

# 基于图像识别技术的垃圾智能分类

刘晓波,宋成明,曹凯源

山西工程科技职业大学,山西 太原 030619

**摘要** 现阶段,如何有效处理日渐增多的生产和生活垃圾已经成为我国亟需解决的问题。为了有效应对垃圾问题,一种以 YOLOv5s 网络为依托的智能垃圾分类模型逐渐受到相关研究人员的重视,并且通过一系列实验证明其在垃圾分类过程中的合理应用,能够在很大程度上提升垃圾分类处理的有效性。鉴于此,将在简要分析图像识别技术及其过程的基础上,展示基于 YOLOv5s 网络的垃圾分类和检测模型的工作过程和实验结果,以期对垃圾分类的智能化提供一定参考。

**关键词** 图像识别技术;垃圾智能分类;深度学习

**中图分类号** TP391.41

**DOI** 10.19769/j.zdhy.2022.02.021

## 0 引言

现阶段,我国大部分地区在开展垃圾分类工作时候,主要使用的是弹跳分选、液体浮选以及尺寸筛选等方法。其中,人工挑选是筛选过程中必不可少的步骤,这一环节的主要目的是将机械不能有效区分的目标物进行再次挑选,并且对机械分选的质量进行严格把控。但是,在利用人工分选垃圾时,通常存在劳动强度大、工作效率低、作业环境差等问题,亟需设计一种智能化和自动化程度相对较高的设备来代替人工开展垃圾分选工作<sup>[1]</sup>。鉴于此,一种以图像识别为基础的智能垃圾分选系统进入人们视野,这一系统的主要工作原理是在垃圾通过时,借助彩色相机技术进行拍照,然后通过图像处理算法自动识别目标物的材质、类型以及具体位置等,同时将该信息发送给后续作业机器人,机器人在接到指令之后进行针对性抓取,进而实现智能垃圾分选。由此可见,在这一系统中,图像识别技术居于核心地位,但是在实际应用此项技术的时候发现,如果垃圾表现出现破损、脏污等,普通的图像识别技术很难准确地识别垃圾类型,严重降低了智能分类的有效性。为了有效解决这一问题,一种基于 YOLOv5s 网络的垃圾分类检测方法被提出,其在垃圾识别和分类方面具有准确率高、鲁棒性好个计算速度快的优点。因此,YOLOv5s 网络在垃圾智能分类中的有效应用,能够帮助垃圾处理单位实现智能化和自动化分拣的同时,还能够保障垃圾分类的准确性,进而有效提升垃圾处理效率。

## 1 图像识别技术相关概述

### 1.1 图像识别技术

图像识别技术在计算机视觉方面占据重要地位,并且这项技术在人们生产生活中的有效应用,为人们解决了众多难题。由于我们的现实生活是复杂多样的,在我们周边无时无刻不围绕着大量的图像信息,让我们难以

从这些信息中高效获取目标信息<sup>[2]</sup>。随着人们对计算机技术应用的不断深化,在一定程度上提升了目标信息的获取效率,同时也不断衍生出新的研究方向。当前所使用的图像识别技术大多是基于深度学习算法中的神经网络框架展开的。通常情况下,图像识别系统由图像预处理、特征提取和图像预测及分类等结构构成,并且基于图像空间和特征空间是现阶段最常用的两种分类办法。在图像空间识别过程中,底层特征差异是主要的识别核心,然后借助 Hu 不变矩等进一步获取图像在颜色、形状等方面的特征信息,并在此基础上进行具体的细化分类。科技的不断进步,推动计算机视觉技术向高精尖方向发展,同时这项技术也开始在人们的生产生活中有效应用。例如,在公共安全方面,利用步态识别技术对行人进行有效检测,能够快速识别出危险分子,保障公众安全;在个人安全方面,可以利用人脸识别技术进行身份验证等。

### 1.2 图像识别技术的过程

图像识别技术过程主要包括对图像的处理和识别两大部分。首先是图像的处理过程。图像处理主要说的是借助相应的计算机软件将原始图像进行有效处理,以减少其在采集期间产生的损耗,提升图片质量,将原始图片进行数字化处理之后,转变成能够被计算机识别的二进制形式。图像的处理过程具体又可以分为图像的采集、增强、复原、编码以及压缩和分割等。其中,采集图像是提取图像信息的重要过程,主要是借助扫描仪、摄像机或者照相机等,获取数字化图像,这同样是图像识别工作的首要步骤。图像增强主要是解决图像在成像、传输、采集或者复制的时候容易发生损耗,进而降低图像质量,导致图像的数字化效果与人们预期之间存在差距的问题,提升图片的清晰度,一般情况下会借助图像增强技术,使图像中事物的轮廓更加清晰,为后期分析图像特征奠定基础。图像的恢复:在采集图像的过程中,可能受到环境因素的影响,导致呈现出来的图像不清晰,因此需要对图像进行恢复操作,以有效提升其清晰度。现阶段,常用的复

