

文章编号: 2095-2163(2023)12-0154-04

中图分类号: TP391

文献标志码: A

基于卷积神经网络的食物识别及实现

颜乾坤^{1,2}, 肖玉芝^{1,3}, 杜秀娟^{1,2,3,4}, 赵建^{1,2}

(1 青海师范大学 计算机学院, 西宁 810008; 2 青海师范大学 青海省物联网重点实验室, 西宁 810008;

3 青海师范大学 藏语智能信息处理及应用国家重点实验室, 西宁 810008; 4 高原科学与可持续发展研究院, 西宁 810008)

摘要: 在超市无人结算服务中,使用电子标签对部分货物如水果、蔬菜等进行结算的成本过高、便捷性不高,至今依然采用人工结算的方式。针对这一问题,本文提出了基于卷积神经网络的食物识别方法。通过自建水果数据集来训练卷积神经网络分类模型;基于训练后的模型构建可视化平台进行食物识别。实验结果表明,利用卷积神经网络的食物识别的预测准确率为96.34%。

关键词: 卷积神经网络; 食物识别; 超市无人结算

Convolutional neural network based food recognition and implementation

YAN Qiankun^{1,2}, XIAO Yuzhi^{1,3}, DU Xiujuan^{1,2,3,4}, ZHAO Jian^{1,2}

(1 Computer Department, Qinghai Normal University, Xining 810008, China; 2 Qinghai Provincial Key Laboratory of IoT, Qinghai Normal University, Xining 810008, China; 3 The State Key Laboratory of Tibetan Intelligent Information Processing and Application, Qinghai Normal University, Xining 810008, China; 4 Academy of Plateau Science and Sustainability, Xining 810008, China)

Abstract: In the unmanned settlement service of supermarkets, the cost of using electronic tags to settle some goods such as fruits and vegetables is too high and the convenience is not high, and manual settlement is still used so far. To address this problem, this paper proposes a food recognition method based on convolutional neural network. A self-built fruit dataset is used to train a convolutional neural network classification model; a visualization platform is constructed based on the trained model for food recognition. The experimental results show that the prediction accuracy of food recognition using convolutional neural network is 96.34%.

Key words: convolutional neural network; food recognition; unmanned checkout in supermarkets

0 引言

食物识别技术可以应用到生活中的很多场景中,能够代替人进行一些枯燥的识别任务。目前,有许多学者研究品质分级、残缺检测、成熟度识别和机器人采摘等技术,多类食物识别的研究较少,而多类食物识别在实践中有着广泛的应用价值^[1]。

如今,许多产业正在逐渐实现无人化,即用机器代替人进行一些枯燥且繁琐的工作。在超市的无人结算服务中,多数大件物品可以利用电子标签等可识别工具对物品进行分类,以供无人结算设备进行

计价。但也有部分货物无法使用电子标签标记,如水果、蔬菜等食物,这些食物体积大小不一,使用电子标签对其进行标码极其不便;由于一些食物数量极多,如水果、大米等,难以使用电子标签对其一一标记。因此,这些食物难以利用电子标签对其进行有效的标记。

近年来,随着卷积神经网络的发展,卷积神经网络在许多领域内都展现其强大的能力^[2]。本文利用卷积神经网络来完成食物的识别,采用轻量级卷积神经网络 MobileNet 作为基础模型,以5种水果:苹果、梨、香蕉、橘子、西红柿为实验对象,利用手机

基金项目: 青海省物联网重点实验室(2022-ZJ-Y21); 国家自然科学基金(61962052)。

作者简介: 颜乾坤(1999-),男,硕士研究生,主要研究方向:目标检测;肖玉芝(1980-),女,博士,教授,硕士生导师,主要研究方向:复杂网络;赵建(1999-),男,硕士研究生,主要研究方向:图像融合、目标检测。

通讯作者: 杜秀娟(1970-),女,博士,教授,博士生导师,主要研究方向:无线网络与安全、物联网技术。Email:124111397@qq.com

收稿日期: 2022-10-16

拍摄的 2 500 张图片作为数据集供模型训练、测试, 并利用该模型来搭建可视化平台, 实现食物识别。

1 MobileNet 网络

1.1 MobileNet 网络概述

卷积神经网络 (Convolutional Neural Networks, CNN) 是一类包含卷积核与卷积计算并具有深度结构的前馈神经网络, 是深度学习 (Deep Learning) 的代表算法之一^[3]。相对于其他的神经网络, CNN 一层或者多层中用卷积运算来替换矩阵乘法运算。CNN 中主要有 3 个基本层: 卷积层、池化层、全连接层, 在不同的卷积神经网络中又有各自的延伸。CNN 是一种深度的监督学习下的机器学习模型, 具有极强的适应性, 善于挖掘数据局部特征、提取全局训练特征、分类等^[4]。CNN 的最大特点是稀疏连接 (局部感受) 和权值共享, 可以减少训练参数, 减少计算复杂度。CNN 的权值共享结构网络类似于生物神经网络, 在模式识别等各个领域都取得了不错的效果^[5]。

