

# 针对小目标垃圾检测的 YOLOv8 网络结构改进与性能验证

doi:10.20269/j.cnki.1674-7100.2026.2013

何达维 吴岳忠  
陈玲姣 韩梅  
湖南工业大学  
交通与电气工程学院  
湖南 株洲 412007

**摘要:**为攻克复杂城市环境下的垃圾检测难题,尤其是在垃圾目标较小、背景繁杂的场景,提出一种基于YOLOv8的改进垃圾检测算法YOLO-SPRL。首先,引入SPD-Conv模块并增加P2检测层,删除P5检测层,以增强模型对小目标垃圾物体的检测能力。其次,在特征金字塔P3层嵌入矩形自校准模块(RCM),抑制背景噪声,消除复杂背景的干扰。最后,在模型颈部引入可分离多级卷积核注意力模块(LSKA),在降低计算复杂度的同时强化深度特征融合能力,在减小模型体积的同时提高检测精度。在VisDrone和自建街景垃圾数据集上的实验结果显示,YOLO-SPRL在VisDrone数据集上的mAP@0.5达到38.91%,相较于基准模型提升了4.22%;在自建数据集上亦展现出良好的鲁棒性和小目标识别能力。本研究所提出的集成改进策略有效提升了复杂场景下的垃圾检测性能,YOLO-SPRL在精度、速度和模型体积之间实现了良好的平衡,具备在嵌入式移动设备上部署的可行性,为城市智能环卫系统的实时监测与管理提供了可靠的技术支撑。

**关键词:**YOLOv8; 小目标检测; 矩形自校准模块; 可分离多级卷积核注意力

**中图分类号:**TP183; TP391.4; X799.3

**文献标志码:**A

**文章编号:**1674-7100(2026)02-0103-07

**引文格式:**何达维,吴岳忠,陈玲姣,等.针对小目标垃圾检测的YOLOv8网络结构改进与性能验证[J].包装学报,2026,18(2):103-109.

## 1 研究背景

随着城市化进程加速和人口持续增长,城市垃圾处理面临着日益严峻的挑战。传统垃圾管理方式效率低下、成本高昂,难以满足现代城市可持续发展的需求。在此背景下,结合计算机视觉与深度学习的目标检测技术为智能垃圾管理提供了新的解决方案,通过对垃圾的自动识别、分类和监测,显著提升了垃圾处理的效率与准确性。基于深度学习的

垃圾分类目标检测方法已成为环境智能管理领域的研究热点。

目标检测是计算机视觉领域的一个重要研究方向,旨在从图像或视频中识别出感兴趣的目标,并确定其位置和大小。这一技术结合了图像分类和定位两个任务,要求模型不仅能够识别出图像中的物体类别,还要能够准确地标注出物体的边界框。目标检测技术的发展经历了从传统方法到深度学习的转变。传统方法主要依赖于手工设计的特征和分类器,如方向

收稿日期:2025-11-04

基金项目:湖南省自然科学基金资助项目(2025JJ70057)

作者简介:何达维,男,硕士生,主要研究方向为机器视觉,E-mail:hdw2459@163.com

通信作者:吴岳忠,男,教授,博士生导师,主要从事智能信息处理研究,E-mail:wuyuezhong@hut.edu.cn

梯度直方图 (histogram of oriented gradient, HOG) + 支持向量机 (support vector machine, SVM)、可变形部件模型 (direct part marking, DPM) 等。然而, 这些方法在复杂场景下的性能有限, 难以应对垃圾处理中的多样性和不确定性。近年来, 随着深度学习技术的兴起, 目标检测取得了突破性进展。以卷积神经网络 (CNN) 为代表的深度学习模型能够自动学习图像中的特征表示, 从而实现对目标的准确检测。基于深度学习的目标检测方法主要分为两类: 基于候选区域的方法和基于回归的方法。前者如 R-CNN 系列, 通过生成候选区域并进行精细分类和回归; 后者如 YOLO (you only look once) 和 SSD (single shot multibox detector) 等, 则直接对图像进行网格划分, 对每个网格进行类别预测和边界框回归<sup>[1-5]</sup>。

