







IA-CIOU: An Improved IOU Bounding Box Loss Function for SAR Ship Target Detection Methods

Pingping Huang , Shihao Tian , Yun Su , *Member, IEEE*, Weixian Tan , *Member, IEEE*, Yifan Dong ,
and Wei Xu , *Member, IEEE*

Abstract—Ship detection in synthetic aperture radar (SAR) images is crucial in both civilian and military fields, offering extensive application prospects. Nonetheless, owing to the distinctive characteristics of SAR imaging, this task confronts numerous challenges. Specifically, ships with high aspect ratios, dense arrangements and small sizes in complex environments frequently yield in suboptimal positioning effects, consequently impacting detection performance. In response to the challenges in ship target detection, this article introduces a novel approach, termed Inner-alpha-CIOU (IA-CIOU), that relies on an enhanced intersection over union (IOU). Primarily, the method introduces Inner IOU, which effectively regulates generation of auxiliary bounding boxes through scale factor r . This ensures a better fit for dimensions of ship target frames, thereby enhancing target detection performance as well as expediting model convergence. Subsequently, this method introduces Alpha IOU, enhancing robustness of small-size ship targets in complex backgrounds by adjusting α . This allows the detector to achieve greater flexibility in ship regression accuracy. Following numerous experimental validations, proposed algorithm consistently outperforms on both SAR-Ship-Dataset, MSAR-1.0 dataset, and SAR ship detection dataset (SSDD) dataset. This groundbreaking innovation not only possesses immeasurable practical worth, but also introduces a fresh perspective together with enlightening insights for future research efforts.

Index Terms—Intersection over union (IOU) loss function, synthetic aperture radar (SAR) images, ship detection, YOLOv5.

I. INTRODUCTION

AS A crucial target for marine surveillance and wartime attack, ships hold significant practical value in both civilian and military applications [1]. Compared with optical ship images, synthetic aperture radar (SAR) ship images play an indispensable role in various fields, including target detection and recognition [2], [3], [4], maritime surveillance [5], disaster prevention [6], and surface segmentation along with

classification [7]. Because it can provide high-quality images regardless of day or night and adverse weather conditions, it has become an essential tool in these fields [8]. Numerous cutting-edge remote sensing technologies have emerged, enabling acquisition of high-resolution, extensive SAR images, greatly aiding related research efforts. However, in complex environments such as near shore, far shore, inland rivers, small islands, as well as reefs, SAR ship detection encounters immense challenges. The presence of small-sized ship targets together with clutter resembling ships makes accurate detection of ships in these scenarios a challenging endeavor. In these environments, background noise, partial occlusions of ships, as well as high aspect ratio of ships all make it challenging for current target detection methods to achieve satisfactory results.

To address challenges of ship detection, researchers have developed various algorithms [9], [10], [11]. From a temporal classification, ship detection algorithms comprise both traditional methods and currently popular deep learning algorithms. Traditional ship detection algorithms encompass sea-land segmentation, preprocessing, prescreening, and identification. Among these, a widely utilized traditional detection method is constant false alarm rate (CFAR) [12], [13]. With false alarm probability held constant, CFAR dynamically adjusts detection threshold in response to changes in clutter, thereby optimizing target detection probability. Nevertheless, balancing the recall rate along with false alarm rate remains a challenge for CFAR. To solve this problem, Migliaccio et al. [14] have created an innovative model that combines CFAR with a physical model. By integrating CFAR with a physical model, the proposed approach effectively suppresses clutter and improves algorithm performance. However, traditional algorithms suffer from poor adaptability, as their migration application capabilities are limited, so the detection results of the model can be affected by background changes.

Deep learning has been widely employed in computer vision for numerous years [15], [16]. This approach has achieved remarkable success in these fields, leading to significant attention from both academic and industrial circles [17], [18], [19]. Therefore, traditional detection methods are being gradually replaced by those based on deep learning, indicating a promising application outlook. The algorithm for ship detection based on deep learning comprises both one-stage and two-stage algorithms. One-stage algorithm achieves target detection through a single network branch, eliminating the intricate feature extraction together with candidate box generation steps present in two-stage

Manuscript received 3 February 2024; revised 1 April 2024; accepted 7 May 2024. Date of publication 17 May 2024; date of current version 5 June 2024. This work was supported in part by the Joint Funds of the National Natural Science Foundation of China under Grant U22A2010, in part by the Inner Mongolia Autonomous Region 2022 Science and Technology Leading Talent Team Project under Grant 2022LJRC0002, in part by the Project Plan for Basic Research Projects under Grant JY20220077 and Grant JY20220174, and in part by the Key Industry Development Special Fund Project of Inner Mongolia. (*Corresponding author: Weixian Tan.*)

The authors are with the College of Information Engineering, Inner Mongolia University of Technology, Hohhot 010051, China, and also with the Inner Mongolia Key Laboratory of Radar Technology and Application, Hohhot 010051, China (e-mail: hwangpp@imut.edu.cn; tsh127140@163.com; suyun@imut.edu.cn; wxtan@imut.edu.cn; yfdong@imut.edu.cn; xuwei1983@imut.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2024.3402540

target detection algorithms. Consequently, it boasts a rapid target detection speed. Notable one-stage detection algorithms include YOLOv1 [20], SSD [21], RetinaNet [22], YOLOv5, YOLOv7 [23], YOLOv8, and the latest YOLOv9 [24]. The two-stage algorithm first extracts object region, then classifies together with identifies the region. This approach typically exhibits high detection accuracy, yet it often exhibits relatively slow detection speeds. Typical two-stage algorithms include R-CNN [25], Faster RCNN [26], Mask R-CNN [27], and Cascade R-CNN [28]. In practical applications, selecting appropriate methods requires careful consideration of specific application scenarios, as well as balancing demands. To address their shortcomings, researchers are continuously optimizing and enhancing these two target detection algorithms. As a result, numerous researchers have integrated deep learning techniques into SAR ship images to enhance ship detection tasks. Based on Mask-RCNN, Nie et al. [29] integrated an attention module, resulting in improved detection results. In order to solve problems of low detection performance in small size under complex background, Zha et al. [30] introduced a novel method that effectively resolved the problems by utilizing multiple feature map transformations and fusion. Li et al. [31] combined diverse network features to effectively solve problems of diverse ship scales. In addition, numerous attention modules have emerged to mitigate noise interference. Notable modules include CBAM [32] and SE [33] modules. These techniques have demonstrated impressive outcomes in SAR ship detection, offering fresh perspectives and approaches for SAR ship target identification.

The methods of deep learning exhibit strong expressive power and excellent nonlinear fitting capabilities, making them highly competitive. In natural scenes, optical images contain color features as well as detailed texture information that can be leveraged for target detection based on optical properties of the objects [34], [35]. However, in SAR images, image values represent electromagnetic scattering information, rendering these methods unsuitable for target detection in SAR images [36]. At the same time, many challenges will be encountered during SAR ship detection. Ship detection has a complex background, which causes ships with high aspect ratios, dense arrangements and small sizes to encounter issues such as land interference, image blurring and noise. These issues ultimately affect SAR ship imaging quality, leading to poor ship target positioning effects during target detection and ultimately impacting detection performance.

To address challenges encountered in ship detection tasks, this study introduces a novel loss function, Inner-alpha-CIOU (IA-CIOU), for YOLOv5 algorithm. This loss function comprehensively considers length and width of ships, as well as position parameters of ship center point. First, Inner-IOU [37] is introduced to enhance IOU. This allows for effective control of auxiliary bounding box size through scale factor r , thereby expediting model regression. Second, based on Inner-intersection over union (IOU), Alpha IOU [38] is introduced, and sensitivity of loss function is adjusted using parameter α . Finally, YOLOv5 is enhanced by replacing complete intersection over union (CIOU) with novel IA-CIOU. In this article, a meticulous series of experiments have been

carefully crafted across three publicly available ship datasets. On SAR-Ship-Dataset, the article undertakes a thorough comparison and analysis of various IOU algorithms alongside conventional detection methods, aiming to comprehensively assess the performance of our proposed approach. Furthermore, ablation experiments were crafted on MSAR-1.0 dataset to examine distinct functions of each component within IA-CIOU, aiming to disclose their individual impacts on overall performance. Finally, cross-experiments were systematically conducted on SSDD dataset to comprehensively demonstrate noteworthy generalization ability and robust effectiveness of IA-CIOU. Through meticulously designed experiments, superior performance and practical applicability of IA-CIOU in realm of ship detection have been unequivocally confirmed. Following are the key contributions.

- 1) The aim of IA-CIOU is to enhance CIOU loss function, enabling faster bounding box regression (BBR) and improved model generalization, ultimately leading to superior algorithm performance.
- 2) The improved algorithm has better detection ability for high aspect ratio, dense arrangement, and small size ship targets in complex background.
- 3) A large number of experiments have been carefully designed to prove the effectiveness of IA-CIOU and achieve significant performance enhancements on public datasets.

In conclusion, IA-CIOU addresses some challenges in ship detection tasks, offering valuable insights for future research in this area.

II. RELATED METHODS

This section aims to review and comprehensively summarize prior research outcomes related to object detection algorithms and optimization of loss functions.

