ShipNet: Post CNN Model-based Ship Extraction from High Resolution Optical Remotely Sensed Images

Zareena Begum[1], Dilip Kumar[2], Ananthoju Chandu[2], Manish Goud Uppalwai[2], Pinaka Sai Akhilesh[2]

[1,2]Department of Information Technology

[1,2] CMR Engineering College, Kandlakoya, Medchal, Hyderabad.

ABSTRACT

The ship extraction from remotely sensed images has attracted much attention. It can supervise fisheries and manage marine traffics to ensure its safety. With the development of satellite and intelligence, automatic ship extraction has replaced the traditional manual ship detection. Ship extraction methods can be divided into two groups by image sources: synthetic aperture radar (SAR) image based and optical image-based methods. The SAR image-based ones have advantages of all-weather time and big difference of ships and sea; therefore, it has extensively been studied. A constant false-alarm rate (CFAR) detector is a usually used ship extraction algorithm, which assumes a certain background distribution for SAR images, such as k-distribution, Gamma distribution, Gauss distribution, and other combination. However, the ship extraction from SAR images also has its limitations, such as low resolution of SAR images, relatively long revisit cycle, and the complicated sea clutter. Optical remotely sensed images have the advantage of high resolution and relatively short revisit cycles, and have more detailed texture, spectral and shape information. Therefore, the ship extraction from optical remotely sensed images were widely studied. Shape features and gray intensity are two mainly used optical image-based ship extraction methods. Shape features based methods use shape and edge information to extract ships. For example, most ships have narrow bow area and parallel hull edges, which are easily detected because of the big difference between ships and water. Many researchers have adopted this idea to detect ships. Due to the big gray intensity distinction between ships, and water, gray intensity was exploited to detect ships by segmenting images. With the development of artificial intelligence, neural network has been widely studied in much research, especially deep learning, and deep neural network.

In this project, a post convolutional neural networks (CNN) method is proposed to extract ships from high resolution optical remotely sensed images. It consists of two parts: ship proposal detection and ship extraction based on CNN. The first part aims to locate possible ships through classification of water and no water, seawater area extraction using mathematical morphology, and ship proposal extraction. Ships are extracted in the second part by implementing a trained CNN on the ship proposals.

Keywords: Ship extraction, Synthetic aperture radar, Constant false-alarm rate.

1. INTRODUCTION

Ship detection on remote sensing images has a wide range of applications in civil areas and defence security. Ship detection with satellite imagery can provide real-time location information for navigation management control and maritime search and rescue, which guarantees the effectiveness and safety of work at sea and on inland rivers, such as ocean transportation supply. It also contributes to the supervision and construction of important coastal zones and harbours, which promotes the protection of the ecology and sea health, offshore areas, and inland rivers.

In view of the existing systems, another approach is to use a target detection algorithm based on high resolution optical remotely sensed images. During the past decades, optical remote sensing images have provided an abundance of shape, outline color, and texture information, and ship detection using

2D object detection algorithms in remote sensing imagery has been extensively studied [1]. The classic methods of ship detection are based on threshold segmentation which requires a favorable condition of the sea surface; however, its detection results are not sufficiently satisfactory. Then, many groups of researchers began to use classifiers such as support vector machine (SVM), AdaBoost, decision trees, etc. [2], which are based on hand-engineered features such as the local binary pattern (LBP), histogram of oriented gradient (HOG), Gabor and so on. In addition, a method based on the mixture of DPMs can detect ships close to each other. However, these classic methods are limited by manually designed image features and templates and encounter bottlenecks when ships vary in size and position. Recently, object detection algorithms based on machine learning, especially deep learning, have been used in both SAR and optical remote sensing. To address the above problem, in this paper, we propose post convolutional neural networks (CNN) method for ship detection on optical remote sensing images.

Problem Statement

The problem statement for ship extraction from high resolution optical remotely sensed image is to develop an automated system that can accurately and efficiently detect and extract ships from large-scale satellite or aerial imagery. The main challenge in ship extraction from high resolution optical remotely sensed images is that ships can have varying sizes, shapes, and orientations, and can be located in complex maritime environments with other objects such as land, clouds, and waves. Additionally, the images themselves may have varying levels of noise, occlusion, and illumination conditions, which can further complicate the ship extraction task.

The goal of ship extraction is to identify the locations, shapes, and sizes of ships in the image, which can be used for a variety of applications, such as maritime surveillance, environmental monitoring, and navigation. To solve this problem, researchers have developed a variety of methods based on deep learning, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs). These methods typically involve preprocessing the images to enhance features such as edges and corners, and then using the trained deep learning model to extract the ships from the image. Other approaches include using hand-crafted features, such as shape and texture, and applying traditional image processing techniques such as thresholding and morphological operations.

2. LITERATURE SURVEY

Zhang et al. adopted the idea of deep networks and presented a fast regional-based convolutional neural network (R-CNN) method to detect ships from high-resolution remote sensing imagery. First, this work choosed GaoFen-2 optical remote sensing images with a resolution of 1 m and preprocess the images with a support vector machine (SVM) to divide the large detection area into small regions of interest (ROI) that may contain ships. Then, this framework applied ship detection algorithms based on a region-based convolutional neural network (R-CNN) on ROI images. To improve the detection result of small and gathering ships, this framework adopted an effective target detection framework, Faster-RCNN, and improve the structure of its original convolutional neural network (CNN), VGG16, by using multiresolution convolutional features and performing ROI pooling on a larger feature map in a region proposal network (RPN).

Li et al. collected the methods of ship detection and classification for practically testing in optical remote sensing images and provided their corresponding feature extraction strategies and statistical data. Basic feature extraction strategies and algorithms are analyzed associated with their performance and application in ship detection and classification. Furthermore, publicly available datasets that can be applied as the benchmarks to verify the effectiveness and the objectiveness of ship detection and classification methods are summarized in this paper.

Hidalgo et al. presented a system for the detection of ships and oil spills using side-looking airborne radar (SLAR) images. The proposed method employed a two-stage architecture composed of three pairs of convolutional neural networks (CNNs). Each pair of networks is trained to recognize a single class (ship, oil spill, and coast) by following two steps: a first network performs a coarse detection, and then, a second specialized CNN obtains the precise localization of the pixels belonging to each class. After classification, a postprocessing stage is performed by applying a morphological opening filter in order to eliminate small look-alikes and removing those oil spills and ships that are surrounded by a minimum amount of coast.

