

# 基于 Faster RCNN 的可回收物自动分类算法研究

李威,李楠,于清玲

(沈阳工业大学信息科学与工程学院,辽宁沈阳 110870)

**摘要:**利用机器视觉对可回收物进行自动分类识别可以提高废品回收再利用的效率。应用改进的 Faster RCNN 算法对办公区四种常见的可回收物,塑料瓶、易拉罐、玻璃瓶和纸张进行自动分类识别。通过对特征提取网络和区域建议网络进行改进,提升 Faster RCNN 算法的检测精度。对第三方专业数据平台科赛网提供的可回收物图像库和自采图像库分别进行训练和测试。对存在相似特征的玻璃瓶和塑料瓶误检率较高问题,引入待检目标的重量特征,提高了对二者的召回率。实验结果表明,结合重量特征的改进 Faster RCNN 算法在四类可回收物识别精度上分别为 92.28%、87.78%、89.25% 和 91.24%,平均检测时间为 0.202 秒,提升了可回收物自动分类识别的效率。

**关键词:**垃圾分类;深度学习;目标检测;Faster RCNN;区域建议网络;锚框

中图分类号:TP391.4

文献标识码:A

文章编号:1673-629X(2021)0100-06

## Research on Automatic Classification Algorithm of Recyclables Based on Faster RCNN

LI Wei, LI Nan, YU Qing-ling

(School of Information Science and Engineering, Shenyang University of Technology, Shenyang 110870, China)

**Abstract:** Using machine vision to automatically classify and identify recyclables can improve the efficiency of waste recycling. Using the improved Faster RCNN algorithm to automatically classify and identify four common recyclables in the office area, plastic bottles, cans, glass bottles and paper. Improve the detection accuracy of Faster RCNN algorithm by improving the feature extraction network and region proposal network. Train and test the recyclables image library provided by the third-party professional data platform Kesai.com, and train and test self-collected image library. For glass bottles and plastic bottles with similar characteristics, the false detection rate is relatively high. The introduction of the weight characteristics of the target to be inspected improves the recall rate of both. The experimental results show that the improved Faster RCNN algorithm combined with weight characteristics has 92.28%, 87.78%, 89.25% and 91.24% in the recognition accuracy of the four types of recyclables, and the average detection time is 0.202 seconds, which improves the efficiency of automatic classification and recognition of recyclables.

**Key words:** garbage classification; deep learning; target detection; Faster RCNN; region proposal network; anchor

## 0 引言

可回收物的分类是当前制约国内环保事业发展的瓶颈之一,也是造成环境污染、资源再利用困难的根源之一<sup>[1]</sup>。人们在办公生活中会产生大量的可回收物,其中包括报废的打印纸张、空的塑料瓶、玻璃瓶和易拉罐。对可回收物实现自动分类既可以节约人力成本,也可以避免物资浪费<sup>[2]</sup>,方便物资的二次回收利用,因此对基于机器视觉的可回收物分类算法进行研究具有一定的实际应用意义。

国外对垃圾分拣技术的研究起步较早。芬兰 ZenRobotics 公司 2012 年研制出一款机器人<sup>[3]</sup>,通过传感器和探测器从成堆的建筑垃圾中分拣出金属、混凝

土、木材和塑料等可回收利用的材料。2017 年,日本垃圾处理公司对 ZenRobotics 的分拣系统进行改进,提高了机器人分拣速度和单臂负荷重量。以上两篇文献是针对建筑垃圾研发的分拣系统。在垃圾分类算法研究方面,文献[4]通过网络检索和自采图像组成 400 张图像的数据集,分别采用 KNN、Nave Bayes、决策树和随机森林四种不同的算法,对无垃圾区和高密度垃圾区进行分类,其中 KNN 算法的准确率高达 98.6%;文献[5]根据区域中垃圾量的多少对街道的污染程度进行 1 到 5 等级的分类,分别采用 KNN、CNN、随机森林和 SVM 算法进行分类,其中使用 CNN 算法的准确率最高,为 89.68%,然而其仅仅对图像中垃圾的多少

