

DOI:10.19651/j.cnki.emt.2417274

面向遥感图像的轻量化小目标检测算法研究^{*}

葛文 邵钰琦 屈乐乐

(沈阳航空航天大学电子信息工程学院 沈阳 110136)

摘要: 针对遥感图像中存在因背景复杂、尺度变化大、小目标密集而导致的目标检测准确率低等问题,提出了一种改进 YOLOv8n 的检测算法:MGL-YOLO。首先,设计 MSConv 减少模型参数,基于 MSConv 重构 C2f 模块,提高多尺度特征提取能力;其次,基于 GLSA 和 GSConv 模块改进 BiFPN 结构,简化颈部网络,增强网络的特征处理能;在头部设计 Light-head 进一步轻量化模型,加强小目标特征提取能力;最后引入 NWD 损失函数替换原损失函数,增强对小目标的定位精度。在 DIOR-R、DOTAv1.0 和 VEDAI 数据集上验证,实验结果表明 MGL-YOLO 模型在 DIOR-R 数据集上准确率和 mAP@0.5 比基准模型提高了 1.7% 和 1.3%,在 DOTAv1.0 数据集上提高了 1.0% 和 1.1%,在 VEDAI 数据集上提高了 2.4% 和 2.1%。参数量降低 47%,计算量降低 32%,与其他小目标检测算法相比也得到了较好的检测性能。

关键词: YOLOv8n; 多尺度卷积; 特征金字塔; 轻量检测头; NWD

中图分类号: TN919.8 文献标识码: A 国家标准学科分类代码: 520.6040

Research on lightweight small target detection algorithm for remote sensing images

Ge Wen Shao Yuqi Qu Lele

(School of Electronic Information Engineering, Shenyang University of Aeronautics and Astronautics, Shenyang 110136, China)

Abstract: Aiming at the problem of low accuracy in object detection caused by complex background, large scale changes, and dense small targets in remote sensing images, an improved YOLOv8n detection algorithm, MGL-YOLO, is proposed. Firstly, design MSConv to reduce model parameters and reconstruct C2f module based on MSConv to improve multi-scale feature extraction capability; secondly, based on GLSA and GSConv modules, the BiFPN structure is improved to simplify the neck network and enhance its feature processing capability. Design a Light head model at the head to further reduce weight and enhance the ability to extract small target features; finally, the NWD loss function is introduced to replace the original loss function and enhance the localization accuracy for small targets. Verified on the DIOR-R, DOTAv1.0, and VEDAI datasets, the experimental results show that the MGL-YOLO model has high accuracy on the DIOR-R dataset mAP@0.5 Improved by 1.7% and 1.3% compared to the benchmark model, 1.0% and 1.1% on the DOTAv1.0 dataset, and 2.4% and 2.1% on the VEDAI dataset. The parameter count has been reduced by 47%, and the computational complexity has been reduced by 32%. Compared with other small object detection algorithms, it has also achieved good detection performance.

Keywords: YOLOv8n; MSConv; feature pyramid; Light-head; NWD

0 引言

遥感技术是航空航天、地球科学、环境监测等众多领域中不可或缺的研究工具。随着近些年遥感技术的迅猛发展,我们能够获得更为精确、详尽的观测数据,这些优质数据大大推动了图像分析技术在遥感领域的应用^[1]。在遥

像的分析过程中,目标检测技术占据了关键地位,它 can 有效识别并定位图像中的地面目标,如车辆、建筑物等,这项技术在城市规划、交通监测、农业监控和环境保护等领域均有广泛的应用价值^[2]。

近年来,深度学习技术的发展显著提升了遥感图像中的目标检测精度,特别是卷积神经网络 CNN 的应用,大大

收稿日期:2024-11-04

* 基金项目:国家自然科学基金(61671310)、辽宁省兴辽英才计划项目基金(XLYC1907134)项目资助

增强了检测的效率与精度。目标检测方法主要分为两类:一类是以 R-CNN^[3]、Fast R-CNN^[4]、Faster R-CNN^[5]为代表的基于区域生成的两阶段方法;另一类是 YOLO 和 SSD 等基于回归的单阶段方法。YOLO 作为 CNN 的代表性算法,在遥感领域应用广泛,表现出色。YOLOv5 结合了跨级局部网络 CSPNet 的特征提取模块和快速空间金字塔池化 SPPF,优化了小目标检测能力,提升了推理速度;而 YOLOv7^[6]通过引入增强型扩展高效层聚合网络模块,改进了特征金字塔网络(feature pyramid network, FPN),使小目标检测更具表现力。Betti 等^[7]则提出了一种专门用于小目标检测的 YOLO 网络,通过小规模特征提取与跳跃连通实现特征复用。Chaurasia 等^[8]提出的轻量级旋转目标检测算法解决了传统基于计算机视觉的探测技术不足。Zhang 等^[9]提出 FFCA-YOLO 遥感图像小目标检测算法,设计特征增强模块(feature enhancement module, FEM),特征提取模块(feature extraction module, FFM),空间上下文感知模块(spatial context awareness module, SCAM)来提高对小目标检测能力。桑雨等^[10]针对遥感图像小目标提出改进 YOLOv7-tiny 检测模型,设计轻量级的空间金字塔池化结构,来减少多余的卷积算子操作,设计单尺度检测头,解决颈部信息冗余问题。梁礼明等^[11]设计多尺度特征聚合模块加强多尺度和小目标的检测准确度,同时引入 Wise-IoU 来提高普通质量锚框,提升算法的检测性能。刘富宽等^[12]对 YOLOv8 进行改进,设计多尺度特征的加权融合,引入 P2 检测层,增加对小目标的检测精度。左露等^[13]对 YOLOX 进行改进,在主干网络加入注意力机制,引入多尺度信息融合模块和 CIoU 来提高小目标检测精度。

但现有方法还存在一定的不足,如小目标不明显,特征提取能力不足。密集分布和小目标聚集的场景检测漏检误检概率高。对于大范围的遥感图像,检测小目标计算资源和时间受限。在低分辨率遥感图像中,小目标的细节信息可能丢失等。基于上述问题,本文提出一种轻量化的遥感图像小目标检测算法:MGL-YOLO,力图在损失较小的前提下提升检测精度。本文的创新点包括:

