A Method for Oriented Small Ship Target Detection in SAR Based on Axial Large Kernels and Dynamic Convolution

Field of the Invention

This invention pertains to the fields of target detection and remote sensing image processing, particularly involving a synthetic aperture radar (SAR) method for detecting directional maritime ship targets based on an axial large kernel feature enhancement module Axiom, a dynamic convolution DynamicConv, and a gated dynamic upsampling module GateDySample.

This invention's method is applicable to maritime traffic monitoring, maritime supervision, maritime law enforcement, maritime rescue, national defense security, and other application scenarios that require ship target detection and identification in complex sea conditions and multi-directional scenarios.

This invention is particularly suitable for tasks of detecting oriented bounding boxes (OBB) for slender, rotating, and multi-scale ship targets in SAR images.

Background to the Invention

Synthetic aperture radar (SAR) has the capability of imaging at all times and under all weather conditions, unaffected by sunlight and clouds. It is an important imaging method for monitoring maritime targets and detecting ships. Compared with visible light images, SAR images have characteristics such as strong speckle noise, low contrast, and complex background clutter.

Traditional SAR ship detection methods are mostly based on constant false alarm rate detection (CFAR), statistical modeling, filtering, and morphology. These methods are highly sensitive to threshold setting, background modeling, and prior assumptions. They often encounter problems of high false alarm rate and high missed detection rate in strong sea clutter, near-shore scenarios, and densely clustered multi-target areas.

With the development of deep learning, target detection methods based on convolutional neural networks (such as the YOLO series, Faster R-CNN, etc.) have gradually been introduced into the SAR ship detection field. Through end-to-end learning, they can alleviate the difficulties of manual feature design and threshold adjustment to some extent.

However, the conventional horizontal-vertical axis-aligned horizontal bounding box (HBB) is difficult to precisely describe slender ships with obvious directionality, often resulting in large redundant bounding box areas and high overlap, which affects the positioning accuracy and the accuracy of subsequent tasks (such as target tracking, heading estimation, etc.).

Therefore, the academic and industrial communities have proposed the oriented bounding box (OBB) detection method, which predicts the center position, long axis length, short axis length, and rotation angle of the target, and more accurately fits the ship body. However, the existing OBB detection methods still have insufficient robustness in SAR scenarios for complex sea clutter, near-shore strong reflection, and multi-scale targets.

Specifically, on the one hand, the conventional convolution kernel receptive field is limited, insufficient for modeling the long-range structure of slender ships along the bow-to-stern direction, and it is difficult to balance local details and overall shape.

On the other hand, the convolution kernel parameters are usually static, lacking adaptability to changes in input features. When dealing with ship targets of different sizes, directions, and background conditions, the fixed convolution kernel is difficult to accommodate the features of small and large targets.

In terms of multi-scale feature fusion, existing structures such as Feature Pyramid Network (FPN) and Path Aggregation Network (PAN) mostly use fixed interpolation or fixed convolution for upsampling and downsampling, which are difficult to adaptively adjust the sampling focus according to the target and sea clutter distribution in SAR images.

In particular, in SAR maritime scenarios, the size of small ship targets is limited, and they are often submerged by strong speckle noise and shoreline reflections. Simply increasing the network depth or width is likely to lead to a sharp increase in parameters and computational complexity, which is not conducive to deployment on limited computing power airborne/shipborne platforms.

Furthermore, some improved methods stack multiple complex modules at the same scale and stage, which is prone to cause unstable training or gradient oscillation, affecting the convergence speed and final performance of the model.

In conclusion, existing technologies urgently need a SAR ship target detection method that can control the parameter quantity and computational complexity while having the capabilities of axial large kernel modeling, dynamic convolution feature adaptation, and lightweight gated upsampling, and is closely integrated with the oriented detection head.

Therefore, how to design a SAR oriented ship detection method that can maintain high detection accuracy and robustness in complex sea conditions, multi-scale, and multi-directional scenarios, while meeting engineering deployment requirements, is a technical problem that researchers in this field urgently need to solve.

