

DOI:10.16136/j.joel.2022.06.0633

基于双通道注意力机制的 ResNet 果实外观品质分类

赵 辉^{1,2}, 乔艳军^{1*}, 王红君¹, 岳有军¹

(1. 天津理工大学 电气工程学院 天津市复杂系统控制理论与应用重点实验室, 天津 300384; 2. 天津农学院 工程技术学院, 天津 300392)

摘要: 为了实现对采摘后的果实进行快速、精确的外观品质分类,并配合分拣生产线完成果实大规模集中分拣,该研究提出了一种基于改进 ResNet 的果实分类方法。首先,将深度残差神经网络(deep residual neural network, ResNet)网络中的残差模块与双通道 SE 模块(dual channel squeeze-and-excitation block, DC-SE Block)结合,增强有效的通道特征并抑制低效或无效的通道特征,提高特征图的表达能力,从而提升识别精度;其次,在原始 ResNet 模型中加入 Inception 模块,将果实不同尺度的特征进行融合,增强对较小缺陷的识别能力;最后,对收集到的 4 类不同外观品质的果实图像进行数据增强并利用迁移学习的方法对模型进行初始化。以苹果为例进行的试验结果表明:经过数据集训练之后的改进模型,在测试集下的准确率达到 99.7%,高于原模型的 98.5%;精确率达到 99.7%,高于原模型的 98.3%;召回率达到 99.7%,高于原模型的 98.7%;在图形处理器(graphic processing unit, GPU)下的平均检测速度达到 32.3 帧/s,略低于原模型的 35.7 帧/s。与 GoogleNet、MobileNet 等几种目前先进的分类方法进行比较并对不同改进模型进行对比试验的结果表明,该方法具有良好的分类性能,对解决果实外观品质的精准分级问题具有重要参考价值。

关键词: 图像增强; 外观品质分类; 深度残差神经网络; 双通道 SE 模块; Inception 模块; 迁移学习中图分类号: TP391.4 文献标识码: A 文章编号: 1005-0086(2022)06-0643-09

ResNet fruit appearance quality classification based on dual channel attention mechanism

ZHAO Hui^{1,2}, QIAO Yanjun^{1*}, WANG Hongjun¹, YUE Youjun¹

(1. Tianjin Key Laboratory of Control Theory & Applications in Complicated Systems, School of Electrical and Electronic Engineering, Tianjin University of Technology, Tianjin 300384, China; 2. College of Engineering and Technology, Tianjin Agricultural University, Tianjin 300392, China)

Abstract: In order to realize the fast and accurate appearance quality classification of picked fruits, and cooperate with the sorting production line to complete the large-scale centralized sorting of fruits, a fruit classification method based on improved ResNet is proposed in this study. Firstly, the residual module in ResNet network is combined with the dual channel squeeze-and-excitation block (DC-SE Block) to enhance the effective channel features, suppress the inefficient or invalid channel features, and improve the expression ability of the feature map, so as to improve the recognition accuracy. Secondly, the Inception module is added to the original ResNet model to fuse the characteristics of different scales of fruit, so as to enhance the recognition ability of small defects. Finally, four kinds of fruit images with different appearance quality are enhanced, and the model is initialized by transfer learning method. Taking apple as an example, the experimental results show that the accuracy of the improved model trained by the data set is 99.7%, which is higher than 98.5% of the original model; The precision rate is 99.7%, which is

* E-mail: qiaoyanjun1996@163.com

收稿日期: 2021-09-06 修订日期: 2021-10-08

基金项目: 天津市科技支撑计划项目(18YFZCNC01120, 19YFZCSN00360)资助项目

higher than 98.3% of the original model; The recall rate reaches 99.7%, which is higher than 98.7% of the original model; The average detection speed under graphic processing unit (GPU) is 32.3 frame/s, which is slightly lower than 35.7 frame/s of the original model. Compared with several advanced classification methods such as GoogleNet and MobileNet, and compared with different improved models, the results show that the proposed method has good classification performance, and has important reference value for solving the problem of accurate classification of fruit appearance quality.

Key words: image enhancement; appearance quality classification; deep residual neural network (ResNet); dual channel squeeze-and-excitation block (DC-SE Block); Inception module; transfer learning

1 引 言

水果分拣生产线可实现分拣工作的自动化,能够很好地解决劳动力不足、人工分拣成本高且效率低的问题。水果分拣生产线的重点和难点在视觉系统,分拣机的工作效率和稳定性取决于其对果实品质分类的速度和准确率,而果实的表面缺陷情况是影响水果品质的重要指标之一,因此,研究一套能够精确识别果实不同外观缺陷的视觉系统,对实现果实自动分拣具有重要研究价值和现实意义^[1,2]。

