# Ship detection using satellite technology

Valentino Portillo
MSc DS
IMT Atlantique
Brest, France
valentin.pg@outlook.com

Tanuj Garg
B.Tech- Mathematics and Computing
IIT Delhi
New Delhi, India
Email: tgarg199@gmail.com

***Abstract*—Ship detection has now become a widely researched area, majorly because of security reasons in marine economic zones of different countries. This is also beneficial for more efficient management of fisheries, and emergency situations like oil spills etc. The techniques have advanced from classical to learning, applied to different types of images like Synthetic Aperture Radar (SAR) and Multispectral Instrument (MSI) satellite images. This letter aims at getting a better understanding of major techniques proposed and algorithms used at present, in this field. This letter briefly explains different procedures employed in detecting ships using the SAR and MSI images, the advantages and disadvantages of MSI images over SAR images. This letter also points out the development of the software Search for Unidentified Maritime Objects (SUMO) developed for SAR images and joint usage of satellite imagery with other available data such as Automatic Identification System (AIS) and procedures to boost highlights from this information like big data.**

**Keywords:** ***SAR, MSI images, AIS, Ship Detection, Ship discrimination, SUMO, Deep Learning, Big Data***

## I. Introduction

With expansion in sea trade, possibilities of illegal sea activities have also increased, raising the need of better maritime security. Also, each country with a coastline, has some region of sea in its territory, whose resources can be used by that country. This results into need for better security and resource management. Ship detection, being a major component of security, has been widely researched over for decades. Besides, resources required to scan the sea and even the Exclusive Economic Zones (EEZ) are very high because of a huge survey area in consideration. However, this is very expensive and inaccurate task because it's automation lacks efficiency and therefore, an accurate automation with minimum human intervention becomes vital and interesting. This includes using different techniques: Classical and learning, and using different types of images: SAR, MSI images and lately, hyperspectral images. This letter attempts to analyse the existing methods used and proposed in recent years.

Over the years, mainly SAR images were used for the purpose of ship detection, because of lack of appropriate technology for MSI images. However, in the last decade, with the increase in sensors for MSI imagery, a huge increase in research on MSI images has been observed.[1]. MSI images are usually referenced as Optical images in literature.

Once ships are detected in sea, they could to be verified to be legal or illegal. For that purpose, some ships have a machinery called the AIS. It's placement in some cases is mandatory but not all every ship has it and it could be turned off. Apart from AIS, Vessel Monitoring System (VMS) is also used for the same purpose.

This review is further structured as follows. In Section II, we introduce the AIS system. Section III deals with the differences between SAR and MSI images. Section IV explains in detail, different algorithms used in ship detection using the MSI images.In Section V, methods used for SAR images are presented. Section VI is focused on a particular algorithm included in a software: SUMO, which is a fully automated software used for ship detection in SAR images. Section VII aims to present the integration of AIS technology with Ship detection methods, in order to identify the ships after their detection. Section VIII is concerned with research directed over by the private sector to get some highlights about their strategies. Finally, we have the discussion in Section IX.

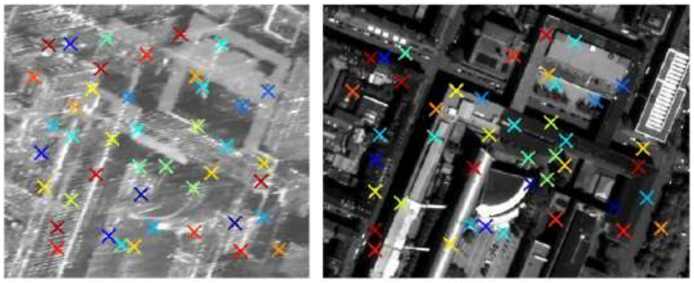
## II. AIS

The AIS is designed to prevent collisions. Thus, each ship transmits its position to the other ships in the vicinity, and everyone is aware of the surrounding maritime traffic. AIS is an interesting way of observing cooperative ships and several highlights could be extracted from the records of this information, specially when AIS data belongs to a satellite receiver. But, as not all ships are cooperative, or simply dont have the required equipment and as we are interested also in detecting IUU (Illegal, Unreported, and Unregulated) activity, satellite imagery analysis plays a huge role. Another limitation of AIS is that in high traffic, packet collisions sent to the satellite can lead to wrong information (incorrect localization, time)

## III. SAR v/s MSI images

There exist different systems and types of sensors for capturing landscapes from distance (mainly aircraft/satellites). Major two of them are Synthetic Aperture Radar (SAR) and Multispectral Instrument (MSI). SAR sequentially send electromagnetic radiations and detect their echos, and store the data for further work (https://en.wikipedia.org/wiki/Synthetic-aperture_radar,*accessed Nov 2018*). MSI measures the Earth's radiance in several spectral bands from Visible and Near Infrared (VNIR) to Short-Wave Infrared (SWIR). There are

however some major differences in the two types of sensors (http://sar.ece.ubc.ca/SARintro/SAR.html, *accessed Nov 2018*). SAR systems work on a wavelength of 1cm to 1m, whereas MSI systems work in the wavelength close to that of visible light (1 micron).
The figure below shows a SAR image to the left, and an MSI image of the same region to the right. It has been taken from [2].



(a) SAR image. (b) Optical image.

Fig. 1. SAR versus MSI image, taken from [2]

SAR systems create their own sources of illumination in the form of radio waves emitting from the antenna, and don't require any form of external illumination. On the contrary, MSI depends on the illumination provided by the sun. Also, MSI face interference by clouds, haze, winds, reflection of sunlight from water and wave crests [1], and therefore, it results in a large number of false alarms.

