

YOLO-MAF: 多尺度特征融合的轻量级 小目标检测算法^{*}

胡 闽 南¹ 赵 鸣²

(1. 南京信息工程大学计算机学院 南京 210044; 2. 无锡学院网络空间安全学院 无锡 214105)

摘 要: 在无人机航拍交通监控中,小目标检测面临特征表达不充分、多尺度融合效率低等挑战。为此,提出改进的 YOLO-MAF 目标检测算法。首先,设计多尺度边缘增强特征提取模块(MSEE),通过自适应多尺度感受野机制增强微小目标边缘信息捕获;其次,构建 SEGE 双向采样增强模块,结合软最近邻插值(SNI)和增强型分组卷积策略(GSConvE),优化特征融合效率;最后,在建立一层小目标检测层的同时结合构造的多尺度自适应空间融合检测头(MASF-Head),采用双重注意力机制动态调节特征融合权重。实验结果表明,在 VisDrone2019 数据集上,YOLO-MAF 的 mAP@0.5 和 mAP@0.5:0.95 分别达到 45.6% 和 29.4%,相比基线方法提升 7.3% 和 6.4%,参数量减少 50%。所提方法有效提升了无人机航拍场景下的小目标检测性能。

关键词: 小目标检测;边缘增强;多尺度特征融合

中图分类号: TP391.4;TN919.8 **文献标识码:** A **国家标准学科分类代码:** 520.60

YOLO-MAF: A lightweight small object detection algorithm based on multi-scale feature fusion

Hu Minnan¹ Zhao Ming²

(1. School of Computer Science, Nanjing University of Information Science and Technology, Nanjing 210044, China;

2. School of Cyberspace Security, Wuxi University, Wuxi 214105, China)

Abstract: In UAV traffic surveillance, small target detection faces challenges such as insufficient feature representation and low efficiency in multi-scale fusion. To address this, this study proposes YOLO-MAF, a lightweight detector. First, the multi-scale edge enhancement (MSEE) module strengthens edge information via an adaptive multi-scale receptive field. Second, the SEGE module combines soft nearest-neighbor interpolation (SNI) and enhanced group convolution (GSConvE) to improve cross-level alignment and fusion. Finally, the MASF-Head adopts dual attention to learn spatial-channel weights for adaptive multi-scale fusion. On VisDrone2019, YOLO-MAF achieves 45.6% mAP@0.5 and 29.4% mAP@0.5:0.95, improving the baseline by 7.3% and 6.4% with 50% fewer parameters, demonstrating effective small-object detection under UAV scenarios.

Keywords: small object detection; edge enhancement; multi-scale feature fusion

0 引 言

随着城市化进程的加快和智能交通系统的发展^[1],基于视觉的目标检测技术在交通监控中发挥着重要作用^[2]。无人机凭借机动性强、覆盖广、部署成本低等优势,在交通流量监测、违章识别和事故管理等方面展现出广阔应用前景^[3-4]。然而,无人机航拍中的目标检测仍面临目标尺寸

小、光照复杂、背景干扰强等挑战,导致检测性能下降^[5]。同时,边缘设备算力受限,小目标的细粒度特征易被削弱,跨尺度对齐误差也会放大检测不稳定性,使得“高精度”与“实时性”难以兼顾。

近年来,深度学习推动了无人机航拍目标检测的显著进展^[6]。例如,基于 YOLOv8 改进的 YOLOv8-ERD 通过引入多尺度特征重构与浅层纹理增强模块,有效减少了小

收稿日期:2025-08-07

^{*} 基金项目:无锡学院 2025 年高层次人才科研启动基金(2025r023)、2023 中国高校产学研创新基金-新一代信息技术创新项目(2023IT072)资助

目标在特征传递过程中的信息损失^[7];改进 YOLOv11 的航拍图像检测算法采用跨层注意力与特征融合策略,以增强模型对细粒度区域的响应能力,提高了小目标的敏感性^[8];基于 CS-YOLOv5s 的无人机航拍小目标检测方法通过通道选择和轻量化特征融合结构强化高低层特征之间的交互,从而提升特征学习效率^[9]。

为提高计算效率,部分研究在 YOLOv8 结构基础上增设轻量化检测层并采用条件卷积以增强特征表达能力^[10];一些轻量化多尺度融合的改进模型在保持网络轻量化的同时实现了较好的检测精度^[11];而 Tiny-YOLO 系列轻量化模型也进一步促进了小目标检测技术在资源受限场景中的实用化^[12]。

此外,PARE-YOLO 通过重构颈部网络并融合多尺度注意力结构以增强跨尺度特征传递能力^[13];TA-YOLO 采用变换注意力机制改善复杂背景下的特征匹配,从而提升目标识别的稳定性^[14];AMFE-YOLO 在无人机图像场景中展现出较高的检测准确性^[15]。另外,一些方法尝试在精度与速度之间取得平衡:如 DAU-YOLO 利用动态注意力上采样结构增强特征表达,但在复杂场景中仍存在适应性不足的问题^[16];SEMA-YOLO 借助浅层增强模块与多尺度特征融合结构提升检测效率,但对极小目标仍较为敏感^[17];LS-YOLO 则结合尺度自适应交并比损失与轻量注意机制,在检测性能与推理速度之间取得了较好的折中^[18]。

针对上述问题,本研究旨在构建一种兼顾检测精度与推理效率的轻量化目标检测框架,重点强化小目标的边缘特征建模与多尺度语义融合能力,以提升无人机航拍中小目标检测的准确性与稳定性。

为此,本研究提出一种多尺度自适应融合 YOLO 算法(multi-scale adaptive fusion YOLO, YOLO-MAF)。该方法在轻量化架构下引入多尺度边缘增强与跨尺度语义融合机制,提升小目标的高频特征表达与多尺度一致性,从而在复杂场景中实现高精度与实时性的平衡。本研究主要贡献包括:

1) 提出多尺度边缘增强模块(multi-scale edge enhancement, MSEE),通过多尺度感受野与边缘增强机制强化微小目标的边缘信息;

2) 设计双向采样增强模块(sampling enhancement with GSConv enhancement, SEGE),结合软最近邻插值(soft nearest-neighbor interpolation, SNI)与增强型组卷积(GSConv enhancement, GSConvE)提高特征融合效率并减少计算冗余;