但多数卷积神经网络对于设备的内存需求大、运算力要求高, 导致其无法在移动设备以及嵌入式设备上高效运行。MobileNet 网络作为一种轻量级的 CNN 网络, 常被应用于移动端或者嵌入式设备中^[6]。相比于传统卷积神经网络, MobileNet 网络在准确率小幅降低的前提下大大减少模型参数与运算量, 与经典的 VGG16 网络相比, MobileNet 网络准确率在减少 0.9% 的前提下, 模型参数只有 VGG 的 1/32^[7]。考虑到嵌入式设备中部署时的硬件要求, 本文采用 MobileNet 网络作为基础网络模型。

1.2 深度可分离卷积

MobileNet 网络最大的特点是使用深度可分离

卷积代替标准卷积。标准卷积每个卷积核的维度与输入维度相同, 每个通道单独做卷积运算后相加; 而深度可分离卷积作为 MobileNet 的主要结构, 深度可分离卷积分为两步操作, 第一步为深度卷积, 用 3×3 大小的卷积核对特征图进行卷积操作, 每个核只负责特征图中的一个通道, 再通过一个 BN (Batch Normalization) 层和一个 ReLU 激活函数后逐点卷积; 第二步为逐点卷积, 逐点卷积的卷积核大小为 1×1×M, 其中 M 为输入数据的维度, 逐点卷积每次仅卷积一个像素的区域, 通过将上一层的特征图在深度方向上进行加权组合, 以产生新的特征图。深度可分离卷积的操作流程图如图 1 所示。



图 1 深度可分离卷积操作流程图

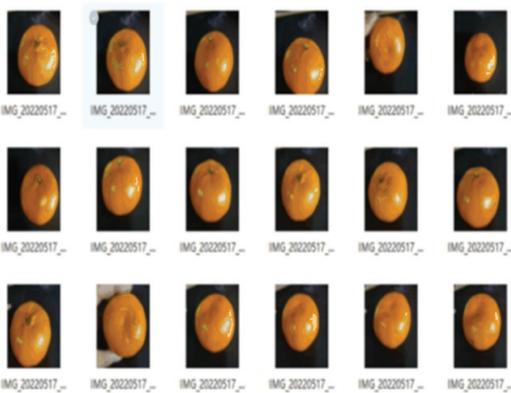
Fig. 1 Flowchart of deep separable convolution operation

