Instituto Superior de Engenharia de Lisboa

MEET

Computationally efficient vessel classification using shallow neural networks on SAR data

November 2023

Supervisor: Paulo Marques, paulo.marques@isel.pt
Student: Álvaro Campos Nº45937, a45937@alunos.isel.pt

Jury:
President: Alessandro Fantoni
Vowels: Paulo Alexandre Carapinha Marques
Andrea Radius

Instituto Superior de Engenharia de Lisboa

MEET

Computationally efficient vessel classification using shallow neural networks on SAR data

November 2023

Supervisor: Paulo Marques, paulo.marques@isel.pt
Student: Álvaro Campos Nº45937, a45937@alunos.isel.pt

Jury:
President: Alessandro Fantoni
Vowels: Paulo Alexandre Carapinha Marques
Andrea Radius

Agradecimentos

Gostaria de expressar a minha sincera gratidão a todas as pessoas que contribuíram para o sucesso da minha tese de mestrado.

Em primeiro lugar, quero agradecer aos meu orientador, Professor Paulo Marques, pela orientação valiosa, apoio constante e pelo incentivo incansável ao longo deste processo. Agradecer pelas suas ideias e discussões que foram fundamentais para o desenvolvimento deste trabalho.

Quero agradecer tambem aos meus amigos por me terem apoiado nos momentos em que foi necessário.

Finalmente, quero agradecer também à minha família e namorada pelo apoio incondicional ao longo dos anos, pelo incentivo constante e pela compreensão nos momentos em que a minha atenção estava inteiramente voltada para esta tese. Obrigado também por estarem ao meu lado em todos os momentos, dando-me forças para continuar em frente, independentemente das dificuldades.

Cada um de vocês desempenhou um papel fundamental nesta jornada e sou profundamente grato por isso.

Muito obrigado.

Abstract

Synthetic aperture radar (SAR) is an active radar that is mounted on a moving platform, simulating a longer antenna length than the physical antenna real length. Similar to a conventional radar, electromagnetic waves are sequentially transmitted and the backscattered echoes are collected by the radar. With the proper signal processing, this kind of system is able to provide high resolution microwave images of a desired target area by synthesising a larger antenna aperture, in virtually all-weather conditions. Nowadays SAR systems have been extensively used for remote sensing. It has various applications such as Earth surface monitoring, charting and military applications. Since it is weather independent and is able to operate whether it is day or night, SAR can be a more reliable source when compared with optical imagery [1].

Ship detection and recognition in SAR images has become an important topic in research in recent years. This thesis presents a computationally efficient algorithm for the classification of vessels in SAR images using Neural Networks (NN) with a reduced number of hidden layers, also called Shallow Neural Networks (SNN). Herein the use of SNN for vessel classification will be divided into two main steps: feature extraction and classification. Feature extraction aims to lessen the burden deep neural networks cause on computational resources by extracting key features beforehand from the SAR image. The low computational requirements make this implementation compatible with onboard vessel systems and real time applications. The classification is implemented using a SNN that uses parameters obtained from feature extraction algorithms to classify the vessel present in the radar image.

In this thesis feature extraction processes data from the Open SAR Ship dataset [2] in order to obtain the vessel's various features, such as ship length, width, mean, standard deviation and the number of scatter points present on the vessel.

Keywords: Vessel classification, Synthetic Aperture Radar, neural networks, SAR image processing.

Resumo

O radar de abertura sintética (SAR) é um radar ativo montado em uma plataforma em movimento, que simula um comprimento de antena maior do que o comprimento real da antena física. De forma semelhante ao radar convencional, ondas eletromagnéticas são transmitidas sequencialmente e os ecos são coletados pelo radar. Com o devido processamento de sinal, este tipo de sistema é capaz de fornecer imagens de micro-ondas de alta resolução de uma área-alvo desejada, em praticamente todas as condições meteorológicas. Atualmente, os sistemas SAR têm sido amplamente utilizados para a detecção remota possuindo várias aplicações, como observação da superfície terrestre, cartografia e aplicações militares. Dado que é independente do clima e pode operar tanto de dia quanto de noite, o SAR pode ser uma fonte mais confiável quando comparado com imagens ópticas [1].

A deteção e reconhecimento de navios em imagens SAR tornou-se um tópico importante de pesquisa nos últimos anos. Esta tese apresenta um algoritmo computacionalmente eficiente para a classificação de embarcações em imagens de SAR usando Redes Neuronais com um número reduzido de camadas, também conhecidas como *shallow neural networks*. A utilização de *shallow networks* para a classificação de embarcações será dividida em duas etapas: extração de características e classificação. A extração de características tem como objetivo reduzir a carga computacional que as *deep neural networks* causam nos recursos computacionais, extraindo antecipadamente características-chave da imagem SAR. Os baixos requisitos computacionais tornam esta implementação compatível com sistemas a bordo de navios e aplicações em tempo real. A classificação é realizada usando uma rede neural com um número reduzido de camadas, que utiliza parâmetros obtidos a partir de algoritmos de extração de características para classificar a embarcação presente na imagem de radar.

O processo de extração de características processa dados do conjunto de dados Open SAR ship [2] para obter várias características da embarcação, como comprimento, largura, média, desvio padrão e o número de pontos de dispersão presentes na embarcação.

Index

List of Figures

List of Acronyms

AI Artificial Intelligence

AE AutoEncoder

AIS Automatic Identification System

ATR Automatic Target Recognition

CA-CFAR Cell Averaging Constant False Alarm Ratio

CFAR Constant False Alarm Ratio

CNN Convolutional Neural Networks

DBN Deep Belief Networks

DNN Deep Neural Networks

EW Extra Wide Swath

FSI Fine Strip-Map 1

FSII Fine Strip-Map 2

GAN Generative Adversarial Networks

GRD Ground Range Detected

GT Global Threshold

IW Interferometric Wide swath

LSTN Long Short Term Neural networks

ML Machine Learning

NN Neural Networks

OCN Ocean products

PCA Principal Component Analysis

PDGS Payload Data Ground Segment

PFA Probability of False Alarm

PDF Probability Density Function

QPSI Full Polarization 1

QPSII Full Polarization 2

RFI Radio Frequency Interferences

RNN Recurrent Neural Network

SAE Stacked AutoEncoder

SAR Synthetic Aperture Radar

SCG Scaled Conjugate Gradient

SLC Single Look Complex

SM Strip Map

SNN Shallow Neural Networks

SVM Support Vector Machine

UFS Ultrafine Strip-Map

WV Wave mode

1 Introduction

Ship target interpretation in SAR images has become an important research topic in recent years. With the improvement of SAR image resolution, the performance of traditional automatic target recognition (ATR) methods increases gradually. This thesis aims to implement a computationally efficient technique for the classification of vessels SAR images using Artificial Intelligence (AI). The major handicap of AI based techniques is the huge computational requirements. Herein we will perform feature extraction before proceeding to classification using NN. Feature extraction aims to lessen the burden deep neural networks cause on computational resources by extracting key features beforehand, making the implementation compatible with onboard vessel systems. The classification of vessels is implemented using a shallow neural network that uses parameters obtained from feature extraction to classify the vessels present in the radar image.

In recent times classifying problems are all resolved with the help of Deep Neural Networks (DNN). The problem with this approach is that DNNs demand high computational power in order to have results in a useful time. Some problems don't need an overly powerful DNN, they can be resolved with a shallow neural network and pre-processing techniques that acquire the best of the given data.

In this thesis, feature extraction uses data from the Open SAR Ship dataset [2] as a base data set for the research. The main features extracted for the classification of ships in SAR images are the length, width, mean, standard deviation and bimodality coefficient of the ship. The first step in feature extraction is identifying the pixels from the ship, in this thesis, two distinct methods for this goal were created. The Global threshold algorithm uses, as the name suggests, a threshold to identify if a pixel is from a ship. The other uses a Cell Averaging Constant False Alarm Ratio (CA-CFAR) [3] algorithm to identify these pixels. Having the pixels from the vessel, the SAR image can be binarized and the length and width may be extracted from the image.

To extract the length and width of the ship two approaches were taken into account, a more general approach supported on the pre-existing *regionprops* function in MATLAB and a more specific approach based on the center and angle of the ship to create lines in both the ship directional axis (length) and the perpendicular axis (width) to acquire the dimensions of the vessel.

After acquiring the length and width, the next parameters are the sta-

tistical values obtained from processing the ship pixels. These parameters are the mean, standard deviation and the bimodality coefficient [4]. The bimodality coefficient is a very important parameter because it can be a major factor in deciding what type of ship it is since it can identify if the data has two distinct peaks or clusters of values, suggesting the presence of two separate distributions within the data. It can unveil if the vessel has one or more scatter points present.

With the parameters obtained beforehand, a SNN is now trained to classify three different types of vessels: Cargo, Tanker and Fishing. Using the pre-processed parameters the shallow network is able to classify these three types of ships with high accuracy.

This thesis will be divided into 5 sections, starting with the state of the art, the existing technology on the topic of vessel detection and classification is discussed and presented in order to create a background on the project. This section is divided into three subsections Conventional Strategies, AI based application and Public SAR Data sets. The proposed approach section is where the algorithm developed will be explained and tested as well as an introduction on Sentinel-1 products. The results section explains the conducted tests and evaluates the overall strategy. Finally, the conclusion and future work sections are presented.

The developed algorithms, datasets and papers can be found on this GIT Repository

2 State of the art

Ship detection and identification has received unprecedented attention due to its significant application value such as fishery management, ship rescue, maritime traffic control, and battlefield awareness [5]. Synthetic aperture radar is an advanced active microwave sensor that can simultaneously achieve high resolution and wide swath [1]. Due to the prominent capacity of all-day and all-weather imaging, it has been extensively used for marine surveillance such as ship detection and identification among other applications. However, there still remain some difficulties to resolve for SAR ship detection, such as near coast ships due to imperfect land masks. Moreover, it is quite difficult to satisfy the requirements for both high-speed and high-accuracy detection in SAR images.

SAR images are different from optical images, and the microwave imaging mechanism is complex. The recognition of SAR image objects is difficult and requires strong domain knowledge. SAR ocean images are heterogeneous and contain ships, upwelling, breaking waves, and many artifacts such as radio frequency interferences (RFIs) and azimuth ambiguities [6]. Quite often the texture and grey information between the ships and the false alarms are indistinguishable. Moreover, the background of the SAR inshore image is more complex due to the impact of multiple interferences from land, which also greatly increases the difficulty of detection.

2.1 Conventional strategies

A lot of research has been devoted to ship detection and identification in SAR images for decades [1]. Traditional methods could be roughly divided into three categories: threshold methods, statistical methods, and transformation methods [7]. They generally use a-priori knowledge to extract features manually through a series of candidate regions. Among these traditional methods, the constant false alarm rate (CFAR) [8] and its improved versions based on threshold are frequently used. As for CFAR detectors, an appropriate statistical model is selected to match the probability density function (PDF) of the ocean clutter, and then an adaptive threshold is calculated with a typical probability of false alarm (PFA). The key to CFAR ship detection is the statistical model of the ocean clutter.

In general, the most commonly used model in real applications are mainly based on Gauss, K, Rayleigh, and Weibull distributions [9]. The most pop-

ular CFAR method is the cell-averaging CFAR with the Gaussian distribution [9]. However land, islands, or some artefacts such as RFI and azimuth ambiguities will affect the final performance of CFAR detectors. These interferences will lead to a mass of false alarms even if the statistical model accurately matches the real sea state. In order to deal with these problems a land mask can be used to reduce the number of false alarms in the SAR image.