收稿日期:2022-02-15

基金项目:2020 年产学研合作协同育人项目,软件工程专业“双师双能型”队伍建设,项目编号:202002165026。

作者简介:刘晓波(1995—),男,汉族,山西太原人,硕士研究生,助教,研究方向软件工程、人工智能、深度学习。

原图像技术是滤波,其能够将不清晰的图像进行有效复原。同时,图像重建也属于一种图像恢复技术,其主要是借助物体横剖面的投影数据来建立新图像。开展编码和压缩操作主要是由于数字化图像的数据量较大,占据的存储空间也相对较大。为了能够提升图像传输的便捷性,需要对图像进行编码或者压缩。<sup>[3]</sup>通过这种操作,不仅能够压缩图像数据量,节省存储空间,还能够有效提升传送速度,减少传送时间。图像的分割主要是将同一张图像分割成多个不同区域,并且这些区域之间不能交互重叠,同时又要具备各自的区域特征,使各个区域共同构成图片像素的连续集,这也是后期识别、分析和处理图像的基础。

其次是图像的识别过程,这一过程主要是选择和抽取已处理图像的特征,并在此基础上进行合理分类。图像识别过程主要包含图像特征的抽取和选择、分类器设计和分类决策 2 个阶段。其中,特征的抽取和选择主要说的是在图像的识别模式中,将图像的所有特征进行分析和抽取,并将其选择出来,为以后的图像信息的分类和汇总提供前提和基础。特征抽取是识别图像的基础,因此其特征抽取正确与否,在很大程度上决定了图像分类是否有效。分类器设计也就是制定图像数据的分类细则,并在相应的细则范围内对图像特征进行有效的分类,以提升识别效果<sup>[4]</sup>。最后,需要将已经提取的特征信息与分类细则进行对比,做好分类决策。同时,还需要有效评估分类结果的准确率以及损失函数。

## 2 基于 YOLOv5s 网络的垃圾分类和检测模型

### 2.1 总体设计方案

此次试验主要涉及的是基于 YOLOv5s 网络的垃圾自动识别和分拣系统。首先需要收集垃圾图像,并对每张图像进行预处理,包括图像大小、归一化和图像增强等。如果收集到的图片数量较少,就需要利用图像增广技术(如反转、拆剪、模糊、旋转、亮度对比度调节等)进行数据扩充,并根据其图像内容和类别为每张图片制作标签集。然后利用数据集训练 YOLOv5 的网络模型,并将训练好的模型部署到开发板上,安置在分拣机械臂的单片机内。

在垃圾分类现场,图像采集设备实时采集垃圾图片,并将这些图片借助无线形式传送到开发板。接着,开发板使用训练好的 YOLOv5 垃圾识别和分类模型来检测图片中所含垃圾的具体类别,并将结果发送给单片机,然后由控制分拣装置,根据模型识别出来的具体垃圾类别,将其投放到与之相对应的区域<sup>[5]</sup>。

### 2.2 YOLOv5 网络模型

YOLOv5 是基于 CNN(卷积神经网络)的图像识别模型框架,其中包括 4 种已经通过 COCO 数据集预训练的模型。如图 1 所示,COCO AP val 表示 mAP(各平均精度的平均值),在拥有 5000 张图像 of COCO val2017 数据集上,在 256 到 1536 的不同推理大小上测量指标。GPU 速度衡量了在 COCO val2017 数据集上使用 AWS

p3.2xlarge V100 实例,在批次大小为 32 时,每张图像的平均推理时间,效率数据来自 google/automl。由此可以看出,YOLOv5s 在准确度和 GPU 处理速度方面的表现都较为理想,因此选用 YOLOv5s 作为本文使用的模型。

