Ship Detection in SAR Images Based on an Improved Detector with Rotational Boxes

Xiaowei Ding[1], Changbo Hou[1,2(✉)], and Yongjian Xu[1]

[1] Harbin Engineering University, Harbin 150001, HLJ, China
houchangbo@hrbeu.edu.cn
[2] Key Laboratory of Advanced Marine Communication and Information Technology, Ministry of Industry and Information Technology, Harbin Engineering University, Harbin 150001, HLJ, China

Abstract. In the SAR ship data under complex backgrounds, especially in the coastal area, the horizontal bounding box detection algorithm makes a large number of coastal noise interference targets feature extraction and bounding box regression. In addition, the horizontal bounding box cannot well reflect the characteristics of large aspect ratio of ships. Therefore, this paper proposes an improved YOLOv3 detection algorithm based on the rotational bounding box, which increases the encoding method of the angle parameter, and generates the prediction bounding box at a fixed angle interval. Different angle intervals will have different effects. Focus loss function is used to solve the problem of positive and negative sample balance and difficult sample feature learning. The experimental results show that the average precision of the R-YOLOv3 algorithm based on the rotational bounding box on the SAR ship data set is 87.3%, which is a 13.5% gain compared with the classic YOLOv3, which reflects the high precision of the ship targets.

Keywords: SAR image · Ship target detection · Convolutional neural network

1 Introduction

Ship detection is one of the main technologies for maritime surveillance, which is extremely important for maintaining maritime security, monitoring maritime transportation and improving the capability of maritime defense and early warning [1]. At present, the data sources of ship detection are mainly optical sensors, infrared sensors and SAR sensors. Synthetic aperture radar (SAR) [2] is an active side microwave imaging sensor. Compared with infrared and optical passive sensors, it has the advantages of penetrating clouds and working day and night. It can collect large-area data anytime and anywhere under any weather conditions such as daytime, nighttime and foggy days, and generate high-resolution images by comprehensive utilization of signal processing, pulse compression and synthetic aperture principle. The traditional methods of target detection in SAR images can be divided into threshold-based methods [3], salient region-based methods [4], texture-based methods [5] and statistical analysis-based methods [6]. Among

J. Xiong et al. (Eds.): MobiMedia 2021, LNICST 394, pp. 824–836, 2021.
https://doi.org/10.1007/978-3-030-89814-4_61

these methods, Constant False Alarm Rate (CFAR) [7] and its variants are most widely used. In recent years, with more and more successful launch of SAR satellites, SAR image acquisition becomes more and more easy and high-resolution image data increases. The development of high-precision and high-efficiency target detection system has attracted much attention. Deep learning algorithm has a broad application prospect in ship target detection and image classification [8, 9] due to its strong autonomous learning ability [10] and feature representation ability [11]. At present, the algorithms based on convolutional neural network are Faster R-CNN [12], YOLOv1 [13], YOLOv3 [14], SSD [15], RetinaNet [16], etc. Subsequently, more and more improved algorithms are applied to SAR ship target detection. For example, in 2017, Kang et al. [17] proposed a region-based R-CNN target detection algorithm based on multi-scale feature fusion, which combines shallow and deep features and is conducive to eliminating false alarms. Li et al. [18] constructed the first public dataset SSDD in the field of ship target detection in SAR images in 2017, and proposed an improved detector based on Faster R-CNN based on multi-technology fusion. Then Kang et al. [19] proposed a new Faster R-CNN detection network by combining CFAR algorithm and Faster R-CNN algorithm. Faster R-CNN uses the sliding window generated by CFAR algorithm as a candidate region to detect small ship targets, and obtains better detection performance.

