HIGH-SPEED SHIP DETECTION IN SAR IMAGES BY IMPROVED YOLOV3

TIANWEN ZHANG[1], XIAOLING ZHANG[1], JUN SHI[1], SHUNJUN WEI[1]

[1]School of Information and Communication Engineering, University of Electronic Science and Technology of China
E-MAIL: twzhang@std.uestc.edu.cn, xlzhang@uestc.edu.cn, shijun@uestc.edu.cn, weishunjun@uestc.edu.cn

Abstract:

Ship detection in synthetic aperture radar (SAR) images plays an important role in marine transportation, fishery management, and maritime disaster rescue. Nowadays, the current researches almost are focusing on improving detection accuracy while detection speed is neglected. However, it is also extraordinarily important to increase the ship detection speed, because it can provide real-time ocean observation and timely ship rescue. Therefore, in order to solve this problem, this paper proposes a high-speed SAR ship detection approach by improved you only look once version 3 (YOLOv3). We experimented on a public SAR ship detection dataset (SSDD) which has been used by many other scholars. Finally, the experimental results indicated that the detection speed of our proposed improved YOLOv3 is faster than current other methods, such as faster-regions convolutional neural network (Faster R-CNN), single shot multi-box detector (SSD), and original YOLOv3 under a same hardware environment. Meanwhile, the detection accuracy remains basically unchanged.

Keywords:

Ship detection; SARyou only look once (YOLO)

1. Introduction

Synthetic aperture radar (SAR), an all-weather and all-time microwave sensor, is one of the most important tools in remote sensing filed. Up to now, many scholars have proposed many SAR ship detection methods, which have greatly promoted the development of SAR image interpretation. Constant false alarm ratio (CFAR) [1-3] is a classical algorithm which can provide a threshold for a specific target by establishing the model of image background and object. Among them, two-parameter CFAR [1], CFAR based on Weibull distribution [2], and optimal CFAR [3] are widely used. However, when CFAR is used, the computation of background clutter distribution is huge, which reduces the detection speed. Another common method is template-based detection [4], which can provide specific target templates according to different scenarios. However, the establishment of template library excessively relies on expert experience, and its generalization ability is weak. Moreover, due to the diversity of templates, it takes more time to select the right template in practical detection, which will also reduce the detection speed.

Nowadays, with the development of artificial intelligence technology, more and more experts begin to study the methods of ship detection based on data-driven and artificial intelligence. Probably a simple explanation to this fact comes from that artificial intelligence methods can automatically extract ship's features, avoiding the manual feature engineering of traditional methods, which greatly improves the detection efficiency. Region-convolutional neural network (R-CNN) [5] is the first to introduce artificial intelligence and deep learning into object detection field, which pioneers deep learning object detection study. However, its detection speed is rather slow because of the large number of calculations. Therefore, Fast R-CNN [6] improved the speed by automatically obtaining Region of Interest (ROI). However, its speed is still not fast enough as the extraction of ROI takes up a large proportion of the time, which reduces detection efficiency. Ultimately, Faster R-CNN [7] was proposed to solve the above problem via region proposal network (RPN). However, the above three methods are all two-stage: 1) extraction of ROI; 2) object detection. Therefore, in order to simplify the detection process, one-stage detector YOLOv3 was proposed by Joseph [8], which removes the process of extraction of ROI. Therefore, the speed gets further improvement by YOLOv3. In short, from the development of the above methods, we can conclude that the detection speed is very important besides accuracy.

However, according to our survey, in SAR ship detection field, many scholars are focusing on improving the accuracy of ship detection while the detection speed is neglected. In fact, it is also extraordinarily important to increase the ship detection speed, because it can provide real-time ocean observation and timely ship rescue. Therefore, in order to solve this problem, this paper proposes a high-speed SAR ship detection approach by improved you only look once version 3 (YOLOv3). We improved YOLOv3 by reducing the size of the network to reduce time consumption which can further increase the

