

DOI:10.19651/j.cnki.emt.2519611

无人机航拍小目标检测研究^{*}

崔博斌 易军凯 谭玲玲

(北京信息科技大学机电工程学院 北京 102206)

摘要: 针对无人机航拍图像中小目标占比较高,目标尺度差异大以及背景复杂等因素导致的目标检测精度不足等问题,并结合边缘设备算力与功耗受限的特点,在YOLOv8n的基础上提出了一种改进的目标检测算法——EGD-YOLO。首先,增加用于小目标检测的P2层并移除用于大目标检测的P5层,同时采用浅层通道扩展策略以增强小目标特征表达能力;其次,设计了多尺度特征融合与加权特征融合级联的自适应多层次全局特征融合架构,实现颈部网络跨尺度语义信息的高效传播与深度整合;最后,采用具有多重注意力机制的DyHead动态检测头,进一步优化模型的小目标检测性能。在VisDrone2019数据集的实验结果表明,所提EGD-YOLO相较于基准模型在mAP_{0.5}和mAP_{0.5:0.95}等指标上分别提升了12.0%和8.6%的同时保持了良好的计算优势;在DOTA数据集的实验结果进一步验证了该方法具有良好的泛化能力,为无人机航拍小目标检测提供了有效的解决方案。

关键词: 无人机;小目标;YOLO;特征融合

中图分类号: TP391.41;TN919.82 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Research on small target detection in UAV aerial photography

Cui Bobin Yi Junkai Tan Lingling

(School of Mechanical and Electrical Engineering, Beijing Information Science and Technology University, Beijing 102206, China)

Abstract: Addressing the issues of insufficient object detection accuracy in UAV aerial images caused by factors such as high proportion of small targets, large scale differences among targets, and complex backgrounds, and considering the limited computational power and power consumption of edge devices, this paper proposes an improved object detection algorithm called EGD-YOLO based on YOLOv8n. First, a P2 layer for small target detection is added while the P5 layer for large target detection is removed, and the shallow channel expansion strategy is adopted to enhance the feature representation capability for small targets. Secondly, a global hierarchical fusion architecture cascading Multi-scale feature fusion and weighted feature fusion was designed to achieve efficient propagation and deep integration of cross-scale semantic information in the neck network. Finally, a DyHead dynamic detection head with multiple attention mechanisms is employed to further optimize the model's small target detection performance. Experiments on the VisDrone2019 dataset demonstrate that the proposed EGD-YOLO achieves improvements of 12.0% in mAP@0.5 and 8.6% in mAP@0.5:0.95 over the baseline while maintaining a clear computational advantage; results on the DOTA dataset further confirm its strong generalization capability, providing an effective solution for small-object detection in UAV aerial imagery.

Keywords: UAV; small object; YOLO; feature fusion

0 引言

随着无人机航拍技术的快速发展及应用成本的不断降低,其在交通监控、智慧农业、应急救援、环境监测等领域得到广泛应用。然而,受飞行高度和地形条件的限制,无人机所拍摄的图像中小目标占比较高且目标尺寸变化较大,这

增加了目标检测任务的难度。此外,无人机作为边缘设备,其算力和功耗限制要求检测算法具备轻量化特性。因此,如何在保持较低计算资源需求的前提下实现高精度的无人机航拍目标检测具有重要研究意义。

