Tech Science Press

Deep Neural Network Based Detection and Segmentation of Ships for Maritime Surveillance

Kyamelia Roy[1], Sheli Sinha Chaudhuri[1], Sayan Pramanik[2] and Soumen Banerjee[2,*]

[1]Department of Electronics and Tele-Communication Engineering, Jadavpur University, Kolkata, 700032, India
[2]Department of Electronics and Communication Engineering, University of Engineering and Management, Kolkata, 700160, India
*Corresponding Author: Soumen Banerjee. Email: prof.sbanerjee@gmail.com

Abstract: In recent years, computer vision finds wide applications in maritime surveillance with its sophisticated algorithms and advanced architecture. Automatic ship detection with computer vision techniques provide an efficient means to monitor as well as track ships in water bodies. Waterways being an important medium of transport require continuous monitoring for protection of national security. The remote sensing satellite images of ships in harbours and water bodies are the image data that aid the neural network models to localize ships and to facilitate early identification of possible threats at sea. This paper proposes a deep learning based model capable enough to classify between ships and no-ships as well as to localize ships in the original images using bounding box technique. Furthermore, classified ships are again segmented with deep learning based auto-encoder model. The proposed model, in terms of classification, provides successful results generating 99.5% and 99.2% validation and training accuracy respectively. The auto-encoder model also produces 85.1% and 84.2% validation and training accuracies. Moreover the IoU metric of the segmented images is found to be of 0.77 value. The experimental results reveal that the model is accurate and can be implemented for automatic ship detection in water bodies considering remote sensing satellite images as input to the computer vision system.

Keywords: Auto-encoder; computer vision; deep convolution neural network; satellite imagery; semantic segmentation; ship detection

1 Introduction

The booming shipping industry with incredible rise in its number of ships and cargoes has led to prolific growth in maritime transportation accounting for 90% of international trade. It not only supports domestic and global manufacturing companies through transportation of their commodities and products; it also helps in delivering goods directly to consumers. Shipping is the lifeline of worldwide economy for export/import of various raw items and consumer products and thus special care needs to be adopted for global maritime safety. Even though maritime transportation has been a dominant support of global trade, its advancement also leads to traffic violations in the waterways and thus in spite of the growing traffic, the access to major trade commodities remains solely the main driving point in setting the maritime

networks. The maritime trade has truly been struck in recent years owing to detection of illegal activities involving maritime terrorism, smuggling and sea-jacking [1]. These menaces put threat to global, regional and national economies; which if not curbed would lead to catastrophic disaster. Hence pledge for safe navigation of sea vessels and safety of sea activities demands surveillance of ocean ships in coastal countries [2,3]. It is herein that the automatic ship detection becomes utmost necessity for maritime security management and surveillance [4]. The primary functions of it are traffic flow monitoring, prevention of marine pollution and detection of illegal fishing and cargo transportation. Intelligent detection of ships by means of automation and computer vision makes the task easier and flexible. Sometimes, infiltration of the intruders through waterways causes big threat to the nation's security. Here, movement of water vessels in irregular pathways is an identification of abnormality in the oceanic perimeter. Thus, computer aided detection system will be efficient in identification of such anomalies. In case of military, automatic ship detection helps in enhancement of maritime security through Intelligence, Surveillance and Reconnaissance (ISR) efforts [5]. A wide range of functionaries in the shipping industry encompasses the defense of territory and naval battles, dynamic harbor surveillance, monitoring of traffic and sea pollution, management of fisheries, etc. [6]. ISR involves development of highly sophisticated sensor systems required for collection of increased volumes of heterogeneous data. Some typical sensors on-board an ISR enabled maritime patrol aircraft includes Electro-Optical/Infrared (EO/IR) camera, radar, Electronic Support Measures (ESM), etc. They facilitate capturing of environmental, individual and conventional signatures and subsequently generate large volumes of collected data.

The critical job involving maritime security and civil management is typically accomplished with the help of Automated Identification System (AIS), which employs radio frequencies in the VHF band to wirelessly broadcast the location of the ships to the nearby receivers on other ships and land based systems [7]. The AIS is effective only when connected with transponder device on the ships; however the functionality ruptures if the transponder is disconnected or not installed on the ships. Satellite imagery under this condition is very helpful in detection and identification of the ships in the water bodies with the aid of machine learning algorithms.

The promising technology for collection of ship/sea vessels related necessary data through high resolution images obtained using satellites and aerial remote sensing devices [8] are recent topics of worldwide research. Moreover during cyclones, large number of ships in the ocean sinks due to the fact that they remain cut-off from the land based systems. Hence, automatic ship identification with satellite images can assist in finding and rescuing the ships in the heavy cyclones; thereby saving many lives. The detection of ships from the high resolution remote sensing images sometimes become tedious owing to the disturbances like clouds, islands, mist, haze, coastlines, tides, etc. However, optical remote sensing images for ship detection are very popular these days owing to a plethora of applications in defense and civil domains. This satellite imagery can provide real-time position information for navigation management control and maritime search and rescue operation ensuring success and work safety at sea and on inland rivers. Additionally, it contributes to the administration and development of important coastal zones and harbors; endorsing ecological protection and sea health.

Ship detection through remote sensing has gained lots of importance owing to the availability of high-resolution Synthetic Aperture Radar (SAR) images suitable for object detection and environment monitoring. These images are capable of providing high-resolution images of the oceans both during day as well as during night. It has proved to be an effective technology for detection of ships and a hot research area throughout the globe [9,10]. These images are weather independent and are most suitable for monitoring maritime activities like ship detection and oil spills. The quad-polarization SAR (QP SAR) mode, amongst the different available polarimetric SAR modes, is found to capture the richest information of the observed area [11]. However, the linear dual-polarization SAR mode having lower system complexity supports a wider swath width [11] while the compact polarimetric SAR (CP SAR) mode provides a

compromise between swath width and scattering information [12]. In [13], the authors have developed a novel algorithm for ships detection using low resolution SAR imagery for ships greater than 35 m length and using high resolution SAR imagery for ships greater than 32 m length. In [14], a segmentation method from CP SAR images is proposed for detection of ships in which pixel-wise detection is based on a fully convolutional network, U-Net. In [15], an enhanced GPU based deep learning method for ship detection using SAR images is reported. The authors proposed a modified deep learning framework called You Only Look Once version 2 (YOLOv2) to model the architecture and train it. Moreover a new architecture with less number of layers called YOLOv2-reduced is also developed. The results exhibited better accuracy in ship detection with appreciable reduction in computational time. In [16], an absolute new detection method is reported to differentiate sea vessels from complex backgrounds from SAR image using proposed local contrast variance weighted information entropy (LCVWIE).

The unique and state-of-the-art performance for detection of ships using optical remote sensing is reported in literature [17,18]. In [17], the authors proposed a fast and robust ship detection algorithm based on deep Convolutional Neural Networks (CNNs). Here, initially deep CNN is used for feature extraction and subsequently a region proposal network (RPN) is applied for discrimination of ship targets. In [18], a Rotation Dense Feature Pyramid Networks (R-DFPN) framework was proposed which was found to detect ships effectively in different scenes including ocean and port. The surveillance video system is also reported to be used for ship detection [19,20]. In [19], the authors applied dynamic fusion technique on background subtraction (BS) and saliency detection (SD) technique results for final boat detection from maritime surveillance videos while in [20], the authors used a vision based autonomous landing algorithm for the same. For smart monitoring concerning effective utilization of port resources, detection and recognition of ships are very vital aspects. However it imposes serious challenges owing to several issues like complex ship profiles, ship background, object occlusion, variations of weather, light conditions, etc. In [21], the authors propose an on-site processing approach called Embedded Ship Detection and Recognition using Deep Learning (ESDR-DL), in which the video stream is processed using embedded devices for accurate ship detection and recognition. In [22], an object detection system based on Histogram of Oriented Gradients (HOG) is proposed for finding ships in maritime videos. The work is further extended in [23], in which the author applied HOG-SVM (Support Vector Machine) detector to a ship detection system for quantitative evaluation. The results were very promising for ships with resolutions larger than 128×64. In [24], a novel multi-level ship detection algorithm is proposed for different offshore ships detection under all possible imaging variations using Multi-Scale Analysis and Fourier HOG Descriptor. In [25], a fully convolutional network (FCN) with task partitioning for inshore ship detection in optical remote sensing images is reported. In [26], the authors proposed novel hierarchical complete and operational ship detection from space-borne optical images (SDSOI) approach based on shape and texture features, which is considered a sequential coarse-to-fine elimination process of false alarms.

