

基于改进YOLOv10和LAMP通道剪枝的 串番茄成熟度检测算法研发

赵丽成^{1,2}, 卢鑫羽², 吴茜², 任妮², 周玲莉², 程雅雯²,
胡安琦², 戚超^{2*}

(1. 淮阴工学院化学工程学院, 江苏淮安 223003, 中国; 2. 江苏省农业科学院农业信息研究所, 江苏南京 210014, 中国)

摘要: [目的/意义] 由于串番茄果实密集重叠、温室光照复杂多变及果实颜色连续渐变等因素, 检测模型在对成熟度进行精准快速识别时面临挑战。因此, 提出了一种基于改进YOLOv10 (You Only Look Once version 10) 的轻量级串番茄成熟度目标检测模型 LampCT-YOLO (Cluster Tomato YOLO with Layer-wise Adaptive Mask Pruning), 以提高检测精度、推理速度和鲁棒性。[方法] 以YOLOv10为基线模型, 在主干网络引入SegNeXt (Segmentation Next) 注意力机制, 通过自适应调整注意力权重, 增强模型对3类不同成熟度串番茄关键区域(如边界、颜色深浅)的特征提取能力; 在训练完成的模型基础上, 通过基于梯度的全局通道重要性方法LAMP (Layer-wise Adaptive Mask Pruning) 通道剪枝, 显著压缩模型体积和计算复杂度, 同时有效保持模型对串番茄3分类成熟度的高检测性能。[结果和讨论] 在NVIDIA A100显卡环境下, LampCT-YOLO模型对串番茄红熟早期、中期、晚期的平均精度均值 (Mean Average Precision, mAP) 分别为84.6%、89.5%、88.4%, 相比YOLOv10分别提升了5.5、7.7和0.9个百分点, 对串番茄3个成熟度类别的mAP₅₀为87.6%, 相比YOLOv10提升了4.7个百分点。应用LAMP通道剪枝技术后, 模型的参数量和计算量分别减少63.07%和50.06%, 推理速度提升23.1%。部署至自主研发的果蔬巡检机器人后, 在边缘设备NVIDIA Jetson AGX Orin上达到9.1帧/s的推理速度, 实现了检测精度与实时性的有效平衡。[结论] 研究成果可为基于巡检机器人的串番茄成熟度精准快速识别提供坚实的技术支撑。

关键词: 串番茄; 成熟度检测; 注意力机制; 通道剪枝; 果蔬巡检机器人; YOLOv10

中图分类号: TP391.41; S641.2; S626

文献标识码: A

文章编号: SA202507045

引用格式: 赵丽成, 卢鑫羽, 吴茜, 任妮, 周玲莉, 程雅雯, 胡安琦, 戚超. 基于改进YOLOv10和LAMP通道剪枝的串番茄成熟度检测算法研发[J]. 智慧农业(中英文), 2026, 8(2): 133-146. DOI: 10.12133/j.smartag.SA202507045

ZHAO Licheng, LU Xinyu, WU Qian, REN Ni, ZHOU Lingli, CHENG Yawen, HU Anqi, QI Chao. An Improved YOLOv10-Based Tomato Ripeness Detection Algorithm with LAMP Channel Pruning[J]. Smart Agriculture, 2026, 8(2): 133-146. DOI: 10.12133/j.smartag.SA202507045 (in Chinese with English abstract)

0 引言

番茄是三大世界性贸易蔬菜之一。中国是世界上最大的番茄生产国, 总产量居于世界首位。串番茄又称为穗番茄, 是中国流行的一类整穗成熟、成串采收上市的番茄品种, 得益于商品性好而深受消费者的欢迎^[1]。成熟度的准确判断对于保证串番茄优越的商品性至关重要, 直接影响串番茄的市场价

值和消费者的购买选择。目前, 串番茄的成熟度判断大多依靠农户经验判断, 个体经验差异、光照条件、观察角度等因素都可能对判断结果造成影响, 具有较强的主观性和不确定性^[2,3]。随着机器人和计算机视觉技术的快速发展, 开发一种高效的串番茄成熟度检测方法迫在眉睫。

近年来, 目标检测技术在农业计算机视觉领域

收稿日期: 2025-07-30

基金项目: 国家自然科学基金(32201664); 江苏省农业科技自主创新资金(CX(24)1021); 农机研发制造推广应用一体化试点专项(JSYTH08)

作者简介: 赵丽成, 硕士研究生, 研究方向为作物表型算法。E-mail: zhao_orange@163.com

*通信作者: 戚超, 博士, 助理研究员, 研究方向为图像处理与机器视觉。E-mail: chaoqi@jaas.ac.cn

copyright©2026 by the authors

(登录 www.smartag.net.cn 免费获取电子版全文)

展现出强大的潜力^[4]。串番茄的目标检测方法主要分为基于传统机器学习的方法和基于卷积神经网络的方法^[5]。对于传统的机器学习方法,毕智健等^[6]对RGB (Red, Green, Blue)图像预处理后转换至HIS (Hue, Intensity, Saturation)和HSV (Hue, Saturation, Value)模型,提取多颜色分量均值,经SPSS (Statistical Package for the Social Sciences)特征筛选后用Matlab判别分析,成熟大番茄在验证集中的识别率为70%,体现出机器学习方式对人工颜色特征依赖强、预处理要求高。2019年,VIL-LASEÑOR-AGUILAR等^[7]提出了一种以RGB颜色模型为基础,借助模糊系统实现番茄成熟度分类,对比不同颜色模型的效果,并通过树莓派完成数据采集与估算,最终为番茄成熟度的自动化、低成本检测提供可行方案。该方法仅在受控环境下完成实验,后续若要应用于实际农田等复杂场景,还需优化以应对光照变化等干扰因素。2022年,王俊平和徐刚^[8]提出了一种基于机器视觉和电子鼻融合的番茄成熟度检测方法,通过机器视觉提取番茄的6个颜色特征,利用电子鼻采集10个气味特征,以这些多源特征为基础建立最小二乘支持向量机模型,模型在测试集识别准确率为84.02%,但该方法可变因素较多,导致其在成熟度识别上存在局限性。传统机器学习方法在串番茄成熟度识别中虽具备实现简单、响应快速等优势,但普遍存在对手工特征依赖性强、对图像预处理要求高、难以适应重叠遮挡与光照变化等复杂场景的局限。

随着深度学习和计算机视觉技术的迅速发展,基于卷积神经网络的相关方法得到了广泛研究与应用。GUO等^[9]研究提出多分辨率检测器用于番茄常见病害识别,通过Dropout (Random Dropout of Neural Network Units)和AdamW (Adam with Weight Decay Fix)优化器缓解过拟合问题,可识别10类番茄病害。该模型虽兼顾实时性,但平均识别准确率仅85.03%,低于同期多数深度学习模型,其精度短板源于多分辨率特征提取时对复杂病害斑点与背景的区分不足。SU等^[10]针对自然温室环境下大番茄成熟度分类任务,提出了一种结合挤压-激励注意力机制 (Squeeze-and-Excitation, SE)与MobileNetV1的改进YOLOv3网络,该模型推理速度达到28帧/s,在复杂遮挡场景下分类准确率为81.3%,且在内存较小的低端嵌入式平台上,推理延迟会增加至150 ms以上,参数规模约为12.8 MB,不利于在资源极度受限的设备上部署。2023年,WANG

等^[11]以YOLOv3-MobileNetV1为基础框架,引入SE注意力机制,强化了模型对串番茄成熟度区分性特征的关注能力。该模型在自然温室环境下,面对叶片遮挡的串番茄时识别率达到82%。

可见,目前主流的目标检测模型在面对串番茄成熟度识别时仍存在计算资源需求高、对复杂背景和光照变化适应性不足,以及模型结构复杂导致推理速度较慢和部署难度大等问题。如何在保持高检测精度的同时,有效压缩模型体积与计算资源消耗,成为当前研究的难点。鉴于此,本研究提出了一种结合层自适应幅值剪枝 (Layer-Adaptive Magnitude-Based Pruning, LAMP)通道剪枝与SegNeXt注意力机制的串番茄成熟度检测模型。该模型在保证高识别精度的前提下,通过LAMP剪枝实现网络瘦身,并利用SegNeXt增强对图像关键区域的关注能力,从而提升模型对串番茄3分类成熟度目标的感知与判别能力。本研究的主要贡献如下:

1) 提出了一种基于LAMP通道剪枝与SegNeXt注意力机制的轻量级串番茄成熟度检测模型,能够实现串番茄3分类成熟度的快速精准判别;

2) 通过引入基于梯度的全局通道重要性评估方法LAMP,实现了模型的高效轻量化;

3) 将LampCT-YOLO模型与YOLOv8、YOLOv10、YOLOv11及YOLOv12模型共计20个版本进行了全面对比,结果验证了所提出模型在串番茄成熟度检测任务中的优越性。

1 材料与方法

1.1 图像采集

本研究于2024年9月—2025年1月,在江苏省农业科学院智慧农业综合试验示范基地,开展3种分类成熟度的串番茄图像数据采集试验。在串番茄种植的成熟阶段,为了准确获取番茄果串的RGB图像数据,并用于后续的成熟度分级识别,本方案采用英特尔RealSense D435i RGB-D相机进行图像采集(图1)。具体采集措施如下。

