doi:10.6041/j.issn.1000-1298.2026.22.027

基于轻量化 SLD－YOLO v8n 的温室番茄成熟度检测方法

奚小波　沈　杨　丁杰源　韩连杰　邹赟涵　张瑞宏
（扬州大学机械工程学院，扬州 225127）

摘要：面向复杂温室环境下番茄成熟度检测中存在的光照波动、枝叶遮挡以及果实密集分布等挑战，本文提出一种轻量化番茄成熟度检测方法。以 YOLO v8n 为基线模型，构建 C2f－SHAB 特征提取模块，改善局部细节与全局语义信息的联合建模效果；引入 DySample 上采样模块，优化多尺度特征融合过程中的高分辨率信息恢复；设计轻量化共享细节感知检测头 LSDA－DHead，提升检测端对复杂边界与遮挡区域的特征表达能力；同时结合 Focaler－DIoU 损失函数优化边界框回归过程。基于自建番茄成熟度数据集开展试验，将番茄划分为未熟期、转色期、转熟期和成熟期 4 个等级。试验结果表明，改进后 SLD－YOLO v8n 模型，与 YOLO v8n 相比，mAP@0.5 提高 7.11 个百分点，参数量、浮点运算量和模型内存占用量分别下降 30.6%、25.9% 和 15.1%。在保证轻量化的同时，有效提升了复杂环境下番茄成熟度检测的精度与鲁棒性，可为智能采摘机器人和边缘端视觉感知系统提供技术支撑。

关键词：番茄；成熟度；机器视觉；YOLO v8n；轻量化
中图分类号：S224.9　　**文献标识码**：A　　**文章编号**：1000-1298(2026)22-0315-11　　**OSID**：

Tomato Maturity Detection Method in Greenhouse Based on Lightweight SLD－YOLO v8n

Xi Xiaobo　Shen Yang　Ding Jieyuan　Han Lianjie　Zou Yunhan　Zhang Ruihong
(*School of Mechanical Engineering*, *Yangzhou University*, *Yangzhou* 225127, *China*)

Abstract: Aiming to address the challenges of illumination fluctuation, branch and leaf occlusion, and dense fruit distribution in complex greenhouse environments for tomato maturity detection, a lightweight tomato maturity detection method was proposed. Taking YOLO v8n as the baseline model, a C2f－SHAB feature extraction module was constructed to improve the joint modeling of local details and global semantic information. The DySample upsampling module was introduced to optimize high-resolution information recovery during multi-scale feature fusion. A lightweight shared detail-aware detection head, LSDA－DHead, was designed to enhance the feature representation capability for complex boundaries and occluded regions at the detection end. Furthermore, the Focaler－DIoU loss function was incorporated to optimize the bounding box regression process. Experiments were conducted on a self-built tomato maturity dataset, in which tomatoes were classified into four stages: immature, turning-color, turning-ripe, and mature. Experimental results showed that the improved SLD－YOLO v8n model achieved a 7.11 percentage point increase in mAP@0.5 compared with the baseline YOLO v8n, while the number of parameters, floating-point operations, and model memory footprint were reduced by 30.6%, 25.9%, and 15.1%, respectively. While ensuring lightweightness, the proposed method effectively enhanced the accuracy and robustness of tomato maturity detection under complex conditions, providing technical support for intelligent harvesting robots and edge-side visual perception systems.

Key words: tomato; maturity; machine vision; YOLO v8n; lightweight

收稿日期：2026－04－07　修回日期：2026－05－26
基金项目：江苏省科技计划现代农业项目(BE2018302)和扬州大学“高端人才支持计划”项目
作者简介：奚小波(1989—)，男，副教授，博士生导师，主要从事智能农业装备研究，E-mail：xbxi@yzu.edu.cn
通信作者：张瑞宏(1960—)，男，教授，博士生导师，主要从事智能农业装备研究，E-mail：zhang-rh@163.com

