

# 基于轻量化卷积神经网络的苹果表皮损伤分级研究

付夏晖, 王菊霞\*, 崔清亮, 张燕青, 王毅凡, 阴 妍

(山西农业大学农业工程学院, 山西太谷 030801)

**摘要:**【目的】苹果在销售过程中,其表皮的损伤情况会直接影响果实的经济价值。运用相机采集苹果表皮的损伤图像,对获取到的图像进行分类和数据预处理,基于迁移学习的方法对苹果表皮损伤进行直接分级研究,为提高苹果表皮损伤分级效率进而更好地指导苹果采收后的分类售卖提供理论依据。【方法】首先对采集到的富士和丹霞两个苹果品种图像进行对比度调整、旋转、翻转、添加噪声等11种批量操作,将数据集扩充到9360张,同时对扩充后的样本集统一调整为224×224像素。针对预处理好的数据集,选取5种20 MB以下的轻量化卷积神经网络在相同超参数设置条件下进行初始化训练、引入迁移学习训练以及在迁移学习基础上增加冻结网络层权重3种方法进行训练对比。【结果】5种网络初始化训练后的测试精度仅为56.32%~71.98%;基于迁移学习的MobileNet-v2模型最终训练精度达99.04%,在轻量级卷积神经网络中,比表现性能最差的EfficientNet-b0模型最终训练精度高18.79%;在基于迁移学习的MobileNet-v2模型基础上冻结不同模块参数,得出模型选择冻结至第1个卷积模块到Bottleneck 3-1模块时均可在缩短模型训练时间的基础上提高模型验证精度,其中在冻结到Bottleneck 3-1模块时比基于迁移学习的MobileNet-v2模型训练时间缩短了29.32%,同时验证精度提高了0.93%,测试精度提升了1.12个百分点达91.58%,检测单张图片所用平均时间为0.14 s,网络大小为8.15 MB,可以满足快速识别需求。【结论】基于迁移学习加冻结训练的MobileNet-v2模型具有较好的鲁棒性和分级性能,可为移动终端和嵌入式设备在苹果损伤直接分级方面提供技术参考。

**关键词:** 苹果表皮; 损伤分级; 轻量化; 迁移学习; 冻结训练

中图分类号: S661.1 文献标志码: A 文章编号: 1009-9980(2023)10-2263-12

## Research on apple epidermal damage grading based on lightweight convolutional neural network

FU Xiahui, WANG Juxia\*, CUI Qingliang, ZHANG Yanqing, WANG Yifan, YIN Yan

(College of Agricultural Engineering, Shanxi Agricultural University, Taigu 030801, Shanxi, China)

**Abstract:** 【Objective】 In the process of apple selling, the damage of its epidermis will directly affect the economic value of the fruit. The presence and severity of apple surface damage directly affect the sales link, and customers often care about the epidermis damage when choosing apples. At present, most studies focus on apple size, color and appearance classification, and the use of high-end instruments to detect the damage inside the apple, while the study on the direct classification of surface damage is rare. The camera was used to collect apple epidermis damage images, classify and preprocess the acquired images, and conduct a direct classification on apple epidermis damage based on transfer learning method, which can provide a theoretical basis for improving the classification efficiency of apple epidermis damage and guiding the classification and apple sale after harvesting. 【Methods】 Firstly, the camera was used to collect the top, side and bottom images of Fuji and Danxia apples to form the first-stage data set. Then, 11 batch of operations, such as contrast adjustment, rotation, flip and noise addition, were carried out to expand the data set to 9360 pieces to form the second-stage data set. At the

收稿日期: 2023-05-24 接受日期: 2023-07-26

基金项目: 国家自然科学基金项目(11802167); 山西省重点研发计划项目(202102020101012); 山西省应用基础研究项目(201901D211364)

作者简介: 付夏晖, 男, 在读硕士研究生, 主要从事机器学习、农产品加工及智能装备研究。Tel: 13294552666, E-mail: 13294552666@163.com

com

\*通信作者 Author for correspondence. Tel: 18634418916, E-mail: wangjuxia79@163.com

same time, the expanded sample set was uniformly adjusted to  $224 \times 224$  pixels to form the final data set. According to the ratio of 7:3:3, the preprocessed data set was divided into training set, verification set and test set. Five lightweight convolutional neural networks less than 20 MB, MobileNet-v2, SqueezeNet, ShuffleNet, NASNet-Mobile and EfficientNet-b0, were selected for initial training, introduction of migration learning training and migration learning under the same super-parameter Settings On a Bottleneck basis, and three methods were added for detailed freezing network layer weights (the MobileNet-v2 network structure is specifically divided into 21 modules for freezing training, which contain 3 convolutional modules, 1 average pooling module, and 17 Bottleneck modules). **【Results】** The test accuracy of the five kinds of networks after initial training was only 56.32%–71.98%. The final training accuracy of MobileNet-v2 model based on transfer learning was 99.04%, 18.79% higher than that of the worst EfficientNet-b0 model among lightweight convolutional neural networks. After freezing different module parameters on the basis of the MobileNet-v2 model were based on transfer learning, it was concluded that models Bottleneck 3-1, when they select to freeze to the first convolutional module, can shorten model training time and improve model validation accuracy. When Bottleneck 3-1 module was frozen, the training time for Bottleneck 3-1 was shortened by 29.32% compared to MobileNet-v2 model based on transfer learning, the verification accuracy increased by 0.93%, and the test accuracy increased by 1.12 percentage points to 91.58%. The average time for detecting a single image was 0.14 s. The network size was 8.15 MB, which can meet the requirements of fast identification. The final training loss value of the MobileNet-v2 model based on transfer learning was less than 0.04, which was 0.5 lower than that of the worst performing EfficientNet-b0 model in lightweight convolutional neural networks. The test results showed the recall rate and precision rate of MobileNet-v2 confusion matrix diagram based on transfer learning and five kinds of lightweight convolutional neural networks were based on transfer learning in the test set. Among them, the MobileNet-v2 model based on transfer learning had the best performance, and the recall rate of 6 types of data in the test set ranged from 89.40% to 100%. The precision ranged from 53.52% to 99.78%. The Grad-CAM visualization comparison of the trained network showed that the SqueezeNet model based on transfer learning had the worst visualization effect and the lowest recognition accuracy. The visualization effect of NASNet-Mobile model based on transfer learning was poor. It can only display a large range of concern areas, and the recognition degree of some pictures was not high. The visualization effect of the MobileNet-v2 model based on transfer learning was obviously better than the previous two models, but the key areas identified by the model were different from the reality. A MobileNet-v2 model based on transfer learning tended to have the best visualization effect on a network that was Bottleneck 3-1 when it was frozen to a Bottleneck 3-1 module, and the key areas identified by the model had the highest compatibility with the actual situation. **【Conclusion】** In this study, five kinds of lightweight models with Bottleneck 3-1 were selected for initialization training and transfer learning training, and it was concluded that MobileNet-v2 model with transfer learning had the best effect. Then, the freezing strategy was used for hierarchical training. The verification accuracy reached 92.23% when Bottleneck 3-1 was frozen. The test accuracy was 91.58%, the average recognition time was 0.14 s, and the network size was 8.15 MB, which can provide technical reference for mobile terminals and embedded devices in the direct classification of apple fruit damage.

