A Real-Time Ship Object Detection in Remote Sensing Image using Faster Region-CNN Model

[1]K. Varalakshmi, [2]Rajapriya S P, [3]R. Arun, [4]S. Gayathri,
[1]Assistant Professor, Department of AI&DS, St. Joseph's Institute of Technology, Chennai, India
[2]Assistant Professor, Sathyabama Institute of Science and Technology, Chennai, Tamilnadu, India
[3]Assistant Professor, Department of Computer and Communication Engineering, Sri Eshwar College of Engineering, Coimbatore- 641202.
[4]Assistant Professor, Sathyabama Institute of Science and Technology, Chennai Tamilnadu, India
[1]varaluckky.2@gmail.com, [2]rajapriya0929@gmail.com, [3]arunr.cbe@gmail.com,
[4]gayukarthick25@gmail.com

Abstract: With the rapid growth of global maritime traffic, effective ship detection from satellite and aerial images has become essential for applications in maritime security, disaster response, and environmental monitoring. Traditional ship detection methods struggle with the challenges posed by the vastness, complexity, and dynamic nature of oceanic environments. To address these challenges, we propose a highly efficient model that combines the power of CNN with the RPN for fast and accurate ship detection. This study introduces a novel approach for real-time ship object detection in RSI images using the Faster R-CNN.

The proposed model is optimized for real-time processing, ensuring the rapid detection of ships even in large-scale satellite imagery. By utilizing a comprehensive dataset of high-resolution RSI images, the system is trained to recognize ships under varying conditions, including different angles, lighting, and sea states. The Faster R-CNN model demonstrates superior performance, offering both high detection accuracy and fast inference times, making it suitable for deployment in real-world maritime surveillance systems. The proposed approach outperformed more conventional R-CNN models in terms of accuracy (99.25 percent for ResNet-50) in experiments conducted on the ship object benchmark dataset.

Keywords: ***Object detection, ship detection, remote sensing image, region based CNN, region proposal network, faster region, aerial image and high resolution image.***

I. INTRODUCTION

The detection and categorization of ships in RSI images is of critical significance for maritime security as well as other applications, such as the monitoring of marine pollution, the surveillance of traffic, and the prevention against illicit fishing [1]. The automated detection and categorization of ships from RSI images is becoming an increasingly important application in both the military and civilian sectors as the volume of satellite image data continues to balloon. In spite of this, the detecting systems are confronted with the demand of virtually real-time reaction capacity as well as the need to handle vast volumes of data that are being received [2]. These standard methods are often only useful for ordinary imagine analysis, and they are not suitable for the task of ship recognition and classification in remote sensing images, which frequently are comprised of enormous amounts of data and a great deal of background noise. However, several significant research have been conducted in this subject. As far as accuracy, performance, and complexity are concerned, the majority of standard procedures are problematic [3].

The use of deep learning has become more promising in recent years across a wide range of real-world domains. Several areas have reported state-of-the-art performance, including text processing, visual object identification, and audio recognition [4]. It is arguable that the network's capacity to learn is one of the main reasons why pattern recognition applications have been so successful recently. Final classification accuracy may be greatly enhanced by scaling up deep learning in terms of either the amount of training instances or the number of model parameters, or both [5]. Consequently, GPUs are a recent breakthrough that allows for the practical training of deep networks with modest sizes. It has been recognized since the beginning of ship detection and classification that building an accurate detection and classification system only by hand is nearly difficult due to the richness and variety of image data. Because ships are not widely dispersed at sea, remote sensing images only show little portions of ship targets [6].

The following is the sequence in which the reminder of this work is structured. In the second section, an overview of the related work that pertains to ship object detection research in remote sensing image is provided. In the third section, the processing of ship candidates extraction is described; and our proposed CNN model for ship detection is explained. In the forth section, tests and analysis of ship object detection results are demonstrated. Finally, the section 6 is provided conclusion and future enhancement of the paper.