3) 在网络结构中增加一层小目标检测层,其次,将骨干网络最后 3 层通道数调整至 512 并构建多尺度自适应空间融合检测头(multi-scale adaptive sparse fusion head, MASF-Head),用动态权重与通道注意力机制实现多尺度特征自适应融合,显著提升检测鲁棒性与模型紧凑性。

1 改进的 YOLOv11 算法

YOLOv11 是一种高效的目标检测算法,通过在主干网络中引入 C3k2 模块实现跨层连接与解耦头实现高效检测^[19-20],但在复杂场景下对小目标边缘与跨尺度对齐仍不足。

在 YOLOv11 基础上构建改进的检测框架,主要包括:首先,在骨干网络中引入 MSEE 模块替换 C3k2 模块,提升对微小目标边缘特征的感知能力;其次,在颈部网络中集成 SNI 软最近邻插值模块和 GSConvE 增强型组卷积模块,优化特征融合效率;最后,再增加 1 层小目标检测层,修改骨干网络的通道数为 512 同时设计 MASF 多尺度自适应特征融合检测头模块,通过动态权重学习机制增强对不同尺度目标的检测鲁棒性。改进的 YOLOv11 网络整体结构如图 1 所示。

为更清晰地展示本研究改进策略的独特性与创新点,表 1 对 YOLO-MAF 与其他典型 YOLOv11 改进算法的网络结构与设计目标进行了对比。可以看出,所提方法不仅在结构设计上更具系统性,还在特征对齐与多尺度融合方面提出了全新的轻量化优化思路。

1.1 MSEE 模块

为提高检测网络对小目标边缘信息的建模能力,提出多尺度边缘信息增强模块 MSEE。该模块由多尺度感受野提取分支、边缘响应增强单元及特征融合结构构成。

MSEE 模块通过自适应多尺度感受野机制增强对微小目标边缘信息的捕获能力。输入特征图通过不同窗口大小($3 \times 3, 6 \times 6, 9 \times 9, 12 \times 12$)的自适应平均池化操作实现多尺度划分,每一路径分别通过 1×1 和 3×3 卷积操作提取局部特征:

$$F_{\text{conv}}^s = \text{Conv}_{3 \times 3}(\text{Conv}_{1 \times 1}(F_{\text{pool}}^s)) \quad (1)$$

并分别接入边缘增强单元(edge enhancer, EE)以强化结构突变区域的响应。结构如图 2 所示。

该模块通过高频残差信号捕获策略获取图像细节区域的响应,即将中间特征 F_{conv}^s 与其通过平均池化形成的平滑版本 $\text{AvgPool}(F_{\text{conv}}^s)$ 进行差分处理并施以卷积非线性映射,具体表达为:

$$F_{\text{edge}}^s = F_{\text{conv}}^s + \text{Conv}(F_{\text{conv}}^s - \text{AvgPool}(F_{\text{conv}}^s)) \quad (2)$$

所有分支的边缘增强特征经上采样恢复至统一分辨率后进行拼接整合,通过卷积融合输出最终特征图:

$$F_{\text{out}} = \text{Conv}(\text{Concat}(F_{\text{edge}}^1, \dots, F_{\text{edge}}^n)) \quad (3)$$

MSEE 模块通过多尺度上下文语义与高频边缘导向机制的协同建模,在保证轻量计算复杂度的同时显著提升模型对微弱轮廓、小尺度结构的辨识能力,特别适用于航拍场景中的边界目标检测。MSEE 模块结构如图 3 所示。

1.2 SEGE 双向采样增强模块

为提升目标检测网络对多尺度特征的融合与对齐能

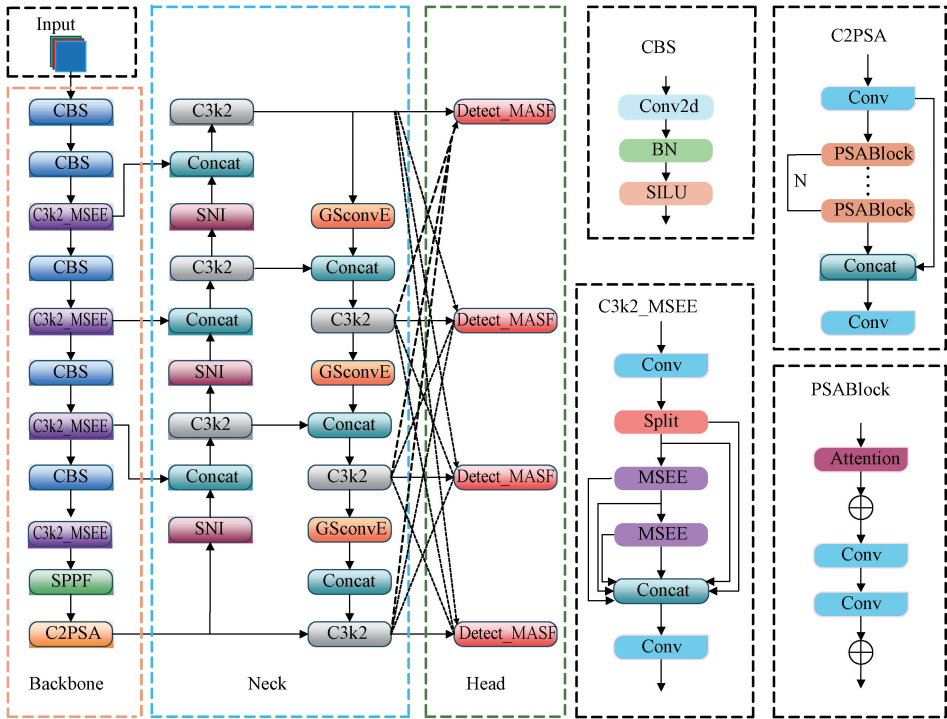


图 1 改进的 YOLOv11 网络结构
Fig.1 Improved YOLOv11 network structure

表 1 YOLO-MAF 与典型 YOLOv11 改进模型的结构差异与创新对比

Table 1 Structural differences and innovations of YOLO-MAF vs representative YOLOv11-based models

