

引用格式: WANG Yanni, SUN Xuesong, YU Lixian. Lightweight Synthetic Aperture Radar Ship Detection Algorithm with Enhanced Receptive Field[J]. Acta Photonica Sinica, 2022, 51(2):0210008

王燕妮, 孙雪松, 余丽仙. 增强感受野的轻量化合成孔径雷达船舶检测算法[J]. 光子学报, 2022, 51(2):0210008

# 增强感受野的轻量化合成孔径雷达船舶 检测算法

王燕妮, 孙雪松, 余丽仙

(西安建筑科技大学 信息与控制工程学院, 西安 710055)

**摘 要:**基于深度学习的合成孔径雷达船舶检测方法大多依赖于强大的图形处理器来实现良好的检测精度,却忽略了检测速度和算法的部署应用。针对上述问题,本文提出一种有效增强感受野的轻量化合成孔径雷达船舶检测算法。首先,使用 ShuffleNetV2 作为主干特征提取网络,有利于减小计算参数量和模型大小;其次,引入改进型空间金字塔池化模块与空间注意力模块,有效扩大模型感受野,进一步挖掘船舶特征信息;然后,采用改进的路径聚合网络,自底向上传递更丰富的船舶定位特征,增加浅层位置信息和多尺度特征,进一步提高模型的特征提取能力。使用 SAR 船舶数据集对本文算法进行验证,实验结果表明,模型大小为 5.3 MB,平均检测精度可达 94.7%,检测速度为 46 FPS,同时具有高精度、高实时性和易移植性等优点。

**关键词:**合成孔径雷达;船舶检测;遥感图像;感受野;特征融合;轻量化网络

中图分类号: TP751

文献标识码: A

doi: 10.3788/gzxb20225102.0210008

## 0 引言

随着人类对海洋资源的过度开采,海洋生态环境持续遭到破坏,世界各国纷纷采取措施应对现状。为保护近海生态环境,我国也开始实施对领海和近岸海域内船舶进行监测。但是在天气、海浪等自然界不可控因素的影响下,派遣海警船等监测手段非常耗费人力物力。合成孔径雷达(Synthetic Aperture Radar, SAR)是一种全天时、全天候、多维度获取信息的主动式微波成像雷达,为海洋上多尺度的船舶检测提供了强有力的数据保障和技术支持,在遥感图像船舶检测领域占有重要地位<sup>[1-2]</sup>。

传统的 SAR 船舶检测方法主要有基于变换域<sup>[3]</sup>,基于视觉显著性<sup>[4]</sup>,恒虚警率方法<sup>[5]</sup>,模板匹配法<sup>[6]</sup>等方法,虽然这些方法在某些特定场景下能够取得较好的效果,但这些方法都是采用手工设计的方法提取船舶特征,计算量大,泛化能力弱,并且手工特征提取也较为困难。

随着人工智能技术的兴起,深度学习技术以其更高的精度、更快的速度、更少的人为干预等优势几乎占据了 SAR 船舶检测领域的主导地位。当前主流的目标检测方法分为三类:一类是基于区域推荐的两阶段检测方法,首先生成输入图像的建议框,然后对建议框进行分类和位置回归操作,最终得到检测结果。典型的两阶段检测算法包括区域卷积神经网络(Region Convolutional Neural Networks, RCNN)<sup>[7]</sup>和 Faster-RCNN<sup>[8]</sup>等,此类方法在精度上有明显提高,但是因其占用较多的计算资源,在实时性上有一定程度的降低。另一类是基于回归的单阶段目标检测方法,此类方法将目标检测问题视为回归问题,从输入图像上直接预测边界框和具体物体类别,YOLO<sup>[9]</sup>和 SSD((Single Shot MultiBox Detector))<sup>[10]</sup>是经典的单阶段目标检测算法,由于不再需要中间选取候选区域的操作,检测速度较快。还有一类是基于关键点的检测方法。该方

基金项目:国家自然科学基金(No. 61803294),陕西省自然科学基金基础研究项目(Nos. 2020JM-499, 2020JQ-684)

第一作者:王燕妮(1975—),女,副教授,博士,主要研究方向为智能信息处理和图像处理。Email:wangyn02@126.com

通讯作者:孙雪松(1995—),男,硕士研究生,主要研究方向为深度学习和目标检测。Email:sxs2049@163.com

收稿日期:2021-07-27;录用日期:2021-09-03

<http://www.photon.ac.cn>

法解决了锚框存在的缺点,不需要锚框,同时可以取得良好的检测的速度和精度,主要算法有 CenterNet<sup>[11]</sup>和 FCOS<sup>[12]</sup>等。

基于深度学习的 SAR 图像的船舶检测算法已取得显著的成果。周雪珂等<sup>[13]</sup>从多尺度特征融合与通道关系校准的角度出发,检测精度在改进模型上达到了 89.4%;JIN Liang 等<sup>[14]</sup>在 SSD 网络中每个特征层添加注意机制挤压激励网络模块,有效提升了检测精度;刘竞升等<sup>[15]</sup>基于 FCOS 引入更多特征层,增加跳跃连接改进特征金字塔,更好地利用低层特征,提升了船舶检测效果;ZHOU Long 等<sup>[16]</sup>结合密集连接、残余连接和组卷积的思想,提出了轻量级的卷积神经网络 Lira-YOLO,降低了网络运行的复杂性,方便部署在移动设备上。

目前应用于海上船舶的 SAR 成像速度不断地提高,算法能够实时快速的检测显得尤为重要。海上船舶检测的实际应用环境难以部署模型复杂的算法。因此,算法模型的计算量,决定了能否实现该算法的轻量化和实时性。如张筱晗等<sup>[17]</sup>提出精简网络结构数据驱动的目标分布聚类算法,结合双向向高低层特征融合机制,减少了训练和测试时间,但检测精度较低,同时模型复杂度较高。

为了能同时兼顾检测精度和速度,便于模型部署,本文提出一种有效增强感受野的轻量级 SAR 图像船舶目标检测算法。首先,采用 ShuffleNetV2<sup>[18]</sup>网络进行主干特征提取,降低了模型尺寸;其次,结合改进的特征金字塔池化模块,有效增强网络的感受野;然后,添加空间注意力模块,增强模型对空间位置信息的关注,提升船舶目标定位能力;再者,引入改进型路径聚合网络,进一步融合船舶位置信息和语义信息;最后,使用轻量化检测头,实现算法的高速和高效的检测。

## 1 算法原理

本文提出一种有效增强感受野的轻量级 SAR 图像船舶目标检测算法。主干网络使用 ShuffleNetV2,对特征信息进行高质量的高效提取。在主干网络之后,引入改进的空间金字塔池化模块(Spatial Pyramid Pooling, SPP),将其命名为 SSPP,该模块有利于增强网络感受野,丰富特征信息。在颈部网络中融入空间注意力模块(Spatial Attention Module, SAM),提高网络对船舶目标的定位能力。采用改进的路径聚合网络(Path Aggregation Network, PAN),进行双向多尺度特征融合,进一步增强网络特征图。在检测头部分,使用轻量级卷积结构,输出船舶的位置和类别置信度信息。本文网络整体结构如图 1 所示。

