

基于注意力机制的轻量级SAR船舶检测器

于楠晶¹, 冯大权¹, 朱颖², 张恒嘉^{3, 4}, 陆平^{3, 4}

(1. 深圳大学电子与信息工程学院, 广东 深圳 518060; 2. 中国信息通信研究院, 北京 100191;

3. 中兴通讯股份有限公司, 广东 深圳 518055; 4. 移动网络和移动多媒体技术国家重点实验室, 广东 深圳 518055)

摘要: 合成孔径雷达 (SAR, synthetic aperture radar) 遥感图像凭借其全天候、全时段优势, 在军事侦察、交通监管等领域得到了广泛的应用。卷积神经网络因其较强的学习能力, 被广泛用于 SAR 图像船舶检测算法。然而, SAR 图像中船舶特征提取难度较大。此外, 计算资源和内存空间受限, 实际应用对算法推理速度需求较高。为此, 提出了一种基于注意力的轻量级船舶检测 (LASD, lightweight attention-based ship detector) 算法。该算法设计了一种新的线性混合注意力残差模块, 先后用全局通道注意力和局部空间注意力在深层特征空间中提取船舶潜在特征; 基于跨阶段部分通道连接的空间金字塔池化模块优化多尺度特征融合质量, 用串联小核池化组替换并行大核池化组以提升算法推理速度; 设计了一种新的基于局部注意力的特征融合策略, 在特征融合阶段利用局部注意力进一步扩大船舶和背景噪声的差异。在公开数据集 SSDD 和 LS-SSDD-v1.0 上的实验数据表明, LASD 算法同时兼顾了检测精度和推理速度, 相比其他先进算法更具竞争力。

关键词: SAR; 船舶检测; 卷积神经网络; 注意力机制; 多尺度特征融合

中图分类号: TP391.4

文献标志码: A

doi: 10.11959/j.issn.2096-3750.2024.00407

Lightweight attention-based SAR ship detector

YU Nanjing¹, FENG Daquan¹, ZHU Ying², ZHANG Hengjia^{3, 4}, LU Ping^{3, 4}

1. College of Electronics and Information Engineering, Shenzhen University, Shenzhen 518060, China

2. China Academy of Information and Communications Technology, Beijing 100191, China

3. Zhongxing Telecom Equipment Corporation, Shenzhen 518055, China

4. State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen 518055, China

Abstract: Synthetic aperture radar (SAR) remote sensing images have been widely applied in military reconnaissance and traffic supervision, owing to their all-weather and all-day abilities. With excellent learning performance, convolutional neural networks are employed in the SAR ship detection algorithms. However, it is difficult to extract features. In practical applications, computing resources and memory space are limited, and high inference speed is required. Therefore, a lightweight attention-based ship detector (LASD) was proposed. A novel linear hybrid attention module was designed which extracted potential ship features from deep-level space by using global channel attention and local spatial attention. A spatial pyramid pooling module based on cross-stage partial connections optimized the quality of multi-scale feature fusion, which replaced the parallel max-pooling group with large kernels with the serial max-poolings with small kernels to improve the inference speed. A novel feature fusion scheme via the local channel attention was suggested which widened the gap between the objects and background noise using local attention during the feature fusion. The results on the public datasets SSDD and LS-SSDD-v1.0 show that LASD achieves the balance of detection precision and inference speed, and is more competitive than the other advanced algorithms.

收稿日期: 2024-06-03; 修回日期: 2024-09-27

通信作者: 朱颖, zhuying@caict.ac.cn

基金项目: 深圳市自然科学基金资助项目 (No. JCYJ20210324095209025)

Foundation Item: The Shenzhen Natural Science Foundation (No. JCYJ20210324095209025)

Key words: SAR, ship detection, convolutional neural network, attention mechanism, multi-scale feature fusion

0 引言

合成孔径雷达 (SAR, synthetic aperture radar) 的高分辨率特性使其成为一种重要的有源微波遥感成像传感器。SAR 遥感图像凭借其全天候、全时段及范围广的优势,被广泛应用于环境监测、地质勘查、军事侦察、灾害评估等领域^[1-3]。船舶检测算法为智能船舶、军事侦察和水上交通监管提供了数据支撑和技术支撑。卷积神经网络 (CNN, convolutional neural network) 的提出和深度学习网络较强的学习能力,推动基于深度学习的船舶检测算法快速发展,不仅减少了人工设置对检测结果的影响,还有利于构建高度智能化的监测系统。不断提升的 SAR 遥感图像分辨率促使基于深度学习的 SAR 船舶检测算法成为近年来的研究热点。然而,由于船舶尺寸、方向的多样性和背景噪声的干扰,基于 CNN 的目标检测算法在 SAR 船舶检测任务中表现得并不理想^[4]。根据工作原理差异,基于 CNN 的目标检测算法主要分为基于区域建议的两阶段目标检测算法和基于回归的一阶段目标检测算法。

两阶段目标检测算法在生成预设候选框后对目标进行回归和分类操作。2014年, Girshick 等^[5]提出 R-CNN,首次将深度学习技术用于目标检测任务,避免了人工设置目标特征带来的误差。随后, SPPNet、Faster R-CNN、Cascade R-CNN 等两阶段算法相继被提出。该类算法凭借良好的检测性能在目标检测领域得到了广泛的应用。针对船舶方向多样性问题, Li 等^[6]提出一种两阶段的空域-频域特征融合网络,在频域空间利用旋转不变性提取船舶特征。针对目标运动带来的图像模糊问题, Zhou 等^[7]提出一种多层特征金字塔网络,并将多普勒中心频率和频率偏移引入船舶运动状态。然而,基于区域建议的目标检测算法模型规模较大、推理速度较慢,实际应用中有有限的内存空间及计算资源限制使其难以部署。

2015年, Redmon 等^[8]提出的 YOLO 算法直接对先验框进行分类、回归操作,打破了分两步的目标检测流程。常见的一阶段目标检测算法包括 SSD、RetinaNet、YOLO 等系列算法。该类算法结

