Maritime Ship Detection using Convolutional Neural Networks from Satellite Images

Jaafar Alghazo
Computer Engineering Department,
Prince Mohammad bin Fahd University,
Al-Khobar, Saudi Arabia
jghazo@pmu.edu.sa

Abul Bashar
Department of Computer Engineering,
Prince Mohammad bin Fahd University,
Al-Khobar, Saudi Arabia
abashar@pmu.edu.sa

Ghazanfar Latif
Department of Computer Science,
Prince Mohammad bin Fahd University,
Al-Khobar, Saudi Arabia.
glatif@pmu.edu.sa

Mohammed Zikria
Department of Computer Science,
Prince Mohammad bin Fahd University,
Al-Khobar, Saudi Arabia
mzikria@pmu.edu.sa

Abstract— The significance of efficient monitoring and control of the marine traffic for the purpose of safety and security of the ships cannot be overemphasized in the current scenario where global trade and commerce is at its pinnacle. Various stakeholders are concerned with serious maritime issues related to hijacking of ships, illegal fishing, encroachments of sea borders, and illicit exchange of sea cargo, accidents, and military attacks. This requires an automated, accurate, fast, and robust sea monitoring system which can avoid or mitigate the negative effects of such issues. This paper proposes, implements, and evaluates a CNN based deep learning model which can accurately identify ships from the images captured from satellite images. Two models CNN Model 1 and CNN Model 2 having different architectures are trained, validated, and tested on the Airbus satellite images dataset. Both classification accuracy and loss functions are measured by varying the number of the epochs. Also, the complexity comparison of the two models is performed by measuring the training time. The paper concludes that the proposed models are automatic, fast and accurate in terms of their performance on the Airbus dataset by achieving a maximum accuracy of 89.7%.

***Keywords: Maritime Ship Detection, Sea Surveillance, Convolutional Neural Networks, Deep Learning, Deep CNN*.**

I. Introduction

Marine monitoring research has increased over lately due to the increased demand for maritime security and safety applications. Various organizations and government bodies need to ensure the safe navigation and the security of all marine activities through border control, fighting pirates, and monitoring ocean pollution. The ocean and seas cover around 71% of the earth's surface, and throughout the ages, water transportation has been important for world economies. However, there are many dangers lurking in the sea, from piracy to the possibility of accidents. In addition, many illegal activities can occur in these waters, from unlawful fishing to the shipment of illicit cargo. The monitoring of these vast oceans can be done through remote sensing satellites. However, the images captured by the satellites require complex image processing algorithms for the detection and subsequent classification of ships and marine transportation vehicles navigating in the seas. In addition to monitoring of the seas for the purpose of protection of the ships, safeguarding against unlawful fishing, and illicit transport of cargo, there is also the military and defense perspective in which countries need to monitor their shores and close by water bodies for possible attack threats from enemy ships, aircraft carriers and submarines.

Airbus, formerly known as the European Aeronautic Defense and Space Company, provides a comprehensive sea surveillance service through a meaningful solution for wide coverage and intensive monitoring of sea regions with high granularity [1]. Airbus provided data to the researchers to design an automated predictive model that can monitor ship traffic, accurately detect all the ships in each satellite image and thereby prevent illegal activities in the sea [2]. The increase in the shipping traffic means that there is an increased probability that irregularities in the sea can occur, such as sea piracy and illicit cargo movement. Airbus desires to build a robust sea monitoring system to oversee ship traffic for the purpose of controlling illegal activities such as piracy, unlawful fishing, and the movement of illicit goods via cargo ships. In addition, such a system would also protect the marine environment by preventing accidents of ships which can cause environmental pollution (e.g., oil spillage or chemical pollutants).

However, based on the current research literature, an optimal solution for ship detection and classification has yet to be developed. The processing speed and classification accuracy of the existing Machine Learning (ML) algorithms still need improvement. The prime objective is to use Artificial Neural Networks (ANNs) for developing predictive algorithms that can achieve higher accuracy in the detection of ships from the images as compared to the previously proposed algorithms and also to speed up the training and decision making process. In order to achieve this objective, this paper contributes by building optimized Convolutional Neural Networks (CNNs) based architectural model for the purpose of ship detection from satellite images.

