

基于卷积神经网络的鲜茶叶分类方法研究

杨会芳, 傅秋椰, 肖黄梅*, 程 静

重庆对外经贸学院大数据与智能工程学院(工业软件产业学院), 重庆

收稿日期: 2025年3月6日; 录用日期: 2025年5月12日; 发布日期: 2025年5月20日

摘 要

针对鲜茶叶分类任务中的技术瓶颈, 本文提出一种基于AlexNet的改进架构, 通过四项核心技术有效突破传统方法的局限性。首先, 在AlexNet模型Conv3~Conv5层后嵌入通道注意力机制(Squeeze-and-Excitation, SE)模块, 自适应地调整通道的特征响应以增强模型对特征的表达能力; 其次, 在全卷积层后添加批归一化(Batch Normalization)层, 使鲜叶数据集的训练收敛轮次由50轮优化至20轮; 另外, 在模型输出层前加入全局平均池化(Global Average Pooling, GAP)层, 成功将模型参数量从178.39 M压缩至14.3 M。实验结果表明, 所提出的改进方案在保证分类精度的前提下, 显著提升了模型的训练效率与泛化能力。

关键词

卷积神经网络, AlexNet, 鲜茶叶分类

Research on Classification Method of Fresh Tea Leaves Based on Convolutional Neural Network

Huifang Yang, Qiuye Fu, Huangmei Xiao*, Jing Cheng

School of Big Data and Intelligent Engineering (Industrial Software Engineering School), Chongqing College of International Business and Economics, Chongqing

Received: Mar. 6th, 2025; accepted: May 12th, 2025; published: May 20th, 2025

Abstract

In order to solve the technical bottleneck in the task of sorting fresh tea leaves, this paper proposes an improved architecture based on AlexNet, which effectively breaks through the limitations of traditional methods through four core technologies. Firstly, the Squeeze-and-Excitation (SE) module was embedded in the Conv3~Conv5 layers of the AlexNet model to adaptively adjust the feature response

*通讯作者。

文章引用: 杨会芳, 傅秋椰, 肖黄梅, 程静. 基于卷积神经网络的鲜茶叶分类方法研究[J]. 人工智能与机器人研究, 2025, 14(3): 548-555. DOI: 10.12677/airr.2025.143054

of the channel to enhance the model's ability to express features. Secondly, a Batch Normalization layer was added after the full convolutional layer, so that the training convergence rounds of the fresh leaf dataset were optimized from 50 rounds to 20 rounds. In addition, the Global Average Pooling (GAP) layer was added before the model output layer, and the model parameters were successfully compressed from 178.39 M to 14.3 M. Experimental results show that the proposed improved scheme significantly improves the training efficiency and generalization ability of the model under the premise of ensuring classification accuracy.

Keywords

Convolutional Neural Network, AlexNet, Classification of Fresh Tea Leaves

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

目前, 基于机器视觉和深度学习的方法在农产品检测[1]、价格预测[2]和分级分选[3]等领域得到了广泛应用。鲜茶叶的精准分类作为茶叶生产加工中的关键环节, 直接影响茶叶的品质分级、品种鉴别及后续加工工艺的优化。传统分类方法依赖人工经验或浅层图像特征, 如颜色、纹理等易受光照变化、叶片重叠及主观因素干扰, 从而导致成本高、效率低、泛化性差。如早期茶叶分级主要依赖物理特征分析, 通过分级机基于叶片尺寸、形状等参数实现粗筛, 此类方法虽能剔除残次叶片, 但难以应对鲜茶叶的形态复杂性, 分类精度受限[4]。随着检测技术的发展, 研究者尝试结合多源信息提升分类效果。黄艳等利用近红外光谱与气相离子迁移谱进行主成分分析[5], 还有通过电子鼻传感器信号构建 PCA 分类模型[6], 虽能满足分类要求, 但这些方法存在硬件成本高、检测流程复杂等推广难题。在特征工程层面, 结合边缘与颜色特征实现 99.45% 的准确率[7], 以及融合形状纹理特征达到 98.5% 的识别率[8], 但严重依赖人工特征提取, 面对茶叶姿态变化、光照差异等复杂场景时泛化能力不足。

