

**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP TP HỒ CHÍ MINH
KHOA CÔNG NGHỆ THÔNG TIN**



**ĐỒ ÁN CUỐI KÌ
DEEP LEARNING**

Người thực hiện: **CHÂU MỸ UYÊN – 20087481**

ĐẶNG QUỐC TOÀN – 20051051

Lớp : KHDL16A

Khoá : 16

Người hướng dẫn: **TS BÙI THANH HÙNG**

THÀNH PHỐ HỒ CHÍ MINH, NĂM 2023

**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP TP HỒ CHÍ MINH
KHOA CÔNG NGHỆ THÔNG TIN**



**ĐỒ ÁN CUỐI KÌ
DEEP LEARNING**

Người thực hiện: **CHÂU MỸ UYÊN - 20087481**

ĐẶNG QUỐC TOÀN - 20051051

Lớp : **KHDL16A**

Khoá : **16**

Người hướng dẫn: **TS. BÙI THANH HÙNG**

THÀNH PHỐ HỒ CHÍ MINH, NĂM 2023

LỜI CẢM ƠN

Để hoàn thành bài báo cáo này, chúng em vô cùng biết ơn và muốn gửi lời cảm ơn chân thành và sâu sắc đến thầy Bùi Thanh Hùng – người thầy đã góp phần trang bị cho chúng em những kiến thức bổ ích và những hành trang vô cùng quan trọng trong quá trình làm báo cáo, đồng thời cũng tạo điều kiện để chúng em có thể hoàn thành được bài báo cáo này.

Trong quá trình thực hiện bài báo cáo, chắc hẳn nhóm chúng em sẽ có nhiều thiếu sót do kiến thức còn nhiều hạn chế. Vì thế, em mong nhận được sự quan tâm và góp ý của thầy cô để bài báo cáo của nhóm em có thể hoàn chỉnh hơn nữa.

Chúng em xin chân thành cảm ơn.

ĐỒ ÁN ĐƯỢC HOÀN THÀNH TẠI TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP TP HỒ CHÍ MINH

Tôi xin cam đoan đây là sản phẩm đồ án của riêng chúng tôi và được sự hướng dẫn của TS. Bùi Thanh Hùng. Các nội dung nghiên cứu, kết quả trong đề tài này là trung thực và chưa công bố dưới bất kỳ hình thức nào trước đây. Những số liệu trong các bảng biểu phục vụ cho việc phân tích, nhận xét, đánh giá được chính tác giả thu thập từ các nguồn khác nhau có ghi rõ trong phần tài liệu tham khảo.

Ngoài ra, trong đồ án còn sử dụng một số nhận xét, đánh giá cũng như số liệu của các tác giả khác, cơ quan tổ chức khác đều có trích dẫn và chú thích nguồn gốc.

Nếu phát hiện có bất kỳ sự gian lận nào tôi xin hoàn toàn chịu trách nhiệm về nội dung đồ án của mình. Trường đại học Công nghiệp TP Hồ Chí Minh không liên quan đến những vi phạm tác quyền, bản quyền do tôi gây ra trong quá trình thực hiện (nếu có).

TP. Hồ Chí Minh, ngày tháng năm

Tác giả

(ký tên và ghi rõ họ tên)

Châu Mỹ Uyên

Đặng Quốc Toàn

PHẦN ĐÁNH GIÁ CỦA GIẢNG VIÊN

Tp. Hồ Chí Minh, ngày tháng năm
(kí và ghi họ tên)

