



DBW-YOLO: A High-Precision SAR Ship Detection Method for Complex Environments

Xiao Tang , Jiufeng Zhang , Yunzhi Xia , and Huanlin Xiao 

Abstract—Synthetic aperture radar (SAR) is widely used for ship target detection with the application of deep learning techniques. However, in certain complex environments, such as near shore or with small ships, the problem of false alarms and missed detections still exists. To address these issues, a high-precision ship target detection method named DBW-YOLO, which builds upon YOLOv7-tiny as its foundational network, is proposed in this article. The proposed method consists of the following main steps. First, a feature extraction enhancement network based on deformable convolution network is introduced to obtain more comprehensive feature representations across various ship types. Second, an adaptive feature recognition method based on BiFormer attention mechanism is proposed to strengthen detection accuracy, which is more beneficial to capture near shore ships and small ships. Third, a wise intersection-over-union based on dynamic nonmonotonic focusing mechanism is proposed to generate the loss function, which improves the convergence speed and generalization ability. Consequently, the DBW-YOLO method trains a more robust model that better utilizes samples from near shore and small ships. To verify the effectiveness of this method, two SAR datasets, HRSID and SSDD, are employed for performance evaluation. Compared to other widely-used methods, the mAP value of DBW-YOLO reaches 88.84% and 99.18% on the HRSID and SSDD datasets, respectively. The findings indicate that DBW-YOLO method outperforms other representative SAR ship detection methods in both accuracy and overall performance.

Index Terms—Deep learning, synthetic aperture radar (SAR), target identification, YOLOv7.

I. INTRODUCTION

SYNTHETIC aperture radar (SAR) stands as an active microwave remote sensing device employing virtual array and pulse compression technologies to acquire high-resolution two-dimensional images of ground objects. Because of its unique system and rich polarization information, SAR images

Manuscript received 28 December 2023; revised 2 February 2024 and 6 March 2024; accepted 9 March 2024. Date of publication 18 March 2024; date of current version 27 March 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 62302205 and Grant 41904163, in part by the Natural Science Foundation of Hunan Province under Grant 2020JJ5483, and in part by the Research Foundation of Education Bureau of Hunan Province under Grant 21B0428. (Corresponding author: Yunzhi Xia.)

Xiao Tang, Jiufeng Zhang, and Huanlin Xiao are with the School of Mechanical Engineering, University of South China, Hengyang 421001, China (e-mail: tangxiao@usc.edu.cn; jiufengzhang@stu.usc.edu.cn; huanlinxiao@stu.usc.edu.cn).

Yunzhi Xia is with the School of Electrical Engineering, University of South China, Hengyang 421009, China, and also with the School of Cyber Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: yzxia@hust.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2024.3376558

find extensive use in various civilian and military applications [1], [2], [3], such as urban planning, environmental and natural disaster monitoring, and ship target detection. Among them, ship target detection serves the purpose of identifying and locating vessels automatically and holds significant importance in monitoring and ensuring the security of coastal areas and waterways. However, precisely detection in complex environments remains a significant challenge, particularly when it comes to detecting near shore ships and small ships.

Traditional SAR ship detection methods is represented by constant false alarm rate (CFAR) [4] and its improvements [5]. These methods generally rely on exploiting the distinction between maritime targets and the background. When the target's prior information is unknown, the CFAR detection method can achieve improved detection outcomes for its stronger scattering echoes. However, due to their sensitivity to background complexity and inadequate adaptability to target variations, it shows poor detection results with the increasing diversity of radar platform parameters and complexity of the model.

In recent years, due to rapid advancements in deep learning [6], [7], [8], [9], SAR ship detection has seen significant improvements. This is because deep learning techniques have demonstrated excellent performance in extracting complex features from SAR images and possess strong generalization capabilities, which significantly improves the detection accuracy and overall performance of SAR ship detection. Deep learning techniques consist of two groups corresponding to the number of stage: single-stage detectors [10] and two-stage detectors [11]. You only look once (YOLO) [12] and single shot multibox detector (SSD) [13] stands out as the two most popular single-stage detectors, while R-CNN [14], Fast R-CNN [15], and Faster R-CNN [16] are representative of the two-stage detectors. However, two-stage detectors suffer from the drawback of slower ship identification, while single-stage detectors offer quicker detection speeds with moderate accuracy.

As a result, numerous constructive enhancements have been developed based on single-stage detectors. For example, Wang et al. [17] introduced an enhanced SAR detection method for SSD model, which incorporates data augmentation and migration techniques. Tang et al. [18] exhibited an N-YOLO model to diminish the effect of noise and coastlines on ship identification. Sun et al. [19] introduced innovative YOLO-based approach for arbitrary-oriented SAR ship detection, which named BiFA-YOLO, to fully use the feature fusion and angular classification. Li et al. [20] presented a YOLOv5 model combined with confluence algorithm to detect ships in complex environments, such

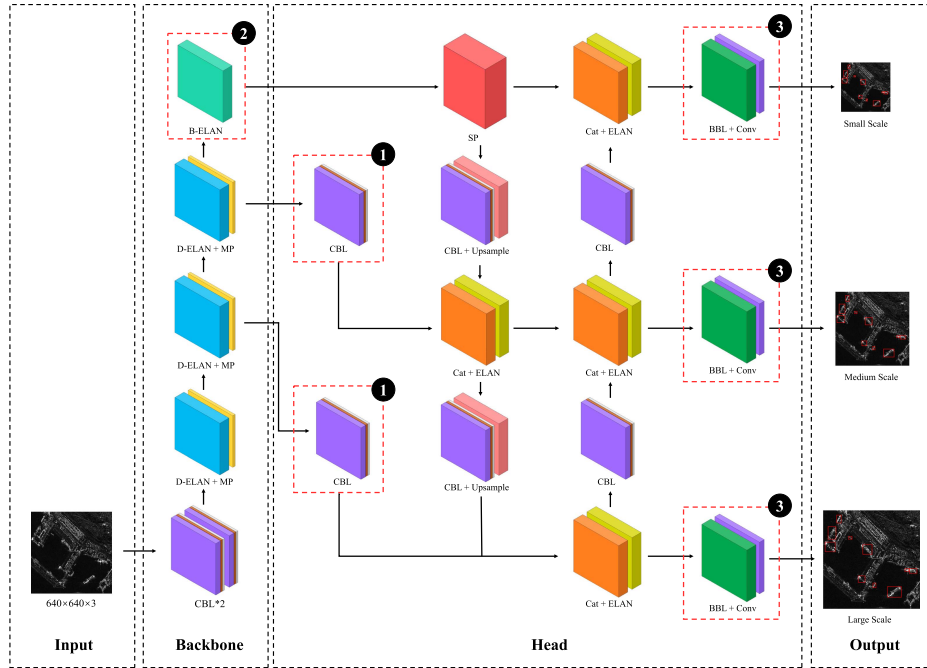


Fig. 1. Network architecture of DBW-YOLO.

as near shore or ships are too dense. Yang et al. [21] introduced an anchor-free detection method named Improved-FCOS to tackle the issues of target detection in multiresolution and near shore environments. Bai et al. [22] introduced an FEPS-Net to optimize ship detection performance with the presence of significant scattered noise and small ships. Chen et al. [23] introduced a CSD-YOLO model, which integrates shuffle attention mechanisms and atrous spatial pyramid pooling, to address the challenge of detecting various ships in intricate environments. Ren et al. [24] introduced YOLO-Lite, an efficient lightweight network addressing challenges related to inaccurate target localization and interference from complex backgrounds in ship detection.

