

基于 ConvN-Classify 的柑橘病害智能诊断系统

林汉源¹ 范子安¹ 陈钦炯¹ 方泽佳¹ 徐竟成¹

LIN Hanyuan FAN Zian CHEN Qinjiong FANG Zejia XU Jingcheng

摘要

柑橘病害的准确分类对于柑橘的科学种植与有效防治至关重要。文章构建了一个柑橘病害智能诊断系统,提出一种可用于嵌入式设备应用的高精度快速柑橘病害诊断模型 ConvN-Classify。采用改进 ConvNeXtV2 轻量化主干结构作为基础架构,结合改进 SPPF 模块提取多尺度特征,最后使用 YOLOv8 分类检测头 Classify 优化模型性能。实验结果表明,模型在柑橘病害分类任务上表现出色,能够准确区分黑斑病、溃疡病和黄龙病等多种病害,分类准确率达到 99.02%,模型计算复杂度和参数量仅为 1.0×10^9 和 0.7×10^6 ,优于其他模型。在实际部署模型后,能够准确识别病害,为柑橘病害的精准诊断和嵌入式设备部署提供了可靠的技术支撑。

关键词

ConvN-Classify; 病害诊断; 嵌入式; ConvNeXtV2; SPPF; Classify

doi: 10.3969/j.issn.1672-9528.2025.07.006

0 引言

2020 年我国柑橘种植总面积 283.15 万 hm^2 ,总产量达 5 121.9 万 $\text{t}^{[1]}$ 。作为全球广泛种植的经济作物,其产业发展对经济和民生具有重要意义。在果园环境中,柑橘容易遭受黑斑病、溃疡病和黄龙病等病害的侵袭,这些病害如果不能正确诊断,将严重影响果实的品质,造成重大经济损失。

在农作物病害诊断领域,早期研究多基于传统图像处理与机器学习技术,通过人工提取特征,再利用分类器进行分类,这些方法模型的鲁棒性与准确度都存在局限性^[2]。随着计算机技术快速发展,基于深度学习的水果病害图像诊断展现出巨大的潜力。王俏等人^[3]通过自注意力精炼机制和行锚分类方法开发了快速葡萄叶片病害检测算法,减少现有方法的时间消耗和人为误差。张帅杰等人^[4]引入 SA 和 ECA 机制后构建 DenseNet201-SA-ECA 模型对甘薯黑斑病害进行分类,分类准确度达 96.8%,分类效果较好。施宇等人^[5]提出 E-Res2Net 模型在柑橘植物叶片病害 8 分类真实图像上的识别准确率达到 86.04%。

上述方法虽然取得了一定的成果,但是对于柑橘病害诊断的相关研究仍然较少,且研究的模型计算程序复杂、内存占用大,难以进行嵌入式设备的部署。因此开发一个高效、可用于嵌入式设备的柑橘病害诊断系统,对于柑橘病害的及时防治和产业的可持续发展具有重要的现实意义。为此,本文设计了一种高效、轻量化的柑橘病害诊断模型 ConvN-

Classify,将其进行部署应用,开发了一种高效的柑橘病害智能诊断系统,旨在实现柑橘病害的快速、准确分类,为柑橘病害防治提供科学依据。

1 ConvN-Classify 模型

ConvN-Classify 模型结构由 ConvNeXtV2、SPPF (空间金字塔池化) 和 YOLOv8 分类检测头 Classify 组成。其具体结构如图 1 所示。

