

阮子行, 黄 勇, 王 梦, 等. 基于改进卷积神经网络的番茄品质分级方法[J]. 江苏农业学报, 2023, 39(4): 1005-1014.
doi: 10.3969/j.issn.1000-4440.2023.04.010

基于改进卷积神经网络的番茄品质分级方法

阮子行¹, 黄 勇^{1,2}, 王 梦², 史 强², 张金玲²

(1. 新疆农业大学机电工程学院, 新疆 乌鲁木齐 830052; 2. 新疆工程学院机电工程学院, 新疆 乌鲁木齐 830023)

摘要: 为了解决番茄人工分级精度低、工作效率低等问题, 基于卷积神经网络提出 1 种用于番茄品质分级的网络结构, 并给予优化改进。设计的卷积神经网络由 7 个权重层(6 个卷积层和 1 个全连接层)和 4 个池化层(3 个最大池化层和 1 个全局平均池化层)构成, 利用批量归一化和压缩激励模块(SE 模块)进行网络结构优化。采用自采集的番茄图像数据集, 通过数据增广将原 1 455 张图片增广至 8 730 张图片并进行训练和测试, 用精确度、召回率、 F_1 值(精确度和召回率的调和平均数)评估模型的各分类差异。优化后的网络模型测试精度为 96.57%, 比未优化的网络模型测试精度提高了 2.58 个百分点。并且与传统经典网络 AlexNet、MobileNet-V2、NasNet-Mobile、ShuffleNet 4 种模型相比, 具有收敛速度更快的优势, 训练时间减少了 22%~96%, 测试精度提高了 0.18~1.89 个百分点, 单张照片测试时间降低了 37%~83%, 计算统一设备架构(CUDA)内存占用比例也得到了一定程度的降低。优化后的网络训练过程更加稳定, 模型注意力更多地集中在整个番茄上, 在一定程度上降低了背景干扰, 提升了算法的鲁棒性与泛化能力。通过探究 3 种优化算法(SGDM、Adam、RMSprop)对模型的影响, 发现与 Adam、RMSprop 优化算法相比, SGDM 优化算法的模型测试精度分别提高了 2.01 个百分点和 2.29 个百分点; 去除数据集的背景后再测试单一背景对模型性能的影响, 发现模型的测试精度达到 96.97%, 相较于未去除背景的数据集, 测试精度提高了 0.40 个百分点。本研究提出的卷积网络结构对简单番茄品质(分类数较少)的分级效果较好, 可为番茄品质分级提供一定的理论支持。

关键词: 番茄; 品质分级; 注意力机制; 改进卷积神经网络; AlexNet

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 1000-4440(2023)04-1005-10

Tomato quality grading method based on improved convolutional neural network

RUAN Zi-hang¹, HUANG Yong^{1,2}, WANG Meng², SHI Qiang², ZHANG Jin-ling²

(1. College of Mechanical and Electrical Engineering, Xinjiang Agricultural University, Urumqi 830052, China; 2. College of Mechanical and Electrical Engineering, Xinjiang Institute of Engineering, Urumqi 830023, China)

Abstract: In order to solve the problems of low accuracy and low work efficiency of tomato manual grading, a network structure for tomato quality grading was proposed and optimized based on convolutional neural network. The designed convolutional neural network consisted of seven weight layers (six convolutional layers and one fully connected layer) and four pooling layers (three max pooling layers and one global average pooling layer). Batch normalization and compressed

excitation module (SE module) were used for network structure optimization. Using the self-collected tomato image data set, the original 1 455 images were augmented to 8 730 images by data augmentation, and then trained and tested. The classification differences of the models were evaluated by accuracy, recall rate and F_1 value (the harmonic mean of accuracy and recall rate). The test accuracy of the optimized network model was 96.57%, which

收稿日期: 2022-06-02

基金项目: 新疆维吾尔自治区科协青年人才托举工程项目 (RCTJ46); 新疆维吾尔自治区高校科研计划自然科学项目 (XJEDU2019Y064)

作者简介: 阮子行 (1996-), 男, 安徽铜陵人, 硕士研究生, 研究方向为图像处理与机器视觉。(E-mail) 1243450926@qq.com

通讯作者: 黄 勇, (E-mail) lishi182@163.com

was 2.58 percentage points higher than that of the unoptimized network model. And compared with the traditional classic network AlexNet, MobileNet-V2, NasNet-Mobile, ShuffleNet, it had the advantage of faster convergence speed. The training time was reduced by 22%–96%, and the test accuracy was improved by 0.18–1.89 percentage points, the test time of a single photo was reduced by 37%–83%, and the compute unified device architecture (CUDA) memory usage was reduced to a certain extent. The optimized network training process was more stable, and the model focused more on the tomato as a whole, which reduced background interference to a certain extent and improved the robustness and generalization ability of the algorithm. The influence of the three optimization algorithms (SGDM, Adam and RMSprop) on the model was explored. Compared with the Adam and RMSprop optimization algorithms, the model test accuracy of the SGDM optimization algorithm was improved by 2.01 percentage points and 2.29 percentage points respectively. After removing the background of the dataset, the effects of a single background on the performance of the model were tested. It was found that the test accuracy of the model reached 96.97%, which was 0.40 percentage points higher than that of the dataset without background removal. The convolutional network structure proposed in this study has better effect on simple tomato quality grading, which provides a certain theoretical support for tomato quality grading.