Statement of Invention

The objective of this invention is to address the shortcomings of traditional SAR ship detection methods and existing deep learning OBB detection methods, and propose a SAR directional ship target detection method based on the Axiom axial large kernel module, Dynamic Convolution, and GateDySample gated dynamic upsampling.

The core technical problems to be solved by this invention include: improving the overall structural perception ability and direction modeling ability for slender ship targets, and enhancing the positioning and angle regression accuracy of directional bounding boxes.

The invention also aims to enhance the adaptive ability of feature extraction for different scales, different directions, and different background conditions, while not significantly increasing the parameter quantity and computational overhead, and balancing the detection performance of small and large targets.

The invention further strives to improve the robustness of multi-scale feature fusion stage against SAR sea clutter and shoreline interference through key sampling and gating enhancement mechanisms, making the network focus more on the ship target area.

Through the above design, this invention expects to achieve significant improvements in indicators such as mAP50, mAP50-95, Precision, Recall, and F1 Score on a typical SAR ship dataset compared to the basic YOLOv11-OBB model without introducing the above modules.

Overview of the technical solution

To achieve the above objectives, this invention proposes a SAR directional ship target detection method based on the YOLOv11-OBB framework, using the yolo11-Hybird-obb.yaml configuration file designed for SAR ship tasks, and mixing the main network and neck structures.

This invention method includes the following steps:

Step S1: Introducing the DynamicConv dynamic convolution module in the YOLOv11 main network to replace some static convolutions or C3/C2f structures to improve the adaptability and expression ability of feature extraction.

Step S2: Embedding the Axiom axial large kernel feature enhancement module on the high-level feature maps of the main network to perform axial large kernel convolution, multi-scale depth separable convolution, axial interaction attention, and local contrast enhancement on multi-scale features, to strengthen the modeling of slender ships and complex backgrounds.

Step S3: Adopting the GateDySample gated dynamic upsampling module in the neck structure to adaptively fuse multi-scale features in the high-resolution space through bilinear baseline, edge-preserving refinement, and global gating.

Step S4: On the modified backbone and neck, using the YOLOv11-OBB directional detection head to jointly predict the category, center position, length and rotation angle of the ship target, and using the loss function adapted to SAR characteristics for end-to-end training.

Step S5: Inputting the detected SAR image into the trained model, output the detection results including directional bounding boxes and confidence, and perform post-processing using rotated non-maximum suppression (Rotated NMS) to obtain the directional ship detection results at sea.

In one preferred implementation method of this invention, the specific network structure and module combination relationships of steps S1-S3 are defined and constrained by the yolo11-Hybird-obb.yaml model configuration file to ensure that the overall parameter scale and computational complexity meet the requirements for engineering deployment.

Principle of the Axiom axial large kernel module

The Axiom axial large core module proposed in this invention is based on the input feature $X \in \mathbb{R}^{C\times H\times W}$. Firstly, it obtains $tildeX = \mathrm{BN}(X)$ through batch normalization, which serves as the unified input for subsequent multi-branch feature enhancement.

The multi-scale deep separable convolution branches apply several different convolution kernel sizes and dilation rates to $\widetilde{X}$ and fuse them along the channel dimension, obtaining multi-scale response $Y^{\mathrm{ms}} = f_{\mathrm{ms}}(\widetilde{X})$, which is used to simultaneously depict the local texture and large-scale structure of slender ships.

The axial interaction branch performs average pooling along the height and width directions to obtain row vectors and column vectors, and then generates axial attention map A through lightweight (1×1) convolution and nonlinear transformation, and performs element-wise modulation on the input feature $Y^{\mathrm{ax}} = \widetilde{X} \odot A$ to enhance the long-range dependency modeling along the bow-to-stern and its vertical direction.

The local contrast enhancement branch estimates the background component through smooth convolution and constructs residual details, which are then sharpened using a convolution initialized with $Laplacian$ thinking to obtain contrast-enhanced feature $Y^{\mathrm{ct}} = f_{\mathrm{ct}}(\widetilde{X})$, thereby highlighting the ship edges and contours and suppressing flat backgrounds and some strong reflection noise.

