

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ KỸ THUẬT TP.HCM

NGÔ ĐỨC ĐẠT

NHẬN DẠNG VÀ PHÂN LOẠI TÀU TRONG CẢNH
GIỚI BỜ BIỂN SỬ DỤNG TRÍ TUỆ NHÂN TẠO

NGÀNH: KỸ THUẬT ĐIỆN TỬ

TÓM TẮT LUẬN ÁN TIẾN SĨ

TP.HCM, Ngày 20 tháng 1 năm 2026

DANH MỤC CÔNG TRÌNH ĐÃ CÔNG BỐ

- Ngo, D.-D., Vo, V.-L., Nguyen, T., Nguyen, M.-H., & Le, M.-H. (2023). Image-Based Ship Detection Using Deep Variational Information Bottleneck. *Sensors*, 23(19), 8093. <https://doi.org/10.3390/s2319>
- Ngo, D.-D., Vo, V.-L., Le, M.-H., Phan, H., & Nguyen, M. H. (2024). Transformer based ship detector: An improvement on feature map and tiny training set. *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, 12(1). <https://doi.org/10.4108/eetinis.v12i1.6794>
- Ngo, D.-D., Nguyen, M.-H., Nguyen, Q.-T.-D., & Le, M.-H. (2021). Clustering based ship classification using radar signal and neuron network. In *Proceedings of the 2021 International Conference on System Science and Engineering (ICSSE)* (pp. 122-127). IEEE. <https://doi.org/10.1109/ICSSE52>
- Ngo, D.-D., Phan, V.-H.-A., Pham, H.-T., Be, T.-T., Nguyen, V.-B., & Le, M.-H. (2023). A vision-based container-code checking system: Case study at international terminal. In *Proceedings of the 2023 International Workshop on Intelligent Systems (IWIS)* (pp. 1–6). IEEE.

<https://doi.org/10.1109/IWIS58789.2023.10284525>

MỞ ĐẦU

1. Lý do chọn đề tài

Trong công tác giám sát và bảo vệ chủ quyền biển đảo, việc phát hiện, nhận dạng và phân loại mục tiêu tàu biển có ý nghĩa đặc biệt quan trọng, nhất là đối với lực lượng Hải quân và các hệ thống radar cảnh giới ven biển. Hiện nay, các hệ thống giám sát bờ biển chủ yếu dựa trên hai loại cảm biến chính là radar và camera giám sát.

Radar đóng vai trò chủ lực trong việc phát hiện và theo dõi mục tiêu ở cự ly xa, hoạt động ổn định trong mọi điều kiện thời tiết và thời gian ngày đêm. Radar cung cấp các thông tin quan trọng như vị trí, tốc độ và hướng di chuyển của mục tiêu, cho phép phát hiện sớm các nguy cơ va chạm, xâm nhập trái phép hoặc hoạt động bất thường trong vùng biển quản lý. Tuy nhiên, tín hiệu radar chỉ phản ánh đặc tính điện từ của mục tiêu, không cung cấp hình ảnh trực quan, do đó gặp nhiều khó khăn trong việc phân biệt chính xác loại tàu và mục đích hoạt động.

Ngược lại, hệ thống camera giám sát, bao gồm camera quang học và camera hồng ngoại, cung cấp hình ảnh trực quan giúp nhận dạng chi tiết hình dáng, kích thước, cờ hiệu và hành vi của tàu thuyền. Camera ngày càng được triển khai rộng rãi tại các trạm radar và trung tâm quan sát ven biển. Tuy nhiên, camera lại bị hạn chế về cự ly quan sát, chịu ảnh hưởng mạnh của điều kiện thời tiết, ánh sáng và tầm nhìn.

Trước những hạn chế riêng lẻ của từng loại cảm biến, xu hướng kết hợp radar và camera trong giám sát bờ biển đã trở nên tất yếu. Radar thực hiện nhiệm vụ phát hiện và chỉ thị mục tiêu từ xa, trong khi camera đảm nhiệm vai trò ghi hình và phân tích chi tiết. Mặc dù vậy, hiện nay việc nhận dạng và phân loại tàu trong các hệ thống radar – camera vẫn chủ yếu dựa vào kinh nghiệm của nhân viên vận hành, dẫn đến tính chủ quan cao và khó đáp ứng yêu cầu giám sát trong môi trường phức tạp, mật độ tàu thuyền lớn.

Vì vậy, việc nghiên cứu ứng dụng trí tuệ nhân tạo nhằm tự động hóa quá trình nhận dạng và phân loại tàu biển dựa trên dữ liệu radar và ảnh camera là yêu cầu cấp thiết, góp phần nâng cao hiệu quả giám sát bờ biển, bảo đảm an toàn hàng hải và bảo vệ chủ quyền quốc gia.

2. Mục đích nghiên cứu

Mục đích của đề tài là nghiên cứu, phát triển phương pháp ứng dụng trí tuệ nhân tạo trong việc nhận dạng và phân loại mục tiêu tàu biển dựa trên tín hiệu phản xạ radar và dữ liệu ảnh từ camera giám sát, nhằm nâng cao độ chính xác, tính khách quan và mức độ tự động hóa trong công tác cảnh giới bờ biển.

3. Nhiệm vụ nghiên cứu

Để đạt được mục đích đề ra, đề tài tập trung vào việc thực hiện các nhiệm vụ nghiên cứu sau:

- Nghiên cứu và xử lý dữ liệu ảnh từ hệ thống camera giám sát phục vụ nhận dạng trực quan và hỗ trợ phân loại mục tiêu tàu biển.
- Phân tích đặc điểm tín hiệu phản xạ radar của các loại tàu biển hoạt động trong vùng biển Việt Nam.

- Xây dựng và đánh giá các mô hình trí tuệ nhân tạo phục vụ phân loại tàu biển dựa trên tín hiệu radar.
- Xây dựng và đánh giá các mô hình trí tuệ nhân tạo phục vụ phân loại tàu biển dựa trên tín hiệu ảnh.

4. Phạm vi nghiên cứu

Phạm vi nghiên cứu của đề tài tập trung vào việc khai thác tín hiệu phản xạ radar và dữ liệu ảnh camera trong điều kiện mô phỏng và thực nghiệm tại một số khu vực ven biển tiêu biểu. Nghiên cứu chủ yếu dừng ở mức phát triển, huấn luyện và đánh giá các mô hình trí tuệ nhân tạo ở dạng nguyên mẫu (prototype), chưa đi sâu vào triển khai hạ tầng phần cứng trên diện rộng.

Đề tài tập trung vào các vấn đề kỹ thuật liên quan đến xử lý tín hiệu radar, xử lý ảnh và hiệu quả thuật toán, không đi sâu vào các khía cạnh pháp lý, chính sách quản lý hay hợp tác quốc tế trong lĩnh vực giám sát biển.

5. Hướng tiếp cận và phương pháp nghiên cứu

5.a Hướng tiếp cận

Đề tài tiếp cận theo hướng kết hợp đa nguồn dữ liệu từ radar và camera giám sát nhằm tận dụng ưu điểm của từng loại cảm biến. Radar đảm nhiệm vai trò phát hiện và theo dõi mục tiêu ở cự ly xa trong mọi điều kiện thời tiết, trong khi camera cung cấp hình ảnh trực quan hỗ trợ nhận dạng chi tiết. Trên cơ sở đó, đề tài ứng dụng các kỹ thuật trí tuệ nhân tạo, học máy và học sâu để tự động trích xuất đặc trưng và phân loại mục tiêu tàu biển, giảm sự phụ thuộc vào kinh nghiệm chủ quan của nhân viên vận hành.

5.b Phương pháp nghiên cứu

Đối với dữ liệu ảnh camera: áp dụng các mô hình học sâu trong bài toán phát hiện và nhận dạng đối tượng, sử dụng các tập dữ liệu chuẩn về tàu biển để huấn luyện và đánh giá mô hình.

Đối với dữ liệu radar: thực hiện tiền xử lý, lọc nhiễu, trích xuất đặc trưng tín hiệu phản xạ, kết hợp các phương pháp phân cụm và phân loại để xây dựng mô hình nhận dạng tàu biển từ tín hiệu radar.

Đánh giá hệ thống dựa trên các tiêu chí: độ chính xác nhận dạng, khả năng phân loại, tốc độ xử lý và độ ổn định trong môi trường biển phức tạp.

CHƯƠNG 1: TỔNG QUAN

Trong chương này nghiên cứu sinh phân tích, tổng hợp, so sánh các công bố liên quan tới giám sát bờ biển trong những năm gần đây. Nhận định các vấn đề còn tồn tại, đồng thời xác định các vấn đề còn có thể phát triển và đề xuất các hướng nghiên cứu của luận án.

CHƯƠNG 2: CƠ SỞ LÝ THUYẾT

Trong chương này, nghiên cứu sinh trình bày ngắn gọn cơ sở lý thuyết về các mô hình phát hiện đối

tượng bằng cảm biến ảnh, các kỹ thuật phân nhóm tín hiệu và phân loại tín hiệu.

