



## 基于深度学习的垃圾分类方法综述

李金玉<sup>1,2,3</sup>, 陈晓雷<sup>1,2,3</sup>, 张爱华<sup>1,2,3</sup>, 李 策<sup>1,2,3</sup>, 林冬梅<sup>1,2,3</sup>

(1. 兰州理工大学 电气工程与信息工程学院, 兰州 730050;

2. 兰州理工大学 甘肃省工业过程先进控制重点实验室, 兰州 730050;

3. 兰州理工大学 电气与控制工程国家级实验教学示范中心, 兰州 730050)

**摘 要:** 垃圾分类是保护生态环境、促进经济发展的有效措施, 利用深度学习进行垃圾分类已成为当前学术界和工业界的研究热点。传统垃圾分类主要由人工进行分拣和分类, 存在劳动强度大、分选效率低、工作环境差等缺点, 急需智能化、自动化的分类方法来替代。近年来研究人员已经开始初步探索利用深度学习技术进行垃圾分类并提出一些有效的方法。从方法、数据集和研究方向等方面分析深度学习垃圾分类方法的研究现状, 介绍不同深度学习模型在垃圾分类中的应用和发展, 研究基于 ResNet 方法、基于 DenseNet 方法、基于单阶段目标检测方法和基于卷积神经网络与迁移学习相结合方法等多种典型方法的性能和特点并对比其优缺点, 对现有的垃圾分类公开数据集进行概述与总结。在此基础上, 分析深度学习在垃圾分类领域面临的挑战, 并对其发展趋势及未来的研究方向进行展望。

**关键词:** 垃圾分类; 深度学习; 卷积神经网络; ResNet 系统; DensenNet 系统; 单阶段目标检测

开放科学(资源服务)标志码(OSID):



中文引用格式: 李金玉, 陈晓雷, 张爱华, 等. 基于深度学习的垃圾分类方法综述[J]. 计算机工程, 2022, 48(2): 1-9.

英文引用格式: LI J Y, CHEN X L, ZHANG A H, et al. Survey of garbage classification methods based on deep learning[J]. Computer Engineering, 2022, 48(2): 1-9.

## Survey of Garbage Classification Methods Based on Deep Learning

LI Jinyu<sup>1,2,3</sup>, CHEN Xiaolei<sup>1,2,3</sup>, ZHANG Aihua<sup>1,2,3</sup>, LI Ce<sup>1,2,3</sup>, LIN Dongmei<sup>1,2,3</sup>

(1. College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China;

2. Key Laboratory of Gansu Province Advanced Control for Industrial Processes, Lanzhou University of Technology, Lanzhou 730050, China; 3. National Demonstration Center for Experimental Electrical and Control Engineering Education, Lanzhou University of Technology, Lanzhou 730050, China)

**[Abstract]** Garbage classification is an effective measure to protect ecological environment and promote economic development. Traditional garbage classification relies heavily on manual work in waste sorting. It is labor-intensive and limited in efficiency, and workers have to suffer from the poor environment. In recent years, intelligent and automated garbage classification methods using deep learning has become a hot research topic. This paper reviews the existing studies of deep learning-based garbage classification from the perspectives of method, dataset and research direction, and introduces the application and development of different deep learning models in garbage classification. The paper analyzes the performance and features of various typical methods, such as the ResNet-based method, DenseNet-based method, single-stage target detection method and the method combining convolution neural network with transfer learning, comparing their advantages and disadvantages. In addition, the paper summarizes the existing public datasets of garbage classification. On this basis, the paper discusses the current challenges faced by deep learning applications in the field of garbage classification, and the development trends as well as future research directions.

**[Key words]** garbage classification; deep learning; Convolutional Neural Network(CNN); ResNet system; DensenNet system; single-stage target detection

DOI: 10.19678/j.issn.1000-3428.0061338

基金项目: 国家自然科学基金(61967012, 61866022)。

作者简介: 李金玉(1996—), 女, 硕士研究生, 主研方向为模式识别、智能系统; 陈晓雷(通信作者), 副教授、博士; 张爱华、李 策, 教授、博士; 林冬梅, 副教授、博士。

收稿日期: 2021-04-01 修回日期: 2021-07-14 E-mail: chenxl703@lut.edu.cn

## 0 概述

随着我国经济的高速发展,城镇人口日益增多,城市生活垃圾迅猛增长。逐年增长的城市生活垃圾,不仅对城市发展带来阻力,而且对生态环境的破坏带来不可估量的影响。因此,2018年由国务院发布的《关于全面加强生态环境保护坚决打好污染防治攻坚战的意见》中要求加快推进垃圾分类实施,到2020年实现所有城市生活垃圾分类处理全覆盖。垃圾分类已成为国家生态环境保护、促进经济发展的重要举措。

由于城市生活垃圾多数为多种废弃物的混合体,无论是采取卫生填埋、堆肥,还是焚烧的处理方式,都不能有效地解决生活垃圾所带来的问题。通过分类改变垃圾的混杂性是实现垃圾处理资源化、减量化、无害化的重要前提。传统的生活垃圾分类方法主要有筛分、重力分选、风力分选、浮力分选、磁力分选以及光点分选等,以上无论是哪种分选方法,人工分选都是不可或缺的一个环节,用于挑选这些传统分类方法无法识别的目标物,以及传统方法分类之后的进一步质量控制。然而,人工垃圾分选存在劳动强度大、监督任务重、分选效率低和工作环境差等问题,急需智能化、自动化的分类方法取代传统分类方法。