1) 相机安装与定位。将深度相机固定于巡检机器设备的机械臂末端,确保相机在移动过程中保持稳定。机械臂的高度稳定在1.2 m左右,以匹配番茄植株的生长高度,避免因高度偏差导致图像畸变或遮挡。根据现场环境(如植株密度、光照条件等)调整相机的俯仰角度($\pm 30^\circ$)和水平偏转角度($\pm 15^\circ$),确保番茄果串始终位于RGB镜头的有效视野范围内。相机与目标果串的最佳拍摄距离控制

在0.5~1.0 m, 以兼顾图像清晰度和视野覆盖范围。若距离过近, 可能导致部分果串超出视场; 若距离过远, 则可能影响细节分辨率。

2) 相机参数设置。将相机的RGB镜头分辨率设置为 $1\,280 \times 720$, 确保图像具有足够的像素信息用于后续分析, 同时避免过高的分辨率导致存储压力。拍摄时采用30帧/s的采集帧率, 保证动态拍摄时图像的流畅性, 避免因帧率过低导致运动模糊。采用外接电源或大容量移动电源进行持续供电, 避免因电量不足中断采集任务。

3) 数据采集。移动巡检机器沿番茄种植行间匀速行走(保持在0.3~0.5 m/s), 确保相机在稳定状态下拍摄, 减少运动导致的使得拍摄图像模糊。采集过程中, 实时监测图像质量, 若发现光照不均、遮挡或失焦等问题, 立即调整相机参数或重新拍摄。将采集后的图像进行保存, 用于串番茄成熟度分级标准识别。

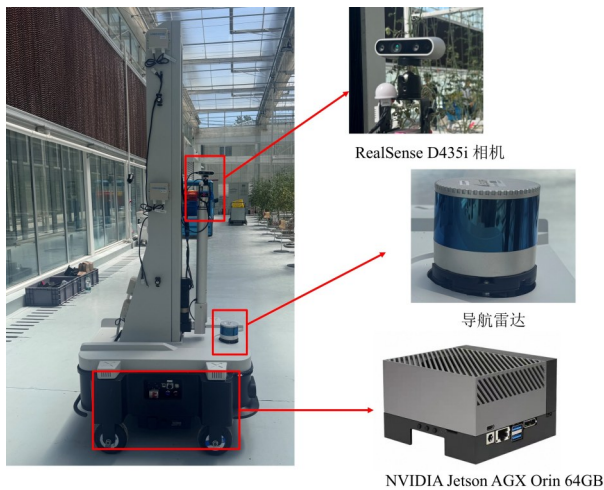


图1 移动巡检机器人现场图

Fig. 1 On-site image of the mobile inspection robot

在采集的串番茄图像中, 结合当前现行国标GHT 1193—2021^[12]和王新等^[13]在串番茄成熟度识别的中的分级标准, 根据串番茄红色番茄颗粒占据总颗粒数的比例划分红熟早期、红熟中期、红熟后期3类成熟度: 红色果实颗粒占据总颗粒数的10% (含)~不足30%为红熟早期, 红色果实颗粒占据总颗粒数的30% (含)~不足70%为红熟中期, 红色果实颗粒占据总颗粒数的70% (含)~100%为红熟晚期(图2)。

1.2 图像处理

为确保后续分析的准确性和效率, 对原始数据集进行筛选, 去除因拍摄抖动或重复拍摄导致的模

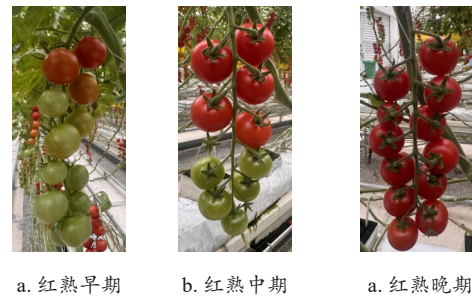


图2 串番茄不同成熟度示例图

Fig. 2 Example images of cluster tomatoes at different maturity stages

糊图像和内容高度相似的冗余样本。利用图像标注工具Labelme对筛选后的1 000张图像进行精细化标注, 生成包含类别标签红熟早期、红熟中期、红熟晚期的结构化数据集(表1)。

表1 串番茄三类成熟度标注样本数量

Table 1 Sample counts for three maturity stages of cluster tomatoes

成熟时期	训练集	验证集	测试集	总计(占比/%)
红熟早期	93	841	148	1 082(32.07)
红熟中期	104	934	166	1 204(35.07)
红熟晚期	107	831	149	1 087(32.86)

为提升模型泛化能力, 采用几何变换技术对标注数据进行增强处理(图3), 通过随机旋转模拟不同拍摄角度, 并应用非对称裁剪(保留70%~90%原始区域)增加果实排列的多样性, 最终将数据规模扩展至2 400张。将增强后的数据集按照8:1:1的比例, 划分为训练集(1 920张)、测试集(240张)和验证集(240张)。

1.3 LampCT-YOLO

1.3.1 LampCT-YOLO模型结构

在实际的串番茄成熟度识别任务中, 由于背景复杂且不同类别间的特征相似性较高, 传统模型在区分成熟度接近的番茄样本时常面临较大挑战^[14]。本研究提出LampCT-YOLO模型(图4), 该模型分为主干网络、特征融合层和检测层, 目标是提升模型在复杂背景下对不同成熟阶段番茄的检测精度, 尤其是对小而密集的番茄目标的识别能力^[15]。输入层采用了非对称裁剪和随机旋转的数据增强策略, 以提升模型对环境变化的适应能力, 并统一图像尺寸以提高训练效率。该策略对于自然光和温室条件下的串番茄成熟度识别尤其重要^[16]。主干网络在最后的C2f模块后插入了SegNeXt注意力机制,

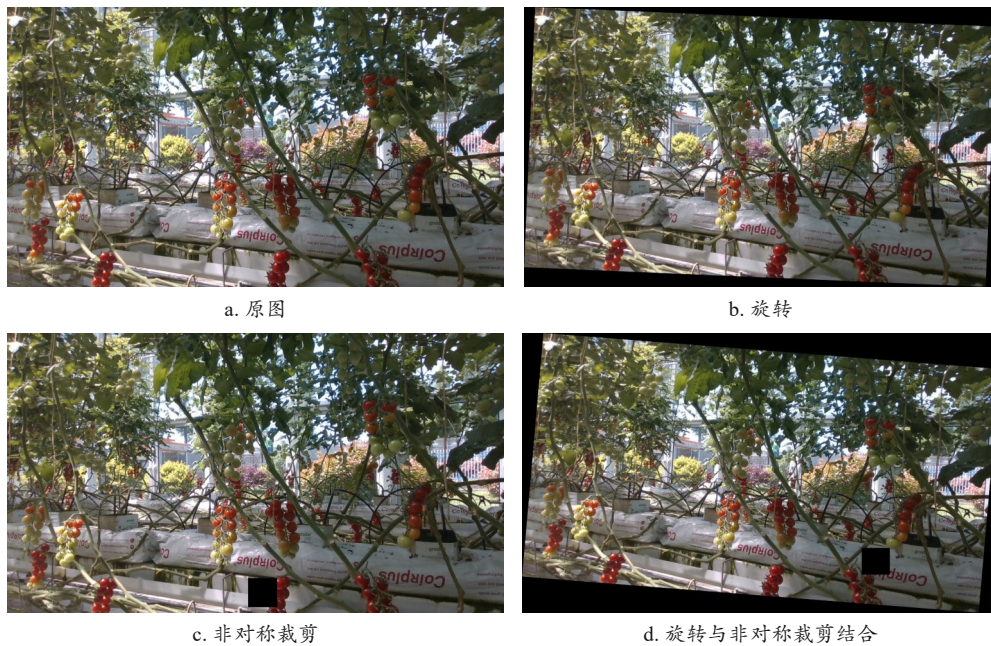


图3 串番茄图像数据增强示例图

Fig. 3 Data augmentation examples for cluster tomato images

提升对番茄细节特征的提取能力, 增强对果实边界、颜色深浅等成熟度关键特征的感知能力, 从而有效提升对成熟番茄和未熟番茄的区分度。特征融合层中, 采用轻量路径聚合网络(Path Aggregation Network, PAN)结构, 实现不同尺度特征图的融合。该设计有助于检测密集分布的串番茄, 尤其是大小不一、遮挡程度不同的果实。检测层以多尺度特征图为输入, 通过多个检测头对不同尺度的番茄果实进行预测, 最终输出统一的检测结果, 以实现对不同大小番茄果实的有效检测。模型通过 anchor-free 的预测方式, 提升对串番茄中不同形态果实的定位精度, 并结合分类与置信度预测, 实现红熟早期、红熟中期、红熟晚期的自动识别。通过上述设计, LampCT-YOLO 模型在保证检测速度的同时, 显著提升了在实际农业场景下的番茄成熟度识别准确性。

