

面向遥感图像舰船识别的轻量化目标检测算法

张帆¹, 王景波¹, 朱杨¹, 刘海英², 郑煜¹, 王文华¹

(1. 吉林大学 仪器科学与电气工程学院, 长春 130061; 2. 吉林东光集团有限公司, 长春 130012)

摘要: 针对现有基于深度学习的目标检测算法因模型复杂而不利于星上部署的问题, 本文提出了一种面向光学遥感图像舰船识别任务的多尺度特征增强轻量化检测算法(MFLDet)。首先, 为减少算法的参数量和计算量, 构造轻量化网络架构 PG-HGNet 作为主干网络, 同时构造轻量化跨尺度特征融合网络实现特征交互融合; 其次, 设计多尺度特征增强模块(MFEM), 以适应遥感图像中舰船目标的尺度差异性, 提高检测精度; 最后, 引入 MPDIoU 边界框损失函数, 以适应预测框与真实框长宽比相同但长宽绝对值不同的情况。在数据集 HRSC2016 上进行的对比实验结果表明, 与基准模型相比, MFLDet 的参数量和计算量分别下降了 55% 和 23.5%, 而平均精确度仅下降 0.2 个百分点, 在降低算法复杂度的同时实现了与精确度的平衡。综合考量, 所提方法在轻量化水平和检测精度方面均优于其他对比方法。

关键词: 计算机视觉; 目标检测; 遥感图像; 特征增强

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 1671-5497(2026)02-0533-10

DOI: 10.13229/j.cnki.jdxbgxb.20240753

Lightweight object detection algorithm for ship recognition in remote sensing image

ZHANG Fan¹, WANG Jing-bo¹, ZHU Yang¹, LIU Hai-ying², ZHENG Yu¹, WANG Wen-hua¹

(1. College of Instrumentation & Electrical Engineering, Jilin University, Changchun 130061, China; 2. Jilin Dongguang Group Corporation, Changchun 130012, China)

Abstract: To solve the problem of existing deep learning-based object detection algorithms being unsuitable for onboard deployment due to their complexity, a multi-scale feature enhanced lightweight detection algorithm (MFLDet) for ship recognition tasks in optical remote sensing images is proposed. Firstly, to reduce the algorithm's parameter and computation load, a lightweight network architecture, PG-HGNet, is constructed as the backbone network, and constructs a lightweight cross-scale feature fusion network for interactive feature integration. Secondly, a multi-scale feature enhancement module (MFEM) is designed to accommodate the scale variability of ship targets in remote sensing images, thereby enhancing detection accuracy. Finally, the MPDIoU bounding box loss function is introduced to adjust for cases where the predicted and actual bounding boxes share the same aspect ratio but differ in absolute

收稿日期: 2024-07-08.

基金项目: 国家自然科学基金青年基金项目(62105119).

作者简介: 张帆(1979-), 男, 副教授, 博士. 研究方向: 航天光学传感器光机结构, 热设计及仿真分析.

E-mail: zhangfan1@jlu.edu.cn

通信作者: 王文华(1982-), 男, 副教授, 博士. 研究方向: 光电成像与定量遥感. E-mail: wangwh900@jlu.edu.cn

dimensions. Comparative experiments conducted on the HRSC2016 dataset demonstrate that, compared to the baseline model, MFLDet achieves a 55% reduction in parameters and a 23.5% reduction in computation, with only a 0.2 percentage points decrease in average precision, thus effectively balancing complexity and accuracy. Overall, the proposed method surpasses other comparative methods in terms of both lightweight level and detection precision.

Key words: computer vision; object detection; remote sensing image; feature enhancement

0 引言

随着遥感技术的发展,光学遥感图像舰船目标检测在军事国防、民用救援等领域发挥着越来越重要的作用。传统检测方式大多是先将遥感数据下传至地面,再在地面完成检测,时效性差;而在轨目标检测省略了数据传输的过程,不仅极大地提高了检测时效性,还缓解了卫星对地数传的压力。但受星上算力的约束,舰船检测算法不能过于复杂,因此在保证算法检测性能的前提下,保证算法较小的复杂度成为实现舰船检测算法在轨部署的关键。

遥感图像背景复杂,且易受光照、云层等因素的影响,传统舰船检测方法通常针对某一特定场景而设计,因此效果不佳、鲁棒性差^[1],而基于深度学习的遥感图像检测算法能够克服上述缺陷。基于深度学习的目标检测方法主要分为基于回归的一阶段算法(如 SSD^[2]、YOLO 系列)和基于区域的两阶段算法(如快速区域卷积神经网络^[3]、级联区域卷积神经网络^[4])。通常情况下,两阶段算法检测精度较高,但计算量大、检测时间长;而一阶段算法相较于两阶段算法具有网络结构简单、速度快的优点,在运算资源受限的应用场景中更具优势,因此本文将一阶段算法作为主要研究对象。

虽然以 YOLO 系列为代表的一阶段算法已经取得了较好的检测效果,但遥感图像具有物体密集分布、目标尺度差异较大等特点,通用的一阶段算法难以直接应用于遥感图像,因此需要根据其特点进行优化。李永刚等^[5]通过在 YOLOv5 算法的特征提取网络中引入可变形卷积神经网络,增强目标的特征提取能力,提高复杂背景下舰船目标的检测率。成倩等^[6]提出了一种基于 YOLOv5 的光学遥感船舶检测算法,通过加入语义信息增强模块和优化检测头,提高了模型对小目标的检测精度。Yi 等^[7]基于 YOLOv8 的基本结

