

DOI:10.19651/j.cnki.emt.2519027

MGD-YOLO: 基于 YOLOv8 的遥感图像目标检测算法*

李菲 陈鹏宇 梁钰漫 郑鑫宇 王烈

(广西大学计算机与电子信息学院 南宁 530004)

摘要: 针对遥感图像中目标尺度差异大、类别多样且分布不均导致的检测不精准、漏检和误检问题,提出一种基于 YOLOv8n 改进的遥感图像目标检测算法 MGD-YOLO。首先,提出了多尺度边缘高斯注意力模块 MEGA,其结合高斯平滑、Scharr 边缘算子与通道注意力机制,有效抑制遥感图像噪声干扰,强化目标轮廓特征表达。其次,设计出 MDPCConv 结构,通过动态加权融合机制和深度可分离卷积相结合,改善传统卷积感受野固定问题,增强模型对多尺度目标的感知能力。最后,检测头部分引入 DLGA 结构,动态分配注意力分支权重,并采用 MLP 融合策略,提升局部细节与全局语义信息的融合效果,从而提高检测性能。实验表明,MGD-YOLO 在 DIOR、DOTA 和 NWPU VHR-10 数据集上与 YOLOv8n 相比,mAP@0.5 分别提升了 1.6%、2.7%、1%,验证了其在遥感图像目标检测中的有效性和良好的鲁棒性。

关键词: 遥感图像;YOLOv8;深度可分离卷积;注意力模块

中图分类号: TN911.73 **文献标识码:** A **国家标准学科分类代码:** 520.6

MGD-YOLO: Target detection algorithm for remote sensing images based on YOLOv8

Li Fei Chen Pengyu Liang Yuman Zheng Xinyu Wang Lie

(School of Computer and Electronic Information, Guangxi University, Nanning 530004, China)

Abstract: To address the issues of imprecise target localization, missed detections, and false alarms caused by significant differences in target scales, diverse categories, and uneven target distribution in remote sensing images, this paper proposes an improved algorithm based on YOLOv8n, named MGD-YOLO. Firstly, the multi-scale edge-gaussian attention module (MEGA) is introduced. By integrating Gaussian smoothing, the Scharr edge operator, and a channel attention mechanism, MEGA effectively suppresses noise and enhances the feature representation of target contours in complex backgrounds. Secondly, the MDPCConv structure is designed, which combines a dynamic weighted fusion mechanism with depthwise separable convolutions to overcome the fixed receptive field problem of traditional convolutions and improve the model's ability to detect targets of varying scales. Lastly, the DLGA structure is introduced in the detection head. By dynamically allocating weights to multiple attention branches and utilizing an MLP fusion strategy, DLGA significantly improves the integration of local and global features, thereby boosting detection performance. Experimental results demonstrate that MGD-YOLO achieves a 1.6%, 2.7% and 1% increase in mAP@0.5 on the DIOR, DOTA and NWPU VHR-10 datasets, respectively, compared to YOLOv8n, thus validating its effectiveness for remote sensing image target detection tasks.

Keywords: remote sensing images; YOLOv8; depth-wise separable convolution; attention module

0 引言

随着卫星和航拍技术的快速发展,遥感图像目标检测的应用场景日益丰富。在军事领域,该技术主要用于目标侦查与导航定位,在民用领域,则广泛服务于地理信息更

新、城市规划建设、灾害监测与应急救援以及自动驾驶等多个方面^[1]。

目标检测算法主要分为传统目标检测算法和基于深度学习的目标检测算法。在基于深度学习的这类算法中,分为两阶段目标检测算法和一阶段目标检测算法。两阶段算

收稿日期:2025-06-06

* 基金项目:广西重点研发项目(桂科 AB24010033)资助

法,如 R-CNN^[2]、Fast R-CNN^[3]、Faster R-CNN^[4]等,需先生成候选区域再进行检测,精度较高但计算复杂、效率低;一阶段算法,如 YOLO 系列^[5]、SSD^[6]和 RetianNet^[7]系列等,无需候选框,通过端到端架构直接预测目标类别与位置,虽精度略低,但计算效率更高、检测速度更快,在实时性要求较高的遥感图像目标检测场景中更具应用潜力,尤其 YOLO 系列因兼顾速度与精度,成为该领域的研究热点。

近年来,国内外针对遥感图像目标检测的研究已取得一系列进展。梁礼明等^[8]提出基于上下文信息的方法,借助上下文模块与空洞空间卷积金字塔模块,实现全局信息和局部信息的融合。李安达等^[9]提出对 YOLOv7 进行改进,将 Siou 应用于损失函数的优化,进而增强了模型在多尺度目标检测方面的性能表现。冒国韬等^[10]提出通过多尺度分割注意力机制优化骨干网络,在瓶颈层嵌入跨层级特征关联模块,利用长程依赖建模增强目标区域表征能力,结合动态权重分配机制自适应调整深浅层特征融合比例,抑制复杂背景干扰并提升多尺度目标检测精度。左露等^[11]提出在 YOLOX 框架中引入注意力机制强化目标特征表达,构建 MSCE-R 多尺度融合结构以缓解目标尺度变化引起的细节信息丢失,采用 Ciou 损失函数优化边界框回归过程,使模型在保持精度的同时满足实时检测需求。Xu 等^[12]提出一种基于局部感知 Swin Transformer 架构的 LPSW 主干网络。该模型融合 Transformer 的全局建模能力与 CNN 的局部特征提取优势,构建具有区域感知增强特性的检测框架,获得更好的检测性能。

尽管现有研究在遥感目标检测领域已取得显著进展,但高分辨率影像的固有复杂性仍带来三大瓶颈。首先,复杂背景与噪声导致目标模糊或遮挡,现有方法难以精准区分前景与背景,造成漏检与误检^[13];其次,传统卷积网络受限于固定感受野或简单多尺度融合,难以动态匹配遥感场景中巨大的目标尺度差异,小目标及尺度剧变目标的检测稳定性不足^[14];此外,现有复杂注意力或融合机制提升精度时常大幅增加模型参数量和计算量,如何在融合多层次特征以提升模糊目标检测能力的同时实现轻量化,仍是亟待解决的难题^[15]。

1 YOLOv8 算法

YOLOv8 是 2023 年初开源的目标检测算法,包含 n、s、m、l、x 五种规模模型。考虑到计算效率和实际应用需求,本研究选择计算量最小、推理速度最快的 YOLOv8n 作为基准模型。

