

文章编号 1004-924X(2025)08-1289-14

## 超融合残差行进几何感知的遥感目标检测

白晨帅, 白晓凤, 邬开俊\*, 王昊雯

(兰州交通大学 电子与信息工程学院, 甘肃 兰州 730070)

**摘要:** 本文提出超融合残差行进几何感知算法, 旨在解决遥感图像目标检测中的多尺度、密集重叠及数据分布不均等挑战。超融合残差行进模块优化网络结构, 其多层次卷积操作利用不同尺度感受野, 能捕捉目标各尺度细节, 增强模型对目标特征的感知能力, 实现小尺度目标特征提取和大尺度目标准确定位。通过计算检测与真实结果的几何相似度精准评估检测效果, 在目标密集重叠场景精细考量契合度, 筛选最终结果, 减少漏检、误检, 提高算法准确性; 设计多路径特征融合模块, 融合不同层次特征信息, 提取更丰富目标特征, 增强网络表示与判别能力, 提升检测准确性与稳定性。在NWPU-VHR-10数据集的实验结果显示, 其 mPrecision, mRecall, mAP 和 mF1 Score 分别提高了 0.041 9, 0.104 0, 0.045 5 和 0.085 0; 在RSOD数据集的实验结果显示, 其 mPrecision, mRecall, mAP 和 mF1 Score 分别提高了 0.022 1, 0.103 4, 0.061 9 和 0.087 5。充分证明了所提出超融合残差行进几何感知算法在遥感图像目标检测领域的有效性和优越性。

**关键词:** 遥感图像; 目标检测; 超融合残差行进模块; 几何相似度; 多路径特征融合

**中图分类号:** TP391 **文献标识码:** A

**doi:** 10.37188/OPE.20253308.1289

**CSTR:** 32169.14.OPE.20253308.1289

## Remote sensing object detection algorithm based on ultra fusion residual marching geometric perception

BAI Chenshuai, BAI Xiaofeng, WU Kaijun\*, WANG Haowen

(School of Electronic and Information Engineering, Lanzhou Jiaotong University,  
Lanzhou 730070, China)

\* Corresponding author, E-mail: wkj@mail.lzjtu.cn

**Abstract:** This paper proposed an ultra-fusion residual marching geometric perception algorithm, which aimed to solve the challenges of multi-scale, dense overlap, and uneven data distribution in remote sensing image object detection. The hyper-fusion residual marching module optimized the network structure, and its multi-level convolution operation used different scale receptive fields to capture the details of each scale of the object, enhance the model's perception of the object features, and achieve small-scale object feature extraction and large-scale object accurate positioning. The detection effect was accurately evaluated by calculating the geometric similarity between the detection and the real results, and the fit was carefully considered in the dense overlapping scene of the object, so as to screen the final results, reduce missed detection

**收稿日期:** 2024-08-30; **修订日期:** 2024-11-08.

**基金项目:** 甘肃省自然科学基金项目(No. 23JRRA913); 甘肃省重点研发项目(No. 25YFGA047); 内蒙古自治区重点研发与成果转化计划项目(No. 2023YFDZ0043, No. 2023YFDZ0054, No. 2023YFSH0043); 甘肃省教育厅: 优秀研究生“创新之星”项目(No. 2025CXZX-632)

and false detection, and improve the mAP of the algorithm. A multi-path feature fusion module was designed to fuse different levels of feature information, extract richer object features, enhance network representation and discrimination capabilities, and improve detection mAP and stability. The experimental results on the NWPU-VHR-10 data set showed that mPrecision, mRecall, mAP and mF1 Score were increased by 0.041 9, 0.104 0, 0.045 5 and 0.085 0, respectively. The experimental results on the RSOD data set show that mPrecision, mRecall, mAP, and mF1 Score are increased by 0.022 1, 0.103 4, 0.061 9, and 0.087 5, respectively. The effectiveness and superiority of the proposed ultra-fusion residual marching geometric perception algorithm in the field of remote sensing image object detection are fully proved.

**Key words:** remote sensing images; object detection; geometric similarity; multipath feature fusion; ultra-fusion residual marching module

## 1 引 言

基于深度学习的方法在遥感图像目标检测领域已取得了显著进展<sup>[1]</sup>。多源数据融合<sup>[2]</sup>、语义信息<sup>[3]</sup>引入等技术不断完善,极大地提升了检测精度和效率。针对小样本目标检测<sup>[4]</sup>和遥感图像超分辨率重建<sup>[5]</sup>等方向也取得了新的突破。结合多模态数据和跨领域知识<sup>[6-7]</sup>,将进一步推动遥感图像目标检测技术的发展。

在这一蓬勃发展的研究进程中,众多学者纷纷提出了各具特色的创新方法。吴等人<sup>[8]</sup>为解决遥感图像检测问题,设计了带通道注意力的主干网络模块与模糊池模块,并且采用无跨步卷积的特征金字塔网络以及结合简单无参数注意力模块(Simple Parameter-free Attention Module, SimAM),以此增强跨尺度融合能力,同时利用空间和尺度感知的加权交并比(Weighted Intersection over Union with Spatial and Scale Awareness, Wise-IoU)优化样本不均衡状况。Li等人<sup>[9]</sup>通过巧妙修改 Transformer 实现全局特征聚合和实例交互建模,并且引入基于注意力机制来有效减少数据集之间的差异,还将数据增强与 Transformer 相结合,进而提升对象检测性能。Zhang 等人<sup>[10]</sup>设计了多模态融合以提取补充信息,同时引入灵活的超分支来学习高分辨率特征表示,从而提高小目标的检测精度。Wan 等人<sup>[11]</sup>利用多层特征金字塔、多检测头策略以及混合注意力模块来改善光学遥感图像中目标检测网络的效果,取得了显著的性能提升。Liu 等人<sup>[12]</sup>在 YOLOv5s 的基础上,引入注意力机制以增强空

间和信道注意力,并运用内容感知的特征重组模块来重新组装特征点。Cao 等人<sup>[13]</sup>提出一种改进的目标检测架构,通过引入完全交并比(Complete-IoU, CIoU)改进先验框生成算法,结合修改后的网络结构和注意力机制以及有效的特征融合方法。Xiao 等人<sup>[14]</sup>通过双向特征融合、特征解耦和定位细化模块的设计,实现了更为精确的对象定位和分类。Chen 等人<sup>[15]</sup>通过引入多核膨胀卷积块、Transformer 块和深度可分离卷积,有力地增强了小物体的本体和相邻空间特征,同时减小了计算成本,旨在有效解决遥感图像中小物体检测和复杂背景下的挑战。Qian 等人<sup>[16]</sup>通过引入语义分割引导的伪标签挖掘模块和实例检测模块,来挖掘高质量的伪地面真实实例并提升模型的定位精度,从而成功解决遥感图像中弱监督目标检测<sup>[17-18]</sup>问题。

然而,当前遥感图像目标检测领域虽创新方法不断涌现,但仍有问题和挑战。(1)部分算法追求高准确性时增加了模型参数和计算量,导致实际应用中难以高效部署和推断。(2)传统评价指标有局限性,无法精准反映检测结果与真实结果的重合程度,易导致漏检和误检频发。为此,本文提出一种超融合残差行进几何感知(YOLOv7 Ultra Fusion Residual Progression Geometric Perception, YOLO-UG)的目标检测算法,首先,该算法利用超融合残差行进模块(Ultra-Fusion Residual Progression Module, UFRPM),其多层次卷积操作利用不同尺度感受野,能捕捉目标各尺度细节,增强模型对目标特征的感知能力,实现小尺度目标特征提取和大尺度目标准确定位。

