doi:10.6041/j.issn.1000-1298.2026.18.028

基于 DPAA - YOLO v8 的复杂逆光场景下柑橘检测方法

李善军[1,2]　宋具辉[1]　吴正开[1]　马佳微[3]　余勇华[1]　鲍秀兰[1,2]

(1. 华中农业大学工学院, 武汉 430070; 2. 农业农村部长江中下游农业装备重点实验室, 武汉 430070;
3. 华中农业大学植物科学技术学院, 武汉 430070)

摘要: 针对自然果园环境中强逆光与果实密集分布、枝叶遮挡等深度耦合导致的目标纹理丢失及多尺度特征混淆问题,本文提出一种面向复杂逆光场景的柑橘检测模型 DPAA - YOLO v8。首先,该模型以 YOLO v8n 为基准在骨干网络设计双路融合交叉部分(Dual-path fusion cross stage partial,DPFCSP)模块,通过双分支并行处理策略分别捕捉上下文依赖与局部细节,显著增强模型在逆光暗区及过曝区域的纹理特征提取能力;其次,在颈部网络引入动态自适应注意力(Dynamic adaptive attention,DAA)机制,通过感知输入特征的语义上下文动态调整卷积核权重与感受野大小,有效聚焦前景果实并抑制逆光产生的高频背景噪声;最后,采用 Wise - IoU(WIoU)损失函数优化边界框回归策略,解决由光晕效应及阴影干扰引起的目标定位漂移问题。试验结果表明,DPAA - YOLO v8 模型平均精度均值(mAP@0.5)达到 96.8%,较 YOLO v8n 提升 3.0 个百分点;小目标(AP_s)与大目标(AP_l)检测精度分别达到 57.2% 与 90.9%,较 YOLO v8n 分别提升 7.8、5.9 个百分点。田间机器人集成试验结果表明,系统在强逆光环境下定位成功率为 92.5%,平均推理耗时 29.2 ms。本文方法在保证轻量化与实时性的前提下显著提升了复杂光照环境下检测鲁棒性,可为柑橘自动化采收提供可靠的技术支持。

关键词: 柑橘检测; 复杂逆光场景; DPAA - YOLO v8; 双路融合; 动态自适应注意力

中图分类号: S24　　**文献标识码**: A　　**文章编号**: 1000-1298(2026)18-0336-10　　**OSID**:

Citrus Detection Method in Complex Backlight Scenarios Based on DPAA - YOLO v8

Li Shanjun[1,2]　Song Juhui[1]　Wu Zhengkai[1]　Ma Jiawei[3]　Yu Yonghua[1]　Bao Xiulan[1,2]

(1. *College of Engineering*, *Huazhong Agricultural University*, *Wuhan* 430070, *China*
2. *Key Laboratory of Agricultural Equipment in Mid-lower Yangtze River*, *Ministry of Agriculture and Rural Affairs*, *Wuhan* 430070, *China*
3. *College of Plant Science and Technology*, *Huazhong Agricultural University*, *Wuhan* 430070, *China*)

Abstract: Aiming to address the problems of target texture loss and multi-scale feature confusion caused by the deep coupling of strong backlight, dense fruit distribution, and branch occlusion in natural orchard environments, a citrus detection model named DPAA - YOLO v8 for complex backlight scenarios was proposed. Based on YOLO v8n, a dual-path fusion cross stage partial (DPFCSP) module was designed in the backbone network. By employing a dual-branch parallel processing strategy to capture context dependencies and local details respectively, the module significantly enhanced the texture feature extraction capability in backlit dark areas and overexposed regions. Secondly, a dynamic adaptive attention (DAA) mechanism was introduced into the neck network. This mechanism dynamically adjusted convolution kernel weights and receptive field sizes by perceiving the semantic context of input features, effectively focusing on foreground fruits and suppressing high-frequency background noise caused by backlight. Finally, the Wise - IoU (WIoU) loss function was adopted to optimize the bounding box regression strategy, resolving the target localization drift caused by halo effects and shadow

收稿日期: 2026 - 02 - 05　修回日期: 2026 - 04 - 28

基金项目: 国家重点研发计划项目(2024YFD2300800)、湖北省重点研发计划项目(2024BBB060)和中央高校基本科研业务费专项资金项目(2662024GXPY006)

作者简介: 李善军(1977—),男,教授,博士,主要从事水果生产机械化技术与智能装备研究,E-mail: shanjunlee@mail.hzau.edu.cn

通信作者: 鲍秀兰(1984—),女,副教授,博士,主要从事果园机器人结构设计与控制研究,E-mail: orchidbaoxl@mail.hzau.edu.cn

interference. Experimental results showed that the mean average precision (mAP@0.5) of the DPAA－YOLO v8 model reached 96.8%, an improvement of 3.0 percentage points over the original YOLO v8n. The detection precision for small targets (AP_s) and large targets (AP_l) reached 57.2% and 90.9%, respectively, representing improvements of 7.8 percentage points and 5.9 percentage points compared with that of the original YOLO v8n. Field robot integration tests indicated that the system's positioning success rate in strong backlight environments was 92.5%, with an average inference time of 29.2 ms. The proposed method significantly improved detection robustness under complex lighting conditions while ensuring lightweight and real-time performance, providing reliable technical support for automated citrus harvesting.

Key words: citrus detection; complex backlight scenarios; DPAA－YOLO v8; dual-path fusion; dynamic adaptive attention

0　引言

柑橘作为全球重要的经济作物，实现其机器人自动化采摘对推动现代智慧农业发展具有重要的战略意义[1－3]。视觉识别系统作为采摘机器人的感知核心，其在非结构化环境中的准确性与鲁棒性直接决定了作业效率[4－5]。随着深度学习技术的发展，以 YOLO 系列和 Faster R－CNN 为代表的卷积神经网络模型，已在农业目标检测领域取得了显著应用进展[6－7]。

然而，真实果园作业环境面临着极具挑战性的“逆光-空间”双重耦合干扰。一方面，强逆光导致图像动态范围压缩与纹理丢失；另一方面，果实密植、遮挡等非结构化特征显著。2 种因素的相互叠加往往诱发 4 种典型干扰情形：逆光团簇因边缘光晕融合导致边界模糊、逆光遮挡因阴影区对比度极低造成特征缺失、逆光密集分布产生复杂的明暗交替噪声干扰特征聚焦，以及逆光远景小目标受光照漫射影响易淹没于背景噪声中。上述复杂逆光场景严重制约了现有视觉算法鲁棒性，成为阻碍柑橘采收机器人全天候作业的关键瓶颈。