YOLOv8 主要由主干网络、颈部网络、检测头三大部分组成。主干网络借助 Mosaic 数据增强提升检测能力,融合相关设计构建 C2f 模块增强特征表达;颈部网络由特征金字塔网络和路径汇聚网络组成,建立跨层级特征交互;检测头通过解耦头分离分类与回归任务,通过独立卷积层实现类别概率预测和位置坐标回归。

2 MGD-YOLO 算法

在处理遥感图像时,受目标尺度变化大、类别多样且分布不均等问题影响,存在检测精度和召回率不足的局面。为解决这一问题,本文对 YOLOv8 的骨干网络、颈部结构及检测头模块进行改进,提出面向遥感图像目标检测的 MGD-YOLO 模型,其整体结构如图 1 所示。

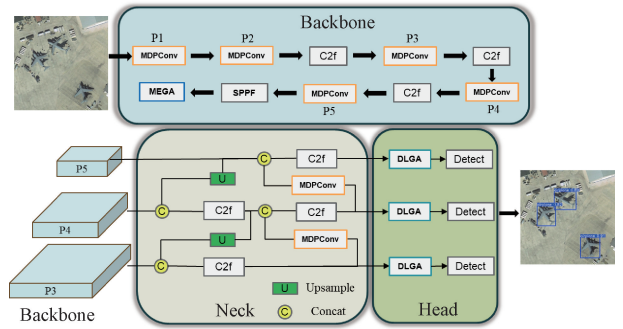


图 1 MGD-YOLO 结构

Fig. 1 MGD-YOLO structural

MGD-YOLO 模型主要改进及贡献如下:

1) 在骨干网络最后一层嵌入多尺度边缘高斯注意力模块,该模块结合高斯平滑、Scharr 边缘算子与通道注意力机制,能有效抑制遥感图像的噪声干扰,强化复杂背景下目标轮廓的特征表达。

2) 采用多尺度动态感知卷积替代基准模型的 Conv 层,通过动态加权融合机制和深度可分离卷积,增强模型对不同尺度目标的感知能力,改善传统卷积在特征提取中感受野固定的问题。

3) 在检测头引入动态轻量化注意力模块,该模块可动态分配多个注意力分支的权重,并采用 MLP 融合策略,显著提升局部细节与全局语义信息的融合效果,进一步提高遥感图像目标和模糊目标的检测性能。

2.1 多尺度边缘高斯注意力模块

遥感图像多采集于卫星或无人机等高空平台,存在分辨率低、成像模糊、噪声干扰、光照不足及目标遮挡等问题。这些因素严重削弱了图像特征的判别能力,导致模型在边界感知上的模糊、响应不稳定,从而降低检测精度。其次,相比自然图像,遥感场景中的目标存在尺度变化大、姿态各异、密集重叠等特性。传统检测方法往往依赖自然图像的先验假设,难以适应这类复杂条件,在低质量图像中表现更为有限。

为了解决上述问题,本文提出多尺度边缘高斯注意力模块(multi-scale edge-Gaussian attention module, MEGA),通过多尺度高斯滤波与平滑处理,有效抑制噪声干扰,并结合 Scharr 算子增强边缘信息,从而显著提升特征质量。此外,MEGA 模块采用不同尺寸和标准差的高斯核,在多个尺度上捕捉目标特征,增强了模型对尺度变化的鲁棒性,有

效应对遥感图像中目标尺度差异大的问题。针对边缘信息提取不足的问题,模块引入 Scharr 边缘检测机制,弥补了传统卷积神经网络在边缘感知方面的不足,提升了目标边界定位的准确性。最后,模块融合 ECA 注意力机制,动态调整各通道权重,使网络能够自适应地关注关键特征通道,进一步提升整体检测性能。MEGA 模块结构如图 2 所示。

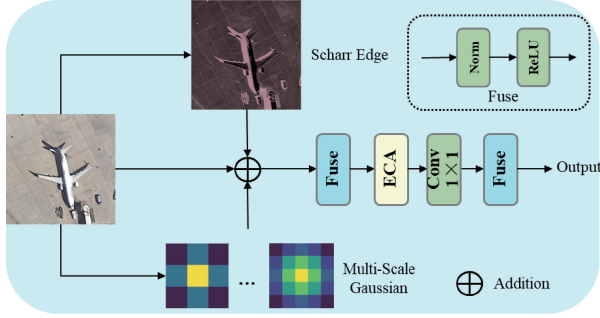


图 2 MEGA 结构
Fig. 2 MEGA structure

假设输入特征图为 $X \in \mathbb{R}^{C \times H \times W}$, 首先使用多尺度高斯滤波器对输入特征图进行处理,对于每一个指定的尺寸和标准差(size, sigma),构建一个对应的高斯核,并应用到输入特征图上:

$$G_i = \text{Gaussian}(X; size_i, \sigma_i) \quad (1)$$

其中, i 表示不同的尺度和标准差组合, G_i 表示第 i 个尺度下的高斯滤波结果。

然后将所有尺度的结果平均得到一个综合的高斯滤波输出:

$$G_{out} = \frac{1}{N} \sum_{i=1}^N G_i \quad (2)$$

其中, N 是使用的不同尺度和标准差组合的数量。

接着使用 Scharr 算子对输入特征图 X 进行边缘检测, Scharr 算子是一种用于计算图像梯度的卷积核,能够有效地突出显示图像中的边缘信息。Scharr 算子分别沿水平方向和垂直方向应用 Scharr 算子,再合并得到最终的边缘图:

$$E_{out} = \sqrt{(\text{Scharr}_X(X))^2 + (\text{Scharr}_Y(X))^2} \quad (3)$$

Scharr_X 和 Scharr_Y 分别表示沿 X 轴和 Y 轴方向的边缘检测结果。

之后将得到的高斯滤波结果 G_{out} 、边缘图 E_{out} 以及原始输入特征图 X 结合起来,经过规范化和激活函数处理后,通过 ECA 注意力进行特征增强,最后再经过 $\text{Conv}1 \times 1$ 、规范化和激活函数处理后得到输出:

$$F = \text{Norm}(\text{ReLU}(X + G_{out} + E_{out})) \quad (4)$$

$$\text{Output} = F(\text{Conv}1 \times 1(\text{ECA}(F))) \quad (5)$$

2.2 多尺度动态感知卷积

遥感图像中的目标往往具有较大的尺度变化范围,从几米大小的小型车辆到几百米长的大型建筑物不等。传统的固定尺度卷积核难以同时捕捉这些多尺度信息,导致模

型对某些尺度的目标检测效果不佳。并且为了提高模型的性能,通常需要增加更多的卷积层和参数,但这会导致计算量大幅增加,影响推理速度和部署成本。

为了解决上述问题,本文提出了多尺度动态感知卷积(multi-scale dynamic perceptual convolution, MDPConv),通过引入动态加权融合机制和深度可分离卷积,有效增强了模型对不同尺度目标的感知能力。MDPConv 模块的核心在于其动态加权融合机制,该机制允许模型根据输入数据自适应地调整每个卷积核的重要性,从而更精确地捕捉不同尺度的目标特征。MDPConv 模块结构如图 3 所示。

