MULTI-SCALE SHIPS DETECTION IN HIGH-RESOLUTION REMOTE SENSING IMAGE VIA SALIENCY-BASED REGION CONVOLUTIONAL NEURAL NETWORK

Zezhong Li, Yanan You, Fang Liu

School of Information and Communication Engineering, Beijing University of Posts and Telecommunication

ABSTRACT

Ship detection is of great significance in both military and civilian application domains. Deep Convolutional Neural Network (DCNN) method with region proposal, e.g. Faster R-CNN, achieves ship detection well. However, for multi-scale target detection in high-resolution remote sensing image, the limitation of accuracy is induced by the region proposal restricted by the training set. Therefore, the mechanism of multi-scale ship detection based on saliency estimation is proposed in our work. Firstly, a saliency estimation algorithm is used to distinguish which image contains large ships and the image pyramid for each input one is established. Then, using a target detection network in different scales of images. The results are merged at the end of network. Finally, accuracy and validity are verified by real data processing.

Index Terms—saliency estimation, region proposal, DCNN, multi-scale ship detection

1. INTRODUCTION

Multi-scale ship detection is of great significance in both military and civilian fields [1]. With the development of deep learning, some researches discuss the object detection method based on Deep Convolutional Neural Network (DCNN). Specially, the method based on region proposal mechanism, such as Faster R-CNN [2], achieves ship detection well. Ren et al [3] modify the Faster R-CNN by adopting top-down and skip connections to accelerate the accuracy of small ships. Kang et al [4] combine the Faster R-CNN and CFAR to detect the ships in SAR images. In recent, a method with image pyramid [5] is utilized to transform the objects to be a uniform scale and reduce sample diversity.

In fact, the ship targets in high-resolution remote sensing image present the characteristics of multi-scale, multi-direction and multi-shape. However, there is a significant difference between the training set and the actual application. The targets are more similar in training dataset. The region proposal determined by the training set are not necessarily suitable for multi-scale targets. Therefore, aiming at the task of multi-scale ship detection, a novel method based on the saliency estimation is proposed in this paper. Firstly, a saliency estimation algorithm, SIM [6] is used to differentiate whether the image contains large targets, and image pyramid is built to compress the large-scale target. Secondly, a Region Proposal Network (RPN) is used to conduct the ship detection in image pyramid. Finally, the performance of our method is compared with the conventional Faster R-CNN by using a high-resolution ship dataset, and the accuracy and validity are verified by real data processing.

The rest part of this paper is organized as follows. Section 2 introduces the phenomenon of multi-scale ships in high-resolution remote sensing image. The saliency-based R-CNN method proposed in our work is demonstrated in Section 3. In section 4, real data experiments are designed to evaluate the proposed method. The related conclusions are found in the end.

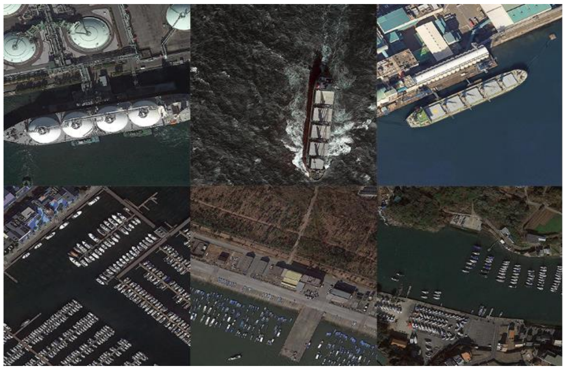

Figure 1 Multi-scale ships in the high-resolution remote sensing images (same resolution)

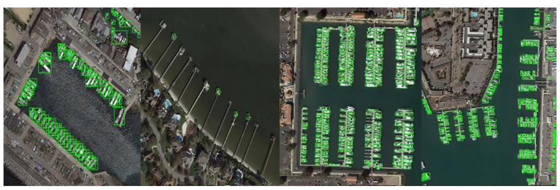

Figure 2 Samples in DOTA dataset.

2. MULTI-SCALE SHIPS

Ship targets in high-resolution remote sensing image, shown in Figure 1, obviously take on different shapes. However, many similar targets, shown in Figure 2, appear in the training dataset, for example, DOTA [7] used in our work. Once the DCNN is driven by this high similarity sample set, it induces the difficulty of multi-scale target detection. Especially for some large-scale targets with a small proportion, the detection accuracy is limited.

