

## 基于 YOLOv5s 网络的垃圾分类和检测

王莉, 何牧天, 徐硕, 袁天, 赵天翊, 刘建飞

(南京工业大学, 南京 211816)

**摘要:** **目的** 为了实现垃圾自动按类处理, 通过研究基于视觉的垃圾检测与分类模型, 实现对垃圾的自动识别和检测。**方法** 采用 YOLOv5s 网络作为垃圾检测与分类的模型, 在自制垃圾分类数据集上对网络进行训练, 利用训练好的 YOLOv5s 网络提取不同种类垃圾图像的特征和位置信息, 实现垃圾分类与检测。**结果** 在真实场景中进行了测试, 基于 YOLOv5s 的垃圾分类检测模型可以有效识别 6 种不同形态的垃圾, 检测 mAP 值为 99.38%, 测试精度为 95.34%, 目标检测速度达到 6.67FPS。**结论** 实验结果表明, 基于 YOLOv5s 网络的垃圾分类检测模型在不同光照、视角等条件下, 检测准确率高, 鲁棒性好、计算速度快。同时, 有助于促进垃圾处理公司实现智能分拣, 提高工作效率。

**关键词:** YOLOv5s 网络; 垃圾分类; 目标检测

**中图分类号:** TP242 **文献标识码:** A **文章编号:** 1001-3563(2021)08-0050-07

**DOI:** 10.19554/j.cnki.1001-3563.2021.08.007

## Garbage Classification and Detection Based on YOLOv5s Network

WANG Li, HE Mu-tian, XU Shuo, YUAN Tian, ZHAO Tian-yi, LIU Jian-fei

(Nanjing Tech University, Nanjing 211816, China)

**ABSTRACT:** To automatically dispose garbage according to its category, the automatic recognition of garbage can be realized by doing research on the vision based garbage detection and classification model. YOLOv5s network is taken as a garbage classification and detection model and trained on the self-made garbage data set. The trained YOLOv5s network extracts the features and location information from different kinds of garbage images, and then recognizes and detects different garbage in the image. The performance of the trained YOLOv5s is validated in a real situation. Garbage classification and detection model which was based on YOLOv5s can identify six different kinds of garbage effectively. The mAP of the trained YOLOv5s is 99.38%, the recognition accuracy is 95.34%, and the speed of target detection reaches 6.67FPS. The experiment results show that the garbage classification model based on YOLOv5s has high accuracy, good robustness and fast speed under different situations such as illumination, camera angle, etc. At the same time, this study can help enterprises to realize intelligent sorting and improve the efficiency.

**KEY WORDS:** YOLOv5s network; garbage classification; object detection

近年来, 全球的垃圾产量急剧增长, 然而大量垃圾并未得到有效处理, 从而导致环境与健康问题日益恶化。垃圾分类可以实现资源的回收利用, 是解决垃圾处理问题的有效方式。住房城乡建设部于 2017 年 12 月发布的《关于加快推进部分重点城市生活垃圾分类工作的通知》<sup>[1]</sup>中也明确了 2019 年全国正式启

动垃圾分类工作。传统的手工垃圾分拣的效率低、成本高, 因此, 如何将垃圾自动按类处理具有重要的研究意义。目前市面上出现了多款基于视觉的垃圾分拣机器人, 如芬兰的 ZenRobotics 垃圾回收机、日本 FANUC 分拣机器人、美国 Max-AI 机器人、AMP 机器人、Google X 日常机器人和 BHS 智能分拣机器人、

收稿日期: 2021-01-16

基金项目: 国家自然科学基金项目 (61973334); 江苏省大学生创新创业训练计划项目 (2024DC0241)

作者简介: 王莉 (1979—), 女, 江苏人, 博士, 南京工业大学副教授, 主要研究方向为模式识别与人工智能。

通信作者: 徐硕 (1996—), 男, 江苏人, 南京工业大学硕士生, 主攻计算机视觉、深度学习。

中国的 Picking-Ai™ 人工智能分拣机器人等，这些机器人都是通过视觉系统判断垃圾的种类，再利用机械装置抓取垃圾并放到指定区域，其中判断垃圾的种类以及在图像中的位置是首要解决的问题。