传统的目标检测算法主要依赖手动提取特征方法,例如方向梯度直方图、尺度不变特征变换和相关滤波等^[1]。

收稿日期:2025-08-15

* 基金项目:国家重点研发计划(2024QY1703)项目资助

这些方法在一些简单场景中表现良好,但在处理复杂场景或者小目标检测任务时存在准确性和鲁棒性不足的问题。

随着计算机性能的不不断提升和深度学习技术的快速发展,基于深度学习的目标检测算法逐渐取代了传统的目标检测算法。这些算法主要分两类:以 YOLO 系列^[2-3]和 SSD^[4]为代表的一阶段检测模型和以 R-CNN^[5]和 Faster R-CNN^[6]为代表的两阶段检测模型。在实际应用中,研究者们不断改进这些基础算法,以进一步提升检测性能。Ye 等^[7]设计了基于 CNN 和 Transformer 的实时目标检测网络 RTD-Net,结合轻量级特征提取模块和高效的卷积 Transformer 模块,有效提升了对复杂背景下遮挡目标和小目标的检测能力。Ding 等^[8]结合单次多框检测器和自适应管道过滤器,通过增强 SSD 的特征层并应用时间滤波来减少误报,进而提升了模型精度。Kang 等^[9]提出一种新的“对齐匹配”策略,该策略在 IoU 的基础上引入纵横比和中心点距离,提升了 SSD 对小目标的检测性能。Li 等^[10]设计的密集嵌套注意力网络 DNA-Net,通过引入密集嵌套互动模块和级联的通道-空间注意力机制,实现渐进式特征交互和自适应特征增强,有效解决了红外小目标在深层网络中容易丢失的问题。Wei 等^[11]基于 RT-DETR 提出了改进的 DV-DETR 模型,采用 ResNet18 作为骨干网络,结合了重校准注意力单元和可变形注意力机制,通过优化小目标检测的尺度感知和空间定位能力,提升了无人机在高视角和动态场景下的目标检测精度。Kong 等^[12]设计了轻量级骨干网络 ESDNet,通过混合池化下采样模块和浅层特征增强模块,结合增强双路径特征融合注意力模块,有效解决了无人机航拍图像中小目标在下采样过程中的特征丢失问题。Ju 等^[13]提出了一种简单高效的网络结构,通过设计扩张模块扩展感受野、应用特征融合改善小目标检测能力、使用通道模块结合精细信息与语义信息,提高小目标的检测效果。Shang 等^[14]提出了基于 YOLOv5s 的小目标检测改进算法。通过添加小目标检测层、采用增强加权双向特征金字塔替代 PANet 网络,以及引入 M-SimAM 注意力模块,改善了 YOLOv5s 在复杂场景中的小目标检测性能。Zuo 等^[15]针对无人机对无人机小目标检测的复杂性,提出了 UAV-STD 模型,通过设计基于注意力机制的小目标检测模块 AMSTD 和具有空间感知尺度感知的 SSP 预测头,结合 N-CIoU 边界框定位损失函数,有效提升了空对空场景下无人机小目标的检测精度和定位敏感性。王灵超等^[16]提出的 MF-YOLO 模型,通过设计多路径特征融合网络结构和引入 EMA 注意力机制,结合密集注意层和挤压激励检测头,有效增强了复杂航拍场景下小目标的特征表达能力和检测精度。王迎龙等^[17]提出 BG-YOLO 算法,通过引入 Biformer 双层路由注意力机制和高分辨率特征层检测头,提升小目标检测精度。冉宁等^[18]针对遥感小目标检测中目标尺寸过小、特征提取困难的问题,提出 SBC-YOLOv8 算法,通过 Space-to-Depth 模块和 BiFPN 多尺度

特征融合,增强了小目标检测能力。

综上所述,现有的改进方法虽有一定效果,但难以兼顾精度与计算效率的双重要求。YOLOv8n 作为 YOLO 系列的轻量化模型,在模型规模、计算资源占用和检测精度等方面具有显著的优势,为无人机边缘部署提供了良好的基础架构。因此,本研究在 YOLOv8n 的基础上提出了 EGD-YOLO 算法。本研究主要工作包括:

1) 移除网络中用于大目标检测的 P5 层,增设了用于小目标检测的 P2 层。引入浅层通道扩展(shallow channel expansion, SCE)策略,扩展主干网络中浅层网络的通道数以获取更多的小目标特征信息。

2) 设计了一种自适应多层次全局特征融合架构(global hierarchical fusion, GHF),通过采用多尺度特征融合(multi-scale feature fusion, MFF)模块与加权特征融合(weighted feature fusion, WFF)模块级联的策略,实现颈部网络中跨尺度语义信息的高效传递与深度融合。

3) 采用了具有多重注意力机制的动态检测头(dynamic head, DyHead),在不同层级、空间和通道维度上自适应地调整响应权重,有效增强小目标的特征表达和定位精度。

1 YOLOv8 算法原理

YOLOv8 是一款实时目标检测框架,主要由主干网络,颈部网络和头部网络 3 部分组成。主干网络采用了改进版的 CSPDarkNet 架构,引入了 C2f 模块替代 YOLOv5 中的 C3 模块。颈部网络采用基于 PANet 改进的 PAN-FPN 结构,通过自上而下和自下而上的双向路径实现高层语义信息与底层空间细节的融合。头部网络采用解耦检测头设计,将分类和回归任务分离,分类任务使用 BCE 损失与 sigmoid 激活函数,回归任务通过 DFL 与 Ciou 联合优化,提高目标定位精度。YOLOv8 的网络结构如图 1 所示。

2 EGD-YOLO 算法设计

本文提出的目标检测模型 EGD-YOLO,旨在解决无人机航拍场景目标尺度差异大以及背景复杂等因素导致的目标检测精度不足等问题,通过优化模型的网络结构,设计高效的特征融合策略,更换具有多重注意力机制的检测头,有效提升了模型复杂场景下的目标检测性能。EGD-YOLO 网络结构如图 2 所示。

2.1 网络结构改进

在小目标检测任务中,大目标检测头的贡献相对有限。为此,本文移除了用于大目标检测的 P5 检测层,增加了用于小目标检测的 P2 检测层。改进后的网络结构如图 3 所示。