目前,国内外在水果外观分类方面的研究已经取得了一定进展。霍迎秋等^[3]通过提取苹果不同病斑的 8 个纹理特征组成训练特征矩阵,利用压缩感知理论实现了对轮纹病、炭疽病和新轮纹病 3 种病害果实的识别与分类,准确率可达 90.0%;NIE 等^[4]通过图像预处理提取果实的目标区域,并提取苹果的颜色、缺陷等外观品质特征,最终利用基于遗传算法优化的 BP(back propagation)神经网络实现了苹果的外观品质分类,准确率为 91.7%。上述传统的分类方法对由人工提取出的特征过于依赖,鲁棒性和泛化性较差,而在实际分拣过程中果实外观的种类繁多且复杂,某些不同品质的果实外观在纹理、颜色等方面具有较强的相似性,同一品质的果实外观也会存在较大的差异性,所以仅凭人工提取出的一种或几种特征然后基于机器学习算法进行分类会导致分类准确率较低,难以满足分拣生产线的实际工作要求^[5,6]。近年来,随着深度学习的发展,卷积神经网络可以通过无监督或半监督的特征学习来提取多层次特征,比人工提取的特征具有更强的泛化能力,其强大的表征能力使图像分类的性能获得了很大提升,推动了图像分类的发展^[7,8]。李善军等^[9]提出了一种改进的单点多盒探测器(single shot multibox detector, SSD)柑橘实时分类检测方法,利用深度残差神经网络(deep residual neural network, ResNet18)作为其特征提取网络,有效提升了识别速度,实现了对 3 种表皮缺陷柑橘的实

时检测,准确率可达 87.9%;AZDC A 等^[10]提出了一种基于 ResNet50 的番茄外部缺陷分类方法,利用迁移学习对模型进行初始化,有效提高了识别精度,准确率为 94.6%;岳有军等^[11]提出了一种基于改进 VGGNet 的苹果分级方法,在 VGG16 的基础上,加入批归一化层和全局平均池化层,并采用联合损失函数对其进行优化,准确率达到 99.3%;MOMENY M 等^[12]提出了一种基于混合池化的樱桃品级分类方法,在卷积层之后添加最大池化和平均池化组成的混合池化层,有效提高了模型的泛化能力,准确率可达 99.4%。

鉴于此,本文以苹果为研究对象,以 ResNet 网络为基础,根据目前果实外观缺陷分类存在的问题,对原网络模型进行优化和改进,并在模拟分拣生产线环境下进行分类试验。

2 数据材料

2.1 图像获取

为了模拟实际分拣生产线环境,本文所用的部分苹果图像拍摄于实验室搭建的图像采集装置(主要包括成像室、LED 照明系统、背景板以及摄像机),图 1 为图像采集装置示意图。相机使用 iPhone12 的 1200 万像素后置摄像头,在拍摄时模拟实际分拣过程,不断翻转苹果以变换拍摄角度,采集到包含了苹果不同位面的图像共计 1600 幅。由于采集到的苹果图像主要为红富士,且外观种类较少,考虑到视觉系统的通用性,又通过互联网搜集了一些不同品种(红富士、嘎啦、蛇果、秦冠等)、不同外观的苹果图

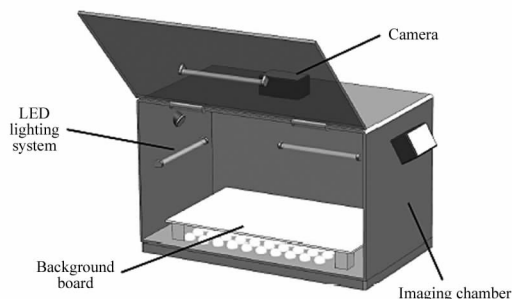


图 1 图像采集装置

Fig. 1 Image acquisition device

像,并将照片的背景统一通过图像预处理算法转换成白色,筛选出了1 000张图像。最终,通过实际拍

摄与互联网收集共计得到 2 600 张苹果图像,图 2 展示了收集到的部分图像。

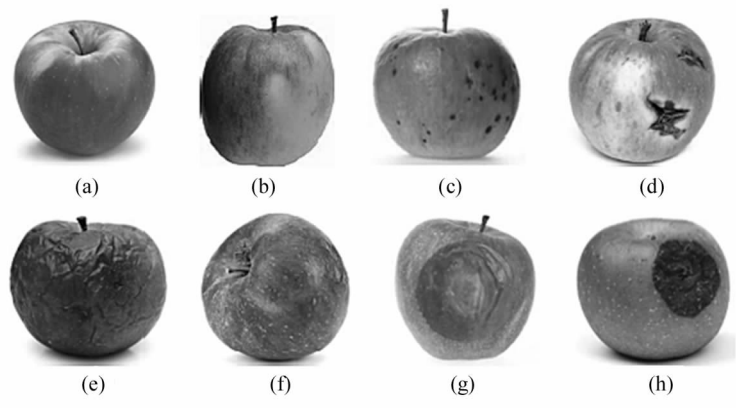


图 2 不同外观的苹果图像:(a) 合格苹果;(b) 着色差苹果;(c) 病斑苹果;(d) 伤疤苹果;(e) 干瘪苹果;(f) 腐烂苹果;(g) 轮纹病苹果;(h) 炭疽病苹果

Fig. 2 Apple images with different appearances:(a) Qualified apple;(b) Poor coloring apple;(c) Diseased apple;(d) Scar apple;(e) Dried apple;(f) Rotten apple;(g) Apple with ring rot;(h) Anthracnose apple