The rapid and continuous increase in the hardware computing power has resulted in the faster development of deep learning based algorithms for object detection. Deep learning algorithms have been extensively adopted for remote sensing image analysis. In [27], the authors proposed a novel detection algorithm called region-based deep forest (RDF) comprising of a deep forest ensemble with a simple region proposal network in order to discriminate ship targets. In [28], a CNN based novel method is proposed which starts from a global search for the relatively distinct ship head with an efficient classification network for inshore ship detection. In [29], an algorithm has been developed by the authors for ship detection by combining CNN with constant false alarm rate (CFAR) in which the proposed CNN is based upon the CFAR global detection algorithm and image recognition with CNN model. In [30], an effective novel rotation-invariant CNN (RICNN) model is proposed for detection of objects through introduction of a new rotation-invariant layer. In [31], determination of presence of ships or not from

aerial image of visible spectrum is presented in which neural codes extracted from CNN is combined with k-Nearest Neighbor method to improve performance. In [32], a novel deep feature-based method is developed in order to detect the presence of ships in very high-resolution optical remote sensing images. The authors have used a regional proposal network in order to generate ship candidates from feature maps produced by deep CNN. In [33], a CNN based effective ship detection framework in remote-sensing images is proposed wherein the framework is so designed so as to predict bounding box of ship with orientation angle information in complex remote-sensing scenes. In [34], the authors proposed a saliency-aware CNN for ship detection using real time captured visual images. The proposed CNN framework consisted of exhaustive ship discriminative features like deep feature, saliency map and coastline prior. The authors in [35], with the help of deep networks, put forward a fast regional-based convolutional neural network (R-CNN) method for ship detection from high-resolution remote sensing imagery. This method proved to be an effective one for offshore and inland river ship detection from high-resolution remote sensing imagery. In [36], a ship detection and segmentation method was proposed based on an improved Mask R-CNN model capable of accurate detection and segmentation of ships up to the pixel level. The proposed model could successfully improve the feature maps in describing the features of the target. In [37], the authors developed a CNN based novel method for SAR image change detection. Here the prime idea focused on producing the classification results directly from the original two SAR images through CNN without the use of any pre-processing operations. The developed algorithm was found to be robust in detecting ship targets under complex conditions such as wave clutter background, target in close proximity, ship close to the shore and multi-scale varieties. Even when applied to the change detection of heterogeneous images, the algorithm yielded satisfactory results. In [38], a new method of detection of ships for the land contained sea area is proposed. Here an island filter is used to minimize the existing false alarms from the island; use of CFAR is done to get candidates from the big map and finally CNN based classifier is used to separate false alarms from ship object. A new pooling called max-mean pooling is introduced in order to extract effective features in CNN flow. In addition to the above methods adopted in [38], a further extension is carried through use of threshold segmentation for quick execution of candidate detection [39]. Thereafter a two-layer lightweight CNN model-based classifier is designed to separate false alarms from ship objects at last visualization is achieved through ship prediction in vertical–horizontal (VH) and vertical–vertical (VV) polarization. The model is found to perform with great accuracy for ship image with size less than 32×32. For small ship detection and to ensure autonomous ship safety, the authors in [40] proposed a novel hybrid deep learning method by combining CNN with a modified Generative Adversarial Network (GAN). In [41], the authors proposed a novel ship detection method from remote sensing images based on multi-layer convolutional feature fusion (CFF-SDN). In the paper [42], the authors proposed a deep learning approach for detecting ships in harbour areas using DenseNet architecture as core CNN based classifier.

In the present paper, the detection of ships is done with deep convolution neural networks. The work has been accomplished primarily by classification of the ships present in each image and thereby localizing its presence in the main frame. The classification of 'Ship' and 'No-Ship' are achieved using a 4-layer 2D CNN and the localization is attained by a bounding box at the specified co-ordinates. Followed by classification, the category of images only with ships is segmented by implementation of an auto-encoder with a pre-processed dataset. Hence the proposed model is a multi-neural network based framework and the results obtained are very satisfactory in terms of parametric studies. The proposed model can be realized for an automated ship monitoring or detection system for maritime applications. The remainder of the paper is organized as follows: Section 2 mainly introduces the proposed methodology comprising of dataset, data pre-processing and the network architecture. Section 3 discusses the results with parametric studies. Conclusions and future work are presented in Section 4.

2 Proposed Methodology

In this paper, the detection of ships is accomplished with visible imagery satellite dataset 'Ships in Satellite Imagery' [43] consisting of 4000 images. As the paper represents a multi-neural network based framework *i.e.*, classification as well as segmentation of the images, so dataset preparation is carried out for the mentioned two tasks. The overall workflow method is depicted in Fig. 1. The labeled dataset comprising ship and no ship images are fed to the CNN classifier training model; the model gets trained generating training accuracy and loss. Followed by the training module, a test image is provided which is normally bigger in size than the training dataset. The big size image is then splitted into multiple frames of the size same as training images and are fed to the model. The model detects the presence or absence of ships in these frames and the location of the frames in the original image is stored in an array. The locations (pixel co-ordinates) are then matched in the original big size image and in the matched location if ship is present, a bounding box is activated for marking the detection of ships. On the other hand, the classified images with the presence of ships are fed to an auto-encoder where segmentation is performed with some pre-processing of the images. The segmented outputs of the ships are marked in blue colour.
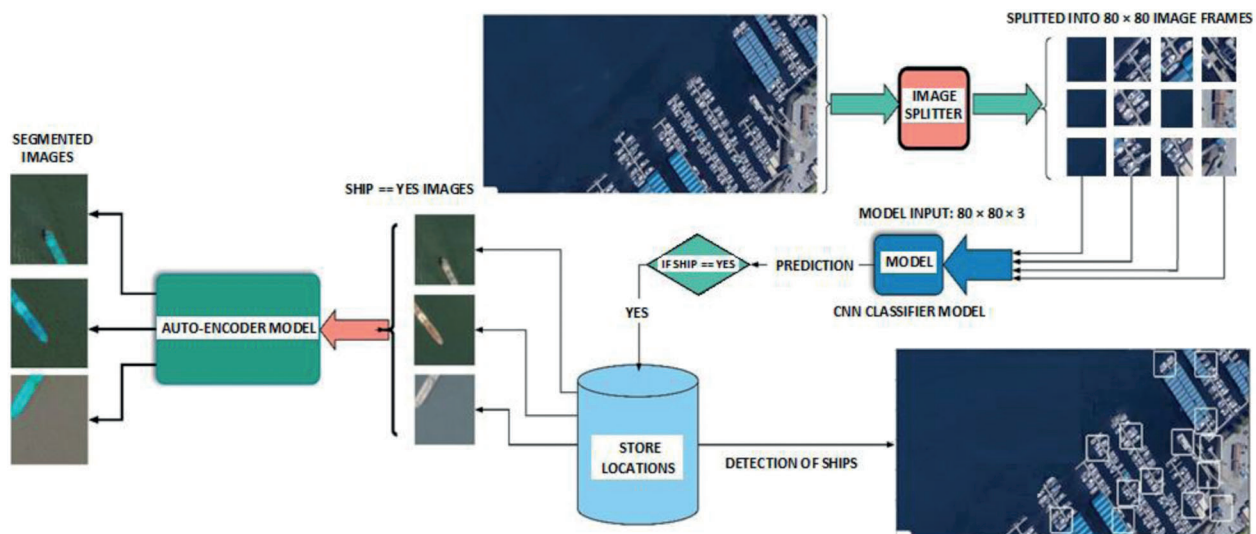
Figure 1: The workflow diagram of the proposed methodology