Wang et al. provided a comprehensive survey of recent developments on ship detection and recognition in optical remotely sensed imagery. The influencing factors are analyzed, and the state-of-the-art methods are reviewed. Furthermore, the remaining problems and future development trends are provided for detecting and recognizing ship targets in optical remotely sensed images.

Mou et al. proposed a hierarchical selective filtering layer to map features in different scales to the same scale space. The proposed method is an end-to-end network that can detect both inshore and offshore ships ranging from dozens of pixels to thousands. This framework tested this network on a large ship data set which will be released in the future, consisting of Google Earth images, GaoFen-2 images, and unmanned aerial vehicle data. Experiments demonstrated high precision and robustness of our method. Further experiments on aerial images show its good generalization to unseen scenes.

Wang et al. used SAR images dataset acquired by Sentinel-1. Experimental results revealed that compared with SSD300, SSD512 achieves lower false alarm and slighter lower in detection accuracy. These results demonstrated the effectiveness of this method.

An et al. aimed at detecting ships in GF-3 SAR images using a new land masking strategy, the appropriate model for sea clutter and a neural network as the discrimination scheme. Firstly, the fully convolutional network (FCN) is applied to separate the sea from the land. Then, by analyzing the sea clutter distribution in GF-3 SAR images, this work chooses the probability distribution model of Constant False Alarm Rate (CFAR) detector from K-distribution, Gamma distribution and Rayleigh distribution based on a trade-off between the sea clutter modeling accuracy and the computational complexity. Furthermore, to better implement CFAR detection, this work also used truncated statistic (TS) as a pre-processing scheme and iterative censoring scheme (ICS) for boosting the performance of detector. Finally, this work employed a neural network to re-examine the results as the discrimination stage.

Zhao et al. develop a coupled CNN for small and densely clustered SAR ship detection. The proposed method mainly consists of two subnetworks: an exhaustive ship proposal network (ESPN) for ship-like region generation from multiple layers with multiple receptive fields, and an accurate ship discrimination network (ASDN) for false alarm elimination by referring to the context information of each proposal generated by ESPN. The motivation in ESPN is to generate as many ship proposals as possible, and in ASDN, the goal is to obtain the results accurately.

Hwang et al. proposed an efficient method for detecting ships from SAR imagery using filtering. This method exploited ship masking using a median filter that considers maximum ship sizes and detects ships from the reference image, to which a non-Local means (NL-means) filter is applied for speckle de-noising and a differential image created from the difference between the reference image and the median filtered image. As the pixels of the ship in the SAR imagery have sufficiently higher values than the surrounding sea, the ship detection process is composed primarily of filtering based on this characteristic.

ISSN- 2394-5125 VOL 10, ISSUE 03, 2023

Li et al. analyzed and summarized previous research on the application of deep learning algorithms in ship detection technologies based on SAR and optical remote sensing images in recent years and has provided suggestions for future studies.

3. PROPOSED SYSTEM

Extracting ships from high-resolution optical remotely sensed images is an important task in maritime surveillance and security. Post-CNN layers can be used to improve the performance of CNNs for ship extraction by enabling the network to learn higher-level features and capture more complex patterns in the input data. In the context of ship extraction from high-resolution optical remotely sensed images, the post-CNN section of the network typically consists of fully connected layers followed by a softmax layer. The fully connected layers perform a global aggregation of the features extracted from the convolutional layers, enabling the network to capture higher-level patterns and relationships in the input data. The output of the fully connected layers is typically fed into a softmax layer, which outputs a probability distribution over the possible classes or categories.

One approach to using post-CNN layers for ship extraction is to train the network on a large dataset of annotated images, with the goal of minimizing the difference between the predicted outputs and the true labels of the training data. During training, the post-CNN layers are fine-tuned to optimize the network's performance on the specific task of ship extraction. Another approach is to use transfer learning, where a pre-trained CNN is used as a feature extractor, and the post-CNN layers are added to the network to perform the final classification task. In this approach, the pre-trained CNN has already learned general features from a large dataset of images, and the post-CNN layers are fine-tuned on a smaller dataset of annotated images for ship extraction.

In recent years, researchers have also explored the use of other types of layers in the post-CNN section of the network for ship extraction, such as attention layers, which enable the network to focus on specific regions of the input data, and recurrent layers, which allow the network to model temporal dependencies in sequences of data. Overall, post-CNN layers can improve the performance of CNNs for ship extraction from high-resolution optical remotely sensed images by enabling the network to learn higher-level features and capture more complex patterns in the input data. By fine-tuning the post-CNN layers, the network can be optimized for the specific task of ship extraction, achieving state-of-the-art performance on this important application in maritime surveillance and security.

In this project, a post convolutional neural networks (CNN) method is proposed to extract ships from high resolution optical remotely sensed images. It consists of two parts: ship proposal detection and ship extraction based on CNN. The first part aims to locate possible ships through classification of water and no water, seawater area extraction using mathematical morphology, and ship proposal extraction. Ships are extracted in the second part by implementing a trained CNN on the ship proposals.

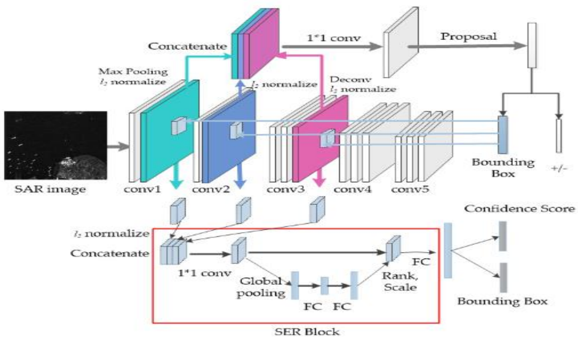


Fig. 1: Proposed architecture for ship extraction from SAR image.

Post CNN

"Post-CNN" is a term that is sometimes used to refer to the layers of a neural network that come after the convolutional layers in a typical CNN architecture. The post-CNN layers in a CNN architecture can be of different types, but they are typically composed of fully connected layers, sometimes followed by other types of layers such as pooling or activation layers. These layers are responsible for transforming the feature maps obtained from the convolutional layers into a form that can be used for the final classification task.

