# Enhancing the Maritime Monitoring by Deep Learning

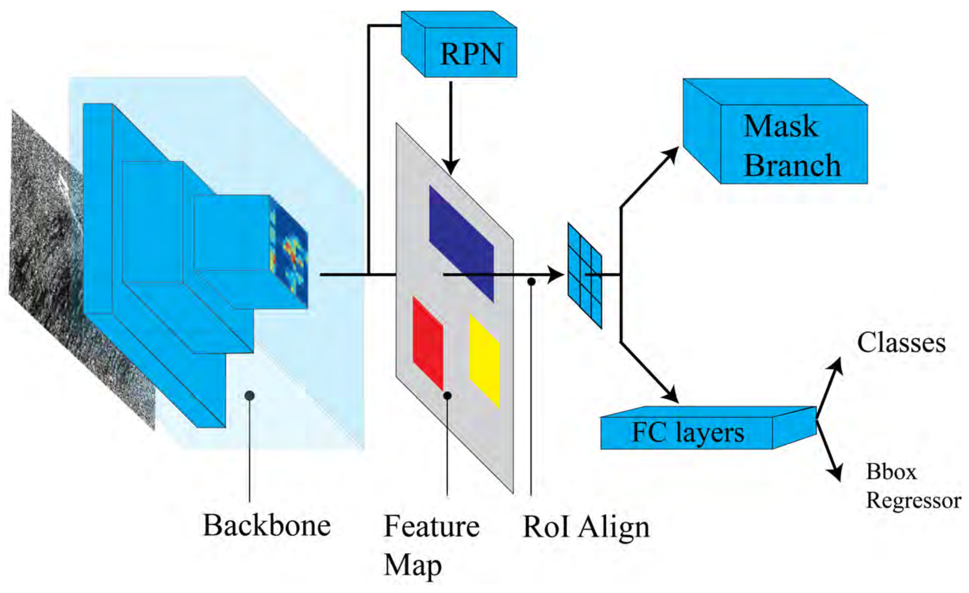
Maria Daniela Graziano, Roberto Del Prete, Alfredo Renga

## Abstract

Artificial intelligence (AI) is paving the way to fully take advantage of the growing amount of Earth Observation (EO) data, mainly driven by the free distribution policy of the Copernicus programme. The rise of Deep Learning (DL) led the possibility of automatically gleaning relevant information from remote sensed images, enhancing the understanding of Earth's complex and wide-scale dynamics. Along with the increase of computational power, convolutional neural networks (CNNs) have indeed emerged as a particularly powerful tool outperforming human experts on conventional tasks. This also allowed a wide variety of unprecedented applications to large set of fields in which image processing is fundamental, from multimedia to medicine and robotics. In the framework of Earth Observation, CNN-based approaches promote extensive application to image understanding (e.g., land use/land cover classification, image retrieval, change detection, semantic labeling), image restoration (e.g., enhancement, denoising, estimation problems) and data fusion. The capability to locate the region of interest (ROI) containing the potential targets (e.g. ships) and recognize those target signatures with high accuracy and efficiency without any manual intervention is referred as automatic target recognition (ATR). This task holds a crucial role into enhancing maritime domain awareness (MDA) preventing illegal, irregular and unreported (IIU) fishing, smuggling and ensuring borders and marine protected areas (MPA) control. Beside multi and hyperspectral sensors, the microwave detection tools offered by Synthetic Aperture Radar (SAR) plays a pivotal role due to its all-weather all-time capabilities. The underlying electromagnetic scattering mechanism and coherent imaging system produce unique target characteristics in the resulting images. This enables the gathering of ship wakes when imaging sea surface in which high returns are caused by enhanced surface roughness or large wave steepness, either via specular reflection or Bragg scattering. Since ship wakes extend for kilometers, wakes are widely employed for ship detection and route estimation from satellite imagery. It is worth noting that the actual maritime monitoring systems rely upon the AIS (Automatic Identification System) transceivers equipped on-board ships. Nevertheless, IMO (International Maritime Organisation) regularizes the usage of AIS only for specific classes of ships (> 300 tons). Besides, this collaborative system can be intentionally turned off during unlawful activities, which according to Frontex occur most of the time night-time. Given this analysis, the duo satellite images and DL can provide better treats at sea. While SAR ship detection by DL is emerging and datasets are being released, the problem of ship wake detection has barely been approached. However, ship wakes can provide useful information even in cases when ships are not imaged (e.g. GO-fast vessels) by the coherent SAR system due to their fiberglass/rubber material and high speed of travelling. The traditional methods assume the wake structure composed by linear feature, therefore the problem is addressed with domain transform as Radon or Hough. Stacking together convolutional and pooling layers, CNN-based methods do not rely on this assumption and can profit of the higher spatial resolutions granted by the existing and future generation of satellites to extract low-, mid- and high-level features. In this framework, an experimental analysis of deep learning methods for automatic wake detection in SAR images will be shown, starting from the preparation of the dataset, inspecting Sentinel-1 images,to the characterization of several networks, until the discussion of results and recommendations for future developments. This work aims at stimulating more attention in this interesting but still under-investigated field of research, discussing methodology and applicability of the proposed implementations.