近年来,深度学习已渗透到各行各业,如何利用深度学习高效地进行垃圾分类已成为当前的研究热点。但是,目前还缺乏针对这一研究领域的综述性研究,本文主要对目前已有的深度学习垃圾分类方法进行综述,分析近年来深度学习相关方法在垃圾分类中的应用,将其归纳整合为基于ResNet的方法、基于DenseNet的方法、基于单阶段目标检测的方法和基于卷积神经网络与迁移学习相结合的方法,在此基础上总结各个方法的分类效果及优缺点,分析并探讨基于深度学习的垃圾分类研究面临的挑战和未来发展方向。

## 1 基于深度学习的垃圾分类方法

深度学习能够有效地解决传统机器学习中手工特征提取不全和选择繁琐的问题,挖掘样本数据的内在规律,并带来性能的改善,广泛应用于计算机视觉<sup>[1]</sup>、自然语言处理<sup>[2]</sup>、数据挖掘<sup>[3]</sup>等领域。因此,基于深度学习的垃圾分类方法为垃圾分类指出了一条新的道路,本节根据不同深度学习模型对现有深度学习垃圾分类方法进行阐述,并给出各种方法在垃圾分类领域的算法及模型的总结对比。

### 1.1 基于ResNet的方法

ResNet<sup>[4]</sup>基本框架如图1所示,其最大的特点是在进行特征提取时,直接将输入信息绕道传到输出端,极大增加网络深度,从最初的十几层增加到后来的150多层,并且不会产生过拟合现象。

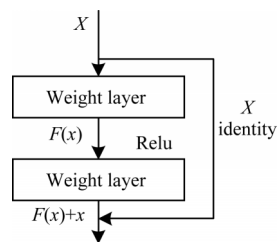


图1 ResNet基本框架

Fig.1 ResNet basic framework

文献[5]提出一种基于ResNet50和支持向量机(SVM)<sup>[6]</sup>的智能废料分类系统,该系统采用ResNet50进行特征提取,通过SVM对所提取到的特征进行分类,在TrashNet数据集上的准确率达到87%。然而,SVM只对小样本数据具有较好的分类效果,故该方法不具有通用性。

文献[7]提出基于ResNet和Inceptionv-4<sup>[8]</sup>的垃圾自动分类DSCR网络,该网络构建一个IR-Block使得模型在多尺度上获得更多的特征,解决了由于使用缩减模块(reduction module)导致模型过于敏感的问题,准确率达到94.38%。文献[9]提出一种注意力机制模型,该模型通过完成局部、全局的特征提取和特征融合机制等手段建立了垃圾图像分类模型GCNet,能获取更加有效完善特征信息,通过有效利用特征信息进而避免梯度消失的现象。

文献[10]提出的自动垃圾检测系统使用ResNet网络作为卷积层,输入数据集图像后通过ResNet的共享卷积层生成特征图,再通过RPN(区域提议网络)层得到输出,生成大量的区域建议,该方法很好地克服了区域误检问题。文献[11]修改了原始ResNext模型,在全局平均池化层之后添加了两个全连接层以提高网络的分类性能,能够快速实现分类模型的稳定和泛化,相比于ResNet-101,该方法对金属、纸张和塑料的分类性能表现较差,如何利用这些样本的独有特征提高分类精度值得深入探索。

文献[12]从输入图像的多特征融合、残差单元的特征重用和新激活函数的设计3个方面对ResNet-34算法进行优化,该方法改进的ResNet-34对小目标物体检测结果较好,但对嵌入式等移动设备而言,由于该方法运算量太大,无法保证垃圾分类的实时性。针对此问题,研究人员通过减少ResNet模型的参数量实现轻量级的垃圾分类,以促使分类任务能够在移动手机端、嵌入式等设备上执行。文献[13]提出一种基于轻量级残差网络MAPMobileNet-18,通过使用最大平均组合池化代替Bottleneck中的最大池化与平均池化,深度可分离卷积替换Bottleneck中的3×3卷积核,模型参数量在理论上减少为原来的1/8,在保持模型精度不变的同时

提高了模型的运算效率。在 ResNet-18 中的 BasicBlock 后串联添加通道注意力机制<sup>[14]</sup>和空间注意力机制<sup>[15]</sup>,用残差块替换了 BasicBlock 中的  $3 \times 3$  卷积核,使网络模型更适合使用移动手机端等计算能力较差的设备进行垃圾分类。

文献[16]基于 ResNet50 设计一种轻量级垃圾分类模型 GA\_MobileNet,利用深度卷积和分组卷积减少计算量和参数,通过通道注意力机制提高模型的准确性。相对于原始模型,该模型分类精度更高,占用内存更少,能够解决垃圾分类在嵌入式设备上的应用问题。

## 1.2 基于 DenseNet 的方法

DenseNet<sup>[17]</sup>为确保网络中最大的信息流通,使每层都与该层之前的所有层相连,即每层的输入是前面所有层的输出,从而可以有效减轻梯度弥散,加强特征的传播和复用,减少模型参数数量。

文献[18]将 AlexNet<sup>[19]</sup>、VGG16、ResNet50、DenseNet169 在数据集 TrashNet 上进行验证,结果表明,DenseNet169 性能更好,检测精度达到 94.9%。文献[20]提出一种新的密集连接模式构建了智能垃圾分类网络 RecycleNet,该网络改变了原始 DenseNet121 密集块内跳跃连接的连接模式,可以将 DenseNet121 网络中的参数数量从 700 万减少到大约 300 万,很大程度上缩短了检测时间。

