

DOI: 10.13382/j.jemi.B2407898

# 面向复杂背景环境下垃圾检测的 YOLOv8n 轻量化改进<sup>\*</sup>

孙世政<sup>1</sup> 何玲玲<sup>1</sup> 郑 帅<sup>2</sup> 徐向阳<sup>1</sup> 陈仁祥<sup>1</sup>

(1. 重庆交通大学机电与车辆工程学院 重庆 400074; 2. 重庆交通大学航运与船舶工程学院 重庆 400074)

**摘要:**垃圾检测与分类对推动绿色经济和实现低碳循环具有重要意义,面向复杂背景环境的垃圾检测模型存在参数量大、计算成本高等问题,限制了模型在资源受限设备上的应用。为解决上述问题,提出一种轻量化的 GCAW-YOLOv8n 模型,旨在平衡模型轻量化与精度检测。首先,在 YOLOv8n 骨干网络中引入 GhostNet 网络中的 C3Ghost 和 GhostConv 模块,有效降低模型参数量;其次,添加上下文锚点注意力机制,增强特征提取能力,提升检测精度;然后,在特征融合阶段,构建渐近特征金字塔网络,提升多尺度目标检测能力;接着,采用 WIoU v3 边界损失函数优化网络边界框回归性能;最后,结合 Taco 数据集和人工采集数据集进行了模型验证实验。实验结果表明,相比原 YOLOv8n 模型,改进后的 GCAW-YOLOv8n 模型在模型参数量 Params 和计算量 FLOPs 分别降低了 14.3% 和 33.3%,而精确度和召回率分别提高了 4.4% 和 1.9%,同时 mAP@0.5 达到了 81.3%,提升了 0.7%。改进模型更好地平衡了模型轻量化和检测精度,对模型部署与应用至边缘端检测装备具有重要的工程意义。

**关键词:**垃圾检测;轻量化 YOLOv8n; GhostNet; 上下文锚点注意力机制; 渐近特征金字塔

**中图分类号:** TP391.41; TN912 **文献标识码:** A **国家标准学科分类代码:** 520.2060

## Lightweight improvement of YOLOv8n for garbage detection in complex background environments

Sun Shizheng<sup>1</sup> He Lingling<sup>1</sup> Zheng Shuai<sup>2</sup> Xu Xiangyang<sup>1</sup> Chen Renxiang<sup>1</sup>

(1. School of Mechatronics and Vehicle Engineering, Chongqing Jiaotong University, Chongqing 400074, China;

2. School of Shipping and Naval Architecture, Chongqing Jiaotong University, Chongqing 400074, China)

**Abstract:** Garbage detection and classification are essential for promoting the green economy and achieving a low-carbon circular economy. However, current models face challenges such as large parameters and high computational costs, limiting their deployment on resource-constrained devices. To address these issues, a lightweight GCAW-YOLOv8n model is proposed that balances model size and detection accuracy. Firstly, the C3Ghost and GhostConv modules from GhostNet are integrated into the YOLOv8n backbone to reduce parameters. Secondly, the context anchor attention is introduced to enhance feature extraction and detection accuracy. Then, the asymptotic feature pyramid network is used to improve multi-scale detection, and the WIoU v3 loss function optimizes bounding box regression. Finally, the improved model is validated using the Taco dataset and a custom dataset. Experimental results show that, compared with the original YOLOv8n model, the GCAW-YOLOv8n model reduces parameters by 14.3% and floating-point operations by 33.3%, while precision and recall increase by 4.4% and 1.9%, respectively. The mAP@0.5 improves to 81.3%, a 0.7% gain. This model achieves a better balance between lightweight design and detection accuracy, making it suitable for deployment in edge devices for garbage detection.

**Keywords:** garbage detection; lightweight YOLOv8n; GhostNet; context anchor attention; asymptotic feature pyramid network

## 0 引言

垃圾检测与分类对推动绿色经济、保障人类健康和实现低碳循环具有重要意义<sup>[1]</sup>。结合机器学习和深度学习算法的检测技术广泛应用于垃圾检测领域。二阶段检测方法存在参数量大、速度慢等缺点,难满足垃圾检测模型在边缘端设备部署的实时性需求<sup>[2-4]</sup>。现阶段垃圾检测的常用模型为一阶段检测方法 YOLO 系列,具有简单高效的特点,适用于边缘端检测装备<sup>[5-7]</sup>。

目标检测模型精度提升通常伴随计算复杂度增加,限制了模型在资源受限边缘端设备上的应用。因此,需在模型轻量化和检测精度之间找到平衡,以满足边缘设备对实时性和计算资源的要求。为解决检测模型在边缘端设备的部署问题,国内外研究人员提出了许多轻量化改进方法。董兆苒等<sup>[8]</sup>提出基于注意力机制和混合膨胀卷积网络的 YOLOv4-tiny 改进模型,实现对小目标特征提取能力和检测精度的提升,但未考虑模型轻量化问题;曾勇杰等<sup>[9]</sup>提出了高效多尺度卷积模块 EMSConv 替代普通卷积 Conv,实现模型参数量和计算量的小幅下降;Ma 等<sup>[10]</sup>将 YOLOv8n 的骨干网络替换为 ShuffleNetv2 网络,在颈部网络引入 Ghost 模块,从而实现模型的轻量化,但相较基线模型平均精度均值 (mAP) 降低了 4%;胡久松等<sup>[11]</sup>将 CBAM 注意力机制融合到 GhostNet 轻量级网络结构中,并在 YOLOv8 的骨干网络中引入卷积块注意力模块 (CBAM),加快了检测速度,但改进后的模型体积仍较大;Zhao 等<sup>[12]</sup>采用轻量级网络 GhostHGNetv2 改进 YOLOv8n 的骨干特征提取网络,虽然模型参数量减少,但 mAP 降低了 1.1%;Wu 等<sup>[13]</sup>通过引入 SEConv 模块改进 YOLOv8s 网络,改进后的模型以增加模型参数量为代价提升了检测精度;曲晨阳等<sup>[14]</sup>提出了添加双向特征金字塔网络 BiFPN 到 YOLOv8s 颈部网络,参数量和计算量小幅降低,相较于 YOLOv8n 略大;肖立中等<sup>[15]</sup>通过结合 GhostConv 卷积模块改进 YOLOv5s,提出了一种轻量化垃圾检测模型 LGD-YOLO,改进后的模型参数量减少了 22.4%,但数据集拍摄环境较为简单。

在垃圾目标检测任务中所针对的复杂背景环境是指包含多种元素、不规则结构和不同光线的场景,例如植被、土壤、路面等,在这些情况下可能混淆垃圾与背景,从而影响检测精度。目前,垃圾检测仍存在以下难题:1) 垃圾数据集背景环境缺乏多样性;2) 目标检测的模型参数量与检测精度难以达到平衡。针对上述问题,通过采用 Taco 数据集<sup>[16]</sup>部分图片以及人工采集图片建立背景多样化的数据集,更加贴合实际应用场景;提出一种基于 YOLOv8n 轻量化改进的垃圾目标检测模型,实现模型轻量化与检测精度的平衡,有利于模型在边缘端设备的部

