

基于改进 YOLOv10 的苹果叶部病害检测

王涛¹ 陈璇¹

WANG Tao CHEN Xuan

摘要

为解决苹果叶部病害能被快速及准确地识别问题,提出并验证了一种改进的 YOLOv10-n 模型。通过引入空间金字塔池化增强局部注意力网络(spatial pyramid pooling efficient lightweight aggregation network, SPPELAN)替换 YOLOv10-n 模型主干网络中的空间金字塔池化-快速模块(spatial pyramid pooling-fast, SPPF)以提升主干网络的多尺度特征融合能力,并采用组洗牌卷积(ghost convolution, GSConv)优化检测头部的计算效率。实验结果表明,改进后的 YOLOv10-n 模型对未参与训练的 711 幅图像的检测速度达 1.4 ms/幅,与原模型相比每幅图像的检测速度提升了 28.6%, mAP@50 提升 2.32%,参数量减少 6.47%。与基于单阶段目标检测框架下的 YOLOv3-SPP、YOLOv6 和两阶段目标检测框架下的 RT-DETR 相比, mAP@50 分别提升 72.95%、6.58% 和 121.92%。

关键词

深度学习;目标检测;改进 YOLOv10;苹果叶部病害识别

doi: 10.3969/j.issn.1672-9528.2025.08.006

0 引言

苹果作为全球重要的经济作物之一,广泛种植于世界各地。目前,中国已成为世界上最大的苹果生产国,种植面积和总产量均占全球的 50% 左右,苹果产业在中国农业经济中占据着重要地位^[1]。近年来,基于卷积神经网络(convolutional neural network, CNN)模型^[2]在农业病害检测中取得了显著成果。Jiang 等人^[3]提出的基于卷积神经网络改进的 INAR-SSD 模型,在苹果叶病数据集上实现了 78.80% 的 mAP 的检测精度,提供了较为高效实时检测解决方案,但该模型计算资源需求高,限制了模型在资源受限的设备上的部署。Sangeetha 等人^[4]的研究利用基于 VGG16 的卷积神经网络模型,实现了 93.3% 的最佳验证准确度,但该研究使用的是 Kaggle 公开数据集,这可能导致缺乏多样性,难以完全反映实际种植环境。Cui 等人^[5]提出了一种基于卷积神经网络的鱼类检测方法,以用于进行实时水下鱼类检测。代宁等人^[6]提出了一种改进的 YOLOv4-disc 算法,用于脊柱磁共振成像影像检测,该算法在平均精度均值(mAP)相比原生 YOLOv4 提升了 3.51 个百分点;这些研究展现了基于卷积神经网络的目标检测技术在多个新兴领域的广泛应用。

尽管卷积神经网络在目标检测方面效果显著,但在农业领域的实际应用中,面临数据集不足、样本不平衡和环境干扰等挑战,限制了其应用。此外,在移动端部署时受到计算和存储资源的限制,不得要求模型更加轻量化。通过改进 YOLOv10-n 模型,优化了特征提取和检测效率,并结合公开

数据集和自然采集的数据进行训练,得到一个检测速度和精度均较为平衡的轻量级模型。

1 实验准备与设置

1.1 数据收集与预处理

数据集的采集工作包括两个主要部分:第一部分数据来源于公开的 PlantVillage 数据集,共计 1 100 幅图像;第二部分数据源自中国贵州省威宁自治县雪山镇谢家社区的苹果种植基地,共计拍摄 2 455 幅图像。整个数据集共包含 3 555 幅图像。

该数据集涵盖了苹果叶部病害的七种类型,分别为斑点落叶病(Alternaria_Blotch)、黑腐病(Black_rot)、褐斑病(Brown_spot)、灰斑病(Grey_spot)、花叶病(Mosaic)、锈病(Rust)和黑星病(Scab)。每种病害对应的标签数量分别为 724 个、1 834 个、2 031 个、547 个、725 个、1 319 个和 2 000 个,总计 9 180 个标注信息,所有标签均使用 Labbing 工具进行标注。为了进行模型训练和验证,该数据集按 8:2 的比例随机划分为训练集和验证集,其中训练集包含 2 844 幅图像,验证集包含 711 幅图像。