Key words: tomato; quality grading; attention mechanism; improved convolutional neural network; AlexNet

中国是亚洲地区番茄产量最大的国家,同时也是世界上番茄产量最大的国家,番茄在中国农作物中的种植面积占比较高,其中在 2020 年仅中国新疆地区的番茄产量便达到 $8.242 \times 10^6 \text{ t}^{[1]}$ 。新疆地区番茄的种植面积连续多年保持在 $46\,667 \text{ hm}^2$ 以上,但是纵观整个中国,番茄的产业化水平仍较低,加工产品的附加值也较低,是中国番茄产业遇到的重要问题^[2],其中番茄精确的品质分级技术是亟待解决的重要问题之一。近年来,为了提高农产品品质分级的作业效率,从而解决劳动力不足的问题,农产品品质分级成为农业的重点工作之一^[3]。高精度的农产品分级不但可以提高生产效率,而且可以提高农产品的整体质量,使其更有利于长期存储^[4-5]。品质检测的手段有很多种,随着图像处理与机器视觉的发展,基于机器视觉和深度学习的品质分级技术已经成为研究热点^[6]。本研究旨在从番茄外部品质分级的角度出发,通过深度学习手段为番茄品质分级提供一种新方法。

目前,国内外学者在农产品品质分级方面做了大量工作,其中 Blasco 等^[7]开发的机器视觉算法基于贝叶斯估计分割水果与背景,并通过尺寸、颜色、茎的位置和外部瑕疵的检测在线评估橙子、桃子和苹果的质量,该分类系统经过在线测试苹果,在批量分类水果时获得了良好的性能,瑕疵检测、尺寸估计的准确度分别为 86%、93%。Sayed 等^[8]提出了 1 种基于海洋捕食者算法(MPA)和卷积神经网络的柑橘病害分类的新混合方法。MPA 用于找到批量大小、退出率、退出期和最大训练批

次,在柑橘病害分类方面的整体准确率可达 100%,表明所提出的基于 MPA 优化的 ResNet50 具有优越性。Bhole 等^[9]基于迁移学习的预训练 SqueezeNet 模型,根据缺陷、形状、大小和成熟度等参数对杧果进行评估分级,2 份数据集分别为杧果果实的 RGB 图像和热图图像,测试结果表明,该系统对 2 份数据集的分类准确率分别为 93.33%、92.27%,RGB 图像、热图图像的训练时间分别为 30.03 min、7.38 min。何进荣等^[10]通过对经典网络模型进行优化,采用经典的卷积结构作为苹果外部特征提取器,并采用批量归一化的方式优化模型,采集了 19 500 张图片,按照 3:1 的比例划分数据集,结果显示,DXNet 模型的分级准确率达到 97.84%。胡发焕等^[11]采用形态学手段将目标脐橙与背景分离,并捕捉脐橙的果面缺陷、尺寸、表面颜色等特征,将其输入支持向量机,训练后的品质分级识别率达到 90.5%,并且实时性较好。

本研究拟以分选台常见的 6 种番茄样本作为研究对象,设计改进网络结构,并通过批量归一化和注意力机制对其进行优化,研究优化后的模型整体性能与批量归一化、注意力机制对模型的影响,以期探究不同优化算法和单一背景数据集对网络模型性能的影响。

1 材料与方法

1.1 数据集

本试验中的所有番茄样本均购自新疆乌鲁木齐番茄市场,番茄品种为樱桃番茄,灯光由普通室内照

明用日光灯提供,采用一加8手机进行图像采集,番茄的外观形态共6种,详见图1。

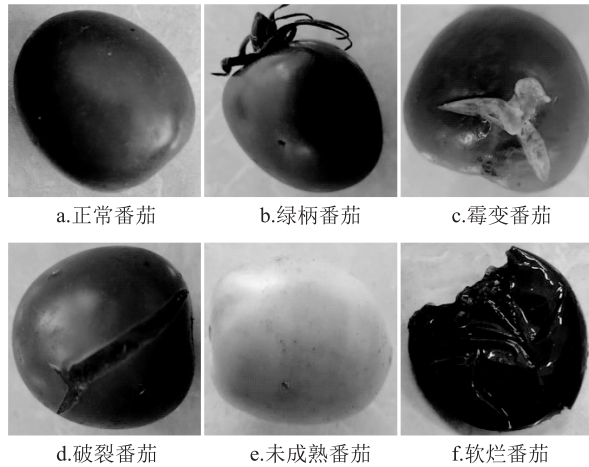


图1 常见番茄的外观形态

Fig.1 Common types of tomatoes

经过筛选后,数据集中共有1 455张图片,各类番茄的具体图片数量见表1。为了获得更好的训练效果,在网络训练前对原始数据进行一些处理。(1)由于不同网络模型中图片数据的输入尺寸不同,所以要根据不同网络模型将图片数据放缩至规定尺寸和格式,本研究中的图片均采用JPG格式^[12]。(2)由于大量数据样本可以增加模型训练的准确度,本研究拟通过任意角度旋转剪切、增强对比度、引入噪声和水平翻转等手段进行数据增广^[13-14]。增强后的各类样本图片数据集数量如下:正常番茄1 758张,绿柄番茄1 650张,霉变番茄1 698张,破裂番茄900张,未成熟番茄1 554张,软烂番茄1 170张,共计8 730张,并用数字1~6给不同外观形态的番茄编号,如表1所示。

