

基于改进的ConvNeXt网络的茶叶病虫害分类识别方法研究

房建兵¹, 储玉文²

¹马鞍山学院电气工程学院, 安徽 马鞍山

²皖江工学院计算机与人工智能学院, 安徽 马鞍山

收稿日期: 2026年6月10日; 录用日期: 2026年7月9日; 发布日期: 2026年7月17日

摘要

随着茶叶种植规模的持续扩大, 病虫害对产量与品质的影响日益显著, 已成为制约茶产业发展的重要因素之一。由于不同病虫害类型对应的防治措施存在差异, 实现对其类别的准确判别对于开展针对性防控具有关键意义。传统方法多依赖人工经验或浅层特征, 存在识别效率低、准确性不足等问题。近年来, 卷积神经网络(CNN)凭借其优异的特征学习能力, 在病虫害图像识别领域得到广泛应用。因此, 本文提出了一种改进的ConvNeXt茶叶病虫害分类模型。在原有ConvNeXt Block中引入融合空间与通道信息的Dual注意力机制, 形成DW-Dual结构, 以增强特征表达能力; 在此基础上进一步嵌入SCConv模块, 构建SCConvNeXt-Dual结构, 从而强化特征建模能力并提升训练效率。实验结果表明, 改进模型在收敛速度与分类性能方面均优于基准模型, 收敛所需迭代次数由约140轮降低至约70轮, 分类准确率由93.7%提升至95.4%。综上, 本文实现了提升茶叶病虫害识别精度的同时兼顾训练效率, 为相关领域的智能化识别与精准防控提供了有益参考。

关键词

茶叶病虫害, 卷积神经网络, ConvNeXt网络, 注意力机制

Research on Tea Pest and Disease Classification Based on an Improved ConvNeXt Network

Jianbing Fang¹, Yuwen Chu²

¹School of Electrical Engineering, Ma'anshan University, Ma'anshan Anhui

²School of Computer Science and Artificial Intelligence, Wanjiang University of Technology, Ma'anshan Anhui

Received: June 10, 2026; accepted: July 9, 2026; published: July 17, 2026

文章引用: 房建兵, 储玉文. 基于改进的 ConvNeXt 网络的茶叶病虫害分类识别方法研究[J]. 计算机科学与应用, 2026, 16(7): 82-91. DOI: 10.12677/csa.2026.167243

Abstract

With the continuous expansion of tea cultivation, pest and disease issues have become increasingly prominent, posing significant threats to both yield and quality and thus constraining the sustainable development of the tea industry. Due to the variation in control strategies for different types of pests and diseases, accurate identification of their categories is essential for implementing targeted management measures. Traditional approaches, which rely heavily on manual experience or shallow features, often suffer from low efficiency and limited accuracy. In recent years, convolutional neural networks (CNNs) have been widely applied in pest and disease recognition tasks owing to their powerful feature extraction capabilities. In this study, an improved ConvNeXt-based model for tea pest and disease classification is proposed. Specifically, a Dual attention mechanism integrating spatial and channel attention is introduced into the ConvNeXt Block to construct the DW-Dual module, thereby enhancing feature representation. Furthermore, the SConv module is embedded into the DW-Dual structure to form the SConvNeXt-Dual module, which strengthens feature modeling capability and improves training efficiency. Experimental results demonstrate that the proposed model outperforms the baseline ConvNeXt in both convergence speed and classification performance. The number of training epochs required for convergence is reduced from approximately 140 to 70, while the classification accuracy is improved from 93.7% to 95.4%. In summary, the proposed method achieves higher recognition accuracy while maintaining efficient training, providing a useful reference for intelligent pest and disease identification and precise control in tea cultivation.

Keywords

Tea Pests and Diseases, Convolutional Neural Networks, ConvNeXt Network, Attention Mechanism

Copyright © 2026 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

茶叶在生长过程中除受光照、温度和湿度等自然因素影响外, 还易遭受茶小翅虫、茶毛虫及多种病害侵扰, 导致产量下降与品质降低, 对产业发展造成不利影响。统计数据显示, 2017 年茶叶产量损失率约为 20%; 2018~2019 年虽有所回升, 但病虫害仍是主要减产因素; 2020 年受其影响, 产量降至约 295 万吨。此外, 病虫害影响面积约占茶园总面积的 25%, 造成数百亿元的经济损失[1]。因此, 实现茶叶病虫害的高效、准确识别具有重要意义。

随着人工智能技术的发展, 卷积神经网络[2] (CNN)在农业病虫害识别领域得到广泛应用。相较于传统方法, CNN 能够自动提取图像特征, 实现识别过程的自动化, 在提升效率的同时降低人工成本; 同时, 其通过持续学习可不断优化模型性能, 提高识别的准确性与稳定性。