构相对简单、推理速度更快,许多学者更倾向于该类船舶检测算法的研究。林鑫伟等^[9]针对信息传递过程中的小尺度船舶特征丢失问题,提出了一种跨连接特征金字塔网络以促进多尺度特征融合,并使用空洞卷积强化船舶特征,但近岸船舶的检测效果依旧不理想。胥小我等^[10]针对复杂场景下正负样本不均衡问题,提出了一种基于自适应锚框和交并比监督的船舶检测算法 A3-IOUS-Net,该算法利用高斯混合模型自适应分配锚框,并在 IoU 预测分支中引入坐标注意力抑制背景杂波,但靠岸和内河的船舶检测效果依旧不理想。谢洪途等^[11]提出了一种基于 CenterNet 的无锚框 SAR 船舶检测方法 Center-SAR,以解决大型算法在小规模数据集上容易出现过拟合的问题,并结合跨阶段局部网络实现对算法的轻量化,但算法参数量和检测效率有待提升。龙昊等^[12]针对 SAR 船舶特征难提取和运行效率较低的问题,提出一种跨尺度融合结构,并结合质量焦点损失、分布焦点损失和广义交并比形成新的损失函数,但检测效果受近岸背景影响较大。邹珺淇等^[13]针对 SAR 图像中噪声干扰和算法轻量化问题,提出一种组合 Ghost 卷积和基于 CBAM 的软阈值化模块,但其推理速度表现不佳。何旭鑫等^[14]为降低 SAR 图像中小尺度船舶的漏检率,在深层网络中引入卷积注意力模块 CBAM,并引入 SIoU 损失函数以提高算法训练效率,但其推理速度在基准模型基础上降低了 50%。考虑 SAR 成像的旁瓣效应和轮廓模糊现象, Zhou 等^[15]设计了一种旁瓣感知模块,利用双池化操作减少旁瓣信号对小尺度船舶的影响。

虽然现有的船舶检测方法已经实现了基于 SAR 图像对船舶进行检测,但由于 SAR 图像中船舶尺寸较小、轮廓不清晰和背景噪点干扰较大等问题,船舶特征提取难度较大,船舶的误检、漏检现象较严重,特别是对近岸船舶的检测效果较差^[16]。因此,学者们主要针对提升算法特征提取能力开展研究,对实际应用中的内存空间、计算资源限制考虑较少,并未同时兼顾检测精度和推理速度需求^[17]。因此,本文提出了一种轻量级、基于注意力的船舶检测器 (LASD, lightweight attention-based ship detector),实现了对 SAR 遥感图像中船舶的精准、快

速检测，主要创新工作如下。

1) 本文设计了一种线性混合注意力残差模块 (LHARM, linear hybrid attention residual module)，在语义信息丰富的深层空间中，利用线性变化替换最大池化操作以保留三维特征信息，先后利用通道注意力和空间注意力提取船舶潜在特征、抑制背景噪声。

2) 本文使用基于跨阶段部分通道连接的快速空间金字塔池化 (SPPF-CSPC, spatial pyramid pooling fast-cross stage partial connections) 模块融合多尺度特征，用串联的小核池化组替换并行大核池化组以提升推理速度，并利用跨阶段部分通道连接结构以增加特征多样性、降低计算量及改善训练过程中的梯度问题。

3) 本文提出了一种基于局部通道注意力的特征融合策略，在特征融合阶段使用基于局部注意力和瓶颈结构的局部通道注意力模块 (LCAM, local channel attention module)，在减少算法复杂度的同时进一步从通道维度强化船舶特征。

1 本文方法

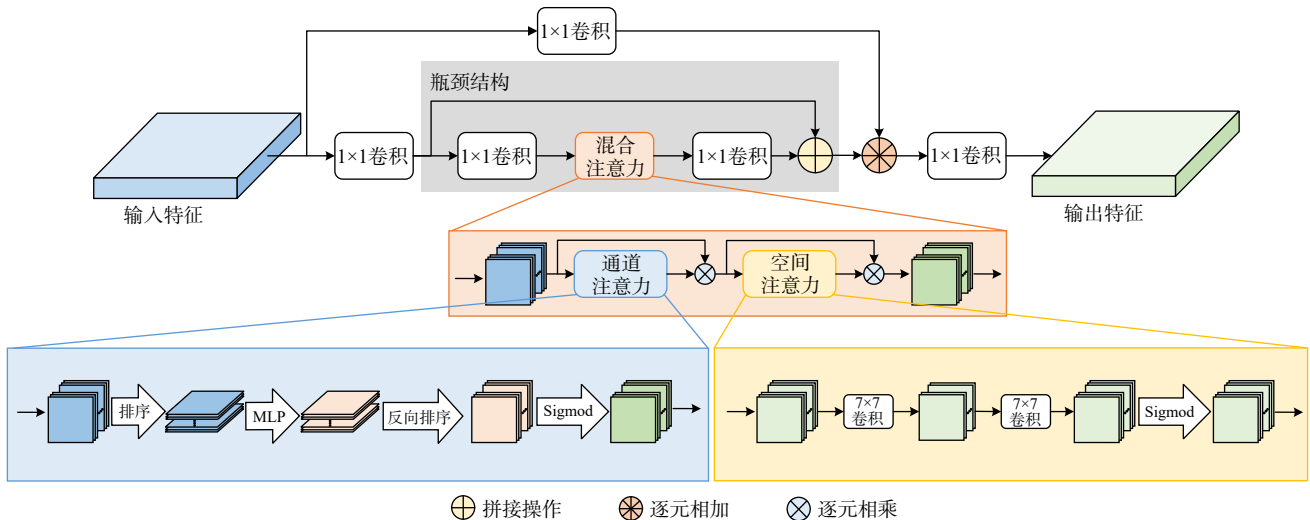
综合考虑实际应用中的检测精度和推理速度需求，本文选择轻量化的YOLOv5算法作为基准算法，所提的LASD算法主要结合CSP-DarkNet^[18]中的瓶颈结构和特征金字塔网络 (FPN, feature pyramid network)^[19]结构提取并融合多尺度特征信息。根据不同的作用，算法可分为3个部分：特征提取、特征融合和目标预测。为优化算法的特征提取能力

和多尺度信息融合性能，实现基于SAR图像对船舶的精准、快速检测，本文在特征提取和特征融合阶段设计了多个模块，本节将对核心模块进行详细介绍。

1.1 线性混合注意力残差模块

基于卷积的CNN结构凭借其优越的特征学习能力在各类深度学习算法中得到了广泛的应用。然而，该类结构对特征图中每个位置进行相同的操作，忽略了不同物体的重要性存在差异。这使得在SAR图像等复杂背景场景下，基于CNN的船舶检测算法无法有效提取船舶的判别特性。此外，CNN通过叠加多层卷积操作获取不同感受野和不同深度的特征信息，而大型SAR场景中存在大量小尺度船舶，信息传递过程中的多层卷积操作容易造成小尺度目标特征信息丢失，进一步导致检测器效果不佳。

考虑SAR图像中船舶特征不明显的问题（主要表现为轮廓及细节信息不显著，从空间维度直接区分船舶和背景难度较大），本文设计了一种线性混合注意力残差模块 (LHARM)，线性混合注意力残差模块结构如图1所示。针对深层空间中的丰富通道信息和SAR图像空间维度的全局相关性较弱的问题，该模块在瓶颈结构和残差结构的基础上，先后利用全局通道注意力和局部空间注意力挖掘船舶潜在特征信息。值得注意的是，LHARM中的局部通道注意力用线性变换（排列、反向排列）代替了传统通道注意力中全局池化和多层感知机 (MLP, multi-layer perceptron) 相结合的方案，不仅



