

基于 Object Detection API 的物流单元货架目标检测

龙健宁¹, 刘 斌¹, 龚德文²

(1. 华南理工大学 聚合物成型加工工程教育部重点实验室 广东省高分子先进制造技术及装备重点实验室 聚合物新型成型装备国家工程研究中心, 广州 510641; 2. 广东昌恒智能科技有限公司, 东莞 523841)

摘要:随着人工智能的兴起,深度学习的方法已经被广泛地应用到各类图像目标的检测当中,并在复杂环境下取得了良好的效果。针对物流仓储环境,该文基于开源框架Tensorflow上的库Object Detection API,选择了Faster R-CNN算法和SSD-MobileNet算法,分别对物流单元货架上摆放的物流周转箱进行目标检测。实验结果表明,相比于Faster R-CNN算法,SSD-MobileNet算法能够同时满足实时性与准确率的要求。将训练所得的SSD-MobileNet模型移植到QT平台,设计了物流单元货架目标检测界面。

关键词:深度学习;物流单元货架;目标检测;Faster R-CNN算法;SSD-MobileNet算法
中图分类号:TP29 **文献标志码:**A **文章编号:**1001-9944(2020)09-0046-05

Object Detection of Logistics Unit Shelf Based on Object Detection API

LONG Jian-ning¹, LIU Bin¹, GONG De-wen²

(1. National Engineering Research Center of Novel Equipment for Polymer Processing, Key Laboratory of Polymer Processing Engineering of Ministry of Education, Guangdong Provincial Key Laboratory of Technique and Equipment for Macromolecular Advance Manufacturing, South China University of Technology, Guangzhou 510641, China; 2. Guangdong Changheng Intelligent Technology Co., Ltd., Dongguan 523841, China)

Abstract: With the rise of artificial intelligence, deep learning method has been widely used in all kinds of image object detection, especially good results has been achieved in the complex environment. Based on the library Object Detection API in Tensorflow, Faster R-CNN algorithm and SSD-MobileNet algorithm are selected in this paper to detect the object of the logistics turnover box on the logistics unit shelf. The experimental results show that compared with Faster R-CNN algorithm, SSD-MobileNet algorithm can meet the requirements of real-time and accuracy at the same time. The SSD-MobileNet model is transplanted to QT platform, and the object detection interface of logistics unit shelf is designed.

Key words: deep learning; logistics unit shelf; object detection; Faster R-CNN algorithm; SSD-MobileNet algorithm

目标检测是以图像识别技术为基础,对图像中的目标对象进行分类和识别,并在目标对象周围绘制预测框进行定位。目前,应用深度学习方法的检测技术已经被广泛地应用到无人驾驶、图片搜索、人脸识别、光学字符识别等方面^[1]。传统的图像检测一般采用“人工设计特征+分类器”的模式,通

过人工设计提取局部特征再借助于向量机等分类器进行图像的分类和识别。人工设计特征容易受到光照变化、位置变化以及角度变化等因素的影响^[2],而深度学习的方法能够自动提取图像特征,克服了传统图像检测的这一缺点,鲁棒性和泛化能力更强,应用也更为广泛。

收稿日期:2020-06-24;修订日期:2020-08-04

作者简介:龙健宁(1995—),男,在读硕士研究生,研究方向为机械工程;刘斌(1969—),男,博士,教授,研究方向为模具CAD/CAE/CAM和材料成型装备及数控技术。

Tensorflow 是谷歌公司发布的基于数据流图的开源的深度学习框架,具有高度的灵活性、通用性、扩展性以及良好的移植性,能够适应实际生产环境,被广泛地应用到图像识别、自然语言识别和语音识别等深度学习领域^[3]。Object Detection API 是 Tensorflow 框架下的目标检测库,集成了 SSD、Faster R-CNN 以及 Mask R-CNN 等主流的目标检测算法,易于构建、训练和部署目标检测模型。深度学习方法在复杂环境下进行目标检测效果卓越。

物流仓储环境条件复杂多变,物流周转箱是物流环境下货物流通的主要载体之一。本文基于 Object Detection API,利用相机拍照分别对单元货架上摆放的物流周转箱进行目标检测,以期为无人叉车或机械臂对物流单元货架上的物流周转箱进行智能搬运提供依据。单元货架目标检测示意图如图 1 所示。