The fully connected layers in the post-CNN section of the network perform a global aggregation of the features extracted from the convolutional layers, enabling the network to capture higher-level patterns and relationships in the input data. The output of the fully connected layers is typically fed into a softmax layer, which outputs a probability distribution over the possible classes or categories. The post-CNN layers are often fine-tuned during the training process to optimize the network's performance on the specific task at hand, such as image classification, object detection, or segmentation.

CNN layer description

According to the facts, training and testing of CNN involves in allowing every source data via a succession of convolution layers by a kernel or filter, rectified linear unit (ReLU), max pooling, fully connected layer and utilize SoftMax layer with classification layer to categorize the objects with probabilistic values ranging from.

Convolution layer is the primary layer to extract the features from a source image and maintains the relationship between pixels by learning the features of image by employing tiny blocks of source data. It's a mathematical function which considers two inputs like source image $I(x, y, d)$ where x and y denotes the spatial coordinates i.e., number of rows and columns. d is denoted as dimension of an image (here d=3 since the source image is RGB) and a filter or kernel with similar size of input image and can be denoted as $F\left(k_x, k_y, d\right)$..

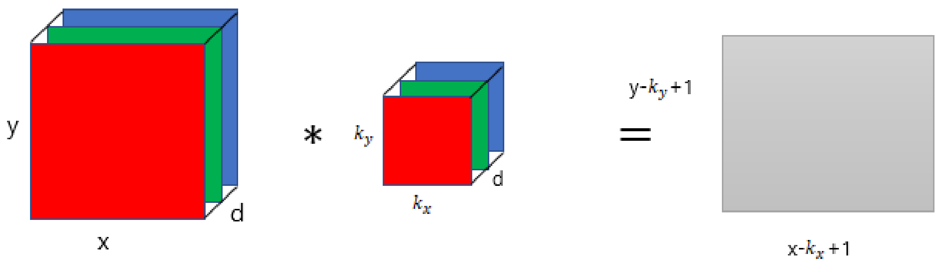


Fig. 2: Representation of convolution layer process.

The output obtained from convolution process of input image and filter has a size of $C\left((x - k_x + 1), (y - k_y + 1), 1\right)$, which is referred as feature map. Let us assume an input image with a size of 5×5 and the filter having the size of 3×3. The feature map of input image is obtained by multiplying the input image values with the filter values.

1	1	1	0	0
0	0	1	1	1
1	1	0	0	1
0	0	0	1	1
1	1	1	0	0

5x5 image

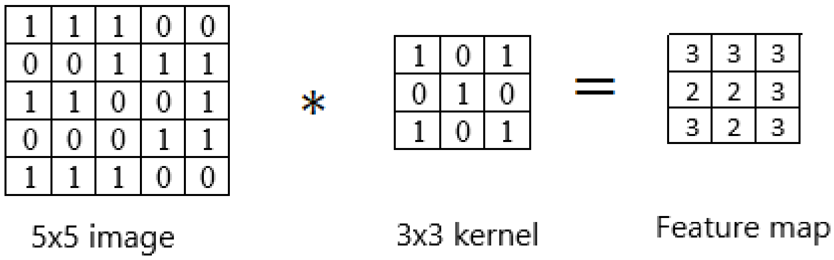
*

1	0	1
0	1	0
1	0	1

3x3 kernel

ISSN- 2394-5125 VOL 10, ISSUE 03, 2023

(a)



(b)

Fig. 3: Example of convolution layer process (a) an image with size 5×5 is convolving with 3×3 kernel (b) Convolved feature map.

3.4.1 ReLU layer

Networks those utilizes the rectifier operation for the hidden layers are cited as rectified linear unit (ReLU). This ReLU function $\mathcal{G}(\cdot)$ is a simple computation that returns the value given as input directly if the value of input is greater than zero else returns zero. This can be represented as mathematically using the function $max(\cdot)$ over the set of 0 and the input x as follows:

$$\mathcal{G}(x) = \max\{0, x\}$$

3.4.2 Max pooing layer

This layer mitigates the number of parameters when there are larger size images. This can be called as subsampling or down sampling that mitigates the dimensionality of every feature map by preserving the important information. Max pooling considers the maximum element form the rectified feature map.

3.5 Softmax classifier

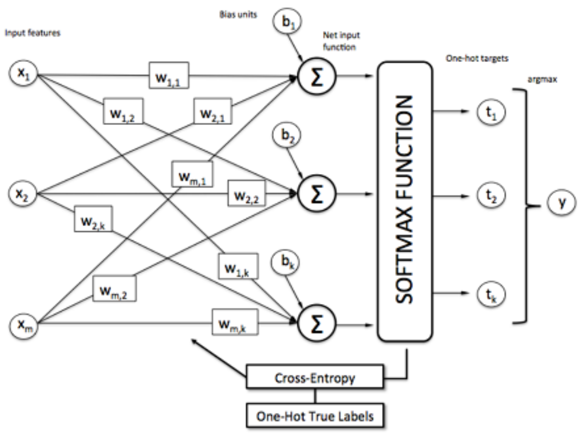


Fig. 4: Crop disease prediction using SoftMax classifier.

Generally, as seen in the above picture softmax function is added at the end of the output since it is the place where the nodes are meet finally and thus, they can be classified. Here, X is the input of all the models and the layers between X and Y are the hidden layers and the data is passed from X to all the layers and Received by Y. Suppose, we have 10 classes, and we predict for which class the given input belongs to. So, for this what we do is allot each class with a particular predicted output. Which means that we have 10 outputs corresponding to 10 different class and predict the class by the highest probability it has.

