

文章编号:1007-2780(XXXX)XX-0001-16

融合混合注意力的轻量化柑橘成熟度检测算法

王文坤¹, 谢辉^{1*}, 姜吴瑾², 李洪兵², 钱楚天¹

(1. 重庆三峡学院 电子与信息工程学院, 重庆 404100;

2. 重庆三峡学院 计算机科学与工程学院, 重庆 404100)

摘要:针对自然环境下果树柑橘检测中面临的成熟度区分困难、枝叶与果实遮挡严重、模型复杂度高及资源部署受限等挑战,本文提出了一种基于改进YOLOv11的轻量化树上柑橘成熟度检测算法YOLO-HiP。首先,采用改进的HGNetv2-L网络作为主干网络,并结合分层特征提取策略,显著提升了模型在复杂场景中的解析能力,同时有效降低了计算复杂度和资源消耗。其次,设计了轻量化混合注意力模块C2PSA_iRMB,通过融合C2PSA机制与iRMB模块,优化了计算开销,并增强了长距离信息的处理能力,提升了模块的灵活性与计算效率。最后,构建了C3k2_PConv模块,通过减少冗余计算和内存访问,进一步提高了空间特征提取的效率。实验结果表明,YOLO-HiP在mAP₅₀上达到了94.3%,较原模型提升了4.7%,参数量仅为5.1M(减少45.7%),计算量降至13.9GFLOPs(降低34.7%),帧率提升至227.4帧/秒(提高25.1%)。该模型在保证检测精度的同时,显著压缩了模型规模,为柑橘采摘机器人等嵌入式系统计算资源有限的平台适用提供了创新且可行的解决方案。

关键词:柑橘;成熟度;YOLOv11;轻量化;目标检测

中图分类号:TP391.4;TN911.73 **文献标识码:**A **doi:**10.37188/CJLCD.2026-0039 **CSTR:**32172.14.CJLCD.2026-0039

Lightweight citrus maturity detection algorithm with hybrid attention fusion

WANG Wenkun¹, XIE Hui^{1*}, JIANG Wujin², LI Hongbing², QIAN Chutian¹

(1. School of Electronic and Information Engineering, Chongqing Three Gorges University,
Chongqing 404100, China;

2. School of Computer Science and Engineering, Chongqing Three Gorges University,
Chongqing 404100, China)

Abstract: To address the challenges faced in citrus detection on fruit trees in natural environments, such as difficulty in maturity differentiation, severe occlusion from branches and leaves, high model complexity, and resource deployment limitations, this paper proposes a lightweight citrus maturity

收稿日期:2026-03-16;**修订日期:**2026-04-08.

基金项目:重庆市自然科学基金(No.2022NSCQ-MSX4084);重庆市教委科学技术研究项目(No.KJZD-M202201204, No.KJZD-M202301203);重庆市教委科学技术研究项目(No.KJQN202401237, No.KJQN202501243)
Chongqing Natural Science Foundation (No.2022NSCQ-MSX4084); Science and Technology Research Program of the Chongqing Municipal Education Commission (No. KJZD-M202201204, No. KJZD-M202301203); Science and Technology Research Program of the Chongqing Municipal Education Commission (No.KJQN202401237, No.KJQN202501243)

*通信联系人, E-mail:20030041@sanxiau.edu.cn

detection algorithm for fruit trees, YOLO-HiP, based on an improved YOLOv11. First, an improved HGNetv2-L network is used as the backbone, combined with a hierarchical feature extraction strategy, significantly enhancing the model's capability to analyze complex scenes while effectively reducing computational complexity and resource consumption. Next, a lightweight hybrid attention module, C2PSA_iRMB, is designed. By integrating the C2PSA mechanism with the iRMB module, the computational cost is optimized, and the ability to process long-range information is enhanced, improving the module's flexibility and computational efficiency. Finally, a C3k2_PConv module is constructed, further improving spatial feature extraction efficiency by reducing redundant computations and memory access. Experimental results show that YOLO-HiP achieves 94.3% mAP50, an improvement of 4.7% over the original model, with only 5.1M parameters (a 45.7% reduction), a computational load of 13.9 GFLOPs (a 34.7% reduction), and a frame rate of 227.4 FPS (a 25.1% increase). This model significantly reduces model size while ensuring detection accuracy, providing an innovative and feasible solution for platforms with limited computational resources, such as citrus-picking robots and other embedded systems.

Key words: citrus; maturity; YOLOv11; lightweight; object detection

1 引 言

中国是全球主要柑橘生产国之一^[1]。柑橘采收质量直接影响果品品质与商品化率,其中成熟度是衡量采收时的关键指标。在精准农业背景下,机器人视觉系统不仅需要快速、准确地定位果实^[2-3],还需实现细粒度成熟度判别,以支撑自动采摘与分级作业,同时满足端侧低计算开销的部署需求。传统基于颜色、形状和纹理等手工特征的识别方法^[4-6],难以适应不同品种及生长阶段带来的表观差异。相比之下,深度学习为复杂场景下的果实检测与成熟度识别提供了更高效的解决方案。

现有基于深度学习的果实检测方法大致可分为两类^[7]:一类是基于候选区域的两阶段检测算法,另一类是将检测任务视为回归问题的单阶段检测算法。两阶段检测算法以 R-CNN 系列为代表,其流程通常包括候选区域生成和目标分类定位两个阶段。Mai 等^[8]基于改进 Faster R-CNN 实现了自然环境下的果实检测,平均精度达到 77.75%;朱旭等^[9]基于改进 Faster R-CNN 实现了蓝莓果实成熟度分级识别,能够同时区分不同成熟度目标,且各类别识别性能较为均衡。尽管两阶段检测算法在精度方面具有优势,但其模型规模较大、内存开销较高,推理速度受限,难以满足实时检测需求^[10]。

相较之下,单阶段检测算法以 YOLO 系列为代表,直接通过卷积神经网络回归目标类别与位置信息,省略候选区域生成过程,在显著提升检测速度的同时,通常会带来一定的精度损失^[11]。由于自然果园场景对实时性要求较高,兼顾速度与精度的单阶段检测方法受到广泛关注。吕石磊等^[12]在 YOLOv3 中采用 MobileNet-v2 作为特征提取网络,并引入 GIoU 损失函数,显著降低了模型参数量,但在遮挡果实场景下检测精度仍有待提升。Zheng 等^[13]基于 YOLOv4 提出多尺度检测模型 YOLO-BP,通过剪枝 CSPDarknet53 并结合 Bi-PANet 融合多层特征,提升了绿色柑橘检测性能,但同时增加了模型计算开销。Lyu 等^[14]基于 YOLOv5s 构建 YOLOv5-CS 模型用于果园绿柑橘识别,但其参数规模较大,限制了在资源受限设备上的应用。李莹等^[15]在 YOLOv8 中引入混合注意力 Transformer 和多头自适应特征融合检测头,以增强成熟度特征表达,但模型复杂度较高、检测效率较低。Lawal 等^[16]在 YOLOv5s 基础上融合特征拼接与注意力机制构建轻量化模型,但其在自然环境下树上柑橘检测中的鲁棒性仍显不足。高新阳等^[17]在 YOLOv5 中引入 ShuffleNetV2 与 SimAM 注意力机制,在实现模型轻量化的同时提升了检测精度,但尚未对不同转色期柑橘目标开展系统验证。杨如强等^[18]在 YOLOv11 框架中引入 GhostH-

GNetV2 主干网络,并对颈部网络和检测头进行多模块优化设计,以提升模型的多尺度特征融合能力,但该方法在模型轻量化方面仍显不足。胡江波等^[19]通过在 YOLOv11 中引入 EGCSN 模块,增强了模型在复杂场景下的目标定位与成熟度分级能力,但其跨场景泛化能力仍有待进一步系统评估。罗一鸣等^[20]通过在 YOLOv11n 中引入 IRSC 模块优化主干网络,并结合 ADown 模块降低模型复杂度,在提升检测精度的同时兼顾了模型轻量化,但其对不同成熟阶段柑橘的细粒度表征能力仍缺乏充分验证。

近年来,YOLO 系列模型因兼具检测精度与实时性,在农业视觉检测中得到广泛应用。作为 Ultralytics 推出的新一代单阶段目标检测模型,YOLOv11 通过引入 C3k2、C2PSA 和 SPPF 等模块,进一步提升了特征提取与多尺度融合能力,在精度、速度和计算效率之间取得了较好平衡。尽管其在树上柑橘成熟度检测中的研究尚少,但仍具有较大应用潜力。然而,面对复杂果园场景,YOLOv11 仍存在局限:主干网络对成熟度细粒度特征表征不足,易受遮挡、密集分布和小目标等因素影响;C2PSA 模块对全局上下文与长距离依赖的建模能力有限,而复杂注意力机制又会增加计算负担;C3k2 模块也存在一定冗余计算,限制了资源受限场景下的实时性。