An experiment was conducted on DOTA dataset with Faster R-CNN, in which, the images were processed to the same image size and the evaluation of AP is 0.53 on validation dataset, Moreover, the trained model is tested on a dataset of high-resolution remote sensing images. Although the targets, which are similar to training set, are well recognized, some large-scale ship are lost just as shown in Figure 3. Therefore, when the sample distribution of validation data is dissimilar to that of the training data, the trained network is not very robust for these unlearned samples. In fact, multi-scale targets must exist in high-resolution remote sensing images.

Figure 3 The large-scale ship is not identified while the small ships are identified.

3. SALIENCY-BASED R-CNN

In order to solve the problem of multi-scale ships detection, a DCNN method based on saliency estimation is proposed. The sample diversity is reduced via pre-process of saliency estimation. Differentiate the target scene with distribution characteristic and build image pyramid to make the target scale consistent with the training dataset, and then use RPN to detect the targets in the processed image.

3.1. Saliency Estimation

The purpose of saliency estimation is to describe the distribution of saliency targets in the image and most of widely used saliency estimation methods are summarized in Cheng`s [8] paper, such as SIM, CA, SEG. In order to distinguish images containing targets of different scales, the saliency estimate of each image is obtained, as shown in Figure 5, in which, a statistical histogram for each result of saliency estimation is calculated, and the statistical results of SIM method have obvious resolvability in different scenes.

Therefore, the following statistical formula is used to distinguish different types of ships:

$$\begin{cases} 1 & if \ \frac{1}{80}\sum_{i=20}^{100} X_i \geq \frac{1}{40}\sum_{i=130}^{170} X_i \geq 2000 \\ 0 & otherwise \end{cases}$$

Where X_i stands for statistical information of saliency feature map.

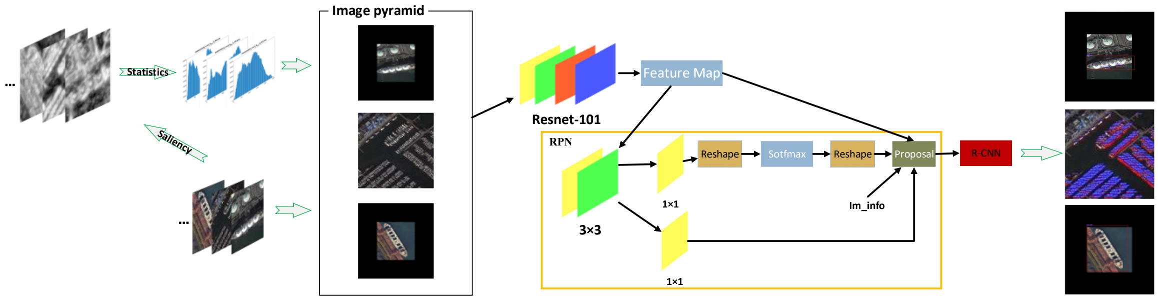

Figure 4 The flowchart of saliency-based R-CNN

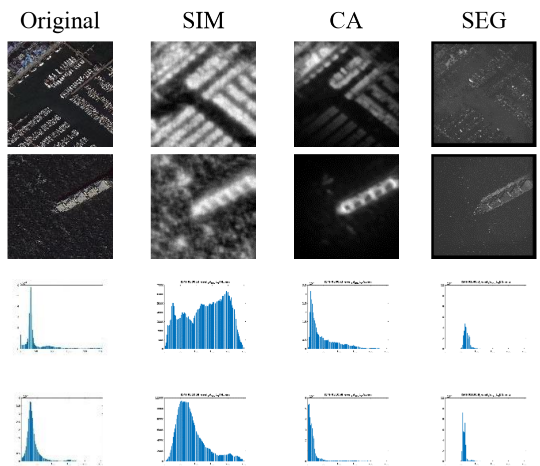

Figure 5 Saliency estimation results and statistical histograms for SIM, CA and SEG.

3.2. Region Proposal Network

In Faster R-CNN, RPN is employed to search different targets in the feature map. RPN adjust the pre-defined anchors coordinates and determine the category for the target detection by training process, which is shown in Figure 6. Especially for small target, the candidate box is limited to a small range. When a large-scale target appears, it will cause the loss of the target, which is shown in Figure 3.

3.3. The Implementation of Saliency-based R-CNN

The saliency-based R-CNN, proposed in this paper, is a combination of image pyramid structure based on saliency estimation and Faster R-CNN, as shown in Figure 4. Firstly, saliency map for each input image is calculated by using SIM algorithm. Then, the images of different target distributions are distinguished according to the statistical result. Finally, image pyramid is established suitable for RPN.