Ship detection will have recognition problems. Usually, the direction of ship targets in SAR images is diverse. Most of the ship detection in the offshore area will only receive the influence of sea clutter, and the detection task is simple. However, in the coastal area, the scene is complex, and the ship arrangement is very dense. The annotation method of horizontal box cannot distinguish the target from the target, as well as the target and the background. The background interference is serious, and it is easy to cause missed detection. It is difficult to distinguish small ship targets from speckle noise, which intensifies the difficulty of ship target detection. Therefore, based on YOLOv3, this paper proposes a ship target detection algorithm based on rotational boxes in SAR image. Firstly, R-YOLOv3 uses ResNet-50 as the backbone network of feature extraction, and improves the extraction of complex features by residual module [20]. Secondly, in order to improve the detection precision of small SAR ships, R-YOLOv3 draws on the idea of FPN [21] to perform multi-scale fusion and independent prediction of the extracted features. Then, in view of the characteristics of ship direction diversity and the problem of large background interference in coastal areas, the detection algorithm adds the encoding method of angle parameters, and generates the prediction boundary box according to the fixed angle interval, and improves the calculation method of rotational boxes cross-parallel ratio to solve the influence of angle change on the detection accuracy. Finally, the algorithm uses the focus loss function to reduce the weight of the samples that are easy to classify, so that the model focuses on learning the characteristics of the foreground objects with less difficult classification and improves the detection precision of the model for ships.

2 Methods

2.1 Overall Scheme of R-YOLOv3

The one-stage detector R-YOLOv3 proposed in this paper uses the idea of residual network structure and RPN multi-scale feature fusion to realize the boundary anchor box regression with angle information. The structure of our method is shown in Fig. 1.

Firstly, the SAR image to be detected is used as the input of the feature extraction network. The feature of the ship target is extracted through the ResNet50 backbone network, and the feature mapping of five layers with different sizes is obtained. The features of three different sizes extracted from the last layers of the feature extraction network, and are fused to achieve the regression of the border on the fused feature map. In the bounding box prediction stage, the box is predicted by the coordinate and angle information of the target. Considering the sensitivity of the intersection ratio to the angle change, a new intersection ratio strategy is adopted. The selection of loss function has an important influence on the performance of the algorithm. The Focal Loss function is used to solve the imbalance of positive and negative samples and the problem of difficult to classify samples, reduce the weight of easy to classify samples, and make the model focus on learning the characteristics of foreground objects with less difficult classification.

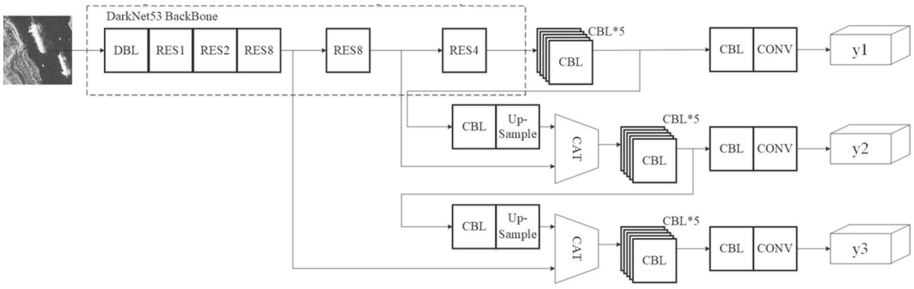

Fig. 1. Structure of R -YOLOv3 based on ResNet-50

2.2 Backbone Network for Feature Extraction

R-YOLOv3 uses ResNet-50 as the backbone network for extracting target features, and its structure is shown in Fig. 2. ResNet-50 is mainly composed of multiple residual units. Each residual unit contains the jump connection from input to output and the output obtained by three convolution operations. The convolution kernel sizes of the three convolution layers are 1×1, 3×3 and 1×1. This residual structure solves the problem of gradient disappearance and model detection accuracy reduction after network deepening. In order to improve the small target detection precision of the model, R-YOLOv3 draws on the idea of FPN feature fusion, and ResNet-50 has carried out five down-samplings. Therefore, the whole feature extraction process is divided into five