表1 番茄样本图片的原始数量及增广后的数量

Table 1 Data from original and augmented samples of tomato

番茄外观	样本标签	原始图片数 (张)	总图片数 (张)
正常	1	293	1 758
绿柄	2	275	1 650
霉变	3	283	1 698
破裂	4	150	900
未成熟	5	259	1 554
软烂	6	195	1 170
总计		1 455	8 730

1.2 试验环境

模型搭建和训练所用试验环境如下:电脑操作系统为64位Win10家庭版,软件平台采用Matlab 2022a中的深度网络设计器(Deep learning toolbox),显卡为NVIDIA GeForce RTX 3050 Laptop(4 G显存),计算统一设备架构(CUDA)版本为11.6.0,CUDNN版本为11.3,采用三星16 G内存、三星512 G固态硬盘。本网络训练均基于Matlab2022a平台,采用图形加速器加速进行训练,试验批次(Batch size)为16次,训练轮数为30轮,迭代次数为9 810次。

2 卷积神经网络模型的构建

2.1 网络模型结构

2.1.1 卷积层结构 本研究设计的卷积网络选取6个卷积层,分别为Conv1、Conv2、Conv3、Conv4、Conv5、Conv6,其中前5个卷积层的作用是捕捉图像信息特征,第6个卷积层采用1×1的卷积核,其主要作用是实现通道维度的变化,减少参数量并结合激活函数引入更多非线性特征^[15-16]。其中卷积核尺寸、步长、各个卷积核数量及Padding参数见表2。

表2 网络中各个卷积层参数

Table 2 Parameters of each convolutional layer in the network

卷积层名	卷积核尺寸	步长	卷积核数量 (个)	Padding
Conv1	5×5	4	96	0
Conv2	5×5	1	96	2
Conv3	3×3	1	128	1
Conv4	3×3	1	128	1
Conv5	3×3	1	256	1
Conv6	1×1	1	512	1

Padding是卷积核的1个修正参数。

本网络中的激活函数均采用ReLU激活函数,采用4个池化层(其中3个为最大池化层,1个为全局平均池化层)和1个全连接层,本研究图片数据集中的番茄共有6种分类类型,因此将最后全连接层节点数设为6个。

2.1.2 全连接层 目前常见的卷积神经网络有很多种,其中LeNet、AlexNet和VGGNet这3种卷积神经网络发布得较早,其网络结构没有分支,相对其他网络结构更为简洁。第1次利用大批量图像数据集完成深层卷积神经网络结构的便是AlexNet网络,它正式打开了深度学习的大门,该网络的权重层有

8 层(5 层卷积层、3 层全连接层),进行 1 000 种分类时模型可计算的参数量为 6.09×10^{17} [17-18]。对 Alex-

Net 网络参数量进行分析,结果见表 3。

表 3 AlexNet 网络各层参数

Table 3 Parameter table of each layer of AlexNet network

层名称	说明	节点数量(个)	参数数量(个)
输入层	$W \times H \times C = 227 \times 227 \times 3$	154 587 (227×227×3)	
CNN1	96 个 $11 \times 11 \times 3$ 卷积核	290 400 (55×55×48×2)	34 944 (11×11×3×96+96)
CNN2	256 个 $5 \times 5 \times 48$ 卷积核	186 624 (27×27×128×2)	307 456 [(5×5×48×128+128)×2]
CNN3	384 个 $3 \times 3 \times 256$ 卷积核	64 896 (13×13×192×2)	885 120 (3×3×256×384+384)
CNN4	384 个 $3 \times 3 \times 192$ 卷积核	64 896 (13×13×192×2)	663 936 [(3×3×192×192+192)×2]
CNN5	256 个 $3 \times 3 \times 192$ 卷积核	43 264 (13×13×128×2)	442 624 [(3×3×192×128+128)×2]
FC1	全连接层	4 096	37 752 832 [(6×6×128×2)×4 096+4 096]
FC2	全连接层	4 096	16 781 312 (4 096×4 096+4 096)
FC3(输出层)	全连接层	1 000	4 097 000 (4 096×1 000+1 000)
合计		813 859	60 965 224

W:图片宽度;H:图片高度;C:图片的通道数。

对 AlexNet 网络各层参数进行统计可以发现,全连接层参数数量为 58 631 144 个,占整个网络参数数量的 96.17%,5 层卷积层的参数量占比不足 4.00%。参数过多使得训练过程对设备的要求更高,造成计算效率降低,更容易出现过拟合,并且也限制了网络在移动平台中的布置 [19-20]。

结合经典网络 AlexNet 的网络参数表可以看出,全连接层中包含整个网络的绝大多数参数信息,需要用其他结构代替,从而进行优化改进。孙俊等 [21]、张华鹏 [22] 针对 AlexNet 模型数据参数庞大的问题,通过全局平均池化层代替全连接层的方式优化了 AlexNet 网络模型参数,使得模型参数的内存需求仅为 2.6 MB。另一种结构采用 1×1 的卷积操作来实现, 1×1 的卷积操作可以在保证参数量为 1 的同时实现网络通道数的变换,并配合非线性激活函数引入更多非线性特性 [23-24]。

结合以上 2 种方式,本研究设计的卷积神经网络结构首先使用 1×1 的卷积操作以增加通道数,经过全局池化层降低参数量,通过全连接层将上一层的全局池化层进行全连接,并传递给 Softmax 分类函数实现分类输出,该模型结构相较于拥有 3 个全连接层的结构,参数量可以得到极大减少,训练时可使图形处理器 (GPU) 中 CUDA 内存占用率由原来的 70%~80% 降至 30%~40%。