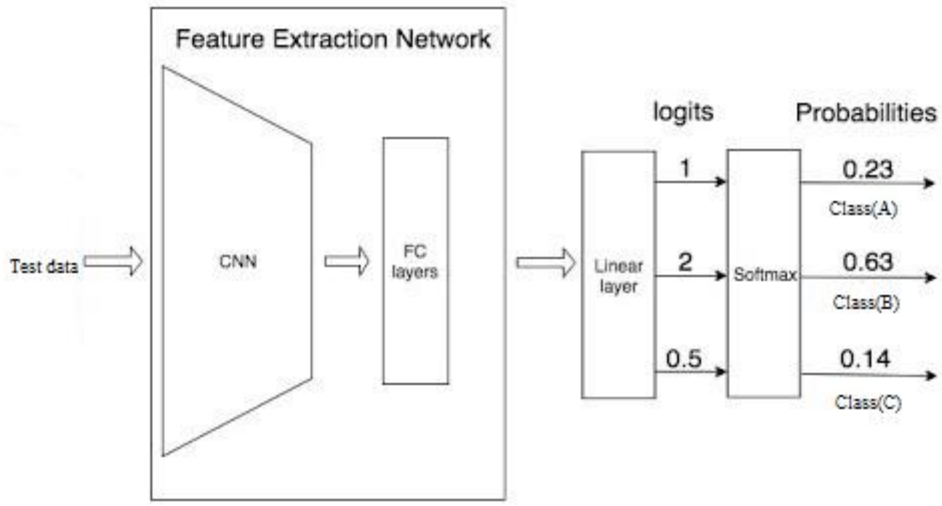


Fig. 5: Example of SoftMax classifier.

In Fig. 4.6, and we must predict what is the object that is present in the picture. In the normal case, we predict whether the crop is A. But in this case, we must predict what is the object that is present in the picture. This is the place where softmax comes in handy. As the model is already trained on some data. So, as soon as the picture is given, the model processes the pictures, send it to the hidden layers and then finally send to softmax for classifying the picture. The softmax uses a One-Hot encoding Technique to calculate the cross-entropy loss and get the max. One-Hot Encoding is the technique that is used to categorize the data. In the previous example, if softmax predicts that the object is class A then the One-Hot Encoding for:

Class A will be [1 0 0]

Class B will be [0 1 0]

Class C will be [0 0 1]

From the diagram, we see that the predictions are occurred. But generally, we don't know the predictions. But the machine must choose the correct predicted object. So, for machine to identify an object correctly, it uses a function called cross-entropy function.

So, we choose more similar value by using the below cross-entropy formula.

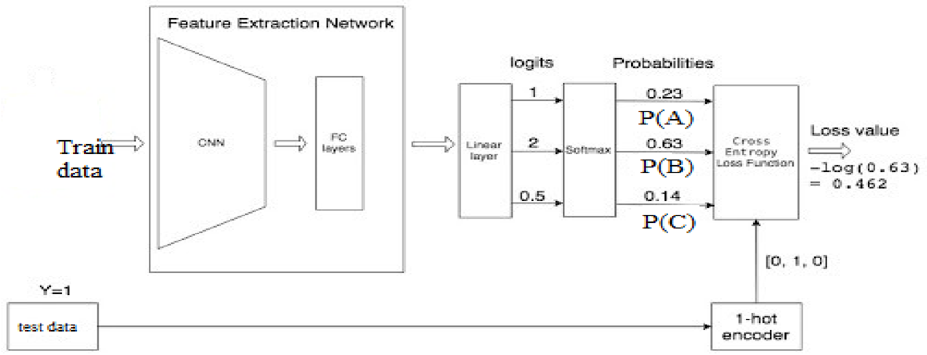


Fig. 6: Example of SoftMax classifier with test data.

In the above example we see that 0.462 is the loss of the function for class specific classifier. In the same way, we find loss for remaining classifiers. The lowest the loss function, the better the prediction is. The mathematical representation for loss function can be represented as: -

$$LOSS = np.sum(-Y * np.log(Y\_pred))$$

4. RESULTS AND DISCUSSION

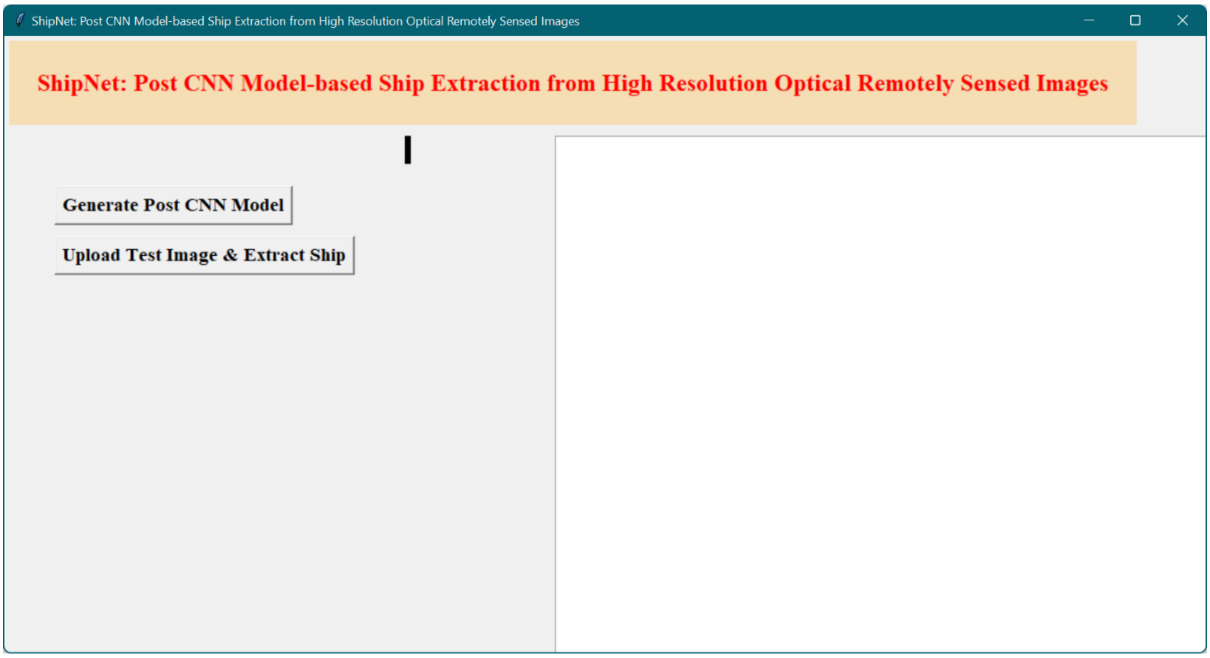
Modules Information

This project consists of following modules:

- Generate post CNN Model: In this module a train post CNN model will be generated using Squeeze and Excitation. Input data to this module is given from 'VGGImageNet.h5py' file. Post CNN extract all features from this file and build a train model, while building model it will read all data from file and then using Squeeze and Excitation technique accept top K features and squeeze irrelevant features. Post CNN internally uses CNN pooling technique to build mode. Below code describe model generation for train images and use five convolution layers.
- Upload Test Image & Detect Ship: In this module we will upload test image and then application extract features from this test image and then apply post CNN train model on that test image to detect ships.

