Detecting Ships from Satellite Images using Deep Learning Technique

MSc Research Project
MSc in Data Analytics

Sneham Mukherjee
Student ID: x19224826

School of Computing
National College of Ireland

Supervisor: Dr. Barry Haycock

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Sneham Mukherjee
Student ID:	x19224826
Programme:	MSc in Data Analytics
Year:	2021
Module:	MSc Research Project
Supervisor:	Dr. Barry Haycock
Submission Due Date:	31/01/2022
Project Title:	Detecting Ships from Satellite Images using Deep Learning Technique
Word Count:	5960
Page Count:	19

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Sneham Mukherjee
Date:	31st January 2022

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Detecting Ships from Satellite Images using Deep Learning Technique

Sneham Mukherjee
x19224826

Abstract

One of the most powerful and successful study areas in the marine surveillance application is ship detection from satellite photos. Beyond that, the suggested research will aid in the development of a powerful navy defense system, the tracking of illicit sea route activities, the tracking of fishing ships, the detection of lost ships, and so on. Detecting ship instances from satellite images is a challenging job. This study will present a Faster R-CNN model architecture based on Feature Engineering. Different types of augmentation techniques like Gaussian filter, Edge detection, horizontal flipping, random cropping, scaling and changing the colour of the image are used along with the Affine transformation. Adam Optimizer is used to construct the model and Binary Crossentropy is employed as a loss function. The model is trained with 50 epochs to get the best accuracy. The findings revealed that data augmentation can aid in the improvement of the model's performance. The accuracy achieved from the model are 98.16% and 97.33% on validation and test data respectively with a precision score of 0.9820 and 0.9736. The method described in this research, when compared to the prior method in the literature review, may considerably enhance the efficiency of ships detection in large aerial photographs while also boosting accuracy in terms of accuracy.

Keywords— Satellite,Defence System, Faster R-CNN,Augmentation,Gaussian filter, Edge detection,Affine Transformation.

1 Introduction

The demand for new strong technologies for ocean surveillance has risen significantly, widely in the field of maritime security and maritime safety. The tracking of nearshore and coastal ships is a tough process for a wide range of security and commercial applications. In the civil sector, ship recognition, for example performs a key supervisory role in monitoring and managing marine traffic, transportation, trash dumping, and illegal trafficking. In the military, monitoring ship geolocation, size, direction, speed, and other facts can be used to detect whether ship cross-border or other aberrant behaviours exist, improve the security of the coast and sea. Remote sensing imaging technology has developed to the extent where it can now track the whole surface of the ocean, sparking a surge in interest in research in aquatic target detection and classification innovation. With the advent of sensor technology, more resources and funding have been invested to further research by the government. Due to many aspects such as weather, climate, ship size, structure, and time expensive, ship detection systems using traditional methods faced a major hurdle. Because optical imaging images are more susceptible to human visual interpretation, they are more successful for observing the earth's dynamic surface.

This deluge of new satellite images is surpassing companies' ability to manually scrutinize each image captured, prompting the use of machine learning, deep learning and computer vision algorithms to aid in the analysis process.

To overcome the situation, Deep learning is introduced to resolve the issue and deliver consistent outcomes. Deep learning boosts accuracy of classification of optical aerial images without the need for manual feature extraction. Deep learning has received greater success in object detection tasks primarily to the powerful feature extraction capabilities of convolutional neural networks. Deep learning is used to address intricate complexities of computer image processing (e.g. image colourization, identification, classification, segmentation and detection). Convolutional Neural Networks for example, have pushed the frontiers of what is attainable by enhancing prediction performance using vast volumes of data and huge computational resources. Problems that were formerly thought to be unsolvable are now being tackled with superhuman ability. Deep learning-based object detection algorithms are categorized into two types: two-stage detection algorithms like Faster R-CNN and single-stage detection algorithms like SSD, RFBNet, and YOLO. Single-stage detection algorithms have an absolute benefit in terms of speed, however two-stage detection algorithms have a high localization accuracy. Both kinds of methods are routinely used in applications such as automated driving, intelligent security, and remote sensing detection. A good example of this is object detection. Since its reintroduction in 2012 by (Krizhevsky et al.; 2012), DL has conquered the domain due to the enhanced operational excellence than traditional methods.

Figure 1: Satellite Image near the Harbour

Traditional ways of extracting artificial features are inefficient and have poor adaptability. Previously, (Chen and Gao; 2018) proposed a Faster RCNN by changing the image aspect ratio and by introducing Soft-NMS and boosted the performance of traditional Faster R-CNN to 77.7 percents in mAP.

This paper proposes deep learning technique Faster R-CNN based on features engineering to detect the ships from the satellite images. This research paper's conclusions will contribute to the creation of a better naval surveillance system with better accuracy and will make it easier to answer the research question.

1.1 Research question

To what extent deep learning algorithm like Faster R-CNN can improve the efficiency in detecting ships object from satellite images with the help of features engineering based on Data Augmentation and Transformation.

1.2 Research Objectives

The study's key goals are as follows:

- A comprehensive examination of existing work on ship detection.
- Detecting the ship and no ship from satellite images using Deep Learning Technique.
- Executing Faster R-CNN model with the help of feature engineering based on data augmentation.
- Fine tuning the model during training phase to obtain a better effective model.
- Measuring number of epochs needed for better model performances.

The following is how this document is organized:

- Section 1 contains Introduction.
- section 2 contains the Related Works.
- section 3 contains the Methodology.
- section 4 contains the Design Specification of the model.
- section 5 contains the Model Implementation.
- section 6 contains the Evaluation and Result.
- section 7 contains Conclusion and Future Work.

2 Related Work

Past experiments employing multiple deep learning models and transfer learning methodologies to detect and categorize ship objects from aerial and satellite photos are detailed below. These studies are optimistic and valuable in bringing research forward in the right path.

2.1 Traditional Machine Learning Approaches for Image Classification and Detection

For maritime security, border control, and the security of numerous commercial sites, surveillance is a major issue. To overcome this issue, (Dugad et al.; 2017) proposed a state-of-the-art with the help of HOG and SVM. Their main motivation is to create model using machine learning concept SVM to detect ships from pirate threats at the seashore. Initially, they have divided the image into positive and negative samples so that SVM can use them to construct a hyperplane which will act as a decision boundary. After that they have used HOG descriptor vector with SVM to calculate the positive and negative samples. With this experiment, the authors have got a good accuracy in detecting ship candidates even in bad weather conditions.

