Discrete Edge Feature Guided Rotation Detection Method for Remote Sensing Ship Wake

Junfei Chen, Zhuhua Hu, *Senior Member, IEEE*, Yaochi Zhao, Wei Wu, Lihang Chen, Yueming Hu, and Ba Huang

***Abstract*—Ship wake rotation detection technology based on remote sensing images has a wide range of application value in ship supervision, trajectory prediction, and maritime rescue. However, objects in natural scenes usually present continuous and smooth features, while the contours and textures of ship wakes appear as discrete lines. This makes the wake have strong edge (discontinuity of image feature distribution) features but lack texture information, so the general detection model is difficult to perceive the wake completely, and low-dimensional edge features may also be ignored in deep networks. In addition, the significant directionality of wake also puts forward higher requirements for the accuracy of angle prediction. To address these issues, an efficient wake rotation detection algorithm is proposed in this study. First, we construct a discrete edge focused feature extraction module. This module features parallel edge branches that capture spatial features of the image while sensing changes in pixel gradient intensity, thereby obtaining a richer feature representation. Second, we construct an edge guided context attention fusion module. It uses the edge information in low-dimensional features as a guide to adaptively fuse contextual features, effectively enhancing the model's global perception capability. Finally, we design a diagonal constrained probabilistic intersection over union. It further enhances the description of rotation angle and shape by considering the relative diagonal distance of minimum external horizontal rectangle, especially when dealing with square like objects. Experimental results show that our algorithm achieves accurate and stable ship wake detection, and has excellent robustness in weak objects and complex backgrounds. The mean average precision of this method for optical remote sensing wake detection reaches 95.1%, and the average error of angle prediction is reduced from 15.35° to 7.16°. In addition, the algorithm also achieves an excellent performance in general remote sensing scenarios, which verifies its excellent adaptability and generalization.**

***Index Terms*—Context information, feature fusion, remote sensing image, rotating object detection, ship wake.**

Received 4 December 2024; revised 15 March 2025; accepted 9 April 2025. Date of publication 11 April 2025; date of current version 2 September 2025. This work was supported in part by the Key Research and Development Project of Hainan Province under Grant ZDYF2022GXJS348 and Grant ZDYF2022SHFZ039, in part by the National Natural Science Foundation of China under Grant 62361024 and Grant 62161010, in part by the Graduate Innovative Talents Training Program of Hainan University under Grant HDJXAL2024Y0007, in part by the Research Project on Education and Teaching Reform of Higher Education in Hainan Province under Grant Hnjg2024-11, and in part by the Education and Teaching Reform Research Project of Hainan University under Grant hdjy2421. *(Corresponding authors: Zhuhua Hu; Yueming Hu.)*

Junfei Chen, Zhuhua Hu, Wei Wu, and Lihang Chen are with the School of Information and Communication Engineering, Hainan University, Haikou 570228, China (e-mail: 22220854000129@hainanu.edu.cn; eagler_hu@hainanu.edu.cn; wuweido@126.com; 23210810000036@hainanu.edu.cn).

Yaochi Zhao is with the School of Cyberspace Security, Hainan University, Haikou 570228, China (e-mail: zhyc@hainanu.edu.cn).

Yueming Hu is with the College of Tropical Agriculture and Forestry, Hainan University, Haikou 570228, China (e-mail: yueminghugis@163.com).

Ba Huang is with the Sanya Institute of South China Sea Geology, Guangzhou Marine Geological Survey, Sanya 572025, China, and also with the Academy of South China Sea Geological Science, China Geological Survey, Sanya 572025, China (e-mail: 273140684@qq.com).

Digital Object Identifier 10.1109/JSTARS.2025.3560200

I. Introduction

REMOTE sensing object detection technology plays an important role in ship monitoring, which has wide application value in ship management, auxiliary law enforcement, and monitoring of fishery activities [1], [2]. Remote sensing images can cover a wide area of the sea, so as to effectively sense ship activities. The object detection technology can quickly and accurately identify the object in the image, providing a powerful tool for ship monitoring. In recent years, with the improvement of satellite resolution and the progress of computer vision algorithms, the parsing ability of remote sensing images has been significantly improved [3]. However, due to the complexity of the marine environment and the tiny size of ships in remote sensing images, there are still challenges in directly detecting ships [4], [5].

In remote sensing images, ships often occupy fewer pixels and are difficult to identify accurately. In contrast, the wake generated by the ship sailing occupies a larger area in the remote sensing image and is easier to detect than the ship itself [6], [7]. Ship wake is a disturbance of water surface caused by ship navigation, and its traces can be observed in remote sensing images [8]. Therefore, it has become an effective method to monitor ships by detecting ship wakes in remote sensing images.

Tail flow detection does not rely on the image information of the ship itself, and can provide effective supplements in situations where the ship target itself is small, blurry, obstructed, and difficult to directly identify the ship

Wake detection does not rely on the image information of the ship itself, and can provide an effective supplement when the ship itself is small, fuzzy, occluded, and other circumstances that are difficult to directly identify the ship [9], [10]. On the basis of detecting ship wake, rotation detection technology can further improve the accuracy of detection and provide the possibility to analyze the morphological characteristics of ship wake. After the wake is detected by the rotation detection method, the direction, length and angle of the wake can be preliminarily judged according to the angle and shape of the prediction box. They can further infer information such as the ship's heading

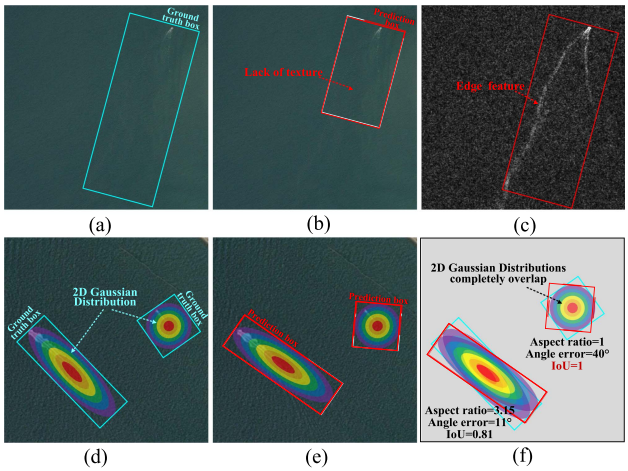

Fig. 1. Qualitative examples of wake detection scenarios. (a) and (b) Images with ground truth and predicted bounding boxes, respectively. (c) Image after edge operator processing, which can reveal subtle wakes. (d) and (e) Images with ground truth and predicted bounding boxes, respectively, while (f) overlays them to compare the overlap of 2-D Gaussian distributions. Notably, the object with a quasi-square shape maintains an IoU of 1 regardless of the angular error.

and speed, and provide important data support for subsequent ship tracking and trajectory prediction [11], [12].

However, there are still some difficulties in the rotation detection of the remote sensing wake, and we describe them in Fig. 1. As shown in Fig. 1(a) and (b), it is difficult for the model to detect the entire wake completely due to the thin wake and lack of texture. In addition, the additional angle parameter produces problems such as boundary discontinuity and complex loss calculation, which are well solved by the method based on 2-D Gaussian distribution [13]. But as shown in Fig. 1(f), if there are two square boxes of the same size as the center point, regardless of the angle difference between them, the intersection over union (IoU) calculated by the 2-D Gaussian distribution is always 1.

In summary, when there is a lot of sea clutter or insufficient wake energy, the wake of a ship is not obvious, discontinuous, and lacks texture. For low aspect ratio wake, 2-D Gaussian distribution cannot accurately determine whether the angle predicted by the model is accurate. In response to these issues, we design a feature extraction module to address strong edges (where pixel information undergoes abrupt changes), a fusion module to recombine low-dimensional and high-dimensional features, and an IoU calculation method with additional constraints.

The main contributions of this article are as follows.

1) We propose discrete edge focused feature extraction module (DEF). This is a multibranch structure that perceives gradient changes of discrete edges through the edge branch while simultaneously capturing spatial features through the convolutional branch. These branches' extracted features are then aggregated to obtain a richer image representation, thereby helping the model better understand the content of images with sparse textures and discontinuities.
2) We propose edge guided context attention fusion module (EGCA). It constructs a weight guidance mechanism through edge operators and attention, and uses this mechanism to guide the weighted reorganization of low-dimensional and high-dimensional features. This forces the model to focus on important edge and gradient features in the image, and improves the model's global perception and robustness against interference.
3) We design diagonal constrained probabilistic IoU (DPIoU). It uses the minimum enclosing horizontal rectangle of a rotated box to describe the rotation and employs its relative diagonal distance to constrain the 2-D Gaussian distribution's IoU method. This improves the model's accuracy in angle prediction. In addition, we design a shape-based dynamic weighting function to focus this constraint on challenging low aspect ratio objects.

The rest of this article is organized as follows. In Section II, we present previous related research in this area. In Section III, we elaborate on the proposed model algorithm. We provide and discuss numerical results in Section IV to verify the effectiveness of the proposed algorithm. Finally, Section V concludes this article.

II. Related Work

Related research work can be divided into three aspects: remote sensing object detection algorithm based on deep learning, ship wake detection algorithm, and rotating object detection algorithm. The following is a summary of the relevant research.

A. Remote Sensing Object Detection Algorithm Based on Deep Learning

With the proposal of convolutional neural network (CNN), many scholars apply deep learning to the research of remote sensing object detection [14], [15]. The current remote sensing object detector is mainly improved from the general detector. Classical two-stage detectors include R-CNN [16] and Fast R-CNN [17]. Classic single-stage detectors include SSD [18] and YOLO [19].

Remote sensing image has the characteristics of small object and complex background, and many improved models have been proposed according to these characteristics. In terms of small object detection, Fan et al. [20] proposed a detection model for remote sensing small objects, which is based on deep large kernel convolution to effectively extract features of small objects. Wu et al. [21] proposed UIU Net, which achieves multilevel and multiscale feature learning through nested U-Net, effectively improving the ability to extract contrast information from small targets. Tan et al. [22] realized rotation detection by designing a pseudodirected label generation algorithm to obtain OBB pseudolabels. In terms of complex backgrounds, Li et al. [23] effectively improved the detection performance of ship targets by embedding pyramid squeeze attention to focus on key features. Li et al. [24] proposed LRR Net by combining low rank representation models with deep learning, achieving efficient detection of anomalous pixels. These methods are based on the general detection model, and then improved according to the characteristics of remote sensing images.