垃圾种类繁多, 尺寸各异, 从微小的电池到庞大的纸箱, 其形状极不规则。故而, 在基于深度学习的垃圾分类检测过程中面临两个核心挑战: 1) 小目标 ( $<32 \times 32$ ) 检测性能受限。垃圾目标尺寸分布高度不均衡, 小目标漏检率较高。2) 复杂背景干扰。垃圾堆叠、异物混杂导致误检率上升。现有研究主要围绕检测精度、实时性及复杂场景适应性展开。主要技术路径聚焦于改进主流算法框架的场景适应性, 尤其是基于单阶段检测器的优化策略占据主导地位。YOLO 系列算法凭借其突出的实时性优势而得到广泛采用<sup>[6-9]</sup>。YOLOv8 采用无锚框 (anchor-free) 机制, 摈弃了对预设锚框的依赖。这种设计不仅增强了模型对不规则目标的泛化检测能力, 还简化了模型训练流程<sup>[10-12]</sup>。在 YOLOv8 的骨干网络中, C2F (cross stage partial faster) 模块是核心创新点, 其优势在于通过丰富梯度流提升网络的特征提取能力。具体而言, 该模块在 CSP 模块的基础上, 引入更多跳跃连接, 能够聚合不同感受野下的浅层细节特征和深层语义特征。此外, YOLOv8 针对检测任务特性进行了针对性优化, 采用解耦检测头设计, 将分类任务和边界框回归任务进行分离, 有效降低了两类任务间的相互干扰, 进而提升模型的分类准确率和定位精度。YOLOv8 强大的开源生态与工程友好性, 显著降低了算法实现与实验迭代的复杂度<sup>[13-17]</sup>。

综上, 本研究针对小目标垃圾检测问题, 基于 YOLOv8 模型, 提出一种改进算法 YOLO-SPRL, 旨在提升垃圾管理效率, 为后续资源化利用奠定基础。

## 2 YOLO-SPRL 模型设计

YOLOv8 网络架构主要由主干网络 (Backbone)、颈部网络 (Neck) 以及检测头 (Head) 3 个核心组件构成。主干网络由 Conv、SPPF 和 C2f 模块组合而成。其中, Conv 模块有卷积层、批量归一化 (batch normalization, BN) 和 SiLU 激活函数, 提取基础特征; SPPF 模块通过堆叠多个最大池化层, 实现多尺度特征处理的加速; C2f 模块引入残差连接, 并融合跨阶段特征, 从而增强梯度流。颈部网络采用双向特征金字塔结构, 借助双向跨尺度特征融合机制, 构建具备多层次、语义增强特性的特征金字塔。检测头采用创新性的解耦头设计, 将目标分类任务与定位回归任务分离。基于此, 对 YOLO-SPRL 实施了如下改进举措: 为解决传统卷积导致的信息丢失问题, 采用空间深度转换卷积模块 (SPD-Conv) 替代步长卷积, 通过空间到深度变换保留完整像素信息; 针对垃圾目标尺寸小的特点, 在骨干网络浅层引入高分辨率 P2 检测层并移除冗余 P5 检测层, 增强对微小目标的细节捕捉能力; 在特征金字塔网络集成矩形自校准模块 (rectangular self-calibration module, RCM), 通过轴向池化捕获全局上下文, 并利用条状卷积动态调整注意力区域, 有效抑制背景噪声; 最后, 采用可分离多级卷积核注意力 (large separable kernel attention, LSKA) 替代传统注意力机制, 在降低计算复杂度的同时, 强化深度特征融合能力。YOLO-SPRL 网络结构见图 1。

