

Original Paper

Oriented Ship Detection Based on Coordinate System Projection in SAR Images

Jiangtao Wang^{1,2,3*} and Mingyang Wang¹

¹*School of Physics and Electronic Information, Huaibei Normal University, Huaibei, China*

²*School of Information, Huaibei Normal University, Huaibei, China*

³*Anhui Province Key Laboratory of Intelligent Computing and Applications, Huaibei, China*

ABSTRACT

Ship detection is a critical and challenging task in aerial images. Due to the special generation method of SAR images, they have unique characteristics. However, different from objects in natural images, ships in SAR images are often distributed with arbitrary orientations and dense distributions. Recently, key points-based anchor-free object detection algorithms have attracted the attention of quite a few researchers. To solve the task of oriented ship detection based on key points, in this paper, we propose an oriented object detection method based on coordinate system projection (CSProjection). In this work, we first detect the key point of the ship, namely, the center point, then establish a coordinate system with the object center point as the base point, and obtain a bounding box of the oriented object through the projection information of the object. Our method can effectively reduce the number of parameters applied to determine the oriented bounding box during training and decrease the network complexity. Experimental results on several SAR ship detection datasets, including SSDD, SRSDD-v1.0 and the optical remote sensing dataset HRSC2016, indicate

*Corresponding author: Jiangtao Wang, jiangtaowang@chnu.edu.cn.

that our method can compete with state-of-the-art algorithms for oriented ship detection, even those with more complex backbones.

Keywords: SAR images, ship detection, coordinate system projection, deep learning.

1 Introduction

SAR systems have unique advantages and play an important role in environmental monitoring, marine monitoring and military applications [9, 33, 40]. Object detection in SAR images has become a research hotspot. Different from objects in natural images, ships in SAR images are often distributed with arbitrary orientations and dense distributions. Therefore, the use of horizontal bounding boxes (HBBs) in detection faces various difficulties, such as complex background information. Horizontal bounding boxes contain not only the information of their corresponding object but also background information and even other object information. In this scenario, oriented bounding boxes (OBBs) can show the advantage of more closely surrounding the object. At the same time, SAR images are different from optical remote sensing images, which is mainly due to the different image generation methods. Generally, SAR ship detection is a challenging task.

Convolutional neural networks promote the research and development of object detection technology due to their powerful feature extraction ability, which also includes object detection in SAR images [37]. At present, object detection algorithms are divided into two-stage and single-stage algorithms. The main difference between them is that the two-stage object detection algorithm usually includes a process of object proposal to find the candidate regions of objects. On the contrary, the single-stage object detection algorithms treat object-detection tasks as regression, and can solve the object detection problem end to end by more straightforward and brute-force means. Two-stage object detection algorithms mainly include RCNN [7], Fast-RCNN [6] and Faster-RCNN [20]. Mainstream object detection algorithms are single-stage algorithms based on deep learning. Such algorithms include anchor-based object detection algorithms and anchor-free object detection algorithms. YOLO [19] and SSD [15] are representative of object detection algorithms based on anchors. These kinds of algorithms detect objects using many anchors. YOLO regards object detection as a regression problem, which makes the structure of the entire network simple. Compared with two-stage algorithms based on candidate regions, such as Faster-RCNN, it pays more attention to the overall information in the training and reasoning process, which is beneficial for enhancing the accuracy of object detection. Similar to YOLO, SSD also regards object detection as a regression problem and performs multiscale prediction, using different feature maps for classi-

fication and position regression, which can further improve the accuracy of detection. The anchor mechanism also has various disadvantages. First, the scale and aspect ratio need to be set according to the object. Second, the number of object instances in each image is limited, but the network will generate many anchor boxes, which will produce many negative samples, resulting in a severe imbalance between positive and negative samples. Object detection algorithms based on key points are anchor-free algorithms. Such algorithms use one or more key points to predict the object. For example, CenterNet [38] is a method for locating an object based on a key point, namely, the center point. If anchors are not needed, the network structure can be made simpler. Therefore, Yi *et al.* [30] extended CenterNet to oriented object detection and realized a fast and accurate detection method called BBAVectors. In their work, in addition to predicting the center point of the object, it is also necessary to predict the four boundary vectors, the outer width and height of the OBB, and the direction map. Hence, this method requires more parameters to predict the oriented bounding box. In addition, the BBAVectors method needs to perform secondary processing when determining the rotated bounding box and then calculate the object rotated bounding box based on the information learned from the network.

To resolve the above issues, in this work, we first select the CenterNet network as the object detection framework, which has a simple structure and faster speed. Second, to reduce the network complexity, we propose a novel structure with fewer parameters to predict the oriented bounding box. The main contributions of this paper are twofold: First, we propose a coordinate system projection (CSProjection) based oriented ship target model. This model uses coordinate system projection to determine an oriented bounding box, which is simpler than the conventional oriented bounding box detection models with less parameters. Then, the proposed CSProjection detection model is realized under the CenterNet object detection framework, and extensive experiments demonstrate that our framework attains higher detection performance on multiple SAR datasets than state-of-the-art methods.

The remainder of this paper is organized as follows. Section 2 describes existing key points-based object detection methods and oriented object detection. The proposed method is then discussed in Section 3. In Section 4, experiment results on three public benchmark data sets are reported and discussed. Finally, we summarize our work in Section 5.