2.1 Dataset

The dataset is taken from 'Ships in Satellite Imagery' [43] available in Kaggle. These satellite captured imagery with the help of a constellation of small satellites are made available through Planet, a new commercial imagery provider. The dataset is extracted from Planet satellite imagery collected over the San Francisco Bay and San Pedro Bay areas of California. It consists of total 4000, $80 \times 80 \times 3$ RGB labeled, images as 'Ship' and 'No-Ship' categories. The category with 'Ship' contains 1000 images and that of 'No-Ship' contains 3000 images.

The images under 'Ship' class are all near-centered on a single ship body. Inclusions of different types of ships in their various sizes, orientation and under different atmospheric conditions are also present in this dataset. However, for 'No-Ship' class, the dataset comprises of (a). images not having any ship portions but having different land cover features like water, vegetation, bare earth, buildings, etc., (b). images covering certain ship portions only *i.e.*, partial-ships, and (c). images mislabeled owing to bright pixels or strong linear features. Further, for the efficient working of the CNN model, more data are generated from the available dataset with the use of data augmentation technique. 16000 total images are considered for

both classification and segmentation purposes. 8000 images generated with 'Ship' and 4800 with 'No-Ship' are taken into account for training purpose while the rest 3200 images comprising of a mixture of ship, no-ship and some land cover areas are used for validation of the CNN model. Some typical dataset sample images are exhibited in Fig. 2.
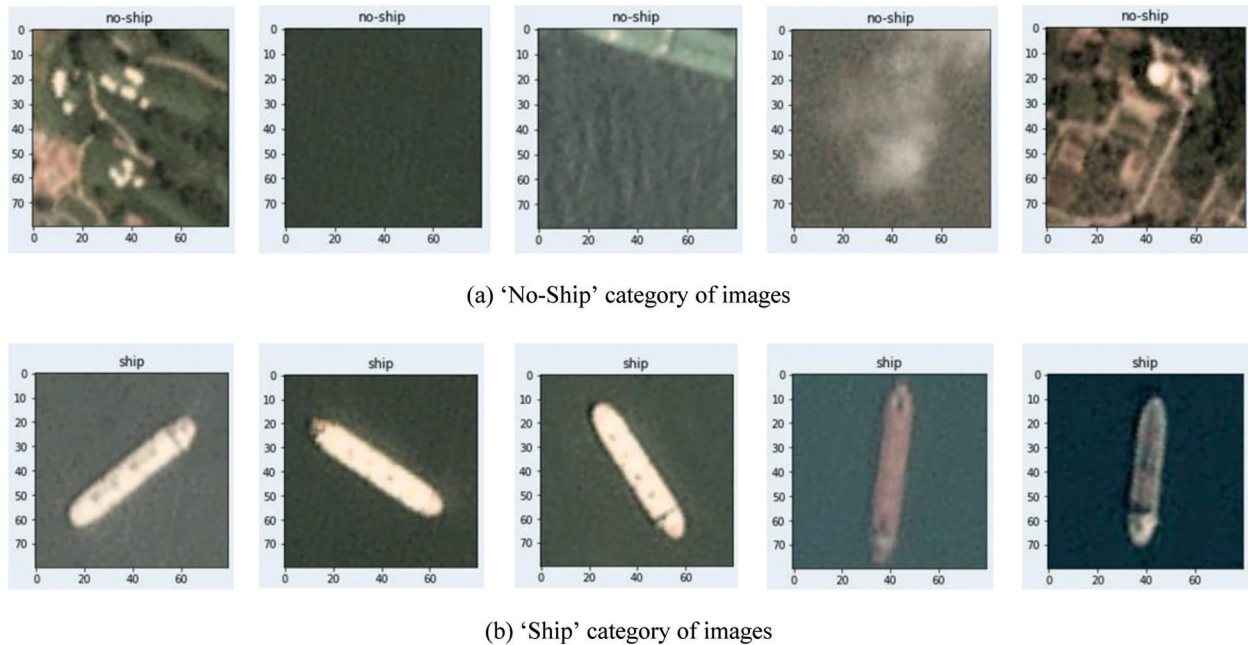
(a) 'No-Ship' category of images

(b) 'Ship' category of images

Figure 2: Typical dataset sample images for ships and no-ships

2.2 Data Pre-Processing

The data pre-processing is undergone in two steps for the two tasks–classification and segmentation. For classification purpose, the RGB satellite images are augmented using different operations like shear, rotate and stretch generating 12800 images for training and the remaining 3200 images for validation. The two classes labeled images are shown in Fig. 2. The RGB images before being fetched by the auto-encoder undergo a processing step as shown in Fig. 3. The RGB images classified as ships (Fig. 4a) are first converted to grey images and then by means of thresholding and binarization, masks are produced [44] as shown in Fig. 4b. Followed by masking, the original ship images and the masks undergo bit XOR operation. After the XOR operation, the images containing ships are marked in blue colour while the images with no-ship remains as it is. These images act as the ground truth data (Fig. 4c) for the auto-encoder. The set of these ground truth images are fed to the training module of the auto-encoder system for segmentation tasks.

2.3 Network Architecture

The present work on detection of ships is accomplished using two architectural models–(a). 4-layer CNN for classification purpose and (b). Auto-encoder for segmentation purpose. The CNN used here acts as a binary classifier producing two classes of outputs as ship and no-ship.
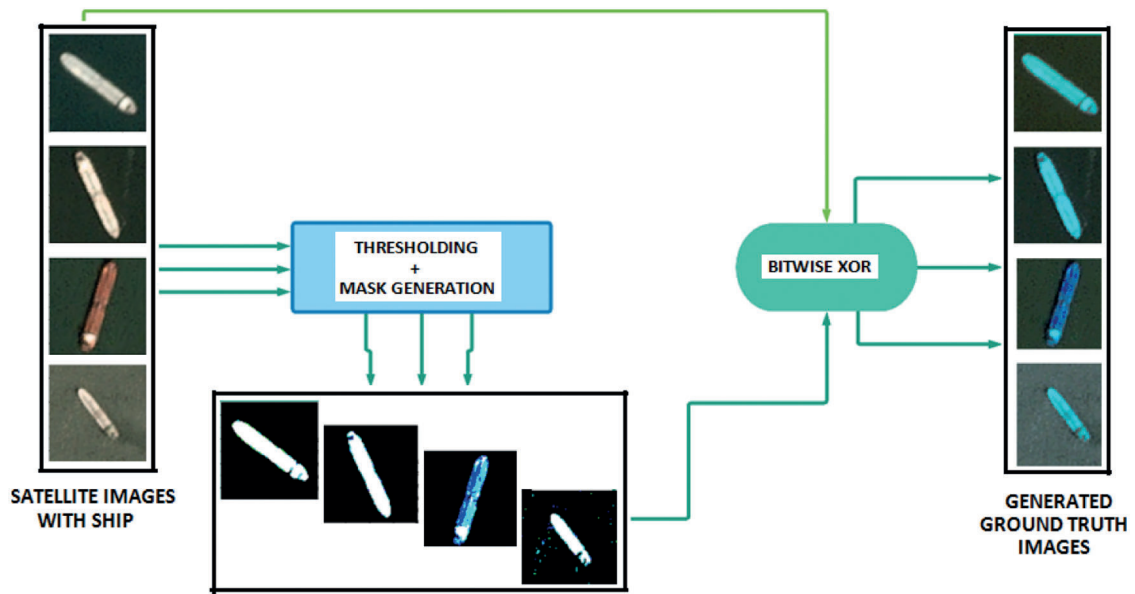
Figure 3: Data preprocessing and generation of ground truth images

