

修士論文概要書

Summary of Master's Thesis

Date of submission: 07 / 27 / 2017 (MM/DD/YYYY)

専攻名（専門分野） Department	Computer Science and Communications Engineering	氏 名 Name	NGENO Kipkemoi Japhet	指 導 教 員 Advisor	Hiroshi Watanabe 印 Seal
研究指導名 Research guidance	Audiovisual Information Processing	学籍番号 Student ID number	CD 5115FG19-8		
研究題目 Title	Research on Ship Classification using Faster Region Convolutional Neural Network for Port Security				

1. Introduction

In our study, we present research work on classification of ships using one of the Deep Learning algorithms for port security. Faster Region Convolution Neural Network (Faster R-CNN) is the choice for our research and is applied to classify various types of ships found in port environment. Faster R-CNN method for classification outperforms many other methods especially for the purpose of automatic classification in a maritime environment. The choice of this method is arrived at after detailed analysis of other state of the art object classification models including the analysis of Faster R-CNN model itself on other types of data. Previously, Faster R-CNN has been applied on the classification of other distinct objects different from the ships.

2. Research objective

This study is influenced by the achievements in object detection approaches based on the rapidly growing field of Deep Learning. Many of these have been attempting to achieve high performance in precision, speed, and efficiency and become more robust. In our research, we first highlight the situations in marine environments including ship monitoring systems, port security system and the challenges faced in marine transport.

From many issues highlighted and the understanding of other approaches, their strengths and weaknesses are vital in deciding on our methodology to obtain improved results.

Having realized the research gap, the purpose of research therefore focuses on ship classification using Faster Region Convolution Neural Network (Faster R-CNN) for port security. The main objectives are as follows:

1. To classify classes of ships using Faster R-CNN for port security
2. To compare performance of Faster R-CNN with other approaches in ship classification
3. To identify area of improvement in future research

3. Methodology

The choice of our approach is made by examining various other approaches which have been used in object classifications including the Faster R-CNN [1] on

other distinct images. The methodology involved dataset preprocessing and writing code to fit to the target attributes of dataset.

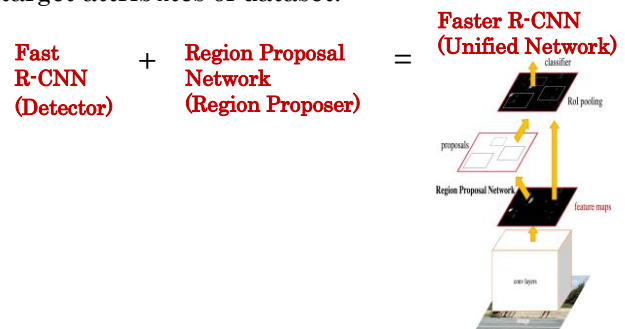


Fig.1 Symbolic illustration of Faster R-CNN.

When all the parameters are set, training and testing followed through two experiments. The aim of the first experiment is to establish the optimum number of images that could be used for the following experiment. The second experiment involves training and testing of the target images of ships for the purpose of classification.

4. Experiment Results and Discussions

The result of our analysis of the performance trends of the recent developments on object classification approaches which included Faster Region Convolution Neural Network model is illustrated in Fig.2. From the analysis, it is evident that the efforts being made to solve image classification problems are experiencing a positive progress in the recent times. The results of the analysis suggests that image classification models are getting better outcome with time.

In the analysis we assumed that the approaches used before the year 2007 has lower precision values while newer models established after 2016 has better performance.

4.1 Experiment 1

The number of images which are suitable for training are selected sequentially while checking on the performance of the approach. By sequentially increasing the number of images at intervals of 50, the AP saturation is achieved at around 400 images as shown in Fig.3.

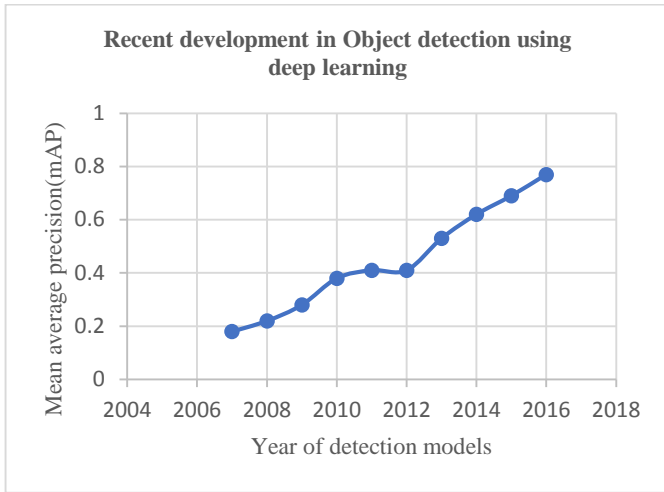


Fig. 2 Object Recognition development using deep learning

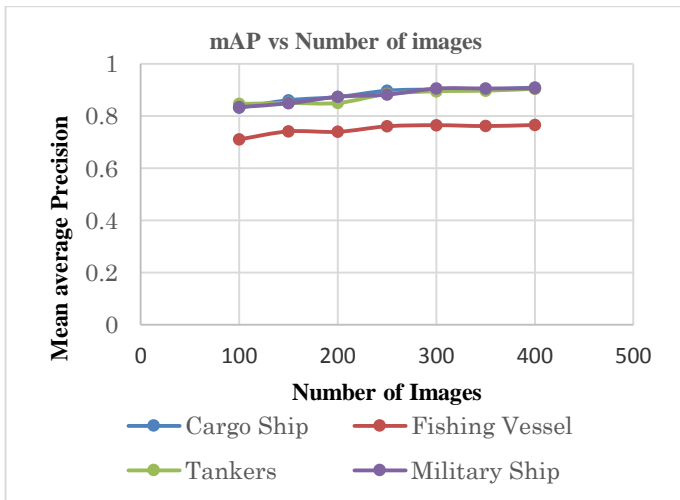


Fig 3. Change in mAP values with the number of training images

4.2 Experiment 2

Based on the experiment 1, we choose 400 images as appropriate number for training and testing. The classification target is 9 types of ships.

Table1. Experiment results showing the Average Precision values of each positively classified images.

Class (IMO Standard Names)	Train Images	Test Images	AP
Container Ship	400	400	0.9091
Fishing Vessel	400	400	0.7659
Tanker	400	400	0.942
Submarine	400	400	0.8896
Tug	400	400	0.8165
Vehicle Carrier	400	400	0.9091
Rescue Ship	400	400	0.8862
Ferry	400	400	0.9091
Destroyer Ship	400	400	0.9069
Mean AP			0.8774



Fig.4 Examples of Output results

The results of the experiment 2 is shown in Table 1. The output results are highly satisfactory and shows mAP 0.8774.

5. Conclusion

It is evident that our approach achieves an average mean precision value of **0.8774** which outperforms other approaches including which have been used in marine vessel classifications. It is quite fast on run time per image compared to Fast R-CNN and CNN methods. In fact, the speed is 25 and 250 times faster than Fast R-CNN and CNN respectively. This may be considered a better classification system for ships which do not need two way communication like the case of AIS. In the future, by incorporating it with the AIS, it may be a suitable method for automatic classification of ships.

Reference

- [1] S. Ren, K. He, R. Girshik and J. Suan, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, 2017.

List of academic achievements

- [1] K. J. Ngeno and H. Watanabe, "Ship Identification for Port Security using Faster Region Convolution Neural Network," in IEICE General Conference, BS-1-12, Nagoya, Japan, March, 2017.
- [2] K. J. Ngeno and H. watanabe, "Ship Classification Using Faster Region Convolution Neural Network (Faster R-CNN) for Automatic Identification of Marine vessels," in Forum on Information Technology, FIT2017, Tokyo, Sep. 2017 (to appear).

**Research on Ship Classification using Faster Region Convolutional
Neural Network for Port Security**

NGENO Kipkemoi Japhet

5115FG19-8

Supervisor

Professor Hiroshi WATANABE

**A Thesis Submitted to the Department of Computer Science and
Communications Engineering,**

the Graduate School of Fundamental Science and Engineering

of Waseda University

in Partial Fulfillment of the Requirements

for the Degree of Master of Engineering

July 27th, 2017

SUMMARY

In our study, we present research work on classification of ships using one of the Deep Learning algorithms for port security. Based on recent analysis of findings on recent researches, the method which was suitable for our objective was Faster Region Convolution Neural Network (Faster R-CNN) which we used it to classify various types of ships found in port environment. Faster R-CNN method for classification outperforms many other methods especially for the purpose of automatic classification in a maritime environment. The choice of this method was arrived at after detailed analysis of other state of the art object classification models including the analysis of Faster R-CNN model itself on other types of data. Previously Faster R-CNN had been applied on the classification of other distinct objects different from ships.

In our approach the classification process was carried out with two experiments using ship images of at least 500 pixels on the shortest side for the purpose of uniformity. Otherwise the approach could take any size of image since its functionality is a scale invariant type of model. The first experiment was useful in determining the appropriate number of images which could give consistent results on correct classification of ships. It also revealed a positive correlation relationship between the number of images and the precision scores of the model. In the second part of the experiment, we used the method to classify the 9 types of ships, which are the target group of ships in the study.

From the results, it is evident that our approach achieved a higher mean precision value which outperforms other approaches which have been used in marine vessel classifications. Apart from the high precision rates, our approach is quite fast on run time per image compared to Fast R-CNN and CNN methods. In fact, the speed is 25 and 250 times faster than Fast R-CNN and CNN respectively. This may be considered a better classification system for ships which do not need two way communication like the case of Automatic Identification System. AIS is a the primary tool for maritime safety for the vessels near the coasts. The AIS equipment is attached on compliant ships and it continuously transmit information about vessel including its, identity, position, course and speed. This system is

only applicable only to compliant vessels. By incorporating Faster R-CNN with the AIS may be a suitable method for automatic classification of ships.

This study has been achieved through various stages which have been presented in form of five chapters. In the beginning various literature was reviewed to identify the research gap and underline the existing challenges facing ship classification methods. In order to arrive at the appropriate approach, the detailed functionalities of the various state of the art methods were critically examined. This was done by examining the architectural designs and doing quality comparison. Based on the results obtained in our approach, the future prospects in the research is illustrated by a conceptual framework of Faster R-CNN-AIS identification system.

Key Words: *Deep learning; Region Proposal Network; Convolution Neural Network; Object Classification; Average Precision; Automatic Identification System.*

DEDICATION

I dedicate my thesis work to my family and many friends. A special feeling of gratitude to my loving wife Roseline and my daughter Claire whose words of encouragement and support have enabled me to successfully undertake the research in harmony despite being far away from them. I express gratitude to all my parents, brothers and my sisters in Kenya who have never forgotten to check on my progress throughout my entire research period in Japan. Finally, I dedicate my work to my young hood friend Wesley Kipkirui “*botum*”. He has consistently kept in touch with me and encouraged me with positive messages throughout my study.

ACKNOWLEDGEMENT

First and foremost, I wish to acknowledge the great contribution of my research supervisor, Professor Hiroshi Watanabe of school of the Fundamental Science and Engineering, Waseda University. I am grateful for his contribution to the development of my thesis right from the conceptualization of the research problem till the end of the research process. He has been quite generous with his expertise and precious time which involved research activities of hours of reflecting, reading, laboratory experiments, and above all being kind and understanding throughout the research period. Moreover, he has ensured that I had a chance to do an after-graduation internship in Japanese company which will expose me more on advanced technological research skills

Secondly, I acknowledge Hideaki Yanagisawa, Ph.D. candidate of Department of Computer Science and Communications Engineering, Hiroshi Watanabe Laboratory, Waseda University, whose guidance and consultations had a direct contribution to the successful completion of the research. I am thankful for the knowledge I acquired from his wide experience and expertise in the area of Deep Learning algorithms.

I also acknowledge the support of Japan International Corporation Agency (JICA) for according me with the opportunity to study in a Japanese university through the African Business Education (ABE) scholarship program. I am thankful for their financial support which has facilitated my study and living in Japan for the entire period of my study. A special thanks to my coordinator from Japan International Corporation Center (JICE), '*Asano San*', who has been monitoring on my study progress on regular basis. I appreciate her words of encouragement which made me feel at home away from home.

Lastly but not least, I acknowledge my employer, the Government of Kenya through the department of the Kenya National Youth Service for nominating me to undertake graduate studies. This made me feel appreciated and recognized as a potential contributor in improving Kenya's economy through research and create business network between Kenya and Japanese companies.

TABLE OF CONTENTS

SUMMARY	ii
DEDICATION	iv
ACKNOWLEDGEMENT	v
TABLE OF FIGURES	viii
LIST OF TABLES.....	ix
CHAPTER 1.....	1
INTRODUCTION	1
1.1 Research Background.....	1
1.1.1 Trending issues on Maritime safety and efficiency	1
1.1.2 Global view on Port security	1
1.1.3 Overview of Automatic Identification System (AIS)	2
1.1.4 Research trends	3
1.1.5 Research Gap.....	4
1.2 Research Objective	6
1.3 Thesis Organization.....	7
CHAPTER 2.....	9
RESEARCH TRENDS.....	9
2.1 Concept overview of Artificial Neural Networks (ANN).....	9
2.1.1 Probabilistic Neural Network (PNN).....	9
2.1.2 Hamming Neural Network (HNN).....	12
2.1.3 Morphological Neural Network	15
2.2 Concept Overview of Convolution Neural Networks (CNNs)	17
2.3 Recurrent Neural Network (RNNs)	18
2.4 Previous researches on ship classification	19
CHAPTER 3.....	22
METHODOLOGY	22
3.1 Proposed Method	22
3.1.1 Region Convolutional Neural Network (R-CNN)	22
3.1.2 Fast Region Convolutional Neural Network (Fast R-CNN).....	23

3.1.3 R-CNN and Fast R- as a unified network of Faster R-CNN	23
3.1.4 Training of Faster R-CNN.....	24
3.2 Experimentation.....	25
3.2.1 Datasets pre-processing	25
CHAPTER 4.....	27
EXPERIMENT RESULTS AND DISCUSSION.....	27
4.1 Analysis on Recent Developments in Object detection approaches	27
4.2 Experiment 1	28
4.3 Experiment 2	30
4.3.1 Sample of Images Used in training and testing for 9 classes.....	31
4.3.2 Sample images of output results	35
CHAPTER 5.....	44
CONCLUSION AND RECOMMENDATION.....	44
5.2 Conclusion	44
5.2 Recommendation	45
5.3 Conceptual framework for Faster R-CNN-AIS ship identification system.....	45
REFERENCES	46
LIST OF ACADEMIC ACHIEVEMENTS	49
Conference Publications.....	49

TABLE OF FIGURES

Fig. 1 Architecture for the Probabilistic Neural Network	11
Fig. 2 Architecture for the Hamming Neural Network.....	12
Fig. 3 Architecture for Morphological Neural Network	16
Fig. 4 CNN architecture model before linking with multiple layers.....	18
Fig. 5 Symbolic structure of Recursive Neural Network [8]	19
Fig. 6 Faster R-CNN as a single, unified network for objection detection.....	23
Fig. 7 RPN framework and examples detection using RPN (from [5])	24
Fig. 8 Object Recognition development using learning	27
Fig. 9 Samples of images used in Preliminary experiment	29
Fig. 10 Change in mAP values with the number of training images	30
Fig. 11 Samples of used in training and testing for 9 classes	32
Fig. 12 Sample results for Container Ship	35
Fig. 13 Sample results for Destroyer Ship.....	36
Fig. 14 Sample results for Ferries	37
Fig. 15 Sample results for Fishing Vessel	38
Fig. 16 Sample results for Rescue Vessel.....	39
Fig. 17 Sample results for Submarine.....	40
Fig. 18 Sample results for Tugs	41
Fig. 19 Sample results for Vehicle Carrier	42
Fig. 20 Sample results for Tanker Ship	43
Fig. 21 Faster R-CNN-AIS conceptual model	45

LIST OF TABLES

Table 1 Properties of PNN	10
Table 2 Properties of HNN	15
Table 3 Properties of MNN.....	16
Table 4 Comparison of R-CNN, Fast R-CNN and Faster R-CNN.....	22
Table 5 Average precision scores for each of 9 class and the overall mean	33
Table 6: why Faster R-CNN is better than other approaches.....	34

CHAPTER 1

INTRODUCTION

1.1 Research Background

1.1.1 Trending issues on Maritime safety and efficiency

Monitoring of ships in a marine environment has become an increasingly serious phenomenon among coastal countries in the world. These countries have been taking steps to improve surveillance and control activities of marine vessels and ports with the aim of solving security and efficiency problems. Maritime shipping activities depend heavily on safety at the sea, ports and voyage routes and has been classified by International Maritime Organization (IMO) as the riskiest industry in the world.