TÓM TẮT

Bài toán Nhận diện tàu của Airbus có thể được hiểu là một bài toán phân đoạn hình ảnh, mục tiêu là phát hiện tất cả các tàu trong hình ảnh vệ tinh càng nhanh càng tốt. Điều này là một thách thức trong lĩnh vực thị giác máy tính, đặc biệt trong việc xử lý hình ảnh đồng thời với sự xuất hiện của nhiều thực thể trên cùng một bức ảnh. Vì mô hình Unet khá được ưa chuộng trong việc xử lý hình ảnh và giải quyết vấn đề đối với dữ liệu có cấu trúc đồ thị nên nhóm chúng em đã quyết định chọn triển khai hai kiến trúc mô hình chính là Unet với các Backbone khác nhau như ResNet50, VGG19, DenseNet169 và RCNN để giải quyết bài toán này. Việc sử dụng Unet trong bài toán Nhận diện tàu của Airbus mang lại kết quả đáng chú ý, Các mô hình này không chỉ giúp tự động hóa quá trình, mà còn nâng cao độ chính xác của quá trình phân loại và định vị, tạo nên cơ hội mới cho áp dụng thực tế trong lĩnh vực quan sát và giám sát từ vệ tinh. Điều này mở ra triển vọng lớn trong việc ứng dụng công nghệ thị giác máy tính để xử lý các thách thức phức tạp liên quan đến nhận diện và định vị tàu trên không gian ảnh.

MỤC LỤC

LỜI CẢM ƠN	i
PHẦN ĐÁNH GIÁ CỦA GIẢNG VIÊN	iii
TÓM TẮT	iv
MỤC LỤC	1
DANH MỤC KÍ HIỆU VÀ CHỮ VIẾT TẮT.....	Error! Bookmark not defined.
DANH MỤC CÁC HÌNH VẼ	2
DANH MỤC CÁC BẢNG	3
1.1 Giới thiệu về bài toán.....	4
1.2 Phân tích yêu cầu của bài toán.....	4
1.2.1 Yêu cầu của bài toán.....	4
1.2.2 Các phương pháp giải quyết bài toán	5
1.2.3 Phương pháp đề xuất giải quyết bài toán.....	6
1.3 Phương pháp giải quyết bài toán	8
1.3.1 Mô hình tổng quát.....	8
1.3.2 Đặc trưng của mô hình đề xuất	9
1.4 Thực nghiệm	10
1.4.1 Dữ liệu	10
1.4.2 Xử lý dữ liệu	10
1.4.3 Công nghệ sử dụng	10
1.4.4 Cách đánh giá.....	11
1.5 Kết quả đạt được	11
1.6 Kết luận.....	12
LÀM VIỆC NHÓM.....	13
TÀI LIỆU THAM KHẢO	14
PHỤ LỤC	15
TỰ ĐÁNH GIÁ	16

DANH MỤC CÁC HÌNH VẼ

Hình 1. Mô hình tổng quát	8
---------------------------------	---

DANH MỤC CÁC BẢNG

THỬ THÁCH NHẬN DIỆN TÀU CỦA AIRBUS - AIRBUS SHIP DETECTION CHALLENGE

1.1 Giới thiệu về bài toán

Lưu lượng vận chuyển đang tăng nhanh. Nhiều tàu hơn làm tăng nguy cơ vi phạm trên biển như tai nạn tàu tàn phá môi trường, cướp biển, đánh bắt trái phép, buôn bán ma túy và vận chuyển hàng hóa bất hợp pháp. Điều này đã buộc nhiều tổ chức, từ các cơ quan bảo vệ môi trường đến các công ty bảo hiểm và cơ quan chính phủ quốc gia, phải giám sát chặt chẽ hơn các vùng biển khơi.

Airbus cung cấp các dịch vụ giám sát hàng hải toàn diện bằng cách xây dựng một giải pháp có ý nghĩa với phạm vi bao phủ rộng, chi tiết rõ ràng, giám sát chuyên sâu, khả năng phản hồi cao cấp và phản hồi giải thích. Kết hợp dữ liệu độc quyền của mình với các nhà phân tích được đào tạo bài bản, họ giúp hỗ trợ ngành hàng hải nâng cao kiến thức, dự đoán các mối đe dọa, đưa ra cảnh báo và nâng cao hiệu quả trên biển.