In the paper, (Huang et al.; 2017) introduced a novel ship detection technique for detecting ship objects from the aerial images using structured forest edge detection method integrated with SVM. The main purpose of this technique is to create a good model on seascape images which are mainly affected by sea-waves. Due to various shape and size of the images, the paper emphasizes more on the region proposal extraction before doing classification. So they designed an edge detection method as it requires a low number training set and then generates contours by suppressing the background. On the images, they applied morphological processing and SVM as a classifier to achieve a good result. The proposed method is evaluated on the basis of Precision, Recall and False Alarm rate and compared with different model. The paper should have used more data for the training and testing purpose in order to achieve a better accuracy. The Precision, Recall and False Alarm Rate scores are showing as 86.67%,87.54% and 13.33% respectively which is more effective on non-uniform satellite images with a tiny loss in recall values.

In the paper, (Zong-ling et al.; 2019) proposed a novel methadology based on ML and morphological matching and can be used in both static and real time ship target. Background noise is suppressed to enhance the gray scale of the objects. The obtained region is detected and then with the help of histogram variance comparison, target region and extracted target region are differentiated. Finally, they applied machine learning network model like VGG16 for detection and classification of the ship objects. The efficiency of morphological matching is 60% whereas the efficiency of machine learning for target recognition is more than 60% with a precision score of 0.97. Because of the complex background and high false rate, this model is able to detect ship objects with good accuracy. But still there is a room for improvement. Due to low resolution rate, there is a chance of overfitting. Also, network pruning is required to reduce the amount of variables and processing time. The use of ship detection algorithm is an important need for states. (Kartal and Duman; 2019) proposed a fast running ship detection system from satellite images with the help of deep learning method and it is an open source and named as Tensorflow Object Detection Application Programming (TODAPI). The system does not require high configured hardware as it can work on a low configured machine also and can be used as object detection API.

In ship detection study, the main challenge is low intensity images or an image with background noises and lead to difficulty in detecting ship objects from satellite images. To rectify the issue, (Morillas et al.; 2015) introduced a new ML technique based on SVM with the help of texture and colour features. They proposed a block divisions that segregates the images into small pixel blocks with the parameter ship and no-ship. SVM is used to classify using texture and colour features of each block. Then the classification is performed and detection method is improvised with the help of reconstruction algorithm. It rectifies inappropriate block and detect the correct one from them. The final model gives a precision score rate up to 96. 98% with an accuracy of 98.14%.

2.2 Use of CNN in Detecting Ship Objects

Along with the exponential rise of deep learning, ship detection using VHR images is now a popular topic recently. A new novel Convolutional Neural Network method based re-normalization method is introduced by the (Wang and Gu; 2018) in their paper. First, CNN is integrated to detect the correct ship objects from the total images on the basis of rotation, location and scale in groups. Now in order to achieve a uniform group, shape information is corrected with the help of renormalization net. The proposed method achieves an accuracy of 95.6% with an IOU of 93% in detecting ship candidates.

In order to predict ships in more diverse weather conditions, (Li et al.; 2021) in their paper introduced a novel CNN approach with the help of rotation bounding box. Some common deficiency with CNN approaches is solved by this method. Initially, Rotated suggestions of excellent quality with a dual-branch regression network are introduced and compared with other models. The authors next introduced a unique multidimensional adaptive pooling strategy to reduce the inconsistent sampling crisis created by classic pooling techniques and to integrate multilevel characteristics, using a spatially unique pooling procedure. Finally, a comprehensive study of the proposed techniques is held out, and the performance of the proposed model was compared against some of the state of the art CNN derived models. Upon comparing the proposed model found out to be more effective and efficient while classifying and detecting ships over the CNN based models.

In the paper, (Wang et al.; 2021) put forwarded a novel solution based on Lightweight CNN, and a unique Multi source Feature Cascade Decision. Firstly, they introduced an iterative segmentation to pre-process the complex weather scenes. Secondly, they extracted the ship objects with the help of multivariate Gaussian distribution to increase the recall. Then they employed optical panchromatic data to facilitate in the training of restricted infrared data. And finally, to eventually minimize false alarms, the specialized attributes from the proposed model were used. Experimental results showed that their proposed method works well with good accuracy and precision even in the bad weather conditions. The solution could be further upgraded by adding more images as training dataset which can help to better generalization of the model resulting in enhancing the accuracy.

A fluid weighted sampling and flexible threshold sliding multifaceted approach is presented by (Zhang et al.; 2021) to solve the prevalent problem of class imbalance in SAR image ship object identification. Fluid weighted sampling is an information approach that only regulates the selection criteria during the learning phase, with no data distribution adjustments made ahead of time. The flexible boundary shifting in the test phase just affects the throughput percentages of Convolutional Neural Network and does not change their properties. As a conclusion, their methodology is suitable for a wide range of the conventional Convolutional Neural Network-based SAR visualization identification models. OpenSARShip dataset demonstrates the accuracy of the author's solution. Classical normalization strategies obtain training accuracy of around 80 percent in the VH frequency and 78 percent with the VV frequency. Also, during the validation phase, the proposed solution also yielded exceptional accuracy with minimal loss.

2.3 Introduction of Transfer Learning and Different Network Architecture in Detecting Ship Candidates

(Li, Guo, Gao and Chen; 2019) in their paper proposed a novel approach based on WHOG features for ship detection. It discovers fixed resolution ships leveraging labeled ships at multiple resolutions. At different resolutions, the training and test samples support different distributions. Without conducting spatial alignment, the JDA approach incorporates statistical modification. The proposed MA-JDA solution is provided as a workaround to overcome issues with detecting ships with low resolution images. To optimize transfer learning performance, the authors have hypertuned the inner parameters of the model and added additional layers to the transfer learning model. While testing the model with fresh images, the test accuracy showed an exceptional accuracy.

(Li, Ding, Zhang, Wang and Chen; 2019) in their project represents a new SAR ship dataset of 2900 frames and 7524 ships under particular scenarios that is considered to be the benchmark for classification and segmentation of boat images. ResNet is regarded as the most advanced and up to date technology when it comes to domain of ship detection from areal images. On a test set involving 580 frames and 1881 ships, their algorithm yields a 94.7 percent AP. In comparison to YOLOv2 and VGG16, their solution delivers higher accuracy and training speed, indicating its efficiency.