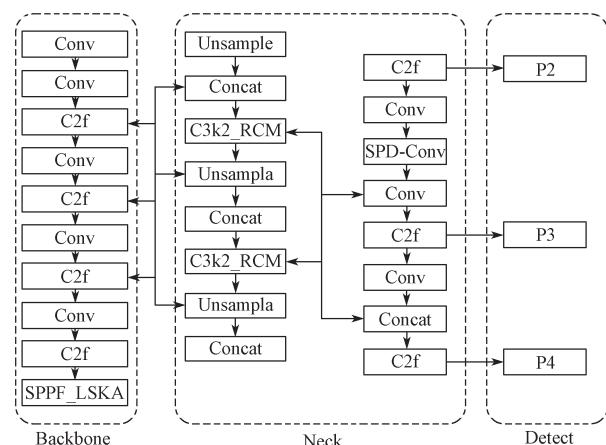


图 1 YOLO-SPRL 网络结构

Fig. 1 YOLO-SPRL network structure

### 2.1 SPD-Conv 模块

SPD-Conv 模块是一种 CNN 构建块, 由一个空

间到深度层 (SPD) 和一个非步长卷积层 (即普通卷积) 组成。SPD 层对特征图  $X$  进行下采样, 但保留通道维度上的所有信息, 避免传统步长操作导致的细粒度信息丢失, 再通过非步长卷积融合特征, 以减少卷积层中的可学习参数的通道数量<sup>[18]</sup>。具体操作步骤如下: 将输入特征图  $X \in \mathbb{R}^{S \times S \times C_1}$  按缩放因子 ( $s=2$ ) 拆分为  $s^2$  个子特征图, 即划分为 4 个子图 (尺寸为  $S/2 \times S/2 \times C_1$ ), 通道数增加  $s^2$  倍, 沿通道维度拼接并生成特征  $X' \in \mathbb{R}^{S/2 \times S/2 \times 4C_1}$ , 接着用步长为 1 的卷积层 (如  $3 \times 3$  卷积) 将高通道特征  $X'$  映射到目标维度  $X'' \in \mathbb{R}^{S/2 \times S/2 \times C_2}$ 。其核心优势为: 1) 空间信息的保留。所有像素都被重组到通道维度, 无信息丢弃。2) 分辨率的适配。输出尺寸缩减为  $1/s$ , 但通道数增加  $s^2$  倍。3) 非步长卷积的使用。采用步长为 1 的卷积能避免因步长大于 1 导致的非对称采样 (如奇数 / 偶数行差异)。通过可学习参数压缩通道数 ( $C_2 < s^2 C_1$ ), 提取判别性特征。该模块能增强模型对小尺寸垃圾的感知能力, 有效缓解低分辨率图像中的特征退化问题。

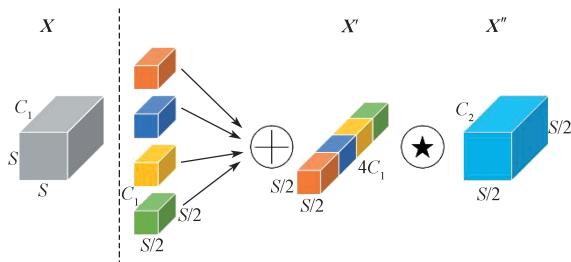


图 2 SPD-Conv 原理

Fig. 2 Principle of SPD-Conv

## 2.2 P2 检测层

在 YOLOv8 网络框架内, 通过 2 倍、4 倍、8 倍、16 倍和 32 倍下采样生成 5 个不同尺度的特征图 ( $P_1 \sim P_5$ )。对于  $640 \times 640$  的输入图像, 这些特征图对应的分辨率分别为  $320 \times 320 (P_1)$ 、 $160 \times 160 (P_2)$ 、 $80 \times 80 (P_3)$ 、 $40 \times 40 (P_4)$  和  $20 \times 20 (P_5)$ 。浅层高分辨率特征 (如  $P_2$  和  $P_3$ ) 包含丰富的空间细节信息, 非常适合于捕捉小目标的轮廓。相比之下, 深层低分辨率特征 (如  $P_4$  和  $P_5$ ) 受益于更大的感受野, 使其更适合检测大目标。尽管融合  $P_3 \sim P_5$  的特征能够覆盖检测分辨率范围在  $8 \times 8 \sim 32 \times 32$  之间的目标, 但其在识别小尺度目标方面仍然存在局限。其原因是深层特征经过多次下采样操作会使细粒度细节丢失。因此, 为提升小尺度垃圾的检测能力, 在骨干网络浅层引入 P2 检测层, 以保留微小垃圾目标更丰富的细

节纹理, 与此同时, 考虑  $P_5$  检测层的贡献有限且可能引入噪声, 故去除  $P_5$  检测层, 以控制计算成本, 改善边界定位效果<sup>[19]</sup>。

