sử dụng để đối sánh. Trong các phương pháp truyền thống, nhiều bộ mô tả tính năng được làm thủ công đạt được kết quả tốt trong hình ảnh vệ tinh, chẳng hạn như SIFT, SURF và ORB. Do biểu diễn tính năng đáng kinh ngạc, các thuật toán học sâu (đặc biệt là mạng nơ ron tích chập) có thể tạo ra nhiều tính năng ứng viên đa cấp độ (bộ mô tả). Do đó, các mạng nơ ron tích chập khác nhau đã được phát triển để trích xuất đối tượng địa lý và đã cải thiện hiệu suất căn chỉnh các đối

tượng ảnh. Ví dụ: Yang, Dan và Yang (2018c) đã sử dụng mạng nơ ron tích chập để tạo các bộ mô tả tính năng đa cấp mạnh mẽ (giữ cả thông tin tích hợp và khả năng định vị) để xử lý các hình ảnh viễn thám đa mục tiêu.

Để tận dụng hơn nữa khả năng học tập từ đầu đến cuối của học sâu, các nhà nghiên cứu đã cố gắng thiết kế các kiến trúc hoàn chỉnh thống nhất việc trích xuất và kết hợp tính năng trong một khuôn khổ vòng lặp kín để học trực tiếp từ sự chuyển đổi hình học giữa hai hình ảnh. Người ta đã thiết kế một mạng ánh xạ trực tiếp, trong đó các lớp ẩn tương ứng với việc trích xuất đối tượng địa lý và lớp đầu ra tương ứng với đối sánh đối tượng địa lý. Kiến trúc end-to-end này tối ưu hóa chức năng ánh xạ học thông qua phản hồi thông tin khi đào tạo, cho phép kết quả đối sánh các tính năng để hướng dẫn quá trình khai thác các hàm chức năng, sau đó làm cho các tính năng đã học trở nên phù hợp hơn.

Ngoài việc căn chỉnh giữa các hình ảnh đồng nhất, các thuật toán học sâu có thể được áp dụng để khai thác dữ liệu viễn thám đa phương thức, ví dụ: hình ảnh quang học và dữ liệu SAR. Chiến lược phổ biến của các phương pháp này là dịch các hình ảnh đầu vào vào cùng một không gian đặc trưng, cho phép hai hình ảnh chia sẻ cường độ hoặc thông tin tính năng tương tự. Sau đó, việc trích xuất và đối sánh tính năng của đối tượng được thực hiện giữa hai hình ảnh này.

Gần đây, các kỹ thuật dựa trên học máy mới nhất đã được đưa vào áp dụng trong việc nắn chỉnh các hình ảnh viễn thám, chẳng hạn như học tăng cường và học đa nhiệm. Một thuật toán căn chỉnh vector-to-raster tự động dựa trên học máy tăng cường đã được giới thiệu để ghim vị trí chính xác của các đối tượng địa lý trên bản đồ được quét. Thuật toán này cũng có thể được áp dụng cho các đối tượng địa lý khác nhau (đường bộ, đường thủy và đường sắt). Các nhà khoa học đã sử dụng tính năng học đa nhiệm để cải thiện hiệu suất của quá trình nắn chỉnh và xử lý hình ảnh (Girard, Charpiat và Tarabalka, 2018).

1.1.2. Xử lý bức xạ (quang phổ)

Xử lý hình ảnh bức xạ hoặc hiệu chỉnh giá trị pixel từ không gian đo này sang không gian đo khác, có hoặc không có dữ liệu tham chiếu, bao gồm quá trình đảo ngược tham số vật lý, hiệu chỉnh bức xạ tuyệt đối và tương đối và chuẩn hóa các ước số bức xạ. Do sự phức tạp của các quá trình vật lý, khi các tham số của mô hình

truyền các sóng bức xạ không tuyệt đối và đồng nhất, các mô hình dựa trên thống kê học (ví dụ: mạng nơ-ron nông, máy vectơ hỗ trợ) đã được phát triển thành các phương pháp áp dụng cho các bài toán thực tế. Đặc biệt, các kỹ thuật học sâu đã thúc đẩy hơn nữa quá trình xử lý bức xạ dựa trên quá trình học tập liên tục. Ogut, Bosch-Lluis và Reising (2019) đã áp dụng kỹ thuật hiệu chuẩn dựa trên học tập sâu để hiệu chuẩn các thiết bị đo bức xạ sóng milimét và vi sóng trong không khí ở tần số cao. Bên cạnh đó cũng đã có một giải pháp dựa trên trí tuệ nhân tạo / học sâu để mô tả đặc điểm của bầu khí quyển tại các điểm thuận lợi khác nhau và để lấy các đặc tính quang phổ của các mục tiêu muốn trích xuất. Một cách tiếp cận giả lập học sâu để hiệu chỉnh khí quyển đã được phát triển, cho thấy rằng một mô hình học sâu có thể được đào tạo để mô phỏng một quá trình vật lý phức tạp của các đối tượng địa lý.

Hiện tại xử lý các bức xạ dựa trên deep-learning vẫn nằm trong lĩnh vực truyền thống của máy học và chưa được áp dụng cho các biểu diễn ngữ nghĩa cấp cao từ mạng dữ liệu sâu của dữ liệu lớn viễn thám; có rất ít nghiên cứu trong lĩnh vực này. Hình ảnh viễn thám được tạo dựng dựa trên cơ sở vật lý và đã có nhiều mô hình thu nhận các bức xạ phản xạ từ các yếu tố ở nhiều dải tần số khác nhau. Các nghiên cứu về xử lý đo bức xạ là mong muốn cho các mô hình quy trình viễn thám xác định được và có thể giải thích được. Tuy nhiên, học sâu dựa trên thống kê được thiết kế để tìm mối tương quan giữa các hình ảnh đầu vào và mục tiêu đang muốn phân tích, điều này không phù hợp với đối tượng của việc xử lý đo bức xạ. Do đó, rất khó để sử dụng học sâu để cải thiện sự hiểu biết về quy trình viễn thám và các thuật toán học sâu hiện tại không phù hợp với xử lý bức xạ (quang phổ).

1.1.3. Mặt nạ đám mây

Mặt nạ đám mây là một phần của quá trình khai thác đối tượng và thuộc lĩnh vực phân tích ảnh viễn thám. Quá trình này lấy mây làm mục tiêu đặc biệt, nhận biết và xử lý để có thể bóc tách riêng mây khỏi các đối tượng địa lý khác trong ảnh. Trong quy trình ứng dụng RS, các vùng được bao phủ bởi đám mây được coi là không hợp lệ và tạo mặt nạ đám mây là một bước quan trọng để khoanh vùng và trích xuất các vùng không có đám mây (hợp lệ) từ hình ảnh viễn thám cho những mục tiêu phân tích sau. Vì vậy, mặt nạ đám mây là một chủ đề nghiên cứu nóng trong các ứng dụng viễn thám. Một mạng nơ-ron tích chập đơn giản với bốn lớp phức hợp và hai lớp được kết nối

đầy đủ đã được thiết kế để tạo mặt nạ đám mây trên các hình ảnh đa phương tiện, cho thấy triển vọng của mạng nơ ron tích chập đối với việc phát hiện, nhận diện và tạo dựng mặt nạ đám mây so với các cách tiếp cận truyền thống. Trong một nghiên cứu khác, một cách tiếp cận đa mục tiêu dựa trên kiến trúc U-Net để bóc tách mặt nạ đám mây đã được áp dụng trên bộ dữ liệu Landsat 8 Biome và SPARCS, cho thấy những cải tiến so với thuật toán FMask (Qiu, Zhu & He, 2019). Các nhà khoa học cũng áp dụng mô hình AE xếp chồng mờ để tích hợp khả năng học tập tính năng của mạng học sâu xếp chồng AE và khả năng phát hiện các chức năng mờ để xử lý các yếu tố mặt nạ đám mây có độ chính xác cao. Bằng cách tích hợp đồng thời các tính năng không gian cấp thấp và các tính năng ngữ nghĩa cấp cao, một mạng phức hợp hoàn chỉnh đã được phát triển để phân biệt mây và tuyết trong ảnh viễn thám. Dựa trên superpixel, Xie, Shi, Shi, Yin và Zhao (2017) đã thiết kế mạng nơ ron tích chập hai nhánh để trích xuất các tính năng đa cấp từ mỗi superpixel và dự đoán superpixel là một trong ba lớp: đám mây dày, đám mây mỏng và không bị mây che. Để có ranh giới chính xác hơn của các đám mây vùng lớn, Wu et al. (2020a) đã giới thiệu phương pháp đào tạo nhóm và tối ưu hóa ranh giới để cải thiện Mặt nạ R-mạng nơ ron tích chập cho việc xây dựng mặt nạ đám mây. Yuan, Meng và Cheng et al. (2017, tháng 9) đã đề xuất một mạng thần kinh học sâu đa nhiệm (hai nhiệm vụ là phân đoạn đám mây và phát hiện vùng cạnh đám mây) để phát hiện đám mây chính xác trong hình ảnh viễn thám.

Trong thực tế, sẽ khó khăn khi tải xuống các hình ảnh có diện tích lớn để tạo và xử lý vấn đề mặt nạ đám mây khi phân tích viễn thám. Sẽ rất hữu ích nếu chúng ta có được các sản phẩm tạo mặt nạ đám mây trực tiếp từ các nền tảng đám mây, chẳng hạn như Google Earth Engine (GEE) và Sentinel Hub.³ Thông qua các dịch vụ GEE, mô hình phát hiện đám mây dựa trên mạng nơ ron tích chập lần đầu tiên được đào tạo tại các máy local sau đó được triển khai đồng bộ trên GEE. Có thể phát hiện trực tiếp và loại bỏ các đám mây đối với hình ảnh Landsat-8 trong GEE (Yin, Ling, & Foody et al., 2020).

1.1.4. Hợp nhất dữ liệu

Vì có những xung đột không tương thích giữa các độ phân giải quang phổ, không gian và thời gian để chụp ảnh cảm biến, nên rất khó để một cảm biến đặc biệt

có thể đồng thời đạt được độ phân giải không gian cao, quang phổ cao và thời gian cao. Ví dụ: dữ liệu MODIS có chu kỳ truy cập ngắn hơn 1 ngày nhưng độ phân giải không gian thấp là 250 m, trong khi hình ảnh Landsat-8 có độ phân giải không gian cao hơn là 30 m nhưng chu kỳ truy cập lại dài hơn tới tận 16 ngày.

