

Bài báo khoa học

Nghiên cứu ứng dụng mô hình học máy trong phân loại hệ sinh thái đất ngập nước nội địa tại Vườn Quốc gia Ba Bể, tỉnh Bắc Kạn

Đặng Kinh Bắc^{1*}, Phạm Hạnh Nguyên², Nguyễn Thành Vĩnh², Trần Ngọc Cường², Nguyễn Văn Quân³, Trịnh Thị Thúy Chinh⁴, Vũ Thị Phương⁵

¹ Khoa Địa lý, trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội; dangkinhbac@hus.edu.vn;

² Cục Bảo tồn Thiên nhiên và Đa dạng Sinh học, Tổng cục Môi trường, Bộ Tài nguyên và Môi trường; nguyenvh.nbca@gmail.com;

³ Công ty TNHH MTV Trắc địa Bản đồ, Cục Bản Đồ, Bộ Tổng tham mưu; quannguyen26795@gmail.com;

⁴ Công ty cổ phần dược phẩm Hoa Anh, 14/1 Pháo Đài Láng, Đống Đa, Hà Nội; trinha2698@gmail.com;

⁵ Đại học Hồng Đức, 565 Quang Trung, Thanh Hóa; vuthiphuong@hdu.edu.vn;

*Tác giả liên hệ: dangkinhbac@hus.edu.vn; Tel.: +84–389542557

Ban Biên tập nhận bài: 12/6/2022; Ngày phản biện xong: 20/7/2022; Ngày đăng bài: 25/7/2022

Tóm tắt: Sự suy thoái và chuyển đổi đất ngập nước đã được cảnh báo trong 10 năm qua, nhưng việc kiểm kê và giám sát những thay đổi này vẫn còn gặp nhiều khó khăn do thiếu khả năng tiếp cận và công nghệ. Do đó, việc trang bị cho các nhà quản lý những công cụ để phân loại và giám sát các hệ sinh thái đất ngập nước theo thời gian thực là điều cần thiết. Mục đích của nghiên cứu này là phát triển mô hình hữu ích phân loại 7 loại đất ngập nước nội địa tại Vườn quốc gia Ba Bể, tỉnh Bắc Kạn cùng các hệ sinh thái lân cận được phân loại trong hệ thống RAMSAR và MONRE (*Ministry of Natural Resources and Environment – Bộ Tài Nguyên và Môi trường*). Từ đó, mô hình ResU–Net (*Deep Residual U–Net*) sử dụng chức năng tối ưu hóa Adadelta đã được sử dụng để phân loại 7 loại đất ngập nước và 3 loại hình hệ sinh thái lân cận ở Vườn quốc gia Ba Bể, dựa trên ảnh vệ tinh Sentinel–2, với độ chính xác cao hơn 85%. Hiệu suất tốt hơn của mô hình này so với các phương pháp *Random forest (RF)* và *Support Vector Machine (SVM)* đã được chứng minh. Mô hình ResU–Net sau khi tối ưu hóa, cũng được sử dụng để lập bản đồ các khu vực đất ngập nước nội địa tại vùng núi Bắc Kạn. Mô hình này có khả năng cập nhật các kiểu đất ngập nước mới tại Việt Nam nhằm giám sát biến động đất ngập nước trong thời gian thực.

Từ khóa: Đất ngập nước; RAMSAR; Vườn quốc gia Ba Bể; U–Net; Viễn thám.

1. Đặt vấn đề

Hiện nay, khoảng 70% dân số thế giới đang sống ở các cửa sông ven biển và xung quanh các vùng nước ngọt nội địa [1–2]. Hệ sinh thái này cung cấp cho nhân loại một số lượng lớn các sản phẩm có giá trị hàng năm. Tuy nhiên các vùng đất ngập nước đang dần bị suy giảm, thế chỗ cho các công trình nhân sinh. Tại Việt Nam, các khu vực đất ngập nước rất đa dạng, có diện tích khoảng 5.810.000 ha, chiếm khoảng 8% tổng diện tích đất ngập nước của châu Á [3]. Hiện nay, cũng như trên thế giới, các hệ sinh thái đất ngập nước này tại nước ta đang đối mặt với sự suy giảm cả chất lượng lẫn diện tích. Mặc dù sự suy giảm này đã được cảnh

báo trong nhiều năm qua, việc đánh giá, kiểm kê và giám sát những thay đổi các vùng đất ngập nước vẫn đang gặp nhiều khó khăn do thiếu khả năng tiếp cận và công nghệ [4]. Do đó, cần thiết trang bị cho các nhà quản lý các công cụ tốt hơn để phân loại và giám sát thường xuyên hệ sinh thái đất ngập nước, đặc biệt tại các khu vực Vườn quốc gia và khu bảo tồn. Vườn quốc gia (VQG) Ba Bể thuộc tỉnh Bắc Kạn với tổng diện tích 7.610 ha, bao gồm khoảng 3.226 ha vùng lõi. Đặc biệt, vào năm 2011 vùng lõi của VQG Ba Bể được Ban Thư ký Công ước RAMSAR công nhận là khu đất ngập nước thứ 1938 có tầm quan trọng trên thế giới. Tuy nhiên, khác với hầu hết các VQG trên cả nước, tại vùng lõi của VQG Ba Bể các hoạt động kinh tế của người dân vẫn diễn ra. Do vậy, việc giám sát đất ngập nước có vai trò quan trọng đối với sinh kế và phát triển kinh tế của người dân tại vùng lõi của VQG Ba Bể.

Gần đây, các mô hình trí tuệ nhân tạo ngày càng phát triển và có chiều sâu, các mạng này có thể cung cấp một dự đoán chính xác hơn so với các phương pháp tính toán viễn thám trước đây. Phương pháp phân đoạn ngữ nghĩa hiện đang bắt đầu được sử dụng trong các mô hình học máy, đặc biệt là trong việc tách các vùng phủ và đất ngập nước cho mỗi đặc điểm lớp phủ trên ảnh viễn thám thành một lớp riêng biệt. Một số mạng sử dụng các phương pháp phân đoạn ngữ nghĩa chẳng hạn như mạng nơ-ron liên kết (CNN) mã hóa thông tin ngữ nghĩa cấp cao với quy trình tách mẫu. Tuy nhiên, phương pháp này không bảo toàn thông tin không gian chi tiết nhất của hình ảnh đầu vào. Do đó, một số mô hình như PSPNet, DeepLab v2 và Deeplab v3 đã được tích hợp với nhiều lớp phức hợp hơn để lưu giữ thông tin không gian trong suốt quá trình. Đặc biệt, mô hình mạng U-Shape có thể khôi phục thông tin không gian của dữ liệu đầu vào sau quá trình mã hóa và giải mã [16]. Hai giai đoạn này được kết nối liên tục và đa chiều, giúp lưu giữ thông tin không gian và thuộc tính đến cùng. Đồng thời để quan sát các loại đất ngập nước trong một khu vực rộng lớn, các ảnh vệ tinh như MODIS, Landsat và Sentinel thường được sử dụng [5–6]. So với các hình ảnh vệ tinh MODIS và Landsat có độ phân giải không gian kém thì Sentinel với khả năng chụp ảnh đa phổ, có thể thu được hình ảnh quang học một cách có hệ thống trên cả khu vực nội địa và ven biển ở độ phân giải không gian cao (10 m đến 60 m). Do đó, trong nghiên cứu này, các tác giả đề xuất mô hình ResU-Net để dự báo lớp phủ đất ngập nước nội địa dựa trên dữ liệu Sentinel đa thời gian tại Vườn Quốc gia Ba Bể, tỉnh Bắc Kạn. Ba câu hỏi nghiên cứu sau liên quan đến phân loại lớp phủ đất ngập nước dựa trên học sâu sẽ được làm rõ trong nghiên cứu này:

