

NGHIÊN CỨU MÔ HÌNH PHÁT HIỆN DÒNG CHẢY XA BỜ TỪ ẢNH UAV SỬ DỤNG MẠNG NƠON TÍCH CHẬP

Nguyễn Đình Công^{1*}

¹Trường Đại học Hồng Đức

* Email: nguyendinhcong@hdu.edu.vn

Ngày nhận bài: 05/9/2022

Ngày nhận bài sửa sau phản biện: 04/11/2022

Ngày chấp nhận đăng: 06/11/2022

TÓM TẮT

Dòng nước chảy xa bờ xuất hiện tại các bãi biển, đặc biệt là các bãi tắm biển du lịch, tiềm ẩn nguy cơ tới an toàn của người bơi do cường độ mạnh, xuất hiện đột ngột và bất thường, đôi khi dòng nước chảy xa bờ có thể kéo người bơi ra xa với vận tốc 1 – 2.5 m/s. Trên thế giới nói chung và ở Việt Nam nói riêng, hằng năm có nhiều trường hợp được ghi nhận tử vong bởi dòng nước chảy xa bờ. Mặc dù vậy, dòng nước chảy xa bờ thường khó nhận biết và phát hiện bằng mắt thường. Nghiên cứu này đề xuất giải pháp sử dụng mô hình mạng nơon tích chập để phát hiện các dòng nước chảy xa bờ trên các bãi tắm thông qua ảnh chụp định kỳ từ các thiết bị bay không người lái. Mô hình phát hiện các dòng nước chảy xa bờ đạt độ chính xác 90.19% trên tập dữ liệu kiểm tra và với tốc độ 15 khung hình trên giây. Kết quả nghiên cứu là tiền đề để xây dựng ứng dụng trong các hệ thống giám sát bờ biển nhằm hỗ trợ cho nhân viên bờ biển phát hiện nhanh dòng nước chảy xa bờ để đưa ra các cảnh báo tới người bơi.

Từ khóa: dòng nước chảy xa bờ, hệ thống giám sát bờ biển, mạng nơon tích chập, thiết bị bay không người lái.

RIP CURRENT DETECTION BASED ON UAV IMAGES AND CONVOLUTIONAL NEURAL NETWORK

ABSTRACT

Rip currents are powerful, narrow seaward flows of water in the surf zone. Rips can appear in any beach environment because they are dependent on wave breaking. Their formations are given by wave characteristics, coastal morphology, and wind. From the beach-safety aspect, rip currents could be considered as one of the most dangerous threats to bathers, pulling swimmers away with a speed of 1 – 2.5 m/s. Hence, it leads to a lot of unfortunate consequences. However, rip currents are often difficult to detect and recognize by naked-eye observation. This paper proposes a new method for detecting rip currents using a convolutional neural network and images collected by an unmanned aerial vehicle. The proposed system achieves an accuracy of 90.19 % on the testing set with the support of 15 frames per second. The results are a premise for building a real application in coastal monitoring systems to detect rip currents.

Keywords: coastal monitoring systems, CNN, rip currents, UAV.

1. ĐẶT VẤN ĐỀ

Dòng nước chảy xa bờ (RC – Rip current) là dòng nước mạnh được tạo thành theo hướng từ bờ hướng ra biển. Thông qua sóng biển, gió và địa chất của bờ biển, RC có thể xuất hiện ở bất kỳ vùng bờ biển nào (Sinnott & Feddersen, 2014). Hình 1 thể hiện một RC đã được sử dụng mục tím để làm nổi bật. Dòng chảy của RC kéo dài từ bờ ra xa gần 40 m.



Hình 1. Biểu diễn RC thực tế (Brander và cs., 2011)

Việt Nam có một vùng bờ biển dài trên 3200 km (không tính vùng bờ biển ở các đảo nhỏ). Vùng bờ biển kéo dài xuyên suốt 28 tỉnh, thành phố và là nơi sinh sống của gần 20 triệu người. Vùng bờ biển đóng vai trò quan trọng trong việc phát triển kinh tế của đất nước, đặc biệt trong phát triển dịch vụ du lịch. Trong việc phát triển du lịch biển, việc giảm thiểu nguy cơ, rủi ro gây ra bởi các dòng chảy xa bờ là cần thiết. Đặc biệt, gần đây, các phương tiện truyền thông liên tục truyền đi các thông tin về các vụ tai nạn gây ra bởi RC. Do vậy, việc xây dựng và phát triển hệ thống tự động phát hiện và cảnh báo về các RC trong khu vực là vấn đề nên được quan tâm.