从实际应用需求来看,树上柑橘成熟度检测的关键挑战在于,如何在复杂自然环境下兼顾成熟度细粒度区分精度与模型轻量化部署性能。一方面,不同成熟阶段的差异主要体现在局部颜色比例、纹理细节和表面光泽等细微变化上,在强光、逆光、遮挡、果实重叠及小目标分布等场景下易出现类别混淆、漏检和误检;另一方面,采摘机器人、移动终端和嵌入式设备又要求模型具有较低的参数量、计算量和较高的实时性。因此,构建一种兼具高精度、强鲁棒性与轻量化特征的树上柑橘成熟度检测模型,是本文着力解决的核心问题。上述关键问题的解决可为智慧农业和智能装备领域提供技术支撑,有助于提升柑橘采摘、分级与产量评估的自动化水平,并为边缘设备上的农业视觉识别任务提供参考。

针对 YOLOv11 在树上柑橘成熟度检测中的不足,本文提出一种融合混合注意力的轻量化

模型 YOLO-HiP。其主要贡献如下:

1) 提出改进的 HGNetv2-L 轻量化主干网络,通过构建层次化特征提取机制,使模型能够在不同尺度和不同语义层级上逐步学习图像表示,从而增强其对复杂模式的表征能力,提升复杂场景下的特征提取效果;同时,在网络结构设计中兼顾轻量化需求,为后续模型部署奠定基础。

2) 提出轻量化混合注意力机制 C2PSA_iRMB,通过引入倒置残差移动块(inverted residual mobile, iRMB)对 C2PSA 结构进行重设计,并融合深度可分离卷积与自注意力机制,在较低计算代价下实现局部空间特征提取与全局依赖关系建模的有效协同,从而增强模型对复杂特征信息和长距离依赖关系的处理能力。

3) 提出 C3k2_PConv 模块,将部分卷积(Partial Convolution, PConv)与 C3k2 模块相结合,通过仅对部分输入通道执行空间特征提取、其余通道采用恒等映射的方式,在保证特征信息有效传递的基础上,减少冗余计算与内存访问开销,从而实现更加高效的特征提取。

2 YOLOv11 及其改进算法

YOLOv11 是一种兼顾检测精度与推理效率的单阶段目标检测模型,其主干网络由卷积层、C3k2、C2PSA 和 SPPF 等模块构成,颈部采用 PAN-FPN 进行多尺度特征融合,检测头则通过解耦设计完成分类与回归任务。本文选取 YOLOv11s 作为基线模型,以满足树上柑橘成熟度检测对精细识别、实时性与轻量化部署的需求。

本文提出的 YOLO-HiP 网络结构如图 1 所示,首先引入改进 HGNetv2-L 的轻量化主干网络,以增强模型对不同尺度和不同语义层级特征的层次化提取能力;其次,设计轻量化混合注意力机制 C2PSA_iRMB,通过融合深度可分离卷积与自注意力机制,提升局部空间信息与全局依赖关系的联合建模能力;最后,提出 C3k2_PConv 模块,通过仅对部分输入通道执行空间特征提取、其余通道采用恒等映射的方式,减少冗余计算与内存访问开销,从而实现模型精度、效率与轻量化性能的协调优化。

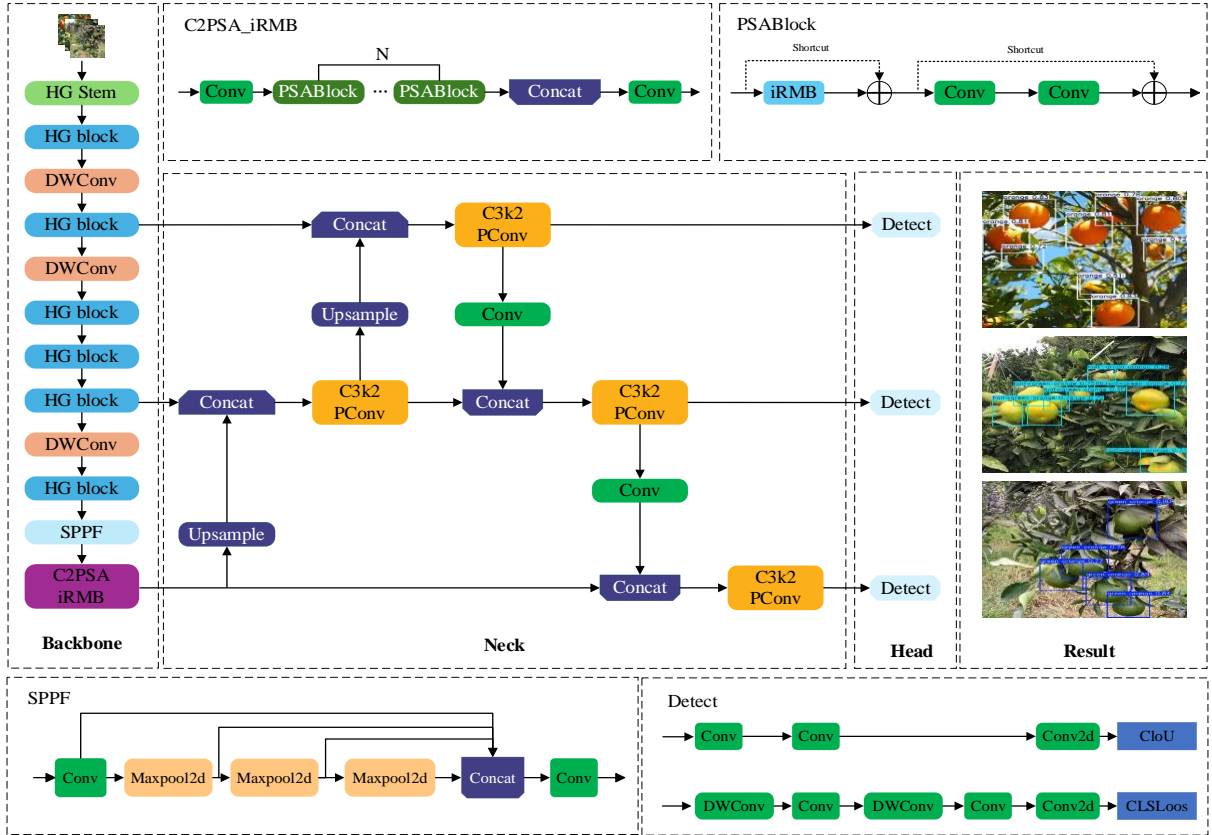


图1 YOLO-HiP网络结构

Fig. 1 YOLO-HiP network structure

3 YOLO-HiP 树上柑橘检测模型

3.1 改进 HGNetv2-L 的轻量化主干网络

为实现高效轻量化网络设计,本文在充分分析 RT-DETR^[22] 中 HGNetv2-L 架构优势的基础上,对其结构进行改进与重构,创新性地设计了一种以 Stem 层和层次图(hierarchical graph, HG)模块为核心的轻量化主干网络,并将其引入 YOLOv11s 以替换原有主干结构。该主干网络通过层次化特征建模机制,在不同尺度与抽象层级上逐级学习复杂模式信息,在显著降低模型参数量与计算复杂度的同时,有效增强了网络对复杂果园图像中多尺度目标和细粒度特征的代表能力,从而为后续检测任务提供更加高效且判别性更强的特征表示。

Stem 层作为网络的输入端,负责对原始像素信息进行初步特征映射。不同于传统结构中采用单一大步幅卷积的方式, HGNetv2-L 对 Stem 层进行了针对性优化,通常由两至三个连

续的 3×3 卷积层组成,并结合归一化层与 HardSwish 激活函数,在保证计算效率的同时增强特征表达能力。HardSwish 激活函数的表达式为:

$$\text{hardswish}(x) = \begin{cases} 0, & \text{if } x \leq -3 \\ x * (x+3) / 6, & \text{if } -3 < x < 3 \\ x, & \text{if } x > 3 \end{cases} \quad (1)$$

式(1)中:当输入 $X \leq -3$ 时,函数输出恒为零,呈现具有明确下界的硬饱和特性,从而提升模型的数值稳定性;当 $X \geq 3$ 时,函数退化为恒等映射,在该区间内保持线性且无上界限制,保证梯度信息的有效反向传播;当 $-3 \leq X \leq 3$ 时,函数在分段点 $X = -3$ 和 $X = 3$ 处保持一阶连续性,形成平滑的非线性过渡区,在增强非线性表征能力的同时有效降低指数运算带来的计算复杂度。

