

# 基于高效聚合特征增强网络的合成孔径 雷达船舰检测方法

单慧琳<sup>1,2</sup>, 刘文星<sup>1</sup>, 王兴涛<sup>1</sup>, 付相为<sup>1</sup>, 李长帅<sup>2</sup>, 张银胜<sup>1,2\*</sup>

<sup>1</sup>南京信息工程大学电子与信息工程学院, 江苏 南京 210044;

<sup>2</sup>无锡学院电子信息工程学院, 江苏 无锡 214105

**摘要** 合成孔径雷达(SAR)由于散射效应以及波长和天线尺寸的分辨率限制,难以获取小尺寸目标的细节和边界信息,因此,检测准确性不高。为了提高SAR船舰检测的准确率以及降低误检率,提出了一种基于高效聚合特征增强网络的SAR船舰检测方法。首先,在主干网络中采用空间通道注意力机制,构建出高效层卷积块作为主要的特征提取模块,以增强模型的特征获取性能,提高模型对船舰目标的识别能力;其次,特征融合部分采用Inception NeXt模块来提高算法效率;最后,在主干网络以及特征提取部分之间构建出一种全局增强特征金字塔分支结构,进一步融合全局特征,避免传输过程中的低维度特征损失,以提升网络的特征融合能力,使其即使对于复杂背景下的小目标仍然能展现出可靠的检测能力。为了证明所提网络的有效性,在SSDD数据集上作了对比实验,实验结果表明,相较于YOLOv7,所提网络的准确率提升了2.5个百分点,召回率提升了9.2个百分点,交并比(IoU)阈值为0.5时的平均精度提升了6.4个百分点,IoU为0.5:0.95时的平均精度提升了9.9个百分点。实验结果证明,所提网络在提升SAR船舰检测精度、改善误检漏检等方面有显著优势,可作为高精度的检测方法来有效应对SAR船舰检测中存在的问题。

**关键词** 深度学习; 目标检测; 高效聚合特征增强网络; 注意力机制; 合成孔径雷达船舰检测

中图分类号 TN959.1+7

文献标志码 A

DOI: 10.3788/AOS231285

## 1 引言

合成孔径雷达(SAR)是利用合成孔径原理实现高分辨的微波成像雷达,具备全天时、全天候、高分辨、大幅宽等多种特点。SAR船舰目标检测技术可以检测海上船舰以及设备,在石油、港口管理、海洋资源开发和海洋科学研究等领域应用广泛,同时海上监测、边防巡逻、海上救援和海上通道的安全保障等方面具有重要的战略意义。因此,利用SAR图像进行舰船检测已成为国内外研究的热点<sup>[1-2]</sup>。

传统的SAR图像目标检测方法包括基于纹理分析法、基于极化特性法、恒定虚警率检测算法(CFAR)等,其中,使用最为广泛的是CFAR<sup>[3]</sup>。虽然该算法在速度上有一定的优势,但该算法的计算复杂度较高,易受复杂背景的影响,导致检测效率及精度低。在实际SAR图像成像过程中,SAR图像背景大多为港口、岛礁及其他建筑,这些背景具有很高的灰度特性和较强的混淆性,因此,对于海上船舰目标检测,存在背景复杂、船舰不规则排列、相似目标误检等不确定因素,为

SAR船舰检测带来了巨大的挑战与困难<sup>[4]</sup>。

随着深度学习的快速发展与成熟,深度学习已经被广泛应用于目标检测当中,基于深度学习的卷积神经网络(CNN)已经成为SAR目标检测算法的重要技术。目前,基于深度学习的目标检测方法分为两大类:双阶段检测模型和单阶段检测模型。双阶段检测模型主要代表有Faster R-CNN<sup>[5]</sup>、Mask R-CNN<sup>[6]</sup>和DETR(end-to-end object detection with Transformers)<sup>[7]</sup>等,这些模型先生成候选区域,再对候选区域进行分类定位。双阶段检测模型具有检测精度高、检测效果良好的优点,但由于双阶段检测模型将候选区域生成与目标分类分开,因此其计算量大、检测速度慢。单阶段检测模型主要代表有SSD(single shot MultiBox detector)<sup>[8]</sup>、YOLO(you only look once)系列<sup>[9-11]</sup>、CenterNet<sup>[12]</sup>、RetinaNet<sup>[13]</sup>等,其主要思路是将目标类型和位置均看作回归问题一次处理,简化了计算流程。其中,YOLO模型具有实时性好和准确性高的特点,可以提高船舶检测精度。2018年4月,YOLOv3模型被提出<sup>[14]</sup>,模型中增加了特征金字塔(FPN)和空间池

收稿日期: 2023-07-18; 修回日期: 2023-08-25; 录用日期: 2023-10-10; 网络首发日期: 2023-10-23

基金项目: 国家自然科学基金(62071240, 62106111)、江苏省一流本科课程(2021YLKC005)、江苏省产教融合型一流课程(2022-133)

通信作者: \*yorkzhang@nuist.edu.cn

化金字塔(SPP)模块,提高了模型对不同尺度目标的检测能力。2022年,Glenn<sup>[15]</sup>在发布的YOLOv5网络模型中引入了切片操作。同年,YOLOv7<sup>[16]</sup>模型被提出,模型引入了多分支堆叠模块、创新的过渡模块和重参数化结构等策略,进一步优化了目标检测效果。

由于近年来海上SAR检测技术的应用越来越广泛,因此,为了提高SAR检测船舰的效率,许多研究人员提出了改进方法。例如:赵鹏飞等<sup>[17]</sup>通过在浅层卷积层中融入ECA(efficient channel attention,起到丰富语义信息的作用),可以增强模型检测小目标的准确率。Zhu等<sup>[18]</sup>则运用重复双边特征金字塔网络(DB-FPN)来增强空间细节和语义细节的融合效果。林志康等<sup>[19]</sup>运用对数双曲余弦比构造差异图,有效提升了图像中产生变化区域与背景的对比度。武钰清等<sup>[20]</sup>利用插值算子与双域滤波处理检测来检测SAR图像变化,通过调整融合系数获得了算法性能的提升。杜艳玲等<sup>[21]</sup>改进了全CNN,利用不同层级卷积层融合策略来提高SAR检测精度。Rostami等<sup>[22]</sup>提出,通过学习共享不变的跨域嵌入空间,电光域图像能够用于训练SAR目标检测模型。Raj等<sup>[23]</sup>提出了一种具有额外预处理输入阶段的深度学习分类模型,以减少错误分类。但是,上述方法的准确率在复杂背景下偏低,并存在漏检误检严重、适应性弱等问题。