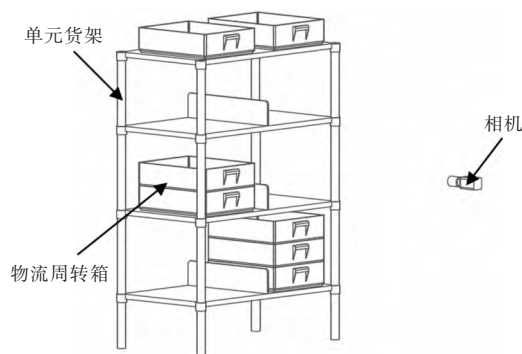


图 1 物流单元货架目标检测示意图

Fig.1 Diagram of object detection of logistics unit shelf

1 目标检测算法

目标检测算法主要分为两大类^[4]。一类是基于分类的目标检测框架(two stage),另一类是基于回归的目标检测框架(one stage)。基于分类的目标检测算法首先生成可能包含被检测目标的候选区域,再对候选区域进行进一步地分类和位置精修,以获得最终检测结果,主要算法包括 R-CNN、Fast R-CNN、Faster R-CNN、Mask R-CNN 及其改进算法等。基于回归的目标检测算法则不需要生成候选区域,直接在原图上生成最终结果,因此速度更快,但是精度有所下降,主要算法包括 YOLO、SSD 及其改进算法等。本文选用 Object Detection API 中集成的 Fast R-CNN 模型和 SSD-MobileNet 模型进行单元货架的目标检测任务。

1.1 Faster R-CNN 模型

Faster R-CNN^[5]是由 Girshick R 等人在 2016 年提出的基于分类的目标检测模型。Faster R-CNN 模型结构如图 2 所示。

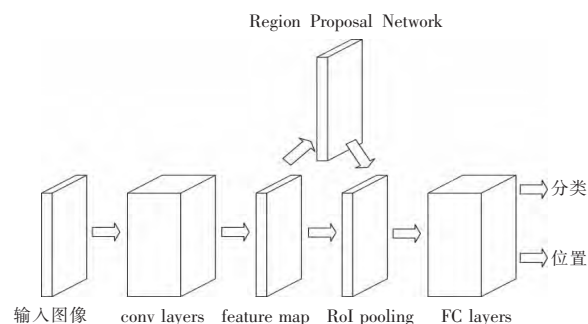


图 2 Faster R-CNN 模型结构

Fig.2 Model structure of Faster R-CNN

Faster R-CNN,首先利用卷积神经网络自动提取输入图像的特征信息,并传递给 RPN(region proposal network)生成多个不同大小和比例的候选区域;接着,ROI Pooling 利用 RPN 生成的候选区域和卷积神经网络提取到的共享特征图(feature map),得到固定大小的候选区域特征图;最后,利用 softmax 对候选区域进行具体类别的分类,利用边框回归获得物体的精确位置,输出检测分类信息和位置信息。

1.2 SSD-MobileNet 模型

SSD^[6]是由 Wei Liu 等人在 2016 年提出的基于回归的目标检测算法。SSD 以 VGG16 卷积神经网络为主体,将两个全连接层改成卷积层,再增加 4 组卷积、8 个卷积层构造网络结构,同时,借鉴了 YOLO 的回归思想和 Faster R-CNN 的 Anchor 机制,各卷层将图像分成若干相同大小的网格(feature map cell),对每个网格设定一系列默认大小长宽比的包围盒(default boxes),最后分别预测 default boxes 的偏移以及类别得分,通过非极大值抑制方法得到检测结果,输出类别和位置。

而 SSD-MobileNet 模型,则是利用轻量化的 MobileNet^[7]神经网络结构替换传统 SSD 模型当中的 VGG16 卷积神经网络结构,在 MobileNet 网络的最后卷积层 conv13 后面再增加 8 个卷积层,分别为 conv14_1, conv14_2, conv15_1, conv15_2, conv16_1, conv17_1, conv17_2, 从中抽取 conv11, conv13, conv14_2, conv15_2, conv16_2 以及 conv17_2 等 6 层进行检测。输入图像归一化为 300×300, conv11,

conv13, conv14_2, conv15_2, conv16_2 以及 conv17_2 的特征尺寸分别为 19×19 , 10×10 , 5×5 , 3×3 , 2×2 以及 1×1 。SSD-MobileNet 综合了 SSD 和 MobileNet 两者的优点,在保证精度的同时,具有计算量小、节约资源、响应迅速等特点。SSD-MobileNet 模型结构如图 3 所示。