Kết hợp dữ liệu là quá trình kết hợp nhiều hình ảnh (được chụp thông qua các cảm biến khác nhau trong các cài đặt thông số khác nhau) để tạo ra một hình ảnh duy nhất kết hợp tất cả thông tin có ý nghĩa từ các hình ảnh riêng lẻ. Do đó, các phương pháp tổng hợp dữ liệu cung cấp hình ảnh tổng hợp (hợp nhất) với thông tin bổ sung. Một ví dụ điển hình của sự kết hợp dữ liệu viễn thám là làm sắc nét theo kiểu lia, cho biết sự kết hợp giữa hình ảnh đa quang độ có độ phân giải thấp và hình ảnh panchromatic (PAN) có độ phân giải cao để đạt được hình ảnh đa khung độ phân giải cao (còn được gọi là sự kết hợp quang phổ không gian). Hợp nhất phổ biến khác là hợp nhất không gian-thời gian, kết hợp hình ảnh có độ phân giải không gian cao với hình ảnh đa vùng để tạo ra hình ảnh có độ phân giải không gian cao đa vùng. Khi dữ liệu viễn thám được thu thập bởi các cảm biến khác nhau với các phương thức hình ảnh khác nhau (bao gồm cả hình ảnh quang học, dữ liệu SAR và LiDAR), quá trình tổng hợp đa phương thức được tiến hành.

(1) hợp nhất không gian-quang phổ

Bắt nguồn từ những tiến bộ gần đây đạt được trong siêu phân giải hình ảnh và khả năng tốt hơn của mạng học sâu trong việc mô tả mối quan hệ phức tạp giữa đầu vào và hình ảnh đích, nhiều phương pháp tổng hợp dữ liệu viễn thám sử dụng kỹ thuật học sâu đã được đề xuất.

Nghiên cứu tiên phong về khả năng làm sắc nét lia dựa trên học tập sâu đã sử dụng một mạng AE nhiều sóng không liên tục để mô hình hóa mối quan hệ từ các bản sửa lỗi hình ảnh có độ phân giải thấp đến độ phân giải cao. Một nghiên cứu được cải tiến đã kết hợp nhiều AE nhiều xếp chồng lên nhau để phát triển số liệu từ một số phương pháp học sâu để học cách nhúng các phần tử lân cận trong đa dạng phức tạp hình học được tinh chỉnh. Ngoài các mạng AE, mạng nơ ron tích chập cũng được sử dụng và hoạt động trong lĩnh vực làm sắc nét toàn bộ các vùng biên. Theo quy trình siêu phân giải cho hình ảnh tự nhiên, Masi, Cozzolino, Verdoliva và Scarpa (2016) đã

đề xuất một phương pháp làm sắc nét lia dựa trên mạng nơ ron tích chập, trong đó các hình ảnh PAN và đa kính ảnh được lấy mẫu xuống được xếp chồng lên nhau và đưa vào mạng và đầu ra của mạng với hình ảnh đa mặt ban đầu đã được so sánh. Một số nhà khoa học đã sử dụng quá trình học tập còn sót lại để phát triển một mạng nơ ron tích chập rất sâu với 11 lớp phức hợp để cải thiện độ chính xác của quá trình tổng hợp sắc nét. Để tái tạo lại các chi tiết không gian trong các hình ảnh đa phổ được lấy mẫu, người ta đã đề xuất phương pháp tiếp cận hai giai đoạn. Trong giai đoạn đầu, mạng nơ ron tích chập đã thu thập các đặc điểm không gian cấp trung và cấp cao từ ảnh PAN. Trong giai đoạn thứ hai, các mô hình chi tiết đã hợp nhất các chi tiết không gian được trích xuất trong giai đoạn đầu và xử lý thành hình ảnh đa phương diện. Liu và cộng sự. (2018c) đã đề xuất một thuật toán dựa trên GAN để làm sắc nét toàn bộ, trong đó kiến trúc tổng hợp hai luồng được thiết kế để tạo ra hình ảnh đa phổ có độ phân giải cao mong muốn và một mạng phức hợp được kết nối đầy đủ đóng vai trò phân biệt đối tượng, và nó được áp dụng để phân biệt hình ảnh đối tượng thực hoặc hình ảnh đa phương diện thông qua các nét chi tiết.

(2) hợp nhất không gian-thời gian

Dữ liệu tuần tự đa vùng là một công cụ quan trọng để quan sát sự thay đổi và biến động của bề mặt Trái đất. Tuy nhiên, chúng thường có độ phân giải không gian thấp, ví dụ: Dữ liệu MODIS. Do đó, kỹ thuật hợp nhất không gian-thời gian được sử dụng để tạo ra các lớp dữ liệu thời gian dày đặc với độ phân giải không gian cao (Li, Li, & He và cộng sự, 2020a). Các kỹ thuật này không chỉ sử dụng thông tin quy mô không gian như trong hợp nhất không gian-phổ và hợp nhất không gian-thời gian mà còn tận dụng tối đa các khuynh hướng thời gian và sự phụ thuộc trong các hình ảnh đa thời gian.

Nghiên cứu đột phá về sự kết hợp không gian-thời gian của học sâu là phương pháp deep-step (Das & Ghosh, 2016). Nó được bắt nguồn từ một mạng xếp chồng sâu, trong đó các tập hợp tính năng phát triển theo thời gian đã được sử dụng để kết hợp sự thay đổi theo thời gian cùng với việc huấn luyện tính năng không gian ở các lớp ẩn để dự đoán các mối tương quan giữa không gian-thời gian. Nhiều nỗ lực nghiên cứu đã được dành để phát triển các khuôn khổ mạng nơ ron tích chập khác nhau cho sự hợp nhất không gian-thời gian bằng cách sử dụng mạng nơ ron tích chập để tìm hiểu ảnh

xạ phi tuyến ở quy mô không gian phức tạp. Trong các mạng này, đầu vào là một cặp hình ảnh không gian thấp thời gian cao và thời gian thấp không gian cao để đào tạo và một cặp hình ảnh không gian thấp thời gian cao khác để dự đoán. Thông tin được hợp nhất dưới dạng các đối tượng địa lý dựa trên mạng nơ ron tích chập được trích xuất và sau đó các đối tượng địa lý đã hợp nhất được tái tạo lại thành hình ảnh dự đoán.

Ngoài ra, để nắm bắt thêm thông tin tạm thời, các khung mạng nơ ron tích chập hai luồng đã được thiết kế. Các nhà khoa học đã đề xuất mạng nơ ron tích chập hai luồng, xem xét sự phụ thuộc theo thời gian và tính nhất quán theo thời gian giữa các chuỗi hình ảnh trong quá trình siêu phân giải dựa trên mạng nơ ron tích chập. Tiến hành kết hợp cả hai nhánh dự báo xuôi và ngược dựa trên mạng nơ ron tích chập vào một khuôn khổ hai luồng để tổng hợp dữ liệu không gian-thời gian. Đó là nơi bản đồ được lập dựa trên thông tin không gian và dựa trên sự thay đổi thời gian được trình bày một cách đồng thời, giải quyết dự đoán về cả lớp phủ đất và hình dạng trạng thái của đối tượng.

(3) Hợp nhất đa phương thức

Khác với các phép hợp nhất ở trên sử dụng các hình ảnh từ cơ chế tạo ảnh giống nhau hoặc tương tự, phép hợp đa phương thức xem xét sự biến đổi và đồng hóa phi tuyến của các thông tin hình ảnh khác nhau trong phép kết hợp đa phương thức, tức là sự hợp nhất giữa hình ảnh quang học và dữ liệu SAR.

Trong hợp nhất quang học-SAR, kỹ thuật phổ biến là biến đổi dữ liệu không đồng nhất thành một không gian tính năng tương tự dựa trên các biểu diễn tính năng của học sâu. Một mạng học sâu đa phương thức kết hợp mạng nơ ron tích chập pseudo-Siamese và mô-đun lựa chọn kênh dựa trên các đặc trưng chủ yếu đã được đề xuất để kết hợp các đặc điểm sâu từ hình ảnh quang học và dữ liệu SAR nhằm phân loại lớp phủ đất. Ngoài ra, một số nghiên cứu đã giới thiệu GAN để thực hiện xử lý trực tiếp từ dữ liệu SAR sang hình ảnh quang học để tạo ra hình ảnh quang học không bị che phủ bởi những đám mây. Hơn nữa, dữ liệu chuỗi thời gian đã được sử dụng để cải thiện hiệu quả và độ ổn định trong chuyển đổi SAR sang dữ liệu quang học.

So với cảm biến quang học và cảm biến SAR, LiDAR đo khoảng cách hoặc vị trí của bề mặt đất. Sự kết hợp giữa hình ảnh quang học và dữ liệu LiDAR sẽ mang lại

lợi ích cho sự hiểu biết về bề mặt đất. Để khắc phục sự khác biệt rất lớn trong dạng thông tin, mạng hai hoặc đa kênh là các thủ thuật thường được sử dụng: một nhánh để trích xuất các đặc trưng quang phổ và nhánh còn lại để học các đặc trưng của sự tương quan không gian. Cuối cùng, đối với các nhiệm vụ tiếp theo, các tính năng của hai nhánh sẽ được làm phẳng và xếp chồng lên nhau.

Với số lượng ngày càng tăng của các mạng cảm biến tạo ra dữ liệu dạng điểm rời rạc khổng lồ, các nghiên cứu gần đây cố gắng kết hợp dữ liệu cảm biến với hình ảnh viễn thám để thu được các bản đồ chuyên đề chính xác và tốt hơn. Tận dụng độ tin cậy của dữ liệu viễn thám và các đặc tính động không gian-thời gian của dữ liệu cảm biến, người ta đã đề xuất phương pháp tiếp cận dựa trên học tập sâu end-to-end để kết hợp dữ liệu viễn thám với dữ liệu xã hội để nhận biết các chức năng của vùng đô thị ở các thành phố có mật độ dân số cao bằng cách sử dụng mạng nơ ron tích chập và LSTM.

1.2. Phân tích dữ liệu lớn viễn thám

1.2.1. Phát hiện đối tượng và trích xuất đối tượng

Trước khi trình bày về việc phát hiện đối tượng và trích xuất đối tượng, điều đầu tiên chúng ta phải chỉ ra sự khác biệt của chúng, vì cả hai đều có các ứng dụng viễn thám tương tự và thường dễ bị nhầm lẫn. Trong nghiên cứu này, phát hiện đối tượng là quá trình tìm kiếm các thể hiện của đối tượng trong hình ảnh đầu vào, ví dụ: các vị trí trượt lở, các đối tượng nhà ở hay các phương tiện như tàu thủy, ô tô, máy bay hoặc máy phát điện gió. Nó tạo ra một hoặc nhiều vùng giới hạn viền với nhãn lớp được gắn vào mỗi vùng giới hạn đó. Tuy nhiên, trích xuất đối tượng, một phần mở rộng hơn nữa của phát hiện đối tượng, mô tả sự hiện diện của một đối tượng (chẳng hạn như vùng nước, đường và khu vực đô thị) thông qua các mặt nạ pixelwise được tạo cho mỗi đối tượng trong hình ảnh, nó gần với kỹ thuật phân đoạn các cá thể trong thị giác máy tính. Nhiệm vụ này khó hơn nhiều so với việc phát hiện đối tượng.