- 1) 在颈部网络中加入 GSLA^[14] 和 GSConv^[15] 模块,改进 BiFPN 结构,增强特征处理能力并轻量化网络。
- 2) 设计多尺度卷积 MSConv 模块改进 Bottleneck 单元,并优化 C2f 模块以减少参数数量与计算量。
- 3) 设计轻量化的 Light-head 检测头,同时强化小目标检测能力。
- 4) 引入 NWD 损失函数以替代 CIoU 损失函数,进一步提高小目标的检测精度。

1 YOLOV8n 算法

Ultralytics 公司在 2023 年 1 月推出了新一代目标检测算法 YOLOv8。该算法主要由主干网络、颈部网络和检测头三部分组成。主干网络负责特征提取,包含 Conv、C2f

和空间金字塔池化等模块。YOLOv8 采用了新型 C2f 结构,通过跨阶段局部网络结构来丰富梯度流信息。SPPF 模块则用于将任意大小的特征图转换为固定尺寸的特征张量。颈部网络用于多尺度特征融合,采用了 PANet 结构,包括 FPN 和路径聚合网络(path aggregation network, PAN),实现上下层特征的全面融合,并通过上采样和通道拼接将三个输出分支传递至 Head 结构。Head 部分则通过解耦头设计,分离回归与预测分支,从而提高了模型的泛化性能^[16]。

2 MGL-YOLO 算法

为应对遥感图像中的小目标检测问题,本文提出了 MGL-YOLO 算法。通过设计 MSConv 改进 C2f 模块,并基于 BiFPN 结合 GLSA 和 GSConv 优化了颈部的 PANet 和 FPN 结构。此外,设计了轻量化检测头,适用于小目标检测的 NWD 损失函数也被引入。其结构如图 1 所示。不仅提高了检测精度,同时显著降低了模型的计算量和参数量,实现了模型的轻量化需求。

2.1 多尺度通道卷积和 C2f 改进

传统卷积操作仅使用单尺度的卷积核,因而在区分不同尺度的目标或场景特征时效果有限。相比之下,多尺度卷积通过使用多种卷积核尺寸,可以更好地提取不同尺度的特征信息,丰富了提取的特征。基于这种思路,本文设计了一种多尺度卷积模块 MSConv。在 MSConv 模块中,首先将输入按照通道数分为两组,第一组采用 1×1 和 1×5 卷积核处理图像,提取基本特征;第二组将未处理的特征进一步分为两部分,分别使用 5×1 和 5×5 卷积核提取特征。随后将所有特征连接,并通过 1×1 卷积层实现融合,从而在实现升维降维的同时增强通道信息的整合。相比传统卷积方法,MSConv 不仅保留了多尺度特征,还有效减少了参数数量。其结构如图 2 所示。

对于大小为 $H \times W \times C$ 的输入特征,采用尺寸为 $k \times k$ 的普通卷积时,所需的参数量和计算量可以通过式(1)来计算,其中 c 表示输入数据的通道数。

$$\begin{cases} par = k \times k \times c^2 \\ FLOPs = h \times w \times k^2 \times c^2 \end{cases} \quad (1)$$

而对于 MSConv 而言,由于其只对部分通道数操作,若使 $k \times k$ 大小的卷积核,所需参数量计算量公式如式 2 所示。

$$\begin{cases} par = k \times k \times \frac{c}{g} \times \frac{c}{g} \times \frac{1}{g} \times \frac{1}{g} \\ FLOPs = h \times w \times k^2 \times \frac{c}{g} \times \frac{c}{g} \end{cases} \quad (2)$$

MSConv 的参数量和计算量相比传统方法减少了约四分之一,因此可以用来替代 YOLOv8n 的 C2f 模块中的标准卷积。替换之后的 Bottleneck 和 C2f 模块结构分别如图 3 和 4 所示。Bottleneck_MSC 模块通过引入更为复杂的