- Lợi ích của việc tích hợp ảnh học sâu và ảnh viễn thám đa thời gian để theo dõi phân loại đất ngập nước nội địa là gì?
- Liệu mô hình học máy sâu có thể thay thế các phương pháp truyền thống trong việc phân loại đất ngập nước nội địa?
- Các loại đất ngập nước nội địa phân bố như thế nào ở khu vực VQG Ba Bể?

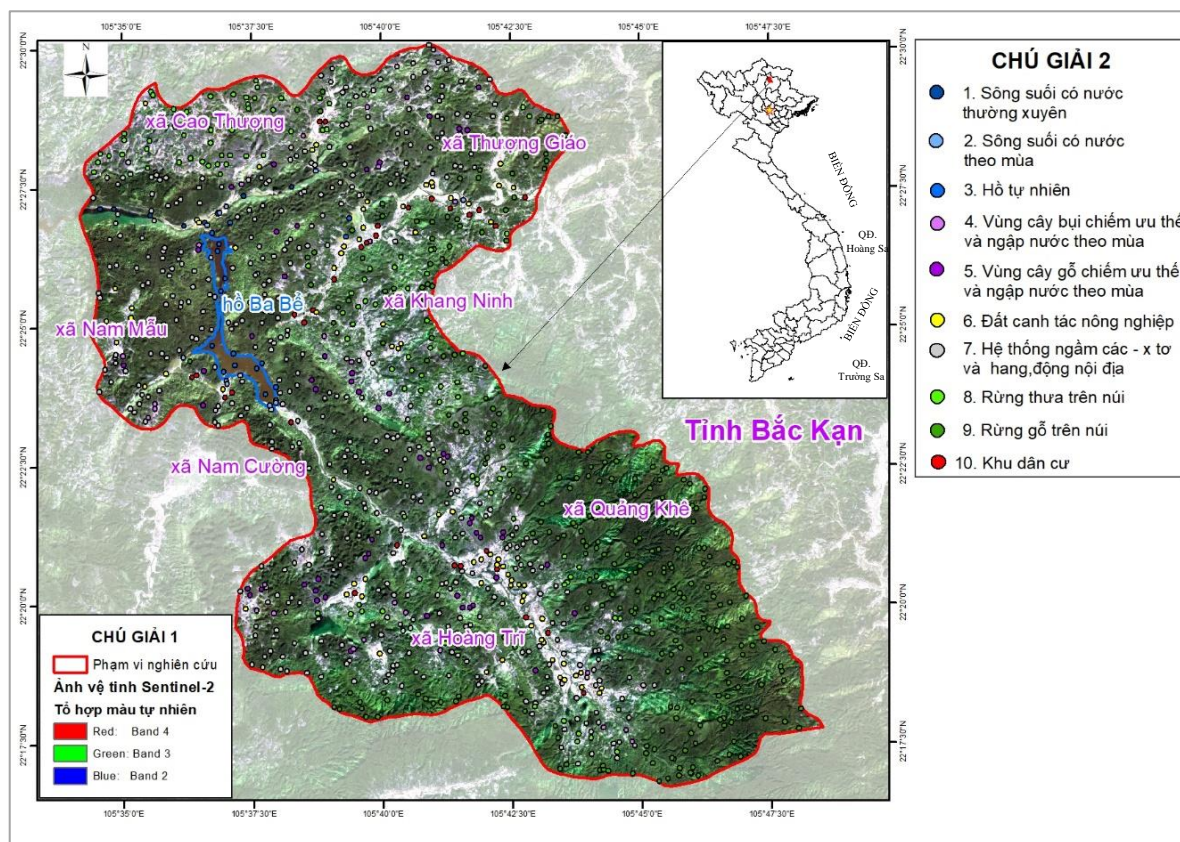
2. Phạm vi và phương pháp nghiên cứu

2.1. Khu vực nghiên cứu

Khu vực trọng tâm trong nghiên cứu này là vùng đất ngập nước vườn Quốc gia Ba Bể thuộc các xã Nam Mẫu, Khang Ninh, Cao Thượng, Thượng Giáo, Quảng Khê, Hoàng Trĩ, Đông Phúc thuộc huyện Ba Bể và xã Nam Cường thuộc huyện Chợ Đồn. Vườn Quốc gia Ba Bể có tổng diện tích 44 ha, trong đó vùng lõi chiếm khoảng 10 ha và vùng đệm là hơn 34 ha [7]. Ngày 2 tháng 2 năm 2011, Ban Thư ký Công ước RAMSAR đã công nhận vùng lõi của VQG Ba Bể là khu đất ngập nước thứ 1938 có tầm quan trọng trên thế giới [8].

Hồ Ba Bể là hệ sinh thái lớn nhất thuộc VQG Ba Bể được hình thành cách đây khoảng 200 triệu năm về trước do một biến động địa chất lớn, làm sụt lún các dãy núi đá vôi tạo thành. Hồ được nhận nước từ ba con sông Tà Han, Nam Cường và Chợ Lèn ở phía nam VQG rồi từ hồ đổ ra sông Năng là con sông chảy qua phía bắc của VQG. Ba Bể là một khu bảo vệ độc đáo trong hệ thống rừng đặc dụng Việt Nam do là nơi có nhiều hệ sinh cảnh nước ngọt đa dạng, nhiều ao tù, vùng đầm lầy. Cấu trúc địa chất, địa mạo khu vực hết sức đa dạng với nhiều đỉnh cao, độ phân

cắt lớn, thung lũng, sông suối với các sườn dốc đứng kéo theo hình thành nhiều hệ sinh thái khác nhau và đồng thời cũng tạo thành phong cảnh nơi đây hết sức ngoạn mục. Theo phân loại của RAMSAR và Bộ Tài nguyên và Môi trường [9], VQG Ba Bể có 7 loại hình đất ngập nước được thống kê trong Bảng 1.