Hơn một thập kỷ qua, mạng nơ-ron tích chập (CNN) đã và đang là một hướng nghiên cứu nổi bật trên thế giới trong lĩnh vực nhận dạng và phát hiện đối tượng. Từ sau khi có sự ra đời của mạng AlexNet (Krizhevsky và cs., 2012) cùng với các hệ thống phân cứng được phát triển bắt kịp với nhu cầu sử dụng. Các mô hình mạng CNN tiếp tục được nâng cấp và cải tiến để đạt được các kết quả vượt trội cả về độ chính xác và thời gian xử lý.

Trong bài toán nhận dạng đối tượng, nhiều mô hình tiêu biểu có thể được liệt kê như YOLOv4 (You Only Look Once version 4)

(Bochkovskiy và cs., 2020), SSD (Single Shot Detector) (W. Liu và cs., 2016), R-CNN (Region Based Convolutional Neural Network) (Girshick và cs., 2014), Fast R-CNN (Girshick, 2015), Faster R-CNN (Ren và cs., 2015).

Một cách đơn giản, để có thể thực hiện một dự đoán từ các ảnh/video đầu vào, các ảnh/video sẽ được chuyển qua mạng CNN đã được huấn luyện để thực hiện các phép biến đổi tích chập qua các lớp tích chập khác nhau để thu được các trọng số trên các bản đồ đặc trưng (feature maps). Các đặc trưng này sau đó sẽ được tiếp tục thực hiện với các phép biến đổi phi tuyến nhằm thu được các đặc trưng cốt lõi nhằm mục đích phân loại và nhận dạng đối tượng.

Trong nghiên cứu này, chúng tôi tiếp tục khai thác mạng CNN để thực hiện bài toán nhận dạng RC từ ảnh chụp bằng các thiết bị bay không người lái (UAV). Đóng góp của bài báo cụ thể là:

- Thảo luận các nghiên cứu gần thực hiện bài toán nhận dạng RC từ ảnh sử dụng trí tuệ nhân tạo (AI) hoặc sâu hơn là CNN.
- Đề xuất cách tiếp cận mới trong việc nhận dạng RC từ ảnh thu được bởi UAV với độ chính xác cao.

2. PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Các nghiên cứu liên quan



Hình 2. Ảnh chụp bờ biển Sầm Sơn thu được bởi UAV của tác giả

UAV đã trở nên phổ biến bởi nhiều ưu điểm như: dễ dàng thực hiện các thao tác chụp tại những địa điểm mà con người khó thao tác; độ phân giải của ảnh chụp cao; có thể bay theo định kỳ để thu thập dữ liệu. Hình ảnh thu được mang tính bao phủ cao và trực quan, như được biểu diễn ở Hình 2.

Nhiều bộ dữ liệu chuẩn cho bài toán nhận dạng đối tượng với ảnh/video thu thập được bởi UAV đã được công bố (Du và cs., 2018; Yu và cs., 2020) cho các hệ thống nhận dạng đối tượng nói chung như: phương tiện giao thông, đường phố, sân ga, v.v..

Trong gần mười năm trở lại đây, nhận dạng các RC trong ảnh/video đã được tập trung nghiên cứu nhiều hơn trong cộng đồng học thuật bởi những ảnh hưởng của các hiện tượng RC tới cuộc sống. Một số nghiên cứu điển hình sẽ được chỉ ra dưới đây.

Nghiên cứu sử dụng thuốc nhuộm để biểu diễn trực quan các dòng RC nhằm huấn luyện cho người bơi cách thoát ra khỏi RC (Brander và cs., 2011) mặc dù khó thực hiện và triển khai nhưng lại là một cách tiếp cận tuyệt vời giúp cho người xem có thể hình dung được cấu tạo, quy mô và diễn biến của các RC tại bờ biển.