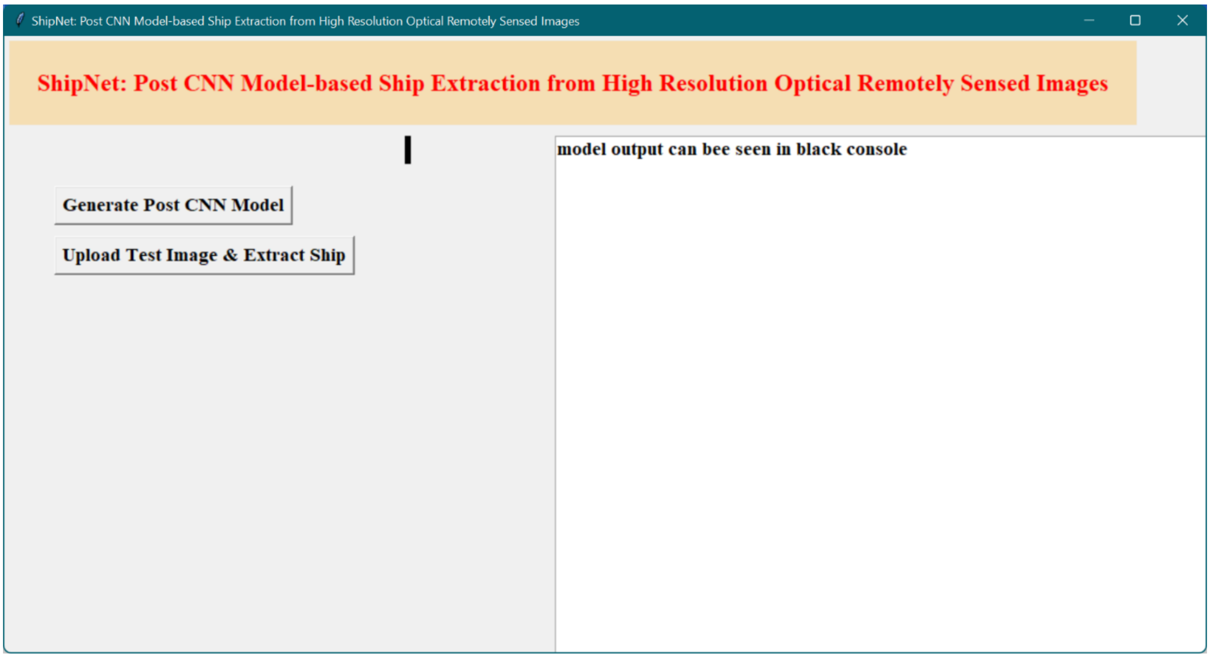
Screen shots

To run this project double click on 'run.bat' file to get below screen

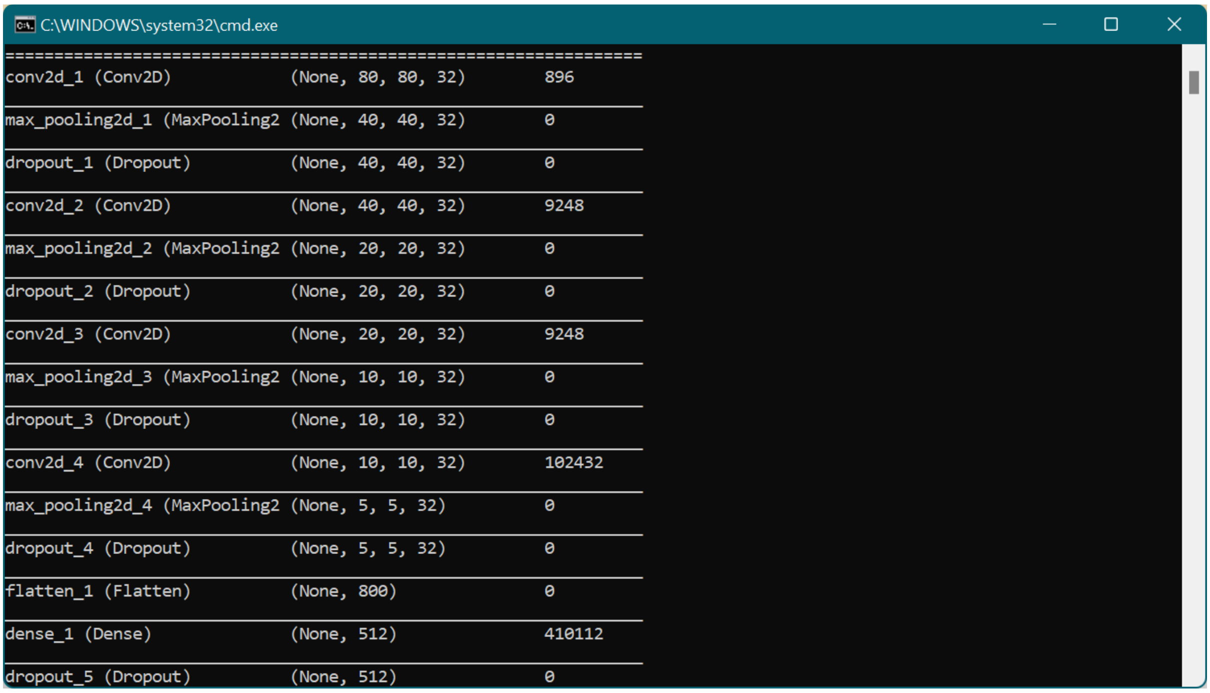


In above screen click on 'Generate post CNN Model' button to train CNN model