The rest of the paper is organized as follows. Section 2 details the literature review related to ship detection from satellite images. Section 3 describes the dataset used for the research experiments. The proposed methodology and model design are presented in Section 4. Experimental results and their discussion are detailed in Section 5. Finally, Section 6 concludes the paper by summarizing the research contributions and pointers to future work.

II. Literature Review

Ship detection is an important application in the field of image processing due to the advantages that the automatic monitoring of the earth's water regions could bring in for the concerned stakeholders. From preventing illegal fishing, illicit export of cargo, sea piracy to advanced military applications such as protecting sea borders from encroachments and naval attacks, the advantages of having an autonomous system that can detect and identify ships is very crucial. Therefore, with the latest advances in technology and recent increased interest in Neural Network research for detection, classification, and segmentation, the applications for the detection of objects of interest in images have also increased. The detection of ships is one of these applications that can bring out great advantages for the concerned stakeholders.

In reference to the Airbus competition previously mentioned, the demand for global trade caused a very rapid growth in ship traffic in the open seas. Thus, the demand for monitoring systems has also increased which can detect ships automatically with improved accuracy and higher speeds. Even though various algorithms have been developed for ship detection, yet an optimal solution with high identification accuracy and speedier decision making is still open for research. It is the aim of this paper to develop an efficient automatic ship detection algorithm with makes identification process faster and more accurate. This is a very challenging task both from the perspective of accuracy in detection and reducing the detection time taken. Detection alone is cumbersome task as various objects in the sea may look like a ship when actually it is not the case. The dataset developed by Airbus consists of various images, including open waters, wharfs, buildings, and clouds [2]. As a matter of fact, most of the images do not contain any ships and many images are obscured with clouds and fog, which increases the difficulty during the detection process. In [3], the authors present a novel technique based on the Generalized Likelihood Ratio Test (GLRT) and have compared it with the traditional Constant False Alarm Rate (CFAR) method using Monte-Carlo simulations. Their approach showed improved performance results as compared to the traditional method. In [4], the authors showcase the importance of maritime safety and security for various countries, and specifically focus on the UK national strategy for maritime security. The document identifies, assesses, and addresses the maritime security challenges in the UK.

There are various methods and techniques to detect ships from satellite images. One technique utilizes the concept of prescreening, which is a filtering strategy, and this is considered as the key ingredient of this algorithm [5]. Ship detection is achieved through reduction of computation time in this technique. This method on the average takes from five to ten minutes for the detection process. There is also ship detection architecture based on an object-oriented methodology that uses different tools to aid in the monitoring task [6]. This system can generate customized reports in HTML format. The ship-detection system has been used for panchromatic VHR satellite images. It can work on Ikonos and QuickBird satellite imagery. The performance of this detection system on the Ikonos scene takes around 30 min to 2 hours, and it takes more than this time for other modules. In fact, the processing time depends on the objects on each scene, so it is still not optimized in terms of its performance. A Novel algorithm based on the wavelet transform that uses satellite-based Synthetic Aperture Radar (SAR) imagery is proposed in [7]. This novel algorithmic approach uses some tools provided by the wavelet theory. This algorithm has been tested over SAR images and over simulated images. The result achieved after following the steps of the algorithm on the satellite images, was that the target was clearly noticeable, and it enlarged the image contrast for better detection [8]. The process of this approach consists of some steps such as background estimation and removal of false detection on RADARSAT images with different modes such as W2 and SNA. Benchmark considered here is based on multi-ship detection systems. Each system has a different performance, some have better performance on certain images, and the others have better results depending on the algorithm and the sense made from an image. For simple images, all systems give 85-95% accuracy, and for complex images, the rate is between 75-99% but sometimes rate becomes as low as 60%. In addition, there is a methodology that has been used that improves ship detection with airborne polarimetric Synthetic Aperture Radar (SAR) data [9]. The key attribute of this method is to compare the data obtained from polarimetric SAR and single-channel SAR data. By applying a statistical decision theory, the detection performance improves by calculating the average performance of different types of radar detection systems. The results show that the polarimetric system gives better performance compared to other systems. In conclusion it can be said that the performance of ship detection can be relatively improved by using polarimetric information.