署与应用。

## 1 YOLOv8n 网络结构

与其他 YOLO 系列目标检测模型相比,YOLOv8 具有检测精度较高,检测速度较快的优点,并且 YOLOv8 拥有自适应锚框和多尺度特征融合机制,在处理尺寸不一、背景复杂的垃圾目标时更具优势。YOLOv8 根据模型尺寸的不同,分为 YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l、YOLOv8x 5 个模型,按照顺序模型的深度和宽度依次增加,检测精度逐步提升,但训练和检测所花费的时间也随之增多,为实现模型轻量化,选用 YOLOv8n 为基准模型。

YOLOv8n 的网络结构主要由输入端 (input)、骨干网络 (backbone)、颈部网络 (neck) 及检测头 (head) 4 个部分构成,如图 1 所示。Input 层将图像数据输入到模型,采用 Mosaic 数据增强方法,提升模型的泛化能力,Backbone 层对输入图像进行特征提取,构建深层次特征提取器;Neck 层整合调整骨干网络提取的特征,融合不同层次的特征,提升网络感知能力;Head 层对特征进行预测,将目标进行定位和分类,获取检测结果。

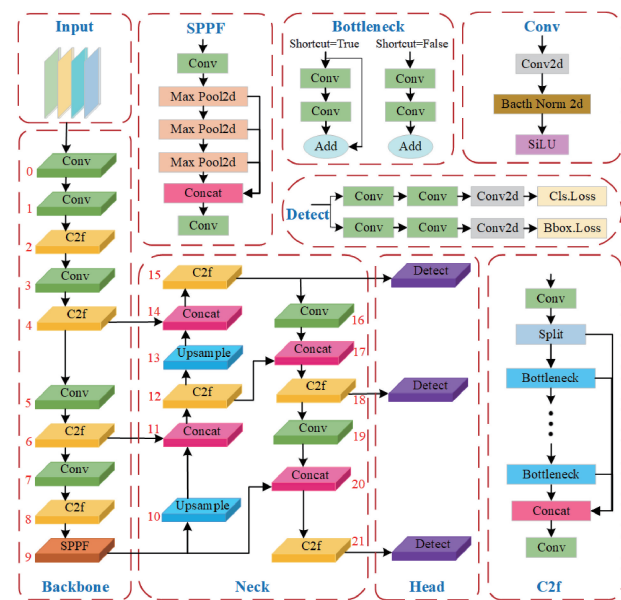


图1 YOLOv8n 网络结构

Fig.1 YOLOv8n network structure

## 2 YOLOv8n 轻量化改进

为实现复杂背景环境下垃圾检测的轻量化,对 YOLOv8n 模型结构进行改进。首先,使用 C3Ghost 和 GhostConv 模块分别替换骨干网络中的普通卷积模块与 C2f 模块,增强模型的表达能力,有效减少模型参数量,

引入上下文锚点注意力机制 (context anchor attention, CAA), 捕获远程上下文信息, 提升检测精度; 其次, 在模型 Neck 部分使用渐近特征金字塔网络 (asymptotic feature pyramid network, AFPN) 进一步减少模型浮点运算数, 提高模型的特征融合能力, 进而提升垃圾目标检测精度; 最后, 采用 WIoU v3 损失函数替代 CIoU 损失函数, 提升网络边界框回归性能和垃圾目标检测效果。改进模型网络结构如图 2 所示。

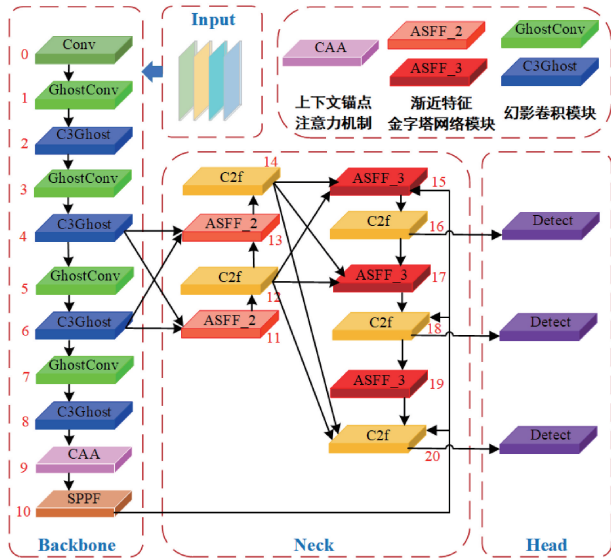


图 2 GCAW-YOLOv8n 网络结构

Fig. 2 GCAW-YOLOv8n network structure

## 2.1 GhostNet 模块

GhostNet 网络的核心思想是使用计算强度较低的 GhostConv 模块生成神经网络的冗余特征图<sup>[17]</sup>。采用 GhostNet 结构可以降低模型参数数量和计算复杂度, 同时降低冗余信息的产生。普通卷积模块通过卷积映射所有通道提取特征信息, 导致参数数量过大, 其结构如图 3(a) 所示。

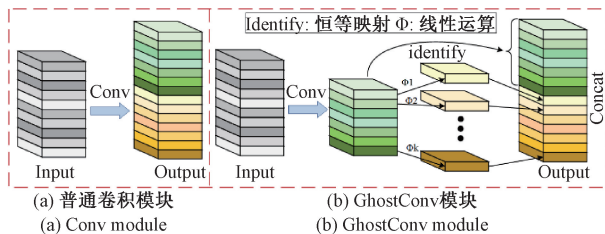


图 3 不同卷积模块网络结构

Fig. 3 Network structure of different Conv modules

GhostConv 卷积模块将普通卷积模块的运算转换为两步, 第 1 步使用部分卷积核对输入端进行特征提取, 得到通道数较少的特征图; 第 2 步将获取的特征图进行轻

量级线性运算, 得到另一部分特征图, 最后将恒等映射和线性运算得到的两部分特征图进行 Concat, 完成输出, GhostConv 卷积模块如图 3(b) 所示。

假设图 3 中输入特征图的大小为  $H \times W \times C$ , 输出特征图的大小为  $H' \times W' \times N$ , 普通卷积核的大小为  $k \times k$ , 各线性卷积核大小为  $d \times d$ , 大小与  $k \times k$  相似。普通卷积的计算量为  $H' \times W' \times N \times C \times k \times k$ , GhostConv 的计算量为  $N/s \times H' \times W' \times N \times C \times k \times k + N/s \times H' \times W' \times (s-1) \times d \times d$ 。

为了确保与普通卷积输出大小相同, GhostConv 卷积第 2 步是对第 1 步输出的特征图进行轻量级线性变换, 如式 (1) 所示。

$$y_{ij} = \Phi(y'_i), \forall i = 1, \dots, m; j = 1, \dots, s \quad (1)$$

式中:  $y'_i$  为第  $i$  个特征映射;  $y_{ij}$  为通过第  $i$  个特征映射的线性运算获得的第  $j$  个特征映射;  $\Phi$  表示线性运算。