构提出了 LAR-YOLOv8,通过在双向特征金字塔网络中设置动态稀疏注意力机制,增强了算法对浅层网络特征的提取能力,从而有助于生成更具鉴别力的信息。车思文等^[8]基于 YOLOv7 的基本结构提出了一种整合多维注意力机制和动态卷积的改进算法,通过设计高效的特征融合网络和多层次超大卷积核层,显著提升了检测精度。上述研究虽然通过结合遥感图像的特点对一阶段算法的结构进行了优化,提高了检测精度,但不可避免地增加了网络的复杂性,不利于在边缘设备上部署。

基于上述分析,为了在遥感图像舰船检测任务中平衡算法的检测精度和模型复杂度,以适应边缘部署的场景,本文以 YOLOv8nano 为基准模型进行改进优化,改进点有:①基于 PP-HGNet 的基础架构,融入 Ghostconv 以减小计算量,将处理后的 PG-HGNet 作为主干网络,降低主干网络复杂度。②基于瓶颈思想和双向特征金字塔网络结构,构造轻量化跨尺度特征融合网络(Lightweight-BiFPN, LW-BiFPN),降低特征融合网络的复杂度并提升检测效果。③设计多尺度特征增强模块(Multi-scale feature enhancement module, MFEM),通过采用更大感受野的卷积核以更好地捕获不同尺度的信息,并采用空洞卷积保证较小的复杂度,改善目标尺度变化对算法检测的不良影响。④引入 MPDIoU 损失函数,以适应检测过程中预测框与真实框具有相同长宽比但长宽绝对值不同的情况,从而更精准地定位目标。对比实验结果表明,综合考量轻量化水平和检测精度,本文算法优于主流目标检测算法。

1 MFLDet 网络

MFLDet 网络的整体架构如图 1 所示,由主干网络、特征融合网络和输出端 3 部分组成。首先,遥感图像被输入主干网络 PG-HGNet 中进行特征提取,输出 3 个不同尺度的特征图。其次,特

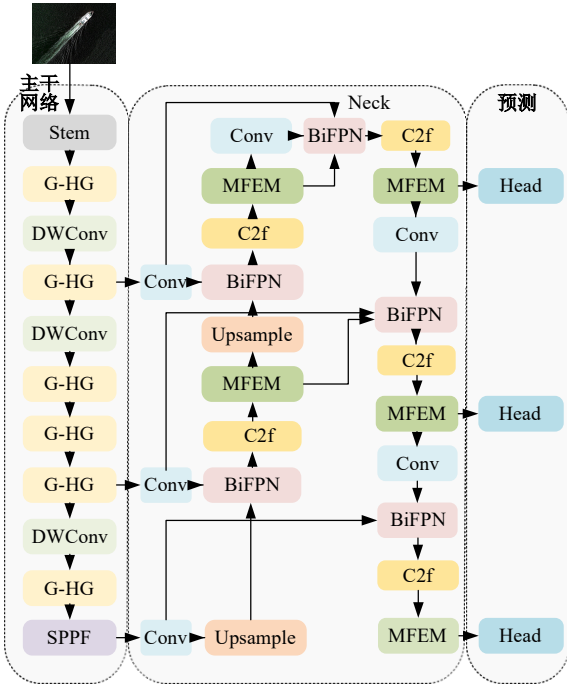


图 1 MFLDet 结构

Fig. 1 Architecture of MFLDet

特征图被输入轻量化特征融合网络中,经通道降维及不同层次特征信息的交互融合后,获得既含有丰富特征语义信息又含有图像细节的融合特征。最后,输出端输出含有预测边界框、中心性和分类的三维张量检测结果,完成检测过程。

1.1 PG-HGNet 主干网络

在目标检测算法中,主干网络作为特征提取的重要组成部分,其感知能力直接影响整个算法的检测性能,同时其参数量与计算量也在很大程度上决定了网络模型的复杂度。为了在保证主干网络检测精度的同时减小模型复杂度,本算法基于高性能 GPU 网络 (High performance GPU net, PP-HGNet) 进行改进。PP-HGNet 是百度飞桨视觉团队自研的高性能骨干网络,该网络在 VOVNet 的基础上引入了可学习的下采样层 (LDS layer),进一步提升了网络性能。本文采用 GhostConv 对网络模型中的主要组成模块 ((HG Block) 进行轻量化处理,得到模块 G-HG Block,并利用该模块构造网络架构 PG-HGNet,将其作为主干网络。

G-HG Block 的结构如图 2 所示。其中, Ghostconv 是在 GhostNet^[9] 中提出的一种卷积模型,其结构如图 3 所示。计算过程中,首先采用逐点卷积将特征通道数压缩至原来的一半,即图 3 中的 (a) 部分;其次根据第一步生成的特征图进行

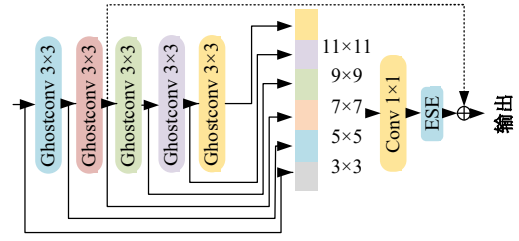


图 2 G-HG Block 结构

Fig. 2 Architecture of G-HG Block

深度卷积,生成等通道数的特征图,即图 3 中的 (b) 部分;最后进行特征图拼接,生成完整的输出特征图。这种方式减少了计算的冗余。相较于 HG Block, G-HG Block 通过引入 Ghostconv 减少了模块的参数量和计算量,因此 G-HG Block 模块将更加轻量化,进而提高了主干网络整体的轻量化水平,减小了网络复杂度。