⊕ 拼接操作 ⊗ 逐元相加 ⊗ 逐元相乘

图1 线性混合注意力残差模块结构

有利于减少算法复杂度，还能保留特征空间的多维信息。空间注意力则仅使用7×7卷积操作，减少了传统的空间注意力中的池化操作带来的计算量和细节信息损失。

对于输入特征空间 $F_1 \in \mathbf{R}^{C \times H \times W}$ ，通道注意力输出和该模块输出可表示为^[20]

$$F_2 = G_c(F_1) \otimes F_1 \quad (1)$$

$$F_3 = G_s(F_2) \otimes F_2 \quad (2)$$

其中， G_c 和 G_s 分别为通道注意力和空间注意力的映射函数， \otimes 为逐元相乘。

1.2 SPP-CSPC 模块

不同尺度特征空间所含信息不同，如深层空间中语义信息较为丰富，浅层空间中位置信息更多。此外，检测任务中船舶目标具有多尺度特点，常使用空间金字塔池化（SPP, spatial pyramid pooling）^[21]模块融合提取到的多尺度特征信息，实现多维信息融合，适应多尺度目标检测任务。然而，池化操作的核尺寸越大，计算耗时越长。因此，快速空间金字塔池化（SPPF, spatial pyramid pooling fast）模块利用3个串联的小核最大池化操作代替SPP中

3个并行的、尺寸不同的最大池化操作，提高了运行速度。当然，小核池化操作也将造成参数量的增加。

此外，跨阶段部分通道连接（CSPC, cross stage partial connections）^[18]结构提出了仅处理部分特征，并与未处理的数据进行融合，改善了训练过程中的梯度爆炸及梯度衰减问题。该结构提高了特征信息的多样性，在一定程度上降低了计算量。

综合考虑多样化船舶特征信息需求和计算资源限制，本文使用SPPF-CSPC模块融合获得的多尺度特征，基于跨阶段部分通道连接的空间金字塔池化模块结构如图2所示。

1.3 基于注意力的特征融合策略

传统的特征融合网络采用特征金字塔结构处理多尺度特征。然而，单一的FPN结构传输方向单一、特征融合不够充分，双向FPN结构可加强语义信息和位置信息的深度融合。此外，FPN主要由堆叠的卷积和池化操作构成，忽略了不同物体重要性的差异。因此，本文设计了一种基于局部注意力和双向FPN的特征融合策略，基于局部注意力的特征融合策略如图3所示。

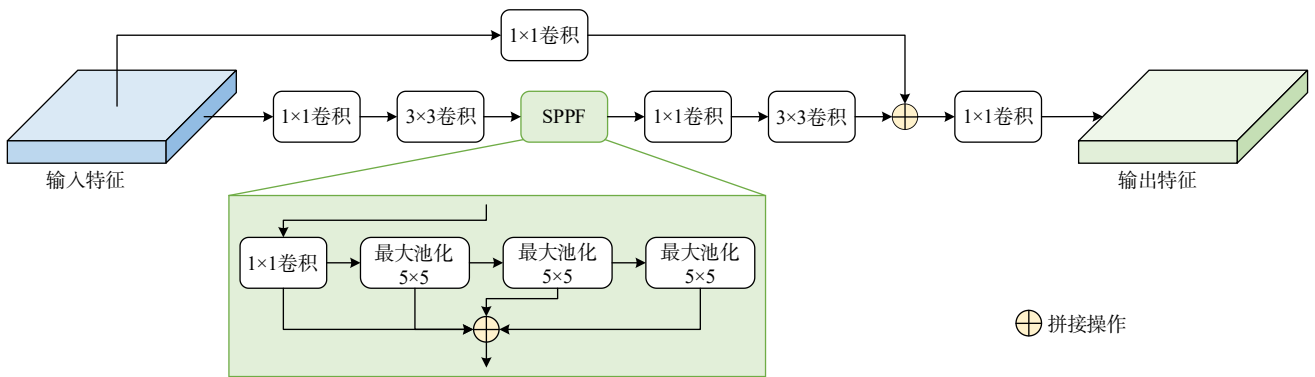


图2 基于跨阶段部分通道连接的空间金字塔池化模块结构

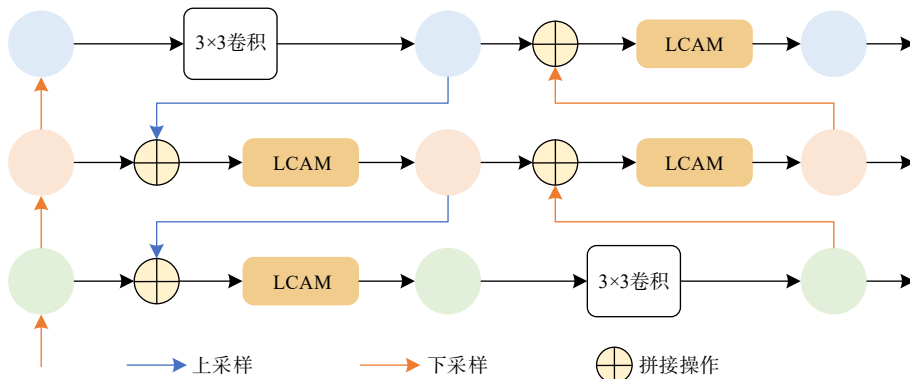


图3 基于局部注意力的特征融合策略

双向 FPN 结构利用两个相反的 FPN 分别传输并深度融合语义信息和位置信息。在特征融合过程中,使用基于局部通道注意力和瓶颈结构的特征融合模块 LCAM (LCAM 结构如图 4 所示),进一步从通道维度扩大船舶特征和背景噪声差异。通常,通道注意力基于所有通道的全局相关性获取权重信息,但实际上相隔较远的通道间相关性较弱。SAR 图像中的噪声影响较大、船舶特征不明显,远距离通道间的相关性对算法性能影响较小。局部通道的范围取决于注意力中的一维卷积核尺寸 k ,与特征通道维度 C 成非线性关系^[22]。 k 为特征通道维度 C 的非线性映射,数学上可表示为

$$k = \lfloor (\ln C + 1) / 2 \rfloor_{\text{odd}} \quad (3)$$

其中, $\lfloor x \rfloor_{\text{odd}}$ 为最接近 x 的奇数。

2 实验验证与结果分析

本文实验采用配置为 Gen Intel(R) Core(TM) i7-12700 2.10 GHz 处理器、NVIDIA GeForce RTX 3080 Ti 显卡的计算机进行,在 Pytorch 框架下使用 Python 语言编程。模型训练过程中,使用梯度下降法优化算法,初始权值使用官方发布的权重文件,预测框和真实框的交并比阈值设置为 0.5,动量因子为 0.937,初始学习率为 0.01,正则化系数为 0.000 5,训练迭代次数为 300 次。