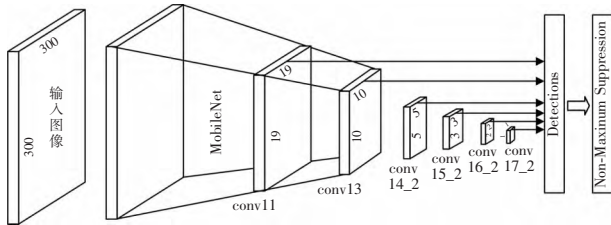


图 3 SSD-MobileNet 模型结构

Fig.3 Model structure of SSD-MobileNet

2 实验准备及目标检测过程

2.1 图像采集及增强

单元货架上物流周转箱的摆放形式如图 4 所示。由此可知,物流周转箱的侧面是检测的目标

对象。

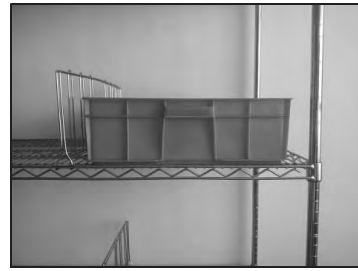
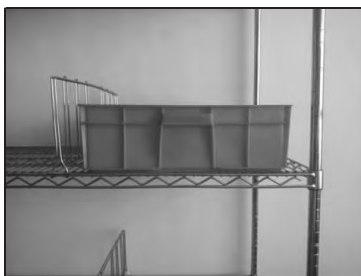


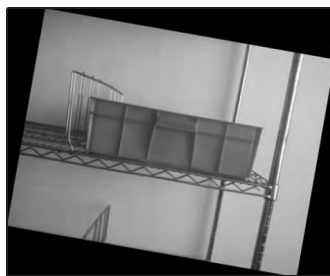
图 4 物流周转箱摆放方式

Fig.4 Style of placing the logistics turnover box

在真实的物流仓储环境下采集了 36 张图像。为了扩充数据,对原图像作顺时针旋转 10° 和逆时针旋转 10° 处理,并对原图像以及旋转图像进行椒盐噪声处理和高斯模糊处理;另外,考虑到光线变化的情况,还对原图像进行了亮度增加和亮度减少的处理。最终一幅原图像能够得到 10 幅扩充图像,如图 5 所示。数据集总共有 $36 \times 11 = 396$ 张图像。



