Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Marine target detection based on Marine-Faster R-CNN for navigation radar plane position indicator images*

Xiaolong CHEN†‡, Xiaoqian MU, Jian GUAN†, Ningbo LIU, Wei ZHOU

Marine Target Detection Research Group, Naval Aviation University, Yantai 264001, China

†E-mail: cxlcxl1209@163.com; guanjian_68@163.com

Received Nov. 6, 2020; Revision accepted Dec. 23, 2020; Crosschecked Dec. 8, 2021

Abstract: As a classic deep learning target detection algorithm, Faster R-CNN (region convolutional neural network) has been widely used in high-resolution synthetic aperture radar (SAR) and inverse SAR (ISAR) image detection. However, for most common low-resolution radar plane position indicator (PPI) images, it is difficult to achieve good performance. In this paper, taking navigation radar PPI images as an example, a marine target detection method based on the Marine-Faster R-CNN algorithm is proposed in the case of complex background (e.g., sea clutter) and target characteristics. The method performs feature extraction and target recognition on PPI images generated by radar echoes with the convolutional neural network (CNN). First, to improve the accuracy of detecting marine targets and reduce the false alarm rate, Faster R-CNN was optimized as the Marine-Faster R-CNN in five respects: new backbone network, anchor size, dense target detection, data sample balance, and scale normalization. Then, JRC (Japan Radio Co., Ltd.) navigation radar was used to collect echo data under different conditions to build a marine target dataset. Finally, comparisons with the classic Faster R-CNN method and the constant false alarm rate (CFAR) algorithm proved that the proposed method is more accurate and robust, has stronger generalization ability, and can be applied to the detection of marine targets for navigation radar. Its performance was tested with datasets from different observation conditions (sea states, radar parameters, and different targets).

Key words: Marine target detection; Navigation radar; Plane position indicator (PPI) images; Convolutional neural network (CNN); Faster R-CNN (region convolutional neural network) method

https://doi.org/10.1631/FITEE.2000611 **CLC number:** TN957.51

1 Introduction

Marine target detection is an important research branch in the field of target detection, as well as an important part of radar signal processing. It is of great significance for maritime transportation, marine environment monitoring, and small target searching and rescue at sea (Daniels, 2010; Yavari et al., 2016). As one of the main methods of target detection, radar has advantages of working continuously in all weathers, and plays a vital role in marine target detection alongside infrared detection and visible light detection (Guan et al., 2012; Chen et al., 2014). However, there are still many technical difficulties in detecting marine targets, such as the complicated marine environment, low echo signal-to-noise/clutter ratio (SNR/SCR), and complex non-uniform backgrounds (Ward et al., 2007; Chen et al., 2015). On one hand, small reflection coefficients of a marine target may result in weak target returns with a relative low SCR, which easily cause missed detection (Trunk and George, 1970). On the other hand, sea spikes generated under

‡ Corresponding author

* Project supported by the Shandong Provincial Natural Science Foundation, China (No. ZR2021YQ43), the National Natural Science Foundation of China (Nos. U1933135 and 61931021), and the Major Science and Technology Project of Shandong Province, China (No. 2019JZZY010415)

ORCID: Xiaolong CHEN, https://orcid.org/0000-0002-1040-1655; Xiaoqian MU, https://orcid.org/0000-0002-5956-3356

high sea states are a kind of strong sea clutter. The intensity of sea spikes is quite similar to that of a target echo, making it easy to register false alarms. Moreover, the motion of the sea surface and the wind cause further interference to the detection of marine targets (Ward et al., 2007).

The main traditional radar marine target detection methods are constant false alarm rate (CFAR) detection and coherent integration (Jalil et al., 2016; Maresca et al., 2019). The detection performance of CFAR is limited by the background distribution model. The generalization ability to the environment is too weak to ensure the stability and high accuracy of detection. The coherent integration cannot match accurately the complex target characteristics and is therefore not versatile. Moreover, traditional radar target detection and classification processes are complex, requiring detection before classification, which is time-consuming. Given that the existing target detection algorithms are not capable of adapting to complex background environments and various target characteristics, new research methods need to be developed to resolve this problem (Tian et al., 2016; Yu et al., 2016). In recent years, with the improvement of information technology, the convolutional neural network (CNN) has been well developed and applied in the field of image detection and recognition. This should provide novel research ideas for rapid, stable, and high-precision radar target detection, especially for marine target detection (Mou et al., 2019a; Wang et al., 2019).

CNN, as the most important theory of deep learning, was proposed by LeCun et al. (1998), and then developed into an important branch of machine learning. It can automatically learn and extract the features of a target, and realize tasks such as intelligent recognition and image segmentation. Since 2012, CNNs have developed rapidly and new models have been proposed, e.g., AlexNet, VGGNet (visual geometry group net), ResNet (He et al., 2016), and DenseNet. CNN-based object detection algorithms include two-stage detection algorithms such as RCNN (region convolutional neural network), Fast R-CNN, Faster R-CNN (Ren et al., 2017), and Mask R-CNN (He et al., 2020), and single-stage detection algorithms such as YOLO (you only look once), SSD (single shot multibox detector), YOLOv2, and YOLOv3. Among these, Faster R-CNN, which has high detection accuracy and good applicability, has been widely used in pedestrian detection and radar image detection (Liu et al., 2020). The most widely used radar fields are synthetic aperture radar (SAR) and inverse SAR (ISAR) image detection, especially for ship detection (Dong et al., 2019). However, SAR and ISAR images are high-resolution radar images and marine targets have significant contour shape information, which enables a high detection probability to be achieved using deep learning methods. While most radars use narrowband signals, the resolution of radar plane position indicator (PPI) images is low, coupled with complex background (e.g., sea clutter) and target characteristics, which greatly increases the difficulty of target detection. Moreover, a marine target in a radar PPI image is a small-scale target compared to the overall PPI image. When applied to this field, Faster R-CNN has the problems of a high false alarm rate and low accuracy and thus needs to be optimized to achieve better performance. While some preliminary optimization work on Faster R-CNN has been done (Mou et al., 2019b), its improvement was limited by a high false alarm rate, and its detection probability needs further improvement. Therefore, more research is needed.

In this paper, taking navigation radar PPI images as an example, a marine target detection method based on an improved Faster R-CNN algorithm, i.e., Marine-Faster R-CNN, is proposed to extract and identify targets in radar PPI images. First, the model is developed to improve the accuracy of detecting marine targets and reduce their false alarm rate, and five aspects of Faster R-CNN are optimized to meet the detection requirements of radar marine targets: a newly designed backbone network named feature fusion network (FFNet), anchor point size, multiple target detection, data sample balance, and scale normalization. Based on this, the Marine-Faster R-CNN based marine target detection algorithm is introduced. Then, JRC (Japan Radio Co., Ltd.) navigation radar is used to collect radar echo data for various observation conditions, and data preprocessing and image labeling are performed to build marine target detection datasets. The results of detection experiments show that the proposed method has better detection accuracy and generalization ability than the traditional Faster R-CNN

and CFAR methods. The main contribution of this study includes two aspects. First, we propose a CNN-based method, named Marine-Faster R-CNN, for marine target image detection of low and medium resolution radars. The method includes the FFNet for improving the feature extraction ability and reducing the false alarm rate, small-scale anchors for improving the ability to detect small targets, and the power non-maximum suppression (P-NMS) method for improving the ability to detect multiple targets. In particular, FFNet achieves target enhancement and sea clutter suppression through feature fusion when extracting image features. Second, we have constructed relatively complete marine target radar image datasets, to test the performance of the new method under different observation conditions.

