

# JOURNAL OF SCIENCE



SAKARYA UNIVERSITY

## Sakarya University Journal of Science

ISSN 1301-4048 | e-ISSN 2147-835X | Period Bimonthly | Founded: 1997 | Publisher Sakarya University |  
<http://www.saujs.sakarya.edu.tr/>

Title: Ship Target Classification in Satellite Images using Deep Convolutional Neural Networks

Authors: Ferhat Ucar, Deniz Korkmaz

Received: 2019-07-05 16:27:51

Accepted: 2019-12-03 10:28:00

Article Type: Research Article

Volume: 24

Issue: 1

Month: February

Year: 2020

Pages: 197-204

How to cite

Ferhat Ucar, Deniz Korkmaz; (2020), Ship Target Classification in Satellite Images using Deep Convolutional Neural Networks . Sakarya University Journal of Science, 24(1), 197-204, DOI: 10.16984/saufenbilder.587731

Access link

<http://www.saujs.sakarya.edu.tr/tr/issue/49430//587731>

New submission to SAUJS

<http://dergipark.gov.tr/journal/1115/submission/start>

## A Ship Detector Design Based on Deep Convolutional Neural Networks for Satellite Images

Ferhat Ucar<sup>\*1</sup>, Deniz Korkmaz<sup>2</sup>

### Abstract

Ship detection and classification systems from satellite images are challenging tasks with their requirements of feature extracting, advanced pre-processing, a variety of parameters obtained from satellites and other types of images, and analyzing of images. The dissimilarity of results, enhanced dataset requirement, the intricacy of the problem domain, general use of Synthetic Aperture Radar (SAR) images and problems on generalizability are some topics of the issues related to ship detection. In this study, we propose a Deep Convolutional Neural Network (DCNN) model for detecting the ships using the satellite images as inputs. Our model has acquired an adequate accuracy value by just using a pre-processed satellite image with a deep learning model built from scratch. The designed CNN model is constructed with a plain and easy to implement form in particular to the preferred satellite image set. Visual and graphical results show that the proposed model provides an efficient detection process with an accuracy of 99.60%.

**Keywords:** deep convolutional neural networks, ship detection, remote sensing, satellite imagery

### 1. INTRODUCTION

Ship detection and classification in remote sensing imagery is one of the important issues for many different military and civilian marine applications such as coast guard, dynamic traffic monitoring, ship rescue, and fishery management [1]–[3]. The increasing interest in ship detection systems makes it more attractive for developing marine systems. Optimization of the routes used by marine transportation companies provides reducing costs. Illegal activities are also followed and national security can be more effective. Sea pollution can be detected with the help of

environmental monitoring. Also, restricted fishing areas can be better controlled by fishery management systems using remote sensing in marine engineering [4].

There are many different methods that can collect knowledge on ship activities. Commonly, these can be categorized into two stages as cooperative and non-cooperative systems. First is that ships give information about themselves. Latter defines that monitoring systems are used to collect the information without any cooperation with ships. While cooperative systems are effective tools for observation, these are only sent a partial situation and ships with illegal activities can turn off these

<sup>\*</sup> Corresponding Author: fucar@firat.edu.tr

<sup>1</sup> Firat University, Department of Electrical and Electronics Engineering, Elazig, Turkey. ORCID: 0000-0001-9366-6124

<sup>2</sup> Malatya Turgut Ozal University, Department of Electrical Engineering, Malatya, Turkey. ORCID: 0000-0002-5159-0659

systems. Also, this data has limited access because of governmental policies [5]. Therefore, such systems are not reliable in some situations for marine applications. From the point of non-cooperative systems in marine engineering, object detection and classification systems have been one of the most popular topics and many researchers have studied in this field and developed different methods in the last decades [5]–[9].