深度学习可以自动从图像中提取特征，避免了手工提取特征的主观性，在目标检测与分类问题中得到了广泛的应用。近年来，许多科研人员将深度学习用于垃圾分类的研究中<sup>[2-10]</sup>。Yang 等<sup>[2]</sup>创建了一个垃圾分类数据集 Trashnet，并使用支持向量机和卷积神经网络分别建立了垃圾分类模型。Mao 等<sup>[3]</sup>采用利用遗传算法对 DenseNet121 网络全连接层参数进行了优化，训练得到垃圾分类模型，分类精度达到 99.60%，Toğaçar 等<sup>[4]</sup>采用自编码网络重新构建了垃圾分类数据集，再使用 CNN 从数据集中自动提取特征。但是这些方法的网络模型体积较大，不适宜部署到嵌入式终端。Rabano 等<sup>[5]</sup>采用 ImageNet 数据集对轻量化网络 MobileNet 网络进行预训练，在垃圾分类数据集上再次训练调整网络参数，将训练好的模型成功部署到三星 Galaxy S6 手机。但该方法只能对垃圾的种类进行判断，不能检测出垃圾的具体位置。

由 Redmon 等<sup>[11]</sup>于 2016 年提出 YOLO 算法，将目标检测的分类和定位用一个神经网络实现，在目标检测领域得到了广泛的应用，目前已经经历了从 v1 到 v5 的发展<sup>[11-14]</sup>。许伟等<sup>[9]</sup>将 Dense 模块嵌入到 YOLOv2 中，从图像中提取浅层和深层特征，在自制的 25 类垃圾数据集上识别准确率达到 84.98%。宁凯等<sup>[10]</sup>对 YOLOv3 网络进行轻量化改进，在自制垃圾数据集上分类准确率达到 90.50%。相比于 YOLOv2，基于 YOLOv3 的垃圾分类检测模型精度有了提升。然而 YOLOv3 是基于 Darknet 框架，在环境配置上比较复杂。然而，移动终端的垃圾检测需求日益增加，这些场景使用的硬件大多为计算能力不强的边缘设

备，YOLOv3 这样较大的检测网络难以部署。Ultralytics 公司 2020 年推出的 YOLOv5 具有体积小、速度快、精度高等优点，且在生态成熟的 PyTorch 中实现，部署实现简单。YOLOv5 借鉴 CutMix<sup>[15]</sup>方法，采用 Mosaic 数据增强方式有效解决了模型训练中最头疼的“小对象问题”。YOLOv5 包括了 YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x 4 个模型，其中 YOLOv5s 体积最小，权重仅仅 27 MB。本文选择 YOLOv5s 作为垃圾检测与分类的模型，在自制垃圾分类数据集上对网络进行训练，利用训练好的 YOLOv5s 网络实现垃圾的自动分类与检测。

### 1 系统总体设计方案

本次实验设计了基于 YOLOv5s 的自动垃圾分拣系统，系统整体框架见图 1。首先在服务器上训练基于 YOLOv5 的垃圾分类模型。然后将训练好的模型部署到 Nvidia Jetson Xavier 开发板上，利用摄像头实时采集现场图片，并将图片通过无线方式传输到 Jetson Xavier。Jetson Xavier 利用训练好的垃圾分类模型检测图片中的是否存在垃圾及相应的类别，Jetson Xavier 将垃圾分类结果通过串口发送给嵌入式终端 STEM32，STEM32 控制分拣装置根据识别的垃圾类别，抓取垃圾投放置到相应区域。

### 2 YOLOv5s 网络模型

YOLOv5s 还是沿用 v3、v4 的整体布局，将整个网络结构分成 Input、Backbone、Neck、Output 4 个部分，网络结构见图 2。输入端：自适应缩放图片，采用 Mosaic 数据增强方式，自动计算数据集的最佳锚框值。Backbone：包含 Focus 结构和 CSPNetCross Stage Partial Networks，跨阶段局部融合网络。Focus