**Key words:** Apple epidermis; Damage classification; Lightweight; Transfer learning; Freezing training

苹果作为人类膳食营养的重要来源之一,其含水量高,口感酸甜,具有促进食欲、降低心血管疾病及冠心病发病率等功效,深受大众喜爱<sup>[1]</sup>。苹果的品质直接决定着果实的价格和销量,但苹果在采摘、运输、包装和售卖等环节中会存在不同程度的挤压、摩擦、碰撞,导致苹果表皮发生不同程度的损伤<sup>[2]</sup>。损伤是影响苹果品质的重要因素之一,苹果表面损伤产生缺陷后,缺陷部位会加速腐烂,散发出更多的催熟激素,造成整批次苹果品质的降低<sup>[3]</sup>。苹果损伤的检测方法,一方面常用人工的方法来辨别苹果的损伤等级,使得分级稳定性较差且不明确<sup>[4]</sup>;另一方面采用光谱仪器、红外热成像仪等高端设备对损伤进行检测,检测成本相对较高。因此,研发苹果表皮快速、低成本、高精度的损伤直接分级检测方法,为丰富苹果表皮品质检测方法提供技术参考,同时对减少商家损失、优化苹果售卖品质等方面具有深远的意义。

国内外学者在苹果的损伤检测和分级方面进行了大量研究,研究主要采用光谱仪器、红外热成像仪、磁共振分析仪等高端设备对苹果的损伤进行检测。在利用光谱仪器研究方面,沈宇等<sup>[5]</sup>和蒋金豹等<sup>[6]</sup>运用高光谱仪器分别基于特征波段和高光谱单元完成对苹果表面的轻微损伤进行无损检测,检测正确率均在90%之上;Lu等<sup>[7]</sup>开发了一种基于液晶可调谐滤光片的多光谱成像系统用于苹果损伤检测,总体检测误差为11.7%~14.2%;Keresztes等<sup>[8]</sup>研发了一种基于HSI的短波红外苹果早期伤痕实时检测系统,能够同时检测30个苹果的早期损伤,识别准确率达到98%;邵志明等<sup>[9]</sup>利用近红外相机和阈值分割法对苹果图像进行分割和缺陷提取达到损伤检测目标,对即时损伤和损伤0.5 h的判别准确率均超过90%;Fan等<sup>[10]</sup>提出了一种基于Yolo v4深度学习算法和近红外图像的苹果缺陷检测方法,在每秒检测5个苹果基础上将平均检测准确率提升至93.9%。在运用红外热成像仪研究方面,门洪等<sup>[11]</sup>借助红外热像仪采集苹果的损伤样本图片,根据苹果不同部位和损伤的温差范围研究其识别方法,其中轻微损伤和重度损伤分别在1.3~2.6 °C和2.6~3.2 °C之间;周其显<sup>[12]</sup>针对苹果热力学图像的不同区域温度和降温曲线两方面特性进行分析,从传热学角度检测苹果的早期损伤,缺陷判别准确率为87.5%。在利用磁共振分析仪检测苹果表皮损伤方面,熊婷<sup>[13]</sup>运用低场核磁共振加权像和图像伪彩色处理技术,结

果表明,在重复时间为1500 ms、回波时间为200 ms时得到的T<sub>2</sub>加权像可正确检测苹果的表皮损伤。

近年来,神经网络被大量研究者运用到农作物的智能识别当中。为提高北方日光温室番茄产量预测结果,尹义志等<sup>[14]</sup>利用小波神经网络预测温室番茄产量,模型预测结果与实际产量平均相对误差仅1.02%;针对番茄生产中作业背景复杂、枝叶遮挡、光照分布不均匀的问题,陈新等<sup>[15]</sup>将MobileNetV3模块引入SSD(single shot multibox detector)算法中,相比原始SSD算法,番茄花果的平均识别率提高了7.9%;为提高苹果损伤判别度,宁景苑等<sup>[16]</sup>通过自主搭建的弛豫光谱采集系统采集光谱信号,使用标准正态变量交换算法(standard normal variable transformation, SNV)优化光谱数据,基于反向传播神经网络算法构建苹果损伤检测模型对苹果机械损伤进行检测,准确率达91.48%;针对现有苹果损伤检测仪器价格高、体积大的问题,Ning等<sup>[17]</sup>提出基于松弛单波长激光和卷积神经网络的富士苹果损伤检测方法,预测准确率达93%;为提高苹果产品质量和生产效率,Ismail等<sup>[18]</sup>研究词袋(bag-of-words, BOW)、空间金字塔匹配(spatial pyramid matching, SPM)和卷积神经网络(convolutional neural networks, CNN)三种图像识别方法用于苹果表皮缺陷的识别和分级,基于支持向量机的空间金字塔匹配算法最优,识别准确率达98.15%。

消费者在购买苹果时会着重挑选无伤和轻微损伤的苹果,因而苹果表皮的损伤情况会直接影响其售卖环节的经济效益。目前针对苹果的检测和分级研究主要是采用高端昂贵仪器对苹果的损伤进行检测,对苹果表皮损伤进行直接检测分级的研究较少。鉴于上述研究,笔者在本研究中提出一种基于迁移学习的苹果表皮损伤分级方法。首先使用相机采集苹果表皮的早期损伤图像,然后对获取到的图像进行分类和数据预处理,最后选择轻量化的卷积神经网络MobileNet-v2,运用迁移学习和冻结训练技术训练模型,完成苹果表皮损伤分级,提升苹果表皮损伤的分级效率,并为表皮损伤检测方面的移动终端和嵌入式设备提供技术参考。

## 1 材料和方法

### 1.1 图像采集

我国苹果品种评价标准中的苹果等级规格把果

面缺陷分为3类:极轻微、轻微和有部分缺陷,而在欧盟的苹果等级标准中进一步说明了果面轻微缺陷的指标:轻微瘀伤面积不超 $1\text{ cm}^2$ ,条状缺陷不超 $2\text{ cm}$ ,其他缺陷总面积不超 $1\text{ cm}^2$ <sup>[19]</sup>。笔者在本研究中以欧盟苹果等级标准为基础,极轻微缺陷代表无伤,有轻微缺陷代表轻伤,超过轻微缺陷标准代表重伤。

试验苹果全部来自于山西省农业科学院果树研

究所,选取丹霞和富士2个品种。图像采集设备为iPhone Xs Max,考虑图像多样性,采集苹果顶部、侧部、底部3个方位图像,分为丹霞无伤185张、丹霞轻伤131张、丹霞重伤121张、富士无伤164张、富士轻伤157张、富士重伤图像22张,共780张,图像分辨率为 $3024\times 3024$ 像素,格式为png,2个品种苹果部分样本实例如图1所示。

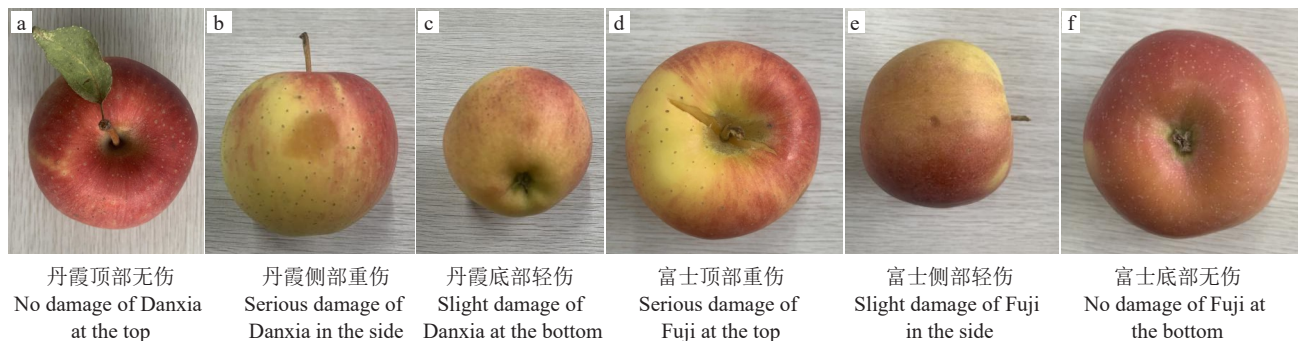


图1 部分苹果样本实例  
Fig. 1 Sample examples of some apples

## 1.2 图像预处理

由于初始数据样本量较少,为防止模型训练发生过拟合,对初始采集的780张图像统一进行左右翻转、上下翻转、随机旋转、随机平移、调整对比度、加入高斯噪声和椒盐噪声等操作将图像扩充至9360张。