CHƯƠNG 3: PHÁT HIỆN TÀU BIỂN TỪ TÍN HIỆU CAMERA

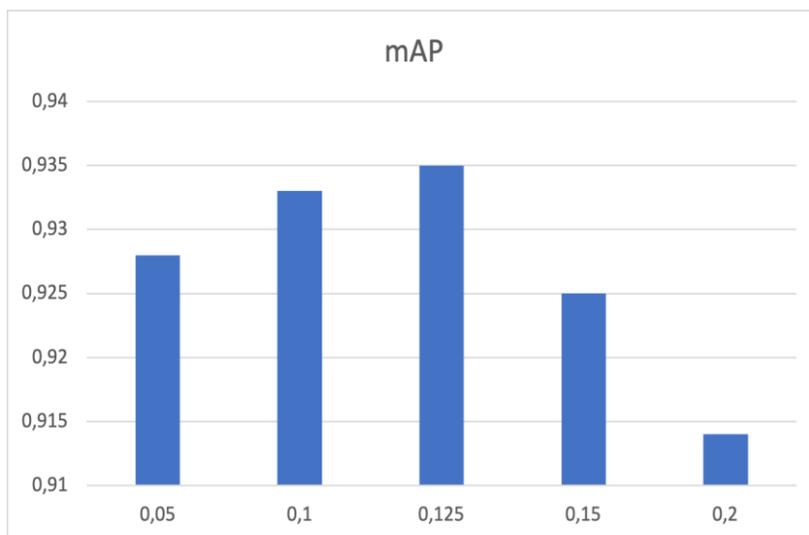
3.1 Bộ dữ liệu và ngữ cảnh thử nghiệm

Để đánh giá chất lượng cho phương pháp đề nghị, nghiên cứu này sử dụng bộ dữ liệu SeaShips [1]. Tổng quan tài liệu cho thấy nhiều công trình nghiên cứu sử dụng 80% dữ liệu đã công bố để huấn luyện/xác nhận và 20% để thử nghiệm. Do đó, chúng tôi chọn D^{Train}_1 bao gồm 5600 hình ảnh để huấn luyện và D^{Test}_1 bao gồm 1400 hình ảnh để thử nghiệm. Ngoài ra, những công trình gần đây [2, 3] cũng sử dụng các ngữ cảnh thách thức hơn, trong đó 50% dữ liệu là tập huấn luyện và phần còn lại là tập kiểm tra. Chúng tôi cũng tuân theo ngữ cảnh này để chuẩn bị D^{Train}_2 và D^{Test}_2 để so sánh. Để đánh giá hiệu suất trên một tập dữ liệu rất nhỏ, chúng tôi chọn ngẫu nhiên một số tập dữ liệu con S_1, S_2, S_3 bao gồm 30%, 70% và 100% mẫu từ D^{Train}_2 để huấn luyện về các thí nghiệm sau này.

3.2 Kết quả thực nghiệm với mạng YoLoX

3.2.1 Chọn siêu tham số α_{KL}

Phần này thảo luận về việc chọn một siêu tham số phù hợp α_{KL} trong phương trình tính loss. Kết quả được mô tả như trong hình 3.1. Kết quả cho thấy giá trị α_{KL} tối ưu là 0.125.



Hình 3.1: Giá trị mAP ứng với các tham số α_{VIB} khác nhau. Trục x có nghĩa mô tả tham số α_{KL} , và trục y mô tả độ chính xác trung bình trung bình trên tất cả các lớp.

Khi không sử dụng kỹ thuật lựa chọn đặc trưng, giá trị mAP chỉ đạt được 0.923; khi α_{KL} bằng 0.05, giá trị mAP tăng lên 0,928. Khi α_{KL} nhận giá trị cao hơn, mAP nhận được ngày càng cao. Tuy nhiên, khi $\alpha_{KL}=0.15$, mAP bắt đầu giảm; và nếu $\alpha_{KL}=0.2$ mAP là 0,914. Giá trị mAP dựa trên $\alpha_{KL} = 0.2$ nhỏ hơn so với khi không sử dụng hàm mục tiêu VIB để lựa chọn đặc trưng. Hiện tượng này là do hàm mục tiêu lựa chọn đặc trưng làm sẽ làm giảm các đặc trưng được chọn cho tác vụ chính. Khi số lượng đặc trưng bị giảm quá nhiều, bộ phân loại phân loại có thể không có đủ thông tin cho nhiệm vụ phân loại. Trong các thí nghiệm tiếp theo, chúng tôi chọn $\alpha_{KL} = 0.125$ cho phương pháp đề xuất của chúng tôi.

3.2.2 So sánh với các phương án tốt nhất trong phát hiện tàu biển

Phần này so sánh phương pháp được đề xuất của chúng tôi với các phương pháp tốt nhất đã được công bố trên chỉ số mAP. Nhiều thiết lập khác nhau đã được sử dụng dựa trên các bài báo trước đó. Dựa trên 70000 hình ảnh được xuất bản từ tập dữ liệu SeaShip [1]; Zhang_2022 [4], và Zhang_2021 [5] đã sử dụng 90% dữ liệu để huấn luyện và xác thực mô hình; 10% dữ liệu còn lại là tập dữ liệu thử nghiệm. Liu_2020 [6], Liu_2022 [7], Han_2021[8], và Light_SDNet [9] sử dụng 80% dữ liệu để huấn luyện và xác thực tập dữ liệu, 20% còn lại là tập dữ liệu thử nghiệm. Để so sánh với các nghiên cứu này, chúng tôi sử dụng D^{Train}_1 để huấn luyện và D^{Test}_1 để kiểm tra. Kết quả trong bảng 3.1 cho thấy phương pháp của chúng tôi tốt hơn các phương pháp khác về độ chính xác trung bình.

Bảng 3.1: So sánh hiệu suất của các phương pháp khác nhau. Các kết quả tốt nhất được in đậm.

Method	Train+Val / Test (in %)	fishing boat	container ship	ore carrier	bulk cargo carrier	passenger ship	general cargo ship	mAP
Zhang_2022 [4]	90/10	0.824	0.940	0.859	0.915	0.787	0.914	0.873
Zhang_2021 [5]	90/10	-	-	-	-	-	-	0.946
Liu_2020 [6]	80/20	-	-	-	-	-	-	0.908
Liu_2022 [7]	80/20	-	-	-	-	-	-	0.964
Han_2021[8]	80/20	-	-	-	-	-	-	0.906
Light_SDNet [9]	80/20	0.986	0.995	0.989	0.990	0.982	0.989	0.988
Proposed method	80/20	0.979	1	0.987	0.994	0.994	0.993	0.991

Yani_2022 (ESDT) [2]	50/50	-	-	-	-	-	-	0.593
Yani_2022 (DETR)[2]	50/50	-	-	-	-	-	-	0.965
Biaohua_2022 [3]	50/50	0.940	0.987	0.966	0.978	0.937	0.972	0.963
Proposed method	50/50	0.970	0.986	0.984	0.991	0.964	0.989	0.98

3.3 Kết quả thí nghiệm với DETR

3.3.1 Lựa chọn siêu tham số

Như đã thảo luận trong Phần 3, số lượng truy vấn có thể ảnh hưởng lớn đến số lượng đầu ra của một bộ phát hiện dựa trên DETR. Do đó, phần này tập trung vào việc lựa chọn tham số phù hợp nhất để kiểm soát mô hình. Chúng tôi tiến hành so sánh hiệu suất khi $n_{queries}$ nhận các giá trị lần lượt trong danh sách [300, 200, 100, 50]. Các tập D^{Train}_2 và D^{Test}_2 được chọn làm tập huấn luyện và tập kiểm thử để đảm bảo một thiết lập bài toán mang tính thử thách cao. Điều này có nghĩa là 50% dữ liệu được sử dụng cho quá trình kiểm thử.

Các kết quả thực nghiệm được trình bày chi tiết trong Bảng 3.2. Trong thư viện mmdetection, giá trị mặc định cho $n_{queries}$ là 300. Quan sát cho thấy việc sử dụng thiết lập mặc định này dẫn đến sự gia tăng đáng kể số lượng phát hiện. Ví dụ, số lượng phát hiện đối với lớp fishing boat là 204.449. Xu hướng này dẫn đến việc giảm độ chính xác trung bình (AP) xuống 0.798, trong khi độ tăng lên 0.913. Ngoài ra, số lượng phát hiện cao hơn đối với tàu cá có thể là do tần suất xuất hiện cao hơn của nhãn này trong tập dữ liệu.

Bằng cách giảm $n_{queries}$, sự thiên lệch trong các phát hiện được giảm thiểu đáng kể. Cụ thể, khi $n_{queries}$ được giảm xuống các mức 200, 100 và 50, số lượng phát hiện tàu cá giảm xuống tương ứng là 208.817, 119.774 và 96.577. Hơn nữa, việc giảm tham số này giúp cân bằng số lượng phát hiện giữa các loại tàu khác nhau. Với $n_{queries} = 300$, số lượng phát hiện thấp nhất thuộc về tàu container (6300). Tuy nhiên, khi $n_{queries} = 50$, số lượng phát hiện thấp nhất tăng lên khoảng 18.949 và sự phân bố giữa tàu container, tàu chở quặng và tàu khách là khá tương đồng.

Bảng 3.2: So sánh hiệu suất của Deformable DETR với số lượng queries khác nhau. Kết quả tốt nhất được in đậm.