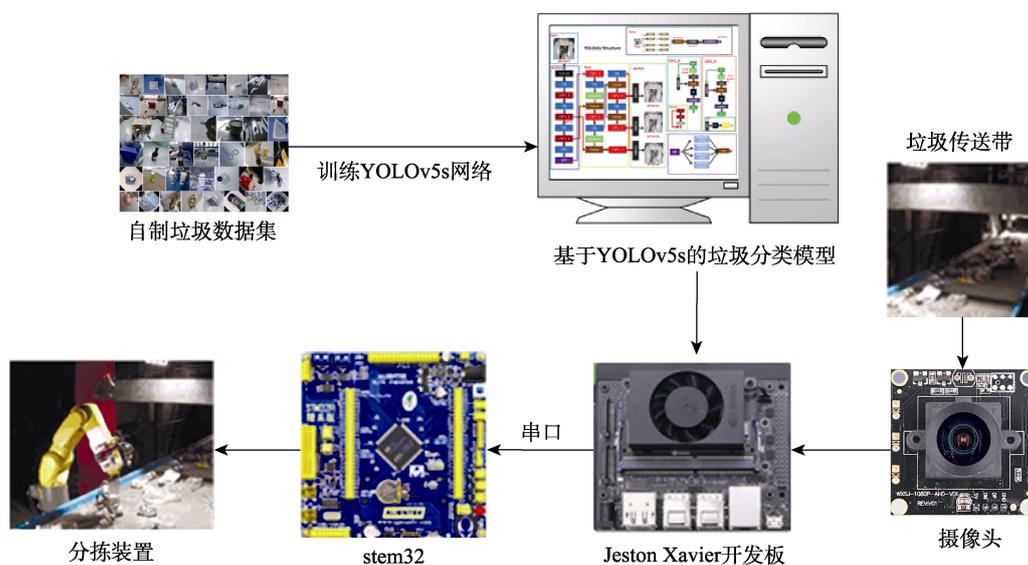


图 1 系统整体框架  
Fig.1 Overall system framework diagram

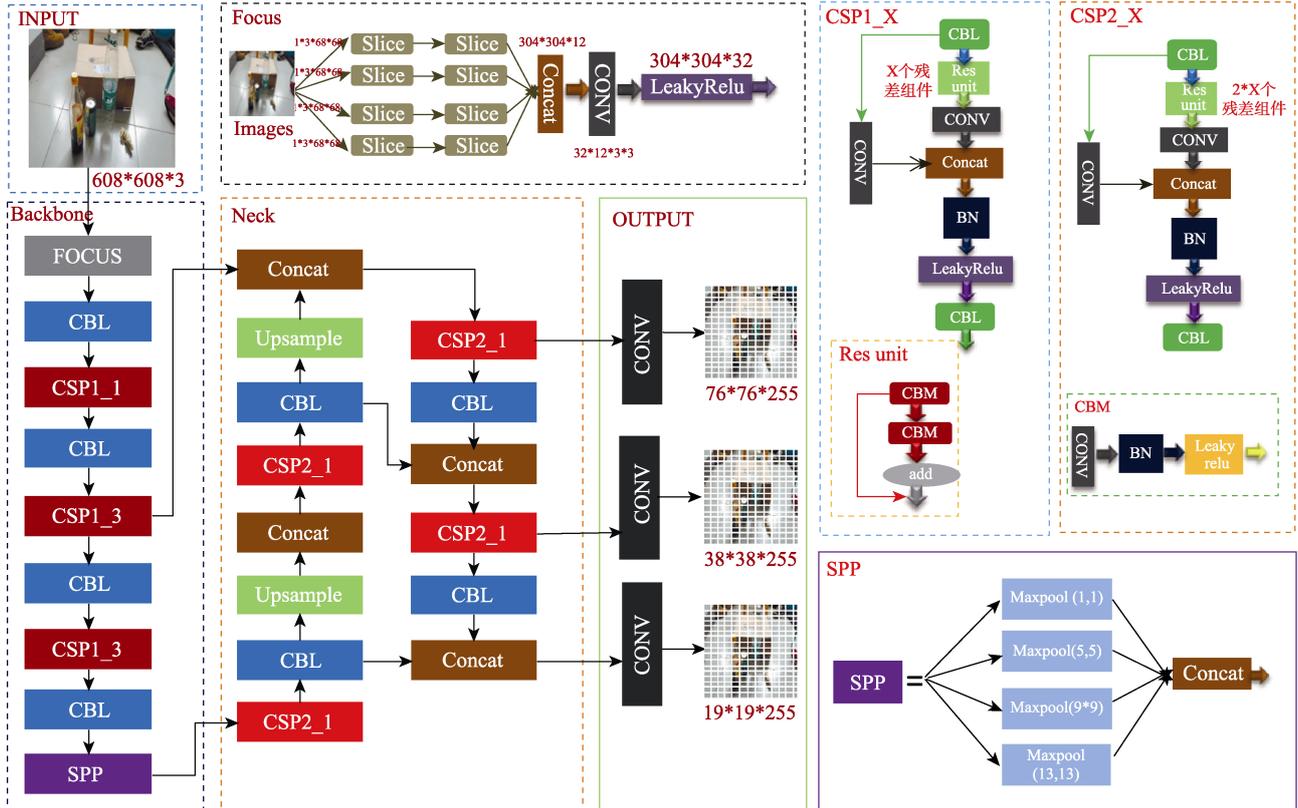


图2 YOLOv5s 模型结构

Fig.2 The structure of YOLOv5s network

结构包含4次切片操作和1次32个卷积核的卷积操作将原始 $608 \times 608 \times 3$ 的图像变成 $304 \times 304 \times 32$ 的特征图。CSPNet仿照Densenet<sup>[16]</sup>密集跨层跳层连接的思想,进行局部跨层融合,利用不同层的特征信息来获得更为丰富的特征图。Neck:包含了PANet和SPP模块。PANet(PathAggregation,路径聚合网络)<sup>[17]</sup>自顶向下将高层特征信息与不同层CSP模块的输出特征进行聚合,再通过自底向上路径聚合结构聚合浅层特征,从而充分融合了不同层的图像特征。空间金字塔池化(space pyramid pooling, SPP)模块先采用4种不同大小的核进行最大池化操作,再进行张量拼接。输出层:使GIOW\_Loss<sup>[18]</sup>代替YOLOv3的IOU\_Loss作为损失函数,增加了相交尺度的衡量,缓解了IOU\_Loss无法优化两个框不相交的情况。