收稿日期:2021-02-14

作者简介:李威(1979-),女,博士,研究生导师,研究方向为图像处理与模式识别;李楠(1996-),女,硕士研究生,研究方向为图像处理与模式识别。

进行分类判断,缺少对垃圾具体类别的标注和分类。以上两篇文献都是对街道的整洁度进行辨识,即根据场景中垃圾数量的多少将场景分类成垃圾区域和清洁区域,并没有对垃圾的具体类别进行标注、划分。文献[6]通过改进TernausNet深度学习框架,设计并实现了一个可以检测并分割图像中垃圾区域的系统,数据集共分为6个类别,分别是非垃圾、塑料、金属、纸张、玻璃和其他,改进后的框架对图像中垃圾区域分割的DICE系数达到0.86,但对不同类别垃圾的检测仍需进一步改进。

国内随着对垃圾分类的重视,文献和发明专利逐渐增多。文献[7]为满足工业分拣系统的实时性要求,设计了一款基于Faster RCNN的分拣机器人,该款机器人仅对目标位置进行预测,并未对目标类别进行详细划分;专利[8]提出一种基于CNN和小波分析的办公垃圾智能分类方法;专利[9]设计了一款针对纸类、瓶类和金属类的自动分类回收环保设备;专利[10]提供了一种办公垃圾分类处理装置的设计思路,通过卷积神经网络对纸类、瓶类和金属类进行分类回收。专利[11]设计了一款按照重量进行垃圾分类且进行预处理的垃圾桶,该款垃圾桶根据重量对重型垃圾和轻垃圾进行分类。以上这些专利没有给出具体待分类的可回收物图片,算法描述也比较简单,但为可回收物自动分类、回收提供了一定的解决思路。

近年来,基于深度学习的目标检测已经成为计算机视觉的一个研究热点<sup>[12]</sup>。相对于传统方法,基于深度学习的目标检测算法精度更高,泛化能力更强。目前,用于目标检测的深度学习框架可分为两类<sup>[13]</sup>,一类是以R-CNN(region convolutional neural network)为代表的二步式目标检测框架(two stage),这一类框架以分类为基础,对候选区域进行再分类和位置矫正

得到结果,代表算法有RCNN、Fast RCNN、Faster RCNN;另一类是以YOLO(you only look once)为代表的一步式目标检测框架(one stage),这一类框架将整幅图像作为输入,直接输出分类结果以及目标位置,代表算法有YOLO、SSD。一步式的目标检测具有更快的检测速度,二步式目标检测具有更高的检测精度和更低的漏检率。二步式目标检测的典型算法Faster RCNN<sup>[14]</sup>引入区域建议网络(region proposal network, RPN)代替选择性搜索来获取目标位置,大大缩短了目标位置检测所耗费的时间。同时,该算法引入锚框<sup>[15]</sup>来应对目标形状出现变化的问题,提升了目标检测的精度。文中对Faster RCNN算法进行了改进,利用残差网络ResNet50<sup>[16]</sup>代替原算法中的VGG16进行特征提取,获得目标图像更深层的语义特征,通过K-means<sup>[17]</sup>聚类算法对锚框尺寸进行改进,使得锚框尺寸更符合文中数据集。利用改进后的Faster RCNN算法对办公区四种常见可回收物,报废的打印纸张和旧书籍、空的塑料瓶、玻璃瓶和易拉罐进行分类检测,并在第三方专业数据平台科赛网提供的可回收物图像库和自采图像库对算法的准确性和实时性方面进行了验证。

## 1 改进的Faster RCNN算法

### 1.1 Faster RCNN算法原理

Faster RCNN算法的整体结构可以分为三部分,分别是特征提取网络、区域建议网络和分类回归网络<sup>[18]</sup>。特征提取网络对待检测图像进行特征提取,区域建议网络确定目标的候选区域,分类回归网络对目标候选区域进行分类和位置回归。Faster RCNN算法进行可回收物分类时的网络结构如图1所示。