Class	300 query	200 query	100 query	50 query
-------	-----------	-----------	-----------	----------

	dets	recal	AP	dets	recal	AP	dets	recal	AP	dets	recal	AP
	l			l			l			l		
fishing boat	20444 9	0,91 3	0,79 8	20881 7	0,98 5	0,97 0	11977 4	0,98 6	0,96 9	9657 7	0,97 2	0,94 9
container ship	6300	0,98 9	0,89 4	11460	0,99 5	0,99 5	31083	0,99 5	0,98 9	1894 9	0,99 8	0,99 7
ore carrier	53520	0,98 7	0,92 3	49362	0,99 7	0,99 2	62122	0,99 6	0,98 9	1894 9	0,99 3	0,98 7
bulk cargo carrier	45541	0,98 5	0,91 2	47754	0,99 6	0,99 2	46490	0,99 7	0,98 6	7623 0	0,99 6	0,98 9
passenger ship	22498	0,96 8	0,61 0	16199	0,97 0	0,95 7	14866	0,96 8	0,93 6	2004 3	0,97 2	0,92 6
general cargo ship	17692	0,98 7	0,93 0	16408	0,99 5	0,99 2	75665	0,99 5	0,99 1	9853 7	0,99 6	0,99 0
mAP			0.84 4			0.98 2			0.97 7			0.97 3

3.3.2 So sánh với các phương pháp tiên tiến nhất (SoTA)

Mục này so sánh phương pháp đề xuất của chúng tôi với các phương pháp tiên tiến nhất (SoTA) dựa trên chỉ số mAP. Đối với mỗi phương pháp, chúng tôi sử dụng tập dữ liệu huấn luyện và kiểm thử tương ứng. Cụ thể, chúng tôi sử dụng D^{1}_{Train} và D^{1}_{Test} để huấn luyện mô hình của mình và so sánh với Zhang_2022, Zhang_2021 [10], Liu_2020, Liu_2022, Han_2021, SDNet_2022, và phương pháp dựa trên DETR. Ngoài ra, chúng tôi sử dụng D^{2}_{Train} và D^{2}_{Test} để huấn luyện một mô hình khác và so sánh với Biaohua_2022, Yani_2022, và phương pháp dựa trên DETR.

Bởi vì chỉ số mAP đạt cực đại khi $n_{queries} = 200$ đối với phương pháp của chúng tôi được thể hiện trong bảng 3.2, chúng tôi lựa chọn thiết lập này cho thí nghiệm này. Kết quả so sánh giữa các phương

pháp SoTA được trình bày trong Bảng 3.3 và một số kết luận có thể được rút ra như sau:

- Khung nền tảng (Baseline framework) là yếu tố then chốt để đạt được kết quả tốt hơn: Cui_2019 và Liu_2020 được phát triển dựa trên YoloV3. Do đó, hiệu năng của chúng không tốt bằng Liu_2022, phương pháp vốn dựa trên khung SSD. Han_2021 dựa trên YoloV4, nhưng hiệu năng không cho thấy sự cải thiện so với Liu_2020 (dựa trên YoloV3). Tập dụng ưu thế của YoloV5, SDNet_2022 đã cải thiện đáng kể so với

Liu_2020. Khung YoloV5 giúp chỉ số mAP tăng tới 8% so với khung YoloV3. phương pháp dựa trên DETR dựa trên YoLoX, còn phương pháp của chúng tôi dựa trên backbone DETR. Gần đây, các khung này đã trở thành những phương pháp ưu việt cho bài toán phát hiện đối tượng. Vì vậy, kết quả đạt được tốt hơn so với các phương pháp khác. Đáng chú ý là các nghiên cứu phát hiện tàu thường tận dụng một khung phát hiện đối tượng làm nền tảng, đi kèm với một số sửa đổi tùy chỉnh. Do đó, việc thừa hưởng các khả năng của một khung mới và mạnh mẽ như vậy tất yếu dẫn đến kết quả được cải thiện.

- Khi số lượng mẫu trong tập huấn luyện giảm, các phương pháp dựa trên DETR có xu hướng hoạt động hiệu quả hơn so với các phương pháp dựa trên CNN. Cụ thể, Trong bảng số liệu, Biaohua_2022 là mô hình phát hiện dựa trên CNN, còn Yani_2022 là mô hình dựa trên DETR. Chỉ số mAP của Yani_2022 và Biaohua_2022 lần lượt là 0,965 và 0,9963. Tuy nhiên, hiệu năng này có thể được nâng cao nếu chúng ta lựa chọn siêu tham số phù hợp. Trong nghiên cứu của chúng tôi, bằng cách thiết lập $n_{queries} = 200$, chỉ số mAP đã được cải thiện lên mức 0,981.
- Phương pháp của chúng tôi tương đương với phương pháp dựa trên CNN tốt nhất hiện nay trong việc phát hiện tàu. Nếu tập dữ liệu huấn luyện có nhiều mẫu hơn tập kiểm thử, phương pháp của chúng tôi tương đương với phương pháp dựa trên Yolo như phương pháp dựa trên DETR. Cụ thể, khi sử dụng D^{Train} và D^{Test} cho quá trình huấn luyện và kiểm thử, chỉ số mAP của cả phương pháp của chúng tôi và phương pháp dựa trên DETR là tương tự nhau. Tuy nhiên, khi số lượng mẫu trong tập huấn luyện bằng với số lượng trong tập kiểm thử, phương pháp của chúng tôi tốt hơn một chút so với phương pháp dựa trên DETR

Bảng 3.3: So sánh hiệu suất của các phương pháp khác nhau. Kết quả tốt nhất được in đậm.

Method	Train+Val / Test (in %)	fishing container ore bulk passenger general mAP						
		boat	ship	carrier	cargo	ship	cargo	ship
Zhang_2022 [13]	90/10	0.824	0.940	0.859	0.915	0.787	0.914	0.873
Zhang_2021 [10]	90/10	-	-	-	-	-	-	0.946
Cui_2019 [9]	80/20	0.900	0.940	0.90	0.910	0.910	0.900	0.910
Liu_2020 [12]	80/20	-	-	-	-	-	-	0.908
Han_2021 [28]	80/20	-	-	-	-	-	-	0.906

Liu_2022 [7]	80/20	-	-	-	-	-	-	0.964
SDNet_2022 [11]	80/20	0.986	0.995	0.989	0.990	0.982	0.989	0.988
phương pháp dựa trên DETR [14]	80/20	0.979	1	0.987	0.994	0.994	0.993	0.991
Ours ($n_{query}=200$)	80/20	0.982	1	0.989	0.991	0.995	0.990	0.991
Yani_2022 (ESDT) [17]	50/50	-	-	-	-	-	-	0.593
Yani_2022 (DETR) [17]	50/50	-	-	-	-	-	-	0.965
Biaohua_2022 [19]	50/50	0.940	0.987	0.966	0.978	0.937	0.972	0.963
phương pháp dựa trên DETR [14]	50/50	0.970	0.986	0.984	0.991	0.964	0.989	0.98
Ours ($n_{query}=200$)	50/50	0.970	0.995	0.992	0.992	0.957	0.992	0.982

3.3.3 Nghiên cứu cắt giảm các hàm mất mát

Phần này thảo luận về một nghiên cứu cắt giảm đối với các hàm mất mát trong quá trình huấn luyện. Deformable DETR sử dụng nhiều hàm mất mát để huấn luyện, bao gồm focal loss, GIoU loss và L1 loss, mỗi hàm phục vụ một mục đích riêng biệt: focal loss dùng cho phân loại, GIoU dùng cho hồi quy hộp giới hạn và L1 dùng cho phát hiện đối tượng. Sự kết hợp của các hàm mất mát này đảm bảo sự thành công cho quá trình huấn luyện. Mặc dù tất cả các hàm mất mát đều đóng vai trò then chốt, nhưng mức độ đóng góp của chúng có thể được điều chỉnh. Theo mặc định, các trọng số được thiết lập ở mức 2.0 cho focal loss, 2.0 cho GIoU loss và 5.0 cho L1 loss. Để đánh giá tác động của từng thành phần, chúng tôi đã giảm trọng số của các hàm mất mát này đi 10 lần (thực hiện lần lượt từng hàm một) và so sánh kết quả với thiết lập mặc định. Các kết quả trong Bảng 3.4 chứng minh rằng L_{cls} (hàm mất mát phân loại) là yếu tố quan trọng nhất; việc giảm trọng số của nó dẫn đến sự sụt giảm hiệu năng đáng kể. Ngược lại, việc giảm hàm mất mát đối tượng gây ra ít tác động hơn, khi hiệu năng vẫn duy trì ở mức gần tương đương với thiết lập ban đầu. Khả năng định vị bị ảnh hưởng nhẹ bởi hàm mất mát GIoU, thể hiện qua việc chỉ số mAP giảm xuống còn 0,931 so với 0,941 trong thiết lập ban đầu.