On the contrary, SAR images suffer from relatively less spatial resolution, high level of speckle noise. Also, it is difficult to recognize false alarms in a SAR image. These images have some limitations with regards to the Radar Cross Section (RCS), so the material of the ship could be an inconvenience (example: wooden boats), which poses an important disadvantage with respect to the MSI images. Also, other limitations are caused by surface conditions, angle of incidence of the radar signal and direction of travel of the vessel [3].
Currently, from a practical point of view, one of the main advantages of SAR is that several agencies give free access to this data such as Copernicus Science Hub (https://scihub.copernicus.eu/ *accessed Nov 2018*) and Alaska Satellite Facility (https://www.asf.alaska.edu/ *accessed Nov 2018*), where we can find free Sentinel satellite type images. Also, depending on the satellite, we can even get $1m$ long pixel images. But one of the most important disadvantages is the frequency of passage of these satellites, which is 1 every 6 days until next revisit of a certain area. (https://sentinel.esa.int/web/sentinel/user-guides/sentinel-1-sar/revisit-and-coverage *accessed Nov 2018*).

The satellites used for this purpose, that are able to revisit "frequently", are included in the Low Earth Orbit satellites and as it's name states, it lets them travel faster than the medium earth orbit. For example, they are able to do around 13 periods per day. Specifically, Sentinel is a group of satellites (Sentinel 1, 2, 3, 4, 5, 6) meant to observe the earth issue from a program carried out by the European Spatial Agency (ESA). Each generation of Sentinel is equipped with different instruments, so each generation is able to perform some tasks, for example Sentinel 1 is dedicated to SAR and Sentinel 2 to MSI.

## IV. Methods used for MSI Images

The task is performed in two major steps: Ship detection and Ship discrimination. Ship detection is performed to identify and determine the candidate regions, where the ships are possibly present. Ship discrimination is done to remove false candidate regions, also called false alarms. The aim of algorithm is to maximize the probability of detection while minimizing the probability of false alarms, which is why the two steps are carried out. Some preprocessing and postscreening steps are also applied in order to enhance the results. Preprocessing steps include land-sea segmentation, and other methods used to reduce the false alarms caused by factors such as clouds and waves. Postprocessing methods include algorithms to exclude the background, remove candidate regions consisting of false alarms. In this section we discuss in detail about the different algorithms used for these processes. We first discuss about the pre-processing methods.The figure shown below has been taken from [1].

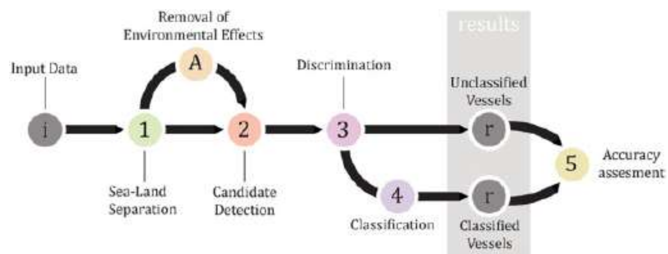



Fig. 2. A common scheme of vessel detection workflows [1].

### A. Preprocessing Methods

Since we wish to detect ships present in the sea, it becomes important to separate land regions from the sea, else different spots on the land region might occur very frequently as false alarms. Two major types of methods are used for this. This is either done by using the image along with the already available geographical information, or by using other methods for segmentation. We majorly discuss the latter. Land-sea segmentation can be performed using spectral fusion on multi-spectral images as in [4]. These images can be captured with the sensors that are used for MSI images as well. This method exploits the fact that the reflection capability of water is less in infrared region than in visible region of wavelength, whereas that of vegetation is higher in infrared region. Sea-Land segmentation is then easily performed by simply obtaining a threshold value for Normalized Difference Water Index (NDWI), using statistics.

$$NDWI(p,q) = \frac{\rho_g(i,j) - \rho_{nir}(i,j)}{\rho_g(i,j) + \rho_{nir}(i,j)} \tag{1}$$

NDWI for any pixel value $(p, q)$ is a quantity that shows the reflectance of the pixel with respect to the green band vis-a-vis reflectance with respect to infrared band [5]. Here, $\rho_g(p, q)$ represents the digital number of a pixel in green band and $\rho_{nif}(p, q)$ represents the digital value of pixel in near infra-red band. Equation$(1)$ has been cited from [4].

This method as used in [4] also helps in removing thick clouds. In [6], a dynamic multi-feature fusion model is used to segment land from sea. Multi feature integration includes the fusion of gray scale and texture features. Gray scale feature for each pixel is the mean of gray scale intensity of all pixels in a fixed sized window around that pixel. Texture feature is 8 directional (for 3x3 window), and hence 8 valued -feature map for each pixel. It is calculated in each direction(of all connected neighbours) by finding second order moments of local Walsh coefficients in for pixels in each window. Local Walsh coefficients are also directional, and are calculated by using Local Walsh transform (LWT). All these maps are next integrated by normalizing them to same scale to form fusion feature map. Then, minimum error segmentation (explained later in the paper) is applied to segment land and sea.

Another Method used in [7] tries to increase the contrast of the target by detecting the edges in an image. This is done by adding a quantity proportional to the Sobel operator at any pixel to the intensity at that pixel. Sobel Operator on an image gives an output that is proportional to the probability of edges existing at each pixel.

$$Mixf(i, j) = f(i, j) + a * Mag(i, j) \tag{2}$$

Here, $Mixf(i, j)$ gives the resultant gray value at any pixel $(i, j)$. $f(i, j)$ gives the original intensity, $Mag(i, j)$ represents the edge magnitude at any pixel, and is calculated using the Sobel operator. Equation $(2)$ has been cited from [7].