Thử thách nhận diện tàu của Airbus mang ý nghĩa quan trọng trong việc ứng dụng công nghệ thị giác máy tính và vệ tinh để nhận diện tàu biển từ không gian. Cuộc thi không chỉ thách thức cộng đồng nghiên cứu mà còn có ý nghĩa lớn trong các lĩnh vực an ninh biển, quản lý giao thông biển, và bảo vệ môi trường. Những tiến bộ đạt được có thể tối ưu hóa giám sát biển, dự đoán hoạt động tàu, và đóng góp vào nỗ lực toàn cầu về an ninh và bảo vệ môi trường biển. Cuộc thi cũng thúc đẩy sự đổi mới trong lĩnh vực thị giác máy tính và khẳng định vai trò của công nghệ trong giải quyết các thách thức toàn cầu.

1.2 Phân tích yêu cầu của bài toán

1.2.1 Yêu cầu của bài toán

Bài toán đặt ra các yêu cầu cụ thể như sau:

1. Dữ liệu đầu vào: các hình ảnh có chứa tàu thuyền và các hình ảnh nhãn tương ứng, trong đó tất cả các pixel thuộc về tàu thuyền đã được gán nhãn.
2. Nhiệm vụ: thực hiện phân đoạn hình ảnh, tức là phân biệt và định vị vị trí của tàu thuyền trong hình ảnh
3. Kết quả đầu ra: hình ảnh có cùng kích thước với hình ảnh đầu vào, trong đó mỗi pixel được gán một giá trị để chỉ ra liệu pixel đó thuộc về tàu thuyền hay không.
4. Nhãn dữ liệu: Cần chuẩn bị nhãn chính xác để mô hình được đào tạo với thông tin chính xác về vị trí của tàu thuyền.
5. Đánh giá hiệu suất: Sử dụng độ đo Dice Score để đánh giá hiệu suất của mô hình trên tập dữ liệu đánh giá và kiểm tra kết quả đầu ra để đảm bảo tính chính xác và hiệu quả của việc nhận diện.

1.2.2 Các phương pháp giải quyết bài toán

1. U-Net Cho Phân Đoạn Hình Ảnh:

- Tiêu đề bài báo : "U-Net: Convolutional Networks for Biomedical Image Segmentation" (2015).
- Tác Giả: Olaf Ronneberger, Philipp Fischer, Thomas Brox.
- Phương Pháp Giải Quyết: Mô hình U-Net được giới thiệu cho việc phân đoạn hình ảnh y tế, nhưng có thể được áp dụng cho nhiều ứng dụng khác.
- Dữ Liệu Thực Nghiệm: Dữ liệu y tế chất lượng cao.
- Kết Quả Đạt Được: Hiệu suất tốt trong việc phân đoạn hình ảnh.

2. RCNN Cho Nhận Diện Đối Tượng:

- Tiêu đề bài báo: "Rich feature hierarchies for accurate object detection and semantic segmentation" (2014).
- Tác Giả: Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik.

- Phương Pháp Giải Quyết: RCNN sử dụng các phương pháp region proposal và deep learning để nhận diện đối tượng.
- Dữ Liệu Thực Nghiệm: Sử dụng tập dữ liệu PASCAL VOC.
- Kết Quả Đạt Được: Đạt được hiệu suất cao trong việc nhận diện đối tượng trong hình ảnh.