2.1 实验数据集

为验证所提算法的有效性,本文在两个公开的 SAR 图像船舶目标检测数据集上进行了实验,分别为 SAR 船舶检测数据集 (SSDD, SAR ship detection dataset)^[23] 和大尺度 SAR 船舶探测数据集 (LS-SSDD-v1.0, large-scale SAR ship detection

dataset-v1.0)^[24]。

SSDD 是首个面向船舶目标检测任务的公开 SAR 图像数据集,共包含 1 160 张图像和 2 540 个船舶目标。该数据集中的图像分辨率为 1~15 m。参考现有工作,本文按照 8:2 的比例将数据集划分为训练集和测试集。

LS-SSDD-v1.0 数据集的图像覆盖范围为 250 km,共有 15 张分辨率为 24 000 px×16 000 px 的原始 SAR 图像。考虑实验平台的算力限制,本文参考现有工作将原始图像分割成 9 000 张分辨率为 800 px×800 px 的子图像,并按照 2:1 的比例划分为训练集和测试集。

2.2 性能评价指标

为全面、直观并定量地展示本文算法性能,从检测精度和计算复杂度两方面,本文选取以下指标对算法检测性能进行评估。检测精度方面,本文选取了 4 个指标:检测精度 (P , precision)、召回率 (R , recall)、平均精度 (AP, aaverage precision) 和 F1 分数。计算速度方面,本文选取模型参数量和每秒的浮点运算次数 (FLOPS, floating point operations per second) 评估算法的空间和时间复杂度,并利用每秒处理帧率 (FPS, frames per second) 和模型尺寸直观地展示算法的推理速度和本地内存占用情况。

P 值是最直观的检测性能评价指标,表达式为

$$P = \frac{TP}{TP + FP} \quad (4)$$

其中, TP 是将船舶识别为船舶的数量, FP 是将背景识别为船舶的数量。

考虑正负样本不均衡对检测精度影响较大, SAR 数据集中正负样本不均衡现象较严重,因此本文对算法的 R 值进行了对比,其表达式为

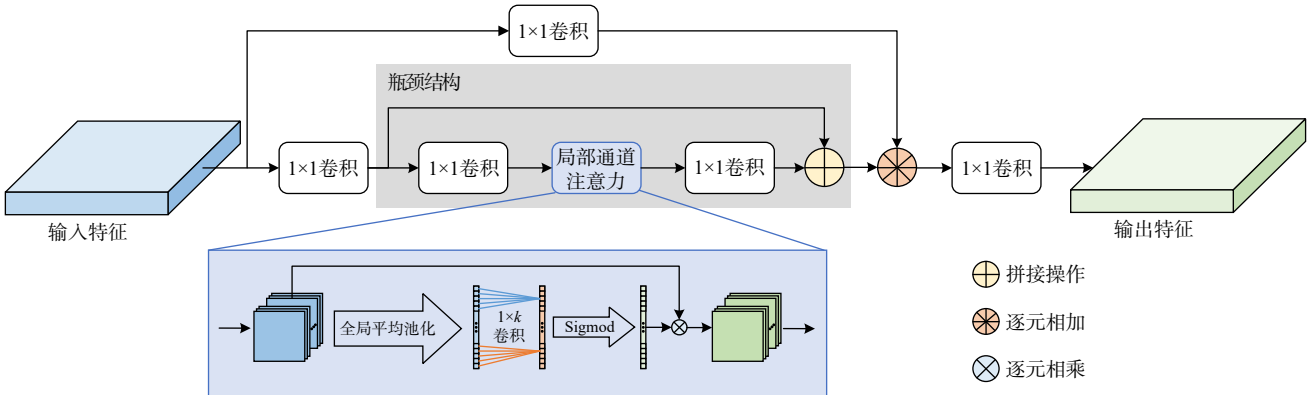


图 4 LCAM 结构

$$R = \frac{TP}{TP + FN} \quad (5)$$

其中，FN为船舶的漏检数量。

AP是算法的准确率-召回率(P-R)曲线与坐标轴所围区域面积，可以更好地展示算法整体检测性能。

$$AP = \int_0^1 P(R) dR \quad (6)$$

F1分数可以全面地反映算法检测精度P和召回率R的情况，常被用于综合评估算法的检测能力，表达式如下

$$F1 = \frac{2 \times P \times R}{P + R} \quad (7)$$

帧率为算法每秒处理图片的数量，可以直观地展示算法的推理速度，表达式为

$$\text{帧率} = \frac{1}{T} \quad (8)$$

其中，T为处理每张图像所需的推理时间成本。

2.3 消融实验

为了验证本文提出的LASD算法和各个模块的有效性，本节在公开数据集SSDD上进行了消融实验，SSDD数据集上的消融实验结果见表1，表中蓝色加粗为最优效果，红色高亮表示指标提升值，绿色高亮表示指标值退化。

SAR图像成像导致的船舶外形特征不明显、噪点较多等问题，使得船舶特征提取难度较大，特别是近岸场景下背景干扰严重，单一注意力模块对不相干信息的抑制效果不明显。考虑深层网络中丰富的通道信息，本文先后使用线性的通道注意力和空间注意力强化船舶特征、抑制背景噪声。实验数据表明，虽然混合注意力模块的使用会带来少量模型尺寸和参数量的增加，但其检测精度得到了大幅提升，特别是算法的AP提升了1.9%。

为减少空间金字塔池化模块使用大核池化操作带来的时延，并提升特征融合质量，本文用串行的小核池化操作替换并行的大核池化组，并嵌入CSPC结构以提高特征多样性并减少计算量。从实验数据可以看出，虽然串行的小核池化组会导致模型尺寸、参数量和FLOPS的增加，但改进后的SPP模块使得检测器的检测精度及推理速度都得到了显著的提升。

此外，针对SAR图像中船舶特征不明显、相似的干扰信息较多及实际应用中较高的检测速度需求，本文在特征融合过程中，使用内嵌局部通道注意力的瓶颈模块LCAM，进一步强化船舶特征、抑制干扰信息。实验数据表明，LCAM模块的引入不仅提升了1.7%的检测精度，还从空间和时间维度降低了算法复杂度，推理速度得到了大幅提升。