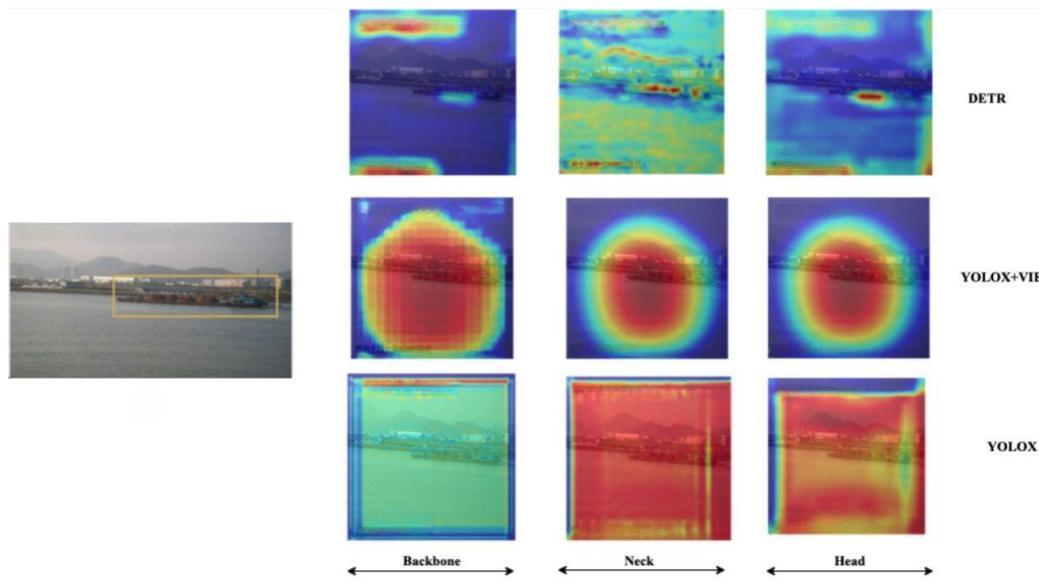
Bảng 3.4: So sánh mAP của Deformable DETR khi giảm các hàm mất mát huấn luyện.

	$LGIoU$	L_1	$Lcls$
fishing boat	0.899	0.899	0.227
container ship	0.976	0.985	0.246
ore carrier	0.953	0.947	0.212
bulk cargo carrier	0.951	0.966	0.151
passenger ship	0.845	0.849	0.0173
general cargo ship	0.964	0.969	0.160
mAP	0.931	0.936	0.178

3.3.4 Phân tích đặc trưng

Các kết quả thực nghiệm trong Mục 3.3.2 chỉ ra rằng phương pháp DETR hoạt động tốt hơn các phương pháp CNN nếu số lượng mẫu trong tập huấn luyện bị hạn chế. Tuy nhiên, việc giải thích tại sao phương pháp DETR có thể đạt kết quả tốt hơn trong kịch bản đặc thù này là rất cần thiết. Sự khác biệt chính giữa hai phương pháp nằm ở mô-đun chú ý trong bộ mã hóa của DETR. Mô-đun này cho phép tương tác phi cục bộ để học các đặc trưng tốt hơn.

Do đó, chúng tôi trực quan hóa các đặc trưng được tạo ra bởi cả hai phương pháp sau các phần backbone, neck và head của mô hình. Với mỗi ảnh đầu vào, các bản đồ đặc trưng được trích xuất sau mỗi mô-đun. Tổng giá trị của một bản đồ đặc trưng đại diện cho mức độ quan trọng của bản đồ đó. Vì vậy, chúng tôi đã chọn 20 bản đồ đặc trưng quan trọng nhất cho mỗi phần backbone, neck và head để tạo ra một bản đồ nhiệt. Bản đồ này được tính trung bình từ tất cả các bản đồ đặc trưng quan trọng và đại diện cho các điểm trọng yếu trên ảnh. Hình 3.2 minh họa các ví dụ về bản đồ nhiệt được tạo ra từ một ảnh đầu vào. Hàng đầu tiên hiển thị các bản đồ đặc trưng từ DETR, hàng thứ hai hiển thị các bản đồ nhiệt từ phương pháp dựa trên DETR (phương pháp sử dụng hàm mất mát lựa chọn đặc trưng để học các đặc trưng), và hàng thứ ba trình bày các bản đồ nhiệt từ YOLOX vốn dựa hoàn toàn vào các mạng CNN. Các kết quả chỉ ra rằng DETR, với cơ chế chú ý của nó, có thể tập trung tốt hơn vào các đối tượng không phải là nền. Ví dụ, sau phần head*, bản đồ đặc trưng làm nổi bật phần văn bản hiển thị điểm số trên hệ thống sinh tồn, và con tàu cũng được làm nổi bật, mặc dù không rõ ràng bằng phần văn bản. Khi các đặc trưng được xử lý qua phần neck, các đặc trưng ngữ nghĩa cấp cao hơn được học, khiến con tàu trở nên nổi bật hơn trong khi sự tập trung vào văn bản giảm dần. Các điểm chính tập trung vào con tàu tại phần head, với sự chú ý vào văn bản giảm đi.



Hình 3.2: Bản đồ đặc trưng tại đầu phân loại, phần cổ và phần xương sống. (Văn bản trên ảnh gốc đã được loại bỏ.)

Ngược lại, phương pháp dựa trên DETR tạo ra các bản đồ nhiệt thừa thớt, nơi nhiều điểm ảnh ở rìa bức ảnh không có phản hồi. Tuy nhiên, các bản đồ này không tập trung chính xác vào đối tượng. Điều này xảy ra bởi vì phương pháp dựa trên DETR sử dụng hàm mất mát chọn lọc đặc trưng để xác định các đặc trưng quan trọng, dẫn đến các bản đồ đặc trưng thừa thớt và có tính phân biệt cao nhưng không nằm chính xác ở tâm đối tượng. Sự phân bố bản đồ nhìn chung khá tương đồng, với những cải thiện nhỏ từ backbone đến head. Nếu không có hàm mất mát chọn lọc đặc trưng, sự phân bố của các bản đồ nhiệt có thể sẽ đồng đều hơn như được hiển thị ở hàng thứ ba.

Chương 4

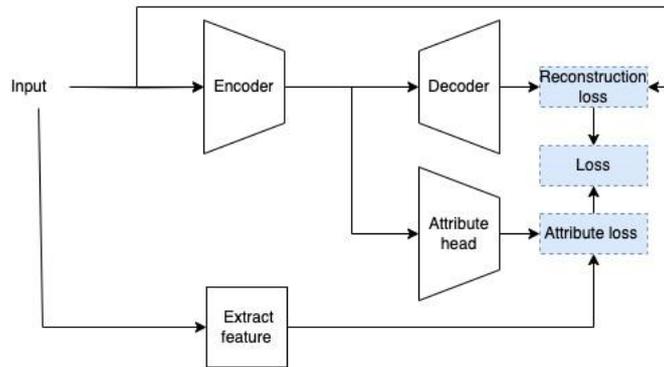
PHÂN TÍCH DỮ LIỆU RADAR BẰNG KỸ THUẬT HỌC SÂU.

4.1 Phương pháp phân cụm dựa trên kiến thức chuyên gia về phân loại tàu biển.

4.1.1 Tổng quan về hệ thống

Tổng quan về hệ thống được trình bày như trong Hình 4.1. Theo đó, bộ dữ liệu về tín hiệu radar được gán các thuộc tính theo kiến thức của chuyên gia. Sau đó mô hình học sâu sẽ được học các đặc trưng để vừa tái tạo lại tốt thông tin của sóng radar phản hồi, vừa dự đoán các thuộc tính được gán nhãn cho

mỗi mẫu. Cuối cùng, các đặc trưng đó sẽ được sử dụng để phân cụm dữ liệu.



Hình 4.1: Tổng quan về các phương pháp được đề nghị.

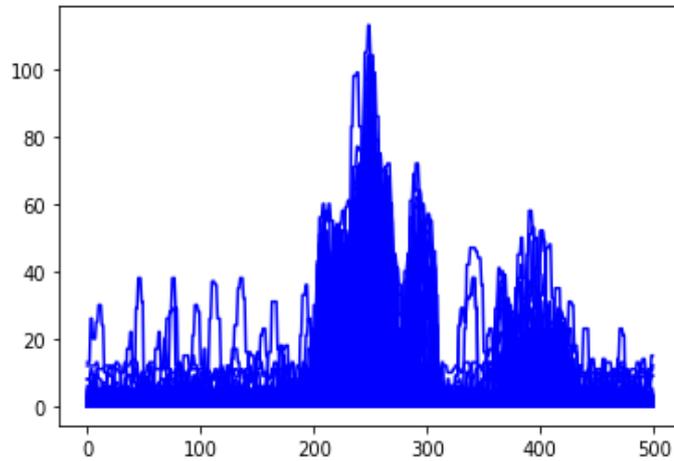
Phần 4.1.2, trình bày cách các thuộc tính mô tả các loại tàu biển và nguyên nhân lựa chọn các thuộc tính ấy. Phần 4.1.3 mô tả cách gán thuộc tính cho mỗi mẫu trong bộ dữ liệu. Phần ?? trình bày kiến trúc của model và hàm mục tiêu.

