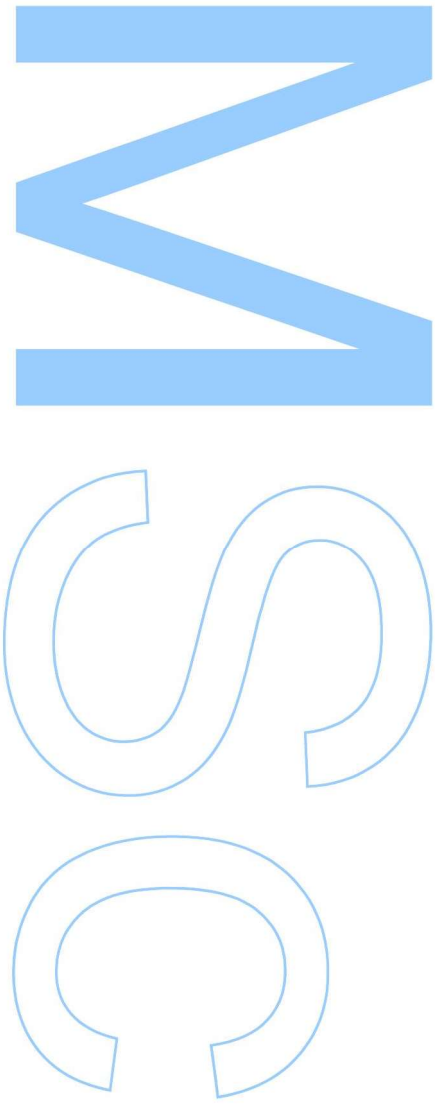
Detection and classification of objects in the ocean

João Carvalho
Mestrado em Estatística Computacional e Análise de Dados
Departamento de Matemática
2022

Orientador
Prof. Dr. André Marçal, Faculdade de Ciências

Local Supervisor
Dr. Miguel Arantes, CEiiA



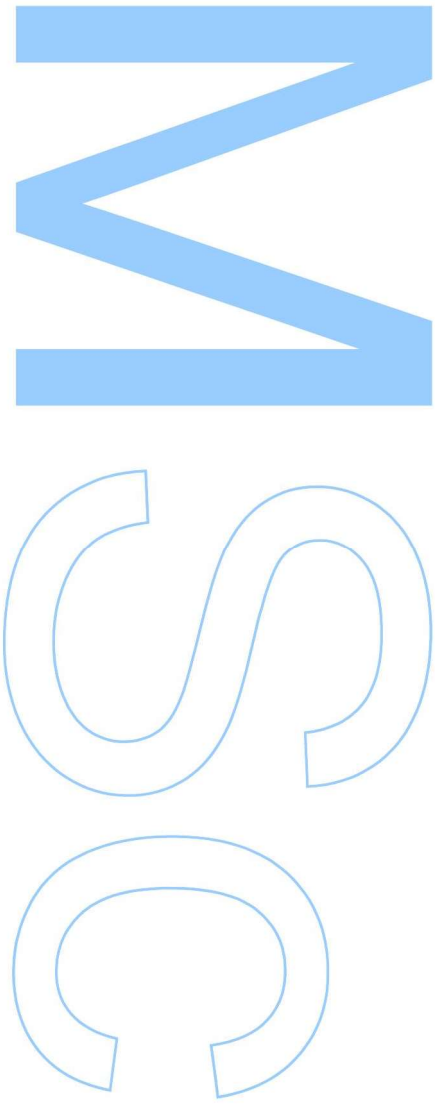
U.PORTO

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, _____/_____/_____



UNIVERSIDADE DO PORTO

MASTERS THESIS

Detection and classification of objects in the ocean

Author:
João CARVALHO

Supervisor:
André MARÇAL

Local Supervisor:
Miguel ARANTES

A thesis submitted in fulfilment of the requirements
for the degree of MSc. Computational Statistics and Data Analytics

at the

Faculty of Sciences of University of Porto
Mathematics Department

January 23, 2023

Sworn Statement

I, João Paulo Teixeira de Carvalho, enrolled in the Master Degree Computational Statistics and Data Analytics at the Faculty of Sciences of the University of Porto hereby declare, in accordance with the provisions of paragraph a) of Article 14 of the Code of Ethical Conduct of the University of Porto, that the content of t this dissertation reflects perspectives, research work and my own interpretations at the time of its submission.

By submiting this dissertation, I also declare that it contains the results of my own research work and contributions that have not been previously submitted to this or any other institution.

I further declare that all references to other authors fully comply with the rules of attribution and are referenced in the text by citation and identified in the bibliographic references section. This dissertation does not include any content whose reproduction protected by copyright laws.

I am aware that the practice of plagiarism and self-plagiarism constitute a form of academic offense.

João Paulo Teixeira de Carvalho
30 September 2022

Acknowledgements

First of all, I would like to thank myself for all the work and resilience I put on in order to conceive this study.

Secondly, I would like to thank both my teacher and supervisor, Dr. André Marçal, Department of Mathematics, Faculty of Sciences of University of Porto and my internship supervisor, Dr. Miguel Arantes, CEiiA - Centre of Engineering and Product Development, for all their support, guidance, patience and dedication, which led me to successfully finish this work.

This would have never been possible without my family's support, in other words, thanks to my mother, my father and my sister.

Amongst my friends and colleagues, a especial thank you to Pinto, Isabel and B, for everything.

Last but not least, a catlike thank you to Sky and Gray for making me company every single day while this work was being produced.

Abstract

Over the years, the number of earth observation satellites has been increasing leading to the production of data capable of being used to model detection algorithms. These are of extremely importance for tracking maritime activity for safety and economical reasons.

In order to contribute to this matter, three segmentation algorithms were built and tested using the band Red. The first one is Region Growing that consisted in threshold technique and arithmetic operations. The second one, Edges Growing used edge detection with threshold, morphological operations techniques and arithmetic operations. The third one, K-means Clustering uses 6 cluster per image with arithmetic operations.
AIS data was used to build labels for the observations as well as to validate the correct object detected based on its coordinates.
For each segmentation algorithm, six classifiers were used: Support Vector Classifier, K-nearest Neighbours, Decision Trees, Bagging, Random Forest and Gradient Boosting. Two evaluating scores, F1 score and Jaccard index were used to compare these combinations. The Edges growing was the best segmentation algorithm while Bagging was the best classifier.

In the future, it would be not only interesting to use other band combinations but also to experiment the YOLO algorithm capable of detection and classification.

Keywords: Vessel, Detection, Segmentation, Classification, Images, Phi-Sat, Satellites

Resumo

Ao longo dos ano, o número de satélites de observação da terra tem estado a aumentar contribuindo para a produção de dados capazes de serem utilizados para modelar algorítmos de deteção. Este são de extrema importância para acompanhar a atividade marítima por motivos económicos e de segurança.

De modo a contribuir para esta temática, foram criados e testados três algorítmos de segmentação utilizando a banda vermelha. A primeira é Region Growing que consiste na técnica binarização com operações aritméticas. O segundo, Edges Growing, consiste na deteção de contornos com binarização, operações morfológicas e operações aritméticas. O terceiro, K-means Clustering utiliza 6 clusters por imagem e operações aritmticas.
Os dados do AIS são utilizados não sõ para construir classes para as observações mas também para validar os objected corretamente detetados com base nas suas coordenadas.
Para cada algorítmo de segmentação, seis classificadores foram utilizados: Máquinas de Suporte Vetorial, K-vizinhos mais próximos, Ãrvores de Decisão, Bagging, Florestas Aleatórias e Gradient Boosting. Foram utilizadas duas métricas de avaliação para comparar estas combinações - F1 e o índice de Jaccard.
O algorítmo Edges Growing foi o melhor algorítmo de segmentação enquanto que o Bagging foi o melhor classificador.

No futuro, seria não apenas interessante utilizar outras combinações de bandas como também de experimentar o algorítmo de deteção e classificação YOLO.

Palavras-chave: Vessel, Deteção, Segmentação, Classificação, Imagens, Phi-Sat, Satélites

Contents

List of Figures

List of Tables

Glossary

ADS-B	Automatic Dependent Surveillance-Broadcast
AIS	Automatic Information Systems
API	Application Programming Interface
AVA	Autonomous Vessel Awareness
BAGG	Bagging
CSA	Canadian Space Agency
CSV	Comma Separated Value (type of file)
CEiiA	Center of Engineering and Product Development
DCT	Decision Trees
EC	European Comission
EDA	Exploratory Data Analysis
EEZ	Exclusive Economic Zone
ESA	European Spacial Agency
EO	Earth Observation
EU	European Union
FP	False Positive
FN	False Negative
GRAD	Gradient Boosting
IMO	International Maritime Organization
KNN	K-nearest Neighbors
LRIT	Long Range Identification and Tracking

NASA	National Aeronautics and Space Administration
NIR	Near Infrared
OI	Original Image
PCA	Principal Component Analysis
RFOR	Random Forests
RGB	Red Green and Blue
ROI	Region of Interest
SAR	Synthetic Aperture Radar
SIFT	Scale Invariant Feature Transform
SWIR	Short-waved Infrared
SVC	Support Vector Classifier
SVM	Support Vector Machine
TP	True Positive
TN	True Negative
VMS	Vessel Monitoring Systems

Chapter 1

Introduction

A topic of significant and growing interest is measuring and tracking marine human activity. Governmental and private players make attempts to stay informed about what is happening in a variety of industries, including fishing, drilling, exploration, or freight transit as well as carrying passengers or tourism.
The number of EO satellites has increased in a considerable way over the last few years which have been generating enormous amount of data, the majority of it yet to be analysed [1].
For security, economic and health reasons, it is important to generate efficient algorithms capable of becoming part of detecting and tracking devices that will provide information in real time to the stakeholders.

Having said that, this study intents to build an algorithm capable of detecting and classifying vessels on the sea. This is the result of an internship program done in CEiiA - Centre of Engineering and Product Development, in the scope of the Master's in Computational Statistics and Data Analytics.

1.1 Contextualization

Prior to 1982, there was no aggregated areas of sea that would be equivalent to the Freedom of Movement (on land) later to be established by the European Union (EU) in 1992. This means, each country would only need to worry about their national area of jurisdiction, which could be controlled by people on shore, and investigated by coastal guards or

the navy, if necessary [2].

In 1982, the UN Convention on the Law of the Sea (UNCLOS) has built the constitution of the oceans. The extensive of EEZ increased the average area of 12 nm national jurisdiction to 200 nm, making it impossible to monitor with naked eye, leading to the investment on marine and aerial patrols. These high monitoring costs could only be maintained by the richest countries enabling criminal actors to take advantage of this situation on less controlled areas.This fact triggered numerous countries in seeking for ways of monitor and survey the vast maritime area, which would a few years later be known as the maritime surveillance system [2].

The maritime surveillance is done with the combination of multiple sensor systems. Some of these systems are cooperative and mostly being global, such as AIS, VMS and LRIT, with the help of satellite or satcom. On the other hand, coastal radars, satellite and airborne EO systems (optical and SAR systems), as well as ground-based cameras and vessel-based cameras are part of the Non-cooperative systems [3].

The automatic identification system (AIS) is an automatic tracking system that uses transceivers on ships and is used by harbour and port authorities [4]. It was build by IMO with the intent of avoiding vessels collisions (similar to ADS-B for aircraft).

Despite the fact that AIS's original purpose was to avoid vessel collisions, many other application have since been developed. Nowadays, it is capable of:

- Collision avoidance
- Fishing fleet monitoring and control
- Maritime security
- Aids to navigation
- Search and rescue
- Accident investigation
- Ocean currents estimates

- Infrastructure protection
- Fleet and cargo tracking
- Statistics and economics

The AIS may be terrestrial-based (T-AIS) [5] or satellite-based (SAT-AIS) [6], and transmits information every 2 to 180 seconds, depending on a vessel's activity [7].
This information is presented in the table 1.1.

MMSI	Maritime Mobile Service Identity	**NAME**	Vessel's name
TIME	Data timestamp	**CALLSIGN**	vessels's callsign
LONGITUDE	Geographical latitude in degrees	**TYPE**	Vessels's type
LATITUDE	Geographical longitude in degrees	**DEVICE**	Positioning devide type
COG	Courser over ground	**A**	DImension to bow, in meters
SOG	Speed over ground	**B**	DImension to stern, in meters
HEADING	Heading of the AIS vessel	**C**	DImension to port, in meters
PAC	Position accuracy	**D**	DImension to starboard, in meters
ROT	Rate of turn	**DRAUGTH**	Draugth multiplied by 10
NAVSTAT	Navigational status	**DEST**	Vessel's destination
IMO	IMO ship identification number	**ETA**	Estimated time of arrival

TABLE 1.1: AIS data.

ESA stands for *European Space Agency* which, as the name suggests, aims to explore the space. It is an intergovernmental organisation composed by 22 states and headquartered in Paris. ESA was a key partner in the collaboration wtih NASA and CSA in the building of James Webb Space Telescope, the world's premier infrared space observatory [8].

Copernicus is the EU's Earth observation programme, implemented by the EC with the support of ESA [9]. This programme aims to monitor and forecast the state of the environment on sea, land and in the atmosphere. This is possible through the Copernicus Sentinel missions (Sentinel 1, 2, 3, 4, 5P, 5, 6) developed by ESA, which carry different technologies such as RADAR and multi-spectral imaging instruments. The first mission, Sentinel-1, launched the Sentinel-1A satellite in 3 April of 2014 and the Sentinel-1B on 25 April 2016 (ended in 2022). The latest mission, Sentinel-6, launched the Sentinel-6A in November 2020. Each mission has its own Sentinel Satellite which may or may not

have specific technology according to the mission's objective. For this study, images from Sentinel-2 satellites were used. These twin satellites provide new images from the same location every 5 days [10]

The Sentinel-2 satellites come with 12 high-resolution multispectral bands. Table 1.2, presents these band's main characteristics.

	Spectral Band	**Centre Wavelength (nm)**	**Band Width (nm)**	**Spatial Resolution (m)**
B1	Coastal,aerosol	443	20	60
B2	Blue,(B)	490	65	10
B3	Green,(G)	560	35	10
B4	Red,(R)	665	30	10
B5	Red-edge,1,(Re1)	705	15	20
B6	Red-edge,2,(Re2)	740	15	20
B7	Red-edge,3,(Re3)	783	20	20
B8	Near,infrared,(NIR)	842	115	10
B8a	Near,infrared,narrow,(NIRn)	865	20	20
B9	Water,vapor	945	20	60
B10	Shortwave,infrared/Cirrus	1380	30	60
B11	Shortwave,infrared,1,(SWIR1)	1910	90	20
B12	Shortwave,infrared,2,(SWIR2)	2190	180	20

TABLE 1.2: The spectral bands and spatial resolutions of Sentinel-2 MSI sensor. [11]

ESA has been developing and deploying different technologies and programmes one of which is called Phi-Sat.

ESA's Phi-sat programmes aim to implement Artifitial Intelligence technologies to nano and mini satellites [12]. This last programme in execution, ϕ-sat-2 consists in EO Cube-Sat plataform capable of running AI Applications than can be developed, deployed on the spacecraft and added/updated also during the satellite flight and operated from the ground using a simple user interface (NMF- Nanosat MO Framework), supporting on-board App uploading.

This mission aims to implement on the CubeSat [13][14] several sophisticated AI-based applications on of which is AVA - Autonomous Vessel Awareness App, which became

responsibility of CEiiA as a direct subcontractor of ESA.

AVA consists in detecting and classifying vessels in optical imagery, determining which images or subsets should be downlinked. It will feature autonomous tasking, providing relevant geographic coordinates and time slots of areas where significant features are detected.

One major issue of maritime domain surveillance is the vast extension of the sea on the Earth's surface which makes observation of ship traffic difficult [2]. Normally, one effective method to obtain information about a ship's current position in near-real-time is the use of satellite-based AIS (automatic identification system). However, when performing illegal activities, the AIS is usually shut down, thus these ships become undetected.
Main oceanic activities requires observation: maritime safeguard is not only essential for carriage goods but also for the integrity of human lives. Piracy, illegal dumping (e.g., tanks cleaning), smuggling and refugee transportation.
Earth observation satellite systems has several advantages to tackle these issues; however, images are processed long after the data is acquired (normally several hours). Hence, the antenna acquisition time is expensive.
By processing and analyzing them on-board and transmitting the product data directly as ship position, heading, and velocity, the delay can be shortened to some minutes.
Figure 1.1 presents the workflow diagram for AVA.

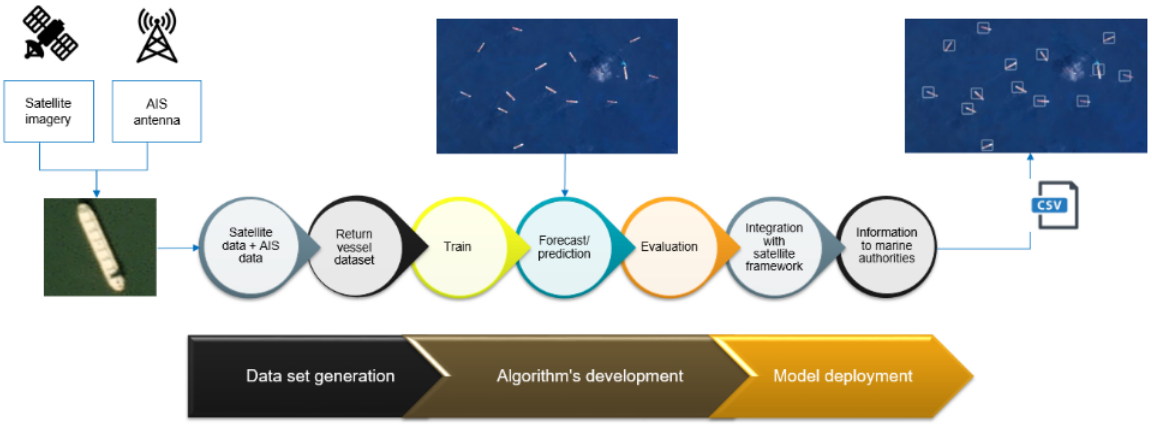


FIGURE 1.1: Workflow, Source: CEiiA.

1.2 Objectives

The main goal of the work presented is to implement different algorithms capable of identifying and classifying vessels on images provided by Sentinel-2, through image segmentation and classification techniques, using Sentinel-2 and AIS data combined.

1.3 Document Structure

This work is divided in seven main Chapters. Chapter 1 contains the introduction and contextualization of the work. Chapter 2 presents the mathematical and statistical concepts. Chapter 3 consists of a literature review about vessel detection techniques. The exploratory data analysis of AIS data is in Chapter 4. Chapter 5 holds the methodology followed by this work's results on Chapter 6. Last but not least, some conclusions along with future work suggestions appear on Chapter 7.

