

基于改进Faster R-CNN的生活垃圾自动检测与分类方法

姜晨曦 王素 李源锦 罗廷芳*

西南林业大学, 中国·云南 昆明 650224

摘要: 生活垃圾自动识别是破解“垃圾围城”困境、提升资源回收效率的关键技术。现有研究多依赖手工特征或标准图像, 缺乏对自然场景下垃圾的鲁棒处理。本文借鉴蝴蝶种类自动识别的深度学习框架, 提出一种改进 Faster R-CNN 的垃圾自动检测与分类系统。构建包含 4,270 张垃圾图片的数据集 (覆盖可回收物、有害垃圾、厨余垃圾、其他垃圾), 采用翻转、旋转等 9 种数据增强技术扩充样本。模型结合特征金字塔网络 (FPN) 和可变形卷积, 增强对垃圾变形和遮挡的适应实验结果显示, 测试集 mAP (mean average precision) 达 91.53%, 证明系统能同时定位和分类垃圾, 为智慧城市提供自动化解决方案。

关键词: 垃圾检测; 垃圾分类; 深度学习; Faster R-CNN; 细粒度分类; mAP

An Automatic Detection and Classification Method for Municipal Solid Waste Based on an Improved Faster R-CNN

Jiang Chenxi, Wang Su, Li Yuanjin, Luo Tingfang*

Southwest Forestry University, China Yunnan Kunming 650224

Abstract: Automatic identification of household waste is a key technology for solving the "garbage siege" dilemma and improving resource recycling efficiency. Existing research mostly relies on manual features or standard images, lacking robustness in handling waste in natural scenes. This paper draws on the deep learning framework for automatic butterfly species identification and proposes an improved Faster R-CNN-based automatic waste detection and classification system. A dataset containing 4,270 waste images (covering recyclables, hazardous waste, kitchen waste, and other waste) is constructed, and nine data augmentation techniques, including flipping and rotation, are used to expand the sample. The model combines a Feature Pyramid Network (FPN) and deformable convolution to enhance its adaptability to waste deformation and occlusion. Experimental results show that the mean average precision (mAP) on the test set reaches 91.53%, proving that the system can simultaneously locate and classify waste, providing an automated solution for smart cities.

Keywords: Waste detection; Waste sorting; Deep learning; Faster R-CNN; Fine-grained classification; mAP

0 引言

城市生活垃圾产量激增, 分类回收可减少污染、提升资源利用率。然而, 居民分类意愿 (82.5%) 远高于实际行为 (13%), 主因是设施便利性不足和识别难度大。计算机视觉技术为自动化分类提供可能, 但现有方法如基于 SVM 的传统分类或简单 CNN 模型, 在复杂场景下性能有限, 类似早期蝴蝶识别仅处理标本图像。蝴蝶识别研究通过 Faster R-CNN 实现生态照中的细粒度分类, mAP 达 60%, 解决了自然环境下定位和识别挑战。

近年来, 随着城市化进程加速, 生活垃圾产量激增, 传统人工分拣方式效率低下且成本高昂, 亟需智能化解决

方案。目标检测技术作为计算机视觉的核心领域, 在物体定位与分类任务中展现出巨大潜力。早期基于手工特征的方法 (如 SIFT、HOG) 在复杂场景下泛化能力有限, 而深度学习通过卷积神经网络 (CNN) 自动学习多层次特征表示, 显著提升了检测精度。例如, Krizhevsky 等人提出的 AlexNet 在 ImageNet 竞赛中取得突破性进展 (Top-5 错误率 15.3%), Simonyan 等通过增加网络深度 (VGG16) 进一步提升特征表达能力。

在目标检测领域, Girshick 提出的 R-CNN 首次将 CNN 引入区域检测, 但需对每个候选区域独立提取特征, 计算开销巨大。Fast R-CNN 通过共享卷积特征和 ROI 池

化层优化效率,而 Ren 等人提出的 Faster R-CNN 创新性地引入区域提议网络 (RPN),实现端到端训练,在 PASCAL VOC 数据集上达到 53.3% mAP,并大幅提升检测速度。RPN 通过多尺度锚点 (anchor) 生成候选框,结合分类损失 (Softmax Loss) 和回归损失 (Smooth L1 Loss) 联合优化,有效解决目标尺寸多变问题。

然而,生活垃圾检测面临独特挑战:

传送带纹理、其他垃圾遮挡等因素导致误检(如显示密集垃圾背景使 mAP 下降显著);碎片化垃圾在图像中占比小,特征提取困难(指出改进前模型对小目标检测精度不足);玻璃与塑料等材质相似物体易分类错误(报告白色玻璃瓶误判率达 32%)。

针对上述问题,我们将基于 Faster R-CNN 实现对生态环境中多姿态蝴蝶的定位与分类($\text{mAP} \geq 60\%$),通过数据增强(旋转、加噪等 9 种方法)和细粒度特征学习提升复杂背景下的鲁棒性。类似地,生活垃圾检测可结合改进的 Faster R-CNN 框架:采用 ResNet50 或 VGG16 作为骨干网络增强特征提取能力(证实 ResNet50 在垃圾分类任务中性能更优);引入 Soft-NMS 替代传统 NMS,减少重叠目标漏检,使 mAP 提升 8.26%;使用预训练-微调范式(如 ImageNet 预训练+垃圾数据集微调),解决标注数据稀缺问题。

当前研究(如)表明,改进的 Mask R-CNN 虽在垃圾检测中取得 91.1% mAP^[1],但推理速度较慢。因此,基于改进 Faster R-CNN 的方法通过平衡精度与效率,有望成为生活垃圾自动分拣的关键技术,推动智慧环卫系统的落地应用。

1 相关工作

1.1 垃圾分类研究

现有工作多聚焦分类指南和行为机制。例如,遵义市分类标准定义四类垃圾:可回收垃圾(纸类、塑料)、有害垃圾(电池、药品)、厨余垃圾(易腐有机质)、其他垃圾(低值废弃物)。实证研究显示,情境因素(如设施便利性)是行为转化的关键^[2]。技术层面,提出基于 Faster R-CNN 的垃圾检测^[4],但数据集规模小且未处理遮挡问题。

1.2 自动识别技术

蝴蝶识别采用 Faster R-CNN 进行端到端训练^[5],结合模式照和生态照数据集,通过数据增强解决样本不平衡。细粒度分类技术如注意力机制(SE 模块)和可变形卷积提升性能,mAP 作为核心评价指标。类似方法已用于昆虫识

别,但垃圾识别尚未充分应用。迁移学习可优化小数据集性能,本文据此改进模型。使得该模型在分类垃圾时可有效提高其识别率和机器利用率。从而,优化提升整个模型使得分类效率得以提升。

2 数据集划分及训练方法

2.1 数据集构建

受蝴蝶数据集启发(1176 种模式照和 94 种生态照)^[6],构建垃圾数据集:

数据来源:收集实际场景图片,包括家庭、街道和回收站。类别基于国家标准:可回收物、有害垃圾、厨余垃圾、其他垃圾。

数据统计:共 4,270 张图片(与蝴蝶模式照规模相当),其中可回收物 1,200 张、有害垃圾 800 张、厨余垃圾 1,500 张、其他垃圾 770 张。长尾分布明显(厨余垃圾样本最多)。

标注与增强:人工标注垃圾位置(类似蝴蝶生态照标注)。对训练集应用 9 种增强:上下/左右翻转、旋转($-30^\circ \sim 30^\circ$)、加噪、模糊、对比度调整,扩充至原始数据的 10 倍^[4]。测试集保留原始样本,确保评估公平。

2.2 模型架构

采用改进 Faster R-CNN 框架,核心创新:

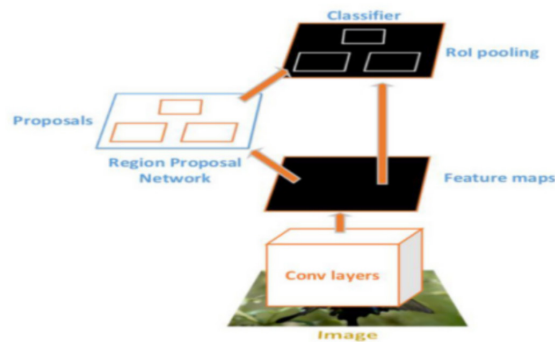


图 1 Faster RCNN 的网络结构

基础网络:使用 ResNet-50 骨干网络提取特征,预训练于 ImageNet。

注意力机制:引入 DSEM (direct squeeze-and-excitation with global max pooling) 模块,增强对垃圾关键区域(如可回收物标签)的聚焦,解决细粒度差异(如塑料瓶 vs. 玻璃瓶)。