将普通卷积与 GhostConv 卷积的计算量进行比值, 可以计算出理论比率  $r_s$ :

$$r_s = \frac{H' \cdot W' \cdot N \cdot C \cdot k \cdot k}{N/s \cdot H' \cdot W' \cdot C \cdot k \cdot k + N/s \cdot H' \cdot W' \cdot (s-1) \cdot d \cdot d} \approx s \quad (2)$$

式中:  $s$  为线性变换的次数且  $s \ll C$ ;  $H$  为输入特征图的高度;  $W$  为输入特征图的宽度;  $C$  为输入特征图的通道数;  $H'$  为输出特征图的高度;  $W'$  为输出特征图的宽度;  $N$  为输出特征图的通道数。

根据式 (2) 可以得出, 普通卷积的计算量为 GhostConv 卷积的  $s$  倍。因此, 基于 GhostConv 卷积设计了 GhostBottleneck 模块, 如图 4(a) 所示, DWConv 表示深度可分离卷积。结合 GhostBottleneck 和 GhostConv 模块的优势, 引入一个轻量级的特征提取结构 C3Ghost, 如图 4(b) 所示, 该模块由卷积层和  $n$  个线性堆叠的 GhostBottleneck 组成, 该结构有效保留原始图像的特征信息, 避免特征丢失。

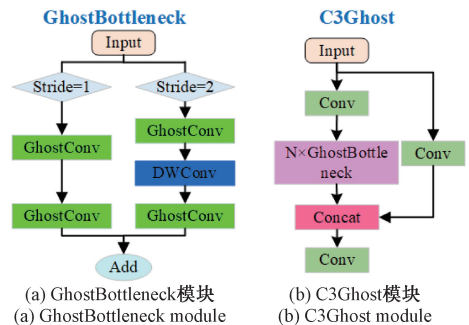


图 4 Ghost 构建块

Fig. 4 Ghost building blocks

## 2.2 CAA 上下文锚点注意力机制

上下文锚点注意力机制 CAA<sup>[18]</sup> 是一种新颖的注意

力机制,通过捕获长距离上下文信息增强中心特征表示,提高目标检测精度。如图5所示,该注意力机制分为3个工作流程:首先,使用平均池化 (average pooling, AvgPool) 和普通卷积操作 Conv 获取局部区域特征;其次,利用两个深度可分离卷积 (depthwise separable convolution, DWConv) 实现轻量化的长距离依赖捕捉;最后,通过生成注意力权重图增强特征表示。CAA 注意力机制主要目的是提高模型在垃圾检测任务中对长距离依赖关系的捕捉能力,增强复杂场景下的垃圾检测性能。

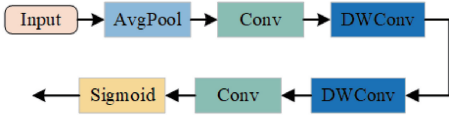


图5 CAA 注意力机制

Fig. 5 Context anchor attention mechanism

通过在 YOLOv8n 模型骨干网络中引入 CAA 注意力机制,使模型在不显著增加计算成本的情况下,更精准地识别多尺度目标,提升检测的鲁棒性和精度。

### 2.3 渐近特征金字塔网络 (AFPN)

原 YOLOv8n 模型采用路径聚合网络 (path aggregation network, PAN) 在目标检测任务中实现多尺度特征融合。经典 PAN 结构难以满足精度要求,易出现漏检和误检;此外,高级特征在与底层低级特征的传播和交互过程中,易出现特征信息丢失或退化,降低非相邻层次的特征融合效果。

改进模型采用 AFPN<sup>[19]</sup>,如图6所示,其中实线箭头表示卷积过程,虚线箭头表示自适应空间融合。渐进式融合过程包括3个阶段:初始阶段融合两个底层特征;后续阶段融合深层特征;最后阶段添加顶层特征。由于非相邻层特征间存在较大的语义差距,尤其是底层和顶层,因此通过渐近特征融合有效缩小语义差距。

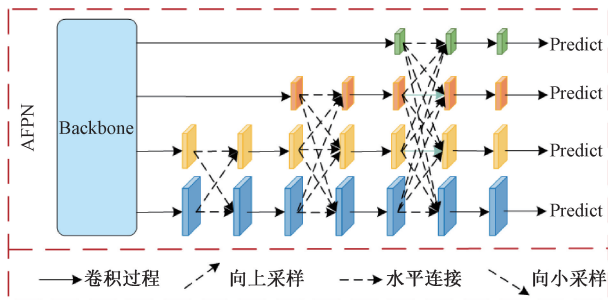


图6 渐近特征金字塔网络结构

Fig. 6 Asymptotic feature pyramid network structure

### 2.4 WIoU v3 损失函数模块

目标检测过程中通常采用交并比 (IoU)<sup>[20]</sup> 衡量预测结果与真实标注区域的相似度,即交集和并集之比。

IoU 原理如图7所示,其中黄色和蓝色矩形分别表示锚框和目标框,  $W_i$  和  $H_i$  分别为锚框和目标框相交的宽和高,  $W_g$  和  $H_g$  分别为锚框和目标框的外矩形宽和高,  $(x, y)$  和  $(x_{gt}, y_{gt})$  分别为锚框和目标框的中心点坐标。

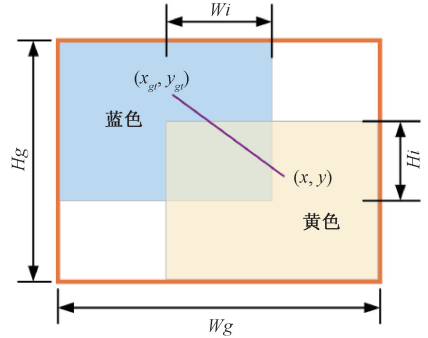


图7 IoU 原理

Fig. 7 IoU schematic

原 YOLOv8n 网络模型采用 CIoU 损失函数计算边界框的回归损失,由于 CIoU 为单调聚焦机制,缺乏对难样本的平衡考虑,导致纵横比和距离等几何指标放大了低质量样本的负梯度,进而影响模型泛化能力。

因此,改进模型采用具有动态非单调聚焦机制的 WIoU v3 损失函数<sup>[21]</sup> 替换 CIoU 损失函数。WIoU v3 损失函数计算公式如式(3)~(7)所示。

$$L_{WIoUv3} = rR_{WIoUv3}L_{IoU} \quad (3)$$

$$r = \frac{\beta}{\delta\alpha^{\beta-\delta}} \quad (4)$$

$$\beta = \frac{L_{IoU}^*}{L_{IoU}} \in [0, +\infty) \quad (5)$$

$$R_{WIoUv3} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*}\right) \quad (6)$$

$$L_{IoU} = 1 - IoU \quad (7)$$

式中:  $r$  为非单调聚焦系数;  $R_{WIoUv3} \in [1, e)$  为惩罚项;  $L_{IoU}^*$  为单调聚焦系数;  $L_{IoU} \in [0, 1]$  为 IoU 损失;  $\beta$  为离群度;  $\delta$  为梯度增益;  $\alpha$  为平衡因子;  $L_{IoU}$  为归一化因子,即动量的滑动平均值;上标 \* 表示进行分离操作。