While exploring numerous promising ideas to improve ship detection performance, the most advanced algorithm models still face the challenge of false alarms and missed detections in environments, such as near shore areas or with small ships. This is because:

- 1) for the approaches proposed in [17], [18], [19], [20], the feature extraction networks of these models fail to generate pure ship features, making them prone to interference from the surrounding environment, especially for targets with complex textures or smaller sizes, leading to a decrease in detection accuracy;
- 2) for references [21], [22], [23], [24], these models introduce a global attention mechanism to the network, resulting in significant resource consumption and computational burden, which hinders the deployment and practical application of the model;
- 3) all the models mentioned above do not take into account the competition between high-quality anchors and the detrimental gradients from low-quality samples when calculating the bounding box.

In the case of detecting near shore ships and small ships, they may exhibit suboptimal performance, leading to localization errors.

To tackle these issues, it is essential and beneficial to design a balanced SAR ship detection method improves detection accuracy while ensure real-time detection and recognition in complex environments. Therefore, we propose a novel ship detection method named DBW-YOLO based on YOLOv7-tiny, which integrates DCNets, BiFormer, and Wise IoU to enhance ship detection accuracy. The main contributions can be outlined as follows.

- 1) To enhance the detection speed, YOLOv7-tiny is chosen as the foundational network for DBW-YOLO method in this article, which achieves faster detection speed while ensuring accuracy requirements.
- 2) To improve the detection accuracy, DCNets are employed in D-ELAN downsampling modules, which improve feature extraction capabilities.
- 3) To enhance the detection accuracy for near shore and small ships, a dynamic sparse attention BiFormer with a two-layer routing mechanism is integrated into the network, which aids in capturing near shore and small ships.
- 4) To increase the generalization ability and convergence speed, WIoU is adopted as the loss function with a dynamic nonmonotonic focusing mechanism, which facilitates more effective network training.

II. METHOD

This section presents a comprehensive overview of the DBW-YOLO method, focusing on its network structure and key components.

A. Network Structure

The method's network primarily comprises a backbone, responsible for extracting image features, and a head, utilized for identifying object categories and anchor boxes. Fig. 1 provides an overview of DBW-YOLO network structure, which is established step by step as follows:

1) *Backbone*: The network of DBW-YOLO backbone integrates convolution block (CBL), D-ELAN module, B-ELAN module, and maxpooling blocks (MP blocks), inspired by YOLOv7-tiny. Two CBL blocks are employed for initial feature extraction of the input image, while MP blocks are used for downsampling operations to decrease the dimension of feature map. The structure of these two blocks remains unchanged with that of YOLOv7-tiny. To fully extract ship feature from the images, we specially designed two novel modules called D-ELAN and B-ELAN in backbone network. D-ELAN module is built upon ELAN module from the original YOLOv7-tiny framework, enhanced with the integration of a DCNet to more effectively extract ship features. B-ELAN module represents a further version of D-ELAN module, which incorporates ELAN module, DcNets and BiFormer attention convolution, enabling a better focus on the detection of near shore ships and small ships. The positions of D-ELAN and B-ELAN inserted in backbone network are shown in Fig. 1.

2) *Head*: The network of DBW-YOLO head consists of spatial pyramid block (SP block), feature pyramid network (FPN) [25] structure and BiFormer block (BBL). SP block is utilized to extract features more comprehensively, preventing image distortion and redundant feature extraction. FPN structure is implemented to enable effective feature fusion, facilitating the identification of targets across multiple scales. These combined efforts notably improve the model's capacity to accurately identify and locate ships. Three BBL blocks, depicted in Fig. 1 at positions 3 are strategically added at the end of feature fusion. These additions improve the feature representation ability and enhance the overall performance and robustness of our method. Consequently, the feature maps at various scales are seamlessly integrated into the detection network.

3) *Additional Details*: The data augmentation techniques of mosaic and mixup methods are proposed to effectively improve the model's generalization and expand the datasets. WIoU is utilized as loss function to optimize the network, resulting in improved convergence speed and detection performance.

B. Components

The DBW-YOLO method is main consists of the following four components, YOLOv7-tiny, DCNets, BiFormer, and WIoU as follows.

1) *YOLOv7-Tiny*: YOLOv7 is one of the latest versions of the YOLO series detector introduced by Wang et al. [26]. Renowned for its efficient real-time target detection ability, it exhibits fast inference speed and low computational complexity, which makes it ideal for real-time applications and presents excellent performance with large datasets [27]. According to the different running environments of the code, including cloud GPU, normal GPU, and edge GPU, it is categorized into three models, namely

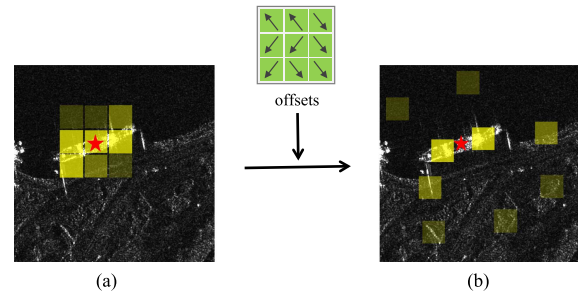


Fig. 2. DCNets mechanism. (a) Standard convolution with a regular sampling grid and (b) deformable convolution with sparse sampling.

YOLOv7-w6, YOLOv7, and YOLOv7-tiny [26]. Compared with YOLOv7 and YOLOv7-w6, YOLOv7-tiny possesses a more compact model size and lower computing requirements, making it ideal for scenarios where computing resources are limited. Moreover, YOLOv7-tiny often outperforms its counterparts in terms of processing speed, rendering it well-suited for scenarios requiring high real-time capabilities. Hence, YOLOv7-tiny is chosen as the foundational model for DBW-YOLO method in this article.

2) *Deformable Convolution Networks (DCNets)*: DCNets were initially introduced by Dai et al. [28]. In contrast to traditional convolutions, which employ fixed-shaped kernel, DCNets incorporate a 2-D offset in standard convolution, allowing network to flexibly adjust its shape, as shown in Fig. 2. The latest version of DCNets were presented by Wang et al. [29].

The feature extraction network of YOLOv7 employs 3 × 3 standard convolutions, sampling the input feature map at fixed positions. However, when dealing with near shore ships, convolutions tend to extract a significant amount of background information, thereby affecting the network's ability to recognize ship targets. In this section, to obtain valuable ship features, two characterization learning enhancement structure, DBL block and D-ELAN module are designed with the assistance of DCNets. This addresses the limitations of traditional convolutions in the context of near-shore ship targets. Through extensive experimental analysis and validation, this network structure proves to be more suitable for feature extraction in SAR ship targets.

The structural details of DBL block and D-ELAN module are illustrated in Fig. 3. By optimizing convolution for various ship targets, DCNets not only enhance the network's feature extraction capabilities but also significantly improve the efficiency and detection performance of model. The main aspects are given as follows.