II. LITERATURE SURVEY

There are a number of benefits associated with RSI images, the most notable of which is that they have a relatively low impact on the weather and the passage of time. Additionally, ship recognition in RSI images has been thoroughly researched [7]. A constant false-alarm rate (CFAR) detector with a specific RSI image background distribution, such as Gauss distribution, k distribution, Gamma distribution, or any other combination, is the foundation of the majority of ship recognition algorithms. These algorithms are generally used to identify vessels. Ship detection methods in polar metric RSI images were the subject of a brief review that Han and Chong [8] conducted.

Using RADARSAT photos captured in a variety of modes, Greidanus et al. [9] conducted a benchmark test to evaluate the performance of eight different ship detection systems that were based on space-borne devices. However, there are certain restrictions associated with ship detection based on RSI. To begin, the revisit cycle is quite lengthy due to the restricted number of SAR satellites, and it is not capable of meeting the requirements of employing real-time ship surveillance. A second issue is that the resolution of the majority of satellite SAR images is sometimes insufficient to extract specific information about ships [10].

There was a significant amount of research conducted on classical approaches for the identification and classification of ships using optical imagery. While Chen and Wang [12] utilized Dynamic Bayesian Network to classify various types of ships, Zhu and Antelo et al. [11] extracted manually designed features from images such as shapes, textures, and physical properties. However, they were unable to overcome the variability of the images and the problems associated with large volumes. An autoencoder-based deep neural network that was merged with an extreme learning machine was recently developed [13], and it surpassed certain other approaches in terms of detection accuracy. This breakthrough came about as a result of the introduction of deep learning architectures.

While keeping the same level of accuracy performance, Huang et al. [14] have introduced their random forest method, which allows for a quicker training pace in comparison to the conventional approach. SR-Hough is a unique approach that was introduced by Yokoya and Iwasaki [15]. This method is dependent on sparse representation and Hough voting. Within the remote sensing pictures, this technique is primarily concerned with locating occurrences of a certain object or a class of objects. A study that was given by Prasad and colleagues in [16] investigates a variety of difficulties that are encountered in maritime surveillance. These difficulties include opacity, differences in orientation and size, a plurality of object classifications, and changes in weather. In addition, Yu et al. [17] presented a machine learning strategy that makes use of a context-driven Bayesian saliency model in order to identify items that are tiny and dim in a Forward Looking Infrared (FLIR) image.

III. PROPOSED SHIP OBJECT DETECTION MODEL

An advanced region-based CNN series, more particularly the Faster Region-CNN algorithm, serves as the foundation for the ship target recognition model that is presented in this study. Figure 1 illustrates how successful the Faster R-CNN algorithm is at recognizing small-scale ship targets, despite the fact that it offers a sufficient detection speed and accuracy for the majority of objects. It is possible for the Faster R-CNN-based model to detect tiny objects like ships in an efficient manner. In order to construct the model, the following stages are included:

- Dataset Collection
- Backbone network (Pre-trained Model)
- Region Proposal Network
- Align Region of Interest
- Class Label and bounding box Generation