模型	改进目标	主要改进内容	特征融合策略	参数量/M	创新特征
S-YOLOv11	小目标检测	改进路径聚合网络、增强特征金字塔网络	多层特征级联融合	5.2	深度增强, 边缘优化不足
SOD-YOLO	轻量化	深度可分卷积	简化型融合路径	6.8	减参明显但小目标召回不足
TA-YOLO	区分特征/抑制背景	多维变换注意力	层间语义融合	5.6	结构新颖, 推理偏慢
YOLO-MAF (本文)	精度+效率	MSEE+SEGE+MASF 协同	自适应多尺度融合(四尺度加权+通道注意)	4.7	兼顾边缘细节与轻量化, 跨尺度鲁棒性最优

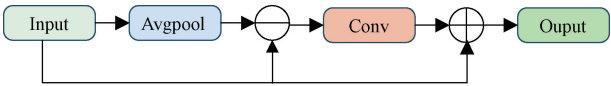


图 2 边缘增强模块结构
Fig.2 Edge enhancer module structure

力,设计了 SEGE 双向采样增强模块。该模块结合软最近邻插值技术(SNI 模块)和增强型分组卷积策略(GSConvE 组件),优化不同层级特征的对齐与融合效率。

传统最近邻插值方法在处理小目标特征时容易产生定位偏差和计算冗余。针对这一问题,SNI 模块引入自适应软化系数 α 来动态调节插值权重,替代传统的硬性插值操作。

$$Y = \alpha \cdot f(X) + (1 - \alpha) \cdot g(Y_{prev}) \quad (4)$$

其中, α 作为动态调节因子,其计算公式为:

$$\alpha = \beta \times \frac{Resolution_X}{Resolution_Y} \times \Psi(F_X, F_Y) \quad (5)$$

其中, β 为可学习的全局调节参数, $\Psi(F_X, F_Y)$ 为特征相似性函数,定义为:

$$\Psi(F_X, F_Y) = \exp(-\gamma \times \|F_X - \text{Interpolate}(F_Y)\|^2) \quad (6)$$

其中, γ 为控制特征差异敏感度的学习参数, F_X 和 F_Y 分别表示高层和低层特征图的特征表示。

这种软化权重分配机制能够更好地保持原始特征的完整性。 α 因子通过特征相似性函数 $\Psi(F_X, F_Y)$ 融入空间相关性信息,具体实现为计算高层特征 F_X 与插值后低层特征的 L2 距离,并通过指数函数进行非线性映射,使得特

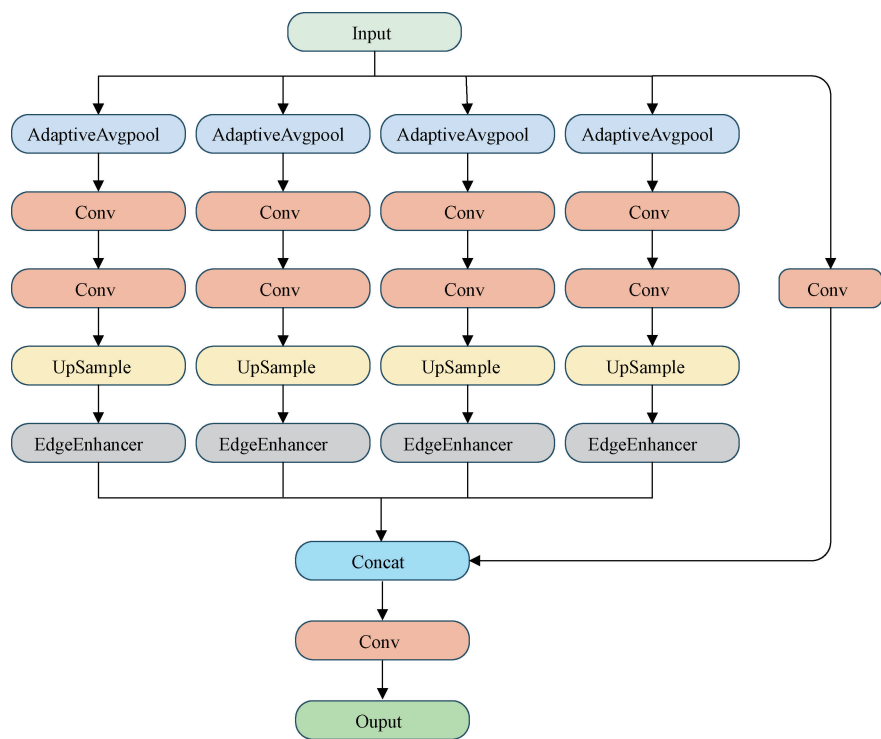


图 3 MSEE 模块结构

Fig. 3 MSEE module structure

征差异较大的区域获得较小的融合权重,从而保持局部特征的空间一致性,在处理小目标和边界模糊区域时表现出显著优势。

GSCovE 组件主要包含线性融合、通道连接、特征重排和非线性激活 4 个核心处理环节。线性融合环节将多尺度特征流进行通道维度的拼接操作,这种通道混合机制

能够打破原有通道分组限制,促进不同特征组之间的信息交互,增强特征表达的多样性。在非线性激活环节中,采用高斯误差线性单元(Gaussian error linear unit,GELU)激活函数,其平滑特性和概率解释能够提供更加稳定的梯度传播,特别适合用于轻量级网络的优化训练。GSCovE 具体的结构图如图 4 所示。