In terms of vessel identification, the present methods used for classification in SAR automatic target recognition (ATR) systems consist of Artificial Intelligence (AI) [9], machine learning (ML) [9] techniques and other methods that do not use ML in the process but the results can be applied to ML [9]. These are template-based methods [10], Statistical methods [11], Sparse representation-based methods [12] and phase-based methods [12]. These methods alone can create state-of-the-art results but when combined they can reach even higher levels of utility in RS and SAR target identification. Template based methods involve using a reference image (template) of a known target to identify similar objects in a SAR image. The process consists of a thoughtful cross-correlation between the template and the image, which pinpoints the position with the highest correlation value. This method is efficient and can detect with great success similar shapes, but it can fail in detecting targets with different orientations or with variations in scale [13].

Statistical methods take place by using mathematical models to extract features from SAR images for target identification. One of the most used methods is Principal Component Analysis (PCA) [7]. Its goal is to lower the size of high dimensional SAR images by finding the principal components that compose the data, in other words, to identify which are the directions in the feature space that explain the most amount of variance in the data. With this PCA can objectively retain only the most important information from the SAR image and this information can be used later on ATR ML methods [11].

Sparse representation-based methods involve representing the SAR image as a sparse linear combination of basis functions and identifying or classifying the object by its sparse representation of the image [7]. Despite this classification method proving its performance in pattern recognition its disadvantage of time consumption can be a challenge, especially in the case of features with huge dimensions. In order to overcome this challenge it is possible to implement a feature vector with reduced dimension making the algorithm run smoothly and with better time consuming perspectives [14].

Phase-based methods take advantage of the phase information in SAR images to perform target recognition. In terms of detection and classification, this method does not bring great results but in the RS field of Co-registration which consists of the adjustment of the position and orientation of the images so that they are precisely overlapped with each other this method makes great advancements. The last mentioned method allows the improvement of the images and can enhance the detection and classification of targets [12].

2.2 AI based application

In recent years the most used techniques for pattern extraction and insights involve machine learning [7], and its use has been growing in remote sensing data exploration. A vast number of machine learning algorithms have been applied to remote sensing problems, which can be attributed to two basic tasks, classification and regression. Machine learning can be explained as a branch of AI that focuses on the development of algorithms and statistical models that enable computers to learn from data. The objective of machine learning is to enable computers to make predictions and identify patterns, without the need for explicit programming. In order to use this process the model needs to be trained using a large data set that represents the problem in question, the model is then tested in making decisions and predictions in new, unseen situations of the same kind. With the training, the model can improve its accuracy over time through the process of trial and error and also from data feedback. There are three main types of machine learning: supervised learning, unsupervised learning and reinforcement learning [15]. In supervised learning, the model is trained on a labelled data set where the correct output is known for each input. Unsupervised learning is the process in which the model is not given any labelled data and must find patterns and relationships within the data on its own. In reinforcement learning, the model receives feedback in the form of rewards or penalties as it interacts with its environment, and must learn from this feedback to make better decisions over time.

Ever since the start of the Deep Learning (DL) era (a branch of ML) the research on basic machine learning techniques has decreased, in any case, there are some methods that retain their importance such as support vector machine (SVM) that maintained its relevance due to its solid theoretical basis and easy expandability to tackle open questions in ML [16] and random forest [17] given its simple operation and solid results.

DL is undoubtedly the most popular method in ML. DL is a specialised form of machine learning that uses neural networks with multiple hidden layers between the input and output, normally having more than three hidden layers, thus being called deep neural networks [18]. Machine learning starts by manually extracting important features from images then these are joined together to build a model that efficiently translates the objects in the image. On the other hand, Deep learning automatically extracts important features from images. More, deep learning does "end-to-end learning" meaning that a network is given raw data and a task to perform, it learns the task and how to perform it and with good training, it can accomplish outstanding results. Another key difference is that deep learning algorithms scale with data. On the other side shallow learning converges with data. Shallow learning refers to ML methods that may decrease the level of performance as more examples are inputted on the network. Finally, with Deep learning, there is no need to select features manually and to add classifiers to sort images since the feature extraction and modelling steps are automatic.

DL in SAR can be divided into three main approaches, sequence-based, generative, and convolutional. With the sequence-based approaches comes the Recurrent Neural Networks (RNN) [19] and Long Short Term Neural (LSTN) [15] networks this approach is particularly useful in target tracking. These neural networks were originally used for modelling the long-range dependencies of sequential data and may be also applied to images and data with large dimensions (data with spectrum and time characteristics like Single Look Complex SAR images). In recent times RNNs have been used to learn features of pixel classification. In terms of SAR utility [15] the RNN are used as an inverse solver for image formation problems in passive SAR [20], for agricultural land cover mapping [21]. Furthermore, spatial-sequential RNN [22] have been developed to include spatial information [7]. LSTM networks are a unique type of RNN capable of learning long-term dependencies. Their most attractive trait is the capability to remember information for longer periods of time [15].

The Generative models are one of the most important branches of ML [7]. Even with its importance, generative models have received less attention in the beginning of Deep learning era since Convolutional Neural Networks (CNNs) [7] have achieved a surprising performance in the remote sensing area. One of the principal Generative models uses is to generate new data that are similar to a specific train data set, thus diminishing the latent need for a great number of samples to train any deep network. There are various

generative models starting with Deep Belief Networks (DBNs) [7] these networks have great usability in terms of pixel classification and object detection does not reach state of the art operations systems because of the following reasons: first DBNs are a fully connected network structure which makes it not as effective as locally connected CNNs, second in DBNs all hidden units must be Boolean in the pre-training creating a limitation on the capacity to tune the optimization. The autoencoder (AE) [7] is a symmetrical network structure trained to copy its input to its output, normally used for unsupervised learning from unlabeled data. In brief, the AE encoder aims to compress data into a reduced dimension while the encoder has the opposite task, to reconstruct the data back to its original state. A stacked autoencoder (SAE) [7] is a deep network that makes use of multiple AEs, where the output of one becomes the input of the next in line. This network finds its finest use in pixel classification and denoising in hyperspectral remote sensing images [23] and in SAR in the domain of polarimetric image classification [15]. Generative adversarial networks (GANs) [15] have become a spotlight in recent research topics. Theoretically, GANs consist of a competing generator and discriminator, the usefulness of GANs in SAR is in the creation of synthetic images of SAR training tokens thus increasing the amount of training data available to further increase the efficiency in target detection and classification [24]. Even as GAN open new opportunities in SAR and RS community for the above-mentioned advantages, GAN generated images are not real and, at times, even unreliable, the misuse of these neural networks may instead of improving the behaviour, cause it to crash and worsen the performance of the network.

Convolutional Neural Networks are the most frequently used in state of the art developments. CNNs have drawn considerable attention for the analysis of SAR data in the fields of automatic target recognition, land cover classification, segmentation, change detection, object detection, and image denoising [15]. The basic structure of a CNN can be described by multiple basic feature extraction stages.

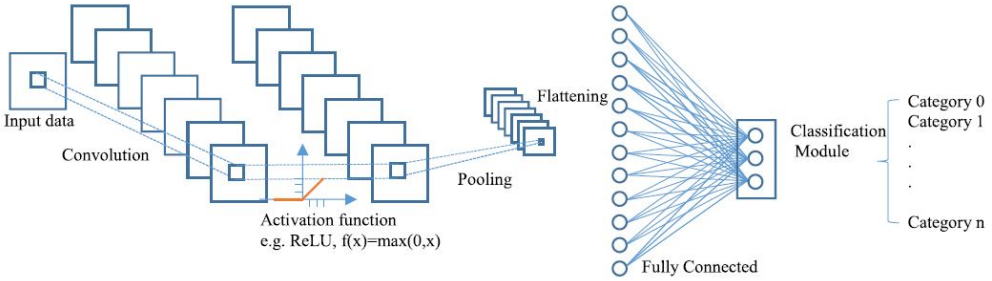


Figure 1: Convolutional neural network, taken from [15].

The main elements of each stage (Figure 1) are the convolutional layer, the activation function and the pooling layer, followed by one or more fully connected layers and a classification module. The convolutional layer is the core part in every stage of CNN. It is a filter bank, where they are applied to extract features from the image. The filters are learned through the training process. The activation function is a simple mathematical function which is applied to the output of the previous layer in order to introduce nonlinearity to the network, therefore, allowing the network to model complex relationships in data. A pooling layer, the objective of this layer is to perform a down-sampling operation on feature maps, this way reducing the size of every map while keeping the most important information. Pooling is normally performed by a sliding window over each map, taking only one value from each window. This process is done to make the representation close to invariant to small alterations in the input. After going through every layer, the output of the last is flattened to a vector and fed to a fully connected layer. This final layer's objective is to learn nonlinear combinations of previous features to classify the samples. One or more fully connected layers can be used. The classification layer, the finishing module of the CNN is used to transform the output of the last fully connected layer to the probability domain. There are various classifiers that are widely used like the softmax classifier for multi class classification [15].

Machine learning is a very powerful technique that provides surprisingly good results on ATR like identifying and classifying vessels in SAR images. However, these algorithms have close to no utility if not properly trained. To efficiently train neural networks a high-quality and well-labelled data set is of the utmost importance. In vessel detection and classification, the most used in the recent state of the art applications involving the European

Space Agency SAR Sentinel 1 satellite is the OpenSARShip data set [2], this data set includes 41 Sentinel-1 images with various environmental conditions, providing 11 346 SAR ship tokens integrated with automatic identification system (AIS) messages. The OpenSARShip owes five essential properties: specificality, large scale, diversity, reliability, and public availability.

To improve the behaviour of machine learning and SAR imagery interpretation, elaborate geometric parameter estimation is essential [25]. Hence, geometric parameter extraction with high accuracy based on Sentinel-1 products is a very crucial and meaningful task for marine surveillance. Size extraction with very high accuracy needs to overcome undesirable effects among the aforementioned factors. Aiming to get accurate size parameters, an extraction procedure generally includes three main stages: 1) binarization; 2) image operations for accurate size extraction; and 3) elaborate geometric parameter estimation. The first stage refers to the processing of dividing pixels into two classes: target and nontarget. Studies developed various detection methods [26] or fused different detection methods [16] to improve the detection performance and the size extraction performance simultaneously. The second and the third stages move forward based on the former results to get the estimation with high accuracy generally from two aspects: image operation and statistical refinement.

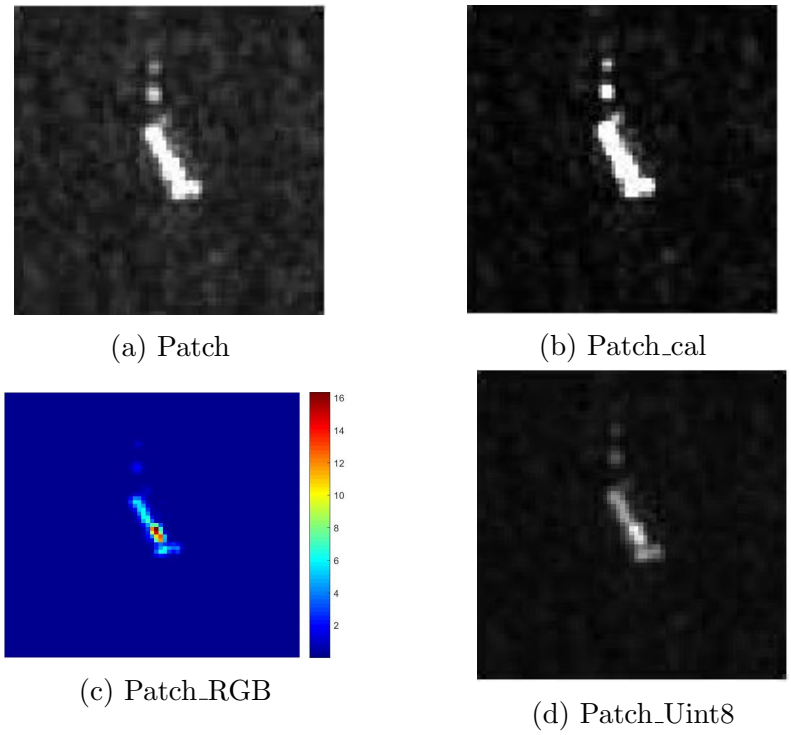
2.3 Public SAR Data Sets

To start the implementation of the project, a thoughtful revision of the OpenSarShip data set [2] was required. The images collected for this data set are all acquired from the sentinel-1 European Space Agency SAR satellite. It contains 41 images with various environmental conditions, providing a round-up of 11000 SAR ship tokens [2] integrated with AIS messages. The data is divided into two types of SAR images Single Look Complex (SLC) and Ground Range Detected (GRD). The difference between these acquisition types is that SLC images are close to the rawest form of SAR images, containing complex-valued data that has not been processed for ground-range correction or multi-looking. They have a high resolution and are useful for advanced processing [27], on the other hand, GRD images have been processed to remove geometric distortions caused by the SAR instrument and to reduce speckle noise. They are usually multi-looked, which means that multiple SLC images are averaged together to improve the signal-to-noise ratio and reduce the appearance of speckle. The result is a smoothed image

that is easier to interpret visually [27].