为了充分发挥 P2 检测层的性能,本文进一步对主干网络的浅层结构进行优化,提出 SCE 策略。通过扩展主干特征提取网络中浅层网络的通道数(相较于基础模型,将主干

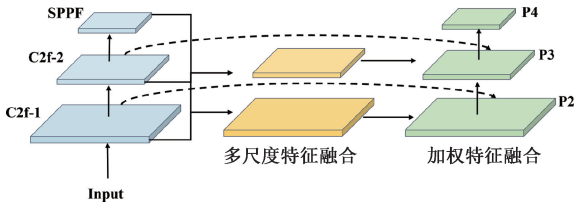


图 4 GHF 网络架构

Fig. 4 GHF network structure

过拼接和卷积操作实现多尺度特征的融合。考虑到融合后的特征常常存在响应不均衡和语义分布混乱等问题,在特征融合后引入卷积注意力(convolutional block attention module, CBAM)模块进行二次优化,通过空间注意力和通道注意力机制增强关键特征的表达能力。多尺度特征融合模块结构如图 5 所示。

CBAM 采用顺序注意力机制^[19],通过先后激活通道注意力模块和空间注意力模块来实现特征增强。通道注意

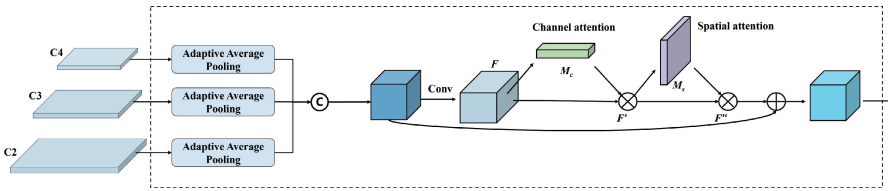


图 5 多尺度特征融合模块结构

Fig. 5 Multi-scale feature fusion module structure

力模块通过全局最大池化和全局平均池化压缩特征图的空间维度,得到两个一维的通道描述向量。接着,使用共享的多层感知机(multi-layer perceptron, MLP)对两个通道描述向量进行非线性变换。最后,采用逐元素相加操作融合两个分支的输出,并通过 Sigmoid 激活函数生成通道注意力权重。CBAM 通道注意力模块结构如图 6 所示。

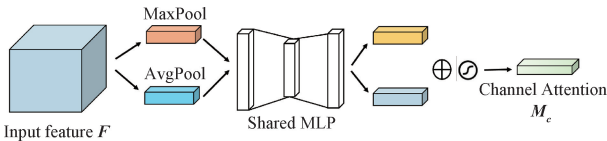


图 6 CBAM 通道注意力模块结构

Fig. 6 CBAM channel attention module structure

在空间注意力模块中,沿通道维度对特征图分别进行平均池化和最大池化得到两个二维空间描述图。随后,将这两个池化结果进行级联操作,并通过 \$7 \times 7\$ 卷积层进行空间特征提取和融合。最后经过 Sigmoid 激活函数生成空间注意力权重。CBAM 空间注意力模块结构如图 7 所示。

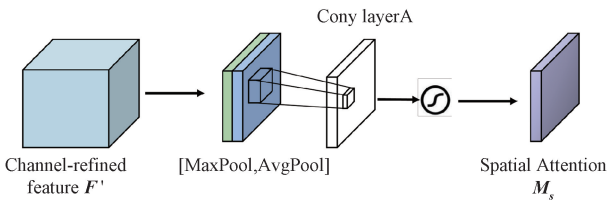


图 7 CBAM 空间注意力模块结构

Fig. 7 CBAM spatial attention module structure

$$O = \sum_i \frac{\omega_i}{\epsilon + \sum_j \omega_j} \cdot I_i \quad (1)$$

其中, \$O\$ 是输出的融合特征, \$I_i\$ 表示第 \$i\$ 个输入特征, \$\omega_i\$ 表示第 \$i\$ 个特征对应的可学习权重, \$j\$ 是所有输入特征的索引。为防止数值不稳定引入小常数 \$\epsilon\$, 通常设为 \$10^{-4}\$。此外,通过对每个 \$\omega_i\$ 应用 ReLU 激活函数来确保权重非负。

2.3 检测头改进

在无人机视角下,同一场景中可能存在尺度差异极大的目标。随着无人机飞行高度的改变,地面目标的视觉尺度会发生显著变化。此外,无人机航拍场景背景复杂,需要准确定位目标空间位置以执行目标检测、分类和定位等多个任务。传统的检测头往往难以同时应对这些挑战。本文采用集成多重注意力机制的 DyHead 检测头。

对于多维特征张量 \$\mathbf{F} \in \mathbf{R}^{L \times S \times C}\$, DyHead 分别在层级维度 \$L\$ 上引入尺度感知注意力;在空间维度 \$S\$ 引入空间感知注意力;在通道维度 \$C\$ 上引入任务感知注意力。最后,将 3 个注意力机制进行统一整合,通过顺序化的注意力函数设计,确保每个注意力模块专注于其特定的任务维度,实现了多维度注意力的解耦与协同^[21]。其计算公式如式(2)所示。

$$\mathbf{W}(\mathbf{F}) = \pi_C(\pi_S(\pi_L(\mathbf{F}) \cdot \mathbf{F}) \cdot \mathbf{F}) \cdot \mathbf{F} \quad (2)$$

