



e-ISSN: 2319-8753 | p-ISSN: 2320-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 3, March 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.512

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com



Inshore and Offshore Ship Detection on High-Resolution Remote Sensing Image Using Mask R-CNN

Lokesh V K, Rajesh K, Santhosh Raj P, Shiny R M

U.G Student, Department of Computer Science and Engineering, Velammal Institute of Technology, Chennai,
Tamil Nadu, India

U.G Student, Department of Computer Science and Engineering, Velammal Institute of Technology, Chennai,
Tamil Nadu, India

U.G Student, Department of Computer Science and Engineering, Velammal Institute of Technology, Chennai,
Tamil Nadu, India

Assistant Professor, Department of Computer Science and Engineering, Velammal Institute of Technology, Chennai,
Tamil Nadu, India

ABSTRACT: Deep convolutional neural network (DCNN) can achieve ship detection mission on the high-resolution remote sensing images. However, the false alarms caused by the onshore ship-like objects may decrease the accuracy and feasibility of these DCNN-based detection frameworks. In our work, an end-to-end method, named as Scene Mask R-CNN, is proposed to reduce the onshore false alarms. The scene mask extraction network (SMEN), as a network branch for scene segmentation, is innovatively introduced into the detection framework. The non-target area is marked out by an inferred scene mask which is used to assist the ship detection. Combining the feature map originated from feature extraction network (FEN) with the inferred scene mask by using the edge probability weighted (EPW) merging method, the false candidate targets in the non-target area are excluded. This novel mechanism of DCNN-based ship detection not only maintains the detection accuracy, but also effectively suppresses the false alarms in the non-target area. Finally, the validity and accuracy of this method are verified on a ship dataset generated by the high-resolution optical remote sensing images.

KEYWORDS: Deep convolutional neural network, False alarms, Scene mask, Ship detection,

I. INTRODUCTION

Deep learning is a class of machine learning algorithms that uses multiple layers to progressively extract higher-level features from the raw input. For example, in image processing, lower layers may identify edges, while higher layers may identify the concepts relevant to a human such as digits or letters or faces.

Deep learning (also known as deep structured learning) is part of a broader family of machine learning methods based on artificial neural networks with representation learning. Learning can be supervised, semi-supervised or unsupervised.

Deep-learning architectures such as deep neural networks, deep belief networks, recurrent neural networks and convolutional neural networks have been applied to fields including computer vision, machine vision, speech recognition, natural language processing, audio recognition, social network filtering, machine translation, bioinformatics, drug design, medical image analysis, material inspection and board game programs, where they have produced results comparable to and in some cases surpassing human expert performance.

Artificial neural networks (ANNs) were inspired by information processing and distributed communication nodes in biological systems. ANNs have various differences from biological brains. Specifically, neural networks tend to be static and symbolic, while the biological brain of most living organisms is dynamic (plastic) and analogue.

The adjective "deep" in deep learning refers to the use of multiple layers in the network. Early work showed that a linear perceptron cannot be a universal classifier, and then that a network with a non polynomial activation function with one hidden layer of unbounded width can on the other hand so be. Deep learning is a modern variation which is concerned with an unbounded number of layers of bounded size, which permits practical application and optimized implementation, while retaining theoretical universality under mild conditions. In deep learning the layers are also permitted to be heterogeneous and to deviate widely from biologically informed connectionist models, for the sake of efficiency, trainability and understandability, whence the "structured" part.

Most modern deep learning models are based on artificial neural networks, specifically convolutional neural networks (CNN), although they can also include propositional formulas or latent variables organized layer-wise in deep generative models such as the nodes in deep belief networks and deep Boltzmann machines.