2 Marine-Faster R-CNN based marine target detection model

2.1 Description of Marine-Faster R-CNN

Faster R-CNN is a popular deep learning target detection algorithm, which has three parts: a shared CNN, a region proposal network (RPN), and a classification and regression network. The shared CNN can extract target features and input the extracted feature maps into the RPN and ROI (region of interest) pooling. RPN implements candidate box generation tasks, which are combined with ROI pooling for feature mapping. The classification and regression network completes final detection and classification.

Although the Faster R-CNN algorithm has achieved good results in conventional target detection, problems of high false alarm rates and low accuracy are encountered when applying it to the field of radar marine target detection. The most serious challenges include the detection of marine targets with complex target characteristics in complex sea backgrounds, missed detection and false detection. To improve the accuracy of marine target detection and reduce the false alarm rate, in this study we propose FFNet as the backbone network of Faster R-CNN. We also improve the four other aspects of Faster R-CNN and optimize the RPN and classification and regression network, to obtain the proposed Marine-Faster R-CNN model. The structure of Marine-Faster R-CNN is shown in Fig. 1, and the backbone network FFNet in Fig. 2.

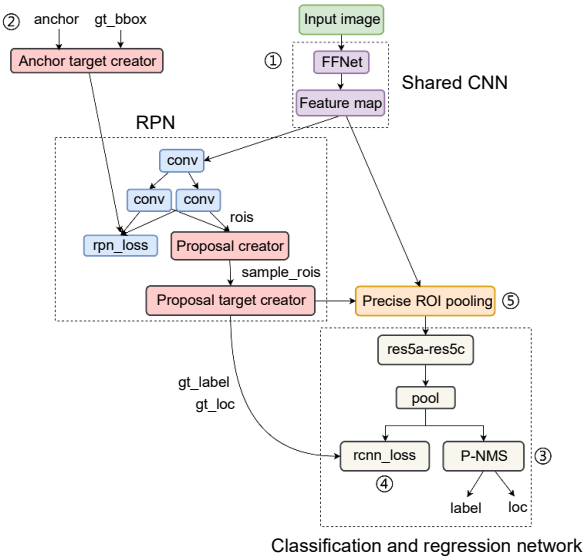

Fig. 1 The structure of Marine-Faster R-CNN

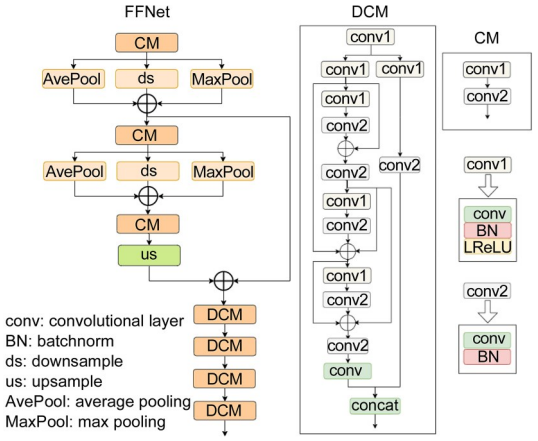

Fig. 2 The structure of FFNet

The specific optimization method of the model is as follows:

1. Improving the accuracy of detection and reducing false alarms: FFNet is proposed as the backbone network of Faster R-CNN, which improves the accuracy of feature extraction and enhances the ability to extract and transfer small target features. This part is identified by ① in Fig. 1.

The image feature extraction process of FFNet is as follows: first, convolve the input image using the convolution module (CM); then perform the first

downsampling followed by feature fusion; after that, perform the second CM; next, perform the second downsampling followed by feature fusion; and then perform the third CM; perform the first upsampling and fusion with the features after the first downsampling; finally, perform feature extraction using four dense convolutional modules (DCMs).

The feature fusion comprises four main aspects: (1) The feature map of the CM output is added and fused after downsampling, average pooling, and maximum pooling; (2) The outputs after the first downsampling and after the first upsampling are added and fused; (3) The outputs of the front layer and the back layer in the DCM are added and fused; (4) The dense convolution network and the sparse convolution network in the DCM are fused.

2. Optimizing the RPN to improve the accuracy of detecting marine targets and reduce the false alarm rate: small-scale anchors are adopted to reduce the problems of missed detection and false alarms caused by small ships in large-scale radar PPI images during the actual detection. In the detection of marine targets in PPI images, the radar image is very large, while a ship is very small, so the target occupies a small proportion of the image size. Thus, the detection of the ship in the image involves detection of a small target. However, the classic Faster R-CNN model is designed for ordinary targets, and its anchor frame size is not suitable for small target detection in large radar images. Therefore, we conduct research and cluster analysis on the anchor and redesign it to meet the detection needs for marine radar image targets. The optimization is identified by ② in Fig. 1.

The clustering using representative (CURE) algorithm is used to cluster labeled truth boxes. The CURE algorithm is an agglomerative clustering method with excellent clustering performance. It has low sensitivity to anomalous data and data input order, and good anti-interference performance. The formula for measuring the distance between truth boxes in the algorithm is as follows:

$$\operatorname{dist}(u,v) = \min \operatorname{dist}(p,q), \quad p,q \in \left\{c_1,c_2,\cdots,c_n\right\}, \tag{1}$$

where u and v are the two closest categories, that is, the two categories to be aggregated, dist(u, v) represents the distance between u and v, $c_i(c_s^i,c_r^i)$ is an arbitrary truth box from the truth box collection $\{c_1, c_2, \cdots, c_n\}$, and the total number of truth boxes is n. The two parameters c_s^i and c_r^i of c_i are the area and the ratio of the length to width of the truth box, respectively. dist(p, q) represents the distance between p and q; here we use the Euclidean distance.

After clustering, the center and representative points of the new class are obtained:

$$\bar{w} = \frac{|u| \cdot \bar{u} + |v| \cdot \bar{v}}{|u| + |v|}, \tag{2}$$

$$w_i = p_i + \alpha \cdot (\bar{w} - p_i), \tag{3}$$

where $\bar{w}$ is the center of the new class, $|u|$ and $|v|$ represent the numbers of truth boxes in classes u and v respectively, $\bar{u}$ and $\bar{v}$ represent the mean of the ground-truth boxes in u and v respectively, w_i (i=1, 2, $\cdots$, n) is the new representative point of the original class p_i after clustering, and α is the clustering regulator.

In the initial state, each truth box is a class, and then the two classes with the smallest distance among all classes are condensed into a new class. $\bar{w}$ is the center of the condensed class, and w_i is the new representative point. According to the above method, clustering continues until the final clusters are formed.

At last, nine categories are obtained after clustering, so there are still nine setting anchors. The scales of the anchors are modified to (73, 42), (64, 61), (55, 40), (37, 27), (32, 32), (28, 36), (42, 53), (47, 48), and (56, 72).

