2022 IEEE 8th International Conference on Cloud Computing and Intelligent Systems (CCIS) | 978-1-6654-7735-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/CCIS57298.2022.10016364

Remote Sensing Ship Detection Based on Three-way Decisions and Multi-granularity Feature Fusion

Li Ying[1,2], Duoqian Miao[1], Zhifei Zhang[2], Qixian Zhang[1,2]

[1]Department of Computer Science and Technology, Tongji University, Shanghai 201804, China
[2]Project Management Office of China National Scientific Seafloor Observatory, Tongji University, Shanghai 200092, China
1910663@tongji.edu.cn, dqmiao@tongji.edu.cn, zhifeizhang@tongji.edu.cn, zhangqx@tongji.edu.cn

Abstract: With the development of remote sensing (RS), ship detection not only plays a vital role in military and civilian applications but also is full of challenges. There are some problems in detecting RS ships, such as image blurring and ship size differences, which seriously affect the accuracy of RS ship detection. To solve the above problems, based on the notion of three-way decisions (3WD) and multi-granularity feature fusion, this paper proposes an efficient RS ship detection network (namely 3WM-Net) consisting of two modules: a fuzzy classification and deblurring (FCD) module and a multi-granularity feature extraction enhancement (MFE^2) module. First, the FCD module aims to solve the problem of unreasonable image blur degree classification and deblurring. It combines 3WD and SVM classifiers to achieve image blur classification and uses the BlurGAN algorithm to deblur the image. Then the MFE^2 module is designed to solve the feature extraction problem for ship objects of various sizes. The feature fusion of multi-granularity layers is performed by adding bottom-up paths and feature enhancement structures. Experiments with FGSD2021 and HRSC2016 datasets demonstrate that our 3WM-Net is superior to other RS ship detection methods and can well solve the problems of image blur and various ship sizes in RS ship detection.

Keywords: RS image; Ship detection; Three-way decisions; Multi-granularity feature fusion

1 Introduction

With recent advances in RS technology, satellites have produced high-resolution images, which provide more abundant ship image information for ship object recognition algorithms. Therefore, how to detect RS ship objects has become a critical study content for RS ship detection.

Current RS of ship object detection research primarily has the following ways. The traditional RS ship detection method mainly performs sea and land segmentation through texture and shape characteristics [1], extracts the region of interest, and then uses the contrast box algorithm [2]. Semi-supervised classification algorithm [3] to obtain candidate object regions and get accurate detection results by screening out candidate regions of false ship objects. The RS ship detection based on deep learning mainly uses rectangular bounding boxes (BBoxes) to define the ship's position. It locates the ship through the regression of the BBoxes parameters. Mainstream ship detection mainly involves two types. One-stage detection consists of SSD [4], YOLO [5], YOLOv3 [6], and RetinaNet [7]. Two-stage detection includes Fast RCNN [8] and Faster RCNN [9].

Due to the problems of image blur and ship scale incoherence in RS ship images, the network detection ability is greatly affected. Therefore, in RS ship detection, it is necessary first to divide the ambiguity of the image and then use the detection method of deep learning to detect the ship. As a classical granular computing model, 3WD [10,11] simulates human cognitive decision-making. It determines the acceptance domain, rejection domain, and uncertainty domain during decision-making [12], which can better divide the degree of image blur [13]. However, in RS ship target detection, due to the different scales of ship targets, the extraction of single-granularity features can no longer meet the needs of ship detection, so it is necessary to consider the extraction of multi-granularity [14,15] features fully.

To enhance the effectiveness of RS ship detection, a method for RS ship detection based on 3WD and multi-granularity feature fusion is proposed. This method is based on the following basic idea: adopting the detection strategy after preprocessing for RS ship images. Our work has contributed mainly to the following points:

(1) We propose an FCD to address the unreasonable problem of RS image blur classification and deblurring, which incorporates 3WD into an SVM classifier.

(2) We design an MFE^2 to address the difficulty of multi-scale feature extraction for RS ships. Combined with multi-granularity thinking, the design of MFE^2 in place of FPN in the network.

(3) Our method performs better than other RS ship detection methods in FGSD2021 and HRSC2016 datasets.