1.3.2 SegNeXt注意力机制

为提升模型的判别能力, 本研究在 YOLOv10 主干网络末端引入了 SegNeXt 注意力机制^[17]。该模块通过构建多尺度空间注意力与通道注意力的协同机制, 使模型能够动态增强与果实成熟度相关的特征响应, 同时有效抑制背景噪声的干扰。SegNeXt 具备出色的多尺度特征建模能力, 尤其适用于农业场景中果实排列多样、互相遮挡等情况, 能够引导模型在不同尺度下精准关注关键区域。与传统注意力机制如 SE^[18] 或 CBAM (Convolutional Block

Attention Module)^[19] 相比, SegNeXt 在计算开销较低的前提下, 能够更好地捕捉局部细节与多尺度特征信息, 对于区分成熟度相近但细节差异微弱的番茄至关重要。此外, SegNeXt 的引入显著增强了模型在复杂环境下的鲁棒性, 有效缓解了光照变化、果实重叠及叶片遮挡等因素对识别准确率的影响。具体来说:

SegNeXt 注意力模块首先通过 1 个 1×1 卷积将 Backbone 输出的高维特征图通道数压缩至固定维度, 以降低计算复杂度并统一特征表征空间。随后, 通过 3 条空洞卷积分支 (空洞率分别为 $r = \{7, 11, 21\}$) 在不同感受野下重组特征, 来兼顾被茎叶遮挡的局部成熟特征与整体形态信息。3 路输出在通道维度上拼接后, 经 1×1 卷积与修正线性单元 (Rectified Linear Unit, ReLU) 激活融合, 实现多尺度语义的非线性增强。接着, 对融合特征分别进行全局平均池化与最大池化, 得到两张统计特征图, 经过 7×7 卷积、BN (Batch Normalization) 与 Sigmoid (Sigmoid Function) 映射生成空间注意力图, 如公式 (1) 所示。

$$\mathbf{M}_s = \sigma(BN(f^{7 \times 7}([AvgPool(\mathbf{F}), MaxPool(\mathbf{F})]))) \quad (1)$$

式中: $\mathbf{M}_s$ 为注意力权重矩阵, 用于衡量特征图不同位置的重要程度; σ 为 Sigmoid 激活函数, 将输入映射到 (0, 1) 区间, 实现注意力权重的归一化; BN 为批归一化操作, 用于对输入特征进行

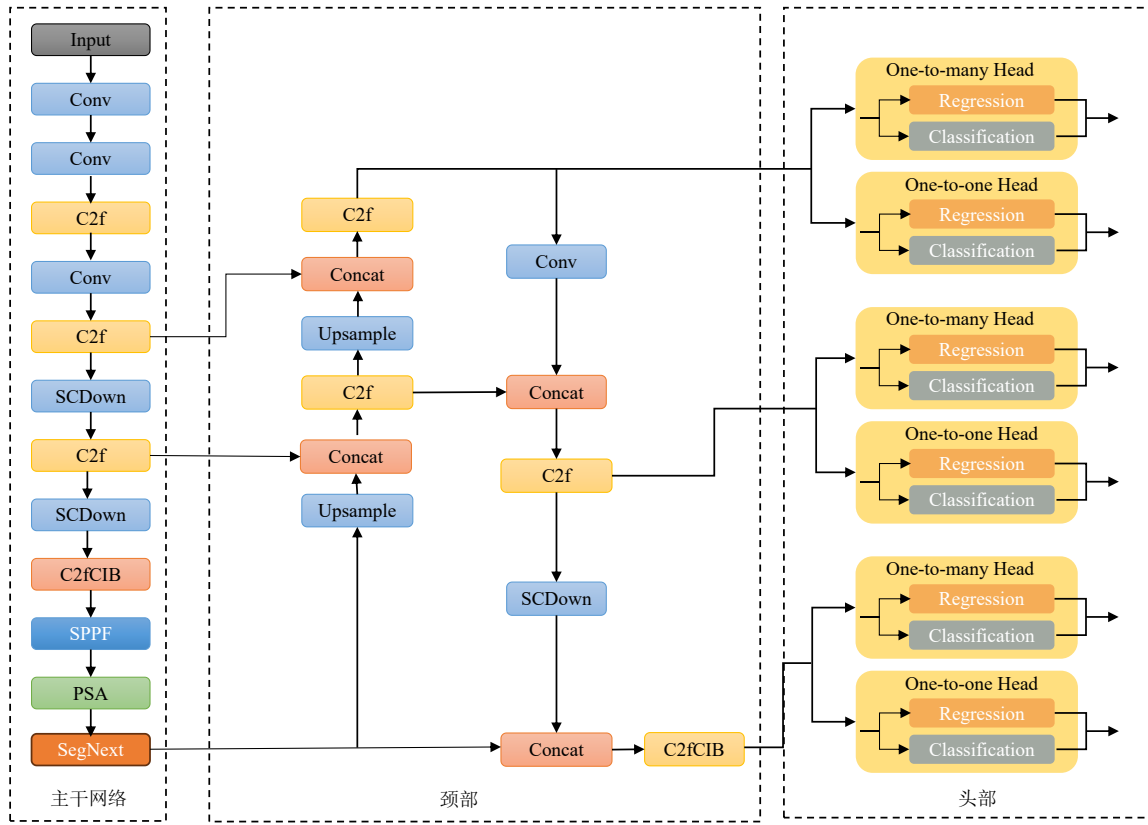


图4 LampCT-YOLO 模型结构图

Fig. 4 Structure diagram of the proposed LampCT-YOLO model

归一化，加速模型训练并提升稳定性； $f^{7 \times 7}$ 为尺寸为 7×7 的卷积操作，用于对输入特征进行特征提取与维度变换； $AvgPool(F)$ 是对特征图 F 进行平均池化操作，用于提取特征图的全局统计信息； $MaxPool(F)$ 是对特征图 F 进行最大池化操作，用于提取特征图的关键局部信息。

将其与融合特征逐元素相乘以突出关键区域，如公式（2）所示。

$$F' = M_s \otimes F \quad (2)$$

式中： F' 为经过注意力机制加权后的输出特征图（矩阵），融合了原始特征信息和注意力权重信息，能够突出重要特征并抑制无关特征。

最后，将加权特征与压缩通道特征残差相加，并经 1×1 卷积恢复通道数，如公式（3）所示。

$$F_{out} = Conv_{1 \times 1}(F' + F_c) \quad (3)$$

式中： F_{out} 为经过 1×1 卷积操作后的输出特征图（矩阵），是该运算的最终结果； $Conv_{1 \times 1}$ 是尺寸为 1×1 的卷积操作，用于对输入特征进行通道维度的变换或信息融合，同时保持特征图的空间尺寸不变； F_c 为另一路输入的特征图（矩阵）。

由此实现了特征的高效增强与稳定传递，为检

测头提供更加聚焦于成熟度判别的语义信息（图5）。

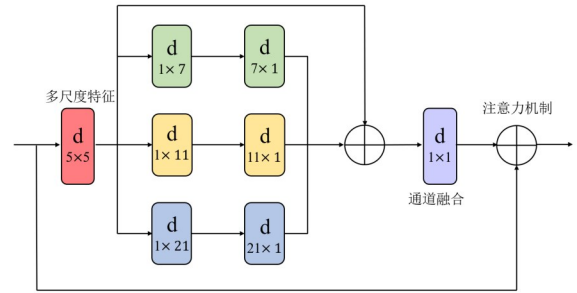


图5 SegNeXt注意力机制网络结构图

Fig. 5 Architecture diagram of SegNeXt attention mechanism network

1.3.3 LAMP通道剪枝

为在边缘设备或嵌入式设备等计算资源相对受限的设备中部署串番茄成熟度识别模型，对串番茄成熟度检测模型进行了轻量化操作。首先，将YOLOv10模型^[20]的卷积替换为深度可分离卷积（Depthwise Separable Convolution, DWConv）^[21]，受限于深度可分离卷积对通道间空间特征融合能力的弱化，在依赖细微颜色变化特征的串番茄成熟度

识别任务中,局部关键特征提取能力受损,导致模型准确率下降。这种效率与性能的权衡可通过数学关系量化表示,如公式(4)所示。

$$\begin{cases} \mathbf{P}' = \alpha \cdot \mathbf{P}, & \alpha < 1, \\ \mathbf{K}' = \beta \cdot \mathbf{K}, & \beta < 1, \\ A' = A - \Delta A, & \Delta A = 12.4\% \end{cases} \quad (4)$$

式中: $\mathbf{P}$ 和 $\mathbf{P}'$ 分别表示基线模型与轻量化模型的参数量; $\mathbf{K}$ 对应基线模型的计算量, $\mathbf{K}'$ 对应轻量化模型的计算量; $A = 87.6\%$ 和 $A' = 75.2\%$ 为模型改进前后的平均识别准确率; α 和 β 为比例因子,反映参数与计算量的压缩程度。试验表明,轻量化操作通过牺牲部分精度($\Delta A > 0$)换取模型效率提升($\alpha, \beta < 1$),为资源受限场景下的部署提供了可行性基础。

基于上述分析,本文引入模型剪枝^[22]方法对YOLOv10模型进行轻量化操作。考虑到串番茄成熟度检测任务对果实表面颜色细微差异具有较高敏感性,本文选择采用LAMP^[23]方法对模型进行结构化剪枝。该方法通过对不同层权重进行归一化处理,并在全局范围内统一评估各通道的重要性,从