2.2 网络模型结构的优化

2.2.1 引入批量归一化层优化 [25] 在卷积神经网络中,当底层网络梯度发生微小改变时,经过激活函数及一系列线性映射后,上层梯度会发生较大幅度的变化,随着网络层数的增加,这种变化会不断被放大,网络也会不断适应梯度的变换,导致网络越来越难以被训练。批量归一化 (Batch normalization, BN) 操作是将这些输入值规范化至 1 个合适范围内,在不影响梯度变换大趋势的前提下降低网络整体大幅度变化带来的不确定性,可以增加数据的稳定性。数据的稳定传递可以加快模型训练的速度,降低网络对参数的敏感程度,使得由超参数造成的网络变化规律更加明显,网络学习更加稳定。在本网络中选择在 Conv2、Conv4 及 Conv6 后增加批量归一化操作,批量归一化公式如下:

$$u = \frac{1}{m} \sum_{i=1}^m X_i \quad (1)$$

$$\delta^2 = \frac{1}{m} \sum_{i=1}^m (X_i - u)^2 \quad (2)$$

$$Y_i = r \frac{X_i - u}{\sqrt{\delta^2 + \varepsilon}} + \beta \quad (3)$$

式中: X 为输入数据; Y 为输出数据; β 、 r 分别是平移参数、缩放参数,在反向传播中训练得到; $\varepsilon > 0$,且是很小的数; m 为输入数据数量; u 为输入数据的均值; δ 为输入数据的标准差。

网络的整体结构如图 2 所示。

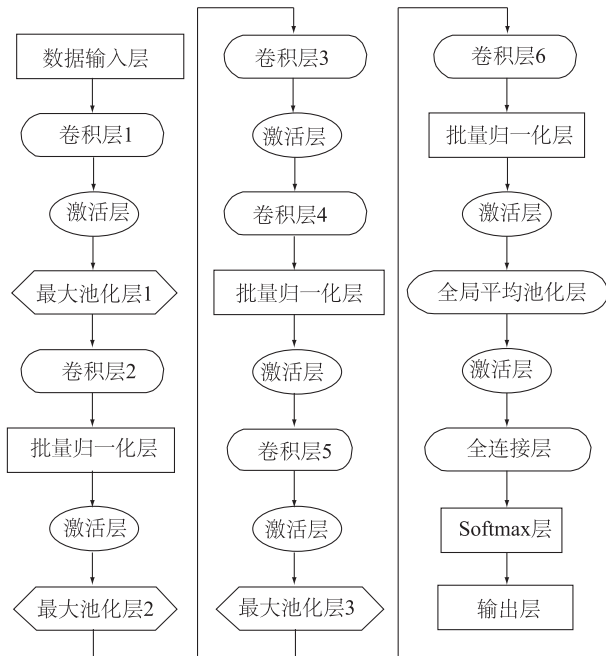
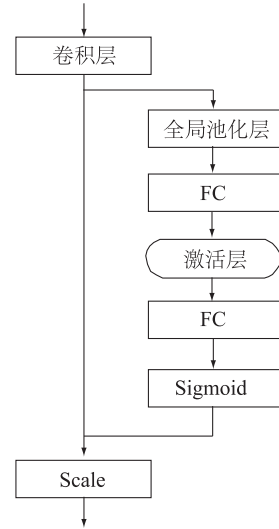


图 2 本研究设计的神经网络结构

Fig.2 Neural network structure designed in this study

2.2.2 通过注意力机制优化^[26] 敏锐度指人视网膜对不同物体具有的信息处理能力,表现为重点关注某些部分而忽视某些部分。在神经网络中,通过给予不同特征不同权重来实现注意力的分配,将更多计算资源分配至更重要的任务中,从而更加高效地实现信息处理^[27]。

图 3 为压缩激励注意力模型结构,该模型学习卷积过程中通道之间的相关性,从而实现注意力分配,具体分为 3 个部分。(1)压缩(Squeeze)。是通过 1 个全局平均池化层将特征图压缩为 $1 \times 1 \times C$ 的向量,再经全连接层(FC)把神经元数量减少,该操作通过减少特征图的宽、高从而降低计算量,除去过多的无用信息,其中降维系数(r 值)选择 16。(2)激励(Excitation)。压缩后 C/r 个通道数的特征图,由激活函数 ReLU 激活,赋予每个通道不同的权重,再经 FC 将 C/r 个通道恢复成 C 个通道,得到 $1 \times 1 \times C$ 尺寸的数据。最后由 Sigmoid 激活函数使各个通道的权重归一化。此时该向量与输入特征图有相同的通道数,并且每个通道的权重不同。(3)Scale 操作。得到 $1 \times 1 \times C$ 向量后,与输入特征图的 $W \times H \times C$ 进行 Scale 操作,赋予输入特征图权重,得到不同权重的 $W \times H \times C$ 特征图。



FC:全连接层;Sigmoid:一种函数名称;Scale:通道权重相乘。

图 3 压缩激励注意力模型结构

Fig.3 The structure of the squeeze excitation attention model

引入压缩激励模块(SE 模块)进一步提升网络模型的精度,在图 2 神经网络结构的基础上,在卷积层 1、卷积层 2、卷积层 3、卷积层 4、卷积层 5 后均加入 SE 注意力机制模块,以此优化整个网络结构。