Chapter 2

Mathematical Fundamentals

2.1 Image Processing

Digital Image processing is a subfield of computer vision and consists in the manipulation of images using computers. These manipulations can be image enhancement, image restoration, image analysis and image compression [15].

Digital Image corresponds to a grid of squares called pixels, which possess a value from 0 (black) to 255 (white), each one. A band on a digital image corresponds to a range of wavelengths measured by a sensor. In the case of human eye, it can only detect 3 wavelengths, that combined, create what it is called a "colour image" [16].

Image segmentation is one field of study that concentrates into representing the images in different regions [17]. These regions may be differentiated due to their shape, pixel intensity, colour among others. The following section mentions some common techniques used in image detection and segmentation. The next two sections, talk about some classification algorithms and evaluating metrics, respectively.

2.2 Detection Methods

2.2.1 Threshold-based methods

Otsu Algorithm

The algorithm proposed by Otsu consists in minimizing the weighted variance between

foreground and background pixels. This is done through iteration over the possible threshold values, leading to the formation of binary images [18].

Multi-level threshold segmentation

On the other hand, it may be interesting to apply different threshold values to at the same time, creating several distinct regions. It a very used method for objects with coloured or complex backgrounds [19].

2.2.2 Salient-based methods

Fourier Transform

Fourier transform decomposes the image into its sine and cosine components. The input is an image (spatial domain) and the output is the image's frequency domain. For a NxN size image, the two dimensional Discrete Fourier Transform are given by:

$$F(k,l) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i,j) e^{-\imath 2\pi \left(\frac{ki}{N} + \frac{lj}{N}\right)}$$

where f(i,j) is the image in the spatial domain and the exponential term is the basis function corresponding to each point F(k,l) in the Fourier space. The value of each point F(k,l) is obtained by multiplying the spatial image with the corresponding base function and summing the result [20].

2.2.3 Methods based on shape and texture

Mathematical morphology

Morphological operations (MO) use basically logical operations even though these are part of the set theory [21]. Four common MO are, dilation, erosion, opening and closing. This is done through a structure element. The dilatation amplifies the size of the 0-valued region, in the binary image. The shrink decreases the size of the 0-values regions. The third, opening, consists in applying firstly the erosion operation and then the dilatation, while for closing, it is first applied the dilatation operation and finally the erosion operation.

Hir-Or-Miss Transformations

This transformation is based on the erosion MO. With two different structural elements, each will be applied on solely on the image, resulting in two images. Then, the final object will be considered as an intersection of these two images [21].

2.2.4 Statistical methods

Principal Component Analysis

PCA is a statistical method that aims to find the main features on a dataset focusing on the total variance. Having a dataset in the shape of a matrix, its rows are the observations and its columns the features. The groups of pixels arranged are in floating point values if the OI is RGB and discrete values if it is a gray scaled image.
First the data is normalized (subtracting the image by its mean value) then, the covariance Matrix is calculated using (X), where m is the number of element.

$$\mathrm{Cov}(X,Y) = \frac{F_{\text{normalized}}\,(x,y) * F_{\text{normalized}}\,(x,y)^T}{m-1}$$

, where m is the number of element.

The eigenvector represent the principal features of the image are typical ordered according to the corresponding eigenvalues, in decreasing way until a quantity that corresponds to a summation of 90% is met. In other words, these chosen features are capable of explaining 90% of the variance of the image. This is useful since it makes data lighter, consuming less computational time [22].

Bayesian Decision Theory

The Bayesian Decision Theory is a statistical approach regarding classification problems which has the characteristic of minimizing a loss function classifier:

$$\lambda(\alpha_i|\omega_j) * P(\omega_j|\overline{x})$$

where, λ is the loss of incorrectly assigning a class, α a set of possible actions, ω a set of categories and, and P stands for probability of the predicted class to be ω for a given entry

feature x. A detailed explanation can be found in [23].

2.2.5 Computer vision methods

Haar-like detector

The rotated haar-like feature detector is based on haar-like detector proposed by Du et all for face detection in 2006 [24]. It consists in selecting a spot on a given image and then proceeds to consider different rectangular regions with the same centroid. After that, the summation of the pixel intensities per region is performed and the difference between these sums is calculated. This will enable a categorization of the subsections generated.

2.3 Classification Algorithms

2.3.1 Decision Trees

The decision trees algorithm was introduced by Hausla and Swain in 1975 [25]. It functions as a regressor or a classifier through multi-stage decision logic. This tree consists in a root node, followed by non-terminal nodes and finally and terminal nodes. The size of the tree will have impact in the algorithm accuracy scoring and computational effort. Some advantages are its easy interpretation and little data preparation.

2.3.2 Support Vector Machines

Support vector machines is the final evolution of the supervised algorithm proposed by Vapnik et all in 1992 [26]. The goal of the SVM method is to construct a ”hyperplane” which do separate the data instances into k classes. The basic technique finds the smallest ”hypersphere” in the kernel space that contains all training instances, and then determines on which side of ”hypersphere” a test instance lies. The maximum margin hyper plane is the one that gives the greatest separation between the classes. The instances that are nearest to the maximum margin hyper plane are called support vectors. This classifier finds the maximum margin hyper plane, and it classifies all training instances correctly by separating them into correct classes through a hyper plane. One of the strengths of

SVMs is enabling different function to be chosen for kernel.

2.3.3 K-nearest neighbours

The concept of the instance-based nearest-neighbour algorithm was first introduced by Aha, Kibler, and Albert (1991) [27]. The KNN algorithm is an example of an instance based learner. In other word, all of the learning models are "instance based," as well, because they start with a set of instances as the initial training information. Sometimes more than one nearest neighbour is used, and the majority class of the closest k neighbours is assigned to the new instance. In this algorithm, distance between two data instances can be calculated in different ways. For continuous attributes, the euclidean distance is very popular while for categorical values matching coefficient is usually used.

2.3.4 Bagging

Bagging ensemble classifier is fast and ti can efficiently handle unbalanced and large databases with thousands of features. It has attracted much attentions due to its simple implementation and improving accuracy. This technique was proposed by Leo Breiman in 1994 [28]. Basically multiple version of the dataset are generated through boostrap method, and each replicate becomes a learning dataset. At last, it is taken the agregate value of the evaluating metric used. The bagging classifier can have different predictor as its estimator being the decision trees one of the most popular ones.

2.3.5 Random Forests

The random forests algorithm is the bagging algorithm with the base estimator as decision trees, this way calculating multiple decision trees results and taking into account the best one.

2.3.6 Gradient Boosting

Gradient boosting is a machine learning technique used in regression and classification tasks, among others. When the response variable is continuous, the Gradient Boosting

Regressor is used, whereas when it is a classification problem, the Gradient Boosting Classifier is chosen. In Gradient Boosting, each predictor tries to improve on its predecessor by reducing the errors. The goal is to minimize the loss function by adding weak learners using gradient descent. Oposite to the bagging algorithm which gives equal weigths to each model generated, the gradient boosting gives a weigth based on the model's scoring [29].

2.4 Evaluation Metrics

For this study only overlap based metrics were considered which are based on the values extracted from the confusion matrix - True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN). The positive values are respect to the Ground Truth (GT), while the negative values concern the Complementary of the Ground Truth (CGT). True or False it means if the pixel is correctly placed within its space (GT or CGT).
THe metrics used are presented next.

Recall or True Positive Rate - relative frequency of positive voxels in GT.

$$Recall = \frac{TP}{TP + FN}$$

Precision or Positive Predicted Value - relative frequency of voxels classified as positive.

$$Precision = \frac{TP}{TP + FP}$$

Jaccard Index is the proportion of TP on the group of False or Positive Values.

$$Jaccard = \frac{TP}{TP + FP + FN}$$

F1_Score is the harmonic mean of the precision and recall for a given class.

$$f_1 = 2 * \frac{Precision * Recall}{Precision + Recall}$$

Chapter 3

Vessel Detection

This chapters presents a literature review regarding vessel detection, consisting in the brief explanation of the methodology used by the authors and their results.

In 1978, McDonnel and Lewis worked with bands 4, 5, 6 and 7 of Landsat-2 MSS to detect some physical attributes of ships above 100 meters in length, using a linear threshold technique, considered the the highest pixel value of band 7 as the vessels's centers together with among other similar transformations regarding the others bands. They were capable of identifying the ships, and depending on the solar elevation angle and the sea state, to be capable of distinguish both the ship and the ship's wake [30].

In 1993, Burgess used the Green and NIR bands from SPOT-5 and Red and NIR bands from Landsat TM to build an automatic ship detector. The author start to masking out the land for each image, generating binary images - 0's for land and 1's for sea, using two different operations regarding the data source, followed by a low pass filter (41x41) to increase the mask's homogeneity and a threshold to regain binary images. Then, a high pass filter (41x41) is applied to both masked images in order to smooth the background due to current flows and sediments. To finally combine the two images, it was used image 1 where it is not non-zero and accept image 2 elsewhere. Once the preprocessing is done, there are 4 stages for object analysis, in order to properly select the ships. First, refinement by area, where objects above 150 square pixels are not considered. Secondly, refinement by shape analysis with the help of moment analysis. Thirdly, refinement by maximum intensity and lastly by spectral signature to separate small pieces of land with ships. The

author concluded that is was possible do detect the ships with some level of accuracy [31].

In 2003, Pegler et. al., used images provided by the satellite IKONOS to develop a marine recreation surveillance system to identify vessels. The author used an algorithms built by Subramanina and Gata [1998], directly on the images without knowing the Ground Truth placement/identification of the vessels, with the intent to understand if it possible identify the vessels on a spatio-spectral template, as the first step of their research. From 11 identified patches, 9 were targeted incorrectly and another 11 were false targets [32].

In 2008, Corbane et. al., built a three-step algorithm for object detection on SPOT-5 optical images of 5 m resolution. First, segmentation (predetection of ship patterns), considering potential ships as bright pixels and the contrast, the sea, as noir-like background, using morphological opening operation together with region growing to connect possible single pixel or very small regions, that alone, are not vessels. Second, using feature extraction, such as large length to width ratio, regular and compact shape, amongst others, to identify the objects that can most likely be a vessels. Third, using neural-network, it was possible to identify which objects were ships. From 28 extracted features, only 8 were considered for the NN. The detection rate was around 60% and 5700% was the corresponding false alarm rate. Reducing the false alarm rate would also reduce the detection rate [33].

In 2009, Wu et. al. used SWIR-1 and 2 from Landsat-5 TM to build a ship detection algorithm for turbid waters. Only one Landsat TM image was used, and only the bands with 30 meter resolution (1-5, 7). Some transformations and projections were applied in order to correct and clean the image. Through visits to Poyang Lake, the authors had a perception where the ships used to stay. When applying a PCA, the first 3 components revealed 180 similar shape objects which matched the size and linear arrangements of the vessels previously observed, which of these 81 were randomly chosen for further analysis. With the reflectance contrast, the authors choose the brightest pixel (highest contrast ratio) to represent the ship object. Finally, ANOVA, normalization and ANCOVA statistical analysis were applyed to the six TM bands in study. While water turbidity was noticeable in the first four bands, the vessels were better identified in bands 5 and 7. Measuring the significance of the relation between the reflectance of ships and that of their adjacent water, at $\alpha < 0.001$, the first 3 bands presented coefficient of determination (R^2) around 0.90,

while the other three bands below 0.60. This way, the author concluded that the ships did not occupied a whole pixel, being the objects mixtures of ships and surrounding waters [34].

In 2011, Huang et. al., have built a ship detection algorithm based on the images from SPOT-5. The authors have divided their method into 3 main groups - Detection of statistical textures, Candidate extraction and Ship verification. In the first one, it is used a texture descriptor to extract features based on local multiple patterns (derived from LMP operator [35]) to build histograms, considering large values of bins as the *quantitized texture features of sea* and small values for ship candidates. The descriptor has been quantified into 100 parts and the histograms were normalized. As for the ship candidates extraction, a confidence map with values from the LMP operator was built, and if the pixel confidence value was above 0.9985, then it would be consider a ship candidate, otherwise, sea, building a binary image. Finally, in order to remove the False Negatives, it was considered ship shape properties such as length width ratio (above 1.35) and ship area (75 m^2). Their method has scored 0.586 of Precision and 0.953 of Recall [36].

In 2012, Fukun et. al., built a detection algorithm based on images also provided by SPOT-5. Attention of Candidate Regions (ACR) is the first stage, to identify the patches with more potential to be ships. This was done using Yu's pulsed cosine transform (PCT) method due to its good performance and fast computational speed. The drawback of Yu's is that it cannot discriminate the ships from other selected salient objects. The Local Context Facilitation (LCF) is the second step. Here, the authors use a neighborhood similarity-based method together with threshold method, to analyse the ships surrounding area, which it has taken as an assumption of having similar texture and intensity. The final stage is called Appearance Identification (AI) which results in the application of SVM radial basis function kernel to the 128 features calculated with the usage of the SIFT descriptor on the patches. The authors, in order to compared results, apply 4 combinations of the algorithm (ACR, ACR+AI, ACR+LCF, their model), for two diferent testing sets (an "easy" one and a "difficult" one), for 2 scoring metrics (Recall and Precision). Their model, had the lowest scores for both measures, regardless of the testing set. The Recall was always higher on the easy testing set, whilst the Precision was always higher in the difficult testing set. [37]

In 2013, satellite images of WorldView-1 and WorldView-2 were used by Máttyus to generate a ship detector (base on a face detector built by Viola and Jones in 2004). This method consists four stages of classification. Watermask with a binary mask. The sliding window will take into account 16 considered directions by the author, which enable the separations of potential ships with land areas on sea. The third stage is the detection grouping which will group vessels with the same position, size and heading, the group with most members reach the final stage. This last stage consists in applying contours for a better measurements extraction, based on the Otsu's method. It was concluded that this method can detect vessels with quality if the watermask is accurate [38].

in 2014, Panagiotis and George, developed an detection algorithm working with images from ASTER VNIR, using the band Green, Red and NIR. Firstly, destriping and reflectance value estimation. Then, separate the sea from its targets using quad tree decomposition. Third, bounding box on the patches to extract ship region, shape and reflectance as well as other ship measures. Lastly, integrate other relative geographical information to predict. Due to 15 meter pixel, the shape of the ships was not well identified. The band Red provided the highest reflectance for the smallest objects. On othe other hand, the maximum size object presented the highest reflectance in band NIR [39].
Also 2014, Yang et. al. used images from Google Earth and Spot-5 to build a ship detection algorithm. The first step consists in the removal of no-candidate regions based on histogram value frequencies and threshold values based on training data. Then ship candidate selection with intensity and texture analysis, based on a self create function. Thirdly, false alarm elimination using the objects shape characteristics such as compactness and length-width ratio. The author made a comparative analysis between quiet sea, texture sea and clutter sea which presented high to low scoring values for precision and recall, respectively. Without step 3, the algorithm ends up detecting lots of false negative objects [40].

in 2015, Tang et. al. used deep neural networks together with extreme learning machine (ELM) to create a ship detection algorithm with images from SPOT-5. It starts with image enchancement for ship extraction and background suppression. Then Sea-land segmentation using otsu's method together with gaussian model, median filter and morphological

dilations and erosion operator. As for the ship location criteria uses object's are, major-minor-axis ratio and compactness to distinguish ships for other objects. In the final stage, deep neural networks are used for object feature extraction on low and high frequencies which serve as an input for the extreme learning machine algorithm fro classification and decision making. For comparison purposes support vector machine algorithm was also used (replacing deep neural networks step). ELM has shown to be crucial for faster feature extraction and higher learning efficiency, presenting the lowest error or missing ratio, and the best accuracy score [40].

Chapter 4

Exploratory Data Analysis

4.1 Data Wrangling

This chapter focuses on analysing the data provided by the Automatic Information System (AIS), extracted from CEiiA's API Ingestor. The main goal is to understand the vessel features and how they can be used to classify the vessels.

The AIS data contains 22 features (presented in 1.1). The graphical representation in figure 4.1 contains 2 barplots for the qualitative variables (*navstat* and *type* and a figure (4.2) containing pairplots with histograms for the quantitative variables.