基于上述分析, 本文提出一种基于改进 ConvNeXt [3]网络的茶叶病虫害分类识别模型。在 ConvNeXt 的深度可分离卷积(DW)模块前引入融合空间注意力与通道注意力的 Dual [4]注意力机制, 以增强特征表达能力; 同时在该模块中嵌入 SConv [5]结构, 通过 SRU 和 CRU 分别建模空间与通道相关性, 构建 SConvNeXt-Dual 模块。实验结果表明, 该模型在保证较快收敛速度的同时, 有效提升了茶叶病虫害的识别准确率。

2. 研究现状

近年来, 随着人工智能、深度学习及图像处理技术的发展, 茶叶病虫害分类识别研究取得了显著进展。国内外学者基于多种方法不断提升识别精度, 为相关应用提供了技术支撑。

2015 年, Chandra Karmokar [6]等提出 TLDR 系统, 结合图像预处理与神经网络集成模型, 实现约 90% 的识别准确率。2018 年, 龙满生[7]等基于卷积神经网络与迁移学习, 将 AlexNet 应用于油茶病害识别, 准确率达 96.53%; 同年, Md. Selim Hossain [8]等采用支持向量机方法, 识别准确率超过 90%。2020 年, 王年等[9]融合 HOG 特征与 Inception V3 模型, 并结合梯度提升树进行分类, 准确率提升至 95%。2022 年, 李子茂等[10]在 DenseNet 中引入 SENet 模块, 并采用迁移学习策略, 在小样本条件下实现 92.66% 的识别准确率。2023 年, 黄铝文等[11]提出 CBAM-TealeafNet 模型, 结合小波变换与 Bneck 结构, 准确率达 98.70%; 吴叶辉[12]等基于 EfficientNet 的分层学习方法, 将准确率提升至 99.21%。

总体来看, 现有研究虽在识别精度上取得较好效果, 但多基于自建数据集, 存在一定的偶然性与区域局限, 难以充分反映模型在实际场景中的泛化能力。同时, 多数方法依赖参数规模较大、计算复杂的模型, 在实际部署中仍面临效率与资源约束问题。

3. 数据与方法

3.1. 实验数据集

本文采用来源于博梅特 Koiwa 地区 Johnstone Boiyon 农场的茶叶病虫害自动分类数据集。该数据集包含 8 类病虫害及健康茶叶图像, 共 885 张, 每类样本数量均在 100 张以上, 且均由专业人员完成标注。实验中按 8:2 比例划分训练集与测试集, 测试集数据不参与模型训练。图 1 是 8 类茶叶病虫害类型: 斑藻类叶斑病、炭疽病、鸟眼斑、褐色枯萎病、灰枯病、正常、红叶斑病、白点及健康茶叶。