The three branches output Y^{ms}, Y^{ax} and Y^{ct} are weighted fused by a set of learnable weights, with $Softmax$ normalization of the weights α, β, γ satisfying $\alpha+\beta+\gamma=1$. Then, $Y^{\mathrm{fuse}}=\alpha Y^{\mathrm{ms}}+\beta Y^{\mathrm{ax}}+\gamma Y^{\mathrm{ct}}$ is obtained.

The fused feature Y^{fuse} is further input into the channel feedforward network (FFN) composed of (1×1) convolutions for nonlinear transformation in the channel dimension, and finally, through residual connections, the module output $Y=X+\mathrm{FFN}(Y^{\mathrm{fuse}})$ is obtained, which retains the original backbone features while superimposing the enhanced information of Axiom, ensuring gradient stability and training convergence.

In this invention method, the Axiom module is used to replace or enhance traditional C3k2/C2f modules at multiple scales (such as mid-level feature maps) in the backbone network. Through different preset combinations (Tiny, Heavy, Lite, Medium), the network can achieve stronger slender target perception and direction modeling capabilities while keeping the overall parameter quantity and computational cost under control.

In summary, Axiom provides a feature representation basis for the SAR directed ship detection method of this invention by adopting a compact structure of "multi-scale deep separable convolution+axial interaction attention+local contrast enhancement+residual feedforward", which takes into account long-range structure, local edges, and complex sea clutter suppression.

Overview of DynamicConv and GateDySample modules

This invention uses the DynamicConv dynamic convolution module as one of the basic convolution units in the backbone network to enhance the adaptability of feature extraction to the input content. Through global average pooling, a channel description vector is obtained, and on this basis, the routing weights of multiple expert convolution kernels are generated to form an equivalent convolution kernel related to the input, achieving adaptive modeling of different scales, directions, and texture patterns.

Compared with Axiom, DynamicConv focuses more on the deformable expression of local and medium-scale structures, such as dynamically re-weighting feature channels during downsampling and feature transformation stages, thereby improving the network's representation ability without significantly increasing the parameter quantity.

This invention designs the GateDySample gated dynamic upsampling module in the upsample path of the neck structure. This module first performs bilinear upsampling on the low-resolution features as a stable baseline output, and then realizes edge refinement in the high-resolution space through depth separable convolution initialized with Laplacian thinking.

GateDySample generates scalar gating coefficients using the global average pooling vector of the input features, scales and superimposes the "difference between the refined result and the bilinear result", so that when the gating value is close to 0, it degenerates to pure bilinear upsampling, and when the gating value is close to 1, it fully utilizes the edge refinement result.

Through this design, GateDySample achieves lightweight enhancement of the ship edges and contours in the high-resolution feature map with extremely low parameters and computational complexity, which is beneficial for improving the detection performance of SAR elongated ships, multi-scale targets, and complex sea clutter environments for the subsequent directed detection head.

Axiom, DynamicConv, and GateDySample do not conflict with each other in the network structure: Axiom mainly acts on the mid-high-level features in the backbone, DynamicConv runs through several stages of the backbone to achieve dynamic convolution expression, and GateDySample focuses on the upsampling stage in the neck. The three form a complementary overall solution.

Through the collaborative design of the above core modules, this invention achieves efficient modeling of SAR elongated ships, multi-scale targets, and complex sea clutter environments while controlling the overall parameters and computational complexity, providing more discriminative feature representations for subsequent directed detection heads.

Brief Description of the Drawings

Drawing 1 is the structure of the network model of this method;

Drawing 2 is the Axiom module;

Drawing 3 is the MultiScaleDepthwise module;

Drawing 4 is the AxialInteraction module;

Drawing 5 is the LocalContrastEnhance module;

Drawing 6 is GateDySample upsampling;

Drawing 7 is the DynamicConv module;

Drawing 8 is the SAR ship image;

Drawing 9 is the experimental comparison of the ablation experiments;

Drawing 10 is the training results of this method on the SSDD+ dataset;

Drawing 11 is the comparison of detection effects between this method and several common YOLO models.