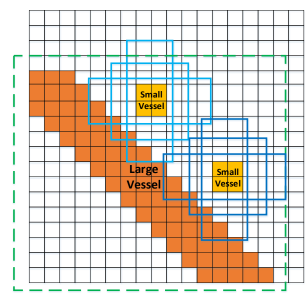

Figure 6 The mechanism of RPN. The green dotted box, covering the large-scale target, will be discarded during the training process while the blue box for small targets can be trained.

4. EXPERIMENTS

4.1. Result and Analysis

To demonstrate the validation of the proposed method, the dataset including 90-labeled high-resolution remote sensing images is used in the experiment. AP (Recall and Precision) on whole dataset and AP on large-scale ship images are evaluated for our method and Faster R-CNN. The result is shown in Table 1. Moreover, the result image of ship detection is shown in Figure 7.

In Table 1, AP for large-scale ship (AP^L) has a significant improvement in our method, which is 0.776 and 0.778 respectively. Anchors size is also changed in this experiment, the result indicates that various anchor set is conducive to target detection. Moreover, the detection results for image pyramid are merged in final step, so that it does not decrease the accuracy of small target.

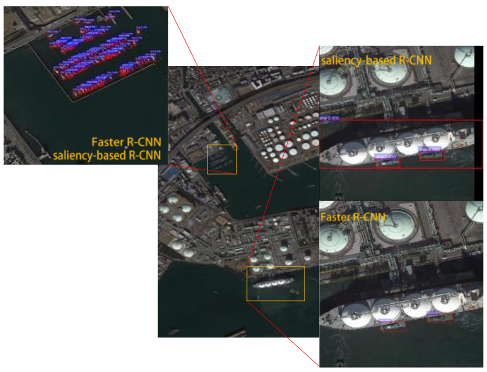

Figure 7 Comparison of Faster R-CNN and saliency-based R-CNN in high-resolution remote sensing image.

Table 1 Assessment results of saliency-based R-CNN and Faster R-CNN

Methods	Anchors	AP	AP^L
Faster R-CNN	8,16,32	0.337	0.300
Ours	8,16,32	0.342	**0.776**
Faster R-CNN	1,2,4,8,16,32,48,64	0.422	0.386
Ours	1,2,4,8,16,32,48,64	0.425	**0.778**

5. CONCLUSION

In this paper, aiming at the task of multi-scale ship detection in high-resolution remote sensing image, we propose a ship detection method based on saliency-based R-CNN. Saliency estimation is calculated for each input image, and its statistical result is used to differentiate images that contain targets of different scales, and build image pyramids to fit the trained RPN. Compared with the Faster R-CNN, this method improves the accuracy of multi-scale ship detection, especially for large-scale targets.

ACKNOWLEDGEMENTS

This work is supported by the “Fundamental Research Funds for the Central Universities” (2018RC09). This work is conducted on the platform of Center for Data Science of Beijing University of Posts and Telecommunications.

REFERENCES

[1] Kanjir, U., Greidanus, H., & Oštir, K, "Vessel detection and classification from spaceborne optical images: A literature survey," Remote sensing of environment, 2018.

[2] Ren, S., He, K., Girshick, R., & Sun, J, "Faster r-cnn: Towards real-time object detection with region proposal networks," In Advances in neural information processing systems, pp. 91-99, 2015.

[3] Ren, Y., Zhu, C., & Xiao, S, "Small Object Detection in Optical Remote Sensing Images via Modified Faster R-CNN," Applied Sciences, 2018.

[4] Kang, M., Leng, X., Lin, Z., & Ji, K, "A modified faster R-CNN based on CFAR algorithm for SAR ship detection," In Remote Sensing with Intelligent Processing (RSIP), 2017 International Workshop on, pp. 1-4, May. 2017.

[5] Singh, B., & Davis, L. S, "An Analysis of Scale Invariance in Object Detection–SNIP," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3578-3587, 2018.

[6] Murray, N., Vanrell, M., Otazu, X., & Parraga, C. A, "Saliency estimation using a non-parametric low-level vision model," In CVPR, pp. 433-440, June. 2011.

[7] Xia, G. S., Bai, X., Ding, J., Zhu, Z., Belongie, S., Luo, J., & Zhang, L, "DOTA: A large-scale dataset for object detection in aerial images," In Proc. CVPR, 2018.

[8] Borji, A., Cheng, M. M., Jiang, H., & Li, J, "Salient object detection: A benchmark," IEEE transactions on image processing, 2015.