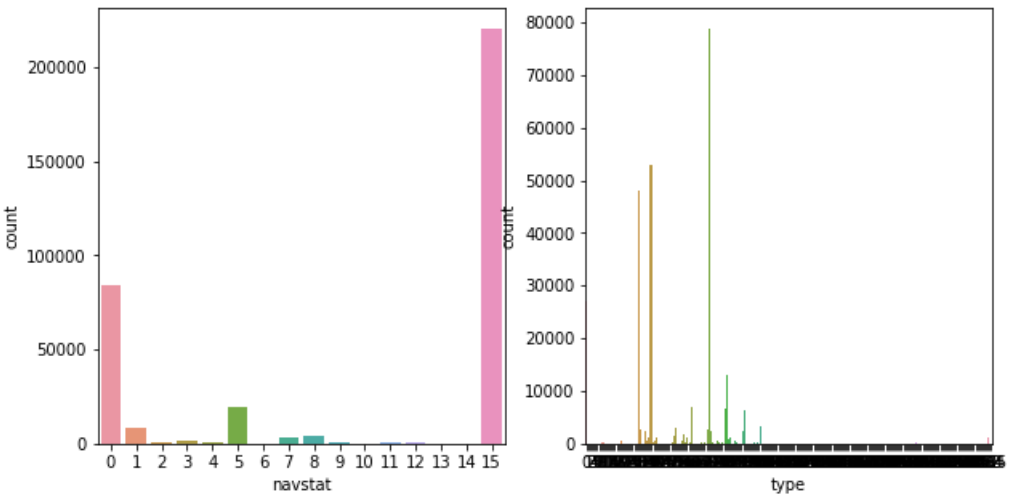


FIGURE 4.1: Barplot for AIS qualilative variables navstat and type

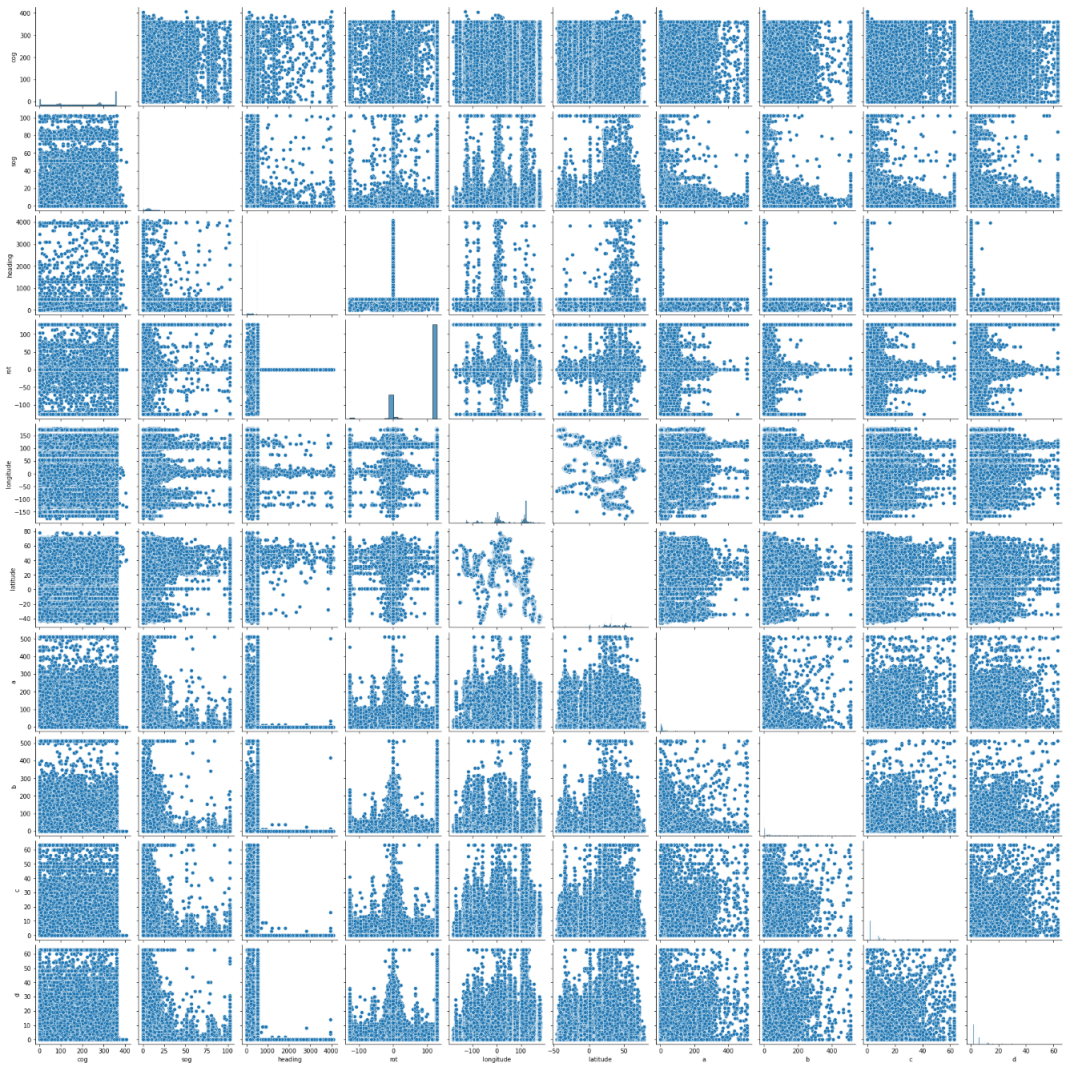
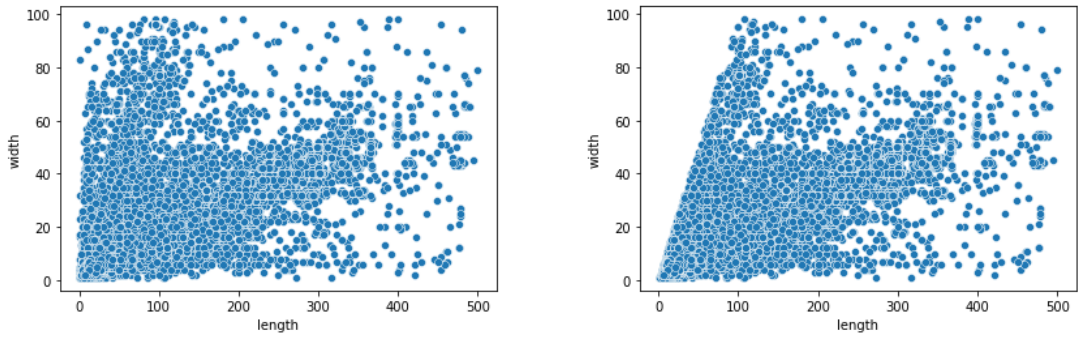


FIGURE 4.2: Pairplot for AIS quantitative variables

Length and Width

The data does not come with a length or width variable, but these can be easily obtained through the variables *A, B, C, D*. The *length* is equal to the sum of A and B and the *width* is equal to the sum of C and D. The product of these two new variables generated a third variable - the *raw_area*. Moreover, the *length* of the vessel cannot exceed 500 meters nor the *width* can exceed 100 meters. Also, the *length* is always higher than the *width*. Taking into account these criteria, two scatter plots of the data before and after these restrictions are imposed, are presented in Figure 4.3.



(A) Before - without restrictions

(B) After - with restrictions

FIGURE 4.3: Scatter plot

Type

The *type* of vessel is the target feature. However, this feature comes in various "subtypes", which makes the data a lot less grouped. That being said, it was necessary to group all *subtypes* into their respective main type. The new variable generated is called *real_type*. From now on, the variable *type* will be referred to as the *subtypes* of the variable *real_type*. The figure 4.4 contains a barplot for vessels per *real_type*.

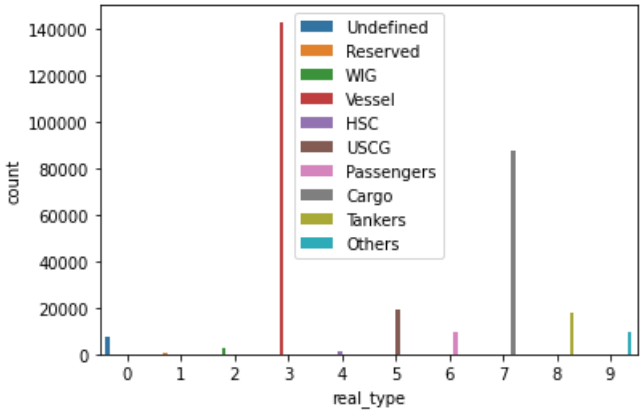
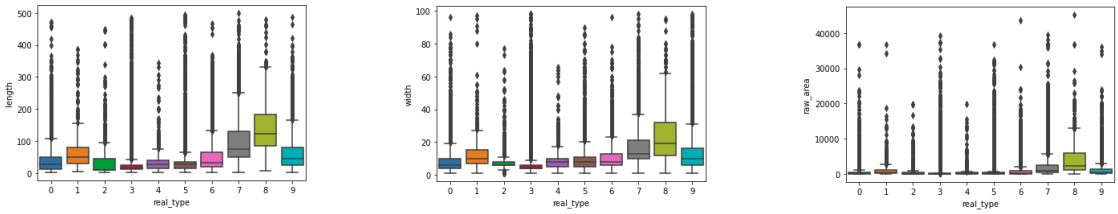


FIGURE 4.4: Barplot - Vessels per real_type



(A) Length (B) Width (C) Raw_Area

FIGURE 4.5: Boxplots for main features on real_type

Among the real_type, the data is very unbalanced, with more than two thirds corresponding to only two types - 3 and 7. The length has a better data division than the width, while the raw_area it is not capable of distinguishing the data so well.

The subtypes

The table 4.1 contains the *real_type*, number of vessels and number of subtypes per each type.

Type	Def	No of vessels	No of subtypes
All	-	298439	215
0	Undefined	7333	0
1	Reserved for future	711	87
2	WIG - wings-in-ground	2448	50
3	Vessel - different variables	142856	11
4	HSC - high-speed-craft	1510	11
5	USCG - vessels in US waters	18970	11
6	Passengers ships	9363	11
7	Cargo ships	87802	11
8	Tanker(s)	18067	11
9	Other types of ships	9379	11

TABLE 4.1: Number of vessels and subtypes per type feature

More than half of the total of subtypes (137) are concentrated in little more than 1% (3159) of the total number of vessels. Almost half of the total number of vessels (142856) is concentrated in a single type

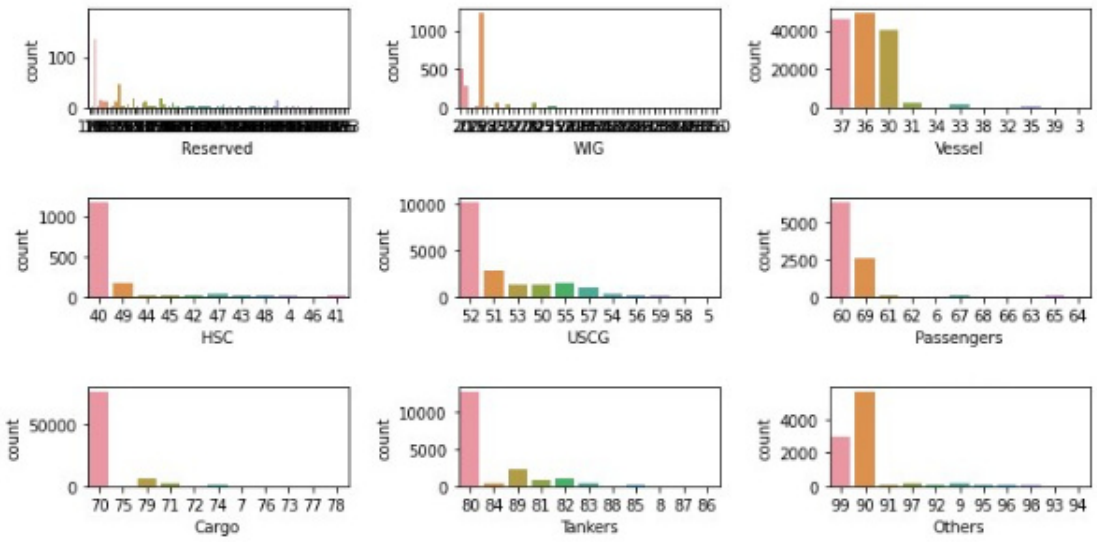


FIGURE 4.6: Barplot of subtypes per real_type

Observing Figure 4.6, it is possible to notice how unbalanced the data is withing each real_type, being the majority of data located at a maximum of three subtypes. Also, several of these subtypes are classified as undefined within its class.

It was considered important to visually understand, by the help of a scatter plot, how these subtypes would position themselves in a two dimensional space, through the combination of pairs of the main features, i.e. Length VS Width, Length VS Raw_Area and Width VS Raw_Area.

These scatterplots are presente in Figures 4.7, 4.8, 4.9, 4.10 and 4.11.

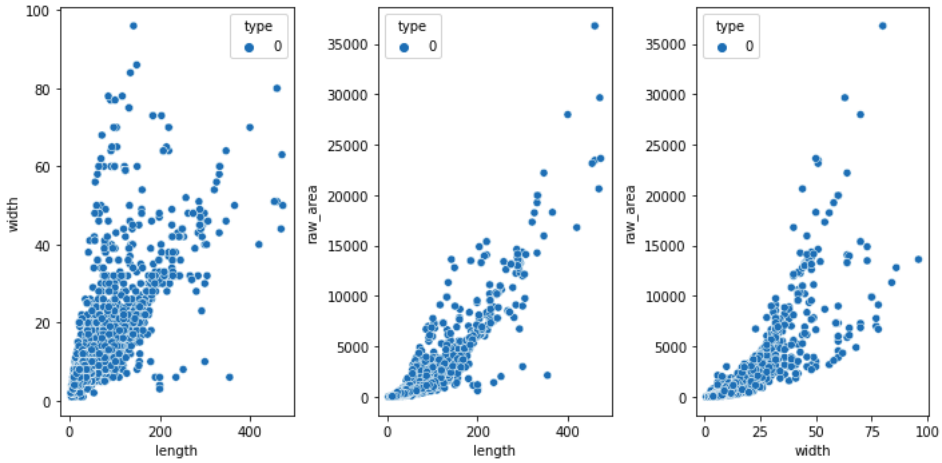


FIGURE 4.7: Scatterplot subtypes per real_type - 0.

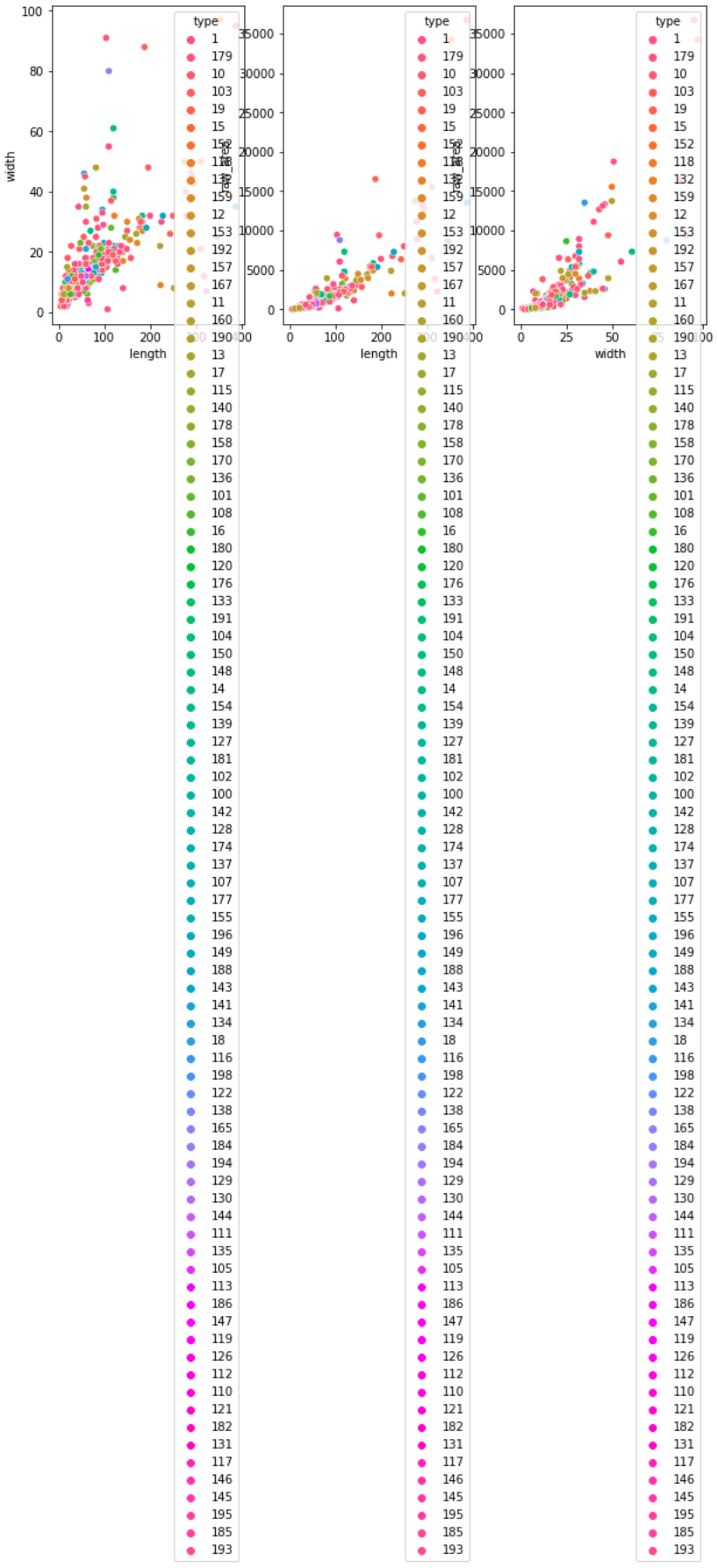


FIGURE 4.8: Scatterplot subtypes per real_type - 1.

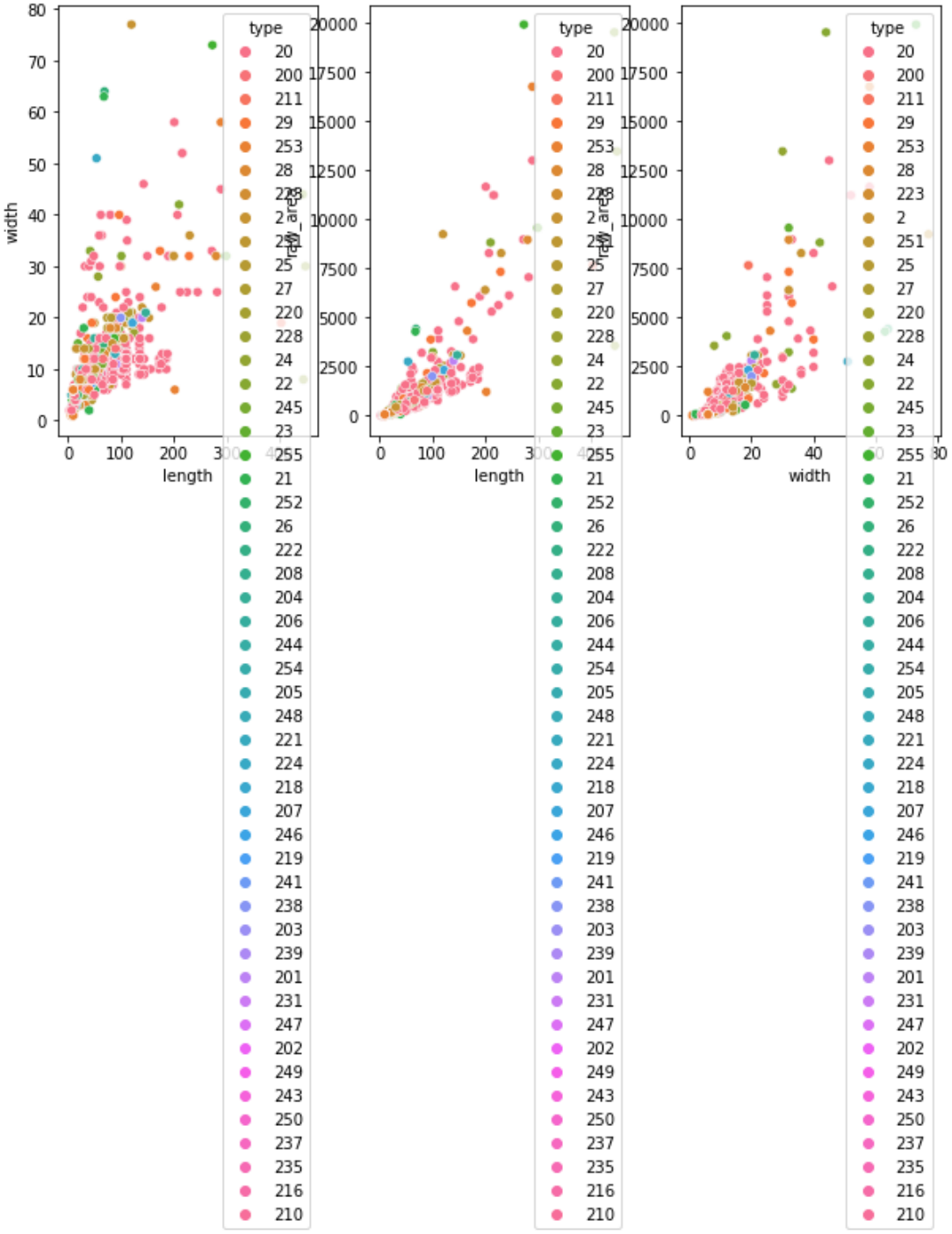


FIGURE 4.9: Scatterplot subtypes per real_type - 2.