其次,利用几何融合指数(Geometric Fusion Index, GFI)来衡量目标检测结果与真实结果之间的重合程度。通过GFI操作,可以更准确地评估目标检测结果,并筛选最终的检测结果,从而减少漏检和误检的情况。最后,利用多路径特征融合(Multipath Feature Fusion, MFF)模块。将不同层次的特征信息进行融合,从而增强网络对于复杂场景的建模能力,通过融合不同层次的特征信息,提取出更丰富的目标特征,提高网络对目标的表示能力和判别能力。

## 2 相关工作

### 2.1 YOLO系列改进算法

随着遥感技术的广泛应用,遥感图像目标检测对精度和效率的要求日益提升。YOLO系列算法因具备快速检测优势在目标检测领域备受瞩目,众多研究者基于此系列开展了针对遥感图像目标检测的改进工作,以提升其在该领域的性能。Shen等人<sup>[19]</sup>基于YOLOv5,浅层加坐标注意力模块强化细节特征提取,深层特定池化,采用随机池化策略融合多尺度特征并优化锚框机制与损失函数,提升不同目标检测能力。Liu等人<sup>[20]</sup>提出基于目标特征增强和边界框辅助回归的方法,含相关模块及损失函数,提升遥感图像准确率并验证有效性。Tian等人<sup>[21]</sup>运用改进K-means++算法获准确锚框,在YOLOv5网络引入注意力机制和双向特征金字塔网络,增强特征提取与融合能力,成果显著。付等人<sup>[22]</sup>针对光学遥感图像检测效果不佳,提出融合多注意力机制方法,设计坐标注意力模块并改进YOLOv5预测头。许等人<sup>[23]</sup>改进YOLOv6算法,骨干网络引入坐标注意力模块与上下文增强模块,颈网络引入自适应空间特征融合,提升多尺度目标检测精度。苗等人<sup>[24]</sup>加入双级路由注意力机制等构建新模块并优化损失函数,提出有效检测方法。王等人<sup>[25]</sup>改进YOLOv7算法,采用尺度交并比(Scaled Intersection over Union, SIoU)损失函数等举措提升小目标检测性能。陈等人<sup>[26]</sup>针对光学遥感图像检测问题,在YOLOv7基础上增加小目标检测层改进,提升多尺度目标检测能力并减少误检、漏检。周等人<sup>[27]</sup>针对航空遥感图像小目标检测设计相关模块增强特征表达、解决语义差

异及增强对重要目标关注。梁等人<sup>[28]</sup>改进YOLOX算法,主干网络引入相关模块增强浅层特征并提取上下文信息、抑制噪声,提升复杂背景下目标检测精度。这些基于YOLO系列的改进研究从特征提取、融合、损失函数优化、网络结构调整等角度,优化了YOLO系列算法在遥感图像目标检测中的应用。

### 2.2 特征增强融合与注意力机制相关算法

在遥感图像目标检测领域,特征增强融合与有效利用注意力机制对提升检测性能意义重大,能应对目标的多样、复杂及尺度变化等问题,许多研究者据此开展了相关算法研究。Zhang等人<sup>[29]</sup>针对遥感小目标检测提出三个轻量级模块,可提升局部感知、多尺度融合、全局关联能力,增强小目标特征表达并抑制背景干扰,提供轻量级有效方案。陈等人<sup>[30]</sup>针对目标尺度小且纵横比例不均衡的检测难题,提出多阶段特征融合方法,一阶段组合拆分通道融合特征,二阶段设计非对称卷积块,提升了检测效果。王等人<sup>[31]</sup>针对多尺度目标检测难题,提出无锚框算法,通过自适应多尺度特征融合与注意力特征增强处理特征,得到检测结果,提供新途径并避锚框设置问题。Jiang等人<sup>[32]</sup>采用YOLOv5模型引入通道注意力机制,显著提升小目标和密集目标检测效果,证明引入该机制可提升特定目标检测性能。Li等人<sup>[33]</sup>针对遥感目标检测难题,提出创新方法,引入全局和局部注意机制并结合,开发多头预测模块捕捉多尺度特征,设计损失函数提升小目标定位精度,提供新思路。这些算法研究通过不同的特征处理方式,有效提升了遥感图像目标检测在不同尺度、比例、目标密度等情况下的性能,为优化算法提供重要参考,展示了运用特征处理技术应对检测难题的可行性。

### 2.3 其他针对性改进算法

遥感图像目标检测面临诸多挑战,除基于YOLO系列改进及特征相关研究外,许多研究者针对不同具体问题提出特色改进算法以全面解决实际问题。Ma等人<sup>[34]</sup>针对小物体检测难题提出模型,含自适应导数掩码模块与样本重要性评估策略,可消除浅层大物体显著响应并解决交并比(Intersection over Union, IoU)损失函数局限,有效提高小物体检测准确性。Xi等人<sup>[35]</sup>通过引

入特征金字塔网络、设计多尺度感受野模块及引入额外浅层特征图等措施,提升了航空遥感目标检测性能。陈等人<sup>[36]</sup>提出级联式逆残差卷积结构,采用深度可分离卷积等操作并引入角度变量,有效提高了检测速度。汪等人<sup>[37]</sup>通过通道拼接融合特征图及相关处理强化目标区域特征,提升了检测性能。Yu等人<sup>[38]</sup>提出新型面向空间的目标检测框架,考虑对象多种特征变化,分两阶段设计,在遥感图像检测中效果显著,提供新检测模式。Sharma等人<sup>[39]</sup>结合YOLO算法优化多模态遥感图像,使其能多尺度检测对象并引入新融合架构,提升了多模态遥感图像目标检测性能。Li等人<sup>[40]</sup>为提高模型性能,结合实例感知的特征蒸馏技术并引入相关模块,有效处理复杂遥感图像,提升了检测模型性能。Chen等人<sup>[41]</sup>引入特定模块解决大尺度对象变化挑战,提升了面对该挑战时的检测性能。Zhang等人<sup>[42]</sup>提出新检测框架,含两个分支以增强特征表示与选择训练样本,解决多种挑战,提升了相应检测性能。Yue等人<sup>[43]</sup>为解决建模问题构建局部校正模块和非局部聚焦模块,提升了目标检测性能。这些针对性改进算法从不同角度出發,针对如小物体检

测、检测速度、复杂环境处理、大尺度对象变化等具体问题進行改进,有效提升了遥感图像目标检测应对各种实际问题的性能,提供了丰富技术手段以全面解决该领域实际问题。

### 3 YOLO-UG 模型

本文提出的YOLO-UG模型基于:(1)采用多尺度特征融合技术,通过在不同层级的特征图中提取信息。(2)自适应注意力机制,使模型能根据目标的重要性动态调整特征图中的注意力分布。(3)采用数据增强技术,通过对原始数据进行旋转、缩放、裁剪等处理,增加了训练样本的多样性,提升了模型的泛化能力。(4)综合利用深度学习技术和传统图像处理方法,对目标检测结果进行后处理,进一步提高了检测精度和稳定性。

如图1所示,通过利用UFRPM、GFI以及MFF模块能使YOLO-UG的遥感图像目标检测算法更加灵活地适应不同尺度和位置的目标,同时提高算法的计算效率,从而在面對遥感图像中目标多样性和复杂性的情况下,提高准确性和鲁棒性。