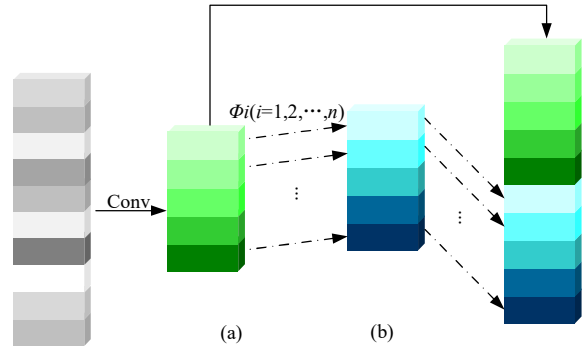


图 3 Ghostconv 结构

Fig. 3 Architecture of Ghostconv

1.2 多尺度特征增强模块 MFEM

针对遥感图像舰船检测中因目标尺度变化大导致检测精度下降的问题,本文设计了 MFEM,以提升网络对不同感受野的特征提取能力,从而提高对多尺度目标的检测精度。在卷积神经网络中,感受野是指某一层输出结果中某个元素所对应的输入层的区域大小。本文将卷积层定义为 $C(p, k, s, c)$, 其中 p, k, s 和 c 分别为补零数、卷积核大小、步幅和输出通道数。第 i 层的输出可表示为:

$$y_i = \sum_{j=0}^{i-1} \sum_{m=1}^{n_{i,j}^m} C_{i,j}^m \left(p_{i,j}^m, k_{i,j}^m, s_{i,j}^m, c_{i,j}^m \middle| y_j \right), i = 1, 2, \dots, n_n \quad (1)$$

式中: $n_{i,j}^m$ 为从第 j 层到第 i 层的卷积核类型数; n_n 为层数; y_0 为输入图像。 $C_{i,j}^m \left(p_{i,j}^m, k_{i,j}^m, s_{i,j}^m, c_{i,j}^m \middle| y_j \right)$ 中一个元素的感受野为^[10]:

$$R_{i,j}^m = R_{y_j} + (k_{i,j}^m - 1) \times \prod_{x=1}^j s_{x,x-1} \quad (2)$$

式中: R_{y_i} 为输入 y_i 中一个元素对应的感受野; $s_{x,x-1}$ 为第 $x-1$ 层到第 x 层的采样步幅。

由此可知,不同的卷积核在经过不同层数的卷积网络后,会具有不同的感受野,当使用大小为 3、步幅为 2 的卷积核时,经过 5 层网络后的感受野分别为 3、7、15、31、63。然而,感受野中并非所有的元素对输出单元有同等影响,有效感受野只占理论感受野的一小部分^[11],因此仅使用大小为 3 的卷积核只能捕捉到目标的局部特征。若使用更大的卷积核,如大小为 7 的卷积核,经过 5 层网络后的感受野依次为 7、19、43、91、187,有效感受

野相应增加,有助于全面覆盖目标对应的像素。因此,通过引入更大尺寸的卷积核,可以扩大有效感受野,进而增强网络对不同尺度目标的提取能力^[12]。

基于上述原理并考虑模型轻量化,本文设计了 MFEM,其结构如图 4 所示。MFEM 通过引入不同扩张率的空洞卷积,在不增加额外参数的情况下,有效扩大了感受野,提高了对不同尺度舰船目标的检测能力。同时,通过嵌入通道注意力机制,进一步优化了特征的表达力和选择性,使模型更好地适应复杂的遥感图像背景。

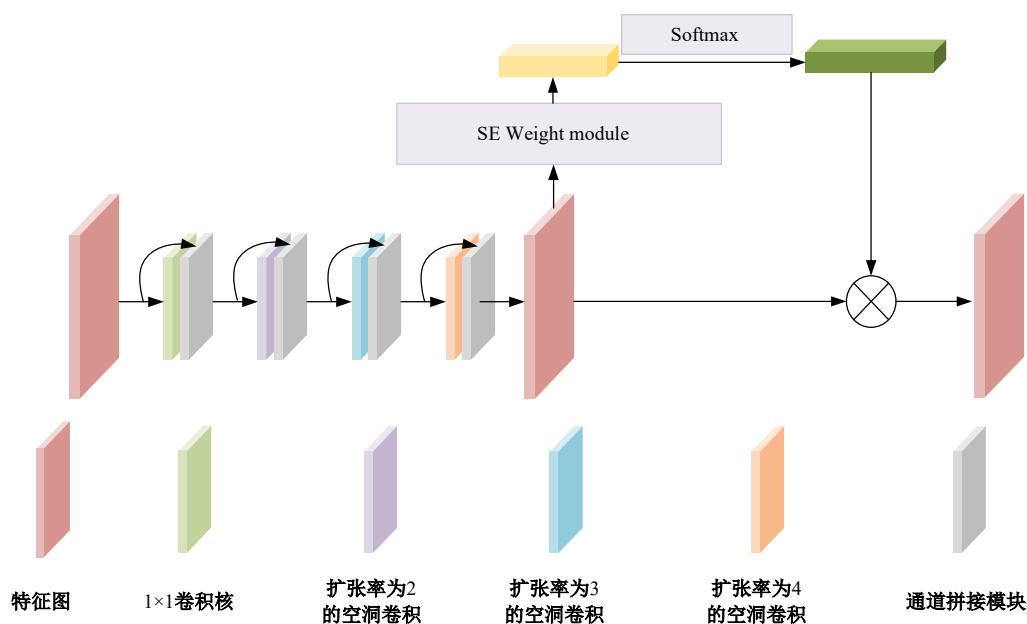


图 4 MFEM 结构

Fig. 4 Architecture of MFEM

在 MFEM 模块的设计过程中,本文结合了文献[13]提出的高效网络设计指导原则。首先,模块通过 1×1 卷积对输入特征图进行降维处理,以减小计算开销;其次,依次采用扩张率为 2、3、4 的空洞卷积逐层处理降维后的特征图,从而实现多尺度信息的有效捕获。具体来说,首先输入特征图通过 1×1 卷积降维,使通道数减少至原来的 $1/8$;其次该特征图经扩张率为 2 的空洞卷积处理,生成新的特征图;最后将这两个特征图在通道维度上进行拼接,形成通道数为输入特征图 $1/4$ 的合成特征图。此过程再重复两次,依次使用扩张率为 3、4 的空洞卷积,最终得到与输入特征图通道数相等的特征图。在此过程中,通过设置不同卷积核大小的卷积层和跨层通道拼接操作,可以有效提取不同尺度和深度的特征,同时通过参