2 Related Works

Compared with horizontal object detection, oriented object detection has become a mainstream research direction. In this section, we introduce some advanced methods for oriented object detection.

2.1 Object Detection Algorithm Based on Key Points

Object detection algorithms based on key points are famous for their simple frameworks and fast running speeds. CornerNet [10] obtains a prediction box by detecting the two key points of the top left corner and the bottom right corner of the object box. This requires grouping the predicted key points in the network to determine which two correspond to the same prediction box. ExtremeNet [39] predicts the four extreme points of each object, namely, the extreme points at the top, left, bottom, and right. In addition, center point prediction is also needed. Similarly, the network also matches these extreme points. Duan’s CenterNet [5] solves the problem that CornerNet does not effectively use the internal features of the object; that is, it cannot perceive the information inside the object, resulting in false detections. Therefore, Duan *et al.* use three key points, namely, the center point, the top left corner point and the bottom right corner point, instead of two points to determine the object, so that the detection network has the ability to perceive the internal information of the object, but as in CornerNet, the key points in Duan’s CenterNet also need to be grouped. Different from the above strategies, Zhou’s CenterNet [30] completely abandons the method of predicting extreme points and corner points; instead, a simpler network detection framework is utilized, which is reflected in the following aspects. First, Zhou’s CenterNet only predicts one key point, namely, the center point of the object, and the center point is used to locate the position of the object. Second, other corner points and extreme points are no longer needed; except the width and height of the object. Finally, the network removes the matching mechanism, making the algorithm simpler.

2.2 Oriented Ship Detection

At present, the research and application of oriented object detection in aerial images has attracted more and more attention. Zhang *et al.* [31] presented a novel object detection network (CAD-Net) that exploits attention-modulated features as well as global and local contexts. The proposed CAD-Net not only learns global and local contexts of objects by capturing their correlations with the global scene but also learns the local neighboring objects or features. Yang *et al.* [29] presented a novel multicategory rotation detector for small, cluttered and rotated objects. It includes a tailored feature fusion structure and a supervised multi-dimensional attention network for small objects. An improved smooth L1 loss is devised to handle of arbitrarily rotated objects, which solved the regression problem of rotating bounding box. Yang and Yan [27] designed a new method to address the boundary problem by transforming angular prediction from a regression problem to a classification task with little accuracy loss. Furthermore, [27] proposed a circular smooth label technique

to handle the periodicity of the angle. Xie *et al.* [26] proposed an effective and simple oriented object detection framework. In the first stage, oriented region proposal network generates high quality-oriented proposals in a nearly cost-free manner whereas oriented R-CNN refines and recognizes oriented regions of interest in the second stage. In [4], Ding *et al.* proposed a Region of Interest (RoI) Transformer to apply spatial transformations on Region of Interests and learn the transformation parameters under the supervision of oriented bounding box annotations to oriented object detection. A novel object detection network was designed in [32], the network designed by Zhang *et al.* learns the global and local context of the object and designs a spatial and scale aware attention module that guides the network to focus on more features. It achieved superior detection performance. Similarly, Han *et al.* [8] proposed a Rotation-equivariant Detector to address the issue that Convolutional Neural Network do not explicitly model the orientation variation by explicitly encoding rotation equivariance and rotation invariance. The network structure of these methods is relatively complicated, and the algorithm complexity is high. To generate more sophisticated feature, [36] proposed an inshore ship detection method in SAR images by using hybrid features. The attention mechanisms were also utilized to improve the performance of SAR ship detection. In [17] a regional attention-based algorithm was designed to locate the regions of interest whereas [22] inserts the spatial and channel convolutional attention mechanism (CBAM) into the YOLOV4 framework to achieve more better performance. Yi *et al.* [30] extended anchor-free-based object detectors to oriented object detection tasks, which reduced the complexity of the network and accelerated the detection speed. However, this method uses many parameters to determine the oriented bounding box, which limits the performance of the one-stage and anchor-free network.

3 The Proposed Methods

3.1 Model Structure

Our network structure is inspired by BBAVectors [30], as shown in Figure 1 (a). Our backbone network chose the classic ResNet structure. The network first performs feature extraction through the backbone network and then performs upsampling on the obtained feature information. The purpose of this step is to expand the size of the feature map to facilitate subsequent convolution to obtain three types of information: the center point position information of the object, the offset of the center point position information and the projection information of the object. The center point information of the object and the offset information jointly determine the position of the ship object, and the projection information is used to determine the oriented bounding box. Finally,