其中, \$\pi(\cdot)\$ 表示注意力函数, \$\pi_C, \pi_S, \pi_L\$ 是分别应用于维度 \$C, S\$ 和 \$L\$ 的 3 种不同的注意力函数。尺度感知注意力 \$\pi_L\$ 通过学习不同尺度特征的重要性,自适应地融合不同层级的特征信息,动态调节各尺度特征的权重分配,使得网络在检测小尺度目标时能够充分利用高分辨率的浅层特征,在处理大尺度目标时充分利用语义丰富的深层特征,从而提升了跨尺度目标检测的准确性和鲁棒性。

空间感知注意力 \$\pi_S\$: 通过可变形卷积实现自适应的特征采样,并结合灵活的特征聚合机制,精准地捕捉目标

2) WFF 模块

WFF 模块通过自适应权重分配使模型自动学习不同输入特征的重要性权重^[20]。同时对权重进行归一化处理,有效避免单一权重过度主导融合过程的问题,从而提升了目标检测的整体性能。加权特征融合计算公式如式(1)所示。

在空间维度上的连续性和不规则形态。同时,这种机制能够动态调整其感受野的大小,聚焦于目标的关键区域,有效应对物体在不同视角下的形状、旋转和位置变化,提升目标识别和定位精度。

任务感知注意力 π_c : 通过引入可学习的激活阈值控制函数,动态调节特征通道的激活状态,该机制使得模型能够动态地生成或调整其目标表示方式,从而提升对特定任务的响应能力。3种注意力机制协同工作,通过融合多尺度特征、选择性关注通道信息以及动态调整感受野,共同增强无人机目标检测任务的性能。DyHead 模块结构如图8所示。

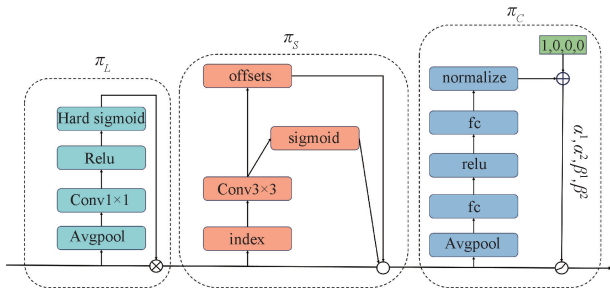


图8 DyHead 模块结构

Fig. 8 DyHead module structure

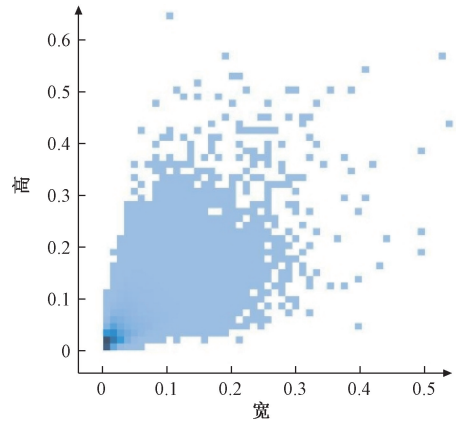


图9 目标宽高分布

Fig. 9 Target width and height distribution

显存,平台内存为 512 G,操作系统为 Ubuntu 22.04.5 系统,深度学习框架为 PyTorch2.10,CUDA 版本为 11.8,Python 版本为 3.9。实验参数设置如表1所示。

表1 模型参数设置

Table 1 Model parameter settings

参数名称	参数值
输入尺寸	640×640
初始学习率	0.04
最终学习率	0.1
迭代次数	200
优化器	SGD
动量	0.937
热身轮数	5

3 实验结果与分析

3.1 数据集介绍

1) VisDrone2019

VisDrone2019 数据集^[22]是由天津大学机器学习与数据挖掘实验室构建的大规模无人机视觉数据集。该数据集由无人机在不同高度、不同场景和地区下拍摄而成,涵盖中国14个不同的城市的各种环境,包括10个类别:行人(站立/行走)、行人(其他姿势)、汽车、货车、公共汽车、卡车、摩托车、自行车、遮阳篷三轮车和三轮车。数据集一共有8629张图片,其中6471张用于模型训练,548张用于模型验证,1610张用于模型测试。图9展示了VisDrone2019数据集的标注框尺寸分布情况,其中横轴和纵轴分别表示标注框的宽度和高度相比于图像总宽度和总高度的比值。从图9中可以观察到,大多数点密集分布在左下角区域,这表明该数据集以小目标为主导特征。这一分布特征与无人机航拍场景的实际情况高度吻合。

2) DOTA 数据集

DOTA 数据集是目标检测领域的大规模遥感数据集,共有2806张图像,涵盖15个类别,188282个标注实例,包含飞机、轮船、大型车辆和小型车辆等常见类别。本文选用DOTA数据集作为模型泛化实验数据集。