- 1) To optimize the extraction ability of ship features, a DBL block is designed, which replaces the original 3×3 standard convolution in CBL block with DCNets. Then, we employ batch normalization to establish weight relationships, and Leaky ReLU activation function to capture more complex features.
- 2) To strengthen the feature learning capability and robustness of the network, a D-ELAN module is created by inserting two DBL blocks into the original ELAN module. These DBL blocks replace the corresponding original

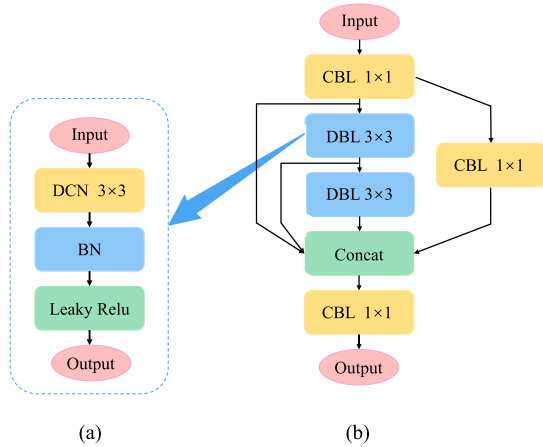


Fig. 3. Details of DCNets. (a) 3×3 DBL block, and (b) D-ELAN module.

CBL blocks. With this design, three ELAN modules in backbone network are replaced with D-ELAN modules to improve the overall performance.

- 3) To determine the appropriate expansion rate for DCNets, rigorous comparative experiments is conducted to obtain the optimal parameters.

3) *BiFormer*: For SAR target detection method, attention mechanisms are commonly incorporated into networks to enhance feature extraction. However, this often leads to significant resource consumption and computational burdens, posing challenges for model deployment and application. In this section, BiFormer attention mechanism [30], [31] is submitted to effectively extract image features. It utilize bilevel routing attention as core building blocks, which reduces huge computational burden and enables flexible calculations. With the applying of BiFormer attention mechanism, another two characterization learning enhancement structure, BBL block and B-ELAN module, are introduced to fully use the distinctive features of ship targets. This improves performance in recognizing features of small targets while simultaneously reducing the computational load. The main aspects are outlined as follows.

- 1) To enhance the convolution's capability to recognize ship features, a BBL block is introduced with BiFormer attention mechanism. Similar to DBL block, the 3×3 standard convolution in CBL block is substituted with a 3×3 BiFormer attention convolution in BBL block, and its detailed structure is shown in Fig. 4(a). As mentioned earlier, after replacing the original CBL block, this block is incorporated into head network at position 3 along with a 1×1 standard convolution.
- 2) To improve the accuracy of backbone network in recognizing near shore ships and small ships, a B-ELAN module is designed based on D-ELAN module and BiFormer attention convolution, as shown in Fig. 4(b). The last 1×1 standard convolution in D-ELAN module is replaced with a 1×1 BiFormer attention convolution, which is more beneficial to capture near shore ships and small ships.
- 3) B-ELAN module, in conjunction with D-ELAN module, collectively constitute the core components of backbone network. These collaborative enhancements notably

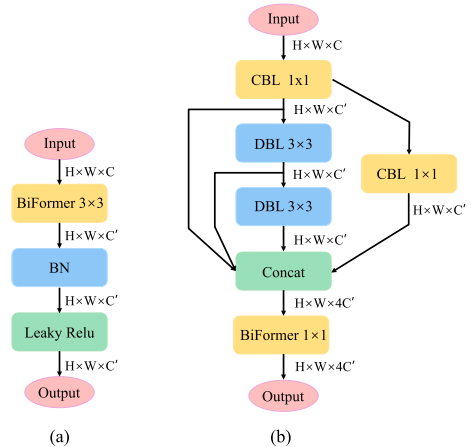


Fig. 4. Details of BBL. (a) BBL block, and (b) B-ELAN module.

improve the model's detection accuracy and overall performance.

4) *WIoU*: The bounding box regression (BBR) loss function holds significant importance in ship detection. Therefore, a WIoU loss function [32] is introduced to identify ships from surrounding background in this section. Unlike the traditional complete-IoU (CIoU) loss function [33], WIoU loss function introduces a dynamic nonmonotonic focusing mechanism, which improves both convergence speed and generalization capability. With the WIoU, the competition between high-quality anchors and the influence of harmful gradients from low-quality samples are significantly reduced in the calculation of target boxes. In terms of probability, WIoU improves the model's generalization. The BBR loss function \mathcal{L}_{IoU} is provided by

$$\mathcal{L}_{IoU} = 1 - IoU = 1 - \frac{W_i H_i}{S_i} \quad (1)$$

where S_i represents the area of predicted box that intersected with the actual box. H_i and W_i represent the height and the weight of the overlapping area, respectively.

In cases where $H_i = 0$ or $W_i = 0$, the existing BBR losses \mathcal{L}_i are designed to penalize \mathcal{R}_i and is denoted as

$$\mathcal{L}_i = 1 - IoU + \mathcal{R}_i. \quad (2)$$

To address the issue of gradient vanishing, researchers have made numerous meaningful explorations in various IoU structures. For example, the Generalized-IoU [34] structure incorporates a penalty term based on minimum bounding box structure. The Distance-IoU (DIoU) [35] structure introduces a penalty term founded on distance measurements, and the CIoU structure augments the DIoU structure with length-width ratio measurements. Especially the SCYLLA-IoU (SIoU) structure, as reported by Gevorgyan [36], exhibits enhanced convergence speed and superior performance by incorporating considerations of cost, distance, and shape.

An effective loss function should incorporate suitable geometric penalties when the predicted bounding box intersects with the ground truth bounding box. Therefore, minimizing intervention during training frequently serves as an effective strategy to upgrade the model's generalization ability. However, low-quality

TABLE I
DESCRIPTION OF EXPERIMENTAL SETUP

Component	Specifications
CPU	Intel(R) Xeon(R) Platinum 8350C CPU@2.60GHz
RAM	16GB
GPU	NVIDIA GeForce RTX 3090
GPU Memory	24GB
Operating System	Ubuntu 20.04
Language	Python 3.8
Framework	PyTorch 1.11.0

samples in the dataset are inevitably present during the training phase. The adverse effects of these samples are compounded by factors, such as aspect ratio and distance, contributing to a decline in the model's generalization capabilities. To tackle these challenges, Tong et al. [32] introduced a WIoU with a two-tier distance attention mechanism and the BBR loss function $\mathcal{L}_{\text{WIoU}_{v1}}$ is denoted as

$$\mathcal{L}_{\text{WIoU}_{v1}} = \mathcal{R}_{\text{WIoU}} \mathcal{L}_{\text{WIoU}} \quad (3)$$

$$\mathcal{R}_{\text{WIoU}} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 - H_g^2)^*}\right) \quad (4)$$

where H_g and W_g represent the height and weight of the smallest enclosing box, respectively, $\mathcal{R}_{\text{WIoU}}$ represents the relative weight of WIoU loss function, which is used to balance the relationship between the overlap degree of the predicted bounding box and the ground truth bounding box. To ensure that $\mathcal{R}_{\text{WIoU}}$ does not produce gradients hindering convergence, the disentanglement of H_g , W_g from the computational graph is represented by the superscript $*$.