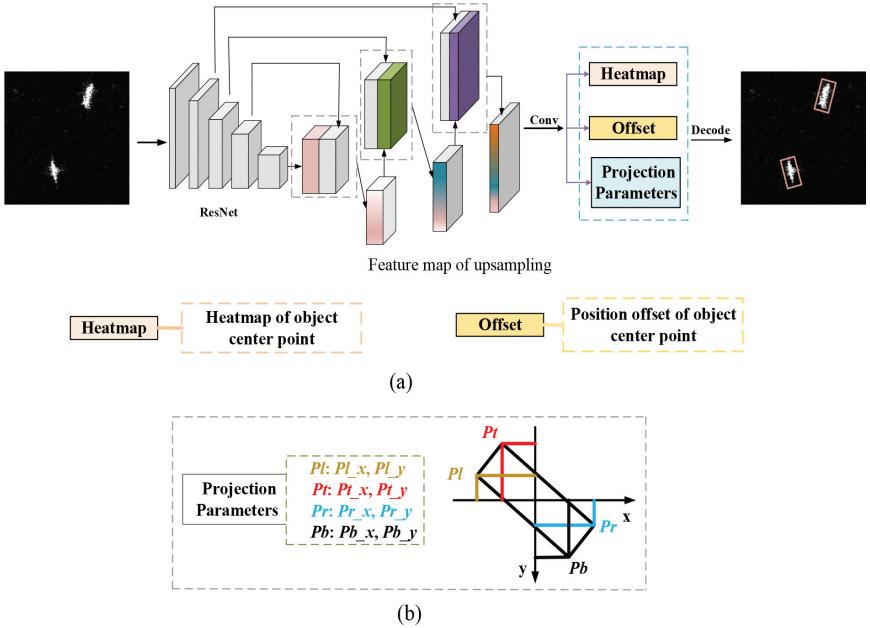


Figure 1: Our network model. We determine the rotated bounding box of the ship object through the coordinate projection vector information of four points.

by decoding the information of these three parts, we can realize oriented ship detection in the SAR image.

3.1.1 Object Location Heatmap

The object heatmap integrates classification and positioning. Different types of location information are integrated into different heatmaps. The value of the heatmap is used as the confidence level. The heatmap is $P \in R^{\frac{W}{S} \times \frac{H}{S} \times C}$, where C is the number of channels and $S = 4$. The key points are distributed in the heatmap according to a Gaussian distribution, and the ground-truth heatmap is $\hat{P} \in R^{\frac{W}{S} \times \frac{H}{S} \times C}$. We use the focal loss for training:

$$L_h = -\frac{1}{N} \sum_i \begin{cases} (1 - p_i)^\alpha \log(p_i) & \text{if } \hat{p}_i = 1 \\ (1 - \hat{p}_i)^\beta (p_i)^\alpha \log(1 - p_i) & \text{otherwise} \end{cases} \quad (1)$$

where \hat{p} and p refer to the ground-truth and the predicted heatmap values, respectively; N is the number of objects; i is the index of the pixel location on the feature map; here $\alpha = 2$ and $\beta = 4$.

3.1.2 Object Location Offset

An offset is used to adjust the position of the target center point to compensate for the error between integer coordinates and floating-point coordinates in the downsampling process. The offset is optimized using the smooth L1 loss:

$$L_o = \frac{1}{N} \sum_{k=1}^N SmoothL1(o_k - \hat{o}_k) \quad (2)$$

$$o = \left(\frac{\bar{c}_x}{s} - \text{floor}\left(\frac{\bar{c}_x}{s}\right), \frac{\bar{c}_y}{s} - \text{floor}\left(\frac{\bar{c}_y}{s}\right) \right) \quad (3)$$

where N is the total number of objects, k is the object index, \hat{o} refers to the ground-truth offset, and (\bar{c}_x, \bar{c}_y) is the ground-truth center point.

3.2 Description of the Oriented Ship Detection Method via Coordinate System Projection

The BBAVectors method provides two ways to determine the rotated bounding box of an object. One is to add a rotation angle to the horizontal bounding box and use the angle information to change the horizontal bounding box into a rotated bounding box. The disadvantages of such an approach are that the change in rotation angle will affect an object with a relatively large aspect ratio and a small angle change will also produce a large change in the IOU value. Moreover, since the angle is periodic, this approach will have an impact on the training and learning of the detection network model. The second approach is to use the BBAVectors method, which considers the vectors on the top, right, bottom and left of the object center point. It also utilizes a direction angle parameter to judge whether the object is better bounded by a horizontal boundary box or a rotated boundary box. BBAVectors uses a total of 11 parameters to determine the directed bounding box. In terms of the final decoding information of the rotated box representing the object in the network, BBAVectors translates the boundary vector information into four points at the upper-left, upper-right, lower-right and lower-left corners of the rotated box as the final decoding information of the rotated box, which requires subsequent operations on the boundary vector information learned from the network. Therefore, the final decoding process of the rotation information is very complicated.

As shown in Figure 1(b), the CSProjection method proposed in this work uses fewer parameters. As a result, it can more effectively determine the rotated boundary boxes of ships in SAR remote sensing images without processing the information learned from detection networks. The CSProjection method consists of 8 parameters, namely, $P = [P_l, P_t, P_r, P_b]$, where P_l represents the projection information of the left boundary point L of the rotated bounding box to the x- and y-axes, P_t represents the projection information of the

upper boundary point T of the rotated bounding box to the x- and y-axes, P_r represents the projection information of the right boundary point R of the rotated box to the x- and y-axes, and P_b represents the projection information of the lower boundary points B of the rotated bounding box to the x- and y-axes. Using the projection information of the four boundary points to represent the rotated boundary box makes the learning of the detection network more effective.

