

自然场景下基于改进 LeNet 卷积神经网络的苹果图像识别技术

Research on apple image recognition technology based on improved LeNet convolution neural network in natural scene

程鸿芳¹

张春友²

CHENG Hong-fang¹ ZHANG Chun-you²

(1. 芜湖职业技术学院, 安徽 芜湖 241000; 2. 内蒙古民族大学机械工程学院, 内蒙古 通辽 028043)

(1. Wuhu Institute of Technology, Wuhu, Anhui 241000, China; 2. College of Mechanical Engineering, Inner Mongolia University for the Nationalities, Tongliao, Inner Mongolia 028043, China)

摘要:针对传统基于内容的识别方法在特征提取方面存在计算复杂、特征不可迁移等问题,为避免光照条件、重叠及其他遮挡等因素对图像识别的影响,以 LeNet 卷积神经网络作为基础,对其结构进行改进,设计了一种基于改进 LeNet 卷积神经网络的苹果目标识别模型,并利用该模型对不同场景的苹果图像进行识别训练与验证。结果表明:该网络模型可有效实现苹果图像的识别,对独立果实、遮挡果实、重叠果实以及相邻果实的识别率分别为 96.25%, 91.37%, 94.91%, 89.56%, 综合识别率达到 93.79%。与其他方法相比,该算法具有较强的抗干扰能力,图像识别速度快、识别率更高。

关键词:图像识别; 目标识别; 卷积神经网络; LeNet

Abstract: In order to avoid the influence of illumination condition, overlap and other occlusion on image recognition, an improved LeNet convolution neural network is used to improve the structure of the traditional content-based recognition method. An Apple target recognition model based on the improved LeNet convolution neural network is designed and used to avoid the influence of illumination condition, overlap and other occlusion factors on image recognition. The model trains and validates Apple images in different scenarios. The results show that the network model can effectively recognize apple images. The recognition

rates of independent fruits, occluded fruits, overlapping fruits and adjacent fruits are 96.25%, 91.37%, 94.91% and 89.56% respectively, and the comprehensive recognition rate is 93.79%. Compared with other methods, this algorithm has stronger anti-jamming ability, faster image recognition speed and higher recognition rate.

Keywords: image recognition; target recognition; convolution neural network; LeNet

随着中国苹果产业不断向机械化、规模化方向发展,利用图像识别技术对果实进行识别、监测及自动化管理是解决目前人工作业效率低,劳动强度大问题的关键^[1-2]。为此,学者们对其进行了深入研究,并取得了较大进展。钱建平等^[3]以图斑为特征变量进行苹果图像特征提取,并建立产量估测模型计算单株果树的产量。许立兵等^[4]针对传统苹果分级检测存在的问题,设计了一种基于 ARM 的苹果采后田间分级检测系统,采用基于 Haar-like 特性的级联 Adaboost 目标检测算法,对苹果果实的大小和缺陷进行识别,实现了苹果的田间预分级检验。田有文等^[5]利用高光谱成像技术提取苹果感兴趣区域光谱信息,对果梗的完整性进行检测,实现了对苹果品质的等级评价。

但是,随着机器视觉研究的发展,数据量剧增,传统基于果实颜色、形状或纹理特征等内容的识别方法,在特征提取方面存在的计算复杂、特征不可迁移等问题愈发突出。特别是对复杂背景环境、光照条件、果实重叠及其他遮挡等情况的图像识别上,存在识别准确率差和运算效率低的问题,不能满足实际需求。近年来,卷积神经网络(Convolutional Neural Networks, CNN)开始逐渐应用于目标识别与定位领域。周云成等^[6]在 VGGNet 基础

基金项目: 2017 年安徽省高校科学研究项目(编号: KJ2017A560); 2017 年高校优秀青年人才支持计划项目(编号: gxyqZD2017141); 2016 年安徽省质量工程项目(编号: 2016ckjh224); 芜湖职业技术学院科技创新团队(编号: Wzykj2018A02)