stages ($conv1$, $conv2$, $conv3$, $conv4$, $conv5$). The output characteristic figure of the last layer of each stage is (C_1, C_2, C_3, C_4, C_5).Because the number of pixels of the target is reflected by the resolution information, the large-scale feature map extracted by the shallow network has high resolution, less semantic information, small network receptive field and accurate location information, which can better detect small targets. The semantic information extracted by the deep network is rich in low-resolution features, and it feels rough about the large position information. Small targets are easy to miss. R-YOLOv3 performs up-sampling and concat respectively on the extracted features at three different scales. The features at different scales are fused with each other, and finally are predicted on the feature maps at three scales. The sizes of the three feature scales output are 32, 16 and 8 times the down-sampling of the input image resolution, and the size of the feature map output is 10×10, 20×20 and 40×40.

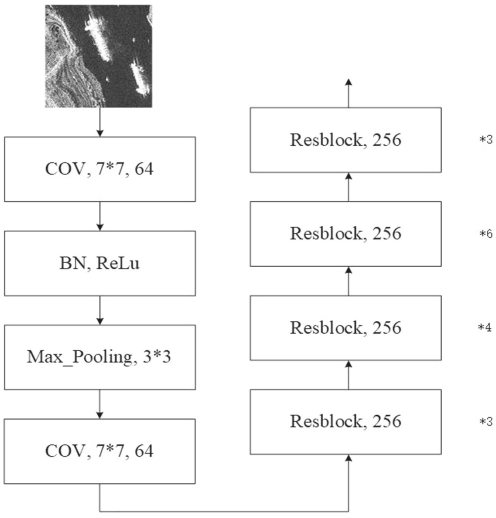

Fig. 2. Network structure of ResNet-50

2.3 Rotational Boundary Box Prediction

In the detection algorithm based on rotational box due to the increase of the angle parameter, the rotational anchor box generated at a certain anchor point is arranged at a certain angle interval, as shown in Fig. 3. Then, by calculating the intersection ratio between the angled anchor box and the ground truth, the anchor box with the largest intersection ratio is assigned to the ground truth. However, there will be a problem with this calculation method. At this time, the anchor box on the matching through the maximum crossover ratio is not the optimal boundary box. Assuming that two anchor boxes A and B, the length and width size of anchor box A is, the length and width size of anchor box B is, and the angle interval set by anchor box is 45°. The size of the real box C is, and its angle is 20°. At this point, when the angle value of anchor box A is 0°, the intersection ratio with

the real box reaches the maximum, which is 0.40; When the angle of anchor box B is 0°, the intersection ratio with the ground truth reaches the maximum, which is 0.45. According to the above matching criteria, the anchor box B matches the real box C at the angle of 0°. However, under the same angle deviation, the size deviation between anchor box B and real box C is 10, while there is no deviation between anchor box A and real box. Therefore, anchor box A is more suitable for allocation to real box C.

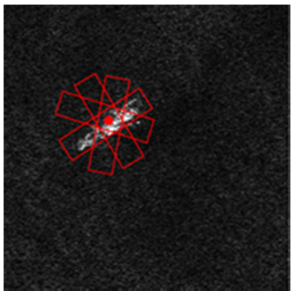

Fig. 3. Multi-angle rotational anchor box

Therefore, the matching result between the anchor box with angle and the ground truth is not completely optimal by directly using the above intersection and union ratio calculation.

$$ArIoU(A,B) = \frac{area\left(\hat{A} \cap B\right)}{area\left(\hat{A} \cup B\right)} |\cos(\theta_A - \theta_B)| \tag{1}$$

Among them, the first half of the formula is IoU between horizontal boxes, the second half is to measure the angle deviation between the two boxes, and the value range after the product is [0, 1].

