

文章编号:1007-2780(XXXX)XX-0001-14

基于 YOLOv11n 的无人机航拍图像目标检测算法

卢永郁^{1,2}, 陈明霞^{1,2*}, 陆俊良^{1,2}, 檀昭亮^{1,2}, 邱琦升^{1,2}

(1. 广西高校先进制造与自动化技术重点实验室(桂林理工大学), 广西 桂林 541006;

2. 桂林理工大学 机械与控制工程学院, 广西 桂林 541006)

摘要:针对无人机航拍图像中小目标密集且尺寸小、背景复杂而导致的漏检与误检问题,本文提出基于 YOLOv11n 的改进 MCF-YOLO 检测算法。首先,设计了一个小目标特征增强模块 MSFE,通过保留高分辨率、加深卷积层并引入残差连接,增强小目标特征表达;其次,在 C3k2 模块中引入上下文锚点注意力(CAA),利用水平-垂直分离卷积捕获长距离空间依赖,提升全局上下文感知能力,解决背景干扰与目标遮挡问题;最后,采用 Focaler-DIoU 损失函数,通过引入动态样本权重调节机制,优化边界框回归过程从而提高小目标定位精度并减少误检和漏检现象。实验结果表明,改进后的算法在 VisDrone2019 数据集上相比基础 YOLOv11n 模型,准确率、召回率, mAP@0.5 和 mAP@0.5:0.95 分别提升了 6.8%、3.2%、6.6% 和 3.9%,并显著优于其他主流方法。此外,在 VEDAI 与 DOTA 数据集上的泛化性测试中,模型的 mAP@0.5 分别提升了 1.2% 和 2.2%,证明了所提方法在不同场景下的有效性和良好的适应能力。

关键词:深度学习;无人机;小目标检测;YOLOv11n;特征增强

中图分类号:TP391.41 文献标识码:A doi:10.37188/CJLCD.2026-0061 CSTR:32172.14.CJLCD.2026-0061

UAV aerial images object detection algorithm based on YOLOv11n

LU Yongyu^{1,2}, CHEN Mingxia^{1,2*}, LU Junliang^{1,2}, TAN Zhaoliang^{1,2}, QIU Qisheng^{1,2}

(1. Key Laboratory of Advanced Manufacturing and Automation Technology
(Guilin University of Technology), Education Department of Guangxi Zhuang Autonomous Region,
Guilin 541006, China;

2. School of Mechanical and Control Engineering, Guilin University of Technology,
Guilin 541006, China)

Abstract: Aiming to address the challenges of missed detections and false positives in UAV aerial imagery caused by densely distributed small objects, small sizes, and complex backgrounds, this paper proposes an improved MCF-YOLO object detection algorithm based on YOLOv11n. First, a Multi-scale Feature Enhancement (MSFE) module is designed to enhance feature representation of small targets through high-

收稿日期:2026-04-15;修订日期:2026-05-16.

基金项目:广西重点研发计划(桂科 AB22080093);梧州市中央引导地方科技发展资金(No.202201001);国家自然科学基金项目(No.61863009)

Supported by Guangxi Key R&D Program (Guike-AB22080093); Wuzhou Central Leading Local Science and Technology Development (No. 202201001); National Natural Science Foundation of China (No.61863009)

*通信联系人, E-mail:765359484@qq.com

resolution preservation, convolutional layer deepening, and residual connections. Second, a Contextual Anchor Attention (CAA) mechanism is integrated into the C3k2 module, leveraging horizontal and vertical separable convolutions to capture long-range spatial dependencies, thereby improving global contextual awareness and mitigating background interference and occlusion issues. Furthermore, the Focaler-DIoU loss function is employed to optimize the bounding box regression process through a dynamic sample weight adjustment mechanism, enhancing localization accuracy for small objects while reducing false positives and missed detections. Tests conducted on the VisDrone2019 dataset show that our method brings substantial performance enhancements relative to YOLOv11n, with increases of 6.8% in precision, 3.2% in recall, 6.6% in mAP@0.5, and 3.9% in mAP@0.5:0.95, outperforming other mainstream approaches. Additional cross-dataset validation on VEDAI and DOTA datasets shows mAP@0.5 improvements of 1.2% and 2.2% respectively, confirming the method's effectiveness and strong adaptability across various scenarios.

Key words: deep learning; unmanned aerial vehicle (UAV); small object detection; YOLOv11n; feature enhancement

1 引言

无人机航拍应用于军事侦察、智慧城市、农业监测等领域,需要高效精准的目标检测算法。但航拍图像背景复杂、目标尺度多变、小目标占比高且易遮挡,使传统方法特征表达不足、小目标漏检率高、复杂度与实时性难平衡^[1-2]。目标检测算法根据流程设计可分为双阶段与单阶段两大类。双阶段算法以 Faster R-CNN 系列^[3]为代表,在检测精度上表现优异,但流程复杂、计算量大,检测速度较慢。单阶段算法以 YOLO 系列^[4]和 SSD^[5]为代表,将检测视为回归问题,无需区域提议,端到端预测类别与位置,推理速度极快,适合实时,但在检测精度上尤其是小目标检测方面存在不足。

近年来无人机航拍小目标检测因目标密集、尺度多变成难点,研究者主要从 YOLO 改进与上下文利用两方面探索。为提升模型性能,许多研究者基于 YOLO 系列模型进行改进。Qi 等^[6]提出 MSFE-YOLO,通过对称 C2f 模块、多尺度注意力及特征融合模块增强小目标特征提取与融合能力。Tao 等^[7]构建的 MIS-YOLOv8 引入多级特征提取结构和高分辨率检测头,融合空间深度卷积与自适应特征融合机制,有效提升了对小目标细节的捕捉能力。雷帮军等^[8]改进 YOLOv8n 特征融合结构并引入跨层连接,新增极小目标检测层,结合动态检测头与 Inner-CIoU 损失

函数,进一步优化了小目标的定位与分类性能。面向海上场景,孔垂乐等^[9]提出小波变换效应卷积(WTEConv)以扩大感受野,构建多分支上采样结构(MUpsample)提升特征质量,并引入动态小目标检测头,实现了检测精度与效率的平衡。罗可心等^[10]提出 PRSU-YOLO,设计风车卷积自适应模块,通过多方向非对称卷积捕捉边缘纹理,并结合通道-空间自适应权重聚焦关键特征,提升小目标特征提取效果。这些改进虽然提升了模型的性能,却使得模型结构更加复杂。

