# VESSEL DETECTION USING IMAGE PROCESSING AND NEURAL NETWORKS

*Konstantina Bereta*[1], *Raffaele Grasso*[2], *Dimitris Zissis*[13]

[1]MarineTraffic, Athens, Greece
[1]o Konstantina.bereta@marinetraffic.com
[2]NATO-STO-CMRE, La Spezia, Italy
[2]raffaele.grasso@cmre.nato.int
[3]University of the Aegean, Ermoupoli, Syros, Greece
[3]dzissis@aegean.gr

## ABSTRACT

The establishment of the Automatic Identification System (AIS) was revolutionary for Maritime Situational Awareness, as it allowed for the tracking of vessels carrying an AIS transponder, which is mandatory for, and not limited to, the majority of the commercial fleet. Despite the benefits of the widespread use of AIS for navigational safety and global maritime security, one cannot depend only on AIS sources in order to obtain the complete maritime situational awareness picture. In this paper we describe a multistage data-centric workflow that integrates satellite optical imagery and AIS data for automatic vessel detection that builds on (i) image processing techniques and (ii) Convolutional Neural networks. The experimental evaluation of our approach shows that our framework achieves an accuracy greater than 95%.

***Index Terms***— Convolutional Neural Networks, MSA

## 1. INTRODUCTION

The establishment of the Automatic Identification System (AIS) was revolutionary for Maritime Situational Awareness (MSA), as it allowed for the tracking of vessels carrying an AIS transponder. In particular, vessels above 300 gross tonnage that are engaged on international voyages, cargo ships above 500 gross tonnage and passenger ships are obliged to have AIS transponders and transmit messages reporting their location and navigational information in intervals that vary from a few seconds to a few minutes depending on their navigational status and speed (e.g., underway using engine at high speed, stopped, etc.). Although the initial aim of the development and establishment of AIS as a standard was to avoid collisions between vessels, it proved to be a very useful source of information regarding the navigational status of the global fleet, especially for commercial vessels, raising the

Maritime Situational Awareness. Soon after the establishment of AIS, a lot of applications started to emerge that make use and analyse AIS data, eventually making available the enriched information to the users. An example of these applications is the MarineTraffic website[1], which presents live information about the navigational status of the global fleet, together with other, arbitrary information, such as weather data, piracy zones, port data, etc.

Tracking a vessel that has AIS installed might not always be possible, however, due to (i) limited AIS coverage, (ii) transponder malfunction, and/or (iii) intentional switch-off of the AIS transponder. More often than not, vessels that are about to engage in illegal activities switch off their AIS transponder intentionally and switch it back on when the activity is over or decide to spoof the information sent (i.e., transmit false location).

Therefore, the motivation of the work described in this paper is to use medium resolution satellite imagery in order to identify vessels that might have their transponders switched-off. The framework that we present relies on the processing of multi-spectral (MS) Sentinel-2 data.

The contributions of the work described in this paper to the state-of-the-art are the following: First, We describe a fully automatic large-scale analysis pipeline for data processing combining on-the-fly AIS data processing and compute insensitive satellite imagery processing tasks. Second, we develop a hybrid approach in the sense that it employs image processing and learning techniques at the same time. Image processing is used for (i) the extraction of features that will be classified by the trained CNN, (ii) the calculation of metrics, such as heading, width, length, speed class (low/high), location (coordinates). The estimation of these metrics makes the correlation with AIS more accurate (instead of using the time and location of vessels as the only criteria to match vessels depicted in a satellite image with their respective AIS signals).

This paper is structured as follows. In Section 2 we discuss related approaches in literature. In Section 3 we describe

[1]http://marinetraffic.com

the architecture of our framework that builds on (i) image processing techniques, and Convolutional Neural Networks, described in Section 4. In Section 5 we describe the experimental evaluation of our approach and, last, in Section 6 we conclude the paper and we also discuss future extensions.