More SAR photos are available than ever before since the debut of space-based satellites, facilitating for dynamic ship surveillance. Object detectors in deep learning, such as ResNet, FPN, YoLo, and SSD Net, enter the top performance. SSD algorithm outperforms YoLo in order to get an enhanced, but it challenges to locate small targets such as SAR images. (Chen et al.; 2020) developed an improved method to tackle this question. On the basis of the proposed model by the researcher, a novel solution was constructed, with two different components. The first component i.e., the deconvolution module was principally involved to incorporate high-level contextual features to the low-level network's feature evidence in order to maximize detection accuracy. Because a predictive component, which is made up of residual networks, can extract rich information and transmit it into classification and regression issues, they named it SSDv2. The SAR-Ship-Dataset is used to test their algorithm. 102 Chinese satellite imagery, around more than 100 Sentinel-1 satellite photos were trained to compile a SAR-Ship-Dataset. It is thought up of 43,819 512-pixel range and azimuth ship chips. After running the model with suitable number of epochs, the proposed model reached a decent score of 92% topping the accuracy score of leading existing solutions in the domain of ship detection.

As a consequence of the rapid advancement of chip technology and the deep learning revolution, many ship detection frameworks for remote sensing radar pictures rooted on Convolution Neural Networks have been designed and have achieved substantial trademarks. However, there are challenges that are delaying their development: 1) Using strong connectivity to retrieve information for SAR ship identification is cost prohibitive since it adds to the compute burden and increases the inference time investment; 2) Techniques with anchor being involved contain a large number of hyperparameters that may be fine-tuned to achieve negative recognition rate. This study designs a low, efficient ship detection network for SAR imaging to fix the problems. To initiate, a basic U-Net is offered as the backbone for extracting the features. Most of the transferred learning models used have 2.37 %, 0.76 %, 0.34 %, 1.01 %, 0.55 %, and 1.07 % learnable weights, respectively. Furthermore, a model which uses no anchors was utilized the results were compared with the first model. Extensive tests are performed and a more comprehensive set of analysis methods were used to gauge the efficiency of their strategy. On the SSDD dataset, the suggested solution obtained sixty-eight percent average precision and sixty-percent average recall, respectively (Mao et al.; 2020).

In the paper, CNN, ResNet, VGG, and DenseNet deep neural networks are used by (Ajay and Raghesh; 2021). Here, a segmentation data pre - processing technique was employed. To identify the optimal model, the F2-score of the model is assessed with and without segmentation. The accuracy of the CNN algorithm that has been trained was 0.73. The accuracy of the transferred learning ResNet50 model with classification, which also was evaluated on the training data, was 0.86. After segmentation, all of the models achieved a higher accuracy metric in boat image classification. Due to the model's exceptional accuracy while working with new and untrained data, this proposed framework has become a go-to tool in the area of satellite image detection.

2.4 Advent of R-CNN in the field of ship detection study

(Wei et al.; 2020) in their paper introduced a mage preprocessing method for remote sensing ship photos. Prior to ship detection, the dark channel analyzes the dehazmg input photo, and then the grey world technique is employed to calibrate the dehazmg image's color stability. The aspect ratio of the anchor has been presented for different for the ship's geometric properties. Dilated convolution enhances the Faster R-CNN detection framework. The proposed strategy enhances the convolution layer's receptive field. Finally, an image preprocessing algorithm and a ship detection network are coupled into a remote sensing ship detection software platform and achieved an accuracy of 70.85%.

Identification approaches built on DCNNs that are currently available in the nearshore ship classifier devote low priority to the reduction of erroneous onshore alarms. Even if the learning algorithm has a high detection accuracy, numerous boat shaped images, such as docks, rooftops, and roadways are all considered as potential subjects of attention with a high possibility of occurring in real life.. In fact, integrating scene information to exclude the non-target area helps the user to locate targets. In this research (You et al.; 2019), a Scene Mask R-CNN-based DCNN-based ship tracking algorithm is designed to decrease onshore erroneous alerts. The R-CNN system is a complete system with 4 segregated systems, each with its own function. The image data is necessary to aid the detection method, which includes non-target zone exclusion based on the scene segmentation sub-network. The model performed decently with training data, however grappled while performing segmentation in naive data.

(Ke et al.; 2021) presented a revised Faster R-CNN which incorporates flexible convolution to boost the network's power to predict geometric transformations of structure ships, yielding in a 2.02 percent increase in mAP over Faster R-CNN. Unlike their model, which only uses resolved convolution in high-level layers, our improved Faster R-CNN uses non - rigid convolution in high-level layers to better extract the feature of shape-changeable ships, which has proven effective in our their experiments and will be further researched with visualization techniques in future work.

A CFAR method is being used in the paper by (Kang et al.; 2017) to enhance the productivity of the Faster R-CNN in the areal image identification challenge. VGG16 collects image features from satellite pictures and stores them in the RPN to provide area suggestions. Second, to calculate classification metrics and modified cluster centres, the Fast R-Convolution Neural Network uses region suggestions and feature maps. Finally, a novel image detector algorithm is used to choose the edges with the lowest classification score. Finally, the detection results include the results of the CFAR detector and targets with a high categorization score. Experiments involving SAR photo shows that integrating pixels-based detection with extensive attributes can enhance the levels of multidimensional boat identification and classification. Faster R-CNN fails to detect some small objects, resulting in a relatively poor detection rate in this study. As a result, one area of future work that can be completed is fine-tuning the model using data augmentation.

3 Methodology

To achieve the goals and answer the questions of this study, the proposed methodology will be used to recognize ship objects from satellite photos as indicated in the figure.

3.1 Data Gathering and Understanding

The accurate knowledge of the data must be acquired before the model framework creation. The dataset was obtained from Kaggle [1], a public resource that includes 4000 (80X80 RGB) pictures in a zip format. The images were taken from Satellite among them 3000 are labeled as "no-ship" and 1000 are labeled as "ship" The pictures are exported as.png files. Together with the data list, each image is stored as a list of 19200 integers. The first 6400, second 6400, and third 6400 entries, respectively, contain red channel, green channel, and blue channel values. The photos are presented in a row-by-row fashion. The ship image dataset also includes elements such as varied shapes, dimensions and weather conditions.