为应对无人机影像中背景复杂、目标被遮挡的问题,研究者聚焦于利用上下文信息捕捉全局信息。彭晏飞等^[11]构建了上下文特征增强模块(CFEM),利用多尺度空洞卷积聚合多维度上下文从而捕捉全局信息,同时提出了 Alpha-EIoU,通过引入幂指数调整项以优化边界框回归的收敛过程。李云红等^[12]则提出了全局感知提取模块(GPE Block),其核心位置感知循环卷积(ParC)实现全局空间上下文交互。Wang 等^[13]在 RT-DETR 架构中融合傅里叶全局特征提取模块与动态位置偏置机制,构建多尺度特征增强金字塔,有效整合频域与空间域信息,提升模型在复杂场景下对小目标的检测鲁棒性。陈崇杨等^[14]设计了自适应选择空间注意力模块(AS-SA),并引入基于最小点距离的边界框回归损失函数 MPDIoU,实现动态调整感受野以融合目标周围不同范围的上下文。张志豪等^[15]提出

YOLOv11在功能上显著提升目标检测的精度与效率,适用于复杂场景下的实时检测任务。其结构设计既保障了特征表达的丰富性也实现了计算资源的节约,通过优化特征提取路径和注意力机制,增强对多尺度目标的识别能力,尤其在低光照、目标遮挡等挑战性场景中表现突出^[17-18]。此外,YOLOv11提供多种规格的模型(如n、s、m、l、x),其中“nano”(n)版本专为资源受限的嵌入式设备设计,在保持较高检测性能的同时实现了模型的轻量化,为无人机航拍、移动端部署等实际应用提供了可靠支持。

2.2 MCF-YOLO 模型简介

尽管YOLOv11在无人机航拍目标检测领域展现出显著的优势,但仍然存在不足:小目标易丢失细节导致漏检;复杂背景干扰多,目标尺度

差异大;同时上下文信息利用不充分及类别样本不均,制约了检测精度与鲁棒性。因此本文在YOLOv11n模型的基础上做了如下改进:

首先在骨干网第3层和检测头第19层引入MSFE模块,通过MSFE模块的高分辨率分支与快捷连接双分支结构,强化小目标细节保留;然后将整个网络中的C3k2替换为C3k2_CAA模块,通过引入上下文锚点注意力机制,捕获长距离空间依赖关系,抑制背景干扰,增强模型对复杂场景中全局上下文信息的感知能力;最后使用Focaler-DIoU损失函数,通过动态调整正负样本的权重,优化边界框回归过程,增强模型对小目标的定位能力,从而提升小目标检测精度。改进后的MCF-YOLO网络结构如图2所示。

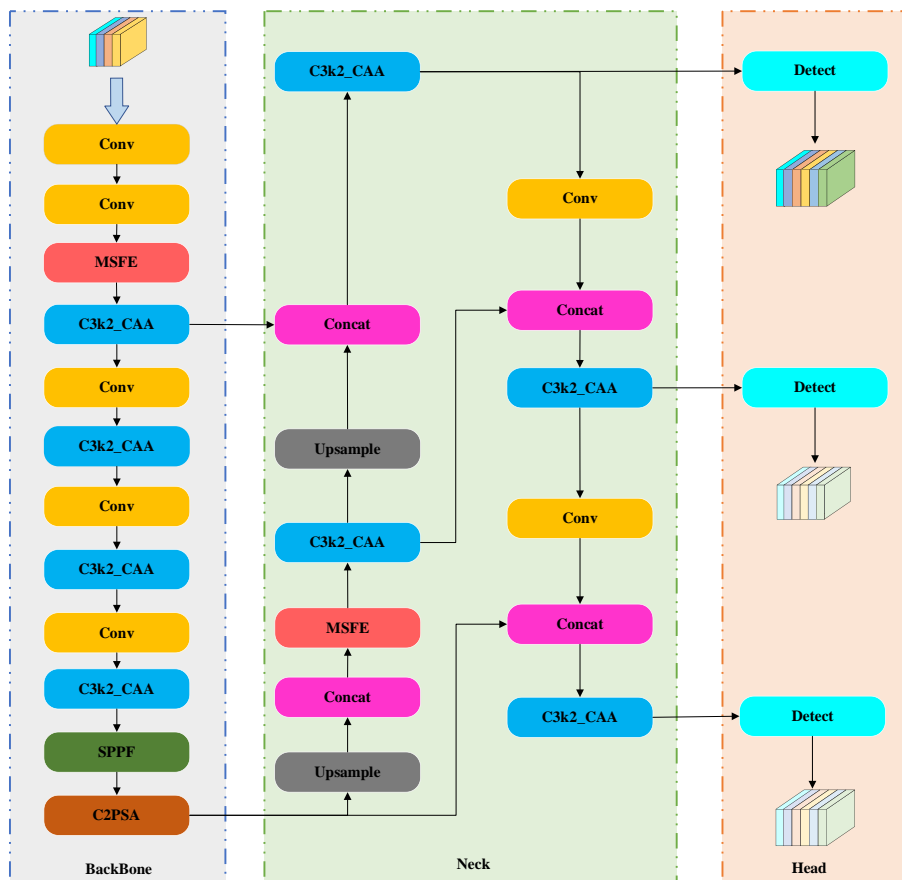


图2 MCF-YOLO网络结构

Fig. 2 Network structure of MCF-YOLO

2.2.1 MSFE 模块

MSFE模块采用精简而高效的双分支架构,即高分辨率分支(HRB)与快捷连接分支

(QCB)。双分支从特征提取与特征保真实现协同增效,针对小目标特征维度低、细节易丢失问题,通过扩大感受野、强化梯度传播实现特征增

强,有效解决小目标检测的核心问题,专门针对小目标检测优化设计。高分辨率分支起始于 1×1 卷积层,负责通道数调整和特征压缩,随后接入批归一化(BN)和SiLU激活函数进行非线性变换,归一化(BN)和SiLU激活函数公式如(1)和(2),该阶段通过调整通道维度适配与非线性表达完成初始化特征转换。

$$BN_n = \gamma_n \cdot \frac{x_n - \mu_n}{\sqrt{\sigma_n^2 + \epsilon}}, \quad (1)$$

$$SiLU = x \cdot \sigma(x), \quad (2)$$

式中, γ_n 为BN缩放系数; x 为分支输入特征图; μ 为均值; σ 为Sigmoid函数; ϵ 为归一化微小值; β 为BN偏移系数。

高分辨率分支核心部分采用深度可分离卷积结构,先通过 3×3 深度卷积进行空间特征提取,再通过 1×1 逐点卷积实现通道间信息交互进行高效空间特征提取。为强化小目标特征表示,模块额外增加了标准 3×3 卷积层,通过更大的感受野捕获局部上下文信息,每层卷积后均配有批归一化和SiLU激活以确保训练稳定性。同时,在高分辨率分支新增一个额外卷积层,其通过深化特征变换和增强局部感受野,强化小目标特征的提取与保留,防止其在网络深层传递中丢失关键细节,从而提升模型对细微特征的捕捉能力和检测精度。高分辨率分支工作完成后,其输出结果输出到相加融合机制进行下一步处理。高分辨率分支(HRB)结构如图3所示。