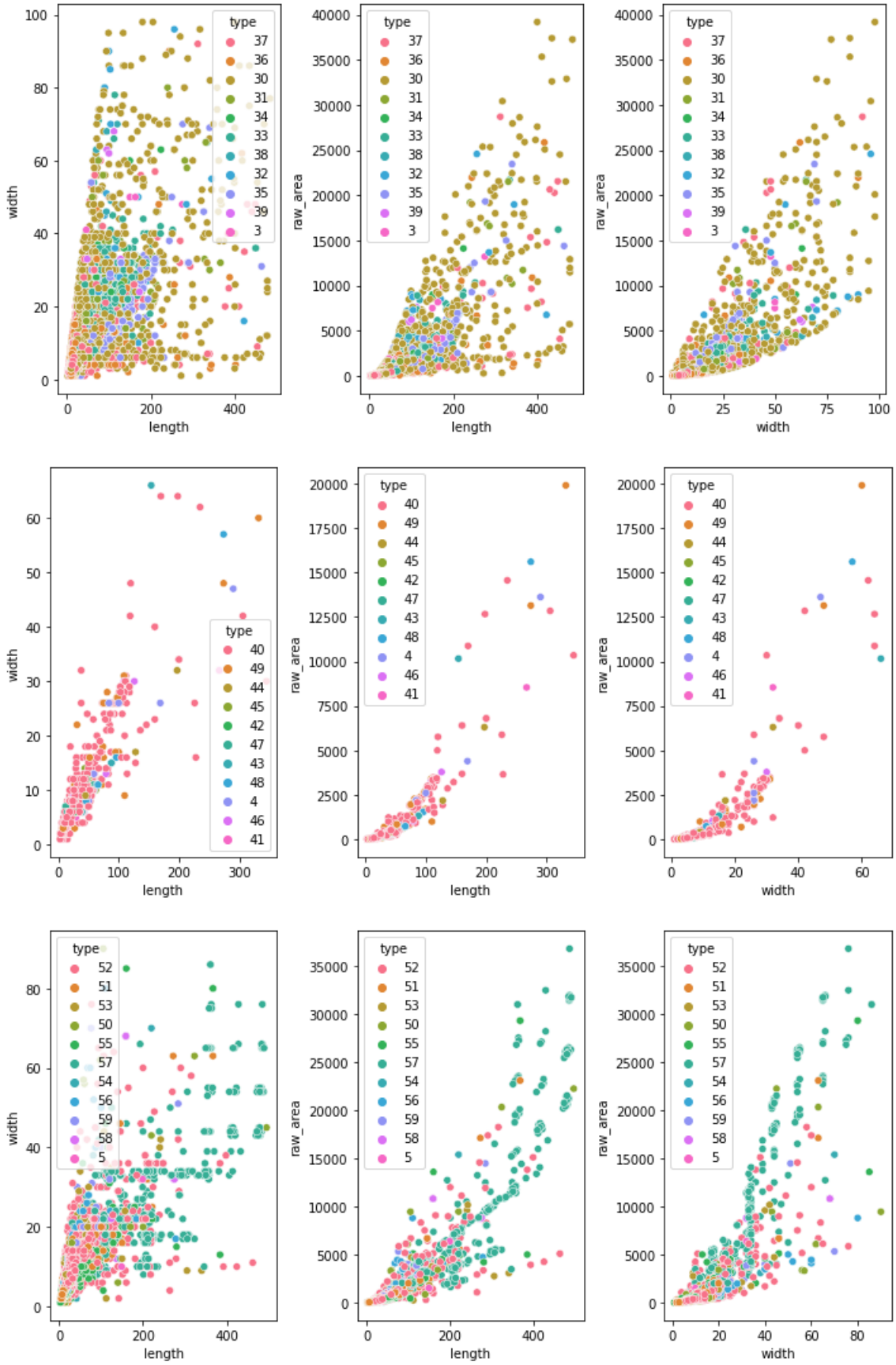


FIGURE 4.10: Scatterplot subtypes per real_type - 3, 4, 5.

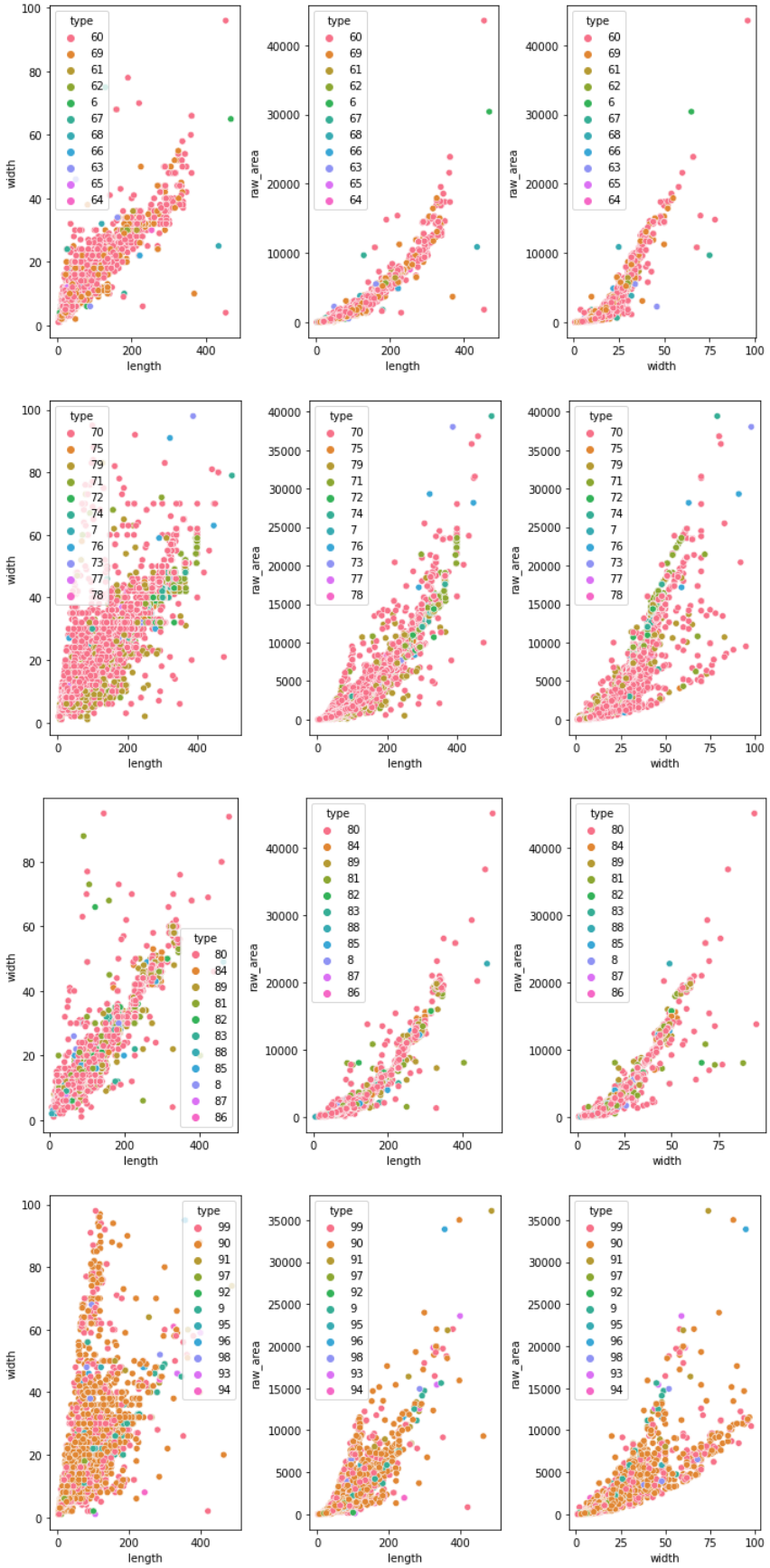


FIGURE 4.11: Scatterplot subtypes per real_type - 6, 7, 8, 9.

The subtypes do not seem to group themselves among the three pairs of dimensions, for any of the real types.

Since the data is very unbalanced within each variable real_type, a visual analysis for the main subtypes was done (referred as *chosen_subtypes*) through the help of a barplot (Figure 4.12) and boxplots (Figure 4.13) for the pair of vessels measure.

These chosen subtypes were selected through their high relative frequency within each real type.

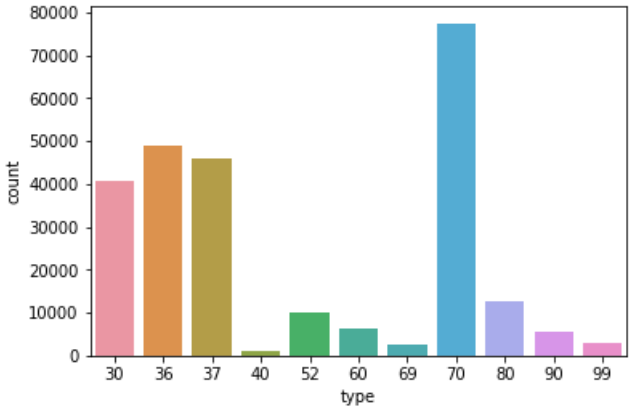
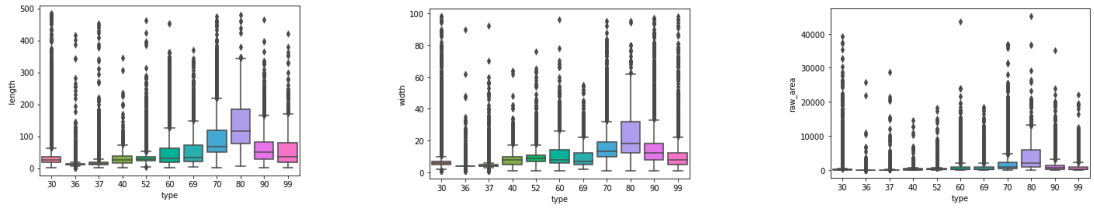


FIGURE 4.12: Barplot - Vessels per chosen_subtype.



(A) Length (B) Width (C) Raw_Area

FIGURE 4.13: Boxplots for chosen_subtypes.

Similarly to the real_type classes, the length enables a much better differentiation of the data than the width, and the raw area is almost incapable of that.

4.2 Cluster Creation

The purpose of choosing the main subtypes was to generate a set of clusters than could be compared with the sets of clusters directly from real type classes. These, were generated based on the real type and chosen types boxplots (Figures 4.5 and 4.13). The real_type gave birth to three sets (169, 7, 8, 2345), (169, 78, 2345) and (1679, 8, 2345) while the chosen

types gave birth to two (30,36,37,40,52 — 60,69,99 — 70,90 — 80) and (30,36,47,40,52 — 60,69,90,99 — 70 — 80). Scatterplots (Figures 4.14 and 4.16) and Boxplots (Figures 4.15 and 4.17) for these sets of clusters are presented in the following pages.

Since class 7 contains a large percentage of vessels, the cluster this class is associate to will have a different behabiour on the scatterplots (Figure 4.14. Among the real_type classes, the set (1679, 8, 2345) presents a better differentiation than the other two sets (Figure 4.15).

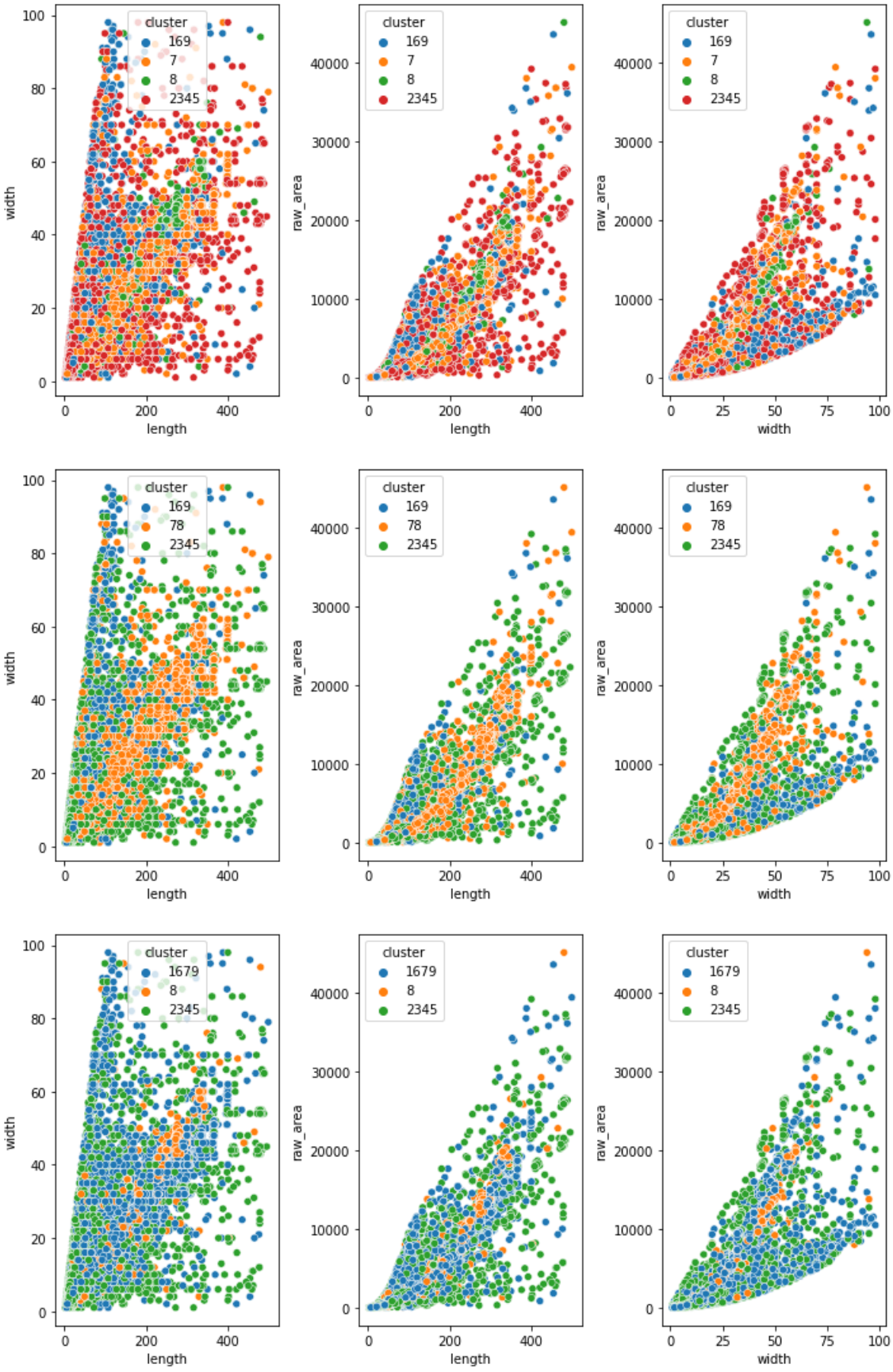


FIGURE 4.14: Sactterplot: cluster of real types.

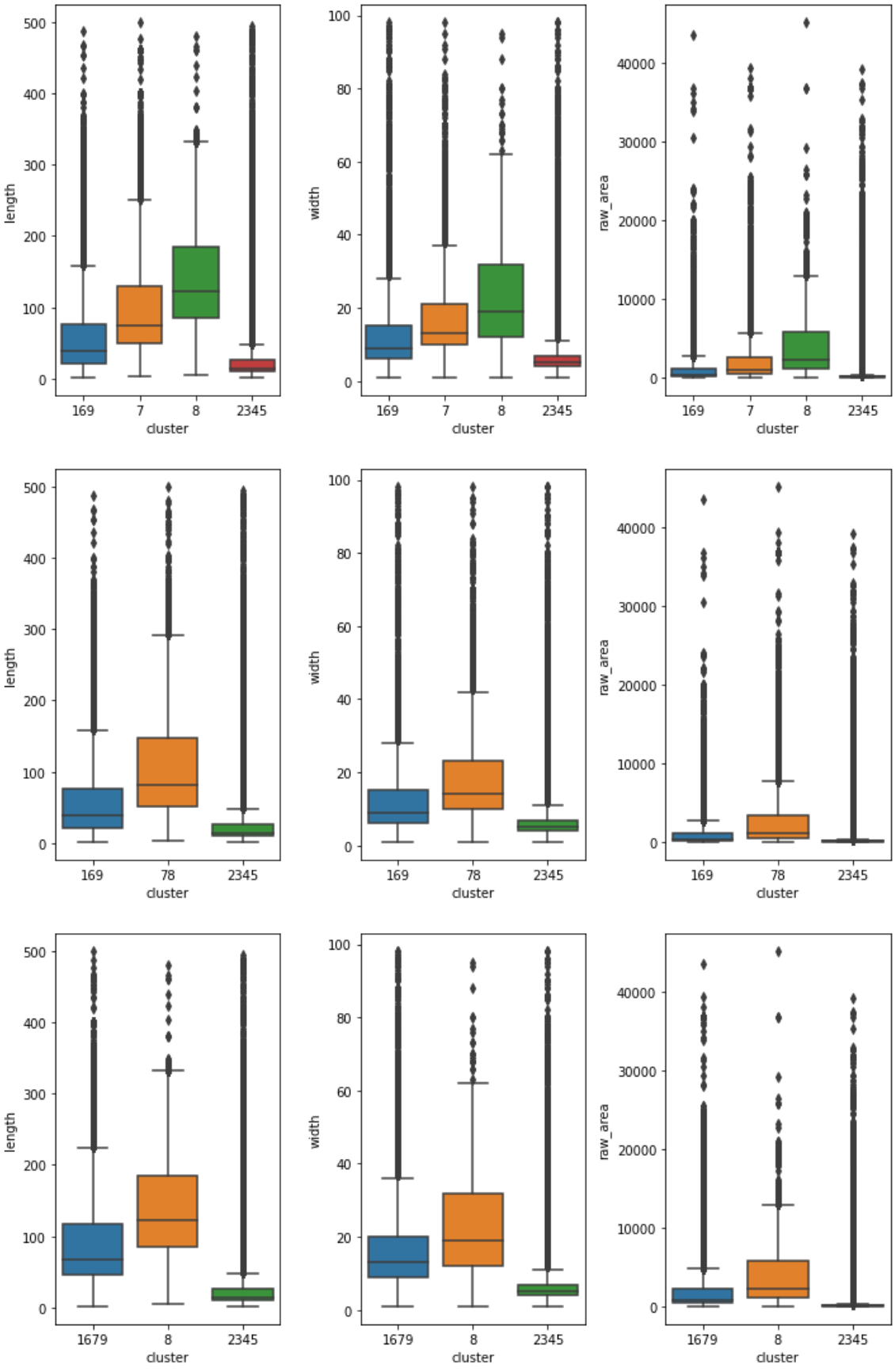


FIGURE 4.15: Boxplot: cluster of real types.