相较于将原始 HGNetv2-L 直接用作通用视觉主干网络,本文结合复杂果园场景下目标检测任务对实时性与轻量化的需求,对其进行了针对

性结构优化。在保留 HGNetv2-L 层次化特征提取思想的基础上,本文从输入端特征编码、主干阶段组织方式和多尺度输出适配三个方面进行了改进。首先,在浅层特征提取阶段,对初始下采样模块进行轻量化重构,采用连续小卷积核堆叠替代高开销下采样操作,在保证感受野扩展的同时降低前端冗余计算。其次,在主干特征提取阶段,对 HGBlock 的堆叠深度、通道配置及信息交互方式进行优化,减少重复特征建模,降低参数冗余与计算复杂度。最后,为适配 YOLOv11s 的颈部网络与检测头结构,重新调整各阶段输出特征图的尺度及通道配置,以提升多层次特征在后续融合与预测中的利用效率。通过上述改进,模型在保证检测精度的同时提升了运

行效率,为后续检测性能优化奠定了基础。

HG 块(HGBlock)是 HGNetv2-L 架构的核心组成,其主要创新体现在高阶特征融合机制上。该模块采用多分支并行结构,引入不同卷积核尺度及恒等映射路径对输入特征进行建模,并通过可学习的门控权重实现自适应加权融合,而非简单的相加或拼接,从而动态整合不同感受野与抽象层次的信息。浅层 HG 块侧重于局部细节与边缘纹理等低级特征建模,深层 HG 块则强化对全局上下文与语义信息的表达。借助这种层次化特征处理策略,HGBlock 能够生成兼具高分辨率与强判别性的特征表示,显著提升模型整体的特征表征能力。HGBlock 模块结构如图 2 所示。

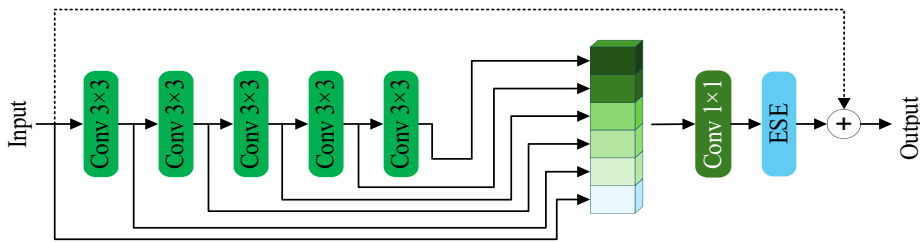


图2 HGBlock网络结构

Fig. 2 HGBlock network structure

3.2 C2PSA_iRMB 混合注意力机制

针对现有 YOLOv11s 网络中 C2PSA 结构在长距离依赖建模能力不足、结构灵活性与计算效率受限,且在移动端密集预测任务中性能下降的问题,本文提出了一种新的轻量化混合注意力建模机制 C2PSA_iRMB,并将其部署于骨干网络末端,以替换原有的 C2PSA 模块。尽管 YOLOv11 在速度与精度之间取得了较好平衡,但在复杂果园场景下仍存在不足。其骨干网络高层的 C2PSA 模块对局部细节与全局依赖的协同建模能力有限:卷积感受野受限,难以有效建模遮挡、密集分布和小目标场景中的长距离关联;原始注意力结构偏重全局响应,对局部纹理与边缘细节保留不足,易削弱目标与复杂背景的区分能力。同时,直接增强全局注意力又会增加参数量和计算开销,不利于轻量化部署。为此,本文提出 C2PSA_iRMB 混合注意力机制,通过在 C2PSA 中引入 iRMB,结合深度可分离卷积、通道注意力与空间注意力,构建“局部建模—全局

交互”协同特征表达方式,在控制计算成本的同时提升了遮挡目标、小目标及密集目标的检测能力。该方法创新性地将倒置残差移动块 iRMB^[23] 引入 C2PSA 结构中,并结合局部与全局混合注意力机制。具体而言,通过融合通道注意力与空间注意力来优化特征的加权与聚焦,从而有效提升长距离信息的建模能力。同时,iRMB 的引入加强了特征重用路径与注意力建模单元的协同作用,既降低了计算开销,又提升了特征交互效率。该混合注意力机制通过有效改善长距离依赖建模与特征交互,增强了模型在移动端和资源受限设备上的密集预测性能,并增强了实际部署的可行性。

iRMB 通过对倒置残差块(inverted residual block, IRB)与 Transformer 中关键有效组件的统一建模,构建了一种面向轻量化网络设计的通用视角,将传统 CNN 的 IRB 扩展至基于注意力机制的特征建模范式,并进一步抽象出残差元移动块(meta mobile block, MMB)。如图 3 所示,

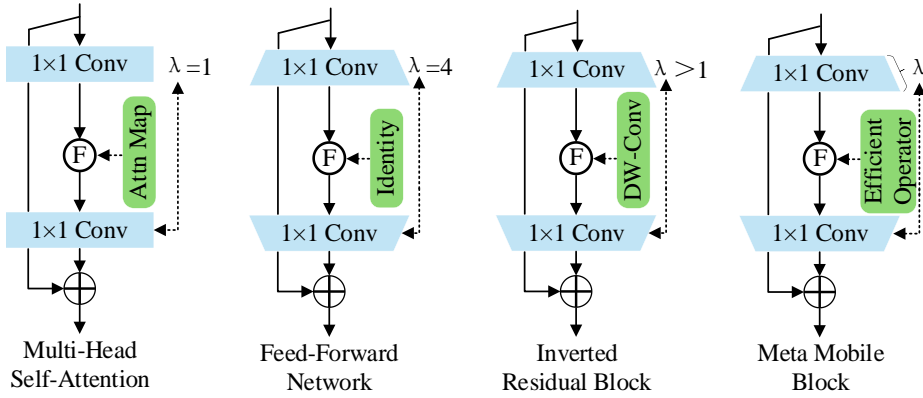


图3 iRMB设计理念和结构

Fig. 3 iRMB design concept and structure

iRMB 的结构设计源于对多头自注意力与前馈网络的模块化整合,通过灵活组合不同扩展比率 λ (expansion ratio) 与高效算子 F (efficient operator), 在显著降低模型复杂度的同时兼顾性能与计算效率,为轻量级模型设计提供了有效支撑。

图 4 展示了 C2PSA_iRMB 的结构范式。它将结合了深度可分离卷积 (3×3 DW-Conv) 的混合网络模块 iRMB 与 C2PSA 注意力机制相融合。 1×1 卷积用于通道数的压缩和扩张,以此优化计算效率。深度可分离卷积 (3×3 DW-Conv) 用于捕捉空间特征,而 C2PSA 注意力机制则用于捕获特征间的全局依赖关系。

设输入特征图为 $X \in R^{H \times W \times C}$, 其中 H, W 表示空间尺寸, C 为通道数。传统标准卷积在空间维度与通道维度上同时进行特征提取与通道融

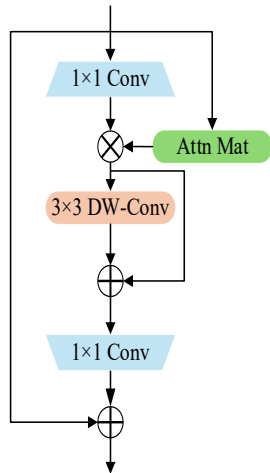


图4 C2PSA_iRMB模块结构

Fig. 4 C2PSA_iRMB module structure

合,其输出 $Y \in R^{H \times W \times C_{out}}$ 可表示为:

$$Y_{i,j,o} = \sum_{c=1}^C \sum_{u=1}^k \sum_{v=1}^k X_{i+u,j+v,c} \cdot K_{u,v,c,o}, \quad (2)$$

其中 k 为卷积核尺寸, $K \in R^{k \times k \times C \times C_{out}}$ 为卷积核参数, o 为输出通道索引。为降低计算开销,深度可分离卷积将上述操作解耦为深度卷积与逐点卷积两步。首先,深度卷积对每个输入通道独立施加空间卷积,不进行跨通道求和。以 3×3 深度卷积为例,其输出 $Z \in R^{H \times W \times C}$ 定义为:

$$Z_{i,j,c} = \sum_{u=1}^3 \sum_{v=1}^3 X_{i+u,j+v,c} \cdot K_{u,v,c}^{DW}, \quad (3)$$

其中 $K^{DW} \in R^{3 \times 3 \times C}$ 为深度卷积核, (i, j) 表示空间位置, c 表示通道索引。该步骤仅建模每个通道内部的空间相关性。随后,为实现通道间信息融合,引入 1×1 逐点卷积将 Z 映射至 $Y \in R^{H \times W \times C_{out}}$:

$$Y_{i,j,o} = \sum_{c=1}^C Z_{i,j,c} \cdot K_{c,o}^{PW}, \quad (4)$$

其中 $K^{PW} \in R^{1 \times 1 \times C \times C_{out}}$ 为逐点卷积参数。与标准卷积相比,深度可分离卷积在保持表征能力的同时显著减少参数量与乘加运算量:标准 3×3 卷积参数量为 $9CC_{out}$, 而深度可分离卷积参数量为 $9C + CC_{out}$, 从而在 C 较大时可获得近似 $O(1/9)$ 量级的计算与参数节省。