由式(4)可知,当离群度  $\beta$  为定值时,预测框将获得最高梯度增益,因此 WIoU v3 在训练过程中使用最优的梯度增益动态分配策略,进而提升模型整体性能。

## 3 实验数据与评价指标

### 3.1 实验数据集

结合 Taco 数据集和人工采集图片开展相关实验,合成包含 2 482 张图片的垃圾分类数据集。Taco 数据集的背景复杂多样,包括室内和室外环境,如瓷砖、人行



道、草地、道路等。合成的数据集涵盖了 8 种常见垃圾(餐盒、易拉罐、果皮、垃圾袋、烟蒂、塑料瓶、瓶盖、纸盒),每张图片的物品标注数量不同,更贴合实际环境,如图 8 所示。

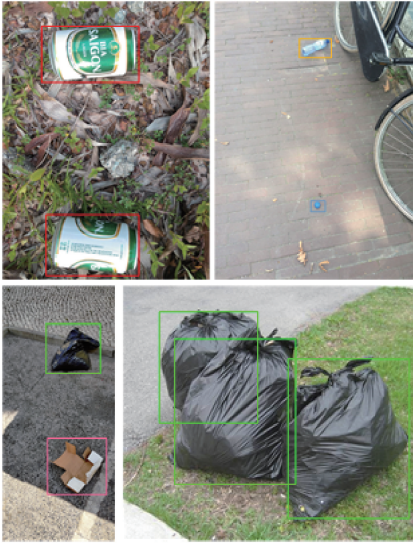


图 8 部分数据集图片

Fig. 8 Some sample images of the dataset

为验证改进模型效果,将数据集按 8 : 1 : 1 比例划分为训练集、验证集和测试集,即 1 986 张图片用于训练,248 张图片用于验证,248 张用于测试。各样本数量如图 9 所示。

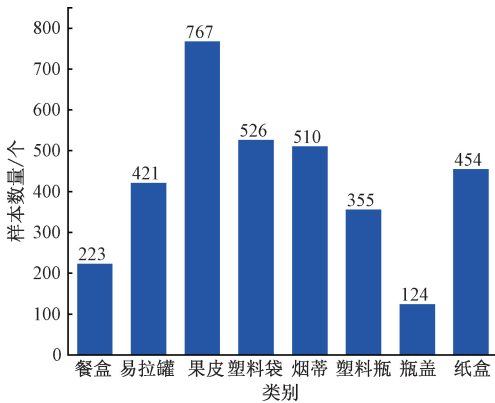


图 9 数据集各类别样本数量

Fig. 9 Number of samples per category in the dataset

### 3.2 评价指标

采用精确度 (precision,  $P$ )、召回率 (recall,  $R$ )、平均精度 (average precision,  $AP$ )、平均精度均值 (mean average precision,  $mAP$ )、每秒检测帧数 (frames per second, FPS)、参数量 (parameters, Params)、浮点运算数 (floating-point operations, FLOPs) 和模型大小 (Size) 作

为评价指标。每秒内可以检测的图片数量是衡量模型检测速度的指标;参数量 Params 表示一个模型包含参数的数量;浮点运算数指每秒浮点运算次数,理解为计算量,用于衡量模型复杂度;模型大小 Size 表示模型所占的存储空间大小,单位为 MB。上述指标的计算式分别为:

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

$$AP = \int_0^1 P(R) dR \quad (10)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (11)$$

式中:  $TP$  为正确检测的样本数;  $FP$  为检测错误的样本数;  $FN$  为漏检的样本数;  $AP$  的值为 Precision-Recall 曲线和坐标轴所围成的面积;  $mAP$  是对所有类别的  $AP$  求均值;  $AP_i$  为数据集某个类别的平均精度;  $N$  为类别数,本文取值为 8。

## 4 实验结果与分析

### 4.1 实验配置

实验所有模型训练环境如下: CPU 为 Intel Xeon Platinum 8255C CPU@2.50 GHz, GPU 为 NVIDIA GeForce RTX2080Ti, 内存为 40 GB, 深度学习框架采用 PyTorch2.0, CUDA 版本为 11.8。输入图像大小为 640×640, 所有实验均训练 300 个 epoch, 批量大小设置为 16, IoU 阈值设置为 0.5, 初始学习率设置为 0.01, 动量为 0.937, 并使用权重衰减率为 0.0005 的随机梯度下降优化器 SGD。

### 4.2 轻量化对比实验

选用 GhostNet、ShuffleNetv2<sup>[22]</sup>、MobileNetv3<sup>[23]</sup> 3 种轻量级网络分别加入到骨干网络进行对比实验, GhostNet 包含 C3Ghost 和 GhostConv 两个模块, 实验结果如表 1 所示, 将 YOLOv8n 骨干网络中的 C2f 模块替换为 C3Ghost 模块, 模型参数量、模型大小和计算量分别减少  $0.52 \times 10^6$ 、1 MB 和 1.4 GFLOPs,  $mAP@0.5$  值减少 3.2%; 引入 GhostConv 模块替代 Conv 模块, 模型参数量、模型大小和计算量分别减少  $0.19 \times 10^6$ 、0.4 MB 和 0.3 GFLOPs,  $mAP@0.5$  值减少 2.6%; GhostNet 引入骨干网络时检测精度更高,  $mAP@0.5$  值分别比 ShuffleNetv2、MobileNetv3 高 1.8% 和 9.8%; 同时, 参数量、模型大小和计算量分别比 YOLOv8n 网络结构低  $0.71 \times 10^6$ 、1.4 MB 和 1.9 GFLOPs, 并且提高了检测速度。实验结果显示, 采用 MobileNetv3 时模型参数量、模型大小和计算量最少, 检测速度最快, 其参数量分别是 ShuffleNetv2 的 0.65 倍、GhostNet 的

0.52 倍,但是其 mAP@0.5 值最低,没有很好的平衡检测精度与模型轻量化。基于实验结果分析,综合考虑模型轻量化、检测速度和精度,选择 GhostNet 作为骨干网络更为合适。

表 1 不同骨干网络的对比实验结果

Table 1 Experimental results of different backbone networks

| 模型                    | P/%  | R/%  | Params/( $\times 10^6$ ) | Size/MB | 浮点数/GFLOPs | mAP@0.5/% | 帧率/FPS |
|-----------------------|------|------|--------------------------|---------|------------|-----------|--------|
| YOLOv8n( Baseline)    | 89.3 | 72.9 | 3.01                     | 6.3     | 8.1        | 80.6      | 118    |
| Baseline+C3Ghost      | 88.1 | 67.2 | 2.49                     | 5.3     | 6.7        | 77.4      | 126    |
| Baseline+GhostConv    | 83.5 | 69.7 | 2.82                     | 5.9     | 7.8        | 78        | 120    |
| Baseline+GhostNet     | 86.9 | 68.1 | 2.30                     | 4.9     | 6.2        | 78.3      | 130    |
| Baseline+ShuffleNetv2 | 79.5 | 69.7 | 1.83                     | 3.9     | 5.1        | 76.5      | 124    |
| Baseline+Mobilenetv3  | 76.9 | 64.2 | 1.19                     | 2.7     | 2.8        | 68.5      | 143    |