A system was proposed based on rotation anchor strategy to foretell the minimum circumscribed rectangle of the object and to reduce the redundant detection region [10]. The dataset of the experiment depends on the remote sensing images from Google earth for detecting ships. Results indicated that the R-DFPN has better performance in ship detection in complex images, specifically in detecting densely arranged ships. In another method for ship detection using satellite images [11], the authors first experimented with K-Nearest Neighbors, Naive Bayes, Random Forest, and Support Vector Machine models. Deep Learning methods were also tested where researchers used pre-trained network architectures as the baseline network for comparison purposes. A comparison between the Deep Learning method and traditional methods showed that the Random Forest model gave the best performance in the traditional methods category, with a 93% detection accuracy rate, while the Deep Learning methods achieved 94% detection accuracy rate. In [13], the authors present an Automatic Identification System (AIS) based on Synthetic Aperture Radar (SAR) images. The AIS ship is responsible for informing the coastal receiver about the position, speed, and identification information and allows tracking of ships up to a distance of 40 kms. The novel Generalized Likelihood Ratio Test (GLRT) algorithm was tested on a dataset acquired at (X-band) and (C-band) in the same area at the same time and achieved higher results compared to other algorithms. Another approach for detecting ships at sea with high frequency (HF) ultrasound Over the Horizon Radar (OTHR) is presented in [14]. This technique can detect the ships at a range of 2000 km or more by using methods from the Radar and Doppler frequency domains.

The optical sensors can detect smaller objects where the Radar cannot, while the radar is able to see what the optical sensors cannot see. By mixing optical satellite data and radar data, further opportunities exist in detecting and monitoring objects. For example, oil spill detection in the case of very low wind appears as dark shapes, and because of the surface roughness of the sea it has a higher probability of being detected from the SAR images as oil spots. Different weather conditions of the oil spill are presented based on TerraSAR-X radar data [15]. Ship detection techniques mainly implement the detection in two stages, namely, candidate selection and classification. In [16], the authors present a technique that implements the candidate selection in various conditions of backgrounds in VNREDSar-1, which is a dataset of panchromatic satellite images used to measure the complex background of the sea surface. The classification stage is implemented using Support Vector Machines, Neural Networks and CART decision tree. In [17], the author presents the methodologies using both radar and optical images of satellite data to improve maritime safety. The author presents sophisticated space technology and its advantages in maritime usage. In [18], the author details in his dissertation, ship detection techniques and proposes various improved methods for ship detection.

III. Expiremental Dataset

Airbus provided the dataset on the Airbus ship detection challenge website [2]. The images were taken by the Airbus satellite. A particular captured image could have one ship, multiple ships, or no ship at all. It was found that in the given dataset, approximately ¼ number of the total images had a presence of at least one ship. Some images have a small portion of a ship, for example, in the corner of the image only the bow of the ship is captured. In addition, ships have a variety of shapes and sizes, such as Aircraft Carrier, Barque, Yacht and there are several other types. The images were taken from different locations, so some of the ships were in the open sea, marinas, or at the dock. The dataset contains 192,556 images with the pixel size of (758px × 758px) and total storage requirement of more than 30GB. To reduce the computational time, 20,000 images are selected randomly for the binary classification problem with equal distribution of images for each class (*ship* or *no ship*). Fig. 1 shows a set of sample satellite images for the purpose of sea surveillance.