针对果园非结构化环境下的检测难题，现有研究主要从复杂空间形态与光照环境变化 2 个维度展开。

在应对果实遮挡、密集分布及多尺度等空间形态挑战方面，研究者多致力于优化特征融合网络与引入注意力机制[8－12]。针对枝叶遮挡与果实重叠，文献[14－15]分别在 Faster R－CNN 与 YOLO v5 中引入注意力机制，有效增强了模型对局部遮挡特征的提取能力。为解决密集与小目标检测难点，顾文娟等[15]与 Liu 等[16]通过优化特征金字塔结构及引入最新 YOLO 架构，显著提升了复杂背景下的召回率。此外，Li 等[17]与苗荣慧等[18]亦通过改进骨干网络，增强了模型在复杂背景下的特征表达。

针对果园自然光照变化及颜色干扰，现有研究聚焦于提升模型在逆光、多角度光照及相近色背景下的鲁棒性。Fu 等[19]利用深度滤波技术去除背景噪声，有效改善了逆光场景下的检测精度；朱立成等[20]提出的 SDE－YOLO 模型则针对性解决了不同光照角度及阴影干扰问题。同时，文献[21－22]致力于模型的轻量化改进，以兼顾光照适应性与实时检测需求。

综上所述，尽管现有研究在单一的复杂空间或特定光照场景下均取得一定成果，但针对强逆光干扰与复杂空间分布深度耦合的柑橘检测场景研究尚未开展。为此，本文提出一种面向复杂逆光场景的检测网络（DPAA－YOLO v8）。设计 DPFCSP 模块与 DAA 机制，以解决逆光导致的纹理丢失与多尺度混淆难题；同时结合 WIoU 损失函数，以解决边缘模糊引发的定位漂移问题。以期为柑橘采收机器人实地作业提供可靠技术支持。

1　数据集构建

1.1　数据采集

现有公开柑橘数据集多采于理想光照，缺乏强逆光与复杂空间分布深度耦合的专项数据，制约了模型对逆光团簇、遮挡及小目标等极端特征的学习。为此，本研究构建了包含多种典型复杂逆光场景的专用数据集，以增强模型在非结构化作业环境中的有效性。

为构建适用于复杂逆光场景下高鲁棒性柑橘检测模型，于 2024 年 9 月—2025 年 10 月在华中农业大学现代标准化柑橘果园（30°28′7″N，114°21′35″E，海拔 38.1 m）完成了数据集采集工作，果园环境如图 1 所示，试验品种和数量为温州蜜柑 25 株、南丰蜜柑 30 株、碰柑 22 株和纽荷尔脐橙 19 株，树龄为 8～9 年，种植株距为 1.2～1.8 m，果树行距为 3.8～4 m，平均树高为 2.5 m，该果园能够充分涵盖研究所涉及的各类复逆光场景，满足图像采集需求。图像采集主要在晴朗天气下进行，使用手持式 Intel

D435i 型相机在自然光照环境中以 1 250 万像素进行系统化拍摄。



(a) 果园俯视图 (b) 果园内部视图

图 1 柑橘果园环境示意图

Fig. 1 Citrus orchard environment schematics

为确保数据异质性与代表性,采集工作严格覆盖日间光照变化显著的 3 个关键时段(上午 08:00—11:00、中午 12:30—13:00、下午 14:00—17:00),以捕捉不同太阳高度角下光影特征。针对复杂逆光环境下检测难点,数据集重点囊括了逆光团簇、逆光遮挡、逆光密集分布以及逆光小目标 4 类典型困难样本,记录了柑橘果实在不同光照条件下的视觉形态,代表性样本如图 2 所示。经筛选,构建包含 1 150 幅有效图像的专用数据集。



(a) 团簇 (b) 遮挡

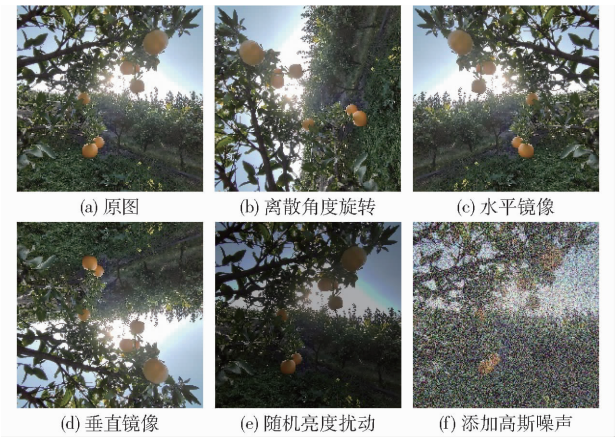
(c) 密集分布 (d) 小目标

图 2 复杂逆光场景典型样本

Fig. 2 Typical samples in complex backlit scenarios

1.2 数据集处理

为缓解小样本场景下深度学习模型面临的过拟合风险,采用系统化的数据增强策略以提升训练集的规模与多样性。通过引入多种图像空间与色彩域的变换方法,包括离散角度旋转(90°、180°、270°)、随机亮度扰动、水平及垂直镜像翻转和添加高斯噪声等操作,实现对原始数据的有效扩增。该策略不仅强化了柑橘图像中全局结构与局部判别性特征的表达,还通过引入足够的类内差异模拟真实场景的复杂性,从而增强深层网络在特征提取与泛化方面的能力。本研究共采集原始有效图像 1 150 幅,经数据增强扩增后,数据集扩充至 3 200 幅,具体分布情况为:逆光遮挡样本 850 幅(占比 26.6%)、逆光团簇样本 780 幅(占比 24.4%)、逆光密集分布样本 720 幅(占比 22.5%)以及逆光小目标样本 550 幅(占比 17.2%),其余为常规光照样本。全部图像按照比例 8∶1∶1分为训练集、验证集和测试集,为模型在复杂逆光果园环境中实现鲁棒识别提供了充分的数据支撑。增强后典型样本示例如图 3 所示。