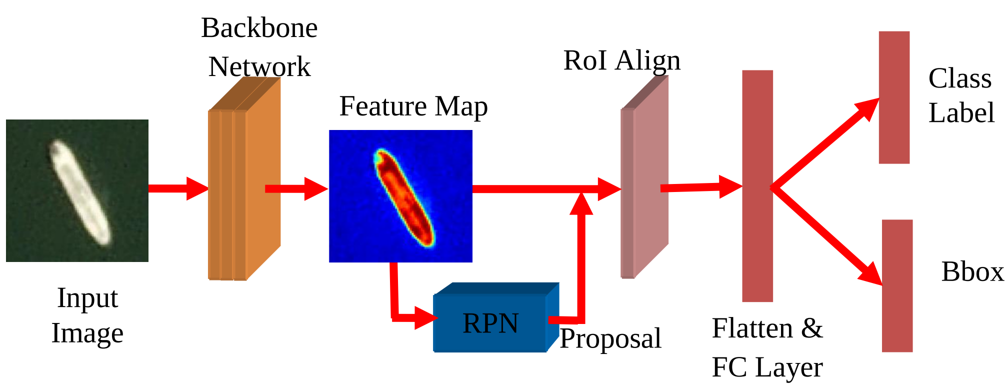

Fig. 1 Architecture of proposed ship object detection model

3.1. Collecting the Dataset

Within the scope of this study, the experiment is carried out on the ship images that were obtained from a satellite images dataset that was compiled by the Kaggle platform [18]. The surface of the world is depicted in these images, which also include highways, farming, houses, and other items. These satellite images of San Pedro Bay and San Francisco Bay, both of which are situated in the districts of California, were taken by PlanetScope. There are two unique classes, "ship" and "non-ship," each of which comprises four thousand RGB images with a pixel resolution of 80 by 80. The total number of images that belong to each of the classes (Ship and Nonship) and sample images is shown in Figure 2 and Table 1. It is necessary for images that feature a ship to depict the ship in its entirety. There is a possibility that ships are positioned in a different direction and have varying sizes. Additionally, it is likely that they are emitting atmospheric noises.

Table 1 Dataset distribution of Ship Object

S. No.	Dataset Type	No. of Images	Training Images	Testing Images
1.	Ship	3000	2700	300
2.	No Ship	3000	2700	300

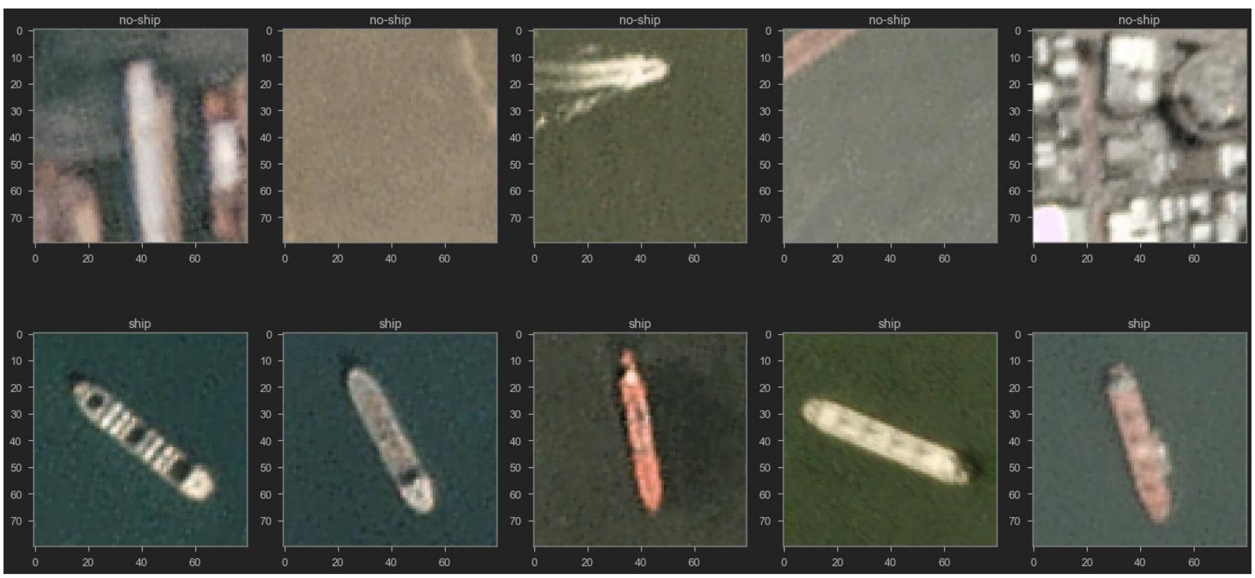

Fig. 2 Sample Ship Object Dataset

3.2. Backbone Network

The architecture incorporates a powerful feature extraction backbone like ResNet-50, fine-tuned specifically for ship detection in RSI images. The backbone extracts multi-scale features, enabling the detection of ships of varying sizes, from small fishing boats to large container ships. The pre-trained weights from large-scale datasets such as ImageNet are adapted to the task, significantly improving convergence rates and detection performance. The features that distinguish ships from the surrounding water, such as sharp edges, distinct shapes, and unique textures, are emphasized through deeper convolutional layers.