Detailed Description

To make the purpose, technical solution and beneficial effects of this invention clearer and more understandable, the following provides a detailed explanation of the method of this invention by combining specific implementation methods.

Data preprocessing and annotation

In the implementation methods, the input is a single-channel or multi-channel SAR maritime scene image. Firstly, the original SAR image is gray-scale normalized, and the pixel values are mapped to a preset range to reduce the statistical offset caused by differences in imaging conditions.

According to the need, the SAR image is subjected to speckle noise suppression and contrast enhancement processing to improve the separability between the ship target and the background clutter, and enhance the visibility of the target edges and structural textures.

The preprocessed SAR image is cropped or scaled to a unified input size of 640×640.

In the annotation format, the OBB four-corner coordinate format is used to describe the ship target. The real box precisely locates the position and deflection of the target by recording the coordinate sequence of the four vertices of the bounding box $(x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4)$, thereby directly participating in the loss calculation of the directed detection network.

While keeping the test set unchanged, the amplitude and probability of the enhancement strategy can be appropriately adjusted according to the specific characteristics of the SAR dataset to avoid introducing too many pseudo samples inconsistent with the real scene.

In this implementation method, the SAR maritime target directed ship dataset SSDD+ and HRSID-R are selected, and both are organized and merged into a single dataset according to the unified annotation format, and are randomly divided into training set, validation set and test set in a ratio of approximately 7:2:1 to ensure the objectivity and statistical significance of the evaluation results.

For the differences in imaging parameters, resolution and sea conditions between the two data sources SSDD+ and HRSID-R, in the preprocessing stage, a unified normalization or standardization strategy is used to align them, thereby reducing the domain differences between different data sources and improving the generalization ability of this invention method in multi-source SAR maritime scenes.

Network structure configuration and module insertion

In the implementation methods, the YOLOv11-OBB network backbone structure is constructed based on the yolo11-Hybird-obb. yaml configuration file. DynamicConv and Axiom modules are introduced in multiple stages of the main body, and the GateDySample module is introduced in the neck structure for feature upsampling and fusion.

In the high-resolution stage close to the input end (such as P2 or P3), AxiomYOLOBlockTiny or AxiomYOLOBlockHeavy with a relatively smaller force can be appropriately introduced to ensure the sensitivity to small ships and local textures, while controlling the computational cost.

In the middle and high-level stages (such as P4, P5), AxiomYOLOBlockLite or AxiomYOLOBlockMedium are applied to obtain a larger effective receptive field and axial structure modeling ability, adapting to medium and large-sized ship targets and complex backgrounds.

At the positions of some convolutional layers or C3/C2f structures in the main network, the DynamicConv dynamic convolution module is introduced, using multiple expert convolution kernels and routing weight mechanisms to enable the network to adaptively select more suitable convolution kernel combinations based on the input features.

To achieve a balance between parameter quantity and computational cost, some standard convolution or lightweight structures can be retained in the shallow stage, and DynamicConv is mainly used in the middle and deep layers to enhance the network's expression ability at high semantic levels.

In the upsampling path of the neck structure, the GateDySample module replaces some traditional upsampling modules (such as only using bilinear interpolation plus convolution), achieving a lightweight upsampling method that combines bilinear baseline, edge-preserving refinement, and global gating.

The GateDySample module first performs bilinear interpolation upsampling on the low-resolution features, then uses the depth separable convolution initialized with $Laplacian$ thinking in the high-resolution space to refine the features, and generates a scalar gate through the gating branch composed of global average pooling and fully connected layers, weighting the refined residuals.

By reasonably setting the insertion positions and numbers of Axiom, DynamicConv, and GateDySample in the network, this implementation improves the network's modeling ability for SAR slender ships, multi-scale targets, and complex sea clutter while maintaining a compact overall structure.