Cloud masking and local contrast enhancement is performed in [15]. Here, local thresholding (explained later in the paper) is performed for different parts of an image, assuming the statistics to be Gaussian. It is able to remove false alarm clouds other than thin clouds. We now discuss about the Ship detection methods.

### *B. Detection Methods*

Different image segmentation methods have been used to extract candidate regions. We start by describing the most basic algorithm used.

Adaptive Segmentation: This method, as presented by Otsu [8], provides a global threshold for an image in an unsupervised manner. This then helps in partitioning an image into two, based on the comparison between each pixel's intensity and this threshold. The threshold is found so as to decrease the similarities between the two classes, and increase the similarities in the intra-class context. The process involves maximizing the inter-class variance, and minimizing the inter-class variance. Since local statistics are usually different in an image, local thresholding is also done for different segments of the image.

This is used in [7] to find a global threshold, as a coarse method of segmentation. [7] also then uses morphological operators, which work as region filling operators [9], and help in removing noise in the segmented image, which can be generated because of ocean waves. This research then uses a finer method of segmentation, which effectively acts as an edge localization algorithm, and helps in refining the countours of the divided regions based on the Chan-Vese Model [10]. This model proposes to minimize the energy between the two segmented regions. The energy in the Chan-Vese Model can be given by the following expression.

$$\begin{aligned} F(C) = F_o(C) + F_b(C) = \iint\limits_{insideC} |u(x, y) - c_o|^2 \, \mathrm{d}x \, \mathrm{d}y \\ + \iint\limits_{outsideC} |u(x, y) - c_b|^2 \, \mathrm{d}x \, \mathrm{d}y \end{aligned} \tag{3}$$

Here, $F(C)$ is the energy when C is the contour of segmentation, $F_o(C)$ is the energy of region outside the curve, $F_b(C)$ is the energy of region inside the curve, $u(x, y)$ is the grey level intensity of a pixel, $c_o$ is the mean of pixel values inside the curve, $c_b$ is the mean if pixel values outside the curve. Equation $(3)$ has been cited from [7].

Simple adaptive segmentation usually doesn't work because of the absence of bimodal statistics even locally. So, usually the segmentation is done on the basis of some features of the image that somehow account to the target being a ship.

In the paper [6], Simple connected regions are taken as candidates that are obtained from the sea-land segmentation. Next for extraction, the Variance feature (VF) is used. by calculating the variance value at each pixel, with respect to a window centered at that pixel. Then, the image is segmented based on variance feature. Then to obtain final candidate regions, AND operation is performed on results of VF and DFM (Dynamic Fusion Model).

Graph based fore/background segmentation: This method proposed in [11], uses Min-cut/Max-flow algorithms as labelling method. These methods [12] label each pixel as foreground or background and are inherently binary. Here, a label is assigned to all pixels of image such that the energy corresponding to that label in that image is minimum. Fore and background act as source and sink for the image. Each pixel acts as a node, connected with source/sink with a weight that amounts to its probability of belonging to source/sink. The nodes are also connected to other nodes in their 4-neighbourhood depending on similarities, by edges with their corresponding weights/costs. The image is then partitioned into source and sink by cutting edges, leading to two components of graph belonging to source and sink. This has to be done so that the cut is a Min-cut with respect to costs. Equation is given below: The graph that is finally constructed is supposed to minimise the energy of each label $f \in [l_{fore}, l_{back}]$

$$E(f) = \sum_{p} D_p(f_p) + \sum_{(p,q)} V_{p,q}(f_p, f_q) \tag{4}$$

where, $D_p(f_p) = (f_p - I_p)^2$. $I_p$ is the intensity of pixel p. Also $V_{p,q}(f_p, f_q)$ equals 1 if $f_p$ equals $f_q$, else 0. Equation $(4)$ has been cited from [12]. This is used with the adaptive segmentation in [11] for the segmentation process.

In [13], active contours based segmentation algorithm was used. It is called GPAC [14].Here, the algorithm starts by distributing small countours (initially circles) which are iteratively evolved, to adapt to the specific characteristics of that image. This finally separates the image into two regions, such that the similarity in the inter region context is minimum, whereas that in intra-region context is high. Equation is shown below:

$$E = \iint_{p_2 \in R_i} \iint_{p_1 \in R_o} w(p_1, p_2)\, \mathrm{d}p_1\, \mathrm{d}p_2 \tag{5}$$

Here, $R_o$, $R_i$, $p_1$, $p_2$ represents outside region, inside region and pixels respectively. $w(p_1, p_2)$ is the measure of similarity between the two pixels. Equation $(5)$ has been cited from [13]. We also see segmentation approaches based on applying connected morphological operators on the images as used in [15]. In these approaches, morphological operators are applied on an image in order to retain or remove the so desired features. Connected Operators are a special class of operators, which interact by finding components of an image, and then either preserve or remove the entire component depending on the feature [16]. We define connected components to be regions of space where the image is constant. Here, first a tree from the image is defined based on its components, and is called the component tree [17].

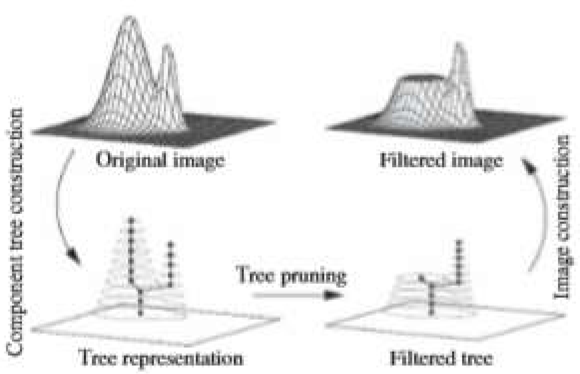



Fig. 3. Component tree and connected operator filtering, taken from [15].