0 引言

番茄成熟度识别是果实适期采收、品质分级与智能采摘决策中关键环节,识别结果直接影响采摘时机、分拣策略及最终商品品质。已有研究指出,番茄成熟度与果实品质和口感密切相关,是生产中采摘与分拣的重要依据[1]。随着设施园艺、农业机器人与智能装备技术的发展,成熟度识别已成为智能采摘系统感知、决策、执行链条中的核心感知任务,对于降低人工依赖、提升采收效率以及推动农业生产向精准化、自动化和智能化转型具有重要意义[2-3]。然而,在设施温室和自然场景中,番茄通常呈簇生分布,果实空间位置与姿态差异较大,并普遍伴随枝叶遮挡、果实重叠、背景干扰、光照波动以及果梗细小短促等问题,不仅增加了成熟度识别与空间定位的难度,也对后续采摘操作精度提出了更高要求[4-6]。传统基于颜色、形状和纹理等人工特征的识别方法对环境依赖较强,多适用于背景较简单的场景,在复杂条件下易出现步骤繁琐、精度下降和鲁棒性不足等问题[7]。近年来,深度学习方法凭借端到端特征学习优势,显著提升了复杂场景下果蔬目标检测与成熟度判别的稳定性和精度,为番茄成熟度自动检测、空间定位及智能采摘提供了更具鲁棒性的技术路径。

近年来,番茄视觉检测研究已由依赖颜色空间和传统分割方法的目标识别,逐步发展为以深度学习为核心的多任务、轻量化和实时化检测体系。传统方法主要利用 HSV 颜色特征及分水岭等算法完成成熟番茄的识别与区域分割,在非均匀光照和复杂背景条件下具有一定适应能力,但对遮挡、重叠及多尺度目标的表征能力仍较有限[2]。随着 YOLO 和 RT-DETR 等深度学习模型的广泛应用,相关研究逐渐聚焦于复杂冠层环境下的果实检测、成熟度分级、采摘点定位和表型参数提取,并通过多任务学习、动态加权、全局上下文建模及多模态特征融合等策略,实现果实、果串和成熟度等信息的协同识别,提高了复杂场景下检测的准确性与实时性[5,8-10]。针对枝叶遮挡、果实重叠、小目标特征易丢失及相邻成熟阶段差异较弱等问题,现有研究进一步从网络结构、特征融合和损失函数等方面进行优化,通过引入 FasterNet、P2 小目标检测分支、BiFPN、高分辨率预测头、坐标注意力、CBAM 及大核卷积注意力等模块,增强多尺度特征提取、细粒度特征表达和目标区域长距离依赖建模能力,同时结合 WIoU、DIoU 等边界框回归损失改善复杂场景下的定位精度。此外,空间金字塔池化、自适应上采样、Soft-NMS 及轻量化结构设计进一步提升了模型在重叠目标、小目标及连续作业环境中的鲁棒性,并推动检测模型向移动端和 Jetson Nano 等边缘设备部署发展[11-16]。

针对农业场景中模型部署资源受限、移动终端算力不足等问题,研究者围绕轻量化设计与边缘端应用开展了大量探索。现有方法通过引入自适应卷积、轻量化卷积、注意力机制及 BiFPN 等特征融合结构,在增强多尺度目标表征能力的同时,有效降低了模型参数量与计算复杂度,从而提升自然环境下番茄成熟度检测的精度与实时性[17-19]。此外,相关研究已进一步拓展至芒果、梨等果实检测与计数任务,通过优化检测网络结构及融合目标跟踪算法,在保证检测精度的基础上兼顾推理速度和计算成本,并实现连续场景下的高效果实计数[20-21]。

现有研究表明,番茄成熟度检测正在由单纯追求识别精度转向精度、速度与部署可行性并重的发展方向。然而,在复杂温室环境下,目标密集、遮挡频繁以及成熟度差异细微等问题仍会削弱模型的稳健性;边缘设备算力和存储资源有限,也对模型轻量化提出了更高要求。基于此,本文以 YOLO v8n 为基础框架,围绕特征提取、多尺度信息恢复、检测头设计及边界框回归优化等关键环节开展协同改进,构建轻量化番茄成熟度检测模型 SLD-YOLO v8n,以期为复杂温室环境下番茄成熟度高精度实时识别及边缘端轻量化部署提供可靠的视觉感知方法与技术支撑。