For every Sentinel-1 SAR image, four formats of ship tokens are provided each on a subfolder, namely:

- Patch - Contains the original data (Figure 2a).
- Patch_cal - Contains the calibrated data (Figure 2b).
- Patch_RGB - Contains the data in pseudo-colour (heat map) for the intensity of each pixel (Figure 2c).
- Patch_Uint8 - Containing the same data visualised in greyscale (Figure 2d).



(a) Patch

(b) Patch_cal

(c) Patch_RGB

(d) Patch_Uint8

Figure 2: Types of image contained in Open SAR Ship data set.

The Sentinel-1 SAR image folder also contains a file where all the AIS data from every vessel is stored, i.e. the latitude, longitude, type, name,

length and width, and so on [28]. These characteristics will be used later in the preliminary tests.

With the aim of better understanding the data set, an algorithm has been implemented whose purpose is to count the number of ship tokens for each type, as shown in Table 1.

Table 1: Number of ship tokens present in OpenSarShip dataset GRD section.

Vessel Type	Cargo	Tanker	Dredging	Fishing	Port tender	Passenger	Search	Tug	Other
Ship tokens	6568	1847	122	120	8	60	15	96	28

The sum of all ship tokens does not add up to 11000 because the results in Table 1 are only for the GRD SAR images.

The SAR ship dataset is a cornerstone in the domain of Ship datasets. In addition to this dataset, a SAR Dataset for Ship Detection under Complex Backgrounds [29] can complement and provide more vessel tokens with complex backgrounds. This labelled dataset was created using 102 Gaofen-3 images and 108 Sentinel-1 images. It consists of 43,819 ship chips of 256 pixels in both range and azimuth.

Gaofen-3 provides images at resolutions ranging from 3 meters to 10 meters, employing the Ultrafine Strip-Map (UFS), Fine Strip-Map 1 (FSI), Full Polarization 1 (QPSI), Full Polarization 2 (QPSII), and Fine Strip-Map 2 (FSII) imaging modes. In the Sentinel-1 SAR images, the imaging modes include Strip-Map (SM) and IW-mode. These characteristics are shown in Table 2.

Table 2: Detailed information for original synthetic aperture radar (SAR) imagery, taken from [29]

Value	**Imaging Mode**	**Resolution Rg. x Az (m)**	**Swath (km)**	**Incidence Angle (º)**	**Polarization**	**Number of Images**
GF-3	UFS	3 x 3	30	20∼50	Single	12
GF-3	FS1	5 x 5	50	19∼50	Dual	10
GF-3	QPSI	8 x 8	30	20∼41	Full	5
GF-3	FSII	10 x 10	100	19∼50	Dual	15
GF-3	QPSII	25 x 25	40	20∼38	Full	5
Sentinel-1	SM	1.7 x 4.3 to 3.6 x 4.9	80	20∼40	Dual	49
Sentinel-1	IW	20 x 22	250	29∼46	Dual	10

3 Proposed approach

This chapter presents an introduction to sentinel-1 products and explains how every component of the project was implemented, starting with SAR image characteristics, preliminary tests on Matlab machine learning tools, an explanation of the binarization algorithms used on the project, length and width acquisition, mathematically calculated parameters like standard deviation, mean and number of scattering points as well as an explanation of the training on shallow neural network used for the project. The diagram of the project is shown in Figure 3.

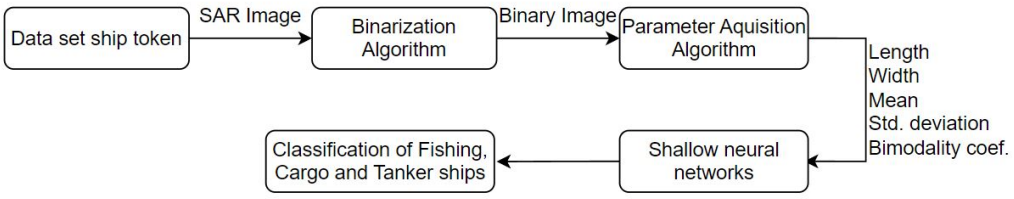


Figure 3: Block diagram of the project.

3.1 SAR image characteristics

All the images that were used in the development of this project originate from the ESA's Sentinel-1 SAR satellite constellation. Sentinel-1 is a constellation of two C-Band SAR identical satellites (centred at 5.495 GHz). These satellites are used for the monitoring of sea, land, ice, surveillance of oil spills and ships, monitoring of marine wind and waves, monitoring of land surface motion risks, and the mapping of forest and water. It is important to mention that there are other types of SAR that work with different frequency bands like X-band. The major difference between C-band and X-band radars is their operating frequency. The first operates at 4 to 8 GHz and the latter at 8 to 12.5 GHz, giving them different characteristics in terms of propagation, penetration, sensitivity and product resolution [30]. C-band has lower product resolution but has better propagation and cloud penetration. On the other hand, X-band can provide a better resolution but in tumultuous weather conditions it may be impossible to surveil the desired location. For

the purpose of this thesis, the images from the Sentinel-1 Open SAR Ship data set [2] are used.

The Sentinel-1 satellites provide four types of acquisition modes [30]:

- Stripmap mode (SM): 80 km swath with a spatial resolution of 5 m x 5 m.
- Interferometric Wide swath mode (IW): 250 km swath, 5 m x 20 m spatial resolution and burst synchronization for interferometry.
- Extra Wide Swath mode (EW): 400 km swath and 25 m x 100 m spatial resolution (3-looks). Six overlapping swathes have to be foreseen to cover the required access range of 375 km.
- Wave mode (WV): low data rate and 5 m x 20 m spatial resolution. Sampled images of 20 km x 20 km at 100 km intervals along the orbit. The Wave mode at VV polarization is the default mode for acquiring data over open ocean.

The acquisition modes can be found in Horizontal-Horizontal and Horizontal-Vertical (HH+HV) polarizations or vertical-vertical and vertical-horizontal (VV+VH) polarizations. For this project the VV polarisation was significantly more used than VH polarisation because in most of the images, the irregularities of the SAR images are felt more strongly in VH polarisation and the side lobe effects are less effective in VV polarization. The acquired radar images have different product types and processing levels. There are three types of sentinel-1 images [27]:

- Single Look Complex (SLC) - Are images in the slant range by azimuth imaging plane, in the image plane of satellite data acquisition. Each image pixel is represented by a complex (I and Q) magnitude value.
- Ground Range Detected - Consists of focused SAR data that has been detected, multi-looked and projected to ground range using an Earth ellipsoid model. Each image pixel is represented as the detected magnitude.
- Ocean - Ocean products consist of several measurements that are focused on observing and studying oceanic and coastal environments.

SAR processing levels refer to different stages of data processing and calibration that transform raw SAR data, into different usable products with increasing levels of refinement, accuracy and information content. Sentinel-1 products are available in three different levels [27]:

- level 0 - products require SAR processing as performed by the Payload Data Ground Segment (PDGS). They can be useful for scientists testing SAR data processing.
- level 1 - These products are created by applying algorithms and calibration to form a baseline engineering product, from which higher levels are derived.
- level 2- consists of geo-located geophysical products derived from Level-1. Level-2 Ocean (OCN) products for wind, wave and currents applications.

Sentinel-1 images using IW acquisition mode from level-1 GRD come in medium and high resolutions and they have the characteristics listed on 3.

Table 3: Parameters for sentinel-1 IW Products [27]

Product	Resolution rg x az (m)	Looks rg x az	Pixel spacing rg x az (m)
GRD Medium resolution	88 x 84	22 x 5	40 x 40
GRD High resolution	20 x 22	5 x 1	10 x 10

The pixel resolution and pixel spacing will be critical further on at 3.2. The difference between these two parameters is resolution represents the distance in which it is possible to distinguish two different scatters and pixel spacing is the distance between pixels on the SAR image.

3.1.1 Preliminary tests

As mentioned before, the data set contains the AIS information on every ship token and in the preliminary test this data will be used to identify some of the challenges the project will face.

For this test, the length and width of the vessels were used to train a neural network with three layers as shown in Figure **??**. To train the network we need the input and output vectors. To create the input the algorithm runs through all the folders and acquires the length, width and type of every

vessel, from the AIS data from every ship, next with the type of vessel the output is created by introducing an array of 1's and 0's according to the ship type. The preliminary test uses only cargo (index 1), tanker (index 2) and fishing (index 3) vessel tokens because there are only two inputs and having too many outputs would increase the confusion on the neural network.

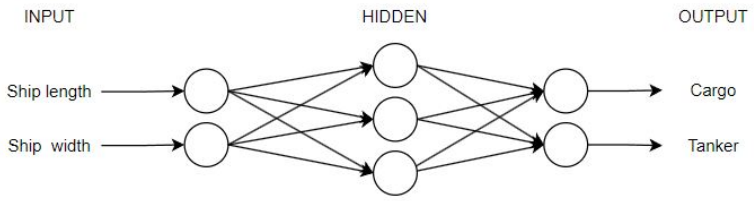


Figure 4: Confusion matrix.

The network is trained using cross-validation with a ratio of 70% of the data to training, 15% to verification and the other 15% to testing. The results are shown in Figure 5 these values are typical for the training of neural networks.

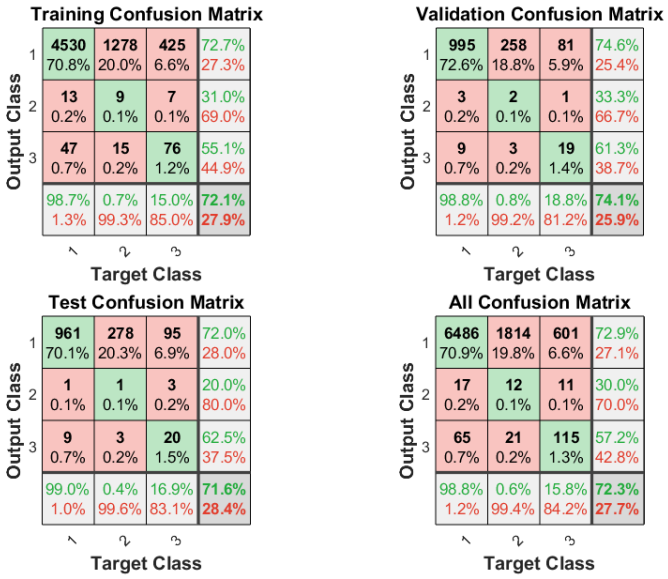


Figure 5: Confusion matrix.

