

基于 MEW - YOLO 的库尔勒香梨目标检测方法

姜彦武^{1,2} 王小谅¹ 陈军¹ 胡广锐³

(1. 西北农林科技大学机械与工程学院, 陕西杨凌 712100; 2. 新疆农业大学机电工程学院, 乌鲁木齐 810052;
3. 西安工业大学设计学院, 西安 710021)

摘要: 库尔勒香梨果实检测是实现智能化低损采摘的关键环节,其检测精度与实时性直接影响采摘机器人的作业效率与稳定性。然而,受制于果实尺寸差异、光照变化、枝叶遮挡及复杂背景等因素,现有检测方法在实际应用中仍面临挑战。为提高库尔勒香梨的识别准确性与作业适应性,本文提出一种基于 YOLO 11n 的轻量化改进检测模型 MEW - YOLO,通过引入混合局部通道注意力模块 (Mixed local channel attention, MLCA)、高效上卷积模块 (Efficient up-convolution block, EUCB) 以及加权交并比损失函数 (Wise intersection over union, WIoU),分别从特征增强、空间细节恢复与边界回归优化 3 个层面提升模型对复杂场景的适应性。其中,MLCA 融合局部与全局通道信息以增强多尺度特征表达;EUCB 在特征融合阶段强化空间信息重建以提升小目标与遮挡目标表征;WIoU 通过样本质量感知加权与聚焦调节优化回归梯度分配,增强遮挡与重叠条件下的定位鲁棒性。在采集的库尔勒香梨自然场景数据集上,MEW - YOLO 相较基准 YOLO 11n 的 mAP@0.5、mAP@0.5:0.95、精确率与召回率分别提升 2.7、3.3、5.3、3.7 个百分点;参数量仅增加约 3.4%,浮点运算量为 6.8×10^9 。研究结果表明,MEW - YOLO 可为库尔勒香梨自然果园场景下的自动化检测与后续采摘作业提供可靠的视觉输入,并为特色水果目标检测模型的轻量化设计提供参考。

关键词: 库尔勒香梨; 智能化采摘; 目标检测; YOLO 11n; 识别方法

中图分类号: S23; TP391.4

文献标识码: A

文章编号: 1000-1298(2026)10-0239-12

OSID:



Object Detection Method for Korla Fragrant Pears Based on MEW - YOLO

JIANG Yanwu^{1,2} WANG Xiaoliang¹ CHEN Jun¹ HU Guangrui³

(1. College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China

2. College of Mechanical and Electrical Engineering, Xinjiang Agricultural University, Urumqi 830052, China

3. School of Design, Xi'an Technological University, Xi'an 710021, China)

Abstract: Accurate detection of Korla fragrant pear fruit is a key prerequisite for intelligent low-damage harvesting, as detection accuracy and real-time performance directly affect the operational efficiency and stability of harvesting robots. However, existing methods still face challenges in practical orchard applications due to variations in fruit size, illumination changes, occlusion by leaves and branches, and complex backgrounds. To improve detection accuracy and operational adaptability for Korla fragrant pear, a lightweight improved detection model, termed MEW - YOLO, was proposed based on YOLO 11n. MEW - YOLO introduced a mixed local channel attention (MLCA) module, an efficient up-convolution block (EUCB), and a wise intersection over union (WIoU) loss function, thereby enhancing model adaptability to complex scenes from three aspects: feature enhancement, spatial detail restoration, and bounding-box regression optimization. Specifically, MLCA integrated local and global channel information to strengthen multi-scale feature representation; EUCB reinforced spatial information reconstruction during feature fusion to improve the representation of small and occluded targets; and WIoU optimized the allocation of regression gradients through sample-quality-aware weighting and focusing regulation, enhancing localization robustness under occlusion and overlap conditions. On the collected Korla fragrant

收稿日期: 2025 - 11 - 14 修回日期: 2025 - 12 - 30

基金项目: 国家自然科学基金项目 (32272001) 和陕西省咸阳市“科学家 + 工程师”队伍项目 (L2024 - CXNL - KJRCTD - DWJS - 0002)

作者简介: 姜彦武 (1989—), 男, 博士生, 新疆农业大学副教授, 主要从事林果智能化采收研究, E-mail: jywxd@nwfau.edu.cn

通信作者: 陈军 (1970—), 男, 教授, 博士生导师, 主要从事林果智能化采收与果园智能作业装备研究, E-mail: chenjun_jdxy@nwsuaf.edu.cn

pear dataset captured under natural orchard conditions, MEW – YOLO outperformed the baseline YOLO 11n by 2.7, 3.3, 5.3, and 3.7 percentage points in mAP@ 0.5, mAP@ 0.5:0.95, precision, and recall, respectively; it used 2.67×10^6 parameters and required 6.8×10^9 of computation. The results indicated that MEW – YOLO can provide reliable visual input for automated detection and subsequent harvesting operations in natural Korla fragrant pear orchards, and offered a reference for the lightweight design of detection models for specialty fruit targets.

Key words: Korla fragrant pear; intelligent harvesting; object detection; YOLO 11n; recognition method

0 引言

在果实采收作业中,目标识别与定位精度直接影响采摘效率与果品完整度^[1]。库尔勒香梨作为中国西部干旱区的代表性经济果种,以其香甜细腻口感、丰富营养和较高商品价值在国内外市场享有盛誉^[2]。目前,香梨采收主要依赖人工完成,劳动强度大、效率低,难以满足现代果园规模化、标准化生产需求。同时,库尔勒香梨的自然果园环境复杂,果实常被枝叶遮挡,果实颜色与背景相近,给果实检测带来了较大挑战^[3-4]。在劳动力成本攀升与供给不足的背景下,面向智能化、低损采摘的自动感知技术需求日益迫切,而果实目标检测作为采摘机器人感知与定位的关键环节,其性能直接决定了采摘作业的准确性与可靠性。

果实目标检测技术经历了从人工特征到数据驱动的演进,早期方法多基于颜色、形态和纹理等显式特征,通过阈值分割或边缘检测实现目标分离,但在自然果园中易受光照变化和背景干扰影响,鲁棒性不足^[5]。随后引入支持向量机、随机森林等机器学习方法在特定场景下提升了识别效果,但仍依赖人工特征提取,难以适应果园环境多变性^[6-7]。近年来,卷积神经网络等深度学习技术凭借特征自学习能力显著提升了农业视觉检测性能^[8]。Faster R – CNN、SSD 和 YOLO 系列,已广泛应用于苹果、番茄、葡萄、猕猴桃等果实的检测任务^[9-12]。其中,YOLO 系列凭借其端到端的训练方式、单阶段结构和良好的实时性表现,在面向作业场景部署时具有较强应用潜力^[13]。然而,针对库尔勒香梨的研究仍相对有限;在小目标、遮挡与背景相似等复杂条件下,现有模型仍面临精度与轻量化难以兼顾的问题^[14]。因此,如何在保证检测精度的同时兼顾实时性和模型轻量化,成为当前库尔勒香梨果实检测研究的关键问题。

基于上述问题,在 YOLO 11n 架构基础上,本文提出一种适用于库尔勒香梨的轻量化改进型检测方法。通过引入混合局部通道注意力模块 (Mixed local channel attention, MLCA)、高效上卷积模块 (Efficient up-convolution block, EUCB) 及加权交并

比损失函数 (Wise intersection over union, WIoU), 以提升模型复杂环境下的检测精度与鲁棒性基础上,保持较高的计算效率与轻量化特性;并通过采集自然环境下果园数据集对模型进行验证,以期智能化低损采摘提供技术支撑。

1 模型架构和改进

1.1 YOLO 11n 基准架构

YOLO 系列推动单阶段检测在精度与实时性之间的协同发展,研究重点逐步转向特征增强、轻量化与结构自适应^[15]。YOLO 11 通过更高效特征交互与多尺度信息融合机制提升复杂背景下小目标感知能力,并提供可调的网络深度与宽度设置以适配不同算力平台^[16]。选用 YOLO 11n 作为基础模型,主要基于其在检测精度、推理速度与资源消耗之间取得较优平衡,适用于端侧/嵌入式平台部署,并能在光照变化、枝叶遮挡与果实重叠等果园复杂条件下保持较好的检测鲁棒性^[17-18]。