元移动块 MMB 通过引入不同扩展比率 λ 与高效算子 F 实现高度模块化设计,在本文所构建的 C2PSA_iRMB 模块中,MMB 的扩展比率 λ 取 2, 高效算子 F 采用 3×3 深度可分离卷积与注意力分支相结合的轻量化算子形式。MMB 模块具体结构如图 5 所示。

具体而言,输入特征首先进入 1×1 卷积进

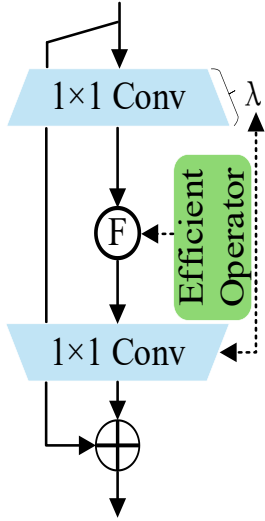


图5 Meta Mobile Block模块结构
Fig. 5 Meta Mobile Block module structure

行通道扩展,将通道数由 C 映射为 λC ,以提升中间特征空间的表达能力;随后,扩展后的特征送入高效算子 F 进行空间信息提取与特征交互,其中局部建模部分采用 3×3 DW-Conv 捕获邻域空间细节,注意力分支则用于补充长距离依赖关系的建模;最后,再通过 1×1 卷积将通道压缩回输出维度,以控制整体参数量与计算复杂度。之所以将 λ 设定为 2,主要基于轻量化检测任务中精度与复杂度平衡的考虑:当 λ 取值过小时,模块中间表征空间受限,特征重构能力不足,不利于复杂果园场景下遮挡目标、小目标及密集目标的判别;而当 λ 进一步增大时,虽然能够增强特征表达能力,但也会显著提升参数量与 FLOPs,削弱本文面向轻量化部署的设计目标。因此,本文采用 $\lambda=2$ 作为兼顾表示能力与推理效率的折中设置。与此同时, F 选用深度可分离卷积而非标准卷积,其依据在于前者能够在保持较强空间建模能力的同时显著降低参数开销,并与 C2PSA 的全局依赖建模形成互补,从而使模块同时具备局部细节提取与全局信息交互能力。通过上述参数设置,MMB 在无需大幅增加模型负担的前提下,提高了特征表达的灵活性与模块的可复用性,也为后续模型复现与性能分析提供了明确依据。

3.3 C3k2_PConv 模块

为在保证特征表达能力的前提下进一步降低冗余计算与内存访问开销,本文创新性地将部

分卷积 PConv^[24] 引入 C3k2 结构中,设计了一种新型高效特征提取模块 C3k2_PConv,并将其部署于颈部网络中,以替换原有的 C3k2 模块,其整体结构如图 6 所示。

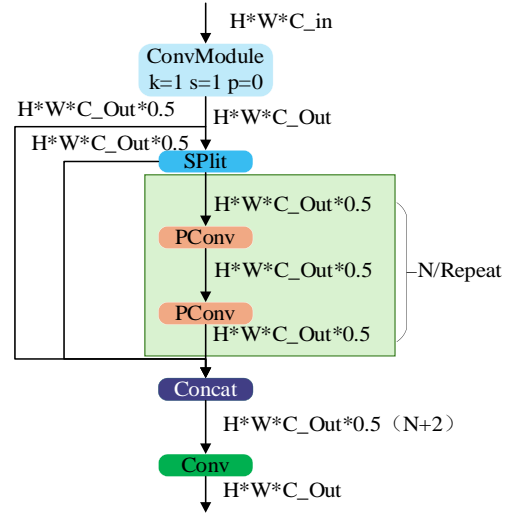


图6 C3k2_PConv2模块结构
Fig. 6 C3k2_PConv2 module structure

图 6 中所示的“N/Repeat”表示 PConv 子结构在 C3k2_PConv2 模块内部的重复堆叠次数,即模块的深度参数 N 。该参数用于控制局部特征提取单元的迭代次数,从而调节模块的特征表达能力与计算复杂度。当 N 取值较小时,模块整体结构更为轻量,参数量和计算量较低,能够满足实时检测任务对推理效率的要求,但其特征提取深度相对有限,对复杂背景下小目标、遮挡目标及密集目标的细粒度表征能力相对不足;当 N 取值增大时,模块对局部纹理信息和空间结构特征的建模能力将得到增强,有助于提升特征重构效果和目标判别能力,但同时也会带来更高的参数开销与推理延迟,不利于轻量化模型的部署应用。

基于上述考虑,本文将 N 作为 C3k2_PConv 2 模块的重要结构超参数,其设置遵循检测精度与计算效率相平衡的原则。采用重复堆叠设计的主要动机在于:单层 PConv 虽然能够以较低的计算代价完成局部特征提取,但其单层建模能力有限,难以充分适应果园场景中目标尺度变化大、遮挡严重和背景干扰强等复杂情况。通过适度增加 PConv 子结构的重复次数,可以在尽量

控制模型复杂度的前提下,逐步增强模块对局部细节信息的挖掘能力,提高特征表示的充分性与稳定性,并进一步改善后续特征融合阶段的输入质量。因此,“N/Repeat”的引入本质上是为了实现模块深度的可调节设计,使 C3k2_PConv2 能够在轻量化约束下兼顾特征提取能力与运行效率。本文实验中,N 采用固定配置,并结合整体网络层次结构进行统一设计,以保证改进模块与主干网络及颈部网络之间具有良好的特征衔接关系,同时避免因过度堆叠而引入冗余计算。

PConv 的核心思想在于充分利用特征图通道间的冗余性,以降低计算量与内存访问开销。具体而言,PConv 仅在部分输入通道上执行常规卷积以提取空间特征,其余通道保持恒等映射。该设计具有三方面优势:一是通过减少参与卷积计算的通道数,显著降低浮点运算量(FLOPs);二是有效减少内存访问次数,对 I/O 受限设备尤为友好;三是在后续逐点卷积阶段实现跨通道信息交互,从而在降低计算开销的同时保持特征信息的完整流动。图 7 给出了 PConv 的结构示意。

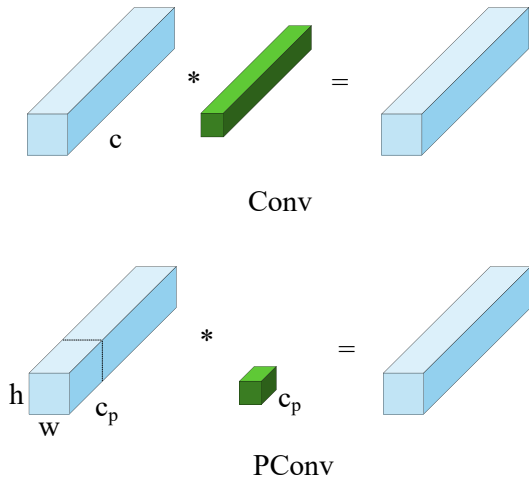


图 7 PConv 模块结构

Fig. 7 PConv module structure

常规卷积和 PConv 模块的浮点运算量 FLOPs 计算公式如下:

$$h * w * k^2 * c^2, \quad (5)$$

$$h * w * k^2 * c_p^2. \quad (6)$$

通过式(5)和式(6)看出,当通道比例 $r = c_p / c = 1/4$ 时,PConv 的浮点运算量仅为常规卷积的 $1/16$,从而显著降低了模型的计算复杂度。为

充分挖掘并融合各通道的特征信息,在 PConv 之后进一步引入常规卷积操作,以实现跨通道特征交互并保证特征表达的完整性。

4 实验与结果分析

4.1 实验环境及树上柑橘数据集

本研究数据集来源于重庆市某水果种植场标准化生产基地,依据统一的采样规范和现场记录流程采集,覆盖柑橘果实未熟期与成熟期两个关键生长阶段。采集过程中,采用专业图像采集设备,在强光、逆光和低照度等多种典型光照条件下,于约 1 m 距离、果树周向多视角进行实地拍摄,图像内容涵盖单果、多果、枝叶遮挡及果实重叠等复杂自然场景。