To address the slow convergence issue of the loss function $\mathcal{L}_{\text{WIoU}_{v1}}$ during late training stages, a monotonic focus coefficient $\mathcal{L}_{\text{IoU}}^{\gamma^*} \in [0, 1]$ is introduced. Then, a new BBR loss function $\mathcal{L}_{\text{WIoU}_{v2}}$, is defined as follows:

$$\mathcal{L}_{\text{WIoU}_{v2}} = \mathcal{L}_{\text{IoU}}^{\gamma^*} \mathcal{L}_{\text{WIoU}_{v1}}, \mathcal{L}_{\text{IoU}}^{\gamma^*} = \left(\frac{\mathcal{L}_{\text{IoU}}^*}{\mathcal{L}_{\text{IoU}}}\right)^\gamma \quad (5)$$

where $\gamma > 0$, $\overline{\mathcal{L}_{\text{IoU}}}$ is the exponential running average with momentum m .

By introducing a nonmonotonic focus coefficient $\beta \in [0, +\infty)$, the last but most needed BBR loss function $\mathcal{L}_{\text{WIoU}_{v3}}$ is defined as

$$\mathcal{L}_{\text{WIoU}_{v3}} = r \mathcal{L}_{\text{WIoU}_{v1}}, r = \frac{\beta}{\delta \alpha^{\beta - \delta}} \quad (6)$$

where hyperparameters $\alpha > 0$, $\beta > 0$, and δ are chosen such that $r = 1$ when $\beta = \delta$. A low outlier degree, indicated by $\beta = \frac{\mathcal{L}_{\text{IoU}}^*}{\mathcal{L}_{\text{IoU}}}$, suggests that the anchor box is of high quality.

III. EXPERIMENT AND RESULTS

A. Experimental Environment

Table I provides a comprehensive overview of the experimental settings utilized in this article.

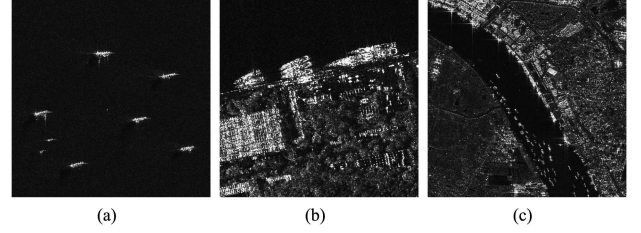


Fig. 5. Samples of HRSID dataset. (a) Multiscale ship samples, (b) inshore ship samples, and (c) small ship samples.

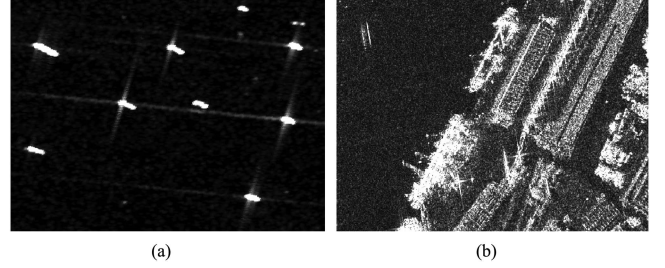


Fig. 6. Samples of Official-SSDD dataset. (a) Offshore ship samples, and (b) inshore ship samples.

B. Datasets

To assess the effectiveness of DBW-YOLO method, two SAR datasets, HRSID [37] and SSDD [38], are employed for comparative analysis.

1) *HRSID Dataset*: The HRSID dataset was presented by [37]. It consists of 5604 SAR images and 16 951 ships, and provides a panoramic view of SAR images with a 25% overlap rate and resolutions ranging from 1 to 5 m. Each image sized at 800×800 pixels. Due to its diverse representation of ship targets, the HRSID dataset is often utilized for multiscale SAR ship detection. Sample images extracted from the HRSID dataset are depicted in Fig. 5.

2) *SSDD Dataset*: The SSDD dataset was presented by [38]. It consists of 1160 images and 2456 ships collected from Sentinel-1, TerraSAR, and RadarSat-2, featuring resolutions range from 1 to 15 m. Each image has an approximate size of 600 pixels in length and width. To make full use of SSDD dataset, Zhang et al. [39] proposed an official version of SSDD, which includes specific criteria for usage. Sample images from the Official-SSDD dataset are illustrated in Fig. 6.

C. Performance Metrics

In assessing the effectiveness of the DBW-YOLO method, the following three key performance metrics are introduced: precision, recall, and average precision (AP). These metrics consists of the following four components: True Positive (TP), false positive (FP), true negative (TN), and false negative (FN). TP refers to positive sample correctly predicted as ships, while FP corresponds to the positive sample incorrectly predicted as ship. TN accounts for the negative sample correctly predicted as background, while FN refers to the negative sample incorrectly

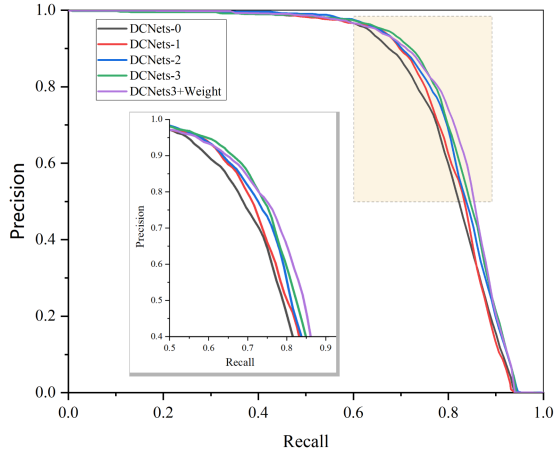


Fig. 7. AP curves of different expansion rates for DCNets.

predicted as background. Precision, denoted as P , represents the ratio of accurately predicted ships and is mathematically expressed as

$$P = \frac{TP}{TP+FP}. \quad (7)$$

The metric recall, denoted as R , represents the total number of retrieved ships and is given by

$$R = \frac{TP}{TP+FN}. \quad (8)$$

AP is computed based on precision and recall, which is articulated by the integral expression

$$AP = \int_0^1 P(R) dR. \quad (9)$$

Here, the function $P(R)$ signifies the precision–recall curve. For a ship detection method like DBW-YOLO, higher AP values indicate a superior overall performance of the ship detection method concerning precision and recall.

D. Experiment Comparison and Analysis

To estimate the performance of DBW-YOLO method, a series of experiments were carried out involving modifications to the activation function and various network structures of the comparative models. These results can further guide the optimization of the model.

In this section, we utilize the HRSID dataset for all experiments, maintaining uniform setups throughout: 200 epochs, the SGD optimizer, an initial learning rate of 0.01, a batch size of 64, and an input size of 640×640 pixels.

1) *Effect of DCNets*: To enhance the convolution’s capability in recognizing ship features, we incorporate an improved version of DCNets into the original ELAN blocks of YOLOv7-tiny. The performance of DCNets is significantly influenced by diverse expansion rates, as depicted in the AP curves illustrated in Fig. 7.

In the results shown in Fig. 7, DCNets-0 (represented by black curve) denotes the baseline model without DCNets, while

TABLE II
EFFECT OF DIFFERENT EXPANSION RATES ON DCNETS

Model	P (%)	R (%)	mAP@0.5 (%)	mAP@0.5:0.95 (%)
DCNets-0	86.31	72.97	81.21	51.73
DCNets-1	87.56	72.26	81.73	51.95
DCNets-2	88.41	71.71	82.21	52.38
DCNets-3	87.55	73.45	82.67	52.64
DCNets3+Weight	87.93	73.61	83.79	53.32

The bold values indicate the maximum value of the evaluation.

