

文章编号: 1005-5630(2023)02-0026-10

DOI: 10.3969/j.issn.1005-5630.2023.002.004

基于注意力机制的水果损伤检测及分类

张杰, 夏春蕾, 张荣福, 哈利扎提·居来提, 刘怡

(上海理工大学 光电信息与计算机工程学院, 上海 200093)

摘要: 水果作为人们日常必备的食物, 其越来越高的消费需求使得行业对自动损伤检测和自动分类的要求日益提高。针对这一需求, 近年来水果损伤自动检测成为研究的热门话题。针对现有的深度学习技术, 即卷积神经网络在水果的特征提取和分类方面的应用进行了探讨, 提出了一种以 ResNet34 作为主干网络, 并在此基础上引入注意力机制 SE 和 CBAM 模块的方法来实现水果损伤的检测和基本分类。在 fruit fresh and rotten for classification 数据集上完成了该方法的验证。经过与 VGG16, GoogLeNet, MobileNetV2 等常见网络的比较分析, 结果显示改进后的模型分类准确度达到 98.8%。通过加入新的苹果数据集, 该模型相比原网络 ResNet34, 在性能方面进一步提升, 有效提高了模型的泛化性。该模型提升了水果损伤检测和分类处理的精确性, 在实际中, 可为复杂的水果图片的多特征分类处理提供借鉴。

关键词: 深度学习; 水果损伤检测; ResNet; 注意力机制

中图分类号: TP 183 **文献标志码:** A

Fruit damage detection and classification based on attention mechanism

ZHANG Jie, XIA Chunlei, ZHANG Rongfu, HALIZHATI Julaiti, LIU Yi

(School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: For the daily essential food of people, automatic damage detection and automatic classification are essential for the increasing consumption of fruit. In view of this demand, automatic detection of fruit damage has become a hot topic in recent years. In this paper, the application of convolutional neural network, an existing deep learning technology, in fruit feature extraction and classification was discussed. A method based on ResNet34 as the backbone network and the introduction of attention mechanism SE and CBAM module was proposed to realize the detection and basic classification of fruit damage. The method was verified on fruit fresh and rotten for classification data set, and compared with VGG16, GoogLeNet, MobileNetV2 and other common networks. The accuracy of fruit damage detection and classification is improved. The classification accuracy reaches 98.8%. By adding the new apple data set, the performance of the

收稿日期: 2022-09-17

第一作者: 张杰 (1997—), 男, 硕士研究生, 研究方向为图像处理。E-mail: 18234509891@163.com

通信作者: 夏春蕾 (1974—), 女, 讲师, 研究方向为图像处理和信号编码。E-mail: xiachunlei@usst.edu.cn

model is further improved, compared with the original network ResNet34, and the generalization of the model is effectively improved, which provides a reference for the complex multi-feature classification of actual fruit images.

Keywords: deep learning; fruit damage detection; ResNet; attention mechanism

引 言

在众多的食物中, 水果是人类日常健康饮食的重要组成部分, 因为水果富含的维生素等营养元素能帮助预防疾病。然而, 很多水果的保质期较短。由于储存不当或者运输过程中的碰撞等原因造成的损伤, 使得水果品质下降甚至变得腐烂而不宜食用。因此, 食品行业在出售水果之前需要对其进行严格的检测。传统的人工检测方案不仅成本高、效率低, 还存在一致性和准确性较差等问题。随着生活水平的提高, 水果的消费量逐年增长, 而不同市场消费者对水果品质、种类等方面的需求呈现出多样化特点, 水果的自动化品质检测和分类也就成为必须要解决的问题。近年来, 随着深度学习和计算机视觉技术的不断发展, 机器学习及卷积神经网络在分类和检测领域的应用研究也越来越多。探索基于深度学习技术的水果损伤检测和分类方法成为了研究热点^[1]。

近几年, 研究人员在腐坏水果的检测方面进行了很多研究。Karakaya 等^[2]研究了多种特征提取技术在水果新鲜度分类上的性能表现。他们在对水果进行特征提取后, 采用支持向量机 (support vector machine, SVM) 分类技术进行实验。在对 1 个共有 1200 张图片, 包含橘子、香蕉和苹果等 3 类水果在内的数据集进行测试后, 证明卷积神经网络搭配 SVM 分类器呈现出最好的性能, 分类准确率达到 97.61%。Wajid 等^[3]提出了 1 种快速判别柑橘状态的方法。他们通过对比分析包括贝叶斯、人工神经网络和决策树的适用性和性能, 发现决策树分类技术对橙色条件的分类效率高于其他方法。该方法的准确度、精确度和灵敏度分别为 93.13%, 93.45% 和 93.24%。Singh 等^[4]运用小波变换、定向梯度直方图、灰度共生矩阵等方法提取苹果的纹理特征, 然后用 SVM、k-NN、逻辑回归、线性判别等多种分