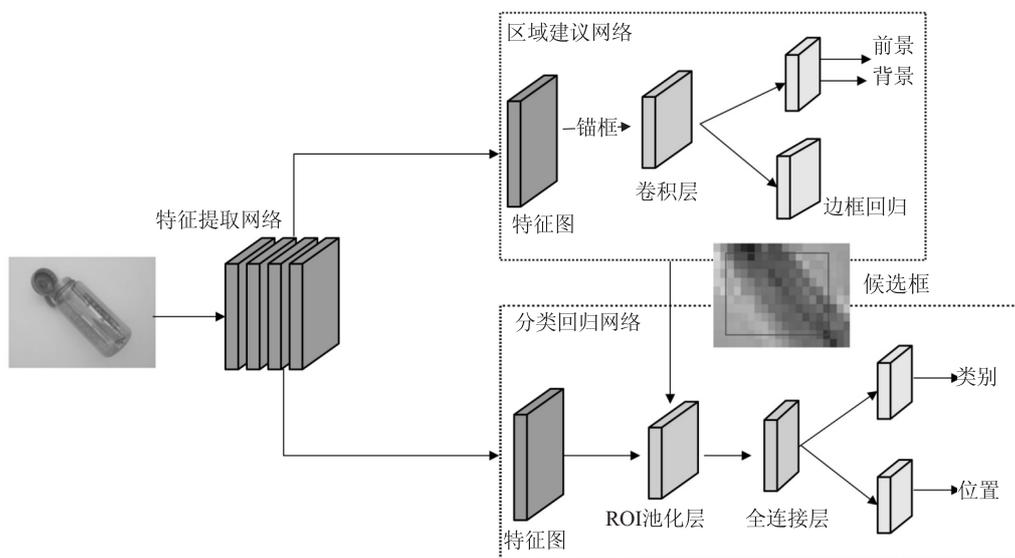


图1 Faster RCNN网络结构

## 1.2 Faster RCNN 算法的改进

### 1.2.1 特征提取网络

原 Faster RCNN 算法采用网络结构较浅的 VGG16 作为特征提取网络, VGG16 更适用于目标物占图像比例小、分辨率较低的情况, 而文中数据集中的可回收物在图像中面积比例大、目标清晰, 为获取深层语义特征, 提高分类精度, 更适用网络结构较深的模型。因此, 选用 ResNet50 网络模型代替原算法中的 VGG16 模型。同时, 由于 ResNet50 网络模型中的残差结构, 有效避免随着网络层数的加深而导致的过拟合现象。通过残差块, 将浅层的纹理、颜色和轮廓等特征与深层语义特征结合, 既避免了网络层数的加深导致的网络退化现象, 同时又提高了识别准确率。

为避免由于网络结构的加深, 造成模型计算量和难度的提高, ResNet50 利用瓶颈结构, 即采用  $1 \times 1$  的卷积核进行升维和降维操作, 因此, ResNet50 相比 VGG16, 尽管网络层数加深, 但计算复杂度并没有提高。瓶颈结构下的残差模块如图 2 所示。

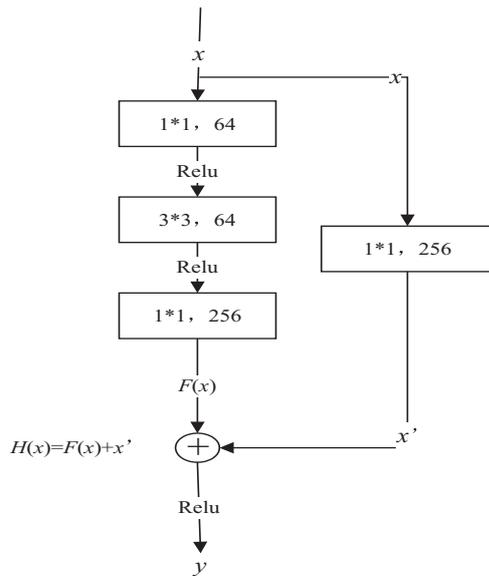


图 2 残差模块结构

其中,  $x$  表示输入,  $x'$  表示  $x$  经过维度变化后的输出,  $F(x)$  表示下一层卷积层的输出, 同时也是原网络模型下一层激活函数的输入,  $H(x)$  为网络训练的目标, 但