In the implementation method, based on the specific computing power of the hardware platform, the depth coefficients, width coefficients, and repetition times of each module in yolo11-Hybird-obb.yaml can be adjusted to trim the network size, meeting real-time or low-power engineering constraints.

The network structure of the invention method has good scalability. Depending on different task requirements and data scale, other auxiliary modules can be added, deleted, or replaced while keeping the core modules Axiom, DynamicConv, and GateDySample unchanged.

Training and inference process

In the training stage, this implementation conducts experiments on a hardware platform based on Ubuntu 22.04 operating system and NVIDIA GeForce RTX 3090 graphics card (with 24GB of video memory). The deep learning environment includes Python 3.10.19, PyTorch 2.9.1, and CUDA 12.6. The target detection framework uses the Ultralytics YOLO version 8.3.174.

The specific training configuration is implemented through the Ultralytics YOLO training interface. In the training script, an OBB detection model is constructed based on the yolo11-Hybird-obb.yaml model configuration file, and the merged SSDD+ and HRSID-R datasets are used as training data.

During the training process, the training epochs epochs=300, batch size batch=32, input size imgsz=640, optimizer is momentum-based stochastic gradient descent (SGD), momentum coefficient momentum=0.937, weight decay weight_decay=0.0005, learning rate is 0.01, and the AMP mixed precision and early stopping strategy with patience=50 are enabled as hyperparameters.

Under the above configuration, by reasonably setting the data loading process number workers=8 and an appropriate learning rate scheduling strategy, the mAP50, mAP50-95, Precision, Recall, F1 Score, etc. indicators are continuously monitored during the training and validation processes, and the model weight with the best performance is selected, and the training strategy is appropriately fine-tuned.

In the inference stage, the SAR image to be detected is preprocessed in the same way as in the training stage and input into the trained network model. It passes through the DynamicConv backbone, the Axiom module, and the GateDySample neck structure for feature extraction and fusion, and finally outputs the ship target prediction result by the directed detection head.

Performing directed non-maximum suppression on all the prediction boxes to filter out the redundant boxes with low confidence and high overlap, and obtaining the final SAR maritime direction-based ship detection results, which can be applied to maritime traffic monitoring, maritime supervision, situational awareness and other scenarios.

During actual deployment, according to the platform computing power and real-time requirements, the input resolution, network scale and post-processing threshold can be adjusted to achieve a balance between detection accuracy and inference speed.

Performance evaluation and assessment indicators

To objectively evaluate the performance of this invention method in the SAR ship detection task, common indicators such as accuracy (Precision), recall (Recall), F1 score (F1 Score), mAP@50 (mAP50) and mAP@50:95 (mAP50-95) are adopted. For all the predicted boxes and true boxes in the test set, after matching according to the given IoU threshold, the true positive number TP, false positive number FP and false negative number FN are obtained.

Accuracy (Precision) is defined as:

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

Among the targets detected as ships, the proportion of those that are actually ships reflects the false detection situation.

The recall rate (Recall) is defined as:

$$\mathrm{Recall} = \frac{TP}{TP + FN}$$

It represents the proportion of actual ship targets that have been successfully detected, reflecting the situation of missed detections.

The F1 score (F1 Score) is the harmonic mean of the accuracy rate and the recall rate:

$$\mathrm{F1} = \frac{2 \cdot \mathrm{Precision} \cdot \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}$$

It is used to comprehensively measure the degree of balance between "few false detections" and "few missed detections" in the detection results.

Given an IoU threshold t, for example $t = 0.5$, a set of (Recall_k, $\mathrm{Precision}_k$) corresponding to different confidence thresholds can be sorted, and interpolation is performed on the recall dimension to obtain the precision-recall curve $\mathrm{Precision}_t(r)$, and the area under this curve is the average precision (AP) at this threshold. In practical implementation, a discrete summation approximation is usually adopted.

mAP50 refers to the average detection performance metric obtained by taking the average of the average precision for all categories at IoU threshold $t = 0.5$:

$\mathrm{mAP50} = \frac{1}{C}\sum_{c=1}^{C}\mathrm{AP}_{t=0.5}^{(c)}$, wherein, C is the number of categories, and in single-class ship detection, this can be regarded as the AP value of this category.

mAP50-95 refers to calculating AP separately on multiple IoU threshold sets $T = \{0.50, 0.55, \ldots, 0.95\}$, and then taking the average of all categories and all thresholds:

$$\mathrm{mAP50} \setminus \mathrm{mbox} - 95 = \frac{1}{C \cdot |T|}\sum_{c=1}^{C}\sum_{t \in T}\mathrm{AP}_{t}^{(c)}$$

This indicator measures the comprehensive detection performance of the method under different overlap requirements more strictly.

