

PHÂN TÁCH TỰ ĐỘNG VÙNG NUÔI THỦY SẢN SỬ DỤNG NGUỒN ẢNH VIỄN THÁM ĐỘ PHÂN GIẢI CAO VÀ THUẬT TOÁN MÁY HỌC TẠI ĐÀM SAM CHUỒN - HÀ TRUNG, THỪA THIÊN HUẾ

**Hà Nam Thắng*, Phạm Hữu Ty, Trần Thị Thúy Hằng, Kiều Thị Huyền,
Trương Văn Đan**

Trường Đại học Nông Lâm, Đại học Huế

*Tác giả liên hệ: hanamthang@hueuni.edu.vn

Nhận bài: 09/10/2022 Hoàn thành phản biện: 18/01/2023 Chấp nhận bài: 03/02/2023

TÓM TẮT

Để quy hoạch hoạt động nuôi trồng thủy sản, việc lập bản đồ chính xác vùng nuôi từ ảnh viễn thám là mục tiêu quan trọng. Tuy nhiên, quá trình này còn hạn chế như độ chính xác phân loại dao động theo nhiều địa điểm, ảnh viễn thám sử dụng có độ phân giải chưa cao, phương pháp phân loại phức tạp. Nghiên cứu này sử dụng phương pháp khoanh vùng hướng đối tượng (Object-based Image Analysis - OBIA) và mô hình máy học Random Forest (RF) giúp lập bản đồ phân bố vùng nuôi thủy sản với độ chính xác và độ tin cậy cao từ ảnh PlanetScope (độ phân giải không gian 3 m) tại đầm Sam Chuồn - Hà Trung (thuộc khu hệ đầm phá Tam Giang - Cầu Hai). Tiến hành so sánh khả năng lập bản đồ với hai nguồn dữ liệu đầu vào khác nhau, kết quả cho thấy so với việc chỉ sử dụng tham số hình học (Precision 0,77, hệ số Kappa 0,471) giá trị phổ trung bình (Precision 0,94, hệ số Kappa 0,928) cải thiện rõ rệt độ chính xác trong phân tách vùng nuôi thủy sản. Tổng diện tích vùng nuôi được ước tính khoảng 1.000 ha, trong đó đầm Sam Chuồn có 454 ha và đầm Hà Trung có 546 ha vùng nuôi thủy sản. Nghiên cứu đóng góp công cụ viễn thám mới hỗ trợ quản lý hoạt động thủy sản một cách chính xác và bền vững.

Từ khóa: Cầu Hai, NTTS, OBIA, PlanetScope, Random Forest, Tam Giang

AUTOMATE CLASSIFICATION OF AQUACULTURE AREA USING HIGH SPATIAL SATELLITE IMAGERY AND MACHINE LEARNING IN SAM CHUON – HA TRUNG LAGOON, THUA THIEN HUE PROVINCE

**Ha Nam Thang*, Pham Huu Ty, Tran Thi Thuy Hang, Kieu Thi Huyen,
Truong Van Dan**

University of Agriculture and Forestry, Hue Univeristy

ABSTRACT

Precision mapping of aquaculture activities is an important target, facilitating a better management of aquaculture. Aquaculture mapping from satellite image is preferred, however remains the challenges of accuracy variance, low resolution imagery, and complex classification method. This study develops an accurate, simple and reliable approach using the Object-based Image Analysis (OBIA) and the machine learning Random Forest model to accurately map the aquaculture areas from the very high resolution PlanetScope (3 m) in Sam Chuon – Ha Trung (belongs to Tam Giang – Cau Hai lagoon). We aimed to compare the geometry and spectral dataset, which indicated the spectral support to map at higher confidence (Precision 0.94, Kappa coefficient 0.928) than the geometry dataset (Precision 0.77, Kappa coefficient 0.471). Aquaculture area was estimated as total 1,000 ha, of which 454 ha and 546 ha for Sam Chuon and Ha Trung lagoon, respectively. Our results contribute the novel approach of remote sensing, supporting a precision and sustainable management of aquaculture activities.

Keywords: Aquaculture, Cau Hai, OBIA, PlanetScope, Random Forest, Tam Giang

1. MỞ ĐẦU

Nuôi trồng thủy sản (NTTS) là hình thức sản xuất thủy sản chính trên các khu vực đầm phá, biên mở, góp phần giải quyết sinh kế và tạo nguồn thu nhập đáng kể cho nông - ngư dân (Naylor et al., 2021). Tuy vậy, sự phát triển quá mức cả về quy mô và số lượng các hình thức NTTS đang đặt ra nhiều thách thức về môi trường, sinh thái, gây ra xung đột giữa phát triển kinh tế với các mục tiêu bảo tồn, phát triển bền vững tại nhiều nơi trên thế giới (Spanou và cs., 2020). Với điều kiện thuận lợi về điều kiện tự nhiên, khí hậu, Việt Nam là quốc gia sản xuất thủy sản nổi tiếng với mức tăng trưởng xuất khẩu đều hàng năm (VASEP, 2021). Đi đôi với sự thành công trong phát triển thủy sản là những câu hỏi cấp thiết về khả năng quản lý diện tích, hình thức, vị trí NTTS nhằm tránh các xung đột tiềm tàng trong tương lai (Oglend, 2020).