在总体样本集中,适当增加验证集和测试集的比例,使用自编函数批量随机抽取图像,将训练集、

验证集和测试集划分为7:3:3,使得验证结果和测试结果更可靠,具体数量分别为5040幅、2160幅、2160幅。如图2所示,对图1-a丹霞顶部无伤图像进行左右翻转、上下翻转、随机旋转、随机平移4个操作;对图1-d富士顶部重伤图像增强对比度、减弱对比度、添加椒盐噪声、添加高斯噪声,得到的图像如图3所示;同时为适应模型训练,将扩充后的图像样本全部调整为 $224\times 224$ 像素。

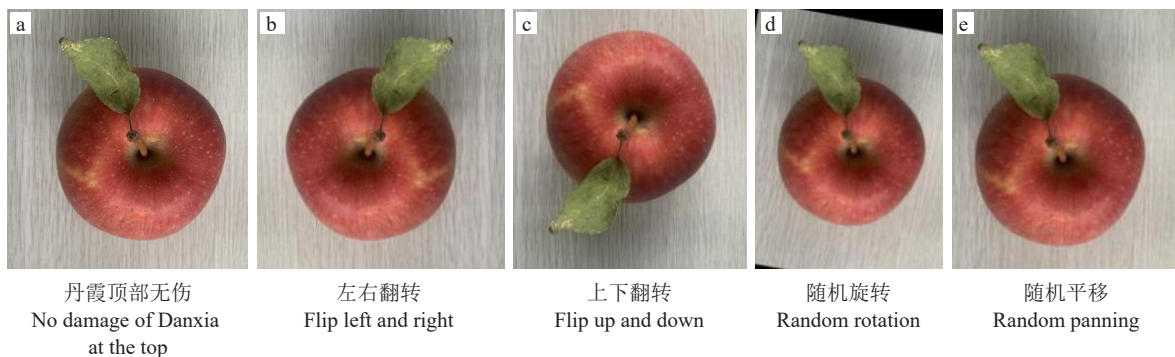


图2 对图1-a左右翻转、上下翻转、随机旋转、随机平移  
Fig. 2 Flip left and right, flip up and down, rotate randomly, and translate randomly for Fig. 1-a

## 1.3 模型构建

### 1.3.1 MobileNet-v2 整体网络结构

MobileNet-v2网络由标准卷积、全局平均池化以及大量的瓶颈结构构成<sup>[20]</sup>(表1),其中 $q$ 代表分类数,在本试验中分类数为6。

### 1.3.2 倒残差结构(瓶颈结构)

图4为MobileNet-v1、MobileNet-v2以及ResNet三种网络的核心模块思想。由图4可知, $C_{in}$ 为输入通道数, $C_{out}$ 为输出通道数,DW为深度卷积,PW为逐点卷积,Standard Conv为标准卷积。

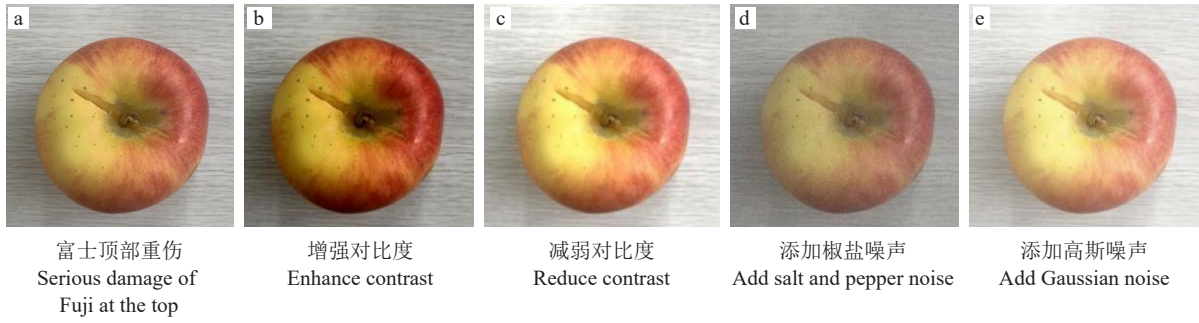


图 3 对图 1-d 增强对比度、减弱对比度、添加椒盐噪声、添加高斯噪声

Fig. 3 Enhance the contrast, reduce the contrast, add salt and pepper noise, and add Gaussian noise for Fig. 1-d

表 1 MobileNet-v2 整体网络结构

Table 1 MobileNet-v2 overall network structure

操作 Operator	输入 Input	扩张倍数 Expansion multiple	输出通道数 Number of output channels	迭代次数 Number of iterations	步长 Stride
Conv 1	224×224×3	–	32	1	2
Bottleneck	112×112×32	1	16	1	1
Bottleneck	112×112×16	6	24	2	2
Bottleneck	56×56×24	6	32	3	2
Bottleneck	28×28×32	6	64	4	2
Bottleneck	14×14×64	6	96	3	1
Bottleneck	14×14×96	6	160	3	2
Bottleneck	7×7×160	6	320	1	1
Conv 2	7×7×320	–	1280	1	1
Avg pool	7×7×1280	–	–	1	–
Conv 3	1×1×1280	–	$q$	–	1

注：“–”表示该操作无此项数据。下同。

Note: “–” represents the operation has no such data. The same below.

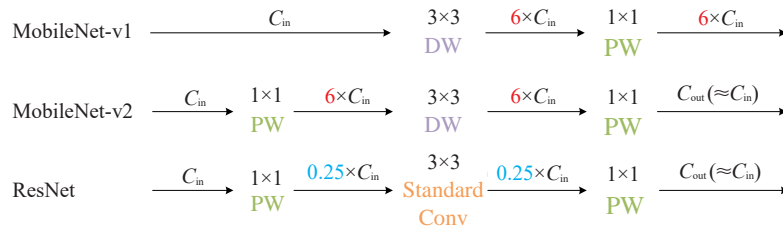


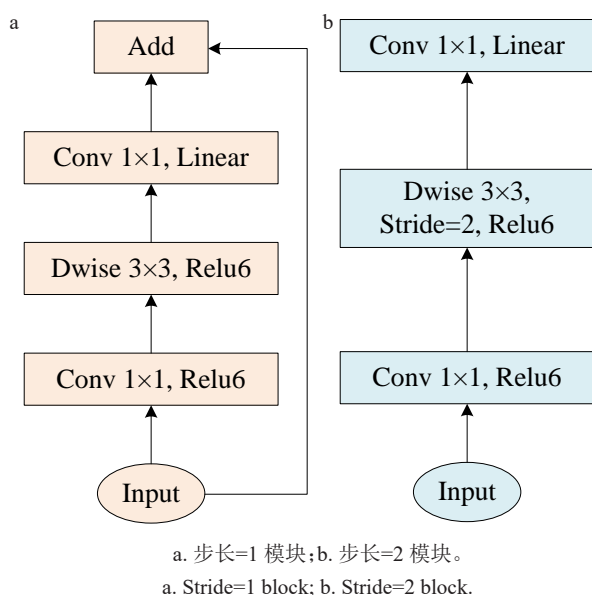
图 4 MobileNet-v1、MobileNet-v2、ResNet 三种网络的核心模块思想