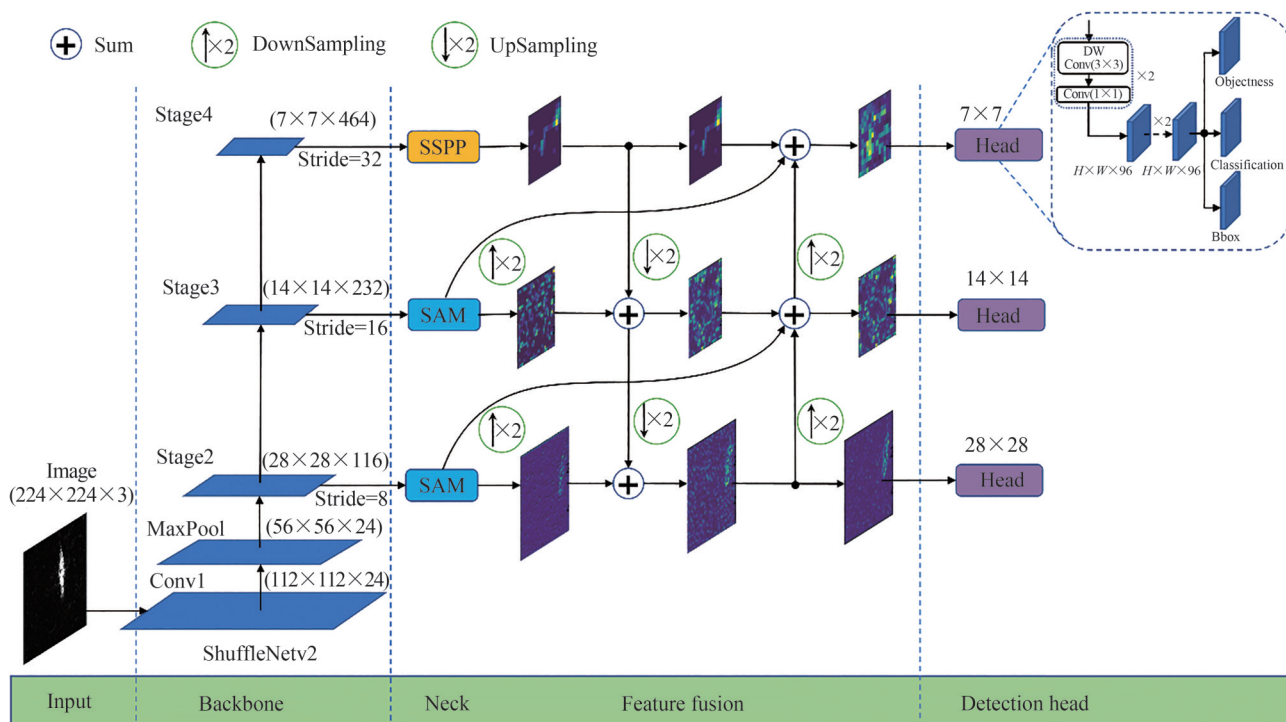


图1 本文网络结构

Fig.1 The proposed network structure

### 1.1 ShuffleNetV2

为了满足嵌入式等移动设备高效运行深层神经网络的同时保持精度的需求,ZHANG Xiangyu等<sup>[19]</sup>引入 ShuffleNetV1模型,该模型基本单元由组卷积和通道洗牌组成。虽然组卷积具有减少参数量的作用,但信息缺少融合。为了克服这个问题,提出了通道洗牌操作,该方法可以实现不同组之间通道信息的交流,操作过程如图2所示。

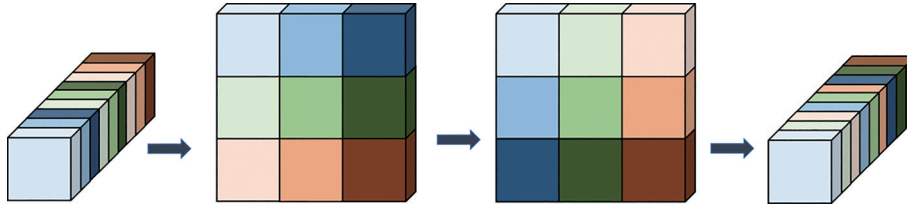


图2 通道洗牌的操作过程  
Fig.2 The operation process of channel shuffle

MA Ningning等<sup>[18]</sup>提出四种高效轻量化网络设计指导方针,并按照指导方针对 ShuffleNetV1进行了改进,引入通道分割结构,重新设计 Block 模块,最终得到 ShuffleNetV2轻量化模型。Block 模块结构如图3所示,图3(a)是步距为1时的基本单元,图3(b)是步距为2时的降采样单元。图1主干网络中的每个 Stage 层由这2个 Block 模块堆叠而成。ShuffleNetV2详细网络参数如表1所示。

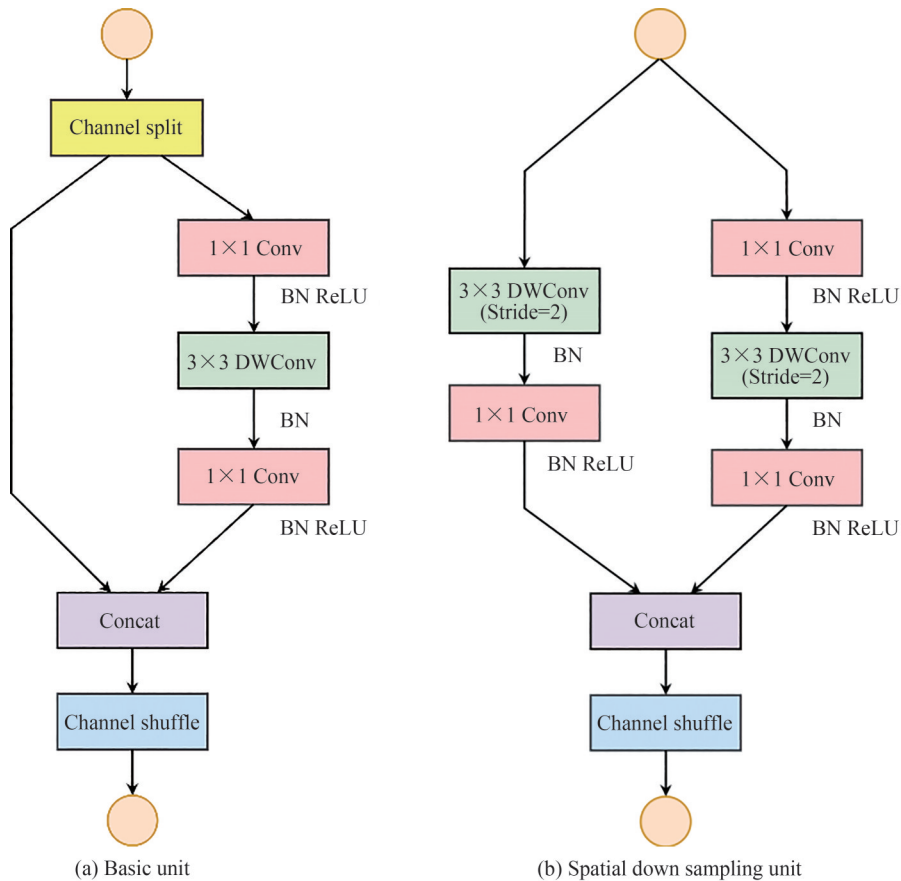


图3 ShuffleNetV2的构建模块  
Fig.3 Building blocks of ShuffleNetV2

表 1 ShuffleNetV2 网络结构  
Table 1 Overall architecture of ShuffleNetV2

Layer	Kernel size	Stride	Repeat	Output size	Output channels (1×)
Image	—	—	—	224×224	3
Conv1	3×3	2	1	112×112	24
MaxPool	3×3	2	1	56×56	24
Stage2	—	2	1	28×28	116
	—	1	3	28×28	
Stage3	—	2	1	14×14	232
	—	1	7	14×14	
Stage4	—	2	1	7×7	464
	—	1	3	7×7	
Conv5	1×1	1	1	7×7	1 024
GlobalPool	7×7	—	—	7×7	—
FC	—	—	—	—	1 000
FLOPs	—	—	—	—	146 M
Weights	—	—	—	—	2.3 M