Ship detection has more challenges due to long and thin shapes of ships. They can be surrounded by complex structures such as other ships and/or ports. Furthermore, detection and classification systems need to analyze and simultaneously process big amounts of incoming remote sensing data. Most approaches use the coarse-to-fine structures, which contain ship candidate extraction and false elimination [9], [10]. The difference between these approaches is the format of the computation distinction. While some approaches determine the distinct patterns [11], others search regions of distinct intensity [10], [12]. Although both types give better performance, they may be unsuccessful in complex scenes. These approaches also require carefully tuned parameters and they are still open to improvement. In recent years, Convolutional Neural Networks (CNNs) have provided significant achievements on the processing of remote sensing images. CNNs can easily learn features from labeled data and their implementations are easy to build. They have also great generalization abilities [3], [13], [14].

In this paper, a deep learning model for ship satellite image detection is proposed. The designed model uses a deep CNN based classifier. In contrary to existing literature, our model uses satellite images as inputs instead of Synthetic Aperture Radar (SAR) images. Working with easy to handle dataset makes our model appropriate for embedded real time applications. The proposed intelligent model of ship detection configuration operates the raw images without any handcrafted feature extraction process. Therefore, the proposed model has advantages in applicability and computational complexity avoiding the coarse-to-fine structure.

Furthermore, the selected dataset includes real-world remote sensing satellite images obtained from sea and land cover scenes.

This paper is organized as follows: In section 2, the methodology of deep CNNs is given briefly. Section 3 presents the ship satellite images dataset in detail. Proposed CNN model architecture and experimental results are analyzed and illustrated in section 4. Finally, a brief conclusion is provided in the last section.

## 2. OUTLINE OF THE METHODOLOGY

CNNs have been preferred for various image data processing applications in two-dimensional form. Because CNNs are formed by the inspiration of the biological processes of the animal visual cortex, there are various neurally-inspired models in the literature [13].

The main logic of a convolutional network relies on the three main neural layer structures, which are described as convolutional layers, pooling layers, and fully connected layers. Every layer in the structure plays its own role in a different way [15].

Figure 1 shows a general scheme of a common CNN architecture for the ship image detection process with layer details. In the training process of the network, there are two main steps called forward and backward stages. The main aim in the forward stage is to define the input image with weight and bias parameters in each layer. Then, the output computes the loss cost with the test labeled data. In the second step, which is the backward stage, computed loss cost is used to achieve the gradients of each parameter based on chain rules. After the updating process based on the gradients, all parameters are operated for the following forward stage. When a sufficient number of forward and backward iteration is reached, the learning of the network can be finished [13].

According to the layer-based structure of the common CNN architecture, there are six special layers as convolution, rectified linear unit (ReLU), batch normalization, pooling, softmax, and fully connected layers.

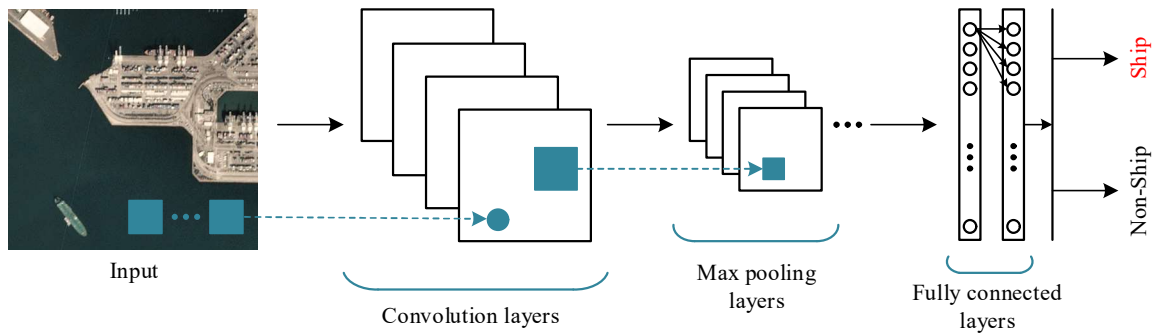


Figure 1. A common CNN architecture scheme

In the convolution layer, a selected kernel matrix is used for the convolution process and thus, a matrix multiplication operation between the kernel and the portion of the input image is the principle of the layer. A convolution layer generates the same number of feature maps as its kernels. In the convolutional layer, there are four hyper parameters as filter size ( $K$ ) with spatial extent ( $F$ ), stride ( $S$ ) and padding ( $P$ ). If the first layer of convolution has a size of  $W_1 \times H_1 \times D_1$  with the hyper parameters, then the second layer of the convolution occurs as  $W_2 \times H_2 \times D_2$  and given as follows;

