International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 5 Issue 1, November-December 2020 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

# Research on Ship Detection in Visible Remote Sensing Images

**Peng Wei, Qian Li, Kaiwen Yang**

Beijing Wuzi University, Beijing, China

**ABSTRACT**

With the continuous development of remote sensing technology and satellite communication technology, the higher is the resolution of remote sensing images, and the richer is information contained in the image. How to obtain information from remote sensing images quickly and accurately is the key problem in the application of remote sensing technology. This paper introduces the automatic detection of ship targets in visible light remote sensing images, analyzes the methods and problems of target detection and semantic segmentation based on deep learning, which improves the accuracy of ship detection. Finally, this paper summarizes the short comings of the existing methods and prospects of the future research trend.

***KEYWORDS:*** *remote sensing images, computer vision, deep learning*

***How to cite this paper:*** Peng Wei | Qian Li | Kaiwen Yang "Research on Ship Detection in Visible Remote Sensing Images" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-5 | Issue-1, December 2020, pp.55-59, URL: www.ijtsrd.com/papers/ijtsrd35823.pdf

**IJTSRD35823**



## I. INTRODUCTION

Remote sensing images [1] are captured the real image of the earth's surface by remote sensing satellites. Through remote sensing images, people can intuitively understand the ground buildings, road conditions, natural environment, and other information. According to the types of sensors, remote sensing images can be divided into synthetic aperture radar (SAR) image, laser radar image, thermal infrared image, optical sensor (including visible light sensor) image, etc. Optical remote sensing mainly refers to the sensors operating in the visible and, that is, 0.38-0.76 micron. Visible images are intuitive and easy to understand with high spatial resolution. When the light and sunny weather conditions are well, high-resolution visible images can provide lots of information, obvious target structure characteristics. In the aspect of remote sensing specific target detection, the visible image has incomparable advantages such as SAR image and multispectral image [2]. Therefore, visible remote sensing images have been paid more and more attention when researchers extract the content of interest, the detection and recognition of specific targets from remote sensing images.

In recent years, with the continuous development of remote sensing technology and satellite communication technology, the higher is the resolution of remote sensing images and the richer is information contained in the image. How to obtain information from remote sensing images quickly and accurately is a key problem in the application of remote sensing technology. The target detection of remote sensing image mainly aims at the given remote sensing image, and automatically identify the interesting objects in the image, such as vehicle, aircraft, ship, port, airport, etc. As an important part of remote sensing, ship detection plays an important role in the military and civil fields such as maritime security, maritime traffic, border control, fishery management, etc. China is rich in marine resources. Ship detection based on remote sensing images is widely used in military and civil fields. For example, in a specific sea area or port, remote sensing technology can realize the monitoring and management of illegal fishing and smuggling ships. Ships may be in distress and loss of contact occur from time to time in bad weather conditions, while remote sensing technology can quickly and accurately detect the location of ships in distress, which is conducive to the rescue work[3]. In recent years, with the rapid development of optical remote sensing imaging technology, "GaoJing-1", "worldview-4", "GaoFen-6" and other optical remote sensing satellites have been launched one after another, which can continuously provide massive high-resolution images and bring new opportunities and challenges to the development of ship detection technology. Therefore, ship detection of remote sensing images has important research significance and broad application prospects.

The traditional target detection algorithm is mainly based on image processing to extract features manually. This method is not accurate when processing complex images. The vigorous development of deep learning technology provides a very effective tool for remote sensing image target detection. Although the target detection technology based on deep learning has achieved remarkable success in natural images, it is difficult to directly apply this method to remote sensing images because of the significant difference between remote sensing images and natural scene images. Compared with the traditional image, the remote sensing image has the characteristics of a large amount of data, complex background, small target size, and so on. At the same time, because the remote sensing image is a top view, different projection directions make the same target has different rotation angles, which makes the angle direction of the target arbitrary. Due to the particularity of remote sensing image, the current mainstream general target detection method based on deep learning cannot play an optimal role in remote sensing image target detection. Therefore, it is of great significance to study how to apply deep learning methods to target detection in satellite visible light remote sensing images to effectively improve the accuracy of target detection and reduce the detection time.