本研究采用 SGD 优化器和余弦退火策略训练。SGD 优化器具有良好的泛化能力和避免过拟合的优势,而余弦退火策略通过逐步降低学习率提高模型精度,减少过拟合。

1.2 数据增强策略与超参数的设置

本文采用的数据增强策略包括添加高斯噪音、椒盐噪音、旋转、增亮、变暗。除此外还有在 HSV 颜色空间变换、旋转、平移、缩放、剪切、透视变换、图像翻转、马赛克、混合增

1. 贵州师范大学物理与电子科学学院 贵州贵阳 550025

强和随机擦除等，同时采用 RandAugment 策略进行自动数据增强^[7]、SGD 优化器和余弦退火策略。

在超参数的设置方面，动量设为 0.937 以加速收敛，并采用权重衰减 0.000 5 来防止模型过拟合。初始动量设置为 0.8，初始偏置学习率则为 0.1。此外，为了平衡不同损失项对训练过程的影响，分别为边界框损失、类别损失、分布式焦点损失（DFL）、姿态损失和关键点目标损失设置了增益值 7.5、0.5、1.5、12 和 1。

1.3 模型选择与优化

在目标检测领域，检测模型主要分为单阶段目标检测框架和两阶段目标检测框架。单阶段目标检测框架直接在网络的前向传播过程中完成目标的识别和定位，无需生成候选区域或进行额外的后处理步骤，这使得模型能够实现快速检测但对小目标或密集目标的检测精度较低。与之相对的两阶段目标检测框架，则是先通过区域提议网络生成一系列候选区域，然后在这些区域上进行分类和边界框的精确回归，虽然精度较高，但计算成本也相对较高。

YOLOv10 是一种单阶段目标检测模型，该模型采用大核心卷积和高效部分自注意力模块，显著提高了精确性和性能，降低了计算开销。后处理阶段引入一致的双重分配方法，实现了低延迟推理。YOLOv10-n 是该系列的轻量级模型，包含精简的主干网络、快速空间金字塔池化、逐点空间注意力和 YOLOv10 检测头^[8-9]。

在苹果叶部病害检测任务中，尽管 YOLOv10 模型的 SPPF 模块通过多尺度特征融合提高了特征表达能力，但在处理具有相似视觉特征的病害时，该模块可能难以有效区分，导致分类准确率降低。此外，尽管 SPPF 提高了特征提取效率，但在移动端或计算资源受限的环境中，其计算开销可能仍然较高，影响模型的实时应用。针对这些问题，引入的 SPPELAN 模块展现出了明显的优势。SPPELAN 结合了空间金字塔池化和轻量级聚合网络的优势^[10-11]，增强了对苹果叶片病害细微特征的捕捉能力，显著提升了模型对局部区域的敏感度，从而增强了特征融合能力，如图 1 所示。

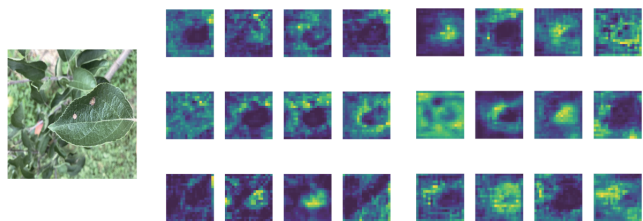


图 1 SPPF 模块与 SPPELAN 模块特征提取可视化

同时，SPPELAN 优化了特征通道聚合过程，不仅减少了计算量和参数数量，还保持了模型的表达能力，提高了检测效率。此外，SPPELAN 的轻量化设计使其更适合在资源受限的设备上部部署，满足了移动端和嵌入式系统的实时应用需求。进一步提升模型的检测性能和效率。采用了组洗牌卷