Quản lý hoạt động NTTS từ ảnh viễn thám là hướng đi được chú ý trên thế giới với các giải pháp chọn vùng sản xuất, quy hoạch diện tích, hình thức sản xuất thủy sản phù hợp cũng như đánh giá tác động môi trường (Belkin, 2021). Các nghiên cứu này chủ yếu sử dụng ảnh viễn thám có độ phân giải thấp (100 - 300 m) đến trung bình - cao (10 - 100 m) với các kỹ thuật nhận diện và lập bản đồ truyền thống, dẫn đến việc tăng thời gian xử lý và giảm độ chính xác của đối tượng được nhận diện từ ảnh. Áp dụng các nguồn ảnh viễn thám có độ phân giải rất cao đến siêu cao (0,1 - 5 m) là một hướng đi mới và được chú trọng nhiều trong quản lý thủy sản, giúp nâng cao độ chính xác của bản đồ chuyên đề. Đối với nguồn ảnh này, kỹ thuật phân tích và nhận diện dựa trên điểm ảnh hoặc hướng đối tượng đều có thể được sử dụng, trong đó nhóm hướng đối tượng được áp dụng rộng rãi hơn giúp nhận diện chính xác các hình dáng đối tượng khác nhau trên nền ảnh có độ phân giải cao. Sau giai đoạn nhận diện đối tượng, các kỹ thuật phân loại ảnh khác nhau có thể được sử dụng để lập bản đồ phân bố như kỹ thuật phân loại phi kiểm định, kỹ thuật phân loại có kiểm định. Hiện tại, việc kết hợp với kỹ thuật phân loại

có kiểm định sử dụng các thuật toán máy học rất tiềm năng trong việc phân loại chính xác các đối tượng đã được nhận diện (Ha và cs., 2021). Tuy nhiên, nghiên cứu chỉ quan sát một số lượng hạn chế các nghiên cứu sử dụng hướng tiếp cận này trong hoạt động phân vùng, quản lý các phương tiện NTTS trên thế giới (Diep và cs., 2019; Rahlf và cs., 2021; Virdis, 2014).

Tại tỉnh Thừa Thiên Huế, hình thức NTTS rất đa dạng, trong đó tập trung chủ yếu vào các hình thức ao nuôi thủy sản vùng cao triều và thấp triều. Ngoại trừ một số khu vực được sắp xếp lại từ dự án IMOLA, vẫn còn tồn tại nhiều nơi đã hoặc đang phát triển tự phát, tự coi nói các ao nuôi với nhiều hình dáng khác nhau như khu vực đầm Sam Chuồn - Hà Trung, thuộc khu hệ đầm phá Tam Giang - Cầu Hai. Hiện tại, mật độ ao nuôi ở đây rất cao (Đàn và cs., 2018), có thể gây ra các tác động tiêu cực đến chất lượng nước, nông hóa nền đáy và hàng trăm héc ta thảm cỏ biển trong vùng, đặt ra nhu cầu cấp thiết trong lập bản đồ phân bố vùng NTTS, giúp bảo vệ môi trường, quản lý chính xác các hoạt động thủy sản và tối ưu hóa việc sử dụng không gian vùng bờ.

Trong nghiên cứu này, nghiên cứu đã thử nghiệm một hướng đi mới, kết hợp nguồn ảnh viễn thám độ phân giải rất cao PlanetScope (3 m), cùng với kỹ thuật phân tích, nhận diện đối tượng tự động hướng đối tượng (Object-Based Image Analysis - OBIA) và thuật toán máy học Random Forest (RF) trong lập bản đồ phân bố chính xác vùng nuôi thủy sản tại đầm Sam Chuồn - Hà Trung, tỉnh Thừa Thiên Huế. Nghiên cứu được kỳ vọng đóng góp các kỹ thuật mới trong xử lý ảnh viễn thám độ phân giải rất cao, lập bản đồ chính xác nhằm hỗ trợ tốt hơn công tác quản lý thủy sản không chỉ ở Việt Nam mà có thể mở rộng đến các khu vực khác trên thế giới.

2. NỘI DUNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

Dữ liệu được phân tích và xử lý qua hai bước (Hình 1), bao gồm bước 1 phân tách đối tượng với kỹ thuật OBIA và bước 2 phân loại đối tượng với mô hình RF.

2.1. Địa điểm nghiên cứu

Đầm Sam Chuồn - Hà Trung (SC-HT) (Hình 1) thuộc khu hệ đầm Thủy Tú, cùng với phá Tam Giang và đầm Cầu Hai tạo nên khu hệ đầm phá Tam Giang - Cầu Hai tại tỉnh Thừa Thiên Huế. SC-HT có độ dài ước tính 19 km với chiều ngang dao động từ 0,5 km (cầu Trường Hà) đến 5,7 km (đầm Sam), là nơi diễn ra mạnh mẽ các hoạt động sinh kế của người dân như NTTS, khai thác thủy sản, dịch vụ du lịch, nông nghiệp, trong đó NTTS là nghề được người dân phát triển mạnh mẽ. Sự gia tăng hoạt động NTTS như đào và xây dựng hồ nuôi, xả thải nước sau nuôi, và số lượng ao nuôi lớn quanh khu vực đầm SC-HT (Đàn và cs., 2018) đã gây ra sức ép lớn về mặt môi trường nước cũng như tác động đến hệ sinh thái xung quanh.

2.2. Thu thập ảnh vệ tinh

Bảng 1. Thông tin thu nhận ảnh PlanetScope dùng trong nghiên cứu

Ngày thu nhận ảnh	Độ phân giải không gian (m)	Độ phủ mây (%)	Số băng phổ	Dải phổ (nm)
31/5/2022	3	0	8	443 - 865

2.3. Hiệu chỉnh khí quyển

Ảnh PlanetScope được hiệu chỉnh khí quyển để tính chuyển giá trị phản xạ từ TOA về giá trị phản xạ tại bề mặt (Surface Reflectance - SR) với công cụ ACOLITE (Vanhellemont, 2016).

