electronics

MDPI

Article

Integrated Anti-Aliasing and Fully Shared Convolution for Small-Ship Detection in Synthetic Aperture Radar (SAR) Images

Manman He [1,†], Junya Liu [2,*,†], Zhen Yang [1,*,†] and Zhijian Yin [1,†]

[1] The School of Information Engineering, Jiangxi Science and Technology Normal University, Road of Xuefu, Nanchang 330013, China; 2022010264@jxstnu.edu.cn (M.H.); 1020100837@jxstnu.edu.cn (Z.Y.)
[2] The School of Foreign Languages, Jiangxi Science and Technology Normal University, Road of Xuefu, Nanchang 330013, China
* Correspondence: 1020150923@jxstnu.edu.cn (J.L.); 1020170701@jxstnu.edu.cn (Z.Y.)
† These authors contributed equally to this work.

Abstract: Synthetic Aperture Radar (SAR) imaging plays a vital role in maritime surveillance, yet the detection of small vessels poses a significant challenge when employing conventional Constant False Alarm Rate (CFAR) techniques, primarily due to the limitations in resolution and the presence of clutter. Deep learning (DL) offers a promising alternative, yet it still struggles with identifying small targets in complex SAR backgrounds because of feature ambiguity and noise. To address these challenges, our team has developed the AFSC network, which combines anti-aliasing techniques with fully shared convolutional layers to improve the detection of small targets in SAR imagery. The network is composed of three key components: the Backbone Feature Extraction Module (BFEM) for initial feature extraction, the Neck Feature Fusion Module (NFFM) for consolidating features, and the Head Detection Module (HDM) for final object detection. The BFEM serves as the principal feature extraction technique, with a primary emphasis on extracting features of small targets, The NFFM integrates an anti-aliasing element and is designed to accentuate the feature details of diminutive objects throughout the fusion procedure, HDM is the detection head module and adopts a new fully shared convolution strategy to make the model more lightweight. Our approach has shown better performance in terms of speed and accuracy for detecting small targets in SAR imagery, surpassing other leading methods on the SSDD dataset. It attained a mean Average Precision (AP) of 69.3% and a specific AP for small targets (AP_S) of 66.5%. Furthermore, the network's robustness was confirmed using the HRSID dataset.

Keywords: SAR; ship detection; CFAR; DL

check for updates

Citation: He, M.; Liu, J.; Yang, Z.; Yin, Z. Integrated Anti-Aliasing and Fully Shared Convolution for Small-Ship Detection in Synthetic Aperture Radar (SAR) Images. *Electronics* **2024**, *13*, 4540. https://doi.org/10.3390/electronics13224540

Academic Editor: Silvia Liberata Ullo

Received: 25 October 2024
Revised: 13 November 2024
Accepted: 15 November 2024
Published: 19 November 2024

1. Introduction

SAR sensors emit radar pulses and, by analyzing the reflected echoes, can generate high-resolution images that remain sharp regardless of weather and light conditions. As the maritime sector and navigation technologies advance swiftly, the identification of ship targets in SAR imagery is gaining more significance for managing maritime traffic and conducting sea rescue operations. Nonetheless, the identification of small vessels is complicated by background noise and false echoes present in SAR images. In recent years, a variety of ship detection techniques in SAR imagery have emerged. These approaches are generally categorized into two principal groups depending on the type of features they employ: classical methods and those that incorporate deep learning techniques. Classical techniques often rely on handcrafted features, while deep learning approaches automatically learn features from large datasets.

The Constant False Alarm Rate (CFAR) algorithm, as described by Tao [1] in his research, is a prevalent technique for detecting ship targets within the realm of standard Synthetic Aperture Radar (SAR) imagery. A detection is deemed successful when the echo signal from a potential target surpasses an established threshold. However, the complex and noisy background of SAR images can hinder the accurate detection of smaller vessels. CFAR has problems with accuracy and real-time performance. The difference between the actual noise environment and the statistical model can reduce the detection accuracy and robustness of CFAR technology. Moreover, the thresholds must be customized to meet the particular demands and features of various applications and targets, which restricts the approach's adaptability and broad applicability.

In contradistinction to the CFAR methodology, the advent of deep learning (DL) [2] has heralded a paradigm shift, endowing the field with advanced mechanisms for feature extraction and exploitation. Consequently, techniques that leverage deep learning (DL) for identifying ship targets in SAR imagery have gained significant traction in scholarly studies. In this field, Convolutional Neural Networks (CNNs) have become a foundational and extensively applied approach. The landscape of detector architectures is bifurcated into two principal categories: those predicated on the utilization of anchor-based mechanisms and those that eschew the reliance on anchors [3], opting for an anchor-free [3] approach. Utilizing a predefined set of anchor boxes, anchor-based object detection algorithms reformulate the object localization challenge into a binary classification task—assessing the presence of an object within each anchor—and a regression problem—estimating the target's class and bounding box coordinates. The primary advantage of this technique is its capability to adapt the dimensions and forms of anchor boxes to accommodate targets of varying sizes, thereby enhancing detection accuracy. However, the effectiveness of anchor-based detection methods hinges on the precise calibration of anchor parameters to align with the particular features of the dataset in use. The quantity and proportional distribution of anchor boxes also exert a significant influence on the detection outcomes, potentially amplifying the algorithm's complexity and the intricacy of its implementation.

In 2022, Gao et al. [4] introduced a RetinaNet-derived algorithm specifically designed for ship detection within compact polarimetric (CP) Synthetic Aperture Radar (SAR) imagery. Utilizing power-entropy (PE) decomposition theory, the method extracts polarization features and employs RetinaNet for both the training and prediction phases. This method surpasses the CFAR algorithm by reducing false alarms while preserving the accuracy of detection. The research highlights the capability of neural networks to enhance ship detection in circular polarization (CP) SAR imagery, outperforming conventional detection techniques. Li, Xiao, and Shi [5] have developed an advanced YOLOv4 model specifically designed for detecting ships in SAR images. This model incorporates a decoupled head and a coordinate attention mechanism to tackle the issues of high model complexity and the difficulty in detecting small, densely packed targets. It features a lightweight backbone network and employs the K-Means algorithm for optimizing anchor boxes. Yue et al. [6] have introduced an innovative two-stage ship detection network designed for SAR images, tackling issues like dense ship distribution and cluttered backgrounds. The network is equipped with a Generating-Anchor Module (GAM) that produces high-quality anchors closely tailored to the features of ships.

In 2023, to tackle the challenges of high intra-class similarity and complex background contexts in object detection within remote sensing imagery, Lin and colleagues [7] presented YOLO-DA, an enhanced YOLO-based detection model. This model enhances detection capabilities by incorporating a decoupled attention head to boost performance. Liu and team [8] have developed an enhanced YOLOv5s model, tailored for detecting small targets in optical remote sensing imagery. The significant improvements include adding a detection layer that works with shallow feature maps, implementing skip connections to improve feature fusion, creating a new Hybrid Spatial Pyramid Pooling (HSPP) module for combining local and global features, incorporating a Coordinate Attention (CA) mechanism to boost feature representation, and utilizing an Efficient Intersection over Union (EIOU) loss function for bounding box regression. Additionally, the model employs the K-means++ algorithm to optimize anchor points. Tackling the difficulty of detecting ships in SAR images, particularly in intricate coastal settings, Qin et al. [9] introduced S2LSDNet. This innovative method incorporates semi-soft labels obtained through self-distillation and employs an Angle-related and Balanced Intersection-over-Union (ArBIoU) loss function to boost detection accuracy, surpassing current methods. Gao and colleagues [10] introduced SFF–YOLOX, a cutting-edge anchor-free model specifically crafted for ship detection within SAR imagery, effectively addressing challenges related to feature integration and the imbalance of training samples. This model incorporates a Salient Region Extraction module and a Salient Feature Fusion mechanism, which bolster the discrimination of targets across different scales. Sun and associates [11] devised an advanced algorithm for identifying ships in SAR imagery, which integrates multilevel superpixel segmentation with fuzzy fusion techniques. This approach entails the segmentation of SAR images into different levels with varying superpixel dimensions, the extraction of pertinent features, the generation of soft detection outcomes via the fuzzy C-means clustering algorithm, and the synthesis of these outcomes to achieve the final detection. This method significantly reduces the dependency on the choice of superpixel size and shows resilience to variations in the dimensions of ship targets, leading to enhanced detection accuracy and a decrease in false alarms within SAR imagery-based ship detection. Hao and Zhang [12] developed a more efficient CNN tailored for ship detection. This network is built upon the YOLOX algorithm and features a newly designed lightweight block called MobileNetV3S, as well as a multiscale feature extraction module known as the F-ESP Block. Additionally, it employs a convolutional block attentional module and an absolute Intersection over Union (IoU) loss function to enhance performance. Zhang and team [13] unveiled MLBR-YOLOX, a sophisticated neural network optimized for ship detection within SAR imagery. The network employs a multi-tiered approach for background suppression, incorporating a dedicated Spatial Patch-based Detector (SSPD) and an advanced Spatial Feature-based Detector (DSFD) to initially identify ships and eliminate non-critical background details, thereby lessening the computational burden. By integrating with the YOLOX framework, MLBR-YOLOX not only matches YOLOX in terms of detection precision but also boasts a markedly reduced computational demand, proving to be highly effective in the context of SAR-based ship detection missions.