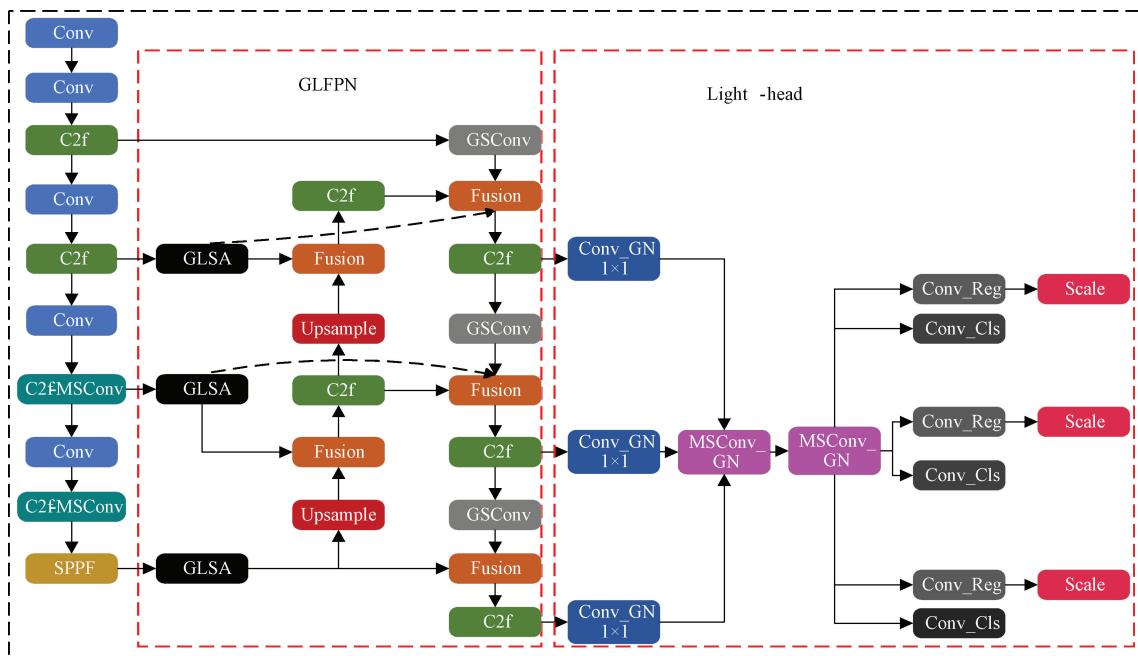


图 1 MGL-YOLO 结构图

Fig. 1 MGL-YOLO structure diagram

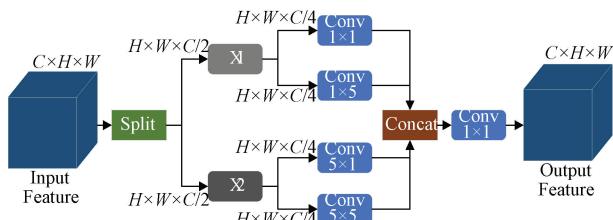


图 2 MSConv 结构图

Fig. 2 MSConv structure diagram

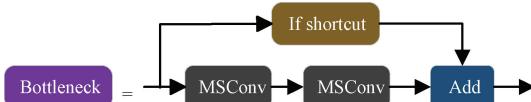


图 3 Bottleneck_MSC 结构图

Fig. 3 Bottleneck_MSC structure diagram

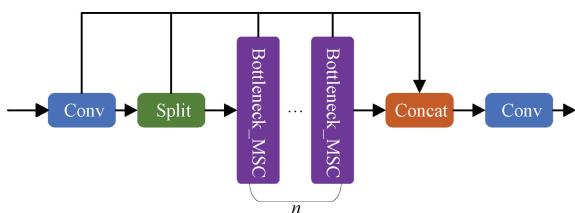


图 4 C2f-MSConv 结构图

Fig. 4 C2f-MSConv structure diagram

特征表示,有效地提升了图像细节信息的提取能力。这种增强的特征提取对小目标检测尤为关键,使模型能够更准确地识别小目标的特征。这两个模块的应用不仅减少了模型的参数量,还显著提高了网络在目标检测任务中的精

确度。

2.2 Neck 结构优化

1) GLSA 模块

LSA 和 GSA 结构图如图 5 所示, GLSA 模块如图 6 由 GSA 和 LSK^[17]模块构成。GSA 强调空间中每个像素的长程关系,可以作为局部空间注意力的补充,生成全局空间注意力图。 F_i^1 为输入特征图,输出表达式如下:

$$Att_G(F_i^1) = \text{Softmax}(\text{Transpose}(C_{1 \times 1}(F_i^1))) \quad (3)$$

$$G_{sa}(F_i^1) = \text{MLP}(Att_G(F_i^1) \otimes F_i^1) + F_i^1$$

$\text{AttG}(\cdot)$ 为注意力运算, $C_{1 \times 1}$ 为 1×1 卷积。 \otimes 表示矩阵乘法。MLP(\cdot)由两个全连接层和一个 ReLU 非线性和规范化层组成。MLP 的第一层将其输入转换为一个扩展比为 2 的高维空间,第二层将其恢复为与输入相同的维数。

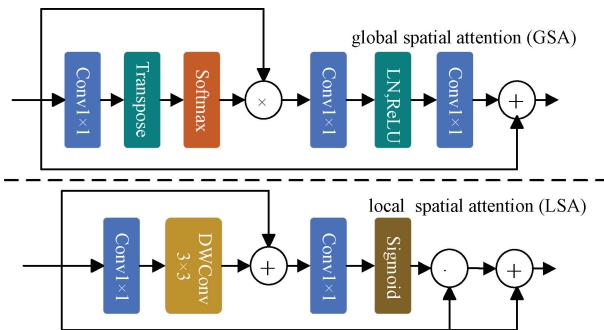


图 5 LSA 和 GSA 结构图

Fig. 5 LSA and GSA structure diagram

LSA 模块在给定的特征图的空间维度上有效提升小

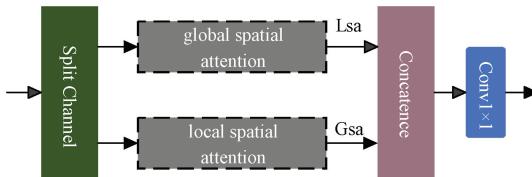


图 6 GLSA 结构图

Fig. 6 GLSA structure diagram

目标的局部特征,计算局部空间注意响应和作为输入 F_i^2 ,如下:

$$\begin{aligned} Att_L(F_i^2) &= \sigma(C_{1 \times 1}(F_c(F_i^2)) + F_i^2) \\ L_{sa} &= Att_L(F_i^2) \cdot F_i^2 + F_i^2 \end{aligned} \quad (4)$$

GLSA 特征映射 F_i 的通道的 g 平均分成两个特征映射组 F_i^1 和 F_i^2 , 分别馈入 GSA 模块和 LSA 模块。这两个注意单元的输出最后被一个 1×1 的卷积层连接起来。将这样一个过程表述为:

$$\begin{aligned} F_i^1, F_i^2 &= Split(F_i) \\ F'_i &= C_{1 \times 1}(Concat(G_{sa}(F_i^1), L_{sa}(F_i^2))) \end{aligned} \quad (5)$$

2) 轻量级卷积 GSConv

GSConv 模块由 Conv、DWConv、Concat 和 Shuffle 等子模块构成,网络结构如图 7 所示。首先, C_1 通道的特征图被分为两部分,一部分使用深度可分离卷积处理,另一部分使用标准卷积。之后,将处理后的两个部分进行合并和特征串联。Shuffle 作为一种信道混合方法,通过在不同信道间传递 SC 模块的特征信息,使其与 DSC 输出相互融合,实现信道间的信息交换。

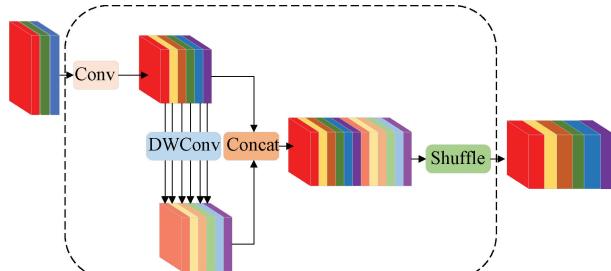


图 7 GSConv 结构图

Fig. 7 GSConv structure diagram

GSConv 对时间的复杂度为:

$$Time_{GSConv} \sim O\left(W \cdot H \cdot K_1 \cdot K_2 \cdot \frac{C_2}{2} \cdot (C_1 + 1)\right) \quad (6)$$