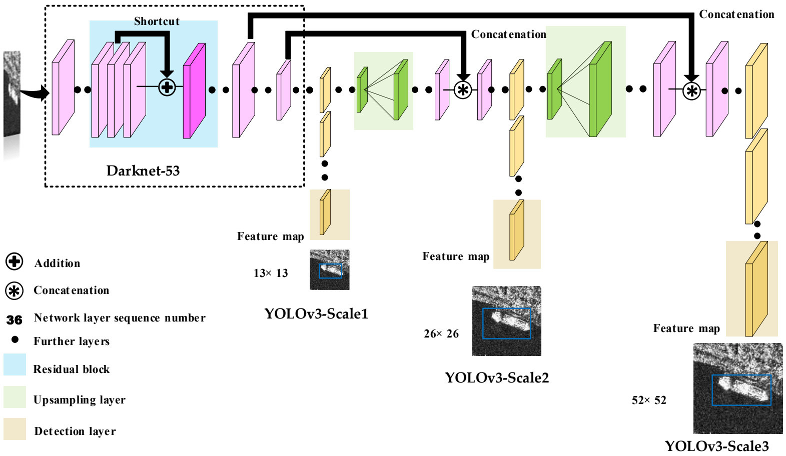

Fig.1 The network structure of original YOLOv3

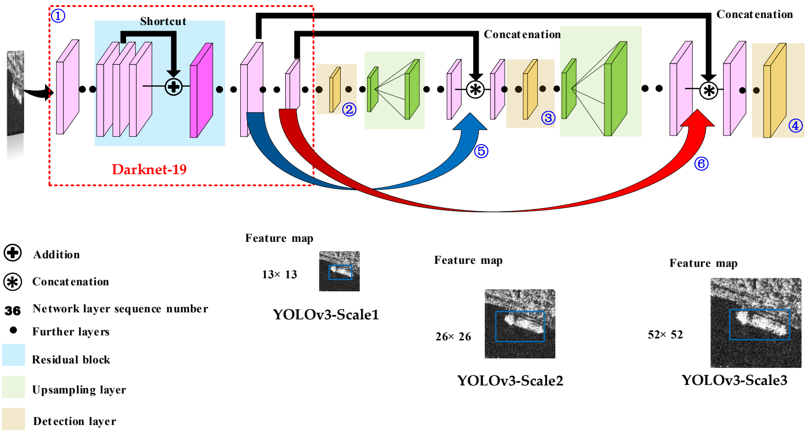

Fig.2 The network structure of improved YOLOv3

detection speed. In order to compare fairly with other studies, we selected an open SSDD dataset for experiments. The experimental results indicated that the detection speed of our proposed improved YOLOv3 is faster than current other methods, such as Faster R-CNN, SSD, and YOLOv3 under a same hardware environment. Our approach achieves high-speed ship detection in SAR images, authentically, requiring only 24ms per image, whose speed is 2.3 times faster than the original YOLOv3, 3.4 times faster than SSD, 8.3 times faster than Faster R-CNN and 12.8 times faster than Fast R-CNN.

2. Methodology

2.1. Original YOLOv3

The network structure of original YOLOv3 is shown in fig.1. From fig.1, YOLOv3 consists of four main parts: Darknet-53, YOLOv3-Scale1, YOLOv3-Scale2, and YOLOv3-Scale3.

2.2. Improved YOLOv3

In order to improve the speed of ship detection, we have improved the original YOLOv3. Different from the original YOLOv3 where 20 types of targets need to be detected, SAR target detection in this paper contains only one class that is ship, so the reduction of network size does not significantly reduce accuracy by our research findings. The network structure of improved YOLOv3 is shown in fig.2. The detailed improvements are shown in the table 1. From fig.2 and table 1, to reduce the size of the network, we will use Darknet-19 as the backbone of the improved YOLOv3, which can reduce detection time. We also deleted repeated layers in YOLOv3-Scale1, YOLOv3-Scale2, and YOLOv3-Scale3. Finally, in order to make full use of the features extracted from the network, we have added two feature concatenation paths, which can improve the detection accuracy.

Table 1 The detailed improvements.

Number	Explanation
1	Change the backbone from Darknet-53 to Darknet-19.
2	Delete repeated layers of YOLOv3-Scale1.
3	Delete repeated layers of YOLOv3-Scale2.
4	Delete repeated layers of YOLOv3-Scale3.
5	Increase a concatenation.
6	Increase a concatenation.