### 3 实验设置与结果分析

#### 3.1 数据集介绍

本次实验自制了一个垃圾数据集,包括纸盒、塑料瓶、易拉罐、玻璃瓶、药瓶、电池、塑料袋、废纸、衣服、砖头、水果和蔬菜12类垃圾,考虑了不同光照、角度、距离和遮挡等情况,各类垃圾示例见图3。采用数据增广的方式将样本个数增至原来的4倍,用labelimg工具对图片中垃圾物体进行了标注,按照8:2的比例将数据集划分为训练集和测试集。数据

集共12000张图像,每张图像的分辨率大小为 $1596 \times 640$ 。

#### 3.2 YOLOv5s 网络训练

本次实验的运行环境:CPU为Intel i7 7800X, GPU为GeForce GTX 1080ti,内存为16G,操作系统为ubuntu16.04,安装CUDA9.0库文件,开发语言为Python,Pytorch框架。

Coco和Voc数据集上的训练结果初始化Yolov5s的网络参数,参数训练采用SGD<sup>[5]</sup>优化算法,参数设置如下:Batch大小为32;最大迭代次数为400;动量因子为0.9;权重衰减系数为0.0005。采用余弦退火策略动态调整学习率,初始学习率为0.01。采用GIOW\_Loss作为损失函数。为了验证本实验方法的有效性,将Yolov5s和Yolov3在自制垃圾数据集上进行实验对比,两种网络在训练集上的损失函数曲线见图4。