为验证改进模型在特征提取和特征融合方面的有效性,将 YOLOv8n 的特征提取骨干结构中的普通卷积 Conv 模块和 C2f 模块分别替换为 GhostConv 模块和 C3Ghost 模块;并将特征融合结构替换为 AFPN 结构。实验结果如表 2 所示。

实验结果显示,引入 GhostNet 结构后,模型的参数

量、模型大小以及计算量分别减少了  $0.71\times 10^6$ 、1.4 MB 以及 1.9 G FLOPs。特征融合结构替换为 AFPN 结构后,降低了计算复杂度,与原始 YOLOv8n 的结构相比,结合轻量级网络的 YOLOv8n 网络结构明显更加紧凑,参数量、模型大小以及计算量分别减小了  $0.57\times 10^6$ 、1 MB 以及 2.8 G FLOPs。

表 2 轻量化对比实验结果

Table 2 Comparative experimental results of lightweighting

| 模型                     | P/%  | R/%  | Params/( $\times 10^6$ ) | Size/MB | 浮点数/GFLOPs | mAP@0.5/% | 帧率/FPS |
|------------------------|------|------|--------------------------|---------|------------|-----------|--------|
| YOLOv8n( Baseline)     | 89.3 | 72.9 | 3.01                     | 6.3     | 8.1        | 80.6      | 118    |
| Baseline+GhostNet      | 86.9 | 68.1 | 2.30                     | 4.9     | 6.2        | 78.3      | 130    |
| Baseline+GhostNet+AFPN | 86.1 | 68.8 | 2.44                     | 5.3     | 5.3        | 78.9      | 130    |

4.3 注意力机制对比实验

为验证 CAA 注意力机制的优越性,在骨干网络引入 GhostNet 结构的基础上,与同一实验条件下的全局注意力机制<sup>[24]</sup>(global attention mechanism, GAM)、CBAM 模块<sup>[25]</sup>、高效多尺度注意力机制<sup>[26]</sup>(efficient multi-scale attention, EMA)、高效通道注意力机制<sup>[27]</sup>(efficient channel attention, ECA)和压缩与激励网络<sup>[28]</sup>(squeeze-and-excitation networks, SE)进行对比实验。实验结果显

示如表 3 所示,添加 CAA 注意力机制后,mAP@0.5 相比基准模型提升了 2.1%,添加 GAM 注意力机制后,模型参数量、模型大小以及计算量分别增加  $1.64\times 10^6$ 、3.3 MB、1.3 GFLOPs,同时 mAP@0.5 下降 0.5%,不利于轻量化部署。其余注意力机制虽然在参数量、模型大型及计算量上与 CAA 相当,但检测效果均不如 CAA,注意力机制对比实验表明添加 CAA 模块后的模型更好的平衡检测精度和轻量化。

表 3 注意力机制对比实验结果

Table 3 Comparison of attention mechanisms with experimental results

| 模型                          | P/%  | R/%  | Params/ $\times 10^6$ | Size/MB | 浮点数/GFLOPs | mAP@0.5/% |
|-----------------------------|------|------|-----------------------|---------|------------|-----------|
| YOLOv8n+GhostNet( Baseline) | 86.9 | 68.1 | 2.30                  | 4.9     | 6.2        | 78.3      |
| Baseline+GAM                | 85   | 70.9 | 3.94                  | 8.2     | 7.5        | 77.8      |
| Baseline+CBAM               | 88.2 | 70.3 | 2.30                  | 4.9     | 6.3        | 77.5      |
| Baseline+EMA                | 85.3 | 70.1 | 2.31                  | 4.9     | 6.3        | 78.5      |
| Baseline+ECA                | 86.7 | 67.7 | 2.30                  | 4.9     | 6.3        | 76.4      |
| Baseline+SE                 | 79.3 | 72.5 | 2.31                  | 4.9     | 6.2        | 77.9      |
| Baseline+CAA                | 83.6 | 74.9 | 2.43                  | 5.2     | 6.3        | 80.4      |

4.4 损失函数对比实验

为验证 WIoU v3 损失函数有效性,在引入 GhostNet、

CAA 和 AFPN 结构的模型基础上将 CIoU、SIoU 以及 WIoU 损失函数做对比实验。

实验结果如表 4 所示,相较于 CIoU 损失函数,采用 WIoU v3 损失函数后 mAP@0.5、精确度和召回率分别提高了 0.4%、0.9% 和 0.8%,同时检测效果也优于其它损失函数,mAP@0.5 分别比 SIoU、WIoU v1 和 WIoU v2 高 11.2%、9%、5.6%。图 10 所示为损失函数实验结果对比。

表 4 损失函数对比实验结果

| Table 4 Experimental results of the loss function |      |      |                          |           |
|---------------------------------------------------|------|------|--------------------------|-----------|
| 模型                                                | P/%  | R/%  | Params/( $\times 10^6$ ) | mAP@0.5/% |
| CIoU                                              | 84   | 74   | 2.58                     | 80.9      |
| SIoU                                              | 78.2 | 59.5 | 2.58                     | 70.1      |
| WIoU v1                                           | 80.4 | 62.9 | 2.58                     | 72.3      |
| WIoU v2                                           | 75.8 | 68.6 | 2.58                     | 75.7      |
| WIoU v3                                           | 84.9 | 74.8 | 2.58                     | 81.3      |

4.5 消融实验

为验证所有改进模块的有效性,以原 YOLOv8n 网络

表 5 消融实验

| Table 5 Ablation experiments |          |          |     |      |         |      |      |                          |            |           |
|------------------------------|----------|----------|-----|------|---------|------|------|--------------------------|------------|-----------|
| 模型                           | Baseline | GhostNet | CAA | AFPN | WIoU v3 | P/%  | R/%  | Params/( $\times 10^6$ ) | 浮点数/GFLOPs | mAP@0.5/% |
| Baseline                     | ✓        |          |     |      |         | 89.3 | 72.9 | 3.01                     | 8.1        | 80.6      |
| A                            | ✓        | ✓        |     |      |         | 86.9 | 68.1 | 2.3                      | 6.2        | 78.3      |
| B                            | ✓        | ✓        | ✓   |      |         | 83.6 | 74.9 | 2.43                     | 6.3        | 80.4      |
| C                            | ✓        | ✓        |     | ✓    |         | 86.1 | 68.8 | 2.44                     | 5.3        | 78.9      |
| D                            | ✓        | ✓        | ✓   | ✓    |         | 84   | 74   | 2.58                     | 5.4        | 80.9      |
| 本文                           | ✓        | ✓        | ✓   | ✓    | ✓       | 84.9 | 74.8 | 2.58                     | 5.4        | 81.3      |