Các nghiên cứu sử dụng các công cụ phần cứng như cảm biến sóng, máy đo giao động sóng âm (Elgar và cs., 2001), hoặc các module nổi sử dụng các bộ cảm biến và định vị GPS để đo đạc và định vị các RC (Castelle và cs., 2014) có thể đánh giá và ước lượng các RC ở độ chính xác cao. Tuy nhiên, một số nhược điểm gặp phải như: chi phí và giá thành cao; tốn thời gian và vận hành bởi các chuyên gia; phải đo đạc hoá địa hình một các kỹ lưỡng mới có thể triển khai điều đó ảnh hưởng tới việc nhân rộng của các mô hình.

Cùng với sự phát triển của ngành khoa học thị giác máy tính (computer vision) và các nền tảng phần cứng mạnh mẽ, nhiều nghiên cứu gần đây đã được thực hiện để xác định các RC trong ảnh và video. Mori và cs. (2022) đã đề xuất sử dụng kỹ thuật phân tích luồng quang (optical flow) để phát hiện các RC trong video. Kết hợp các kỹ thuật trực quan hoá sự di chuyển của các luồng RC, phương

pháp đạt độ ổn định không những về độ chính xác trong nhận dạng mà còn về độ rõ ràng và trực quan.

Rashid và cs. (2021) đề xuất mô hình RipNet sử dụng mô hình tích chập tự giải mã (autoencoder-based method) với mạng cơ sở (base model) sử dụng mô hình Tiny-YOLOv3 được tiền huấn luyện trên mô hình COCO2017.

RC có thể tồn tại ở nhiều hình thù khác nhau do cấu tạo của địa tầng tại mỗi vùng biển nhất định. Một vài loại RC xuất hiện với màu sắc tương đối đặc trưng của lớp bùn trầm tích của đáy biển (sediment plumes. Chính nhờ điều này, Y. Liu & Wu (2019) đã sử dụng mô hình không gian màu HSV (Hue, Saturation, Value) để nhận dạng RC. Phương pháp này chỉ đạt hiệu quả cao trong một số trường hợp cụ thể chứ không mang tính tổng quát cao.

Nghiên cứu của Maryan và cs. (2019) nỗ lực tìm ra các đặc trưng chính của các dòng RC thông qua việc sử dụng các đặc trưng Haar (Haar features) và đề xuất kết hợp đặc trưng mới tìm được của RC với 5 đặc trưng khác của Haar cổ điển. Sau đó sử dụng một số thuật toán học máy (machine learning) để huấn luyện và phân loại các RC trong ảnh. Một số phương pháp học máy đã được sử dụng như: Cây quyết định (Decision tree), Rừng ngẫu nhiên (Random forest), SVM, mạng nơron (Neural network). Tuy nhiên, đặc trưng Haar mới mà nghiên cứu đề xuất là tương đối nhạy cảm với các điều kiện khác nhau về ảnh thu nhận được từ môi trường.