DenseNet121 全连接层的分类能力是非常明显的,但全连接层的参数占整个网络参数的 80%。为了在提高分类性能和减少网络参数之间获得平衡,文献[21]利用遗传算法优化了 DenseNet121 全连接层的超参数以提高精度,实验结果表明,使用两个全连接层作为 DenseNet121 的分类器,相比配备有全局平均池化和 softmax 分类器的原始 DenseNet121,在垃圾分类任务上表现更好。

## 1.3 基于单阶段目标检测的方法

单阶段目标检测方法以结构简单、计算高效,同时具备较高检测精度的优势,在实时目标检测领域中具有广泛的研究和应用价值<sup>[22]</sup>。文献[23]采用改进的 YOLOv2 网络模型进行装修垃圾检测和识别,并通过目标盒维数聚类 and 分类网络预训练来提高 YOLOv2 模型的性能,使用 MobileNet<sup>[24]</sup>代替 YOLOv2 原来的 Darknet-19 进行垃圾检测,将整个改进的模型移植到嵌入式模块中得到较好的分类效果。

文献[25]提出一种 YOLOv2-dense 网络,该网络能够有效解决随着深度卷积深度增加而带来的梯度消失问题,并且还能够很好地复用与融合图像浅层特征和深层特征。改进后的 YOLOv2-Dense 算法可

以保留更多的浅层图像信息,提取目标特征的能力显著提高,可在不同光照、背景、视角与分辨率情况下进行精确实时的垃圾定位和检测。

文献[26]研究 YOLOv3 算法在垃圾分类中的应用,在自制数据集上的 mAP 值达到 94.99%。文献[27]构建了面向瓶罐识别分类的网络模型并进行实验验证,针对 YOLOv3 对小目标不敏感的问题,用 GIOU 替代传统的 IOU,通过 K-means 聚类算法来优化锚框尺寸。实验结果表明,改进后的 YOLOv3 算法识别精确率比原始算法提高约 4%。文献[28]提出一种新的基于 YOLO 的变分自动编码器神经网络模型,该模型由卷积特征提取器、卷积预测器和解码器组成。经过训练过程,该模型以 3 210 万个参数和 60 frame/s 的处理速度达到 69.70% 的准确率,使其可应用于现实生活的垃圾回收场景。

文献[29]介绍一种 LSSD 算法,该算法克服 SSD<sup>[30]</sup>算法存在小目标丢失和用不同大小框同时检测同一目标的缺点,建立一个在所有类型的尺度上都具有更强语义的新特征金字塔,将 Focal Loss<sup>[31]</sup>函数代替传统损失函数解决了单阶段目标检测方法正负样本比例严重失衡的问题。

## 1.4 基于卷积神经网络与迁移学习相结合的方法

深度学习的性能提高主要依赖于大量的训练数据集,与其他领域的大规模数据集(如 ImageNet 数据集)等相比,当前公开的垃圾数据集数量和种类很少,有限地标注垃圾数据集已成为深度学习方法在垃圾图像分类中的应用瓶颈。因此,研究人员利用迁移学习来解决垃圾分类的小样本数据集问题。

文献[32]提出一种基于迁移学习和 Inception-v3<sup>[33]</sup>的垃圾图像分类方法,该方法保留了 Inception-v3 模型出色的特征提取能力,可以提供有效的计算机辅助检测,在图像数据不足时具有较高的识别精度。文献[34]在“华为云人工智能大赛·垃圾分类挑战杯”上,发挥深度迁移学习的优势,构建一种端面向细粒度垃圾图像分类的迁移模型 GANet。实验结果表明,该方法能较好地应对多分辨率、多尺度、多粒度垃圾图像的特征提取,有效缓解过拟合,提高收敛速度,减少训练参数。

文献[35]提出基于 Inception-v3 网络特征提取模型和迁移学习相结合的垃圾分类方法,该方法使用 2 个  $1 \times n$  和  $n \times 1$  的卷积核取代  $n \times n$  的卷积核以减小计算量,RMSprop(Root Mean Square prop)替代 SGD(Stochastic Gradient Descent)对模型进行优化以提取更多特征,利用 LSR(Label Smoothing Regularization)方法降低过拟



合。文献[36]提出一种扩展特定网络层分支的新方法来拓宽Xception<sup>[37]</sup>网络结构,将网络底层的信息映射到网络高层,传输梯度可以无损,从而解决了反向传播的负面效应。实验结果表明,该方法对遮挡环境下的垃圾分类效果较好。

文献[38]在Inception-v3模型的基础上,使用迁移学习来训练一个专门识别垃圾类别的模型,通过对算法的研究和模型的修改,在实际效果下的识别率较高。文献[39]开源一个基于TACO数据集的数据集AquaTrash,在此数据集上应用了基于深度学习的对象检测模型AquaVision,该模型可以检测和分类漂浮在海洋和海岸上的不同污染物和有害废物,平均精度达到81.48%。

文献[40]设计一个新的增量学习框架

GarbageNet,以解决垃圾分类面临的缺乏足够数据、高成本的类别增量和噪声标签等问题,使用增量学习方法使模型不断从新样本中学习和更新,通过AFM(Attentive Feature Mixup)消除噪声标签的影响,在华为云垃圾数据集上以可接受的推理速度达到了96.96%的最佳性能。