The projection information of our coordinate system is expressed in Equation (4), where point C is the center point of the ship while Cx and Cy are the x-axis and y-axis information.

$$\begin{aligned} Pl &= [Lx - Cx, Ly - Cy] & Pt &= [Tx - Cx, Ty - Cy] \\ Pr &= [Rx - Cx, Ry - Cy] & Pb &= [Bx - Cx, By - Cy] \end{aligned} \quad (4)$$

As the loss function of the CSProjection method, the smooth L1 loss is still used in the experiment:

$$L_P = \frac{1}{N} \sum_{k=1}^N SmoothL1(P_k - \hat{P}_k) \quad (5)$$

where P_k and \hat{P}_k are the predicted and ground-truth parameter values, respectively.

To express our method more clearly, we compare the proposed CSProjection method with the Center + wh + angle and BBAVectors methods in Figure 2. When Center + wh + angle addresses rotated objects, due to the periodicity of the angle, the learning process of the model may be disordered, affecting the learning performance of the model. BBAVectors perfectly solves this problem using boundary-aware vectors at the cost of more parameters required to determine the object orientation, which affects the performance of the model. Therefore, we propose a method called CSProjection, which not only solves the problem of having many parameters but also can directly represent the rotated box without information postprocessing.

4 Results

In our experiments, to fully evaluate our proposed SAR rotated ship detection algorithm, we select the SSDD and SRSDD-v1.0 datasets for SAR rotated object detection. We also conduct an extended experiment, in which training and testing are conducted on the optical remote sensing image dataset HRSC2016 to further verify the scalability of the algorithm.

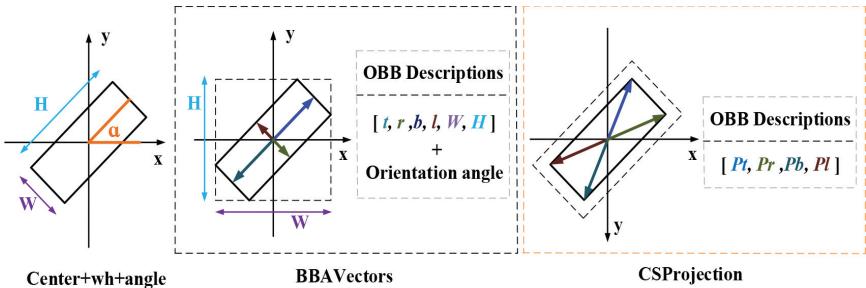


Figure 2: Comparison of CSProjection with other methods.

4.1 Datasets

SRSDD-v1.0 [11] is a dataset dedicated to SAR ship detection. All data in the dataset were collected using GF-3 Spotlight (SL) mode with a 1 m resolution. SRSDD-v1.0 has a total of six categories and 2884 ships, and it is a challenging dataset for SAR rotated object detection. The ships are divided into six categories: ore-oil ships (ore-oil), bulk cargo ships (bulk cargo), fishing boats (fishing), law enforcement ships (LawEnforce), dredger ships (dredger), and container ships (container).

SSDD [12] is also a publicly available SAR image dataset. In SSDD , there are a total of 1160 images and 2456 ships. SSDD is a publicly available dataset that can be used to train and test detection algorithms.

HRSC2016 [16] dataset is a relatively difficult dataset in ship detection. The images in the dataset are from Google Earth. The image sizes range from 300×300 to 1500×900 , and the dataset contains 436 training images, 181 validation images, and 444 testing images.

4.2 Experimental Details

All experiments are run on the Ubuntu16.04 system, using PyTorch to build the detection network model and using the Adam optimization algorithm during training. The initial learning rate is set to 3.25e-4. We use GeForce GTX 1070 for the experiment on the SAR dataset. The input image size of SSDD is 400×500 , and the input image size of SRSDD-v1.0 is 1024×1024 .

4.3 SRSDD-v1.0

SRSDD-v1.0 is divided into six object categories, but there are severe imbalances in the numbers of samples among the object categories of SRSDD-v1.0, which poses a challenge to the performance of the object detection algorithm. Therefore, our experiments on SRSDD-v1.0 are conducted in two parts. One

part is the network training and testing of the six object categories. Notably, we increase the training proportion of some categories in the part of the training. The other part is classification of the six categories into one category, namely, the ship category. This part of the training still uses the original training dataset.

Table 1 compares the detection data on the six object categories in SRSDD-v1.0, mainly including the mAP value of the object detection algorithm, the speed of the detection network and the model size of the network. As the model of the algorithm, we use a shallow convolutional neural network with few network parameters and low computational complexity. However, in terms of detection accuracy, our proposed algorithm obtains a higher mAP value and has advantages in detection speed.

Table 1: Data comparison for the six ship categories in SRSDD-v1.0.

Algorithm	Model Category	Object Categories	mAP	FPS	Model Size
R-RetinaNet[21]	One-stage	6	32.73	10.53	277MB
R3Det[28]	One-stage	6	39.12	7.69	468 MB
BBAVectors[30]	Anchor-free	6	45.33	3.26	829MB
R-FCOS[24]	Anchor-free	6	49.49	10.15	244 MB
CSPProjection	Anchor-free	6	50.23	15.48	157MB

Table 2 presents the AP values for each object category. From the perspective of network detection accuracy, the LawEnforce category has too few training samples, resulting in a poor detection performance of the detection network for this category. Our proposed CSPProjection method realizes the highest AP values on the ore-oil, Dredger and Container categories, and its detection accuracies on bulk-cargo and Fishing are lower than those of other algorithms. Figure 3 presents detection results on SRSDD-v1.0.