3. Experiments and results

3.1. Dataset

When using the method of deep learning to detect ships, the first factor is to build an effective and typical dataset where real ships (ground truth) have been labeled. SAR ship detection dataset (SSDD) is an open dataset, created by Li [9], which has been used by many other scholars.

There are 1160 SAR images in SSDD and there are 2358 ships in all these images. In order to ensure the rationality of SSDD dataset, these 1160 SAR images come from different satellites, different polarization modes and different resolutions. Moreover, the backgrounds of ships are various.

3.2. Experimental Procedures

Data preparation. In order to adapt to the input of YOLOv3 network, the size of all image is resized to 416×416.

Divide SSDD dataset. We divided the 1160 images into training set, verification set and test set according to the ratio of 7:2:1, randomly. Training set is used to train. The verification set is responsible for adjusting the parameters to avoid over-fitting in the training process. The test set is used to test the performance of ship detection.

Formulate training strategies. We set the learning rate to 0.001 in advance, and dynamic adjustments are made according to the training logs, as well. Because of the large amount of training data, we set batch size to 8 in order to quickly realize the iteration and adjustment of parameters. Adaptive moment estimation (Adam) [10], an algorithm similar to stochastic gradient descent (SGD), is used to update the weights and biases in the network, with the advantages of efficient calculation and less memory. Adam is suitable for solving problems with high noise or sparse gradient, which fully accords with the characteristics of SAR image and ship sparse distribution. Finally, when one epoch of training is completed, the network will automatically verify on the verification set to avoid the occurrence of over-fitting, according to the loss of verification set.

Start training. According to training strategies, we started training and fine tuning. The weights in each iteration are saved in order to select the one with best performance to detect ships in SAR images.

3.3. SAR Ship Detection Results

After some reasonable trainings, when the loss reaches the minimum, we get the final detection model. Then, we perform the actual SAR ship detection on the test set. The detection results of some samples are shown in the fig.3.

From fig.3, almost all real ships can be detected correctly, and the original YOLOv3 and the improved YOLOv3 have similar detection accuracy.

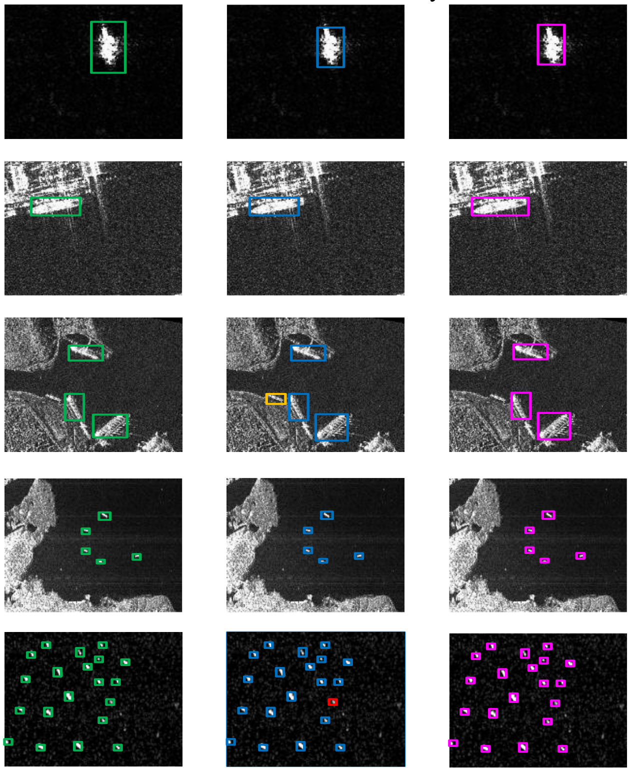

Fig.3 SAR ship detection results. (a) Ground truth. (b) Original YOLOv3. (c) Improved YOLOv3. Red is miss-detection and yellow is false alarm

Table 2 is the comparison of their quantitative evaluation indicators. From fig.3 and talbe 2, the performance of the improved YOLOv3 is slightly better than the original YOLOv3, but the gap is narrow. More importantly, on the premise of keeping the accuracy basically unchanged, the improved network is smaller and lighter, which can reduce the detection time. In addition, due to the added two concatenation, the detection accuracy can maintain unchanged.

Table 2 The detailed improvements.