The rest of this article is structured as follows. Section 2 describes the current state of research in RS. Section 3 introduces the basic idea of our 3WM-Net. Section 4

978-1-6654-7735-2/22/$31.00 ©2022 IEEE

analyzes the experimental effect. Section 5 draws some conclusions.

2 Preliminaries

The main works on horizontal object detection and SR ship detection are reviewed in this section.

2.1 Horizontal object detection

Continuous development of deep convolutional neural networks has led to remarkable results in object detection in the horizontal direction. Law et al. [16] proposed CornerNet. Zhou et al. [17] proposed CenterNet. Tian et al. [18] proposed FCOS. These methods can only produce horizontal BBoxes, which affects their applicability.

2.2 RS ship detection

RS ship image has the characteristics of fuzzy and inconsistent ship scale. Guo et al. [19] used rotational angle information to upgrade the R-CNN network in response to the inaccurate detection of ships. Yang et al. [20] designed a novel significant segmentation framework to handle scene clutter and inconsistent ship size. Liu et al. [21] designed a spatial network of approximately closed rotating BBoxes for RS ship extraction with complex scenes. The above methods cannot completely alleviate some of the problems existing in RS ship detection.

3 The 3WM-Net model

We illustrate 3WM-Net based on RetinaNet. Let us sequentially introduce the two components of our 3WM-Net (shown in Fig.1), namely the FCD module, and the MFE^2 module.

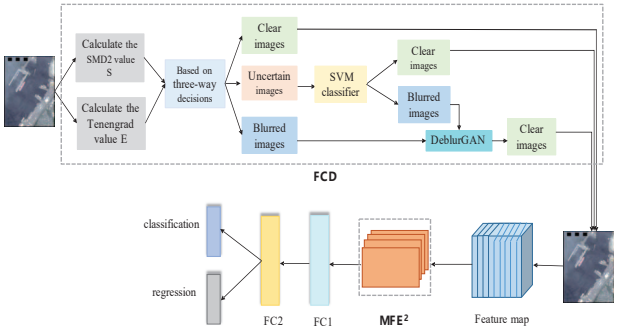
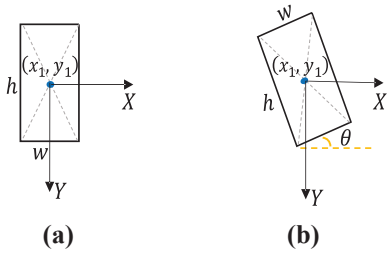


Figure 1 The structure of our 3WM-Net

3.1 RetinaNet as Baseline

In general ship object detection, RetinaNet outputs horizontal BBoxes (shown in Fig.2(a)). To be compatible with RS ship detection in arbitrary-oriented, we change the RetinaNet regression output to rotating BBoxes (shown in Fig.2(b)).



(a) **(b)**

Figure 2 Different BBoxes. (a) Horizontal BBoxes (x_1, y_1, w, h), center point (x_1, y_1), width w and height h. (b) Rotate BBoxes (x_1, y_1, w, h, θ), θ is the acute angle the box makes with the x-axis$(\theta \in [-(\pi/2), 0])$.

3.2 FCD module

In RS ship image processing, fuzzy images will affect the image sharpness [22] and reduce the detection rate of the model. Therefore, to enhance the robustness of the model, this paper proposes a 3WD-based image fuzzy classification algorithm, which can reasonably classify blurry and clear images. Algorithm 1 mainly selects two categories of ambiguity evaluation algorithms. One is the SMD2 function value. The other is the Tenengrad function value. The training is carried out directly on the clear image, and the BlurGAN [23] algorithm deblur the blurred image before training.

- **SMD2 function**

$$SMD2(z) = \sum_y \sum_x |z(x,y) - z(x+1,y)| * |z(x,y) - z(x,y+1)| \quad (1)$$

Where $z(x,y)$ represents the grayscale of the image z corresponding to the pixel (x,y). The larger $SMD2(z)$ is, the less image blur, and vice versa, the more noise.

- **Tenengrad function**

$$Tenengrad(z) = \sum_y \sum_x |H(x,y)| \#(2)$$

$$H(x,y) = \sqrt{H_x^2(x,y) + H_y^2(x,y)} \#(3)$$

Where H_x and H_y are the convolutions of the Sobel horizontal and vertical edge detection operators at the pixel (x,y), respectively. The larger $Tenengrad(z)$ is, the less image blur, and vice versa, the more noise.