而能够在控制剪枝强度的同时,较好地保持关键特征表达能力,因而适用于串番茄成熟度分级场景。

LAMP剪枝的基本步骤如下:首先,对第1层第*i*个通道对应的权重向量进行 L_2 归一化,以消除不同层之间的尺度差异,其计算方式如公式(5)所示。

$$\tilde{\mathbf{w}}_{l,i} = \frac{\mathbf{w}_{l,i}}{\|\mathbf{w}_{l,i}\|_2} \quad (5)$$

式中: $\mathbf{w}_{l,i}$ 表示第1层第*i*个通道的原始权重向量; $\tilde{\mathbf{w}}_{l,i}$ 表示归一化后的权重向量。

然后通过计算每个权重的重要性得分*s*,基于其幅度的平方和,如公式(6)所示。

$$s_i^{(l)} = \sum_j (\tilde{\mathbf{w}}_{l,i,j})^2 \quad (6)$$

式中: $\tilde{\mathbf{w}}_{l,i,j}$ 为归一化后的权重矩阵元素,由原始权重经归一化得到。最后将所有层的通道重要性得分合并进行排序,确保全局剪枝一致性。根据目标稀疏度和层敏感度,自动分配各层剪枝比例,移除重要性低的通道,同步调整相邻层的输入/输出通道数,保持网络连贯性,具体如图6所示。

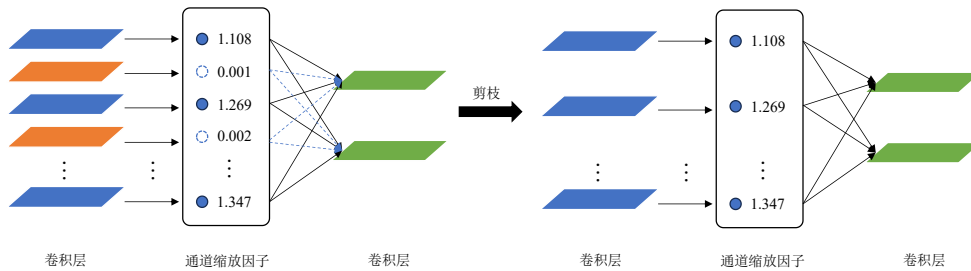


图6 LAMP通道剪枝示意图

Fig. 6 Schematic diagram of LAMP channel pruning

1.3.4 评估指标

为全面评估LampCT-YOLO模型的性能,选取 mAP_{50} (Mean Average Precision, mAP)、推理速度 (Frames per Second, FPS)、计算量GFLOPs、模型权重文件大小,以及参数量 (Parameters) 5项核心指标,从精度、速度、复杂度和部署成本等多个维度综合反映模型的优劣,具体定义如下:

mAP_{50} 用于衡量模型在识别不同成熟度等级串番茄目标时的检测准确率; N 表示类别数; AP_i (Average Precision of Class) 表示第*i*类在IoU=0.5下的平均精度,具体计算如公式(7)所示。

$$\text{mAP}_{50} = \frac{1}{N} \sum_{i=1}^N \text{AP}_i \quad (7)$$

式中: mAP_{50} 为平均精度均值,是目标检测任务的核心评价指标,衡量模型在IoU阈值为50%时

对所有类别的综合检测精度; AP_i 是衡量模型对单个类别检测性能的指标,通过计算该类别Precision-Recall曲线下的面积得到。

FPS的计算如公式(8)所示。

$$\text{FPS} = \frac{1}{T_{\text{inference}}} \quad (8)$$

式中: $T_{\text{inference}}$ 表示单帧图像的平均推理时间。

FLOPs和GFLOPs的计算如公式(9)和公式(10)所示。

$$\text{FLOPs} = 2 \times C_{\text{in}} \times C_{\text{out}} \times K_h \times K_w \times H_{\text{out}} \times W_{\text{out}} \quad (9)$$

$$\text{GFLOPs} = \frac{\sum \text{FLOPs}}{10^9} \quad (10)$$

式中: FLOPs为浮点运算次数,是衡量模型计算复杂度的核心指标,仅统计乘法和加法运算;

C_{in} 、 C_{out} 为输入、输出特征图的通道数； K_h 、 K_w 为卷积核的高度、宽度； H_{out} 、 W_{out} 为输出特征图的高度、宽度；GFLOPs 用于衡量 LampCT-YOLO 模型在执行 1 次完整识别任务时所需的计算量。

模型权重文件大小表示模型在存储设备中所占的空间，直接影响模型在资源受限的智能农机装备上的部署便捷性与加载速度。模型参数量则指 LampCT-YOLO 模型中需要学习和存储的参数数量，代表模型结构复杂度与内存资源消耗水平。

通过以上多维度指标的综合评估，可全面分析模型在串番茄成熟度识别任务中的准确性、效率与轻量化表现。

2 结果与分析

2.1 试验设置

本研究试验在配备 NVIDIA A100 系列 GPU、Intel i7 处理器及 32 GB 内存的工作站上完成，操作系统为 Ubuntu 18.04，深度学习框架基于 PyTorch 2.0.1 与 CUDA 11.8。

在稀疏训练阶段，采用了 LAMP 稀疏正则化策略，其主要目的是通过稀疏化模型中的冗余通道参数，提升推理效率并降低计算成本。具体的设置包

括：稀疏正则化权重 0.04，正则衰减因子 0.05，并将训练轮数设置为 500 轮。该过程能够有效促使网络学习到更加紧凑的特征表示。

在超参数配置上，采用了 AdamW 优化器，初始学习率为 0.001，并使用余弦退火学习率^[24]调度 (Cosine Annealing LR Scheduler)。损失函数方面，检测部分使用 CIOU Loss (Complete Intersection over Union Loss) 进行边界框回归，分类和置信度分支则使用了 Focal Loss，以解决类别不平衡问题。Batch Size 设为 16，并使用了 0.000 1 的权重衰减。为了防止过拟合，引入了早停机制，当训练集的损失在连续 50 轮内未进一步下降时，训练自动终止。

2.2 消融试验

为验证改进模型在串番茄成熟度识别任务中的效果，本研究设计了两组消融试验。第 1 组试验探讨在基线模型 YOLOv10 基础上引入 SegNeXt 注意力机制的有效性 (表 2)；第 2 组试验比较了深度可分离卷积与所提出的 LAMP 通道剪枝方案对模型轻量化与检测性能的综合影响 (两者目的都是在减少模型的计算量和参数量，但优化方式不同)。两组试验通过检测精度 mAP_{50} 、推理速度，以及权重文件大小指标进行对比分析 (表 3)。

表 2 串番茄成熟度检测的 SegNeXt 注意力机制消融试验

Table 2 Ablation experiment of SegNext attention mechanism for cluster tomato maturity detection

模型	红熟早期/%	红熟中期/%	红熟晚期/%	$mAP_{50}/\%$	FPS/(帧/s)	计算量/GFLOPs	参数量/MB	权重文件大小/MB
YOLOv10	79.1	81.8	87.5	82.8	37.37	106.700	45.811	107.7
YOLOv10+SegNeXt	84.6(+5.5)	89.5(+7.7)	88.4(+0.9)	87.5(+4.7)	35.8(-1.57)	228.765	53.522	107.6(-0.1)

注：括号内数字表示该模型相较于基础模型 YOLOv10 的性能变化量，“+”表示提升，“-”表示下降。

试验结果表明，在第 1 组消融试验中，引入轻量级 SegNeXt 注意力机制后，模型在红熟早期、红熟中期和红熟晚期的识别精度分别提升了 5.5、7.7

和 0.9 个百分点， mAP_{50} 提高至 87.5%，相较原始 YOLOv10 提升了 4.7 个百分点，表明注意力机制有助于增强模型对细粒度特征的感知能力。

表 3 Lamp 模型剪枝消融试验

Table 3 Ablation experiment on the pruning of the lamp model

模型	红熟早期/%	红熟中期/%	红熟晚期/%	$mAP_{50}/\%$	FPS/(帧/s)	计算量/GFLOPs	参数量/MB	权重文件大小/MB
YOLOv10+SegNeXt	84.6	89.5	88.4	87.5	35.8	228.7	53.5	107.6
YOLOv10+SegNeXt +DWConv	75.4 (-9.2)	75.0 (-14.5)	82.2 (-6.2)	77.5 (-10.0)	37.7 (+1.9)	115.2 (-113.5)	21.1 (-32.4)	52.2 (-55.5)
YOLOv10+SegNeXt +Lamp	81.4 (-3.2)	84.3 (-5.2)	91.9 (+3.5)	85.9 (-1.6)	66.9 (+23.1)	114.2 (-114.5)	19.7 (-33.7)	40.0 (-67.7)

注：括号内数字表示该模型相较于基础模型 YOLOv10 的性能变化量；“+”表示提升；“-”表示下降。

在第 2 组试验中，为提升模型轻量化程度并兼顾检测性能，试验对比了深度可分离卷积与 LAMP 通道剪枝策略。结果显示，DWConv 虽有效减小了

模型体积 (权重文件减小至 52.2 MB，参数量降至 21.1 MB，计算量减少 32.4 GFLOPs)，但检测精度下降明显， mAP_{50} 仅为 77.5%。相比之下，LAMP

