



AN IMAGE RECOGNITION ALGORITHM APPROACH TO
DISTINGUISHING AND CATEGORISING THE SMALL-BOAT FLEET
IN THE GULF OF CALIFORNIA

A DISSERTATION
SUBMITTED TO UCL ENERGY INSTITUTE
OF UNIVERSITY COLLEGE LONDON
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

UCL Energy Institute
The Bartlett School of Environment, Energy and Resources
The Bartlett Faculty of the Built Environment
University College London

UCL Candidate Code: KNCM2

August 2021

Word count: 9906

Coursework Submission

Title “An Image Recognition Algorithm Approach to Distinguishing and Categorising
the Small-Boat Fleet in the Gulf of California”

Student name KNCM2

Student number (This can be found on your student identity card) 19068274

Module title MSc ESDA Dissertation

Module Code BENV0096

Submitted to UCL Energy Institute, Bartlett School of Energy, Environment and Resources

Statement of authorship:

I certify that the attached coursework exercise has been completed by me and that each and every quotation, diagram or other piece of exposition which is copied from or based upon the work of other has its source clearly acknowledged in the text at the place where it appears; all field work and/or laboratory work has been carried out by me with no more assistance from the members of the department than has been specified; and all additional assistance which I have received is indicated and referenced in the report .

Signature of Student

KNCM2

Date Submitted:

31 Aug 2021

You are reminded that the normal ESDA regulations apply when applying penalties for late submission – see the ESDA General Moodle page for details.

Abstract

Teaching machines to recognize and even understand satellite images is one of the most elusive and persistent challenges in artificial intelligence. This thesis tackles satellite image recognition: using the data and models needed to train models to identify small boats and measure their length. On the one hand, detecting small boats in satellite images is a major test of artificial intelligence algorithms. But, on the other hand, if we can build efficient object detection models, it will become a key technology for building carbon inventories in the shipping industry.

This thesis focuses on convolutional neural network models: a class of neural network models built on top of convolutional operations. These end-to-end neural models are more effective in learning media content (e.g., texts, images, and videos) and substantially improve performance on all object detection benchmarks than traditional feature-based models.

This thesis consists of two parts. In the first part, our goal is to cover every aspect of object detection algorithm and show our efforts in building effective object detection models and, more importantly, to understand what convolutional neural network models have actually learned. In the second part of this thesis, I show my detection results and analyze the various possibilities behind the results.

In particular, I discuss two new research directions at the end of this thesis: 1) whether a CAD model can replace the data from the training model, and 2) whether the algorithm has the ability (memory and understanding) to identify the same ship that appears at different times. I believe that they hold great promise for future satellite image analysis.

Acknowledgments

A special thanks to my supervisor Dr Santiago Suarez de la Fuente. He was always very insightful and visionary about the field, but he was also very detail-oriented and understood the nature of the issues very well. More importantly, Dr Suarez de la Fuente is an extremely caring and supportive mentor. He always gave positive feedback in each weekly meeting, and I couldn't ask for more.

Finally, I want to thank other academic colleagues inside and outside UCL, my fellow schoolmates, family and friends. Without them, this project would not be finished smoothly: Andrea Grech La Rosa, UCL Research Computing, Edward Gryspeerdt, Tom Lutherborrow, Jeff Jinfeng Guo, Mia Stieglitz-Courtney, Shiyi, Hao, Ju, and Shunying.

Contents

Abstract	iv
Acknowledgments	v
1 Introduction	1
1.1 Motivation	1
1.1.1 Energy Crisis, Resources and Climate Change	1
1.1.2 Small-Boat Fleet and Emission Inventory	5
1.1.3 Bringing Convolutional Neural Networks in Satellite Images Detection	7
1.2 Research Questions	8
1.3 Thesis Outline	9
1.4 Contributions	10
2 Literature Review	11
2.1 Small Boat Fleet and Shipping Decarbonisation	11
2.2 Convolutional Neural Networks in Image Recognition	13
2.3 Noise Removal for Image in the Shipping Sectors or Similar Applications .	17
3 Methodology	19
3.1 Artificial Intelligence Framework	19
3.1.1 The Algorithm: Deep Learning Architecture (CNN/ConvNet) . . .	19
3.1.2 Data Source and Improvements	24
3.1.3 Computing Power: CPU or GPU	24

3.2	Algorithms for Detection and Classification	25
3.2.1	Target Areas in the Gulf of California	25
3.2.2	Image Kernel and Image Sharpening	30
3.2.3	Detecting the Length of Small boats	31
3.2.4	Classification of Small Boats and Big Boats	32
3.2.5	Classification of Entertainment Boats and Fishing Boats	34
3.3	Workflow	34
4	Results	35
4.1	Train Custom Data: Weights, Biases Logging, Local Logging	35
4.2	Detection Results and Vessel Composition	38
4.3	Entertainment Boats in the Gulf of California	43
5	Future Work	46
6	Conclusions	49

List of Tables

3.1	Dates when Google Earth Pro has images from 2018 to 2021 in Baja California, Meico.	26
3.2	Dates when Google Earth Pro has images from 2018 to 2021 in Sinaloa, Meico.	26
3.3	Dates when Google Earth Pro has images from 2018 to 2021 in Colima, Meico.	26
3.4	Dates when Google Earth Pro has images from 2018 to 2021 in Nayarit, Meico.	27
3.5	Dates when Google Earth Pro has images from 2018 to 2021 in Jalisco, Meico.	27
3.6	Dates when Google Earth Pro has images from 2018 to 2021 in Sonora, Meico.	27
3.7	Specific dates when Google Earth Pro has images from 2018 to 2021 in Baja California Sur, Meico.	28
4.1	Small ships and large ships in Guaymas in 2019.	41
4.2	Small ships and large ships in Guaymas in 2020.	41
4.3	Small ships and large ships in Guaymas in 2021.	41
4.4	Small ships and large ships in Loreto in 2019.	41
4.5	Small ships and large ships in Loreto in 2020.	41
4.6	Small ships and large ships in Loreto in 2021.	41
4.7	Small ships and large ships in Santa Rosalia in 2019.	42
4.8	Small ships and large ships in Santa Rosalia in 2020.	42
4.9	Small ships and large ships in Santa Rosalia in 2021.	42

4.10	Small white boats in Guaymas in 2019.	43
4.11	Small white boats in Guaymas in 2020.	43
4.12	Small white boats in Guaymas in 2021.	44
4.13	Small white boats in Loreto in 2019.	44
4.14	Small white boats in Loreto 2020.	44
4.15	Small white boats in Loreto in 2021.	44
4.16	Small white boats in Santa Rosalia in 2019.	44
4.17	Small white boats in Santa Rosalia 2020.	44
4.18	Small white boats in Santa Rosalia in 2021.	44

List of Figures

1.1	Share of electricity production from renewables.	2
1.2	Fossil fuel price index, 1987 to 2015.	3
1.3	The evidence that atmospheric CO ₂ has increased since the Industrial Revolution began	4
1.4	Atmospheric CO ₂ concentration.	4
1.5	The hillshade image of the sonar or lidar output by the model.	7
2.1	Map of the density of human population, boats, and fishing offices in the Gulf of California.	12
2.2	YOLOv5 model in a driving video.	15
2.3	Faster RCNN model in a driving video.	15
2.4	YOLOv5 model in an NBA game video.	16
2.5	Faster RCNN model in an NBA game video.	16
3.1	Letter X in a 7x7 image.	20
3.2	Letter X in a 7x7 matrix.	20
3.3	A 3x3 convolution kernel.	20
3.4	A 5x5 feature map.	20
3.5	A 3x3 feature map after pooling.	22
3.6	A 3x3 feature map after activating with sigmoid function.	22
3.7	Average Precision (AP) vs. GPU Speed in the 6th generation of YOLOv5 model under COCO data set.	23
3.8	Approximate location of Santa Rosalia, Loreto, and Guaymas in the Gulf of California	29

3.9	Detections before and after sharpening. Detection accuracy increased by 26.32%.	30
3.10	An image from Google Earth Pro for Zurich Lake on 16 Aug 2018 when eye alt is 200 meters.	32
3.11	A cargo (March 2021, Guaymas) overflowed the image and will be excluded from the statistics.	33
3.12	The algorithm can detect all the boats (May 2021, Guaymas) well if no big boats overflowed the image.	33
3.13	The workflow of detection algorithm.	34
4.1	Average model precision when IOU is larger than 0.5; Average model precision when IOU is between 0.5 and 0.95; The model precision; The model recall rate.	35
4.2	The box loss rate of the model; The class loss rate of the model; The object loss rate of the model.	37
4.3	Test result of a trained model for detecting ships.	37
4.4	Test the length of boat.	38
4.5	When small boats are moored closely together in the harbour, the model may recognise two small boats as one. Image is from Guaymas, Jan 2020. .	40
4.6	When the cargo ship is full of cargo, the ship looks like a rectangular jetty from above and loses the normal shape of a ship. Image is from Guaymas, Jan 2020.	40
4.7	Models may have difficulty detecting small boats moored on the beach. Image is from SantaRosalia, Feb 2021.	40

Chapter 1

Introduction

1.1 Motivation

1.1.1 Energy Crisis, Resources and Climate Change

The energy crisis is one of the most essential and critical crises in the 21st century. Due to the growth of population and the increase of energy intensity per capita, a shortage of energy supply has occurred frequently in recent years, which often causes an energy crisis, usually involving scarcity of oil, electricity or other natural resources.

Nowadays, non-renewable resource still consists of a large proportion in the energy system. However, according to the BP Statistical Review of World Energy & Ember (see Figure 1.1), the share of electricity production from renewables has continued growing since 2007. In 2020, the share of electricity production from renewables was around 29% (see Figure 1.1). In other words, non-renewables are still the majority sources of electricity production today, although their share is shrinking.

Although we might not meet the complete depletion of non-renewable resources in the future 50 years, based on Hotelling's "Economics of Exhaustible Resources", David Ricardo proposed that as the historical production stock accumulates, higher grade ores get depleted, and the producer resorts to lower grade ores, sustaining greater extraction costs (Devarajan and Fisher, 1981). It means the extraction costs rise, and the price of the products based on ores will rise as well.

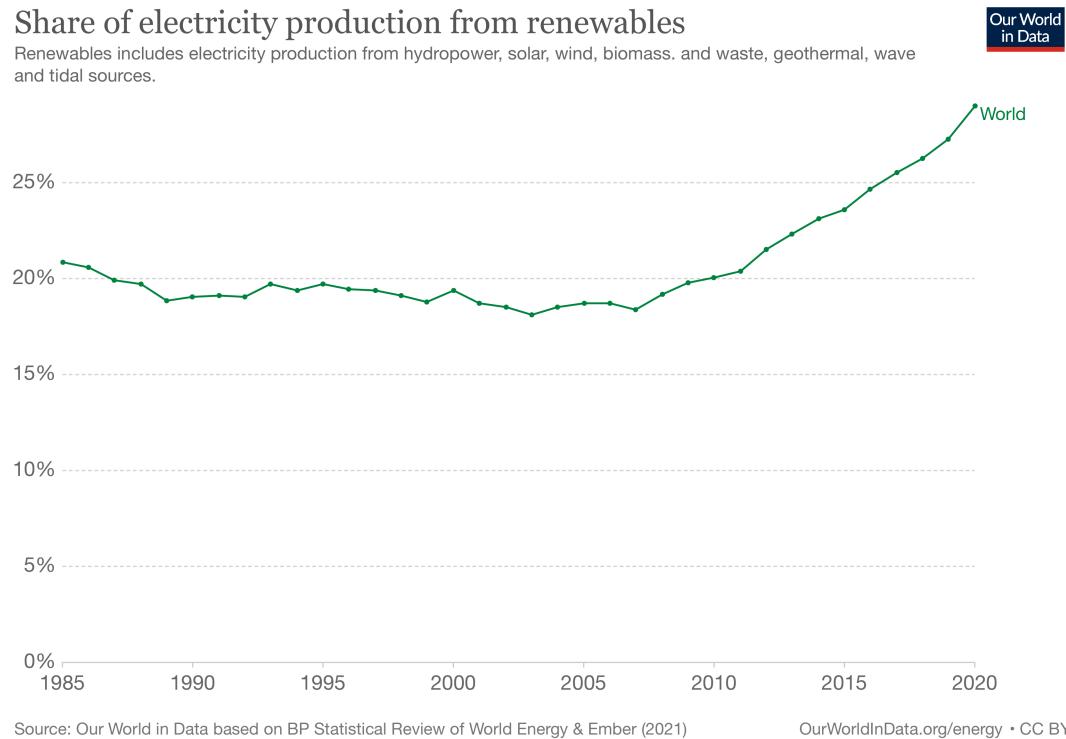


Figure 1.1: The share of electricity production from renewables increased from around 18% to 21% during 1985 to 2007 (BP, 2021). The share of electricity production from renewables has been continually growing since 2007. In 2020, the share of electricity production from renewables was around 29%.

Thus, we can assume that the price of most non-renewable resources, like oil, coal and gas, will rise since these have similar properties with ores. In fact, according to BP Statistical Review 2016 (BP, 2016), from 1987 to 2015 (from 1989 to 2015 for natural gas), the price of oil, coal and natural gas rose by approximately 36%, 81%, and 53% overall (Figure 1.2). It is worth mentioning that the final price of any fuel is a complex portfolio of extraction costs, offers and demand affected by different geopolitical events.

In the past 650,000 years, there were seven-cycle glaciers to advance and retreat. However, the climate in the past 70 years has been changed differently from other periods (Parmesan and Yohe, 2003). It has already had effects on the environment around us. Glaciers are shrinking, and ice is breaking up earlier on the lakes and rivers. Most climate scientists agree that it is human activities that cause global warming (Pachauri et al., 2014). The



Figure 1.2: Prices of different fossil fuels rose from 1987 to 2015 overall ([BP, 2016](#)).

climate on the earth is changing throughout history.

As we know, atmospheric CO₂ is a significant component of the atmosphere. However, atmospheric CO₂ had never been above 300 parts per million until 1950 (Figure 1.3).

The atmospheric concentration of CO₂ has been risen for around 36% from 1914 to 2018 (Figure 1.4). More than a third has increased atmospheric CO₂ concentration since the Industrial Revolution began ([Pachauri et al., 2014](#)). More importantly, atmospheric CO₂ has exceeded the highest level in the past 400,000 years (Figure 1.3), and it was 408.52 ppm in 2018 (Figure 1.4).