3.2 实验环境配置以及超参数设置

本实验所使用的硬件环境:CPU为Intel(R) Xeon(R) Platinum 8269CY, GPU为4张NVIDIA RTX4090 24 G

3.3 实验评估指标

为了全面评估模型性能,本研究采用以下3个关键指标:平均精度均值(mean average precision, mAP)、精确度(precision, P)和召回率(recall, R)。其中,P表示被模型预测为正例的样本中,真正例所占的比例;R表示所有真正例中,被模型成功检测到的比例;mAP由所有类别的AP求平均值得到。计算公式如式(3)~(5)所示。

$$P = \frac{TP}{TP + FP} \quad (3)$$

$$R = \frac{TP}{TP + FN} \quad (4)$$

$$mAP = \sum_{i=1}^M AP_i / M \quad (5)$$

其中,TP(true positives)表示模型预测结果为正例,真实结果也为正例;TN(true negative)表示模型预测结果为反例,真实结果也为反例;FP(false positive)模型预测结果为正例,真实结果为反例;FN(false negative)表示模型预测结果为反例,真实结果为正例。M表示类别总数,AP_i表示第i类的平均精度值。

3.4 对比实验

为了验证本研究所提模型的先进性,选取了多种目标检测模型在进行对比实验,包括经典的两阶段检测算法(Faster R-CNN),单阶段检测算法(YOLOv8s),以及近年来针对无人机场景优化的较为先进的检测算法(Drone-YOLO、Sparse DETR、BDP-YOLOs 等)。所有算法的训练和测试均在 VisDrone2019 数据集上进行。实验结果如表 2 所示。与经典检测算法 Faster R-CNN 相比,EGD-YOLO 在检测精度上具有显著优势,mAP@0.5 提升了 9.2%。基于 Transformer 架构的 Sparse DETR 虽然在检测精度上有所提升,但其计算复杂度约为 EGD-YOLO 的 5.3 倍,难以满足实时检测需求。Vectorized IOU-

YOLOv5 和 YOLO-DA 表现与 EGD-YOLO 接近,但 Vectorized IOU-YOLOv5 参数量约为 EGD-YOLO 的 8.4 倍,YOLO-DA 的参数量和计算量分别约为 EGD-YOLO 的 28 倍和 6.4 倍,模型复杂度过高。对于同为轻量化小目标检测的 YOLOv7-tiny、YOLOv8s、BDP-YOLOs 和 LRDS-YOLO,YOLOv7-tiny 虽然计算开销相对较低,但其参数量仍为 EGD-YOLO 的 2.6 倍,且检测精度明显不足。YOLOv8s、BDP-YOLOs 和 LRDS-YOLO 在模型规模上有所优化,但计算复杂度和参数量仍高于 EGD-YOLO,同时检测精度仍有差距。综合实验结果表明,EGD-YOLO 在保持优异检测精度的同时,实现了模型复杂度的最优控制。

表 2 在 VisDrone2019 数据集上不同算法对比实验

Table 2 Comparative experiments of different algorithms on the VisDrone2019 dataset

模型	mAP@0.5/%	mAP@0.5:0.95/%	Params/M	GFLOPs
Faster R-CNN	35.8	19.7	—	—
Sparse DETR	42.5	27.3	—	121.0
Vectorized IOU-YOLOv5 ^[23]	44.6	26.6	19.30	—
UN-YOLOv5s ^[24]	40.5	22.5	—	37.4
YOLOv7-tiny	35.0	18.5	6.04	13.3
YOLOv8s ^[25]	43.8	—	10.90	26.9
Drone-YOLO (large) ^[26]	40.7	—	76.20	—
YOLO-DA ^[27]	44.7	25.3	63.80	147.9
EdgeYOLO ^[28]	44.8	—	40.50	109.1
BDP-YOLOs ^[29]	45.0	27.4	5.80	36.7
LRDS-YOLO ^[30]	43.6	26.6	4.07	23.7
EGD-YOLO	45.0	27.9	2.30	23.0

3.5 消融实验结果

为了验证本研究所提改进方法的有效性,在 VisDrone2019 数据集上进行了消融实验。实验以 YOLOv8n 为基线模型,逐一加入或者替换本文所提到的改进方法。消融实验结果如表 3 所示。从消融实验结果可以看出,加入 P2 检测层后,小目标检测效果显著提升。模型的 mAP@0.5 提升

了 4.1%,mAP@0.5:0.95 提升了 3.0%。表明 P2 层能够提取更丰富的空间位置信息和细节特征,有效提升小目标检测的性能。删除 P5 检测层之后,参数量显著减少,有效降低了计算与存储开销,整体检测性能仅出现了轻微下降。表明 P5 层在小目标检测任务中的贡献相对有限,移除 P5 检测层能使模型在保持较高精度的同时,大幅提升模型的轻量化程度。