2.2 数据集制作

将收集到的 2 600 张苹果图像按照其外观分为 4 类,分别为合格苹果,类别标签设定为 Qualified_apple、颜色缺陷苹果(着色差、成熟度不够等),类别标签设定为 Color_defect_apple、表皮缺陷苹果(病斑、机械损伤、伤疤等),类别标签设定为 Surface_defect_apple、不合格苹果(腐烂、失水干瘪、轮纹病、炭疽病等),类别标签设定为 Unqualified_apple。其中,合格苹果为优等苹果可正常售卖;缺陷苹果其外观存在瑕疵,但大部分该类苹果内质并未受到损坏,可以进行再加工处理或低价售卖,存在一定价值;不合格苹果无法食用,一般没有利用价值,通常会被丢弃。在标定时,部分苹果可能既存在病变或腐烂,又存在表皮缺陷,对于该类苹果将其标定为不合格苹果,类别优先级从低到高分别为合格苹果、颜色缺陷苹果、表皮缺陷苹果以及不合格苹果,当某个苹果果实同时符合多个类别时,则以优先级高的类别作为该苹果的类别;此外,为增加训练后模型的容错率,对于苹果表皮缺陷特征极其不明显的,如只在边缘处有极少量缺陷的苹果,此类苹果将被标定为合格苹果。最后,随机选取其中 2 000 张图像作为训练集,300 张图像作为验证集,300 张图像作为测试集,并将分好的图像根据训练集或验证集以及不同类别苹果的类别名称分别放置到对应的文件夹中,完成数据集的制作,用于后续训练。

2.3 数据增强

由于收集到的图像数量有限,并且不同类别的苹果图像数量不同,合格苹果数量较多,缺陷苹果和不合格苹果数量较少。因此,为了增加数据的多样性,降低网络过拟合的风险,并避免由于类别间样本不平衡导致对少样本类别分类精度低的问题,采用随机翻转、旋转、平移、添加噪声(椒盐噪声和高斯噪声)等 4 种方法的随机组合对不同类别的图像进行有针对性的扩增,使数据集的分布尽可能地均衡,从而提高训练效果,进一步增强模型的泛化性和鲁棒性^[13]。增强后数据集的样本数量如表 1 所示。

表 1 数据增强后的苹果数据集数量

Tab. 1 Number of apple data sets after data enhancement				
Data set	Train set	Validation set	Test set	Total
Qualified	3 867	670	660	5 197
Color defect	3 728	632	649	5 009
Surface defect	3 672	676	640	4 988
Unqualified	3 712	640	672	5 024

3 基于 ResNet 苹果分类模型的改进

3.1 双通道 SE 模块

ResNet50 的特征提取网络中大量使用了残差连接,其网络结构主要由一系列残差块(residual block)串联而成,残差块结构如图 3 所示。然而,在实际分拣任务中,果实外观品质分类存在类间差异小,类内

差异大的问题,限制了模型的分类精度^[14]。在分拣过程中,实时性是最重要的性能指标之一,因此,在不增加网络计算量的前提下,提升模型的参数利用率,从而进一步提升分类精度,是非常必要的。

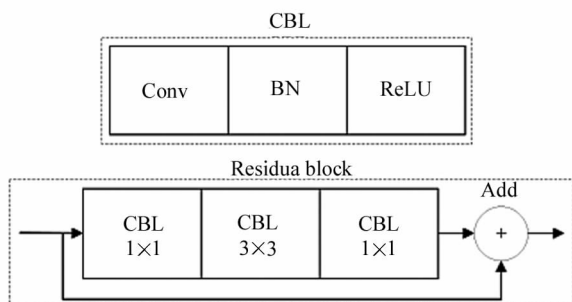
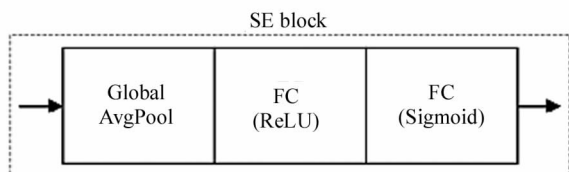


图3 ResNet模型的残差块

Fig. 3 Residual block of ResNet model

SENet^[15] (squeeze-and-excitation networks) 是一种先进的注意力机制模型,其核心是 SE 模块 (squeeze-and-excitation block, SE Block),如图4所示。SE 模块是一种加在通道维度上的注意力机制,其结构主要包括一个全局平均池化层和两个全连接层。这种注意力机制可以通过学习的方式自动获取到每个特征通道的重要程度,然后依照这个重要程度去增强有效的通道特征,并抑制对当前任务用处不大的低效或无效的通道特征,即学习通道之间的相关性,从而对不同通道进行权重分配,有助于提高特征图的表达能力,进而提高模型的识别精度。



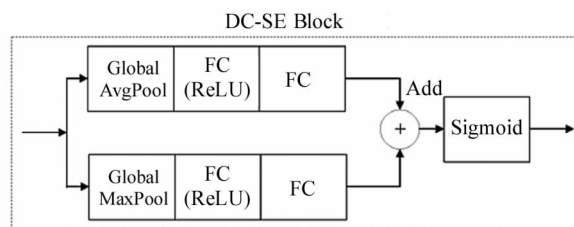
Global AvgPool: global average pooling;
FC: the fully connected layers;
ReLU and Sigmoid: the activation functions
used by the current fully connected layers

图4 SE模块

Fig. 4 SE module

本文在 SE 模块的基础上,提出了一种双通道 SE 模块 (dual channel squeeze-and-excitation block, DC-SE Block),如图5所示。与 SE 模块的不同之处在于 DC-SE Block 同时采用了全局平均池化以及全局最大池化;平均池化能更好地保留果实的整体特征,而最大池化可以更好保留果实的局部特征,因