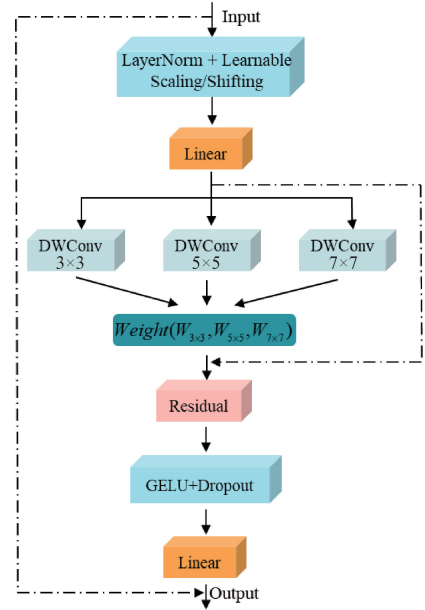


图 3 MDPConv 结构
Fig. 3 MDPConv structure

输入特征图记为 X , MDPConv 卷积并行使用 3 个不同尺寸的深度可分离卷积处理,提取不同尺度的空间特征:

$$Y_k = \text{DWConv}_{k \times k}(X), k \in \{3, 5, 7\} \quad (6)$$

其中, $\text{DWConv}_{k \times k}$ 为 $k \times k$ 的深度可分离卷积。

为了根据输入内容动态选择最合适的尺度,引入 3 个可学习标量权重 $W_{3 \times 3}, W_{5 \times 5}, W_{7 \times 7}$, 对 3 条支路进行加权融合:

$$Z = \frac{W_{3 \times 3} Y_3 + W_{5 \times 5} Y_5 + W_{7 \times 7} Y_7}{W_{3 \times 3} + W_{5 \times 5} + W_{7 \times 7} + 10^{-6}} \quad (7)$$

接着,引入一个轻量级残差块进一步强化特征。首先对融合特征 Z 执行 1×1 卷积与 GELU 激活:

$$H = \text{GELU}(\text{Conv}_{1 \times 1}(Z)) \quad (8)$$

然后将结果通过 3×3 深度卷积,并与 Z 相加:

$$R = Z + \text{DWConv}_{3 \times 3}(H) \quad (9)$$

其中, $\text{Conv}_{k \times k}$ 为 $k \times k$ 卷积, GELU 为激活函数。

最后,为防止过拟合,在残差块末尾加入 Dropout,并与原始输入直接相加,完成残差连接:

$$\text{Output} = X + \text{Dropout}_{0.1}(\text{Conv}_{1 \times 1}(R)) \quad (10)$$

深度可分离卷积降低了参数量和计算复杂度,使得多尺度并行计算可行;动态加权融合则让网络随输入内容实时调整各尺度的贡献,提升对多尺度目标的感知。残差块借助跳跃连接和 GELU 激活增强梯度流动与非线性表达,缓解梯度消失;Dropout 正则化进一步提高泛化能力,使网络在关注关键通道的同时保持鲁棒性。

2.3 动态轻量级组注意力块

在遥视觉任务中,模型面临多尺度特征提取和注意力机制适配的挑战。首先,遥感图像中物体尺度差异显著、种类繁多,现有轻量化骨干网络多基于自然图像设计,难以提取多尺度对象细粒度特征,限制了任务的准确率。其次,现有注意力机制在处理遥感图像多尺度特征时存在不足,如特征提取不均衡、细粒度捕捉能力弱及计算复杂度过高等,且缺乏针对遥感任务特性的优化,导致模型在不同尺度对象处理上的兼顾能力不足。

因此,本文引入了动态轻量化注意力模块(dynamic lightweight group attention block, DLGA),通过 4 个分支(PA、LA、MRA、GA)捕捉不同尺度的特征。PA 分支专注于提取微小目标和细粒度特征,LA 分支捕捉局部特征的邻域空间关系,MRA 分支增强中等尺度复杂结构对象的表达,GA 分支根据当前阶段动态选择合适的全局注意力模块,捕捉全局特征和长距离依赖关系。此外,DLGA 模块采用动态通道分配机制,依据输入特征自适应生成通道权重,优化各分支的权重分配,从而强化关键特征的表达。最后,通过 MLP 融合和残差连接,DLGA 模块在多尺度信息建模中兼顾局部细节与全局上下文,显著提升了特征的鲁棒性。DLGA 模块结构如图 4 所示。

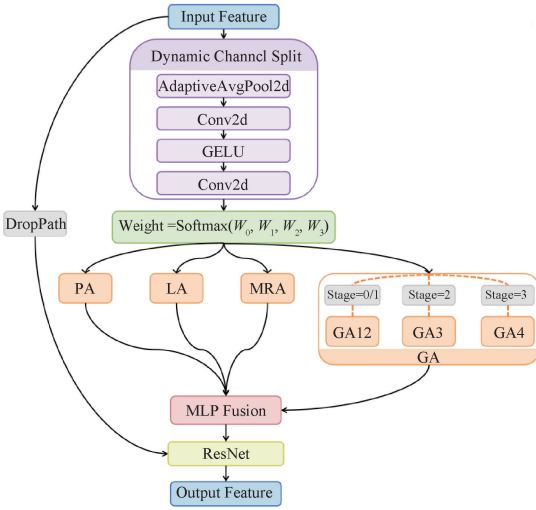


图 4 DLGA 结构

Fig. 4 DLGA structure

首先通过 Dynamic Channel Split 子模块对输入特征图 X 进行动态通道分配,依据输入特征自适应生成 4 个通道权重。该过程可以表示为:

计算池化后的特征:

$$P = AdaptiveAvgPool2d(X) \quad (11)$$

其中, AdaptiveAvgPool2d 为自适应全局平均池化。

对池化后的特征进行卷积和激活:

$$G = GELU(Conv2d(P)) \quad (12)$$

其中, Conv2d 为 1×1 卷积, GELU 为激活函数。

计算权重:

$$weight = Softmax(Conv2d(G)) \quad (13)$$

其中, $weight$ 表示生成的权重向量,形状为 $(B, 4)$,用于动态调整各分支所使用的通道重要性。

然后,将输入特征图 X 复制 4 份,并分别乘以生成的权重向量中的对应权重:

$$X_i = X \times weight[:, i].view(B, 1, 1, 1) \quad (14)$$

其中, $i \in \{0, 1, 2, 3\}$, 表示权重向量中的对应权重。

将加权后的特征图分别送入 4 个分支(PA、LA、MRA、GA)进行特征提取,得到各自的输出特征:

$$\begin{cases} X_{PA} = PA(X_0) \\ X_{LA} = LA(X_1) \\ X_{MRA} = MRA(X_2) \\ X_{GA} = GA(X_3) \end{cases} \quad (15)$$

对于不同阶段,GA 分支的处理方式如下:在阶段 0 和 1 时,则采用全局注意力模块(GA12),在保留一定特征细节的同时进行全局特征提取;阶段 3 时,使用下采样全局注意力模块(D_GA),先对特征图进行下采样以减少计算量,再应用全局注意力机制进行特征提取;阶段 4 时,直接使用全局注意力模块(GA4),以便更直接地捕捉广泛的全局特征。

最后,将 4 个分支的输出特征进行拼接形成融合特征,并通过一个高效的多层感知机(MLP)进行深度特征融合,该 MLP 包含两个 1×1 卷积层,中间加入归一化层和激活函数,能有效整合不同分支的互补特征并增强非线性表达能力。添加 DropPath 跳跃连接以增强训练稳定性,将融合后的特征与输入特征 X 相加,形成残差连接,得到输出特征。这种 MLP 融合策略通过逐点卷积的非线性变换,优化了多尺度特征的整合效果,同时残差结构确保了关键信息的无损传递,使模型能够自适应地强化重要特征并抑制冗余信息。具体的计算过程如下:

$$L = \alpha_{PA} \cdot X_{PA} + \alpha_{LA} \cdot X_{LA} + \alpha_{MRA} \cdot X_{MRA} + \alpha_{GA} \cdot X_{GA} \quad (16)$$

$$X = X + MLP(L) \quad (17)$$

其中, α_{PA} 、 α_{LA} 、 α_{MRA} 、 α_{GA} 表示各个分支可学习的融合系数。

3 实验结果与分析

3.1 实验环境以及配置

实验使用 64 位 Windows10 操作系统,编译器是 Python3.8.20, PyTorch 2.4.1 和 CUDA11.8。显卡为

NVIDIA GeForce RTX 4090。实验参数情况如表 1 所示。

表 1 实验参数设计

Table 1 Experimental parameter design

参数	设定值
训练次数 (epochs)	300
批量大小 (batch size)	8
图片尺寸 (image size)	640×640
初始学习率 (lr0)	0.01
权重衰减系数 (weight_decay)	0.000 5
学习率动量 (momentum)	0.937
自动混合精度训练 (amp)	False
优化器 (optimizer)	SGD

3.2 数据集

为了充分验证本文改进模型的有效性,实验采用 DIOR、DOTA 以及 NWPU VHR-10 数据集进行验证。

DOTA 数据集^[16]由 2 806 幅航拍图像组成,每张图像的像素尺寸在 800×800 到 4 000×4 000 的范围内,空间分辨率从 0.5 到 30 m 不等。数据集分为 1 411 个训练集、458 个验证集和 937 个测试集。它包含 15 种常见的目标类别:飞机、船只、储油罐、桥梁、停车场、港口、网球场、棒球场、田径场、游泳池、瀑布、车辆、篮球馆、直升机和风力发电机。这些目标以任意方向和大小分布,标注为定向边界框(OBB),适用于目标检测和识别任务。

DIOR 数据集^[17]由 23 463 幅光学遥感图像组成,每幅图像大小为 800 pixel×800 pixel,空间分辨率从 0.5 到 30 m 不等。数据集分为 10 000 个训练集、8 000 个验证集和 5 463 个测试集。该数据集包含 20 种常见的陆地物体类别:飞机、机场、棒球场、篮球场、桥梁、烟囱、水坝、高速公路服务区、高速公路收费站、高尔夫球场、地面田径场、港口、立交桥、船

舶、体育场、储油罐、网球场、火车站、车辆和风力发电厂。

NWPU VHR-10 数据集^[18]由 800 幅超高分辨率(VHR)遥感图像组成,其中 715 张彩色图像来自 Google Earth,空间分辨率在 0.5~2 m 之间,85 张全色锐化彩色红外图像来自 Vaihingen 数据,空间分辨率为 0.08 m。数据集包含 10 类地理空间物体:飞机、轮船、储油罐、棒球场、网球场、篮球场、地面跑道、港口、桥梁和车辆。

3.3 模型评价指标

为了准确评估改进后模型的性能,本文研究主要选择以下指标进行实验结果的评估:精度(precision,P)、召回率(recall,R)、平均精度均值(mean average precision,mAP)。

其中,精确率和召回率是通过混淆矩阵中的真正例(TP)、假正例(FP)和假负例(FN)计算得到的。

$$P = \frac{TP}{TP + FP} \quad (18)$$

$$R = \frac{TP}{TP + FN} \quad (19)$$

坐标轴与 P-R 曲线围成的面积表示 AP 值。平均精度均值(mAP)是对所有 AP 值的平均,用于评估整个目标检测网络的性能。mAP@0.5 表示在 IoU(交并比)为 0.5 时的平均精度,而 mAP@0.5:0.95 则是在 IoU 从 0.5 变化到 0.95 的范围内,以 0.05 为步长计算的平均精度。

$$AP = \int_0^1 p(r) dr \quad (20)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n P_{AP,i} \quad (21)$$