A. Target Detection Algorithm

As technology progresses, numerous target detection algorithms have emerged. Among traditional methods, sliding window method and feature pyramid method (FPN) [39] are particularly prevalent. Sliding window method slides windows of various sizes and aspect ratios across image, and employs a classifier to determine whether window contains target. However, this method incurs significant computational overhead and struggles to effectively handle targets of varying scales and proportions. FPN detects targets of various sizes by constructing image pyramids of varying scales to process images across different scales. Since introduction of concept of deep learning by researchers, target detection algorithms have undergone significant changes. When processing images, detection model of deep learning can be divided into two categories according to processing stage. One category includes one-stage algorithms that achieve detection through direct forward propagation. While these algorithms are fast, they may sacrifice accuracy. Other algorithm is a two-stage approach. Initially, candidate boxes are utilized in image to identify potential object locations. Subsequently, each candidate box

undergoes classification and positioning to produce final detection outcome. While two-stage approach offers high accuracy, it comes with a high computational complexity and relatively slow speed. To exploit advantages of both algorithms, several hybrid methods have been gradually introduced. These methods combine direct regression with candidate region generation techniques to achieve faster speeds and enhanced accuracy. In practical applications, choosing appropriate algorithm requires careful consideration of specific requirements, computational resources, and performance needs.

After thoroughly analyzing strengths and limitations of both traditional methods and deep learning methods, as well as various algorithms within deep learning, this article made informed decision to adopt YOLOv5 algorithm for further enhancement. YOLOv5 excels in predicting position of target boxes and their corresponding categories through direct regression, making it an effective choice for our needs. Network architecture encompasses backbone network, neck segment and head segment. To enhance prediction precision and model generalization, several preprocessing techniques are utilized, including image augmentation and adaptive anchor frame. Additionally, Mosaic data enhancement technology is employed. Backbone network employs CSPDarknet53 network, while neck segment utilizes PANet [40]. PANet is an advanced instance segmentation framework that utilizes bottom-up path enhancement and low-level positioning information to enhance overall feature representation, thereby optimizing information flow. Head segment generates three distinct scales of feature maps, a prediction box is generated based on features at varying scales. It undergoes nonmaximum suppression processing. Network structure of YOLOv5 is displayed in Fig. 1.

B. Optimization of Loss Function

The two core tasks of detector are classification and regression of objects. Classification task aims to classify detected items, while regression task aims to accurately localize detected items, such as ships. Therefore, total loss for each anchor frame comprises two components: one is classification loss, which is closely linked to category recognition, and the other is regression loss, which is intimately connected to location positioning. In target detection, mathematical expression for total loss typically reads as follows [41]:

$$\text{LOSS}_{\text{all}}(p, p^*, l, l^*) = \alpha_1 \text{LOSS}_{\text{cls}}(p, p^*) + \beta \text{LOSS}_{\text{reg}}(l, l^*). \quad (1)$$

In the formula, l stands for position of prediction box, l^* represents position of ground truth box, p represents probability of being classified into a specific category for prediction box, and p^* is corresponding category probability for ground truth box. Classification loss LOSS_{cls} and regression loss LOSS_{reg} together constitute total loss in target detection. These two types of losses collaboratively determine performance. α_1 and β are two hyperparameters utilized to regulate weights of class loss and regression loss. By adjusting these two parameters, emphasis of model on classification accuracy and regression accuracy can be evenly balanced.

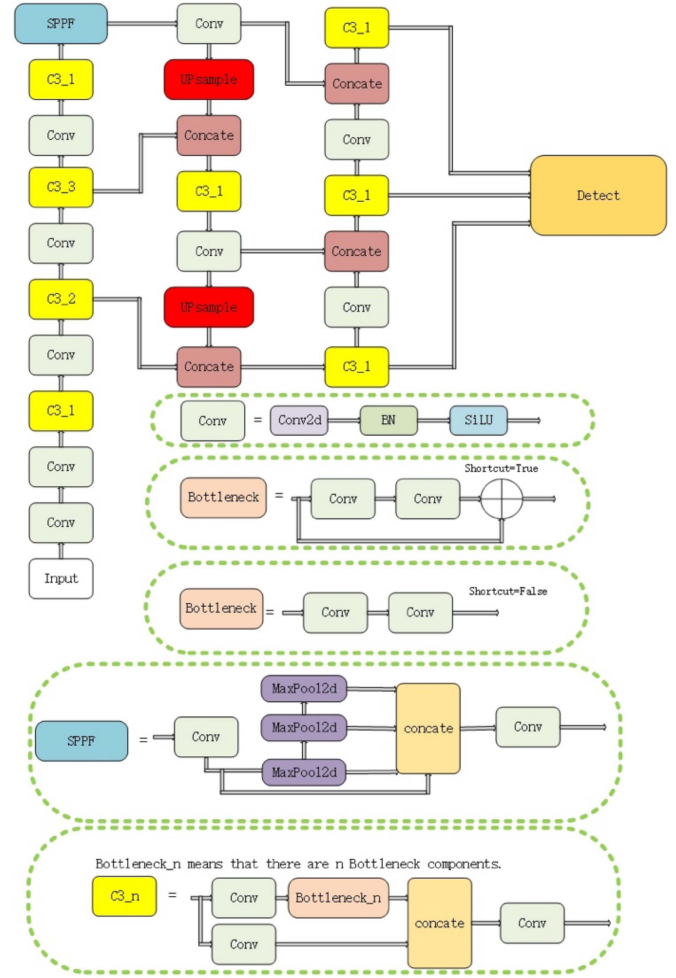


Fig. 1. YOLOv5 network structure diagram.

In target detection, majority of methods rely on BBR module to accurately identify target location. To ensure the effectiveness of BBR, it is essential to develop an appropriate loss function. Currently, there are two main categories of BBR loss function: norm-based [42] and IOU-based [43]. While initial norm-based loss function was simple, it was highly susceptible to scale variations. To overcome limitations of norm-based loss function, researchers have introduced refined techniques. For instance, YOLOv1 effectively mitigates impact of large bounding boxes on target detection by adjusting their size and implementing square root transformation. In addition, YOLOv3 [44] introduces a novel loss function that aims to enhance target detection performance by adjusting offset and scale of anchor box. These norm-based functions calculate regression loss for four corners of bounding box independently, but without considering their correlation, they tend to bias model toward larger objects. While four corners may appear to be independent, they are actually interrelated.

To address the aforementioned limitations, researchers proposed IOU loss as an alternative to ln-norm loss. IOU loss treats each of four corners of candidate box holistically, ensuring that values remain nonnegative, satisfy identity, exhibit symmetry, and satisfy triangle inequality. In comparison with

In-norm loss, IOU loss offers scale invariance, meaning its output remains within range of [0, 1], thereby enabling more accurate predictions of regression results. IOU loss is calculated by considering intersection and union ratios and incorporating coordinate information of rectangular box. Its expression is as follows:

$$\text{inter} = \max((\min(x_2^{gt}, x_2^p) - \max(x_1^{gt}, x_1^p)), 0) * \\ \times \max((\min(y_1^{gt}, y_1^p) - \max(y_2^{gt}, y_2^p)), 0) \quad (2)$$

$$\text{union} = (x_2^{gt} - x_1^{gt}) * (y_1^{gt} - y_2^{gt}) \\ + (x_2^p - x_1^p) * (y_1^p - y_2^p) - \text{inter} \quad (3)$$

$$\text{IOU} = \frac{\text{inter}}{\text{union}} \quad (4)$$

$$L_{\text{IOU}} = 1 - \text{IOU}. \quad (5)$$

Here, the coordinates of top-left and bottom-right corners of prediction box b^p are (x_1^p, y_1^p) and (x_2^p, y_2^p) , respectively. Similarly, top-left and bottom-right coordinates of ground truth box b^{gt} are (x_1^{gt}, y_1^{gt}) and (x_2^{gt}, y_2^{gt}) , respectively. By employing IOU loss, similarity between two bounding boxes can be more accurately evaluated, thereby improving accuracy and robustness of model. The range of IOU loss spans from 0 to 1. As IOU approaches 1, the more similar two frames are, resulting in more precise positioning. Consequently, corresponding loss function value nears 0, leading to smaller penalties. Conversely, as IOU decreases, penalty increases accordingly. Nevertheless, IOU Loss has some limitations: when there is no common area between two boxes, IOU value is 0, which prevents effective gradient backpropagation by loss function. Consequently, in certain scenarios, model fails to effectively update its parameters via gradient descent method, thereby compromising its training and performance.

To address this limitation, GIOU [45] enhances IOU loss by introducing a minimum bounding box term. DIOU [46] uses minimum center point distance between two boxes as a penalty term, effectively addressing issue where GIOU fails to distinguish relationship between two boxes when they are inclusive. CIOU loss [46] considers impact of shape similarity on BBR and incorporates shape loss term in DIOU loss. Since area and center distance are already considered in DIOU Loss, CIOU Loss introduces an additional penalty term to adjust aspect ratio of prediction box. Definition of CIOU loss is as follows:

$$L_{\text{CIOU}} = 1 - \text{IOU} + \frac{\rho^2}{C^2} + \alpha_2 v \quad (6)$$

$$\alpha_2 = \frac{v}{1 - \text{IOU} + v} \quad (7)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w_g}{h_g} - \arctan \frac{w_p}{h_p} \right)^2. \quad (8)$$

In CIOU function, ρ signifies distance between center points of two boxes, while C denotes diagonal length of smallest bounding box. Furthermore, w_g and h_g stand for width and height of ground truth box, respectively, whereas w_p and h_p represent width and height of predicted box. CIOU aims to maintain identical aspect ratios and diagonal angles between

frames to ensure consistent shapes. During optimization, CIOU Loss not only focuses on common area and center distance between two boxes, but it also emphasizes adjusting length-width ratio of prediction box. By incorporating shape loss, CIOU aims for more precise interframe matching, ultimately enhancing positioning accuracy and model performance.