The posterior probability P of each data, the formula is as follows:

$$P(z_j = i | t_j) = \frac{\beta_i p(t_j | v_i, \Sigma_i)}{\sum_{l=1}^{c} \beta_i p(t_j | v_l, \Sigma_l)} \#(4)$$

Where $p(t_j | v_i, \Sigma_i)$ denotes the probability density function of each blend element in t_j.

Update covariance matrix $\Sigma_i^{'}$, the mean vector $v_i^{'}$, and mixing coefficient $\beta_i^{'}$, the formula is as follows:

$$v_i^{'} = \frac{\sum_{j=1}^{n} \eta_{ji} t_j}{\sum_{j=1}^{n} \eta_{ji}} \#(5)$$

$$\Sigma_i^{'} = \frac{\sum_{j=1}^{n} \eta_{ji} \left(t_j - v_i^{'}\right)\left(t_j - v_i^{'}\right)^T}{\sum_{j=1}^{n} \eta_{ji}} \#(6)$$

$$\beta_i^{'} = \frac{\sum_{j=1}^{n} \eta_{ji}}{n} \#(7)$$

The calculation formula of cluster mark l_j is as follows:

$$l_j = argmax\, \eta_{ji} (i \in 1,2,...,c\,) \#(8)$$

Algorithm 1 Automatic classification of image blur level based on 3WD

Input: All images $Q = (q_1, q_2, \cdots, q_n)\ (j = 1,2,\cdots,n)$
Output: Clear image Q_C, Blurred image Q_B
1: For all images, calculate the SMD2 and Tenengrad values of each image, get $S_i = SMD2(q_i)$, $E_i = Tene ngrad(q_i)$, $T_i = [S_i, E_i]$, note $T = (t_1, t_2, \cdots, t_n)\ (j = 1,2,\cdots,n)$
2: Based on 3WD, initialize the parameters $\{(\beta_i, v_i, \Sigma_i)\}$ $(1 \le i \le c, c = 3)$
3: Repeat
4: for $j = 1,2,\cdots,n$ do
5: Calculate the posterior probability generated by each mixed component of t_j according to formula (4), that is $\eta_{ji} = P(z_j = i|t_j)\ (1 \le i \le c, c = 3)$
6: end for
7: for $i = 1,2,\cdots,c$ do
8: Update $v_i^{'}$, $\Sigma_i^{'}$, and $\beta_i^{'}$ according to eq. (5), eq. (6), and eq. (7)
9: end for
10: Until the stop condition is met: the current β_i, v_i, and Σ_i remain unchanged
11: $C_i = \phi(1 \le i \le c, c = 3)$
12: for $j = 1,2,\cdots,n$ do
13: Calculate the cluster mark l_j of t_j according to eq. (8) and assign t_j to the corresponding cluster $C_{l_j} = C_{l_j} \cup \{t_j\}$
14: end for
15: Get blurred image C_1, uncertain image C_2, and clear image C_3. The uncertain image C_2 is divided into blurred image C_{11} and clear image C_{33} using a SVM classifier, to obtain the blurred image $Q_B = C_1 \cup C_{11}$, the clear image $Q_C = C_3 \cup C_{33}$.

3.3 MFE² module

In the MFE² module, on the one hand, a bottom-up feature fusion layer (BF²L) (shown in Fig.3(b)) is added, which can easily transfer information from the bottom to the top. On the other hand, a feature enhancement layer (FEL) (shown in Fig.3(a)) is added to improve the high-level feature representation and reduce high-level information loss.

In the FEL (shown in Fig.3(a)), scale-invariant adaptive pooling is first used in the C_5 layer to generate feature maps of multiple scales $\{\beta_1 \times S, \beta_2 \times S, \ldots, \beta_n \times S\}$, and then use a 1 × 1 convolution kernel to generate 256-channel output, then upsample to the scale of S through bilinear interpolation, and finally fuse into F_6 by summation. Combine F_6 and F_5 by summation, propagate to fuse with other low-level features, and append a 3×3 convolutional layer to each feature map to construct a feature pyramid $\{P_2, P_3, P_4, P_5\}$. Where '$conv1 \times 1$' refers to 1×1 convolution, '$conv3 \times 3$' refers to 3×3 convolution, '$2 \times up$' refers to bilinear difference for upsampling, and '+'means addition.