剪枝在进一步压缩模型参数(至19.7 MB)和计算量(降至114.2 GFLOPs)的同时,显著提升了推理速度(66.90 帧/s, 提升23.1%),仍保持较高的检测性能(早期下降3.2%、中期下降5.2%,而晚期成熟果实提升3.5%,平均 mAP_{50} 达85.9%),验证了其在模型压缩与性能保持之间的优越权衡能力。

2.3 引入SegNeXt可视化结果对比

引入的SegNeXt模块在保证低计算开销的前提下,有效增强了对复杂背景中成熟度差异细微区域的感知能力,使得模型在不同成熟阶段串番茄目标区分上更具鲁棒性,与基线模型进行可视化对比可以清晰地观察到SegNeXt注意力机制对模型性能的多方面提升(图7)。

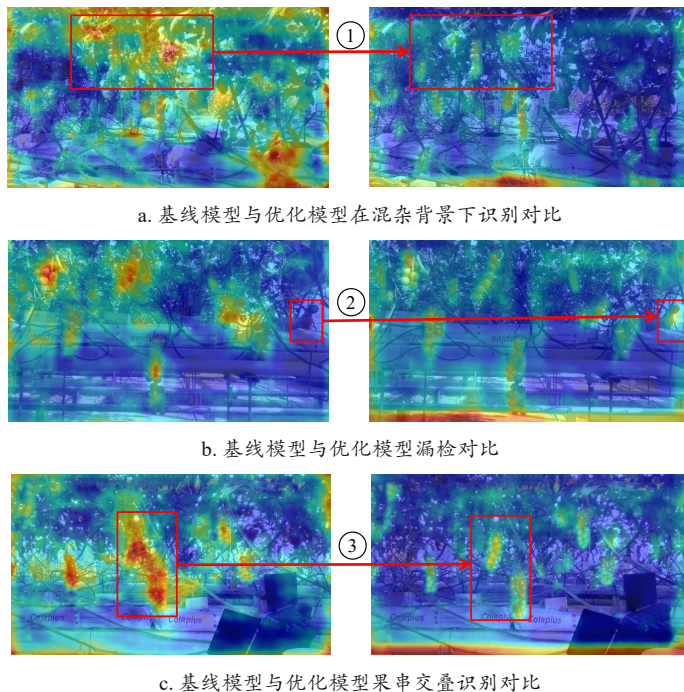


图7 LampCT-YOLO模型优化前后注意力可视化对比

Fig. 7 Visual comparison of attention mechanism before and after optimization of LampCT-YOLO model

基线模型倾向于将目标识别为较为宽泛的“植物”或“区域异常点”,而引入SegNeXt注意力机制后的优化模型能够更精准识别出目标物体的类别(图7),这表明该注意力机制有效增强了模型对关键语义特征的提取和判别能力。在检测置信度方面,改进后的模型在红色高置信度区域更集中、范围更大且颜色更深,增加的注意力机制引导模型聚焦于与检测任务最相关的图像区域,抑制了背景信息的干扰,提高了模型对串番茄不同成熟度检

测结果的准确率。具体而言,在图7区域①中,基线模型虽显示了目标检测结果,但未能准确识别具体物体类别,边界多围绕枝干、叶片及果实,类别划分不明确且未集中于番茄果串,导致结果难以直接解读;优化后的模型则有效聚焦于番茄果串,显著减少了叶片和枝干的干扰。在图7区域②中,基线模型因注意力未能集中于番茄果串,出现漏检现象;而引入SegNeXt注意力机制模型成功避免了此类漏检。在图7区域③中,果串间存在明显粘连,颜色与位置重叠,基线模型可能将多个粘连果串误判为单一目标,影响检测准确性。通过增强对关键特征的关注,优化后的模型能够更有效地识别并分离粘连果串,降低误判与漏判的概率。

2.4 剪枝试验

为了提升模型的部署效率与推理速度,本研究采用LAMP通道剪枝方法对模型结构进行优化。LAMP剪枝策略通过引导网络自动识别各层通道的重要性,实现通道级别的精细化裁剪。在稀疏训练阶段,采用了LAMP稀疏正则化策略,其主要目的是通过稀疏化模型中的冗余通道参数(图8)提升推理效率并降低计算成本。

剪枝优化后,模型的计算量由原始的228.765 GFLOPs显著下降至114.252 GFLOPs,计算复杂度降低约50.06%;参数量从53.522 MB减少至19.765 MB,压缩幅度高达63.07%。在检测性能方面,虽然 mAP_{50} 略微下降至85.9%,仅减少了1.6%,但考虑到模型在推理速度和资源消耗方面的显著提升,该精度损失在串番茄成熟度检测这一应用场景中是可以接受的,尤其适用于资源受限的机器人与边缘计算设备。图8展示了剪枝前后模型各层参数量的变化情况,显示大多数网络层在剪枝后均实现了显著的参数压缩,特别是第6层和第7层,参数量分别从约1300万和700万大幅下降至约400万和250万,表明模型冗余主要集中在中间特征融合与通道扩展模块中。其余各层也普遍呈现出压缩趋势,说明LAMP策略能够有效识别全网络范围内的重要通道,并在保持整体性能的前提下剔除冗余结构。

为验证剪枝前后模型性能差异的显著性,对LampCT-YOLO模型进行了10次独立实验(图9)。其平均 mAP_{50} 为87.549%,标准差仅约0.031%,表明模型性能非常稳定。95%置信区间为(87.527%,

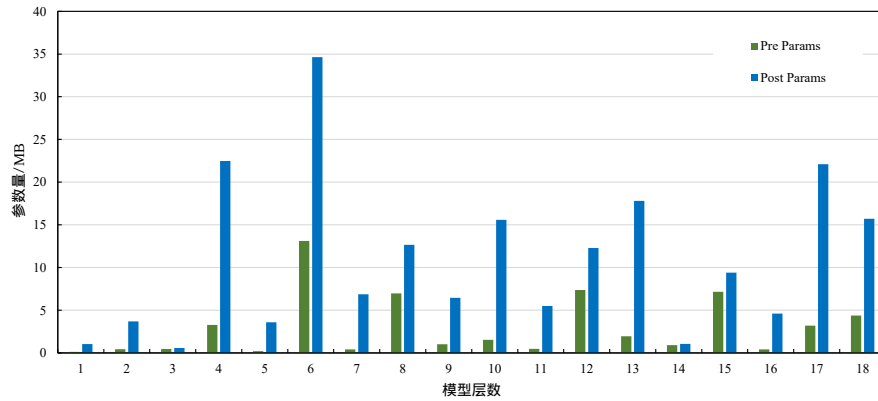


图8 LampCT-YOLO 模型各层剪枝前后参数对比图

Fig. 8 Comparison chart of parameters of each layer of the LampCT-YOLO model before and after pruning

87.571%), 区间范围极窄, 进一步说明实验结果具有高度可靠性与重复性。

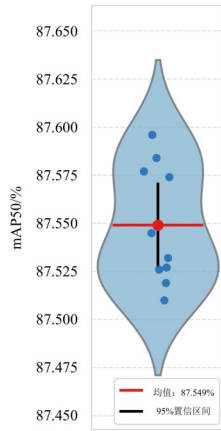


图9 LampCT-YOLO 模型置信区间与分布小提琴图

Fig. 9 Confidence interval and distribution violin plot of LampCT-YOLO model

试验结果证明该方法不仅能够保持较高的检测准确率, 还显著提升了模型在计算资源受限场景下的部署可行性, 展现出良好的实用价值和推广潜力。

2.5 与主流目标检测模型的比较

为全面评估所提出的改进模型的性能, 本试验选取了当前主流的 YOLO 系列模型作为对比对象, 包括 SSD (Single Shot MultiBox Detector)、Faster R-CNN (Faster Region-based Convolutional Neural Networks)、YOLOv7、YOLOv8^[25]、YOLOv10、YOLOv11^[26] 和 YOLOv12^[27]。其中 YOLOv8 至 YOLOv12 模型均选取 nano、small、medium、large、extra-large 这 5 个版本 (代表不同的网络深度和宽度), 以保证比较的全面性。

根据表 4 可知优化后的模型优于主流 YOLO 系

列模型。精度方面, 其 mAP_{50} 达 87.5%, 高于所有对比模型, 表明在串番茄成熟度检测中改进后的模型具有更强的识别能力。推理速度方面能够达到 66.2 帧/s, 相比其他模型普遍处于 30~45 帧/s 的水平, 展现出显著的实时性优势。模型参数量为 19.77 MB, 权重大小为 40.0 MB, 保持了良好的部署友好性。综合来看, 该模型在精度、速度和规模之间实现了最优权衡, 适用于农业场景中的高效部署。

同时, 为探究不同版本 YOLOv10 模型在引入 SegNeXt 注意力机制与 Lamp 模块后的性能表现, 本研究分别在 YOLOv10n 与 YOLOv10l 上进行了对比。结果如表 5 所示。