3.3. Region Proposal Network

The RPN generates region proposals, which are potential areas of interest where ships might be located. The predefined anchor boxes are configured with diverse scales and aspect ratios to accommodate the varying dimensions of ships in remote sensing images. Each anchor box is assigned an objectness score, reflecting the likelihood of containing a ship. Low-score anchors are discarded to reduce computational overhead. The positions and dimensions of anchor boxes are fine-tuned to better align with the actual ship boundaries, ensuring precise localization. To minimize false positives, particularly in challenging background regions like water, sandbanks, or wave patterns, the RPN is trained to distinguish between ships and non-ship regions using hard negative samples.

3.4. RoI Align

The Region of Interest (ROI) Pooling module transforms the variable-sized region proposals into fixed-size feature maps, enabling consistent processing. Each RoI is classified into one of two categories—"Ship" or "Background"—with confidence scores.

3.5. Bounding Box Generation

A regression layer adjusts the coordinates of the detected regions, enhancing the accuracy of ship localization. If the application requires ship categorization (e.g., cargo, fishing, military), the classification layer is extended to include multiple ship types.

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

The purpose of this part is to conduct an analysis of the performance and efficacy of RSI ship object identification using the Faster R-CNN model. In the beginning, we showed the benchmark datasets for RSI ship object detection. After that, we examined the performance of Faster R-CNN, and lastly, we gave the experimental results for the proposed Faster R-CNN. The Python and Anaconda integrated development environment (IDE) technologies were utilized in the development of the suggested model.

4.1. Performance Metrics

We have utilized four conventional performance measures in order to evaluate the performance of a proposed model. These metrics are the Accuarcy, Average Precision (AP), the Precision-Recall Curve (PRC), and the Intersection of Union (IoU). The precision may be evaluated by dividing the total number of samples in a class by the number of objects that were correctly identified. Equation (1) may be used to demonstrate the precision value of the class c, Pc. In this equation, tc represents the total number of objects in class c that have been correctly identified, and nc represents the total number of samples in class c.

$$\mathrm{P_C} = \frac{\mathrm{t_c}}{\mathrm{n_c}} \tag{1}$$

One method for determining the recall is to divide the total number of relevant samples in the associated class by the number of objects that were correctly identified. Rc is the recall value of the class, and it is represented by equation (2). In this equation, tc represents the total number of object samples that have been correctly recognized in class c, and kc represents the number of samples that have been identified as representative of class c.

$$\mathrm{R_C} = \frac{\mathrm{t_c}}{\mathrm{k_c}} \tag{2}$$

As a result, a larger AP indicates a better performance. The AP is calculated by calculating the average value of the accuracy across the interval from Pc = 0 to Rc = 1, which is also known as the area under the PRC. According to the IoU, which is a typical evaluation metric, the accuracy of object identification and the overlap rate of the predicted bounding box (Bp) and ground truth (Bgt) generated by the model are both evaluated.

$$IoU = \frac{B_p \cap B_{gt}}{B_p \cup B_{gt}} \geq a0 \tag{3}$$

4.2. Result and Discussions

For the purpose of evaluating the proposed approach, a Faster Region-CNN model was trained using the PlanetScope ship target object identification dataset. In order to train and carry out all of the tests, a personal computer equipped with Intel® i7 processors operating at 3.40 GHz and 16 GB of RAM was utilized. We have trained the network for 50 epochs using a learning rate of 0.0001 and Adam optimizer. The backbone of the feature extraction network that we have used is ResNet-50. Figure 3 illustrates the training, validation accuracy, and loss of any ship object identification model that may have been used.

Fig. 3 Training, Validation accuracy and loss of ship object detection

In a similar manner, we have used the three characteristics of AP, AP50, and AP75 to evaluate the object identification capabilities of various models. IoU thresholds ranging from 0.50 to 0.95 are represented by the AP in this context. A threshold of 0.50 for the IoU is represented by the AP50, while a threshold of 0.75 for the IoU is represented by the AP75. On display in Figure 4 is an example depiction of the detection of ship objects through the utilization of area proposal networks.

