

文章编号:1005-0523(2025)06-0042-09



ES-YOLO: 基于细节特征增强与冗余特征抑制的小目标检测方法

朱志亮¹, 黄欣荣¹, 刘 怡¹, 罗文俊², 朱碧堂², 张小刚³

(1. 华东交通大学信息与软件工程学院, 江西南昌 330013; 2. 华东交通大学土木建筑工程学院, 江西南昌 330013;
3. 中国铁路武汉局集团有限公司电务部, 湖北武汉 430071)

摘要:针对低空小目标在多层次下采样过程中细节特征丢失的问题,提出一种基于细节特征增强与冗余特征抑制的小目标检测模型ES-YOLO。该方法以轻量化YOLOv5s为基础,构建由空间细节增强模块(SDE)与冗余特征抑制模块(RFS)组成的双重特征优化机制。SDE通过动态上采样与反卷积上采样协同实现尺度自适应的空间细节精细恢复与结构一致性重建,增强小目标纹理与边界信息;RFS从通道与空间多维度建模特征依赖关系,抑制背景噪声与冗余响应,提高特征纯净度与目标显著性。实验结果表明,ES-YOLO在VisDrone2019数据集上的mAP@0.5与mAP@[0.5:0.95]较YOLOv5s分别提升12.97个百分点与9.22个百分点,计算量GFLOPs仅为YOLOv8m的38.59%。

关键词:小目标检测;细节特征增强;冗余特征抑制;YOLO

中图分类号:TP311

文献标志码:A

本文引用格式:朱志亮,黄欣荣,刘怡,等. ES-YOLO: 基于细节特征增强与冗余特征抑制的小目标检测方法[J]. 华东交通大学学报, 2025, 42(6):42-50.

ES-YOLO: Small Object Detection Method Based on Detail Feature Enhancement and Redundant Feature Suppression

Zhu Zhiliang¹, Huang Xinrong¹, Liu Yi¹, Luo Wenjun², Zhu Bitang², Zhang Xiaogang³

(1. School of Information and Software Engineering, East China Jiaotong University, Nanchang 330013, China;
2. School of Civil Engineering and Architecture, East China Jiaotong University, Nanchang 330013, China;
3. Electric Department of China Railway Wuhan Bureau Group Co., LTD., Wuhan 430071, China)

Abstract: To address the problem of detail feature loss of low-altitude small objects during multi-layer down-sampling, a small object detection model ES-YOLO is proposed, based on detail feature enhancement and redundant feature suppression. The method is built upon the lightweight YOLOv5s framework and constructs a dual-feature optimization mechanism consisting of spatial detail enhancement (SDE) and redundant feature suppression (RFS) modules. SDE collaborates dynamic upsampling with transposed convolution upsampling to achieve scale-adaptive fine recovery of spatial details and structural consistency reconstruction, enhancing small object texture and boundary information. RFS models feature dependencies across both channel and spatial dimensions to suppress background noise and redundant responses, improving feature purity and object saliency. Experimental results show that ES-YOLO achieves improvements of 12.97 percentage point and 9.22 percentage point in mAP@0.5

收稿日期:2025-11-11

基金项目:国家重点研发计划项目(2023YFB2603900);中国铁路武汉局集团有限公司科技研究开发计划课题(24D03)

and mAP@[0.5:0.95], respectively, compared to YOLOv5s on the VisDrone2019 dataset. The proposed model requires only 38.59% of the GFLOPs of YOLOv8m, achieving a significant reduction in computational cost.

Key words: small object detection; detail feature enhancement; redundant feature suppression; YOLO

Citation format: ZHU Z L, HUANG X R, LIU Y, et al. ES-YOLO: Small object detection method based on detail feature enhancement and redundant feature suppression[J]. Journal of East China Jiaotong University, 2025, 42(6): 42–50.

在计算机视觉领域,目标检测一直是基础且关键的研究方向,广泛应用于自动驾驶^[1]、路面病害监控^[2]和遥感图像处理^[3]等实际场景。随着深度学习的发展,目标检测算法的性能得到了显著提升。主流检测算法可分为两阶段(Two-stage)与单阶段(One-stage)两类。两阶段检测方法以R-CNN系列^[4]为代表,先生成候选区域再进行分类与回归,具有较高精度但计算复杂度较大;单阶段方法如SSD^[5]、RetinaNet^[6]、YOLO系列^[7-9]则通过端到端的预测框架在速度上具有明显优势,成为实时检测的主流方案。然而,在复杂背景下的小目标检测仍面临显著挑战。