(a) 原图 (b) 离散角度旋转 (c) 水平镜像

(d) 垂直镜像 (e) 随机亮度扰动 (f) 添加高斯噪声

图 3 数据增强处理示例

Fig. 3 Data augmentation examples

2 模型改进

针对自然果园环境中复杂逆光场景导致的图像对比度低、目标边缘模糊以及强光斑干扰等难题,在 YOLO v8n 基础上提出了一种高性能柑橘检测网络 DPAA - YOLO v8,整体架构如图 4 所示。该网络主要由 Backbone、Neck 和 Head 3 部分构成:①在 Backbone 部分,为了在逆光造成的“暗区”和“过曝区”中有效提取纹理特征,将原有的 C2f 模块替换为本文设计的 DPFCSP 模块,利用其双通路特征融合机制,在保持模型轻量化的同时显著增强了对低对比度特征的捕捉能力。②在 Neck 特征融合阶段,针对逆光条件下多尺度目标易产生的尺度混淆问题,在上采样与特征拼接的关键节点引入 DAA 模块,通过根据输入特征的语义信息动态调整感受野大小,有效抑制了背景光噪干扰。③在 Head 部分,针对逆光光晕导致的目标边缘模糊及定位不准问题,在延续解耦头设计的基础上引入 WIoU 损失函数替代原有的 CIoU,利用其动态非单调聚焦机制智能分配梯度增益,从而确保模型在复杂光照变化下的鲁棒性与高精度检测。

2.1 DPFCSP 模块

在强逆光条件下,柑橘表面色彩信息往往严重丢失,且与背景树叶的对比度极低。传统的卷积操

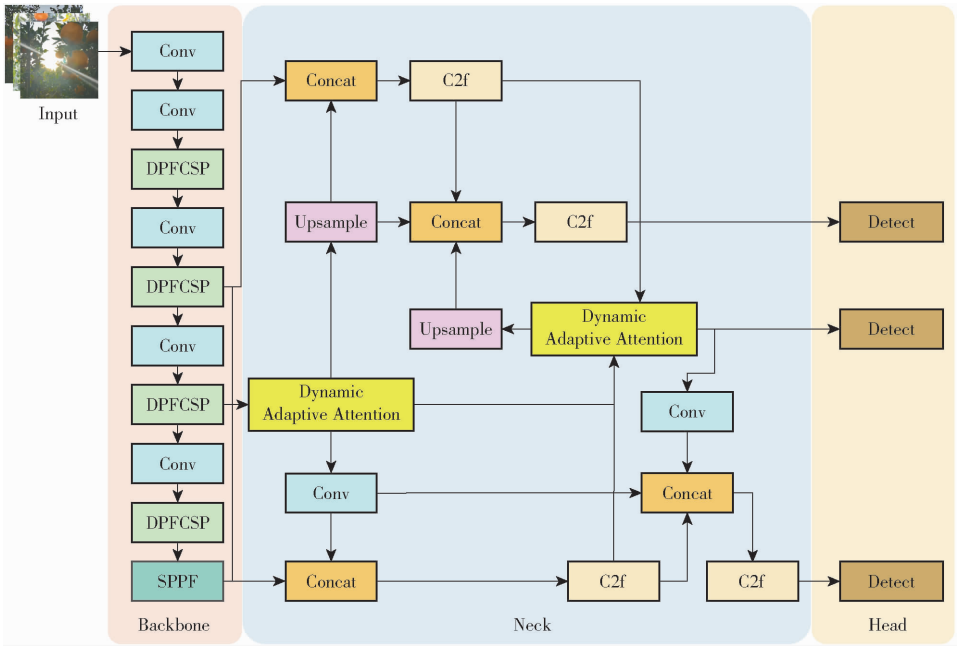


图 4　DPAA－YOLO v8 模型结构

Fig. 4　Structure of DPAA－YOLO v8 model

作在处理此类图像时，往往难以区分目标特征与背景噪声。为了解决这一问题，提出了 DPFCSP 模块，如图 5 所示，核心思想是利用双分支并行处理策略分别捕捉上下文依赖与局部细节。

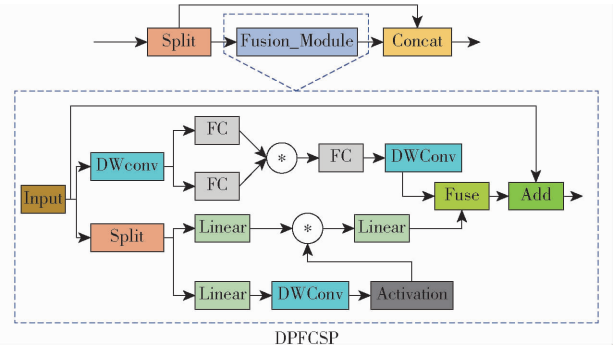


图 5　双路融合交叉模块结构

Fig. 5　Architecture of dual-path fusion cross stage partial module

DPFCSP 模块摒弃了单一路径的特征流向，而是将输入特征 $X \in \mathbf{R}^{C\times H\times W}$ 通过 Split 操作在通道维度上分为 X_{upper}（上分支的输入特征图）和 X_{lower}（空间门控分支的输入特征图）2 部分，分别送入不同处理分支，以降低特征冗余并提高计算效率。

上下文感知分支如图 5 上半部分所示，主要负责“特征重校准”。由于逆光下颜色特征不可靠，需更多地依赖全局上下文信息推断目标类别。该分支通过深度可分离卷积 DWConv 提取初步特征后，利用由全连接层 FC 构成的注意力机制，显式地建模通道间的依赖关系。这类似于种“软阈值”操作，能够自动识别并抑制那些响应强光噪声的通道，同时增强包含柑橘形状纹理的通道权重。其数学表达为

$$Y_{\text{dw}} = \text{DWConv}(X_{\text{upper}}) \tag{1}$$

$$Y_{\text{attn}} = \sigma(\text{FC}_1(Y_{\text{dw}}) \odot \text{FC}_2(Y_{\text{dw}})) \tag{2}$$

$$Y_{\text{upper}} = \text{DWConv}(\text{FC}_{\text{out}}(Y_{\text{attn}})) \tag{3}$$

式中　Y_{dw}——上分支经深度可分离卷积提取的中间特征

Y_{attn}——经通道注意力机制重校准后的特征权重图

Y_{upper}——上下文感知分支处理后的最终输出特征

σ——Sigmoid 函数　　$\odot$——逐元素乘法

通过这种方式，Y_{upper} 成功编码了对逆光环境更具鲁棒性的上下文特征。

如图 5 下半部分所示，空间门控分支引入门控机制，以实现对局部空间细节的精准恢复。在逆光图像中，目标边缘往往与背景光斑混合。该分支利用门控线性单元的变体，通过学习 1 个空间掩码过滤信息。输入特征被分为两路，一路用于生成内容信息，另一路经过激活函数生成“门控信号”。只有当门控信号为高响应时，对应的空间细节才会被保留。有效滤除了树叶间隙的强光干扰，即

$$Y_{\text{gate}} = \text{Linear}(X_{l1}) \odot \sigma(\text{DWConv}(\text{Linear}(X_{l2}))) \tag{4}$$