## 2. RELATED WORK

Due to the fact that AIS data is not always available, several approaches that fuse satellite data with AIS data have been proposed in literature. However, the majority of these approaches use SAR imagery. For example, the approach introduced in [1] uses SAR data matched with AIS data via a simulated annealing process. The authors of [2] used a Constant False Alarm Rate (CFAR) algorithm to detect ships and oil platforms in TerraSAR images. The CFAR algorithm was also used in the case described in [3], for detecting ships in Sentinel-1 SAR imagery, as well as in the case of [4]. Another approach that uses both AIS and SAR satellite data is described in [5], which also targets the problem of the time difference between the acquisition time of the satellite image and the timestamp of the AIS signals by performing a spatiotemporal interpolation of the vessels to the image acquisition time and taking into account historical data and a similar problem was also addressed in [6]. [7, 8] presents an approach for multi-class vessel detection in VHR satellite images. We differentiate from this approach in that (i) we produce the training dataset automatically using AIS data, (ii) we employ a CNN only once, and (iii) we do not use VHR satellite imagery, which the approach proposed in [7] is suitable for. The solution described in [9] is also very relevant to our approach, as it performs image segmentation and calculation of metrics of the derived features. However, it does not use neural networks to perform classification of features.

In some of the aforementioned approaches, AIS data is used as ground truth, as in our case. Although the approach that we present in this paper works also for SAR data (e.g., Sentinel-1 data), in the context of this paper we emphasize in the use of multi-spectral optical imagery (i.e., Sentinel-2 data). The majority of documented approaches rely on semi-automatic data investigation and processing methods. Our approach relies on data-driven methods to accelerate the full path of data use, with data-stream workflows orchestrating on-the-fly data processing and analysis tasks. Also, in all of the discussed solutions employ either image processing algorithms or learning approaches (i.e., neural networks are used). In this paper we present a hybrid approach that uses both image processing and neural networks.

## 3. ARCHITECTURE

The workflow of our approach is depicted in Figure 1 is described as follows: First, we download a Sentinel-2 image
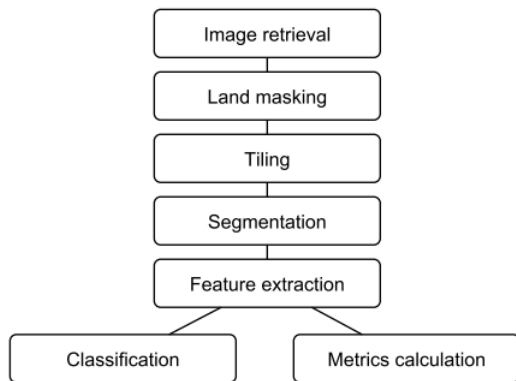


**Fig. 1**. Architecture overview

of the area of interest. The boundaries of the area are provided as a Well-known-text polygon. To perform this step we use the Python API for the Copernicus Open Access Hub[2], named SentinelSat[3]. Then, we perform *land masking*. We use an external source that contains the coastline geometries and we mask out the pixels of the image that intersect with these geometries. This step is performed so that we minimise the land parcels contained in the extracted features. Next, we split the land masked image into tiles, so that each tile is processed separately and *concurrently*. For each tile, we employ thresholding in order to distinguish the background (i.e., the sea), from the foreground (i.e., features). Best results were observed using a combination a variation of the Otsu filter (i.e., producing multiple thresholds) [10] and the Yen filter [11]. This process results in producing an initial pool of features for each tile. For each feature, we calculate a number of metrics using pixel-level calculations: width, length, heading, and the existence of waves (used to estimate the navigational status). Then, we classify the feature using the trained model of the CNN presented in Section 4. The reason why we use a CNN is because convolutional neural networks have significant results in image classification tasks [12]. If it is a vessel, we correlate it with the closes AIS signal spatially (vessel location), as well as temporally (e.g, taking into account AIS signal received around the acquisition time of the image). Apart from the spatio-temporal criteria, we also take into account the heading and the dimensions of the vessel, as calculated above. If the vessel depicted in the satellite image cannot be associated with an AIS message, then it is marked as a possible dark target.

For the training phase we use a labelled dataset automatically created by correlating MarineTraffic AIS daata with satellite imagery.