近年来, 机器视觉与深度学习技术在农产品检测与分选领域展现出显著优势, 其中, 卷积神经网络 (Convolutional Neural Networks, CNN) 因其强大的特征提取能力在农业图像识别领域得到广泛应用, 其在茶叶分类研究中也逐渐成为热点。很多研究者尝试通过改进网络结构与训练策略优化分类效果, 如采用 ResNet 进行迁移学习, 利用 ImageNet 预训练权重解决小样本问题[9], 但通用图像特征与鲜茶叶的细粒度形态特征存在显著差异而导致深层特征适应性弱。在特征融合方面, 张金炎等在传统 HOG 特征的基础上引入 CNN 深层语义特征, 构建多尺度特征金字塔, 增强叶片细节表征能力[10], 以及针对新鲜茶叶的形态特异性, 将轮廓角点检测模块与 CNN 结合, 通过距离矩阵相似度计算实现几何特征的自适应提取[11]。上述方法在一定程度上可以提高新鲜茶叶的分类精度, 不过需注意, 依赖手工设计规则是不能实现几何特征与 CNN 深层语义特征的端到端联合优化, 会导致特征对齐误差与信息冗余。也有学者通过深度可分离卷积、通道剪枝等技术压缩模型规模, 将参数量降低[12]。但关注参数量压缩, 容易忽略实际部署中内存占用、实时推理延迟等关键指标, 难以满足茶园边缘设备的资源约束需求。

针对以上问题, 本文从模型架构出发, 根据经典模型 AlexNet 在鲜叶数据集上表现出的局限, 如需大量计算资源、收敛速度慢以及全连接层参数冗余等局限, 提出一种改进的 AlexNet 模型, 通过注意力机制引导特征增强、加入批归一化提升训练速度, 并使模型轻量化等策略的适配, 构建高效、鲁棒的鲜茶叶分类模型。

2. 数据采集与预处理

2.1. 数据采集

本研究中的鲜茶叶样本数据采集自重庆市永川区茶山竹海核心产茶区，数据采集工作于 2024 年 9 月秋季茶芽萌发旺盛期进行，使用华为荣耀 30 智能手机在自然光条件下进行标准化采集，重点采集茶树顶端嫩梢，拍摄距离 30 cm 左右。本次共采集 123 张图像，均为 RGB 图。

2.2. 数据预处理

1) 图像裁剪

为方便鲜茶叶分类，需要对采集的茶叶图像进行预处理。由于图像中干扰信息太多，背景区域占比太大，鲜茶叶目标太小，不利于目标的提取，所以需要图像进行裁剪以去除图中的无关背景，并增强目标信息，以实现鲜叶与背景分离。在裁剪过程中，由于人工裁剪速度慢，工作量大，所以为了提高裁剪效率，使用 makesense.ai 在线网页工具对这三类数据中的目标进行画框裁剪，将画框图片输入 YOLOv5 模型进行训练，再把剩下的图像全部输入训练好的模型进行目标检测以及目标裁剪，经过裁剪得到三种不同形状的新鲜茶叶，并根据其形状分为工厂中常见的单芽、一芽一叶和一芽二叶，如图 1 所示。裁剪后的数据集包括单芽 104 张，一芽一叶 112 张，一芽二叶 104 张。