Table 2: Comparison of AP values for each ship category in SRSDD-v1.0.

Algorithm	ore-oil	bulk-cargo	Fishing	LawEnforce	Dredger	Container
R-RetinaNet[21]	30.37	35.79	11.47	2.07	67.71	48.94
R3Det[28]	44.61	42.98	18.32	1.09	54.27	73.48
BBAVectors[30]	54.33	34.84	21.03	1.09	82.21	78.51
R-FCOS[24]	54.88	47.36	25.12	5.45	83.00	81.11
CSPProjection	69.37	33.20	16.61	0.20	88.47	93.54

Because SRSDD-v1.0 has the problem of sample imbalance, to further test our proposed algorithm, we combine the original six object categories into one

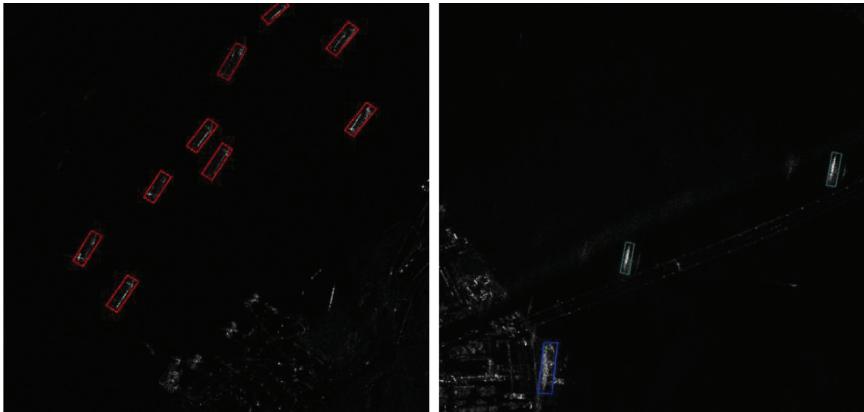


Figure 3: Examples of partial detection results on SRSDD-v1.0.

category, namely, the ship category. Compared with the BBAVectors method in Table 3, when the backbone network is ResNet-18, the CSProjection method has a higher detection accuracy than BBAVectors. On the other hand, because we use fewer parameters when determining the rotated object bounding box, CSProjection has advantages in terms of detection speed. Even compared with ResNet-34, our proposed method is still very competitive when the backbone network is ResNet-18.

Table 3: Data comparison for the ship category in SRSDD-v1.0.

Algorithm	Model Category	Backbone Network	Object Categories	mAP	FPS
BBAVectors	Anchor-free	ResNet-18	1	30.74	15.63
BBAVectors	Anchor-free	ResNet-34	1	35.06	13.12
CSProjection	Anchor-free	ResNet-18	1	41.03	15.82

The performance of the object detection algorithm is further analyzed in terms of Recall, Precision and F1 score in Table 4. From the data, it can conclude that the CSProjection method outperforms the BBAVectors method.

4.4 SSDD

On SSDD, two types of bounding boxes are utilized: horizontal and oriented bounding boxes. In the experiment, we use rotated boxes for training and testing. On the test data, the listed horizontal box detection algorithms achieve a maximum accuracy of 83.4%. As for the Oriented Bounding Box

Table 4: Comparison of recall, precision and F1 values on SRSDD-v1.0.

Algorithm	Backbone Network	Object Categories	Recall	Precision	F1
BBAVectors	ResNet-18	1	44.28	31.44	36.77
BBAVectors	ResNet-34	1	46.32	31.05	37.17
CSPProjection	ResNet-18	1	52.11	42.04	46.53

based methods, it obtained higher AP value than the Horizontal Bounding Box methods. The detection performance of the proposed method is not only better than that for horizontal detection but also competitive with that of the oriented detection algorithm for rotated bounding box detection. Compared with the second highest AP of BBAVetors, our proposed method with backbone of ResNet18 improves the AP value by 4.1%. To demonstrate the performance of our proposed method, Figure 4 shows the test results of the network on SSDD.

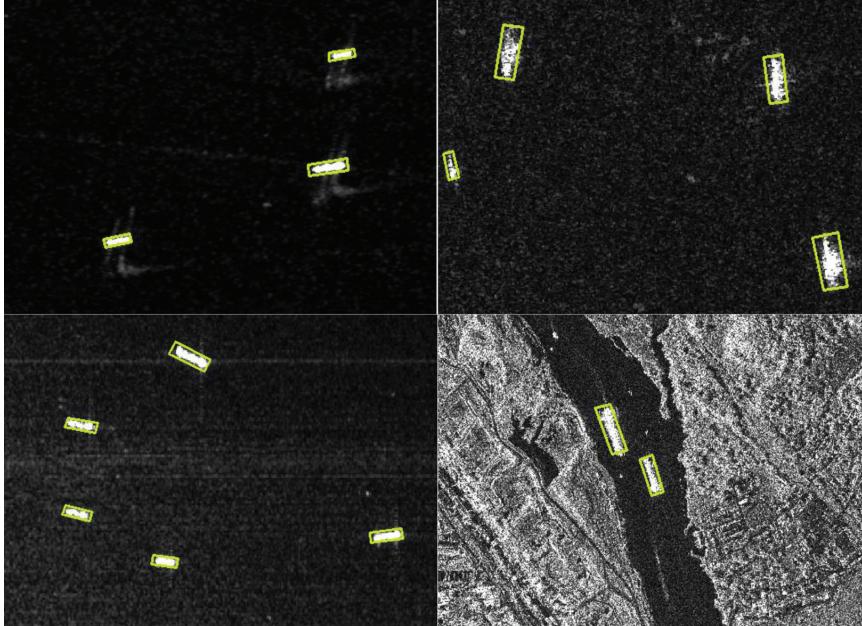


Figure 4: Examples of detection results on SSDD.