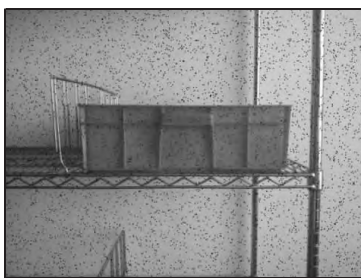
(a)原图



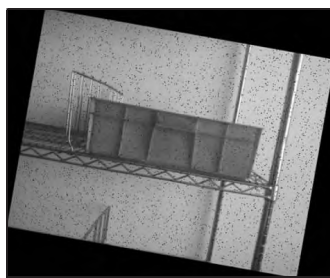
(b)顺时针旋转 10°



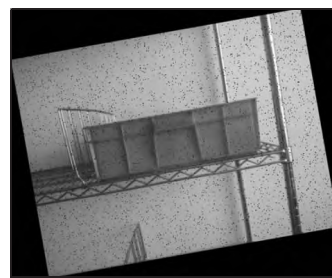
(c)逆时针旋转 10°



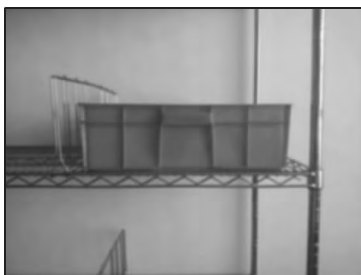
(d)椒盐噪声



(e)椒盐噪声+顺时针旋转 10°



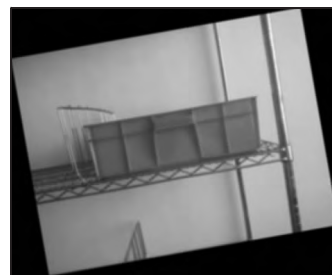
(f)椒盐噪声+逆时针旋转 10°



(g)高斯模糊



(h)高斯模糊+顺时针旋转 10°



(i)高斯模糊+逆时针旋转 10°

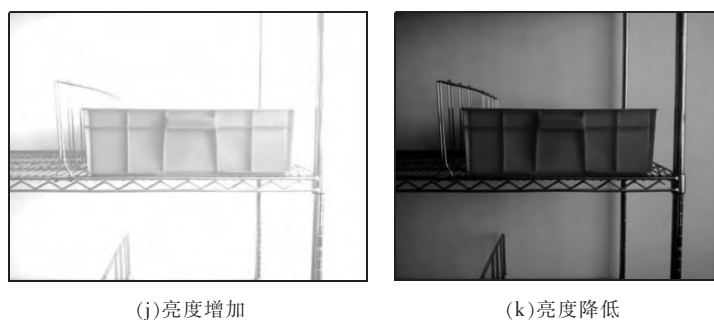


图5 图像增强效果

Fig.5 Effect of image enhancement

2.2 图像标注及转换

基于 Faster R-CNN 和 SSD-MobileNet 的目标检测,都属于监督学习,利用 labelImg 标注工具获得图像中目标对象的类别和像素位置坐标,保存成 xml 格式文件。

将获得的数据集 xml 文件划分为训练集、验证集和测试集,比例分别为 75%,15%和 10%;然后将训练集、验证集和测试集转换成 csv 格式文件,再通过 Object Detection API 下的 generate_tfrecord.py 将 csv 格式文件转换成 Tensorflow 能够调用的 tfrecord 格式文件。

2.3 模型训练

在 Object Detection API 下训练模型十分方便,以训练 SSD-Mobilenet 模型为例。为了加快训练的速度,实验中采用了预训练模型 ssd_mobilenet_v1_coco_11_06_2017,修改配置文件,将初始学习率设置为 0.004,衰减步数和衰减系数分别设置为 100 和 0.95,训练步数设置为 10000 步。实验采用 GPU 训练的方式来加快训练速度,利用 Object Detection API 下的 train.py 开始训练,经过约 3 h 的训练,得到训练模型。训练结束后,利用 export_inference_graph.py 可以将训练得到了包含模型结构和参数的临时文件转化为可独立运行的 pb 格式文件,方便模型的移植与集成。

2.4 结果对比分析

训练结束后,利用 Tensorboard 可以方便地得到模型的损失率变化情况。Faster R-CNN 模型和 SSD-Mobilenet 模型的 loss 曲线分别如图 6 和图 7 所示。

由图 6 和图 7 可知,loss 随着迭代次数的增加而下降,经过 10000 次迭代,Faster R-CNN 模型的 loss 逐渐稳定在 0.1 左右,SSD-MobileNet 模型的

loss 逐渐稳定在 1.5,而 Faster R-CNN 模型的收敛速度比 SSD-MobileNet 模型快。

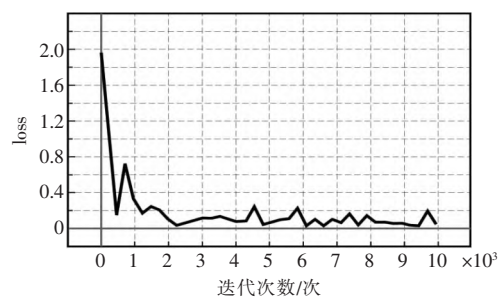


图6 Faster R-CNN 模型 loss 曲线

Fig.6 Loss curve of Faster R-CNN

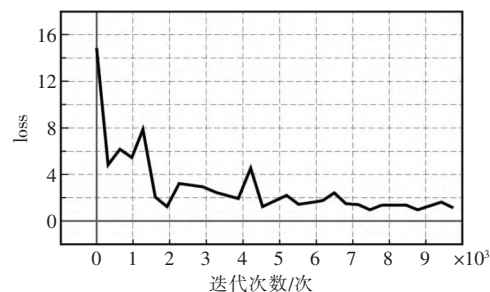


图7 SSD-MobileNet 模型 loss 曲线

Fig.7 Loss curve of SSD-MobileNet

利用 Object Detection API 下的 eval.py 可以对训练结果进行验证,获得检测准确率;另外,对同一张图像进行目标检测,可以获得 Faster R-CNN 模型和 SSD-MobileNet 模型的检测时间。二者的检测准确率和检测时间对比如表 1 所示。

表1 检测准确率和检测时间对比

Tab.1 Comparison of detection accuracy and detection time

模型	准确率/(%)	时间/(ms)
Faster R-CNN	100	6185.80
SSD-MobileNet	98.5	166.94

由表 1 可知,对于采集的物流周转箱数据集,Faster R-CNN 模型的检测准确率可以达到 100%,

而 SSD-MobileNet 模型的检测准确率也达到了98.5%。而在检测时间方面,二者相差巨大,通过对比可知,相比于 Faster R-CNN 模型,SSD-Mobilenet 模型更能满足物流环境下单元货架目标检测的实时性需求。