2.4. Dữ liệu định vị Global Positioning System (GPS)

Máy định vị GPS Extrex - 30 (độ chính xác ± 2 m) được sử dụng để ghi vị trí và hình dáng các ao nuôi thủy sản cao triều tại địa điểm nghiên cứu (đầm Sam Chuồn - Hà Trung) và sát với thời điểm thu nhận ảnh. Mỗi ao nuôi được định vị 4 lần với 4 điểm ở 4 góc của ao, trong đó điểm định vị ở góc ao được xác định là điểm góc của ao, nơi gần với mép nước trong ao nhất và có

Một cảnh PlanetScope được tải về từ trang <https://www.planet.com/products/explorer/> theo chương trình Giáo dục và Nghiên cứu (Education and Research Program <https://www.planet.com/markets/education-and-research/>) ở mức xử lý L3B trong dải phổ từ 443 - 865 nanometer (nm), bao gồm các băng ảnh vùng bờ (444 nm), xanh blue (492 nm), xanh green 1 (533 nm), xanh green 2 (566 nm), cam (612 nm), đỏ 1 (666 nm), đỏ 2 (707 nm), cận hồng ngoại (866 nm). (Bảng 1). Ảnh được chọn không có mây và vào mùa nắng (31/5/2022) để có thể tiến hành công tác thực địa (10/5 - 30/5/2022) gần nhất với thời điểm thu nhận ảnh, đã được nắn chỉnh địa hình và hiệu chỉnh khí quyển đến giá trị phản xạ ở thượng tầng khí quyển (Top of Atmospheric - TOA).

thể đứng cố định được. Trong giới hạn đề tài, có khoảng 20% - 30% số lượng các ao nuôi thủy sản cao triều được định vị làm dữ liệu đầu vào cho quá trình huấn luyện và kiểm định mô hình.

2.5. Phương pháp phân tách đối tượng sử dụng kỹ thuật Object Based Image Analysis (OBIA)

Trong nghiên cứu này, kỹ thuật nhận diện tự động OBIA được thiết lập với các tham số tối thiểu theo phương pháp thử - sai (thử nghiệm tham số đến khi đạt kết quả phù hợp) trong môi trường phần mềm mã nguồn mở GRASS GIS để thực hiện quá trình khoanh vùng tự động các đối tượng trên ảnh PlanetScope dựa trên thuật toán *regional growing* (Bảng 2).

Bảng 2. Tham số sử dụng với thuật toán *regional growing* của phương pháp OBIA dùng trong nghiên cứu

Tham số	Band width	Neighbourhood	Distance	Variance	Similar threshold
Giá trị	10	4	Feature space	0.001	0.001

Sau giai đoạn này, các vùng (22.341 đa giác) được trích xuất các tham số thống kê về đặc tính hình học, bao gồm các tham số mô tả diện tích của vùng (*area*), chiều dài của đường bao quanh tạo nên hình dáng của vùng (*perimeter*), không gian *compact* so với một hình vuông (*compact_square*), không gian *compact* so với một hình tròn (*compact_circle*), một đặc tính tổng hợp từ tham số *area* và *perimeter* (*fractal dimension* (fd)) và giá trị phổ trung bình của mỗi vùng. Quá trình phân tách và tính toán tham số thống kê cho các vùng được thực hiện trong môi trường phần mềm GRASS GIS.

2.6. Phương pháp phân loại đối tượng sau phân tách sử dụng kỹ thuật Random Forest

Random Forest (RF) (Breiman, 2001) là mô hình máy học có tính ổn định và độ tin cậy cao trong quá trình hoạt động với nhiều thành công trong các nghiên cứu thực tiễn (Belgiu & Drăguț, 2016). RF sử dụng cấu trúc rừng cây quyết định gồm nhiều cây quyết định, được xem như một *weak learner* (thành phần học yếu) với quá trình trích xuất mẫu ngẫu nhiên có lặp lại từ bộ mẫu, trong đó 2/3 được dùng để huấn luyện và 1/3 để kiểm định mô hình. Mỗi cây quyết định sử dụng một bộ mẫu trích xuất

ngẫu nhiên khác nhau với các tham số chính như độ sâu tối đa (*maximum depth*), số lượng cây (*number of estimator*), số lượng biến tối đa (*maximum features*), số lượng mẫu lá (*minimum sample leaf*) và lần chia tối thiểu (*minimum sample split*) để đưa ra dự báo từ dữ liệu đầu vào. Phương pháp trung bình được áp dụng để chọn kết quả dự báo tốt nhất trong mô hình RF.

Giai đoạn phân loại đối tượng gồm hai bước. Bước 1, chúng tôi sử dụng dữ liệu GPS của hai lớp gồm lớp 1 (không phải vùng nuôi thủy sản) và lớp 2 (vùng nuôi thủy sản) nhằm xác định và chọn các đa giác trên nền ảnh tính chuyển từ kỹ thuật OBIA và tạo bộ dữ liệu dùng để huấn luyện và kiểm định mô hình RF. Do việc phân loại trên dữ liệu vector khá tốn kém về mặt tính toán và chưa có nhiều công cụ chuẩn hỗ trợ quá trình này, chúng tôi chuyển ảnh tính chuyển từ kỹ thuật OBIA dạng vector thành định dạng raster và tiến hành phân loại với hai lớp 1 và 2, sử dụng mô hình RF (bước 2). Quá trình phân loại được thực hiện trong môi trường dòng lệnh Python với một bộ tham số được dùng chung (Bảng 3) cho hai nguồn dữ liệu, bao gồm tham số hình học và giá trị phổ trung bình (chỉ khác giá trị *max features* do sai khác về số lượng băng ảnh đầu vào ở mỗi nguồn dữ liệu).

Bảng 3. Tham số dùng trong mô hình phân loại RF

Tham số	Giá trị	Tham số	Giá trị
Bootstrap	True	Max features	Theo nguồn dữ liệu (5, 8)
Max depth	4	Min sample leaf	1
Number of estimator	50	Min sample split	3

Dữ liệu đầu vào gồm 198.105 điểm ảnh, được chia ngẫu nhiên (sử dụng thư viện sklearn (Pedregosa và cs., 2011) trong môi trường lập trình Python) theo tỉ lệ 50% (99.052 điểm ảnh) dữ liệu để huấn luyện mô

hình và 50% (99.052 điểm ảnh) dữ liệu để kiểm định mô hình.