Hình 1. Phạm vi khu vực nghiên cứu (Ảnh Sentinel-2).

Bảng 1. Phân loại đất ngập nước VQG Ba Bể dựa trên RAMSAR, MONRE.

STT	HST	Loại hình đất ngập nước	RAMSAR	MONRE	VQG Ba Bể
1		Sông, suối có nước thường xuyên	X	X	X
2		Sông, suối có nước theo mùa	X	X	X
3		Hồ tự nhiên	X	X	X
4	Đất ngập nước tự nhiên	Vùng đất than bùn có rừng, cây bụi hoặc không có thực vật che phủ	X	X	
5		Vùng ngập nước có cây bụi chiếm ưu thế và ngập nước theo mùa	X	X	X
6		Vùng ngập nước có cây gỗ chiếm ưu thế và ngập nước theo mùa	X	X	X
7		Suối, điểm nước nóng, nước khoáng	X	X	
8		Hệ thống thủy văn ngầm karst và hang, động nội địa	X	X	X
9		Ao, hồ, đầm nuôi trồng thủy sản nước mặn, lợ	X	X	
10		Ao, hồ, đầm nuôi trồng thủy sản nước ngọt	X	X	
11	Đất ngập nước nhân sinh	Đất canh tác nông nghiệp	X	X	X
12		Đồng cỏ	X	X	
13		Đồng muối	X	X	
14		Ao, hồ chứa và xử lý nước thải	X	X	
15		Moong khai thác khoáng sản	X	X	
16		Hồ chứa nước nhân tạo	X	X	
17		Sông đào, kênh, mương, rạch	X	X	

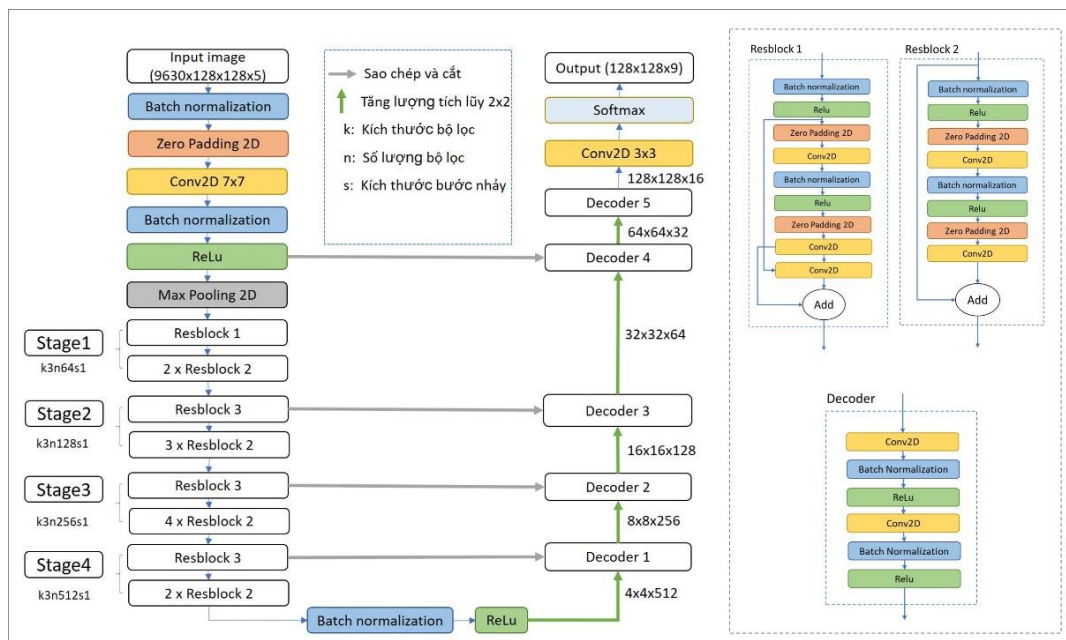
2.2. Thu thập dữ liệu và mẫu

2.2.1. Chuẩn bị dữ liệu đầu vào

Để thu thập thông tin về lớp phủ đất ở khu vực Vườn Quốc gia Ba Bể, ảnh Sentinel-2 (hiệu chỉnh ở mức 2 với độ phân giải 10 mét) đã được chọn sử dụng. Dựa vào các ảnh vệ tinh có độ phân giải trung bình, nghiên cứu phân biệt được các vùng đất ngập nước vĩnh viễn và tạm thời. Đối với vùng có hệ thủy văn ngầm karst và hang động, nghiên cứu bổ sung dữ liệu địa chất-địa mạo để xác định khu vực có thành tạo karst, ghép với các hệ sinh thái trên bề mặt để chuẩn hóa dữ liệu đầu ra từ bước giải đoán. Các hình ảnh vệ tinh thu được năm 2021 và 2020 được sử dụng để xác minh các lớp đầu vào để đào tạo mô hình ResU-net. Khảo sát thực địa được thực hiện vào tháng 11/2021 để kiểm chứng các loại hình đất ngập nước đã được giải đoán tại phòng. Ngoài ra, công trình nghiên cứu sử dụng thêm 5 hình ảnh Sentinel-1 và 2 vào các năm 2014, 2016, 2017, 2018 để đánh giá những thay đổi đất ngập nước trong Vườn Quốc gia Ba Bể sau khi mô hình ResU-net được hoàn thành. Trên cơ sở đó, nghiên cứu tính toán chỉ số MNDWI nhằm đánh giá sự biến đổi chất lượng môi trường nước trên các hệ sinh thái có được từ hệ thống ResU-Net, đặc biệt tại khu vực hồ Ba Bể.

2.2.2. Thiết lập ResU-Net để phân loại đất ngập nước

ResU-Net là một kiến trúc tận dụng các mạng thần kinh còn sót lại sâu với 34 lớp [1,10] và U-Net [11]. Kiến trúc của ResU-Net được đề xuất thể hiện trong Hình 2. Các mạng ResU-Net tích hợp các khối xây dựng còn lại (viết tắt là ResBlock) trong một mặt mã hóa của các mô hình U-Net, trong khi mặt giải mã của chúng vẫn được giới thiệu trong kiến trúc U-Net cũ [12]. Các mạng Resnet34 đã được sử dụng trong phân loại đối tượng, nhận dạng hình ảnh và các tác vụ phi thị giác máy tính. Dựa trên những lợi thế này, kiến trúc ResU-Net được chọn làm xương sống của mạng trong nghiên cứu này.



Hình 2. Cấu trúc ResU-Net để đào tạo mô hình phân loại các loại hệ sinh thái đất ngập nước nội địa.