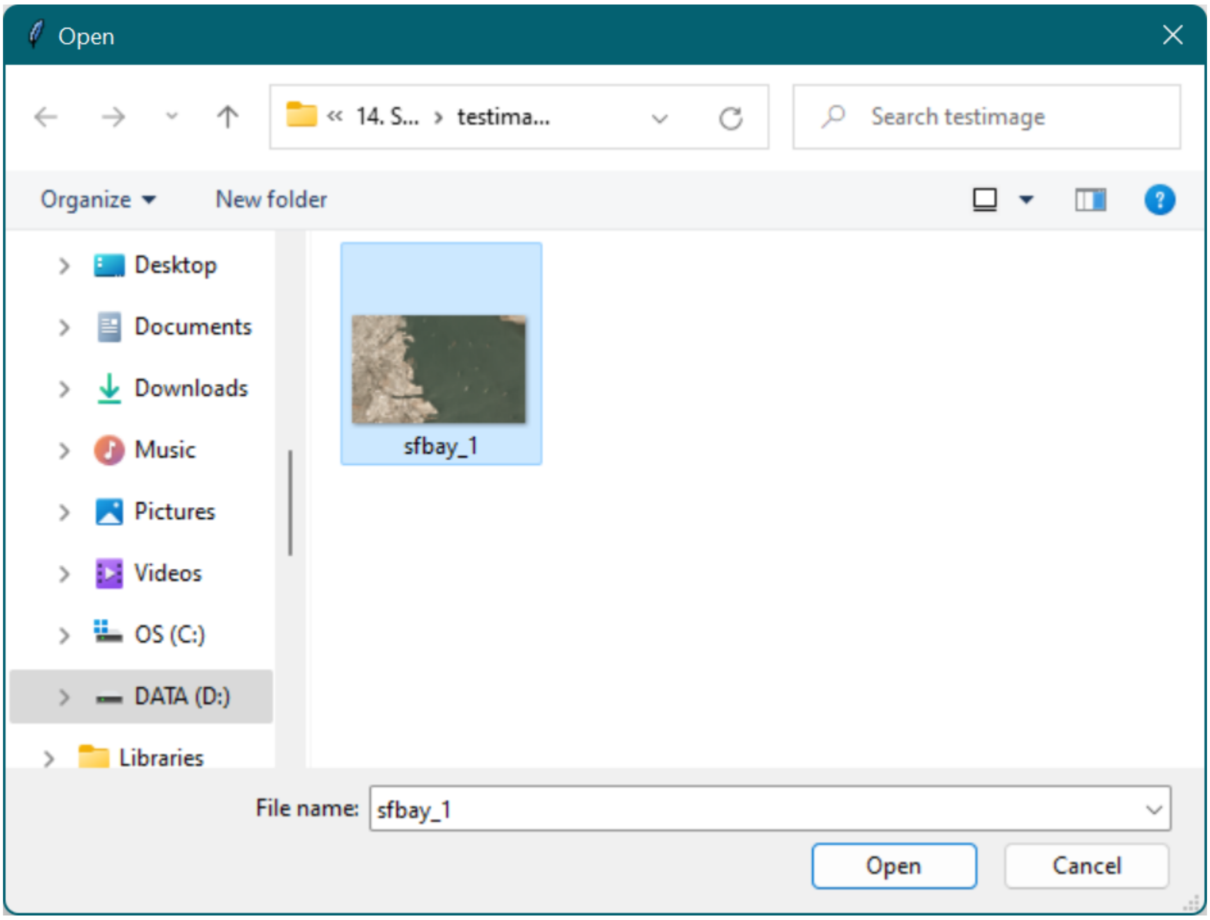
ISSN- 2394-5125 VOL 10, ISSUE 03, 2023



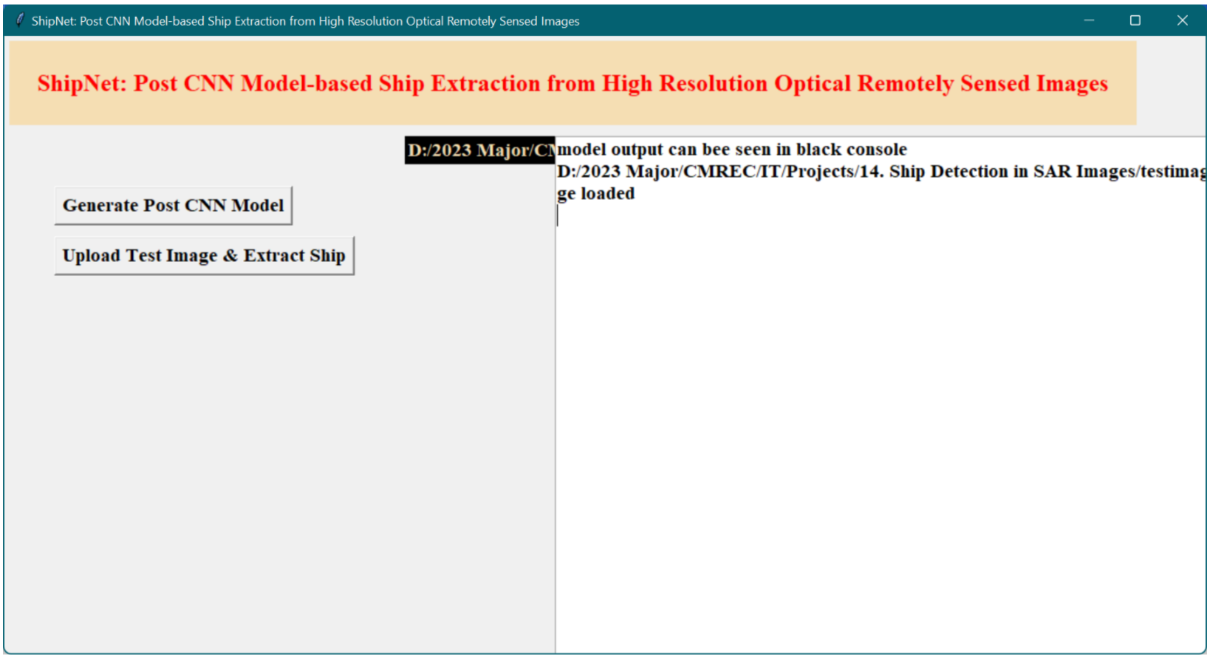
After generating model we will get above screen and if we want to see Model details then see black console screen below