In 2024, to overcome the challenges of identifying small objects in satellite imagery—such as the loss of details for these objects during the fusion of features and the heightened sensitivity of loss functions to them, Lei and Liu [14] introduced SOLO-Net, an innovative detection framework. The framework has been tested on the RSOD and DIOR datasets, with the results showing considerable enhancements in detection precision. SOLO-Net has surpassed current deep-learning-based detection techniques, proving its efficacy in addressing the challenges of identifying small objects within intricate remote-sensing images. Huang and colleagues [15] presented IA-CIOU, an approach that integrates Inner IOU and Alpha IOU to refine detection accuracy and model convergence, especially for small ships in cluttered backgrounds. Zhou and colleagues [16] developed EGTB-Net, an efficient and compact network specifically designed for ship detection within Synthetic Aperture Radar (SAR) imagery. The network leverages the power of transformers coupled with feature amplification methods. It features an innovative Ghost-ECA module for effective feature extraction and a transformer block to capture long-range dependencies and introduces a novel SIoU loss function to enhance localization accuracy. Yang et al. [17] introduced LT-Net, a deep learning model optimized for PolSAR ship detection that integrates radar polarization theory to reduce computational demands. They also present the FPSD dataset, comprising 853 PolSAR images with ship targets, to foster research in this area. Jia and co-workers [18] introduced a swift progressive ship detection algorithm tailored for extensive, full-scene SAR imagery, merging conventional approaches with deep learning (DL) methodologies. This method features a Candidate Region Retrieval Module (CRM) that swiftly pinpoints potential ship sites, an Oriented Ship Detector (OSD) that ensures precise detection, and a False Alarm Discrimination Module (FDM) intended to minimize false positives. Chen and co-workers [19] developed LiteSAR-Net, an efficient SAR ship detection model aimed at boosting the detection performance in SAR images. Zhou and team [20] introduced BFEA, a deep learning architecture crafted for identifying ships of varying sizes within SAR images. This model integrates WEF-Net, leveraging attention mechanisms to harmonize multi-scale feature information, along with a refined loss function that improves the identification of smaller targets.

Additionally, the model incorporated a linear polynomial loss function to bolster robust representation learning. Li and colleagues' [21] study presents AID-YOLO, an advanced YOLOv8s-based detector designed for small target detection in aerial imagery. It incorporates a RepNCSPELAN4 module and a C2FCBAM module to boost feature extraction and target focus, along with a NWD-CIoU Loss function to tackle small object sensitivity.

Despite recent advances in CNN-based [22] SAR ship detection, its practical applications still face many challenges, such as more complex backgrounds (e.g., port facilities, strong sea clutter, and unstable sea conditions) and more diverse ships, which in turn can easily make themselves yield unsatisfactory detection performance. To address these limitations, we introduce AFSC, an innovative neural network that integrates anti-aliasing techniques with fully shared convolutional layers, enhancing the detection of small ship targets in SAR imagery. The network is structured with the Backbone Feature Extraction Module (BFEM) at its core for extracting feature information, emphasizing the capture of critical details. Anti-aliasing techniques are integrated to refine the depiction of small targets, thus improving the model's efficacy. The Neck Feature Fusion Module (NFFM) functions as the feature integrator, and the Head Detection Module (HDM) serves as the decision-making unit of the detection model, responsible for producing the final detection results. We present a straightforward yet potent detector, AFSC, as depicted in Figure 1. In summary, this paper contributes in three key areas.

1. Backbone Feature Extraction Module (BFEM): The BFEM is the core component of the network designed for extracting detailed features from small targets in SAR images. It integrates sophisticated methods to mitigate the impact of intricate backdrops and bolsters the network's proficiency in detecting and analyzing the features of small targets.
2. Neck Feature Fusion Module (NFFM): The Neck Feature Fusion Module (NFFM) plays a pivotal role in the network's process of integrating features. It integrates the WaveletPool module [23] to balance the location and semantic information, ensuring that the network can effectively merge features, especially those related to small targets within SAR imagery, which is crucial for improving the model's robustness and accuracy.
3. Head Detection Module (HDM): The HDM is a novel detection module that significantly reduces the model's parameter count by employing shared convolutions [21]. This module is crafted to elevate the accuracy and resilience of the detection procedure, rendering it more dependable across the spectrum of complex SAR imaging conditions.

The subsequent sections of this paper are organized as follows: Section 2 details the methodology, Section 3 reports on the experiments and offers a comparison with existing research, and Section 4 delivers an analysis of the experimental results. The paper concludes with Section 5, which presents final thoughts and suggests potential directions for future research.