## 1.2 空间注意力机制模块

在认知科学中,人类视网膜的不同部位具有不同的敏感性。当看到事物时,人类会有选择地将注意力集中在关键信息上,而忽略其他无用信息,这种特性被称为注意力机制。当前,注意力机制被广泛应用于图像分割、图像识别等各种计算机视觉任务中,并取得了显著的效果。卷积块注意力模块(Convolutional Block Attention Module, CBAM)<sup>[20]</sup>是由通道注意力与空间注意力串联而成的混合域注意力机制。通道注意力增强了网络对有意义的输入特征的关注度,有利于提高卷积通道间资源分配的粒度。空间注意力在将原始图像中的空间信息转换到另一个空间时保留关键信息,有助于网络更多地关注特征位置信息。

考虑到本文检测只有一个类别,对网络模型的分类型能力要求较低,但是定位能力要求较高。结合轻量化思想,本文只使用CBAM中的空间注意力模块(Spatial Attention Module, SAM),如图4所示。SAM是对输入特征图在通道维度上进行最大池化和平均池化操作生成两个2维的空间特征图矩阵,将两个特征图在通道维度上进行拼接,之后使用7×7卷积层进行权重的优化,优化后的特征图经过Sigmoid激活函数获得空间注意力图,最后将两者逐点相乘便可得到空间注意力的新特征。SAM表示为

$$M_s(F) = \sigma \left\{ f^{7 \times 7} \left\{ [P_{\text{avg}}(F); P_{\text{max}}(F)] \right\} \right\} \quad (1)$$

式中, $P_{\text{max}}$ 和 $P_{\text{avg}}$ 分别表示最大池化和平均池化操作, $f^{7 \times 7}$ 表示7×7的卷积层, $\sigma()$ 表示Sigmoid激活函数, $F$ 代表输入特征图, $M_s$ 表示空间注意力图。

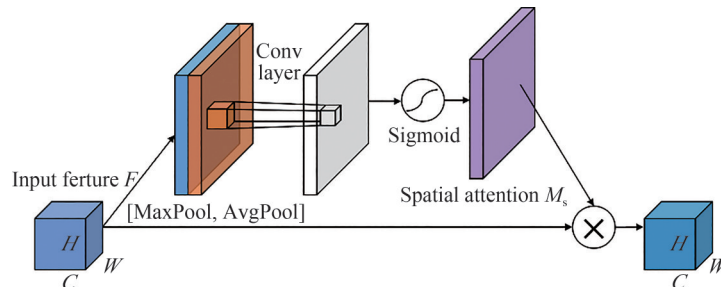


图4 空间注意力模块  
Fig.4 Spatial Attention Module

将SAM添加在主干网络8倍和16倍下采样输出的特征图处,船舶图片经过SAM处理的热力图如图5所示。其中,图5(a)为输入图像,图5(b)为普通卷积输出的热力图,图5(c)为SAM热力图。可以看出,与普通卷积输出热力图对比,经过空间注意力机制的热力图,船舶目标位置信息被显著放大,背景噪声被明显抑

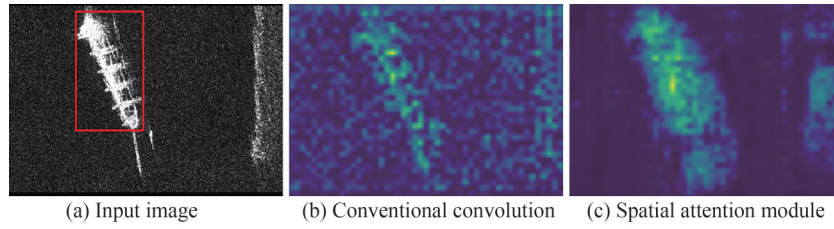


图5 SAM 热力图  
Fig.5 SAM heatmap

制,有利于网络学习到关键目标的信息。

### 1.3 改进型感受野增强模块

SAR图像中远海的船舶目标尺寸相对较小,近岸的船舶目标尺寸较大,对于多尺度船舶的检测一直是研究的重点。一般来说,高层特征的感受野表达强烈语义信息,而低层特征的感受野具有良好的空间位置信息和高分辨率表达能力。文献[21]提出了一种SPPNet,是一种通过多尺度聚合操作将特征映射融合为一个固定长度特征向量的结构,解决了信息丢失和尺度不统一的问题。为了能够提高不同尺度的船舶检测,增加高层语义特征的感受野,本文在主干网络32倍下采样后添加改进的SPP和SAM组合而成的SSPP模块,如图6所示。

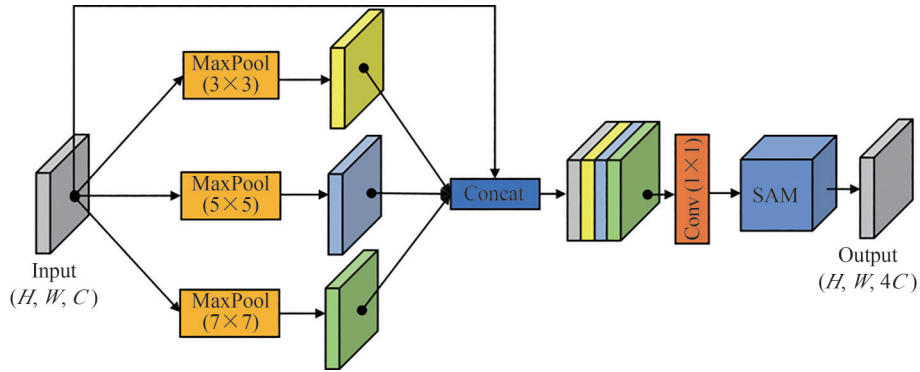


图6 SSPP 模块结构图  
Fig.6 The structure of SSPP

由于远海小尺寸船舶数量相对较多,空间金字塔池化结构的最大的池化核要尽可能与需要池化的特征图的大小相等,因此对原SPP的最大池化核改进为7、5和3,最大池化操作可以在不同尺度上保留最显著的特征,增加了局部区域特征图的感受野,使得网络获得更丰富的局部特征信息,捕获更多的船舶特征。在SSPP模块,输入特征图( $H, W, C$ )分别用 $3 \times 3$ 、 $5 \times 5$ 和 $7 \times 7$ 三种不同大小的核做最大池化操作,池化步长为1,在特征图周围进行padding操作来保证特征图输入输出尺寸不变,把经过最大池化后的多尺度局部感受野特征图和输入特征图在通道上用Concat方式拼接到一起,再输入SAM对船舶空间位置进行优化输出结果( $H, W, 4C$ )。图7为SSPP模块处理后的特征图,图7(a)是原图,图7(b)是普通卷积输出特征图,图7(c)是SSPP处理后特征图。