1 材料与方法

1.1 数据采集

数据采集于扬州育才生物科技有限公司蒋王镇种植基地,采集设备为 iPhone 16 型手机,图像统一保存为 JPG 格式。采集过程覆盖番茄不同生育阶段及多种自然干扰条件,包括光照强弱变化、枝叶遮挡、果实重叠和多成熟度混杂分布等典型场景,以保证数据对实际温室作业环境的代表性。依据 GH/T 1193—2021《番茄》,并结合果皮颜色变化及采收应用需求,将番茄成熟度归并为未熟期、转色期、转熟期和成熟期 4 个等级,如图 1 所示。其中,未熟期包含未熟果和绿熟果,转色期单独作为后熟潜力阶段,转熟期由红熟前期与红熟中期合并而成,成熟期对应红熟后期。最终共获得原始图像 766 幅,标注目标 2 208 个,其中未熟期 623 个、转色期 652 个、转熟期 724 个、成熟期 209 个。

1.2 数据集构建

采用在线标注工具 Make Sense 对采集图像进行人工框选与类别赋值,无需额外本地部署,可直接


(a) 未熟果　(b) 转色果　(c) 转熟果　(d) 成熟果

图 1　番茄成熟度分级

Fig. 1　Tomato maturity grading

导出符合 YOLO 系列训练要求的标注文件，便于后续数据整理与模型训练。为提高标注一致性，统一规定：当目标遮挡程度超过其可辨识区域的三分之二时，不纳入标注；其余目标依据果皮主导颜色及阶段性视觉特征进行成熟度判定。通过统一标注口径，尽量减少人工判断差异对模型训练带来的干扰。

考虑到原始样本规模有限且场景变化复杂，对数据集进行多种形式的增强处理，以提升模型的泛化能力并缓解过拟合风险。如图 2 所示，增强方式主要包括亮度与对比度调整、模糊、滤波、旋转、平移、局部遮挡、噪声扰动以及镜像翻转等。经扩增后，数据集总规模增至 3 672 幅。按照比例 8∶1∶1将全部样本划分为训练集、验证集和测试集，划分后对训练集进行数据增强，以满足模型训练、调参与性能评估的需要。


(a) 原图　(b) 亮度与对比度调整　(c) 模糊、滤波

(d) 旋转、局部遮挡　(e) 噪声扰动　(f) 镜像翻转

图 2　数据增强形式

Fig. 2　Data enhancement form

2　番茄成熟度检测模型

2.1　SLD－YOLO v8n 模型

选择 YOLO v8n 作为基线模型，主要是因其在检测精度、推理速度与模型轻量化之间具有较好的综合平衡，能够为后续结构改进提供稳定且具有代表性的对比基础。相较于早期目标检测模型，YOLO v8n 采用无锚框检测机制，减少了锚框设计对检测性能的影响，有助于提升模型对不同尺度目标的适应能力；同时，解耦检测头将分类与回归任务分离，有利于增强模型的优化效率和检测精度。主干网络中引入的 C2f 模块在保证特征表达能力的同时有效降低了参数冗余，增强了梯度流动和特征复用能力，使模型在轻量化条件下仍具备较强的特征提取能力。YOLO v8n 结合多尺度特征融合策略，能够较好地应对番茄目标在温室环境中普遍存在的小目标、遮挡和密集分布等问题。作为 YOLO v8 系列中的轻量化版本，YOLO v8n 参数量较小、计算开销较低，更适合后续移动端或边缘设备部署需求。因此，以 YOLO v8n 作为基线模型，不仅能够体现改进方法在精度和效率上的提升效果，也有助于验证所提方法在实际农业场景中的应用潜力。

为提升复杂环境下番茄成熟度检测的精度与部署效率，在 YOLO v8n 基础上构建了轻量化改进模型 SLD－YOLO v8n。针对果实密集分布、枝叶遮挡明显以及相邻成熟度类别易混淆等问题，在主干和颈部网络中引入 C2f－SHAB 模块，以增强局部细节提取与全局语义建模能力；采用 DySample 上采样模块，以改善多尺度特征融合过程中高分辨率信息的恢复效果；在检测端设计 LSDA－DHead，以降低检测头冗余并强化对轮廓边界、颜色过渡和遮挡区域的表征能力；同时结合 Focaler－DIoU 损失函数优化边界框回归过程。通过上述改进，模型在保持轻量化特性的同时进一步提升了复杂场景下番茄成熟度检测精度与鲁棒性。上述改进模型命名为 SLD－YOLO v8n，结构如图 3 所示。