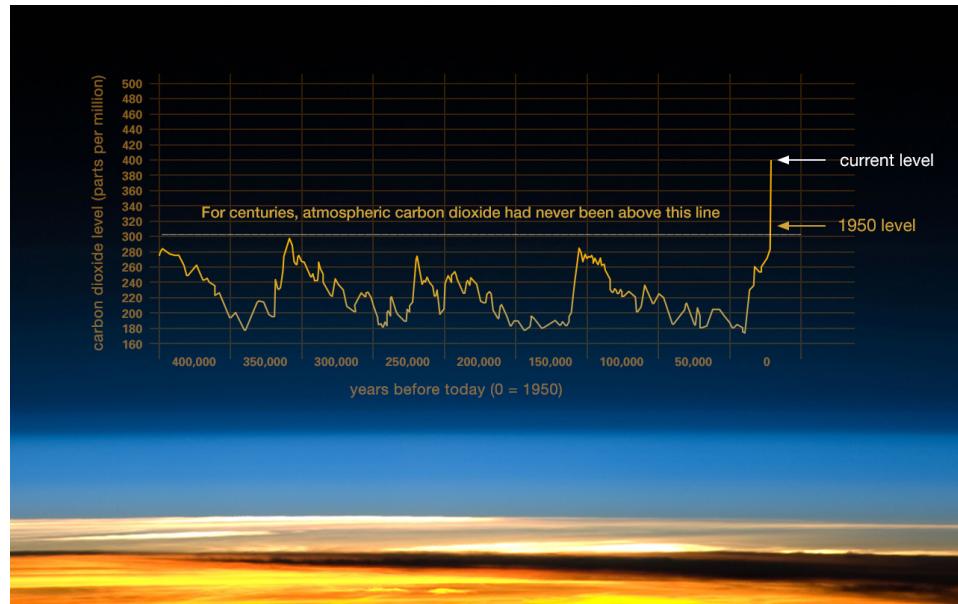


Figure 1.3: The evidence that atmospheric CO₂ has increased since the Industrial Revolution began. Image courtesy: <https://climate.nasa.gov/evidence>

Atmospheric CO₂ concentration

Global average long-term atmospheric concentration of carbon dioxide (CO₂), measured in parts per million (ppm). Long-term trends in CO₂ concentrations can be measured at high-resolution using preserved air samples from ice cores.

Our World
in Data



Figure 1.4: Atmospheric CO₂ concentration in 1914: 300.17ppm; Atmospheric CO₂ concentration in 2018: 408.52ppm.

In summary, the rising energy demand and lack of energy supply may cause a short-term energy crisis. The finite resources of fossil energy will drive the price of electricity and consumer goods to rise. Moreover, the burning of fossil energy will cause excessive emissions of greenhouse gases, leading to global warming. However, a door opens for cheaper but unpredictable renewable energy due to more expensive fossil fuels. According to Paris Agreements (UNFCCC, 2015), it is vital to reduce the emissions of fossil energy on a large scale in every sector of human activities. To respond to climate change, supporting the United Nations Sustainable Development Goals and taking urgent action to address climate change and its impact, the International Maritime Organization has formulated a timetable to reduce greenhouse gas emissions from international shipping. (Joung et al., 2020) It pointed out that between 2030 and 2050, the carbon intensity of the fleet will be reduced by at least 70% (Joung et al., 2020). Before 2050, the total annual greenhouse gas emissions will be reduced by at least 50%, which requires a reduction of approximately 85% of carbon dioxide per ship (Joung et al., 2020).

1.1.2 Small-Boat Fleet and Emission Inventory

Emission inventories for small vessels, including small fishing vessels are not developed for all years hence it is necessary to make the assumption of the emission growth for this class. In 2018, total shipping CO₂ emissions increased to 1056 million tonnes compared to 962 million tonnes in 2012 (IMO, 2021). In 2016, total CO₂ emissions of the industrial fishing sector were 159 million tonnes, and the small-scale fishing sector emitted 48 million tonnes (Greer et al., 2019). Suppose the increase rate of CO₂ in 2018 was the same as the rate in 2016, and the ratio of the number of small boats and the number of boats can be approximate as the ratio of the number of small fishing boats and the number of fishing boats. Then in 2018, the total CO₂ emissions for the small boat fleet can be calculated as 318.8 million tonnes.

Small vessels are classified as those smaller than 24 meters (UK Government, 2021). Knowing the shipping sector's emissions inventory can help understand what measures need to be taken to enable the industry to start the road to full decarbonization. Although it is possible to calculate large vessels from the international registry system and use the

satellite data sent from the ship's transponder to account for the large vessels (IMO, 2021), the small vessels depend on the national registration system, and their operation is assumed. In addition, there are many types of small vessels fleets such as machinery (e.g. fishery, people carrier, etc.), hull shape and structure, and the activities of owners and operators. The diverse small boat fleet operational profile is increasing the challenge of accurately accounting for their emission inventories.

Emissions from the global fishing industry grew by 28% between 1990 and 2011, with a minor coinciding increase in production; however, marine fisheries are typically excluded from international assessments of CO₂ or are generalized based on a limited number of case studies (Parker et al., 2018). Developed economies such as the UK have a national registry (UK Ship Register, 2021) that allows to have a sense of the level of small boat activity and hence infer the CO₂ emissions.

However, in developing countries, it tends to be a mixed bag on the level of precision and availability. For instance, in Mexico, only fishing vessels are counted into registry (Mexico National Aquaculture and Fisheries Commission, 2021). Still, it is difficult to know where they are located. Besides, the rest of the small-boat fleets are not considered. In all, Mexico does not have a regional CO₂ inventory considering the small-boat fleet. Therefore, quantifying the number of small-boat fleets will allow a better precision of where the emissions are being emitted and will be the focus of understanding the emission inventory of the shipping sector.

Observing the shipping activity in the Gulf of California is essential due to its unique geographical location, conformation, and biophysical environment (Lluch-Cota et al., 2007; Munguia-Vega et al., 2018; Marinone, 2012). In the Gulf of California, there is the largest fish producing state (Sonora) in Mexico (Meltzer and Chang, 2006) and the most prominent sports fishing destination (Los Cabos, Baja) in Mexico (Hernández-Trejo et al., 2012). Besides, the Gulf of California has a faster shipping route to mainland Mexico than from Yucatán Peninsula.

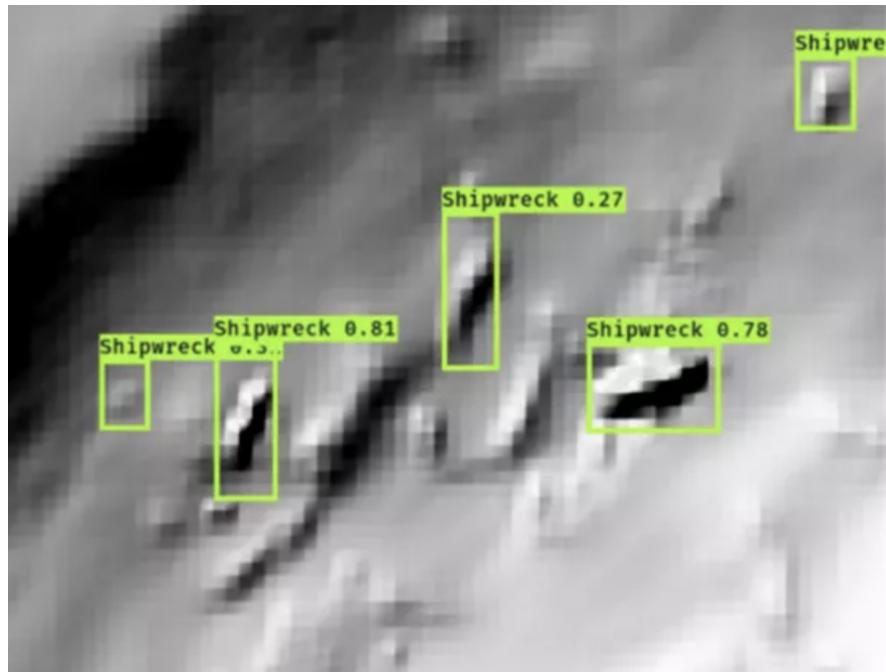


Figure 1.5: The hillshade image of the sonar or lidar output by the model.

1.1.3 Bringing Convolutional Neural Networks in Satellite Images Detection

Bringing convolutional neural networks to the field of satellite image recognition is important. First, the threshold of satellite image recognition is gradually decreasing. As the quality and quantity of global satellite images improve, obtaining the same or even better detection results with reduced parameters and complexity of convolutional neural networks is possible. Second, satellite image recognition is not new. A current PhD at the University of Texas at Austin focuses on using machine learning and remote sensing imagery to discover undersea shipwrecks (Character et al., 2021). However, as Figure 1.5 shows, discovering submarine wrecks does not require algorithms to describe the location and size of the wreck very precisely, as opposed to precisely identifying and measuring the length of a small boat. Therefore, in this thesis, accurately measuring the dimensions of small ships is definitely a challenge and a rewarding thing to do.

1.2 Research Questions

This article aims to determine whether image recognition such as convolutional neural networks can be used to successfully find boats with a length of 24 meters or less on the sea surface, and classify the boats according to their attributes.

Although the object detection technology based on convolutional neural networks is now mature, it is not easy to measure accurately to the centimeter level. Should convolutional neural network or image recognition technology be an effective solution? If effective, can this method be extended to a larger sea area to detect more different types of boats? The research questions to be completed in this paper are detailed as follows:

- (a) What is the small-boat fleet numbers and composition around the Gulf of California?
- (b) How accurate is the image machine-learning algorithm that recognises the small boat?
- (c) How this algorithm can be scaled up for the rest of Mexico/west coast of the U.S.A.?

1.3 Thesis Outline

Following the three research questions that we just discussed, this thesis consists of four parts — **CHAPTER 2 LITERATURE REVIEW**, **CHAPTER 3 METHODOLOGY**, **CHAPTER 4 RESULTS**, and **CHAPTER 5 CONCLUSIONS AND FUTURE WORK**.

In Chapter 2, I will begin with an overview of recent developments in identifying small boats and shipping decarbonisation. Next, I will briefly discuss developing convolutional neural networks and algorithms for cloud removal for satellite images.

In Chapter 3, I will talk about the algorithms I used in this project and their mathematical foundations, such as convolutional neural networks. Then, I will discuss the data sources. For example, how to improve the original data specification to fit an existing deep learning framework. Likewise, I will briefly introduce the advantages of GPUs. This will enable the reader to understand the mathematics behind this decision to use GPUs. Then, I will talk about everything about detection and classification. First, I will start by defining the scope of recognition. I will explain why I selected only some of the cities in the Gulf of California. Then, I will discuss the quality of the data provided by Google Earth Pro and the principles and effects of sharpening images using image kernels. Finally, I will explain the principle of detecting the length of small boats and the thinking behind classifying small boats. At the end of this chapter, I will include a workflow on the algorithm.

In Chapter 4, I will discuss the logs and the length of the detected boats when training the convolutional neural network. Importantly, I will also analyze the composition of the small boats in the three port cities and the potential reasons for this.

In Chapter 5 and Chapter 6, I will discuss a few pressing issues to be addressed in the future and summarize the conclusions reached and the progress made.

1.4 Contributions

The contributions of this thesis are summarized as follows:

- The accuracy of my trained object detection algorithm can reach 96% to 98%.
- I proposed and implemented an algorithm for measuring the length of an object (small boat) based on object detection, which showed excellent accuracy in tests.
- I proposed and implemented the algorithm to determine whether the detected small boat is a domestic recreational small boat, dividing small boats into domestic recreational small boats and fishing boats. Although the algorithm has flaws, this is an attempt to use a pure computer vision algorithm rather than a neural network in identifying objects. Nevertheless, I believe this is the first but important step in constructing a way to identify the types of small boats. Based on this, I answered questions about the composition of small boats in the Gulf of California.
- I developed Python scripts that can count the number of small boats, the number of large boats, and the number of small boats for home recreation. This automated the statistics and saved a great deal of time in counting.

Chapter 2

Literature Review

2.1 Small Boat Fleet and Shipping Decarbonisation

Current literature related to estimating small-scale vessels without machine learning methods includes using statistical factors or measures. Johnson et al. (2017) used Kernel Density Estimation (KDE) to distribute data on the population, the number of ships, and the average annual total catch for the entire population, and finally showed that their forecasts could accurately predict the landing of fisheries in the bay. In their paper, they explained the source of their data (López-Sagastegui et al., 2017). Lopez-Sagastegui et al. worked with fishers to generate and record information about fish catches, fishing efforts, profits, species breeding seasons, and spatial patterns of fishing activities. However, one of the points against this work is that it only focuses on fishing vessels. While these are the majority, it still misses the other ship types. Notably, the authors provided information (Figure 2.1) of the density of human population and boats in the Gulf of California that can save much time in creating data sets of the Gulf of California for training machines learning models.

Johansson et al. (2018) proposed a new model (FMI-BEAM) to describe the emissions of the leisure boat fleet in the Baltic Sea region with over 3000 dock locations, national small boat registry, AIS data and vessel survey results. However, the method cannot cover countries with no national registry for small boats. Besides, small boats are not just leisure boats. Ugé et al. (2020) estimated global ship emissions with the help of data from the Automatic Identification System (AIS). They set up movement information relating to ship size

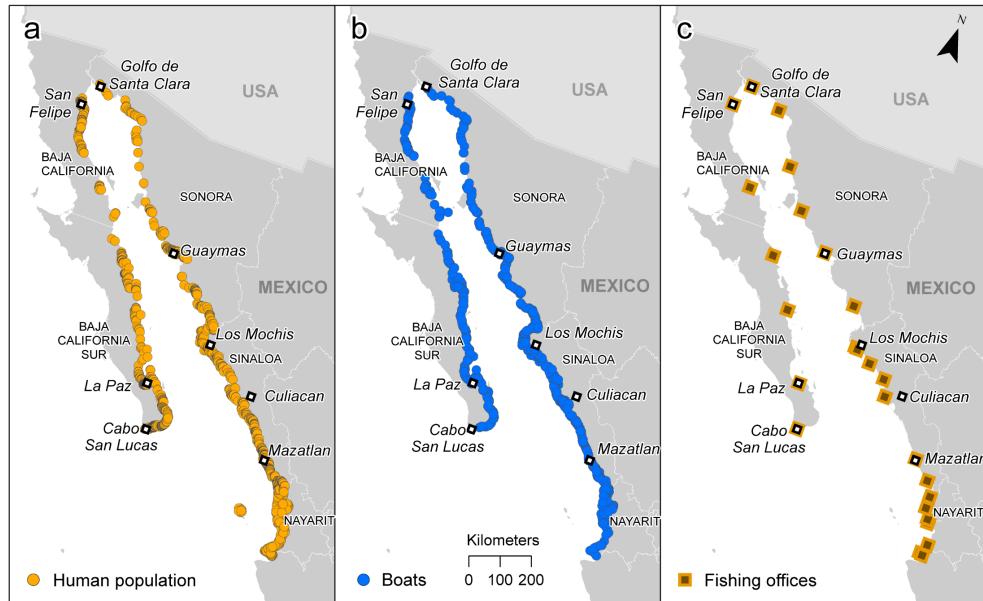


Figure 2.1: Map of the density of human population, boats, and fishing offices in the Gulf of California (Johnson et al.).

and speed and meteorological and marine environmental conditions. Over 3,000,000,000 daily AIS data records from hundreds of owners and thousands of partner AIS base stations and detailed ship data. However, this method is highly dependent on AIS data which is impossible for unregistered small boats. Traut et al. (2013); Johansson et al. (2016); Mabunda et al. (2014); Hensel et al. (2020); Han et al. (2016) have proposed the use of AIS to monitor the carbon emissions of the fleet as well.