其中, $p(r)$ 是召回率 r 对应的查准率。

3.4 消融实验

为验证每种改进对模型的增益,本文在 NWPU VHR-10 和 DIOR 两个数据集上分别进行了消融实验,结果如表 2 和表 3 所示。

表 2 在 NWPU VHR-10 数据集的消融实验

Table 2 Ablation experiment on the NWPU VHR-10 dataset

YOLOv8n	MEGA	DLGA	MDPConv	mAP@0.5	mAP@0.5:0.95	P	R	%
✓				86.9	55.7	88.1	82.0	
✓	✓			88.3	55.8	88.8	81.9	
✓		✓		87.8	55.7	86.4	82.0	
✓			✓	88.6	57.8	91.0	82.6	
✓	✓	✓		88.7	56.7	86.8	84.5	
✓	✓	✓	✓	87.9(+1)	58.7(+3)	88.7(+0.6)	82.7(+0.7)	

由表 2 可知,在 NWPU VHR-10 数据集消融实验中,基准模型引入 MEGA 后 mAP@0.5 提升至 88.3%,P 也提高 0.7%,单独插入 DLGA 时 mAP@0.5 提升至 87.8%,保持召回率 82%,仅嵌入 MDPConv 则使 mAP@0.5:0.95 大幅增至 57.8%,P 更突破 90%。当 MEGA、

DLGA、MDPConv 全部融合后,模型取得 87.9% mAP@0.5 与 58.7% mAP@0.5:0.95 刷新所有评价指标,证明 3 种改进协同互补,可在 NWPU VHR-10 上实现更精准、更鲁棒的遥感目标检测。

由表 3 可知,DIOR 数据集上的消融实验同样印证了

表3 在DIOR数据集的消融实验

Table 3 Ablation experiment on the DIOR dataset

YOLOv8n	MEGA	DLGA	MDPConv	mAP@0.5	mAP@0.5:0.95	P	R	%
✓				77.6	54.8	85.7	69.5	
✓	✓			77.7	55.2	85.9	69.6	
✓		✓		77.5	55.3	86.3	68.8	
✓			✓	78.3	56.0	85.7	70.3	
✓	✓	✓		78.2	56.0	86.3	69.8	
✓	✓	✓	✓	79.2(+1.6)	57.3(+2.5)	86.4(+0.7)	70.7(+1.2)	

3项改进的递进价值。相比于基准模型,融合全部改进点后,mAP@0.5、mAP@0.5:0.95分别为77.6%、54.8%,整体提升1.6%、2.5%,P、R同步增长0.7%、1.2%。跨数据集的共同抬升表明,所提改进兼顾精度与鲁棒性,具备优良泛化能力。

3.5 对比实验

为了全面评估MGD-YOLO模型在遥感图像目标检测任务中的性能,将其与多种经典目标检测算法以及其他遥感图像目标检测算法进行对比。实验基于3个数据集:DIOR、DOTA和NWPU VHR-10,评价指标包括mAP@0.5、mAP@0.5:0.95、P(准确率)和R(召回率)。对比的结果如表4~6所示。

表4 在DIOR数据集的对比实验

Table 4 Comparative experiments on the DIOR dataset

模型	%			
	mAP@0.5	mAP@0.5:0.95	P	R
YOLOv5n	76.3	53.2	85.8	67.9
YOLOv6n	74.2	53.3	86.2	65.9
YOLOV9t	77.6	56.7	86.3	69.8
YOLOV10n	76.2	54.3	85.4	68.1
YOLOv11n	76.9	54.9	86.8	68.3
Faster-RCNN+SGIAF ^[19]	69.8	49.3	76.6	65.3
CSYOLOv7 ^[20]	74.0	—	83.1	68.3
TPH-YOLOv5 ^[21]	70.9	41.8	75.4	67.6
YOLOv8n	77.6	54.8	85.7	69.5
MGD-YOLOv8	79.2	57.3	86.4	70.7

从表4~6可知,本文所改进的算法模型MGD-YOLO在DIOR、DOTA及NWPUVHR-10数据集上的检测精度明显优于YOLO系列模型及其他遥感图像目标检测算法。在DIOR数据集上,MGD-YOLO相较于YOLO系列的基础版本和其他对比模型,在mAP@0.5上提高了1.6%,达到了79.2%;在mAP@0.5:0.95上提升了2.5%,达到

表5 在DOTA数据集的对比实验

Table 5 Comparative experiments on the DOTA dataset

模型	%			
	mAP@0.5	mAP@0.5:0.95	P	R
YOLOv5n	78.0	54.6	82.2	72.1
YOLOv6n	74.8	51.9	82.1	69.3
YOLOV9t	79.3	56.4	82.8	73.7
YOLOV10n	77.5	54.6	81.3	71.4
YOLOv11n	78.5	54.7	82.7	72.3
RSD-YOLO ^[22]	73.0	50.4	76.6	68.7
YOLOv11n改进模型 ^[23]	72.5	47.9	81.7	70.5
RT-DETR+PConvBlockr ^[24]	72.3	48.5	77.6	69.6
TPH-YOLOv5 ^[25]	71.4	—	80.0	67.6
FEW-YOLOv8 ^[26]	69.3	—	73.5	66.4
YOLOv8-LGA ^[27]	73.6	—	81.6	68.8
YOLOv8n	79.4	56.3	83.7	73.4
MGD-YOLOv8	82.1	59.6	84.3	76.2

表6 在NWPU VHR-10数据集的对比实验

Table 6 Comparative experiments on the

NWPU VHR-10 dataset

模型	%			
	mAP@0.5	mAP@0.5:0.95	P	R
YOLOv5n	86.9	55.6	85.9	81
YOLOv6n	86.7	54.6	88.6	79.1
YOLOV9t	84.4	54.0	86.3	77.8
YOLOV10n	80.3	50.6	78.9	75.6
YOLOv11n	85.6	54.0	86.4	80.3
YOLOv8n	86.9	55.7	88.1	82.0
MGD-YOLOv8	87.9	58.7	88.7	82.7