Trong quá trình phát triển ResU-Net, độ chính xác của cả dữ liệu đào tạo và xác nhận đã được kiểm tra để tránh hiện tượng mô hình quá khít hoặc không khớp với mẫu đầu vào. Mạng ResU-Net tốt nhất sẽ được chọn nếu khả năng dự đoán các loại hình đất ngập nước phù hợp với các nhãn được chỉ định từ dữ liệu huấn luyện trong dữ liệu thô. Mô hình ResU-Net được phát triển dựa trên API python mô hình Segmentation trong khuôn khổ Keras, dưới dạng API được

thiết kế để phân đoạn hình ảnh dựa trên Tensorflow. Chu kỳ đào tạo ResU-Net được giới hạn ở 200 vòng (chu kỳ), nhưng nếu hệ số trên tập dữ liệu thử nghiệm hội tụ, chu kỳ có thể bị dừng lại nếu tất cả các giá trị chính xác không thay đổi sau 20 chu kỳ.

2.2.3. Phương pháp kiểm tra hiệu suất

Chức năng đánh giá mức độ mất mát dữ liệu trong quá trình huấn luyện được tính toán liên tục để đánh giá hiệu suất hoạt động của mô hình. Nếu như mô hình có độ chính xác cao nhưng mất đi thông tin dữ liệu ban đầu thì mô hình đó cũng không có ý nghĩa [13–14]. Do số lượng mẫu cho bảy đối tượng đất ngập nước nội địa không cân bằng trong bộ dữ liệu đầu vào, chỉ số đánh giá mức độ mất mát đã được chọn trong nghiên cứu này là F1 để đánh giá mức độ mất thông tin phân loại đa lớp. Nó làm giảm sự mất cân bằng của các bộ dữ liệu đào tạo giữa các đối tượng, đặc biệt khi diện tích đất ngập nước nội địa chiếm phần lớn dữ liệu đầu vào [14–15]. Ngoài ra, trong nghiên cứu này, hai giá trị độ chính xác khác sẽ được tính toán, bao gồm tổng độ chính xác ACC và IoU như các công thức sau:

$$ACC = \frac{2TP}{2TP+FP+FN} \quad (1)$$

$$IoU = \frac{TP}{TP+FP+FN} \quad (2)$$

Trong đó TP là giá trị dương thực sự, FP là giá trị dương tính giả và giá trị âm sai FN giữa dự đoán và sự thật cơ bản. Mô hình được đào tạo có giá trị thấp nhất trong tất cả các chức năng mất mát sẽ là mô hình tốt nhất để phân loại các khu vực đất ngập nước mới.

2.2.4. Phương pháp tối ưu hóa

Các phương pháp tối ưu hóa được sử dụng rộng rãi để xây dựng các mạng thần kinh dựa trên thuật toán SGD để giảm chức năng chi phí. Cách tiếp cận này để thay đổi trọng số mô hình theo hướng gradient âm cải thiện độ chính xác của các mạng thần kinh đủ điều kiện và giảm thiểu sự mất mát. Các lỗi của các mô hình được đào tạo (hoặc chức năng mất mát) đã được tính toán trong các chu kỳ tối ưu hóa. Qua các vòng lặp tính toán, dữ liệu di chuyển qua các lớp của các mô hình ResU-Net [1, 15]; và cập nhật trọng số sau mỗi vòng lặp là cần thiết để giảm giá trị mất mát cho đánh giá tiếp theo. Ba thuật toán tối ưu hóa đã được sửa đổi tuần tự trong nghiên cứu này bao gồm Adam (Ước tính khoảng khắc thích ứng), Adagrad (Thuật toán Gradient thích ứng), Adadelata trong quá trình phát triển ResU-Net. Nói chung, cách tiếp cận tối ưu hóa các trọng số trong mô hình tạo ra độ chính xác cao nhất và giá trị mất mát thấp nhất.

Bảng 2. Ba thuật toán tối ưu hóa để đào tạo các thông số của kiến trúc ResU-Net để phân loại đất ngập nước.

Công thức	Phương pháp Optimizer	Thuật toán
03	Adam	$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t + \epsilon}} \hat{m}_t$
04	Adagrad	$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{G_t + \epsilon}} g_t$
05	Adadelata	$\theta_{t+1} = \theta_t - \frac{\sqrt{E[\Delta\theta^2]_{t-1} + \epsilon}}{\sqrt{E[g^2]_t + \epsilon}} \cdot g_t$

Trong đó θ : giá trị tham số ở đâu; η : là tỷ lệ học tập; t là bước thời gian; $\epsilon = 10^{-8}$ g_t : là gradient; $E[g]$: trung bình động của gradient bình phương; m , v là ước tính của khoảng khắc thứ nhất và thứ hai; g_t : hoạt động tối đa.

Trong đó, Adam là thuật toán cho phép tính tốc độ học thích ứng với mỗi trọng số. Thuật toán Adagrad thay đổi tùy thuộc vào trọng số: tốc độ học thấp đối với các trọng số tương ứng với

các đặc trưng phổ biến và tốc độ học cao tương ứng với các trọng số đặc trưng ít phổ biến hơn. Còn với Adadelta là một biến thể khác của Adagrad nhằm mục đích khắc phục tình trạng giảm tốc độ học của Adagrad.

2.2.5. So sánh mô hình

Trong phần này, kết quả dự đoán của ba mô hình ResU–Net được so sánh với kết quả từ hai mô hình chuẩn, bao gồm RF và SVM. Tổng cộng có 9630 điểm ngẫu nhiên đã được chọn ở khu vực trong quá trình thực địa lấy mẫu. Các loại đất ngập nước được giải thích từ năm mô hình và mặt nạ được chỉ định cho 9630 điểm này. Kết quả giải thích từ năm mô hình được so sánh với thông tin ban đầu từ mặt nạ để kiểm tra hiệu suất của từng mô hình được đào tạo. Hai giá trị đánh giá được chọn là độ chính xác tổng thể (ACC) và các giá trị hệ số kappa. Mô hình tốt nhất sẽ đạt được các giá trị ACC và kappa cao nhất.