小目标在图像中所占像素有限,其特征在深层下采样过程中易被削弱甚至丢失。因此,传统的目标检测模型难以捕捉关键细节信息,从而降低检测精度与稳定性。YOLO系列在多目标快速检测方面虽表现优异,但对于小目标的特征表达与细节保持仍显不足。为解决这一问题,已有研究主要从上采样重建与特征增强两个方向展开。在上采样方面,Wang等^[10]通过融合注意力与可变形卷积提升特征恢复效果;Park等^[11]在特征金字塔中引入语义注意以减少上采样过程中的语义损失;He等^[12]与徐洪俊等^[13]使用CARAFE替代传统上采样算子以增强细节捕捉能力。但这些方法在提升特征还原的同时,往往增加了计算复杂度,并可能引入噪声响应,影响检测稳定性。在特征增强方面,Tang等^[14]、Zhang等^[15]、Kang等^[16]通过增加微小检测头以提升高分辨率特征的利用率;Shang等^[17]与Liu等^[18]采用多层特征融合策略以改善检测性能,但仍存在过度增强背景细节的副作用。Li等^[19]、Miao等^[20]与Sun等^[21]进一步引入注意力机制强化特征选择能力,但未充分利用多维特征,导致冗余特征抑制不足。

针对低空小目标在多层下采样过程中细节特征丢失的问题,本文提出一种基于细节特征增强与冗余特征抑制的小目标检测方法ES-YOLO(detail feature enhancement and redundant feature suppression YOLO)。该方法以轻量化YOLOv5s为基础,构建由SDE与RFS组成的双重特征优化机制。CARAFE和Deformable/Dynamic upsampling等方法主要聚焦于局部区域的细节恢复,但在全局结构重建方面仍有提升空间。SDE模块不仅通过动态上采样恢复细节,还通过基于反卷积上采样的空间结构一致性重建来恢复小目标的纹理和边界信息,进一步提升细节恢复的丰富度。同时,区别于CBAM/ECA等注意力机制单纯通过加权通道来增强特征,RFS模块的多维度冗余特征过滤机制,从通道和空间两个维度进行冗余信息抑制,突出小目标的显著结构特征。因此,ES-YOLO的双重特征优化机制在细节恢复与冗余抑制的协同作用下,可以表现出更加优异的检测性能。

1 ES-YOLO模型结构及原理

1.1 ES-YOLO方法结构

如图1所示,ES-YOLO整体框架以轻量化YOLOv5s为基线模型,在颈部结构上将原三尺度融合改为四尺度特征融合,即采用PANet^[22]将P2(160×160)、P3(80×80)、P4(40×40)、P5(20×20)进行自顶向下与自底向上的双向聚合,充分利用浅层细节与深层语义的互补性。针对小目标在下采样过程中易丢失细节特征的问题,本文提出了以下3方面的改进。

1) 设计SDE:为了恢复下采样过程中小目标的纹理与边缘信息,SDE模块通过动态上采样与反卷积上采样协同工作。动态上采样根据局部特征自适应分配插值权重,从而精确恢复细节;而反卷积

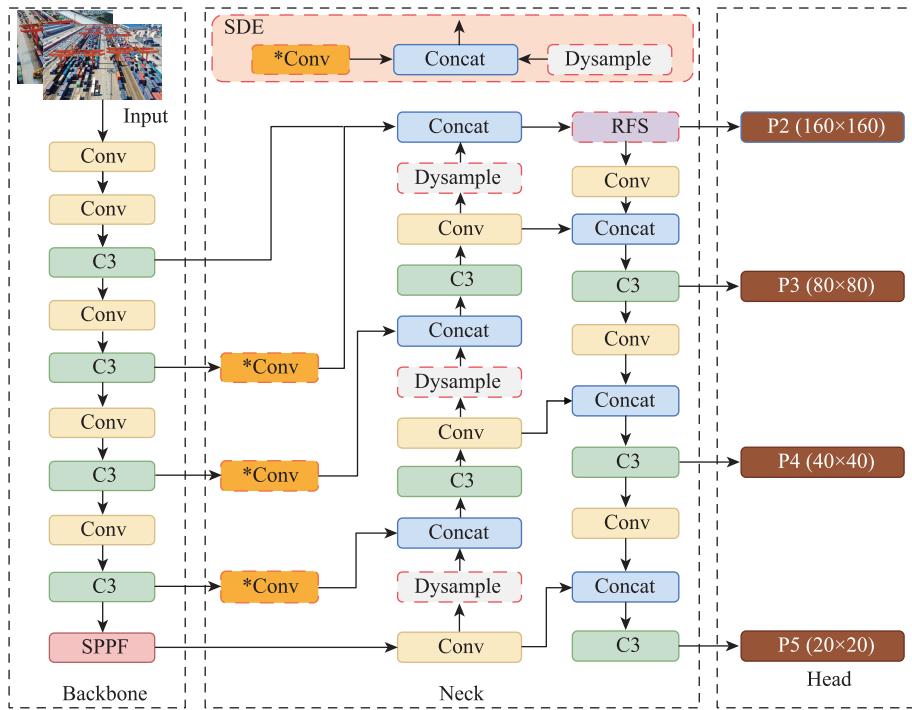


图1 ES-YOLO方法结构
Fig. 1 Structure of the ES-YOLO methodology

上采样利用可学习卷积核进一步重建空间结构与边界信息,显著提高细节恢复的精度。两者协同作用下,SDE模块能够有效地丰富特征图中的细节信息,增强小目标的检测能力。

2) 新增高分辨率检测头(P2)与RFS:在YOLOv5基线的三尺度检测头(P3, P4, P5)之外,新增高分辨率检测头P2。P2从主干网络第1个C3模块引出,并与SDE模块提取的浅层高分辨率特征融合作为输入。与第1个C3模块的特征相比,此时P2的细节特征更加丰富,但也伴随更多冗余特征。为了抑制冗余特征对深层次特征提取的干扰,设计了RFS。RFS模块通过从通道与空间维度建模特征依赖关系,采用加权筛选与响应强化的方法,抑制冗余特征,突出小目标的显著结构特征。