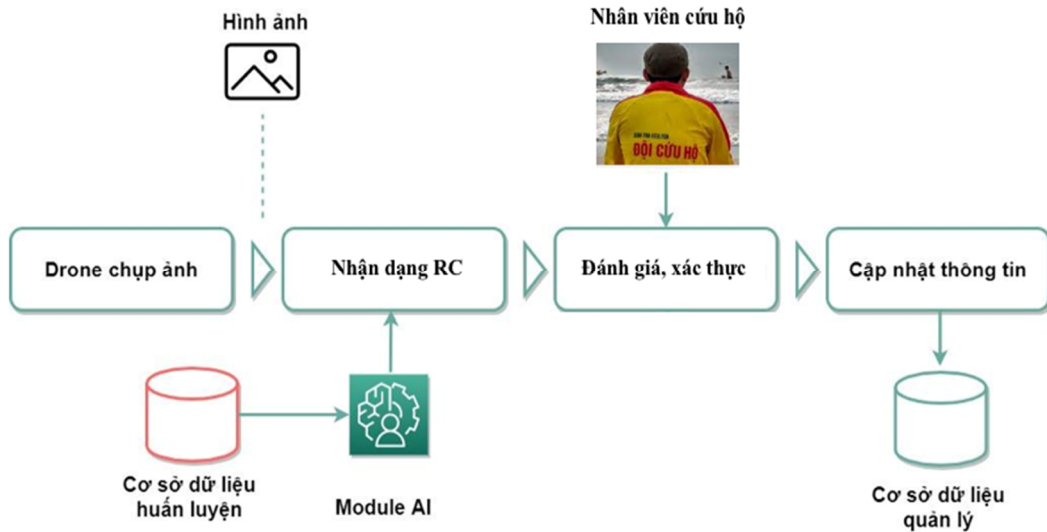
Luận án của Sun (2020) sử dụng dữ liệu thu được từ các thiết bị UAV sau đó áp dụng phương pháp phân tích luồng quang (optical flow) thông qua các thuật toán Lucas-Kanade, HOR Lucas-Kanada, và Horn-Schunck để xác định sự chuyển động bất thường của các điểm trên bề mặt nước biển của vùng chứa RC và các vùng nước biển thông thường. Tuy nhiên, phương pháp này đạt độ chính xác tương đối thấp, chỉ dưới 70%. Ngoài ra, giải pháp này được đánh giá là ít có khả năng áp dụng rộng rãi trên các vùng biển khác nhau.

Nhằm cung cấp thêm tính đa dạng trong nghiên cứu về các mô hình nhận dạng RC từ ảnh, trong nghiên cứu này, chúng tôi đề xuất phương án tiếp cận mới giải quyết bài toán về nhận dạng RC sử dụng UAV và xây dựng giải pháp chuyển đổi số cho công tác đảm bảo an toàn tại các bãi biển du lịch. Một thực tế không thể phủ nhận khi tìm hiểu thực trạng các bãi tắm hiện tại ở Việt Nam, phần lớn

công tác đảm bảo an toàn, phát hiện các nguy cơ, rủi ro là dựa vào lực lượng nhân viên bờ biển quan sát bằng mắt thường.

2.2. Đề xuất cách tiếp cận mới

Trong nghiên cứu này, chúng tôi đề xuất giải pháp nhận dạng các RC tại bờ biển một cách tự động thông việc sử dụng các thiết bị UAV. Mô hình đề xuất được biểu diễn ở Hình 3.



Hình 3. Mô hình đề xuất

Đầu tiên, thiết bị UAV mà chúng tôi đề xuất sử dụng là dòng DJI Mavic¹ với một số ưu thế: có thể truyền tín hiệu hình ảnh thu thập được trực tiếp tới đa tới 15 km về bộ xử lý trung tâm (server); hiệu năng sử dụng pin cao và có thể bay trong phạm vi 30 km, điều này phù hợp trong việc giám sát các bãi tắm. Nhiệm vụ của thiết bị UAV là bay theo định kỳ để giám sát các khu vực tắm và được cài đặt thời gian bay bởi người quản lý cho phù hợp dựa vào tình hình thực tế.

Thiết bị UAV sẽ truyền hình ảnh thu thập về máy chủ một cách trực tiếp và tại đây, chúng tôi đề xuất sử dụng mô hình mạng YOLOv4 (Bochkovskiy và cs., 2020). Tuy nhiên, việc xử lý liên tục các video với độ phân giải cao được truyền theo thời gian thực về máy chủ gây một áp lực lớn lên hiệu năng vận hành của hệ thống. Do vậy, chúng tôi đề xuất giải pháp chỉ chọn 5 khung hình trên mỗi

giây của video để thực hiện quá trình nhận dạng. Đây là một kỹ thuật thường được sử dụng khi xử lý các video có độ phân giải cao. Thông thường các video này có tốc độ khoảng 25 – 35 khung hình/giây. Việc lấy mẫu theo bước nhảy là 5 giảm sự quá tải trong quá trình xử lý trong khi vẫn đảm bảo về nội dung của các khung hình.