表1对以上基于深度学习的垃圾分类方法从所使用的基础网络、数据集、是否进行数据增强、使用场景、测试精度、优缺点等方面进行了详细总结,其中使用场景“单”表示单目标,“多”表示多目标,背景单一指图像背景是固定的,背景复杂则指图像的背景处于不同环境下,(1)表示引用Trash垃圾数据集<sup>[42]</sup>, (2)表示引用华为云人工智能大赛,见<https://competition.huaweicloud.com/information/1000007620/introduction?track=107>。

表1 基于深度学习的垃圾分类方法对比

Table 1 Comparison of garbage classification methods based on deep learning

方法	基础网络	数据增强	使用场景	测试精度/%		优点	缺点
				Trash <sup>(1)</sup> /华为云 <sup>(2)</sup>	自建数据集		
文献[5]方法	ResNet50+SVM	是	单/背景单一	87.00 <sup>(1)</sup>	—	通过设置include_top=False移除了顶部分类层,特征被直接传递给多类SVM模型,减少了网络模型参数,分类过程更快	SVM对于小样本数据具有较强的分类效果,故该方法对华为云等类别多的垃圾数据集性能欠佳,不具有通用性
文献[7]方法	ResNet+Inception-v4	是	单/背景单一	—	94.38	该网络构建的IR-Block使得模型在多尺度上获得更多的特征	DSCR网络改进了原始模型,但它也带来了两个问题,即预测速度慢和大量的可训练参数,若要获得高速度和小参数数量,则必须以牺牲精度作为代价
文献[9]方法	ResNet101	是	单/背景复杂	96.73 <sup>(2)</sup>	—	该网络所构建的注意力机制能提取局部和全局的特征,特征融合机制能将不同层级、尺寸的特征进行融合	该网络在网络复杂度方面未进行改进,增加的注意力机制和特征融合机制在提升性能的同时会增加网络复杂度与计算量,故而不适用于实时性垃圾分类
文献[10]方法	ResNet+Faster RCNN	否	单&多/背景复杂	—	89.00	克服了区域误检测问题,适用于多分辨率和背景复杂情况下的垃圾分类	小区域目标容易被遗漏,且网络结构复杂,参数较大,检测时间长,无法进行实时性检测
文献[11]方法	ResNext	是	单/背景复杂	94.00 <sup>(1)</sup>	98.00	在两种垃圾数据集中鲁棒性和适应性较强	对网络自建的数据集中的医疗垃圾的识别效果较差,且在有机和无机垃圾相似时容易混淆
文献[12]方法	ResNet-34	否	单/背景单一	—	89.96	有效的特征融合策略对小目标的检测结果准确	网络的参数量和计算量大,无法适应于计算能力差的嵌入式设备
文献[13]方法	ResNet-50	否	单/背景单一	—	86.34	轻量级网络,参数量少,计算时间短,对计算机性能要求低	分类精度不够理想,且提出的最大平均池化对全局和局部的特征提取理论研究不够完善
文献[16]方法	ResNet50	是	单/背景单一/适合嵌入式设备	—	88.34	利用深度卷积和分组卷积减少计算量和参数,参数数量比ResNet50减少280倍;浮点运算量减少了31.7倍;准确率提高约3.1%;运行时间减少50%	随着模型结构越来越深,模型中会有大量的数据冗余,需要通过微调模型进一步优化和减少冗余,以克服计算机计算资源和内存的巨大浪费
文献[20]方法	DenseNet-121	是	单/背景单一	81.00 <sup>(1)</sup>	—	提出的密集块内跳跃连接的模式将计算参数减少1/2	需要依赖大量数据集进行训练,对小范围内数据集的分类任务难以达到期望结果