B. Ship Wake Detection Algorithm

Traditional ship wake detection algorithms mainly rely on image processing technologies, such as histogram of oriented gradients and morphological processing. But these methods often do not perform well when dealing with images with complex backgrounds and noise. With the rapid development of deep learning, the wake detection algorithm based on CNN has gradually become the mainstream.

In the traditional ship wake detection algorithm, Biondi et al. [8] separated low-rank components from sparse components to reduce sea surface clutter, and supplemented by Radon transform to detect sparse objects, accurately estimate wake inclination angle and Kelvin spectrum analysis. Similarly, Ai et al. [1] enhanced data and realized ship wake detection through SCR enhancement and normalized Hough transform. Jiang et al. [2] used polarization enhancement to preprocess simulated SAR image scenes to improve ship wake detection. These traditional methods are strict in parameter setting and have poor adaptability.

In the method of deep learning, Liu et al. [7] used remote sensing images to locate and classify ships and wakes. They quickly locate the candidate hulls through the phase spectrum of the Fourier transform, and then classify the ships using a fuzzy classifier that combines hull and wake information. Liu et al. [9] proposed a new technique for detecting ship wake from remote sensing images. They first obtained the candidate wake by normalized Radon transformation of hull-centered images, and then removed the false wake by pixel value verification, turbulent wake identification, angle verification, and contrast verification. These methods based on deep learning greatly improve the detection accuracy of wake.

For the selection of image categories, there are not many SAR wake samples available for researchers to use. Del et al. [25] first used labeled SAR datasets to train deep learning detectors, producing only a dataset of about 250 wake samples. In the practice of SAR wake detection based on CNN, Kang et al. [10] only had hundreds of wake samples. Compared with SAR satellites, optical satellites have the advantages of more publicly available data, higher coverage, and fewer revisit cycles. Xue et al. [26] produced a wake dataset called SWIM based on optical remote sensing images, which contains 11 600 images and 15 356 wake instances. It provides powerful data support for the study of wake based on optical image.

C. Rotating Object Detection Algorithm

Traditional detectors mainly detect objects surrounded by horizontal bounding box (HBB). When faced with objects of different sizes, proportions, and orientations, the predicted HBB often contains many unnecessarily complex backgrounds that affect detection performance. Therefore, the rotation detector is more suitable for fine detection, and the most commonly used rotation detection methods include rotation invariance and loss function (as shown in Fig. 2).

The method based on rotation invariance is to detect the object by extracting its rotation invariance feature. First, scholars have studied the method of directional anchor. Liu et al. [27] defined

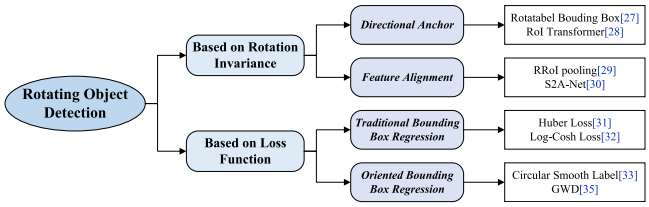

Fig. 2. Classification of rotating object detection methods. The methods can be divided into the methods based on rotation invariance and the methods based on loss function.

a new rotatable bounding box. The network first determines whether the object is a detection object and then generates multiple anchor boxes of different angles and sizes. This method presets many different rotation boxes, so detection is slow. Ding et al. [28] converted horizontal RoI (HRoI) into rotational RoI (RRoI), enabling the model to learn RRoI from horizontal feature mapping HRoI. This method transforms the HBB into a directional bounding box, and improves the accuracy of the model effectively. Although there is a directional box around the object in the abovementioned directional anchor method, there is still a problem of feature misalignment in detection. To solve this problem, Zhang et al. [29] proposed the RRoI pooling operation. It enhances the orientation angle information and boundary information of the anchor point, and makes the RoI feature extraction more precise. Han et al. [30] designed an S2A-Net composed of feature alignment module and directional detection module, which effectively alleviated the problem of inconsistency between classification score and positioning accuracy.

For another kind of rotation detection algorithm based on loss function. Traditional boundary box regression loss functions include Huber loss [31] and log-cosh loss [32]. In addition, Yang et al. [33] proposed circular smooth label (CSL), which transforms angle prediction from a regression task to a classification task and solves the periodicity problem. Although these methods add an additional direction parameter, they can only reduce the loss of angle regression, and cannot reflect the connection between angle parameters and IoU [34]. Therefore, IoU method based on 2-D Gaussian distribution has attracted the attention of scholars. Yang et al. [35] designed a new IoU loss based on 2-D Gaussian distribution and Gaussian Wasserstein distance to solve the problem of inconsistent loss calculation and measurement metric. Later, Kullback–Leibler divergence (KLD) [36] and ProbIoU [13] also used different mathematical distance formulas to calculate the distance between 2-D Gaussian distributions.

These methods have effectively improved detection accuracy, but the 2-D Gaussian distribution also makes it difficult for them to accurately predict the angles of low aspect ratio objects.

III. Proposed Methods

This section describes our proposed approach in detail. The proposed method is applied to YOLOv8 network for validation,

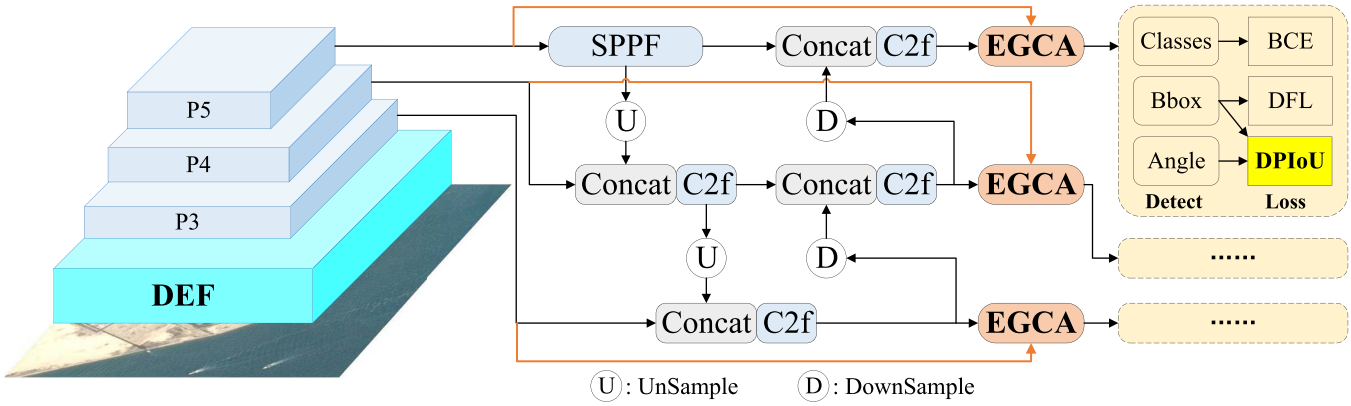

Fig. 3. Overall network framework diagram. We introduce DEF at the forefront of the network to perform the first feature extraction on the image. Then EGCA is introduced to the end of the feature pyramid, allowing the model to complete a contextual feature fusion before final detection. Finally, we introduce DPIoU into the loss function, which will participate in the calculation of Bbox and angle losses.

which is one of the general networks with an excellent performance in remote sensing object detection. It is worth noting that the modules we designed are universal and can also be widely applied in other networks. We will cover the DEF in detail in Section III-B and the ECAF in Section III-C. In Section III-D, we delve into the 2-D Gaussian distribution based IoU method and propose DPIoU.

A. Overall Network Structure

We adopt YOLOv8n as the basic model, and introduce the 2 modules and 1 loss function in this article into it. The overall structure of the obtained network is shown in Fig. 3.

We introduce the DEF module at the forefront of the network, which first performs the first feature extraction on the image. It will extract spatial features and additional edge features of the image, allowing the model to better perceive the gradient changes of pixels. Second, we introduce the EGCA module to the end of the feature pyramid network (FPN). It enables the model to complete a fusion of high-dimensional and low dimensional features before final detection, preventing deep networks from ignoring low dimensional edge features. Finally, we introduce DPIoU into the loss function for IoU calculation. It improves the accuracy of the model's angle prediction by considering more comprehensive constraints.

B. Discrete Edge Focused Feature Extraction Module

In optical remote sensing images, ship wakes often have blurred contours or incomplete textures due to insufficient energy or interference by sea clutter. Meanwhile, they do not have a continuous texture like objects in natural scenes, which makes it more difficult for neural networks to learn their features [26]. Therefore, this section aims to study how to extract the features of the wake more effectively and how to make the model perceive the wake more keenly.

We show in Fig. 4 the wakes in optical remote sensing images, whose contours and textures are almost all composed of multiple discrete lines. This is completely different from objects in natural scenes, where objects are continuous, gradual, and point-blocky.

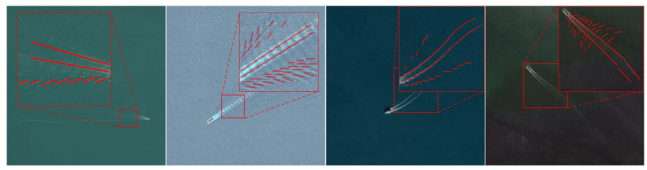

Fig. 4. Wake in optical remote sensing image. We traced the outline and texture of the wake with red lines, and it can be observed that most of them are composed of discrete lines.

However, the wake is discrete, abrupt, and linear, therefore it has strong gradient changes.