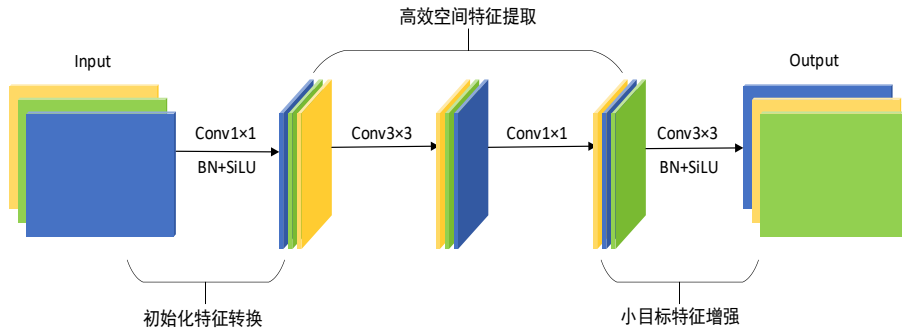


图3 HRB结构图

Fig. 3 Structure of HRB

快捷连接分支(QCB)是当输入输出通道相同时采用恒等映射(Identity Mapping, IM),直接将输入特征传递到输出端;通道不相同则通过 1×1 卷积与批归一化进行维度匹配。最终通过残差相加实现特征重用,有效缓解梯度消失问题。快捷连接分支(QCB)结构如图4所示。

恒等映射计算如公式(3)所示。

$$identity = \begin{cases} x, & c1 = c2 \\ BN_s(Conv_{1 \times 1}(x) + b_s), & c1 \neq c2 \end{cases}, \quad (3)$$

式中, $Conv_{1 \times 1}$ 为 1×1 卷积操作; $BN_s(\cdot)$ 为批量归一化操作。

高分辨率分支和快捷连接分支在每次前向传播过程中同时协同工作,高分辨率分支始终承担主要的特征提取与增强功能,通过多层卷积序列对输入特征进行深度处理以强化小目标的细节信息;而快捷连接分支则并行地建立输入到输出的直接通路,在输入输出通道匹配时采用恒等

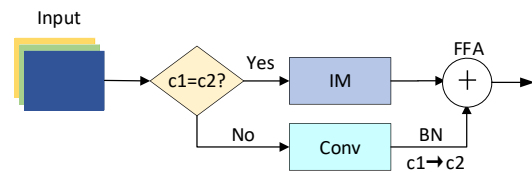


图4 QCB结构图

Fig. 4 Structure of QCB

映射保持特征原貌,在通道不匹配时通过 1×1 卷积进行维度调整,最终两个分支的输出通过相加融合(Feature Fusion Addition, FFA)机制结合,通过将高分辨率分支提取的增强特征与快捷连接的原始特征进行逐元素相加,实现特征重用与梯度稳定,同时平衡特征增强与信息完整性。MSFE模块结构如图5所示。

MSFE模块在YOLOv11中部署于两个关键位置,共同构建了小目标检测的增强体系。首先

在骨干网络的第3层,MSFE作用于早期特征提取阶段,利用此时特征图分辨率较高的优势,强化小目标的细节信息保留,防止后续下采样过程中的信息衰减。其次在头部网络的第19层,MSFE位于检测头之前,将处理后的特征图进行最终优化,用于提升小目标的检测精度。这两个位置形成特征提取到检测优化的协同机制:前者确保基础特征质量,后者保障最终检测性能,共同构建从特征提取到输出的小目标增强体系。MSFE模块采用残差连接,其恒等映射可将原始特征无损传递至输出,而高分辨率分支仅针对小目标进行上下文建模,因此模块在提升小目标性能的同时不会对大目标检测造成负面影响。

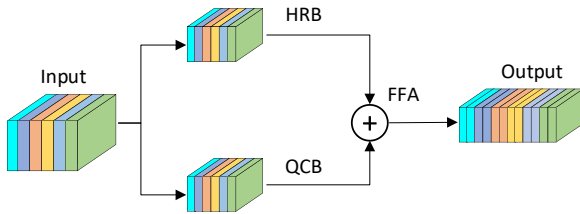


图5 MSFE模块结构图

Fig. 5 Structure of MSFE module

2.2.2 C3k2_CAA 模块

在无人机目标检测任务中,目标常常被树木、建筑物等其他目标干扰,进而影响检测结果。为解决这一问题,本研究在YOLOv11模型的C3k2中引入上下文锚点注意力机制(Context Anchor Attention, CAA)^[19]来捕捉长距离的上下文信息。CAA作为一种基于注意力机制的新型模块,主要用于优化目标检测中的特征提取流程。该模块通过引入上下文锚点来捕获图像中的关键区域或像素之间的远程依赖关系,从而自适应地调整模型对重要特征的关注,并抑制背景噪声或无关信息的干扰,利用全局平均池化和1D条形卷积来增强中心区域的特征。它能够有效解决复杂场景下的检测难题,通过捕捉全局上下文信息提升模型对目标特征的表达能力和检测精度。CAA结构如图6所示,CAA的结构主要包括平均池化、 1×1 卷积、深度可分离条带卷积和基于注意力的加权增强操作。

CAA计算过程如公式(4)所示:首先,通过

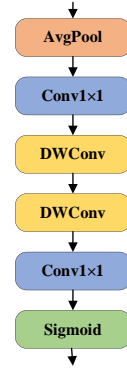


图6 CAA结构图

Fig. 6 Structure of CAA

平均池化和 1×1 卷积提取全局上下文特征,生成池化后的特征:

$$F_{pool} = Conv_{1 \times 1} [P_{avg}(F_1)], \quad (4)$$

式中, F_{pool} 为池化后的特征; P_{avg} 表示平均池化操作; $Conv$ 为卷积。接着使用两个深度可分离条带卷积分别沿宽度和高度方向捕获远程像素关系,生成中间特征 F_w 和 F_H ,计算如公式(5)和(6)所示:

$$F_w = DWConv_{1 \times k_h}(F_{pool}), \quad (5)$$

$$F_H = DWConv_{k_w \times 1}(F_w), \quad (6)$$

式中, F_w 、 F_H 分别表示宽度、高度方向的特征; $DWConv$ 为深度可分离卷积操作。最后,通过Sigmoid激活函数和 1×1 卷积输出加权的注意力特征图,计算如公式(7)所示:

$$F_2 = Sigmoid[Conv_{1 \times 1}(F_H)]. \quad (7)$$

C3k2_CAA模块结构如图7所示。C3k2_CAA模块是将原有的Bottleneck模块替换为CAA注意力模块,并继承C2f基础架构。该模块通过串联多个C3k或CAA子模块来增强特征提取能力,其中C3k块用于加强局部特征交互,而CAA(Context Anchor Attention)则通过水平和垂直方向的卷积捕获长距离空间依赖关系,增强模型对全局上下文的感知。