数设置保证了模块内部卷积层的输入通道数与输出通道数相等,满足文献[13]所提输入通道数和输出通道数相等时降低内存访问成本的要求。最后,对经过多尺度特征提取的特征图施加通道注意力,通过在通道维度上进行信息交互,有效识别不同通道中特征的重要性差异,获取特征在不同通道维度上的重要性差异,从而使网络选择性地聚焦于关键信息通道,突出更为重要的特征,进一步增强网络的多尺度特征提取能力。

1.3 基于 BiFPN 的轻量化特征融合模块

深度学习网络中不同层之间的信息传播至关重要。深层网络的感受野较大,特征语义信息更为丰富,但特征图的分辨率低,细节不充分;浅层网络的感受野较小,特征图分辨率高,细节充分,但缺乏丰富的特征语义信息。因此,需要对不同

层次的特征进行融合。传统特征金字塔网络FPN(结构如图5(a)所示)通过一条自上而下的通路实现特征融合。路径聚合网络PAN(结构如图5(b)所示)基于传统特征金字塔网络进行改进,增加了一条自下而上的特征融合通路,通过提取更多的深层与浅层特征,取得了更好的检测效果,但增加了网络模型的复杂度。双向特征金字塔网络BiFPN(结构如图5(c)所示)在路径聚合网络的基础上进一步改进:首先删除仅有一个输入且贡献度较低的节点,有效降低网络的复杂性;其次引入跳跃连接机制,在同一尺度特征间增加特征融合路径,缓解网络层级过多造成的特征信息丢失问题,更好地将低层次的位置信息和高层次的语义信息进行融合,提升了网络的表征能力。

瓶颈结构被广泛应用于深度学习网络模型中,用于减小模型的计算复杂度,其核心策略是运用 1×1 卷积调整特征通道数,从而降低特征提取过程中的计算量^[14]。以Resnet中的Bottleneck Block为例(如图6所示),首先采用 1×1 卷积对输入特征图进行降维,降维后的特征图通过 3×3 卷积完成特征提取;其次通过 1×1 卷积恢复维度,并将结果与输入特征图进行融合,以此减少参数量并保持高计算效率。本文基于此思想和BiFPN结构,构造了LW-BiFPN,通过在输入端设置 1×1 卷积,对输入特征图进行通道降维,减少特征融合网络中的计算量,实现网络轻量化,同时增强网络的非线性处理能力。

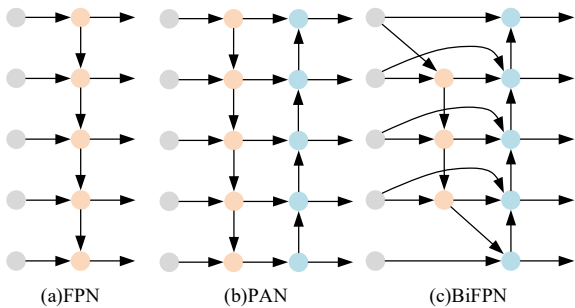


图5 特征融合网络结构

Fig. 5 Architecture of feature fusion network

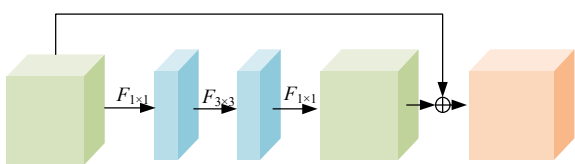


图6 Bottleneck Block结构

Fig. 6 Architecture of Bottleneck Block

1.4 MPDIoU 边框损失函数

在目标检测算法中,损失函数对网络模型的性能发挥着重要作用,合适的损失函数可以有效降低网络模型的定位误差,进而改善模型的检测精度。YOLOv8模型的损失函数主要由3类损失组成,即边框损失、类间损失、目标损失。其中,边框损失采用CIoU损失函数,该损失函数综合考虑了检测框的中心点距离、重叠比、长宽比等因素,其定义如下:

$$\text{CIoU} = \text{IoU} - \left(\frac{\rho^2(b, b^{\text{gt}})}{c^2} + \alpha v \right) \quad (3)$$

$$\alpha = \frac{v}{(1 - \text{IoU}) + v} \quad (4)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{\text{gt}}}{h^{\text{gt}}} - \arctan \frac{w}{h} \right)^2 \quad (5)$$

式中: b^{gt} 为真实框的中心点; b 为预测框的中心点; c 为覆盖预测框和真实框的最小封闭框的对角线长度; ρ 为真实框中心点到预测框中心点的欧氏距离; α 为权衡系数; v 用于衡量长宽比的一致性; w, h 和 $w^{\text{gt}}, h^{\text{gt}}$ 分别为预测框的长宽与真实框的长宽。