3. Phương Pháp Kết Hợp U-Net và RCNN:

- Tiêu đề bài báo: “ Mask-RCNN and U-net Ensembled for Nuclei Segmentation” (2019)
- Tác giả: Aarno Oskar Vuola, Saad Ullah Akram, Juho Kannala
- Phương pháp giải quyết: kết hợp hai kiến trúc phân đoạn là U-Net và Mask-RCNN để giải quyết bài toán phân đoạn hạt nhân trong hình ảnh y học.
- Dữ liệu thực nghiệm: tập hình ảnh y học chứa hạt nhân (nuclei).
- Kết quả đạt được: vượt qua cả hai kiến trúc gốc với hiệu suất đáng kể

1.2.3 Phương pháp đề xuất giải quyết bài toán

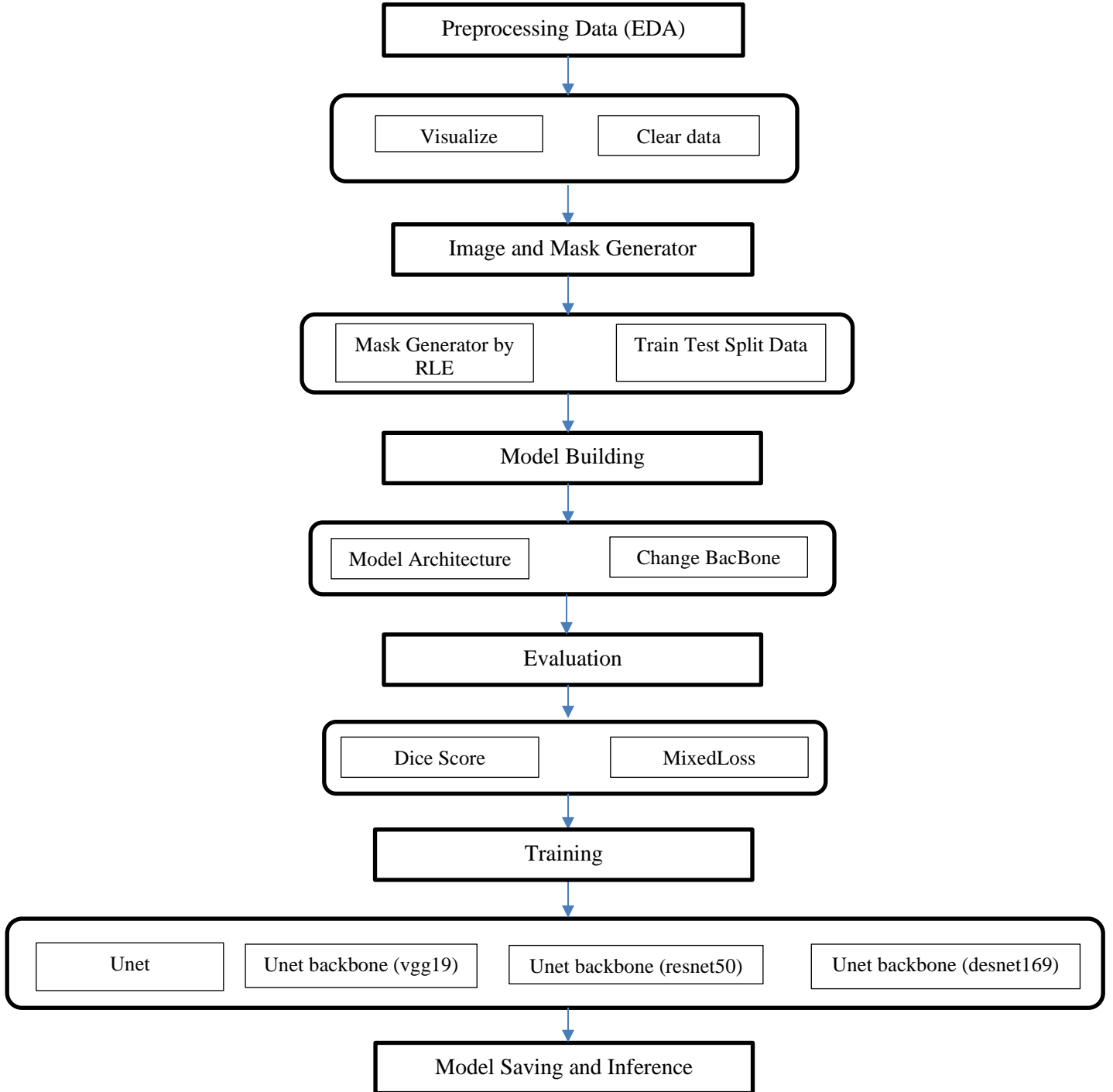
Nhóm quyết định sử dụng UNet để giải quyết bài toán nhận diện tàu của Airbus vì những lí do như sau:

- Chuyên Biệt Cho Phân Đoạn Hình Ảnh: U-Net là một mô hình chuyên biệt được thiết kế đặc biệt cho nhiệm vụ phân đoạn hình ảnh. Kiến trúc của nó giúp giữ lại thông tin chi tiết trong quá trình tái tạo hình ảnh.
- Kiến Trúc Encoder-Decoder: U-Net sử dụng kiến trúc Encoder-Decoder, giúp mô hình trích xuất và tái tạo đặc trưng một cách hiệu quả. Điều này đặc biệt hữu ích khi đối mặt với bài toán nhận diện tàu thuyền, nơi cần xác định đối tượng và đường biên của nó.

- **Chất Lượng Phân Đoạn Chi Tiết:** U-Net có khả năng tái tạo chi tiết cao, quan trọng để xác định chính xác các vùng chứa tàu thuyền trong hình ảnh. Điều này đáp ứng yêu cầu cụ thể của bài toán nhận diện tàu thuyền.
- **Linh Hoạt Cho Dữ Liệu Ảnh:** U-Net thường linh hoạt và hiệu quả khi áp dụng cho dữ liệu ảnh, đồng thời giảm thiểu lỗi và mất mát thông tin.

1.3 Phương pháp giải quyết bài toán

1.3.1 Mô hình tổng quát



Hình 1. Mô hình tổng quát

1.3.2 Đặc trưng của mô hình đề xuất

1. Tiền xử lý Dữ liệu (Preprocessing Data)

- Làm sạch Dữ liệu (Data Cleaning): ở bước này nhóm em đọc hiểu dữ liệu và xóa đi những dữ liệu trong file csv có Encodedpixels là giá trị null tức là không có mặt nạ hay còn gọi là tàu trong ảnh, ngoài ra nhóm còn clean đi những dữ liệu chứa size nhỏ hơn <35 kb bởi vì đa số dữ liệu dưới 35kb là hình ảnh không có tàu
- Visualize: Nhóm sử dụng các thư viện của Matplotlib hay Seaborn để biểu diễn data (hình ảnh) để có một cách nhìn trực quan

2. Xây dựng quá trình tạo mặt nạ thông qua hình ảnh (Image and Mask Generator)

- RLE: Phương pháp nhóm hướng đến là Run-Length Encoding kỹ thuật mã hóa dùng để biểu diễn các mặt (mask) đối tượng trong ảnh
- Train Test Split: Sau khi áp dụng phương pháp RLE cho việc mã hóa hình ảnh và mặt nạ được lưu trữ riêng biệt thì nhóm chia tập dữ liệu thành các tập train và tập test để dùng cho huấn luyện mô hình

3. Xây dựng Mô hình (Model Building)

- Model Architecture: Sử dụng mô hình Unet làm nền tảng, với các điều chỉnh tùy chỉnh như số lượng và kích thước của các lớp hidden, cũng như các tỷ lệ dropout để tránh overfitting.
- Change Backbone: Sử dụng mô hình Unet và thay đổi các lớp trích xuất đặc trưng để so sánh đánh giá với mô hình kiến trúc

4. Đánh giá (Evaluation)

- Dice Score: Trong bối cảnh phân đoạn hình ảnh, Dice Score thường được sử dụng để đánh giá chất lượng của mô hình phân đoạn bằng cách so sánh mặt nạ (mask) dự đoán và mặt nạ thực tế. Dice Score chạy từ 0 đến 1, với 0 thể hiện không có sự giống nhau và 1 thể hiện sự giống nhau hoàn toàn giữa hai

tập hợp. Trong ngữ cảnh phân đoạn hình ảnh, một Dice Score cao thường chỉ ra rằng mô hình phân đoạn có khả năng dự đoán chính xác các vùng của đối tượng trong hình ảnh.