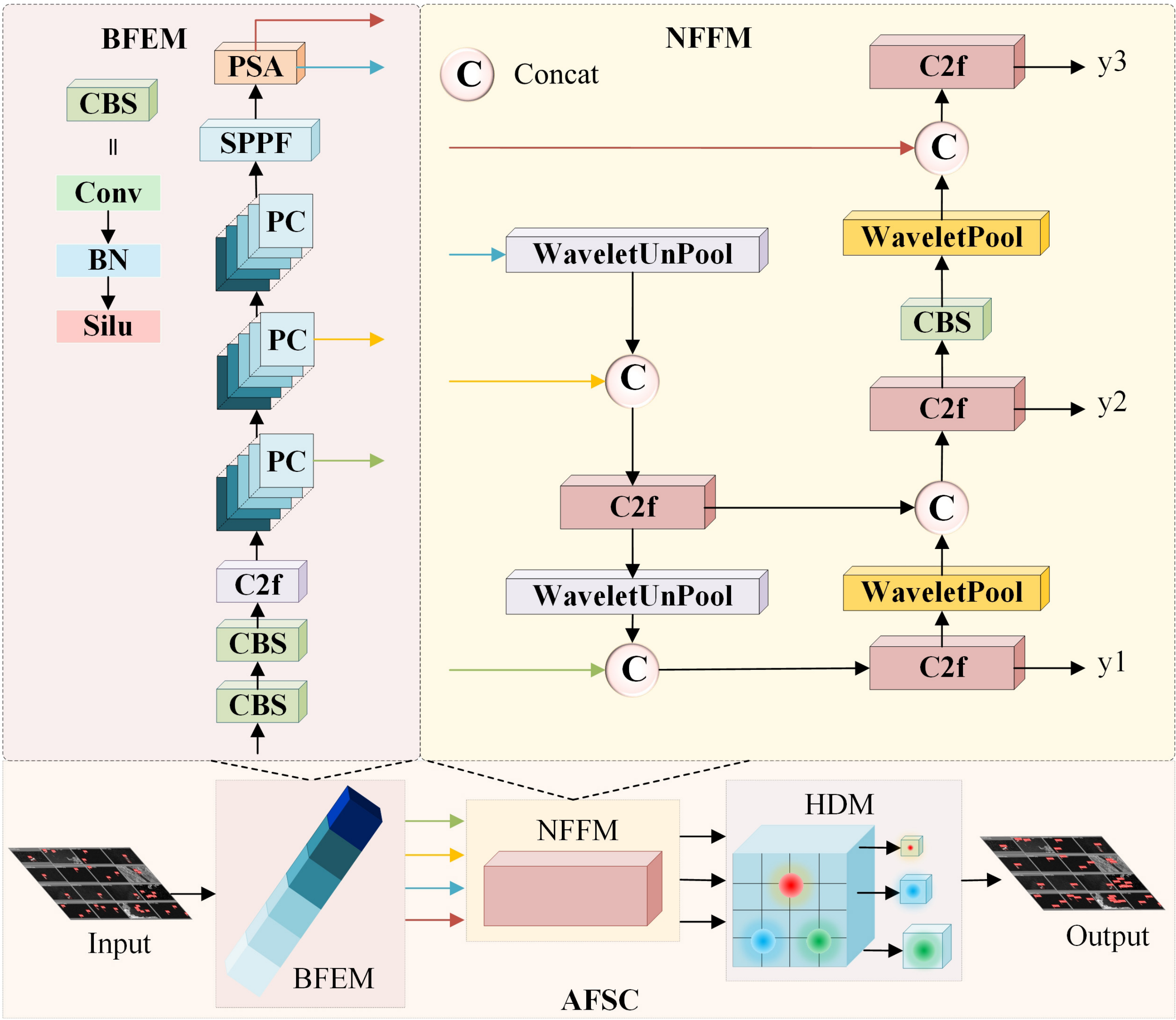


Figure 1. Architecture of the proposed network AFSC, including BFEM, NFFM, and HDM.

2. Methodology

In this section, we will delve into and explain the complexities of the proposed detector, AFSC, which is mainly composed of three essential modules: BFEM, NFFM, and HDM.

2.1. Backbone Feature Extraction Module (BFEM)

To address this challenge, we have developed the BFEM module, which is tailored to mitigate the impact of background noise and enhance the model's capacity to identify features within smaller targets. This is primarily achieved by utilizing the PC sub-module, as illustrated in Figure 2.

PC Sub-Module

In contrast to typical optical images, SAR images provide a wide-area view from an aerial perspective, resulting in intricate background details. The significant noise in the background can obscure or reduce the details of the objects, which negatively impacts the accuracy of detection. To tackle these issues, we have developed the PC sub-module.

Anti-aliasing is key to detecting small targets, and before the use of anti-aliasing techniques, spotting small targets was often tough because they were commonly mistaken or missed in the process. The incorporation of anti-aliasing methods has effectively resolved this issue, boosting the network's effectiveness in spotting small targets. Moreover, to further boost the network's capacity to pull out features from small targets, we have incorporated the $C2f$ module, which is meticulously designed to amplify feature extraction prowess.

By integrating the WaveletPool and $C2f$ modules, we propose a novel module, PC, which focuses more intently on the feature extraction of small targets. This module can effectively reduce model complexity and offers improved detection performance for small targets in the SAR imagery of ships.

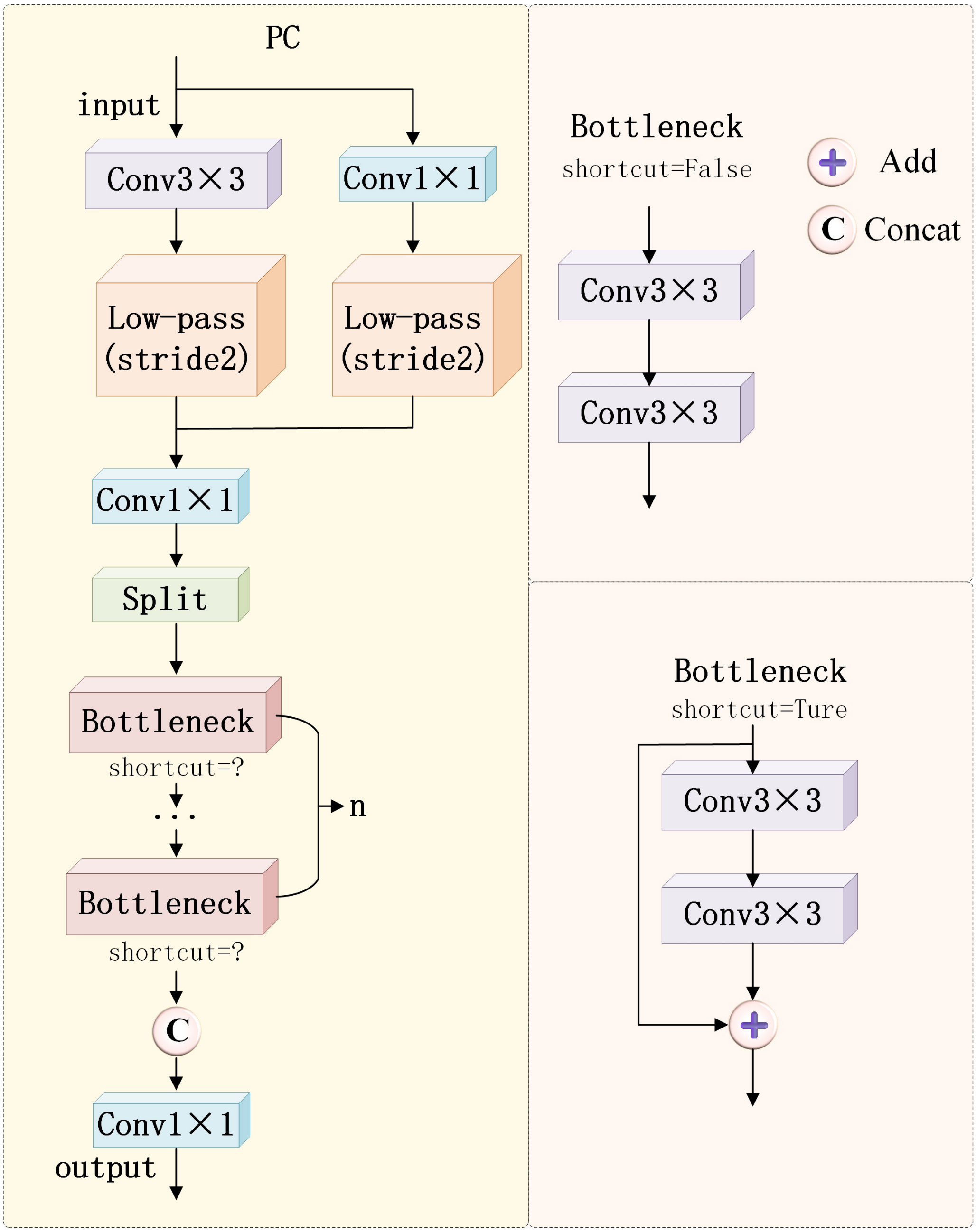


Figure 2. Architecture of the proposed sub-module PC.

At the outset, the input feature map X is subjected to Wavelet Pooling, which decomposes the image into elements that span various scales and directions:

$$W_{\phi} = DWT_{\phi}(X) \tag{1}$$

$$W_{\psi} = DWT_{\psi}(X) \tag{2}$$

Following this, a pooling operation is conducted on the approximation coefficients:

$$P = Downsample(W_{\phi}) \tag{3}$$

Subsequently, the feature map is reconstructed through the application of the inverse wavelet transform:

$$X' = IDWT(P, 0, 0) \tag{4}$$

The output from the Wavelet Pooling, X', is subsequently processed through the $C2f$ module, which consists of a sequence of convolutions and bottleneck operations designed to further refine and fuse features:

$$Y_a, Y_b = Conv(X'; filters = 2c, kernel\_size = 1, stride = 1) \tag{5}$$

$$Y_{bottleneck} = Bottleneck(Y_b; c, c, shortcut, g, k) \tag{6}$$

$$Y_{merged} = Concat(Y_a, Y_{bottleneck}) \tag{7}$$

Combining the above steps, the complete formula for the integration of the PC module can be expressed as

$$Y_{out} = Conv(Y_{merged}; filters = c2, kernel\_size = 1) \tag{8}$$

This equation represents the complete sequence of processing steps that begins with the input feature map X and culminates in the output feature map Y_{out}, which encompasses wavelet transformation, pooling, inverse wavelet transformation, feature mapping, splitting, bottleneck processing, feature merging, and final feature mapping.