DCNets-1 (represented by red curve) corresponds to an expansion rate of (2, 4, 8, 16). In addition, DCNets-2 (represented by blue curve) and DCNets-3 (represented by green curve) represent expansion rates two and four times higher than DCNets-1, respectively. It is worth noting that DCNets-3+Weights (represented by purple curve) signifies DCNets-3 with pretrained weights, which is obtained by training 300 epochs on the FUSAR-ship dataset [40]. The pretraining weights refer to the learned weight parameters during the training process, which are used to adjust the model’s focus on input data. The comparative analysis reveals that the purple curve, representing DCNets-3+Weights, encompasses the largest area, indicating superior detection capability in comparison to other curves.

Table II presents the performance evaluation of DCNets under various expansion rates. The table displays precision, recall, mean Average Precision at IoU 0.5 (mAP@0.5), and mean Average Precision ranging from IoU 0.5 to 0.95 (mAP@0.5:0.95).

The original model, DCNets-0, achieves a precision of 86.31% and a recall of 72.97%, resulting in an mAP@0.5 of 81.21% and an mAP@0.5:0.95 of 51.73%. As we increase the expansion rate in DCNets-1, a slight improvement in precision and mAP@0.5 are observed, although the Recall remains similar. DCNets-2 exhibits a similar performance to DCNets-1. A notable performance enhancement is seen in DCNets-3, where recall, and mAP value show significant improvement over DCNets-0. However, the precision remains slightly increase. The most remarkable results are observed in the case of DCNets-3+Weights, where precision, recall, mAP@0.5, and mAP@0.5:0.95 reach their peak values, 87.93%, 73.61%, 83.79%, and 53.32%, respectively. These findings suggest that higher expansion rates, particularly in conjunction with pretrained weights, have a substantial positive impact on the detection performance of DCNets, enhancing both precision and recall, as indicated by the increased mAP values.

2) *Effect of BiFormer*: In order to enhance the FPN, the BiFormer attention mechanism is strategically integrated at different points within the original YOLOv7-tiny model. These specific insertion positions for the BiFormer attention mechanism are highlighted within dashed boxes in Fig. 1. Fig. 8 illustrates AP curves of the comparative models under different conditions, and Table III provides their comparative results in detail.

Results in Fig. 8 reveal that the original YOLOv7-tiny model (depicted by black curve) exhibits the smallest area under AP curve, indicating poor detection performance. When attention

TABLE III
ABLATION EXPERIMENTS FOR BiFORMER

Model	Position 1	Position 2	Position 3	P (%)	R (%)	mAP@0.5 (%)	mAP@0.5:0.95 (%)
YOLOv7-tiny				87.93	73.61	83.79	53.32
YOLOv7-tiny+1+2	✓	✓		87.69	77.01	85.17	53.53
YOLOv7-tiny+1+3	✓		✓	87.28	76.90	84.53	53.45
YOLOv7-tiny+2+3		✓	✓	90.98	76.31	86.58	56.29
YOLOv7-tiny+1+2+3	✓	✓	✓	87.95	77.81	85.46	54.68

The bold values indicate the maximum value of the evaluation.

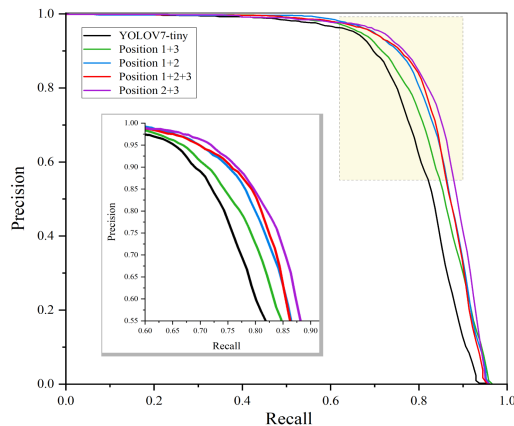


Fig. 8. AP curves of ablation experiments for BiFormer.

mechanisms are introduced at various positions, the area under AP curve significantly increases for position 1+3 (depicted by green curve), position 1+2 (depicted by blue curve), position 2+3 (depicted by purple curve), and position 1+2+3 (depicted by red curve). These findings demonstrate a substantial enhancement in detection performance achieved by the attention mechanisms. Notably, the purple curve representing position 2+3 encloses the largest area, indicating a superior detection capability compared to the others.

More experimental details are presented in Table III. The findings reveal that, in comparison with the original model, the precision of position 2+3 significantly increases to 90.98%, Recall to 76.31%, mAP@0.5 to 86.58%, and mAP@0.5:0.95 to 56.29%, while the performance of Position 1+2, Position 1+3, and Position 1+2+3 shows slight improvement. These indicate that excessive use of BiFormer attention mechanism can potentially diminish the efficacy of feature extraction and detection capabilities. For the DBW-YOLO method, placing the BiFormer attention mechanism at position 2+3 produces the optimal detection performance.

Besides, to assess the effectiveness of BiFormer model when inserted at position 2+3, a comparative analysis was performed involving three alternative attentional mechanism models, squeeze-and-excitation (SE) [41], coordinate attention (CA) [42], and convolutional block attention module (CBAM) [43], all integrated at the same positions. Detailed experimental findings are outlined in Table IV.

The performance analysis showcases the BiFormer model's outstanding results, with a precision of 90.98%, a recall

TABLE IV
COMPARATIVE RESULTS OF DIFFERENT ATTENTION MECHANISM MODELS

Model	P (%)	R (%)	mAP@0.5 (%)	mAP@0.5:0.95 (%)
CBAM	87.32	73.93	84.03	54.66
CA	86.11	76.88	84.17	54.25
SE	89.90	74.09	84.94	53.61
BiFormer	90.98	76.31	86.58	56.29

The bold values indicate the maximum value of the evaluation.

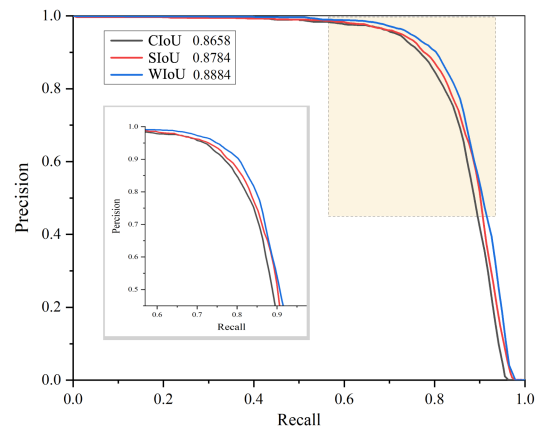


Fig. 9. AP curves of different BBR loss function.

of 76.31%, an excellent mAP@0.5 of 86.58%, and a high mAP@0.5:0.95 of 56.29%. The precision of BiFormer model has witnessed substantial enhancements of 3.66%, 4.87%, and 1.08% in comparison to CBAM, CA, and SE models, respectively. For mAP@0.5, it has been increased by 2.55%, 2.41%, and 1.64%, respectively. These findings highlight the substantial improvement in detection performance achieved by BiFormer attention mechanism compared to other evaluated models. They emphasize the effectiveness of BiFormer attention mechanism, which improves the overall performance concerning Precision and mAP.