3. Improving the accuracy of detection of marine targets: the P-NMS algorithm is proposed to resolve the problem of missed detection in multiple targets areas. The optimized part is identified by ③ in Fig. 1. In actual sea surface target detection scenes, multiple targets are often encountered. Dense multiple targets are more likely to cause missed detection, which greatly limits the generalization ability of the algorithm and the further improvement of detection accuracy. P-NMS is based on improvement of the NMS algorithm. The core idea of NMS is to remove all candidate frames that overlap with the optimal candidate frame. The formula of NMS is as follows:

$$s_i = \begin{cases} s_i, & \operatorname{IOU}(M,b_i) < N_t, \\ 0, & \operatorname{IOU}(M,b_i) \geq N_t. \end{cases} \tag{4}$$

However, in the actual detection of multiple marine targets, the candidate frames of multiple targets may overlap, and the NMS algorithm suppresses all candidate frames whose overlap ratios are higher than a certain threshold. This may cause individual targets to miss detection. Considering that ships are different from other types of targets in that they should not get too close when sailing or berthing, there is no high overlap of candidate frames. Based on these considerations, we design a power function method called P-NMS, which is suitable for the detection of multiple targets with a moderate overlap in candidate frames. The formula is as follows:

$$s_i = \begin{cases} s_i, & \mathrm{IOU}(M,b_i) < N_\mathrm{t}, \\ s_i(\mathrm{IOU}(M,b_i))^2, & \mathrm{IOU}(M,b_i) \geq N_\mathrm{t}, \end{cases} \tag{5}$$

where s_i is the score of the i^{th} candidate frame, M denotes the current highest scoring frame, b_i is the frame to be processed, N_t is the detection threshold, and IOU is the ratio of the intersection to the union between the prediction frame and the labeled frame.

4. Optimizing the classification and regression networks to improve the accuracy of detecting marine targets and reduce the false alarm rate: focal loss is used as the classification loss function of the RCNN. This solves the imbalance problem on positive and negative samples, simple and difficult samples, and improves the training results. Deep learning target detection models are known to have the problem of sample imbalance. In terms of the imbalance of positive and negative samples, there will be far fewer positive samples containing marine targets than negative samples containing clutter. When the clutter intensity is strong, it will greatly aggravate the probability that some negative samples containing sea clutter are falsely detected as targets, resulting in false alarms. In terms of the imbalance of simple and difficult samples, if there are too many simple samples which are easily detected, the model will be overfitted, which reduces the generalization ability and accuracy of the model for marine target detection. Therefore, we optimize the sample imbalance problem to reduce the false alarm rate of marine target detection, enhance the model's generalization ability and accuracy for marine target detection, and improve the detection performance of the model. The optimization is identified by ④ in Fig. 1. Its formula is as follows:

$$\mathrm{FocalLoss}(p_\mathrm{t}) = -\alpha_\mathrm{t}(1-p_\mathrm{t})^\gamma \ln p_\mathrm{t}, \tag{6}$$

where α_t is the balance factor, γ is the modulation coefficient, and p is the probability of the training results being correctly classified, i.e.,

$$p_\mathrm{t} = \begin{cases} p, & y = 1, \\ 1-p, & y = -1. \end{cases} \tag{7}$$

In the above, $y \in \{\pm 1\}$ specifies the ground-truth class, which indicates whether it is in this class.

Focal loss is based on the cross entropy function. The formula of the cross entropy function is

$$\mathrm{CrossEntropy}(p_\mathrm{t}) = -\ln p_\mathrm{t}. \tag{8}$$

Focal loss adds modulation coefficients γ, so that the loss value of easy-to-separate samples in the total classification loss is reduced (too many easy-to-separate counter examples will lead to invalid learning and detection of only the background), while increasing the loss of difficult-to-separate samples over the proportion of the total classification loss value. This ensures the useful training of the difficult samples on the model. A balance factor α_t is added to resolve the problem of the uneven proportion of positive and negative samples.

5. Optimizing classification and regression networks to improve the accuracy of detection of marine targets: precise ROI pooling is used as the scale normalization function to improve the accuracy of the quantization operation, reduce the accuracy loss of the scale normalization process, and improve the integrity and accuracy of the features input to the classification and regression network. In marine target detection based on navigation radar PPI images, marine targets are small targets relative to the entire image. Achieving accurate detection requires a particularly high accuracy of feature extraction. ROI pooling involves normalizing the ROI of the RPN and the feature extraction network, and its accuracy directly affects the classification and regression network. Therefore, we optimize ROI pooling to improve its accuracy. This will improve the detection accuracy of the entire model for small marine

target detection. The optimization is identified by ⑤ in Fig. 1. The formula is shown as follows:

$$\mathrm{PrPool}(\mathrm{bin}) = \frac{\int_{y_1}^{y_2}\int_{x_1}^{x_2} f(x,y)\,\mathrm{d}x\mathrm{d}y}{(x_2 - x_1)(y_2 - y_1)}, \tag{9}$$

where

$$f(x,y) = \sum_{i,j} \mathrm{IC}(x,y,i,j)\,w_{i,j}, \tag{10}$$

$$\mathrm{IC}(x,y,i,j) = \max(0,1 - |x - i|)\max(0,1 - |y - j|), \tag{11}$$

where (x_1, y_1) and (x_2, y_2) are the upper-left corner and lower-right corner of one bin area, respectively (a bin is the area divided before pooling). The IC function calculates the product of each successive x- and y-direction offset, and the deviation is within one pixel, using pixel (i, j) as representation. Then multiplying the product of the offset by the pixel value $w_{i,j}$ of (i, j) obtains $f(x, y)$. Perform integration $f(x, y)$ from (x_1, y_1) to (x_2, y_2) to obtain the sum of pixels in the entire bin areas, and take the average of the integral results to obtain the output of the bin area. Finally, each bin area outputs a value to form a feature map matrix. Precise ROI pooling makes the pixel values have a gradient transfer, which improves the accuracy of pooling.

2.2 Detection flowchart using the Marine-Faster R-CNN model

A flowchart of the radar image marine target detection algorithm based on Marine-Faster R-CNN is shown in Fig. 3. The flowchart includes two stages: the offline training stage and the actual detection stage. The offline training procedure is as follows:

1. Use navigation radar to transmit long and short pulses with different sampling frequencies to collect radar echo data under different conditions (different weather conditions and sea states) to ensure the diversity of the collected data samples. At the same time, collect complex samples (samples that are not easily detected) to improve the generalization performance of the dataset in different situations.

2. Convert the range-azimuth raw echo data obtained by the radar scan into PPI images.

3. Pre-process the images, e.g., crop and label the images to build a marine target dataset.

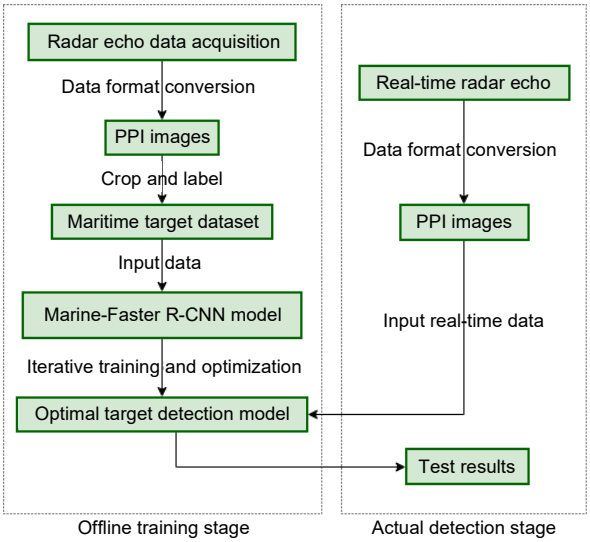

Fig. 3 Flowchart of the marine target detection algorithm

4. The Faster R-CNN model is optimized in five respects to be more suitable for radar detection, and the Marine-Faster R-CNN model is obtained.