In this case, it has become a general concern to various stakeholders to enhance the culture of safety and ease of operations within the maritime transportation industry and shipping organizations. The safety requirement in sea transport is attracting more new approaches to solve the emerging threats of piracy, illegal fishing, human trafficking, pollution and terrorism. In recent times, there are numerous incidences reported on piracy along the horn of Africa, near the coastlines of embattled country of Somalia which has affected import and export businesses in east and central parts of Africa through Kenyan ports.

1.1.2 Global view on Port security

Concerns on improving security and efficiency on ports across the world have led to various inputs through research and modification of ship monitoring systems. Most of the recent developments have resulted from various maritime conventions and safety regulations with a global perspective. A few examples of these recommendations include International Convention for Safety of Life at Sea (SOLAS), International Convention for Prevention of Pollution of Ships (MARPOL), International Regulations for Preventing Collisions at Sea (COLREG), International Convention on Load Lines (LOADLINE), The International

Management Code for the Safe Operation of Ships and Pollution Prevention (ISM code), and The International Ship and Port Facility Security Code (ISPS code).

Based on these international conventions, recent researches have shifted focus to understanding ship image-based security systems using ship images. This is critical to modern maritime issues surrounding prediction of threats. Ports have been widening in size and the complexity of maritime scenes caused by many vessel activities, waves, small vessel size, and occlusions. For instance, the United States of America is served by 360 commercial ports and is facing real threats in form of smuggled contraband, Water Borne Improvised Explosive Devices (WBIED), and other disruptive actions [1].

In a port environment, ships are monitored from a control area by security personnel. These officers are tasked to keep watch and to recognize situation of ship channels during voyage of own vessels as well as the situation of other ships' behaviors. Understanding ships' behaviors is therefore a fundamental requirement for enhancing security in port environment. However, it is insufficient to rely on human on watchtowers as it will be constrained by two factors. First, human is incapable of making clear observation around the vessels especially from the rare beams of ships. Secondly, human watch relies on visibility. These are the reasons why various technological approaches about detecting, recognizing and classification are employed in a marine environment.

1.1.3 Overview of Automatic Identification System (AIS)

For a long time, Automatic Identification System (AIS) and radar have been the most typical approaches used in monitoring ships with the primary reason of ensuring safety in port and the entire marine navigations by ships. AIS can accurately classify the behavior of compliant ships of which they have AIS installed in them. The International Maritime Organization, long-established that not all ships are installed with the AIS. This means that the ships without AIS installed in them are difficult to identify, and monitoring of their activities and other suspicious behaviors is a challenge.

It is envisaged that in the future the AIS will acquire new capabilities in vessel tracking tools and technology and is expected to give marine transport industry real time solutions to the many maritime issues. The incorporation of newer technologies may lead to moving the sea transport beyond the “only compliant” ships where there is manual aggregation and analysis of data for identification insights and intelligent gatherings. AIS can achieve highly dependable results to the marine transport industry questions surrounding ship monitoring and port traffic management. In this case, incorporation of new technologies for ship classification should improve its operation in providing real time classification potentials. The new applications can bring efficiency in port operations while providing continuous visibility of port and ship activities which in the long run cut management costs and improve on security decisions.

Modernization and expansion of ports is heading towards proactive approaches rather than reactive ways to ship incident management by taking advantage of integration of AIS data with newer approaches. Previously, AIS data were very limited to incident response worldwide where only real-time vessel positions could provide immediate visibility to key vessel information. AIS and newer classification capabilities are becoming an integral part of real time plans and continuous applications which could transform port security systems.

1.1.4 Research trends

There are efforts made in the past through research as an attempt to address the technology gap in ship monitoring and surveillance. Here discussed are a few of them which highlights the research trend in this field.

Xuefeng *et. al.* [2] proposed to use ship behavior recognition algorithm based on video analysis, and verified it with a ship-borne video. Their analysis is based on the behavior of ship's silhouette by considering its size and shape with respect to time. They drew a conclusion that if the size of the observed ship's silhouette changed and the shape did not change, then the focused ship was headed towards the observer. At this point, their conclusion was that the observer treated the approaching ship as a threat and therefore needed to signal

a warning. However, in this approach, they reported a difficulty in meeting their algorithm thresholds which included time interval, field angle of camera, size of unit pixels and the resolution constrains. Jiang et. al. [3] presented a novel method for ship classification by using Synthetic Aperture Radar (SAR) image to distinguish ships based on superstructure scattering features. Their approach comprised of three independent steps which included: ship isolation from the sea, parametric vector estimation and categorization using a support vector machine. The approach tested the classification by using RadarSat-2 images while the ground truth information was obtained from the AIS. The results of the approach showed that accuracy was up higher than the previous state of the earth approaches.

Alex-Net Deep Convolution Neural Networks have been used to classify marine vessel images with different configurations [4]. This was done by measuring the top-1 and top-5 accuracy rates. It involved tuning specific range of vessels which depended on commodity hardware and size of images. This method used a dataset of 130,000 images of maritime vessels and labelled 35 classes. The method registered 80.39% and 95.43% accuracy rates for top-1 and top-5 accuracy rates respectively. In the case of Faster Region Convolutional Neural Network (Faster R-CNN), the ground truth of image source with a varied lighting conditions do not have adverse effect on performance.

1.1.5 Research Gap

Ships activity analysis and identification requires detailed understanding of ship location, movement, and identification system capabilities. To obtain these attributes, there is a need to consider feature based systems which can be used to detect, recognize, and track vessels. Many approaches that have been used previously, have limitations in automatic classifications since most rely on two-way transmission of data. Most problems relate to the limited use of Automatic Identification System (AIS) data which can only be utilized on compliant ship. It is therefore preferable to consider ways which can address such limitations. The continuous security threats in marine environment has triggered many ideas in tackling ship monitoring and surveillance issues.

The gaps in the state of the art approaches are majorly on automatic classifications, the speed of detection and the average precision values.

Our study uses the Faster R-CNN method for ship classification for port security, a method which has been found to have higher performance both in accuracy and runtime comparisons. A comprehensive Faster R-CNN method is an efficient, accurate, and consistent in performance method using Region Proposal Networks (RPNs) in region proposal generation. It is nearly a cost-free method because it operates by sharing convolutional features with the detection networks [5]. In addition, the learned RPN improves the region proposal quality and hence the improved accuracy on object detection accuracy. Our aim is to obtain automatic classification of different types of ship in a port environment.

1.2 Research Objective

This study is influenced by the achievements in object detection approaches based on the rapidly growing field of Deep Learning. Many of these have been attempting to achieve high performance in accuracy, speed, efficiency and become more robust. The highly-cluttered scenes in marine environment are triggering further progress in this sphere of identification and classification of vessels. There is growth in the amount of data and evolving technologies that are dealt with in automatic classification fields. In our study, we focus on understanding the dynamics surrounding ship classification approaches with the view of contributing towards security solutions in ports and marine industry. In this section, we discuss our research objectives and a preview of methodology used in carrying out the experiments.

In our research, we first highlighted the situations in marine environments including ship monitoring systems, port security system and the challenges faced in marine transport. We highlighted the security incidences brought about by the failures in the identification systems in place. We also highlighted emerging new types of incidences which have been shared globally through the International Maritime Organization (IMO) over time. This clarifies the real problems affecting the ports of ships are the key components.

The study covered on research trends related to ship detection, identification and classification with consideration of past, recent, current and to point out the prospects in emerging research contributions. This was a fundamental step to reveal the research gap which would then inform on the decision to take on choosing an appropriate approach. Upon understanding of other approaches, their strengths and weaknesses were vital in deciding on our methodology to obtain improved results.

For the study to be realistic in operations, clear understanding of neural network approaches had to be done. In this study, therefore, the progressive development of Deep Neural Networks from single layer network to Deep Convolution Neural Network was done to ensure assess the capabilities of many approaches and to compare with our method.

Having realized the research gap, the purpose of research therefore focused on ship classification using Faster Region Convolution Neural Network (Faster R-CNN) for port security. The main objectives were as follows:

1. To classify classes of ships using Faster R-CNN for port security
2. To compare performance of Faster R-CNN with other approaches in ship classification
3. To identify area of improvement in future research on ship classification for port security

1.3 Thesis Organization

The study thesis has been organized into five chapters of introduction in Chapter 1 and research trends in Chapter 2. Details of our approach is presented in Chapter 3 while experimentation and results discussions are presented in Chapter 4. Based on discussions of the experiment results, conclusions and recommendations are provided in Chapter 5. The last sections include references and academic achievements of the author. The details of the thesis organization are discussed as follows;

In Chapter 1, introduction of the study is elaborated to include existing information related to the study objectives, existing application of AIS and its challenges, trending researches on ship classification and identifying research gaps that exist in the literature. First, information concerning marine transport have been highlighted based on global emerging issues. The information provides a clear understanding of issues of port security and its impacts on coastal countries which depend on ports for trade. A highlight of major conventions is made to reveal the ever-changing needs in providing proper ship monitoring system for ship monitoring.

The chapter also gives a brief description on the use of AIS in monitoring of ships and its limitations on successful implementation on all kinds of ships. Here, we track its application over time and possibilities of enriching its performance with integration of newer approaches. The diminishing capacity of AIS is revealed and linked to advancement in port operations due to modernization in most countries. A few researches related to improving the functionality of the existing system, the AIS, are also presented to give a direction in which newer technologies may be leading to. This also helps to identify the existing research gaps and challenges that triggered the decision of using our approach of Faster R-CNN method for ship classification. The study approach is then aligned to information provided in the introduction.

In Chapter 2, research trends on image classification models are discussed with review of past, recent, and current researches being focused. In the beginning, an elaboration of object recognition models is presented to understand their operations, relationships and their properties. This begins with the overview of Artificial Neural Networks (ANN) followed by Convolutional Neural Networks and finally Recurrent Convolutional Neural Networks. Introductory note on Deep Convolutional Networks is briefly mentioned since much of the information is provided while describing our study approach of Faster R-CNN in Chapter 3. This chapter also provides information on machine learning and Deep Learning models that have been used in classification of ships which forms the basis of performance comparison with our approach.

This thesis provides detailed information of our approach of Faster R-CNN in Chapter 3 with focus on its development from the Fast Region Convolutional Neural Networks with incorporation of region proposal. A comparison is made on between other models in the same chapter with mention of its successes on various datasets (mostly of distinct images) with different attributes from ship images.

In Chapter 4, the methodology used in carrying out the experiment is discussed. This includes analysis of performances of various models on ship classification over the recent years and predicted the future potential of Deep Learning approaches in better classification of ships. Description of datasets used and preprocessing are also elaborated. The step by step experimentation is highlighted and results provided in form of tables and figures. Description of respective results data is also given to explain the nature of the results.

Summary of the experiment findings are provided in Chapter 5 where conclusions and recommendations from the research findings are made. Conclusions are provided to reveal the strong and weak points of our research and to recommend on future research work. In addition, a conceptual framework of Faster RCNN/AIS identification model is presented as a futuristic prospects of our study.

CHAPTER 2

RESEARCH TRENDS

2.1 Concept overview of Artificial Neural Networks (ANN)

Object detection, identification and classification models have developed more profoundly with the emergence of Deep Learning approaches. The development of neural networks has inspired many researchers which started with the interests in the neural processes and neurons of a human being. Scientists describe human beings as self-intelligent based on these neural processes. This human feature formed a baseline in the introduction of neural networks in machine learning models. The neural networks in machine learning consist of series of neurons and nodes, which are interconnected to perform some processes resulting in an output. The connection between neurons are weighted, based on the intended function and learned for a collection of data including images, characters and even facial features of humans and other creatures.

In this section, we present various classification systems based on the neural network models. To begin with, we discuss the architecture of Artificial Neural Networks and their properties. There are several classification systems based on ANN which have been developed over years. Some of them include Probabilistic Neural Network, Hamming Neural Network and Morphological Neural Network. The details of these networks are discussed and illustrated in this section. Much information on these models have been provided based on previous researches in [6], [7], [8], [9], [10] and [11].

2.1.1 Probabilistic Neural Network (PNN)

Probabilistic Neural Network (PNN) is one of the first single layer networks which performs well in data classification than the old state of the earth models. The classification method for this network follows the rule of Bayes maximum *a posteriori* classification scheme and the Parzen kernel PDF estimation model. Its architecture is made of four layers as shown in Fig. 1 below. The input pattern vectors, of L dimensions are received as input layer \mathbf{X} and are normalized. This means that every input pattern vector may belong to one of the c classes. The \mathbf{w} layer possess the number of weights which store components of the reference patterns. The neurons w_{cn} , where each belongs to one of the ω_c classes ($1 \leq c \leq C$), computes a kernel function of its reference patterns, and is stored in weights of \mathbf{w} and the present input, in accordance with the formula;

$$w_{cn}(\mathbf{x}) = K\left(\frac{\mathbf{x} - \mathbf{x}_{cn}}{h_c}\right), \text{ for } 1 \leq c \leq C, \text{ and } 1 \leq n \leq N_c, \quad (1)$$

Where \mathbf{x}_{cn} represents the n -th prototype pattern from an c -th class of ω_c ,

h_c Represents a parameter that controls the effective width of the kernel for the respective class hence acting as a zone of influence.