Once again, the class 7 will affect how visually the clusters will be displayed on the plot (Figure 4.16. As for the boxplots on Figure 4.17, these seem very similar, but less capable of differentiate the date when compared to the three later mentioned.

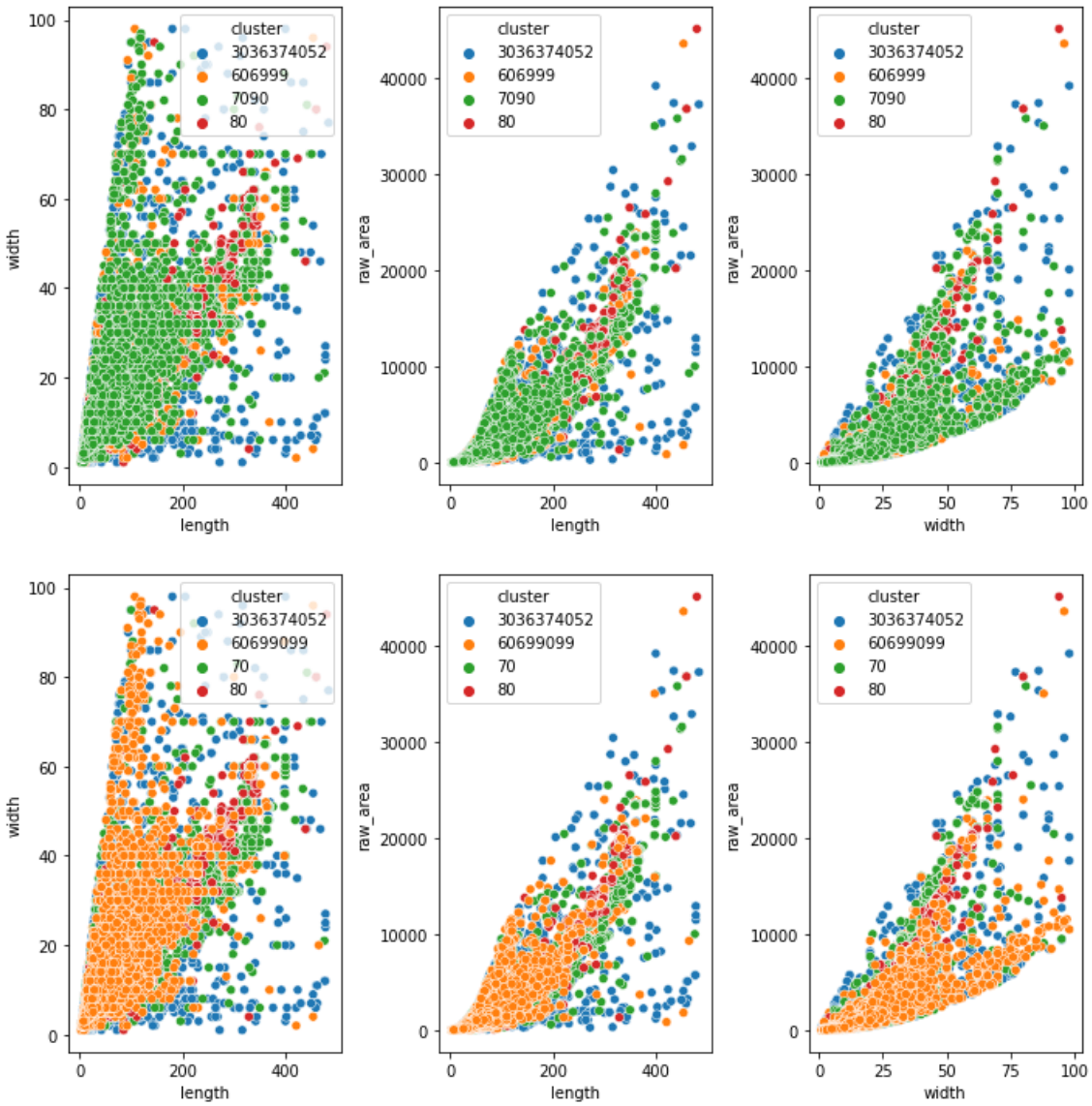


FIGURE 4.16: Scatterplot: cluster of chosen subtytpes.

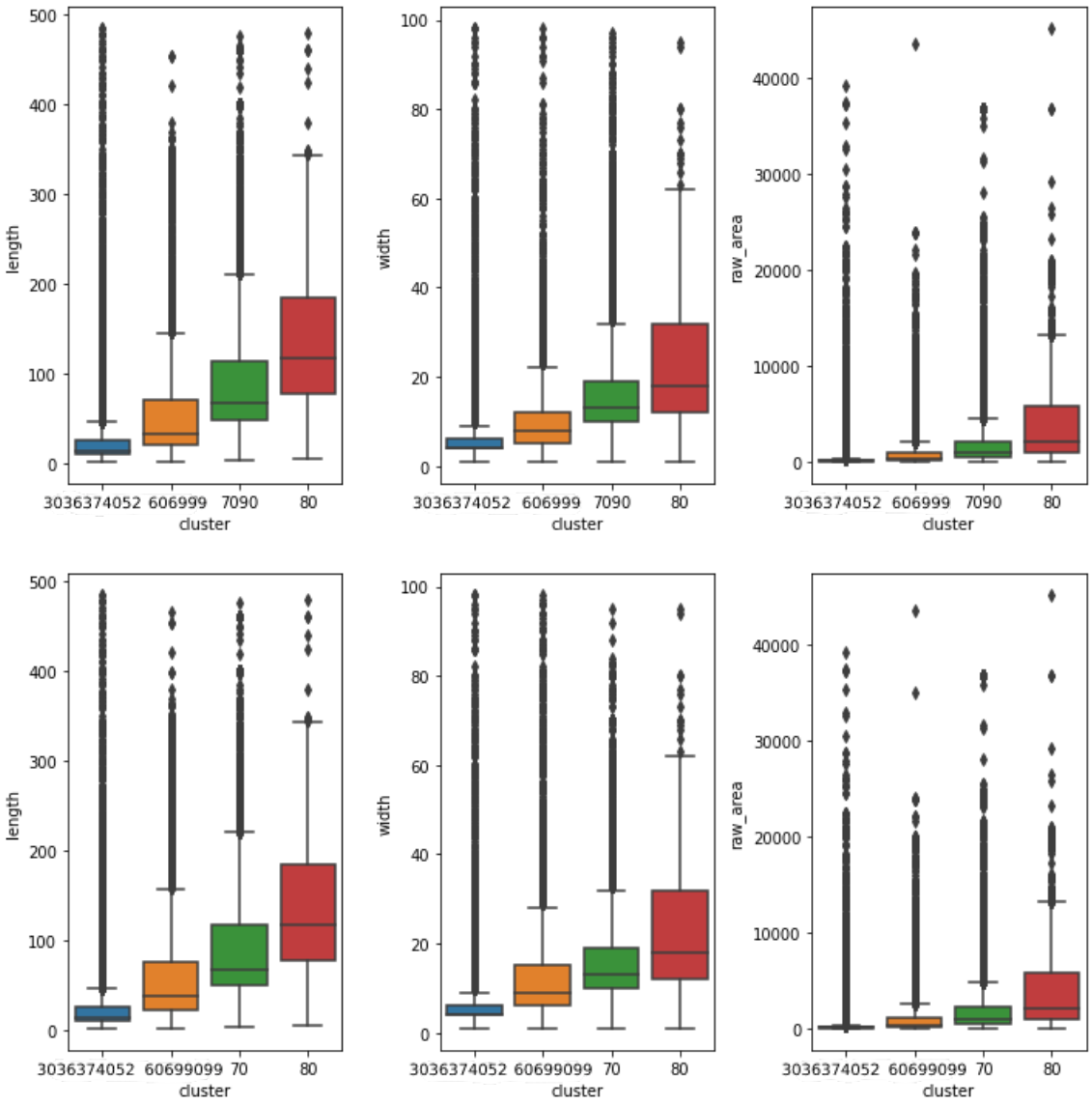


FIGURE 4.17: Boxplot: cluster of chosen subtypes.

4.3 Cluster Validation

Now that there are five sets of cluster, one must be chosen to become the response variable for classification purposes. That being said, a monte carlo simulation [41] with linear discriminant analysis (LDA) [42] was used to understand how the observations would be correctly put into their respective clusters. The accuracy scoring was chosen as the evaluating metric.

Three sets of simulations were done. Two of them performed in balanced data through undersampling, one of which uses LDA with one component and the other LDA with

two components. As for the last set, which is highly unbalanced data, some simulations among the lower sample size would only capture observation from two or even one cluster in extreme cases. For this reason, only LDA with one component was used.

For the two unbalanced simulation, sample sized from 10 to 10.000 observation in steps of 100 were used, resulting in 100 iterations. For the unbalanced data, it were used sample sizes from 10 to 100.000 in steps of 1000. The Figure 4.18 contains each of the three simulation described, where the x-axis corresponds to the number of iteration and y-axis to the accuracy scoring. Table 4.2 presents the scoring values on the last iteration for each simulation.

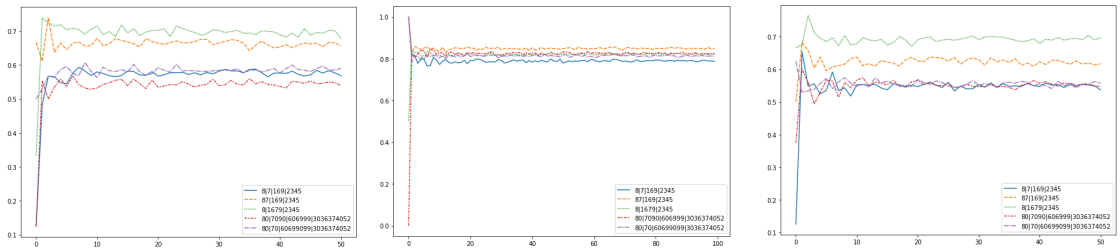


FIGURE 4.18: Simulation: Balanced LDA 1 — Balanced LDA 2 — Unbalanced LDA 1

Set of clusters	**Balanced - LDA 1**	**Balanced - LDA 2**	**Unbalanced - LDA 1**
8 — 7 — 169 — 2345	0.537088	0.568432	0.786739
87 — 169 — 2345	0.616883	0.656177	0.848298
8 — 1679 — 2345	0.696978	0.676823	0.819665
80 — 70,90 — 60,69,99 — 30,36,37,40,52	0.547328	0.53996	0.824159
80 — 70— 60,69,90,99 — 30,36,37,40,52	0.557942	0.590035	0.810928

TABLE 4.2: Simulation - Results.

Regarding the unbalanced data, the values are very close throughout the simulation and most sets keep the same positioning. As for the balanced data, the second and third set stand out reaching the 60% accuracy mark. Since the data is typically balanced and scaled, for classification purposes, the third set formed by three clusters (8), (1679) and (2345) performed better, thus being the chosen one.

Chapter 5

Methodology

The CEiiA's API ingestor contains, at present (Set. 2022), around 18.000 images. These images are a result of a pre-processing method that consisted in the cropping on a sentinel image, based on the coordinates of a vessel provide by the AIS. This results in 1000x1000 pixel images, with each pixel corresponding to 10x10 meters. Also, at the point, the images only contain 3 bands R, G e B).
It is important to mention that several programs were created to make it possible to work with the data. One to extract the data from the CEiiA's API ingestor. A second one to allocate the images that came in JP2000 format to a single directory. A third to convert a set of JP2000 images from a given directory to JPEG. A fourth to cut the images based on its center with a given length and directory. A fifth to generate a CSV file based on a set of images in a given directory with selected features. It was also developed a program that can misalign the bands of each image from a given directory. The later was one of the ESA's requests to apply on the training images in order to build a model with a better prediction, based on the misalignment error that may occur on the satellite. However, this has not been applied in this work due to computational issues.

5.1 Algorithms

The current section describes the steps and assumptions that lead to the creation of the segmentation algorithms.

5.1.1 Image Analysis

As previously mentioned, the original images have a size of 1000 pixels in height and width, with a pixel measurement of 10x10 meters. This makes it very difficult to identify vessels because this relative area on the images is very small (only a few pixels). The images were cropped based on its center to 512 pixels in height and width, making the vessels relative area higher, enabling a potentially better detection and segmentation.

In order to build the segmentation algorithms, a sample of 12 very different images were chosen, as shown in Figure 5.1. This way, it extends the potential of the model making it more general.

FIGURE 5.1: The 12 Selected Original Images.

Then, single band and multiple band displays were done in order to help finding the best band combination in which the model will be based on. By visual analysis, the single band RED was selected.
Each column meaning:

- oi - original image
- oi_gray - original image grayscaled (average)
- r - RED band
- g - GREEN band
- b - BLUE band
- rg_gray - RED and GREEN band combination grayscaled

- rb_gray - RED and BLUE band combination grayscaled
- gb_gray - GREEN and BLUE band combination grayscaled

For a much better visualization, the 12 images were divided into sets of 4 and displayed per page, on landscape mode, in Figures 5.2, 5.3 and 5.4 . It is also important to mention that due to unknown reasons, some random images are displayed with a colour filter, instead of appearing as gray-scaled images.