3. Kết quả

3.1. Đánh giá độ chính xác mô hình học máy

Độ chính xác của việc phân loại kiểu hệ sinh thái ĐNN trong VQG Ba Bể dựa trên việc giải thích năm mô hình học máy được đào tạo thể hiện trong Bảng 2. Theo đó, hầu hết các hệ sinh thái đất liền đều được dự đoán hiệu quả trong cả ba mô hình U–Net. Các hệ thống thủy văn ngầm karst và hang, động nội địa được giải đoán có phần kém nhất do tính chất karst không được thể hiện đặc trưng trên bề mặt. Bên cạnh đó, các hệ sinh thái biến động theo mùa vụ cũng khó được phát hiện như sông suối theo mùa và đất canh tác nông nghiệp. Mô hình Unet–Adadelta–256–64 phát hiện hầu hết các sông suối thường xuyên và hồ tự nhiên. Các khu vực sông suối theo mùa không đồng nhất bằng mô hình Unet–Adam–256–32 và Unet–Adagrad–256–64, trong khi sự phân bố của kiểu hệ sinh thái này dường như đồng nhất hơn trong mô hình Unet–Adadelta–256–64. Trong kết quả thu được từ hai mô hình điểm chuẩn, mô hình RF không phát hiện được các hệ sinh thái ngập nước theo mùa. Kết quả của mô hình RF giải đoán tốt các khu vực có hồ tự nhiên và sông suối thường xuyên. Cả hai mô hình điểm chuẩn đều dự đoán rằng rất khó để tách các khu vực còn lại. Có thể thấy sự khác biệt giữa kết quả của tất cả các mô hình U–Net và hai mô hình điểm chuẩn là tương đối lớn.

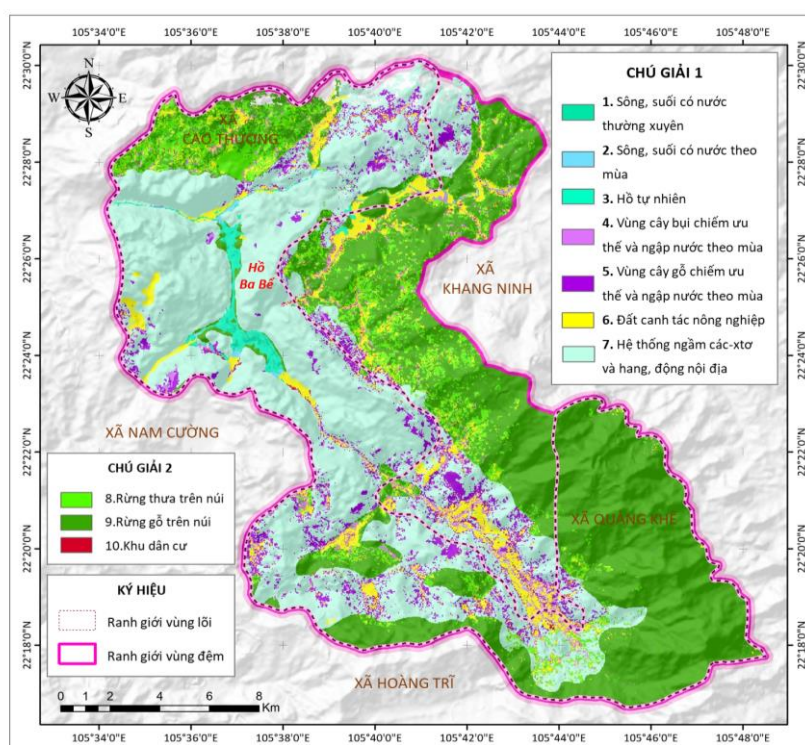
Bảng 3. Xác nhận chéo của ba mô hình U–Net được đào tạo và hai mô hình điểm chuẩn để phân loại kiểu hệ sinh thái ĐNN.

	Độ chính xác mô hình					
	Số mẫu	UNet–Adam–256–32	UNet–Adagrad–256–64	UNet–Adadelata–256–64	SVM	RF
	Kiểu HST đất ngập nước					
Sông, suối có nước thường xuyên	997	97,4	97,8	99,4	98,5	97,2
Sông, suối có nước theo mùa	962	74,8	76,0	79,5	77,5	26,8
Hồ tự nhiên	983	85,8	88,3	93,4	85,4	85,0
Vùng ngập nước có cây bụi chiếm ưu thế và ngập nước theo mùa	939	80,9	80,4	90,4	76,3	19,2
Vùng ngập nước có cây gỗ chiếm ưu thế và ngập nước theo mùa	990	97,4	97,2	98,6	97,8	97,2
Hệ thống thủy văn ngầm karst và hang, động nội địa	852	27,0	30,4	64,9	21,8	3,2
Đất canh tác nông nghiệp	992	77,6	85,0	98,7	59,1	48,1
Tổng số	9,630	Độ chính xác (%)				
		73,0	77,0	86,6	66,7	50,2
		Kappa Coefficient				
		0,7	0,8	0,9	0,6	0,5

So sánh độ chính xác giữa ba mô hình U-Net và hai mô hình điểm chuẩn với các dự đoán mới được trình bày trong Bảng 2. Cả ba mô hình U-Net có thể phát hiện bốn kiểu hệ sinh thái ĐNN trong VQG gồm sông suối có nước theo mùa, hồ tự nhiên, vùng ngập nước có cây bụi và cây gỗ ngập nước theo mùa và đất canh tác nông nghiệp. Mô hình UNet-Adadelta-256-64 là mô hình tốt nhất để phân loại hầu hết các kiểu hệ sinh thái này với tổng độ chính xác là 86,6 % và chỉ số Kappa là 0,9. Hai mô hình U-Net khác khả năng xác định các khu vực thủy văn karst ngầm và hang động nội địa với độ chính xác thấp. Trong hai mô hình điểm chuẩn, RF chỉ đạt được độ chính xác là 50 % với chỉ số Kappa là 0,5. Mặc dù SVM có thể giải đoán hồ và sông suối tự nhiên với độ chính xác cao hơn 80 %, nó không thể được sử dụng để giải đoán các hệ sinh thái khác. Do đó, có thể dễ dàng khẳng định rằng kết quả từ tất cả các mô hình U-Net có độ chính xác cao hơn so với kết quả từ hai mô hình truyền thống trước đây.