(1) Phát hiện đối tượng

Khởi nguồn từ sự thành công của các thuật toán học sâu để phát hiện đối tượng trong các hình ảnh tự nhiên, nhiều phương pháp phát hiện đối tượng dựa trên học sâu

đã được phát triển cho dữ liệu RS. Nó cho thấy những lợi thế khác biệt so với các phương pháp truyền thống vì chúng có khả năng học và được huấn luyện tốt hơn các biểu diễn ngữ nghĩa cấp cao. Tuy nhiên, có một số thách thức đối với việc phát hiện đối tượng, bao gồm trị số của cặp giá trị độ lệch và tỷ lệ vùng nghiên cứu và số lượng mẫu được gắn nhãn để đào tạo là có giới hạn.

Để giảm nhu cầu về các mẫu đào tạo, một số phương pháp phát hiện đã chuyển các mạng nơ ron tích chập được đào tạo trước để phát hiện đối tượng. Zhou, Cheng, Liu, Bu và Hu (2016) đã trình bày một khung học tập được giám sát yếu để đào tạo một hệ thống dò đối tượng, sử dụng mô hình mạng nơ ron tích chập được đào tạo trước để trích xuất các tính năng và một lược đồ với vị trí khởi tạo ban đầu âm để hội tụ nhanh hơn. Ví dụ các nghiên cứu của Li, Zhang, Lei, Wang và Guo (2020c) đã sử dụng phương pháp học chuyển giao và tinh chỉnh để đào tạo ra ba mô hình dựa trên mạng nơ ron tích chập được thiết lập với các giá trị trọng số để phát hiện các đối tượng nhà kính trong nông nghiệp từ các hình ảnh viễn thám đa nguồn.

(2) Trích xuất đối tượng

Việc trích xuất đối tượng có thể được hình thành dưới dạng phân loại biclass hoặc phân loại đối tượng qua quá trình phân đoạn, nhằm dự đoán lớp pixelwise của mỗi đối tượng trong ảnh viễn thám. Những đối tượng này có thể là các tòa nhà, đường xá, vùng chứa nước và các khu vực quan tâm khác.

Trích xuất các tòa nhà là mục tiêu quan trọng trong dữ liệu viễn thám có độ phân giải không gian cao. Nhiều mạng học sâu được thiết lập cho hình ảnh tự nhiên đã được cải thiện để xây dựng quá trình trích xuất. Ngoài ra, một số thủ thuật hữu ích đã được sử dụng để đạt được kết quả tốt hơn, bao gồm học có giám sát, học tập đa nhiệm, hợp nhất dữ liệu theo mô không gian hình nón và mô hình tích chập lan rộng.

Các vùng chứa nước cũng là mục tiêu phổ biến trong ảnh viễn thám. Nhiều khuôn khổ học tập sâu khác nhau đã được phát triển để trích xuất các khu vực chứa nước. Điều đáng nói là các vùng nước bao gồm hồ, sông và biển. Do đó, khai thác vùng nước từ ảnh viễn thám cũng gặp phải các vấn đề tương tự như khai thác xây dựng và đường, chẳng hạn như mục tiêu nhiều cấp, mất cân bằng lớp và độ phân giải pixel của ảnh.

1.2.2. Phân loại hình ảnh sử dụng đất và lớp phủ

Phân loại sử dụng đất / lớp phủ (LULC) là phân tích nền tảng cho dữ liệu RS, đã nâng cao từ pixelwise sang mô hình phân tích dựa trên đối tượng. Phân loại Pixelwise LULC, gán nhãn loại đất cho mỗi pixel trong một hình ảnh, có quy trình kỹ thuật tương tự như phân đoạn ngữ nghĩa trong thị giác máy tính. Phân loại LULC pixelwise gần đây sử dụng mạng phân đoạn ngữ nghĩa đã đạt được những tiến bộ lớn. Đặc biệt, để khám phá thêm thông tin tô pô không gian và biểu đồ, GNN đã được sử dụng trong các nhiệm vụ phân loại LULC trên ảnh siêu kính, hình ảnh vệ tinh có độ phân giải rất cao, và hình ảnh thu nhận theo chuỗi thời gian.

So với phân tích hình ảnh pixelwise, phân tích hình ảnh dựa trên đối tượng sử dụng đầy đủ các đặc điểm khác nhau của đối tượng (vùng hoặc mảng), chẳng hạn như các đặc điểm quang phổ, kết cấu và hình học. Từ quan điểm này, bằng cách học các đại diện khu vực đa tỷ lệ, các thuật toán học sâu bẩm sinh rất thích hợp để phân loại LULC dựa trên đối tượng. Thủ tục truyền thống của phân loại dựa trên đối tượng thường bao gồm tạo đối tượng và phân loại đối tượng. Các nghiên cứu dựa trên deep-learning hiện tại tập trung vào phân loại đối tượng theo giả định của bản đồ đối tượng có sẵn. Lấy các bản vá có kích thước 5×5 làm đối tượng, Sharma, Liu, Yang và Shi (2017) đã đề xuất một khung mạng nơ ron tích chập dựa trên bản vá sâu cho dữ liệu viễn thám có độ phân giải trung bình. Fu, Ma, Li và Johnson (2018) đã tạo ra các bản sửa đổi bằng cách sử dụng trung tâm của các đối tượng được phân đoạn (kết quả từ phân đoạn đa phân giải) làm trung tâm và cố định cửa sổ ở 32×32 pixel và 64×64 pixel và sau đó phân loại các bản bổ sung này thông qua mạng nơ ron tích chập. Một chiến lược khác kết hợp học sâu với phân loại dựa trên đối tượng, dựa vào mạng nơ ron tích chập để dự đoán loại đất ở cấp pixel bằng cách phân loại theo từng mảng và sau đó bỏ phiếu để xác định loại đối tượng được phân đoạn trên cấp đối tượng bằng các nhãn giả của pixel. Một số nghiên cứu cố gắng sử dụng dữ liệu bổ sung để tạo các đối tượng. Huang, Zhao và Song (2018) đã sử dụng một phương pháp phân hủy dựa trên khung xương sử dụng mạng lưới đường để chia hình ảnh thành các vùng thông thường và đưa chúng vào một mạng nơ ron phức hợp sâu để phân loại sử dụng đất đô thị.

Khi xem xét cẩn thận các phân loại LULC dựa trên đối tượng này bằng cách sử dụng học sâu, chúng tôi nhận thấy rằng các đối tượng từ phân đoạn truyền thống hoặc các nguồn dữ liệu khác gặp phải các vấn đề về phân đoạn dưới / giám sát và ranh giới thô, khiến không thể đại diện chính xác các tính năng của LULC. Một số nghiên cứu tiên phong đã cố gắng sử dụng các thuật toán học sâu để tạo ra các đối tượng mặt đất có ranh giới chính xác. Ví dụ, Waldner và Diakogiannis (2020) đã sử dụng một mạng nơ-ron tích tụ sâu với một đường trục U-Net được kết nối đầy đủ có các tính năng chập giãn và suy luận có điều kiện để trích xuất ranh giới trường từ hình ảnh viễn thám. Liu và cộng sự. (2020d) đã kết hợp ba mạng nơ-ron tích chập trong một sơ đồ phân cấp để trích xuất các thửa đất nông nghiệp chính xác ở các khu vực miền núi. Tuy nhiên, ranh giới hoặc các cạnh trong hình ảnh là các đối tượng địa lý cấp thấp. Rất khó để các thuật toán học sâu với tích chập và tổng hợp để tạo ra các cạnh pixel. Một số thủ thuật đã được giới thiệu để cải thiện độ chính xác của ranh giới, bao gồm tích chập giãn, lấy mẫu tăng và xử lý hậu kỳ trên các cạnh được trích xuất bằng cách sử dụng trường ngẫu nhiên có điều kiện và các thuật toán phân tích không gian (Liu và cộng sự, 2019a; Marmanis và cộng sự, 2018; Pan, Zhao, & Xu, Năm 2020).

1.2.3. Phát hiện các yếu tố thay đổi và biến động

Phát hiện thay đổi là quá trình xác định những thay đổi trên bề mặt đất (chẳng hạn như lũ lụt, sạt lở đất, thảm thực vật phát triển và xây dựng đô thị) bằng cách sử dụng nhiều dữ liệu viễn thám thu được tại các thời điểm khác nhau. Với việc tăng cường các hoạt động của con người và lượng dữ liệu viễn thám có độ phân giải thời gian cao không lồ sẵn có, việc phát hiện thay đổi đã được nghiên cứu rộng rãi trong những thập kỷ gần đây.

Phương pháp chủ đạo để phát hiện thay đổi là xác định các thay đổi trên bề mặt đất từ các ảnh nguồn. Khi xếp chồng hai hình ảnh lại với nhau, vấn đề phát hiện thay đổi có thể được chuyển thành phân loại biclass, khoanh vùng vùng bị thay đổi thông qua các thuật toán học sâu. Nó thường bao gồm hai bước chính, bao gồm biểu diễn tính năng và phân đoạn ngữ nghĩa. Đối với các hình ảnh đồng nhất thu được từ các cảm biến tương tự hoặc giống nhau (ví dụ: hình ảnh quang học), các mạng nơ-ron tích chập giống nhau có thể được áp dụng để trích xuất các đại diện tính năng đa tỷ lệ và khoanh vùng các vùng thay đổi. Đối với dữ liệu viễn thám không đồng nhất (ví dụ:

hình ảnh quang học và dữ liệu SAR), mạng biến đổi trước tiên đã chuyển đổi một hoặc hai điểm dữ liệu vào cùng một không gian đặc trưng. Sau đó, các tính năng này được đưa vào các mạng ở trên được thiết kế cho hình ảnh đồng nhất để xác định các vùng thay đổi.