By observing Figure 5 it is possible to conclude that using only the length and width of the vessel, won't be enough for the neural network to differentiate between cargo and tanker vessels. The aforementioned results happen because cargo and tankers have similar sizes. The other reason is the data set has 6486 cargo ship tokens and only 1814 tanker tokens this disproportion causes the network to fail to identify the tanker ship since it has too many cargo ship tokens leading to a biased decision. The fishing vessels are easily identified since their size differs from tanker and cargo ships but still do not obtain the desired accuracy with only 57% of the correct classifications.

In order to solve the discrepancy between Cargo, Tanker and fishing vessel tokens a downsample of the data set was done. The number of ship tokens was defined by the type with lesser training data, i.e., the fishing vessel that has a total of 120 tokens. The test was conducted again with the new parameters and the result is shown in 6

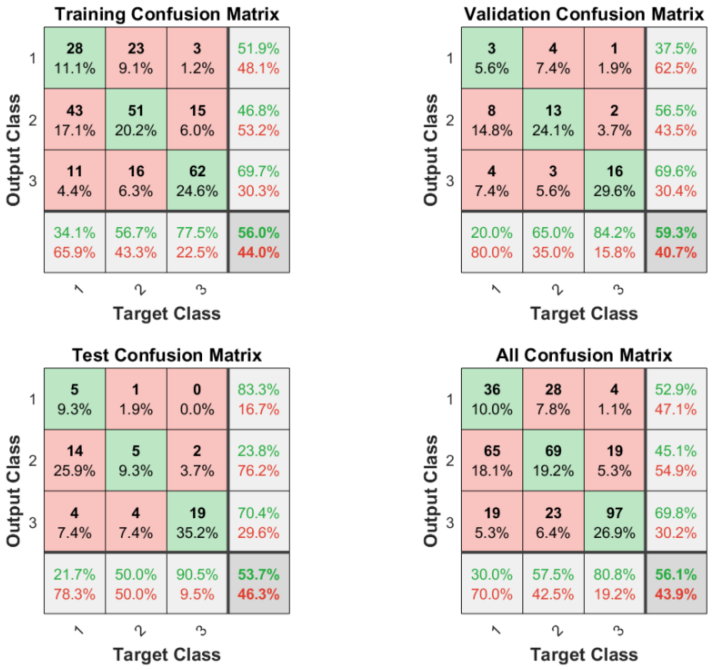


Figure 6: Confusion matrix.

The results present in Figure 6 are much better than the ones previously shown, making the downsampling of the data an indispensable step for less

biased learning and better results.

3.2 SAR image binarization

A binarization algorithm is a technique employed in image processing to transform an image into a binary representation of itself, where every pixel is shown with only two values: 1 and 0 (black and white).

Dividing an image into sections of interest based on the intensity is the main purpose of binarization. Binarization algorithms can be divided into three categories Threshold selection, Pixel Classification and Output image creation

- **Threshold Selection**: The first step is to choose a threshold value that will be used in order to classify pixels as either pixels of interest or background. The threshold value can be selected based on various methods, such as statistical analysis of pixel intensities, global or local histogram analysis or by simple trial and error.
- **Pixel Classification**: Each pixel's intensity value is compared to the chosen threshold. If the pixel value is greater than or equal to the threshold, it is classified as a pixel of interest; otherwise, it is classified as background.
- **Output Image Creation**: Based on the pixel classified as of interest, a new binary image is generated, where the detected pixels values are replaced by "1" (white) and the background pixels by 0 (black).

During the development of the project, two different approaches were implemented to binarize the SAR images: the standard binarization using a threshold composed of the minimum and maximum value of the image pixels which will be mentioned by the Global threshold (GT) approach and a second algorithm that uses the Constant False Alarm Ratio to detect the pixels from the vessel.

3.2.1 Global threshold

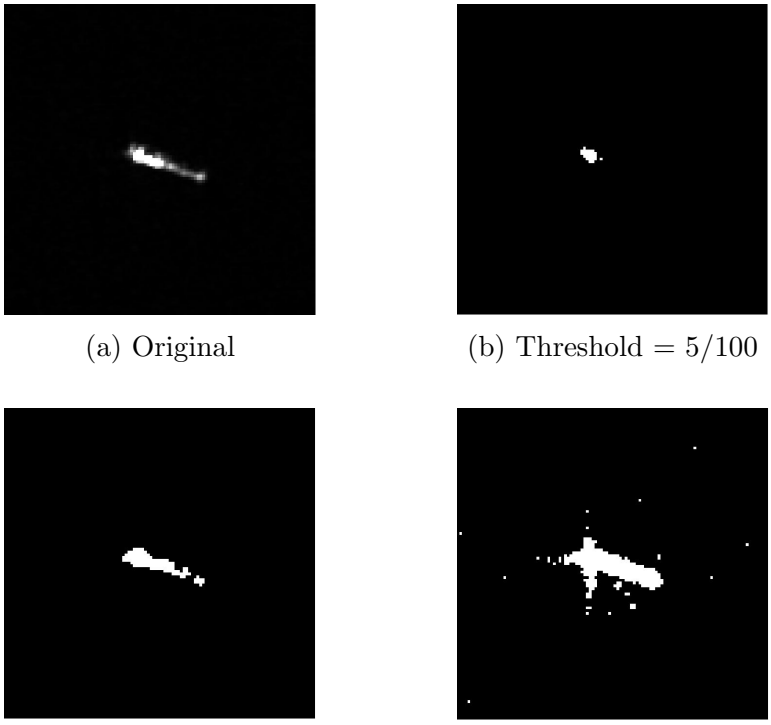
As previously mentioned, the Global threshold binarization algorithm consists of using equation 1 that compares the maximum (I_{max}) and minimum (I_{min}) pixel values within the image multiplied by a threshold ($Tresh$) and normalized into the Imin value.

$$I_{Th} = Tresh(I_{max} - I_{min}) - I_{min} \tag{1}$$

$$Binar(I_{pixel}) = \begin{cases} 1 & , I_{pixel} \geq I_{Th} \\ 0 & , I_{pixel} < I_{Th} \end{cases} \tag{2}$$

The *Binar* equation (2) takes the intensity value of the pixel under test and compares it with I_{Th} . If its value is greater than or equal to I_{Th} the pixel is saved as a pixel of interest therefore if it has a smaller intensity the pixel is ignored and considered as part of the background.

In order to demonstrate the results of the Global threshold algorithm a few tests were conducted. These tests aim to discover the best threshold parameter to have the most accurate pixel detection in comparison to the real ship pixels observed in the SAR image.



(a) Original

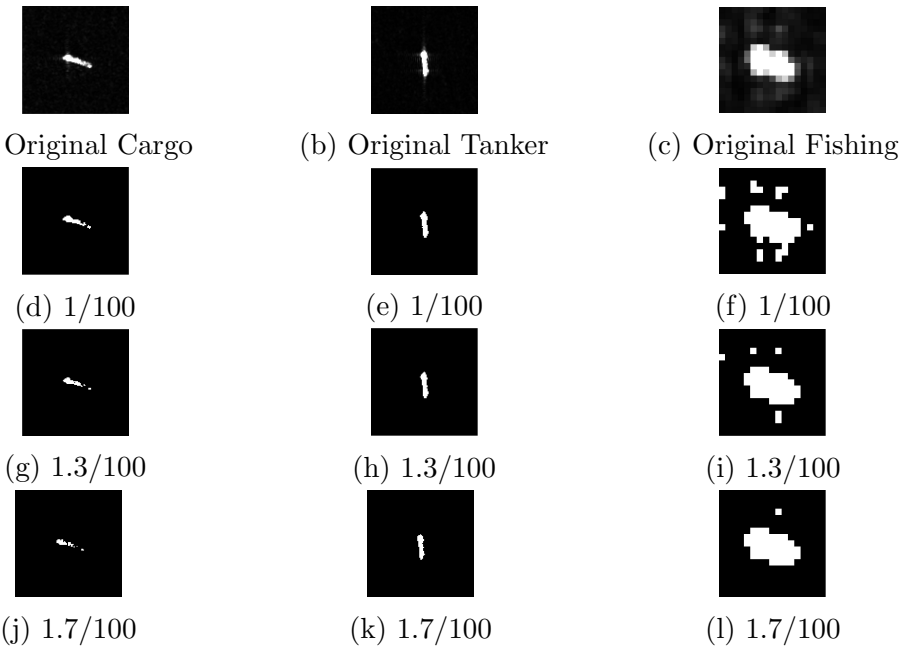
(b) Threshold = 5/100

(c) Threshold = 1/100

(d) Threshold = 0.5/100

Figure 7: Application of the Global threshold algorithm with different thresholds.

Comparing the threshold test results present in figure 7 we come to the conclusion that the optimal threshold sits in the 1/100 threshold mark and that the number of detections is proportionally inverse to the value selected as the threshold. We have to keep in mind that this test was conducted using only one vessel token so in order to better hone the threshold value a test on multiple vessel tokens of different types was conducted.



(a) Original Cargo (b) Original Tanker (c) Original Fishing

(d) 1/100 (e) 1/100 (f) 1/100

(g) 1.3/100 (h) 1.3/100 (i) 1.3/100

(j) 1.7/100 (k) 1.7/100 (l) 1.7/100

Figure 8: Application of the Global threshold algorithm with different thresholds and ship types. The image is organized by: Lines - Different thresholds and Columns - Ship type.

Through image comparison, it is possible to verify that the threshold value has a great impact on false alarms caused by rough sea activity (Figure 8f) and that by using a higher threshold it is possible to minimise these false alarms (Figure 8f and Figure 8i). To determine the best working threshold for the algorithm, a compromise was made, between having a high number of false alarms and lowering the number of ship pixels detected. As a result, the selected threshold for the Global Threshold algorithm was 1.3/100. The outputs of this function are the binarized image and the coordinates of every pixel flagged as of interest.

3.2.2 CFAR binarization

This subsection is divided by a brief explanation of the implemented Cell Averaging CFAR (CA-CFAR) algorithm and afterwards how it is used to binarize the ship token.

CA-CFAR is an algorithm that uses the mean from the background to declare a pixel as a ship candidate if its intensity is above the threshold. For that, it uses two boxes (figure 9), both surrounding the pixel being classified to determine if it's a pixel of interest and, therefore a part of a ship.

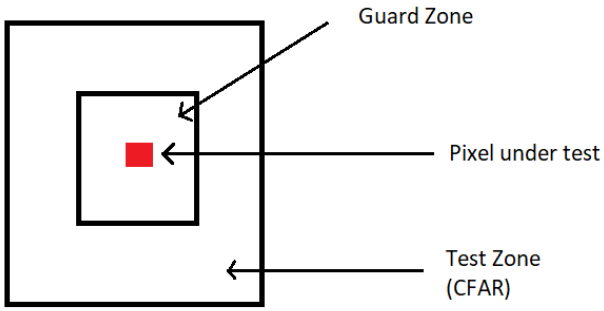


Figure 9: CFAR algorithm idea

There is a larger box and a small one (figure 9). The larger box is the outer limit of pixels and the small box is the guard zone. The pixels in between these boxes are called the Test Zone and they will be used for the calculation of the mean and used for the equation 4 to define if the pixel being tested is a pixel of interest [31].

In this project the dimensions of the outer and inner box are selected by an adaptive limit. This method was chosen due to the fact that every image from the data set has varying dimensions ranging from 9x9 for some Fishing vessels to 204x204 Cargo and Tanker ships and because of this peculiarity the same outer and inner box sizes can't be used for every token or else the detection would heavily downgrade since the test zone could be inside the limits of the ship causing the algorithm to detect false negatives. In order to keep a good detection the ratio between small and medium images is kept the same.