Numerous researchers have introduced various detection algorithms, and bounding box loss function has undergone constant innovation in recent years. For instance, EIOU penalty term [47] is derived from CIOU penalty term, and shape loss is separated to independently optimize length and width of candidate box. SIOU [48] introduces angle between two boxes as a novel constraint on BBR loss, leading to a faster convergence rate. MPDIOU [49] uses corner distance between two boxes as a penalty term, simplifying calculation process and significantly enhancing detection performance. Furthermore, WIOU [50] not only takes into account overlap between two frames, but also places emphasis on region between bounding boxes, exhibiting dynamic nonmonotonic FM property. Inner IOU loss function series proposed by Zhang Hao not only takes into account aforementioned geometric constraints, but also emphasizes rationality of IOU itself. By utilizing auxiliary bounding box calculations for regression acceleration, Inner IOU loss function effectively expedites convergence speed, thereby further enhancing detection performance.

Throughout their research, numerous experts have integrated innovative loss functions into ship target detection algorithms. For instance, Guo et al. [51] analyzed angle information of ship, introduced SIOU into model, and successfully completed ship detection task. Additionally, Yu et al. [52] developed ECIOU, an advancement upon CIOU. Their primary objectives were to address two key challenges: fragility of penalty term v in terms of outliers and limited value range of tangent function, which did not meet normalization requirements of loss function. To achieve numerical normalization, CIOU incorporates unique coefficients, leading to increased computational complexity. Distance loss function proposed by Zhou et al. [53] integrates IOU for comprehensive evaluation and utilizes distance between two coordinate points to enhance ship target detection accuracy. These studies continue to drive progress of ship detection task.

III. PROPOSED METHODS AND MODEL ARCHITECTURE

This section will explore the working principle of Inner IOU. Later, it will separately and in depth explore Inner CIOU and IA-CIOU.

A. Inner IOU

To address the issue of existing IOU loss that exhibits poor generalization and slow convergence speed in various detection tasks, this study introduces an innovative approach: utilizing auxiliary bounding boxes for loss computation. This technique significantly expedites BBR process. In Inner IOU, scaling factor r is used to adjust size of auxiliary box. By utilizing auxiliary bounding boxes of various scales in various datasets and detectors, limitations of existing methods in generalization have been successfully overcome.

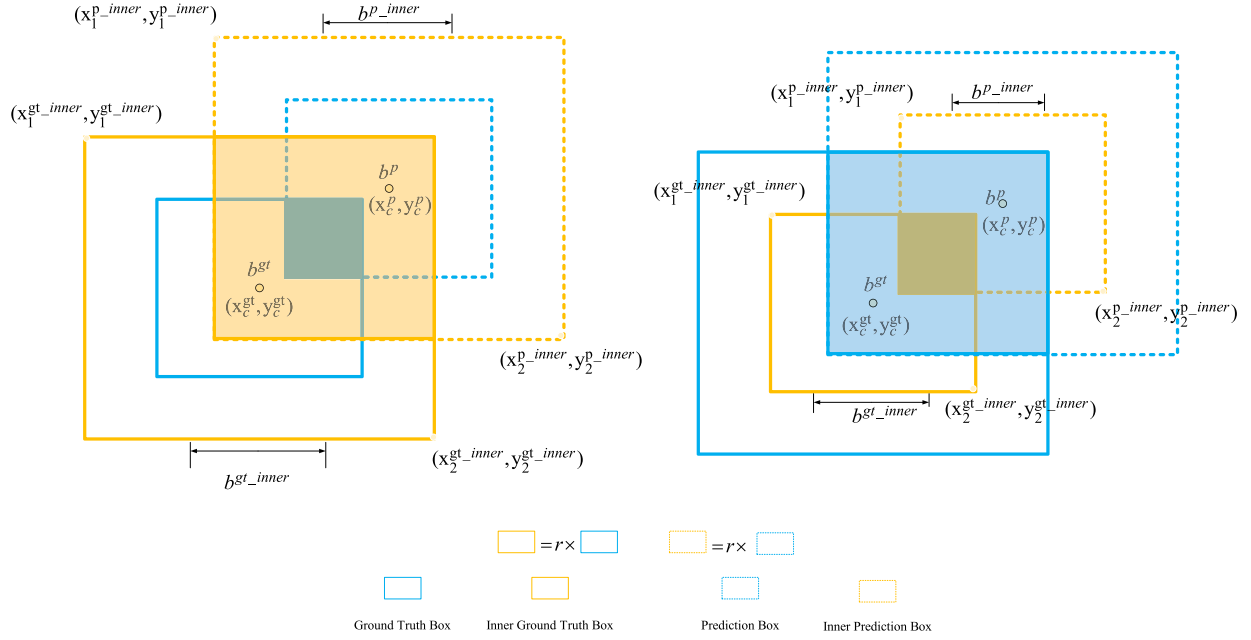


Fig. 2. Auxiliary boxes are aligned precisely with center points of both ground truth box and prediction box. By using scaling factor r , size of auxiliary box can be adjusted to meet needs of actual scene. On the left side of figure, it is evident that when value of scale factor r exceeds 1, auxiliary box enlarges proportionally. Conversely, on the right side of figure, a scale factor r below 1 result in auxiliary box contracting.

As demonstrated in Fig. 2, algorithm generates two auxiliary boxes, b^{gt_inner} and b^{p_inner} . Auxiliary box b^{gt_inner} created by ground truth box b^{gt} has vertex coordinates $(x_1^{gt_inner}, y_1^{gt_inner})$ and $(x_2^{gt_inner}, y_2^{gt_inner})$, respectively, while auxiliary box b^{p_inner} created by prediction box b^p has vertex coordinates $(x_1^{p_inner}, y_1^{p_inner})$ and $(x_2^{p_inner}, y_2^{p_inner})$. The expression for Inner IOU is as follows:

$$x_1^{gt_inner} = x_c^{gt} - \frac{(x_2^{gt} - x_1^{gt})}{2} * r \quad (9)$$

$$x_2^{gt_inner} = x_c^{gt} + \frac{(x_2^{gt} - x_1^{gt})}{2} * r \quad (10)$$

$$y_1^{gt_inner} = y_c^{gt} + \frac{(y_1^{gt} - y_2^{gt})}{2} * r \quad (11)$$

$$y_2^{gt_inner} = y_c^{gt} - \frac{(y_1^{gt} - y_2^{gt})}{2} * r \quad (12)$$

$$x_1^{p_inner} = x_c^p - \frac{(x_2^p - x_1^p)}{2} * r \quad (13)$$

$$x_2^{p_inner} = x_c^p + \frac{(x_2^p - x_1^p)}{2} * r \quad (14)$$

$$y_1^{p_inner} = y_c^p + \frac{(y_1^p - y_2^p)}{2} * r \quad (15)$$

$$y_2^{p_inner} = y_c^p - \frac{(y_1^p - y_2^p)}{2} * r. \quad (16)$$

In the formula, auxiliary box b^{gt_inner} and ground truth box b^{gt} share a central point (x_c^{gt}, y_c^{gt}) , and auxiliary box b^{p_inner} and prediction box b^p also share a central point (x_c^p, y_c^p) . Scaling factor r determines size of auxiliary box. If r exceeds 1, auxiliary box expands. In the case of low IOU samples, absolute value of

IOU gradient surpasses absolute value of gradient pertaining to ground truth box IOU. Conversely, when r is less than 1, auxiliary box contracts. For high IOU samples, absolute IOU gradient of smaller auxiliary box exceeds absolute value of gradient of ground truth box IOU. Based on this analysis, employing auxiliary boxes of smaller scales for IOU loss calculation enhances regression of high IOU samples, thus facilitating faster convergence. Conversely, use of auxiliary boxes of larger scales for IOU loss calculations expedites regression process of low IOU samples. Inner IOU is defined as follows:

$$\text{winter}^{inner} = \max \left(\left(\min \left(x_2^{gt_inner}, x_2^{p_inner} \right) - \max \left(x_1^{gt_inner}, x_1^{p_inner} \right) \right), 0 \right) \quad (17)$$

$$\text{hinter}^{inner} = \max \left(\left(\min \left(y_1^{gt_inner}, y_1^{p_inner} \right) - \max \left(y_2^{gt_inner}, y_2^{p_inner} \right) \right), 0 \right) \quad (18)$$

$$\text{inter}^{inner} = \text{winter}^{inner} * \text{hinter}^{inner} \quad (19)$$

$$\text{gtunion}^{inner} = \left(x_2^{gt_inner} - x_1^{gt_inner} \right) * \left(y_1^{gt_inner} - y_2^{gt_inner} \right) \quad (20)$$

$$\text{punion}^{inner} = x_2^{p_inner} - x_1^{p_inner} * \left(y_1^{p_inner} - y_2^{p_inner} \right) \quad (21)$$

$$\text{union}^{inner} = \text{gtunion}^{inner} + \text{punion}^{inner} - \text{inter}^{inner} \quad (22)$$

$$\text{IOU}^{inner} = \frac{\text{inter}^{inner}}{\text{union}^{inner}}. \quad (23)$$

Inner IOU loss exhibits its unique advantages by inheriting characteristics of IOU loss. Inner IOU, like IOU loss, has a value range of [0, 1]. Due to similarity in scale between two frames, calculation of Inner IOU is identical to that of IOU loss. Given significant proportion of high IOU samples within ship dataset, this article specifically assigns a value of $r = 0.8$. By adopting this configuration, absolute value of IOU gradient exhibited by auxiliary box of a smaller scale surpasses absolute value of IOU gradient at ground truth box. This particular setting contributes to enhancing convergence rate of model, ultimately leading to an improvement in its detection performance.