Vì việc thu thập đủ các mẫu được dán nhãn để đào tạo có giám sát thường tốn nhiều thời gian và công sức, nhiều nỗ lực đã được thực hiện để đạt được khả năng phát hiện thay đổi dựa trên học sâu theo cách thức không giám sát hoặc bán giám sát. Mã hóa các hình ảnh đa mục tiêu như biểu đồ, nhiều nhà khoa học đã áp dụng GCN để truyền tải thông tin có nhãn từ một số mẫu đào tạo sang các mẫu không được gắn nhãn để phát hiện thay đổi được giám sát bán chu kỳ. Du, Ru, Wu và Zhang (2019) đã sử dụng lý thuyết phân tích tính năng chậm để loại bỏ các thành phần không thay đổi và làm nổi bật các thành phần đã thay đổi của các tính năng được biến đổi từ mạng sâu đối xứng để tìm các pixel không thay đổi với độ tin cậy cao làm mẫu huấn luyện. Góc khác thì một số nhà khoa học đã sử dụng các nhãn sai được tạo ra bởi phân cụm không gian mờ để chuyển đổi quy trình đào tạo có giám sát thành quy trình học tập không được giám sát để phát hiện các thay đổi chưa được đánh giá.

1.2.4. Phân tích viễn thám đa thời gian

Với lượng dữ liệu viễn thám ngày càng tăng, phân tích đa mục tiêu đã trở thành một kỹ thuật quan trọng trong các ứng dụng viễn thám động và tuần hoàn. Do sự giống nhau giữa dữ liệu tuần tự viễn thám với video và âm thanh, nhiều khuôn khổ học sâu trong thị giác máy tính có thể được sử dụng để phân tích đa mục tiêu viễn thám.

Với khả năng nắm bắt các phụ thuộc dài hạn, mạng nơ ron hồi quy đã được sử dụng rộng rãi để phân tích đa mục tiêu trong các ứng dụng viễn thám. Liu và cộng sự. (2020d) đã thiết kế mạng LSTM hai lớp để phân loại trình tự chuỗi thời gian. Sun, Di và Fang (2019) đã thiết kế một khuôn khổ dựa trên LSTM để tận dụng mô hình thời gian của sự phát triển cây trồng trong chuỗi thời gian hình ảnh để cải thiện độ chính xác của việc lập bản đồ cây trồng. Dựa trên bản đồ thửa đất, Zhou et al. (2019) đã đề xuất phương pháp phân tích chuỗi thời gian dựa trên LSTM sử dụng dữ liệu SAR đa mục tiêu để phân loại cây trồng, đạt được độ chính xác tổng thể cải thiện 5,0% so với các phương pháp truyền thống. Hơn nữa, Zhao et al. (2019) đã so sánh một số kiến

trúc Mạng nơ ron hồi quy-tương đối để phân loại loại cây trồng bằng cách sử dụng dữ liệu Sentinel-2 đa mục tiêu, bao gồm mạng nơ ron tích chập 1D, LSTM, GRU và một mạng tự chú ý mới. Trong các nghiên cứu này, các ứng dụng dựa trên mạng nơ ron hồi quy bao gồm hai bước tách biệt: trích xuất các đặc trưng không gian và xây dựng dữ liệu chuỗi thời gian và phân tích chuỗi thời gian. Thiếu các khuôn khổ end-to-end.

Mạng LSTM tích hợp (ConvLSTM) kết nối mạng tích hợp với mạng lặp lại được đề xuất để học đồng thời cách biểu diễn đặc điểm không gian và thời gian để phân loại và dự đoán không gian-thời gian. Trong các ứng dụng của hình ảnh siêu kính, các nhà nghiên cứu đã lập mô hình phụ thuộc dài hạn trong miền quang phổ và sử dụng mạng ConvLSTM để trích xuất các đặc điểm phổ không gian phân biệt hơn để đạt được các khuôn khổ end-to-end. Sự thành công gần đây của cơ chế chú ý trong dữ liệu tuần tự đã truyền cảm hứng cho các ứng dụng của nó trong phân tích đa mục tiêu. Ví dụ, Feng et al. (2020) đã đề xuất một mạng nơ-ron phức hợp lặp lại dựa trên sự chú ý để lập bản đồ thực vật chính xác từ hình ảnh màu đỏ-xanh lục-xanh lam của các thiết bị bay UAV đa vùng.

Từ quan điểm của định dạng dữ liệu, dữ liệu chuỗi thời gian có thể được coi là một trường hợp đặc biệt của hình ảnh, trong đó chiều cao = 1. Do đó, mạng nơ ron tích chập và các biến thể được sử dụng trong miền hình ảnh có thể dễ dàng chuyển để phân tích đa mục tiêu, ví dụ: mạng nơ-ron tích chập tạm thời (TCN). TCN là một khuôn khổ sử dụng các biến đổi chập và giãn nở thông thường và do đó thích ứng với dữ liệu tuần tự với tính thời gian và các trường tiếp nhận lớn. Yan, Chen, Chen và Liang (2020a) đã đề xuất một phương pháp dự đoán theo chuỗi thời gian sử dụng TCN để thực hiện dự đoán nhiều bước về lớp phủ đất từ các hình ảnh viễn thám của chuỗi thời gian dày đặc.

1.2.5. Các nhiệm vụ phân tích viễn thám khác

Ngoài các lĩnh vực ứng dụng thông thường ở trên trong phân tích hình ảnh RS, học sâu đã thành công trong nhiều lĩnh vực thú vị, chẳng hạn như:

+ Kiểm đếm: Trong một khu vực quy mô lớn, ví dụ: trên toàn cầu, rất khó để đếm số lượng vật thể đặc biệt, chẳng hạn như cây cối, động vật hoang dã (Peng, Wang, Liao, Shao, Sun, Yue & Ye, 2020) và cá voi (Guirado, Tabik, Rivas, Alcaraz-Segura,

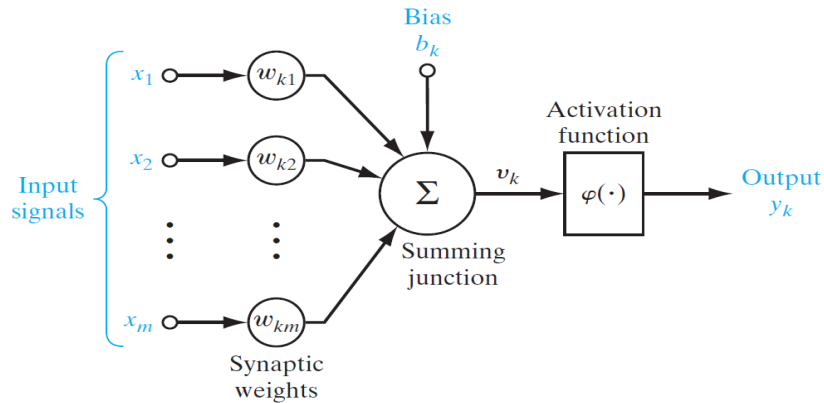
& Herrera, 2019), trong hình ảnh viễn thám cho các phương pháp truyền thống. Brandt và cộng sự. (2020) đã sử dụng học sâu để lập bản đồ kích thước phủ rộng của mỗi cây có kích thước hơn 3 m² trên diện tích đất rộng 1,3 triệu km² ở Tây Phi Sahara, Sahel và các vùng ẩm thấp, khám phá vai trò của chúng trong việc giảm thiểu suy thoái, biến đổi khí hậu và nghèo đói.

+ Phân tích các chỉ số vĩ mô. Một số nghiên cứu đã áp dụng các kỹ thuật học sâu trên dữ liệu lớn viễn thám và vệ tinh quan sát để truy xuất thông tin về dung tích của các bể chứa dầu, các nhà máy điện than và các nhà máy điện mặt trời để tính toán các chỉ số kinh tế vĩ mô, đo lường thương mại quốc tế, cường độ sản xuất và dự trữ nguyên liệu.

CHƯƠNG 2:

CÁC MÔ HÌNH HỌC SÂU PHỔ BIẾN

Lần đầu tiên xuất hiện vào những năm 1960, perceptron là cơ sở của các mạng nơ-ron (Neural Networks – NN) nhân tạo sớm nhất. Đây là một mô hình kỹ thuật sinh học cơ bản ứng dụng phân loại nhị phân nhằm mục đích mô hình hóa về mặt toán học cách hoạt động của một tế bào thần kinh sinh học.



Hình 2.1: Mô hình mạng Nơ ron nhân tạo

Tuy nhiên, perceptron không thể thực hiện bằng phương pháp phân loại phi tuyến. Từ đó, thuật toán lan truyền ngược bắt đầu đào tạo mạng nơ-ron nhân tạo nhiều lớp (NN) vào những năm 1980, thúc đẩy sự phát triển vượt bậc của mạng nơ-ron (Rumelhart, Hinton, & Williams, 1985).

Mạng nơ-ron có thể giải các bài toán phi tuyến, nhưng thông thường nó chỉ chứa một lớp ẩn (được gọi là NN nông) vì khó đào tạo một mạng nơ-ron nhiều lớp. Cho đến năm 2006, các phương pháp luận phức tạp hơn đã được đề xuất để đào tạo mạng nơ-ron với nhiều hơn một lớp ẩn (gọi là mạng nơ-ron sâu) (Hinton & Salakhutdinov, 2006). Cho đến nay, các mạng nơ-ron trải qua các quá trình huấn luyện học máy và đã bước vào kỷ nguyên học sâu với rất nhiều các mô hình huấn luyện cho kết quả chính xác cao.

Từ đó, các mô hình phân tích và đào tạo trong học máy lần lượt được ra đời. Từ các lưới pixel thông thường trên ảnh vệ tinh thích hợp cho các phép toán tích chập, các mạng nơ-ron tích chập (Convolutional neural networks) đã thành công trong các nhiệm vụ xử lý các hình ảnh nói chung và ảnh vệ tinh nói riêng. Bên cạnh đó, khắc phục những yếu điểm của quá trình tính toán qua các lớp chập, mạng nơ-ron hồi quy

(Recurrent neural networks) đã được nghiên cứu và áp dụng, và trở thành phương án tốt nhất cho các dạng dữ liệu tuần tự có thêm các yếu tố thời gian. Hay trong một khuôn khổ học không giám sát, các mã tự động sinh ra từ mạng nơ ron tự lập mã(AE) có thể được xếp chồng lên nhau để tìm hiểu các biểu diễn sâu hơn của các lớp dữ liệu từ đó cho phép tạo ra một tập dữ liệu ban đầu có tính tương đồng cao. Trong khi đó, mô hình mạng nơ ron đối luyện (GAN) tìm hiểu đặc điểm chung của phân phối dữ liệu thông qua các phương pháp đối nghịch (Gui, Sun, & Wen và cộng sự, 2020). Mạng nơ-ron đồ thị (GNN) thì phù hợp để xử lý dữ liệu có cấu trúc không đồng đều thông qua việc phân bố dữ liệu trên miền đồ thị(Zhou, Cui, & Zhang và cộng sự, 2018). Bên cạnh những mô hình phổ biến này, cũng có rất nhiều biến thể và sự lai tạo của từ những mô hình kiến trúc này tạo nên một hệ đa dạng các mô hình học máy ứng dụng. Dưới đây là danh sách những mạng nơ ron phổ biến và được ứng dụng trong phổ biến trong việc giải đoán và phân tích ảnh vệ tinh.