In the perception of image gradient change, edge detection operator can achieve a good effect. Image edges refer to the discontinuities in local image characteristics, where abrupt changes in pixel or structural information typically occur. Scharr operator is a classical discrete differential operator for edge detection, which combines Gaussian smoothing and differential differentiation. The operator consists of two sets of 3×3 matrices, which can be convolved with the image, respectively, to obtain the approximate brightness difference values of the horizontal and vertical. Their formulas can be expressed as

$$G_x = \begin{bmatrix} -3 & 0 & +3 \\ -10 & 0 & +10 \\ -3 & 0 & +3 \end{bmatrix} * A,\ G_y = \begin{bmatrix} +3 & +10 & +3 \\ 0 & 0 & 0 \\ -3 & -10 & -3 \end{bmatrix} * A \tag{1}$$

where A represents the original image, G_x and G_y represent the results of horizontal and vertical edge detection, respectively. Then, we add G_x and G_y to get the final edge feature graph G

$$G = G_x + G_y. \tag{2}$$

We test Scharr operator in several scenarios to verify the possibility of using it to extract edge features. The wake in Fig. 5(a) and (b) has almost no texture information, only vague and incomplete "V-shaped" lines. After they are processed by Scharr operator, more obvious contours can be obtained, which will reduce the difficulty of wake detection by the model. The wake in Fig. 5(c)

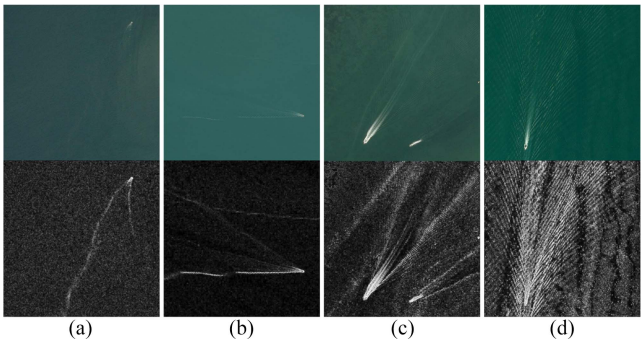

Fig. 5. Comparison between the original image of optical remote sensing and the image processed by Scharr operator. The first line is the original image, and the second line is the processed image, in which the discrete lines that make up the wake can be more clearly observed. (a) and (b) has almost no texture information, only vague and incomplete "V-shaped" lines. (c) and (d) has a more pronounced "V-shaped" line, but its texture is difficult to observe.

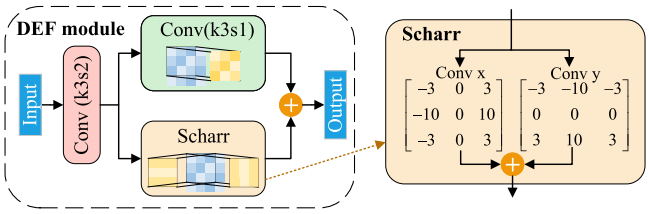

Fig. 6. Structure diagram of the DEF module. It mainly extracts features by edge branches and convolution branches in parallel, so as to obtain rich feature representations that fuse edge information and spatial information.

and (d) has a more pronounced “V-shaped” line, but its texture is difficult to observe. After they are processed by Scharr, the texture information composed of multiple discrete lines can be clearly seen. This is because discrete line textures have strong gradient variations similar to edge features.

Therefore, we design DEF with the edge operator as one of the components, aiming to improve the sensitivity of the model to the discrete edge texture specific to the wake. As shown in Fig. 6, we use the edge operator as one of the feature extraction branches to combine with the convolution branch. They will be used as the front-end components of the model to sense the gradient changes of pixel and the spatial structure of image in parallel, so as to learn more abundant image features.

As the front-end component of the model, DEF module extracts the features initially through convolution with step size 2, and then further refines the captured information through edge branches and convolution branches. The convolution branch can effectively extract the spatial features of the image, but it is not sensitive to the contour shape of the image [37]. The edge branches in the module can effectively make up for this deficiency. It can sense gradient changes in pixels to effectively capture contour and texture features made up of discrete lines in an image. Finally, the module will aggregate the refined spatial information and edge information to get a richer representation of image features, so as to help the model understand the image content more comprehensively.

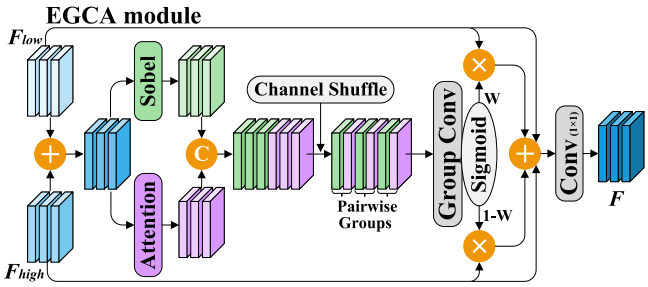

Fig. 7. Structure diagram of the EGCA module. The module gets the initial weight from the attention, and then lets the edge features guide it to grow into the final weight. Finally, the weights are used to recombine the low and high dimensional features.

The abovementioned design makes DEF module more sensitive to abrupt changes in image information, can effectively extract edge features and reduce the model’s dependence on texture information, thus improving the model’s performance in remote sensing scenarios such as ship wake detection.

C. Edge Guided Context Attention Fusion Module

For object detection models, images gradually change from low-dimensional features to high-dimensional features after multiple feature extraction in backbone [38]. After these features are fused through the feature pyramid, they will become more abstract semantic information [39]. At this time, the high-dimensional features and low-dimensional features have great differences in morphology and characteristics. Low-dimensional features provide rich edge, texture, and other details, while high-dimensional features provide abstract feature representation.

In the introduction to Section III-A, we explain in detail the advantages of focusing on more wake edge features in the image. However, edge operators are difficult to apply to high-dimensional feature maps, because high-dimensional features are abstract and irregular, and their gradient mutations do not represent the edge features in the image. Therefore, we hope to fuse low-dimensional feature maps containing a lot of edge information with high-dimensional feature maps, so that the model can have better generalization ability when facing different types and scales.

Therefore, we design a three-step context feature fusion module based on the idea of paying attention to edge features, which guides the whole fusion process through edge feature information. Specifically, we first use the attention mechanism to generate initial feature importance weights $W_{\text{ini}} \in \mathbb{R}^{C\times H\times W}$ (C is the number of channels, H is the length, and W is the width) for each channel separately, and then use edge features to guide the initial weights to grow into final weights $W \in \mathbb{R}^{C\times H\times W}$. Finally, the weights are used to reorganize the low-dimensional and high-dimensional feature maps. The structure of the entire EGCA module is shown in Fig. 7.

The attention branch in the module is a flexible component that can use any attention with the same input and output dimensions depending on the image scene. In this study, we

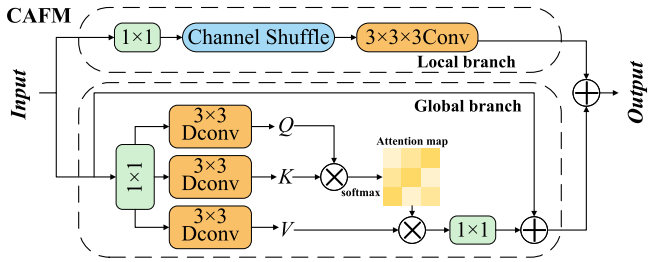

Fig. 8. Structure diagram of CAFM module. It mainly uses local branches and global branches to jointly calculate the attention weight.

used convolution and attention fusion module (CAFM) [40] as a component. CAFM was proposed in the related research of remote sensing images, which can make up for the shortcomings of CNN in perceiving global features. It consists of local and global branches, as shown in Fig. 8.

In the local branch of CAFM, the channel shuffle operation is first used to mix various channels, and then 3-D convolution is used to extract features, so as to obtain local features and achieve denoising. In the global branch, three $3\times$ 3 deep convolution are first used to generate query (Q), key (K), and value (V), respectively. Then, calculate the mutual attention map of Q and K, and finally multiply it with V to get the output.

While the EGCA module generates the initial feature importance weights $W_{\text{ini}} \in \mathbb{R}^{C\times H\times W}$ through attention, an edge branch extracts the edge feature G in parallel. In the edge branch, we use Sobel operator to extract features. Sobel is a very similar edge detection operator to Scharr, it is more conservative in the parameters of the convolution matrix than Scharr, so it has lower sensitivity. This can effectively reduce the noise caused by the high-dimensional feature graph, and we will prove this in the ablation experiment. The vertical and horizontal detection formulas of the Sobel operator can be expressed as

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * A,\ G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A \quad (3)$$

where A represents the original image, G_x and G_y represent the results of horizontal and vertical edge detection respectively.

In order to get the final feature importance weight $W \in \mathbb{R}^{C\times H\times W}$, we take the feature G extracted from the edge branch as the guide, and then adjust the initial weight $W_{\text{ini}} \in \mathbb{R}^{C\times H\times W}$ of each corresponding channel. Specifically, we first concatenate (Concat) W_{ini} with G, and then group them in pairs along the channel by channel shuffle. This allows each group to have a W_{ini} and its corresponding channel G. After that, we use Group Conv [41] to convolve each group separately, with G as the guide to adjust the W_{ini}. Finally, after the adjusted result is calculated by activation function Sigmoid, the final usable weight W is obtained. This process can be expressed by

$$W = \text{Sigmoid}(\text{GC}_{n=C}(\text{CS}[G, W_{\text{ini}}])) \quad (4)$$

where G is the feature map extracted by the edge branch and W_{ini} is the initial weight obtained by the attention mechanism. CS indicates the channel shuffle operation. GC refers to Group Conv, and its group number n is the number of channels C of $W_{\text{ini}} \in \mathbb{R}^{C\times H\times W}$.

In the third step of the module, we recombine the low-dimensional feature map F_{low} and the high-dimensional feature map F_{high} adaptively by weight W. In this process, we also use skip connections similar to the residual structure. The problem of gradient disappearance and network degradation in the neural network is alleviated by inputting the feature map to be fused again. Finally, after a $1\times$ 1 convolution completes the aggregation and adjustment, the final fusion result F is obtained. The whole process can be expressed by

$$F = \text{CBS}_{k=1}[F_{\text{low}} \cdot (W+1) + F_{\text{high}} \cdot (2-W)] \quad (5)$$

where F is the final result of fusion, F_{low} is the input low-dimensional feature map, and F_{high} is the input high-dimensional feature map. W is the feature importance weight obtained from formula (4). $\text{CBS}_{k=1}$ is a convolution of size $1\times$ 1.