The N_c represents the number of available prototypes for the respective classes.

The outputs of the neurons from the layer \mathbf{w} are then fed into summation layer, which is composed of C neurons. For respective outputs, each follows the Eq. (2) which is the Parzen method for non-parametric density estimation.

$$g_c(\mathbf{x}) = \alpha_c \sum_{n=1}^{N_c} K\left(\frac{\mathbf{x} - \mathbf{x}_{cn}}{h_c}\right), \quad (2)$$

Where α_c represents a scaling parameter. Gaussian kernel is selected for K which leads to the function;

$$g_c(x) = \frac{1}{N_c} \sum_{n=1}^{N_c} e^{-\frac{\|\mathbf{x} - \mathbf{x}_{cn}\|^2}{2\sigma^2}}, \quad (3)$$

In this Eq. (3), the same parameter σ is substituted for h_c for all the classes, and the common multiplicative constant is omitted. Therefore, one neuron of the summation layer with all its preceding neurons composes the Kernel PDF estimation path, specific to a given class. The final layer Ω selects one $g_c(x)$, which gives the maximal response. Its index j indicates the class to which the input pattern has been classified by the network;

$$\omega_{c'} = \arg \max(g_c(x)) \quad 1 \leq c \leq C \quad (4)$$

The whole network realizes the mean average precision (MAP) classification rule which follows the maximum membership strategy.

This approach has been widely used in computer vision problems for instance in gesture recognition and tracking. Input images are first projected to low dimensional space, then used to train the algorithm. The classification is based on affine moment invariants which allow moderate affine variations of the input images of pictograms. Table 1 shows the summary of the properties of PNN method.

Table 1 Properties of PNN

Merits	Demerits
1. Fast in training	1. Slow in response
2. Bayes optical classifier	2. High memory requirements
3. Parallel structure	3. No easy simple method for σ selection

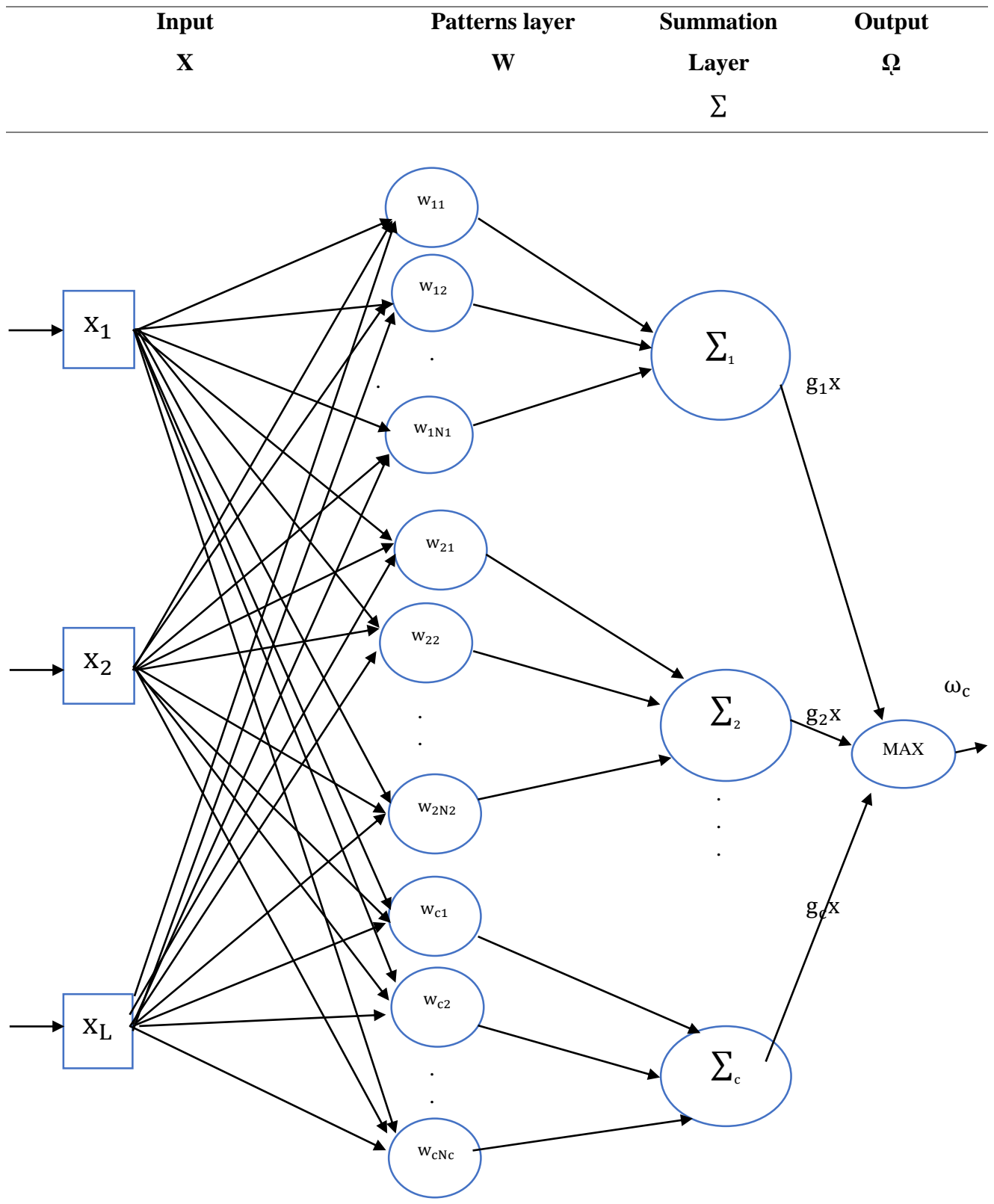


Fig. 1 Architecture for the Probabilistic Neural Network

2.1.2 Hamming Neural Network (HNN)

The second model of artificial neural networks is Hamming Neural Network (HNN) which allows classification of patterns whose features can be measured with hamming distance. This approach was first proposed by [6]. This method directly realizes the one-nearest neighbor classification rule. Its structure is as shown in Fig. 2 below;

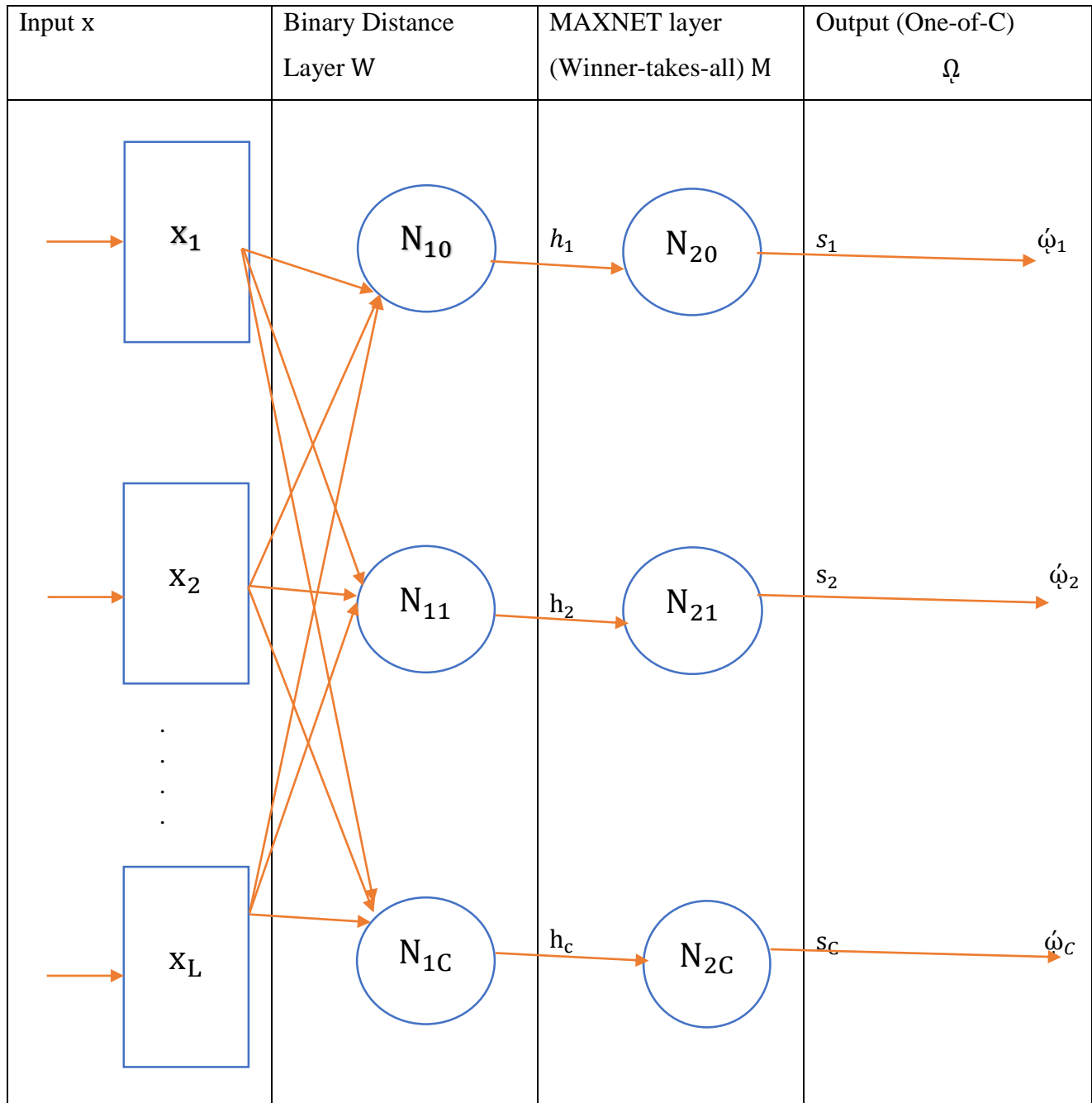


Fig. 2 Architecture for the Hamming Neural Network

The Hamming Neural Network contains four auto associative versions and five hetero-associative layers of neurons. It has larger pattern capacity than the probabilistic neural network. It is also faster in training and recognition than PNN. It classifies test pattern represented by a feature vector x into one of the C classes. It consists of setting row matrix W with the following prototype patterns;

$$w_i = X_i \quad (5)$$

Where;

$1 \leq i \leq C$ is the number of prototype pattern X_i . Each has a length of length L ,

w_i is the i^{th} row of the matrix W , of dimension $C \times L$.

The computation time is linear with the size and the number of input patterns C . MAXNET selects a winning neuron indicating the class input pattern. MAXNET is initialized by assigning negative values to the square matrix M of dimensions $C \times C$ except of the main diagonal which is 1.0 since it implies self-excitation of a neuron. The values used to initialize M is determined by

$$m_{kl} = \begin{cases} -\frac{1}{C-1} + \xi_{kl}, & \text{for } k \neq l \\ 1, & \text{for } k = l \end{cases} \quad (6)$$

Where;

$1 \leq k, l \leq C, C > 1$, ξ is a sufficiently a small random value for which $|\xi| \ll 1/(C - 1)$. However its modification, assigns the same value ξ_k to all m_{kl} for $k \neq l$. This is explained in the equation:

$$m_{kl} = m_k = \begin{cases} -\frac{1}{C-t} + \xi_{kl}, & \text{for } k \neq l \\ 1, & \text{for } k = l \end{cases} \quad (7)$$

This computation transforms the values of M to be near-Optimal in terms of the network convergence. The complexity for the general case is of order $O(C \log(LC))$ for general case and for the unique prototype vector is $O(C \log(L))$. The unique prototype vector therefore becomes the nearest to the input pattern. The advantage of this modification is significant reduction in memory as the matrix M reduces to a vector of length C . However, the most efficient and convergent solution consists of setting equal weights for all neurons in the MAXNET layer as follows

$$m_{kl} = m_k = \begin{cases} -\frac{1}{C-t} + \xi_{kl}, & \text{for } k \neq l \\ 1, & \text{for } k = l \end{cases} \quad (8)$$

The t element is a time step in the classification process. The m_k values must be modified in each stage from which the convergence is achieved in $p - 1 - r$, for $r > 1$ indicates several nearest prototypes stored in W . By choosing the two models, it will influence the speed of convergence of the classification stage which is determined by the type of stored prototypes. The run-time stage neurons of the distance layer, the binary hamming distance between the input patterns x and the prototype patterns already stored in W is computed as in Eq. (9) below.

$$h_i(x, W) = 1 - \frac{1}{L} D_H(x, W_i) \quad (9)$$

Where $1 \leq i \leq C$, $b_i \in [0,1]$ is a value of an i^{th} neuron in this layer, $D_H(x, W_i) \in \{0,1, \dots, L\}$ is a Hamming distance between the input patterns x and i^{th} stored prototype pattern, W_i . Usually all feature vectors are assumed to have coefficients from the set $\{-1, +1\}$ then reduces to

$$h_i(x, W) = \frac{1}{2} \left(\frac{W_i x}{L} + 1 \right) = \frac{1}{2} \left(\frac{1}{L} \sum_{r=1}^L W_{ir} x_r + 1 \right) \quad (10)$$

For the input vectors with values from the set $\{0,1\}$ the hamming distance can then be determined as

$$h_i(x, W) = 1 - \frac{W_i x}{L} = 1 - \frac{1}{L} \sum_{r=1}^L W_{ir} x_r \quad (11)$$

Extension the allowable values from the binary set $\{-1, +1\}$ to the ternary $\{-1, 0, +1\}$ where value 0 indicates an additional state whose role is to express certain condition of “don’t know” state. During the classification, the MAXNET layer performs recursive computations to select a winner neuron in accordance with the scheme

$$s_i[t + 1] = \theta \left(\sum_{j=1}^L m_{ij} s_j[t] \right) = s_i[t] + \sum_{j=1, i \neq j}^L m_{ij} s_j[t] \quad (12)$$

Where, $s_i[t]$ is an output of the i^{th} neuron in MAXNET at the iteration step t , while θ denotes a function which suppresses all negative values to 0 as follows;

$$\theta(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (13)$$

This means that the output of a neuron from this layer goes beyond 0 its signal s_i is also set to 0. The role of iterative process is to proceed until only one neuron has a value different from 0. Hence, this neuron is a winner of the process and its index indicates the determined class of the input pattern. The properties of Hamming Neural Network are summarized in the Table 2 below:

Table 2 Properties of HNN

Advantages	Disadvantages
<ol style="list-style-type: none"> 1. Fast training 2. Fast response 3. High Capacity 	<ol style="list-style-type: none"> 1. Binary input 2. Only Binary distance (Hamming) 3. Iterative process

2.1.3 Morphological Neural Network

Morphological Neural Network (MNN) consists of a group of neural networks which exhibit many desirable properties such as high capacity, resistance to the erosive and dilative type of noise, as well as ability to respond by just one step process. These properties make it more popular than HNN and PNN on pattern classification community, especially on real-time systems. Its establishment was to the need for very fast neural solutions. The basic concepts of MNN utilizes the mathematical lattice theory. The Fig. 3 shows the MNN model which operates in respect to the Eq. (14) below.

$$y_k = \theta(p_k \bigvee_{i=1}^L r_{ik}(r_{ik} + w_{ik})) \quad (14)$$

Where r_{ik} is a pre-synaptic response which transfers excitatory ($r_{ik} = +1$), or inhibitory ($r_{ik} = -1$), incitation of an i^{th} neuron, p_k is the post-synaptic response of a k^{th} neuron to the total input signal, \vee denotes a max product, and finally θ is a saturation function. The MNN model structure is illustrated in the Fig. 3.