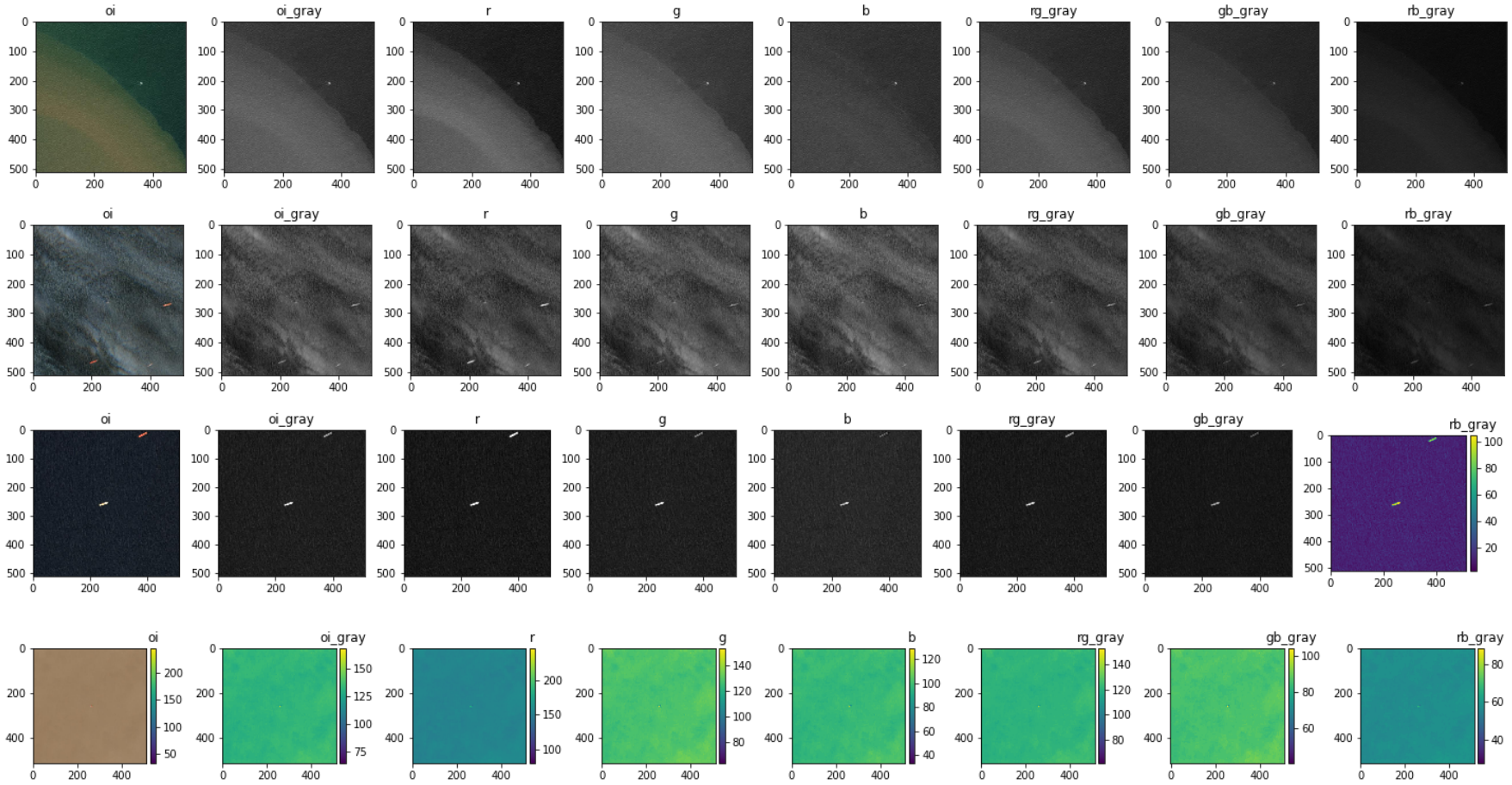


FIGURE 5.2: Centered 512x512 px images - 1 to 4

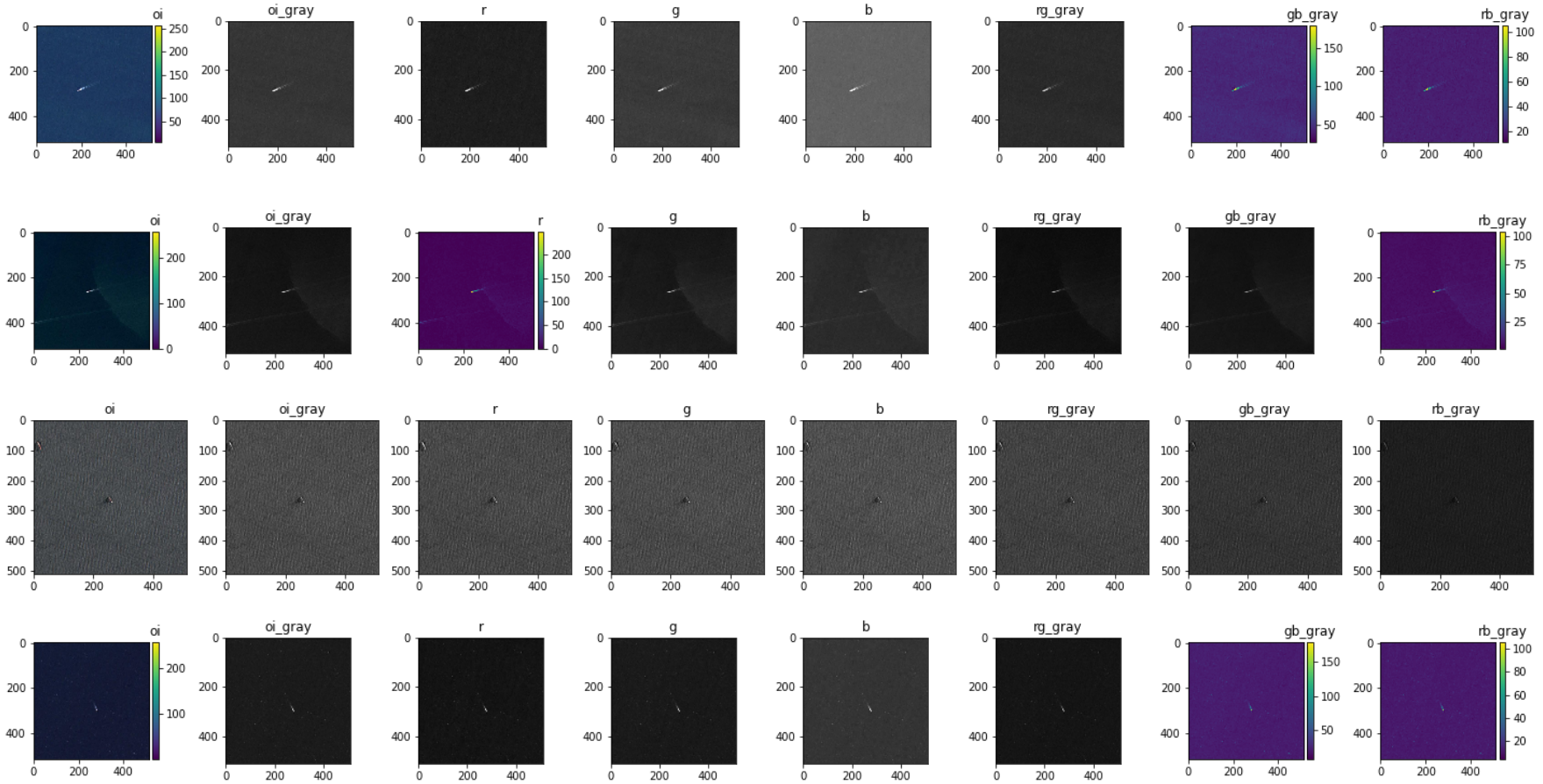


FIGURE 5.3: Centered 512x512 px images - 5 to 8

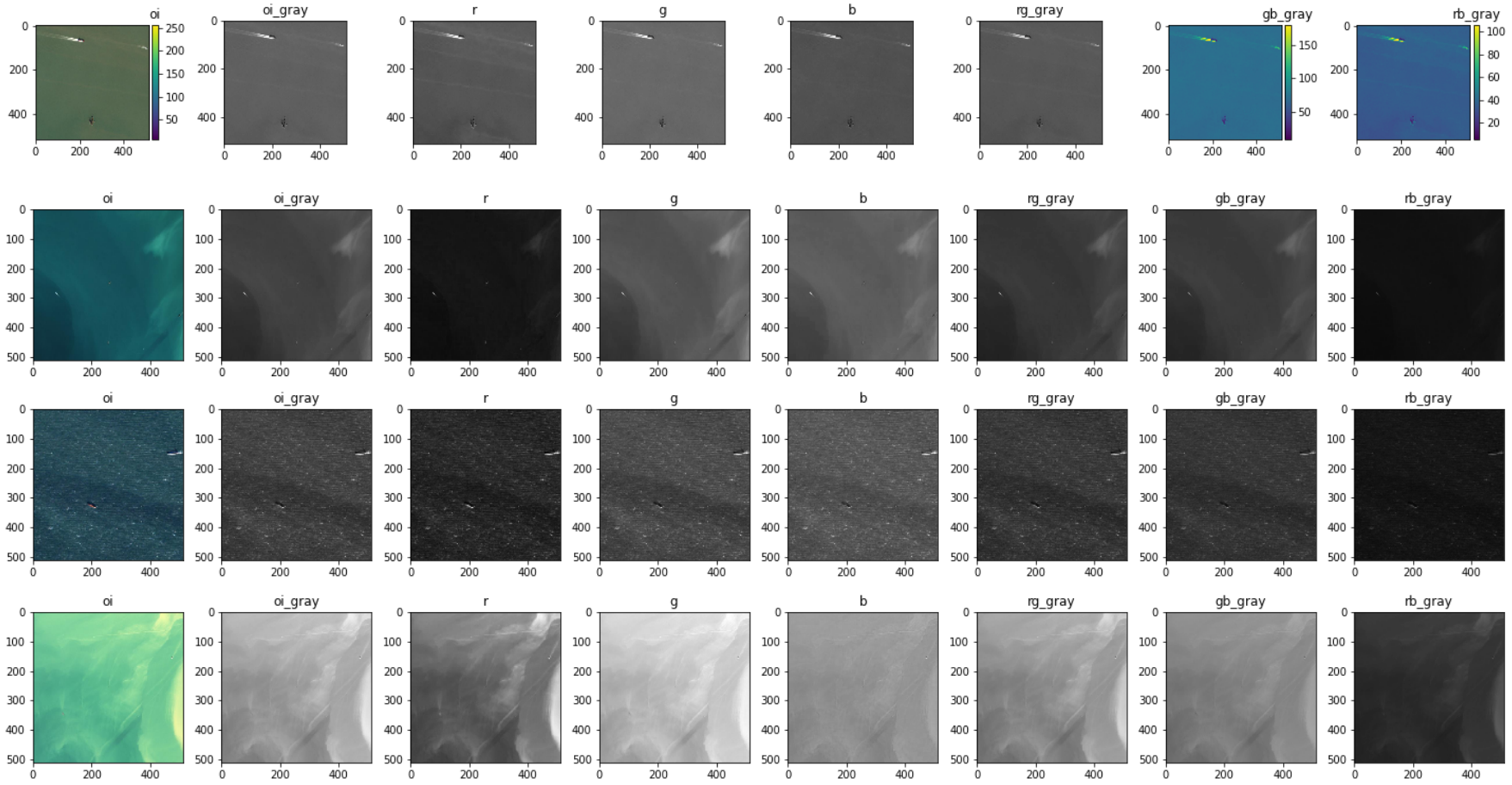


FIGURE 5.4: Centered 512x512 px images - 9 to 12

5.1.2 Region Growing

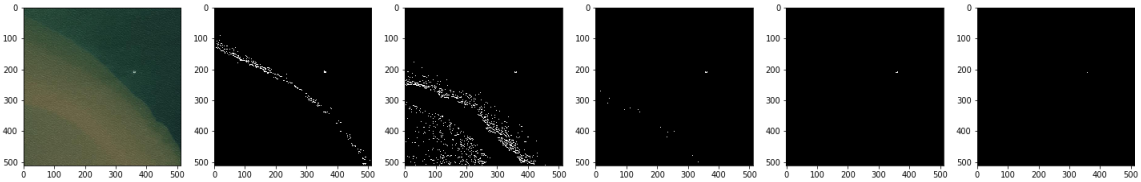
Region growing consists in marking, manually or automatically, a certain number of seeds on the image, and based on a similarity criteria, the adjacent pixels join the pixel-seed, until no more pixels fulfil the criteria, resulting on a object segmented image.

However, it is not feasible to manually select the seeds for big sample sizes, in other words, it takes a lot of time to check, register and compute these seeds on the correct position of the target object if the sample size is big. With the usage of six different threshold values - 80, 100, 120, 140 and 160 - five binary images were generated, visually enabling how each image would react regarding each threshold value, as shown in the two examples presented in Figure 5.5. The visual representation of the rest of the images can be seen in the Appendix - A.1.

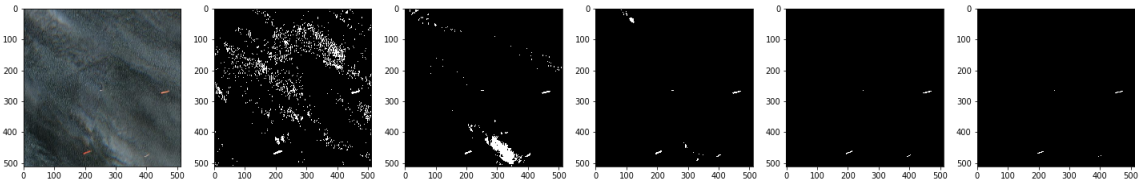
On one hand, it is noticeable that the binary objects resembled a more accurate shape (compared to its respectively RGB) with a lowest threshold (BILTV). On the other hand, on the binary image with the highest threshold value (BIHTV) it is possible to detect the positing of the vessels, minimizing the number of wrongly detected objects.

In order to attain the best ROI, bounding boxes over all the objects presented in both binary (BILTV and BIHTV) were generated, and their object features - such as object centroid coordinates, length and width of the object, area of the bounding box amongst others were extracted to a csv file (one csv file for each binary image).

Then, the objects on BILTV were filtered by the equivalent objects on the BIHTV. This was done through a similarity criteria using some of the object features, choosing the object that minimizes the sum of the absolute value of the difference per each feature. In some cases, small objects distinguished on BIHTV may belong to the same vessel, thus leading to the same object chosen on BILTV. Due to this particular situation, the removal of duplicated objects was implemented on the algorithm.

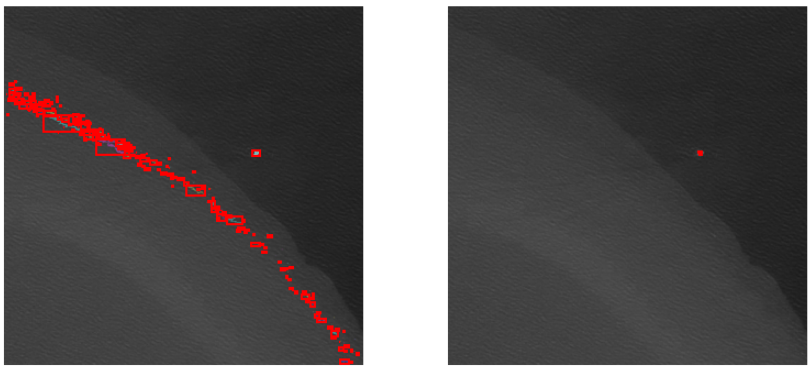


(A) Example 1



(B) Example 2

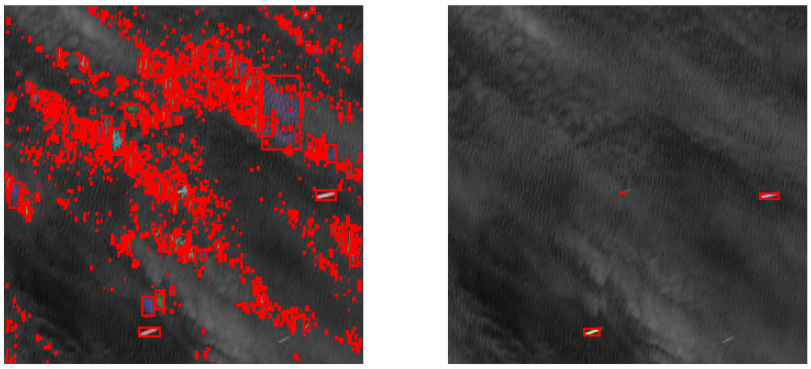
FIGURE 5.5: Region Growing seed images for 2 images (RGB, left), obtained by global thresholding with different threshold values (80, 100, 120, 140, 160) binary images from left to right



(A) Lowest Threshold as Seed

(B) Highest Threshold as Seed

FIGURE 5.6: Region Growing - Bounding Box - Example 1



(A) Lowest Threshold as Seed

(B) Highest Threshold as Seed

FIGURE 5.7: Region Growing - Bounding Box - Example 2

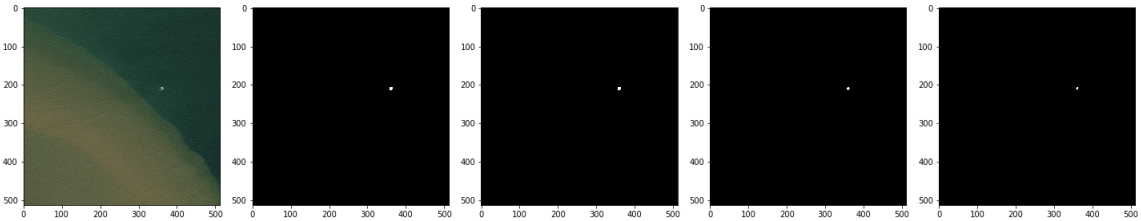
5.1.3 Edges Growing

This Edges Growing algorithm utilizes edge detection together with a linear thresholding and the morphological operation closing in order to segment the objects.

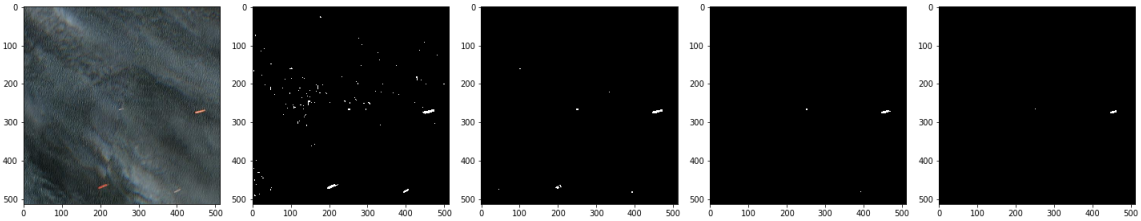
Once the edge detection is applied, instead of using absolute threshold values in order to generate the binary images, it was used 4 relative values - $\frac{3}{8}$, $\frac{4}{8}$, $\frac{5}{8}$ and $\frac{6}{8}$ based on the range of values (maximum pixel value - minimum pixel value) of each image. Just like with the region growing, values close to the extreme values were not taken into consideration.

After generating four binary images for each original image, they are intersected in a unique image (equivalent to BIHTV in region growing), and those objects are considered the vessels. From this point forward, the procedure is analogous to the one described on 5.1.2.

The figure 5.8 presents 4 binary images for 2 example images. The same images have their BIHTV and BILTV equivalents presented on Figures 5.9 and 5.10, respectively.

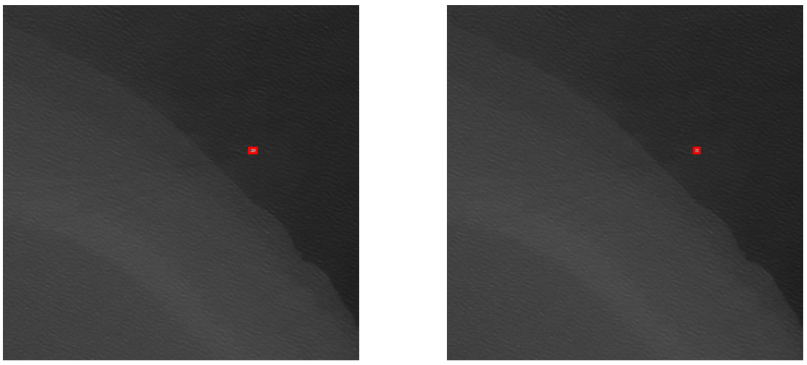


(A) Example 1



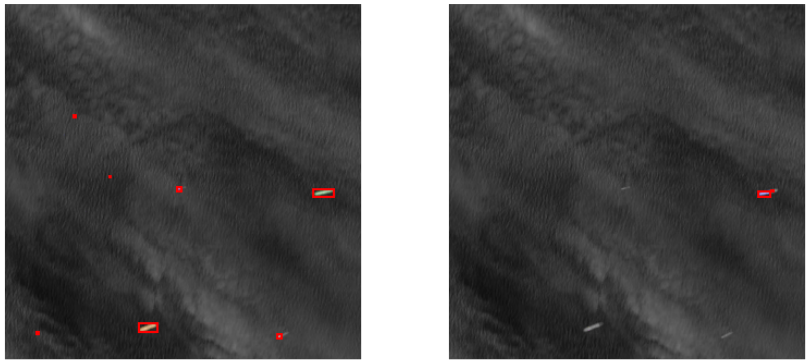
(B) Example 2

FIGURE 5.8: Edges Growing seed images for 2 images (RGB, left), obtained by global thresholding with different threshold values ($\frac{3}{8}$, $\frac{4}{8}$, $\frac{5}{8}$ and $\frac{6}{8}$) binary images from left to right



(A) Lowest Threshold (B) Intersection Binary

FIGURE 5.9: Edges Growing - Bounding Box - Example 1



(A) Lowest Threshold (B) Intersection Binary

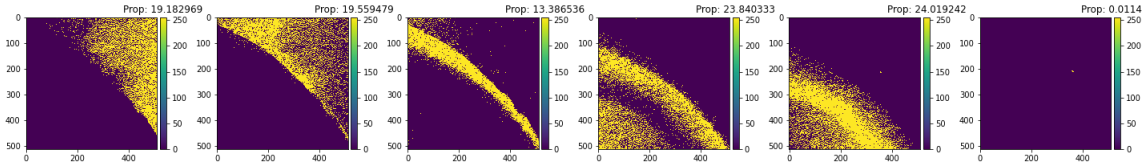
FIGURE 5.10: Edges Growing - Bounding Box - Example 2

5.1.4 Kmeans Clustering

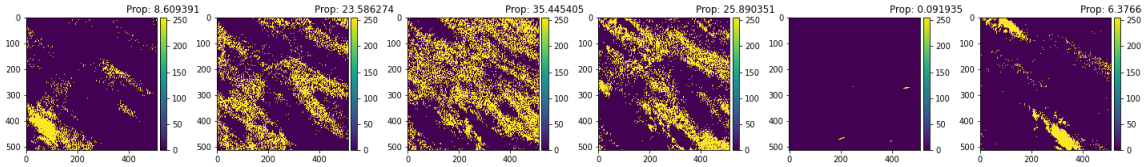
The last algorithm used was kmeans clustering. The number of K was 6.
Through a visual analysis of these images, it seemed that the clusters whose proportion, based on the quantity of image pixels (total of $262.144 = 512^2$), was less than 0.016%, were placed right on top of the vessels position, in the majority of the cases.

