

多尺度交互融合的轻量级小目标检测算法^{*}

郝江鹏 牛芳琳 于玲 韩驰

(辽宁工业大学电子与信息工程学院 锦州 121000)

摘要: 针对无人机航拍图像中目标尺寸小、尺度变化显著、背景复杂以及计算资源有限等挑战,提出了一种改进YOLOv11的算法模型LMH-YOLO。首先,提出C3k2_PMDRB,以增强级联扩展路径生成的多尺度语义特征并减少模型参数量;其次,设计LCFI_Net网络架构,实现跨层信息的高效融合,并精准感知小目标的空间位置信息;再次,构建轻量化检测头LGDHead,降低计算开销以提升计算效率;最后,提出 W^3 -MPDIoU损失函数,优化模型收敛并降低漏检率。实验结果表明,在VisDrone2019数据集,LMH-YOLO的mAP@50提升了4.8%,参数量减少了42.3%,模型大小缩减了32.7%;在TinyPerson极小目标数据集上,mAP@50提升了6.7%,参数量和模型大小分别下降了42.3%和30.8%。改进算法实现了性能与轻量化的最佳权衡,更适用于无人机航拍小目标检测任务。

关键词: 目标检测;YOLOv11;轻量化;多尺度特征;网络架构;损失函数

中图分类号: TP391.4;TN911.73 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Lightweight small object detection algorithm based on
multi-scale interactive fusion

Hao Jiangpeng Niu Fanglin Yu Ling Han Chi

(School of Electronic and Information Engineering, Liaoning University of Technology, Jinzhou 121000, China)

Abstract: To address the challenges of small target sizes, significant scale variations, complex backgrounds and limited computational resources in UAV aerial images, this paper proposes an improved YOLOv11-based algorithm, termed LMH-YOLO. First, we introduce C3k2_PMDRB to enhance multi-scale semantic features generated by cascade expansion paths while reducing the model's parameter count. Second, we design the LCFI_Net architecture to enable efficient cross-layer information fusion and accurately capture the spatial locations of small targets. Subsequently, a lightweight detection head, LGDHead, is developed to reduce computational overhead and improve efficiency. Finally, the W^3 -MPDIoU loss function is proposed to optimize model convergence and mitigate missed detections. Experimental results demonstrate that on the VisDrone2019 dataset, LMH-YOLO achieves a 4.8% improvement in mAP@50, along with a 42.3% reduction in parameter count and a 32.7% decrease in model size. On the TinyPerson dataset for extremely small targets, mAP@50 increases by 6.7%, while parameter count and model size decrease by 42.3% and 30.8%, respectively. These results indicate that LMH-YOLO attains an optimal balance between performance and model compactness, making it particularly well-suited for small target detection in UAV aerial imagery.

Keywords: object detection; YOLOv11; lightweight design; multi-scale features; network architecture; loss function

0 引言

无人机(unmanned aerial vehicle, UAV)具备机动性强、部署灵活等优势,适合在复杂的场景中执行多样化任务。现已广泛应用于农作物监测、交通监管、输电线路巡检、搜索与救援和国防军事等领域^[1]。然而,在实际应用

中,无人机航拍检测面临目标尺寸小、尺度变化显著、背景复杂、易受遮挡等挑战,同时边缘部署平台资源有限,使得模型检测性能与轻量化难以兼顾。为应对上述挑战,本文设计了跨层多尺度特征提取网络架构并采用高效的通道压缩策略,使模型在检测精度和轻量化之间取得平衡,对提升航拍小目标的检测精度和实现模型结构的高效轻量化具有

重要意义。

目前,基于深度学习的目标检测算法主要分为两类:以 SSD(single shot multibox detector)^[2] 和 YOLO(you only look once)^[3] 为代表的单阶段检测算法,以及以区域卷积神经网络(region-based convolutional neural networks, R-CNN)^[4] 系列为代表的双阶段检测算法。近年来,针对无人机小目标检测任务,学者们在此领域提出了多种改进策略。田红鹏等^[5]在 RT-DETR 中提出 Focaler-MPDIoU(focaler-minimum point distance intersection over union)损失函数以提高边界框的回归精度并减少模型的漏检率,但其计算开销较大,难以在边缘设备上部署;张永宏等^[6]在 MFSF-DETR 中设计了跨尺度特征移位融合网络(cross-scale feature shift fusion, CFSF),该融合网络通过通道移位操作对非相邻层语义信息进行有效交互,但其性能在一定程度上受限于固定移位策略及对特征金字塔层数的依赖;Liu 等^[7]在 YOLOv5 中通过引入坐标注意力(coordinate attention, CA)以增强通道与位置信息之间的关联来提升检测精度,但在全局占比极低的小目标场景中仍易产生漏检;李泽胤等^[8]在 YOLOv8 中提出任务对齐动态检测头(task align dynamic detection head, TADD)以提升目标的定位和分类性能并降低误检率,但对小目标的检测精度提升有限;贾亮等^[9]在 UCM-YOLOv8 中将卷积加性自注意力机制(conv additive self-attention, CASA)融合到 C2f 模块,以降低背景无关信息的干扰,但检测精度提升幅度有限,且模型的泛化能力尚未得到充分验证;张彪等^[10]在 LASW-YOLOv8 中提出了自适应旋转卷积模块(adaptive rotation convolution kernel, ARConv),该模块根据缺陷方向与形状信息动态调整卷积核的方向与权重分布,实现对多方向缺陷特征的自适应建模;冉宁等^[11]在 EDB-YOLOv11 中提出基于 C3k2 改进的 EMSE(effective mbconv squeeze-excitation)以提升模型的特征提取能力,但缺乏对小目标细粒度空间信息的有效建模能力;崔克彬等^[12]在 YOLO11 中设计 HCDNet(HSFPN with coordinate attention and dysample module network)增强模型的多尺度特征表达能力,并在特征图上引入位置信息以精准捕获小目标的空间分布;张红民等^[13]在 SDA-YOLO11 中构建 HSFPN(heterogeneous selective feature pyramid network),通过深浅层特征的交互来实现多尺度信息的高效融合,但计算复杂度的大幅增加在一定程度上限制了模型的实时推理性能。

尽管上述方法在不同程度上提升了对小目标的检测性能,但在无人机视角下的目标检测受限于目标尺度小且分布密集、背景复杂、易受遮挡以及部署平台的计算资源有限,其检测任务仍面临挑战。为此,本文提出一种多尺度层间交互融合的轻量化模型(lightweight multi-scale hierarchical fusion-YOLO, LMH-YOLO)。主要贡献为:

1) 提出 C3k2_PMDRB (C3k2_partial multi-scale

depth-wise residual block),将输入特征的部分通道引入多尺度深度可分离卷积,更有效地建模多尺度的空间信息,且降低模型参数量与计算开销。

2) 设计轻量化跨层级特征交互网络架构(lightweight cross-stage feature interaction_network, LCFI_Net),通过通道压缩策略实现轻量化,并将低层的细粒度特征通过跳跃连接实现与高层的语义特征进行高效融合以提升检测精度。

3) 提出轻量级分组深度可分离卷积检测头(lightweight grouped depthwise head, LGDHead),进一步降低模型参数量与计算量,使模型更高效。

4) 提出 W^{v3} -MPDIoU(wisev3-minimum point distance intersection over union)回归损失,通过动态梯度增益机制自适应提升小目标样本的贡献度,从而增强定位精度并促进模型收敛。

1 YOLOv11 及其改进算法

YOLOv11 是以 Backbone-Neck-Head 为架构的单阶段目标检测算法,Backbone 主干网络集成 C3k2 和 C2PSA(cross-stage partial selective attention);Neck 颈部网络采用 PAN-FPN(path aggregation network-feature pyramid network)^[14-15]网络架构构建可高效融合不同层级语义信息的双向特征金字塔;检测头 Head 采用解耦式,并以完整交并比损失(complete intersection over union, CIoU)与分布损失(distribution focal loss, DFL)作为定位损失函数。本文选取 YOLOv11n 作为基线模型,旨在保持合理检测性能的同时更好地满足无人机小目标检测对实时航拍高效率与低算力的要求。

本文设计的 LMH-YOLO 网络结构如图 1 所示。提出 C3k2_PMDRB 模块以增强模型对小目标的多尺度特征提取能力。设计全新的 LCFI_Net 网络架构,实现轻量化跨层级特征信息的高效融合。采用 LGDHead,实现模型精度与效率的平衡。最后,将 Wise-IoUv3 与 MPDIoU 相结合提出 W^{v3} -MPDIoU 损失函数,促进模型收敛并提升检测性能。