YOLOv4 là phiên bản cải tiến thứ tư của mô hình mạng nơ-ron tích chập được đề xuất bởi nhóm nghiên cứu từ Google. Phiên bản được biết tới với độ chính xác và thời gian xử lý nhanh. Cấu trúc của YOLOv4 được chia làm bốn phần chính: backbone (xương sống), neck (cổ), dense prediction (dự đoán dày đặc) – sử dụng các mô hình một trạng thái như các phiên bản YOLO cũ, spare prediction (dự đoán thưa thớt) – sử dụng các mô hình hai trạng thái như R-CNN (Girshick, 2015).

¹ <https://www.dji.com/products/mavic>

Ngoài các cải tiến về mô hình dự đoán, nhằm nâng cao tính đa dạng trong huấn luyện mô hình, YOLOv4 sử dụng rất nhiều các cải tiến về tăng cường dữ liệu như: cắt và trộn (cutout/mixout), kỹ thuật Mosaic kết hợp 4 ảnh đầu vào vào một ảnh nhằm tạo sự phong phú cho nội dung ảnh. Mặc dù đã có một số phiên bản nâng cấp của mô hình YOLO như YOLOv5, YOLOv7 tuy nhiên mô hình YOLOv4 vẫn được đánh giá cao bởi tính ổn định vào hiệu quả cao trong bài toán nhận dạng.

Sau khi sử dụng mô hình YOLOv4 để nhận dạng việc có hay không xuất hiện của RC tại khung hình hiện tại, do tính chất quan trọng của việc phát hiện các RC nên cần một quản trị viên là nhân viên cứu hộ xác thực lại các cảnh báo để tránh việc mô hình đưa ra các cảnh báo sai (false alarm).

Trong trường hợp phát sinh các RC trong khung hình gửi về bởi thiết bị bay UAV, các biện pháp nghiệp vụ sẽ được đưa ra bởi lực lượng giám sát bờ biển nhằm tối thiểu hoá nguy cơ gây ra bởi các RC. Ngoài ra, các sự kiện này sẽ được cập nhật lại trong hệ thống quản lý nhằm phục vụ cho công tác quản lý và đánh giá trong cơ sở dữ liệu. Trong phạm vi nghiên cứu của bài toán, trong phần tiếp theo, chúng tôi sẽ đánh giá hiệu năng của mô hình với việc dành chủ yếu cho mô hình YOLOv4 trong nhận dạng các dòng RC từ ảnh.

Đóng góp của nghiên cứu là đề xuất quy trình nhận dạng các RC từ ảnh được thu thập bởi UAV. Chúng tôi đánh giá tính khả thi của việc sử dụng mô hình YOLOv4 trong các kịch bản này.

Ngoài ra, để tăng cường độ chính xác của việc huấn luyện mô hình, chúng tôi đề xuất sử dụng các kỹ thuật tăng cường dữ liệu trong tập dữ liệu có sẵn và dữ liệu tự thu thập, gắn nhãn. Chi tiết được trình bày ở mục 3.1.

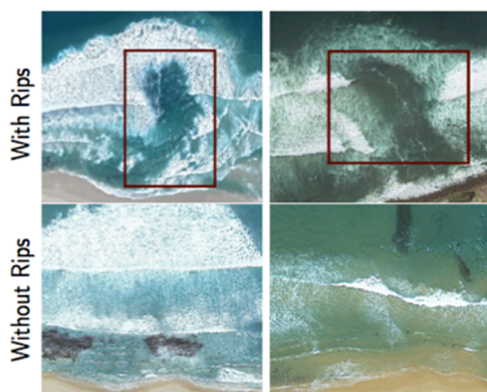
3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Dữ liệu

Để chuẩn bị dữ liệu cho việc huấn luyện mô hình, chúng tôi đã sử dụng hai giải pháp:

- Sử dụng các ảnh thu thập một cách thủ công từ các công cụ tìm kiếm trên website để tăng cường dữ liệu huấn luyện mô hình. Thông qua sử dụng các công cụ tìm kiếm, chúng tôi đã thu thập được hơn 200 ảnh biểu diễn các RC. Sau đó, sử dụng công cụ gắn nhãn LabelImg² để khoanh vùng của các RC trong ảnh.