3.2. Phân bố vùng đất ngập nước trong VQG Ba Bể

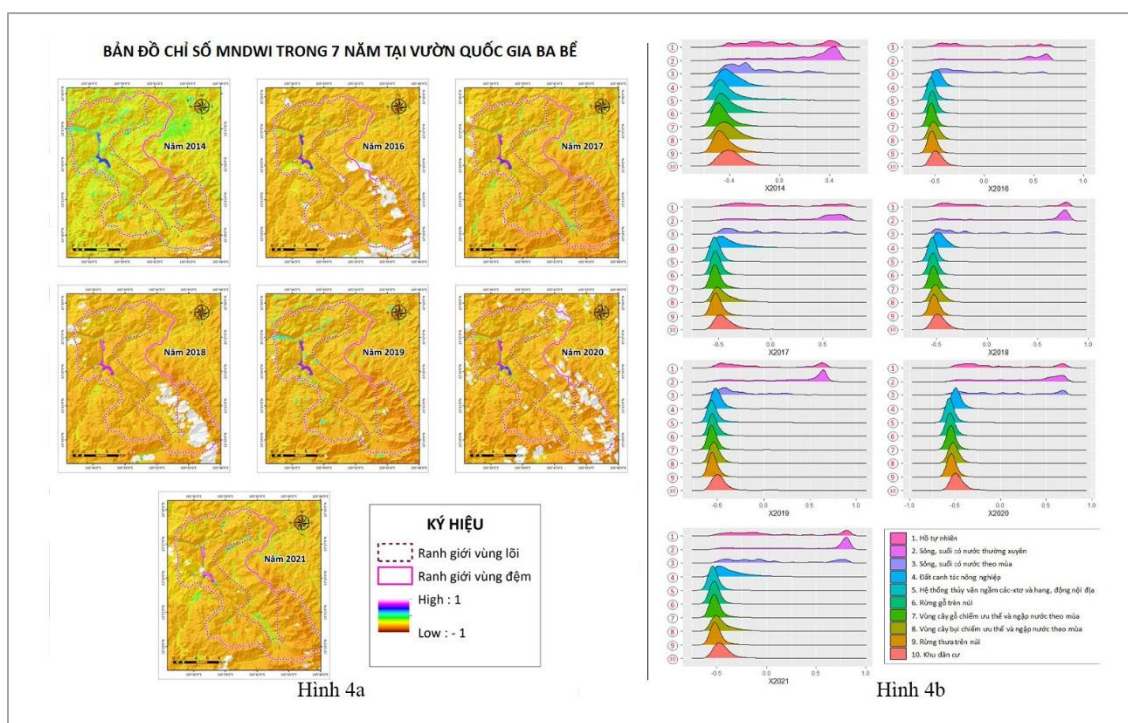
Khu vực VQG Ba Bể có 10 loại hình lớp phủ, trong đó có 7 loại hình đất ngập nước lục địa được thống kê theo danh sách phân loại của RAMSAR và MONRE (Hình 3). Dân cư sinh sống rải rác dọc theo đường quốc lộ và một số điểm du lịch quanh khu vực Hồ Ba Bể và trung tâm các xã vùng đệm. Ngoài các hệ sinh thái sông suối và hồ tự nhiên ngập nước thường xuyên trong năm, khu vực VQG có nhiều hệ sinh thái ĐNN ngập nước theo mùa như vùng cây bụi, cây gỗ và đất canh tác nông nghiệp. Ba lớp phủ rừng thưa và gỗ trên núi tập trung chủ yếu tại khu vực núi cao phía đông khu vực vùng đệm, xã Quảng Khê và Cao Thượng. Đặc biệt, khu vực vùng lõi chủ yếu được hình thành bởi hệ thống đá vôi hệ tầng Mia Lé. Bản thân hồ Ba Bể được hình thành trên một vùng sụt, nằm giữa vùng karst lớn. Tại đây phát hiện được nhiều hang động lớn nhỏ. Hệ thống hang động kì vĩ đã và đang phát triển dưới tác động của quá trình kiến tạo bao gồm hệ thống hang khô, hệ thống hang nước và hệ thống sông ngầm. Do đó, toàn bộ hệ thống lớp phủ rừng gỗ và cây bụi có nền địa chất karst đều được phân tách thành loại lớp phủ phù hợp với phân loại của RAMSAR và MONRE là “hệ thống ngầm các-tơ và hang, động nội địa”. Một số hang động được kể đến như động Puông, động Hua Mạ, hang Thẩm Phầy và hang Nà Phoong.



Hình 3. Bản đồ phân bố các kiểu đất ngập nước khu vực VQG Ba Bể.

3.3. Biến động chỉ số mặt nước khu vực nghiên cứu trong 7 năm

Kết quả tính toán chỉ số mặt nước khu vực VQG Ba Bể cho thấy chỉ số MNDWI (Modified Normalized Difference Water Index) thấp hơn tại các năm ảnh được thu vào mùa khô như 2014 và 2019 (Hình 4a). Mặc dù vậy, do đây là khu vực miền núi đá vôi, lượng nước mưa sau khi được đổ xuống nhanh chóng bị tiêu thoát theo các hệ thống mạch nước, khiến diện tích nước trên bề mặt không nhiều, chỉ tập trung chủ yếu tại khu vực hồ Ba Bể. Theo đó, biến động diện tích hồ sẽ được phân tích ở mục sau. Các khu vực có hệ thống rừng gỗ hoặc cây bụi vẫn giữ chỉ số MNDWI thấp dưới 0. Trong khi các khu vực mặt nước có chỉ số MNDWI lớn hơn 0, riêng các khu vực sông suối và hồ ngập nước thường xuyên có chỉ số gần ngưỡng 1. Riêng khu vực canh tác nông nghiệp có phần biến động nhất do thời gian canh tác 1 vụ trong năm (Hình 4b). Toàn vùng lõi có 448 ha đất nông nghiệp (bình quân 0,73 ha/hộ và 0,12 ha/người) chủ yếu là canh tác nông nghiệp 1 vụ, ngoài ra một vài diện tích đất trồng màu tại các bãi ven sông suối và cửa sông suối đổ vào hồ. Hầu hết diện tích này chỉ canh tác được 1 vụ và thường xuyên bị cát bồi lấp nên năng suất rất thấp.

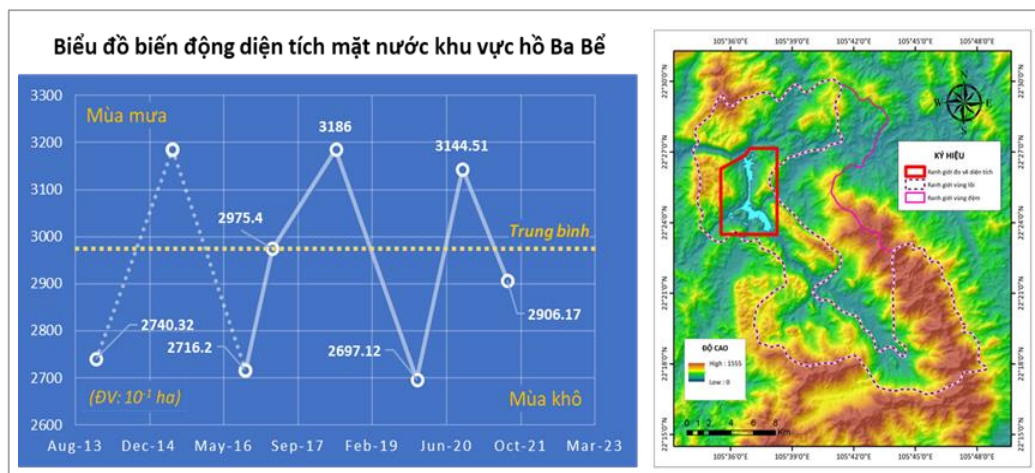


Hình 4. Kết quả tính toán chỉ số MNDWI của từng hệ sinh thái ĐNN và toàn khu vực VQG Ba Bể trong 7 năm.