3 基于 QT 的物流单元货架检测界面设计

SSD-MobileNet 模型能够检测出物流周转箱在单元货架上的位置,并画出包围盒,通过包围盒的像素高度可以得到物流周转箱的层数。所用的单元货架分为4层,每一层分为左右位置,最上面一层的左右位置分别最多能放置一个物流周转箱,其余三层的每个位置最多能叠放3个物流周转箱。将整个货架划分为8个区域,分别编号01,02,1,2,3,4,5,6,如图8所示。



图8 单元货架划分区域

Fig.8 Division of logistics unit shelf

当相机拍照时,相机与单元货架的位置相对固定,这样就可以根据检测到的包围盒的中心点的像素坐标,判断包围盒所在的单元货架区域,继而得到物流周转箱在单元货架上的位置。

QT 是跨平台的 C++图形用户界面开发框架,利用训练转换得到的 SSD-MobileNet 模型的 pb 文件并结合 OpenCV 库,在 QT 平台上设计了物流单元货架检测界面,如图9所示。

物流单元货架检测界面由相机状态显示模块、原图及检测结果显示模块、设置模块以及货架状态显示模块组成,各模块的功能如下:

(1)相机状态显示模块,主要用于显示相机连接状态以及获取相机名称;

(2)原图及检测结果显示模块,用于显示采集到的图像以及检测后的图像;

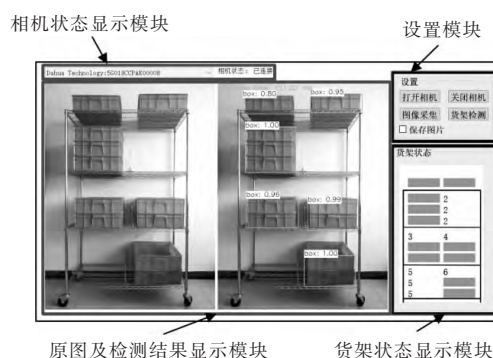


图9 物流单元货架目标检测界面

Fig.9 Interface of object detection of logistics unit shelf

(3)设置模块,主要包括“打开相机”按钮、“关闭相机”按钮、“图像采集”按钮和“货架检测”按钮以及“保存图片”选项;

(4)货架状态显示模块,能够直观地反映检测结果,显示单元货架上每一层每个位置的物流周转箱的层数。

4 结语

本文将深度学习的方法应用到物流环境下的单元货架目标检测,基于 Tensorflow 框架下的 Object Detection API,通过采集实地图像进行模型训练,综合对比了 Faster R-CNN 模型和 SSD-MobileNet 模型的检测准确率和检测时间,实验发现二者的检测准确率均能满足要求,而 SSD-MobileNet 的检测时间仅为 166.94 ms,因此,选择了兼顾准确性与实时性的 SSD-MobileNet 模型来检测单元货架上的物流周转箱的摆放位置与叠放层数,并在 QT 平台上设计开发了基于 SSD-MobileNet 算法的物流单元货架目标检测界面。该检测界面能够方便直观地查看物流单元货架的状态。

参考文献:

- [1] 郑泽宇,梁博文,顾思宇.TensorFlow:实战 Google 深度学习框架(第二版)[M].北京:电子工业出版社,2018.
- [2] 童星,张激.基于 SSD-MobileNet 模型的 ROS 平台目标检测[J].计算机系统应用,2019,28(1):94-99.
- [3] Abadi M,Barham P,Chen J,et al. TensorFlow: A system for large-scale machine learning[C]//12th USENIX Conference on Operating Systems Design and Implementation,Savannah,2-4 November 2016:265-283.
- [4] 曹燕,李欢,王天宝.基于深度学习的目标检测算法研究综述[J].计算机与现代化,2020(5):63-69.