表 1 不同锚框尺寸结果

参数	锚框长宽比(像素)										平均重叠度/%
原算法	128 : 128	91 : 182	182 : 91	256 : 256	181 : 362	362 : 181	512 : 512	362 : 724	724 : 362		67.89
K=9	325 : 350	342 : 143	427 : 192	409 : 284	457 : 378	208 : 93	127 : 184	187 : 314	297 : 236		83.54

根据表 1 可知, 通过 K-means 聚类算法重新生成的锚框尺寸, 与数据集中图像的平均重叠度可达到 83.54%, 相较原算法提高了 15.65%。因此文中采用重新生成的 9 种尺寸的锚框进行模型训练。

由于残差网络中的跳跃连接, 下一层激活函数的输入变为  $F(x) + x'$ , 因此训练目标由原网络的  $F(x) = H(x)$  转换为学习一个残差函数, 即  $F(x) = H(x) - x'$ 。若  $F(x) = 0$ , 则为恒等映射  $H(x) = x'$ , 若  $F(x) \neq 0$ , 则  $F(x) = H(x) - x'$  作为残差模块需要学习的残差部分。

根据图 2 可知, 一个残差块的输出  $y$  为:

$$y = \sigma(H(x)) = \sigma(F(x) + x') = \sigma(W_3 \sigma(W_2 \sigma(W_1 x)) + W_s x) \quad (1)$$

式中,  $\sigma$  表示激活函数 Relu,  $W_1$  表示第一层的网络权值,  $W_2$  表示第二层的网络权值,  $W_3$  表示第三层的网络权值, 由于  $x$  与  $F(x)$  的相加为逐元素相加,  $W_s$  为线性映射, 使得  $x$  与  $F(x)$  维度保持一致。

### 1.2.2 K-means 聚类算法重新定义锚框尺寸

原算法的锚框尺寸根据 PASCAL VOC 数据集得出, 即 3 个面积大小  $128^2$ 、 $256^2$ 、 $512^2$ , 3 种比例  $1 : 1$ 、 $1 : 2$ 、 $2 : 1$ , 共 9 种尺寸。然而, 原算法锚框与文中数据集的平均重叠度仅为 67.89%, 即原算法锚框尺寸并不适用于文中数据集的目标。重叠度 IoU 的计算公式为:

$$IoU = \frac{S(A \cap G)}{S(A \cup G)} \quad (2)$$

其中,  $S(A \cap G)$  为锚框与标定框相交的面积,  $S(A \cup G)$  为锚框与标定框相并的面积。

故根据文中数据集中可回收物目标的长宽比, 采用 K-means 聚类算法重新生成 9 种尺寸的锚框, 即在所有目标的长宽比数据中, 迭代确定 9 个中心, 使得聚类后的损失函数最小, 即各个数据与所属类中心点的距离平方和最小, 损失函数的计算公式为:

$$L(c, \mu) = \sum_{i=1}^N \|x_i - \mu_{c_i}\|^2 \quad (3)$$

其中,  $N$  是数据总数,  $x_i$  是第  $i$  个样本,  $c_i$  是  $x_i$  所属的类,  $\mu_{c_i}$  是类对应的中心点。

计算重新生成的锚框与数据集的平均重叠度, 并与原锚框的平均重叠度进行比较。文中使用的数据集中图像尺寸均为  $512 \times 384$  像素, 通过 K-means 聚类算法得到的锚框尺寸与原算法锚框尺寸的平均重叠度结果如表 1 所示。

## 2 实验数据集

### 2.1 数据采集

实验使用的数据集分为两个部分: (1) 第三方专

业数据平台科赛网提供的可回收物图像数据库;(2)利用本研究团队开发设计的智能垃圾桶对报废的打印纸张、旧书报、空的塑料瓶、玻璃瓶和易拉罐进行采样,自建了包括四种常见办公可回收物的数据图像库。

数据图像库(1)原有图像2390张。选出玻璃瓶、易拉罐、纸张、塑料瓶四类可回收物的图像,并对碎纸

张、碎玻璃、不规则金属和塑料等不合格数据进行筛选删除。筛选后的数据集包含521张图像,图像尺寸均为512\*384像素,均是在白板上在日光/室内光源下进行拍摄的。数据图像库(1)的图像示例及说明如表2所示。