虽然CIoU引入了边框的长宽值,但是以长宽比的形式使用,而不是绝对值。因此,在遥感图像舰船检测场景中,由于舰船目标的尺度变化比较大,容易出现预测框与真实框长宽比一致但长宽绝对值不相等的情况。此时, $v=0$ 失去惩罚作用,损失函数难以被优化,使得在非极大值抑制过程中容易造成漏检现象,降低了边界框回归的准确性。为解决此问题,本文引入基于最小点距离的交并比(Minimum point distance based IoU, MPDIoU)^[15]损失函数替换CIoU计算边框损失,其示意图如图7所示。

图7中,红色框表示真实框,黄色框表示预测框, $(X_1^{\text{prd}}, Y_1^{\text{prd}})$ 和 $(X_2^{\text{prd}}, Y_2^{\text{prd}})$ 分别为预测框的左上角坐标和右下角坐标, $(X_1^{\text{gt}}, Y_1^{\text{gt}})$ 和 $(X_2^{\text{gt}}, Y_2^{\text{gt}})$ 分别为真实框的左上角坐标和右下角坐标, h, w 分别为输入图像的长和宽。MPDIoU的损失函数定义如下:

$$\text{MPDIoU} = \text{IoU} - \frac{d_1^2}{h^2 + w^2} - \frac{d_2^2}{h^2 + w^2} \quad (6)$$

$$\begin{cases} d_1 = (x_1^{\text{gt}} - x_1^{\text{prd}})^2 + (y_1^{\text{gt}} - y_1^{\text{prd}})^2 \\ d_2 = (x_2^{\text{gt}} - x_2^{\text{prd}})^2 + (y_2^{\text{gt}} - y_2^{\text{prd}})^2 \end{cases} \quad (7)$$

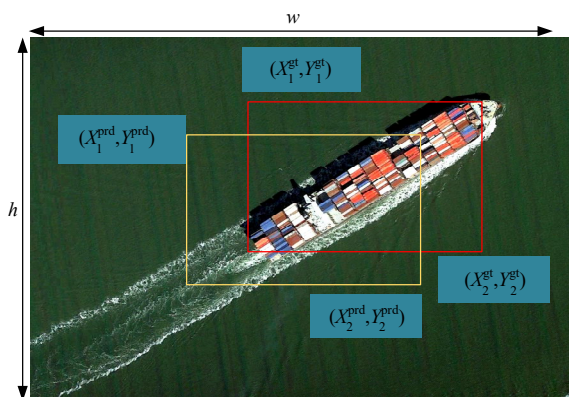


图7 MPDIOU示意图

Fig. 7 MPDIOU schematic

式中: d_1 为真实框与预测框左上角的距离; d_2 为真实框与预测框右下角的距离;IoU为真实框与预测框的交并比。

从式(6)和式(7)中可见,作为一种基于点距离的度量方法,MPDIOU损失函数解决了CIoU损失函数无法有效优化预测边界框与真实边界框具有相同长宽比但长宽绝对值不同的情况,能够更好地适应遥感图像舰船检测过程尺度差异大的应用场景,提高了模型检测准确率。

2 实验与分析

2.1 数据集

本实验以HRSC2016数据集^[16]作为网络模型性能验证的基础数据集,该数据集由从GoogleEarth中6个重要港口收集的高分辨率光学卫星图像数据构成,共1 061幅图像,包括2 976个样本,图像尺寸在 300×300 像素到 $1\,500 \times 900$ 像素之间。本文将数据集划分为训练集、验证集和测试集3部分,包含的图片数量分别为436、181和444张。除特别说明外,本文实验部分均基于该数据集进行。此外,为验证算法的泛化性能,本文还从由西北工业大学提供的大规模公开遥感目标检测数据集DIOR中提取了2 690张ship类别图片,并按照7:2:1的比例将所提取舰船图像数据集分为训练集1 882张、验证集538张、测试集270张。

2.2 实验环境配置及参数设置

本文实验在Ubuntu 22.04操作系统下进行,CPU采用AMD EPYC 9654 96-Core处理器,内存为60 GB,GPU为RTX 4090,显存为24 GB,所有深度学习模型均在Pytorch框架下构建,Python版本为3.8,CUDA版本为11.8。

网络训练阶段,训练轮次设置为300轮,每个批次样本量设置为16,采用随机梯度下降优化器,momentum设置为0.937,训练过程中采用Warm-Up方法预热学习率,初始学习率设置为0.01,学习率衰减权重为0.000 5,采用余弦退火学习率衰减方法自动调整学习率,不使用预训练权重。

2.3 实验内容

为确定MFEM模块融入网络中的最佳位置,设置不同融入位置的对比实验确定最佳融入位置;为直观展现MFEM模块对模型检测能力的提升效果,设置添加该模块前后的检测效果热力图对比实验;为探究所提算法中各模块对整个算法的贡献,进行消融实验;为验证所提优化算法的有效性,通过实验与其他6种常用的目标检测算法进行对比;为验证所提优化算法的有效性,通过实验分别与6种主流目标检测算法和6种先进遥感图像目标检测算法进行对比实验;为验证所提算法的泛化性,在Dior-ship数据集上进行泛化性实验,并与其他主流目标检测算法的性能效果进行对比。实验过程中,控制设备、训练超参数、迭代次数等实验条件保持一致,得出实验结果并进行分析。采用mAP@0.5和mAP@0.5:0.95作为衡量模型精度的指标,同时采用模型参数量、浮点数计算量(Giga floating-point operations per second, GFLOPs)作为衡量轻量级的指标。