类器对新鲜苹果和腐烂苹果进行对比实验, 结果发现, SVM 分类器的性能为 98.8%, 优于其他分类器。Chakraborty 等^[5]使用卷积神经网络提取水果图像的特征, 并使用 Max pooling、Average pooling 和 MobileNetV2 架构对图像进行分类。在 Kaggle 数据集上的性能测试结果显示: MobileNetV2 在训练集和验证集中的准确率分别达到了 99.46% 和 99.61%; Max pooling 的训练集准确率达到 94.49%, 验证集准确率达到 94.97%; Average pooling 的训练集准确率为 93.06%, 验证集准确率为 93.72%。结果表明, 该研究所提出的卷积神经网络模型能够区分新鲜水果和腐烂水果。

上述方法主要是一些主流的分类网络在水果数据集上的直接应用, 研究者对网络本身并无太多的改进^[6]。结合以上论文的研究成果, 通过与经典的分类网络对比, 本文选择 ResNet 模型作为分类模型, 并在此基础上尝试加入两个不同的注意力机制进行改进, 从而将水果分类所依据的特征信息予以进一步优化, 获取有利于水果分类的重要特征; 通过可视化算法 Grad-CAM^[7]将模型提取的效果进行展示, 最终获得更高的分类准确率; 利用自制的新数据集达到图像增强效果, 提升了网络的泛化性能, 对复杂的水果图像也有不错的识别效果。

1 模型构架

1.1 ResNet 网络模型

ResNet^[8] 残差网络在 2015 年同时斩获 ImageNet 竞赛中分类任务第一名和目标检测第一名^[9]。这个网络最初由微软工作室提出, 其亮点是: 解决了网络层数不断加深而导致的梯度爆炸或者梯度消失问题; 拥有超深的网络结构, 不影响其性能; 提出了 Residual 模块, 解决了网络

退化问题；使用了 Batch Normalization 加速训练等。该网络被广泛使用，在图像分割和目标检测

方向也取得了很好的效果。ResNet34 的结构如图 1 所示。

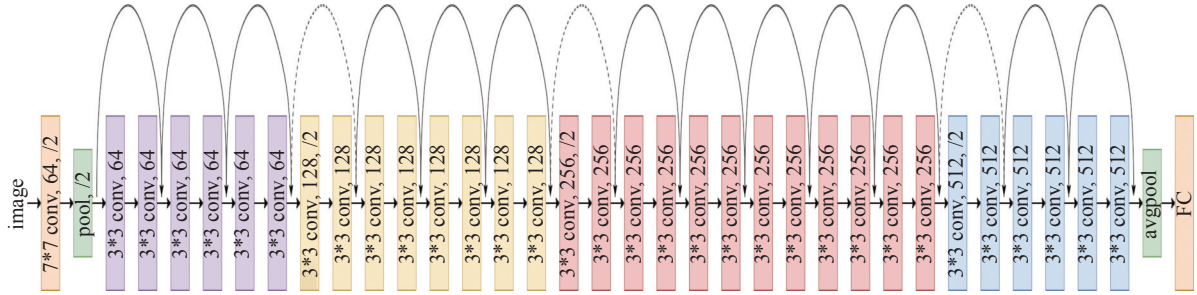


图 1 ResNet34 结构图

Fig. 1 ResNet34 network structure

为了克服神经网络层次越深越难训练而导致的网络退化问题，ResNet 模型引入了残差模块，很好地解决了这一问题^[10]。残差模块如图 2 所示，对于 1 个堆叠层结构，当输入为 x 时，其学习到的特征记为 $H(x)$ ，残差 $F(x) = H(x) - x$ 。残差学习相比原始特征直接学习更加容易。当残差为 0 时，堆叠层仅仅做了恒等映射，网络

性能不会下降。实际上残差不会为 0，这也使得堆叠层在输入特征基础上学习到新的特征，从而拥有更好的性能。

1.2 SE 模块

SENet^[11] 是 Squeeze-and-Excitation Networks 的简称，该网络获得了 ILSVRC2017 分类比赛的冠军。SE 模块由 Squeeze 和 Excitation 两部分组成，如图 3 所示。Squeeze 表示顺着空间维度进行特征压缩，将每个二维的特征通道变成 1 个实数。这个实数某种程度上具有全局的感受野。输出的维度和输入的特征通道数相匹配^[12]。