由于公开柑橘成熟度检测数据集较少,且在场景复杂性、光照变化、遮挡形式和成熟度细粒度标注等方面难以满足本文研究需求,本文构建了面向实际果园环境的自建数据集。同时,为提高数据多样性与模型泛化能力,补充选取公开数据集 CitDet 中的部分图像,并通过网络采集获取部分自然环境下的树上柑橘图像。经人工筛选并剔除模糊、重复及低质量样本后,最终获得 5 057 张高质量原始图像;结合 Mosaic、随机翻转、旋转、裁剪及图像拼接等数据增强方法,构建了包含 11 233 张图像的训练数据集。该数据集目前尚未公开发布,但在采集与构建过程中已尽可能覆盖多光照、多视角、多遮挡及多目标分布等真实场景,以增强样本多样性和模型对复杂环境的适应能力。

依据农业行业标准 NY/T 716-2003^[26] 中关于柑橘采摘成熟度的规定,并结合产量评估与实际采摘需求,本文将柑橘样本划分为三类:1)果面无转色区域(green orange);2)果面存在转色且比例不超过 $2/3$ (half-green orange);3)果面转色面积超过 $2/3$ (orange)。基于该分类准则,采用 LabelImg 工具对图像进行人工标注,并生成 YOLO 格式的 TXT 标注文件。部分样本如图 8 所示。随后,按 8:1:1 的比例将数据集随机划分为训练集、验证集和测试集,分别包含 8 986、1 123 和 1 124 张图像,构建了自然环境下树上柑橘成熟度检测数据集,如表 1 所示;模型训练环境及参数配置如表 2 所示。

表1 数据集划分

Tab. 1 Dataset partitioning

类型	训练集	测试集	验证集	总计
无转色区域	2 991	371	375	3 737
转色不超 2/3	2 999	377	375	3 751
转色超过 2/3	2 996	375	374	3 745

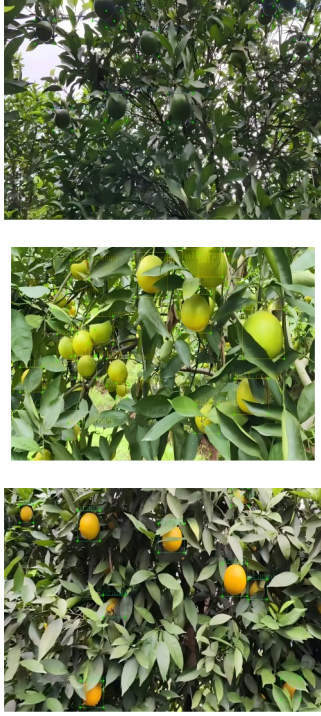


图8 树上柑橘成熟度分类数据集的部分图像样本

Fig. 8 Partial image samples from the citrus maturity classification dataset on trees

表2 实验环境及参数配置

Tab. 2 Experimental environment and parameter configuration

名称	参数
操作系统	windows 10
CPU	酷睿 i5 12600KF
GPU	NVIDIA GeForce RTX 4060TI
显存	16G
编程语言	python3.8
深度学习框架	Pytorch 2.2.2+CUDA 12.1
优化函数	SGD
训练轮数	100
初始学习率	0.01
批量	16
输入图像尺寸	640*640

4.2 模型评价指标

实验使用参数量 (Params)、浮点运算次数 (giga floating point operations, GFLOPs) 来综合衡量模型轻量化水平,使用帧率 (frames per second, FPS) 衡量模型的检测速度,选取 IoU 阈值为 0.5 的平均精度均值 mAP_{50} (mean average precision, mAP), IoU 在 0.5~0.95 (其中步长为 0.05) 加权平均后的平均精度即 $mAP_{50:95}$ 来评估模型的检测准确度。使用 P、R 曲线图和 mAP 图来评估模型的整体性能。其计算公式如下所示:

$$P = \frac{TP}{TP + FP}, \quad (7)$$

$$R = \frac{TP}{TP + FN}, \quad (8)$$

$$F1 = 2 \times \frac{P \times R}{P + R}, \quad (9)$$

$$AAP = \int_0^1 PdR, \quad (10)$$

$$mmAP = \frac{1}{N} \sum_{i=1}^N AAP_i, \quad (11)$$

式中: TP (true positives) 表示正确识别树上柑橘目标的数量; FP (false positives) 表示错误识别树上柑橘目标的数量; FN (false negative) 表示未识别出树上柑橘目标的数量。如果柑橘目标的真实框和预测框的 $IOU \geq 0.5$ 则检测框被标记为 TP, 反之记为 FP。 A_{AP} (average precision, AP) 是 P-R 曲线与坐标轴围成的面积。

4.3 消融实验

为验证改进算法的有效性及各模块贡献,在自然环境下树上柑橘成熟度分类数据集上开展 8 组消融实验。对比不同网络结构配置,进一步系统评估各改进模块对模型性能的影响及其在整体框架中的作用。消融实验检测结果如表 3 所示,检测结果对比如图 9 所示。

针对自然环境下树上柑橘成熟度分类检测任务,本文对所提出的各项改进模块进行了系统而全面的评估。实验结果表明,在以改进 HGNetv2-L 轻量化主干网络替换原有主干结构后,模型的 mAP_{50} 由 0.896 提升至 0.903, 提升幅度为 0.7%; 同时模型参数量减少 21.3%, GFLOPs 降低 4.1 G, 推理速度提升 22.3 FPS, 表明改进后的主干网络能够有效减少特征提取过程中的冗余计算,在保证性能的同时显著降低计算开销,并进一步提升整体推理效率。在原主

表3 树上柑橘成熟度分类数据集上消融对比实验

Table 3 Ablation comparison experiments on the dataset for citrus fruit maturity classification on trees

模型	改进HGNetV2-L	C2PSA_i RMB	C3k2_P Conv2	mAP ₅₀ /%	mAP _{50,95} /%	Para/M	GFLOPs/G	FPS
YOLOv11s				89.6	47.8	9.4	21.3	181.8
YOLOv11s-H	✓			90.3	48.5	7.4	17.2	204.1
YOLOv11s-i		✓		92.6	56.7	9.5	21.4	144.9
YOLOv11s-P			✓	91.4	53.1	8.2	20.1	62.1
YOLOv11s-Hi	✓	✓		92.2	55.7	7.4	17.3	113.6
YOLOv11s-HP	✓		✓	92.0	54.9	6.8	16.8	208.3
YOLOv11s-iP		✓	✓	92.8	57.1	8.9	21.0	204.1
YOLO-HiP	✓	✓	✓	94.3	59.8	5.1	13.9	227.4

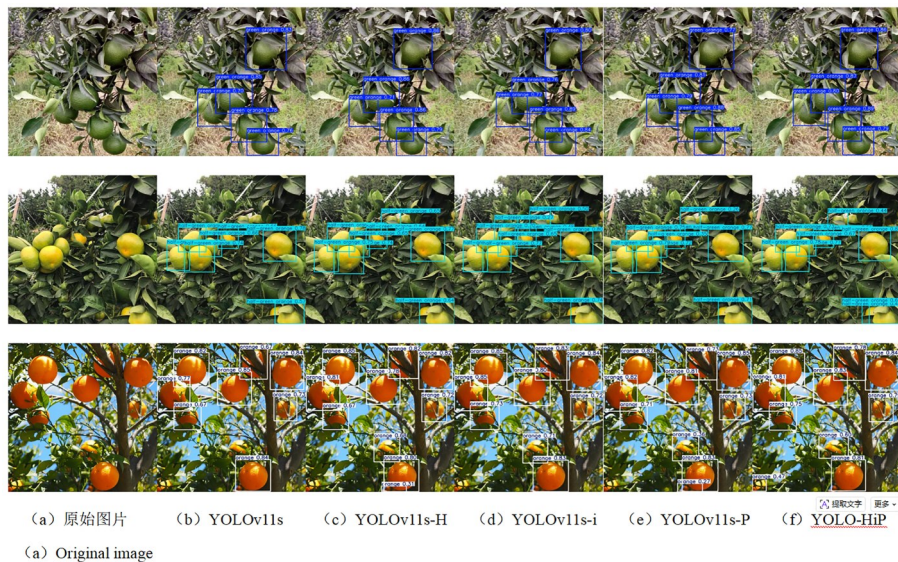


图9 消融实验检测结果对比图

Fig. 9 Comparison of ablation experiment detection results

干网络末端引入C2PSA_iRMB模块后,模型mAP₅₀提升至0.926,较基线提高3.0%,验证了该模块在长距离依赖建模与关键信息捕获方面的有效性,对检测精度提升具有显著贡献。在采用了C3k2_PConv卷积模块后,模型mAP₅₀提升至0.914,提升1.8%,同时参数量减少12.8%,GFLOPs降低1.2G,表明该模块在显著压缩模型规模的同时仍能保持较高检测精度,在轻量化与性能之间实现了良好平衡,并有效削减了特征图中的冗余信息。

综合实验结果,YOLO-HiP同时引入改进后的HGNetv2-L主干网络、C2PSA_iRMB模块与C3k2_PConv结构后,模型mAP₅₀达到94.3,较基线整体提升4.7%,参数量降至5.1M,减少