In the specific experiments, the invention method can calculate Precision, Recall, F1, mAP50 and mAP50-95 based on the test set of the SAR vessel dataset (such as SSDD+, HRSID-R, etc.), and compare it with the basic YOLOv11-OBB model without using Axiom, DynamicConv, and GateDySample.

The experimental results show that, under the same training data and training strategy, the invention method has significant improvements in indicators such as mAP50, mAP50-95, and F1 scores, verifying the effectiveness of the collaborative design of Axiom, DynamicConv, and GateDySample in the SAR directed vessel detection task.

Experimental results and analysis

Table 1 Module ablation experiment of SSDD+

YOLOv11n-OBB	Axiom	DynamicConv	GateDySample	P	R	mAP50	mAP50-95	F1	Parameters	GFLOPS
√				94.86	91.70	96.77	75.47	93.25	2.66M	6.6
√	√			96.70	96.61	98.95	79.11	96.65	2.65M	8.1
√		√		96.86	96.17	98.66	78.73	96.52	2.42M	5.9
√			√	94.42	96.89	98.95	78.94	95.63	2.67M	6.7
√	√	√		96.00	96.65	98.91	79.40	96.32	2.41M	7.4
√	√		√	95.31	96.15	98.69	79.25	95.73	2.65M	8.1
√		√	√	**96.91**	97.51	98.95	78.68	97.21	2.42M	**6.0**
√	√	√	√	96.88	**98.54**	**99.19**	**79.77**	**97.70**	**2.41M**	7.4

From the table, it can be seen that the Axiom module increases P, R, mAP50, and mAP50-95 by 1.84%, 4.91%, 2.18%, and 3.64% respectively. The DynamicConv module increases P, R, mAP50, and mAP50-95 by 2.0%, 4.47%, 1.89%, and 3.26% respectively. Notably, the DynamicConv module improves the accuracy while reducing the parameter size from 2.66M to 2.42M and the GFLOPs from 6.6 to 5.9, effectively reducing the model complexity and computational cost. The GateDySample module shows significant improvements in R and mAP, increasing by 5.19% and 3.47% respectively.

Meanwhile, the multi-module ablation experiment results show that the final model integrating the above modules (the last row in the table) has the most significant improvement in accuracy P, R, mAP50, and mAP50-95, which is 2.02%, 6.84%, 2.42%, and 4.30% higher than the baseline model (Baseline), respectively. The F1 score also increases from 93.25 to 97.70. The experiment proves that the collaborative effect of Axiom, DynamicConv, and GateDySample can significantly improve the detection accuracy (especially the recall rate R), while compressing the parameter size by approximately 9.4%, achieving a balance between model lightweighting and high performance.

Contrasting case

Table 2 compares various object detection models on the SSDD+ dataset

Method	P	R	mAP50	F1	Parameters	GFLOPS
R-FasterRCNN	92.83	91.36	90.15	92.09	41.13M	198.4
O-RCNN	88.42	89.28	87.65	88.85	41.13M	198.54
ROI-TRANSFORMER	83.55	87.73	82.91	85.59	41.4M	220.0
R-FCOS	84.41	85.64	83.72	85.02	32.2M	200.0
R3Det	86.56	88.19	87.61	87.37	41.67M	330.71
R-RetinaNet	94.93	92.54	94.86	93.72	33.3M	156.0
R-LRBPNet	92.83	91.36	90.15	92.09	3.2M	8.5
YOLOv8n-OBB	92.94	94.74	96.50	93.83	3.08M	8.4
YOLOv11n-OBB	94.86	91.70	96.77	93.25	2.66M	**6.6**
YOLOv12n-OBB	92.29	91.51	96.57	92.29	2.64M	6.7
Ours	**96.88**	**98.54**	**99.19**	**97.70**	**2.41M**	7.4

As shown in Table 2, the experimental results on the SSDD+ dataset indicate that our model has achieved the best performance in all core performance indicators except for GFLOPS.