B. Inner CIOU

By integrating Inner IOU and CIOU, this article has successfully developed Inner CIOU. Inner CIOU Loss introduces a novel penalty term to further optimize shape of predicted box, beyond what is considered in DIOU Loss. This penalty term ensures that shapes of both boxes remain consistent. The mathematical expression for Inner CIOU Loss is as follows:

$$\begin{aligned} L_{\text{inner-CIOU}} &= L_{\text{CIOU}} + \text{IOU} - \text{IOU}^{\text{inner}} \\ &= 1 - \text{IOU}^{\text{inner}} + \frac{\rho^2}{C^2} + \alpha_2 v \end{aligned} \quad (24)$$

$$\alpha_2 = \frac{v}{1 - \text{IOU} + v} \quad (25)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w_g}{h_g} - \arctan \frac{w_p}{h_p} \right)^2. \quad (26)$$

The Inner CIOU Loss is designed to achieve a more harmonious shape between two boxes by balancing their width-to-height ratio and diagonal tilt angle. This ensures a more precise alignment of their shapes. Not only does Inner CIOU uphold rationality of IOU, but it also significantly enhances convergence speed of model by incorporating an auxiliary box mechanism. Furthermore, during optimization, Inner CIOU Loss takes into account not only common area and center distance between two boxes but also strives to maintain a consistent shape between them through shape loss, thus adjusting shape of prediction box.

C. IA-CIOU

This study introduces Alpha IOU, a novel target detection loss function, aimed at enhancing generalization performance of detection algorithms across diverse tasks. By introducing an additional parameter α , it can be flexibly adjusted within a certain range. Alpha IOU introduces a weighted combination in traditional IOU calculations, enabling flexible adjustment of loss function. This allows model to better adapt to scale and shape of objects. Design concept of Alpha IOU aims to address limitations of existing IOU losses in various tasks, providing a more flexible and effective loss calculation method for enhancing development of target detection. The expression of IA-CIOU is as follows:

$$L_{\text{IA-CIOU}} = 1 - (\text{IOU}^{\text{inner}})^\alpha + \left(\frac{\rho^2}{C^2} \right)^\alpha + (\alpha_2 v)^\alpha \quad (27)$$

TABLE I
EXPERIMENTAL ENVIRONMENT

Software and Hardware	Version or Model
Operating system	Windows10
CPU	Intel(R) Core (TM) i5-11400F
GPU	NVIDIA GeForce RTX 3060
Display memory	12G
CUDA	12.2
CUDNN	V8.0.5
Pytorch version	1.11
Python version	3.7
Software	PyCharm2021

$$\alpha_2 = \frac{v}{1 - \text{IOU} + v} \quad (28)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w_g}{h_g} - \arctan \frac{w_p}{h_p} \right)^2. \quad (29)$$

Alpha IOU possesses a distinct characteristic as an adjustable index parameter. When its value exceeds 1, it effectively magnifies the weight of high IOU targets within loss function. This attribute enables detector to concentrate more intently on high IOU targets and become more responsive to their associated losses, ultimately enhancing regression precision of these targets. Given preponderance of high IOU samples within ship dataset, this study specifically sets Alpha IOU parameter to 3. By bolstering loss and gradient contributions of high IOU samples, BBR accuracy has undergone a notable enhancement. Therefore, IA-CIOU takes into account not just shared area and distance between centers of two boxes, but also incorporates shape loss to maintain congruity of their shapes. By aptly adjusting scaling factor r and parameter α , IA-CIOU can expedite convergence rate of high IOU samples within ship dataset, thereby enhancing ship detection performance even further.

IV. EXPERIMENTS

In this part, this article designs a large number of experiments on three public ship datasets. The results clearly show that IA-CIOU model performs well in mAP, precision, and recall, which verifies the effectiveness of the method.

A. Experimental Environment and Parameters

In Table I, this article presents a comprehensive overview of training environment and fundamental parameters used in this experiment. The entire experiment was conducted on a GPU, utilizing PyTorch framework for model training. Refer to Table II for a detailed display of parameters used in this experiment.

B. Data Set

To thoroughly assess the performance of IA-CIOU method introduced in this article, this chapter undertakes experimental

TABLE II
EXPERIMENTAL PARAMETERS

Parameter Name	Version or Model
weights	Yolov5s.pt
img-size	640×640
epochs	300
batch-size	16
max-det	1000
conf-thres	0.25
iou-thres	0.45
α	3
r	0.8

verification across multiple public datasets, encompassing SAR-Ship-Dataset [54], SSDD dataset [55], and MSAR-1.0 dataset [56]. SAR-Ship-Dataset, in particular, comprises a substantial collection encompassing 102 scenes from GF-3 satellite and 108 scenes from Sentinel-1 SAR satellite. This compilation boasts a comprehensive total of 43 819 ship images, averaging approximately 256×256 pixels in size. Ship scenes within this dataset exhibit remarkable diversity, encompassing both optimal and adverse sea conditions, as well as intricate landing scenarios and simpler offshore environments. Notably, a total of 59 535 ship targets have been meticulously annotated within this dataset. These targets range significantly in size, with smallest ship target measuring merely 6 pixels wide and 4 pixels tall, encompassing a mere 24 pixels in total. Conversely, largest ship target spans 129 pixels in width and 207 pixels in height, encompassing a whopping 26 703 pixels. To ensure precision and credibility of our experiments, this article has randomly selected 6000 images from this extensive dataset as representative samples. These large samples encapsulate a broad array of ship usage scenarios, including both distant offshore and near-shore environments, thereby enhancing universality and representativeness of our experimental outcomes.

MSAR-1.0 dataset boasts an extensive collection of SAR image resources, totaling 28 449 detection slices. These slices are sourced from both Haisi-1 and Gaofen-3 satellites, ensuring multisource diversity and comprehensiveness of dataset. Breadth of scenarios represented is immense, ranging from airports and ports to inshore environments, islands, open seas and urban areas, offering a rich tapestry of materials for diverse application scenarios. As for target types, dataset encompasses four categories: aircraft, oil tanks, bridges and ships. Detailed data breakdown includes 1851 bridges, 39 858 ships, 12 319 oil tanks and 6368 aircraft, providing ample labeled data for target detection tasks.

SSDD dataset comprises 1160 SAR image samples sourced from Radar Sat-2, Terra SAR-X and Sentinel-1 satellites. Each image averages approximately 484 × 329 pixels in size. These images not only capture ships in various environmental contexts, ranging from calm sea surfaces to turbulent conditions, but also

present both intricate shoreline scenarios and simpler offshore views. Dataset is extremely rich in content, encompassing a total of 2587 ships. Among them, smallest ship target is astonishingly minute, measuring just 4 pixels wide and 5 pixels tall, totaling a mere 20 pixels. In stark contrast, largest ship target is immense, boasting a width of 308 pixels and a height of 180 pixels, amounting to a whopping 55 440 pixels. This staggering size difference is remarkable, with largest ship target being a colossal 2772 times larger than smallest one.

C. Evaluation Metrics

The evaluation of model performance is crucial, and this article uses metrics such as precision and recall and mAP to evaluate the performance of IA-CIOU model. Their respective formulas are as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (30)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (31)$$

In the given formula, TP represents count of positive samples that have been accurately predicted, FP denotes number of negative samples that have been falsely predicted, and FN signifies count of positive samples that have been incorrectly predicted. In this article, IOU is used as judgment criterion. When IOU surpasses a pre-determined threshold, predicted box is considered a positive sample. Efficacy of model is then assessed by computing average precision (AP) and mAP. The detailed calculation formula is outlined as follows:

$$\text{AP} = \int_0^1 P \times R dR \quad (32)$$

$$\text{mAP} = \frac{1}{C} \sum_{i=1}^c \text{AP}_i \quad (33)$$

In the provided formula, C represents type of sample, P and R stand for precision and recall rates, respectively. These rates combine to form AP, which has a value range between 0 and 1. When AP reaches 1, it means model performs best in given task.

D. Experiments on SAR-Ship-Dataset

Many researchers have conducted extensive research on this dataset [57], [58]. To showcase the effectiveness of IA-CIOU, this paper executed rigorous experiments, utilizing same datasets and equipment conditions. Through detailed comparisons of various performance indicators, excellent performance of algorithm in various complex environments has been confirmed. This article compares popular IOU algorithms and object detection models, and thoroughly tests uniqueness and rationality of IA-CIOU algorithm.

1) *Compare the Performance of Different Inner-Alpha-IOUs:* This article conducted an experimental comparison of common IOU models, specifically selecting GIOU and DIOU as benchmark models for comparison. In order to ensure rationality of experiment, these algorithms were improved to Inner Alpha GIOU and Inner Alpha DIOU, and they were compared with

TABLE III
PERFORMANCE COMPARISON DIAGRAM OF DIFFERENT IOUS

Algorithm	R (%)	P (%)	mAP (%)
GIOU	92.0	91.6	95.5
Inner-Alpha-GIOU	92.6	93.1	96.1
DIOU	89.8	93.0	96.1
Inner-Alpha-DIOU	93.2	92.1	96.3
CIOU	92.7	90.4	95.4
IA-CIOU	92.7	91.8	96.4

IA-CIOU algorithm proposed in this paper in depth. During this process, key evaluation metrics such as recall, precision, and mAP were used to comprehensively measure performance of each algorithm. Table III provides a detailed summary of these experimental results, clearly demonstrating superior performance of IA-CIOU algorithm in various indicators.

As evident from Table III, enhanced IOU algorithm exhibits superior performance within YOLOv5 framework. Precisely, mAP value of Inner-Alpha-GIOU surpasses that of GIOU. Similarly, Inner-Alpha-DIOU and IA-CIOU exhibit higher mAP values than DIOU and CIOU, respectively. These comparisons not only firmly establish efficacy of Inner IOU and Alpha IOU in enhancing model detection accuracy but also underscore their vast applicability. Notably, IA-CIOU emerges as top performer among all improved algorithms, surpassing both Inner-Alpha-DIOU and Inner-Alpha-GIOU. This is primarily attributed to comprehensive consideration of crucial factors like overlapping area, center point distance, and length–width ratio in its design of IA-CIOU, leading to its exceptional detection performance. These discoveries further validate the effectiveness of the proposed algorithm.