其中, W 是输出特征图的宽度, H 是高度, K_1 和 K_2 是卷积核的大小, C_1 是每个卷积核的通道数,也是输入特征图的通道数, C_2 是输出特征图的通道数。

GSConv 可以以一种更简单的方式完成通道洗牌,而不需要额外的 FLOPs。对于轻量化模块来说,GSConv 的优势很明显,通过增加 DSC 层和通道洗牌,非线性表达能力得到了增强。

3) GLPFN 结构

与传统的 FPN 结构不同,BiFPN 在特征金字塔的相邻层级之间引入了双向连接,使得信息能够在高层特征和低层特征之间双向流动。因此,针对遥感图像中小目标数量多、背景复杂的特点,本文设计了一种改进的 BiFPN 结构,命名为 GLFPN 全局到局部特征融合结构,如图 8 所示。该结构通过引入 GLSA 全局到空间聚合模块,并将 BiFPN 中的标准卷积替换为轻量化的 GSConv 卷积。GSConv 的轻量化设计减少了计算量和参数量,加速了复杂背景下的特征融合,并实现了更深层次的特征提取。GLSA 模块替换了原有的标准卷积,提升了全局到局部的特征提取能力,并提高了检测的准确率。GLFPN 结合了全局到局部的特征融合、轻量化的 GSConv 卷积和 BiFPN 结构,显著提高了对遥感图像中小目标的检测精度,降低了误检和漏检的几率,增强了模型的鲁棒性。

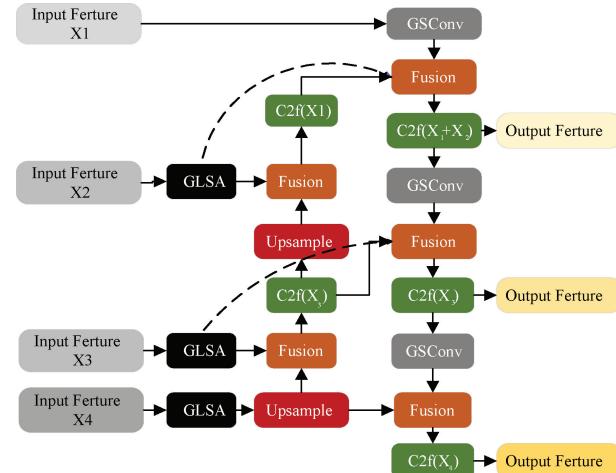


图 8 GLFPN 结构图

Fig. 8 GLFPN structure diagram

2.3 轻量化检测头

原始的 YOLOv8 检测头存在一定限制。首先,检测头的参数量较大,占据算法计算量的约 1/5。每个检测头均需通过两次 3×3 卷积和一次 1×1 卷积来提取信息,导致整体参数量显著增加。其次,原始算法采用单尺度预测结构,无法有效处理多尺度目标,因为它仅利用一个特征映射尺度进行预测,忽视了其他尺度的特征贡献。为了解决这些问题,本文设计了新的 Light-head 检测头,引入了 GroupNorm 思想,这一方法在 FCOS^[18] 中已被证明可以显著提升定位性能。其结构如图 9 所示。Light-head 设计的核心在于利用共享的 GroupNorm 卷积和 MSConv 共享卷积来替代 3 个检测头中的两个常规卷积。此外,为了解决不同尺度目标检测不一致的问题,引入了尺度层对特征进行调整。该结构设计有效减少了参数量,同时赋予检测头更强的多尺度感知能力,以提高复杂环境下的检测准确率。

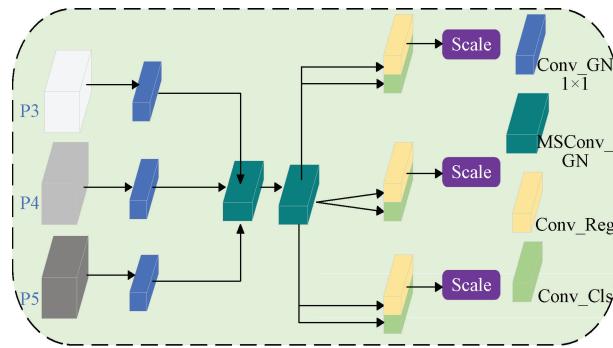


图 9 Light-head 结构图

Fig. 9 Light-head structure diagram

2.4 优化损失函数

遥感图像中的小目标物体数量较多, 检测这些小目标时往往遇到困难。例如, 预测的边界框可能无法与真实的框重叠, 或是完全包围真实框。此时, 传统的 IoU(交并比)损失函数通常难以提供充分的梯度信息, 从而无法有效地引导网络优化, 导致训练和预测时易发生漏检与误检。这类问题直接影响了模型的检测性能和精度, 使得小目标的准确检测面临挑战。鉴于小目标检测的独特性, 依赖 IoU 的损失函数并不足够, 因此本文引入了基于 Wasserstein 距离的 NWD 损失函数。对于两个二维高斯分布 μ_1 和 μ_2 , 其二阶 Wasserstein 距离的定义为:

$$W_2^2(\mu_1, \mu_2) = \|m_1 - m_2\|_2^2 + Tr \\ (\sum 1 + \sum 2 - 2(\sum_i^{1/2} \sum_j^{1/2})^{1/2}) \quad (7)$$

对于高斯分布 N_a 和 N_b , 它们由边界框 $A = (cx_a, cy_a, w_a, h_a)$ 和 $B = (cx_b, cy_b, w_b, h_b)$, 由式(7)可进一步简化为:

$$W_2^2(N_a, N_b) = \left\| \left[cx_a, cy_a, \frac{w_a}{2}, \frac{h_a}{2} \right], \left[cx_b, cy_b, \frac{w_b}{2}, \frac{h_b}{2} \right] \right\|_2^2 \quad (8)$$

因此, 利用它的指数形式归一化, 得到了新的度量, 称为归一化 Wasserstein 距离:

$$NWD = (N_a, N_b) = \exp\left(-\frac{\sqrt{W_2^2(N_a, N_b)}}{C}\right) \quad (9)$$