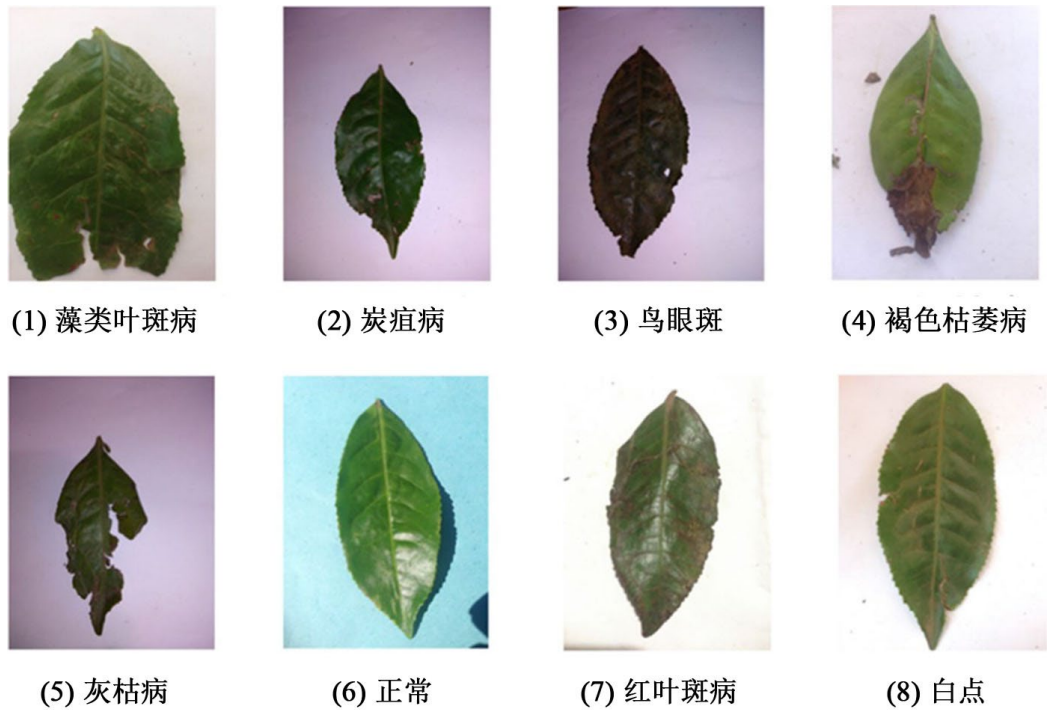


Figure 1. Types of tea diseases and pests
图 1. 茶叶病虫害类型

各类别的样本分布如表 1 所示:

Table 1. System resulting data of standard experiment

表 1. 标准试验系统结果数据

标识	类型	数量(张)	占比(%)
1	斑藻类叶斑	113	12.77
2	炭疽病	100	11.30
3	鸟眼斑	100	11.30
4	褐色枯萎病	113	12.77
5	灰枯病	100	11.30
6	正常	74	8.36
7	红叶	143	16.16
8	白点	142	16.05

3.2. 数据预处理

为提升模型的鲁棒性及对复杂场景的适应能力, 本文对数据集进行预处理与数据增强, 包括尺寸调整、随机旋转、随机裁剪与缩放、随机水平翻转及色彩抖动等操作。所有图像统一调整为 224×224 像素, 并进行归一化处理。图 2 展示了各类别样本及增强效果示例。

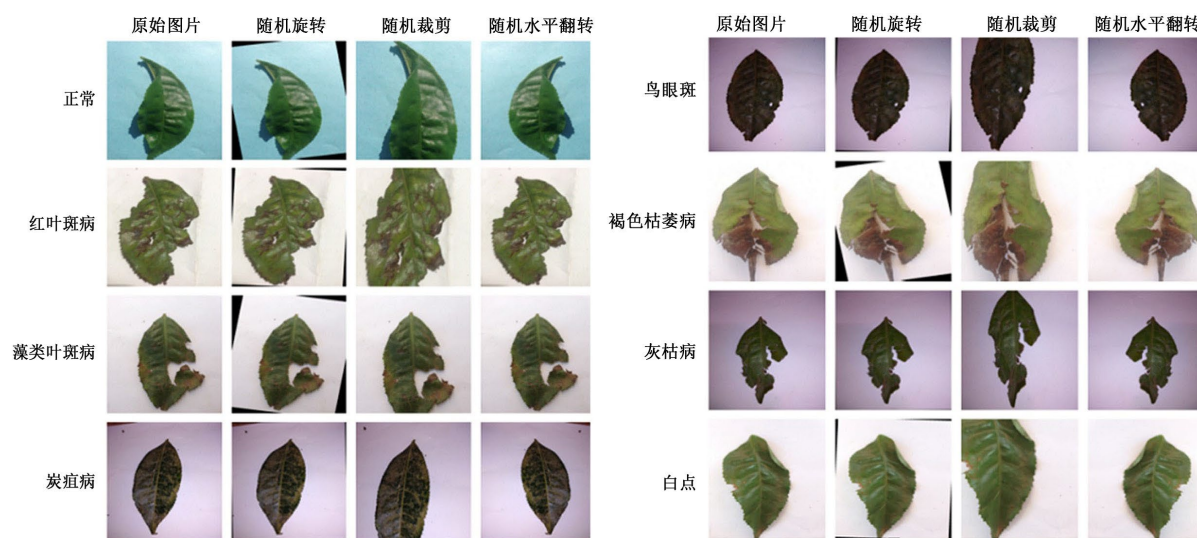


Figure 2. Sample images of each category and data augmentation examples

图 2. 各个类别的样本以及数据增强效果示例

针对原始数据集中类别分布不均的问题, 本文采用加权重采样策略进行处理。首先统计各类别样本数量, 并按其反比例分配样本权重; 随后利用 “Weighted Random Sampler” 采样器在训练过程中进行重采样, 使各类样本被等概率抽取, 从而减弱数据不平衡带来的模型偏置。经过数据预处理与平衡处理后, 构建了一个更加均衡、多样的训练数据集。图 3 和图 4 分别展示了处理前后的数据分布情况。