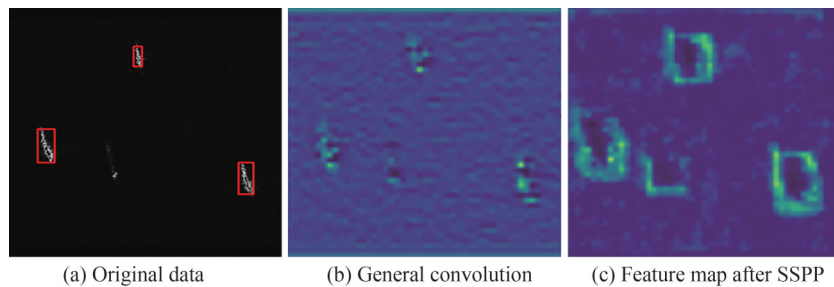


图7 SSPP 特征图  
Fig.7 Feature map of SSPP

从图7可以看出,采用SSPP模块的特征图,相比普通卷积输出的特征图感受野有明显增强,船舶边缘轮廓信息和主要位置特征更加清晰,背景噪声也有显著抑制。SSPP模块能够有效增强网络感受野,提炼出更关键的船舶目标上下文特征,为后面的路径聚合网络输入更加丰富的特征信息,提高对不同尺度船舶的检测效果。

#### 1.4 改进型路径聚合网络

在目标检测任务中,为了使模型学习多样性的特征,提高检测性能,颈部网络对骨干网络提取的特征进行重新处理,将不同尺度的学习分散到多个层次的特征图中,并将不同感受野大小的特征图进行融合。本文引入了双向路径聚合的特征金字塔网络PANet<sup>[22]</sup>,如图8所示。图8中,FPN层是融合路径a,通过自上而下采样与高层特征信息进行融合,增强了网络高层强大的语义特征,但没有增强定位信息;在FPN层后面添加了一个自下而上的特征金字塔b,使得低层特征提供的更准确位置信息通过下采样方式与高层特征进行融合,聚合了不同主干层的参数,提高了网络的特征提取性能。

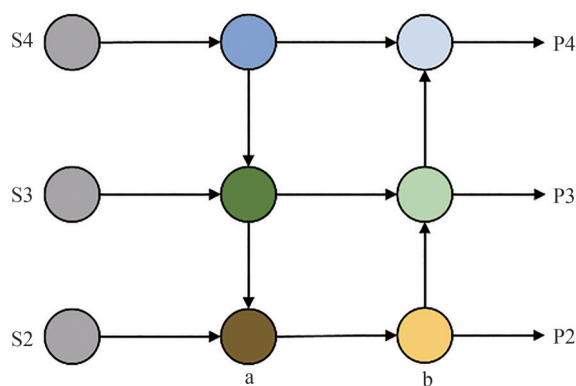


图8 PANet结构  
Fig.8 PANet structure

本文从轻量化和特征提取能力出发,对PANet进行改进,如图9所示。在自下而上的路径b上,融合经过SAM处理后的特征,如图9中的虚线箭头,SAM重点关注船舶位置特征信息,通过与低层特征图的定位信息融合,使得网络在检测小尺度船舶时对低层高分辨率信息有更好的利用,对船舶目标的定位能力更加精确,进一步提高了小目标船舶的检测能力。融合方式采用element sum将多尺度的特征图直接相加,相比Concat操作,没有改变通道数,更节省计算量。为了将检测头轻量化,本文采用了深度可分离卷积,相比普通卷积,该操作可以在保持网络精度的同时使用更少的计算成本。

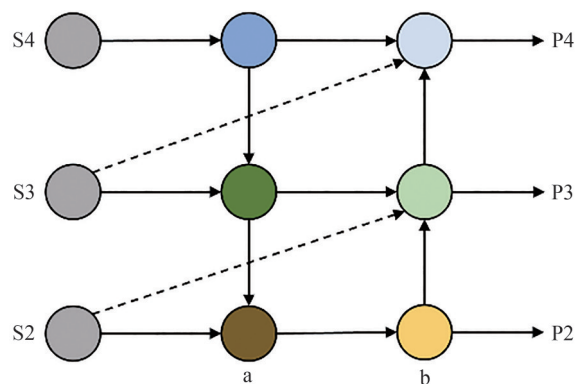


图9 改进的 PANet 结构  
Fig.9 Improved PANet structure

## 2 实验与分析

### 2.1 实验数据

本文使用的数据集是国内外公开的首个专门用于SAR图像船舶目标检测的数据集(SAR Ship

Detection Dataset, SSDD)<sup>[23]</sup>,共1160幅图像,2456条船舶,平均每幅图像有2.12条船舶。该数据集仿照PASCAL VOC数据集的格式构造,使用LabelImg软件对所有船舶图像进行xml格式标注,每个船舶目标的框表示为 $(x,y,w,h)$ ,如图10所示。该数据集包括不同分辨率、极化、海况、大片海域和靠岸等条件的船舶SAR图像,船舶的多样性能较好地验证算法的有效性。本文将SSDD数据集按照7:2:1的比例划分为训练集、验证集和测试集,按照船舶的尺度变化和密集程度的原则进行随机划分,确保划分的数据集在远海和近海的不同尺度的船舶目标均匀分布,以保证实验的有效性。

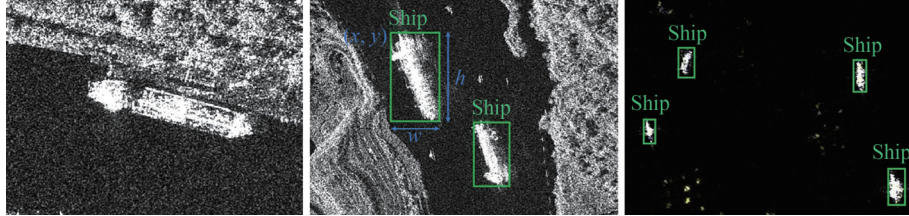


图10 SSDD的部分样本和标签  
Fig.10 Sample images and labels in SSDD

## 2.2 评价指标

本文采用的算法评价指标为准确率(Precision)、召回率(Recall)、平均精度均值(mAP)、每秒帧数(FPS)、运算浮点数(FLOPs)、参数量(Parameters)和模型大小(Model size)。本文仅有船舶一类目标,因此 $mAP = AP$ ,取IoU阈值为0.5时的mAP值。mAP计算公式为

$$mAP = AP = \int_0^1 P(R) dR \quad (2)$$

式中, $P$ 表示准确率, $R$ 表示召回率,mAP表示所有AP值的平均值,衡量模型在所有类别上训练效果的好坏。FPS表示检测速度的快慢,FLOPs用来衡量算法模型的复杂度,Parameters表示网络中的参数数量的多少。

## 2.3 训练过程

本文实验配置环境为:CPU为AMD Ryzen7 3750H、GPU为NVIDIA GeForce GTX1060Ti(4 GB显存)、内存8 GB、硬盘500 GB、Ubuntu 18.04操作系统。使用Python 3.7语言,PyTorch 1.2深度学习框架,并使用cuda 10.0驱动实现图像矩阵的加速训练。