2 LMH-YOLO

2.1 C3k2_PMDRB

传统 C3k2 对输入特征的全部通道采用单一尺度卷积核而难以充分捕获小目标的多尺度上下文信息,同时导致计算量与参数量增加。基于上述不足,本文通过采用部分卷积(partial convolution, PConv)^[16]并结合多尺度深度残差模块(multi-scale depth-wise residual block, MDRB)^[17]提出 C3k2_PMDRB,如图 2 所示。

1) MDRB

MDRB 以多尺度深度可分离卷积(multi-scale depth-wise convolution, MDC)为核心结构,通过在多个尺度上并

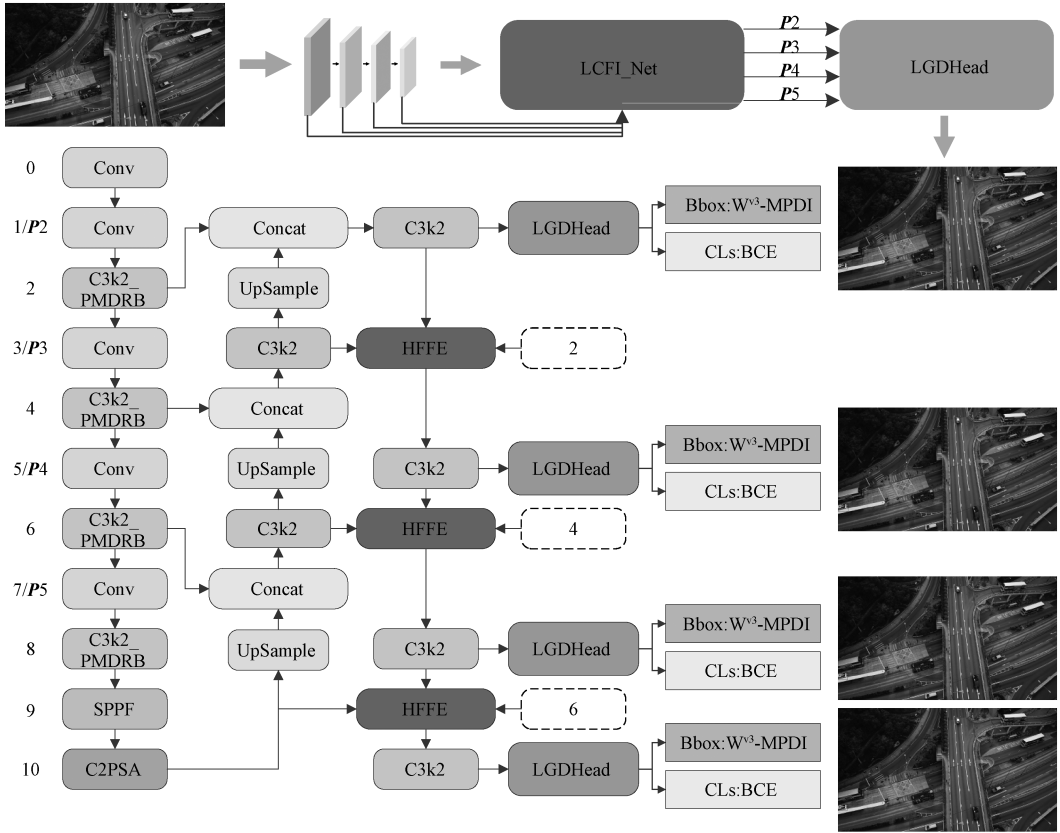


图 1 LMH-YOLO 模型结构

Fig. 1 The network architecture of LMH-YOLO

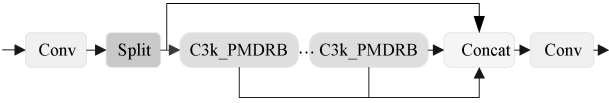


图 2 C3k2_PMDRB 模块结构

Fig. 2 C3k2_PMDRB module architecture

行执行深度可分离卷积来高效建模不同感受野下的空间结构语义,从而强化小目标间的细粒度空间表征,显著提升了模型对复杂小目标场景的感知能力与识别精度,其结构如图 3 所示。

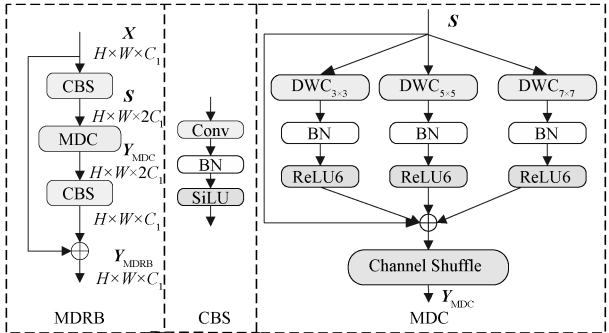


图 3 MDRB 模块结构及其子模块结构

Fig. 3 Architecture of the MDRB module and its sub-modules

MDRB 模块公式表示为:

$$\mathbf{Y}_{\text{MDRB}}(\mathbf{X}) = \mathbf{X} + \text{CBS}(\text{MDC}(\text{CBS}(\mathbf{X}))) \quad (1)$$

其中, $\mathbf{X}$ 为输入, $\mathbf{Y}_{\text{MDRB}}$ 为输出。初始 CBS(conv batch normalization silu)单元通过扩展通道数以增加特征维度, MDC 单元用于捕获多尺度的上下文信息,第 2 个 CBS 单元将通道数恢复至原始通道以确保残差结构的可行性。MDC 单元并通过 Channel Shuffle 通道混洗机制在不同组间进行重新排列通道,促进跨组信息流动。具体操作:

$$\mathbf{Y}_{\text{MDC}} = \Gamma(\mathbf{S} + \sum_{k \in \mathcal{K}} \text{ReLU6}(\text{BN}(\text{DWConv}_{k \times k}(\mathbf{S})))) \quad (2)$$

其中, $\mathbf{S}$ 为经过初始 CBS 扩展通道的输入特征, $\mathbf{Y}_{\text{MDC}}$ 为输出特征, $\mathcal{K} = \{3, 5, 7\}$ 为卷积核大小, $\Gamma(\cdot) = \text{ChannelShuffle}(\cdot)$ 表示通道混洗机制, BN (batch normalization) 为批量归一化, ReLU6 为激活函数。

2) PMDRB

为实现模型的轻量化,本文基于部分卷积 PConv 的工作机制,提出了 PMDRB 模块。该模块设计如图 4 所示,针对 c 个通道,选取部分通道 c_p 进行 MDRB 运算,从而实现多尺度空间特征提取以生成具有补充性的增量特征,未参与 MDRB 运算的其余 $c - c_p$ 通道保持原始特征不变,以确保部分原始信息在全通道上的连续流动和语义完整性。

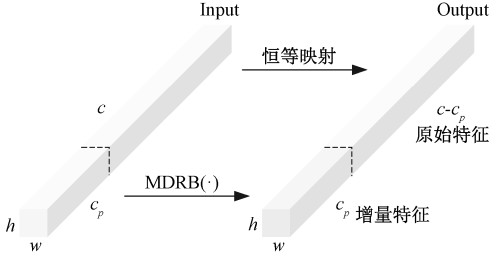


图 4 PMDRB 模块结构

Fig. 4 PMDRB module architecture

对于给定的输入为 $I \in \mathbf{R}^{h \times w}$, 将 MDRB 的运算量记为 $\text{FLOPs}_{\text{MDRB}}(h, w, c)$, 则 PMDRB 的运算量为:

$$\text{FLOPs}_{\text{PMDRB}} = \text{FLOPs}_{\text{MDRB}}(h, w, c_p) \quad (3)$$

部分比例超参数 r :

$$r = \frac{c}{c_p} \quad (4)$$

相较于常规的 MDRB, PMDRB 的核心优势在于仅对 c_p 个通道进行多尺度深度卷积建模并保留剩余通道的原生语义, 该设计通过“增量特征+原始特征”的协同表征机制来实现模型对小目标进行高效地多尺度特征提取, 显著增强小目标特征的可分辨性。

由式(3)、(4)可知, PMDRB 的计算量相较 MDRB 显著降低, 同时, 通过在部分通道所学习的空间结构增量特征与原始语义通道的残差信息相互补充, 增强模型对局部纹理、结构细节与多尺度上下文的联合建模能力。进一步将 PMDRB 集成至 C3k2 而得到 C3k2_PMDRB, 该模块能够在降低计算效率的同时, 有效地捕捉复杂的空间关系并专注于显著区域, 使模型取得了表征能力与计算效率之间的绝佳平衡。