$$\begin{aligned} W_2 &= \frac{(W_1 - F + 2P)}{S} + 1, \\ H_2 &= \frac{(H_1 - F + 2P)}{S} + 1, \\ D_2 &= K. \end{aligned} \quad (1)$$

Here, the  $W$ ,  $H$ , and  $D$  are the width, height, and depth of the convolutions. The width and height values are calculated equally by symmetry. ReLu layer performs an activation function operation which may be described as a threshold operation for each element of the input ( $x$ ) set as;

$$R(x) = \max(0, x). \quad (2)$$

ReLu layer sets the values to zero which are less than zero. The batch normalization layer also performs the normalization of its input channels using a mini-batch. To achieve speed training and to reduce the sensitivity to the initialization of the network, batch normalization layers are located between the convolution layers and nonlinearities, such as ReLu layers. The pooling layer is used for reducing the size of the feature

maps by means of a down sampling. There are two types of pooling operations as max pooling and average pooling. Both of them use a rectangular pooling region where max pooling computes the maximum and the average pooling computes the average value. The pooling layer also enhances the network robustness. In the softmax layer, a softmax activation function is applied to the input where the output vector represents the probabilities of the possible outcomes. The predicted probability of  $y=j$  can be defined as:

$$P(y = j | x) = e^{x^T w_j} / \sum_{k=1}^K e^{x^T w_k}, \quad (3)$$

where  $x$  is an input vector and  $w$  shows the weight vector. In a classification process, softmax is the final function which is followed by a cost computation. It is noted that fully connected layer operation is the same as conventional neural networks and it multiplies the inputs by weight parameters beside a bias vector [15], [16].

### 3. SHIP SATELLITE IMAGES DATASET

The selected dataset consists of 2 categories as “Ship” and “Non-Ship (NShip)” with 4000 80×80 RGB images. In the dataset, class label Ship contains 1000 images, centered on the ship body. NShip class also contains 3000 images. NShip class consists of three parts. The first part images are a random sampling of different land cover without any portion of a ship. The images of the second part are partial ships which include only some portion of a ship. Those samples are not enough to clarify the Ship class. The last part of the images includes bright pixels or strong linear

features [17]. According to the distribution of the NShip class, the second and last parts are intentionally given to provide the confusing operation of the classifier in the training process. In the proposed ship detection model, cropped  $80 \times 80$  images are used as training inputs. Selected sample images for each class are shown in Figure 2.