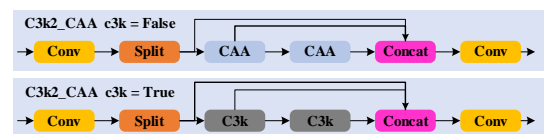


图7 C3k2_CAA结构图

Fig. 7 Structure of C3k2_CAA

在特征交互效率上,CAA模块通过深度可分离条形卷积与全局平均池化,以线性复杂度高效建模长距离依赖,减少传统自注意力的平方复杂度开销。C3k模块则专注于局部特征交互,通过小卷积核精细提取邻近区域的细节信息。CAA捕捉全局上下文与远距离语义关联,C3k保留局部纹理与边缘特征,二者形成互补,共同提升信息流动效率。

2.2.3 Focaler-DIoU 损失函数

在目标检测中,边界框回归的准确性直接决定了模型的定位能力。然而传统 CIoU 损失在处理小尺寸目标或具有复杂形状的物体时存在明显不足:当预测框与真实框的重叠程度较低时,损失函数对中心点距离和长宽比一致性的优化贡献容易被低 IoU 所压制,造成梯度信号微弱、优化方向不清晰,进而导致训练收敛缓慢,影响模型的整体检测性能。为解决这一难题,本文采用 Focaler-DIoU^[20]来替代原有的 CIoU 损失。Focaler-DIoU 主要通过线性区间映射策略,实现对原始 IoU 损失进行重构,使其能够根据不同任务需求灵活调整对各类回归样本的关注程度。其计算如公式(8)所示:

$$IoU^{focaler} = \begin{cases} 0, IoU < d \\ \frac{IoU - d}{u - d}, d \leq IoU \leq u, \\ 1, IoU > u \end{cases} \quad (8)$$

式中, IoU 为预测框与真实框之间的原始交并比值; d 为设定的下限阈值,当 IoU 值低于该阈值时,认为预测结果与真实框的匹配质量较差,属于定位困难的样本; u 为设定的上限阈值,当 IoU 值超过该阈值时,则认为两者匹配质量较高,属于容易学习的正样本。通过调节 d 和 u 的取值,可以灵活控制损失函数对不同质量回归样本的关注程度,从而实现了对难易样本的差异化学习。

DIoU Loss 则进一步考虑到预测框与真实框之间的欧氏距离和重叠面积,能够更精准地优化边界框的空间位置关系,尤其在小目标或部分遮挡情况下显著降低定位偏差,计算如公式(9)所示:

$$DIoU = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2}, \quad (9)$$

式中: ρ 表示预测框与真实框中心点之间的欧氏距离; c 为两框最小包围矩形的对角线长度; IoU

为交互比。

将 Focaler-IoU 损失引入至基于 IoU 的边界框回归损失计算过程,公式如式(10)所示。

$$L^{Focaler-DIoU} = L^{DIoU} + IoU - IoU^{Focaler}, \quad (10)$$

式中: L^{DIoU} 表示预测框与真实框之间在重叠度上的损失项,用于评估两者的空间匹配程度; IoU 为两个边界框交集面积与并集面积的比值,反映其原始重叠率; $IoU^{Focaler}$ 为基于原始的 IoU 值,通过线性区间映射等重构策略调整后得到的增强型相似性度量,旨在更合理地分配不同质量样本的优化权重,提升模型对关键样本的学习能力。

Focaler-DIoU 引入动态样本权重调节机制,通过学习参数实现动态样本权重自适应调节,增强低 IoU 正样本梯度、降低简单样本权重,避免模型过早收敛。有效提升难例(尤其小目标)学习能力,优化检测定位精度、抑制误检漏检,在复杂航拍图像中表现优异,鲁棒性与泛化性突出。并通过线性区间映射重构 IoU ,损失了区间 $[d, u]$ 外 IoU 的精细差异及区间内原始非线性梯度,但保留了 $DIoU$ 的中心距离惩罚。在实验中,它适配难易样本分布,在难样本,如小目标密集场景下大幅提升检测性能,易样本场景小幅优化。阈值设置不当可能存在潜在风险,但该问题可通过合理调整规避。

3 实 验

3.1 实验数据集和实验环境

VisDrone2019^[21]是一个广泛应用于无人机航拍目标检测研究的权威基准数据集,该数据集由天津大学 AISKYEYE 团队构建,通过无人机平台在多个城市的不同场景下采集。数据集总共包含 10 209 张静态航拍图像,并被划分为三个部分:训练集(6 471 张图像)、验证集(548 张图像)和测试集(1 610 张图像)。数据集中标注了超过 260 万个目标边界框,涵盖行人、汽车、自行车、卡车等 10 个常见类别。其包含大量小目标,其中像素面积小于 32×32 的目标占比高达 44.7%,包含了丰富的城市、乡村环境以及多样的光照和天气条件。

本实验在 Windows 11 操作系统环境下进行,硬件配置为 Intel(R) Core(TM) i7-13650HX 处理器(主频 2.2GHz)和 NVIDIA GeForce

RTX 4060 Laptop GPU, 配备 8GB 内存。软件环境基于 Anaconda 构建, Python 版本为 3.8.20, 深度学习框架采用 PyTorch2.0.1+cu118, 开发工具为 PyCharm, CUDA 版本为 11.8。详细实验设置信息见表 1。

表 1 实验参数设置

Tab. 1 Experimental parameter setting

参数	设置
epochs	200
lr0	0.01
patience	100
batch	8
images	640
workers	8
momentum	0.937
optimizer	SGD

3.2 评价指标

为评估模型的整体性能, 本实验使用的评估指标为平均精度均值(mAP)、精度(Precision, P)、召回率(Recall, R)、每秒帧数(Frames Per Second, FPS)和模型大小(Params)。精确度(P)和召回率(R)、分别如公式(11)和(12)所示:

$$P = \frac{TP}{TP + FP}, \quad (11)$$

$$R = \frac{TP}{TP + FN}, \quad (12)$$

其中 TP 表示真正例(True Positives); FP 表示假正例(False Positives); FN 表示假反例(False Negatives)。