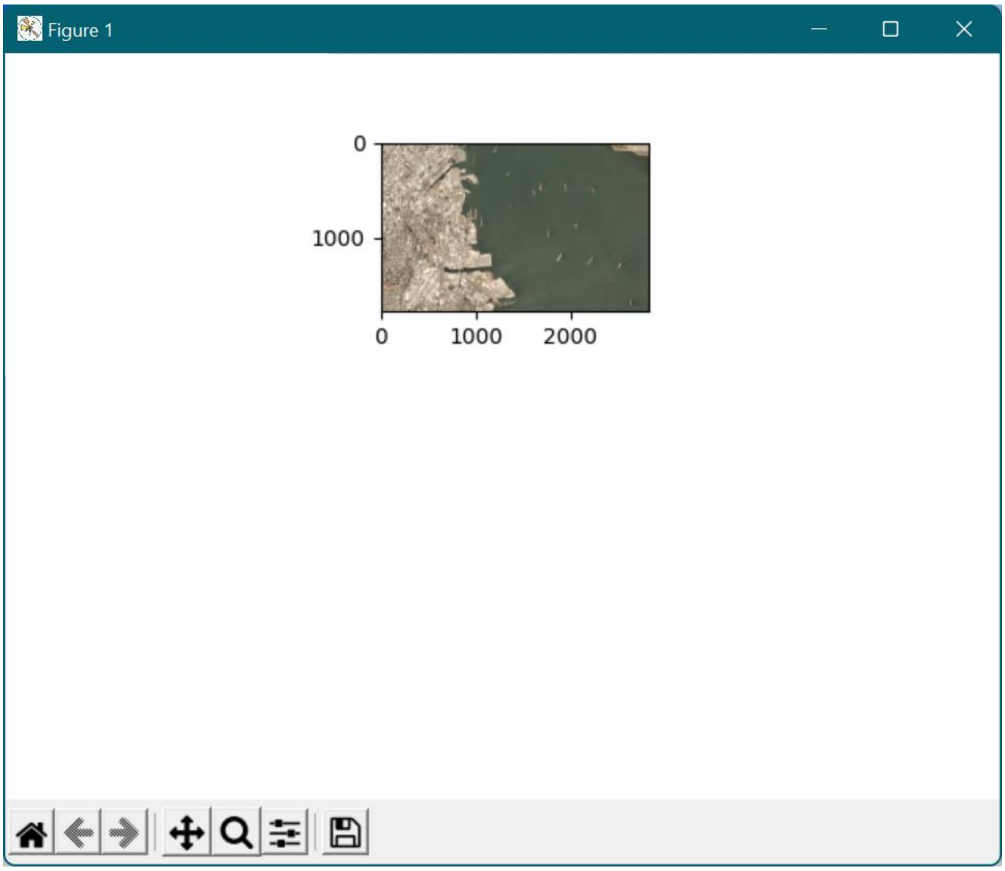


From above screen we can see total numbers of layers generated by proposed post CNN model from images to build train model and in bottom we can see from VGG ImageNet it has used how many numbers of train images. Now click on ‘Upload Test Image & Detect Ship’ button to upload test image and then application will detect ship from input image.



In above screen uploading one sea image which has some ships and now click on 'Open' button and wait for few seconds to get below screen



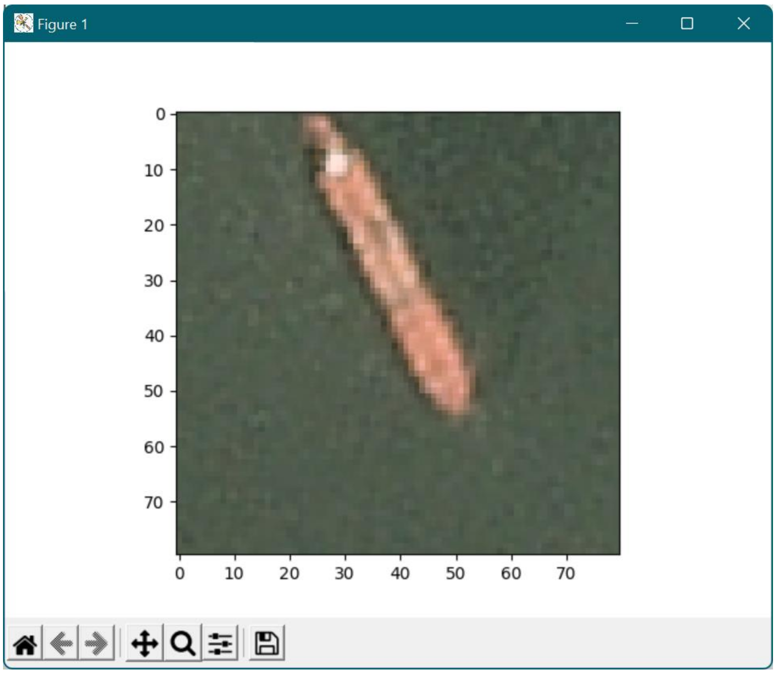


After getting above screen just close image screen and backside screen to allow image to extract features from uploaded image and then to detects ship.