4.1.2 Các thuộc tính được đề xuất bởi chuyên gia.

Để có thể phân tích các thuộc tính có thể hỗ trợ cho việc phân cụm tàu biển, các mẫu tàu cá, tàu vận tải và tàu quân sự đã được phân tách và hiển thị như trong Hình 4.2, 4.3, 4.4. Các tàu cá có kích thước nhỏ và đặc điểm vỏ tàu khiến sóng phản hồi có biên độ nhỏ. Bên cạnh đó do các tàu cá thường đi cùng với nhau nên có thể có nhiều đỉnh sóng trong một sóng phản hồi như trong Hình 4.2. Trong Hình 4.3, ta thấy nếu tàu quân sự ở xa trạm radar thì sóng phản hồi sẽ có biên độ đỉnh sóng nhỏ như trong Hình 4.3a; ngược lại nếu tàu quân sự ở gần trạm radar thì sóng phản hồi sẽ có biên độ đỉnh sóng và biên độ đáy sóng lớn như trong Hình 4.3b. Đối với tàu vận tải; nếu tàu vận tải đang ở xa và có xu hướng chạy ngang bờ biển, thì sóng phản hồi có biên độ thấp và độ rộng đỉnh sóng cao hơn do có ảnh hưởng của các thùng hàng như hình 4.4a. Nếu tàu vận tải chạy gần bờ nhưng không hướng vào cảng, thì sóng phản hồi có biên độ lớn hơn, và phần đáy sóng cũng có biên độ lớn; ngoài ra phần nhấp nhô của

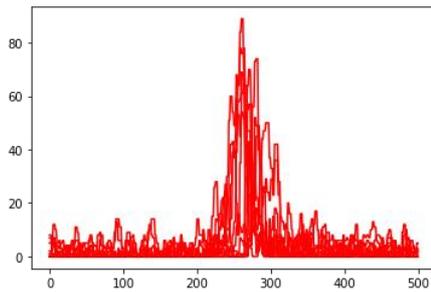
phần đỉnh sóng chuyển về vãn lớn như trong hình 4.4b. Nếu tàu vận tải chạy gần bờ nhưng có xu hướng hướng vào cảng, thì biên độ sóng lớn, nhưng độ rộng của sóng chỉ ở trung bình do lúc này chỉ

có phần mũi tàu tiếp xúc với sóng radar như trong Hình 4.4c.

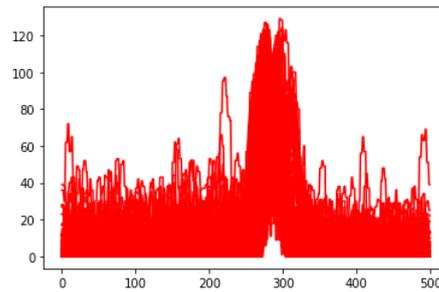


Hình 4.2: Dạng sóng của tàu cá, chỉ xuất hiện trong nhóm 1.

Từ những phân tích trên, các đặc điểm sau sẽ được sử dụng để đánh giá loại tàu dựa vào sóng phản hồi của radar.

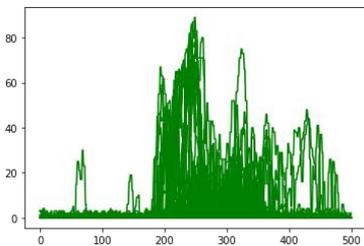


(a) Từ nhóm 1

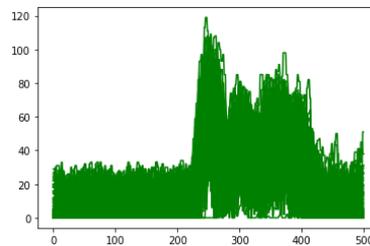


(b) Từ nhóm 2

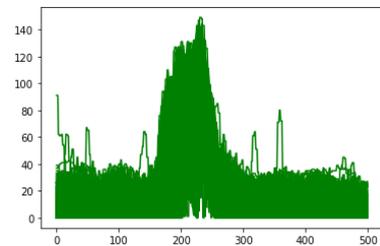
Hình 4.3: Dạng sóng của các tàu quân sự.



(a) Từ nhóm 1



(b) Từ nhóm 2

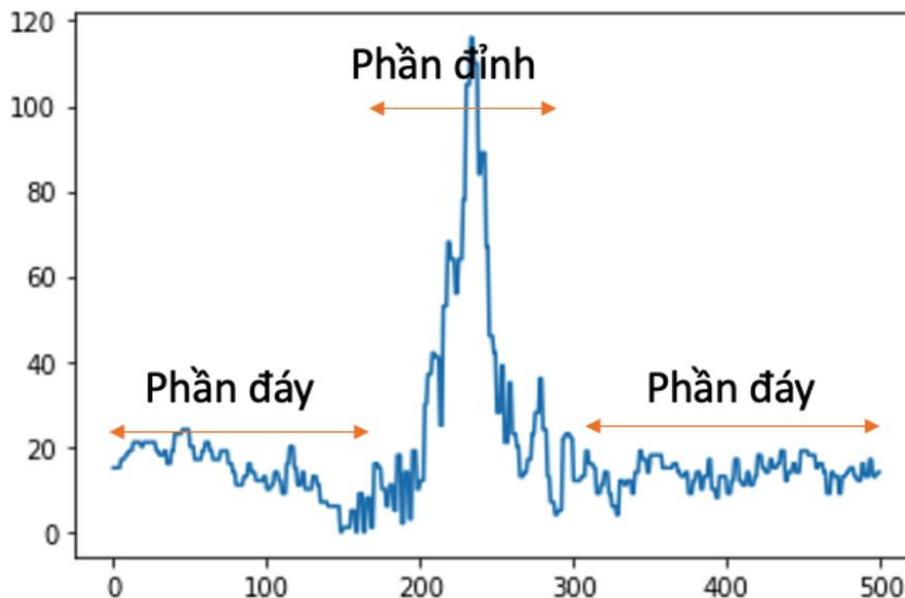


(c) Từ nhóm 3

Hình 4.4: Dạng sóng của các tàu vận tải.

- Số lượng đỉnh của sóng phản hồi: thông thường, dạng sóng phản hồi từ radar chỉ có một đỉnh nhưng đối với tàu cá, do đối tượng nhỏ và đôi khi bị nhiễu bởi các tàu cá lân cận nên có thể xuất hiện nhiều hơn một đỉnh. Nhưng nhìn chung số lượng đỉnh không quá nhiều đối với một sóng phản hồi.
- Biên độ phần đỉnh của sóng phản hồi: Đây là giá trị lớn nhất của phần đỉnh sóng, thường ảnh hưởng rất lớn tới kết quả phân cụm nếu chỉ sử dụng tín hiệu cảm biến thô.
- Độ rộng phần đỉnh của sóng phản hồi: đặc trưng này rất nhỏ với tín hiệu phản hồi từ tàu cá. Đối với tàu quân sự, giá trị này có thể nhỏ hoặc trung bình tùy thuộc vào khoảng cách tàu. Trong khi đó, độ rộng đỉnh sóng của tàu vận tải rất lớn do đặc thù nhiều thùng hàng được xếp lên tàu. Mặc dù vậy, đặc trưng này chỉ nhận giá trị trung bình nếu tàu hướng vào bờ biển.
- Độ dao động phần đỉnh của sóng phản hồi: Đặc trưng này mô tả độ dao động ở phần đỉnh của dạng sóng. Nếu là tàu cá hoặc tàu quân sự, giá trị này có xu hướng nhỏ hơn. Trong khi tàu vận tải thường có giá trị lớn khi mô tả bằng đặc trưng này.
- Biên độ phần đáy của sóng phản hồi: Đặc trưng này mô tả phần đáy của một dáng sóng. Nếu là tàu cá, giá trị này rất nhỏ vì kích thước của tàu cá nhỏ. Tàu vận tải hoặc tàu quân sự có giá trị lớn hơn vì đặc thù của vỏ tàu.
- Độ dao động phần đáy của sóng phản hồi: Với đặc trưng này, tàu cá thường có giá trị nhỏ, và tàu quân sự, vận tải thường có giá trị lớn.

4.1.3 Rút trích thuộc tính.



Hình 4.5 Đỉnh và đáy của dạng sóng radar phản hồi

Gọi tín hiệu sóng radar phản hồi về là x , quy trình rút trích các thuộc tính trong phần 4.1.2, được mô tả như sau:

1. Áp dụng bộ lọc Butterworth bậc 1 để rút trích thành phần thông thấp x^L và thông cao x^H .
2. Áp dụng thuật toán phát hiện đỉnh [10] để phát hiện số đỉnh n_{peak} của dạng sóng x .
3. Sử dụng tín hiệu tần số thấp x^L , có thể tách phần đỉnh và đáy của dạng sóng dựa vào thời điểm bắt đầu B và kết thúc E của đỉnh sóng cao nhất. Mô tả của các thời điểm B và E được mô tả như Hình 4.5. Gọi $min_{30}(x)$ là tập hợp 30% giá trị nhỏ nhất của x , giá trị $Base$ trong Hình 4.5 được xác định bằng công thức $Base = mean(min_{30}(x))$.
4. Dựa vào các thời điểm bắt đầu B và kết thúc E của đỉnh sóng cao nhất, rút trích thông tin phần đỉnh và phần đáy của thành phần tần số thấp x^L và thành phần tần số cao x^H . Cụ thể x_{BE}^L là phần đỉnh của tín hiệu tần số thấp; trong khi $x_{\sim BE}^L$ là phần đáy của tín hiệu tần số thấp.
5. Gọi N là số phần tử trong một mẫu tín hiệu radar phản xạ. Các đặc trưng về độ rộng phần đỉnh sóng, biên độ phần đỉnh sóng, độ dao động phần đỉnh sóng, biên độ phần đáy sóng, độ dao động

phần đáy sóng được tính bằng các phương trình 4.1, 4.2, 4.3, 4.4, và 4.5 một cách tương ứng.

$$L_{peak} = \frac{E-B}{N} \quad (4.1)$$

$$A_{peak} = mean(max_{30}(x_{BE}^L)) \quad (4.2)$$

$$S_{peak} = mean[x_{BE}^H{}^2] \quad (4.3)$$

$$A_{base} = mean(x_{\sim BE}^L) \quad (4.4)$$

$$S_{base} = mean[x_{\sim BE}^H{}^2] \quad (4.5)$$

4.1.4 Mô hình rút trích đặc trưng và hàm mục tiêu.

Kiến trúc của mô hình sử dụng trong bài báo được mô tả như sau:

Mô hình encoder: $f_{\theta^E}(\cdot)$ hình này được sử dụng để rút trích đặc trưng. Khối này là sự nối tiếp của các khối trong danh sách $[nn.linear, nn.relu,]$.