In [7], MNN structure was modelled for classification of road sign recognition. In this version of MNN, a set of N input/output pairs are given as $(x_i, y_i), \dots, (x_N, y_N)$. In this case, x is the linear version of an image of a sign and y is a binary version of a pattern's class. The input X , therefore, contains a binarized pictogram of a tracked sign. This robustness of MNN model made binarization and sampling possible to be carried out in accordance to classes of pictograms which is useful in image processing. The properties for MNN are summarized in Table 3 below;

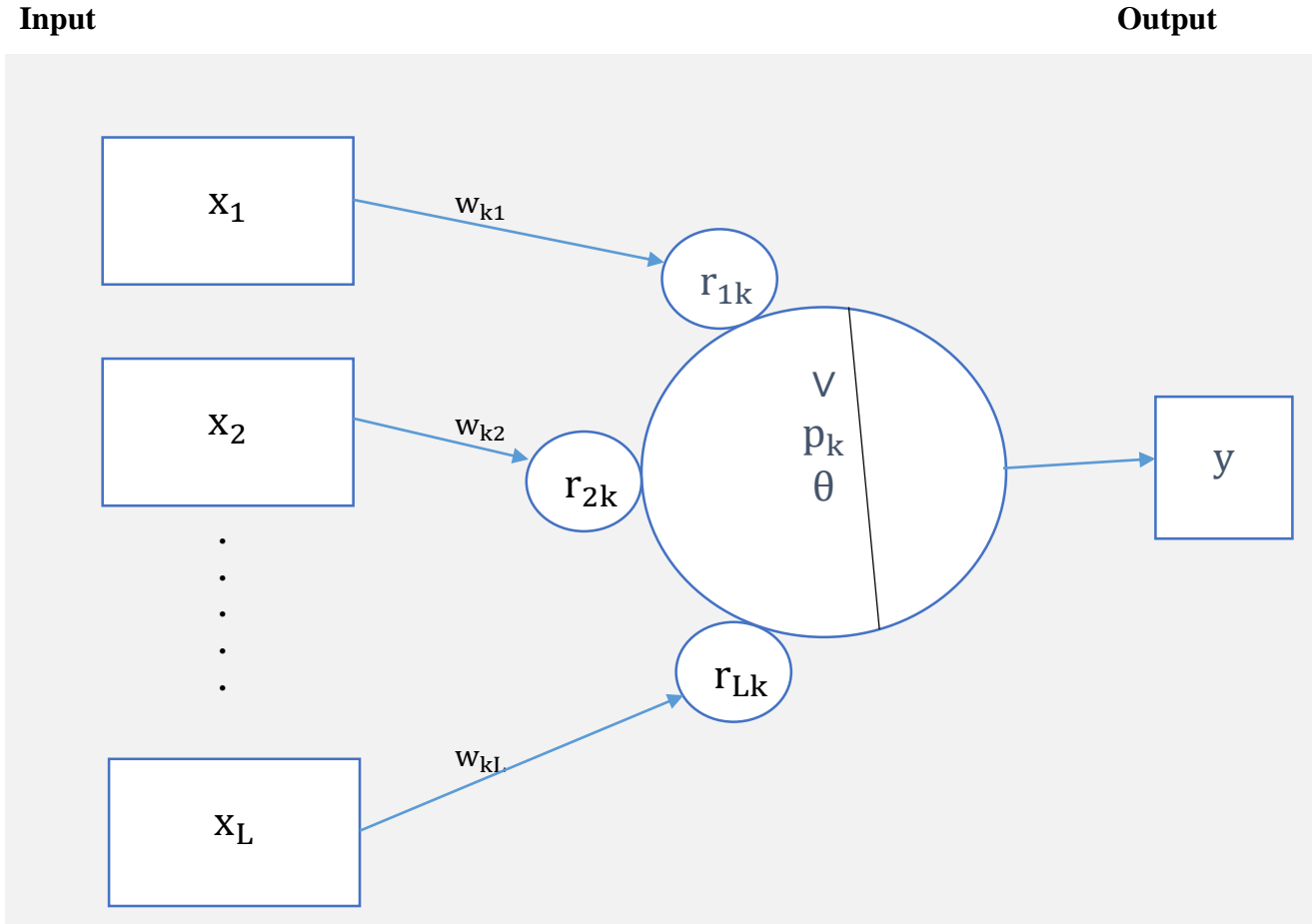


Fig. 3 Architecture for Morphological Neural Network

The max product V for two matrices A_{pq} and B_{qr} is a matrix C_{pr} , with elements C_{ij} is defined as

$$C_{ij} = V_{k=1}^q (a_{ik} + b_{kj}) \quad (15)$$

and similarly, the min operator \wedge is defined as

$$C_{ij} = \wedge_{k=1}^q (a_{ik} + b_{kj}) \quad (16)$$

Table 3 Properties of MNN

Advantages	Disadvantages
<ol style="list-style-type: none"> 1. Fast running 2. Fast response 3. High capacity 4. Associative Memory resistant too 5. The morphological noise 	<ol style="list-style-type: none"> 1. Lattice based (requires order relation on data)

2.2 Concept Overview of Convolution Neural Networks (CNNs)

Many Researches related to computer vision using Deep Learning have achieved better results compared to artificial neural networks. Some of the outstanding features in Deep Convolution Neural Network approaches include non-utilization of manual features of target data. That means Deep Learning has capability in extracting and hence learning features with high independence. Compared to Artificial Neural Network, these approaches contains many layers with nonlinear processing units of which each layer transform inputs from the previous layer leading to formation of hierarchical data. This implies that high level features are more abstract than lower level features.

Basically, CNNs have two types of layers: convolutional layers and pooling layers. The convolutional layer takes feature maps of the previous layer as inputs which are processed as 2-dimensinal convolution and learnable filters. The next layer receives them as stack of new feature maps. Thus, is expressed mathematically in [12] as

$$X_n = f(\sum_m W_n^m * X^m + b_n), \quad (17)$$

Where X_n represents nth feature maps, W_n^m are the filters which operate on convolution process. b_n shows value that corresponds to each feature map. In CNNs the convolutional layers only allow neurons to connect with a local region of input hence reducing the number of parameters in the model. This makes CNN to accept lager dimensional inputs which is not the case in ordinary neural networks. These processes are well illustrated in [8] the CNN model architecture. In this model, the convolutional layer is preceded by the pooling layer which summarizes the neighboring feature detectors thereby reducing the number of features for the next layer. This is achieved by taking the optimal values of a neighboring feature patch and passed to the next layer. This is a unique feature for CNNs which allows the creation of Deep Neural Network, by stacking multiple layers and alternated between convolutional and pooling layers. Fig. 4 shows the CNN model before it is linked with multiple layer perceptron for classification.

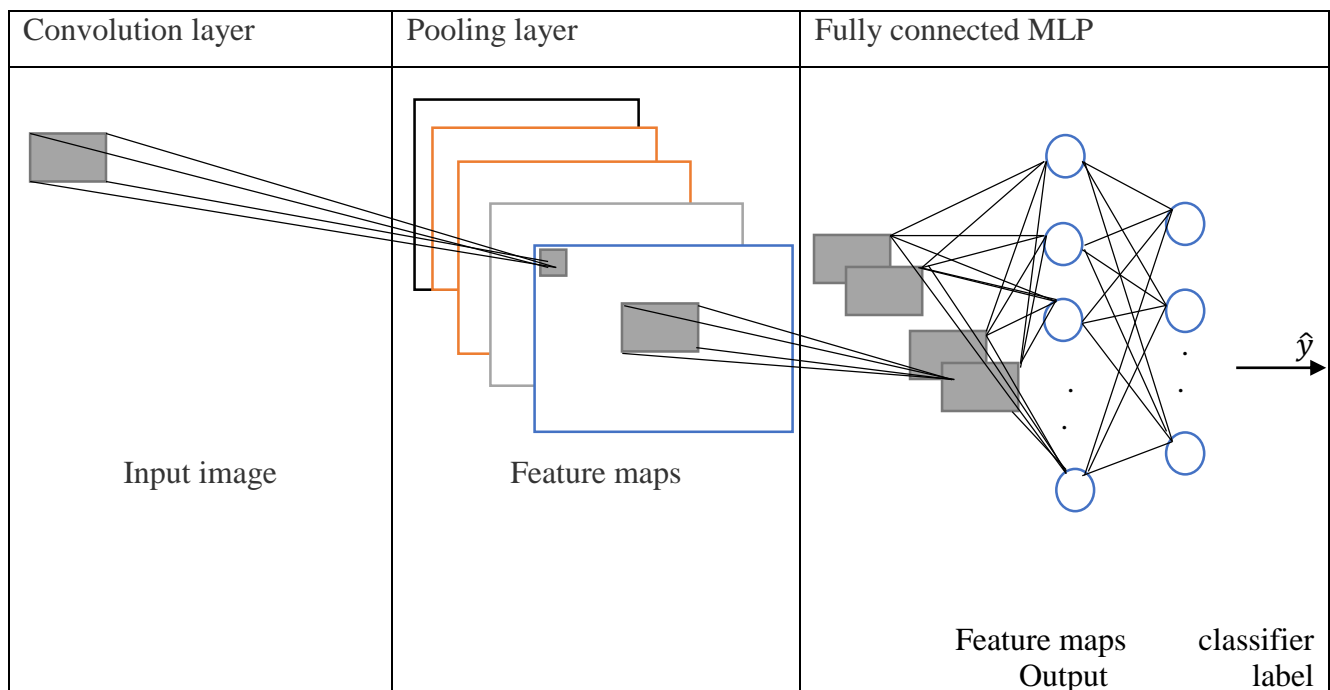


Fig. 4 CNN architecture model before linking with multiple layers

2.3 Recurrent Neural Network (RNNs)

In Recurrent Neural Networks, the time component is introduced. It acts like feedforward neural networks alongside feedback loops through time. In this case, the neurons activate other neurons which fire at a later point in time. This time component and robust structure enables the network to use the available inputs as well as other inputs encountered earlier. RNNs are attributed to better results on speech recognition, language translation, and connected handwriting recognition but are difficult to train as compared to other recognition models. A symbol structure of RNN is illustrated in Fig. 5.

Based on the various models in Deep Learning, there are many important concepts which has attracted more researches on the use of machine learning approaches. The research development trends started with simple single layer models and now we have multiple neurons to neural network architectures which are being applied in more complex problems like classification of ships.

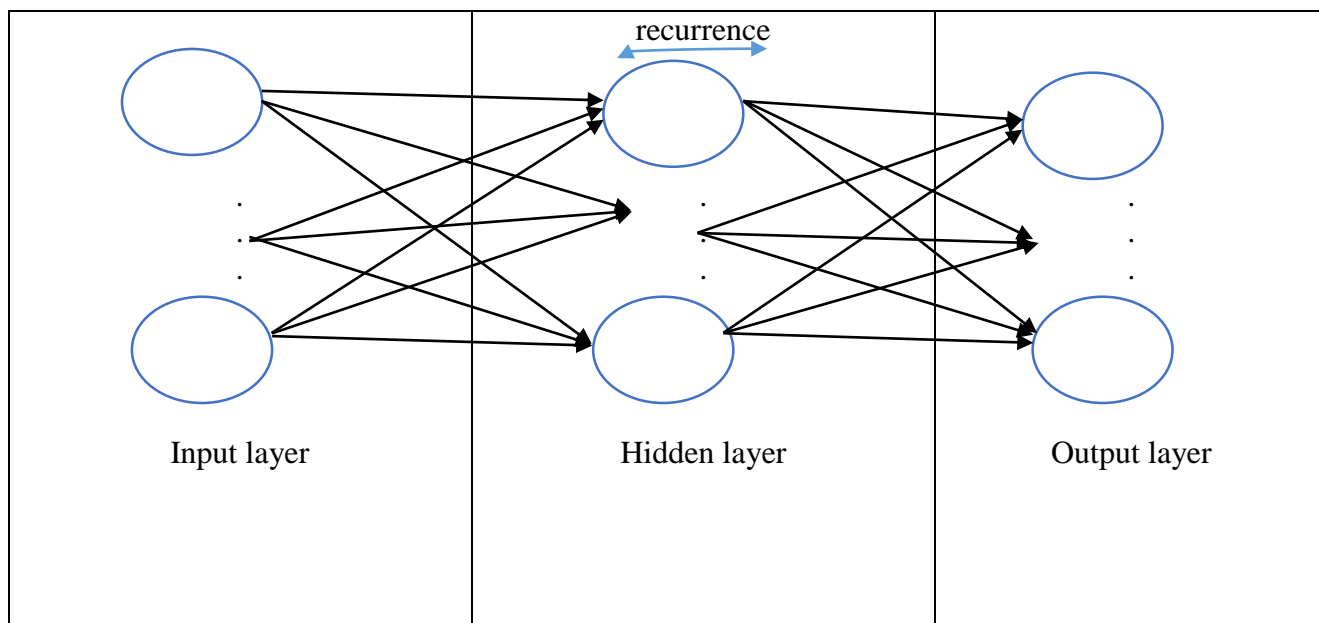


Fig. 5 Symbolic structure of Recursive Neural Network [8]

2.4 Previous researches on ship classification

About a decade ago, modern approaches for ship detection, identification and classification began to emerge alongside with growth in space related researches. The work of [12] relied on the remote sensing data with vessel position data to detect ships from space. However, classification of ships remained difficult according to prospects then. The main limitation was because of weather on sensors and their approach was less suitable for wide area surveillance. Their study revealed that, the satellite radar imagery provided limited amount of information on ships and the traffic conditions. Therefore, classification of vessels was very limited making even identification nearly impossible. However, their approach suggested that satellite imagery could be successfully used in combination of other methods.