表2 数据集(1)的图像示例和说明

垃圾	示例	说明
玻璃瓶		包含玻璃杯、啤酒瓶、玻璃罐等多种形状的玻璃瓶
易拉罐		包含易拉罐、金属桶以及经过挤压的易拉罐
纸张		包含多种颜色纸张,但纸张均近似为矩形,且占面积比较大
塑料瓶		包含塑料瓶、塑料桶等多种形状的空塑料瓶

数据图像库(2)包括100张数据图像,同时还引入了可回收物的目标重量数据,从而解决了塑料瓶和玻璃瓶不易区分的问题。数据集图像尺寸为512\*384像素。为避免由于数据量不足产生的过拟合和泛化性能低等问题,采用数据增强来增加数据量,通过对

图像上下翻转、左右翻转、随机裁剪、缩放、亮度变换、对比度变换和人工噪声扰动等,将数据集扩大8倍,即自采数据图像库共包含800张图像,数据图像库(2)的图像示例及说明如表3所示。

表3 数据集(2)的图像示例和说明

垃圾	示例	说明
玻璃瓶		包含多种形状的玻璃瓶,重量最轻为200克,最重为510克
易拉罐		图像中目标均为易拉罐
纸张		包含多种颜色的纸张、书本,且目标均为矩形
塑料瓶		包含多种形状和大小的空塑料瓶,重量最轻为18克,最重为45克

## 2.2 数据集标定

利用LabelImg软件对数据集中的数据图像进行标定。使用矩形标定框确定图像中的垃圾目标位置,记录每张图像中目标的位置和类别。进行图像标定时,采用外接矩形的方式确定垃圾目标的位置,减少背景信息所占区域,避免算法实现过程中因背景信息造成模型训练难收敛以及误检率高的情况。标定后可以得到每张图像的位置和类别信息,包括标定框的四点坐标以及垃圾目标的类别,采用PASCAL VOC数据集的保存格式,即xml格式,后经程序转换为训练模型时可识别的tfrecord格式。

## 3 实验结果与分析

文中实验所使用的硬件平台为英特尔I5-4590四核CPU,开发环境为Windows10(64位)系统下的python3.6语言,实验模型基于Tensorflow1.13.1深度学习框架进行搭建。

使用数据集(1)进行模型训练,对模型进行评估,分别计算四种类别垃圾的精确率 $P$ ,召回率 $R$ 和F1值,计算公式为:

$$P = \frac{TP}{TP + FP} \times 100\% \quad (4)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (5)$$

$$F1 = \frac{2PR}{P + R} \quad (6)$$

其中,TP 表示真正例,即被判定为正样本,实际也为正样本;FP 表示假正例,即被判定为正样本,实际却为负样本;FN 表示假反例,即被判定为负样本,实际却为正样本。

测试不同类别垃圾的分类检测能力,从而有针对性地提高其准确率,训练后模型对于不同类别的垃圾

分类检测评估结果如表 4 所示。在数据集(1)中,由于玻璃瓶图像和塑料瓶图像包含较多相似特征,因此在分类时容易混淆,导致召回率不高。实验中,会有 13.1% 的玻璃瓶被误判为塑料瓶,11.49% 的塑料瓶被误判为玻璃瓶。而尽管易拉罐的召回率较高,但精确率却较低,分析实验结果,由于数据集(1)中易拉罐类别的图像大部分为矩形,且包含图像较少,会有 4.73% 的塑料、3.57% 的玻璃和 5.43% 的纸张误判为易拉罐。

表 4 不同类别垃圾分类检测评估

垃圾类别	精确率 P / %	召回率 R / %	F1
玻璃瓶	71.58	80.95	0.76
易拉罐	72.84	86.76	0.79
纸张	97.54	89.59	0.93
塑料瓶	86.96	81.08	0.84

为进一步提高玻璃瓶和塑料瓶的召回率,文中在数据集(2)中引入了待分类目标的重量数据。通过对空玻璃瓶和空塑料瓶的重量进行统计,空玻璃瓶的平均重量为 450 g,空塑料瓶的平均重量为 30 g(2 L 的塑料瓶重量为 45 g)。因此,首先通过模型对待检测目标进行分类,对分类为塑料瓶或玻璃瓶的目标,利用重量阈值进一步细分类,若目标重量大于重阈值,则将其分类为玻璃瓶,若目标重量小于重量阈值,则将其分

