中图法分类号:TP3　　**文献标识码**:A　　**文章编号**:1006-8961(2025)01-0297-12

论文引用格式:Ying L, Zhang Z F, Miao D Q and Zhao C R. 2025. Multi-granularity context network for SAR ship detection. Journal of Image and Graphics, 30(01):0297-0308(应丽, 张志飞, 苗夺谦, 赵才荣. 2025. 多粒度上下文网络的SAR船舶检测. 中国图象图形学报, 30(01):0297-0308)[DOI:10.11834/jig.230838]

多粒度上下文网络的SAR船舶检测

应丽[1,2],张志飞[1,2*],苗夺谦[1],赵才荣[1]

1. 同济大学电子与信息工程学院,上海　201804; 2. 同济大学国家海底科学观测系统项目办公室,上海　200092

摘　要:**目的**　合成孔径雷达(synthetic aperture radar, SAR)是一种主动式微波传感器,能够获取高分辨率的遥感图像,在海上船舶检测中至关重要。然而,SAR船舶检测主要面临两个挑战:复杂背景和船舶目标尺寸的多样性。为此,提出了适用于SAR船舶检测的多粒度上下文网络。**方法**　首先,设计了多粒度通道注意力(multi-granularity channel attention, MCA)模块,对全局和局部的不同粒度的上下文信息进行加权,以增强对船舶目标重要信息的关注,降低复杂背景对检测结果的干扰。然后,设计了多粒度空洞自适应空间特征融合(multi-granularity atrous adaptive spatial feature fusion, MAASFF)模块,采用自适应空间特征融合的方法,将3种不同扩张率(1、2和3)的空洞卷积提取的特征图进行融合,以减少特征图生成中的上下文信息损失,增强特征金字塔的表示能力,从而提高不同尺寸船舶的检测效果。**结果**　实验结果表明,本文方法在SAR-Ship-Dataset和SSDD(SAR ship detection dataset)两个数据集上与其他9种方法进行比较均取得最佳的检测结果,平均精度分别达到了96.1%和97.0%,进一步验证了该网络在SAR船舶检测任务中出色的性能表现。**结论**　本文提出了一种多粒度上下文网络,旨在抑制复杂背景干扰并增强对多尺寸船舶特征的提取能力,有效提高了SAR船舶检测的准确性。

关键词:SAR图像;船舶检测;多粒度;通道注意力;特征融合

Multi-granularity context network for SAR ship detection

Ying Li[1,2], Zhang Zhifei[1,2*], Miao Duoqian[1], Zhao Cairong[1]

1. School of Electronic and Information Engineering, Tongji University, Shanghai　201804, China;

2. Project Management Office of China National Scientific Seafloor Observatory, Tongji University, Shanghai　200092, China

Abstract: **Objective** Synthetic aperture radar (SAR), as an active microwave sensor, can acquire high-resolution remote sensing images, which is crucial in marine ship detection. Nevertheless, two primary challenges confront ship detection in SAR images. First and foremost, SAR images frequently encompass complex backgrounds, incorporating turbulent sea waves, islands, and various forms of clutter. These complex backgrounds can significantly hinder the accurate identification of ship targets. The diverse spectrum of ship target sizes within SAR images presents another significant challenge. Traditional detection methods struggle to adapt to the broad range of ship sizes encountered in real-world scenarios. In recent years, with the extensive application of deep learning models and attention mechanisms, researchers have successfully improved the performance of SAR ship detection methods and effectively overcome the detection challenges of SAR

收稿日期:2023-12-12;**修回日期**:2024-06-18;**预印本日期**:2024-06-25

*通信作者:张志飞　zhifeizhang@tongji.edu.cn

基金项目:国家重点研发计划资助(2022YFB3104700);国家自然科学基金项目(62376198,62076182,62076040);上海市自然科学基金项目(22ZR1466700)

Supported by: National Key R&D Program of China(2022YFB3104700); National Natural Science Foundation of China(62376198, 62076182, 62076040); Natural Science Foundation of Shanghai, China(22ZR1466700)

ship targets of different sizes in complex backgrounds. However, these methods either have limitations in detection accuracy or require extensive computing resources. To address these challenges, this paper proposes a multi-granularity context network for SAR ship detection. **Method** First, a multi-granularity channel attention (MCA) module is designed to weight the global and local contextual information of different granularities. The primary function of the MCA module is to focus on important characteristics of ship targets and minimize interference caused by complex backgrounds on detection results. Pointwise convolution replaces traditional convolution as an aggregator of global and local channel context information to ensure a lightweight MCA module. This substitution not only trims computational overhead but also streamlines the process. Furthermore, the MCA module is integrated into the first layer of the backbone network feature extraction of the YOLOv5s framework. The fusion of pointwise convolution and integration within the network architecture collectively strengthens our capability for accurate and efficient SAR ship detection. Then, a multi-granularity atrous adaptive spatial feature fusion (MAASFF) module is designed to reduce the loss of contextual information in feature map generation and enhance the representation capability of feature pyramids, thereby improving the detection performance of ships at different sizes. In the process of fusing features of different granularities, the MAASFF module avoids ignoring the differences between ship target features of different sizes and reduces unnecessary computational overload. It primarily employs an adaptive spatial feature fusion method to merge the feature maps extracted using three different atrous rates (1, 2, and 3) of atrous convolutions. This design effectively captures features at different spatial granularities, enhancing the feature representation capability for ships of different sizes. **Result** Compared with nine other methods on two datasets, SAR-Ship-Dataset and SSDD, our method achieves the best detection results. On the SAR-Ship-Dataset, compared with the two-stage methods Faster R-CNN, DAPN, CR2A-Net, KCPNet, and BL-Net, our method can improve model detection accuracy by approximately 1.9% to 6.0%. Compared with common one-stage methods such as YOLOv4, CenterNet++, CRDet, and YOLOv5s, our method can enhance performance by 2.9%, 1.2%, 0.9%, and 1.8%, respectively. Experimental results indicate that our method achieves the best performance on the SAR-Ship-Dataset, reaching 96.1% mAP and thus outperforming all compared methods. On the SSDD dataset, our method improves the performance by approximately 8.7% (Faster R-CNN), 6.9% (DAPN), 7.2% (CR2A-Net), 4.5% (KCPNet), 1.8% (BL-Net), 0.9% (YOLOv4), 4.3% (CenterNet++), 0.6% (CRDet), and 1.6% (YOLOv5s) while maintaining a speed similar to the baseline YOLOv5s. These results show our method has good generalization ability and achieves the best performance with 97.0% mAP, further verifying its excellent effect and performance in SAR ship detection tasks. **Conclusion** This paper proposes a multi-granularity context network, which aims to suppress complex background interference and enhance the ability to extract features of multi-sized ships, effectively improving the accuracy of SAR ship detection.