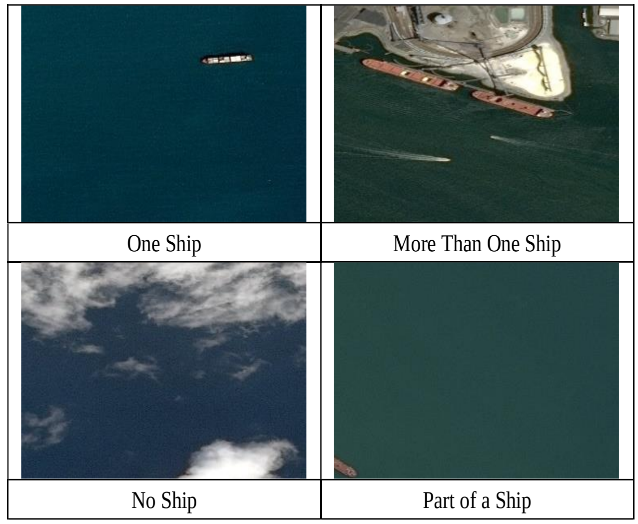

Figure 1. Sample satellite images for sea surveillance

IV. Proposed Method

The Convolutional Neural Network (CNN) is the cutting-edge approach for image recognition and processing which is applied on wide variety of images in various domains [19-20]. CNN is basically a multilayer perceptron having a regularized form. Multilayer recognitions are more often than not complex systems that are completely associated, i.e., each neuron in one layer is associated with all neurons within the following layer. These systems are also termed as "full network" which makes them inclined to information over fitting. However, CNN takes a distinctive approach to regularization by utilizing fewer and less complex calculations to achieve the desired results. Hence, CNN is subsequently rated at the lower end on the scale of network and complexity. CNN models can be made to work in a parallel framework, just like the native neural systems, where the adjoining neural units are permitted to communicate with each other. CNN has broad applications in scientific domain ranging from face recognition, action recognition to text classification and Natural Language Processing [21-22]. In comparison to hand-engineering low-level highlights, CNN's ubiquity is due to its predominant multi-scale high-level data representations. In any case, assessing millions of profound CNN parameters requires an expansive number of commented on tests, which as of now, avoids the application of numerous lower profound CNNs (such as Alex Net, VGG, ResNet) to restricted preparing information issues. One of CNN's employments is tally objects in still pictures, and video could be a well-defined cognitive assignment in which individuals outflank machines to an extraordinary degree.

A deep convolutional neural network is commonly used in several image classification major problems as an important identification algorithm [20]. With the fast development of large-performance computing devices as well as parallel computing devices, the neural network is also increasingly attracting more exposure from several researchers in this field.

CNN is a regularized type of multi-level perceptron and typical shapes of regularization. By using smaller and less complicated equations, CNN takes a unique approach to regularization, taking advantage of the different leveled information techniques, and building more complex designs. In recent years, in several computer vision functions, the convolutional neural network (CNN) has gained great success. Directly inspired by neuroscience CNN explores the neural visual system with several other properties. A major difference is that CNN is normally a feed-forward architecture, whereas repeated links are common in the visual system.

Table 1 and Table 2 summarize the proposed CNN models. Model 1 consists of total 13 layers including 3 convolutional layers while Model 2 consists of total 16 layers with 4 convolutional layers. At the end of both models, average pooling layer, flatten layer are applied before getting the output class using the dense layer.

TABLE 1. PROPOSED CNN MODEL 1

#	Input	Output	Parameters
1	Input Image	100, 100, 3	
2	Conv2d	100, 100, 16	448
3	Batch Normalization	100, 100, 16	64
4	Activation	100, 100, 16	0
5	Conv2d	50, 50, 32	4,640
6	Batch Normalization	50, 50, 32	128
7	Activation	50, 50, 32	0
8	Conv2d	25, 25, 64	18,496
9	Batch Normalization	25, 25, 64	256
10	Activation	25, 25, 64	0
11	Average Pooling	3, 3, 64	0
12	Flatten	576	0
13	Dense (Output)	2	1154

TABLE 2. PROPOSED CNN MODEL 2

#	Input	Output	Parameters
1	Input Image	100, 100, 3	
2	Conv2d	100, 100, 16	448
3	Batch Normalization	100, 100, 16	64
4	Activation	100, 100, 16	0
5	Conv2d	100, 100, 64	1088
6	Batch Normalization	100, 100, 64	256
7	Activation	100, 100, 64	0
8	Conv2d	50, 50, 128	8320
9	Batch Normalization	50, 50, 128	512
10	Activation	50, 50, 128	0
11	Conv2d	25, 25, 256	33024
12	Batch Normalization	25, 25, 256	1024
13	Activation	25, 25, 256	0
14	Average Pooling	3, 3, 256	0
15	Flatten	2304	0
16	Dense (Output)	2	4610

V. EXPERIMENTAL RESULTS

In this section we now present the results from the experiments which were performed on the proposed system. The proposed models which were described in Section IV, were trained with the Airbus images dataset. The overall dataset was partitioned into three smaller sets, namely, training (60%), validation (20%) and testing (20%) sets. Training data had both types of images, namely, with a ship captured in them and no ship captured in them. This consideration was essential to make the model learn both type of classes (*ship* and *no ship*) to work well when it would be utilized for decision making in the testing phase.