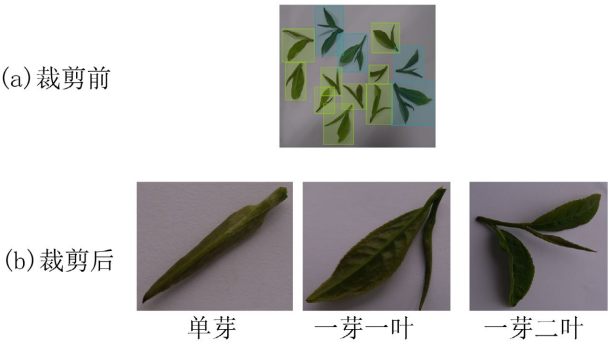


Figure 1. Data cropping
图 1. 数据裁剪

2) 数据增强

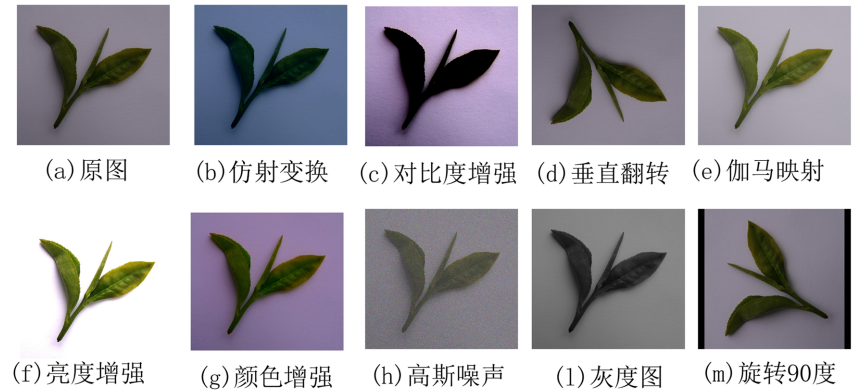


Figure 2. Data augmentation
图 2. 数据增强

为了防止因数据量少而导致过拟合现象，以及保证数据的平衡性，将对训练数据集进行数据增强处理，所谓的数据增强就是针对现有数据进行数据扩展，增加数据数量。通过随机变换来获得多种可信图像，从而提升模型的泛化能力。本文通过将所有训练样本进行仿射变换、3 倍对比度增强、垂直翻转、伽马映射、2.5 倍亮度增强、3 倍颜色增强、高斯噪声、灰度处理以及 90 度逆时针旋转等预处理，最终获得了 3200 张训练数据图像。数据增强后的部分样本图像如图 2 所示。将增强后的鲜茶叶数据图像按照 8:2 的比例分别划分为训练集和测试集，共有训练集 2560 张和测试集 640 张，鲜茶叶分类数据集具体见表 1。

Table 1. Classification dataset of fresh tea leaves

表 1. 鲜茶叶分类数据集

数据类别	训练集/张	测试集/张
单芽	835	205
一芽一叶	875	245
一芽二叶	850	190
合计	2560	640

3. 模型架构

3.1. AlexNet 模型

AlexNet 是由 Hinton 和他的学生在 2012 年设计的网络结构，如图 3 所示。AlexNet 结构主要包括五个卷积层、三个最大池化层和两个全连接层，同时通过重复使用 RELU 激活函数，可获得更多的线性变换，从而进一步提高学习能力。卷积层在卷积神经网络模型中扮演着重要的角色，它的重要功能就是提取输入模型结构的图像的不同特征。池化层是对数据进行下采样处理，减少参数的数量，防止神经网络模型在训练过程中产生过拟合的现象，常见的方式有最大池化和平均池化。全连接层主要设置在模型结构的末端，主要是将前一层的全部神经元与后一层的每个神经元进行全连接，它的功能是整合模型结构中经过若干卷积层和池化层处理后带有区分不同种类的局部信息数据。同时，AlexNet 网络在训练过程中加入了使神经元随机失活的 Dropout 层，使得它的性能能够超过传统的机器学习，成为第一个经典的卷积神经网络模型。