YOLO 11n 的网络架构主要由输入层 (Input)、主干网络 (Backbone)、特征颈部 (Neck) 与检测头 (Head) 4 个核心部分构成^[19]。输入层负责图像归一化与尺寸标准化,为网络提供分布一致的数据输入。主干网络作为特征提取核心,以标准卷积层起始,通过多层 C3k2 模块实现浅层特征提取与深层语义表达的协同优化;在末端引入快速空间金字塔池化模块 (Spatial pyramid pooling-fast, SPPF), 通过多尺度池化扩展感受野,以捕获全局上下文信息。特征颈部采用路径聚合网络 (Path aggregation network, PANet) 结构,通过上采样与特征拼接 (Concatenation, Concat) 实现深浅层特征融合。检测头部分采用解耦式设计,将分类与边界框回归任务分离处理,以缓解任务竞争、提升定位精度并加快模型收敛^[20]。整体而言,YOLO 11n 在检测精度与推理速度之间取得了良好平衡,为后续结构改进提供了良好的扩展基础。

1.2 模型优化与自适应模块设计

YOLO 11n 在处理库尔勒香梨特有的复杂视觉特征时,仍存在检测精度和边界回归稳定性受限问题。为提升模型在香梨果园环境下的检测性

能,在其基础上进行了结构优化与模块改进,提出了改进型检测模型 MEW-YOLO,引入 3 项关键改进:①混合局部通道注意力模块 (MLCA),用于增强特征聚焦与显著性响应能力。②高效上卷积模块 (EUCB),通过轻量化上采样与深度可分离卷积实现空间细节重建,提升小目标与遮挡目标的检测鲁棒性。③加权交并比损失函数 (WIoU),以强化边界回归精度与模型收敛稳定性。三者特征提取、空间重建与损失优化层面形成协同作用,使模型在检测精度、鲁棒性与实时性方面实现全面提升。

不同于传统的单模块增强方式,研究将多层注意力机制与卷积特征融合结构相结合,构建面向复

杂农业视觉场景的自适应优化框架。在该框架下,网络在不同层级间实现了语义与空间特征的动态交互,并通过多尺度通道融合机制提升了目标区域的显著性响应。同时,在特征融合路径中嵌入的上卷积与注意力协同结构,有助于重建被遮挡或细粒度目标的空间信息,改善检测边界的连续性与定位稳定性。为进一步提升模型的训练收敛效率与预测精度,在损失函数设计中引入了加权交并比机制,以自适应权重调节方式平衡正负样本贡献,从而强化模型的边界回归鲁棒性。该优化策略实现了从局部特征增强到全局信息协调的结构优化,为 MEW-YOLO 模型提供了更强的自适应检测能力与泛化性能。改进整体网络结构如图 1 所示。

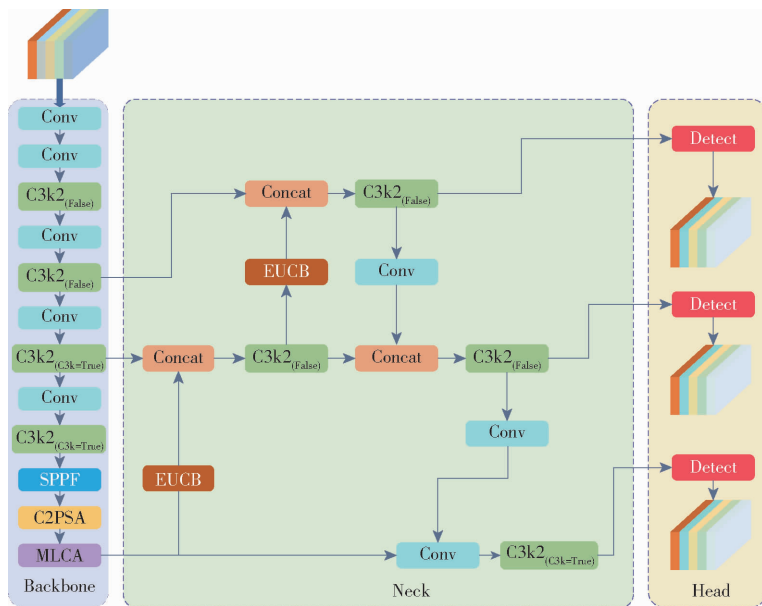


图 1 基于 YOLO 11n 的改进网络结构示意图

Fig. 1 Architecture of improved YOLO 11n-based network

1.2.1 混合局部通道注意力 (MLCA)

在自然采摘场景中,香梨目标易受枝叶遮挡、光照变化及背景纹理干扰,导致特征提取阶段的局部显著信息被稀释,从而影响小目标与遮挡目标的识别稳定性。现有通道注意力机制(如 SENet、ECA-Net)多基于全局平均池化(Global average pooling, GAP)构建通道描述,能够刻画通道间整体依赖,但全局汇聚会弱化局部差异,使模型对局部显著区域的响应不足。为兼顾全局统计与局部细节,引入混合局部通道注意力模块 (MLCA),在 GAP 基础上并行引入局部平均池化(Local average pooling, LAP),通过融合全局与局部通道统计信息实现通道权重标定。

设输入特征映射由 C 个通道组成,表示为

$$X = \{x_c \in \mathbf{R}^{H \times W} | c = 1, 2, \dots, C\} \quad (1)$$

式中 H, W ——特征图高与宽

x_c ——第 c 个通道二维特征图

GAP 对每个通道计算全局均值作为通道统计描述,即

$$z_c^{GAP} = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j) \quad (2)$$

式中 $x_c(i, j)$ ——空间位置 (i, j) 处特征响应 ($i = 1, 2, \dots, H; j = 1, 2, \dots, W$)

为增强局部感知, LAP 采用固定尺度的局部汇聚:将特征图在空间维度上划分为 $k_s \times k_s$ 个子区域, $k_s \in \{3, 5, 7\}$, 定义第 (i, j) 个子区域对应的像素索引集合为 $\Omega_{k_s}(i, j)$, 其像素点数为 $|\Omega_{k_s}(i, j)|$ 。LAP 在该子区域内计算局部均值,即

$$z_c^{LAP}(i, j) = \frac{1}{|\Omega_{k_s}(i, j)|} \sum_{(u, v) \in \Omega_{k_s}(i, j)} x_c(u, v) \quad (3)$$

式中 (u, v) ——子区域内空间索引

由式(3)可知, LAP 对每个通道输出一个 $k_s \times k_s$ 的局

部统计图,因此整体 LAP 输出尺寸为 (C, k_s, k_s) 。

为融合全局与局部信息,MLCA 模块对 GAP 与 LAP 的结果进行拼接

$$\mathbf{Z}_c = (z_c^{GAP}, z_c^{LAP}) \tag{4}$$

式中, $\mathbf{Z}_c$ 为第 c 个通道融合后特征描述向量,由全局平均池化输出 z_c^{GAP} 与局部平均池化输出 z_c^{LAP} 拼接形成,用于同时表征通道的全局与局部响应信息,得到综合通道描述矩阵 $\mathbf{Z} \in \mathbf{R}^{C \times k_s}$, $\mathbf{Z}$ 为所有通道的特征描述向量 $\mathbf{Z}_c$ 构成,反映整体通道特征分布。

随后,采用一维卷积算子对通道间的依赖关系进行建模,即

$$\hat{\mathbf{z}}_c = \sigma(\text{Conv1D}(\mathbf{Z}_c, \mathbf{W})) \tag{5}$$

式中 σ ——sigmoid 激活函数

$\mathbf{W}$ ——可学习卷积核参数

Conv1D——在通道维度上进行的一维卷积操作,用于捕获相邻通道间的特征依赖关系

$\hat{\mathbf{z}}_c$ ——经卷积加权后的通道注意力向量,反映通道重要性权重

得到的通道注意力权重向量 $\hat{\mathbf{z}}_c$ 经维度恢复 (UNAP) 后,与输入特征 X_c 逐通道相乘,实现特征重

标定,计算式为

$$Y_c = \hat{\mathbf{z}}_c \otimes X_c \tag{6}$$

式中 $\otimes$ ——通道加权操作

Y_c ——加权后第 c 通道输出特征

输出特征 $Y \in \mathbf{R}^{C \times H \times W}$ 作为增强后的特征映射输入至后续检测层,用于优化香梨在复杂环境下的检测性能。