$$Y_{\text{lower}} = \text{Linear}_{\text{out}}(Y_{\text{gate}}) \tag{5}$$

式中　Y_{gate}——空间门控分支生成的掩码特征，用于抑制背景噪声

Linear——线性投影层

X_{l1}、X_{l2}——下分支特征 X_{lower} 进一步分割后的 2 个子特征流

最终,2 个分支的输出在 Fusion Module 中进行拼接与深度融合,并叠加残差连接,形成富含细节且抗干扰能力强的特征输出Y_{out},即

$$Y_{out} = \mathrm{Add}(\mathrm{Concat}(Y_{upper}, Y_{lower}), X) \tag{6}$$

式中 Add——特征图逐元素相加操作,实现残差连接

Concat——通道拼接操作,将不同分支的特征在通道维度上合并

Y_{out}——DPFCSP 模块的最终输出特征图

2.2 DAA 模块

在复杂果园环境中,拍摄距离变化和光照角度偏移会使得柑橘目标呈现显著的多尺度特性。对于固定尺寸的卷积核,如 3×3,难以同时兼顾对小目标的精细捕捉和对大目标的全局感知。为此,本文设计了 DAA 模块,结构如图 6 所示,赋予网络动态卷积核选择的能力。

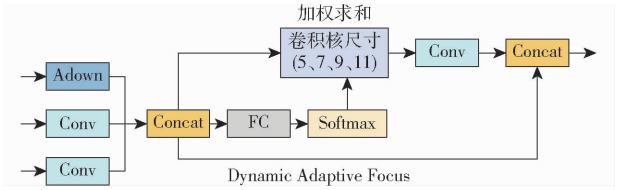


图 6 动态自适应注意力机制结构

Fig. 6 Architecture of dynamic adaptive attention mechanism

该模块工作流程模拟人眼在观察不同距离物体时的变焦过程。具体而言,模块首先汇聚来自 Backbone 和 Neck 的多尺度特征$\{F_{s1}, F_{s2}, F_{s3}\}$,以获得全面视野。随后,通过 1 个轻量级的全连接网络感知输入图像的全局语义,并利用 Softmax 函数生成 1 组归一化自适应权重$\boldsymbol{\alpha}_w$。这组权重表示网络对于当前输入图像的“尺度偏好”,即

$$F_{in} = \mathrm{Concat}(F_{s1}, F_{s2}, F_{s3}) \tag{7}$$

$$\boldsymbol{\alpha}_w = \mathrm{Softmax}(\mathrm{FC}(F_{in})) \tag{8}$$

式中 F_{in}——经拼接融合后的多尺度全局上下文特征

其中,$\boldsymbol{\alpha}_w = (\alpha_5, \alpha_7, \alpha_9, \alpha_{11})$对应于预设的 4 种不同尺寸的卷积核。如果输入包含大面积的模糊光斑或近距离的大目标,网络会自动增大α_{11},以扩大感受野,利用周围环境信息辅助推断;反之,若目标细小清晰,则增大α_5以保留边缘细节。基于生成的权重,模块对静态卷积核组进行加权聚合,生成形态可变的动态卷积核$K_{dynamic}$即

$$K_{dynamic} = \sum_{k \in \{5,7,9,11\}} \alpha_k K_k \tag{9}$$

式中 α_k——对应卷积核尺寸的重要性权重系数,$k \in \{5,7,9,11\}$

K_k——尺寸为 $k \times k$ 的预设静态卷积核

这种软选择机制使得卷积核的参数不再是训练后固定的,而是随着每幅输入图像的内容实时变化的。与传统注意力机制(如 SE 模块)通过逐元素相乘对特征图通道进行显式重加权不同,本文 DAA 机制的“软选择”在数学本质上是对不同感受野的卷积核参数矩阵本身进行加权融合,从而在物理层面实现了网络有效感受野的动态变焦。最后,利用动态核进行特征提取并融合原始信息即

$$F_{out} = \mathrm{Concat}\left(\sum_{k \in \{5,7,9,11\}} \alpha_k \mathrm{Conv}_k(F_{in}; K_k), F_{in}\right) \tag{10}$$

通过这一机制,DAA 模块有效解决了逆光下多尺度目标检测难的问题,显著提升了模型泛化能力。

2.3 基于 WIoU 的边界框回归损失函数

为抑制逆光环境下边缘模糊样本引发的惩罚梯度过载,引入 WIoU 损失函数优化边界框回归策略。WIoU 摒弃了 CIoU 僵化的几何聚合,转而利用基于离群度β的非单调聚焦机制进行梯度增益重分配。其核心距离注意力R_{WIoU}引入梯度截断操作[23],通过限制最小包围框尺寸的梯度回传确保收敛稳定性,具体定义为

$$R_{WIoU} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*}\right) \tag{11}$$

式中 (x, y)——预测框中心点的坐标

(x_{gt}, y_{gt})——真实框中心点的坐标

W_g、H_g——同时包围预测框与对应真实框的最小外接矩形的宽度、高度

结合聚焦系数 r,最终损失L_{WIoU}实现了对样本质量的动态感知,计算公式为

$$L_{WIoU} = rR_{WIoU}L_{IoU} \tag{12}$$

其中

$$r = \frac{\beta}{\delta\alpha^{\beta-\delta}} \tag{13}$$

离群度β为当前样本的基础 IoU 损失与其滑动平均值的比值;$\alpha\delta$为控制聚焦系数峰值分布的超参数。该机制使网络自动规避极其模糊或非常清晰的极端样本,转而专注于优化中等质量目标。这一改进有效降低了逆光环境下因边缘不确定性导致的定位偏差,显著提升了检测模型在复杂光照条件下鲁棒性。

3 试验结果与分析

3.1 试验环境与参数配置

所有模型训练和推理检测环境配置:CPU 为英特尔酷睿 i5 - 12600KF 3.7 GHz,GPU 为 NVIDIA GeForce RTX 3080,运行内存为 32.0 GB RAM,操作系统选用 Windows 10,Python 版本为 3.8.10,CUDA

版本为 11.7，利用 Pytorch 1.10.11 深度学习框架进行训练。训练过程中，设置训练轮次为 200，选用 SGD 作为优化器，初始学习率为 0.01，动量设置为 0.937，输入图像尺寸为 640 像素×640 像素。

3.2　评价指标

采用精确率（Precision，P）、召回率（Recall，R）、平均精度均值（mAP@0.5 和 mAP@0.5:0.95）、模型参数量和浮点运算量作为评价指标[24-25]。