因此,本文提出一种高效聚合特征增强网络(EAFENet),用于解决当前SAR船舰检测所存在的普遍问题。首先,构建高效层卷积块(EL-CB)来作为主干网络的主要特征提取模块,以捕获船舰的几何特征信息,使网络更加注重细节特征;并通过多残差级联层获取不同尺度信息,同时进行特征融合,从而缓解梯度消失以及浅层特征丢失等问题,使网络更好地关注船舰位置,提高检测准确率。其次,在特征提取网络部分引入Inception NeXt模块,其并行的四分支结构可以有效提高网络对上下文的感知能力,进一步增强网络对目标的关注能力,提高特征提取网络的效率。最后,引入全局增强特征金字塔网络(GE-FPN)结构,通过特征增强和跳跃连接等方式减少主干网络与特征提取网络传输过程中的信息丢失,并提高网络的稳定性;通过对主干网络浅层信息与特征提取网络部分的融合,使主干网络的EL-CB以及特征提取网络的Inception NeXt模块进行更加高效的协同作用,进而使网络具有更好的适应性,对船舰目标的捕捉更加敏感。在SSDD数据集上对所提网络进行了效果验证,实验结果表明,EAFENet模型在多种背景下检测SAR船舰图像时均有精准稳定的检测效果。

## 2 EAFENet模型

EAFENet结构主要由输入端(input)、主干提取网络(backbone)、特征提取网络(neck)和探测头(head)

组成。输入端的主要作用是将需要手动处理的图像通过Mosaic数据增强以及自适应计算锚框和自适应图片缩放等预处理操作裁剪成统一尺寸,使图像满足主干特征提取网络的输入要求。然后主干网络与特征提取网络协作进行特征融合处理,输出的图像作为探测头的输入。探测头由大、中、小三种目标尺寸的IDetect检测头以及Rep Conv组成。再通过CIoU(complete intersection over union)损失函数和非线性极大值抑制处理,得到最终的预测结果。EAFENet首先在输入端将SAR船舰图像传输至主干网络中,然后经过主干网络的CBS(convolution+batch normalization+SiLU)模块和EL-CB处理,再与特征提取网络部分所提取到的特征进行融合。其中,CBS模块作为网络的主要构成模块,主要由卷积层、卷积层归一化(BN)以及SiLU激活函数组成。特征提取网络部分采用了性能高效的SPPCSPC(spatial pyramid pooling convolutional neural network and cascading spatial pooling convolutional neural network)作为池化金字塔,Inception NeXt作为特征提取模块,并使用跳跃连接的方式增加结构中FPN的特征融合能力。EAFENet结构如图1所示,EAFENet FPN模型如图2所示。

### 2.1 EL-CB

高效层聚合网络(ELAN)在主干网络中起到特征提取、增强网络泛化能力的作用。微波反射和散射特性导致SAR图像产生几何失真,使图像尺寸差异较大,据此,加入通道空间注意力机制(CBAM)<sup>[24]</sup>,使网络具有较强的特征提取和细节关注能力,可以高效地应对图像中不同目标形状和大小差异较大的问题,以获得更好的检测效果和性能。图3为CBAM示意图,其中:Maxpool与Avgpool分别为最大池化和平均池化;Sig为Sigmoid激活函数;convolution scaling为卷积缩放; $\oplus$ 为特征融合; $\otimes$ 为特征数值相乘。

EL-CB的示意图如图4所示,EL-CB用leaky ReLU<sup>[25]</sup>作为CBAM的激活函数,多个CBS组成的多残差级联层不仅避免了神经元死亡的现象,而且能有效解决梯度消失和梯度爆炸的问题,增强特征的传递和重用。CBS使用BN与SiLU激活函数,使模型具有更好的平滑性和非线性表达能力。在逐渐深入到更高层次的特征表示的末端,以leaky ReLU作为激活函数的CBAM能够全面准确地对高级特征进行加权和筛选,在特征融合的过程中起到平衡和优化的作用,提高模型对重要特征的捕捉能力,同时可以对前面提取的特征进行精细调整,避免对所有特征图都进行全局的注意力加权,从而减少计算负载。利用多尺度特征的信息融合主支路不同层级的特征输出,有效提升了不同层级特征的互补性。这种特征融合能够增大模型的感受野,提升模型的表达能力。