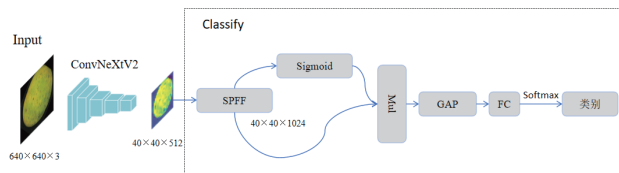


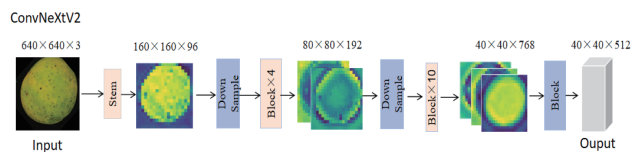
图 1 ConvN-Classify 模型结构图

输入尺寸为 $640\times 640\times 3$ 的柑橘病害图像在经过 ConvNeXtV2 主干结构提取特征后,进入 SPPF 进一步提取多尺度特征,增强模型对不同大小病害的适应性。同时将提取的特征以固定维度的特征表示,得到 $40\times 40\times 1\,024$ 的特征图,方便后续的分类。最后由 YOLOv8 分类检测头 Classify 完成分类任务,提升柑橘病害分类的准确性和可靠性。

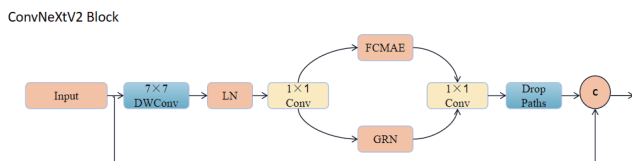
1.1 ConvNeXtV2 概述

ConvNeXtV2^[6] 结合了自监督学习技术和架构创新的一种新型轻量化的卷积神经网络架构,在图像分类的性能上表现优越。其框架基于非对称的编码器-解码器设计,图 2 展示的 ConvNeXtV2 的具体结构。

1. 华南农业大学电子工程学院 (人工智能学院)
广东广州 510642



(a) ConvNeXtV2 主干结构



(b) Block 结构图

图 2 ConvNeXtV2 结构图

ConvNeXtV2 引入了全卷积掩码自编码器框架 (FCMAE) 和全局响应归一化 (GRN) 层两大核心创新。FCMAE 通过随机掩码输入图像的部分区域, 并让模型预测被掩码的部分从而学习更鲁棒的特征表示。GRN 层是 ConvNeXtV2 的另一大创新, 其通过全局特征聚合、特征归一化和特征校准, 最终通过 L2 范数计算归一化分数并校准输入响应, 最终解决特征冗余问题, 节省了计算复杂度和参数量。与此同时, 对主干进行缩放处理, 减少了计算复杂度。

1.2 SPPF 概述

SPPF^[7] 是一种快速空间金字塔池化, 其核心思想是将不同尺寸的输入特征图, 通过构建不同尺度的池化窗口, 对特征图进行池化操作, 提取不同尺度的特征信息, 转化为固定长度的特征向量输出, 从而解决了传统卷积神经网络对输入图像尺寸固定的要求。

本文在 SPPF 的基础上独立于不同尺度的最大池化过程, 引入一个 5×5 的平均池化层来减少噪声的敏感性, 抑制局部极端值的影响, 提升模型在复杂背景下的检测能力和鲁棒性。具体改进结构如图 3 所示。

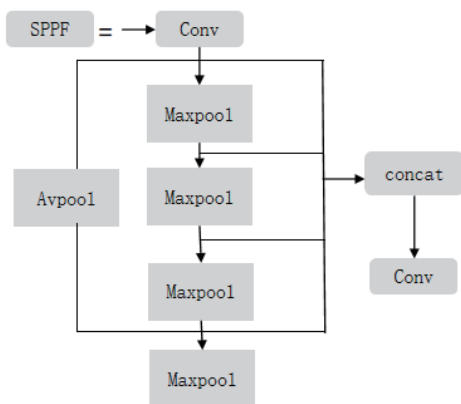


图 3 改进 SPPF 结构图

将改进 SPPF 模块安置于特征提取网络之后, 可以有效提取多尺度特征, 丰富模型细节, 起到承上启下的作用, 为后续分类提供较为准确的特征提取图像。

1.3 Classify 概述

Classify^[8] 是用于 YOLO 系列多类别分类任务的核心组件。负责对输入图像提取的特征进行类别概率预测, 输出每个类别的置信度得分。其由全局池化层 + 全连接层构成, 设计轻量化以平衡精度与速度。