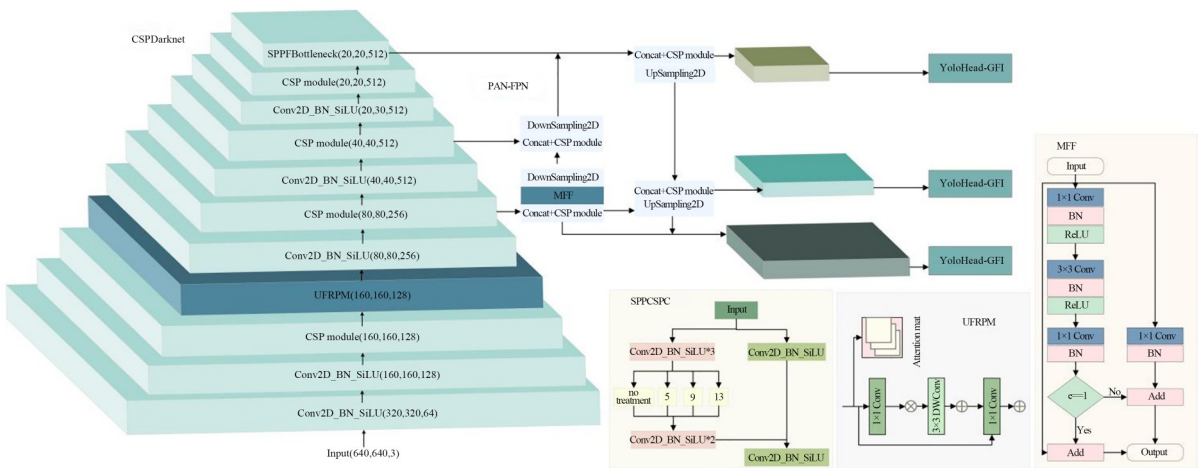


图1 YOLO-UG 总体网络架构  
Fig. 1 YOLO-UG overall network architecture

#### 3.1 超融合残差行进模块

如图2所示,UFRPM在处理多尺度目标方面具有关键作用。传统网络模块在面對尺度变化大的目标时,提取特征效果差,而UFRPM模块通过独特结构设计,聚焦于不同尺度目标的关

键特征。超融合残差行进模块基于归纳元移动块设计,从微观角度构建高效模块结构。归纳元移动块能充分利用不同尺度的感受野,通过多层次卷积操作对输入特征图进行细致特征提取。这种设计可有效捕捉目标在不同尺度下的细节

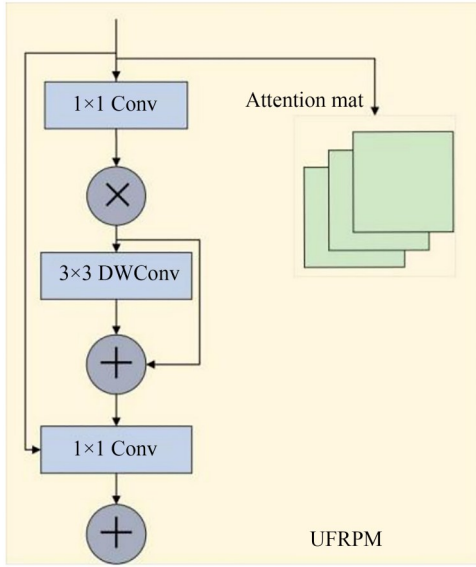


图 2 超融合残差行进模块

Fig. 2 Ultra-fusion residual progression module

信息,增强模型对目标的局部特征和全局上下文信息的感知能力,无论是小尺度目标的微弱特征提取,还是大尺度目标的准确定位和识别都能有效实现。

在设计原则方面,UFRPM中的 $F$ 被建模为级联多头自注意力网络(Multi-Head Self-Attention, MHSA)和卷积操作,公式为:

$$F(\cdot) = Conv(MHSA(\cdot)). \quad (1)$$

本设计吸收了类似卷积神经网络(Convolutional Neural Network, CNN)的局部特征建模效率和类似Transformer的动态建模能力来学习远距离交互。 $\lambda$ 一般大于1,中间维数是输入维数的倍数,导致参数和计算量二次 $\lambda$ 递增。因此, $F$ 的分量应该是独立的或线性依赖于通道的数量。MHSA的FLOPs与总图像像素的平方成正比。

为了优化模型成本和精度,采用了窗口高效的多头自注意力(Window Multi-Head Self-Attention, WMHSA)和深度卷积(Depth-Wise Convolution, DWConv),并采用跳跃连接来权衡模型成本和精度<sup>[44]</sup>。W-MHSA中获取Q、K的参数和FLOPs是通道的平方,因此采用未展开的 $X$ 更有效地计算注意力矩阵,即 $Q=K=X(\in R^{C \times H \times W})$ ,而展开的值 $X_e$ 为 $V(\in R^{C \times H \times W})$ 。这种改进被称为扩展窗口多头注意力机制(Expanded Window Multi-Head Self-

Attention, EWMHSA),表述如下:

$$F(\cdot) = (DWConv, Skip)(EWMHSA(\cdot)). \quad (2)$$

这种级联方式可以提高感受野的扩展速度,并将模型的最大路径长度减小到 $O(2W/(k-1+2W))$ 。

MHSA通常用于通道一致投影( $\lambda=1$ ),即注意矩阵乘以扩展 $X_e(\lambda>1)$ 的FLOPs增加 $\lambda-1$ 。利用从 $X$ 到扩展的 $V(X_e)$ 信息流只涉及线性运算,即 $MLP_e(\cdot)$ ,因此可以推导出一个等价命题:“当 $MLP_e$ 的组等于WMHSA的头数时,交换顺序的乘法结果保持不变。”为了减少FLOPs,默认情况下在 $MLP_e$ 之前使用矩阵乘法。

### 3.2 几何融合指数

如图3所示,GFI模块是解决目标密集重叠问题的关键。在目标密集场景下,传统检测方法易出现误检和漏检情况。GFI模块通过衡量目标检测结果与真实结果之间的重合程度,为检测结果评估提供更精细依据,从而有效筛选出准确的检测结果,减少因目标重叠而产生的误判。GFI通过计算两个边界框(Bounding Box)的交集面积除以它们的并集面积来衡量其重叠程度。

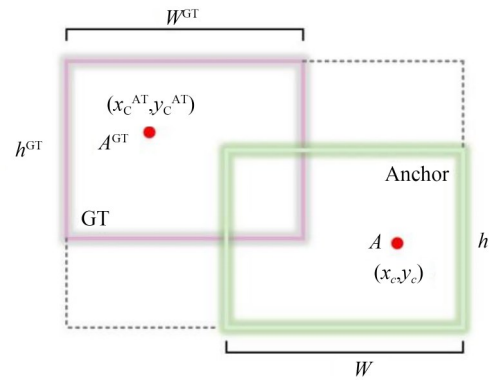


图 3 GFI直观图

Fig. 3 GFI intuitive graph

$$FI = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|}. \quad (3)$$

还涉及一些与尺度因子、水平和垂直方向权重系数相关的计算<sup>[45]</sup>。其中scale为尺度因子,与数据集中目标的尺度有关,水平方向权重系数如式(4),垂直方向权重系数如式(5),同时还有式(6)~式(8):

$$\omega\omega = \frac{2 \times (\omega^{\text{gt}})^{\text{scale}}}{(\omega^{\text{gt}})^{\text{scale}} + (h^{\text{gt}})^{\text{scale}}}, \quad (4)$$

$$hh = \frac{2 \times (h^{\text{gt}})^{\text{scale}}}{(\omega^{\text{gt}})^{\text{scale}} + (h^{\text{gt}})^{\text{scale}}}, \quad (5)$$

$$\text{distance}^{\text{shape}} = hh \times \frac{(x_c - x_c^{\text{gt}})^2}{c^2} + \omega\omega \times \frac{(y_c - y_c^{\text{gt}})^2}{c^2}, \quad (6)$$

$$\Omega^{\text{shape}} = \sum_{t=w,h} 1 - e^{-\omega t^2} \theta = 4, \quad (7)$$

$$\begin{cases} \omega_\omega = hh \times \frac{|\omega - \omega^{\text{gt}}|}{\max(\omega, \omega^{\text{gt}})} \\ \omega_h = \omega\omega \times \frac{|h - h^{\text{gt}}|}{\max(h, h^{\text{gt}})} \end{cases}. \quad (8)$$