续表							
方法	基础网络	数据增强	使用场景	测试精度/%		优点	缺点
				Trash <sup>(1)</sup> /华为云 <sup>(2)</sup>	自建数据集		
文献[21]方法	DenseNet-121	是	单/背景单一	99.60 <sup>(1)</sup>	—	用遗传算法优化的全连接层对分类精度影响大	过分依赖数据增强对分类任务的影响,且图像中的阴影部分对分类任务影响巨大
文献[23]方法	YOLOv2	否	多/背景复杂且有注释	—	89.10	轻量级的 YOLO-v2 网络,具有复杂度低,参数量少的特点,在嵌入式设备上表现良好	属于有监督学习,对数据标注依赖性太大,现阶段的装修垃圾分类需要人工标注,标记数据较少时容易出现过拟合现象
文献[25]方法	YOLOv2	是	单/背景复杂	—	84.98	对浅层特征和深层特征的融合和复用方法能适应不同光照、背景、视角与分辨率的情况	对于透明类液体附着在地面上的污渍,由于透明形状不确定,特征提取困难,识别难度大,有漏检误检的情况
文献[27]方法	YOLOv3	是	单/背景单一	—	84.02	改进的算法克服了对小目标不敏感的问题,开发的系统支持图像识别和视频识别,能极大提升识别效率,满足实际生产	改进的算法虽然在识别准确率和速度上满足需求,但在位置上的精确率仍需要改进,可以参考 Faster R-CNN 来提升位置定位的准确性
文献[28]方法	YOLOv1	是	单&多/背景复杂	—	69.70	改进的算法相比于原始 YOLOv1 速度更快,鲁棒性更好,能有效防止过拟合,容易部署在便携式、实时和节能的边缘计算设备上	因为每个预测器只预测以小网格为中心的物体的一个边界框,所以对彼此靠近的边缘不够敏感,可考虑通过机械臂将成堆的垃圾分离开来
文献[29]方法	LSSD	否	单/背景单一	—	83.48	能有效结合低分辨率语义强的特征与高分辨率语义弱的特征,对小目标检测性能优良	需要手动识别类别标志,工作量大且对垃圾种类表达能力较差,与实际应用还有较大差距
文献[32]方法	Inceptionv-3	否	单/背景复杂	—	95.33	在标注数量和训练数据不足的情况下具有低成本、高精度的特点	迁移能力有限,容易出现负迁移的情况,需要牺牲训练时间或准确率来消除负迁移对目标领域的任务学习效果
文献[34]方法	EfficientNet-B5	否	单/背景复杂	95.12 <sup>(2)</sup>	—	使用的数据集种类更贴合实际生活,提出的注意力机制能实现对图像的像素级遮罩,相比于池化后直接进行分类,能有效缓解过拟合,减少参数和改善收敛速度	在不同的光照条件和终端采集的复杂条件下有局限性
文献[35]方法	Inceptionv-3	否	单/背景单一	—	99.00	适用于垃圾分类这种公开数据集少的领域,对标注较少和训练样本少的数据友好	加深了网络深度势必增加网络复杂度和错误率,检测算法对小目标的垃圾识别精度较低
文献[36]方法	Xception	是	单/背景单一	94.34 <sup>(1)</sup>	—	扩展分支拓宽的网络能充分利用特征信息,物体被遮挡时鲁棒性好,对难预测的单类的分类准确率高	以牺牲计算成本获取提高网络性能,难以实现实时垃圾分类,更难转移到手机等移动设备上
文献[38]方法	Inceptionv-3	是	单/背景复杂	—	93.20	在 Inception-v3 上修改的全连接层对新的分类,即垃圾分类的适应性好,且得到的高验证精度基本满足常见生活垃圾的分类识别任务	由于过期药物的塑料瓶和装有药丸的纸盒跟其他垃圾和可回收垃圾很像,因此不能很好地识别和分类过期药物
文献[39]方法	RetinaNet	否	多/背景复杂且有注释	—	81.48	所提出的模型是为非水中图像训练的,在水中图像上也仍然能够检测和分类垃圾对象。模型是在非常少的图像上训练的,应用于随机图像时也展现出了很好的结果	由于数据集需要手动标注才能保证正确性,因此对数据集的要求高,尤其对于类别多的数据集更具有挑战性
文献[40]方法	ResNext101	是	单/背景复杂	96.69 <sup>(2)</sup>	—	该网络使用弱监督迁移学习保证了特征提取器的能力,使用的增量学习对未训练过的新类别垃圾识别能力较强,适用于当前垃圾数据集标签少或标签混乱的现状	该网络是针对单目标的识别和分类,并未对多目标或相互遮挡的目标进行识别和验证

2 垃圾分类数据集

由于开源的垃圾数据集较少,研究人员一般采用自行拍摄和互联网查找的方式建立数据集。

TrashNet数据集<sup>[41]</sup>由斯坦福大学的YANG等创建,通过将物体放置在白色海报板上并使用自然光或室内照明来采集图像,所有图像均调整为512×384的空间分辨率。该数据集包含6类垃圾的RGB图像,其中每个图像仅包含一种类型的垃圾,分别是:玻璃、纸张、纸板、塑料、金属和普通垃圾。当前,该数据集由2 527张图像组成,每类包含的图像数量具体为:玻璃501张,纸张594张,纸板403张,塑料482张,金属410张和普通垃圾137张。

GINI<sup>[42]</sup>数据集包含2 561张垃圾图像,其中956张图像通过在互联网上查询“路边垃圾”、“市场垃圾”等垃圾相关的关键词获得。该数据集的每个图像还标注了其严重程度和可生物降解性的水平。

2020年深圳举办“华为云人工智能大赛-垃圾分类

挑战杯”数据应用创新大赛发布了一组生活垃圾图像数据集,该数据集有4个大类和44个小类,总计14 964张图像,数据集的数据结构是标准的VOC格式,见[https://modelarts-competitions.obs.cn-north-1.myhuaweicloud.com/garbage\\_classify/dataset/garbage\\_classify.zip](https://modelarts-competitions.obs.cn-north-1.myhuaweicloud.com/garbage_classify/dataset/garbage_classify.zip)。

TACO<sup>[43]</sup>是一个用于垃圾分类和检测的数据集,包括1 500张图像和4 784个标注,尽管该数据集非常小,但是在此数据集上可以进行垃圾分类和垃圾的边缘检测。

AquaTrash<sup>[39]</sup>数据集通过对TrashNet数据集和TACO数据集进行比较研究后,发现两个数据集都有一些缺点,例如TrashNet数据集没有注释,TACO数据集的注释不可靠等,因此提出AquaTrash数据集,该数据集包含来自与各种垃圾物品相关的4个不同类别的369张图像。其中所有图像都经过手动注释以获得结果的准确性。

垃圾分类公开数据集如表2所示。

表2 现有垃圾分类公开数据集  
Table 2 Existing garbage classification public dataset

数据集	类别	类别数	标注模式	标注数	背景	发布年份	处理难度
AquaTrash <sup>[39]</sup>	4	369	手动注释	—	随机	2020	大
Trash <sup>[41]</sup>	6	2 527	Labeled Data	—	白板	2016	小
GINI <sup>[42]</sup>	—	2 561	—	1 494	随机	2016	大
华为云	4大类44小类	14 964	Labeled Data	—	随机	2019	较大
TACO <sup>[43]</sup>	28	1 500	Crowd-source	4 784	随机	2020	大