Ship recognition using optical sensors and database of previously obtained ship images was elaborated in [13]. They describe a criterion of recognizing ships using SIFT features with consideration of cluttered backgrounds in maritime environments. The approach used local features which were extracted from optical imagery for automatic ship classification in port surveillance. Local features were found to be tolerant on cluttered scene. The SIFT technique seemed to be reliable for better classification results when proper local interest points are detected and good feature matching is provided. It contributed a lot in incorporating multiple views of target ships from the image database and in verification of geometric relationships of the matched features. In comparison with other methods, SIFT

achieved excellent recognition rates but could not solve real time problems as compared to modern methods like Deep Learning methods.

Port surveillance has attracted researchers into focusing on detection, identification and classification of marine vessels which vary greatly from each other depending on size, shape, vessels activity and transition data. IMO is the specialized agency under United Nations with responsibility for the safety and security of shipping and prevention of marine pollution. International shipping plays a big role in global trading and the world relies on safe, secure and efficient shipping industry. It is the important factor that has triggered many researchers to focus in ship monitoring and surveillance technologies. In this section, we present various works related to our study that are baseline for improving performance of various approaches.

The difficulties faced in protecting sea surface and busy ports, the work of [14] presented a state of the art solution for ship intrusion detection using image processing and Support Vector Machine (SVM). The aim of their contribution was to detect ships, which crossed over designated spaces. CFAR algorithm [15] was applied to improve fast R-CNN in SAR ship detection task. The role of the Fast R-CNN was to obtain classification scores and refined bounding boxes from region proposals and feature maps. This approach however could not detect small sized targets hence leading to a low detection rate. Data augmentation was a constraint factor for consideration to realize better performance.

An attempt was carried out to automatically detect ships based on S-CNN method with proposals designed from a combination of ship model and an improved saliency detection method [16]. Models with “V” ship head and “||” ship body were used to localize the ship proposals from the line segments of test images. The proposals are fed to the trained CNN for efficient detection which proved suitable for better application on remote sensing images with different kinds of ships. This resulted in 91.1% and 97.9% recall performance for in-shore and off-shore ships respectively, while, 95.9% and 99.1 % precision rates respectively. In this regard, Deep Learning still is an appropriate method for detection, however S-CNN may not produce better results on classification due to problems affecting space images like weather as well as being slower in detection rate.

Surveillance is a paramount problem for harbor protection, border control and security of various commercial facilities. It is particularly challenging to protect the vast near coast sea surface and busy harbor areas from intrusions of unauthorized marine vessels, such as trespassing boats and ships. In their project, they presented a state-of-the-art solution for ship intrusion detection using image processing and Support Vector Machine (SVM). The main aim was to detect the ships, which cross over the border and secured industrial spaces. Using the interworking mechanisms of these two techniques,

we can detect the intruding ship from the constantly changing sea environment. SVM can be used as a machine learning to train the system by exposing it to different seashore environments. Hence, it can be used as a real-time security system at seashore areas. The approach was an integration of machine learning with intrusion detection system to teach the system to learn the highly-clustered environment. This contribution obtained high accuracy on detection on intruder vessels but could not classify the target ship images.

CHAPTER 3

METHODOLOGY

3.1 Proposed Method

Faster R-CNN is an advancement of Fast R-CNN [17] object detection model. The advancement into the model resulted from previous research focusing on solving emerging challenges in computer vision. Faster R-CNN is a unique Deep Learning model which makes use of the successes in the region proposal methods of selective search and the Region Convolution Neural Networks. The Region Proposal Networks (RPNs) are very slow as compared to the faster R-CNN. However, the RPN which shares the convolutional layers plays a big role in Faster R-CNN as it makes the model a nearly cost free with the computational speed and a test time of about 3-13ms. The RPN also makes Faster R-CNN to support any scale in image dataset hence making it scale invariant. Compared to the state of the art models, Faster R-CNN seems to produce excellent results with minimal testing times and with high accuracy rates. A summary of its capabilities compared to other models is illustrated in the table below.

Table 4 Comparison of R-CNN, Fast R-CNN and Faster R-CNN

	R-CNN	Fast R-CNN	Faster R-CNN
Testing time per image (Seconds)	~50	~2	~0.2
Speed capacity	1 ×	~25 ×	~250 ×
Mean average precision (mAP) using VOC 2007	≤ 0.66	≤ 0.669	≥ 0.66

To understand the operation of Faster R-CNN, it is useful to follow its development from the state of the art methods and then visualize its potentials in classification of ships. In this section, we discuss the architectural relationship between R-CNN, Fast R-CNN and the unified network of the two defining the operating principles of Faster R-CNN.

3.1.1 Region Convolutional Neural Network (R-CNN)

In R-CNN, the sequence of input images with region proposals are fed to and run by CNN. The proposed regions are referred to as those containing target object with high probability. These regions are provided by extra proposals like the selective search [18]. For a typical input image, it usually has approximately 2000 region proposals. This is a huge number and it makes the R-CNN computationally expensive due to its huge complexity. It therefore requires a long time of detection per image. From various applications, it has shown that it takes approximately 50 seconds of computation per input image.

3.1.2 Fast Region Convolutional Neural Network (Fast R-CNN)

The basic principle of operation in R-CNN differs from that of Fast R-CNN because of the sequence of the processes used in completing the task of object classification. For R-CNN it starts with running the region proposals with CNN. This first step is done only once then followed by the calculation of Region of Interest (RoI). The reason for RoI calculation is to estimate the location of the region proposals. Generally Convolutional Neural Networks have the convolutional layer, activation layer and the max pooling layer. The specifications of the Convolutional Neural Network are important since it determines the location from the input image and in the output of CNN. The operational difference of fast R-CNN and faster R-CNN is that, Faster R-CNN does not need external proposals anymore.

3.1.3 R-CNN and Fast R- as a unified network of Faster R-CNN

Faster R-CNN (Unified Network)

[5]

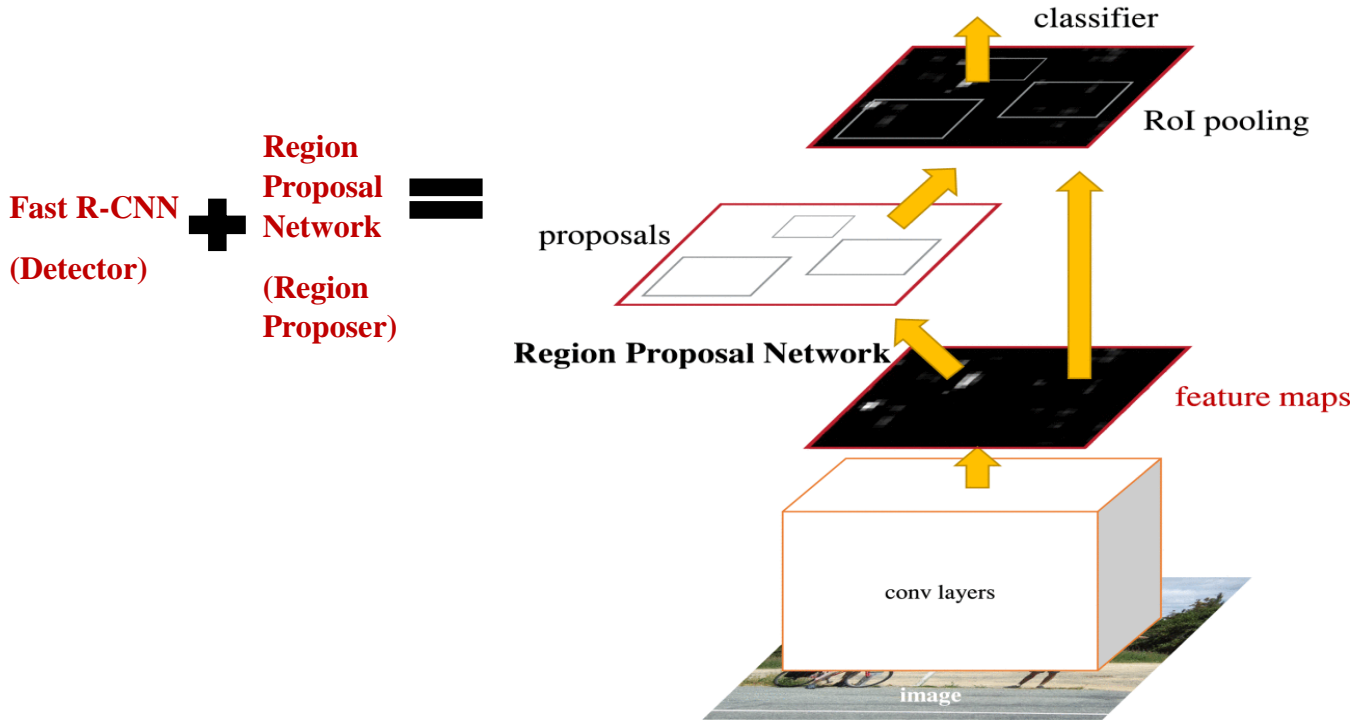


Fig. 6 Faster R-CNN as a single, unified network for objection detection

In Faster R-CNN architecture, both the Convolutional Neural Network and Fast R-CNN act as a unified network. This involves the introduction of Region Proposal Network (RPN) which lets the unified network to locate where to look. The symbolic expression of this unified network is shown in Fig. 6 above. It is a two-module system where a deep fully convolutional network proposes regions while the second module is a Fast R-CNN [17] detector which uses proposed regions. The role of RPN is to decide where the second module of the system will look for in the Region of Interest.

In this model, the role of Region Proposal Network has been explained in detail by [5] with a symbolic model shown in the Fig. 7 below. Since RPN takes any size of images as input, it makes the model to be scale invariant. The RPN is designed to output a set of rectangular object proposals with respective probability object scores as shown.

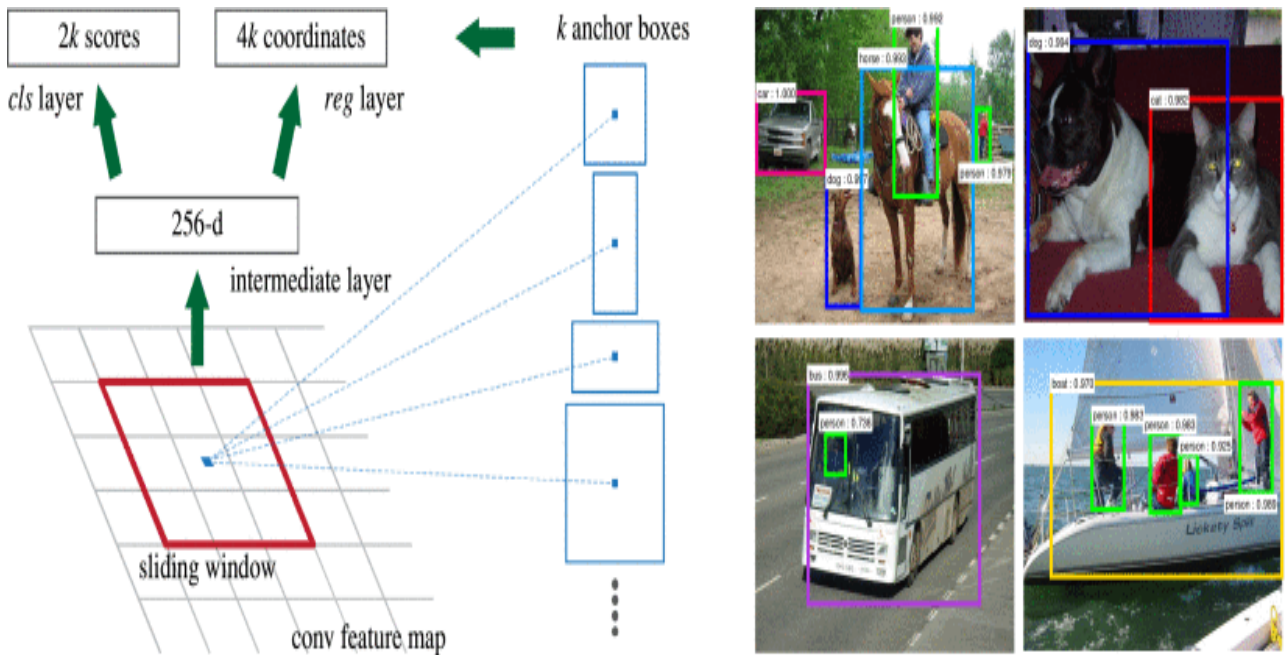


Fig. 7 RPN framework and examples detection using RPN (from [5])

Advantages of Faster R-CNN, include the computational capacity, speed, accuracy and reliability of performance, which enables it to be suitable for our study on classification of ships for port security.

3.1.4 Training of Faster R-CNN

Technically, training of Faster R-CNN depends on the arrangement in the network itself which makes it a four-stage process. From the process sequence followed, the RPN should be the first network to

train. This is done by initializing with pre-trained models of ImageNet. The detector network (Fast R-CNN) is the second stage of training. This is done by the proposal generated by the first training in RPN. This is followed by fine tuning the unified layers to RPN after being initialized in step 2. The last training involves fixing convolutional layers and fine tuning of full convolutional layers of fast R-CNN. The overall training of faster R-CNN depends on some properties as explained in [9]. These properties include:

- a.) Weighting values initializations using Gaussian equations
- b.) Learning rates which should be decreasing in the subsequent batches
- c.) Learning update scheme and
- d.) The weight decay which is explained in Eq. (18) below

$$L = \frac{1}{L} \sum_i L_i + \lambda R(W) + \lambda \sum_k \sum_l W_{k,l}^2 \quad (18)$$

Where, L , the weight decay, is the objective function,

$\frac{1}{L} \sum_i L_i$ represents data loss, and

$\lambda R(W)$ denotes the regularization loss.

3.2 Experimentation

3.2.1 Datasets pre-processing

The aim of our study is to classify various classes of ships using the faster R-CNN method and with focus on improving performance than the previous methods on ship classification. Our dataset therefore constitutes of images of ships obtained from *shipspotting.com* [19]. As described in the unified network structure of the Faster R-CNN model, all images are re-scaled such that the shorter side was ≥ 500 pixels for the purpose of uniformity, since the approach is scale invariant. All the images are checked individually to ascertain their attributes in readiness for training. This leads to the determination of specific models of ship images as illustrated in Fig. 11. Some ship images are not considered for training as they had extremely abstract shapes despite belonging to specific class and model.