Zhang et al. (2019) included unidentified vessels in the AIS-based vessel emission inventory. They developed an AIS-instrumented emissions inventory, including both identified and unidentified vessels. In particular, missing vessel parameters for unidentified vessels were estimated from a classification regression of vessels with similar vessel types and sizes in the AIS database. However, the authors do not discuss whether the regression model applies to vessels in most coastal areas of the planet. In addition, the authors do not discuss whether the vessel data in the AIS database is regionally diverse. Finally, if there is a diversity of vessels in the AIS database, the authors did not discuss whether this diversity would produce more significant errors in the predictions for small vessels in a single region (e.g. the Gulf of California, Mexico).

2.2 Convolutional Neural Networks in Image Recognition

Literature tends to be inaccurate for emission inventories for the small boat fleet. The above literature review has demonstrated that there is still a lot of work to understand how the small boat fleet is being operated, what fuels they are using, and the level of activity for this shipping sector. This master thesis project intends to use image recognition algorithms to detect small boats in any sea area, significantly reducing the time to calculate small vessel emission inventories. Besides, it will be in the national interest for the small fleet to account for and control these emissions within the powers of the state, incentivising the energy-efficient technologies and fuel change. Further, if countries meet their ambitious net-zero carbon emissions targets, they cannot afford to ignore the small boat fleet emission inventories that can help governments account for carbon emissions from small boats more quickly.

The goal of target detection, which is to determine the location of a target in an image based on a large number of predefined classes, is one of the most fundamental and challenging problems in computer vision. Deep learning techniques, which have emerged in recent years, are a powerful method for learning features directly from data and have led to significant breakthroughs in the field of target detection. Furthermore, with the rise of self-driving cars and face detection, the need for fast and accurate object detection is growing.

In 2012, AlexNet, a deep convolutional neural network (DCNN) proposed by Krizhevsky et al. (2012), achieved record accuracy in image classification at the ImageNet Large-Scale Visual Recognition Challenge (ILSRVC), making convolutional neural networks the dominant paradigm for image recognition. Then, Girshick et al. (2014) introduces Region-Based Convolutional Neural Networks (R-CNN), the first convolutional neural network (CNN)-based object detection method. The R-CNN algorithm represents a two-step approach, where a region proposal is generated firstly, and then a CNN is used for recognition and classification. Compared to the traditional sliding convolutional window to determine the possible regions of objects, R-CNN uses selective search to pre-extract some candidate regions that are more likely to object to avoid computationally costly classification and object search, which is much faster and significantly less expensive (Uijlings et al., 2013;

Girshick et al., 2014). Overall, the R-CNN approach is divided into four steps.

- Generating candidate regions;
- Feature extraction using CNN on the candidate regions;
- Feeding the extracted features into a support vector machine (SVM) classifier;
- Finally, the object positions are corrected using a regressor;

However, R-CNN also has many drawbacks: the selective search method is very slow in generating positive and negative sample candidate regions for the training network, which affects the overall speed of the algorithm; the CNN needs to perform feature extraction once for each generated candidate region separately, and there are a large number of repeated operations, which limits the algorithm performance (Huang et al., 2017).

Since its inception, R-CNN has undergone several developments and iterations: Fast R-CNN, Faster R-CNN and Mask R-CNN (Girshick, 2015; Ren et al., 2015; He et al., 2017). The improvement of Fast R-CNN is the design of a pooling layer structure for ROI (Region of Interest) Pooling, which effectively solves the operation of R-CNN algorithms that must crop and scale image regions to the same size. Faster R-CNN replaces the selective search method with RPN (Region Proposal Network) (Ren et al., 2015). The selection and judgment of candidate frames are handed over to the RPN for processing, and the candidate regions after RPN processing are subjected to multi-task loss-based classification and localization.

Several convolutional neural network-based object detection frameworks have recently emerged that can run faster, have a higher detection accuracy, have cleaner results and are easier to develop. Compared to the Faster RCNN model, the YOLO model can better detect smaller objects, for example, by easily detecting smaller traffic lights at a distance (Dwivedi, 2020) (Figure 3.7, Figure 3.8), which is important when detecting satellite images. Also, the YOLO model has a faster end-to-end runtime and detection accuracy than the Faster RCNN. For example, the figures below is from a YouTube clip of an NBA game. As seen in Figure 3.9 and Figure 3.10, the detection accuracy under the YOLO model is more accurate than that of the Faster RCNN. This helps to monitor the number of small boats in real-time using the YOLO model in the future.



Figure 2.2: YOLOv5 model in a driving video (Dwivedi, 2020).



Figure 2.3: Faster RCNN model in a driving video (Dwivedi, 2020).

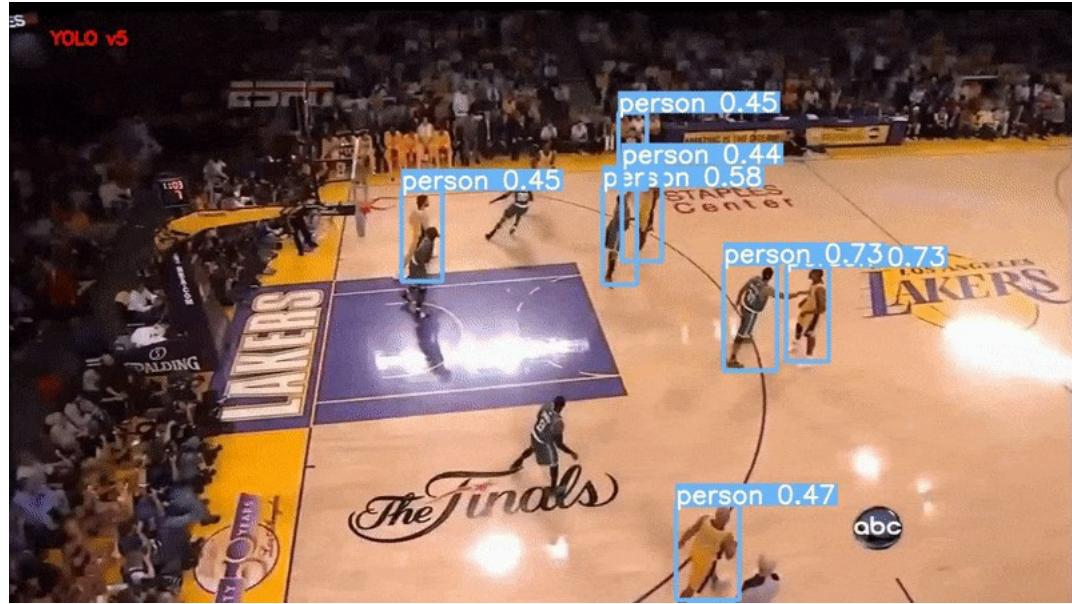


Figure 2.4: YOLOv5 model in an NBA game video (Dwivedi, 2020).

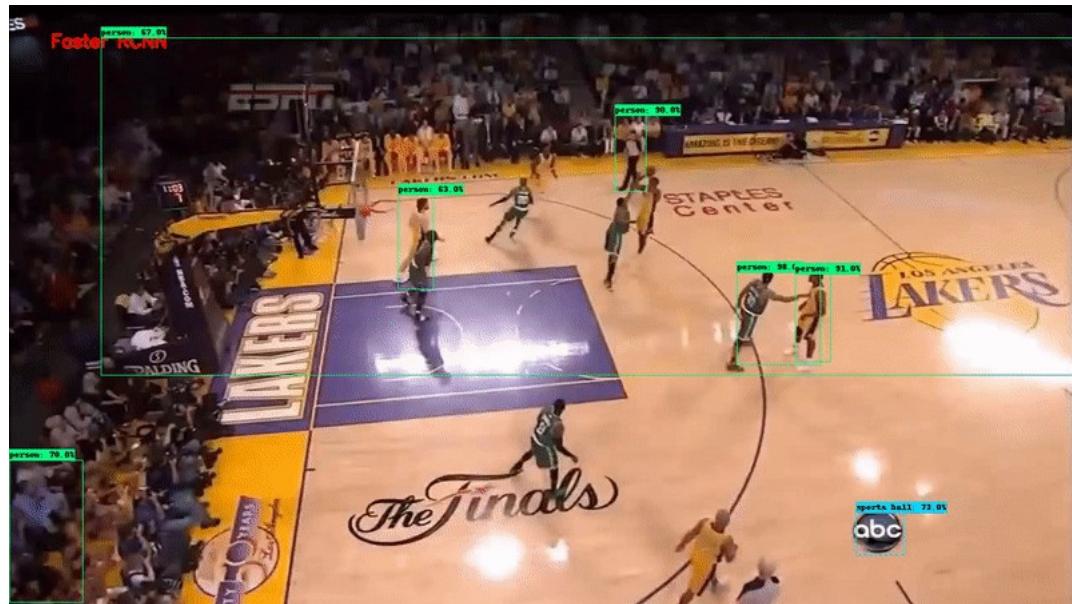


Figure 2.5: Faster RCNN model in an NBA game video (Dwivedi, 2020).

Mask R-CNN upgrades the ROI Pooling layer of Fast R-CNN to an ROI Align layer and adds a branching FCN layer, the mask layer, to the bounding box recognition for semantic mask recognition (He et al., 2017). Thus, the Mask R-CNN is essentially an Instance Segmentation algorithm, compared to Semantic Segmentation. Instance Segmentation is a more fine-grained segmentation of similar objects than Semantic Segmentation.

However, even traditional CNNs can be very useful for large-scale image recognition. Simonyan and Zisserman (2015) from the University of Oxford and Google Deep-Mind researched the effect of convolutional network depth on its accuracy in the large-scale image recognition setting. Their research found out that even they used very small (3x3) convolution filters, a significant improvement can be achieved by pushing the depth to 16 to 19 weight layers.

2.3 Noise Removal for Image in the Shipping Sectors or Similar Applications

Satellite images often have noises that should not be there, such as shadows cast by water on the sea surface due to sunlight or clouds in the atmosphere. These noises can make the training data inaccurate and often cause problems for the correctness of the model. He et al. (2009) proposed a simple but effective image prior-dark channel before removing haze from a single input image. The dark channel prior is a kind of statistics of outdoor haze-free images. Based on critical observation, most local patches in outdoor haze-free images contain some pixels whose intensity is very low in at least one colour channel. Using this before the haze imaging model, the thickness of the haze can be estimated, and a high-quality haze-free image can be recovered. Moreover, a high-quality depth map can also be obtained as a byproduct of haze removal.

In summary, past literature on carbon inventories of shipping has not focused on small vessels. However, with the development and maturation of a range of computer vision techniques such as convolutional neural networks, it may be possible to identify small vessels from open satellite imagery accurately. In the next section, I will first explain

how object detection models can be trained through a convolutional neural network-based YOLO architecture.

Chapter 3

Methodology

3.1 Artificial Intelligence Framework

3.1.1 The Algorithm: Deep Learning Architecture (CNN/ConvNet)

Neural networks originate from the human perception of the brain. In 1943, American neuroscientists McCulloch and Pitts proposed a theory that every neuron is a multiple-input single-output structure ([McCulloch and Pitts, 1943](#)). And there are only two possibilities for this output signal, it is either zero or one, which is very similar to a computer.

In image recognition, if we have a 7x7 image, this 7x7 image has 49 elements. If we write down an ‘X’ in this grid, as shown in Figure 3.1, in the computer’s view, it is actually a series of numbers. If each cell is either black or white, for example, black is 1 and white is 0, so what it represents could be a 7x7 matrix, as shown in Figure 3.2.

After feeding the program as much data as available, the program will be trained to find parameters to determine if it is an ‘X’ or not. For example, if it is a grey-scale picture, each number is not 0 nor 1, but a grey-scale value from 0 to 255. If it is a colour image, then it is RGB colours. Essentially, no matter what the image is, the image can end up replacing itself with a bunch of numbers, and that a bunch of numbers can be an input of the training neural network. The goal of training is to find the parameters that make the loss function ¹ smallest.

¹Loss functions measure how far an estimated value is from its true value.

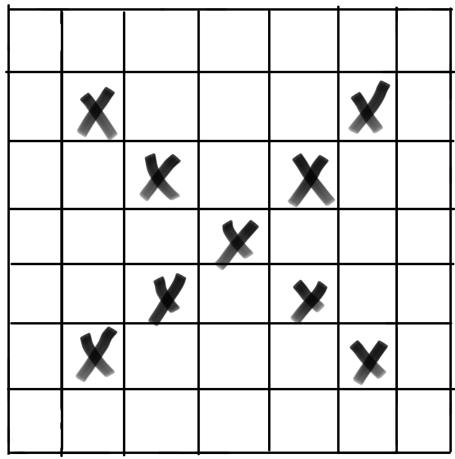


Figure 3.1: Letter X in a 7x7 image.

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Figure 3.2: Letter X in a 7x7 matrix.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Figure 3.3: A 3x3 convolution kernel.

$$\begin{bmatrix} 2 & 0 & 1 & 0 & 1 \\ 0 & 3 & 0 & 1 & 0 \\ 1 & 0 & 3 & 0 & 1 \\ 0 & 1 & 0 & 3 & 0 \\ 1 & 0 & 1 & 0 & 2 \end{bmatrix}$$

Figure 3.4: A 5x5 feature map.