2.2. Neck Feature Fusion Module (NFFM)

SAR images encompass a variety of distinct features. Typically, the receptive field of shallow features is more confined, and the overlap between these receptive fields is minimal, enabling the network to seize more granular details. The shallow-level information possesses higher resolution and is richer in locational and detailed data, yet it contains less semantic content and is more prone to noise due to fewer convolutional layers. Conversely, deep features have a larger receptive field, increasing the overlap between these fields, compressing the image, and yielding a holistic representation rich in semantic information. However, this comes at the cost of very low resolution and poor detail perception. To enhance the detection capabilities for ship targets within Synthetic Aperture Radar (SAR) imagery, we have crafted an NFFM (Non-Local Feature Fusion Module) aimed at integrating features. This module amalgamates information from various features, thereby bolstering the model's robustness against diverse conditions. Leveraging an array of features diminishes the risk of model overfitting, as they provide a more diverse dataset, consequently boosting the model's capacity to generalize. Essentially, the NFFM feature fusion module enables the model to gain a deeper understanding of the data, which subsequently elevates its performance and generalization capabilities. As illustrated in Figure 1, it mainly comprises two submodules: WaveletPool [23] and $C2f$ [21].

The NFFM process can be expressed as

$$Y_1 = C2f(Y_{\alpha}, Additional\_Features) \tag{9}$$

$$Y_2 = C2f(Y_{\beta}) \tag{10}$$

$$Y_3 = C2f(Y_\delta) \tag{11}$$

In the formula, Y_α, Y_β and Y_δ, respectively, represent:

$$Y_\alpha = C2f(postprocess(concat(predict(WaveletUnPool(PC(X_i)))_{i=1}^{n}), max\_det, nc)) \tag{12}$$

$$Y_\beta = Concat(Wavelet\ Pooling\ Process(X), C2f\ Module\ Process(X)) \tag{13}$$

$$Y_\delta = PSAConcat(WaveletPool(X), OtherFeatures) \tag{14}$$

In Equations (9) and (12): Y_α is the output from the object detection process. $Additional\_Features$ represents additional feature information that comes from different parts of the same network.

The detailed steps are as follows.

Object detection process: From the input feature map X, after a series of processing steps (including point-wise convolution, wavelet unpooling, convolutional processing, prediction layers, decoding operations, and post-processing), Y_α is obtained.

$C2f$ fusion: $C2f(\cdot)$ represents the operation of the $C2f$ module that takes the concatenated output and produces the final output. The features denoted as Y_α and $Additional\_Features$ are merged within the $C2f$ module to generate the ultimate output Y_1.

In Equations (10) and (13), concatenate the outputs from Wavelet Pooling and the $C2f$ module to form Y_β. Process Y_β through the $C2f$ module to refine and finalize the features and detections. Obtain the final output Y_2, which represents the ultimate result of the object detection and feature fusion pipeline.

In Equations (11) and (14), accept the concatenated output Y_δ from the previous stage, which includes the results from Wavelet Pooling and PSA Concat. Process Y_δ through the $C2f$ module, which may involve additional convolutional operations, feature fusion, and activation functions to refine the features. Obtain the final output Y_3, which represents the ultimate result of the feature enhancement and decision-making pipeline.

2.3. Head Detection Module (HDM)

Due to the distinct imaging process of SAR imagery, there is an increased level of interference noise in the background. Traditional structures of detection algorithms are often complex, leading to slower processing times and inadequate performance in detecting small targets. This situation may result in a high rate of false detections and missed targets. To enhance the precision and speed of detecting small ships in SAR imagery, we have engineered a novel detection module called HDM, specifically designed to tackle these difficulties.

The architectural diagram of the HDM is displayed in Figure 3. Briefly, the contributions of HDM are threefold:

1. Group Normalization (GN), particularly when applied in the context of Convolutional layers as ConvGN, has demonstrated its efficacy in enhancing the accuracy of object localization and classification tasks within detection headers, as empirically validated in the seminal work on FCOS [24].
2. Employing shared convolutional operations greatly diminishes the model's parameter count, substantially cutting down on computational demands and making the model practical for use on resource-constrained platforms.
3. To address the variance in the scales of targets detected by individual detection heads within a shared convolutional framework, a scale layer is integrated to uniformly adjust the feature magnitudes, thereby ensuring consistency in the subsequent analysis and processing stages.

In conclusion, the detection head has been refined to minimize parameter count and computational load, while still aiming to preserve the highest level of accuracy, thus striking a balance between efficiency and effectiveness.

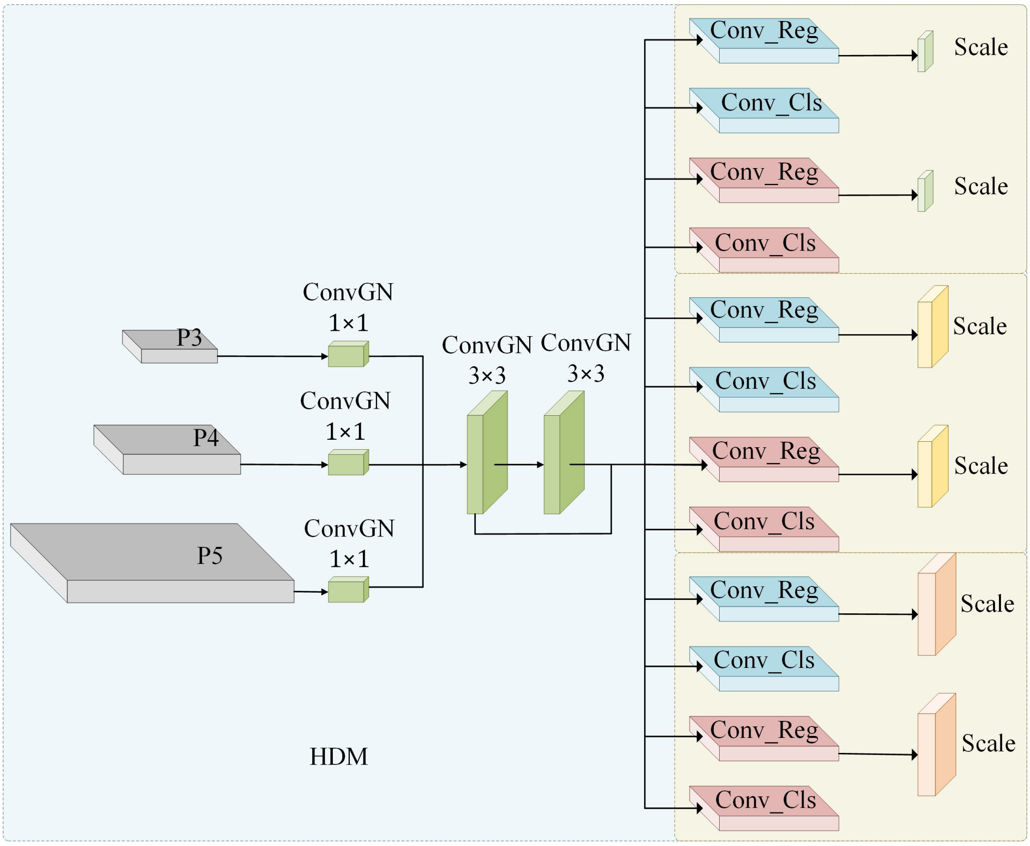


Figure 3. Architecture of the proposed module HDM.

The Head Detection Module (HDM) plays a pivotal role in object detection networks, tasked with producing the ultimate detection results from the feature maps that have been processed by preceding layers. This module integrates normalization, shared convolutions, and scaling operations to optimize detection accuracy and computational efficiency.

Group Normalization: Group Normalization (GN) is utilized to secure the training procedure and augment the model's efficacy in generalizing across different data sets:

$$Y_{\text{gn}} = GN(X) \tag{15}$$

where X signifies the input feature map and Y_{gn} indicates the output subsequent to the application of group normalization.

Shared Convolution: A conventional convolutional operation is employed to reduce the model's parameter count and mitigate computational demands.

$$Y_{\text{shared}} = CONV_{\text{shared}}(Y_{\text{gn}}) \tag{16}$$

where $CONV_{\text{shared}}$ denotes the convolutional operation shared across the module.