Equation 3 is the mathematical explanation for the adaptative limits, It uses the property of all images being squares and by acquiring the dimension

of the same ($I_{dimensions}$) adapts the Test zone to better detect the pixel of interest:

$$TestZone(I_{dimensions}) = \begin{cases} \frac{I_{dimension}}{3} & , I_{dimension} \leq 68 \\ \frac{I_{dimension}}{3} & , 68 < I_{dimension} \leq 136 \\ \frac{I_{dimension}}{6} & , I_{dimension} > 136 \end{cases} \quad (3)$$

In order to threshold the program to identify the pixels surrounded by water we used the next equation:

$$x_t > \mu_b t \quad (4)$$

Equation (4) where:

x_t - The pixel being evaluated.

μ_b - Background mean.

t - User defined constant to control false alarm rates.

In the following block diagram (figure 10) the flow of the algorithm is presented, starting from the data set SAR image, until saving the coordinates of every pixel of interest.

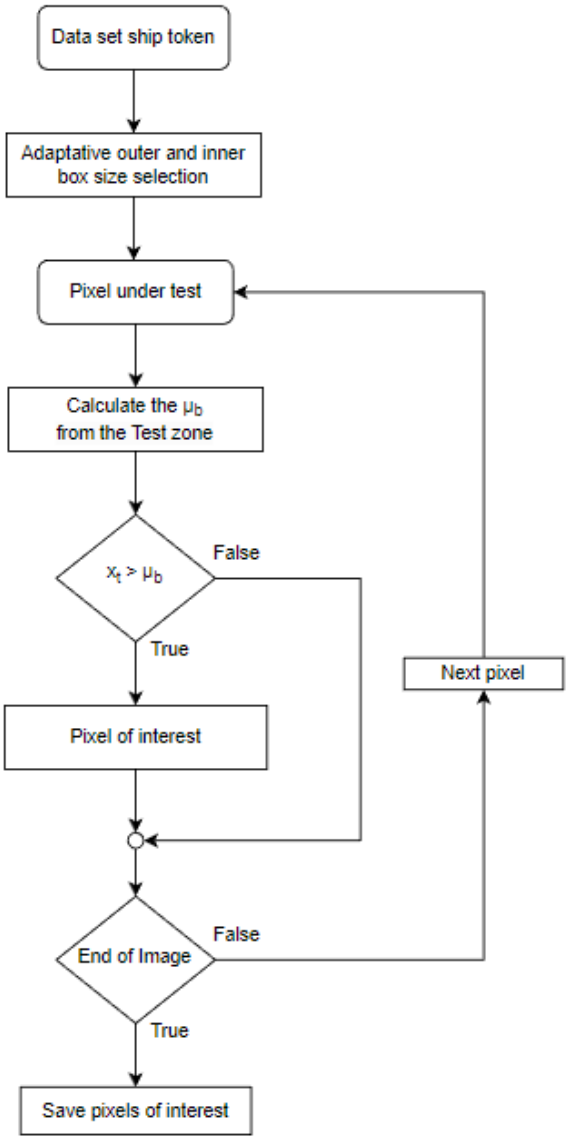
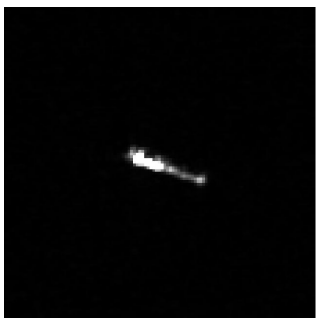


Figure 10: CFAR algorithm Flow Chart

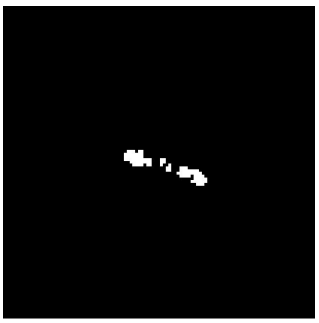
After saving every detected pixel, the binary image is created by using an $I_{dimensions}$x$I_{dimensions}$ bi-dimensional array filled with zeros and replacing the detected pixels location for ones. The outputs for this function, like the

Threshold algorithm, are the binarized image and the coordinates of every pixel flagged as of interest.

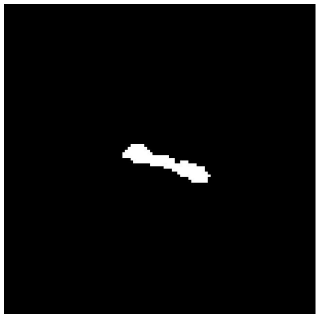
The user defined constant "t" was experimentally chosen. To select the best value for "t", with best false alarm to detection ratio, the following tests were carried out (Figure 11 and Figure 12). These tests consist of applying different values for "t" and observing if the binarized images are as close as possible to the original SAR image and how the threshold works on different types of vessel tokens.



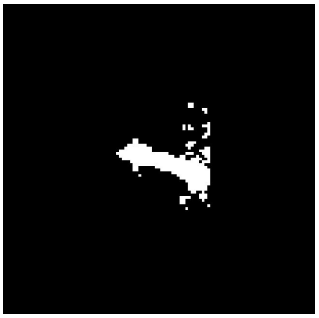
(a) Original



(b) Threshold = 1



(c) Threshold = 0.4



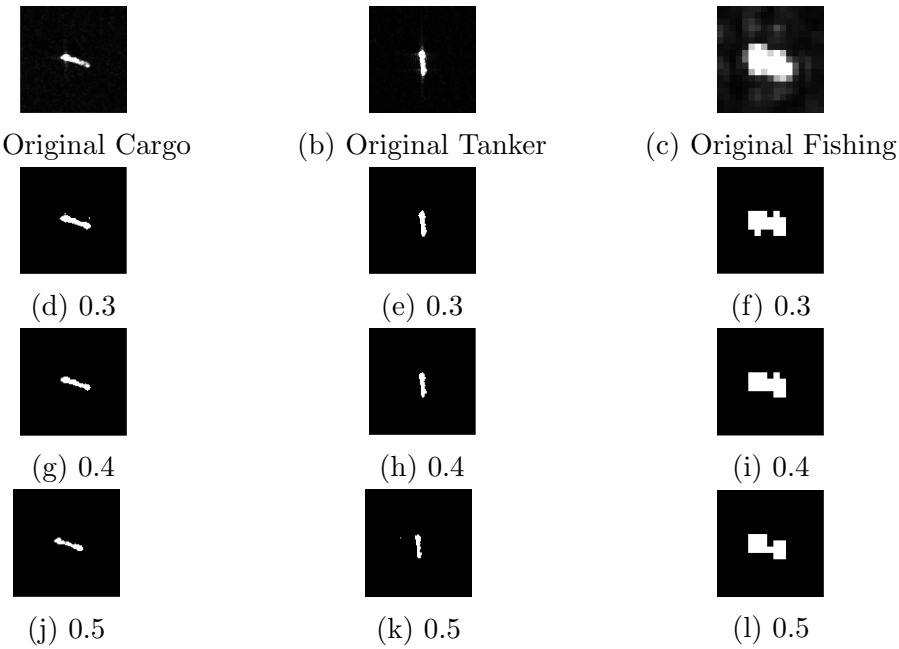
(d) Threshold = 0.1

Figure 11: Application of the Global threshold algorithm with different thresholds.

Later on in this subsection, the efficiency in binarization between the Global Threshold and the CFAR binarization algorithms is compared, in order to keep the integrity of the tests the same ship tokens are used on both tests.

Comparing the images present in Figure 11 it's possible to perceive that Figure 11b is too high of a threshold since some pixels in the middle of the vessel are not flagged as of interest. Figure 11c shows every section of the boat and not detecting any of the background pixels as part of the ship on the other hand using a threshold of 0.1 (Figure 11d) causes a large number of false detections on the background. One can conclude that the best threshold for a good detection rate is around the value of 0.4 (Figure 11c).

The second test is able to identify the best threshold value for a good detection of all types of ships in the data set.



(a) Original Cargo (b) Original Tanker (c) Original Fishing

(d) 0.3 (e) 0.3 (f) 0.3

(g) 0.4 (h) 0.4 (i) 0.4

(j) 0.5 (k) 0.5 (l) 0.5

Figure 12: Application of the CFAR algorithm with different thresholds and ship types. The image is organized by: Lines - Different thresholds and Columns - Different ship types.

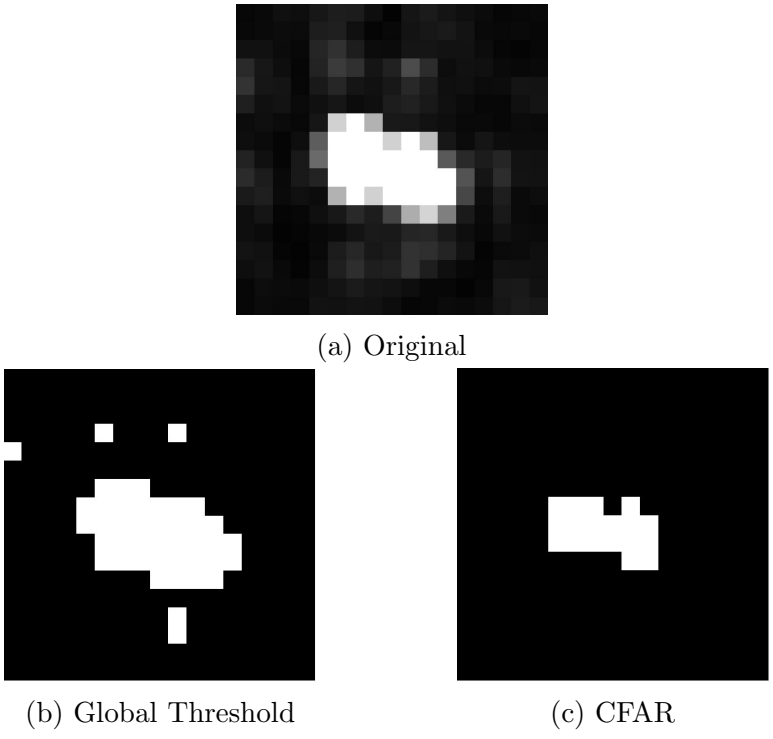
Within the experimentation using the CFAR algorithm this test was conducted on three different ship tokens (Cargo, Tanker and Fishing) using the thresholds of 0.3, 0.4 and 0.5 the variance between these values is not very significative but is a way to refine what threshold is best for ship pixel detection. It is important to note that the differences in pixel detection aren't always substancial. In particular, for Figures 12i and 12l we see a decrease

in pixels flagged as of interest and comming to the conclusion that the 0.5 threshold mark is not the best for this scenario. Furthermore, by comparing Figures 12b and 12c we have an increase in falsely detect pixels, this feature may not be so simple to spot due to the small scaled images.

In summary, this test strengthens the decision to use 0.4 as the final threshold for this algorithm.

3.2.3 Algorithm Comparison

With the objective of determining the best algorithm, some of the binary images were chosen to compare the efficiency of each of them as shown in Figures 14c and 14b.