## II. Related work on target detection

Target detection [4] is a basic research task in computer vision, which is mainly used to identify the types and locations of targets in images. Target detection technology is widely used in industrial product detection, airport security inspection, robot navigation, video monitoring, target tracking, remote sensing, and many other fields. Through using computer technology to replace the traditional manual detection, it is an important practical significance for improving production efficiency and reducing production cost. Therefore, target detection has become a hot topic in the academic and industrial circles this year. At the same time, as an important branch of computer vision, target detection technology also plays an important role in the following target tracking, semantic segmentation [5,6], instance segmentation [7].

The traditional target detection algorithm is mainly based on SIFT [8] and HOG [9] to extract features manually. In 2012, Alexnet [10] made a breakthrough in the ILSVRCimage classification competition, and researchers began to use the deep convolution neural network method to solve computer vision problems. At present, the common network structures include Google net [11], VGG [12], Darknet [13], ResNet [14], DenseNet [15], etc., which are designed to improve the accuracy and speed of feature extraction from different aspects. At present, the general target detection methods based on convolutional neural networks can be divided into two categories: the one-stage detection method and the two-stage detection method. In the two-stage detection method, candidate regions are generated from the image firstly, and then the candidate regions are classified. The main methods are R-CNN, Fast R-CNN, FasterR-CNN [16].One-stage detection methods mainly include YOLO, YOLO9000, YOLOv3[17], SSD[18], etc.

## III. Ship detection methods

With the development of remote sensing technology, people can understand the earth from a better perspective. In recent years, with the improvement of remote sensing image resolution, ship detection in optical remote sensing images is a research hotspot, and many different methods have been proposed. Generally speaking, these methods can be divided into two categories: traditional methods and deep learning methods. The traditional ship detection methods are divided into three steps: the first step is the separation of land and sea, which aims at separating the target from the ocean, the second step is candidate detection, which shows that pixels represent possible ships, and the last step is classification, which identifies one of all the detected candidates in the previous step of the real target. Traditional methods are mainly based on geometric features of images.

Yang et al [19] proposed a detection method based on sea surface analysis and used a newlinear function combining pixel and region features to select candidate ships. Shi et al [20] extended the existing HOG features and some additional information to improve the ability of ship identification. Qi et al [21] used visual saliency to extract candidate regions and improved the hand feature based on HOG. Traditional detection methods mainly use manual features to detect ships. These methods have achieved good results in open waters, but it is difficult to detect in complex backgrounds. With the development of deep learning, the methods based on convolutional neural networks have surpassed the traditional methods in many computer vision tasks. Inspired by the successful application of convolutional neural networks in natural scenes, many detection methods based on deep learning have been applied to ship detection in remote sensing images. To improve the detection accuracy, Huang et al [22] proposed a squeeze excitation hop connection path network (SESPNets) using a path-level hop connection structure to improve the feature extraction ability. Zhang et al [23] improved the original Fast R-CNN structure to improve the detection results of small-scale distributed ships. You et al [24] proposed a DCNN framework to deal with multi-scale ship detection. You et al. [25] proposed an end-to-end method called scene mask R-CNN to reduce onshore false-positives. These methods based on CNN improve the efficiency and detection accuracy and liberate researchers from the tedious feature processing. However, these methods can only generate horizontal boundary boxes, which are not suitable for targets placed in any direction in remote sensing images.

Due to the great differences between remote sensing images and natural scenery images, it is still difficult to apply it directly to remote sensing images. Different from natural images, remote sensing images are taken with an aerial view. Targets may exist in any direction. When detecting density targets, especially with a high length-width ratio, ordinary targets often are missed in the NMS step by relying on the horizontal bounding box method. In contrast, the rotating bounding box with an arbitrary rotation angle is more suitable for ship detection. To solve this problem, researchers began to introduce a rotating bounding box, which was initially used for text detection. Yang et al. [26] proposed a framework called rotation dense feature pyramid network (R-DFPN), which can effectively detect ships in different scenes such as ocean and port. Ma et al. [27] proposed a two-stage CNN ship detection method based on ship center and azimuth prediction, which can accurately detect ships in any direction in optical remote sensing images. Tian et al. [28] proposed a framework integrating multi-scale feature fusion network, rotating region recommendation network, and context pool, which realized accurate positioning and false alarm suppression. Zhang et al. [29] proposed a rotating area recommendation network (R2PN), which uses the azimuth information of ships to generate multi-directional suggestions. At present, most of the methods are two-stage. These methods can predict the direction of the target and improve detection accuracy. However, due to the existence of RPN network, these methods use complex network structures and affect the detection velocity.