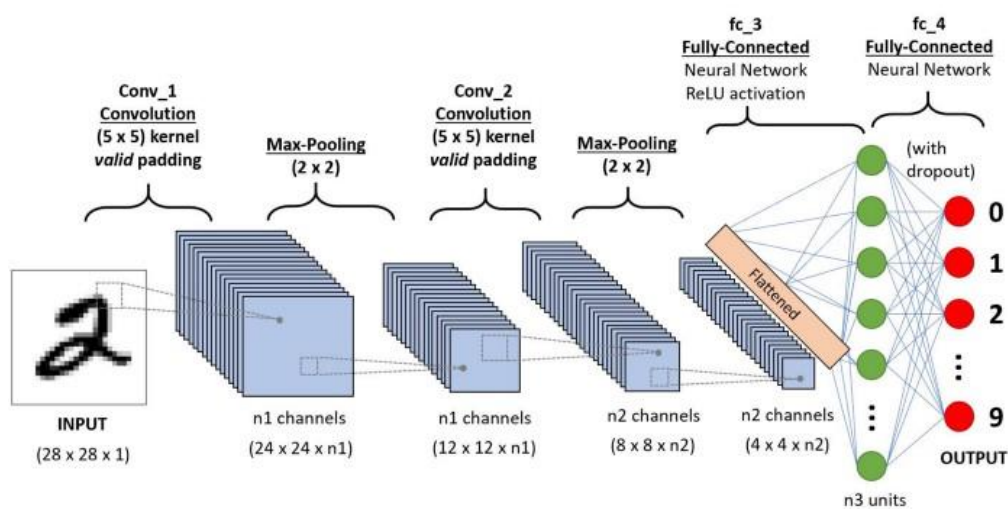
Model	Features and advantages	Variants or implementation
CNN	Image processing and analysis	LeNet, R-mạng nơ ron tích chập, VggNet, ResNet, U-Net
RNN	Sequential data and time-series analysis	LSTM, BiLSTM, GRU
AE	Feature extraction and representation	Stacked AE, Sparse AE, Variational AE, Denoising AE
GAN	Data generation, unsupervised learning	Conditional GAN, f-divergence GAN, Cycle-consistent GAN
GNN	network analysis of irregular data	RGNN, GCN, GAT, GAE, STGNN

Bảng 2.1 Các mô hình mạng nơ ron phổ biến nhất được sử dụng trong viễn thám

2.1. Mạng nơ-ron tích chập (Convolutional neural networks - CNN)

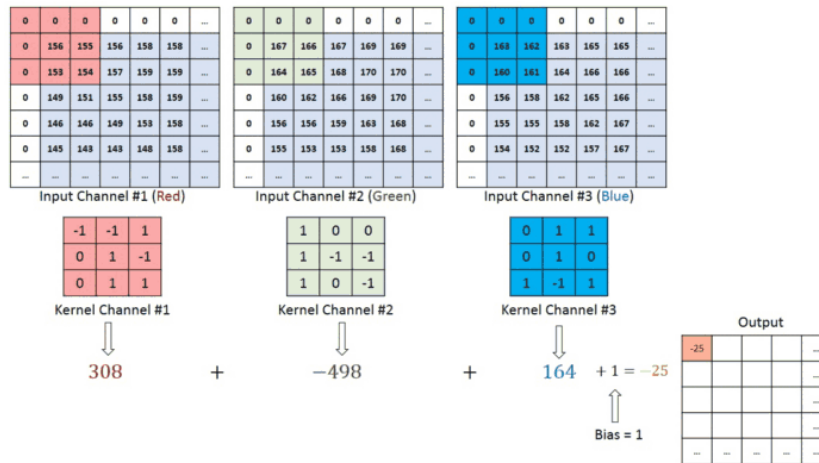
Mạng nơ ron tích chập được thiết kế để tận dụng cấu trúc dạng lưới đa chiều của hình ảnh đầu vào (LeCun, Bengio, & Hinton, 2015). Vì vậy, mô hình này đã cực kỳ thành công trong việc xử lý dữ liệu ảnh và ảnh viễn thám. Mô hình này sử dụng sự chia sẻ trọng số của các ô pixel và kết nối cục bộ thông qua mạng lưới trượt. Mạng nơ ron tích chập thiết lập các quá trình liên kết sâu để tìm hiểu các tính năng từ cấp pixel đến cấp ngữ nghĩa cho các ứng dụng cao hơn. Mạng nơ ron tích chập tiêu chuẩn

thường bao gồm ba loại lớp thần kinh chính: lớp chập, lớp gộp và lớp kết nối đầy đủ. Mỗi lớp của mạng nơ ron tích chập chuyển đổi khối lượng đầu vào thành khối lượng kích hoạt các nơ-ron đầu ra, cuối cùng dẫn đến một tổng thể với các lớp trong mô hình được kết nối đầy đủ, dẫn đến ánh xạ dữ liệu đầu vào từ các ô lưới đa chiều thành vectơ đặc trưng một chiều (1D). Đối với một nhiệm vụ đặc biệt, tính năng 1D được đưa vào một lớp perceptron để gán nhãn lớp hoặc tính toán xác suất của lớp đã cho trong đầu vào.



Hình 2.2: Mô hình kiến trúc của mạng nơ ron tích chập - CNN

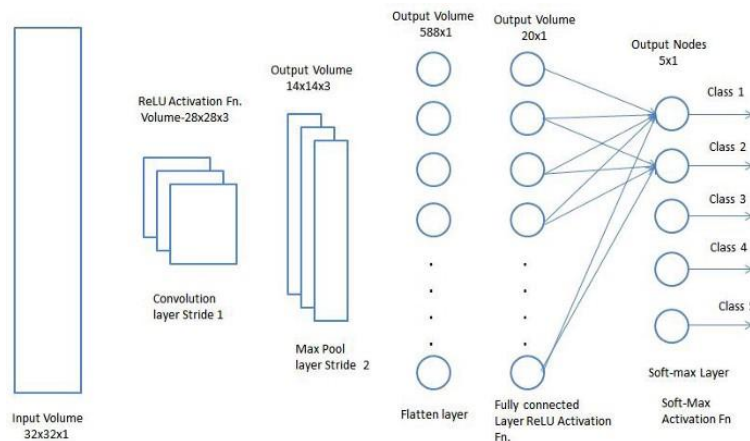
Lớp chập là một phần thiết yếu của mạng nơ ron tích chập, chuyển đổi dữ liệu đầu vào thành một biểu diễn dữ liệu ở mức trừu tượng hơn. Nó sử dụng các phép nhân tích chập để trượt khung qua hình ảnh đầu vào, thực hiện phép toán tích chập giữa từng vùng đầu vào và nhân lớp trượt, từ đó tạo ra các bản đồ đặc trưng. Thông thường, sẽ là một hàm tính toán phi tuyến tính toán và trượt theo chiều kim đồng hồ được áp dụng cho các kết quả tính toán tuyến tính của các lớp chập để cải thiện khả năng biểu diễn phi tuyến của mạng nơ ron tích chập.



Hình 2.3: Mô phỏng lớp tích chập sử dụng cửa sổ 3x3

Thông thường, sau một lớp phức hợp, một lớp tổng hợp được sử dụng để giảm kích thước không gian (chiều rộng và chiều cao) của bản đồ các đối tượng địa lý. Nó tóm tắt các đối tượng địa lý từ lớp phức hợp đến cấp độ trừu tượng hơn bằng cách trượt bộ lọc hai chiều (2D) qua từng dải của bản đồ đối tượng địa lý và ghi lại các giá trị sau tính toán (ví dụ: tìm giá trị lớn nhất và trung bình) của các đối tượng địa lý trong khu vực được bộ lọc bao phủ. Hoạt động như vậy dẫn đến mất thông tin không gian nhưng đồng thời cải thiện tính tổng quát và mạnh mẽ của mạng nơ ron tích chập.

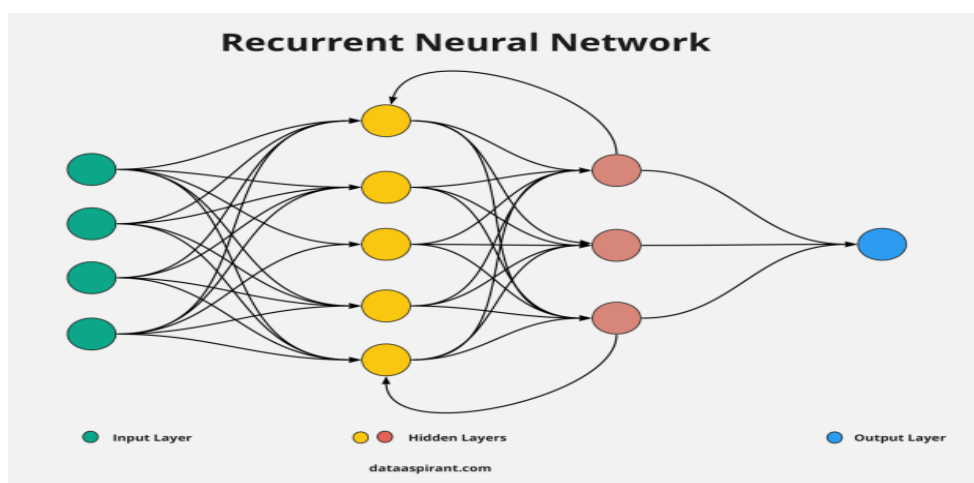
Lớp quan trọng thứ 3 trong mô hình này là lớp kết nối đầy đủ, thường được thêm vào giữa lớp gần cuối và lớp đầu ra để mô hình hóa thêm các mối quan hệ phi tuyến của các đặc trưng đầu vào. Trong một lớp kết nối đầy đủ, các tế bào thần kinh có đầy đủ các kết nối với tất cả các hoạt động của nó trong lớp trước đó. Các lớp được kết nối đầy đủ cuối cùng chuyển đổi bản đồ đối tượng địa lý 2D thành vectơ đối tượng địa lý 1D.