3 结果与分析

3.1 混淆矩阵分析

选用数据增广后的 6 类番茄样本数据,共计 8 730 张图片,按照 4 : 1 的比例划分为训练集与测试集,并按照网络输入大小裁剪为 227 像素 \times 227 像素 \times 3 (宽 \times 高 \times 通道数) 的尺寸,图 4 为本研究设计的卷积神经网络对 6 种不同番茄外观形态进行分类的混淆矩阵。

对样本进行测试后分别计算精确度(P)、召回率(R)、精确度和召回率的调和平均数(F_1),其公式如下:

$$P = \frac{TP}{TP + FP} \times 100\% \quad (4)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (5)$$

$$F_1 = \frac{R \times P \times 2}{R + P} \times 100\% \quad (6)$$

式中: TP 代表预测为正的正样本; FP 代表预测为正的负样本; FN 代表预测为负的正样本。

表 4 为 6 类番茄外观形态的 3 种分级评估指标参数,其中未成熟番茄的外观形态分级效果最精确,软烂番茄的外观形态分级效果最差,6 类番茄外观

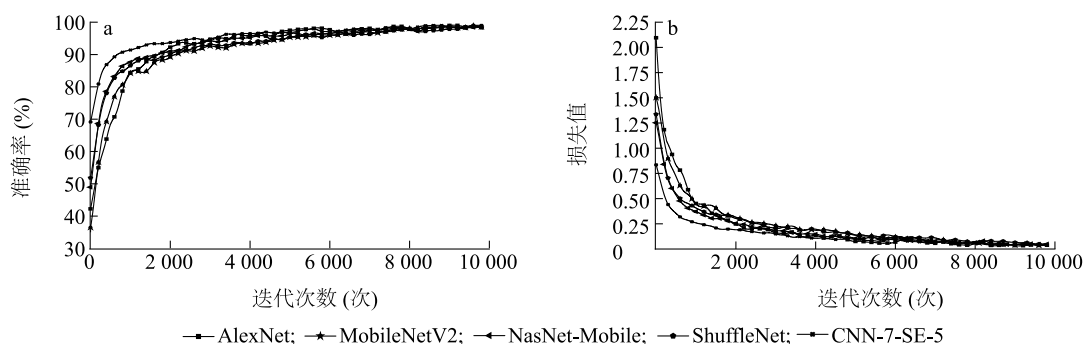
真实分类	1	344	5		3			97.7%	2.3%	
	2	11	319					96.7%	3.3%	
	3			320	2		18	94.1%	5.9%	
	4		5	3	171		1	95.0%	5.0%	
	5					311		100.0%	0	
	6			9	3		222	94.9%	5.1%	
		96.9%	97.0%	96.4%	95.5%	100.0%	92.1%			
		3.1%	3.0%	3.6%	4.5%	0	7.9%			
		1	2	3	4	5	6			
		预测分类								

图4 本研究模型的混淆矩阵

Fig.4 The confusion matrix of the model constructed in this study

形态的综合分类精确度(表5中的测试精度)达到96.57%,表明本研究设计的神经模型对番茄各个外观形态的识别能力均较强。

本研究设计并优化后的网络以 CNN-7-SE-5 作为记号。将目前常见的4种神经网络模型输出节点数微调后,用以自建番茄数据集训练网络,并与本研究设计的网络模型的训练数据进行对比。从图5a可以看出,该网络模型在收敛速度上相较于 AlexNet、MobileNet-V2、NasNet-Mobile、ShuffleNet 4种模型更快,迭代500次左右就优先达到90%的训练精度(准确率),并且整体训练精度变化稳定,没有大幅度的变化,最终5种模型的训练精度都在95%以上。图5b可以看出,本研究设计的模型也是收敛最快的,迭代1000次左右便最先将损失值降至0.25,最终5种模型的损失值均在0.13左右。



a: 训练精度曲线; b: 训练损失曲线。

图5 5种模型训练曲线

Fig.5 Five model training curves

表4 6类番茄评估指标的对比

Table 4 Comparison of evaluation indicators of six kinds of tomatoes

番茄外观	精确度 (%)	召回率 (%)	F_1
正常	96.9	97.7	0.97
绿柄	97.0	96.7	0.97
霉变	96.4	94.1	0.95
破裂	95.5	95.0	0.95
未成熟	100.0	100.0	1.00
软烂	92.1	94.9	0.93

F_1 : 精确度和召回率的调和平均数。

本研究设计的模型结构,以自建的6种番茄外观形态分类数据为数据集进行模型训练,得到表5中的数据。可以看出,该模型在识别精度上相较于经典模型并没有明显优势,但是在参数量、单张照片测试时间及训练时间上的优势明显,训练时间减少了22%~96%,测试精度提高了0.18~1.89个百分点,单张照片测试时间减少了37%~83%,此外,CUDA内存占用比例也得到了一定程度的降低。

3.2 试验模型性能探讨与分析

3.2.1 优化后操作对网络模型的影响 分别对优化后的网络模型(引入SE模块、批量归一化)和未优化的网络模型进行训练与测试,对比在相同数据集、环境、超参数条件下2种模型的训练精度和损失值随迭代次数的变化趋势。如图6所示,在模型训练过程中共迭代9810次,加入SE模块和批量归一化后的网络模型在前2000次迭代中的收敛速度明显优于未优化网络的模型,训练好的模型在测试集中的测试精度也由未优化的93.99%提升至优化后的96.57%,提高了2.58个百分点。