实验结果显示,采用 GhostNet 网络结构改进 YOLOv8n 的骨干网络后,模型参数量和计算量分别从  $3.01\times 10^6$ 、8.1 GFLOPs 降至  $2.3\times 10^6$ 、6.2 GFLOPs,但是 mAP@0.5 也随之降低,原因在于 GhostNet 模块采用简单的线性运算代替部分较为复杂的卷积运算,显著减少模型在卷积运算时的参数量和计算量。在骨干网络中添加 CAA 注意力机制,尽管该注意力机制使模型参数量和计算量分别从  $2.3\times 10^6$ 、6.2 GFLOPs 增加到  $2.43\times 10^6$ 、6.3 GFLOPs,但 mAP@0.5 提高了 2.1%。为了进一步提升模型检测精度,在改进骨干网络模型的基础上,添加 AFPN 网络结构到颈部网络中,mAP@0.5 从 80.4% 提升至 80.9%,虽然模型参数量略微增加,从  $2.43\times 10^6$  增加至  $2.58\times 10^6$ ,但计算量从 6.3 GFLOPs 下降至 5.4 GFLOPs。最后添加 WIoU v3 损失函数,改进模型相较基线模型,其参数量和计算量分别下降了 14.3%、33.3%,mAP@0.5 提升了 0.7%,改进模型在平衡检测精度的同时实现模型的显著轻量化,改进后模型的消融实验结果对比如图 11 所示。

为了展示改进后模型对不同类别垃圾目标的检测性能,将原模型与改进后模型在自建垃圾数据集上进行对

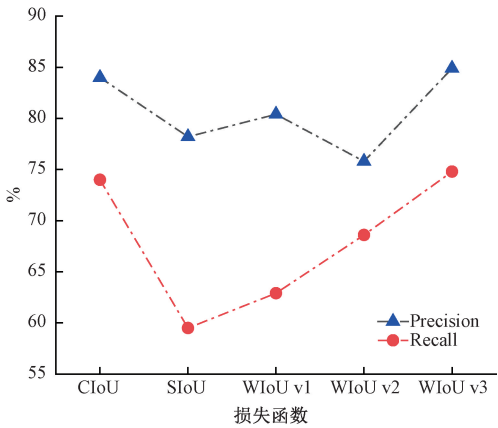


图 10 损失函数实验结果对比

Fig. 10 Comparison of loss function experiment results

模型为基线模型进行消融实验,并以精确度、召回率、参数量、计算量和 mAP@0.5 作为评价指标,通过逐一消融的方式进行实验,实验结果如表 5 所示。

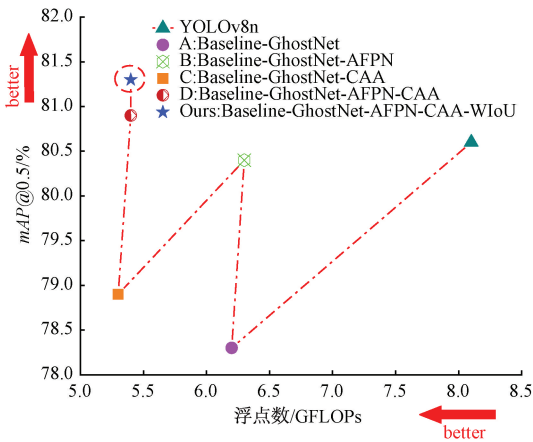


图 11 消融实验结果对比

Fig. 11 Comparison of ablation experiment results

比实验,得到精确度、召回率和 mAP 等相关指标,实验结果如表 6 所示。从表 6 可以看出改进模型对 7 类垃圾的平均检测精度值均有所提升,其中对烟蒂和瓶盖等小目标的提升效果最佳,均提高 3%,实验结果表明改进后模型能有效提高对复杂场景下各类垃圾的检测精度。

表 6 原模型与改进模型对各类垃圾检测对比

| Table 6 Comparison between the original model and the improved model for the detection of various types of garbage (%) |          |          |           |              |          |           |
|------------------------------------------------------------------------------------------------------------------------|----------|----------|-----------|--------------|----------|-----------|
| Category                                                                                                               | YOLOv8n  |          |           | GCAW-YOLOv8n |          |           |
|                                                                                                                        | <i>P</i> | <i>R</i> | mAP@ 0. 5 | <i>P</i>     | <i>R</i> | mAP@ 0. 5 |
| Lunch boxes                                                                                                            | 92       | 93. 3    | 97. 9     | 98. 6        | 99. 1    | 98. 7 ↑   |
| Can                                                                                                                    | 85. 2    | 81. 2    | 84. 3     | 87. 8        | 78. 1    | 85. 4 ↑   |
| Peel                                                                                                                   | 89       | 83. 8    | 90. 6     | 87. 8        | 84. 6    | 91. 1 ↑   |
| Garbage bags                                                                                                           | 93. 9    | 84. 7    | 90. 1     | 93. 6        | 86. 4    | 90. 5 ↑   |
| Butt                                                                                                                   | 87       | 59. 3    | 66. 1     | 89. 4        | 57. 1    | 69. 1 ↑   |
| Plastic bottles                                                                                                        | 89. 1    | 64. 4    | 78. 8     | 92. 1        | 68. 4    | 79. 8 ↑   |
| Caps                                                                                                                   | 88. 3    | 50. 5    | 57. 7     | 78           | 53. 3    | 60. 7 ↑   |
| Carton                                                                                                                 | 89. 9    | 66       | 79. 1     | 91. 3        | 66. 7    | 75. 5     |

4. 6 不同模型对比实验

为综合验证改进模型 GCAW-YOLOv8n 的优越性,选取 YOLOv5s、YOLOv5n、YOLOv7-tiny、YOLOv10n、YOLOv8s、YOLOv8n 与改进模型进行对比实验,实验结果如表 7 所示。

实验结果显示, YOLOv5n 模型参数量、模型大小及计算量最小,但检测精度太低,其 mAP@ 0. 5 相较于改进模型低 10. 9%。YOLOv8s 模型的检测精度最高,但其参数量、模型大小和计算量分别是改进模型的 4. 3、3. 8 和 5. 3 倍。YOLOv7-tiny 模型虽然检测速度最快,但模型轻

量化和检测精度两方面均比不上改进模型,模型参数量、模型大小和计算量是改进模型的 2. 3、2 和 2. 4 倍, mAP@ 0. 5 比改进模型低 0. 7%。改进模型虽然在检测速度上略低于原 YOLOv8n 模型,但能够满足实时检测的需求,改进模型的参数量、模型大小和计算量指标均优于其余检测模型,检测精度达到了 81. 3%,相比原 YOLOv8n 模型提高了 0. 7%。综合各项指标来看,改进模型与其他主流垃圾检测模型相比,在复杂背景环境下的垃圾检测中表现更为优秀。