2) *Compare With the Mainstream Algorithms:* This article conducted a thorough comparative analysis of IA-CIOU algorithm, not merely confining ourselves to extensive experiments with other leading target detection models, including SSD, Faster R-CNN, YOLOv7, YOLOv8, and YOLOv9. In addition, by integrating IA-CIOU algorithm into YOLOv7 and YOLOv8, algorithm was further analyzed for a more detailed comparison. To ensure breadth and accuracy of evaluation, key indicators such as precision, recall, and mAP were used for comprehensive evaluation. As evident in Table IV, IA-CIOU algorithm clearly demonstrates superior performance in detection metrics when compared to other prominent target detection models.

Table IV clearly demonstrates that IA-CIOU algorithm significantly enhances detection indicators of YOLOv5, YOLOv7, and YOLOv8 models, particularly in the case of YOLOv5 model. This remarkable effect strongly validates widespread applicability and efficient performance of IA-CIOU algorithm. Additionally, when compared to mainstream algorithms like SSD and Faster R-CNN, IA-CIOU algorithm exhibits notable advantages in detection metrics. Integrating IA-CIOU into YOLOv5 results in a 1.6% increase in mAP compared to SSD and a

TABLE IV
PERFORMANCE COMPARISON DIAGRAM OF DIFFERENT DETECTION MODELS

Algorithm	R (%)	P (%)	mAP (%)
SSD	88.5	91.5	94.8
Faster R-CNN	98.3	62.9	95.2
YOLOv9	91.4	91.5	95.8
YOLOv7(CIOU)	90.7	91.7	95.0
YOLOv7(IA-CIOU)	92.8	90.9	95.3
YOLOv8(CIOU)	92.2	91.9	95.4
YOLOv8(IA-CIOU)	94.1	91.8	96.0
YOLOv5(CIOU)	92.7	90.4	95.4
YOLOv5(IA-CIOU)	92.7	91.8	96.4

4.2% enhancement in recall. Furthermore, precision of YOLOv5 algorithm with IA-CIOU integrated surpasses Faster R-CNN by 28.9%. These substantial improvements reconfirm efficacy of IA-CIOU algorithm. It is noteworthy that even when contrasted with the latest YOLOv9 algorithm, integrating IA-CIOU into YOLOv5 still demonstrates distinct advantages, further emphasizing advancement and robust competitiveness of IA-CIOU algorithm. The significant advantage observed can be attributed to abundant ship images in this dataset, featuring high IOU samples. Thanks to combined effects of Inner IOU and Alpha IOU, convergence speed of the model has undergone a remarkable enhancement. Additionally, CIOU incorporates the shape factor of ship, ensuring more precise positioning by model, thereby further enhancing its detection capabilities. Collectively, these advantages contribute to outstanding performance of IA-CIOU algorithm in ship detection tasks.

3) *Label Visualization and Experimental Results Analysis:* In Fig. 3, this article presents a carefully chosen set of 12 ship images, representing two distinct environments: offshore and open sea. These images illustrate various scenarios, including single ships, multiple ships and ships arranged in a dense formation. Through a comprehensive analysis of ship visualization, performance of original CIOU training was compared with that of IA-CIOU training.

In contrast, IA-CIOU shows significant advantages over CIOU in terms of detection performance. Specifically, as observed in subplot (a), (d), (g), and (j) in Fig. 3, IA-CIOU has high confidence in detection of small ships with high aspect ratio in open sea environment and offshore complex background. Furthermore, as seen in subplot (b), (e), (h), and (k) in Fig. 3, under identical environmental backgrounds, IA-CIOU significantly reduces false alarm rate for such small-sized ships. Lastly, as observed in subplot (c), (f), (i), and (l) in Fig. 3, IA-CIOU exhibits a higher detection rate for this type of ship within same context. Therefore, regardless of whether it is situated in a solitary environment in far-reaching sea or within intricate backdrop of nearshore, IA-CIOU demonstrates a lower false

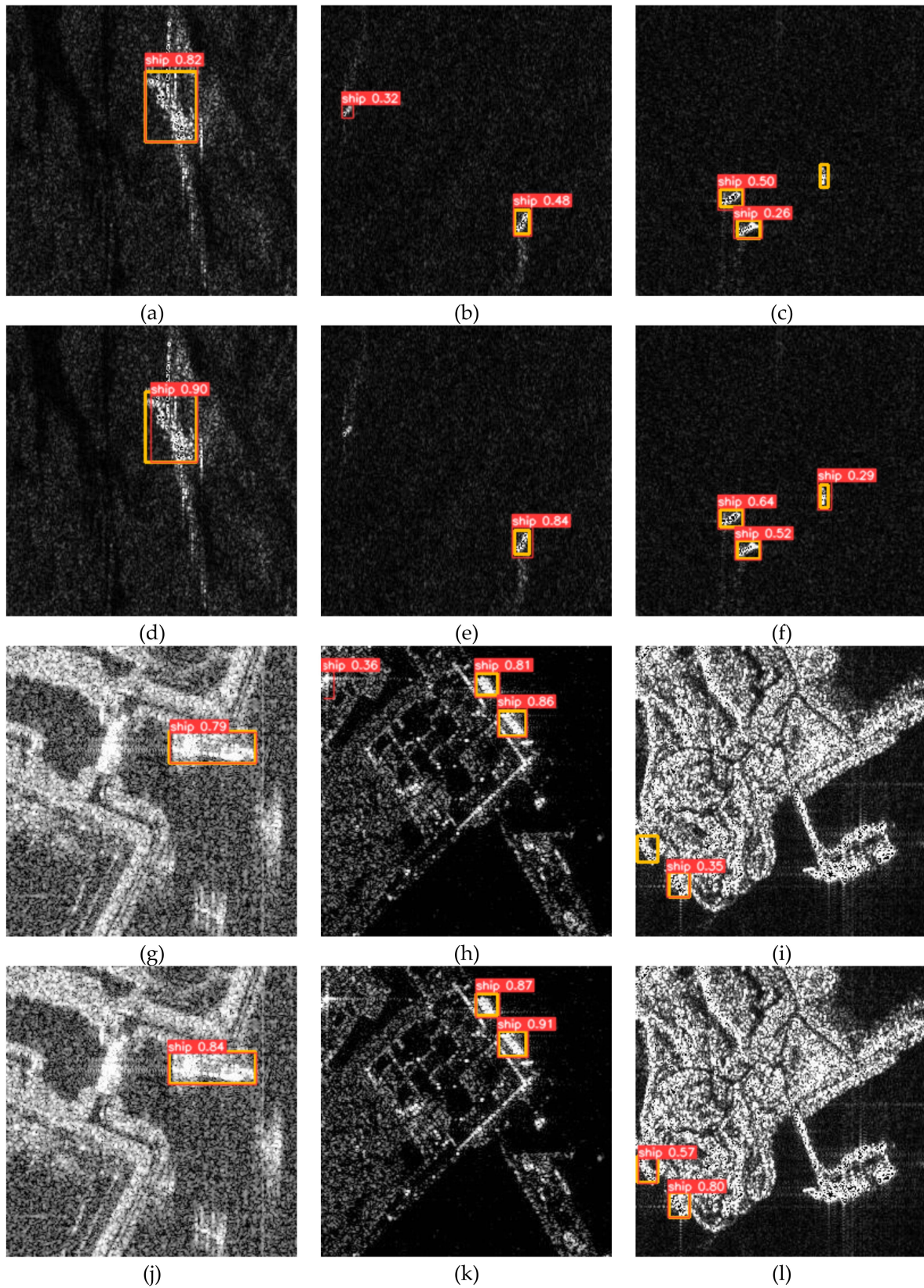


Fig. 3. Visualization results of SAR-ship-dataset. Among the 12 images, three images (a), (b), and (c) represent ship detections achieved by CIOU method in open sea environment, while three images (d), (e), and (f) demonstrate detections achieved by IA-CIOU method in same environment. Similarly, the three images of (g), (h), and (i) present ship detections using CIOU in offshore environment, and the three images of (j), (k), and (l) display detections achieved by IA-CIOU in offshore environment. In these images, the yellow box represents actual ship frame, and the red box represents predicted ship frame.

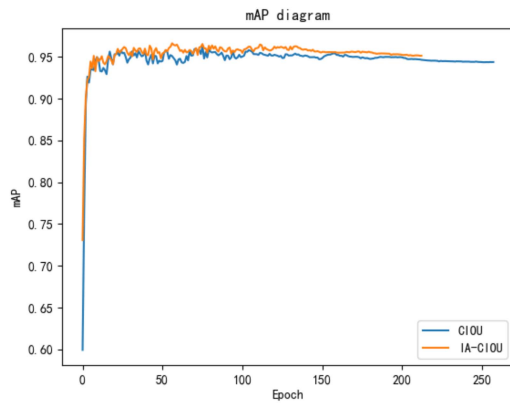


Fig. 4. Comparison chart of mAP.

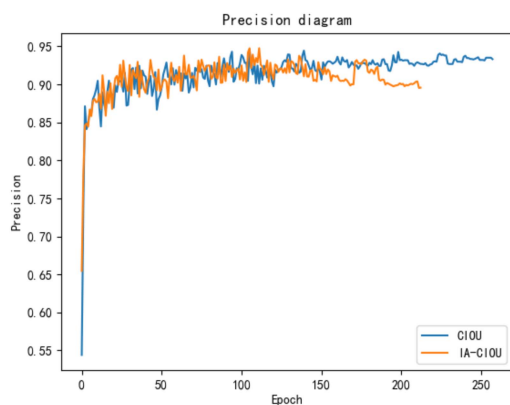


Fig. 5. Comparison chart of precision.