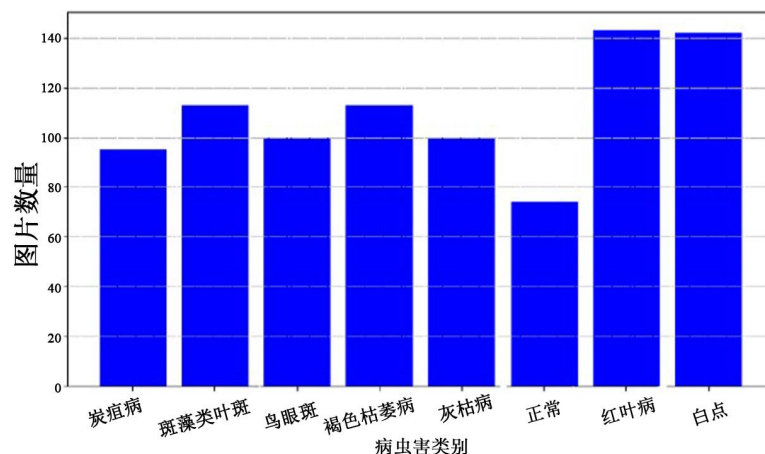


Figure 3. Class distribution in the original dataset

图 3. 原始数据集中各类别的分布情况

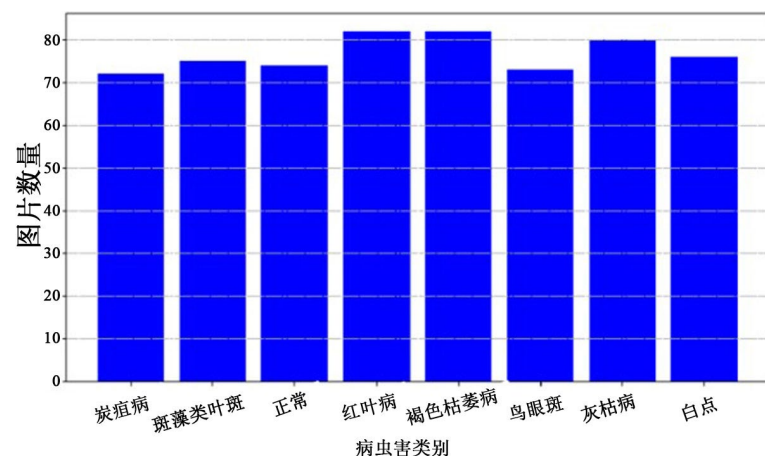


Figure 4. Class distribution in the training set after data augmentation and class balancing

图 4. 数据增强和数据平衡处理后训练集每个类别的数据分布情况

3.3. 改进的 ConvNeXt 网络模型

本文选取 ConvNeXt 作为基础网络结构。该模型在设计上融合了 ResNet 的层级结构思想与 Vision Transformer (ViT) 的部分特征表达方式, 通过对卷积形式、归一化方法及激活函数的改进, 在保持较高计算效率的同时, 进一步提升了模型的特征提取能力与泛化性能。

在实际实验过程中发现, 针对茶叶病虫害图像中局部差异不明显的情况, 原始 ConvNeXt 在关键特征的捕捉与表达方面仍存在一定不足, 易导致部分细粒度特征被忽略。为增强模型对有效特征的关注能力, 本文引入注意力机制对网络结构进行优化。通过对不同注意力机制的对比分析, 选用 Dual 注意力机制, 并将其嵌入至 ConvNeXt Block 中的深度可分离卷积(DW)之前, 从而构建 DW-Dual 模块, 以提升特征表达的针对性。

此外, 考虑到 ConvNeXt 网络参数规模较大、计算开销较高, 易影响模型收敛速度, 本文进一步引入 SCConv 模块, 以加强对空间信息与通道信息的联合建模能力。在此基础上, 将 SCConv 嵌入 DW-Dual 模块中, 形成 SCConvNeXt-Dual 结构。相关消融实验结果表明, 该改进策略在提升模型收敛性能的同时, 有效提高了分类识别精度。改进后的整体网络结构如图 5 所示。

