

## 基于注意力机制的水果新鲜度检测可解释模型

张寅升<sup>1</sup>, 宋曾林<sup>2</sup>, 王海燕<sup>1\*</sup>

<sup>1</sup>浙江工商大学 食药质量安全工程研究院 杭州 310018

<sup>2</sup>浙江工商大学管理工程与电子商务学院 杭州 310018)

**摘要** 近年来,基于机器视觉和深度学习的水果新鲜度检测成为主流方法之一。针对现有的深度学习技术,对卷积神经网络在水果的特征提取方面的应用进行探讨,在 ResNet34 主干网络中引入卷积注意力模块(CBAM),以实现水果新鲜度的检测,并采用类激活图(CAM)系列技术对于能够反映图片关键特征的像素进行热力图可视化。在水果公开数据集上,引入注意力机制前、后的 ResNet34 分类准确率分别为 96.80%和 99.71%。同时,CAM 热力图反映注意力模型能够更加准确地捕获水果图像中变质腐烂的区域,表明提出的模型改善了深度学习特征提取的能力,不仅提高了模型的泛化能力,而且增强了模型的可解释性。

**关键词** 水果新鲜度检测;深度学习;注意力机制;残差网络

**文章编号** 1009-7848(2024)10-0028-09 **DOI**: 10.16429/j.1009-7848.2024.10.003

水果的品质管理直接关系公众的健康状况和经济行为,水果的新鲜度判别显得尤为重要。目前,在传统机器学习方面,大多数水果新鲜度分类是通过人工设计特征提取后,经由 K-近邻(KNN)、支持向量机(Support vector machine, SVM)、浅层人工神经网络等模型解决分类问题<sup>[1]</sup>。Ghazal 等<sup>[2]</sup>提出一种结合色调(Hue)、彩色尺度不变特征变换(Color scale-invariant feature transform, Color-SIFT)、离散小波变换(Discrete wavelet transform, DWT)和 Haralick 统计量的特征提取方法,将其与朴素贝叶斯(Naive bayes)、线性判别分析(Linear discriminant analysis, LDA)、决策树(Decision trees, DT)等 6 种有监督机器学习技术相结合,分类精度在 99%~100%之间。Kumari 等<sup>[3]</sup>使用颜色和纹理特征进行水果分类。在 HSV【色调(Hue)、饱和度(Saturation)、明度(Value)】颜色空间中进行阈值分析以提取感兴趣的区域(ROI)。然后,在应用三级离散小波变换后,从色相和饱和通道(Hue and saturation channels)中提取颜色特征,从亮度通道(Luminance channel)中提取纹理特征。使用 SVM 分类器对超市农产品数

据集中的 10 种水果类别进行分类。系统的准确率为 95.3%。Goma 等<sup>[4]</sup>从 15 个类别的水果图像中提取颜色、大小、形状和纹理特征,通过 KNN 取得 81.9%的分类准确率。Bhargava 等<sup>[5]</sup>对 4 种水果的类型和质量进行分级。首先,使用高斯滤波、模糊聚类以及抓取切割法(Grab-Cut)对图像预处理;然后,从特征向量中提取和选取统计特征、颜色特征、纹理特征、几何特征、梯度直方图和离散小波变换等各种特征;最后,利用 SVM 取得 97.63%的分类识别率。Patel 等<sup>[6]</sup>以苹果、西瓜和芒果等 6 种水果为研究对象,使用阈值和形态学处理来查找感兴趣的区域。提取的特征包括面积、颜色、质心、区域、周长、大小和圆度。通过 SVM 取得 91.67%的分类准确率。Nossseir 等<sup>[7]</sup>基于颜色和纹理特征进行水果类型和新鲜度识别,准确率分别达到 96.3%和 98%。以上这些传统机器视觉方法在水果新鲜度分类中表现出良好的性能,然而,这些模型依赖人为设计的视觉特征提取方法,受目标对象的影响较大,具有一定的专业门槛和非通用性。

近年来,随着计算机视觉领域的不断发展,基于深度学习的水果质量判别开始得到广泛应用<sup>[8]</sup>。深度学习不需要人为进行特征提取,在部分领域弥补了传统方法的不足。Gai 等<sup>[9]</sup>提出一种改进的 YOLO-V4 深度学习算法来检测樱桃果实,准确率达到了 94.7%。Suhajito 等<sup>[10]</sup>提出一种结合迁移学习以及 float16 量化的模型 EfficientNetB0,对油棕进

收稿日期: 2024-10-28

基金项目: 国家重点研发计划项目(2023YFD1000400);  
国家自然科学基金项目(62376249,91746202)