表 7 不同模型对比实验结果

| Table 7 Comparison of experimental results of different models |             |             |                                    |         |            |             |        |
|----------------------------------------------------------------|-------------|-------------|------------------------------------|---------|------------|-------------|--------|
| 模型                                                             | <i>P</i> /% | <i>R</i> /% | <i>Params</i> /(×10 <sup>6</sup> ) | Size/MB | 浮点数/GFLOPs | mAP@ 0. 5/% | 帧率/FPS |
| YOLOv5s                                                        | 69. 1       | 59. 9       | 7. 04                              | 14. 4   | 15. 8      | 66. 2       | 106    |
| YOLOv5n                                                        | 78. 9       | 59. 2       | 1. 77                              | 3. 9    | 4. 3       | 70. 4       | 120    |
| YOLOv7-tiny                                                    | 81. 8       | 73. 6       | 6. 03                              | 12. 3   | 13. 2      | 80. 6       | 140    |
| YOLOv10n                                                       | 87. 6       | 71. 8       | 2. 7                               | 5. 6    | 8. 2       | 81          | 123    |
| YOLOv8s                                                        | 86. 6       | 75. 7       | 11. 1                              | 22. 5   | 28. 7      | 82. 2       | 116. 5 |
| YOLOv8n                                                        | 89. 3       | 72. 9       | 3. 01                              | 6. 3    | 8. 1       | 80. 6       | 118    |
| 本文                                                             | 84. 9       | 74. 8       | 2. 58                              | 6. 0    | 5. 4       | 81. 3       | 114    |

为了直观验证改进模型的目标检测有效性,将原 YOLOv8n 模型与改进模型进行可视化对比,分别选取昏暗光线、小目标、多目标、复杂场景和遮挡等不同情况下的样本图片进行对比,图 12(a)表示人工标注图片,图 12(b)表示原 YOLOv8n 模型的检测结果,图 12(c)表示改进模型的检测结果,对不同垃圾类别分别用不同颜色的锚定框来表示,检测结果对比如图 12 所示。

由图 12 可知,在复杂背景环境下目标检测容易出现漏检、误检的问题,尤其是针对瓶盖和烟蒂这种小目标检测。相比于原模型,改进模型 GCAW-YOLOv8n 具有更高效的边界框回归性能,垃圾目标检测效果更好,并且能够识别出更多有效的检测框,虽然在个别场景下出现漏检现象,但在提高置信度的基础上有效减少了误检现象,避免了冗余目标检测的产生,从而降低了后续处理的复杂性,减少了资源浪费。

5 结 论

针对面向复杂背景环境的垃圾检测模型存在参数量大、计算成本高等问题,限制了模型在资源受限设备上的应用,提出一种轻量化的 GCAW-YOLOv8n 模型。基于 YOLOv8n 模型进行轻量化改进,在骨干网络引入 GhostConv 和 C3Ghost 结构,并添加 CAA 注意力机制,降低模型参数量的同时提升特征提取能力;在颈部网络构建渐近特征金字塔 AFPN,增强多尺度目标检测能力;采用 WIoU v3 损失函数替换 CIoU 损失函数。改进模型在复杂背景环境下的垃圾图像数据集上 mAP@ 0. 5 达到了 81. 3%,与原 YOLOv8n 模型相比,模型参数量和计算量分别下降了 14. 3%、33. 3%,更好地平衡了检测精度与模



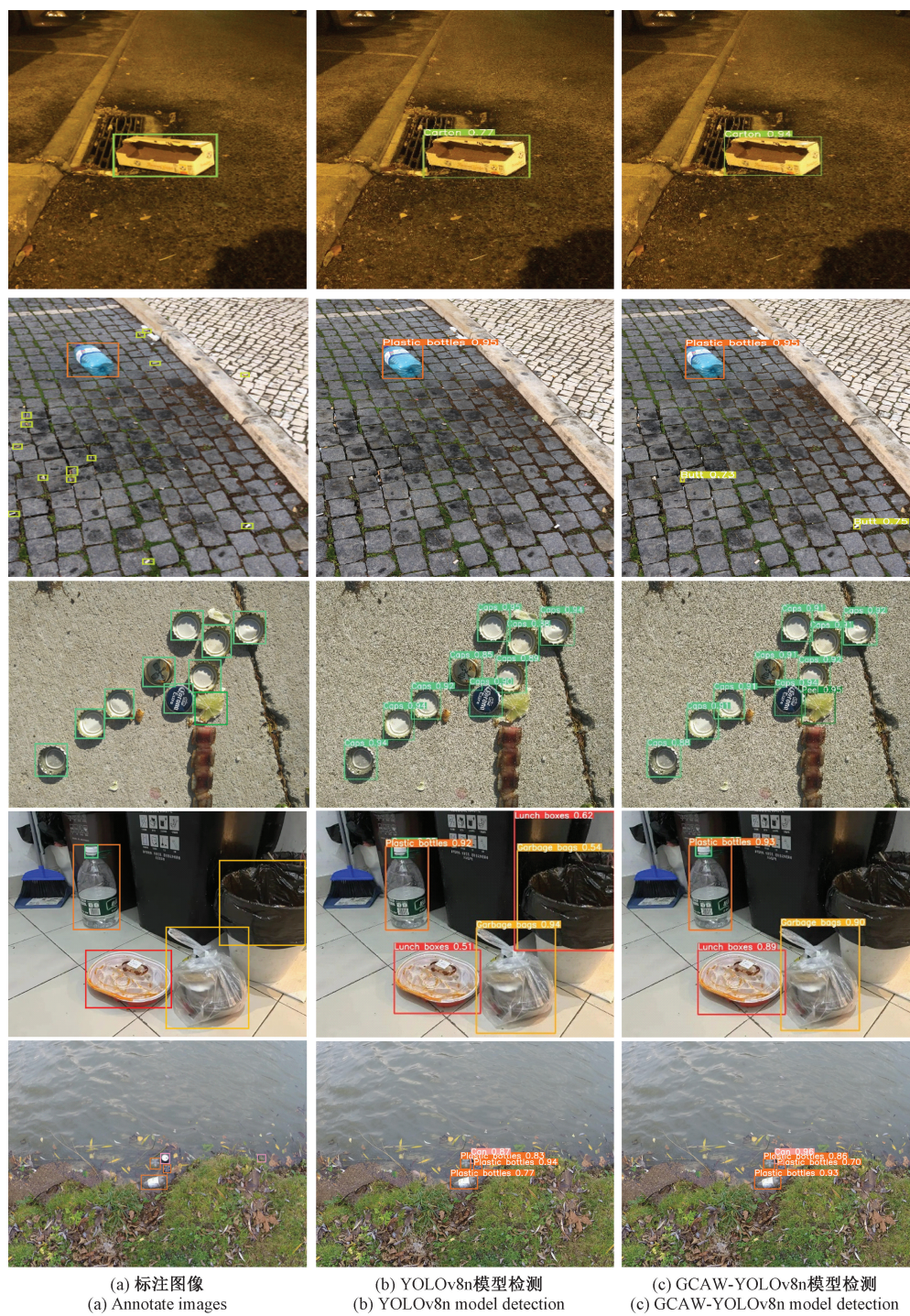


图 12 检测结果可视化对比  
Fig. 12 Visual comparison of test results

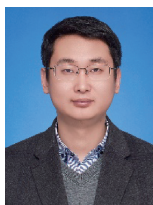
型轻量化,对模型部署与应用至边缘端设备具有重要的工程意义。然而,该模型目前仅在 8 种常见垃圾识别任务中取得了较好的检测效果。日常生活中的垃圾种类繁多,下一步研究将致力于丰富垃圾样本信息,进一步提升识别种类的多样性和小目标检测的精确性。

参考文献  
[ 1 ] KHAN S, ANJUM R, RAZA S T, et al. Technologies for municipal solid waste management: Current status, challenges, and future perspectives[J]. Chemosphere, 2022, 288: 132403.