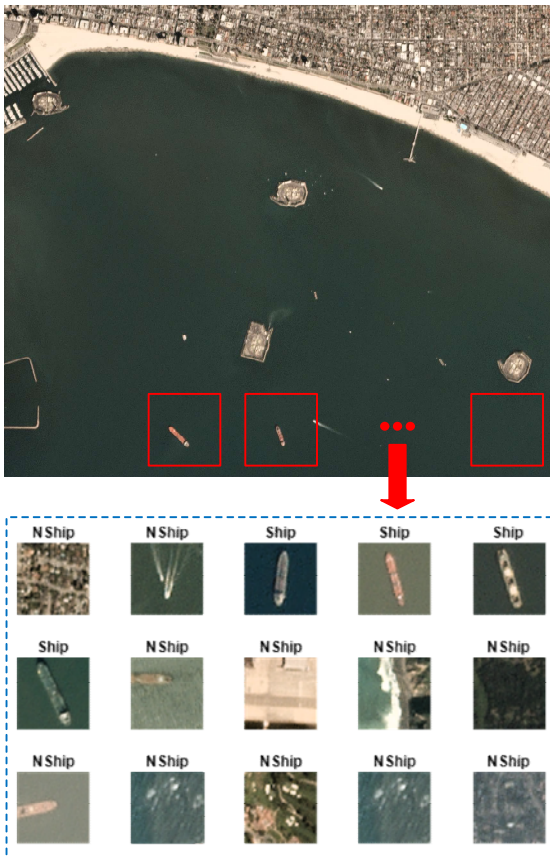


Figure 2. Sample images of ship satellite dataset

In the satellite dataset, regions of the ships have small, thin and sharp parts with cloudy and sea wave scenes. Furthermore, the distribution of ships is sparse on the sea and around the ports. These dataset specifications make the detection process challenging in remote sensing. Considering to computational complexity, raw satellite images were cropped to  $80 \times 80$  for better learning performance and more efficient use of GPU capacity in the deep learning process. Thus, the proposed model is able to perform in embedded real time applications without a need for any upper-level GPU hardware.

## 4. EXPERIMENTAL RESULTS

### 4.1. Proposed CNN Model and Architecture

The general architecture of the proposed 2D CNN is presented in Figure 3. Description of the layers, filter size, input and output size are given in Table 1. Our proposed method consists of 1 input layer, 12 convolution and ReLu layers with batch normalization, 2 max pooling layers, 1 average pooling layer, 2 fully connected layers, 1 softmax, and 1 classification layer. It contains totally of 43 layers. In the designed architecture, the input image size is  $80 \times 80$  and network depth is 3. All kernels used in convolution layers are  $3 \times 3$  and pooling operation works with  $2 \times 2$  matrix size. As shown in Figure 3, after pooling layer operations, image sizes reduce in half.