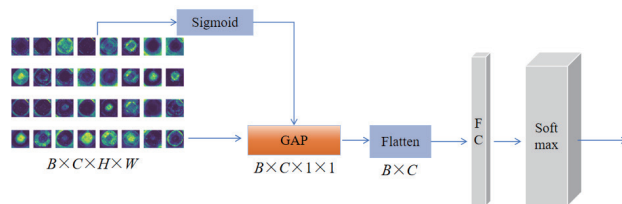


图 4 Classify 结构图

图 4 为 Classify 结构图, 将特征图的空间维度 ($H \times W$) 压缩为 1×1 , 保留通道维度 (C), 减少参数量并防止过拟合。然后将特征图进行展平, 为全连接层准备一维特征向量。最后经过全连接层 FC 将通道维数映射到类别数, 并用 Softmax 激活函数进行分类。这种分类头结构相对简洁, 通过全局池化和全连接层的组合, 能够快速地对特征进行处理和分类, 减少了计算量和参数数量, 提高了模型的运行效率, 使其在实时性要求较高的应用场景中具有优势。

2 数据采集与处理

2.1 数据采集

本文数据集是来自 Kaggle 上公布的柑橘病害数据集 Orange diseases dataset。包含有一类健康柑橘与三类常见的病害柑橘共 1 090 张图像。为增强数据集的泛化性, 另外从现实中用一加 Ace2 型号手机采集了 80 张实际的图像, 添加进数据集中共 1 170 张, 部分样本如图 5 所示。

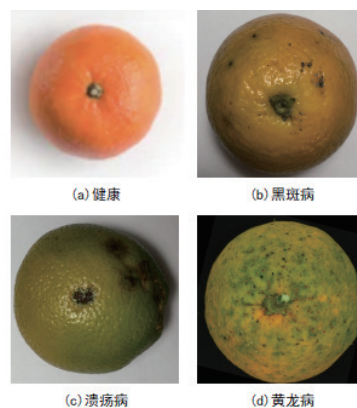


图 5 各类柑橘病害图片

2.2 数据预处理

数据根据类别进行划分, 将之转化为标准的文件夹目录, 并按照类别为文件夹命名。按照 8:2 的比例划分训练集和验证集, 共得到 936 张训练集图像和 234 张验证集图像。表 1

展示的是各个柑橘类别对应的标签。

表 1 各类别柑橘的标签

序号	类别	标签名称
1	健康	Normal
2	黑斑病	Blackspot
3	溃疡病	Canker
4	黄龙病	Grenning

3 实验结果与分析

3.1 实验环境配置

本文基于 PyTorch 深度学习框架来进行模型的训练和验证。采用 Sto - 随机梯度下降（SGD）优化器进行网络训练。为确保实验的可靠性，本文所有研究均处于同一环境下。具体实验环境如表 2 所示。

表 2 实验基本配置

Parameters	Experimental Configuration
CUDA	11.8
GPU	NVIDIA GeForce RTX 2060 SUPER
Operating system	Window 10 64 bit
Deep learning framework	PyTorch2.3.1
Programming software	Python3.9.20
CPU	12 th Gen Intel(R) Core(TM) CPU i5-12490F 3.00 GHz

3.2 实验结果与分析

3.2.1 混淆矩阵分析

混淆矩阵^[9]是一种常用的评估分类模型性能的工具，以矩阵的形式展示了模型在每个类别上的预测结果与真实标签之间的关系。本次实验选用混淆矩阵作为评估指标之一，设置输入图像尺寸为 640 px×640 px，模型迭代次数为 200，批量大小为 8，初始学习率为 0.01 进行模型的训练。得到混淆矩阵如图 6 所示。