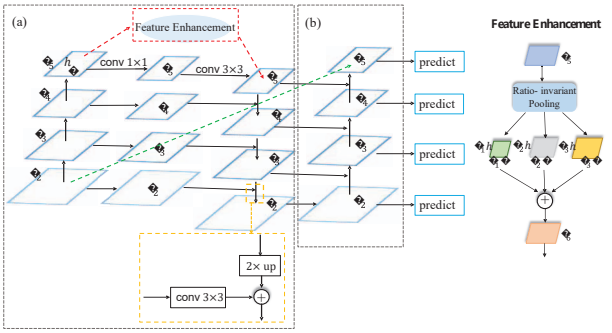


Figure 3 The overall structure of the MFE² module

In the BF²L, a total of four feature maps $\{G_2, G_3, G_4, G_5\}$ are obtained by adding a bottom-up path (shown by the green dotted line in Fig.3(b)).

4 Experiments and Results

4.1 Datasets

HRSC2016 [24] is an RS image ship detection dataset with 1061 images. Each image size is 300 × 300 to 1500 × 900, and the training, validation, and test set contain 436, 181, and 444 images, respectively.

FGSD2021 [25] was derived from the open Google Earth acquisition of high-resolution satellite imagery to capture 636 normalized GSD images. The image width is 157-7789 pixels, the average width is 1202 pixels, and the height is 224-6506 pixels. It was split into 424 training and 212 test images.

4.2 Experimental Settings and Evaluation Criterions

The operating system Ubuntu16.04 was used in the experiment, graphics card model Tesla V100 PCIe 32GB. The SGD optimizer is used with an initial learning rate of 0.01, the learning rate is divided by 10 for each decay step, and the batch size is 4. Momentum and weight decay are 0.9 and 0.0001, respectively. For HRSC2016, a total of 36 epochs. For FGSD2021, a total of 72 epochs.
Selected indicators [7,8] to assess experimental results of different algorithms, including R, P, AP, and mAP.

4.3 Ablation study

To discuss the impact of different modules on 3WM-Net, this section conducts a series of experiments on FGSD2021and HRSC2016 datasets.

4.3.1 Effectiveness of FCD

To discuss the effect of FCD in the experiments, we detect the RS ship image before deblurring and the RS ship image after deblurring, respectively. As shown in Table I, the ship detection accuracy of the deblurred FGSD2021 is 1.42% higher than that of the unblurred FGSD2021. The ship detection accuracy of the deblurred HRSC2016 is 1.71 % higher than that of the unblurred HRSC2016.

4.3.2 Effectiveness of the MFE^2

In the FEL, we choose β to be treated as 0.1, 0.2, and 0.3, respectively. As shown in Table I, the model with the MFE^2 module achieves a more compelling performance. For HRSC2016 and FGSD2021, the improvements are 4.98% and 4.81%, respectively. It shows that our use of the MFE^2 module can significantly enhance the performance of the generated FPN.

Table I Performance of different module combination strategies on FGSD2021 and HRSC2016

	FCD	MFE^2	mAP
FGSD2021			
Baseline			73.45
Different Settings of 3WM-Net	✓		74.87
		✓	78.26
	✓	✓	**80.08**
HRSC2016			
Baseline			80.81
Different Settings of 3WM-Net	✓		82.52
		✓	85.79
	✓	✓	**87.15**

4.4 Comparison with other algorithms

This section compares and analyzes this algorithm with other typical RS ship detection algorithms (See Tables II and III. The results show that our mAP is much higher than others. The reasons are: (1) Based on 3WD，3WM-Net performs a reasonable fuzzy classification and deblurring of RS ship images for subsequent image detection. (2) Combined with the idea of multi-granularity, 3WM-Net enhances the multi-granularity feature fusion of RS ship objects, which is helpful for effectively extracting RS ship object features.