- Sử dụng bộ dữ liệu chuẩn được cung cấp bởi (de Silva, 2021) bao gồm 1740 ảnh biểu diễn RC, 700 ảnh không phải là các RC các ảnh có độ phân giải dao động từ 1086×916 tới 234×234 pixels là các ảnh chụp trên không (aerial images) do đó phù hợp với ngữ cảnh và kích bản trong nghiên cứu này. Một số hình ảnh trong bộ dữ liệu được trình bày lại ở Hình 4.



Hình 4. Minh họa ảnh trong bộ dữ liệu (de Silva và cs., 2021)

Bộ dữ liệu chúng tôi tự thu thập sẽ được sử dụng để tiền huấn luyện mô hình. Sau đó, bộ dữ liệu của de Silva và cs. (2021) sẽ được cấu hình theo tỉ lệ 50: 10: 40 tương ứng với dữ liệu huấn luyện, dữ liệu xác thực, dữ liệu kiểm tra.

3.2. Giao thức và thang đo

Để thực hiện huấn luyện cho mô hình, chúng tôi sử dụng các dữ liệu được mô tả ở mục 3.1. Cấu hình file huấn luyện với một số

² <https://github.com/heartexlabs/labelImg>

tham số chính như sau:

- Kích thước ảnh: 608×608
- Số lượng kênh: 3
- Momentum: 0.949
- Learning rate: 0.01
- Batch_size: 32
- Epochs: 250

Chúng tôi sử dụng Google Colab để huấn luyện mô hình bởi đây là môi trường được cung cấp hoàn toàn miễn phí của Google với cấu hình máy chủ đảm bảo cho việc huấn luyện các mô hình CNN.

Để thực hiện việc đánh giá mô hình, chúng tôi sử dụng thang đo là mAP (mean average precision), thang đo đánh giá độ chính xác chung của mô hình nhận dạng. Nếu chỉ số này càng cao chứng tỏ hiệu quả làm việc của mô hình trong nhận dạng càng tốt. Công thức mô tả ở phương trình

$$mAP = \frac{\sum_{i=1}^n P_i(R_i) dR_i}{N}$$

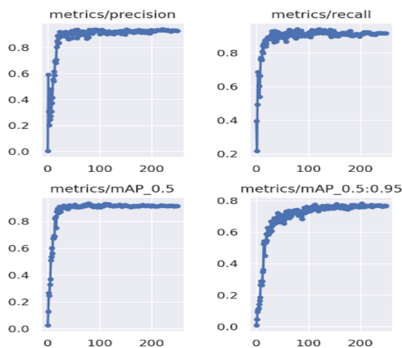
$$P = \frac{TP}{TP + FP}$$

$$N = \frac{TP}{TP + FN}$$

trong đó TP là số lượng các mẫu nhận dạng đúng đối tượng, FP là số lượng các mẫu nhận dạng nhầm đối tượng, FN là số lượng các mẫu bị bỏ sót trong quá trình nhận dạng, N là tổng số lượng mẫu nhận dạng.

3.3. Kết quả thực nghiệm

Kết quả của quá trình huấn luyện của mô hình được biểu diễn ở Hình 5.



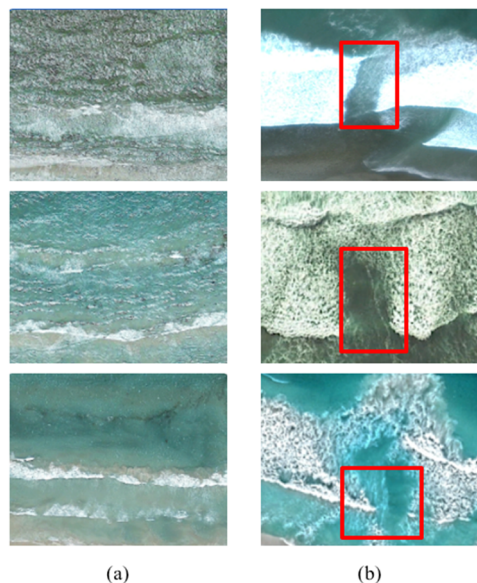
Hình 1. Biểu diễn quá trình huấn luyện mô hình

Nhằm tăng tính thuyết phục cho mô hình đề xuất sử dụng, chúng tôi sử dụng thêm một số mô hình YOLOv3 (Redmon & Farhadi, 2018), Fast R-CNN (Girshick, 2015).