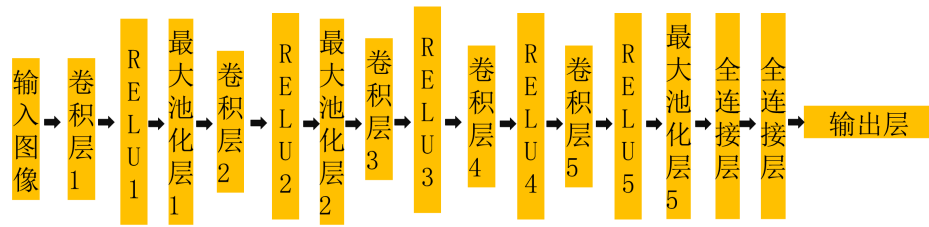


Figure 3. AlexNet model

图 3. AlexNet 模型

3.2. 改进 AlexNet 模型

本文改进的 AlexNet 模型结构如图 4 所示，改进内容如下：

1) 通道注意力增强

由于卷积神经网络中不同通道往往有不同的特征响应，在 AlexNet 模型架构中，Conv3~Conv5 层包

含了丰富的语义信息, 以及特征图数量多和特征信息复杂, 故在深层卷积层(Conv3~Conv5)后嵌入 SE 模块(Squeeze-and-Excitation)。首先对特征图进行空间特征压缩, 在空间维度上实现全局平均池化, 得到 $1 \times 1 \times C$ 的特征图, 其次通过全连接层学习, 得到具有通道注意力的特征图, 它的维度还是 $1 \times 1 \times C$, 最后将通道注意力 $1 \times 1 \times C$ 的特征图和原始输入 $H \times W \times C$ 的特征图进行逐通道乘以权重系数, 最终输出具有通道注意力的特征图, 见图 5。SE 通道注意力机制能够自适应地调整每个通道的特征响应权重, 并学习对当前任务更为重要的通道特征, 然后增强这些重要通道特征表示, 从而让网络更加聚焦于关键信息。

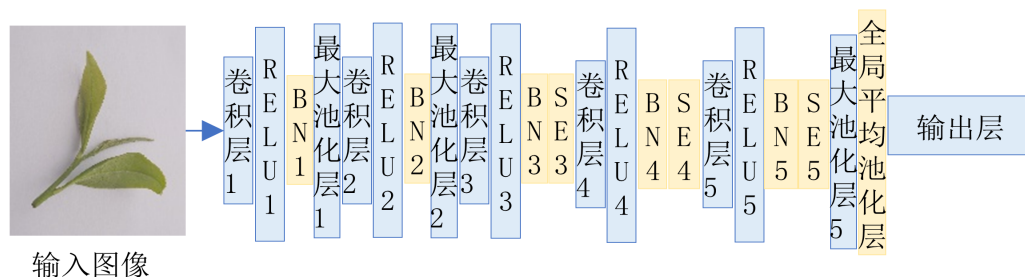


Figure 4. Improved AlexNet structure figure

图 4. Improved AlexNet 结构图

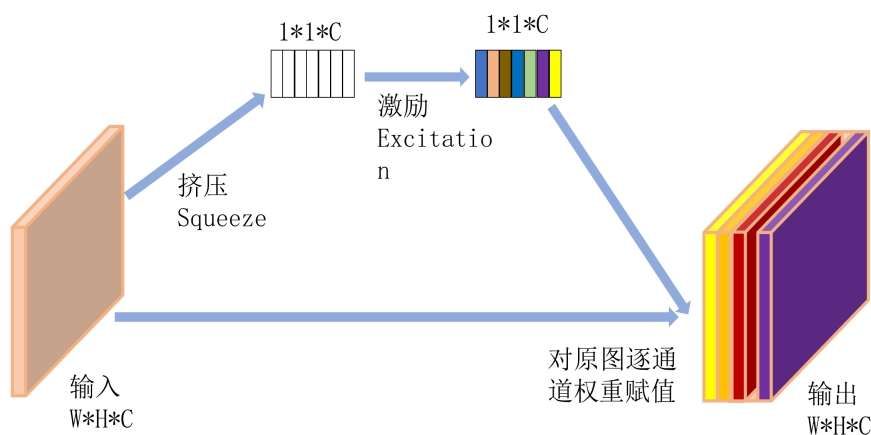


Figure 5. Squeeze-and-excitation module

图 5. SE 模块

2) 批归一化层

在每个卷积层后面添加一层批归一化(Batch Normalization, BN)层, 批归一化层(BN)是计算每个样本的均值和方差, 相当于在每一层输入前先作一个批归一化处理。由于深度网络经过多层次的运算后可能会使数据的分布变化变大, 通过采用批归一化, 将中间层数据统一到同一量纲, 可以减少对学习率的要求, 促进网络模型的训练效果, 缓解模型在训练过程中产生过拟合的问题, 并且能大幅度地提高模型训练的速度。

3) 池化层重构

在原始的 AlexNet 网络中, 是将全连接层设置在最后一层池化层的后面, 但如果网络模型中涉及的神经元多, 就会使全连接层中的参数量所占比重增大, 从而导致模型的训练时间延长。所以本文将在模型结构的尾部使用全局平均池化层(Global Average Pooling, GAP)代替原模型中的两层全连接层。全局平均池化层的主要作用是降维, 大幅压缩参数量并抑制过拟合。它的实现过程如图 6 所示, 主要是对经历

