

MEM-YOLO11:一种多特征选择机制的无人机小目标检测模型

徐世周,杨红,张梦洁,张钰昊

(河南师范大学 光电工程学院,河南 新乡 453007)

摘要:针对遥感图像目标具有多尺度、多旋转角度、背景复杂等特点,提出了一种多特征选择机制融合改进模型 MEM-YOLO11.首先,在 backbone 中设计了一种特征提取模块,结合了多尺度特征提取和边缘信息增强,提高了对特征信息的整合效果.其次,设计了一种高效的卷积网络模块,由原始图像生成更高分辨率的图像,增强了颈部网络中全局到局部的特征学习能力.并且,重新设计了一个多头部特征提取模块,减少标准卷积中广泛存在的空间和信道冗余,减少了参数量和计算量.最后,优化 MEM-YOLO11 模型的超参数,缓解分类和定位不相关问题.在 Vis-Drone2019 数据集上精度、召回率、 $mAP50$ 分别提升 7.6%、5.6%、6.5%; Wider Person 数据集上精度、召回率、 $mAP50$ 分别提升 6.8%、4.2%、5.6%.

关键词:小目标检测;多尺度特征提取;边缘增强;超参数

中图分类号:TP391.41

文献标志码:A

文章编号:1000-2367(2026)04-0091-07

随着观测技术的不断进步,遥感技术能够捕捉到大量细节,在重要领域得到了广泛的应用,包括军事侦察、城市规划、环境监测、土壤科学和海上船舶监视等^[1].基于卷积神经网络的深度学习技术通过对大量数据的训练,能够实现高精度的识别和定位拍摄图像中的各种物体,已成为物体检测的主流方法.基于深度学习的目标检测模型算法主要分为单阶段算法和两阶段算法.两阶段算法^[2]推理精度高,面临着模型复杂度高的问题.单阶段算法^[3]通过直接回归生成检测框,精度稍低,但具有更高的推理速度和帧率,更适合部署在资源受限的无人机平台.

1 相关工作

无人机遥感拍摄的小目标图像分辨率有限、尺度变化大且背景复杂^[4],传统的卷积神经网络在处理这些小目标时往往表现不佳,各种因素会导致模型的准确率下降,影响检测的实时性和可靠性.并且受硬件平台限制,面临着漏检、误检以及准确率低的问题.Li 等^[5]提出了一种基于增强通道空间融合的小目标检测网络,充分利用浅层特征中丰富的空间细节信息来定位小目标的位置,但是增加了模型的复杂度.Yi^[6]等提出了一种基于 YOLOv7 的网络,通过融合注意力机制、增强 FPN 网络拓扑结构等方式,集中小目标更关键的特征信息,然而复杂的结构导致模型过拟合训练数据,泛化能力不足.Guan 等^[7]基于 YOLOv10 引入双向特征金字塔网络等模块,使模型能够更准确地处理复杂信息,然而泛化能力较差.

受现有工作启发,本文提出了一种多特征选择机制融合改进模型 MEM-YOLO11(multi-module enhanced

收稿日期:2025-05-12;**修回日期:**2025-06-05.

基金项目:国家自然科学基金(11904085);河南省科技攻关项目(202102210299).

作者简介:徐世周(1985—),男,河南南阳人,河南师范大学副教授,博士,研究方向为容错控制策略、三电平逆变器、调制策略、大功率三电平逆变器建模、控制与改进等,E-mail:xushizhousiee@163.com.

通信作者:杨红,E-mail:13781254305@163.com.