其对应的边界盒回归损失为:

$$L_{\text{GFI}} = 1 - \text{IOU} + \text{distance}^{\text{shape}} + 0.5 \times \Omega^{\text{shape}}. \quad (9)$$

通过这些计算和参数, GFI 模块能更准确地评估目标检测结果在密集重叠情况下的质量, 进而提高检测准确性。

### 3.3 多路径特征融合

如图 4 所示, 在处理多尺度和数据分布不均问题时, MFF 模块发挥了关键作用。MFF 模块被添加在 YOLOv7 目标检测算法 FPN 层的 P3 的下采样拼接中间部分。数据分布不均会使目标特征呈现出多样性和复杂性, 而 MFF 模块的设计旨在缓解这些问题对目标检测的影响。

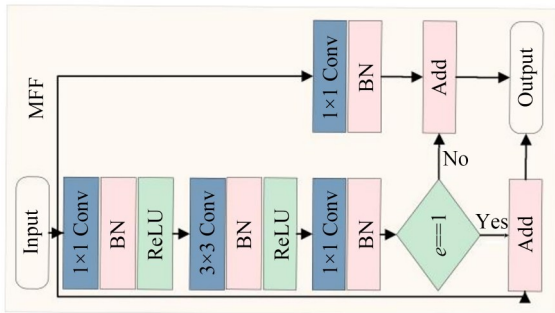


图 4 多路径特征融合模块

Fig. 4 Multipath feature fusion module

在处理多尺度目标方面, MFF 模块有着独特的机制。模块通过巧妙融合网络中不同层次的特征信息来应对多尺度问题。不同层次的特征图对不同尺度的目标有着不同的感知。MFF 模块中的卷积层和批归一化层的组合操作实现

了这种融合。Conv1: 使用  $1 \times 1$  的卷积核对输入特征图  $X$  进行卷积, 将输入通道数转换为输出通道数。Conv2: 对经过 Conv1 处理后的特征图进行卷积操作。  $3 \times 3$  的卷积核可以覆盖局部区域, 有效提取更丰富的特征信息。Conv3: 再次使用  $1 \times 1$  的卷积核进行卷积操作, 将特征图深度调整为输出通道数。

在 MFF 模块中, 每个卷积操作后都配备了批归一化层 (Batch Normalization, BN)。批归一化在解决数据分布不均问题上发挥了重要作用。在训练过程中, 数据分布不均可能导致模型训练不稳定, 收敛速度慢等问题。BN 层通过对经过卷积操作后的特征进行归一化处理, 使得数据在每个批次内的分布更加稳定和均匀。恒等映射根据参数  $e$  的取值, 调整输入和输出特征图的融合方式。若  $e = 1$ , 则直接将输入特征图与经过一系列卷积和批归一化操作后的输出特征图相加。若  $e \neq 1$ , 先对输入特征图通过卷积和批归一化操作进行调整后再与输出特征图相加。最后, 通过 ReLU 激活函数对经过上述处理得到的特征图进行激活, 得到最终的输出。

假设输入的特征图为  $X$ , 经过模块处理后得到的输出特征图为  $Y$ 。路径融合过程, 通过三个卷积操作和批归一化层, 可以表示为:

$$\begin{aligned} z_1 &= \text{ReLU}(\text{BN}(\text{Conv1}(X))), \\ z_2 &= \text{ReLU}(\text{BN}(\text{Conv2}(z_1))), \\ z_3 &= (\text{BN}(\text{Conv3}(z_2))), \end{aligned} \quad (10)$$

其中: ( $\text{Conv1}, \text{Conv2}, \text{Conv3}$ ) 分别表示三个卷积操作。 ( $z_1, z_2, z_3$ ) 分别表示三个中间特征图。

恒等映射: 如果  $e = 1$ , 则执行恒等映射, 直接将输入特征图与最终的输出特征图相加, 表示为:

$$Y = z_3 + X. \quad (11)$$

如果  $e \neq 1$ , 则先对输入特征图执行卷积和批归一化操作, 然后再与最终的输出特征图相加, 表示为:

$$Y = z_3 + \text{BN}(\text{Conv}(X)). \quad (12)$$

最后, 将输出特征图  $Y$  输入到 ReLU 激活函数中, 得到最终的输出:

$$\text{Output} = \text{ReLU}(Y). \quad (13)$$

## 4 实验结果与分析

### 4.1 评价指标介绍

Precision(查准率):查准率是指分类器预测为正例的样本中,实际为正例的比例,它衡量了模型在预测为正例的样本中的准确性。计算公式如下:

$$Precision = TP / (TP + FP), \quad (14)$$

其中:TP代表真正例(True Positive),FP代表假正例(False Positive)。

Recall(查全率):查全率是指实际为正例的样本中,被分类器正确预测为正例的比例,它衡量了模型对正例样本的查全率。计算公式如下:

$$Recall = TP / (TP + FN), \quad (15)$$

其中:TP代表真正例(True Positive),FN代表假反例(False Negative)。

平均精度均值(mean Average Precision, mAP)用于衡量模型在多个类别上的平均性能。首先计算每个类别的平均精度(Average Precision, AP),然后对所有类别的AP取平均值得到mAP,mAP值越高,表明模型性能越好。

F1 Score:F1 Score是Precision和Recall的综合评价指标,它是二者的调和平均值。F1 Score综合考虑了分类器的准确性和查全率,并且对Precision和Recall给予相同的权重。计算公式如下:

$$F1 = \frac{2 \times (Precision \times Recall)}{(Precision + Recall)}. \quad (16)$$

F1 Score的取值范围在0和1之间,值越接近1表示模型的性能越好。F1 Score适用于不平衡类别分布的情况,当分类器在正例和反例上的

表现差异较大时,F1 Score能更好地评估模型的性能。

参数量(Params):在遥感图像目标检测中,参数量指的是检测模型中需要训练的参数总数。这些参数包括权重(weights)和偏置(biases)等,用于在训练过程中学习数据的特征表示。

复杂度(complexity):十亿次乘加运算(Giga Multiply-Accumulate operations, GMac)是一种衡量模型计算复杂度的指标,表示十亿次乘加运算。

### 4.2 数据集介绍

如图5所示,NWPU-VHR-10与RSOD遥感图像数据集,分别应用于不同的目标检测任务。训练集:验证集:测试集=8:1:1。NWPU-VHR-10数据集于2014年由西北工业大学发布,包含800张高分辨率光学遥感图像,分为650张正样本(每张至少含一个目标)和150张负样本(不含给定目标)图像。图像包括715幅采集自谷歌地球、分辨率0.5~2 m的RGB图像和85幅来自Vaihingen数据、分辨率0.08 m的pan-锐化彩色红外图像。其目标类别有757架飞机、390个棒球场、159个篮球场、124座桥梁、224个港口、163个田径场、302艘轮船、655个储罐、524个网球场、477辆汽车10种,共3775个对象实例。RSOD数据集由武汉大学于2017年发布,采集自Google Earth、天地图等平台,共976幅图像,用于目标检测任务。其目标类别包含飞机、油罐、立交桥和操场,其中飞机有4993架,分布在446幅图像中;操场有191个,在189幅图像;立交桥有180座,位于176幅图像;油罐有1586个,处于165幅图像。