2.3.1 添加MFEM模块数量及位置的对比实验

为确定MFEM融入网络中的最佳数量与位置,实现改进算法对舰船检测的最佳效果,本文开展了对比实验进行效果验证。

对比实验中将MFEM模块分别融入网络中的不同位置,开展4组对比实验。实验1在主干特征提取网络SPPF模块之前融入1个MFEM模块;实验2在主干网络3个输出端之后分别融入1个MFEM模块;实验3为实验1和实验2的集合,在主干特征提取网络SPPF模块之前和主干网络3个输出端之后均添加MFEM模块;实验4在特征融合网络中不同尺度信息的交汇处融入MFEM模块。实验结果如表1所示,从表1的结果中可以看到,相较于其他3组实验,实验4在检测精度及网络复杂度上都优于其他3组实验。因此,本文选择将MFEM融入特征融合网络中不同尺度信息的交汇处,以获得最佳性能表现。

表 1 MFEM 位置对比实验

Table 1 Comparative experiment of MFEM position

实验编号	mAP _{0.5} /%	mAP _{0.5:0.95} /%	参数量/M	FLOPs/G
1	89.8	77	1.50	6.0
2	89.7	75.2	1.76	7.3
3	90.2	76.5	1.97	7.4
4	90.9	78.8	1.35	6.2

2.3.2 MFEM 热力图可视化对比实验

为直观地展示 MFEM 模块的有效性,用 GradCAM 绘制加入 MFEM 模块前后的热力图进行可视化对比,结果如图 8 所示。热力图中,红色

区域的深浅程度反映算法对特定区域的关注度,颜色越深,表示该区域在算法中的权重越大,对最终检测结果的贡献越显著。在未加入 MFEM 模块的情况下,如图 8(b)所示,热力图中的部分红色区域分布在背景上,未能避免背景干扰,在背景上投入了一定关注度。图 8(c)为加入 MFEM 模块后的热力图,可以看出,热力图的红色区域几乎都集中在目标上,说明 MFEM 模块能够更好地捕获全局信息,能够更加精确地聚焦于检测船只,有效降低背景噪声的干扰。

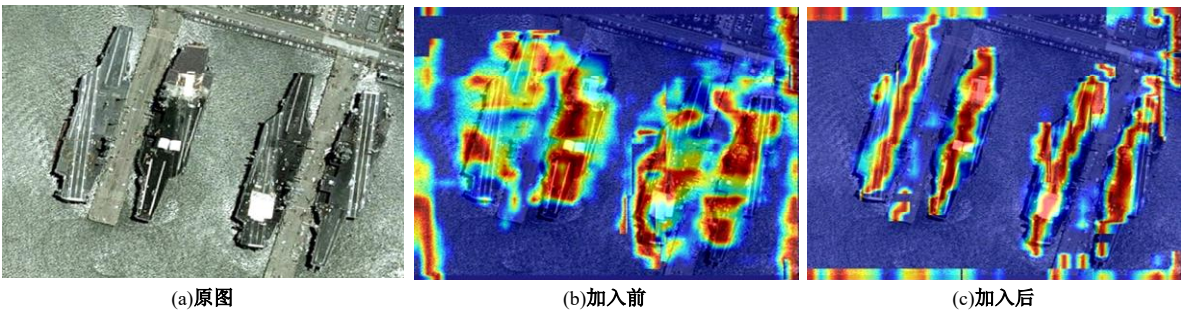


图 8 热力图对比实验

Fig. 8 Comparison experiment of heatmap

2.3.3 消融实验

为验证本文所提改进模块的有效性,设计了多组消融实验,消融实验结果见表 2。本文以原始 YOLOv8nano 模型作为验证和实验控制的基线。在实验 2 和实验 3 中,分别引入 PG-HGNet 和 LW-BiFPN,评估 PG-HGNet 和 LW-BiFPN 在模型轻量化方面的有效性。在实验 4 和实验 5 中,分别引入 MFEM 和 MPDIoU,评估 MFEM 和 MPDIoU 对精度的提升效果。实验 6 将实验 2 和实验 3 相结合,评估同时引入 PG-HGNet 和 LW-BiFPN 对模型复杂度的综合影响。实验 7 将实验

4 和实验 5 相结合,评估同时加入 MFEM 和 MPDIoU 对检测性能的影响。实验 8 将实验 6 和实验 7 相结合,评估本文提出的所有改进模块对模型的复杂度和检测精度的综合影响。

上述 8 组消融实验的结果证明,本文方法在保证网络检测精度的前提下减小网络复杂度的有效性。由表 2 可知,实验 2 通过将基线模型的主干网络替换为 PG-HGNet,在精度小幅下降的情况下,模型参数量和计算量分别下降 23.18% 和 16%,表明 PG-HGNet 能够在减少网络复杂度的同时保持良好检测性能。实验 3 通过引入 LW-BiFPN,在精度小幅度上升的同时,参数量和计算量分别下降 33.7% 和 12.5%,表明 LW-BiFPN 在降低网络复杂度和提升检测效果方面的优越性。实验 4 通过引入本文设计的 MFEM 模块,在参数量和计算量小幅度上升的同时提升了网络的检测精度,证明本文设计的 MFEM 模块对提升网络多尺度特征表征能力的有效性。实验 5 引入了 MPDIoU 损失函数,mAP@0.5 和 mAP@0.5:0.95 分别提升了 0.8 个百分点和 1.1 个百分点,表明其可以有效缓解遥感图像舰船检测过程中尺度变化带来的检测精度下降问题。实验 6 通过同时应用 PG-HGNet 和 LW-BiFPN 模块,参数量