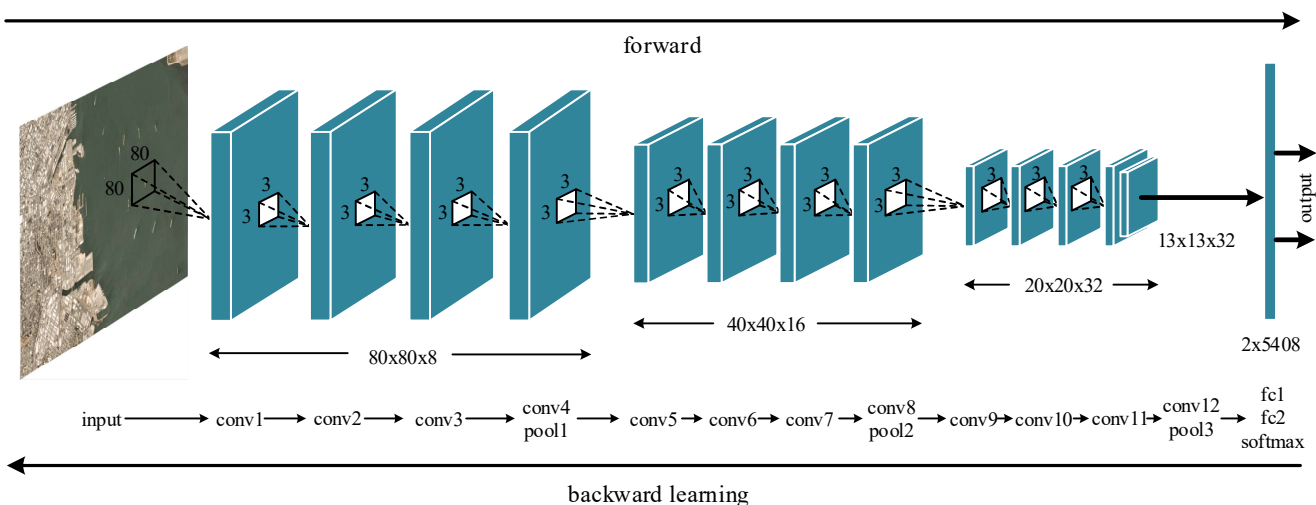


Figure 3. Proposed 2D CNN architecture with detailed layers



Table 1. 2D CNN architecture details for the proposed ship detection model

Layers	Filter Size	Output Size
Image Input	-	80×80×3
Conv1+ReLU	(3,3)×8	80×80×8
Conv2+ReLU	(3,3)×8	80×80×8
Conv3+ReLU	(3,3)×8	80×80×8
Conv4+ReLU+MaxPool	(3,3)×8	40×40×8
Conv5+ReLU	(3,3)×16	40×40×16
Conv6+ReLU	(3,3)×16	40×40×16
Conv7+ReLU	(3,3)×16	40×40×16
Conv8+ReLU+MaxPool	(3,3)×16	20×20×16
Conv9+ReLU	(3,3)×32	20×20×32
Conv10+ReLU	(3,3)×32	20×20×32
Conv11+ReLU	(3,3)×32	20×20×32
Conv12+ReLU+AvgPool	(3,3)×32	13×13×32
FC1+FC2+Softmax	-	1×1×2
Classification Output	-	-

All the parameters of the proposed model are empirically stated as a result of several experiments in order to obtain the best test results. It should be noted that the training process of the designed network includes 10-fold cross-validation method.

## 4.2. Results

In the experiments, code design is performed in MATLAB environment running on a workstation, which has dual Intel Xeon E5 with 2.1 GHz clock, Quadro M4000 8 GB GPU and 32 GB, RAM memory. In the detection stage, we use data split with rates of 87.5% for training and 12.5% for testing. In this way, 3500 training and 500 testing images are selected randomly.

Figure 4 shows the visualization of the layer activations. This activation illustration provides that each feature map includes different forms and features to be learned using the same input image. In addition, one can also say that each layer provides deep filtering to the feature maps.

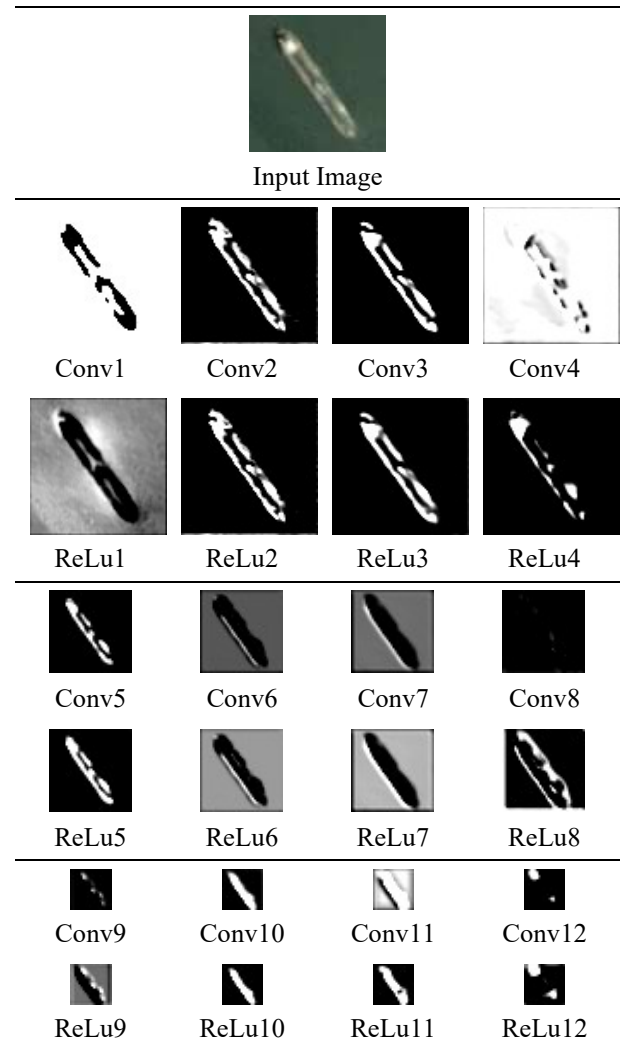


Figure 4. Visualization of the layer activations from convolution 1+ReLU 1 to convolution 12+ReLU 12

In order to achieve a better evaluation, we use the most known metrics over the test data. Considering the un-balance situation, the recall and specificity metrics have an important role besides the accuracy value. F1-score and Area Under the Curve (AUC) is also obtained to support the overall performances. Table 2 lists the test performances of the designed ship detection deep learning model. The trained CNN model is tested using randomly chosen 500 sample images in the test dataset, which includes 370 of NShip class and 130 of Ship class. According to Table 2, the proposed model generates **0.9960** accuracy value, which nearly provides the perfect detection. From the experiments performed to increase the generalization, precision, recall, and specificity values are obtained as highly acceptable.