- [2] ZHANG H, CHANG H, MA B, et al. Dynamic R-CNN: Towards high quality object detection via dynamic training[C]. Computer Vision-ECCV 2020: 16th Euro-pean Conference, 2020: 260-275.
- [3] 张睿萍, 宁芊, 雷印杰, 等. 基于改进 Mask R-CNN 的生活垃圾检测[J]. 计算机工程与科学, 2022, 44(11): 2003-2009.
- ZHANG R P, NING Q, LEI Y J, et al. Household waste detection based on improved Mask R-CNN[J]. Computer Engineering & Science, 2022, 44(11): 2003-2009.
- [4] 蒋占军, 吴佰靖, 马龙, 等. 多尺度特征和极化自注意力的 Faster-RCNN 水漂垃圾识别[J]. 计算机应用, 2024, 44(3): 938-944.
- JIANG ZH J, WU B J, MA L, et al. Faster-RCNN drifting garbage recognition based on multi-scale features and polarized self-attention [J]. Journal of Computer Applications, 2024, 44(3): 938-944.
- [5] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]. 2017 IEEE International Conference on Computer Vision (ICCV), 2017: 2999-3007.
- [6] 陈健松, 蔡艺军. 面向垃圾分类场景的轻量化目标检测方案[J]. 浙江大学学报(工学版), 2024, 58(1): 71-77.
- CHEN J S, CAI Y J. A lightweight object detection solution for garbage classification scenarios [J]. Journal of Zhejiang University (Engineering Science), 2024, 58(1): 71-77.
- [7] ZAILAN N A, MOHD KHAIRUDDIN A S, HASIKIN K, et al. An automatic garbage detection using optimized YOLO model[J]. Signal, Image and Video Processing, 2024, 18(1): 315-323.
- [8] 董兆苒, 董明利, 何彦霖, 等. 血管介入手术导丝末端检测方法研究[J]. 仪器仪表学报, 2023, 44(2): 221-229.
- DONG ZH R, DONG M L, HE Y L, et al. Study on the detection method of putting guide wire endinvascular interventional surgery [J]. Chinese Journal of Scientific Instrument, 2023, 44(2): 221-229.
- [9] 曾勇杰, 范必双, 杨涯文, 等. 改进 YOLOv8 算法在风机叶片缺陷检测上的应用[J]. 电子测量与仪器学报, 2024, 38(8): 26-35.
- ZENG Y J, FAN B SH, YANG Y W, et al. YOLOv8 algorithm is improved in the defect detection of wind turbine blades applications [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38 ( 8 ): 26-35.
- [10] MA B L, HUA ZH X, WEN Y CH, et al. Using an improved lightweight YOLOv8 model for real-time detection of multi-stage apple fruit in complex orchard environments[J]. Artificial Intelligence in Agriculture, 2024, 11: 70-82.
- [11] 胡久松, 刘张驰, 余谦, 等. 融入 GhostNet 和 CBAM 的 YOLOv8 烟雾识别算法[J]. 电子测量与仪器学报, 2024, 38(8): 201-207.
- HU J S, LIU ZH CH, YU Q, et al. YOLOv8 smoke detection algorithm integrated with GhostNet and CBAM [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(8): 201-207.
- [12] ZHAO X F, ZHANG W W, ZHANG H, et al. ITD-YOLOv8: An infrared target detection model based on yolov8 for unmanned aerial vehicles [J]. Drones, 2024, 8(4): 161.
- [13] WU T, DONG Y. YOLO-SE: Improved YOLOv8 for remote sensing object detection and recognition [J]. Applied Sciences, 2023, 13(24): 12977.
- [14] 曲晨阳, 吕进, 卫策. 基于改进 YOLOv8s 的无人机目标检测算法[J]. 国外电子测量技术, 2024, 43(7): 14-23.
- QU CH Y, LYU J, WEI C. UAV target detection algorithm based on improved YOLOv8s [J]. Foreign Electronic Measurement Technology, 2024, 43 ( 7 ): 14-23.
- [15] 肖立中, 胡凡. 基于 LGD-YOLO 高精度轻量化目标检测网络的垃圾检测研究 [J]. 环境工程, 2024, 42(6): 169-177.
- XIAO L ZH, HU F. Research on garbage detection based on LGD-YOLO high-precision lightweight object detection network [J]. Environmental Engineering, 2024, 42(6): 169-177.
- [16] PROENÇA P F, SIMÕES P. Taco: Trash annotations in context for litter detection [J]. ArXiv preprint arXiv: 2003. 06975, 2020.
- [17] HAN K, WANG Y H, TIAN Q, et al. Ghostnet: More features from cheap operations [C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 1577-1586.
- [18] CAI X, LAI Q, WANG Y, et al. Poly kernel inception network for remote sensing detection [C]. 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024: 27706-27716.
- [19] 张辉, 李博涛, 尚景浩, 等. 基于多尺度卷积注意力机制的输电线路防振锤缺陷检测[J]. 电工技术学报, 2024, 39(11): 3522-3537.
- ZHANG Y, LI B T, SHANG J H, et al. Transmission line anti-vibration hammer defect detection based on

- multi-scale convolutional attention mechanism [J]. Transactions of China Electrotechnical Society, 2024, 39(11): 3522-3537.
- [20] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146-157.
- [21] TONG Z, CHEN Y, XU Z, et al. Wise-IOU: Bounding box regression loss with dynamic focusing mechanism [J]. ArXiv preprint arXiv: 2301.10051, 2023.
- [22] MA N, ZHANG, X, ZHENG H T, et al. Shufflenet v2: Practical guidelines for efficient CNN architecture design[C]. European Conference on Computer Vision (ECCV), 2018: 116-131.
- [23] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3 [C]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019: 1314-1324.
- [24] LIU Y CH, SHAO Z R, HOFFMANN N. Global attention mechanism: Retain information to enhance channel-spatial interactions[J]. ArXiv Preprint arXiv: 2112.05561, 2021.
- [25] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3-19.
- [26] OUYANG D L, HE S, ZHANG G ZH, et al. Efficient multi-scale attention module with cross-spatial learning[C]. ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2023: 1-5.
- [27] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11534-11542.
- [28] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.

### 作者简介



孙世政(通信作者), 2012 年于重庆理工大学获得硕士学位, 2015 年于合肥工业大学获得博士学位, 现为重庆交通大学教授、硕士生导师, 主要研究方向为精密仪器与机械、智能传感技术。

E-mail: ssz091011@163.com

**Sun Shizheng** (Corresponding author) received his M.Sc. degree from Chongqing University of Technology in 2012, Ph. D. degree from Hefei University of Technology in 2015. Now, he is a professor and master's supervisor at Chongqing Jiaotong University. His main research interests include precision instruments & machinery and intelligent sensing technology.