Fig. 4 The core module ideas of MobileNet-v1, MobileNet-v2 and ResNet

MobileNet-v1 网络于 2017 年问世,其主要思想是采用深度可分离卷积(depthwise separable convolution)来缩减参数量和运算量,而在结构上没有采用残差思想<sup>[21]</sup>。ResNet 模型的残差结构,运用 1×1 的卷积核对输入矩阵先降维再升维<sup>[22]</sup>,但在 MobileNet-v2 模型中,采用与 ResNet 模型相反的倒残差结构,先升维,提取特征后再降维。MobileNet-v2 模型维度的波动类似瓶颈,所以其倒残差结构也称为瓶颈结构<sup>[23-24]</sup>; MobileNet-v2 倒残差结构分为两种

形式(图 5),当输入与输出大小相同且步长为 1 时,将输入与输出通过一条残差边(shortcut)相连进行直接相加;当步长为 2 时不使用残差边连接。

1.3.3 迁移学习 通过迁移学习<sup>[25]</sup>,在 ImageNet 数据集上训练得到的预训练网络,确立目标任务为苹果表皮损伤分级,使源领域的知识通过苹果数据集和 MobileNet-v2 网络得到重用,完成对苹果表皮的损伤分级。在基于迁移学习的 MobileNet-v2 模型基础上引入目标检测任务中的冻结训练策略<sup>[26]</sup>,比较



a. 步长=1 模块; b. 步长=2 模块。  
a. Stride=1 block; b. Stride=2 block.

图5 MobileNet-v2 倒残差结构

Fig. 5 MobileNet-v2 inverted residual structure

冻结不同网络层参数在训练后的效果。

1.3.4 评价指标 选择识别准确率(accuracy)、精确率(precision)、召回率(recall)、损失值(loss)、训练时间(training time)、识别时间(recognition time)以及网络大小(network size)作为模型训练结果的评价指标。

准确率  $A$  是指识别图像结果中识别正确的数量占整体识别数量的比重,能够直接表达模型的训练效果<sup>[27]</sup>;精确率  $P$  代表结果中对正样本的预测准确程度;召回率  $R$  是以真实类为基础,正样本中被正确识别的比例<sup>[28]</sup>。其运算公式为:

$$A = \frac{T_p + T_n}{T_p + T_n + F_p + F_n}, \quad (1)$$

$$P = \frac{T_p}{T_p + F_p}, \quad (2)$$

$$R = \frac{T_p}{T_p + F_n}. \quad (3)$$

其中:  $T_p$  为正样本被模型识别正确的数量;  $T_n$  为负样本被模型识别正确的数量;  $F_p$  代表被模型预测为正样本实际为负样本的数量;  $F_n$  代表被模型预测为负样本实际为正样本的数量。

损失值  $L$  能够表现出模型预测值与真实值之间的偏差情况,损失值小,表明模型的预测值与真实值之间相近,反之则相远<sup>[29]</sup>,其运算公式为:

$$L = -\frac{1}{m} \sum_i \log \left( \frac{e^{f_i}}{\sum_j e^{f_j}} \right). \quad (4)$$

其中:  $i$  表示对应样本;  $y_i$  表示第  $i$  个样本对应的标签;  $j$  是求和变量;  $m$  是样本总量;  $f$  为模型输出函数。

训练时间指模型训练完成所需时间;识别时间具体表示检测单张图片平均所用时间;网络大小为训练完成后所得网络占用的空间。

1.3.5 试验环境 试验所用计算机版本为 Windows 10 专业版;CPU 为 Intel 酷睿 i7-7700HQ,频率为 2.80 GHz;16 GB 运行内存;系统类型为 64 位操作系统,基于 x64 的处理器;NVIDIA GeForce GTX 1050 Ti 显卡,4 GB 显存。模型训练环境为 MATLAB R2021b 版本。

1.3.6 超参数设置 试验统一超参数设置,网络训练采用带动量的随机梯度下降法(stochastic gradient descent with momentum, SGDM),动量因子(momentum)设为 0.90;尺寸大小(batch size)设为 16,使用 L2 正则化,正则化参数  $\lambda=0.0005$ ;迁移学习需要保留预训练网络中迁移层的权重,降低学习速率,可将初始学习速率(initial learning rate)设为较小的值 0.0001;考虑数据集体量较小以及进行迁移学习时所需的训练轮数(epoch)相对较少,将训练轮数设为 20,同时采用每轮打乱一次的数据打乱策略;因训练周期中每轮的迭代次数(iteration)为 314,将验证频率(validation frequency)也设置为同样大小,即每 314 次迭代验证 1 次。

## 2 结果与分析

### 2.1 5 种轻量级模型对比

为了使试验结论更具有说服力,选择具有代表性的适合于移动终端和嵌入式设备的 MobileNet-v2、SqueezeNet、ShuffleNet、NASNet-Mobile 及 EfficientNet-b0 共 5 种 20 MB 以下的轻量级模型,首先在相同超参数设置的条件下进行初始化训练得出 5 种模型的测试精度在 56.32%~71.98%之间,难以满足分级需求,所以继续利用迁移学习技术进行 5 种轻量级模型的训练与验证。

训练精度曲线和训练损失曲线分别如图 6、图 7 所示。模型训练精度曲线表达的是迭代次数和准确率之间的关系。由图 6 可知 MobileNet-v2、SqueezeNet、ShuffleNet 及 NASNet-Mobile 模型的最终训练精度均在 90%以上, EfficientNet-b0 模型训练精度最低(80.25%);5 种模型中,收敛速度最快、

效果最好的是 MobileNet-v2 模型,在迭代第 1837 次时已经达到了 90.01%的精度,最终训练精度达到了 99.04%。

损失曲线表达的是迭代次数和损失值之间的关系。由图 7 可知,除 EfficientNet-b0 模型外,其余 4 种模型的最最终训练损失值均在 0.50 以下;其中收敛速度最快、效果最好的是 MobileNet-v2 模型,损失值

在迭代第 988 次时已经降到了 0.50 以下,最终训练损失值降到了 0.04。

如图 8 所示,基于迁移学习 MobileNet-v2 模型利用测试集得出的混淆矩阵,可以直观地看出每一分类的识别情况以及各分级的召回率。表 2、表 3 分别为基于迁移学习的 5 种网络模型在各分类上的召回率、精确率的对比。根据上述试验结论,可选择精