3) 独立小目标检测头的引入:受Zhu等^[23]与Liu等^[18]的研究启发,本文在新增的高分辨率检测层P2引入独立的小目标检测头,用于输出该层的分类与回归结果。该检测头针对极小目标的尺度特征进行了轻量化设计,能够充分利用浅层高分辨率特征,实现对极小目标的精细检测与高置信度识别,从而进一步提高检测精度。

1.2 SDE

YOLOv5s的双线性插值方法虽然计算高效,但在复杂场景中无法有效保留整体边缘和局部纹理信息。为了解决这一问题,本文设计了SDE,通过动态上采样与反卷积的协同作用,实现尺度自适应的空间细节恢复与结构一致性重建。

1) 动态上采样层(Dysample)。Liu等^[24]提出了一种轻量级的动态上采样方法Dysample,通过内容感知的方式有效恢复细节。其基本思想是通过点采样代替传统的内核卷积操作来实现上采样,从而避免了动态卷积的高计算开销。具体来说,Dysample通过线性投影生成偏移量,并利用PyTorch的grid_sample函数重新采样输入特征图。

偏移量是通过一个线性层生成的,线性层的输入是经过卷积处理的特征图,输出为每个上采样点的偏移量。该偏移量用于调整原始特征图中每个采样点的位置,重新定义其采样区域,从而增强特征图的空间细节。基于生成的偏移量,结合原始的网格采样位置,形成新的采样集。该采样集通过将原始采样网格与偏移量叠加得到

$$\begin{cases} O = \text{linear}(X) \\ S = G + O \end{cases} \quad (1)$$

式中: X 为输入特征图; G 原始的网格采样位置;

O 为生成的偏移量; S 为新的采样位置。

使用 `grid_sample` 函数基于采样集重新采样特征图, 从而得到高分辨率的特征图。这一过程通过以下公式表示为

$$X' = \text{grid_sample}(X, S) \quad (2)$$

通过这种方式, Dysample 实现了比传统上采样方法更灵活的采样策略, 同时保持了计算效率, 避免了额外的计算负担。

2) 反卷积层(*Conv)。在动态上采样获得初步放大特征后, 反卷积通过可学习的卷积核参数实现空间细节重建与边界结构修复。相比单纯的插值操作, 反卷积能在尺度恢复过程中引入局部空间关联建模能力, 使特征映射更具结构一致性。

3) 特征融合。最终, 通过 Concat 操作沿通道维度融合两支输出

$$F_{\text{sdem}} = \text{Concat}(Y_{\text{dyn}}, Y_{\text{conv}}) \quad (3)$$

动态上采样与反卷积通路被赋予同等的权重, 二者以并行、等比例的方式协同工作, 而非通过可学习的权重进行自适应融合。这种设计的初衷是充分利用两种不同上采样机制的优势——动态上采样能自适应地生成上采样核以保留细节, 而反卷积则通过固定的学习核来恢复结构信息——从而从不同角度共同增强空间细节。两条通路分别独立处理输入特征, 然后将输出的特征图进行逐元素相加。这种直接的融合方式确保了两种上采样结果的特征得以直接互补, 为后续检测层提供更精确的细节输入。

1.3 冗余特征抑制模块(RFS)

在 SDE 模块通过动态上采样和反卷积增强细节后, 生成的最高分辨率特征(P2)包含了丰富的细节信息和背景噪声。由于这一层特征图融合了多层次信息, 冗余特征与背景噪声较为显著, 因此需

要进行冗余特征抑制。为了有效地去除冗余特征, 确保小目标的显著特征能够被准确提取, 提出了 RFS 模块。RFS 模块通过从通道、空间和原始特征维度建模特征依赖关系, 采用加权筛选和响应强化的方法, 抑制冗余特征的影响, 突出小目标的结构特征。RFS 模块的工作原理是首先对最高分辨率的特征图进行多维度优化, 去除无关的冗余信息, 确保特征图的“纯净度”, 然后通过增强有效特征的响应, 使模型更加专注于小目标的显著特征。