第一作者: 张寅升,男,博士,副教授

通信作者: 王海燕 E-mail: njue2010@163.com

行检测,准确率达 89.3%。Kang 等<sup>[11]</sup>使用融合两种不同架构(ResNet-50 和 ResNet-101)的多任务深度卷积神经网络的瓶颈特征的集成模型,对包含多种新鲜和腐烂水果的数据集进行检测,模型在新鲜度分类和水果分类上分别实现 98.50% 和 97.43% 的平均准确率。Fahad 等<sup>[12]</sup>利用 VGG-16 卷积神经网络和 YOLO-v5 两种方法,对一个包含 11 种水果和蔬菜的 3 种新鲜度分类的数据集进行检测,平均准确率达到 84%。Amin 等<sup>[13]</sup>提出一种高效且准确的水果新鲜度分类方法,该方法首先对数据集进行颜色统一、图像缩放、数据增强和图像标记,随后,基于 AlexNet 模型进行分类预测。该模型在 3 个公开数据集上的准确率分别为 98.2%、99.8% 和 99.3%。Fu 等<sup>[14]</sup>利用 ResNet、VGG 和 GoogLeNet 等多种深度学习方法作为基础网络,使用 YOLO 从数字图像中提取感兴趣区域(ROI),完成了对苹果、香蕉等多种水果的新鲜度分级。Ananthanarayana 等<sup>[15]</sup>使用轻量级的卷积神经网络对水果新鲜度进行判别,取得 97% 的准确率。Kazi 等<sup>[16]</sup>使用多种经典卷积神经网络架构和一种残差卷积神经网络来进行水果新鲜度分类,根据测试得分评价最佳表现模型,试验结果表明残差网络表现最优,其测试准确率超过 99%。上述研究采用的深度学习方法在一定程度上有效解决了水果新鲜度分类的问题,然而,这些模型大多数侧重考虑其在性能上的表现,对于模型可解释性的研究和特征提取有效性的验证较少。

基于上述相关工作,本文在经典残差神经网络模型的基础上,尝试加入注意力机制进行改进,以增强模型对于水果新鲜度特征信息的关注和提取能力。此外,对模型注意力机制进行可视化,增强模型的可解释性和特征提取的验证能力。

## 1 理论和方法

### 1.1 残差网络模型

残差网络(Residual network, ResNet)是由 He 等<sup>[17]</sup>在微软研究院提出的一种深度学习架构。其亮点包括: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;解决了网络退化问题<sup>[18]</sup>,具有超深的网络结构,而不牺牲其性能;使用批归一化(Batch normalization)加速训练等。该

网络现已得到广泛使用,尤其是在图像分割和目标检测方向取得很好的效果。ResNet 残差模块如图 1 所示,当输入为  $x$  时,其学习到的特征记为  $H(x)$ ,残差为  $F(x)=H(x)-x$ 。相较于直接学习原始特征,残差学习更加容易。当残差为 0 时,网络退化为恒等映射,确保了信息留存的下限和网络性能的稳定。实际使用中,残差不会为 0,这意味着模型能够在原有输入特征的基础上学习到新特征,从而提升性能。

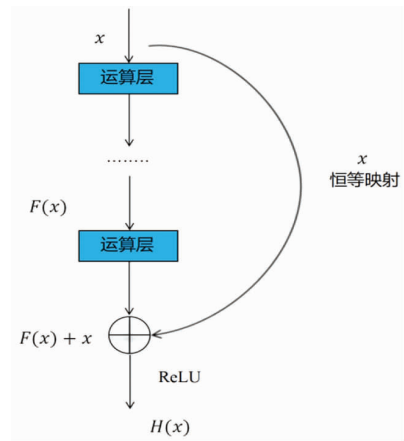


图 1 残差模块

Fig.1 The residual module

### 1.2 CBAM 模块

2018 年, Woo 等<sup>[19]</sup>提出了卷积块注意力模块(Convolutional block attention module, CBAM)。CBAM 由 2 个注意力子模块组成:通道注意力模块(Channel attention module, CAM)和空间注意力模块(Spatial attention module, SAM)。其结构如图 2 所示。

通道注意力机制(CAM)首先进行平均池化和最大池化操作聚合特征图的空间信息,生成平均池化特征( $F_{avg}^c$ )和最大池化特征( $F_{max}^c$ )。然后,这两个特征被送入多层感知器(MLP)共享网络,以生成通道注意力图并通过逐元素求和来合并输出特征向量  $M_c$ 。通道注意力计算如下:

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F)))$$

$$= \sigma(W_1(W_0(F_{avg}^c)) + W_1(W_0(F_{max}^c))) \quad (1)$$

式中,  $\sigma$  为 sigmoid 函数;  $F$  为输入特征;  $W_0$

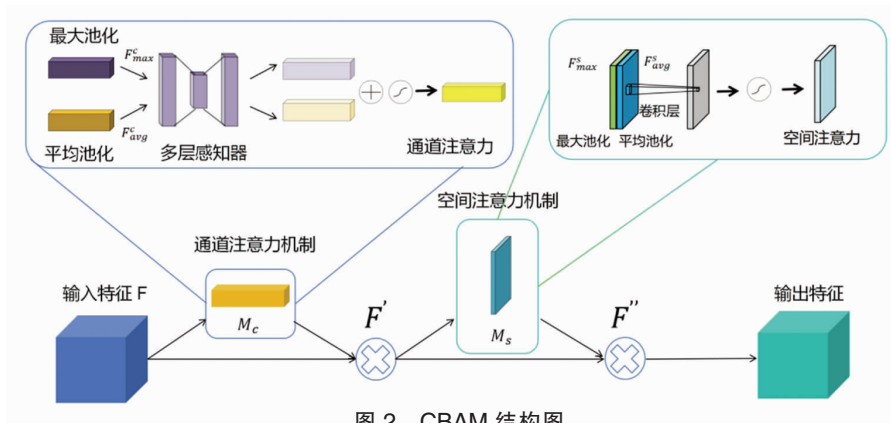


图2 CBAM 结构图

Fig.2 The structure of CBAM

和  $W_1$  为 MLP 权重,  $W_0 \in R^{C \times C}$ ,  $W_1 \in R^{C \times C \times r}$ ;  $C$  为通道数;  $r$  为缩减比率。CAM 可视作对压缩和激励 (Squeeze-and-excitation, SE) 结构<sup>[20]</sup>的扩展。经典的 SE 只有平均池化 (AvgPool) 通道, 而 CAM 增加了一个并行的最大池化 (MaxPool) 通道。平均池化重在聚合全局低频信息, 而最大池化可以额外捕获边缘和梯度等局部高频信息, 因此 CAM 的信息提取能力比 SE 更加全面。

空间注意力机制 (SAM) 首先通过使用 2 个池化操作来聚合特征图的通道信息, 生成 2 个二维图: 整个通道的最大池化特征 ( $F_{\max}^s$ ) 和平均池化特征 ( $F_{\text{avg}}^s$ ), 然后, 这 2 个特征通过卷积层, 生成二维空间注意力图。空间注意力计算公式如下:

$$M_s(F) = \sigma(f(\text{AvgPool}(F); \text{MaxPool}(F)))$$

$$= \sigma(f([F_{\text{avg}}^s; F_{\max}^s])) \quad (2)$$

式中,  $\sigma$  为 sigmoid 函数;  $F$  为输入特征;  $f$  为卷积运算。

给定输入特征  $F$ , CBAM 模块依次沿着通道注意力和空间注意力子模块产生注意力权重图, 通过将注意力权重图与输入特征相乘得到赋权的新特征。CBAM 是轻量级通用模块, 可以作为即插即用的模块集成到 ResNet 等主干网络架构中<sup>[21]</sup>。

### 1.3 模型结构

本文设计的 ResNet34-CBAM 模型结构如图 3 所示, CBAM 被附加到每个残差模块末端<sup>[21]</sup>, 保证了原始模型的主体拓扑结构, 确保预训练参数的可迁移性。移除网络中的 CBAM 则退化为经典的 ResNet34 模型。

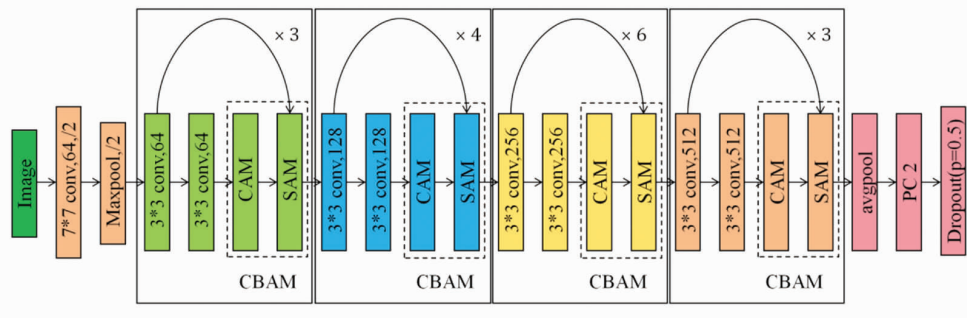


图3 ResNet34-CBAM 结构图

Fig.3 ResNet34-CBAM structure diagram

### 1.4 可视化方法

目前深度学习可解释性方法包括类激活图 (Class activation mapping, CAM)<sup>[22]</sup>、梯度加权类激

活映射 (Gradient-weighted class activation mapping, Grad-CAM)<sup>[23]</sup>、局部可解释模型 (Local interpretable model-agnostic explanations, LIME)<sup>[24]</sup>、加和