图5 数据集样例图

Fig. 5 Sample diagram of datasets

### 4.3 定量分析

表 1 展示在平均查准率 (mean Precision, mPrecision)、平均查全率 (mean Recall, mRecall)、mAP 和平均 F1 值 (mean F1 Score, mF1 Score) 指标的表现。像 DETR, SSD 等经典算法, 在各项指标上与一些较新的算法相比存在一定差距。例如 DETR 在 NWPU-VHR-10 数据集上的 mAP 仅为 0.572 5, 在 RSOD 数据集上 mF1 Score 为 0.792 5, 整体效果相对有限。而 YOLO 系列算法不断迭代更新, YOLOv5, YOLOv7 等都展现出不错的性能, 不过在关键指标上仍有提升空间。例如 YOLOv8 在 NWPU-VHR-10 数据集上 mAP 为 0.9683, 在 RSOD 数据集上 mAP 为 0.949 5, 已

处于较高水平。NWPU-VHR-10 数据集, YOLO-UG 在 mPrecision 上达到了 0.961 6, 远超其他模型, 确保了目标识别的准确性和可靠性; 其 mRecall 也高达 0.915 8, 有效减少了数据集中目标的漏检情况; mAP 值 0.972 4 和 mF1 Score 值 0.934 0, 均证明了 YOLO-UG 在处理该数据集时的综合能力和卓越性能。RSOD 数据集, YOLO-UG 在 mPrecision 高达 0.967 5, 再次验证了其卓越的目标识别准确性; mRecall 值 0.922 6 确保了更广泛的目标覆盖, 减少了漏检; mAP 值 0.969 9 和 mF1 Score 值 0.942 5, 进一步突出了 YOLO-UG 在目标检测任务中保持高精度与高查全率之间卓越平衡的能力。

表 1 对比实验结果

Tab. 1 Experimental results contrasting

Models \ Metrics	NWPU-VHR-10 data set				RSOD data set			
	mPrecision	mRecall	mAP	mF1 Score	mPrecision	mRecall	mAP	mF1 Score
DETR	0.569 4	0.656 6	0.572 5	0.607 0	0.731 2	0.884 5	0.863 5	0.792 5
SSD	0.851 4	0.757 7	0.843 9	0.781 0	0.947 4	0.768 3	0.912 2	0.800 0
YOLOv5	0.912 1	0.896 2	0.922 9	0.901 0	0.932 3	0.783 2	0.906 4	0.792 5
HIC-YOLOv5 <sup>[46]</sup>	—	—	0.876 0	—	—	—	0.896 0	—
Drone-YOLO <sup>[46]</sup>	—	—	0.927 0	—	—	—	0.906 0	—
Xu et al <sup>[46]</sup>	—	—	0.944 0	—	—	—	0.930 0	—
YOLOv7	0.919 7	0.811 8	0.926 9	0.849 0	0.945 4	0.819 2	0.908 0	0.855 0
YOLOv7-tiny	—	—	0.921 0	—	—	—	0.945 0	—
Tang et al <sup>[47]</sup>	—	—	0.971 6	—	—	—	0.962 7	—
Wang et al <sup>[48]</sup>	—	—	0.916 5	—	—	—	0.901 6	—
Cheng et al <sup>[49]</sup>	—	—	0.938 7	—	—	—	0.862 9	—
YOLOv8	0.953 9	0.914 1	0.968 3	0.930 0	0.850 5	<b>0.925 6</b>	0.949 5	0.877 5
Zhang et al <sup>[50]</sup>	0.934 0	0.902 0	0.940 0	0.917 0	0.963 0	0.927 0	0.944 0	<b>0.971 0</b>
YOLOv9 <sup>[51]</sup>	—	—	—	—	0.934 5	0.896 5	0.920 5	0.915 3
YOLOv10 <sup>[51]</sup>	—	—	—	—	0.944 7	0.910 2	0.933 5	0.927 3
SPPF-Mamba <sup>[52]</sup>	—	—	0.901 0	—	—	—	0.945 3	—
Li et al <sup>[53]</sup>	—	—	0.926 0	—	—	—	0.951 0	—
MLA-YOLO <sup>[54]</sup>	—	—	0.946 8	—	—	—	0.939 3	—
YNet <sup>[55]</sup>	—	—	0.889 0	—	—	—	0.962 0	—
PR-Deformable DETR <sup>[56]</sup>	—	—	0.883 0	—	—	—	0.951 0	—
LMATDet <sup>[57]</sup>	—	—	0.954 3	—	—	—	0.949 8	—
SEB-YOLO <sup>[58]</sup>	—	—	0.935 0	—	—	—	0.939 0	—
YOLO-UG	<b>0.961 6</b>	<b>0.915 8</b>	<b>0.972 4</b>	<b>0.934 0</b>	<b>0.967 5</b>	0.922 6	<b>0.969 9</b>	0.942 5

表2和表3消融实验评估UFRPM, GFI, MFF模块对性能的影响。NWPU-VHR-10数据集, UFRPM模块:引入后复杂度从94.35 GMac提升至105.06 GMac, mPrecision, mRecall, mAP, mF1 Score分别提高0.011 9, 0.053 2, 0.023 7, 0.046 0, 增强了目标检测精度和识别能力。GFI模块:融合后mPrecision等指标适度增长,在优化特征表示、提升检测稳定性和全面性方面有作用。MFF模块:加入后参数量从50.60 M增加到50.88 M,复杂度从105.06 GMac提升到106.88 GMac, mPrecision等指标显著提升,在复杂特征融合和多尺度目标检测中起关键作用。RSOD数据集, UFRPM模块:融合后复杂度从

94.28 GMac提升至105.00 GMac, mPrecision略有下降0.044 3, 但mRecall显著提高0.089 6, mAP提高0.046 3, 在增强模型检测全面性贡献, mF1 Score小幅下降0.047 5体现查准率和查全率的权衡。GFI模块:mPrecision恢复并提升0.010 9, mRecall保持较高水平, mAP和mF1 Score微小变化,在优化特征表示、减少误检方面有作用。MFF模块:融合后参数量从50.56 M增加到50.84 M,复杂度从105.00 GMac提升到106.81 GMac,所有性能指标显著提升, mPrecision大幅提高0.055 5, mRecall再次提升0.043 7, mAP和mF1 Score显著增长,体现了YOLO-UG算法在目标检测任务中的卓越性能。

表2 NWPU-VHR-10数据集消融实验结果

Tab. 2 Results of ablation experiment on NWPU-VHR-10 data set

UFRPM	GFI	MFF	mPrecision	mRecall	mAP	mF1 Score	Params/M	Complexity/GMac
×	×	×	0.919 7	0.811 8	0.926 9	0.849 0	50.60	94.35
✓	×	×	0.931 6	0.865 0	0.950 6	0.895 0	50.60	105.06
✓	✓	×	0.941 8	0.871 9	0.956 9	0.901 0	50.60	105.06
✓	✓	✓	0.961 6	0.915 8	0.972 4	0.934 0	50.88	106.88

表3 RSOD数据集消融实验结果

Tab. 3 Results of ablation experiment on RSOD data set

UFRPM	GFI	MFF	mPrecision	mRecall	mAP	mF1 Score	Params/M	Complexity/GMac
×	×	×	0.945 4	0.819 2	0.908 0	0.855 0	50.56	94.28
✓	×	×	0.901 1	0.908 8	0.954 3	0.902 5	50.56	105.00
✓	✓	×	0.912 0	0.878 9	0.956 3	0.887 5	50.56	105.00
✓	✓	✓	0.967 5	0.922 6	0.969 9	0.942 5	50.84	106.81