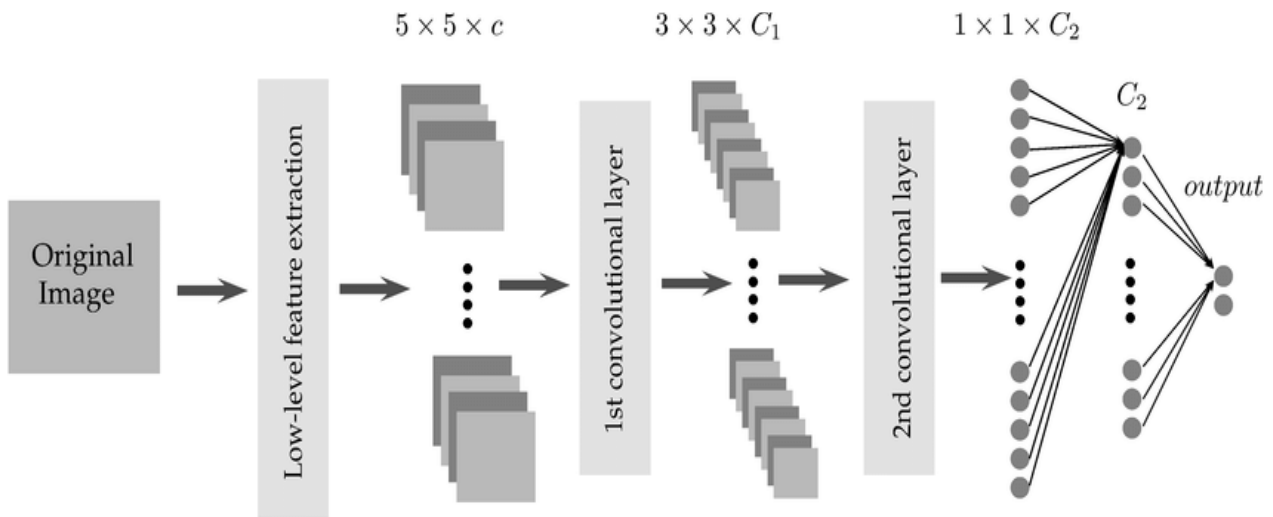


Fig 1.1 Architecture of Deep Learning

II. LITERATURE SURVEY

In year 2011, Matthew D. Zeiler, Graham W. Taylor and Rob Fergus proposed a Adaptive Deconvolutional Networks for Mid and High Level Feature Learning. In this study, a hierarchical model that learns image decompositions via alternating layers of convolutional sparse coding and max pooling. When trained on natural images, the layers of our model capture image information in a variety of forms: low-level edges, mid-level edge junctions, high-level object parts and complete objects. To build our model we rely on a novel inference scheme that ensures each layer reconstructs the input, rather than just the output of the layer directly beneath, as is common with existing hierarchical approaches. This makes it possible to learn multiple layers of representation and we show models with 4 layers, trained on images from the Caltech-101 and 256 datasets. Features extracted from these models, in combination with a standard classifier, outperform SIFT and representations from other feature learning approaches.

In year 2015, Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun proposed a Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. In this study, Existing deep convolutional neural networks (CNNs) require a fixed-size (e.g., 224×224) input image. This requirements "artificial" and may reduce the recognition accuracy for the images or sub-images of an arbitrary size/scale. In this work, we equip the networks with another pooling strategy, "spatial pyramid pooling", to eliminate the above requirement. The new network structure, called SPP-net, can generate a fixed-length representation regardless of image size/scale. Pyramid pooling is also robust to object deformations. With these advantages, SPP-net should in general improve all CNN-based image classification methods. On the Image Net 2012 dataset, we demonstrate that SPP-net boosts the accuracy of a variety of CNN architectures despite their different designs. On the Pascal VOC 2007 and Caltech101 datasets, SPP-net achieves state-of-the art classification results using a single full-image representation and no fine-tuning. The power of SPP-net is also significant in object detection. Using SPP-net, we compute the feature maps from the entire image only once, and then pool features in arbitrary regions (sub-images) to generate fixed-length representations for training the detectors. This method



avoids repeatedly computing the convolutional features. In processing test images, our method is 24-102 faster than the R-CNN method, while achieving better or comparable accuracy on Pascal VOC 2007. In ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2014, our methods rank #2 in object detection and #3 in image classification among all 38 teams. This manuscript also introduces the improvement made for this competition.

In year 2015, Karen Simonyan_ & Andrew Zisserman proposed a VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION. In this study, the effect of the convolutional network depth on its accuracy in the large-scale image recognition setting. Our main contribution is a thorough evaluation of networks of increasing depth using an architecture with very small (3x3) convolution filters, which shows that a significant improvement on the prior-art configurations can be achieved by pushing the depth to 16–19 weight layers. These findings were the basis of our ImageNet Challenge 2014 submission, where our team secured the first and the second places in the localization and classification tracks respectively. We also show that our representations generalise well to other datasets, where they achieve state-of-the-art results. We have made our two best-performing ConvNet models publicly available to facilitate further research on the use of deep visual representations in computer vision