Table II Comparison results on the FGSD2021 dataset

Method	Backbone	Image Size	mAP
CSL [26]	Resnet50	512x512	73.73
SCRDet [27]	Resnet50	512x512	75.90
R^{2}CNN [7]	Resnet50	512x512	78.09
DCL [28]	Resnet50	512x512	73.34
R^{3}Det [29]	Resnet50	512x512	70.47
RSDet [30]	Resnet50	512x512	73.74
3WM-Net(ours)	Resnet50 -MFE^2	512x512	**80.08**

Table III Comparison results on the HRSC2016 dataset, 07 means using the 2007 evaluation metric.

Method	Backbone	Image Size	mAP (07)
R^{2}CNN [7]	Resnet101	800x800	73.07
RRPN [19]	Resnet101	800x800	79.08
R^{2}PN [31]	VGG16	-	79.6
ROI-Trans [21]	Resnet101	512x800	86.20
RSDet [30]	Resnet152	-	86.5
3WM-Net(ours)	Resnet101 -MFE^2	512x512	**87.15**

5 Conclusions

This paper fully uses the 3WD theory and the idea of multi-granularity and proposes a simple and effective 3WM-Net for RS ship object detection. In the proposed 3WM-net, the FCD module achieves reasonable blur classification and deblurring of RS ship images. The MFE^2 module effectively extracts ship object features of different sizes by enhancing the fusion of multi-granularity features. Extensive experiments show that our 3WM-Net has better detection accuracy than other ship detectors on datasets HRSC2016 and FGSD2021. In the future, we hope to explore a lightweight single-stage object detector to improve the detection efficiency of RS ships.

Acknowledgements

The work is supported by the National Key Research and Development Program (Gran No.2022YFB3104700), the National Nature Science Foundation of China (Grant Nos. 61976158, 61906137, 62076040, 61976160, 62076182, 62163016, 62006172), the Jiangxi “Double Thousand Plan”, the Jiangxi Provincial Natural Science Fund (No. 20212ACB202001), the China National Scientific Sea-floor Observatory, the Natural Science Foundation of Shanghai (Grant No. 22ZR1466700), and the Interdisciplinary Project in Ocean Research of Tongji University.

References

[1] Tsung-Yi Lin, Priya Goyal, Ross Girshick, et al. Focal Loss for Dense Object Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 42(2):318-327.
[2] Yindong Yu, Xubo Yang, Shuangjiu Xiao, et al. Automated Ship Detection from Optical Remote Sensing Images. Key Engineering Materials, 2012, 500:785-791.