R-YOLOv3 uses K-Means clustering algorithm to calculate the clustering results of the data set annotation box as the target prior box. The feature map of each scale matches three prior boxes, and nine prior boxes are obtained by three dimensional copolymerization classes. In the prediction boundary box stage, each cell on each feature map predicts three boundary boxes. The regression network learns the boundary box offset of the target according to the input characteristics. In addition to the four values of the center point coordinates, the angle parameter is also added. The predictive value conversion formula of rotational boundary box is defined as.

$$\begin{aligned} b_x &= \sigma(t_x) + c_x \\ b_y &= \sigma\left(t_y\right) + c_y \\ b_w &= p_w e^{t_w} \\ b_h &= p_h e^{t_h} \\ b_\theta &= (t_\theta + i) p_\theta \end{aligned} \tag{2}$$

$b_x, b_y, b_w, b_h, b_\theta$ is the center coordinate, width, height and angle of the prediction boundary box. c_x, c_y is the offset of the grid where the target center is located and the upper left corner of the feature map. i represents the angle interval of the anchor box.

2.4 Loss Function

In the R-YOLOv3 model, the size of the input image is 320 × 320.6300 prediction boxes will be generated. Most of them do not contain ship targets, and the distribution of positive and negative samples is unbalanced. In the training process, there will be a problem of uneven distribution of positive and negative samples, and the background contains a lot of noise interference. Compared with the foreground ship targets, it is very small and difficult to classify the samples. It is difficult to effectively train the model. Although the positive and negative sample imbalance problem is adjusted by setting the Ignore _ thread threshold and reducing the confidence of the boundary box that does not contain the target. However, the imbalance between positive and negative samples and the difficulty in classifying samples still exist. Therefore, the R-YOLOv3 algorithm further uses the Focal Loss function [22] to solve this problem, reduce the weight of the easy-to-class samples, make the model focus on learning the characteristics of the foreground objects with less difficult-to-class numbers, and improve the detection accuracy of the model for ships. The loss function of YOLOv3 algorithm is the sum of coordinate loss, confidence loss and classification loss function.

$$loss = bboxloss + confidenceloss + classloss \tag{3}$$

The class loss function used in YOLOv3 is the direct summation of the cross entropy loss function [23] of various training samples. It is defined as.

$$CE(p_t) = -\log_a(p_t) \tag{4}$$

Where, $p_t = \begin{cases} p & y = 1 \\ 1 - p & y = 0 \end{cases}$, p predicts the probability of the sample output category, y represents the label of the category.

In order to extract the features with more information from the model, the cross-entropy loss function is usually multiplied by a weight coefficient α inversely proportional to the probability of target existence, which weakens the contribution of a large number of negative samples to the model and increases the weight proportion of positive samples.

$$CE(p_t) = -\alpha_t \log_a(p_t) \tag{5}$$

$$\alpha_t = \begin{cases} \alpha & y = 1 \\ 1 - \alpha & y = 0 \end{cases} \tag{6}$$

Although the weight parameter solves the problem of imbalance between positive and negative samples in the training process, the ship targets in the distant waters in the ship data set are easy to detect, while the background noise in the coastal area is large,

and it is difficult to distinguish the ship from the background. Therefore, in the Focal Loss calculation formula, a modulation parameter is introduced to reduce the weight of simple samples through this super parameter, so that the model focuses on the learning of difficult-to-class samples.

$$FL(p_t) = -(1 - p_t)^{\gamma} \log_a(p_t) \tag{7}$$

Where, $\gamma \geq 0$. It is the standard cross entropy loss function when $\gamma = 0$.

3 Experiments and Results

3.1 Dataset

The experimental data in this paper are SAR Ship Detection Dataset (SSDD). SSDD data set is the first open dataset for SAR ship target detection created by Naval Aeronautics and Astronautics University in 2017. The dataset has a total of 1160 images, containing 2456 ship targets, only including the ship category. The experimental data set mainly adopts the rotational box labeling method. Different from the horizontal box data that are represented by the upper left and the lower right, the parameters of the rotational box are the coordinates of four points. For the ship target near the coast, the rotational box divides the target and background pixels, reduces the interference of artificial targets in the coastal area on the ship target detection, and the length-width ratio of SAR ship target is larger than other targets. The annotation method of the rotation box well reflects the real shape and size of the ship, and the length-width ratio and size of the horizontal box and the real shape of the ship, as shown in Fig. 4. In this paper, the partition ratio of the experimental training set and the test set is 8: 2.