This tree is used for better representation of different components of an image and the links that exist between them, where each node represents its absolute gray level. Tree pruning/filtering is performed in order to remove some of the connected components from the tree, by deciding different nodes to be active/non-active. Each component is either selected/rejected on the basis of different attributes of that component such as the Height and area of the component in the 2-D image. The entire process is shown in Fig.3,[15]. Appropriate thresholds are found for these different attributes, and applied henceforth.

Wavelet Transform: Whereas the Fourier transform characterizes a signal in frequency domain, WT does it in both time and frequency domains. Here, the signal (image) is convoluted with a wavelet, that is a function of two variables, scale and translation. [4] uses extended wavelet transform (EWT), which is a novel method proposed here, instead of discrete wavelet transform, because the DWT reduces spatial resolution of the image. Here, EWT is coupled with the Phase Saliency Map (PSM) method. EWT is helpful in increasing the contrast of the targets as compared to the background and is then forwarded to PSM.

Phase Saliency Map(PSM): Visual saliency is the concept that mimics how human vision works, and attempts to highlight the areas that catch maximum attention of human eyes, called as salient signals. Saliency is the quality of any pixel(here), by which it stands out from its neighbours (https://en.wikipedia.org/wiki/Salience_(neuroscience), *November 2018*). Phase Saliency can be computed using Phase spectrum of fourier transform, as in [18]. The equations have been given below.

$$f = F(I) \tag{6}$$

$$p = P(F(I)) \tag{7}$$

$$S = g * ||F^{-1}(e^{ip})||^2 \tag{8}$$

$I$ here refers to the image. $F[.]$ and $F^{-1}[.]$ represent Fourier and inverse Fourier transforms respectively.P[.] represents the phase spectrum of the Fourier Transform. Equations $(6), (7), (8)$ have been cited from [19] .Adaptive Threshold based segmentation is then simply applied on the phase saliency map to give binary maps as their outputs. Then, each connected area in this binary map is bounded by rectangles and is called candidate region. This method is also used in [4] coupled with EWT,[19],[18]. Wavelet Transform along with Phase Saliency Map is used in [4] to perform candidate extraction.

Bottom-up Saliency Method: This method aims again at mimicking the visual search mechanism of the human eye and has been employed in [20]. The algorithm starts by downsampling the original scene $M_{ori}$ to $M_{mid}$ which is further downsampled to $M_{low}$. Thus three stages mimicking visual search are set. The first step is called Attention of Candidate Regions (ACR), wherein candidate regions are extracted from $M_{low}$ using Pulsed cosine Transform(PCT). Equations are given below:

$$P = sign(C(M_{low})) \tag{9}$$

$$F = absC^{-1}(P) \tag{10}$$

$$SM_{low} = G * F^2 \tag{11}$$

Here $C[.], C^{-1}[.], sign(.), abs(.)$ represent 2-D discrete cosine transform, inverse cosine transform, signum function and absolute value function respectively. Equations $(9), (10), (11)$ have been cited from [20]. Further, adaptive thresholding is again used to provide us with initial candidate regions. In the next step, context of each candidate region is analyzed (as in visual process), called as Local Context Facilitation (LCF). Here, a neighbourhood method is applied to judge if the ship has similar characteristics as neighbouring sea or not, as other false alarms don't satisfy this. This is done by drawing 8 blocks neighbouring the candidate region in $M_{mid}$ with the same size as that of candidate region, and different features from each blocks are extracted and compared, such as mean, variance, texture based features. Texture based features are measured using the gray-level co-occurence matrix (GLCM). If the similarity level is above a certian threshold, candidate region gets passed for the third level in $M_{ori}$ called as Appearance Identification (AI). This stage is intended to remove the false alarms which have same characteristics as their surroundings. Here, The candidate region is extracted from the $M_{ori}$ and the Scale-invariant feature transform (SIFT) descriptor [21] is used, and finally this process is merged with ship discrimination and classification by using the learning method of Support Vector Machine (SVM), which will be discussed later on in the paper. It is proposed by Bar [22] that the three stages-ACR, LCF, AI mimic the functions of three types of neuronal cells-Prefrontal Cortex (PFC), parahippocampal cortex (PHC) and inferior temporal cortex (ITC) respectively.

All the methods mentioned above are majorly unsupervised, and have disadvantages in terms of their efficiency in the presence of clouds and in case of import ships [23]. Moreover, the parameters for the post-processing methods are used by hand, and hence are unreliable. So far, we have seen pre-determined feature descriptors, for which thresholds were obtained using machine learning. We now see a system where features are also learned by the system.

Convolutional Neural Networks(CNN): These are deep, feed-forward artificial neural networks, that mimic the functioning of human neural system (https://en.wikipedia.org/wiki/Convolutional_neural_network *accessed Nov 2018*). The CNN consist of many layers, each layer representing mapping mechanism for certain features, which keep on getting complex with the layer number, Here, these features are also learned by a training set. In CNN, there are some convolutional layers, some non linear layers, which finally gives the ship candidates. Self feature learning is done by the help of many algorithms. This method was employed in [23], called as SVDNets, because this method used Singular value Decomposition (SVD) algorithm for self-feature learning along with CNN. Here, the CNN has 3 Convolutional layers(CL) and 3 non linear layers, last one of which (CL) is useful in suppressing the backgrounds. At the end of all these layers, we get an image of the same size as original, with each pixel representing the probability of that pixel to have a ship. SVD (unsupervised) is only used for learning features of the first and second convolutional layers Third layer is used only for highlighting the ship candidates and suppressing unwanted backgrounds. Fig.4 shows different layers of CNN.