通过将 LAP 的局部感知能力与 GAP 的全局统计信息相结合,MLCA 模块能够有效提升模型在复杂场景下的特征选择性与语义表达能力,其整体结构如图 2 所示^[21]。MLCA 采用双分支并行架构以协同建模全局语义依赖与局部空间细节:全局分支首先对输入特征进行 GAP,得到通道级全局描述,并在通道维度上进行的一维卷积操作 (Conv1D) 学习相邻通道依赖,生成全局通道权重;局部分支对输入特征执行 LAP,将空间维汇聚为 $k_s \times k_s$ 的局部统计图 ($k_s \in \{3, 5, 7\}$),随后经 Reshape 与通道维 Conv1D 建模,在局部尺度上获得通道权重图。两分支输出在 (C, k_s, k_s) 尺度上进行融合,并经 UNAP 恢复至 (C, H, W) 后与输入特征逐通道相乘,实现对通道响应的自适应重标定,从而增强目标相关通道响应并抑制背景干扰。

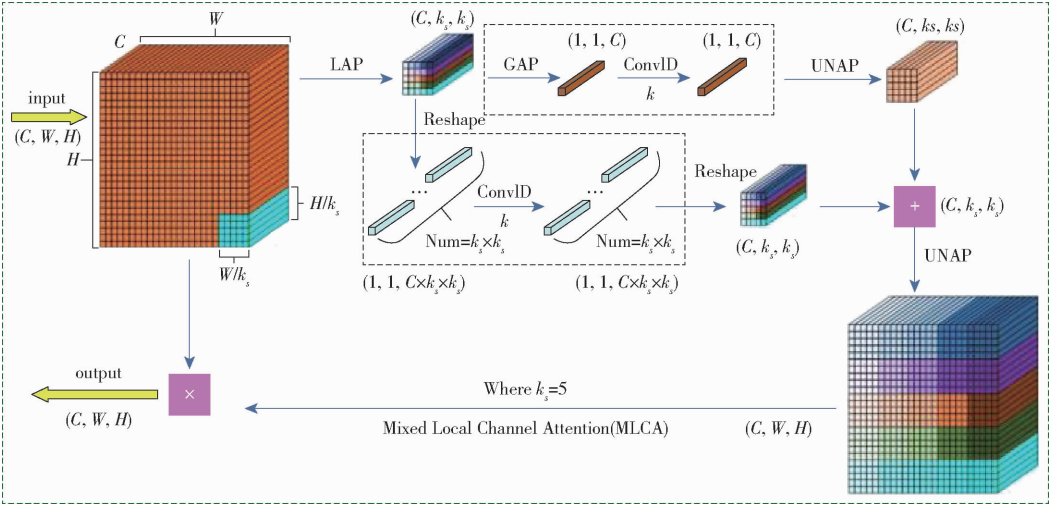


图 2 混合局部通道注意力模块 (MLCA) 结构示意图

Fig. 2 Architecture of mixed local channel attention (MLCA) module

与传统仅依赖全局统计的通道注意力机制相比,MLCA 通过可调局部尺度 k_s 引入局部统计信息,实现多尺度局部感知,使模型在目标尺度变化及部分遮挡条件下仍能保持对关键区域的稳定响应;同时,其核心计算采用通道一维卷积对通道间依赖进行轻量化建模,在参数量与浮点运算量增幅可控的前提下嵌入 YOLO 11n 并保持推理效率;此外,模块以全局分支与局部分支协同融合的方式兼顾全局语义与局部细节,有助于提升复杂光照与枝叶遮挡等自然条件下的特征表达鲁棒性。基于良好的模块化

特性,MLCA 可按网络结构需求部署于 Backbone 的特征融合阶段,以提升特征表达质量并在检测精度与实时性之间取得合理权衡。

1.2.2 高效上卷积 (EUCB)

在目标检测任务中,高层语义特征在上采样与融合过程中易出现空间分辨率不足与边界细节弱化,尤其在果实部分遮挡或光照变化显著的场景下,易导致目标轮廓与局部纹理表征不稳定,进而影响检测头的定位回归精度。为改善上采样阶段的细节恢复能力,在 YOLO 11n 的特征融合阶段引入高效

上卷积模块 (EUCB), 在参数量与计算量增量较小的前提下提升特征重建与融合质量。

EUCB 结构如图 3 所示, 由上采样 (Upsample $2\times$)、深度卷积 (DWC 3×3)、批归一化 (BN)、非线性激活函数 ReLU 以及压缩卷积 (Conv 1×1) 组成^[22]。模块计算过程可表示为

$$F_{out} = \text{Conv}_{1\times 1}(\text{ReLU}(\text{BN}(\text{DWC}_{3\times 3}(\text{Up}(F_{in})))) \quad (7)$$

式中 F_{in} 、 F_{out} ——模块输入与输出特征
 Up ——上采样操作, 采用双线性插值方法, 将空间尺寸由 $H\times W$ 扩展至 $2H\times 2W$, 以恢复高层语义特征的空间细节
 $\text{DWC}_{3\times 3}$ ——深度可分离卷积, 用以极低参数量提取局部结构与边缘特征
 BN ——归一化操作, 用于稳定特征分布、抑制梯度震荡并加速网络收敛
 ReLU ——激活函数, 用于提升模型的非线性表达能力与特征稀疏性
 $\text{Conv}_{1\times 1}$ ——压缩卷积层, 通过通道融合与维度重构减轻上采样带来的通道冗余问题

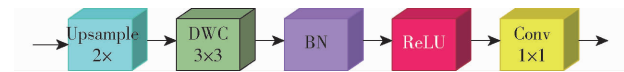


图 3 高效上卷积模块 (EUCB) 结构示意图

Fig. 3 Schematic of efficient up-convolution block (EUCB)

从特征传递路径看, EUCB 先通过上采样恢复空间分辨率, 再以深度卷积补充局部细节; 经 BN 与 ReLU 稳定并增强特征表征; 最终通过 1×1 卷积完成通道压缩与融合, 有助于提高上层特征的空间连续性与边缘表征能力, 从而为后续特征融合与定位回归提供更稳定的高分辨率语义特征输入。EUCB 与前述 MLCA 在作用维度上形成互补: 前者在空间维度上实现高分辨率语义特征的重建, 后者在通道维度上强化特征选择与加权表达。二者协同提高复杂背景条件下的特征质量与定位稳定性, 为模型在果园场景中的检测性能提升提供结构支撑。

1.2.3 加权交并比损失函数 (WIoU)

在目标检测任务中, 边界框回归损失函数直接影响模型的定位精度与训练稳定性。传统的 IoU 类损失虽能度量预测框与真实框的重叠程度, 但在小目标、部分遮挡或低质量样本占比较高的情况下, 回归梯度分配易失衡, 导致收敛效率下降或定位偏移加剧^[23]。针对库尔勒香梨自然果园环境中果实密集、遮挡重叠与形态差异显著等特点, 在回归项中引入 WIoU 损失, 以样本质量感知的动态加权与非单调聚焦机制对不同质量样本的回归贡献进行自适应

调节, 从而提升边界回归的一致性与训练过程的稳定性。

设预测框与真实框的交并比为 IoU, 则传统 IoU 损失定义为

$$L_{IoU} = 1 - IoU \quad (8)$$

WIoU 在 IoU 框架上引入动态权重 ω , 构造加权回归损失, 即

$$L_{WIoU} = \omega(1 - IoU) \quad (9)$$

其中 ω 由样本几何偏差与回归难度共同决定, 公式为

$$\omega = \exp\left(-\frac{d_{center}^2}{\delta^2}\right) f_{focus}(IoU) \quad (10)$$