Key words: synthetic aperture radar (SAR) image; ship detection; multi-granularity; channel attention; feature fusion

0 引 言

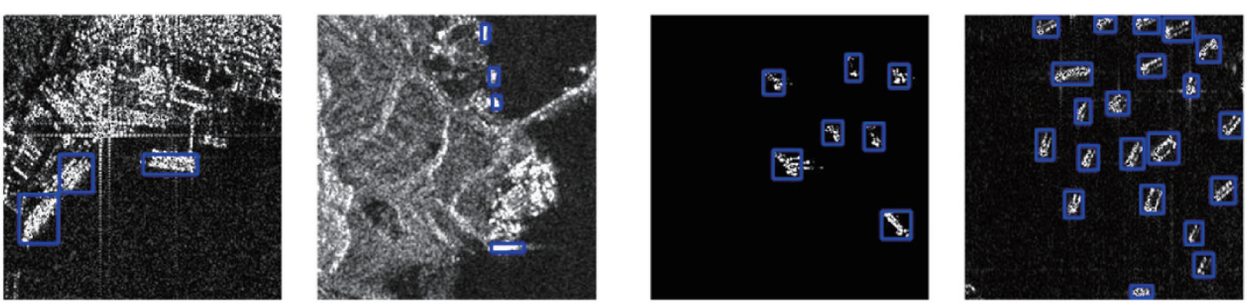
随着国家海洋战略的推进,海上船舶已成为海上监测、海事管理及战时追踪的关键目标,对船舶检测的要求不断提升。合成孔径雷达(synthetic aperture radar, SAR)采用独特的成像技术,具备全天候、全天时特点,不受自然因素影响,从而获得高分辨率的遥感图像(Oh等,2021),成为海上船舶检测的重要手段。然而,由于杂波、海洋和陆地等复杂背景和不同尺寸的船舶目标的影响(龚声蓉 等,2022;阮晨等,2021),SAR图像中船舶检测的难度大大增加。因此,在SAR图像中准确检测船舶目标是一项具有挑战性的任务。

SAR船舶检测方法分为传统方法和深度学习方法。由于SAR成像原理和船舶目标的多样性,传统的恒虚警率方法(Gao和Shi,2017)受到限制。近年来,学者们开始采用深度学习技术提升SAR船舶检测性能。基于深度学习的检测方法主要分为双阶段和单阶段两类。双阶段方法包括提议区域和目标分类与定位两个阶段,如R-CNN(Girshick等,2014)和Faster R-CNN(Ren等,2017)。虽然双阶段方法通常具有更高的准确性,但是需要更多的计算成本。而单阶段方法则将目标检测视为回归问题,直接利用卷积神经网络(convolutional neural network,CNN)完成整个检测过程,如SSD(single shot multibox detec-

tor)(Liu 等,2016)、YOLOv4(you only look once version 4)(Bochkovskiy 等,2020)和 YOLOv5(you only look once version 5)(Jocher,2021)。相较于双阶段,单阶段具有更快的速度和更简单的结构,广泛应用于各种实际任务中。

尽管深度学习方法在SAR船舶检测中取得了显著的进展,但仍然面临着复杂背景和不同尺寸船舶目标的检测挑战。为了应对这些挑战,近年来,研究人员开始结合流行的深度学习模型和注意力机制来提升SAR船舶检测算法的性能。Guo等人(2021)提出了CenterNet++,使用注意力机制和特征融合,以解决SAR图像中小船舶检测和复杂背景。Yu等人(2022)设计了CR2A-Net(cascade rotated anchor-aided network),能够在SAR图像中高效地检测不同尺寸的船舶。Zhang等人(2021)提出了BL-Net(balance learning network),通过引入平衡学习方法来解决不同尺寸SAR船舶检测中的数据不平衡问题,从而提升检测性能。Cui 等人(2019)设计了 DAPN(dense attention pyramid network),结合注意力机制和金字塔特征提取有效检测不同尺寸船舶。此外,Yang等人(2022)提出了CRDet(coordinate receptive detector),引入多尺度特征融合和背景抑制实现准确鲁棒的SAR船舶检测。Han等人(2023)设计了KCPNet(knowledge-driven context perception network),利用知识驱动的上下文感知网络和注意力机制提升SAR船舶检测性能。

上述这些方法在检测性能上表现出色,但在SAR船舶检测的应用中仍面临挑战,主要包括:1)在复杂背景的SAR图像中,船舶目标常常受到太多噪声的干扰。这是因为SAR的特殊工作原理,图像中含有大量杂波和斑点噪声(如图1(a)所示)。此外,图像背景还包括海洋、陆地和城市等,增加了背景的复杂性。这些因素会使真正的船舶目标难以辨别,产生许多漏检和虚警。2)SAR图像中船舶呈现出各种不同的尺寸特征。船舶的吨位和多分辨率成像模式的变化,使得船舶在SAR图像中显示出多样的尺寸特征(如图1(b)所示)。特别是小型船舶的存在,增加了虚警和漏检率,降低了模型的检测性能。



(a) 复杂背景　　(b) 不同尺寸船舶

图1　SAR-Ship-Dataset中的SAR图像呈现不同的背景和大小

Fig. 1　The SAR images in the SAR-Ship-Dataset present different backgrounds and sizes

((a) complex backgrounds; (b) ships of different sizes)

为了克服SAR船舶检测面临的上述挑战,本文提出了适用于SAR船舶检测的多粒度上下文网络,旨在提高在复杂背景下对不同尺寸船舶的检测准确性和可靠性。具体而言,本文引入了多粒度思想,设计了多粒度通道注意力(multi-granularity channel attention,MCA)模块,以减轻复杂背景对模型的影响,增强模型的鲁棒性和稳定性。同时,本文还设计了多粒度空洞自适应空间特征融合(multi-granularity atrous adaptive spatial feature fusion,MAASFF)模块,通过增强不同尺寸船舶目标的特征信息,尤其是小型船舶,来提升检测性能。

本文的主要贡献如下:1)设计了一个MCA模块,融入主干网络,减少复杂背景对船舶检测的干扰。该模块使用全局通道注意力和局部通道注意力,捕获多粒度的上下文信息,有效区分目标和背景,适应复杂背景下的船舶检测。2)提出了一个MAASFF模块,旨在增强对不同尺寸船舶目标的检测能力。该模块使用多个不同粒度的空洞卷积,并应用自适应空间特征融合来捕捉船舶特征。这增强了特征金字塔的多粒度特征表达,改进了对不同尺寸船舶的检测准确性,从而减少漏检和虚警。3)提出了在MCA模块中使用点卷积替代传统卷积,以及

在MAASFF模块中采用空洞卷积替代传统卷积，以降低模型参数和计算量。同时，通过在SAR-Ship-Dataset和SSDD（SAR ship detection dataset）数据集上进行大量实验，验证了本文方法的有效性。

1 方 法

在SAR船舶检测中，本文以YOLOv5作为基线模型，引入了MCA和MAASFF两个模块，以解决复杂背景和多尺寸船舶目标的问题，从而增强了模型性能。本文方法框架如图2所示。

1.1 YOLOv5s作为基线

YOLOv5是一种高效的单阶段目标检测算法，相较于YOLOv4，它引入了一些新的改进，显著提升了检测速度和精度。尽管YOLO系列推出了更高版本，包括YOLOv6（Li等，2022）、YOLOv7（Wang等，2023）和YOLOv8（Ultralytics，2023），但鉴于YOLOv5在SAR船舶检测领域得到广泛应用，并且在性能和稳定性方面表现良好，本文选择YOLOv5作为基线。此外，考虑到YOLOv5包括4个模型：YOLOv5s、YOLOv5m、YOLOv5l和YOLOv5x，其中YOLOv5s是最小模型，具有最小的尺寸、计算量和参数量，因此选为本文的基线模型。