表 5 5 种模型训练参数的对比

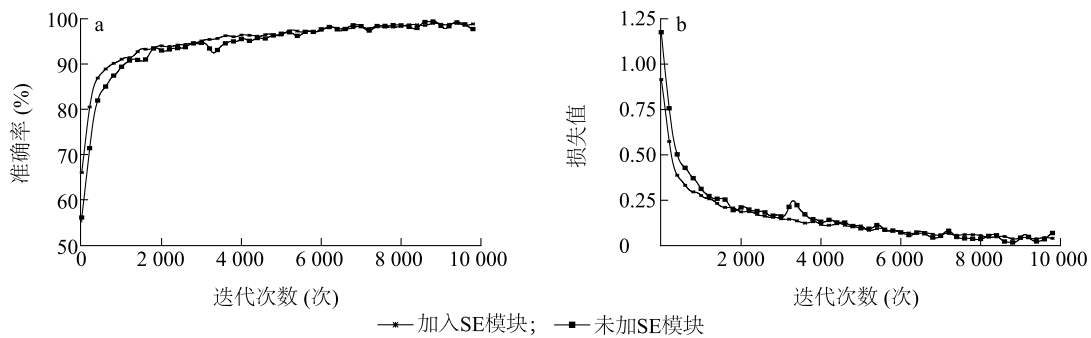
Table 5 Comparison of training parameters of five models

网络名称	可学习参数总数	训练时间 (s)	测试精度 (%)	单张照片测试时间 (s)	CUDA 内存占用比例 (%)
AlexNet	56 800 000	1 634	95.02	0.001 538 4	70~80
MobileNet-V2	2 200 000	5 748	94.68	0.003 394 9	50~60
NasNet-Mobile	4 200 000	30 465	96.39	0.005 766 3	25~75
ShuffleNet	864 900	3 640	95.66	0.001 683 4	50~60
CNN-7-SE-5	942 700	1 268	96.57	0.000 976 7	30~40

CUDA:计算统一设备架构。

图 6 显示,与优化后的网络相比,未优化的网络在训练过程中准确率与损失值变化波动起伏较大,通过本研究中的优化操作可以使网络模型的训练更

加稳定,更有利于网络模型的建立,在测试精度、网络训练稳定性上都能得到一定程度的提升,说明本研究的优化方式是十分有效的。



a: 训练精度曲线; b: 训练损失曲线。

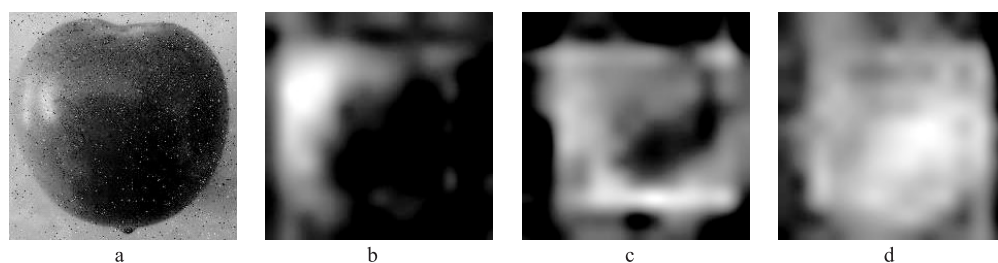
图 6 模型优化前后性能对比曲线

Fig.6 Performance comparison curves before and after model optimization

3.2.2 注意力机制对网络模型性能的影响 为了验证注意力机制 SE 模块对具体目标的影响,对比番茄外观形态原图和 3 种网络的 Conv5 层中激活区域最大特征的通道图(AlexNet 网络、本研究设计的未加 SE 模块的网络、本研究设计的增加 SE 模块后的网络)。由于 AlexNet 网络与本研究设计的网络同为采用 5 层卷积层提取特征的网络,因此选用 AlexNet 网络进行对比。图 7b 为 AlexNet 网络 Conv5 层中激活区域最大的特征图,对比原图(图 7a)可以明显看出,正向激活区域主要集中在番茄左上角高光区域及周边区域。由图 7c 可以看出,本研究设计的网络在未加入 SE 注意力机制时,Conv5 层最大激活区域图中整个正向区域相较于图 7b 明显增加,包含在整个番茄四周。由图 7d 可以看出,加入 SE 注意力机制的 Conv5 层最大激活区域中的激活区域得到进一步增加,整个番茄区域中绝大部分都被正向激活,使得目标特征得到更多关注,也避免了番茄表面局部特征和背景噪音的干扰,能够增强网络