3.3　注意力机制检测性能测试

为验证本文提出的 DAA 机制在复杂逆光场景下的优越性，并解释模型未采用传统注意力机制的原因，开展了不同注意力机制横向对比试验。试验选取在骨干网络中引入 DPFCSP 的改进 YOLO v8n 模型作为基准模型，在颈部网络相同位置（Concat 拼接操作之后）分别嵌入目前主流的注意力机制，包括压缩-激励（Squeeze-and-excitation，SE）、卷积块注意力（Convolutional block attention module，CBAM）以及坐标注意力（Coordinate attention，CA）机制，并与 DAA 机制进行性能对比。试验结果如表 1 所示。

由表 1 可知，在基准模型中嵌入注意力机制均能提升检测性能。然而，相比于传统的静态注意力机制（如 SE 仅关注通道加权，CBAM 与 CA 感受野固定），DAA 机制展现出最优的综合性能，mAP@0.5 达到 95.6%，mAP@0.5:0.95 达 59.8%；同时，DAA 模块参数量低于 CBAM 与 CA，兼顾了高精度与轻量化。这是因为 DAA 机制能够感知输入语义并动态调整有效感受野，在复杂逆光场景下既能精准聚焦微小的远景目标，又能有效抑制高频背景光噪。充分证明了本文将 DAA 作为核心特征融合组件的合理性。

表 1　不同注意力机制检测性能测试结果

Tab.1　Test results of detection performance across different attention mechanisms

注意力模块	mAP@0.5/%	mAP@0.5:0.95/%	参数量
无	94.7	55.5	2.78×10^6
SE	95.1	56.2	2.85×10^6
CBAM	95.4	58.1	2.92×10^6
CA	95.2	57.7	2.88×10^6
DAA	95.6	59.8	2.87×10^6

3.4　消融试验

为了验证 DPAA－YOLO v8 中各改进组件的有效性，在 YOLO v8n 基础上开展了 8 组消融试验。分别考察 DPFCSP 模块、DAA 模块以及 WIoU 损失函数的引入对模型性能的贡献，以评估改进策略合理性。消融试验结果如表 2 所示。

表 2　不同改进模块消融试验结果

Tab.2　Ablation experiment results with different improved modules

模型	精确率/%	召回率/%	mAP@0.5/%	mAP@0.5:0.95/%	参数量	浮点运算量
YOLO v8n	90.8	87.4	93.8	53.3	3.00×10^6	8.1×10^9
YOLO v8n + DPFCSP	91.5	89.3	94.7	55.5	2.78×10^6	7.5×10^9
YOLO v8n + DAA	91.9	90.1	95.4	56.8	3.10×10^6	9.6×10^9
YOLO v8n + WIoU	91.3	88.2	96.1	54.6	3.00×10^6	8.1×10^9
YOLO v8n + DPFCSP + DAA	92.6	91.2	95.6	59.8	2.87×10^6	9.0×10^9
YOLO v8n + DPFCSP + WIoU	92.2	89.6	95.9	57.2	2.77×10^6	7.5×10^9
YOLO v8n + DAA + WIoU	92.4	91.0	95.6	58.1	3.10×10^6	9.6×10^9
DPAA－YOLO v8	93.2	92.3	96.8	62.2	2.88×10^6	9.0×10^9

由表 2 可知，YOLO v8n mAP@0.5 为 93.8%，展现出不错的基础性能。引入 DPFCSP 模块后，模型精确率与 mAP@0.5 分别提升至 91.5% 和 94.7%，且参数量降低约 7.3%，证实了该模块在剔除冗余特征、提升计算效率方面的优势。引入 DAA 模块重点强化了对多尺度目标的捕获能力，使召回率提升至 90.1%，有效降低了漏检率。最终，融合 DPFCSP、DAA 及 WIoU 损失函数的 DPAA－YOLO v8 性能大幅提升：精确率、召回率、mAP@0.5 分别达到 93.2%、92.3% 和 96.8%，mAP@0.5:0.95 提至 62.2%。试验结果充分表明，各改进模块之间具有良好的兼容性与互补性，组合模型在未显著增加计算负担的情况下，显著提升了逆光场景下的检测鲁棒性与定位精度。

类激活图能够将神经网络抽象特征转换为可视化关注热点，是分析模型分类依据的重要工具。利用该技术对“Detect”层上游的卷积层进行了可视化重构，如图 7 所示，并用红色圆圈标记未识别柑橘。通过观察生成的特征热力图，可以清晰地识别出模型高响应区域，从而验证模型在检测决策时的注意力指向。在复杂逆光环境下，原始 YOLO v8n 模型受限于传统卷积固定的几何感受野，往往难以有效平衡由于光照不均导致的“前景过暗”与“背景过曝”问题，极易造成注意力发散，将高亮的枝叶间隙

误判为目标,或因暗部纹理缺失而发生漏检。相比之下,图 7 热力图直观展示了 DPAA - YOLO v8 的改进成效,得益于 DPFCSP 模块对深层语义特征的强化提取,以及 DAA 模块对关键区域的动态聚焦,改进后的模型成功修正了原有的注意力偏差。从图 7 可见,高置信度的激活区域紧密包裹住阴影中的柑橘果实,而背景中的强光干扰被彻底抑制。这一对比有力证明了 DPAA - YOLO v8 弥补了基准模型在极端光照下特征提取能力不足的缺陷。此外,通过观察对小目标果实的检测场景,可以看出本方法生成的类激活图在低像素、低对比度的微小区域依然表现出清晰的聚类响应,展示出 DPAA - YOLO v8 模型在逆光小目标检测方面的显著提升。

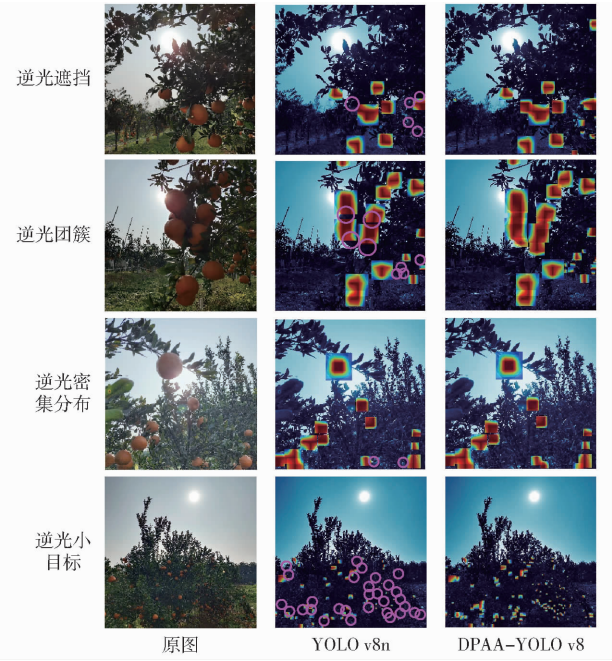


图 7 不同模型热力图结果

Fig. 7 Heatmap results of different models

3.5 改进前后模型性能测试

为评估 YOLO v8 系列模型在目标检测任务中的性能差异,对轻量级模型 YOLO v8n 与综合性能模型 DPAA - YOLO v8 开展了目标检测性能测试。试验以相同软硬件环境,将检测目标按像素面积划分为小目标(s,面积 $<32\times32$)、中目标(m,$32\times32\leqslant$面积$\leqslant96\times96$)、大目标(l,面积 $>96\times96$)3 类,采用细分尺度平均精度(AP_s、AP_m、AP_l)与平均召回率(AR_s、AR_m、AR_l)作为核心评价指标。试验结果如表 3 所示。