EGCA module reintegrates rich low-dimensional features and forces the model to pay attention to important edge and gradient information in the image, which improves the model's global awareness and anti-interference ability.

D. Diagonal Constrained Probabilistic IoU

Rotating object detection is an important topic in computer vision and has a wide range of applications, especially in remote sensing scenes. Compared with the horizontal box, the rotating box can more accurately describe the pose and position of the object, so as to improve the detection accuracy [35]. Due to the distinct directional features of ship wakes, this article uses rotation detection methods to identify ships and obtain richer ship heading information.

The rotating box has one more angle parameter than the horizontal box, but this angle parameter has the problem of boundary discontinuity. This is because angles are periodic, and regression-based methods can easily lead to predictions outside the defined range. CSL [33] sidesteps this problem by converting angle regression into angle classification. But there is a new problem where the loss (using class loss) and metric (using IoU) are not aligned. With the development of rotating object detection technology, the IoU calculation method based on 2-D Gaussian distribution has been proposed. In this class of methods, any rotated rectangle is approximated by a 2-D Gaussian distribution (like an inner ellipse) and then compute the mathematical distance between the distributions. They sidestep the problems of boundary discontinuity and index inconsistency well.

A rotating box can be expressed as (x_0, y_0, w, h, θ), which are the center point coordinates x_0 and y_0 of the box, the length w and height h, and the rotation angle θ, respectively. To determine a 2-D Gaussian distribution, we need to compute the mean μ and the covariance matrix Σ. Where μ is $(x_0, y_0)^T$, Σ has three free quantities a, b, and c, and their transformation with the parameters of the rotating box can be calculated by

$$\sum = \begin{bmatrix} a & c \\ c & b \end{bmatrix} = R_\theta \begin{bmatrix} w & 0 \\ 0 & h \end{bmatrix} R_\theta^T$$

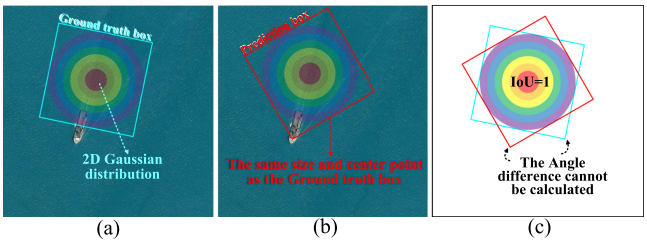

Fig. 9. Deficiency of 2-D Gaussian distribution on square-like anchor boxes. (a) Ground truth box. (b) Prediction box. (c) Superimposed image of the true and predicted boxes. When two anchor boxes have the same size and center point, the IoU calculated by 2-D Gaussian distribution is always 1, no matter what their angle difference is.

$$= \begin{bmatrix} w\cos^2\theta + h\sin^2\theta & \frac{1}{2}(w-h)\sin 2\theta \\ \frac{1}{2}(w-h)\sin 2\theta & w\sin^2\theta + h\cos^2\theta \end{bmatrix} \tag{6}$$

where R_θ is a 2-D rotation matrix. In other words, a rotating box p with argument (x_0, y_0, w, h, θ) can be expressed as $p \sim \mathrm{N}(\mu, \Sigma)$ with

$$\mu = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \Sigma = \begin{bmatrix} w\cos^2\theta + h\sin^2\theta & \frac{1}{2}(w-h)\sin 2\theta \\ \frac{1}{2}(w-h)\sin 2\theta & w\sin^2\theta + h\cos^2\theta \end{bmatrix}. \tag{7}$$

Next we take the ProbIoU [13] as an example. It considers the characteristics of 2-D Gaussian distribution, satisfies all distance metrics, can represent the true distance between different distributions, and is differentiable everywhere. ProbIoU first computes the Hellinger Distance between two 2-D Gaussian distributions (p and q)

$$\mathrm{HD}(p, q) = \frac{1}{2}\int \left(\sqrt{p(x)} - \sqrt{q(x)}\right)^2 dx \tag{8}$$

where $\mathrm{HD}(p, q) \in [0, 1]$. And $\mathrm{HD}(p, q) = 0$ if and only if the two distributions are the same. Then, we can get the specific calculation formula of ProbIoU

$$\mathrm{ProbIoU} = 1 - \mathrm{HD}(p, q). \tag{9}$$

However, this kind of IoU method based on 2-D Gaussian distribution does not perform well when facing square-like. As shown in Fig. 9, if there are two squares with the same size and center, they will always have the same 2-D Gaussian distribution regardless of the angle difference between them.

We mathematically perform the analysis of the 2-D Gaussian distribution $p \sim \mathrm{N}(\mu, \Sigma)$. If we set $t = \frac{w}{h} - 1$ (the difference between the anchor's aspect ratio and 1), the covariance matrix Σ becomes

$$\begin{aligned} \Sigma &= \begin{bmatrix} w\cos^2\theta + h\sin^2\theta & \frac{1}{2}(w-h)\sin 2\theta \\ \frac{1}{2}(w-h)\sin 2\theta & w\sin^2\theta + h\cos^2\theta \end{bmatrix} \\ &= h\begin{bmatrix} 1 - t\sin^2\theta & \frac{t}{2}\sin 2\theta \\ \frac{t}{2}\sin 2\theta & 1 - t\cos^2\theta \end{bmatrix}. \end{aligned} \tag{10}$$

From formula (10), it can be found that when t approaches 0 (when the anchor box aspect ratio approaches 1), the covariance matrix Σ is calculated almost without θ participation. In this case, the mathematical distance between 2-D Gaussian distributions can only be calculated from the mean μ (the position of the center point) and the height h of the anchor box.

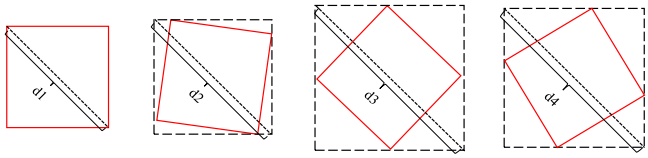

Fig. 10. Change in the diagonal distance of the minimum bounding horizontal rectangle when the rectangular frame rotates. Because the diagonal geometric distance of the bounding rectangle undergoes continuous and bounded changes, it can be used as a constraint condition for IoU calculation.

Therefore, we want to introduce an additional constraint to make the IoU have better performance when facing the class square. We observe that when a rectangle rotates, its minimum bounding horizontal rectangle changes. At this point, the diagonal distance of the minimum external horizontal rectangle also undergoes a transformation. And the change is continuous and bounded, there will be no boundary discontinuity caused by angle periodicity. The change of the external rectangle caused by the rotation of the anchor box is shown in Fig. 10.

Based on the ProbIoU method, we use the minimum external horizontal rectangle of the anchor box to describe the rotation of the anchor box. The diagonal distance of the external rectangle is then used as a constraint to calculate the IoU. This diagonal distance d can be obtained by the following formula:

$$d = \sqrt{(w\cos\theta)^2 + (w\sin\theta + h\cos\theta)^2}, (w \geq h). \tag{11}$$

We refer to the d values of ground truth box and predicted box as d_t and d_p, respectively, and then divide their difference by d_t to obtain the relative distance difference. Multiplying this relative distance by a dynamic coefficient k can serve as the diagonal distance constraint. At this point, the formula for DPIoU can be obtained

$$\mathrm{DPIoU} = \mathrm{ProbIoU} - k\frac{|d_t - d_p|}{d_t}. \tag{12}$$

Next, we need to discuss the dynamic coefficient k. We want the main object of diagonal distance constraints to be square-like objects, since other rectangles are perfectly capable of distinguishing rotation angles by 2-D Gaussian distributions. Based on this, we design a shape-based dynamic weighting function to adaptively determine the coefficient k. The function is as follows:

$$k = \alpha \cdot \exp\left[\frac{-(\frac{w}{h} - 1)^2}{2\sigma^2}\right], (w \geq h). \tag{13}$$

This is a function based on Gaussian distribution, where σ is the standard deviation (which determines the width of the crest) and α determines the peak of the function. It reaches its maximum value when the aspect ratio is 1, and gradually decreases on both sides of the peak. In the ablation experiments in this article, the best properties can be obtained when α and σ are 0.08 and 0.25, respectively. This function forces diagonal constraints to focus on anchor box with low aspect ratios.

We can express the ground truth box and predicted box as $t(x)$ and $p(x)$, respectively, with 2-D Gaussian distribution, and

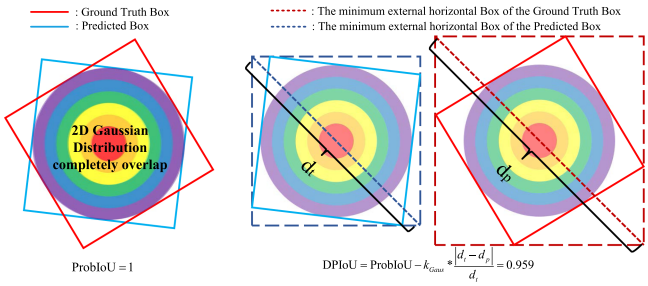

Fig. 11. ProbIoU compared to DPIoU when faced with two anchor boxes with the same size and center point. At this point, the result of ProbIoU's calculation is 1, which means that it considers the two boxes are completely coincident. DPIoU correctly detects that the two boxes are different.