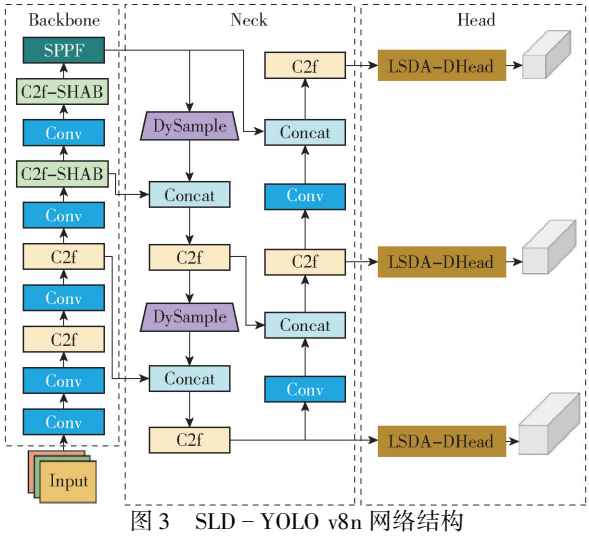


图 3　SLD－YOLO v8n 网络结构

Fig. 3　SLD－YOLO v8n network structure

2.1.1　C2f－SHAB 特征提取模块

番茄成熟度检测任务对特征表达能力具有较高要求，难点不仅在于目标检测本身，还体现在不同成熟阶段果实之间颜色差异细微、局部纹理变化不明显，以及复杂环境中普遍存在的枝叶遮挡、光照不均

和果实密集分布等复杂干扰因素。在此背景下,传统卷积结构虽然能够提取局部纹理与边缘信息,但受限于感受野范围,对远距离依赖关系和全局上下文信息的建模能力不足,易导致模型在复杂背景下对番茄关键区域关注不充分,进而影响成熟度判别精度。为此,引入 SHSA Block[22]对特征提取单元进行改进。如图 4 所示,SHSA Block 由深度卷积、单头自注意力和前馈网络组成,其中深度卷积主要负责提取局部空间细节信息,单头自注意力机制通过对部分通道特征构建查询、键和值映射,建立不同空间位置之间的相关性,从而实现全局上下文信息的聚合,增强模型对关键区域和长程依赖关系的感知能力;随后,前馈网络进一步完成特征的非线性映射与通道交互,以提升整体表达能力。

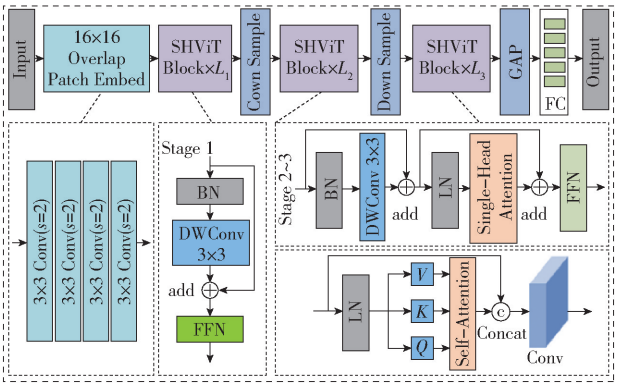


图 4 SHSA Block 结构

Fig. 4 SHSA Block structure

基于此,在 C2f 结构中嵌入 SHSA Block,构建了 C2f-SHAB 模块,如图 5 所示。该模块在继承 C2f 结构良好梯度流传递能力和特征复用能力的基础上,进一步融合了 SHSA Block 在局部细节提取与全局语义建模方面的优势。一方面,模块能够有效提取番茄表面颜色过渡、边缘轮廓及局部纹理等成熟度相关特征;另一方面,又可从全局视角建模不同空间位置之间的关联,缓解因叶片遮挡、背景相似或果实重叠造成的特征混淆问题。相比传统多头自注意力结构,SHSA Block 采用单头设计,在降低参数量和计算复杂度的同时,仍能较好地捕获长程依赖信息,更符合番茄成熟度检测在边缘设备部署中对轻量化与实时性的要求。因此,C2f-SHAB 模块能够在复杂环境下兼顾局部细节增强与全局语义建模,对提升番茄目标检测精度及多阶段成熟度识别性能具有较强的适用性。