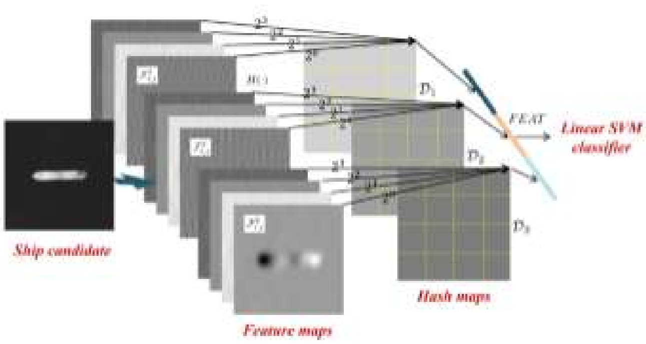



Fig. 4. Convolutional Neural Network, taken from [23].

### C. Postscreening Methods

We now talk about the Postscreening methods.

[19] uses a homogenisation filter. This filter homogenizes similar area of a candidate region, which can get affected because of segmentation (PSM Method). The filter for each pixel of the image works in an iterative convex manner. Otsu segmentation is further applied again.

[7] uses simple shape Analysis to eliminate evident false alarms. Different parameters come into play here. Since ships must usually be small, big pixel are candidates which probably consist of clouds and islands are done away with, with the help of appropriate thresholds.Also, since ships are usually bar shaped, their length to width ratio must be greater than a certain minimum.

[11] uses the CFAR (Constant False Alarm Rate) Scheme to detect some parts of a ship that go unnoticed in segmentation result because of shadow of parts of ship on the sea and ship itself. Here, first a convex hull of the ship is obtained, through which we can obtain the shadow mask for that image. The shadow mask is then again graph-segmented and the remaining parts of shadow are judged by the help of statistics of the shadow,sea and ship to be either sea or ship. This is done in the same fashion as the CFAR scheme.

### D. Discrimination Methods

We now talk about ship discrimination methods.

This step is applied in order to remove the false alarms that have so far been included in the candidate regions. Here, most research papers use supervised methods based on several features to classify the target as ship/non-ship. To do the job, several feature descriptors have been used that are suitable to show the presence of a real ship in the candidate region. We now list the several features used. Most commonly used features are the length, width, area, length to width ratio (called eccentricity)(used in [7], [4],[6]) are commonly used features. Other shape based features used are Compactness(ratio of square of perimeter to the area of region)(used in [7], [4]), Convexness (ratio of area of the region to the area of the convex hull of that region)(used in [7]), Rectangularity (maximum value of the ratio of area of the region to the area of rectangle bounding the candidate region)(used in [7], [4], [6]

) and Moment invariants (used in [7]). Several other Texture based features such as mean,variance, moments,entropy are used in [7],[4].

Other than shape and texture features, spectral features such as Wavelet Transform (discussed above)(used in [7],[15]), Radon Transform(used in [15]), Differential features such as Histogram of Shape index(used in [7]), Scale-invariant feature Transform(SIFT)(used in [20]), features used to determine difference in texture distribution such as Local Binary Patter(LBP)(used in [4]) and Local Multiple Pattern(LMP)(used in [7]) are used.

Histogram of Shape index: Shape index is a curvature based feature defined at every pixel of any region, equation is given below.

$$S(p,\sigma) = arctan\frac{\kappa+\mu}{\kappa-\mu}(p,\sigma) \tag{12}$$

Here $\kappa$ is the isophote curvature and $\mu$ is the flowline curvature of the intensity surface [7]. Equation $(12)$ has been cited from [7].

Radon Transform (RT) : This transform is useful in linear features detection. It takes a function $f$ defined on a plane to a function $Rf$ defined on 2-D space of all lines in the plane, such that its value at any particular line equals the line integral of function at that line [15]. RT enhances the ship and its wake. Relative Height of the peak obtained after RT is taken as a feature. Equation $(13)$ given below has been cited from [15].

$$R(p,\tau)[f(x,y)] = \int_{-\infty}^{\infty} f(x,\tau+px)\,\mathrm{d}x \tag{13}$$

Methods such as [23] use features from the CNN layers obtained.

Now, once all the features on the basis of which classification is to be done are obtained, we need to employ suitable Machine learning methods to actually classify them. The most commonly used classifier is Support Vector Machine (SVM) (used in [6], [7], [20], [4], [23]). SVM takes different features of images( as described above), then divides the points by an optimal hyperplane (SVM is supervised), in order to then discriminate if candidate is ship or not (https://en.wikipedia.org/wiki/Support_vector_machine , *November 2018*). Several other classifiers such as Bayesian (used in [24], Neural network, logistic regression (used in [15]), Random Forest, k-Nearest Neighbours are used.

Work done in [25] compares the performances of several of above mentioned binary classifiers for several cases and also implements the CNN method for classification, and it is concluded that CNN is a superior method for image classification.

Some researchers also use unsupervised feature descriptors for the job as in [19]. Here, a novel form of Histogram of Gradients(HOG) is proposed, called as S-HOG, or Ship HOG. This feature takes into account the fact that the gradients on the sides of a ship are symmetric with respect to the major axis of the ship, which is assumed to be symmetrical here. Also that the major component of these gradients would be present in perpendicular to the minor axis direction, because of the bar like shape a ship or a boat. Gradient Directions are divided into groups to take into account the above stated facts, and analysis is done.

## V. Methods used for SAR images

In order to be able to accurately detect vessels in the sea, there are several methods. A great part of them use the CFAR as the ship detector algorithm for excellence, and propose methods to improve its accuracy.