3) *Effect of Loss Functions*: Following the enhancement of FPN and backbone, this section analyzes the comparison results of loss functions. We setup three BBR loss functions, CIoU [33], SIoU [36], and WIoU [32], for comparative analysis with the established model. Results in Fig. 9 reveal that the measured mAP@0.5 for WIoU is 88.84%, which is increased by 2.26% and 1.00%, respectively, compared with that of CIoU and SIoU. This proves that WIoU exhibits an excellent performance in terms of detection accuracy.

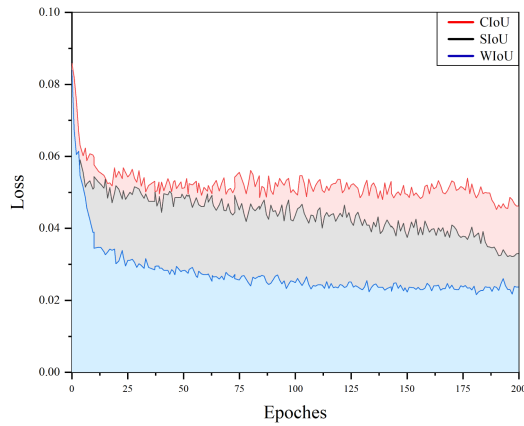


Fig. 10. Convergence of different BBR loss function.

TABLE V
PERFORMANCE EVALUATION OF VARIOUS MODELS ON THE HRSID DATASET

Model	Dataset	P (%)	R (%)	mAP@0.5 (%)
SSD	HRSID	92.05	64.31	84.60
Faster R-CNN	HRSID	86.20	63.96	74.78
YOLOv5-s	HRSID	86.42	76.14	84.25
YOLOv7-tiny	HRSID	88.41	71.71	81.41
DBW-YOLO	HRSID	90.04	80.47	88.84

The bold values indicate the maximum value of the evaluation.

TABLE VI
PERFORMANCE EVALUATION OF VARIOUS MODELS ON THE SSDD DATASET

Model	Dataset	P (%)	R (%)	mAP@0.5 (%)
SSD	SSDD	94.83	82.41	92.79
Faster R-CNN	SSDD	89.09	75.32	87.53
YOLOv5-s	SSDD	97.39	88.71	96.53
YOLOv7-tiny	SSDD	89.65	88.68	94.30
DBW-YOLO	SSDD	98.41	97.64	99.18

The bold values indicate the maximum value of the evaluation.

To assess other aspects of the performance of WIoU loss function, the convergence curve of different BBR loss function is also studied in this section, as shown in Fig. 10. The findings suggest that WIoU loss function achieves the lowest loss value and the fastest convergence speed, making DBW-YOLO particularly effective in ship detection within SAR images. This underlines the superior detection performance achieved by WIoU in terms of convergence speed and generalization capability.

4) *Comparison With Other Ship Detection Models:* For the purpose of evaluating strong detection capabilities of DBW-YOLO on SAR images, we conduct a comparative analysis with four widely-used SAR ship detection models, SSD, Faster R-CNN, YOLOv5-s, and YOLOv7-tiny, with identical conditions and parameter settings. Two datasets, HRSID and SSDD, are utilized to evaluate performance. The comparative outcomes are presented in Tables V and VI, and a revised performance comparison of DBW-YOLO and other models on the HRSID dataset is shown in Fig. 11, and a similar comparison on the SSDD dataset is presented in Fig. 12.

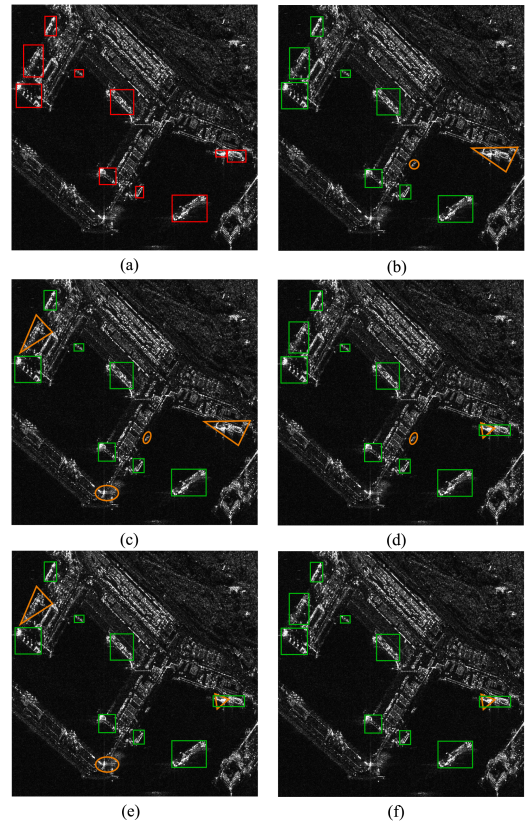


Fig. 11. Revised comparison chart: Performance of DBW-YOLO and other models on the HRSID dataset. (a) Original label image, (b) SSD visualization results, (c) Faster R-CNN visualization results, (d) YOLOv5-s visualization results, (e) YOLOv7-tiny visualization results, and (f) DBW-YOLO visualization results.

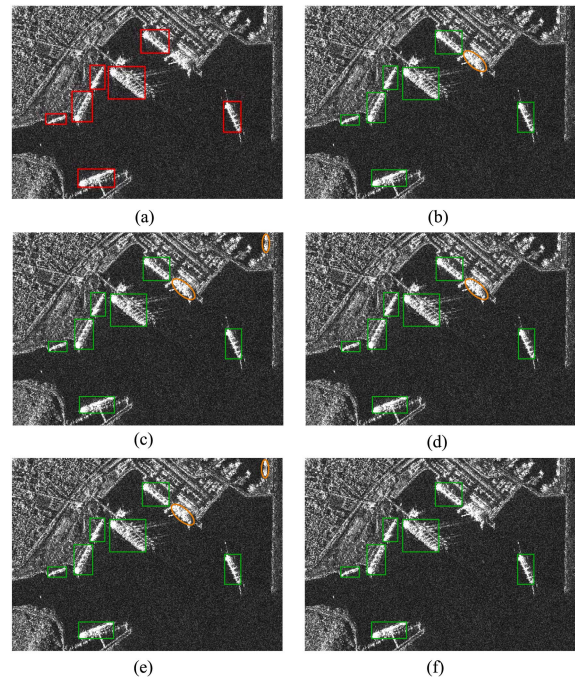


Fig. 12. Revised comparison chart: Performance of DBW-YOLO and other models on the SSDD dataset. (a) Original label image. (b) SSD visualization results. (c) Faster R-CNN visualization results. (d) YOLOv5-s visualization results. (e) YOLOv7-tiny visualization results. (f) DBW-YOLO visualization results.

Table V provides performance comparison of various models on the HRSID dataset. The results reveal that DBW-YOLO model exhibits the highest precision (90.04%), while SSD shows the highest recall (64.31%). DBW-YOLO also achieves the best $mAP@0.5$ performance, with a value of 88.84%. Notably, it exhibits superior performance in terms of precision and $mAP@0.5$.