Table 2. Model performance for test data

Evaluation Metrics	Test Performance
Accuracy	0.9960
Precision	0.9973
Recall	0.9973
Specificity	0.9922
F1-Score	0.9973
AUC	0.9948

F1-score, which is the harmonic average of precision and recall values, shows a good balance between precision and recall with a value of **0.9973**. High F1-score value also supports the accuracy of the proposed model in the unbalanced distribution of the dataset. The AUC value of **0.9948** obtained from the ROC curve is illustrated in Figure 5. This value shows that the designed model performs well in the ship detection process.

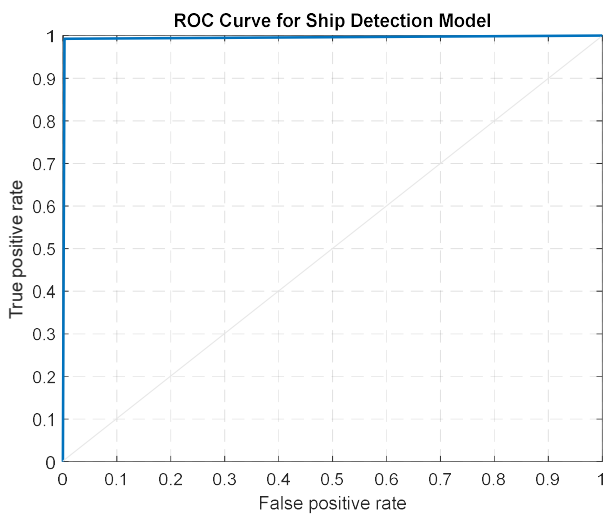


Figure 5. ROC curve demonstration of the test results

For a detailed evaluation, we also give the confusion matrix of the test results. Figure 6 shows the class distribution in terms of predicted and true classes. As we can see in Figure 6, the total number of misclassification is 2; consequently, a single image per class is confused in the test data.

Besides the binary classification results, the CNN model also gives the probabilities of each test image. A probability value indicates which class

a test image belongs to in percent. Figure 7 shows randomly chosen test images together with their probability values.

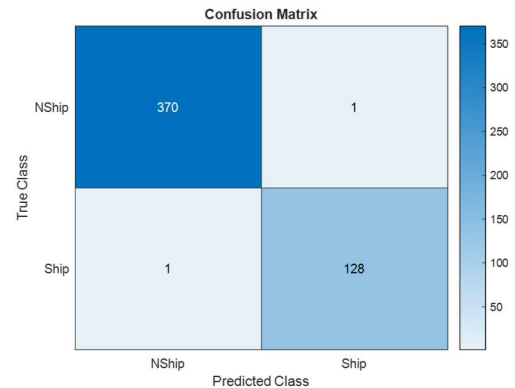


Figure 6. Confusion matrix for the testing results

As seen in Figure 7, the proposed model classifies the partial ship images as NShip according to the dataset definition.



Figure 7. Randomly chosen test results with probabilities

All those results show that the designed deep CNN architecture overcomes the ship detection for satellite images despite the irregular structure of the ship image dataset. Also, the testing results gives the superior performance in ship detection for remote sensing field.

## 5. CONCLUSION

In this paper, we proposed a robust and reliable method for a ship detector in satellite images using a deep CNN. Building a deep network

consisting of convolution and fully connected layers makes the ship detection task from ship satellite images effective and applicable. Thus, the detection of the ships can be realized easily with a minimum computational effort. Without needing any handcrafted feature extraction process is also an advantage of the designed model. The proposed model shows significant performance in both accuracy and other evaluation metrics, which proof the generalizability and robustness.