积（GSConv）重构了 YOLOv10-n 模型头部的卷积层，以提升模型的检测性能和效率。GSConv 通过引入组卷积和通道切换机制，有效减少了计算量和参数量，同时提升了特征信息的互通性。

2 实验结果与分析

2.1 模型精度对比

通过比较改进后的 YOLOv10-n 模型与基于单阶段目标检测框架的 YOLOv3-SPP、YOLOv6，以及基于两阶段目标检测框架的 RT-DETR 在苹果叶部病害数据集上的性能表现，可以看出，改进后的 YOLOv10-n 模型在该数据集上的 mAP@50 值明显更高，这表明其在苹果叶部病害检测任务中具有更优的检测性能，如图 2 所示。

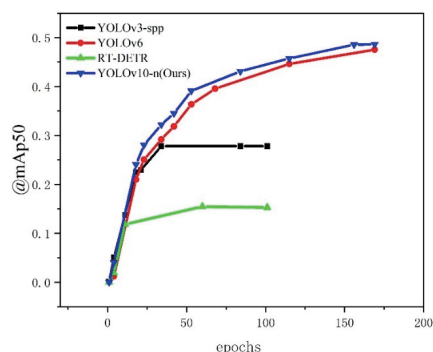


图 2 模型 mAP@50 精度对比

改进后的 YOLOv10-n 模型的 mAP@50 值为 0.486，相较于 YOLOv3-spp 模型的 0.281，提升了约 72.95%；相较于 YOLOv6 模型的 0.456，提升了约 6.58%；相较于 RT-DETR 的 0.219，提升了显著的 121.92%。这些结果表明，改进后的 YOLOv10-n 在特征提取、目标定位方面得到了有效改进，提升了检测性能。

2.2 模型性能分析

通过对 YOLOv10-n 模型进行了优化配置，并采用 F_1 得分、平均精度均值（mAP@50）和参数量评估模型性能。如表 1 所示，优化前的 F_1 值为 0.169，精度为 0.475，参数量为 2.71×10^6 。替换 SPPF 为 SPPELAN 后， F_1 值和精度保持不变，但参数量减少 5.81%。优化头部卷积为 GSConv 后， F_1 值增加 8.88%，精度增加 0.21%，参数量减少 0.65%。综合优化（替换 SPPF 为 SPPELAN 并优化头部卷积为 GSConv）后， F_1 值略微下降 1.18%，mAP@50 精度显著提升 2.32%，参数量减少 6.47%。还对比了不同优化配置下 YOLOv10-n 模型的推理时间。优化前，总推理时间为 1.8 ms。更换 SPPF 为 SPPELAN 后，总时间增至 2.5 ms；重构头部卷积为 GSConv 后，总时间仍为 2.5 ms。综合优化后，总时间显著减少至 1.4 ms。优化后的 YOLOv10-n 模型在推理和后处理阶段的时间显著减少，提升了实时性和资源利用率。

表 1 不同优化下 YOLOv10-n 模型的性能

优化配置	F_1	mAP@50	参数量/ 10^6	总时间/ms
优化前	0.169	0.475	2.71	1.8
仅对 SPPF 优化	0.169	0.475	2.55	2.5
仅重构头部卷积层	0.184	0.476	2.69	2.5
同时优化	0.167	0.486	2.53	1.4

通过检测效果对比图,如图 3 所示。可以看出,基线模型即原 YOLOv10-n 模型和优化后的模型在验证集上的表现,优化后的模型在验证推理结果中展现出更优异的性能。