表 3 消融实验结果

Table 3 Ablation experiment results

方法	P2	Delete P5	SCE	GHF	DyHead	P/%	R/%	mAP@0.5/%	mAP@0.5:0.95/%	Params/M	GFLOPs
YOLOv8n						44.5	32.8	33.0	19.3	3.0	8.9
A	✓					48.7	35.6	37.1	22.3	2.9	12.4
B	✓	✓				47.2	35.7	36.6	21.9	1.0	10.6
C	✓	✓	✓			50.3	37.1	38.7	23.4	1.0	12.8
D	✓	✓		✓		48.8	36.2	37.9	22.8	1.1	12.5
E	✓	✓			✓	52.4	39.9	41.7	25.3	2.0	17.4
F	✓	✓	✓	✓		51.6	40.4	42.0	25.4	1.2	16.2
G	✓	✓	✓	✓	✓	54.5	42.0	45.0	27.9	2.3	23.0

在此基础上,引入 SCE 策略后,模型的各项指标均有显著提升,其中 $mAP@0.5$ 和 $mAP@0.5:0.95$ 提升了 2.1% 和 1.5%。引入 GHF 策略后,模型的 $mAP@0.5$ 和 $mAP@0.5:0.95$ 分别提升了 1.3% 和 0.5%。同时引入 SCE 和 GHF 之后,模型的 $mAP@0.5$ 和 $mAP@0.5:0.95$ 分别提升了 5.4% 和 3.5%,表现出明显的协同增强效应。表明 SCE 策略可以提供更丰富的细节特征基础,而全局特征融合机制则将这些增强的细节特征与高层语义信息进行深度整合,使得模型既具备精细的局部感知能力,又拥有全局的语义理解能力。

替换 DyHead 检测头之后,模型的 $mAP@0.5$ 和 mAP

@0.5:0.95 分别提升了 3.0% 和 2.5%。表明具有多重注意力机制的检测头可以使模型更好地捕捉小目标细节,增强了小目标的检测准确性和鲁棒性。

3.6 泛化实验

为了进一步验证模型的泛化性,本文在 DOTA 数据集进行了对比实验,实验结果如表 4 所示。EGD-YOLO 相比于 YOLOv8n 和 YOLOv8s 在 $mAP@0.5$ 上分别提升了 5.7% 和 2.4%,相较于最新的 YOLOv11n 和 YOLOv11s 分别提升了 6.9% 和 2.6%,表明所提算法在不同条件、不同场景下均有较强的泛化性。

表 4 在 DOTA 数据集上不同算法对比实验

Table 4 Comparative experiments of different algorithms on the DOTA dataset

模型	$mAP@0.5/\%$	$mAP@0.5:0.95/\%$	Params/M	GFLOPs
YOLOv8n	40.9	24.6	3.0	8.9
YOLOv11n	39.7	24.6	2.6	6.6
YOLOv8s	44.2	27.2	11.2	28.8
YOLOv11s	44.0	27.7	9.5	21.7
EGD-YOLO	46.6	28.6	2.3	23.0

3.7 可视化实验

为了直观展示改进算法的检测效果,从 VisDrone2019 测试集中选取了几个典型场景下的检测结果进行可视化对比,如图 10~12 所示。红色矩形虚线框矩形虚线框为

EGD-YOLO 提升较为明显的区域。从图像中可以看到本研究方法在各种场景下均展现出更强的检测能力,证明了本研究改进方法的有效性 与鲁棒性,能够满足实际任务需求。