Squeeze 通过全局池化将每个通道的二维特征 ($H \times W$) 压缩为一个实数，其公式为

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (1)$$

给定 1 个输入, $u_c \in \mathbf{R}^{H \times W \times C}$, H, W, C 分别为提取特征后的高、宽和通道。通过全局池化，将每个通道对应的空间信息 ($H \times W$) 压缩到对应通道中变为 1 个数值。此时 1 个像素表示 1 个通

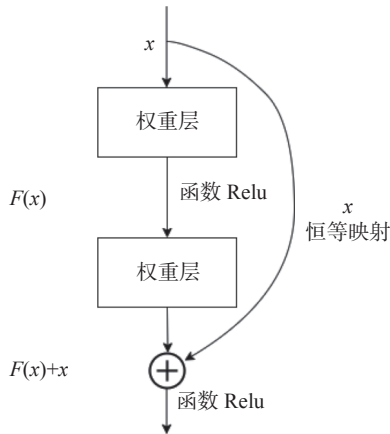


图 2 残差模块

Fig. 2 Residual module

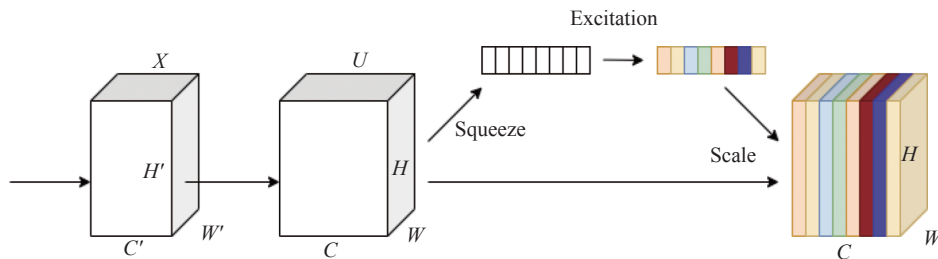


图 3 SE 模块结构图

Fig. 3 Structure of SE module

道, 最终维度变为 $1 \times 1 \times C$, 所以输出 $z_c \in \mathbf{R}^{1 \times 1 \times C}$ 最终成为 1 个向量。

Excitation 表示基于特征通道间的相关性, 每个特征通道生成 1 个权重, 用来代表特征通道的重要程度。为了实现这一目标, 选择了门控机制和 Sigmoid 激活函数, 其公式为

$$s_c = \sigma(W_2 \delta(W_1 z)) \quad (2)$$

式中: δ 为激活函数 Relu; σ 为激活函数 Sigmoid; W_1 和 W_2 为线性层; 最终得到的 s_c 为权重值。

最后, 利用 Scale 将上一步得到的权重值 s_c 与原始特征图 u_c 进行操作, 将前面得到的归一化权重加权到每个通道的特征上, 逐个通道乘加

权系数, 这样就在通道维度中引入了注意力机制, 表达式为

$$\tilde{x}_c = F_{\text{Scale}}(u_c, s_c) = s_c \times u_c \quad (3)$$

式中: u_c 为 u 的 1 个通道; s_c 为通道的权重。

1.3 CBAM 模块

卷积注意力模块^[13](convolutional block attention module, CBAM)是 1 种轻量通用注意力模块, 如图 4 所示, 能同时在空间和通道上进行特征的注意力机制, 所以该模块由两部分组成, 通道注意力模块和空间注意力模块^[14]。

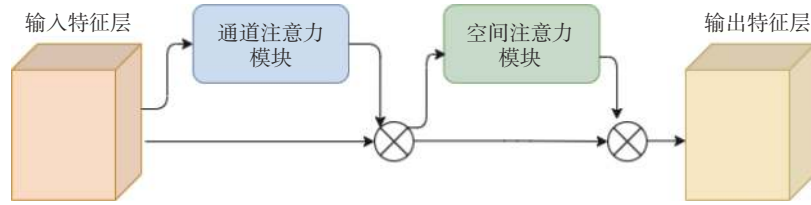


图 4 CBAM 模块结构图

Fig. 4 Structure of CBAM module

通道注意力模块首先将特征图分别采用平均池化和最大池化操作来聚合特征映射的空间信息, 然后将其放入到多层感知机, 最后将两者合并, 通过激活函数得到结果 M_c , 其公式为

$$M_c(F) = \sigma(\text{MLP}(\text{Avgpool}(F)) + \text{MLP}(\text{Maxpool}(F))) \quad (4)$$

式中: σ 为 Sigmoid 函数; MLP 为多层感知机; M_c 为通道注意力模块。

空间注意力模块的步骤跟上面类似, 先两个池化操作聚合特征图, 然后通过卷积操作和激活函数得到结果 M_s , 其公式为

$$M_s(F) = \sigma(f^{7 \times 7}([\text{Avgpool}(F); \text{MaxPool}(F)])) \quad (5)$$