Scale Layer: A scale layer adjusts the feature responses according to the varying scales of targets within the feature map:

$$Y_{\text{scale}} = SCALE(Y_{\text{shared}}) \tag{17}$$

where $SCALE$ is the scaling operation tailored to normalize feature magnitudes.

Detection Operation: The final detection operation synthesizes the processed features to output the bounding boxes, class labels, and confidence scores:

$$Y_{\text{final}} = DETECT(Y_{\text{scale}}) \tag{18}$$

where $DETECT$ encapsulates the operations for bounding box regression and class classification.

Complete Formula: The comprehensive formula for the Head Detection Module, integrating all the aforementioned steps, is expressed as

$$Y_{\text{final}} = DETECT(SCALE(CONV_{\text{shared}}(GN(\mathcal{X})))) \tag{19}$$

This formula outlines the entire process from the input feature map X to the ultimate detection results Y_{final}, which includes normalization, shared convolutions, and scaling to enhance detection performance.

2.4. Integrated Anti-Aliasing and Fully Shared Convolution (AFSC)

The AFSC substantially enhances the detection capabilities for small targets in SAR imagery by integrating anti-aliasing and fully shared convolution methods. This approach maintains computational efficiency while reducing the loss of feature dimensions and improving detection accuracy. The cooperative functioning of the different modules enables our detection system to excel over existing top-tier detectors in detecting small-sized targets.

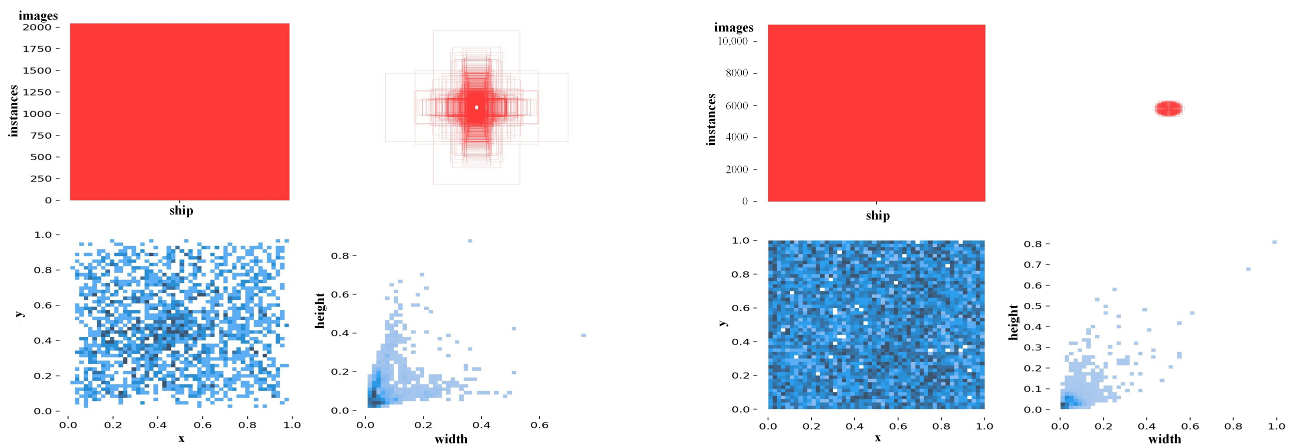
3. Experiments and Results

In this part, we will conduct a thorough assessment of the AFSC's performance. Initially, we will outline the datasets we employed, the experimental configuration, and the metrics for evaluation. Following that, we will present a comparative analysis, comparing the performance of the AFSC with that of other leading detectors currently in use. In summary, we carry out a variety of ablation studies to confirm the individual contributions and effectiveness of the proposed modules within the framework.

3.1. Datasets and Experimental Setup

3.1.1. Datasets

We leveraged two substantial real-world datasets from a public repository: the Synthetic Aperture Radar Ship Detection dataset (SSDD) [25] and the high-definition Synthetic Aperture Radar image dataset (HRSID) [26]. These datasets were utilized to confirm the validity and generalizability of our proposed method. Additionally, the distribution of bounding box sizes for both datasets is depicted in Figure 4.



(**a**) SSDD. (**b**) HRSID.

Figure 4. Details regarding the dimensions of the bounding boxes within both datasets. The **top-left** graph of each dataset is the number of data in the training set and how many there are in each category; the **top right** figure is the size and number of boxes; the **lower-left** figure depicts the location of the object's center relative to the entire image; and the **lower-right** figure shows the object's height-to-width ratio in comparison to the entire image.

The SSDD dataset, released in 2017, consists of 1160 images depicting intricate scenes along with 2456 ships of diverse sizes. Each image was resized to a resolution of 512 by 512 pixels. To ensure uniformity in variables, all models within this experiment were trained and tested on the same divided datasets.

To evaluate the model's ability to generalize, we conducted further experiments using the HRSID dataset. Released in 2020, the HRSID dataset consists of 5604 images, with each image measuring 800 by 800 pixels and containing 16951 ship instances. The images are provided in spatial resolutions of 0.5 m, 1 m, and 3 m. The dataset encompasses ship targets located in wide oceanic areas as well as coastal locales. Furthermore, 98% of the ships within the dataset are classified as small to medium-sized, making it particularly well-suited for high-resolution detection tasks.

3.1.2. Experimental Equipment

For the experimental configuration, the GPU computing environments were provisioned with CUDA 11.6 and CUDNN 8.8.0. Also, the specifics of the experimental hardware are presented in Table 1.

Table 1. Experimental equipment.

Item	Description
Python version	3.10
Pytorch version	1.11.0
CPU type	Intel® CoRE ™ I5-11600K CPU
GPU type	NVIDIA GeForce RTX 3060 GPU
Windows	Windows11, a 64-bit operating system
CUDA and CUDNN version	CUDA 11.6 + CUDNN 8.8.0

3.1.3. Evaluating Metrics

Experimental evaluation indexes are shown in Table 2.

Table 2. Experimental evaluation indexes.

Metric	Equation
P	$P = \frac{TP}{TP+FP}$
R	$R = \frac{TP}{TP+FN}$
F1	$F1 = 2 \times \frac{P \cdot R}{P+R}$
AP	$AP = \int_0^1 p(r)dr$

The oriented Intersection over Union (IoU) quantifies the overlap between the predicted bounding box and the ground-truth bounding box as a fraction of their combined area. If the IoU exceeds a specified threshold t, the predicted bounding box is considered a correct detection.

This is illustrated by the subsequent equation:

$$IOU = \frac{B_p \cap B_{gt}}{B_p \cup B_{gt}} \tag{20}$$

Furthermore, in order to make a better comparison between various methods, we also adopted COCO metrics [3] to evaluate the results.

3.2. Results and Analyses

Comparison with SOTA Methods

In this subsection, the SOTA algorithms Faster-RCNN [27], MobileNet [7], SSD [28], RetinaNet [29], YOLOv7 [20], YOLOv8 [21] and YOLOv10 [21] are chosen as the comparison methods.

AFSC showed the best detection results in AP, AP_{75}, AP_S, and AP_L, and these results show that our network performs best on small ship targets in SAR images. AFSC achieved 68.1%, 82.3%, 66.5% , and 70.8% results for AP, AP_{75}, AP_S and AP_L, respectively, surpassing the SOTA detector for the best results. As shown in Table 3.

To ascertain the reliability and adaptability of the proposed detection model across various datasets, a subsequent analysis of its performance on the HRSID dataset is undertaken.

Table 4 presents the quantitative outcomes of the various ship detectors on the HRSID dataset. It is evident that, when compared to other SAR ship detectors, our model exhibits a significant advantage across all evaluation metrics.

Table 3. AFSC versus other networks on the SSDD dataset.

Methods	AP (%)	AP_{50} (%)	AP_{75} (%)	AP_S (%)	AP_M (%)	AP_L (%)	Speed (ms)
Faster R-CNN [27]	65.04	96.12	77.47	61.68	71.54	66.07	85.9
MobileNet [7]	41.88	79.89	38.97	35.65	56.9	29.09	57.97
SSD [28]	52.94	88.07	60.08	49	61.26	52.5	98.2
RetinaNet [29]	60.76	90.9	70.46	57.32	67.62	61.77	87.1
YOLOv7 [20]	56.8	89.9	64.3	55	63.3	32.2	15.0
YOLOv8 [21]	68.0	97.2	81.5	66.3	72.3	63.8	3.4
YOLOv10 [21]	67	96.2	80.3	65.2	71.6	51.2	2.9
AFSC (Ours)	68.1	96.2	82.3	66.5	71.1	70.8	3.2

Table 4. AFSC versus other networks on the HRSID dataset.