式中 d_{center} ——预测框与真实框中心点的欧氏距离
 δ ——调节常数

用于控制几何权重随中心偏差增大而衰减的速率, 使中心偏差较大的样本贡献被抑制, 从而降低由严重偏移样本引入的梯度扰动。

为进一步实现动态非单调聚焦, WIoU 采用基于 IoU 分布的聚焦函数对样本贡献进行调节, 即

$$f_{focus}(IoU) = \alpha \exp(-\delta_f(IoU - \mu)^2) \quad (11)$$

式中 μ ——IoU 分布均值

δ_f ——控制曲线陡峭程度

α ——归一化系数

该函数呈近似高斯形态, 使中等质量样本获得更高梯度响应, 而极高或极低质量样本的权重均被抑制, 从而在强调有效学习样本的同时降低噪声样本对训练稳定性的影响。

在损失整合方面, 采用分类损失、目标存在性损失与 WIoU 回归损失的加权和作为总体优化目标, 即

$$L_{total} = \lambda_1 L_{cls} + \lambda_2 L_{dfl} + \lambda_3 L_{WIoU} \quad (12)$$

式中 λ_1 、 λ_2 、 λ_3 ——3 种损失的权重系数, 以平衡检测精度与边界框拟合能力

L_{cls} ——分类损失

L_{dfl} ——目标存在性预测损失

L_{WIoU} ——回归损失

为保证训练过程的可复现性与对比公平性, 本文在所有试验中固定采用 $\lambda_1 = 0.5$, $\lambda_2 = 1.0$, $\lambda_3 = 2.0$, 并保持其余训练超参数一致。

1.2.4 模块间协同机制

MEW-YOLO 引入混合局部通道注意力模块 (MLCA)、高效上卷积模块 (EUCB) 以及加权交并比损失函数 (WIoU), 形成特征增强、空间重建、回归优化的层级协同框架。其中 MLCA 用于增强多尺度特征的通道选择性与信息聚合, EUCB 用于提升特征融合阶段的空间细节恢复质量, WIoU 用于在回归学习

中对不同质量样本的梯度贡献进行自适应调节。

2 试验与分析

2.1 数据集构建与试验设置

2.1.1 数据集构建

为验证所提方法的有效性,构建了1个面向库尔勒香梨的专用目标检测数据集。数据采集在新疆巴音郭楞蒙古自治州库尔勒铁门关市园艺一连香梨标准化建设园中完成,果树种植模式为纺锤形,种植株行距为4 m×1.5 m,水肥管理条件一致,所选试验园总面积约2.67 hm²。采集时沿果树行连续拍摄,并对每行分别从正向与反向取景,以覆盖不同朝向与光照条件下的外观差异。拍摄时相机与果树距离为1~2 m,镜头正对果树主体,以模拟真实采摘场景。采集时间为2024年9月5—10日(采摘期),覆盖晴朗与微阴等典型自然光照;为降低风致摆动造成的模糊与噪声,拍摄选择静风/微风条件,并在拍摄前确保邻近果枝相对静止。采集设备为ZED型双目相机(1 920像素×1 080像素),在自然光下拍摄;采集过程中综合考虑顺/逆光、果实姿态(正面、侧面、部分遮挡)以及果实与枝叶背景的交互关系,以提高样本复杂度并用于评估模型泛化能力。香梨图像采集过程如图4所示。拍摄过程中使用三脚架固定相机并保持姿态与高度一致,以减少外界扰动引入的视觉偏差。



图4 库尔勒香梨图像采集过程

Fig.4 Image acquisition process of Korla fragrant pears

为确保数据采集条件具有一致性与可复现性,对光照与遮挡条件进行了操作性划分。光照条件依据相机光轴与主要入射光方向的相对关系进行判定:顺光指入射光主要来自相机后方,果面受光相对均匀、明暗反差较小;逆光指入射光主要来自相机前方,果面整体亮度下降且存在明显明暗反差。遮挡类别与标注一致:无遮挡(Pear)指果实轮廓与表面纹理信息完整可辨;遮挡(Occlusion)指果实被枝叶/枝条遮挡导致轮廓边界或纹理信息缺失。数据标注由单一标注员按统一准则完成,并通过两轮自

校核,对遮挡类别判定与边界框贴合度进行复核,以降低主观偏差、保证标注一致性。为避免同一果树或相邻区域样本同时出现在不同子集造成数据泄露,数据划分采用按拍摄行划分,即同一行内图像归入同一数据子集,从而保证训练、验证与测试集在空间上的相对独立。

最终获得2 000幅原始图像,其中顺光与逆光各1 000幅,采用LabelImg进行人工矩形框标注,类别分为2类:完整可见香梨与被遮挡香梨。数据按比例7:2:1划分为训练集、验证集与测试集,分别用于模型训练与性能评估。采集样本已包含光照方向变化、遮挡形态及果实密集分布等典型自然场景,具有较好的多样性,因此本文未引入自定义增强操作。此外,为公平比较不同结构改进模块的效果,各模型统一采用一致的数据处理流程与训练配置,尽量减少增强策略引入的混杂因素,使性能差异更能反映结构改进本身的贡献。

2.1.2 试验环境与参数配置

试验在Windows 11环境下完成。硬件平台为12核Intel Xeon Platinum 8255C(2.50 GHz)、NVIDIA GeForce RTX 2080 Ti、43 GB内存及SSD存储。软件环境采用PyTorch 2.2.2,基于CUDA 12.1与cuDNN进行GPU加速,Python版本为3.9.19。

为评估所提方法的检测性能与部署可行性,从检测准确性、综合检测性能与计算开销3个方面选取评价指标^[24-25]。在检测准确性方面,采用精确率(Precision, P)与召回率(Recall, R)衡量误检抑制与漏检控制能力。在综合性能评估方面,采用平均精度均值(mean Average Precision, mAP)作为核心衡量指标,采用mAP@0.5作为核心指标,并进一步采用mAP@0.5:0.95(IoU 阈值在[0.5, 0.95]内以0.05步长取均值)以更严格衡量不同定位精度要求下的检测表现。在效率与资源消耗方面,采用模型参数量与浮点运算量表征模型复杂度与计算开销,用于评估其在智能采摘场景中的实时性与轻量化适配能力。

为确定训练配置参数,对训练轮数、输入分辨率及优化器进行了对比试验,结果如表1所示。训练轮数由150提升至200(640像素×640像素,SGD)时,mAP@0.5与mAP@0.5:0.95均提升,说明增加训练轮数有助于提升检测性能。输入分辨率由640像素×640像素降至520像素×520像素(200轮,SGD)后,mAP下降,表明降低分辨率会削弱细节表征并影响检测效果。在相同轮数与分辨率(200轮,640像素×640像素)条件下,SGD的mAP@0.5和mAP@0.5:0.95均优于Adam。

表 1 不同训练配置下检测性能

Tab.1 Detection performance under different training configurations

参数			精确率/	召回率/	mAP@ 0. 5/	mAP@ 0. 5;0. 95/	参数量	浮点运算量
训练轮数	分辨率/(像素 × 像素)	优化器	%	%	%	%		
200	640 × 640	SGD	77. 5	79. 2	79. 5	64. 4	2. 67 × 10 ⁶	6. 8 × 10 ⁹
150	640 × 640	SGD	71. 2	73. 7	78. 9	60. 7	2. 67 × 10 ⁶	6. 8 × 10 ⁹
200	520 × 520	SGD	70. 8	73. 4	78. 0	59. 9	2. 67 × 10 ⁶	6. 8 × 10 ⁹
200	640 × 640	Adam	75. 2	77. 0	77. 8	62. 1	2. 67 × 10 ⁶	6. 8 × 10 ⁹

综合对比结果,采用 640 像素 × 640 像素、200 轮与 SGD 作为统一训练配置,初始学习率为 0. 01,并使用余弦退火进行衰减;动量与权重衰减分别设为 0. 937 和 0. 000 5。