2.7. Đánh giá độ chính xác

Nghiên cứu sử dụng các phép đo tiêu chuẩn, được sử dụng rộng rãi trong các

nghiên cứu về ứng dụng viễn thám trên thế giới để đánh giá độ chính xác quá trình phân loại ảnh. Các phép đo này bao gồm giá trị Precision (độ chính xác theo lớp phân loại), Recall (mức độ bỏ sót trong phân loại), F_1

(tham số trung bình của Precision và Recall), độ chính xác chung (overall accuracy - OA) và hệ số Kappa (công thức (1) – (5)) (Hossin & Sulaiman, 2015).

$$Precision = \frac{t_p}{t_p + f_p} \quad (1)$$

$$Recall = \frac{t_p}{t_p + f_n} \quad (2)$$

$$F_1 = \frac{2 * precision * recall}{precision + recall} \quad (3)$$

trong đó:

t_p : true positive là dự báo đúng

f_p : false positive là dự báo nhầm

f_n : false negative là dự báo bị bỏ sót

$$OA(y, y_{pred}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} 1(y_{pred} = y_i) \quad (4)$$

trong đó y_{pred} là giá trị dự báo và y là giá trị đo thật

$$Kappa = \frac{p_o - p_e}{1 - p_e} \quad (5)$$

trong đó:

p_o : giá trị đồng thuận quan sát

p_e : xác suất giả định của khả năng đồng thuận

3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Thông tin chung về ao nuôi thủy sản tại khu vực nghiên cứu

Ao nuôi tại đầm SC-HT có hình dáng và kích thước rất đa dạng (Hình 3d - 3f), phân bố trải dài theo hình dáng của đầm. Diện tích các ao biến thiên lớn từ 1.000 - 5.400 m² với hình dáng không cố định (thay đổi từ hình chữ nhật đến gần vuông), phụ thuộc vào diện tích và không gian có sẵn lúc xây dựng ao. Về phân bố không gian, hầu hết các ao được xây dựng tại khu vực cao triều và một số ít được người dân bồi đắp và kiến tạo ở khu vực hạ triều (Hình 3a, 3c, 3e). Ao được bố trí liền kề thành các khu vực liên tục, với nhiều hơn các khu vực mới được coi nói để xây dựng ao nuôi.

3.2. Lập bản đồ phân bố vùng nuôi thủy sản sử dụng mô hình RF

Lập bản đồ phân bố vùng nuôi thủy sản sử dụng mô hình RF và tham số hình học

Sử dụng nguồn dữ liệu thống kê mô tả đặc tính hình học của các vùng, mô hình cho thấy có thể nhận diện vùng nuôi thủy sản với độ chính xác khá (Precision 0,77) nhưng độ tin cậy của kết quả phân loại còn thấp với giá trị F_1 và hệ số Kappa chỉ đạt lần lượt là 0,68 và 0,47 (Bảng 4). Đối với nguồn dữ liệu này, mô hình RF chỉ ra tham số fd có mức đóng góp cao nhất (64,6%), tiếp theo là *perimeter* (11,5%), *area* (10,2%) và *compact_square* (9,38%); thấp nhất là tham số *compact_circle* với 4,31% (Hình 4).

Lập bản đồ phân bố vùng nuôi thủy sản sử dụng mô hình RF và giá trị phổ trung bình

Với nguồn dữ liệu đầu vào là giá trị phổ trung bình, kết quả cho thấy có sự cải thiện rõ rệt về cả độ chính xác và mức độ tin cậy trong kết quả giải đoán. Vùng nuôi thủy sản có thể được nhận diện với độ chính xác cao (Precision 0,94) và hầu như không bỏ sót khu vực nuôi nào (Recall 0,98). Giá trị cao của tham số F_1 (0,964) và hệ số Kappa

(0,924) cho thấy tính ổn định và độ tin cậy cao của kết quả giải đoán (Bảng 4). Khi xem xét mức đóng góp của các băng phổ khác nhau, chúng tôi nhận thấy tầm quan trọng của các băng là không đồng đều. Các băng cận hồng ngoại (Băng 8 cận hồng ngoại, 864 nm), băng đỏ 2 (Băng 7, 707 nm) và băng xanh 2 (Băng 4, 566 nm) có mức đóng góp cao nhất lần lượt là 74,2%, 15,5% và 6,75% trong khi các băng còn lại tham gia không đáng kể vào quá trình nhận diện vùng nuôi và chỉ chiếm khoảng 3% (Hình 5).

Bảng 4. Độ chính xác kết quả lập bản đồ phân bố vùng nuôi từ các nhóm dữ liệu khác nhau

	Độ chính xác (Precision)	Mức độ bỏ sót (Recall)	F_1	Độ chính xác chung (OA)	Hệ số Kappa
Tham số hình học	0,77	0,60	0,68	0,745	0,471
Giá trị phổ trung bình	0,94	0,98	0,96	0,964	0,928

Với độ chính xác cao nhất, mô hình RF sử dụng 8 băng ảnh giá trị phổ trung bình được sử dụng để xuất bản đồ vùng nuôi tại đầm SC-HT (Hình 6). Tổng diện tích vùng nuôi ước tính khoảng 1.000 ha, trong đó diện tích vùng nuôi ở đầm Hà Trung là 546 ha và đầm Sam Chuồn là 454 ha.

Như vậy, với hai nhóm dữ liệu được thử nghiệm, việc chỉ sử dụng các tham số hình học (*area*, *perimeter*, *compact_square*, *compact_circle*, *fd*) khiến quá trình nhận diện vùng nuôi không đạt được độ tin cậy cao (hệ số Kappa 0,471) và nhiều vùng nuôi bị bỏ sót trong quá trình phân loại (Precision 0,77, Recall 0,60), trong khi bổ sung giá trị phổ trung bình của các băng giúp nhận diện tốt hơn hẳn các vùng nuôi (hệ số Kappa 0,928, Precision 0,94). Điều này có thể được lý giải từ sự đa dạng về hình dáng và diện tích của ao nuôi (Hình 3), dẫn đến các tham số thống kê hình học không thể mô tả đúng nhất hình dáng của ao hoặc vùng nuôi ở độ phân giải 3 m của ảnh PlanetScope. Tham số *fd*, *perimeter*, và *area* mô tả độ phức tạp, chiều dài và diện tích của đối tượng có mức ảnh hưởng lớn nhất đến quá trình nhận diện vùng nuôi từ tham số hình học, trong đó tham số *fd* giúp nhận diện hình dáng vùng nuôi tốt nhất. Có được điều này là do *fd* được tính tổng hợp từ hai tham số

perimeter và *area*, giúp tìm ra được sự khác biệt trong hình dạng và kích thước của vùng trong khi *perimeter* và *area* có thể trả về nhiều kết quả giống nhau do các đối tượng khác nhau có thể giống nhau về diện tích và chiều dài.