训练过程的优化参数为:训练的epoch为120,设置每批次输入网络的图像数为16,训练过程中,epochs为5的倍数或者超过100时,对模型进行验证以保存验证结果最优的模型。初始学习率设为0.001,训练开始采用warm-up学习率策略,让学习率从0缓慢地增加到学习率0.001,有助于保持模型深层的稳定性。然后采用余弦退火衰减优化策略更新学习率,权重衰减参数设置为0.00005,学习率最终减小到0.0001。采用多尺度训练的策略,提升网络对多尺度船舶检测的能力。本文算法使用迁移学习的方式进行训练,使用ShuffleNetV2在ImageNet上的预训练权重初始化网络的参数,对于后面改进的网络层进行权重正态随机初始化,再对网络整体进行训练。

为了使模型的鲁棒性更强,对单通道的SAR图像的进行数据增强,主要采用了水平翻转、加噪、旋转、尺度变化、改变对比度等随机方案。同时为了解决小数据集训练时泛化能力不足的问题,使用了Mosaic在线数据增强,丰富了数据集中的小目标船舶,提升小尺度船舶的检测性能。Mosaic数据增强方案的效果如图11。

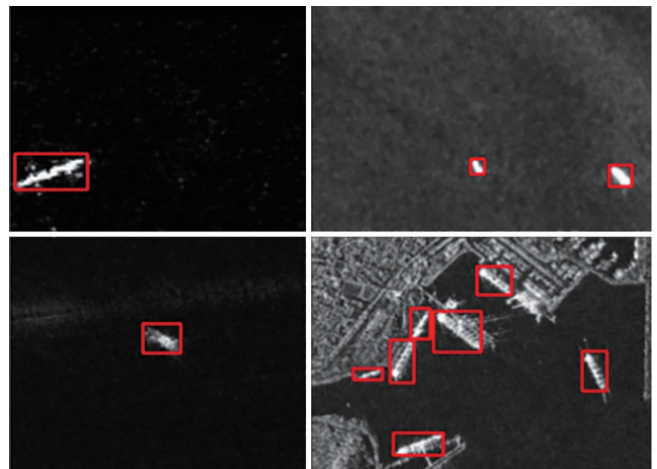


图11 马赛克数据增强  
Fig.11 Mosaic data enhancement

## 2.4 实验结果

为了验证本文改进算法的有效性,在测试集上选取了四种不同的场景、分辨率和船舶尺度的典型SAR图像来展示其检测性能,检测效果如图12,图中绿色标记表示真实框,紫色标记表示预测结果。图12(a)和(b)是近岸码头、近岸海湾停靠的船舶目标检测结果,图12(c)是远海水域的小船舶目标检测结果,图12(d)是噪声干扰下导致图像质量降低的船舶图像检测结果。

从图12可以发现,图12(a)在陆地背景下近岸不同尺度的船舶能够被有效检测识别。图12(b)在海湾中紧挨着排列放置的小船舶目标同样能够被准确识别出来。图12(c)和图12(d)在远海宽阔水域稀疏分布、目标较小的船舶和强噪声干扰下分辨率降低的SAR图像也可以准确地对船舶目标进行定位识别。图13是本文算法在SSDD数据集上的精度召回率图(P-R曲线图),改进模型的mAP可达94.7%,由图中可以发现船舶类别的P-R曲线下面积较大,平衡点取值较高,检测性能较优。从以上结果可以看出,本文算法对近海复杂背景的大小目标、远海水域小目标以及由于噪声干扰下的船舶目标都具有良好的检测能力。

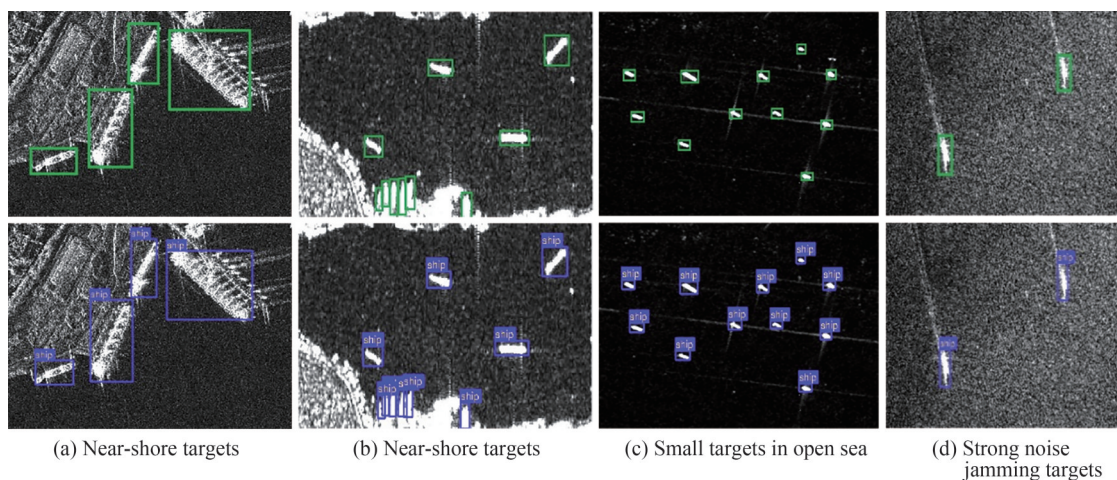


图12 本文算法在不同场景下的目标检测结果

Fig.12 Target detection results of the improved algorithm in different scenarios

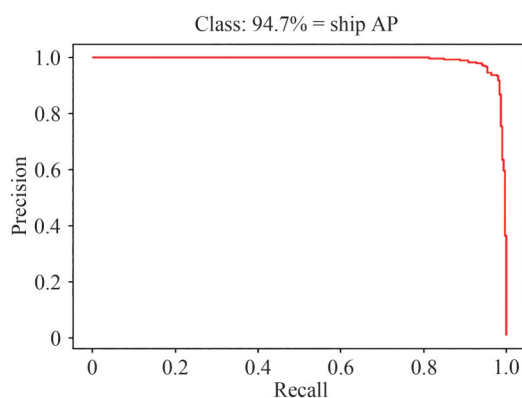


图13 P-R曲线(阈值从0.05变为0.95)

Fig.13 P-R curve (the threshold changed from 0.05 to 0.95)

## 2.5 实验对比分析

### 2.5.1 消融实验

本文提出的轻量化SAR图像船舶目标检测算法,包含4个改进,分别是:采用轻量化ShuffleNetV2作为主干网络,引入SSPP模块,引入SAM模块,以及使用轻量化的PAN。为了评估本文算法中改进的模块对SAR图像船舶目标的检测影响,以轻量化ShuffleNetV2主干网络为基础,在SSDD数据集上进行了消融实验,定量分析了各模块的检测性能,结果如表2所示。

从表2可以看出,在不加任何改进模块的情况下,基础的主干特征提取网络ShuffleNetV2在SSDD测试



表2 不同模块精度对比结果  
Table 2 Accuracy comparison results of different modules

SSPP	SAM	PAN	mAP/%	Speed/FPS
×	×	×	85.6	52
√	×	×	90.2	50
√	√	×	91.5	48
√	×	√	93.4	47
√	√	√	94.7	46