3 未来研究展望

深度学习以精度高、速度快、稳定性强、特征可重用性高等优势在垃圾图像分类任务领域中表现优异,但深度学习在垃圾图像分类应用中还存在着一些无法规避的问题亟待解决:1)可用数据集少,对数据依赖性高;2)现有的模型训练与预测耗费大量时间和成本,需设计轻量级网络;3)现有方法关注单目标和少目标数据且背景简单,难以适应现实生活中垃圾种类多、数量大、背景复杂的特点,需设计复杂情况下鲁棒性好的多目标检测网络。

3.1 数据集依赖问题

目前,垃圾分类领域公开的数据集较少,构建数量庞大、内容丰富的垃圾数据集需要耗费大量的资源和成本。而主动学习(Active Learning)<sup>[44]</sup>、零样本学习<sup>[45]</sup>、迁移学习<sup>[46]</sup>等方法放宽了对数据集的要求和依赖性。将深度学习与主动学习、零样本学习和迁移学习相结

合很可能成为消除或减少数据依赖的突破口。

主动学习能通过自适应查询策略选择大量未标注的数据交由专门人员进行标记,用尽可能少的高质量标注数据训练高性能的模型,主要有基于预设计策略和基于学习损失的主动学习方法,但目前的多数主动学习方法主要是针对特定任务而预设计采样策略,如何在垃圾图像分类领域设计采样策略是主动学习应用的难点。

零样本学习框架如图2所示,零样本学习在图像分类中主要包括基于空间嵌入和基于生成模型两种方法,能将可见类学习到的属性迁移到不可见类上,建立可见类与不可见类的耦合关系,实现在没有学习不可见类标签样本的前提下对不可见类的分类,对垃圾分类模型的长久可用性具有重要意义,如何消除零样本学习在垃圾分类领域中出现的领域漂移、枢纽点和语义鸿沟等问题有着巨大的研究前景。



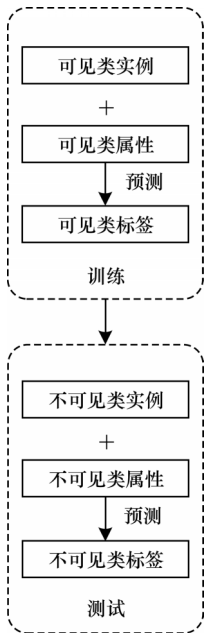


图 2 零样本学习框架  
Fig.2 Framework of zero-shot learning

深度学习与迁移学习结合的方法已经有了一些成果<sup>[47]</sup>,可以通过迁移对所传输的子网络进行微调以适应新的数据,再进行深度特征的挖掘,能有效解决垃圾图像小规模数据集的问题。随着深度学习在各个领域

的广泛应用,大量的深度迁移学习方法被提出,利用深度学习在无监督或半监督学习中进行知识的迁移将会受到越来越多的关注。预计未来的研究将集中在使用很少的标注数据以获得更好的结果。

3.2 轻量级网络模型

尽管目前提出的轻量级网络方法很多,但这些模型通常只针对通用物体检测,当前较少有专注于垃圾检测和分类的轻量级模型,深度学习在垃圾分类领域依旧面临运算复杂、时效性差、成本高等问题。

目前主流的网络轻量化方法主要有:

- 1)首先通过逐点卷积 (Pointwise Convolution, PW)对数据进行降维,然后进行常规卷积核的卷积,最后根据 PW 对数据进行升维;
- 2)多尺度卷积和多个小卷积代替大卷积;
- 3)权值量化;
- 4)使用 depth-wise 卷积等操作,起到减少网络计算量的作用。

因此,针对垃圾图像特性设计提出轻量级模型,降低运算成本和复杂度,使其能够在实时垃圾检测和分类领域上带来新的突破。主流轻量化方法的优缺点对比如表 3 所示。

表 3 主流轻量化方法的优缺点对比

Table 3 Comparison of advantages and disadvantages of mainstream lightweight methods

轻量化方法	优点	缺点
升维-常规操作-降维	两个 1×1 卷积有效地减少了卷积的参数个数和计算量	更深的网络需要更大显存的 GPU 来训练
多尺寸和多个小卷积代替大卷积	在连通性不变的情况下大幅降低了参数个数和计算复杂度,加深网络的同时减少网络参数	对于特别稀疏的数据无法表示其特征
权值量化	模型变小,运行速度变快	效果不太明显
使用 depth-wise 卷积	比标准的卷积操作减少不少参数,同时能得到更好的效果	需要的 I/O 读取次数是普通卷积的 100 倍,对 GPU 的依赖大,且会发生信息流通不畅的情况

3.3 多目标网络模型

多目标检测和多目标跟踪在交通监控、自动驾驶以及虚拟现实等领域已有广泛应用<sup>[48]</sup>。为适应我国垃圾复杂度高、数量大的特点,多目标垃圾检测和分类在垃圾焚烧发电厂等垃圾处理的末端有很大的应用前景,可采用在算法中增加更精细的特征提取层和大尺度特征层上的检测输出,改进现有的深度学习模型。

值得注意的是如纽扣电池等小体积垃圾的正确分类非常重要,此类小目标垃圾像素少,图像模糊,难以提取有效的特征,一直是目标检测领域中的热点和难点,常用的方法是用不同分辨率的分类器确定小目标物体的关键点,利用空洞卷积、特征金字塔<sup>[49]</sup>(FPN)、极端尺寸检测<sup>[50]</sup>(SNIP)等进行多尺度