2 数据集构建与模型训练

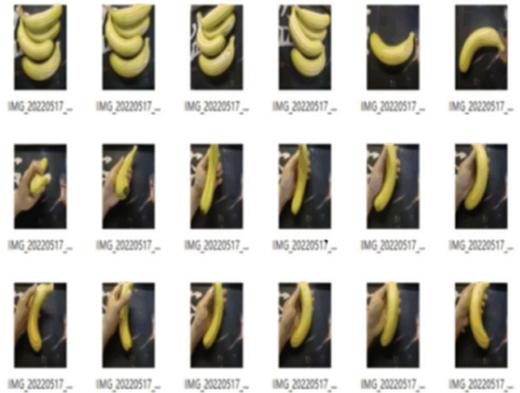
2.1 数据集的制作

2.1.1 拍摄照片

使用手机拍摄苹果、橘子、西红柿、香蕉、梨这 5 种水果各 500 张图片。部分数据集图片如图 2 所示。



(a) 部分橘子数据集图片



(b) 部分香蕉数据集图片

图 2 部分数据集图片

Fig. 2 Selected datasets images

2.1.2 数据集清洗

人工剔除模糊、残缺照片,清洗后的数据集分布情况见表1。

表1 数据集分布情况
Table 1 Distribution of datasets

类别	训练集	测试集
苹果	400	100
梨	400	100
香蕉	400	100
橘子	400	100
西红柿	400	100

2.1.3 数据集制作

数据集清洗过后,再对数据集进行预处理及训练集与测试集的随机分类。训练集、测试集的比例为8:2,并给每一个图片标注一个标签,标签内容

为:香蕉(0)、西红柿(1)、橘子(2)、苹果(3)、梨(4)。

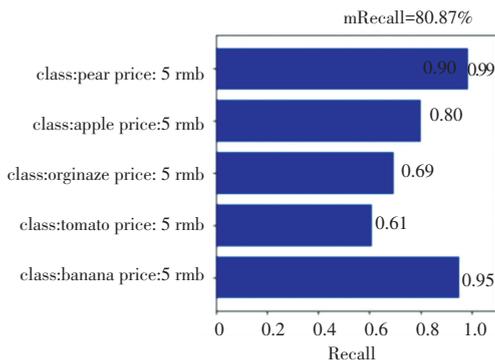
2.2 食物识别模型的训练

使用训练集对模型进行训练,共50轮。训练过程中,损失率不断减小、精准度不断提升。迭代次数近40次时,损失函数基本收敛,接近于0。

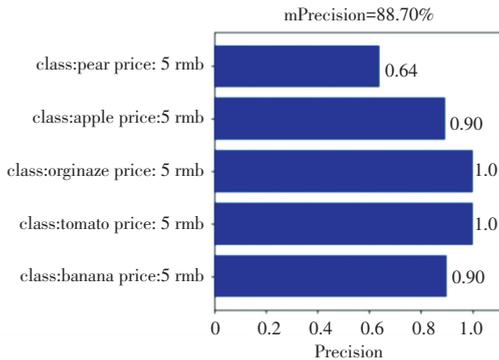
2.3 食物识别模型的评估

在训练好模型后,用与训练集图像不同的测试集对模型进行测试与评估,得到精准率、召回率。模型训练第10轮时对测试集的评估图如图3所示。由图3可见,经过10轮的训练后召回率为80.77%,精准率为88.70%,仍不理想。

继续训练模型训练到第50轮,结果如图4所示,精准率从88.70%提高到93.46%,召回率从80.87%提高到97.30%,提升明显。



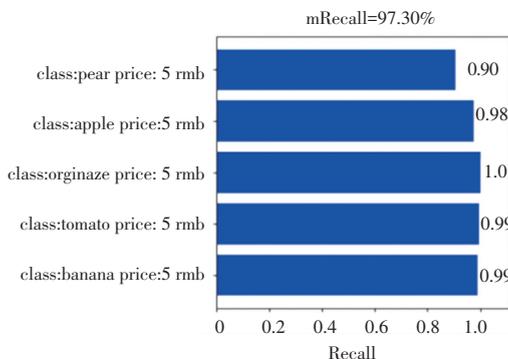
(a) 第十轮召回率



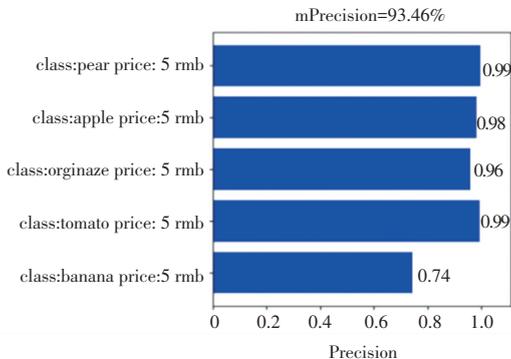
(b) 第十轮精准率

图3 模型训练第10轮评估图

Fig. 3 Model training round 10 evaluation chart



(a) 第50轮召回率



(b) 第50轮精准率

图4 模型训练第50轮评估

Fig. 4 Model training round 50 evaluation

3 实验

本文利用开源的计算机视觉框架OpenCV来构建可视化平台。OpenCV是Intel公司资助的开源计

算机视觉库,具有丰富的图像处理函数、良好的跨平台性、库更新速度快等优势^[8]。利用OpenCV提供的各种图像处理算法,将从摄像头采集的图像视频信息,通过图像滤波、图像分割等算法,能提取到丰

富的图像信息;再对其进行轮廓特征、运动信息分析,得出图像视频的识别图像内容,并转化为计算机图像信息,传递给识别模型,得到识别结果,从而实现计算机识别物体的目的。实时识别的示例图如图 5 所示,摄像头实时捕捉当前画面并且传递到模型进行识别,如果画面中出现了模型所训练的食物种类,则给出相关种类名称、价格信息。



图 5 动态摄像头实时识别示例图

Fig. 5 Example diagram of live motion camera recognition

4 结束语

本文针对超市无人结算服务中,使用电子标签对部分货物如水果、蔬菜等进行结算的成本过大、便捷性不高的问题,结合轻量级卷积神经网络

MobileNet 搭建了一个小型食物识别系统。通过自制数据集训练模型,获得了很好的识别效果,训练好的模型通过 OpenCV 库实现了食物的识别,但识别率以及识别响应时差对于零售行业来说显然不够,对于模型的优化和设计仍需完善。