为实现基于SAR图像对船舶的精准、快速检测，本文提出的LASD算法同时引入上述3种模块，其检测精度可达97.2%，处理帧率为333.33 FPS。相较基准模型，本文算法的AP提升了2.1%、推理速度提升了125.00 FPS。此外，LASD的F1分数达到了0.95。

2.4 不同算法检测性能对比实验

为展示本文所提算法检测性能的优越性，本节基于SSDD数据集，对比了LASD算法和其他先进算法的检测性能。SSDD数据集上不同算法检测性能对比见表2，FIERNet算法和Quad-FPN算法在船舶检测任务中分别拥有最佳的P值和R值。虽然LASD算法的P值和R值并未达到最优水平，但其平均检测精度AP达到了97.20%，F1分数达到了0.950，推理速度达到333.33 FPS，均比其他先进算法更具优势。

表1

SSDD数据集上的消融实验结果

LHARM	SPPF-CSPC	LCAM	AP	P	R	F1	模型尺寸/MB	参数量/MB	FLOPS/GB	帧率/(f·s ⁻¹)
×	×	×	95.1%	92.6%	91.3%	0.93	13.7	6.73	16.3	208.33
√	×	×	97.0%	97.4%	91.1%	0.94	14.1	6.92	17.0	222.22
×	√	×	96.9%	97.0%	90.7%	0.94	25.9	12.85	21.5	243.90
×	×	√	96.8%	94.8%	93.6%	0.94	11.7	5.75	14.2	270.27
√	√	×	96.9%	96.9%	92.6%	0.95	26.4	13.04	22.1	208.33
√	×	√	96.6%	98.2%	91.6%	0.95	12.1	5.94	14.9	232.56
×	√	√	96.7%	96.4%	93.4%	0.95	24.0	11.88	19.4	256.41
√	√	√	97.2% (+2.2%)	95.1% (+2.5%)	95.1% (+3.8%)	0.95 (+0.02)	24.4 (+10.7)	12.07 (+5.34)	20.0 (+3.7)	333.33 (+125.00)

表2 SSDD数据集上不同算法检测性能对比

算法	输入图像尺寸	AP	P	R	F1	参数量/MB	帧率/(f·s ⁻¹)
FIER-Net ^[25]	448	94.10%	92.70%	94.00%	0.934	32.50	24.94
Quad-FPN ^[26]	512	95.20%	89.50%	95.70%	0.923	—	11.37
YOLOv5 ^[27]	512	95.10%	92.60%	91.30%	0.930	6.73	208.33
YOLOv7 ^[28]	512	96.40%	90.50%	93.60%	0.920	34.79	185.19
SAR-ShipNet ^[29]	512	89.08%	95.12%	76.30%	0.850	134.00	82.00
FIERNet ^[30]	512	94.14%	98.16%	87.89%	0.930	—	—
CenterNet++ ^[31]	512	92.70%	92.60%	94.50%	0.936	20.30	46.51
CR2A-Net ^[32]	512	89.80%	94.00%	87.80%	0.908	88.60	14.88
DAPN ^[33]	512	90.10%	87.60%	91.40%	0.890	63.80	28.99
GFB-Net ^[34]	608	93.00%	85.60%	94.00%	—	41.36	31.50
本文算法 LASD	512	97.20%	95.10%	95.10%	0.950	12.07	333.33

为了直观地展示本文算法检测效果，分别对比了真值图、基准算法、YOLOv7和本文算法在SSDD数据集上对不同尺度船舶的检测结果。SSDD数据集大尺度船舶检测结果对比如图5所示，SSDD数据集包含大量位于港口尺度较大的船舶目标，该类目标纵横比较大、轮廓特征不明显，船舶角度多样，易受港口复杂背景的干扰，对算法的特征提取能力要求较高。基准算法和YOLOv7算法在此类船舶的检测任务中常出现误检、漏检和重检问题。LASD算法在特征提取和特征融合阶段利用注意力模块和瓶颈结构强化船舶的特征信息，强化了算法

的特征捕获能力，优化了对大尺度船舶的检测性能。本文算法不仅减少了误检、漏检和重检问题，其检测精度也得到了提升。SSDD数据集小尺度船舶检测结果对比如图6所示，SSDD数据集中还有离岸较远的小尺度船舶，该类目标的细节信息较少、尺度较小，岛屿和SAR图像成像生成的噪点干扰较大，算法的潜在信息提取能力至关重要。基准算法和YOLOv7算法在此类船舶的检测任务中，受干扰较为严重，将岛屿、噪声点误检为船舶的现象较为明显，而本文提出的LASD算法检测效果较好，减少了误检现象的发生。