(a) Original

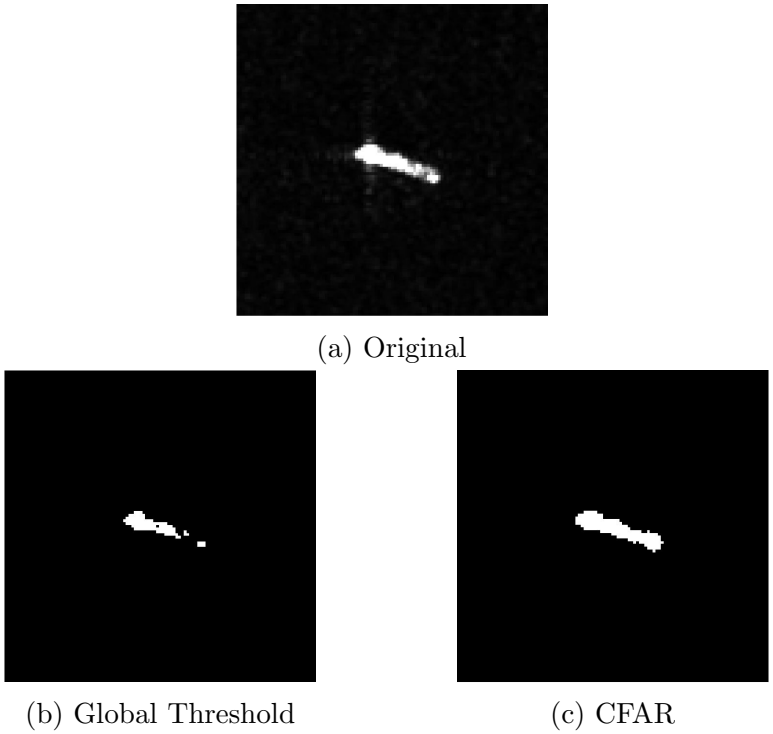
(b) Global Threshold

(c) CFAR

Figure 13: Comparison between both algorithms output binary image on heterogeneous sea.

The images on Figure 13 exemplify a binarization with a heterogeneous sea. This test is very important in particular because it tests the algorithm

on harsh conditions, proving that it gives accurate output even on difficult situations.



(a) Original

(b) Global Threshold

(c) CFAR

Figure 14: Comparison between both algorithms output binary image on vessel pixel detection.

The experiments done above demonstrate that the CFAR algorithm is more efficient in binarizing SAR images than the Global threshold algorithm. The CFAR algorithm adapts very well to detecting the correct pixels in every token type and even with a heterogeneous sea (Figure 13). The value of accurate pixel detection is higher as well on the CFAR algorithm (Figure 14).

The CFAR algorithm maintained an accurate detection therefore it will be the algorithm used for SAR image binarization for further tests.

3.3 Vessel measurements algorithm

Now that the binarization of SAR ship tokens is complete, the next step is to obtain the measurements from the specific vessel. To get a good idea of what vessel is present in the image, the length and width parameters are of the most importance. These measurements can provide an easy comparison for the neural network to distinguish big from small ships and to identify what are the average sizes of each ship type.

The algorithm to acquire the measurements of the ship can be explained as follows, we have a binarized image of the ship and the first step is to identify the center and the rotation angle in which the boat is turned. To obtain these parameters a Matlab function named *regionprops* [32] is used, it calculates the centroid and the angle of rotation of the vessel, and with this information, it is possible to measure the length and width of the ship. To calculate these values two algorithms were tried: one algorithm that uses the size of the ship obtained by the *regionprops* function and another one that uses basic mathematics to identify the parameters. Both algorithms only identify the size in pixel so in order to obtain the calculated length of the ship first we have to multiply it by 10 since it is the Sentinel-1 pixel spacing for high-quality GRD images (Table 3).

$$Length = 10 * P_{Laxys} \tag{5}$$

$$Width = 10 * P_{Waxys} \tag{6}$$

Equation 5 and 6 where:

P_{Laxys} - Pixels under the line on the direction axys.

P_{Waxys} - Pixels under the line perpendicular to the direction axys.

3.3.1 Cluster of pixels identification

Regionprops is a Matlab function that identifies a cluster of pixels, in a binary image, as an object and obtains the radius from the ellipse that covers the cluster area, like presented in Figure 15 [32], and it obtains the angle of the ship by comparing the bigger radius of the ellipse with the x axis.

Figure 15: Elipsoid example extrated from [32].

By using the ellipse method to identify the angle of the ship, it can also estimate the length and width of the vessel.

3.3.2 Linear equation algorithm

The algorithm presented in this section is based on the use of the *region-props* function to obtain the centre and angle of the ship, but it uses a linear equation to determine the ship's directional axis and the perpendicular axis. Using the angle (θ) and the centre of the ship, it's possible to create a line equation similar to 7.

$$y = ax + b \tag{7}$$

This simple equation consists of slope (a) and y-interception (b). The slope can be calculated using $\tan(\theta)$ and b by entering the x and y coordinates of the centre, now we have the equation for the ship's heading axis line. To find the equation for perpendicular line you need to calculate $a_\perp$ and $b_\perp$. It is possible to calculate these parameters using the property of perpendicular lines, given in equation 8.

$$a_\perp = \frac{-1}{a} \text{ Where } a = \text{Slope of ships direction axis} \tag{8}$$

With this property, the value of $a_\perp$ can be easily calculated by using equation 7. The centre coordinates can be used to determine the value of $b_\perp$. Having both the equation we achieve the result present in Figure 16.

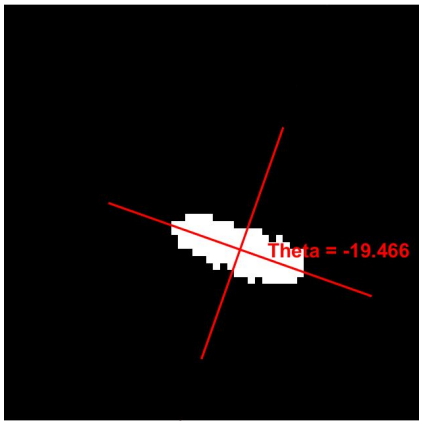


Figure 16: Ship token with the determined lines.

Furthermore, after drawing the lines it is possible to obtain the coordinates and value of the pixels under the line for both the directional axis and the perpendicular axis, making it possible to count the number of pixels with value equal to "one". With this, the algorithm is finished and the number of "one" pixels can be fed to equation 5 and 6 to determine the ship size.

3.4 Statistical aproach

SAR image processing needs a deeper exploration of mathematical parameters to decipher the intricate characteristics of vessels at sea. Among these parameters, standard deviation and mean emerge as indispensable tools in the realm of ship classification, offering an approach to distinguish between ship types.

Classifying ships from SAR imagery is a difficult challenge. Vessels vary in size, shape, orientation, and operational states, contributing to a diverse range of radar signatures. Moreover, environmental factors such as sea state, wind, and the presence of other objects in the scene introduce complexities that must be taken into consideration. Mathematical parameters play a very important role in facilitating this complexity, providing metrics for discriminating between different ship classes in particular for this project differences between Cargo, Tanker and Fishing vessels.

3.4.1 Mean

The main objective of calculating the mean is to measure the central tendency of the vessel and reveal the average intensity of the pixels. With this we obtain a new parameter to enter the network in order to better identify the ship types. The mean of every vessel is calculated using equation 9 having as data all the pixels previously detected as from the ship.

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} a_i \tag{9}$$

The mean can sometimes be a good parameter for vessel classification but must be treated with caution, due to its values depending exclusively on the angle of incidence of the radar electromagnetic waves and the polarization used to acquire the reflected signal.

3.4.2 Standard deviation

While the mean provides information about central tendencies, the standard deviation quantifies the variability or dispersion of radar intensity values within ship regions. This parameter is a powerful tool for capturing the nuances of ship behaviour and conditions. High standard deviations suggest variations in ship radar signatures, which may be indicative of structural differences, damage, or unique operational states. By assessing standard deviation alongside the mean, analysts can discern subtle variations that hold the key to ship classification.

The standard deviation equation is presented at 10.

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \overline{x})^2} \tag{10}$$

3.4.3 Bimodality function to identify the number of scatter points

Bimodality refers to a data type where it exists two distinct peaks or clusters of values, suggesting the presence of two separate distributions within the data. This parameter can be very useful for ship classification due to some ships having distinct scatter points. For instance, on tanker ships it's expected to have only 1 on the main bridge (also referred to as pilothouse).

This function returns a bimodality flag and a bimodality coefficient and calculates these values using a straightforward approach as it only requires three numbers: the sample size, the skewness of the distribution of interest, and its excess kurtosis.

Skewness is a statistical measure that quantifies the asymmetry or lack of symmetry in the distribution of data. It provides information about the direction and degree of asymmetry in a dataset's distribution. In other words, skewness helps us understand whether the data is skewed to the left (negatively skewed), skewed to the right (positively skewed), or roughly symmetric (approximately normally distributed).

Kurtosis is another important statistical measure that provides information about the shape of the distribution of data. Specifically, kurtosis measures the "tailedness" or the degree to which the data distribution has outliers or extreme values (values significantly far from the mean). It indicates whether the data has heavy tails or light tails compared to a normal distribution.

To identify the number of scattering points equation 11 is used [4] in this equation "s" refers to the skewness of the distribution and k to its excess kurtosis.

$$BC = \frac{s^2 + 1}{k + 3.\frac{(n-1)^2}{(n-2)(n-3)}} \tag{11}$$

The BC of the given distribution is then compared to a benchmark value of $BCcrit = 5/9 \approx 0.555$ [33] that would be expected for a uniform distribution higher numbers point toward bimodality whereas lower numbers point toward unimodality. In certain cases the bimodality coefficient can be infinite, in which case the coefficient is set to 0.

A test was conducted in order to identify how the bimodality coefficient functions and to identify in which cases it is detected as being bimodal or unimodal (Figure 17).

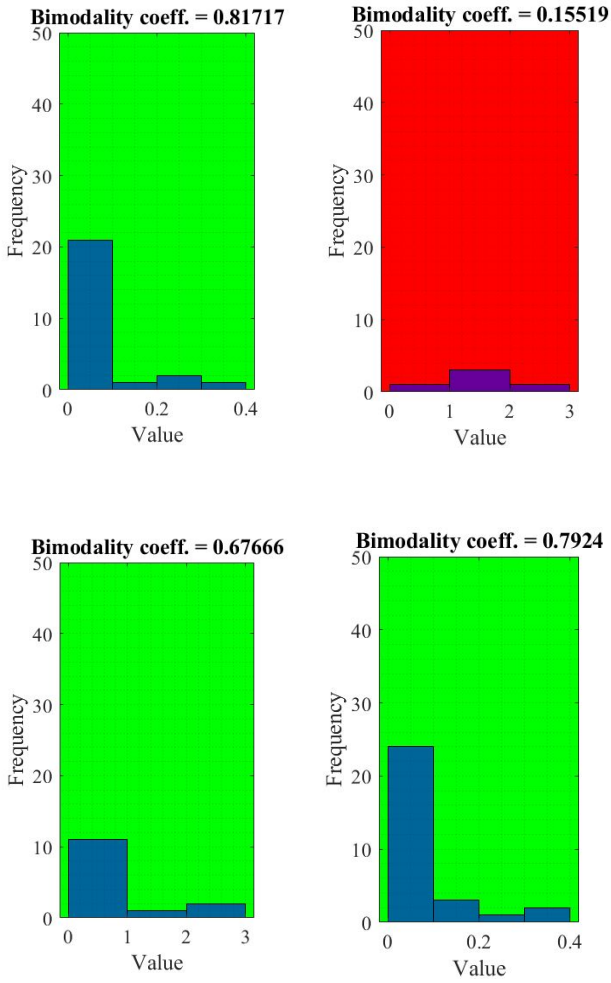


Figure 17: Example of the bimodality function output.

Whenever the distribution of the values, present on the directional axis, have two different peaks of frequency it is classified as bimodal and on the contrary when it only has one peak it's classified as an unimodal distribution, in some cases the value c.

