

基于图像识别的果蔬自助结算系统

段中兴, 李伟哲, 张亚俐, 周 孟, 丁青辉

(西安建筑科技大学 信息与控制工程学院, 西安 710055)

摘要: 实现对缺少条形码的水果蔬菜的识别与结算, 是超市自助结算的一大难题; 为在资源有限的结算终端设备上实现超市果蔬的识别与分类, 提出了一种基于神经网络的果蔬识别算法; 通过增加网络宽度的方法改进 Alex Net, 提升识别性能; 结合压力传感器、摄像头等硬件设备, 在树莓派上进行实验, 完成了果蔬自助结算系统的搭建; 经实验测试, 系统对果蔬的平均识别准确率可达 98.25%, 单次结算总耗时约 7.48 s, 仅为人工结算耗时的 1/4, 满足果蔬自助结算系统的实际应用需求。

关键词: 图像识别; 果蔬分类; Alex Net; 自助结算系统

Fruit and Vegetable Self-service Settlement System Based on Image Recognition

DUAN Zhongxing, LI Weizhe, ZHANG Yali, ZHOU Meng, DING Qinghui
(School of Information and Control Engineering, Shaanxi Xi'an University of Architecture and
Technology, Xi'an 710055, China)

Abstract: Realizing the identification and settlement of fruits and vegetables that lack barcodes is a major problem in supermarket self-service settlement. In order to realize the identification and classification of fruits and vegetables in supermarkets on the settlement terminal equipment with limited resources, a fruit and vegetable identification algorithm based on neural network is proposed. Improve Alex Net by increasing the network width to improve recognition performance. Combined with hardware devices such as pressure sensors and cameras, experiments were conducted on the Raspberry Pi to complete the construction of a fruit and vegetable self-service settlement system. After experimental tests, the average recognition accuracy of the system for fruits and vegetables can reach 98.25%, and the total time for a single settlement is about 7.48 seconds, which is only 1/4 of the time for manual settlement, which meets the actual application requirements of the fruit and vegetable self-service settlement system.

Keywords: image recognition; fruit and vegetable classification; Alex Net; self-service settlement system

0 引言

网络技术和电子商务的快速发展, 提高了线上购物的便捷性。由于在线上选购商品到实际收到商品所用时间较长, 而蔬果类商品的保质期较短, 因此消费者多选择在线下商店选购此类商品。而目前线下商店售卖蔬果类商品大多需员工称重后张贴条形码^[1-2], 然后由消费者前往收银台或自助结算终端扫描条形码完成支付, 无形中增加了商家的人力资源投入和消费者的购物流程, 人流量较多时果蔬称重区也会出现较严重的排队问题。因此如何实现果蔬自动识别是当前超市自助结账系统面临的一个亟待解决的问题。

图像识别技术的发展为果蔬自助结算提供了新的解决方案。罗承成等^[3]通过改进单色块及其领域算法提出了一种超市农产品图像识别方法, 能够识别不同光照背景下的果蔬图像; S. Arivazhagan 等^[4]通过提取水果的颜色和纹理特征提出了一种水果图像识别方法, 平均识别准确率达到 86%; S. R. Dubey 等^[5]提出一种基于图像灰度值的纹理

特征分类方法, 实验结果表明, 图像灰度值的纹理特征分类效果更好。但上述方法均需要人工设计图像特征, 其识别分类的效果往往依赖于算法工程师的领域知识和设计经验。

自 Alex Net^[6]取得 Imagenet 竞赛冠军以来, 卷积神经网络^[7]以无需人工设计特征的特点而得到海内外学者的广泛关注并迅速发展, 在人脸识别^[8-9]、农作物识别^[10]、交通标志识别^[11-12]等领域得到了广泛研究。在果蔬图像识别领域, 巨志勇等^[13]利用柔性注意力算法改进 Inception V3 建立了果蔬识别模型, 樊帅昌等^[14]采用迁移学习方法训练了一种基于 Res Net 的蘑菇识别模型, 均得到了较高的识别精度, 但深层卷积神经网络的计算量较大, 模型的响应时间也较长。

综上所述, 本文提出了一种基于改进 Alex Net 的超市果蔬识别方法, 并结合硬件设计实现了果蔬自助结算系统。与 Alex Net 相比, 该方法能够提取不同尺度下的果蔬图像

收稿日期: 2021-05-13; 修回日期: 2021-06-18。

基金项目: 国家自然科学基金项目(51678470)。

作者简介: 段中兴(1969-), 男, 湖南株洲人, 博士生导师, 主要从事智能检测与机器视觉方向的研究。

引用格式: 段中兴, 李伟哲, 张亚俐, 等. 基于图像识别的果蔬自助结算系统[J]. 计算机测量与控制, 2021, 29(12): 195-203.

特征,不同尺度的特征融合后表征能力更强;与 Inception V3^[15-16]、Res Net^[17]等网络相比,改进 Alex Net 的参数数量和计算量较小,对硬件设备的处理器性能等要求较低,识别速率更快,更适合应用于果蔬自助结算系统。与现有自助结算设备相比,基于图像的自助结算系统能够在较短时间内完成果蔬的识别、称重和计价,大大减少了果蔬类商品称重标记的时间以及资源投入。

1 系统结构及原理

1.1 系统整体方案

果蔬自助结算系统无需人工操作即可识别出待检测商品的类别及单价,并通过压力传感器等硬件采集商品重量信息,然后基于商品的类别、单价以及重量等信息通过硬件内置算法计算得出结算信息并在显示器上显示,最后消费者出示付款码由扫描枪扫描完成支付,信息有误时可通知管理人员修改错误信息,系统的整体方案见图 1。

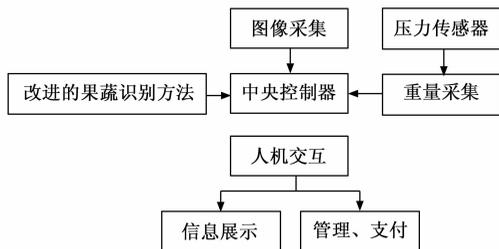


图 1 系统方案示意图

系统由重量采集、图像采集、中央控制器和人机交互共 4 部分组成,其中中央控制器搭载有改进的果蔬识别方法。结算时系统的处理器首先检测到重量采集部分接口的电平变化,之后启动程序准备接收重量信号。检测到重量采集部分接口电平变化的同时,系统启动图像采集部分,获取待结算商品的图像。之后中央控制器运行改进的果蔬识别方法判断果蔬类别,将类别与重量信息融合得到商品总价并通过信息展示部分向消费者展示。若信息无误则消费者可使用支付设备扫描付款码完成支付;若商品信息有误,可通知管理人员,由管理人员使用相应设备修正。