But if we use the method described above to train real-world images, it is time-consuming and computationally expensive. Besides, once the image is a little bit deflated and rotated, or changed, then the algorithm will not recognize it. Contrary to that is our eyes that are very efficient: if I see a car and a motorcycle once, I can immediately tell the difference between them, and the next time I see the motorcycle, even if the motorcycle has changed direction, position, or is broken, I can still recognize it as a motorcycle and not a car.

In 1981, the Nobel Prize in Physiology or Medicine was awarded to two neuroscientists, David Hubel and Wiesel. These two scientists experimented with cats, inserting electrodes into their brains and then showing them a variety of different images to study the results of their brain responses (Hubel and Wiesel, 1962). They found that there are two types of cortical fields in the brain related to vision. The first one is called ‘simple cortical fields’, which is characterized by sensitivity to certain lines. After the appearance of lines in a certain direction, these fields will be more sensitive and can see it. There are ‘complex cortical fields’ as well. These complex fields are not only able to respond to lines, but it is also able to respond to the movement of the objects.

Later, inspired by them, a Japanese scientist named Kunihiko Fukushima proposed a model called the Neocognitron Model, which explains how a person can see whether the object is a cat or a dog (Fukushima and Miyake, 1982). He proposed that there are many layers in the human brain. The visual signals are processed layer by layer. After light enters the eye from the outside, it enters the first layer, then the second layer, then the third layer, and so on. Each layer processes the signal differently. In the beginning, when the light enters the retina, the human eye actually sees a large number of pixels. In the first layer, these pixels abstract some features, such as edges. Then the next layer combines these features to form the outline of the object and more details of the object. Finally, the contours and details are combined into a whole to make a final judgment.

Based on this principle, Yann LeCun invented a practical method for image recognition, called convolutional neural network (LeCun et al., 1995). The role of convolution is to use a mathematical method to extract these features from the image. The way to extract the features is to use a convolution kernel to do the convolution operation. The convolution kernel is also a matrix and is usually a 3x3 or 5x5 matrix. For instance, if we have a convolution kernel which has a 3x3 kernel, and the numbers in it are shown in Figure 3.3, then

a convolution operation will be done with this kernel and the matrix shown in Figure 3.2. The operation result is shown in Figure 3.4. This result is also known as a feature map.

In fact, the feature map reinforces the features of the convolution kernel. If you look carefully, you will see that this convolution kernel (Figure 3.3) only has three oblique blocks of pixels being ones. So if the original matrix (Figure 3.2) also has oblique pixel blocks of ones, the number would be extensive when they do the convolution operation. That means we have extracted this feature. The smaller the value of the pixel block in the other positions, the less it satisfies the feature. In a word, with different convolution kernels, we can get different feature maps.

The next step after convolution is pooling. The pooling method can reduce the size of the feature map and maintain similar features to the feature map before pooling. The largest elements in the 2x2 matrix range can be selected in left-to-right, top-down order. For example, for the feature map in Figure 3.4, the numbers in the first 2x2 matrix are 2, 0, 0, 3, then the largest number is 3. The numbers in the second 2x2 matrix are 1, 0, 0, 1, and then the largest number is 1. There are only two numbers, 1 and 0, in the third set of the 2x2 matrix, then the largest number is 1. Finally, we can extract the largest numbers from the 5x5 matrix in Figure 3.4, and the 5x5 matrix be pooled into a relatively small feature map (Figure 3.5).

The step after pooling is activation. The essence of the activation function is to introduce nonlinear factors to solve problems that a linear model cannot solve. Suppose the activation function is a sigmoid function, i.e.,

$$S(x) = \frac{1}{1 + e^{-x}}. \quad (3.1)$$

After activating the sigmoid function, each element in the feature map would be between 0 to 1, as shown in Figure 3.6.

$$\begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

Figure 3.5: A 3x3 feature map after pooling.

$$\begin{bmatrix} 0.95 & 0.73 & 0.73 \\ 0.73 & 0.95 & 0.73 \\ 0.73 & 0.73 & 0.88 \end{bmatrix}$$

Figure 3.6: A 3x3 feature map after activating with sigmoid function.

It is worth noting that the initial convolution kernel may be artificially set. But later on in the process of machine learning, it will go backwards to adjust and find the most suitable convolution kernel based on its own data, just like the above-mentioned method of using training to adjust the parameters is no different. Since a picture will generally have many features, there will be many corresponding convolution kernels. After many convolutions and poolings, features can be found, such as the slanted lines of the image, the contours, and the colour features. We take this information and then feed it into the fully connected network for training, and finally, it is possible to determine what this image really is.

In Sec 2.2, recent literature and development of Convolutional Neural Networks, including the YOLO model, were discussed. According to the official YOLOv5 tutorial, YOLOv5 has four different categories of models, YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x. They have 7.3 million, 21.4 million, 47 million and 87.7 million parameters respectively. The official performance charts are also available, as shown in Figure 3.7.

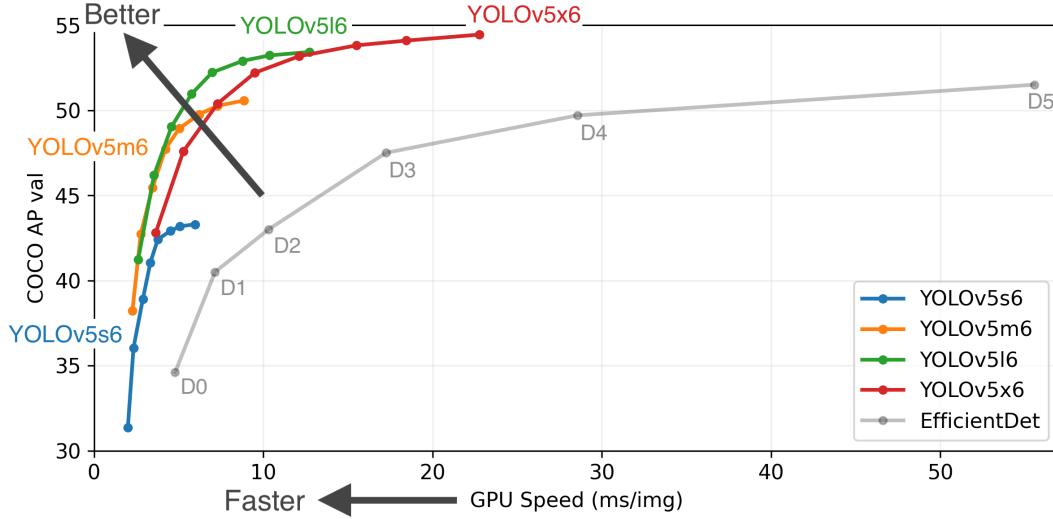


Figure 3.7: Average Precision (AP) vs. GPU Speed in the 6th generation of YOLOv5 model under COCO data set (Jocher, 2020).

Thus, the YOLOv5l model is able to achieve higher average precision with the same faster computing speed. Thus, in this thesis, the YOLOv5l model was chosen as the model for the training dataset.

3.1.2 Data Source and Improvements

Artificial intelligence algorithms need to be complemented by a large amount of data. An open-source data from a former TU Berlin researcher Tom Lutherborrow at Kaggle was used in this thesis. This open-source data contains 794 high-resolution images of ships from Google Earth ([Lutherborrow et al., 2018](#)). However, this dataset is larger than expected. Each image is larger than 9 megabytes, which is an efficiency burden for the training of neural networks, especially when there are few objects to be detected. With the tool RoboFlow, each training image can be resized to a 416x416 pixel image whose image size is approximately 10 KB to 40 KB.

3.1.3 Computing Power: CPU or GPU

Another important foundation of artificial intelligence is computing power. For example, in this convolutional neural network algorithm that we have just done, each calculation is actually not very complicated: it is just addition and multiplication. But the amount of computation is huge. For instance, there is an image that is 800x600, but due to the RGB colours, there are around 1.44 million pixels in total. Using a 3x3x3 convolution kernel will take about 13 million multiplications and about 12 million additions. It is simple addition and multiplication, but it is a vast number of operations. Besides, that's just using a convolution kernel for a simple image. There are actually thousands of images convolved thousands of times during the training process, so that's a vast number of operations.

The graphics processing unit (GPU) was originally meant to be a graphics processing unit, and it has a parallel structure that allows for more efficient computing than a central processing unit (CPU). For instance, if a link in a neural network is being computed simultaneously with another link, a large number of small cores can be used simultaneously to speed up the computation. In this thesis, all ConvNet models were trained using a YOLOv5 framework with Tesla P100 GPU in the Google Colab.

3.2 Algorithms for Detection and Classification

3.2.1 Target Areas in the Gulf of California

The Gulf of Californian was chosen as an area of study for this thesis. Ideally, in order to analyse a sufficient amount of satellite image data, the ports of each of the major harbour cities in the Gulf of California would need to be included in the scope of our study. Thus, it was worth finding out if there was enough data for the area in Google Earth Pro. The results of the statistics can be seen in Table 3.1 to Table 3.7.

In fact, Google Earth Pro provided very little image data than was expected. Most of the Mexican cities in the Gulf of California do not have satellite data for 2021. A few cities have satellite data from 2018. More cities have satellite data for 2019 and 2020. This reflects that:

- a steady increase in the collection of satellite data in the Gulf of California.
- the free, high-quality satellite data available from Google Earth Pro is not immediately available for publication.
- differences in satellite data are still evident. For example, some cities, such as Guaymas, will have access to about 15 satellite images a year in 2019 or 2020. However, some cities, such as La Ventana, did not appear in Google Earth Pro in 2019 and 2020, when almost all other cities have been photographed by satellite.

For this reason, it is clear that continuing with the previous strategy of analysing the satellite data for each city in the Gulf of California (from Table 3.1 to Table 3.5) would lead to a relatively large information bias and thus would not achieve an accurate result of the composition of the boats of the Gulf of California. Therefore, the following three cities with the most data in Google Earth Pro were finally chosen as the target areas for this thesis: Santa Rosalia, Loreto, Guaymas. Figure 3.8 shows the approximate location of Santa Rosalia, Loreto, and Guaymas in the Gulf of California.

City	2021	2020	2019	2018
San Felipe	NA	12/30; 12/9; 11/1; 4/19	12/26; 11/17; 1/22	6/17; 2/5; 10/23; 9/29; 3/3; 3/1; 2/26; 2/5
Bahia de los Angeles	NA	5/15	6/17	11/10
Isla Smith	NA	NA	6/17	NA

Table 3.1: Dates when Google Earth Pro has images from 2018 to 2021 in Baja California, Meico.

City	2021	2020	2019	2018
Los Mochis	2/1; 1/31	12/2; 9/30; 9/15; 9/12; 4/20; 1/29; 1/18	12/8; 8/3; 7/12; 6/13; 4/30; 2/7	11/6; 10/20; 7/3
La Cruz	1/16	10/31; 2/25; 2/21	10/29; 8/1; 7/10; 6/20; 4/19; 1/27	11/3; 10/16; 8/9; 1/17
Mazatlan	NA	10/31; 10/1; 9/19; 9/10; 8/29; 7/21; 2/28; 1/26; 1/18	12/7; 11/30; 8/3	11/3; 10/30; 10/27; 10/24; 10/12; 8/4; 6/27; 5/11; 4/19; 1/12

Table 3.2: Dates when Google Earth Pro has images from 2018 to 2021 in Sinaloa, Meico.

City	2021	2020	2019	2018
Manzanillo	NA	12/26; 9/14; 5/31; 3/3; 1/12	12/9; 3/7; 3/6	NA

Table 3.3: Dates when Google Earth Pro has images from 2018 to 2021 in Colima, Meico.

City	2021	2020	2019	2018
San Blas	NA	1/6	2/27	10/11; 5/7; 4/10; 2/27

Table 3.4: Dates when Google Earth Pro has images from 2018 to 2021 in Nayarit, Meico.

City	2021	2020	2019	2018
Puerto Vallarta	NA	5/1; 3/28	NA	3/27; 3/21; 3/19

Table 3.5: Dates when Google Earth Pro has images from 2018 to 2021 in Jalisco, Meico.

City	2021	2020	2019	2018
Golfo de Santa Clara	1/9	NA	6/24	NA
Puerto Penasco	1/26	8/27; 8/19; 7/28; 2/15	11/30; 11/7; 10/15; 10/12; 10/1; 6/19; 6/17; 2/15; 2/12; 2/7; 1/27; 1/25	10/10; 9/29; 9/13
Bahia de Kino	1/10	NA	6/25; 6/18	NA
Guaymas	2/10	12/9; 11/28; 11/9; 10/24; 9/24; 8/23; 8/17; 8/9; 7/26; 7/10; 5/27; 5/13; 4/23; 4/14; 3/6; 2/6; 2/1; 1/29; 1/7	11/13; 11/1; 9/15; 7/15; 6/21; 6/20; 6/13; 6/9; 4/24; 4/19; 3/9; 2/4; 1/19	11/1; 7/20; 3/16; 1/22

Table 3.6: Dates when Google Earth Pro has images from 2018 to 2021 in Sonora, Meico.

City	2021	2020	2019	2018	
Santa Rosalia	NA	12/21; 12/9; 11/28; 11/1; 9/11; 8/11; 8/6; 7/17; 7/14; 7/10; 7/6; 6/22; 6/17; 6/9; 5/26; 5/18; 5/8; 4/26; 4/18; 4/14; 3/15; 2/14; 2/3; 1/28; 1/27; 1/22; 1/13; 1/5	12/19; 12/14; 12/11; 12/3; 10/31; 10/20; 9/24; 9/20; 9/9; 8/26; 8/21; 8/11; 7/8; 7/5; 4/17; 1/30	11/9; 7/4; 6/7; 4/2; 2/23	
Mulege	1/18	NA	7/8	NA	
Loreto	2/12; 1/18	11/25; 11/23; 11/22; 11/4; 9/26; 9.13; 8/25; 8/17; 8/14; 7/29; 7/25; 7/16; 7/14; 7/10; 6/25; 6/3; 4/30; 4/23; 4/22; 4/18; 4/11; 3/16; 2/25; 1/29; 1/18; 1/2	12/30; 12/11; 11/1; 10/31;	10/2; 9/26; 8/14	
La Paz	NA	1/18	NA	4/4	
La Ventana	NA	NA	NA	4/4	
San Jose del Cabo	NA	NA	NA	3/28	
Cabo San Lucas	NA	3/28	NA	3/28	

Table 3.7: Specific dates when Google Earth Pro has images from 2018 to 2021 in Baja California Sur, Mexico.