若干卷积和池化层而得到的 N 个特征图的像素分别进行求平均，可定义为：

$$A^k = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n B_{ij}^k \tag{1}$$

其中， B_{ij}^k 为第 k 张鲜茶叶特征图值， A^k 为鲜茶叶特征图对应输出的特征值。

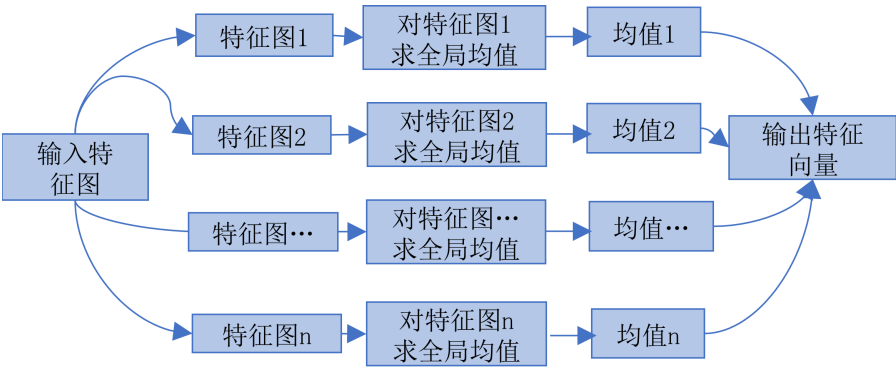


Figure 6. Global average pooling
图 6. 全局平均池化

4. 结果与分析

4.1. 实验设置

硬件环境：操作系统是 Windows 11 家庭版 64 位(Version 21H2/DirectX 12)，处理器是 AMD Ryzen 7 8845HS w/Radeon 780M Graphics，内存 32 GB，显卡为 AMD Radeon 780M 集成显卡。软件环境：采用基于 Tensorflow 为后端运行的 Keras 深度学习框架，汇编语言为 Python 3.8，科学计算库为 Numpy 1.23，数据处理包为 Pandas 1.5，数据可视化包为 Matplotlib 3.6。在实验中，将 batch 的大小设置为 16，训练轮数设置为 50，AlexNet、ResNet50、MobileNetV1、ShuffleNetV1 均选择 SGD 作为优化器，而改进的 AlexNet 模型采用优化器 Adam，学习率设置为 0.001。

4.2. 分类结果

对预处理后的 3200 张鲜茶叶图像进行了试验。本文采用测试集的分类准确率(Accuracy)和损失函数值(Loss)作为评价指标，准确率越高，损失值越小，模型学习的效果越好。

Table 2. Model training result
表 2. 模型训练结果

模型名称	输入尺寸	时间(min)	总迭代数 Epoches	准确率(%)	损失值	模型大小(MB)
AlexNet	224 × 224	65.30	50	94.71	0.16	178.39
AlexNet + BN	224 × 224	48.52	50	96.44	0.13	178.39
AlexNet + GAP	224 × 224	39.44	50	93.40	0.18	14.30
AlexNet + SE	224 × 224	35.40	50	97.53	0.10	178.55
Improved AlexNet	224 × 224	19.03	50	98.84	0.05	3.67

由图 7 准确率图可知，改进的 AlexNet 准确率起步较高，快速上升，后趋于平稳接近 1；AlexNet 前

期波动大, 上升较曲折, 后期趋于平稳但略低于改进模型。在损失值图中, 改进的 AlexNet 下降较平滑, 而 AlexNet 波动大, 并且改进的 AlexNet 模型比 AlexNet 提前 30 轮进入收敛。另外, 从表 2 可知, 改进的 AlexNet 较于 AlexNet 来说, 训练时间快了 46.27 分钟, 准确率提高了 4.13 个百分点, 模型大小减小了 174.71 MB。