了45.7%,GFLOPs降低7.4G,推理速度提升45.6FPS。结合图9的检测可视化结果可知,YOLO-HiP在枝叶颜色相近、果实密集遮挡及强曝光等复杂场景条件下,对不同成熟度的柑橘目标均表现出更加稳定且准确可靠的检测能力。综上,所提出的YOLO-HiP在显著降低模型复杂度的同时,仍能有效提升检测精度,真正实现了模型轻量化与高性能的统一与兼顾,为其在资源受限设备上的实际部署、稳定运行与长期应用提供了更有力支撑,并有助于降低系统维护与迭代成本。

4.4 不同目标检测算法对比试验

为全面评估本文所提出YOLO-HiP算法在树上柑橘成熟度检测任务中的性能优势,本文将与其与多种具有代表性的先进目标检测方法进行

了对比实验,包括 Faster R-CNN^[27]、RT-DETR^[28]、YOLOv12s^[29]、文献[15]、文献[30]及文献[31]中提出的方法。通过横向对比不同模型

在检测精度、计算复杂度与推理效率等方面的表现,系统验证了YOLO-HiP在复杂自然环境下的综合性能优势。实验结果如表4所示。

表4 树上柑橘成熟度分类数据集上对比实验

Tab. 4 Comparative experiments on the citrus maturity classification dataset

模型	mAP ₅₀ /%	mAP _{50:95} /%	Para/M	GFLOPs/G	FPS
YOLOv11s	89.6	47.8	9.4	21.3	181.8
Faster R-CNN	84.7	39.9	136.7	401.7	31.4
RT-DETR	92.0	55.4	20.6	60.8	71.6
YOLOv12s	92.7	57.0	9.1	19.3	144.9
文献[15]	91.2	49.6	18.9	45.0	90.4
文献[30]	88.2	42.9	36.5	103.2	92.2
文献[31]	90.5	49.6	2.3	21.6	76.9
YOLO-HiP	94.3	59.8	5.1	13.9	227.4

与主流单阶段检测算法 Faster R-CNN、RT-DETR、YOLOv11s 和 YOLOv12s 相比, YOLO-HiP 的 mAP₅₀ 分别提升 9.6%、2.3%、4.7% 和 1.6%, mAP_{50:95} 分别提升 19.9%、4.4%、12.0% 和 2.8%。与此同时,模型参数量分别减少 96.3%、75.2%、45.7% 和 44.0%,GFLOPs 分别降低 387.8 G、46.9 G、7.4 G 和 5.4 G,帧率分别提升 196.0 帧/秒、155.8 帧/秒、45.6 帧/秒和 82.5 帧/秒。表明 YOLO-HiP 在平均精度、模型参数规模、计算复杂度及检测速度等关键指标上均表现出更优的综合性能。进一步与最新提出的文献[15]和文献[30]中的改进方法对比, YOLO-HiP 的 mAP₅₀ 分别提高 3.1% 和 6.1%, mAP_{50:95} 分别提升 10.2% 和 16.9%,同时模型参数量分别减少 73.0% 和 86.0%,GFLOPs 分别下降 31.1 G 和 89.3 G,帧率分别提升 137.0 帧/秒和 135.2 帧/秒,显示出更优的精度与复杂度之间的权衡能力。最后,与最新文献[31]中的改进方法对比, YOLO-HiP 的 mAP₅₀ 提高 3.8%, mAP_{50:95} 提升 10.2%,同时 GFLOPs 下降 7.7 G,帧率提升 150.5 帧/秒,不仅显示出其更强的检测精度,而且表明在实时处理能力上也有了显著的提高。综合对比结果表明,在所有参评算法中, YOLO-HiP 在 mAP₅₀ 和 mAP_{50:95} 指标上均取得最高值,分别达到 94.3% 和 59.8%,且模型参数量仅为 5.1 M,计算量仅为 13.9 GFLOPs,显著小于其他对比模型。同时,其检测速度也达

到最高的 227.4 FPS,能够充分满足树上柑橘成熟度检测对实时性的要求。

结合文献[15]、[30]和[31]的实验结果,本文进一步从模型改进的核心策略层面对其与 YOLO-HiP 之间的差异进行分析。文献[15]、[30]与[31]分别从不同技术路径对目标检测模型进行了优化:文献[15]基于 YOLOv8s,通过引入 HAT 模块与 FASFF 检测头,从增强全局建模能力与多尺度特征融合角度显著提升柑橘成熟度识别精度,但以较高参数量与计算开销为代价,体现了“性能优先”的优化思路;文献[30]基于 YOLOv5,采用 FECov 与 FEACov 卷积核强化局部特征提取能力,并结合多模态融合实现生育期分类,在提升检测精度的同时拓展了任务能力,但同样带来较高的模型复杂度,体现了“特征增强与多模态融合”的技术路径;文献[31]则基于 YOLOv11s,通过 EMA 注意力机制替换、CSPPC 模块重构以及 MPDIoU 损失函数优化,侧重减少卷积运算以实现模型极致轻量化,但在检测精度方面存在一定局限,体现了“极致压缩”的设计取向。相比之下, YOLO-HiP 通过改进主干网络,并引入 C2PSA_iRMB 注意力模块与 C3k2_PConv 轻量化结构,在强化全局特征建模能力的同时有效降低参数量与计算开销,实现了检测精度与推理效率的协同提升,体现了“精度—效率平衡”的优化理念。综合上述实验结果与分析可以看出,本文提出的 YOLO-HiP 算法在显

著提升检测精度的同时,能够有效降低模型参数规模与计算复杂度,因而在模型轻量化设计、边缘设备部署可行性以及应用成本控制等方面均表现出较为明显的综合优势。

4.5 算法检测效果展示

为进一步验证本文所提出 YOLO-HiP 算法的有效性与实际检测性能,将其在测试集上的检测结果与 YOLOv11s、YOLOv12s、文献[15]、文献[30]及文献[31]中方法进行了对比分析,从可视化角度评估各算法在复杂自然环境下对树上柑橘成熟度目标的检测效果差异。不同算法检测结果 mAP 曲线如图 10 所示,不同算法检测效果对比如图 11 所示。

由图 10 可知, YOLO-HiP 的整体检测性能显著优于其他对比算法。其在训练初期 mAP 提

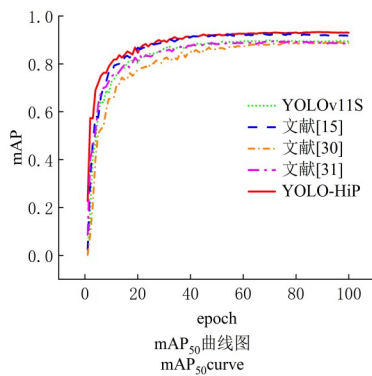
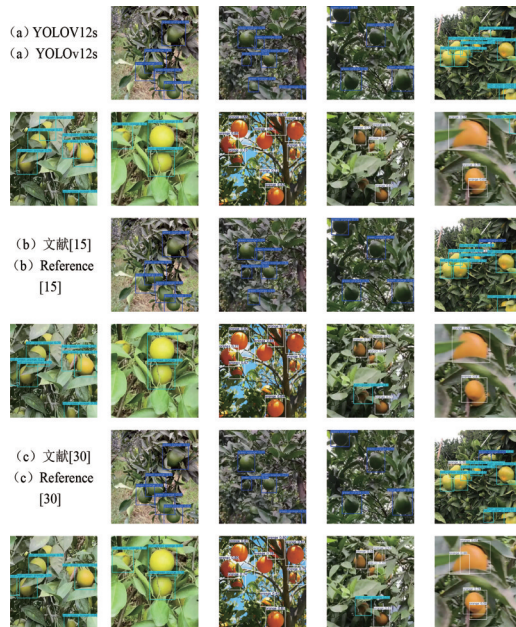


图 10 算法检测结果 mAP 曲线图

Fig. 10 The mAP curve of the algorithm detection results

升较快、收敛速度更高,说明改进模块增强了特征表达能力并提高了训练效率;在训练后期, mAP 稳定保持最高水平,取得了最优检测精度。同时,曲线波动较小、变化平稳,表明模型具有良好的稳定性与鲁棒性,能够适应复杂自然环境。综上, YOLO-HiP 不仅整体检测性能优于基线方法,相较于其余三种相关方向的最新算法也表现出明显优势,验证了所提改进策略的有效性。