1.2 本文方法

本文方法由输入层、多粒度主干网络、多粒度颈部层、头部层和输出层构成，如图2所示。首先，将1幅SAR图像输入到融合了MCA模块的主干网络中，以获取复杂背景下全局和局部船舶目标的上下文信息，进而获得更具代表性的特征图。接着，将MAASFF模块融入特征金字塔网络，建立M-Neck，以增强多粒度上下文信息，提高特征金字塔对不同尺寸的船舶的表征。最后，通过Head输出最终的检测结果，其中分类和回归分别采用BCE（binary cross entropy）损失和CIoU（complete intersection over union）损失（Zheng等，2020）。Conv2d、BN（batch normalization）、SiLU（sigmoid linear unit）、MaxPool、Concat和add分别表示卷积、批次归一化、SiLU激活函数、最大池化、张量拼接和张量逐元素相加。

本文方法在YOLOv5s框架的基础上嵌入MCA和MAASFF模块，构建多粒度主干网络和多粒度颈部层。由于YOLOv5s是本文方法的基线，因此本文方法的检测层与YOLOv5s相同。

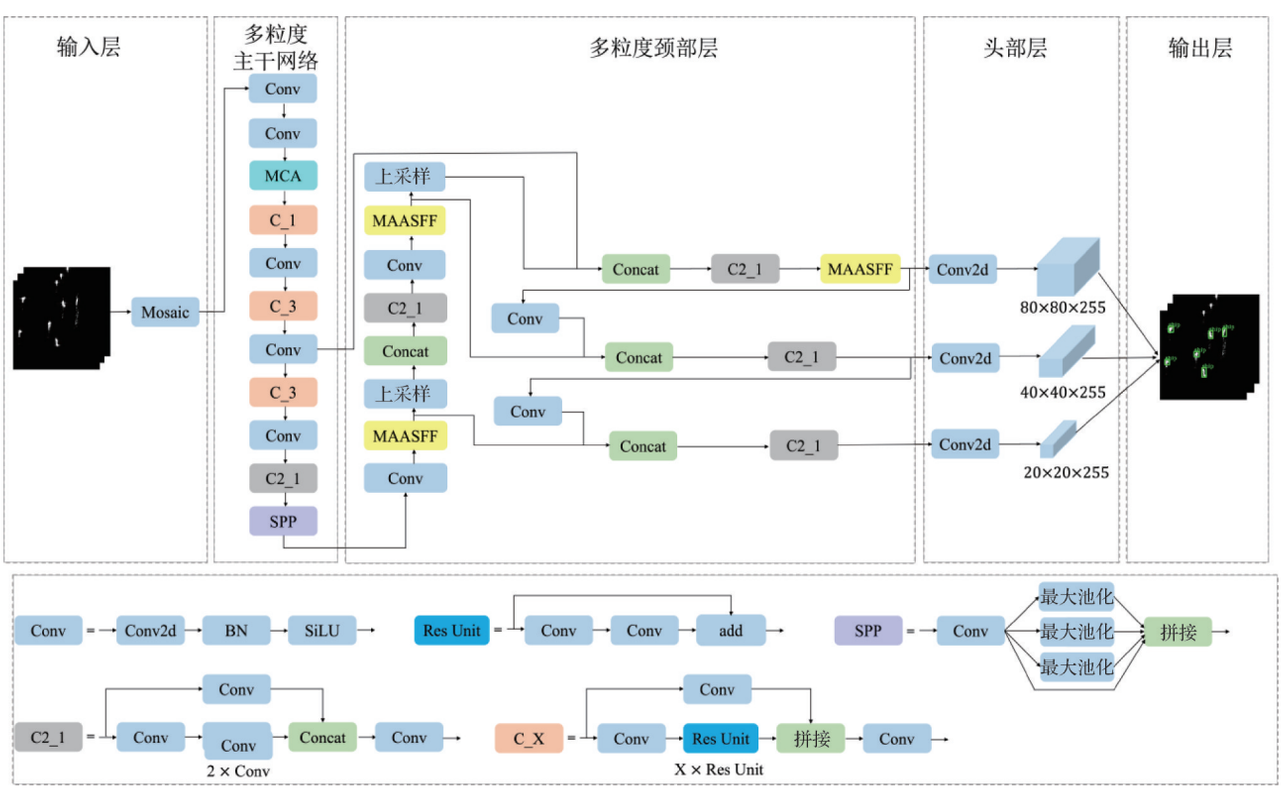


图2 本文方法的总体框架

Fig. 2 Overall framework of the proposed method

1.3 多粒度通道注意力(MCA)模块

SAR图像通常呈现复杂背景，这是由于SAR的独特工作原理、陆地和海洋等因素导致图像中存在大量斑点和杂波噪声，从而严重影响了SAR船舶检测的准确性。为了解决这一问题，本文引入了MCA模块，包括局部通道和全局通道注意力，如图3所示。该模块主要借鉴多粒度的思想，通过引入从全局到局部的通道注意力机制，以缓解SAR图像复杂背景对检测结果的影响。为了使MCA模块更加高效和轻量化，本文选择使用点卷积(Chollet，2017)替代传统卷积，以聚合全局和局部通道的上下文信息。点卷积不仅减少模型的参数量和计算成本，还增强模型对特征的捕捉能力，从而提高模型的性能。