[1] https://www.kaggle.com/rhammell/ships-in-satellite-imagery

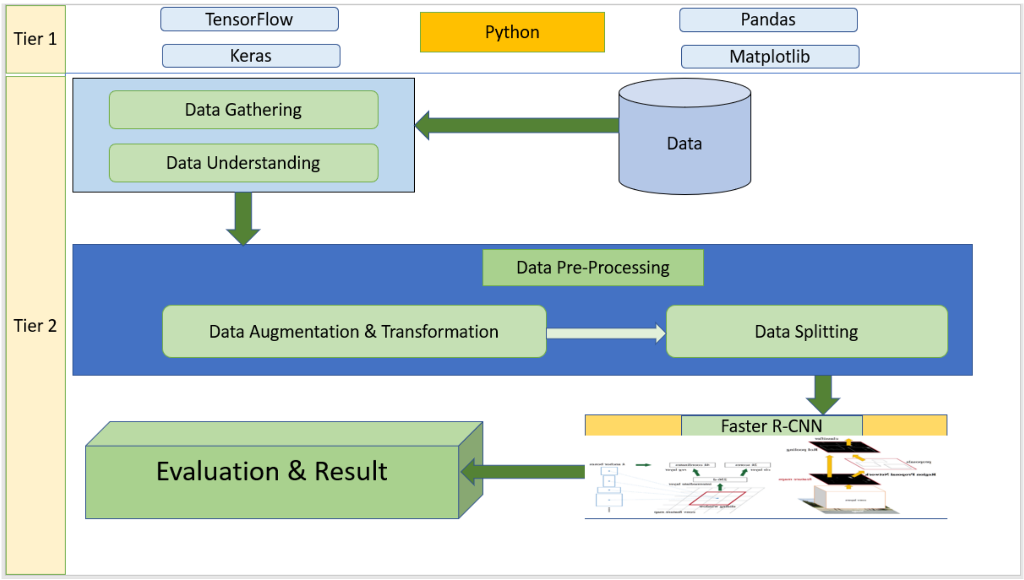

Figure 2: Research Methodology

3.2 Data Pre-Processing

Satellite pictures are impacted by weather and environment. As a result, the image has a lot of background noise by default. For this reason, Gaussian filter is used to eliminate background noise.

To keep the model from overfitting and underfitting, data augmentation is used. Horizontal flipping, random cropping, scaling, edge detection and changing the color of the image will increase the diversity of the training sample for low detection data. In terms of visualization, affine transformation is used to extend satellite images with diverse angles.

One hot encoder is used on labels numpy array with the help of to_categorical using Keras. This eliminates any unwanted bias in the dataset by bringing all categories on anequal footing in terms of labels.

3.3 Data Splitting-Training, Validation and Testing

The pictures and labels arrays are randomly mixed using the same seed value of 42 instead of train test split. This retains the pictures and their labels linked even after they've been shuffled.

The training and validation datasets are used to train the model, whereas the testing dataset is used to test the model on data that hasn't been seen before. Because the model has never seen this data before, it is used to realistically simulate prediction. It enables the developers to assess the model's robustness.

Splitting of Data

- Training-70 percent
- Validation-20 percent
- Testing-10 percent

4 Design Specification

Deep Learning technique Faster R-CNN is implemented to carry out the research work. This section explain the architecture of the model. Python Jupyter Notebook is employed as the fundamental tool to carry out the research for its ease of use and the utilization of different libraries such as Keras, TensorFlow, Seaborn, tqdm, Sklearn, and Matplotlib that are used to efficiently obtain the ability and build models. Jupyter Notebook is used for pre-processing and implementation. The output is then represented using the python matplotlib tools, Seaborn tools to help us to efficiently examine the model.

4.1 Faster R-CNN with RPN

The feature map produced by the convolutional layers is used for image detection in images using Faster R-CNN, i.e., the model drops the layers after the final convolutional block and passes the generated feature map to a Regional Proposal Network, which can either use a vanilla CNN model with fully connected layers or a Logistic Regression, Support Vector Machines, or Random forests.

- conv_block: The convulotional layer, batch normalization, and activation layers are all components of this function. The developer defines the number of filters, kernel size, and measures to be taken. This allows a developer to design the model without having to manually write the same lines. It also incorporates Python's OOPs principles, which are encouraged over C-style programming.

- basic_model: Using the previous algorithm, max pooling layers and dropouts, this function builds the model. A flatten layer is added after the requisite number of convolutional layers, along with dense layers, so that the image may be identified. The flatten layer transforms the feature map provided by the convolutional layers into a single classifying column.

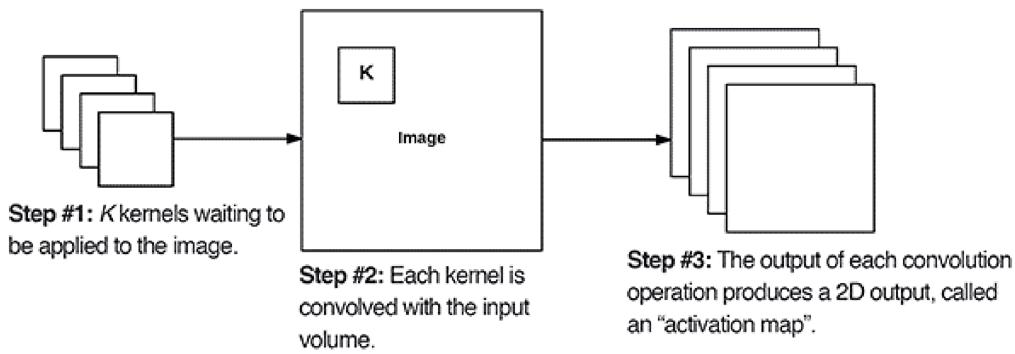

Figure 3: Working principle of Kernel in CNN

4.2 Explanation of Features

- Conv2D- The number of filters dictates what the convolutional layer learns in this 2-dimensional convolutional layer. The larger the number of filters deployed, the more data is obtained.

- MaxPooling2D- This minimizes the feature space of the convolutional layer's feature map without reducing range information. This boosts the model's stability significantly.

- Dropout- This removes a percentage of links between neurons in multiple stages that the user specifies. This makes the model extremely stable. It may be used in both fully linked and completely convolutional layers.

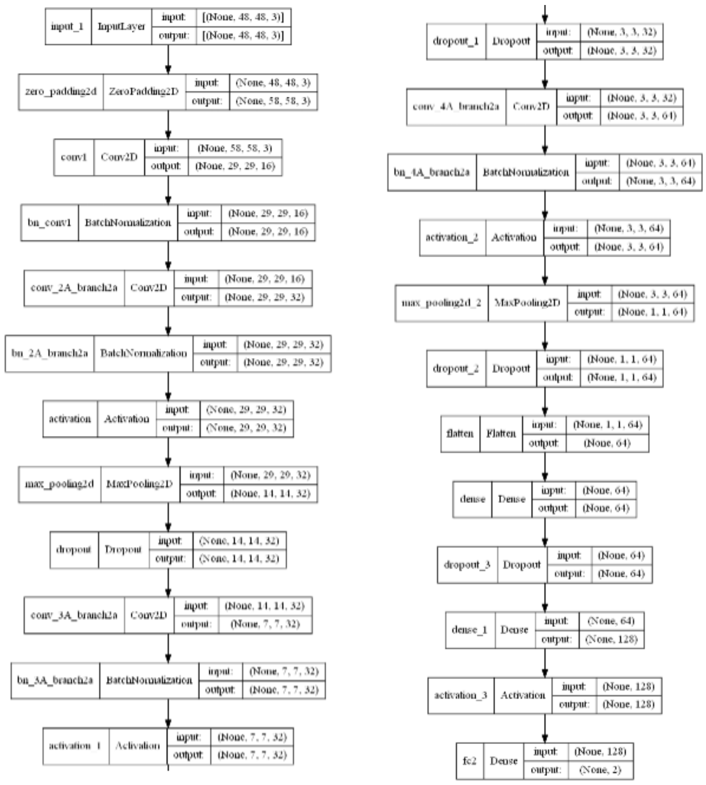

Figure 4: Faster R-CNN Model Architecture for Ship Detection

- BatchNormalization- The values in the hidden component of the neural network are standardized in this layer. This is comparable to how machine learning algorithms implement MinMax/Standard scaling.
- Padding- This applies zeros to the feature map/input image, keeping border features to persist.