In subsequent further work, a larger dataset will be used to improve the model learning process for ship classification and detection in satellite images with a wide range of remote sensing. Also, a combination of different layer organizations and types of CNN models will be planned to use in the classification and detection process.

## REFERENCES

- [1] Y. Liu, H.-Y. Cui, Z. Kuang, and G.-Q. Li, "Ship Detection and Classification on Optical Remote Sensing Images Using Deep Learning," *ITM Web Conf.*, vol. 12, p. 05012, 2017.
- [2] Z. Zou and Z. Shi, "Ship Detection in Spaceborne Optical Image with SVD Networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 5832–5845, 2016.
- [3] W. Liu, L. Ma, and H. Chen, "Arbitrary-Oriented Ship Detection Framework in Optical Remote-Sensing Images," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 6, pp. 937–941, 2018.
- [4] E. Ross, B. Arifin, and Y. Brodsky, "An Information System for Ship Detection and Identification," *Int. Geosci. Remote Sens. Symp.*, pp. 2081–2084, 2011.
- [5] U. Kanjir, H. Greidanus, and K. Oštir, "Vessel Detection and Classification From Spaceborne Optical Images: A Literature Survey," *Remote Sens. Environ.*, vol. 207, no. November 2016, pp. 1–26, 2018.
- [6] 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, 2015.
- [7] H. Greidanus, M. Alvarez, C. Santamaria, F. X. Thoorens, N. Kourti, and P. Argentieri, "The SUMO Ship Detector Algorithm for Satellite Radar Images," *Remote Sens.*, vol. 9, no. 3, 2017.
- [8] F. Xu, J. Liu, M. Sun, D. Zeng, and X. Wang, "A Hierarchical Maritime Target Detection Method for Optical Remote Sensing Imagery," *Remote Sens.*, vol. 9, no. 3, pp. 1–23, 2017.
- [9] Y. Yin, N. Liu, C. Li, W. Wan, and T. Fang, "Coarse-to-Fine Ship Detection using Visual Saliency Fusion and Feature Encoding for Optical Satellite Images," *Int. Conf. Audio, Lang. Image Process.*, pp. 705–710, 2017.
- [10] F. Yang, Q. Xu, and B. Li, "Ship Detection From Optical Satellite Images Based on Saliency Segmentation and Structure-LBP Feature," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 5, pp. 602–606, 2017.
- [11] Zhenwei Shi, Xinran Yu, Zhiguo Jiang, and Bo Li, "Ship Detection in High-Resolution Optical Imagery Based on Anomaly Detector and Local Shape Feature," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4511–4523, 2013.
- [12] J. Tang, C. Deng, G. Bin Huang, and B. Zhao, "Compressed-Domain Ship Detection on Spaceborne Optical Image using Deep Neural Network and Extreme Learning Machine," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1174–1185, 2015.
- [13] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, "Deep Learning for Visual Understanding: A Review," *Neurocomp.*, vol. 187, pp. 27–48, 2016.



- [14] H. Lin, Z. Shi, and Z. Zou, "Fully Convolutional Network with Task Partitioning for Inshore Ship Detection in Optical Remote Sensing Images," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1665–1669, 2017.
- [15] R. Pradhan, R. S. Aygun, M. Maskey, R. Ramachandran, and D. J. Cecil, "Tropical Cyclone Intensity Estimation Using a Deep Convolutional Neural Network," *IEEE Trans. Image Process.*, vol. 27, no. 2, pp. 692–702, 2018.
- [16] M. H. Beale, M. T. Hagan, and H. B. Demuth, "Deep Learning Toolbox, User's Guide." Mathworks, 2018.
- [17] D. Moraite, "Satellite Images." [Online]. Available: <https://medium.com/dataseries/detecting-ships-in-satellite-imagery-7f0ca04e7964>. [Accessed: 01-Jul-2019].