3.5 The neural network

On this topic, the reasoning behind the decision to use a shallow neural network will be presented. Shallow neural networks offer a compelling choice for a great number of tasks, mainly in the classification of vessels, due to its simplicity and efficiency. A key reason for using shallow networks is their ease of training and interpretation. These networks have a fewer number of layers and input parameters, making them less prone to overfitting and an excellent choice for smaller datasets or when computational resources are limited.

Additionally, shallow neural networks can be easily interpreted since they have fewer layers than deep neural networks. Furthermore, shallow networks can be quicker to train, deploy, and maintain, making them a pragmatic choice for rapid prototyping and real-time applications.

The neural network created for this project in order to classify between Cargo, Tanker and Fishing vessels has the following architecture (Figure 18)

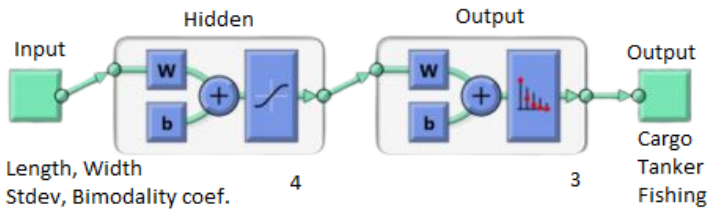


Figure 18: Neural network architecture.

It has 1 hidden layer with 4 perceptrons and 1 output layer with 3 perceptrons and for the input, we use the previously mentioned characteristics of length, width, standard deviation and bimodality coefficient. Every single one of these parameters is obtained by processing the SAR images of every ship individually.

The neural network at work uses the Scaled Conjugate Gradient (SCG) optimization algorithm for training. This algorithm is used to identify the optimal value of weights and biases in the model which minimizes the number of errors in the classification.

The primary objective of SCG in training a neural network is to adjust the model's parameters during the training process to minimize the error between predicted outputs and the actual target outputs in the training data [34].

The cross-entropy loss function is used to ensure network performance. Cross-entropy is widely used for classification problems [35]. The function for multiclass classification can be translated in equation 12.

$$H(Y, P) = -\sum_{i=1}^{n} Y_i \log(P_i) \tag{12}$$

Where:

Y - Represents the true class labels.

P - Represents the predicted probability of the predicted class.

n - Number of classes.

In this project, the value for n will go up to 3 since we only have three possible outputs Cargo, Tanker or Fishing vessels.

4 Results

In this chapter, the results of the vessel classification algorithm will be presented and analyzed. In order to start the discussion, firstly it's needed to explain the data that will be used in the results.

As it was previously mentioned in chapter 3.1.1 if all ship tokens are used in the training of the neural network a fatal problem will occur. The network will be highly biased towards the vessel type in which there are more tokens, in this case Cargo vessels, causing it to only identify this type of vessel. To prevent this biased result, the data set was downsampled to 120 tokens of each type selected for the detection, so that the final data set consists of 120 Cargo, 120 Tanker and 120 fishing vessel tokens. These tokens were not randomly selected since some of the tokens present in the Open Sar Ship data set contain various ships in the same token. For the binarization of the SAR vessel images the algorithm that will be used is the CFAR binarization with a threshold of 0.4. Briefly explained this algorithm uses a cell averaging CFAR to identify if a pixel is part of the vessel and all the pixels flagged by the CFAR are replaced by 1's and the ones not flagged are replaced by 0's, creating the binary image. This algorithm keeps all the pixel values from the ship to be processed to acquire the statistical parameters for the neural network. After obtaining the binary image of the ship the length and width are estimated using the Linear equation algorithm presented in section 3.3.2. Statistical parameters are all obtained by processing the beforehand stored ship pixel values.

In terms of the network used for the classification of vessels, it will have the following distribution (Figure 21): four inputs, one hidden layer and three outputs this number of hidden layers was chosen due to the objective of the thesis to greatly minimize the computational requirements to execute the classification network.

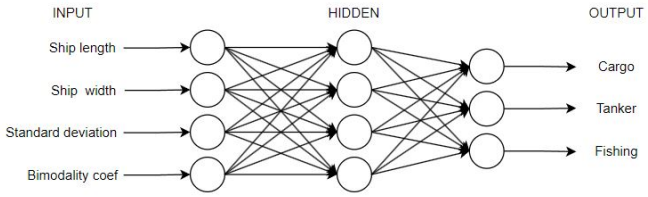


Figure 19: Diagram of the neural network for cargo, tanker and fishing vessels.

This distribution was chosen taking into consideration the number of inputs and the objective of minimising the computational requirements needed to execute the algorithm. Since all the variables are defined, testing of the neural network can start. Firstly by training the network with the parameters as input and a target vector of the expected value for every token.

Figure 20: Confusion matrix of training, validation and test steps.

Figure 20 shows the confusion matrix of results for the training, validation and test. The confusion matrix contains a considerable amount of information so in order to provide a clear and objective explanation of the results, a brief explanation is necessary. On the confusion matrix Figures, the rows correspond to the predicted class (Output Class) and the columns correspond to the true class (Target Class). Since the objective of this neural network is to classify between Cargo, Tanker and Fishing vessels we have 3 classes of output tre correspondence between number and ship type can be clarified by Table 4.

Table 4: Corresponding ship type

Value	Corresponding ship type
1	Cargo
2	Tanker
3	Fishing

The cells along the diagonal indicate accurately classified observations, whereas the cells elsewhere represent those that are misclassified. Both the number of observations and the percentage of the total number of observations are shown in each cell. The column on the far right side of the matrix presents all the vessels predicted to belong to each class that were correctly or incorrectly classified for each class, referred to as the positive predictive value. The row at the bottom of the matrix shows the percentages of all the examples belonging to each class that are correctly and incorrectly classified, referred to as true positive rates. The cell in the bottom right corner shows the overall accuracy of the network.

Figure 21: Confusion matrix of the first neural network.

The neural network's classification results are displayed in Figure 21, the true positive rates are the values to which the accuracy of the network will be compared. For the classification of Cargo vessels we have a 96 correct decisions on the 120 samples given to the network which gives a total of 80.0% accuracy in classifying Cargo ships, for the Tanker vessel type we have only 62 right decisions that lead to a 51.7% ratio and for fishing vessels, we have 102 right decisions and a ratio of correct classification of 85% and an overall accuracy of 72.2%.

These results show that using these parameters the network isn't capable of discerning with good accuracy Tanker vessels but on the other hand we have a very promising result in the classification of fishing vessels and a good result for Cargo vessels as well.

To overcome this inaccuracy problem in the classification of Cargo and Tanker vessels, another network with the same characteristics as the one previously presented was created to differentiate Cargo from Tanker vessels. The difference between these networks is as shown in Figure 22.

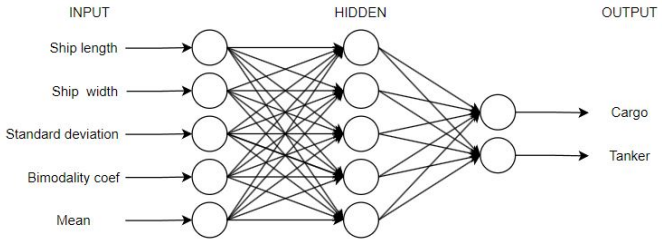


Figure 22: Diagram of the neural network for cargo and tanker vessels.

The characteristic that distinguishes this network from the one previously mentioned is the number of inputs and outputs. By doing various tests using only the length, width, standard deviation and bimodality coefficient the results were not very promising with a successful classification rate of around 63%. Then with the objective of improving the classification, another parameter was selected to be inserted into the network, this parameter was the mean of the pixels from the ship.

Starting with the training of the network by entering into it the input and target vectors of every ship token that was not classified as fishing on the previous network.

Figure 23: Confusion matrix of training, validation and test steps.

By undertaking this approach, we obtain the outcomes illustrated in Figure 24.

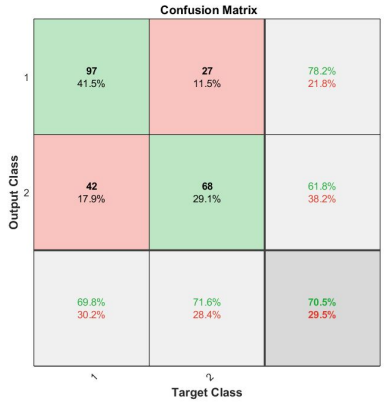


Figure 24: Confusion matrix of the second neural network.

The results of the analysis reveal that the second network has better results for the classification of tanker vessels. Having a 71.2% and 70.5% accuracy rate on classifying Cargo and Tanker ships respectively.

By joining these two networks better results were obtained with an overall accuracy for Cargo ships of 80.0%, for tanker ships 70.5% and for fishing vessels 85.0%. These results are summarised in table 5.

Table 5: Approach final results.

Network	Cargo (%)	Tanker (%)	Fishing (%)
1	80.0	51.7	85.0
2	71.2	70.5	-
Final results	80	70.5	85

5 Conclusions

The objective of this thesis was to create a computationally efficient solution to classify vessels from SAR images using a shallow neural network alongside pre-processed inputs, such as length, width, mean, standard deviation and the bimodality coefficient with the objective of obtaining state of the art results using only a fraction of the computational resources that a Deep Neural Network would use. A shallow neural network consists of a machine learning algorithm with a maximum of three hidden layers. All the data employed for implementing this approach was sourced from the Open SAR Ship dataset [2].

Prior to acquiring these parameters, a preliminary processing of the SAR images is required. To start this process firstly the identification of which pixels are from the vessel is of the utmost importance. Two distinct algorithms that identify the pixels of the ship and also binarize the image were developed. We implemented the Global threshold algorithm that uses a global threshold with the minimum and maximum values of pixels from the image alongside a user defined threshold. Other algorithm that uses a Cell Averaging Constant Alarm Ratio (CA-CFAR) to identify the pixels of interest was also implemented. Both approaches produce good results but the CA-CFAR is an algorithm that adapts more efficiently to rough sea activity in SAR images.

Feature extraction uses a binarized version of the original ship token, using the pixels extracted beforehand, to obtain the dimensional parameters from the ship token. Two algorithms were tested in order to select the one with the best results since accurate vessel length and width estimation using SAR data is a very complex topic. These algorithms consist of using the Matlab function *regionprops* that identifies the centre and angle of the ship, this algorithm can also obtain the length and width of the vessel but due to it being a very generalised function, its outcome isn't the best for this situation. So another algorithm was developed that uses the centre and angle of the ship to create lines in the axis that corresponds to the heading of the ship and the perpendicular one to obtain the corresponding length and width of the vessel. This algorithm has an error of approximately 10 meters due to the resolution of Sentinel-1 GRD images being 20x22 and the pixels spacing 10x10 causing a scatter to overlap the pixels surrounding it. These results were deemed acceptable due to it affecting every vessel the same way.

SAR image processing needs a deeper exploration of mathematical pa-

rameters to decipher the intricate characteristics of vessels at sea, statistical parameters are good for unveiling hidden patterns for ship classification. The fundamental parameters acquired in this project are mean, standard deviation and the bimodality coefficient. The bimodality coefficient refers to data in where it exists two distinct peaks of values, suggesting the presence of two separate distributions within the data. This parameter can be very revealing in ship classification due to certain ships exhibit distinctive scatter points.

The previously estimated parameter were used to train a shallow network to differentiate cargo, tanker and fishing vessels. The training of the first network led to the following accuracy rates: 80.0%, 51.7% and 85.0%, for cargo, tanker and fishing vessels respectively. These results show that this network classifies fishing vessels the best and cannot distinguish cargo and tanker vessels with the same accuracy so, in order to improve the detection rates a new network was trained to classify only cargo and tanker vessels. This second neural network had an accuracy rate of 71.2% and 70.5% for cargo and tanker vessels respectively. By joining both networks and using the first to classify fishing and cargo ships and the second to classify tanker vessels we came to a correct classification rate of 80.0% for cargo, 70.5% for tankers and 85.0% for fishing ships.