表 3 改进前后模型性能测试结果

Tab. 3 Test results of model performance before and after improvement %

模型	AP_s	AP_m	AP_l	AR_s	AR_m	AR_l
YOLO v8n	49.4	66.7	85.0	55.9	70.8	85.0
DPAA - YOLO v8	57.2	79.2	90.9	62.8	82.7	92.5

由表 3 可知,DPAA - YOLO v8 AP_s、AP_m、AP_l分别为 57.2%、79.2%、90.9%,AR_s、AR_m、AR_l分别达到 62.8%、82.7%、92.5%;相较而言,YOLO v8n 分别为 49.4%、66.7%、85.0% 和 55.9%、70.8%、85.0%。由此可见,DPAA - YOLO v8 在小、中、大尺度目标的检测精度与召回率上均呈显著优势,综合检测性能更为突出。

3.6 不同模型性能测试

为探究不同目标检测模型在性能与效率维度的表现差异,选取 YOLO v5n、YOLO v6n、RTDETR、YOLO v8n 及 DPAA - YOLO v8 模型开展试验。试验使用相同的硬件设备和数据集进行训练和测试,训练轮次为 200 轮,输入图像尺寸为 640 像素 ×640 像素。试验结果如表 4 所示。

表 4 不同模型性能测试试验结果

Tab. 4 Test results of different model performance

模型	精确率/%	召回率/%	mAP@0.5/%	mAP@0.5:0.95/%	参数量	浮点运算量
YOLO v5n	88.9	84.7	91.2	48.2	2.500×10^6	7.10×10^9
YOLO v6n	88.9	84.0	91.1	48.0	4.230×10^6	1.18×10^{10}
RTDETR	88.4	86.1	92.3	50.2	1.988×10^7	5.69×10^{10}
YOLO v8n	90.8	87.4	93.8	53.3	3.000×10^6	8.10×10^9
DPAA - YOLO v8	93.2	92.3	96.8	62.2	2.880×10^6	9.00×10^9

由表 4 可知,DPAA - YOLO v8 模型在所有对比模型中表现最优,在精确率、召回率、mAP@0.5 和 mAP@0.5:0.95 均领先。具体而言,DPAA - YOLO v8 的 mAP@0.5 达到 96.8%,相较次优的 YOLO v8n (93.8%)提升 3.0 个百分点;在 mAP@0.5:0.95 上优势较为突出,达到 62.2%,远超 YOLO v8n (53.3%)。表明了本文所集成改进模块的有效性,模型在保持高定位精度的同时,对不同 IoU 阈值下的检测鲁棒性得到了极大增强。在与其他系列模型的对比中,YOLO v5n 与 YOLO v6n 在精度指标上相近,但均低于 YOLO v8n 基线模型。值得注意的是,基于 Transformer 架构的 RTDETR 模型,虽其召回率和 mAP@0.5:0.95 优于 YOLO v5n 和 YOLO v6n,展现了结构上的某些优势,但参数量与

计算复杂度远高于其他对比模型，限制了其在资源受限场景下应用。在模型复杂度方面，DPAA－YOLO v8 展现出优异的效率。其参数量不仅低于 YOLO v8n，更远低于 RTDETR 与 YOLO v6n。尽管计算量因特征融合的增强而略高于 YOLO v5n 和 YOLO v8n，但考虑到其带来的巨大性能提升，这种计算成本的增加是可接受的。这表明 DPAA－YOLO v8 模型成功地在模型精度与计算效率之间取得了更佳的平衡。

3.7　视觉大模型部署对比与讨论

随着视觉大模型发展，Grounding DINO、OWL－ViT 等模型在通用检测中表现优异。为验证本文方法在边缘部署时的选型合理性，在统一硬件平台下使用 500 幅复杂逆光柑橘图像进行了离线对比测试。测试使用模型参数量、mAP@0.5、召回率、单帧平均推理耗时和推理帧率作为评价指标。其中，逆光遮挡、逆光团簇、逆光密集分布和逆光小目标图像占比各为 25%，试验结果如表 5 所示。

表 5　视觉大模型与本文模型推理性能

Tab.5　Inference performance of vision large models and proposed model

模型	参数量	mAP@0.5/%	召回率/%	单帧平均推理耗时/ms	推理帧率/(f·s^{-1})
Grounding DINO	2.32×10^{10}	75.86	65.36	175.68	5.69
OWL－ViT	1.53×10^{10}	70.38	53.66	67.20	14.88
DPAA－YOLO v8	3.01×10^{8}	96.50	91.90	6.83	146.47

由表 5 可知，大模型虽具零样本泛化优势，但其庞大架构带来了较高计算开销。Grounding DINO 参数量达 2.32×10^{10}，桌面端帧率为 5.69 f/s，难以满足边缘设备视觉伺服的实时性刚需（>30 f/s）。此外，面对逆光＋密集等极端场景，大模型因缺乏微弱纹理修复机制，容易出现一定程度的漏检，Grounding DINO 与 OWL－ViT 召回率分别为 65.36% 和 53.66%。相比之下，本文模型参数量仅为 3.01×10^{8}，在召回率与 mAP 上均取得更优表现，且帧率达146.47 f/s，较好地平衡了特定场景高精度与边缘部署的高实时性要求。

3.8　可视化对比试验

为进一步验证 DPAA－YOLO v8 算法的检测性能，选用 YOLO v5n、YOLO v8n 和本文模型，在柑橘检测中进行可视化对比。试验通过随机选择逆光遮挡、逆光团簇、逆光密集分布和逆光小目标的场景下进行测试。选取置信度阈值为 0.5，IoU 阈值为 0.7。试验结果如图 8 所示，图中蓝色矩形框为预测目标框，红色矩形框为漏检目标框。