- **Metric Tracking:** Ghi lại và theo dõi các chỉ số đánh giá sau mỗi epoch, bao gồm độ chính xác, loss, Binary Accuracy và Dice Score, cùng các biểu đồ tiến trình để đánh giá và tối ưu hóa mô hình qua thời gian.

1.4 Thực nghiệm

1.4.1 Dữ liệu

Dữ liệu lấy từ cuộc thi trên Kaggle có tên là “Airbus Ship Detection Challenge”.

Dữ liệu đầu vào gồm:

1. Hình ảnh vệ tinh: các bức ảnh vệ tinh chứa tàu hoặc không.
2. Mặt nạ phân đoạn: các khu vực trong hình ảnh được đánh dấu là tàu.

Dữ liệu được chia thành 2 phần: tập huấn luyện chứa ảnh vệ tinh chứa tàu và mặt nạ phân đoạn, còn tập kiểm tra cũng chứa ảnh vệ tinh chứa hình nhưng không có thông tin về mặt nạ phân đoạn.

1.4.2 Xử lý dữ liệu

Tiền xử lý dữ liệu thường là một phần quan trọng của quy trình đào tạo mô hình máy học, bao gồm cả việc sử dụng mô hình U-Net cho bài toán nhận diện tàu thuyền. Các bước tiền xử lý dữ liệu trong bài toán gồm:

- Xử lý giá trị thiếu(NaN)
- Loại bỏ những dữ liệu không liên quan tới tàu thuyền

1.4.3 Công nghệ sử dụng

Ngôn ngữ lập trình sử dụng: python

Các thư viện sử dụng : numpy, pandas, seaborn, tensorflow, matplotlib.pyplot

Các công cụ sử dụng giải quyết bài toán: Kaggle

1.4.4 Cách đánh giá

$$Dice\ Score = \frac{2 \times intersection + smooth}{union + smooth}$$

Dice score dao động từ 0 đến 1, 1 là hiệu suất tốt nhất. Cho phép đo độ tương đồng dựa trên kích thước của phần chung so với kích thước tổ hợp của 2 tập hợp.

1.5 Kết quả đạt được

Trong quá trình thực nghiệm, mô hình đã được huấn luyện qua 5 epoch với các tham số cấu hình như sau:

BATCH_SIZE = 48

EDGE_CROP = 16

GAUSSIAN_NOISE = 0.1

UPSAMPLE_MODE = 'SIMPLE'

NET_SCALING = (1, 1)

IMG_SCALING = (3, 3)

VALID_IMG_COUNT = 900

MAX_TRAIN_STEPS = 9

MAX_TRAIN_EPOCHS = 5

AUGMENT_BRIGHTNESS = False

Kết quả các epochs sau mỗi lần huấn luyện được thể hiện như sau:

Epoch 1	Epoch 2	Epoch 3	Epoch 4	Epoch 5	Best Binary Accuracy
0.9927	0.9993	0.9929	0.9930	0.9932	0.9993

1.6 Kết luận

1.6.1 Tóm tắt kết quả đạt được

Trong quá trình triển khai và đánh giá mô hình Unet để giải quyết bài toán nhận diện tàu thì kết quả nhóm đạt được lại khá tốt với binary accuracy lên tới 0.993 tuy nhiên khi thử nghiệm sử dụng lọc trích xuất với thêm lớp Resnet 50 hay vgg19,... thì kết quả cho ra không được tốt lắm

1.6.2 Hạn chế

Mục dù đạt được kết quả khá tốt với mô hình Unet tuy nhiên khi thêm lọc trích xuất thì nó có xu hướng overfitting.