Before the testing phase the model was validated with another unseen dataset consisting of 20% of the total images. Validation step was incorporated to check for the model going to the over fitting or under fitting regions, which would adversely affect the model performance during the testing phase. Another purpose of the validation phase was to choose among the different CNN models to optimize the choice of

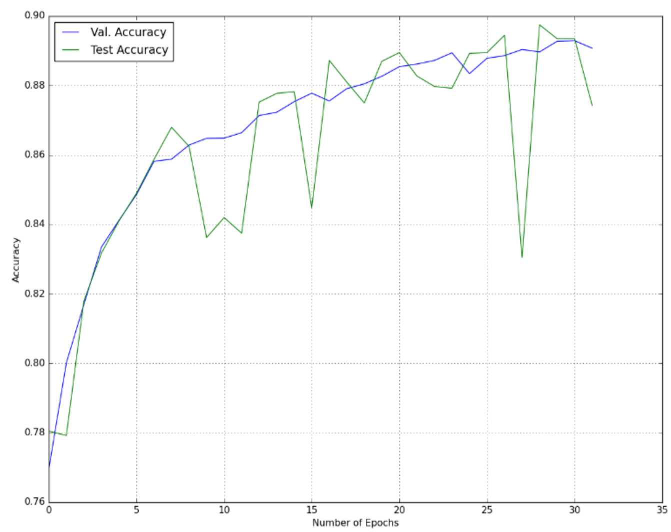

Figure 2. Accuracy curve of the proposed CNN Model 1 for upto 32 epochs

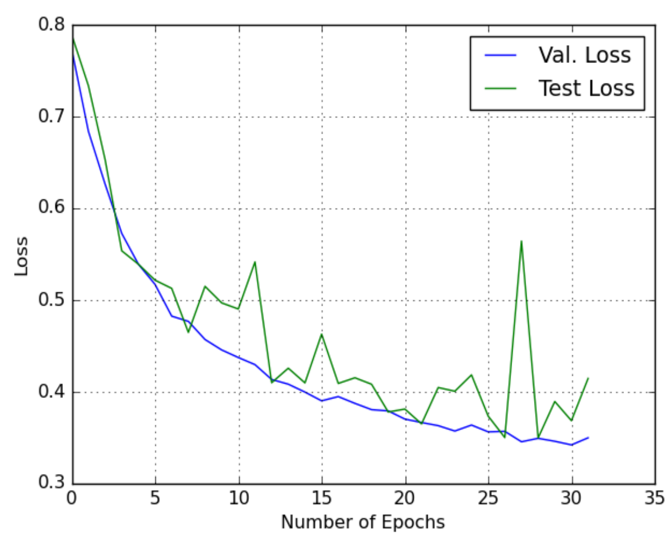

Figure 3. Loss curve of the proposed CNN Model 1 for upto 32 epochs

the various parameters which are best suited for the said problem of image classification. The testing dataset was then finally utilized for obtaining the classification accuracy. Apart from the classification accuracy measurement, the loss function was also measured. Actually, loss measurement of the CNN model is the complement of the accuracy measurement, implying that the higher the accuracy the lower is the loss value. The experiments were conducted with the aim of studying the effect of the number of epochs on the accuracy and loss values of the CNN models. Also, the experiments were conducted on two models (Model 1 and Model 2) as described earlier in Section IV.

Fig. 2 shows the plot of validation accuracy and test accuracy for CNN Model 1 for the epochs ranging from 0 to 35. As observed, both the accuracies increase by an increase in the number of epochs. It is to be noted that the validation accuracy has a more of smoother progression as compared to the test accuracy which has more of an irregular increasing