1.2 系统硬件组成及结构

现有自助结算设备通过扫描条形码读取商品信息,这一过程只需要处理少量数据即可实现商品信息的准确获取及结算,因此现有自助结算设备的处理器性能往往比较低,而基于图像识别果蔬类别需要经过大量计算,因此难以直接将果蔬识别方法部署于现有的自助结算设备。

而关于图像识别的研究大多在实验室等环境完成,该环境可以为识别模型的训练与运行提供性能优越的处理器,大幅减少了模型训练与运行的时长。但高性能处理器往往价格昂贵,用于自助结算系统具有极低的性价比,导致系统的应用价值降低。因此该研究一方面通过优化识别模型的网络结构以降低对处理器性能的要求;另一方面综合考虑处理器性能、性价比等因素对系统的硬件进行筛选。

考虑到实际应用时自助结算系统除了识别果蔬类别之外还需要补充多种商品识别方式和支付途径,因此中央控制器需预留一定量的拓展空间;而图像识别过程会涉及大量计算过程,因此中央控制器的处理器(CPU)要具有较高的性能。树莓派 4B 集成了 4 核处理器 Arm Cortex-A72,提供有 Micro-HDMI、USB、CSI、GPIO 等多种拓展接口^[18],具有无线/有线网络和蓝牙等连接方式,作为自助结算系统的中央控制器可以同时满足拓展性和计算资源的需求,树莓派 4B 扩展配件接口(GPIO)的具体分配方式见表 1。

表 1 GPIO 接口分配方式

功能名	物理引脚 BOARD 编码		功能名
3.3 V	1	2	5 V
SDA. 1	3	4	5V
SCL. 1	5	6	GND
GPIO. 7	7	8	TXD
GND	9	10	RXD
GPIO. 0	11	12	GPIO. 1
GPIO. 2	13	14	GND
GPIO. 3	15	16	GPIO. 4
3.3 V	17	18	GPIO. 5
MOSI	19	20	GND
MISO	21	22	GPIO. 6
SCLK	23	24	CE0
GND	25	26	CE1
SDA. 0	27	28	SCL. 0
GPIO. 21	29	30	GND
GPIO. 22	31	32	GPIO. 26
GPIO. 23	33	34	GND
GPIO. 24	35	36	GPIO. 27
GPIO. 25	37	38	GPIO. 28
GND	39	40	GPIO. 29

摄像头是系统图像采集部分最重要的硬件组成,其采集的图像质量与识别的精度息息相关。系统使用的 Raspicam 摄像头通过 CSI 接口与树莓派连接,启用摄像头的支持后可直接通过命令行驱动,相应更快且工作稳定性好,资源消耗低,图像质量好,非常适用于图像识别系统。

系统重量采集部分的输出信号是启动结算程序和计算商品总价的重要信息。重量信息的获取首先由压力传感器将压力值转换为电量信号,之后再通过数模转换芯片将之转换成易处理的数字信号。应变式压力传感器由受压时电阻值会改变的材料和电阻值固定的材料组成图 2 所示的惠斯登桥,惠斯登桥的输入为稳定的电压源 U_{IN} ,无外界压力时惠斯登桥处于平衡状态;受到压力时, R_1 和 R_2 的电阻值发生改变,惠斯登桥退出平衡状态,输出电压值 U_{OUT} 发生改变,经放大器、补偿电路后产生与压力值成线性对应关系的标准输出信号。

考虑到人们在超市内消费时普遍不会一次性购入大量果蔬,因此该研究选用的应变式压力传感器的量程为 5 kg,

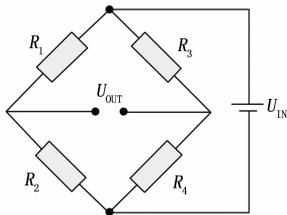


图 2 惠斯登桥

不受压条件下惠斯登桥电阻材料的阻值为 1 kΩ, 电源输入通过 HX711 数模转换芯片与树莓派的 3.3 V 接口连接, 即 $U_{IN}=3.3\text{ V}$ 。

HX711 数模转换芯片^[19]可以将传感器输出的模拟信号转变为数字信号, 应用于自助结算系统的重量测采集模块可以很大程度上降低成本, 提高系统性能和可靠性。芯片共有 4 个输入输出引脚, 分别为 VCC、GND、SCK、DT, 其中信号传输主要通过 SCK 和 DT 引脚实现。将传感器的模拟信号转换为数字信号后, DT 引脚将从高电平降至低电平, 表明芯片已准备输出重量的数字信息, 之后 SCK 引脚输出一串包含重量信息的时钟信号。系统根据 DT 引脚电平的变化情况确定系统的工作状态, 使用 SCK 引脚的输出信号计算商品总价, 芯片与树莓派的引脚连接方式见表 2。

表 2 HX711 与树莓派连接方式

HX711	GPIO 接口
VCC	1(3.3 V)
GND	6(GND)
SCK	13(GPIO. 2)
DT	15(GPIO. 3)

系统人机交互部分包括用于展示商品信息的显示器、用于支付的扫描枪和用于修正信息的键盘鼠标等硬件设备。系统硬件连接的原理如图 3 所示, 图中隐去了硬件设备的内部电路。

2 系统软件设计

系统软件部分由果蔬识别算法和主程序组成, 果蔬识别算法包括算法模型和模型参数两部分, 算法模型在 Alex

Net 的基础上进行改进, 使用不同尺寸的卷积核增加特征的丰富度, 使用“捷径”连接方式增加特征的完整性, 改进后模型能够在性能相对较低的树莓派等微型主机上运行并在较短时间内得到精准的识别结果; 模型参数的训练使用自建数据集以交叉熵值最小化原则训练得到, 训练过程中根据交叉熵值的变化情况动态调整学习率。主程序按一定的时序将系统的外围硬件和内部计算算法串联起来, 充分利用结算过程产生的不同信号, 检测托盘上是否存在待结算商品, 并自行完成识别计价过程, 无需人工参与。消费者完成支付后主程序可自行返回待机状态, 监测托盘上是否存在待结算商品; 若商品信息或总价有误, 可通过管理人员修正支付信息, 消费者支付成功后主程序重新进入待机状态, 系统软件部分的设计思路见图 4。