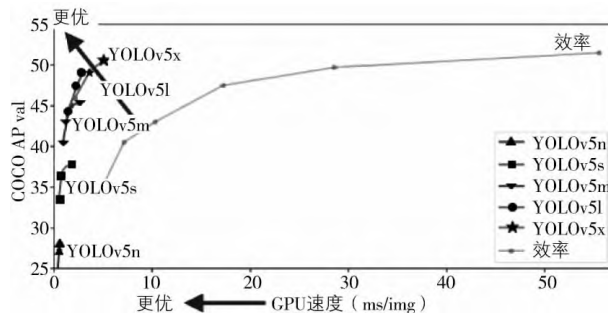


图 1 YOLOv5 各模型性能对比图

YOLOv5s 模型包括 Input、Backbone、Neck、Output 四部分(如图 2 所示)。其中,输入端能够进行自适应,并自主缩放图片,并且主要采用 Mosaic 数据增强模式,自动计算数据集的最佳锚框值。Backbone 主要包括 Focus 结构和 CSPNetCross Stage Partial Networks,能够跨阶段局部融合网络。在 Focus 的结构中主要通过 4 次切片和一次 32 个卷积核操作来将原始图片由  $608 \times 608 \times 3$  转变成  $304 \times 304 \times 32$  的特征图。CSPNet 则仿照 Densenet 密集跨层跳层连接思路,实现局部跨层的融合。在 Neck 中主要包含 2 个模块,分别是 PANet、SPP。PANet 主要是通过从上往下的方式将模型高层的特征信息与不同层次的 CSP 模块输出特征信息进行有效融合,然后再通过从下往上的传播方式实现集合结构的聚合浅层特征,从而实现不同层图像特征的有效融合。空间金字塔池化模块主要是借助大小不同的 4 种核开展最大池化操作,并在此基础上进行张量拼接。输出层主要是利用 GIOU\_Loss 代替 IOU\_Loss 成为损失函数,以增加相交尺度衡量,有效解决 IOU\_Loss 不能优化 2 个不相交框的问题。

### 2.3 实验设置和结果分析

一是介绍数据集。此次实验所使用的数据集属于自制,并且数据集在收集时充分考虑了角度、遮挡和光照等因素,并综合考量了图像内容拥挤或者距离较远等特殊状况。在实验开展过程中主要利用数据增广技术,将样本数量提升至原来的数倍,利用 labeling 工具标注图像中的垃圾,并按照 8:2 的比例将数据集具体划分为训练集和测试集 2 部分。其中数据集共有图像 10 000 张,每张图片的分辨率为  $960 \times 640$ 。

二是训练 YOLOv5s 网络。此次实验的具体运行环境如下:GPU 是 Tesla K80 GPU,显卡内存是 12.69 G,操作系统是 ubuntu 18.04.3,安装的文件库是 CUDA 10.1,开发语言是 Python3,框架是 Tensorflow。YOLOv5s 模型的参数训练采用 SGD(随即梯度下降)优化算法,参数设置如下:Batch 大小为 32;最大迭代次数为 100;动量因子为 0.8;权重衰减系数为 0.000 1。在实验过程中,使用余弦退火策略动态调整学习率,并且初始学习率为 0.01。