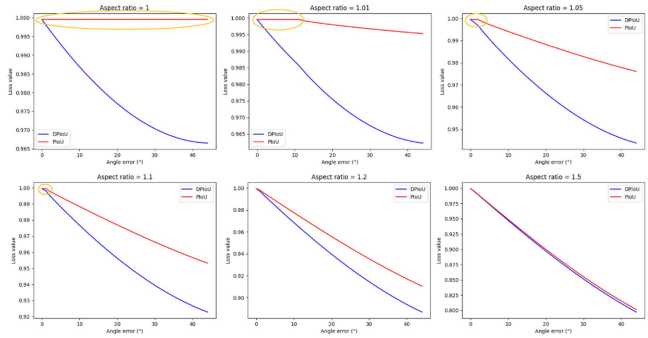

Fig. 12. Relationship curve between angle error and IoU under different aspect ratios. The red curve represents PIoU, the blue curve represents DPIoU, and the yellow circle indicates a clear unreasonable situation where IoU=1.

then combine the formula (11), (12), and (13) to get the final DPIoU formula

$$\mathrm{DPIoU} = 1 - \frac{1}{2}\int \left(\sqrt{t(x)} - \sqrt{p(x)}\right)^2 dx - k_t \frac{|d_t - d_p|}{d_t}. \tag{14}$$

For comparison, we apply ProbIoU and DPIoU to a set of anchor boxes with the same size and center point, and the calculated IoU is shown in Fig. 11. ProbIoU in the figure believes that the two anchor boxes with different angles overlap completely (IoU=1), while DPIoU effectively avoids this incorrect situation.

In order to more clearly demonstrate the difference between PIoU and DPIoU, we designed an experiment. First, construct a set of identical ground truth boxes and prediction boxes, then rotate the predicted boxes along the center point and calculate the IoU at different rotation angles. Then, change the aspect ratio of the anchor frame and repeat the abovementioned experiment. Finally, we obtained the relationship between angle error and IoU at different aspect ratios, as shown in Fig. 12.

From Fig. 12, it can be observed that when the aspect ratio of the anchor frame is relatively small, PIoU is insensitive to angle errors. Especially when the aspect ratio is 1, regardless of the angle error, PIoU always remains at 1. When DPIoU faces anchor boxes with low aspect ratios, the constraint on diagonal distance increases, thereby alleviating the problem of insensitivity to angle errors caused by the same 2-D Gaussian distribution. When DPIoU faces anchor boxes with high aspect ratios, the dynamic weighting function reduces the constraint on diagonal distance, thereby achieving an excellent performance similar to PIoU.

Therefore, DPIoU improves the rationality of calculating the angle error by introducing additional constraints. Especially for the target with low aspect ratio, the reliability of the model to judge whether the angle prediction is correct is greatly improved.

TABLE I
MAIN EXPERIMENTAL HARDWARE AND SOFTWARE PARAMETERS, AS WELL AS THE SETTINGS DURING MODEL TRAINING

Environment	Configuration	Hyp options	Setting
Operating System	Ubuntu 18.04.1	Input resolution	640×640
GPU	Nvidia Tesla V100	lr0	0.01
GPU Driver	CUDA 12.0	lrf	0.01
DL Framework	Pytorch 1.18	Batch_size	16
DL Language	python 3.11	Epochs	300

IV. EXPERIMENTAL RESULTS

A. Experimental Environment

All experiments in this study were performed on a 64-b Ubuntu operating system computer. It is equipped with an Nvidia Tesla V100 GPU, and the related environment and training configuration are shown in Table I.

We chose the SWIM public dataset as the benchmark dataset for training and testing in this study. The SWIM dataset is an optical remote sensing dataset constructed for ship wake detection. It contains images of coastal areas around Asia, Europe, Africa, North America, and Oceania. The annotated objects include the wake of ships, from small yachts to large container ships, with backgrounds including open seas, ports, straits, and canals. The spatial resolution of the image ranges from 2.5 to 0.5 m, with a size of 768 × 768 pixels. SWIM contains 11 600 images and 15 356 wake instances. We divided it into training set: validation set: test set in a ratio of 7:2:1.

B. Evaluation Metrics

Mean average precision (mAP) is a commonly used comprehensive performance evaluation metric in the field of computer vision and object detection. It is calculated from the precision and recall metrics. We denote true positive, false positive, and false negative as TP, FP, and FN, respectively. Then, the formulas for precision and recall are as follows:

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{15}$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}. \tag{16}$$

In this case, the calculation formula of average precision (AP) can be expressed as

$$\mathrm{AP} = \int_0^1 P(R)dR \tag{17}$$

where P refers to precision and R refers to recall. The mAP with an IoU threshold of 0.5 is denoted as mAP_{50}. The mAP for IoU

TABLE II
EXPERIMENTAL RESULTS OF DIFFERENT NETWORK METHODS ON SWIM DATASET (ROTATING BOX)

Methods	Backbone	mAP_{50}	$mAP_{50:95}$	FPS
R-RetinaNet [42]	ResNet50-FPN	73.6	53.2	26
R-Faster R-CNN [17]	ResNet101-FPN	79.2	56.8	27
ROI Transformer [28]	Swin Tiny-FPN	86.5	64.1	21
S2ANet [30]	ResNet152-FPN	88.2	63.5	23
YOLOv5n-CSL [43]	CSPDarkNet-C3	85.7	62.8	98
YOLOv7-KLD [44]	CSPDarkNet-ELAN	90.5	66.5	112
YOLOv8n-obb [45]	CSPDarkNet-C2f	92.8	68.3	**142**
Ours	CSPDarkNet-C2f	**95.1**	**70.4**	135

mAP is measured in % and FPS is measured in frames/second.
The best score for each indicator has been highlighted in bold.

thresholds ranging from 0.5 to 0.95 (in steps of 0.05) is denoted as $mAP_{50:95}$. mAP is defined as follows:

$$\text{mAP} = \frac{\sum_{i=1}^{N} \text{AP}_i}{N} \tag{18}$$

where N refers to the number of categories and AP_i represents the AP value of the i th class. In addition, we use the frames per second (FPS) metric as a measure of the real-time performance of the model. It represents the number of frames processed by the model per second, with higher FPS indicating better real-time performance of the model.

C. Results and Analysis

To verify the effectiveness of the algorithm proposed in this study, we trained and tested our method, R-RetinaNet [42], R-Faster R-CNN [17], ROI Transformer [28], S2ANet [30], YOLOv5n-CSL [43], YOLOv7-KLD [44], and YOLOv8-obb [45] on the SWIM dataset, respectively. Among them, Li et al. introduced CSL [33] into the horizontal box detector YOLOv5n [46] to obtain YOLOv5n-CSL with rotation detection ability. Liu et al. introduced KLD [36] into the horizontal box detector YOLOv7 [47] to obtain YOLOv7-KLD with rotation detection ability. YOLOv8n-obb is a rotation detection version obtained by the authors of YOLOv8 through the introduction of ProbIoU [13]. The experimental results of these models are shown in Table II.

As can be seen from the table, our model achieves the best scores on both mAP_{50} and $mAP_{50:95}$ metrics. Its mAP_{50} reaches 95.1% and $mAP_{50:95}$ reaches 70.4%, which indicates that it has the best comprehensive detection capability and can detect the wake accurately and stably. In terms of the real-time performance, it achieves 135FPS, which is only slightly lower than YOLOv8n-obb, but it is enough to meet the real-time requirements of remote sensing image detection.

In order to verify the effect of the improved angle prediction and edge features of this research method, we show and compare the detection results of the model proposed in this study, YOLOv8-obb, YOLOv7-KLD, and YOLOv5n-CSL in different scenarios.

Fig. 13 shows the results of different methods in simple scenarios, which facilitates us to compare the angle prediction accuracy of different methods. From Fig. 13(a) and (b), it can be found that the angle can be accurately predicted by the method of this study when the object anchor box is close to a square.

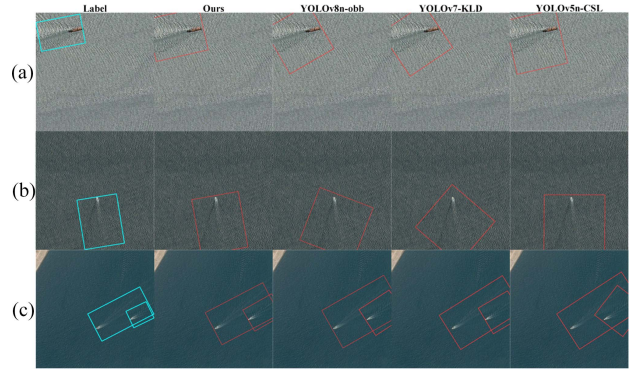

Fig. 13. Detection results of different network methods in simple scenarios. They can easily compare the accuracy of angle prediction. (a), (b), and (c) are the model detection results in different scenarios.

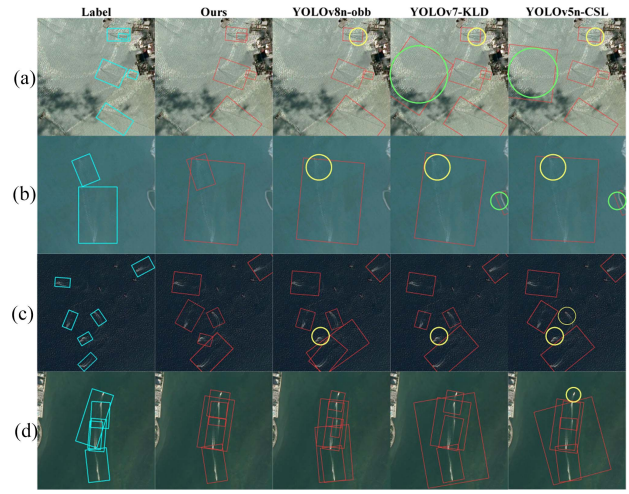

Fig. 14. Detection results of different network methods in complex scenes.Yellow circles represent missed detections and green circles represent false detections. (a)–(d) are the model detection results in different scenarios.

However, YOLOv8-obb and YOLOv7-KLD produced a large angle prediction error, and YOLOv5-CSL also produced a small angle prediction error. This is because both YOLOv8-obb and YOLOv7-KLD use IoU loss based on 2-D Gaussian distribution, which causes them to fail to accurately judge whether the angle prediction is good or bad when facing square-like. The method in this study avoids this problem by introducing diagonal constraints. YOLOv5-CSL classifies angles and replaces regression prediction with category prediction. So it has no obvious defect when facing square-like. However, its comprehensive performance is not as good as the method using 2-D Gaussian distribution [35], which is the reason why only YOLOv5-CSL produces a large angle prediction error in Fig. 13(c). The object in Fig. 13(c) has a high aspect ratio, where the method using 2-D Gaussian distribution can fully exert its performance and achieve accurate angle prediction.