2.1 果蔬识别方法

超市果蔬种类繁多, 颜色纹理轮廓等特征多种多样且部分品种间存在大量相似特征, 因此识别网络需具备足够的复杂度以确保所提取的特征具备足够的表征能力。同时由于自助结算系统的中央控制器的计算资源并不丰富, 因此网络模型的计算量和参数量应当尽可能小。对比 Alex Net、Inception V3、Res Net 等经典网络模型发现, Alex Net 的网络结构简单, 特征的完整性与多样性相对不足, 但多数研究表明该网络在图像识别领域具有一定的应用潜力; 而 Inception V3、Res Net 等网络结构复杂, 特征表征能力更强, 但参数量和计算量巨大, 对硬件设备的性能要求较高且网络进一步优化改进的空间较小。因此选择在 Alex Net 的基础上进行改进, 使网络模型能够以较小的计算资源需求完成对常见果蔬的准确识别。

2.1.1 并行卷积通道

卷积神经网络通过卷积计算提取图像特征, 得到输入图像的特征图, 特征图尺寸与输入尺寸之间的关系见式 (1):

$$S_{out} = \frac{S_m + 2P - k}{s} + 1 \quad (1)$$

其中: S_m 为输入图像的尺寸, S_{out} 是输出的尺寸, P 是图像填充的尺寸, k 为卷积核尺寸, s 为卷积的步长。

假设输入图像的尺寸为 5×5 , 卷积步长设置为 1, 图像

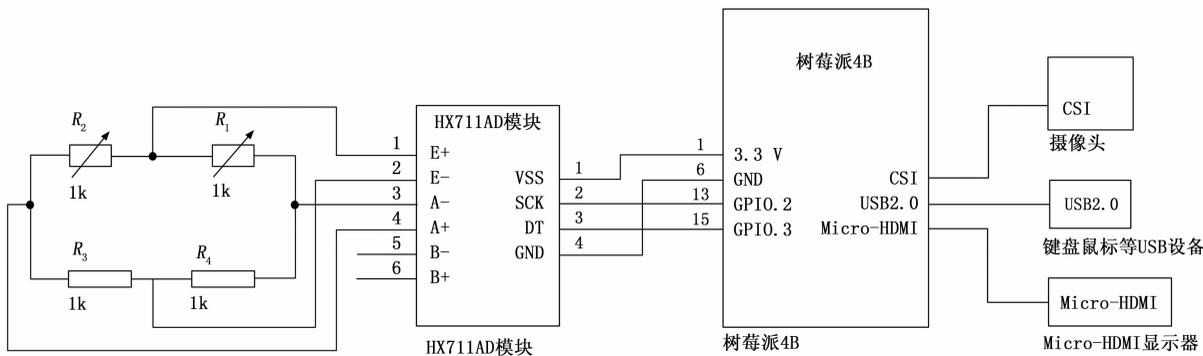


图 3 硬件结构原理图

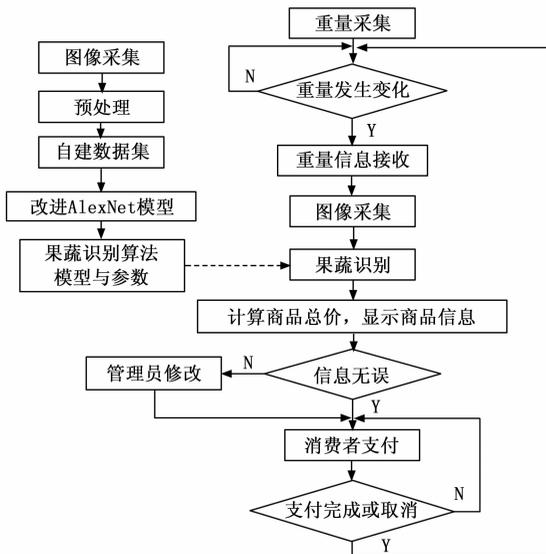


图 4 软件流程图

填充方式为无填充，使用 3×3 大小的卷积核提取图像特征时，输出尺寸 $S_{out} = 3$ ，如图 5 (a) 所示。输出特征图中数值与输入图像的固定像素值区域一一对应，如图 5 (a) 输入图像深色框线标注的 3×3 像素值区域与输出特征图的深色区域，特征图数值对应的像素值区域即为卷积核的感受野，可见 3×3 卷积核的感受野大小为 3×3 ；同理，使用 5×5 大小的卷积核时输出的特征图尺寸为 1×1 ，该数值与全部的 5×5 像素值区域相关，即感受野大小为 5×5 ，如图 5 (b) 所示。

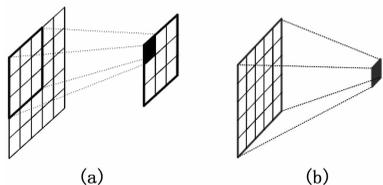


图 5 感受野与卷积核对应关系

果蔬识别过程中，感受野的大小直接影响特征图对不同尺度特征信息的表达。感受野较大时，特征图数值对应较大的图像区域，在这一区域，背景图像与前景图像的像素值差距远大于前景图像中纹理与叶片之间的像素值差距，因此更为细节的纹理等信息对特征图输出值的影响很小，甚至可以忽略不计，导致大感受野下图像特征图更多用来表达轮廓等较为明显的特征而忽视细节纹理特征；感受野较小时，特征图数值对应的图像区域较小，多数区域不会同时包含前景、背景等内容，因此图像轮廓等信息难以在小感受野内充分体现，而纹理等细节信息将对特征值产生较大影响，如图 6 所示。

因此在设计果蔬识别方法时，感受野过大或过小都会导致识别方法损失部分图像特征，从而影响识别精度。Alex Net 的卷积核尺寸较为单一，易丢失不同感受野的图像特征，难以保证特征的丰富性和表征能力。受 C. Szegedy