平均精度均值(mAP)分别如公式(13)所示:

$$mAP@50 = \frac{1}{N} \sum_{i=1}^N AP_i@0.5, \quad (13)$$

其中 N 为类别总数, $AP_i@0.5$ 表示第 i 个类在 IoU 阈值为 0.5 时的平均精度。 mAP 体现了模型在各类别上综合权衡精确率与召回率的整体检测能力。 $mAP@0.5:0.95$ 则是对 IoU 从 0.5 到 0.95 以 0.05 为步长计算每个阈值下的 mAP 值并取平均。

3.3 消融实验

为验证本文各项改进策略对检测精度的提升作用, 在 VisDrone2019^[16] 数据集上使用 YOLOv11n 作为基线模型进行消融实验, 每组实验采用相同的训练策略, \checkmark 表示添加了该模块。实验结果如表 2 所示。

由表 2 可知, 本文引入的三项改进策略, 均有效提升了基础模型的检测性能。首先加入 MSFE 模块, 通过增强小目标特征显著性, 增强网络对小目标的特征提取能力。实验结果表明, $mAP@0.5$ 与 $mAP@0.5:0.95$ 分别提升 4.4% 和 2.9%, 精确率与召回率分别提升 5% 和 2.7%, 模型参数量仅增加 0.75M。其次, 引入 CAA 上下文锚点注意力机制重构 C3k2 (C3k2_CAA), 提高了模型全局上下文联系能力。模型加入 C3k2_CAA 后 $mAP@0.5$ 提高了 0.1%, 召回率提高 0.1%, 虽然精确率和 $mAP@0.5:0.95$ 略有下降, 但是参数量减少了 0.1M。最后, 加入 Focal-DIoU 损失函数, 模型 $mAP@0.5$ 和 $mAP@0.5:0.95$ 分别提高 0.4% 和 0.3%, 召回率提高 0.6%。

根据表中实验数据分析, 相较于 YOLOv11n 基准模型, 本文方法在检测性能上实现了显著提升。精确度与召回率分别增长 6.8% 和 3.2%, $mAP@0.5$ 与 $mAP@0.5:0.95$ 两项关键指标则分别提升 6.6% 和 3.9%。虽然本文改进策略带来了模型参数量的增加和推理速度的轻微下降,

表 2 消融实验

Tab. 2 Results of ablation studie

BaseLine	MSFE	C3k2_CAA	Focaler-DIoU	P/%	R/%	mAP@0.5/%	mAP@0.5:0.95/%	Params/M	FPS
\checkmark				43.6	33.0	32.5	18.8	2.59	136.7
\checkmark	\checkmark			48.1	35.4	36.5	21.3	3.34	114.5
\checkmark		\checkmark		44.3	34.1	33.6	19.5	2.48	117.8
\checkmark			\checkmark	43.9	33.6	33.1	19.1	2.59	136.7
\checkmark	\checkmark	\checkmark		49.9	35.9	37.6	22.4	3.23	103.2
\checkmark	\checkmark	\checkmark	\checkmark	50.3	36.1	38.1	22.7	3.23	103.2

但在实际部署中,模型依然能够稳定满足实时检测要求并保持卓越的检测精度与高精度检测性能,充分验证了改进策略的有效性与实用性。

3.4 对比实验

3.4.1 算法对比

为了验证本文所提出改进算法的有效性,将

其与该领域常用的其他算法进行了对比分析,包括 SSD、Faster-R-CNN、YOLOv5s、YOLOv6、YOLOv8n、YOLOv9t、YOLOv10n、YOLOv11n、YOLOv12n、PRSU-YOLO^[10]、改进YOLOv11^[22]、PE-YOLO^[23]以及RSC-YOLO^[24]算法。结果如表3所示。

表3 不同算法对比实验

Tab. 3 Comparison test of different algorithms

方法	mAP@0.5/%	P/%	R/%	Params/M	GFLOPs
SSD	22.5	36.8	26.5	24.6	89.5
Faster-R-CNN	32.8	44.3	37.1	41.1	203.3
YOLOv5s	32.1	37.8	33.2	7.08	16.7
YOLOv6	29.4	39.6	30.2	4.23	11.9
YOLOv8n	33.1	43.4	33.2	3.00	8.9
YOLOv9t	33.1	44.5	32.5	2.00	8.5
YOLOv10n	32.3	43.1	32.0	2.69	8.4
YOLOv11n	32.5	43.6	33.0	2.59	6.6
YOLOv12n	32.5	43.4	32.7	2.55	6.7
改进YOLOv11	34.2	44.9	34.2	3.18	—
PE-YOLO	37.1	49.7	34.4	2.86	7.2
PRSU-YOLO	37.4	49.9	37.9	—	53.4
RSC-YOLO	37.7	42.5	—	2.70	8.7
MCF-YOLO(Ours)	38.1	50.3	36.1	3.23	9.6

从表3可以看出,两阶段算法Faster-R-CNN与单阶段算法SSD不仅推理速度偏慢,参数规模也显著较高,其mAP@0.5较本文提出的MCF-YOLO分别低5.3%和15.6%;相较于YOLO系列主流轻量化模型,MCF-YOLO的mAP@0.5均实现明显提升,其中比YOLOv8n、YOLOv10n、YOLOv12n分别高出5.0%、5.8%、5.6%,优于同类型轻量化模型的检测精度;与PRSU-YOLO、改进YOLOv11、PE-YOLO、RSC-YOLO等优化算法相比,MCF-YOLO的mAP@0.5仍保持领先优势,分别提升0.4%、3.9%、1.0%、0.4%。在参数量与计算量方面,MCF-YOLO的参数量为3.23M,计算量为9.6GFLOPs,虽略高于部分轻量化模型,但远低于Faster-R-CNN和SSD,且其精确度和召回率在所有对比算法中表现出突出检测精度与计算效率。

综上,本文提出的MCF-YOLO算法在确保较低计算开销与合理参数规模的基础上,显著提

升了无人机目标检测的准确率,实现了检测精度与速度的高效平衡,相较于当下主流及改进检测算法更具综合优势。虽然本文所提方法在一定程度上提升了模型参数量与计算量,但仍可满足工程实际中的实时检测需求,同时保障了检测精度,证明了本文改进策略的有效性与实用性。

3.4.2 特征增强模块对比

为验证MSFE特征增强模块的优势,以MCF-YOLO模型为基础,将本文设计的MSFE模块与其他的特征增强模块进行对比,主要有DSAM^[25]、PKIBlock^[19]和CAF^[26],实验结果如表4所示。

由表中结果可以看出,模型加入MSFE模块后在精确率P上较DSAM、PKIBlock与CAF分别提升3.5%、2.0%和1.2%,mAP@0.5分别提升2.3%、1.6%和0.9%,虽然模型加入CAF模块后在参数量与计算量上表现最佳,但综合精度与效率,MSFE模块较其他模块表现更