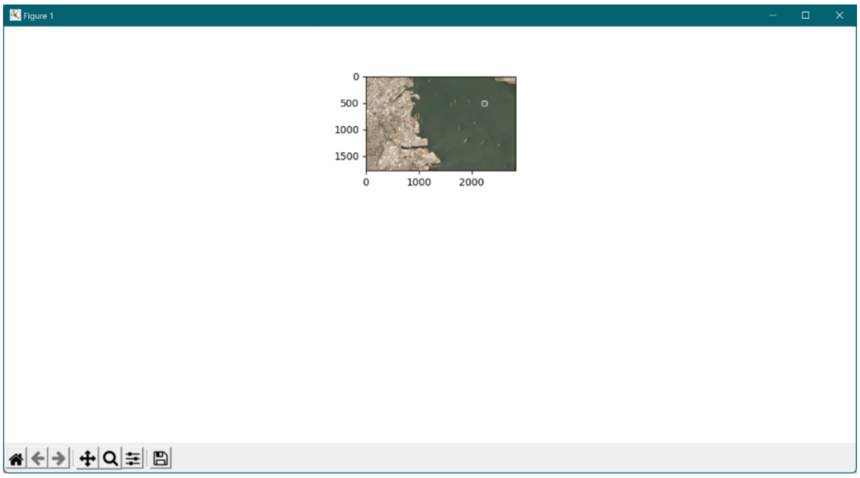
(Note: Please close above both screens only not black screens, once we close above both screens then in black console we can see features extraction process and once it detect ships then it will display detected ship)

```
C:\WINDOWS\system32\cmd.exe
_________________________________________________________________
dropout_2 (Dropout)          (None, 20, 20, 32)        0
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 20, 20, 32)        9248
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 10, 10, 32)        0
_________________________________________________________________
dropout_3 (Dropout)          (None, 10, 10, 32)        0
_________________________________________________________________
conv2d_4 (Conv2D)            (None, 10, 10, 32)        102432
_________________________________________________________________
max_pooling2d_4 (MaxPooling2 (None, 5, 5, 32)          0
_________________________________________________________________
dropout_4 (Dropout)          (None, 5, 5, 32)          0
_________________________________________________________________
flatten_1 (Flatten)          (None, 800)               0
_________________________________________________________________
dense_1 (Dense)              (None, 512)               410112
_________________________________________________________________
dropout_5 (Dropout)          (None, 512)               0
_________________________________________________________________
dense_2 (Dense)              (None, 2)                 1026
=================================================================
Total params: 532,962
Trainable params: 532,962
Non-trainable params: 0
_________________________________________________________________
57
92
X:1050 Y:60
```

In above black console in bottom X and Y features you can see and it will keep on processing till it detect ship, you just wait till you get below screen



In above screen we can see one ship is detected and then you close above ship screen to get another screen



In above screen we can see one ship is detected and one rectangle bounding box is surrounded around that ship.

5. CONLUSION AND FUTURE SCOPE

Ship extraction from high resolution optical remotely sensed images is a challenging task that requires advanced image processing and computer vision techniques. One approach to this problem is to use a post-CNN model, which involves a combination of convolutional neural network (CNN) layers and fully connected layers to extract features and classify ships in the image. The post-CNN model has shown promising results in ship extraction, achieving high accuracy and efficiency in detecting and classifying ships from large-scale satellite or aerial imagery. This approach leverages the power of deep learning and neural networks to automatically learn and extract relevant features from the image, and then use them to identify the presence and location of ships. Overall, ship extraction from high resolution optical remotely sensed images using a post-CNN model is a promising research area that has the potential to provide accurate and reliable information for various maritime applications. With further development and optimization, these systems can be a valuable tool for maritime surveillance, environmental monitoring, and navigation.

Future scope

The accuracy of the post-CNN model can be further improved by fine-tuning the model parameters and optimizing the training process. Additionally, incorporating other advanced techniques such as attention mechanisms, recurrent neural networks, and ensemble methods can further enhance the performance of the ship extraction system.

REFERENCES

[1] Nie, T.; He, B.; Bi, G.; Zhang, Y.; Wang, W. A Method of Ship Detection under Complex Background. ISPRS Int. J. Geo-Inf. 2017, 6, 159.

[2] Dong, C.; Liu, J.; Xu, F. Ship Detection in Optical Remote Sensing Images Based on Saliency and a Rotation-Invariant Descriptor. Remote Sens. 2018, 10, 400.

[3] Zhang S, Wu R, Xu K, Wang J, Sun W. R-CNN-Based Ship Detection from High Resolution Remote Sensing Imagery. Remote Sensing. 2019; 11(6):631. https://doi.org/10.3390/rs11060631.

[4] LI, Bo & XIE, Xiaoyang & WEI, Xingxing & TANG, Wenting. (2020). Ship detection and classification from optical remote sensing images: A survey. Chinese Journal of Aeronautics. 34. 10.1016/j.cja.2020.09.022.

[5] M. N. Hidalgo, A. -J. Gallego, P. Gil and A. Pertusa, "Two-Stage Convolutional Neural Network for Ship and Spill Detection Using SLAR Images," in IEEE Transactions on Geoscience and Remote Sensing, vol. 56, no. 9, pp. 5217-5230, Sept. 2018, doi: 10.1109/TGRS.2018.2812619.

[6] Wang, Y.-Q & Ma, L. & Tian, Y.. (2011). State-of-the-art of ship detection and recognition in optical remotely sensed imagery. Zidonghua Xuebao/Acta Automatica Sinica. 37. 1029-1039. 10.3724/SP.J.1004.2011.01029.

[7] Q. Li, L. Mou, Q. Liu, Y. Wang and X. X. Zhu, "HSF-Net: Multiscale Deep Feature Embedding for Ship Detection in Optical Remote Sensing Imagery," in IEEE Transactions on Geoscience and Remote Sensing, vol. 56, no. 12, pp. 7147-7161, Dec. 2018, doi: 10.1109/TGRS.2018.2848901.

[8] Y. Wang, C. Wang and H. Zhang, "Combining single shot multibox detector with transfer learning for ship detection using Sentinel-1 images," 2017 SAR in Big Data Era: Models, Methods and Applications (BIGSARDATA), 2017, pp. 1-4, doi: 10.1109/BIGSARDATA.2017.8124924.

[9] An Q, Pan Z, You H. Ship Detection in Gaofen-3 SAR Images Based on Sea Clutter Distribution Analysis and Deep Convolutional Neural Network. Sensors (Basel). 2018 Jan 24;18(2):334. doi: 10.3390/s18020334. PMID: 29364194; PMCID: PMC5855143.

[10] Zhao, J., Guo, W., Zhang, Z. et al. A coupled convolutional neural network for small and densely clustered ship detection in SAR images. Sci. China Inf. Sci. 62, 42301 (2019). https://doi.org/10.1007/s11432-017-9405-6.

[11] Hwang, JeongIn, Daeseong Kim, and Hyung-Sup Jung. "An efficient ship detection method for KOMPSAT-5 synthetic aperture radar imagery based on adaptive filtering approach." Korean Journal of Remote Sensing 33.1 (2017): 89-95.

[12] X. Li, P. Chen and K. Fan. Overview of Deep Convolutional Neural Network Approaches for Satellite Remote Sensing Ship Monitoring Technology, IOP Conf. Series: Materials Science

and Engineering 730 (2020) 012071, IOP Publishing, doi:10.1088/1757-899X/730/1/012071.