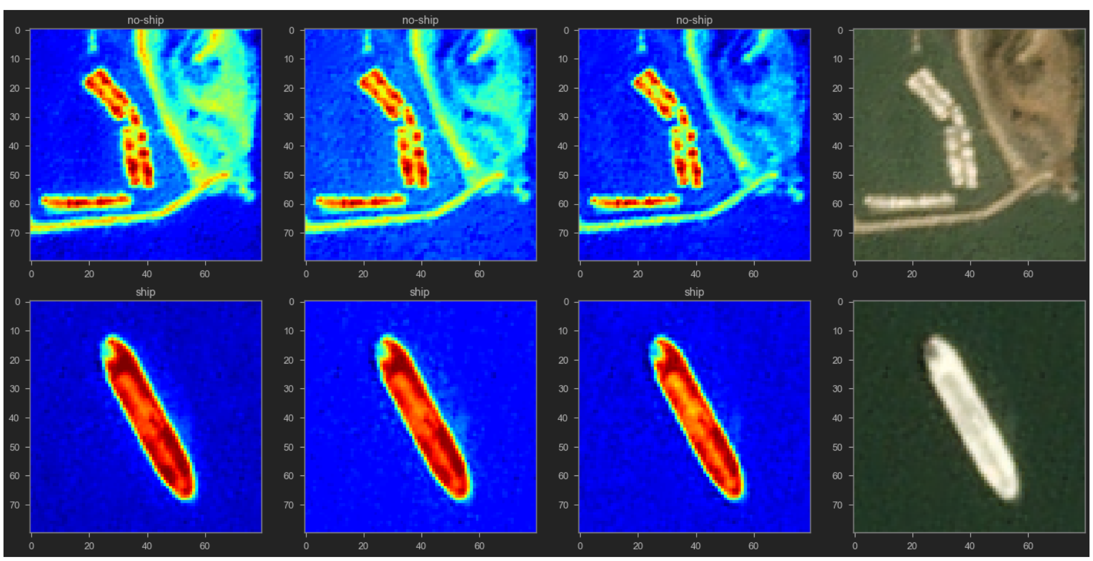

Fig. 4 Sample Ship Object Visualization Results

The Faster R-CNN model that we suggested was successful in detecting the ship object in remote sensing images with more accuracy and efficiency than existing object detection models. Based on the findings shown in Table 2 and Figure 6, we were able to determine that the proposed Faster R-CNN model is capable of accurately identifying ship objects in remote sensing images.

Figure 5 illustrates the outcome of application of the Faster R-CNN model to the detection of ship objects in RSI images.

Table 2 Performance Analysis of ship object detection

S. No.	Model	Acc.	Prec.	Rec.
1.	R-CNN	89.15	88.90	89.1
2.	Fast R-CNN	93.45	93.15	93.40
3.	**Faster R-CNN**	**99.25**	**99.18**	**99.20**