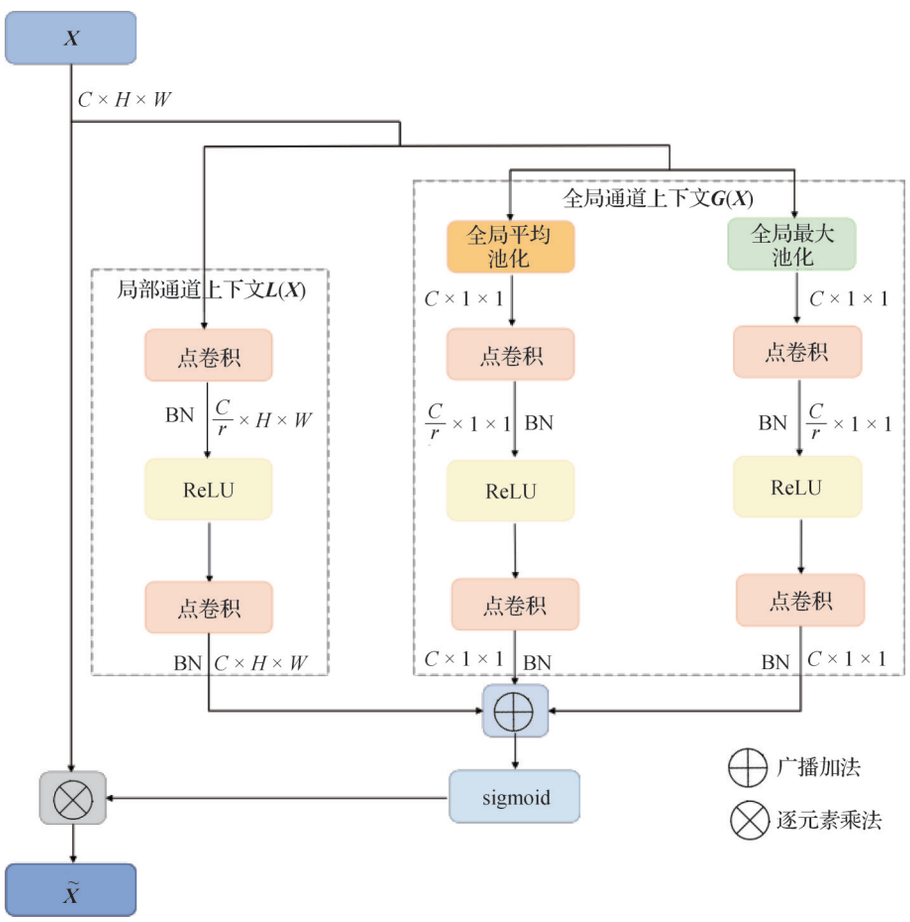


图3 MCA模块的结构

Fig. 3 Structure of MCA module

局部通道上下文$\boldsymbol{L}(\boldsymbol{X})\in\mathbf{R}^{C\times H\times W}$通过瓶颈结构计算，具体为

$$\boldsymbol{L}(\boldsymbol{X})=B(P_2(\delta(B(P_1(\boldsymbol{X})))))\tag{1}$$

式中，C，H，W分别表示通道数、高度和宽度。P_1和P_2分别表示核为$C/r\times C\times 1\times 1$和$C\times C/r\times 1\times 1$的点卷积。$r$表示缩减比($r=32$)，$B$和$\delta$分别表示批次归一化(BN)和ReLU(rectified linear unit)激活函数。全局通道上下文$\boldsymbol{G}(\boldsymbol{X})\in\mathbf{R}^{C\times H\times W}$通过瓶颈结构计算，具体为

$$\boldsymbol{G}_1(\boldsymbol{X})=B(P_2(\delta(B(P_1(GAP(\boldsymbol{X})))))) \tag{2}$$

$$\boldsymbol{G}_2(\boldsymbol{X})=B(P_2(\delta(B(P_1(GMP(\boldsymbol{X})))))) \tag{3}$$

式中，GAP和GMP分别表示全局平均池化和全局最大池化。P_1和P_2分别表示核为$C/r\times C\times 1\times 1$和$C\times C/r\times 1\times 1$的点卷积。$B$和$\delta$分别表示BN和ReLU激活函数。

因此，通过MCA模块获得的精细特征$\tilde{\boldsymbol{X}}\in\mathbf{R}^{C\times H\times W}$的过程为

$$\tilde{\boldsymbol{X}}=\boldsymbol{X}\otimes(\boldsymbol{L}(\boldsymbol{X})\oplus\boldsymbol{G}_1(\boldsymbol{X})\oplus\boldsymbol{G}_2(\boldsymbol{X}))\tag{4}$$

式中，$\oplus$和$\otimes$分别表示广播加法和逐元素乘法。

1.4 多粒度空洞自适应空间特征融合(MAASFF)模块

在SAR图像中，船舶目标呈现多样的尺寸，尤

其是小型船舶数量众多，这给船舶检测任务带来了巨大的挑战。因此，在SAR船舶检测中充分利用多粒度信息变得至关重要。为了在不引入过多参数的前提下充分利用多粒度信息，本文设计了MAASFF模块，如图4所示。该模块由多粒度卷积层和自适应空间特征融合层（Liu等，2019）构成。3×3、BN和ReLU分别表示核大小为3的空洞卷积、BN和ReLU激活函数。具体而言，多粒度卷积层由3个并行分支组成，每个分支使用相同大小的卷积核，但具有不同扩张率的空洞卷积（Chen等，2018）。这种设计通过引入不同扩张率的空洞卷积，有效扩大了感受野，使模型能够在不同空间粒度上捕捉更丰富的特征，从而增强了模型对不同尺寸船舶目标的特征表达能力。与传统卷积相比，空洞卷积在保持计算效率的同时提供了更大的感受野，有助于模型在不增加过多参数的情况下提升性能。

假设输入特征图为$\boldsymbol{Y}\in\mathbf{R}^{C\times H\times W}$，在这个模块中，首先使用3个空洞卷积分别对每个通道进行独立的卷积操作，然后生成3个特征图$\boldsymbol{Y}_0$、$\boldsymbol{Y}_1$和$\boldsymbol{Y}_2$作为输出。这些操作可以表示为

$$\boldsymbol{Y}_0=\delta(B(D_1(\boldsymbol{Y}))) \tag{5}$$

$$\boldsymbol{Y}_1=\delta(B(D_2(\boldsymbol{Y}))) \tag{6}$$

$$\boldsymbol{Y}_2=\delta(B(D_3(\boldsymbol{Y}))) \tag{7}$$

式中，D_1，D_2，D_3分别表示空洞卷积的核大小为3×3，扩张率为1，2，3。B和δ分别表示BN和ReLU激活函数。

此外，在融合不同粒度特征的过程中，本文采用了自适应空间特征融合的方法，以避免忽略不同尺寸目标特征之间的差异，并减少不必要的计算开销。这进一步提升了不同尺寸的SAR船舶检测性能，特别对小型船舶。具体而言，将输入特征图$\boldsymbol{Y}_0$，$\boldsymbol{Y}_1$，$\boldsymbol{Y}_2$分别设置为级别0、级别1和级别2。然后，通过学习每个输入层的空间特征权重来进行不同粒度特征图的融合计算，最终得到融合后的特征图$\boldsymbol{Y}'$。计算过程可写为

$$\boldsymbol{Y}'=\boldsymbol{\omega}^l\times\boldsymbol{Y}^{0\to l}+\boldsymbol{\mu}^l\times\boldsymbol{Y}^{1\to l}+\boldsymbol{\varsigma}^l\times\boldsymbol{Y}^{2\to l} \tag{8}$$

$$\boldsymbol{\omega}^l+\boldsymbol{\mu}^l+\boldsymbol{\varsigma}^l=1(\boldsymbol{\omega}^l,\boldsymbol{\mu}^l,\boldsymbol{\varsigma}^l\in[0,1]) \tag{9}$$

$$\boldsymbol{\omega}^l=\frac{\mathrm{e}^{\lambda_{\omega}^l}}{\mathrm{e}^{\lambda_{\omega}^l}+\mathrm{e}^{\lambda_{\mu}^l}+\mathrm{e}^{\lambda_{\varsigma}^l}} \tag{10}$$

式中，$\boldsymbol{\omega}^l$，$\boldsymbol{\mu}^l$，$\boldsymbol{\varsigma}^l$为来自不同粒度层的权重，$\boldsymbol{Y}^{0\to l}$，$\boldsymbol{Y}^{1\to l}$，$\boldsymbol{Y}^{2\to l}$为来自不同粒度层的输出。$\boldsymbol{\omega}^l$，$\boldsymbol{\mu}^l$，$\boldsymbol{\varsigma}^l$的和为1，并通过softmax压缩在[0，1]之间。参数λ_{ω}^l，λ_{μ}^l，λ_{ς}^l是通过在网络过程中使用反向传播方法计算softmax层的参数而获得。