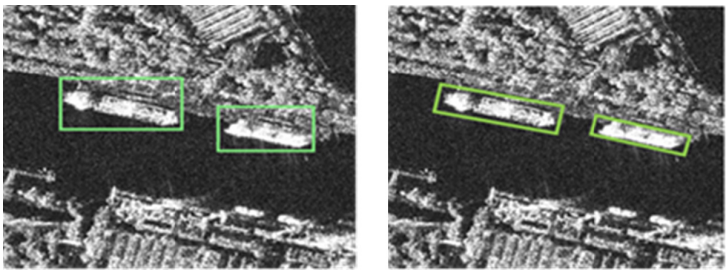

Fig. 4. Two target labeling methods

3.2 Evaluation Indicators

In the experiment of this article, three typical evaluation indicators in the target detection algorithm are used to evaluate the performance of different ship detection algorithms. They are accuracy, recall and average precision (AP).

(1) Accuracy. Also known as the precision rate, it is expressed as the probability of the true positive sample among all the samples predicted to be positive samples, reflecting the correctness of the detection target.

$$precision = \frac{TP}{TP + FP} \tag{8}$$

(2) Recall. Also known as recall rate, it is expressed as the probability that all positive samples that are correctly detected account for all positive samples.

$$recall = \frac{TP}{TP + FN} \tag{9}$$

(3) Average accuracy. Because both the precision rate and the recall rate have single-point value limitations and cannot reflect the complete performance of a detection model, the average precision is generally used to evaluate the pros and cons of the model. Use the accuracy rate of all detected images as the value of the ordinate and the recall rate as the value of the abscissa to draw the accuracy-recall rate curve (PR curve). The average accuracy is the area enclosed by the PR curve and the two coordinate axes.

$$AP = \int_0^1 P(R)dR \tag{10}$$

3.3 Detection Results

Since the size of the final extracted feature map of the R-YOLOv3 feature network is 32, 16 and 8 times of the down-sampling of the original image, the size of the input image should be an integer multiple of 32. The image size of the SSDD + dataset used in this section is not fixed. Therefore, first, the long side is cut into a square according to the short side of the image, and then the side length is filled with resizing to form a fixed size. This method can retain the original size of the ship as much as possible, especially the small target of the ship, reduce the deformation of the target and reduce the loss of feature information. Considering the distribution characteristics of the simple background and the complex background of the data set, the test set selects the images with the number 1 and 9 at the end of the image name, so as to ensure the uniform distribution of the target in the simple distant sea area and the target in the offshore complex background, and better evaluate the performance of the detection model. Before the training of the model, nine prior boxes were generated by K-Means clustering according to the true values of the tilt boundary box with angle information marked by the dataset, which were (5, 7), (7, 12), (8, 18), (9, 24) (11, 36), (15, 33), (17, 55), (27, 87) and (42, 142).

The picture size is 320 × 320. The batch size set by the training is 4, and the initial learning rate is 0.001. The SGD algorithm is used to optimize the model. The momentum coefficient is 0.9, and the weight attenuation rate during training is 0.0005. The IoU threshold set by the NMS is 0.5, and the confidence threshold is 0.01. The angle interval of the generated anchor box is 45°. All the samples in the training set achieve one-time training, which is called epoch. Each experiment trains 50 epochs.