2.2 LCFI_Net

在 YOLOv11 中, PAN-FPN 网络架构缺乏跨层语义选择机制, 导致小目标语义信息表达不足, 层间语义融合低效且计算冗余。为此, 本文设计了全新的 LCFI_Net 网络架构, 如图 5 所示。该架构将非相邻层的语义信息通过跨层连接进行有效融合, 显著增强模型对强浅层细粒度特征的建模能力。为进一步提升特征融合效率, 设计了层间特征融合增强模块 (hierarchical feature fusion enhancement, HFEE) 以增强层间特征的高效信息交互。同时本文在颈部至检测头的特征传递过程中引入通道压缩策略, 将输出通道数压缩至 256。在保证模型特征表达能力的前提下, 该策略显著降低了模型的计算复杂度, 使参数量减少 26.9%。LCFI_Net 高效促进了跨层语义交互并显著提升了模型的检测精度。

HFEE 对不同特征图进行跨层高效融合, 并通过信息集成注意力机制 (information integration attention, IIA) [18] 信息融合机制实现信息加权整合。如图 6 所示。

部分输入特征图通过下采样 (down sample, DSAMPLE)

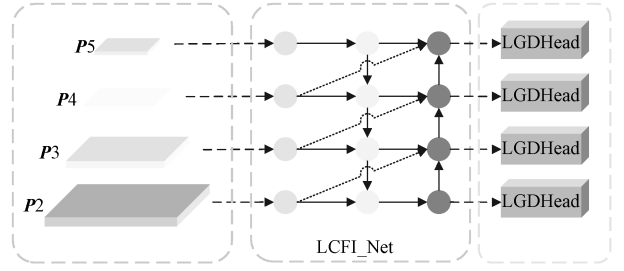


图 5 LCFI_Net 网络架构

Fig. 5 Architecture of the LCFI_Net network

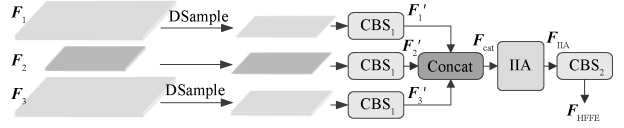


图 6 HFEE 模块结构

Fig. 6 HFEE module architecture

进行尺寸对齐, CBS₁ 模块对特征图的通道进行线性降维映射, Concat 操作对跨层特征进行拼接以促进特征信息的高效交互。

$$\mathbf{F}'_k = \text{CBS}_1(\text{Align}(\mathbf{F}_k)), k \in \{1, 2, 3\} \quad (5)$$

$$\mathbf{F}_{\text{cat}} = \text{Concat}(\mathbf{F}'_1, \mathbf{F}'_2, \mathbf{F}'_3) \quad (6)$$

其中, 特征图的输入为 $\mathbf{F}_k \in \mathbf{R}^{H_k \times W_k \times C}$, $k \in \{1, 2, 3\}$, Align($\cdot$) 特征对齐操作, 用于将不同尺度特征对齐到统一空间尺寸。 $\mathbf{F}'_k \in \mathbf{R}^{H \times W \times C_r}$, $C_r < C$, 实现通道线性降维投影。Concat 表示特征拼接操作 $\mathbf{F}_{\text{cat}} \in \mathbf{R}^{H \times W \times 3C_r}$ 。

为了增强跨层聚合特征图 $\mathbf{F}_{\text{cat}}$ 的语义信息, 本文采用一种 IIA 来重建特征图中小目标的空间信息并捕获相对位置关系, 该机制在多尺度条件下捕获不同尺寸目标的全局信息, 其结构如图 7 所示。

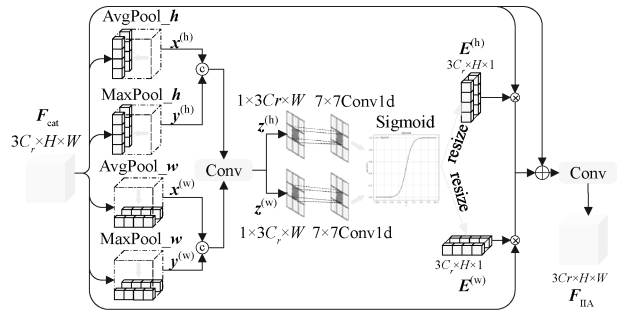


图 7 IIA 模块结构

Fig. 7 IIA module architecture

$\mathbf{F}_{\text{cat}}$ 进入 IIA 之后, 对其空间维度进行调整, 沿高度与宽度方向建立序列依赖, 生成两个新的特征图 $\mathbf{h}$ 和 $\mathbf{w}$ 。

$$\mathbf{h}, \mathbf{w} \leftarrow \text{Resize}(\mathbf{F}_{\text{cat}}) \quad (7)$$

其中, $\mathbf{h} \in \mathbf{R}^{W \times 3C_r \times H}$, $\mathbf{w} \in \mathbf{R}^{H \times 3C_r \times W}$ 。

新特征图 $\mathbf{h}$ 和 $\mathbf{w}$ 沿高度与宽度方向独立执行平均池

化和最大池化以压缩冗余特征并增强关键特征区域。其输出可表示为:

$$\mathbf{x}^{(h)} = \text{AvePool}(\mathbf{h}) \quad (8)$$

$$\mathbf{y}^{(h)} = \text{MaxPool}(\mathbf{h}) \quad (9)$$

$$\mathbf{x}^{(w)} = \text{AvePool}(\mathbf{w}) \quad (10)$$

$$\mathbf{y}^{(w)} = \text{MaxPool}(\mathbf{w}) \quad (11)$$

将经过池化操作后的特征图分别进行拼接,并调整通道数,以获得最终的信息表达形式。

$$\mathbf{z}_h = \text{Conv}(\text{Concat}(\mathbf{x}^{(h)}, \mathbf{y}^{(h)})) \quad (12)$$

$$\mathbf{z}_w = \text{Conv}(\text{Concat}(\mathbf{x}^{(w)}, \mathbf{y}^{(w)})) \quad (13)$$

$$\text{其中, } \mathbf{z}_h \in \mathbf{R}^{1 \times 3C_r \times H}, \mathbf{z}_w \in \mathbf{R}^{1 \times 3C_r \times W}.$$

采用一维卷积 Conv1d 对特征 $\mathbf{z}_h$ 和 $\mathbf{z}_w$ 进行处理,以增强位置信息的融合效果,并通过 Sigmoid 激活函数 $\sigma(\cdot)$ 生成沿高和宽维度的空间位置权重 $\mathbf{E}_h$ 和 $\mathbf{E}_w$:

$$\mathbf{E}_h = \sigma(\text{Conv1d}(\mathbf{z}_h)) \quad (14)$$

$$\mathbf{E}_w = \sigma(\text{Conv1d}(\mathbf{z}_w)) \quad (15)$$

输入特征 $\mathbf{F}_{\text{cat}}$ 分别与高、宽注意力权重进行逐元素加权融合 $\otimes$, 并与初始特征残差相加生成输出特征 $\mathbf{F}_{\text{IIA}}$ 。

$$\mathbf{F}_{\text{IIA}} = \mathbf{F}_{\text{cat}} + (\mathbf{F}_{\text{cat}} \otimes \mathbf{E}_h) + (\mathbf{F}_{\text{cat}} \otimes \mathbf{E}_w) \quad (16)$$

CBS₂ 对 $\mathbf{F}_{\text{IIA}}$ 的输出进行通道恢复,以确保其与主干网络后续模块在通道维度上保持一致。HFPE 的最终输出表示为:

$$\mathbf{F}_{\text{HFPE}} = \text{CBS}_2(\mathbf{F}_{\text{IIA}}) \quad (17)$$

$$\text{其中, } \mathbf{F}_{\text{IIA}} \in \mathbf{R}^{H \times W \times 3C_r}, \mathbf{F}_{\text{HFPE}} \in \mathbf{R}^{H \times W \times C}.$$

HFPE 通过跨层融合与 IIA 自适应增强的协同作用,有效提升了跨层语义的交互能力,并增强了模型在复杂场景下的对小目标判别能力。

2.3 LGDHead

YOLOv11 检测头采用解耦式结构,将分类和回归任务独立建模,不可避免地会带来更大的计算开销以及更高的参数量。针对 Detect_LADHead (detect lightweight, asymmetric decoupled detection head)^[19] 在全局上下文建模方面的局限,通过在特征预处理阶段引入分组卷积 (group convolution, GConv)^[20], 据此设计了改进的 LGDHead 检测头。如图 8 所示。