解释 (SHapley additive explanations, SHAP)<sup>[25]</sup> 等多种方法。本文着重考虑 CAM 系列可视化方法。CAM 是 Zhou 等<sup>[22]</sup> 提出的深度学习可视化技术。该方法通过利用模型最后一层卷积层的特征图与最终输出之间的关系来生成热力图。该热力图通过显示输入图像中对类别预测贡献最大的区域, 以此来提高模型的可解释性。Selvaraju 等<sup>[23]</sup> 在 CAM 方法<sup>[22]</sup> 的基础上提出 Grad-CAM 方法, 克服了 CAM 方法中替换分类器后需要重新训练模型的缺点。Omeiza 等<sup>[26]</sup> 在 Grad-CAM 和 SmoothGrad<sup>[27]</sup> 的基础上建立了 SmoothGradCAM++ 方法。该方法是通过减少冗余和增强重要性权重, 来提高激活图的质量。它通过在多个扰动的输入上计算梯度, 然后采用一个更加精细的加权方案来组合这些梯度, 从而生成更加清晰和准确的激活热力图。Selvaraju 等<sup>[28]</sup> 提出 Grad-CAM++ 方法, 该方法通过使用更精细的梯度加权方案来改进激活图的生成, 以解决原始 Grad-CAM 方法在某些情况下可能无法精确反映特征图对模型决策的实际贡献问题。Wang 等<sup>[29]</sup> 提出一种新的可视化方法 Score-CAM, 该方法不需要梯度信息, 而是直接基于输出分数来估计特征映射的重要性, 这使得它更加通用且不易受梯度消失或爆炸问题的影响。Wang 等<sup>[30]</sup> 在 Score-CAM 的基础上提出 SSCAM 方法, 旨在通过减少噪声并提供更平滑的响应, 从而生成更清晰、更可解释的激活图。SSCAM 的主要创新在于引入“平滑”(smoothing) 技术, 这种方法通过在生成 CAM 前对输入图像进行多次轻微随机扰动来实现。Fu 等<sup>[31]</sup> 提出 XGrad-CAM 方法, 该方法在敏感性和保守性方面是对 Grad-CAM 的一种增强版本。它比 Grad-CAM 可实现更好的可视化性能, 同时相较于 Grad-CAM++ 和 Ablation-CAM, 还具有类别区分能力和易于实现的优点。Jiang 等<sup>[32]</sup> 提出 Layer-CAM 方法, 该方法主要在于对深度神经网络 (DNNs) 中不同层的特征图进行加权求和, 以生成特定类别的热力图, 显示输入图像中与分类决策相关的区域。

图 4 展示上述 CAM 技术在青提公开数据集图片上的热力图生成效果。图中暖色调较深 (红色) 的区域表示当前位置权重较高, 是决定当前分类的关键像素特征。该方法可以用来解释和诊断

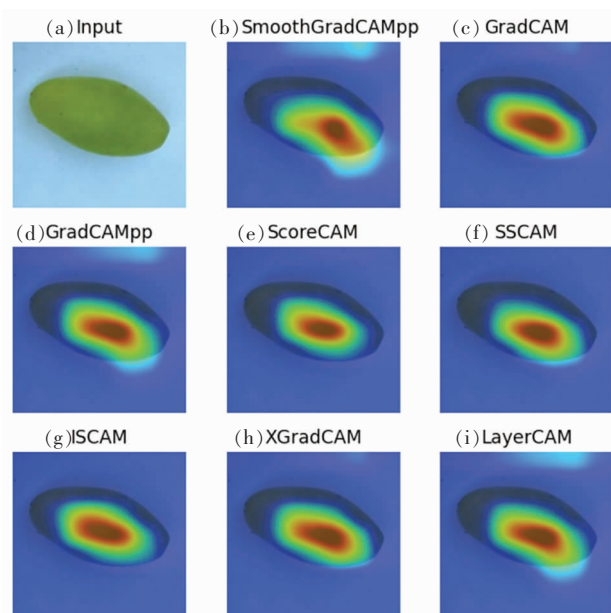


图 4 8 种 CAM 可视化方法的示意图

Fig.4 Examples of the eight CAM visualization methods

深度学习模型。

## 2 试验及结果分析

### 2.1 试验平台和数据集

试验运行环境为 Python 3.8.10, 深度学习框架为 Pytorch, 版本为 1.13.1。硬件 CPU 型号为 i7-13700H, GPU 型号为 RTX 4060, 显存为 8 GB。本文使用的数据集是飞桨 AI Studio 网站上的青提公开数据集, 共 2 类, 分别是新鲜青提和腐烂青提 (如图 5 所示)。数据集大小为 545 MB, 图片数量为 1 502 张, 按 4:1 划分训练集和测试集。