此,两种不同池化层的并行连接可以提取到更加丰富的高层次特征。首先,将输入的特征图分别经过全局最大池化和全局平均池化,得到两个 $1 \times 1 \times C$ (C 为特征图通道数) 的特征向量;然后,将它们分别送入两个两层神经网络中建模通道之间的相关性,第一层神经元个数为 C/r (r 为减少率),激活函数为 ReLU,第二层神经元个数为 C ;最后,将两个全连接层输出的特征向量进行加和操作,再经过 Sigmoid 激活函数进行激活,生成最终的权值向量。



Global MaxPool: global maximum pooling

图5 双通道 SE 模块

Fig. 5 Dual channel SE block

将 ResNet 的残差块与 DE-SE Block 结合,构成 DC-SE 残差块 (dual channel squeeze-and-excitation residual block, DC-SE Residual Block),如图6所示。改进后的残差块将 3×3 卷积得到的特征图首先经过 DC-SE Block 得到权值向量;然后,将得到的特征通道的权重通过乘法操作逐通道加权到原来的特征图上,完成在通道维度上的原始特征重标定;最终,将增强之后的特征图送入到后续网络中进行传播。

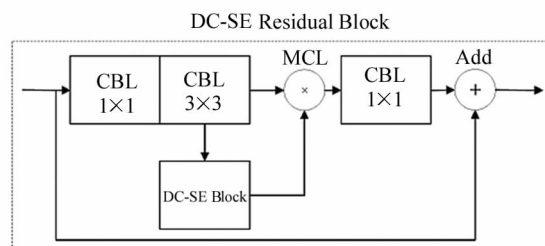


图6 DC-SE 残差块

Fig. 6 DC-SE Residual Block

3.2 Inception 模块

当果实的病斑、伤疤、腐烂等缺陷面积较大或较为密集时,模型可以比较容易地实现品质分类;而在实际分拣任务中,存在一些缺陷面积较小或较稀疏的情况,使得缺陷或不合格果实与合格果实在外观上有着极大的相似性,从而导致模型错误分类。基于此,在 ResNet 网络中加入 Inception 模块,这是 1

个由一系列卷积层和 1 个最大池化层组成的结构,如图 7 所示。Inception 模块将输入特征图分别经过 3 个不同大小的卷积组以及 1 个池化核大小为 3×3 的最大池化层进行卷积运算与池化,其中 1×1 卷积起到降维的作用,然后,将得到的 4 个特征图进行按通道拼接操作,最后,将拼接好的特征图传入到后续网络中,可以有效地融合多尺度特征,丰富感受野,提高最终特征图的表达能力,模型对较小病斑、伤疤等不明显缺陷的识别能力被进一步增强。

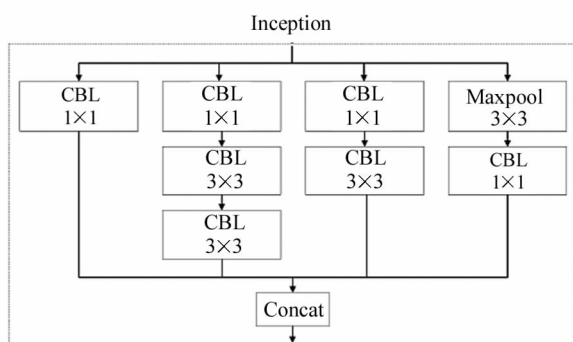


图 7 Inception 模块

Fig. 7 Inception module

3.3 网络模型改进

改进后的模型结构如图 8 所示。其网络结构主要由一系列残差块与 DC-SE 残差块堆叠而成,每个

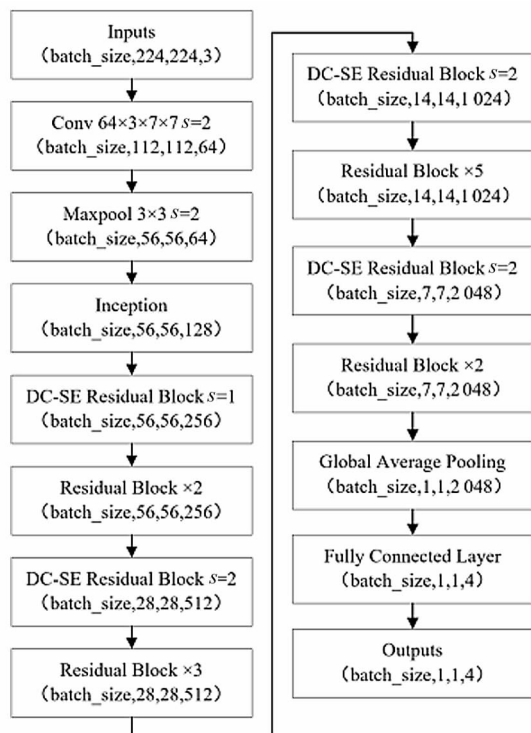


图 8 改进后的网络结构图

Fig. 8 Improved network structure diagram

残差块和 DC-SE Residual Block 又包含 3 个 CBL 模块, CBL 模块是卷积层、BN 层以及激活层的组合,激活函数使用 ReLU 激活函数。首先经过 1 个卷积层和 1 个最大池化层(步长 s 均为 2),进行 4 倍下采样;然后,经过 1 个 Inception 模块和 1 个 stride 为 1、3 个 stride 为 2 的 DC-SE Residual Block(其中 stride 为 1 和 2 代表的是 DC-SE Residual Block 中 3×3 卷积的步长)以及一系列传统残差块,每经过 1 个 stride 为 2 的 DC-SE Residual Block,特征图的尺寸将缩小为原来的一半,在经过一系列层结构进行 32 倍下采样之后,得到了大小为 $7 \times 7 \times 2048$ 的预测特征图;接着,经过 1 个全局平均池化层和 1 个全连接层,生成尺度为 $1 \times 1 \times 4$ 的预测特征向量(本模型需要预测 4 个类别,所以最终生成的预测特征向量维度为 4);最后,将预测特征向量送入 Softmax 分类器中,完成最终的类别预测。如果是训练过程,则每次都会将一个 Batch 的图像同时输入到网络模型中,进行前向传播,然后将得到的预测结果与真实值一同送入损失函数中,进行损失计算和反向传播,并基于梯度利用优化器调整模型得参数,对模型进行优化。