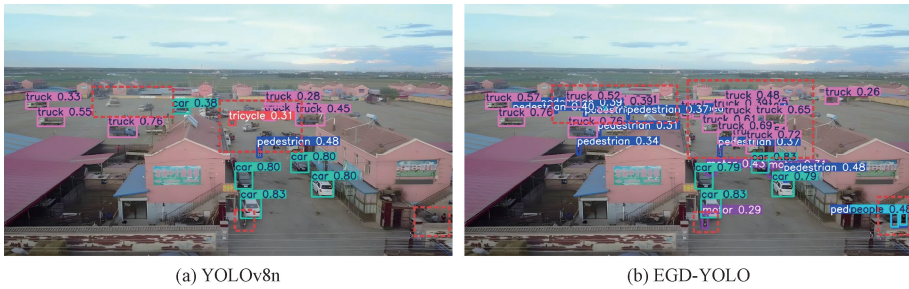


图 10 密集场景下检测结果对比

Fig. 10 Comparison of detection results in dense scenes

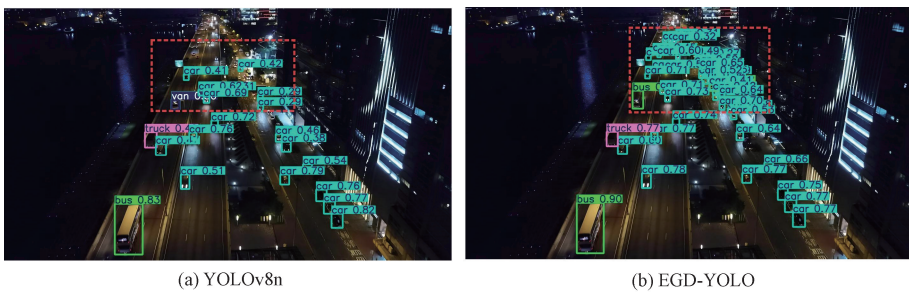


图 11 夜间环境下检测结果对比

Fig. 11 Comparison of detection results in nighttime environments

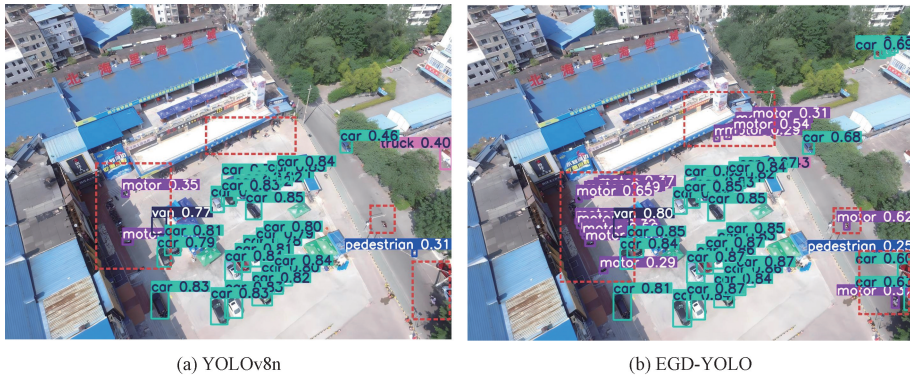


图 12 强光环境下检测结果对比
Fig. 12 Comparison of detection results in strong light environments

4 结 论

针对无人机航拍图像目标检测任务中的小目标检测困难、背景复杂、硬件资源受限,以及大多数现有检测算法检测精度不足、难以满足实际应用需求等关键问题,本文在 YOLOv8n 的基础上提出了改进的目标检测模型 EGD-YOLO。在网络结构方面移除了用于大目标检测的 P5 层,增加了用于小目标检测的 P2 层,使模型更适应无人机航拍场景中小目标占主导的特点。为获取更丰富的小目标信息,引入了 SCE 策略,提升模型对细粒度特征的特征能力。提出 GHF 架构替换原有的 PAN-FPN 结构,通过 MFF 与 WFF 模块级联的融合策略,实现多层次特征的全局整合。引入具有尺度感知注意力、空间感知注意力和任务感知注意力机制的 DyHead 检测头,进一步优化模型的检测性能。经过系列对比实验和消融实验验证,改进后的模型 mAP@0.5 比基线模型提升了 12.0%,并优于多个经典模型,相比于 EdgeYOLO 和 Vectorized IOU-YOLOv5 等算法在提升精度的同时具有良好的计算优势。同时,在 DOTA 数据集上的对比实验进一步验证其泛化性,mAP@0.5 比基线模型提升了 5.7%。未来可以继续研究如何在提高模型检测精度的同时,进一步减少模型的参数量和计算复杂度,以更好地适应边缘设备的部署需求。

参考文献

[1] ZOU ZH X, CHEN K Y, SHI ZH W, et al. Object detection in 20 years: A survey[J]. IEEE, 2023, 111(3): 257-276.

[2] 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, 2016: 779-788.

[3] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 7263-7271.

[4] LIU W, ANGUILOV D, ERHAN D, et al. SSD:

Single shot multibox detector [C]. European Conference on Computer Vision, 2016: 21-37.

[5] AMJOU D A B, AMROUCH M. Object detection using deep learning, CNNs and vision transformers: A review[J]. IEEE Access, 2023, 11: 35479-35516.

[6] 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.

[7] YE T, QIN W Y, ZHAO Z Y, et al. Real-time object detection network in UAV-vision based on CNN and transformer[J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1-13.

[8] DING L H, XU X, CAO Y, et al. Detection and tracking of infrared small target by jointly using SSD and pipeline filter[J]. Digital Signal Processing, 2021, 110: 102949.

[9] KANG S H, PARK J S. Aligned matching: Improving small object detection in SSD[J]. Sensors, 2023, 23(5): 2589.

[10] LI B Y, XIAO CH, WANG L G, et al. Dense nested attention network for infrared small target detection[J]. IEEE Transactions on Image Processing, 2022, 32: 1745-1758.

[11] WEI X L, YIN L, ZHANG L L, et al. DV-DETR: Improved UAV aerial small target detection algorithm based on RT-DETR[J]. Sensors, 2024, 24(22): 7376.

[12] KONG Y N, SHANG X F, JIA SH J. Drone-DETR: Efficient small object detection for remote sensing image using enhanced RT-DETR model[J]. Sensors, 2024, 24(17): 5496.