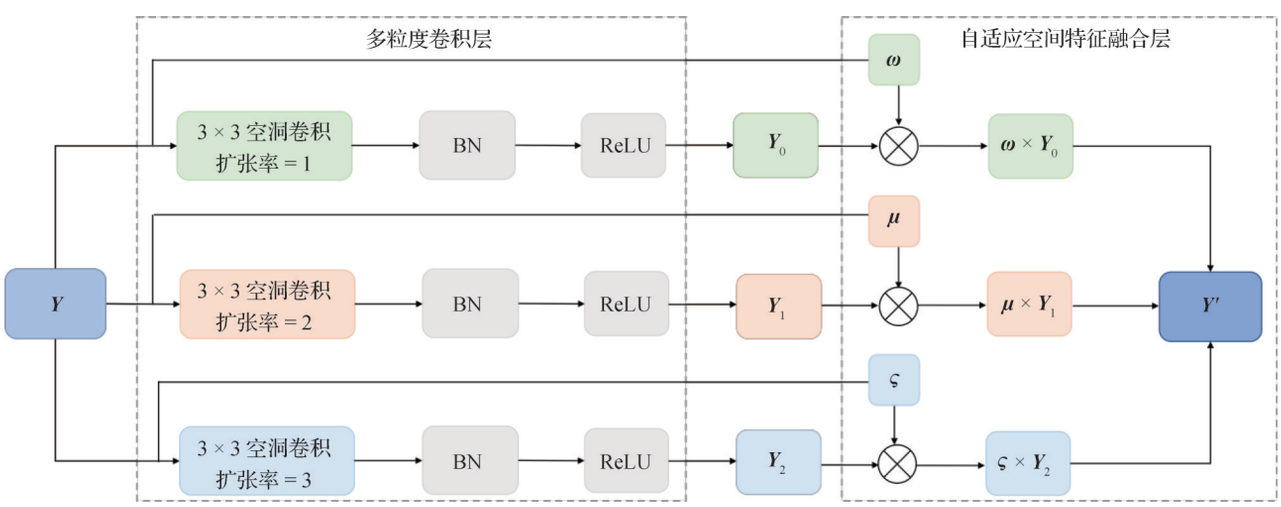


图4　MAASFF模块的结构

Fig. 4　Structure of MAASFF module

2　实　验

2.1　数据集

SAR-Ship-Dataset（Wang等，2019）是一个包含43 819幅船舶切片的数据集，来源于我国高分3号SAR数据和Sentinel-1 SAR数据。它的船舶切片种类丰富、背景多样，适用于多种SAR图像应用。为了进行模型的训练和评估，该数据集随机按照7∶2∶1的比例分为训练集、验证集和测试集，并将图像尺寸裁剪为256×256像素。

SSDD（Li等，2017）是一个用于SAR船舶检测的

数据集，包含1 160幅图像和2 464个船舶实例。这些图像的难度级别丰富，涵盖了各种海洋环境和不同季节的图像。测试集由编号最后数字为1和9的图像组成，其余被视为训练集，并调整图像尺寸为512×512像素。

2.2 实验设置和评价指标

本文在Ubuntu 16.04系统上使用PyTorch 1.7、CUDA 10.2、CUDNN 8.0.2和2个Tesla V100 GPUs（32 GB内存）进行实验，在SAR-Ship-Dataset和SSDD上进行所有消融实验和参数分析实验。由于SAR图像的独特性，模型从头开始训练且没有加载预训练权重。使用随机梯度下降（stochastic gradient descent，SGD）优化器，初始学习率为0.01，每次衰减时学习率除以10，批大小为16。动量和权重衰减分别为0.9和0.000 1。同时，根据实验分别在SAR-Ship-Dataset和SSDD上进行100和80个epoch的训练。

实验使用了AP_{50}、精确率、召回率、F1值、运行时间、模型参数量和每秒浮点运算次数（floating point operations per second，FLOPS）作为评价指标。其中，平均精度（average precision，AP）是精度—召回曲线下的面积，AP_{50}是交并比（intersection over union，IoU）阈值为0.5时的AP，精确率（precision，P）、召回率（recall，R）和F1值在置信度分数为0.3时进行计算。

2.3 参数分析

2.3.1 缩减比r对MCA模块的影响

在1.3节提出了MCA模块并引入一个超参数缩减比r。不同r值对于SAR船舶检测的性能有一定影响。为了确定最佳的r值，本文进行了一系列的实验，并记录了不同r值下的性能和参数量，实验结果如表1所示。为了更好地平衡参数量和性能，将r设为32以获得较好的性能，模型在SAR-Ship-Dataset和SSDD数据集上的AP_{50}分别达到95.4%和96.5%。

表1 比较不同的缩减比r对MCA模块的影响

Table 1 Comparing the effect of different reduction ratios r on the MCA module

r	AP_{50}/%		参数量/M
	SAR-Ship-Dataset	SSDD	
8	94.6	95.6	11.8
16	95.0	96.0	9.7
32	**95.4**	**96.5**	8.5
64	94.8	96.0	3.9

注：加粗字体表示各列最优结果。

2.3.2 不同分支数N对MAASFF模块的影响

在1.4节提出了一个MAASFF模块，其中多粒度层的分支数N对于检测性能有重要影响。本文根据实验来确定最佳的N值，以获得更好的检测结果。实验结果如表2所示，发现随着N的增加，模型参数量也不断增加，但性能表现出先增后减。为了在参数量和性能之间取得更好的平衡，将N设为3以实现相对较少的参数量和较好的性能表现。该模型的AP_{50}在SAR-Ship-Dataset和SSDD数据集上分别达到95.0%和96.2%。

表2 比较不同的分支数N对MAASFF模块的影响

Table 2 Comparing the effect of different branch numbers N on the MAASFF module

N	空洞卷积核大小	AP_{50}/%		参数量/M
		SAR-Ship-Dataset	SSDD	
1	3×3，扩张率=1	94.3	95.5	7.4
2	3×3，扩张率=1 3×3，扩张率=2	94.4	95.7	7.6
3	3×3，扩张率=1 3×3，扩张率=2 3×3，扩张率=3	**95.0**	**96.2**	8.0
4	3×3，扩张率=1 3×3，扩张率=2 3×3，扩张率=3 3×3，扩张率=4	95.0	96.0	8.5
5	3×3，扩张率=1 3×3，扩张率=2 3×3，扩张率=3 3×3，扩张率=4 3×3，扩张率=5	94.3	95.5	9.3

注：加粗字体表示各列最优结果。

2.4 消融实验

本节通过消融实验探究MCA和MAASFF两个模块的有效性，实验结果如表3所示。

1）基线。相比双阶段检测器，YOLOv5s不需要预定义锚点，能够快速、精确地识别目标。如表3所示，YOLOv5s在SAR-Ship-Dataset和SSDD数据集上的AP_{50}分别为94.3%和95.4%，达到了竞争性能

水平。

2）MCA。相对于基线，在SAR-Ship-Dataset上，MCA模块可以将AP_{50}提高1.1%，从94.3%提高到95.4%。同时，在SSDD上，MCA模块可以将AP_{50}提高1.1%。该模块包括全局通道注意力和局部通道注意力两个关键设计。全局通道注意力可以让网络更加注重全局的空间信息，提高全局位置信息的捕捉能力；而局部通道注意力则可以让网络更加注重与船舶目标相关的局部空间位置信息，提高细节信息的提取能力，从而有效减少了复杂背景的干扰，提高了检测准确率。