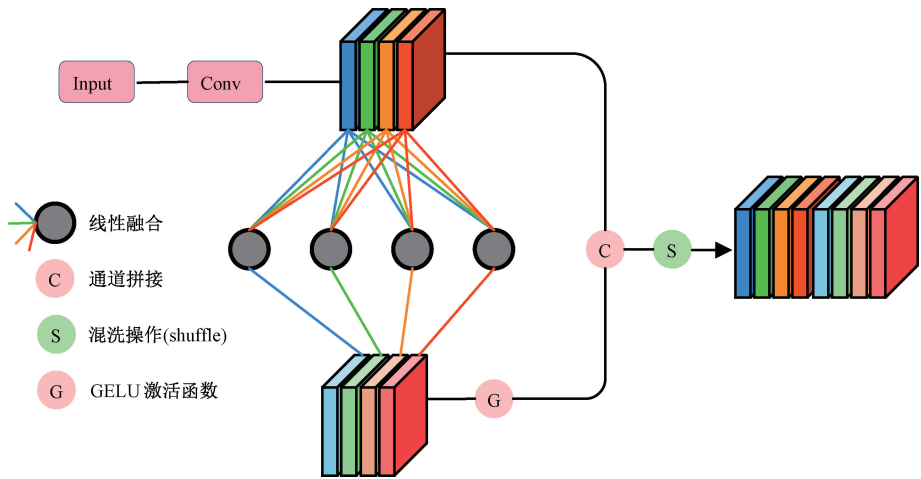


图 4 GSCovE 模块结构

Fig. 4 GSCovE module structure

通过联合优化,SEGE 模块在保持轻量化的同时显著提升多尺度特征的交互性与表达能力,并减少计算冗余,适合复杂背景下的小目标检测任务。

1.3 MASF 检测头模块

在添加了一层小目标检测层后,为增强多尺度特征的稳定性与检测鲁棒性,在 YOLOv11 检测头基础上提出

MASF-Head。具体的结构如图5所示。

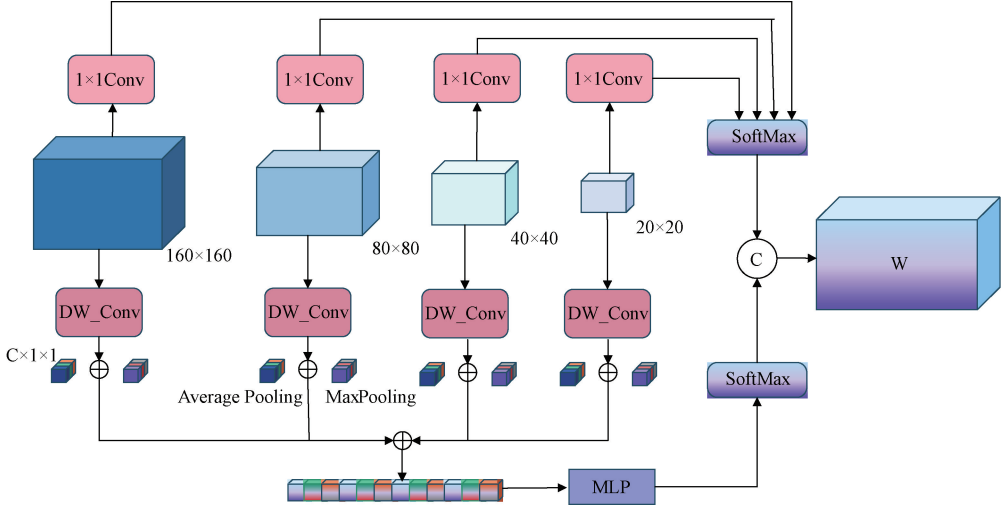


图5 MASF模块结构

Fig. 5 MASF module structure

feature fusion, ASFF)的思想,能够动态学习来自不同尺度特征层的融合权重,将各个尺度的特征进行尺寸对齐后进行自适应整合。与传统方法不同,采用4个尺度输入,并引入4个学习权重参数 $\alpha_s^l, \beta_s^l, \gamma_s^l$ 和 δ_s^l 分别用于控制各尺度特征在目标层级中的融合比例。

将 $x^{n \rightarrow l}$ 定义为缩放的特征向量, n, l 分别表示源层和目标层的索引,通过卷积运算使其大小均匀。为保证原始特征在融合过程中占据较高权重,使用softmax函数(softmax)对4个控制参数 $\alpha_s^l, \beta_s^l, \gamma_s^l, \delta_s^l$ 进行归一化处理,从而得到最终的融合权重。各尺度特征经卷积统一尺寸后进行加权融合,实现多尺度信息的有效整合并提升检测精度:

$$\alpha_s^l = \frac{e^{\lambda_a^l}}{e^{\lambda_a^l} + e^{\lambda_\beta^l} + e^{\lambda_\gamma^l} + e^{\lambda_\delta^l}} \quad (7)$$

$$\beta_s^l = \frac{e^{\lambda_\beta^l}}{e^{\lambda_a^l} + e^{\lambda_\beta^l} + e^{\lambda_\gamma^l} + e^{\lambda_\delta^l}} \quad (8)$$

$$\gamma_s^l = \frac{e^{\lambda_\gamma^l}}{e^{\lambda_a^l} + e^{\lambda_\beta^l} + e^{\lambda_\gamma^l} + e^{\lambda_\delta^l}} \quad (9)$$

$$\delta_s^l = \frac{e^{\lambda_\delta^l}}{e^{\lambda_a^l} + e^{\lambda_\beta^l} + e^{\lambda_\gamma^l} + e^{\lambda_\delta^l}} \quad (10)$$

确保约束条件:

$$\alpha_s^l + \beta_s^l + \gamma_s^l + \delta_s^l = 1 \quad (11)$$

为了调整信道的重要性和抑制无关信息,使用信道注意力图来获得另一组权重参数,从而提高模型的性能。具体操作为:

$$s(i) = \text{DW_Conv}(\text{MaxPool}(f(i)) + \text{AvgPool}(f(i))) \quad (12)$$

该方法借鉴了自适应空间特征融合(adaptively spatial

fusion, ASFF)的思想,能够动态学习来自不同尺度特征层的融合权重,将各个尺度的特征进行尺寸对齐后进行自适应整合。与传统方法不同,采用4个尺度输入,并引入4个学习权重参数 $\alpha_s^l, \beta_s^l, \gamma_s^l$ 和 δ_s^l 分别用于控制各尺度特征在目标层级中的融合比例。

$$c(i) = \text{MLP}(s(i)) \quad (13)$$

再通过对各尺度特征 $c(2), c(3), c(4), c(5)$ 施加softmax变换,获得通道权重系数 $\alpha_c^l, \beta_c^l, \gamma_c^l$ 和 δ_c^l 。最终的加权参数 α^l 的计算公式为:

$$\alpha^l = w_1 \alpha_s^l + w_2 \alpha_c^l \quad (14)$$

其中, w_1 和 w_2 是可学习的参数。对于 β^l, γ^l 和 δ^l 的计算方法类似。使用这些学习到的加权融合参数,得到一个新的多尺度特征映射张量 y^l :

$$y^l = \alpha^l x^{P2 \rightarrow l} + \beta^l x^{P3 \rightarrow l} + \gamma^l x^{P4 \rightarrow l} + \delta^l x^{P5 \rightarrow l} \quad (15)$$