由图 8a 可看出，当受逆光影响，同时遮挡较严重时，3 种模型均出现漏检情况，其中 YOLO v5n 和 YOLO v8n 漏检 12 个柑橘，而本文模型仅漏检 2 个

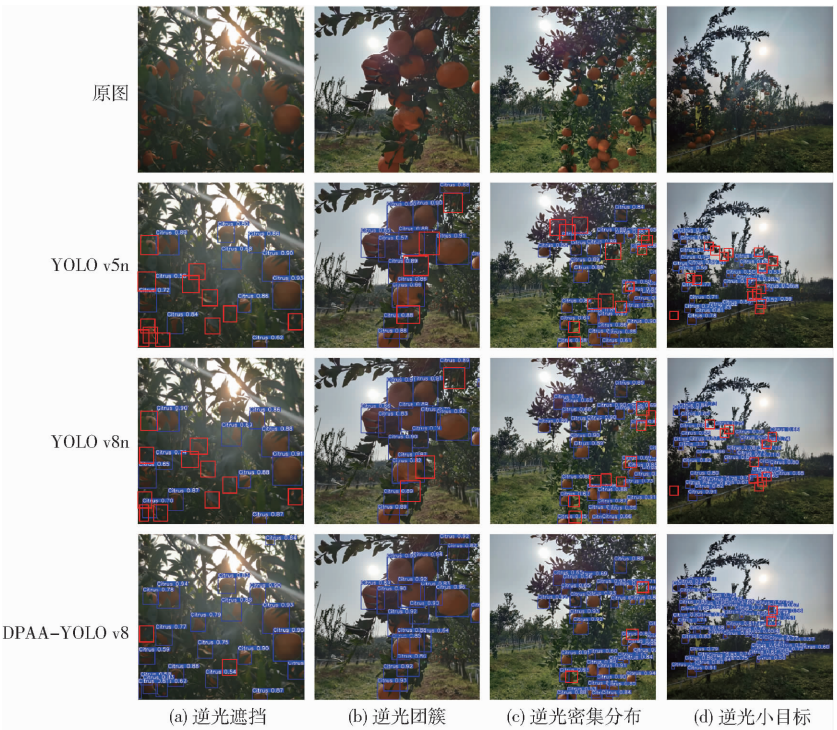


图 8　不同模型检测可视化结果

Fig.8　Visual results of detection by different models

柑橘,说明本文模型在逆光遮挡场景下柑橘目标检测能力具有较强提升;由图 8b 可看出,本文模型在逆光团簇场景下没有出现漏检情况,而 YOLO v5n 漏检 5 个柑橘,YOLO v8n 漏检 3 个柑橘,表明本文模型在此场景下具有较强的检测能力;由图 8c 可看出,YOLO v5n 漏检 13 个柑橘,YOLO v8n 漏检 10 个柑橘,而本文模型漏检 3 个柑橘,可以看出本文模型在此场景下检测能力显著提升;由图 8d 可看出,YOLO v5n 漏检 16 个柑橘,YOLO v8n 漏检 11 个柑橘,本文模型仅漏检 2 个柑橘,同样说明本文模型在逆光小目标检测场景下也具有较强能力提升。

根据各组漏检对比结果可看出,本文模型相较于现有方法在漏检数量上具有显著优势,同时,结合各组检测置信度对比情况可以看出本提出模型检测效果较优。本文模型在鲁棒性与跨场景泛化能力方面均取得显著提升。

4 田间试验

为全面验证 DPAA - YOLO v8 在逆光-空间耦合场景下的作业性能,本研究构建了柑橘采摘机器人试验平台进行田间模型性能验证。选取 200 个典型柑橘样本,划分为晴朗顺光(对照组)及逆光遮挡、逆光团簇、逆光密集分布、逆光小目标(试验组)共 5 个测试组(每组 40 个)进行定位引导试验。图 9 为 2025 年 12 月在华中农业大学柑橘果园验证算法可靠性的试验现场,柑橘采摘机器人搭载了 UFACTORY xarm6 六自由度机械臂、英特尔 D435i 型深度相机、夹剪一体式采摘末端执行器和升降平台等。

结果表明,系统总体定位成功率为 92.5%,平均耗时 29.2 ms。其中,对照组成功率高达 97.5%;逆光遮挡场景受 DPFCSP 模块对暗区特征的增强,成功率达到 95.0%;逆光团簇场景经 WIoU 优化边界回归,成功率为 92.5%;而在极具挑战的逆光密集分布与逆光小目标场景中,DAA 机制虽有效抑制了光噪,但受限于像素占比过小与强背景干扰,成功率分别为 90.0% 与 87.5%。结果表明,尽管极端场景下性能略有波动,但改进的模型整体仍能有效克服强逆光与复杂空间分布的叠加干扰,满足采摘机器人在非结构化环境下的全天候作业需求。



图 9 田间试验

Fig. 9 Field experiments

1. 履带底盘 2. 电气柜 3. 果框 4. 英特尔 D435i 型深度相机 5. 夹剪一体式采摘末端执行器 6. UFACTORY xarm6 六自由度机械臂 7. 升降平台

5 结论

(1)针对逆光-空间双重耦合干扰,提出 DPAA - YOLO v8 检测模型。通过 DPFCSP 模块提取暗区纹理,结合 DAA 动态聚焦特征与 WIoU 优化边界回归,有效克服了多尺度混淆及光斑引起的漏检、误检问题。

(2)DPAA - YOLO v8 在保持轻量化的同时,mAP@0.5 达 96.8%,较 YOLO v8n 提升 3.0 个百分点,且 AP_s 和 AP_l 分别达 57.2% 与 90.9%,综合性能优于同级别主流网络。

(3)田间集成试验结果表明,强逆光环境下平均定位成功率为 92.5%,推理耗时 29.2 ms,满足采摘机器人边缘设备的部署与全天候高精度实时作业需求。同时,尽管本文模型综合性能优异,但在面对极端逆光且目标极微小的场景时,性能仍有一定波动,成功率 87.5%。

参 考 文 献

[1] Yan B, Fan P, Lei X, et al. A real-time apple targets detection method for picking robot based on improved YOLO v5[J]. Remote Sensing, 2021, 13(9): 1619.