## 1. Methodology

Contrarily to the traditional programming paradigm, a CNN autonomously learns the most valuable patterns to classify a target by extracting weights and biases from sets of data; the process is referred as supervised learning. However, the encoder architecture only allows to classification task, to achieve object detection the CNN architecture must be broadened. Several architecture have been proposed in literature, the most popular detectors are the R-CNN, based on a two-stage proposal-driven mechanism. The first stage generates a sparse set of candidate object locations and the second one classifies these locations trough the CNN. This two-stage framework has been submitted to a sequence of advances (e.g. Faster R-CNN, Mask R-CNN, Cascade Mask R-CNN) achieving top accuracy on the COCO benchmark. In contrast, one-stage detectors (e.g. YOLO and SSD) are applied over a regular, dense sampling of object locations, scales and aspect ratios. Usually, this class of detectors privileges the computing speed trailing in accuracy even with a larger compute budget.




## 2. Dataset Preparation

As transfer learning is a common practice in DL, the model training has been performed fine-tuning the pre-trained COCO weights on our custom dataset. The dataset has been manually constructed with more than 250 wake chips extracted from Interferometric Wide (IW) swath Sentinel-1 SAR images. The images are obtained from Copernicus Open Access Hub and correspond to Level-1 Ground Range Detected (GRD) products with a pixel spacing of about 10 m x 10 m (ground range x azimuth). All images are gathered in vertical polarization to enhance the wake appearance. The dataset contains mainly turbulent and narrow-V wakes. Very few samples of Kelvin wakes are distinguishable and not included both in annotation and training of the network. Consequently, the proposed approach does not enable the detection of the Kelvin wake. Data labelling has been performed in MS COCO format specifically excluding ships from polygons. This is purposely done because the proposed methodology aims to recognize wakes also when ships are not mapped in the SAR image, due to the coherent mechanism of SAR focusing. Besides, also practical issues prevent the combined ship and wake labelling such as the azimuth-displacement. Class imbalance problems were not faced but the wake detection problem is still intrinsically tricky to solve due to the coherent speckle-noise always present both in the clutter and in the wake.

## 3. Model Training

Model training has been performed for 5000 epochs using the standard SGD (Stochastic Gradient Descent) optimizer with a small learning rate of 0.001. A hook function has been implemented to evaluate the models on the test set at each 50 epochs. Our training strategy indeed involves early-stopping the model to avoid overfitting. Due to the limited amount of annotated date, for crack this problem also a data augmentation pipeline (i.e., resizing, shifting, flipping, and rotating) has been adopted.

Other relavant parameters related to the Region Proposal Network (RPN) are listed in the Table 1.

Table.1 RPN training parameters

| | |
|---|---|
| BATCH_SIZE_PER_IMAGE | 256 |
| BBOX_REG_LOSS_TYPE | smooth_l1 |
| BBOX_REG_LOSS_WEIGHT | 1.0 |
| BBOX_REG_WEIGHTS | (1.0, 1.0, 1.0, 1.0) |
| BOUNDARY_THRESH | -1 |
| HEAD_NAME | StandardRPNHead |
| IN_FEATURES | ['res4'] |
| IOU_LABELS | [0, -1, 1] |
| IOU_THRESHOLDS | [0.3, 0.7] |
| LOSS_WEIGHT | 1.0 |
| NMS_THRESH | 0.7 |
| POSITIVE_FRACTION | 0.5 |
| POST_NMS_TOPK_TEST | 1000 |
| POST_NMS_TOPK_TRAIN | 2000 |
| PRE_NMS_TOPK_TEST | 6000 |
| PRE_NMS_TOPK_TRAIN | 12000 |
| SMOOTH_L1_BETA | 0.0 |

## 4. Preliminary Results

Some preliminary results of the proposed methodology are here reported. The segmentation masks are completed by the confidence score and bounding box, output of the instance segmentation algorithm.