分类的泛化能力与鲁棒性。

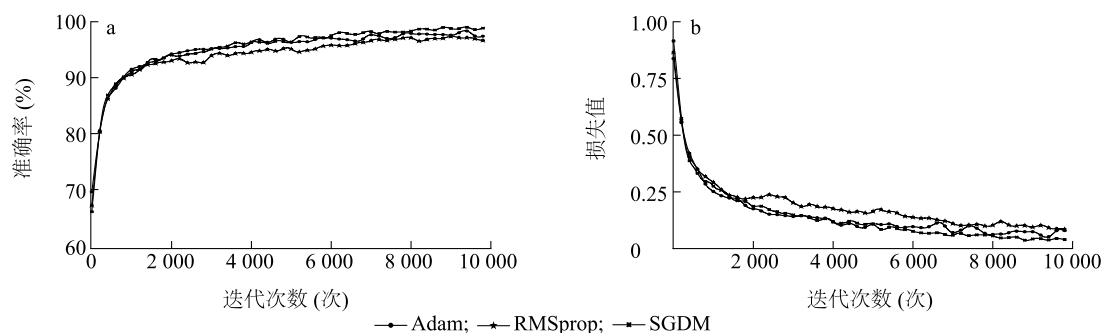
3.2.3 优化算法对模型性能的影响 在相同数据集、参数设置的情况下,分别采用 SGDM、Adam、RMSprop 这 3 种常见的优化算法对改进后的卷积神经网络进行网络训练对比。从图 8 可以看出,在前 1 000 次迭代中,3 种优化算法的训练精度曲线和损失曲线基本保持重合;在迭代次数达到 1 000 次后,3 条曲线的轨迹开始发生变化,其中 SGDM 优化算法的训练精度曲线增加幅度和速率都高于采用 Adam、RMSprop 优化算法的卷积模型;损失值降低幅度和速率也高于采用 Adam、RMSprop 优化算法的卷积模型。采用 SGDM、Adam、RMSprop 3 种优化算法的最终模型测试精度分别为 96.57%、94.56%、94.28%。由此可见,SGDM 优化算法相较于 Adam、RMSprop 优化算法更适合本模型,测试精度分别提高了 2.01 个百分点和 2.29 个百分点,因此在本研究中,SGDM 优化算法是更好的选择。



a:原图;b:AlexNet 网络激活;c:未加 SE 模块激活;d:加 SE 模块激活。

图 7 3 种模型的 Conv5 层激活区域对比

Fig.7 Comparison of the activation regions of the Conv5 layer of the three models



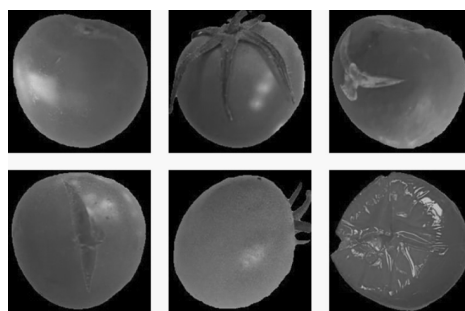
a:训练精度曲线;b:训练损失曲线。

图 8 3 种优化算法对应的训练曲线

Fig.8 Training curves corresponding to the three optimization algorithms

3.2.4 单一背景数据集对网络性能的影响 通过传统的图像处理手段,先进行滤波降噪,消除噪声干扰,再通过 RGB 和 HSV 颜色空间的不同算子实现了 6 种类型番茄与背景的分割,其中注意避免高亮区域对背景分离的干扰,结合 RGB 空间和 YIQ 空间两个空间信息特点,利用布尔运算实现目标的完整分割,大部分情况分割效果较好,如图 9 所示。通过对比两组数据集在相同模型下的训练效果可以发现,单一黑色背景数据集训练出的模型测试精度为 96.97%,相较于未去除背景的数据集测试精度提高了 0.40 个百分点,可以推测背景中存在影响学习分类的噪声,影响了模型的学习,移除背景后减少了背景干扰,使得分类精度得到提升。

消除背景后,虽然测试精度得到小幅度提升,但是提升幅度并不明显,可能是仍有部分数据并未很好地实现背景分割导致,详见图 10。由于光线照射原因导致背景明暗程度不一致,使得番茄目标不能从背景中完全分割出来,整体数据样本存在小部分数据不一致,存在的部分背景在训练过程中会引入不确定的干扰因素,导致训练精度提升并不明显。



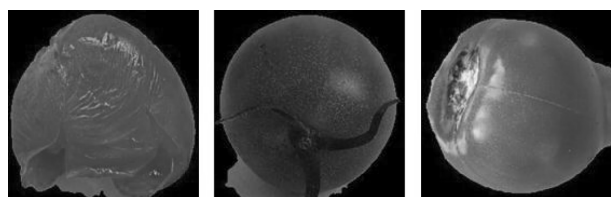
第 1 行的 3 张图从左到右依次为去除背景的正常番茄、去除背景的绿柄番茄、去除背景的霉变番茄,第 2 行的 3 张图从左到右依次为去除背景的破裂番茄、去除背景的未成熟番茄、去除背景的软烂番茄。

图 9 消除背景的 6 种番茄

Fig.9 Six tomato samples without background

4 结论

以自采集的 6 类番茄外观形态数据集为研究对象,通过数据增广对数据进行 6 倍扩增,构建并优化卷积神经网络算法,对 6 类不同品质番茄图像进行检测与分类,讨论网络模型训练效果并与其他模型



3 张图从左到右依次为未完全去除背景的破裂番茄、未完全去除背景的青柄番茄、未完全去除背景的霉变番茄。

图 10 背景未完全分割的番茄样品

Fig. 10 Tomato sample image with incompletely segmented background

进行对比,探究归一化操作、注意力 SE 机制、3 种优化算法、数据单一背景等对模型性能的影响,得到如下结论:

(1)本研究设计的卷积神经网络结构中使用 1×1 的卷积层和全局平均池化层并通过 1 个全连接层实现所有神经元的全连接,并传递给 Softmax 分类函数实现分类输出,与 3 个全连接层相比,此结构的参数量可以得到极大降低,训练时使得 GPU 中的 CUDA 内存占用比例得到了一定程度的降低。