表 2 消融实验

Table 2 Ablation experiment

模型	mAP _{0.5} /	mAP _{0.5:0.95} /	参数量/	FLOPs/
	%	%	M	G
基线	91.1	78.7	3.01	8.1
+PG-HGNet	89.8	76.1	2.31	6.8
+LW-BiFPN	91.7	79.1	1.99	7.1
+MFEM	92.2	80.7	3.33	8.7
+MPDIoU	91.9	79.8	3.01	8.1
+PG-HGNet+ LW-BiFPN	90.3	77.2	1.29	5.8
+MFEM+ MPDIoU	92.5	81.1	3.33	8.7
+all	90.9	78.8	1.35	6.2

和计算量分别下降了 57.1% 和 28.4%,而精度仅下降 0.8 个百分点,证明算法轻量化改进的有效性。实验 7 通过同时应用 MFEM 和 MPDIoU 边框损失函数,mAP@0.5 和 mAP@0.5:0.95 分别提升了 1.4 个百分点和 2.4 个百分点,证明其对模型精度提升的有效性。实验 8 同时应用所有改进模块,相比基线模型,mAP@0.5 下降 0.2 百分点,mAP@0.5:0.95 提升 0.1 个百分点,参数量下降 55%,计算量下降 23.5%,在保持网络检测精度的同时,极大减小了网络复杂度,证明本文算法在复杂度和精度平衡上的有效性。

2.3.4 对比实验

为验证本文算法在检测精度与网络复杂性上的优越性,与当前主流目标检测算法和先进遥感目标检测算法进行对比实验。第一组实验选取当前主流目标检测算法 Faster R-CNN、SSD、Yolov7-tiny、Yolov10nano 等进行对比实验。第二组对比实验中,将本文算法与其他先进遥感目标检测算法进行比较。

从表 3 的实验结果可知,本文算法在 HRSC2016 测试集上的 mAP@0.5 和 mAP@0.5:0.95 分别达到 90.9% 和 78.8%,mAP@0.5 相较于 Faster R-CNN、SSD、Yolov5nano、Yolov7-tiny、Yolov10nano 算法分别提高 0.7、26.4、0.1、0.3、3.8 个百分点,与 Yolov3-tiny 持平,mAP@0.5:0.95 相较于 Faster R-CNN、SSD、Yolov3-tiny、Yolov5nano、Yolov7-tiny、Yolov10nano 算法分别提高 13.4、41.1、4.2、1.5、7.6、4.6 个百分点,参数量分别为 Faster R-CNN、SSD、Yolov3-tiny、Yolov5nano、Yolov6n、Yolov7-tiny、Yolov10nano 的 0.99%、5.7%、11.15%、54.02%、31.94%、21.72%、50.2%,计算量分别为 Yolov3-tiny、Yolov5nano、Yolov6nano、Yolov7-ti-

表 3 主流目标检测算法对比实验

Table 3 Comparative experiment of mainstream object detection algorithm				
算法	mAP _{0.5} /	mAP _{0.5:0.95} /	参数量/	
	%	%	M	G
Faster R-CNN	90.2	65.4	136.69	369.7
SSD	64.5	37.7	23.61	60.8
Yolov3-tiny	90.9	74.6	12.13	18.9
Yolov5nano	90.8	77.3	2.50	7.1
Yolov7-tiny	90.6	71.2	6.23	13.9
Yolov10nano	87.1	74.2	2.69	8.2
本文	90.9	78.8	1.35	6.2

ny、Yolov10nano 的 32.8%、87.3%、52.5%、44.6%、75.6%。综合来看,本文算法在光学遥感舰船检测任务上实现了检测精度和网络模型复杂度的平衡,综合性能优于主流目标检测算法。

从表 4 与当前较为先进遥感目标检测算法的对比实验可知,在综合考量算法精度与复杂度的情况下,本文算法具有显著的优越性。SOCDet 是用于在轨目标检测的轻量化模型,其检测精度比本文算法高 3.38 个百分点,但其参数量和计算量分别是本文算法的 7.4 倍和 2.7 倍,轻量化水平不如本文算法。AOPG、LO-Det 的检测精度及轻量化水平均劣于本文算法。PP-PicoDet-S 的参数量略低于本文算法,但其计算量是本文算法的 2.25 倍,不具备优势。VOD-YOLOv7 的检测精度比本文算法高 1.2 个百分点,但其计算量是本文算法的 18.7 倍,在轻量化方面远逊于本文算法。

为验证本文算法的泛化能力,在 DIOR-ship 数据集上进行验证,实验结果如表 5 所示。表 5 所示实验结果与在 HRSC2016 数据集上所得结果类似,本文算法的检测效果略逊于 Yolov5nano 和基线模型 Yolov8nano(平均精确度差≤0.5 个百分点),优于其他对比方法,轻量化水平远优于其他方法,进一步验证了本文算法的通用性和有效性。

表 4 先进模型对比实验

Table 4 Comparative experiment of advanced model			
算法	mAP _{0.5} /%	参数量/M	FLOPs/G
SOCDet ^[17]	94.28	10.01	16.72
AOPG ^[18]	90.58	33.87	164.33
PP-PicoDet-S ^[19]	90.89	1.08	13.92
LO-Det ^[20]	90.17	6.78	10.98
VOD-YOLOv7 ^[21]	92.1	7.45	116
本文	90.9	1.35	6.2

表 5 基于 Dior-ship 数据集的泛化性实验

Table 5 Generalization experiment based on the Dior-ship dataset				
算法	mAP _{0.5} /	mAP _{0.5:0.95} /	参数量/	
	%	%	M	G
Yolov3-tiny	91.3	56.8	12.1	18.9
Yolov5nano	97.1	61	2.5	7.1
Yolov7-tiny	96.4	61.3	6.23	13.9
Yolov8nano	96.8	61.7	3	8.1
Yolov10nano	96.5	61.1	2.69	8.2
本文	96.6	60.9	1.35	6.2