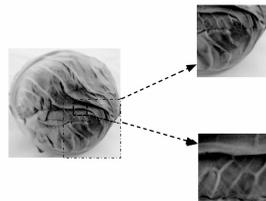


图 6 卷积核尺寸对提取特征的影响

等^[15-16]的启发对 Alex Net 进行改进，在 Alex Net 的每个卷积层上额外增加一条或多条卷积通道，通道使用不同尺寸的卷积核与原来的卷积核组成并行多通道卷积层，增加了卷积核尺寸的丰富度。将不同通道的特征图叠加 (concatenate) 作为该层卷积的特征图输出，输出中同时包含不同感受野的图像特征，提高了输出特征的完整性和表征能力。

将卷积神经网络第 l 层卷积的输出定义为 x_l ，则改进前网络各卷积层的输入输出关系见式 (2)，改进后的输入输出关系见式 (3)。

$$x_l = f(w, h, x_{l-1}) \tag{2}$$

$$x_l = g(f(w_1, h_1, x_{l-1}), f(w_2, h_2, x_{l-1}), \dots) \tag{3}$$

其中： f 代表卷积操作， w, h 分别为卷积核的宽度和高度。 g 代表特征图叠加操作，即将不同卷积通道的输出在特征图厚度（即特征图的通道）上进行堆叠。

鉴于输入图像的尺寸相对较大，若网络首层使用的卷积核尺寸过小，则易导致细节信息不能完整表达，使提取的特征带有大量无效信息，网络首层的改进为在原网络结构的基础上增加一条卷积核尺寸为 7×7 的卷积通道，改进前首层卷积的输入输出见式 (4)，改进后的输入输出关系式见式 (5)：

$$x_1 = f(11, 11, x_0) \tag{4}$$

$$x_1 = g(f(11, 11, x_0), f(7, 7, x_0)) \tag{5}$$

其中： x_0 为网络输入图像，尺寸为 $224 \times 224 \times 3$ 。

之后使用池化卷积降低输出特征图的尺寸，公式见式 (6)：

$$x_{p1} = pool(x_1, 3, 3, 2, valid) \tag{6}$$

其中： x_{p1} 为池化后第一层卷积的输出特征， $pool$ 代表池化运算，包括最大池化和平均池化两种，分别取池化范围内的最大值或平均值作为输出特征。 $pool$ 的输入为第一层卷积的输出 x_1 ，之后两个参数为池化卷积核尺寸的宽度和高度，第 4 个参数为卷积步长，本次池化的像素值区域与下次池化区域之间间隔的特征图行列数，最后一个参数为图像填充方式，包括 same padding 和 valid padding 两种，same padding 也叫 0 填充，即在图像边缘填充 0 像素，可以保持输入与输出维度相同；valid padding 即不填充，输出尺寸小于输入尺寸。

经过一次池化后，特征图尺寸已经相对较小，因此第二个卷积层取消了 7×7 等相对较大的卷积核，在原网络 5×5 卷积的基础上增加了更小的 3×3 、 1×1 两种小卷积核，用于提取不同感受野的特征信息，叠加后使用池化层再次

降低输出特征图的尺寸。改进网络第二层卷积的计算公式见式 (7)、(8):

$$x_2 = g(f(5, 5, x_{p1}), f(3, 3, x_{p1}), f(1, 1, x_{p1})) \quad (7)$$

$$x_{p2} = \text{pool}(x_2, 3, 3, 2, \text{valid}) \quad (8)$$

此时网络输出的尺寸已经较小, 因此后三层卷积在 3×3 卷积的基础上补充了 1×1 卷积, 以满足特征图丰富度的要求。最后三层卷积的运算公式见式 (9):

$$x_l = g(f(3, 3, x_{l-1}), f(1, 1, x_{l-1})) \quad (9)$$

其中: $l = (3, 4, 5)$, $l = 3$ 时卷积层的输入为 x_{p2} 。

并行卷积通道的存在使网络使用了更多的卷积核, 导致网络参数量增加, 不利于网络在资源有限的终端设备上部署, 为在保证感受野丰富度的条件下降低网络的参数量, 该研究使用多个连续的 3×3 卷积核替代较大卷积核。例如在图 5 中, 若再次使用 3×3 的卷积核对图 5 (a) 输出的特征图进行卷积计算, 其输出维度将降低至 1×1 , 与图 5 (b) 的输出特征图维度相同, 且最终输出值与输入图像的 5×5 维数据相关, 因此连续两个 3×3 卷积核叠加后感受野与 5×5 卷积核的感受野大小相同。而两个 3×3 卷积核的参数量为 $3 \times 3 \times 2 = 18$, 一个 5×5 卷积核的参数量为 $5 \times 5 \times 1 = 25$, 替换后参数量可降低 28%, 表明使用多个小卷积核替代大卷积核能够在保持卷积感受野不变的前提下降低网络参数量的目标。

2.1.2 “捷径”连接

传统 Alex Net 的卷积层仅与上下层连接, 逐层的卷积操作将浅层卷积所提取的边缘纹理等低层次特征逐渐映射为抽象的高层次特征, 最后使用这些高层次特征进行分类。这种分类方法没有考虑到低层次特征对分类结果的影响, 用于分类的特征信息缺乏完整性; 同时, 网络深度的增加也使网络在训练过程中易出现梯度消失或梯度爆炸问题, 导致网络的训练难度增大。

K. M. He 等^[17]在网络中引入了“捷径”连接方式, 即将本层网络的输出通过恒等映射等方式与更深层网络的输出相融合, 使深层网络的输出特征中存在浅层特征的映射。前向传播时, 这种连接方式能够直接将低层次特征映射到深层网络所提取的高层次特征中, 提高了特征信息的完整性; 反向传播时, 深层网络的梯度也可以通过“捷径”连接直接传递到浅层, 可以有效避免梯度消失和梯度爆炸等问题。与传统网络相比, 这种连接方式有效利用了对边缘纹理等信息保存较好的低层次特征, 使得网络输出的高层次特征中的边缘和细节信息得到一定程度的增强, 网络的输出特征也更加完整。