(a) Original RGB Images

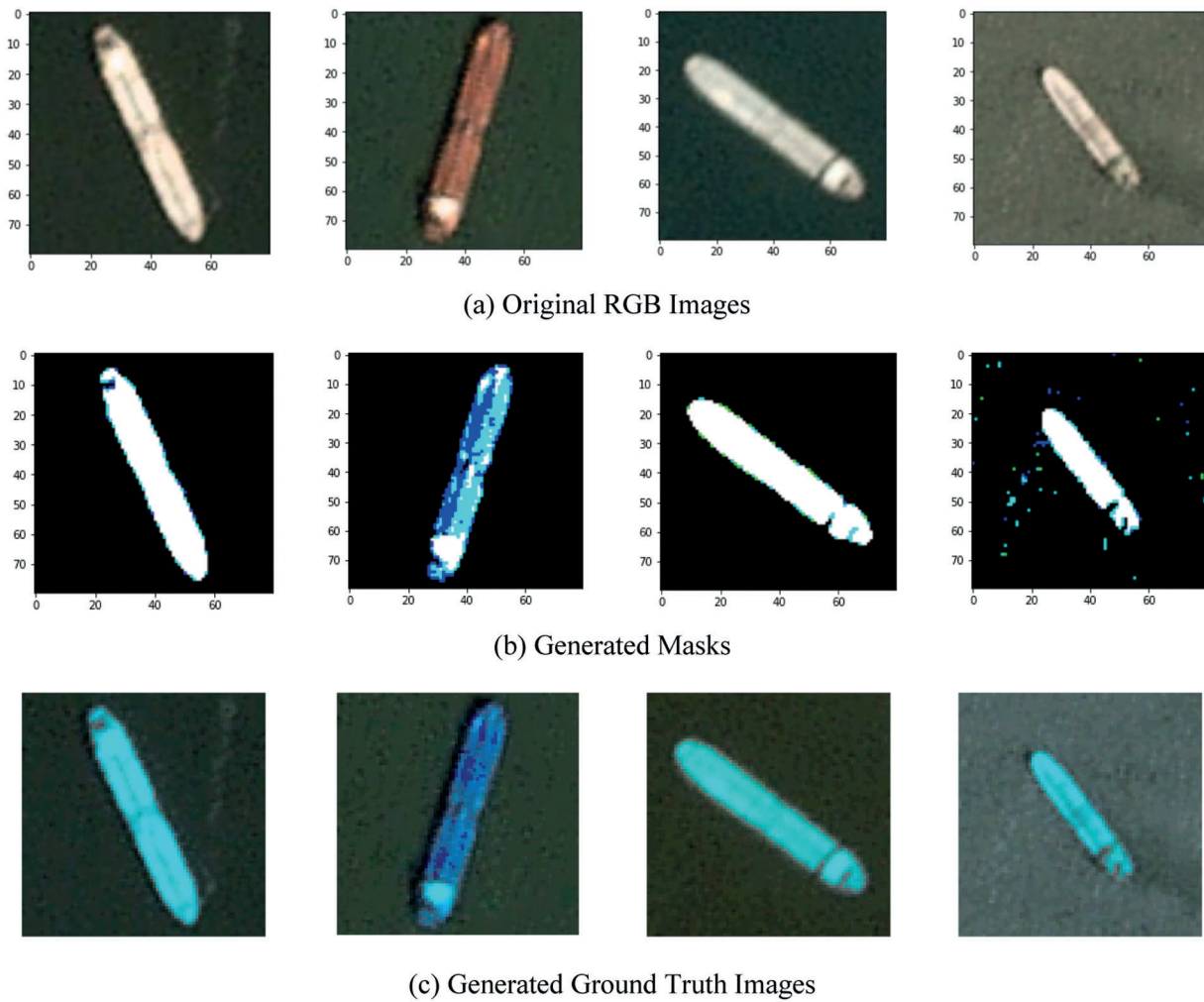(b) Generated Masks

(c) Generated Ground Truth Images

Figure 4: Created dataset: Conversion of RGB (original images) to corresponding binary masks and ground truth images

2.3.1 4-layer CNN for Classification

A 4-layer CNN is implemented for binary classification comprising of CNN 2D, Maxpool 2D and dropout in each layer as shown in Fig. 5. The output layer comprises of the flattening and the dense layer generating the classified outputs. The CNN classifier has been trained with a typical dataset [43]. The dataset has been augmented and prepared for training and validation. The classifier is trained with the augmented dataset and for the testing procedure; a new test input image is fed to the model. The size of the test input image is not the same as that of the training dataset of $(80 \times 80 \times 3)$ size and hence the big size test input image is splitted into multiple grids of size $80 \times 80 \times 3$. The portions of the big sized image are now fed to the system model and the system verifies for the presence or absence of ships in the particular grid. If the classification result comes true then the particular location of the grid in the original frame is stored in an array. Followed by accumulation of the co-ordinates (the grids), the localization of ships in the original image is accomplished by bounding boxes in the locations (stored in the array) of the original test image. The bounding boxes established at different points in the image shows the presence of any ship or vessel in the water body. The network architecture of the CNN classifier is illustrated in Tab. 1.
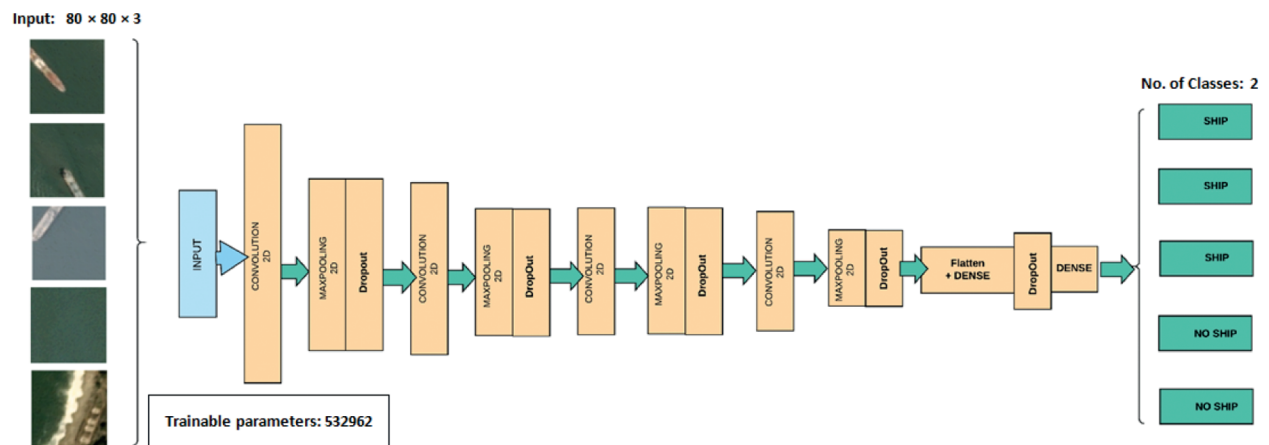