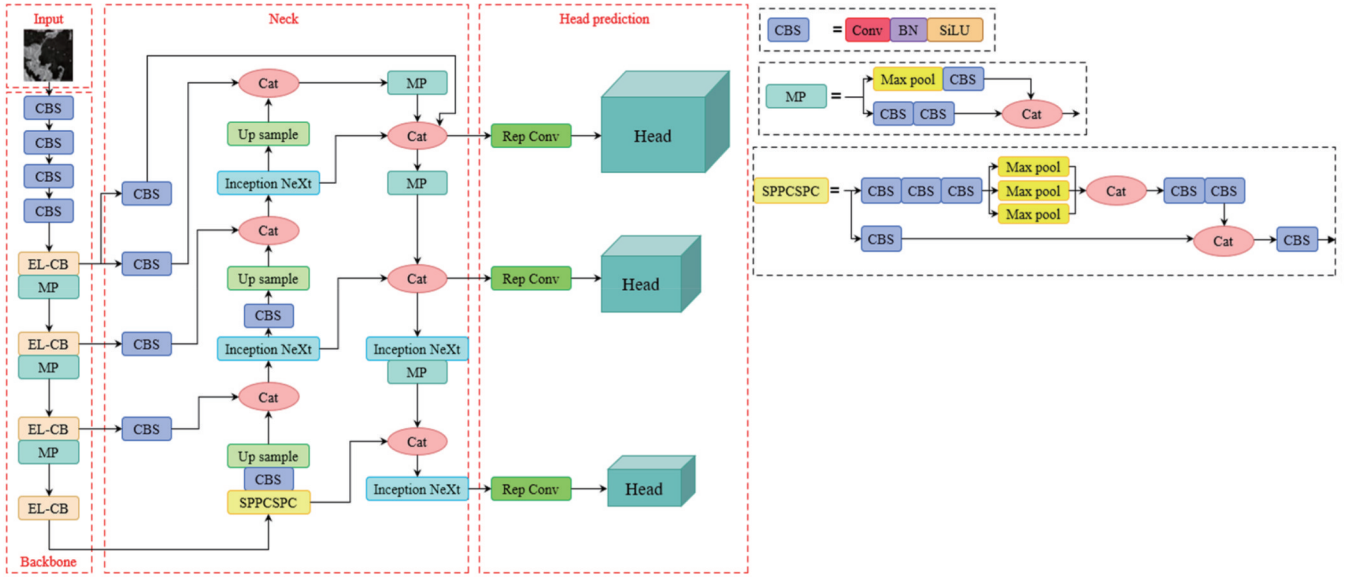


图 1 EAFENet 结构  
Fig. 1 EAFENet structure

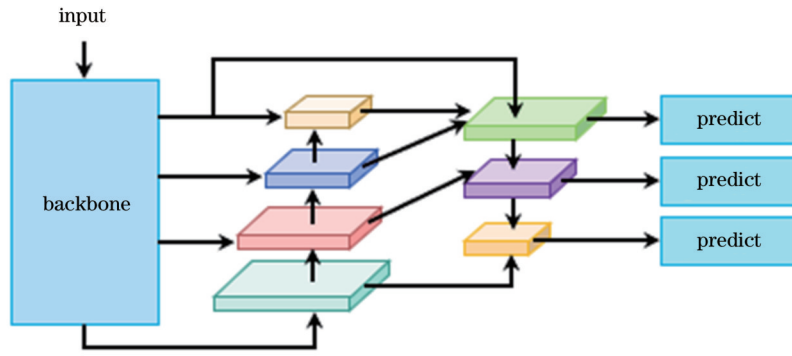


图 2 EAFENet FPN 模型  
Fig. 2 EAFENet FPN model

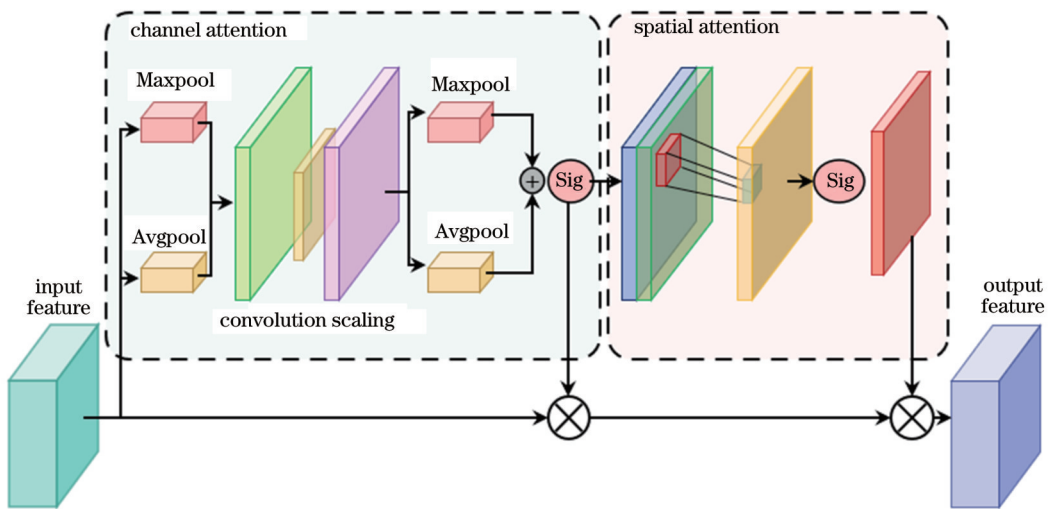


图 3 CBAM 示意图  
Fig. 3 Schematic diagram of CBAM

### 2.2 Inception NeXt 模块

在 SAR 舰船检测中,需要对图像进行高效而准确

的特征提取,以便从岸边、海岛、排列有模糊不清物体等各种复杂测量场景中分离出舰船目标。Inception

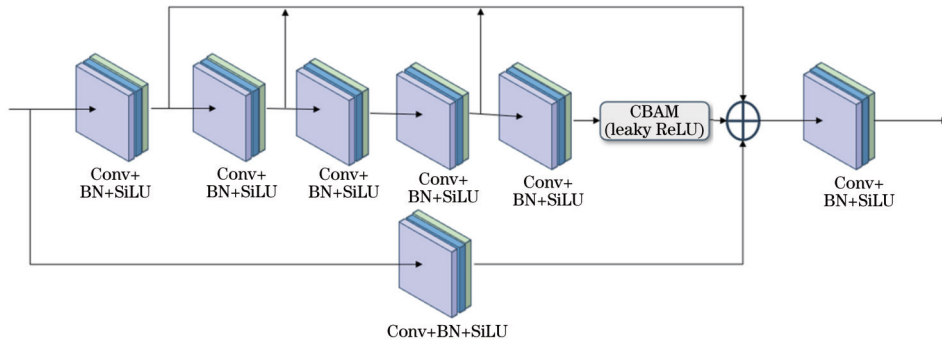


图 4 EL-CB 示意图

Fig. 4 Schematic diagram of EL-CB

NeXt 模块具有多个分支和密集连接的特点,使用的深度可分离卷积(DWConv)可以在有效控制计算量的同时,提取更为丰富和准确的特征。利用多尺度感受野、并行操作以及特征交互等技术,Inception NeXt 模块能够高效地处理大量的波长敏感信息。此外,该模块利用多路特征提取,使模型能够更好地理解输入图像中的多尺度信息,从而提高模型对细节和上下文的感知能力。采用多路特征信息融合可以提高模型对全局和

局部特征的识别能力,实现对图像特征更全面、更多尺度的感知和理解,进一步提升模型对复杂场景的理解能力。

在 neck 部分采用 Inception NeXt 模块来提高 SAR 船舰检测的精度和鲁棒性,同时尽可能地降低计算复杂度。在所提模型中,将 Inception NeXt 模块与 CBS 模块进行串联以取得更好的检测效果,示意图如图 5 所示。

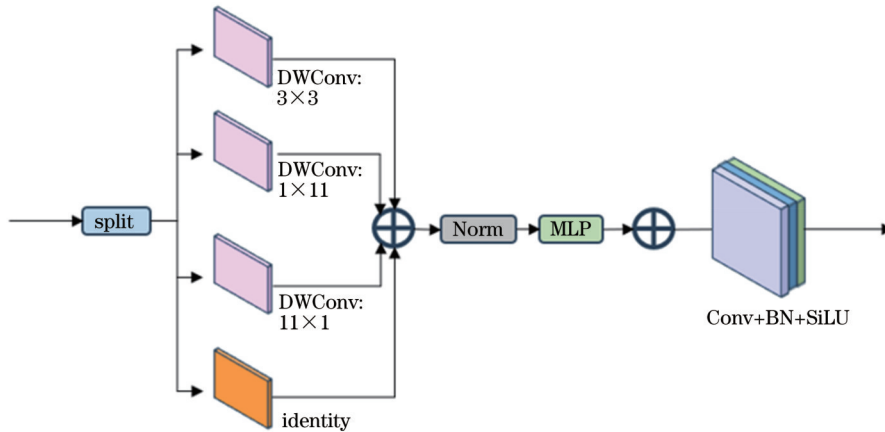


图 5 Inception NeXt 模块

Fig. 5 Inception NeXt module

### 2.3 GE-FPN 分支结构

由于 SAR 图像的特殊性,为了更好地融合特征,需要引入更低维度的特征进行融合,同时需要避免低维度特征传输过程中的损失。因此,设计了 GE-FPN 分支结构,其网络模型如图 6 所示。GE-FPN 的核心思想为融合更低维度的主干网络特征,使 neck 部分获得更多特征。其组成部分 CBS 模块通过对主干网络的 EL-CB 进行连接,结合不同尺度的特征图,将低级和高级特征进行融合,并进一步利用注意力机制来增强模型对重要特征的关注,从而提高模型对特征的表达能。最大池化操作可以帮助模型提取更全局的上下文信息,而上采样操作则可以将低分辨率的特征图还原至原始图像尺寸,以进一步增强特征的丰富性。级联操作作用于不同尺度特征图的连接,以形成合并后

的 FPN。通过引入这些组件,GE-FPN 在 EAFENet 中实现了一种更强大的特征融合和表达能力;使用残差思想进行跳跃连接,增强了模型的学习能力以及泛化性。原特征融合部分的信息主要依赖于主干特征提取网络的三个 EL-CB,然而只选取部分提取到的信息进行特征加强的方式在一定程度上会忽略浅层信息,从而影响模型识别目标的精度。为了让特征加强部分融入更多的浅层信息,将 backbone 部分中提取到的部分信息也融入 FPN<sup>[26]</sup>过程中,在原有的金字塔模型上进行修改,使其更加细致地进行特征融合。