如图 2 所示, RFS 模块包括以下 4 个步骤。

1) 特征通道扩展与多尺度特征提取。如图 2 所示, RFS 模块首先通过 1×1 卷积扩展输入特征图的通道数, 确保特征信息更加丰富。接着, 将扩展后的特征图划分为 5 组, 其中 X_1 保持原始特征, X_2 到 X_5 通过不同大小的倒置瓶颈层(分别为 $3 \times 3, 5 \times 5, 7 \times 7, 9 \times 9$ 卷积核)进行处理。这些卷积核有助于捕捉不同尺度的特征信息, 提升对多种目标尺度的适应性。这些特征图的数学表达式可以统一表示为

$$Y_i = \text{BottleNeck}_k(X_i), \quad i = 2, 3, 4, 5 \quad (4)$$

式中: k 为不同的卷积核大小; Y_i 为通过倒置瓶颈层处理后的特征图。

2) 多尺度特征融合。 Y_i 经过 1×1 卷积后进行多尺度拼接融合, 生成包含不同语义信息的综合特征图。通过这种融合方式, 特征图能够同时保留浅层细节信息与深层语义信息, 为后续的目标检测任务提供更为丰富的输入。

3) 跨维度依赖关系建模。RFS 模块接着对融合后的特征通过 3 个分支对不同维度的依赖关系进行建模, 确保特征间的有效交互, 去除冗余信息。

第 1 个分支专注于捕捉通道维度与水平空间维度(高度)之间的依赖关系。为了实现这一目标, 输

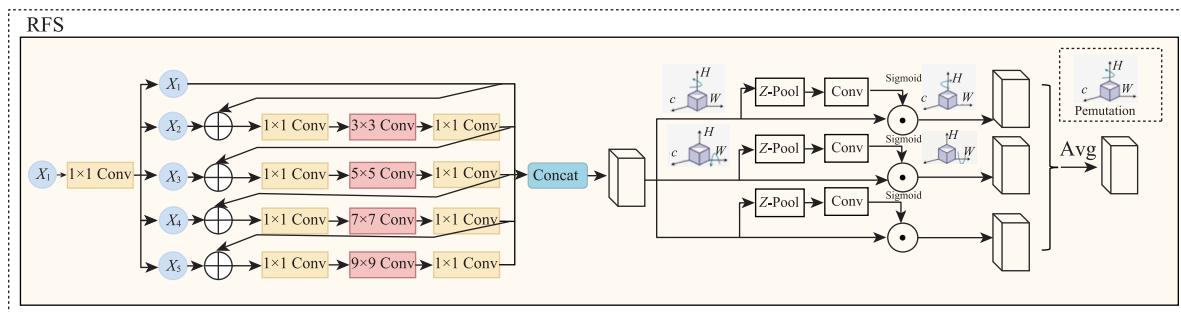


图 2 RFS 结构
Fig. 2 RFS structure

入特征图首先沿高度轴(H 轴)旋转90°,使得原本的水平空间信息转变为“垂直”方向。接着,通过Z-pool层(结合全局平均池化和全局最大池化)对旋转后的特征图 $R_h(X)$ 在高度和通道维度上进行平均和最大池化,生成蕴含跨维度信息的特征图。这些特征图随后经过一个标准卷积层和Sigmoid激活函数,生成用于调整原始特征图权重的注意力系数,并应用于原始特征图,以实现对通道-水平空间依赖关系的精细调整。该过程可以用公式表示为

$$A_h = \text{Sigmoid}(\text{Conv}(Z\text{-pool}(R_h(X)))) \quad (5)$$

第2个分支则负责捕捉通道维度与垂直空间维度(宽度)之间的依赖关系。其操作与第一个分支类似,但不同之处在于,这里特征图是沿宽度轴(W 轴)旋转90°进行处理。通过这一旋转,原本的垂直空间信息被转变为“水平”方向,便于捕捉通道与垂直空间之间的交互。生成的注意力权重同样应用于原始特征图,以强调或抑制特定的特征。该过程可以用公式表示为

$$A_w = \text{Sigmoid}(\text{Conv}(Z\text{-pool}(R_w(X)))) \quad (6)$$

第3个分支与传统的空间注意力机制相似,它捕捉特征图在水平和垂直空间维度之间的依赖关系。尽管其操作与通道-水平空间交互分支有相似之处,但此分支更侧重于纯粹的空间维度间的交互。该过程可以用公式表示为

$$A_s = \text{Sigmoid}(\text{Conv}(Z\text{-pool}(X))) \quad (7)$$

4) 注意力机制与特征加权。经过以上3个分支的处理后,将得到的3个注意力权重 A_h , A_w , A_s 分别应用于输入特征图,从而对冗余信息进行抑制,并突出小目标的显著结构特征。最终,3个分支的输出特征图会进行加权平均,得到最终的输出特征图 Y 。这一权过程确保了冗余信息被有效过滤,同时增强了小目标的显著特征,减少背景噪声的干扰,从而进一步提高目标的检测精度。该过程可以表示为

$$Y = \frac{1}{3}(A_h X + A_w X + A_s X) \quad (8)$$

1.4 损失函数

YOLOv5s在每个尺度的特征图上对网格位置进行预测,并将预测与标注进行匹配以计算损失,从而指导网络收敛。为同时优化定位、目标性与类别判别,本文采用由定位损失 L_{loc} 、置信度(目标性)损失 L_{obj} 与分类损失 L_{cls} 加权组成的总损失

$$L = \alpha L_{\text{loc}} + \beta L_{\text{obj}} + \gamma L_{\text{cls}} \quad (9)$$