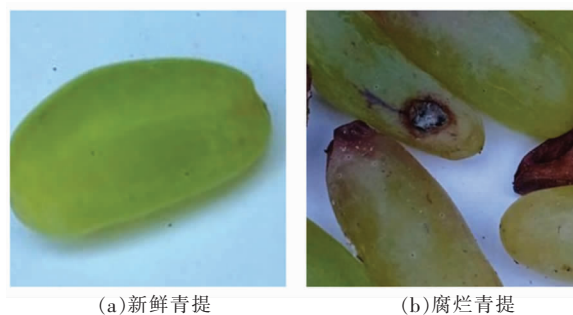


图 5 青提数据集

Fig.5 The grape dataset

## 2.2 模型训练与参数调整

表1显示ResNet34和ResNet34-CBAM模型的参数规模。后者整体参数略有增加,实测模型的前向推理速度并未受到显著影响。表2显示模型训练过程使用的超参数,2个模型采用相同的超参数配置。Shuffle操作用于随机打乱训练样本次序,提高网络的泛化性能。经过学习曲线调试,最终确定训练超参数组合是:训练轮次(Epoch)为30,批大小(Batch size)为24,优化器为自适应矩估计(Adaptive moment estimation,Adam),学习率为0.001。

## 2.3 模型评估指标

为了评价模型的性能,采用准确率(Accuracy)、精确率(Precision)、召回率(Recall)、特异度(Specificity)<sup>[33]</sup>以及模型提取特征的能力作为评价指标。其中,模型提取特征的能力指模型在进行分类预测时,准确提取关键特征作为分类依据的能力。本文通过可视化技术来呈现。

准确率为模型正确分类样本数占总样本数比例,计算公式:

$$\text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (3)$$

精确率衡量的是预测为正类的样本中有多少比例实际上是正类,计算公式:

$$\text{Pre} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (4)$$

召回率衡量的是实际为正类的样本中有多少比例被正确地预测为正类,计算公式如下:

$$\text{Rec} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (5)$$

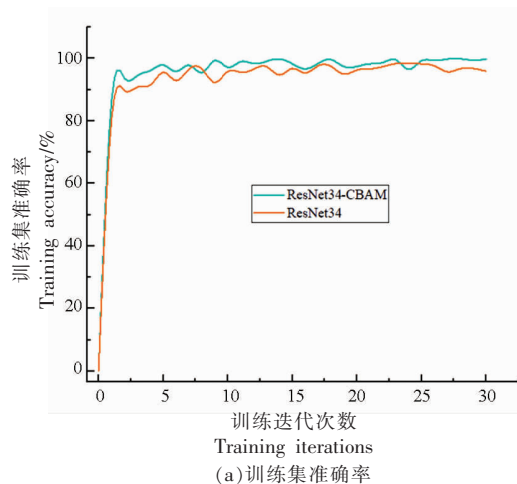


表1 两种ResNet34模型参数比较

Table 1 Comparison of the two ResNet34 models' parameters

模型	参数规模
ResNet34	21 285 698
ResNet34-CBAM	21 444 450

表2 两种ResNet34模型训练超参数

Table 2 Hyper parameters for training the two ResNet34 models

参数	属性
是否随机(Shuffle)	True
训练轮次(Epochs)	30
批大小(Batch size)	24
优化器(Optimizer)	Adam
学习率(Learning rate)	0.001

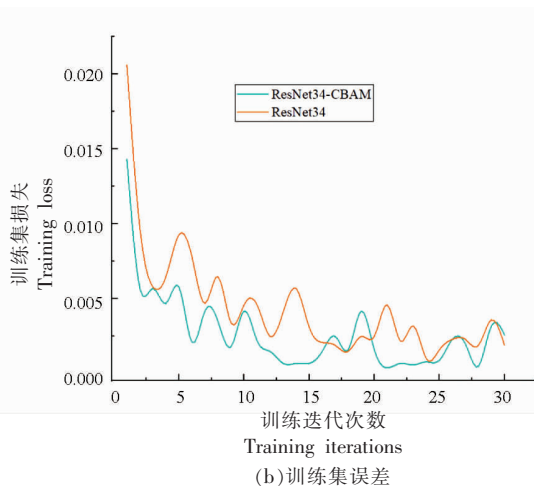
特异度衡量的是实际为负类的样本中有多少比例被正确地预测为负类,计算公式:

$$\text{Spe} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (6)$$

式中,TP(True positives)是真实类别为正类且被模型正确分类为正类的样本数;TN(True negatives)是真实类别为负类且被模型正确分类为负类的样本数;FP(False positives)是真实类别为负类而被模型错误分类为正类的样本数;FN(False negatives)是真实类别为正类而被模型错误分类为负类的样本数。