## 3 实验结果与分析

### 3.1 实验环境

本实验以及消融实验配置设备均如表 1 所示。

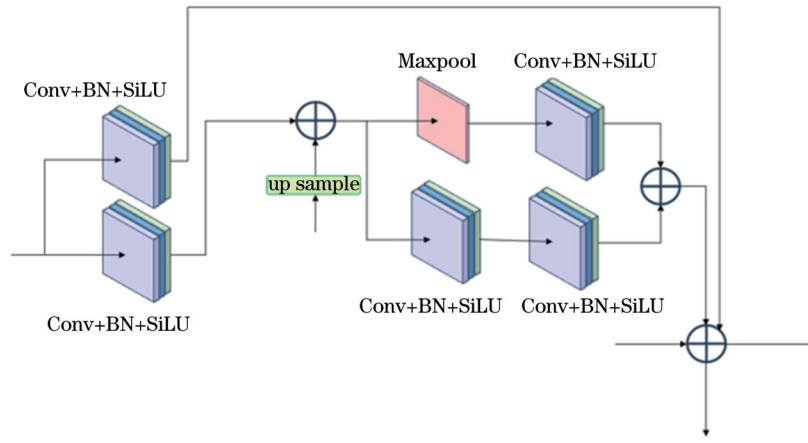


图 6 GE-FPN 模型

Fig. 6 GE-FPN model

表 1 硬件平台和软件平台的参数配置

Table 1 Parameter configurations of hardware and software platform

Configuration	Model
CPU (central processing unit)	Intel(R) Xeon(R) CPU E5-2696 v3 @ 2.30 GHz
GPU (graphic processing unit)	NVIDIA GeForce RTX 3060
Operating system	Windows 10
Running memory	64 Gb
Video memory	12 Gb
Language and framework	Anconda3 + Python3.8 + PyTorch1.13.0

### 3.2 实验数据集

实验使用的数据集为 SSDD 数据集<sup>[27]</sup>, 该数据集由中国人民解放军海军航空大学李建伟于 2017 年发布, 共有 1160 张 SAR 图像, 图像中有 2456 个舰船, 平均每张图片的舰船数量为 2.12 个, 数据集中、背景丰富, 包含了海洋中、近岸等各种排列场景。SSDD 数据集旨在为船舶目标检测算法提供具有挑战性的真实数据集, 同时可以用于比较各种算法的性能。数据集包含大量的 SAR 图像和对应的标注, 可以作为这一研究领域的重要数据资源。实验将 SSDD 数据集按照 7:2:1 的比例随机划分成训练集、验证集和测试集, 具体训练参数如表 2 所示。

### 3.3 评价指标

所采用的实验评价指标为准确率 ( $P$ )、召回率 ( $R$ )、mAP0.5 [代表交并比 (IoU) 检测阈值为 0.5 时的

表 2 训练参数配置

Table 2 Training parameter configurations

Training parameter	Configuration
Momentum parameter	0.937
Initial learning rate	0.001
Batch size	8
Epoch	300
Optimizer	SGD (stochastic gradient descent optimizer)

均值平均精度 (平均准确率)] 以及 mAP0.50:0.95 (代表 IoU 检测阈值从 0.50 到 0.95 时的平均准确率), 它们的公式分别表示为

$$P = \frac{T_p}{T_p + F_p}, \quad (1)$$

$$R = \frac{T_p}{T_p + F_n}, \quad (2)$$

$$V_{AP} = \int_0^1 P(R) dR, \quad (3)$$

式中:  $T_p$  表示真正的正样本 (true positive);  $F_p$  表示错误的正样本 (false positive);  $F_n$  表示未能检测出的正样本 (false negative)。

### 3.4 消融实验结果分析

为了验证所提 EAFENet 模型的有效性, 在 SSDD 数据集上进行了 6 组消融实验, 实验环境以及参数等均相同, 实验结果如表 3 所示。

在表 3 中, 以 YOLOv7 为基准模型, 作为后面 5 组实验的对比基准, YOLOv7 的准确率为 92.9%, 召回率为 88.5%, mAP0.5 为 92.5%, mAP0.50:0.95 为 55.7%。当只将 CBAM 引入主干网络的特征提取模块中时, 与基准模型相比, 准确率下降少许, 但是 mAP 以及召回率的数据均提升约 4~6 个百分点。这是因为在使用 CBAM 后引入了更多的上下文信息, 模型可以更好地学习到目标区域的特征, 并对目标区域进行更准确的检测和定位, 因此, 一些较小的目标被更好地

检测到,从而提高了召回率。同时,对于一些复杂背景干扰的情况,CBAM可以帮助网络更好地筛选出重要的特征,减少误检的情况,从而提高了mAP。但是由于CBAM引入了少量的计算和参数,因此网络的准确率有所下降,需要更多的计算资源来进行训练和推理,可能会增加一些误差。

当使用EL-CB作为主干网络的特征提取模块时,准确率、召回率、mAP等评估指标均有不小的提升,不仅准确率超过了原模型,而且召回率、mAP等指标均提升了约5~7个百分点。这是因为引入EL-CB后,CBS模块、CBAM以及leaky ReLU激活函数的结合增加了网络对目标特征的敏感性,有助于缓解梯度消失和特征丢失问题。通过使用残差连接,EL-CB可以更好地学习到特征的细节和细微变化,从而提高检测器的性能。通过融合主支路不同层级的输出特征以及残差支路的特征,模型可以有效地提取更丰富和更多样化的特征信息。

当仅引入Inception NeXt模块作为neck的特征提取模块时,模型的准确率、召回率、mAP等指标同样有所提升,相对于基准模型,mAP0.5提升了4个百分点。当Inception NeXt作为neck的特征提取模块时,通过引入多个分支实现了多尺度特征提取,提高了模型对不同尺度目标的感知能力;通过多尺度特征提取、

特征融合等方式,提升了模型的性能。

以EL-CB作为主干网络的特征提取模块,并引入Inception NeXt模块作为neck的特征提取模块后,模型的准确率和召回率等再次提升,性能显著超越了基准模型,mAP0.5提升了5.4个百分点,mAP0.50:0.95提升了近10个百分点。EL-CB和Inception NeXt模块都采用了多分支结构和融合策略,可以更全面、更多样化地表达特征,这样的设计使得模型能够更好地捕捉到不同尺度和细节的特征信息,注重多尺度的信息交互和整合,增强其感知能力和泛化能力,更好地适应不同场景和目标,提高检测的准确性。

在EAFENet模型中,不仅引入了EL-CB和Inception NeXt,还使用了GE-FPN结构,增强了金字塔的特征提取能力,令各项指标再次有了显著的提升。与基准模型相比,准确率提升了近3个百分点,召回率与mAP0.50:0.95均提升了近10个百分点,mAP0.5提升了6.4个百分点。EL-CB和Inception NeXt模块能够提供更强大的特征表达能力和更高效的特征提取过程,GE-FPN结构则进一步增强了金字塔的特征提取能力,将主干网络中提取到的浅层特征与FPN进行融合,此外,通过特征加强和跳跃连接等方式,能够获得更准确、更丰富的特征。上述结果证明了每一个改进点的有效性。