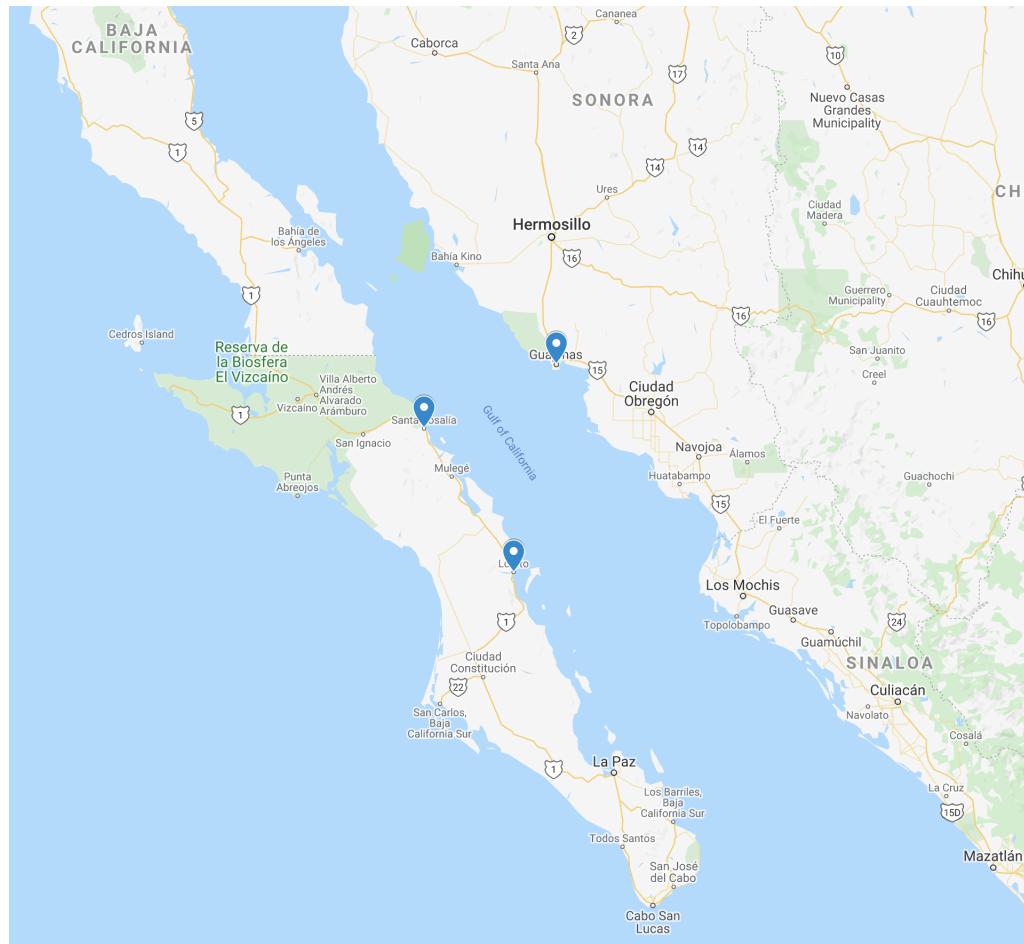


Figure 3.8: Approximate location of Santa Rosalia, Loreto, and Guaymas in the Gulf of California

3.2.2 Image Kernel and Image Sharpening

Similar to the principle of using convolution kernels, specific image kernels can sharpen the image. While the sharpening kernel does not make the image higher resolution, it emphasises the differences in adjacent pixel values which makes the image look more vivid. Overall, with the same model weight, as shown in Figure 3.9, sharpening an image can significantly improve the recognition accuracy of an image with a 5x5 image kernel.



(a) Detection before sharpening.



(b) Detection after sharpening.

Figure 3.9: Detections before and after sharpening. Detection accuracy increased by 26.32%.

3.2.3 Detecting the Length of Small boats

Measuring the length of a ship is one of the most difficult topics in this thesis. As Google Earth Pro does not provide an application programming interface (API) for accurate scales, manually measuring the size of a particular scale became the core of measuring the size of a ship.

As the dataset for the training model was created with each edge tangent to the edge of the detected object, we can roughly treat the boat's length as the length of the diagonal of the detection box. Secondly, since the scale is central to the production of the bay dataset that follows, i.e., the satellite image at this scale:

- cannot be too large: they need to contain the full complement of small boats (if any), and the complexity of producing the dataset must also be considered;
- cannot be too small: otherwise, there is a tendency for multiple small boats docked together to be detected as a whole party;
- be sufficiently clear: the model needs to be allowed to accurately identify the length of small boats to reduce errors in scale.

Based on the above rules, the eye altitude was finally determined to be 200m. Meanwhile, I used a satellite image of Zurich Lake, Switzerland, on 16 August 2018 as the standard image for defining the scale (Figure 3.10). Compared with other regions, the satellite image of Zurich Lake is clearer and is very suitable as a standard for measuring the length of small boats.

After measurement, the vessel in Figure 3.10 with a YOLO length of 0.434517 is actually 55.17m. Then, with an eye altitude of 200 metres, the ratio of the real length to the YOLO length is approximately 127. Finally, after several verifications, this ratio is within the margin of error and can be used as a scaling ratio. And having the same ratio and the same eye altitude is not enough. We also need to keep the resolution of each image the same as well. For this purpose, all datasets involved in the detection will maintain a resolution of 3840x2160 pixels.

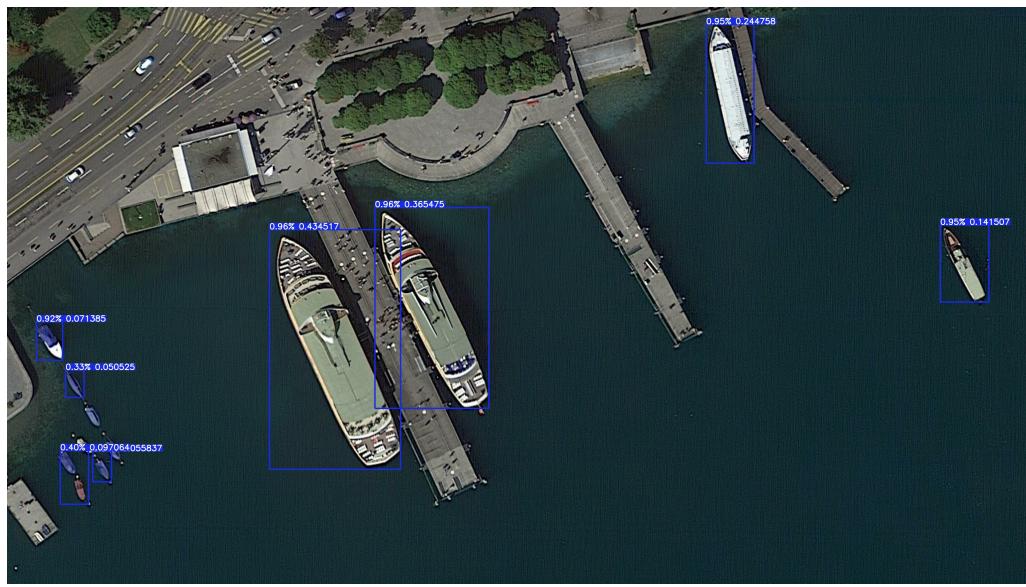


Figure 3.10: An image from Google Earth Pro for Zurich Lake on 16 Aug 2018 when eye alt is 200 meters.

3.2.4 Classification of Small Boats and Big Boats

In a certain sense, large vessels (e.g. cargo ships or warships) and small vessels (e.g. small boats for domestic use for recreation and fishing) are distinguished when creating the dataset for the area. However, due to the scaling, the eye altitude of the image is only 200 meters. Therefore, some large vessels, such as general cargo ships, do not even appear fully in an image with a resolution of 3840x2160 pixels (such as Figure 3.11), i.e., they do not appear in the statistical results. On the other hand, some of the larger vessels, slightly shorter in length, are identified correctly by the algorithm and counted as part of the number of large vessels in the area (Figure 3.12).



Figure 3.11: A cargo (March 2021, Guaymas) overflowed the image and will be excluded from the statistics.

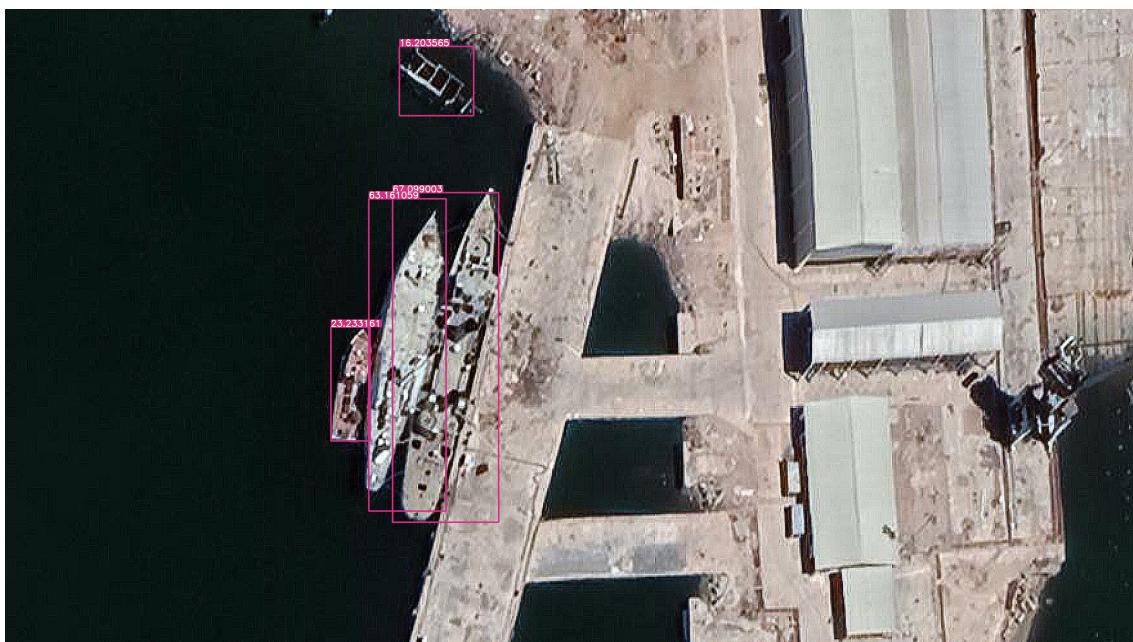


Figure 3.12: The algorithm can detect all the boats (May 2021, Guaymas) well if no big boats overflowed the image.

3.2.5 Classification of Entertainment Boats and Fishing Boats

After distinguishing between large and small boats, it is necessary to distinguish between small recreational boats for domestic use and fishing boats without a roof. In this thesis, whether or not the colour shown on satellite images of detected small boats is mostly white becomes a feature that distinguishes recreational boats from fishing boats. To do this, I need to find the range of small boats previously identified by the algorithm, i.e. the four coordinates of the anchor box. Then, analyse whether the colour within the anchor frame is white or close to white. Finally, the number of white boats is counted.

3.3 Workflow

The flow chart from the training data to the final statistics made is shown in Figure 3.13. When training the object detection model, Tesla P100D GPU was used provided by Google Colab. Google Earth Pro was the platform for creating the data set of the Gulf of California. Python package *pandas*, *numpy*, and *PIL* were used to do the statistics work. Python wrote all programs in this thesis.

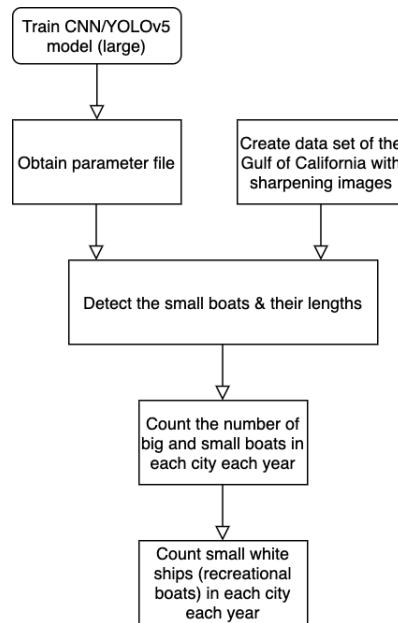


Figure 3.13: The workflow of detection algorithm.

Chapter 4

Results

4.1 Train Custom Data: Weights, Biases Logging, Local Logging

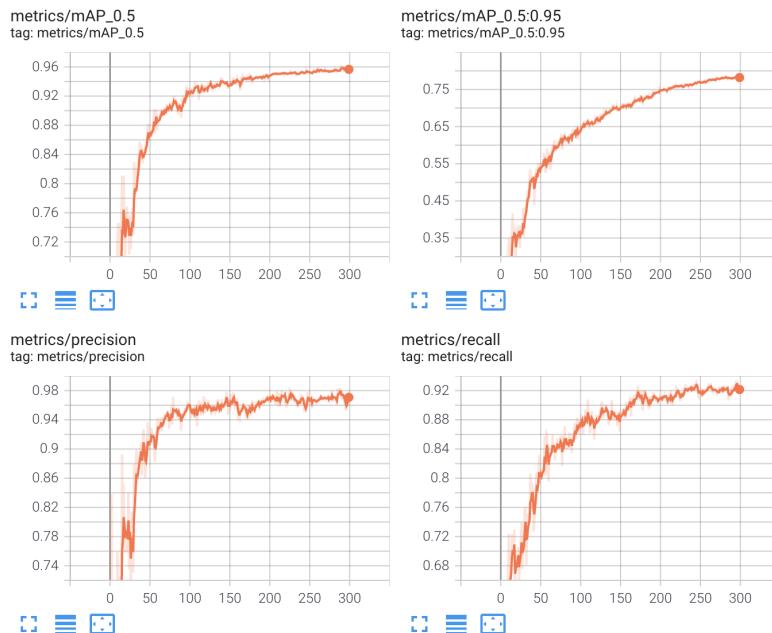


Figure 4.1: Average model precision when IOU is larger than 0.5; Average model precision when IOU is between 0.5 and 0.95; The model precision; The model recall rate.

As described in Sec 3.1, 694 satellite images with a resolution of 4800x2908 were rescaled to a size of 416x416. This can greatly improve the efficiency of training the model. It is a fact that with the training of a model with 47 million parameters, the precision and recall of the training results improved significantly with the number of training sessions.

As shown in Figure 4.1 above, the average accuracy of the model, the precision of the model and the recall of the model all show a significant increase with the number of times the model is trained when the IOU¹ (Intersection over Union) is between 0.5 and 0.95. In particular, the precision of the model can eventually reach a level close to 98%. However, this does not necessarily mean that the model will also fit satellite imagery of the Gulf of California. First, such high accuracy results only tell us that the model can achieve a relatively high recognition accuracy, which gradually increases and eventually reaches 98% after 300 training repetitions. If the algorithm needs to be trained for this area, then consideration needs to be given to purposefully selecting many small boats in or near the area as a source of data for training the model. Secondly, there is also an inequity in the level of resolution of the training and test data.