4 模型训练

4.1 评价指标

使用精确率(Precision, P)、召回率(Recall, R)、特异度(Specificity, S)以及准确率(Accuracy, Acc)这 4 个性能指标对模型进行性能评估,如式(1)~(4)所示。

$$P = \frac{TP}{TP + FP}, \quad (1)$$

$$R = \frac{TP}{TP + FN}, \quad (2)$$

$$S = \frac{TN}{TN + FP}, \quad (3)$$

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}, \quad (4)$$

式中, P 为模型预测的所有正样本中,预测正确的比例; R 为所有真实正样本中,模型预测正确的正样本比例; S 为所有真实负样本中,模型预测正确的负样本比例; Acc 为模型正确分类样本数占总样本数的比例; TP 为预测为正,真值为正的样本数量; FP 为预测为正,真值为负的样本数量; FN 为预测为负,真值为正的样本数量; TN 为预测为负,真值为负的样本数量。

4.2 训练过程

训练开始前,为了减少模型训练所需要的时间

和资源,并降低过拟合的风险,增强模型的泛化能力和训练效果,利用 ImageNet 大规模数据集对改进后的模型进行预训练,并使用预训练好的权重对模型初始化^[16]。使用 Adam 优化器对网络进行优化,在训练时使用 Warmup,即在训练初期先使用一个较小的学习率初始值,使模型在一开始对数据分布理解为 0 时,不至于对 Batch size 的小批量图像过拟合,本文将初始值设置成 0.000 1,在前 3 轮训练过程中,逐渐上升到 0.001,并从第 4 轮开始逐渐衰减至 0.000 1, Batch size 设置为 32, Epoch 设置为 100, DC-SE Block 中的减少率 r 设置为 4。每训练完一个 Epoch 保存一次网络的权重文件,并利用 Tensorboard 工具对模型进行实时监测,记录模型的损失函数值和准确率,训练共耗时 2 h 45 min。

4.3 训练数据分析

训练过程中的损失曲线如图 9 所示,在训练初期,6 种算法的损失值迅速减小至一个较小值,然后缓慢下降,最终趋于稳定。本文算法和 ResNet50 相较于其他 4 种算法收敛速度更快,变化更平稳,损失值更低。此外,本文算法损失值略低于改进前的 ResNet50,表现出了最优的训练效果。

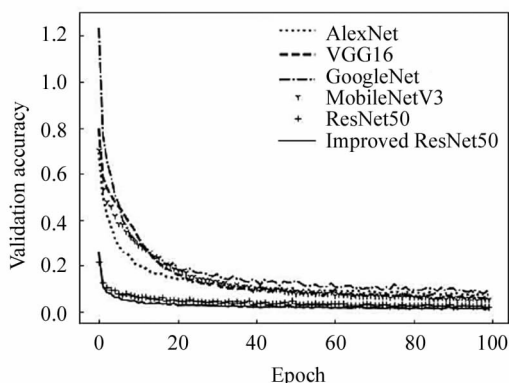


图 9 损失曲线

Fig. 9 Loss curve

训练过程中验证集的准确率变化曲线如图 10 所示,在训练初期,6 种算法的准确率变化幅度较大,但是总体趋势是增长的,在之后的训练过程中,逐渐趋于稳定,在小范围内振荡。本文方法相较于其他方法可以在更短的时间内达到更高的分类准确率,且变化平稳,波动较小。将模型取得准确率最大值时对应的训练轮次所得到的权重作为最终的模型权重,本文算法在第 51 个 Epoch 得到最高的准确率,值为 99.8%。

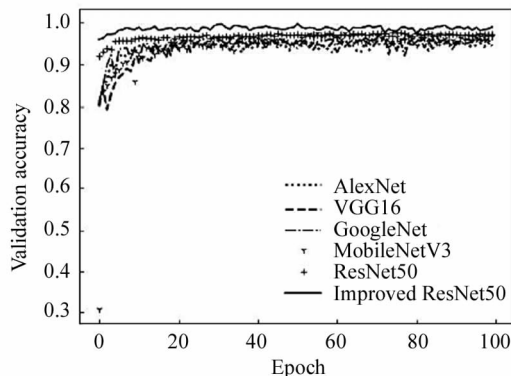


图 10 验证集准确率曲线

Fig. 10 The accuracy curve of validation set

5 试验结果

5.1 性能评估

利用测试集对最终获得的模型进行测试,本文的测试集中包含了 4 类苹果图像,基本符合整体数据集的分布。混淆矩阵可以直观地评估模型的优劣,如图 11 所示,对角线上颜色较深的部分是分类正确的样本数量,对角线之外的部分是分类错误的样本数量。可以看出,2 621 个测试样本中,共有 8 个样本被误分类,准确率 Acc 为 99.7%,其中,两个颜色缺陷苹果被误认为是合格苹果,1 个合格苹果被误认为是颜色缺陷苹果,这两类存在混淆;两个不合格苹果被分别误分类成合格苹果和表皮缺陷苹果,还有 3 个表皮缺陷苹果被误分类成合格苹果和不合格苹果,这 3 类也会发生混淆。分类错误的主要原因就是这些苹果的外观及其相似,类间差异较小。