2. 2 模型性能试验

为全面评估 MEW - YOLO 模型在库尔勒香梨

目标检测任务中的综合性能,选取当前主流的单阶段与双阶段检测模型,包括 SSD、Faster R - CNN、RT - DETR、DEIM、YOLO v8n、YOLO 12、YOLO 13 以及基准模型 YOLO 11n 进行对比试验,所有模型在同一数据集与试验条件下训练。结果如表 2 所示。

表 2 不同模型在库尔勒香梨检测任务中的性能

Tab.2 Performance of different models on Korla fragrant pear detection task

模型	精确率/%	召回率/%	mAP@ 0. 5/%	mAP@ 0. 5;0. 95/%	参数量	浮点运算量
SSD	84. 00	83. 90	23. 51	11. 53	2. 628 × 10 ⁷	6. 270 × 10 ¹⁰
Faster R - CNN	32. 10	54. 80	32. 62	18. 21	1. 371 × 10 ⁸	3. 702 × 10 ¹¹
RT - DETR	67. 10	71. 20	74. 39	58. 34	3. 200 × 10 ⁸	1. 034 × 10 ¹¹
DEIM	66. 90	70. 30	72. 13	56. 18	4. 001 × 10 ⁷	3. 000 × 10 ⁹
YOLO v8n	68. 40	76. 90	78. 30	60. 80	3. 006 × 10 ⁶	8. 100 × 10 ⁹
YOLO 12	69. 60	70. 50	73. 20	56. 00	2. 509 × 10 ⁶	5. 800 × 10 ⁹
YOLO 13	67. 80	75. 90	76. 70	58. 30	2. 448 × 10 ⁶	6. 100 × 10 ⁹
YOLO 11n	72. 20	75. 50	76. 80	61. 10	2. 582 × 10 ⁶	6. 300 × 10 ⁹
MEW - YOLO	77. 50	79. 20	79. 50	64. 40	2. 669 × 10 ⁶	6. 800 × 10 ⁹

2. 2. 1 总体性能提升分析

从表 2 可知,MEW - YOLO 在多个指标取得较大幅度提升。其中,mAP@ 0. 5 提升 2. 7 个百分点,mAP@ 0. 5;0. 95 提升 3. 3 个百分点,体现出模型在多尺度目标与复杂背景条件下的检测性能与定位一致性得到增强。召回率提升 3. 7 个百分点,说明其在枝叶遮挡、光照不均等自然场景中对果实目标的检出更为充分,精确率提高 5. 3 个百分点,表明模型在抑制误检倾向的同时提升了目标检出完整性。与此同时,模型参数量仅增加约 3. 4%,浮点运算量为 6. 8 × 10⁹,在精度提升与计算代价之间保持了较好的平衡。

2. 2. 2 主流检测模型性能对比

两阶段模型 Faster R - CNN 的 mAP@ 0. 5 较低,主要与其基于候选区域的两阶段处理流程有关。在果实密集分布或枝叶遮挡较多场景中,候选框生成与前景/背景判别更易受干扰,导致误检与漏检风险增加;同时其计算复杂度较高,推理开销较大,不利于农业现场嵌入式平台实时部署。

单阶段轻量模型中,SSD 的精确率、召回率较高但 mAP@ 0. 5 明显偏低,主要与其多尺度特征表达

与边界回归能力对复杂背景和小目标场景的适配不足有关:在香梨数据集中存在小目标占比高、遮挡与重叠频繁等情况时,预测框排序与定位一致性下降,导致 PR 曲线面积较小,从而出现精确率、召回率不低但 mAP 偏低的现象。该结果体现了模型结构与任务场景差异带来的性能局限。YOLO 12 参数与计算量较低,但在更严格阈值下的定位指标仍低于基准与改进模型,说明其边界拟合精度与重叠目标区分能力仍有提升空间。

Transformer 检测器 RT - DETR 与 DEIM 具有全局建模优势,但在本任务中未体现出对轻量 YOLO 框架的综合优势:RT - DETR 参数量大且计算量较高,推理代价不利于实时应用;DEIM 计算量较低,但检测精度与基准模型接近。MEW - YOLO 在保持轻量结构的前提下取得更高 mAP@ 0. 5。

YOLO 系列模型中,YOLO v8n 的 mAP@ 0. 5 较高,但其 mAP@ 0. 5;0. 95 低于改进模型,表明在更严格定位要求下仍存在边界一致性差距;YOLO 13 的 mAP@ 0. 5 与基准接近,但 mAP@ 0. 5;0. 95 仍低于改进模型,反映其回归精度与重叠场景区分能力仍受限。MEW - YOLO 在保持轻量化参数规

模与较低计算复杂度的前提下,实现了更优综合检测性能。

2.2.3 性能协同与实用性评价

MEW - YOLO 性能提升源于特征增强、空间重建与回归优化三者协同作用的结果,使模型在特征表达、空间解析度及边界精度等方面得到系统性提升,从而增强在复杂果园环境中的检测稳定性与泛化能力。综合而言,MEW - YOLO 在精度提升的同时维持轻量化参数与计算量水平,具备在农业自动采摘等资源受限平台部署的可行性。为进一步明确各模块在性能提升中的具体作用机制,开展了系统

的消融试验分析,以验证模块协同的实际效果与机制合理性。

2.3 消融试验

为验证所提出的混合局部通道注意力模块、高效上卷积模块及加权交并比损失函数在提升库尔勒香梨检测性能中的实际作用,基于 YOLO 11n 框架进行了系统的消融试验。试验采用逐步叠加与组合对比的方式,分析各模块对检测精度、定位一致性及模型复杂度的影响。所有试验在相同训练环境与超参数设置下进行,仅改变模块配置,其余训练与评估流程保持一致。结果如表 3 所示。

表 3 消融试验结果
Tab.3 Results of ablation study

模型	精确率/%	召回率/%	mAP@0.5/%	mAP@0.5;0.9/%	参数量	浮点运算量
YOLO 11n	72.2	75.5	76.8	61.1	2.58×10^6	6.3×10^9
YOLO 11n - EUCB	74.8	76.4	78.3	61.4	2.67×10^6	6.8×10^9
YOLO 11n - MLCA	75.6	76.9	78.5	62.3	2.58×10^6	6.4×10^9
YOLO 11n - WIoU	73.4	77.0	77.5	61.2	2.58×10^6	6.3×10^9
YOLO 11n - EUCB - MLCA	76.8	78.2	78.7	63.7	2.67×10^6	6.8×10^9
YOLO 11n - EUCB - WIoU	68.8	78.1	79.2	61.8	2.67×10^6	6.8×10^9
YOLO 11n - MLCA - WIoU	76.0	78.7	79.0	63.1	2.58×10^6	6.4×10^9
YOLO 11n - MLCA - EUCB - WIoU	77.5	79.2	79.5	64.4	2.67×10^6	6.8×10^9

2.3.1 单模块性能

从表 3 可知,3 种改进在单独引入时均带来不同程度的性能增益。EUCB 在 mAP@0.5 提升较为明显,说明其通过上采样与深度卷积的组合能够改善特征融合阶段的空间细节恢复,从而有利于遮挡与小目标场景下的检出与定位。MLCA 对 mAP@0.5 与 mAP@0.5;0.95 分别提升 1.7、1.2 个百分点,表明引入局部-全局联合的通道加权后,模型在复杂背景与尺度变化条件下的特征选择性得到增强。WIoU 在保持网络结构不变的情况下带来召回率与 mAP 的小幅提升,说明其通过样本质量感知的动态加权机制能够在一定程度上改善回归学习的梯度分配,降低遮挡、重叠等低质量样本对训练稳定性的影响。

2.3.2 双模块组合效应

不同模块组合表现出互补性。EUCB 与 MLCA 联合使用时,mAP@0.5 与 mAP@0.5;0.95 相较基准模型分别提升 1.9、2.6 个百分点,表明空间特征重建与通道特征增强在网络中形成互补关系:前者强化空间结构连续性,后者提高通道间的特征选择性,共同提升特征表达能力。在 EUCB 与 WIoU 组合中,mAP@0.5 在双模块组合中表现最优,同时 mAP@0.5;0.95 相较基准模型仍有提升,但相较 EUCB + MLCA 与 MLCA + WIoU 等组合在高 IoU 阈