式中: α , β , γ 分别为定位、目标性、分类三项的权重系数。

2 实验结果与对比分析

2.1 系统软硬件配置及参数设置

实验在搭载Intel(R) Core(TM) i9-9900KF CPU @ 3.60 GHz 和 NVIDIA GeForce RTX 4080 GPU 的工作站上完成。算法基于PyTorch框架实现,并使用CUDA 12.2进行加速训练。本文基于YOLOv5s为基线进行实验验证,将输入图像大小设置为640×640, Batch size设置为16,训练Epoch设置为300,深度和宽度分别设置为0.33和0.50,使用提前停止策略来避免过度拟合。其中Patience设置为15,优化器使用Adam,初始学习率为 1×10^{-3} ,损失函数的权重分别设置为0.05(定位),0.50(目标)和0.25(分类),对应式(9)中的 α , β , γ 。其他参数采用YOLOv5s默认参数。

2.2 性能比较与结果分析

为全面评估ES-YOLO在不同场景下的小目标检测性能,选用VisDrone 2019公开数据集进行实验。模型性能通过Precision(P)、Recall(R)、mAP@0.5、mAP@[0.5:0.95]和GFLOPs指标进行评估。

在VisDrone 2019测试集上(表1),ES-YOLO的mAP@0.5为40.54%,较YOLOv5s提升12.97个百分点;mAP@[0.5:0.95]为23.65%,较YOLOv5s提

表1 VisDrone 2019测试集上的算法比较

Tab.1 Comparison of the algorithms on the VisDrone 2019 test set

方法	mAP@0.5/%	mAP@[0.5:0.95]/%	GFLOPs/G
YOLO3-tiny	16.20	6.30	11.7
YOLOX-tiny	23.10	12.40	3.2
YOLOv5n	23.10	11.50	4.2
YOLOv5s	27.57	14.43	16.5
YOLOv8s	34.20	19.90	28.7
YOLOv8m	39.20	23.10	79.3
RetinaNet ^[25]	29.00	16.90	21.0
ESOD-YOLO ^[26]	29.30	16.60	14.3
SF-YOLOv5 ^[18]	34.30	18.20	—
GBS-YOLOv5 ^[27]	35.30	20.00	—
Li等 ^[28]	36.40	20.10	—
HIC-YOLOv5 ^[14]	36.95	20.85	31.2
ES-YOLO	40.54	23.65	30.6

升9.22个百分点。同时,ES-YOLO的GFLOPs仅为30.60 G,仅为YOLOv8m的38.59%,说明本模型在低计算成本下仍能保持高检测精度。

与其他改进算法对比:SF-YOLOv5^[18]、RetinaNet^[25]、ESOD-YOLO^[26]、GBS-YOLOv5^[27]、HIC-YOLOv5^[14]及Li等^[28]的方法均聚焦特征融合优化,提升了特征层间语义交互,但未充分考虑冗余特征的抑制。ES-YOLO在此基础上同时关注特征细节增强与多维度冗余过滤,在有效提升小目标特征表达力的同时减少噪声干扰,从而在精度与鲁棒性上均表现最优。

2.3 视觉对比分析

图3对比了YOLOv5、HIC-YOLOv5与ES-YOLOv5在VisDrone图像上的检测结果。由图3可见,ES-YOLOv5能识别出更多被遮挡或远距离的小目标,尤其在红色箭头标注区域中,HIC-YOLOv5未

能检测出的目标被准确识别,表明本模型在复杂场景下具有更强的细节感知能力。

2.4 消融实验

为验证本文提出的双重细节特征增强策略中各模块的有效性,在VisDrone 2019数据集上进行了多组消融实验。实验通过逐步添加或移除不同模块的方式,分析各组件对整体性能的影响,其中“√”表示模型中包含该模块。

1) 基于细节特征增强与冗余特征抑制方法的整体验证。为验证本文提出的先增强后抑制框架的整体作用,设计了4组实验(表2)。所有实验均在已添加高分辨率检测层(P2)的前提下进行:①组1:仅添加P2层(不含增强模块);②组2:在组1基础上添加SDE;③组3:在组1基础上添加RFS;④组4:同时添加SDE与RFS。