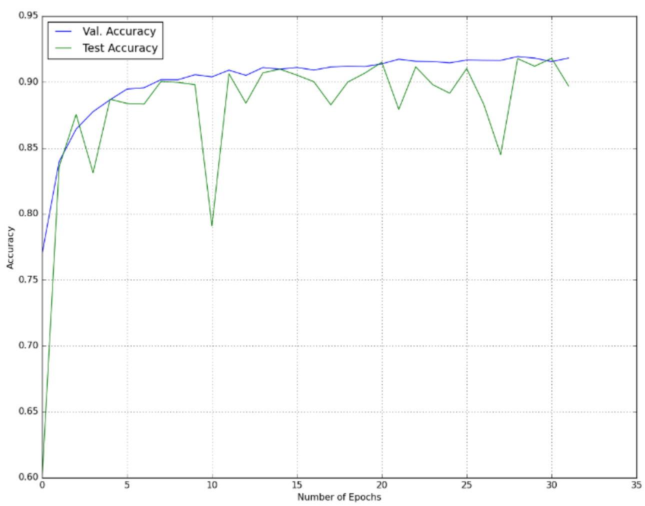

Figure 4. Accuracy curve of the proposed CNN Model 2 for upto 32 epochs

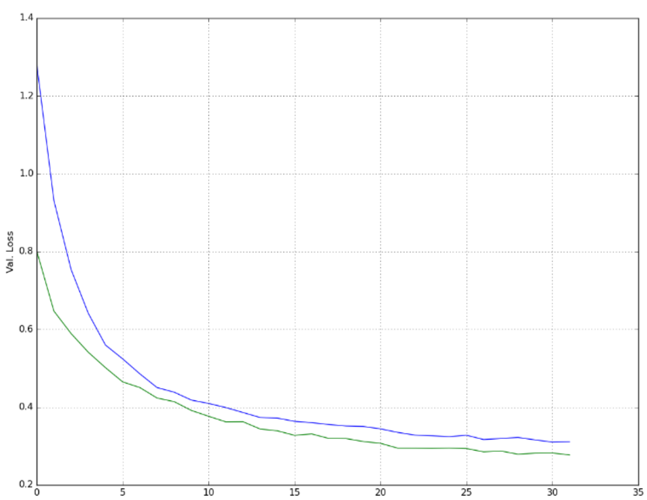

Figure 5. Loss curve of the proposed CNN Model 2 for upto 32 epochs

pattern. However, on an average we see that both the accuracies saturate after 25 epochs, whereby further increase in the number of epochs does not increase the accuracies further. Thus, we get a maximum validation accuracy of 89.35% and testing accuracy of 87.42% respectively. Similar discussing is also applied to Fig. 4 for CNN Model 2 where we achieve a maximum validation accuracy of 91.79% and the testing accuracy of 89.70% respectively. As it is clear that the CNN Model 2 performs better in terms of both accuracies as compared to CNN Model 1. This is because CNN Model 2 has more hidden layers in it which results in a better classification accuracy. However, the price which we must pay for this improved performance is the time and complexity involved during the training phase. As can be seen from Table I, the computational time for CNN Model 2 is 2131 secs, whereas it is less for CNN Model 1 (which is only 1882 secs).

Now we look at the results for the loss functions for both the models by looking at the validation and testing phases. From Fig. 3 and Fig. 4, we observe that the loss functions have

TABLE III. COMPARISON OF THE PROPOSED CNN MODEL 1 AND MODEL 2 ACCURACIES, LOSS AND COMPUTATIONAL TIME.

	Evaluation Technique	**Proposed Model 1**	**Proposed Model 2**
Training	Accuracy	89.29%	91.82%
	Loss	0.342	0.277
	Computational Time (Sec.)	1882.16	2131.68
Validation	Accuracy	**89.35%**	91.79%
	Loss	0.368	0.291
Testing	Accuracy	87.42%	89.70%
	Loss	0.415	0.317

decreasing trends when the number of epochs are increased as before for the accuracy experiments. This is expected as we know that the loss function has a complementary behavior to accuracy. It was found that for CNN Model 1 the validation loss was 0.368 and the testing loss was 0.415, respectively. Whereas in the CNN Model 2, these losses were having relatively lesser values of 0.291 for validation loss and 0.317 for the testing loss. These values again confirm the superior performance of CNN Model 2 as compared to the CNN Model 1.

VI. CONCLUSION

It is evident through the study that the method of Deep Learning Architecture in Convolutional Neural Networks (CNN) is effective, and more hidden layers construct a better rate of recognition. The results depict that having 2 hidden layers enhanced the rate of recognition obtained. However, including more hidden layers, the additional more the network and enhanced time of computation is required to generate better outcomes. For the aim of better outcomes, a rate of recognition of 89.35% would be sufficient, and just including more hidden layers and enhancing the duration to generate outcomes would not be justified.

REFERENCES

[1] Petrescu, R. V., Aversa, R., Akash, B., Corchado, J., Apicella, A., & Petrescu, F. I. (2017). Home at airbus. *Journal of Aircraft and Spacecraft Technology*, *1*(2).

[2] Kaggle. Dataset For Airbus Ship Detection Challenge. https://www.kaggle.com/c/airbus-ship-detection/data, 2018.

[3] Iervolino, P., & Guida, R. (2017). A novel ship detector based on the generalized-likelihood ratio test for SAR imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(8), 3616-3630.