这种加权融合方法不仅有效保留了不同尺度特征的优势,还增强了模型提取通道信息的能力,从而提升了整体检测性能。MASF具体的结构图如图5所示。

2 实验验证与分析

2.1 实验数据集

实验采用VisDrone2019数据集^[21]验证算法有效性。该数据集由天津大学联合AISKYEYE团队构建,涵盖中国14个城市的航拍图像,具有丰富的场景多样性。

VisDrone2019包含10类目标(行人、汽车、卡车等)。数据集共8599幅静态图像,按官方标准划分为训练集6471张、验证集548张和测试集1580张。

2.2 实验环境配置

实验所使用的相关硬件、软件及实验环境详细信息及训练参数设置如表2、3所示。其他策略和超参数遵循基

准模型。

表 2 实验环境配置

Table 2 Experimental environment configuration	
配置项	详细信息
操作系统	Windows 11
CPU	Intel ® Core™ i9-14900 HX
GPU	RTX 4060Ti (16 GB)
Python 版本	3.10.14
CUDA	12.1
深度学习框架	PyTorch 2.2.2

表 3 训练参数设置

Table 3 Training parameter settings	
参数名称	具体信息
输入图像尺寸	640×640
批处理大小	16
最大训练轮数	200
优化器	SGD
初始学习率	0.001

2.3 实验评估指标

为评估改进算法的检测性能,采用目标检测领域的标准评价指标:精确率(precision,P)、召回率(recall,R)、平均精度(average precision,AP)、平均精度均值(mean average precision,mAP)。

精确率表示预测正样本中实际正样本的比例,计算公式为:

$$P = \frac{TP}{TP + FP}$$
(16)

召回率表示实际正样本中被正确识别的比例,计算公式为:

$$R = \frac{TP}{TP + FN}$$
(17)

其中,TP为真正例(正确检测目标数),FP为假正例(误检目标数),FN为假负例(漏检目标数)。平均精度(AP)为不同召回率下精确率的平均值,综合反映各阈值下的检测性能:

$$AP = \int_0^1 P(R) dR$$
(18)

平均精度均值(mAP)为所有类别 AP 的算术平均:
$$mAP = \frac{1}{C} \sum_{i=1}^C AP_i$$
(19)

其中,C表示目标类别总数。

评估过程中采用交并比(intersection over union,IoU)在阈值 0.5 的 mAP(记为 mAP@0.5)和 0.5-0.95 的 mAP(记为 mAP@0.5:0.95)进行评估,后者更严格评估定位精度。

此外,采用参数量(parameters,Param)、计算复杂度(giga floating-point operations,GFLOPs)和帧率(frames per second,FPS)衡量模型的轻量化与实时性。

2.4 消融实验

为验证提出的各个改进模块的有效性,设计并完成了 7 组消融实验。实验结果如表 4 所示。表 4 中,第 1 组为基础 YOLOv11s 模型的实验结果,表中“√”标记代表在对应改进点基础上进行的实验。每组实验分别针对基础模型的各个组件进行消融验证,旨在量化各独改进策略对整体性能的贡献。

表 4 YOLO-MAF 模型消融实验

Table 4 Ablation experiment of YOLO-MAF model									
模型	MSEE	SEGE	MASF	P/%	R/%	mAP@0.5/%	mAP@0.5:0.95/%	Param/M	GFLOPs/G
YOLOv11s				49.5	37.3	38.3	23.0	9.4	21.7
A	√			50.1	39.1	39.6	23.7	9.2	21.4
B		√		50.3	38.6	40.5	24.4	9.2	18.3
C			√	52.5	41.3	42.6	24.6	5.0	25.0
D	√	√		52.7	39.9	41.1	25.5	4.9	20.6
E	√		√	53.3	41.8	43.8	26.6	9.0	23.7
F		√	√	53.5	42.7	44.6	27.1	4.8	22.8
G	√	√	√	54.7	43.1	45.6	29.4	4.7	21.9

由表 4 可知,各模块在不同组合下均对检测性能产生积极影响。

单独引入 MSEE 模块(实验 A)后,模型的 mAP@0.5 从 49.5%提升至 50.1%,召回率从 37.3%提升至 39.1%,参数量保持在 9.2 M,表明多尺度边缘增强机制有效改善小目标特征提取。

单独引入 SEGE 模块(实验 B)使 mAP@0.5 达到 50.3%,召回率升至 40.5%,同时参数量与计算复杂度分别降低至 9.2 M 与 18.3 G,验证了 SNI 插值与 GSConvE 组卷积在特征融合中的效率优势。

增加一层小目标检测层,其次,将骨干网络最后 3 层通道数调整至 512,最后采用 MASF 检测头(实验 C)后,精

度提升至 52.5%，召回率 41.3%，mAP@0.5 达 42.6%，在参数 5.0 M、复杂度 25.0 G 的条件下实现较优的精度-效率平衡。

组合实验进一步验证了各模块的互补作用：

MSEE + SEGE (实验 D) 使 mAP@0.5 提升至 41.1%，参数量降至 4.9 M；MSEE+MASF(实验 E)精度达 53.3%，mAP@0.5 为 43.8%；SEGE+MASF(实验 F)精度提升至 53.5%，mAP@0.5 增至 44.6%，参数量仅 4.8 M。

集成所有模块的 YOLO-MAF 模型(实验 G)取得最佳性能：精度 54.7%，召回率 43.1%，mAP@0.5 达到 45.6%，参数量 4.7 M，计算复杂度 21.9 G，结果充分说明

各改进模块的协同作用显著提升了检测精度与模型紧凑性。

2.5 对比实验

为验证改进算法的有效性，在 VisDrone2019 数据集上与多种主流检测算法进行了对比实验，包括 Faster R-CNN^[22]、YOLOv5n^[23]、YOLOv8n^[24]、YOLOv10n^[25]、RT-DETR^[26]以及最新的 YOLOv12^[27]等方法。采用 mAP@0.5、mAP@0.5:0.95、参数量(Param)、GFLOPs 与 帧率(FPS)等指标综合衡量精度与轻量化性能。对未公开 FPS 的模型，均基于官方权重复现实测。为确保对比的公平性，所有方法均在统一硬件与软件环境下复现。实验结果如表 5 所示。