Hình 2.4: Kiến trúc mô hình lớp kết nối đầy đủ

2.2. Mạng nơ ron hồi quy (Recurrent Neural Networks RNN)

Bằng cách khai thác sự phụ thuộc cục bộ của thông tin hình ảnh, mạng nơ ron tích chập đã chứng minh kết quả thiết lập kỷ lục trong nhiều ứng dụng hình ảnh. Tuy nhiên, mạng nơ ron tích chập dựa trên giả định về sự độc lập giữa các lớp chập. Điều này là một tính toán chưa đầy đủ đối với dữ liệu có chứa các yếu tố liên quan đến thời gian hoặc không gian. Ví dụ như các dữ liệu video, âm thanh và văn bản. Mạng nơ ron hồi quy là mô hình kết nối với khả năng truyền thông tin có chọn lọc qua các bước trình tự liên tục (Graves, Mohamed & Hinton, 2013). Do đó, chúng có thể mô hình hóa sự phụ thuộc tuần tự và thời gian trên nhiều cấp độ đo. Đặc điểm điển hình của kiến trúc mạng nơ ron hồi quy là kết nối tuần hoàn, cho phép mạng nơ ron hồi quy có khả năng cập nhật trạng thái hiện tại dựa trên các trạng thái trong quá khứ và dữ liệu đầu vào ở thời điểm hiện tại. Do đó, mạng nơ ron hồi quy được ứng dụng rộng rãi cho các tập dữ liệu tuần tự.



Hình 2.5 Mô hình kiến trúc mạng nơ ron hồi quy - RNN

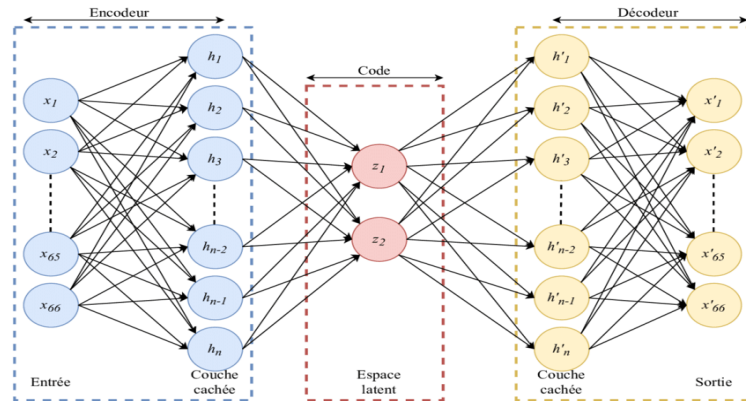
Tuy nhiên, các mạng nơ ron hồi quy bao gồm các hàm phụ thuộc sigma hoặc tanh không thể huấn luyện và kết nối thông tin liên quan khi chuỗi vector đầu vào là quá lớn. Để khắc phục và xử lý vấn đề “các bộ nhớ phụ thuộc dài hạn”, chuỗi các bộ nhớ phụ thuộc ngắn hạn (long short-term memory - LSTM) đã được đề xuất nghiên cứu và đưa vào áp dụng (Hochreiter & Schmidhuber, 1997), trong đó các “cổng” được giới thiệu để kiểm soát các phụ thuộc. Mỗi khối LSTM bao gồm một ô, chiếm giữ một phần bộ nhớ của khối và ba cổng: một cổng vào, một cổng ra và một cổng ẩn (còn gọi là cổng giữ) (Gers, Schraudolph, & Schmidhuber, 2002). Ba cổng kiểm soát luồng

thông tin đi qua nút xử lý. Một đơn vị LSTM có thể nhớ và lưu giữ các giá trị trong các khoảng thời gian tùy ý. Thiết kế LSTM này kết hợp các điều khiển phi tuyến, phụ thuộc vào dữ liệu xử lý bởi Mạng nơ ron hồi quy, có thể được huấn luyện để đảm bảo rằng độ hội tụ của hàm mục tiêu đối với gradient trạng thái không bị mất. Do đó, LSTM thể hiện khả năng chứa dung lượng bộ nhớ lớn dữ liệu và giải quyết vấn đề gradient biến mất trước khi quá trình hội tụ. Nó đã được sử dụng rộng rãi trong các loại nhiệm vụ tuần tự khác nhau, chẳng hạn như phân loại và dự đoán chuỗi thời gian.

Khả năng tự học của LSTM vượt trội so với khả năng huấn luyện của các perceptron hồi quy tiêu chuẩn. Tuy nhiên, việc bổ sung các tham số sau mỗi lần học làm tăng gánh nặng tính toán. Do đó, nút đơn vị hồi quy (GRU) đã được tạo ra nhằm bổ sung và hỗ trợ cho các quá trình tính toán (Chung, Gulcehre, & Cho et al., 2014). Nút đơn vị hồi quy GRU này tích hợp cổng ẩn và cổng đầu vào của ô LSTM để tạo giá trị ở cổng cập nhật. Tuy nhiên, nút GRU chỉ có hai cổng: một cổng cập nhật và một cổng đặt lại nên ô GRU đơn lẻ được tạo phát huy những khả năng huấn luyện kém hơn so với LSTM ban đầu.

2.3. Mạng Nơ ron tự lập mã (Autoencoder neural networks – AE)

Mạng nơ ron tự lập mã được thiết kế để học biểu diễn nhằm khám phá cấu trúc ẩn trong dữ liệu đầu vào và để lựa chọn tính năng và giảm kích thước cho các tập dữ liệu huấn luyện. Nó bao gồm ba thành phần: bộ mã hóa, bản mã và bộ giải mã (Dong, Liao, Liu và Kuang, 2018). Bộ mã hóa nén đầu vào để tạo bản mã và bộ giải mã sẽ tạo lại đầu vào dựa trên bản mã tương ứng. Đặc tính chung của AE là kích thước của lớp đầu vào và đầu ra giống như một kiến trúc đối xứng (Hinton & Salakhutdinov, 2006). Ý tưởng cơ bản là học cách ánh xạ từ một mẫu đầu vào x sang một mã hóa mới $c = h(x)$, trong điều kiện lý tưởng là xuất ra cùng một mẫu mà nó sử dụng cho đầu vào, tức là $x \approx y = g(c)$. Do đó, bộ bản mã c , thường có kích thước thấp hơn x , cho phép chúng ta tái tạo (hoặc mã hóa) x với tốc độ và quy mô tính toán được tiết kiệm hơn.



Hình 2.6: Sơ đồ kiến trúc mạng Nơ ron tự lập mã - AE

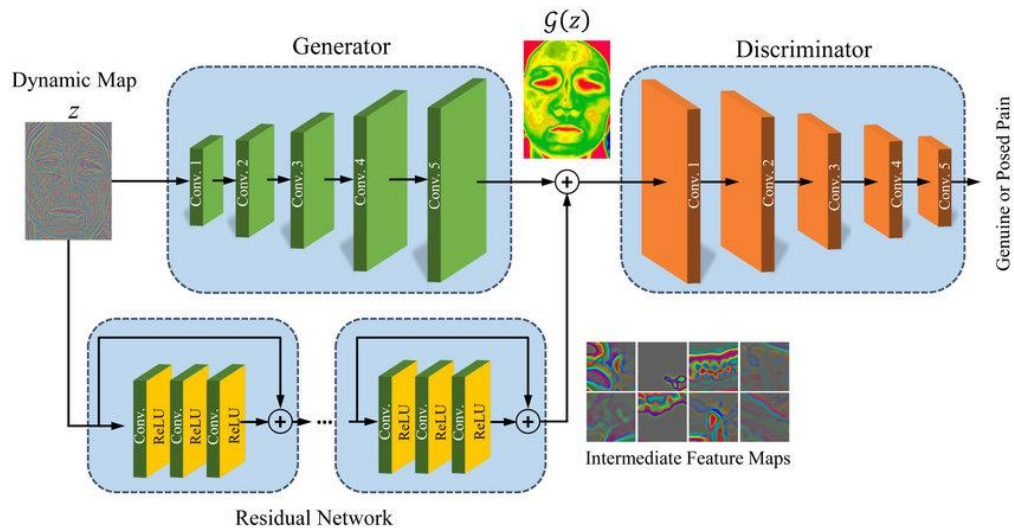
Bằng cách sử dụng các quy chuẩn khác nhau trong kiến trúc tự động mã hóa chung, nhiều biến thể trung gian đã được đưa ra để biểu diễn quá trình lập mã đảm bảo quá trình tạo mã tạo ra một không gian có ý nghĩa và có thể khái quát được dữ liệu gốc ban đầu. Khi biểu diễn mạng nơ ron tự lập mã, các giá trị nhiễu ngẫu nhiên được cố tình thêm vào dữ liệu đầu vào, từ đó mô hình buộc phải xây dựng lại dữ liệu ban đầu như chưa được sửa đổi. Điều này cho phép tìm kiếm một biểu diễn tối ưu nhất có khả năng đại diện tất cả các cấu trúc ổn định dưới dạng các hàm phụ thuộc và đặc trưng cơ bản nhất của tập dữ liệu ban đầu.

Mạng nơ ron tự lập mã được hiểu đơn giản là mô hình ánh xạ đầu vào x thành một biểu diễn ngầm thông qua một lớp mã hóa ẩn (Zhu và cộng sự, 2017). AE thường được phân loại dưới dạng xếp chồng hoặc AE theo chiều sâu (SAE) và là một mạng nơ-ron bao gồm nhiều lớp AE, trong đó các đầu ra của mỗi lớp AE ở trên (hoặc lớp trước) sẽ được nối với các đầu vào của lớp AE kế tiếp trong chuỗi huấn luyện của mình.

2.4. Mạng Nơ ron đối luyện (Generative adversarial networks – GAN)

Mạng nơ ron đối luyện tìm hiểu mô hình chung của tập dữ liệu huấn luyện ban đầu thông qua các phương pháp đối nghịch (Goodfellow, Pouget-Abadie, & Mirza et al., 2014). Lấy cảm hứng từ cấu trúc của lý thuyết trò chơi có tổng bằng không, GAN bao gồm hai mô hình: một bộ tạo và một bộ phân biệt đối nghịch (bộ so sánh). Trình tạo sẽ cố gắng nắm bắt sự phân phối của các mẫu kiểm tra đúng để tạo ra bộ dữ liệu tương đồng mới. Bộ đối nghịch (bộ so sánh) thường là bộ phân loại nhị phân, phân biệt các mẫu được tạo ra từ các mẫu kiểm tra, để từ đó tìm ra mẫu đúng càng chính xác

với thực tế càng tốt. Quá trình tối ưu hóa GAN là một bài toán tối ưu hóa cực tiểu. Việc tối ưu hóa kết thúc tại điểm yên ngựa là điểm tối thiểu liên quan đến bộ tạo và tối đa đối với bộ phân biệt đối nghịch. Sau đó, bộ tạo có thể được coi là đã nắm bắt được sự phân phối chuẩn của các mẫu kiểm tra thực tế để có thể mô phỏng và tạo ra các đối tượng tương đồng. Qua đó, hai mô hình này có thể được triển khai với bất kỳ dạng mô hình chức năng khác biệt nào có sử dụng ánh xạ dữ liệu từ không gian này sang không gian khác, chúng thường được thực hiện bởi các mạng nơ-ron sâu (ví dụ: mạng nơ-ron tích chập, LSTM..).