Figure 5: Block diagrammatic representation of 4-layers CNN

2.3.2 Auto-Encoder

A very useful property which the auto-encoder possesses is the automatic learning from the data examples. Auto-encoders do not require any innovative engineering but requires an appropriate training data. The auto-encoder undergoes mainly the compression and decompression functions implemented by the neural networks. It performs segmentation tasks effectively and shows promising results. In this paper, along with classification of ships, the semantic segmentation in ships is also carried out using auto-encoder model. In semantic segmentation, different clusters of pixels are assigned to a specific class. A processed dataset has been created for training of the auto-encoder as shown in Fig. 4. Creation of such masked dataset is of great importance while undergoing semantic segmentation task with auto-encoders. The images which have been classified in the earlier step as 'ship' are now segmented implementing the auto-encoder for a detailed representation. The auto-encoder in general includes the encoder and the decoder blocks. The encoder after fetching the input image data maps the input data into compressed data or latent space where similar points remain close together. On the contrary, the decoder blocks maps those data points to target size output. The decoder block performs the up-sampling of the data points at different layers and reaches to the softmax layer. The working model output is shown in Fig. 6. The network architecture of the auto-encoder is illustrated in Tab. 2.

Table 1: The architecture of 4-layer 2D CNN

Layer (type)	Output shape	Parameter num.
conv2d_10 (Conv2D)	(None, 80, 80, 32)	896
max_pooling2d_6 (MaxPooling2D)	(None, 40, 40, 32)	0
dropout_5 (Dropout)	(None, 40, 40, 32)	0
conv2d_11 (Conv2D)	(None, 40, 40, 32)	9248
max_pooling2d_7 (MaxPooling2D)	(None, 20, 20, 32)	0
dropout_6 (Dropout)	(None, 20, 20, 32)	0
conv2d_12 (Conv2D)	(None, 20, 20, 32)	9248
max_pooling2d_8 (MaxPooling2D)	(None, 10, 10, 32)	0
dropout_7 (Dropout)	(None, 10, 10, 32)	0
conv2d_13 (Conv2D)	(None, 10, 10, 32)	102432
max_pooling2d_9 (MaxPooling2D)	(None, 5, 5, 32)	0
dropout_8 (Dropout)	(None, 5, 5, 32)	0
flatten_1 (Flatten)	(None, 800)	0
dense_2 (Dense)	(None, 512)	410112
dropout_9 (Dropout)	(None, 512)	0
dense_3 (Dense)	(None, 2)	1026

Total parameters: 532,962
Trainable parameters: 532,962
Non-trainable parameters: 0
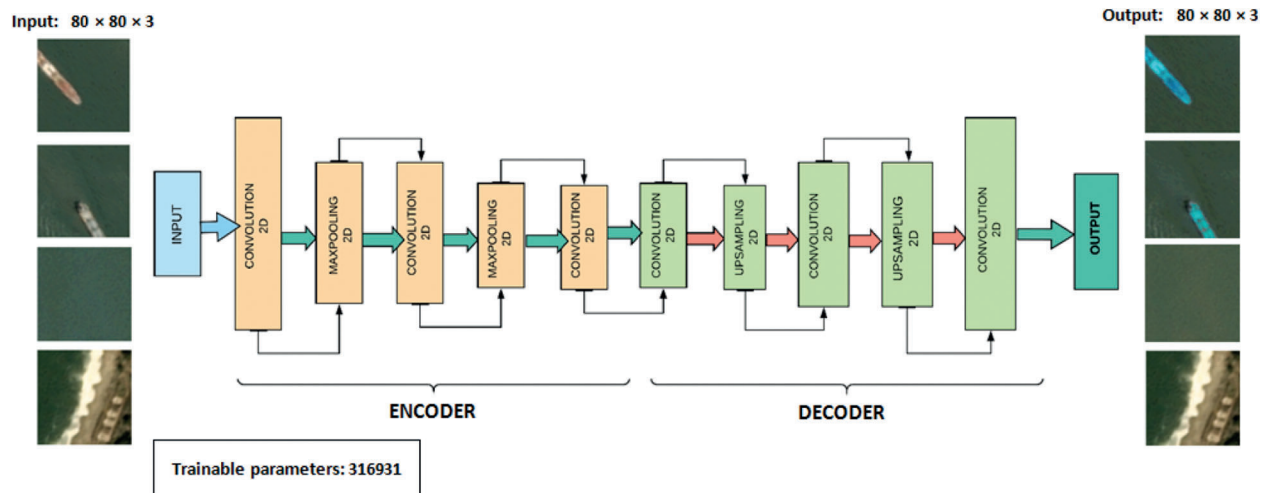
Figure 6: Block diagrammatic representation of auto-encoder

Table 2: The architecture of auto-encoder

Layer (type)	Output shape	Parameter num.
input_2 (InputLayer)	(None, 80, 80, 3)	0
conv2d (Conv2D)	(None, 80, 80, 32)	896
max_pooling2d (MaxPooling2D)	(None, 40, 40, 32)	0
conv2d_1 (Conv2D)	(None, 40, 40, 64)	18496
max_pooling2d_1 (MaxPooling2D)	(None, 20, 20, 64)	0
conv2d_2 (Conv2D)	(None, 20, 20, 128)	73856
conv2d_3 (Conv2D)	(None, 20, 20, 128)	147584
up_sampling2d (UpSampling2D)	(None, 40, 40, 128)	0
conv2d_4 (Conv2D)	(None, 40, 40, 100)	115300
up_sampling2d_1 (UpSampling2D)	(None, 80, 80, 100)	0
conv2d_5 (Conv2D)	(None, 80, 80, 3)	2703
Total parameters: 358,835		
Trainable parameters: 358,835		
Non-trainable parameters: 0		

3 Results and Discussion

An extensive study has been carried out for the proposed work based on simulation of the deep learning architecture. In this paper two types of model architectures have been considered for classification and segmentation procedures. The architecture models have been executed in a system with 4GB DDR5 Graphics Memory and GPU-NIVIDA 740 mx. The implemented 4-layer CNN classifier incorporates the stochastic gradient descent (SGD) optimizer with a learning rate of 0.01 and momentum of 0.9. Cross-entropy specifies the loss function of this typical model. The total number of trainable parameters demonstrated by the model is 532962. The model summary has been reported in Tab. 1. The classification model has been trained and validated against the RGB satellite images. The CNN model has generated very promising results in terms of training and validation accuracy and loss. 99.5% of validation accuracy and 99.2% of training accuracy are achieved with the model for classification task. Apart from training and validation accuracies, other performance parameters have also been studied viz., Precision, Recall, F-score and were proved to be very impressive in terms of their values. The accuracy and loss graphs of the classifier model are shown in Figs. 7a and 7b respectively and the parametric studies are reported in Fig. 8. The confusion matrix for the binary classifier model is also shown in Fig. 8 predicting 590 times true-positive (TP), 2 false-negative (FN), 1 false-positive (FP) and 210 true-negative (TN). From the confusion matrix it is quite evident that the classification for ship and no-ship has been highly accurate with higher values of TP and TN. The Precision, Recall and F-score have also come up with satisfying values validating the work. The output image with the detected ships and the corresponding original image are shown in Fig. 9.