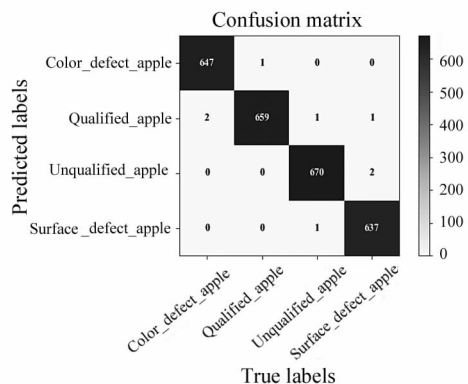


图 11 本文算法在测试集上的混淆矩阵

Fig. 11 Confusion matrix of the paper algorithm on test set

根据混淆矩阵和性能指标计算公式,可得到模型在测试集上 4 个类别的各项性能指标,如表 2 所示。两类缺陷苹果的精确率最高,可达到 99.8%,合

格苹果的精确率最低,为 99.4%;召回率最高可达到 99.8%,为合格苹果类别,表皮缺陷苹果召回率最低,为 99.5%;特异度除合格苹果为 99.8%以外,其他 3 个类别均可达到 99.9%;平均识别速度为 32.3 帧/s。本文提出的算法对 4 个类别的分类精度和分类速度均满足工业需求,故可以应用于实际分拣生产线环境中。

表 2 本文算法在测试集上的各项性能指标

Tab. 2 Each performance index of the paper algorithm on test set

Classification	P/%	R/%	S/%	Recognition speed /frame · s ⁻¹
Qualified_apple	99.4	99.8	99.8	33.7
Color_defect_apple	99.8	99.7	99.9	32.5
Surface_defect_apple	99.8	99.5	99.9	31.4
Unqualified_apple	99.7	99.7	99.9	31.7
Average	99.7	99.7	99.9	32.3

5.2 不同算法的分类结果

利用本文制作的数据集对另外 5 种目前应用较多的分类算法进行训练,得到最优模型之后,在测试集上进行测试,6 种方法的准确率 Acc、精确率 P、召回率 R、特异度 S 以及识别速度如表 3 所示。在识别精度方面,本文方法的精确率 P 为 99.7%、召回率 R 为 99.7%、特异度 S 为 99.9%、准确率 Acc 为 99.7%,均高于其他 5 种算法,比 ResNet50 的 Acc 提升了 1.2 个百分点;在识别速度方面,本文方法的平均分类速度为 32.3 帧/s,仅比 ResNet50 低了 3.4 帧/s,能够对生产线上的果实进行实时分类,满足实际分拣工作的实时性要求。综上,在所有方法中,本文所提出的方法表现出最好的综合分类性能。

表 3 6 种算法在测试集上的各项性能指标

Tab. 3 Each performance index of six algorithms on test set

Algorithm	P/%	R/%	S/%	Acc/%	Recognition speed /frame · s ⁻¹
AlexNet	96.1	95.9	98.7	96.2	58.8
VGG16	96.9	97.0	98.9	97.0	45.5
GoogleNet	97.1	97.1	99.1	97.4	33.3
MobileNetV3	97.6	97.2	99.2	97.6	35.8
ResNet50	98.3	98.7	99.5	98.5	35.7
Proposed algorithm	99.7	99.7	99.9	99.7	32.3

5.3 消融试验

本节建立了不同改进模型的对比试验,按其结构不同,分为原始 ResNet50、只加入 Inception 模块、只加入 DC-SE Block、只加入 SE 模块以及 Inception

模块分别和两种注意力机制的组合模型,将上述 6 种模型在测试集上进行测试,得到各项性能指标,如表 4 所示。与原始 ResNet50 相比,在仅加入 Inception 模块的情况下,Acc 提升了 0.2 个百分点,各项性能指标提升较小;两种仅加入注意力机制的模型都取得了较好的分类结果,各项性能指标提升明显,其中,DC-SE Block Acc 提升了 0.9 个百分点,SE 模型 Acc 提升了 0.5 个百分点,验证了在模型中加入注意力机制的有效性,也说明了本文提出的 DC-SE Block 可以比原始 SE 模块提取到更加丰富的特征,从而实现更好的分类效果;两种组合模型的 Acc 分别提升了 1.2 个百分点和 0.7 个百分点,其中 Inception 模块与 DC-SE Block 的组合模型显示出更好的分类性能,也说明了两种模块的组合比单独加入其中任何一种模块的效果都要好;此外,6 种模型的分

类速度并无太大差距,均满足实时性要求。综合结果表明,5 种改进模型整体性能均优于未改进的 ResNet50,且本文提出的改进模型具有最佳的分类性能,验证了所提出改进方法的有效性。