其中, C 为数据集密切相关的常数, $W_2^2(N_a, N_b)$ 是一个权重参数, 用于衡量真实框损失与预测框损失之间的关系。NWD 损失函数通过引入二维高斯分的 Wasserstein 距离能够更好地反映出预测框与真实框的重叠情况, 增强了小目标和微小目标的检测能力。

3 算法实验与分析

3.1 数据集及预处理

DIOR 数据集是遥感图像研究中广泛使用的一个大型公开数据集。该数据集由西北工业大学于 2019 年制作, 包含 190 288 个目标实例和 23 463 张遥感图像。在此基础上, 本文对这些目标实例进行了重新边框标注, 生成了

DIOR-R 数据集。同时, 将训练集、测试集和验证集按 7:2:1 划分。

DOTAv1.0 数据集是一个面向目标检测的公开遥感图像数据集, 包含 2 806 张图片。其标注方式分为水平标注和八参数标注, 图像分辨率在 $800 \times 800 \sim 4 000 \times 4 000$ 像素之间。本文将这些图像裁剪为 $1 024 \times 1 024$ 的尺寸, 按 7:2:1 划分训练集、测试集和验证集。

VEDAI 数据集是一个针对小目标的高分辨率遥感图像数据集。该数据集包含来自不同地理区域的多光谱和全色图像, 分辨率可达到亚米级, 包含 9 个类别, 能够提供丰富的地面细节信息。本文将其按 7:2:1 划分训练集、测试集和验证集。

针对遥感图像中小目标检测的挑战, 应用 Mosaic 和 MixUp 数据增强方法。Mosaic 通过将四张随机选择的图像拼接为一张新图像, 丰富了训练数据的背景多样性和目标分布, 使模型能够更有效地学习复杂场景下的特征。此外, Mosaic 在小目标检测任务中尤其有效, 因为拼接后的小目标比例相对增大, 能够显著增强模型对小目标的感知能力。另一方面, MixUp 通过对两张图像及其标签进行线性插值, 生成包含混合特征的新样本, 平滑了类别决策边界, 进一步提升模型的泛化能力和鲁棒性。在遥感图像场景中, Mosaic 和 MixUp 的结合可以有效扩展有限数据集的样本分布, 提高对小目标检测的准确性, 同时缓解过拟合问题。

3.2 实验环境配置

本实验平台为 ubuntu20.04 操作系统, 实验基于深度学习框架 Pytorch1.13.1, CPU 为 14 vCPU Intel(R) Xeon (R) Platinum 8362 CPU @ 2.80 GHz GPU 主机配备了 NVIDIA GeForce RTX 3090 24 G 显卡, 开发语言为 python3.9, CUDA 版本为 11.8, 设置训练轮数 DIOR-R 为 200, DOTAv1.0 为 300, VEDAI 为 400。批量大小为 8, 输入图像大小为 640×640 , Batch 为 8, 设置网格学习率为 0.01。

3.3 评价指标

用浮点运算量(floating point operations, FLOPs)衡量算法的复杂程度, 用参数量(number of parameters, Params)衡量模型轻量化程度, 用精确率(precision, P)、召回率(recall, R)和均值平均精度(mean average precision, mAP)作为对比评价网络性能的指标。一切预测得到的结果均记为正样本, 相应计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (10)$$

$$R = \frac{TP}{TP + FN} \quad (11)$$

$$AP = \int_0^1 P(R) dR \quad (12)$$

$$mAP = \frac{1}{\lceil n_c \rceil} \sum_{i=1}^{n_c} AP(i) \quad (13)$$

3.4 消融实验

为验证本文改进模块对于模型检测能力提升效果,在 DIOR-R 数据集上将不同模块对于模型检测性能进行消融实验。以 YOLOv8n 为基线模型,针对该模型逐个添加改进模块,首先应用 C2f-MSC(A 模块),然后添加 GLFPN(B 模块)和 Light-head(C 模块),最后引入 NWD(D 模块)。通过相同的实验条件评估不同模块对检测效果的影响,相关消融实验结果如表 1 所示。应用改进模块对小目标检测任务的 mAP@0.5 指标均取得了提升。在基线模型上应用 C2f-MSC 模块后,mAP@0.5 指标提升了 0.4%,参数量计算量均有减少,说明 C2f-MSC 模块增强了对小目标的检测能力;在颈部网络引入 GLFPN 模块后,mAP@0.5 较基线模型分别提高了 0.5%,参数量和计算量略有下降;引入 Light-head 检测头后 mAP@0.5 相比基线模型提升了 0.8%,参数量降低 47%,计算量下降 32%,模型轻量化同时进一步提升对小目标的检测定位能力;最后加入 NWD 损失函数 mAP@0.5 指标提高到 74.9%,所提改进方法相比基线模型 YOLOv8n 均有一定程度的提升。结果表明,同时应用 4 种改进模块能够促进检测性能,符合轻量级网络设计并对遥感图像小目标检测任务是有效的。

表 1 消融实验结果表

Table 1 Results of ablation experiment table

模块	P/%	mAP50/%	Param/M	FLOPs/G
baseline	83.3	73.6	3.0	8.1
A	83.1	74.0	2.8	7.9
A+B	83.2	74.1	2.7	7.9
A+B+C	84.5	74.4	1.6	5.5
A+B+D	83.5	74.6	2.7	7.9
A+C+D	84.0	74.3	2.2	6.1
A+B+C+D	85.0	74.9	1.6	5.5

3.5 对比实验

1) 损失函数对比

为评估 NWD 损失函数性能,将 NWD 与 CIoU(基线)、SIoU^[19]、ShapeIoU^[20]、WIoU^[21]、DIoU、EIoU 进行对比试验,实验结果如表 2 所示,NWD 损失函数的 map@0.5 达到 73.9%,略优于其他损失函数,验证了改进损失函数对遥感图像小目标检测精度有所提升。

2) 不同检测头对比

为验证 Light-head 检测头的检测性能,如表 3 所示,给出 Efficient-head、SEAM-head、CBAM-head 和 Light-head 在 DIOR-R 数据集相同基准模型下的比较结果。本文设计的 Light-head 检测头更轻量化,在检测性能上优于其他算法的检测头。这得益于 GroupNorm 和设计的多尺度卷积共享,大大减少参数量和计算量的同时,提高多尺度检测能力,增强对小目标特征的关注度。