5. Input the dataset into the Marine-Faster R-CNN model for training and optimization.

6. Adjust the initial training parameters of the model and perform iterative optimization, and select the model with the best training result as the final marine target detection model for detection and classification tasks.

The actual detection stage is as follows:

1. Collect real-time navigation radar data.

2. Convert the original data into PPI images.

3. Input the PPI images to the Marine-Faster R-CNN, and finally the detection results are obtained.

3 Construction of a radar image dataset for marine target detection

3.1 Data collection and sampling

X-band non-coherent JRC-5312 navigation radar was used to transmit electromagnetic signals, and data acquisition experiments were conducted near Yantai Port, China, to obtain radar returns of ships under different sea states and weather conditions. Radar returns were stored as datasets in DAT format. The marine environment of the Yantai Port area is shown in Fig. 4. The area is close to the main international

waterway and is located in the core area of the Northeast Asia International Economic Circle.

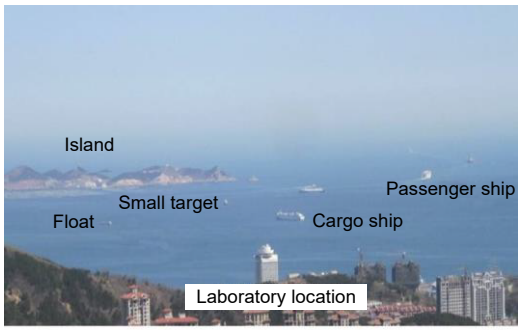

Fig. 4 Yantai Port sea area

Not only are there a lot of passenger and cargo ships entering and leaving the port every day, but also fish breeding areas around and fishing boats, all of which can provide great convenience for the collection of marine targets. The data collection process is shown in Fig. 5. After the electromagnetic waves are emitted by the radar antenna, the echoes of ships are collected by the radar signal acquisition equipment, and then output to the terminal display to obtain the PPI images after signal processing. The position information of the ships obtained by the automatic identification system (AIS) device can be displayed in the PPI images as verification of the target detection results. The acquisition software interface and real-time PPI are shown in Fig. 6, in which the land clutter, sea clutter, and ship's echo are displayed. The parameters of the JRC navigation radar are shown in Table 1.

Radar datasets were collected under various observation conditions including different radar parameters, weather conditions, and sea states, and different types of maritime targets, to ensure the diversity of the collected data samples (Table 2). To improve the generalization ability of the dataset to complex environments, and ameliorate the imbalance between the numbers of simple and difficult samples, it is necessary to collect difficult samples. Difficult samples are samples that cannot be easily distinguished by classifiers. Such samples help improve the recognition ability and detection performance of the training model. Therefore, the collection of target samples in complex weather such as rain and snow, high wind and wave weather, and small-scale target samples,

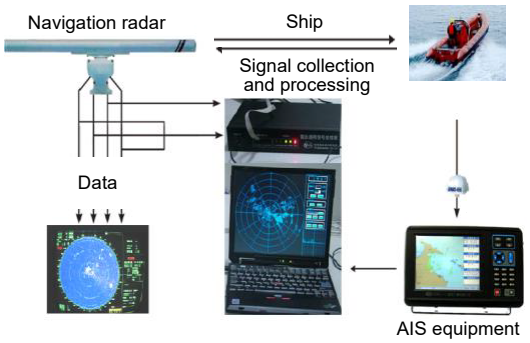

Fig. 5 Data collection diagram

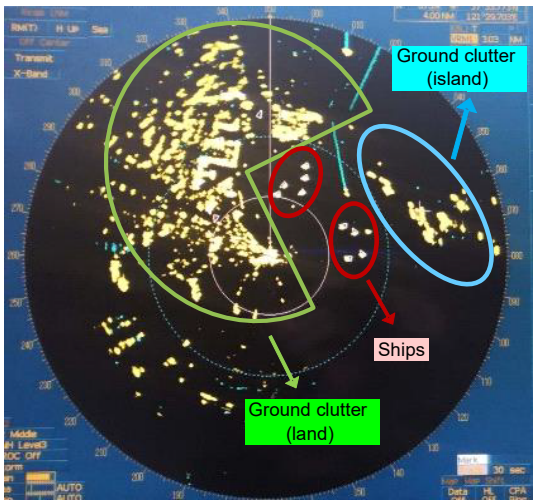

Fig. 6 Real-time radar PPI images

Table 1 JRC navigation radar parameters

Parameter	Value
Name	JRC navigation radar
Type	JMA-5312
Range scope	0.125–48 n mile
Transmit power	10 kW
Working frequency	9410±30 MHz
Rotation speed	24 r/min

Table 2 Specific implementation of sample diversity

Data type	Description/Value
Radar parameters	
Pulses	Long, short
Sampling frequency	30, 60 MHz
Environments	
Weather	Sunny, cloudy, rainy, snowy
Sea states	First, second, third, fourth
Marine targets	
Sizes	Big, small
Number	Single, dual, multiple
Types	Passenger ship, cargo ship, fishing boat, etc.

can not only further increase the diversity of data samples, but also resolve the problem of imbalance between the numbers of simple and difficult samples. After the data collection was completed, the collected content was recorded (Fig. 7).

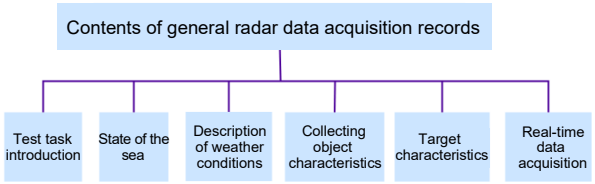

Fig. 7 Content of data acquisition records

3.2 Conversion of data form

Matlab was used to convert the DAT format raw data collected by the PCI (peripheral component interconnect) card to MAT format, and to read the format data to generate PPI images. Samples of ship targets under various background environments were collected through various collection methods, and the diversity of the datasets was guaranteed. A typical background environment image and a part of the datasets for different pulses and frequencies, numbers of targets, sea states, and complex samples are shown in Figs. 8 and 9.

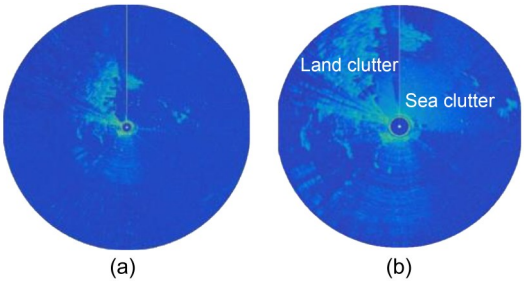

Fig. 8 Normal sample PPI (clear weather) (a) and difficult sample PPI (complex weather background) (b)

3.3 Data preprocessing (cropping)

Because the PPI image is large (greater than 4000×4000 pixels) and there are no targets in most areas, training with the whole PPI image may take up too much memory and too much time, so the original PPI images need to be cropped. Fig. 6 shows that the marine targets are located mainly in the upper right part of the PPI image, that is, in the sea area east of the Yantai Port. Therefore, this part was cropped and the training and testing datasets were constructed based on this PPI image. An example of the cropped area images is shown in Fig. 10.