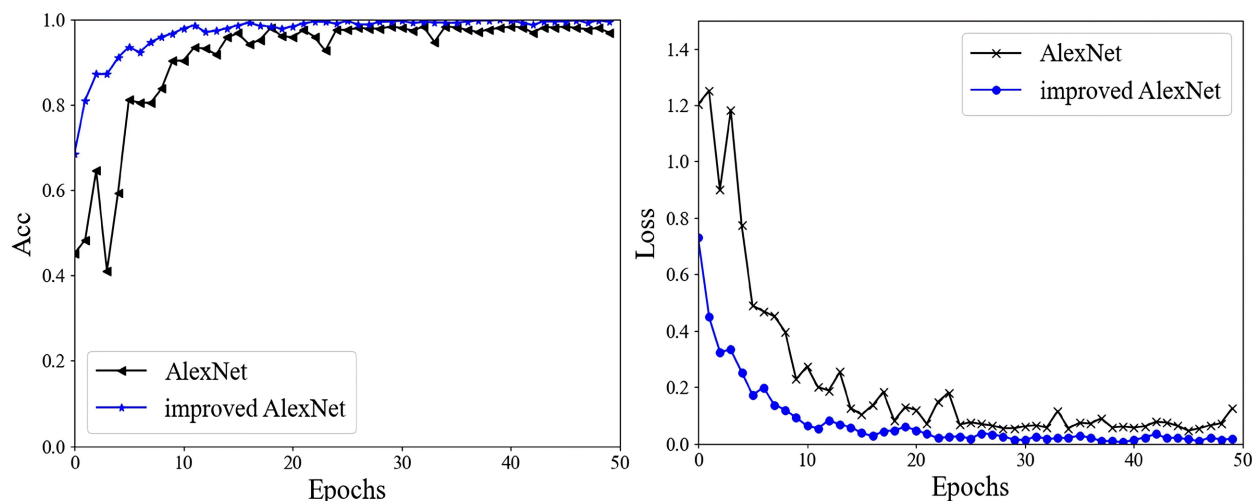


Figure 7. Plot of accuracy and loss values for AlexNet and Improved AlexNet

图 7. AlexNet 和 Improved AlexNet 的准确率和损失值图

4.3. 算法比较

随着迭代次数的增加, 各个网络的准确率值整体上越来越高, 损失值整体上越来越低。如图 8 可知, 改进的 AlexNet 网络准确率都比其他网络收敛得更快, 所有模型在前 20 轮中准确率都在不断提高, 且改进的 AlexNet 网络在准确率和损失值曲线图中收敛得最早并最稳定。其他网络模型在整个损失值曲线图中始终起伏震荡, 不稳定, 即训练集和验证集结果有明显差异, 出现轻微的过拟合现象。另外, 从表 3 可知, 改进的 AlexNet 网络相较其他轻量级网络来说, 模型更小, 损失值更小, 准确率更高。综合分析, 改进的 AlexNet 网络总体性能偏好, 可以满足对鲜茶叶的分类识别。

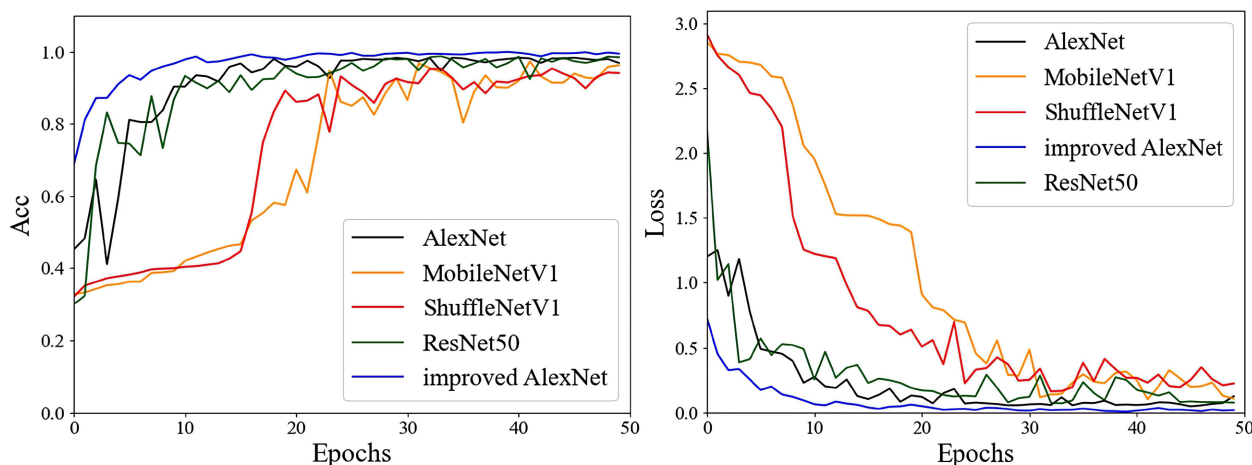


Figure 8. Graphs of accuracy and loss values

图 8. 准确率曲线图和损失值图

Table 3. Accuracy and loss values of the different models on the test set
表 3. 测试集上不同模型的准确率和损失值