式中: σ 为 Sigmoid 函数; $f^{7 \times 7}$ 为 7×7 的卷积核操作。

最后将这两个模块按照先通道后空间的顺序排列, 就组成了完整的 CBAM 模块。

1.4 模型结构

为了进一步提升 ResNet34 模型的分类效果, 使网络更加注意到新鲜水果与缺陷水果的特征, 加强图像有效特征信息, 从而提高模型的鲁

棒性, 本文提出了将 SE 模块和 CBAM 模块嵌入到 ResNet 网络中。由于该实验的数据集样本偏少, 数据特征相对简单, 因此在模型训练过程中容易产生过拟合的现象, 从而导致模型泛化性能变差。所以将 SE 模块嵌入到 ResNet 网络的每一层 Residual Block 中, 网络的每一个 Residual Block 都将提升网络对图像重要区域的捕捉。一共引入了 16 层的 Residual Block, 如图 5 所示。将 CBAM 模块嵌入到第 1 层 Residual Block 之前与最后 1 层 Residual Block 之后, 相当于将该模块添加到整个模型的开头和结尾, 一共 2 次。其原因是 CBAM 更善于捕捉图像的位置信息。先确定好想要捕捉的方位, 这将更有针对性地对图像的具体细节进行特征提取, 结构如图 6 所示。两种不同注意力机制可以相互作用。其中, SE 模块可以在不引入新的空间维度情况下显示构建特征通道之间的依赖关系, 通过在图像的每个通道施加 1 个权重, 加大水果图像对通道信息的差异。而 CBAM 模块可使得模型对空间信息更加敏感, 抑制无效特征层的影响, 提升模型的准确率^[15]。

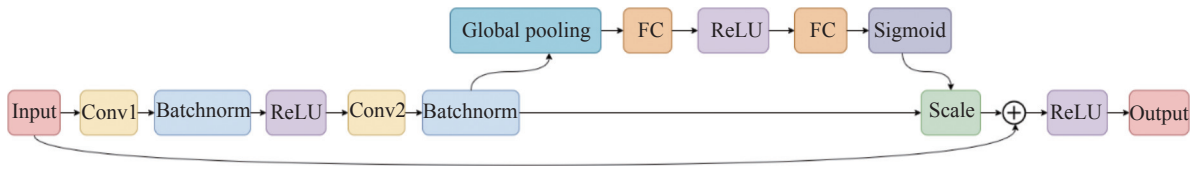


图 5 Residual Block-SE 结构图
Fig. 5 Structure of Residual Block-SE

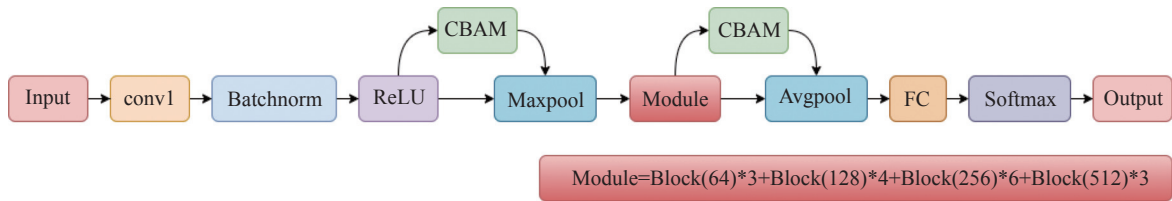


图 6 改进后的模型结构
Fig. 6 Improved model structure

2 实验及结果分析

2.1 实验平台

本实验在 Google Colabs 平台上进行，GPU 型号为 Tesla K80，显存为 11 441 MiB，Python 版本为 3.7.12，深度学习框架为 Pytorch，版本为 1.10.0。

2.2 数据集

本文使用的数据集是 Kaggle 网站上的 fruit fresh and rotten for classification 公开数据集。数据集大小为 1.95 GB，分为训练集和测试集，每个集里又分为 6 类，分别是新鲜苹果、新鲜香蕉、新鲜橙子、腐烂苹果、腐烂香蕉和腐烂橙子（如图 7 所示），图片数量一共有 13 600 多张。

2.3 模型训练与参数调整

在模型训练过程中，将图像缩放至 224×224×3 作为网络输入尺寸，由于数据集本身已经进行了加入椒盐噪声、旋转角度等数据增强处理，所以仅需要将数据集转化为张量并进行归一化。为了实验数据的准确性，实验中所有模型的超参数设置都保持一致，训练时批量大小设置为 16，训练迭代次数为 10 次，采用 Adam 梯度下降法来更新参数和优化模型，其学习率设置