集上的mAP值是85.6%。通过添加SSPP模块,有效扩大了模型感受野,显著地分离了船舶的上下文特征,使得mAP在ShuffleNetV2基础上提升了4.6%。通过添加空间注意机制模块(SAM),提高了船舶“在哪里”的定位能力,减少了小尺度船舶的漏检,mAP提高了1.3%。最后添加轻量化的路径聚合网络(PAN)做特征融合,增加自底向上的细节信息流通过路,提高了船舶目标框坐标的回归精度,增强了对小尺度船舶的检测,相比只使用SSPP模块,mAP提升了3.2%。本文算法在ShuffleNetV2网络的基础上经过3个方面的改进,检测精度相比基础模型提升了9.1%,检测速度达到46 FPS,满足实时性要求,在检测精度和速度上均达到了令人满意的结果。

### 2.5.2 不同算法两种场景对比

鉴于SAR图像近岸船舶通常会有复杂背景,远海船舶目标尺度相对较小的特点,将SSDD测试集的船舶目标分为近岸(Near shore)和远海(Open sea)两类,在SSD、YOLOv3<sup>[24]</sup>、YOLOv3-tiny<sup>[24]</sup>、YOLOv4<sup>[25]</sup>、YOLOv4-tiny<sup>[25]</sup>和本文算法上分别测试其准确率(Precision)与召回率(Recall),对比算法均采用相同的数据划分和优化参数,使用在ImageNet上的预先训练好的权重进行训练,测试结果如表3所示。

表3 不同检测算法在近岸与远海场景下的性能对比  
Table 3 Comparison of different detection algorithms in near shore and open sea scenes

Algorithm		SSD	YOLOv3	YOLOv4	YOLOv3-tiny	YOLOv4-tiny	Ours
Near shore	Precision/%	74.2	75.3	78.4	69.3	72.3	80.6
	Recall/%	91.6	92.4	91.5	78.5	83.8	94.3
Open sea	Precision/%	87.3	92.3	96.7	87.5	91.2	97.8
	Recall/%	93.4	95.4	98.2	83.4	85.6	97.2

从表3可以看出,在复杂背景的近岸场景下,本文算法检测效果相比其它五种算法都有一定能力的提高。在远海船舶目标检测的场景下,六种算法的准确率与召回率都较优,这是因为在远海船舶样本的背景都较为简单,只是在尺度大小上有所差异,小目标船舶所占比例较多,SSD算法对小尺度目标船舶特征描述不充分导致检测效果较差;YOLOv3使用的网络结构是Darknet53,准确率和召回率都较为理想,但是因其网络的感受野不够大,对小目标特征的学习能力不足,故对小目标检测效果不佳;YOLOv4采用了较为复杂的骨干网络CSPdarknet53,能够提取更丰富的特征信息,对远海场景下的小目标船舶的检测效果较好,但在近岸复杂场景下船舶检测的准确率和召回率均低于本文改进算法;YOLOv3-tiny简化了骨干网络,特征信息提取不充分,并且只有2个尺度的预测分支,对大目标和中等目标的检测效果较好,对小尺度船舶目标的定位效果较差造成漏检,两种场景下的准确率和召回率都偏低;YOLOv4-tiny的骨干网络CSPdarknet53\_tiny结构较为简单,网络泛化性不足,对远海场景下的小目标和复杂背景下密集排列的目标漏检较多,故召回率较低。针对Yolov3-tiny和Yolov4-tiny网络难以识别小目标船舶,后续研究引入Focal Loss损失函数进行改进,由于小尺度的目标船舶所占据的像素区域相对于整个SAR图像而言比例很小,在训练过程中会出现正负样本不平衡的问题,采用Focal Loss函数,可以在网络的训练过程中自动对正、负样本,以及简单、复杂样本的不平衡进行调整,减少正负样本分布不平衡所造成的误差,使得网络更加关注难以训练的小目标船舶样本,从而提高小目标的检测精度。

本文算法以轻量化ShuffleNetV2网络为基础,在网络的颈部添加改进的空间金字塔池化模块,极大地增强了网络的感受视野,降低了网络对船舶尺度的敏感性,使用空间注意力机制和改进路径聚合网络,使得

模型更加关注船舶目标的位置,提高了船舶目标的特征定位能力,经过各改进模块的作用,模型对船舶目标的特征表达能力有一定的提高,检测效果表现较为优异。

### 2.5.3 不同算法检测效果对比

为了更直观地比较不同算法的检测结果,从SSDD中选择近岸密集排列停靠的船舶和远海小目标船舶两种典型场景,分别使用SSD算法、YOLOv3-tiny算法和本文算法进行检测,对比效果图如图14和图15所示。图中黄色圆形框表示漏检的目标,红色矩形框表示虚警(误检)的目标。

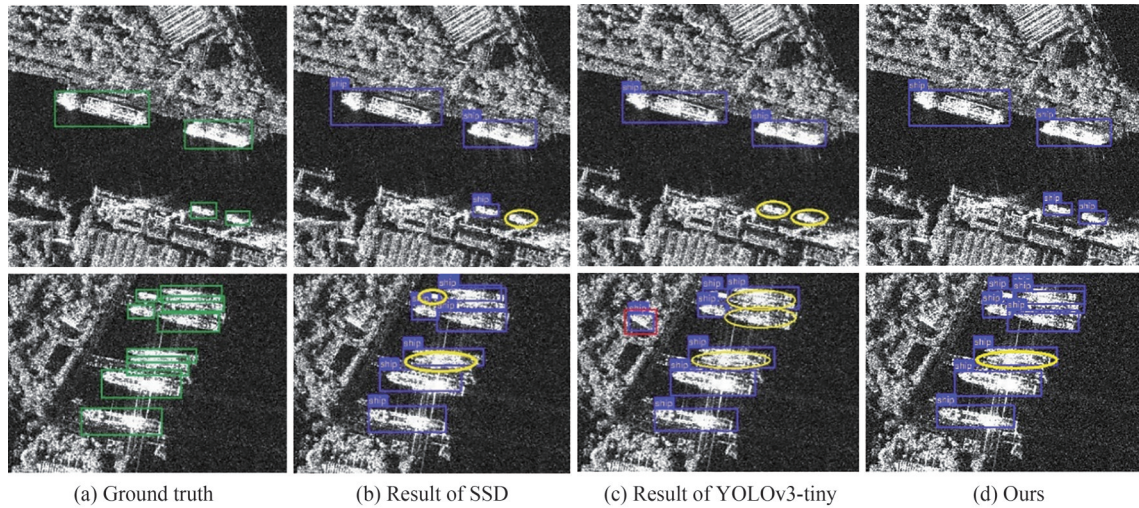


图14 近岸船舶复杂背景的检测算法对比

Fig.14 Comparison of detection algorithms for complex background of near-shore ships