表 5 展示了不同版本的 YOLOv10 模型在引入 SegNeXt 注意力机制及 Lamp 模块后性能的对比结果。在 nano 版本中引入 SegNeXt 后精度虽略有提升, 但整体性能受模型容量限制, 提升幅度有限; 进一步引入 Lamp 模块后, 虽然推理速度有所提高, 但检测精度出现一定下降。相比之下, large 版本在引入 SegNeXt 注意力机制后精度显著提升 (mAP_{50} 由 82.6% 提升至 87.5%), 表明 large 版本模型具有更强的特征提取与注意力响应能力。此外, 在融合 Lamp 模块后, large 版本在保持较高精度的同时大幅提升了推理速度 (FPS 由 35.80 帧/s 上升至 66.90 帧/s, 提升 46.49%), 验证了本研究提出模型在性能与效率间取得了较好的平衡。

此外, 为了进一步验证模型在稀疏训练阶段的稳定性, 本研究引入了 LAMP 稀疏正则化机制, 并绘制了训练损失曲线 (图 10)。从图 10 中可以看出, 3 类损失 (box_loss、cls_loss、dfl_loss) 在稀疏训练阶段均呈平稳下降趋势: 初期下降迅速, 后期波动较小并逐渐趋于稳定, 表明 LAMP 正则化并

表4 增加注意力机制后模型与主流目标检测模型对比试验

Table 4 Comparison of the attention-enhanced model with mainstream object detection models

模型名称	mAP ₅₀ /%	FPS/(帧/s)	计算量/GFLOPs	模型参数量/MB	权重文件大小/MB
SSD	77.3	43.22	37.8	21.472	91.6
Faster RCNN	80.7	15.75	177.3	115.684	108.3
YOLOv7	76.0	48.25	51.6	36.490	71.3
YOLO8n	84.5	39.89	8.1	3.006	6.3
YOLO8s	84.0	43.82	28.4	11.127	22.5
YOLO8m	82.5	45.41	78.7	25.841	52.0
YOLO8l	83.6	39.91	164.8	43.609	87.7
YOLO8x	82.4	37.54	257.4	68.126	136.7
YOLOv10n	82.8	43.57	6.0	2.207	5.6
YOLOv10s	81.9	44.49	21.4	7.219	16.5
YOLOv10m	82.5	44.57	58.9	15.315	33.5
YOLOv10l	82.6	37.37	228.7	53.522	107.7
YOLOv10x	79.0	41.60	160.0	29.399	130.4
YOLOv11n	85.9	43.99	6.3	2.583	5.5
YOLOv11s	85.3	43.12	21.3	9.414	19.2
YOLOv11m	85.3	37.89	67.7	20.032	40.5
YOLOv11l	85.9	37.15	194.4	56.830	114.4
YOLOv11x	84.0	37.21	194.4	58.362	114.4
YOLOv12n	84.1	39.71	5.8	2.509	11.2
YOLOv12s	82.0	40.57	19.3	9.074	18.6
YOLOv12m	84.6	41.98	59.5	17.579	39.7
YOLOv12l	79.9	36.23	82.1	26.396	53.7
YOLOv12x	81.3	33.39	184.1	59.248	119.5
YOLOv10 l + SegNeXt	87.5	66.20	114.2	19.765	40.0

表5 YOLOv10不同规模模型在SegNeXt与Lamp模块作用下的检测性能分析

Table 5 Detection performance analysis of different YOLOv10 models with SegNeXt and Lamp modules

模型名称	mAP ₅₀ /%	FPS/(帧/s)	计算量/GFLOPs	模型参数量/MB	权重文件大小/MB
YOLOv10n	82.8	43.57	6.000	2.207	5.6
YOLOv10n+ SegNeXt	85.5	41.50	8.600	2.450	6.3
YOLOv10n+ SegNeXt+Lamp	84.8	51.20	6.500	2.000	5.1
YOLOv10l	82.6	37.37	106.700	45.811	107.7
YOLOv10 l + SegNeXt	87.5	35.80	228.765	53.522	107.6
YOLOv10l+ SegNeXt+Lamp	85.9	66.90	114.252	19.765	40.0

未导致训练发散或收敛不稳定。尤其是定位损失的平滑下降趋势，说明在参数稀疏化的同时，模型的边界框回归能力保持良好；而分类损失同样顺利收敛，进一步证明模型在稀疏化过程中仍能保持较强的特征判别能力。

为了进一步验证 LampCT-YOLO 模型在实际应用中的检测效果，本研究用 YOLOv10、YOLOv10 增加注意力机制后的模型和 LampCT-YOLO 模型，在多组不同环境下的番茄图像上进行可视化对比，包括理想状态（果串无遮挡、光线均匀、背景简

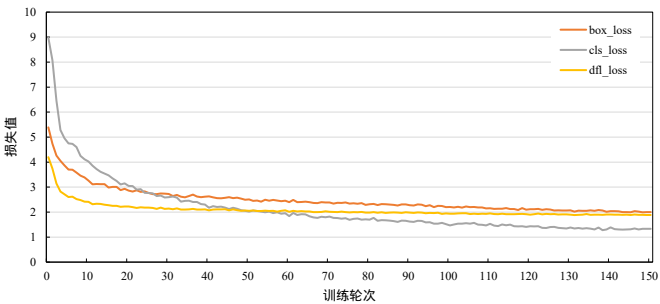


图10 LampCT-YOLO模型训练损失下降趋势图
Fig.10 Training loss convergencecurve of the LampCT-YOLO model

单)、光线过暗过曝(光线Lux值低于30为过暗,光照度Lux值高于1 000为过曝)、果串被部分遮挡(果串被叶片、枝条或其他果串部分遮挡,遮挡率超过10%)、果串交叠(多个果串交叠,果实之间

边界不清晰,交叠区域超过10%)等典型场景。通过对比不同模型的检测结果,可以更直观地观察LampCT-YOLO在目标定位精度、漏检率,以及边界框鲁棒性方面的优势(图11)。

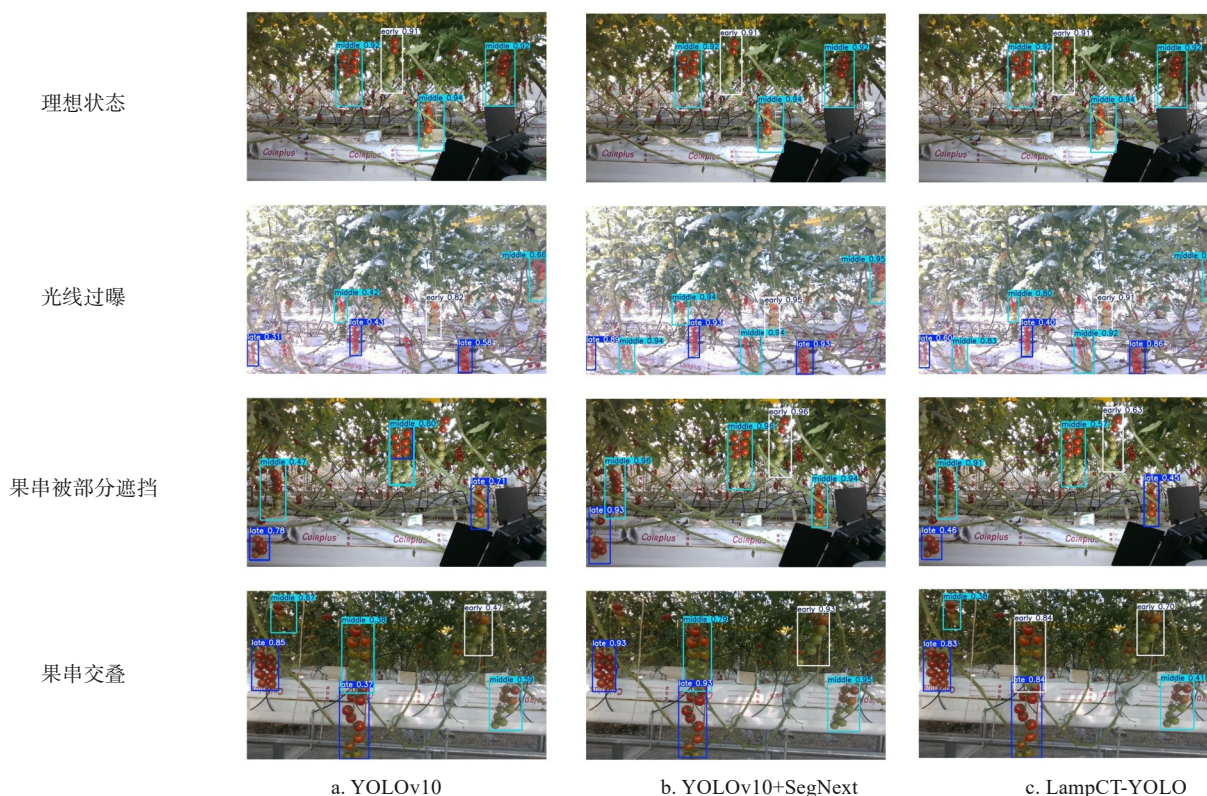


图11 复杂情况下串番茄检测示例图

Fig. 11 Example images of cluster tomato detection under complex conditions

2.6 果蔬巡检机器人部署

为了推动串番茄成熟度精准快速识别在农业生产中的落地应用,本研究将自主研发的轻量化串番茄成熟度检测模型LampCT-YOLO部署至设施果蔬智能巡检机器人中。部署流程分为3个主要阶段,如图12所示,包括自主导航至目标位置、开始巡检任务,以及现场识别与反馈。巡检机器人通过自主导航系统精确定位,行驶至果园指定起始位置(图12a);机器人启动巡检任务,实时采集周围图像并传输至NVIDIA Jetson AGX Orin进行分析(图12b);经过模型LampCT-YOLO处理,机器人识别出不同成熟度的串番茄,并在可视化界面实时显示成熟度及置信度(图12c),直观展示模型在真实温室环境下的检测准确性与鲁棒性。本次部署验证了LampCT-YOLO模型在设施果蔬巡检机器人上应用的可行性和高效性。