引用本文:徐世周,杨红,张梦洁,等.MEM-YOLO11:一种多特征选择机制的无人机小目标检测模型[J].河南师范大学学报(自然科学版),2026,54(4):91-97.(Xu Shizhou, Yang Hong, Zhang Mengjie, et al. MEM-YOLO11: A UAV small object detection model with multi-feature selection mechanism[J]. Journal of Henan Normal University (Natural Science Edition), 2026, 54(4): 91-97. DOI:10.16366/j.cnki.1000-2367.2025.05.12.0002.)

mechanism-YOLO11),集成了多尺度特征提取(MFEE)、高效特征学习(ELFR)和多头感受野增强(MRFE),旨在提升复杂场景下的检测精度与鲁棒性,主要贡献如下:

- ①设计了一种多尺度特征提取和边缘增强模块 MFEE(multi-scale feature extraction and edge enhancement),从不同尺度上提取特征,并将这些多尺度特征整合到一起,突出了小目标的边缘信息,增强了小目标的特征输出.
- ②设计了一种高效特征学习模块 ELFR(effective learning of feature representations),P2 特征层经过 SPDConv 切片操作,能够有效降低参数量,更大程度地保留小目标的关键信息,得到富含小目标信息的特征图.
- ③设计了一个多头部感受野增强模块 MRFE(multi-head receptive field enhancement),通过多视角特征融合增强模型的鲁棒性和泛化能力,促进了更大感受野内特征的聚合.
- ④调整 topK、 α 、 β 的阈值大小,缓解分类和定位不相关的问题.

2 提出的方法

2.1 MFEE 模块

融合 CSPNet^[8]思想,MFEE 模块结构如图 1 所示,结合了多尺度特征提取^[9]、边缘信息增强^[10]和卷积操作.首先经过自适应平均池化^[11](AdaptiveAvgPod)操作提取输入图片多尺度的信息,经过两个 C3K2 卷积进行多尺度特征间的语义互补,随后经过一个上采样模块增加特征图的分辨率,再经过 EdgeEnhancer 模块.EdgeEnhancer 模块用于提取边缘信息,增强网络对边缘信息的敏感度,架构如图 2 所示.二维平均池化^[12]操作(AvgPool2d,使用 3×3 核,步长为 1,padding=1)用于提取特征图的低频信息.将原始输入特征图与平滑后的特征图进行相减,得到增强的高频信息.其次利用 C3k2 模块,精细化高频特征并提升模型对小目标的表征能力.最后将提取到的信息与原始输入特征图相加,形成增强后的输出.

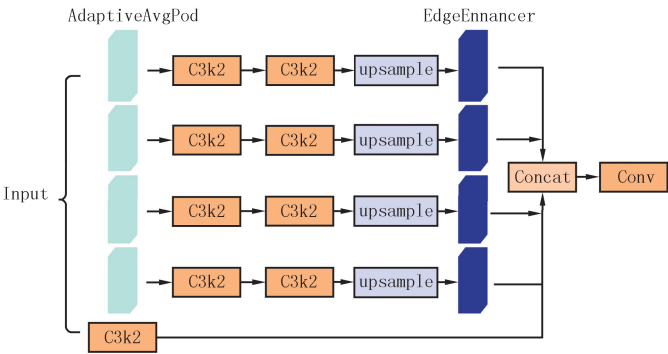


图1 MFEE模块结构
Fig.1 The MFEE module

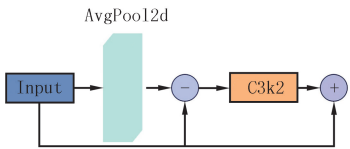


图2 EdgeEnhancer模块结构
Fig.2 EdgeEnhancer module

MFEE 模块有 4 个不同自适应平均池化参数的分支,对应不同尺度的特征图的信息.4 层设计覆盖了从全局到局部的完整特征层次,避免单一池化尺度导致的特征缺失.仅用大尺度池化会丢失细节,仅用小尺度池化缺乏语义引导,进行多尺度的池化,提取不同大小的局部信息,有助于捕捉图像的多层次特征.自适应平均池化模块会指定固定尺寸的输出(H,W).第 1 个分支经过平均池化后,把特征图的尺度缩放成 3×3,第 2 个分支缩放成 6×6,第 3 个分支缩放成 9×9,第 4 个分支缩放成 12×12.为了弥补第 1 个分支的特征图较小,增加了一个局部卷积,以达到信息交互的目的.将不同尺度下提取的特征通过插值操作对齐到同一尺度,然后将它们拼接在一起,最后经过卷积层融合成统一的特征表示输出,有效提高了模型对多尺度特征的感知.