结果显示,组1相较原始YOLOv5显著提升了

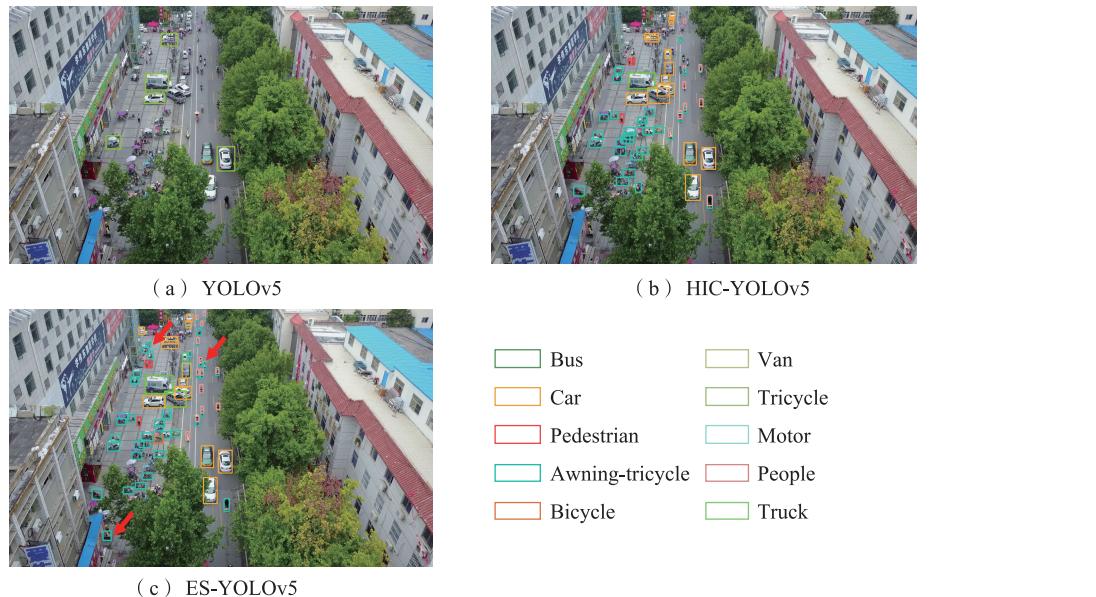


图3 YOLOv5, HIC-YOLOv5与ES-YOLOv5检测效果的比较
Fig. 3 Comparison of the detection effects of YOLOv5, HIC-YOLOv5 and ES-YOLOv5

表2 基于细节特征增强与冗余特征抑制的小物体目标检测方法消融实验

Tab.2 Ablation experiments on small object detection methods based on detail feature enhancement

and redundant feature suppression

%

组别	SDE	RFS	P	R	mAP@0.5	mAP@[0.5:0.95]
YOLOv5	—	—	37.59	31.52	27.57	14.43
组1	—	—	45.18	37.89	35.88	19.94
组2	√	—	49.33	42.28	39.52	22.90
组3	—	√	46.80	41.57	40.06	23.15
组4	√	√	49.81	43.25	40.54	23.56

mAP指标,证明了引入小目标检测层的有效性。进一步比较可见,组2中加入SDE后,mAP@0.5由35.88%提升至39.52%,说明SDE能够显著强化特征细节表达。组3在此基础上引入RFS,mAP@[0.5:0.95]提升3.21个百分点,表明其有效抑制了冗余特征、增强了特征判别性。组4联合使用两模块时性能最佳,mAP@0.5提升4.66个百分点,mAP@[0.5:0.95]提升3.62个百分点,充分验证了双重增强策略的协同效应。

2) SDE的组件分析。为探究SDE内部结构的有效性,对其两种核心机制——动态上采样与反卷积进行了组合测试(表3)。实验默认包含RFS模块作为支撑:①组1:无动态上采样与反卷积;②组2:仅使用反卷积;③组3:仅使用动态上采样;④组4:同时启用两者(完整SDE)。

表3 SDE消融实验
Tab.3 Ablation study of SDE

组别	反卷积	动态上采样	mAP@0.5	mAP@[0.5:0.95]	%
组1	—	—	40.06	23.15	
组2	√	—	40.13	23.22	
组3	—	√	40.35	23.25	
组4	√	√	40.54	23.56	

实验结果表明,SDE完整配置(组4)取得最高性能($mAP@0.5=40.54\%$, $mAP@[0.5:0.95]=23.56\%$)。单独使用任一子模块虽有提升,但幅度有限。由此可见动态上采样在内容自适应插值方面提升空间一致性,而反卷积在结构重建上强化细节表达。两者协同使特征的尺度恢复与空间重构兼具适应性与层次性,从而显著提升小目标检测精度。

3) RFS在不同检测层下的适用性。为验证RFS在不同尺度特征层中的适用性,设计了在P2~P5共4个检测层间的组合实验(表4)。所有组

别均包含SDE模块,以隔离RFS的影响。

结果表明,单纯在多个检测层中叠加RFS并不能带来持续增益。组1(4层均添加)性能反而下降,而组7(仅在P2层添加)性能最优($mAP@0.5=40.54\%$, $mAP@[0.5:0.95]=23.56\%$)。这说明RFS的作用在于强化低层细粒度特征,而过度叠加会导致高层语义信息被反复过滤,产生特征损耗。进一步比较组4~组7的单层插入实验结果可知,RFS在P2层(160×160)的增益最明显。这是因为P2层专注于浅层高分辨率特征,对小目标敏感度最高;RFS的多维度特征过滤机制能有效提取有益的局部细节并抑制背景干扰。因此,将RFS模块部署于P2层是最优方案,可在性能与复杂度之间取得最佳平衡。

2.5 基于DOTAv1.0和DOTAv1.5数据集的泛化性实验

为了评估提出模型在小物体目标检测任务中的泛化性能,本文在DOTAv1.0和DOTAv1.5数据集上进行泛化性实验。使用的评价指标与对比实验中的不同,主要原因是DOTA数据集要求将模型的检测结果上传至官方评测网站进行测试,而该网站仅提供了特定的评价指标。为了确保结果的可比性并符合DOTA数据集的评估规范,在泛化实验中采用了这些标准指标进行性能评估,结果如表5所示。