3）MAASFF。相对于基线，在SAR-Ship-Dataset上，MAASFF模块可以将AP_{50}提高0.7%，达到了95.0%。此外，在SSDD上，MAASFF模块可以将AP_{50}提高0.8%，达到了96.2%。这说明该模块可以有效地增强不同粒度层的上下文信息。具体而言，MAASFF模块能够在捕获不同粒度层的特征时增加上下文信息的融合，从而更好地处理小型船舶的检测，进一步提高了船舶检测的准确性。

将MCA和MAASFF两个模块结合后，模型在SAR-Ship-Dataset和SSDD数据集上的AP_{50}分别达到96.1%和97.0%，取得最佳的检测结果。这意味着，这两个模块不仅单独应用时有效提高船舶检测的性能，而且二者的结合能进一步提升模型性能。

表3　提出的不同模块在SAR-Ship-Dataset和SSDD上的实验结果

Table 3　Experimental results of different proposed modules on SAR-Ship-Dataset and SSDD

基线	MCA	MAASFF	SAR-Ship-Dataset		SSDD	
			AP_{50}/%	运行时间/ms	AP_{50}/%	运行时间/ms
√	–	–	94.3	**3.5**	95.4	**6.7**
√	√	–	95.4	4.1	96.5	7.8
√	–	√	95.0	3.7	96.2	7.2
√	√	√	**96.1**	4.4	**97.0**	8.5

注：加粗字体表示各列最优结果，“√”表示采用，“–”表示未采用。

2.5　与目前先进方法的对比

本节在SAR-Ship-Dataset和SSDD两个数据集上验证了本文方法的有效性和泛化能力，并将其与目前先进的方法进行了对比，结果如表4和表5所示。

1）SAR-Ship-Dataset。在SAR-Ship-Dataset上将本文方法与其他船舶检测方法进行对比，实验结果如表4所示。与双阶段方法Faster R-CNN、DAPN、CR2A-Net、KCPNet和BL-Net相比，本文方法可以将模型的检测精度AP_{50}提高约1.9%～6.0%。对于常见单阶段的算法YOLOv4、CenterNet++、CRDet和YOLOv5s，本文方法AP_{50}分别可以提高2.9%、1.2%、0.9%和1.8%。实验结果表明，本文方法在SAR-Ship-Dataset上达到最佳性能，AP_{50}达到96.1%，优于所有比较的方法。同时，它的速度保持在与基线YOLOv5s相近的水平。

2）SSDD。根据表5的结果，本文方法在与其他经典算法和改进方法的比较中表现出色，具有一定的竞争力和优越性。相较于这些算法，本文方法以与基线YOLOv5s相近的速度将AP_{50}分别提高了8.7%（Faster R-CNN）、6.9%（DAPN）、7.2%（CR2A-Net）、4.5%（KCPNet）、1.8%（BL-Net）、0.9%（YOLOv4）、4.3%（CenterNet++）、0.6%（CRDet）和1.6%（YOLOv5s）。这些结果表明，本文方法取得显著的性能提升，AP_{50}达到97.0%的最佳性能。

2.6　可视化结果及分析

为了更好地展示本文方法的优越性和泛化性，选取了SAR-Ship-Dataset上的可视化结果。图5展示了在复杂背景下5种不同方法的检测结果。其中，YOLOv5s的漏检最严重，DAPN的虚警最严重，其他方法都有不同程度的虚警，而本文方法可以准确地检测出复杂背景下的所有船舶。图6展示了5种不同方法对小型船舶目标的检测结果，其中，YOLOv5s和DAPN对于小型船舶存在较多的漏检和虚警，CenterNet++和CRDet对于小型船舶存在较少的虚警和漏检，而本文方法对小型船舶实现了精确的定位。

为了更加生动地展示本文方法的优越性，图7和图8分别展示了本文方法在不同数据集上的检测

表 4　不同方法在SAR-Ship-Dataset上评估指标的比较

Table 4　Comparison of evaluation metrics for different methods on the SAR-Ship-Dataset

	方法	主干网络	精确率/%	召回率/%	F1/%	AP_{50}/%	运行时间/ms	参数量/M	FLOPs/G
双阶段	Faster R-CNN	ResNet101	91.0	91.0	91.0	91.0	25.1	60.1	31.1
	DAPN	ResNet101	91.0	91.4	91.2	91.9	27.8	63.8	39.7
	CR2A-Net	ResNet101	91.7	92.2	91.9	90.1	41.7	88.6	59.7
	KCPNet	ResNet101	89.9	90.7	90.2	91.2	29.4	67.1	48.5
	BL-Net	ResNet101	94.0	95.8	94.9	94.2	139.5	47.8	151.6
单阶段	YOLOv4	CSPDarknet53	85.7	92.7	89.1	93.2	6.9	64.3	11.3
	CenterNet++	DAL34	85.4	93.5	89.3	94.9	13.8	20.3	6.7
	CRDet	ResNet101	93.8	94.4	94.1	95.2	22.1	73.5	13.8
	YOLOv5s	CSPDarknet53	92.1	93.0	93.0	94.3	**3.5**	**7.3**	**4.4**
	本文	CSPDarknet53	**94.1**	**94.9**	**94.5**	**96.1**	4.4	9.2	5.9

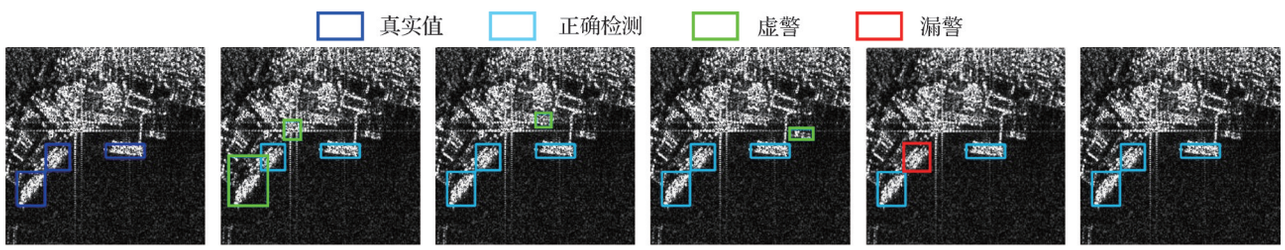
注：加粗字体表示各列最优结果。

表 5　不同方法在SSDD上评估指标的比较

Table 5　Comparison of evaluation metrics for different methods on the SSDD

	方法	主干网络	精确率/%	召回率/%	F1/%	AP_{50}/%	运行时间/ms	参数量/M	FLOPs/G
双阶段	Faster R-CNN	ResNet101	90.9	87.6	89.2	88.3	30.2	60.1	82.7
	DAPN	ResNet101	87.6	91.4	89.5	90.1	34.5	63.8	117.2
	CR2A-Net	ResNet101	94.0	87.8	90.8	89.8	67.2	88.6	112.0
	KCPNet	ResNet101	90.4	92.7	91.5	92.5	34.6	67.1	121.4
	BL-Net	ResNet50	91.2	96.1	93.6	95.2	190.4	47.8	417.8
单阶段	YOLOv4	CSPDarknet53	93.6	94.0	93.8	96.1	12.9	64.3	45.3
	CenterNet++	DAL34	92.6	94.5	93.5	92.7	21.5	20.3	26.6
	CRDet	ResNet101	95.1	94.5	94.8	96.4	24.4	73.5	54.9
	YOLOv5s	CSPDarknet53	92.9	93.5	93.2	95.4	**6.7**	**7.3**	**10.7**
	本文	CSPDarknet53	**95.3**	**95.7**	**95.5**	**97.0**	8.5	9.2	16.9