为验证移动设备的部署对模型的检测精度的影

响,本研究对456串番茄进行成熟度分类检测,检测速度能够达到9.1帧/s,将移动端检测结果与人工计数进行了对比,结果如表6所示。

从表6中可以看出,移动设备端的检测结果与人工计数总体一致,但存在一定程度的误检与漏检。结果显示,巡检机器人成功识别了红熟早期、红熟中期、红熟晚期串番茄各78串、61串、248串,检测准确率分别为84.8%、87.1%、84.4%,平均准确率达85.4%。同时,对于红熟早期、红熟中期、红熟晚期的串番茄分别误检了5串、7串、10串,分别漏检了9串、2串、36串。相比于LampCT-YOLO模型测试得到的 mAP_{50} 为85.9%,移动端检测精度略有下降(约0.5%)。该差异主要受限于移动设备算力较低、图像采集光照变化,以及实时检测过程中的压缩误差等因素。综上,移动设备的部署对模型检测精度影响较小,模型在移动设备端侧仍能保持较高的检测性能与稳定性,具有良

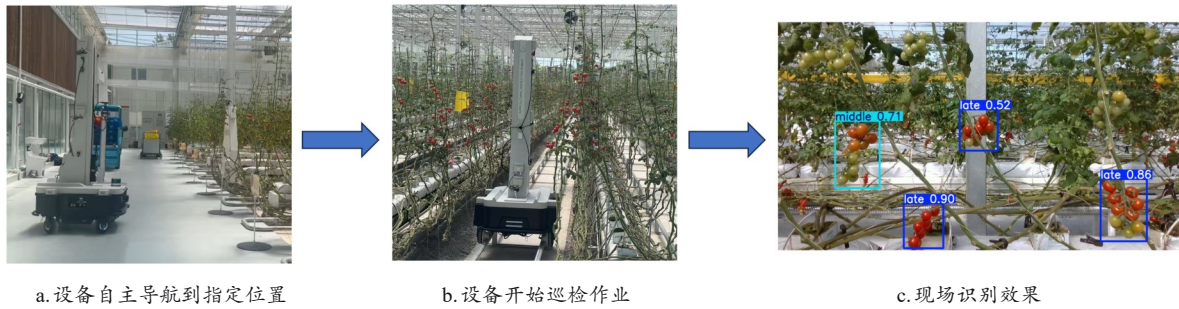


图 12 果蔬巡检设备现场运行图

Fig. 12 On-site operation diagram of intelligent inspection equipment

表 6 LampCT-YOLO 模型部署移动设备检测结果与人工计数的对比

Table 6 Comparison between mobile device detection results of the LampCT-YOLO model and manual counts

	红熟早期/串	红熟中期/串	红熟晚期/串	误检/串	漏检/串
人工计数	92	70	294	0	0
设备检测	78	61	248	22	47

好的应用可行性。本研究验证了LampCT-YOLO模型在设施果蔬巡检机器人上应用的可行性和高效性。

3 结 论

本研究针对串番茄果实密集重叠和复杂光照带来的成熟度精准快速检测挑战,提出了融合LAMP通道剪枝与SegNeXt注意力机制的串番茄成熟度轻量化识别模型LampCT-YOLO,该模型有效提升了不同成熟度番茄的识别准确率并降低了模型复杂度,实现了在资源受限场景中的有效部署。试验结果显示,LampCT-YOLO通过增强模型对关键目标区域的感知能力,提升识别精度(mAP_{50} 为85.9%),通过剪枝技术大幅度压缩模型体积,减少了63.07%的参数量和50.06%的计算量,同时推理速度提升了46.49%,部署于巡检机器人后,所提出模型有效平衡了果蔬巡检机器人对实时性与精度的需求,展现出良好的落地推广前景。

剪枝后模型 mAP_{50} 仍较原模型略有下降(约1.6%)。从类别层面分析可见,早期(下降3.2%)与中期(下降5.2%)果实的识别精度下降更为明显,而晚期成熟果实(提升3.5%)反而略有提升。这表明剪枝在一定程度上削弱了模型对细微颜色变化与边界模糊区域的特征敏感性,尤其是对于颜色渐变不明显的早期和中期果实。为弥补这部分性能损失,可以尝试在剪枝后引入微调训练或知识蒸馏策略,以恢复模型的特征表达能力。

然而,该方法仍存在两点缺陷:一是模型对于光照剧烈变化或果实严重遮挡的极端场景适应性有限;二是针对非红色品种串番茄(成熟后也为青色、黑色、黄色等),不具有普适性。针对复杂环境适应性问题,可引入多模态信息(如融合深度图像或光谱图像)以增强模型对被遮挡果实的补全能力;多模态信息能够提供更丰富的空间和光谱特征,有助于提升模型在复杂环境中的检测鲁棒性和准确性。针对模型对不同颜色品种串番茄的泛化能力,可引入视觉语言模型(Visual-language Model),通过结合图像和文本语义信息,视觉语言模型能够更好地理解“成熟”“未熟”等语义概念,从而提升模型对不同品种、多样外观的串番茄果实的识别能力,实现跨品种、跨场景的鲁棒检测。

综上所述,本研究工作为果蔬成熟度智能识别领域提供了可落地的轻量化目标检测新路径。未来研究可围绕多模态感知融合、结合视觉语言模型来增强模型的语义理解能力等方向展开,进一步推动智能农业在边缘计算平台的普及与应用。

利益冲突声明: 本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献:

- [1] 汉永乾,孙步功,张鹏,等.农业采摘机器人研究进展[J].林业机械与木工设备,2023,51(4):4-8.
HAN Y Q, SUN B G, ZHANG P, et al. Research progress of agricultural picking robot[J]. Forestry Machinery & Woodworking Equipment, 2023, 51(4): 4-8.
- [2] MA J, LI M J, FAN W P, et al. State-of-the-art techniques for fruit maturity detection[J]. Agronomy, 2024, 14(12): 2783.
- [3] RIZZO M, MARCUZZO M, ZANGARI A, et al. Fruit ripeness classification: A survey[J]. Artificial Intelligence in Agriculture, 2023, 7: 44-57.
- [4] 张东彦,高玥,程涛,等.基于YOLOv8-FECA的密集场景下小麦赤霉病孢子目标检测[J].农业工程学报,2024,40(21):127-136.
ZHANG D Y, GAO Y, CHENG T, et al. Detection of wheat scab spores in dense scene based on YOLOv8-FECA[J]. Transactions of the Chinese Society of Agricultural Engineering, 2024, 40(21): 127-136.