2.2 ELFR 模块

小目标检测的困难根源是小目标覆盖的像素点少,导致漏检误检等问题.本文设计由原始图像生成更高分辨率的图像,对更高分辨率的图像进行目标检测.小目标在 P3、P4、P5 检测层上略显吃力,比较传统的做

法是加 P2 检测层来提升小目标的检测能力^[13].然而这种做法会导致计算量过大、后处理更加耗时等问题.这一矛盾促使研究者探索兼顾精度与效率的小目标优化特征金字塔结构.

本文根据 CSP 思想以及基于 Omni-Kernel^[14]改进,设计加入在 YOLOv11n 中引入改进的 P2 检测层,改进的金字塔架构如图 3 所示.为了有效降低参数量,改进的 P2 特征层经过 SPDCov 切片操作,更大程度地保留小目标的关键信息,然后与 P3 特征层进行融合,经过 ELFR 模块进行特征整合.

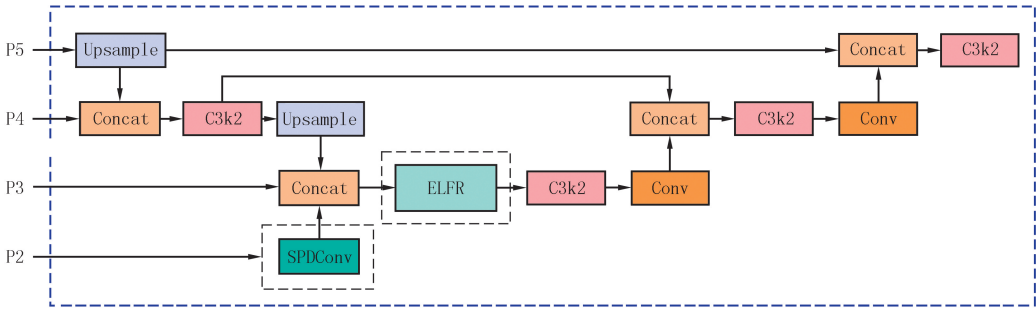


图3 改进的金字塔结构
Fig.3 The improved pyramid architecture

ELFR 模块结构如图 4 所示,先对输入特征图进行一个卷积操作,然后按照 1 : 3 的比例划分通道数,为了更大程度地保留原始信息,只有 25 %比例通道数经过 OmniKernel 模块,最后将两部分拼接在一起后进行一个卷积操作输出.OmniKernel 模块结构如图 5 所示.ELFR 模块增强了颈部网络中全局到局部的特征学习能力的同时有效解决了计算量过大、后处理耗时的问题.

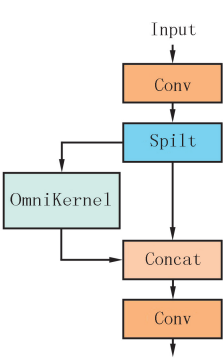


图4 ELFR模块结构
Fig.4 The ELFR structure

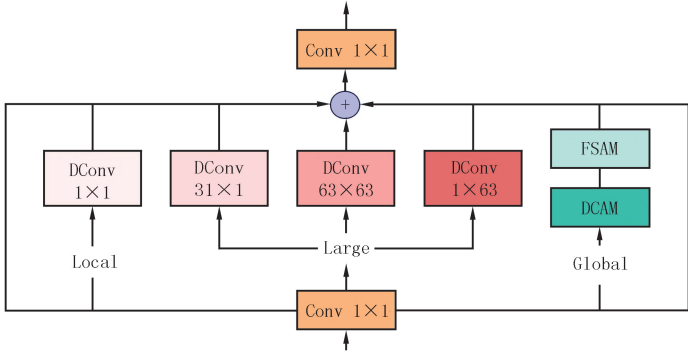


图5 Omni-Kernel模块结构
Fig.5 The Omni-Kernel structure

2.3 MRFE 模块

为了解决复杂背景下小目标分布重叠检测困难的问题,受到 YOLO-FaceV2^[15] 的启发,本文重新设计了一个多头部感受野增强模块 MRFE 来补偿被遮挡的信息的损失,MRFE 模块结构如图 6 所示.