#### 4.4 定性分析

图6展示了在NWPU-VHR-10和RSOD数据集不同模型(YOLOv7, YOLOv7-UFRPM, YOLOv7-UFRPM-GFI, YOLO-UG)的目标检测效果情况。通过该图可以直观地对比各模型在特定数据集下对于目标检测任务的表现差异,以便评估不同改进或添加模块对模型性能的影响。

模块添加后存在负面效果,包括漏检问题和准确率变低。在漏检方面,添加模块后,特定的目标会出现未被检测出来的情况,因为新模块在融合特征或处理过程中丢失了对这些目标至关重要的信息,致使模型无法准确识别。在准确率

方面,部分目标的检测准确率有所下降,新模块在优化整体性能时对某些局部特征处理产生偏差。

尽管模块添加后存在如漏检和准确率降低等负面问题,但从整体性能来看有一定提升。在多尺度目标处理能力方面,UFRPM通过多层次卷积操作,借助不同尺度感受野捕捉目标各尺度细节,在多数情况下有助于提升模型对小尺度目标的特征提取能力和大尺度目标的定位能力,减轻了目标尺度差异带来的检测困难。在目标密集重叠问题的缓解上,GFI通过衡量检测结果与真实结果的重合程度筛选检测结果以减少误检, MFF融合不同层次特征信息,增强网络对复杂

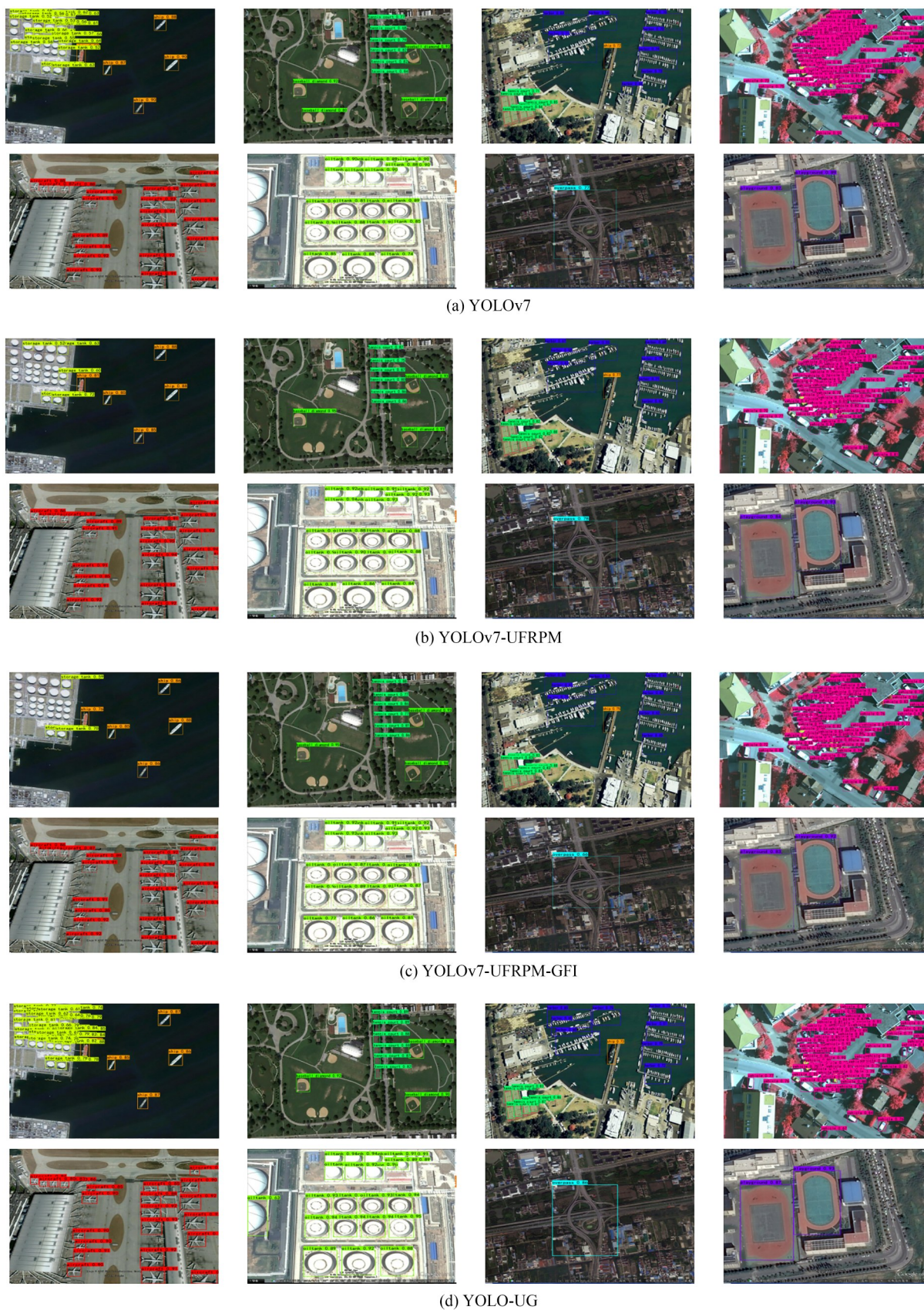


图6 NWPU-VHR-10和RSOD数据集消融实验目标检测效果图

Fig. 6 NWPU-VHR-10 and RSOD data set ablation experiment target detection effect diagram

场景的建模能力,便于在目标密集场景中更好地分离和识别目标。同时,这些模块的综合作用也改善了数据分布不均的问题,通过特征融合和优化使模型能更好地适应不同类型数据的分布,增强了在各类数据场景下的鲁棒性。

## 5 结 论

本文采用超融合残差行进模块、几何融合指数和多路径特征融合模块相结合的策略,解决遥感图像目标检测问题。超融合残差行进模块优化网络结构,增强对目标特征的感知,利于小尺度目标特征提取和大尺度目标准确定位。几何融合指数能更准确评估检测效果,通过计算几何相似度筛选结果,减少漏检和误检,提高算法准确性。多

路径特征融合模块融合不同层次特征信息,提取更丰富目标特征,提升网络对目标的表示和判别能力,增强算法对遥感图像变化和挑战的适应性,提高检测准确性和稳定性。未来可优化算法、利用更多数据、尝试多模态融合、改进硬件部署及完善评估指标体系,提升目标检测性能。

### 作者贡献声明:

白晨帅:对文章所使用的技术或方法以及研究理念做出了重要贡献;

白晓凤:对文章的设计做出了重要贡献;

邬开俊:对文章的批判性修改做出了重要贡献;