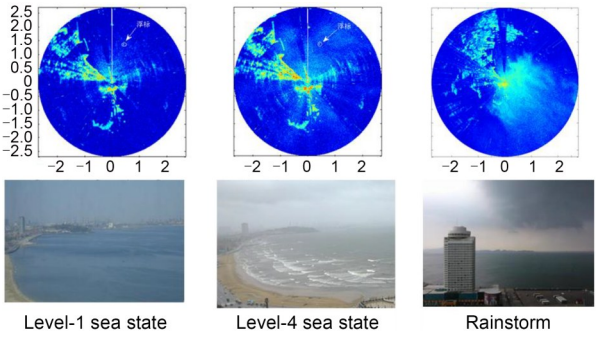

Fig. 9 Comparison of different background environments

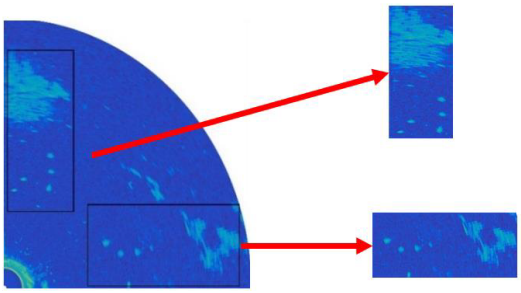

Fig. 10 An example of the cropped area image

3.4 Image annotation for the dataset

We used Pascal VOC (visual object classes) format to make the dataset. First, use the software to annotate the image and generate an annotation file in XML format; then, set 80% of the images in the dataset for training, 20% for testing and verification. The four text files, trainval.txt, train.txt, val.txt, and test.txt, serve as an index directory for the training validation, training, validation, and test sets. The training dataset includes 2000 images and the test dataset 500 images. A diagram showing the use of annotation software to annotate the cropped sub-image is shown in Fig. 11, in which the five rectangular boxes are the truth boxes after labeling.

4 Training and experimental results analysis

In this section, the training results of the Marine-Faster R-CNN model trained on the surface target

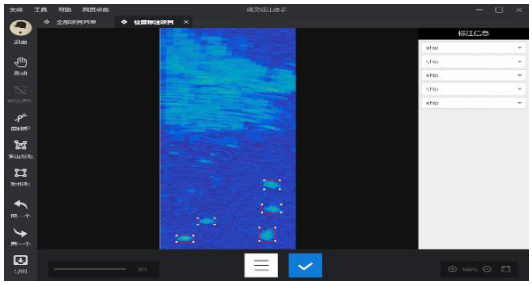

Fig. 11 Diagram of image annotation

dataset are shown first. Then, five evaluation indexes were selected based on the evaluation criteria of radar target detection and computer image detection. These indexes were used to analyze and evaluate the experimental results. Finally, detection results are compared with the classic Faster R-CNN and CFAR.

The computer configuration was as follows: Intel® Core™ i7-8700 K processor, 16 GB memory, GeForce GTX 1080 Ti graphics card with 11 GB memory. The system environment was Ubuntu 18.04 and the framework was Pytorch. The experimental parameters were set as follows: the gradient descent algorithm was the Adam, the initial learning rate was 0.001, and the number of iterations was set to 20 000.

4.1 Model training and analysis of the results

The shared CNN (backbone network) model is essentially an extractor of target features. We proposed FFNet as the backbone network for the Marine-Faster R-CNN model. Fig. 12 shows the loss value of the training, which consists of the classification loss (rpn-cls-loss) and regression loss (rpn-bbox-loss) of the RPN, and the classification loss (rcnn-cls-loss) and regression loss (rcnn-bbox-loss) of the RCNN. It can be seen that the training loss value stabilized at about 0.1 as the number of iterations increased.

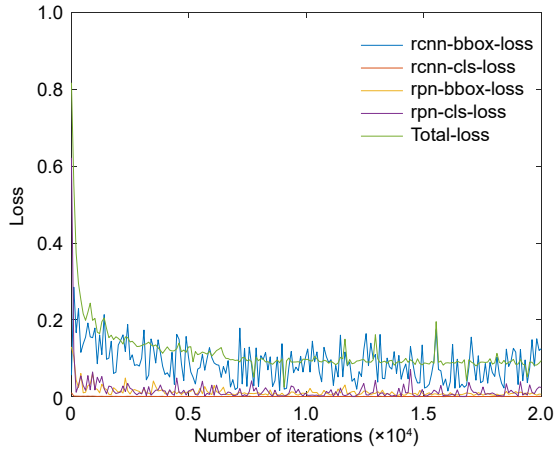

Fig. 12 Training loss curves of FFNet
References to color refer to the online version of the figure

Further analysis and evaluation of the training and detection results were carried out. Combining the evaluation parameters of deep learning target detection and traditional radar target detection, three evaluation parameters including the recall rate, accuracy rate, and false alarm rate were employed, and the pros and cons of the model and detection results were compared accordingly. The three evaluation parameters are as follows:

$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \tag{12}$$

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \tag{13}$$

$$\mathrm{FA} = 1 - P, \tag{14}$$

where R stands for recall, which indicates the recall rate, P denotes precision, which indicates the accuracy rate, and FA is the false alarm rate. False positive (FP) indicates a negative sample predicted to be positive, false negative (FN) indicates a positive sample predicted to be negative, and true positive (TP) indicates a positive sample predicted to be positive.

An example of target detection results is shown in Fig. 13. In this image, there were five ships near

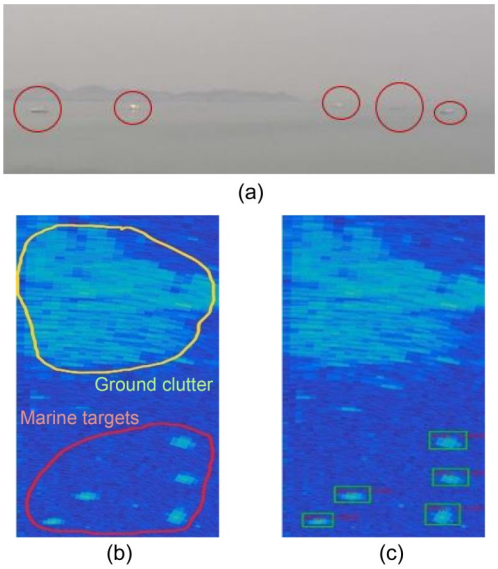

Fig. 13 Targets in real scenes (a), targets in PPI images (b), and detection result images (c)

an island in the low sea state in the Yantai Port. The proposed model could recognize all five ships and achieve good detection performance.

4.2 Verification of the effect of optimizations

1. The effect of the focal loss function. An experiment was performed to verify how the focal loss function improves the target detection results of the original model. The Faster R-CNN model, using FFNet as the backbone network, was compared with another sample balance optimization method, the online hard example mining (OHEM) algorithm. The results are shown in Table 3, where the frame rate is the number of images (frames) recognized in 1 s. The OHEM algorithm is also a method for optimizing sample imbalances. The core principle is similar to that of the boosting algorithm of machine learning. Samples that are easily misclassified by the classifier are repeatedly trained to improve the model. The ratio of positive to negative samples was 1: 3. This fixed ratio led to poor generalization ability, which limits the improvement of detection accuracy.