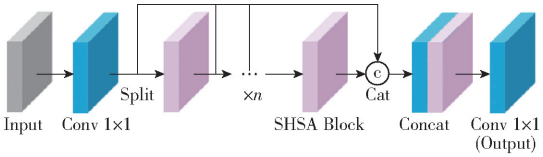


图 5 C2f-SHAB 结构

Fig. 5 C2f-SHAB structure

2.1.2 DySample 上采样

番茄成熟度检测任务对多尺度特征融合质量具有较高要求,尤其在复杂环境中,番茄目标常表现出尺度变化大、分布密集、相互遮挡明显等特点,同时不同成熟阶段之间的判别往往依赖于果皮颜色过渡、边缘轮廓及局部纹理等细粒度信息。在此背景下,传统上采样方法如最近邻插值或双线性插值虽然计算开销较低,但其本质上属于固定规则的特征重建方式,缺乏对空间内容的自适应建模能力,易在特征恢复过程中造成边缘模糊、纹理细节丢失及局部语义错位,从而影响后续检测头对番茄目标的分类与定位精度。为此,引入 DySample 上采样模块对多尺度特征重建过程进行优化。如图 6、7 所示,其中,B 为批大小,C 为通道数,(H,W) 为输入特征图空间尺寸,(H',W') 为上采样后空间尺寸,δ 为可学习缩放因子。DySample[23]的核心思想是通过学习式动态采样替代固定插值,在上采样过程中根据输入特征内容自适应生成采样偏移,使重建后的高分辨率特征更加贴合原始目标结构。

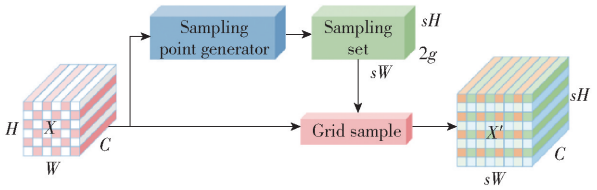


图 6 DySample 结构

Fig. 6 DySample structure

具体而言,DySample 以输入的低分辨率特征图 $X\in\mathbf{R}^{B\times C\times H\times W}$ 为起点,通过线性变换层 Linear1 生成动态范围调整因子 δ,将值域被限制在[0,0.5],以控制采样范围。随后,利用另一线性变换层 Linear2 生成初始偏移量 $O_{\text{int}}\in\mathbf{R}^{B\times 2\times H'\times W'}$。通过结合动态调整因子 δ 与初始偏移量 O_{int},生成动态偏移量 O_{dyn},即

$$O_{\text{dyn}}=O_{\text{int}}\delta \tag{1}$$

将动态偏移量 O_{dyn} 与固定网格位置 G 相加,生成动态采样器 $S\in\mathbf{R}^{B\times 2\times H'\times W'}$,定义所有采样点新二维坐标

$$S=G+O_{\text{int}}+\delta O_{\text{dyn}} \tag{2}$$

动态采样器能够自适应调整采样点分布,使其更精准地覆盖目标区域的语义特征。基于采样器 S,通过网格采样操作 gridsample 对输入特征图 X 进行特征重组,生成高分辨率特征图 $X'\in\mathbf{R}^{B\times 2\times H'\times W'}$,即