In this paper, it has already been mentioned that the work is accomplished using two approaches, *i.e.*, classification and segmentation. The system is designed with a robust systematic approach to localize as well as segment the anomalies on water bodies from low resolution satellite or drone images. The auto-encoder's performance in terms of accuracy and loss are shown in Fig. 10 attaining 84.2% training and 85.1% validation accuracies which are quite remarkable.
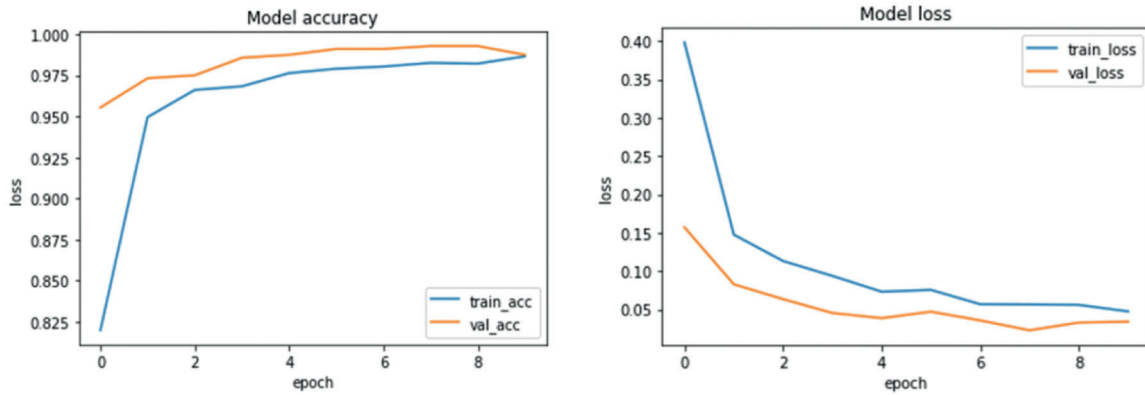
(a) Training and Validation Accuracy of CNN classifier   (b) Training and Validation Loss of CNN classifier

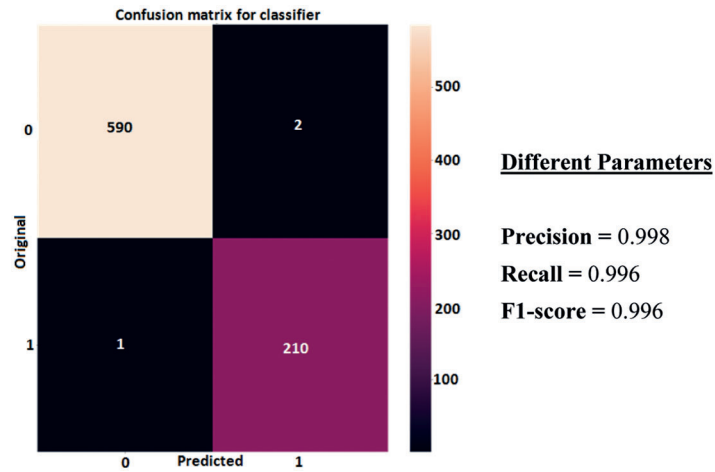**Figure 7:** Accuracy and loss plot of CNN classifier model
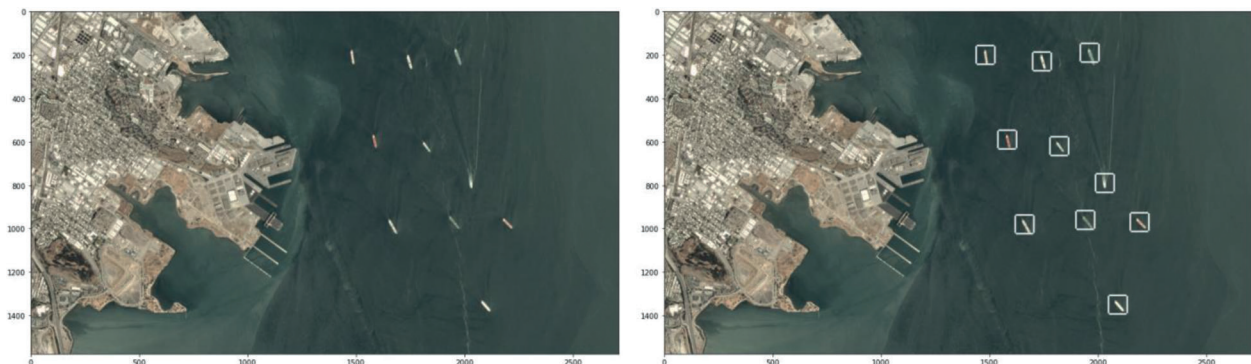

Figure 8: Confusion matrix of CNN classifier model
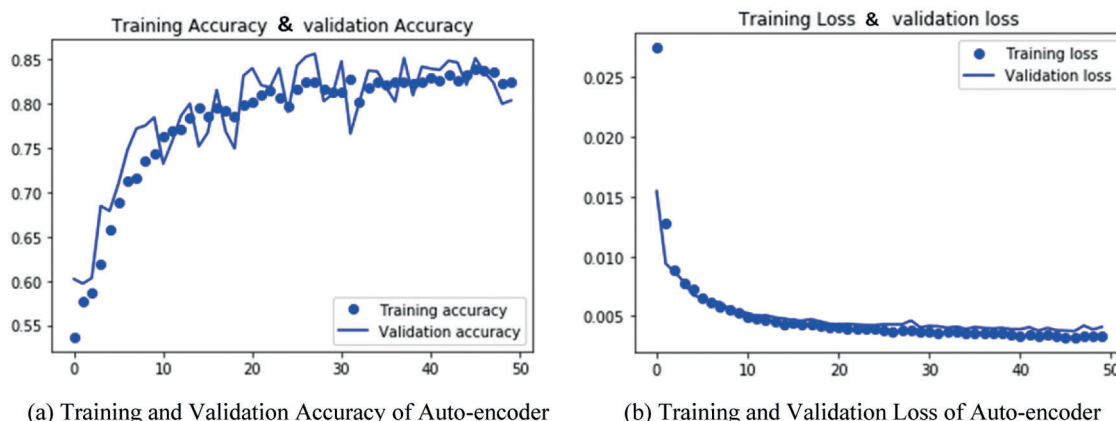

Figure 9: Detection of ships in a test input image

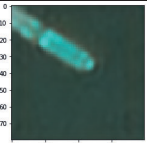(a) Training and Validation Accuracy of Auto-encoder   (b) Training and Validation Loss of Auto-encoder
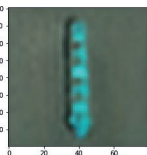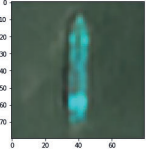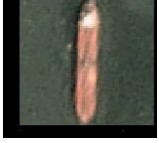
Figure 10: Accuracy and loss plot of auto-encoder model

IoU metric, also termed as Jaccard Index, is very effective and commonly used for evaluation of any semantic segmentation model. The segmented image should have a match of the ground tru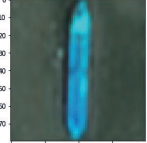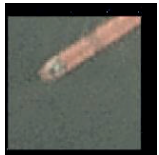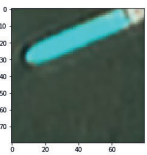th masks from the satellite input image data. The output obtained after segmentation best matches with the generated masks depicting notable output for the segmentation of the ship as shown in Tab. 3. The IoU score is calculated for each sample output as witnessed from Tab. 3. The best IoU achieved is 0.77 with this dataset using the auto-encoder model. For the auto-encoder, the total trainable parameters are 358835. Sigmoid activation function is exercised in the model at the last stage. The model has been compiled in 50 epochs with batch size of 128. The time complexity of the model is 2.5 h approximately for the training with the mentioned specification of the GPU.