模型名称	总迭代数 Epochs	准确率(%)	损失值	模型大小(MB)
AlexNet	50	94.71	0.16	178.39
ResNet50	50	91.50	0.34	90.00
MobileNetV1	50	87.42	0.52	12.33
ShuffleNetV1	50	85.33	0.67	11.26
Improved AlexNet	50	98.84	0.05	3.67

5. 结论

本研究提出改进的 AlexNet 模型来对 3 种鲜茶叶进行分类, 试验中分别从三方面对传统的 AlexNet 网络进行了改进: 一是在 AlexNet 结构的 Conv3~Conv5 层后插入注意力机制 SE 模块; 二是在 AlexNet 网络结构的每层卷积层后面添加一层批归一化层; 三是在模型结构输出层前使用全局平均池化层, 并去掉原始网络中的两个全连接层。通过实验证明, 改进的 AlexNet 模型相比原来的 AlexNet 模型, 训练时间快 46.27 分钟, 准确率提高 4.13 个百分点, 模型大小减小 174.71 MB。说明改进的 AlexNet 模型能够实现鲜茶叶的分类。同时, 也证明深度卷积神经网络对鲜茶叶分类的可行性。

基金项目

2024~2025 年重庆对外经贸学院科研项目: 人工智能图像识别技术助力农产品品质分级与分拣(项目编号: KYZK2024028); 2024~2025 年重庆对外经贸学院《智能算法应用》课程教学改革(项目编号: KG2024067)。

参考文献

- [1] 吴玉生, 李安虎, 万亚明, 等. 基于机器视觉的烟草在线检测技术研究进展[J]. 激光与光电子学进展, 2024, 61(8): 39-52.
- [2] 王润周, 张新生, 王明虎. 基于信号分解和深度学习的农产品价格预测[J]. 农业工程学报, 2022, 38(24): 256-267.
- [3] 蔡杰, 孙希东, 李更新, 等. 专业化分级散叶成件收购模式探索与研究[J]. 现代农业研究, 2020, 26(4): 50-51.
- [4] 袁海波, 滑金杰, 邓余良, 等. 基于 YJY-2 型鲜叶分级机的机采茶叶分级分类工艺优化[J]. 农业工程学报, 2016, 32(6): 276-282.
- [5] 黄艳, 罗玉琴, 张灵枝, 等. 基于近红外光谱和气相离子迁移谱的白茶等级评价研究[J]. 食品工业科技, 2023, 44(21): 348-357.
- [6] 张红梅, 邹光宇, 王淼森, 等. 基于传感器阵列多特征优化融合的茶叶品质检测研究[J]. 传感技术学报, 2018, 31(3): 491-496.
- [7] 万广, 陈忠辉, 方洪波, 等. 基于特征融合的随机森林模型茶鲜叶分类[J]. 华南农业大学学报, 2021, 42(4): 125-132.
- [8] 米福, 陆佳平, 孙旋, 等. 基于 LBP 的特征融合猴魁茶叶叶尖和叶柄识别研究[J]. 包装与食品机械, 2022, 40(4): 14-19.
- [9] 唐宇峰, 杨泽林, 胡光忠, 等. 基于 SE-ResNet 和迁移学习的轴承故障诊断方法研究[J]. 四川轻化工大学学报(自然科学版), 2024, 37(4): 19-27.
- [10] 张金炎, 曹成茂, 李文宝, 等. 基于多特征融合的茶叶鲜叶等级识别的方法研究[J]. 安徽农业大学学报, 2021, 48(3): 480-487.
- [11] 毛腾跃, 黄印, 文晓国, 等. 基于多特征与多分类器的鲜茶叶分类研究[J]. 中国农机化学报, 2020, 41(12): 75-83.
- [12] 马利, 刘新宇, 李皓宇, 等. 应用空洞卷积的神经网络轻量化方法[J]. 计算机工程与应用, 2022, 58(5): 85-93.