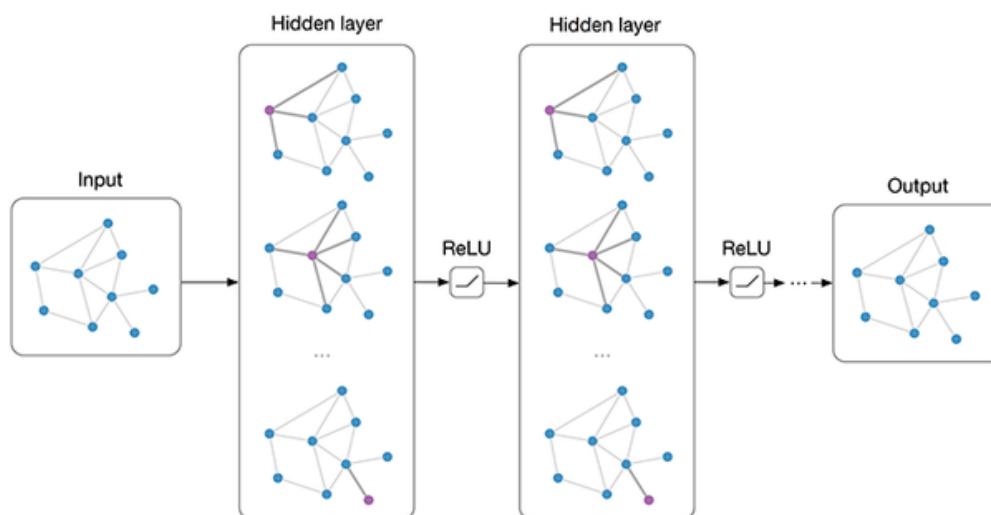
Hình 2.7: Mô hình kiến trúc mạng nơ-ron đối luyện - GAN

GAN là một trong những mô hình sinh dữ liệu thông minh tuyệt vời. Đã có rất nhiều những nghiên cứu dành cho mô hình này nhằm thu được GAN có giá trị tối ưu tốt nhất thông qua các phương pháp tối ưu hóa khác nhau, chẳng hạn như GAN có điều kiện, GAN tích chập sâu, GAN phân kỳ f và GAN phù hợp với chu kỳ. Những ưu điểm này cho phép GAN và các biến thể của chúng được áp dụng rộng rãi cho nhiều ứng dụng của viễn thám và phân tích ảnh vệ tinh, chẳng hạn như kết hợp dữ liệu (Jiang et al., 2019; Ma et al., 2020), xử lý độ che phủ đám mây (Sun et al., 2017).

2.5. Mạng nơ-ron đồ thị (Graph neural networks – GNN)

Sự thành công của học sâu (ví dụ: mạng nơ-ron tích chập và Mạng nơ-ron hồi quy) trong nhiều lĩnh vực một phần được cho là do biểu diễn các tính năng ẩn trong mô hình thông qua các tính toán dữ liệu Euclide. Tuy nhiên, cách biểu diễn như vậy sẽ làm mất đi tính hiệu quả trong dữ liệu đồ thị. Vì đồ thị có thể không đều, nên đồ thị có

thể có kích thước nút thay đổi và các nút từ đồ thị có thể có số lượng lân cận khác nhau, dẫn đến một số phép toán quan trọng (ví dụ: tích chập) để thực hiện trong không gian Euclide nhưng khó áp dụng cho miền đồ thị (Wu, Pan, & Chen và cộng sự, 2020b; Zhou và cộng sự, 2018).



Hình 2.8: Mô hình kiến trúc mạng nơ ron đồ thị (GNN)

Mạng nơ ron đồ thị được đề xuất để học tốt hơn các biểu diễn trên đồ thị thông qua cơ chế truyền và tổng hợp đặc trưng của dữ liệu. Nhiều nỗ lực đã được thực hiện trong việc truyền tải và tổng hợp các tính năng của dữ liệu ban đầu. GNN có thể được phân loại chung thành năm nhóm: mạng nơ ron đồ thị hồi quy (RGNN), mạng tích chập đồ thị (GCN), mạng đồ thị (GAT), mạng đồ thị tự lập mã (GAE) và mạng đồ thị không gian-thời gian (STGNN).

Do có tính thuyết phục và khả năng diễn giải cao ứng dụng cho đa dạng dữ liệu đầu vào khác nhau, GNN gần đây đã được áp dụng rộng rãi trong một loạt các lĩnh vực có tập dữ liệu lớn thông qua việc tạo biểu đồ ngữ cảnh, phân loại ngữ cảnh (Gao, Shi, & Li et al., 2021; Liang, Deng, & Zeng, 2020), phân đoạn / phân loại đám mây điểm, phân loại văn bản, dự báo lưu lượng và phát hiện các sự kiện thay đổi. Tuy nhiên, các ứng dụng của GNN trong lĩnh vực viễn thám còn khá là mới mẻ và chưa được áp dụng rộng rãi do những đặc tính ràng buộc và xây dựng ngữ cảnh cho các tập dữ liệu ban đầu.

CHƯƠNG 3:

TÌM HIỂU CÁC MÔ HÌNH HỌC MÁY TRONG NHẬN DIỆN TRƯỢT LỖ TỪ DỮ LIỆU ẢNH VỆ TINH

Như đã trình bày ở nội dung trước, các thuật toán Machine Learning đã được ứng dụng rộng rãi trong các lĩnh vực khoa học khác nhau, đặc biệt là trong khoa học trái đất để khắc phục các vấn đề phức tạp nhằm giảm thiểu chi phí và tăng khả năng phân tích của các đối tượng trong không gian. Học máy được coi là một lĩnh vực cụ thể chính của trí tuệ nhân tạo, nó bắt chước hoạt động của bộ não con người trong việc giải quyết vấn đề và ra quyết định. Khi làm như vậy, nó sử dụng nhiều thuật toán khác nhau, như mạng nơ-ron nhân tạo (ANN) trong quá trình học tập và huấn luyện. Dưới đây là những đánh giá về các mô hình học máy được ứng dụng trong việc nhận diện các đối tượng trượt lở thông qua ảnh vệ tinh quan sát trái đất.

3.1 Mô hình mạng nơ-ron đa lớp (MLP-NN)

Mạng nơ-ron đa lớp là một phương pháp phát triển từ mạng nơ-ron nhân tạo đã được sử dụng rộng rãi trong mô hình hóa các yếu tố địa không gian. Hiệu suất của phương pháp này bị ảnh hưởng bởi các yếu tố như cấu trúc của mô hình, các loại hàm khởi tạo và hàm tính toán được áp dụng và sử dụng các trọng số được cập nhật theo phương thức nào. Nói chung MLP-NN bao gồm một lớp đầu vào, một hoặc nhiều lớp ẩn và một lớp đầu ra. Việc mô hình hóa các khu vực có nguy cơ sạt lở đất cao, lớp đầu vào bao gồm các nơ-ron giống như các yếu tố ảnh hưởng đến sạt lở đất như lượng mưa, lớp phủ, vỏ phong hóa ... và số lượng lớp ẩn phụ thuộc vào dữ liệu huấn luyện và độ phức tạp của vấn đề. Trong nghiên cứu này, thuật toán lan truyền ngược (BPA) là phương pháp huấn luyện chính được sử dụng trong mạng nơ-ron, nó được sử dụng để cập nhật các trọng số. Để chạy phương pháp, các trọng số ban đầu được khởi tạo các giá trị ngẫu nhiên bởi BPA và được gán cho mỗi nơ-ron. Sau đó phương pháp được tối ưu hóa liên tục dựa trên quá trình lặp đi lặp lại để tìm tỷ lệ lỗi giữa giá trị đầu ra và giá trị mong đợi, cho đến khi tỷ lệ lỗi đạt được điều kiện ổn định (xem Công thức (1)):

$$y = (\sum_{i=1}^n w_i * x_i + b) \quad (1)$$

Trong đó W là vectơ trọng số, X là vectơ đầu vào của các đặc trưng của đối tượng và b là độ lệch. Ngoài ra, hàm tính toán sigmoid được sử dụng trong nghiên cứu này có thể được mô phỏng dưới dạng phương trình sau (2):

$$f(z) = \frac{1}{1+e^{-z}} \quad (2)$$

Trong đó f(z) là đầu ra của hàm kích hoạt và nó có giá trị trong khoảng từ 0 đến 1.

3.2. Mô hình Hồi quy logistic (LR)

Hồi quy logistic là một trong những phương pháp phân tích đa biến cho phép chúng ta tạo ra một kết nối hồi quy đa biến giữa một tập hợp các biến độc lập và một biến phụ thuộc. LR là một phương pháp mạnh mẽ để dự đoán sự hiện diện của một sự kiện bằng cách điều chỉnh mô hình tuyến tính tốt nhất dựa trên các giá trị biến độc lập và nó là một phương pháp thống kê thường được sử dụng áp dụng trong phân tích nhận biết và dự báo trượt lở đất. Một trong những ưu điểm đáng kể của LR là trong phương pháp này, các biến độc lập có thể vừa rời rạc vừa liên tục, hoặc có thể sử dụng hỗn hợp các loại biến này. Trong trường hợp này, biến phụ thuộc được giới thiệu dưới dạng giá trị nhị phân của 0 và 1, theo đó giá trị 0 cho thấy sự vắng mặt của sự kiện trượt lở đất và giá trị 1 cho thấy sự hiện diện của một vụ trượt lở. Định nghĩa toán học của LR được xác định bởi công thức (3):

$$p = \frac{1}{(1+e^{-z})} \quad (3)$$

Trong đó, p là xác suất xảy ra trượt lở đất, giá trị thay đổi trong khoảng từ 0 đến 1. Z là tổ hợp tuyến tính với các biến độc lập (các trọng số điều kiện gây trượt lở) được thể hiện trong công thức (4).

$$z = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n \quad (4)$$

khí b_0, b_i và x_i tương ứng là hệ số giới hạn mô hình, hệ số thống kê và các biến độc lập.