图 7 水果数据集

Fig. 7 Fruit data set

为 0.0001。

图 8 给出了本文改进后模型的训练过程可视化展示，该可视化使用了 Grad-CAM 方法。在该模型过程的可视化分析中，选取新鲜苹果和腐烂苹果图片各 1 张，图中列出的 layer 1—layer 4 分别对应模型中 Block(64)、Block(128)、Block(256)、Block(512)。由图像效果可知，训练模型所学习到的特征随着层数的增加而逐渐显著，最终完成的模型也因此更容易捕捉到图像的关键信息。

2.4 模型评估指标

为了评价模型的性能，采用了准确率（accuracy）、精确率（precision）、召回率（recall）、特异度（specificity）作为评价指标^[16]。其中，准确率 A 为模型正确分类样本数占总样

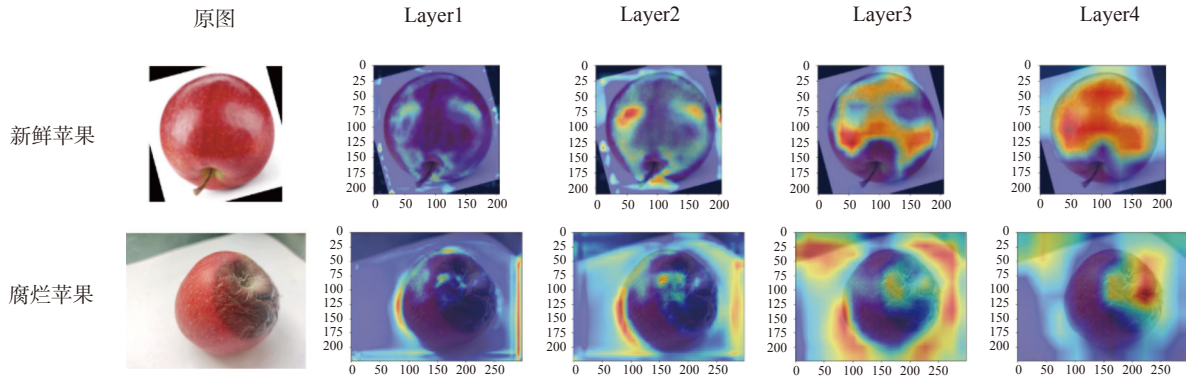


图 8 模型提取各层特征可视化热力图

Fig. 8 Visualization thermodynamic diagram of each layer feature extracted by mode

本数比例, 公式为

$$A = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \quad (6)$$

精确率 P 为模型预测的所有阳性样本中, 模型预测正确的比例, 公式为

$$P = \frac{T_p}{T_p + F_p} \quad (7)$$

召回率 R 为所有真实的阳性样本中, 模型预测正确的正样本比例, 公式为

$$R = \frac{T_p}{T_p + F_n} \quad (8)$$

特异度 S 为所有真实的阴样本中, 模型预测正确的负样本比例, 公式为

$$S = \frac{T_n}{T_n + F_p} \quad (9)$$

式中: T_p 为真阳性样本数量; F_p 为假阳性样本数量; T_n 为真阴性样本数量; F_n 为假阴性样本数量。

F_1 -Score 为综合评价指标, 公式为

$$F_1 = \frac{2R \times P}{R + P} \quad (10)$$

2.5 结果与分析

先将 VGG16^[17]、GoogLeNet^[18]、ResNet34、MobileNetV2^[19] 模型在该数据集进行实验, 结果如图 9、图 10 所示。ResNet34 网络训练的准确率和损失值都优于其他网络, 其准确率为 97.9%, 故将 ResNet34 作为主干网络。

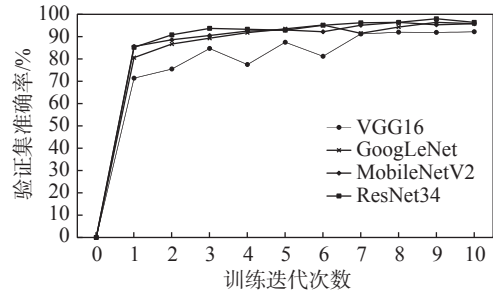


图 9 各网络验证集准确率结果对比

Fig. 9 Comparison of accuracy results of each network model for the validation set

然后分别测试了 ResNet34、ResNet34+SE、ResNet34+CBAM 和 ResNet34+SE+CBAM 网络的实验准确率, 并进行对比。结果显示, 改进后的 ResNet34+SE+CBAM 网络准确率最高, 为 98.8%, 比 ResNet34 提高了 0.9%, 比 ResNet34+SE 提高了 0.1%, 比 ResNet34+CBAM 提高了 0.5%。在精确率上, ResNet34+SE+CBAM 网络比 ResNet34 提高了 0.9%, 比 ResNet34+SE 提高了 0.1%, 比 ResNet34+CBAM 提高了 0.4%。