表3 消融实验各个改进点效果对比

Table 3 Comparison of the effects of various improvement points in ablation experiment

Model	P / %	R / %	mAP0.5 / %	mAP0.50:0.95 / %
YOLOv7	92.9	88.5	92.5	55.7
YOLOv7+CBAM	92.2	93.2	96.4	61.9
YOLOv7+EL-CB	93.6	93.4	97.2	62.3
YOLOv7+Inception NeXt	93.6	93.3	96.5	61.2
YOLOv7+EL-CB+Inception NeXt	94.4	96.1	97.9	65.0
YOLOv7+EL-CB+InceptionNeXt+GE-FPN	95.4	97.4	98.9	65.6

所提模型EAFENet与YOLOv7的检测效果如图7所示。从图7能够看出,相较于YOLOv7,EAFENet在各种复杂背景下均具有良好的检测效果,误检率显著降低,检测精度提升较大。从第三组检测结果可以看出,即使对于复杂背景下的船舰不规则排列在一起的情况,EAFENet模型检测出的船舰位置与个数也与真实情况相同,即有效地检测出了船舰目标个数以及船舰的位置信息。

### 3.5 对比实验分析

为了进一步对比所提模型的性能,与当前主流算法进行了对比实验,实验环境均相同,且均采用相同的训练集和测试集。CenterNet、SSD、Faster R-CNN、YOLOv5、YOLOv7与所提EAFENet的各项指标如表4所示,其中,FPS代表每秒帧数,单位为frame/s。

在准确率等方面,EAFENet模型相对于其他主流算法更加突出,准确率达到95.40%,其次是

YOLOv5和YOLOv7,准确率分别达到了93.32%和92.90%,而SSD和Faster R-CNN的准确率分别为84.10%和82.70%。这是因为EAFENet使用了更高效的特征提取模块,在一定程度上减少了误判,而SSD等主流算法在特征提取等方面较为薄弱,且在特征融合过程中没有更加深入地融合浅层特征,导致预测结果相对不够准确。在平均精度方面,EAFENet依旧表现最好,mAP0.5达到了98.90%,其次是YOLOv5和YOLOv7,mAP0.5分别达到了94.25%和92.50%,SSD和Faster R-CNN的mAP0.5分别为86.01%和89.17%。然而,EAFENet在网络结构上进行了更深层次的融合导致FPS有些许下降,但是总的来说,相对于其他经典算法,所提EAFENet在速度上具有较大优势,而且EAFENet极低的误检率也能够满足实时检测的基本需求。图8为EAFENet、SSD、Faster R-CNN和CenterNet的检测效果对比图。

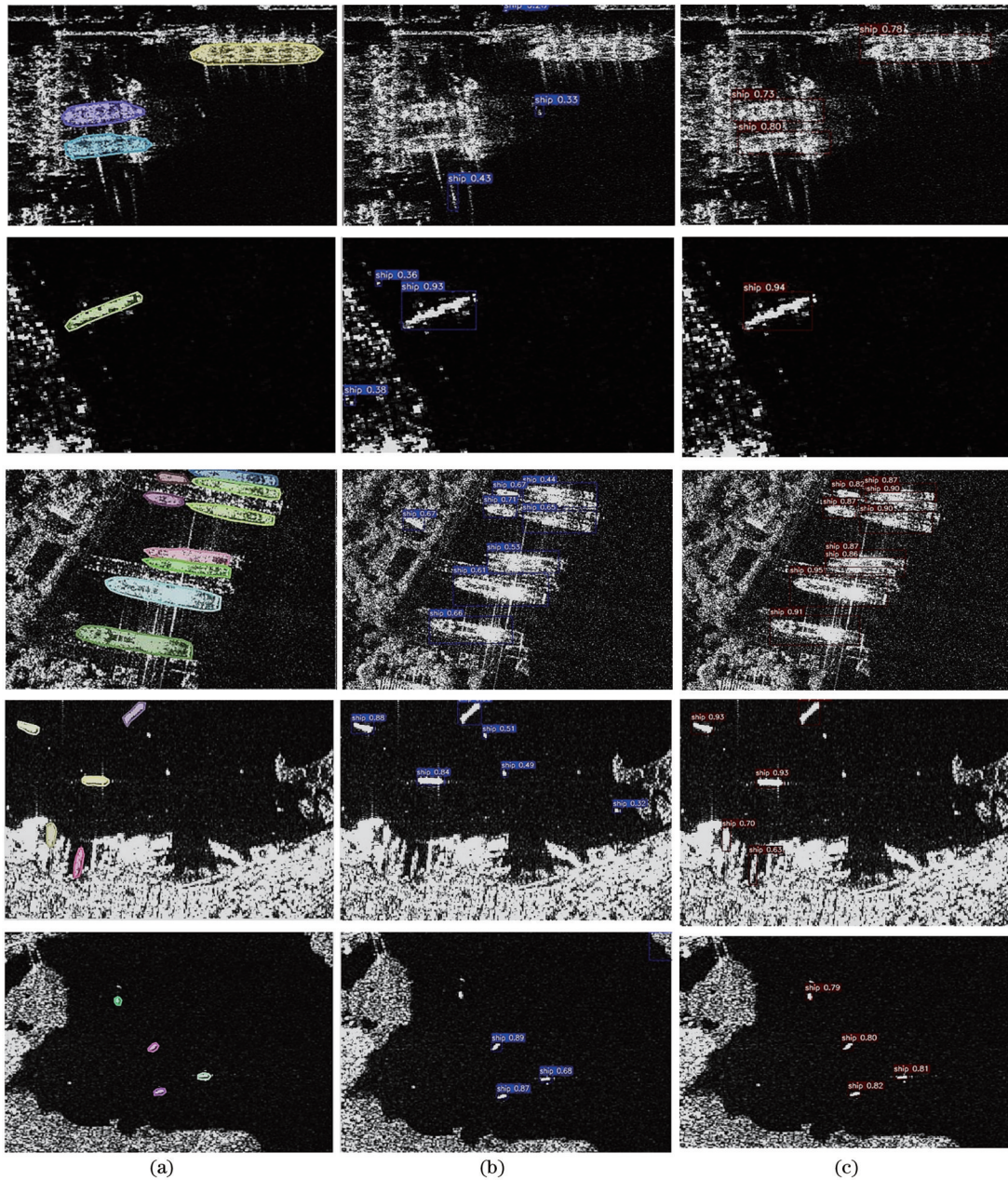


图 7 YOLOv7 与 EAFENet 真实检测效果对比。(a) 真实值; (b) YOLOv7 检测结果; (c) EAFENet 检测结果  
 Fig. 7 Comparison of real detection effects between YOLOv7 and EAFENet. (a) Real values; (b) YOLOv7 detection results; (c) EAFENet detection results

表 4 所提改进模型与主流模型的对比结果

Table 4 Comparison results of improved model and mainstream models

Model	$P / \%$	$R / \%$	mAP0.5 / %	FPS / (frame/s)
CenterNet	87.20	78.62	90.34	23.15
SSD	84.10	89.34	86.01	46.13
Faster R-CNN	82.70	88.62	89.17	24.47
YOLOv5	93.32	89.96	94.25	33.88
YOLOv7	92.90	88.50	92.50	36.18
Ours	95.40	94.70	98.90	31.40