表4 不同特征增强模块对比实验

Tab. 4 Comparison of different feature enhancement modules

MCF-YOLO	$P/\%$	$R/\%$	mAP@0.5/%	Params/M	GFLOPs
+DSAM	46.8	34.3	35.8	2.87	15.8
+PKIBlock	48.3	35.2	36.5	3.77	10.5
+CAFEM	49.1	35.7	37.2	3.09	7.6
+MSFE	50.3	36.1	38.1	3.23	9.6

好,证明了MSFE模块在平衡精度与计算成本方面的优越性。

3.4.3 下文注意力模块对比

为了验证上下文锚点注意力CAA的可行

性,以MCF-YOLO模型为基础,本文将其与其他上下文注意力模块进行对比,包括ContextGuided^[27]、ContextAggregation^[28]、与MSCA^[29],实验结果如表5所示。

表5 不同上下文注意力模块对比实验

Tab. 5 Comparison test of different context attention modules

MCF-YOLO	$P/\%$	$R/\%$	mAP@0.5/%	mAP@0.5:0.95/%
+ContextGuided	49.3	35.5	37.2	21.1
+ContextAggregation	49.8	35.3	37.5	21.9
+MSCA	49.5	35.8	37.8	22.2
+CAA	50.3	36.1	38.1	22.7

由表中结果可以看出,模型加入CAA模块后较加入ContextGuided、ContextAggregation与MSCA后在精确度P上分别提升了1.0%、0.5%和0.8%,mAP@0.5分别提升0.9%、0.6%和0.3%,mAP@0.5:0.95分别提升1.6%、0.8%和0.5%,同时其召回率也为所有模块中最高,证明了引入的上下文锚点注意力CAA模块在模型中的有效性与优越性。

3.4.4 损失函数对比

为了验证Focaler-DIoU损失函数的有效性,本文将其与其他损失函数进行对比,包括CIoU、Focaler-GIoU与Focaler-CIoU,实验结果如表6所示。

表6 不同损失函数对比实验

Tab. 6 Comparison test of different losses

MCF-YOLO	$P/\%$	$R/\%$	mAP@0.5/%	mAP@0.5:0.95/%
+CIoU	49.9	35.9	37.6	22.4
+Focaler-GIoU	49.4	36.1	37.3	22.1
+Focaler-CIoU	48.5	36.7	37.6	22.3
+Focaler-DIoU	50.3	36.1	38.1	22.7

由表中结果可以看出,虽然模型加入Focaler-CIoU损失函数后在召回率上表现最佳,但模型加入Focaler-DIoU损失函数后在精确度P上较CIoU、Focaler-GIoU与Focaler-CIoU分别提升了0.5%、0.9%和1.8%,mAP@0.5分别提升了0.5%、0.8%和0.5%,mAP@0.5:0.95分别提升了0.3%、0.6%和0.4%,证明了Focaler-DIoU在模型中发挥重要作用。

3.5 可视化实验

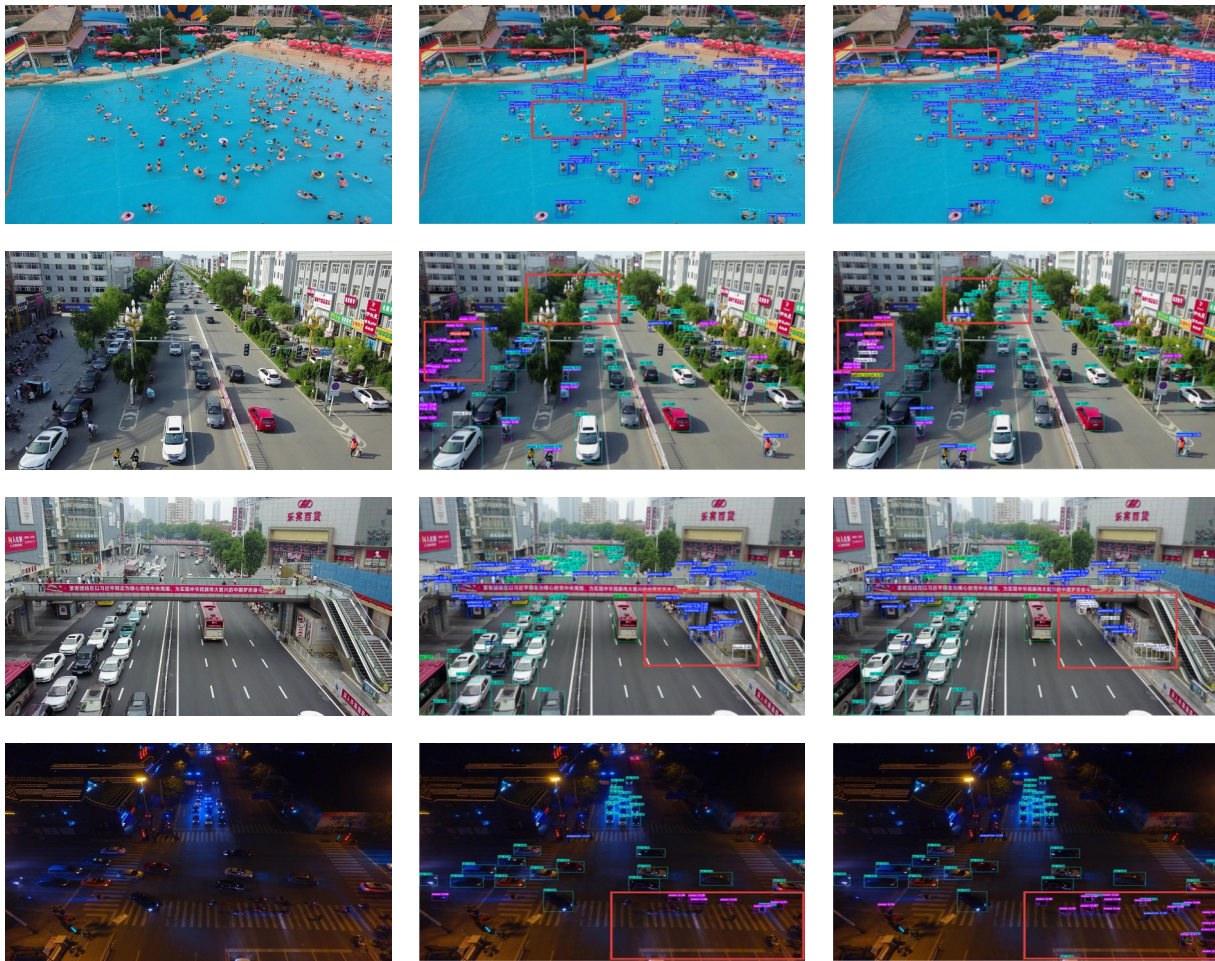
为直观对比本文所提改进算法MCF-YOLO与基准模型YOLOv11n在无人机航拍场景下的性能差距,在VisDrone2019数据集上选取四类典型复杂场景开展可视化对比试验,包括小目标密集分布、光照条件剧烈变化、多尺度目标场景以及夜间场景。其中,小目标密集分布与光照条件剧烈变化场景均为航拍图像目标检测的常见难点场景;多尺度目标场景则表示同一场景内同时存在尺寸悬殊的目标(如行人、小汽车与货车),对模型浅层细节捕捉与深层语义理解的综合能力提出更高要求;夜间场景是指夜间光线不足易致航拍图像特征模糊,对模型的特征提取带来较大困难。试验结果如图8所示,左侧为原始航拍