III. PROPOSED METHODOLOGY AND DISCUSSION

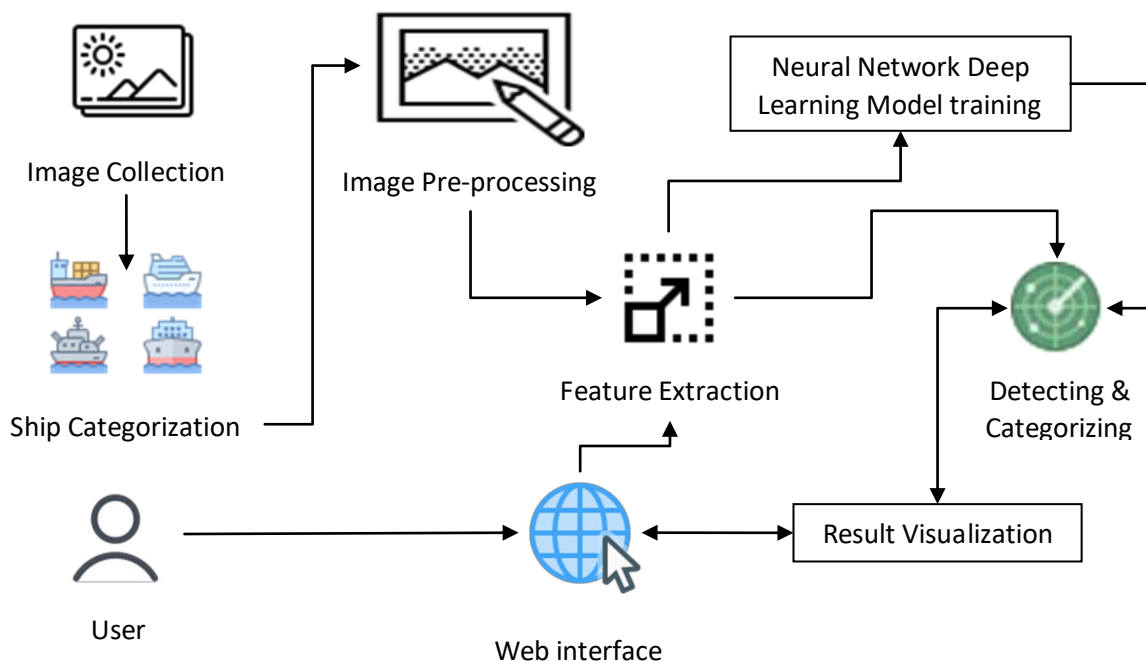
EXISTING SYSTEM

Computer-aided ship detection methods greatly free up human resources and typically include two steps: extracting image features, and then using classifiers for classification and localization. These methods can produce stable results under calm sea conditions. However, when disturbances such as waves, clouds, rain, fog, and reflections happen, the extracted low-level features are not robust. Besides, manual selection of features is time-consuming and strongly dependent on the expertise and characteristics of the data itself.

PROPOSED SYSTEM

A system is proposed to automate the detection of presence of ships in the given image using Machine Learning and Deep Learning Algorithms. We are proposing along with ship detection, a ship classification based on the type and category of the ships. The proposed system will not only detect a ship but also categorize as war ship, container ship etc.

Fig 2: Architecture of the proposed system





Data collection, categorizing and pre-processing

A raw dataset was collected with type of ships and images for each type of ships. Since we did not have more number of ships images classified based on the categories, we use python to web crawl images from internet. The web scraped images were stored in different folders. For this project we were using four categories. We have downloaded ship images for the following categories, Accommodation, Container, War and Cruise. Python with vast text processing and networking libraries is the best tool to write one-off web scrapper.

Deep Learning using CNN

Deep Learning is a subfield of machine learning concerned with algorithms inspired by the structure and function of the brain called artificial neural networks. Deep learning is a class of machine learning algorithms that use a cascade of multiple layers of nonlinear processing units for feature extraction and transformation. Each successive layer uses the output from the previous layer as input. We use convolutional Neural Network in our project for training and detecting the ships in the given image. We use multilayer neural network and each layer output is given as input to the next layer.

Detecting and categorizing Ships in images

The CNN trained model is used to detect the availability and the category of the ship in the given image. The output of the training module is saved as an object file using python. When the user inputs the images to for detection, the prediction method in the trained model is used to detect the presence of the ship, and also the category of the ship. A new image file is created with the category written on top of the actual image.