表 4 不同改进模型在测试集上的各项性能指标

Tab. 4 Each performance index of different improved models on test set

Algorithm	P/%	R/%	S/%	Acc/%	Recognition speed /frame · s ⁻¹
ResNet50	98.3	98.7	99.5	98.5	35.7
Inception	98.5	98.8	99.6	98.7	35.5
DC-SE	99.3	99.4	99.7	99.4	32.6
SE	98.9	98.9	99.7	99.0	32.7
Inception+DC-SE	99.7	99.7	99.9	99.7	32.3
Inception+SE	99.1	99.2	99.8	99.2	32.5

5.4 特征图可视化

为了更好地了解本文所提出模型的特征提取能力,本节试验可视化了其部分层结构的输出特征图,以一个不合格苹果样本为例,如图 12 中(a)所示,将图像输入到模型中进行测试,然后依次提取出 Inception 模块以及第 2、第 5、第 10、第 12 个残差块的前 12 个通道的输出特征图,分别如图 12 中(b)~(f)所示。根据特征图的明暗程度,可以理解不同卷积核关注的特征信息,特征图中亮度越高的部分就是卷积核越感兴趣的部分。通过可视化结果,可以清晰地看到一些特征图对果实整体特征和腐烂区域特征有很好的响应,说明模型能够提取到丰富的全局特征和局部特征,并且可以增强有效的通道特征(特征图明亮),抑制低效或无效的通道特征(特征图较

暗)。通过同一模块不同通道的输出特征图可以看出,同一个层结构中的不同卷积滤波器可以提取目标对象不同部分的特征,叠加多个卷积核后,会得到更加丰富的果实特征信息,目标属性的表达会更加充分和全面。此外,随着网络的加深,所得到的特征信息也越来越抽象,其中,(b)图主要是在提取果实

的形状和轮廓等浅层特征;(c)图和(d)图则更加关注果实的纹理、边缘等更为细节的特征;最后,(e)图和(f)图是更加抽象的高级语义特征,其中(f)图是模型最终得到的输出特征图,用于全局平均池化后,输入到全连接层和分类器中,完成最终的果实分类。试验结果表明,该模型可以提取出非常丰富的果实特

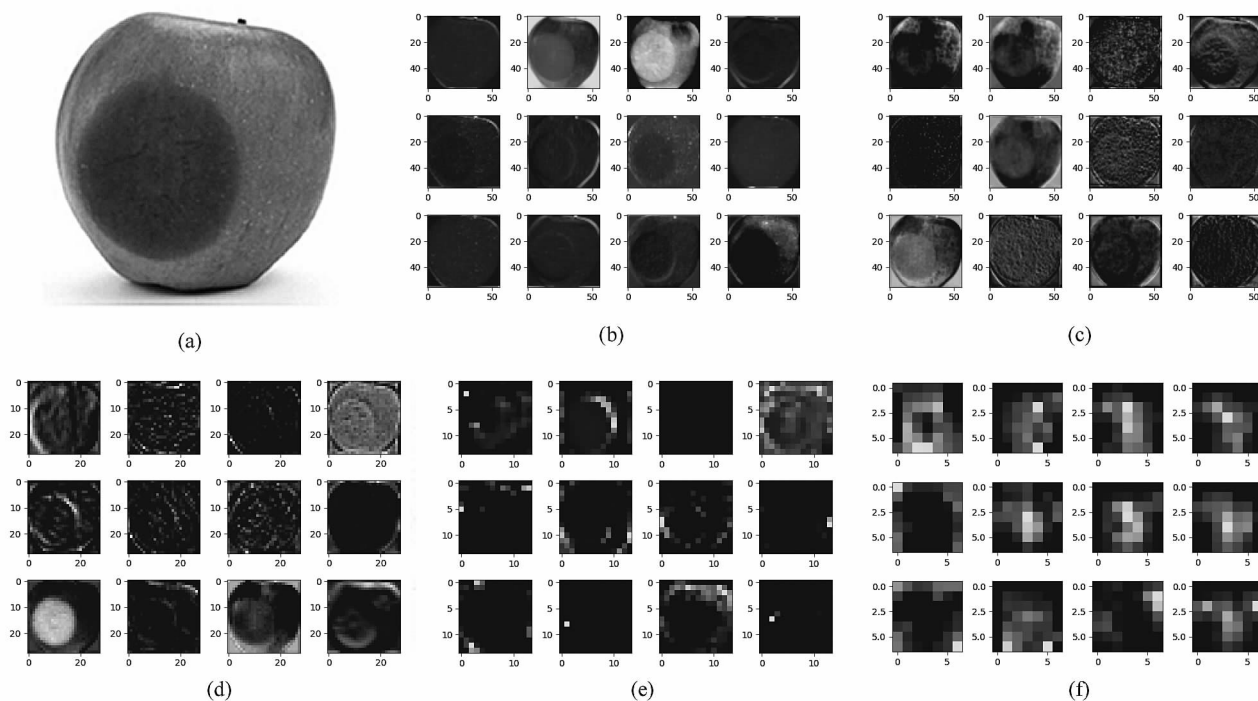


图 12 不同层结构的输出特征图:(a) 不合格苹果;(b) Inception 模块输出特征图;(c) 第 2 个残差块输出特征图;(d) 第 5 个残差块输出特征图;(e) 第 10 个残差块输出特征图;(f) 第 12 个残差块输出特征图

Fig. 12 Output feature map of different layer structures:(a) Unqualified apple;(b) Output feature map of Inception module;(c) Output feature map of the second residual block;(d) Output feature map of the fifth residual block;(e) Output feature map of the tenth residual block;(f) Output feature map of the twelfth residual block