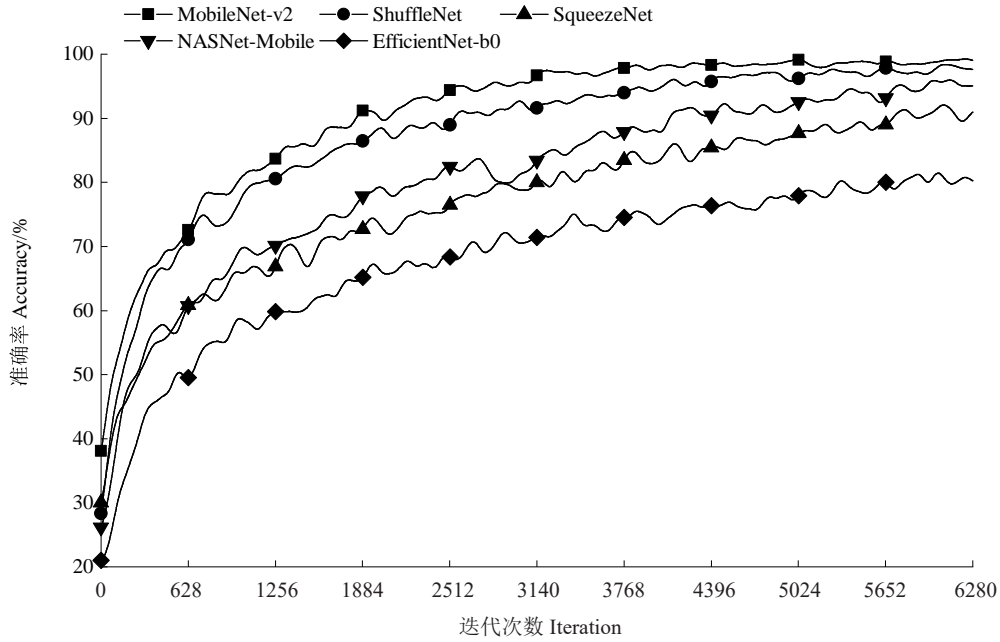


图 6 基于迁移学习的 5 种模型训练精度曲线

Fig. 6 Training accuracy curves of five models based on transfer learning

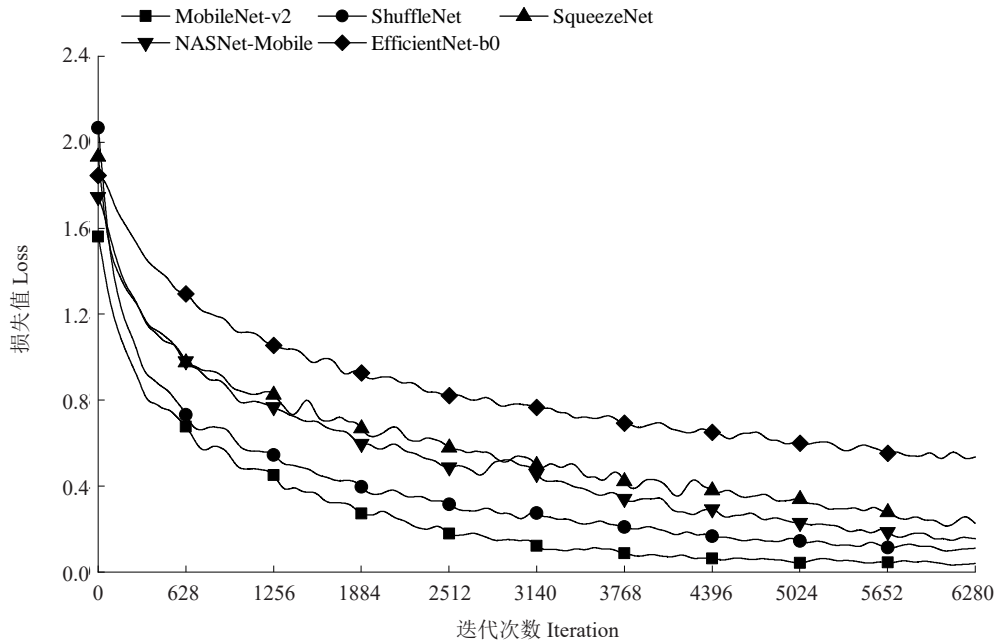
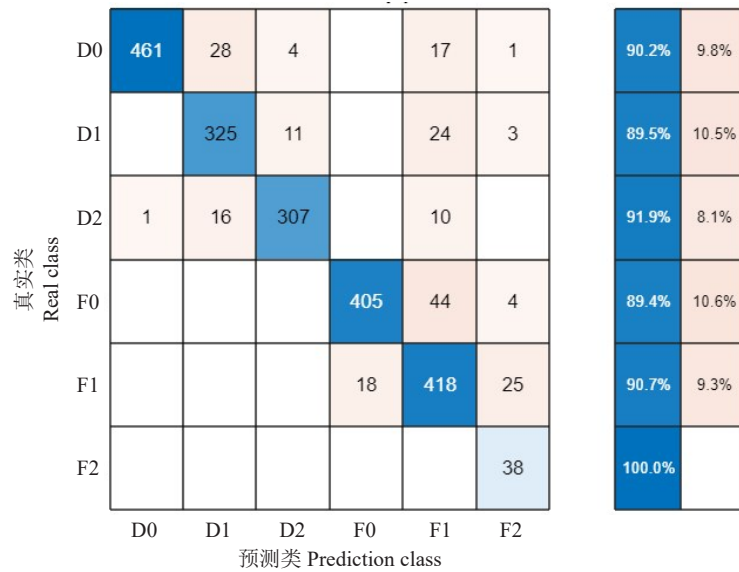


图 7 基于迁移学习的 5 种模型训练损失曲线

Fig. 7 Training loss curves of five models based on transfer learning



D0~D2 表示丹霞无伤、丹霞轻伤和丹霞重伤;F0~F2 表示富士无伤、富士轻伤和富士重伤。

D0-D2 represent no damage of Danxia, slight damage of Danxia and serious damage of Danxia, respectively; F0-F2 represent no damage of Fuji, slight damage of Fuji and serious damage of Fuji, respectively.

图 8 基于迁移学习的 MobileNet-v2 模型利用测试集得出的混淆矩阵

Fig. 8 Confusion matrix based on transfer learning MobileNet-v2 model using test set

表 2 基于迁移学习的模型各分类召回率对比

Table 2 Comparison of recall rates of different types of models based on transfer learning

%

基于迁移学习训练的模型 Transfer learning training model	丹霞无伤 No damage of Danxia	丹霞轻伤 Slight damage of Danxia	丹霞重伤 Serious damage of Danxia	富士无伤 No damage of Fuji	富士轻伤 Slight damage of Fuji	富士重伤 Serious damage of Fuji
MobileNet-v2	90.22	89.53	91.92	89.40	90.67	100.00
NASNet-Mobile	85.33	83.68	87.12	83.62	85.41	96.33
ShuffleNet	84.18	84.26	84.93	83.90	84.27	92.00
SqueezeNet	76.25	76.11	77.43	76.07	78.20	87.25
EfficientNet-b0	75.60	74.59	76.45	75.12	78.19	86.73

表 3 基于迁移学习的模型各分类精确率对比

Table 3 Comparison of precision rates of different types of models based on transfer learning

%

基于迁移学习训练的模型 Transfer learning training model	丹霞无伤 No damage of Danxia	丹霞轻伤 Slight damage of Danxia	丹霞重伤 Serious damage of Danxia	富士无伤 No damage of Fuji	富士轻伤 Slight damage of Fuji	富士重伤 Serious damage of Fuji
MobileNet-v2	99.78	88.08	95.34	95.74	80.69	53.52
NASNet-Mobile	94.38	82.32	90.37	89.55	76.01	51.56
ShuffleNet	93.11	82.89	88.10	89.85	75.00	49.24
SqueezeNet	84.34	74.87	80.32	81.47	69.59	46.70
EfficientNet-b0	83.62	73.38	79.30	80.45	69.60	46.42

度和各分类召回率最高的基于迁移学习的 MobileNet-v2 模型继续进行冻结训练研究,进一步探究是否可以加快网络训练时间以及提升精度。

2.2 基于迁移学习的 MobileNet-v2 模型冻结训练

表 4 为基于迁移学习的 MobileNet-v2 模型冻结训练。具体展开 MobileNet-v2 网络结构中的 Bottle-

neck 模块,在基于迁移学习的 MobileNet-v2 模型基础上冻结不同模块参数。

不冻结任何参数权重的验证精度为 91.30%,训练时间为 7816 s。从表 4 中可以看出模型从浅层至深层冻结时,训练时间和模型的验证精度总体呈下降趋势,冻结至 Bottleneck 7 时模型验证精度已经低