- 127-136.
- [5] 张勤, 陈建敏, 李彬, 等. 基于RGB-D信息融合和目标检测的番茄串采摘点识别定位方法[J]. 农业工程学报, 2021, 37(18): 143-152.
ZHANG Q, CHEN J M, LI B, et al. Method for recognizing and locating tomato cluster picking points based on RGB-D information fusion and target detection[J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(18): 143-152.
- [6] 毕智健, 张若宇, 齐妍杰, 等. 基于机器视觉的番茄成熟度颜色判别[J]. 食品与机械, 2016, 32(12): 133-136.
BI Z J, ZHANG R Y, QI Y J, et al. Tomato maturity color discrimination based on machine vision[J]. Food & Machinery, 2016, 32(12): 133-136.
- [7] VILLASEÑOR-AGUILAR M J, BOTELLO-ÁLVAREZ J E, PÉREZ-PINAL F J, et al. Fuzzy classification of the maturity of the tomato using a vision system[J]. Journal of Sensors, 2019, 2019: 3175848.
- [8] 王俊平, 徐刚. 机器视觉和电子鼻融合的番茄成熟度检测方法[J]. 食品与机械, 2022, 38(2): 148-152.
WANG J P, XU G. Research on tomato maturity detection method based on machine vision and electronic nose fusion[J]. Food & Machinery, 2022, 38(2): 148-152.
- [9] GUO D, LIU J, WANG X. On development of multi-resolution detector for tomato disease diagnosis[J]. Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology, 2021, 41(6): 6461-6471.
- [10] SU F, ZHAO Y P, WANG G H, et al. Tomato maturity classification based on SE-YOLOv3-MobileNetV1 network under nature greenhouse environment[J]. Agronomy, 2022, 12(7): 1638.
- [11] WANG S, JIANG H P, YANG J X, et al. Lightweight tomato ripeness detection algorithm based on the improved RT-DETR[J]. Frontiers in Plant Science, 2024, 15: 1415297.
- [12] 中华全国供销合作社. 中华人民共和国供销合作行业标准: GH/T 1193-2021[S]. 北京: 中华全国供销合作社, 2021.
All China Supply and Marketing Cooperatives. Supply and Marketing Cooperation Industry Standard of the People's Republic of China: GH/T 1193-2021[S]. Beijing: All China Supply and Marketing Cooperatives, 2021
- [13] 王新, 唐灿, 朱建新, 等. 基于深度增强与特征抗噪的夜间串番茄成熟度识别方法[J]. 农业机械学报, 2025, 56(4): 363-374.
WANG X, TANG C, ZHU J X, et al. Maturity identification method of night string tomatoes based on depth enhancement and feature noise suppression[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(4): 363-374.
- [14] SA I, GE Z Y, DAYOUB F, et al. DeepFruits: A fruit detection system using deep neural networks[J]. Sensors, 2016, 16(8): 1222.
- [15] MUREŞAN H, OLTEAN M. Fruit recognition from images using deep learning[J]. Acta Universitatis Sapientiae, Informatica, 2018, 10(1): 26-42.
- [16] BRAHIMI M, BOUKHALFA K, MOUSSAOUI A. Deep learning for tomato diseases: Classification and symptoms visualization[J]. Applied Artificial Intelligence, 2017, 31(4): 299-315.
- [17] GUO M H, LU C Z, HOU Q B, et al. SegNeXt: Rethinking convolutional attention design for semantic segmentation[EB/OL]. arXiv: 2209.08575, 2022.
- [18] JIN X, XIE Y P, WEI X S, et al. Delving deep into spatial pooling for squeeze-and-excitation networks[J]. Pattern Recognition, 2022, 121: 108159.
- [19] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]// Proceedings of the 15th European Conference on Computer Vision (ECCV 2018). Munich, Germany: Springer, 2018: 3-19.
- [20] CHEN H, CHEN K, DING G G, et al. YOLOv10: Real-time end-to-end object detection[C]// The Thirty-Eighth Annual Conference on Neural Information Processing Systems (NeurIPS 2024). Vancouver, BC, Canada: Neural Information Processing Systems Foundation, Inc. (NeurIPS), 2024: 107984-108011.
- [21] KC K, YIN Z D, WU M Y, et al. Depthwise separable convolution architectures for plant disease classification[J]. Computers and Electronics in Agriculture, 2019, 165: 104948.
- [22] JIANG Y A, WANG S Q, VALLS V, et al. Model pruning enables efficient federated learning on edge devices[J]. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34(12): 10374-10386.
- [23] NOTOMI T, MORI Y, TOMITA N, et al. Loop-mediated isothermal amplification (LAMP): Principle, features, and future prospects[J]. Journal of Microbiology, 2015, 53(1): 1-5.
- [24] JOHNSON O V, XINYING C, KHAW K W, et al. Ps-CALR: Periodic-shift cosine annealing learning rate for deep neural networks[J]. IEEE Access, 2023, 11: 139171-139186.
- [25] XIE Z R, LIU W T, LI Y H, et al. Enhanced litchi fruit detection and segmentation method integrating hyperspectral reconstruction and YOLOv8[J]. Computers and Electronics in Agriculture, 2025, 237: 110659.
- [26] YOU S C, LI B H, CHEN Y J, et al. Rose-Mamba-YOLO: An enhanced framework for efficient and accurate greenhouse rose monitoring[J]. Frontiers in Plant Science, 2025, 16: 1607582.
- [27] 姚晓通, 曲绍业. 基于改进YOLOv12s的辣椒叶片病虫害轻量化检测方法[J]. 智慧农业(中英文), 2026, 8(1): 1-14.
YAO X T, QU S Y. Lightweight detection method for pepper leaf diseases and pests based on improved YOLOv12s[J]. Smart Agriculture, 2026, 8(1): 1-14.

An Improved YOLOv10-Based Tomato Ripeness Detection Algorithm with LAMP Channel Pruning

ZHAO Licheng^{1,2}, LU Xinyu², WU Qian², REN Ni², ZHOU Lingli²,
CHENG Yawen², HU Anqi², QI Chao^{2*}

(1. School of Chemical Engineering, Huaiyin Institute of Technology, Huai'an 223003, China;

2. Institute of Agricultural Information, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China)

Abstract: [Objective] As a major crop in protected horticulture, cluster tomatoes grow in clusters with dense overlapping fruits. In greenhouse environments, light conditions are complex and variable, and the fruit color transitions continuously from green to red across different ripening stages, showing continuous gradation characteristics. These factors result in the low efficiency and strong subjectivity of traditional manual recognition methods. Meanwhile, deep learning-based detection models often suffer from decreased

detection accuracy, large localization errors, and slow inference speed when facing complex backgrounds and color interference, making it difficult to meet the dual requirements of real-time performance and high precision in practical applications. Therefore, to meet the practical application requirements of high accuracy, high real-time performance, and strong robustness for cluster tomato ripeness detection, this paper proposes a lightweight target detection model for cluster tomato ripeness, namely LampCT-YOLO (Cluster Tomato YOLO with LAMP pruning), which is based on improved YOLOv10. Through structural optimization and lightweight transformation of the baseline model, the detection accuracy, inference speed, and robustness are effectively improved, providing a novel technical solution for cluster tomato ripeness detection. **[Methods]** Taking YOLOv10 as the baseline model, first, the issue of insufficient feature extraction capability in complex scenarios was addressed by introducing the SegNeXt attention mechanism into the backbone network. By adaptively adjusting attention weights and calculating the correlation matrix between different feature channels, the mechanism automatically identified color channels strongly associated with the three ripeness levels of cluster tomatoes and assigned them higher attention weights, while suppressing feature responses from irrelevant background channels such as greenhouse frames, soil, and irrigation pipes. To achieve lightweight deployment of the model and meet the real-time detection requirements of edge devices, a gradient-based global channel importance method—LAMP channel pruning technology—was introduced after model training. The core principle of this technology was to evaluate the contribution of each channel to the model's detection performance by calculating the gradient magnitude of channels in each network layer, thereby eliminating redundant channels. This significantly reduced the model size and computational complexity while effectively maintaining the model's high detection performance for the three-category ripeness classification of cluster tomatoes. **[Results and Discussions]** Experiments showed that under the environment of NVIDIA A100 graphics card, for 240 cluster tomato images in the test set, the LampCT-YOLO model exhibited excellent detection performance. The mean average precision at 50 intersection over union (mAP50) for the early ripe, mid-ripe, and late ripe stages of cluster tomatoes was 84.6%, 89.5%, and 88.4%, respectively, which represented increases of 5.5, 7.7, and 0.9 percentage points compared with YOLOv10. The average mAP50 for the three ripeness categories of cluster tomatoes reached 87.6%, a 4.7 percentage points improvement over YOLOv10, demonstrating outstanding performance in both detection accuracy and stability. In addition, the model was found to maintain high recognition accuracy when facing variations in light intensity, fruit occlusion ratio, and background complexity, indicating good robustness and environmental adaptability. Regarding the lightweight effect, after applying the LAMP channel pruning technology, the number of model parameters and computational complexity were reduced by 63.07% and 50.06%, respectively, while the inference speed was improved by 23.1%. This effectively met the requirements of edge computing devices for real-time detection and low power consumption, alleviating the trade-off between model accuracy and inference speed. To verify the practical application value of the LampCT-YOLO model, the model was deployed on a self-developed fruit and vegetable inspection robot, which conducted field tests on 456 clusters of tomatoes in a real greenhouse environment. The results showed that the inspection robot successfully identified 78, 61, and 248 clusters of early ripe, mid-ripe, and late ripe cluster tomatoes, respectively, with detection accuracies of 84.8%, 87.1%, and 84.4%, and an average accuracy of 85.4%. Meanwhile, there were 5, 7, and 10 false detections, as well as 9, 2, and 36 missed detections for the early ripe, mid-ripe, and late ripe stages respectively, which to a certain extent reflected the practical application potential of the model. **[Conclusions]** The optimized LampCT-YOLO model not only significantly improves the recognition accuracy of cluster tomatoes at different ripening stages but also greatly reduces the model complexity, successfully achieving efficient deployment in resource-constrained scenarios. This model effectively balances the dual requirements of detection accuracy and real-time performance for inspection robots, and further constructs a reusable technical framework for the ripeness detection of protected horticultural fruits and vegetables. It provides strong support for the transformation of protected agriculture from labor-intensive to technology-intensive, and injects key innovative impetus into the large-scale and diversified implementation of smart agriculture.

Keywords: cluster tomato; ripeness detection; attention mechanism; channel pruning; fruit and vegetable inspection robot; YOLOv10

Foundation items: National Natural Science Foundation of China (32201664); Jiangsu Provincial Agricultural Science and Technology Independent Innovation Fund (CX(24)1021); Special Pilot Program for the Integrated R&D, Manufacturing, Promotion and Application of Agricultural Machinery (JSYTH08)

First author: ZHAO Licheng, E-mail: zhao_orange@163.com

***Corresponding author:** QI Chao, E-mail: chaoqi@jaas.ac.cn