CFAR is a stochastic approach. In order to calculate it, some parameters related to the K distribution (input) need to be established. After this, Probability Density Functions ($PDFs$) (regarding mean, width measures, amplitude, distributed sea clutter) will be used to get the CFAR threshold according to the values of the parameters, the known amplitude mean and variance. Other parameter that should be considered is the $f$ which is meant to compensate the difference between the real clutter $PDF$ and the fitted K distribution. Its value depends on the polarization used for the SAR image and the intensity of the sea clutter, and is given as an input [26].

Some of the algorithms to process SAR images will be mentioned. As preprocessing to CFAR , there is sub-look analysis that is meant to increase the contrast between ships and sea if it is low (in case of cluttered sea) , proposed as a complement for intensity detectors. Sub-look analysis is based on the idea that portions of the sea clutter single-look complex spectrum are uncorrelated, while vessels usually preserve high correlation along the spectrum [27]. Majority of alternative algorithms apply deep learning approach in one or more steps of the process. As an example, one of the papers uses "Caffe" as learning Framework [27]. This approach includes three steps: making a training set, training the network, completing the activity for which the algorithm was trained. In order to describe the logic of these frameworks, we are going to list them and its main methods:

- Land masking using a Fully Convolutional Network (FCN) + CFAR as detection algorithm + false alarms discrimination with a Neural Network [28].
- A process inspired by the human vision system that applies a random-forest-based hierarchical sparse model (HSM) in order to select the candidate regions, then the CFAR with a contour saliency model to do the filtering of false alarms [29].
- CFAR + Support Vector Machine (SVM) for classification [30].

Also, some algorithms that don't use CFAR during their process. They are listed below

- A Texture Feature based Deep Learning Algorithm. Here. a Back Propagation Neural Network (BNN) for the analysis is proposed [31].
- Sea Land Segmentation Based Convolutional Neural Network (SLS-CNN) [32].

Some of the above algorithms are also compared in their papers. [28] is faster (1s - 2s) than [32] (10s) and [31]

(550s) and for the land masking it outperforms the OTSU and Entropy methods. The [32] states that it outperforms the CFAR algorithm: 95% against 40% of precision.
In general, for the time complexity, we have not found an algorithm that takes more than half an hour for analysis of an image.

## VI. SUMO

Search for Unidentified Maritime Objects (SUMO) is a ship detection algorithm and it is an open source software (https://github.com/ec-europa/sumo, *accessed Nov 2018*) developed by the Joint Research Centre (JRC), which belongs to the Europian Union (EU). This software accept images from several satellites such as Sentinel-1, it does the land masking and it analyses the input images by applying a Constant False Alarm Rate (CFAR), calibrated with the sea clutter K distribution which makes it suitable for a wide variety of background. At the same time, configuration plays an important role in accurate detection. Besides, this software can fuse information with the AIS (Automatic Identification System), LRIT (Long-Range Identification and Tracking) or VMS (for fishing ships) [26].

There are other papers studying this software and it is found that the performance of SUMO depends on the conditions where it is applied. The algorithm has problems near the coastal line because of masking. And if compared to the other algorithms, it continues to be better than softwares developed by other agencies [33]. So, from this research we can conclude that SUMO needs one of the masking algorithms proposed for the SAR images in order to improve. Also, the performance of the algorithm is optimal so even mass detection can be done [34].

## VII. Ship Detection integrated with AIS

The usage of AIS can turn more interesting for other applications when image detection systems are boosted with AIS, since the usage of both can lead to a big data approach where several type of studies can be obtained, such as economical activities and IUU activities. For example: a ship may have its AIS system on until it gets into a restricted zone where it turns it off and so, alerts of this type can be obtained and a follow up of the ship could be performed (with its limits) using satellite imagery. Concerning this approach, there are private companies that are already capitalizing the advantages of this joint information such as CLS (Collective Localisation Satellites) working with researchers and developing an application that would be able to analyze, fuse this information and throw different type of alerts to its users [35].

Furthermore, as mentioned before, MSI images and SAR images can have their own advantages and disadvantages, but neither of both is able to perform an optimum satellite image processing for ship detection in solitude. So another option is to use a combination of both and even another satellite technology called Hyperspectral imaging, as proposed in [36]. For example: Hyperspectral detection algorithm, maximum likelihood classifier (for MSI) and CFAR scheme for SAR. This opens opportunities to a precise follow up of a vessel, by being able to have an image every 3h and/or being able to obtain data (image) from each source.

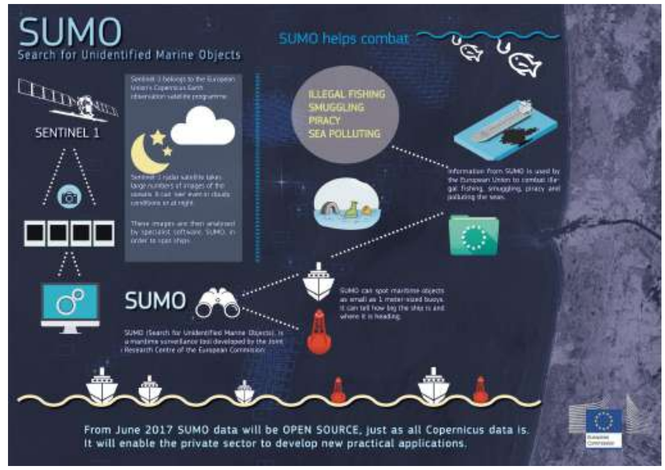



Fig. 5. SUMO presentation slide (https://phys.org/news/2017-10-jrc-ship-detection-software-source.html *accessed Nov 2018*)

## VIII. Private Sector

In order to understand better what could be the most accurate methods, we will see what the private sector is doing in order to approach this problem for more practical feasible. As an international enterprise, Telespazio is using machine learning paradigms developed by Orbital Insight in order to process the satellite data (http://www.telespazio.com/-/briefing_paris *assessed Nov 2018*). Similarly, the IRT Saint Exupery (research center) is implementing big data and artificial intelligence for same diverse purposes including ship detection and follow up and also interesting, commercial trading and economical indicators based on this information.
Apart from SUMO, the European Commission, Joint Research Center (JRC) has developed other tools that allow massive analysis [34] and big data approaches including fusing AIS data [37]: SUMO + JRC Earth Observation Data and Processing Platform (JEODPP) in order to improve performance of analysis + Blue Hub to fuse data: AIS and SAR. As a note, this software can be used freely by European institutions.