从图4中可以看出,YOLOv5s初始的损失函数值较小,YOLOv5在训练100轮左右达到稳定,而YOLOv3在训练了300轮左右才逐渐稳定,可见YOLOv5的初始参数值更加合理,收敛速度较YOLOv3更快。

#### 3.3 测试结果与分析

在不同光照、不同背景、不同视角下、不同尺度和遮挡情况下对训练好的模型进行了测试,实验结果

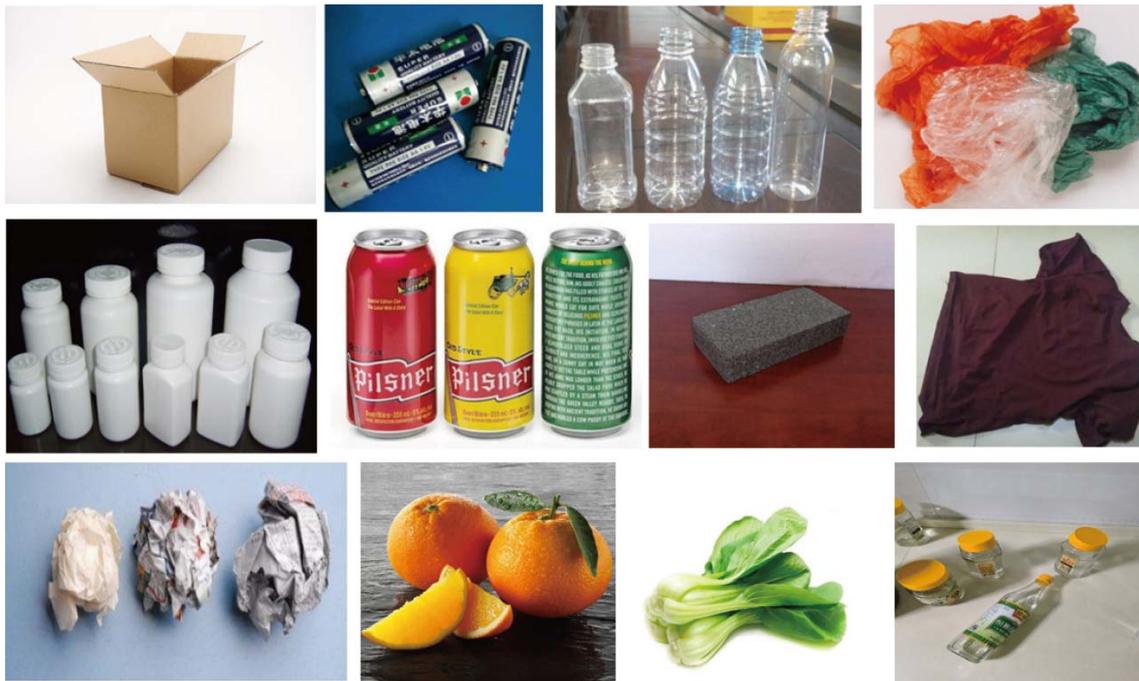


图 3 自制垃圾数据集

Fig.3 Garbage data set made by ourselves

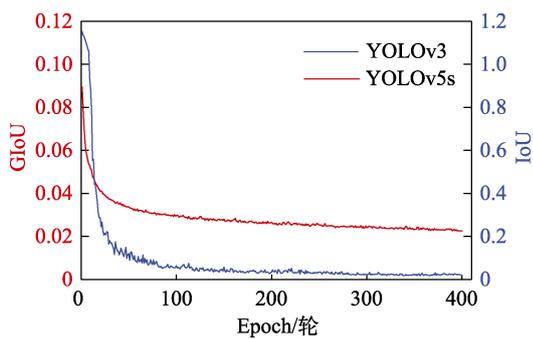


图 4 YOLOv3 与 YOLOv5s 损失函数对比

Fig.4 Comparison of Yolov3 and Yolov5s loss function

见图 5—8 (其中 a 为 YOLOv3; b 为 YOLOv5s), 图上的数值为置信度。从检测结果可以看出, 在各种不同条件下, YOLOv3、YOLOv5s 都能检测到各类目标垃圾, YOLOv5s 的置信度整体略高于 YOLOv3。光照条件越好, 整体的检测置信度越高, 在遮挡情况下, 检测目标的置信度均有所下降。