Method	Recall	Precision	mAP	F1-score
Original YOLOv3	89.12%	88.57%	90.14%	0.89
Improved YOLOv3	91.10%	89.36%	90.08%	0.90

4. Discussion

Under the same computer hardware condition, we compared the detection time of a same SAR image with different methods, such as Fast R-CNN, Faster R-CNN, SSD, the original YOLOv3 and the improved YOLOv3. The results are shown in the fig.4.

From fig.4, the detection time of our proposed improved YOLOv3 is the least. In other words, the detection speed is the fastest, so the high-speed SAR ship detection is realized truthfully. The detection speed of our improved YOLOv3 is 2.3 times faster than the original YOLOv3, which means that our improvements are correct and effective.

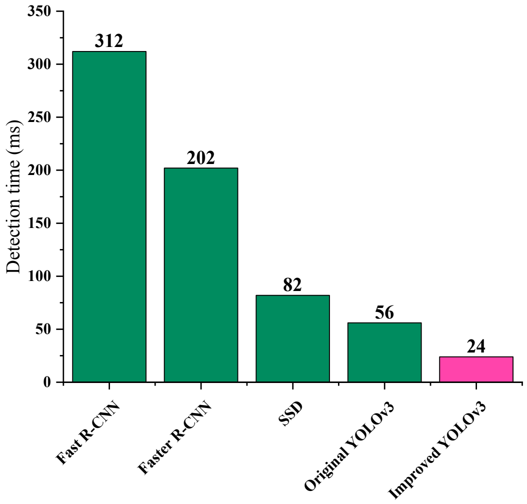

Fig.4 Detection time of different methods.

5. Conclusions

Aiming at the problem that the detection speed is neglected at present, a high-speed SAR ship detection method is proposed in this paper by improved YOLOv3. Experiments on an open SSDD dataset demonstrated the correctness and effectiveness of our proposed method. The experimental results indicated that the detection speed of our proposed improved YOLOv3 is faster than current other methods, such as Faster R-CNN, SSD, and YOLOv3 under a same hardware environment. Our approach achieves high-speed ship detection in SAR images, authentically, requiring only 24ms per image, whose speed is 2.3 times faster than the original YOLOv3, 3.4 times faster than SSD, 8.3 times faster than Faster R-CNN and 12.8 times faster than Fast R-CNN.

Acknowledgements

This work was supported in part by the National Key R&D Program of China under Grant 2017YFB0502700 and in part by the National Natural Science Foundation of China under Grants 61571099, 61501098, and 61671113.

References

[1] Raj, Nithin , R. Sethunadh , and P. R. Aparna. Object detection in SAR image based on bandlet transform. Academic Press, Inc. 2016.

[2] Kuttikkad, S.; Chellappa, R. Non-gaussian CFAR techniques for target detection in high resolution SAR images. In Proceedings of the IEEE International Conference on Image Processing (ICIP-94.), Austin, TX, USA, 13–16 November 1994.

[3] Anastassopoulos, V. , and G. A. Lampropoulos. Optimal CFAR detection in Weibull clutter. IEEE Transactions on Aerospace and Electronic Systems 31.1(1995):52-64.

[4] Zhu, Jiwei , et al. Projection Shape Template-Based Ship Target Recognition in TerraSAR-X Images. IEEE Geoscience and Remote Sensing Letters 14.2(2017):222-226.

[5] Girshick, R. , Donahue, J. , Darrelland, T. , & Malik, J. . (2014). Rich feature hierarchies for object detection and semantic segmentation. 2014 IEEE Conference on Computer Vision and Pattern Recognition. IEEE.

[6] Girshick, R. Fast R-CNN. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 1440–1448.

[7] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: towards real-time object detection with region proposal networks. International Conference on Neural Information Processing Systems.

[8] Joseph, R.; Ali, F. YOLOv3: An Incremental Improvement. arXiv 2018, arXiv:1804.02767.

[9] Li, Jianwei , C. Qu , and J. Shao. Ship detection in SAR images based on an improved faster R-CNN. 2017 SAR in Big Data Era: Models, Methods and Applications (BIGSARDATA) IEEE, 2017.

[10] Kingma, D.; Ba, J. Adam: A Method for Stochastic Optimization. arXiv 2014, arXiv:1412.6980.