SRU 模块用于抑制空间冗余,CRU 模块用于减少信道冗余.搭配使用有效限制了特征冗余,不仅减少了模型参数的数量和 FLOPs,而且提高了特征表示的能力.最大池化(mix pooling)相对于平均池化(average pooling),能够更大程度地保留输入图像的信息,在训练时提高模型的泛化能力.

MRFE 模块通过多视角特征融合增强了模型的鲁棒性和泛化能力,补偿了被遮挡信息的损失,进一步增强了多尺度特征融合,并且减少了冗余特征,显著降低了复杂度和计算成本,适用于部署在如无人机等边缘设备上.这个过程促进了更大感受野内特征的聚合,最终增强了模型多特征融合的能力.

2.4 Task Aligned Assigner 超参数调整

目标检测承担着分类和定位两个任务,由于分类和定位学习机制的不同,两个任务学习的特征的空间分布可能会不同,当同时使用两个独立的分支进行预测时,会导致一定程度的分类和定位失调^[16].正负样本标签分配的依据如下:

$$t = s^{\alpha} \times u^{\beta}, \tag{1}$$

其中, s 和 u 分别表示分类得分和 IoU, α 和 β 是权重系数. 调整 $\text{top}K$ 、 α 、 β 的阈值大小, 可以缓解分类和定位不相关的问题.

3 实验结果与分析

3.1 实验环境和细节

为了保证实验的公平性, 实验过程中没有使用预训练的模型. 实验中使用初始学习率为 0.01 的 SGD 优化器, 为达到较好的收敛效果并且避免内存溢出, 本文将动量设置为 0.937, 并使用 $5\text{e-}4$ 的权重衰减系数, $\text{epochs}=300$, $\text{batch}=8$, $\text{workers}=4$, 学习速率设置为 0.01, 采用 GPU 进行模型训练.

3.2 数据集介绍

为了保证实验数据的合理性, 本文选取了小目标检测领域具有代表性公开数据集 VisDrone2019 数据集^[17]和 The WiderPerson 数据集^[18]进行实验.

VisDrone2019 数据集共包含 8 629 张图像. 该数据集包含 10 个类别的日常场景: 行人(P1)、人群(P2)、自行车(B1)、汽车(B2)、面包车(C)、卡车(V)、三轮车(T1)、遮阳棚三轮车(T2)、公共汽车(A)和摩托车(B2). 在该数据集中, 各类分布的比例不平衡, 大多数图像包含小目标具有多尺度、多旋转角度、背景复杂等特点.

WiderPerson 数据集是比较拥挤场景的行人检测基准数据集, 其图像是从多种场景中选择的, 不再局限于交通场景. 行人遮挡严重、目标密集. 考虑到计算机算力局限, 经过合理筛选之后选取 7 000 幅图片, 其中 4 800 幅图片作为训练集, 1 200 幅图片作为验证集, 1 000 幅图片作为测试集.

3.3 评价指标

在本实验中, 采用精度(P)、召回率(R)、平均均值精度(mAP) 等性能指标来评估模型的检测能力. 精度(P) 和召回率(R) 是根据真阳性(TP)、假阳性(FP)、真阴性(TN) 和假阴性(FN) 计算:

$$P = \frac{TP}{TP + FP}, \tag{2}$$

$$R = \frac{TP}{TP + FN}. \tag{3}$$

通过在精度-召回率曲线(P - R)上绘制精度和召回率, 定义曲线下面积(AP), 并将平均精度(mAP) 计算为所有类别的 AP 平均值. 计算公式见下式:

$$AP = \int_0^1 P(R) dR, \tag{4}$$

$$mAP = \frac{1}{N} \sum AP. \tag{5}$$

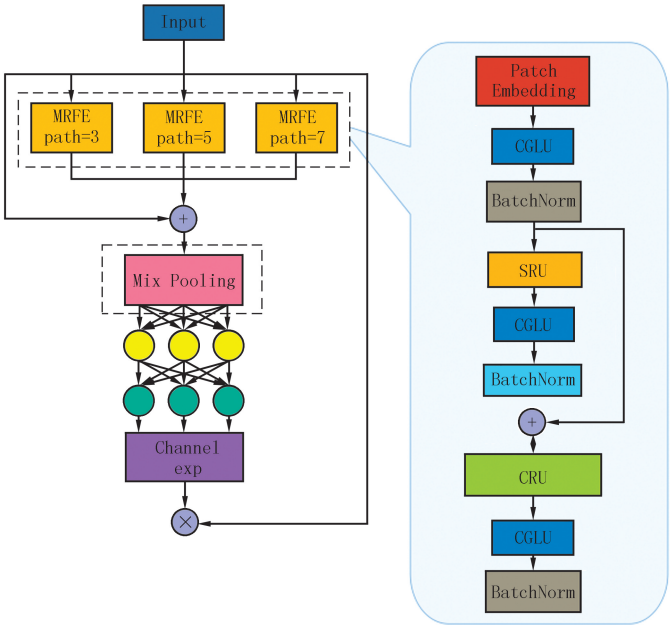


图6 MRFE 模块结构

Fig.6 The architecture of the MRFE module

3.4 实验结果和分析

基于 VisDrone2019 数据集,本文提出的 MEM-YOLO11 模型,精度(P)达到了 45.3%,相比 YOLOv11n 模型提高了 7.6%,召回率(R)提高了 5.6%, $mAP50$ 提高了 6.5%, $mAP50-95$ 提高了 4.4%,并且计算参数有所下降.

为了验证本文设计模块的有效性,基于 VisDrone2019 数据集进行了消融实验.附录表 S1 消融实验结果表明,同时加入 MFEE 模块、ELFR 模块、MRFE 模块的效果最好,显示了它们对提高模型整体检测性能的重大贡献.

本文调整了 MEM-YOLO11 模型的超参数,由附录表 S2 对比实验结果可得,本模型采用 $\text{topK}=7$ 、 $\alpha=0.5$ 、 $\beta=5$ 时,性能表现最佳.

本文对 MEM-YOLO11 模型和其他广泛使用的模型在 VisDrone2019 验证集上进行了对比试验,评估了数据集中 10 个类别的检测准确率($mAP50\%$).表 1 所示的实验结果表明,与其他流行的模型相比,MEM-YOLO11 模型具有优异的检测性能,特别是对于行人、人、自行车、面包车、卡车、三轮车和公共汽车对象检测效果最佳.

表 1 不同模型在 VisDrone2019 数据集上的检测精度对比试验
Tab. 1 Performance comparison of different models on the VisDrone2019 dataset %

方法	all	P1	P2	B1	C	V	T1	T2	A	B2	M
Faster R-CNN	21.5	19.3	14.6	5.8	51.2	32.1	20.9	13.5	6.8	30.7	20.8
Cascade R-CNN	22.2	19.5	13.5	7.6	52.6	31.5	20.6	14.8	8.6	32.9	20.4
RetinaNet	13.4	14.1	8.3	1.5	40.7	19.2	12.5	6.9	5.7	15.4	10.6
SSD300	20.0	13.1	11.3	9.1	43.6	25.2	26.1	14.2	6.7	37.9	12.6
RSF-SSD	27.4	14.1	13.0	10.5	56.8	41	32.1	13.5	20.1	51.5	17.9
CenterNet	26.0	28.0	11.6	9.0	51.0	36.2	27.9	20.1	19.9	37.7	21.0
CornerNet	17.7	20.0	6.6	4.6	40.9	20.2	20.5	14.0	9.3	24.4	12.1
YOLOv3-tiny	23.7	17.8	16.6	3.1	49.2	11.9	9.8	7.9	3.5	11.6	17.6
YOLOv5n	22.0	30.1	22.6	2.9	65.7	15.2	14.4	8.2	6.3	31.9	22.4
YOLOv5s	31.3	39.5	30.1	9.3	70.1	34.2	22.3	17.2	9.3	40.3	40.5
YOLOv6n	25.5	27.3	19.7	3.1	66.9	32.5	20.5	15.1	7.1	35.4	27.6
YOLOx-n	30.7	33.9	29.6	7.9	71.1	36.2	24.7	18.1	10.2	38.2	37.4
YOLOv7-tiny	31.7	38.1	36.3	8.1	75.1	34.2	29.6	18.7	8.6	41.2	27.1
YOLOv8n	27.0	22.5	12.6	6.16	65.4	29.9	33.0	13.0	15.0	50.2	22.1
YOLOv9-t	31.1	29.8	24.5	8.1	73.8	35.4	22.6	20.8	11.7	45.1	40.8
YOLOv10n	27.3	24.7	14.1	7.2	67.3	30.0	29.0	12.0	13.5	48.8	26.4
YOLO11n	27.7	24.1	12.5	6.9	68.1	30.2	31.1	14.5	13.2	51.6	25.0
Ours	36.1	40.7	20.4	11.3	77.6	36.6	34.5	15.2	20.9	54.3	31.5