## IX. Discussion

In this letter, we have briefly explained different algorithms used in the entire process of ship detection. We have observed that the process consists majorly of four steps: Preprocessing, Candidate Extraction, Postscreening and discrimination. The discriminating step is performed so as to minimize the probability of false alarms being recognized as ships, whereas Candidate Extraction is performed so as to maximize the probability of ships to selected as candidate regions. We have also seen that MSI images are largely, less researched than SAR images, and are less practically used. However, there is an increase in the research on MSI images in recent years because of availability of more MSI sensors. We also discussed about the fusion of AIS, VMS with Ship detection in order to serve as a bigger ship identification process. Apart from using SAR and MSI images in solitude, we also saw the use of a combination of algorithms used on different types of images for the same region, leading to better results. Finally, we also discussed the development of an automated software called SUMO, for ship detection using SAR images.

After this research, we can get the image of a system benefiting from all of the resources (SAR, MSI, AIS) complimented with data analysis in order to develop a robust automated software where images interpretation by a user would be reduced to a minimum and alerts would allow preventive and corrective actions. This as a possible application for IUU activities, and thinking beyond, economical and commercial highlights. Of course, this system would be an application so need of a database, back-end and front-end, and algorithms for data analysis as the SESAME framework proposed.

Nevertheless, access to the resources may be limited mainly because of the cost, and developing algorithms for this proposal imply time and a possibly a strong investment. So, as a starting point, we could approach from the most economical and high impact solution which is SAR: independent of sunlight and weather, and free images available. There is also an already existing software (SUMO) and to which improvements could be done for land-masking and discrimination so for the procedures before and after the detector algorithm is applied. In order to tackle this parts, deep learning algorithms are advised. Even if SUMO is not taken into account, different deep learning based algorithms would be the answer. Furthermore, for MSI images process, same recommendation: deep learning algorithms outperforms the others.

Lastly, the different methods studied in this paper are not directly compared since each uses different images and different precision evaluation methods. This issue suggests that a work flow could be proposed: a set of images to test with the different algorithms and the same accuracy evaluation method.

## REFERENCES


[1] Urka Kanjira, Harm Greidanusb,, Kritof Otirc,*"Vessel detection and classication from spaceborne optical images: A literature survey "*,Remote Sensing of Environment 207 (2018) 126

[2] Chunping Qiua, Michael Schmitta, Xiao Xiang Zhua,b,*"Towards automatic SAR-optical stereogrammetry over urban areas usingvery high resolution imagery"*,C. Qiu et al. / ISPRS Journal of Photogrammetry and Remote Sensing 138 (2018) 218231

[3] Alberto Moreira, Pau Prats-Iraola, Marwan Younis,Gerhard Krieger, Irena Hajnsek, and Konstantinos P. Papathanassiou, *"A tutorial on Synthetic Aperture Radar"*, Microwaves and Radar Institute of the German Aerospace Center (DLR),

[4] Ting Nie 1,2,*, Bin He 2, Guoling Bi 2, Yu Zhang 2 and Wensheng Wang 1,2 ,*"A Method of Ship Detection under Complex Background"* ,ISPRS Int. J. Geo-Inf. 2017, 6, 159

[5] Zhang, D.; Ni, Q.; Fang, D.; Li, J.H.; Yao, W.; Yuan, S.G.*" Application of multispectral remote sensing technology in surface water body extraction."* In Proceedings of the 2016 International Conference on Audio, Language and Image Processing, Shanghai, China, 1112 July 2016

[6] Y. Xia, S. H. Wan, and L. H. Yue, *"A novel algorithm for ship detection based on dynamic fusion model of multi-feature and support vector machine"*, in Proc. ICIG, 2011, pp. 521-526.

[7] 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 Trans. Geosci. Remote Sens., vol. 48, no. 9, pp. 3446-3456, Sep. 2010.

[8] N. Otsu, *A threshold selection method from gray-level histograms*, IEEE Trans. Syst., Man, Cybern., vol. SMC-9, no. 1, pp. 6266, Jan. 1979.

[9] R. C. Gonzalez and R. E. Woods,*" Digital Image Processing"*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 2002.

[10] F. T. Chan and L. Vese, *Active contours without edges*, IEEE Trans. Image Process., vol. 10, no. 2, pp. 266277, Feb. 2001.

[11] F. Chen, W. Yu, X. Liu, K. Wang, L. Gong, and W. Lv, *Graph-based ship extraction scheme for optical satellite image*, in Proc. IEEE IGARSS, 2011, pp. 491-494.

[12] Yuri Y. Boykov,Marie-Pierre Jolly,*"Interactive Graph Cuts for Optimal Boundary Region Segmentation of Objects in N-D Images"* ,Proceedings of Internation Conference on Computer Vision, Vancouver, Canada, July 2001 vol.I, p. 105

[13] J. Antelo, G. Ambrosio, J. Gonzalez alez, and C. Galindo, *Ship detection and recognition in high-resolution satellite images*, in Proc. IEEE IGARSS, 2009, pp. 514-517.