To train the model faster, I reduced the resolution of the images by a factor of about 70 (from 4800x2908 pixels to 416x416 pixels). Otherwise, the training could take half a month if used images with a resolution of 4800x2908 pixels as the data source! However, each image from the Gulf of California was maintained at a resolution of 4K (3840x2160 pixels). This would result in the algorithm not framing the target perfectly when it detects it, i.e., it would not be perfectly tangential to the edges of the target.

Similarly, as shown in Figure 4.2 below, the loss rate of the box can eventually reach 1% as the number of training sessions increases. Since this thesis has defined only one class of object, i.e. boat, this means that the probability that the detection box does not detect that it is a boat at all is 1%. Similarly, because there is only one class, the class loss rate is 0. Figure 4.3 below shows the prediction results during the training of the model, it can be seen that the model is able to detect the occurrence of boats 100% of the range tested and gives the corresponding range box. Most of the detected boats were mostly considered to have a 90% probability of being boats. Since only one class was set, some were also

¹Intersection over Union is an evaluation metric used to measure the accuracy of an object detector on a particular dataset.

considered to have a 100% probability of being boats.

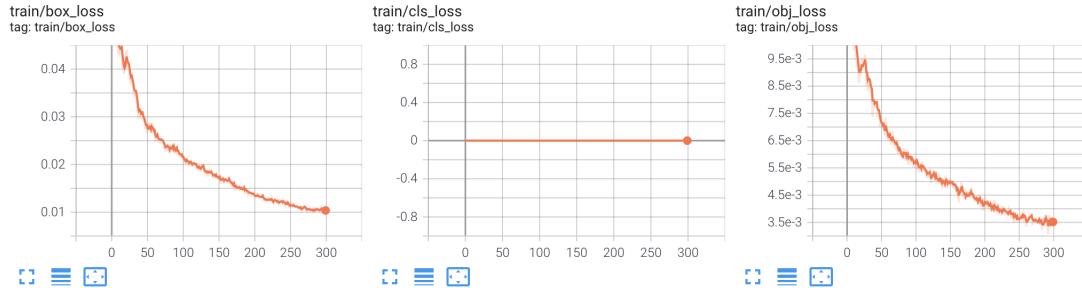


Figure 4.2: The box loss rate of the model; The class loss rate of the model; The object loss rate of the model.

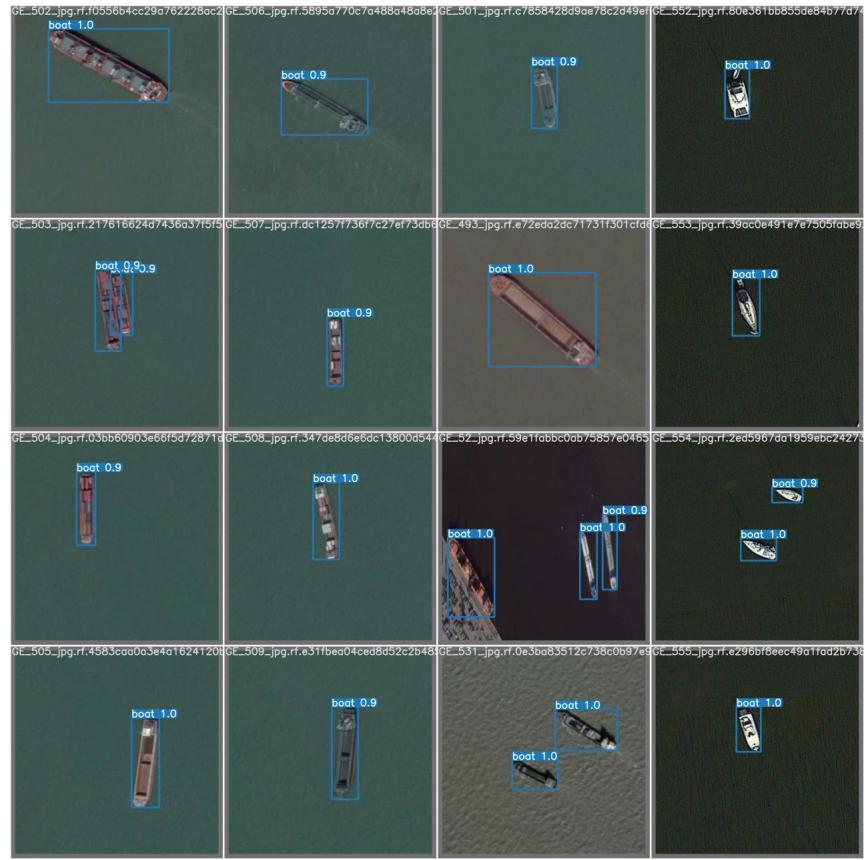


Figure 4.3: Test result of a trained model for detecting ships.

4.2 Detection Results and Vessel Composition

In Sec 3.2.3, I discussed the method of detecting small boats. Most for the result, I will test the length of two small boats on a certain sea. The results are shown in Figure 4.4. A smaller boat shown in a Google Earth Pro image measured 6.98 meters, while the algorithm detected a length of 6.744 meters. The error between them is 3.38%. The other larger boat Google Earth Pro measured 41.38 meters in length, and the algorithm detected 40.984 meters. The error between them is 0.96%. After testing, the length of the boat detected by the algorithm was within an acceptable error range.

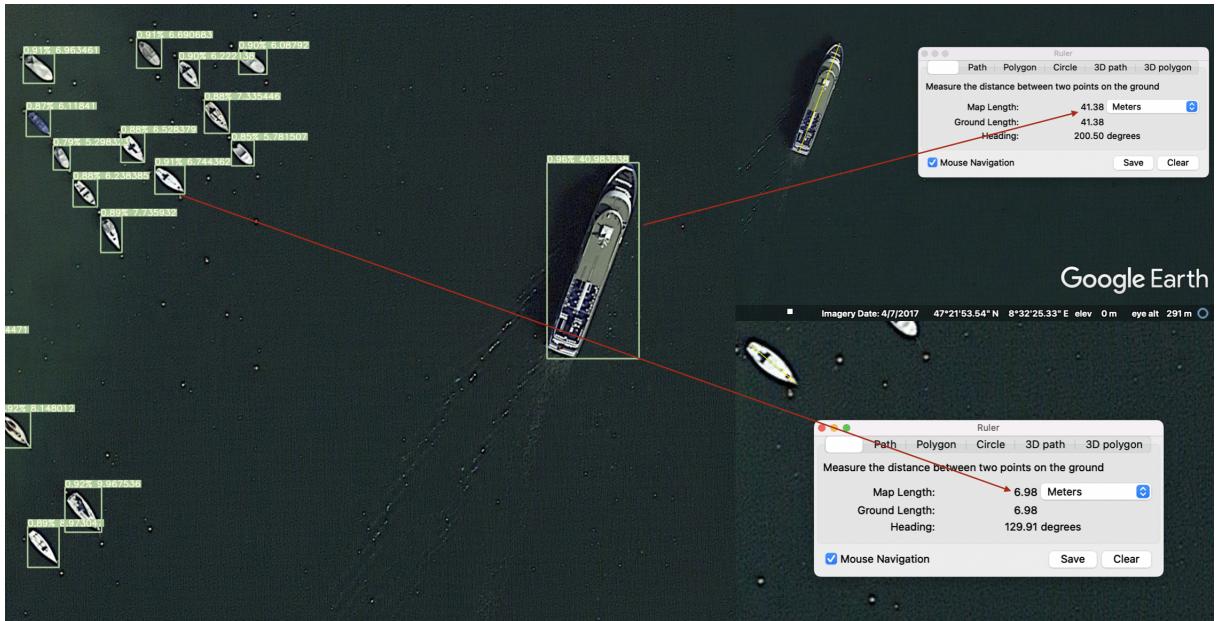


Figure 4.4: Test the length of boat.

As explained in Sec 3.2.1, it was unwise to select the entire region for the study due to the uneven amount of data that Google Earth Pro has provided for the Gulf of California over the past three years. Therefore, three areas with more data were chosen: Santa Rosalia, Loreto, and Guaymas. Ultimately, satellite images of these three areas were found for 2019, 2020, and 2021, for a total of 690 images. Each of the 690 images is a satellite image with an eye altitude of 200 metres and a resolution of 3840x2160 pixels.

However, as stated in Sec 3.2.2, some of the slightly earlier satellite images had inferior

detail representation capabilities, which resulted in the model not being very good at accurately detecting the features of the target, so that many small boats in the Gulf of California could not actually be detected. However, after enhancement with the 5x5 sharpening kernel, the recognition rate of the model was significantly improved. However, the following situations still occur.

- (a) Figure 4.5: When the detailed representation of the image is indigent and two or three small boats are moored together, the model is very likely to recognise the two or three boats as a whole. There are two reasons for this problem. One is that the training data is mostly a ‘fuzzy’ data source. Thus, when there are two or three small boats moored together, the model cannot easily detect the features of each small boat individually. In contrast, it may seem more reasonable to the model that the two or three boats as a whole have the same features. The second reason is that most data sources are individual boats on the surface or boats docked close to each other. As the data sources do not fully consider the fuzzy nature of the detail needed to detect the data and the fact that they are too close together, the model naturally does not recognise such cases.
- (b) Figure 4.6: When a cargo ship is moored on the shore, the ship and cargo appear as a ‘rectangle’ from the sky, much like a long jetty, and therefore sometimes cannot be detected. This is because small ships with a rectangular shape are not common in the preparation of data sources. This also applies to uncommon vessels such as battleships.
- (c) Figure 4.7: The recognition rate is also significantly lower when the boat is sometimes parked on the beach rather than on the water. This is also because most of the data was based on data on the water rather than on boats on the beach when the model was trained.

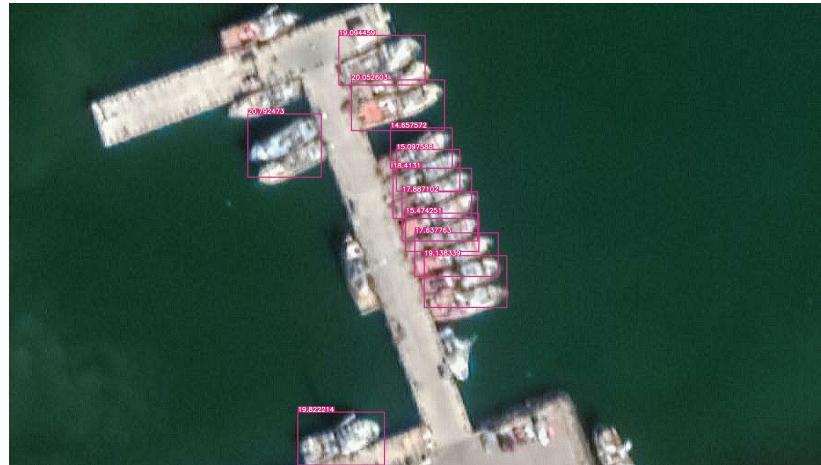


Figure 4.5: When small boats are moored closely together in the harbour, the model may recognise two small boats as one. Image is from Guaymas, Jan 2020.

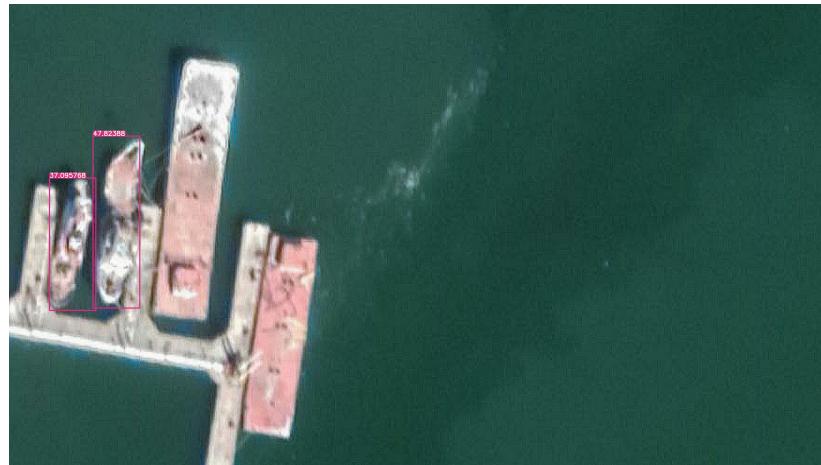


Figure 4.6: When the cargo ship is full of cargo, the ship looks like a rectangular jetty from above and loses the normal shape of a ship. Image is from Guaymas, Jan 2020.

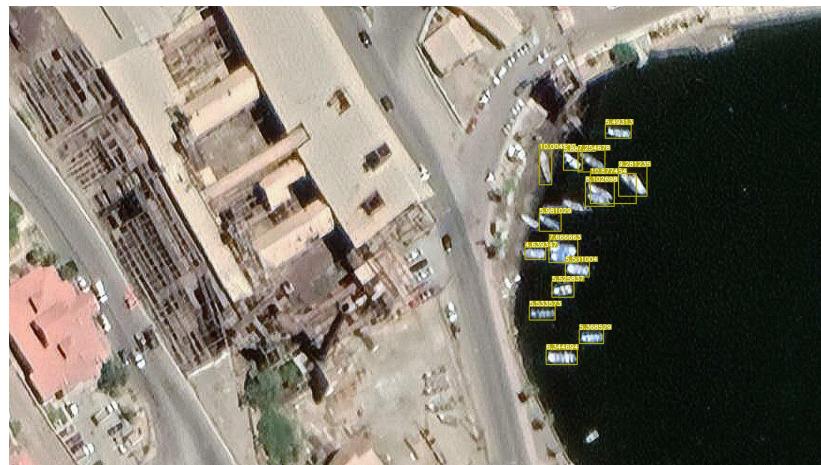


Figure 4.7: Models may have difficulty detecting small boats moored on the beach. Image is from SantaRosalia, Feb 2021.

Nevertheless, as Figure 4.5, Figure 4.6, Figure 4.7 demonstrate, the model still detects most of the small boats in the poorly detailed satellite images. A Python script was then designed to count the number of small and large boats between regions. The results (Table 4.1 to Table 4.9) are shown below.

	Jan	Feb	Mar	Apr	Jun	July	Sep	Nov
Boats length <24 meters	115	83	99	131	82	107	90	98
Boats length \geq 24 meters	31	24	21	24	22	22	21	19

Table 4.1: Small ships and large ships in Guaymas in 2019.

	Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep	Nov	Dec
Boats length <24 meters	63	101	159	52	145	85	66	121	28	74	113
Boats length \geq 24 meters	18	13	15	13	21	4	13	22	15	28	15

Table 4.2: Small ships and large ships in Guaymas in 2020.

	Feb	Mar	Apr	May
Boats length <24 meters	161	167	144	119
Boats length \geq 24 meters	19	33	21	22

Table 4.3: Small ships and large ships in Guaymas in 2021.

	Jan	Apr	Jun	July	Aug	Sep	Nov	Dec
Boats length <24 meters	40	60	18	49	33	42	38	17
Boats length \geq 24 meters	0	1	0	0	1	0	1	0

Table 4.4: Small ships and large ships in Loreto in 2019.

	Jan	Feb	Mar	Apr	Jun	July	Aug	Sep	Nov
Boats length <24 meters	18	26	18	30	32	33	20	43	43
Boats length \geq 24 meters	0	0	0	0	0	0	0	1	1

Table 4.5: Small ships and large ships in Loreto in 2020.

	Jan	Feb	Apr
Boats length <24 meters	16	33	33
Boats length \geq 24 meters	0	0	0

Table 4.6: Small ships and large ships in Loreto in 2021.

	Jan	July	Aug	Sep	Dec
Boats length <24 meters	27	41	36	29	22
Boats length \geq 24 meters	0	2	1	0	2

Table 4.7: Small ships and large ships in Santa Rosalia in 2019.

	Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep	Nov	Dec
Boats length <24 meters	50	43	37	48	8	44	66	58	58	43	57
Boats length \geq 24 meters	1	1	1	1	0	1	3	1	1	1	1

Table 4.8: Small ships and large ships in Santa Rosalia in 2020.

	Apr	May
Boats length <24 meters	32	46
Boats length \geq 24 meters	1	2

Table 4.9: Small ships and large ships in Santa Rosalia in 2021.

As can be seen from the nine tables above, Guaymas, Loreto and Santa Rosalia can be classed by two different types of port cities:

1. Loreto and Santa Rosalia have a much smaller number of ships and almost no large ships larger than 24 metres in length.
2. The port of Guaymas was probably much larger than that of Loreto and Santa Rosalia. There are about four times more small ships in Guaymas than in Loreto and about three times more small ships than in Santa Rosalia. The port of Guaymas has around 20 large ships per day, while Loreto has almost no large ships, and Santa Rosalia probably has at least one large ship per day.

In addition to the number of ships in each port, another interesting finding is that the number of ships in Guaymas's port, in similar seasons, increases compared to the numbers in the previous year. In February, for example, the number of ships in 2019 is less than the number of ships in 2020, and the number of ships in 2020 is less than the number of ships in 2021. The reason for this could be the rapid development of the maritime industry in Guaymas in just three years. However, another potential reason is that the quality of

satellite imagery in Guaymas in 2021 is much better than in 2020, resulting in almost 50% more vessels being detected by the model in 2021 than in 2020.

Similarly, the quality of satellite imagery in the Guaymas area in 2020 is also better than in 2020, resulting in almost 25% more ships being detected by the model in 2020 than in 2019. The same thing happened in Santa Rosalia. For instance, in July, there were nearly 50% more small ships in Santa Rosalia in 2020 than in 2019.

However, this does not always happen. In fact, as the satellite images of Loreto have not improved significantly in the last three years, such a conclusion does not hold. Besides, it is also difficult to encapsulate the 2021 data in this conclusion due to the large blank in the 2021 Santa Rosalia data.

4.3 Entertainment Boats in the Gulf of California

According to the statement in Sec 3.2.5, determining whether a boat is white or not can be used as a criterion to determine whether a boat is a small boat for family recreation or a fishing boat. Therefore, a Python script was designed to count the small white boats. The results are shown in the tables below.

	Jan	Feb	Mar	Apr	Jun	July	Sep	Nov
Boats length <24 meters	115	83	99	131	82	107	90	98
Boats length \geq 24 meters	31	24	21	24	22	22	21	19
Small White Boats	112	83	99	131	79	105	89	96

Table 4.10: Small white boats in Guaymas in 2019.

	Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep	Nov	Dec
Boats length <24 meters	63	101	159	52	145	85	66	121	28	74	113
Boats length \geq 24 meters	18	13	15	13	21	4	13	22	15	28	15
Small White Boats	63	101	138	52	142	79	66	119	25	66	108

Table 4.11: Small white boats in Guaymas in 2020.

	Feb	Mar	Apr	May
Boats length <24 meters	161	167	144	119
Boats length \geq 24 meters	19	33	21	22
Small White Boats	147	158	144	106

Table 4.12: Small white boats in Guaymas in 2021.

	Jan	Apr	Jun	July	Aug	Sep	Nov	Dec
Boats length <24 meters	40	60	18	49	33	42	38	17
Boats length \geq 24 meters	0	1	0	0	1	0	1	0
Small White Boats	40	55	18	49	33	41	37	16

Table 4.13: Small white boats in Loreto in 2019.

	Jan	Feb	Mar	Apr	Jun	July	Aug	Sep	Nov
Boats length <24 meters	18	26	18	30	32	33	20	43	43
Boats length \geq 24 meters	0	0	0	0	0	0	0	1	1
Small White Boats	18	26	17	28	30	32	20	43	39

Table 4.14: Small white boats in Loreto 2020.

	Jan	Feb	Apr
Boats length <24 meters	16	33	33
Boats length \geq 24 meters	0	0	0
Small White Boats	16	33	33

Table 4.15: Small white boats in Loreto in 2021.

	Jan	July	Aug	Sep	Dec
Boats length <24 meters	27	41	36	29	22
Boats length \geq 24 meters	0	2	1	0	2
Small White Boats	25	41	35	29	22

Table 4.16: Small white boats in Santa Rosalia in 2019.

	Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep	Nov	Dec
Boats length <24 meters	50	43	37	48	8	44	66	58	58	43	57
Boats length \geq 24 meters	1	1	1	1	0	1	3	1	1	1	1
Small White Boats	37	42	37	43	6	42	65	54	58	41	56

Table 4.17: Small white boats in Santa Rosalia 2020.

	Apr	May
Boats length <24 meters	32	46
Boats length \geq 24 meters	1	2
Small White Boats	32	46

Table 4.18: Small white boats in Santa Rosalia in 2021.

As can be seen from the table above, most of the boats in the three regions are white, i.e., all can be classified as recreational boats for domestic use. However, this conclusion is imperfect. The algorithm does not consider the uncertainty of future work, for example, black leisure boats. Therefore, it is necessary to include an appropriate sensitivity analysis in the algorithm described above. Although there are many uncertainties in detecting the colour of the boats, the algorithm also takes into account situations where the colour of the boats is not pure white due to atmospheric refraction, weather interference, cloud interference etc., i.e. the algorithm also takes into account situations where the colour of the boats is light. This is still acceptable from the point of view of algorithm complexity, results, and detecting data quality.

Chapter 5

Future Work

Although the location and size of small boats have been successfully identified, and the boats have been successfully but roughly classified into domestic recreational and fishing categories, more work is still needed to refine and improve them. In the future,

1. in determining whether the boats are white or not, I created a function with RGB as the independent variable. I required the output of this function to be greater than a particular value (threshold). However, adjusting the parameters of this function can, in a certain sense, only be based on experience. So, does this threshold accurately identify the boat as being white in colour?
2. if it is possible to roughly determine the type of small boat based on colour alone, is it possible that a dark coloured fishing boat is not detected on dark coloured seas? Would the number of fishing boats, in fact, then be greater?
3. the satellite images capture only a moment in time during the year, are the fishing vessels operating slightly further out to sea and not captured by the satellite due to the nature of the fishing vessels?
4. can data providers provide scholars with clear and close satellite imagery for free? Data is one of the triads of artificial intelligence and also the most overlooked factor. In fact, AI algorithms are highly dependent on data. It is as if the algorithms are always hungry, and all need a constant stream of data to feed their hunger. For object

detection algorithms, poor image detail representation means less high-quality data. For the algorithm, fewer data often brings poor output results. Therefore, both the data used in building the model and the data used for detection should be kept at the highest level. Otherwise, training AI models loses its relative meaning.

5. will Google Earth Pro provide a deep integration tool about the zoom scale in the future? Google Earth Pro provides very many images with details. However, Google Earth Pro does not offer similar integrated tools for zooming rulers as Google Maps does. If such a scaling environment were available, it would mean that every photo taken on Google Earth Pro would contain its scaling ratio. This means that we do not need to define a fixed zoom scale. It would be easy to know the object's length on each screenshot, regardless of whether the eye altitude is 200 meters or not. This leads to the fact that when some giant cargo ships or cruise ships need to be detected, we can reduce the scale to get an overall picture of the giant ship. When some tiny ships (e.g., a small ship under 5 meters in length) need to be inspected, we can zoom in to get the overall shape of the small ship.
6. will small boats have a more precise classification in the future? Using colour to distinguish whether a boat is a recreational boat or a fishing boat may not be a solution that is acceptable to everyone. However, to solve this problem with a purely deep learning approach, one needs to train the data with all the types of boats that one wants to detect. However, doing the data classification and training in a realistic environment, i.e. Google Earth Pro, would be labour-intensive and costly. Furthermore, when looking for the classification category and the data under this category, is it possible to guarantee that the object's environment is the same as the object's environment under other categories due to the AI fairness principle? For example, is it possible to guarantee that the proportion of large cruise ships appearing on the beachside is the same as that of small recreational boats appearing on the beachside? The reason for this is that we do not want all large cruise ships to be offshore and all small recreational boats to be on the beach. If this is the case, the deep neural network does not need to know the contours or characteristics of the different boats. It just needs to analyze the background of the boats (i.e., the colour of the sea level, the

colour of the beach, etc.) to predict whether a large cruise ship is likely to be a large cruise ship in the first place. One of the potential solutions to this problem is using computer simulation techniques to create CAD models of various ships. These CAD ships are then placed evenly into a variety of different background environments. Finally, these data are trained with deep neural networks, and then the trained models are put into real-life situations to detect some real ships.

7. will it be possible in the future to quickly scale up the algorithm to other regions? As well, will it be possible to analyze real-time satellite video in the future? Again, the key to expanding the algorithm to other regions and analyzing real-time satellite video is to have fast and timely access to the data in the region. Frankly, it is whether the data providers can provide the data in the region. With real-time satellite video, it will be possible to get a better picture of maritime traffic in any area to monitor and control the carbon emissions of shipping in that area.
8. is it possible to add more recognition objects, such as harbour and sea ripples, to prevent the recognition of harbour or sea ripples as small boats? The purpose of this is that once the algorithm can successfully identify the harbour or sea ripples, they will not be identified as small boats.
9. can the timeframes be better? It is a question of optimization, a trade-off question on cloud storage, a question of the algorithm's speed, and a question of the accuracy of the result. Besides, can the algorithm distinguish the same boat?
10. Can we get a better idea of the type of engine on the boat? The connection of the ship is rightly allocated but still possible that it is difficult to know what type of engine or power is installed onboard. For this, a better understanding between the ship size and type and the levels of power needed is needed. This suggestion opens a new door of research.

Chapter 6

Conclusions

In this thesis, I introduce the reader to satellite image recognition and how I can make a modest contribution to the development of the field.

In Chapter 2, I review the traditional research literature on estimating the number of small boats. Researchers use kernel density estimation to distribute population and vessel numbers and ultimately show that their predictions can accurately predict the number of vessels in the Gulf fisheries. Some scholars then focused on describing the combination of national small vessel registries and surveyed vessel discharge results. However, the methodology does not cover states without a national small boat registry, and small boats are not just recreational vessels.

Obviously, these traditional approaches do not extract information well from a large amount of data. Although as early as the 1990s, researchers such as Yann LeCun realized that convolutional neural networks could recognize images. However, with the exponential growth in computing power in the last five years, scientists and engineers have only been able to apply convolutional neural networks to topics such as large-scale image recognition and object detection. Since 2012, the field has been reinvented by creating large-scale supervised datasets and developing neural object detection network models. Although it has only been nine years since then, the field has evolved at an amazing rate. Innovations in building better datasets and more effective models have alternated, and both have contributed to the field's growth.

In Chapter 3, I first explain how convolutional neural networks can detect a letter ‘X’

and describe the mathematical principles involved. Next, I introduce the two most commonly used neural network frameworks today: Faster RCNN and YOLO and cite Dwivedi's work to illustrate to the reader the advantages of the YOLO model in recognizing small objects advantage of higher accuracy in recognizing objects in videos. Next, I show the data needed in training the model and the preparation needed to train the model faster. Finally, to speed up the model's training, I use a GPU to train the model and use Google Colab, Google Drive and GitHub as my testing and development tools.

In the subsequent section, I highlight the inequality in the amount of data available on satellite images for the Gulf of California. Thus, I took the approach of selecting one of the cities with a large amount of satellite data every year and then analyzing it. However, even this does not avoid the fact that the quality of the satellite images in 2019 is worse compared to the following two years. In response, I took the approach of sharpening the images to bring out the details of the images. As a result, the model can improve the object recognition rate by around 26%.

Then, I designed algorithms to detect the length of small boats. Since I did not have access to the zoom scale of Google Earth Pro, I had to define a zoom scale myself to find the relationship between the length of the real boat and the boat in the image. Finally, by using the same eye altitude and resolution for all the photos in the test dataset, the algorithm can automatically detect the boat's length. In the test, the boat's length detected by the algorithm is similar to the length measured from Google Earth Pro.

In Chapter 4, I show the detection results that can combine open high-resolution satellite data and convolutional neural networks. The results are not as high as the ideal recognition rate, but such detection results are still acceptable due to the inferior quality of the monitoring data.

The results show a divergence. In 2019, there were on average 100.63 small boats in Guaymas. Although this value dropped to 91.55 in 2020, it quickly rises again to 147.75 in 2021. Similarly, in 2019, 2020 and 2021, there are 31, 46.55 and 39 small boats in Santa Rosalia, respectively. However, Loreto did not show such an upward trend in general. In 2019, 2020 and 2021, there are 37.13, 29.22 and 27.33 small boats in Loreto, respectively. On the one hand, there is a large port like Guaymas with many small ships. On the other hand, small ports like Loreto and Santa Rosalia have almost no big ships and fewer small

ships. But it is interesting to note that small boats in Guaymas, Loreto and Santa Rosalia are almost certainly small boats for family use for recreation. The analysis of satellite images shows that most small boats are docked in the harbour, which means it is rare to see a large number of small boats floating on the sea. Another interesting point is that the number of small boats identified increases as the year increases. Considering that most of the boats are in the harbour and that the models used for the tests are identical, the only difference is that the detail of the satellite images is better realizable each year, i.e., each year we have an image with higher quality. Then, it is reasonable to believe that the increase in the quality of satellite photos can improve the quality of object detection.