To establish the most suitable quality dataset we first, investigate the effect of the number of images on performance after training through the Faster R-CNN method. As a preliminary experiment, a few classes (four) of ship types are sampled for classification by Faster R-CNN method. Images of ships used for the preliminary classification begin with a few, then the number is increased subsequently while checking the accuracy rates and consistency of the approach. The aim of this preliminary

experiment is to determine the effect of number of training images on the performance of the approach and also to determine the appropriate number of training images with steady Average Precision (A.P.) values. Upon the determination of saturation level, the appropriate number of images are then used in the classification process of the target nine classes of ships.

The performance of our approach is evaluated based on run time, precision and ability to classify ships on near real-time situation with considerations of the highly cluttered marine environment. A comparison with related state of the art approaches is made in order to highlights the strengths weakness and possible improvemet methods in the future. The success of our study would be an important contribution in automatic classification of ships for improvig security and efficiency in port and other marine environments.

CHAPTER 4

EXPERIMENT RESULTS AND DISCUSSION

4.1 Analysis on Recent Developments in Object detection approaches

In our study, we began by analysing the performance trends of the recent developments on object classification approaches which include Faster Region Convolution Network model. From the analysis, it is evident that the efforts being made to solve image classification problems are experiencing a positive progress in the recent times. The state of the art approaches have been aiming to achieve better performance in term of accuracy, learning rates, speed and reliability of models. There is a tendency on accuracy improvement and reliability on application of Deep Learning approaches on image classification. The results of the analysis is illustrated in Fig. 8 below.

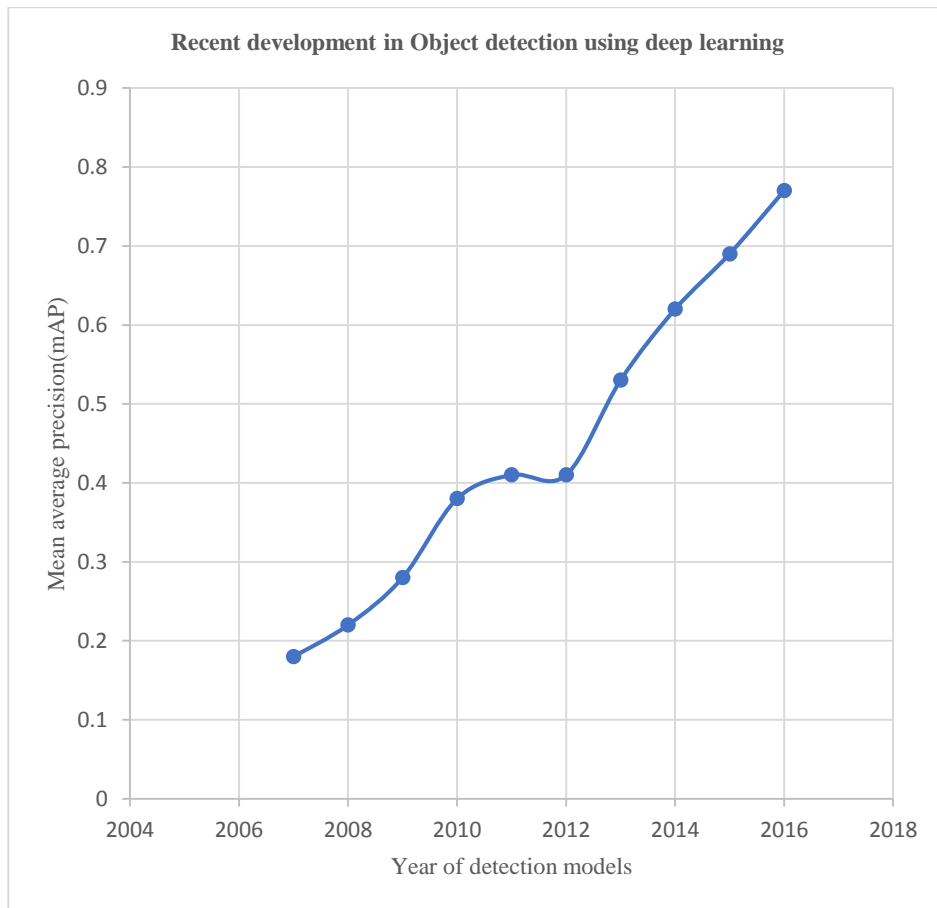


Fig. 8 Object Recognition development using learning

Form the graph in Fig. 8 above, the analysis assumes that the approaches used before the year 2007 had lower precisiion values while newer models eastablished after 2016 would have better performance. In our study, the analysis results on the performance of deep learning approaches in [18], [7], [9], [17],

[5], [11] and [10] for older object detection approaches (before Deep Learning) are considered. The data used for this analysis is based on selected results of old models and those using deep convolutional network models from the year 2006 upto the year 2016. From the analysis, Faster R-CNN appeared to be the most suitable for ship classification particularly because it is attributed to high accuracy and recall.

4.2 Experiment 1

The first aim in our experimentation is to investigate the effect of the number of training images on the mean Average Precision (mAP) values of our approach when applied on classification of ships. In this case, the number of images which are suitable for training are selected sequentially while checking on the performance of the approach. Saturation of AP values in output is a key determinant of the respective number of ship images to be used per class. The samples of the ship images used are shown in Fig. 9.

For the preliminary experiment, 4 types of ships (Cargo, Tanker, Fishing Vessel and Military) are classified using Faster R-CNN. The choice of these 4 types of ship images is based on representation rule for large versus small, shape difference and functionalities of each type. We started with training a 100 images for each of the 4 classes then increase sequentially upto 400 images at every 50 images. The results of mean average precision are shown in Fig. 10. The results indicate that there exists a strong positive correlation between the number of images used in training and the AP values. The AP values in our approach tend to improve with increase in number of images, and saturates at around 400 images.

Since the average precision values tend to saturate when the number of images per class are 400, this number is then selected for use in our experiment 2. The target is to classify nine classes of ships which are common in a port environment.





<p>a. Cargo</p> <p>It is the most common class of ships along maritime trade routes and in major ports across the world</p>	
<p>b. Fishing Vessel</p> <p>Fishing Vessel is common along the routes of other ships and mostly has varied modification features (additions by users to its original shape)</p>	
<p>c. Tanker</p> <p>Tanker is generally bigger in size than most many other kinds of ship</p>	
<p>6. Military ship</p> <p>It has unique shape protruding upwards which makes it unique in all other types</p>	

Fig. 9 Samples of images used in Preliminary experiment

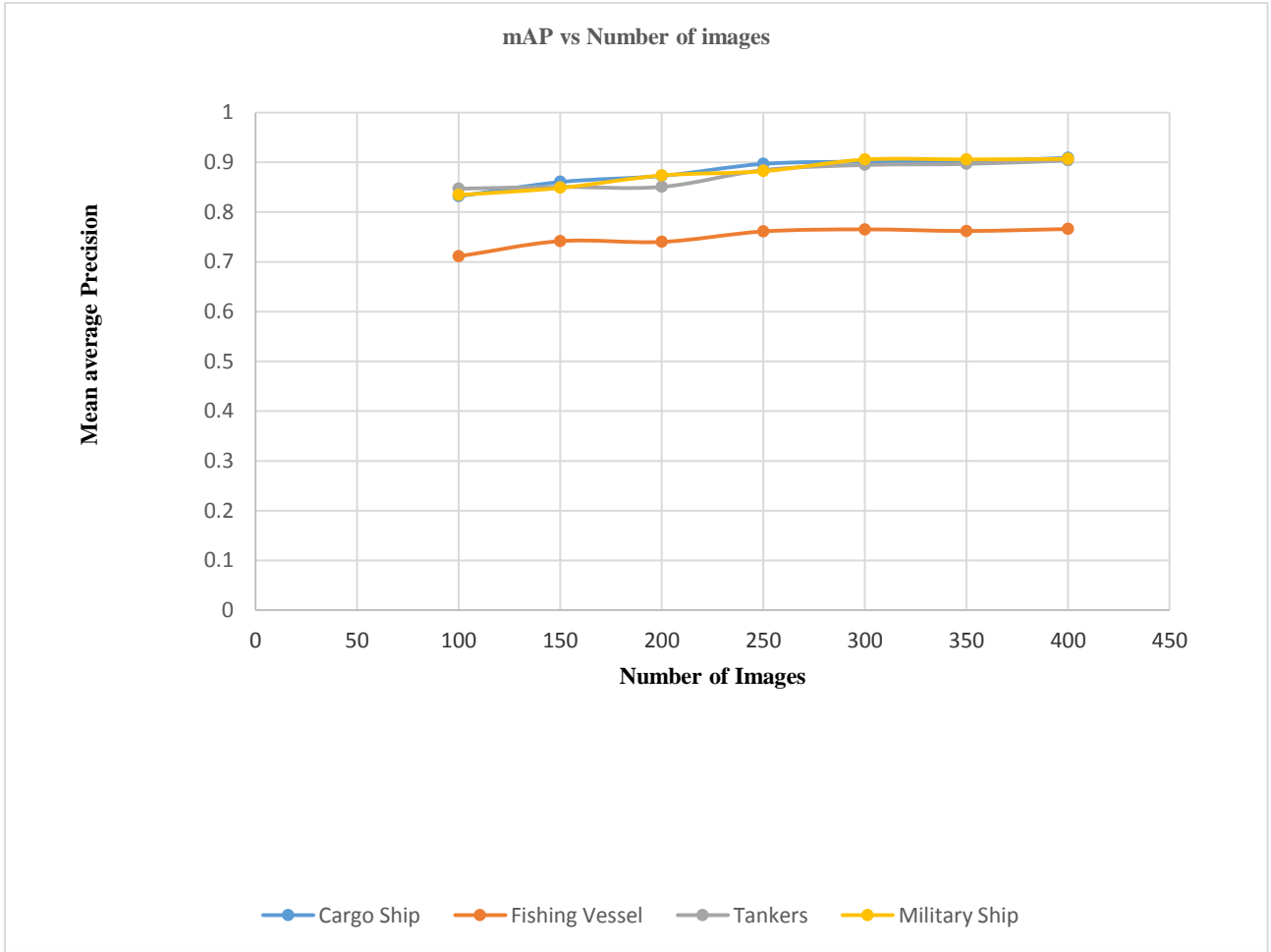










Fig. 10 Change in mAP values with the number of training images

4.3 Experiment 2

In the second experiment, 400 images from each of the 9 classes are chosen for training and another 400 for testing with the intention of classifying the target ship images using Faster R-CNN method. All images used are obtained from *shipspotting.com* [19] and are pre-processed by resizing to have the shortest side with 500 pixel for the purpose of uniformity and to avoid risks in poor results based on biased image sizes. However, our approach of faster R-CNN is scale invariant and this risk would still be minimal. The mean average precision value over all class predictions is used as metric measure of performance. In addition, run time per image and number of iterations are noted for the purpose of confirming the performance of the method. Details on average precisions scores for each class of ship images and the overall performance are shown in Table 5.

4.3.1 Sample of Images Used in training and testing for 9 classes

Class	Images used in Training	Images used in testing
a.) Fishing Vessel: Mostly altered in physical features due to modification of users		
b.) Tanker: Generally large in size and with limited variations in shape.		
c.) Submarine: The protruding part in the middle makes it unique from other ships.		
d.) Tug: Mostly found along side other bigger ships and are generally small in size.		

<p>e.) Vehicle Carrier: Nearly identical in shape</p>		
<p>f.) Rescue: Almost similar to Tug but more features</p>		
<p>g.) Ferry: Mostly fully floatin unlike other ships</p>		
<p>h.) Destroyer Ship: The provision of comprehensive communication structures makes it different from other types of ships</p>		

Fig. 11 Samples of used in training and testing for 9 classes

Table 5 Average precision scores for each of 9 class and the overall mean

Class (IMO Standard Names)	Train Images	Test Images	AP
Container Ship	400	400	0.9091
Fishing Vessel	400	400	0.7659
Tanker	400	400	0.942
Submarine	400	400	0.8896
Tug	400	400	0.8165
Vehicle Carrier	400	400	0.9091
Rescue Ship	400	400	0.8862
Ferry	400	400	0.9091
Destroyer Ship	400	400	0.9069
Mean AP			0.8774

There is the minimal variation of AP values on individual classes except for fishing vessels which registered lowest score of AP (0.7659). The AP for Tanker ship is the highest with the score of 0.942. Container ship, Vehicle Carrier and Ferry also scored highly, each with average precision score of 0.9091. Another contender for higher score is the Destroyer Ship with average precision value of 0.9069 while others had AP scores below 0.9. These are Submarine, Rescue Ship and Tag Boat with Average precision scores of 0.8896, 0.8862 and 0.8165 respectively.

Though the classification of Fishing Vessels is the lowest, the value is still high compared to other models which cannot score beyond 0.6 of AP. compared to the nature of ship images the Fishing Vessel could not have been the best probably because of its physical features. Most of images of the Fishing Vessels have huge transformations due to modification by users hence the complexity [20]. It envisaged that by understanding the complexity on the object features in Fishing Vessels the average precision

value may be improved for better classification. The run time per image during classification processes is as expected for the approach compared to its performance on distinct objects like in [5]. For the first batches the run time is approximately 0.3 seconds while in the preceding batches the average run time per image is about 0.2 seconds. This suggests that the speed is relatively high than other approaches like [4], [3], [16] and [13]

It is evident that our approach achieved an average mean precision value of 0.8774 which outperforms other approaches that have been used in marine vessel classifications. Samples of the results output are shown in Fig. 5. The Tanker Ship which is usually physically big in size scores highly in AP value compared to other classes. This may suggest that the physical appearance of a ship can be a factor for consideration in ship classification. This high performance by using Faster R-CNN of an average mean precision value of 0.8774 is an excellent performance and it confirms that the method can be utilised in marine vessel classifications which can address the port security challenges. The potential of our approach may be related to outperform many others and even solve their weak points including the following:

Table 6: Why Faster R-CNN is better than other approaches

Approach	Limitations
Ship classification based on super-structure features in SAR images[3]	Difficult in setting up the algorithm threshold
Maritime vessel images classification using Convolution Neural Network[4]	Slowest with more complex configurations of network. Accuracy at an average of 0.8 only on simple configurations.
S-CNN based ship detection[16]	Higher recall rate but slowest in training making it not applicable on real-time scenario
Ship recognition using optical Imagery[13]	The region of interest is not possible for smaller ships. Lower recall capability.