[13] JU M R, LUO J N, ZHANG P P, et al. A simple and efficient network for small target detection[J]. IEEE Access, 2019, 7: 85771-85781.

- [14] SHANG J CH, WANG J S, LIU SH B, et al. Small target detection algorithm for UAV aerial photography based on improved YOLOv5s[J]. *Electronics*, 2023, 12(11): 2434.
- [15] ZUO G B, ZHOU K, WANG Q. UAV-to-UAV small target detection method based on deep learning in complex scenes[J]. *IEEE Sensors Journal*, 2025, 25(2): 3806-3820.
- [16] 王灵超, 沈学利, 艾强, 等. 多路径特征融合的YOLOv8 航拍图像检测算法[J]. *电子测量技术*, 2025, 48(17): 160-168.
WANG L CH, SHEN X L, AI Q, et al. YOLOv8 aerial image detection algorithm with multi-path feature fusion[J]. *Electronic Measurement Technology*, 2025, 48(17): 160-168.
- [17] 王迎龙, 孙备, 丁冰, 等. BG-YOLO: 复杂大视场下低慢小无人机目标检测方法[J]. *仪器仪表学报*, 2025, 46(2): 255-266.
WANG Y L, SUN B, DING B, et al. BG-YOLO: A low-altitude slow-moving small UAV targets detection method in complex large field of view[J]. *Chinese Journal of Scientific Instrument*, 2025, 46(2): 255-266.
- [18] 冉宁, 施高朗, 张少康, 等. 基于YOLOv8的遥感小目标检测算法[J]. *电子测量与仪器学报*, 2025, 39(5): 197-207.
RAN N, SHI G L, ZHANG SH K, et al. Remote sensing small target detection algorithm based on YOLOv8[J]. *Journal of Electronic Measurement and Instrumentation*, 2025, 39(5): 197-207.
- [19] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]. *European Conference on Computer Vision(ECCV)*, 2018: 3-19.
- [20] TAN M X, PANG R M, LE Q V. EfficientDet: Scalable and efficient object detection[C]. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020: 10781-10790.
- [21] DAI X Y, CHEN Y P, XIAO B, et al. Dynamic head: Unifying object detection heads with attentions [C]. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021: 7373-7382.
- [22] DU D W, ZHU P F, WEN L Y, et al. VisDrone-DET2019: The vision meets drone object detection in image challenge results[C]. *IEEE/CVF International Conference on Computer Vision Workshops*, 2019: 213-226.
- [23] LU SH, LU H Y, DONG J, et al. Object detection for UAV aerial scenarios based on vectorized IOU[J]. *Sensors*, 2023, 23(6): 3061.
- [24] GUO J M, LIU X CH, BI L Y, et al. UN-YOLOv5s: A uav-based aerial photography detection algorithm[J]. *Sensors*, 2023, 23(13): 5907.
- [25] 党宏社, 高宛蓉, 孟饶辰, 等. 融合多尺度特征的轻量级航拍目标检测算法[J]. *电子测量技术*, 2024, 47(16): 92-99.
DANG H SH, GAO W R, MENG R CH, et al. Target detect algorithm of lightweight in aerial images based on multi-scale feature fusion [J]. *Electronic Measurement Technology*, 2024, 47(16): 92-99.
- [26] ZHANG ZH X. Drone-YOLO: An efficient neural network method for target detection in drone images[J]. *Drones*, 2023, 7(8): 526.
- [27] 苏雨蕾, 黄丹丹, 刘智, 等. 无人机对地小目标检测方法研究[J]. *电子测量与仪器学报*, 2024, 38(9): 144-154.
SU Y L, HUANG D D, LIU ZH, et al. Research on detection methods of small targets on ground by UAV[J]. *Journal of Electronic Measurement and Instrumentation*, 2024, 38(9): 144-154.
- [28] LIU SH H, ZHA J L, SUN J, et al. EdgeYOLO: An edge-real-time object detector[C]. *2023 42nd Chinese Control Conference(CCC)*, 2023: 7507-7512.
- [29] CHAO M X, PENG C, YUN L J, et al. A lightweight small object detection model for UAV images based on deep semantic integration [J]. *Scientific Reports*, 2025, 15(1): 31888.
- [30] HAN Y Q, WANG CH CH, LUO H, et al. LRDS-YOLO enhances small object detection in UAV aerial images with a lightweight and efficient design[J]. *Scientific Reports*, 2025, 15(1): 22627.

作者简介

崔博斌, 硕士研究生, 主要研究方向为基于无人机平台的目标识别、目标追踪。

E-mail: 1447191515@qq.com

易军凯(通信作者), 教授, 博士生导师, 主要研究方向为大数据分析、数据安全检测、网络空间安全、人工智能安全。

E-mail: yijk@bistu.edu.cn

谭玲玲, 副教授, 硕士生导师, 主要研究方向为隐藏信息深度挖掘、网络感知与动态控制。

E-mail: tanlingling@bistu.edu.cn