增加“捷径”连接后卷积层的输入输出关系见式 (10):

$$x_l = f(w, h, x_{l-1}) + x_{l-1} \quad (10)$$

与式 (2) 相比, 式 (10) 增加了上一层输出的恒等映射, 在该层的输出中可以在一定程度上保留上层信息。

改进后的 Alex Net 网络虽然增加了并行卷积通道, 但依旧仅包含 5 个卷积层, 网络层数较浅, 特征图前向传播

和误差反向传递过程较为简单, 因此在改进网络中设置“捷径”连接时主要考虑了前向传播过程中低层次特征向深层网络的映射问题。改进后的 Alex Net 前 2 层卷积感受野尺寸较为丰富, 所提取特征包含不同尺度下的边缘、纹理等低层次特征; 后面 3 层卷积结构较为单一, 功能上主要负责将上层输出映射为更高层次的特征。因此改进网络中通过“捷径”连接方式和“与”操作将改进 Alex Net 的第一层、第二层网络的输出映射至第五层网络的输出, 实现特征在不同尺度上的融合, 增加特征信息的完整性。补充恒等映射后的网络输出 x_{Res5} 与第五层卷积的输出特征图 x_5 之间的关系式见式 (11):

$$x_{\text{Res5}} = \text{Add}(x_{p1}, x_{p2}, x_5) \quad (11)$$

其中: Add 表示“与”操作, 实现低层次特征与高层次特征在相同维度内的融合。

改进 Alex Net 的网络结构见图 7, 参数设置见表 3。

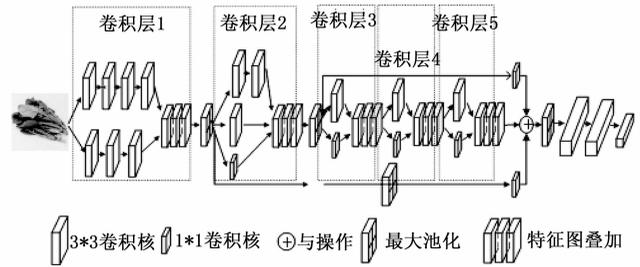


图 7 改进 Alex Net 网络结构

2.2 时序程序设计

时序主程序负责将系统的各组成部分按一定顺序串联, 其设计合理与否直接决定系统是否能够有效利用结算过程中产生的各种信息自动运行, 并尽可能减少冗余信息, 同时避免组成部分间的时序冲突, 是系统能否稳定有效运行的基础。功能上主要包括监测重量信号、获取商品图像、识别商品类别、确认结算完成等部分, 软件的具体实现流程见图 8。

使用该系统结算时, 首先需要商家在消费者结算前启动主程序, 预加载识别算法模型及参数, 之后主程序进入待机状态, 时刻监测 GPIO. 3 接口电平信号的变化情况。

消费者结算时, 将商品放置于称重托盘, 托盘下方的压力传感器受压后输出模拟电量信号, 经 HX711 数模转换芯片转换为数字信号, 通过 GPIO. 2、GPIO. 3 接口输出, 输出时 GPIO. 3 接口首先由高电平状态降低至低电平, 主机确认接收之后 HX711 芯片的 SCK 接口发送存储重量信息的数字时钟信号至树莓派。

树莓派主机检测到 GPIO. 3 接口的电平变化后, 首先开始接收 GPIO. 2 接口的重量信息, 之后打开摄像头, 采集待结算商品的图像, 并使用 Image 函数库将图像尺寸调整为 $224 \times 224 \times 3$ 。之后将图像输入到果蔬识别算法, 经处理器运算得到商品最大概率的所属类别, 融合商品重量和类别信息计算出商品总价并在显示器上展示相关信息。

表 3 改进 Alex Net 参数设置

网络层	网络层输入尺寸	卷积规格 (长×宽×通道)	池化规格	步长	填充模式	网络层输出尺寸
输入层	224×224×3	—	—	—	—	—
卷积层 1 分支 1.1	224×224×3	3×3×64	—	2	—	—
卷积层 1 分支 1.2	—	3×3×64	—	2	0 填充	—
卷积层 1 分支 1.3	—	3×3×64	—	1	—	—
卷积层 1 分支 1.4	—	3×3×64	—	1	0 填充	55×55×64
卷积层 1 分支 2.1	224×224×3	3×3×32	—	2	—	—
卷积层 1 分支 2.2	—	3×3×32	—	2	0 填充	—
卷积层 1 分支 2.3	—	3×3×32	—	1	—	55×55×32
卷积层 1 输出叠加层	—	—	—	—	—	55×55×96
最大池化层 1	55×55×96	—	3×3	2	—	27×27×96
卷积层 2 分支 1.1	27×27×96	3×3×160	—	1	0 填充	—
卷积层 2 分支 1.2	—	3×3×160	—	1	0 填充	27×27×160
卷积层 2 分支 2.1	27×27×96	3×3×64	—	1	0 填充	27×27×64
卷积层 2 分支 3.1	27×27×96	1×1×32	—	1	0 填充	27×27×32
卷积层 2 输出叠加层	—	—	—	—	—	27×27×256
最大池化层 2	27×27×256	—	3×3	2	—	13×13×256
卷积层 3 分支 1.1	13×13×256	3×3×320	—	1	0 填充	13×13×320
卷积层 3 分支 2.1	13×13×256	1×1×64	—	1	0 填充	13×13×64
卷积层 3 输出叠加层	—	—	—	—	—	13×13×384
卷积层 4 分支 1.1	13×13×384	3×3×320	—	1	0 填充	13×13×320
卷积层 4 分支 2.1	13×13×384	1×1×64	—	1	0 填充	13×13×64
卷积层 4 输出叠加层	—	—	—	—	—	13×13×384
卷积层 5 分支 1.1	13×13×384	3×3×192	—	1	0 填充	13×13×192
卷积层 5 分支 2.1	13×13×384	1×1×64	—	1	0 填充	13×13×64
卷积层 5 输出叠加层	—	—	—	—	—	13×13×256
与操作输入支路 1 池化层	27×27×96	—	3×3	2	—	13×13×96
与操作输入支路 1 卷积层	—	1×1×156	—	1	0 填充	13×13×256
与操作输入支路 2	13×13×256	1×1×256	—	1	0 填充	13×13×256
与操作输入支路 3	13×13×256	—	—	—	—	—
与操作	—	—	—	—	—	13×13×256
最大池化层 3	13×13×256	—	3×3	2	—	6×6×256
全连接层 1	6×6×256	—	—	—	—	4096×1
全连接层 2	4096×1	—	—	—	—	4096×1
输出层	4096×1	—	—	—	—	1000×1