方法处理和利用生成对抗网络、提高小物体检测率将是更具挑战和发展前景的研究方向。

4 结束语

本文总结垃圾分类领域中深度学习应用的研究现状,从算法特点及模型方面介绍经典的垃圾分类方法并进行分类、评估、对比不同方法和模型在数据集上的分类效果。通过分析深度学习在垃圾分类应用中的有效性,指出基于深度学习的垃圾分类方法是未来发展的重要方向,但面对准确度高、实时性强的应用需求,垃圾分类的发展仍面临较多的挑战。因此,研究人员应针对减少对标注数据集的依赖、轻量级网络模型设计和适应多目标网络模型进行研究,提升深度学习技术在垃圾分类领域内的应用能力。

## 参考文献

- [1] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [2] DENG L, YU D. Deep learning: methods and applications[J]. Foundations and Trends in Signal Processing, 2014, 7(3/4): 197-387.
- [3] JAISAL E K. Data-mining and analytics: rising concerns over privacy and people's security[C]//Proceedings of the 42nd Indian Social Science Congress Conference. New Delhi, Indian: [s. n.], 2018: 346-455.
- [4] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2016: 770-778.
- [5] ADEDEJI O, WANG Z. Intelligent waste classification system using deep learning convolutional neural network[J]. Procedia Manufacturing, 2019, 35: 607-612.
- [6] PLATT J. Sequential minimal optimization: a fast algorithm for training support vector machines[EB/OL]. [2020-03-01]. <https://www.researchgate.net/publication/2624239>.
- [7] SONG F, ZHANG Y, ZHANG J. Optimization of CNN-based garbage classification model[C]//Proceedings of the 4th IEEE International Conference on Computer Science and Application Engineering. Washington D. C., USA: IEEE Press, 2020: 1-5.
- [8] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4, inception-ResNet and the impact of residual connections on learning[EB/OL]. [2020-03-01]. <https://arxiv.org/pdf/1602.07261.pdf>.
- [9] 董子源, 韩卫光. 基于卷积神经网络的垃圾图像分类算法[J]. 计算机系统应用, 2020, 29(8): 199-204.
- [10] DONG Z Y, HAN W G. Garbage image classification algorithm based on convolutional neural network[J]. Computer Systems & Applications, 2020, 29(8): 199-204. (in Chinese)
- [11] WANG Y, ZHANG X. Autonomous garbage detection for intelligent urban management[C]//Proceedings of the 2nd International Conference on Electronic Information Technology and Computer Engineering. Shanghai: EDP Sciences, 2018: 140-147.
- [12] VO A H, SON L H, VO M T, et al. A novel framework for trash classification using deep transfer learning[J]. IEEE Access, 2019, 7: 178631-178639.
- [13] KANG Z, YANG J, LI G, et al. An automatic garbage classification system based on deep learning[J]. IEEE Access, 2020, 8: 140019-140029.
- [14] 袁建野, 南新元, 蔡鑫, 等. 基于轻量级残差网络的垃圾图片分类方法[J]. 环境工程, 2021, 39(2): 110-115.
- [15] YUAN J Y, NAN X Y, CAI X, et al. Junk image classification method based on lightweight residual network[J]. Environmental Engineering, 2021, 39(2): 110-115. (in Chinese)
- [16] 梁斌, 刘全, 徐进. 基于多注意力卷积神经网络的特定目标情感分析[J]. 计算机研究与发展, 2017, 54(8): 1724-173.
- [17] LIANG B, LIU Q, XU J. Target-specific sentiment analysis based on multi-attention convolutional neural network[J]. Computer Research and Development, 2017, 54(8): 1724-173. (in Chinese)
- [18] ZHU X, CHENG D, ZHANG Z, et al. An empirical study of spatial attention mechanisms in deep networks[C]//Proceedings of IEEE International Conference on Computer Vision. Washington D. C., USA: IEEE Press, 2019: 6688-6697.
- [19] GAN Y Q, GE B. Research on lightweight convolutional neural network in garbage classification[C]//Proceedings of IOP Conference on Earth and Environmental Science. [S. l.]: IOP Publishing, 2021: 320-231.
- [20] HUANG G, LIU Z, VAN DER M L, et al. Densely connected convolutional networks[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2017: 4700-4708.
- [21] SUSANTH G S, LIVINGSTON L M J, LIVINGSTON L G. Garbage waste segregation using deep learning techniques[C]//Proceedings of IOP Conference on Materials Science and Engineering. [S. l.]: IOP Publishing, 2021: 204-213.
- [22] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[EB/OL]. [2020-03-01]. <https://web.cs.ucdavis.edu/~yjee/teaching/ecs289g-winter2018/alexnet.pdf>.
- [23] BIRCANOGLU C, ATAY M, BESER F, et al. Recyclenet: intelligent waste sorting using deep neural networks[C]//Proceedings of 2018 IEEE Innovations in Intelligent Systems and Applications. Washington D. C., USA: IEEE Press, 2018: 1-7.
- [24] MAO W L, CHEN W C, WANG C T, et al. Recycling waste classification using optimized convolutional neural network[J]. Resources, Conservation and Recycling, 2021, 164(1): 105-112.
- [25] 刘俊明, 孟卫华. 基于深度学习的单阶段目标检测算法研究综述[J]. 航空兵器, 2020, 27(3): 44-53.
- [26] LIU J M, MENG W H. A survey of research on single-stage target detection algorithms based on deep Learning[J]. Aero Weaponry, 2020, 27(3): 44-53. (in Chinese)
- [27] LIU Y, GE Z, LÜ G, et al. Research on automatic garbage detection system based on deep learning and narrowband internet of things[J]. Journal of Physics, 2018, 1069: 120-132.
- [28] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: efficient convolutional neural networks for mobile vision applications[EB/OL]. [2020-03-01]. <https://arxiv.preprint.org/abs/1704.04861>.
- [29] 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]. 中国图象图形学报, 2019, 24(8): 1358-1368.
- [30] NING K, ZHANG D B, YIN F, et al. Garbage detection and