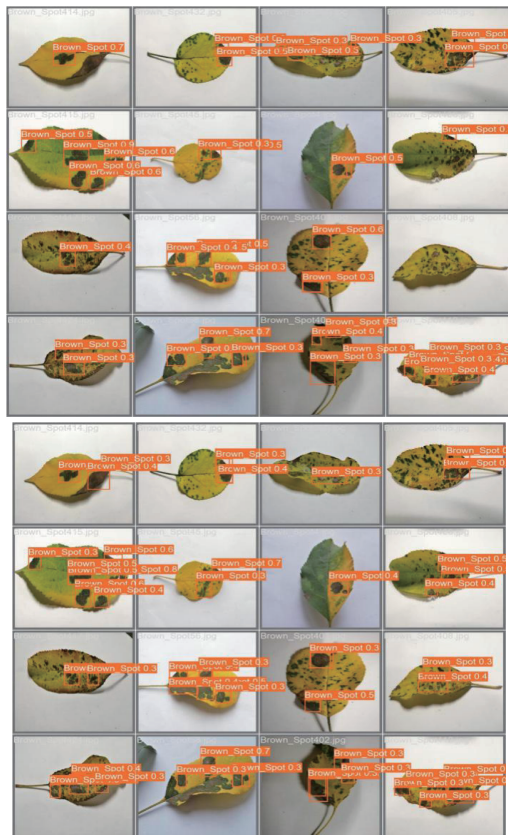


图 3 检测效果对比图

3 结论

为了提升苹果叶部病害早期发现和定位的准确性及速度,本文提出并验证了一种改进的 YOLOv10-n 模型。通过引入空间金字塔池化增强局部注意力网络(SPELAN)和组洗牌卷积(GSConv),改进模型在特征融合能力和效率上均有显著优化。实验结果表明,改进后的模型在未参与训练的 711 幅苹果叶部病害图片上表现优异,检测速度提升至 1.4 ms/幅, mAP@50 精度提升了 2.32%,参数量减少了 6.47%。与 YOLOv3-SPP、YOLOv6 和 RT-DETR 模型相比,改进后的 YOLOv10-n 模型在 mAP@50 精度上分别提升了 72.95%、6.58% 和 121.92%。该改进模型在准确性和效率上均有提升,为深度学习技术在农业病害监测中的应用提供新的优化方案。

参考文献:

- [1] 王壹. 中国苹果产量稳居世界第一 [N]. 农民日报, 2023-11-18(7).
- [2] LECUN Y, BOSER B, DENKER J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural computation, 1989, 1(4): 541-551.
- [3] JIANG P, CHEN Y H, LIU B, et al. Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks[J]. IEEE access, 2019, 7: 59069-59080.
- [4] SANGEETHA K, VISHNU R P, PRANESH K P, et al. Apple leaf disease detection using deep learning[C]//2022 6th International Conference on Computing Methodologies and Communication (ICCMC), Piscataway:IEEE, 2022: 1063-1067.
- [5] CUI S X, ZHOU Y, WANG Y H, et al. Fish detection using deep learning[J]. Applied computational intelligence and soft computing, wiley online library, 2020: 1-13.
- [6] 代宁, 谷玉海, 张志成, 等. 基于改进 YOLOv4 模型的脊柱磁共振成像影像检测 [J]. 激光与光电子学进展, 2023, 60(6): 0617002.
- [7] CUBUK E D, ZOPH B, MANE D, et al. AutoAugment: learning augmentation strategies from data[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Piscataway:IEEE, 2019: 113-123.
- [8] JOCHER G. YOLOv5 release v7.0[EB/OL]. [2024-03-15]. <https://github.com/ultralytics/yolov5/tree/v7.0>.
- [9] ZHAO H S, ZHANG Y, LIU S, et al. PSANet: point-wise spatial attention network for scene parsing[C]//Computer Vision (ECCV), Berlin:Springer, 2018: 267-283.
- [10] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 37(9): 1904-1916.
- [11] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]// 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2023: 7464-7475.

【作者简介】

王涛(1998—),男,贵州兴义人,硕士研究生,研究方向:计算机视觉与图像检测。

陈璇(1984—),女,江西上饶人,博士研究生,副教授,研究方向:神经网络稳定性分析与控制。

(收稿日期: 2025-03-05 修回日期: 2025-07-29)