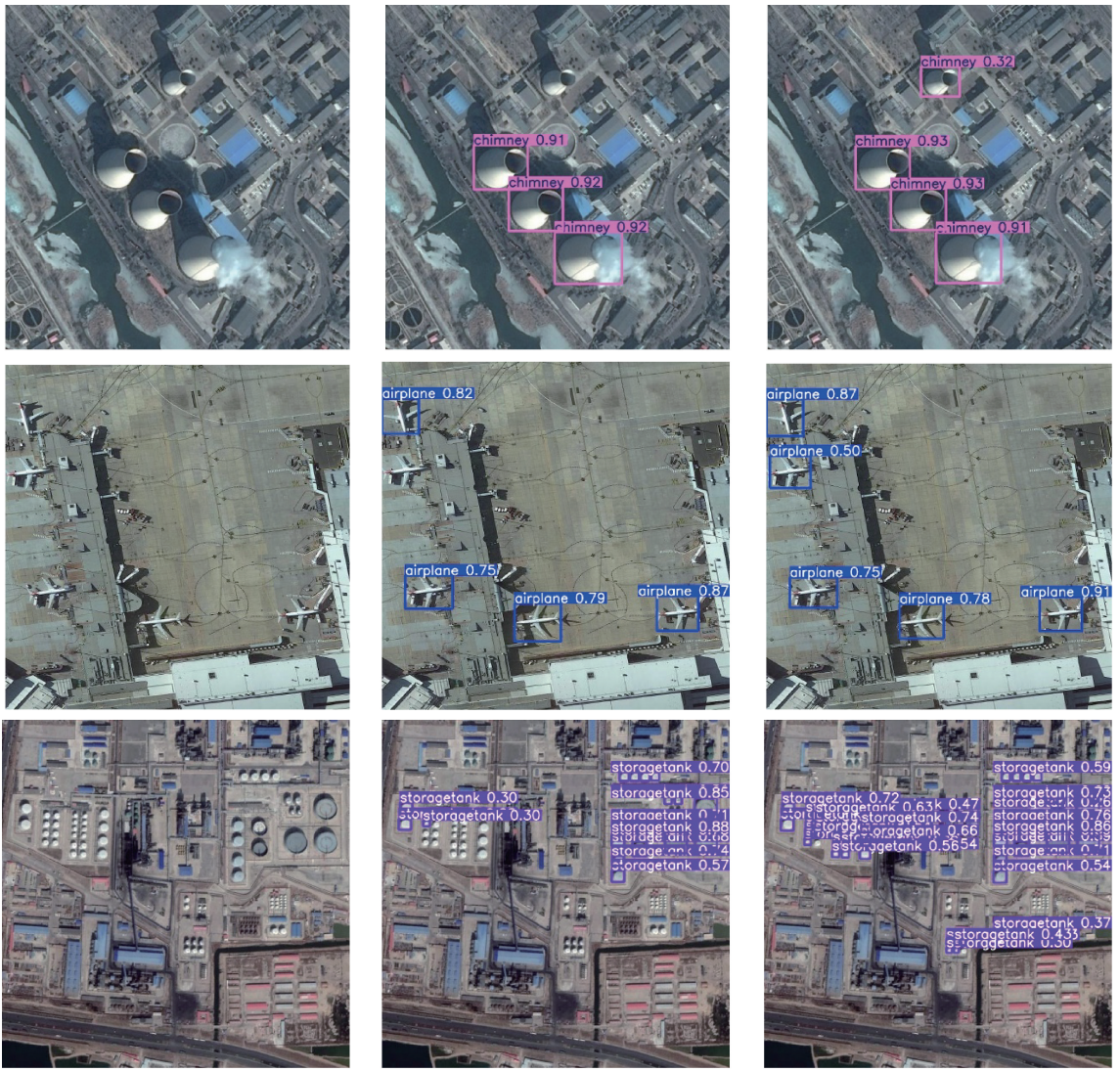
57.3%。此外,P值和R值分别有0.7%和1.2%的提升。

这表明 MGD-YOLOv8 在处理多尺度目标时具有更优秀的检测精度和召回能力。对于包含更多复杂场景和多样尺度目标的 DOTA 数据集, MGD-YOLOv8 展示了更为显著的改进效果。具体而言, $mAP@0.5$ 提升至 82.1%, 相比 YOLOv8n 提高了 2.7%; $mAP@0.5:0.95$ 达到 59.6%, 增加了 3.3%。同时, P 值和 R 值也分别实现了 0.6% 和 2.8% 的增长, 这反映出 MGD-YOLOv8 在复杂背景下对不同尺度目标的识别能力得到了有效增强。最后, 在 NWPU VHR-10 数据集上, 尽管各模型间的性能差异相对较小, 但 MGD-YOLO 依然表现出色。它在 $mAP@0.5$ 上获得了最高分数 87.9%, 相较于 YOLOv8n 提高了 1%, 且 $mAP@0.5:0.95$ 提升至 58.7%, 增加了 3%。P 值和 R 值也有轻微提升, 证明了该模型在高分辨率遥感图像

中细小目标检测的有效性。

3.6 实验结果可视化分析

为了更清晰地展示模型改进的效果, 现将 YOLOv8n 模型和 MGD-YOLO 模型的检测效果以及热力图进行对比。如图 5 所示, 图 5(b) 为 YOLOv8n 的检测效果, 图 5(c) 为 MGD-YOLO 的检测效果。可以看出 MGD-YOLO 算法可以检测出更多特征不明显的遥感图像目标以及降低误检率, 以适应不同环境中的复杂问题, 展现了良好的鲁棒性与实用性。如图 6 所示, MGD-YOLO 算法在提升遥感图像目标检测精度方面显著优于 YOLOv8n。MGD-YOLO 模型的注意力机制更加集中在目标的中心点上, 使得预测更加精准, 从而显著提升了模型的整体检测性能。