Ngoài ra việc bộ dữ liệu cũng khá nặng nên để mô hình chạy tốt thì val loss phải < -2 mới xem như là mô hình đạt kết quả mong muốn nhất nhưng với tài nguyên hạn chế nên việc triển khai cũng gặp rất nhiều trở ngại

1.6.3 Hướng phát triển trong tương lai

Dựa trên những kết quả đạt được nhóm đề xuất ra một số hướng phát triển trong tương lai:

1. Tối ưu hóa mô hình: Tiếp tục tìm kiếm các cách tối ưu hóa mô hình để giảm thiểu overfitting, cũng như cải thiện khả năng tổng quát hóa trên tập dữ liệu đa dạng.
2. Nghiên cứu sâu hơn về dữ liệu: Phân tích sâu hơn về tập dữ liệu để hiểu rõ hơn về các yếu tố có thể ảnh hưởng đến hiệu suất của mô hình.
3. Triển khai thử nghiệm trên tập dữ liệu lớn hơn: Để kiểm tra và cải thiện khả năng tổng quát hóa của mô hình, việc triển khai thử nghiệm trên tập dữ liệu lớn hơn sẽ là bước quan trọng tiếp theo nhưng đi đôi với nó cũng cần một môi trường tài nguyên hợp lý
4. Tối ưu hóa về mặt tính toán:

Tìm cách fix được những lỗi từ trích xuất đặc trưng và tìm hiểu sâu hơn về mô hình Unet để giảm thiểu yêu cầu tính toán của mô hình

LÀM VIỆC NHÓM

STT	Công việc thực hiện	Người thực hiện	Trạng thái công việc
1	Xử lý dữ liệu	Toàn, Uyên	100%
2	Xây dựng mô hình	Toàn, Uyên	100%
3	Huấn luyện mô hình	Toàn	100%
4	Viết báo cáo	Uyên	100%

Nhóm 5 làm việc bằng hình thức online và cả offline

Thời gian offline không cụ thể, thời gian online là mỗi tối thứ 3, thứ 5, thứ 7 sau 22h

Tổng thời gian gặp nhau là khoảng 20 buổi, 50 giờ

TÀI LIỆU THAM KHẢO

1. [\[1505.04597\] U-Net: Convolutional Networks for Biomedical Image Segmentation \(arxiv.org\)](#)
2. <https://www.youtube.com/watch?v=pY3HaHFB7yA&t=3735s>
3. <https://www.kaggle.com/code/sidelmeril/airbus-ship-detection#Imports>
4. <https://lmb.informatik.uni-freiburg.de/people/ronneber/u-net/>
5. <https://www.geeksforgeeks.org/r-cnn-region-based-cnns/>
6. <https://www.youtube.com/watch?v=lgV1O4fHg6A&t=926s>
7. [\[1901.10170\] Mask-RCNN and U-net Ensembled for Nuclei Segmentation \(arxiv.org\)](#)
8. [\[1901.10170\] Mask-RCNN and U-net Ensembled for Nuclei Segmentation \(arxiv.org\)](#)

PHỤ LỤC

TỰ ĐÁNH GIÁ

STT	Nội dung	Điểm chuẩn	Tự chấm	Ghi chú
1 (8.5)	1.1 Giới thiệu về bài toán	0.5	0.5	
	1.2 Phân tích yêu cầu của bài toán	1	1	
	1.3 Phương pháp giải quyết bài toán	1.5	1.5	
	1.4 Thực nghiệm	4	2	
	1.5 Kết quả đạt được	1	0.5	
	1.6 Kết luận	0.5	0.5	
2 (1)	Báo cáo (chú ý các chú ý 2,3,4,6 ở trang trước, nếu sai sẽ bị trừ điểm nặng)	1đ	0.5	
3 (0.5)	Điểm nhóm (chú ý trả lời các câu hỏi trong mục làm việc nhóm)	0.5đ	0.5	
Tổng điểm			7	