作者简介: 程鸿芳,女,芜湖职业技术学院副教授,硕士。

通信作者: 张春友(1980—),男,内蒙古民族大学副教授,博士。

E-mail: zcyimun@163.com

收稿日期: 2019-01-03

上,设计了一种基于深度卷积神经网络的番茄主要器官检测器,在检测速度和精度上优于同类算法。马永杰等^[7]提出了改进深度卷积神经网络 AlexNet 的分类识别算法模型,对交通标志进行识别,结果达到预期。龙满生等^[8]利用卷积神经网络强大的特征学习和特征表达能力,对油茶病害特征进行提取,并借助迁移学习方法,将 AlexNet 模型在 ImageNet 图像数据集上,学习得到的知识迁移到油茶病害识别任务,从而提高了图像的分类精度。

可见,卷积神经网络可直接将原始图像作为输入,避免传统识别算法中繁琐的数据处理过程,实现图像特征的自动提取,并且能进行自我学习,对图像具有极强的数据表征能力,是一种高效的识别方法,经过检索尚未发现有其他研究人员将此技术用于复杂背景环境、光照条件、苹果果实重叠及其他遮挡等情况的图像识别上。本试验拟以 LeNet 卷积神经网络作为基础,对其池化方法、激活函数等参数结构进行改进,设计一种基于改进 LeNet 卷积神经网络的苹果目标识别模型,以此模型为基础,深入分析研究卷积神经网络对果实目标的识别效果。同时,对照其他研究方法在自然场景下对不同位置形态的苹果果实进行识别训练及测试验证,并对果实重叠及其他遮挡等复杂情况进行重点分析,以期建立一种抗干扰能力强,识别成功率高,识别速度快,对枝叶遮挡、光线变化等影响均具有相对较好鲁棒性的方法。

1 卷积神经网络

CNN 模型主要包括输入层、卷积层、池化层、全连接层和输出层。其中卷积层和池化层是模型核心部分,主要用于图像特征的提取,全连接层能够实现图像从高维特征域到图像类别域的映射^[9-10]。

1.1 卷积层

卷积层包含多个卷积核,通过卷积计算得到输入图像的特征图,其表达式为^[11-13]:

$$x_j^L = f(\sum_c X_c^{L-1} * W_{ij}^L + b_j^L), \quad (1)$$

式中:

x_j^L ——第 L 卷积层中的第 j 个特征图;

W_{ij}^L ——卷积核;

$f(\cdot)$ ——激活函数;

b_j^L ——偏置项。

1.2 池化层

池化层又叫下采样层,本质上是图像的一种聚合操作,用于对原始的特征信息进行提取,减小特征维数和数据的空间大小,可以在一定程度上控制模型的过拟合现象,其表达式为:

$$x_j^L = f[\beta_j^L \text{down}(x_j^{L-1}) + b_j^L], \quad (2)$$

式中:

$\text{down}(\cdot)$ ——下采样函数;

β_j^L ——采样系数。

2 试验数据

2.1 图像采集

分别在阴天和晴天的顺光和逆光条件下,在某苹果园自然场景下采集苹果图像,以保证样本的多样性。共采集原始图像 600 幅(晴天和阴天各 300 幅),格式为 JPEG,分辨率为 2 992 像素×2 992 像素。

2.2 样本处理与试验方法

对图像进行预处理,可以增强数据样本,减少后续试验运行时间,提高模型性能。包括:① 对数据集进行数据增强处理,以更好地提取图像特征,避免出现过拟合;② 图像进行空间尺度变换,即将图像进行随机缩放、平移、旋转、镜像、随机裁剪;③ 进行图像均值与方差均衡,即样本归一化。预处理后图像尺寸为 32 像素×32 像素。处理中,共得到苹果图像 2 435 幅,在不同天气情况下随机选取 487 幅图像作为测试,其他图像用于训练,二者之间无重叠。表 1 为测试集样本信息。