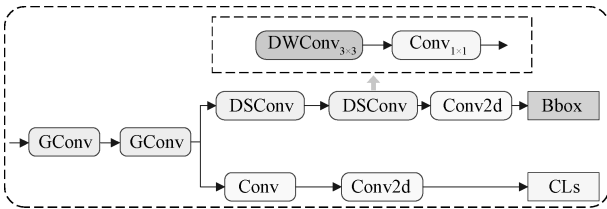


图 8 LGDHead 模块结构

Fig. 8 LGDHead module architecture

1) 分组卷积

组卷积通过将输入特征划分为 G 个互不重叠的子组,并让每个子组独立生成对应的输出,有效减少卷积核的参

与范围,同时显著降低了参数规模与计算量。如图 9 所示,左侧为标准卷积 (Standard Convolution),右侧为 GConv。

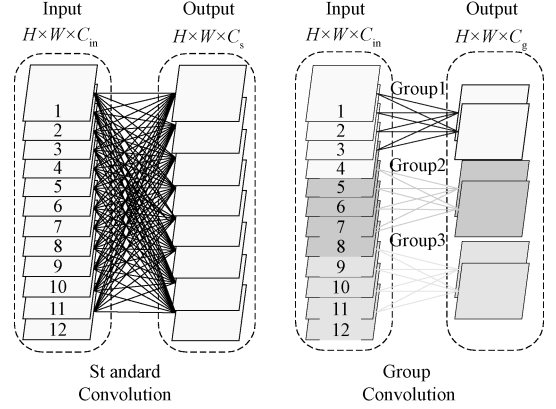


图 9 GConv 模块结构

Fig. 9 GConv module architecture

标准卷积计算公式为:

$$\text{Params}_{\text{standard}} = C_{\text{out}} \times C_{\text{in}} \times K^2 \quad (18)$$

$$\text{FLOPs}_{\text{standard}} = C_{\text{out}} \times C_{\text{in}} \times K^2 \times H \times W \quad (19)$$

分组卷积计算公式为:

$$\text{Params}_{\text{group}} = G \times \left(\frac{C_{\text{out}}}{G} \times \frac{C_{\text{in}}}{G} \times K^2 \right) =$$

$$\frac{C_{\text{out}} \times C_{\text{in}} \times K^2}{G} \quad (20)$$

$$\text{FLOPs}_{\text{group}} = \frac{C_{\text{out}} \times C_{\text{in}} \times K^2 \times H \times W}{G} \quad (21)$$

其中,输入特征图的空间分辨率为 $H \times W$, C_{in} 和 C_{out} 分别为输入与输出通道数, K 为卷积核的空间尺寸, G 为组卷积的分组数量。作为特征预处理层,分组卷积通过分组计算将参数量和运算量降低至标准卷积的 $1/G$,从而减少通道间冗余耦合,并实现适度的通道混合。

2) 深度可分离卷积

深度可分离卷积 (depthwise separable convolution, DSConv)^[21] 分为逐通道卷积 (Depthwise Convolution) 和逐点卷积 (Pointwise Convolution)。如图 10 所示。

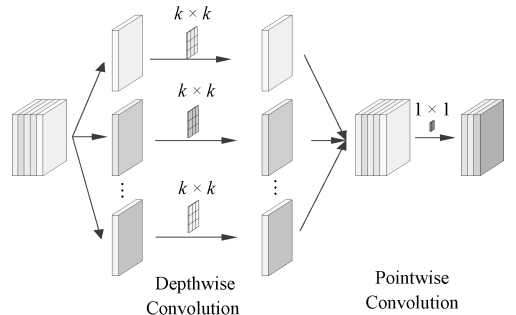


图 10 DSConv 模块结构

Fig. 10 DSConv module architecture

逐通道卷积对输入特征的每个通道分别进行卷积,而逐点卷积是对逐通道卷积的输出特征使用 1×1 卷积进行通道融合,从而实现通道间的信息整合和维度变换。DSConv 的参数量 $\text{Params}_{\text{DSConv}}$ 和运算量 $\text{FLOPs}_{\text{DSConv}}$ 分别为:

$$\text{Params}_{\text{DSConv}} = C_{\text{in}} \times K^2 + C_{\text{in}} \times C_{\text{out}} \quad (22)$$

$$\text{FLOPs}_{\text{DSConv}} = C_{\text{in}} \times H \times W \times (K^2 + C_{\text{out}}) \quad (23)$$

由式(18)~(21)可知, $\text{Params}_{\text{DSConv}}$ 和 $\text{FLOPs}_{\text{DSConv}}$ 分别是标准卷积的 $1/C_{\text{out}} + 1/K^2$ 。通过逐通道卷积和逐点卷积的组合,DSConv 在保持多层次空间信息建模能力的同时,显著降低了模型的计算开销和参数量。

因此, LGDHead 在保持模型轻量化的同时,通过 GConv 和 DSConv 的协同作用显著增强了全局空间信息的建模能力。

2.4 损失函数 W^3 -MPDIoU

基线模型 YOLOv11 采用的 CIoU 对中等质量样本关注不足,影响边界框回归的整体精度和鲁棒性,且无法动态调整梯度以适应训练数据的实际情况。为此,本文设计 W^3 -MPDIoU 损失函数,通过 MPDIoU^[22] 进一步优化边界框回归的几何约束,结合 Wise-IoUv3^[23] 引入的动态梯度增益 R_{WIoU} 调节梯度分配,增强模型对中等质量锚框的关注,促进网络收敛并提升回归精度。

MPDIoU 旨在最小化预测框与真实框在左上角与右下角坐标点之间的欧氏距离误差,并在保持几何约束有效性的同时简化了计算过程。其损失函数定义为:

$$\text{MPDIoU} = \text{IoU} - \frac{\rho^2(P_1^{\text{pred}}, P_1^{\text{gt}})}{\omega^2 + h^2} - \frac{\rho^2(P_2^{\text{pred}}, P_2^{\text{gt}})}{\omega^2 + h^2} \quad (24)$$

$$L_{\text{MPDIoU}} = 1 - \text{MPDIoU} \quad (25)$$

其中, ω, h 为输入图像的宽和高, $P_1^{\text{pred}}, P_2^{\text{pred}}$ 和 $P_1^{\text{gt}}, P_2^{\text{gt}}$ 分别是预测框和真实框的左上角与右下角的点, $\rho^2(\cdot)$ 表示欧氏距离。

为进一步提升不同质量样本的梯度调控能力,引入基于离群度 β (outlier degree) 的动态增益 R_{WIoU} :

$$R_{\text{WIoU}} = \frac{\beta}{\delta \alpha^{\beta-\delta}} \quad (26)$$

$$\beta = \frac{L_{\text{IoU}}^*}{L_{\text{IoU}}} \in [0, +\infty) \quad (27)$$

其中, δ 和 α 为超参数。锚框的离群程度 β 用于动态衡量样本匹配质量的异常程度, L_{IoU}^* 表示当前样本的 IoU 损失, $\overline{L_{\text{IoU}}}$ 表示历史 IoU 损失的指数加权移动平均。

锚框的离群程度 β 由当前样本 IoU 损失 L_{IoU}^* 与历史 IoU 的指数加权移动平均 $\overline{L_{\text{IoU}}}$ 之间的相对幅度差异计算得到,用于动态衡量样本匹配质量的异常程度。

根据式(25)、(26),将动态增益权重项 R_{WIoU} 和 MPDIoU 相乘,得到 W^3 -MPDIoU 损失函数:

$$L_{W^3\text{-MPDIoU}} = R_{\text{WIoU}} L_{\text{MPDIoU}} \quad (28)$$

该损失函数将由最小点距离引导的几何约束与动态梯度调控机制相结合,使梯度分配始终与训练数据的实际质量相适应,有助于加速收敛并提升模型的整体性能。

3 实验结果与分析

3.1 数据集

本文使用的数据集 VisDrone2019-DET^[24] 是由天津大学 AISKEYEYE 团队从不同高度和视角收集的大量无人机航拍图像。此数据集采自无人机视角的城市与乡村区域,涵盖目标密集、环境遮挡以及不同光照条件的多种场景。