## 2.4 结果与分析

ResNet34和ResNet34-CBAM两种模型训练的准确率和损失函数如图6所示。从学习曲线来



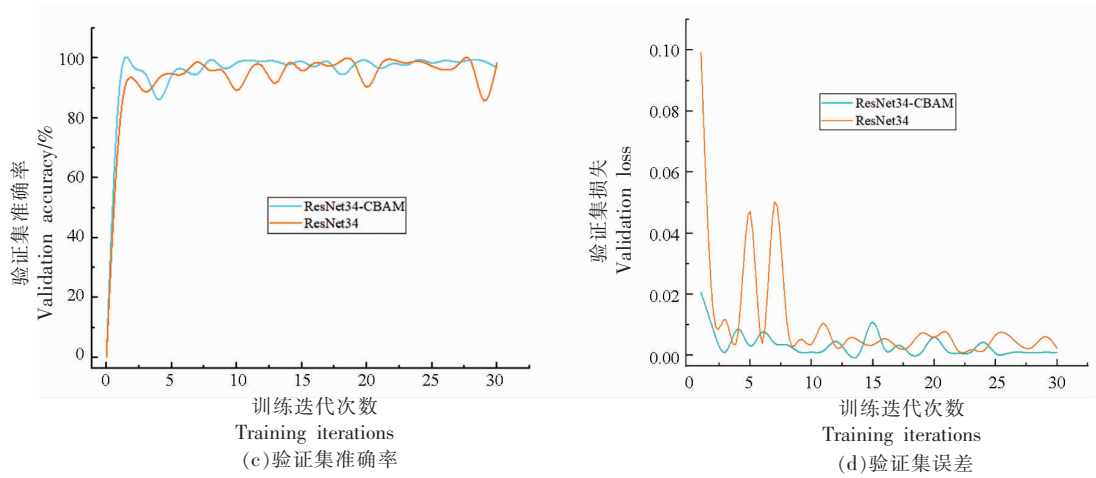


图 6 ResNet34 和 ResNet34-CBAM 两种模型训练过程中的准确率和误差

Fig.6 Classification accuracy and loss of ResNet34 and ResNet34-CBAM during model training

表 3 模型性能对比

Table 3 Comparison of model performance

模型	准确率/%	精确率/%	召回率/%	特异度/%
ResNet34	96.8	96.5	96.5	93
ResNet34-CBAM	99.71	99.0	99.0	99.0

看, 训练过程中, ResNet34-CBAM 的损失整体上低于 ResNet34, 前者准确率也高于后者。完成训练后, 测试两种模型在预留测试集中的准确率。如表 3 所示, ResNet34-CBAM 准确率为 99.71%, 比 ResNet34 提高了 3.01%, 精确率提高了 2.59%, 召回率提高了 2.59%, 特异度提高了 6.45%。

接下来, 分别测试两种模型提取特征的能力, 即能否准确识别出新鲜青提及腐烂青提的感兴趣区域(Region of interest, ROI)。选取新鲜青提和腐烂青提图片各 3 张, 然后采用两种模型进行预测, 之后采用 CAM 可视化方法画出热力图, 选择典型效果图进行对比分析。

图 7 是两种模型对新鲜青提进行处理后的结果图。每一排代表同一张图片两种模型的处理结果。从图中可以看出两种模型都能够识别出图片中的新鲜青提, 并且能够大致准确地将其标注出来。然而, 与 ResNet34 相比, ResNet34-CBAM 赋予青提主体更高的权重, ROI 的识别更加精准。而 ResNet34 受到边缘特征的干扰, 导致 ROI 部分覆盖到青提主体和背景之间的过渡区域。

图 8 为两种模型对于腐烂青提的处理结果。

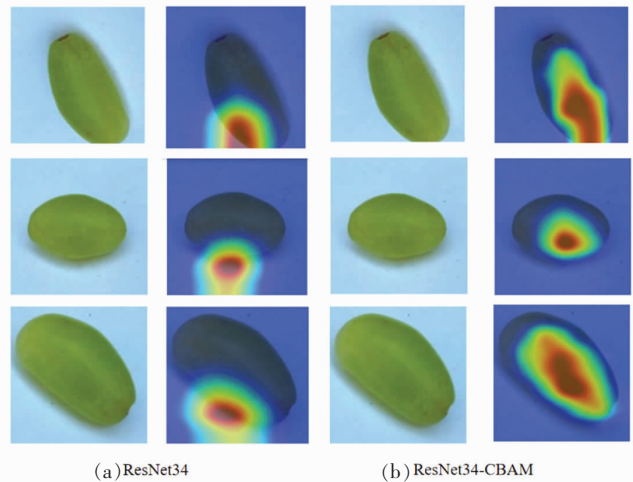


图 7 新鲜青提可视化结果对比

Fig.7 Comparison of visualization results of fresh grapes

对比图 8 中的两种模型, ResNet34 模型无法有效关注对象主体, 受背景和高梯度边缘影响较大。ResNet34-CBAM 效果更优, 提取到的 ROI 基本覆盖青提腐烂区域。以上结果直观地表明注意力模块的引入有利于模型捕获关键视觉特征, 使得推理结果更具可解释性, 推理过程更有说服力。