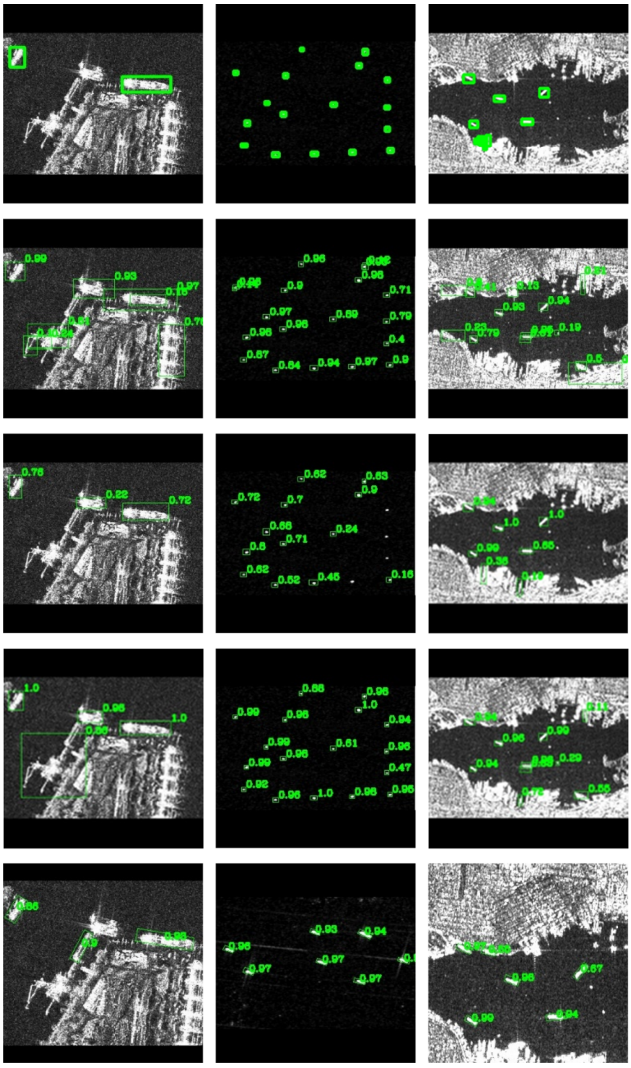

Fig. 5. SAR ship detection results. (a) Ground truth. (b) SSD. (c) Faster R-CNN. (d) Original YOLOv3. (e) R-YOLOv3

The detection visualization results of YOLOv3, SSD, Faster R-CNN and R-YOLOv3 are shown in Fig. 5. The observation results show that the three detection algorithms based on horizontal box have false detection, especially the interference of coastal background has a negative impact on the detection accuracy. In addition, Faster R-CNN has a good detection effect on large ships, but there is a missed detection of small ship targets. R-YOLOv3 separates the ship target from the coastal background, clusters the real target information to generate a more accurate priori box, provides more accurate

target location information, reduces the risk of missed detection and false detection, and improves the detection accuracy of the model.

Table 1. Comparison of evaluation metrics of different methods

Model	Backbone	Precision	Recall	AP
SSD	VGG-16	62.8	76.2	70.7
Faster R-CNN	VGG-16	66.0	80.1	74.6
YOLOv3	DarkNet-53	64.5	78.0	73.8
R-YOLOv3	ResNet-50	80.8	90.1	87.3

The quantitative results of the four network models are shown in Table 1. It can be seen from the comparison of the results in the table that the detection algorithm of the rotating box not only retains the aspect ratio and shape characteristics of the ship target, but also reduces the interference of the redundant background near the coast and reduces the difficulty of detecting the ship target near the coast. R-YOLOv3 improves the accuracy of 13.5% compared with the algorithm YOLOv3 based on the horizontal box, and obtains 12.7% and 16.6% gains respectively compared with Faster R-CNN and SSD. The ships in SSDD data set have the characteristics of multi-direction. The grid on each feature map extracted by the R-YOLOv3 model predicts the boundary box at different angles, which will have different effects on the detection model. The detection results of different angle intervals are shown in Table 2. Since the ship ‘ s orientation angles are mostly near the SSDD dataset, the model has better performance when the angle interval is set. There are fewer horizontal and vertical ship targets in the data set, so when the angle parameter is set to be, the accuracy of the predicted boundary box information is low, and the detection performance of the model is poor.