- classification of intelligent sweeping robot based on visual perception[J]. Chinese Journal of Image and Graphics, 2019, 24(8): 1358-1368. (in Chinese)
- [26] KUMAR S, YADAV D, GUPTA H, et al. A novel YOLOv3 algorithm-based deep learning approach for waste segregation: towards smart waste management[J]. Electronics, 2021, 10(1): 14.
- [27] 吴子沛. 面向垃圾分类的瓶罐识别分类系统[D]. 邯郸: 河北工程大学, 2020.
- WU Z P. Recognition and classification system of bottles and cans for garbage classification[D]. Handan: Hebei University of Engineering, 2020. (in Chinese)
- [28] YE A, PANG B, JIN Y, et al. A YOLO-based neural network with VAE for intelligent garbage detection and classification [C]//Proceedings of the 3rd IEEE International Conference on Algorithms, Computing and Artificial Intelligence. Washington D. C., USA: IEEE Press, 2020: 1-7.
- [29] MA W, WANG X, YU J. A lightweight feature fusion single shot multibox detector for garbage detection[J]. IEEE Access, 2020, 8: 188577-188586.
- [30] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//Proceedings of European Conference on Computer Vision. Berlin, Germany: Springer, 2016: 21-37.
- [31] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of 2017 IEEE International Conference on Computer Vision. Washington D. C., USA: IEEE Press, 2017: 2999-3007.
- [32] FENG J, TANG X. Office garbage intelligent classification based on inception-v3 transfer learning model[C]//Proceedings of the 4th IEEE International Conference on Control Engineering and Artificial Intelligence. Washington D. C., USA: IEEE Press, 2020: 8-15.
- [33] SZZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2016: 2818-2826.
- [34] 高明, 陈玉涵, 张泽慧, 等. 基于新型空间注意力机制和迁移学习的垃圾图像分类算法[J]. 系统工程理论与实践, 2021, 41(2): 498-512.
- GAO M, CHEN Y H, ZHANG Z H, et al. Garbage image classification algorithm based on new spatial attention mechanism and transfer learning[J]. System Engineering Theory and Practice, 2021, 41(2): 498-512. (in Chinese)
- [35] 康庄, 杨杰, 郭濠奇. 基于机器视觉的垃圾自动分类系统设计[J]. 浙江大学学报(工学版), 2020, 54(7): 1272-1280.
- KANG Z, YANG J, GUO H Q. Design of automatic garbage classification system based on machine vision[J]. Journal of Zhejiang University (Engineering Science Edition), 2020, 54(7): 1272-1280. (in Chinese)
- [36] SHI C, XIA R, WANG L. A Novel multi-branch channel expansion network for garbage image classification[J]. IEEE Access, 2020, 8: 154436-154452.
- [37] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2017: 1251-1258.
- [38] CAO L, XIANG W. Application of convolutional neural network based on transfer learning for garbage classification[C]//Proceedings of the 5th IEEE Information Technology and Mechatronics Engineering Conference. Washington D. C., USA: IEEE Press, 2020: 1032-1036.
- [39] PANWAR H, GUPTA P K, SIDDIQUI M K, et al. Aqua vision: automating the detection of waste in water bodies using deep transfer learning[J]. Case Studies in Chemical and Environmental Engineering, 2020: 100-116.
- [40] YANG J, ZENG Z, WANG K, et al. GarbageNet: a unified learning framework for robust garbage classification[J]. IEEE Transactions on Artificial Intelligence, 2021, 58(1): 3214-3226.
- [41] YANG M, THUNG G. Classification of trash for recyclability status[EB/OL]. [2020-03-01]. <http://cs229.stanford.edu/proj2016/report>.
- [42] MITTAL G, YAGNIK K B, GARG M, et al. Spotgarbage: smartphone app to detect garbage using deep learning[C]//Proceedings of 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. New York, USA: ACM Press, 2016: 940-945.
- [43] PROENCA P F, SIMOES P. TACO: trash annotations in context for litter detection[EB/OL]. [2020-03-01]. <https://arxiv.org/abs/2003.06975>.
- [44] YOO D, KWEON I S. Learning loss for active learning[C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2019: 93-102.
- [45] CHANGPINYO S, CHAO W L, GONG B, et al. Synthesized classifiers for zero-shot learning[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2016: 5327-5336.
- [46] PAN S J, TSANG I W, KWOK J T, et al. Domain adaptation via transfer component analysis[J]. IEEE Transactions on Neural Networks, 2011, 22(2): 199-210.
- [47] TAN C, SUN F, KONG T, et al. A survey on deep transfer learning[C]//Proceedings of International Conference on Artificial Neural Networks. Berlin, Germany: Springer, 2018: 270-279.
- [48] 李志华, 于杨. 基于检测的多目标跟踪算法综述[J]. 物联网技术, 2021, 11(4): 20-24.
- LI Z H, YU Y. Survey of detection-based multi-target tracking algorithms[J]. Internet of Things Technology, 2021, 11(4): 20-24. (in Chinese)
- [49] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2017: 2117-2125.
- [50] SINGH B, DAVIS L S. An analysis of scale invariance in object detection snip[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2018: 3578-3587.