Interpretation by the neural network about the decision making regarding classification and segmentation is very important in terms of generated output. Heat maps are the result of the interpretation made by neural network about the accomplishment of the task done by the networks. It is very trivial to know that where the network is looking into or on what features the network is relying for making a decision. Hence in practical real world cases it is absolutely essential to monitor the visualization of the neural network. A sliding window technique is used here and the probability of predictions at each point is calculated. A heat map of the corresponding image is generated using those probability values. The heat maps generated for the sample images are shown in Fig. 11. A comparison of the proposed model with other models for detection of ships as available in literature is enlisted in Tab. 4.

Table 3: IoU metric of the segmented images

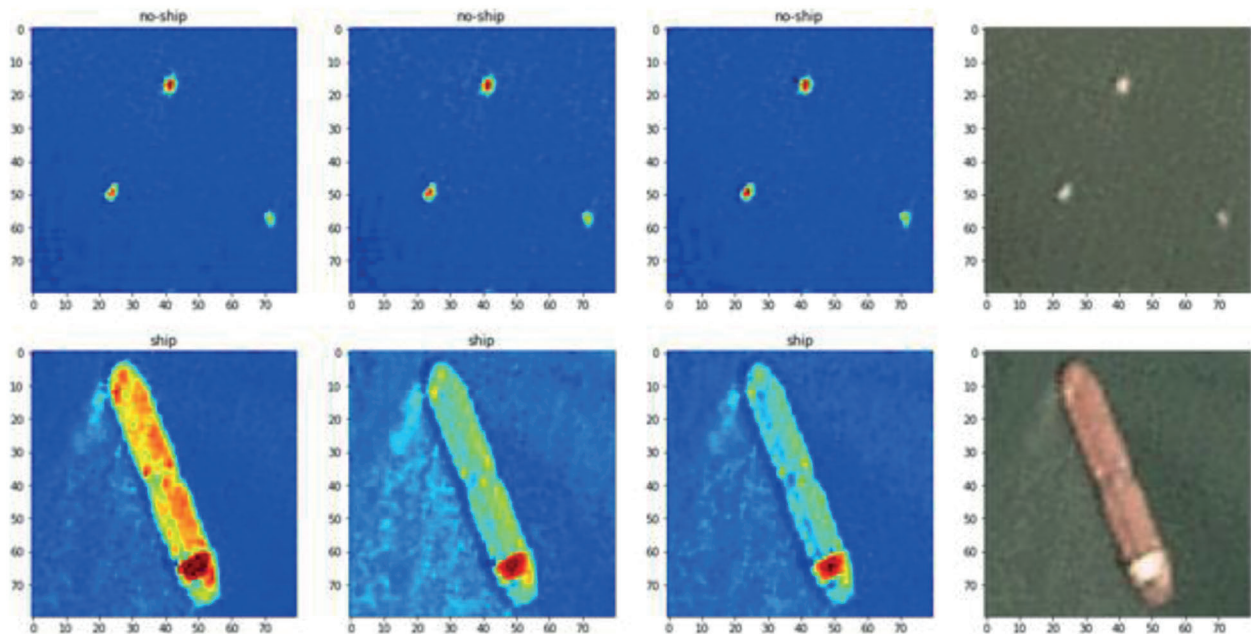| Original Image	Ground Truth Image	Segmented Output	IoU
			0.4
			0.77
			0.35
			0.39
			0.4

Figure 11: Typical heat maps of the sample images

Table 4: Performance analysis of the proposed model with that of other models for ship detection (Parameters: A–Accuracy, P–Precision, R-Recall)

Detection of ships	Architecture/Dataset	Parameters
Ship detection based on YOLOv2 [15]	Faster R-CNN/SSDD YOLOv2/SSDD	A = 70.63%, P = 70.63% A = 90.05%, P = 90.05%
	Faster R-CNN/DSSDD YOLOv2/DSSDD	A = 68.43% A = 89.13%
Ship detection in optical remote sensing images [25]	FCN/Optical remote sensing dataset (Google Earth & GaoFen-2 satellite)	P = 96.1%, R = 91.7%
Ship detection using satellite imagery [42]	DenseNet/Satellite imagery	A = 96.93%
Proposed model	CNN/Satellite imagery	A = 99.5%, P = 99.8%, R = 99.6%

4 Conclusion

Detection and segmentation of ship are carried out in the water body based on satellite images (remote sensing imagery). The satellite RGB images of ships and no-ships are assembled over the San Pedro Bay and San Francisco Bay areas in California. The set of images are fed to the model for classification and thereafter localization and segmentation of the ships in the water body is realized. Results obtained from the CNN classifier model illustrate a validation accuracy of 99.5% and that of the auto-encoder is 85.1%. Other parameters like precision, recall and F1-score have also been studied and extracted in case of the classifier model with very good convincing values. The best IoU score attained by the segmentation model (auto-encoder) is 0.77. The ability of the proposed model can be very much useful in the areas of maritime security management and surveillance, monitoring of traffic, finding and rescuing ships in heavy cyclones, maritime search and rescue operations, navigation management control, etc. The future scope of the present work primarily concerns about the superimposition of the segmented ship images over the original images. Also ship detection can be materialized through remote sensing using high resolution SAR images. But as these images suffer from noise, hence these images can be first de-noised using deep learning approach, often used in dehazing technique, and they can be then further processed to yield the desired result. Hence this approach can also be considered as a future scope of the present work. The sophistication of the model lies in its novel design, ability to classify, locate and segment the object of interest. Real time implementation of such automated system design will reduce the complexity in terms of time, cost, installation and maintenance.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] C. Bueger, "What is maritime security?" *Marine Policy*, vol. 53, no. 1, pp. 159–164, 2015.

[2] P. Iervolino, R. Guida, P. Lumsdon, J. Janoth, M. Clift *et al.,* "Ship detection in SAR imagery: A comparison study," in *Proc. IGARSS*, Fort Worth, TX, USA, pp. 2050–2053, 2017.

[3] S. Brusch, S. Lehner, T. Fritz, M. Soccorsi, A. Soloviev *et al.,* "Ship surveillance with TerraSAR-X," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 3, pp. 1092–1103, 2010.

[4] K. Ward, R. Tough and S. Watts, "Sea clutter: Scattering, the K distribution and radar performance," in *IET: Radar, Sonar and Navigation*. Stevenage: Michael Faraday House, 2013.

[5] D. J. Crisp and T. Keevers, "Comparison of ship detectors for polarimetric SAR imagery," in *Proc. IEEE OCEANS-2010*, Seattle, USA, pp. 1–8, 2010.

[6] He H., Lin Y., Chen F., Tai H-M. and Yin Z., "Inshore ship detection in remote sensing images via weighted pose voting," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 6, pp. 1–17, 2017.

[7] J. F. Vesecky, K. E. Laws and J. D. Paduan, "Using HF surface wave radar and the ship automatic identification system (AIS) to monitor coastal vessels," in *Proc. IGARSS*, Cape Town, South Africa, pp. 761–764, 2009.

[8] H. Su, S. Wei, S. Liu, J. Liang, C. Wang *et al.,* "HQ-ISNet: High-quality instance segmentation for remote sensing imagery," *Remote Sensing*, vol. 12, no. 6, pp. 989, 2020.

[9] D. J. Crisp, *The state-of-the-art in ship detection in synthetic aperture radar imagery*. Salisbury, Australia: Defence Science and Technology Organization Salisbury (Australia) Info Sciences Lab, 2004.

[10] M. Kang, K. Ji, X. Leng and Z. Lin, "Contextual region-based convolutional neural network with multilayer fusion for SAR ship detection," *Remote Sensing*, vol. 9, no. 8, pp. 860, 2017.