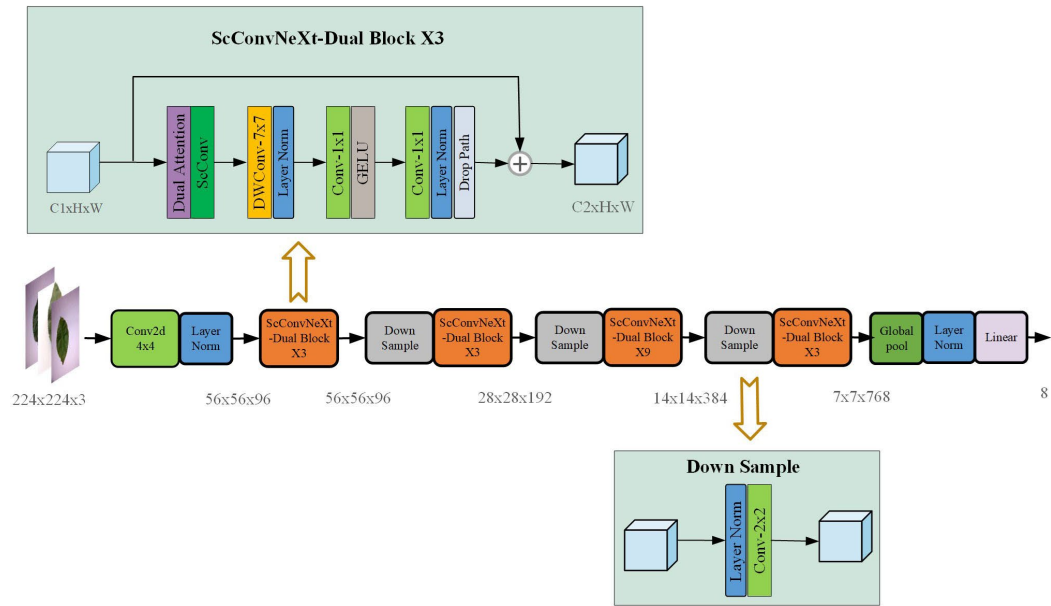


Figure 5. Improved ConvNeXt network
图 5. 改进的 ConvNeXt 网络

4. 实验结果及分析

4.1. 实验环境及设置

本文实验在 Linux 平台下进行, 运行环境为配备 NVIDIA Quadro RTX A6000 GPU 的服务器。基于 Anaconda3 环境, 采用 PyTorch 1.13.0 框架与 Python 3.8 进行模型训练与测试。训练过程中使用 Adam 优化器, 初始学习率设为 5×10^{-4} , 具体配置见表 2。

Table 2. Experimental environment
表 2. 实验配置环境

名称	参数/版本
CPU	Intel(R) Xeon(R) Silver 4214
GPU	NVIDIA Quadro RTX A6000
内存	256 GB
GPU 显存	48 GB/张
CUDA	11.8
Pytorch	1.13.0
Epoch	300
学习率	$5e-4$

4.2. 评价指标

本章旨在实现对多类别茶叶病虫害的有效识别。为此, 将数据集按比例划分为训练集与测试集, 其中 80% 用于模型训练, 其余 20% 用于性能评估。在数据划分过程中, 对样本进行严格隔离, 确保训练阶段所使用的图像数据不会出现在测试集中, 从而保证评估结果的客观性与可靠性。后续各项实验结论均

基于与训练过程完全独立的测试集获得。

在模型性能评价方面, 选取准确率(Accuracy, Acc)作为主要衡量指标, 用以反映模型对整体样本的判别能力。其计算公式如下所示:

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP}$$

其中, TP (True Positive)表示被正确判定为正类的样本数, FP (False Positive)表示被错误判定为正类的样本数, FN (False Negative)表示被错误判定为负类的样本数, TN (True Negative)表示被正确判定为负类的样本数。该指标能够从整体上反映模型分类结果的正确程度。

此外, 为全面评价模型的训练性能, 本文进一步对模型收敛性能进行分析。本文将验证集准确率(Acc)趋于稳定、验证集损失(Loss)基本不再下降, 且连续多个 Epoch 内变化幅度较小时视为模型达到收敛状态, 并以达到该状态所需的训练轮数作为衡量模型收敛速度的依据。

4.3. 实验结果分析

在上述实验条件下, 对比分析改进前后 ConvNeXt 模型在测试集上的分类结果, 其混淆矩阵如图 6 所示。其中, 图 6(a)对应原始 ConvNeXt 模型, 图 6(b)为改进模型的结果。由图可知, 在斑藻类叶斑病类别中, 两种模型均实现了全部 22 张样本的正确识别。在炭疽病类别上, 原模型存在 3 例误判, 分别被错误归类为褐色枯萎病(1 张)和灰枯病(2 张); 相比之下, 改进模型的误分类数量减少至 2 例, 分别被误判为斑藻类叶斑病和灰枯病。在正常茶叶与红叶斑病类别中, 两种模型均未出现误分类, 识别结果保持一致且准确率达到 100%。

综合来看, 改进后的 ConvNeXt 模型在茶叶病虫害分类任务中表现出更优性能, 整体准确率达到 95.4%, 较原模型提升 1.7 个百分点。结果表明, 所提出的改进策略在降低误分类率方面具有一定效果, 尤其在部分易混淆类别的判别能力上得到提升。