[3] Changren Zhu, Hui Zhou, Runsheng Wang, et al. A Novel Hierarchical Method of Ship Detection from Spaceborne Optical Image Based on Shape and Texture Features. IEEE Transactions on Geoscience and Remote Sensing, 2010, 48(9):3446-3446.
[4] Wei Liu, Dragomir Anguelov, Dumitru Erhan, et al. SSD: Single Shot Multibox Detection. European Conference on Computer Vision. 2016:21-37.
[5] Joseph Redmon, Santosh Kumar Divvala, Ross B. Girshick, et al. You Only Look Once: Unified, Real-Time Object Detection. IEEE Conference on Computer Vision and Pattern Recognition. 2016:779-788.
[6] Joseph Redmon, Ali Farhadi. YOLOv3: An Incremental Improvement. IEEE Conference on Computer Vision and Pattern Recognition. 2018:878-886.
[7] Zikun Liu, Jingao Hu, Lubin Weng, et al. Rotated Region Based CNN for Ship Detection. IEEE International Conference on Image Processing. 2017:900-904.
[8] Ross Girshick. Fast R-CNN. IEEE International Conference on Computer Vision. 2015:1440-1448.
[9] Shaoqing Ren, Kaiming He, Ross Girshick, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(6):1137-1149.
[10] Can Gao, Jie Zhou, Duoqian Miao, et al. Three-way Decision with Co-training for Partially Labeled Data. Information Sciences.2021, 544:500-518.
[11] Xiaodong Yue, Yufei Chen, Duoqian Miao, Hamido Fujita. Fuzzy Neighborhood Covering for Three-way Classification. Information Sciences. 2020,507:795-808.
[12] Guangming Lang, Duoqian Miao, Hamido Fujita. Three-way Group Conflict Analysis Based on Pythagorean Fuzzy Set Theory. IEEE Transactions on Fuzzy Systems. 2020, 28(3): 447-461.
[13] Yuebing Zhang, Zhifei Zhang, Duoqian Miao, et al. Three-way Enhanced Convolutional Neural Networks for Sentence-level Sentiment Classification. Information Sciences, 2019, 477(3):55-64.
[14] Jie Zhou, Zhihui Lai, Duoqian Miao, et al. Multigranulation Rough-fuzzy Clustering Based on Shadowed Sets. Information Sciences. 2020,507:553-573.
[15] Shengdan Hu, Duoqian Miao, Witold Pedrycz. Multi granularity Based Label Propagation with Active Learning for Semi-supervised Classification. Expert Systems with Applications, 2021: 116276.
[16] Hei Law, Jia Deng. CornerNet: Detecting Objects as Paired Keypoints. International Journal of Computer Vision. 2020, 128:642-656.
[17] Xingyi Zhou, Dequan Wang, Philipp Krähenbühl. Objects as Points. 2019, arXiv:1904.07850. [Online]. Available: http://arxiv.org/abs/1904.07850
[18] Zhi Tian, Chunhua Shen, Hao Chen, et al. FCOS: Fully Convolutional One-Stage Object Detection. International Conference on Computer Vision. 2019:9626-9635.
[19] Haoyuan Guo, Xi Yang, Nannan Wang, et al. A Rotational Libra R-CNN Method for Ship Detection. IEEE Transactions on Geoscience and Remote Sensing,2020, 58(8):5772-57
81.
[20] Feng Yang, Qizhi Xu, Bo Li. Ship Detection from Optical Satellite Images Based on Saliency Segmentation and Structure-LBP Feature. IEEE Geoscience and Remote Sensing Letters, 2017, 14(5): 602-606.
[21] Zikun Liu, Hongzhen Wang, Lubin Weng, et al. Ship Rotated Bounding Box Space for Ship Extraction from High-Resolution Optical Satellite Images with Complex Backgrounds. IEEE Geoscience and Remote Sensing Letters, 2016, 13(8):1074-1078.
[22] Chunwei Tian, Menghua Zheng,Wangmeng Zuo, et al. Multi-stage image denoising with the wavelet tranform. Pattern Recognition, 2022: 109050.
[23] Orest Kupyn, Tetiana Martyniuk, Junru Wu, et al. DeblurGAN-v2: Deblurring (Orders-of-Magnitude) Faster and Better. International Conference on Computer Vision. 2019:8877-8886.
[24] Liu Z, Yuan L, Weng L, et al. A High Resolution Optical Satellite ImageDataset for Ship Recognition and Some New Baselines. ICPRAM. 2017: 324-331.
[25] Feng Zhang, Xueying Wang, Shilin Zhou, et al. Arbitrary-Oriented Ship Detection through Center-Head Point Extraction. IEEE Transactions on Geoscience and Remote Sensing. 2021:3120411.
[26] Xue Yang, Junchi Yan. Arbitrary-Oriented Object Detection with Circular Smooth Label. European Conference on Computer Vision. 2020:677-694.
[27] Xue Yang, Jirui Yang, Junchi Yan, et al., SCRDet: Towards more robust detection for small, cluttered, and rotated objects. European Conference on Computer Vision. 2019:8231-8240.
[28] Xue Yang, Liping Hou, Yue Zhou, et al. Dense Label Encoding for Boundary Discontinuity Free Rotation Detection. IEEE Conference on Computer Vision and Pattern Recognition. 2021:15819-15829.
[29] Xue Yang, Junchi Yan, Ziming Feng, et al. R^3Det: Refined Single-Stage Detector with Feature Refinement for Rotating Object. 2019, arXiv:1908.05612. [Online]. Available: http://arxiv.org/abs/1908.05612
[30] Wen Qian, Xue Yang, Silong Peng, et al. Learning Modulated Loss for Rotated Object Detection. Association for the Advancement of Artificial Intelligence. 2021:1-11.
[31] Zenghui Zhang, Weiwei Guo, Shengnan Zhu, et al. Toward Arbitrary-Oriented Ship Detection with Rotated Region Proposal and Discrimination Networks. IEEE Transactions on Geoscience and Remote Sensing Letters, 2018, 5(11):1745-174
9.