### 3.6 性能曲线

EAFENet 与 YOLOv7 迭代过程中的 loss 曲线、mAP0.5 曲线以及 mAP0.50:0.95 曲线如图 9~11 所示。

图 9 展示了 EAFENet 与 YOLOv7 的 loss 曲线对比效果,可以看出,一开始二者差距不大,但是在训练 30 轮左右之后,所提模型的 loss 曲线下降得更快、更稳定,且收敛值更低。这主要是因为 CBAM 的加持以及 GE-FPN 的特征融合使模型可以更加精准地捕捉目标特征,同时减少传输过程中的信息丢失,从而降低目标检测中的损失。

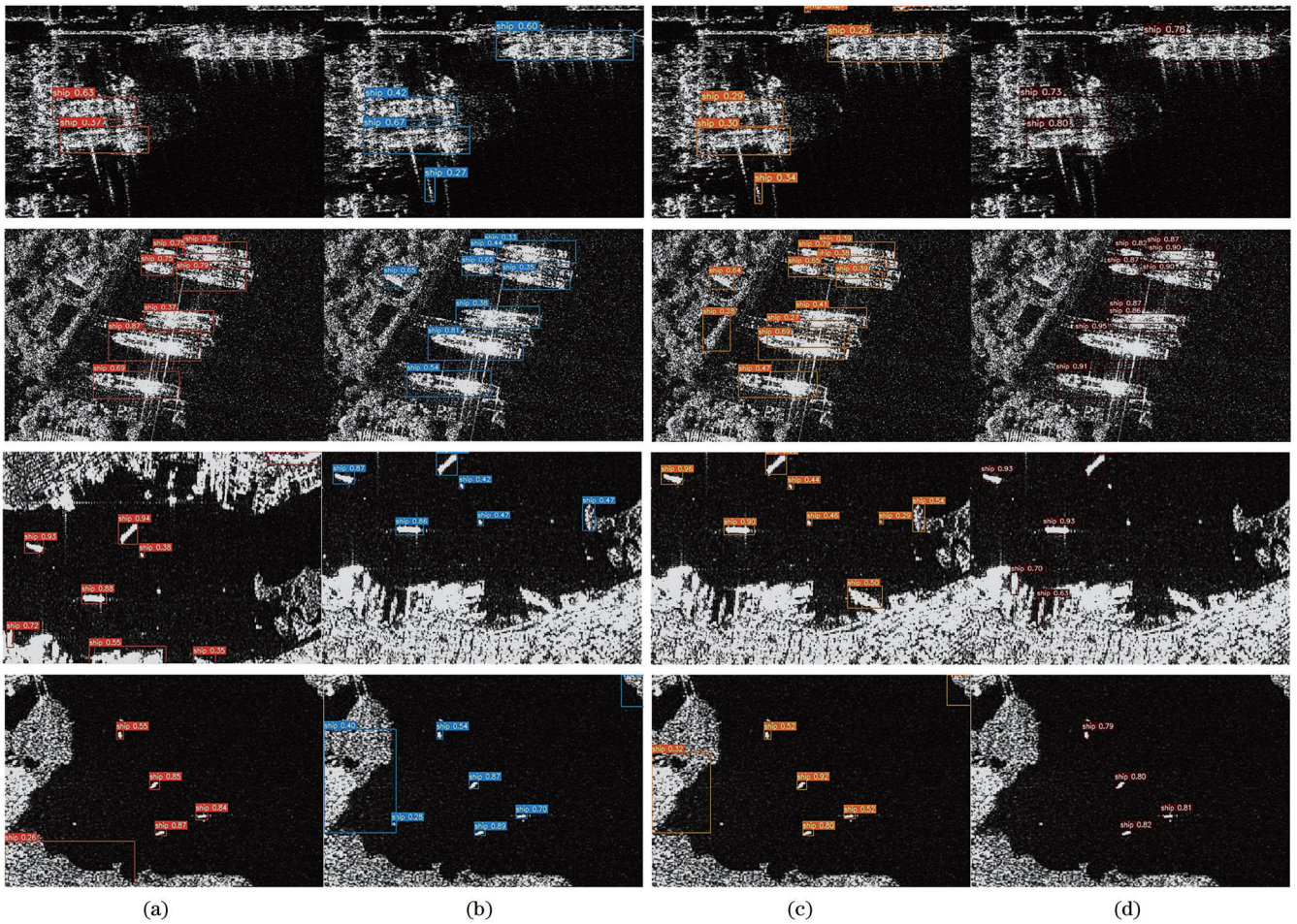


图 8 算法检测效果对比。(a)SSD;(b)Faster R-CNN;(c)CenterNet;(d)EAFENet

Fig. 8 Comparison of detection effects by different algorithms. (a) SSD; (b) Faster R-CNN; (c) CenterNet; (d) EAFENet

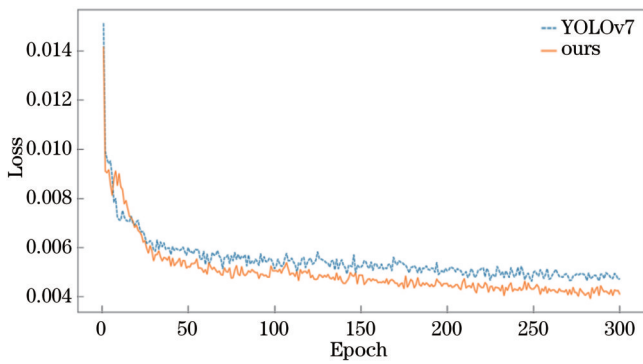


图 9 EAFENet 与 YOLOv7 损失对比

Fig. 9 Comparison of loss for EAFENet and YOLOv7

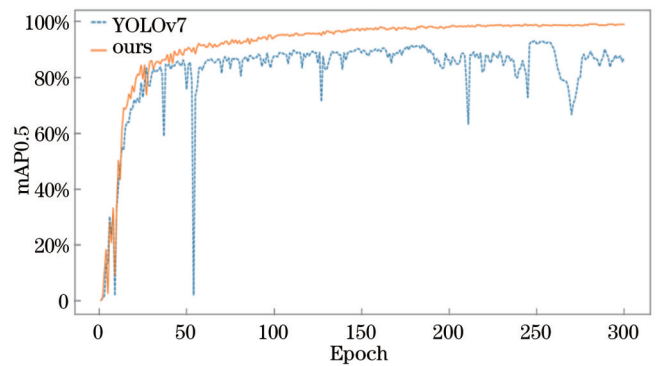


图 10 EAFENet 与 YOLOv7 的 mAP0.5 对比

Fig. 10 Comparison of mAP0.5 for EAFENet and YOLOv7

图 10 展示了 EAFENet 与 YOLOv7 的 mAP0.5 效果对比。从图 10 可以看出,与 YOLOv7 相比,所提模型的 mAP0.5 始终平稳上升,尤其是在 50 轮之后, mAP0.5 已经在 90% 附近,始终比 YOLOv7 有着更加显著的提,且收敛更快,准确度更高。这是 EL-CB、Inception NeXt 和 GE-FPN 相互协同产生的效果,多残差级联层以及并行结构使网络模型在训练的迭代更新过程中更加稳定,效率更高。