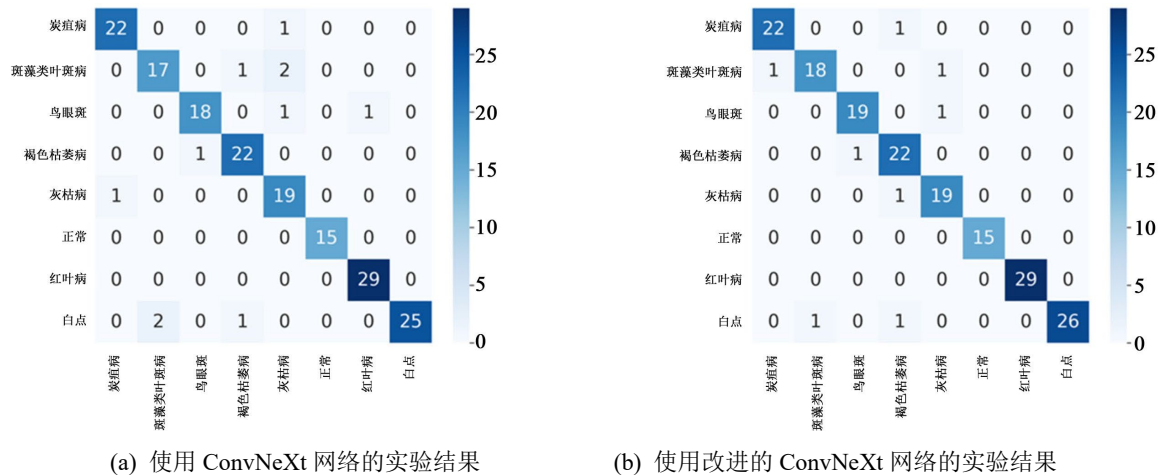


Figure 6. Confusion matrices before and after model improvement
图 6. 模型改进前后输出结果的混淆矩阵

图 7 与图 8 分别给出了原始 ConvNeXt 模型与改进模型在训练过程中的 Acc 及 Loss 变化曲线。两种模型均设置为 300 个 Epoch 进行训练。由曲线变化趋势可以看出, 原始 ConvNeXt 模型在约第 140 个 Epoch 时趋于稳定, 收敛过程相对较慢; 而改进模型在约第 70 个 Epoch 时已基本完成收敛。对比结果表

明, 引入改进结构后, 模型在训练早期即可达到稳定状态, 收敛速度明显提升, 有助于降低训练时间成本并提高训练效率。

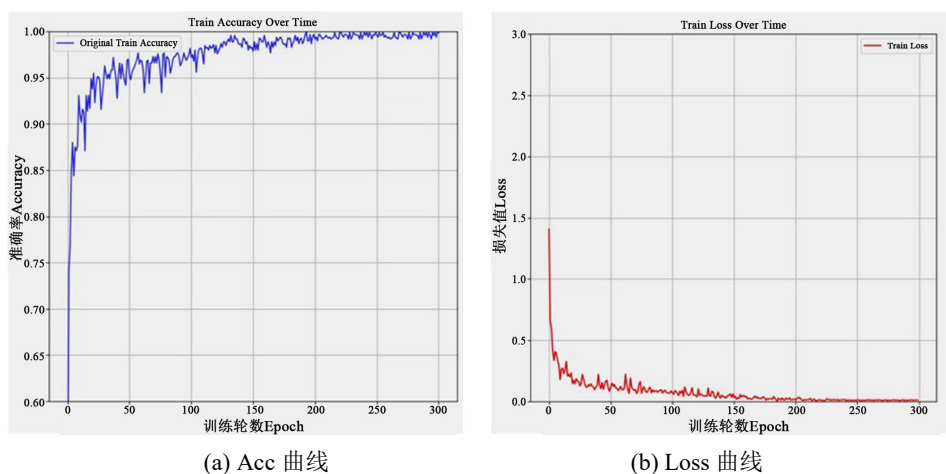


Figure 7. Acc and Loss curves of the original ConvNeXt network
图 7. 采用原 ConvNeXt 网络进行分类的 Acc 与 Loss 曲线