表 4 基于迁移学习的 MobileNet-v2 模型冻结训练  
Table 4 MobileNet-v2 model freeze training based on transfer learning

操作 Operator	输入 Input	输出通道数 Number of output channels	扩张倍数 Expansion multiple	步长 Stride	冻结至该操作后的训练时间 Freeze to the training time after the operation/s	冻结至该操作后的验证精度 Freezing the verification accuracy after the operation/%
Conv 1	224×224×3	32	-	2	7627	92.78
Bottleneck 1	112×112×32	16	1	1	7799	91.81
Bottleneck 2-1	112×112×16	16	6	2	7685	91.53
Bottleneck 2-2	56×56×16	24	6	1	6695	91.44
Bottleneck 3-1	56×56×24	24	6	2	5524	92.23
Bottleneck 3-2	28×28×24	24	6	1	4951	89.63
Bottleneck 3-3	28×28×24	32	6	1	4988	89.50
Bottleneck 4-1	28×28×32	32	6	2	5203	88.71
Bottleneck 4-2	14×14×32	32	6	1	4481	89.59
Bottleneck 4-3	14×14×32	32	6	1	5292	87.51
Bottleneck 4-4	14×14×32	64	6	1	3952	89.08
Bottleneck 5-1	14×14×64	64	6	1	4079	87.23
Bottleneck 5-2	14×14×64	64	6	1	3702	87.51
Bottleneck 5-3	14×14×64	96	6	1	4005	87.27
Bottleneck 6-1	14×14×96	96	6	2	3712	83.06
Bottleneck 6-2	7×7×96	96	6	1	3715	80.93
Bottleneck 6-3	7×7×96	160	6	1	3173	73.48
Bottleneck 7	7×7×160	320	6	1	3056	69.18
Conv 2	7×7×320	1280	-	1	-	-
Avg pool	7×7×1280	-	-	-	-	-
Conv 3	1×1×1280	1000	-	1	-	-

于 70%，所以不再继续冻结网络的最后 3 层。同时从表 4 还可知，选择冻结至第 1 个卷积模块到 Bottleneck 3-1 模块时均可在缩短模型训练时间的基础上提高模型验证精度。在冻结第 1 个卷积层参数时比不冻结任何参数权重模型验证精度提高了 1.48%，训练时间缩短了 189 s。在冻结到 Bottleneck 3-1 模块时比不冻结任何参数权重模型训练时间缩短了 29.32%，验证准确度提高了 0.93%。

如表 5 所示，分别是基于迁移学习和冻结至 Bottleneck 3-1 模块后训练的 MobileNet-v2 模型 (MobileNet-v2\*)、基于迁移学习的 MobileNet-v2 模型、基于迁移学习的 NASNet-Mobile 模型、基于迁移学习的 ShuffleNet 模型、基于迁移学习的 SqueezeNet 模型和基于迁移学习的 EfficientNet-b0 模型基于测试集在测试精度、识别时间、网络大小三方面的对比，基于迁移学习的 MobileNet-v2 模型在冻结到 Bottleneck 3-1 模块后训练所得网络测试精度最高，达 91.58%，识别时间 0.14 s，网络大小 8.15 MB，能够在保证最高识别精度的前提下较好地适配于移动终端和嵌入式设备。

表 5 不同模型评价指标对比  
Table 5 Comparison of evaluation indexes of different models

基于迁移学习的模型 Transfer learning based model	测试精度 Testing accuracy/%	识别时间 Recognition time/s	网络大小 Network size/MB
MobileNet-v2*	91.58	0.14	8.15
MobileNet-v2	90.46	0.15	8.17
NASNet-Mobile	85.13	0.39	15.50
ShuffleNet	83.62	0.12	3.27
SqueezeNet	75.96	0.04	2.63
EfficientNet-b0	73.97	0.16	14.60

### 2.3 Grad-CAM 可视化

Grad-CAM 翻译为梯度加权类激活映射 (gradient-weighted class activation map)，用于对神经网络的输出进行可视化，可直观地展示出卷积神经网络做出分类决策的重要部位<sup>[30-31]</sup>。选择富士重伤侧部原图、添加高斯噪声的富士轻伤侧部图、增强对比度的丹霞重伤顶部图、添加椒盐噪声的丹霞无伤底部图以及顺时针旋转 30° 的丹霞轻伤侧部图，使用 Grad-CAM 技术得出热力图的可视化展示，如图 9 所示，从