Mô hình decoder: $f_{\theta^D}(\cdot)$ Mô hình này được sử dụng để khôi phục lại thông tin gốc. Khối này là sự nối tiếp của các khối trong danh sách $[nn.linear, nn.relu,]$.

Mô hình dự đoán thuộc tính $f_{\theta^A}(\cdot)$: Khối này dùng để dự đoán các thuộc tính từ các đặc trưng được rút trích từ khối encoder. Các lớp trong khối dự đoán thuộc tính được mô tả như sau [*nn.linear, nn.relu*]

Gọi x là vector 500 phần tử mô tả tín hiệu vào của cảm biến và \hat{x} là dạng sóng được khôi phục lại,

hàm mục tiêu khôi phục dữ liệu được mô tả như phương trình 4.6.

$$L_R(x, \hat{x}) = \|x - \hat{x}\|^2 \quad (4.6)$$

Gọi a là thuộc tính được gắn nhãn từ phần 4.1.3 và \hat{a} là thuộc tính được dự đoán từ mô hình dự đoán

thuộc tính. Hàm mục tiêu xác thuộc tính được mô tả như phương trình 4.7.

$$L_A(a, \hat{a}) = \|a - \hat{a}\|^2 \quad (4.7)$$

Tham số α điều khiển sự cân bằng của hai ràng buộc $L_R(x, \hat{x})$ và $L_A(a, \hat{a})$. Kết quả hàm mục tiêu để

huấn luyện bộ rút trích đặc trưng sẽ được mô tả như sau:

$$Loss(x, a) = L_R\left(x, f_{\theta^D}\left(f_{\theta^E}(x)\right)\right) + \alpha L_A\left(a, f_{\theta^A}\left(f_{\theta^E}(x)\right)\right) \quad (4.8)$$

4.2 Kết quả phân đoạn cho tín hiệu tàu biển

4.2.1 Kết quả thí nghiệm với các phương pháp truyền thống.

Trong phần này, các kết quả thí nghiệm dựa trên các thuật toán phân cụm truyền thống được trình bày. Thuật toán K-means được sử dụng với các đặc trưng khác nhau như FFT [11], DCT [12], DWT [13]. Bên cạnh đó, các phương án chuẩn hoá dữ liệu cũng được áp dụng lên các đặc trưng để bù lại sự phi tuyến khi chuyển tín hiệu từ miền thời gian sang miền tần số. Kết quả thí nghiệm được mô tả trong Bảng 4.1. Kết quả cho thấy đặc trưng trong miền thời gian cho kết quả tốt hơn đặc trưng fft hoặc dct.

Bảng 4.1: Kết quả phân cụm dựa vào Kmeans và các đặc trưng.

Đặc trưng	chuẩn hoá	MI	A_MI	NorMI	A_RS	Complex	Flow
Raw	TRUE	14.41	17.48	17.61	03.05	24.78	51.59
Raw	FALSE	76.26	71.27	71.30	72.11	70.32	81.93
FFT	TRUE	49.11	46.22	46.28	46.96	46.00	65.86

FFT	FALSE	60.71	57.35	57.40	54.57	57.22	70.83
DCT	TRUE	60.81	59.58	59.63	47.23	61.75	67.72
DCT	FALSE	57.81	55.71	55.76	47.55	56.75	67.23
DWT	TRUE	44.63	45.79	45.86	34.65	50.05	61.95
DWT	FALSE	76.26	71.27	71.30	72.11	70.32	81.93

4.2.2 So sánh với các phương pháp học sâu tiên tiến.

Phương pháp	MI	A_MI	NorMI	A_RS	Complex	Flow
AE [14]	77.28	72.19	72.22	73.03	71.19	82.51
VAE [15]	79.18	74.16	74.19	76.56	73.33	84.84
DEC [16]	77.59	72.47	72.50	73.44	71.46	82.27
DeepCluster [17]	75.33	70.39	70.43	87.12	71.75	69.45
Proposed method	84.37	80.01	80.03	84.98	80.07	90.39
Only attribute	28.48	26.72	26.81	19.92	26.60	26.61

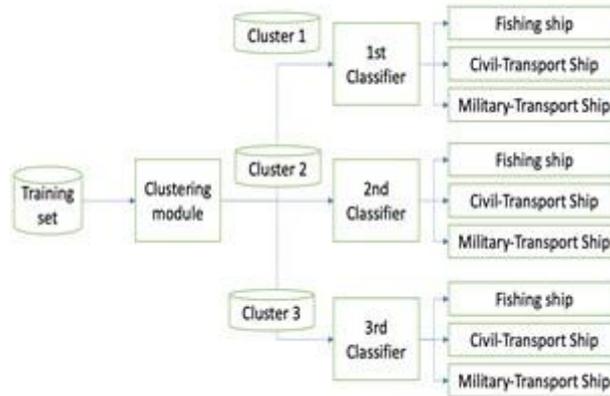
Bảng 4.2: Kết quả phân cụm bằng kỹ thuật học sâu

Phần này so sánh phương pháp đề nghị với các phương pháp học sâu cho phân đoạn bao gồm AE [14], VAE [15], DEC [16] và DeepCluster [17]. Trong đó AE [14] và VAE [15] là hai phương pháp kinh điển mà khối rút trích đặc trưng và khối phân cụm dữ liệu được tách biệt với nhau. Khác với các phương pháp học sâu kinh điển, phương pháp được đề nghị có sự hỗ trợ của thông tin thuộc tính để cải thiện bộ rút trích đặc trưng. Trong điều kiện dữ liệu còn ít, việc tận dụng kiến thức chuyên gia để huấn luyện mô hình rút trích đặc trưng là một phải pháp phù hợp. Theo đó, bộ rút trích đặc trưng không chỉ giúp tái tạo lại thông tin ban đầu mà đôi khi còn hữu ích trong việc dự đoán các thuộc tính của một mẫu. Ta có thể thấy giá trị MI được tăng đáng kể lên 84.13 % ứng với $\alpha = 0.1$. Cần lưu ý là nếu chỉ tiến hành phân cụm dựa vào thông tin của thuộc tính, thì giá trị phân cụm cho các chỉ số MI, A_MI, N_MI, RS, A_RS, CS, và FM_S sẽ lần lượt là 0.2848, 0.2672, 0.2680, 0.5083, 0.1992, 0.2660, và 0.4838. Qua đó có thể thấy việc gắn attribute không thực sự giúp ích trực tiếp cho việc phân cụm; nhưng khi kết hợp với hàm mục tiêu tái tạo tín hiệu, thì các đặc trưng học được sẽ hỗ trợ tìm được các đặc trưng tốt hơn cho việc phân cụm. Chính vì lý do này mà khi thiết kế khối phân cụm, ta cần sử

dùng nhiều hơn một lớp mạng neuron để giảm bớt ảnh hưởng từ việc dự đoán attribute đối với các đặc trưng dùng cho phân cụm.

4.3 Phân loại tàu biển thông qua kỹ thuật phân nhóm

4.3.1 Tổng quan hệ thống.



Hình 4.6: Tổng quan hệ thống

Do đặc thù bộ dữ liệu thu được vẫn còn ít, trong tiểu luận này, học viên chỉ sử dụng các phương pháp truyền thống để thực hiện việc phân loại tàu. Thuật toán phân cụm là một thuật toán chia một tập hợp các mẫu đầu vào thành các cụm khác nhau. Cụm là một nhóm các điểm dữ liệu tương tự nhau dựa trên mối quan hệ của chúng với các điểm dữ liệu xung quanh. Một trong những thuật toán phân cụm đơn giản và dễ thực hiện nhất là kỹ thuật K-means.

Thuật toán K-means được sử dụng tốt nhất trên các tập dữ liệu nhỏ hơn vì nó lặp lại trên tất cả các điểm dữ liệu. Điều đó có nghĩa là sẽ mất nhiều thời gian hơn để phân loại các điểm dữ liệu nếu có một lượng lớn chúng trong tập dữ liệu. Với tập dữ liệu $X = [x_1, x_2, \dots, x_N] \in R^{dxN}$ bao gồm N mẫu được biểu thị bằng d đặc trưng, thuật toán K-Means sẽ phân tách tập dữ liệu này thành K cụm. Trong mỗi cụm, các thành viên của nó phải càng giống nhau càng tốt. Mỗi cụm được đại diện bởi một tâm cụm là $m_k \in R^{dx1} (k = 1 \sim K)$; và đối với mỗi điểm trong tập dữ liệu, một nhãn phải được gán để xác định mẫu thuộc về cụm nào.