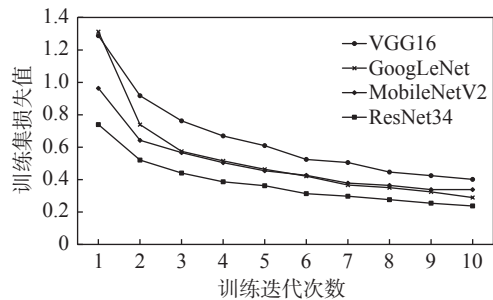


图 10 各网络训练集损失值对比

Fig. 10 Comparison of the loss value of each network model for the train set

在召回率上，ResNet34+SE+CBAM网络比ResNet34提高了1.1%，比ResNet34+SE提高了0.2%，比ResNet34+CBAM提高了0.6%。结果证明，改进后模型性能最佳，具体结果见表1。

表 1 各网络模型分类性能对比
Tab. 1 Comparison of classification performance of each network model

模型	准确率/%	精确率/%	召回率/%
ResNet34	97.9	97.9	97.9
ResNet34+SE	98.7	98.7	98.8
ResNet34+CBAM	98.3	98.4	98.4
ResNet34+SE+CBAM	98.8	98.8	99.0

表 2 为改进后的模型 ResNet34+SE+CBAM 在测试集中各品种预测精确率、召回率、特异度结果。该网络每类品种的特异度均超过 99.5%，说明模型的误分类率很小，但在腐烂苹果这个品类上精确率和召回率最低，说明模型的性能还有

提升的空间。

表 2 测试集各品种预测结果
Tab. 2 Results of each variety in test set

类别	精确率/%	召回率/%	特异度/%
新鲜苹果	98.2	99.3	99.5
新鲜香蕉	100.0	99.8	100.0
新鲜橙子	98.8	99.3	99.8
腐烂苹果	98.3	97.7	99.5
腐烂香蕉	99.6	99.3	99.9
腐烂橙子	99.5	98.2	99.9

表 3 为改进后的模型 ResNet34+SE+CBAM 与文献 [5] 的实验结果数据对比。该文献采用 MobileNetV2 模型进行实验，且使用与本文相同的数据集。通过对比可知，改进后模型的大部分实验数据都优于文献 [5] 报道的实验结果，说明改进后模型的识别效果显著提升。

表 3 测试集实验结果对比
Tab. 3 Comparison of experimental test set results

类别	精确率/%		召回率/%		F_1 -Score/%	
	ResNet34+SE+CBAM	文献[5]	ResNet34+SE+CBAM	文献[5]	ResNet34+SE+CBAM	文献[5]
新鲜苹果	98.2	98	99.3	99	98.7	97
新鲜香蕉	100.0	99	99.8	98	99.8	97
新鲜橙子	98.8	99	99.3	99	99.0	99
腐烂苹果	98.3	98	97.7	99	97.9	98
腐烂香蕉	99.6	99	99.3	98	99.4	98
腐烂橙子	99.5	99	98.2	99	98.8	99

由于公开的水果数据集中新鲜与腐烂水果的图片差异较大，特征区分较容易等原因，上述各网络在分类性能测试中精确率都很高^[20]。但实际在做水果分类时，一些损坏不那么明显的水果也经常需要予以剔除，如图 11(d) — (f) 所示：分别为虫蛀、缺水缩皱、挤压伤等。对于这类损伤，现有的数据集中相关样本较少，从而使得训练后模型难以分辨现实水果的腐坏程度，可能出现分类错误。为了测试各模型对这类水果损伤的分类效果，新拍摄了 334 张苹果图片作为补充数

据，其中 296 张为腐烂苹果，38 张为新鲜苹果，图 11 为拍摄样例，借此讨论新增数据对网络模型识别效果的提升作用。

由于新增的水果数据集数量较少，为了讨论新数据集加入后对实验结果的影响，在原数据集保持不变的基础上，将新数据集加入到验证集作为结果测试使用，而训练集保持不变。实验结果表明，经过训练之后，各个网络模型在验证集中测得的准确率均有所下降，特别是在苹果这个类别中的各项指标均有明显地下降，但改进后的