[4] Britain, G. (2014). *The UK national strategy for maritime security*. HM Government.

[5] Corbane, C., Pecoul, E., Demagistri, L., & Petit, M. (2008, December). Fully automated procedure for ship detection using optical satellite imagery. In *Remote Sensing of Inland, Coastal, and Oceanic Waters* (Vol. 7150, p. 71500R)

[6] Willhauck, G., Caliz, J. J., Hoffmann, C., Lingenfelder, I., & Heynen, M. (2005). Object-oriented ship detection from VHR satellite images. *Semana Geomática*.

[7] Tello, M., López-Martínez, C., & Mallorqui, J. J. (2005). A novel algorithm for ship detection in SAR imagery based on the wavelet transform. *IEEE Geoscience and remote sensing letters*, *2*(2), 201-205.

[8] Greidanus, H., Clayton, P., Indregard, M., Staples, G., Suzuki, N., Vachoir, P., ... & Ringrose, R. (2004, September). Benchmarking operational SAR ship detection. In *Proceedings of 2004 IEEE International Geoscience and Remote Sensing Symposium.* pp. 4215-4218. IEEE.

[9] Liu, C., Vachon, P. W., & Geling, G. W. (2005). Improved ship detection with airborne polarimetric SAR data. *Canadian Journal of Remote Sensing*, *31*(1), 122-131.

[10] Yang, X., Sun, H., Fu, K., Yang, J., Sun, X., Yan, M., & Guo, Z. (2018). Automatic ship detection in remote sensing images from google earth of complex scenes based on multi-scale rotation dense feature pyramid networks. *Remote Sensing*, *10*(1), 132.

[11] Chen, Y., Zheng, J., & Zhou, Z. Airbus Ship Detection-Traditional vs Convolutional Neural Network Approach, *Stanford.*

[12] Touzi, R., Charbonneau, F. J., Hawkins, R. K., & Vachon, P. W. (2004). Ship detection and characterization using polarimetric SAR. *Canadian Journal of Remote Sensing*, *30*(3), 552-559.

[13] Iervolino, P., Guida, R., Lumsdon, P., Janoth, J., Clift, M., Minchella, A., & Bianco, P. (2017, July). Ship detection in SAR imagery: A comparison study. In *2017 in proceeding of IEEE International Geoscience and Remote Sensing Symposium*, pp. 2050-2053. IEEE.

[14] Barnum, J. (1986). Ship detection with high-resolution HF skywave radar. *IEEE Journal of Oceanic Engineering*, *11*(2), 196-209.

[15] Singha, S., Ressel, R., Velotto, D., & Lehner, S. (2016). A combination of traditional and polarimetric features for oil spill detection using TerraSAR-X. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(11), 4979-4990.

[16] Lưu Việt, H. (2017). Operational Detection and Management Of Ships in Vietnam Coastal Region using Vnredsat-1 Image (Doctoral dissertation).

[17] Weintrit, B. (2017). Earth Observation Opportunities to Enhance Maritime Safety. *TransNav, the International Journal on Marine Navigation and Safety of Sea.*

[18] Iervolino, P. (2016). *Ship detection with SAR: modelling, designing and real data validation* (Doctoral dissertation, University of Surrey (United Kingdom).

[19] Latif, G., Iskandar, D. N. F. A., Alghazo, J., & Butt, M. M. (2020). Brain MR Image Classification for Glioma Tumor detection using Deep Convolutional Neural Network Features. *Current medical imaging*.

[20] Latif, G., Alghazo, J., Maheswar, R., Vijayakumar, V., & Butt, M. Deep learning based intelligence cognitive vision drone for automatic plant diseases identification and spraying. *Journal of Intelligent & Fuzzy Systems*, (Preprint), 1-12.

[21] Al-Faraj, S., Al-Bahrani, M., Al-Ghamdi, S., Rafie, M., Bashar, A., & Latif, G. CNN-Based Alphabet Identification and Sorting Robotic Arm. In *International Conference on Communication, Computing and Electronics Systems: Proceedings of ICCCES 2020* (p. 349). Springer Nature.

[22] Alufaisan, S., Albur, W., Alsedrah, S., & Latif, G. Arabic Braille Numeral Recognition Using Convolutional Neural Networks. In *International Conference on Communication, Computing and Electronics Systems: Proceedings of ICCCES 2020* (p. 87). Springer Nature.