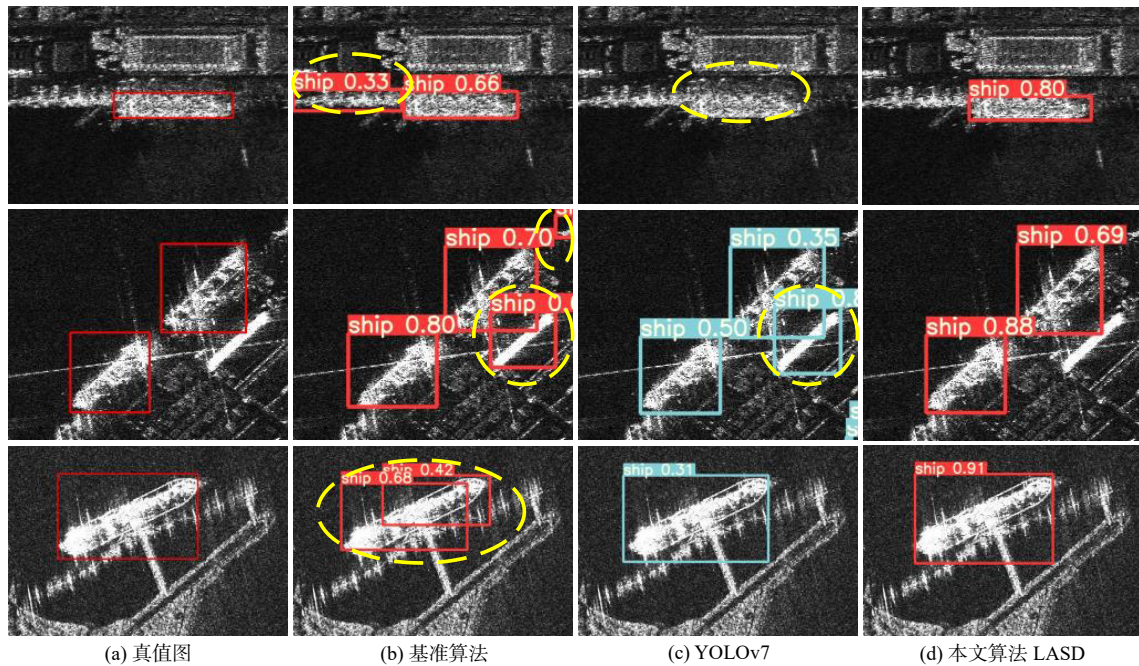


图5 SSDD数据集大尺度船舶检测结果对比

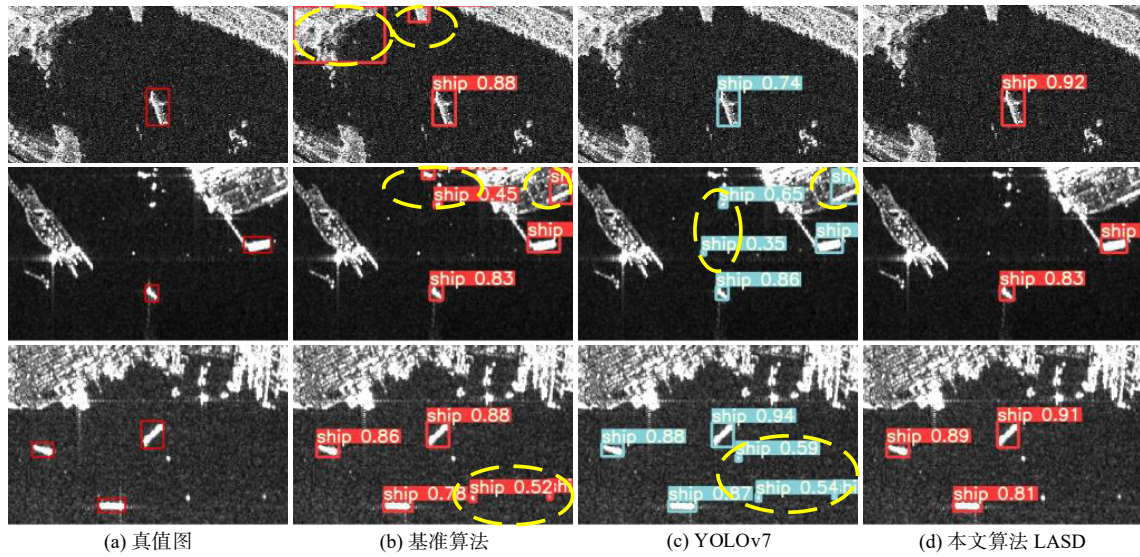


图6 SSDD数据集小尺度船舶检测结果对比

2.5 泛化性实验

为了验证本文算法的泛化性，本节在公开的LS-SSDD-v1.0数据集上进行了性能验证实验，并与其他先进的检测器进行性能对比，LS-SSDD-v1.0数据集上的对比结果见表3。

虽然Free anchor算法和FoveaBox算法分别有最佳召回率和最佳精度，但上述方法并未同时兼顾算法的召回率和精度，F1分数均较低，算法复杂度较高。Libra R-CNN算法有0.750的最佳F1分数，但参数量较高、推理速度较慢。

虽然本文提出的LASD算法F1分数为0.740，略低于Libra R-CNN算法，但其检测精度AP为74.70%，在计算复杂度方面拥有20.00 GB的FLOPS、12.07 MB的参数量和136.98 (f·s⁻¹)的推理速度。实验数据表明，LASD算法可以更好地兼顾检测精度和推理速度，相比其他先进算法更具优势。

为了更直观地展示LASD算法检测性能的优越性，本节还在LS-SSDD-v1.0数据集上进行了检测结果可视化分析。LS-SSDD-v1.0数据集检测结果对比如图7所示，LS-SSDD-v1.0数据集中的船舶尺寸极小，船舶分布较为不均，如港口船舶分布较为密集、远岸船舶分布较为分散。此外，该数据集中噪声点和背景干扰问题更加严重，对算法的特征提取能力要求更高。基准算法的误检问题较为严重，主要表现为将较大较亮的光斑、礁石和近岸的背景误判为船舶。虽然YOLOv7算法极大地优化了误检问题，但在抑制背景噪声时也抑制了部分近岸船舶信息，造成了近岸船舶的漏检。虽然本文所提算法抑制干扰信息效果不如YOLOv7算法，依旧存在少量将礁石误判为船舶的现象，但优化了对近岸船舶的检测性能，且船舶的检测精度明显优于YOLOv7算法。

表3 LS-SSDD-v1.0数据集上的对比结果

算法	AP	P	R	F1	FLOPS/GB	参数量/MB	帧率/(f·s ⁻¹)
Libra R-CNN ^[35]	73.68%	73.50%	76.70%	0.750	162.18	41.62	9.63
EfficientDet ^[36]	61.35%	62.14%	67.49%	0.650	107.52	39.40	4.57
Free anchor ^[37]	71.04%	55.30%	77.67%	0.650	127.82	36.33	11.47
FoveaBox ^[38]	52.32%	97.50%	53.03%	0.680	126.59	36.24	11.47
BANet ^[39]	67.30%	25.10%	35.50%	0.294	—	—	—
L-YOLO ^[40]	73.01%	84.78%	63.96%	0.730	8.10	—	16.31
YOLOv5 ^[27]	74.30%	84.70%	65.10%	0.740	16.30	6.73	118.68
YOLOv7 ^[28]	66.80%	79.00%	62.20%	0.700	103.20	34.79	175.44
本文算法 LASD	74.70%	82.90%	66.20%	0.740	20.00	12.07	136.98