---

[2]https://scihub.copernicus.eu/dhus/
[3]https://sentinelsat.readthedocs.io/en/stable/#

## 4. CONVOLUTIONAL NEURAL NETWORKS

Figure 2 shows the neural network used in this work to built the ship target classifiers. In general, the network can process input data coming from $M$ different sources, or modalities, and integrate information at the feature level by means of a combination layer. Each branch of the system has a multi-layer convolutional block followed by a dense layer entering the combination layer. Dropout layers can be added to regularize the network. The combination layer is followed by a multi-layered fully connected sub-net with $K$ outputs, where $K$ is the number of target classes. These outputs provide the probits to the final max-rule layer to choose the target class having the maximum probability. A prior probability can be associated to the weights of all the layers of the network, or part of them, and the network is trained by using a Bayesian variational inference method by maximizing the evidence lower bound (ELBO) loss function [13].
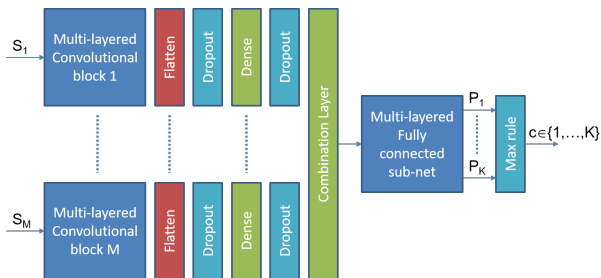


**Fig. 2**. General structure of the CNN

The specific network used to classify ship targets from MSI images has two branches, one for each resolutions, $\Delta R = 10m$ and $\Delta R = 20m$, considered as two different acquisition modalities. The image channels having the same resolution are assigned to a branch of the network. The system structure has a high degree of flexibility and can be configured as needed by including or excluding branches, depending on the available modalities, and considering different channel combinations. In this work, the full multi-modal structure is tested, while in future work, the performance of different system configurations will be compared to evaluate the minimum number of branches and channels needed to achieve a given level of performance.

## 5. EXPERIMENTS

The results of the experiments on two classifiers are reported below. In particular, the first classifier discriminate between static and non-static ships while the second one classifies the type of the ship. A third classifier, not reported in this paper, has been trained to discriminate between ship and no-ship targets, achieving a performance similar to the other two.

A labeled data set of about 6000 ship target multi-spectral images has been used to train the classifiers. Each image has

a size of 46x46 pixels and includes the MSI channels at 10m and 20m of resolution. A set of attributes, such as ship type and speed over ground (SOG), is associated to each ship to select targets having given properties. The data set is representative for targets of type "Cargo" and "Tanker" with about 3000 and 2000 contacts per type, respectively. Concerning the navigational status, ships under way are roughly 3800 while those observed at anchor are about 2000. The data set has been automatically labeled by fusing image data with AIS data, as described in [14].

### 5.1. Ship SOG classifier

The first classifier is trained to distinguish static or almost static ships from sailing ships by learning the features associated to the ship wake. The training data set contains 4042 ship images. The static class is obtained by considering ships having SOG less than 1kn, while the ships having the SOG greater than 2kn are included in the sailing ship class. The number of samples of the first class is 1761 while the samples of the second class are about 2281. The length of the ships included in the training data set is greater than 100m. The 80% of the data set is used to train the classifier, while the remaining 20% is used for the validation.

### 5.2. Ship type classifier

The second classifier discriminates the type of ship among three broad classes: "cargo", "tanker", "other". The ship type attribute associated to a ship image is used to build the training data set from the initial one. The number of samples of the three classes is 3066, 2071 and 1231, respectively, for a total of 6368 samples. The data set is split in 80% training set and 20% validation set. Figure 3 shows the scatter plots of the features learned by the classifier and calculated on the training set. Even in this case, the three classes are very well separated allowing to reach a classification accuracy of about 97% for both the training and the validation sets.
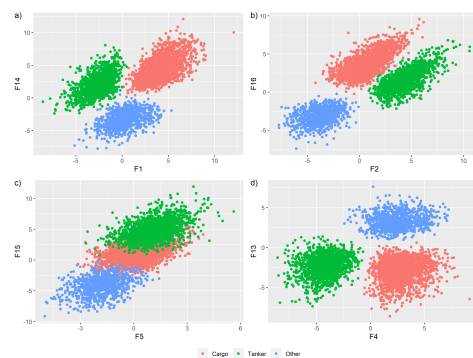