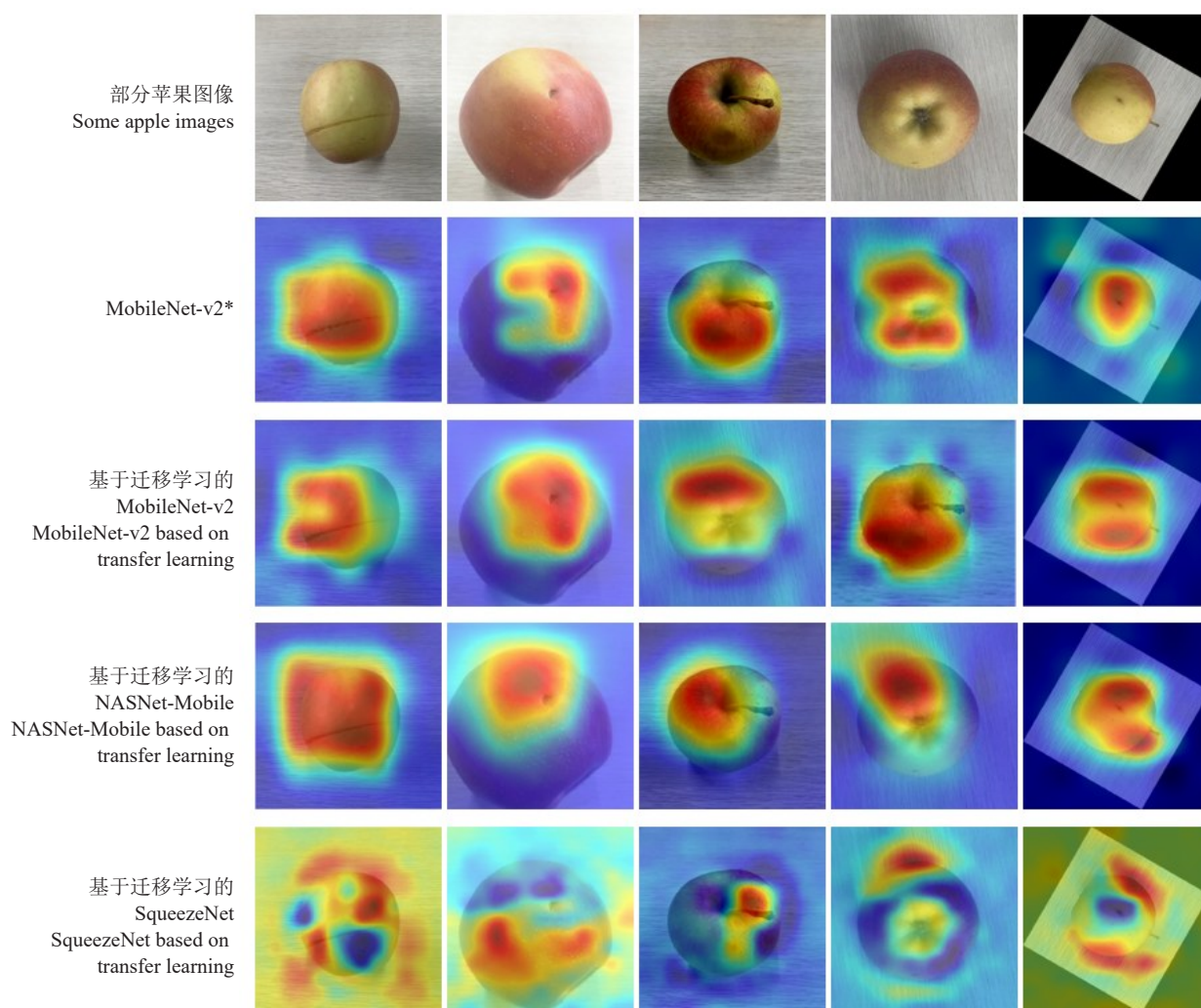


图9 Grad-CAM 可视化  
Fig. 9 Grad-CAM visualization

蓝色到红色颜色的加深代表模型对图片区域关注度的增加。从图9可以看出,基于迁移学习的SqueezeNet模型可视化效果最差,同时也验证了该模型的识别精度最低;基于迁移学习的NASNet-Mobile模型可视化效果较差,只能显示大范围的关注区域,部分图片识别度不高;基于迁移学习的MobileNet-v2模型可视化效果明显优于前两种模型,但模型识别的重点区域与实际有所偏差;基于迁移学习的MobileNet-v2模型在冻结到Bottleneck 3-1模块后训练所得网络的可视化效果最优,模型识别图片的重点区域与实际吻合度最高,可以直观地显示出该模型在识别苹果表皮图像时的稳定性与精确度。

### 3 讨 论

众所周知,苹果的种质资源丰富,其果实的大小、硬度、表皮颜色等特征不尽相同,这将直接影响

着苹果表皮损伤识别的准确度;同时采集损伤图片比较困难、识别难度较大是对苹果表皮损伤研究较少的主要原因之一。目前,普遍使用光谱仪器、红外热成像仪等高端设备对苹果损伤进行检测和分级,成本相对较高且不适合大众消费者使用;而使用普通相机采集苹果表皮的损伤图像进行分级的研究较少,普通相机采集图像有成本低、硬件兼容性强、获取图片来源广等优点,Ismail等<sup>[18]</sup>选取了5个品种共550张苹果图像进行试验分析,虽然品种多样,但数据量较少,容易发生训练过拟合,且得出的模型泛化能力较弱。笔者在本研究中选取山西地区具有代表性的富士和丹霞2个苹果品种,共采集780张图像,并将初始数据集扩充到9360张;专注于卷积神经网络的研究,使用较少品种、大量数据的思想防止模型训练过拟合,同时提升了测试集和验证集的比例,让结果更贴近实际;在选取网络上,挑选5个20 MB以

下的轻量级卷积神经网络进行试验对比,目的是为针对移动端和嵌入式设备的研究者提供技术参考,从而推进苹果商家的分类售卖;在模型的训练上,使用迁移学习技术的同时增加冻结训练策略,在缩短模型训练时间的基础上将验证精度提升到了92.78%。其中,在冻结到Bottleneck 3-1模块时比基于迁移学习的MobileNet-v2模型训练时间缩短了29.32%,同时验证精度提高了0.93%,测试精度为91.58%,能够满足基本的分级需求,具有较强的泛化能力。

在进一步的研究中,可增加苹果的种类,扩大损伤部位采集的数据量,探寻轻量级卷积神经网络的优化或对深度卷积神经网络进行压缩,提高识别的速度和压缩网络大小;探索目标检测领域,对比优化目前流行的Yolo算法和Faster R-CNN算法,提高定位和分级效果。

## 4 结 论

笔者在本研究中选用5种20 MB以下的轻量化模型进行初始化训练与迁移学习训练对比,得出经过迁移学习的MobileNet-v2模型效果最优,随后使用冻结策略进行分层训练,其中在冻结到Bottleneck 3-1模块时验证精度达92.23%,测试精度为91.58%,平均识别时间为0.14 s,网络大小为8.15 MB,为移动终端和嵌入式设备在苹果损伤直接分级方面提供技术参考。

### 参考文献 References:

- [1] 王皎,李赫宇,刘岱琳,宋新波,於洪建. 苹果的营养成分及保健功效研究进展[J]. 食品研究与开发,2011,32(1):164-168.  
WANG Jiao, LI Heyu, LIU Dailin, SONG Xinbo, YU Hongjian. Research progress of apple nutrition components and health function[J]. Food Research and Development, 2011, 32(1): 164-168.
- [2] 屠鹏,边红霞,石萍,黎虹,杨晰. 压力损伤对苹果贮藏期品质的影响[J]. 食品工业科技,2018,39(14):239-243.  
TU Peng, BIAN Hongxia, SHI Ping, LI Hong, YANG Xi. Effect of pressure injury on the quality during storage period in apple[J]. Science and Technology of Food Industry, 2018, 39(14):239-243.
- [3] 王艳颖,胡文忠,庞坤,朱蓓薇,范圣第. 机械损伤对富士苹果生理生化变化的影响[J]. 食品与发酵工业,2007,33(7):58-62.  
WANG Yanying, HU Wenzhong, PANG Kun, ZHU Beiwei, FAN Shengdi. Effect of mechanical damage on the physiology and biochemistry in Fuji apple[J]. Food and Fermentation Industries, 2007, 33(7):58-62.
- [4] 赵康. 苹果表面损伤光学检测系统设计与实现[D]. 哈尔滨:黑龙江大学,2021.  
ZHAO Kang. Design and implementation of optical detection system for apple surface damage[D]. Harbin: Heilongjiang University, 2021.
- [5] 沈宇,房胜,郑纪业,王风云,张琛,李哲. 基于高光谱成像技术的富士苹果轻微机械损伤检测研究[J]. 山东农业科学,2020,52(2):144-150.  
SHEN Yu, FANG Sheng, ZHENG Jiye, WANG Fengyun, ZHANG Chen, LI Zhe. Detection of slight mechanical damage of fuji apple fruits based on hyperspectral imaging technology[J]. Shandong Agricultural Sciences, 2020, 52(2): 144-150.
- [6] 蒋金豹,尤笛,汪国平,张政,门泽成. 苹果轻微机械损伤高光谱图像无损检测[J]. 光谱学与光谱分析,2016,36(7):2224-2228.  
JIANG Jinbao, YOU Di, WANG Guoping, ZHANG Zheng, MEN Zecheng. Study on the detection of slight mechanical injuries on apples with hyperspectral imaging[J]. Spectroscopy and Spectral Analysis, 2016, 36(7):2224-2228.
- [7] LU Y Z, LU R F. Development of a multispectral structured illumination reflectance imaging (SIRI) system and its application to bruise detection of apples[J]. Transactions of the ASABE, 2017, 60(4): 1379-1389.
- [8] KERESZTES J C, GOODARZI M, SAEYS W. Real-time pixel based early apple bruise detection using short wave infrared hyperspectral imaging in combination with calibration and glare correction techniques[J]. Food Control, 2016, 66:215-226.
- [9] 邵志明,王怀彬,董志城,原育慧,李军会,赵龙莲. 基于近红外相机成像和阈值分割的苹果早期损伤检测[J]. 农业机械学报,2021,52(S1):134-139.  
SHAO Zhiming, WANG Huaibin, DONG Zhicheng, YUAN Yuhui, LI Junhui, ZHAO Longlian. Early bruises detection method of apple surface based on near infrared camera imaging technology and image threshold segmentation method[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(S1):134-139.
- [10] FAN S X, LIANG X T, HUANG W Q, ZHANG V J, PANG Q, HE X, LI L J, ZHANG C. Real-time defects detection for apple sorting using NIR cameras with pruning-based YOLOV4 network[J]. Computers and Electronics in Agriculture, 2022, 193: 106715.
- [11] 门洪,陈鹏,邹丽娜,朱斌. 基于主动热红外技术的苹果损伤检测[J]. 中国农机化学报,2013,34(6):220-224.  
MEN Hong, CHEN Peng, ZOU Lina, ZHU Bin. Apples' damage detection based on active thermal infrared technique[J]. Journal of Chinese Agricultural Mechanization, 2013, 34(6): 220-224.
- [12] 周其显. 苹果早期机械损伤的红外热成像检测研究[D]. 南昌:华东交通大学,2011.  
ZHOU Qixian. Research on infrared thermography technique for detecting early chanical damage in apples[D]. Nanchang: East