5 Implementation

This section offers details on how the Faster R-CNN model is used to recognize ship and non-ship objects in satellite photos. Due of the lack of ship object photos, training deep learning models from scratch demands a vast quantity of data. A feature engineering approach is used to train both the models and boost performance.

5.1 Environment

The project's data loading, pre-processing, feature extraction, and evaluation are done in Jupyter Notebook.The host machine is Lenovo Yoga laptop with processor Intel(R) Core(TM) i7-1065G7 CPU @ 1.30GHz 1.50 GHz, RAM 16.0 GB and 64-bit operating system. And the important library used are matplotlib,seaborn,sklearn and tensorflow.keras.

5.2 Data Augmentation and Transformation

Bar plot and Pie Chart are drawn to find the count of the images per class and to find the percentage of class distribution respectively in the dataset.

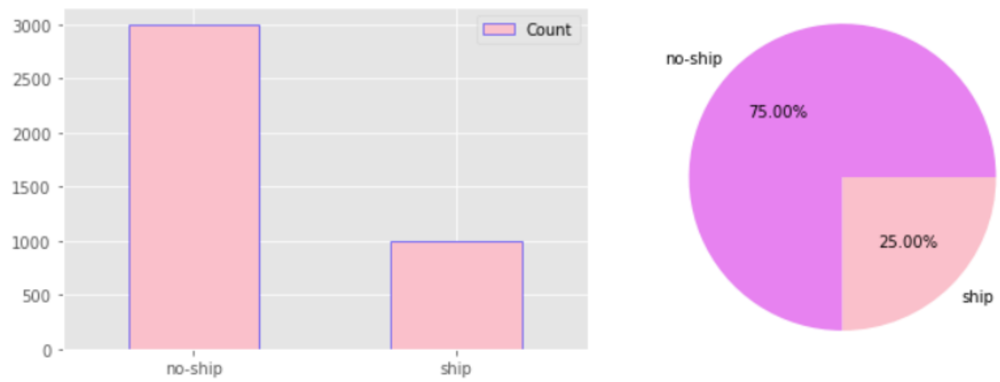

Figure 5: Bar Plot and Pie Plot of the unbalanced dataset

Because of imbalance in the dataset, the minority class is upsampled by randomly duplicating frames until the two sets have a normal proportion in the dataset. After then, the dataset will be partitioned into training, validation and testing sets by shuffling them randomly and then splitting them.

Another approach is to provide class weights to each individual class. The exact class weight is applied to each class. The greater the penalty, the higher the class weight. A larger penalty is applied to classes with a lower percentage. This enables the model to aggressively penalize itself if the class detected is erroneous.

With the help of iaa library in python, below tasks as a part of augmentation have been done:

Augmentation/Transformation	Action	Range/Value
Fliplr	Horizontally flip 50% of all images.	0.5
EdgeDetect	Search in some images either for all edges or for directed edges. These edges are then marked in a black and white image and overlayed with the original image.	Alpha= (0,0.7)
Crop	Random crop some of the images by 0-10% of their height/width.	Percent= (0,0.1)
LinearContrast	Strengthen or weaken the contrast in each image.	(0.75,1.5)
Multiply	In 20% of all cases, we sample the multiplier once per channel, which can end up changing the colour of the images.	per_channel=0.2
AdditiveGaussianNoise	For 50% of all images, we sample the noise once per pixel. For the other 50% of all images, we sample the noise per pixel AND channel. This can change the colour (not only brightness) of the pixels.	scale= (0.0,0.05*255), per_channel=0.5
Affine	Scale/zoom them, translate/move them, rotate them and shear them.	scale={'x':(0.8,1.2), "y":(0.8,1.2)}, translate_percent={"x":(-0.2,0.2),"y":(-0.2,0.2)}, rotate= (-25,25), shear= (-8,8)

The visuals in the ship class are augmented and then stored in the dataset to verify that almost all of the classes are represented equally. The current class ratio is 1:3, which means that for every image in the ship class, there are three images in the no-ship class. To counteract this, two enhanced images for each original image of the ship class are created. The dataset is balanced as a result of this.

Set AUGMENTATION to True if dataset augmentation is required. This will help to balance the dataset by adding minority classifications. Set AUGMENTATION to False to train with class weights.

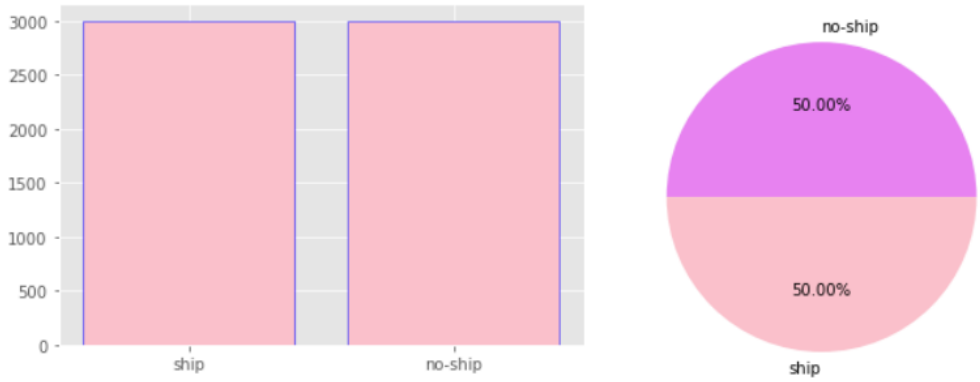

Figure 6: Pie chart after augmentation of the dataset

From the bar and pie plot as shown in figure 6, it can be seen that the dataset is balanced. The Up-sampling technique helps in balancing the data equally before split into training, validation and test data. Each class contain 3000 images of ship and no-ship images.

One hot encoder is used on labels numpy array with the help of to_categorical using Keras. This eliminates any unwanted bias in the dataset by bringing all categories on an equal footing in terms of labels.