表中：“—”表示无直接输入、无卷积运算、无池化运算、无图像填充等。

若商品类别及单价信息无误，消费者结算后主程序返回待机状态，等待下次结算；若商品信息有误，可通过管理人员修改商品信息并结算，之后程序待机，等待下次结算。

使用该主程序完成果蔬自助结算系统设计，系统运行时仅最初启动时加载一次识别算法的模型和参数，后续结算时可直接调用识别，大大减少了识别过程的时长；消费者结算时仅需将果蔬等商品放置于称重托盘，之后等待系统出示结算信息，极大程度上简化了消费者的结算流程，与当前超市普遍使用的结算方式相比更加方便快捷。

3 实验结果与分析

3.1 果蔬数据集构建

搭建果蔬自助结算系统，除了要设计基于图像的果蔬识别方法之外，还需要有相应的图像数据集训练识别算法。训练集数据与实际情况的相符程度对识别方法的精度等有着重要影响，但目前关于果蔬识别的研究较少，国内外相应的数据库还未建成，因此研究过程中需设计图像采集方案建立相应超市果蔬数据集。对超市常见果蔬种类进行调研，筛选出 24 类常见果蔬（包括特征相似的国产香蕉、进

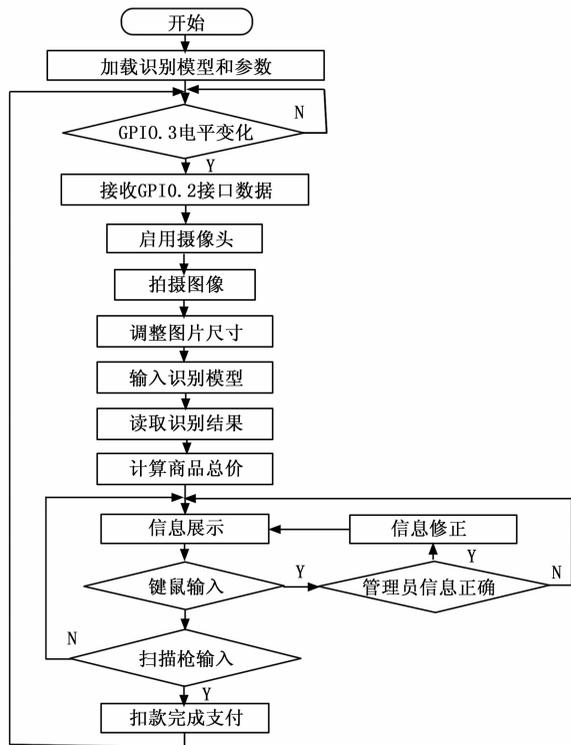


图 8 主程序流程图

口香蕉等类别) 作为采集目标, 使用模拟实际场景搭建的硬件设备采集果蔬图像以增强数据集图像与实际应用场景的相符程度。

实际应用过程中结算环境的光照等外部条件的变化以及结算时果蔬在托盘上摆放方式的不同等可能会影响果蔬识别的准确率, 因此在采集果蔬图像时设置了正常光照、强光照、弱光照 3 种光照条件, 分别在 3 种条件下采集果蔬不同摆放角度的图像, 预训练后建立数据集。

采集时使用额外光源、遮光窗帘等设备调节现场的光照强度, 使用曲别针等制作简易小型支架, 使果蔬可以处于特定倾斜角度。采集时将果蔬放置于托盘固定位置, 每旋转一定角度采集一次图像, 之后使用支架使果蔬处于一定倾斜角度, 再次旋转采集图像, 直至充分采集果蔬任意角度下的图像。最终采集到图像共 43 989 张, 各光照条件下的图像数量在 14 000 张以上, 相同光照条件下各类果蔬的图像均在 500 张以上。

拍摄的原始图像尺寸较大, 存在较多无关信息, 会影响网络的特征提取效果, 需要通过图像预处理方法将目标图像与无关信息分割开, 即实现图像前景区域和背景部分的分割。其分割过程如图所示。首先将原始图像转换为灰度图像, 使用自适应阈值分割法^[20]将图像的前景与背景分割开, 之后使用填充处理操作得到前景图像, 见图 9 (a)、图 9 (b) 和图 9 (c)。发现由于光照等条件的影响, 图像的部分前景区域呈现出与背景区域相近的灰度, 导致分割后的前景图像不完整, 因此需要使用形态学膨胀处理方法进一步处理, 之后再使用填充处理操作得到完整的前景图像,

见图 9 (d) 和图 9 (e)。

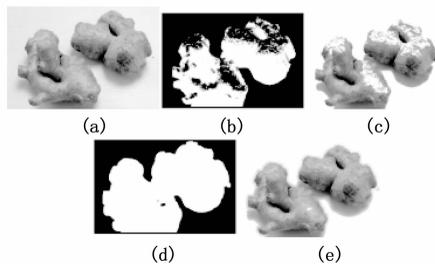


图 9 数据预处理过程图解

3.2 模型训练与对比

为测试改进 Alex Net 的效果, 设计了模型训练和测试方法, 使用超市果蔬数据集分别训练改进 Alex Net、Inception V3 和 Res Net, 并记录训练和测试过程的指标参数。实验平台使用 Windows10 家庭中文版操作系统, 中央处理器为 Intel Core i5-8300H。

首先将数据集按照 7 : 2 : 1 的比例划分为训练集、验证集和测试集; 设置初始学习率、最大迭代次数等超参数, 使用训练集数据训练模型参数。

之后依照交叉熵值最小化原则调节模型参数。模型参数保存条件为保留最近一次交叉熵值降低时的模型参数。训练过程超参数的调节依据为交叉熵值连续未降低的迭代次数, 若少于 3 次则保持超参数不变, 直接开始下次迭代; 大于或等于 3 次且小于 7 次则将学习率调整至当前学习率的 0.5 倍; 连续 7 次未降低时提前结束训练过程, 保存最后一次交叉熵值降低时的模型参数。3 种网络的训练过程如图 10 所示。