本文将ES-YOLO模型与基线YOLOv5s模型和HIC-YOLOv5模型在两个数据集上复现的结果进行了对比。ES-YOLO模型在各项指标上均优于其他模型,展现了其在不同小物体目标检测任务中的优秀泛化能力。在DOTAv1.0数据集上,ES-YOLO模型的AP50得分为61.54%,AP75得分为41.40%,平均精度(mAP)达到38.86%,展示了其在处理不同目标大小和复杂背景时的鲁棒性。同样,

表4 RFS消融实验

组别	P2	P3	P4	P5	P	R	mAP@0.5	mAP@[0.5:0.95]	%
组1	√	√	√	√	47.20	40.07	38.08	21.82	
组2	√	√	√	—	48.21	41.47	39.10	22.10	
组3	√	√	—	—	49.85	43.99	39.40	22.92	
组4	—	√	—	—	48.33	42.83	39.94	23.24	
组5	—	—	√	—	47.64	42.61	39.61	22.55	
组6	—	—	—	√	49.66	39.92	38.74	22.55	
组7	√	—	—	—	49.81	41.57	40.54	23.56	

表5 DOTA 数据集上不同模型的性能对比

Tab.5 Performance comparison of different models

		on the DOTA dataset			%
数据集	模型	AP50	AP75	mAP	
DOTAv1.0	YOLOv5s	55.30	34.13	33.54	
	HIC-YOLOv5 ^[14]	59.86	38.59	37.02	
	ES-YOLO (本文)	61.54	41.40	38.86	
DOTAv1.5	YOLOv5s	46.89	27.57	27.85	
	HIC-YOLOv5 ^[14]	49.47	29.38	29.04	
	ES-YOLO (本文)	50.66	30.38	30.17	

在DOTAv1.5数据集上, ES-YOLO模型的mAP达到30.17%,有效提高了极小物体的检测性能。这些结果验证了本文提出的改进方法的有效性,并提升了模型在多样化和具有挑战性的环境中的表现。

3 结论

1) 本文提出的ES-YOLO模型通过创新的SDE模块和RFS模块,解决了小目标检测中的细节丢失和冗余特征干扰问题,显著提升了检测精度。实验结果表明,ES-YOLO在多个标准数据集上表现优异,尤其在复杂背景下的小目标检测中,展现出较高的鲁棒性和计算效率。这使得该模型在交通基础设施智能建造中具有广泛的应用潜力,能够为智能监控、实时检测和施工进度管理提供强有力的支持。

2) 未来,ES-YOLO将在无人化施工和智能建造的融合应用中发挥重要作用,尤其是在交通基础设施建设领域。通过集成到无人机巡检、自动化施工机器人、施工现场实时监控系统等场景中,ES-YOLO能够实现自动检测施工进度、实时识别结构缺陷、确保安全隐患预警等功能。随着5G、物联网和边缘计算的不断发展,ES-YOLO将能够在更广泛的无人化施工环境中部署,实现更加智能化、高效的施工与监控,推动交通基础设施建设向智能化、自动化方向迈进,提升建设效率与安全性。

参考文献:

- LI B Y, OUYANG W L, SHENG L, et al. GS3D: an efficient 3D object detection framework for autonomous driving[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019: 1019-1028.
- 杨烨,徐霖,徐峰. 基于改进YOLOv8n的轻量化路面裂缝检测算法[J]. 华东交通大学学报, 2025, 42(3): 117-126.
- YANG Y, XU P, XU F. Lightweight road crack detection algorithm based on improved YOLOv8n[J]. Journal of East China Jiaotong University, 2025, 42(3): 117-126.
- 程路,刘家伟,周庆忠,等. 基于改进RetinaNet的遥感图像目标检测算法[J]. 华东交通大学学报, 2024, 41(6): 74-80.
- CHENG L, LIU J W, ZHOU Q Z, et al. Remote sensing image object detection algorithm based on improved RetinaNet[J]. Journal of East China Jiaotong University, 2024, 41(6): 74-80.
- SAFFARINI R, KHAMAYSEH F, AWWAD Y, et al. Dynamic generative R-CNN[J]. Neural Computing and Applications, 2025, 37(10): 7107-7120.
- LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[M]//Computer Vision-ECCV 2016. Cham: Springer International Publishing, 2016: 21-37.
- LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 2999-3007.
- MURAT A A, KIRAN M S. A comprehensive review on YOLO versions for object detection[J]. Engineering Science and Technology, an International Journal, 2025, 70: 102161.
- CHEN Y M, YUAN X B, WANG J B, et al. YOLO-MS: rethinking multi-scale representation learning for real-time object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2025, 47(6): 4240-4252.
- DOHERTY J, GARDINER B, KERR E, et al. BiFPN-YOLO: one-stage object detection integrating Bi-Directional Feature Pyramid Networks[J]. Pattern Recognition, 2025: 111209.
- WANG Y, WEI Y S, WU Z Z, et al. Adaptive convolutional neural network for aluminum surface defect detection[J]. Computational Materials Science, 2023, 227: 112262.
- PARK H, PAIK J. Pyramid attention upsampling module for object detection[J]. IEEE Access, 2022, 10: 38742-38749.
- HE Y S, SHEN L R, HU Y. SCW-YOLO: an improved algorithm for fall detection based on deep learning[M]//Advanced Intelligent Computing Technology and Applications. Singapore: Springer Nature Singapore, 2024: 408-418.
- 徐洪俊,唐自强,张锦东,等. 钢材表面缺陷检测的YOLOv5s算法优化研究[J]. 计算机工程与应用, 2024, 60(7): 306-314.
- XU H J, TANG Z Q, ZHANG J D, et al. Research on optimization of YOLOv5s detection algorithm for steel surface defect[J]. Computer Engineering and Applications,