Một yếu tố khác, theo chúng tôi, cũng tác động lớn đến khả năng nhận diện các ao nuôi từ tham số hình học đó là bờ ngăn giữa các ao có bề rộng (khoảng 1,6 m) dưới kích thước điểm ảnh (3 m) và do vậy, không đủ để mô hình nhận diện ao riêng lẻ trong quá trình khoanh vùng từ phương pháp OBIA. Việc sử dụng giá trị phổ trung bình, vùng nuôi được nhận diện tốt hơn từ các băng phổ cận hồng ngoại (băng 8), băng cận đỏ 2 (băng 7) và băng xanh 2 (băng 4). Điều này có thể giải thích do đặc trưng phản xạ phổ tại vùng nuôi là phản xạ phổ của nước (trong ao nuôi), nền cát và đất trống (bờ ao, xung quanh ao), khác với phản xạ phổ từ vùng sản xuất nông nghiệp và khu dân cư xung quanh vùng nuôi. Tuy nhiên, do kích thước bờ ao không đủ lớn nên việc phân loại dừng lại ở mức nhận diện chính xác vùng nuôi, khu vực tập hợp nhiều ao nuôi liền kề nhau.

Nghiên cứu cho thấy việc sử dụng phương pháp OBIA kết hợp với phương pháp phân loại máy học (mô hình RF) trong

nhận diện vùng nuôi nói chung và ao nuôi nói riêng từ nguồn ảnh độ phân giải rất cao có độ chính xác cao và hoàn toàn đáng tin cậy (Bảng 4). OBIA có thể giúp nhận diện nhanh và chính xác dựa trên hình dáng và phân bố không gian của vùng nuôi (Hình 3), người dùng từ đó có thể trích xuất nhiều giá trị thống kê khác nhau của vùng nuôi và sử dụng như nguồn dữ liệu đầu vào cho quá trình phân loại. Với đầm SC-HT, mô hình RF thể hiện khả năng phân loại tốt với các tham số ở mức thấp (số cây quyết định *number of estimator* 50, độ sâu của mô hình *max depth* 4, mức độ phân nhánh của cây quyết định *min sample leaf* 1), giúp mô hình tính toán nhanh hơn và giảm khả năng phân loại sai. Bên cạnh đó, nghiên cứu duy trì bộ tham số chung (Bảng 3) của mô hình RF khi tiến hành so sánh hai nguồn dữ liệu đầu vào, giúp duy trì tính khách quan khi so sánh và đồng thời cho thấy RF có tính khái quát hóa rất cao khi hoạt động ổn định với các nguồn dữ liệu khác nhau. So với các nghiên cứu gần tương tự trên thế giới, như sử dụng ảnh độ phân giải siêu cao WorldView - 2 lập bản đồ phân bố lồng nuôi cá từ kỹ thuật Nearest Neighbor Classification (NNC) với độ chính xác từ 87% - 92% (Fu và cs., 2019), Gao-Feng 1 với độ chính xác 67% - 100% sử dụng kỹ thuật khoanh vùng nhiều mức độ cho khu vực NTTS nói chung (Zhu và cs., 2019), hay (Fu và cs., 2021) và (Liu và cs., 2019) áp dụng kỹ thuật học sâu Convolution Neural Network (CNN) và Richer Convolution Features (RCF) nhận diện khu vực NTTS trên biển đạt độ chính xác chung 93% - 95% thì kết quả của nghiên cứu này (độ chính xác 94%, hệ số Kappa 0,928) rất đáng khích lệ và tiềm năng trong mở rộng phạm vi áp dụng. Kết quả trong nghiên cứu này cũng cao hơn một số kết quả khác đạt được độ chính xác 84% - 92% và 87% - 94% với ảnh Sentinel - 1 (Ottinger và cs., 2017) và Landsat (Ren và cs., 2019) trong khoanh vùng NTTS. Tiến hành so sánh với một số kết quả từ công trình trong nước sử dụng ảnh Landsat (77% - 91%, (Diep và cs., 2019; Nguyễn và cs., 2018)), (84% - 92% (Béland và cs., 2006)) trong phân tách đất

NTTS, chúng tôi nhận thấy độ chính xác còn khá dao động theo khu vực nghiên cứu và phương pháp phân tích ảnh chưa thực sự có độ tin cậy cao so với hướng tiếp cận trong nghiên cứu này. Gần nhất, (Virdis, 2014) áp dụng kỹ thuật khoanh vùng OBIA trong đánh giá phân bố vùng NTTS từ ảnh vệ tinh SPOT 5 và WordView - 1 với độ chính xác chung giao động từ 84% - 93% và 90% - 95% tại đầm phá Tam Giang - Cầu Hai. Kết quả của (Virdis, 2014) có thể nói khá tương đương với kết quả trong nghiên cứu tại đầm SC-HT, tuy nhiên phương pháp do (Virdis, 2014) đề xuất là phức tạp và đòi hỏi kỹ năng chuyên sâu trong phân tích ảnh viễn thám của người sử dụng. Phương pháp OBIA và RF, như được đề xuất trong nghiên cứu này khá đơn giản, dễ hiểu và có thể được thực hiện với các phần mềm mã nguồn mở miễn phí (như QGIS, SAGA GIS, Orfeo Toolbox) hoặc với môi trường lập trình mở như ngôn ngữ Python, do vậy phương pháp hoàn toàn có thể lặp lại với kỹ thuật tương tự.

