

基于深度学习的盲人辅助购物 APP 的设计与实现

张 洁 刘芬芬 谢 宏 邵秋丹 林 建*

(贵州大学 计算机科学与技术学院, 贵州 贵阳 550025)

摘 要: 针对盲人群体在超市购物存在视觉障碍的问题, 笔者设计一款基于深度学习的盲人辅助购物 APP。设计的功能包括商品种类识别、商品颜色识别、商品标签识别以及语音播报。该 APP 采用轻量级神经网络 MobileNet 和单阶段目标检测算法 SSD 完成商品种类识别, 借助 Tesseract-OCR 技术完成商品标签识别, 利用讯飞语音 SDK 完成语音播报, 拥有广阔的应用前景。

关键词: 盲人辅助购物; 深度学习; MobileNet 神经网络

中图分类号: TH77; TP18; TP391.41 **文献标识码:** A **文章编号:** 1003-9767 (2021) 04-103-02

Design of Blind Assistant Shopping APP Based on Deep Learning

ZHANG Jie, LIU Fenfen, XIE Hong, SHAO Qiudan, LIN Jian*

(College of Computer Science and Technology, Guizhou University, Guiyang Guizhou 550025, China)

Abstract: Aiming at the visual impairment of blind people shopping in supermarkets, the author designed an assisted shopping APP for blind people based on deep learning. The designed functions include product type recognition, product color recognition, product label recognition, and voice broadcast. The APP uses lightweight neural network MobileNet and single-stage target detection algorithm SSD to complete product category recognition, Tesseract-OCR technology to complete product label recognition, and iFLYTEK Voice SDK to complete voice broadcast. It has broad application prospects.

Keywords: blind assisted shopping; deep learning; MobileNet neural network

0 引言

盲人在超市购物存在诸多困难, 如跑错商品区域或者误拿所需商品等问题。在软件方面, 虽然市场上的无障碍辅助软件已经出现多年, 但是仍存在较大局限^[1]; 在硬件方面, 全国盲人协会与计算机咨询有限公司合作设计出一种称为“扫描读出”的便携式装置, 但这种便携式装置对于盲人来说并不方便且设备价格较高。在此背景下, 本文设计一种盲人辅助购物软件, 在深度学习的基础上通过手机摄像头识别商品并播报商品的种类、颜色、标签信息, 帮助盲人解决在超市中购物难的问题。

1 相关技术

盲人辅助购物 APP 系统使用的相关技术有轻量级神经网络 MobileNet、目标检测算法 SSD、OpenCV 计算机视觉库、

光学字符识别技术 OCR 及语音合成技术。MobileNet 使用深度可分离卷积, 能够大幅度减少模型的参数量和计算量^[2]。SSD (Single Shot MultiBox Detector) 是单阶段目标检测算法, 通过主干网络完成特征抽取后, 直接进行区域回归和目标分类^[3]。OpenCV 视觉库提供了很多图像处理算法^[4]。光学字符识别技术 OCR (Optical Character Recognition) 是一种用于识别图像和字符的识别技术^[5]。语音合成技术 (Speech to Text, STT) 是自动通过文本生成声音的过程。

2 盲人辅助购物 APP 的设计

2.1 整体功能设计

盲人辅助购物 APP 的整体功能包括商品种类识别、商品颜色识别、商品标签识别以及语音播报。系统的整体功能设

基金项目: 贵州省高等学校大学生创新创业训练计划项目“基于人工智能的盲人辅助购物 APP 的设计与实现”(项目编号: 202010657047); 贵州大学高等教育研究项目“本科导师制人才创新能力培养模式研究”(项目编号: Gdgj2016008)。