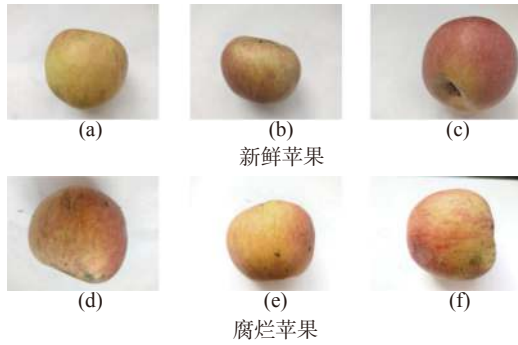


图 11 苹果数据集

Fig. 11 Apples data set

ResNet34+SE+CBAM 网络分类效果依然优于未改进前网络。表 4、表 5 分别展示了新数据加入验证集前后的各网络效果对比。

表 4 新数据加入验证集前改进前后的网络性能对比

Tab. 4 Comparison of performance effects before and after network improvement before the data is added into the validation set

类别	模型	精确率/%	召回率/%	特异度/%
新鲜苹果	ResNet34	97.8	98.8	99.2
	ResNet34+SE	98.0	99.2	99.4
	ResNet34+CBAM	97.9	98.9	99.3
	ResNet34+SE+CBAM	98.2	99.3	99.5
腐烂苹果	ResNet34	97.1	97.2	99.1
	ResNet34+SE	98.0	97.5	99.5
	ResNet34+CBAM	97.7	97.2	99.4
	ResNet34+SE+CBAM	98.3	97.7	99.5

表 5 新数据加入验证集后改进前后的网络性能对比

Tab. 5 Comparison of performance effects before and after network improvement after the data is added into the validation set

类别	模型	精确率/%	召回率/%	特异度/%
新鲜苹果	ResNet34	93.8	91.3	99.0
	ResNet34+SE	97.1	91.4	99.3
	ResNet34+CBAM	94.1	91.4	99.1
	ResNet34+SE+CBAM	97.3	91.6	99.5
腐烂苹果	ResNet34	96.0	89.2	97.6
	ResNet34+SE	96.5	92.6	97.4
	ResNet34+CBAM	96.1	95.1	97.8
	ResNet34+SE+CBAM	96.7	96.2	97.9

表 4 为原始数据集中的 10 次迭代测试结果, 因为原始数据集中腐烂水果图片损坏程度都

比较严重, 所以分类效果好。表 5 则加入了一些损坏不严重的样本, 可以看到网络模型的相关测试数据均有所下降, 但改进后模型分类效果仍有优势。

表 6 为训练集中加入新数据后, 网络改进前后的实验结果分析。经过 10 次训练迭代后, 从各个网络在苹果类别的各项指标中可以看出, 加入新数据后, ResNet34+SE+CBAM 网络的性能仍优于 ResNet34 网络, 损失值降低 0.052, 准确率提高 1.5%, 同样优于加入单个注意力机制模块的网络性能。说明在新增数据集后, 注意力机制模块仍有助于提升网络的识别能力。

表 6 网络改进前后在数据增强后的性能对比

Tab. 6 Comparison of performance effects after data enhancement before and after network improvement

类别	损失值	准确率/%
ResNet34	0.245	97.4
ResNet34+SE	0.195	98.0
ResNet34+CBAM	0.235	97.8
ResNet34+SE+CBAM	0.193	98.9

将新的苹果数据加入训练集, 经过损坏程度较低、不易区分的水果数据训练后, 各模型在苹果类别的指标都有所提升, 包括新鲜苹果和腐烂苹果两个类别, 如图 12、图 13 所示。结果对比显示, 注意力机制模块加入原网络是有效的, 改进后的 ResNet34+SE+CBAM 模型效果更佳。

训练集数据扩增之后, 在新鲜苹果和腐烂苹果这两个类别的对比实验中, ResNet34+SE+CBAM 网络分别比 ResNet34 网络的精确率提高了 1.2% 和 2.0%, 比 ResNet34+SE 网络提高 1.2% 和 0.4%, 比 ResNet34+CBAM 网络提高

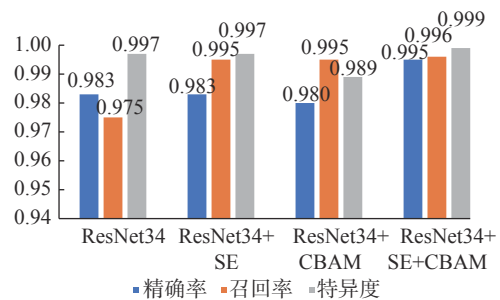


图 12 各网络在新鲜苹果类别的对比图

Fig. 12 Comparison of network in the fresh apples category

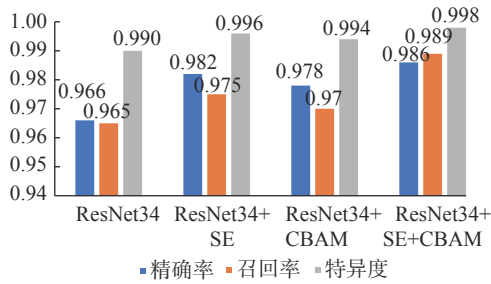


图 13 各网络在腐烂苹果类别的对比图

Fig. 13 Comparison of network in the rotten apples category

1.5% 和 0.8%；在召回率上，比 ResNet34 提高 2.1% 和 2.4%，比 ResNet34+SE 提高了 0.1% 和 1.4%，比 ResNet34+CBAM 提高 0.1% 和 1.9%。结果证明，在数据增强后，双注意力机制模块的加入使改进后的模型表现出较好的实验效果。