图,中间为YOLOv11n检测效果,右侧为MCF-YOLO检测效果,二者差异以红色框标注,便于直观观察比较。

在小目标密集场景下,MCF-YOLO相比YOLOv11n明显减少了漏检情况,即便目标相互遮挡堆叠,也能更精准地定位并识别出更多微小目标;在光照大幅波动的环境中,MCF-YOLO可有效削弱背景噪声干扰,同时完整保留目标关键判别特征,比基准模型检出更多有效目标,光照适应能力更突出;针对多尺度目标共存场景,

MCF-YOLO具备良好的尺度适配性,可精准提取行人、车辆等不同大小目标的特征,明显提升常规目标类别检测精度;在夜间低照度环境下,MCF-YOLO有效改善了暗光下识别困难的问题,较YOLOv11n检测出更多行人与非机动车,且置信度更高,弱光检测性能更具优势。

结果表明,MCF-YOLO在各类复杂无人机航拍场景下均能实现更优的检测效果,尤其在光照环境多变的条件下展现出更为出色的鲁棒性。



(a) 原图 (b) YOLOv11n模型 (c) MCF-YOLO模型
 (a) Original images (b) YOLOv11n (c) MCF-YOLO

图8 可视化检测结果对比

Fig. 8 Comparison of visualized detection result

3.6 泛化实验

为验证提出模型的泛化能力,本实验分别将基线算法YOLOv11n和MCF-YOLO改进算法在

VEDAI数据集^[30]和DOTA数据集^[31]进行实验。在实验中,已将DOTA数据集的旋转边界框(OBB)转换为水平边界框(HBB)。实验结果如表7所示。

表7 数据集泛化实验

Tab. 7 Comparison test of dataset generalization

数据集/方法	P/%	R/%	mAP@0.5/%	mAP@0.5:0.95/%
VEDAI/YOLOv11n	52.7	64.7	64.0	38.1
VEDAI/MCF-YOLO	55.3	66.6	65.2	39.7
DOTA/YOLOv11n	52.3	54.3	50.2	33.1
DOTA/MCF-YOLO	53.6	55.9	51.5	34.2

泛化实验结果表明,MCF-YOLO改进算法在VEDAI数据集上的精确度P、召回率R、mAP@0.5和mAP@0.5:0.95较YOLOv11n分别提升了2.6%,1.9%,1.2%和1.6%,在更为复杂的DOTA数据集上则分别提升了1.3%,1.6%,1.3%和1.1%,在两个数据集的所有评估指标上均一致性地超越了YOLOv11n基线模型,有力证明了我们所提方法具有更优的泛化能力和鲁棒性。

4 结 论

针对无人机航拍图像中小目标密集、背景干

扰严重导致检测精度下降的问题,本文提出基于YOLOv11的改进算法MCF-YOLO,分别从构建MSFE特征增强模块,引入上下文锚点注意力CAA和改进Focaler-DIoU损失函数三方面进行改进。在VisDrone2019数据集上,该方法在精确率、召回率、mAP@0.5和mAP@0.5:0.95等指标上显著优于YOLOv11n,分别提升了6.8%、3.2%、6.6%和3.9%。同时,在VEDAI与DOTA等具有不同特点的遥感图像数据集上的泛化实验进一步表明,该算法具备优秀的泛化性能与场景适应性,为无人机遥感图像中小目标检测提供了一种可行的技术路径。

参 考 文 献:

- [1] 吴一全, 童康. 基于深度学习的无人机航拍图像小目标检测研究进展[J]. 航空学报, 2025, 46(3): 030848.
Wu Y Q, Tong K. Research advances on deep learning-based small object detection in UAV aerial images[J]. *Acta Aeronautica et Astronautica Sinica*, 2025, 46(3): 030848. doi:10. 7527/S1000-6893. 2024. 30848. (in Chinese)
- [2] 金黎威, 徐望明, 李焱翔. 基于自适应切片辅助推理的航拍图像目标检测方法[J]. 液晶与显示, 2025, 40(03): 472-480.
Jin L W, Xu W M, Li Y X. Object detection method for aerial images based on adaptive slicing aided inference[J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(3): 472-480. (in Chinese)
- [3] Ren S, 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] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]// *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 779-788.
- [5] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single S-hot MultiBox Detector. [J]. *CoRR*, 2015, abs/1512.02325.
- [6] Qi S H, Song X F, Shang T F, et al. MSFE-YOLO: An Improved YOLOv8 Network for Object Detection on Drone View[J]. *IEEE Geoscience and Remote Sensing Letters*, 2024, 21: 1-5.
- [7] Tao S, Shengqi Y, Haiying L, et al. MIS-YOLOv8: An Improved Algorithm for Detecting Small Objects in UAV Aerial Photography Based on YOLOv8 [J]. *IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT*, 2025, 74.
- [8] 雷帮军, 余翔, 吴正平, 等. 改进YOLOv8n的无人机航拍小目标检测算法[J]. 现代电子技术, 2025, 48(3): 26-34.
Lei B J, Yu A, Wu Z P, et al. Improved small object detection algorithm based on YOLOv8n for UA-V aerial photography[J]. *Modern Electronics Technique*, 2025, 48(3): 26-34. (in Chinese)