数据集 TinyPerson^[25] 专门针对背景杂乱、远景模糊等实际场景中的极小尺寸行人目标,行人目标主要分布在海上和陆地。

3.2 实验环境

实验平台为 Ubuntu 22.04 操作系统, CUP 为 Intel Xeon Platinum 8470Q, 显卡为 RTX 4090 (24 GB 显存)。环境配置: Python 版本为 3.10.18, Torch 版本为 2.8.0, CUDA 版本为 12.8。参数设置如表 1 所示。

表 1 实验参数设置

Table 1 Experimental parameter settings

参数	取值
学习率(lr0)	0.01
图像尺寸/Pixel(imgsz)	640×640
动量(momentum)	0.937
优化器(optimizer)	SGD
批次大小(batch)	16
线程数(workers)	8
训练轮数/Epochs	200

3.3 评价指标

本文采用多个计算机视觉中公认的评估算法性能的主要指标,包括准确率(precision, P)、召回率(recall, R)、平均精度(mean average precision, mAP)、浮点运算(GFLOPs)和参数量(Parameters)。

$$P = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (29)$$

$$R = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (30)$$

$$\text{AP} = \int_0^1 P(R) dR \quad (31)$$

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^N \text{AP}_i \quad (32)$$

其中, TP 表示被预测正确的正样本; FP 表示被预测错误的正样本; FN 表示被预测错误的负样本; AP 表示单类别的平均检测精度; mAP 表示数据集中所有类别的平均精度。

3.4 消融实验

为评估 LMH-YOLO 算法各创新模块的检测性能,本文在 VisDrone2019 无人机小目标和 TinyPerson 超小目标

数据集上进行消融实验,实验结果如表 2 和 3 所示。其中,A、B、C、D 分别表示 C3k2_PMDRB、LCFI_Net、LGDHead、W³-MPDIoU。

表 2 基于 VisDrone2019 无人机小目标消融实验
Table 2 Ablation experiments on small object detection using the VisDrone2019 dataset

方法	A	B	C	D	P/%	R/%	mAP@50/%	mAP@50:95/%	Params/(×10 ⁶)	GFLOPs	Size/MB
(1)					43.6	32.5	32.3	18.7	2.6	6.3	5.2
(2)	✓				42.4	33.0	32.5	18.7	2.5	6.2	5.0
(3)		✓			44.8	36.7	36.7	21.6	1.9	11.1	4.1
(4)			✓		43.4	31.7	31.9	18.6	2.3	5.2	4.8
(5)				✓	45.2	34.2	34.1	19.3	2.6	6.3	5.2
(6)	✓	✓			47.2	35.0	35.8	21.1	1.7	11.0	3.9
(7)	✓	✓	✓		47.4	35.7	36.4	21.6	1.5	7.0	3.5
(8)	✓	✓	✓	✓	48.5	36.4	37.1	21.9	1.5	7.0	3.5

表 3 基于 TinyPerson 无人机超小目标消融实验
Table 3 Ablation experiments on UAV tiny object detection using the TinyPerson dataset

方法	A	B	C	D	P/%	R/%	mAP@50/%	mAP@50:95/%	Params/(×10 ⁶)	GFLOPs	Size/MB
(1)					37.8	21.8	19.4	7.4	2.6	6.3	5.2
(2)		✓			41.1	30.0	26.8	10.8	1.9	11.1	4.2
(3)	✓	✓			41.7	29.1	26.8	11.1	1.7	10.9	4.0
(4)	✓	✓	✓		42.6	28.5	25.9	10.0	1.5	6.9	3.6
(5)	✓	✓	✓	✓	42.2	28.7	26.1	10.2	1.5	6.9	3.6

由表 2 可知,方法(2)在降低参数量和模型大小的同时,检测精度也获得提升,表明 C3k2_PMDRB 能够在保持轻量化前提下强化小目标特征表示;方法(3)带来最显著的检测性能提升:mAP@50 从 32.3%提升到 36.7%,mAP@50:95 从 18.7%提升到 21.6%,尽管 GFLOPs 增加到 11.1,但参数量下降至 1.9 M,这表明 LCFI_Net 显著提升了模型特征交互效率;引入新的检测头 LGDHead 虽然使模型性能略有变化,但有效降低了模型的参数量与计算开销,使模型更加轻量化;方法(5)实现了 mAP 的提升,并未增加计算开销,表明 W³-MPDIoU 损失能有效提升模型对小目标的定位能力;与单独使用 B 相比,A+B 虽在 mAP@50 上略有下降,但模型的参数量和计算量均有所降低,这表明 C3k2_PMDRB 有效优化了跨层级交互模块的资源开销;A+B+C 的 mAP@50 进一步提升至 36.4%,同时计算量从 11.0 GFLOPs 降至 7.0 GFLOPs,模型大小降至 3.5 MB,说明轻量化检测头有效降低计算成本,并显著提升模型整体效率;方法(8)在计算开销不增加的前提下,进一步增强了模型的定位与回归质量。改进算法的 mAP@50 和 mAP@50:95 分别提升至 37.1%和 21.9%,同时参数量和模型大小分别下降了 42.3%和 32.7%,实现了性能与效率的综合最优。

无人机小目标消融实验结果表明,本文所提出的 LMH-YOLO 算法在 VisDrone 数据集的密集小目标检测任务上表现出显著的模块协同增益和高性能。

为进一步验证改进算法在检测小目标任务上的有效性,本文在超小目标数据集 TinyPerson 上进行了消融实验。方法(2)增强了模型对极小目标的空间位置信息提取与上下文信息捕获能力,使 mAP@50 提升至 26.8%,mAP@50:95 提升至 10.8%,表明特征交互模块对 TinyPerson 极小目标场景尤为关键;组合 A 与 B 后,模型性能得到进一步提升,mAP@50:95 提升至 11.1%,同时参数量和运算量有效降低,这体现了浅层结构优化与跨层特征交互模块之间的协同作用;在 A+B 的基础上引入轻量化检测头 LGDHead,尽管检测性能略有下降,但计算量从 10.9 GFLOPs 大幅降至 6.9 GFLOPs,模型大小降至 3.6 MB,极大地提升了模型的部署效益;最终,改进模型在保持 1.5×10⁶ 的极少参数量和 3.6 MB 的轻量模型体积的前提下,mAP@50 和 mAP@50:95 分别达到 26.1%和 10.2%,实现了性能与轻量化的良好平衡。

在 Visdrone2019 和 TinyPerson 数据集上进行的消融实验结果表明,LMH-YOLO 在小目标及超小目标检测任务中表现出优异性能,并在检测精度和模型轻量化之间实现了良好平衡。

3.5 对比实验

1)C3k2_PMDRB 部分比例超参数 r 对比实验
为评估 C3k2_PMDRB 中部分比例超参数 r 对基线模型性能的影响,本文对 r 进行系统对比实验,其结果如表 4 所示。

表 4 部分比例超参数 r 对比实验

Table 4 Comparison experiments of the partial ratio hyperparameter r

r	$P/\%$	$R/\%$	mAP@50/%	mAP@50:95/%	Params/($\times 10^6$)	GFLOPs	Size/MB
1	42.8	32.7	32.1	18.5	2.68	6.5	5.5
2	42.8	32.9	32.3	18.6	2.49	6.3	5.1
4	42.4	33.0	32.5	18.7	2.46	6.2	5.0
8	42.7	31.6	31.8	18.6	2.43	6.2	5.0
16	42.0	32.5	31.5	18.3	2.41	6.1	5.0

随着 r 的逐渐增大,模型的参数量与 GFLOPs 呈持续下降趋势,呈现良好的轻量化趋势。但过大的 r 会削弱有效特征通道的表征能力,导致模型对小目标的建模能力不足,检测性能会随之下降。结合综合评价指标可见,当 $r=4$ 时,模型的检测性能最佳;模型的参数量下降 8.2%,GFLOPs 降低 4.6%,模型大小缩减 9.1%,mAP@50 达到 32.5%,实现了精度与轻量化之间的最佳折中。

2)分组卷积超参数 G 对比实验

为了在检测精度和轻量化之间取得最佳平衡,本文对 LGDHead 中的分组卷积超参数 G 进行调参, G 分别取 4、

8、16、32 和 64 进行对比实验。

实验结果如表 5 所示,当 G 较小时(4、8),模型参数量和计算量显著下降,但检测精度略低于原始低于基线模型;随着 G 增大至 16 和 32,检测精度逐步提升,且模型的运算量低于基线模型;进一步 G 增大至 64,虽然检测精度高于基线模型,但模型的参数量和运算量也显著增加,不利于模型轻量化。综合对比 $G=16$ 和 $G=32$,在 mAP 接近的前提下, $G=16$ 的 FLOPs 更低。因此,综合考虑精度和轻量化效果,选取 $G=16$ 能够在保证模型性能的同时控制复杂度,实现精度与轻量化之间的较优折中。