3 结 论

本文将 ResNet34 网络应用于缺陷水果分类数据集中，并将其与 VGG16、GoogLeNet、MobileNetV2 网络作对比。实验结果表明，ResNet34 网络在准确率较高的情况下，损失下降得最快，效果最好。为了优化模型，使其能在相同的迭代次数下达到较好的效果，在网络中加入了注意力机制 SE 和 CBAM，增强了模型特征提取过程中关注显著信息的能力，同时抑制无关的特征信息。实验对比发现：改进后的模型相比之前准确率提高了 0.9%，而且比分别单独加入一个注意力机制模块的效果更好，准确率分别提高了 0.1% 和 0.5%；在加入新的苹果数据集之后，改进后网络在新鲜苹果类别的精确率提高了 1.2%，在腐烂水果类别的精确率提高了 2.0%。在未来的工作中，将进一步研究更加细化的水果品质分类，用更优的分类网络来帮助实现更高的实验目标。

参考文献：

[1] MIAH M S, TASNUVA T, ISLAM M, et al. An advanced method of identification fresh and rotten fruits using different convolutional neural networks[C]//2021 12th International Conference on

Computing Communication and Networking Technologies (ICCCNT). Kharagpur: IEEE, 2021: 1 – 7.

- [2] KARAKAYA D, ULUCAN O, TURKAN M. A comparative analysis on fruit freshness classification[C]//2019 Innovations in Intelligent Systems and Applications Conference (ASYU). Izmir: IEEE, 2019: 1 – 4.
- [3] WAJID A, SINGH N K, JUNJUN P, et al. Recognition of ripe, unripe and scaled condition of orange citrus based on decision tree classification[C]//2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET). Sukkur: IEEE, 2018: 1 – 4.
- [4] SINGH S, SINGH N P. Machine learning-based classification of good and rotten apple[M]//KHARE A, TIWARY U, SETHI I, et al. Recent Trends in Communication, Computing, and Electronics. Singapore: Springer, 2019: 377 – 386.
- [5] CHAKRABORTY S, SHAMRAT F M J M, BILLAH M M, et al. Implementation of deep learning methods to identify rotten fruits[C]//2021 5th International Conference on Trends in Electronics and Informatics (ICOEI). Tirunelveli: IEEE, 2021: 1207 – 1212.
- [6] 苏宝峰, 沈磊, 陈山, 等. 基于注意力机制的葡萄品种多特征分类方法 [J]. 农业机械学报, 2021, 52(11): 226 – 233, 252.
- [7] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization[C]//Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 618 – 626.
- [8] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770 – 778.
- [9] 江紫薇. 基于神经网络的自适应梯度下降法的改进研究 [D]. 哈尔滨: 哈尔滨工业大学, 2020: 36 – 37.
- [10] 陈娟, 陈良勇, 王生生, 等. 基于改进残差网络的园林害虫图像识别 [J]. 农业机械学报, 2019, 50(5): 187 – 195.
- [11] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 7132 – 7141.
- [12] 蒋恺彦程. 基于图卷积网络的 miRNA-disease 关联预

- 测 [D]. 杭州: 浙江理工大学, 2021: 12 – 13.
- [13] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV). Munich: Springer, 2018: 3 – 19.
- [14] 薛瑞晨, 郝媛媛, 张振, 等. 基于改进 YOLOv3 的头盔佩戴检测算法 [J]. 电子测量技术, 2021, 44(12): 115 – 120.
- [15] 叶冲, 杨晶东. 基于 CBAM-EfficientNet 的垃圾图像分类算法研究 [J]. 智能计算机与应用, 2021, 11(5): 218 – 222.
- [16] 顾兆军, 杨文瑾, 周景贤. 基于迁移学习的小样本 DGA 恶意域名检测方法 [J]. 计算机工程与应用, 2021, 57(14): 103 – 109.
- [17] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]//3rd International Conference on Learning Representations. San Diego: ICLR, 2015.
- [18] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions[C]//Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015: 1 – 9.
- [19] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 4510 – 4520.
- [20] 朱冬梅. 基于深度学习的水果分类研究 [D]. 赣州: 赣南师范大学, 2020: 26 – 28.

(编辑: 李晓莉)