4.5 Extended Experiments

To fully validate the performance of our proposed algorithm, we conduct extended experiments in this section.

We conduct experiments not only on the SAR dataset but also on the optical remote sensing HRSC2016 dataset, which can more effectively demonstrate the excellent network performance of our proposed method. The results are compared with those of state-of-the-art algorithms on the HRSC2016 dataset, as presented in Table 6. Zhang *et al.* [35] proposed a rotated region proposal network for generating multioriented proposals with ship orientation angle information, which reached 79.6% AP. Liao *et al.* [14] used the rotation-sensitive regression detector to reach an AP value of 84.3%. Wang *et al.* [25] used a variety of methods, including a novel feature fusion module, a specific weighted loss and a normalization strategy for the representation of rotated bounding boxes, and the AP value reached 85.4%. Sun *et al.* [23] used ResNet-50 as the

Table 5: Comparison of detection results on SSDD.

Method	Object Type	AP (Ship)	Time (ms)
LFO-Net [34]	Horizontal Bounding Box	80.12	9.28
The method of [2]	Horizontal Bounding Box	89.80	-
The method of [1]	Horizontal Bounding Box	79.78	28.40
DAPN [13]	Horizontal Bounding Box	83.40	-
BBAVectors	Oriented Bounding Box	89.79	16.60
RAD [17]	Oriented Bounding Box	87.78	92.00
CSPROJECTION	Oriented Bounding Box	93.89	15.30

Table 6: Detection results of the extended experiment compared with the results of other models.

Method	Backbone	Image size	FPS	AP(Ship)
R ² PN [35]	VGG16	-	-	79.60
RRD [14]	VGG16	384 × 384	-	84.30
SARD [25]	ResNet-101	800 × 800	1.50	85.40
ROI Trans. [4]	ResNet-101	512 × 800	5.90	86.20
R ⁴ Det [23]	ResNet-50	800 × 800	8.60	88.17
BBAVectors+r [30]	ResNet-101	608 × 608	12.88	88.20
BBAVectors+rh [30]	ResNet-101	608 × 608	11.69	88.60
SDA-Net[3]	ResNet-101	-	-	89.21
Point RCNN[18]	ReR50-ReFPN	800 × 512	-	90.53
CSPROJECTION	ResNet-18	608 × 608	50.05	90.53

backbone network, and the AP value was increased to 88.17%. The AP value of the BBAVectors method was found to reach 88.6%. Even though the latest model of Point RCNN [18] which presented in year of 2022 achieves AP value of 90.53%, but both the image size and backbone of this framework are larger than CSProjection. Our proposed CSProjection method achieves an AP value of 90.53% when the backbone network is ResNet-18. By reducing the large numbers of parameters and calculations, the accuracy of detection is improved. Figure 5 shows some test results on the HRSC2016 dataset.

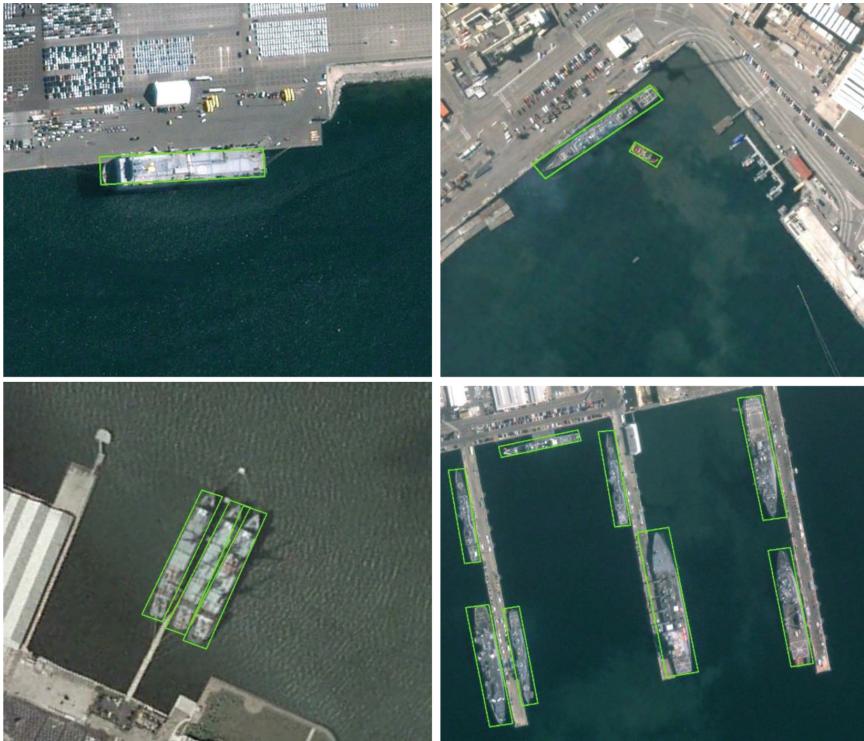


Figure 5: Example of detection results on the HRSC2016 dataset.