Methods	AP (%)	AP_{50} (%)	AP_{75} (%)	AP_S (%)	AP_M (%)	AP_L (%)	Speed (ms)
Faster R-CNN [27]	56.7	81.1	65.3	43.6	74.5	46.3	78.5
MobileNet [7]	35.4	59.1	39	16.1	62	34.6	64.7
SSD [28]	43.5	67.1	49.8	23.3	70.4	52.7	116.3
RetinaNet [29]	50.7	83.1	54	36.2	72.1	63.2	72.7
YOLOv7 [20]	54	85.4	58.6	40.5	73.2	70	18.7
YOLOv8 [21]	57.7	83.2	65	42.8	77.3	59.3	4
YOLOv10 [21]	57.3	83.8	65.1	43.8	75.8	41.7	2.9
AFSC (Ours)	59.8	85.5	68.8	47	76.8	63.3	2.9

The value indicates that in AP, it is 2.1% higher than the best detector YOLOv8. In terms of AP_{50} and AP_S, both are 2.2% and 1.5% higher than the best detector Fater-RCNN. In terms of AP_M, it is 1.9% more than the best detector, YOLOv8. In terms of AP_L, it is more than 8.6% more than RetinaNet.

AFSC demonstrated the highest performance in AP, AP_{50}, AP_{75}, AP_S, AP_L, and speed, all outperforming SOTA detector; AFSC outperformed the second-place detector YOLOv8 by 2.1% in AP, and also achieved the best detection result in AP_{50}. It outperformed the second detector Faster R-CNN by 3.5%. It was 3.2% ahead of the second-place detector, YOLOv10, in AP_S. It also performed best in AP_L. At the same time, AFSC showed excellent advantages in speed. These figures indicate that AFSC achieves the highest precision in detecting small ship targets within SAR imagery.

The results from the HRSID dataset confirm that our model exhibits strong robustness and excellent generalization capabilities.

Additionally, the visual outcomes on the SSDD dataset are illustrated in Figure 5, providing a visual representation of the detection results. These visual outcomes clearly indicate that our model is exceptionally adept at detecting ships in SAR imagery. The results also imply that our model is capable of identifying ships in both open sea and coastal areas, even in the presence of intricate backgrounds. Additionally, the results indicate that our detector excels at identifying small targets.

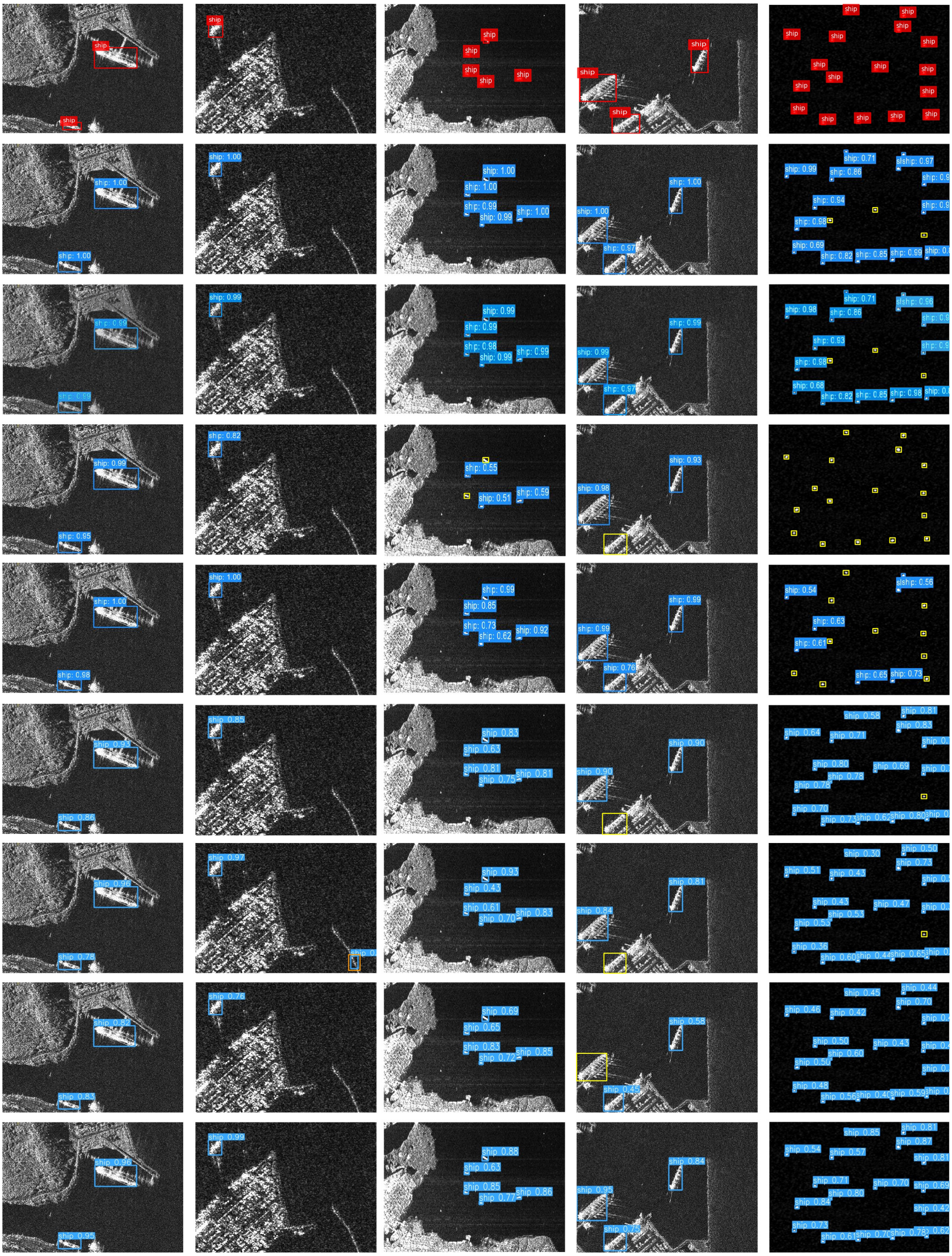


Figure 5. The visual results on the SSDD dataset are displayed. Red bounding boxes indicate the actual positions, while blue bounding boxes denote the predicted locations, and yellow bounding boxes indicate the missed detections and the orange box indicates false detection. From the first line to line 9 are Ground Truth, Faster R-CNN, MobileNet, SSD, RetinaNet, YOLOv7, YOLOv8, YOLOv10 and AFSC.

4. Discussion

To gain a more exhaustive understanding of the detection performance of each model, we will proceed to conduct a range of ablation studies. Taking YOLOv10 as the base, the experimental setup is presented in Table 5. An × indicates that the module is not used, and a ✓ indicates that the module is used. Case 1 only adds the HDM module. Case 2 only introduces BFEM. Case 3 only adds NFFM. Case 4 introduces HDM and BFEM at the same time, Case 5 introduces HDM and NFFM at the same time, Case 6 introduces BFEM and NFFM at the same time, and Case 7 adds HDM on the basis of Case 6. For specific experimental conditions, please refer to Table 5. The data from Case 1 to Case 3 showcase the efficacy of each module, each playing a role in improving the detection of small targets in SAR images to different extents. Case 1 shows that the integration of the HDM module greatly improves the model's precision in spotting small targets in SAR imagery and also quickens the detector's overall processing speed. Experimental data show that the detection speed increased from the original 2.9 ms to 2.5 ms, and the overall speed increased by 13.8%. Case 2 represents a drop in GFLOPs from 6.5 to 6.2, which means that the complexity of the model has decreased, making the overall model more lightweight. And it still has a very good effect on small-target ship detection. Case 2 indicates that when utilizing solely the BFEM module, the total number of parameters in our detector was reduced to 74.5% of its original count, consequently enhancing the overall detection speed of the model. Case 3 greatly improves the AP_S of the whole model, which proves that Case 3 has a better detection effect on small targets.