(下转第 55 页)

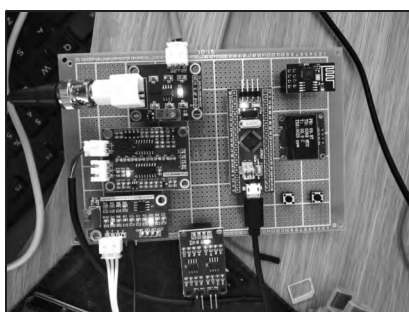


图 12 下位机运行

Fig.12 Operation of lower computer

为试验地点,接取 3 杯自来水,然后向这 3 杯自来水中由少到多加入 pH 值为 4.0 的缓冲液、含磷化肥、干燥的泥土及 90 ℃开水作为待测溶液,将待测溶液充分搅拌后静置 10 s,再分别使用标准检测仪器、本文系统进行检测,得到的试验数据见表 1。

表 1 水质参数测试数据

Tab.1 Water quality parameter test data

参数	测量值	标准值	误差
pH	7.86	7.79	0.8%
pH	7.12	7.16	0.5%
pH	6.65	6.59	0.9%
TDS	28.5×10^{-6}	27.8×10^{-6}	2.5%
TDS	29.1×10^{-6}	29.7×10^{-6}	2.0%
TDS	29.6×10^{-6}	30.3×10^{-6}	2.3%
浑浊度/NTU	275	283	2.8%
浑浊度/NTU	323	331	2.4%
浑浊度/NTU	490	499	1.8%
温度/℃	18.9	18.8	0.5%
温度/℃	32.3	32.5	0.6%
温度/℃	46.5	46.8	0.6%

注:NTU(nephelometric turbidity unit)为散射浊度单位。

表中,由标准检测仪器检测的数据为标准值,由本系统测量得到的数据为测量值;每种参数的 3 个数据自上而下对应于污染程度逐渐增加的待测溶液。

经过试验验证,采用所设计系统测量出来的水质参数与标准值间的误差均在 3% 以内,且上位机运行显示结果与下位机相同,由此表明该系统可以稳定地检测水质各项参数。上位机运行界面

如图 13 所示。

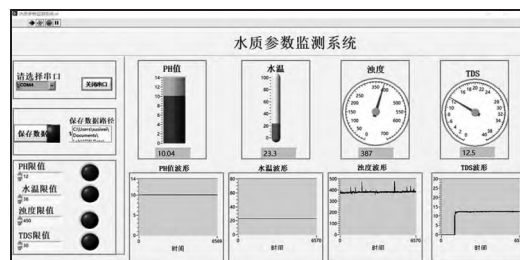


图 13 上位机运行界面

Fig.13 Host computer operation interface

5 结语

通过试验验证,所设计的基于物联网多参数水质监控系统能够稳定运行,可以实现多种方式对水质参数的实时监测,且系统造价低,测量精度较高。另外,LabVIEW 上位机人机交互界面操作简单,功能较为丰富,手机 App 显示界面较为直观,该系统能够满足用户的绝大部分需求。该系统还可以应用于水产养殖等需要监测水质的领域,具有较强的应用前景^[8]。

参考文献:

- [1] 安宝明,徐从海,黄永军.基于地理信息技术的典型地下水污染评析[J].工业安全与环保,2019,45(12):88-92.
- [2] Zhang J,Wei Z,Jia H F,et al.Factors influencing water quality indices in a typical urban river originated with reclaimed water [J].Frontiers of Environmental Science & Engineering(环境科学与工程前沿:英文版),2017,11(4):73-82.
- [3] 吴鹏,汪善盛.水质监测系统的研究[J].现代电子技术,2019,42(13):122-126.
- [4] 韩剑,莫德清.基于 LabVIEW 的智能家居多参数监控系统设计[J].现代电子技术,2020,43(4):48-50.
- [5] 杨琦.消防水池水质和水量监测指标与物联网系统构建的研究[J].给水排水,2019,45(11):103-107.
- [6] 曹俊.基于物联网的大棚环境监控系统研究[J].农机化研究,2019,41(12):212-215.
- [7] 刘子洋,刘忠富,赵宏宇,等.基于 STM32 及 GSM 技术的智能插座设计[J].自动化与仪表,2020,35(1):95-99.
- [8] 张琴,戴阳,杨胜龙,等.基于 LoRa 的低功耗水产养殖水质监测系统[J].传感器与微系统,2019,38(11):96-99.

(上接第 50 页)

- [5] Ren S,He K,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):1137-1149.
- [6] Liu W,Anguelov D,Erhan D,et al. SSD:Single shot MultiBox

detector[C]// Proc of European Conference on Computer Vision (s. 1.):Springer International Publishing,2016:21-37.

- [7] Howard A G,Zhu M,Chen B,et al.MobileNets:Efficient convolutional neural networks for mobile vision applications [J].Proceedings of the IEEE,2017,105(12):2295-2329.