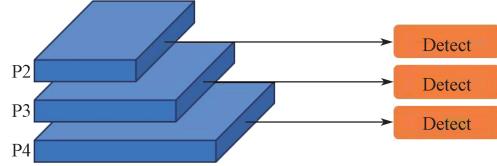


图 3 检测层示意图

Fig. 3 Schematic diagram of detection layer

## 2.3 RCM 模块

RCM 模块 (见图 4) 主要通过矩形自校准注意力机制 (residual cross-attention, RCA) 和可学习的形状自校准机制, 实现高效且精确的前景目标空间建模。该模块首先对输入特征分别进行水平与垂直方向的全局池化, 获取两个轴向的全局上下文向量, 再通过广播加法将两个轴向量融合为一个矩形注意力区域。随后, 形状自校准函数<sup>[20]</sup>利用解耦的大核条状卷积 ( $1 \times K$  和  $K \times 1$ ) 动态调整该矩形区域的形状, 使其更紧密地贴合前景物体的实际轮廓, 有效优化垃圾特征提取并抑制复杂背景噪声。接下来, 通过  $3 \times 3$  深度可分离卷积提取输入特征的局部细节, 通过特征融合将校准后的注意力特征加权到精炼的局部特征上, 以同时保留全局上下文与局部细节。最后, 结合批量归一化与多层感知机 (multilayer perceptron, MLP) 进一步提炼特征, 并通过残差连接促进信息流动, 在保持计算轻量化的同时, 显著提升模型对前景目标的定位与表示能力。

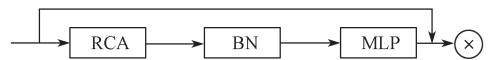


图 4 RCM 组成

Fig. 4 Components of RCM

## 2.4 LSKA 机制

CNN 的大感受野通常通过两种方式获取: 堆叠多层卷积操作或使用大尺寸卷积核。前者可能导致模型规模膨胀, 后者则因内存与计算需求过高而造成成本难以承受。然而, 近期研究表明, 通过连接级联的深度卷积与空洞的深度卷积可等效实现大尺寸卷积核功能, 且不会引发计算量与内存占用的平方级增长。基于大核注意力 (large kernel attention, LKA) 的视觉注意力网络 (visual attention network, VAN) 在图像分类、目标检测及语义分割任务中的表现已超越 Vision Transformers (ViTs) 与 CNN 模型。

但是, LKA 仍存在高计算与内存开销问题, 随着核尺寸增大, 模型效率显著下降。因此, 本研究采用使用深度可分离卷积的 LSKA (见图 5)。LSKA 通过将一个  $k \times k$  的二维卷积核分解为级联的水平一维核 ( $1 \times k$ ) 和垂直一维核 ( $k \times 1$ )。这种分解显著降低了模型的计算复杂度, 从  $O(k^2)$  降低到  $O(2k)$ , 有效解决了 LKA 的计算效率瓶颈。LSKA 在保持大感受野来建模长距离依赖关系的同时, 引导模型产生更强的形状偏置, 提升模型在各种图像损坏和扰动下的鲁棒性<sup>[21]</sup>。

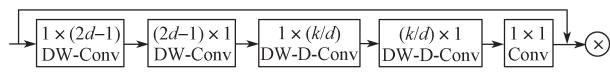


图 5 LSKA 原理

Fig. 5 Principle of LSKA

### 3 实验结果与分析

#### 3.1 数据集

本研究采用两个互补的数据集进行实验验证: VisDrone 基准数据集和自建街景垃圾数据集。VisDrone 作为权威的公开数据集, 提供了用无人机拍摄的 10 209 张高质量图像, 涵盖 10 个通用目标类别, 确保了方法对比的公平性和结果的可复现性; 而自建数据集则是专门针对环保监测场景, 用无人机拍摄的 771 张街景垃圾图像, 包括垃圾袋、烟蒂、包装盒等 5 个类别。两个数据集的结合验证可证明算法在通用场景下的性能, 以及在特定应用任务中的实用价值。



图 6 自建数据集示例图

Fig. 6 Example from custom dataset

#### 3.2 基准模型

为选取适合嵌入式部署的垃圾分类模型, 本研究对轻量级 YOLO 系列模型进行对比评估, 结果如表 1 所示。相较于其他目标检测模型, YOLOv8n 采用 Anchor-Free 机制、C2f 模块和解耦检测头等先进架构, 在检测精度 (mAP@0.5 和 mAP@0.5:0.95 (B)) 上表