- [9] 孔垂乐,孟昱煜,火久元,等.改进YOLOv11的无人机海上小目标检测算法[J/OL].计算机工程与应用,1-15 [2025-10-26].<https://link.cnki.net/urlid/11.2127.TP.20250708.1531.004>.
- Kong C L, Meng Y Y, Huo J Y, *et al.* Improved UAV Maritime Small Target Detection Algorithm for YOLOv11 [J/OL]. *Computer Engineering and Applications*, 1-15 [2025-10-26]. <https://link.cnki.net/urlid/11.2127.TP.20250708.1531.004>. (in Chinese)
- [10] 罗可心,李松江,王鹏,等.面向无人机影像小目标检测的轻量化算法[J].液晶与显示,2026,41(02):253-266.
- Luo K X, Li S J, Wang P, *et al.* Lightweight algorithm for small object detection in UAV images [J]. *Chinese Journal of Liquid Crystals and Displays*, 2026,41(02):253-266. (in Chinese)
- [11] 彭晏飞,赵涛,陈炎康,等.基于上下文信息与特征细化的无人机小目标检测算法[J].计算机工程与应用,2024,60(05):183-190.
- Peng Y F, Zhao T, Chen Y K, *et al.* UAV Small Object Detection Algorithm Based on Context Information and Feature Refinement [J]. *Computer Engineering and Applications*, 2024,60(05):183-190. (in Chinese)
- [12] 李云红,张富星,苏雪平,等.增强上下文特征交互的实时无人机影像分割算法[J/OL].北京航空航天大学学报,1-12[2025-10-26].<https://doi.org/10.13700/j.bh.1001-5965.2023.0830>.
- Li Y H, Zhang F X, Su X P, *et al.* Real-time UA-V image segmentation algorithm with enhanced contextual featureinteraction [J/OL]. *Journal of Beijing University of Aeronautics and Astronautics*, 1-12 [2025-10-26]. <https://doi.org/10.13700/j.bh.1001-5965.2023.0830>. (in Chinese)
- [13] Wang H, Yu Y, Tang Z. FDM-RTDETR : A Multi-Scale Small Target Detection Algorithm [J]. *IEEE ACCESS*, 2025,13:88747-88761.
- [14] 陈崇杨,彭力,杨杰龙.基于特征增强与上下文融合的无人机小目标检测算法[J/OL].计算机科学,<https://link.c-nki.net/urlid/50.1075.TP.20241225.1945.022>.
- Chen C Y, Peng L, Yang J L. UAV Small Object Detection Algorithm Based on Feature Enhancement and Context Fusion [J/OL]. *Computer Science*, <https://link.c-nki.net/urlid/50.1075.TP.20241225.1945.022>. (in Chinese)
- [15] 张志豪,厉小润,陈淑涵.基于改进YOLO11的无人机航拍图像小目标检测算法[J].液晶与显示,2025,40(06):915-930.
- Zhang Z H, Li X R, Chen S H. Small object detection algorithm in UAV aerial images based on improved YOLO11 [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(6) : 915-930. (in Chinese)
- [16] Rahima K, Muhammad H. YOLOv11: an Overview of the Key Architectural Enhancements [J]. *CoRR*, 2024, abs/2410.17725.
- [17] 吕学涵,李富,祁铭瑞,等.基于改进YOLOv11s的无人机小目标检测算法[J].液晶与显示,2025,40(11):1744-1756.
- LÜ X H, Li F, Qi M R, *et al.* Target detection algorithm based on improved YOLOv11s UAV aerial image [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(11) : 1744-1756. (in Chinese)
- [18] 张志豪,杜丽霞,侯越,等.跨层注意力交互下的多特征交叉无人机图像检测[J].光学精密工程,2024,32(24):3616-3631.
- Zhang Z H, Du L H, Hou Y, *et al.* Multi-feature cross UAV image detection algorithm under cross-layer attentional interaction [J]. *Optics and Precision Engineering*, 2024,32(24):3616-3631. (in Chinese)
- [19] Cai X H, Lai Q X, Wang Y W, *et al.* Poly Kernel Inception Network for Remote Sensing Detection [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024:27706-27716.
- [20] ZHANG H, ZHANG S J. Focaler-IoU: more focused intersection over union loss [J/OL]. *arXiv*, 2024: 2401.10525.
- [21] DU D W, ZHU P, WEN L, *et al.* VisDrone-DET2019: The Vision Meets Drone Object Detection in Image Challenge Results [C]//2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), 2019, :213-226.
- [22] 朱嘉浩,黄俊,刘兆金.基于YOLOv11的小目标检测模型[J].激光杂志,2025,46(11):49-56.
- Zhu J H, Huang J, Liu Z J. Small object detection model based on YOLOV11 [J]. *LASER JOURNAL*, 2025,46

- (11):49-56.
- [23] 景婷婷,曹玉东,陈鑫,等.改进YOLOv11的无人机航拍小目标检测算法[J].计算机工程与应用,2025:1-12. (in Chinese)
Jing T T, Cao Y D, Chen X, *et al.* Improved YOLOv11 Algorithm for Aerial Small Target Detection in UAVs [J]. *Computer Engineering and Applications*, 2025:1-12. (in Chinese)
- [24] 梁秀满,张永胜,吴楠,等.基于RSC-YOLO的无人机目标检测算法[J].红外技术,2025,47(12):1491-1501.
Liang X M, Chen Y S, Wu N, *et al.* UAV Object Detection Algorithm Based on RSC-YOLO [J]. *Infrared Technology*, 2025,47(12):1491-1501. (in Chinese)
- [25] BAOLONG N, ZHANG Chenyu, SHI Yuzhi, *et al.* DeBiFormer: vision transformer with deformable agentbi-level routing attention[M]//Computer Vision - ACCV2024. Singapore: Springer Nature Singapore, 2024: 445–462.
- [26] Hu S, Gao F, Zhou X, *et al.* Hybrid Convolutional and Attention Network for Hyperspectral Image Denoising[J]. *IEEE GEOSCIENCE AND REMOTE SENSING LETTERS*, 2024,21.
- [27] Wu T, Tang S, Zhang R, *et al.* CGNet: A Light-Weight Context Guided Network for Semantic Segmentation[J]. *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 2021,30:1169-1179.
- [28] Gao P, Lu J, Li H, *et al.* Container: Context aggregation network[J]. *arXiv preprint arXiv:2106.01401*, 2021.
- [29] Guo M H, Lu C Z, Hou Q, *et al.* Segnext: rethinkingconvolutional attention design for semanticsegmentation[C]. Advances in Neural InformationProcessing Systems, 2022: 1140-1156.
- [30] Mostofa M, Ferdous S N, Riggan B S, *et al.* Joint-SRVDNet: Joint Super Resolution and Vehicle Detection Network[J]. *IEEE Access*, 2020,8:82306-82319.
- [31] XIA G S, BAI X, DING J, *et al.* DOTA: a large-scale dataset for object detection in aerial images [C]// *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT, USA:IEEE,2018:3974-3983.

作者简介:



卢永郁,男,硕士研究生,2024年于广西桂林理工大学获得学士学位,主要从事遥感图像处理方面的研究。
E-mail:1976537204@qq.com



陈明霞,女,学士,教授级高级工程师,1993年于大连理工大学获得学士学位,主要从事装备自动化技术、过程控制系统和运动控制方面的研究。
E-mail: 765359484@qq.com