注：加粗字体表示各列最优结果。

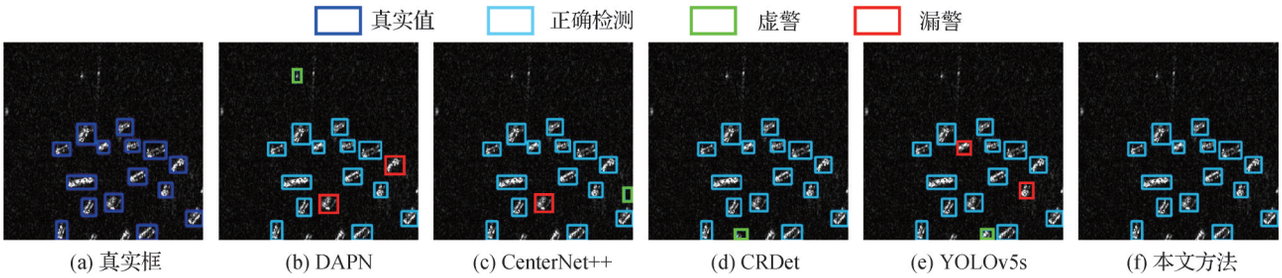


(a) 真实框　(b) DAPN　(c) CenterNet++　(d) CRDet　(e) YOLOv5s　(f) 本文方法

图 5　不同方法在复杂背景下的检测结果对比

Fig. 5　Comparison of detection results using different methods in complex backgrounds

((a) ground truth; (b) DAPN; (c) CenterNet++; (d) CRDet; (e)YOLOv5s; (f) ours)



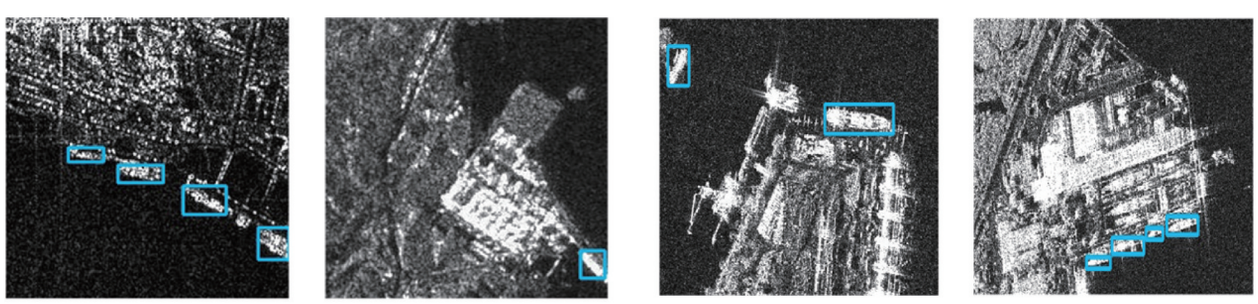
(a) 真实框　(b) DAPN　(c) CenterNet++　(d) CRDet　(e) YOLOv5s　(f) 本文方法

图6　不同方法在小型船舶中的检测结果对比

Fig. 6　Comparison of detection results using different methods for small ship

((a) ground truth; (b) DAPN; (c) CenterNet++; (d) CRDet; (e)YOLOv5s; (f) ours)

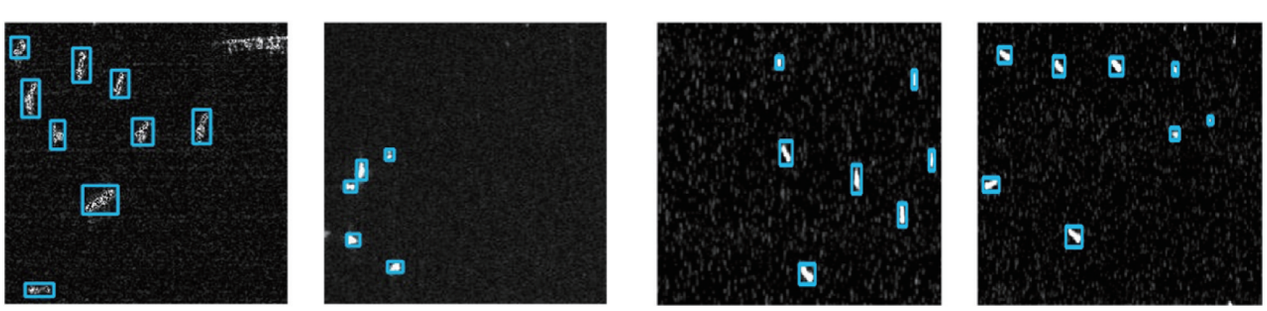
效果。其中，图7(a)(b)分别表示在复杂背景下SAR-Ship-Dataset和SSDD的检测结果。由于本文设计的MCA可以缓解图像中存在复杂背景的问题，实现了复杂图像下的船舶的精确定位和识别。图8(a)(b)分别表示SAR-Ship-Dataset和SSDD的多尺寸船舶检测的结果，尤其是小型船舶。通过图8可以看出，本文方法通过对捕捉的多粒度的上下文信息进行增强，以取得令人满意的检测结果。总而言之，可视化检测结果直观地反映了本文方法在应对存在复杂背景和多尺寸船舶目标(尤其是小型船舶)这两个挑战方面表现出了很好的效果，进一步说明了本文方法是一个非常有潜力的检测模型。



(a) SAR-Ship-Dataset　(b) SSDD

图7　本文方法在各种SAR数据集的复杂背景下的检测结果

Fig. 7　The detection results of our method in the complex background of various SAR datasets ((a) SAR-Ship-Dataset; (b) SSDD)



(a) SAR-Ship-Dataset　(b) SSDD

图8　本文方法在两个不同的SAR数据集上检测不同尺寸船舶的一些结果(特别是小型船舶)

Fig. 8　Some results of our method for detecting ships of different sizes on two different SAR datasets (especially small ships)

((a) SAR-Ship-Dataset; (b) SSDD)

3　结　论

本文提出了一种适用于SAR船舶检测的多粒度上下文网络，旨在解决SAR图像中存在的复杂背景干扰和船舶尺寸大小不一问题。该方法融合了多粒度思想，设计了MCA和MAASFF两个模块。一方面，将MCA模块嵌入到主干网络中，利用全局和局