(a) 原图
(a) Original image

(b) YOLOv8n

(c) MGD-YOLO

图 5 检测效果对比

Fig. 5 Detection effect comparison

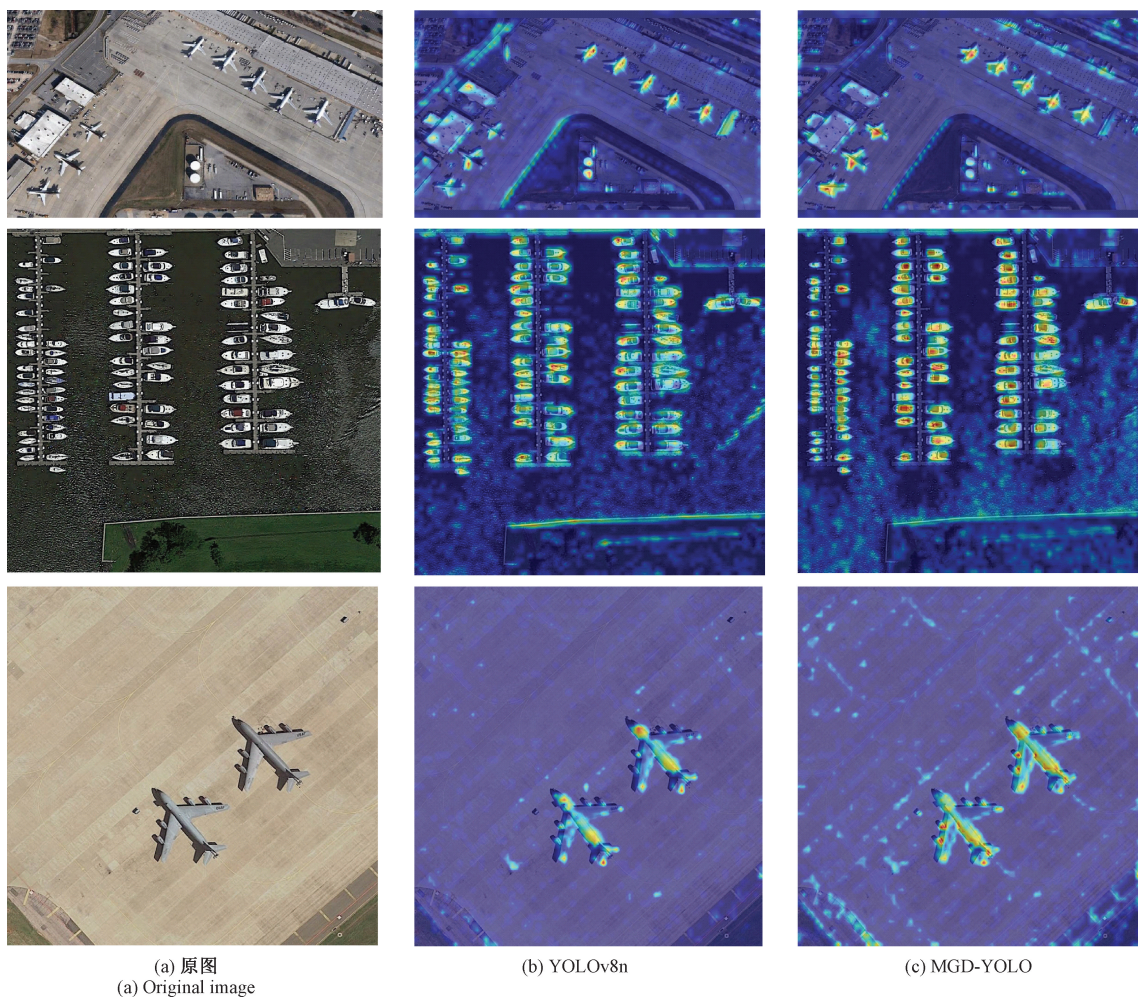


图6 热力图对比

Fig. 6 Heat map comparison

4 结 论

本文提出了一种基于YOLOv8的改进算法MGD-YOLO,以解决遥感图像中目标尺度差异大、类别多样且分布不均等问题。提出MEGA多尺度边缘高斯注意力模块,解决遥感图像中噪声干扰和目标轮廓特征表达不足的问题,有效增强复杂背景下目标轮廓的识别精度;设计MDPCnv多尺度动态感知卷积,改善传统卷积在特征提取过程中感受野固定、难以适应多尺度目标的问题,提高了模型对不同尺度目标的感知能力;引入DLGA动态轻量化注意力模块,解决检测头中局部细节与全局语义信息融合不足的问题,通过动态分配多个注意力分支的权重,并采用MLP融合策略,提升特征融合效果,进一步提高遥感图像目标和模糊目标的检测性能。

实验表明,MGD-YOLO在DIOR数据集上 $mAP@0.5$ 提升了1.6%, $mAP@0.5:0.95$ 提升了2.5%,精确率提升了0.7%,召回率提升了1.2%。在DOTA数据集上, $mAP@0.5$ 提升了2.7%, $mAP@0.5:0.95$ 提升了

3.3%,精确率提升了0.6%,召回率提升了2.8%。在NWPU VHR-10数据集上, $mAP@0.5$ 提升了1%, $mAP@0.5:0.95$ 提升了3%,精确率和召回率也有所提升。这些结果验证了MGD-YOLO在遥感图像目标检测任务中的优越性和良好的鲁棒性。

然而,MGD-YOLO在计算复杂度和对极端尺度变化及复杂背景场景的适应能力方面仍有提升空间。后续研究将致力于优化算法,降低计算复杂度,并探索更高效的特征提取与融合策略,以进一步增强模型的检测性能和泛化能力。

参考文献

- [1] 武腾辉,邓炳光. LS-YOLO:基于改进YOLOv8n的航拍小目标检测算法[J]. 电讯技术, 2026, 66(2): 221-228.
WU T H, DENG B G. A small target detection algorithm for aerial images based on improved YOLOv8n: LS-YOLO[J]. Telecommunication Engineering, 2026,