Table 3 The results of different sample balance optimizations

Method	Recall	Frame rate (frame/s)
Faster R-CNN (FFNet)	91.34%	3.95
Faster R-CNN (FFNet)+OHEM	91.70%	3.78
Faster R-CNN (FFNet)+focal loss	91.96%	4.02
Faster R-CNN (FFNet)+Soft-NMS	91.82%	3.92
Faster R-CNN (FFNet)+P-NMS	91.90%	3.94
Faster R-CNN (FFNet)+ROI alignment	91.72%	3.95
Faster R-CNN (FFNet)+precise ROI pooling	91.85%	3.94
Faster R-CNN (FFNet)+small anchors	92.01%	3.99
Marine-Faster R-CNN (all)	93.65%	3.97

The OHEM algorithm achieved an improvement in the probability of detection, but caused a decrease in the detection speed (Table 3). Our method with focal loss takes both the probability of detection and speed into account: the probability of detection was increased by 0.62%, and the detection speed for each picture decreased by 0.07 frames/s, thereby achieving improvements in both aspects.

2. The effect of P-NMS. Table 3 shows that P-NMS improved the probability of detection more than the NMS algorithm. The 0.56% improvement achieved by P-NMS was higher than that achieved by Soft-NMS. The P-NMS algorithm is therefore more suitable for the detection of dense multi-targets on the sea surface.

3. The effect of precise ROI pooling. Table 3 shows that with the refinement of the normalization of the ROI region scale, the probability of detection of the model was gradually improved. For example, compared with the ROI pooling algorithm, the probability of detection of precise ROI pooling was increased by 0.51%. However, the increased complexity of the algorithm affects the speed of training.

4. The effect of small anchors. Table 3 shows that small-scale anchors increased the detection probability by 0.67% compared with the original anchors. This illustrates the importance of anchor point scale adaptation to the detection target scale, and proves the effectiveness of the optimization.

To sum up, Table 3 shows that small-scale anchors made the largest contribution to the optimization of the Faster R-CNN algorithm, increasing the detection probability by 0.67%, followed by focal loss with a contribution of 0.62%, and then P-NMS, 0.56%, and finally precise ROI pooling, 0.51%. The total contribution was 2.36%. However, the overall detection performance improvement was 2.31%, which shows that the improvements by the four optimization methods are not a simple summation of the individual's. In general, the changes had a good effect on the optimization of Faster R-CNN.

4.3 Comparisons with classic Faster R-CNN

To verify the performance of the model proposed in this paper, we compared it with the classic Faster R-CNN algorithm. In the experiment, the training set of the marine target dataset was used as the input data for training, and the test set was used for testing and verification. The test results are shown in Table 4, and the verification results in Figs. 14–18. The experiments showed that the performance of the Marine-Faster R-CNN algorithm was better than that of the classic Faster R-CNN algorithm, with higher accuracy and a lower false alarm rate, and was suitable for marine target detection in complex backgrounds.

Fig. 14 shows that Faster R-CNN (VGG16), Faster R-CNN (Resnet101), and Marine-Faster R-CNN all

Table 4 Comparison of test results

Method	Recall	Precision	False alarm	Frame rate (frame/s)
Faster R-CNN (VGG16)	89.77%	97.34%	2.66%	6.32
Faster R-CNN (Resnet101)	92.59%	97.46%	2.54%	2.37
Marine-Faster R-CNN	93.65%	99.02%	0.98%	3.97

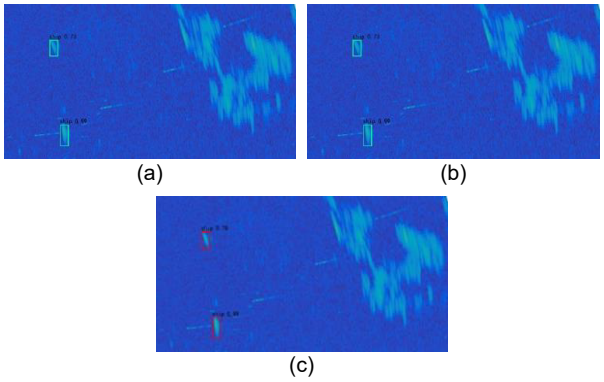

Fig. 14 Results under the level-2 sea state for small targets: (a) Faster R-CNN (VGG16); (b) Faster R-CNN (Resnet101); (c) Marine-Faster R-CNN

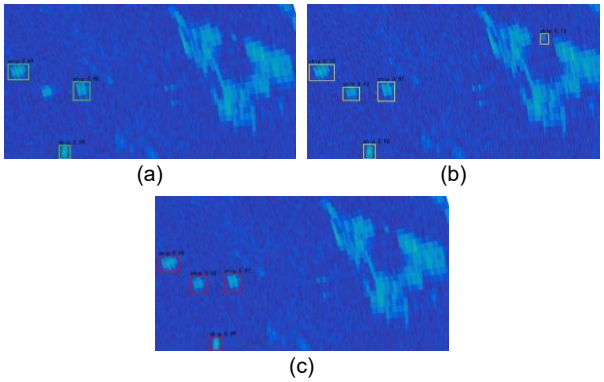

Fig. 15 Results under the level-2 sea state for large and medium multiple targets: (a) Faster R-CNN (VGG16); (b) Faster R-CNN (Resnet101); (c) Marine-Faster R-CNN

had good detection results for sparse targets in a simple background. Fig. 15 shows that Faster R-CNN (VGG16) missed detections while Faster R-CNN (Resnet101) gave false alarms for multi-target detection under the level-2 sea state. Figs. 16 and 17 show that Faster R-CNN (VGG16) and Faster R-CNN (Resnet101) were more likely to cause false alarms than Marine-Faster R-CNN for multi-target detection

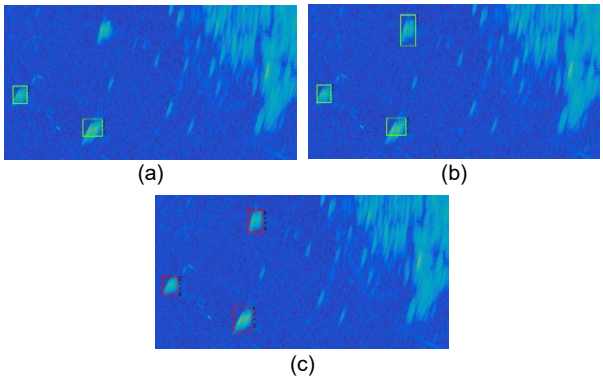

Fig. 16 Results under the level-3 sea state for large multiple targets: (a) Faster R-CNN (VGG16); (b) Faster R-CNN (Resnet101); (c) Marine-Faster R-CNN

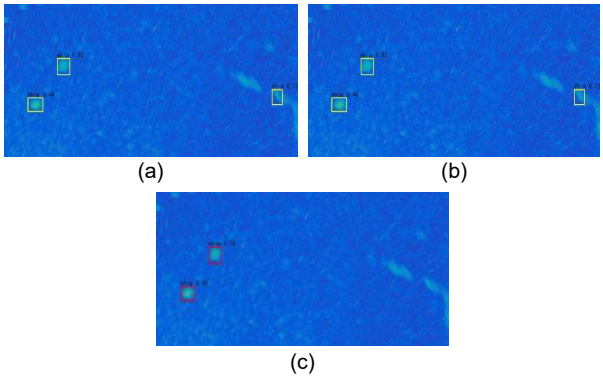

Fig. 17 Results under the level-3 sea state for medium targets: (a) Faster R-CNN (VGG16); (b) Faster R-CNN (Resnet101); (c) Marine-Faster R-CNN