值下的增益相对有限,表明该组合对定位精细度仍有提升空间。MLCA 与 WIoU 的联合模型在多尺度表征与定位鲁棒性之间取得较好平衡,体现出在更严格 IoU 阈值条件下的边界拟合优势。

2.3.3 三模块协同效应

集成 3 种改进模块后的 MEW - YOLO 在综合性能上取得最优结果,相较 YOLO 11n,mAP@0.5 与 mAP@0.5;0.95 分别提升 2.7、3.3 个百分点,精确率、召回率亦同步提升,说明三者 in 整体性能上形成了有效互补。结合表 3 可见,三模块联合配置在主要指标上均获得稳定增益,表明其协同提升并非由单一模块主导,而是对检测任务中特征增强、空间细节重建、边界回归优化不同环节的共同强化。该结果从试验层面印证了前文关于模块协同机制的分析。

2.3.4 模块冗余性

从性能增益、计算复杂度与结构层面 3 个维度对所提改进模块的冗余性进行检验。由表 3 可知,单模块引入均带来稳定的精度增益:EUCB、MLCA 与 WIoU 的 mAP@0.5 稳定提升,说明三者均具有独立贡献。在两两组合配置下 EUCB + MLCA、EUCB + WIoU 与 MLCA + WIoU 的 mAP@0.5 相较基准模型进一步提升,未出现叠加无效或性能饱和不增的情况,表明模块间不存在明显功能

重复。需指出的是,不同组合在 $mAP@0.5$ 与 $mAP@0.5:0.95$ 上提升幅度存在差异,反映了各模块对检出能力与高 IoU 定位一致性的促进侧重点不同,但不改变其在组合配置下仍能形成有效互补的结论。在计算复杂度方面,引入 EUCB 会带来参数量与计算量的小幅增加,其余组合在轻量化水平上保持稳定;模型总体精度获得稳定提升,未出现明显成本增加而收益有限的冗余现象。结构层面上,MLCA 作用于特征提取阶段以增强通道选择性,EUCB 作用于特征融合阶段以改善空间细节恢复,WIoU 作用于回归优化阶段以调整样本对回归学习的贡献分配,分别对应不同环节,作用位置清晰、层次互补。

2.4 综合检测结果与稳定性

为系统验证 MEW - YOLO 模型在库尔勒香梨目标检测任务中的整体性能表现,从训练收敛性、检测结果可视化、PR 曲线及注意力热力图 4 个方面对

模型表现进行分析。

2.4.1 训练收敛性

为评估 MEW - YOLO 的训练收敛性与稳定性,对训练与验证阶段的关键指标(包括框回归损失 box_loss、分类损失 cls_loss、分布焦点损失 dfl_loss 等损失函数以及主要性能指标)进行监测,变化趋势如图 5(横坐标为训练轮次,纵坐标为损失值或指标值)所示。从图 5 可见,训练阶段 box_loss、cls_loss 与 dfl_loss 在训练初期快速下降,随后逐步趋于平稳;验证损失与训练损失曲线变化趋势一致,未出现明显发散或反向上升,表明训练过程收敛稳定且过拟合迹象不显著。随训练推进,精确率、召回率稳步提升并在 80 ~ 100 轮次后趋于收敛, $mAP@0.5$ 与 $mAP@0.5:0.95$ 同步上升并分别收敛至约 0.80 与 0.64。整体曲线变化平滑、波动幅度较小,表明模型在当前设置下具有良好的可优化性与训练稳定性。

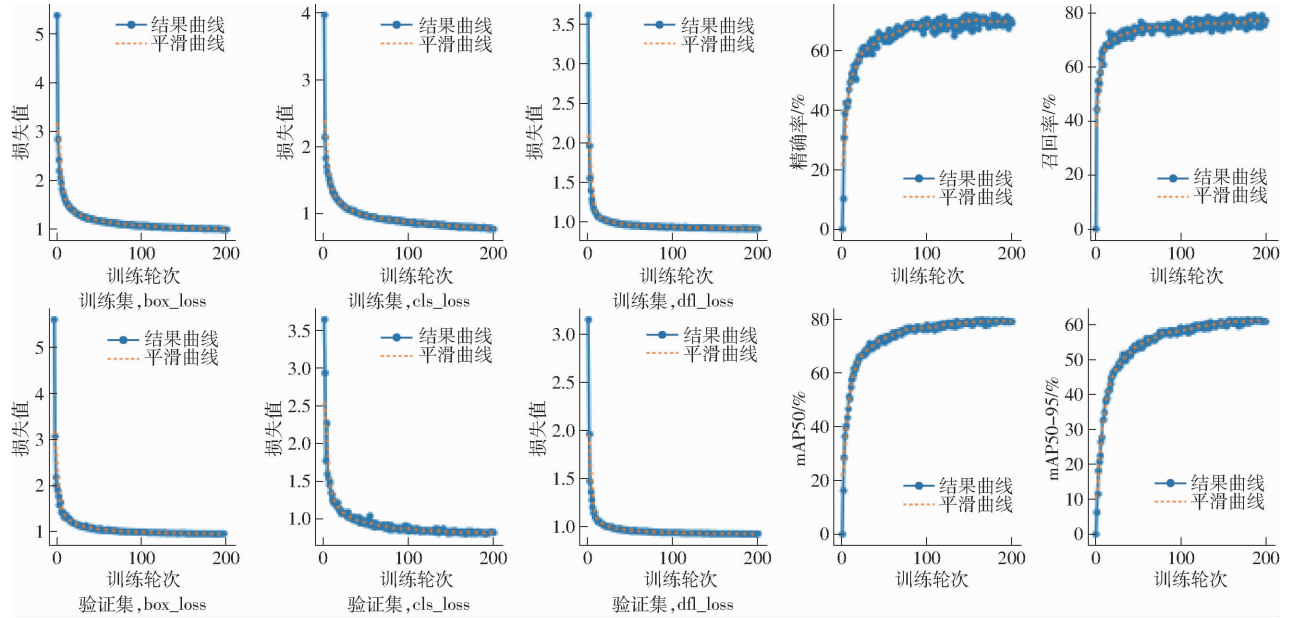


图 5 MEW - YOLO 训练与验证阶段损失与性能指标变化曲线

Fig. 5 Curves of training losses and evaluation metrics for MEW - YOLO during training

2.4.2 检测结果可视化与稳定性

为直观展示 MEW - YOLO 在复杂果园环境下的检测表现,选取顺光、逆光、枝条遮挡、叶片遮挡与果实重叠等典型样本,对基准模型 YOLO 11n 与改进模型 MEW - YOLO 的检测结果进行对比,如图 6 所示。顺光条件下,两模型均能检出主要目标,但 YOLO 11n 在局部目标上存在漏检,而 MEW - YOLO 可补充检出,并呈现更完整的框选范围与更贴合果实轮廓的定位结果。逆光及遮挡场景中,MEW - YOLO 对果实区域的框选更连续、边界更稳定,背景干扰引起的定位偏差相对减弱。尤其在遮挡与重叠样本中,MEW - YOLO 的检测框更集中于果实主体,

漏检与定位波动减少;相较之下,YOLO 11n 更易出现漏检或框选不稳定。可视化结果表明 MEW - YOLO 在光照变化、遮挡与重叠等复杂条件下具有更好的检出稳定性与定位一致性,其趋势与表 3 的定量对比结果一致,为后续实际部署提供了直观佐证。

2.4.3 精确率-召回率关系

为比较模型在不同置信度阈值下精确率-召回权衡,绘制了 MEW - YOLO 与基准模型 YOLO 11n PR 曲线,如图 7 所示。由图 7 可见,MEW - YOLO PR 曲线在大多数召回率区间均略高于 YOLO 11n,且在中高召回率段仍能保持相对更高的精确率,曲