现卓越, 同时保持了较低的计算复杂度 (GFLOPs), 在精度与效率之间取得了良好平衡, 并且其能够轻松实现每秒上百帧的检测速度, 契合实时处理需求。因此, 本研究将 YOLOv8n 作为垃圾分类的基准模型, 可以满足垃圾分类应用场景对模型的实时性和准确性的高要求。

表 1 YOLO 系列模型性能对比  
Table 1 Comparative data of YOLO models

| 模型       | mAP@0.5/% | mAP@0.5:0.95 (B)/% | GFLOPs/ $10^9$ |
|----------|-----------|--------------------|----------------|
| YOLOv5n  | 33.05     | 19.44              | 7.2            |
| YOLOv6n  | 30.59     | 18.01              | 11.9           |
| YOLOv8n  | 34.69     | 20.40              | 8.2            |
| YOLOv9n  | 32.32     | 18.97              | 7.9            |
| YOLOv10n | 34.49     | 20.15              | 8.4            |

#### 3.3 消融实验

基于 YOLOv8n, 依次引入 4 个关键模块: SPD-Conv、P2、RCM 和 LSKA。为验证每个模块的独立贡献及其组合效果, 进行了一系列消融实验。所有实验均在相同硬件环境下进行, 以 mAP 作为核心评估指标, 在不同数据集上的消融实验结果如表 2~3 所示。由表 2~3 可知, 所引入的组件均为模型性能带来正向收益, 证明了其设计的有效性。SPD-Conv 是基础, 为后续改进提供了信息更完整的高质量特征图。P2 对小目标检测的提升效果最为显著。RCM 通过优化特征融合路径, 稳健地提升了多尺度检测能力。LSKA 作为“性能助推器”, 通过扩大感受野捕获全局上下文, 最终将模型性能推至顶峰。尽管各项改进在参数量与计算量方面呈现小幅增长, 但当运用 4 个模块共同对 YOLOv8n 进行改进时, 其在检测精度 (mAP@0.5 和 mAP@0.5:0.95 (B)) 方面表现优异。

#### 3.4 基于热力图的模型改进效果

梯度加权类激活映射 (gradient-weighted class activation mapping, Grad-CAM) 是一种用于可视化 CNN 决策过程的关键技术, 其通过生成类激活热力图, 揭示模型在图像识别任务中所关注的区域。而漏检率作为一项定量指标, 可用于评价目标检测算法的有效性。为验证算法改进的有效性, 本研究运用 Grad-CAM 热力图分析与漏检率相结合的方式, 对改进前后的模型在不同数据集上的目标检测效果展开分析, 结果如图 7~8 所示。图中, 箭头指向小目标高响应区域对比, 虚线框为背景噪声抑制区域对比。

**表 2 在 VisDrone 数据集上的消融实验数据****Table 2 Ablation study data from the VisDrone dataset**

| SPD-Conv | P2 | RCM | LSKA | mAP@0.5/% | mAP@0.5:0.95 (B)/% |
|----------|----|-----|------|-----------|--------------------|
|          |    |     |      | 34.69     | 20.40              |
| ✓        |    |     |      | 36.72     | 22.43              |
| ✓        | ✓  |     |      | 37.86     | 23.03              |
| ✓        | ✓  | ✓   |      | 38.62     | 23.33              |
| ✓        |    |     |      | 38.14     | 23.17              |
| ✓        |    | ✓   |      | 37.92     | 22.89              |
| ✓        | ✓  |     |      | 37.34     | 22.50              |
| ✓        |    |     |      | 35.24     | 20.80              |
|          |    |     | ✓    | 35.28     | 21.05              |
| ✓        | ✓  | ✓   | ✓    | 38.91     | 23.62              |