Table 2. Detection result of different angle intervalls

Angle	Precision	Recall	AP
30	81.2	88.0	84.2
45	80.8	90.1	87.3
60	80.0	86	82.7
90	80.7	84.8	80.2

3.4 Detection Results of Different Backbone Networks

In order to test the advantages and disadvantages of different backbone networks and the influence on the detection performance of the model, this section completed the R-YOLOv3 detection model based on the extraction characteristics of different backbone

networks. The extracted backbone networks mainly include DarkNet-53, MobileNetv2, ResNet-50 and ResNet-101.

In addition, the experiment also verifies the ability of the focus loss function to solve the imbalance problem of positive and negative samples and the classification problem of difficult samples.

Table 3. Comparison of evaluation metrics of different backbone networks

Backbone	Loss	Precision	Recall	AP
DarkNet-53	CE	70.7	84.3	81.7
DarkNet-53	Focal Loss	73.2	89.8	85.3
MobileNetv2	Focal Loss	63.1	87.0	81.8
ResNet-50	Focal Loss	80.8	90.1	87.3
ResNet-101	Focal Loss	83.7	90.8	87.8

According to the results shown in Table 3, it can be concluded that in the R-YOLOv3 network model based on the same backbone network DarkNet-53, the focus loss function reduces the probability of negative samples by improving the balance of positive and negative samples, and accelerates the convergence rate of the model. At the same time, it enhances the learning ability of the model for SAR ship samples with complex backgrounds, especially for offshore targets, so that the model is effectively trained. From the evaluation index, the AP value of the R-YOLOV3 detection model using the focus loss function is nearly 4% higher than that of the model based on the cross entropy function, and more accurate results are obtained in the detection accuracy. The recall rate of the detection model is greatly improved, and the leakage rate of the ship target in the complex background is reduced, indicating the effectiveness of the focus loss function. In the R-YOLOv3 network model based on different backbone network extraction features, compared with the lightweight network MobileNetv2, the multi-scale fusion detection structure of DarkNet-53 combines the low-level features and high-level features well, enhances the feature expression ability, improves the detection ability of the model for small ship targets, and obtains a 4.7% gain. The R-YOLOv3 framework based on the residual network ResNet-50 and ResNet-101 shows that the expansion of network depth makes the model easier to train and learn the target characteristics. ResNet50 and ResNet101 improve the detection performance by 2% and 2.5% respectively compared with the model based on DarkNet-53.

4 Conclusion

The characteristics of large length and width ratio and direction diversity of ships in the sar image make it possible to detect the rotating frame. This paper presents an improved YOLOv3 algorithm for SAR ship detection. In order to improve the accuracy of small target detection, ResNet-50 is selected as the feature extraction network of YOLOv3. In

view of the large aspect ratio and background interference of SAR ships, the boundary box prediction with angle information is introduced. The improved rotational box cross-parallel ratio strategy is adopted to reduce the influence of angle parameters, and the Focal Loss function is used to solve the imbalance between positive and negative samples exacerbated by the anchor box and the difficulty of sample learning. The experimental results show that the average accuracy of the improved YOLOv3 algorithm based on the rotational box on the SAR ship data set has been improved to some extent, but there are still missing and false positives for the difficult samples in the coastal area. In the future, the detection algorithm needs to be improved to improve the detection precision.

Acknowledgments. This work is supported by the National Key Research and Development Program of China under Grant 2018AAA0102702; the National Natural Science Foundation of China (62001137); the Natural Science Foundation of Heilongjiang Province (JJ2019LH2398); the Fundamental Research Funds for the Central Universities (3072020CFT0801).