Table VI presents a performance comparison of the same models on the SSDD dataset. On this dataset, DBW-YOLO achieves the highest precision (98.41%), recall (97.64%), and $mAP@0.5$ (99.18%), surpassing the other models. This result highlights the superior performance of DBW-YOLO in detecting ships on the SSDD dataset.

The findings presented in Tables V and VI highlight the superior performance of our method across both the HRSID and SSDD datasets, surpassing that of other models. This can be attributed to its improved backbone and well-suited FPN, which allows for the efficient extraction of ship characteristics, thereby improving the model's prediction performance. Furthermore, the utilization of WIoU loss function enhances the accuracy of target box predictions, which leads to improved performance of generalization ability and robustness.

In order to better demonstrate the performance of each model on various datasets, the accuracy performance of the same model on different datasets can be obtained by comparing the results in Tables V and VI. It indicates that the precision, recall, and mAP values of all five models on the SSDD dataset surpass those on the HRSID dataset. The discrepancy between two datasets can be attributed to the SSDD dataset's reputation for high-quality ship annotations, which ensures accurate ship labeling. Moreover, the dataset usually comprises a balanced distribution of positive and negative examples, where the number of ship and nonship instances is relatively equal, which contributes to improved model training and performance. Conversely, the HRSID dataset presents heightened challenges as it includes high-resolution SAR images containing finer details and potential noise.

Fig. 11 presents a revised performance comparison of DBW-YOLO and other models on the HRSID dataset, where Fig. 11(a) depicts the original ground truth image, Fig. 11(b) illustrates SSD visualization results, Fig. 11(c) displays Faster R-CNN visualization results, Fig. 11(d) exhibits YOLOv5-s visualization results, Fig. 11(e) demonstrates YOLOv7-tiny visualization results, and Fig. 11(f) showcases DBW-YOLO visualization results. In the label image, ten ship targets are visible, including three small ships. The Faster R-CNN model correctly detected seven ships (indicated by green rectangle) but produced two FP detections (indicated by yellow circles) and missed two ships (indicated by yellow triangles). The YOLOv7-tiny model correctly detected eight ships, but had one FP detection and missed two ships. These two models exhibited the lowest detection accuracy, as they failed to identify near shore ships and misclassified non-ship objects. The SSD model correctly detected eight ships but had one FP detection and one missed detection. The YOLOv5-s model correctly detected nine ships but had one FP detection and one missed detection. Overall, these two models achieved slightly higher detection accuracy than the previous two. However, small targets were incorrectly labeled in all four of the preceding models. The model presented in this paper effectively detected and correctly labeled all nine ship

targets, with only one ship missed. Consequently, the proposed model demonstrated superior capabilities in detecting near shore and small ships compared to the other four comparison models, indicating its suitability for ship detection in HRSID dataset.

Fig. 12 presents a revised performance comparison of DBW-YOLO and other models on the SSDD dataset, where Fig. 12(a) depicts the original ground truth image, and Fig. 12(b)–(f), respectively, showcases the visualization results of SSD, Faster R-CNN, YOLOv5-s, YOLOv7-tiny, and DBW-YOLO. In the label image, seven ship targets are visible, and all of them are inshore ships. The results show that despite occasional FP detections (illustrated by yellow circles), all models correctly identified seven actual ships (illustrated by green rectangles), this also shows that the SSDD dataset has better adaptability and recognition accuracy compared to HRSID dataset. Among them, the Faster R-CNN model correctly detected seven ships but produced two FP detections. The YOLOv7-tiny model had the same detection results as the Faster R-CNN model. These two models exhibited the lowest detection accuracy for misclassifying non-ship targets near the shore as ships. The SSD model precisely identified seven ships but registered one FP detection. Similarly, the YOLOv5-s model demonstrated detection results equivalent to those of the SSD model. Overall, these two models achieved slightly higher detection accuracy than their predecessors. The model presented in this paper effectively detected and labeled seven ship targets, with no ship missed or false detections. Consequently, the proposed model displays superior capabilities in near shore detection compared to the other four models, underscoring its suitability for SAR-based ship detection in SSDD dataset.

In summary, the performance of the DBW-YOLO is noteworthy, as it consistently outperforms other models on both the HRSID and SSDD datasets for complex environments, which demonstrates its effectiveness in SAR ship detection.

IV. CONCLUSION

This paper introduces DBW-YOLO, a detection method designed for SAR ship detection. It exhibits good performance in the detection of complex environments such as near shore ships and small ships. The performance of DBW-YOLO was compared with other widely-used ship detection methods, namely SSD, Faster R-CNN, YOLOv5-s, and YOLOv7-tiny, using the HRSID and SSDD datasets. Specifically, the mAP of DBW-YOLO reached 88.84% on the HRSID dataset, and 99.18% on the SSDD dataset, respectively, which proves the superiority of the DBW-YOLO method in terms of detection accuracy. Additionally, the Precision and Recall of DBW-YOLO show significant improvements compared to the other models, which proves that our method obtains superior performance in terms of robustness. Overall, DBW-YOLO exhibits strong performance, particularly in detecting ships near the shore and small vessels in SAR images, making it highly suitable for tasks demanding high-precision detection capabilities. However, we also recognize that DBW-YOLO method suffers from the issue of excessive computational burden and insufficient detection accuracy in resource-constrained environments. Therefore, we will pay

more attention to the lightweight design of the model and the anchor free algorithm in future work.