作者简介: 张洁 (1999—), 女, 湖北襄阳人, 本科。研究方向: 深度学习、计算机视觉。

通信作者: 林建 (1994—), 男, 贵州黔西人, 硕士研究生。研究方向: 深度学习、推荐系统。E-mail: jian_lin0802@163.com。

计如图1所示。

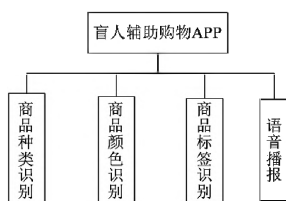


图1 系统功能图

考虑到目标用户是盲人，功能图标以圆形样式展示在手机屏幕中间，以方便盲人点击功能按钮，同时语音播报盲人点击的功能按钮信息。功能界面如图2所示。

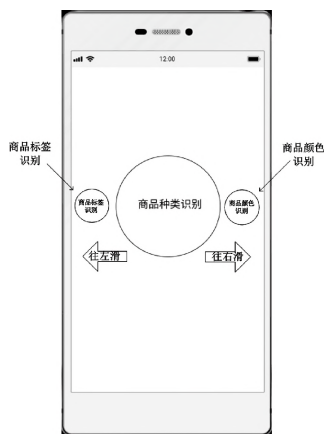


图2 功能界面图

2.2 商品种类识别功能设计

商品种类识别主要将深度学习模型部署在移动端，通过处理手机摄像头捕获的视频帧数据，完成对商品的种类识别。商品种类识别的过程如图3所示。考虑到移动端计算及内存资源有限，APP采用轻量级神经网络模型MobileNet来提高运算性能和内存的利用率。MobileNet使用深度可分离卷积替换标准卷积操作，同时引入宽度乘数和分辨率乘数，使得模型的计算量和参数量大大减少。同时，为提高检测视频流的实时性和分类的准确度，结合MobileNet神经网络模型和SSD算法框架的优势，在保证检测精度的前提下，加快神经网络模型的收敛速度。

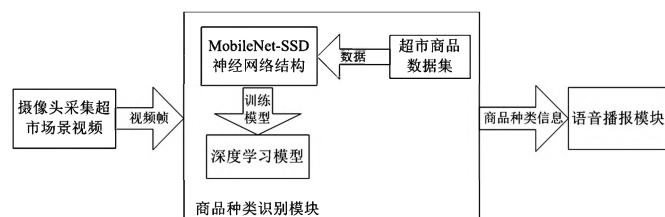


图3 商品种类的识别处理过程

2.3 商品颜色识别功能设计

商品颜色识别是基于OpenCV库函数对摄像头取景范围内的目标物体颜色进行识别。商品颜色识别的处理过程如图4所示，首先通过手机摄像头获取视频，然后利用OpenCV函数解析视频获取图像序列帧，将图片的颜色空间转换到

HSV (Hue Saturation Value) 颜色空间，再进行直方图均衡化和颜色检测，通过分量值的范围来判断商品颜色，最后将颜色信息输入语音播报模块中，把文字信息转换为语音信息，完成商品颜色识别的过程。

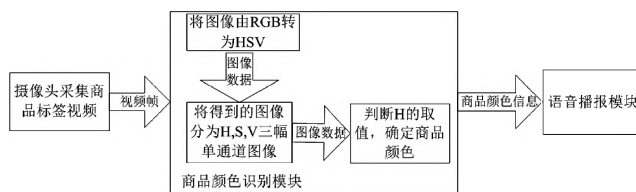


图4 商品颜色识别的处理过程

2.4 商品标签识别功能设计

商品标签识别采用光学字符识别OCR技术进行识别。商品标签识别的过程如图5所示，通过手机摄像头获得含有商品标签的视频，利用OpenCV函数解析视频获取视频帧，然后将使用了OCR引擎Tesseract的开源项目Tess-two部署到移动端，用于识别标签图像上的文字，最后把识别的文字信息输入语音播报模块中，把文字信息转换为语音信息，完成商品标签的识别过程。

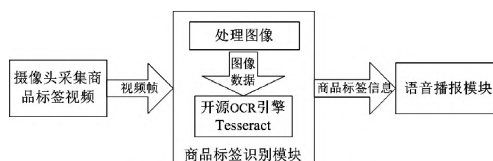


图5 商品标签识别的处理过程

2.5 语音播报功能设计

语音播报功能是借助讯飞的语音SDK完成语音播报的。首先使用SpeechSynthesizer类调用createSynthesizer函数创建语音合成对象，然后通过设置参数调节朗读的语速和音调，最后调用startSpeaking函数实现语音朗读功能。语音播报模块能够接收其他模块识别的商品信息，利用讯飞提供的语音合成技术实现文字信息到语音信息的转换。

3 结语

随着社会的快速发展和技术的不断进步，各界愈发关注盲人的生活。针对盲人在生活中面临的问题，本文在深度学习的基础上设计了一款盲人辅助购物APP，能够帮助盲人轻松完成购物。

参考文献

- [1] 倪赞民, 张寒凝. 盲人移动购物趋势分析研究[J]. 设计, 2017(13):60-61.
- [2] 罗莎莎. 基于MobileNet-SSD模型的道路目标识别研究及其安卓应用开发[D]. 广州: 华南理工大学, 2019.
- [3] 孙元辉, 徐智勇, 张建林, 等. 基于深度学习的单步目标检测器特征增强算法[J]. 半导体光电, 2019,40(1):108-111.
- [4] 任志敏. 基于Android及OpenCV平台的形状及颜色识别系统设计[J]. 山西电子技术, 2019(4):51-53,62.
- [5] 张婷婷, 马明栋, 王得玉. OCR文字识别技术的研究[J]. 计算机技术与发展, 2020,30(4):85-88.