多种垃圾的检测效果见图 9。从图 9 中可以看出, 两种模型均能准确检测到纸盒、药品和废纸 3 种垃圾, YOLOv3 (见图 9a) 虽然能够检测到塑料瓶, 但其置信度较低, 仅为 61%, 而 YOLOv5s (见图 9b) 的置信度为 86%, 优于 YOLOv3。

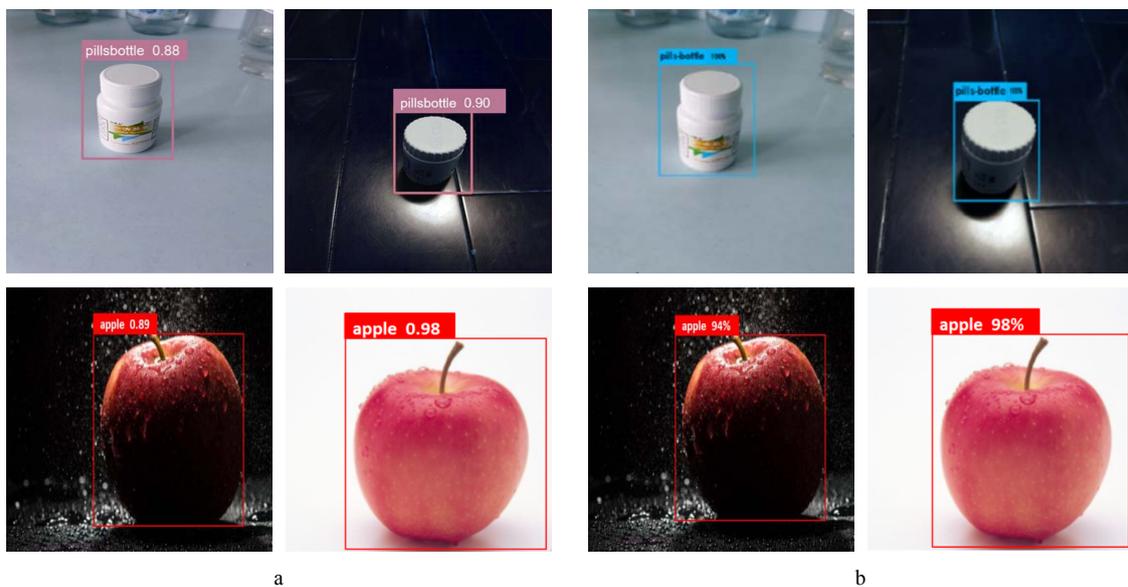


图 5 不同光照情况下的检测效果对比

Fig.5 Comparison of detection results under different lighting conditions



图6 不同尺度情况下的检测效果对比  
Fig.6 Comparison of detection results under different viewing angles



图7 不同视角下的检测效果对比  
Fig.7 Comparison of detection results under different viewing angles



图8 带遮挡情况下检测效果对比  
Fig.8 Comparison of Yolov3 and Yolov5s detection results with occlusion