部通道注意力来捕捉从细到粗的上下文信息，增强特征的表征能力。这有助于提高模型对复杂背景干扰的鲁棒性，进而提高检测的精度。另一方面，利用MAASFF模块中多粒度空洞卷积层和自适应空间特征融合层来增强模型对不同粒度特征的提取能力。这使得模型更好地处理SAR图像中不同尺寸的船舶，并提升了对小型船舶的检测能力，从而进一步提高检测的准确性和稳定性。从大量的实验结果来看，与先进方法相比，本文方法在SAR船舶检测方面取得了更好的性能。在未来的工作中，考虑将本研究中的模型思路扩展到其他YOLO系列模型，从而进一步提高SAR船舶检测性能，以满足更广泛的场景和需求。

参考文献（References）

Bochkovskiy A，Wang C Y and Liao H Y M. 2020. YOLOv4：optimal speed and accuracy of object detection［EB/OL］.［2023-12-12］. https://arxiv.org/pdf/2004.10934v1.pdf

Chen L C，Papandreou G，Kokkinos I，Murphy K and Yuille A L. 2018. DeepLab：semantic image segmentation with deep convolutional nets，atrous convolution，and fully connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence，40(4)：834-848［DOI：10.1109/TPAMI.2017.2699184］

Chollet F. 2017. Xception：deep learning with depthwise separable convolutions//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu，USA：IEEE：1800-1807［DOI：10.1109/CVPR.2017.195］

Cui Z Y，Li Q，Cao Z J and Liu N Y. 2019. Dense attention pyramid networks for multi-scale ship detection in SAR images. IEEE Transactions on Geoscience and Remote Sensing，57(11)：8983-8997［DOI：10.1109/TGRS.2019.2923988］

Gao G and Shi G T. 2017. CFAR ship detection in nonhomogeneous sea clutter using polarimetric SAR data based on the notch filter. IEEE Transactions on Geoscience and Remote Sensing，55(8)：4811-4824［DOI：10.1109/TGRS.2017.2701813］

Girshick R，Donahue J，Darrell T and Malik J. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus，USA：IEEE：580-587［DOI：10.1109/CVPR.2014.81］

Gong S R，Xu S J，Zhou L F，Zhu J and Zhong S. 2022. Deformable atrous convolution nearshore SAR small ship detection incorporating mixed attention. Journal of Image and Graphics，27(12)：3663-3676（龚声蓉，徐少杰，周立凡，朱杰，钟珊. 2022. 融入混合注意力的可变形空洞卷积近岸SAR小舰船检测. 中国图象图形学报，27(12)：3663-3676）［DOI：10.11834/jig.210866］

Guo H Y，Yang X，Wang N N and Gao X B. 2021. A CenterNet++ model for ship detection in SAR images. Pattern Recognition，112：#107787［DOI：10.1016/j.patcog.2020.107787］

Han Y Q，Liao J W，Lu T S，Pu T and Peng Z M. 2023. KCPNet：knowledge-driven context perception networks for ship detection in infrared imagery. IEEE Transactions on Geoscience and Remote Sensing，61：#5000219［DOI：10.1109/TGRS.2022.3233401］

Jocher G. 2021. YOLOv5［EB/OL］.［2023-12-12］. https://github.com/ultralytics/yolov5

Li C Y，Li L L，Jiang H L and Weng K H. 2022. YOLOv6：a single-stage object detection framework for industrial applications［EB/OL］.［2023-12-12］. https://arxiv.org/pdf/2209.02976v1.pdf

Li J W，Qu C W and Shao J Q. 2017. Ship detection in SAR images based on an improved faster R-CNN//2017 SAR in Big Data Era：Models，Methods and Applications. Beijing，China：IEEE：1-6［DOI：10.1109/BIGSARDATA.2017.8124934］

Liu S T，Huang D and Wang Y H. 2019. Learning spatial fusion for single-shot object detection［EB/OL］.［2023-12-12］. https://arxiv.org/pdf/1911.09516.pdf

Liu W，Anguelov D，Erhan D，Szegedy C，Reed S，Fu C Y and Berg A C. 2016. SSD：single shot multibox detector//Proceedings of the 14th European Conference on Computer Vision. Amsterdam，the Netherlands：Springer：21-37［DOI：10.1007/978-3-319-46448-0_2］

Oh J，Youm G Y and Kim M. 2021. SPAM-net：a CNN-based SAR target recognition network with pose angle marginalization learning. IEEE Transactions on Circuits and Systems for Video Technology，31(2)：701-714［DOI：10.1109/TCSVT.2020.2987346］

Ren S Q，He K M，Girshick R and Sun J. 2017. Faster R-CNN：towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence，39(6)：1137-1149［DOI：10.1109/TPAMI.2016.2577031］

Ruan C，Guo H and An J B. 2021. SAR inshore ship detection algorithm in complex background. Journal of Image and Graphics，26(5)：1058-1066（阮晨，郭浩，安居白. 2021. 复杂背景下SAR近岸舰船检测. 中国图象图形学报，26(5)：1058-1066）［DOI：10.11834/jig.200266］

Ultralytics. 2023. YOLOv8［EB/OL］.［2023-12-12］. https://github.com/ultralytics/ultralytics

Wang C Y，Bochkovskiy A and Liao H Y M. 2023. YOLOv7：trainable bag-of-freebies sets new state-of-the-art for real-time object detectors//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver，Canada：IEEE：7464-7475［DOI：10.1109/CVPR52729.2023.00721］

Wang Y Y，Wang C，Zhang H，Dong Y B and Wei S S. 2019. A SAR dataset of ship detection for deep learning under complex backgrounds. Remote Sensing，11(7)：#765［DOI：10.3390/rs11070765］

Yang X，Zhang X，Wang N N and Gao X B. 2022. A robust one-stage

detector for multiscale ship detection with complex background in massive SAR images. IEEE Transactions on Geoscience and Remote Sensing, 60: #5217712 [DOI: 10.1109/TGRS.2021.3128060]

Yu Y, Yang X, Li J and Gao X B. 2022. A cascade rotated anchor-aided detector for ship detection in remote sensing images. IEEE Transactions on Geoscience and Remote Sensing, 60: #5600514 [DOI: 10.1109/TGRS.2020.3040273]

Zhang T W, Zhang X L, Liu C, Shi J, Wei S J, Ahmad I, Zhan X, Zhou Y, Pan D C, Li J W and Su H. 2021. Balance learning for ship detection from synthetic aperture radar remote sensing imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 182: 190-207 [DOI: 10.1016/j.isprsjprs.2021.10.010]

Zheng Z H, Wang P, Liu W, Li J Z, Ye R G and Ren D W. 2020. Distance-IoU loss: faster and better learning for bounding box regression//Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York, USA: AAAI, 12993-13000 [DOI: 10.1609/aaai.v34i07.6999]

作者简介

应丽,女,博士研究生,主要研究方向为计算机视觉和深度学习。E-mail:1910663@tongji.edu.cn

张志飞,通信作者,男,副研究员,博士生导师,主要研究方向为模式识别和大数据分析。E-mail:zhifeizhang@tongji.edu.cn

苗夺谦,男,教授,博士生导师,主要研究方向为机器学习、数据挖掘、大数据分析、粒计算、人工智能和文本图像处理。E-mail:dqmiao@tongji.edu.cn

赵才荣,男,教授,博士生导师,主要研究方向为人工智能、计算机视觉、行人再识别、场景文字检测与识别、隐私保护与安全防御、垂直领域基础模型。E-mail:zhaocairong@tongji.edu.cn