王昊雯:对文章的数据收集做出了重要贡献。

### 参考文献:

- [1] ZHAO Z Q, ZHENG P, XU S T, *et al.* Object detection with deep learning: a review [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2019, 30(11): 3212-3232.
- [2] YAGER R R. A framework for multi-source data fusion [J]. *Information Sciences*, 2004, 163(1/2/3): 175-200.
- [3] BAR-HILLEL Y, CARNAP R. Semantic information [J]. *The British Journal for the Philosophy of Science*, 1953, 4(14): 147-157.
- [4] HU X D, WANG X Q, MENG F J, *et al.* Gabor-CNN for object detection based on small samples [J]. *Defence Technology*, 2020, 16(6): 1116-1129.
- [5] WANG Y, BASHIR S M A, KHAN M, *et al.* Remote sensing image super-resolution and object detection: Benchmark and state of the art [J]. *Expert Systems with Applications*, 2022, 197: 116793.
- [6] CHEN H, LI Y F, SU D. Multi-modal fusion network with multi-scale multi-path and cross-modal interactions for RGB-D salient object detection [J]. *Pattern Recognition*, 2019, 86: 376-385.
- [7] HE X, TANG C, ZOU X, *et al.* Multispectral object detection Via cross-modal conflict-aware learning [C]. *Proceedings of the 31st ACM International Conference on Multimedia. Ottawa ON Canada. ACM*, 2023: 1465-1474.
- [8] 吴建成,郭荣佐,成嘉伟,等. 注意力特征融合的快速遥感图像目标检测算法 [J]. *计算机工程与应用*, 2024, 60(1): 207-216.
- [9] WU J C, GUO R Z, CHENG J W, *et al.* Fast remote sensing image object detection algorithm based on attention feature fusion [J]. *Computer Engineering and Applications*, 2024, 60(1): 207-216. (in Chinese)
- [10] LI Q Y, CHEN Y S, ZENG Y. Transformer with transfer CNN for remote-sensing-image object detection [J]. *Remote Sensing*, 2022, 14(4): 984.
- [11] ZHANG J Q, LEI J, XIE W Y, *et al.* SuperYOLO: super resolution assisted object detection in multimodal remote sensing imagery [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 5605415.
- [12] WAN D H, LU R S, WANG S L, *et al.* YOLO-HR: improved YOLOv5 for object detection in high-resolution optical remote sensing images [J]. *Remote Sensing*, 2023, 15(3): 614.
- [13] LIU P F, WANG Q, ZHANG H, *et al.* A lightweight object detection algorithm for remote sensing images based on attention mechanism and YOLOv5s [J]. *Remote Sensing*, 2023, 15(9): 2429.
- [14] CAO X, ZHANG Y W, LANG S, *et al.* Swin-transformer-based YOLOv5 for small-object detection in remote sensing images [J]. *Sensors*, 2023, 23(7): 3634.
- [15] XIAO J S, YAO Y T, ZHOU J, *et al.* FDLR-

- Net: a feature decoupling and localization refinement network for object detection in remote sensing images [J]. *Expert Systems with Applications*, 2023, 225: 120068.
- [15] CHEN J J, HONG H S, SONG B, *et al.* MDCT: multi-kernel dilated convolution and transformer for one-stage object detection of remote sensing images [J]. *Remote Sensing*, 2023, 15 (2): 371.
- [16] QIAN X L, LI C, WANG W, *et al.* Semantic segmentation guided pseudo label mining and instance re-detection for weakly supervised object detection in remote sensing images [J]. *International Journal of Applied Earth Observation and Geoinformation*, 2023, 119: 103301.
- [17] ZHANG D W, HAN J W, CHENG G, *et al.* Weakly supervised object localization and detection: a survey [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(9): 5866-5885.
- [18] CHENG G, YANG J Y, GAO D C, *et al.* High-quality proposals for weakly supervised object detection [J]. *IEEE Transactions on Image Processing*, 2020.
- [19] SHEN L Y, LANG B H, SONG Z X. CA-YOLO: model optimization for remote sensing image object detection [J]. *IEEE Access*, 2023, 11: 64769-64781.
- [20] LIU B, JIANG W Q. RS-YOLOx: target feature enhancement and bounding box auxiliary regression based object detection approach for remote sensing [J]. *Journal of Applied Remote Sensing*, 2024, 18 (1): 016514-016514.
- [21] TIAN Z W, HUANG J, YANG Y, *et al.* KCFS-YOLOv5: a high-precision detection method for object detection in aerial remote sensing images [J]. *Applied Sciences*, 2023, 13(1): 649.
- [22] 付宏建, 白宏阳, 郭宏伟, 等. 融合多注意力机制的光学遥感图像目标检测方法 [J]. *光子学报*, 2022, 51(12): 312-320.
- FU H J, BAI H Y, GUO H W, *et al.* Object detection method of optical remote sensing image with multi-attention mechanism [J]. *Acta Photonica Sinica*, 2022, 51(12): 312-320. (in Chinese)
- [23] 许德刚, 王再庆, 邢奎杰, 等. 改进 YOLOv6 的遥感图像目标检测算法 [J]. *计算机工程与应用*, 2024, 60(3): 119-128.
- XU D G, WANG Z Q, XING K J, *et al.* Remote sensing image target detection algorithm based on improved YOLOv6 [J]. *Computer Engineering and Applications*, 2024, 60(3): 119-128. (in Chinese)
- [24] 苗茹, 岳明, 周珂, 等. 基于改进 YOLOv7 的遥感图像小目标检测方法 [J]. *计算机工程与应用*, 2024, 60(10): 246-255.
- MIAO R, YUE M, ZHOU K, *et al.* Small target detection method in remote sensing images based on improved YOLOv7 [J]. *Computer Engineering and Applications*, 2024, 60(10): 246-255. (in Chinese)
- [25] 王家宝, 程塔, 谢星星, 等. 多元信息监督的遥感图像有向目标检测 [J]. *遥感学报*, 2023, 27 (12): 2726-2735.
- WANG J B, CHENG G, XIE X X, *et al.* Multi-information supervision in optical remote sensing images [J]. *National Remote Sensing Bulletin*, 2023, 27(12): 2726-2735. (in Chinese)
- [26] 陈鑫钰, 陈媛媛, 廉永健, 等. 基于 PSN-YOLOv7 的光学遥感图像目标检测 [J]. *现代电子技术*, 2024, 47(21): 119-124.
- CHEN X Y, CHEN Y Y, LIAN Y J, *et al.* Optical remote sensing image object detection based on PSN-YOLOv7 [J]. *Modern Electronics Technique*, 2024, 47(21): 119-124. (in Chinese)
- [27] 周沁坤, 周华平, 孙克雷, 等. ARST-YOLOv7: 用于航空遥感图像的小目标检测网络 [J/OL]. *计算机工程与应用*, 1-15 [2024-11-11].
- ZHOU Q K, ZHOU H P, SUN K L, *et al.* An Object Detector for Remote Sensing Small Weak Object Detection in Aerial Images [J]. *Computer Engineering and Applications*, 1-15 [2024-11-11]. (in Chinese)
- [28] 梁燕, 饶星晨. 改进 YOLOX 的遥感图像目标检测算法 [J]. *计算机工程与应用*, 2024, 60(12): 181-188.
- LIANG Y, RAO X C. Remote sensing image object detection algorithm with improved YOLOX [J]. *Computer Engineering and Applications*, 2024, 60(12): 181-188. (in Chinese)
- [29] ZHANG Y, YE M, ZHU G Y, *et al.* FFCA-YOLO for small object detection in remote sensing images [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5611215.
- [30] 陈立, 张帆, 郭威, 等. 面向遥感图像的多阶段特征