三是分析测试结果。如图 3 所示,训练好的 YOLOv5s 模型在使用训练集时的表现良好,准确率趋于 98%,且损

YOLOv5模型总览

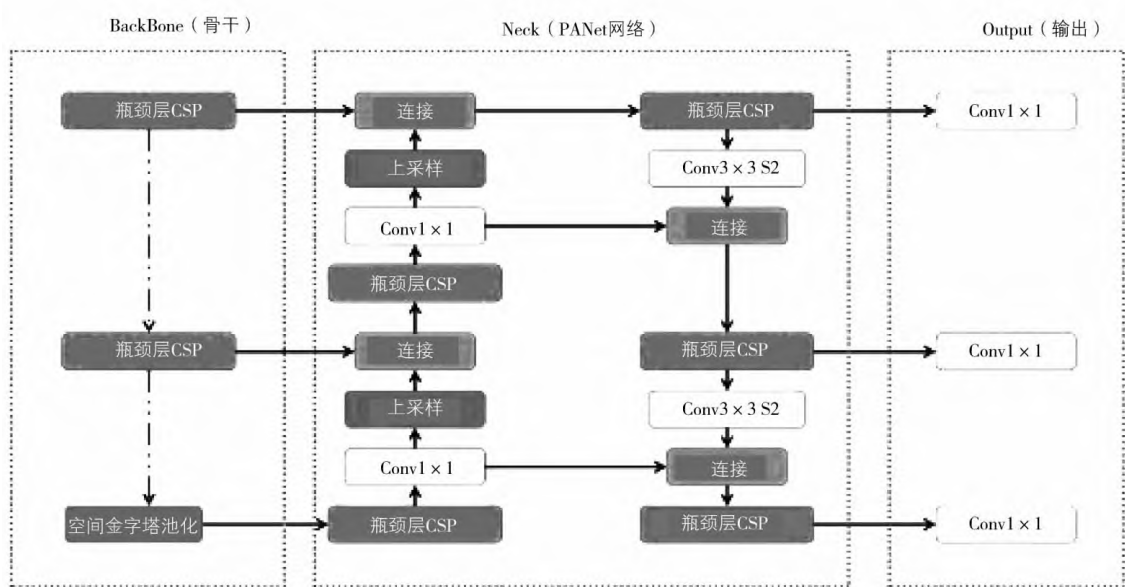


图2 YOLOv5 模型总览

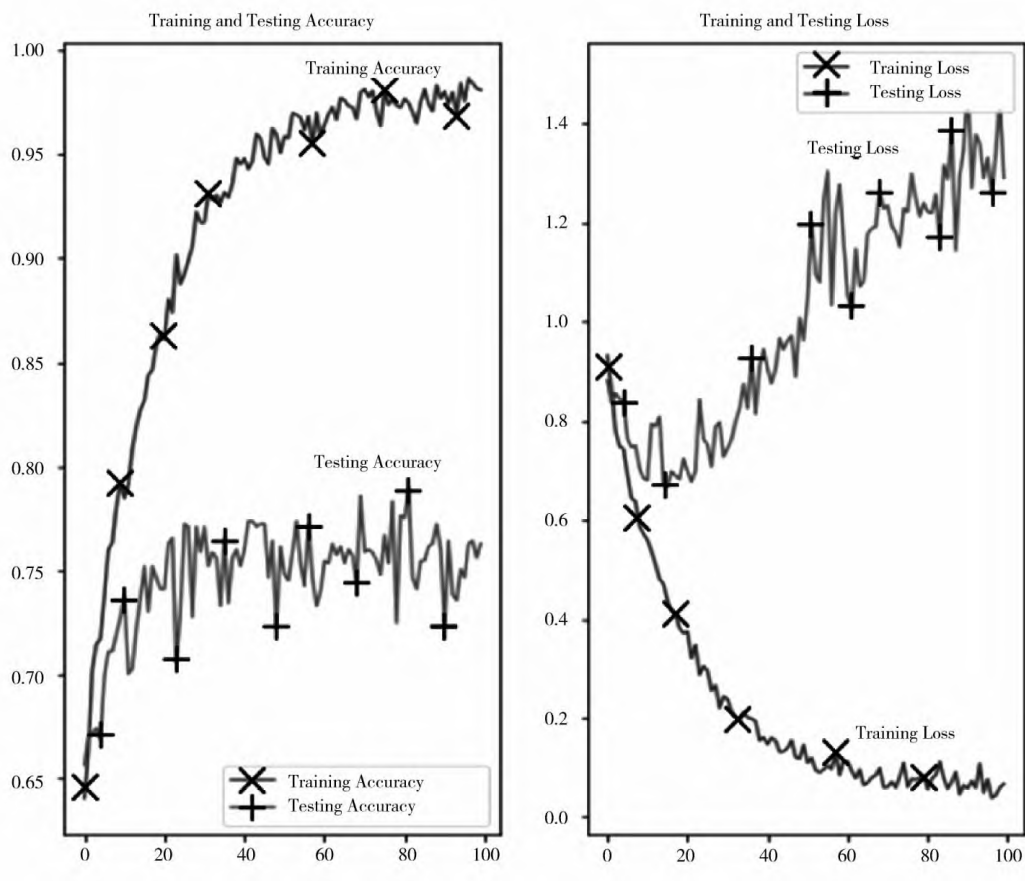


图3 测试结果

失函数数值收敛于 0.1 上下。但是,在使用测试集时其表现明显比训练集差,准确率是 75% 左右,损失函数数值虽然超过了 1.2,但是有向更小数值收敛的趋势。整体而

言,训练好的 YOLOv5s 模型具有较好的垃圾图像分类和处理能力,但仍存在很大的提升空间。

(下转第 84 页)