表 1 测试集样本数量

Table 1 Test set sample size

晴天		阴天	
顺光	逆光	顺光	逆光
136	118	127	106

3 基于改进 LeNet 卷积神经网络的苹果图像识别

3.1 卷积神经网络的构建

以 LeNet 作为基础网络架构在 Matlab 中建立卷积神经网络模型,并对其重要的结构参数进行改进,为了更好地提取图像的低阶特征和有效的传递误差梯度,采用 Re LU 作为激活函数。在池化操作中使用最大池化方法代替平均池化方法,以降低特征提取时因估计均值偏移所引起的误差,如图 1 所示。

为保证所建立的神经网络模型具有较好的性能,综合考虑模型结构参数,包括 3 个卷积层 C1、C3 和 C5,2 个

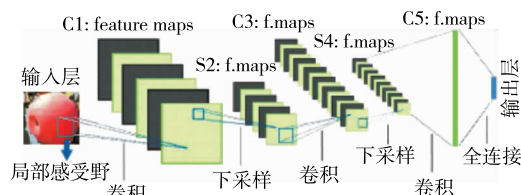


图 1 基于改进 LeNet 的苹果图像识别网络结构

Figure 1 Apple Image Recognition Network Structure Based on Improved LeNet

池化层 S2 和 S4,1 个全连接层 F6,1 个输出层。

输入到第 1 卷积层 C1 的是大小为 32×32 的图像,取 6 个尺寸为 5×5 的卷积核,卷积运算后得到 6 个特征图,C1 层中包含的训练参数为 $6 \times (5 \times 5 + 1)$,即得到的特征图大小为 28×28 ;第 1 池化层 S2 中包含 6 个大小为 14×14 的特征图,与 C1 中所对应的大小为 2×2 特征图的领域相连接;第 2 卷积层 C3 中采用 16 个大小为 5×5 的卷积核,计算方法同 C1 类似,得到 16 个 10×10 大小的特征图;第 2 池化层 S4 采用大小为 2×2 的采样核,特征图个数为 16 个,经过池化操作后输出 16 个 5×5 大小的特征图;第 3 卷积层 C5 中每个神经元与 S4 中的 5×5 邻域相连接,输出 120 个 1×1 的特征图,使池化层 S4 与卷积层 C5 之间构成全连接;全连接层 F6 中拥有 84 个隐含节点,与卷积层 C5 构成全连接;最后在输出层中输出分类结果。各层具体参数如表 2 所示。

表 2 LeNet 神经网络各层参数

Table 2 LeNet neural network parameters

层	运算方式	图像大小	特征图数量
C1 卷积层	5×5	28×28	6
S2 池化层	2×2	14×14	6
C3 卷积层	5×5	10×10	16
S4 池化层	2×2	5×5	16
C5 卷积层	5×5	1×1	120
F6 全连接层	1×1	1×1	84
输出层	Softmax 分类	1×1	10

3.2 图像识别步骤

采用所提的改进 LeNet 卷积神经网络模型对苹果图像进行识别,具体方法参照文献[14]。

- (1) 对预处理后的图像进行筛选,达到训练要求。
- (2) 随机选取图像,建立数据集,初始化网络结构得到卷积核的权值。
- (3) 将图像在 C1 卷积层中进行卷积计算,获得相应的特征图。
- (4) 利用最大池化法将特征图在 S2 池化层中进行下采样。
- (5) 重复步骤(3)和(4),将得到的特征图在 C3 卷积层和 S4 池化层中进行 2 次卷积和采样,获得新的特征图。
- (6) 重复步骤(3)和(4),对输出的特征图进行 3 次卷积。
- (7) 将最后得到的特征图转换成列向量,作为全连接层的输入,通过反向传播算法更新网络参数;
- (8) 利用训练得到的网络模型对测试图像进行分类,通过 Softmax 分类器结合多尺度滑动窗算法显示识别结