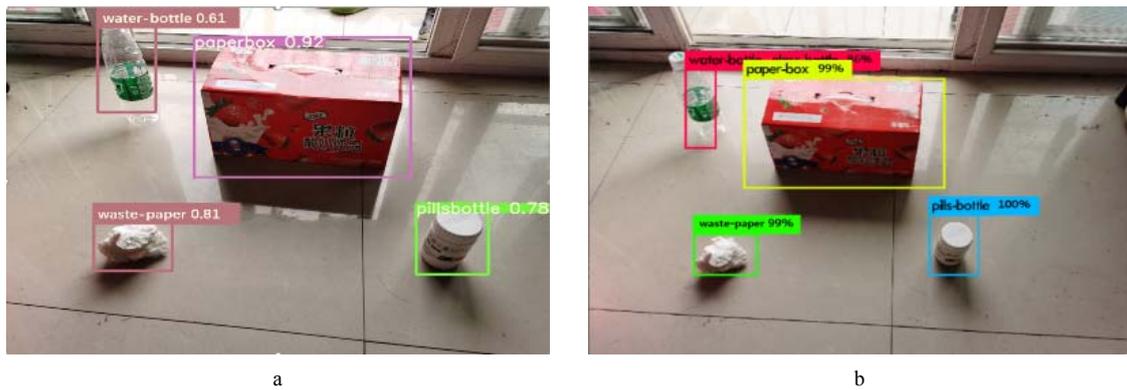


图 9 多目标检测效果对比  
Fig.9 Multi-target detection effect comparison

表 1 网络对比实验结果

Tab.1 Network comparison experiment results

模型	测试精度/%	mAP/%	训练时间/h	帧数/FPS
YOLOv3	89.43	91.47	48	2.98
YOLOv5s	95.34	99.38	36	6.67

#### 4 结语

为了实现垃圾的自动分类处理，基于 YOLOv5s 网络在自制垃圾分类数据集上进行训练，得到垃圾分类与检测模型。本次实验的垃圾分类检测模型网络参数量较少，可以从图像提取丰富的特征信息。实验结果表明，本次实验构建的基于 YOLOv5s 的垃圾分类检测模型整体性能优于 YOLOv3，检测精度高、鲁棒性好、体积小、速度快。

致谢：本文得到南京工业大学高性能计算中心的计算支持。

#### 参考文献：

[1] 中华人民共和国住房和城乡建设部. 关于加快推进部分重点城市生活垃圾分类工作的通知[EB/OL]. (2017-12-20)[2021-01-15]. <https://wenku.baidu.com/view/4cbafa81cdbff121dd36a32d7375a417866fc190.html>. Ministry of Housing and Urban-Rural Development of People Republic of China. Notice on Accelerating the Classification of Domestic Waste in Some Key Cities [EB/OL]. (2017-12-20)[2021-01-15]. <https://wenku.baidu.com/view/4cbafa81cdbff121dd36a32d7375a417866fc190.html>.

[2] YANG M, THUNG G. Classification of Trash for Recyclability Status[EB/OL]. (2017-09-22)[2021-01-15]. <http://cs229.stanford.edu/proj2016/report/>.

[3] MAO W L, CHEN W C, WANG C T, et al. Recycling Waste Classification Using Optimized Convolutional Neural Network[EB/OL]. (2020-09-08)[2021-01-15]. <http://doi.org/10.1016/j.resconrec.2020.105132>.

[4] TOĞAÇAR M, ERGEN B, CÖMERT Z. Waste Classification Using Auto Encoder Network with Integrated Feature Selection Method in Convolutional Neural Network Models[EB/OL]. (2019-12-31)[2021-01-15]. <http://doi.org/10.1016/j.measurement.2019.107459>.

[5] RABZNO S L, CABATUAN M K, SYBINGCO E, et al. Common Garbage Classification Using Mobile Net[C]. BaguioPhilippines: International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management, 2018.

[6] LIU Y, GE Z, LV G, et al. Research on Automatic Garbage Detection System Based on Deep Learning and Narrowband Internet of Things[J]. Journal of Physics: Conference Series, 2018: 1069-1076.

[7] ARAL R A, KESKINE Ş R, KAYA M, et al. Classification of Trashnet Dataset Based on Deep Learning Models[C]. Seattle: IEEE International Conference on Big Data (Big Data), 2018.