Table 5. Basic information of ablation experiment.

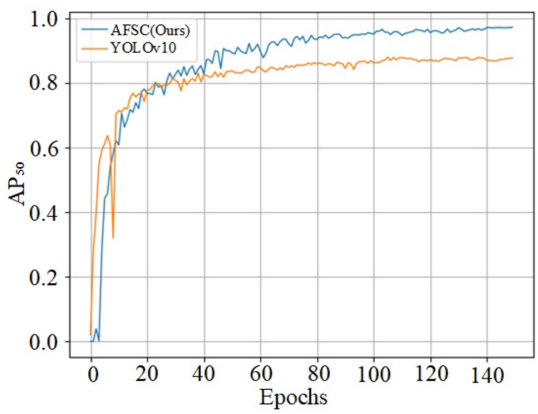
Case	HDM	BFEM	NFFM	P (%)	R (%)	AP_{50}(%)	$AP_{50:95}$(%)	F1 (%)	GFLOPs	Speed (ms)
Base	×	×	×	94.5	90.9	97.2	68.9	92.7	6.5	2.9
Case 1	✓	×	×	90.9	90.3	96.5	69.2	90.6	6.3	2.5
Case 2	×	✓	×	94.8	89.5	97	68.8	92.1	6.2	3.2
Case 3	×	×	✓	91.8	93.2	97.1	68.5	92.5	6.4	3.4
Case 4	✓	✓	×	92	86.9	95.6	68.8	89.4	6.0	3.0
Case 5	✓	×	✓	92.7	89.9	96.1	67.8	91.3	6.1	2.9
Case 6	×	✓	✓	92.6	93.6	96.9	69.2	93.1	6.1	3.4
Case 7	✓	✓	✓	92.6	91.6	97	69.3	92.1	5.8	3.2

Cases 4 to 7 illustrate the effective collaboration among all components of the detector, which equips our detector with superior capabilities for detecting small targets in SAR images. Keep in mind that, to save space, the ablation studies were only carried out using the SSDD dataset. It is evident that each module provides some level of improvement for the SAR ship detection task. Therefore, the modules developed in this paper are advantageous for ship detection. The basic information of the ablation experiment is shown in Table 5, and the situation under the COCO metrics of the ablation experiment is shown in Table 6.

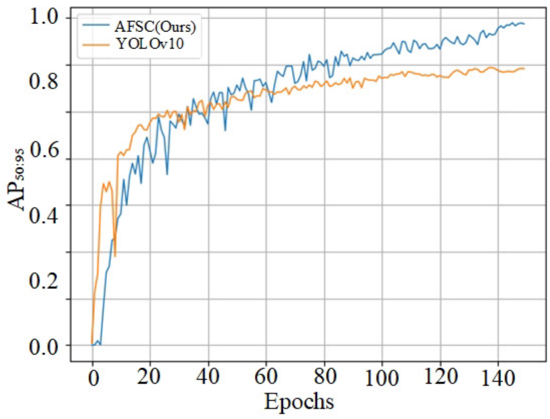
Table 6. The situation under the COCO metrics of the ablation experiment.

Case	AP(%)	AP_{50}(%)	AP_{75}(%)	AP_S(%)	AP_M(%)	AP_L(%)
Base	57.3	83.8	65.1	43.8	75.8	41.7
Case 1	60	85.9	68.7	47	76.8	66.9
Case 2	67.2	96.3	81.8	65	71.9	55.2
Case 3	66.8	96.3	81.6	64.4	72.15	54.9
Case 4	67.9	95.1	82.3	66	71.3	72.7
Case 5	66.8	95.5	81.5	65.3	70.3	69.1
Case 6	56.1	84.1	62.9	42.9	75.5	28.3
Case 7	59.8	85.5	68.8	47	76.8	63.3

To vividly demonstrate the effectiveness of our method, Figure 6 presents a comparison between the baseline detection performance curve of the standard YOLOv10 approach and the performance curve of our AFSC method. Our method has a faster learning time. (a) is the evaluation index of AP_{50}, and (b) is the evaluation index of $AP_{50:95}$. It is evident that AFSC can attain satisfactory outcomes within a few epochs.



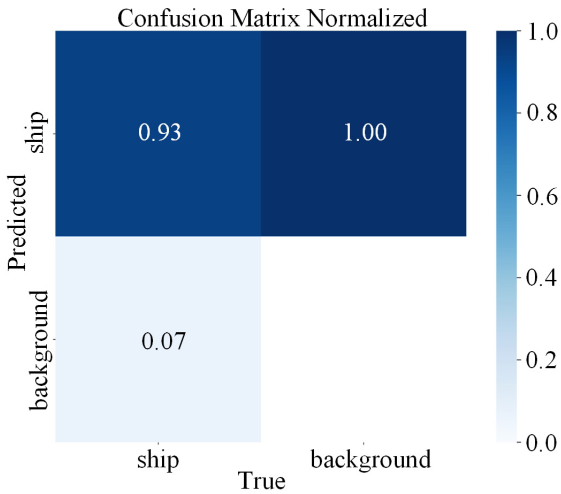
(**a**) AP_{50} comparison curve.



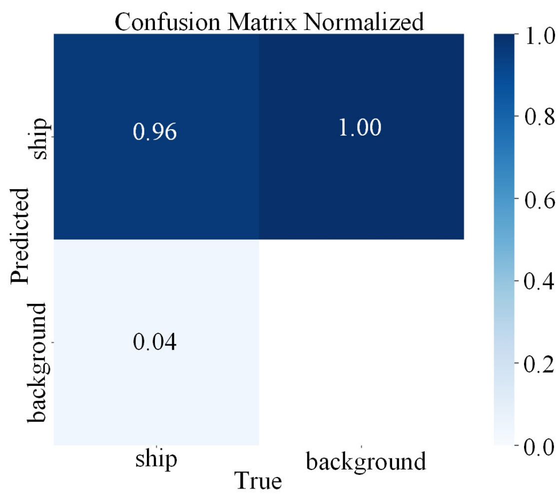
(**b**) $AP_{50:95}$ comparison curve.

Figure 6. The curves of the experimental results, with the x-axis indicating the number of epochs and the y-axis representing the corresponding quantitative results.

To more effectively present our experimental results, we utilize the confusion matrix. The confusion matrix offers a tabular summary of the prediction outcomes for classification tasks. To vividly present our experimental results, we utilize a confusion matrix. This matrix offers an organized overview of the classification predictions. This matrix makes it easy to see if the machine has confused two different classes, mistaking one for the other. The confusion matrix not only enables us to easily grasp the mistakes made by the classification model, but more crucially, it helps us identify the specific types of errors occurring. This detailed breakdown of results overcomes the shortcomings of relying solely on overall classification accuracy (from general to specific). Figure 7 displays the confusion matrices for both the base model and the AFSC model. The True Positive (TP) value for the base model is 0.93, while for the AFSC model, it is 0.96, indicating that AFSC achieves superior accuracy.



(**a**) Base confusion matrix.



(**b**) AFSC confusion matrix.

Figure 7. Based on the actual and predicted categories from the classification model, the confusion matrix organizes the dataset's records into a matrix format. In this matrix, the rows correspond to the true categories, and the columns correspond to the model's predicted categories. The above are the confusion matrices for the base model and the AFSC model, respectively.

5. Conclusions

This paper introduces the AFSC detector, an innovative solution designed to improve ship detection within Synthetic Aperture Radar (SAR) imagery. The AFSC detector integrates anti-aliasing techniques with fully shared convolutional layers to improve detection accuracy and robustness. By reducing noise and preserving target details, the AFSC detector outperforms traditional methods, as demonstrated by extensive evaluations on real-world SAR datasets.