(2)利用批量归一化和压缩激励模块(SE 模块)进行网络结构优化后,与未优化的网络模型相比,优化后的网络模型测试精度提高了 2.58 个百分点。并且对比传统经典 AlexNet、MobileNet-V2、NasNet-Mobile、ShuffleNet 4 种网络模型,具有收敛速度更快、模型体积更小、计算量更少等优势,训练时间减少了 22%~96%,测试精度提高了 0.18~1.89 个百分点,单张照片测试时间降低了 37%~83%。

(3)采用批量归一化和注意力机制优化后的网络训练过程更加稳定,模型注意力更多集中在番茄整体上,正向激活区域也更多,在一定程度上降低了背景干扰,提升了算法的鲁棒性和泛化能力。

(4)探究 3 种优化算法 SGDM、Adam、RMSprop 对模型的影响,结果表明,SGDM 优化算法的模型测试精度相较于 Adam、RMSprop 优化算法分别提高了 2.01 个百分点和 2.29 个百分点,可见 SGDM 优化算法更适合该网络模型与数据集。为了探究单一背景对模型性能的影响,通过图像处理去除数据集背景,重新训练后模型的准确率提高到 96.97%,相较于未去除背景的数据集测试精度提高了 0.40 个百分点,可见去除背景在一定程度上降低了背景噪点的干扰,使得模型准确度得到提升。

参考文献:

- [1] 张超,李晓艳,汪晶晶. 绿色贸易壁垒对新疆番茄及其制品出口影响分析[J].北方园艺,2022(9):135-141.
- [2] 李凯,谭丹. 新疆番茄产业国际竞争力分析[J].现代商业,2018(35):65-68.
- [3] 侯新科,常英. 基于机器视觉的樱桃番茄在线分级系统设计[J].计算机与数字工程,2020,48(7):1798-1803.
- [4] 纪莹莹. 基于计算机视觉的西红柿品质分级技术研究[D].长春:吉林农业大学,2021.
- [5] 马高尚. 基于机器视觉的梨品质检测分级系统研究[D].秦皇岛:河北科技师范学院,2021.
- [6] 田有文,吴伟,卢时铅,等. 深度学习在水果品质检测与分级分类中的应用[J].食品科学,2021,42(19):260-270.
- [7] BLASCO J, ALEIXOS N, MOLTÓ E. Machine vision system for automatic quality grading of fruit[J]. Biosystems Engineering, 2003, 85(4): 415-423.
- [8] SAYED G I, HASSANIEN A E, TANG M. A novel optimized convolutional neural network based on marine predators algorithm for citrus fruit quality classification[M]//LISS Proceedings of the 11th International Conference on Logistics. Singapore: Springer, 2022: 682-692.
- [9] BHOLE V, KUMAR A. Mango quality grading using deep learning technique: perspectives from agriculture and food industry[C]// Proceedings of the 21st Annual Conference on Information Technology Education. USA, 2020: 180-186.
- [10] 何进荣,石延新,刘斌,等. 基于 DXNet 模型的富士苹果外部品质分级方法研究[J].农业机械学报,2021,52(7):379-385.
- [11] 胡发焕,董增文. 基于机器视觉的脐橙品质在线分级检测[J].湖北农业科学,2014,53(9):2160-2164.
- [12] 赵辉,曹宇航,岳有军,等. 基于改进 DenseNet 的田间杂草识别[J].农业工程学报,2021,37(18):136-142.
- [13] 郭小清,范涛杰,舒欣. 基于改进 Multi-Scale AlexNet 的番茄叶部病害图像识别[J].农业工程学报,2019,35(13):162-169.
- [14] 樊湘鹏,周建平,许燕,等. 基于改进卷积神经网络的复杂背景下玉米病害识别[J].农业机械学报,2021,52(3):210-217.
- [15] 马崇鼻,唐娉,赵理君,等. 深度学习图像数据扩增方法研究综述[J].中国图象图形学报,2021,26(3):487-502.
- [16] 唐爽. 基于深度神经网络的微表情识别[J].电子技术与软件工程,2017(3):93-95.
- [17] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [18] 刘晓龙. 面向开放集的模糊域自适应图像分类及应用[D].无锡:江南大学,2021.
- [19] 孙颖异,李健,时天,等. 基于改进的 AlexNet 卷积神经网络

- 络的植物叶片识别[J].种子,2020,39(2):77-81.
- [20] 叶涛涛. 基于深度卷积神经网络的 X 光影像分析研究[D].杭州:浙江工业大学,2020.
- [21] 孙 俊,谭文军,毛罕平,等. 基于改进卷积神经网络的多种植物叶片病害识别[J].农业工程学报,2017,33(19):209-215.
- [22] 张华鹏. 基于深度学习的苹果病害识别与施药辅助决策系统研发[D].泰安:山东农业大学,2021.
- [23] LIN M, CHEN Q, YAN S. Network in network[J]. arXiv preprint arXiv:1312.4400, 2013.
- [24] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv, 2014, 1409:1556.
- [25] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//International Machine Learning Society, IMLS. International Conference on Machine Learning. Lille, France, 2015: 448-456.
- [26] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Computer Vision Foundation/IEEE Computer Society. Proceedings of the IEEE conference on computer vision and pattern recognition. Salt Lake City, UT, USA, 2018: 7132-7141.
- [27] 王文冠,沈建冰,贾云得. 视觉注意力检测综述[J].软件学报,2019,30(2):416-439.

(责任编辑:徐 艳)