5.3 Model Creation

- Input Block: The Input Keras layer is used to establish an input layer, and it defines the number of neurons in the input layer. To ensure that border characteristics are not lost, ZeroPadding is applied to the input image.
- Block 2: The first convolutional layer commences with 16 filters and a kernel size of (3,3), as well as strides (2,2). The output does not change spatially until the next block in which MaxPooling occurred since the padding is maintained constant.
- Block 3-4: Both have the same structure, with a convolutional layer, MaxPooling, and Dropout layers.
- Output Block: The feature map developed by the previous convolutional layers is flattened and then categorised using a Dense layer with the number of classes included in the dataset and sigmoid as the activation function.

5.4 Training Faster R-CNN

Faster R-CNN model is trained using the training and validation set. The Adam optimizer is used to construct the model, and the learning rate is set to 1e-3. Because there are just two classes, Binary Crossentropy is employed as a loss function. With a batch size of 16, the model is trained for 50 epochs. TensorBoard logs are preserved in the logs directory and the best model weights are maintained in the file. The learning rates are determined after a substantial percentage of trail runs and an investigation of the model's losses. The best epoch was chosen after training the model for 50 epochs on the basis of training accuracy, training loss, validation accuracy and validation loss.

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_1 (InputLayer)         [(None, 48, 48, 3)]       0

zero_padding2d (ZeroPadding  (None, 58, 58, 3)         0
2D)

conv1 (Conv2D)               (None, 29, 29, 16)        448

bn_conv1 (BatchNormalizatio  (None, 29, 29, 16)        64
n)

conv_2A_branch2a (Conv2D)    (None, 29, 29, 32)        4640

bn_2A_branch2a (BatchNormal  (None, 29, 29, 32)        128
ization)

activation (Activation)      (None, 29, 29, 32)        0

max_pooling2d (MaxPooling2D  (None, 14, 14, 32)        0
)

dropout (Dropout)            (None, 14, 14, 32)        0

conv_3A_branch2a (Conv2D)    (None, 7, 7, 32)          25632

bn_3A_branch2a (BatchNormal  (None, 7, 7, 32)          128
ization)

activation_1 (Activation)    (None, 7, 7, 32)          0

max_pooling2d_1 (MaxPooling  (None, 3, 3, 32)          0
2D)

dropout_1 (Dropout)          (None, 3, 3, 32)          0

conv_4A_branch2a (Conv2D)    (None, 3, 3, 64)          18496

bn_4A_branch2a (BatchNormal  (None, 3, 3, 64)          256
ization)

activation_2 (Activation)    (None, 3, 3, 64)          0

max_pooling2d_2 (MaxPooling  (None, 1, 1, 64)          0
2D)