果。按照图像中果实的相对位置将果实分为 4 类:独立果实、遮挡果实、重叠果实以及相邻果实。

4 结果与分析

4.1 模型训练

利用训练集样本进行训练,取网络初始权重的初始标准差为 0.01,均值为 0 的高斯分布,设样本迭代次数为 2 000 次,权重参数的初始学习速率为 0.001,动量因子为 0.9。训练结果如图 2 所示。

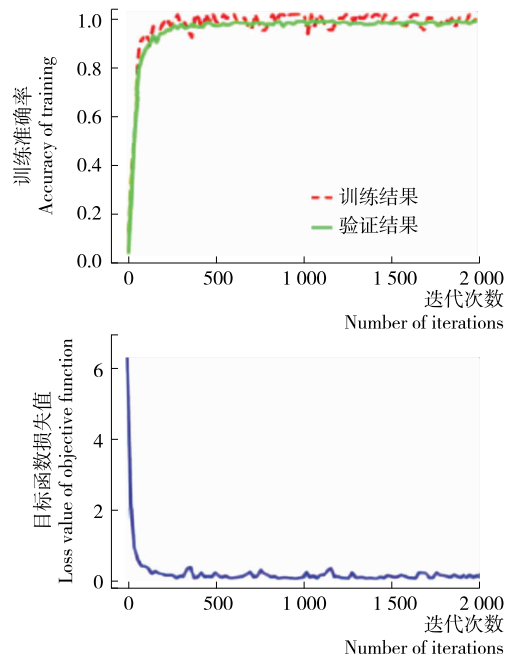


图 2 CNN 训练准确率与损失值变化曲线

Figure 2 CNN training accuracy and loss value curve

由图 2 仿真结果可以看出,所设计算法的训练准确率随着迭代次数的增加而迅速升高,且趋于稳定,训练集和验证集两者的分类结果非常接近,准确率达到 98.5%。目标函数的损失值迅速降低,在迭代 2 000 次时,目标函数的损失值收敛在 0.05 左右,基本保持稳定,表明所设计的网络模型是有效的,达到了预期的训练效果。

4.2 模型验证

网络的最终输出为识别到苹果目标的概率,为了验证改进 LeNet 卷积神经网络模型的有效性,对测试集的图像进行识别。在测试集中随机选取 150 幅图像,阴天和晴天各 75 幅,共有果实图像 4 252 个。选用识别目标与真实目标重合的比率重叠系数作为试验结果有效性的评价指标,网络的最终输出为识别到的苹果目标的概率,结果仅保留重叠系数 $> 75\%$ 的区域,即当重叠系数 $> 75\%$ 时,认为识别准确,识别结果示例和具体识别结果分别如表 3 所示。

由表 3 可知,独立果实、重叠果实、相邻果实以及遮

表 3 果实识别具体结果

Table 3 Fruit identification specific results

果实	实际果实数/个	正确识别的果实数/个	识别率/%
独立果实	961	925	96.25
重叠果实	510	466	91.37
相邻果实	1 986	1 885	94.91
遮挡果实	795	712	89.56
总计	4 252	3 988	93.79

挡果实的识别率分别为 96.25%，91.37%，94.91%，89.56%，综合识别率达到 93.79%。其中因遮挡、重叠以及光线等原因,使果实轮廓不全或不清晰的果实,容易出现漏识别现象,从而影响了整体的识别精度。但整体来看,所设计的卷积神经网络模型具有很好的识别效果。

4.3 与其他方法对比分析

为进一步验证算法的优越性,将本算法与基于 ResNet-44 的 R-FCN 网络、基于 ResNet-101 的 R-FCN 网络以及基于 ResNet-50 的 R-FCN 网络进行对比^[15],结果如表 4 所示。与其他网络算法相比,本文设计的基于改进 LeNet 卷积神经网络的苹果识别方法具有较强的抗干扰能力,识别成功率高,识别速度快,对枝叶遮挡、光线变化等影响均具有相对较好的鲁棒性。