表 5 分组卷积超参数 G 对比实验

Table 5 Comparison experiments of the group convolution hyperparameter G

G	$P/\%$	$R/\%$	mAP@50/%	mAP@50:95/%	Params/($\times 10^6$)	GFLOPs
YOLO11	43.6	32.5	32.3	18.7	2.60	6.3
4	42.5	32.1	31.5	18.2	2.25	4.9
8	42.3	32.1	31.8	18.5	2.28	5.0
16	43.4	31.7	31.9	18.6	2.34	5.2
32	43.6	32.2	32.1	18.7	2.47	5.7
64	45.2	32.4	32.6	18.9	2.73	6.5

3)轻量化检测头对比实验

为研究不同轻量化检测头对无人机小目标检测性能

的影响,本文在基线模型 YOLOv11n 上进行对比实验,结果如表 6 所示。

表 6 不同检测头对比实验

Table 6 Comparison experiments with different detection heads

检测头	$P/\%$	mAP@50:95/%	Params/($\times 10^6$)	GFLOPs	Size/MB
Detect(baseline)	43.6	18.7	2.6	6.3	5.2
Detect_LSCD	42.4	19.1	2.3	6.0	4.9
Detect_Seam	41.9	18.1	2.5	5.8	5.1
Detect_Efficient	42.2	18.6	2.3	5.1	4.7
Detect_LADH	41.8	17.9	2.2	5.0	4.6
LGDHead	43.7	18.6	2.3	5.2	4.8

Detect_LSCD(detect lightweight shared convolutional detection)^[26] 的准确率低于 LGDHead,同时计算开销高于 LGDHead;Detect_Seam(detect seamless attention enhanced)^[27] 的准确率、mAP@50:95 指标上均低于 LGDHead,且计算

开销较高;Detect_Efficient^[28] 与 LGDHead 的计算开销相近,但 LGDHead 在准确率上优于 Efficient;Detect_LADH 虽计算开销低于 LGDHead,但准确率和 mAP@50:95 均低于 LGDHead,表明 LGDHead 在保持轻量化优势的同时

实现了更优的检测性能。

与基线模型相比,LGDHead 仅以 2.3×10^6 参数量与 5.2 GFLOPs 的计算开销,使 mAP@50:95 提升至 18.6%。显著提高了模型的推理效率,实现了模型复杂度与性能之间的最佳平衡。

4)W^{v3}-MPDIoU 对比实验

本文旨在评估 W^{v3}-MPDIoU 中不同权重超参数在基线模型中的影响,并确定能够提升小目标检测性能的最优参数组合,以进行对比实验。实验结果如表 7 所示。

表 7 W^{v3}-MPDIoU 不同权重超参数对比实验

Table 7 Comparison experiments with different weight hyperparameters of W^{v3}-MPDIoU

α	δ	$P/\%$	mAP@50/%	mAP@50:95/%
2.0	2.7	42.9	31.5	18.3
1.5	2.7	42.9	32.5	18.5
1.5	3.0	44.5	32.6	18.6
1.7	2.0	43.9	32.3	18.7
1.7	2.7	45.2	34.1	19.3

由式(26)可知, β 表示当前锚框的离群程度; δ 表示非单调梯度曲线中的性点(归一化点),即 $R_{wIoU}=1$ 时的 β 值; α 用于调节梯度放大的幅度。 $\alpha=2,\delta=2.7$ 参数组合对高质量框的放大最强,但当对低质量锚框抑制过于激进,导致模型训练稳健性下降及小目标检测性能最低; $\alpha=1.5,\delta=2.7$; $\alpha=1.5,\delta=3.0$ 以及 $\alpha=1.7,\delta=2.0$ 这三组参数组合(图 11 中非实线)对高质量框的梯度增益放大作用较弱,模型的收敛速度难以加快,性能均不及 $\alpha=1.7,\delta=2.7$ 参数组合; $\alpha=1.7,\delta=2.7$ 参数组合(图 11 中黑色实线)兼顾高质量预测框的梯度增益放大与低质量预测框的平缓抑制,在准确率和 mAP 上均取得最佳性能。实验表明, $\alpha=1.7,\delta=2.7$ 的 W^{v3}-MPDIoU 可在保持计算开销的前提下,有效加速模型收敛并提升模型检测性能。

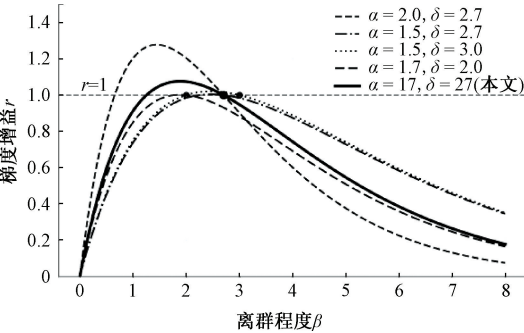


图 11 W^{v3}-MPDIoU 不同参数组合对比

Fig. 11 Comparison of different parameter combinations of W^{v3}-MPDIoU

5)损失函数对比实验

为研究不同定位回归损失函数对无人机小目标检测

的性能影响,本文以 YOLOv11 为基线模型,将 CIoU^[29]、SIoU^[30]、Inner-IoU^[31]、Focaler-IoU^[32]、Wise-IoU、MPDIoU 以及 W^{v3}-MPDIoU 依次引至基线模型中进行验证。根据表 8 和图 12 的实验结果,相较于原始的 CIoU,W^{v3}-MPDIoU 在准确率、mAP@50 和 mAP@50:95 上分别提升了 1.6%、1.8%和 0.6%,在无人机小目标检测任务中取得最佳检测性能。

表 8 损失函数对比实验

Table 8 Comparison experiments of different loss functions

损失函数	$P/\%$	mAP@50/%	mAP@50:95/%
CIoU	43.6	32.3	18.7
SIoU	43.9	32.8	18.8
Inner-IoU	42.4	31.6	18.2
Focaler-IoU	42.6	32.5	18.7
Wise-IoU	43.7	32.1	18.2
MPDIoU	41.7	31.5	18.2
W ^{v3} -MPDIoU	45.2	34.1	19.3

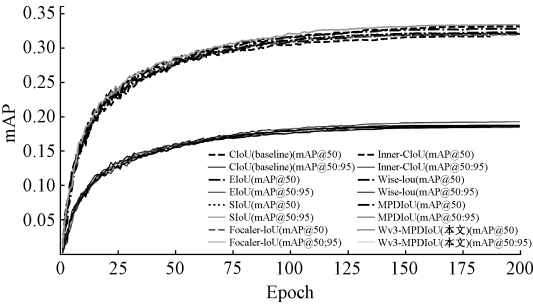


图 12 损失函数 mAP 对比

Fig. 12 mAP comparison of different loss functions

6)经典算法对比实验

为了验证改进算法 LMH-YOLO 在 VisDrone 数据集上的检测性能和轻量化优势,本文选取当前目标检测领域内的具有代表性的主流算法开展横向对比实验,如表 9 所示。

为全面评估 LMH-YOLO 的性能,本文选取经典检测器 Faster-RCNN、SSD 及 YOLO 系列模型进行对比实验。双阶段检测算法 Faster-RCNN 在 mAP@50 指标上取得较高的检测精度,但其 41.7×10^6 参数量和 215.6 GFLOPs 的计算开销限制了其在资源受限设备上的部署。SSD 虽将参数量和计算量分别降至 25.7×10^6 和 88.8 GFLOPs,但其检测精度显著下降,难以满足小目标密集场景的检测需求。YOLO 系列基础模型在性能与模型复杂度之间呈现出持续优化的趋势。相比于主流的 YOLOv8n 和 YOLOv12n,LMH-YOLO 将参数量降低至 1.5×10^6 ,实现模型的轻量化,同时在准确率、召回率、mAP@50、mAP@50:95 等评价指标上取得最优表现。与基准模型 YOLOv11n