4.6 Discussion

When training and testing the SRSDD-v1.0 dataset, we used two methods. When training six object categories, we chose the ResNet-18 network in the experiment. Compared with the deep network, the feature extraction ability of the shallow network was weak. According to the mAP values in Table 1, the algorithm proposed in this paper still demonstrated advantages compared

with similar algorithms. However as presented in Table 2, we found that the detection accuracy rate of the algorithm decreased on some categories compared with similar algorithms. The main reasons for this problem were that the number of object samples was too small and the dataset had severe sample balance problems. Moreover, the ResNet-18 network cannot be used to extract deep features of certain categories.

The proposed algorithm has the following shortcomings. For example, the predicted detection box may not be a standard rectangular box. However, the proposed method is a competitive algorithm for the detection of rotated objects, as was well demonstrated by the experimental results.

5 Conclusions

In this article, we proposed oriented ship detection based on coordinate system projection in SAR images, namely, CSProjection. This method is anchor-free with fewer parameters to describe the oriented bounding box. It can directly use the information learned by the network to describe the rotated bounding box. The results of experiments on SSDD, SRSDD-v1.0 and the HRSC2016 dataset show the advanced nature of our proposed CSProjection method.

Acknowledgments

This research was funded in part by the Project of HuaiBei Normal University (2023ZK041, 2023ZK042) and in part by the Key Project of University Natural Science Research of Anhui Province (2023AH050319, 2023AH010044).

References

- [1] C. Chen, C. He, C. Hu, H. Pei, and L. Jiao, “A Deep Neural Network Based on an Attention Mechanism for SAR Ship Detection in Multiscale and Complex Scenarios”, *IEEE Access*, 7, 2019, 104848–63.
- [2] Z. Cui, Q. Li, Z. Cao, and N. Liu, “Dense Attention Pyramid Networks for Multi-Scale Ship Detection in SAR Images”, *IEEE Transactions on Geoscience and Remote Sensing*, 57, 2019, 8983–97.
- [3] Z. Cui, H. Sun, and R. Yin, “SDA-Net: a detector for small, densely distributed, and arbitrary-directional ships in remote sensing images”, *Applied Intelligence*, 52, 2022, 12516–32.

- [4] J. Ding, N. Xue, Y. Long, G. Xia, and Q. Lu, “Learning RoI Transformer for Oriented Object Detection in Aerial Images”, in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, 2019, 2844–53.
- [5] K. Duan, S. Bai, L. Xie, H. Qi, Q. Huang, and Q. Tian, “Centernet: Keypoint triplets for object detection”, in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, 6569–78.
- [6] R. Girshick, “Fast R-CNN”, in *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, Chile, 2015, 1440–8.
- [7] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation”, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, 2014, 580–7.
- [8] *ReDet: A Rotation-equivariant Detector for Aerial Object Detection*, arXiv:2103.07733, 2021.
- [9] R. W. Jansen, R. G. Raj, L. Rosenberg, and M. A. Sletten, “Practical Multichannel SAR Imaging in the Maritime Environment”, *IEEE Transactions on Geoscience and Remote Sensing*, 56, 2018, 4025–36.
- [10] H. Law and J. Deng, “CornerNet: Detecting objects as paired keypoints”, in *Proceedings of the European Conference on Computer Vision*, Munich, Germany, 2018, 734–50.
- [11] S. Lei, D. Lu, X. Qiu, and C. Ding, “SRSDD-v1.0: A High-Resolution SAR Rotation Ship Detection Dataset”, *Remote Sensing*, 13, 2021, 5104.
- [12] J. Li, C. Qu, and J. Shao, “Ship detection in SAR images based on an improved faster R-CNN”, in *2017 SAR in Big Data Era: Models, Methods and Applications (BIGSARDATA)*, Beijing, China, 2017, doi: [10.1109/BIGSARDATA.2017.8124934](https://doi.org/10.1109/BIGSARDATA.2017.8124934).
- [13] Q. Li, R. Min, Z. Cui, Y. Pi, and Z. Xu, “Multiscale Ship Detection Based on Dense Attention Pyramid Network in Sar Images”, in *IGARSS 2019–2019 IEEE International Geoscience and Remote Sensing Symposium*, Yokohama, Japan, 2019, doi: [10.1109/IGARSS.2019.8899062](https://doi.org/10.1109/IGARSS.2019.8899062).
- [14] M. Liao, Z. Zhu, B. Shi, G. Xia, and X. Bai, “Rotation-sensitive Regression for Oriented Scene Text Detection”, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 2018, 5909–18.
- [15] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Fu, and A. Berg, “SSD: Single Shot MultiBox Detector”, in *Proceedings of the European Conference on Computer Vision*, Amsterdam, Netherlands, 2016, 21–37.
- [16] Z. Liu, L. Yuan, L. Weng, and Y. Yang, “A High Resolution Optical Satellite Image Dataset for Ship Recognition and Some New Baselines”, in *Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods*, Porto, Portugal, 2017, 324–31.