For each image, two binary images were formed: one based on the intersection of individual cluster whose proportion was inferior to 0.016%, and the other based on the union of the same clusters.
The binary image from intersection is the region's BILTV equivalent whilst the binary image from union is the region's BIHTV equivalent.
That being said, the further procedures are analogous to the ones on 5.1.2.

The figure 5.11 presents 6 binary images for 2 example images. The same images have their BIHTV and BILTV equivalents presented on Figures 5.12 and 5.13, respectively.

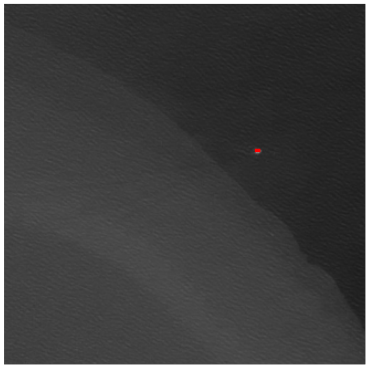


(A) Example 1

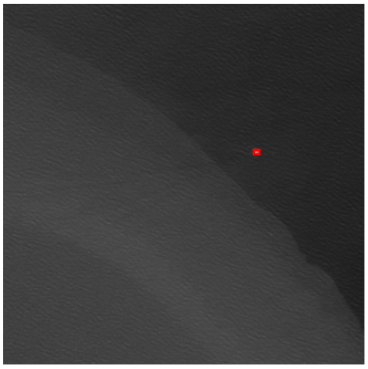


(B) Example 2

FIGURE 5.11: Kmeans Clustering for 2 images (RGB, left), with k=6, gener binary imating 6ages from left to right

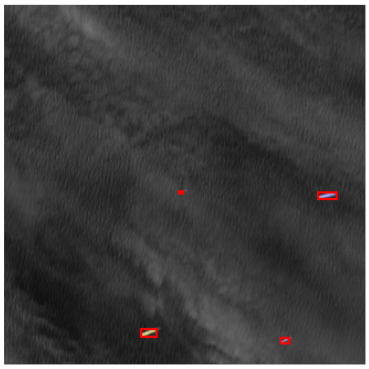


(A) Union Binary

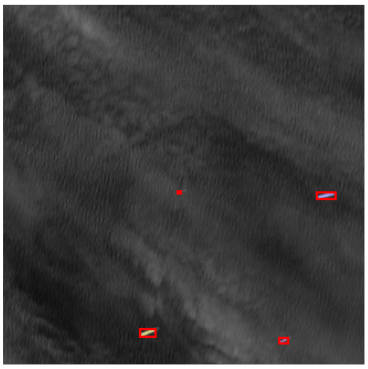


(B) Intersection Binary

FIGURE 5.12: Kmeans Clustering - Bounding Box - Example 1



(A) Union Binary



(B) Intersection Binary

FIGURE 5.13: Kmeans Clustering - Bounding Box - Example 2

5.1.5 Sample Results

By visually comparing the objects detected on BILTV (and its equivalent for Edges Growing and Kmeans Clustering) with the original images, it was possible to find not only the ground truth values, in other words, the real number of vessels per image but also the TP, FP, TN and FN values as well. These values had enable the calculation of four overlap evaluation metrics - Recall, Precision, F1_Score and Jaccard index.

Tables 5.1, 5.2 and 5.3 present all values and metrics described included the overall metric value, for each of the three segmentation algorithms. The Edge Growing algorithm has the best scores for all the metrics, with 90% precision. This suggests that it is the best algorithm for segmentation amongst the 3 in analysis.

Ground Truth	**TP**	**FP**	**TN**	**FN**	**Recall (TPR)**	**Precision (PPV)**	**F1_score**	**Jaccard**
1	1	0	0	0	1	1	1	1
2	2	0	0	0	1	1	1	1
4	3	0	0	1	0.75	1	0.8571	0.75
1	0	0	0	1	0	0	0	0
1	1	0	0	0	1	1	1	1
1	1	0	0	0	1	1	1	1
1	1	1	0	0	1	0.5	0.6667	0.5
2	0	0	0	2	0	0	0	0
5	0	0	0	5	0	0	0	0
2	1	8	0	1	0.5	0.1111	0.1818	0.1
2	2	0	0	0	1	1	1	1
5	1	0	0	4	0.2	1	0.3333	0.2
Overall score					**0.6208**	**0.6343**	**0.5866**	**0.5458**

TABLE 5.1: Region - Sample Evaluation Metrics

Ground Truth	TP	FP	TN	FN	Recall (TPR)	Precision (PPV)	F1_score	Jaccard
1	1	0	0	0	1	1	1	1
2	2	0	0	0	1	1	1	1
4	1	0	0	3	0.25	1	0.4	0.25
1	1	0	0	0	1	1	1	1
1	1	0	0	0	1	1	1	1
1	1	0	0	0	1	1	1	1
1	1	0	0	0	1	1	1	1
2	2	1	0	0	1	0.6667	0.8	0.6667
5	1	0	0	4	0.2	1	0.3333	0.2
2	1	1	0	0	1	0.5	0.6667	0.5
2	2	0	0	0	1	1	1	1
5	1	0	0	4	0.2	1	0.3333	0.2
Overall score					**0.804**	**0.9306**	**0.7944**	**0.7347**

TABLE 5.2: Edges - Sample Evaluation Metrics

Ground Truth	TP	FP	TN	FN	Recall (TPR)	Precision (PPV)	F1_score	Jaccard
1	1	0	0	0	1	1	1	1
2	2	1	0	0	1	0.6667	0.8	0.6667
4	4	0	0	0	1	1	1	1
1	1	0	0	0	1	1	1	1
1	1	0	0	0	1	1	1	1
1	1	0	0	0	1	1	1	1
1	1	1	0	0	1	0.5	0.6667	0.5
2	1	1	0	1	0.5	0.5	0.5	0.3333
5	0	0	0	5	0	0	0	0
2	0	0	0	2	0	0	0	0
2	2	0	0	0	1	1	1	1
5	1	0	0	4	0.2	1	0.3333	0.2
Overall score					**0.7250**	**0.7222**	**0.6917**	**0.6417**

TABLE 5.3: Kmeans - Sample Evaluation Metrics

5.2 Multi-vessel Detection

In this section it is explains how to access the ground truth values for the potential segmented vessels from the algorithms previously described.

First step

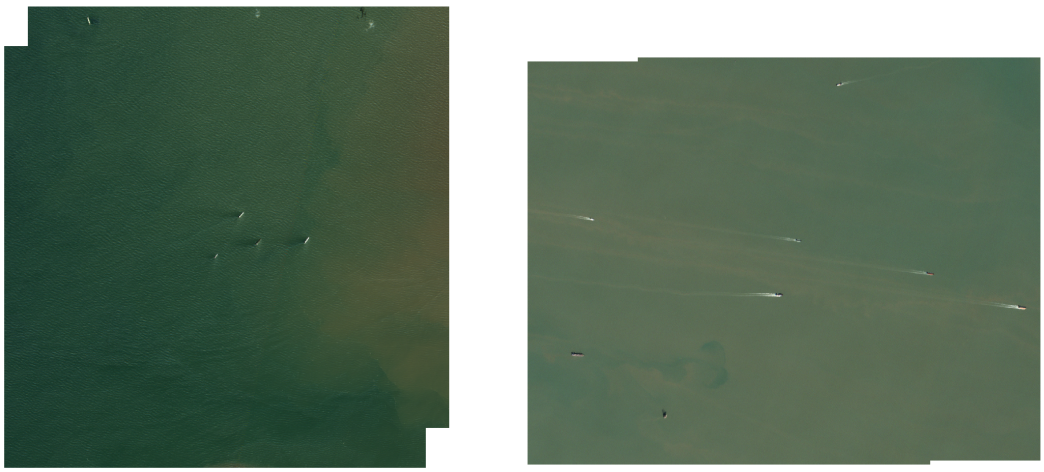
Each photo contains one metadata file with information about the image and information about the vessel associated with that same image. That same vessel may be registered on another photo. Together with the geographical referencing information gathered from each photo, it is possible to highlight the following relevant features:

- p_ts - photo timestamp
- p_min_la - photo min latitude
- p_max_la - photo max latitude
- p_min_lo - photo min longitude
- p_max_lo - photo max longitude
- v_ts - vessel timestamp
- v_la - vessel latitude
- v_lo - vessel longitude

The first step of this algorithm is to iterate the vessels information with the image information. If the three following criteria are fulfilled, then that vessel's metadata will also be associated to that respective image. Otherwise, it will not be considered.
The criteria used is:

- $p\_min\_la < v\_la < p\_max\_la$
- $p\_min\_lo < v\_lo < p\_max\_lo$
- $|p\_ts - v\_ts| < 1$ minute

Figure 5.14 presents two examples of groups of vessels located in different images, with 2 images overlap in both cases.



(A) Example 1 (B) Example 2

FIGURE 5.14: Overlap images for the same group of vessels

Second step

Now it is important to highlight other relevant information for the this next step where information from segmented objects and vessel's information are combined. This crucial information is:

- v_x - vessel's x coordinate
- v_y - vessel's y coordinate
- so_x - segmented object's x coordinate
- so_y - segmented object's y coordinate

The second step of this algorithm is to iterate every object from the segmentation algorithm against the vessels information. If the following criterion is fulfilled, that vessel's information is associated to that of the segmented object, otherwise, it is not.
The criteria is:

- $v\_x - 15 \text{ pixels} < so\_x < v\_x + 15 \text{ pixels}$
- $v\_y - 15 \text{ pixels} < so\_y < v\_y + 15 \text{ pixels}$

Chapter 6

Results

The region, edges and K-Means algorithms were capable of extracting 1068, 1057 and 937 different observations, respectively. The first stage of the multi-vessel algorithm generated 2918 different observations, which is more than 200% of the initial data (1273). Matching the two previous sets of data, like the 2nd stage indicates, made it possible to find the ground truth real_type values for the objects, enabling the attribution of the respective class. Table 6.1 presents the data for each algorithm.

	2345	**8**	**1679**	**Total**
Region	8	38	166	212
Edges	8	43	160	211
KMeans	13	37	112	162

TABLE 6.1: classes per segmentation algorithm

The first row indicates the clusters and the following rows indicate the absolute value of objects classified as being part of those clusters regarding the segmentation algorithm.

The data is very unbalanced so, before applying the classification algorithms, new data was also generated through oversampling and undersampling, giving a total of three different dataset per segmentation algorithm. Then, the data was scalled and used 5-fold cross-validation [] which provides a better estimate. The cross-validation method was part of the hiper-parameter tunning, where multiple combinations of the classifiers parameters are done in order to choose the combination which performs best, i.e. that has the highest scoring metrics within the classifier. Tables - 6.2, 6.3 and 6.4 - present the f1 and jaccard scoring metrics for every classifier used, for the unbalanced, oversampling

and undersampling datasets.

Region Growing

Unbalanced data shows to have better results, compared to the both type of balanced data, and oversampling data as always better results than undersampling data this might be due to the possibility of the test set containing only two or even one class. Support vector machines had the best performance for each dataset and overall. KNN had the lowest result of around 9% for precision and 4% for jaccard.

Scoring	*f1_score*					
Model	SVC	KNN	DCT	BAGG	RFOR	GRAD
Unbalanced	0.8372	0.8140	0.8140	0.8140	0.8134	0.7907
Oversampling	0.8140	0.6047	0.6977	0.8140	0.5814	0.6744
Undersampling	0.6512	0.0930	0.5349	0.6512	0.3721	0.5581
Overall	0.7674	0.5039	0.6822	0.7597	0.5891	0.6744

Scoring	*jaccard_index*					
Model	SVC	KNN	DCT	BAGG	RFOR	GRAD
Unbalanced	0.7200	0.6863	0.6863	0.6863	0.6863	0.6538
Oversampling	0.6863	0.4333	0.5357	0.6863	0.4098	0.5088
Undersampling	0.4828	0.0487	0.3651	0.4828	0.2286	0.3870
Overall	0.6297	0.3895	0.5290	0.6184	0.4416	0.5166

TABLE 6.2: f1 and jaccard metrics for region growing

Edges Growing

For the edges algorithm, the unbalanced data obtained between scoring results as well, followed by oversampling and then undersampling. In this case Bagging has the best overall scoring for f1 and Gradient Descending for jaccard. The lowest scoring values are both obtained by Random Forest on oversampling with 0.53 for f1 and 0.36 for jaccard.

Scoring	*f1_score*					
Model	SVC	KNN	DCT	BAGG	RFOR	GRAD
Unbalanced	0.7907	0.7907	0.7442	0.8372	0.8140	0.8372
Oversampling	0.7209	0.7907	0.6744	0.7442	0.5349	0.8140
Undersampling	0.6279	0.6279	0.5814	0.6947	0.6928	0.6047
Overall	0.7132	0.7364	0.6667	0.7587	0.6406	0.7519

Scoring	*jaccard_index*					
Model	SVC	KNN	DCT	BAGG	RFOR	GRAD
Unbalanced	0.6538	0.6538	0.5926	0.7200	0.6863	0.7200
Oversampling	0.5636	0.6538	0.5088	0.5926	0.3651	0.6863
Undersampling	0.4576	0.4576	0.4098	0.4333	0.4576	0.4333
Overall	0.5584	0.5884	0.5037	0.5820	0.5030	0.6132

TABLE 6.3: f1 and jaccard metrics for edges growing

K-Means Clustering

At last there is the K-Means algorithm. This one performed the worst. Its majority of jaccard scores are below 30%. The unbalanced data gathers the best scoring metrics. F1's lowest scoring goes to KNN on undersampling which happens to be also the lowest for jaccard. Random Forest had the best scores for each f1 and jaccard, regardless of the dataset.

Scoring	*f1_score*					
Model	SVC	KNN	DCT	BAGG	RFOR	GRAD
Unbalanced	0.5758	0.6667	0.6667	0.5152	0.6061	0.5758
Oversampling	0.3333	0.3333	0.3333	0.4545	0.4848	0.3939
Undersampling	0.4545	0.1515	0.3636	0.4545	0.5758	0.5152
Overall	0.4545	0.3838	0.4545	0.4747	0.5556	0.4949

Scoring	*jaccard_index*					
Model	SVC	KNN	DCT	BAGG	RFOR	GRAD
Unbalanced	0.4043	0.5000	0.5000	0.3469	0.4348	0.404255
Oversampling	0.2000	0.2000	0.2000	0.2941	0.3200	0.245283
Undersampling	0.2941	0.0820	0.2222	0.2941	0.4043	0.3469
Overall	0.2995	0.2607	0.3074	0.3117	0.3863	0.3321

TABLE 6.4: f1 and jaccard metric for kmeans

Bagging and SVC performed better on Region and Edges while Random Forest performed better on K-Means. Taking into account the sample scoring results on these segmentation algorithms, it seems to exist a correlation between segmentation scoring and classification scoring. In other words, the better the detection and segmentation of the vessels, the more accurate the vessel's shape attributes will be extracted resulting in a better data for modeling and predicting.

Chapter 7

Conclusions

The work developed indicates that it is possible to detect vessels on images as well as to classify the vessel in a number of abroad types.
The metadata from each image, which contained data regarding the image and the associated vessel, enabled the cross between both data, resulting in multiple vessel's information for each image, which consisted in an increase over Z-fold in observations.
There had been used a combinations of three segmentation algorithms with six classification algorithms. The Edges growing algorithm presented a more potential than the other two, and in fact, the classifier had a better scoring for it, rather than for the others.
Region and Edges had an overall scoring among unbalanced data above 80% while K-Means only reached 60%, for f1. For any dataset, kmeans performed worse than the other two, having an average scoring of the overall dataset per classifier of 0.47 for f1 and 0.32 for jaccard. These values are very low compared to the 0.66 and 0.52 on Region and 0.72 and 0.56 on Edges. KNN and Decision Trees were among the worst classifiers while Bagging was among the best.

For future work, it should be considered the manual identification, at least the number of vessel for each image or even also their coordinates, building the ground truth data for the 1273 images, in order to fully evaluate the segmentation algorithms.
It would be interest to study the results of these classifiers and these segmentation algorithms in another bands or combinations of RGB bands. Once the NIR bands are available, it will be possible to better segment and detect the vessels using, for instance, the NDVI index.
Finally, the usage of on of the most recent detection algorithms such as YOL O might be

of extremely relevance, due to its not only detection but also classification capabilities.