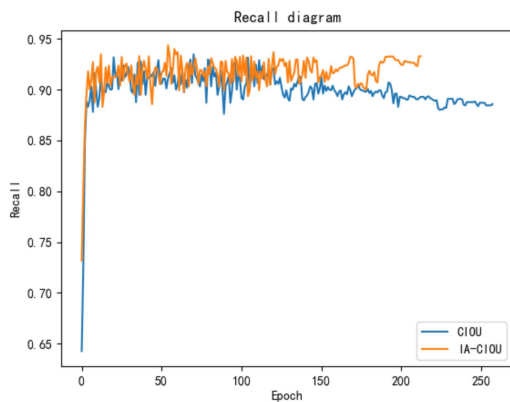


Fig. 6. Comparison chart of recall.

alarm rate and superior detection performance for ships with dense arrangements and an elevated aspect ratio.

4) *Visualization of Detection Indicators and Analysis of Experimental Results:* This article has displayed detection performance of CIOU and IA-CIOU in form of pictures, comparing mAP, recall, and precision metrics before and after refinement. A cursory examination of Figs. 4–6 shows that refined model outperforms across all indicators.

By juxtaposing outcomes obtained prior to and following enhancement, it is evident that optimized IA-CIOU has witnessed

TABLE V
TABLE OF ABLATION EXPERIMENTS

Algorithm	R(%)	P(%)	mAP(%)
IOU	86.8	93.1	92.0
CIOU	91.5	93.0	94.6
Alpha CIOU	92.7	91.8	94.9
Inner CIOU	91.8	93.1	94.9
IA-CIOU	92.0	93.5	95.4

significant improvements across key metrics such as mAP, recall, and precision. This substantial improvement offers irrefutable evidence of efficacy and superiority of the refined model.

E. Experiments on MSAR-1.0 Dataset

Numerous researchers have conducted extensive studies on this dataset [59]. In experimental section, this article conducted in-depth research on IA-CIOU algorithm. In order to more accurately understand individual contributions of each component in IA-CIOU, ablation experiments were conducted. Table V presents these findings. For ship detection tasks, CIOU loss function in YOLOv5 was used as a reference to visually demonstrate the performance differences before and after enhancement.

The analysis of Table V reveals that optimized model has demonstrated remarkable enhancements in key performance metrics during training process, including precision, recall and mAP. Notably, within YOLOv5 framework, IA-CIOU approach has yielded a 3.4% increase in mAP compared to IOU, a 5.2% boost in recall, and a 0.4% enhancement in precision. Significant improvement can be attributed to CIOU algorithm, which takes into account not only the length and width of the ship, but also the distance of its center point. This consideration has significantly enhanced detection accuracy. Furthermore, this article has innovatively combined Alpha IOU with CIOU to create Alpha CIOU. This innovation enables model to excel in detecting and identifying ship targets with greater accuracy, particularly in small ship datasets with intricate backgrounds. Additionally, integration of Inner IOU and CIOU, known as Inner CIOU, has also significantly improved the performance of model. It not only elevates mAP, recall, and precision values but also markedly accelerates convergence speed of model, thereby further optimizing its detection capabilities.

The IA-CIOU algorithm significantly outperforms original CIOU algorithm. By incorporating an auxiliary box mechanism, it heightens sensitivity of model to positional information, thereby bolstering detection accuracy. Furthermore, introduction of Alpha IOU not only elevates generalization capabilities of model but also ensures consistent performance across diverse scenarios. Given significant proportion of ship images with high IOU samples in dataset, synergy between Inner IOU and Alpha IOU substantially accelerates convergence speed of model. Notably, CIOU algorithm also factors in shape of ship, enhancing positioning accuracy and further refining detection

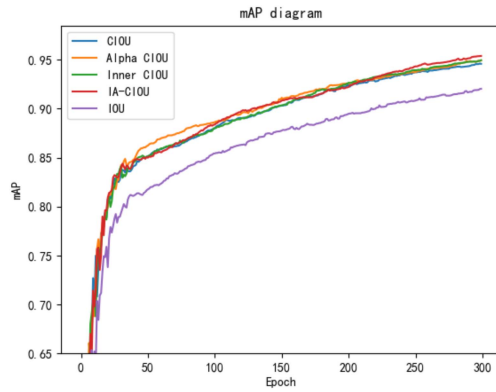


Fig. 7. Ablation experiment mAP figure.

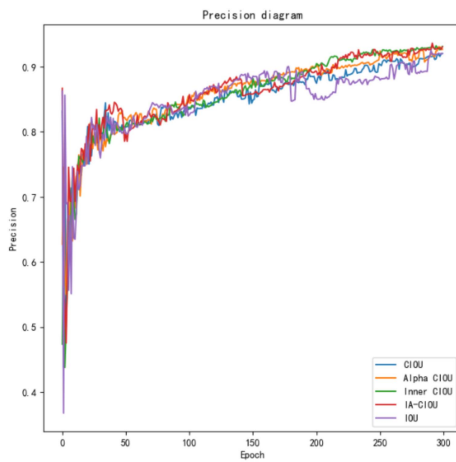


Fig. 8. Ablation experiment precision figure.

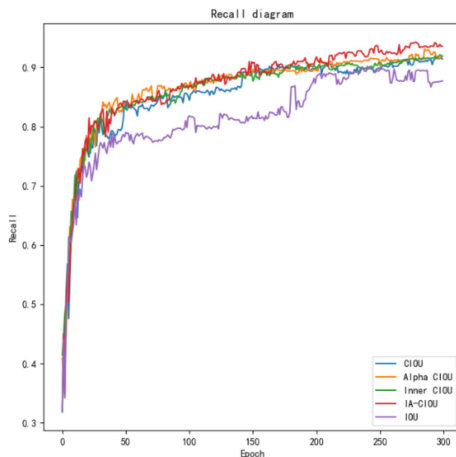


Fig. 9. Ablation experiment recall figure.

performance. Collectively, these advantages propel IA-CIOU algorithm to excel in ship detection tasks. Figs. 7–9 provide a vivid demonstration of superior performance of IA-CIOU algorithm across various performance metrics.

TABLE VI
TABLE OF CROSS EXPERIMENTS

Algorithm	R (%)	P (%)	mAP(%)
CIOU (M weight)	72.3	83.6	79.9
IA-CIOU (M weight)	72.8	84.4	81.9
CIOU (S weight)	84.6	70.0	74.9
IA-CIOU (S weight)	81.7	71.8	77.1

F. Experiments on SSDD Dataset

In the experiment, the study conducted a complex analysis of IA-CIOU. In order to provide a more comprehensive and accurate evaluation of its detection ability, a cross-validation experiment was designed. Specifically, using IA-CIOU weights trained on SAR-Ship-Dataset and MSAR-1.0 dataset, the study further evaluated their performance on a randomly selected subset of 500 ship images in SSDD dataset. Table VI summarizes this information. In order to gain a more intuitive understanding of performance enhancement achieved by IA-CIOU, this article specifically chose CIOU loss function in YOLOv5 as our benchmark, which was trained on SAR-Ship-Dataset and MSAR-1.0 dataset. By comparing these results, performance differences of IA-CIOU before and after optimization can be clearly distinguished.

M weight represents the weight obtained through training on MSAR-1.0 dataset, whereas S weight corresponds to the weight derived from training on SAR-Ship-Dataset. As evident from comparison results presented in Table VI, mAP values of M weights trained on CIOU and IA-CIOU surpass mAP values of S weights when tested on SSDD dataset. This superiority can be attributed to rich diversity of ship types present in MSAR-1.0 dataset, along with a substantial number of training samples. Furthermore, on SSDD dataset, mAP value of M weight trained using IA-CIOU is 2.0% higher than that of M weight trained on CIOU. Similarly, mAP value of S weight trained with IA-CIOU exceeds mAP value of S weight trained on CIOU by 2.2%. These observations underscore enhanced detection performance of IA-CIOU weights trained on both SAR-Ship-Dataset and MSAR-1.0 dataset, thus conclusively demonstrating superior generalization capabilities of IA-CIOU.

V. CONCLUSION

The technology of SAR ship target detection holds a crucial position in various fields, including marine monitoring, ship traffic management, and military reconnaissance. However, traditional algorithms often falter when dealing with SAR images that feature dense ship distributions and intricate backgrounds. This can result in inaccurate positioning, thereby compromising the overall effectiveness of detection. To address this challenge, this article introduces a groundbreaking SAR ship target detection algorithm aimed at significantly improving the detection performance of ship targets.

A novel loss function has been designed, utilizing the unique characteristics of ships. By seamlessly integrating it into the YOLOv5 algorithm and enhancing its key components, an enhancement in detection performance was observed. To further improve ship positioning, this article introduces IA-CIOU to ensure higher accuracy. Specifically, the introduction of a scaling factor r into Inner IOU can effectively control the generation of auxiliary boxes, thereby accelerating the convergence process of the model. Additionally, the inclusion of Alpha IOU aims to increase the generalization performance of the model. This can be achieved by carefully adjusting the parameters r and α , accelerating the regression of high IOU samples, and ultimately improving the detection performance of ships.