We show the detection results in more complex scenes in Fig. 14, which is used to compare the comprehensive performance of different methods. It can be seen from the figure that the proposed method can accurately identify the ship wakes

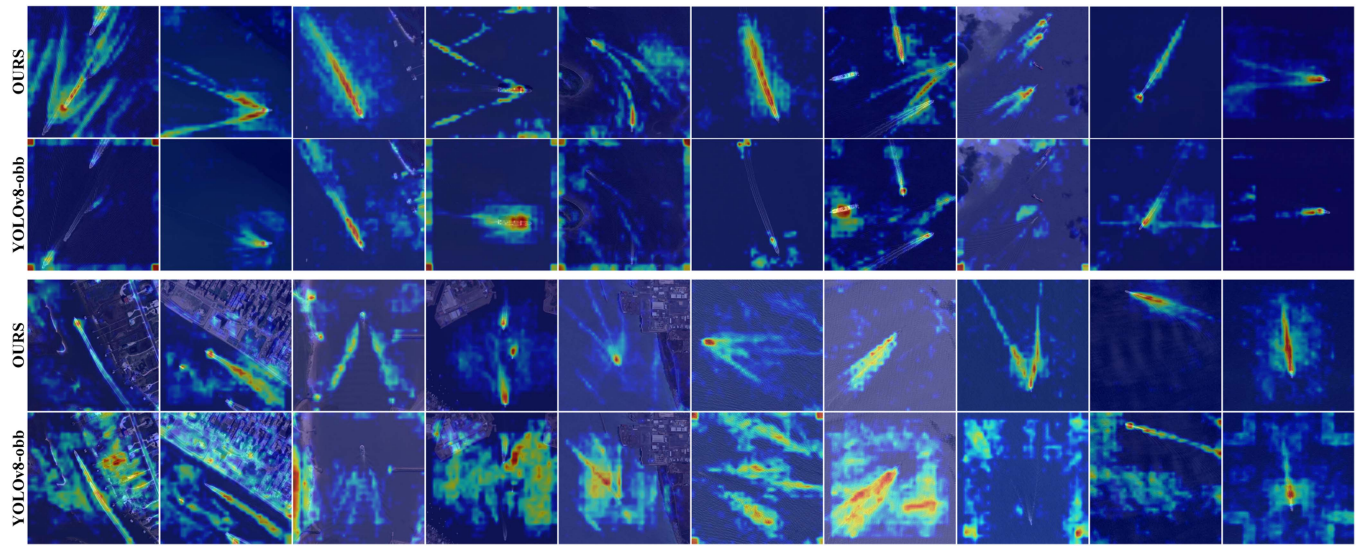

Fig. 15. Heatmaps of our approach and the original network on the SWIM dataset. In the images, regions that are more red indicate areas where the model pays greater attention.

in each scenario, while YOLOv8n-obb is prone to miss detection when facing overlapping wakes. For YOLOv7-KLD and YOLOv5n-CSL, they both misjudge the water surface clutter and island edges as wakes, and produce more missed detections than YOLOv8n-obb. In Fig. 14(d), where the wake overlap is very severe, our method obtains the result closest to the labeled box. Other methods, after detecting multiple small wakes, will combine them again as one large wake. This may be caused by the network not sufficiently learning the discrete line-like texture features.

In addition, we validate the effectiveness of our method by comparing heatmaps. As seen in Fig. 15, our model focuses primarily on the wake, accurately identifying its contours. In contrast, the baseline model pays more attention to the ship itself, demonstrating weaker wake detection capabilities and being more susceptible to interference from complex backgrounds.

Therefore, the experimental results show that our method can achieve competitive performance in a variety of complex scenes, and significantly improve the angle prediction accuracy when facing square-like objects.

D. Ablation Experiment

We successively introduce the proposed method into YOLOv8n-obb to verify the effectiveness of each module. In addition to using mAP as the evaluation metric of the comprehensive performance, we also evaluate the improvement of angle prediction accuracy of the model through mean angle prediction error (MAPE). MAPE is the average of the error between the predicted angle and the actual angle for all correctly predicted objects, and it reflects the overall accuracy of the angle prediction. The results of the ablation experiments are shown in Table III.

As can be seen from Table III, both the modules and loss functions proposed in this study can effectively improve the accuracy of the model, and they will achieve a better performance when combined. In addition, DPIoU can significantly improve the accuracy of angle prediction. The method in this study reduces the average angle prediction error from 15.35° to 7.16°, which can help us predict the ship's heading more accurately.

TABLE III
ABLATION EXPERIMENTAL RESULTS OF INTRODUCING THE THREE METHODS PROPOSED IN THIS STUDY INTO YOLOV8N-OBB, RESPECTIVELY

DEF	EGCA	DPIoU	mAP_{50}(%)	$mAP_{50:95}$(%)	MAPE(°)
			92.8	68.3	15.35
√			93.8	68.8	14.58
	√		94.1	69.1	13.32
		√	94.3	69.2	9.55
√	√		94.5	70.2	12.43
√		√	94.5	69.4	8.23
	√	√	94.7	69.4	7.78
√	√	√	**95.1**	**70.4**	**7.16**

The best score for each indicator has been highlighted in bold.

TABLE IV
EXPERIMENTAL RESULTS OF USING SCHARR AND SOBEL OPERATORS IN DEF AND EGCA MODULES RESPECTIVELY

mAP(%)	DEF&Scharr		DEF&Sobel	
	mAP_{50}	$mAP_{50:95}$	mAP_{50}	$mAP_{50:95}$
EGCA&Scharr	93.2	69.0	93.1	68.8
EGCA&Sobel	**94.5**	**70.2**	94.2	69.9

The best score for each indicator has been highlighted in bold.

We use the Scharr operator when designing the DEF module, and the Sobel operator when designing the EGCA module. The main difference between the two operators is that Scharr has a larger convolution kernel coefficient and is, therefore, more sensitive. However, this high sensitivity tends to produce more noise in high-dimensional features. We design experiments to verify which operator should be used in these two modules. The relevant experimental results are shown in Table IV.

It can be seen from the experimental results that the best performance can be achieved when DEF uses Scharr and EGCA uses Sobel. At the same time, for EGCA module, the

TABLE V
RESULTS OF HYPERPARAMETER ABLATION EXPERIMENTS FOR DPIOU

$mAP_{50}(\%)$	σ=0.1	σ=0.2	σ=0.25	σ=0.3	σ=0.4
α=0.07	94.0	94.1	94.2	94.1	93.7
α=0.08	94.0	94.2	**94.3**	94.1	93.8
α=0.09	93.9	94.0	94.1	94.0	93.7

We only introduce DPIoU into YOLOv8n-obb, and then experiment with different values of the hyperparameters α and σ in DPIoU.
The best score for each indicator has been highlighted in bold.

TABLE VI
EXPERIMENTAL RESULTS OF DIFFERENT NETWORK METHODS ON OPENSARWAKE DATASET (ROTATED BOX)

Methods	$mAP_{50}(\%)$	$mAP_{50:95}(\%)$
ROI transformer [28]	36.7	13.6
S2ANet [30]	34.1	12.3
R-CNN*o* [48]	42.8	15.9
YOLOv5n-CSL [43]	52.8	20.4
YOLOv7-KLD [44]	56.4	21.8
YOLOv8n-obb [45]	62.8	26.6
Ours	**66.2**	**29.4**

The best score for each indicator has been highlighted in bold.

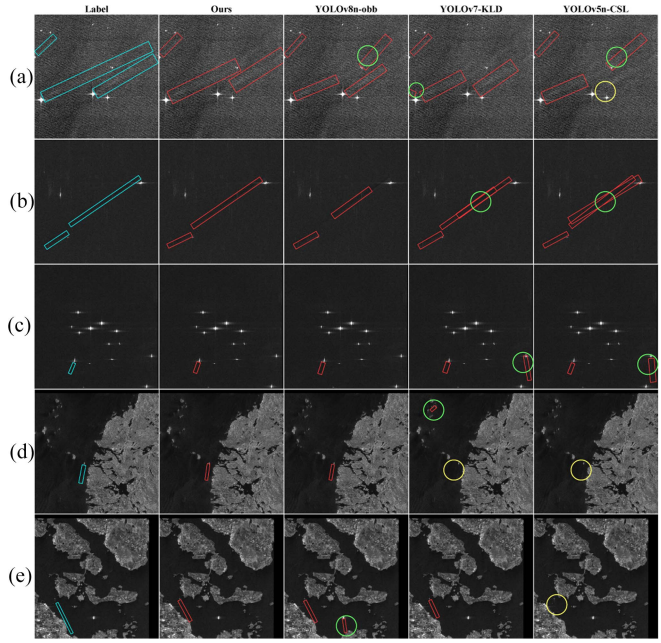

Fig. 16. Detection results of different network methods on OpenSARWake dataset. Yellow circles represent missed detections and green circles represent false detections. (a)–(e) are the model detection results in different scenarios.

performance of Scharr operator is significantly lower than that of Sobel operator.

In this study, a shape-based dynamic weighting function [formula (13)] is introduced in the design of DPIoU. The function uses α (which determines the peak) and σ (which determines the crest width) as hyperparameters to adjust the range of values. We conducted ablation experiments to verify which values of α and σ can achieve the best results, and the relevant experimental results are shown in Table V.

According to the experimental results in the table, the optimal performance can be achieved in our method when the hyperparameters α and σ are set to 0.08 and 0.25, respectively. In addition, compared to the α that determines the peak, σ that determines the peak width has a greater impact on performance. This is consistent with our design that ensures the constraint is focused on the low aspect ratio object, further verifying the rationality of the DPIoU designed in this study.

E. Generalization Experiments

In order to verify the generalization, we conducted experiments on the OpenSARWake [6] public dataset. OpenSARWake is a dataset for ship wake detection based on SAR images. This dataset contains SAR images in L, C, and X bands with a spatial resolution of 1.25–12.5 m. It has 3973 images and 4096 instances, and the size of the images is 1024× 1024 pixels.