dropout_2 (Dropout)          (None, 1, 1, 64)          0

flatten (Flatten)            (None, 64)                0

dense (Dense)                (None, 64)                4160

dropout_3 (Dropout)          (None, 64)                0

dense_1 (Dense)              (None, 128)               8320

activation_3 (Activation)    (None, 128)               0

fc2 (Dense)                  (None, 2)                 258

=================================================================
Total params: 62,530
Trainable params: 62,242
Non-trainable params: 288
_________________________________________________________________
```

Figure 7: Summary of Faster R-CNN Model

6 Evaluation

It is critical to evaluate the model's performance after it has been correctly implemented. This section focuses on the model's study as well as all of the aspects that have been fine-tuned in order to obtain the optimal model for this study research. During the assessment phase, the losses and accuracy of the process are evaluated for each epoch for the model. The charts of accuracy and loss are shown. To acquire a true positive and true negative result, confusion metrics are computed.

6.1 Training Accuracy, Training Loss, Validation Accuracy, Validation Loss

Table 1: Training Accuracy, Training Loss, Validation Accuracy, Validation Loss

Epoch	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1st	0.7934	0.4475	0.7942	0.4443
10th	0.9274	0.1825	0.9600	0.1211
20th	0.9529	0.1298	0.9808	0.0773
30th	0.9648	0.0966	0.9667	0.0878
40th	0.9743	0.0770	0.9808	0.0507
48th	0.9765	0.0841	0.9817	0.0621

From the table, it can be seen that at 1st epoch, the training accuracy and validation accuracy was 0.7934 and 0.7942 respectively and they are increasing after every 10 epochs as shown. Whereas the training and validation loss are decreasing steadily. At 48th epoch, the highest accuracy for both training and validation is achieved with a score of 0.9765 and 0.9817 respectively with a lowest loss. Thus it can be said that the model fitted well.

6.2 Plotting Accuracy VS Loss

Accuracy is a criterion for assessing the classification results. In most cases, it's given as a percentage. The frequency of forecasts when the calculated return is equal to the true value is known as accuracy. For a particular section, it is binary. During the learning phase, accuracy is typically graphed and monitored, nevertheless the value is frequently related with the overall or final model accuracy. Loss is more hard to comprehend than accuracy.

A loss function, also known as a cost function, examines the certainty or probability of a predictions based on how much it ranges from the true value. This produces a more nuanced picture of the model's efficiency.

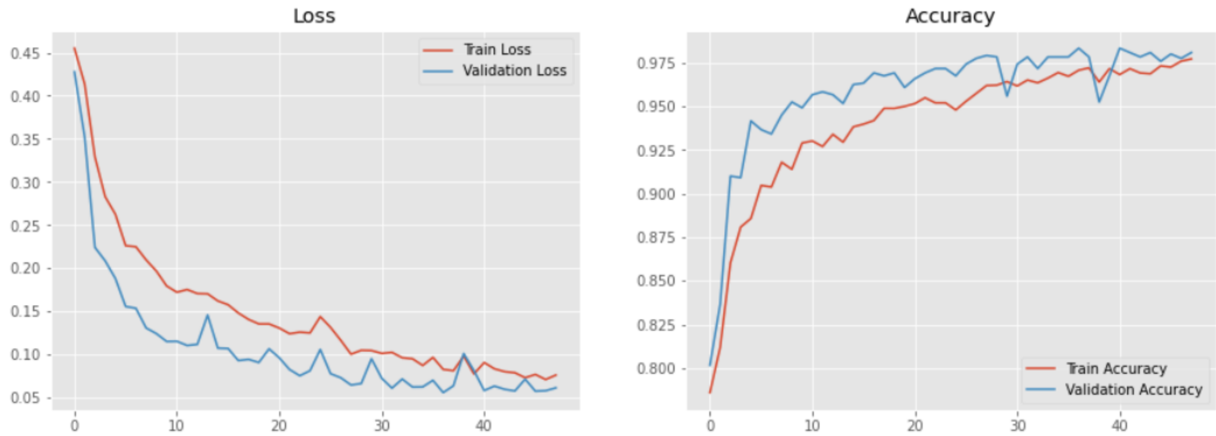

Figure 8: Training and Validation Accuracy and Loss with respect to Epochs

From Figure 8, it can be seen that Train loss and Validation Loss are decreasing steadily after each epoch. Also at the same time Train Accuracy and Validation Accuracy are increasing steadily which means the model fitted well and there is no chance of overfitting and underfitting.

6.3 Confusion Matrix for Validation data and Testing data

A confusion matrix shows how well a classifier performs when given some truth values or cases. True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) are the four components of a confusion matrix. We need to specify some threshold (say alpha)based on IoU to define all the components.

- TP: This is an example of the classifier correctly predicting positive when the truth is positive.
- FP: This is a wrong positive detection
- FN: This is an actual incident that the classifier failed to detect.
- TN: Given that the actual case is also negative, this measure predicts a negative detection. This metric does not apply to object detection because there are numerous possible predictions that should not be spotted in an image. As a result, TN includes all probable inaccurate intercepts that were missed.

Figure 9 depicts the confusion matrix of validation and test data. It can be seen that most of the prediction value are in True Positive and True Negative for both validation and test data which means most of the time, the model is making true prediction rather than false prediction.

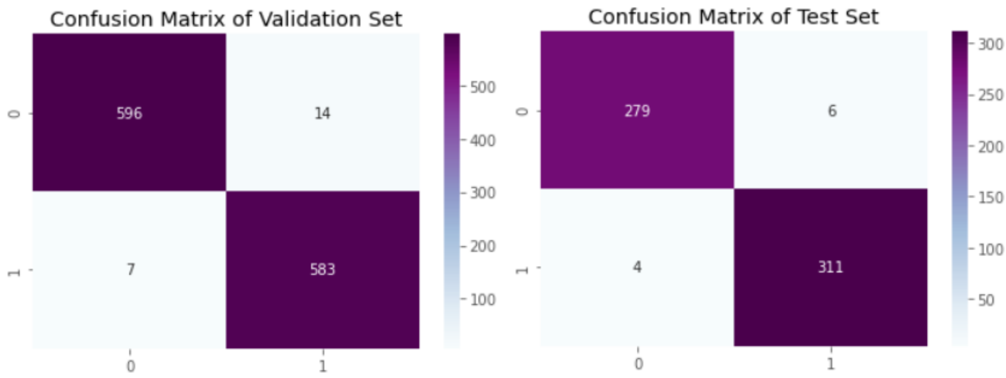

Figure 9: Confusion Matrix for Validation data and Testing data

6.4 Recall,Precision and F1 Score [2]

- Precision refers to a classifier's ability to recognize only relevant things. It's the percentage of positive predictions that are correct.
- Recall is a metric that gauges a classifier's ability to find all relevant examples (that is, all the ground-truths). It's the percentage of true positives found in all ground-truths.
- F1 score is harmonic mean of recall and precision.

Table 2: Recall,Precision and F1 Score

Metrics	Validation	Test
Accuracy	0.9816666666666667	0.9733333333333334
Weighted F1 Score	0.9816674814905351	0.9733107241267752
Weighted Precision Score	0.982017368403507	0.9736361502609786
Weighted Recall Score	0.9816666666666667	0.9733333333333334

From the table, it can be said that model is able to detect and identify ships from the satellite images accurately with few wrong detection and thus the model is well balanced and can be considered well fitted model.

[2] https://towardsdatascience.com/confusion-matrix-and-object-detection

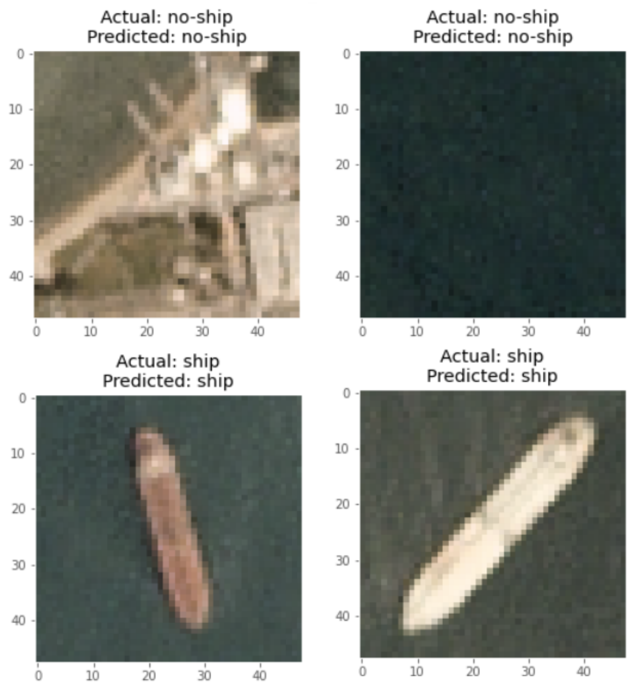

Figure 10: Actual vs Prediction

6.5 Detection of Ships on Random Images

During the testing and validation phase, the model was able to detect ships successfully. Following are some examples of outcomes generated by our algorithm during testing and validation. As seen in Figure 11, this model can handle both small and large ships in the test dataset. It can also recognize many ships in a single image. Ship detection on the validation set is also performing well. In a few photos, the model generates a few false alarms due to irregular surface and waves.

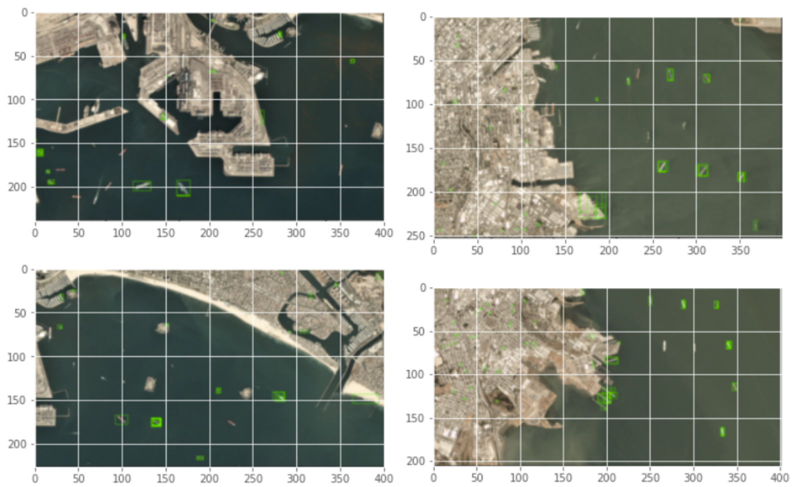

Figure 11: Model testing on random photos

6.6 Discussion