类为塑料瓶。

塑料瓶和玻璃瓶的召回率与重量阈值如图 3 所示,当重量阈值为 100 克时,塑料瓶和玻璃瓶的召回率分别为 92% 和 93%,当重量阈值为 150 克时,塑料瓶和玻璃瓶的召回率分别为 92% 和 91.8%,因此对于数据集(2),确定重量阈值为 100 克。

改进后模型对四个类别的可回收物检测结果如图 4 所示。

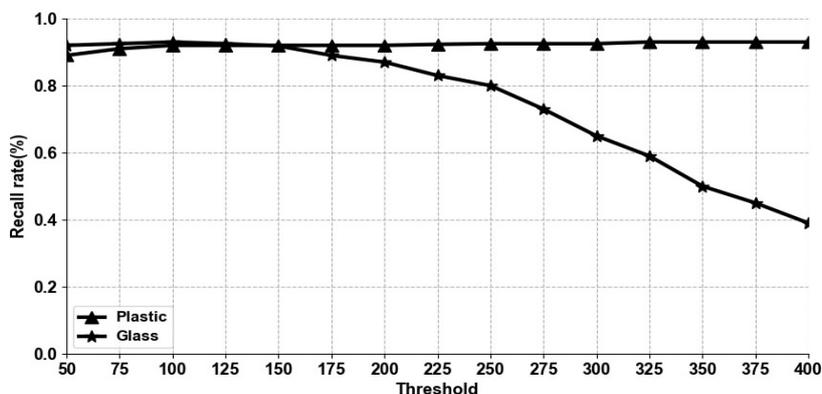


图 3 重量阈值与召回率

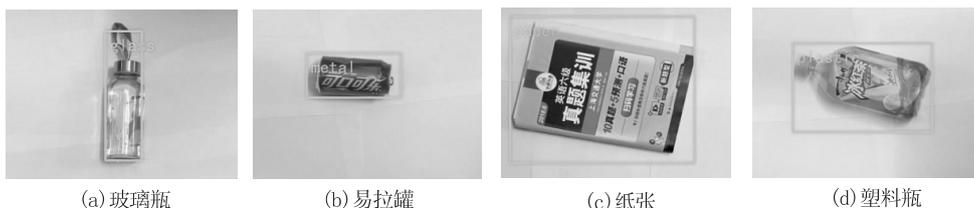


图 4 模型检测结果

利用传统目标检测算法,即 SVM 算法进行分类,选择性搜索算法进行位置检测。并对传统目标检测算法,原 Faster RCNN 模型、改进后 Faster RCNN 模型、改进后模型与重量阈值结合,在数据集(2)上进行测

试,得到四个类别的准确率、平均准确率、检测时间以及检测框与标定框的平均重叠度,如表 5 所示。

根据表 5 可知,基于 Faster RCNN 的目标检测算法相较传统算法具有明显优势,而改进后 Faster

RCNN的模型较原Faster RCNN模型,平均准确率提高了3.86%,检测时间缩短了0.066秒每张图像,与重量阈值相结合的模型相较原Faster RCNN模型,平均准确率提高了8.43%,检测时间缩短了0.05秒每张图像,但由于增加了重量阈值的判断,检测时间相较改进后模型增加了0.016秒每张图像,通过改进锚框尺

寸,平均重叠度也有原先的76.59%增加至91.25%,使得对目标位置的判断更为准确。但文中对办公区可回收垃圾分类检测仍存在许多不足,对数据集依赖性较强,且垃圾种类局限,因此需要不断扩充数据集来增强模型泛化能力。

表5 改进前后模型评估

模型	准确率/%				平均准确率/%	检测时间/(秒/张)	平均重叠度/%
	玻璃瓶	易拉罐	纸张	塑料瓶			
SVM+选择性搜索	67.30	75.46	78.31	69.48	72.64	0.190	74.30
原Faster RCNN	78.57	84.49	86.07	77.70	81.71	0.252	76.59
改进Faster RCNN	82.55	87.78	89.25	82.68	85.57	0.186	91.25
改进Faster RCNN+重量阈值	92.28	87.78	89.25	91.24	90.14	0.202	91.25