[14] ManjunathB.S. and SumengenB., *Graph partitioning active contours(GPAC) for image segmentation*, IEEE Trans.PatternAnal.Mach.Intell.,vol.28,no.4,2006

[15] C. Corbane, L. Najman, E. Pecoul, L. Demagistri, and M. Petit, *A complete processing chain for ship detection using optical satellite imagery*, Int. J. Remote Sens., vol. 31, no. 22, pp. 5837-5854, Jul. 2010.

[16] P. Salembier and J. Sem,*"Flat Zones Filtering, Connected Operators, and Filters by Reconstruction "*,IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 4, NO. 8, AUGUST I995

[17] NAJMAN, L. and COUPRIE, M., 2006, *"Building the component tree in quasi-linear time"*, IEEE Transactions on Image Processing, 15, pp. 35313539.

[18] C. Guo, Q. Ma, and L. Zhang, *Spatio-temporal saliency detection using phase spectrum of quaternion Fourier transform,* in Proc. CVPR, 2008, pp. 1 8.

[19] S. Qi, J. Ma, J. Lin, Y. Li, and J. Tian, *Unsupervised ship detection based on saliency and S-HOG descriptor from optical satellite images*, IEEE Geosci. Remote Sens. Lett., vol. 12, no. 7, pp. 1451-1455, Jul. 2015.

[20] F. K. Bi, B. C. Zhu, L. N. Gao, and M. M. Bian, *A visual search inspired computational model for ship detection in optical satellite images*, IEEE Geosci. Remote Sens. Lett., vol. 9, no. 4, pp. 749-753, Jul. 2012.

[21] D.Lowe,*Distinctive image features from scale-invariant keypoints*, Int. J. Comput. Vis., vol. 60, no. 2, pp. 91110, Nov. 2004.

[22] M. Bar, *Visual objects in context*, Nat. Rev. Neurosci., vol. 5, no. 8, pp. 617629, Aug. 2004.

[23] Zou, Z., Shi, Z., 2016. *"Ship detection in Spaceborne optical image with SVD network"*, IEEE Trans. Geosci. Remote Sens. 54, 58325845. http://dx.doi.org/10.1109/TGRS. 2016.2572736.

[24] Nadia Proia and Vincent Pag,*"Characterization of a Bayesian Ship Detection Method in Optical Satellite Images "*, IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, VOL. 7, NO. 2, APRIL2010

[25] Yifan Li, Huadong Zhang, Qianfeng Guo, Xiaoshi Li,*"Machine Learning Methods for Ship Detection in Satellite Images "* ,write down Germany, IEEE, 2013

[26] Harm Greidanus *, Marlene Alvarez, Carlos Santamaria, Franois-Xavier Thoorens, Naouma Kourti and Pietro Argentieri, *"The SUMO Ship Detector Algorithm for Satellite Radar Images"*, Remote Sensing, 2017

[27] Armando Marino, Maria J. Sanjuan-Ferrer, Irena Hajnsek and Kazuo Ouchi, *"Ship Detection with Spectral Analysis of Synthetic Aperture Radar: A Comparison of New and Well-Known Algorithms"*, Remote Sensing, 2015

[28] Quanzhi An, Zongxu Pan and Hongjian You, *"Ship Detection in Gaofen-3 SAR Images Based on Sea Clutter Distribution Analysis and Deep Convolutional Neural Network"*, Sensors, 2018

[29] ShigangWang, Min Wang, Shuyuan Yang, and Licheng Jiao, *"New hierarchical Saliency Filtering for Fast Ship Detection in High-Resolution SAR Images"*, IEEE, Jan 2017

[30] R. G. V. Meyer, W. Kleynhans, C. P. Schwegmann, *"Small ships don't shine: classification of ocean vessels from low resolution, large swath area SAR acquisitions"*, IEEE, 2016

[31] E. Khesali, H. Enayati, M. Modiri, M. Mohseni Aref, *"Automatic ship detection in Single-Pol SAR Images using texture features in artificial neural networks"*, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-1/W5, 2015

[32] Yang Liu, Peng Xu, Miao-hui Zhang, Zheng-wei Guo, *"SAR ship detection using Sea-Land Segmentation-Based Convolutional Neural Network"*, IEEE, 2017

[33] Mattia Stasolla, Jordi J. Mallorqu, Member, IEEE, Gerard Margarit, Carlos Santamaria, and Nick Walker, *"A Comparative Study of Operational Vessel Detectors for Maritime Surveillance Using Satellite-Borne Synthetic Aperture Radar"*, IEEE, JUNE 2016

[34] Carlos Santamaria, Marlene Alvarez, Harm Greidanus, Vasileios Syrris, Pierre Soille and Pietro Argentieri, *"Mass Processing of Sentinel-1 Images for Maritime Surveillance"*, Remote Sensing, 2017

[35] Ronan Fablet, Nicolas Bellec, Laetitia Chapel, Chlo Friguet, Ren Garello, et al.., *"Next Step for Big Data Infrastructure and Analytics for the Surveillance of the Maritime Traffic from AIS Sentinel Satellite Data Streams."* BiDS 2017 - Conference on Big Data from Space, Nov 2017, Toulouse, France. pp.1-4. ¡hal-01671323¿

[36] Kyung-Ae Park, Jae-Jin Park, Jae-Cheol Jang, Ji-Hyun Lee, Sangwoo Oh, and Moonjin Lee, *"Multi-Spectral Ship Detection Using Optical, Hyperspectral, and Microwave SAR Remote Sensing Data in Coastal Regions"*, Sustainability, MDPI, 2018

[37] P. Soille *, A. Burger, D. De Marchi, P. Kempeneers, D. Rodriguez, V. Syrris, V. Vasilev,*"A versatile data-intensive computing platform for information retrieval from big geospatial data"*, Future Generation Computer Systems, Elsevier, 2017