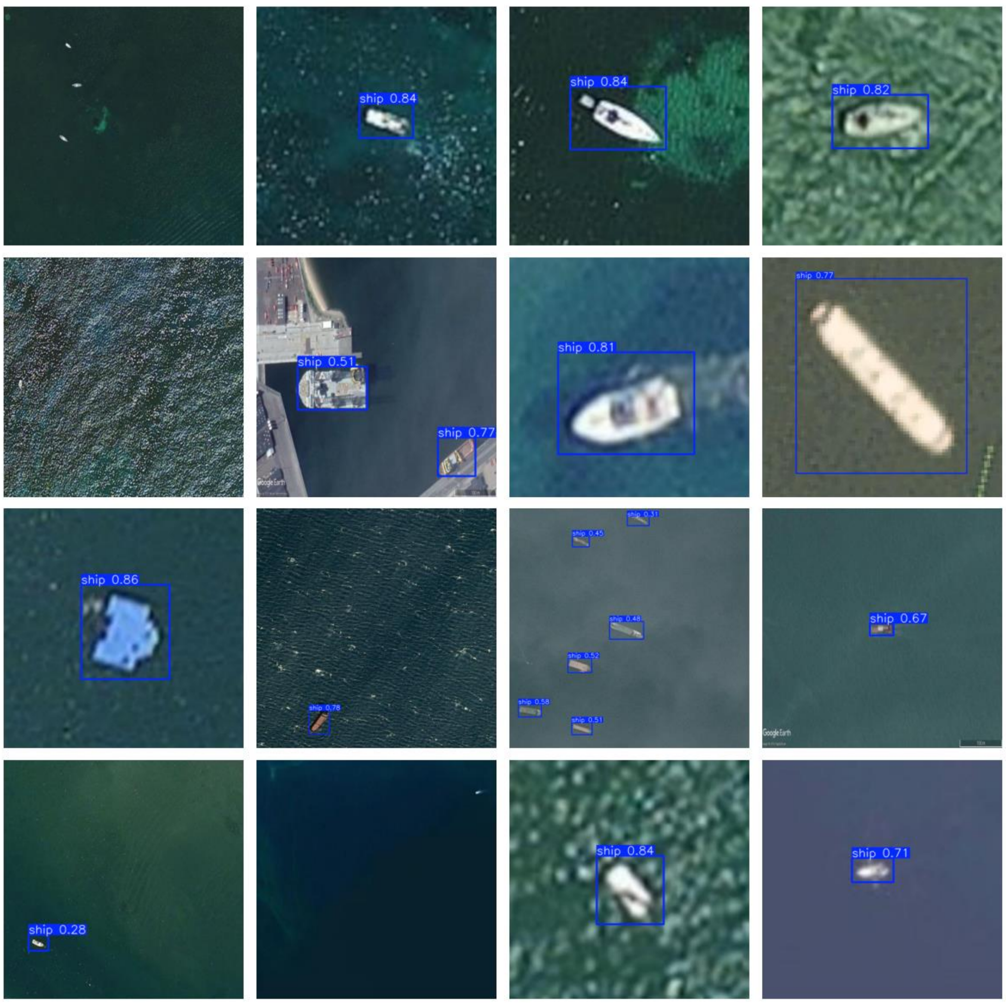

Fig. 5 Sample Ship Object Detection Prediction Results

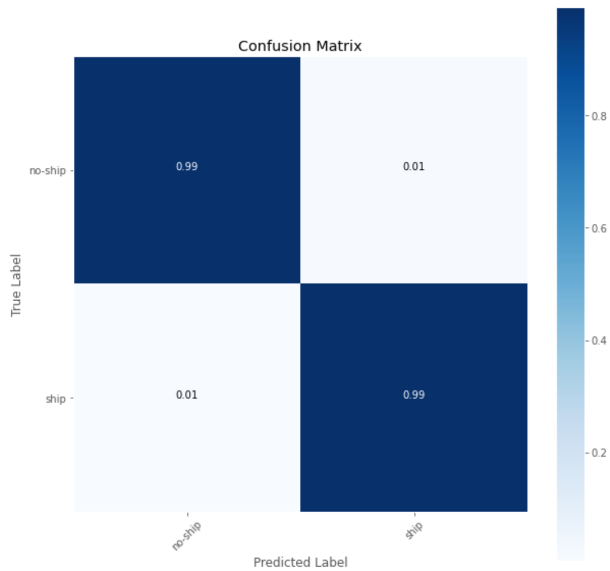

Fig. 6 Test result of ship object detection

V. CONCLUSION

In this research, the proposed Faster R-CNN-based ship detection model demonstrates significant advancements in addressing the challenges of real-time ship detection in satellite and aerial imagery. By leveraging the strengths of CNN and RPN, the model achieves remarkable accuracy and speed, making it a highly effective solution for maritime security, disaster response, and environmental monitoring applications. Its ability to adapt to varying conditions and large-scale datasets further underscores its suitability for real-world deployment. The superior performance metrics, including a detection accuracy of 99.25% with ResNet-50, validate the model's potential to set new benchmarks in remote sensing and maritime surveillance systems, offering a robust and scalable tool for global maritime operations.

VI. REFERENCES

1. P.Deepan and L.R. Sudha, "Object Classification of Remote Sensing Image Using Deep Convolutional Neural Network", The Cognitive Approach in Cloud Computing and Internet of Things Technologies for Surveillance Tracking Systems, pp.107-120, 2020. https://doi.org/10.1016/B978-0-12-816385-6.00008-8.
2. Ghantasala, G. P., Sudha, L. R., Priya, T. V., Deepan, P., & Vignesh, R. R. An Efficient Deep Learning Framework for Multimedia Big Data Analytics. Multimedia Computing Systems and Virtual Reality, 99.
3. Stofa, M.M.; Zulkifley, M.A.; Zaki, S.Z.M. A deep learning approach to ship detection using satellite imagery. IOP Conf. Ser. Earth Environ. Sci. 2020, 540, 012049.