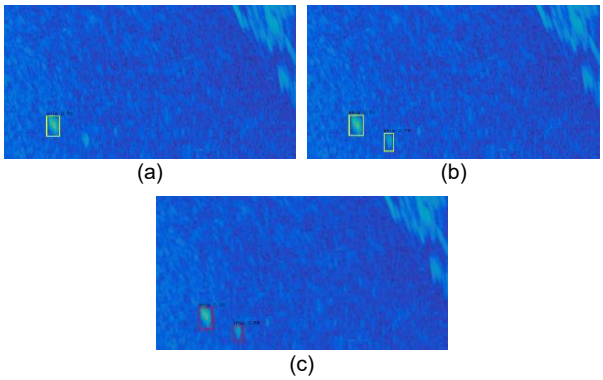

Fig. 18 Results under the level-4 sea state for large and small targets: (a) Faster R-CNN (VGG16); (b) Faster R-CNN (Resnet101); (c) Marine-Faster R-CNN

under the level-3 sea state, and Faster R-CNN (VGG16) was more likely to have missed detection. Fig. 18 shows that Faster R-CNN (VGG16) missed detection for target detection under the level-4 sea state, while

Faster R-CNN (Resnet101) and Marine-Faster R-CNN had good detection performance.

In summary, through the proposal of a new backbone network named FFNet and the adaptation of small-scale anchor points, the improved balance of samples, the refined scale normalization, and the new algorithm for NMS, Marine-Faster R-CNN showed a great improvement in performance compared with the classic Faster R-CNN. The Marine-Faster R-CNN method has a higher detection probability and a lower false alarm rate. The new method has good generalization ability for different environments and is better than the classic Faster R-CNN.

4.4 Comparisons with CFAR algorithms

To further verify the effectiveness of the proposed model, we compared it with traditional target detection methods, including the dual-parameter CFAR detection method and the classic two-dimensional cell averaging constant false alarm rate (CA-CFAR) methods widely used in radar image detection. The core idea of CFAR detection is to judge whether there is a target in the detection unit by use of a decision threshold, which is estimated by the power level of the clutter background. A block diagram is shown in Fig. 19. Taking CA-CFAR as an example, select N reference units x_1, x_2, $\cdots$, x_N, and obtain the total clutter background power level Z. Then multiply Z by the background threshold decision factor T to determine the presence or absence of the target. The judgment criterion formula is as follows:

$$\begin{cases} \mathrm{H}_1, & D \geq TZ, \\ \mathrm{H}_0, & D < TZ, \end{cases} \tag{15}$$

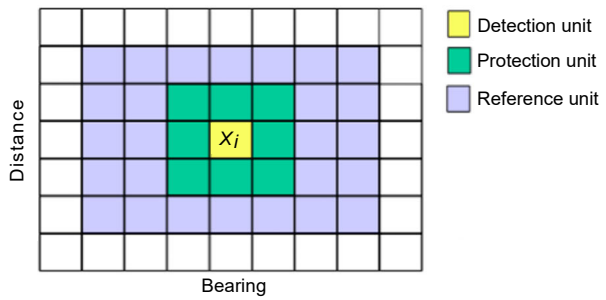

Fig. 19 The diagram of two-dimensional cell averaging constant false alarm rate (CA-CFAR)
References to color refer to the online version of the figure

where D is the detection unit, H_1 indicates that a target exists in the detection unit, H_0 indicates that there is no target in the detection unit, and T is a threshold decision factor.

The core idea of dual-parameter CFAR is to use the mean value and variance of the reference unit in the rectangular reference window to estimate the clutter power level, and to obtain the decision threshold to detect the detection unit. The figure of the reference window of the dual-parameter CFAR is shown in Fig. 20.

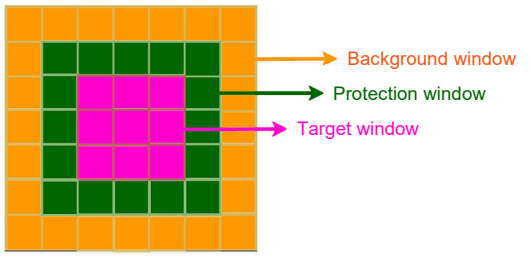

Fig. 20 Dual-parameter constant false alarm rate (CFAR) reference window

Select the target window, protection window, and background window. m_{A} and m_{B} are the mean values of all pixels in the target window and the background window, respectively, and σ_{B}^2 is the variance of the background unit in the background window. The decision criterion formula is as follows:

$$\begin{cases} \mathrm{H}_1, & \dfrac{m_{\mathrm{A}} - m_{\mathrm{B}}}{\sigma_{\mathrm{B}}} \geq K, \\ \mathrm{H}_0, & \dfrac{m_{\mathrm{A}} - m_{\mathrm{B}}}{\sigma_{\mathrm{B}}} < K, \end{cases} \tag{16}$$

where K is the threshold decision factor.

Assume that the clutter distribution type is a Rayleigh distribution. In the two-dimensional CA-CFAR, the number of reference units for the upper, lower, left, and right of the detection unit was set to 20, and the number of protection units for the upper, lower, left, and right of the detection unit was set to 2. In the dual-parameter CFAR, the length and width of the target window was 25, the protection window was 10, and the reference window was 50. The detection results of dual-parameter CFAR and CA-CFAR with the probabilities of false alarm of 10^{-4}, 10^{-3}, and 10^{-2}

were selected for comparison with the Marine-Faster R-CNN algorithm. The test results are shown in Table 5.

Table 5 Comparison with CFAR

Method	Probability of detection		
	Pfa=10^{-4}	Pfa=10^{-3}	Pfa=10^{-2}
Two-dimensional CA-CFAR	19.16%	39.24%	62.51%
Dual-parameter CFAR	18.23%	42.88%	64.07%
Marine-Faster R-CNN	–	90.52%	93.71%

Table 5 shows that the detection performance of the CFAR algorithm was unstable and greatly affected by the setting of the probability of false alarm (Pfa). The reason is that it is limited by the clutter distribution model of the background environment, and a single clutter distribution model cannot adapt to diverse background environments. The model is susceptible to interference by ground clutter (island, land) and sea clutter when detecting the target, and the false alarm rate was high. The Marine-Faster R-CNN method proposed in this paper had a higher detection probability under the same false alarm rate, and the detection probability was less affected by the false alarm rate compared with the CFAR algorithms. The target detection results of the three methods under the level-4 sea state are shown in Fig. 21.