- China Jiaotong University, 2011.
- [13] 熊婷. 基于低场磁共振技术的果品无损检测[D]. 杭州: 中国计量学院, 2014.  
XIONG Ting. Nondestructive detection of fruit quality based on low-field magnetic resonance technology[D]. Hangzhou: China Jiliang University, 2014.
- [14] 尹义志, 王永刚, 张楠楠, 刘宇航. 基于小波神经网络的温室番茄产量预测[J]. 中国瓜菜, 2020, 33(8): 53-59.  
YIN Yizhi, WANG Yonggang, ZHANG Nannan, LIU Yuhang. Study on tomato yield prediction in greenhouse based on wavelet neural network[J]. China Cucurbits and Vegetables, 2020, 33(8): 53-59.
- [15] 陈新, 伍萍辉, 祖绍颖, 徐丹, 张云鹤, 董静. 基于改进 SSD 轻量化神经网络的番茄疏花疏果农事识别方法[J]. 中国瓜菜, 2021, 34(9): 38-44.  
CHEN Xin, WU Pinghui, ZU Shaoying, XU Dan, ZHANG Yunhe, DONG Jing. Study on identification method of thinning flower and fruit of tomato based on improved SSD lightweight neural network[J]. China Cucurbits and Vegetables, 2021, 34(9): 38-44.
- [16] 宁景苑, 叶海芬, 孙雨妃, 熊思怡, 梅正昊, 蒋晨豪, 黄科涛, 张苏婕, 朱哲琛, 李昱权, 惠国华, 易晓梅, 郗园园, 吴鹏. 一种基于弛豫动态光谱技术的苹果损伤检测方法研究[J]. 传感技术学报, 2022, 35(8): 1150-1156.  
NING Jingyuan, YE Haifen, SUN Yuqi, XIONG Siyi, MEI Zhenghao, JIANG Chenhao, HUANG Ketao, ZHANG Sujie, ZHU Zhechen, LI Yuquan, HUI Guohua, YI Xiaomei, GAO Yuanyuan, WU Peng. Study of apple slight damage detecting method based on dynamic relaxation spectrum[J]. Chinese Journal of Sensors and Actuators, 2022, 35(8): 1150-1156.
- [17] NING J Y, YE H F, SUN Y Q, ZHANG J Y, MEI Z H, XIONG S Y, ZHANG S J, LI Y Q, HUI G H, YI X M, GAO Y Y, WU P. Study on apple damage detecting method based on relaxation single-wavelength laser and convolutional neural network[J]. Journal of Food Measurement and Characterization, 2022, 16(5): 3321-3330.
- [18] ISMAIL A, IDRIS M Y I, AYUB M N, POR L Y. Vision-based apple classification for smart manufacturing[J]. Sensors, 2018, 18(12): 4353.
- [19] 李泰, 卢士军, 黄家章, 陈磊, 范协裕. 苹果品质评价标准研究进展[J]. 中国农业科技导报, 2021, 23(11): 121-130.  
LI Tai, LU Shijun, HUANG Jiazhang, CHEN Lei, FAN Xieyu. Research progress of apple quality evaluation standards[J]. Journal of Agricultural Science and Technology, 2021, 23(11): 121-130.
- [20] 轩勃娜, 李进, 宋亚飞, 马泽煊. 基于改进 MobileNetV2 的恶意代码分类方法[J]. 计算机应用, 2023, 43(7): 2217-2225.  
XUAN Bona, LI Jin, SONG Yafei, MA Zexuan. Malicious code classification method based on improved MobileNetV2[J]. Journal of Computer Applications, 2023, 43(7): 2217-2225.
- [21] HOWARD A G, ZHU M L, CHEN B, KALENICHENKO D, WANG W J, WEYAND T, ANDREETTO M, ADAM H. MobileNets: efficient convolutional neural networks for mobile vision applications[Z]. ArXiv, 2017. DOI: 10.48550/arXiv.1704.04861.
- [22] HE K M, ZHANG X Y, REN S Q, SUN J. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 770-778.
- [23] SANDLER M, HOWARD A, ZHU M L, ZHMOGINOV A, CHEN L C. MobileNetV2: inverted residuals and linear bottlenecks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE, 2018: 4510-4520.
- [24] 葛淑伟, 张永茜, 秦嘉欣, 李雪, 王晓. 基于优化 SSD-MobileNetV2 的煤矿井下锚孔检测方法[J]. 采矿与岩层控制工程学报, 2023, 5(2): 66-74.  
GE Shuwei, ZHANG Yongqian, QIN Jiabin, LI Xue, WANG Xiao. Rock bolt borehole detection method for underground coal mines based on optimized SSD-MobileNetV2[J]. Journal of Mining and Strata Control Engineering, 2023, 5(2): 66-74.
- [25] PAN S J, YANG Q. A survey on transfer learning[J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10): 1345-1359.
- [26] BROCK A, LIM T, RITCHIE J M, WESTON N. FreezeOut: accelerate training by progressively freezing layers[Z]. ArXiv, 2017. DOI: 10.48550/arXiv.1706.04983.
- [27] 孔德宇. 基于 Kinect 骨骼时空特征的步态人物识别研究[D]. 南宁: 广西大学, 2022.  
KONG Deyu. Research on gait recognition based on kinect skeletal spatiotemporal features[D]. Nanning: Guangxi University, 2022.
- [28] 谢圣桥, 宋健, 汤修映, 白阳. 基于迁移学习和残差网络的葡萄叶部病害识别[J]. 农机化研究, 2023, 45(8): 18-23.  
XIE Shengqiao, SONG Jian, TANG Xiuying, BAI Yang. Identification of grape leaf diseases based on transfer learning and residual networks[J]. Journal of Agricultural Mechanization Research, 2023, 45(8): 18-23.
- [29] 高博, 陈琳, 严迎建. 基于 CNN-MGU 的侧信道攻击研究[J]. 信息安全, 2022, 22(8): 55-63.  
GAO Bo, CHEN Lin, YAN Yingjian. Research on side channel attack based on CNN-MGU[J]. Netinfo Security, 2022, 22(8): 55-63.
- [30] SELVARAJU R R, COGSWELL M, DAS A, VEDANTAM R, PARIKH D, BATRA D. Grad-CAM: Visual explanations from deep networks via gradient-based localization[C]. Italy: 2017 IEEE International Conference on Computer Vision (ICCV), 2017: 618-626.
- [31] PAPANDRIANOS N I, FELEKI A, MOUSTAKIDIS S, PAPA-GEORGIOU E I, APOSTOLOPOULOS I D, APOSTOLOPOULOS D J. An explainable classification method of SPECT myocardial perfusion images in nuclear cardiology using deep learning and grad-CAM[J]. Applied Sciences, 2022, 12(15): 7592.