3.4. Biến động diện tích hồ Ba Bể

Dựa trên kết quả tính toán chỉ số MNDWI, nghiên cứu tách thông tin mặt nước khu vực quanh hồ Ba Bể để đánh giá biến động hồ Ba Bể trong 7 năm (Hình 5). Theo đó, diện tích ngập nước trung bình quanh hồ Ba Bể là 290 ha/năm, trong mùa khô đạt trên 270 ha và trong mùa mưa đạt gần 320 ha. Trong 7 năm tính toán, dữ liệu ảnh mùa khô được tính vào các năm 2014, 2016 và 2019 và mùa mưa được tính toán vào các năm 2018 và 2020. Ảnh được chụp vào các tháng chuyển tiếp của các năm 2017 và 2021 thường chỉ mang lại diện tích ở mức trung bình trong năm. Trong vài năm gần đây, xuất hiện việc sử dụng chất độc và chất nổ để đánh bắt cá trong hồ Ba Bể, nhưng việc đánh bắt cá bằng các phương thức hủy diệt hiện đã hầu như được kiểm soát. Tuy nhiên, tàu thuyền du lịch đi lại trong hồ là nguồn xả chất thải rắn và cặn bã xăng dầu vào hồ. Một vấn đề môi trường lớn là hiện tượng bồi lắng lòng hồ do việc canh tác ở các vùng đầu nguồn của ba con sông nhánh chính đổ vào hồ. Mặc dù vậy, kết

qua tính toán mặt hồ không thể hiện rõ ràng điều này. Diện tích mặt nước có xu thế giảm nhẹ từ năm 2014 đạt 274 ha nhưng đến năm 2019 chỉ đạt 270 ha vào mùa khô. Ngay cả khi phỏng vấn người dân bản địa và các nhà quản lý khu vực cũng cho rằng diện tích mặt hồ có sự tác động của quá trình bồi lắng lòng hồ, nhưng sự suy giảm diện tích mặt hồ là không đáng kể.



Hình 5. Biến động diện tích mặt hồ khu vực Ba Bể trong 7 năm.

4. Thảo luận

So với các hệ thống phân loại đất ngập nước của RAMSAR và MONRE [9], nghiên cứu này tập trung vào 7 loại hình đất ngập nước lục địa. Mặc dù các mô hình phân loại đất ngập nước đã được phát triển trong một số nghiên cứu và các phương pháp tách lọc lớp phủ theo phương pháp truyền thống đã có nhiều trước đây, tuy nhiên các phương pháp này đòi hỏi nhiều thời gian, công sức, sai số lớn hơn nếu thực hiện trong nhiều thời điểm, thời gian khác nhau. Trong khi, mô hình Unet–Adadelta–256–64 lại thể hiện ưu thế về độ chính xác, tiết kiệm thời gian hơn rất nhiều cho các nhà nghiên cứu.

Ngoài ra, việc sử dụng dữ liệu viễn thám đã được tối ưu hóa trong nghiên cứu này. Mô hình được đào tạo sử dụng hình ảnh vệ tinh Sentinel–2 chất lượng cao (không có mây che phủ) được thu thập hai đến ba lần mỗi năm, có thể được sử dụng hiệu quả để theo dõi việc sử dụng đất ngập nước, lớp phủ thực vật thường xuyên.

5. Kết luận

Nghiên cứu này đã chứng minh lợi ích của việc kết hợp dữ liệu học sâu và viễn thám để theo dõi các kiểu đất ngập nước trong khu vực VQG Ba Bể. Mô hình U–Net được đào tạo đã sử dụng để xác định 7 kiểu hệ sinh thái ĐNN trong khu vực VQG trong vòng 7 năm (2014–2021). Trong đó, 7 loại hình đất ngập nước phù hợp với hệ thống phân loại của RAMSAR và MONRE. Ngoài các hệ sinh thái ngập nước theo mùa, hệ thống thủy văn dòng chảy karst là loại hình lớp phủ đặc biệt trong khu vực đã được xác định nhờ dữ liệu địa chất–địa mạo. Tuy nhiên, do nước tại vùng karst có khả năng tiêu thoát nước nhanh, mức độ phủ mặt nước khu vực tại một thời điểm không cao. Tập trung chủ yếu chỉ tại khu vực hồ Ba Bể.

Sau 7 năm, các hệ sinh thái đất ngập nước không thay đổi nhiều. Trong thời gian này, diện tích mặt nước suy giảm lớn nhất được ghi nhận vào năm 2019 trong khoảng 270 ha, với độ chính xác 87%. Mặc dù các thống kê cho thấy hiện tượng bồi lấp lòng hồ diễn ra mạnh, kết quả phân tích cho thấy hiện tượng này xuất hiện chủ yếu do hoạt động nông nghiệp ven lòng hồ. Hoạt động nông nghiệp bị thu hẹp nhiều vào những thời điểm nước rút, khiến người dân không thể phát triển sinh kế nông nghiệp tại đây. Các nhà quản lý sử dụng đất có thể sử dụng các mô hình trí tuệ nhân tạo để theo dõi động thái của hệ sinh thái đất ngập nước hoặc mặt hồ mỗi mùa. Các nhà khoa học có thể đào tạo lại mô hình học máy dễ dàng với các mẫu bổ sung trong tương lai và sử dụng nó để phân loại các vùng sinh thái đất ngập nước nội địa khác.

Đóng góp của tác giả: Xây dựng ý tưởng nghiên cứu: Đ.K.B., P.H.N.; Lựa chọn phương pháp nghiên cứu: Đ.K.B., N.V.Q., V.T.P.; Xử lý số liệu: T.T.C.; Phân tích mẫu: V.T.P.; Lấy mẫu: N.V.Q.; Viết bản thảo bài báo: Đ.K.B., P.H.N., N.T.V.; Chỉnh sửa bài báo: N.T.V., T.N.C.

Lời cảm ơn: Bài báo được thực hiện trong khuôn khổ đề tài khoa học và công nghệ: “Nghiên cứu, đề xuất giải pháp duy trì đặc tính sinh thái và cải thiện sinh kế cộng đồng tại một số khu Ramsar ở Việt Nam”, mã số TNMT.2020.562.08 của Bộ Tài nguyên và Môi trường.

Lời cam đoan: Tập thể tác giả cam đoan bài báo này là công trình nghiên cứu của tập thể tác giả, chưa được công bố ở đâu, không được sao chép từ những nghiên cứu trước đây; không có sự tranh chấp lợi ích trong nhóm tác giả.