As can be seen in Table VI, our method achieves the best comprehensive performance on the OpenSARWake dataset compared to other methods. Its mAP_{50} achieves 66.2%, while $mAP_{50:95}$ achieves 29.4% . In addition, we present the detection results of the abovementioned methods in different scenarios, as shown in Fig. 16.

In Fig. 16(a) and (b), there is a long wake that is not obvious, and there are a few places where the texture is blurred in the middle of this long wake. Among all the methods, the predicted box of our method is closest to the Ground Truth Box. When there is an unobvious texture in a long wake, YOLOv8n-obb is easy to identify the long wake into two short wakes, or only part of them can be detected. Duplicate detection occurs for YOLOv7-KLD and YOLOv5-CSL, which consider the long wake to be composed of multiple wakes stacked together.

In Fig. 16(c)–(e), there are many nonship interference objects (such as cage, reef, island, etc.). At this time, our method can still accurately identify the ship wake without missing or false detection. All the other methods have missed or false detections. This shows that the algorithm in this study has excellent robustness, and it has learned the characteristics of the wake more clearly, so that it is not disturbed by other objects.

The abovementioned experiments prove that for ship wake detection, this method not only has advantages in optical remote sensing images, but also has competitive performance in SAR images. Finally, we will further test the model in a generic scenario. We conducted experiments on the DIOR-R public dataset [49] using our research method and other advanced methods. DIOR-R is a large-scale optical remote sensing object detection dataset annotated with rotation boxes. It covers 20 object classes, consisting of 23 463 remote sensing images and 190 288 target instances.

From Table VII, it can be seen that our method achieved the best overall performance. This means that in addition to accurately detecting ship wakes, it can also achieve excellent performance in general remote sensing scenarios.

F. Discussions

We conduct experiments on both public optical remote sensing wake datasets and SAR wake datasets. Experiments show that the proposed method has an excellent comprehensive detection performance and generalization. It has excellent perception ability for weak objects, and can also achieve excellent angle prediction accuracy for difficult square-like objects.

TABLE VII
COMPARISONS WITH ADVANCED OOD METHODS ON THE DIOR-R DATASET

Method	APL	APO	BF	BC	BR	CH	DAM	ETS	ESA	GF	GTF	HA	OP	SH	STA	STO	TC	TS	VE	VM	mAP
S2ANet [30]	65.4	42.0	75.2	83.9	36.0	72.6	28.0	65.1	75.1	75.6	80.5	35.9	52.1	82.3	65.9	66.1	84.6	54.1	48.0	69.7	62.9
RoI-Trans [28]	63.3	37.9	71.8	87.5	40.7	72.6	26.9	78.7	68.1	69	82.7	47.7	55.6	81.2	78.2	70.3	81.6	54.9	43.3	65.5	63.9
AOPG [49]	62.4	37.8	71.6	87.6	40.9	72.5	31.1	65.4	78	73.2	81.9	42.3	54.5	81.2	72.7	71.3	81.5	60.0	52.4	70	64.4
DODet [50]	63.4	43.4	72.1	81.3	43.1	72.6	33.3	78.8	70.8	74.2	75.5	48	59.3	85.4	74.0	71.6	81.5	55.5	51.9	66.4	65.1
PIIDet-101 [51]	77.9	31.8	79.4	90.0	45.0	72.7	30.7	79.2	80.1	76.6	81.3	44.3	57.9	89.0	69.3	73.1	59.5	57	59.9	62.4	67.4
YOLOv8 [45]	89.3	73.7	94.3	90.9	**65.9**	90.5	**56.2**	92.3	82.1	84.9	**91.7**	**73**	72.1	94.6	94.3	82.5	94.5	73.5	69.9	89.8	82.8
OURS	**93.0**	**75.8**	**97.4**	**94.6**	64.7	**94.0**	54.7	**94.3**	**84.7**	**87.3**	90.1	71.6	**74.5**	**97.8**	**97.0**	**86.7**	**97.7**	**78.9**	**73.5**	**93.7**	**85.1**

The indicators used in the table are mAP_{50}(%).
The best score for each indicator has been highlighted in bold.

But there are still some limitations to this approach. The sea clutter also has discrete line characteristics similar to wake, so the performance of the method will be affected when the clutter is large. This may be improved by introducing something like a filter, which can mitigate the noise caused by clutter. In addition, when facing a square rotated box and its box after flipping horizontally or vertically, the method based on 2-D Gaussian distribution and the method in this study will also consider them to be the same. This is because they have the same 2-D Gaussian distribution and the same minimum external horizontal rectangle. This may require more comprehensive constraints to improve this.

V. CONCLUSION

In this study, we construct DEF based on the characteristics of the wake, such as being indistinct, lacking texture, and having discrete line shapes. As the front end module of the network, it senses the gradient intensity changes of pixels in parallel through additional edge branches, which effectively improves the perception ability of the model to faint textures and contours. In addition, we create EGCA. It forces the model to focus on the edge features in the context information, preventing low-dimensional line texture information from being ignored by the deep network. Finally, we design DPIoU. It further enhances the ability to describe the rotation angle and shape, and solves the problem that the IoU method based on 2-D Gaussian distribution cannot accurately predict the angle of square-like objects. In the experiments, our method achieved a mAP_{50} of 95.1% and a $mAP_{50:95}$ of 70.4% on the optical remote sensing image dataset (SWIM). It also reached 135 FPS and reduced the average angle prediction error from 15.35° to 7.16°. In addition, our method demonstrated competitive performance on the SAR remote sensing image dataset (OpenSARWake), with a mAP_{50} of 66.2% and a $mAP_{50:95}$ of 29.4% . This demonstrates that our method offers excellent accuracy and adaptability, effectively detecting wakes in various scenarios.

In future research, we hope to further optimize the accuracy of angle prediction. In practical scenarios, wake can be subdivided into turbulent wake and Kelvin wake, and their shapes may also be diverse, such as curved or S-shaped. These different wake shapes will affect the accuracy of angle prediction, and we will attempt to design more comprehensive IoU calculation methods to alleviate this issue. In addition, we also want to focus on the issue of uncertain wake boundaries. In addition to helping locate ships, wakes also have rich information such as angle, length, and direction, which can be used to estimate the ship's heading and speed. The prerequisite for doing so is to accurately and completely identify the entire wake, but the contour of the wake always gradually disappears in the image, making it difficult to determine the accurate boundary of the wake. We will attempt to introduce additional threshold methods or filtering modules to alleviate this issue.

REFERENCES

[1] A. Jiaqiu et al., "A novel ship wake CFAR detection algorithm based on SCR enhancement and normalized Hough transform," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 681–685, Jul. 2011.

[2] Y. Jiang, Z. Yang, K. Li, and T. Liu, "Pre-processing of simulated synthetic aperture radar image scenes using polarimetric enhancement for improved ship wake detection," *Remote Sens.*, vol. 16, no. 4, 2024, Art. no. 658.

[3] Z. Zang, C. Lin, C. Tang, T. Wang, and J. Lv, "Zero-shot aerial object detection with visual description regularization," in *Proc. AAAI Conf. Artif. Intell.*, 2024, vol. 38, no. 7, pp. 6926–6934.

[4] R. Zhang et al., "A benchmark and frequency compression method for infrared few-shot object detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, 2025, Art. no. 5001711.

[5] R. Zhang, Z. Cao, Y. Huang, S. Yang, L. Xu, and M. Xu, "Visible-infrared person re-identification with real-world label noise," *IEEE Trans. Circuits Syst. Video Technol.*, to be published, doi: 10.1109/TCSVT.2025.3526449.

[6] C. Xu and X. Wang, "OpenSARWake: A large-scale SAR dataset for ship wake recognition with a feature refinement oriented detector," *IEEE Geosci. Remote Sens. Lett.*, vol. 21, 2024, Art. no. 4010105.

[7] Y. Liu, R. Zhang, R. Deng, and J. Zhao, "Ship detection and classification based on cascaded detection of hull and wake from optical satellite remote sensing imagery," *GIScience Remote Sens.*, vol. 60, no. 1, 2023, Art. no. 2196159.

[8] F. Biondi, "Low-rank plus sparse decomposition and localized radon transform for ship-wake detection in synthetic aperture radar images," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 1, pp. 117–121, Jan. 2018.

[9] Y. Liu, J. Zhao, and Y. Qin, "A novel technique for ship wake detection from optical images," *Remote Sens. Environ.*, vol. 258, 2021, Art. no. 112375.

[10] K.-M. Kang and D.-J. Kim, "Ship velocity estimation from ship wakes detected using convolutional neural networks," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 11, pp. 4379–4388, Nov. 2019.

[11] Y. Gu, Z. Hu, Y. Zhao, J. Liao, and W. Zhang, "MFGTN: A multi-modal fast gated transformer for identifying single trawl marine fishing vessel," *Ocean Eng.*, vol. 303, 2024, Art. no. 117711.

[12] Y. Zhao, S. Chen, S. Liu, Z. Hu, and J. Xia, "Hierarchical equalization loss for long-tailed instance segmentation," *IEEE Trans. Multimedia*, vol. 26, pp. 6943–6955, 2024.

[13] J. M. Llerena, L. F. Zeni, L. N. Kristen, and C. Jung, "Gaussian bounding boxes and probabilistic intersection-over-union for object detection," 2021, *arXiv:2106.06072*.

[14] R. Zhang et al., "Detail-aware network for infrared image enhancement," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, 2025, Art. no. 5000314.

[15] R. Zhang, L. Xu, Z. Yu, Y. Shi, C. Mu, and M. Xu, "Deep-IRtarget: An automatic target detector in infrared imagery using dual-domain feature extraction and allocation," *IEEE Trans. Multimedia*, vol. 24, pp. 1735–1749, 2022.

[16] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 2980–2988.