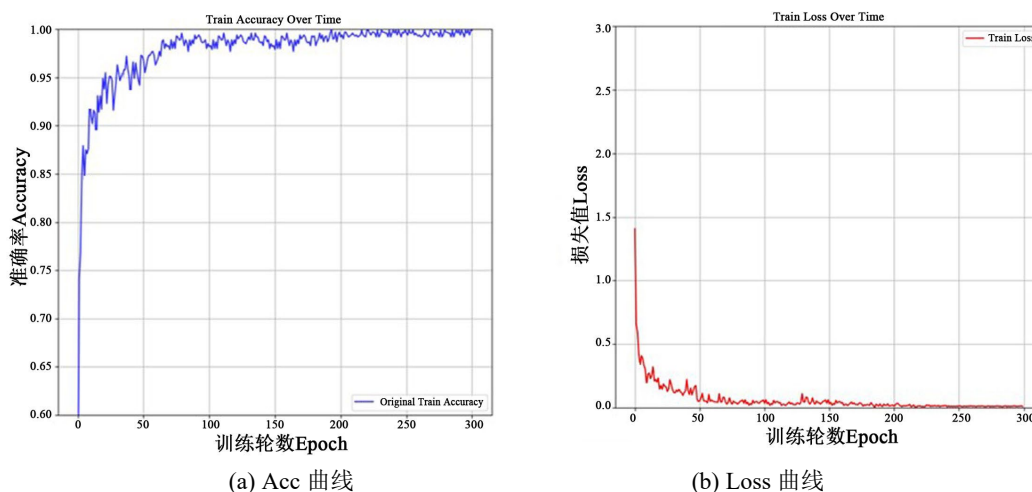


Figure 8. Acc and Loss curves of the improved ConvNeXt network
图 8. 采用改进的 ConvNeXt 网络进行分类的 Acc 与 Loss 曲线

为验证所提出改进结构的合理性及其在茶叶病虫害分类任务中的有效性, 本文围绕 ConvNeXt 模型开展了系统实验分析。首先, 对比分析 ConvNeXt 融合多种注意力机制时的分类效果差异, 其次, 通过消融实验考察 Dual 注意力机制与 SCConv 模块在不同组合方式下对模型性能的影响, 相关实验结果以表格形式呈现, 并以识别准确率作为主要评价指标进行分析。

(1) 对比实验

表 3 对不同注意力机制在本研究任务中的应用效果进行了比较。其中, SE 模块[13]通过压缩空间维度获取全局上下文信息, 并对通道进行重标定; ECA [14]在此基础上引入一维卷积以实现通道权重的自适应分配, 降低了计算复杂度; CBAM 从通道与空间两个维度对特征进行联合建模; CA [15]则通过刻画通道间依赖关系以强化关键信息表达。

Table 3. Comparison of experimental results for different attention mechanisms
表 3. 不同注意机制的实验结果比较

方法	准确率(%)
SE Block	92.0
ECA Block	93.1
CBAM Block	93.7
CA Block	94.3
Dual Block	94.8

实验结果表明, 在 ConvNeXt 中引入 SE 模块后, 模型准确率降至 92.0%, 低于基准模型的 93.7%; ECA 与 CBAM 分别达到 93.1%和 93.7%, 整体未体现出性能提升。相比之下, 采用 CA 模块后, 准确率提升至 94.3%, 表现出一定的改进效果。在此基础上进一步引入 Dual 注意力机制, 模型准确率达到 94.8%, 较基准模型提升 1.1%, 同时优于 CA 方案。

综合分析可知, Dual 注意力机制在本研究数据集及网络结构下具有更好的适配性, 能够从多维角度增强特征表征能力, 并在一定程度上提升模型的泛化性能与稳定性。

(2) 消融实验

在消融实验中, 表 4 给出了不同模块配置对模型性能及计算复杂度的影响。其中, “+Dual”表示在原 ConvNeXt 网络中引入 Dual 注意力机制, “+SCConv”表示引入 SCConv 模块。由实验结果可知, 加入 Dual 注意力后, 模型分类准确率由 93.7%提升至 94.8%; 引入 SCConv 模块后, 准确率提升至 94.3%, 均优于基准模型。在此基础上, 同时融合 Dual 注意力机制与 SCConv 模块后, 模型最终准确率达到 95.4%, 较基准模型提高了 1.7 个百分点。

随着两种改进模块的引入, 模型参数量由 87.6 M 增加至 90.6 M, 浮点运算次数(FLOPs)由 15.42 G 增加至 16.09 G, 推理时间由 8.21 ms/image 增加至 8.86 ms/image, 总占用内存由 598.03 M 增加至 638.59 M, 表明模型计算复杂度有所提升。然而, 相较于计算开销的小幅增加, 模型分类精度获得了更为明显的提升, 说明所提出的改进方法能够在保证推理效率的前提下有效增强模型特征表达能力, 在识别性能与计算成本之间取得了较好的平衡。

Table 4. Ablation study results
表 4. 消融实验结果

方法	准确率(%)	总参数(M)	FLOPs (G)	总占用内存(M)	Times (ms/image)
ConvNeXt	93.7	87.6	15.42	598.03	8.21
+Dual	94.8	89.3	15.81	611.86	8.47
+SCConv	94.3	88.9	15.67	624.76	8.63
+Dual+SCConv	95.4	90.6	16.09	638.59	8.86