[8] RUIZ V, SANCHEZ Á, VELEZ J F, et al. Automatic Image-based Waste Classification[C]. Almeria, Spain,: International Work Conference on the Interplay between Natural and Artificial Computation, 2019.

[9] 许伟, 熊卫华, 姚杰,等. 基于改进 YOLOv3 算法在垃圾检测上的应用[J]. 光电子·激光, 2020, 31(9): 928-937.

XU Wei, XIONG Wei-hua, YAO Jie, et al. Application of Garbage Detection Based on Improved YOLOv3 Algorithm[J]. Journal of Optoelectronics·Laser, 2020, 31(9): 928-937.

[10] 宁凯, 张东波, 印峰,等. 基于视觉感知的智能扫地机器人的垃圾检测与分类[J]. 中国图像图形学报, 2019, 24(8) :1358-1368.

NING Kai, ZHANG Dong-bo, YIN Feng, et al. Garbage Detection and Classification of Intelligent Sweeping Robot Based on Visual Perception[J]. Journal of Image and Graphics, 2019, 24(8): 1358-1368.

[11] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-time Object Detection [C]. Las Vegas: IEEE Conference on Computer Vision

- and Pattern Recognition, 2016.
- [12] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger[C]. Honolulu, USA,: IEEE Conference on Computer Vision and Pattern Recognition, 2017.
- [13] REDMON J, FARHADI A. YOLOv3: An Incremental Improvement[EB/OL]. (2018-04-08)[2021-01-15]. <https://arxiv.org/1804.02767>.
- [14] BOCHKOVSKIY A, WANG C Y, LIAO H M. YOLOv4: Optimal Speed and Accuracy of Object Detection[EB/OL]. (2020-04-23)[2021-01-15]. <https://arxiv.org/abs/2004.10934>.
- [15] YUN S, HAN D, Oh S J, et al. CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features[EB/OL]. (2019-08-07)[2021-01-15]. <https://arxiv.org/abs/1905.04899v2>.
- [16] HUANG G, LIU Z, LAURENS V D M, et al. Densely Connected Convolutional Networks[C]. Honolulu: Conference on Computer Vision and Pattern Recognition, 2017.
- [17] TRINH H C, LE D H, KWON Y K. PANET: A GPU-Based Tool for Fast Parallel Analysis of Robustness Dynamics and Feed-Forward/Feedback Loop Structures in Large-Scale Biological Networks[J]. PLoS ONE, 2014, 9(7): e103010.
- [18] REZZTOFIGHI H, TSOI N, GWAK J Y, et al. Generalized Intersection over Union: A Metric and A Loss for Bounding Box Regression[C]. Long Beach, USA,: IEEE Conference on Computer Vision and Pattern Recognition, 2019.

---

(上接第49页)

- [11] CHEN S C, LI S H, LI C Y. Recent Related Research in Technology Acceptance Model: a Literature Review[J]. Australian Journal of Business and Management Research, 2011, 1(9): 124.
- [12] 何人可, 王玥虹, 马超民. 家用服务机器人人机交互亲密度意象研究[J]. 包装工程, 2021, 2(22): 1-8.  
HE Ren-ke, WANG Yue-hong, MA Chao-min. Intimacy and Perceptual Imagery of Home-use Robot Interaction[J]. Packaging Engineering, 2021, 2(22): 1-8
- [13] HAIR J F, RISHER J J, SARSTEDT M, et al. When to Use and How to Report the Results Of Pls-Sem[J]. European Business Review, 2019, 31(1): 2-24
- [14] DIJKSTRA T K, HENSELER J. Consistent Partial Least Squares Path Modeling[J]. MIS Quarterly, 2015, 39(2): 297-316.
- [15] FORNELL C, LARCKER D F. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error[J]. Journal of Marketing Research, 1981, 18(1): 39-50.
- [16] HENSELER J, RINGLE C M, SARSTEDT M. A New Criterion for Assessing Discriminant Validity in Variance-based Structural Equation Modeling[J]. Journal of the Academy of Marketing Science, 2015, 43(1): 115-135.