Tuy kết quả thể hiện rất tốt trong việc nhận diện vùng nuôi, việc phân tách được riêng từng ao nuôi vẫn còn là một khó khăn với hệ thống nuôi ở đầm SC-HT. Sự đa dạng trong hình dáng, diện tích và khoảng cách bờ giữa các ao rất nhỏ khiến phương pháp OBIA sử dụng thuật toán *regional growing* chưa thể tách được riêng từng ao với độ phân giải 3 m của ảnh PlanetScope. Trong các nghiên cứu tiếp theo, chúng tôi sẽ thử nghiệm các phương pháp khoanh vùng khác nhau (phương pháp *mean-shift*, *connected components*, *watershed* và *phương pháp dựa trên hình thái*) với nguồn ảnh PlanetScope hoặc thử nghiệm phương pháp *regional growing* với ảnh độ phân giải siêu cao (< 1 m, WorldView, Pleiades, GeoEye, SkySat). Với kết quả hiện tại, nghiên cứu đóng góp kỹ thuật nhận diện vùng nuôi có độ chính xác và độ tin cậy cao, rất tiềm năng trong hỗ trợ quá trình đánh giá phân bố và quản lý vùng nuôi một cách bán tự động, giảm được rất nhiều chi phí, nhân lực và thời gian so với đi điều tra thực địa toàn vùng.

4. KẾT LUẬN

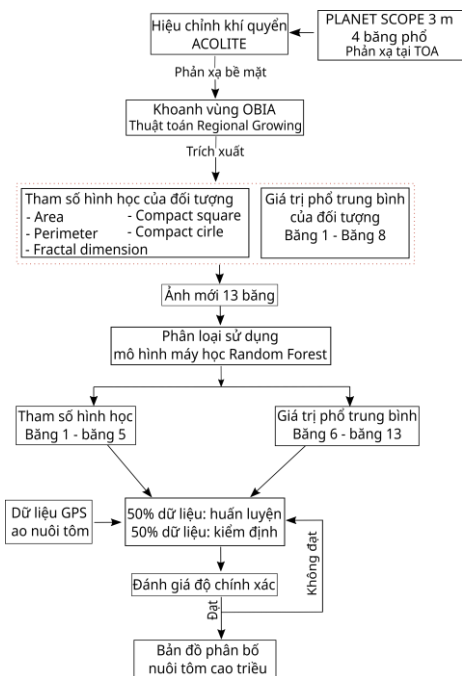
Nuôi trồng thủy sản là một ngành sản xuất quan trọng, đóng góp giá trị gia tăng rất lớn vào nền kinh tế nói chung. Sự gia tăng nhanh chóng về số lượng ao nuôi đặt ra yêu cầu cấp bách trong việc phát triển các phương pháp lập bản đồ vùng nuôi nhanh, có độ tin cậy và sử dụng kỹ thuật tiên tiến. Nghiên cứu này đã thử nghiệm phương pháp khoanh vùng OBIA kết hợp với phương pháp phân loại máy học sử dụng mô hình RF giúp nhận diện thành công vùng nuôi ở độ chính xác và độ tin cậy cao (Precision 0,94, hệ số Kappa 0,928) sử dụng dữ liệu giá trị phổ trung bình. Với hệ thống vùng nuôi ở đầm SC-HT, việc nhận diện chỉ từ các tham số hình học chưa có độ tin cậy cao (hệ số Kappa 0,471) so với việc sử dụng giá trị phổ trung bình. Với phương pháp đề xuất, tổng diện tích vùng nuôi ở đầm SC-HT được ước tính khoảng 1.000 ha với diện tích xấp xỉ bằng nhau ở đầm Sam Chuồn

(454 ha) và Hà Trung (546 ha).

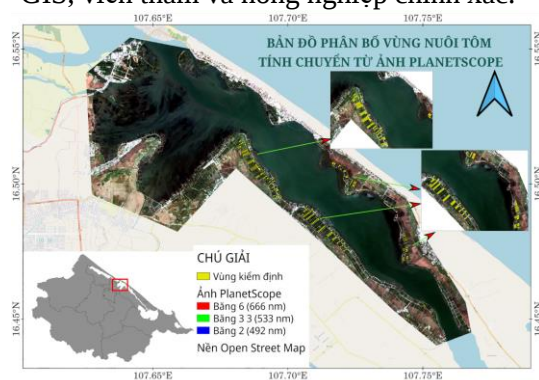
Phương pháp sử dụng OBIA sử dụng thuật toán *regional growing* và RF có thể kết hợp với bất kỳ ảnh vệ tinh có độ phân giải rất cao giúp lập bản đồ phân bố các vùng nuôi tại khu vực ven bờ biển hoặc đầm phá trên lãnh thổ Việt Nam và thế giới. Chúng tôi đề xuất nhân rộng, áp dụng rộng rãi phương pháp này ở các vùng nuôi sử dụng ao ở các khu vực bờ biển, đầm phá của Việt Nam đồng thời giúp đẩy nhanh quá trình số hóa dữ liệu vùng NTTS ở phạm vi cấp tỉnh và cấp quốc gia trong các năm tiếp theo.

LỜI CẢM ƠN

Nhóm tác giả xin chân thành gửi lời cảm ơn đến trường Đại học Nông Lâm (Đại học Huế) đã cung cấp kinh phí để thực hiện đề tài thông qua quỹ thực hiện đề tài cấp cơ sở và quỹ của chương trình nghiên cứu mạnh với mã số NCM.ĐHNL.2021.03: GIS, viễn thám và nông nghiệp chính xác.



Hình 1. Sơ đồ các bước thực hiện trong nghiên cứu



Hình 2. Đầm Sam Chuồn- Hà Trung, địa điểm thực hiện nghiên cứu (Ảnh vệ tinh PlanetScope tổ hợp màu 6-4-3 trên nền bản đồ Open Street Map. Hình chữ nhật màu đỏ chỉ vị trí vùng nghiên cứu trên địa bàn tỉnh Thừa Thiên Huế). Đường màu xanh sáng xác định ranh giới tương đối giữa đầm Sam Chuồn (phần phía trên) và Hà Trung (phần phía dưới). Hai ảnh nhỏ là ảnh phóng lớn hai khu vực với dữ liệu kiểm định (đa giác màu vàng)



(a)



(b)