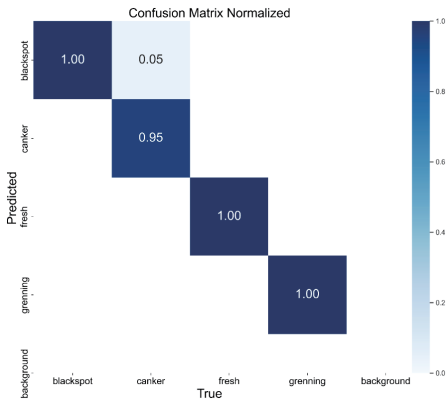


图 6 ConvN-Classify 的标准化混淆矩阵

图中对角线代表的是模型分类的精度，可以看出健康、黑斑病、溃疡病的检测精度均在 100%，黄龙病的检测精度在 95% 以上，说明模型对柑橘病害的诊断效果较好。值得注意的是，ConvN-Classify 模型在处理复杂病害特征时具备优势，在黄龙病和黑斑病的分类任务上，模型能够准确把握不同病害之间的细微差异，准确区分相似类别病害。

3.2.2 模型对比

为验证改进 ConvN-Classify 算法的有效性，基于同一数据集和同一实验环境进行主流模型的性能对比，包括 VGG16^[10]、Resnet50^[11]、YOLOv8^[12]、YOLOv11^[13]，采用分类精度、浮点运算次数（GFLOPs）、参数量（Parameters）以及内存占比作为模型评价指标。实验结果如表 3 所示。

表 3 模型精度对比

模型	分类精度 /%	FLOPs /×10 ⁹	Parameters /×10 ⁶	内存 /MB
VGG16	97.97	15.37	41.45	158.00
Resnet50	97.98	4.13	24.03	91.90
YOLOv8	98.96	3.30	1.40	2.84
YOLOv11	94.94	3.20	1.50	3.05
ConvN-Classify	99.02	1.00	0.70	1.41

由表 3 可以看出提出的 ConvN-Classify 模型在柑橘病害诊断上的表现优于其他经典的分类模型，在主干网络缩放比例为 0.25 的情况下，分类准确率达到 99.02%，较 YOLOv11 提升 4.08%，高于其他比对模型。这表明在柑橘病害分类任务中，ConvN-Classify 模型对病害类别的判断准确性更高，能更精准地区分不同的柑橘病害类型。

值得注意的是，提出的 ConvN-Classify 模型的 FLOPs 和参数量仅为 1.0×10⁹ 和 0.7×10⁶，在保证识别精度优于 VGG16 的情况下，计算复杂度和参数量显著下降。同时内存占比仅为 1.41 MB，能够在大部分的嵌入式设备上投入使用，为轻量化模型部署提供了可靠的技术支持。

4 模型部署

为验证提出的 ConvN-Classify 模型的实际应用效能，本文进行移动设备的模型部署。选用 Android Studio 2024 版本作为应用开发工具，设备为一加 ACE2 型号智能手机，安卓版本为 14.0。在部署之前进行模型优化操作，将 ConvN-Classify 模型从浮点型转换为整型，然后转换自己训练的 pt 权重为 onnx 格式，输入模型进行推理，最后将诊断结果展示在界面上。实际诊断效果如图 7 所示。可以看出模型在设备上 FPS 在 100 帧 /s 左右，速度较快，同时在陌生数据集上能够正确分辨健康柑橘和病害柑橘。根据测试结果，