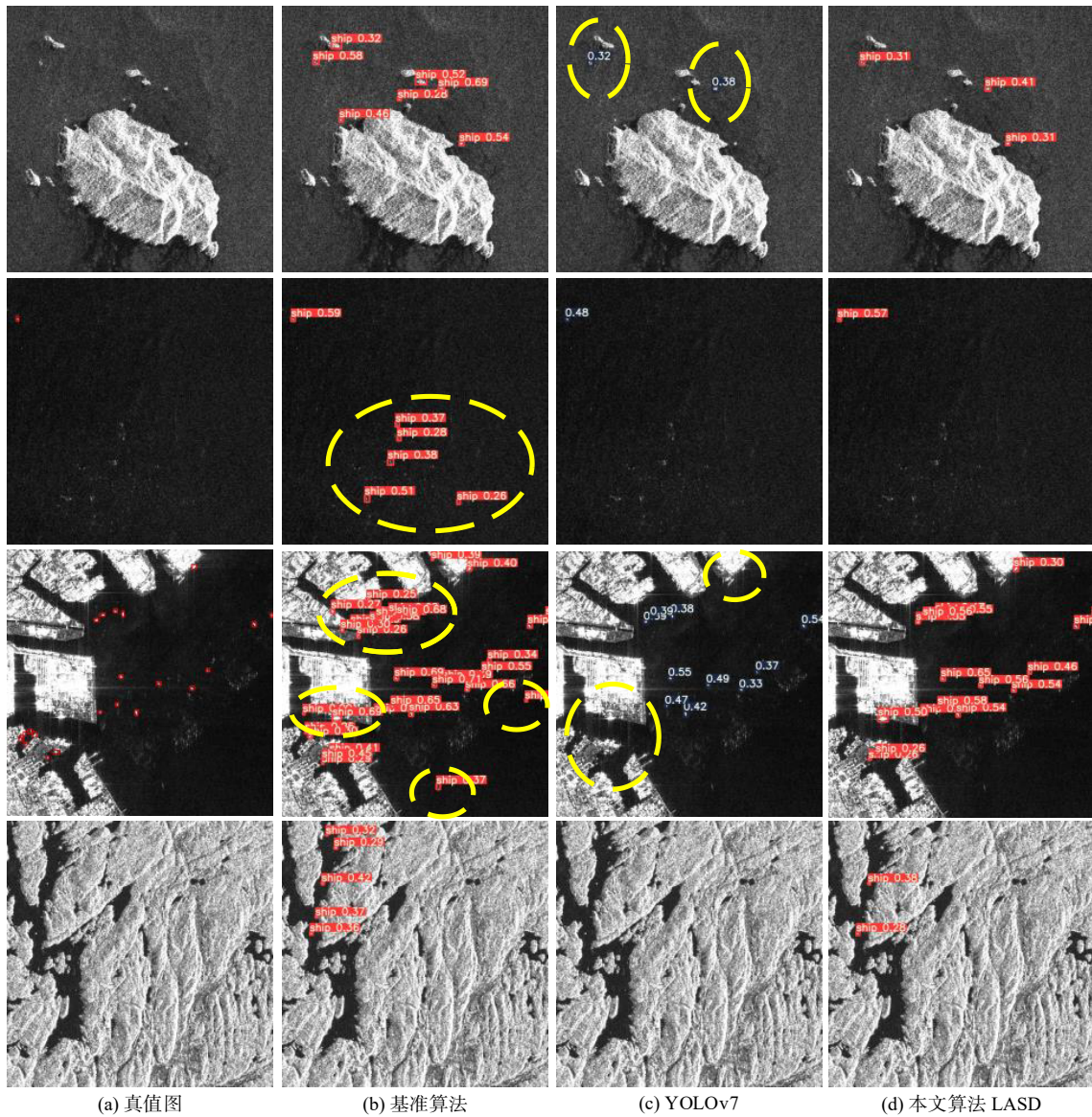


图 7 LS-SSDD-v1.0 数据集检测结果对比

3 结束语

为了基于 SAR 图像对船舶进行精准、快速检测，本文提出了一种基于注意力的轻量级船舶检测算法 LASD。首先，针对 SAR 图像中船舶尺寸较小、船舶判别特征不显著、背景噪声干扰较大等问题，本文设计了一种线性的混合注意力模块 LHARM，在瓶颈结构和残差结构中嵌入线性的混合注意力模块，先后从通道和空间维度提取深层空间中的船舶潜在特征。针对空间金字塔池化模块在融合多尺度特征的过程中大核池化操作带来的时延问题，本文利用串联小核池化替换并行大核池化组，并结合了 CSPC 结构，最终实现了对多尺度特

征的高效融合。此外，本文在检测器的瓶颈部分设计了一种基于局部注意力和双向 FPN 的特征融合策略，根据局部通道之间的相关性进一步扩大船舶特征和背景噪声之间的差异，并降低了算法复杂度。在 SSDD 数据集和 LS-SSDD-v1.0 数据集上的实验结果表明，在 SAR 图像船舶检测任务中，本文提出的 LASD 算法检测性能及推理速度相比其他先进算法更具竞争力。

然而，内河和港口水域背景干扰严重、船舶尺寸较小，本文算法对该区域船舶的检测效果有待进一步提升，应通过提升算法特征提取能力、利用前期图像处理技术等方式，扩大船舶与背景信息差异，进一步优化内河、港口水域的船舶检测效果。