In experimental stage, this article thoroughly compared IA-CIOU with traditional detection models on three datasets. Specifically, ablation experiments were conducted on MSAR-1.0 dataset, carefully examining the role of each component in IA-CIOU and ultimately validating the effectiveness of method. Our detection results were presented on SAR-Ship-Dataset, indicating that IA-CIOU has a low false alarm rate and superior detection performance in both high seas and complex marine environments. In addition, this article conducted cross experiments and found that IA-CIOU weights trained on SAR-Ship-Dataset and MSAR-1.0 dataset showed enhanced detection performance on SSDD dataset. This discovery strongly corroborates remarkable generalization capabilities of IA-CIOU. Our findings indicate that our algorithm excels at precise ship positioning and achieving outstanding detection results in challenging environments, particularly for ships with high aspect ratios, dense arrangements and small sizes.

While IA-CIOU has indeed demonstrated improvement in detection performance, there remains ample opportunity for further exploration. Our future endeavors will focus on developing novel loss functions, particularly pertaining to auxiliary frame scaling factor r within IA-CIOU methodology. This article aims to meticulously investigate strategies that seamlessly integrate scaling factor r with IOU of ship, enabling a more rational and adaptive approach to auxiliary box generation. This approach will enhance adaptability of our algorithm to varying ship target sizes. By implementing these advancements, this article anticipates a substantial enhancement in detection performance, ultimately leading to more precise target detection.

ACKNOWLEDGMENTS

The authors are grateful to J. Chen et al. from Interdisciplinary Intelligent Computing Laboratory (CICG) of School of Electronic and Information Engineering, Anhui University and Y. Wang et al. for their kind assistance in data collection and labeling process. The authors appreciate the valuable comments and suggestions made by anonymous reviewers.

REFERENCES

- [1] T. Zhang, S. Quan, Z. Yang, W. Guo, Z. Zhang, and H. Gan, "A two-stage method for ship detection using PolSAR image," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5236918, doi: [10.1109/TGRS.2022.3216532](https://doi.org/10.1109/TGRS.2022.3216532).
- [2] X. Ke, X. Zhang, T. Zhang, J. Shi, and S. Wei, "Sar ship detection based on swin transformer and feature enhancement feature pyramid network," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2022, pp. 2163–2166, doi: [10.1109/IGARSS46834.2022.9883800](https://doi.org/10.1109/IGARSS46834.2022.9883800).
- [3] 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, Oct. 2021, Art. no. 4209, doi: [10.3390/rs13214209](https://doi.org/10.3390/rs13214209).
- [4] T. Zhang and X. Zhang, "High-speed ship detection in SAR images based on a grid convolutional neural network," *Remote Sens.*, vol. 11, no. 10, Jan. 2019, Art. no. 1206, doi: [10.3390/rs11101206](https://doi.org/10.3390/rs11101206).
- [5] W.-M. Boerner, "Recent advances in extra-wide-band polarimetry, interferometry and polarimetric interferometry in synthetic aperture remote sensing and its applications," *IEE Proc., Radar Sonar Navig.*, vol. 150, no. 3, pp. 113–124, 2003, doi: [10.1049/ip-rsn:20030566](https://doi.org/10.1049/ip-rsn:20030566).
- [6] J. Jiang and S. Cao, "Real-time disaster monitoring system by using SAR," *Proc. SPIE*, vol. 2313, pp. 91–95, 1994.
- [7] F. Ma, F. Zhang, Q. Yin, D. Xiang, and Y. Zhou, "Fast SAR image segmentation with deep task-specific superpixel sampling and soft graph convolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5214116, doi: [10.1109/TGRS.2021.3108585](https://doi.org/10.1109/TGRS.2021.3108585).
- [8] C. Mao, L. Huang, Y. Xiao, F. He, and Y. Liu, "Target recognition of SAR image based on CN-GAN and CNN in complex environment," *IEEE Access*, vol. 9, pp. 39608–39617, 2021, doi: [10.1109/ACCESS.2021.3064362](https://doi.org/10.1109/ACCESS.2021.3064362).
- [9] K. Zheng, Y. Dong, W. Xu, Y. Su, and P. Huang, "A method of fusing probability-form knowledge into object detection in remote sensing images," *Remote Sens.*, vol. 14, no. 23, Dec. 2022, Art. no. 6103, doi: [10.3390/rs14236103](https://doi.org/10.3390/rs14236103).
- [10] P. Huang, R. Wei, Y. Su, and W. Tan, "Swin-YOLO for concealed object detection in millimeter wave images," *Appl. Sci.*, vol. 13, no. 17, Aug. 2023, Art. no. 9793, doi: [10.3390/app13179793](https://doi.org/10.3390/app13179793).
- [11] Y. Su et al., "Enhancing concealed object detection in active millimeter wave images using wavelet transform," *Signal Process.*, vol. 216, Mar. 2024, Art. no. 109303, doi: [10.1016/j.sigpro.2023.109303](https://doi.org/10.1016/j.sigpro.2023.109303).
- [12] C. C. Wackerman, K. S. Friedman, W. G. Pichel, P. Clemente-Colón, and X. Li, "Automatic detection of ships in RADARSAT-1 SAR imagery," *Can. J. Remote Sens.*, vol. 27, no. 5, pp. 568–577, Oct. 2001, doi: [10.1080/07038992.2001.10854896](https://doi.org/10.1080/07038992.2001.10854896).
- [13] T. Li, Z. Liu, R. Xie, and L. Ran, "An improved superpixel-level CFAR detection method for ship targets in high-resolution SAR images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 1, pp. 184–194, Jan. 2018, doi: [10.1109/JSTARS.2017.2764506](https://doi.org/10.1109/JSTARS.2017.2764506).
- [14] M. Migliaccio, F. Nunziata, A. Montuori, and C. E. Brown, "Marine added-value products using RADARSAT-2 fine quad-polarization," *Can. J. Remote Sens.*, vol. 37, no. 5, pp. 443–451, Oct. 2011.
- [15] Y. Ji, H. Zhang, Z. Zhang, and M. Liu, "CNN-based encoder-decoder networks for salient object detection: A comprehensive review and recent advances," *Inf. Sci.*, vol. 546, pp. 835–857, Feb. 2021, doi: [10.1016/j.ins.2020.09.003](https://doi.org/10.1016/j.ins.2020.09.003).
- [16] L. Liang, S. Zhang, and J. Li, "Multiscale DenseNet meets with Bi-RNN for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 5401–5415, 2022, doi: [10.1109/JSTARS.2022.3187009](https://doi.org/10.1109/JSTARS.2022.3187009).
- [17] Y. Sun, Z. Wang, X. Sun, and K. Fu, "SPAN: Strong scattering point aware network for ship detection and classification in large-scale SAR imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 1188–1204, 2022, doi: [10.1109/JSTARS.2022.3142025](https://doi.org/10.1109/JSTARS.2022.3142025).
- [18] J. Wang, Z. Cui, T. Jiang, C. Cao, and Z. Cao, "Lightweight deep neural networks for ship target detection in SAR imagery," *IEEE Trans. Image Process.*, vol. 32, pp. 565–579, 2023, doi: [10.1109/TIP.2022.3231126](https://doi.org/10.1109/TIP.2022.3231126).
- [19] S. Liu, Q. Shi, and L. Zhang, "Few-shot hyperspectral image classification with unknown classes using multitask deep learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 6, pp. 5085–5102, Jun. 2021, doi: [10.1109/TGRS.2020.3018879](https://doi.org/10.1109/TGRS.2020.3018879).
- [20] 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, doi: [10.1109/CVPR.2016.91](https://doi.org/10.1109/CVPR.2016.91).
- [21] W. Liu et al., "SSD: Single shot MultiBox detector," in *Proc. Comput. Vis.*, 2016, pp. 21–37, doi: [10.1007/978-3-319-46448-0_2](https://doi.org/10.1007/978-3-319-46448-0_2).
- [22] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 2999–3007, doi: [10.1109/ICCV.2017.324](https://doi.org/10.1109/ICCV.2017.324).
- [23] 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," *Jul. 2022*, Accessed: Jan. 02, 2024. [Online]. Available: <http://arxiv.org/abs/2207.02696>