表 9 经典算法对比实验

Table 9 Comparison experiments with classical algorithms

Model	<i>P</i> /%	<i>R</i> /%	mAP@50/%	mAP@50:95/%	Params/($\times 10^6$)	GFLOPs
Faster-RCNN	—	—	35.8	20.8	41.7	215.6
SSD	—	—	25.2	12.7	25.7	88.8
Yolov5n	42.7	31.2	31.2	17.7	2.2	5.8
Yolov6n	41.3	29.4	29.2	16.8	4.2	11.5
Yolov8n	43.6	32.6	31.9	18.3	2.7	6.8
Yolov9t	43.8	33.1	33.2	19.1	2.0	7.6
Yolov10n	42.9	32.4	32.7	18.9	2.3	6.5
Yolov11n	43.6	32.5	32.3	18.7	2.6	6.3
Yolov12n	43.9	33.2	32.6	18.7	2.6	6.3
LMH-YOLO(本文)	48.5	36.4	37.1	21.9	1.5	7.0

相比,LMH-YOLO 的准确率、召回率、mAP@50 和 mAP@50:95 分别提高了 4.9%、3.9%、4.8%和 3.2%,同时参数量下降了 42.3%。各主流模型的实验结果表明,LMH-YOLO 在 1.5×10^6 的最小参数量和 7.0 GFLOPs 的适中计算开销下,准确率、召回率和 mAP 均取得最佳性能,在轻量化与计算效率之间取得最佳平衡,验证了其模型轻量化设计与多层次特征融合策略的有效性,进一步提升了小目标检测性能。

7)改进算法对比实验

为进一步验证 LMH-YOLO 的有效性,本文选取了近期针对无人机航拍图像目标检测的方法在 VisDrone 数据集上进行对比实验,包括基于 Transformer 架构的 MGEF-DETR^[33] 以及多种基于 YOLOv11 改进的算法模型(文献[34]、RSC-YOLO^[35] 和 PC-YOLO^[36])。实验结果如表 10 所示。

表 10 改进模型在 VisDrone 数据集上的对比实验

Table 10 Comparative experiments of the improved model on the VisDrone dataset

模型	<i>P</i> /%	<i>R</i> /%	mAP@50/%	mAP@50:95/%	Params/($\times 10^6$)	GFLOPs
MGEF-DETR ^[33]	63.6	49.8	51.3	31.9	17.20	68.2
文献[34]	45.1	34.8	34.5	—	3.18	—
RSC-YOLO ^[35]	47.4	—	37.7	—	2.70	8.7
PC-YOLO ^[36]	46.8	35.4	36.1	21.5	1.90	—
LMH-YOLO	48.5	36.4	37.1	21.9	1.50	7.0

从实验结果可以看出,MGEF-DETR 在检测精度上表现优异,mAP@50 达 51.3%,mAP@50:95 达 31.9%,但计算开销偏高,难以满足无人机等边缘设备的实时部署需求。RSC-YOLO 侧重浅层细节与跨尺度融合,在 mAP@50 达到 37.7%,但其在复杂度和精度权衡方面仍存在优化空间;文献[34]与 PC-YOLO 通过引入注意力机制或调整检测层结构,在检测精度与参数规模之间取得一定平衡,但其整体性能仍有提升空间。

相比之下,LMH-YOLO 在保持较低参数量和计算开销的同时,在准确率和 mAP 指标上也表现出一定优势,在检测精度与模型复杂度之间实现了更优的综合平衡,更适合用于无人机航拍场景下的实时小目标检测。

3.6 实验可视化

为了更直观地验证 LMH-YOLO 在复杂场景中的检测效果,本文在 VisDrone 数据集中选取典型复杂场景进行热力图可视化检测对比,如图 13 所示。

在小目标密集的交通流场景中,YOLOv8 和 YOLOv10 热力图呈现分散且模糊的响应,对单个目标的热力集中度较低;LMH-YOLO 则保持较高的特征集中,能够精确聚焦于每个单独的小目标中心,表明其具有更卓越的小目标特征提取和抗干扰能力。在目标尺度差异显著的场景中,YOLOv8 在远端小目标的热力图强度较低,以至于部分小目标未能被准确感知;LMH-YOLO 在近景大目标和远景小目标保持显著的热力强度,表明改进算法采用的多尺度特征融合策略显著有效,能够提升不同尺寸目标的识别能力。在树木遮挡场景中,LMH-YOLO 能够对可见区域实现精准聚焦,并能对被遮挡部分进行有效检测,表明改进算法增强了模型对遮挡目标特征的鲁棒性。在昏暗夜晚场景中,LMH-YOLO 对低对比度中的目标具有更强的感知与识别能力。在反光场景中,LMH-YOLO 的特征提取器具有更强的纹理和语义分辨能力,能够精确提取高强度光照条件下的小目标语义信息。

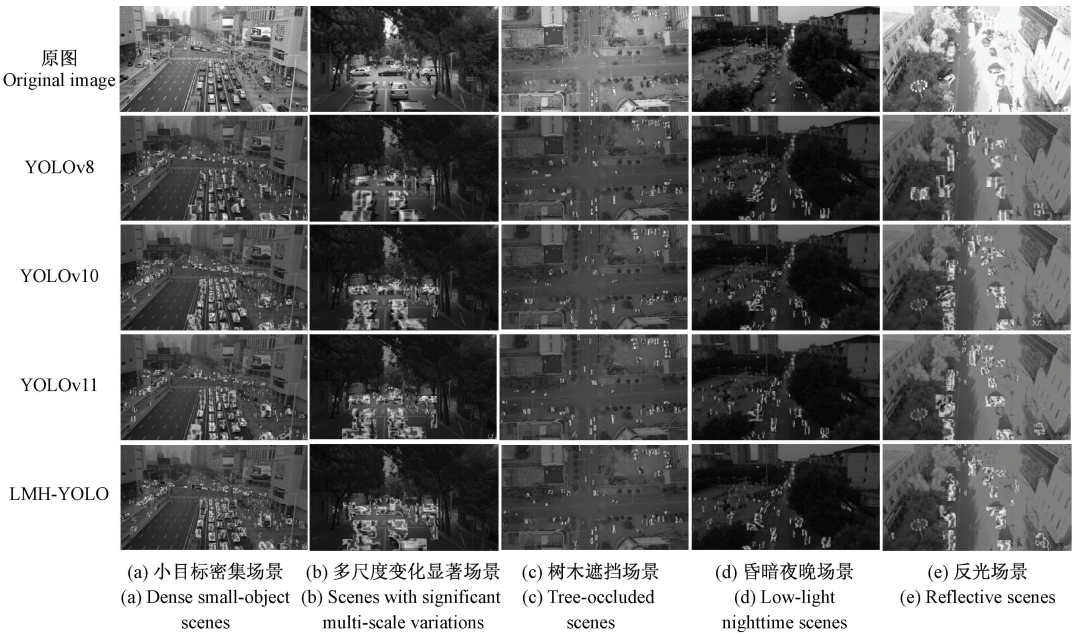


图 13 热力图结果对比

Fig. 13 Comparison of heatmap results

4 结 论

为解决无人机航拍图像中目标检测的精度与计算资源难以兼顾的问题,本文提出了一种基于多尺度交互融合的轻量化改进算法模型 LMH-YOLO。

LMH-YOLO 从多尺度特征表达、跨层级交互融合、轻量化检测头和定位优化损失 4 个方面,对小目标检测进行了系统性改进。通过 C3k2_PMDRB 增强多尺度表征能力,设计 LCFI_Net 以提升跨层特征融合效率并降低模型复杂度,采用轻量化检测头进一步降低模型参数量与计算开销,并构建 W^3 -MPDIoU 损失函数以加速模型收敛并提升精度。实验结果表明,LMH-YOLO 在小目标及超小目标检测任务中取得显著优势,在实现模型轻量化的同时显著提升了检测精度。

未来将致力于将模型部署于无人机等资源受限的平台,以发挥在无人机航拍检测任务中的实用价值。

参考文献

[1] 罗旭东,吴一全,陈金林. 无人机航拍影像目标检测与语义分割的深度学习研究方法研究进展[J]. 航空学报, 2024,45(6):241-270.
LUO X D, WU Y Q, CHEN J L. Research progress on deep learning methods for object detection and semantic segmentation in UAV aerial images[J]. Acta Aeronautica et Astronautica Sinica, 2024, 45(6): 241-270.

[2] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]//European

Conference on Computer Vision, 2016: 21-37.

[3] 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.

[4] REN S 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, 2016, 39(6): 1137-1149.

[5] 田红鹏,李志强,杨赛. 改进 RT-DETR 的航拍图像小目标检测算法[J/OL]. 计算机工程,1-14[2025-11-19]. <https://doi.org/10.19678/j.issn.1000-3428.0252661>.
TIAN H P, LI ZH Q, YANG S. An improved algorithm for small object detection in UAV aerial images based on RT-DETR [J/OL]. Computer Engineering, 1-14 [2025-11-19]. <https://doi.org/10.19678/j.issn.1000-3428.0252661>.