4.4 Một số kết quả thí nghiệm

4.4.1 So sánh các đặc trưng trong miền tần số và trong miền thời gian.

Một cách truyền thống, các bài toán phân loại tín hiệu 1D [19],[20] sẽ phải phân đoạn tín hiệu trước khi tiến hành phân loại. Bởi vì tín hiệu trong miền thời gian dễ bị ảnh hưởng bởi việc phân đoạn, các đặc trưng tần số của tín hiệu 1D sẽ được rút trích để loại bỏ ảnh hưởng này. Mặc dù vậy, trong ứng dụng này, sóng radar phát ra được điều chế bởi một xung vuông nên bài toán phân đoạn đã được giải quyết tốt. Do đó, tín hiệu trong miền thời gian cũng có thể là một tín hiệu đáng tin cậy để thực hiện các thuật toán học máy. Trong thí nghiệm này, học viên khảo sát ảnh hưởng của các đặc trưng trong miền tần số và trong miền không gian đối với các thuật toán học máy. Hai bài toán học máy cơ bản là phân nhóm và phân loại sẽ được áp dụng với các đặc trưng FFT và đặc trưng trong miền thời gian. Ứng với mỗi bài toán, học viên sử dụng 33% dữ liệu để kiểm tra và 66% dữ liệu cho huấn luyện. Các thuật toán được thực thi

dựa trên thư viện sklearn và tham số random seed được cố định để đảm bảo việc lựa chọn tập huấn luyện và tập đánh giá là giống nhau qua mỗi bước thí nghiệm.

Đối với bài toán phân nhóm, học viên sử dụng thuật toán K-means để phân chia dữ liệu thành 3 nhóm.

Kết quả phân nhóm được trình bày trong bảng 4.3.

Bảng 4.3: Kết quả phân loại với các đặc tính khác nhau của các loại tàu.

Quá trình huấn luyện				
		Tàu cá	Tàu vận tải quân sự	Tàu vận tải thường
FFT	C0	418	14	21
	C1	0	13	205
	C2	1	186	232
FFT + Scaling	C0	0	18	299
	C1	399	33	28
	C2	20	162	201
FFT + Log	C0	367	9	9
	C1	2	45	308
	C2	50	159	141
Time domain	C0	0	0	341
	C1	419	7	28
	C2	0	206	89
Quá trình kiểm tra đánh giá				
FFT	C0	192	6	10
	C1	0	9	114

	C2	0	100	106
FFT + Scaling	C0	0	16	116
	C1	186	15	12
	C2	6	84	102
FFT + Log	C0	175	2	6
	C1	1	32	160
	C2	16	81	84
Time domain	C0	0	0	180
	C1	192	4	14
	C2	0	111	36

Trong quá trình đánh giá, đưa vào một mẫu để kiểm tra, bộ phân nhóm ở trên sẽ được sử dụng để đưa dữ liệu kiểm tra vào một trong ba nhóm đã được tạo ra trong quá trình đào tạo. thuộc nhóm mà kiểm tra mẫu được chỉ định.

Bảng 4.4: Tóm tắt các cấu hình mạng nơ ron và thông số huấn luyện

Parameter	Solvver	Hệ số Regulization	Hidden Layer Size	Random State
S1	lbfgs	1e-5	(300,200,100,50)	100
S2	lbfgs	1e-5	(300,100,50)	100
S3	lbfgs	1e-5	(200,75)	100
S4	lbfgs	1e-5	(100,75)	100
S5	lbfgs	1e-5	(75,50)	100

Không chỉ thực hiện phân nhóm bằng các đặc trưng khác nhau, học viên thực hiện việc phân loại tàu trực tiếp dựa trên các đặc trưng đã được sử dụng trong bài toán phân nhóm. Học viên sử dụng ba cấu hình mạng nơ ron (S1,S3,S5) để đánh giá vai trò của các đặc trưng. Trong các bộ phân loại này, cấu hình S1 có độ phức tạp cao nhất vì gồm nhiều lớp và số nút ở mỗi lớp đều cao. Mô hình S5 có độ phức tạp thấp nhất vì chỉ có 2 lớp ẩn và số lượng nút cho mỗi lớp chỉ đều thấp. Chi tiết của các bộ phân loại này được mô tả trong bảng 4.4. Kết quả thí nghiệm trong bảng 4.5 cho thấy đặc trưng FFT không thích hợp để thực hiện việc phân loại tàu. Kết quả phân loại rất thấp chứng tỏ mô hình không thể hội tụ tốt để phân loại tàu. Nếu tiến hành bù lại các yếu tố phi tuyến bằng chuẩn hoá dữ liệu [21] và hàm log, mô hình có thể học nhưng độ chính xác không cao. Giá trị tốt nhất có thể đạt được trên tập test là 96% trong khi độ chính xác khi phân loại trên miền thời gian là 98%. Kết quả này thống nhất với kết quả khi thực hiện thuật toán phân nhóm khi chỉ ra rằng tín hiệu trong miền thời gian là đủ tốt để xử lý tín hiệu radar trên biển.

Bảng 4.5: Kết quả phân loại bằng các đặc trưng khác nhau với tỷ lệ huấn luyện/kiểm tra là 2.

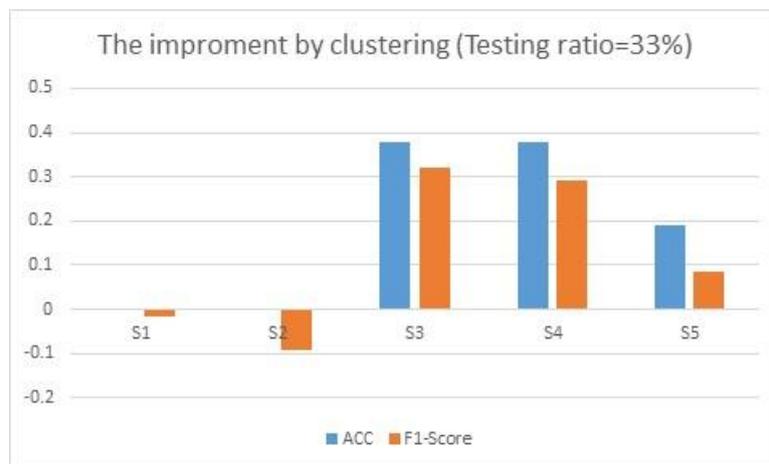
Pha test			
		ACC	F1-Score
FFT	S1	0.1195	0.2364
	S3	0.5680	0.6312
	S5	0.3910	0.4357
FFT + Scaling	S1	0.9606	0.9608
	S3	0.9606	0.9608
	S5	0.9492	0.9497

FFT + Log	S1	0.9414	0.9422
	S3	0.9243	0.9255
	S5	0.0945	0.9255
Time domain	S1	0.9813	0.9814
	S3	0.9795	0.9795
	S5	0.9795	0.9795

4.4.2 Ảnh hưởng của thuật toán phân nhóm lên bài toán phân loại

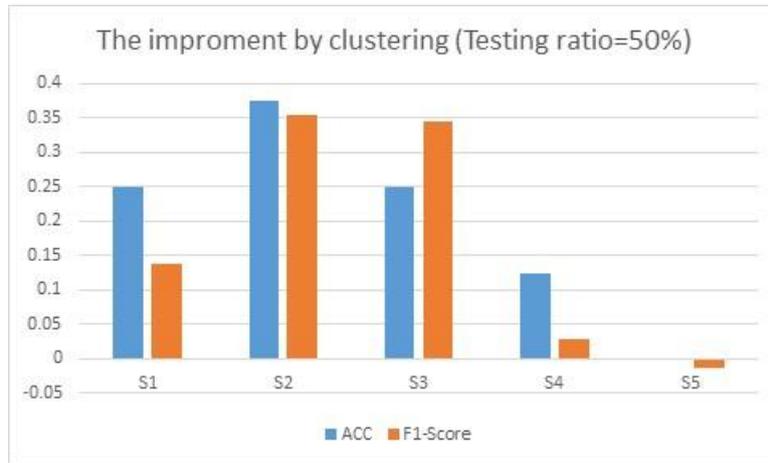
Trong phần này, học viên chứng minh tính hiệu quả của phương pháp được đề nghị khi sử dụng thuật toán phân nhóm để tiền xử lý dữ liệu trước khi đưa vào bài toán phân loại. Giống như các bài toán phân loại truyền thống, học viên sử dụng accuracy và F1-score để đánh giá hiệu quả của hệ thống. Mặc dù vậy, để nhấn mạnh sự đóng góp của quá trình phân nhóm, học viên sử dụng tỷ lệ tăng cường của ACC và F1-score như trong phương trình (4.13).

$$E = \frac{P_{clus} - P}{P} 100 \quad (4.13)$$



Hình 4.7: Độ tăng cường của ACC và F1-Score khi tỷ lệ huấn luyện/testing là 66%/33%

Ở đây, P là hiệu năng của hệ thống khi không có quá trình tiền xử lý phân nhóm; P_{clus} là hiệu năng của hệ thống khi có quá trình tiền xử lý phân nhóm. Hiệu năng ở đây bao gồm độ chính xác (ACC) hoặc điểm F1 (F1-score). Nếu giá trị này lớn hơn 0, ta có thể kết luận phép tiền xử lý bằng cách phân nhóm giúp cải thiện chất lượng của việc phân loại. Ngược lại, nếu giá trị này nhỏ hơn 0, có nghĩa là phép phân nhóm không giúp tăng cường performance. học viên tiến hành thí nghiệm với nhiều cấu hình nơ ron khác nhau từ rất dày đặc (dense) (S1) tới rất thưa thớt (sparse) (S5). Chi tiết của các cấu hình nơ ron và các thông số setting được liệt kê như trong bảng 4.4. Hình 4.6 thể hiện mức độ tăng cường của độ chính xác và F1-Score khi lựa chọn tỷ lệ huấn luyện/testing là 66%/33%.