4. P.Deepan and L.R. Sudha, “Deep Learning and its Applications related to IoT and Computer Vision”, Artificial Intelligence and IoT: Smart Convergence for Eco-friendly Topography, Springer Nature, pp. 223-244, 2021, https://doi.org/10.1007/978-981-33-6400-4_11.
5. Gallego, A.-J.; Pertusa, A.; Gil, P. Automatic Ship Classification from Optical Aerial Images with Convolutional Neural Networks. Remote Sens. 2018, 10, 511.
6. P.Deepan and L.R. Sudha, "Scene Classification of Remotely Sensed Images using Ensembled Machine Learning Models", Proceedings in Lecturer Notes on Electrical Engineering, Springer Nature, pp.535-550, 2021, https://doi.org/10.1007/978-981-16-0289-4_39
7. P.Deepan and L.R. Sudha, “Comparative Analysis of Remote Sensing Images using Various Convolutional Neural Network”, EAI End. Transaction on Cognitive Communications, 2021. ISSN: 2313-4534, doi: 10.4108/eai.11-2-2021.168714.
8. Z. Y. Han and J. S. Chong, “A Review of Ship Detection Algorithms in Polarimetric SAR Images,” International Conference on Signal Processing (ICSP 04), IEEE press, vol. 3, Sept. 2004, pp. 2155-2158, doi: 10.1109/ICOSP.2004.1442203.
9. H. Greidanus, P. Clayton, N. Suzuki, and P. Vachon, “Benchmarking Operational SAR Ship Detection,” International Geoscience and Remote Sensing Symposium (IGARSS 04), IEEE press, vol. 6, Dec. 2004, pp. 4215-4218, doi: 10.1109/IGARSS.2004.1370065.
10. C.C.Wackerman, K.S.Friedman, and X. Li, “Automatic Detection of Ships in RADARSAT-1 SAR Imagery,” Canadian Journal of Remote Sensing, vol. 27, Jul. 2014, pp. 568-577.
11. J. Antelo, G. Ambrosio, and C. Galindo, “Ship Detection and Recognition in High-resolution Satellite Images,” International Geoscience and Remote Sensing Symposium (IGARSS 09), IEEE press, vol. 4, Feb. 2010, pp. 514-517, doi: 10.1109/IGARSS.2009.5417426.
12. 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 Transactions on Geoscience and Remote Sensing, vol. 48, Sept. 2010, pp. 3446-3456, doi: 10.1109/TGRS.2010.2046330.
13. P.Deepan and L.R. Sudha, “Remote Sensing Image Scene Classification using Dilated Convolutional Neural Networks”, International Journal of Emerging Trends in Engineering Research, Vol. 8, No.7, pp.3622-3630, 2020, ISSN: 2347-3983.
14. Huang, X.; Zhang, B.; Perrie, W.; Lu, Y.; Wang, C. A novel deep learning method for marine oil spill detection from satellite synthetic aperture radar imagery. Mar. Pollut. Bull. 2022, 179, 113666.
15. Yokoya N and Iwasaki A, Object detection based on sparse representation and Hough voting for optical remote sensing imagery IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. Vol. 8, PP. 2053-2062, 2015.
16. Prasad D K, Prasath C K, Rajan D, Rachmawati L, Rajabaly E and Quek C, Challenges in video based object detection in maritime scenario using computer vision CoRR, pp 1-6, 2016.

17. Yu J G, Xia G, Deng J and Tian J, Small object detection in forward-looking infrared images with sea clutter using context-driven Bayesian saliency model Infrared Phys. Technol. 73, PP. 175-83, 2015.
18. Ships in Satellite Imagery, “Shipsnet”. Kaggle. 2018. Available online: https://www.kaggle.com/rhammell/ships-in-satelliteimagery (accessed on 28 October 2021).