#### 4 结束语

文中提出基于Faster RCNN算法的办公区可回收垃圾分类,并对目标检测算法进行两点改进,首先采用ResNet50作为特征提取网络,其次在区域建议网络中通过K-means聚类算法重新定义锚框尺寸。通过自采数据集,引入目标重量特征,设定重量阈值为100克,降低对玻璃瓶和塑料瓶的误检率,对自采数据集进行模型训练,结果表明优化后的算法的准确率提高了8.43%,检测速度缩短了0.05秒每张图像。该可回收物分类检测模型可搭配智能分拣系统进行可回收物的自动分拣。

#### 参考文献:

- [1] 周滢慧. 基于机器视觉的生活垃圾智能分拣系统的设计与实现[D]. 哈尔滨:哈尔滨工业大学,2018.
- [2] 千承辉,侯天远,肖若兰,等. 基于机器视觉的垃圾分拣机器人设计[J]. 单片机与嵌入式系统应用,2019,19(9):39-42.
- [3] LUKKA T J, TOSSAVAINEN T, KUJALA J V, et al. Zen-Robotics recycler - robotic sorting using machine learning [C]//Proceedings of the international conference on sensor-based sorting (SBS). Helsinki, Finland: [s. n.], 2014:1-8.
- [4] GHANSHALA T, TRIPATHI V, PANT B. A machine learning based framework for intelligent high density garbage area classification [C]//Proceedings of the future technologies conference (FTC) 2020. [s. l.]: Springer, 2020.
- [5] BOSE C, PATHAK S, AGARWAL R, et al. A computer vision based approach for the analysis of acuteness of garbage [C]//Advances in computing and data sciences. Singapore: Springer, 2020.
- [6] DESPOTOVSKI A, DESPOTOVSKI F, LAMESKI J, et al. Towards cleaner environments by automated garbage detec-

tion in images[C]//Machine learning and applications. [s. l.]: Springer, 2020:56-63.

- [7] 孙雄峰,林 浒,王诗宇,等. 基于改进Faster RCNN的工业机器人分拣系统[J]. 计算机系统应用,2019,28(9):258-263.
- [8] 程龙君,胡 锋,卞 凯,等. 一种基于CNN和小波分析的办公室垃圾智能分类方法:中国,10246052.2[P]. 2020-07-28.
- [9] 石永军,赵学进,李县法,等. 一种垃圾自动分类资源回收环保设备:中国,10704949.5[P]. 2020-09-25.
- [10] 丁 泉,胡 峰,卞 凯,等. 一种办公室垃圾智能化分类处理装置:中国,10246118.8[P]. 2020-07-14.
- [11] 何洪森. 一种按照重量进行垃圾分类且进行预处理的垃圾桶:中国,202010673381.5[P]. 2020-07-14.
- [12] 段仲静,李少波,胡建军,等. 深度学习目标检测方法及其主流框架综述[J]. 激光与光电子学进展,2020,57(12):59-74.
- [13] 李明熹,林正奎,曲 毅. 计算机视觉下的车辆目标检测算法综述[J]. 计算机工程与应用,2019,55(24):20-28.
- [14] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6):1137-1146.
- [15] 黄继鹏,史颖欢,高 阳. 面向小目标的多尺度Faster-RCNN检测算法[J]. 计算机研究与发展,2019,56(2):319-327.
- [16] 徐义鑫,贺 鹏,任 东,等. 基于改进faster RCNN的木材运输车辆检测[J]. 计算机应用,2020,40(S1):209-214.
- [17] 杜玉红,董超群,赵 地,等. 改进Faster RCNN模型在棉花异性纤维识别中的应用[J]. 激光与光电子学进展, 2020,57(12):132-141.
- [18] FAN Q, BROWN L, SMITH J. A closer look at Faster R-CNN for vehicle detection[C]//2016 IEEE intelligent vehicles symposium (IV). Gothenburg: IEEE, 2016:124-129.