Hình 4.8: Độ tăng cường của ACC và F1-Score khi tỷ lệ huấn luyện/testing là 66%/50%

Chương 5

KẾT LUẬN VÀ KIẾN NGHỊ

Giám sát và bảo đảm an ninh, an toàn trên biển và vùng ven bờ có vai trò đặc biệt quan trọng đối với phát triển kinh tế biển, bảo vệ chủ quyền quốc gia và duy trì trật tự hàng hải. Trong bối cảnh hoạt động hàng hải ngày càng gia tăng về quy mô và mức độ phức tạp, các hệ thống giám sát dựa trên một loại cảm biến đơn lẻ bộc lộ nhiều hạn chế, đặc biệt trong bài toán phát hiện, nhận dạng và phân loại tàu biển. Xuất phát từ yêu cầu thực tiễn đó, luận án tập trung nghiên cứu các phương pháp giám sát bờ biển dựa trên sự kết hợp giữa tín hiệu radar xung và dữ liệu ảnh camera, với sự hỗ trợ của các kỹ thuật trí tuệ nhân tạo hiện đại.

Thông qua khảo sát có hệ thống các công bố trong và ngoài nước, luận án chỉ ra rằng phần lớn các nghiên cứu hiện nay vẫn tiếp cận dữ liệu radar và dữ liệu ảnh một cách tách biệt. Trong khi radar có ưu thế về khả năng phát hiện và theo dõi mục tiêu trong mọi điều kiện thời tiết, camera lại cung cấp thông tin ngữ nghĩa phong phú phục vụ nhận dạng và phân loại. Việc tích hợp hai loại cảm biến này trong bối cảnh dữ liệu ven bờ thực tế, nhiều nhiễu và biến động, vẫn còn là một khoảng trống nghiên cứu đáng kể.

Đối với dữ liệu ảnh, luận án đề xuất hai mô hình phát hiện tàu biển dựa trên học sâu. Mô hình thứ nhất dựa trên mạng nơ-ron tích chập kết hợp các khối lựa chọn đặc trưng nhằm loại bỏ thông tin dư thừa và làm nổi bật các đặc trưng mang tính phân biệt cao. Kết quả thực nghiệm cho thấy mô hình đạt độ chính xác cao hơn so với nhiều phương pháp tiên tiến, đặc biệt hiệu quả trong điều kiện dữ liệu huấn luyện hạn chế. Mô hình thứ hai dựa trên kiến trúc transformer, khai thác cơ chế chú ý để tập trung vào các vùng thông tin quan trọng trong ảnh. Mô hình này thể hiện khả năng khái quát tốt với tập dữ liệu nhỏ và cung cấp các bản đồ chú ý có ý nghĩa, hỗ trợ phân tích và giải thích kết quả.

Đối với tín hiệu radar xung, luận án xây dựng một bộ dữ liệu thực tế thu thập từ trạm radar giám sát ven bờ tại Việt Nam, khắc phục hạn chế của các nghiên cứu trước đây vốn chủ yếu dựa trên dữ liệu mô phỏng. Trên cơ sở bộ dữ liệu này, luận án kết hợp tri thức chuyên gia với các phương pháp học máy để rút trích các đặc trưng phản ánh đặc tính vật lý và động học của mục tiêu. Phương pháp phân nhóm và phân loại tàu biển dựa trên tín hiệu radar được đề xuất cho thấy khả năng phân biệt các nhóm tàu chính, đồng thời chỉ ra những hạn chế của các phương pháp truyền thống khi chịu ảnh hưởng mạnh bởi yếu tố khoảng cách quan sát.

Tổng hợp các kết quả đạt được, luận án đóng góp cả về mặt lý thuyết và thực tiễn cho lĩnh vực giám sát bờ biển. Về học thuật, luận án làm rõ vai trò của học sâu và biểu diễn đặc trưng có chọn lọc trong bài toán nhận dạng tàu biển với dữ liệu hạn chế. Về ứng dụng, các mô hình và bộ dữ liệu được xây dựng tạo nên tảng cho việc phát triển các hệ thống giám sát bờ biển thông minh, hướng tới tự động hóa, nâng cao độ tin cậy và giảm sự phụ thuộc vào con người trong vận hành thực tế.

Tài liệu tham khảo

- [1] Z. Zhang, L. Zhang, Y. Wang, P. Feng, and R. He, “Shipsimagenet: A large-scale fine-grained dataset for ship detection in high-resolution optical remote sensing images,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 8458–8472, 2021.
- [2] Y. Zhang, M. J. Er, W. Gao, and J. Wu, “High performance ship detection via transformer and feature distillation,” in *2022 5th International Conference on Intelligent Autonomous Systems (ICoIAS)*, 2022, pp. 31–36.
- [3] B. Ye, T. Qin, H. Zhou, J. Lai, and X. Xie, “Cross-level attention and ratio consistency network for ship detection,” in *2022 26th International Conference on Pattern Recognition (ICPR)*, 2022, pp. 4644–4650.
- [4] Q. Zhang, Y. Huang, and R. Song, “A ship detection model based on yolox with lightweight adaptive channel feature fusion and sparse data augmentation,” in *2022 18th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 2022, pp. 1–8.
- [5] T. Liu, B. Pang, L. Zhang, W. Yang, and X. Sun, “Sea surface object detection algorithm based on yolo v4 fused with reverse depthwise separable convolution (rdsc) for usv,” *Journal of Marine Science and Engineering*, vol. 9, no. 7, 2021. [Online]. Available: <https://www.mdpi.com/2077-1312/9/7/753>
- [6] T. Liu, B. Pang, S. Ai, and X. Sun, “Study on visual detection algorithm of sea surface targets based on improved yolov3,” *Sensors*, vol. 20, no. 24, 2020. [Online]. Available: <https://www.mdpi.com/1424-8220/20/24/7263>
- [7] J. Zheng and Y. Liu, “A study on small-scale ship detection based on attention mechanism,” *IEEE Access*, vol. 10, pp. 77940–77949, 2022.
- [8] X. Han, L. Zhao, Y. Ning, and J. Hu, “Shipyolo: An enhanced model for ship detection,” *Journal of Advanced Transportation*, vol. 2021, pp. 1–11, 06 2021.
- [9] M. Zhang, X. Rong, and X. Yu, “Light-sdnet: A lightweight cnn architecture for ship detection,” *IEEE Access*, vol. 10, pp. 86647–86662, 2022.
- [10] M. Lueken, W. ten Kate, J. P. Batista, C. Ngo, C. Bollheimer, and S. Leonhardt, “Peak detection algorithm for gait segmentation in long-term monitoring for stride time

estimation using inertial measurement sensors,” in *2019 IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*, 2019, pp. 1–4.

- [11] J. S. Walker, *Fast Fourier transforms*. Boca Raton, Fla.: CRC Press, 1996. [Online]. Available: <http://www.amazon.co.uk/gp/search?index=books&linkCode=qs&keywords=9780849371639>
- [12] N. Ahmed, T. Natarajan, and K. Rao, “Discrete cosine transform,” *IEEE Transactions on Computers*, vol. C-23, no. 1, pp. 90–93, 1974.
- [13] M. Fu, H. Liu, Y. Yu, J. Chen, and K. Wang, “Dw-gan: A discrete wavelet transform gan for nonhomogeneous dehazing,” 2021.
- [14] D. Bank, N. Koenigstein, and R. Giryes, “Autoencoders,” 2021.
- [15] D. P. Kingma and M. Welling, “An introduction to variational autoencoders,” *Foundations and Trends® in Machine Learning*, vol. 12, no. 4, p. 307–392, 2019. [Online]. Available: <http://dx.doi.org/10.1561/22000000056>
- [16] J. Xie, R. Girshick, and A. Farhadi, “Unsupervised deep embedding for clustering analysis,” 2016.
- [17] M. Caron, P. Bojanowski, A. Joulin, and M. Douze, “Deep clustering for unsupervised learning of visual features.” *CoRR*, vol. abs/1807.05520, 2018. [Online]. Available: <http://dblp.uni-trier.de/db/journals/corr/corr1807.html#abs-1807-05520>
- [18] S. Rajkamal, “Selecting reviewers for research by clustering proposals using expectation maximization clustering algorithm,” in *2017 International Conference on Technical Advancements in Computers and Communications (ICTACC)*, 2017, pp. 56–60.
- [19] A. E. Vincent and K. Sreekumar, “A survey on approaches for ecg signal analysis with focus to feature extraction and classification,” in *2017 International Conference on Inventive Communication and Computational Technologies (ICICCT)*, 2017, pp. 140–144.
- [20] S.-S. Kim and T. Kasparis, “A modified domain deformation theory on 1-d signal classification,” *IEEE Signal Processing Letters*, vol. 5, no. 5, pp. 118–120, 1998.

- [21] M. S. Azmi, N. A. Arbain, A. K. Muda, Z. A. Abas, and Z. Muslim, "Data normalization for triangle features by adapting triangle nature for better classification," in *2015 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)*, 2015, pp. 1–6.