The BFEM is an essential module specifically designed for extracting features of small targets in SAR imagery. It skillfully mitigates the impact of complex backgrounds and noise on the detection of small targets by integrating advanced feature extraction techniques, thereby enhancing the model's ability to extract meaningful features. Subsequently, we introduce the NFFM and incorporate the WaveletPool module to enable the neural network to more effectively balance locational information and semantic information during feature fusion. Ultimately, the HDM module is crafted to boost the detection capabilities for small vessels. The experimental outcomes indicate that our proposed AFSC algorithm outperforms current ship detection models. Moreover, considering that the incorporation of multiple modules in AFSC could potentially escalate the computational demands of ship inspection models, our future research will concentrate on developing lightweight ship inspection models.

Author Contributions: Conceptualization, Z.Y. (Zhen Yang) and M.H.; methodology, Z.Y. (Zhen Yang); software, M.H.; validation, J.L., Z.Y. (Zhen Yang) and Z.Y. (Zhijian Yin); formal analysis, J.L.; investigation, Z.Y. (Zhijian Yin); resources, Z.Y. (Zhen Yang); data curation, J.L.; writing—original draft preparation, M.H.; writing—review and editing, M.H.; visualization, J.L.; supervision, Z.Y. (Zhijian Yin); project administration, J.L.; funding acquisition, Z.Y. (Zhijian Yin) and Z.Y. (Zhen Yang). All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (No. 62261026, 62061019, 62201343), the Jiangxi Natural Science Foundation (No. 20232ACB212006).

Data Availability Statement: The experimental data used for verification in this article can be publicly obtained on the extreme mart, with the identifier https://github.com/13672027370/ELE, accessed on 1 July 2024.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhang, T.; Ji, J.; Li, X.; Yu, W.; Xiong, H. Ship Detection From PolSAR Imagery Using the Complete Polarimetric Covariance Difference Matrix. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 2824–2839. [CrossRef]
2. Xu, C.; Li, Y.; Ji, C.; Huang, Y.; Wang, H.; Xia, Y. An improved CFAR algorithm for target detection. In Proceedings of the 2017 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), Xiamen, China, 6–9 November 2017; pp. 883–888.
3. Sun, K.; Liang, Y.; Ma, X.; Huai, Y.; Xing, M. DSDet: A Lightweight Densely Connected Sparsely Activated Detector for Ship Target Detection in High-Resolution SAR Images. *Remote Sens.* **2021**, *13*, 2743. [CrossRef]
4. Gao, S.; Liu, H. RetinaNet-Based Compact Polarization SAR Ship Detection. *IEEE J. Miniaturization Air Space Syst.* **2022**, *3*, 146–152. [CrossRef]
5. Li, Q.; Xiao, D.; Shi, F. A Decoupled Head and Coordinate Attention Detection Method for Ship Targets in SAR Images. *IEEE Access* **2022**, *10*, 128562–128578. [CrossRef]
6. Yue, T.; Zhang, Y.; Liu, P.; Xu, Y.; Yu, C. A Generating-Anchor Network for Small Ship Detection in SAR Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 7665–7676. [CrossRef]
7. Lin, J.; Zhao, Y.; Wang, S.; Tang, Y. YOLO-DA: An Efficient YOLO-Based Detector for Remote Sensing Object Detection. *IEEE Geosci. Remote Sens. Lett.* **2023**, *20*, 6008705. [CrossRef]
8. Liu, Z.; Zhang, W.; Yu, H.; Zhou, S.; Qi, W.; Guo, Y.; Li, C. Improved YOLOv5s for Small Ship Detection with Optical Remote Sensing Images. *IEEE Geosci. Remote Sens. Lett.* **2023**, *20*, 8002205. [CrossRef]
9. Qin, C.; Wang, X.; Li, G.; He, Y. A Semi-Soft Label-Guided Network with Self-Distillation for SAR Inshore Ship Detection. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5211814. [CrossRef]

10. Gao, Y.; Wu, C.; Ren, M. A Novel Anchor-Free Model with Salient Feature Fusion Mechanism for Ship Detection in SAR Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2023**, *16*, 9089–9105. [CrossRef]
11. Sun, Q.; Liu, M.; Chen, S.; Lu, F.; Xing, M. Ship Detection in SAR Images Based on Multilevel Superpixel Segmentation and Fuzzy Fusion. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5206215. [CrossRef]
12. Hao, Y.; Zhang, Y. A Lightweight Convolutional Neural Network for Ship Target Detection in SAR Images. *IEEE Trans. Aerosp. Electron. Syst.* **2024**, *60*, 1882–1898. [CrossRef]
13. Zhang, J.; Sheng, W.; Zhu, H.; Guo, S.; Han, Y. MLBR-YOLOX: An Efficient SAR Ship Detection Network with Multilevel Background Removing Modules. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2023**, *16*, 5331–5343. [CrossRef]
14. Lei, M.; Liu, X. SOLO-Net: A Sparser but Wiser Method for Small Object Detection in Remote-Sensing Images. *IEEE Geosci. Remote Sens. Lett.* **2024**, *21*, 6001005. [CrossRef]
15. Huang, P.; Tian, S.; Su, Y.; Tan, W.; Dong, Y.; Xu, W. IA-CIOU: An Improved IOU Bounding Box Loss Function for SAR Ship Target Detection Methods. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 10569–10582. [CrossRef]
16. Zhou, S.; Zhang, M.; Wu, L.; Yu, D.; Li, J.; Fan, F.; Zhang, L.; Liu, Y. Lightweight SAR Ship Detection Network Based on Transformer and Feature Enhancement. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 4845–4858. [CrossRef]
17. Yang, Z.; Zhang, P.; Wang, N.; Liu, T. A Lightweight Theory-Driven Network and Its Validation on Public Fully Polarized Ship Detection Dataset. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 3755–3767. [CrossRef]
18. Jia, H.; Pu, X.; Liu, Q.; Wang, H.; Xu, F. A Fast Progressive Ship Detection Method for Very Large Full-Scene SAR Images. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5206615. [CrossRef]
19. Chen, C.; Zhang, Y.; Hu, R.; Yu, Y. A Lightweight SAR Ship Detector Using End-to-End Image Preprocessing Network and Channel Feature Guided Spatial Pyramid Pooling. *IEEE Geosci. Remote Sens. Lett.* **2024**, *21*, 4003605. [CrossRef]
20. Zhou, L.; Wan, Z.; Zhao, S.; Han, H.; Liu, Y. BFEA: A SAR Ship Detection Model Based on Attention Mechanism and Multiscale Feature Fusion. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 11163–11177. [CrossRef]
21. Li, Y.; Zheng, J.; Li, S.; Wang, C.; Zhang, Z.; Zhang, X. AID-YOLO: An Efficient and Lightweight Network Method for Small Target Detector in Aerial Images. *Electronics* **2024**, *13*, 3564. [CrossRef]
22. Williams, T.; Li, R.Y. Wavelet Pooling for Convolutional Neural Networks. In Proceedings of the International Conference on Learning Representations, Vancouver, BC, Canada, 30 April–3 May 2018. Available online: https://api.semanticscholar.org/CorpusID:92998064 (accessed on 1 July 2024).
23. Ning, J.; Spratling, M. The Importance of Anti-Aliasing in Tiny Object Detection. *arXiv* **2023**, arXiv:2310.14221.
24. Tian, Z.; Shen, C.; Chen, H.; He, T. FCOS: Fully Convolutional One-Stage Object Detection. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019; pp. 9626–9635. Available online: https://api.semanticscholar.org/CorpusID:91184137 (accessed on 1 July 2024).
25. Zhang, T.; Zhang, X.; Li, J.; Xu, X.; Wang, B.; Zhan, X.; Xu, Y.; Ke, X.; Zeng, T.; Su, H.; et al. Sar ship detection dataset (ssdd): Official release and comprehensive data analysis. *Remote Sens.* **2021**, *13*, 3690. [CrossRef]
26. Wei, S.; Zeng, X.; Qu, Q.; Wang, M.; Su, H.; Shi, J. HRSID: A High-Resolution SAR Images Dataset for Ship Detection and Instance Segmentation. *IEEE Access* **2020**, *8*, 120234–120254. [CrossRef]
27. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [CrossRef]
28. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.-Y.; Berg, A.C. *SSD: Single Shot MultiBox Detector*; Springer International Publishing: Berlin/Heidelberg, Germany, 2016; pp. 21–37. [CrossRef]
29. Lin, T.-Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal loss for dense object detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *42*, 318–327. [CrossRef]