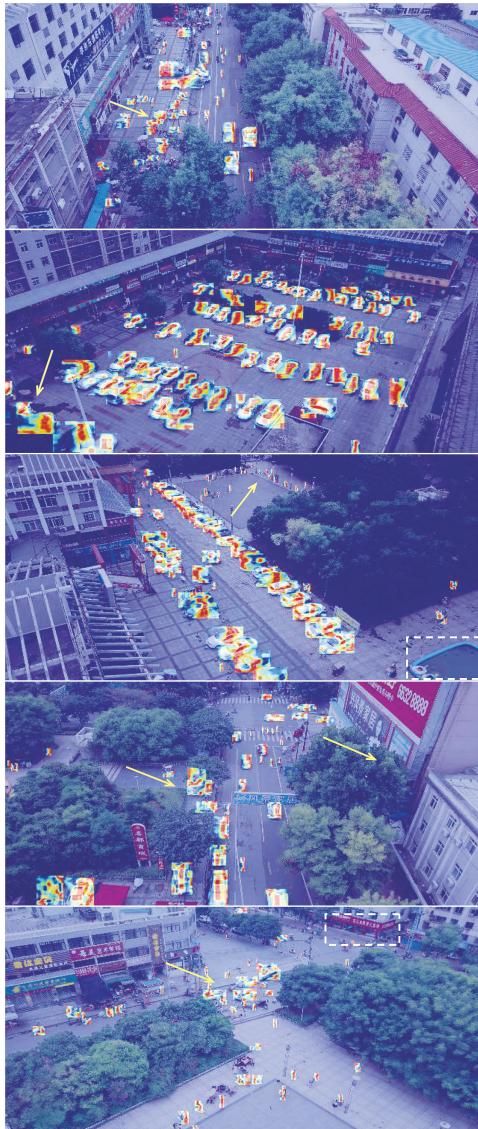
a) 原 YOLOv8n

图 7 改进前后模型在 VisDrone 数据集上对同一复杂场景的热力图对比

Fig. 7 Heatmap comparison of the same complex scene before and after model improvement from the VisDrone dataset

**表 3 在自建数据集上的消融实验数据****Table 3 Ablation study data from the custom dataset**

| SPD-Conv | P2 | RCM | LSKA | mAP@0.5/% | mAP@0.5:0.95 (B)/% |
|----------|----|-----|------|-----------|--------------------|
|          |    |     |      | 31.29     | 15.54              |
| ✓        |    |     |      | 35.30     | 16.48              |
| ✓        | ✓  |     |      | 37.97     | 17.21              |
| ✓        | ✓  | ✓   |      | 40.22     | 17.48              |
| ✓        |    |     |      | 42.94     | 19.23              |
| ✓        |    | ✓   |      | 45.88     | 18.22              |
| ✓        | ✓  |     |      | 44.49     | 18.65              |
| ✓        |    |     | ✓    | 35.49     | 16.74              |
|          |    |     | ✓    | 41.31     | 17.29              |
| ✓        | ✓  | ✓   | ✓    | 48.40     | 18.11              |



b) 本模型

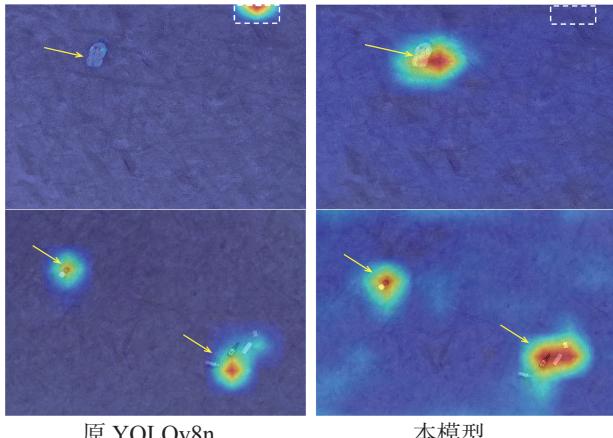


图8 改进前后模型在自建数据集上对同一复杂垃圾场景的热力图对比

Fig. 8 Heatmap comparison of the same complex garbage scene before and after model improvement from the custom dataset