IV. PROPOSED SYSTEM ALGORITHMS

Regional CNN

In this chapter we will talk about the Region based methods for object detection. In particular, R-CNN (Regional CNN), the original application of CNNs to this problem, along with its descendants Fast R-CNN, and Faster R-CNN .In classification, there's generally an image with a single object as the focus and the task is to say what that image is. But when we look at the world around us, we carry out far more complex tasks. We see complicated sights with multiple overlapping objects, and different backgrounds and we not only classify these different objects but also identify their boundaries, differences, and relations to one another. The goal of a Region based CNN is to take an image input to detected and localize the object in the image. All three above mentioned architectures were used for object detection and localization for a driving scenario to detect cars, pedestrians and traffic signs etc. and performed a time study to conclude which of the three architectures can be used as a detector for a real-world driving scenario for detection of objects present on road. We will also discuss their architecture and describe how they perform the task of object detection and localization given the input image.

R-CNN Architecture

In R-CNN we input an image and then using the basic feature detection techniques such as edge detection etc. we get Region Proposals Also referred as Regions of interest. This process is also known as selective search.

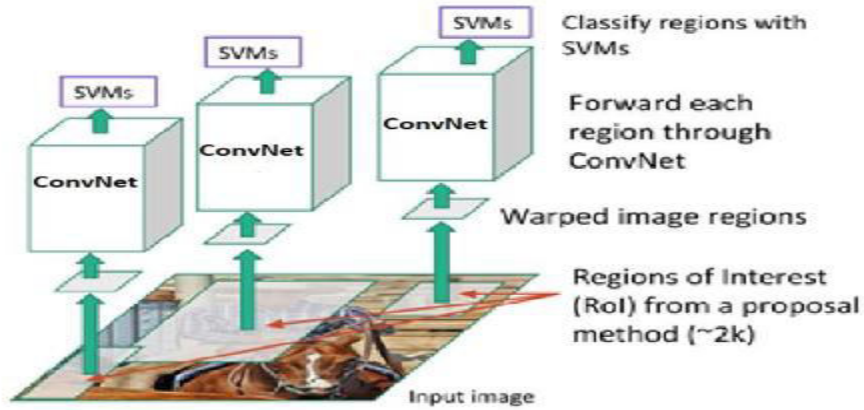


Fig 3: R-CNN Architecture

In the next step after these region proposals are extracted these wrapped image regions go through a trained CNN in our implementation AlexNet and then on the final layer the classification is done Using a Support Vector Machine(SVM) classifier which classifies whether this is an object and if so what. Pictorial representation of above process is shown in Figure.