## IV. Semantic segmentation methods

The semantic segmentation combines image classification, target detection, and image segmentation. Image semantic segmentation is to classify every pixel in the image. By semantic segmentation of remote sensing images, we can separate the specific semantic pixels and obtain boundary information to improve the accuracy of target detection. Image semantic segmentation methods include traditional methods and methods based on convolution neural networks. The traditional semantic segmentation methods mainly include threshold-based segmentation, edge-based segmentation, region-based segmentation, and so on.

### A. Image segmentation methods based on threshold

The threshold-based segmentation is to calculate one or more gray thresholds based on the gray characteristics of the image, compare the gray value of each pixel with the threshold, and finally classify the target into appropriate categories according to the pixel comparison results.

The advantages of threshold-based segmentation are simple, high efficiency, and speed. The global threshold can effectively segment different targets and backgrounds with a great difference in the gray level. Local threshold or dynamic threshold is more suitable for the targets with little difference. Although the threshold-based segmentation method is simple and efficient, it also has some limitations. The method only considers the gray value of the pixel itself, generally does not consider the spatial characteristics, so it is very sensitive to noise. In practical application, the threshold-based methods are usually combined with other methods.

### B. Image segmentation methods based on edge

The edge refers to the set of continuous pixels on the boundary of two different regions in the image, which reflects the discontinuity of local features and the mutation of image characteristics such as gray, color, texture, etc. The edge-based segmentation method detects the edge according to the gray value and divides the image into different parts. It is based on the observation that the gray value of the edge will show a step change.

### C. Image segmentation technology based on Region

The image is divided into different regions according to the similarity criterion. It mainly uses the local space information of the image and can avoid the defect of small segmentation space brought by other algorithms. However, this kind of segmentation method is slow in large area segmentation and has poor anti-noise performance, which often results in meaningless region segmentation or over-segmentation of image. In general, it will be combined with other methods to give full play to their advantages to obtain a better segmentation effect.

With the rapid development of deep learning, it is found that image semantic segmentation based on deep learning can greatly improve accuracy. Most of the traditional segmentation methods consider the visual information of the image pixel itself without training. Although the time complexity is not high, the accuracy is low, and it cannot effectively process the scene with complex backgrounds. The primary difference between the semantic segmentation method based on convolutional neural networks and the traditional semantic segmentation method is that the network can automatically learn the features of the image and carry out end-to-end classification learning, which greatly improves the accuracy of semantic segmentation. The main idea of the image semantic segmentation method based on deep learning is that a large number of original image data are directly inputted into the deep network without artificial design features. According to the designed deep network algorithm, the image data is processed complexly to obtain high-level abstract features. The output is no longer a simple classification category or target positioning but the segmentation image with pixel category label.

Long et al. [30] proposed a semantic segmentation model of a full convolution network. Firstly, the features were extracted by convolution operation, and then the feature map was up sampled and restored to the input size. The result of up sampling is fuzzy. It is not sensitive to the details of the image, and the segmentation accuracy is poor. Vijay et al. [31] proposed the SegNet symmetric semantic segmentation model, which can reduce the loss of pixel information, improve the resolution, and accurately locate the image segmentation boundary. Chen et al. [32] proposed the Deeplab model and added CRF based on FCN to improve the boundary segmentation accuracy. Ronneberger et al. [33] proposed a U-net network to improve the detection accuracy by adding feature fusion. When the convolution network is used for semantic segmentation, the input image usually needs to be adjusted to the input size of the model for the fixed-size model. After semantic segmentation, it is restored to the original size. Because the modification of image size will lose useful information, especially for the small target, it increases the difficulty of target segmentation.

## V. Conclusion

Due to the great difference between remote sensing image and natural scenery image, it is still difficult to apply it directly to the remote sensing image. Different from natural images, remote sensing images are taken from an aerial view. Targets in remote sensing images may exist in any direction. When detecting densely packed objects, especially targets with a large aspect ratio, ordinary targets oftenare missed targets in the NMS step by relying on the horizontal bounding box method, as shown in Figure 1. The method retains the boundary box A with the maximum probability, removes other boundary boxes B and C, which results in the omission of the real target.