4.3.2 Sample images of output results



Fig. 12 Sample results for Container Ship

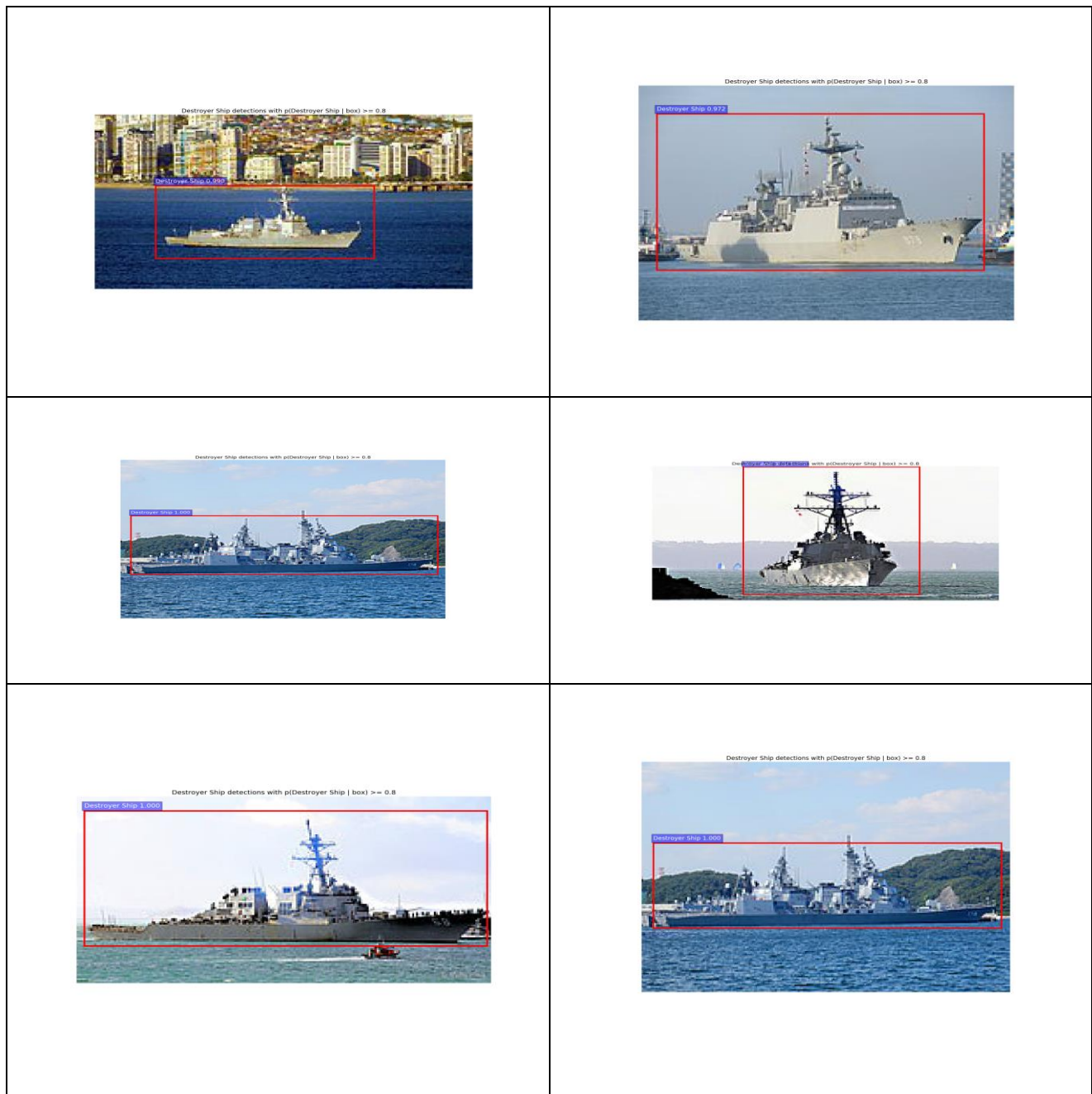


Fig. 13 Sample results for Destroyer Ship

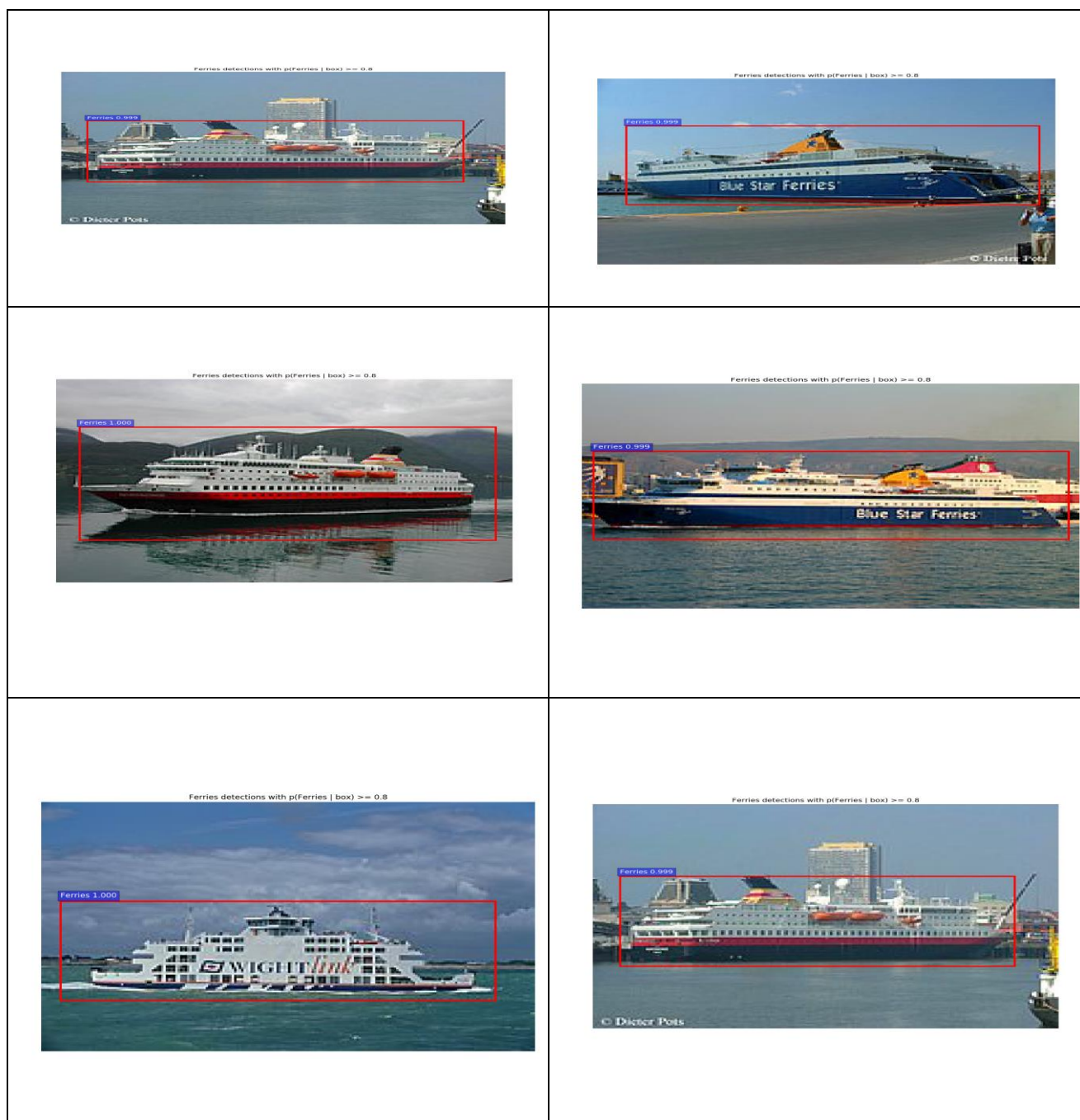


Fig. 14 Sample results for Ferries



Fig. 15 Sample results for Fishing Vessel

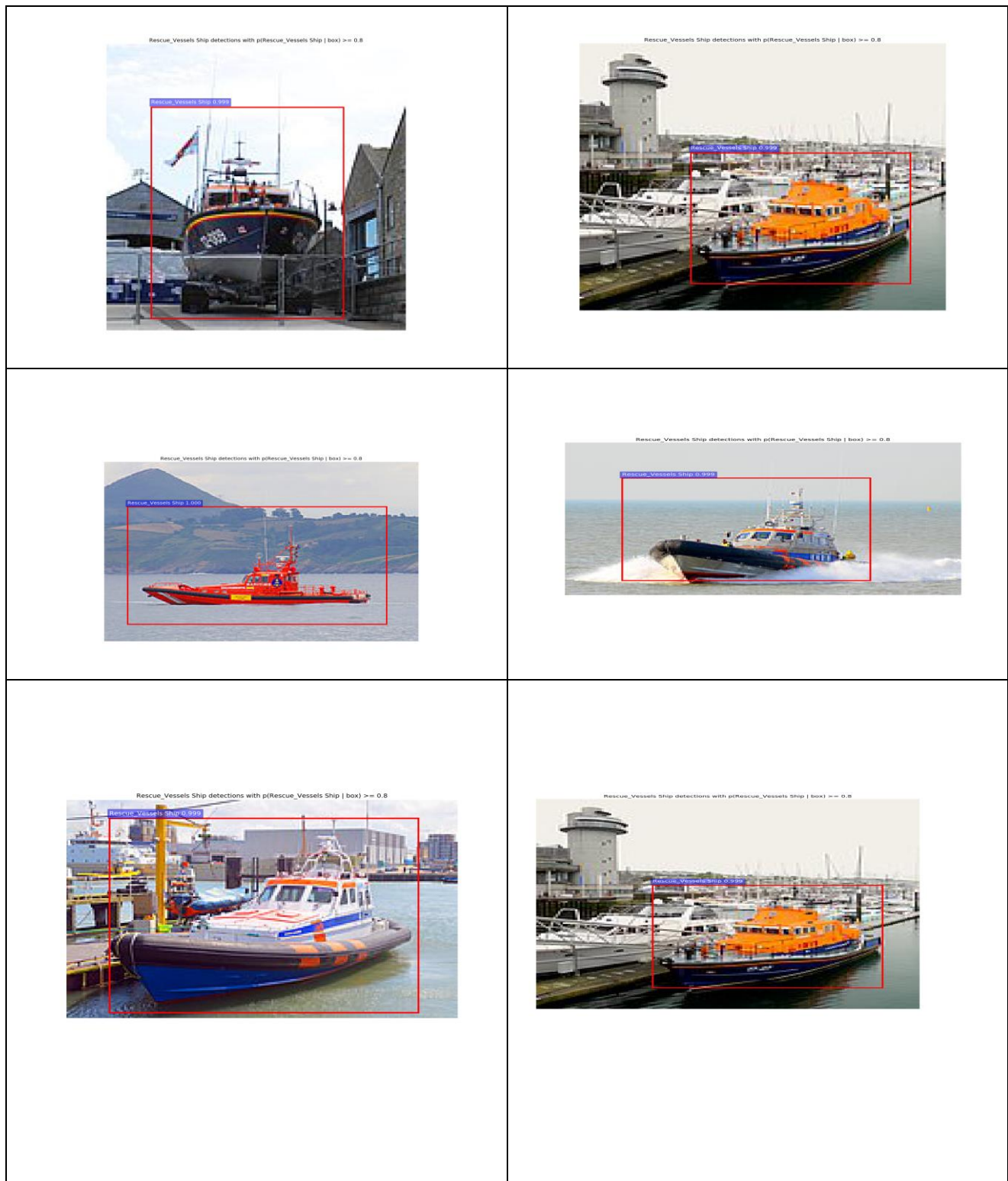


Fig. 16 Sample results for Rescue Vessel

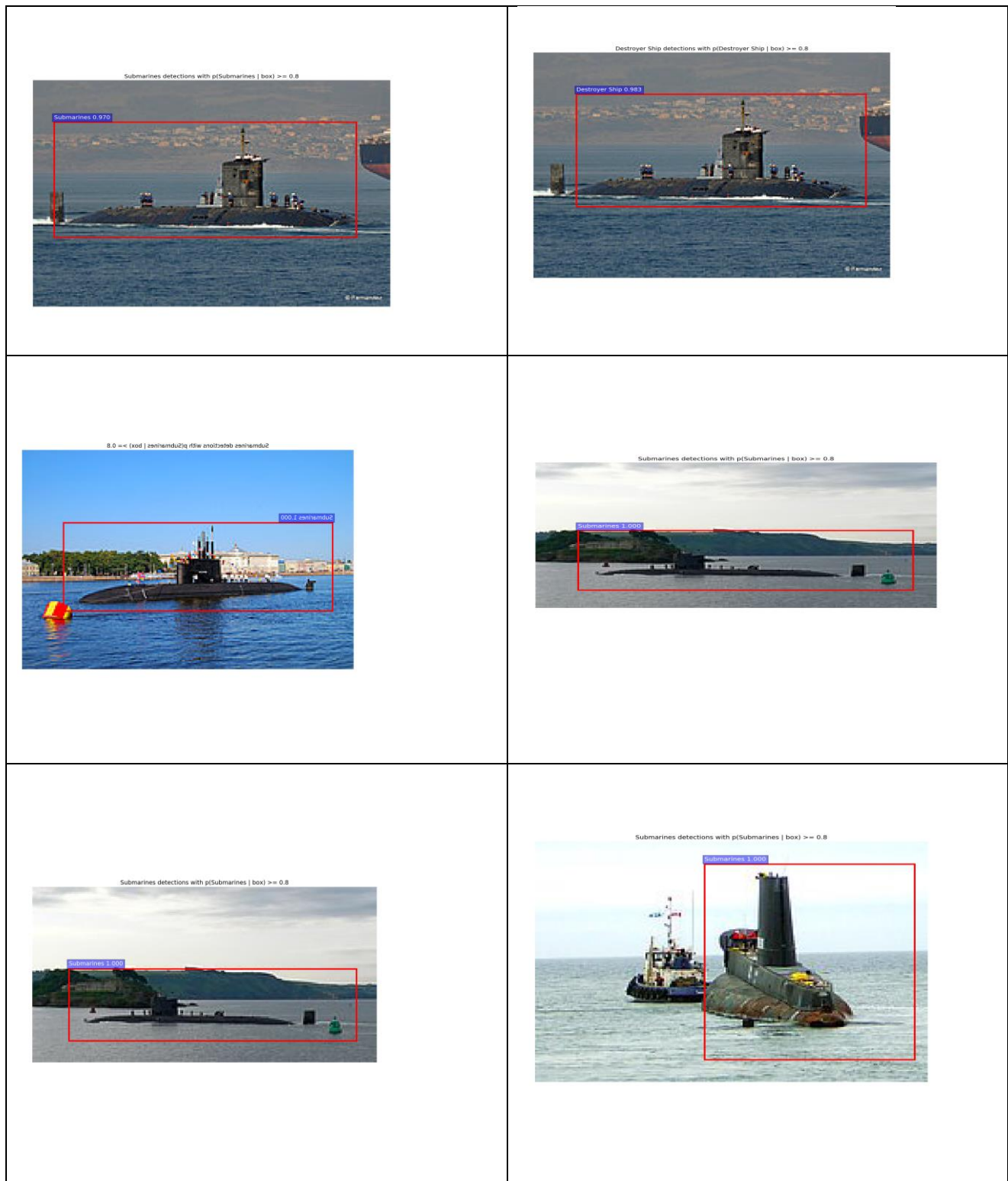


Fig. 17 Sample results for Submarine



Fig. 18 Sample results for Tugs

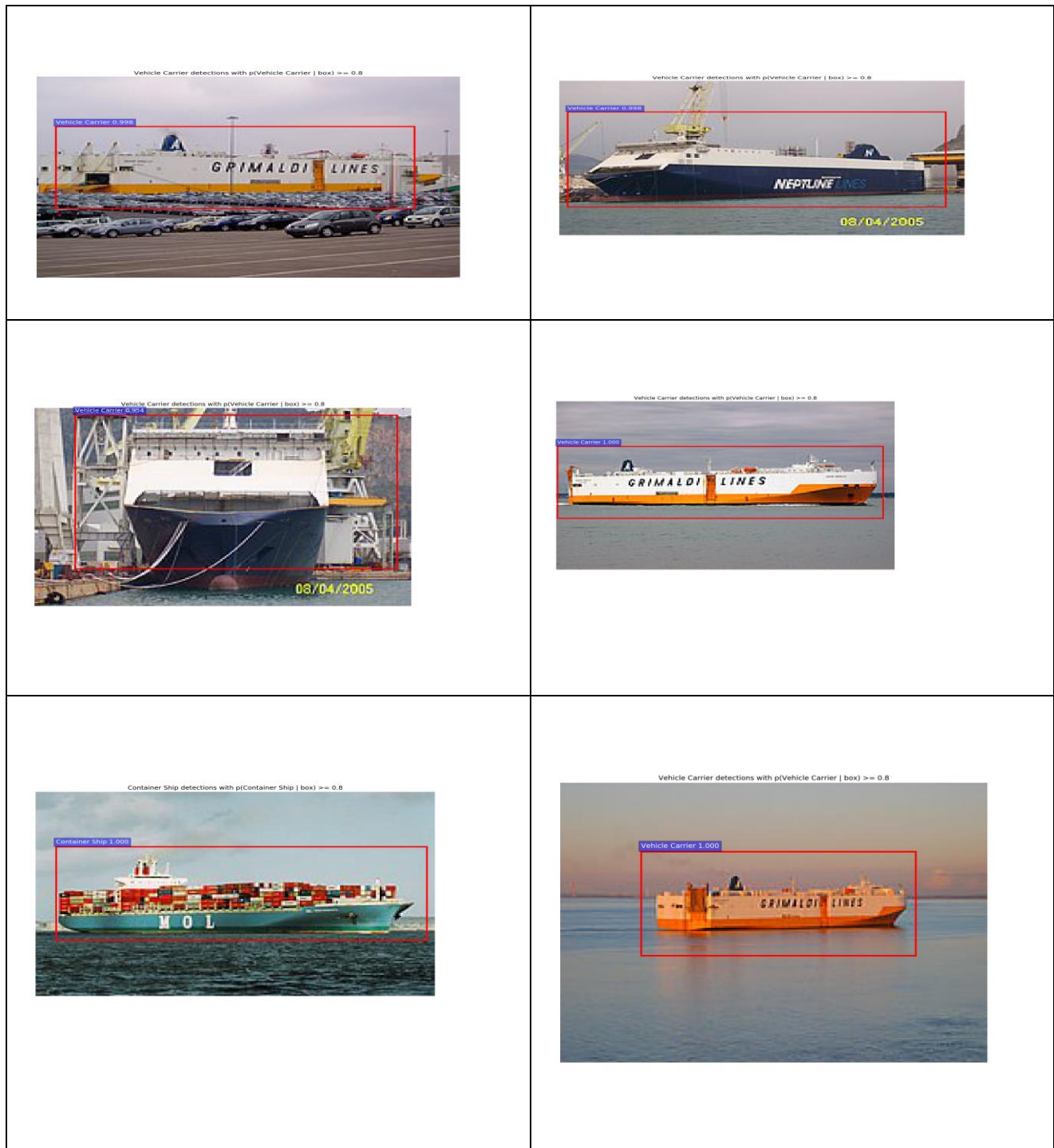


Fig. 19 Sample results for Vehicle Carrier



Fig. 20 Sample results for Tanker Ship

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.2 Conclusion

In our study, we present the application of Faster R-CNN method for classification of 9 classes of ships for the purpose of automatic classification in a maritime environment. The choice of this method is arrived at after detailed analysis of other state of the art object classification models including the analysis of Faster R-CNN model on other distinct objects apart from the ships. The classification process was done following two experiments. In each of the experiments, ship images of at least 500 pixels on the shortest side for the purpose of uniformity were used. Otherwise the approach could take any size of image since its functionality is a scale invariant type of model. The first experiment is useful in determining the appropriate number of images which could give consistent results on correct classification of ships. It is also important consequence that the number of images have a positive correlation to the accuracy scores of the model. In the second part of the experiment, we use the method to classify the 9 classes of ships, which are the target group of ships in the study.

From the results, it is evident that our approach achieves an average mean precision value of **0.8774** which outperforms other approaches including [4], [3], [16] and [13] which have been used in marine vessel classifications. Apart from the high accuracy rates, our approach is quite fast on run time per image compared to Fast R-CNN and CNN methods. In fact, the speed is 25 and 250 times faster than Fast R-CNN and CNN respectively. This may be considered a better classification system for ships which do not need two way communication like the case of AIS. By incorporating it with the AIS, it may be a suitable method for automatic classification of ships.

In the case of fishing vessels which had relatively lower score on classification precision, more understanding of ship image features may be necessary to correct the poor results. In addition, the marine environment backgrounds are highly cluttered and this could be the cause of AP scores less than one or 100% which is an ideal case reference. Our approach can be improved further by utilizing the infra-red camera images to address issues of image quality caused by weather effects and night times. This is therefore useful outcome which can be applied in marine technology for port security management.

5.2 Recommendation

Based on our research results, the following recommendations are highlighted for future considerations in research;

1. Mechanisms for obtaining full object features of ships regardless of weather conditions, day and night acquisition of ship images and the motion factor in ships should be considered in future research
2. Incorporation of Deep Learning approaches to AIS should be focused in future researches to enable real-time classification of ships for automatic identification.
3. It might be possible to develop an electronic marine highway which can solve challenges facing ports and marine environments.

5.3 Conceptual framework for Faster R-CNN-AIS ship identification system

Input images

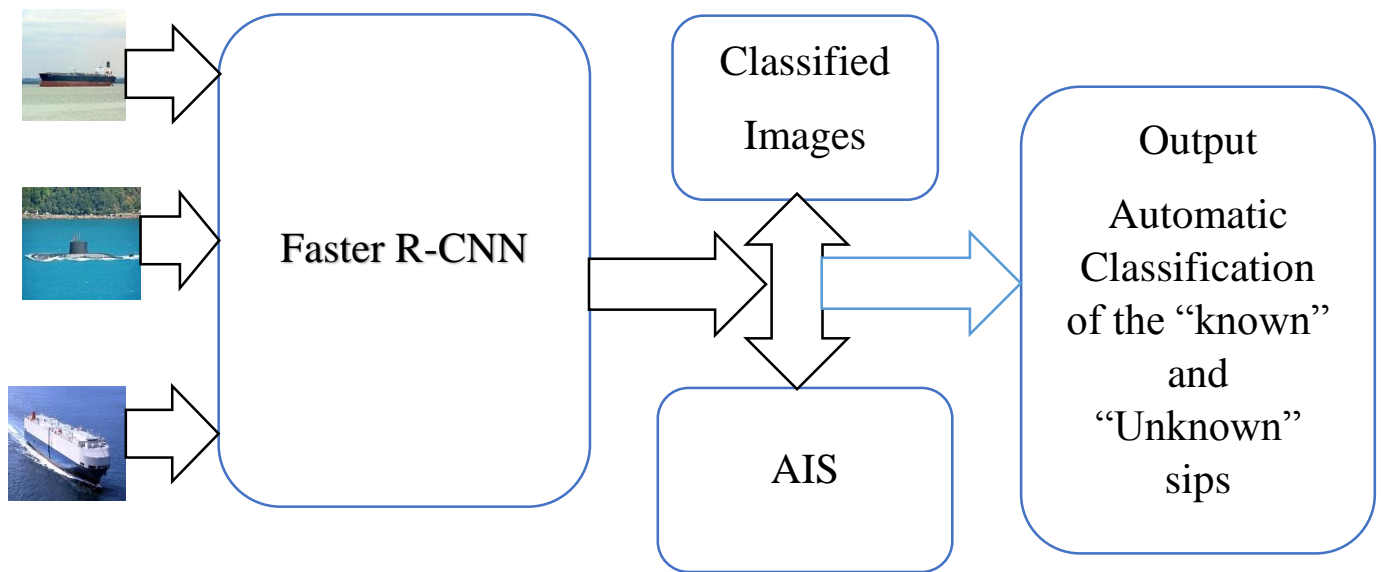


Fig. 21 Faster R-CNN-AIS conceptual model

The study proposes the integration of Faster R-CNN approach with AIS in order to achieve real time identification of both “*known*” and “*unknown*” ships. The input may be obtained from both stationary cameras and infrared cameras to cater for day and night data. Further research is also recommended to solve challenges facing the image qualities due to the highly clustered marine environment. The conceptual framework may be utilized in monitoring and controlling maritime traffic around the ports hence improving security and efficiency. This is an important baseline for future development of electronic highway which is envisaged to solve many problems faced by ports as well as sea transport industry.