References

1. Liu, L., Ouyang, W., Wang, X., et al.: Deep learning for generic object detection: a survey. Int. J. Comput. Vis. **128**(2), 261–318 (2020)
2. Su, H., Wei, S., Liu, S., et al.: HQ-ISNet: high-quality instance segmentation for remote sensing imagery. Remote Sens. **12**(6), 989 (2020)
3. Yadav, S., Biswas, M.: Threshold-based clustering of SAR image using gaussian kernel and mean-shift methods (2019)
4. Zhang, Q., Wu, Y., Zhao, W., et al.: Multiple-scale salient-region detection of SAR image based on gamma distribution and local intensity variation. IEEE Geosci. Remote Sens. Lett. **11**(8), 1370–1374 (2014)
5. Ressel, R., Lehner, S.: Texture-based sea ice classification on TerraSAR-X imagery. In: Proceedings of the 22 IAHR International Symposium on ICE 2014 (IAHR-ICE 2014) (2014)
6. Eltoft, T., Doulgeris, A., Anfinsen, S.N.: Model-based statistical analysis of PolSAR data. In: IEEE International Geoscience & Remote Sensing Symposium. IEEE (2009)
7. Robey, F.C., Fuhrmann, D.R., Kelly, E.J., et al.: A CFAR adaptive matched filter detector. IEEE Trans. Aerosp. Electron. Syst. **28**(1), 208–216 (1992)
8. Wu, Q., Li, Y., Lin, Y., et al.: Weighted sparse image classification based on low rank representation. Comput. Mater. Continua **56**(1), 91–105 (2018)
9. Tu, Y., Lin, Y., Hou, C., et al.: Complex-valued networks for automatic modulation classification. IEEE Trans. Vehicular Technol. (99), 1 (2020)
10. Lin, Y., Tu, Y., Dou, Z.: An improved neural network pruning technology for automatic modulation classification in edge devices. IEEE Trans. Veh. Technol. **69**(5), 5703–5706 (2020)
11. Yu, J., Hu, A., Li, G., et al.: A robust RF fingerprinting approach using multi-sampling convolutional neural network. IEEE Internet Things J. (99), 1
12. Ren, S., He, K., Girshick, R., et al.: Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal. Mach. Intell. **39**(6) (2015)
13. Qinggang, W., Xueming, Z.: Remote sensing object detection via an improved YOLO network. Int. J. Performability Eng. **16**(11), 1803 (2020)
14. Redmon, J., Farhadi, A.: YOLOv3: an incremental improvement. In: IEEE Conference on Computer Vision and Pattern Recognition, Utah: arXiv preprint: 1804. 0276 (2018)
15. Liu, W., Anguelov, D., Erhan, D., et al.: SSD: single shot multibox detector. In: European Conference on Computer Vision **6**, 21–27 (2016)

16. Lin, T.Y., Goyal, P., Girshick, R., et al.: Focal loss for dense object detection. IEEE Trans. Pattern Anal. Mach. Intell. **99**, 2999–3007 (2017)
17. Kang, M., Ji, K., Leng, X., et al.: Contextual region-based convolutional neural network with multilayer fusion for SAR Ship detection. Remote Sens. **9**(8), 860 (2017)
18. Li, J., Qu, C., Shao, J.: Ship detection in SAR images based on an improved faster R-CNN. Sar in Big Data Era: Models, Methods & Applications. IEEE (2017)
19. Kang, M., Leng, X., Lin, Z., et al.: A modified faster R-CNN based on CFAR algorithm for SAR ship detection. In: 2017 International Workshop on Remote Sensing with Intelligent Processing (RSIP). IEEE (2017)
20. He, K., Zhang, X., Ren, S., et al.: Deep residual learning for image recognition. In: IEEE Conference on Computer Vision & Pattern Recognition. IEEE Computer Society (2016)
21. Lin, T.Y., Dollar, P., Girshick, R., et al.: Feature pyramid networks for object detection. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society (2017)
22. Lin, T.Y., Goyal, P., Girshick, R., et al.: Focal loss for dense object detection. IEEE Trans. Pattern Anal. Mach. Intell. (99), 2999–3007 (2017)
23. Rampun, A., López-Linares, K., Morrow, P.J., et al. Breast pectoral muscle segmentation in mammograms using a modified holistically-nested edge detection network. Med. Image Anal. 57 (2019)