改进后的模型在3个方面实现了提升：1) 注意力集中度的提升。改进前，模型的热力图响应较为弥散，注意力分散在目标物体以及部分背景区域。而改进后，模型的热力图响应则高度集中于待检测的垃圾目标本体，背景噪声得到有效抑制。这表明改进后的模型能够更精准地定位目标的核心特征区域。2) 特征利用效率的增强。对于小目标，改进前的模型响应微弱甚至无响应，直接导致了漏检。改进后，模型对这些小目标生成了清晰且强烈的响应焦点。这证明了改进策略（如增强的特征金字塔或注意力机制）有效提升了模型对小尺度特征的提取与利用能力。3) 对遮挡与复杂背景的鲁棒性。在堆叠、部分遮挡的场景中，改进前的模型可能只关注到物体的可见部分，导致特征不全而漏检。改进后的模型可借助上下文信息对被遮挡部分的轮廓加以推断，构建更为完整的注意力区域，进而实现目标的成功检测。漏检率的降低直接反映了模型“无法识别目标”问题的改善。综上所述，本模型针对原本响应微弱或存在错误的小目标以及遮挡目标构建了更强的注意力机制，进而成功地将这些目标从背景中分离并实现检测。注意力集中度的提高有效减少了误激活现象，使模型不易受到复杂背景的干扰，从而能够将更多的计算资源聚焦于真实目标，进一步降低了因误判背景而引发的漏检情况。

#### 4 结语

针对复杂城市环境中垃圾目标小、分布杂乱导致的检测难题，本研究提出了一种基于YOLOv8n的改

进算法YOLO-SPRL。引入SPD-Conv模块增强小目标特征提取，增加P2检测层提升小目标检测能力，集成RCM模块抑制背景噪声，并应用LSKA强化多尺度特征融合。实验结果表明，本算法在VisDrone数据集上的mAP@0.5达到38.91%，较基准模型提升了4.22%；在自建数据集上的mAP@0.5达到48.40%，较基准模型提升了17.11%。在多种复杂场景中，YOLO-SPRL展现出卓越的鲁棒性与小目标识别能力，在精度、速度以及模型体积之间达成了良好的平衡，适用于嵌入式移动设备的部署，为智能垃圾管理提供了有效的技术解决方案。未来，需进一步探究基于多模态融合的垃圾材质构成识别方法，并以此为依据评估其可回收价值，为后续的回收利用提供支撑。

#### 参考文献：

- [1] MAITY S, CHAKRABORTY T, PANDEY R, et al. YOLO (You Only Look Once) Algorithm-Based Automatic Waste Classification System[J]. Journal of Mechanics and Continuous Mathematics Science, 2023, 18: 25–35.
- [2] ZHANG Y, ZHANG H F, HUANG Q Q, et al. DSP-YOLO: An Anchor-Free Network with DsPAN for Small Object Detection of Multiscale Defects[J]. Expert Systems with Applications, 2024, 241: 122669.
- [3] JIANG C C, REN H Z, YE X, et al. Object Detection from UAV Thermal Infrared Images and Videos Using YOLO Models[J]. International Journal of Applied Earth Observation and Geoinformation, 2022, 112: 102912.
- [4] XUE C, XIA Y L, WU M J, et al. EL-YOLO: An Efficient and Lightweight Low-Altitude Aerial Objects Detector for Onboard Applications[J]. Expert Systems with Applications, 2024, 256: 124848.
- [5] MA W Q, CHEN H, ZHANG W K, et al. DSYOLO-Trash: An Attention Mechanism-Integrated and Object Tracking Algorithm for Solid Waste Detection[J]. Waste Management, 2024, 178: 46–56.
- [6] HE Y H, LI J F. TSRes-YOLO: An Accurate and Fast Cascaded Detector for Waste Collection and Transportation Supervision[J]. Engineering Applications of Artificial Intelligence, 2023, 126: 106997.
- [7] KIM M, JEONG J, KIM S, et al. ECAP-YOLO: Efficient Channel Attention Pyramid YOLO for Small Object Detection in Aerial Image[J]. Remote Sensing, 2021, 13(23): 4851.
- [8] DAS D, DEB K, SAYEED T, et al. Outdoor Trash Detection in Natural Environment Using a Deep Learning Model[J]. IEEE Access, 2023, 11: 97549–97566.
- [9] FAN J H, CUI L Z, FEI S M, et al. Waste Detection System Based on Data Augmentation and YOLO\_Ec[J]. Sensors, 2023, 23(7): 3646.
- [10] YANG Y Z, LI Y X, TAO M H, et al. FE-YOLO: A Lightweight Model for Construction Waste Detection