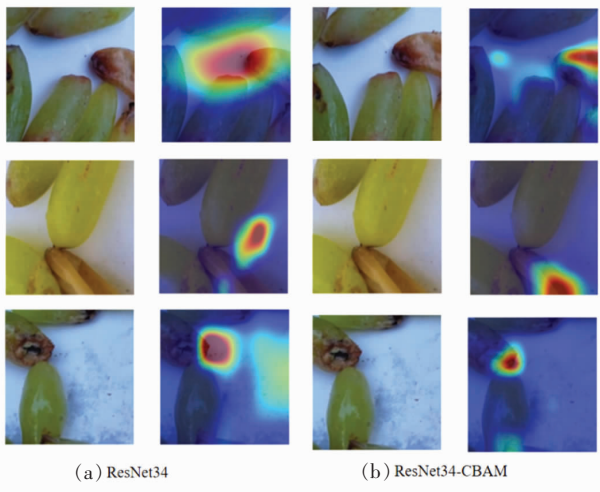


图8 腐烂青提可视化结果对比

Fig.8 Comparison of visualization results of rotten grapes

### 3 结论

本文将 ResNet34 模型应用于水果新鲜度检测,并在此基础上引入注意力机制,构建了 ResNet34-CBAM 模型。试验结果显示,ResNet34-CBAM 模型在分类任务中的准确率达到 99.71%,相较于原始 ResNet34 提高了 3.01%。此外,在精确率、召回率、特异度方面也有提升。通过对比两种模型的 CAM 结果,发现引入 CBAM 注意力机制不仅增强了模型的关键特征提取能力,还显著提升了整体性能和泛化能力,验证了注意力机制在机器视觉应用的有效性和实用性。后续有望将该方法拓展到其它食品对象的新鲜度检测和质量安全管理场景中。

### 参 考 文 献

- [1] 鄢紫,陈良艳,刘卫华,等.基于YOLO-FFD的水果品种和新鲜度识别方法[J].食品与机械,2024,40(1):115-121.  
YAN Z, CHEN L Y, LIU W H, et al. Fruit variety and freshness recognition method based on YOLO-FFD[J]. Food and Machinery, 2024, 40(1): 115-121.
- [2] GHAZAL S, QURESHI W S, KHAN U S, et al. Analysis of visual features and classifiers for fruit classification problem[J]. Computers and Electronics in Agriculture, 2021, 187: 106267.
- [3] KUMARI R S S, GOMATHY V. Fruit classification using statistical features in SVM classifier[C]. 2018 4th International Conference on Electrical Energy Systems (ICEES). Chennai: IEEE, 2018: 526-529.
- [4] GOMA J D, QUILAS C A M, VALERIO M A B, et al. Fruit recognition using surface and geometric information[J]. Telecommunication Electronic Comput Eng, 2018, 10(1): 39-42.
- [5] BHARGAVA A, BANSAL A, GOYAL V. Machine learning-based detection and sorting of multiple vegetables and fruits Food Anal[J]. Methods, 2022, 15(1): 228-242.
- [6] PATEL C C, CHAUDHARI V K. Comparative analysis of fruit categorization using different classifiers [J], Advanced Engineering Optimization Through Intelligent Techniques, 2019, 949: 153-164.
- [7] NOSSEIR A, AHMED S E A. Automatic classification for fruits' types and identification of rotten ones using k-NN and SVM[J]. International Journal of Online and Biomedical Engineering, 2019, 15: 47-61.
- [8] 张杰,夏春蕾,张荣福,等.基于注意力机制的水果损伤检测及分类[J].光学仪器,2023,45(2):26-35.  
ZHANG J, XIA C L, ZHANG R F, et al. Fruit damage detection and classification based on attention mechanism[J]. Optical Instruments, 2023, 45(2): 26-35.
- [9] GAI R L, CHEN N, YUAN H. A detection algorithm for cherry fruits based on the improved YOLOGv4 model[J]. Neural Computing and Applications, 2023, 35(19): 13895-13906.
- [10] SUHARJITO, ELWIREHARDJA G, PRAYOGA J. Oil palm fresh fruit bunch ripeness classification on mobile devices using deep learning approaches [J]. Computers and Electronics in Agriculture, 2021, 188: 106359.
- [11] KANG J, GWAK J. Ensemble of multi-task deep convolutional neural networks using transfer learning for fruit freshness classification[J]. Multimedia Tools and Applications, 2022, 81: 22355-22377.
- [12] FAHAD L G, TAHIR S F, RASHEED U, et al. Fruits and vegetables freshness categorization using deep learning[J]. Computers, Materials & Continua, 2022, 71(3): 5083-5098.
- [13] AMIN U, SHAHZAD M I, SHAHZAD A, et al. Automatic fruits freshness classification using CNN