5. 结论

针对传统茶叶病虫害监测方法存在效率低、漏检率高等问题, 本文提出了基于改进的 ConvNeXt 网络的茶叶病虫害分类识别模型。通过在 DW 卷积前引入 Dual 注意力机制, 构建 DW-Dual 模块, 有效增强了模型对空间与通道关键信息的表达能力; 进一步结合 SCConv 模块, 形成 SCConvNeXt-Dual 结构,

在强化特征建模能力的同时提升了模型的收敛效率。实验结果表明, 改进模型在分类性能上优于原 ConvNeXt 网络, 整体识别准确率达到 95.4%, 较原模型提升 1.7 个百分点。同时, 训练过程收敛更快, 模型稳定性得到改善。消融实验与对比实验结果进一步验证了 Dual 注意力机制与 SCConv 模块在提升模型性能方面的有效性。综上, 本文所提出的改进方法在保证计算开销可控的前提下, 实现了识别精度与训练效率的协同提升, 可为茶叶病虫害的智能化监测与精准防控提供一定的技术参考。

参考文献

- [1] 周曙东, 吴新林. 中国茶叶生产布局变迁特征及影响分析[J]. 茶叶通讯, 2020, 47(3): 496-501.
- [2] Wang, P., Li, W., Gao, Z., Zhang, J., Tang, C. and Ogunbona, P.O. (2016) Action Recognition from Depth Maps Using Deep Convolutional Neural Networks. *IEEE Transactions on Human-Machine Systems*, **46**, 498-509. <https://doi.org/10.1109/thms.2015.2504550>
- [3] Fu, Y., et al. (2022) OTFPF: Optimal Transport-Based Feature Pyramid Fusion Network for Brain Age Estimation with 3D Overlapped ConvNeXt. arXiv: 2205.04684.
- [4] Fu, J., Liu, J., Tian, H., Li, Y., Bao, Y., Fang, Z., et al. (2019) Dual Attention Network for Scene Segmentation. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 3141-3149. <https://doi.org/10.1109/cvpr.2019.00326>
- [5] Li, J.F., Wen, Y. and He, L. (2023) SCConv: Spatial and Channel Reconstruction Convolution for Feature Redundancy. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, 17-24 June 2023, 6153-6162. <https://doi.org/10.1109/CVPR52729.2023.00596>
- [6] ChandraKarmokar, B., Samawat Ullah, M., Kibria Siddiquee, M. and Md. Rokibul Alam, K. (2015) Tea Leaf Diseases Recognition Using Neural Network Ensemble. *International Journal of Computer Applications*, **114**, 27-30. <https://doi.org/10.5120/20071-1993>
- [7] 龙满生, 欧阳春娟, 刘欢, 等. 基于卷积神经网络与迁移学习的油茶病害图像识别[J]. 农业工程学报, 2018, 34(18): 194-201.
- [8] Hossain, S., Mou, R.M., Hasan, M.M., Chakraborty, S. and Razzak, M.A. (2018) Recognition and Detection of Tea Leaf's Diseases Using Support Vector Machine. 2018 *IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA)*, Penang, 9-10 March 2018, 150-154. <https://doi.org/10.1109/cspa.2018.8368703>
- [9] 王年, 孟树林, 吴洛天, 等. 基于改进关系网络的小样本学习[J]. 安徽大学学报(自然科学版), 2020, 44(4): 38-44.
- [10] 李子茂, 徐杰, 郑禄, 等. 基于改进 DenseNet 的茶叶病害小样本识别方法[J]. 农业工程学报, 2022, 38(10): 182-190.
- [11] 黄铝文, 关非凡, 谦博, 等. 基于 2D DWT 与 MobileNetV3 融合的轻量级茶叶病害识别[J]. 农业工程学报, 2023, 39(24): 207-214.
- [12] 吴叶辉, 李汝嘉, 季荣彪, 等. 基于随机增强 Swin-Tiny Transformer 的玉米病害识别及应用[J]. 吉林大学学报(理学版), 2024, 62(2): 381-390.
- [13] Hu, J., Shen, L. and Sun, G. (2018) Squeeze-and-Excitation Networks. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 7132-7141. <https://doi.org/10.1109/cvpr.2018.00745>
- [14] Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W. and Hu, Q. (2020) ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 11531-11539. <https://doi.org/10.1109/cvpr42600.2020.01155>
- [15] Hou, Q., Zhou, D. and Feng, J. (2021) Coordinate Attention for Efficient Mobile Network Design. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 13708-13717. <https://doi.org/10.1109/cvpr46437.2021.01350>