3.3 Mô hình Rừng Ngẫu nhiên (RF)

Rừng ngẫu nhiên là một phương pháp học máy được giám sát mạnh mẽ, được đề xuất và đã được sử dụng rộng rãi trong các ứng dụng viễn thám và GIS, chẳng hạn như phân loại xử lý ảnh vệ tinh và lập bản đồ tính hiện trạng và dự báo trượt lở đất. Phương pháp mô hình này dựa trên cơ sở cây quyết định và hoạt động bằng cách xây dựng vô số cây quyết định trong quá trình đào tạo, điều này làm cho nó ít bị ảnh hưởng hơn với các vấn đề quá mức. Trong phương pháp RF, mỗi cây quyết định tạo ra kết quả đầu ra và sau đó xuất ra các trọng số, được lấy từ các phiếu bầu, là tận tụy. Ưu điểm của RF là dễ áp dụng vì nó yêu cầu ít tham số hơn và nó mang lại độ chính xác cao hơn so với các phương pháp ML khác do quá trình đóng gói các dữ liệu huấn luyện. Ngoài ra, nó có thể xử lý các cấu trúc dữ liệu nhiều chiều và phức tạp. Do tính đơn giản và mang lại hiệu quả cao trong quá trình thực hiện, nên phương thức này được đa số các nhà khoa học sử dụng trong việc xây dựng các bản đồ hiện trạng trượt lở và cảnh báo nguy cơ trượt lở.

3.4 Phương pháp mô hình xếp chồng

Ý tưởng chính xuất phát từ việc mong muốn tổng hợp các phương pháp trong Machine Learning nhằm cải thiện hiệu suất của kết quả cuối cùng. Bằng cách kết hợp nhiều phương pháp khác nhau để xây dựng một cơ chế học mạnh mẽ phục vụ quá trình dự đoán hoặc phân loại một tập dữ liệu chức năng. Điểm mạnh của việc sử dụng các phương pháp tổng hợp là chúng ta có thể giảm phương sai bằng cách kết hợp một số phương pháp đơn lẻ, riêng biệt lại với nhau để mang lại hiệu suất một cách tốt nhất. Có ba phương pháp học máy tổng hợp chính: đóng gói, tăng cường và xếp chồng. Trong chuyên đề này chúng tôi đánh giá đại diện thông qua phương pháp học máy xếp chồng này.

Trong phương thức học máy xếp chồng, có hai cấp độ, cụ thể là cấp độ 0 và cấp độ 1. Ở cấp độ 0, các phương pháp mô hình đơn lẻ tạo ra một tập hợp các dự đoán dựa trên dữ liệu đào tạo và các kết quả đầu ra này (dự đoán) sau đó được sử dụng làm đầu vào cho một cơ chế học meta. Ở cấp độ 1, cơ chế học máy Meta là một phương pháp duy nhất nhằm tổng hợp các kết quả qua quá trình thực hiện một loạt các hoạt động huấn luyện nhằm đưa ra các dự đoán mới. Do đó, trong trường hợp này các phương

pháp học máy như LR, MLP-NN và RF được đào tạo như các phương pháp đơn lẻ ở cấp độ 0 và mô hình LR được tái sử dụng sau đó được như một siêu học viên của ở cấp độ 1 để đưa ra các dự đoán cuối cùng.

3.5 Mô hình lai tích hợp giữa mạng nơ ron đa lớp và OBIA

Trong phát hiện trượt lở đất thông thường và lập bản đồ dự báo các khu vực có nguy cơ trượt lở bằng cách sử dụng các mô hình học máy các kết quả nghiên cứu thường được thực hiện trên tập dữ liệu raster n chiều. Cụ thể là các phép tính toán diễn ra ở cấp pixel và các giá trị pixel trong mỗi lớp được thay đổi sau mỗi lần thực hiện. Việc áp dụng các thuật toán học máy cho các bộ dữ liệu lớn hoặc có độ phân giải rất cao ở mức pixel đòi hỏi hệ thống tính toán mạnh mẽ và tốn nhiều thời gian, đặc biệt, khi phạm vi không gian của khu vực nghiên cứu là rất đáng kể. Trong khi đó, trong OBIA các phân tích được thực hiện ở cấp đối tượng và các quy trình phân loại và phân đoạn ảnh tương đối nhanh so với các phương pháp dựa trên pixel. Thay vì xử lý từng pixel trong một tập hợp có cùng chung giá trị, thì OBIA xử lý trên một tập các pixel có giá trị trung bình của đối tượng kéo theo một tập hợp các pixel tương tự sẽ được phân tích. Ngoài các giá trị trung bình của các đối tượng trong mỗi lớp (ví dụ: dải hình ảnh, DEM, NDVI, v.v.), các đặc điểm thống kê và hình học khác, chẳng hạn như SD, chế độ, chỉ số hình dạng, vị trí và kết cấu, có thể được tính toán và sử dụng trong phân loại và trích xuất tính năng của đối tượng. Do đó, trong cách tiếp cận tích hợp này bước đầu tiên là tạo ra các đối tượng hình ảnh thông qua quá trình phân đoạn đa phân giải bằng cách sử dụng các hệ số tỷ lệ khác nhau và sau đó tính toán các đặc tính thống kê và hình học cho từng đối tượng.

KẾT LUẬN

Trong nội dung báo cáo “Viễn thám và dữ liệu lớn, nghiên cứu và vận dụng dữ liệu lớn từ ảnh viễn thám” tác giả đã nêu lên khái niệm và dữ liệu lớn từ viễn thám và các vấn đề ứng dụng các mô hình học máy trong nghiên cứu dữ liệu viễn thám, trong đó chúng tôi có nêu về trượt lở và nhận diện các đối tượng trượt lở đất.. Việc xây dựng cơ sở hiện trạng ban đầu của các khu vực trượt lở rất quan trọng và nó được áp dụng phương pháp phân tích và xử lý ảnh vệ tinh quan sát trái đất thông qua các mô hình học máy nhằm làm gia tăng tốc độ tính toán và giảm thiểu các chi phí thống kê theo phương pháp truyền thống khác. Việc ứng dụng các mô hình học máy thông minh làm cho kết quả trở nên nhanh chóng và chính xác hơn. Tuy nhiên, việc sử dụng các mô hình phụ thuộc chủ yếu vào các yếu tố của dữ liệu huấn luyện ban đầu và dữ liệu cơ sở ảnh vệ tinh. Độ phân dải và các dải tần quang phổ của ảnh vệ tinh càng cao thì tính chính xác của mô hình càng tốt, dữ liệu trích xuất các đối tượng trượt lở đất càng gần với thực tế, cũng như dữ liệu đã được kiểm nghiệm ngoài thực tế càng đầy đủ chính xác thì kiểm nghiệm mô hình càng trở nên hiệu quả hơn. Việc kết hợp các nguồn dữ liệu làm cho nền dữ liệu viễn thám trở nên đa dạng, phức tạp và khối lượng dữ liệu cực kỳ lớn, đòi hỏi các nghiên cứu ứng dụng dữ liệu này cần có tính khoa học và kết hợp đa thuật toán, đa lĩnh vực để tận dụng và khai thác triệt để dữ liệu, phục vụ cho phát triển và nghiên cứu khoa học khác.

ThS. Trương Xuân Bình

DANH MỤC TÀI LIỆU THAM KHẢO

- [1] Xin Zhang, Ya'nan Zhou & Jiancheng Luo of Deep learning for processing and analysis of remote sensing big data, 2021
- [2] Gong Cheng, Lei Guo, Tianyun Zhao, Junwei Han, Huihui Li & Jun Fang of Automatic landslide detection from remote-sensing imagery using a scene classification method based on BoVW and pLSA, 2012.
- [3] Viet-Ha Nhu, Nhat-Duc Hoang, Mahdis Amiri, Tinh Thanh Bui, Phuong Thao T. Ngo, Pham Viet Hoa, Pijush Samui, Long Nguyen Thanh, Tu Pham Quang & Dieu Tien Bui of An approach based on socio-politically optimized neural computing network for predicting shallow landslide susceptibility at tropical areas, 2021
- [4] Abolfazl Abdollahi 1, Biswajeet Pradhan 1,2,* , Nagesh Shukla 1, Subrata Chakraborty 1 and Abdullah Alamri 3 Deep Learning Approaches Applied to Remote Sensing Datasets for Road Extraction, 5-2020
- [5] Ghorbanzadeh, O.; Blaschke, T.; Gholamnia, K.; Meena, S.R.; Tiede, D.; Aryal, J. Evaluation of Different Machine Learning Methods and Deep-Learning Convolutional Neural Networks for Landslide Detection
- [6] Ghorbanzadeh, O.; Blaschke, T.; Gholamnia, K.; Meena, S.R.; Tiede, D.; Aryal, J. Evaluation of Different Machine Learning Methods and Deep-Learning Convolutional Neural Networks for Landslide Detection.
- [7] Hou, X., Wang, B., Hu, W., Yin, L., & Wu, H. (2019). *SolarNet: A deep learning framework to map solar power plants in China from satellite imagery*.
- [8] Pourghasemi, H.; Gayen, A.; Park, S.; Lee, C.-W.; Lee, S. Assessment of landslide-prone areas and their zonation using logistic regression, logitboost, and naïve bayes machine-learning algorithms, 2018.
- [9] Goetz, J.; Brenning, A.; Petschko, H.; Leopold, P. Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. *Comput. Geosci.* 2015
- [10] Meena, S.R.; Tavakkoli Piralilou, S. Comparison of Earthquake-Triggered Landslide Inventories: A Case Study of the 2015 Gorkha Earthquake, Nepal. *Geosciences* 2019.
- [11] Mezaal, M.; Pradhan, B.; Rizeei, H. Improving Landslide Detection from Airborne Laser Scanning Data Using Optimized Dempster–Shafer, 2018.
- [12] Aryal, J.; Josselin, D. Environmental Object Recognition in a Natural Image: An Experimental Approach Using Geographic Object-Based Image Analysis (GEOBIA). 2014.
- [13] Rajbhandari, S.; Aryal, J.; Osborn, J.; Lucieer, A.; Musk, R. Leveraging Machine Learning to Extend Ontology-Driven Geographic Object-Based Image Analysis (O-GEOBIA): A Case Study in Forest-Type Mapping. 2019.
- [14] Martha, T.R.; Kerle, N.; van Westen, C.J.; Jetten, V.; Kumar, K.V. Segment optimization and data-driven thresholding for knowledge-based landslide detection by object-based image analysis, 2011