All in all, I am very excited about the progress I have made in this area over the past four months and are pleased to contribute to the field. At the same time, I am convinced that there is still a long way to get AI to the point where it can detect objects beyond the human level, and we still face huge challenges and many outstanding questions that need to be addressed in the future. One key challenge is that we still do not have a good way to handle deeper object detection — those problems that require understanding photos or video inference — for example, the question of whether the two boats are one boat. In the future, we will also have to address the complex problem of integrating object detection and natural language processing to reach a level that allows AI to understand video.

We also hope to encourage more researchers to work on applying satellite image detection to new areas. We believe it will lead us to build better agents that can understand media and hope to see these ideas implemented and developed in industry applications.

Bibliography

- BP. 2016. Statistical review of world energy 2016. *BP Statistical Review, London, UK*.
- BP. 2021. Statistical review of world energy 2021. *BP Statistical Review, London, UK*.
- Leila Character, Agustin Ortiz Jr, Tim Beach, and Sheryl Luzzadder-Beach. 2021. Archaeologic machine learning for shipwreck detection using lidar and sonar. *Remote Sensing*, 13(9):1759.
- Shantayanan Devarajan and Anthony C Fisher. 1981. Hotelling’s” economics of exhaustible resources”: Fifty years later. *Journal of Economic Literature*, 19(1):65–73.
- Priya Dwivedi. 2020. Yolov5 compared to faster rcnn. who wins? <https://towardsdatascience.com/yolov5-compared-to-faster-rcnn-who-wins-a771cd6c9fb4>. Accessed: 2021-08-30.
- Kunihiro Fukushima and Sei Miyake. 1982. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. In *Competition and cooperation in neural nets*, pages 267–285. Springer.
- Ross Girshick. 2015. Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 1440–1448.
- Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 580–587.

- Krista Greer, Dirk Zeller, Jessika Woroniak, Angie Coulter, Maeve Winchester, M.L. Deng Palomares, and Daniel Pauly. 2019. Global trends in carbon dioxide (co₂) emissions from fuel combustion in marine fisheries from 1950 to 2016. *Marine Policy*, 107:103382.
- Weiwei Han, W. Yang, and Suixiang Gao. 2016. Real-time identification and tracking of emission from vessels based on automatic identification system data. *2016 13th International Conference on Service Systems and Service Management (ICSSSM)*, pages 1–6.
- Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. 2017. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969.
- Kaiming He, Jian Sun, and X. Tang. 2009. Single image haze removal using dark channel prior. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1956–1963.
- T. Hensel, C. Ugé, and C. Jahn. 2020. Green shipping: using ais data to assess global emissions. *Sustainability Management Forum — NachhaltigkeitsManagementForum*, 28:39–47.
- VA Hernández-Trejo, G Ponce-Díaz Germán, D Lluch-Belda, and LF Beltrán-Morales. 2012. Economic benefits of sport fishing in e relative los cabos, mexico: is the abundance a determinant. *Journal of Sustainable Tourism*, 161:165–176.
- Jonathan Huang, Vivek Rathod, Chen Sun, Menglong Zhu, Anoop Korattikara, Alireza Fathi, Ian Fischer, Zbigniew Wojna, Yang Song, Sergio Guadarrama, et al. 2017. Speed/accuracy trade-offs for modern convolutional object detectors. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7310–7311.
- David H Hubel and Torsten N Wiesel. 1962. Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. *The Journal of physiology*, 160(1):106–154.
- IMO. 2021. Fourth greenhouse gas study 2020.
- Glenn Jocher. 2020. ultralytics/yolov5: v3.1 - Bug Fixes and Performance Improvements. <https://github.com/ultralytics/yolov5>.

- L. Johansson, J. Jalkanen, E. Fridell, I. Maljutenko, Erik Ytreberg, M. Eriksson, E. Roth, and Vivian Fischer. 2018. Modeling of leisure craft emissions.
- L. Johansson, J. Jalkanen, and J. Kukkonen. 2016. A comprehensive modelling approach for the assessment of global shipping emissions.
- Andrew Frederick Johnson, Marcia Moreno-Báez, Alfredo Giron-Nava, Julia Corominas, and Brad Erisman. Exequiel ezcurra, octavio aburto-oropeza. 2017. a spatial method to calculate small-scale fisheries effort in data poor scenarios. *Plos One*, 12(4):e0174064.
- Andrew Frederick Johnson, Marcia Moreno-Báez, Alfredo Giron-Nava, Julia Corominas, Brad Erisman, Exequiel Ezcurra, and Octavio Aburto-Oropeza. 2017. A spatial method to calculate small-scale fisheries effort in data poor scenarios. *PLoS One*, 12(4):e0174064.
- Tae-Hwan Joung, Seong-Gil Kang, Jong-Kap Lee, and Junkeon Ahn. 2020. The imo initial strategy for reducing greenhouse gas (ghg) emissions, and its follow-up actions towards 2050. *Journal of International Maritime Safety, Environmental Affairs, and Shipping*, 4(1):1–7.
- Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2012. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105.
- Yann LeCun, Yoshua Bengio, et al. 1995. Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10):1995.
- Salvador E. Lluch-Cota, Eugenio A. Aragón-Noriega, Francisco Arreguín-Sánchez, David Aurioles-Gamboa, J. Jesús Bautista-Romero, Richard C. Brusca, Rafael Cervantes-Duarte, Roberto Cortés-Altamirano, Pablo Del-Monte-Luna, Alfonso Esquivel-Herrera, Guillermo Fernández, Michel E. Hendrickx, Sergio Hernández-Vázquez, Hugo Herrera-Cervantes, Mati Kahru, Miguel Lavín, Daniel Lluch-Belda, Daniel B. Lluch-Cota, Juana López-Martínez, Silvio G. Marinone, Manuel O. Nevárez-Martínez, Sofia Ortega-García, Eduardo Palacios-Castro, Alejandro Parés-Sierra, Germán Ponce-Díaz, Mauricio

- Ramírez-Rodríguez, Cesar A. Salinas-Zavala, Richard A. Schwartzlose, and Arturo P. Sierra-Beltrán. 2007. The gulf of california: Review of ecosystem status and sustainability challenges. *Progress in Oceanography*, 73(1):1–26.
- Tom Lutherborrow, Amgad Agoub, and Martin Kada. 2018. Ship detection in satellite imagery via convolutional neural networks.
- Catalina López-Sagástegui, Ismael Mascareñas-Osorio, Brad Erisman, Marcia Moreno-Báez, Victoria Jiménez-Esquivel, and Octavio Aburto-Oropeza. 2017. Comparing two fishing communities in the upper gulf of california. <https://datamares.org/stories/comparing-two-fishing-communities-in-the-upper-gulf-of-california/?lang=comparing-two-fishing-communities-in-the-upper-gulf-of-california>. Accessed: 2021-08-30.
- A. Mabunda, Abdelali Astito, and S. Hamdoune. 2014. Estimating carbon dioxide and particulate matter emissions from ships using automatic identification system data. *International Journal of Computer Applications*, 88:27–31.
- S.G. Marinone. 2012. Seasonal surface connectivity in the gulf of california. *Estuarine, Coastal and Shelf Science*, 100:133–141. Recent advances in biogeochemistry of coastal seas and continental shelves.
- Warren S McCulloch and Walter Pitts. 1943. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133.
- Lorayne Meltzer and Jason Oliver Chang. 2006. Export market influence on the development of the pacific shrimp fishery of sonora, mexico. *Ocean & Coastal Management*, 49(3):222–235.
- Mexico National Aquaculture and Fisheries Commission. 2021. Registered vessels. <https://www.conapesca.gob.mx/wb/cona/embarcacionesRegistradas>. Accessed: 2021-08-30.

- Adrian Munguia-Vega, Alison L Green, Alvin N Suarez-Castillo, Maria Jose Espinosa-Romero, Octavio Aburto-Oropeza, Andrés M Cisneros-Montemayor, Gabriela Cruz-Piñón, Gustavo Danemann, Alfredo Giron-Nava, Ollin Gonzalez-Cuellar, et al. 2018. Ecological guidelines for designing networks of marine reserves in the unique bio-physical environment of the gulf of California. *Reviews in Fish Biology and Fisheries*, 28(4):749–776.
- R. K. Pachauri, M. R. Allen, V. R. Barros, J. Broome, W. Cramer, R. Christ, J. A. Church, L. Clarke, Q. Dahe, P. Dasgupta, N. K. Dubash, O. Edenhofer, I. Elgizouli, C. B. Field, P. Forster, P. Friedlingstein, J. Fuglestvedt, L. Gomez-Echeverri, S. Hallegratte, G. Hegerl, M. Howden, K. Jiang, B. Jimenez Cisneroz, V. Kattsov, H. Lee, K. J. Mach, J. Marotzke, M. D. Mastrandrea, L. Meyer, J. Minx, Y. Mulugetta, K. O'Brien, M. Oppenheimer, J. J. Pereira, R. Pichs-Madruga, G.-K. Plattner, Hans-Otto Pörtner, S. B. Power, B. Preston, N. H. Ravindranath, A. Reisinger, K. Riahi, M. Rusticucci, R. Scholes, K. Seyboth, Y. Sokona, R. Stavins, T. F. Stocker, P. Tsakert, D. van Vuuren, and J.-P. van Yperle. 2014. *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC, Geneva, Switzerland.
- Robert WR Parker, Julia L Blanchard, Caleb Gardner, Bridget S Green, Klaas Hartmann, Peter H Tyedmers, and Reg A Watson. 2018. Fuel use and greenhouse gas emissions of world fisheries. *Nature Climate Change*, 8(4):333–337.
- Camille Parmesan and Gary Yohe. 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421(6918):37.
- Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. *arXiv preprint arXiv:1506.01497*.
- K. Simonyan and Andrew Zisserman. 2015. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556.
- M. Traut, A. Bows, Conor J. Walsh, and R. Wood. 2013. Monitoring shipping emissions via ais data? certainly.

- C. Ugé, Tina Scheidweiler, and C. Jahn. 2020. Estimation of worldwide ship emissions using ais signals. *2020 European Navigation Conference (ENC)*, pages 1–10.
- Jasper RR Uijlings, Koen EA Van De Sande, Theo Gevers, and Arnold WM Smeulders. 2013. Selective search for object recognition. *International journal of computer vision*, 104(2):154–171.
- UK Government. 2021. Operational standards for small vessels. <https://www.gov.uk/operational-standards-for-small-vessels>. Accessed: 2021-08-30.
- UK Ship Register. 2021. Uk small ship registration. <https://www.ukshipregister.co.uk/registration/leisure/>. Accessed: 2021-08-30.
- UNFCCC. 2015. The paris agreement. In *The 21st meeting of the Conference of the Parties*, pages 1–25.
- Yiqi Zhang, J. Fung, J. Chan, and A. Lau. 2019. The significance of incorporating unidentified vessels into ais-based ship emission inventory. *Atmospheric Environment*, 203:102–113.

Step A6 – MSC ESDA Dissertation Ethics Declaration

Statement of Risk Assessment & Ethics Approval Requirements	
Student Candidate Number:	<u>KNCM2</u>
Student Name:	[REDACTED]
Student UCL Email Address:	[REDACTED]
Supervisor Name:	<u>Santiago Suarez De La Fuente</u>
Supervisor UCL Email Address:	<u>santiago.fuente.11@ucl.ac.uk</u>
Dissertation Research Proposal:	<ul style="list-style-type: none"> • Title / Topic: <u>An Image Recognition Algorithm Approach to Distinguishing and Categorising the Small-Boat Fleet in the Gulf of California</u> • Research Question(s) / Aims & Objectives: <u>To develop a machine-learning method that can be trained to distinguish small boats from Google Earth pictures of the gulf's coastline and categorise them by boat type (e.g., fishing, transport or leisure)</u> • Data & source (specify all data to be used; if none, explain why): <u>Google Earth</u> • Method(s) (specify all methods to be used): <u>Image recognition algorithm</u>
I have read and understood Step A1 'Does the research require a Risk Assessment?' and: [DELETE ONE STATEMENT]:	<ul style="list-style-type: none"> • This planned research does NOT require a risk assessment.
I have read and understood Step A2 'Does the research require External research ethics approval?' and: [DELETE ONE STATEMENT]:	<ul style="list-style-type: none"> • This planned research does NOT require external ethics review.
External ethics approval is <u>not required</u> and I have read and understood Step A3 'Is the research Exempt from the need for ethics approval?' and: [DELETE ONE STATEMENT]:	<ul style="list-style-type: none"> • This planned research IS EXEMPT from the need for research ethics approval.
I confirm that:	<ul style="list-style-type: none"> • the information I have provided is accurate to the best of my knowledge. • if the answers to any of these questions changes, I will go through this protocol again.

NEXT STEPS:

- STUDENT: Copy the text of the completed statement above into an email and email it to your supervisor.
- SUPERVISOR: Reply to the email confirming your approval of the completed statement, attaching the form and copying in the Dissertation Tutor (Despina Manouseli).
- STUDENT:
 - Include this A6 Statement as a Dissertation Appendix after you have BLACKED OUT YOUR NAME & EMAIL ADDRESS so the second marker can mark anonymously.

The Dissertation mark sheet asks the second marker whether this form was filled out correctly and, if not, what % mark deduction they recommend.

Deleted: [FILL IN]

Deleted: : [FILL IN]

Deleted: [FILL IN]

Deleted: : [FILL IN]

Deleted: : [FILL IN]

Deleted: [FILL IN]

Deleted: <#>This planned research DOES require a risk assessment and appropriate approval will be secured before undertaking any of the activities covered by the risk assessment. ¶

Deleted: <#>This planned research DOES require external ethics review and appropriate external ethics approval will be secured before the data collection starts. ¶

Deleted: ¶
This planned research is NOT EXEMPT from the need for research ethics approval.

Deleted: The research is not exempt from the need for ethics approval and
I have read and understood Step A4 'Does the research require High Risk ethics approval?' and:
[DELETE ONE STATEMENT]:
This planned research IS deemed high risk and approval from the UCL Research Ethics Committee will be secured before the research starts. ¶
This planned research is NOT deemed high risk. ... [1]