The goal of this work is to evaluate the newly proposed Faster R-CNN employing Feature Engineering based on Data Augmentation and Transformation, which was inspired by the studies in the SAR detection area that were reviewed in the literature review. The investigation begins with the gathering information from satellite photos of ships and their objects. The most relevant study which is related to this research work was conducted by (Wei et al.; 2020) and (Kang et al.; 2017) . As per (Kang et al.; 2017), Faster R-CNN failed to detect some small objects, resulting in a relatively poor detection rate in their study. The major difference between proposed model and the model implemented by (Kang et al.; 2017) is the lack of use of data augmentation and Transformation. Research conducted by (Wei et al.; 2020) achieved an accuracy of 70.85%.A deep learning model's efficacy can be boosted by data augmentation. In the project, data augmentation like Gaussian filter, Edge detection, horizontal flipping, random cropping, scaling and changing the colour of the image are used to have better accuracy. Also, Affine transformation is used to correct the geometric distortions that occur with non-ideal camera angle. The proposed model achieved an validation and Test accuracy of 98.16% and 97.33% respectively with 50 epochs which is much higher than the study conducted by (Wei et al.; 2020). Overall, the model is able to perform well in detecting ships from satellite images and answer the research questions.

7 Conclusion and Future Work

In this paper, Feature Engineering based on Data Augmentation and Transformation is employed to improve the performance of Faster R-CNN in detection of Ship images. Different types of augmentation data like Gaussian filter, Edge detection, horizontal flipping, random cropping, scaling and changing the colour of the image are used along with the Affine transformation. The findings revealed that data augmentation can aid in the improvement of our model's performance. The proposed model achieved an accuracy of 98.16% and 97.33% on validation and test data respectively with a precision score of 0.9820 and 0.9736. The suggested study's findings will aid in the development of a robust marine surveillance system with high detection efficiency.

As a part of future work, more data needs to be used for training and testing purpose and thus high configuration machine is needed. Different types of optimizer needs to be implemented to minimize model loss at training and validation phase. Also, my target is to detect moving ship objects using Deep Learning techniques.

8 Acknowledgment

I would like to acknowledge and give my warmest thanks to my supervisor Dr. Barry Haycock who made this work possible. His guidance and advice carried me through all the stages of writing my project.

References

Ajay, M. N. V. and Raghesh, K. K. (2021). A segmentation based approach for improved detection of ships from sar images using deep learning models, *2021 2nd International Conference on Smart Electronics and Communication (ICOSEC)*, pp. 1380–1385.

Chen, Y., Yu, J. and Xu, Y. (2020). Sar ship target detection for ssdv2 under complex backgrounds, *2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL)*, pp. 560–565.

Chen, Z. and Gao, X. (2018). An improved algorithm for ship target detection in sar images based on faster r-cnn, *2018 Ninth International Conference on Intelligent Control and Information Processing (ICICIP)*, pp. 39–43.

Dugad, S., Puliyadi, V., Palod, H., Johnson, N., Rajput, S. and Johnny, S. (2017). Ship intrusion detection security system using image processing amp; svm, *2017 International Conference on Nascent Technologies in Engineering (ICNTE)*, pp. 1–7.

Huang, J., Jiang, Z., Zhang, H., Cai, B. and Yao, Y. (2017). Region proposal for ship detection based on structured forests edge method, *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 1856–1859.

Kang, M., Leng, X., Lin, Z. and Ji, K. (2017). A modified faster r-cnn based on cfar algorithm for sar ship detection, *2017 International Workshop on Remote Sensing with Intelligent Processing (RSIP)*, pp. 1–4.

Kartal, M. and Duman, O. (2019). Ship detection from optical satellite images with deep learning, *2019 9th International Conference on Recent Advances in Space Technologies (RAST)*, pp. 479–484.

Ke, X., Zhang, X., Zhang, T., Shi, J. and Wei, S. (2021). Sar ship detection based on an improved faster r-cnn using deformable convolution, *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pp. 3565–3568.

Krizhevsky, A., Sutskever, I. and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks, *in* F. Pereira, C. J. C. Burges, L. Bottou and K. Q. Weinberger (eds), *Advances in Neural Information Processing Systems*, Vol. 25, Curran Associates, Inc.

Li, H., Guo, B., Gao, T. and Chen, H. (2019). A transfer learning method of ship identification based on weighted hog features, *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 1302–1305.

Li, L., Zhou, Z., Wang, B., Miao, L. and Zong, H. (2021). A novel cnn-based method for accurate ship detection in hr optical remote sensing images via rotated bounding box, *IEEE Transactions on Geoscience and Remote Sensing* **59**(1): 686–699.

Li, Y., Ding, Z., Zhang, C., Wang, Y. and Chen, J. (2019). Sar ship detection based on resnet and transfer learning, *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 1188–1191.

Mao, Y., Yang, Y., Ma, Z., Li, M., Su, H. and Zhang, J. (2020). Efficient low-cost ship detection for sar imagery based on simplified u-net, *IEEE Access* **8**: 69742–69753.

Morillas, J. R. A., García, I. C. and Zölzer, U. (2015). Ship detection based on svm using color and texture features, *2015 IEEE International Conference on Intelligent Computer Communication and Processing (ICCP)*, pp. 343–350.

Wang, N., Li, B., Wei, X., Wang, Y. and Yan, H. (2021). Ship detection in spaceborne infrared image based on lightweight cnn and multisource feature cascade decision, *IEEE Transactions on Geoscience and Remote Sensing* **59**(5): 4324–4339.

Wang, T. and Gu, Y. (2018). Cnn based renormalization method for ship detection in vhr remote sensing images, *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, pp. 1252–1255.

Wei, S., Chen, H., Zhu, X. and Zhang, H. (2020). Ship detection in remote sensing image based on faster r-cnn with dilated convolution, *2020 39th Chinese Control Conference (CCC)*, pp. 7148–7153.

You, Y., Cao, J., Zhang, Y., Liu, F. and Zhou, W. (2019). Nearshore ship detection on high-resolution remote sensing image via scene-mask r-cnn, *IEEE Access* **7**: 128431–128444.

Zhang, Y., Lei, Z., Zhuang, L. and Yu, H. (2021). A cnn based method to solve class imbalance problem in sar image ship target recognition, *2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, Vol. 5, pp. 229–233.

Zong-ling, L., Lu-yuan, W., Ji-yang, Y., Bo-wen, C., Liang, H., Shuai, J., Zhen, L. and Jian-feng, Y. (2019). Remote sensing ship target detection and recognition system based on machine learning, *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 1272–1275.