Ship Targets Detection in Remote Sensing Images Based on Improved Faster-RCNN

Dawei Liu*, Shujing Gao
Information Engineering College, Engineering University of PAP, Xi'an, Shaanxi, 710086, China
*Corresponding author’s e-mail: wjmicheal@163.com

Abstract. An improved algorithm is proposed to solve the problems of inaccurate recognition and low recall of Faster-Regions with Convolutional Neural Network (Faster-RCNN) algorithm for the detection of ship targets in remote sensing images. The algorithm is based on the Faster-RCNN network framework. Aiming at the small size and dense distribution of ship targets in remote sensing images, the feature extraction network is improved to enhance the detection ability of small targets. ResNet50 is used as the basic feature extraction network of the algorithm, and the hole residual block is introduced for multi-layer feature fusion to construct a new feature extraction network, which improves the feature extraction capability of the algorithm. The experimental results show that compared with the Faster-RCNN algorithm, this algorithm can learn more abundant target features in smaller pixel areas, thereby effectively improving the detection accuracy of ship targets.

1. Introduction
Remote sensing image targets detection is an important content of remote sensing image interpretation. It refers to automatically extracting the key targets of interest for a specific interpretation purpose, measuring and analyzing their attribute information, so as to provide evidence for further interpretation. As important military targets and maritime transportation carriers, ship targets detection in remote sensing image plays an important role in military and civil fields, such as fishery management, vessel traffic services, illegal oil spills, naval warfare, and maritime activities. Ship targets detection based on remote sensing image has important research significance and application value. The traditional ship target detection methods in remote sensing images include template matching[1], knowledge guidance[2-3] and supervised classification[4]. However, due to the difficulty in extracting the deep features of the image and the complexity of the remote sensing image itself, the traditional detection algorithms have poor robustness and low detection efficiency, which is difficult to meet the requirements of large-scale automation.

In recent years, the targets detection methods based on deep learning have made a breakthrough in the field of computer vision. Their excellent performance has brought new opportunities to the task of ship targets detection in remote sensing images. Among many targets detection algorithms, Faster-RCNN proposed by Ren et al.[5] uses region proposal network (RPN) to replace the selective search algorithm, reduces the number of candidate boxes, realizes end-to-end detection, and achieves high detection accuracy while ensuring detection efficiency.

Although the Faster-RCNN algorithm has achieved good results in the targets detection of conventional natural images, there are still some problems in the direct application of the algorithm to the ship targets detection of remote sensing images, such as inaccurate recognition, low recall rate,
large detection frame offset and so on. It is mainly affected by the following factors: first, the ship targets in the remote sensing image have the characteristics of small size and dense arrangement, which increases the complexity of targets detection task; second, the remote sensing images are changeable and vulnerable to factors such as illumination, clouds and atmospheric conditions, and the speckle noise is serious, which will lead to missed detection due to the unbalanced distribution of positive and negative samples in the training process. In order to solve the problems faced by ship targets detection in remote sensing images, based on Faster-RCNN algorithm, this paper proposes an improved algorithm to solve the problem of low ship target detection rate by optimizing the feature extraction network.

2. Improved Faster-RCNN target detection algorithm

2.1. Basic principle of Faster-RCNN

Faster-RCNN is a region based convolutional neural network framework, which is mainly composed of four parts: feature extraction layer, RPN layer, region of interest (ROI) pooling layer and detection network. The network structure is shown in Figure 1. The feature extraction layer extracts the features of the input image through convolution neural network; The RPN performs convolution operation on the obtained image feature map to generate candidate boxes that may contain the target objects; The ROI pooling layer converts the region feature map corresponding to different candidate boxes into a feature map with the same size; The detection network performs target recognition and candidate frame position correction on the region of interest, and obtains the final detection result.

The detection process of Faster-RCNN is as follows: ① firstly, the training image is scaled and input into the network, and the image feature map is extracted through the feature extraction layer; ② Then, the feature map is transmitted to the RPN layer, which generates candidate boxes according to the intersection over Union (IOU) threshold; ③ Then, the candidate boxes generated by the RPN layer and the feature map generated in step① are introduced into the ROI pooling layer to obtain a fixed scale(7×7) mapping feature map of candidate boxes; ④ Finally, it is input into the detection network, and the candidate boxes passing through the ROI pooling layer is regressed to obtain the final boxes positions. At the same time, the target categories are identified to generate the classification results.
3. Experimental results and analysis
The text of your paper should be formatted as follows:

The use of sections to divide the text of the paper is optional and left as a decision for the author. Where the author wishes to divide the paper into sections the formatting shown in table 2 should be used.

3.1. Experimental data
The experimental data set used in this paper is 215 remote sensing images in HRSC2016[6] and DOTA data set[7]. The HRSC2016 data set consists of remote sensing images of six world-famous ports, with image resolution between 0.4m and 2.0m and image size between 300×300 and 1500×900 pixels. The size of remote sensing image in DOTA data set is about 4096×4096 pixels, which has similar resolution to HRSC2016 data set. The remote sensing images in the experimental data set are distributed under different lighting conditions and cloud occlusion. The data set contains a large number of ship targets with various shapes and directions.