为了更直观地表示 MEM-YOLO11 模型对无人机拍摄的小目标对象的检测能力,绘制的热力图如附录图 S1 所示.与原始模型相比,MEM-YOLO11 模型在多种小目标密集分布场景的识别中有效降低了漏检率,相比原始模型感受野区域有效增大.此外,本文绘制了损失和精度对比曲线,直观反映了本模型有效减少了特征信息损失.

为了验证 MEM-YOLO11 的泛化能力和通用性,进一步在 The WiderPerson 数据集上进行了实验验证.表 2 实验结果表明,MEM-YOLO11 在精确率(P)、召回率(R)、 $mAP50$ 和 $mAP50-95$ 分别提升了 6.8%、4.2%、5.6%、4.7%.与 YOLOv11n 模型相比验证了 MEM-YOLO11 模型的通用性与可行性.

表 2 MEM-YOLO11 与 YOLOv11n 在 The Wider Person 数据集上检测结果

Tab. 2 Detection result between our method and YOLOv11n on The Wider Person dataset

方法	$P/\%$	$R/\%$	$mAP50/\%$	$mAP50-95/\%$	参数量	GFLOPs
YOLO11n	75.7	61.6	70.2	43.2	2 584 102	6.3
Ours	82.5	65.8	75.8	48.9	2 214 945	5.4

4 结论

本文提出的 MEM-YOLO11 模型建立在 YOLOv11n 基线网络之上,引入了 MFEE、ELFR 和 MRFE 模块,并且调整了 $\text{top}K$ 、 α 、 β 参数的阈值大小,实验结果表明,与其他的检测模型相比,本文提出的 MEM-YOLO11 模型在密集场景小目标检测方面,精度有较高提升,并且参数量有效降低,是部署无人机等边缘设备进行目标检测实现优良检测性能的最佳选择.

附录见电子版(DOI:10.16366/j.cnki.1000-2367.2025.05.12.0002).

参 考 文 献

[1] Li L Z, Han L, Ye Y X, et al. Deep learning in remote sensing image matching: a survey[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2025, 225: 88-112.

[2] 叶栢铖, 朱允攀, 周永康, 等. 轻量级目标检测算法综述[J]. 红外技术, 2025, 47(3): 289-298.

Ye B C, Zhu Y P, Zhou Y K, et al. Review of lightweight target detection algorithms[J]. Infrared Technology, 2025, 47(3): 289-298.

[3] 王宁, 智敏. 深度学习下的单阶段通用目标检测算法研究综述[J]. 计算机科学与探索, 2025, 19(5): 1115-1140.

Wang N, Zhi M. Review of one-stage universal object detection algorithms in deep learning[J]. Journal of Frontiers of Computer Science and Technology, 2025, 19(5): 1115-1140.

[4] 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): 174-200.

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): 174-200.

[5] Li J L, Zheng C H, Chen P, et al. Small object detection in UAV imagery based on channel-spatial fusion cross attention[J]. Signal, Image and Video Processing, 2025, 19(4): 302.

[6] Yi W G, Wang B. Research on underwater small target detection algorithm based on improved YOLOv7[J]. IEEE Access, 2023, 11: 66818-66827.