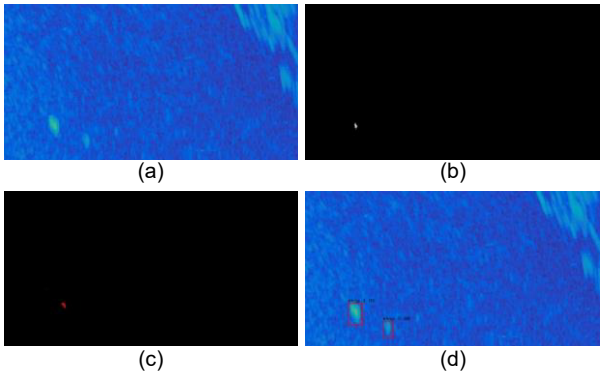

Fig. 21 The detection results under the level-4 sea state: (a) original image; (b) two-dimensional CA-CFAR; (c) dual-parameter CFAR; (d) Marine-Faster R-CNN

In summary, the method proposed in this paper is more accurate and robust. Compared with the classic Faster R-CNN detection method, it can reduce missed detection and false alarms with more robustness and generalization ability, and can be applied to the marine target detection of navigation radar. Compared with the traditional constant false alarm detection method, it can overcome the limitations of a low false alarm rate and low detection probability in complex background environments, greatly improve the detection probability under the same false alarm rate, and is more robust.

5 Conclusions

Statistical model based detection methods can achieve their best performance only under the assumed statistical model. However, it is difficult to adapt to diverse and complex backgrounds and multi-type target scenarios. The classic Faster R-CNN target detection algorithm has problems when dealing with low-resolution radar PPI images. Such problems include missed detection of multiple targets and rough scale normalization. A new marine radar target detection method based on the Marine-Faster R-CNN was proposed to alleviate the problem that the existing algorithms cannot adapt to complex background environments and various target characteristics. The measured data verified the validity of the new model. The main work can be summarized as follows:

1. The Faster R-CNN algorithm was optimized in five respects: a new backbone network, anchor point scale, data balance, scale normalization, and dense target detection, and a Marine-Faster R-CNN model was established.

2. By collecting data based on different radar parameters and under different weather conditions, a dataset of radar images for marine target detection was constructed.

3. Experiments with measured data proved that the Marine-Faster R-CNN has higher detection accuracy and better generalization ability than the classic Faster R-CNN and CFAR, and can achieve better marine target detection.

The proposed method should provide support for widespread applications in radar image detection. In the future, different CNN models will be used to further improve the detection performance.

Contributors

Xiaolong CHEN and Xiaoqian MU designed the research. Ningbo LIU and Wei ZHOU provided the radar data.

Xiaoqian MU processed the data. Xiaolong CHEN drafted the paper. Xiaolong CHEN, Xiaoqian MU, and Jian GUAN revised and finalized the paper.

Compliance with ethics guidelines

Xiaolong CHEN, Xiaoqian MU, Jian GUAN, Ningbo LIU, and Wei ZHOU declare that they have no conflict of interest.

References

Chen XL, Guan J, Bao ZH, et al., 2014. Detection and extraction of target with micromotion in spiky sea clutter via short-time fractional Fourier transform. *IEEE Trans Geosci Remote Sens*, 52(2):1002-1018. https://doi.org/10.1109/TGRS.2013.2246574

Chen XL, Guan J, Li XY, et al., 2015. Effective coherent integration method for marine target with micromotion via phase differentiation and radon-Lv's distribution. *IET Radar Sonar Navig*, 9(9):1284-1295. https://doi.org/10.1049/iet-rsn.2015.0100

Daniels DJ, 2010. Radar systems. In: Daniels DJ (Ed.), EM Detection of Concealed Targets. Wiley-IEEE Press, Hoboken, USA, p.164-213. https://doi.org/10.1002/9780470539859.ch5

Dong RC, Xu DZ, Zhao J, et al., 2019. Sig-NMS-based faster R-CNN combining transfer learning for small target detection in VHR optical remote sensing imagery. *IEEE Trans Geosci Remote Sens*, 57(11):8534-8545. https://doi.org/10.1109/TGRS.2019.2921396

Guan J, Chen XL, Huang Y, et al., 2012. Adaptive fractional Fourier transform-based detection algorithm for moving target in heavy sea clutter. *IET Radar Sonar Navig*, 6(5): 389-401. https://doi.org/10.1049/iet-rsn.2011.0030

He KM, Zhang XY, Ren SQ, et al., 2016. Deep residual learning for image recognition. Proc IEEE Conf on Computer Vision and Pattern Recognition, p.770-778. https://doi.org/10.1109/CVPR.2016.90

He KM, Gkioxari G, Dollár P, et al., 2020. Mask R-CNN. *IEEE Trans Patt Anal Mach Intell*, 42(2):386-397. https://doi.org/10.1109/TPAMI.2018.2844175

Jalil A, Yousaf H, Baig MI, 2016. Analysis of CFAR techniques. Proc 13th Int Bhurban Conf on Applied Sciences and Technology, p.654-659. https://doi.org/10.1109/IBCAST.2016.7429949

LeCun Y, Bottou L, Bengio Y, et al., 1998. Gradient-based learning applied to document recognition. *Proc IEEE*, 86(11):2278-2324. https://doi.org/10.1109/5.726791

Liu ZG, Lyu Y, Wang LY, 2020. Detection approach based on an improved faster RCNN for brace sleeve screws in high-speed railways. *IEEE Trans Instrum Meas*, 69(7): 4395-4403. https://doi.org/10.1109/TIM.2019.2941292

Maresca S, Bogoni A, Ghelfi P, 2019. CFAR detection applied to MIMO radar in a simulated maritime surveillance scenario. Proc 16th European Radar Conf, p.157-160.

Mou XQ, Chen XL, Su NY, et al., 2019a. Motion classification for radar moving target via STFT and convolution neural network. *J Eng*, 2019(19):6287-6290. https://doi.org/10.1049/joe.2019.0179

Mou XQ, Chen XL, Guan J, et al., 2019b. Marine target detection based on improved faster R-CNN for navigation radar PPI images. Proc Int Conf on Control, Automation and Information Sciences, p.1-5. https://doi.org/10.1109/iccais46528.2019.9074588

Ren SQ, He KM, Girshick R, et al., 2017. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans Patt Anal Mach Intell*, 39(6): 1137-1149. https://doi.org/10.1109/TPAMI.2016.2577031

Tian C, Tian YH, Ma HW, et al., 2016. Small target detection for solid-state marine radar. Proc CIE Int Conf on Radar, p.1-4. https://doi.org/10.1109/RADAR.2016.8059302

Trunk GV, George SF, 1970. Detection of targets in non-Gaussian sea clutter. *IEEE Trans Aerosp Electron Syst*, AES-6(5):620-628. https://doi.org/10.1109/TAES.1970.310062

Wang L, Tang J, Liao QM, 2019. A study on radar target detection based on deep neural networks. *IEEE Sens Lett*, 3(3):7000504. https://doi.org/10.1109/LSENS.2019.2896072

Ward KD, Tough RJA, Watts S, 2007. Sea clutter: scattering, the K distribution and radar performance. *Waves Random Compl Med*, 17(2):233-234. https://doi.org/10.1080/17455030601097927

Yavari E, Boric-Lubecke O, Yamada S, 2016. Radar principles. In: Boric-Lubecke O, Lubecke VM, Droitcour AD, et al. (Eds.), Doppler Radar Physiological Sensing. Wiley, Hoboken, USA, p.21-38. https://doi.org/10.1002/9781119078418.ch2

Yu XH, Chen XL, Hu WC, et al., 2016. An overview of marine moving target detection via high-resolution sparse representation. Proc CIE Int Conf on Radar, p.1-5. https://doi.org/10.1109/RADAR.2016.8059231