参考文献：

- [1] ZHANG C J, LIU P, WANG H P, et al. NPA2Net: a nested path aggregation attention network for oriented SAR ship detection[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024, 17: 9772-9789.
- [2] ZHANG J D, SHENG W X, ZHU H R, et al. MLBR-YOLOX: an efficient SAR ship detection network with multilevel background removing modules[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2023, 16: 5331-5343.
- [3] LV Z S, LU J, WANG Q, et al. ESP-LRSMD: a two-step detector for ship detection using SLC SAR imagery[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5233516.
- [4] CHEN B J, YU C R, ZHAO S, et al. An anchor-free method based on transformers and adaptive features for arbitrarily oriented ship detection in SAR images[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2012, 17: 2012-2028.
- [5] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//*Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Piscataway: IEEE Press, 2014: 580-587.
- [6] LI D, LIANG Q H, LIU H Q, et al. A novel multidimensional domain deep learning network for SAR ship detection[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5203213.
- [7] ZHOU Y C, FU K, HAN B, et al. D-MFPN: a Doppler feature matrix fused with a multilayer feature pyramid network for SAR ship detection[J]. *Remote Sensing*, 2023, 15(3): 626.
- [8] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//*Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Piscataway: IEEE Press, 2016: 779-788.
- [9] 林鑫伟, 徐志京, 黄海. 复杂背景下的SAR图像多尺度舰船检测[J]. *中国航海*, 2023, 46(2): 17-24, 32.
LIN X W, XU Z J, HUANG H. Multi-scale detection of ship target against complex background out of SAR image[J]. *Navigation of China*, 2023, 46(2): 17-24, 32.
- [10] 胥小我, 张晓玲, 张天文, 等. 基于自适应锚框分配与IOU监督的复杂场景SAR舰船检测[J]. *雷达学报*, 2023, 12(5): 1097-1111.
XU X W, ZHANG X L, ZHANG T W, et al. SAR ship detection in complex scenes based on adaptive anchor assignment and IOU supervise[J]. *Journal of Radars*, 2023, 12(5): 1097-1111.
- [11] 谢洪途, 姜新桥, 王国倩, 等. 基于改进CenterNet的轻量级无锚框SAR图像多尺度舰船检测算法[J]. *哈尔滨工程大学学报*, 2024, 45(3): 504-516.
XIE H T, JIANG X Q, WANG G Q, et al. Lightweight and anchor-free frame method of multiscale ship detection based on improved CenterNet in SAR images[J]. *Journal of Harbin Engineering University*, 2024, 45(3): 504-516.
- [12] 龙昊, 张思佳, 周晶, 等. 基于改进YOLOv5-ResNet的海上舰船SAR图像快速检测[J]. *宇航计测技术*, 2024, 44(2): 52-59.
LONG H, ZHANG S J, ZHOU J, et al. Rapid detection of SAR images of naval vessels based on improved YOLOv5-ResNet[J]. *Journal of Astronautic Metrology and Measurement*, 2024, 44(2): 52-59.
- [13] 邹琨溟, 任酉贵, 冷芳玲, 等. LW-YOLOv7SAR: 轻量SAR图像目标检测方法[J]. *小型微型计算机系统*, 2024, 1-9.
ZHOU J H, REN Y G, LENG F L, et al. LW-YOLOv7SAR: lightweight SAR image object detection method[J]. *Journal of Chinese Computer Systems*, 2024, 1-9.
- [14] 何旭鑫, 吴建平, 余咏, 等. 复杂背景下SAR舰船目标检测算法的改进[J]. *计算机技术与发展*, 2023, 33(11): 41-49.
HE X X, WU J P, YU Y, et al. Improvement of SAR ship target detection algorithm in complex background[J]. *Computer Technology and Development*, 2023, 33(11): 41-49.
- [15] ZHOU Y S, LIU H C, MA F, et al. A sidelobe-aware small ship detection network for synthetic aperture radar imagery[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61, 1-16.
- [16] BAI L, YAO C, YE Z, et al. Feature enhancement pyramid and shallow feature reconstruction network for SAR ship detection[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2023, 16: 1042-1056.
- [17] YANG Y G, JU Y W, ZHOU Z Y. A super lightweight and efficient SAR image ship detector[J]. *IEEE Geoscience and Remote Sensing Letters*, 2023, 20: 1-5.
- [18] WANG C Y, MARK LIAO H Y, WU Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C]//*Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Piscataway: IEEE Press, 2020: 1571-1580.
- [19] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//*Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Piscataway: IEEE Press, 2017: 936-944.
- [20] LIU Y, SHAO Z, HOFFMANN N. Global attention mechanism: retain information to enhance channel-spatial interactions[J]. *arXiv preprint*, 2021, arXiv:2112.05561.
- [21] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904-1916.
- [22] WANG Q L, WU B G, ZHU P F, et al. ECA-net: efficient channel attention for deep convolutional neural networks[C]//*Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Piscataway: IEEE Press, 2020: 11531-11539.
- [23] ZHANG T W, ZHANG X L, LI J W, et al. SAR ship detection dataset (SSDD): official release and comprehensive data analysis[J]. *Remote Sensing*, 2021, 13(18): 3690.
- [24] ZHANG T W, ZHANG X L, KE X, et al. LS-SSDD-v1.0: a deep learning dataset dedicated to small ship detection from large-scale

- sentinel-1 SAR images[J]. Remote Sensing, 2020, 12(18): 2997.
- [25] FU J M, SUN X, WANG Z R, et al. An anchor-free method based on feature balancing and refinement network for multiscale ship detection in SAR images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(2): 1331-1344.
- [26] ZHANG T W, ZHANG X L, KE X. Quad-FPN: a novel quad feature pyramid network for SAR ship detection[J]. Remote Sensing, 2021, 13(14): 2771.
- [27] JOCHER G, NISHIMURA K, MINEEVA T, et al. YOLOv5 (minor version 1.0)[EB].2022..
- [28] WANG C, BOCHKOVSKIY A, LIAO H. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[J].arXiv preprint, 2022, arXiv: 2207.02696.
- [29] DENG Y W, GUAN D H, CHEN Y Y, et al. Sar-shipnet: sar-ship detection neural network via bidirectional coordinate attention and multi-resolution feature fusion[C]//Proceedings of the ICASSP 2022—2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway: IEEE Press, 2022: 3973-3977.
- [30] YU J M, WU T, ZHOU S B, et al. An SAR ship object detection algorithm based on feature information efficient representation network[J]. Remote Sensing, 2022, 14(14): 3489.
- [31] GUO H Y, YANG X, WANG N N, et al. A CenterNet++ model for ship detection in SAR images[J]. Pattern Recognition, 2021, 112: 107787.
- [32] YU Y, YANG X, LI J, et al. A cascade rotated anchor-aided detector for ship detection in remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 60: 5600514.
- [33] CUI Z Y, LI Q, CAO Z J, et al. Dense attention pyramid networks for multi-scale ship detection in SAR images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(11): 8983-8997.
- [34] YAO C, BAI L, XUE D L, et al. GFB-net: a global context-guided feature balance network for arbitrary-oriented SAR ship detection[C]//Proceedings of the 2022 7th International Conference on Image, Vision and Computing (ICIVC). Piscataway: IEEE Press, 2022: 166-171.
- [35] PANG J M, CHEN K, SHI J P, et al. Libra R-CNN: towards balanced learning for object detection[C]//Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2019: 821-830.
- [36] TAN M X, PANG R M, LE Q V. EfficientDet: scalable and efficient object detection[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2020: 10778-10787.
- [37] ZHANG X S, WAN F, LIU C, et al. Learning to match anchors for visual object detection[C]//Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence. Piscataway: IEEE Press, 2022: 3096-3109.
- [38] KONG T, SUN F C, LIU H P, et al. FoveaBox: beyond anchor-based object detection[J]. IEEE Transactions on Image Process-

ing, 2020, 29: 7389-7398.

- [39] HU Q, HU S H, LIU S Q. BANet: a balance attention network for anchor-free ship detection in SAR images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-12.
- [40] XU X W, ZHANG X L, ZHANG T W, et al. On-board ship detection in SAR images based on L-YOLO[C]//Proceedings of the 2022 IEEE Radar Conference (RadarConf22). Piscataway: IEEE Press, 2022: 1-5.

[作者简介]



于楠晶(1998-), 女, 深圳大学电子与信息工程学院博士生, 主要研究方向为目标检测算法、点云编解码算法等。



冯大权(1986-), 男, 博士, 深圳大学电子与信息工程学院副教授、博士生导师, 主要研究方向为沉浸式通信、VA/AR、频谱共享技术。



朱颖(1987-), 女, 中国信息通信研究院高级工程师, 主要研究方向为移动通信射频测试技术。



张恒嘉(1998-), 男, 中兴通讯股份有限公司战略规划师、XRExplore平台工程师, 移动网络和移动多媒体技术国家重点实验室成员, 主要研究方向为深度学习、计算机视觉、SLAM、实时云渲染、空间计算等。



陆平(1971-), 男, 博士, 移动网络和移动多媒体技术国家重点实验室副主任、正高级工程师, 中兴通讯股份有限公司副总裁, 主要研究方向为云计算、AR、媒体大数据。