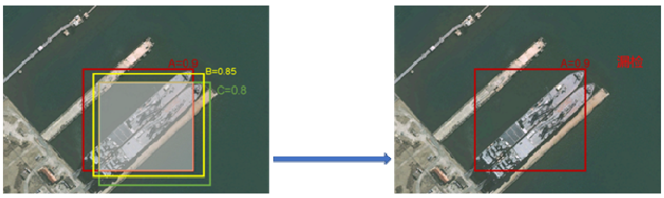


**Figure 1falsedetection caused by NMS**

When the full convolution network is used for semantic segmentation, the input image usually needs to be adjusted to the input size of the model for a fixed size model. After semantic segmentation, it is restored to the original size. The process of semantic segmentation is shown in Figure 2.

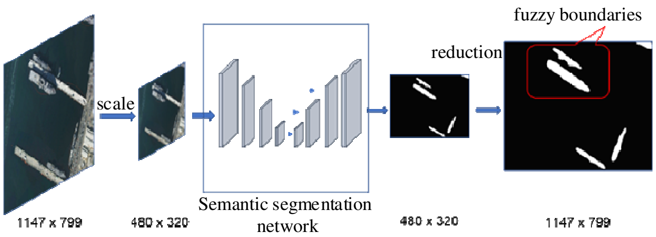


**Figure 2 semantic segmentation flow based on convolutional neural network**

For example, the above network limits the input image to 480 * 320. Because the modification of image resolution will lose useful information in the image, especially for small targets. The size in the feature map is very small after down sampling, and it is difficult to include all the information about the target, which makes it more difficult to segment the target. For targets with complex boundaries, the detection results are restored to the original image. And the boundary is not smooth enough, and it is not sensitive to the details of the image.

At present, the methods based on deep learning in remote sensing target detection achieve some achievements, which are mainly from the improvement of the target detection algorithm in natural scenes. There are significant differences between remote sensing images and natural scene images, especially in the aspects of target rotation, scale change, and complex background. Although researchers have explored and studied the ship target detection and recognition technology in optical remote sensing image, there are still many problems and challenges in ship target detection algorithm based on remote sensing image

### 1. The environment in optical remote sensing image is complex

Due to the transmission characteristics of optical sensors, camera angles, light intensity, weather, and sea background, and other factors, the target will be blocked, or incomplete and the contrast is low, which makes it more difficult to detect the target.

### 2. Detection of inshore ships

Different from the ship on the sea, because the inshore ship is often connected with the coast, which has similar grayscale and texture of the buildings on the shore and is affected by the shadow and side-by-side berthing. Thus, the conventional ship detection method cannot achieve the ideal detection results for the inshore ship. It is also difficult to extract the ship target quickly and accurately from the complex port environment.

### 3. Real-time performance of ship detection algorithm

At present, some target detection algorithms have a high recognition rate. But because of its high complexity and slow detection speed, it cannot meet the real-time requirements. With the development of deep learning technology, remote sensing image target detection technology is still an open problem to be further studied.

**Reference:**

[1] Yao L. Q., Hu X., Lei H..Object Detection in Remote Sensing Images Using Multiscale Convolutional Neural Networks. *Acta Optica Sinica*, 2019. 39(11): 346-353.

[2] Kanjir U., H. Greidanus, K. Ostir. Vessel detection and classification from spaceborne optical images: A literature survey. *Remote Sens. Environ.*, 2018. 207: 1-26.

[3] Yu D. X., et al., Joint salient feature and convolutional neural network for ship detection in remote sensing images. *Journal of Image and Graphics*, 2018. 23(12): 1947-1958.

[4] Fang L. P., He H. J., ZhouG. M..Research overview of object detection methods. *Computer Engineering and Applications*, 2018. 54(13): p. 11-18+33.

[5] Liang X. Y., Quan J. CH., Xiao K. H., Gao W. J..Research on Progress of Image Semantic Segmentation Based on Deep Learning. *Computer Engineering and Applications*, 2020. 56(02): p. 18-28.