[6] 张永宏,许鑫豪,尹贺峰,等. MFSF-DETR:一种基于多尺度特征移位融合的 PCB 缺陷与元件检测算法[J]. 仪器仪表学报,2025,46(8):266-285.
ZHANG Y H, XU X H, YIN H F, et al. MFSF-DETR: A PCB defect and component detection algorithm based on multi-scale feature shift fusion[J]. Chinese Journal of Scientific Instrument, 2025, 46(8): 266-285.

[7] LIU S Z, CAO L H, LI Y. Lightweight pedestrian detection network for UAV remote sensing images

- based on strideless pooling [J]. Remote Sensing, 2024, 16(13): 2331.
- [8] 李泽胤,李栋,房建东,等.改进的YOLOv8n遥感图像轻量化检测模型[J].电子测量技术,2025,48(6): 130-142.
- LI Z Y, LI D, FANG J D, et al. An improved YOLOv8n lightweight detection model for remote sensing images [J]. Electronic Measurement Technology, 2025, 48(6): 130-142.
- [9] 贾亮,陈茂辉,王琪,等.改进YOLOv8n的航拍多尺度目标检测模型[J/OL].电子测量技术,1-10[2026-01-07]. <https://link.cnki.net/urlid/11.2175.TN.20250828.1602.004>.
- JIA L, CHEN M H, WANG Q, et al. An improved multi-scale target detection model for aerial photography based on YOLOv8n[J/OL]. Electronic Measurement Technology, 1-10 [2026-01-07]. <https://link.cnki.net/urlid/11.2175.TN.20250828.1602.004>.
- [10] 张彪,荀荣科,许家忠.基于深度学习的工业轴承缺陷检测算法研究[J].仪器仪表学报,2025,46(4): 136-149.
- ZHANG B, XUN R K, XU J ZH. Research on industrial bearing defect detection algorithm based on deep learning [J]. Chinese Journal of Scientific Instrument, 2025, 46(4): 136-149.
- [11] 冉宁,施高朗,张昊昊,等.基于YOLO11的轻量级遥感图像检测算法[J/OL].电子测量与仪器学报,1-12[2026-01-07]. <https://link.cnki.net/urlid/11.2488.tn.20251224.1643.074>.
- RAN N, SHI G L, ZHANG H Y, et al. A lightweight remote sensing image detection algorithm based on YOLO11 [J/OL]. Journal of Electronic Measurement and Instrumentation, 1-12 [2026-01-07]. <https://link.cnki.net/urlid/11.2488.tn.20251224.1643.074>.
- [12] 崔克彬,吕思怡,杨立然.基于HCDNet-YOLO11的输电线路关键部位及缺陷检测算法[J/OL].电子测量技术,1-13[2026-01-07]. <https://link.cnki.net/urlid/11.2175.TN.20251210.1842.010>.
- CUI K B, LYU S Y, YANG L R. Key parts and defect detection of transmission line based on HCDNet-YOLO11 [J/OL]. Electronic Measurement Technology, 1-13 [2026-01-07]. <https://link.cnki.net/urlid/11.2175.TN.20251210.1842.010>.
- [13] 张红民,冯雪,陈清,等.SDA-YOLO;改进YOLOv11的无人机航拍图像小目标检测方法[J/OL].计算机工程与应用,1-13[2025-11-19]. <https://link.cnki.net/urlid/11.2127.tp.20251024.1429.014>.
- ZHANG H M, FENG X, CHEN Q, et al. SDA-YOLO: An improved YOLOv11-based method for small object detection in UAV images [J/OL]. Computer Engineering and Applications, 1-13 [2025-11-19]. <https://link.cnki.net/urlid/11.2127.tp.20251024.1429.014>.
- [14] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation [C]//IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [15] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]//IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117-2125.
- [16] CHEN J R, KAO S, HE H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.
- [17] RAHMAN M M, MUNIR M, MARCULESCU R. Emcad: Efficient multi-scale convolutional attention decoding for medical image segmentation[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 11769-11779.
- [18] LI L P, YI J ZH, FAN H, et al. A lightweight semantic segmentation network based on self-attention mechanism and state space model for efficient urban scene segmentation [J]. IEEE Transactions on Geoscience and Remote Sensing, 2025, 63: 4703215.
- [19] FAN ZH P, WU Y Y, LIU W, et al. LG-YOLOv8: A lightweight safety helmet detection algorithm combined with feature enhancement [J]. Applied Sciences, 2024, 14(22): 10141.
- [20] HUANG G, LIU S CH, VAN DER MAATEN L, et al. CondenseNet: An efficient DenseNet using learned group convolutions [C]//IEEE Conference on Computer Vision and Pattern Recognition, 2018: 2752-2761.
- [21] CHOLLET F. Xception: Deep learning with depthwise separable convolutions[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1251-1258.
- [22] MA S L, XU Y. MPDIoU: A loss for efficient and accurate bounding box regression[J]. ArXiv preprint arXiv:2307.07662, 2023.
- [23] TONG Z J, CHEN Y H, XU Z W, et al. Wise-IoU: Bounding box regression loss with dynamic focusing mechanism[J]. ArXiv preprint arXiv:2301.10051, 2023.
- [24] 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.
- [25] YU X H, HAN Z J, GONG Y Q, et al. The 1st tiny object detection challenge: Methods and results [C]//European Conference on Computer Vision, 2020: 315-323.
- [26] NING S T, TAN F, CHEN X, et al. Lightweight corn leaf detection and counting using improved YOLOv8[J]. Sensors, 2024, 24(16): 5279.
- [27] SUN Y, GUO H R, CHEN X A, et al. YOLOv8n-SSDW: A lightweight and accurate model for barnyard grass detection in fields [J]. Agriculture, 2025, 15(14): 1510.
- [28] 张政银, 向玮, 刘子锋, 等. 基于树皮纹理的轻量化 YOLOv11 树种识别方法[J]. 北京林业大学学报, 2025, 47(8): 134-148.
- ZHANG ZH Y, XIANG W, LIU Z F, et al. Lightweight YOLOv11 tree species recognition method based on bark texture[J]. Journal of Beijing Forestry University, 2025, 47(8): 134-148.
- [29] ZHENG Z H, WANG P, REN D W, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation[J]. IEEE Transactions on Cybernetics, 2021, 52(8): 8574-8586.
- [30] GEVORGYAN Z. SiOv loss: More powerful learning for bounding box regression [J]. ArXiv preprint arXiv:2205.12740, 2022.
- [31] ZHANG H, XU C, ZHANG S J. Inner-IoU: More effective intersection over union loss with auxiliary bounding box [J]. ArXiv preprint arXiv: 2311.02877, 2023.
- [32] ZHANG Y F, REN W Q, ZHANG Z, et al. Focal and efficient IoU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146-157.
- [33] 侯林杰, 卢承方, 崔艳荣. MGEF-DETR: 多尺度门控增强融合的无人机目标检测算法[J/OL]. 电子测量技术, 1-17 [2026-01-10]. <https://link.cnki.net/urlid/11.2175.tn.20251117.1700.002>.
- HOU L J, LU CH F, CUI Y R. MGEF-DETR: Multi-scale gated enhancement fusion for UAV object detection algorithm[J/OL]. Electronic Measurement Technology, 1-17 [2026-01-10]. <https://link.cnki.net/urlid/11.2175.tn.20251117.1700.002>.
- [34] 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 ZH J. Small object detection model based on YOLOv11 [J]. Laser Journal, 2025, 46(11): 49-56.
- [35] 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, ZHANG Y SH, WU N, et al. UAV object detection algorithm based on RSC-YOLO[J]. Infrared Technology, 2025, 47(12): 1491-1501.
- [36] WANG Z, SU Y T, KANG F, et al. PC-YOLO11s: A lightweight and effective feature extraction method for small target image detection[J]. Sensors, 2025, 25(2): 348.

作者简介

郝江鹏, 硕士研究生, 主要研究方向为机器视觉、目标检测。

E-mail: 15532063758@163.com

牛芳琳(通信作者), 博士, 教授, 主要研究方向为机器视觉、图像处理。

E-mail: dx_niufli@lnut.edu.cn

于玲, 博士, 副教授, 主要研究方向为图像识别。

E-mail: yl_lg@163.com

韩驰, 硕士研究生, 主要研究方向为目标检测。

E-mail: 1744309707@qq.com