3 结束语

本文提出了一种面向光学遥感图像舰船检测的多尺度特征增强轻量化检测算法(MFLDet)。首先,以PG-HGNet作为主干网络,并采用基于BiFPN结构的轻量化多尺度特征融合网络减小模型复杂度,使网络的参数量和计算量大幅下降。其次,通过引入自行设计的多尺度特征增强模块MFEM及MPDIoU损失函数,增强网络对遥感图像尺度差异大这一应用场景的检测能力,提升检测精度。改进后,本文算法在测试集上mAP@0.5达到90.9%,参数量为1.35 MB,计算量为6.5 GFlops;相较于基准模型,参数量下降55%,计算量下降23.5%。综合考量轻量化水平及检测精度,本文算法优于其他对比方法,适宜星上在轨舰船检测的应用场景。后续可对模型进行剪枝操作,进一步提高模型的轻量化水平。

参考文献:

- [1] 董超,刘晶红,徐芳,等. 光学遥感图像舰船目标快速检测方法[J]. 吉林大学学报:工学版, 2019, 49(4): 1369-1376.
Dong Chao, Liu Jing-hong, Xu Fang, et al. Fast ship detection in optical remote sensing images[J]. Journal of Jilin University (Engineering and Technology Edition), 2019, 49(4): 1369-1376.
- [2] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector[C]//14th European Conference on Computer Vision-ECCV 2016. Cham, Switzerland: Springer International Publishing, 2016: 21-37.
- [3] Ren S Q, He K, 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.
- [4] Cai Z W, Vasconcelos N. Cascade R-CNN: delving into high quality object detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2018: 6154-6162.
- [5] 李永刚,朱卫纲,黄琼男,等. 复杂背景下SAR图像近岸舰船目标检测[J]. 系统工程与电子技术, 2022, 44(10): 3096-3103.
Li Yong-gang, Zhu Wei-gang, Huang Qiong-nan, et al. Near shore ship target detection with SAR images in complex background[J]. Systems Engineering and Electronics, 2022, 44(10): 3096-3103.
- [6] 成倩,李佳,杜娟. 基于YOLOv5的光学遥感图像舰船目标检测算法[J]. 系统工程与电子技术, 2023, 45(5): 1270-1276.
Cheng Qian, Li Jia, Du Juan. Ship target detection algorithm of optical remote sensing image based on YOLOv5[J]. Systems Engineering and Electronics, 2023, 45(5): 1270-1276.
- [7] Yi H, Liu B, Zhao B, et al. Small object detection algorithm based on improved YOLOv8 for remote sensing[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2023, 17: 1734-1747.
- [8] 车思文,汪宇玲. 一种改进的YOLOv7光学遥感图像舰船目标检测算法[J]. 电光与控制, 2024, 31(5): 34-39.
Che Si-wen, Wang Yu-ling. An improved YOLOv7-based ship target detection algorithm for optical remote sensing images[J]. Electronics Optics & Control, 2024, 31(5): 34-39.
- [9] Han K, Wang Y H, Tian Q, et al. GhostNet: More features from cheap operations[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2020: 1580-1589.
- [10] Le H, Borji A. What are the receptive, effective receptive, and projective fields of neurons in convolutional neural networks?[J/OL]. [2024-09-11]. <https://arxiv.org/abs/1705.07049>
- [11] Luo W, Li Y, Urtasun R, et al. Understanding the effective receptive field in deep convolutional neural networks[J/OL]. [2024-09-11]. <https://arxiv.org/abs/1701.04128>
- [12] Ding X H, Zhang X Y, Han J G, et al. Scaling up your kernels to 31×31 : revisiting large kernel design in CNNs[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2022: 11953-11965.
- [13] Ma N N, Zhang X Y, Zheng H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design[C]//Proceedings of the European Conference on Computer Vision. Cham, Switzerland: Springer Interaction Publishing, 2018: 116-131.
- [14] Xie X R, Sun F X, Lin J, et al. Fast-ABC: a fast ar-

- chitecture for bottleneck-like based convolutional neural networks[C]//2019 IEEE Computer Society Annual Symposium on VLSI. Piscataway, NJ: IEEE, 2019: 1-6.
- [15] Ma S L, Xu Y. MPDIoU: a loss for efficient and accurate bounding box regression[J/OL]. [2024-09-11]. <https://arxiv.org/abs/2307.07662>
- [16] Liu Z K, Yuan L, Weng L B, et al. A high resolution optical satellite image dataset for ship recognition and some new baselines[C]//International Conference on Pattern Recognition Applications and Methods. New York: INSTICC, 2017, 2: 324-331.
- [17] Pang Y H, Zhang Y M, Kong Q L, et al. SOCDet: a lightweight and accurate oriented object detection network for satellite on-orbit computing[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 1-15.
- [18] Cheng G, Wang J B, Li K, et al. Anchor-free oriented proposal generator for object detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-11.
- [19] Li H, Zhang S, Li X, et al. DetectorNet: Transformer enhanced spatial temporal graph neural network for traffic prediction[C]//Proceedings of the 29th International Conference on Advances in Geographic Information Systems. New York: ACM, 2021: 133-136.
- [20] Zhou L, Wei H R, Li H, et al. Arbitrary-oriented object detection in remote sensing images based on polar coordinates[J]. IEEE Access, 2020, 8: 223373-223384.
- [21] Yang J R, Wu B, Zhang H Y, et al. Rotating frame labeling and improved YOLOv7-based algorithm for vehicle object detection[C]//Proceedings of the 2023 IEEE 6th International Conference on Pattern Recognition and Artificial Intelligence. Piscataway, NJ: IEEE, 2023: 170-175.