**Fig. 3**. Ship type classifier scatter plots

## 6. CONCLUSIONS AND FUTURE WORK

In this paper, we presented an approach for vessel detection and classification in Sentinel-2 multi-spectral images. Our approach employs image processing techniques for image segmentation, feature extraction and calculation of metrics for all identified features. For the classification task, we constructed a Convolutional Neural Network and trained it using AIS data as ground truth. The produced model is able to decide whether a feature illustrated in a satellite image is a vessel or not. We evaluated our approach using a multi-modal CNN trained by maximizing the variational ELBO, achieving an accuracy that is greater than 95%. Future directions of the proposed solution are the following: (i) definition of classifier confidence metrics by using the learned CNN weight distributions and (ii) benchmarking different configurations of the network to evaluate the optimal combination of branches and channels needed to achieve optimal performance.

## 7. REFERENCES

[1] Mazzarella, Fabio and Alessandrini, Alfredo and Greidanus, Harm and Alvarez, Marlene and Argentieri, Pietro and Nappo, Domenico and Ziemba, Lukasz, "Data fusion for wide-area maritime surveillance," 06 2013.

[2] Paes, Rafael and Lorenzzetti, Joao and Gherardi, Douglas, "Ship detection using terra-sar-x images in the campos basin (brazil)," *Geoscience and Remote Sensing Letters, IEEE*, vol. 7, pp. 545–548, 07 2010.

[3] Grover, A and Kumar, Shashi and Kumar, A, "Ship detection using sentinel-1 sar data," *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. IV-5, pp. 317–324, 11 2018.

[4] Achiri, Lotfi and Guida, Raffaella and Iervolino, Pasquale, "Sar and ais fusion for maritime surveillance," 09 2018.

[5] F. Mazzarella and M. Vespe and C. Santamaria, "Sar ship detection and self-reporting data fusion based on traffic knowledge," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 8, pp. 1685–1689, Aug 2015.

[6] Fan Wu and Chao Wang and Hong Zhang and Bo Zhang, "Vessel detection and analysis combining SAR images and AIS information," in *2016 IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2016, Beijing, China, July 10-15, 2016*. pp., 6633–6636, IEEE.

[7] Sergey Voinov and Frank Heymann and Ralf Bill and Egbert Schwarz, "Multiclass vessel detection from high resolution optical satellite images based on deep neural networks," in *2019 IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2019, Yokohama, Japan, July 28 - August 2, 2019*. pp., 166–169, IEEE.

[8] Sergey Voinov and Detmar Krause and Egbert Schwarz, "Towards automated vessel detection and type recognition from VHR optical satellite images," in *2018 IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2018, Valencia, Spain, July 22-27, 2018*. pp., 4823–4826, IEEE.

[9] Marco Gianinetto and Martina Aiello and Andrea Marchesi and Francesco Topputo and Mauro Massari and Riccardo Lombardi and Francesco Banda and Stefano Tebaldini, "OBIA ship detection with multispectral and SAR images: A simulation for copernicus security applications," in *2016 IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2016, Beijing, China, July 10-15, 2016*, pp., 1229–1232.

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

[11] Jui-Cheng Yen and Fu-Juay Chang and Shyang Chang, "A new criterion for automatic multilevel thresholding," *IEEE Transactions on Image Processing*, vol. 4, no. 3, pp. 370–378, March 1995.

[12] D. Ciregan and U. Meier and J. Schmidhuber, "Multicolumn deep neural networks for image classification," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp., 3642–3649.

[13] Bishop, Christopher M., *Pattern Recognition and Machine Learning*, Springer, 2006.

[14] Raffaele Grasso, "Ship classification from multispectral satellite imaging by convolutional neural networks," in *27th European Signal Processing Conference, EUSIPCO 2019, A Coruña, Spain, September 2-6, 2019*, pp., 1–5.