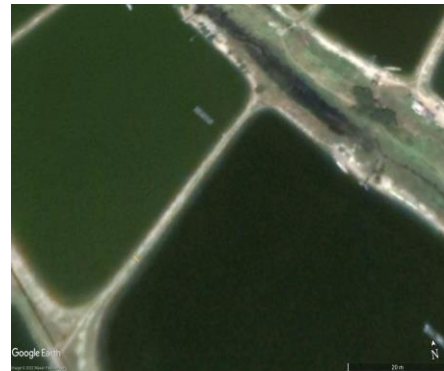
(c)



(d)

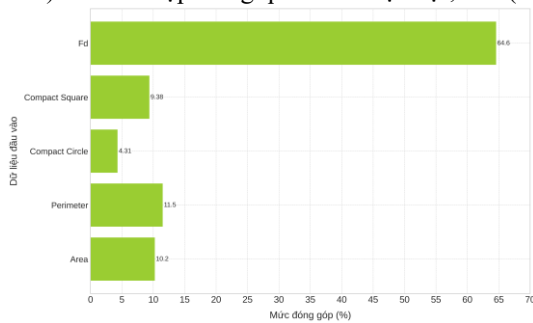


(e)

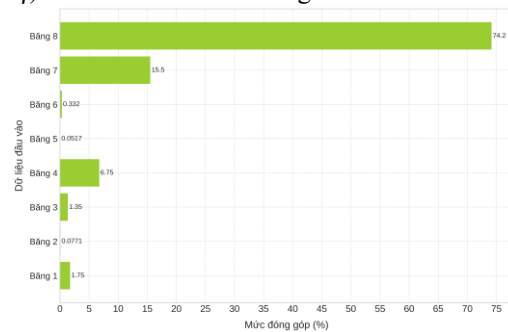


(f)

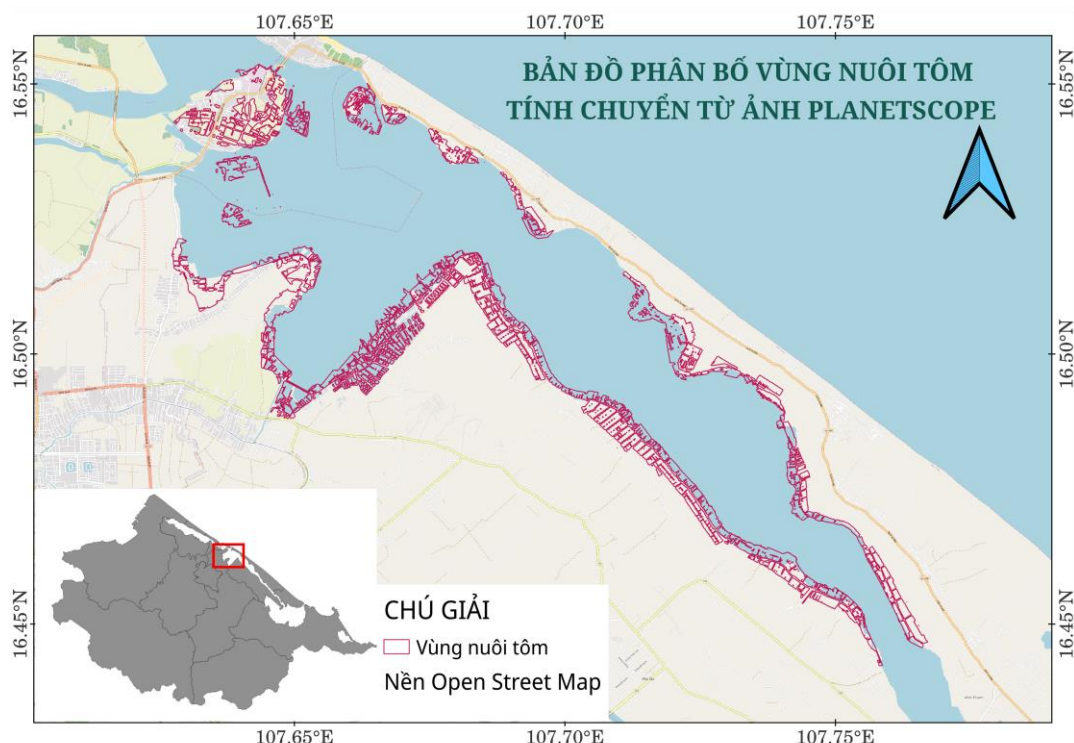
Hình 3. Ao nuôi thủy sản tại khu vực đầm Sam Chuồn - Hà Trung: (a - b): bờ ao và góc ao; (c) một góc khu vực nuôi; (d) ao nuôi tại đầm Hà Trung; (e) ao nuôi tại đầm Sam Chuồn; (f) đường bờ ao (màu trắng, phân cách giữa 2 ao nuôi) có chiều rộng 1,6 m, dưới kích thước điểm ảnh PlanetScope); Ảnh (3a - 3c) là ảnh chụp trong quá trình thực địa, ảnh (3d - 3f) là ảnh trích xuất từ Google Earth



Hình 4. Mức đóng góp của các băng ảnh khi sử dụng tham số hình học



Hình 5. Mức đóng góp của các băng ảnh khi sử dụng tham số giá trị phổ trung bình



Hình 6. Phân bố vùng nuôi thủy sản xây dựng từ ảnh PlanetScope với thuật toán khoanh vùng region growing và phân loại Random Forest

TÀI LIỆU THAM KHẢO

1. Tài liệu tiếng Việt

- Đàn T. V., Ut V. N., Thắng H. N., Niêm P. T. A., Luân N. T., & Bình M. N. (2018). Điều tra hiện trạng nuôi trồng thủy sản lợi nhuận cao triều ở xã Phú Mỹ, huyện Phú Vang, tỉnh Thừa Thiên Huế, Việt Nam. *Can Tho University, Journal of Science*, 54(7), 126. <https://doi.org/10.22144/ctu.jvn.2018.149>
- Nguyễn, T. H. Đ., Nguyễn, Q. X., Nguyễn, K. K., & Nguyễn, H. (2018). Tích hợp ảnh viễn thám Landsat 8 xây dựng bản đồ hiện trạng thủy sản vùng ven biển tỉnh Sóc Trăng. *Tạp Chí Khoa Học Đại Học Sư Phạm TP Hồ Chí Minh*, 15(11b), 116–123.