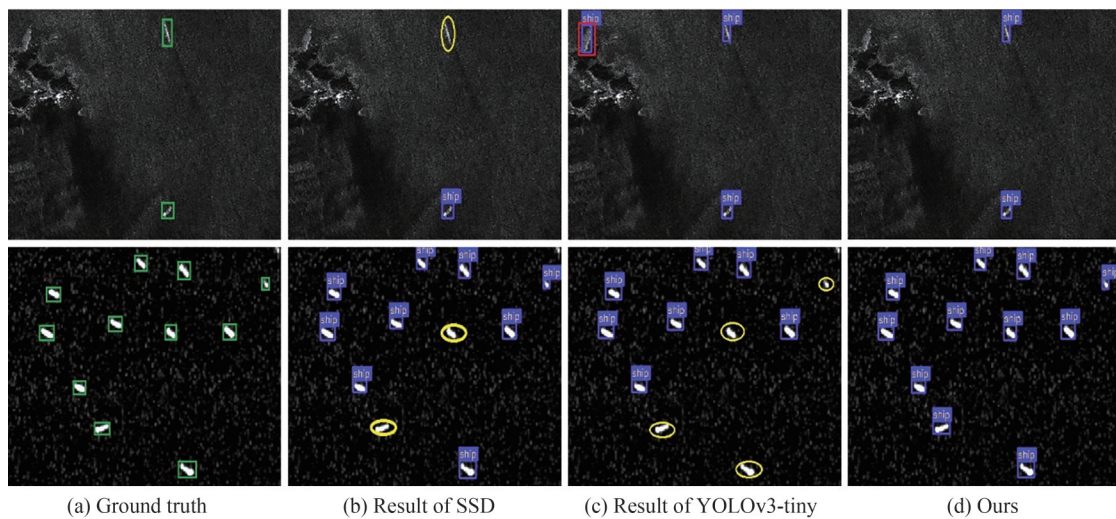


图15 远海小目标船舶的检测算法对比

Fig.15 Comparison of detection algorithms for small target ships in the open sea

从图14的对比结果可以看出,在近岸复杂背景下,SSD算法对小尺度船舶特征提取不充分导致检测精度较低,图14(b)中的小目标船舶均有漏检。YOLOv3-tiny算法由于采用的网络去掉了一些特征层,结构较为简单,只保留了2个独立预测分支,对小尺度目标不敏感,泛化性较差造成了一定的虚警和小目标的漏检,如图14(c)所示。本文算法通过改进的空间金字塔池化操作,能够提取出更丰富的语义信息并且将多种尺度特征融合,有效提高了模型的感受野,同时结合空间注意力机制,提高了船舶目标的定位能力。在图14(d)中本文算法对上一行图中的4艘靠岸船舶目标能够全部检出,对下一行图中的密集排列船舶目标有一个漏检,图14(b)和(c)也存在相同的漏检问题,分析其原因,主要是由于该艘船舶与相邻船舶是平行停靠的,模型会把它检测成单一的船舶目标。通过对比,本文算法能够减少复杂背景下船舶检测中的漏检率和误报率,检

测效果明显更优。

在图15的对比结果中,如图15(b)和(c)所示,SSD算法和YOLOv3-tiny算法均对远海区域的小尺度目标船舶目标不敏感,都有漏检情况,YOLOv3-tiny算法网络结构不是很复杂,泛化性较差,导致漏检率较高并且有虚警情况的发生。本文算法,通过使用特征融合模块路径聚合网络,增加自下而上的信息流通过程,来提供坐标回归所需要低层的细节信息,提高船舶特征的定位能力。图15(d)有效降低了小目标的漏检数量,同时也消除了虚警现象,说明其对小目标船舶的检测效果也非常出色。

#### 2.5.4 不同算法的性能对比

为了进一步验证本文算法的目标检测性能,将本文算法与5种具有代表性的单阶段目标检测算法SSD、YOLOv3、YOLOv4、YOLOv3-tiny以及YOLOv4-tiny进行对比实验。均采用相同的数据划分和优化参数进行训练,实验结果的比较如表4所示。

表4 不同检测算法的性能对比  
Table 4 Performance comparison of different detection algorithms

Algorithm	Backbone	FLOPs	Params	Speed/FPS	Model size/MB	mAP/%
SSD	VGG16	130.9 B	99.6 M	28	100	89.3
YOLOv3	Darknet53	156.3 B	246.1 M	26	236	92.3
YOLOv4	CSPDarknet53	128.5 B	255.7 M	27	244	95.6
YOLOv3-tiny	Darknet53_tiny	5.62 B	8.86 M	35	33.4	90.3
YOLOv4-tiny	CSPdarknet53_tiny	6.96 B	6.06 M	38	22.6	93.8
Ours	ShuffleNetV2	1.52 B	2.34 M	46	5.3	94.7

由表4可知,相比于其它检测算法,本文提出的轻量化SAR船舶检测算法能取得很好的检测效果。本文算法在输入为 $416 \times 416$ 分辨率的情况下,整个模型的浮点运算量Flops只有1.52 B,相比YOLOv3-tiny和YOLOv4-tiny降低了近75%。同时网络的参数量也只有2.34 M,训练完成之后的模型权重文件只有5.3 MB,非常适合在移动端部署。尽管模型非常的轻量,但是性能却依旧强劲,在相同分辨率输入时,能够达到94.7%的mAP,比SSD和YOLOv3-tiny分别高5.4%和3.4%,只比YOLOv4低0.9%。

表4中,YOLOv3和YOLOv4算法都有较高和较快的检测精度,但是其模型尺寸较大、网络参数量较多,难以做到移动端的部署。而本文算法虽然检测精度比YOLOv4低0.9%,但浮点运算量、网络参数和模型尺寸都是最少的,非常适合FPGA或移动端的移植。

本文使用轻量化网络方法对SAR图像船舶进行识别检测,将训练后的模型部署在移动手机端测试验证。测试手机型号为红米K30,移动CPU处理器为联发科MT6885,CPU主频为2.6 GHz,结果表明,在近岸复杂背景和远海小目标场景下,能准确对船舶目标进行精准识别,识别准确率为85.7%,部署到手机后模型前向推理计算的时间只要15 ms左右,加上手机端调用摄像头、图片预处理、检测框后处理以及绘制检测框的时间,测试实时性达到32 FPS,由于手机性能、图像尺寸等因素的影响,并且在移动端的CPU和GPU性能不及训练网络的电脑设备端,检测精度和推理速度有所降低,但仍然满足实时性的要求。

### 3 结论

为解决当前SAR图像船舶检测模型难以部署在便携式移动设备和嵌入式平台的问题,本文提出一种有效增强感受野的轻量级SAR图像船舶目标检测算法。主干网络使用ShuffleNetV2轻量化网络,参数少、复杂度低并且特征表达能力较强。为增强感受野引入改进SSPP模块,加强轻量化网络整体特征提取能力,并与空间注意力模块SAM结合,提高船舶目标的定位能力。同时还采用改进的路径聚合网络PAN做多尺度特征融合,有效利用多层浅层网络特征,丰富特征信息,提高模型检测精度。基于SSDD数据集实验,本文算法平均检测精度为94.7%,模型大小为5.3 MB,检测速度为46 FPS,与当前主流的检测方法相比,具有更少的参数量、浮点运算次数和更小的模型尺寸,在保证高检测精度的同时,实现了快速检测。在未来工作中,将继续针对近岸复杂场景密集排列的船舶目标检测问题进行深入研究。