表 5 YOLO-MAF 与主流模型对比实验

Table 5 Comparative experiment between YOLO-MAF and mainstream models

模型	P%	R%	mAP@0.5/%	mAP@0.5:0.95/%	Param/M	GFLOPs/G	FPS
Faster R-CNN	51.8	30.8	38.3	17.9	136.90	369.7	—
CenterNet	82.6	32.7	42.8	25.4	32.60	70.2	40
YOLOv3-tiny	39.1	24.3	23.6	13.2	9.50	14.3	17
YOLOv5n	42.3	33.2	32.9	19.1	2.10	5.8	23
YOLOv5s	51.1	38.1	39.3	23.4	7.80	18.8	26
YOLOv5m	47.7	36.8	39.4	23.0	20.90	48.0	14
YOLOv5l	50.7	38.6	41.4	24.6	46.10	107.8	8
YOLOv6s	40.3	30.5	30.2	17.7	4.10	11.5	—
YOLOv7-tiny	47.6	37.3	35.8	18.8	6.00	13.3	33
YOLOv8n	45.0	33.0	38.1	19.2	2.68	6.8	26
YOLOv8s	50.7	37.9	39.1	23.4	9.80	23.4	19
YOLOv8m	53.3	41.1	42.5	26.0	23.20	67.5	8
YOLOv9n	47.3	34.5	40.4	20.1	4.20	18.2	37
YOLOv10n	45.0	34.5	34.5	19.9	2.20	6.5	36
YOLOv10s	49.9	38.7	39.4	23.6	21.30	9.4	32
RT-DETR	57.2	40.0	41.4	25.1	57.00	20.0	60
YOLOv12s	47.6	36.8	37.3	22.1	9.20	21.7	32
YOLOv11s	49.5	37.3	38.3	23.0	9.40	21.7	34
Ours	54.7	43.1	45.6	29.4	4.70	21.9	45

相较于 Faster R-CNN，本方法在 mAP@0.5 上提升了 7.3%，且参数量与 GFLOPs 分别下降约 96.6%与 94.1%，体现出显著的轻量化优势；与轻量级模型 YOLOv5n 与 YOLOv8n 相比，YOLO-MAF 的 mAP@0.5、mAP@0.5:0.95 均显著更高，且 FPS 提升至 45 (YOLOv8n 为 26)，说明在实时性方面同样占优；与中等规模模型 YOLOv8m 相比，YOLO-MAF 的 mAP@0.5:0.95 更高，但参数量仅为其 20.3%，实现了更优的精度—参数量平衡。

总体来看，YOLO-MAF 在保持较低计算复杂度与高帧率的前提下，实现了检测精度的显著提升。特别是在小

目标检测场景中，模型的性能优势更加明显，这得益于所提出的多尺度边缘信息增强与自适应特征融合机制的协同作用。实验结果充分验证了所提方法在无人机航拍目标检测任务中的有效性与实时性，尤其适用于边缘计算等资源受限场景的实际部署需求。

2.6 与其他改进算法对比实验

为进一步验证所提方法的优势，选取 S-YOLOv11^[28]、SOD-YOLO^[29]、DAU-YOLO 等在 YOLOv11 框架上的代表性改进模型进行对比，这些算法均基于 YOLOv11 改进，具有较强代表性。为保证公平性，所有实验都在同一条件下进行。对比结果如表 6 所示。

表 6 YOLO-MAF 与其他改进模型对比实验

Table 6 Comparative experiment between YOLO-MAF and other improved models							
模型	P%	R%	mAP@0.5/%	mAP@0.5:0.95/%	Param/M	GFLOPs/G	FPS
S-YOLOv11	51.4	40.2	43.8	27.3	5.2	23.6	31
SOD-YOLO	48.9	39.6	42.5	25.8	6.8	25.4	29
DAU-YOLO	52.3	41.0	44.1	27.9	4.9	22.7	39
本文	54.7	43.1	45.6	29.4	4.7	21.9	45

由表 6 可知,YOLO-MAF 在 6 项核心指标(P、R、mAP@0.5、mAP@0.5:0.95、参数量、GFLOPs)及 FPS 上均取得最优;相对于最强对比模型 DAU-YOLO^[16],其在 mAP@0.5 与 mAP@0.5:0.95 上分别提升 1.5%,同时以更低参数量(4.7 M)与更小计算量(21.9 G)实现更高速(45 fps),体现出更佳精度-效率-复杂度平衡。

S-YOLOv11 强化特征金字塔网络(feature pyramid network,FPN)但边缘信息不足;SOD-YOLO 通过深度可分卷积减小通道表达受限;DAU-YOLO 动态上采样有利于跨层一致性但在密集/极小目标上仍有瓶颈。YOLO-MAF 得益于 MSEE 的边缘增强、SEGE 的对齐融合与 MASF 的自适应加权,在较低复杂度下实现更优精度-一速

度平衡。

2.7 可视化分析

为了验证改进后的算法在边缘小目标的检测能力,对改进前后的模型进行了热力图对比。从图 6 的对比图可以看出在黄框标注的人群密集区域,基线模型的特征激活相对模糊,热力图显示该区域的目标识别边界不够清晰,存在目标定位不精确的问题。改进算法在相同区域展现出明显优势:黄框内的目标检测更加精准,热力图显示出更强烈且集中的特征响应。算法能够准确捕获人群聚集区域的边界信息,实现了更精细的目标定位。该对比充分说明了 YOLO-MAF 算法在处理复杂城市场景时的技术优势。