最后重新载入模型, 测试参数加载时间, 输入测试集图像测试模型的总体平均识别准确率和单张图片识别用时。测试结果见表 4。

表 4 模型对比实验结果

模型	交叉熵值	Top_1 准确率/%	参数量/KB	模型加载用时/s	单张图片识别速度/s
Inception V3	0.136 2	97.35	85 280	4.62	3.71
Res Net	0.019 7	98.14	92 738	3.95	2.69
改进 Alex Net	0.074 5	97.68	16 168	2.74	2.48

对比图 10 准确率和交叉熵值的变化过程发现, 虽然三种网络最终的识别准确率接近, 交叉熵值也很低, 但 Inception V3 和 Res Net 训练过程中存在训练集准确率和交叉熵值平滑过渡但验证集准确率和交叉熵值大幅变化的情况, 认为在训练 Inception V3 和 Res Net 时网络模型出现了过拟合现象, 表明在当前应用情形下模型复杂度高于实际所需, 模型计算过程存在冗余的计算空间而部署果蔬识别方法的终端设备的处理器性能低于训练平台, 运行深层网络存在一定难度, 因此训练过程较为平滑的改进 Alex Net 更为适合果蔬自助结算系统。结合表 4 测试结果发现改进 Alex

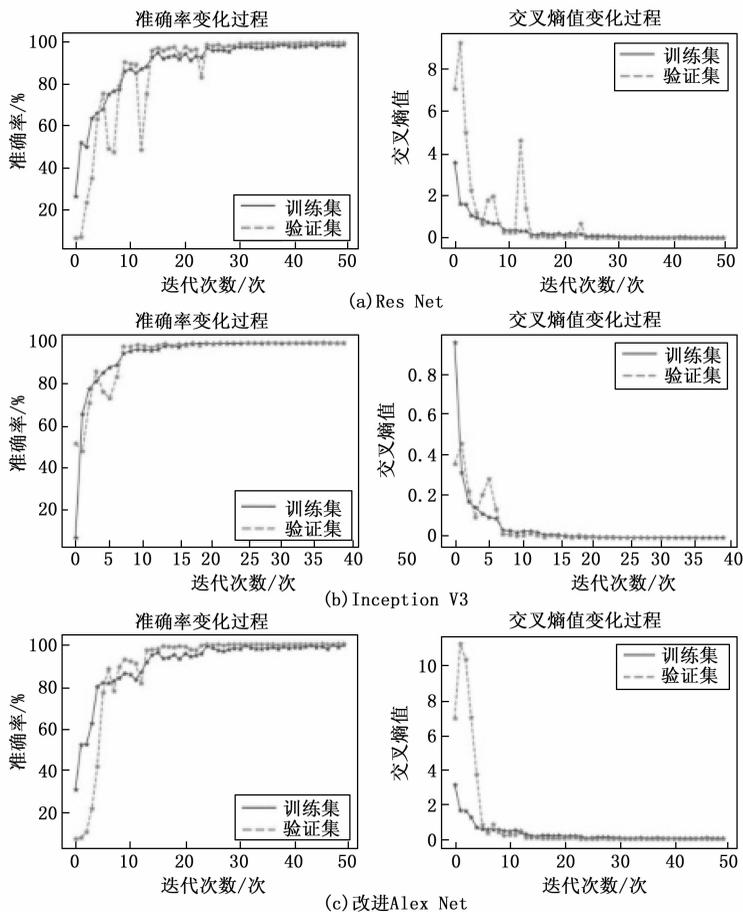


图 10 网络训练过程

Net 的平均识别精度与 Inception V3 和 Res Net 接近, 但参数量规模更小, 模型加载和识别也更为快速。因此, 基于改进 Alex Net 的果蔬识别方法更适合用于解决超市自助结账系统面临的果蔬识别问题。

3.3 系统测试结果与分析

经测试认为改进的果蔬识别方法能够在资源需求更小的前提下保持相对较高的识别精度, 符合果蔬自助结账系统的使用条件, 因此将该方法部署到果蔬自助结账系统的树莓派中, 运行主程序测试系统单次的结算时长以及对各类果蔬的识别准确率。系统的运行环境如图 11 所示。

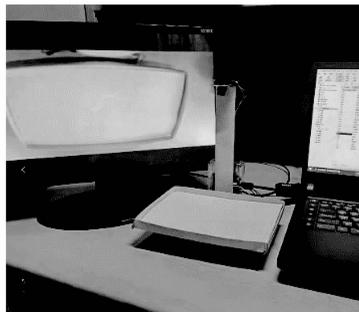


图 11 系统运行环境

系统测试步骤如下:

1) 配置程序运行环境, 程序进入待机状态。接通电源, 打开树莓派, 启动自助结账主程序。程序初次启动时需要加载 TensorFlow 深度学习框架、改进果蔬识别方法的模型参数、OpenCV 函数库和 WiringPi 软件包, 作用分别是构建识别模型、加载模型参数、摄像头调用和图片处理、GPIO 接口功能配置与实现。运行环境配置完成后程序进入循环部分, 即待机状态, 此时程序通过监测 GPIO. 3 接口的电平状态判断程序是否进入识别结算流程。

2) 读取商品重量信息。消费者将待结算商品放置到称重托盘, 压力传感器受压后输出商品重量信息, 此时 GPIO. 3 接口由高电平变为低电平, 程序进入循环的识别结算部分。之后读取 GPIO. 2 接口的数字时钟信号, 换算后得到商品的重量值, 单位为 g, 精度可达 0.000 1, 考虑到实际应用场景, 计算后的重量数据仅保留到克。

3) 采集待结算商品的图像。GPIO. 2 接口的数字时钟信号接收完成后, 使用预配置的 OpenCV 函数库调用摄像头模块, 保存当前图像并将图像尺寸处理为识别方法所需尺寸维度。

4) 识别商品类别。将采集图像输入果蔬识别算法模型, 经特征提取与计算后得到待结算商品的类别。

5) 计算商品总价, 展示结算信息。根据识别结果读取该类商品的单价等信息, 与商品重量相乘得到商品总价, 通过 Image 函数库的相关函数将所采集的商品图像、识别结果、商品编号及单价总价等信息在显示器上显示, 见图 12。