表 4 不同网络识别效果对比

Table 4 Comparison of different network recognition effects

算法	识别成功率/%	单个果实识别时间/s
R-FCN(ResNet-50)	89.50	0.197
R-FCN (ResNet-101)	89.50	0.228
R-FCN(ResNet-44)	90.20	0.187
修正算法	93.79	0.183

5 结论

针对传统图像识别方法在特征提取方面存在计算复杂、特征不可迁移等问题,为避免光照条件、重叠及其他遮挡等因素给图像识别带来的影响,利用卷积神经网络图像特征提取能力强的特点,提出基于卷积神经网络的苹果图像识别方法。以 LeNet 卷积神经网络作为基础网络架构,并对其结构和参数进行优化和改进,以更好地对图像特征进行提取。同时,利用该模型对不同果实位置差别的数据集进行训练、验证和结果分析,并与同类文献进行识别精度对比。分析结果表明:本试验所建立的网络模型可有效实现苹果图像的识别,对独立果实、重叠果实、相邻果实以及遮挡果实的识别率分别为 96.25%，91.37%，94.91%，89.56%，综合识别率达到 93.79%。与

其他方法相比,本算法果实识别速度快、识别率更高,尤其对重叠及其他遮挡的果实图像具有较好的识别能力。

参考文献

[1] 夏雪,周国民,丘耘,等. 自然环境下苹果作业机器人双果重叠目标侦测方法[J]. 中国农业科技导报, 2018, 20(7): 63-73.

[2] 贾伟宽. 基于智能优化的苹果采摘机器人目标识别研究[D]. 苏州: 江苏大学, 2016: 27-29.

[3] 钱建平,李明,杨信廷,等. 基于双侧图像识别的单株苹果树产量估测模型[J]. 农业工程学报, 2013, 29(6): 132-138.

[4] 许立兵,朱启兵,黄敏. 基于 ARM 的苹果采后田间分级检测系统设计[J]. 计算机工程与应用, 2015, 51(16): 234-238.

[5] 田有文,赖兴涛,张芳,等. 基于高光谱成像的苹果果梗完整性识别方法研究[J]. 沈阳农业大学学报, 2018, 49(2): 234-241.

[6] 周成云,许童羽,郑伟,等. 基于深度卷积神经网络的番茄主要器官分类识别方法[J]. 农业工程学报, 2017, 33(15): 219-226.

[7] 马永杰,李雪燕,晓凤. 基于改进深度卷积神经网络的交通标志识别[J]. 激光与光电子学进展, 2018, 55(12): 250-257.

[8] 龙满生,欧阳春娟,刘欢,等. 基于卷积神经网络与迁移学习的油茶病害图像识别[J]. 农业工程学报, 2018, 34(18): 194-201.

[9] BARGOTI S, UNDERWOOD J P. Image segmentation for fruit detection and yield estimation in apple orchards[J]. Journal of Field Robotics, 2017, 34(6): 1 039-1 060.

[10] REN Shao-qing, HE Kai-ming, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1 137-1 149.

[11] ALWZWAZY H, ALBEHADILI H, ALWAN Y, et al. Handwritten digitrecognition using convolution neural networks[J]. International Journal of Innovative Research in Computer & Communication Engineering, 2016, 4(2): 1 101-1 106.

[12] 高震宇,王安,刘勇,等. 基于卷积神经网络的鲜茶叶智能分选系统研究[J]. 农业机械学报, 2017, 48(7): 53-58.

[13] LU Jun. Detecting citrus fruits and occlusion recovery under natural illumination conditions [J]. Computers & Electronics in Agriculture, 2015, 110: 121-130.

[14] 傅隆生,冯亚利,ELKAMIL Tola,等. 基于卷积神经网络的田间多簇猕猴桃图像识别方法[J]. 农业工程学报, 2018, 34(2): 205-211.

[15] 王丹丹,何东健. 基于 R-FCN 深度卷积神经网络的机器人蔬果前苹果目标的识别[J]. 农业工程学报, 2019, 35(3): 156-163.