The detection accuracy P, recall rate R, mAP50, and F1 score of this model reached 96.88%, 98.54%, 99.19%, and 97.70% respectively. Compared with classic models such as R-FasterRCNN, O-RCNN, R-FCOS, and R-RetinaNet, our model demonstrates a significant advantage in detection accuracy.

Compared with the lightweight design of YOLOv12n-OBB, our method further compressed the parameter quantity by approximately 8.7% (from 2.64M to 2.41M), while the mAP50 indicator increased by 2.62%. The parameters and GFLOPs indicators of this model are 2.41M and 7.4 respectively, which are far lower than those of large classic models such as R-FasterRCNN (41.13M and 198.4). Although the GFLOPs of the model (7.4) is slightly higher than YOLOv11n-OBB (6.6), it has improved the recall rate (R) by 6.84% and the mAP50 by 2.42%, achieving the best balance between accuracy and efficiency.

Table 3 Comparison of various target detection models on the HRSID-R dataset

Method	P	R	mAP50	F1	Parameters	GFLOPS
R-FasterRCNN	80.62	81.04	77.87	80.83	41.13M	198.4
O-RCNN	85.35	84.61	80.2	84.98	41.13M	198.54
ROI-TRANSFORMER	84.19	82.48	78.76	83.33	41.4M	220.0
R-FCOS	78.77	74.91	73.15	76.79	32.2M	200.0
R3Det	79.86	75.29	77.45	77.51	41.67M	330.71
R-RetinaNet	83.72	81.45	80.23	82.57	33.3M	156.0
R-LRBPNet	91.35	87.59	88.74	89.43	3.2M	8.5
YOLOv8n-OBB	92.96	85.91	94.10	89.30	3.08M	8.4
YOLOv11n-OBB	92.34	86.45	93.96	89.30	2.66M	**6.6**
YOLOv12n-OBB	91.21	86.43	93.71	88.76	2.64M	6.7
Ours	**95.15**	**90.82**	**96.95**	**92.93**	**2.41M**	7.4

Furthermore, we verified the robustness of this method on the more challenging HRSID-R dataset. As shown in Table 3, in the face of the challenge of complex nearshore backgrounds and the coexistence of multi-scale targets, the recall rate and high precision indicators (mAP50-95) of the benchmark model YOLOv11n-OBB easily encountered bottlenecks. The experimental results demonstrated that this method exhibited outstanding generalization ability on the HRSID-R dataset. Thanks to the fine reconstruction of the features of small targets by GateDySample, the model achieved a qualitative leap in the Recall indicator, from 86.45% to 90.82%, effectively solving the problem of missed detections in complex port scenarios.

Meanwhile, the anti-interference ability provided by DynamicConv enabled mAP50 to reach a new high of 96.95%. Compared with YOLOv11n-OBB and YOLOv12n-OBB, it increased by 2.99% and 3.24% respectively. It is worth noting that in the high standard positioning indicator mAP50-95, this method achieved an excellent result of 92.93%, proving that the model can still achieve high-precision directional bounding box regression in complex environments.

This invention provides an implementation method for detecting oriented small targets of ships in synthetic aperture radar images based on deep learning. There are many methods and approaches to specifically implement this technical solution. The above-mentioned are only the preferred implementation methods of this invention. It should be pointed out that for ordinary technicians in this technical field, without departing from the principle of this invention, they can make some improvements and refinements. These improvements and refinements should also be regarded as within the protection scope of this invention. In this implementation method, the various components that are not explicitly described can be realized using existing technologies.