[17] 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.
[18] W. Liu et al., "SSD: Single shot multibox detector," in *Proc. 14th Eur. Conf. Comput. Vis.*, 2016, pp. 21–37.
[19] 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.
[20] X. Fan, Z. Hu, Y. Zhao, J. Chen, T. Wei, and Z. Huang, "A small-ship object detection method for satellite remote sensing data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, pp. 11886–11898, 2024.
[21] X. Wu, D. Hong, and J. Chanussot, "UIU-Net: U-Net in U-Net for infrared small object detection," *IEEE Trans. Image Process.*, vol. 32, pp. 364–376, 2023.
[22] Z. Tan, Z. Jiang, C. Guo, and H. Zhang, "WSODet: A weakly supervised oriented detector for aerial object detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5604012.
[23] J. Li, M. Chen, S. Hou, Y. Wang, Q. Luo, and C. Wang, "An improved S2A-Net algorithm for ship object detection in optical remote sensing images," *Remote Sens.*, vol. 15, no. 18, 2023, Art. no. 4559.
[24] C. Li, B. Zhang, D. Hong, J. Yao, and J. Chanussot, "LRR-Net: An interpretable deep unfolding network for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5513412.
[25] R. Del Prete, M. D. Graziano, and A. Renga, "First results on wake detection in SAR images by deep learning," *Remote Sens.*, vol. 13, no. 22, 2021, Art. no. 4573.
[26] F. Xue, W. Jin, S. Qiu, and J. Yang, "Rethinking automatic ship wake detection: State-of-the-Art CNN-based wake detection via optical images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5613622.
[27] L. Liu, Z. Pan, and B. Lei, "Learning a rotation invariant detector with rotatable bounding box," 2017, *arXiv:1711.09405*.
[28] J. Ding, N. Xue, Y. Long, G.-S. Xia, and Q. Lu, "Learning RoI transformer for oriented object detection in aerial images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 2844–2853.
[29] Z. Zhang, W. Guo, S. Zhu, and W. Yu, "Toward arbitrary-oriented ship detection with rotated region proposal and discrimination networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 11, pp. 1745–1749, Nov. 2018.
[30] J. Han, J. Ding, J. Li, and G.-S. Xia, "Align deep features for oriented object detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5602511.
[31] P. J. Huber, "Robust estimation of a location parameter," in *Breakthroughs in Statistics: Methodology and Distribution*. Berlin, Germany: Springer, 1992, pp. 492–518.
[32] R. R. Muller and W. H. Gerstacker, "On the capacity loss due to separation of detection and decoding," *IEEE Trans. Inf. Theory*, vol. 50, no. 8, pp. 1769–1778, Aug. 2004.
[33] X. Yang and J. Yan, "Arbitrary-oriented object detection with circular smooth label," in *Proc. 16th Eur. Conf. Comput. Vis.*, 2020, pp. 677–694.
[34] L. Wen, Y. Cheng, Y. Fang, and X. Li, "A comprehensive survey of oriented object detection in remote sensing images," *Expert Syst. Appl.*, vol. 224, 2023, Art. no. 119960.
[35] X. Yang, J. Yan, Q. Ming, W. Wang, X. Zhang, and Q. Tian, "Rethinking rotated object detection with Gaussian Wasserstein distance loss," in *Proc. Int. Conf. Mach. Learn.*, 2021, pp. 11830–11841.
[36] X. Yang et al., "Learning high-precision bounding box for rotated object detection via Kullback-Leibler divergence," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 34, pp. 18381–18394, 2021.
[37] L. Huang, B. Jiang, S. Lv, Y. Liu, and Y. Fu, "Deep-learning-based semantic segmentation of remote sensing images: A survey," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, pp. 8370–8396, 2024.
[38] H. Yi, B. Liu, B. Zhao, and E. Liu, "Small object detection algorithm based on improved YOLOv8 for remote sensing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, pp. 1734–1747, 2024.
[39] M. Guan et al., "DiffWater: Underwater image enhancement based on conditional denoising diffusion probabilistic model," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, pp. 2319–2335, 2024.
[40] S. Hu, F. Gao, X. Zhou, J. Dong, and Q. Du, "Hybrid convolutional and attention network for hyperspectral image denoising," *IEEE Geosci. Remote Sens. Lett.*, vol. 21, 2024, Art. no. 5504005.
[41] Y. Ioannou, D. Robertson, R. Cipolla, and A. Criminisi, "Deep roots: Improving CNN efficiency with hierarchical filter groups," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 5977–5986.
[42] R. Yang, Z. Pan, X. Jia, L. Zhang, and Y. Deng, "A novel CNN-based detector for ship detection based on rotatable bounding box in SAR images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 1938–1958, 2021.
[43] X. Li, Z. Cai, and X. Zhao, "Oriented-YOLOv5: A real-time oriented detector based on YOLOv5," in *Proc. 7th Int. Conf. Comput. Commun. Syst.*, 2022, pp. 216–222.
[44] K. Zhao, R. Lu, S. Wang, X. Yang, and F. Lian, "Rotating target detection for nearshore SAR ships based on improved YOLOv7," *Proc. SPIE*, vol. 12960, 2023, Art. no. 1296002.
[45] G. Jocher, A. Chaurasia, and J. Qiu, "YOLOv8," 2023. [Online]. Available: https://github.com/ultralytics/ultralytics
[46] G. Jocher, K. Nishimura, and T. Mineeva, "YOLOv5," 2020. [Online]. Available: https://github.com/ultralytics/yolov5
[47] 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.
[48] X. Xie, G. Cheng, J. Wang, X. Yao, and J. Han, "Oriented R-CNN for object detection," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2021, pp. 3500–3509.
[49] G. Cheng et al., "Anchor-free oriented proposal generator for object detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5625411.
[50] G. Cheng et al., "Dual-aligned oriented detector," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5618111.
[51] T. Zhang et al., "Posterior instance injection detector for arbitrary-oriented object detection from optical remote-sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5623918.

Junfei Chen received the B.Eng. degree in communication engineering in 2018 from the School of Information and Communication Engineering, Hainan University, Haikou, China, where he is currently working toward the M.Eng. degree in artificial intelligence with the School of Information and Communication Engineering.

His main research interests include image object detection and digital image processing.

Zhuhua Hu (Senior Member, IEEE) received the B.Eng. degree in biological and agricultural engineering and M.Eng. degree in computer software and theory from Jilin University, Changchun, China, in 2002 and 2005, respectively, and the Ph.D. degree in information and communication engineering from Hainan University, Haikou, China, in 2019.

He was a Software Engineer with Ningbo BIRD Research Institute of China, Ningbo, China, from 2005 to 2006. He was a Software Engineer with the Nanjing Research Institute of ZTE, Nanjing, China, from 2006 to 2007. He was a Minister of the Software Department, Shanghai Aoxun Information Technology Company, Ltd., Shanghai, China, from 2007 to 2009. He has been a Professor and Doctorial Tutor with the School of Information and Communication Engineering, Hainan University since 2020. He has authored or coauthored more than 110 academic papers in journals such as IEEE TRANSACTIONS ON MULTIMEDIA, IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, IEEE TRANSACTIONS ON COMPUTER-AIDED DESIGN OF INTEGRATED CIRCUITS AND SYSTEMS, IEEE JOURNAL OF OCEANIC ENGINEERING, and *Computers and Electronics in Agriculture*, authorized 13 patents, and hosted more than ten large-scale commercial projects that have been successfully implemented. His current research interests include artificial intelligence, signal, and information processing.

Dr. Hu is currently an IEEE Senior Member and CCF Senior Member. He is currently a high-level talent in Hainan Province. He led the "Multimodal information intelligent processing and decision control" innovation team. He was the Reviewer for IEEE INTERNET OF THINGS JOURNAL, *Ocean Engineering*, IEEE-ACM TRANSACTIONS ON NETWORKING, *Engineering Applications of Artificial Intelligence*, *Image and Vision Computing*, IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTING, 2023 International Conference on Acoustics, Speech, and Signal Processing, and 2024 IEEE International Conference on Multimedia and Expo.

Yaochi Zhao received the M.S. degree in pattern recognition and intelligent system from Central South University, Changsha, China, in 2005 and the Ph.D. degree in computer application technology from Tianjin University, Tianjin, China, in 2023.

She was with Ningbo BIRD Research Institute, Ningbo, China, and Shanghai Wingtech Communication Company, Ltd., Shanghai, China, for three years. Later, she was engaged in teaching and research work with the College of Information Science and Technology, Hainan University, Haikou, China. She is currently an Associate Professor with the School of Cyberspace Security, Hainan University. Her current research interests include image processing and computer vision.

Wei Wu received the B.Eng. degree in electronic information engineering from the Beijing University of Chemical Technology, Beijing, China, in 2003, and the M.Eng. degree in software engineering from Tsinghua University, Beijing, China, in 2008. He is currently working toward the Ph.D. degree in communication and information engineering with the School of Information and Communication Engineering, Hainan University, Haikou, China.

His main research interests include digital image processing and computer systems.

Lihang Chen received the B.Eng. degree in intelligent science and technology in 2019 from the School of Information and Communication Engineering, Hainan University, Haikou, China, where he is currently working toward the M.Eng. degree in information and communication engineering with the School of Information and Communication Engineering.

His main research interests include image object detection and multimodal information processing.

Yueming Hu received the B.S. degree in forestry from Central South Forestry University, Changsha, China, in 1984, the M.S. degree in soil science from Northwest Agricultural University, Xianyang, China, in 1990, and the Ph.D. degree in soil science from Zhejiang Agricultural University, Hangzhou, China, in 1997.

He is currently a Professor with the College of Tropical Agriculture and Forestry, Hainan University, Haikou, China. He has authored more than 400 research papers in which 100 papers indexed by SCI. His research interests include monitoring and evaluation of land resources and GIS applications.

Ba Huang received the M.Eng. degree in communication and information systems from the Changsha University of Science and Technology, Changsha, China, in 2012.

She has over ten years of extensive experience in big data analysis and satellite remote sensing applications. Her current research interests include data analysis, maritime target identification, and more. She was responsible for the Hainan Provincial Key R&D Project—the "Precision Intelligent Joint Control Integration" combat system and demonstration application subproject for offshore vessels, serving as the person in charge of the subproject. She was in charge of extracting remote sensing data information to create corresponding products and identifying maritime vessels.