征,对不同类别的果实具有很好的识别性能。

6 结 论

本文针对传统方法在分拣生产线环境下对果实外观品质分类的准确率和实时性较低、鲁棒性较差等问题,提出了一种基于改进 ResNet 的分类方法,该方法在测试集下的精确率 P 为 99.7%,召回率 R 为 99.7%,特异度 S 为 99.9%,准确率 Acc 为 99.7%,检测速度可达 32.3 帧/s,整体性能优于其他 5 种算法,能够达到苹果外观品质分类准确性和实时性的要求,通过更改数据集,本文模型可应用于其他各类水果分拣生产线的视觉系统中,具有良好的泛化性。

为了进一步验证本文算法的有效性和可行性,

设计了不同改进模型的性能对比试验。结果表明,Inception 模块和 DC-SE Block 的组合改进模型拥有比其他改进模型更好的分类性能,并且本文提出的 DC-SE Block 对模型性能的提升也要优于普通 SE 模块,证明了所提出改进方法的有效性。此外,通过可视化特征图,进一步评估了所提出模型的特征提取能力,以更好地理解卷积层如何变换其输入以及每个卷积滤波器的含义。

参考文献:

- [1] HUO Y Q, TANG J L, YIN X Z, et al. Apple disease identification method based on compressed sensing theory[J]. Journal of Agricultural Machinery, 2013, 44(10): 227-232. 霍迎秋,唐晶磊,尹秀珍,等.基于压缩感知理论的苹果

- 病害识别方法[J]. 农业机械学报, 2013, 44(10): 227-232.
- [2] NIE M, ZHAO Q, BI S, et al. Apple external quality analysis based on BP neural network[C]//2019 1st International Conference on Industrial Artificial Intelligence (IAI), July 23-27, 2019, Shenyang, China. New York: IEEE, 2019: 1-5.
- [3] LI S J, HU D Y, GAO S M, et al. Real-time classification and detection of citrus based on improved single short multibox detector[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(24): 307-313.
- 李善军, 胡定一, 高淑敏, 等. 基于改进 SSD 的柑橘实时分类检测[J]. 农业工程学报, 2019, 35(24): 307-313.
- [4] AZDC A, HE Hf A, JAF B. Computer vision based detection of external defects on tomatoes using deep learning-science direct[J]. Biosystems Engineering, 2020, 190: 131-144.
- [5] YUE Y J, TIAN B K, WANG H J, et al. Application of improved VGG model in apple appearance classification[J]. Science Technology and Engineering, 2020, 20(19): 7787-7792.
- 岳有军, 田博凯, 王红君, 等. 改进 VGG 模型在苹果外观分类中的应用[J]. 科学技术与工程, 2020, 20(19): 7787-7792.
- [6] MOMENY M, JAHANBAKHSI A, JAFARNEZHAD K, et al. Accurate classification of cherry fruit using deep CNN based on hybrid pooling approach[J]. Postharvest Biology and Technology, 2020, 166: 111204.
- [7] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.
- [8] LIU T Z, TENG G F, YUAN Y C, et al. Fruit recognition method of winter jujube in natural scene based on improved YOLOv3[J]. Journal of Agricultural Machinery, 2021, 52(5): 17-25.
- 刘天真, 滕桂法, 苑迎春, 等. 基于改进 YOLOv3 的自然场景下冬枣果实识别方法[J]. 农业机械学报, 2021, 52(5): 17-25.
- [9] XIAO J W, TIAN J W, WANG Q, et al. Fruit disease classification method based on improved residual network[J]. Computer Engineering, 2020, 46(9): 221-225.
- 肖经纬, 田军委, 王沁, 等. 基于改进残差网络的果实病害分类方法[J]. 计算机工程, 2020, 46(9): 221-225.
- [10] ZHANG L, JIA J, GUI G, et al. Deep learning based improved classification system for designing tomato harvesting robot[J]. IEEE Access, 2018, 6(99): 67940-67950.
- [11] LIU B, DING Z, TIAN L, et al. Grape leaf disease identification using improved deep convolutional neural networks[J]. Frontiers in Plant Science, 2020, 11: 1082.
- [12] XING S, LEE M, LEE K K. Citrus pests and diseases recognition model using weakly dense connected convolution network[J]. Sensors (Basel, Switzerland), 2019, 19(14): 3195.
- [13] ORANO J, MARAVILLAS E A, ALIAC C. Jackfruit fruit damage classification using convolutional neural network[C]//2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), November 29-December 01, 2019, Laoag, Philippines. New York: IEEE, 2019: 1-6.
- [14] WANG S H, CHEN Y. Fruit category classification via an eight layer convolutional neural network with parametric rectified linear unit and dropout technique[J]. Multimedia Tools and Applications, 2020, 79(4): 15117-15133.
- [15] PONCE J M, AQUINO A, ANDUJAR J M. Olive-fruit variety classification by means of image processing and convolutional neural networks[J]. IEEE Access, 2019, 7: 147629-147641.
- [16] WU L, ZHANG H, CHEN R, et al. Fruit classification using convolutional neural network via adjust parameter and data enhancement[C]//2020 12th International Conference on Advanced Computational Intelligence (ICACI), August 14-16, 2020, Dali, China. New York: IEEE, 2020: 294-301.

作者简介:

乔艳军 (1996—), 男, 硕士研究生, 主要从事机器视觉、数字图像处理、模式识别、农业机器人方面的研究。