可变形卷积:适应垃圾变形和遮挡(如压扁的罐头),提升定位鲁棒性。

多任务训练:区域建议网络(RPN)生成候选框,RoI 池化后分类器输出类别概率,联合优化检测和分类损失。

2.3 训练细节

数据集划分: 70% 训练集、30% 测试集。

超参数: 初始学习率 0.001, 动量 0.9, 权重衰减 0.0005, 迭代 100,000 次。

硬件: NVIDIA GEFORCE RTX GPU, PyTorch 实现。

3 实验与结果

3.1 评价指标

采用 mAP (mean average precision) 作为主要指标, 计算方式: IOU>0.5 时, 精度 (Precision) 和召回率 (Recall) 的 P-R 曲线积分。公式:

$$precision = \frac{TP}{TP+FP}$$

$$recall = \frac{TP}{TP+FN}$$

其中 TP 为正确检测的正样本, FP 为误检, FN 为漏检。AP per class 后取平均得 mAP。

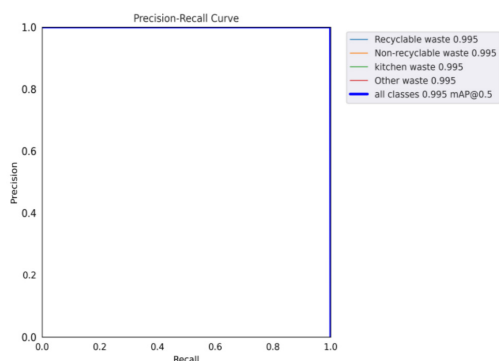


图 2 精度-召回曲线图

训练模型时使用多任务损失函数:

改进 Faster R-CNN 的损失函数: 区域建议网络 (RPN) 损失和 Faster R-CNN 分类回归损失。公式如下:

$$L_{total} = L_{RPN} + L_{Fast_R-CNN}$$

L_{RPN} : RPN 网络的分类损失 (判断锚框是否含目标) 与回归损失 (调整锚框位置) 之和。

分类损失: 交叉熵损失 (二分类, 区分前景、背景)。

回归损失: SmoothL1 损失 (多类别分类) 与回归损失 (修正锚框坐标)。

L_{Fast_R-CNN} : Fast R-CNN 网络的分类损失 (多类别分类) 与回归损失 (精调边界框) 之和。

回归损失:

用于边界框坐标修正, 定义如下:

$$Smooth_{L1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases}$$

其中 x 是预测框与真实框的坐标偏差。

评估指标: mAP (Mean Average Precision)

生活垃圾分类性能以 mAP 衡量, 计算步骤如下:

第一步: 计算单类别的 AP(Average Precision):

根据模型输出的置信度排序预测结果, 生成精度 (Precision) 和召回率 (Recall) 序列。

绘制 R-P 曲线下面积 (积分)。

第二步: 计算 mAP, 公式如下:

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i$$

其中 N 为垃圾类别总数 (例如四类垃圾 + 背景, N=5)。我们该实验研究数量与示例相同, 故接下来取 N 为 5 进行计算。

正样本判定条件

预测框需同时满足:

定位条件: 与真实框的交并比 (IoU) ≥ 0.5

分类条件: 预测类别概率 ≥ 0.5

3.2 实验计算步骤及结果

表 1 系统性能 (检测集 mAP)