图 11 是 EAFENet 和 YOLOv7 的 mAP0.50:0.95 对比图,从图 11 可以看出,EAFENet 具有更好的稳定性,可以面对各种场景下的 SAR 船舰检测。

综合以上结果可以看出,EAFENet 模型比其他模型的表现更为出色,无论是在准确率、召回率还是 mAP 指标上,EAFENet 模型均取得了最好的表现,并且评价指标显著优于其他模型。



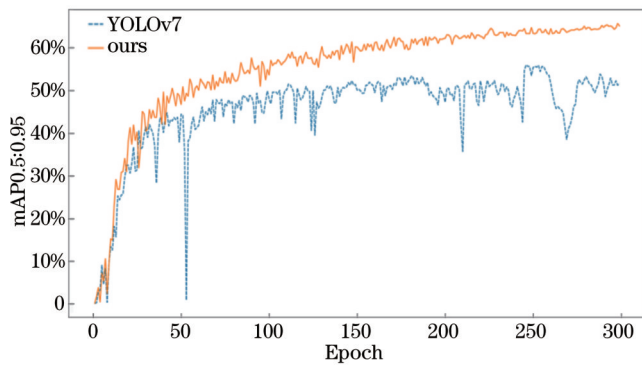


图 11 EAFENet 与 YOLOv7 的 mAP0.50:0.95 对比

Fig. 11 Comparison of mAP0.50:0.95 for EAFENet and YOLOv7

## 4 结 论

针对 SAR 舰船检测精度低、误检率高等问题,提出了一种基于 EAFENet 的 SAR 舰船检测方法。该方法通过空间通道注意力构建出一种 EL-CB 作为主干网络的特征提取模块,以 Inception NeXt 作为 neck 的特征提取部分,提高了算法效率,使网络模型更好地理解多尺度信息和细节。在网络结构中通过将深层次的特征与低层次的特征相融合的方式构建出 GE-FPN 分支结构,使得特征提取网络可以同时兼顾低层次信息与深层次信息,有效增强特征的获取能力,确保模型对复杂背景下舰船的检测稳定性更高。实验结果表明,与当前多种检测算法相比,所提模型的检测精度更高,同时能够满足实时检测的需求。在未来的研究中将进一步优化网络结构,提升检测精度和检测效率。

## 参 考 文 献

- [1] 胡炎, 单子力, 高峰. 基于 Faster-RCNN 和多分辨率 SAR 的海上舰船目标检测[J]. 无线电工程, 2018, 48(2): 96-100.  
Hu Y, Shan Z L, Gao F. Ship detection based on Faster-RCNN and multiresolution SAR[J]. Radio Engineering, 2018, 48(2): 96-100.
- [2] 蒋明哲. SAR 图像舰船检测与分类方法研究[D]. 合肥: 合肥工业大学, 2016.  
Jiang M Z. Research on ship detection and classification in SAR images[D]. Hefei: Hefei University of Technology, 2016.
- [3] Keel B. Constant false alarm rate detectors[M]. London: IET Digital Library, 2010.
- [4] 王昌安. 遥感影像中的近岸舰船目标检测和细粒度识别方法研究[D]. 武汉: 华中科技大学, 2019.  
Wang C A. Research on target detection and fine-grained recognition of nearshore ships in remote sensing images[D]. Wuhan: Huazhong University of Science and Technology, 2019.
- [5] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [6] He K M, Gkioxari G, Dollár P, et al. Mask R-CNN[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2980-2988.
- [7] Carion N, Massa F, Synnaeve G, et al. End-to-end object

detection with transformers[M]//Vedaldi A, Bischof H, Brox T, et al. Computer vision - ECCV 2020. Lecture notes in computer science. Cham: Springer, 2020, 12346: 213-229.

- [8] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M]//Leibe B, Matas J, Sebe N, et al. Computer vision - ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9905: 21-37.
- [9] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 779-788.
- [10] Redmon J, Farhadi A. YOLO9000: better, faster, stronger [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6517-6525.
- [11] Redmon J, Farhadi A. YOLOv3: an incremental improvement [EB/OL]. (2018-04-08) [2022-08-20]. <https://arxiv.org/abs/1804.02767>.
- [12] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23) [2022-08-20]. <https://arxiv.org/abs/2004.10934>.
- [13] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.
- [14] Redmon J, Farhadi A. Yolov3: An incremental improvement[J]. arXiv preprint arXiv: 1804.02767, 2018.
- [15] Glenn J. YOLOv5 releases v6.1[CP/OL]. (2022-02-07) [2022-03-10]. <https://github.com/ultralytics/yolov5/releases/tag/v6.1>.
- [16] Wang C Y, Bochkovskiy A, Liao H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[EB/OL]. (2022-07-06) [2022-08-20]. <https://arxiv.org/abs/2207.02696>.
- [17] 赵鹏飞, 谢林柏, 彭力. 融合注意力机制的深层次小目标检测算法[J]. 计算机科学与探索, 2022, 16(4): 927-937.  
Zhao P F, Xie L B, Peng L. Deep small object detection algorithm integrating attention mechanism[J]. Journal of Frontiers of Computer Science and Technology, 2022, 16(4): 927-937.
- [18] Zhu H Z, Xie Y, Huang H H, et al. DB-YOLO: a duplicate bilateral YOLO network for multi-scale ship detection in SAR images[J]. Sensors, 2021, 21(23): 8146.
- [19] 林志康, 刘伟, 牛朝阳, 等. 基于对数双曲余弦比差异图构造与多区域特征卷积极限学习机的合成孔径雷达图像变化检测[J]. 光学学报, 2023, 43(12): 1228001.  
Lin Z K, Liu W, Niu C Y, et al. Synthetic aperture radar image change detection based on difference image construction of log-hyperbolic cosine ratio and multi-region feature convolution extreme learning machine[J]. Acta Optica Sinica, 2023, 43(12): 1228001.
- [20] 武钰清, 徐青, 马京振, 等. 基于空频双域滤波的 SAR 变化检测算法[J]. 光学学报, 2023, 43(12): 1228009.  
Wu Y Q, Xu Q, Ma J Z, et al. SAR change detection algorithm based on space-frequency dual-domain filtering[J]. Acta Optica Sinica, 2023, 43(12): 1228009.
- [21] 杜艳玲, 崔建华, 魏泉苗, 等. 基于改进 FCN 的多极化 SAR 影像海上溢油检测[J]. 激光与光电子学进展, 2022, 59(4): 0415005.  
Du Y L, Cui J H, Wei Q M, et al. Marine oil-spill detection in multi-polarization image-based SAR on improved FCN[J]. Laser & Optoelectronics Progress, 2022, 59(4): 0415005.
- [22] Rostami M, Kolouri S, Eaton E, et al. Deep transfer learning for few-shot SAR image classification[J]. Remote Sensing, 2019, 11(11): 1374.
- [23] Raj J A, Idicula S M, Paul B. SAR target identification using SAR-COM technique[C]//2019 9th International Conference on

- Advances in Computing and Communication (ICACC), November 6-8, 2019, Kochi, India. New York: IEEE Press, 2020: 92-96.
- [24] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision - ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 3-19.
- [25] Maas A L, Hannun A Y, Ng A Y. Rectifier nonlinearities improve neural network acoustic models[EB/OL]. [2022-08-20]. [http://ai.stanford.edu/~amaas/papers/relu\\_hybrid\\_icml2013\\_final.pdf](http://ai.stanford.edu/~amaas/papers/relu_hybrid_icml2013_final.pdf).
- [26] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 936-944.
- [27] 李健伟, 曲长文, 彭书娟, 等. 基于生成对抗网络和线上难例挖掘的 SAR 图像舰船目标检测[J]. 电子与信息学报, 2019, 41(1): 143-149.
- Li J W, Qu C W, Peng S J, et al. Ship detection in SAR images based on generative adversarial network and online hard examples mining[J]. Journal of Electronics & Information Technology, 2019, 41(1): 143-149.