后续还将进一步优化模型和应用代码,确保满足实际使用场景的性能要求和精度要求。

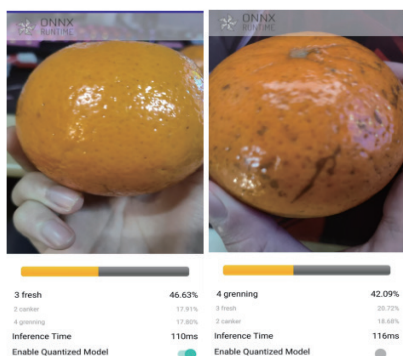


图7 诊断效果图

5 结论

本文基于 ConvN-Classify 开发了一套可用于嵌入式设备部署的高精度快速柑橘病害诊断模型,在 Kaggle 上选取包含柑橘病害的公开数据集 Orange diseases dataset 和自采数据集进行研究,实现了对柑橘病害的精准快速诊断。在主干结构上,使用改进 ConvNeXtV2 主干网络,使用 0.25 的缩放系数,减少特征冗余,保持高效率运算。在主干提取特征之后,引入平均池化层改进 SPPF 模块,进一步提取多尺度特征。最后使用 Classify 分类头进行分类,简化模型,提高了模型的分类效率,使其在实时性要求较高的应用场景中具有优势。结果表明,提出的 ConvN-Classify 模型在橘子病害数据集上的分类准确度达到 99.02%,较 YOLOv11 提升 4.08%,高于其他比对模型。参数量和计算复杂度分别下降至 0.7×10^6 和 1.0×10^9 ,内存仅占 1.41 MB,满足大部分嵌入式场景的应用。

为了验证提出的 ConvN-Classify 模型的实际应用效能,将模型进行部署应用,其帧率达到 100 帧/s 左右,满足实时检测的要求,在陌生数据集上的诊断效果也较为准确,模型鲁班性较好。后续的研究将进一步扩大数据集的规模,增加更多的病害种类和样本数量,以提高模型对各种病害的识别能力。同时进一步优化部署,提升实际应用的分类准确率。

参考文献:

- [1] 蒋怡,李宗南,董秀春,等. 1978—2020 年中国柑橘面积与产量格局演变研究[J]. 中国农业信息, 2023,35(1):43-54.
- [2] 邵峻青,魏霖静. 基于深度学习的农业病害图像识别的研究[J]. 计算机与数字工程, 2024,52(10):3131-3135.
- [3] 王俏,张彪,刘鑫. 基于改进锚分类的快速葡萄叶片病害检测算法[J]. 江苏农业科学, 2024,52(23):206-213.
- [4] 张帅杰,陈思思,戴丹,等. 基于密集卷积网络的甘薯黑斑

病害程度分类[J]. 中国农业大学学报, 2025,30(4):51-66.

- [5] 施宇,李敏. 基于卷积神经网络的柑橘病害叶片分类算法[J]. 信息与电脑(理论版), 2022, 34(3):47-51.
- [6] WOO S, DEBNATH S, HU R H, et al. ConvNeXtV2: co-designing and scaling ConvNets with masked autoencoders[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2023: 16133-16142.
- [7] LI H C, XIONG P F, AN J, et al. Pyramid attention network for semantic segmentation[EB/OL]. (2018-11-25)[2024-05-25]. <https://doi.org/10.48550/arXiv.1805.10180>.
- [8] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23)[2023-06-05]. <https://arxiv.org/abs/2004.10934>.
- [9] BANAN A, NASIRI A, TAHERI-GARAVAND A. Deep learning-based appearance features extraction for automated carp species identification[J]. Aquacultural engineering, 2020, 89: 134-139.
- [10] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521: 436-444.
- [11] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016: 770-778.
- [12] LOU H T, DUAN X H, GUO J M, et al. DC-YOLOv8: small-size object detection algorithm based on camera sensor[J]. Electronics, 2023, 12(10): 2323.
- [13] KHANAM R, HUSSAIN M. YOLOv11: an overview of the key architectural enhancements[EB/OL]. (2024-10-23)[2024-12-26]. <https://doi.org/10.48550/arXiv.2410.17725>.

【作者简介】

林汉源(2005—),男,广东汕尾人,本科,研究方向:数字图像处理、缺陷检测、深度学习。

范子安(2004—),男,湖南岳阳人,本科,研究方向:图像处理。

陈钦炯(2003—),男,广东普宁人,本科,研究方向:嵌入式开发。

方泽佳(2004—),男,广东揭阳人,本科,研究方向:图像处理。

徐竟成(2003—),男,广东佛山人,本科,研究方向:嵌入式开发。

(收稿日期: 2025-02-27 修回日期: 2025-07-01)