[6] Tian X., Wang L., Ding Q..Review of Image Semantic Segmentation Based on Deep Learning. *Journal of Software*, 2019. 30(2): p. 440-468.

[7] He, K., et al. Mask R-CNN. *in Proceedings of the IEEE international conference on computer vision*. 2017, 42: 386-397.

[8] Lowe, D. G.. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 2004. 60(2): 91-110.

[9] Dalal N., B. Triggs. Histograms of oriented gradients for human detection. *in Proceedings of the IEEE conference on computer vision and pattern recognition*, 2005, 1: 886-893.

[10] [10]Krizhevsky A., I. Sutskever, G. Hinton. ImageNet Classification with Deep Convolutional Neural Networks. *Advances in neural information processing systems*, 2012, 25(2): 1097–1105.

[11] Szegedy, C., et al. Going deeper with convolutions. *in Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, 1:1-9.

[12] Simonyan, K. A.J.a.p.a. Zisserman. Very deep convolutional networks for large-scale image recognition. *in Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016.

[13] Redmon J., A. Farhadi. YOLO9000: better, faster, stronger. *in Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, 1: 6517-6525.

[14] He K., et al. Deep residual learning for image recognition. *in Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, 1: 770-778.

[15] Huang, G., et al. Densely connected convolutional networks. *in Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, 1: 2261-2269.

[16] Ren S., et al. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2017, 39(6): 1137 - 1149.

[17] Redmon, J. A. Farhadi. Yolov3: An incremental improvement. 2018.

[18] Liu, W., et al.. SSD: Single shot multibox detector. *in European conference on computer vision*, 2016, 26-27.

[19] Yang G., et al.. Ship detection from optical satellite images based on sea surface analysis. IEEE Geoscience and Remote Sensing Letters, 2013. 11(3): 641-645.

[20] Shi, Z., et al.. Ship detection in high-resolution optical imagery based on anomaly detector and local shape feature. *IEEE Transactions on Geoscience and Remote Sensing*, 2013. 52(8): 4511-4523.

[21] Qi S., et al.. Unsupervised ship detection based on saliency and S-HOG descriptor from optical satellite images. *IEEE Geoscience and Remote Sensing Letters*, 2015. 12(7): 1451-1455.

[22] Huang G., et al., Ship detection based on squeeze excitation skip-connection path networks for optical remote sensing images. *Neurocomputing*, 2019. 332: 215-223.

[23] Zhang S., et al.. R-CNN-Based Ship Detection from High Resolution Remote Sensing Imagery. Remote Sensing, 2019. 11(6):631.

[24] You Y., et al.. Broad Area Target Search System for Ship Detection via Deep Convolutional Neural Network. *Remote Sensing*, 2019. 11(17):1965.

[25] YouY., et al.. Nearshore Ship Detection on High-Resolution Remote Sensing Image via Scene-Mask R-CNN. *IEEE Access*, 2019. 7: 128431-128444.

[26] Yang X., 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*, 2018. 10(1): 132.

[27] Ma J., et al.. Ship Detection in Optical Satellite Images via Directional Bounding Boxes Based on Ship Center and Orientation Prediction. *Remote Sensing*, 2019. 11(18): 2173-2173.

[28] Tian T., et al.. Arbitrary-Oriented Inshore Ship Detection based on Multi-Scale Feature Fusion and Contextual Pooling on Rotation Region Proposals. *Remote Sensing*, 2020, 12(2): 339-358.

[29] Zhang Z., et al.. Toward arbitrary-oriented ship detection with rotated region proposal and discrimination networks. *IEEE Geoscience Remote Sensing Letters*, 2018. 15(11): 1745-1749.

[30] Long J., E. Shelhamer, T. Darrell. Fully convolutional networks for semantic segmentation. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2015, 79, 1337-1342.

[31] Badrinarayanan V., A. Kendall, R. Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 2017. 39(12): 2481-2495.

[32] Chen L.-C., et al.. Semantic image segmentation with deep convolutional nets and fully connected CRFs.*Computer Science*, 2014(4): 357-370.

[33] Ronneberger O., P. Fischer, T. Brox. U-net: Convolutional networks for biomedical image segmentation. *Lecture Notes Comput Sci*, 2015, 9351:234–241.