由图 11 中多组可视化对比结果可见,不同算法在树上柑橘成熟度检测任务中均能完成基本目标识别,但在复杂自然场景下的检测稳定性存在明显差异。对于果面未转色、果面未 2/3 转色和果面已 2/3 转色等不同成熟度目标, YOLO-HiP 在多数场景中均表现出更好的检测完整性与类别判别一致性。相较于 YOLOv12s 及文献[15]、文献[30]、文献[31]方法, YOLO-HiP 在枝叶遮挡、果实密集分布、尺度变化以及多成熟度目标共存条件下,能够更准确地定位目标区域,并有效减少漏检、误检及类别混淆现象。尤其在成熟度过渡阶段样本中,由于不同类别之间外观差异较小、局部颜色变化不均,部分对比方法容易出现识别偏差,而 YOLO-HiP 仍能保持较好的分类稳定性;在小目标及边缘目标场景下,其检测框覆盖更完整,对弱显著性目标的响应更充分。综合来看, YOLO-HiP 不仅在简单场景下具备可靠检测能力,而且在复杂背景和细粒度成熟



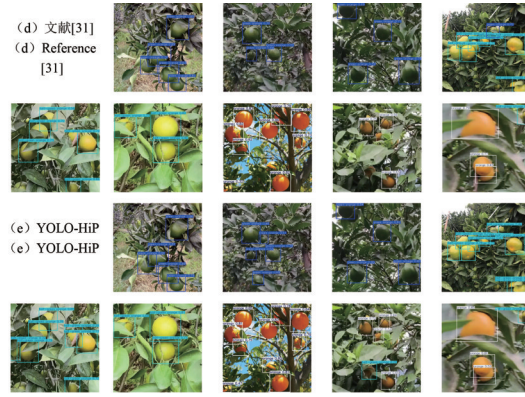


图 11 算法检测效果对比图

Fig. 11 Comparison of algorithm detection results

度区分任务中展现出更强的鲁棒性与泛化能力,说明本文提出的轻量化主干网络、混合注意力机制以及高效卷积结构能够有效提升模型对复杂果园环境中柑橘成熟度目标的表征与识别能力。

4.6 不同成熟度类别的检测性能分析

为分析不同成熟度类别的检测差异,本文统计了YOLOv11s与YOLO-HiP在green orange、half-green orange和orange三类上的Precision、Recall、 mAP_{50} 和 $mAP_{50:95}$,结果如表5所示。从基线模型结果看,orange类检测性能最佳,其Precision、 mAP_{50} 和 $mAP_{50:95}$ 分别为87.6%、91.6%和48.8%,说明成熟果实颜色特征更明显、与背景区分度更高,因此更易识别。相比之下,green orange与half-green orange的检测性能较低。其中,green orange的 $mAP_{50:95}$ 仅为46.2%,主要由于其颜色与枝叶背景接近,易受复杂光照和遮挡影响,导致边界定位不准;half-green orange的Recall为86.3%,但Precision仅为81.7%,表明该类虽较易检出,却更容易误检。这是因为其处

于成熟度过渡阶段,果面转色比例和分布差异较大,类内差异明显,且与green orange和orange两类相似性较强,更易产生类别混淆。

引入YOLO-HiP后,各类别性能均有所提升。对于green orange, Precision、Recall、 mAP_{50} 和 $mAP_{50:95}$ 分别提升5.6、5.7、5.2和7.2个百分点,说明改进模型对低对比度目标和复杂背景具有更强鲁棒性。对于half-green orange, mAP_{50} 和 $mAP_{50:95}$ 分别提升至92.8%和53.9%,表明改进模型能够有效缓解过渡态果实的类别混淆问题。对于orange,改进效果最为显著,其 $mAP_{50:95}$ 提升23.2个百分点至72.0%,说明YOLO-HiP在显著目标上的定位精度提升更明显。

总体来看,green orange的识别难点主要在于背景相似导致的低可分性, half-green orange的难点主要在于过渡态样本的类内差异和类间混淆,而orange整体最易识别。YOLO-HiP在三类目标上均取得稳定增益,说明其对不同成熟度柑橘具有较好的适应能力和泛化性能。

表 5 不同成熟度类别的检测性能分析

Tab. 5 Detection performance analysis Across different maturity categories

Categories	Models	P/%	R/%	$mAP_{50}/\%$	$mAP_{50:95}/\%$
Green orange	YOLOv11s	82.1	84.0	88.6	46.2
Green orange	YOLO-HiP	87.7	89.7	93.8	53.4
Half-green orange	YOLOv11s	81.7	86.3	88.7	48.4
Half-green orange	YOLO-HiP	85.7	85.7	92.8	53.9
Orange	YOLOv11s	87.6	83.9	91.6	48.8
Orange	YOLO-HiP	94.3	92.3	96.4	72.0

4.7 不同数据集结果对比

为进一步验证所提模型的泛化能力,本文进一步选取了由 University of Texas at Arlington 的 Robotic Vision Lab 与美国农业部 USDA Agricultural Research Service 相关研究人员联合发布的 CitDet 数据集,以及由奥地利 Graz University of Technology (TU Graz) 研究团队发布的 FruitVegDB 数据集,对模型进行补充评估。考虑到 CitDet 数据集原始样本数量较少,难以充分反映模型性能,本文首先对其进行了数据增强,再开展相关实验;同时,从 FruitVegDB 中筛选出柑橘相关子集,并重新标注成熟度检测

框及成熟度类别,在相同训练策略下完成实验。实验结果如表 6 所示。结果表明,在采用真实果园场景采集且果实尺度变化较大的 CitDet 数据集上,改进模型的 mAP_{50} 提升了 3.7%, $mAP_{50:95}$ 提升了 7.1%;而在具有层级化细粒度分类特征、且与移动终端应用场景高度相关的 FruitVegDB 数据集上,尽管图像在白平衡、色彩还原、锐化强度及噪声控制等方面存在一定差异,模型的 mAP_{50} 仍提升了 3.9%, $mAP_{50:95}$ 提升了 10.3%。上述实验结果进一步表明,改进后的模型在不同数据分布和图像特征条件下均表现出良好的泛化能力。

表 6 不同数据集的检测结果

Tab. 6 Detection results on different datasets

Datasets	Models	$mAP_{50}/\%$	$mAP_{50:95}/\%$	Para/M	GFLOPs/G
Ours	YOLOv11s	89.6	47.8	9.4	21.3
Ours	YOLO-HiP	94.3	59.8	5.1	13.9
CitDet	YOLOv11s	84.6	47.2	9.4	21.3
CitDet	YOLO-HiP	88.3	54.3	5.1	13.9
FruitVegDB	YOLOv11s	91.2	52.1	9.4	21.3
FruitVegDB	YOLO-HiP	95.1	62.4	5.1	13.9

5 结 论

在自然环境中进行树上柑橘成熟度检测时,面临着在特征提取、多尺度信息融合和计算效率之间难以实现理想平衡的挑战。为解决这一问题,本文以 YOLOv11s 为基础进行多项针对性的结构创新与策略优化,提出一种兼具轻量化与高检测效率的树上柑橘成熟度检测方法 YOLO-HiP。首先,采用重新构建的 HGNetv2-L 轻量化网络替代 YOLOv11s 的主干网络,通过层次化特征提取方法,提升了网络处理复杂图像数据的能力,显著降低了计算复杂度。其次,设计了轻量化混合注意力机制 C2PSA_iRMB,结合 C2PSA 和 iRMB,减少计算开销的同时增强了长距离信息处理能力,提升了检测精度和鲁棒性。最后,结合 PConv 和 C3k2 模块,提出了新的 C3k2_PConv 结构,进一步减少冗余计算和内存访问,优化了空间特征提取效率。

实验结果表明,相较于 YOLOv11s 基准模

型, YOLO-HiP 模型在减少 45.7% 的参数量和 34.7% 的 GFLOPs 的同时,帧率提升了 45.6 帧/秒,且 mAP_{50} 提升了 4.7 个百分点。该模型仅需 5.1M 的参数量和 13.9 GFLOPs,显著优化了计算效率和检测精度。通过减小参数量和计算复杂度, YOLO-HiP 实现了轻量化设计,使其更加适用于计算资源受限的环境。与此同时,在复杂自然环境下, YOLO-HiP 展现出优异的树上柑橘成熟度识别能力,尤其在逆光和部分遮挡情况下,能够准确识别柑橘目标,有效降低漏检率。未来可进一步面向嵌入式设备与智能采摘平台开展模型部署优化研究,以提升方法的工程适配性与实时应用能力,从而推动其在智慧农业中的落地与推广。除柑橘成熟度检测外,本文方法在其他农业视觉任务及资源受限场景中同样具有应用潜力,可扩展至作物成熟度识别、果实计数、病害检测及采摘定位等任务。与此同时,本文提出的轻量化结构设计也可为工业缺陷检测、物流分拣识别和移动端视

觉感知等任务提供参考,具有一定的推广价值 与应用前景。

参 考 文 献:

- [1] 李玉萍,叶露,梁伟红,等. 我国柑橘类“一村一品”发展现状及对策研究[J]. 中国南方果树, 2025, 54(05): 205-210+217.
LI Y P, YE L, LIANG W H, *et al.* Research on the current situation and countermeasures of ‘One Village, One Product’ for citrus in China[J]. *South China Fruits*, 2025, 54(05): 205-210+217. (in Chinese)
- [2] 李泽慧,张琳,山显英,等. 基于多尺度时序建模与动态空间特征融合的视频摘要模型[J]. 液晶与显示, 2025, 40(11):1729-1743.
LI Z H, ZHANG L, SHAN X Y, *et al.* Video summarization model based on multi-scale temporal modeling and dynamic spatial feature fusion[J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(11): 1729-1743. (in Chinese)
- [3] LI L, SHUANG Y, MA Q, *et al.* Intelligent metasurface imager and recognizer[J]. *Light: science & applications*, 2019, 8(1): 97.
- [4] JI C, ZHANG J, YUAN T, *et al.* Research on key technology of truss tomato harvesting robot in greenhouse[J]. *Applied Mechanics and Materials*, 2014, 442: 480-486.
- [5] JALALI B, MACPHEE C. VEVID: Vision enhancement via virtual diffraction and coherent detection[J]. *eLight*, 2022, 2(1): 24.
- [6] 王伯霄,宋延嵩,董小娜. 难点注意力感知红外小目标检测网络[J]. 中国光学(中英文), 2024, 17(03): 538-547.
WANG B X, SONG Y S, DONG X N. Indistinguishable points attention-aware network for infrared small object detection[J]. *Chinese Optics*, 2024, 17(3): 538-547. (in Chinese)
- [7] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. *arXiv preprint arXiv:2004.10934*, 2020.
- [8] MAI X, ZHANG H, JIA X, *et al.* Faster R-CNN With Classifier Fusion for Automatic Detection of Small Fruits [J]. *IEEE Transactions on Automation Science and Engineering*, 2020, PP(99): 1-15.
- [9] 朱旭, 马溪, 姬江涛, 等. 基于 Faster R-CNN 的蓝莓冠层果实检测识别分析[J]. 南方农业学报, 2020, 51(06): 1493-1501.
ZHU X, MA H, JI J T, *et al.* Detecting and identifying blueberry canopy fruits based on Faster R-CNN[J]. *Journal of Southern Agriculture*, 2020, 51(6): 1493-1501. (in Chinese)
- [10] 杨彦萍, 高军伟, 刘兆龙, 等. 基于改进 YOLO11 的印刷电路板表面缺陷检测算法[J]. 液晶与显示, 2025, 40(10):1545-1556.
YANG Y P, GAO J W, LIU ZH L, *et al.* Surface defect detection on printed circuit boards based on improved YOLO11[J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(10): 1545-1556. (in Chinese)
- [11] 山宏刚,倪奕麟,蔡建刚,等. FFD-YOLO:面向小目标与复杂背景的液晶屏缺陷检测[J]. 液晶与显示, 2026, 41(02):208-221.
SHAN H G, NI Y L, CAI J G, *et al.* FFD-YOLO: LCD screen defect detection for small targets and complex backgrounds[J]. *Chinese Journal of Liquid Crystals and Displays*, 2026, 41(2): 208 - 221. (in Chinese)
- [12] 吕石磊, 卢思华, 李震, 等. 基于改进 YOLOv3-LITE 轻量级神经网络的柑橘识别方法[J]. 农业工程学报, 2019, 35(17): 205-214.
LV S L, LU S H, LI ZH, *et al.* Orange recognition method using improved YOLOv3-LITE lightweight neural network [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2019, 35(17): 205 - 214. (in Chinese)
- [13] ZHENG Z, XIONG J, LIN H, *et al.* A method of green citrus detection in natural environments using a deep convolutional neural network[J]. *Frontiers in Plant Science*, 2021, 12: 705737.
- [14] LYU S, LI R, ZHAO Y, *et al.* Green citrus detection and counting in orchards based on YOLOv5-CS and AI edge system[J]. *Sensors*, 2022, 22(2): 576.
- [15] 李莹, 刘梦莲, 何自芬, 等. 基于改进 YOLOv8s 的柑橘果实成熟度检测[J]. 农业工程学报, 2024, 40(24): 157-164.
LI Y, LIU M L, HE Z F, *et al.* Detecting citrus fruit maturity using improved YOLOv8s[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2024, 40(24): 157-164. (in Chinese)

- [16] LAWAL O M, ZHU S, CHENG K. An improved YOLOv5s model using feature concatenation with attention mechanism for real-time fruit detection and counting[J]. *Frontiers in Plant Science*, 2023, 14: 1153505.
- [17] 高新阳, 魏晟, 温志庆, 等. 改进 YOLOv5 轻量级网络的柑橘检测方法[J]. *计算机工程与应用*, 2023, 59(11): 212-221.
GAO X Y, WEI SH, WEN ZH Q, *et al.* Citrus Detection Method Based on Improved YOLOv5 Lightweight Network[J]. *Computer Engineering and Applications*, 2023, 59(11): 212-221. (in Chinese)
- [18] 杨如强, 赵霞, 张鑫. 基于改进 YOLOv11 模型的柑橘叶片病害检测[J]. *江苏农业学报*, 2026, 42(01): 99-109.
YANG R Q, ZHAO X, ZHANG X. An improved YOLOv11 model for citrus leaf disease detection[J]. *Jiangsu Journal of Agricultural Sciences*, 2026, 42(01): 99-109. (in Chinese)
- [19] 胡一波, 陈婷婷, 黄子涛, 等. 基于改进 YOLOv11 的柑橘成熟度检测方法[J]. *大众科技*, 2025, 27(06): 9-13.
HU J B, CHEN T T, HUANG Z T, *et al.* Citrus maturity detection method based on an improved YOLOv11[J]. *Popular Science & Technology*, 2025, 27(06): 9-13. (in Chinese)
- [20] 罗一鸣, 达列雄, 张鹏超, 等. 基于改进 YOLOv11n 的复杂环境下柑橘识别[J]. *农业工程学报*, 2026, 42(3): 253-262.
LUO Y M, DA L X, ZHANG P CH, *et al.* Recognizing citrus in complex environments using improved YOLOv11n[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2026, 42(3): 253-262. (in Chinese)
- [21] KHANAM R, HUSSAIN M. Yolov11: An overview of the key architectural enhancements[J]. *arXiv preprint arXiv:2410.17725*, 2024.
- [22] ZHAO Y, LV W, XU S, *et al.* Detsr beat yolos on real-time object detection[C]//*Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2024: 16965-16974.
- [23] ZHANG J, LI X, LI J, *et al.* Rethinking mobile block for efficient attention-based models[C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE Computer Society, 2023: 1389-1400.
- [24] CHEN J, KAO S, HE H, *et al.* Run, don't walk: chasing higher FLOPS for faster neural networks[C]//*Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2023: 12021-12031.
- [25] JAMES J A, MANCHING H K, MATTIA M R, *et al.* Citdet: A benchmark dataset for citrus fruit detection[J]. *IEEE Robotics and Automation Letters*, 2024, 9(12): 10788-10795.
- [26] 农业农村部. 柑橘采摘技术规范: NY/T 716-2003[S]. 北京: 中国标准出版社, 2004.
- [27] REN S, HE K, GIRSHICK R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2016, 39(6): 1137-1149.
- [28] ZHAO Y, LV W, XU S, *et al.* Detsr beat yolos on real-time object detection[C]//*Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2024: 16965-16974.
- [29] TIAN Y, YE Q, DOERMANN D. Yolov12: Attention-centric real-time object detectors[J]. *arXiv preprint arXiv:2502.12524*, 2025.
- [30] 黄蕊, 李贤辉, 杨彦鑫, 等. 基于 FEACConv-YOLOv5 的柑橘果实目标检测与识别研究[J]. *江苏农业科学*, 2025, 53(05): 54-61.
HUANG R, LI X H, YANG Y X, *et al.* Citrus fruit target detection and recognition based on FEACConv-YOLOv5 [J]. *Jiangsu Agricultural Sciences*, 2025, 53(5): 54-61. (in Chinese)
- [31] LIAO Y, LI L, XIAO H, *et al.* YOLO-MECD: citrus detection algorithm based on YOLOv11[J]. *Agronomy*, 2025, 15(3): 687.

作者简介:



王文坤, 男, 硕士研究生, 2023 年于安阳工学院获得学士学位, 主要从事机器视觉及目标检测研究。E-mail: 18738090147@163.com



谢辉, 女, 硕士, 教授, 2007 年于重庆大学获得硕士学位, 主要从事人工智能及图像处理方面的研究。E-mail: 20030041@sanxiau.edu.cn