表 2 损失函数对比实验

Table 2 Loss function comparison experiment %

方法	P	R	mAP50	mAP50~95
CIOU	83.3	68.6	73.6	51.9
SIOU	83.6	68.9	73.5	51.6
ShapeIOU	84.1	69.2	73.8	52.3
WIUO	84.2	69.1	73.7	52.4
DIOU	83.9	68.5	73.4	51.6
EIOU	84.2	69.6	73.2	52.2
NWD	84.6	69.8	73.9	53.0

表 3 不同检测头对比实验

Table 3 Comparison experiment of different detection heads

方法	P/%	mAP50/%	Params/M	FLOPs/G
Efficient-head	83.4	73.7	3.83	8.1
SEAM-head	82.6	74.0	2.82	7.0
CBAM-head	84.0	73.9	4.92	7.3
Light-head	84.4	74.2	1.71	5.5

3) 主流算法对比实验

为探究本文算法的检测性能,在 DIOR-R 数据集上进行对比实验。选取一些当前主流的目标检测算法 Faster-RCNN、VOD-YOLOv7^[22]、RT-DETR-L^[23]、YOLOX^[24]等进行对比实验,实验结果如表 4 所示,相比于其他主流目标检测方法,具备更高的检测性能,mAP@0.5、Precision 和 Recall 分别达到了 74.9%、85% 和 70.1%。在 mAP@0.5 上,本文算法相比于 Faster-RCNN 提升 5.1%、VOD-YOLOv7 提升 1%、YOLOv5 提升 2.2%、RT-DETR-L 提升 3.3%、YOLOX 提升 1%、YOLOv7-tiny 提升 0.8%、YOLOv8n 提升 1.3%。在准确率上比基准模型高出 1.7%,略高于其他主流模型,在检测上有一定优势。召回率上比基准模型高出 1.5%,略高于其他主流的模型。计算量和参数量最小,符合轻量化的同时,在遥感图像小目标检测上具有较好的性能。

表 4 主流算法对比实验

Table 4 Comparison experiment of mainstream algorithms

模型	P/%	R/%	mAP50/%	Params/M	FLOPs/G
Faster-RCNN	79.7	66.2	69.8	41.2	206.7
VOD-YOLOv7	83.9	68.8	73.9	44.4	116.0
YOLOv5	82.1	66.9	72.7	2.5	7.2
RT-DETR-L	80.6	67.7	71.6	32.8	100.9
YOLOX	82.8	68.4	73.9	8.9	26.7
YOLOv7-tiny	84.1	69.3	74.1	6.0	13.3
YOLOv8n	83.3	68.6	73.6	3.0	8.1
MGL-YOLO	85.0	70.1	74.9	1.6	5.5

4)先进模型对比实验

为验证模型良好的检测性,与现在较为先进的目标检测算法进行对比实验,结果如表 5 所示,本文算法在检测精度上高于 FFCA-YOLO、YOLOv10n 和 YOLO11n,进一步验证模型良好的检测性能。

表 5 先进模型对比实验

Table 5 Advanced model comparison experiment %

模型	P	R	mAP50	mAP50~95
YOLOv8n	83.3	68.6	73.6	51.9
FFCA-YOLO	83.5	69.1	73.9	52.6
YOLOv10n	84.5	69.5	74.2	53.1
YOLO11n	84.9	69.7	74.5	53.3
MGL-YOLO	85.0	70.1	74.9	53.8

3.6 模型的泛化性验证

1) DOTA v1.0 数据集泛化性验证

为验证本文算法模型在其他遥感图像目标数据集上泛化性，在公开数据集 DOTA v1.0 上进行泛化性实验。实验结果如表 6 所示，改进模型准确率、召回率和 mAP@0.5 相比基线模型和 YOLOv1 n 模型都有所提高，验证了 MGL-YOLO 模型良好的泛化性。

表 6 泛化性对比实验

Table 6 Generalization comparison experiment

模型	P	R	mAP50	mAP50~95
YOLOv8n	74.4	68.9	71.9	48.1
YOLO11n	75.1	69.5	72.4	50.8
MGL-YOLO	75.4	70.1	73.0	51.2

如图 10 所示为改进算法前后对比图,左侧为 YOLOv8n 右侧为本文算法的检测结果,可看出本文算法在检测小目标时优于 YOLOv8n 算法,在背景复杂的情况下,提高了检测精度和准确度,验证了改进模型的效果^[25]。

2) VEDAI 数据集泛化性验证

为验证本文算法在高分辨率和复杂环境下模型的泛化能力,在 VEDAI 数据集上进行泛化性验证,实验结果如表 7 和图 11 所示,由结果可知,本文算法在高分辨率遥感图像上的检测性能表现良好,对小目标有更高的检测精度。

表 7 泛化性对比实验

Table 7 Generalization comparison experiment (%)

模型	P	R	mAP50	mAP50~95
YOLOv8n	55.3	49.0	56.5	30.6
YOLO11n	56.8	49.6	57.7	31.2
MGL-YOLO	57.7	50.2	58.6	32.9

3.7 可视化分析

GradeCAM 用于生成 YOLOv8n 和 MGI-YOLO 的

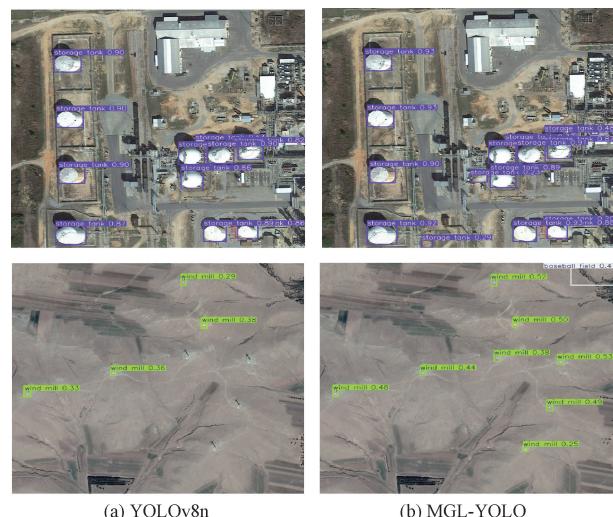


图 10 改进模型可视化对比

Fig. 10 Visual comparison of improved models

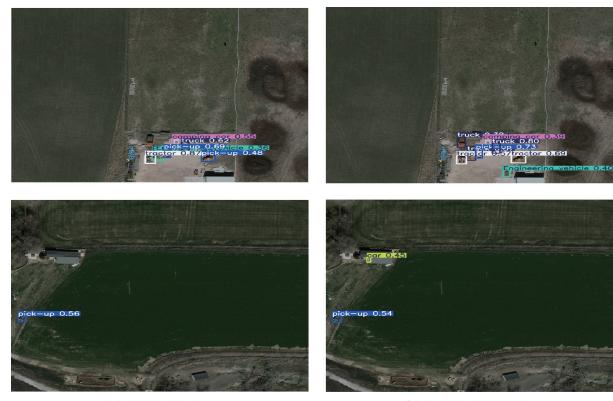


图 11 改进模型可视化对比

Fig. 11. Visual comparison of improved model.