Appendix A

Appendix A - Algorithms

A.1 Region Growing

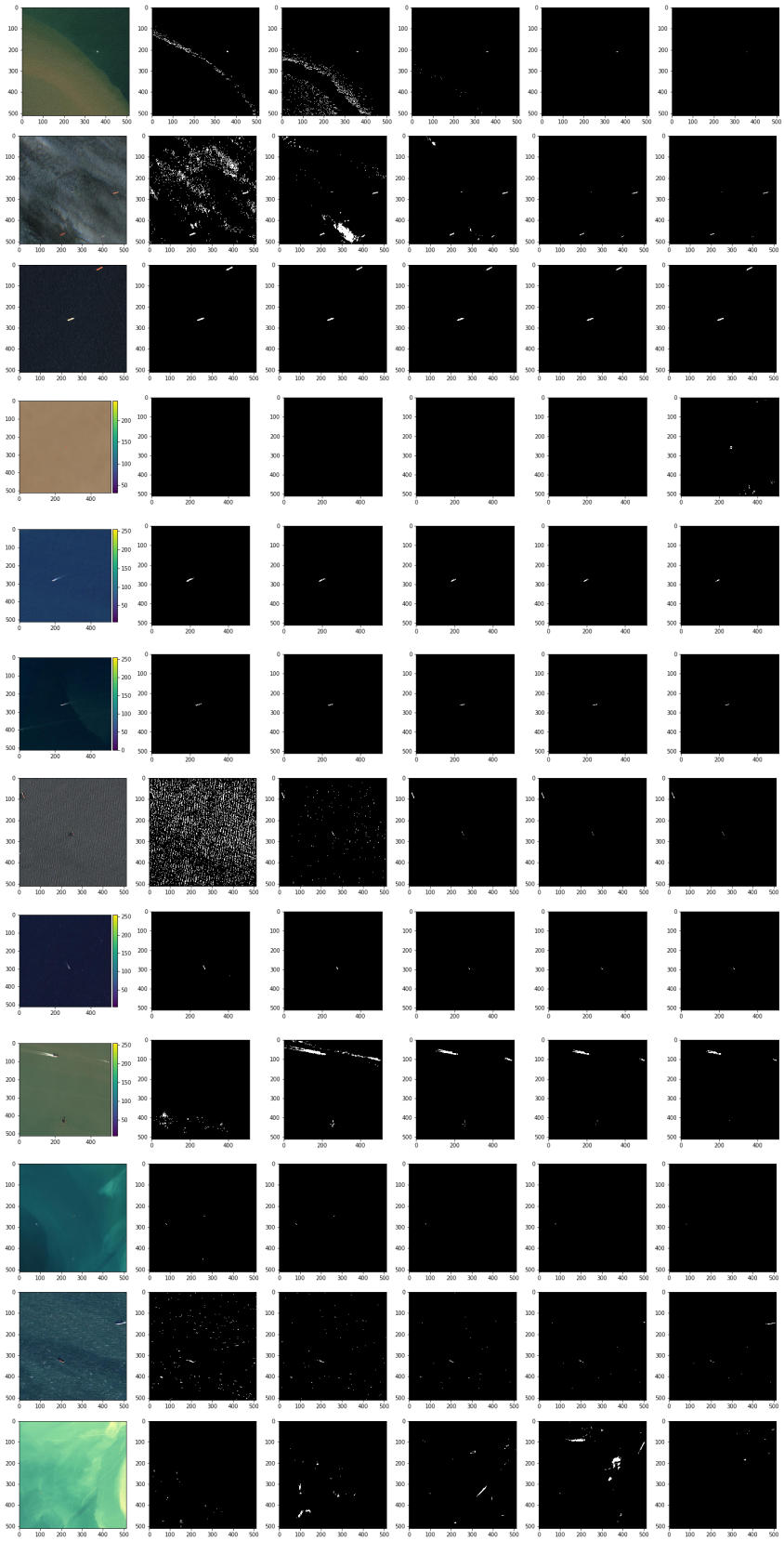


FIGURE A.1: Region Growing - binary images for chosen 12

A.2 Edges Growing

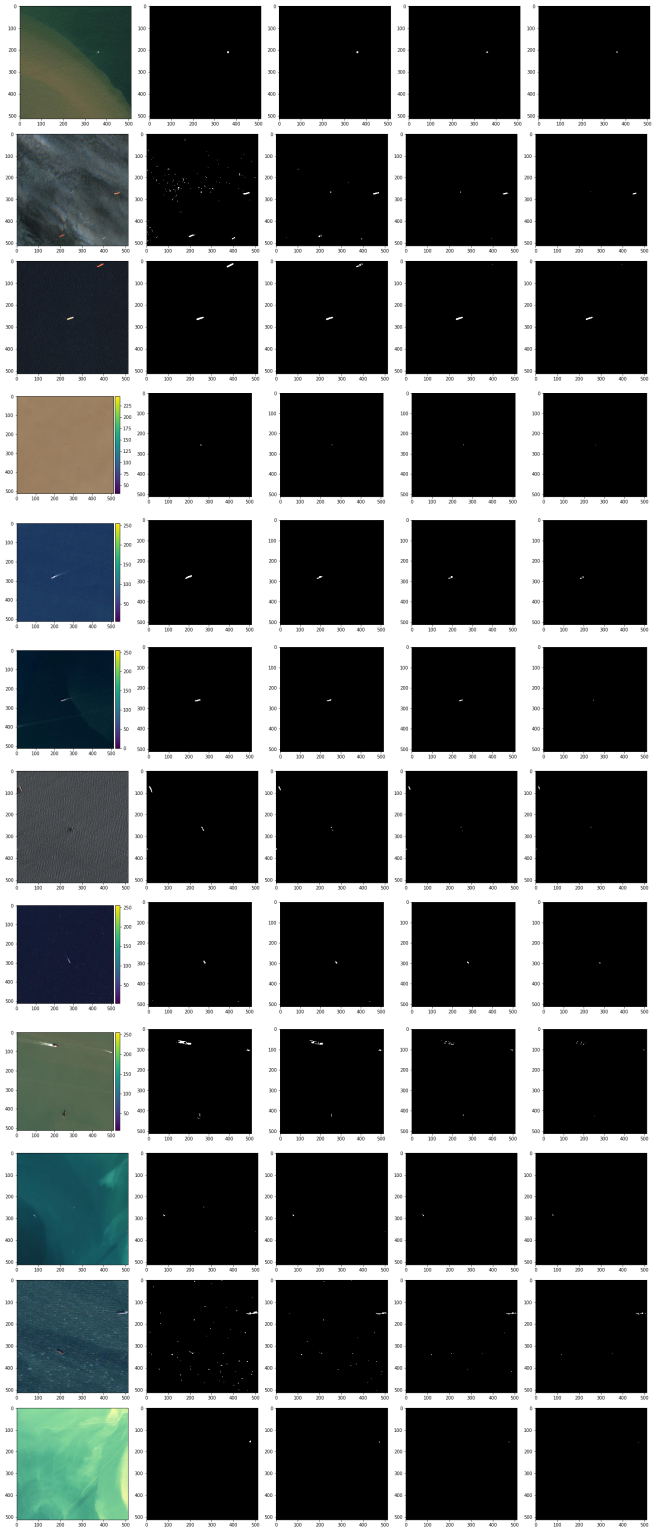


FIGURE A.2: Edges Growing - Binary images for chosen 12

A.3 Kmeans Clustering

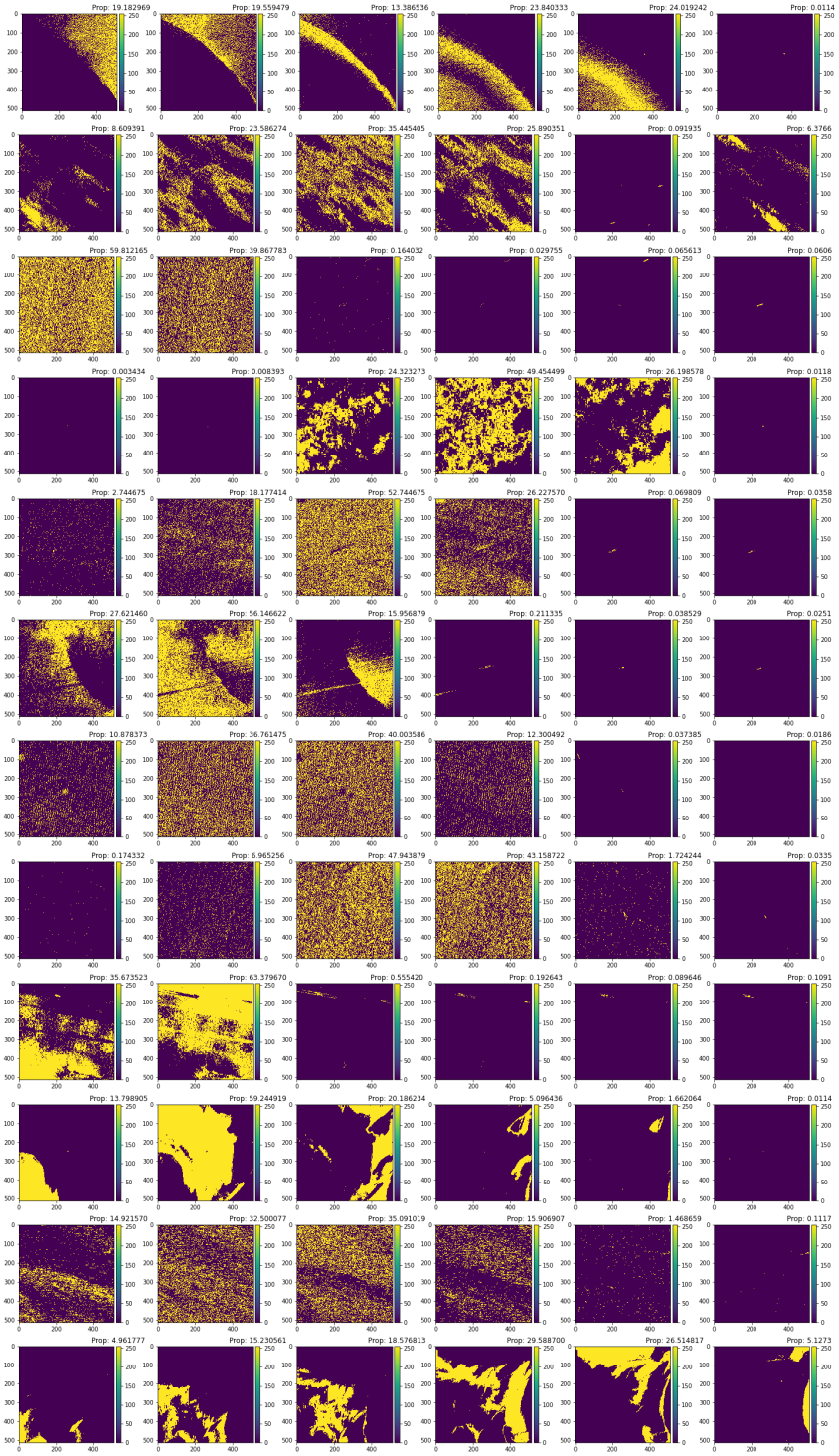


FIGURE A.3: Kmeans Clustering - Binary images for chosen 12

Bibliography

[1] G. Denis, H. de Boissezon, S. Hosford, X. Pasco, B. Montfort, and F. Ranera, "The evolution of Earth Observation satellites in Europe and its impact on the performance of emergency response services," *Acta Astronautica*, vol. 127, pp. 619–633, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0094576515303337 [Cited on page 1.]

[2] J. Larsen and C. Nissen, *Reconciling international priorities with local needs: Denmark as a new security actor in the Gulf of Guinea*. DIIS Report, 2018, no. 2018: 08. [Cited on pages 2 and 5.]

[3] R. Dekker, H. Bouma, E. den Breejen, B. van den Broek, P. Hanckmann, M. Hogervorst, A. Mohamoud, R. Schoemaker, J. Sijs, R. Tan, A. Toet, and A. Smith, "Maritime situation awareness capabilities from satellite and terrestrial sensor systems," 03 2013. [Cited on page 2.]

[4] "Webservice url." [Online]. Available: https://www.aishub.net/api [Cited on page 2.]

[5] L. Goldsworthy and B. Goldsworthy, "Modelling of ship engine exhaust emissions in ports and extensive coastal waters based on terrestrial AIS data – An Australian case study," *Environmental Modelling & Software*, vol. 63, pp. 45–60, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S136481521400262X [Cited on page 3.]

[6] R. Challamel, T. Calmettes, and C. N. Gigot, "A european hybrid high performance satellite-ais system," in *2012 6th Advanced Satellite Multimedia Systems Conference (ASMS) and 12th Signal Processing for Space Communications Workshop (SPSC)*, 2012, pp. 246–252. [Cited on page 3.]

[7] ESA, "Satellite – automatic identification system (sat-ais) overview." [Online]. Available: https://artes.esa.int/satellite-%E2%80%93-automatic-identification-system-satais-overview [Cited on page 3.]

[8] ——, "History of infrared astronomy." [Online]. Available: https://www.esa.int/Science_Exploration/Space_Science/Herschel/History_of_infrared_astronomy [Cited on page 3.]

[9] ——, "What is copernicus," Sep 2022. [Online]. Available: https://emergency.copernicus.eu/mapping/ems/what-copernicus [Cited on page 3.]

[10] ——, "Sentinel-2." [Online]. Available: https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2 [Cited on page 4.]

[11] Q. Zheng, W. Huang, X. Cui, Y. Shi, and L. Liu, "New Spectral Index for Detecting Wheat Yellow Rust Using Sentinel-2 Multispectral Imagery," *Sensors*, vol. 18, p. 868, 2018. [Cited on pages xv and 4.]

[12] ESA, "Next artificial intelligence mission selected." [Online]. Available: https://www.esa.int/Applications/Observing_the_Earth/Ph-sat/Next_artificial_intelligence_mission_selected [Cited on page 4.]

[13] CubeSat, "About." [Online]. Available: https://www.cubesat.org/about [Cited on page 4.]

[14] C. P. SLO, "Cubesat design specification rev 13." [Online]. Available: https://static1.squarespace.com/static/5418c831e4b0fa4ecac1bacd/t/56e9b62337013b6c063a655a/1458157095454/cds_rev13_final2.pdf [Cited on page 4.]

[15] E. A. B. da Silva and G. V. Mendonça, "4 - Digital Image Processing," in *The Electrical Engineering Handbook*, W.-K. CHEN, Ed. Burlington: Academic Press, 2005, pp. 891–910. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780121709600500645 [Cited on page 7.]

[16] K. R. Castleman, *Digital image processing*. Prentice Hall Press, 1996. [Cited on page 7.]

[17] G. Stockman and L. G. Shapiro, *Computer Vision*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2001. [Cited on page 7.]

[18] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979. [Cited on page 8.]

[19] S. Arora, J. Acharya, A. Verma, and P. K. Panigrahi, "Multilevel thresholding for image segmentation through a fast statistical recursive algorithm," *Pattern Recognition Letters*, vol. 29, no. 2, pp. 119–125, 2008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167865507002905 [Cited on page 8.]

[20] K. Somasundaram and S. P. Gayathri, "Brain segmentation in magnetic resonance images using fast fourier transform," in *2012 International Conference on Emerging Trends in Science, Engineering and Technology (INCOSET)*, 2012, pp. 164–168. [Cited on page 8.]

[21] R. Srisha and A. Khan, "Morphological operations for image processing : Understanding and its applications," 12 2013. [Cited on pages 8 and 9.]

[22] S. C. Ng, "Principal component analysis to reduce dimension on digital image," *Procedia Computer Science*, vol. 111, pp. 113–119, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050917311900 [Cited on page 9.]

[23] J. C. Harsanyi, "Bayesian decision theory, subjective and objective probabilities, and acceptance of empirical hypotheses," *Synthese*, vol. 57, no. 3, pp. 341–365, Dec 1983. [Online]. Available: https://doi.org/10.1007/BF01064702 [Cited on page 10.]

[24] S. Du, N. Zheng, Q. You, Y. Wu, M. Yuan, and J. Wu, "Rotated Haar-Like Features for Face Detection with In-Plane Rotation," vol. 4270, 2006, pp. 128–137. [Cited on page 10.]

[25] P. H. Swain and H. Hauska, "The decision tree classifier: Design and potential," *IEEE Transactions on Geoscience Electronics*, vol. 15, no. 3, pp. 142–147, 1977. [Cited on page 10.]

[26] M. A. Hearst, "Support vector machines," *IEEE Intelligent Systems*, vol. 13, no. 4, p. 18–28, jul 1998. [Online]. Available: https://doi.org/10.1109/5254.708428 [Cited on page 10.]

[27] W. Aha, D. Kibler, and M. Albert, "Instance-Based Learning Algorithms," *Machine Learning*, vol. 6, pp. 37–66, 1991. [Cited on page 11.]

[28] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, Aug 1996. [Online]. Available: https://doi.org/10.1007/BF00058655 [Cited on page 11.]

[29] C. Bentéjac, A. Csörgő, and G. Martínez-Muñoz, "A comparative analysis of gradient boosting algorithms," *Artificial Intelligence Review*, vol. 54, no. 3, pp. 1937–1967, Mar 2021. [Online]. Available: https://doi.org/10.1007/s10462-020-09896-5 [Cited on page 12.]

[30] M. J. Mcdonnell and A. J. Lewis, "Ship detection from landsat imagery," *Photogrammetric Engineering and Remote Sensing*, vol. 44, 1978. [Cited on page 15.]

[31] D. Burgess, "Automatic ship detection in satellite multispectral imagery," vol. 59, pp. 229–237, 1993. [Cited on page 16.]

[32] K. H. Pegler, D. J. Coleman, Y. Zhang, and R. P. Pelot, "The Potential for Using Very High Spatial Resolution Imagery for Marine Search and Rescue Surveillance," *Geocarto International*, vol. 18, no. 3, pp. 35–39, 2003. [Online]. Available: https://doi.org/10.1080/10106040308542279 [Cited on page 16.]

[33] C. Corbane, F. Marre, and M. Petit, "Using SPOT-5 HRG data in panchromatic mode for operational detection of small ships in tropical area," *Sensors (Basel)*, vol. 8, no. 5, pp. 2959–2973, May 2008. [Cited on page 16.]

[34] G. Wu, J. de Leeuw, A. K. Skidmore, Y. Liu, and H. H. T. Prins, "Performance of Landsat TM in ship detection in turbid waters," *International Journal of Applied Earth Observation and Geoinformation*, vol. 11, no. 1, pp. 54–61, 2009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0303243408000482 [Cited on page 17.]

[35] C. Zhu, H. Zhou, R. Wang, and J. Guo, "A novel hierarchical method of ship detection from spaceborne optical image based on shape and texture features," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 9, pp. 3446–3456, 2010. [Cited on page 17.]

[36] G. Huang, Y. Wang, Y. Zhang, and Y. Tian, "Ship detection using texture statistics from optical satellite images," in *2011 International Conference on Digital Image Computing: Techniques and Applications*, 2011, pp. 507–512. [Cited on page 17.]

[37] F. Bi, B. Zhu, L. Gao, and M. Bian, "A visual search inspired computational model for ship detection in optical satellite images," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 4, pp. 749–753, 2012. [Cited on page 17.]

[38] G. Mattyus, "Near real-time automatic vessel detection on optical satellite images," *ISPRS Hannover Workshop*, vol. Volume, pp. 233–237, 2013. [Cited on page 18.]

[39] P. Partsinevelos and G. C. Miliaresis, "Ship extraction and categorization from aster vnir imagery," in *International Conference on Remote Sensing and Geoinformation of Environment*, 2014. [Cited on page 18.]

[40] J. Tang, C. Deng, G.-B. Huang, and B. Zhao, "Compressed-domain ship detection on spaceborne optical image using deep neural network and extreme learning machine," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 3, pp. 1174–1185, 2015. [Cited on pages 18 and 19.]

[41] S. Raychaudhuri, "Introduction to monte carlo simulation," in *2008 Winter Simulation Conference*, 2008, pp. 91–100. [Cited on page 35.]

[42] A. J. Izenman, *Linear Discriminant Analysis*. New York, NY: Springer New York, 2008, pp. 237–280. [Online]. Available: https://doi.org/10.1007/978-0-387-78189-1{_}8 [Cited on page 35.]