- [24] C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, "YOLOv9: Learning what you want to learn using programmable gradient information," Feb. 2024, Accessed: Mar. 16, 2024. [Online]. Available: <http://arxiv.org/abs/2402.13616>
- [25] 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, doi: [10.1109/CVPR.2014.81](https://doi.org/10.1109/CVPR.2014.81).
- [26] 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, 2017.
- [27] Z. Cai, Q. Fan, R. S. Feris, and N. Vasconcelos, "A unified multi-scale deep convolutional neural network for fast object detection," in *Proc. Comput. Vis.*, 2016, pp. 354–370, doi: [10.1007/978-3-319-46493-0_22](https://doi.org/10.1007/978-3-319-46493-0_22).
- [28] Z. Cai and N. Vasconcelos, "Cascade R-CNN: Delving into high quality object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 6154–6162, doi: [10.1109/CVPR.2018.00644](https://doi.org/10.1109/CVPR.2018.00644).
- [29] X. Nie, M. Duan, H. Ding, B. Hu, and E. K. Wong, "Attention mask R-CNN for ship detection and segmentation from remote sensing images," *IEEE Access*, vol. 8, pp. 9325–9334, 2020, doi: [10.1109/ACCESS.2020.2964540](https://doi.org/10.1109/ACCESS.2020.2964540).
- [30] M. Zha, W. Qian, W. Yang, and Y. Xu, "Multifeature transformation and fusion-based ship detection with small targets and complex backgrounds," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 4511405, doi: [10.1109/LGRS.2022.3192559](https://doi.org/10.1109/LGRS.2022.3192559).
- [31] D. Li, Q. Liang, H. Liu, Q. Liu, H. Liu, and G. Liao, "A novel multidimensional domain deep learning network for SAR SHIP DETECTION," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5203213, doi: [10.1109/TGRS.2021.3062038](https://doi.org/10.1109/TGRS.2021.3062038).
- [32] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Comput. Vis.*, 2018, pp. 3–19, doi: [10.1007/978-3-030-01234-2_1](https://doi.org/10.1007/978-3-030-01234-2_1).
- [33] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-excitation networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 8, pp. 2011–2023, Aug. 2020, doi: [10.1109/TPAMI.2019.2913372](https://doi.org/10.1109/TPAMI.2019.2913372).
- [34] J. Yue et al., "Optical remote sensing image understanding with weak supervision: Concepts, methods, and perspectives," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 2, pp. 250–269, Jun. 2022, doi: [10.1109/MGRS.2022.3161377](https://doi.org/10.1109/MGRS.2022.3161377).
- [35] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778, doi: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90).
- [36] P. Eichel and R. W. Ives, "Compression of complex-valued SAR images," *IEEE Trans. Image Process.*, vol. 8, no. 10, pp. 1483–1487, Oct. 1999, doi: [10.1109/83.791978](https://doi.org/10.1109/83.791978).
- [37] H. Zhang, C. Xu, and S. Zhang, "Inner-IoU: More effective intersection over union loss with auxiliary bounding box," Nov. 2023, Accessed: Jan. 2, 2024. [Online]. Available: <http://arxiv.org/abs/2311.02877>
- [38] J. He, S. Erfani, X. Ma, J. Bailey, Y. Chi, and X.-S. Hua, "Alpha-IoU: A family of power intersection over union losses for bounding box regression," in *Proc. Adv. Neural Inf. Process. Syst.*, 2021, vol. 34.
- [39] 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. 936–944, doi: [10.1109/CVPR.2017.106](https://doi.org/10.1109/CVPR.2017.106).
- [40] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path aggregation network for instance segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 8759–8768, doi: [10.1109/CVPR.2018.00913](https://doi.org/10.1109/CVPR.2018.00913).
- [41] Y. Xiao et al., "A review of object detection based on deep learning," *Multimedia Tools Appl.*, vol. 79, no. 33–34, pp. 23729–23791, Sep. 2020, doi: [10.1007/s11042-020-08976-6](https://doi.org/10.1007/s11042-020-08976-6).
- [42] H. Wang, F. Nie, and H. Huang, "Robust distance metric learning via simultaneous 1-norm minimization and maximization," in *Proc. 31st Int. Conf. Mach. Learn.*, 2014, pp. 3853–3861.
- [43] J. Yu, Y. Jiang, Z. Wang, Z. Cao, and T. Huang, "UnitBox: An advanced object detection network," in *Proc. ACM Multimedia Conf.*, 2016, pp. 516–520, doi: [10.1145/2964284.29672-74](https://doi.org/10.1145/2964284.29672-74).
- [44] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018.
- [45] H. Rezatofighi, 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, doi: [10.1109/CVPR.2019.00075](https://doi.org/10.1109/CVPR.2019.00075).
- [46] 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.*, 2020, vol. 34, pp. 12993–13000.
- [47] Y.-F. Zhang, W. Ren, Z. Zhang, Z. Jia, L. Wang, and T. Tan, "Focal and efficient IOU loss for accurate bounding box regression," *Neurocomputing*, vol. 506, pp. 146–157, Sep. 2022, doi: [10.1016/j.neucom.2022.07.042](https://doi.org/10.1016/j.neucom.2022.07.042).
- [48] Z. Gevorgyan, "SIoU loss: More powerful learning for bounding box regression," 2022. [Online]. Available: <http://dx.doi.org/10.48550/arXiv.2205.12740>
- [49] M. Siliang and X. Yong, "MPDIoU: A loss for efficient and accurate bounding box regression," Jul. 2023, Accessed: Jan. 3, 2024. [Online]. Available: <http://arxiv.org/abs/2307.07662>
- [50] Z. Tong, Y. Chen, Z. Xu, and R. Yu, "Wise-IoU: Bounding box regression loss with dynamic focusing mechanism," Apr. 2023, Accessed: Jan. 3, 2024. [Online]. Available: <http://arxiv.org/abs/2301.10051>
- [51] Y. Guo, S. Chen, R. Zhan, W. Wang, and J. Zhang, "LMSD-YOLO: A lightweight YOLO algorithm for multi-scale SAR ship detection," *Remote Sens.*, vol. 14, no. 19, Sep. 2022, Art. no. 4801, doi: [10.3390/rs1419-4801](https://doi.org/10.3390/rs1419-4801).
- [52] J. Yu, T. Wu, X. Zhang, and W. Zhang, "An efficient lightweight SAR ship target detection network with improved regression loss function and enhanced feature information expression," *Sensors*, vol. 22, no. 9, Apr. 2022, Art. no. 3447, doi: [10.3390/s22093447](https://doi.org/10.3390/s22093447).
- [53] Y. Zhou, H. Liu, F. Ma, Z. Pan, and F. Zhang, "A sidelobe-aware small ship detection network for synthetic aperture radar imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5205516, doi: [10.1109/TGRS.2023.3264231](https://doi.org/10.1109/TGRS.2023.3264231).
- [54] Y. Wang, C. Wang, H. Zhang, Y. Dong, and S. Wei, "A SAR dataset of ship detection for deep learning under complex backgrounds," *Remote Sens.*, vol. 11, no. 7, Mar. 2019, Art. no. 765, doi: [10.3390/rs11070765](https://doi.org/10.3390/rs11070765).
- [55] T. Zhang et al., "SAR ship detection dataset (SSDD): Official release and comprehensive data analysis," *Remote Sens.*, vol. 13, no. 18, Sep. 2021, Art. no. 3690, doi: [10.3390/rs13183690](https://doi.org/10.3390/rs13183690).
- [56] R. Xia et al., "CRTransSar: A visual transformer based on contextual joint representation learning for SAR ship detection," *Remote Sens.*, vol. 14, no. 6, Mar. 2022, Art. no. 1488, doi: [10.3390/rs14061488](https://doi.org/10.3390/rs14061488).
- [57] 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, pp. 8910–8927, 2022, doi: [10.1109/JS-TARS.2022.3213583](https://doi.org/10.1109/JS-TARS.2022.3213583).
- [58] 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, Jul. 2023, Art. no. 3771, doi: [10.3390/rs15153771](https://doi.org/10.3390/rs15153771).
- [59] S. Yang, W. An, S. Li, S. Zhang, and B. Zou, "An inshore SAR ship detection method based on ghost feature extraction and cross-scale interaction," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, 2023, Art. no. 3507305, doi: [10.1109/LGRS.2023.3307793](https://doi.org/10.1109/LGRS.2023.3307793).



Pingping Huang was born in Shandong, China, in 1978. He received the B.S. degree in electronic communication engineering from the College of Information Engineering, Shandong University of Technology, Zibo, China, in 2003, the M.S. degree in detection technology and automation equipment from the College of Information Engineering, Inner Mongolia University of Technology, Hohhot, China, in 2007, and the Ph.D. degree in communication and information engineering from the Graduate University of Chinese Academy of Sciences, Beijing, China,

in 2010.

He is with the College of Information Engineering, Inner Mongolia University of Technology. His major research interests include signal processing, polarimetric interferometry, and spaceborne synthetic aperture radar system design.



Shihao Tian was born in Henan, China, in 1999. He received the B.S. degree in communication engineering from the College of Electronic Information and Electrical Engineering, Anyang Institute of Technology, Anyang, China, in 2021. He is currently working toward the M.S. degree in electronics and communications engineering in the Inner Mongolia University of Technology, Hohhot, China.

His research interests include object detection and computer vision in remote sensing.



Yun Su (Member, IEEE) was born in Inner Mongolia, China, in 1988. He received the B.S. degree in electronic information engineering from the Harbin Institute of Technology, Harbin, China, in 2010, and the M.S. degree in optical communication and mobile communication from Beijing Jiaotong University, Beijing, China, in 2013.

He is with the College of Information Engineering, Inner Mongolia University of Technology, Hohhot, China. His research interests include signal processing and computer vision in remote sensing.



Yifan Dong was born in Hohhot, China, in 1987. He received the B.S. degree in applied physics from the Department of Modern Physics, University of Science and Technology of China, Hefei, China, in 2010, and the Ph.D. degree in particle physics and nuclear physics from the University of Chinese Academy of Sciences, Beijing, China, in 2015.

He is with the College of Information Engineering, Inner Mongolia University of Technology, Hohhot, China. His research interests include radar system development and computer vision in remote sensing.



Weixian Tan (Member, IEEE) was born in Hubei, China, in 1981. He received the Ph.D. degree in signal and information processing from the University of Chinese Academy of Sciences, Beijing, China, in 2009.

From 2009 to 2014, he was an Associate Researcher with the Science and Technology on Microwave Imaging Laboratory, Institute of Electronics, Chinese Academy of Sciences, Beijing. Since 2015, he has been a Professor with the College of Information Engineering, Inner Mongolia University

of Technology, Hohhot, China. His main research interests include airborne synthetic aperture radar (SAR)/GB-SAR/3-D SAR systems, signal processing and applications.



Wei Xu (Member, IEEE) was born in Jiangsu, China, in 1983. He received the M.S. degree in information and communication engineering from the Nanjing Research Institute of Electronics Technology, Nanjing, China, in 2008, and the Ph.D. degree in communication and information engineering from the Graduate University of Chinese Academy of Sciences, Beijing, China, in 2011.

In 2011, he joined the Department of Spaceborne Microwave Remote Sensing System, Institute of Electronics, Chinese Academy of Sciences, Beijing.

Since 2018, he has been a Professor with the College of Information Engineering, Inner Mongolia University of Technology, Hohhot, China. His research interests include spaceborne/airborne synthetic aperture radar (SAR) technology for advanced modes, SAR raw signal simulation, and SAR signal processing.