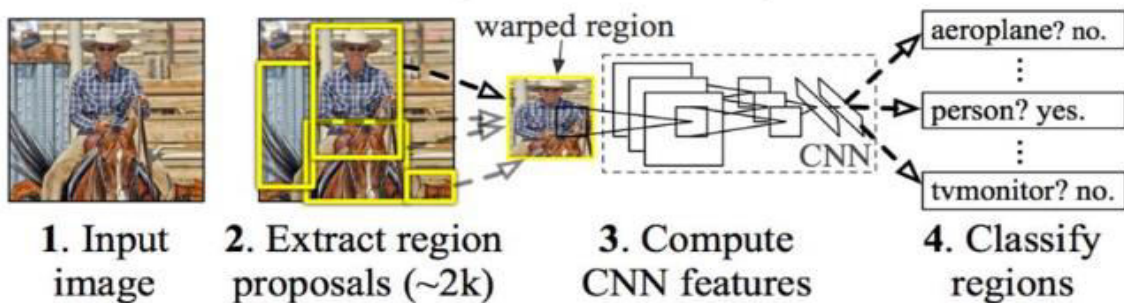


Fig 4: R-CNN Layers

IV. EXPERIMENTAL RESULTS:

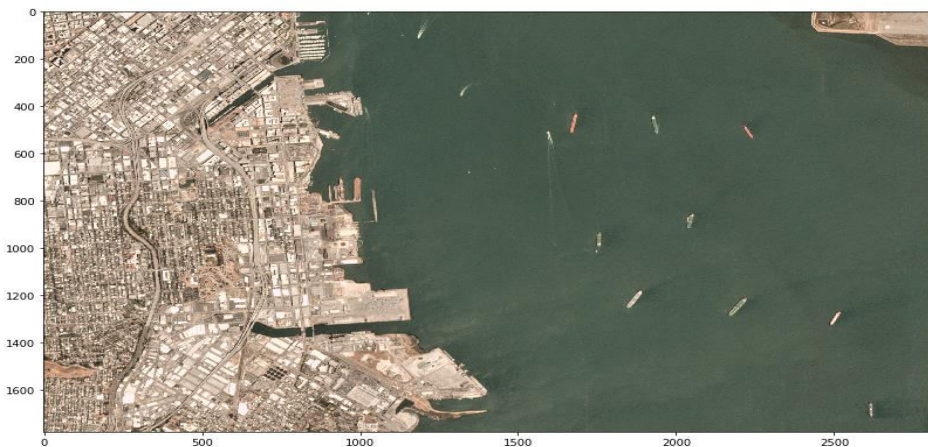


Fig 5: Image Visualization

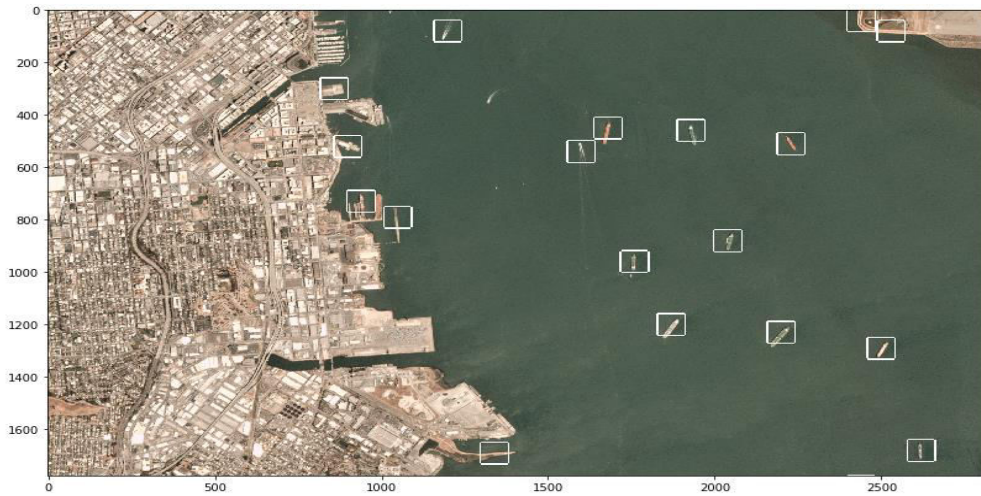


Fig 6: Detection of Object



Fig 7: Ship Detection Webpage



Fig 8: Prediction Page



V. CONCLUSION AND FUTURE ENHANCEMENT:

In this project, we analyze the limitations of dense distribution and different sizes. Consequently, we propose R-Libra R-CNN, a rotational Libra R-CNN for ship detection. BFP module and IoU-BS module are introduced with the motivation of extracting discriminative features to overcome the impact of different sizes. Furthermore, the rotational region detection with balanced L1 loss is finally proposed to be robust against the impact of dense distribution. Our method can gain consistent improvements in accuracy and visualization. In addition, extensive experiments on the DOTA show that the proposed method can gain 3.43% than R2CNN and 4.09% than Libra R-CNN.

In future work, we will explore more efficient strategies for extremely small objects, e.g., super resolution reconstruction. In addition, we will focus on learning R-Libra R-CNN model by introducing a sparse distribution estimation scheme to improve the computational efficiency.

REFERENCES

- [1] H. Lang, Y. Xi, and X. Zhang, "Ship detection in high-resolution SAR images by clustering spatially enhanced pixel descriptor," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5407–5423, Aug. 2019.
- [2] P. Salembier, S. Liesegang, and C. Lopez-Martinez, "Ship detection in SAR images based on maxtree representation and graph signal processing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 5, pp. 2709–2724, May 2019.
- [3] T. Zhang, J. Ji, X. Li, W. Yu, and H. Xiong, "Ship detection from PolSAR imagery using the complete polarimetric covariance difference matrix," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 5, pp. 2824–2839, May 2019.
- [4] X. Leng, K. Ji, S. Zhou, and X. Xing, "Ship detection based on complex signal kurtosis in single-channel SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6447–6461, Sep. 2019.
- [5] 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," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 12, pp. 7147–7161, Dec. 2018.
- [6] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006.
- [7] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.
- [8] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, and S. Reed, "SSD: Single shot multibox detector," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, in Lecture Notes in Computer Science, vol. 9905. Amsterdam, The Netherlands: Springer, Oct. 2016, pp. 21–37.
- [9] Y.-L. Chang, A. Anagaw, L. Chang, Y. C. Wang, C.-Y. Hsiao, and W.-H. Lee, "Ship detection based on YOLOv2 for SAR imagery," *Remote Sens.*, vol. 11, no. 7, p. 789, Apr. 2019.
- [10] Z. Lin, K. Ji, X. Leng, and G. Kuang, "Squeeze and excitation rank faster R-CNN for ship detection in SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 5, pp. 751–755, May 2019.



INNO  **SPACE**
SJIF Scientific Journal Impact Factor

**Impact Factor:
7.512**

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details