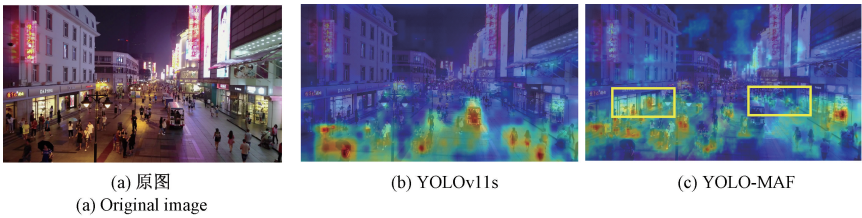
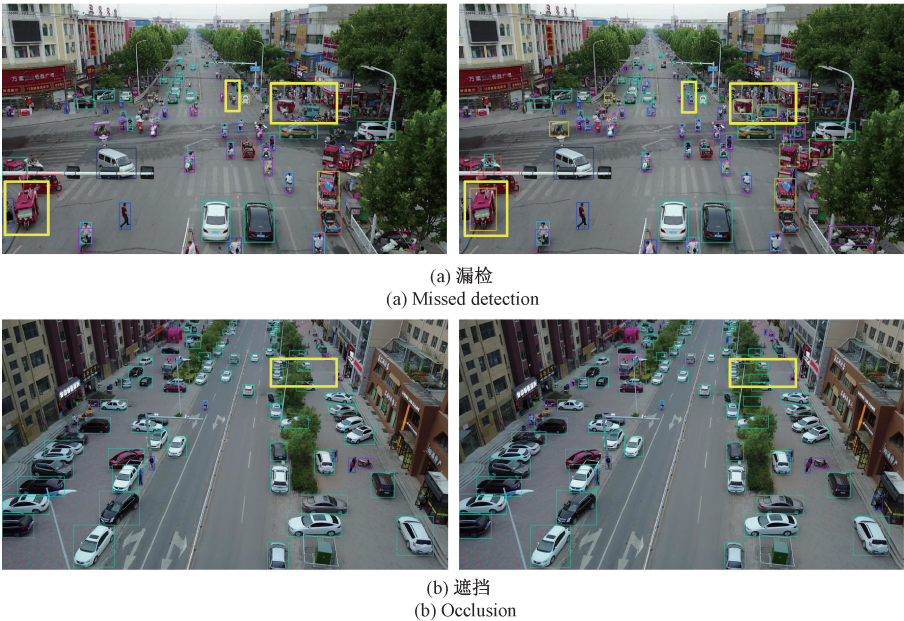


图 6 YOLOv11s 和 YOLO-MAF 两种算法的热力图对比
Fig. 6 Comparison of heatmaps between YOLOv11s and YOLO-MAF algorithms

为全面评估所提方法的实际应用效果,在 VisDrone2019 数据集上构建了 4 组不同复杂度的测试场景

景进行对比实验,具体如图 7 所示。

在小目标密集场景中,YOLOv11s 对目标区域的人员



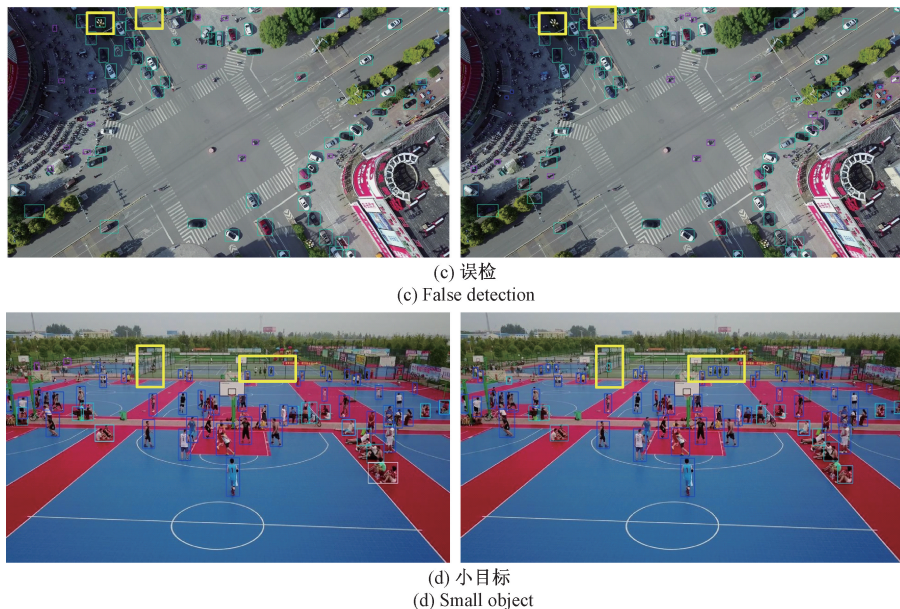


图7 YOLOv11s和YOLO-MAF模型在4种不同复杂场景下检测结果对比

Fig. 7 Comparison of detection results between YOLOv11s and YOLO-MAF models under four different complex scenarios

目标检测存在大量遗漏,而YOLO-MAF能够精准捕获微小目标,显著提升检测覆盖率。在车辆遮挡场景中,基准算法在处理相互遮蔽的车辆时表现不佳,检测准确率明显不足,改进算法通过增强的特征学习能力有效解决了遮挡检测难题。在复杂交通场景中,YOLOv11s出现严重的目标丢失现象,多个车辆未能被有效识别,而YOLO-MAF实现了近乎完美的检测覆盖。在环境干扰场景中,基准模型频繁将道路设施、光影变化误判为目标,造成较高虚检率,改进模型通过优化的判别机制大幅提升了检测准确性。在小目标场景中,YOLO-MAF在远景对于模糊的小目标的识别率要明显高于YOLOv11s。综合分析表明,YOLO-MAF在目标召回率、环境适应性和检测稳定性等关键指标上均实现显著提升,证明了其在复杂航拍场景下的优越性能。

3 结 论

针对无人机图像中小尺寸目标检测精度低的问题,本研究提出轻量级目标检测算法YOLO-MAF,有效解决了多尺度融合效率低和小目标检测精度不足的问题。设计了MSEE、SEGE与MASF 3个模块,并验证其协同作用。

在VisDrone2019数据集上,YOLO-MAF的 $mAP@0.5$ 和 $mAP@0.5:0.95$ 分别达到45.6%和29.4%,相比YOLOv11s分别提升7.3%和6.4%。同时参数量减少50%,结果表明所提方法在提高小目标检测精度的同时有效降低模型复杂度。未来工作将结合模型剪枝与知识蒸馏等轻量化优化技术,进一步提高模型推理速度,并探索其在边缘设备上的高效部署与实时应用。