[2] 何宇. 四川柑橘产业供给侧结构性改革绩效评价及推进路径研究[D]. 成都: 四川农业大学, 2022.
He Yu. Performance evaluation and promotion path of supply side structural reform of citrus industry in Sichuan[D]. Chengdu: Sichuan Agricultural University, 2022. (in Chinese)

[3] 鲍秀兰, 包有刚, 马萧杰, 等. 自然环境下柑橘采摘机器人避障规划研究[J]. 农业机械学报, 2025, 56(2): 420 - 428.
Bao Xiulan, Bao Yougang, Ma Xiaojie, et al. Obstacle avoidance planning of citrus picking robot in natural environment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(2): 420 - 428. (in Chinese)

[4] 毕松, 高峰, 陈俊文, 等. 基于深度卷积神经网络的柑橘目标识别方法[J]. 农业机械学报, 2019, 50(5): 181 - 186.
Bi Song, Gao Feng, Chen Junwen, et al. Detection method of citrus based on deep convolution neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(5): 181 - 186. (in Chinese)

[5] Zhang Y, Li L, Chun C, et al. Multi-scale feature adaptive fusion model for real-time detection in complex citrus orchard

environments[J]. Computers and Electronics in Agriculture, 2024, 219: 108836.
[6] 徐洪丽, 安凤, 史宇轩. 基于改进 YOLO v8 的苹果叶片病害检测研究[J]. 农业机械学报, 2025, 56(10): 530－538.
Xu Hongli, An Feng, Shi Yuxuan. Apple leaf disease detection method based on improved YOLO v8[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(10): 530－538. (in Chinese)
[7] 修贤超, 费士祺, 黄文倩, 等. 基于轻量化 Mamba－YOLO 模型的梨表面缺陷检测方法[J]. 智慧农业(中英文), 2026, 8(2): 147－157.
Xiu Xianchao, Fei Shiqi, Huang Wenqian, et al. Lightweight Mamba－YOLO based approach for pear surface defect detection [J]. Smart Agriculture, 2026, 8(2): 147－157. (in Chinese)
[8] Zhang B, Xia Y, Wang R, et al. Recognition of mango and location of picking pointon stem based on a multi-task CNN model named YOLOMS[J]. Precision Agriculture, 2024, 25(3): 1454－1476.
[9] Li L, Huang J, Gu Z, et al. Positioning of mango picking point using an improved YOLO v8 architecture with object detection and instance segmentation[J]. Biosystems Engineering, 2024, 247: 202－220.
[10] Li H, Chen J, Gu Z, et al. Optimizing edge-enabled system for detecting green passion fruits in complex natural orchards using lightweight deep learning model[J]. Computers and Electronics in Agriculture, 2025, 234: 110269.
[11] Aharon N, Orfaig R, Bobrovsky B Z. BoT－SORT: robust associations multi-pedestrian tracking[EB/OL]. (2022－06－29)[2026－04－28]. https://arxiv.org/abs/2206.14651.
[12] Wang C Y, Bochkovskiy A, Liao H Y M. YOLO v7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2023: 7464－7475.
[13] 李翠明, 杨柯, 申涛, 等. 基于改进 Faster R－CNN 的苹果采摘视觉定位与检测方法[J]. 农业机械学报, 2024, 55(1): 47－54.
Li Cuiming, Yang Ke, Shen Tao, et al. Vision detection method for picking robots based on improved Faster R－CNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(1): 47－54. (in Chinese)
[14] 李丽, 梁继元, 张云峰, 等. 基于改进 YOLO v5 的复杂环境下柑橘目标精准检测与定位方法[J]. 农业机械学报, 2024, 55(8): 280－290.
Li Li, Liang Jiyuan, Zhang Yunfeng, et al. Accurate detection and localization method of citrus targets in complex environments based on improved YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(8): 280－290. (in Chinese)
[15] 顾文娟, 刘浩状, 魏金, 等. 基于 FPBW－YOLO v8 的复杂场景下番茄果实识别方法[J]. 农业机械学报, 2025, 56(8): 467－478.
Gu Wenjuan, Liu Haozhuang, Wei Jin, et al. Tomato fruit recognition in complex scenes based on FPBW－YOLO v8[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(8): 467－478. (in Chinese)
[16] Liu Zifu, Abeyrathna R M R D, Sampurno R M, et al. Faster－YOLO－AP: a lightweight apple detection algorithm based on improved YOLO v8 with a new efficient PDWConv in orchard[J]. Computers and Electronics in Agriculture, 2024, 223: 109118.
[17] Li Meng, Ding Xue, Wang Jinliang, et al. CMD－YOLO: a lightweight model for cherry maturity detection targeting small object[J]. Smart Agricultural Technology, 2025, 12: 101513.
[18] 苗荣慧, 李港澳, 黄宗宝, 等. 基于 YOLO v7－ST－ASFF 的复杂果园环境下苹果成熟度检测方法[J]. 农业机械学报, 2024, 55(6): 219－228.
Miao Ronghui, Li Gang'ao, Huang Zongbao, et al. Maturity detection of apple in complex orchard environment based on YOLO v7－ST－ASFF[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(6): 219－228. (in Chinese)
[19] Fu L, Majeed Y, Zhang X, et al. Faster R－CNN-based apple detection in dense-foliage fruiting-wall trees using RGB and depth features for robotic harvesting[J]. Biosystems Engineering, 2020, 197: 245－256.
[20] 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)
[21] 刘进一, 晏伏山, 董赫, 等. 基于改进 MBS－YOLO v8 的火龙果目标检测与定位方法[J]. 农业机械学报, 2025, 56(5): 425－432.
Liu Jinyi, Yan Fushan, Dong He, et al. Pitaya fruit target detection and localization method based on improved MBS－YOLO v8[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(5): 425－432. (in Chinese)
[22] 罗志聪, 何陈涛, 陈登捷, 等. 基于轻量化 YOLO v8s－GD 的自然环境下百香果快速检测模型[J]. 农业机械学报, 2024, 55(8): 291－300.
Luo Zhicong, He Chentao, Chen Dengjie, et al. Passion fruit rapid detection model based on lightweight YOLO v8s－GD[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(8): 291－300. (in Chinese)
[23] Tong Z, Chen Y, Xu Z, et al. Wise－IoU: bounding box regression loss with dynamic focusing mechanism[EB/OL]. (2023－01－24)[2026－04－28]. https://arxiv.org/abs/2301.10051.
[24] Jin Shouxiang, Zhou Lei, Zhou Hongping, et al. CO－YOLO: a lightweight and efficient model for *Camellia oleifera* fruit object detection and posture determination[J]. Computers and Electronics in Agriculture, 2025, 235: 110394.
[25] Zeng Tiwei, Tong Jintao, Sun Xudong, et al. YOLO－FCAP: an improved lightweight object detection model based on YOLO v8n for citrus yield prediction in complex environments[J]. Smart Agricultural Technology, 2025, 12: 101590.