热力图，模型类别的置信度梯度通过 GradeCAM 反向传播获得，其中高梯度的像素在热力图中显示为深红色，而低梯度的像素则呈现为深蓝色。实验结果如图 12 所示。YOLOv8n 对小目标的关注度较低，且对复杂场景不敏感。相对而言，本文模型对背景噪声有更强的抑制效果，对小

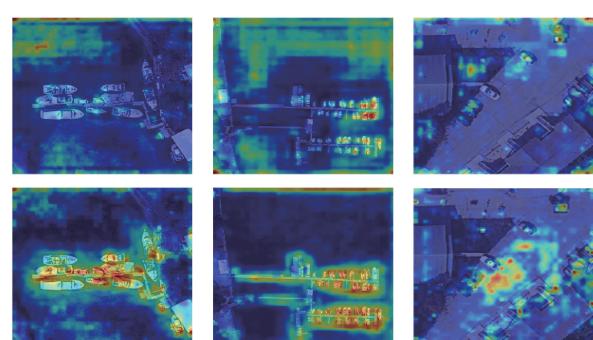


图 12 热力图可视化对比

Fig. 12 Visual comparison of heat maps

目标的关注度更高,并且模型注意力集中于物体的中心,使得预测的边界框更为精确,从而增强了整体检测性能。

为了清晰展示本算法的改进效果,在测试集图片上进行了逐张检测,并选取具有代表性的样本,如图 13(a)~(d)

所示,展示了背景复杂且目标多集中和高低分辨率地形复杂的检测效果。四组实验结果表明,改进后的模型可以有效避免漏检,并能更精确地定位遥感图像中的小目标。

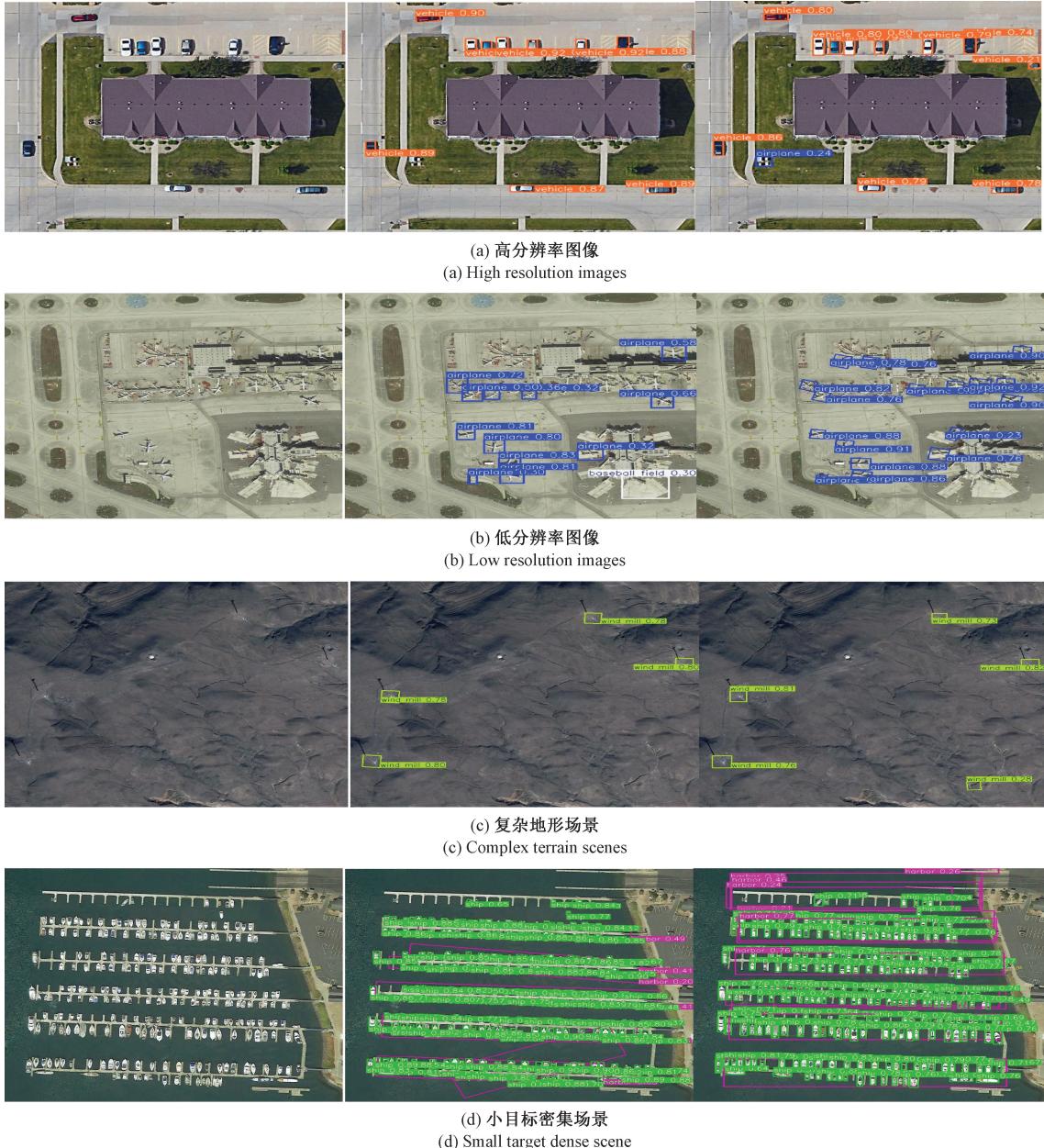


图 13 不同场景模型可视化结果对比

Fig. 13 Comparison of visualization results for different scenario models

4 结 论

针对遥感图像背景复杂、小目标分布密集、目标尺度多样、易受环境因素影响等问题,提出了 MGL-YOLO 算法。首先在 C2f 中用 MSConv 卷积替换 Bottleneck 块中的标准卷积,提高对遥感图像目标的检测准确率;其次在颈