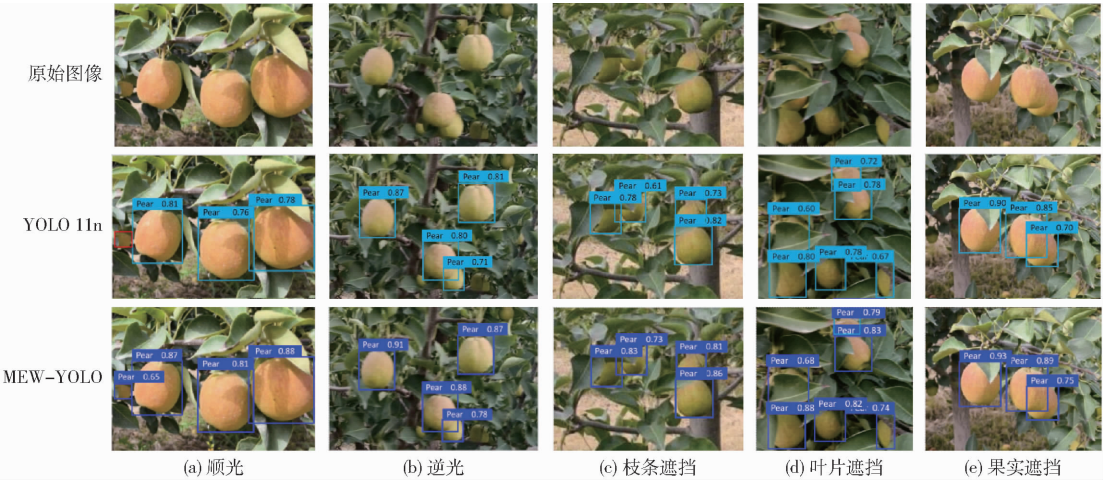


图 6 复杂果园场景下检测结果

Fig. 6 Detection results under complex orchard scenarios

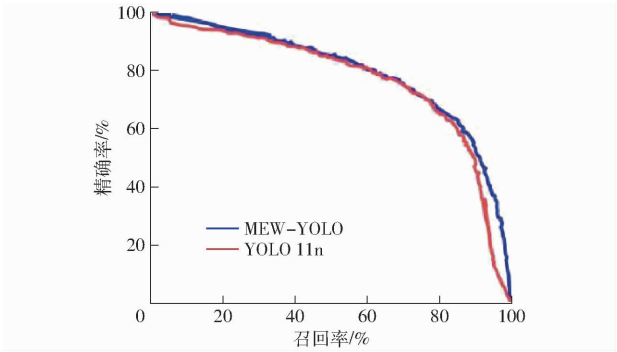


图 7 MEW-YOLO 模型与 YOLO 11n 的精确率-召回率 (PR) 曲线

Fig. 7 Precision-recall (PR) curves between MEW-YOLO and baseline YOLO 11n model

线下下降趋势更为平缓,说明改进模型对阈值变化的敏感性较低,输出一致性更好。相比之下,YOLO 11n 在高召回率区域精确率下降更明显,表明其在提高召回率的同时更易引入误检。总体而言,MEW-YOLO 的曲线下面积更大,反映其在不同阈值设置下具有更稳定的综合检测表现,该结果与前述收敛曲线及可视化结果所呈现的趋势一致。

2.4.4 改进模块注意力热力图

为从特征响应角度解释改进结构对目标聚焦能力的影响,采用 Grad-CAM 对 YOLO 11n 与 MEW-YOLO 进行可视化对照。Grad-CAM 是反映对模型判别贡献较大的空间区域,并不直接等同于检测框质量,因此在同一检测框架内对比更具可比性。选取 1 幅同时包含光照变化、枝叶遮挡与多尺度果实的典型样本进行分析,结果如图 8 所示,其中蓝色、黄色、红色分别表示低、中、高响应。由图 8 可以观察到,YOLO 11n 的高响应区域分布较为分散,除果实外在枝叶与局部高亮背景上亦存在较强响应,反映其对非目标干扰的抑制不足,易在逆光与遮挡条件下引入误检或置信度波动。相比之下,MEW-

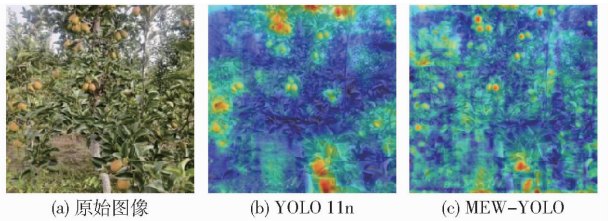


图 8 YOLO 11n 与 MEW-YOLO 的 Grad-CAM 注意力热力图

Fig. 8 Grad-CAM attention heatmaps between YOLO 11n and MEW-YOLO

YOLO 的高响应主要集中于果实主体及其边缘轮廓,背景高响应相对减弱,呈现更清晰的目标聚焦、背景抑制模式;该特征响应差异与前述可视化结果中所见的定位更连贯、误检更少的现象相互印证,从而为性能差异提供了可解释性支撑。

3 讨论

3.1 模型性能改进与结构协同分析

基于性能试验、消融结果及可视化分析,MEW-YOLO 的优势可从特征增强、空间重建与回归优化 3 条关键链路进行解释。通道层的局部-全局交互增强了目标相关响应并抑制背景激活,从而降低复杂背景下的误检风险,这与热力图中响应更集中、背景高响应减弱的现象相互印证。上采样路径的结构化改造增强了融合特征对边缘与局部结构的表达连续性,在严格 IoU 条件下使定位输出更具一致性,这与高阈值评价指标的改善趋势相一致。回归项的样本质量感知机制对不同难度样本的梯度贡献进行调节,有助于减弱低质量样本带来的训练扰动并提升收敛稳定性。上述 3 条链路分别作用于响应聚焦、细节传递与梯度优化环节,使改进并非简单叠加,而是在轻量化约束下形成面向复杂场景的有效分工与互补支撑。

3.2 模型轻量化与可用性

本文数据采集于自然果园,样本覆盖果园采摘中常见的光照变化、遮挡重叠与小目标等典型干扰情形,可较好表征实际采摘环境。面向采摘任务中漏检将造成目标遗漏,改进模型在更严格 IoU 阈值下 (mAP@0.5:0.95) 能更敏感地表征定位一致性与边界稳定性。由表 2 可知,MEW - YOLO 模型在复杂干扰条件下仍能保持较稳定的检出与定位输出,可为后续三维定位与末端执行提供可靠视觉输入。在部署约束方面,MEW - YOLO 在获得精度提升的同时保持轻量化结构,符合移动采摘平台端侧算力与功耗受限条件下对推理延迟与资源占用的要求。

3.3 模型稳定性与泛化能力

在复杂自然场景中,模型稳定性除受检测精度影响外,还与不同场景下特征表征的可靠性密切相关。结合训练收敛过程(图 5)与多场景测试结果(图 6)可见,MEW - YOLO 在光照变化、果实密集与遮挡等条件下的输出更为稳定;训练/验证曲线变化平滑,未呈现明显的过拟合特征,表明模型具有较好的泛化表现。基于消融试验结果可进一步说明,改进方法同时提升了特征表达的有效性与回归学习的稳定性,从而减弱复杂场景扰动与低质量样本对训练过程及定位结果的影响。因此,MEW - YOLO 在多类场景下呈现更一致的检测输出,可为轻量化检测模型在果园复杂场景中的鲁棒性提升提供可解释依据与设计参考。

4 结论

- (1)在 YOLO 11n 架构基础上提出轻量化改进检测模型 MEW - YOLO,围绕特征增强、空间重建与边界优化 3 个层面,提出了 MLCA、EUCB、WIoU 3 项改进:MLCA 融合全局与局部通道信息增强多尺度特征选择性并提升复杂光照与部分遮挡条件下的响应稳定性;EUCB 强化特征融合阶段的空间细节恢复以改善遮挡与重叠样本的边界拟合一致性;WIoU 基于样本质量自适应调节回归梯度,抑制低质量样本干扰并提升回归稳定性。
- (2)在自建库尔勒香梨数据集上,MEW - YOLO 相较 YOLO 11n 的 mAP@0.5、mAP@0.5:0.95 与召回率分别提升 2.7、3.3、3.7 个百分点;精确率提升 5.3 个百分点。模型参数量仅增加约 3.4%,浮点运算量为 6.8×10^9 ,在保持轻量化的同时实现检测精度与计算效率的同步提升。相较 YOLO v8n 等其他对比模型,MEW - YOLO 在较低计算量条件下获得更优的精度-复杂度权衡。
- (3)热力图分析结果表明,改进模型在果实区域的响应更集中、背景干扰响应减弱且边缘响应更连续,与定量结果及消融趋势一致,为性能提升提供了机制层面的解释支撑。
- (4)所提 MEW - YOLO 在检测精度与计算效率之间实现了较优平衡,适用于自然果园场景下的果实识别任务,可为智能采摘机器人视觉感知模块轻量化设计提供参考。

参 考 文 献

[1] DU X, CHENG H, MA Z, et al. DSW - YOLO: a detection method for ground-planted strawberry fruits under different occlusion levels[J]. Computers and Electronics in Agriculture, 2023, 214: 108304.