## Synthetic Aperture Radar Ship Detection Method Based on Highly Efficient Aggregated Feature Enhancement Network

Shan Huilin<sup>1,2</sup>, Liu Wenxing<sup>1</sup>, Wang Xingtao<sup>1</sup>, Fu Xiangwei<sup>1</sup>, Li Changshuai<sup>2</sup>, Zhang Yinsheng<sup>1,2\*</sup>

<sup>1</sup>*School of Electronics & Information Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, Jiangsu, China;*

<sup>2</sup>*School of Electronic & Information Engineering, Wuxi University, Wuxi 214105, Jiangsu, China*

### Abstract

**Objective** Synthetic aperture radar (SAR) is a microwave imaging radar that utilizes the principle of synthetic aperture to achieve high resolution. It has various characteristics such as all day, all weather, high resolution, and wide bandwidth. It is not affected by weather, day, and night and can obtain high-quality, high-resolution, large-scale, and long-distance images. SAR ship target detection technology can provide important technical support in industries such as ocean, oil, port management, marine resource development, and marine scientific research, as it can detect ships and equipments on the sea and detect potential safety risks in advance. At the same time, ship target detection technology has important strategic significance for strengthening maritime monitoring, border patrol, maritime rescue, and safety assurance of maritime channels. We aim to improve the accuracy of SAR ship detection, reduce false positives, and enhance the adaptability of the model.

**Methods** Traditional SAR image target detection methods include texture analysis, polarization characteristics, and constant false alarm rate (CFAR) algorithms. Among them, the most widely used is the CFAR detection algorithm, which has certain advantages in speed, but its drawbacks are high computational complexity and susceptibility to complex backgrounds, resulting in unsatisfactory detection efficiency. In the actual SAR imaging process, the backgrounds of SAR images are mostly ports, islands, reefs, and other buildings. These backgrounds have high grayscale characteristics and strong confusion. Therefore, for the detection of ship targets on the sea, multiple complex backgrounds, various irregular arrangements of ships, similar target misdetection, and other uncertain factors should be considered. The target features of uncertain factors have a certain degree of similarity with ships. Therefore, we propose an efficient aggregation feature enhancement network (EAFENet) to solve the problems of low accuracy, serious false detections, and unstable effects in current SAR ship target detection. The core idea is to efficiently aggregate stacking modules and introduce residual structures to effectively transmit gradient and feature information and alleviate the problem of gradient vanishing and feature loss. The combination of the CBS (convolution+batch normalization+SiLU) module, CBAM (channel spatial attention mechanism), and leaky ReLU activation function increases the sensitivity of the network to target features and introduces low-dimensional feature fusion. Through multi-layer feature pyramid connections, the expression of features is further extended and enhanced, and the residual idea is used for skip connections, enhancing the learning ability and generalization of the model.

**Results and Discussions** In this article, qualitative and quantitative experiments are conducted on EAFENet and other mainstream models for detecting SAR ships, as well as ablation experimental analysis. To demonstrate the effectiveness of each improvement point in this article, the YOLOv7 network model is used as a benchmark, and six sets of experiments

are conducted on the SSDD dataset, with the same environment and parameters. The detected images include multiple targets, few targets, and complex backgrounds. As shown in Table 3, the effect is not ideal only when attention is used alone, and the effect is significantly improved when the mentioned EL-CB (efficient layer convolutional block) is used. The proposed global enhanced feature pyramid network branch structure is used to improve the performance of the feature pyramid and enhance the fusion of shallow features. The accuracy is improved by nearly three percentage points; the recall rate and  $mAP_{0.5 : 0.95}$  are both improved by nearly 10 percentage points, and  $mAP_{0.5}$  is improved by 6.4 percentage points, proving the effectiveness of each module in this article. In order to further compare the performance of the proposed model, the improved algorithm is used for comparative experiments with the current mainstream algorithms. The experimental environment is the same, and the same training and testing sets are used. The indicators of Faster R-CNN, SSD, YOLOv5, YOLOv7, CenterNet, and our algorithm are shown in Table 4. In terms of accuracy, the EAFENet model is more prominent than other mainstream algorithms. EAFENet performs the best with an accuracy of 95.40%, followed by YOLOv5 and YOLOv7, with an accuracy of 93.32% and 92.90%, respectively. The accuracy of SSD and Faster R-CNN is 84.10% and 82.70%, respectively. Compared with other algorithms, EAFENet uses a more efficient feature extraction module, which to some extent reduces misjudgment. However, mainstream algorithms such as SSD have relatively weak designs in feature extraction and other aspects, as well as a lack of deeper fusion of shallow features in the feature fusion process, resulting in relatively inaccurate prediction results. Therefore, when considering the mAP value, EAFENet still performs the best, reaching 98.90% of  $mAP_{0.5}$ , followed by YOLOv5 and YOLOv7, reaching 94.25% and 92.50%, respectively. The  $mAP_{0.5}$  of SSD and Faster R-CNN is 86.01% and 89.17%, respectively. However, the proposed algorithm has undergone deeper fusion in the network structure, resulting in a slight decrease in FPS (frame per second). Overall, compared with other classic algorithms, the proposed algorithm still has significant advantages in speed, and the greatly reduced false detection rate can meet the basic needs of real-time detection.

**Conclusions** In response to the problems of low accuracy and high false detection rate in SAR ship detection, we propose a SAR ship detection method based on an EAFENet. An EL-CB is constructed through spatial channel attention as the feature extraction module of the backbone network, and Inception NeXt is used as the feature extraction part of neck to improve algorithm efficiency, enabling the network model to better understand multi-scale information with detail perception ability. In the network structure, a global enhanced feature pyramid branch structure is constructed by fusing deep-level features with low-level features. This enables the feature extraction network to simultaneously consider both low-level and deep-level information, effectively enhancing the ability to obtain features and ensuring better stability for ship detection in complex backgrounds. The experimental results show that compared with various current detection algorithms, the proposed algorithm has higher detection accuracy and can meet the needs of real-time detection. In future research, the network structure will be further optimized to improve detection accuracy and efficiency.

**Key words** deep learning; target detection; highly efficient aggregated feature enhancement network; attention mechanism; synthetic aperture radar ship detection