REFERENCES

- [1] M. Conry, A. Keefe and W. Ober, "security; BIOSwimmer; Enabling technology for port," *Technologies for Homeland Security (HST), 2013 IEEE International Conference on*, no. 12-14 Nov. 2013, pp. 364-368, 2013.
- [2] Y. Xuefeng, Z. Yingjun and L. Wen, "Ship behavior recognition based on infrared video analysis in a maritime environment," *ITS Telecommunications (ITST), 2015 14th International Conference on*, no. 2-4, pp. 80-84, 2015.
- [3] M. Jiang, X. Yang, Z. Dong, S. Fang and J. Meng, "Ship Classification Based on Superstructure Scattering Features in SAR Images," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 5, pp. 616 - 620, 2016.
- [4] C. Dao-Duc, H. Xiahui and O. Morere, "Maritime Vessel Images Classification Using Convolution Neural Networks," *Proceeding. SoICT 2015 Proceedings of the Sixth International Symposium on Information and Communication technology*, pp. 276-281, 2015.
- [5] S. Ren, K. He, R. Girshik and J. Suan, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, 2017.
- [6] R. Lippmann, "An introduction to computing with neural nets," *IEEE Signal Processing Magazine.*, vol. 4, no. 2, pp. 4-22, 1987.
- [7] C. L. Zitnick and P. Dollar, "Edge Boxes: Locating Object Proposals from Edges," in *ECCV*, Zurich, 2014.
- [8] S. Raschka, *Python Machine Learning*, Birmingham: Packt Publishing, 2015.
- [9] R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, 2014.

- [10] K. E. A. V. D. Sande, J. R. R. Uijlings and T. Gevers, "Selective Search for Object Recognition," in *Computer Vision (ICCV), 2011 IEEE International Conference on*, Barcelona, 2012.
- [11] P. F. Felzenszwalb, R. B. Girshick, D. McAllester and D. Ramanan, "Object Detection with Discriminatively Trained Part-Based Models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1627 - 1645, 2010.
- [12] S. Peng, H. Jiang, H. Wang, H. Alwageed and Y.-D. Yao, "Modulation classification using convolutional Neural Network based deep learning model," in *2017 26th Wireless and Optical Communication Conference (WOCC)*, Newark, NJ, USA, 2017.
- [13] P. A. Feineigle, D. D. Morris and F. D. Snyder, "SHIP RECOGNITION USING OPTICAL IMAGERY," in *Proceedings of Association for Unmanned Vehicle Systems International (AUVSI)*, Washington DC, 2007.
- [14] S. Dugad, V. Puliyadi, H. Palod, N. Johnson and S. J. Simran Rajput, "Ship intrusion detection security system using image processing & SVM," in *2017 International Conference on Nascent Technologies in Engineering (ICNTE)*, Vashi, Navi Mumbai, India, 2017.
- [15] M. Kang, X. Leng, Z. Lin and K. Ji, "A modified faster R-CNN based on CFAR algorithm for SAR ship detection," in *2017 International Workshop on Remote Sensing with Intelligent Processing (RSIP)*, 18-21 May 2017.
- [16] R. Zhang, J. Yao, K. Zhang, C. Feng and J. Zhang, "S-CNN-BASED SHIP DETECTION FROM HIGH-RESOLUTION REMOTE SENSING IMAGES," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vols. XLI-B7, pp. 423-430, 2016.
- [17] R. Girshick, "Fast R-CNN," in *2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, 2015.
- [18] K. E. A. v. d. Sande, J. R. R. Uijlings, T. Gevers and A. W. M. Smeulders, "Segmentation as selective search for object recognition," in *Computer Vision (ICCV), 2011 IEEE International Conference on*, Barcelona, Spain, 2011.

- [19] SHIPSPOTTING.com, "Ship Photos and Free AIS! The worlds largest online ship image archive.," shipspotting.com , 24 July 2017. [Online]. Available: <http://www.shipspotting.com>. [Accessed 24 July 2017].
- [20] K.J. Ngeno and H. Watanabe, "Ship Identification for Port Security using Faster Region Convolution Neural Network," in *IEICE General Conference, BS-1-12*, Nagoya, Japan, March, 2017.
- [21] S. Raschka, "training Artificial Neural Network for Image Recognition," in *Python Machine Learning*, Birmingham, Packt Publishing, pp. 383-384, 2015.
- [22] K.J. Ngeno and H. watanabe, "Ship Classification Using Faster Region Convolution Neural Network (Faster R-CNN) for Automatic Identification of Marine vessels," in *Forum on Information Technology, FIT*, Tokyo, Japan, Sept, 2017.(to appear)

LIST OF ACADEMIC ACHIEVEMENTS

Conference Publications

1. K. J. Ngeno and H. Watanabe, "Ship Identification for Port Security using Faster Region Convolution Neural Network," in *IEICE General Conference, BS-1-12*, Nagoya, Japan, March, 2017.
2. K. J. Ngeno and H. watanabe, "Ship Classification Using Faster Region Convolution Neural Network (Faster R-CNN) for Automatic Identification of Marine vessels," in *Forum on Information Technology, FIT2017*, Tokyo, Sep. 2017 (to appear).