Tài liệu tham khảo

1. Dang, K.B.; Nguyen, M.H.; Nguyen, D.A.; Phan, T.T.H.; Giang, T.L.; Pham, H.H.; Nguyen, T.N.; Van Tran, T.T.; Bui, D.T. Coastal Wetland Classification with Deep U-Net Convolutional Networks and Sentinel-2 Imagery: A Case Study at the Tien Yen Estuary of Vietnam. *Remote Sens.* **2020**, *12*, 1–26. doi:10.3390/rs12193270.
2. Sannigrahi, S.; Chakraborti, S.; Joshi, P.K.; Keesstra, S.; Sen, S.; Paul, S.K.; Kreuter, U.; Sutton, P.C.; Jha, S.; Dang, K.B. Ecosystem Service Value Assessment of a Natural Reserve Region for Strengthening Protection and Conservation. *J. Environ. Manage.* **2019**, *244*, 208–227. doi:10.1016/j.jenvman.2019.04.095.
3. Cục Bảo vệ Môi trường Việt Nam. Tổng Quan Hiện Trạng Đất Ngập Nước Việt Nam Sau 15 Năm Thực Hiện Công Ước Ramsar. Cục Bảo vệ Môi trường Việt Nam: Hà Nội, Việt Nam, 2005.
4. Huyền, N.T.T.; Minh, P.T.N.; Thủy, Đ.T.T. Bảo Tồn Đa Dạng Sinh Học Biển và Đất Ngập Nước (Kinh nghiệm từ một số Dự án nhỏ GRF). Viện Nghiên cứu GEF: Hà Nội, 2015.
5. He, Z.; He, D.; Mei, X.; Hu, S. Wetland Classification Based on a New Efficient Generative Adversarial Network and Jilin-1 Satellite Image. *Remote Sens.* **2019**, *11*, 1–16. doi:10.3390/rs11202455.
6. Ozesmi, S.L.; Bauer, M.E. Satellite Remote Sensing of Wetlands. *Wetl. Ecol. Manag.* **2002**, *10*, 381–402. doi:10.1023/A:1020908432489.
7. Đặng, T.H. Du Lịch Sinh Thái Vườn Quốc Gia Ba Bể, Huyện Ba Bể, Tỉnh Bắc Kan. *Tạp chí kinh tế sinh thái* **2017**, *54*, 12.
8. Thủ tướng Chính phủ. QĐ 83/TTg, Quyết Định Phê Duyệt Luận Chứng Kinh Tế – Kỹ Thuật Đầu Tư Xây Dựng Vườn Quốc Gia Ba Bể (Giai Đoạn I), 1992.
9. MONRE. Thông Tư 07/2020/TT-BTNMT Ngày 31/08/2020 về Hướng Dẫn Điểm c Khoản 1 Điều 31 Nghị Định 66/2019/NĐ-CP về Bảo Tồn và Sử Dụng Bền Vững Các Vùng Đất Ngập Nước, Hà Nội, 2020.
10. Perez, H.; Tah, J.H.M.; Mosavi, A. Deep Learning for Detecting Building Defects Using Convolutional Neural Networks. *Sensors* **2019**, *19*. doi:10.3390/s19163556.
11. Stoian, A.; Poulain, V.; Inglada, J.; Poughon, V.; Derksen, D. Land Cover Maps Production with High Resolution Satellite Image Time Series and Convolutional Neural Networks: Adaptations and Limits for Operational Systems. *Remote Sens.* **2019**, *11*, 1–26. doi:10.3390/rs11171986.
12. Giang, T.L.; Dang, K.B.; Toan Le, Q.; Nguyen, V.G.; Tong, S.S.; Pham, V.M. U-Net Convolutional Networks for Mining Land Cover Classification Based on High-Resolution UAV Imagery. *IEEE Access* **2020**, *8*, 186257–186273. doi:10.1109/access.2020.3030112.
13. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollar, P. Focal Loss for Dense Object Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 318–327. doi:10.1109/TPAMI.2018.2858826.

14. Choung, Y.J.; Jo, M.H. Comparison between a Machine–Learning–Based Method and a Water–Index–Based Method for Shoreline Mapping Using a High–Resolution Satellite Image Acquired in Hwado Island, South Korea. *Sensors* **2017**, 2017. doi:10.1155/2017/8245204.
15. Pouliot, D.; Latifovic, R.; Pasher, J.; Duffe, J. Assessment of Convolution Neural Networks for Wetland Mapping with Landsat in the Central Canadian Boreal Forest Region. *Remote Sens.* **2019**, 11. doi:10.3390/rs11070772.
16. Diakogiannis, F.I.; Waldner, F.; Caccetta, P.; Wu, C. ResUNet–a: A deep learning framework for semantic segmentation of remotely sensed data. *ISPRS J. Photogramm. Remote Sens.* **2020**, 162, 94–114. doi:10.1016/j.isprsjprs.2020.01.013

Application of artificial intelligence model in classification of inland wetland ecosystems in Ba Be National Park, Bac Kan province

Dang Kinh Bac^{1*}, Pham Hanh Nguyen², Nguyen Thanh Vinh², Tran Ngoc Cuong² Nguyen Van Quan³, Trinh Thi Thuy Chinh⁴, Vu Thi Phuong⁵,

¹ Faculty of Geography, University of Natural Sciences, Vietnam National University, Hanoi; dangkingbac@hus.edu.vn

² Department of Nature Conservation and Biodiversity, Vietnam Environment Administration, Ministry of Natural Resources and Environment; nguyennh.nbca@gmail.com

³ Survey and Aerial Mapping One Member Limited Liability Company, Department of Mapping, General Staff; quannguyen26795@gmail.com

⁴ Hoa Anh Pharmaceutical Joint Stock Company, 14/1 Lang Fort, Dong Da, Hanoi; Trinha2698@gmail.com

⁵ Hong Duc University, 565 Quang Trung, Thanh Hoa; vuthiphuong@hdu.edu.vn

Abstract: The degradation of wetland areas has been reported for the past 10 years, but the monitoring process of these changes is still a big challenge due to a lack of accessibility and technology. It is therefore essential to equip managers with tools to classify and monitor wetland ecosystems in real time. The aim of this study is to develop a useful model to classify inland wetlands in Ba Be National Park, Bac Kan Province and surrounding ecosystems based on the RAMSAR and MONRE classification systems. As a result, ResU–Net model using Adadelta optimization functions were used to classify seven types of wetlands and three types of neighboring ecosystems in Ba Be National Park, based on images obtained from the Sentinel–2 satellite, with more than 85% accuracy. The better performance of these models compared to Random forest (RF) and Support Vector Machine (SVM) methods has been demonstrated. The ResU–Net model, after optimization, is also used to map inland wetland areas in the Bac Kan mountains. This model is capable of updating new wetland types in Vietnam to monitor real–time wetland changes in the future.

Keywords: Wetland; RAMSAR; Ba Be National Park; U–Net; Remote sensing.