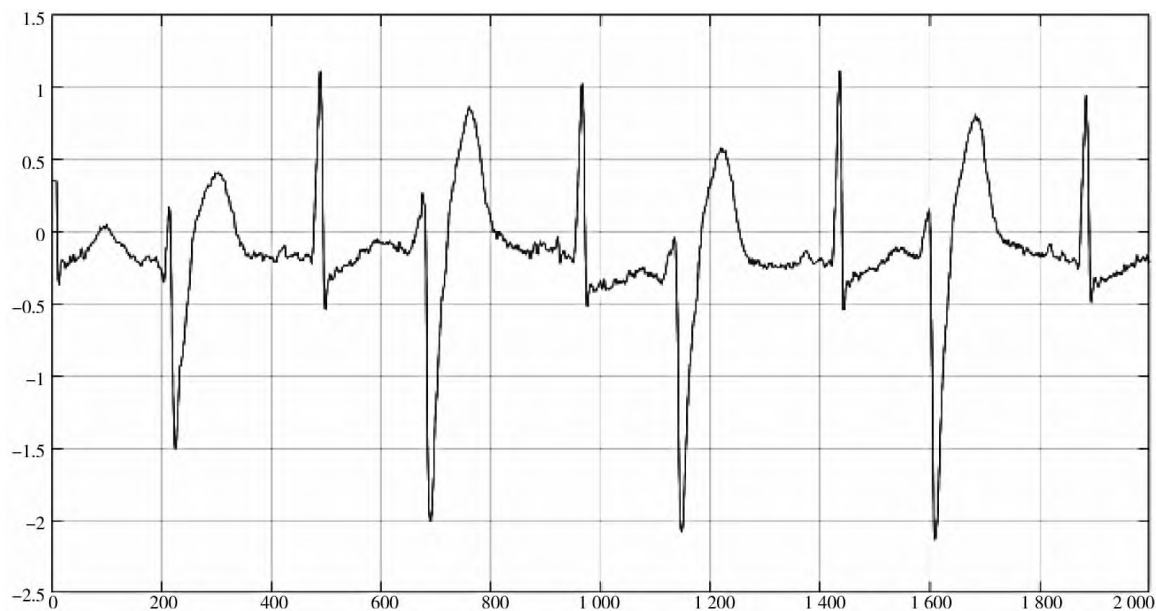


图8 第200组MLII导联M文件解码波形

## 5 结语

MIT-BIH 数据库中的动态心电图能提供其波形,但由于其对心电电压信号特有的数据编码方式不易使用,而 MATLAB 为常用的数据分析处理平台,为广大科研工作者所使用。本文提出了一种基于 MATLAB 平台的解码的 M 文件设计,实现将 MIT-BIH 数据库特殊的动态心电图数据编码转换成 MATLAB 平台能识别和使用的数据编码,为广大科研工作者进一步深入分析动态心电图数据提供了基础。

(上接第 79 页)

## 3 结语

随着科技的进步和环保理念的深入人心,人们逐渐加大了对垃圾处理方式的探究,一种基于图像识别的垃圾智能分类技术逐渐得到人们的关注,这项技术的应用,不仅在一定程度上降低了垃圾分类带来的人力资源损耗,还在一定程度上提升了垃圾分类效率。为了进一步提升垃圾智能分类的有效性,在未来的垃圾智能分类系统中可以尝试增加垃圾图像训练数据集的图像数量,或者在 GPU 性能和开发板性能允许的情况下使用 mAP 更高的 Yolov5l、Yolov5m、Yolov5x 模型,以实现垃圾智能分类系统的不断完善和优化,切实提升垃圾识别的准确

## 参考文献

- [1] 宋喜国,邓亲恺.MIT-BIH 心率失常数据库的识读及应用[J].中国医学物理学杂志,2004(4):231-232.
- [2] 宋春丽.怎样识读 MIT-BIH 中的心电信号[J].科技资讯,2010(9):27.
- [3] 黄少罗,闫聪聪.MATLAB2020 从入门到精通[M].北京:机械工业出版社,2021.
- [4] 沈再阳.MATLAB 信号处理[M].北京:机械工业出版社,2017.

率和处理效率,让智能垃圾分类技术更好地服务于人们的生产生活。

## 参考文献

- [1] 李天千,陈志鑫,黄桂鑫,等.基于计算机图像识别的垃圾智能分类[J].现代信息科技,2021,5(17):92-94,99.
- [2] 叶志祥.智能分类垃圾箱设计研究[J].中国资源综合利用,2019,37(04):191-193.
- [3] Redmon J, Farhadi A. YOLOv3: An Incremental Improvement[J]. arXiv e-prints,2018.
- [4] 任蓓蓓,倪蓉,曹瑞雪,等.“足不出户”生活垃圾智能分类建筑构造设计前景分析[J].居舍,2021(15):109-110.
- [5] 张秀芳,龚小梅,占小洋,等.垃圾分类之自动识别技术探索[J].科技创新与应用,2021(1):178-180,184.