参考文献

- [1] 彭红星,黄博,邵园园,等. 自然环境下多类水果采摘目标识别的通用改进 SSD 模型[J]. 农业工程学报,2018,34(16):155-162.
- [2] 卢宏涛,张秦川. 深度卷积神经网络在计算机视觉中的应用研究综述[J]. 数据采集与处理,2016,31(1):1-17.
- [3] PAN Ning. A sensor data fusion algorithm based on suboptimal network powered deep learning [J]. Alexandria Engineering Journal,2022,61(9):7129-7139.
- [4] 周飞燕,金林鹏,董军. 卷积神经网络研究综述[J]. 计算机学报,2017,40(6):1229-1251.
- [5] 张顺,龚怡宏,王进军. 深度卷积神经网络的发展及其在计算机视觉领域的应用[J]. 计算机学报,2019,42(3):453-482.
- [6] 陈智超,焦海宁,杨杰,等. 基于改进 MobileNet v2 的垃圾图像分类算法[J]. 浙江大学学报(工学版),2021,55(8):1490-1499.
- [7] 邵伟平,王兴,曹昭睿,等. 基于 MobileNet 与 YOLOv3 的轻量化卷积神经网络设计[J]. 计算机应用,2020,40(S1):8-13.
- [8] PRADEEP S, Yogesh Kumar Sharma. Using openCV for machine learning in real time computer vision and image processing [J]. International Journal of Recent Technology and Engineering,2019,8(1):1846-1848.
- [9] pressure estimation from appropriate and inappropriate PPG signals using a whole-based method[J]. Biomed Signal Process Control, 2019, 47:196-206.
- [10] MOU H L, YU J S. CNN-LSTM prediction method for blood pressure based on pulse wave [J]. Electronics, 2021, 10(14):1664.
- [11] 宋义林,张彤,高树枝. 基于容积脉搏波振幅梯度的血压检测法的研究[J]. 生物医学工程研究,2018,37(3):281-285.
- [12] 徐可欣,王继寸,余辉,等. 脉搏波时域特征与血压相关性的研究[J]. 中国医疗设备,2009,24(8):42-45.
- [13] 汤池. 利用脉搏波无创连续检测血压的实验研究[D]. 西安:第四军医大学,2005.
- [14] 叶海民,戎蒙恬,邓晓东,等. 基于绝对值差的归一化波形匹配混合算法[J]. 信息技术,2011,35(8):89-93.
- [15] 董云先,王承亮,宋岩,等. 火电厂凝汽器真空实时寻优数据挖掘算法研究及测试[J]. 科技创新导报,2019,16(16):78-82.
- [16] 文传源. 人工神经网络及其应用[J]. 计算机仿真,1997,14(2):4.
- [17] 陆琳. 不确定信息车辆路径问题及其智能算法研究[D]. 南京:南京航空航天大学,2007.
- [18] 毛健,赵红东,姚婧婧. 人工神经网络的发展及应用[J]. 电子设计工程,2011,19(24):4.
- [19] 黄文柔,欧阳婧. 基于人工神经网络的宽带滤波器设计[J]. 电视技术,2020,44(11):55-57.

(上接第 153 页)

为了促进中医脉诊的客观化发展,将中医的精华理论内化到人工智能中,本文采用机器学习方法,使用脉搏传感信号数据集通过不同的机器学习模型预测人体血压,经过模型的测试与评估,平均绝对误差可达到 0.087,从而证实脉搏传感信号可以准确的预测出人体血压。

参考文献

- [1] 高一凡. 中医舌脉象信息提取与分析及辅助诊断系统的设计 [D]. 郑州:郑州大学,2020.
- [2] 魏红,徐刚. 从中医整体、动态、平衡观论脉诊客观化研究[J]. 中医杂志,2014,55(1):25-27.
- [3] 中国医师协会高血压专业委员会等. 阻塞性睡眠呼吸暂停相关性高血压临床诊断和治疗专家共识[J]. 国呼与危重监护杂志,2013,12(5):435-441.
- [4] 胡琪鑫,徐亚. 地球物理信号特征识别与解释的机器学习方法及应用综述[J]. 地球物理学进展,2022,37(6):2395-2407.
- [5] 汤池,倪安胜,王跃民,等. 利用脉搏波连续检测家兔血压的实验研究[J]. 中国医学物理学杂志,2009,26(1):1019-1021,1042.
- [6] 樊艳梦. 基于 PPG 信号的无袖带血压检测方法研究[D]. 合肥:合肥工业大学,2021.
- [7] MOUSAVI S S, FIROUZMAND M, CHARMI M, et al. Blood