- 2024, 60(7): 306-314.
- [14] TANG S Y, ZHANG S, FANG Y N. HIC-YOLOv5: improved YOLOv5 for small object detection[C]//2024 IEEE International Conference on Robotics and Automation (ICRA), May 13-17, 2024, Yokohama, Japan. New York: IEEE, 2024: 6614-6619.
- [15] ZHANG M, YIN L J. Solar cell surface defect detection based on improved YOLOv5[J]. IEEE Access, 2022, 10: 80804-80815.
- [16] KANG M, TING C M, TING F F, et al. BGF-YOLO: enhanced YOLOv8 with multiscale attentional feature fusion for brain tumor detection[M]//Medical Image Computing And Computer Assisted Intervention-MICCAI 2024. Cham: Springer Nature Switzerland, 2024: 35-45.
- [17] SHANG J C, WANG J S, LIU S B, et al. Small target detection algorithm for UAV aerial photography based on improved YOLOv5s[J]. Electronics, 2023, 12(11): 2434.
- [18] LIU H Y, SUN F Q, GU J, et al. SF-YOLOv5: a light-weight small object detection algorithm based on improved feature fusion mode[J]. Sensors, 2022, 22(15): 5817.
- [19] LI X R, ZHANG J J, SUN L H. ECA-YOLOv5: multi-scale infrared salient target detection algorithm based on anchor free network[C]//Ninth Symposium on Novel Photoelectronic Detection Technology and Applications. November 2-4, 2022. Hefei, China. SPIE, 2023: 24.
- [20] MIAO L Z, LI N, ZHOU M L, et al. CBAM-Yolov5: improved Yolov5 based on attention model for infrared ship detection[C]//International Conference on Computer Graphics, Artificial Intelligence, and Data Processing (ICCAID 2021). December 24-26, 2021. Harbin, China. SP-IE, 2022: 33.
- [21] SUN Y, WANG W N, ZHANG Q Y, et al. Improved YOLOv5 with transformer for large scene military vehicle detection on SAR image[C]//2022 7th International Conference on Image, Vision and Computing (ICIVC), July 26-28, 2022, Xi'an, China. New York: IEEE, 2022: 87-93.
- [22] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 8759-8768.
- [23] ZHU X K, LYU S C, WANG X, et al. TPH-YOLOv5: improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios[C]//2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). October 11-17, 2021. Montreal, BC, Canada. New York: IEEE, 2021: 2778-2788.
- [24] LIU W Z, LU H, FU H T, et al. Learning to upsample by learning to sample[C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV), October 1-6, 2023, Paris, France. New York: IEEE, 2023: 6004-6014.
- [25] LI Y X, DUA A, REN F B. Light-weight RetinaNet for object detection on edge devices[C]//2020 IEEE 6th World Forum on Internet of Things (WF-IoT). June 2-16, 2020. New Orleans, LA, USA. New York: IEEE, 2020: 1-6.
- [26] LUO J, LIU Z C, WANG Y B, et al. Efficient small object detection you only look once: a small object detection algorithm for aerial images[J]. Sensors, 2024, 24(21): 7067.
- [27] LIU H Y, DUAN X H, LOU H T, et al. Improved GBS-YOLOv5 algorithm based on YOLOv5 applied to UAV intelligent traffic[J]. Scientific Reports, 2023, 13: 9577.
- [28] LI X L, BAO Y F. Small target detection algorithm for UAV aerial photography based on improved YOLO V8n[C]//2024 6th International Conference on Data-driven Optimization of Complex Systems (DOCS), August 16-18, 2024, Hangzhou, China. New York: IEEE, 2024: 870-875.



第一作者:朱志亮(1988—),男,副教授,博士,硕士生导师,研究方向为计算机视觉和人机交互。E-mail: rj_zzl@ecjtu.edu.cn。



通信作者:罗文俊(1979—),女,教授,博士,博士生导师,国务院特殊津贴专家、江西省主要学科学术及技术带头人、江西省青年井冈学者、江西省杰出青年人才,研究方向为铁路基础设施建造、运维及防灾减灾。E-mail: lwj06051979@163.com。

(责任编辑:吴海燕)