- 融合目标检测方法[J]. 电子学报, 2023, 51(12): 3520-3528.
- CHEN L, ZHANG F, GUO W, *et al.* Multi-stage feature fusion object detection method for remote sensing image[J]. *ACTA ELECTRONICA SINICA*, 2023, 51(12): 3520-3528. (in Chinese)
- [31] 王坤, 文武, 铁菊红, 等. 自适应融合多尺度特征的无锚框遥感图像目标检测算法[J]. 计算机辅助设计与图形学学报, 2023, 35(9): 1405-1416.
- WANG K, WEN W, TIE J H, *et al.* Anchor-free object detection method in remote sensing image via adaptive multi-scale feature fusion[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2023, 35(9): 1405-1416. (in Chinese)
- [32] JIANG S, HUANG H, YANG J X, *et al.* Innovative research on small object detection and recognition in remote sensing images using YOLOv5[J]. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2024, XLVIII-4/W10-2024: 77-83.
- [33] LI Y Y, ZHOU Z G, QI G Q, *et al.* Remote sensing micro-object detection under global and local attention mechanism[J]. *Remote Sensing*, 2024, 16(4): 644.
- [34] MA W P, WANG X T, ZHU H, *et al.* Significant feature elimination and sample assessment for remote sensing small objects' detection[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5615115.
- [35] XI L H, HOU J W, MA G L, *et al.* A multiscale information fusion network based on PixelShuffle integrated with YOLO for aerial remote sensing object detection[J]. *IEEE Geoscience and Remote Sensing Letters*, 2024, 21: 7501505.
- [36] 陈立, 张帆, 郭威, 等. 基于级联式逆残差网络的遥感图像轻量目标检测算法[J/OL]. 电子学报, 2023: 1-11. (2023-03-06). <http://kns.cnki.net/KCMS/detail/detail.aspx?filename=DZXU2023030300Q&dbname=CJFD&dbcode=CJFQ>.
- CHEN L, ZHANG F, GUO W, *et al.* Remote Sensing Image Lightweight Target Detection Algorithm based on Cascade Inverse Residual Network [J/OL]. *China Industrial Economics*, 2023: 1-11. (2023-03-06). <http://kns.cnki.net/KCMS/detail/detail.aspx?filename=DZXU2023030300Q&dbname=CJFD&dbcode=CJFQ>. (in Chinese)
- [37] 汪西莉, 梁正印, 刘涛. 基于特征注意力金字塔的遥感图像目标检测方法[J]. 遥感学报, 2023, 27(2): 492-501.
- WANG X L, LIANG Z Y, LIU T. Feature attention pyramid-based remote sensing image object detection method[J]. *National Remote Sensing Bulletin*, 2023, 27(2): 492-501. (in Chinese)
- [38] YU D W, JI S P. A new spatial-oriented object detection framework for remote sensing images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 60: 4407416.
- [39] SHARMA M, DHANARAJ M, KARNAM S, *et al.* YOLOrs: object detection in multimodal remote sensing imagery[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020, 14: 1497-1508.
- [40] LI C, CHENG G, WANG G X, *et al.* Instance-aware distillation for efficient object detection in remote sensing images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 5602011.
- [41] CHEN S L, ZHAO J Q, ZHOU Y, *et al.* InfoFPN: an Informative feature pyramid network for object detection in remote sensing images[J]. *Expert Systems with Applications*, 2023, 214: 119132.
- [42] ZHANG C, LAM K M, WANG Q. CoF-net: a progressive coarse-to-fine framework for object detection in remote-sensing imagery [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 5600617.
- [43] YUE C K, YAN J H, ZHANG Y, *et al.* SCF-Net: Semantic correction and focus network for remote sensing image object detection [J]. *Expert Systems with Applications*, 2023, 224: 119980.
- [44] ZHANG J, LI X, LI J, *et al.* Rethinking mobile block for efficient attention-based models [C]. 2023 *IEEE/CVF International Conference on Computer Vision (ICCV)*. *IEEE Computer Society*, 2023: 1389-1400.
- [45] ZHANG H, ZHANG S J. Shape-IoU: More Accurate Metric considering Bounding Box Shape and Scale[EB/OL]. 2023:arXiv:2312.17663.
- [46] 许思源, 吴伟林. 多尺度特征融合的遥感图像目标检测算法研究[J/OL]. 计算机工程与应用, 2024: 1-10. (2024-06-27). <http://kns.cnki.net/KCMS/detail/detail.aspx?filename=JS->

- GG2024062400A&dbname=CJFD&dbcode=CJFQ.
- XU S Y, WU W L. Research on Target Detection Algorithm of Remote Sensing Image based on Multi-Scale Feature Fusion[J/OL]. *China Industrial Economics*, 2024: 1-10. (2024-06-27). <http://kns.cnki.net/KCMS/detail/detail.aspx?filename=JSGG2024062400A&dbname=CJFD&dbcode=CJFQ>. (in Chinese)
- [47] 唐锦云. 基于深度学习的遥感图像目标检测算法研究[D]. 桂林: 广西师范大学, 2023.
- TANG J Y. *Research on Remote Sensing Image Target Detection Algorithm based on Deep Learning* [D]. Guilin: Guangxi Normal University, 2023. (in Chinese)
- [48] 王晨. 基于神经网络结构搜索的遥感图像目标检测研究[D]. 西安: 西安电子科技大学, 2023.
- WANG C. *Research on Remote Sensing Image Object Detection Based on Neural Architecture Search* [D]. Xian: Xidian University, 2023. (in Chinese)
- [49] 程德强, 马尚, 寇旗旗, 等. 基于 YOLOv4 改进特征融合及全局感知的目标检测算法[J]. *智能系统学报*, 2024, 19(2): 325-334.
- CHENG D Q, MA S, KOU Q Q, *et al.* Target detection algorithm for improving feature fusion and global perception based on YOLOv4 [J]. *CAAI Transactions on Intelligent Systems*, 2024, 19(2): 325-334. (in Chinese)
- [50] 张朝阳, 张上, 胡益民, 等. 动态聚焦多维注意力遥感弱小目标检测[J/OL]. *无线电通信技术*, 1-17 [2024-11-13].
- ZHANG C Y, ZHANG S, HU Y M, *et al.* Dynamic Focusing Multidimensional Attention Remote Sensing Weak Target Detection[J/OL]. *Radio Communications Technology*, 1-17 [2024-11-13]. (in Chinese)
- [51] YAO G Z, ZHU S D, ZHANG L, *et al.* HP-YOLOv8: high-precision small object detection algorithm for remote sensing images [J]. *Sensors*, 2024, 24(15): 4858.
- [52] TANG Y, ZHAO Z, HUANG O. Mamba enhanced you only look once network with multiscale spatial attention for remote sensing object detection [J]. *Available at SSRN* 5059656.
- [53] LI C Z, LI Y L, CHEN X H, *et al.* Concerning imbalance and bounding box loss to detect small targets in remote sensing[J]. *IEEE Sensors Journal*, 2024, 24(17): 27631-27639.
- [54] ZHAO W, XUE Y J, SONG Q Z. MLA-YOLO: enhancing remote sensing object detection with multi-attention mechanisms[C]. *Proceedings of the 2024 Guangdong-Hong Kong-Macao Greater Bay Area International Conference on Digital Economy and Artificial Intelligence*, 2024: 738-743.
- [55] SONG X, GAO E H. YGNet: a lightweight object detection model for remote sensing[J]. *IEEE Geoscience and Remote Sensing Letters*, 2024, 22: 7501305.
- [56] CHEN Y P, LIU B J, YUAN L Y. PR-deformable DETR: DETR for remote sensing object detection[J]. *IEEE Geoscience and Remote Sensing Letters*, 2024, 21: 2506105.
- [57] ZHAO Y Q. LMATDet: Lightweight Multi-scale Aerial TargetDetection with attention and efficient loss[J]. 2024.
- [58] HUI Y, YOU S J, HU X H, *et al.* SEB-YOLO: an improved YOLOv5 model for remote sensing small target detection[J]. *Sensors*, 2024, 24(7): 2193.

## 作者简介:



白晨帅(1998—),男,陕西渭南人,博士研究生,2023年于兰州交通大学获得硕士学位,主要从事图像处理、计算机视觉方面的研究。E-mail: 13230071@stu.lzjtu.edu.cn

## 通讯作者:



邬开俊(1978—),男,山东莒南人,博士生导师、教授,2017年于兰州交通大学获得博士学位,主要从事图像处理、计算机视觉方面的研究。E-mail: wkj@mail.lzjtu.cn