[11] F. Charbonneau, B. Brisco and R. Raney, "Compact polarimetry overview and applications assessment," *Canadian Journal of Remote Sensing*, vol. 36, no. sup2, pp. 298–315, 2010.

[12] B. Zhang, X. Li, W. Perrie and O. Garcia-Pineda, "Compact polarimetric synthetic aperture radar for marine oil platform and slick detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 3, pp. 1407–1423, 2017.

[13] C. C. Wackerman, K. S. Friedman, W. G. Pichel, O. N. P. Clemente-Col and X. Li, "Automatic detection of ships in RADARSAT-1 SAR imagery," *Canadian Journal of Remote Sensing*, vol. 27, no. 5, pp. 568–577, 2001.

[14] Q. Fan, F. Chen, M. Cheng, S. Lou, R. Xiao *et al.,* "Ship detection using a fully convolutional network with compact polarimetric SAR images," *Remote Sensing*, vol. 11, no. 18, pp. 2171, 2019.

[15] Y-L. Chang, A. Anagaw, L. Chang, Y. C. Wang, C-Y. Hsiao *et al.,* "Ship detection based on YOLOv2 for SAR imagery," *Remote Sensing*, vol. 11, no. 7, pp. 786, 2019.

[16] W. Huo, Y. Huang, J. Pei, Q. Zhang, Q. Gu *et al.,* "Ship detection from ocean SAR image based on local contrast variance weighted information entropy," *Sensors*, vol. 18, no. 4, pp. 1196, 2018.

[17] Y. Yao, Z. Jiang, H. Zhang, D. Zhao and B. Cai, "Ship detection in optical remote sensing images based on deep convolutional neural networks," *Journal of Applied Remote Sensing*, vol. 11, no. 4, pp. 042611, 2017.

[18] X. Yang, H. Sun, K. Fu, J. Yang, X. Sun *et al.,* "Automatic ship detection in remote sensing images from google earth of complex scenes based on multiscale rotation dense feature pyramid networks," *Remote Sensing*, vol. 10, no. 1, pp. 132, 2018.

[19] T. Tran and T. Le, "Vision based boat detection for maritime surveillance," in *Proc. ICEIC*, Danang, Vietnam, pp. 1–4, 2016.

[20] J. L. Sanchez-Lopez, J. Pestana, S. Saripalli and P. Campoy, "An approach toward visual autonomous ship board landing of a VTOL UAV," *Journal of Intelligent and Robotic Systems*, vol. 74, no. 1–2, pp. 113–127, 2014.

[21] H. Zhao, W. Zhang, H. Sun and B. Xue, "Embedded deep learning for ship detection and recognition," *Future Internet*, vol. 11, no. 2, pp. 53, 2019.

[22] R. Wijnhoven, K. van Rens, E. G. Jaspers and P. H. de With, "Online learning for ship detection in maritime surveillance," in *Proc. SITE 2010*, San Diego, California, USA, pp. 73–80, 2010.

[23] Y. Matsumoto, "Ship image recognition using HOG," *The Journal of Japan Institute of Navigation*, vol. 129, pp. 105–112, 2013.

[24] C. Dong, J. Liu, F. Xu and C. Liu, "Ship detection from optical remote sensing images using multi-scale analysis and fourier HOG descriptor," *Remote Sensing*, vol. 11, no. 13, pp. 1529, 2019.

[25] H. Lin, Z. Shi and Z. Zou, "Fully convolutional network with task partitioning for inshore ship detection in optical remote sensing images," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 10, pp. 1665–1669, 2017.

[26] 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, no. 9, pp. 3446–3456, 2010.

[27] F. Yang, Q. Z. Xu, B. Li and Y. Ji, "Ship detection from thermal remote sensing imagery through region-based deep forest," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 3, pp. 449–453, 2018.

[28] F. Wu, Z. Q. Zhou, B. Wang and J. L. Ma, "Inshore ship detection based on convolutional neural network in optical satellite images," *IEEE Journal of Selected Topics in Applied Earth-observation and Remote Sensing*, vol. 11, no. 11, pp. 4005–4015, 2018.

[29] R. F. Wang, J. Li, Y. P. Duan, H. J. Cao and Y. J. Zhao, "Study on the combined application of CFAR and deep learning in ship detection," *Journal of the Indian Society of Remote Sensing*, vol. 46, no. 9, pp. 1413–1421, 2018.

[30] G. Cheng, P. C. Zhou and J. W. Han, "Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 12, pp. 7405–7415, 2016.

[31] A. J. Gallego, A. Pertusa and P. Gil, "Automatic ship classification from optical aerial images with convolutional neural networks," *Remote Sensing*, vol. 10, no. 4, pp. 511, 2018.

[32] Q. P. Li, L. C. Mou, Q. J. Liu, Y. H. Wang and X. X. Zhu, "HSF-Net: Multiscale deep feature embedding for ship detection in optical remote sensing imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 12, pp. 7147–7161, 2018.

[33] W. C. Liu, L. Ma and H. Chen, "Arbitrary-oriented ship detection framework in optical remote-sensing images," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 6, pp. 937–941, 2018.

[34] Z. Shao, L. Wang, Z. Wang, W. Du and W. Wu, "Saliency-aware convolution neural network for ship detection in surveillance video," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 3, pp. 781–794, 2020.

[35] S. Zhang, R. Wu, K. Xu, J. Wang and W. Sun, "R-CNN-based ship detection from high resolution remote sensing imagery," *Remote Sensing*, vol. 11, no. 6, pp. 631, 2019.

[36] X. Nie, M. Duan, H. Ding, B. Hu and E. K. Wong, "Attention mask R-CNN for ship detection and segmentation from remote sensing images," *IEEE Access*, vol. 8, pp. 9325–9334, 2020.

[37] Y. Li, C. Peng, Y. Chen, L. Jiao, L. Zhou *et al.,* "A deep learning method for change detection in synthetic aperture radar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 8, pp. 5751–5763, 2019.

[38] Z. Wang, T. Yang and H. Zhang, "Land contained sea area ship detection using spaceborne image," *Pattern Recognition Letters*, vol. 130, no. 2, pp. 125–131, 2020.

[39] X. Geng, L. Shi, J. Yang, P. Li, L. Zhao *et al.,* "Ship detection and feature visualization analysis based on lightweight CNN in VH and VV polarization images," *Remote Sensing*, vol. 13, no. 6, pp. 1184, 2021.

[40] C. Zhijun, C. Depeng, Z. Yishi, C. Xiaozhao, Z. Mingyang *et al.,* "Deep learning for autonomous ship-oriented small ship detection," *Safety Science*, vol. 130, no. 9, pp. 104812, 2020.

[41] Y. Zhang, L. Guo, Z. Wang, Y. Yu, X. Liu *et al.,* "Intelligent ship detection in remote sensing images based on multi-layer convolutional feature fusion," *Remote Sensing*, vol. 12, no. 20, pp. 3316, 2020.

[42] M. M. Stofa, M. A. Zulkifley and S. Z. M. Zaki, "A deep learning approach to ship detection using satellite imagery," in *Proc. IGRSM*, Malaysia, 2020.

[43] Ships in Satellite Imagery: [Online]. Available: https://www.kaggle.com/rhammell/ships-in-satellite-imagery.

[44] K. Roy, S. S. Chaudhuri and S. Pramanik, "Deep learning based real-time Industrial framework for rotten and fresh fruit detection using semantic segmentation," *Microsystem Technologies*, vol. 27, no. 9, pp. 3365–3375, 2021.