[7] Guan S T, Lin Y M, Lin G Y, et al. Real-time detection and counting of wheat spikes based on improved YOLOv10[J]. Agronomy, 2024, 14(9): 1936.

[8] Lv L J, Li X J, Mao F J, et al. A deep learning network for individual tree segmentation in UAV images with a coupled CSPNet and attention mechanism[J]. Remote Sensing, 2023, 15(18): 4420.

[9] 王文润, 党建武, 王阳萍, 等. 基于复杂背景的多尺度特征融合手-物交互检测方法[J]. 兰州交通大学学报, 2024, 43(5): 94-102.

Wang W R, Dang J W, Wang Y P, et al. A multi-scale feature fusion method for hand-object interaction detection based on complex background[J]. Journal of Lanzhou Jiaotong University, 2024, 43(5): 94-102.

[10] 倪康, 孙笠焜, 邹旻瑞. 基于边缘特征引导学习的 SAR 目标检测[J]. Journal of Data Acquisition & Processing, 2025, 40(3): 612-624.

[11] 颜豪男, 吕伏, 冯永安. 特征级自适应增强的无人机目标检测算法[J]. 计算机科学与探索, 2024, 18(6): 1566-1578.

Yan H N, Lü F, Feng Y A. Feature-level adaptive enhancement for UAV target detection algorithm[J]. Journal of Frontiers of Computer Science and Technology, 2024, 18(6): 1566-1578.

[12] 张浩, 左杭, 刘宝华. 视觉与二维激光雷达的目标检测方法[J]. 电子测量与仪器学报, 2022, 36(3): 79-86.

Zhang H, Zuo H, Liu B H. Target detection method for visual and 2D laser radar[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(3): 79-86.

[13] 黄骏, 郭颖, 严舒. 基于多感受野与动态特征细化的遥感图像检测算法[J]. 激光与光电子学进展, 2024, 61(22): 2228004.

Huang J, Guo Y, Yan S. Remote sensing image detection algorithm based on multi-receptive field and dynamic feature refinement[J]. Laser & Optoelectronics Progress, 2024, 61(22): 2228004.

[14] Cui Y N, Ren W Q, Knoll A. Omni-kernel network for image restoration[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2024, 38(2): 1426-1434.

[15] Yu Z P, Huang H B, Chen W J, et al. YOLO-FaceV2: a scale and occlusion aware face detector[J]. Pattern Recognition, 2024, 155: 110714.

[16] 梁嘉杰,李星星.特定任务上下文解耦的遥感图像目标检测方法[J].计算机工程与应用,2025,61(2):293-303.
Liang J J,Li X X.Task-specific context decoupling object detection method for remote sensing images[J].Computer Engineering and Applications,2025,61(2):293-303.

[17] Du D,Zhu P,Wen L,et al.VisDrone-DET2019:The vision meets drone object detection in image challenge results[EB/OL].[2025-04-17]. <https://www.semanticscholar.org/paper/VisDrone-DET2019%3A-The-Vision-Meets-Drone-Object-in-Zhu-Wen/80edf721b4ab5f4f4e24db69131c360d3343eae>.

[18] Zhang S F,Xie Y L,Wan J,et al.WiderPerson:a diverse dataset for dense pedestrian detection in the wild[J].IEEE Transactions on Multimedia,2020,22(2):380-393.

MEM-YOLO11:A UAV small object detection model with multi-feature selection mechanism

Xu Shizhou, Yang Hong, Zhang Mengjie, Zhang Yuhao

(School of Optoelectronic Engineering, Henan Normal University, Xinxiang 453007, China)

Abstract: To address the challenges of multi-scale objects, varied orientations, and complex backgrounds in remote sensing images, we propose MEM-YOLO11, an enhanced model integrating a multi-feature selection mechanism. First, a feature extraction module is designed within the backbone network, combining multi-scale feature extraction and edge information enhancement to improve feature integration. Second, an efficient convolutional network module is developed to generate higher-resolution images from raw inputs, strengthening the global-to-local feature learning capability in the neck network. Additionally, a reconstructed multi-head feature extraction module reduces spatial and channel redundancy inherent in standard convolutions, significantly decreasing parameters and computational costs. Finally, hyperparameters of MEM-YOLO11 are optimized to alleviate the discrepancy between classification and localization tasks. Evaluations demonstrate significant improvements: on the VisDrone2019 dataset, precision, recall, and *mAP*50 increase by 7.6%, 5.6%, and 6.5%, respectively; on the Wider Person dataset, gains of 6.8%, 4.2%, and 5.6% are achieved for the same metrics.