Bảng 1: Kết quả đánh giá của mô hình đề xuất sử dụng so với mô hình khác

Mô hình	Độ chính xác trung bình mAP (%)
Mô hình đề xuất YOLOv4	90.19 %
YOLOv3	83.76 %
Fast R-CNN	86.22 %

Bảng 1 thể hiện kết quả của việc đánh giá mô hình YOLOv4 với các mô hình khác. Mô hình đề xuất hoạt động trên tập kiểm tra với mAP (%) là 90.19%, tốt hơn so với mô hình Fast R-CNN gần 4% và với YOLOv3 là gần 7%. Cũng lưu ý thêm, độ chính xác trung bình mAP được tính tại giá trị lấy độ đo IoU³ (Intersection over Union) bằng 0.5. Một số kết quả trực quan của quá trình nhận dạng được nhìn rõ ở Hình 6. Hình 6 (a) minh họa cho các ảnh không chứa RC, trong khi đó Hình 6 (b) biểu diễn cho các RC được nhận dạng bởi mô hình và biểu diễn bằng các hình chữ nhật đỏ bao quanh.



Hình 2. Hình ảnh biểu diễn kết quả nhận dạng của mô hình

³ shorturl.at/drvyz

Để triển khai mô hình lên hệ thống của chúng tôi với hệ thống máy chủ có cấu hình Intel(R) Core (TM) i5-10400F CPU @ 2.90 GHz có RAM 16 GB, 4 x NVIDIA T4 GPU with 16GB, chúng tôi đề xuất mô hình hoạt động với tốc độ gần 15 FPS (frame per second). Sau khi nhúng mô hình vào hệ thống hoàn chỉnh của chúng tôi, tốc độ nhận dạng của mô hình là xấp xỉ 3.5 FPS. Điều này vẫn đảm bảo tốc độ xử lý kịp thời cho các cảnh báo nếu xuất hiện các RC thông qua hình ảnh gửi về từ các UAV.

4. KẾT LUẬN VÀ HƯỚNG NGHIÊN CỨU MỚI

Trong nghiên cứu này, chúng tôi đã đề xuất mô hình hoạt động tự động cho bài toán nhận dạng RC xảy ra ở các bãi biển. Trong đó, chúng tôi sử dụng mô hình mạng YOLOv4 để thực hiện nhiệm vụ nhận dạng RC xuất hiện trong các hình ảnh được gửi về từ thiết bị UAV. Mô hình đạt độ chính xác 90.19% trong tập dữ liệu kiểm tra với tốc độ gần 15 FPS và đạt tốc độ 3.5 FPS khi triển khai trên toàn bộ hệ thống.

Do điều kiện thời gian eo hẹp, trong nghiên cứu này, chúng tôi chưa thu thập được nhiều dữ liệu thực tế tại các bãi biển ở Việt Nam. Do đó, trong tương lai gần, chúng tôi sẽ tiếp tục xây dựng bộ dữ liệu chuẩn cho các bãi biển ở Việt Nam nhằm củng cố độ chính xác của mô hình phát hiện các RC để hệ thống có thể triển khai và lắp đặt thực tế.

Một hướng nghiên cứu tiếp theo là xây dựng các module sử dụng kỹ thuật thực tế ảo (virtual reality) để tạo ra các kịch bản nhằm mục đích giáo dục thanh thiếu niên về RC và cách tự xử lý khi rơi vào trong vùng của RC.

TÀI LIỆU THAM KHẢO

Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv: Computer Vision and Pattern Recognition*. <https://www.scinapse.io/papers/3018757597>

Brander, R., Bradstreet, A., Sherker, S., & MacMahan, J. (2011). Responses of Swimmers Caught in Rip Currents: Perspectives on Mitigating the Global Rip Current Hazard. *International Journal of Aquatic Research and Education*, 5(4). <https://doi.org/10.25035/ijare.05.04.11>