图 12 单次运行结果

最后一步为信息确认与结算。这一阶段程序读取扫描枪和键盘鼠标的信息。若商品信息无误, 消费者出示付款码, 扫描枪扫描后将结果返回树莓派, 程序返回循环顶端, 即重新进入待机状态; 若信息有误可使用键盘鼠标等交互硬件选择取消或联系管理员修改商品信息, 修改信息前需输入正确的管理员信息, 修改完成后消费者可通过扫描枪完成支付。

测试过程中系统启动后可自动运行, 需要人工操作的部分仅有放置待结算商品和出示付款码两部分, 与现有自

助结算设备相比结算流程更简单;在实际应用过程中,系统完成结算的总用时约 7.48 秒,仅为人工结算所需时长的 $1/4^{[21]}$ 。

调整环境光照条件和果蔬摆放角度,重复上述实验,测得系统在实际应用过程中对各类果蔬的识别率及全部类别的平均识别率,如表 5 所示。

表 5 各类果蔬的识别率及数据集平均识别率

类别	准确率/%	类别	准确率/%
红富士	99.71	菠菜	93.86
红星	98.83	花菜	95.77
金冠苹果	99.91	卷心菜	97.18
国产香蕉	99.52	西红柿	99.93
进口香蕉	96.08	土豆	97.85
白洋葱	97.84	火龙果	99.89
紫洋葱	95.78	牛油果	98.33
红洋葱	99.95	猕猴桃	99.95
白鸡蛋	98.47	柠檬	99.91
鸡蛋	94.99	草莓	99.47
红辣椒	99.89	山竹	90.49
绿辣椒	98.72	姜	97.03
总体平均识别准确率/%		98.25	

从表 5 中可以发现,网络模型对国产香蕉、进口香蕉等特征相似度较高的果蔬均取得了较高的识别准确率,证明网络能够精确区分差异较小的果蔬品种;网络的平均识别精度可以达到 98.25%,表明网络模型能在实际应用中取得较好结果。

4 结束语

本文提出了一种基于改进 Alex Net 的果蔬自动识别方法并设计了果蔬自助结算系统。通过增加卷积核尺寸不同的卷积通道和使用“捷径”连接两种方式改进网络结构,增加了网络提取特征的丰富性和完整性,优化了特征的表征能力;使用 3×3 卷积核替代较大的卷积核,进一步降低模型的参数量,使模型在资源有限的终端应用时加载、响应时间更短。搭建硬件平台对果蔬自助结算系统进行实现,在实际运行环境下测试了系统性能,测得系统的整体识别率高于 95%,除需要加载识别模型及参数的启动时间之外系统响应时间不超过 8 s。表明该方法及系统与人工结算相比更加方便快捷,与使用条形码的自助结算设备相比具有更加广阔的应用空间以及发展前景。

参考文献:

[1] 虞玲. 基于计算机视觉的果蔬自动分类技术中的特征分析的研究 [D]. 南京:东南大学, 2015.

[2] 翟雅萌, 刘晓东, 陈智元, 等. 基于图像处理的智能无人收银服务系统设计 [J]. 计算机与数字工程, 2017, 45 (10): 2037-2040, 2062.

[3] 罗承成, 李书琴, 唐晶磊. 基于多示例学习的超市农产品图像识别 [J]. 计算机应用, 2012, 32 (6): 1560-1562, 1566.

[4] ARIVAZHAGAN S, SHEBIAH R N, NIDHYANANDHAN S S, et al. Fruit recognition using color and texture features [J]. Journal of Emerging Trends in Computing and Information Sciences, 2010, 1 (2): 90-94.

[5] DUBEY S R, JALAL A S. Robust approach for fruit and vegetable classification [J]. Procedia Engineering, 2012, 38.

[6] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [C] // Advances in Neural Information Processing Systems, 2012: 1097-1105.

[7] 邱津怡, 罗俊, 李秀, 等. 基于卷积神经网络的多尺度葡萄图像识别方法 [J]. 计算机应用, 2019, 39 (10): 2930-2936.

[8] 刘芾, 李茂军, 胡建文, 等. 基于低像素人脸图像的表情识别 [J]. 激光与光电子学进展, 2020, 57 (10): 89-96.

[9] 吴晶晶, 程武山. 融合全局和局部特征并基于神经网络的表情识别 [J]. 计算机测量与控制, 2018, 26 (6): 172-175.

[10] 刘洋, 冯全, 王书志. 基于轻量级 CNN 的植物病害识别方法及移动端应用 [J]. 农业工程学报, 2019, 35 (17): 194-204.

[11] 李晓琳, 庞保孟, 曹银杰, 等. 基于微控制器改进 Squeeze-Net 交通标志识别的研究 [J]. 计算机测量与控制, 2020, 28 (5): 88-92.

[12] 屈洽华, 邵毅明, 邓天民, 等. 复杂光照条件下的交通标志检测与识别 [J]. 激光与光电子学进展, 2019, 56 (23): 133-140.

[13] 巨志勇, 马素萍. 改进的 Inceptionv3 果蔬识别算法 [J]. 包装工程, 2019, 40 (21): 30-35.

[14] 樊帅昌, 易晓梅, 李剑, 等. 基于深度残差网络与迁移学习的毒莓图像识别 [J]. 传感技术学报, 2020, 33 (1): 74-83.

[15] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1-9.

[16] SZEGEDY C, VANHOUCHE V, IOFFE S, et al. Rethinking the inception architecture for computer vision [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2818-2826.

[17] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.

[18] 罗顺元, 李志强, 马文颢. 基于树莓派的无人值守信息处理系统 [J]. 计算机测量与控制, 2018, 26 (9): 227-231.

[19] 谢照, 吴正平. 基于样条函数的高精度电子秤设计 [J]. 计算机测量与控制, 2017, 25 (6): 167-169, 173.

[20] 周向阳, 罗雪梅, 王霄. 结合边缘信息的二维 Otsu 阈值分割算法研究 [J]. 智能计算机与应用, 2020, 10 (6): 19-24, 30.

[21] 陈登峰, 周瑶, 段优, 等. 基于计算机视觉的智能果蔬结算系统 [J]. 信息化研究, 2019, 45 (2): 65-70.