Keywords: small object detection; Multi-scale feature extraction; edge enhancement; hyperparameter

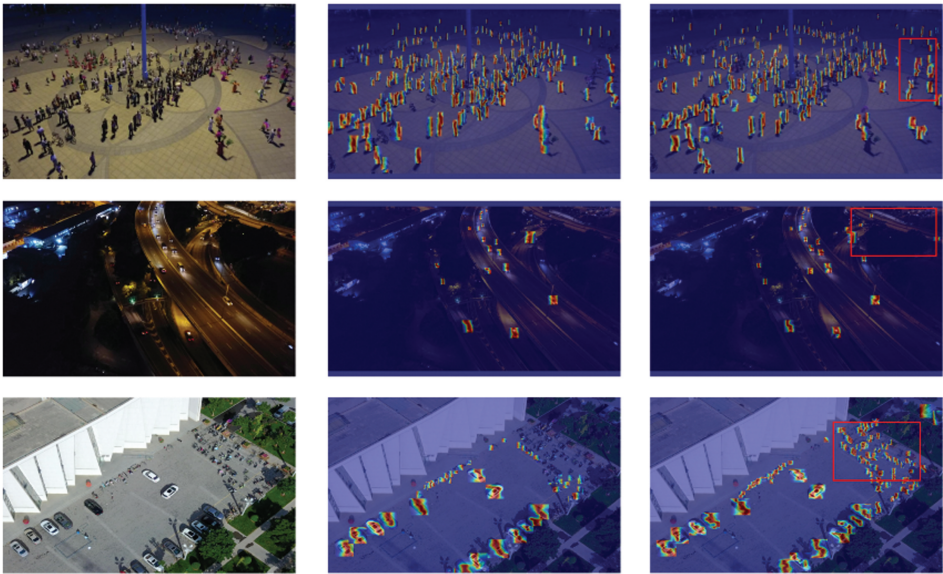
[责任编辑 陈留院 杨浦]

表 S1 消融实验
Tab. S1 Ablation experiments

Yolo11	MFEE	ELFR	MRFE	$P/\%$	$R/\%$	$mAP50/\%$	$mAP50-95/\%$
✓				37.7	29.5	29.6	15.1
✓	✓			39.7	29.7	30.5	15.6
✓	✓	✓		40.1	30.5	31.8	16.3
✓	✓	✓	✓	44.1	34.6	34.7	18.1

表 S2 超参数对比试验
Tab. S2 Comparative experiments of hyperparameters

TAL Para	$P/\%$	$R/\%$	$mAP50/\%$	$mAP50-95/\%$
topK = 13, $\alpha = 0.5, \beta = 6$	44.1	34.6	34.7	18.1
topK = 10, $\alpha = 0.5, \beta = 6$	44.8	33.8	34.9	18.5
topK = 7, $\alpha = 0.5, \beta = 6$	44.9	34.2	35.2	18.6
topK = 5, $\alpha = 0.5, \beta = 6$	44.1	34.5	35.2	19.3
topK = 7, $\alpha = 0.75, \beta = 6$	44.8	34.9	35.5	19.1
topK = 7, $\alpha = 0.5, \beta = 5$	45.3	35.1	36.1	19.5
topK = 7, $\alpha = 0.5, \beta = 4$	45.1	35.0	35.6	19.4



(a) 原始图像 (b) YOLOv11n检测热力图 (c) MEM-YOLO11检测热力图

图S1 MEM-YOLO11与YOLOv11n在VisDrone2019数据集上的热力图

Fig. S1 Heatmap comparison of YOLOv11n and MEM-YOLO11 on the VisDrone2019 dataset