REFERENCES

- [1] J. Li, J. Chen, P. Cheng, Z. Yu, L. Yu, and C. Chi, "A survey on deep-learning-based real-time SAR ship detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, no. 3, pp. 3218–3247, Mar. 2023.
- [2] Y. Yuan, Y. Wu, P. Feng, Y. Fu, and Y. Wu, "Segmentation-guided semantic-aware self-supervised denoising for SAR image," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, no. 10, Oct. 2023, Art. no. 5218416.
- [3] Y. Yuan, Y. Wu, Y. Fu, Y. Wu, L. Zhang, and Y. Jiang, "An advanced SAR image despeckling method by bernoulli-sampling-based self-supervised deep learning," *Remote Sens.*, vol. 13, no. 18, 2021, Art. no. 3636.
- [4] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208–216, Jan. 1992.
- [5] O. Pappas, A. Achim, and D. Bull, "Superpixel-level CFAR detectors for ship detection in SAR imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 9, pp. 1397–1401, Sep. 2018.
- [6] W. Li, H. Liu, Y. Wang, Z. Li, Y. Jia, and G. Gui, "Deep learning-based classification methods for remote sensing images in urban built-up areas," *IEEE Access*, vol. 7, pp. 36274–36284, 2019.
- [7] W. Li et al., "Classification of high-spatial-resolution remote sensing scenes method using transfer learning and deep convolutional neural network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, no. 5, pp. 1986–1995, May 2020.
- [8] Y. Wu, H.-N. Dai, and H. Tang, "Graph neural networks for anomaly detection in industrial Internet of Things," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 9214–9231, Jun. 2022.
- [9] H. Cao, H. Tang, Y. Wu, F. Wang, and Y. Xu, "On accurate computation of trajectory similarity via single image super-resolution," in *Proc. Int. Joint Conf. Neural Netw.*, 2021, pp. 1–9.
- [10] Z. Tian, C. Shen, H. Chen, and T. He, "FCOS: Fully convolutional one-stage object detection," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2019, pp. 9627–9636.
- [11] H. Zhang, Y. Wang, F. Dayoub, and N. Sunderhauf, "Varifocalnet: An iou-aware dense object detector," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 8514–8523.
- [12] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 779–788.
- [13] W. Liu et al., "SSD: Single shot multibox detector," in *Proc. 14th Eur. Conf. Comput. Vis.*, 2016, pp. 21–37.
- [14] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 580–587.
- [15] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 1440–1448.
- [16] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.
- [17] Z. Wang, L. Du, J. Mao, B. Liu, and D. Yang, "SAR target detection based on SSD with data augmentation and transfer learning," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 1, pp. 150–154, Jan. 2019.
- [18] G. Tang, Y. Zhuge, C. Claramunt, and S. Men, "N-Yolo: A SAR ship detection using noise-classifying and complete-target extraction," *Remote Sens.*, vol. 13, no. 5, 2021, Art. no. 871.
- [19] Z. Sun, X. Leng, Y. Lei, B. Xiong, K. Ji, and G. Kuang, "BiFA-YOLO: A novel YOLO-based method for arbitrary-oriented ship detection in high-resolution SAR images," *Remote Sens.*, vol. 13, no. 21, 2021, Art. no. 4209.
- [20] Y. Li, W. Zhu, C. Li, and C. Zeng, "SAR image near-shore ship target detection method in complex background," *Int. J. Remote Sens.*, vol. 44, no. 3, pp. 924–952, 2023.
- [21] S. Yang, W. An, S. Li, G. Wei, and B. Zou, "An improved FCOS method for ship detection in SAR images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, no. 10, pp. 8910–8927, Oct. 2022.
- [22] L. Bai, C. Yao, Z. Ye, D. Xue, X. Lin, and M. Hui, "Feature enhancement pyramid and shallow feature reconstruction network for SAR ship detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, no. 1, pp. 1042–1056, Jan. 2023.
- [23] Z. Chen, C. Liu, V. Filaretov, and D. Yuhimets, "Multi-scale ship detection algorithm based on YOLOv7 for complex scene SAR images," *Remote Sens.*, vol. 15, no. 8, 2023, Art. no. 2071.
- [24] X. Ren, Y. Bai, G. Liu, and P. Zhang, "YOLO-Lite: An efficient lightweight network for SAR ship detection," *Remote Sens.*, vol. 15, no. 15, 2023, Art. no. 3771.
- [25] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 2117–2125.
- [26] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2023, pp. 7464–7475.
- [27] Q. Guo, J. Liu, and M. Kaliuzhnyi, "YOLOX-SAR: High-precision object detection system based on visible and infrared sensors for SAR remote sensing," *IEEE Sensors J.*, vol. 22, no. 17, pp. 17243–17253, Sep. 2022.
- [28] J. Dai et al., "Deformable convolutional networks," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 764–773.
- [29] W. Wang et al., "InternImage: Exploring large-scale vision foundation models with deformable convolutions," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2023, pp. 14408–14419.
- [30] L. Zhu, X. Wang, Z. Ke, W. Zhang, and R. W. Lau, "BiFormer: Vision transformer with bi-level routing attention," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2023, pp. 10323–10333.
- [31] A. Vaswani et al., "Attention is all you need," in *Proc. 31st Int. Conf. Neural Inf. Process. Syst.*, 2017, pp. 6000–6010.
- [32] Z. Tong, Y. Chen, Z. Xu, and R. Yu, "Wise-IoU: Bounding box regression loss with dynamic focusing mechanism," 2023, *arXiv:2301.10051*.
- [33] Z. Zheng et al., "Enhancing geometric factors in model learning and inference for object detection and instance segmentation," *IEEE Trans. Cybern.*, vol. 52, no. 8, pp. 8574–8586, Aug. 2022.
- [34] H. Rezaatofghi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese, "Generalized intersection over union: A metric and a loss for bounding box regression," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 658–666.
- [35] Z. Zheng, P. Wang, W. Liu, J. Li, R. Ye, and D. Ren, "Distance-IoU loss: Faster and better learning for bounding box regression," in *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 07, 2020, pp. 12993–13000.
- [36] Z. Gevorgyan, "SIoU loss: More powerful learning for bounding box regression," 2022, *arXiv:2205.12740*.
- [37] S. Wei, X. Zeng, Q. Qu, M. Wang, H. Su, and J. Shi, "HRSID: A high-resolution SAR images dataset for ship detection and instance segmentation," *IEEE Access*, vol. 8, pp. 120234–120254, 2020.
- [38] J. Li, C. Qu, and J. Shao, "Ship detection in SAR images based on an improved faster R-CNN," in *Proc. SAR Big Data Era: Models, Methods Appl.*, 2017, pp. 1–6.
- [39] T. Zhang et al., "Sar ship detection dataset (SSDD): Official release and comprehensive data analysis," *Remote Sens.*, vol. 13, no. 18, 2021, Art. no. 3690.
- [40] X. Hou, W. Ao, Q. Song, J. Lai, H. Wang, and F. Xu, "FUSAR-Ship: Building a high-resolution SAR-AIS matchup dataset of Gaofen-3 for ship detection and recognition," *Sci. China Inf. Sci.*, vol. 63, pp. 1–19, 2020.
- [41] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 7132–7141.
- [42] Q. Hou, D. Zhou, and J. Feng, "Coordinate attention for efficient mobile network design," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 13713–13722.
- [43] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 3–19.



Xiao Tang received the B.S. degree in physics from Lanzhou University, Lanzhou, China, in 2012 and the M.S. and Ph.D. degrees in optical engineering from the Changchun Institute of Optics, Fine Mechanics and Physics, University of Chinese Academy of Sciences, Changchun, China, in 2014 and 2017, respectively.

He is currently an Associate Professor with the University of South China, Hengyang, China. His current research includes object detection for synthetic aperture radar (SAR) images, remote sensing satellite calibration, remote sensing big data and error analysis, and Internet of Things (IoT) remote sensing.



Jiufeng Zhang received the B.S. degree in mechanical design, manufacturing and automation from the School of Mechanical and Manufacturing Automation, Xi'an University, Xi'an, China, in 2019. He is currently working toward the M.S. degree in mechanical engineering with the School of Mechanical Engineering, University of South China, Hengyang, China.

His research interests include object detection for synthetic aperture radar (SAR) images, digital image processing, and deep learning in remote sensing applications.



Huanlin Xiao received the B.S. degree in mechanical design, manufacturing and automation from the Changsha University of Science and Technology, Changsha, China, in 2021. He is currently working toward the M.S. degree in mechanical engineering with the School of Mechanical Engineering, University of South China, HengYang, China.

His research interests include remote sensing satellite calibration, remote sensing big data and error analysis, and deep learning in remote sensing applications.



Yunzhi Xia received the M.S. degree in signal and information processing from the School of Information and Communication, Guilin University of Electronic Technology, Guilin, China, in 2014. She is currently working toward the Ph.D. degree in cyber science and engineering with the School of Cyber Science and Engineering, Huazhong University of Science and Technology (HUST), Wuhan, China.

She is currently an Assistant Professor with the University of South China, Hengyang, China. Her research interests include object detection for synthetic aperture radar (SAR) images, deep learning in remote sensing applications, and Internet of Things (IoT) Remote Sensing.