Castelle, B., Almar, R., Dorel, M., Lefebvre, J.-P., Senechal, N., Anthony, E. J., Laibi, R., Chuchla, R., & Penhoat, Y. du. (2014). Rip currents and circulation on a high-energy low-tide-terraced beach (Grand Popo, Benin, West Africa). *Journal of Coastal Research*, 70(sp1), 633–638. <https://doi.org/10.2112/SI70-107.1>

de Silva, A., Mori, I., Dusek, G., Davis, J., & Pang, A. (2021). Automated rip current detection with region based convolutional neural networks. *Coastal Engineering*, 166, 103859. <https://doi.org/10.1016/j.coastaleng.2021.103859>

Du, D., Qi, Y., Yu, H., Yang, Y., Duan, K., Li, G., Zhang, W., Huang, Q., & Tian, Q. (2018). The Unmanned Aerial Vehicle Benchmark: Object Detection and Tracking. Trong V. Ferrari, M. Hebert, C. Sminchisescu, & Y. Weiss (B.t.v), *Computer Vision – ECCV 2018* (Vol 11214, tr 375–391). Springer International Publishing. https://doi.org/10.1007/978-3-030-01249-6_23

Elgar, S., Raubenheimer, B., & Guza, R. T. (2001). Current Meter Performance in the Surf Zone. *Journal of Atmospheric and Oceanic Technology*, 18(10), 1735–1746. [https://doi.org/10.1175/15200426\(2001\)018<1735:CMPITS>2.0.CO;2](https://doi.org/10.1175/15200426(2001)018<1735:CMPITS>2.0.CO;2)

Girshick, R. (2015). Fast R-CNN. *2015 IEEE International Conference on Computer Vision (ICCV)*, 1440–1448. <https://doi.org/10.1109/ICCV.2015.169>

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. *2014 IEEE Conference on Computer*

- Vision and Pattern Recognition*, 580–587. <https://doi.org/10.1109/CVPR.2014.81>
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, 1097–1105.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016). SSD: Single Shot MultiBox Detector. Trong B. Leibe, J. Matas, N. Sebe, & M. Welling (B.t.v), *Computer Vision – ECCV 2016* (tr 21–37). Springer International Publishing. https://doi.org/10.1007/978-3-319-46448-0_2
- Liu, Y., & Wu, C. H. (2019). Lifeguarding Operational Camera Kiosk System (LOCKS) for flash rip warning: Development and application. *Coastal Engineering*, 152, 103537. <https://doi.org/10.1016/j.coastaleng.2019.103537>
- Maryan, C., Hoque, M. T., Michael, C., Ioup, E., & Abdelguerfi, M. (2019). Machine learning applications in detecting rip channels from images. *Applied Soft Computing*, 78, 84–93. <https://doi.org/10.1016/j.asoc.2019.02.017>
- Mori, I., de Silva, A., Dusek, G., Davis, J., & Pang, A. (2022). Flow-Based Rip Current Detection and Visualization. *IEEE Access*, 10, 6483–6495. <https://doi.org/10.1109/ACCESS.2022.3140340>
- Rashid, A. H., Razzak, I., Tanveer, M., & Robles-Kelly, A. (2021). RipDet: A Fast and Lightweight Deep Neural Network for Rip Currents Detection. *2021 International Joint Conference on Neural Networks (IJCNN)*, 1–6. <https://doi.org/10.1109/IJCNN52387.2021.9533849>
- Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement (arXiv:1804.02767). arXiv. <https://doi.org/10.48550/arXiv.1804.02767>
- Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *Advances in Neural Information Processing Systems*, 28.
- Sinnett, G., & Feddersen, F. (2014). The surf zone heat budget: The effect of wave heating. *Geophysical Research Letters*, 41(20), 7217–7226. <https://doi.org/10.1002/2014GL061398>
- Sun, A. (2020). *UAV-Video-Based Rip Current Detection in Near-Shore Areas—University of Miami* [Master of Science (MS), Thesis, University of Miami].
- Yu, H., Li, G., Zhang, W., Huang, Q., Du, D., Tian, Q., & Sebe, N. (2020). The Unmanned Aerial Vehicle Benchmark: Object Detection, Tracking and Baseline. *International Journal of Computer Vision*, 128(5), 1141–1159. <https://doi.org/10.1007/s11263-019-01266-1>