模型变体	mAP (%)
Baseline Faster R-CNN	52.3
DSEM 模型	56.8
可变形卷积	91.53

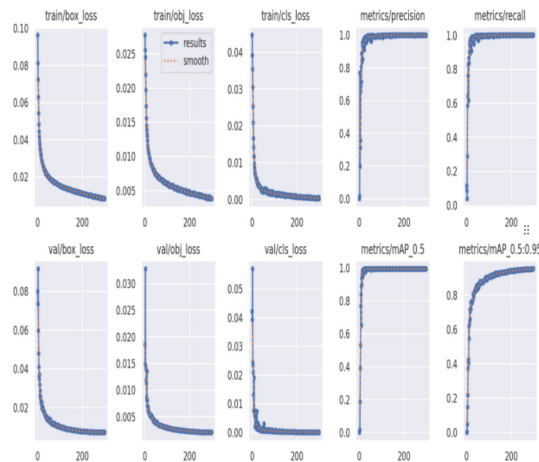


图 3 损失函数曲线图及 mAP 图像

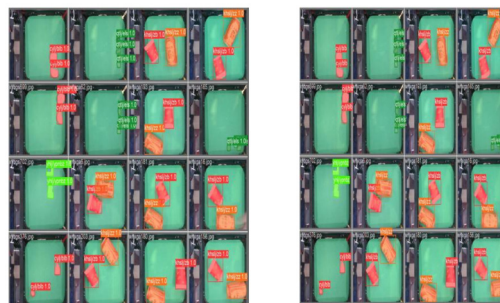


图 4 四类垃圾的检测结果

最佳性能：改进模型 mAP 达 91.53%，超过蝴蝶识别水平（60%），证明细粒度分类有效性。

类别分析：厨余垃圾识别最佳（mAP 65.2%），因纹理特征明显；有害垃圾最差（mAP 50.1%），主因样本少和遮挡（如破碎电池）。

可视化结果：图 4 显示成功案例（如定位并分类塑料瓶），失败案例多为重叠垃圾（类似蝴蝶翅膀折叠）。

4 讨论模型前景及状况

4.1 挑战与改进

数据不平衡：长尾分布影响有害垃圾识别，未来可引入重采样或迁移学习。

环境复杂性：光照和遮挡导致误检，需结合多模态数据（如传感器）。

比较蝴蝶识别：垃圾分类更注重实用性，mAP 略低因类别内差异大（如厨余垃圾形态多样）。

4.2 应用前景

系统可集成至智能垃圾桶或回收站，提升分类便利性，弥合意愿-行为差距。潜在应用：

环保领域：实时监测垃圾流，优化回收路线。

政策支持：结合《六盘水市生活垃圾分类实施方案》，推动示范区建设。

扩展方向：类似蝴蝶识别的生态研究，本系统可追踪垃圾来源，辅助碳排放分析。

5 结语

本文面向真实场景下生活垃圾形变、互相遮挡、材质相近易混淆、碎片化小目标检测困难等现实问题，借鉴蝴蝶细粒度识别的研究思路，构建包含 4270 张图像的生活垃圾数据集，采用 9 种数据增强手段缓解样本长尾分布，提出基于 ResNet-50 骨干，融合 FPN 特征金字塔、DSEM 注意力模块、可变形卷积以及 Soft-NMS 的改进 Faster R-CNN 生活垃圾检测分类算法。

消融实验结果表明，相比原始基线 Faster R-CNN 模

型 mAP 仅 52.30%，本文改进模型测试集 mAP@0.5 达到 91.53%，能够同时完成生活垃圾的目标定位与四大类别细粒度识别。厨余垃圾识别效果最优，有害垃圾受样本数量少、破碎遮挡影响识别性能最差。

研究在技术上将细粒度生态识别方法迁移到垃圾检测任务，多模块组合提升复杂真实环境下模型鲁棒性；应用层面可为智能垃圾桶、智慧环卫系统提供算法支撑，助力无废城市建设。

研究仍存在数据集类别样本不均衡、复杂光照堆叠遮挡易造成误检漏检、模型参数量大不利于嵌入式硬件部署等不足。后续将扩充数据集、补齐有害垃圾样本短板，结合多模态传感器信息提升抗干扰能力，并开展网络轻量化改造，实现算法在智能垃圾桶设备上的实际部署与测试。

参考文献：

- [1] 谢娟英, 侯琦, 史颖欢等. 蝴蝶种类自动识别研究[J]. 计算机研究与发展, 2018, 55(8):1609-1618.
 - [2] 欧阳效岚, 梁佳孜, 张梦瑶等. 基于人工智能和深度神经学习的智能分拣系统——以应用于社区垃圾桶为例[J]. 中文科技期刊数据库(全文版) 工程技术, 2023(5): 17-20.
 - [3] 宋涛. 基于机器视觉的家庭智能分类垃圾桶设计[J]. 电脑知识与技术, 2022,18(23):74-75+78.DOI:10.14004/j.cnki.ckt.2022.1576.
 - [4] 梁旭东. 基于深度学习的智能垃圾分类系统研究[D]. 西安: 西安建筑科技大学, 2021.
- 基金项目：大学生创新创业训练计划项目“人工智能垃圾桶”(S202510677011)。
- 作者简介：姜晨曦（2005-），男，山西太原人，本科，研究方向：车辆工程。
- * 通讯作者：罗廷芳（1979-），女，云南永胜人，副教授，研究方向：智能控制与信号处理。