- 66(2):221-228.
- [2] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [3] GIRSHICK R. Fast R-CNN[C]. IEEE International Conference on Computer Vision, Santiago: IEEE, 2015: 1440-1448.
- [4] REN SH Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [5] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 779-788.
- [6] LIU W, ANGUOLOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. 14th European Conference on Computer Vision. Berlin: Springer, 2016: 21-37.
- [7] ZHANG SH F, WEN L Y, BIAN X, et al. Single-shot refinement neural network for object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 4203-4212.
- [8] 梁礼明, 李仁杰, 董信, 等. 基于上下文信息的遥感图像目标检测[J]. 电光与控制, 2023, 30(10): 89-94.
- LIANG L M, LI R J, DONG X, et al. Target detection in remote sensing images based on context information[J]. Electronics Optics & Control, 2023, 30(10): 89-94.
- [9] 李安达, 吴瑞明, 李旭东. 改进 YOLOv7 的小目标检测算法研究 [J]. 计算机工程与应用, 2024, 60(1): 122-134.
- LI AN D, WU R M, LI X D. Research on improving YOLOv7's small target detection algorithm [J]. Computer Engineering and Applications, 2024, 60(1): 122-134.
- [10] 冒国韬, 邓天民, 于楠晶. 基于多尺度分割注意力的无人机航拍图像目标检测算法[J]. 航空学报, 2023, 44(5): 273-283.
- MAO G T, DENG T M, YU N J. Object detection in UAV images based on multi-scale split attention[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(5): 273-283.
- [11] 左露, 牛晓伟, 朱春惠. 基于改进 YOLOX 的航拍遥感图像检测模型[J]. 电子测量技术, 2023, 46(16): 179-186.
- ZUO L, NIU X W, ZHU CH H. Aerial remote sensing image detection model based on improved YOLOX [J]. Electronic Measurement Technology, 2023, 46(16): 179-186.
- [12] XU X K, FENG ZH J, CAO CH Q, et al. An improved swin transformer-based model for remote sensing object detection and instance segmentation[J]. Remote Sensing, 2021, 13(23): 4779.
- [13] 梁奔. 复杂遥感环境下的目标检测算法研究[D]. 石家庄: 河北科技大学, 2024.
- LIANG B. Research on object detection algorithms in complex remote sensing environments[D]. Shijiazhuang: Hebei University of Science and Technology, 2024.
- [14] 刘洋, 战荫伟. 基于深度学习的小目标检测算法综述[J]. 计算机工程与应用, 2021, 57(2): 37-48.
- LIU Y, ZHAN Y W. Survey of small object detection algorithms based on deep learning [J]. Computer Engineering and Applications, 2021, 57(2): 37-48.
- [15] 马梁, 苟于涛, 雷涛, 等. 基于多尺度特征融合的遥感图像小目标检测[J]. 光电工程, 2022, 49(4): 49-65.
- MA L, GOU Y T, LEI T, et al. Small object detection based on multi-scale feature fusion using remote sensing images [J]. Opto-Electronic Engineering, 2022, 49(4): 49-65.
- [16] XIA G S, BAI X, DING J, et al. DOTA: A large-scale dataset for object detection in aerial images[C]. IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City: IEEE, 2018: 3974-3983.
- [17] LI K, WAN G, CHENG G, et al. Object detection in optical remote sensing images: A survey and a new benchmark[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 159: 296-307.
- [18] LI K, CHENG G, BU SH H, et al. Rotation-insensitive and context-augmented object detection in remote sensing images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(4): 2337-2348.
- [19] 刘洋, 时富斌, 王竹筠, 等. 基于显著性引导增强与改进 Faster-RCNN 的遥感图像目标检测方法[J]. 沈阳航空航天大学学报, 2024, 41(6): 50-60.
- LIU Y, SHI F B, WANG ZH Y, et al. Saliency guided enhancement and improved Faster-RCNN for object detection method in remote sensing images[J]. Journal of Shenyang Aerospace University, 2024, 41(6): 50-60.
- [20] 邵凯, 李浩刚, 梁燕, 等. 跨尺度特征融合的遥感微小目标检测算法[J]. 系统工程与电子技术, 2025, 47(5): 1421-1431.

- SHAO K, LI H G, LIANG Y, et al. Remote sensing small target detection algorithm based on cross-scale feature fusion [J]. *Systems Engineering and Electronics*, 2025, 47(5):1421-1431.
- [21] 廖宁生,曹天秀,刘科言,等.复合特征与多尺度融合的无人机小目标检测算法[J].*计算机工程与应用*,2025,61(3):111-120.
- LIAO N SH, CAO T X, LIU K Y, et al. Small target detection algorithm for UAV based on composite feature and multi-scale fusion [J]. *Computer Engineering and Applications*, 2025, 61(3): 111-120.
- [22] 汤伟博,方强,李沛根,等.基于RSD-YOLO的无人机航拍图像的小目标检测[J/OL].*计算机工程*,1-15[2025-09-19].<https://doi.org/10.19678/j.issn.1000-3428.0070151>.
- TANG W B, FANG Q, LI P G, et al. RSD-YOLO-based small target detection in UAV aerial images [J/OL]. *Computer Engineering*,1-15[2025-09-19].<https://doi.org/10.19678/j.issn.1000-3428.0070151>.
- [23] 罗显志,汪航.跨尺度特征融合的无人机小目标检测算法[J].*计算机工程与应用*,2025,61(14):135-147.
- LUO X ZH, WANG H. Small target detection algorithm for UAV based on cross scale feature fusion [J]. *Computer Engineering and Applications*, 2025, 61(14): 135-147.
- [24] 刘思元,高凯,雍龙泉.改进RT-DETR的航拍小目标检测算法[J].*计算机工程与应用*,2025,61(4):272-281.
- LIU S Y, GAO K, YONG L Q. Improved RT-DETR algorithm for aerial small object detection [J]. *Computer Engineering and Applications*, 2025, 61(4): 272-281.
- [25] ZHU X K, LYU SH CH, WANG X, et al. TPH-YOLOv5: Improved yolov5 based on transformer prediction head for object detection on drone-captured scenarios [C]. 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). Seoul: IEEE, 2021:2778-2788.
- [26] 席阳丽,屈丹,王芳芳,等.基于FEW-YOLOv8遥感图像目标检测算法[J].*郑州大学学报*,2025,46(4):62-69.
- XI Y L, QU D, WANG F F, et al. Target detection algorithm based on FEW-YOLOv8 remote sensing images[J]. *Journal of Zhengzhou University*, 2025, 46(4):62-69.
- [27] 刘勇,吕丰顺,李学琨,等.基于YOLOv8-LGA的遥感图像目标检测算法[J].*光电子·激光*,2025,36(9):928-937.
- LIU Y, LYU F SH, LI X K, et al. Remote sensing image target detection algorithm based on YOLOv8-LGA[J]. *Journal of Optoelectronics · Laser*, 2025, 36(9):928-937.

作者简介

李菲,本科,主要研究方向为目标检测、自然语言处理。

E-mail:17267054751@163.com

陈鹏宇,硕士研究生,主要研究方向为深度学习、目标检测。

E-mail:619847874@qq.com

梁钰塏,硕士研究生,主要研究方向为深度学习、语音识别处理。

E-mail:yuman_china@163.com

郑鑫宇,本科,主要研究方向为机器视觉、自然语言处理。

E-mail:18859131863@163.com

王烈(通信作者),博士,教授,硕士生导师,主要研究方向为FPGA、图像处理。

E-mail:lwang@gxu.edu.cn