## 参考文献

- [1] ZHANG Tianwen, ZHANG Xiaoling. ShipDeNet-20: An only 20 convolution layers and <1-MB lightweight SAR ship detector[J]. IEEE Geoscience and Remote Sensing Letters, 2020, 18(7):1234-1238.
- [2] WEI Shunjun, SU Hao, MING, Jing, et al. Precise and robust ship detection for high-resolution SAR imagery based on HR-SDNet[J]. Remote Sensing 2020, 12(1): 167.
- [3] ARMANDO M, MARIA S F, IRENA H, et al. Ship detection with spectral analysis of synthetic aperture radar: a comparison of new and well-known algorithms[J]. Remote Sensing, 2015, 7(5):5416-5439.
- [4] ZHAI Liang, LI Yu, SU Yi. Inshore ship detection via saliency and context information in high-resolution SAR images[J]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(12):1870-1874.
- [5] AI Jiaqiu, YANG Xuezhi, SONG Jitao, et al. An adaptively truncated clutter-statistics-based two-parameter CFAR detector in SAR imagery[J]. IEEE Journal of Oceanic Engineering, 2017, 43(1):267-279.
- [6] ZHU Jiwei, QIU Xiaolan, PAN Zongxu, et al. Projection shape template-based ship target recognition in terraSAR-X images[J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(2):222-226.
- [7] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. IEEE Computer Society, IEEE Computer Society, 2013.
- [8] REN S, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6):1137-1149.
- [9] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR), Las Vegas, NV, USA, June 27-30, 2016: 779-788.
- [10] LIU Wei, ANGELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]. Proceedings of the European Conference on Computer Vision, Berlin, Germany, October 11 - 14, 2016: 21-37.
- [11] ZHOU Xingyi, WANG Dequan, KRÄHENBÜHL P, Objects as points [DB/OL]. 2019. <https://arxiv.org/abs/1904.07850v2>.
- [12] TIAN Zhi, SHEN Chunhua, CHEN Hao, et al. FCOS: Fully convolutional one-stage object detection[C]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE, 2020.
- [13] ZHOU Xueke, LIU Chang, ZHOU Bin. Ship detection in SAR images based on multiscale features fusion and channel relation calibration of features[J]. Journal of Radars, 2021, 10(4): 531-543.  
周雪珂,刘畅,周滨.多尺度特征融合与特征通道关系校准的SAR图像船舶检测[J].雷达学报, 2021, 10(4): 531-543.
- [14] JIN Liang, LIU Guodong. An approach on image processing of deep learning based on improved SSD[J]. Symmetry, 2021, 13(3): 495-495.
- [15] LIU Jingsheng, WU Xing, WANG Honggang, et al. Improved FCOS two-stage SAR ship detection algorithm [J]. Computer Engineering and Applications, 2021, 57(24): 144-151.  
刘竞升,伍星,王洪刚,等.改进FCOS的二阶段SAR舰船检测算法[J].计算机工程与应用, 2021, 57(24): 144-151.
- [16] ZHOU Long, WEI Suyuan, CUI Zhongma, et al. Lira-YOLO: a lightweight model for ship detection in radar images[J]. Journal of Systems Engineering and Electronics, 2020, 31(5): 950-956.
- [17] ZHANG Xiaohan, YAO Libo, LV Yafei, et al. Data-adaptive single-shot ship detector with a bidirectional feature fusion module for SAR images[J]. Journal of Image and Graphics, 2020, 25(9): 1943-1952.  
张筱晗,姚力波,吕亚飞,等.双向特征融合的数据自适应SAR图像舰船目标检测模型[J].中国图象图形学报, 2020, 25(9):1943-1952.
- [18] MA Ningning, ZHANG Xiangyu, ZHENG Haitao, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design [C]. Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany: Springer Press, 2018: 116-131.
- [19] ZHANG Xiangyu, ZHOU Xinyu, LIN Mengxiao, et al. Shufflenet: an extremely efficient convolutional neural network for mobile devices[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA: IEEE Press, 2018: 6848-6856.
- [20] WOO S, PARK J, LEE J Y, et al. Cbam: convolutional block attention module [C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3-19.
- [21] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 37(9):1904-1916.
- [22] LIU Shu, QI Lu, QIN Haifeng, et al. Path aggregation network for instance segmentation [C]. IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 8759-8768.
- [23] LI Jianwei, QU Changwen, PENG Shujuan, et al. Ship detection in SAR images based on convolutional neural network[J]. Systems Engineering and Electronics, 2018, 40(9):1953-1959.  
李健伟,曲长文,彭书娟,等.基于卷积神经网络的SAR图像舰船目标检测[J].系统工程与电子技术, 2018, 40(9):

1953-1959.

- [24] REDMON J, FARHADI A. Yolov3: An incremental improvement[DB/OL]. 2018. <https://arxiv.org/abs/1804.02767>.
- [25] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection[DB/OL]. 2020. <https://arxiv.org/abs/2004.10934>.

## Lightweight Synthetic Aperture Radar Ship Detection Algorithm with Enhanced Receptive Field

WANG Yanni, SUN Xuesong, YU Lixian

(School of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China)

**Abstract:** The marine environment is polluted and the marine ecological environment continues to be destroyed because of the over-exploitation and utilization of marine resources. Countries around the world have taken measures to solve the problem. In order to protect the coastal ecological environment, China has also begun to implement measures to monitor ships in the territorial sea and inshore waters. Ship detection is the focus of current research. Synthetic aperture radar has been widely used in the field of marine remote sensing monitoring because of its advantages such as all-time and all-weather monitoring, which provides strong data guarantee and technical support for multi-scale ship detection. With the improvement of the algorithm, researchers put more energy on the detection accuracy, while ignoring the detection speed and landing application of the algorithm. At present, the popular detection algorithms basically rely on the powerful graphics processing unit, so they can not be deployed in the front line of ocean monitoring. In order to solve the above problems, a lightweight SAR ship detection algorithm which can effectively enhance the receptive field is proposed in this paper. Firstly, ShuffleNetV2 is used as the backbone network. It is helpful to reduce the number of calculation parameters and memory consumption. Secondly, the improved space pyramid pool module is introduced. It not only expands the receptive field of the model effectively, but also makes the ship feature information be further mined. Simultaneously, the spatial attention module is added to enhance the model's attention to the spatial location information and improve the target positioning ability. Then, the improved path aggregation network is used to transfer more abundant ship positioning features from bottom to top. It is helpful to increase the shallow position information and multi-scale features. Finally, the experimental results on the SAR ship detection dataset show that the model size is 5.3 MB, the average detection accuracy is 94.7%, and the detection speed is 46 FPS. Compared with the current mainstream detection methods, it has fewer parameters, floating-point operations and smaller model size. It not only ensures high detection accuracy, but also realizes fast detection. The deployment verification is carried out based on the mobile phone. The results show that the ship can be accurately identified under the inshore complex background and offshore small target scene, the recognition accuracy is 85.7%, and the test speed is 32 FPS. Compared with the computer equipment with excellent GPU performance, the detection accuracy and reasoning speed are all reduced, but still meet the real-time requirements. This method is helpful to transplant to FPGA or embedded mobile devices. It has great practical application value in real-time maritime safety monitoring and protection of marine ecological environment.

**Key words:** SAR; Ship detection; Remote sensing images; Receptive field; Features fusion; Lightweight network

**OCIS Codes:** 100.3008; 110.2970; 280.6730; 100.4996