- and transfer learning[J]. *Applied Science*, 2023, 13(14): 8087.
- [14] FU Y, NGUYEN M, WANG Q Y. Grading methods for fruit freshness based on deep learning[J]. *SN Computer Science*, 2022, 3(4): 264.
- [15] ANANTHANARAYANA T, PTUCHA R, KELLY S C. Deep learning based fruit freshness classification and detection with CMOS image sensors and edge processors[J]. *Electronic Imaging*, 2020, 13(3): 298.
- [16] KAZI A, PANDA S P. Determining the freshness of fruits in the food industry by image classification using transfer learning[J]. *Multimedia Tools and Applications*, 2022, 81: 7611–7624.
- [17] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]. *Conference on Computer Vision and Pattern Recognition, CVPR, Las Vegas, NV, USA: IEEE*, 2016: 770–778.
- [18] 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.
- CHEN J, CHEN L Y, WANG S S, et al. Image recognition of garden pest based on improved residual network[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2019, 50(5): 187–195.
- [19] 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.
- [20] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. *Proceedings of the IEEE conference on computer vision and pattern recognition*. Salt Lake City, UT, USA: IEEE, 2018: 7132–7141.
- [21] WANG K, CAO B, ZHANG J, et al. Explainable deep learning study for leaf disease classification[J]. *Agronomy*, 2022, 12(5): 1035.
- [22] ZHOU B, KHOSLA A, LAPEDRIZA A, et al. Learning deep features for discriminative localization [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 2921–2929.
- [23] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-cam: Visual explanations from deep networks via gradient-based localization[C]// *The IEEE International Conference on Computer Vision*. Venice, Italy: IEEE, 2017: 618–626.
- [24] RIBEIRO M T, SINGH S, GUESTRIN C. ‘Why Should I Trust You?’: Explaining the predictions of any classifier [C]. *The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco, CA, USA: Association for Computational Linguistics, 2016: 1135–1144.
- [25] LUNDBERG S M, LEE S. A unified approach to interpreting model predictions[C]// *Neural Information Processing Systems*, Red Hook, NY, USA: NIPS, 2017: 4768–4777.
- [26] OMEIZA D, SPEAKMAN S, CINTAS C, et al. Smooth Grad-CAM++: An enhanced inference level visualization technique for deep convolutional neural network models[C]. In *the Intelligent Systems Conference*. London: SAI, 2019.
- [27] SMILKOV D, THORAT N, KIM B, et al. Smoothgrad: Removing noise by adding noise[C]. *The International Conference on Machine Learning (ICML) Workshop on Visualization for Deep Learning*. Sydney, Australia: Curran Associates, 2017.
- [28] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM++: Improved visual explanations for deep convolutional neural networks[J]. *Pattern Recognition Letters*, 2019, 117: 182–190.
- [29] WANG H, JIANG Y, WANG Y, et al. Score-CAM: Score-weighted visual explanations for convolutional neural networks[C]. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, WA, USA: IEEE, 2020: 24–25.
- [30] WANG H, NAIDU R, MICHAEL J, et al. SS-CAM: Smoothed score-CAM for sharper visual feature localization[C]// *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, USA: IEEE, 2020.
- [31] FU R, HU Q, DONG X, et al. Axiom-based Grad-CAM: Towards accurate visualization and explanation of CNNs[J]. *BMVC*, 2020, 32(2): 134–152.
- [32] JIANG P T, ZHANG C B, HOU Q, et al. Layer-CAM: Exploring hierarchical class activation maps for localization[J]. *IEEE Transactions on Image Processing*, 2021, 30: 5875–5888.
- [33] 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.



GU Z J, YANG W J, ZHOU J X. Small sample  
based on transfer learning DGA Malicious domain

name detection method[J]. Computer Engineering and  
Applications, 2021, 57(14): 103–109.

## Fruit Freshness Detection Explainable Model Based on Attention Mechanism

Zhang Yinsheng<sup>1</sup>, Song Zenglin<sup>2</sup>, Wang Haiyan<sup>1\*</sup>

(<sup>1</sup>Zhejiang Food and Drug Quality & Safety Engineering Research Institute,  
Zhejiang Gongshang University, Hangzhou 310018

<sup>2</sup>School of Management and E-Business, Zhejiang Gongshang University, Hangzhou 310018)

**Abstract** In recent years, fruit freshness detection based on machine vision and deep learning has become one of the mainstream methods. This study explores the application of deep learning technologies, particularly convolutional neural networks (CNNs), in feature extraction for fruit freshness identification. This paper introduces the CBAM (Convolutional block attention module) attention mechanism module into the ResNet34 (34-layer residual network) backbone network to achieve fruit freshness detection. Class activation mapping (CAM) techniques are employed to visualize the heatmaps of pixels that reflect the critical features in the images. On a public fruit dataset, the classification accuracy of the ResNet34 network before and after introducing the attention mechanism is 96.80% and 99.71%, respectively. The CAM heatmaps show that the attention model can more accurately capture the regions of interest in the fruit images, indicating that the proposed model improves the feature extraction capability of deep learning, not only enhancing the model's generalization ability but also increasing its interpretability.

**Keywords** freshness detection; deep learning; attention mechanism; residual network