部网络引入了 GLFPN 模块,轻量化模型提高模型的检测性能,引入 NWD 损失函数,加强复杂情况下的小目标检测精度;最后设计轻量检测头,再次轻量化的同时提高小目标的检测能力。通过在 DIOR-R、DOTA v1.0 和 VEDAI 数据集上验证,mAP@0.5 和准确率都有所提高,并且参数量降低 47%,计算量降低 32%。说明本文提出的轻量化

方法在遥感图像小目标检测任务上具备一定的优势。

参考文献

- [1] LI A B, GUO H, QI CH, et al. Dense object detection in remote sensing images under complex background [J]. Computer Engineering and Applications, 2023, 59(8): 247-253.
- [2] LAVANYA K, KARNICK S, GHALIB M R, et al. A novel method for vehicle detection in high-resolution aerial remote sensing images using YOLT approach[J]. Multimedia Tools and Applications, 2022, 81: 23551-23566.
- [3] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014: 580-587.
- [4] GIRSHICK R. Fast R-CNN [C]. 2015 IEEE International Conference on Computer Vision. Santiago: IEEE, 2015: 1440-1448.
- [5] REN SH Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal net works[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [6] WANG CH Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag of freebies sets new state-of-the-art for realtime object detectors[C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023: 7464-7475.
- [7] BETTI A, TUCCI M. YOLO-S: A lightweight and accurate yolo-like network for small target detection in aerial imagery[J]. Sensors, 2023, 23(4): 1865.
- [8] CHAURASIA D, PATRO B D K. Detection of objects in satellite and aerial imagery using channel and spatially attentive YOLO-CSL for surveillance [J]. Image and Vision Computing, 2024, 147: 105070.
- [9] ZHANG Y, YE M, ZHU G, et al. FFCA-YOLO: Object detection method for small objects in remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-15.
- [10] 桑雨,李立权,李铁. 轻量化YOLOv7-tiny的遥感图像小目标检测[J]. 科学技术与工程, 2024, 24(18): 07726-07.
- [11] SANG Y, LI L Q, LI T. Lightweight YOLOv7 tiny object detection in remote sensing images[J]. Science, Technology and Engineering, 2024, 24 (18): 07726-07.
- [12] 刘富宽,罗素云,何佳,等. FVIT-YOLO v8: 基于多尺度融合注意机制的改进 YOLO v8 小目标检测[J]. 红外技术, 2024, 46(8): 912-922.
- [13] LIU F K, LUO S Y, HE J, et al. FVIT-YOLO v8: Improved YOLO v8 small object detection based on multi scale fusion attention mechanism [J] Infrared Technology, 2024, 46(8): 912-922.
- [14] 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.
- [15] 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.
- [16] EEI L T, QI M H, JIN F W, et al. DuAT: Dual-aggregation transformer network for medical image segmentation [J]. ArXiv preprint arXiv: 2212.11677, 2022.
- [17] LI H, LI J, WEI H, et al. Slim-neck by GSconv: A better design paradigm of detector architectures for autonomous vehicles [J]. ArXiv preprint arXiv: 2405.03021, 2024.
- [18] 赵志宏,郝子晔. 改进 YOLOv8 的航拍小目标检测方法:CRP-YOLO[J]. 计算机工程与应用, 2024, 60(13): 209-218.
- [19] ZHAO ZH H, HAO Z Y. Improved aerial small detection method for YOLOv8: CRP-YOLO [J]. Computer Engineering and Applications, 2024, 60(13): 209-218.
- [20] YU X L, QI B H, ZHAO H ZH, et al. Large selective kernel network for remote sensing object detection[J]. ArXiv preprint arXiv:2023.0903, 2023.
- [21] TIAN ZH, SHEN CH, CHEN H. FCOS: fully convolutional one stage object detection [C]. 2019 IEEE/CVF International Conference on Computer Vision, 2019:9627-9636.
- [22] ZHENG ZH, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression [C]. AAAI conference on artificial intelligence, 2020, 34(7): 12993-13000.
- [23] ZHANG H, ZHANG SH. Shape-IoU: more accurate metric considering bounding box shape and scale[J].

- ArXiv preprint arXiv:2312.17663, 2023.
- [21] TONG ZH, CHEN Y, XU ZH, et al. Wise-IoU: Bounding box regression loss with dynamic focusing mechanism [J]. ArXiv preprint arXiv: 2301.10051, 2023.
- [22] YANG J, WU B, ZHANG H, et al. Rotating frame labeling and improved YOLOv7-based algorithm for vehicle object detection [C]. 2023 IEEE 6th International Conference on Pattern Recognition and Artificial Intelligence(PRAI), 2023:170-175.
- [23] LYU W, XU SH, ZHAN Y, et al. Detrs beat yolos on real time object detection [J]. ArXiv preprint arXiv: 2304.08069, 2023.
- [24] 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 CH. Improving YOLOX remote sensing image object detection algorithm[J] Computer
- Engineering and Applications, 2024, 60 (12): 181-188.
- [25] 邝先验,程福军,吴翠琴,等. 基于改进YOLOv7-tiny的高效轻量遥感图像目标检测方法[J]. 电子测量与仪器学报,2024,38(7):22-33.
KUANG X Q, CHENG F J, WU C Q. et al. Efficient and lightweight remote sensing image object detection method based on improved YOLOv7 tiny[J] Journal of Electronic Measurement and Instrumentation, 2024, 38(7): 22-33.

作者简介

葛雯,副教授,主要研究方向为图像处理与传输技术。

邵钰琦(通信作者),硕士研究生,主要研究方向为图像处理与传输技术。

E-mail:1215476176@qq.com

屈乐乐,教授,主要研究方向为雷达信号处理与成像技术。