参考文献

- [1] 赵佰亭,张晨,贾晓芬. ECC-YOLO:一种改进的钢材表面缺陷检测方法[J]. 电子测量与仪器学报, 2024, 38(4): 108-116.
ZHAO B T, ZHANG CH, JIA X F. ECC-YOLO: An improved method for detecting surface defects in steel[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(4): 108-116.
- [2] BAKIRCI M. Utilizing YOLOv8 for enhanced traffic monitoring in intelligent transportation systems(ITS) applications[J]. Digital Signal Processing, 2024, 152: 104594.
- [3] AFRIN T, YODO N, DEY A, et al. Advancements in UAV-enabled intelligent transportation systems: A three-layered framework and future directions[J]. Applied Sciences, 2024, 14(20): 9455.
- [4] KABASHKIN I, KULMURZINA A, NADIMOV B, et al. Synchronized multi-point UAV-based traffic monitoring for urban infrastructure decision support[J]. Drones, 2025, 9(5): 370.
- [5] 刘熹,陈晨,双丰. 基于改进YOLOv7-tiny的多种类绝缘子检测算法[J]. 仪器仪表学报, 2024, 45(9): 102-110.
LIU X, CHEN CH, SHUANG F. Multi-type insulator detection algorithm based on improved YOLOv7-tiny[J]. Chinese Journal of Scientific Instrument, 2024, 45(9): 102-110.
- [6] BAKIRCI M. Vehicular mobility monitoring using remote sensing and deep learning on a UAV-based mobile computing platform[J]. Measurement, 2025, 244: 116579.

- [7] 蓝慧玲,刘琼.基于 YOLOv8-ERD 的无人机航拍小目标检测算法[J].电子测量技术,2025,48(18):150-158.
LAN H L, LIU Q. UAV aerial small target detection algorithm based on YOLOv8-ERD [J]. Electronic Measurement Technology, 2025, 48(18): 150-158.
- [8] 李珺,丁彬彬,史维娟,等.改进 YOLOv11 的无人机航拍图像检测算法[J].电子测量技术,2025,48(18):111-121.
LI J, DING B B, SHI W J, et al. Improved UAV aerial image detection algorithm based on YOLOv11 [J]. Electronic Measurement Technology, 2025, 48(18): 111-121.
- [9] 翁俊辉,成乐,黄曼莉,等.基于 CS-YOLOv5s 的无人机航拍图像小目标检测[J].电子测量技术,2024,47(7):157-162.
WENG J H, CHENG L, HUANG M L, et al. Small object detection in UAV aerial images based on CS-YOLOv5s[J]. Electronic Measurement Technology, 2024, 47(7): 157-162.
- [10] XU L Y, ZHAO Y F, ZHAI Y H, et al. Small object detection in UAV images based on YOLOv8n [J]. International Journal of Computational Intelligence Systems, 2024, 17: 223.
- [11] CHENG Q, WANG Y, HE W, et al. Lightweight air-to-air unmanned aerial vehicle target detection model[J]. Scientific Reports, 2024, 14: 2609.
- [12] FENG F, YANG L, ZHOU Q X, et al. A lightweight small object detection algorithm for UAV aerial imagery [J]. IET Image Processing, 2025, 19(1): e13314.
- [13] ZHANG H, XIAO P, YAO F, et al. Fusion of multi-scale attention for aerial images small-target detection model based on PARE-YOLO[J]. Scientific Reports, 2025, 15: 4753.
- [14] 孙志刚,马俊燕,李文强.面向遥感图像的变换注意力目标检测算法[J].模式识别与人工智能,2024,37(11):1045-1056.
SUN ZH G, MA J Y, LI W Q. Object detection algorithm with transform attention for remote sensing images [J]. Pattern Recognition and Artificial Intelligence, 2024, 37(11): 1045-1056.
- [15] WANG Q, YU CH X. AMFE-YOLO: A small object detection model for drone images [J]. IET Image Processing, 2025, 19(1): e70110.
- [16] WAN ZH Y, LAN Y ZH, XU ZH D, et al. DAU-YOLO: A lightweight and effective method for small object detection in UAV images[J]. Remote Sensing, 2025, 17(10): 1768.
- [17] WU Z C, ZHEN H, ZHANG X X, et al. SEMA-YOLO: Lightweight small object detection in remote sensing image via shallow-layer enhancement and multi-scale adaptation [J]. Remote Sensing, 2025, 17(11): 1917.
- [18] YU Y X, HUANG M H, WANG K, et al. LS-YOLO: A lightweight small-object detection framework with region scaling loss and self-attention for intelligent transportation systems [J]. Signal, Image and Video Processing, 2025, 19: 1005.
- [19] KHANAM R, HUSSAIN M. YOLOv11: An overview of the key architectural enhancements [J]. ArXiv preprint arXiv:2410.17725,2024.
- [20] HE L H, ZHOU Y Z, LIU L, et al. Research on object detection and recognition in remote sensing images based on YOLOv11 [J]. Scientific Reports, 2025, 15(1): 96314.
- [21] ZHU P F, WEN L Y, DU D W, et al. Detection and tracking meet drones challenge[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(11): 7380-7399.
- [22] 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.
- [23] KHANAM R, HUSSAIN M. What is YOLOv5: A deep look into the internal features of the popular object detector [J]. ArXiv preprint arXiv: 2407.20892,2024.
- [24] YASEEN M. What is YOLOv8: An in-depth exploration of the internal features of the next-generation object detector[J]. ArXiv preprint arXiv: 2408.15857,2024.
- [25] WANG A Q, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection [J]. ArXiv preprint arXiv:2405.14458,2024.
- [26] ZHAO Y AN, LYU W Y, XU SH L, et al. DETRs beat YOLOs on real-time object detection (RT-DETR) [J]. ArXiv preprint arXiv:2304.08069,2023.
- [27] TIAN Y J, YE Q X, DOERMANN D. YOLOv12: Attention-centric real-time object detectors[J]. ArXiv preprint arXiv:2502.12524,2025.
- [28] WANG CH SH, SONG X Q, WANG J, et al. An improved YOLOv11 algorithm for small object detection in UAV images[J]. Signal, Image and Video Processing, 2025, 19: 507.
- [29] XIAO Y Z, DI N. SOD-YOLO: A lightweight small object detection framework [J]. Scientific Reports, 2024, 14: 77513.

作者简介

胡闽南,硕士,主要研究方向为目标检测与图像融合。

E-mail:2224292552@qq.com

赵鸣(通信作者),教授,博士,硕士研究生导师,主要研究方向为模式识别与图像处理,智能系统,智能数据处理与分析。

E-mail:hitmzhao@gmail.com