- Based on Improved YOLOv8 Model[J]. Buildings, 2024, 14(9): 2627.
- [11] REN Y M, LI Y Z, GAO X Y, et al. An MRS-YOLO Model for High-Precision Waste Detection and Classification[J]. Sensors, 2024, 24(13): 4339.
- [12] LIU C H, XIE N, YANG X X, et al. A Domestic Trash Detection Model Based on Improved YOLOX[J]. Sensors, 2022, 22(18): 6974.
- [13] 朱泽宇, 肖满生, 徐萌, 等. 一种改进 YOLOv8n 的 PCB 板表面缺陷检测算法 [J]. 湖南工业大学学报, 2026, 40(1): 92–101.
- [14] 王晓红, 张微. MNTH-YOLOv8: 一种用于食品包装中蚊虫高效检测的深度学习方法 [J]. 包装学报, 2024, 16(3): 91–98.
- [15] 吴林, 曹雯. YOLO-SCDI: 基于改进 YOLOv8 的车辆检测算法 [J]. 电子与信息学报, 2025, 47(10): 3967–3978.
- [16] 范博淦, 王淑青, 陈开元. 基于改进 YOLOv8 的航拍无人机小目标检测模型 [J]. 计算机应用, 2025, 45(7): 2342–2350.
- [17] LIU Y W, WU Y Z, CHEN L J, et al. Research on Urban Road Obstacle Detection Method Based on ERLM-YOLOv8[C]//Advanced Intelligent Computing Technology and Applications. Singapore: Springer, 2025: 462–473.
- [18] SUNKARA R, LUO T. No More Strided Convolutions or Pooling: A New CNN Building Block For Low-Resolution Images and Small Objects[C]//Machine Learning and Knowledge Discovery in Databases. Cham: Springer, 2023: 443–459.
- [19] LIU S L, CHEN R H, YE M H, et al. EcoDetect-YOLO: A Lightweight, High-Generalization Methodology for Real-Time Detection of Domestic Waste Exposure in Intricate Environmental Landscapes[J]. Sensors, 2024, 24(14): 4666.
- [20] NI Z L, CHEN X H, ZHAI Y J, et al. Context-Guided Spatial Feature Reconstruction for Efficient Semantic Segmentation[C]//European Conference on Computer Vision. Cham: Springer, 2025: 239–255.
- [21] LAU K W, PO L M, REHMAN Y A U. Large Separable Kernel Attention: Rethinking the Large Kernel Attention Design in CNN[J]. Expert Systems with Applications, 2024, 236: 121352.

(责任编辑: 邓彬)

## Improved YOLOv8 Structure for Detecting Small Garbage Objects and Its Performance Validation

HE Dawei, WU Yuezhong, CHEN Lingjiao, HAN Mei

( School of Transportation and Electrical Engineering, Hunan University of Technology, Zhuzhou Hunan 412007, China )

**Abstract:** To address the challenges in garbage detection in complex urban environments, particularly the difficulties caused by small target sizes and cluttered backgrounds, an improved algorithm named YOLO-SPRL is proposed for small target garbage detection based on YOLOv8. First, the SPD-Conv module is incorporated and a P2 detection layer is added while the P5 detection layer is removed to enhance the detection capability for small garbage objects. Second, a rectangular self-calibration module (RCM) is integrated into the P3 layer of the feature pyramid to suppress background noise and eliminate interference from complex backgrounds. Finally, a learnable separable kernel attention (LSKA) module is introduced in the model's neck, which reduces computational complexity while strengthening deep feature fusion capability. This optimization reduces model size while maintaining improved accuracy, laying the theoretical groundwork for deployment on embedded devices. Experiments on the VisDrone and custom street view garbage datasets demonstrate that YOLO-SPRL achieves an mAP@0.5 of 38.91%, representing an improvement of 4.22% compared to the baseline model. Additionally, the algorithm exhibits excellent robustness and strong small target recognition capability on a custom drone garbage dataset. The integrated improvement strategy proposed in this study effectively enhances garbage detection performance in complex scenarios. YOLO-SPRL achieves an optimal balance among accuracy, speed, and model size, demonstrating the feasibility for deployment on embedded mobile devices, and providing a reliable technical solution for real-time monitoring and management in urban intelligent sanitation systems.

**Keywords:** YOLOv8; small target detection; rectangular self-calibration module; large separable kernel attention