3.2. Training strategy
The data set is randomly divided into training set, verification set and test set according to the ratio of 8:1:1. During the training, the batch size is set to 4. A total of 100 epochs are trained using Adam optimizer. The initial learning rate, momentum and weight attenuation parameters are set to 0.001, 0.9 and 0.0005 respectively, and the learning rate is reduced to 0.0001 after 10000 iterations. The data enhancement strategy of the algorithm includes pixel level data enhancement method, simulated target occlusion method and multiple image mixed data enhancement method.
3.3. Analysis of experimental results

Firstly, the algorithm in this paper is compared with the Faster-RCNN algorithm. Figure 3 shows the comparison results of the test images detected by the two algorithms. Figure 3(a) shows the remote sensing image of a port, and ship targets in the image have the characteristics of small size and dense arrangement. Figure 3(b) shows the remote sensing image of the open sea, and image quality is poor due to the influence of clouds and illumination. In the detection results, the red boxes indicate the correctly detected target, the blue boxes indicate the incorrectly detected target, and the green boxes indicate the missed target. It can be found that compared to Faster-RCNN algorithm, this algorithm can detect more correct targets, and have fewer error detection targets and missed detection targets.

![Detection Results Comparison](image)

(a) Densely arranged small targets
(b) Targets affected by clouds

The proposed algorithm FasterR-CNN

Figure 3. Comparison of detection results of two algorithms

In order to fully illustrate the effectiveness of the algorithm, multiple quantitative evaluation indexes are adopted in the experiment to compare the detection effect of the algorithm with that of yolov3, yolov4 and FasterR-CNN in the test data set. The results are shown in Table 1. The calculation formulas of evaluation indexes are as follows:

\[
\text{Precision} = \frac{TP}{TP + FP} \quad (1)
\]

\[
\text{Recall} = \frac{TP}{TP + FN} \quad (2)
\]

in equations (1) and (2), TP (true positives) represents the number of correct detections, FN (false negatives) represents the number of missed detections, and FP (false positives) represents the number of error detections. In practical application, Recall and Precision cannot have both. When Precision is
high, recall is low, and when Recall is high, Precision is low. F1-score considers both accuracy rate and recall rate to find a balance between them. Its calculation formula is:

$$F1-score = \frac{2*Precision*Recall}{Precision + Recall}$$ (3)

AP (average precision) represents the average precision. It is obtained by calculating the accuracy under different recall rates and then averaging.

Table 1. Comparison of evaluation indicators

| Algorithms          | Precision | Recall | F1-score | AP |
|---------------------|-----------|--------|----------|----|
| YOLOv3              | 0.80      | 0.78   | 0.79     | 0.68|
| YOLOv4              | 0.82      | 0.81   | 0.81     | 0.74|
| FasterR-CNN         | 0.84      | 0.81   | 0.82     | 0.75|
| The proposed algorithm | 0.91      | 0.89   | 0.90     | 0.80|

Experimental results show that compared with other algorithms, the proposed algorithm can effectively improve the detection accuracy of ship targets. This is because in other algorithms, grid cells are generally responsible for larger pixel areas, and larger pixel areas need to predict more possible targets. Therefore, in the case of high density, overlapping and small targets, prediction errors are more likely to occur. The network unit of this algorithm can predict the target in a small pixel area, and can overcome the adverse impact of poor remote sensing image quality to a certain extent, which is more conducive to target prediction.

4. Conclusion

Aiming at the problem of ship targets detection in remote sensing images, a detection method based on improved FasterR-CNN network is proposed to overcome the problems of missed detection and false detection caused by poor image quality and small targets. The experimental results show that compared with the methods of YOLOv3, YOLOv4 and FasterR-CNN, the proposed algorithm can effectively improve the ship detection accuracy and reduce the false detection probability.

In the next step, we will consider how to unify the lightweight and precision of the model, effectively extract the target features, reduce the network level of the model and reduce the complexity of the model; And further discuss how to improve the effect of ship target detection based on limited data sets in the case of small samples.

Acknowledgments

This work is supported in part by the Fundamental Research Funds of Engineering University of People’s Armed Police under Grant WJY202102, and in part by the Military Theory Research Project of Engineering University of People’s Armed Police under Grant JLY2021052.

References

[1] ZHU J, QIU X, PAN Z, et al. Projection shape template-based ship target recognition in TerraSAR-X images[J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(2): 222-226.
[2] QI S, JIE M, JIN L, et al. Unsupervised ship detection based on saliency and S-HOG descriptor from optical satellite images[J]. IEEE Geoscience & Remote Sensing Letters, 2015, 12(7):1451-1455.
[3] TANG J, DENG C, HUANG G B, et al. Compressed-domain ship detection on spaceborne optical image using deep neural network and extreme learning machine[J]. IEEE Transactions on Geoscience & Remote Sensing, 2014, 53(3):1174-1185.
[4] DONG Y, DU B, ZHANG L. Target detection based on random forest metric learning[J]. IEEE Journal of Selected Topics in Applied Earth Observations& Remote Sensing, 2017, 8(4):1830-1838.
[5] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(6): 1137-1149. doi: 10.1109/TPAMI.2016.2577031.

[6] Z. Liu, L. Yuan, L. Weng, et al. “A high resolution optical satellite image dataset for ship recognition and some new baselines” in Proc. 6th Int. Conf. Pattern Recognit. Appl. Methods, 2017, 324–331.

[7] G.-S. Xia, X. Bai, J. Ding, et al. “DOTA: A large-scale dataset for object detection in aerial images” in Proc. IEEE Comput. Conf. Comput. Vis. Pattern Recognit, Jun. 2018, 3974–3983.