6 Future Work

6.1 Improvement of the vessel size acquisition algorithm

Size acquisition is a very important topic for the accurate detection and classification of ships. It can improve the number of correct classifications done by the neural network by a great amount since it can obtain more accurate values. This aspect of the project can be improved by using Dual-Polarization Fusion and applying a nonlinear regression approach [25] creating a more complex approach in ship dimension extraction.

6.2 Mean exploration on the neural network

In the approach presented in this thesis, the mean is used to train the neural network that classifies Cargo and Tanker vessels. This parameter improves the detection rate, of these types of vessels but it was not empirically explained. As future work, the neural network can be upgraded to an auto-explanatory network and by comparing its weights and biases the reason behind this improvement can be unveiled, improving network knowledge.

6.3 Vessel detection

The developed algorithm can classify vessels but only in a controlled environment. In order to test the applicability of the classification algorithm another algorithm to detect vessels can be concatenated. Creating a functioning detection and classification algorithm and testing how the developed algorithm would respond to all types of ships and verifying the degree of accuracy in the classification of vessels with AIS Data.

References

[1] Alberto Moreira et al. "A Tutorial on Synthetic Aperture Radar". en. In: ieee Geoscience and remote sensing magazine (2013), p. 39.

[2] Lanqing Huang et al. "OpenSARShip: A Dataset Dedicated to Sentinel-1 Ship Interpretation". en. In: 11.1 (Jan. 2018), pp. 195–208. ISSN: 1939-1404, 2151-1535. DOI: 10.1109/JSTARS.2017.2755672. URL: https://ieeexplore.ieee.org/document/8067489/.

[3] M. Weiss. "Analysis of Some Modified Cell-Averaging CFAR Processors in Multiple-Target Situations". en. In: IEEE Trans. Aerosp. Electron. Syst. AES-18.1 (Jan. 1982), pp. 102–114. ISSN: 0018-9251. DOI: 10.1109/TAES.1982.309210. URL: http://ieeexplore.ieee.org/document/4102622/ (visited on 10/17/2023).

[4] Hristo Zhivomirov. "Bimodality Coefficient Calculation with Matlab". en. In: (2023). URL: https://www.mathworks.com/matlabcentral/fileexchange/84933-bimodality-coefficient-calculation-with-matlab.

[5] Stephan Brusch et al. "Ship Surveillance With TerraSAR-X". In: 49.3 (Mar. 2011). Conference Name: IEEE Transactions on Geoscience and Remote Sensing, pp. 1092–1103. ISSN: 1558-0644. DOI: 10.1109/TGRS.2010.2071879.

[6] Noboru Oishi and Kei Suwa. "Azimuth Ambiguity Detection and Suppression in SAR Images". en. In: IGARSS 2019 - 2019 IEEE International Geoscience and Re Yokohama, Japan: IEEE, 2019, pp. 676–679. ISBN: 978-1-5386-9154-0. DOI: 10.1109/IGARSS.2019.8898514.

[7] Lefei Zhang and Liangpei Zhang. "Artificial Intelligence for Remote Sensing Data Analysis: A review of challenges and opportunities". en. In: IEEE Geoscience and Remote Sensing Magazine 10.2 (June 2022), pp. 270–294. ISSN: 2168-6831, 2473-2397, 2373-7468. DOI: 10.1109/MGRS.2022.3145854. URL: https://ieeexplore.ieee.org/document/9756442/.

[8] Javad Karimi Gahfarrokhi and Mojtaba Abolghasemi. "Fast VI-CFAR Ship Detection in HR SAR Data". en. In: Tabriz, Iran, Aug. 2020, pp. 1–5. ISBN: 978-1-72817-296-5. DOI: 10.1109/ICEE50131.2020.9260974. URL: https://ieeexplore.ieee.org/document/9260974/.

[9] Pape Sanoussy Diao et al. “A review of Radar Detection Fundamentals”. en. In: (2022), pp. 1–1. ISSN: 0885-8985, 1557-959X. DOI: 10.1109/MAES.2022.3177395. URL: https://ieeexplore.ieee.org/document/9781853/.

[10] Deyu Song et al. “A Novel Hog-Based Template Matching Method for SAR and Optical Image”. en. In: Kuala Lumpur, Malaysia, July 2022, pp. 951–954. ISBN: 978-1-66542-792-0. DOI: 10.1109/IGARSS46834.2022.9884706. URL: https://ieeexplore.ieee.org/document/9884706/.

[11] Su Liu and Gong Zhang. “Euler 2D-PCA for SAR target recognition”. en. In: Hong Kong, China: IEEE, Aug. 2016, pp. 1–6. ISBN: 978-1-5090-2708-8. DOI: 10.1109/ICSPCC.2016.7753684. URL: http://ieeexplore.ieee.org/document/7753684/.

[12] Jianwei Fan et al. “SAR Image Registration Using Phase Congruency and Nonlinear Diffusion-Based SIFT”. en. In: 12.3 (Mar. 2015), pp. 562–566. ISSN: 1545-598X, 1558-0571. DOI: 10.1109/LGRS.2014.2351396. URL: https://ieeexplore.ieee.org/document/6894582.

[13] Zhou Jianxiong et al. “Automatic Target Recognition of SAR Images Based on Global Scattering Center Model”. en. In: 49.10 (Oct. 2011), pp. 3713–3729. ISSN: 0196-2892. URL: http://ieeexplore.ieee.org/document/6018300/.

[14] Ayoub Karine et al. “Target Recognition in Radar Images Using Weighted Statistical Dictionary-Based Sparse Representation”. en. In: 14.12 (Dec. 2017), pp. 2403–2407. ISSN: 1545-598X, 1558-0571. DOI: 10.1109/LGRS.2017.2766225. URL: http://ieeexplore.ieee.org/document/8107684/.

[15] Amir Hosein Oveis et al. “A Survey on the Applications of Convolutional Neural Networks for Synthetic Aperture Radar: Recent Advances”. en. In: IEEE Aerospace and Electronic Systems Magazine 37.5 (May 2022), pp. 18–42. ISSN: 0885-8985, 1557-959X. DOI: 10.1109/MAES.2021.3117369. URL: https://ieeexplore.ieee.org/document/9656541/.

[16] Giorgos Mountrakis, Jungho Im, and Caesar Ogole. “Support vector machines in remote sensing: A review”. en. In: 66.3 (May 2011), pp. 247–259. ISSN: 09242716. DOI: 10.1016/j.isprsjprs.2010.11.

001. URL: https://linkinghub.elsevier.com/retrieve/pii/S0924271610001140.

[17] Pall Oskar Gislason, Jon Atli Benediktsson, and Johannes R. Sveinsson. "Random Forests for land cover classification". en. In: 27.4 (Mar. 2006), pp. 294–300. ISSN: 01678655. DOI: 10.1016/j.patrec.2005.08.011. URL: https://linkinghub.elsevier.com/retrieve/pii/S0167865505002242.

[18] What Is Deep Learning? — How It Works, Techniques & Applications. en. Library Catalog: www.mathworks.com. URL: https://www.mathworks.com/discovery/deep-learning.html.

[19] Lazhar Khelifi and Max Mignotte. "Deep Learning for Change Detection in Remote Sensing Images: Comprehensive Review and Meta-Analysis". en. In: 8 (2020), pp. 126385–126400. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.3008036. URL: https://ieeexplore.ieee.org/document/9136674/.

[20] Bariscan Yonel, Eric Mason, and Birsen Yazici. "Deep Learning for Passive Synthetic Aperture Radar". en. In: 12.1 (Feb. 2018), pp. 90–103. ISSN: 1932-4553, 1941-0484. DOI: 10.1109/JSTSP.2017.2784181. URL: https://ieeexplore.ieee.org/document/8214209/.

[21] Emile Ndikumana et al. "Deep Recurrent Neural Network for Agricultural Classification using multitemporal SAR Sentinel-1 for Camargue, France". en. In: Remote Sensing 10.8 (Aug. 2018), p. 1217. ISSN: 2072-4292. DOI: 10.3390/rs10081217. URL: http://www.mdpi.com/2072-4292/10/8/1217.

[22] Xiangrong Zhang et al. "Spatial Sequential Recurrent Neural Network for Hyperspectral Image Classification". en. In: 11.11 (Nov. 2018), pp. 4141–4155. ISSN: 1939-1404, 2151-1535. DOI: 10.1109/JSTARS.2018.2844873. URL: https://ieeexplore.ieee.org/document/8399509/.

[23] Yushi Chen et al. "Deep Learning-Based Classification of Hyperspectral Data". en. In: 7.6 (June 2014), pp. 2094–2107. ISSN: 1939-1404, 2151-1535. DOI: 10.1109/JSTARS.2014.2329330. URL: https://ieeexplore.ieee.org/document/6844831/.

[24] Jiayuan Kong and Fangzheng Zhang. "SAR Target Recognition with Generative Adversarial Network (GAN)-based Data Augmentation". en. In: Oct. 2021, pp. 215–218. ISBN: 978-1-66543-188-0. DOI: 10.1109/ICAIT52638.2021.9701974. URL: https://ieeexplore.ieee.org/document/9701974/.

[25] Boying Li et al. "Ship Size Extraction for Sentinel-1 Images Based on Dual-Polarization Fusion and Nonlinear Regression: Push Error Under One Pixel". en. In: IEEE Trans. Geosci. Remote Sensing 56.8 (Aug. 2018), pp. 4887–4905. ISSN: 0196-2892, 1558-0644. DOI: 10.1109/TGRS.2018.2841882. URL: https://ieeexplore.ieee.org/document/8392506/ (visited on 09/06/2023).

[26] "A Small Ship Target Detection Method Based on Polarimetric SAR". en. In: (), p. 2938. DOI: 10.3390/rs11242938.

[27] User Guides - Sentinel-1 SAR - Sentinel Online - Sentinel. URL: https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-1-sar.

[28] The Information Sent through AIS. URL: https://help.marinetraffic.com/hc/en-us/articles/205426887-What-kind-of-information-is-AIS-transmitted-.

[29] Yuanyuan Wang et al. "A SAR Dataset of Ship Detection for Deep Learning under Complex Backgrounds". In: Remote Sensing 11.7 (2019). ISSN: 2072-4292. DOI: 10.3390/rs11070765. URL: http://www.mdpi.com/2072-4292/11/7/765.

[30] Copernicus: Sentinel-1. URL: https://www.eoportal.org/satellite-missions/copernicus-sentinel-1#sentinel-1b-status.

[31] Paulo A C Marques. "Lecture 4 - Ship processing". en. In: Warsaw University of Tecnology (2013).

[32] The RegionProps Matlab Function. URL: https://www.mathworks.com/help/images/ref/regionprops.html?s_tid=doc_ta.

[33] Roland Pfister et al. "Good things peak in pairs: a note on the bimodality coefficient". en. In: Front. Psychol. 4 (2013). ISSN: 1664-1078. DOI: 10.3389/fpsyg.2013.00700. URL: http://journal.frontiersin.org/article/10.3389/fpsyg.2013.00700/abstract (visited on 09/18/2023).

[34] Xiao-Bo Jin et al. “Stochastic Conjugate Gradient Algorithm With Variance Reduction”. In: (May 2019), pp. 1360–1369. DOI: 10.1109/TNNLS.2018.2868835.

[35] Shie Mannor, Dori Peleg, and Reuven Rubinstein. “The cross entropy method for classification”. en. In: Bonn, Germany, 2005, pp. 561–568. DOI: 10.1145/1102351.1102422.