[2] JIANG Y, CHEN J, WANG Z, et al. Research progress and trend analysis of picking technology for Korla fragrant pear[J]. Horticulturae, 2025, 11(1): 90.

[3] 刘凯越,吴建军,李智慧,等. YOLO - Vega: 一种面向复杂环境的轻量级番茄检测模型[J]. 计算机工程与应用,2025, 61(21):94 - 104.

LIU Kaiyue, WU Jianjun, LI Zhihui, et al. YOLO - Vega: a lightweight tomato-detection model for complex environment[J]. Computer Engineering and Applications, 2025,61(21):94 - 104. (in Chinese)

[4] JIANG Y, CHEN J, HU G, et al. Mechanical properties of Korla fragrant pear fruiting branches and pedicels: implications for non - destructive harvesting[J]. Horticulturae, 2025, 11(8): 880.

[5] 赵晓,杨梦婷,张懿丹. 基于改进 YOLO v10n 的苹果叶片病害检测方法[J]. 农业机械学报,2025,56(9):657 - 666,728.

ZHAO Xiao, YANG Mengting, ZHANG Yidan. Apple leaf disease detection method based on improved YOLO v10n[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(9): 657 - 666,728. (in Chinese)

[6] 庄煜,徐珉林,李美琪,等. 基于改进 YOLO v11n 的轻量化葡萄穗梗采摘点定位方法与试验[J]. 农业机械学报,2025, 56(9):618 - 627.

ZHUANG Yu, XU Kunlin, LI Meiqi, et al. Lightweight grape cluster stalk picking point localization method and test based on improved YOLO v11n[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(9): 618 - 627. (in Chinese)

[7] 朱立成,王文贝,赵博,等. 基于 SDE - YOLO 的矮砧密植化果园苹果检测方法[J]. 农业机械学报,2025,56(9):638 - 647.

ZHU Licheng, WANG Wenbei, ZHAO Bo, et al. Detection method of apples in dwarf rootstock densification orchards based on SDE - YOLO[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(9): 638 - 647. (in Chinese)

[8] 黄华,张昊,胡晓林,等. 基于改进 YOLO v5 的复杂环境下花椒簇识别与定位方法[J]. 农业机械学报,2024,55(3):243–251.
HUANG Hua, ZHANG Hao, HU Xiaolin, et al. Recognition and localization method for pepper clusters in complex environments based on improved YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(3): 243–251. (in Chinese)

[9] 吕强,林刚,蒋杰,等. 基于改进 YOLO v5s 模型的自然场景中绿色柑橘果实检测[J]. 农业工程学报,2024,40(18):147–154.
LYU Qiang, LIN Gang, JIANG Jie, et al. Detecting green citrus fruit in natural scenes using improved YOLO v5s model[J]. Transactions of the CSAE, 2024, 40(18): 147–154. (in Chinese)

[10] 张震,周俊,江自真,等. 基于改进 YOLO v7 轻量化模型的自然果园环境下苹果识别方法[J]. 农业机械学报,2024,55(3):231–242.
ZHANG Zhen, ZHOU Jun, JIANG Zizhen, et al. Lightweight apple recognition method in natural orchard environment based on improved YOLO v7 model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(3): 231–242. (in Chinese)

[11] 刘天真,滕桂法,苑迎春,等. 基于改进 YOLO v3 的自然场景下冬枣果实识别方法[J]. 农业机械学报,2021,52(5):17–25.
LIU Tianzhen, TENG Guifa, YUAN Yingchun, et al. Winter jujube fruit recognition method based on improved YOLO v3 under natural scene[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(5): 17–25. (in Chinese)

[12] 雷志龙,刘畅,王权. 设施番茄采摘机器人识别定位与采摘方法研究[J]. 农业机械学报,2025,56(7):219–225.
LEI Zhilong, LIU Chang, WANG Quan. Recognition-localization and picking methods of facility tomato picking robots[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(7): 219–225. (in Chinese)

[13] 任锐,孙海霞,张淑娟,等. 基于改进 YOLOv8n 的不同栽培模式下玉露香梨轻量化检测[J]. 农业工程学报,2025,41(5):145–155.
REN Rui, SUN Haixia, ZHANG Shujuan, et al. Lightweight detection method for ‘Yuluxiang’ pear under different cultivation modes based on improved YOLOv8n[J]. Transactions of the CSAE, 2025, 41(5): 145–155. (in Chinese)

[14] 谭厚森,马文宏,田原,等. 基于改进 YOLOv8n 的香梨目标检测方法[J]. 农业工程学报,2024,40(11):178–185.
TAN Housen, MA Wenhong, TIAN Yuan, et al. Improved YOLOv8n object detection of fragrant pears[J]. Transactions of the CSAE, 2024, 40(11): 178–185. (in Chinese)

[15] 李茂,肖洋轶,宗望远. 基于改进 YOLO 11 的多尺度板栗果实识别方法[J]. 农业机械学报,2025,56(5):443–454.
LI Mao, XIAO Yangyi, ZONG Wangyuan. Multi-scale chestnut detection method based on improved YOLO 11 model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(5): 443–454. (in Chinese)

[16] XIE X, ZHANG R, GUO J, et al. Strawberry disease detection algorithm based on YOLO11–Strawberry[J]. Food Quality and Safety, 2025, 9: 027.

[17] 王铁君,张泽宇,郭晓然,等. MCFA–UNet:结合多尺度融合与注意力机制的图像生成网络[J]. 计算机工程与应用, 2025,61(12):222–231.
WANG Tiejun, ZHANG Zeyu, GUO Xiaoran, et al. MCFA–UNet: combining multiscale fusion and attention for image generation networks[J]. Computer Engineering and Applications, 2025, 61(12): 222–231. (in Chinese)

[18] XUE S, LI Z, WANG D, et al. YOLO–ALDS: an instance segmentation framework for tomato defect segmentation and grading based on active learning and improved YOLO11[J]. Computers and Electronics in Agriculture, 2025, 238: 110820.

[19] YAO J, LI Y, XIA Z, et al. WTAD–YOLO: a lightweight tomato leaf disease detection model based on YOLO11[J]. Smart Agricultural Technology, 2025, 12: 101349.

[20] HOU C, ZHANG X, TANG Y, et al. Detection and localization of citrus fruit based on improved you only look once v5s and binocular vision in the orchard[J]. Frontiers in Plant Science, 2022,13:972445.

[21] WAN D, LU R, SHEN S, et al. Mixed local channel attention for object detection[J]. Engineering Applications of Artificial Intelligence, 2023, 123: 106442.

[22] ZHOU Q, WANG Z, ZHONG Y, et al. Efficient optimized YOLOv8 model with extended vision[J]. Sensors, 2024, 24(20): 6506.

[23] WANG D, HE D. Apple detection and instance segmentation in natural environments using an improved Mask Scoring R–CNN Model[J]. Frontiers in Plant Science, 2022, 13:1016470.

[24] SHI H, LIU C, WU M, et al. Real-time detection of Chinese cabbage seedlings in the field based on YOLO11–CGB[J]. Frontiers in Plant Science, 2025, 16: 1558378.

[25] RAHIMA K, MUHAMMAD H. Yolov11: an overview of the key architectural enhancements[J]. arXiv Preprint, arXiv: 2410.17725, 2024.