2. Tài liệu tiếng nước ngoài

- Béland, M., Goïta, K., Bonn, F., & Pham, T. T. H. (2006). Assessment of land-cover changes related to shrimp aquaculture using remote sensing data: A case study in the Giao Thủy District, Vietnam. *International Journal of Remote Sensing*, 27(8), 1491–1510. <https://doi.org/10.1080/01431160500406888>
- Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>

- Belkin, I. M. (2021). Remote Sensing of Ocean Fronts in Marine Ecology and Fisheries. *Remote Sensing*, 13(5), 883. <https://doi.org/10.3390/rs13050883>
- Breiman, L. (2001). Random Forest. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Diep, N. T. H., Korsem, T., Can, N. T., Phonphan, W., & Minh, V. Q. (2019). Determination of aquaculture distribution by using remote sensing technology in Thanh Phu district, Ben Tre province, Vietnam. *Vietnam Journal of Science, Technology and Engineering*, 61(2), 35–41. [https://doi.org/10.31276/VJSTE.61\(2\).35-41](https://doi.org/10.31276/VJSTE.61(2).35-41)
- Fu, Y., Deng, J., Wang, H., Comber, A., Yang, W., Wu, W., You, S., Lin, Y., & Wang, K. (2021). A new satellite-derived dataset for marine aquaculture areas in China's coastal region. *Earth System Science Data*, 13(5), 1829–1842. <https://doi.org/10.5194/essd-13-1829-2021>
- Fu, Y., Deng, J., Ye, Z., Gan, M., Wang, K., Wu, J., Yang, W., & Xiao, G. (2019). Coastal Aquaculture Mapping from Very High Spatial Resolution Imagery by Combining Object-Based Neighbor Features. *Sustainability*, 11(3), 637. <https://doi.org/10.3390/su11030637>
- Ha, N.-T., Manley-Harris, M., Pham, T.-D., & Hawes, I. (2021). Detecting Multi-Decadal Changes in Seagrass Cover in Tauranga

- Harbour, New Zealand, Using Landsat Imagery and Boosting Ensemble Classification Techniques. *ISPRS International Journal of Geo-Information*, 10(6), 371. <https://doi.org/10.3390/ijgi10060371>
- Hossin, M., & Sulaiman, M. N. (2015). A Review on Evaluation Metrics for Data Classification Evaluations. *International Journal of Data Mining & Knowledge Management Process*, 5(2), 01–11. <https://doi.org/10.5121/ijdkp.2015.5201>
- Liu, Y., Yang, X., Wang, Z., Lu, C., Li, Z., & Yang, F. (2019). Aquaculture area extraction and vulnerability assessment in Sanduao based on richer convolutional features network model. *Journal of Oceanology and Limnology*, 37(6), 1941–1954. <https://doi.org/10.1007/s00343-019-8265-z>
- Naylor, R. L., Hardy, R. W., Buschmann, A. H., Bush, S. R., Cao, L., Klinger, D. H., Little, D. C., Lubchenco, J., Shumway, S. E., & Troell, M. (2021). A 20-year retrospective review of global aquaculture. *Nature*, 591(7851), 551–563. <https://doi.org/10.1038/s41586-021-03308-6>
- Oglend, A. (2020). Challenges and opportunities with aquaculture growth. *Aquaculture Economics & Management*, 24(2), 123–127. <https://doi.org/10.1080/13657305.2019.1704937>
- Ottinger, M., Clauss, K., & Kuenzer, C. (2017). Large-Scale Assessment of Coastal Aquaculture Ponds with Sentinel-1 Time Series Data. *Remote Sensing*, 9(5), 440. <https://doi.org/10.3390/rs9050440>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Rahlf, J., Hauglin, M., Astrup, R., & Breidenbach, J. (2021). Timber volume estimation based on airborne laser scanning—Comparing the use of national forest inventory and forest management inventory data. *Annals of Forest Science*, 78(2), Article 2. <https://doi.org/10.1007/s13595-021-01061-4>
- Ren, C., Wang, Z., Zhang, Y., Zhang, B., Chen, L., Xi, Y., Xiao, X., Doughty, R. B., Liu, M., Jia, M., Mao, D., & Song, K. (2019). Rapid expansion of coastal aquaculture ponds in China from Landsat observations during 1984–2016. *International Journal of Applied Earth Observation and Geoinformation*, 82, 101902. <https://doi.org/10.1016/j.jag.2019.101902>
- Spanou, E., Kenter, J. O., & Graziano, M. (2020). The Effects of Aquaculture and Marine Conservation on Cultural Ecosystem Services: An Integrated Hedonic – Eudaemonic Approach. *Ecological Economics*, 176, 106757. <https://doi.org/10.1016/j.ecolecon.2020.106757>
- Vanhellemont, Q. (2016). ACOLITE For Sentinel-2: Aquatic Applications of MSI imagery. *Proceedings of the 2016 ESA Living Planet Symposium*, 8. https://odnature.naturalsciences.be/downloads/publications/vanhellemonttruddick_esa_lps2016_coastalapplications_final_header.pdf
- VASEP. (2021, September 26). *Vietnam exports of marine products in the first 8 months of 2021 reached 2.1 billion USD*. Vietnam. <http://seafood.vasep.com.vn/key-seafood-sectors/other-marine-products/vietnam-exports-of-marine-products-in-the-first-8-months-of-2021-reached-2-1-billion-usd-22865.html>
- Virdis, S. G. P. (2014). An object-based image analysis approach for aquaculture ponds precise mapping and monitoring: A case study of Tam Giang-Cau Hai Lagoon, Vietnam. *Environmental Monitoring and Assessment*, 186(1), 117–133. <https://doi.org/10.1007/s10661-013-3360-7>
- Zhu, H., Li, K., Wang, L., Chu, J., Gao, N., & Chen, Y. (2019). Spectral Characteristic Analysis and Remote Sensing Classification of Coastal Aquaculture Areas Based on GF-1 Data. *Journal of Coastal Research*, 49–57.