- [17] X. Qi, P. Lang, X. Fu, R. Qin, J. Dong, and C. Liu, “A regional attention-based detector for SAR ship detection”, *Remote Sensing Letters*, 13, 2022, 55–64.
- [18] Z. Qiang and Y. Chaohui, “Point RCNN: An Angle-Free Framework for Rotated Object Detection”, *Remote Sensing*, 14, 2022, 2605.
- [19] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection”, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, 2016, 779–88.
- [20] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39, 2017, 1137–49.
- [21] Rotated-RetinaNet, Available online: <https://github.com/ming71/Rotated-RetinaNet> (accessed on 10/28/2022).
- [22] B. Sun, X. Wang, and H. Li, “Small-target ship detection in SAR images based on densely connected deep neural network with attention in complex scenes”, *Applied Intelligence*, 53, 2023, 4162–79.
- [23] P. Sun, Y. Zheng, Z. Zhou, W. Xu, and Q. Ren, “R4Det: Refined single-stage detector with feature recursion and refinement for rotating object detection in aerial images”, *Image and Vision Computing*, 103, 2020, 104036.
- [24] Z. Tian, C. Shen, H. Chen, and T. He, “FCOS: Fully Convolutional One-Stage Object Detection”, in *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, Korea, 27 October–2 November, 2019, 9626–35.
- [25] Y. Wang, Y. Zhang, Y. Zhang, L. Zhao, X. Sun, and Z. Guo, “SARD: Towards Scale-Aware Rotated Object Detection in Aerial Imagery”, *IEEE Access*, 7, 2019, 173855–65.
- [26] X. Xie, G. Cheng, J. Wang, X. Yao, and J. Han, “Oriented R-CNN for Object Detection”, in *IEEE/CVF International Conference on Computer Vision*, arXiv:2108.05699, Montreal, Canada, 2021.
- [27] X. Yang and J. Yan, “Arbitrary-oriented object detection with circular smooth label”, in *Proceedings of the European Conference on Computer Vision*, Glasgow, US, 2020, 677–94.
- [28] X. Yang, J. Yan, Z. Feng, and T. He, “R3det: Refined single-stage detector with feature refinement for rotating object”, in *Proceedings of the AAAI conference on artificial intelligence*, 2021, 35, 3163–71.
- [29] X. Yang, J. Yang, J. Yan, Y. Zhang, T. Zhang, Z. Guo, X. Sun, and K. Fu, “SCRDet: Towards More Robust Detection for Small, Cluttered and Rotated Objects”, in *IEEE/CVF International Conference on Computer Vision*, Long Beach, CA, USA, 2019, doi: [10.1109/ICCV.2019.00832](https://doi.org/10.1109/ICCV.2019.00832).

- [30] J. Yi, P. Wu, B. Liu, Q. Huang, H. Qu, and D. Metaxas, “Oriented Object Detection in Aerial Images with Box Boundary-Aware Vectors”, in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, Snowmass Village, Colorado, United States, 2020, 2150–9.
- [31] G. Zhang, S. Lu, and W. Zhang, “CAD-Net: A Context-Aware Detection Network for Objects in Remote Sensing Imagery”, *IEEE Transactions on Geoscience and Remote Sensing*, 57, 2019, 10015–24.
- [32] G. Zhang, S. Lu, and W. Zhang, “CAD-Net: A Context-Aware Detection Network for Objects in Remote Sensing Imagery”, *IEEE Transactions on Geoscience and Remote Sensing*, 57, 2019, 10015–24.
- [33] T. Zhang and X. Zhang, “High-Speed Ship Detection in SAR Images Based on a Grid Convolutional Neural Network”, *Remote Sensing*, 11, 2019, 1206.
- [34] X. Zhang, H. Wang, C. Xu, Y. Lv, C. Fu, H. Xiao, and Y. He, “Lightweight Feature Optimizing Network for Ship Detection in SAR Image”, *IEEE Access*, 7, 2019, 141662–78.
- [35] Z. Zhang, W. Guo, S. Zhu, and W. Yu, “Toward Arbitrary-Oriented Ship Detection with Rotated Region Proposal and Discrimination Networks”, *IEEE Geoscience and Remote Sensing Letters*, 15, 2018, 1745–9.
- [36] T. Zheng, P. Lei, and W. J., “A Hybrid Features Based Detection Method for Inshore Ship Targets in SAR Imagery”, *Journal of Geodesy and Geoinformation Science*, 6, 2023, 95–107.
- [37] L. Zhou, S. Wei, Z. Cui, J. Fang, X. Yang, and W. Ding, “Lira-YOLO: A lightweight model for ship detection in radar images”, *Journal of Systems Engineering and Electronics*, 31, 2020, 950–6.
- [38] X. Zhou, D. Wang, and P. Krähenbühl, “Objects as points”, arXiv:1904.07850, 2019.
- [39] X. Zhou, J. Zhuo, and P. Krähenbühl, “Bottom-up object detection by grouping extreme and center points”, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, 2019, 850–9.
- [40] Y. Zhou, W. Wang, Z. Chen, Q. Zhao, H. Zhang, Y. Deng, and R. Wang, “High-Resolution and Wide-Swath SAR Imaging Mode Using Frequency Diverse Planar Array”, *IEEE Geoscience and Remote Sensing Letters*, 18, 2020, 321–5.