$$X'=\text{gridsample}(X,S) \tag{3}$$

动态采样机制通过动态调整采样点的位置,使

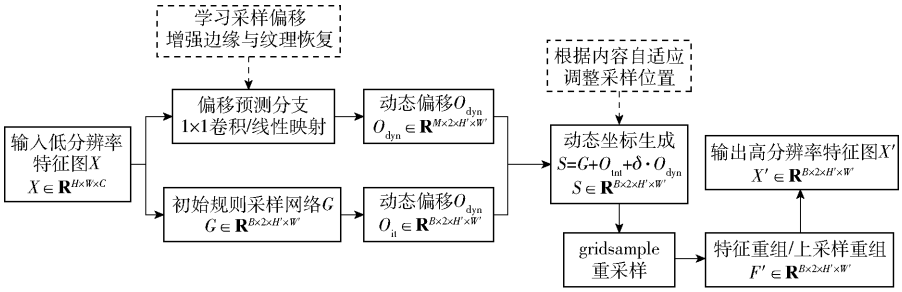


图 7　DySample 点采样流程

Fig. 7　DySample point sampling flowchart

特征图重组过程更加高效，与最近邻插值相比，避免了对所有像素进行固定规则的操作。

2.1.3　LSDA－DHead 检测头

为提升复杂环境下番茄成熟度检测的分类与定位性能，在 YOLO v8n 检测端设计了轻量化共享细节感知检测头（Lightweight shared detail-aware detection head，LSDA－DHead）。如图 8 所示，该检测头的核心思想是在保持轻量化特性的基础上，通过共享式细节感知卷积增强多尺度特征对局部纹理、边缘轮廓及关键区域信息的表征能力。具体而言，来自不同尺度的输入特征首先经过 1×1 卷积进行通道映射与维度统一，以减少冗余信息并实现跨尺度特征对齐；随后，各尺度特征进入共享的细节感知卷积模块进行深层特征提炼。为进一步增强检测头对局部高频信息和细粒度结构的捕获能力，本文在该共享卷积模块中引入细节增强卷积（Detail-enhanced convolution，DEConv）。如图 9 所示，与传统卷积主要侧重通用局部模式提取不同，DEConv[24] 更加关注边缘轮廓、纹理变化及局部结构差异等细节信息的保留与强化，能够有效增强特征图中与成熟度判别密切相关的颜色过渡区域、果实边界及遮挡轮廓的响应强度。

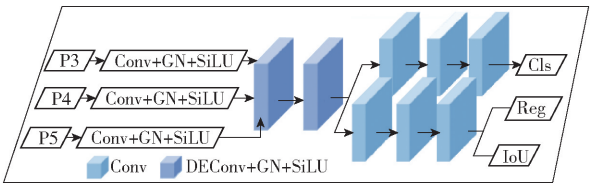


图 8　LSDA－DHead 结构

Fig. 8　LSDA－DHead structure

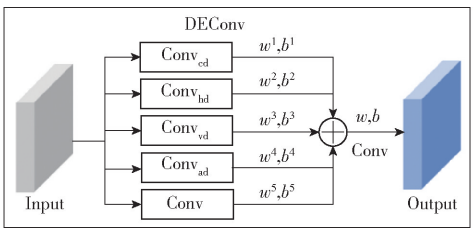


图 9　DEConv 结构

Fig. 9　DEConv structure

对于番茄成熟度检测任务，不同成熟阶段之间往往仅表现出细微的颜色和纹理差异，而复杂环境中的枝叶遮挡、光照不均及背景干扰又进一步增加了判别难度，因此引入 DEConv 有助于提升检测头对关键细节特征的敏感性，从而为后续分类与定位提供更加可靠的表征基础。在此基础上，细节感知卷积模块通过增强局部结构响应，使检测头能够更加敏锐地捕获番茄表面颜色过渡、轮廓边界及遮挡区域等与成熟度判别密切相关的细粒度信息。同时，为提高小批量训练条件下的稳定性，模块中采用归一化策略对特征分布进行调节，从而增强模型在复杂背景与有限样本条件下的收敛能力。在输出端，LSDA－DHead 采用分类分支与回归分支相结合的方式完成目标类别预测与边界框定位，其中回归分支结合分布式边界框回归思想对目标位置进行更精细的建模，从而提升对密集目标、遮挡目标及小目标的定位精度。相较于传统检测头为各尺度分别构建独立卷积分支的方式，LSDA－DHead 通过共享卷积参数有效降低了模型复杂度和计算开销，减少了多尺度检测中的特征冗余，更符合边缘设备部署对轻量化和实时性的要求。

2.1.4　Focaler－DIoU 损失函数

番茄成熟度检测任务不仅要求模型准确识别果实类别，还需要在枝叶遮挡、目标重叠及小目标分布密集等复杂条件下实现精确定位，因此边界框回归性能对检测结果具有重要影响。传统 IoU 类损失虽然能够反映预测框与真实框之间的重叠程度，但在复杂场景下对难样本的关注不足，易导致定位不准或回归收敛缓慢。为此，引入 Focaler－DIoU 损失函数对边界框回归过程进行优化。

该损失在 DIoU[25] 的基础上，综合考虑预测框与真实框之间的重叠程度及中心点距离，同时通过 Focaler 思想对不同质量样本赋予差异化权重，使模型在训练过程中更加关注定位误差较大或回归难度较高的样本。

设预测框为 b，真实框为 b^{gt}，二者交并比（Intersection over Union，IoU）定义为

$$IoU = \frac{|b \cap b^{gt}|}{|b \cup b^{gt}|} \tag{4}$$

式中 $|b \cap b^{gt}|$——预测框与真实框交集面积

$|b \cup b^{gt}|$——预测框与真实框并集面积

Focaler - IoU 的核心是通过线性映射重构 IoU 值,使其能够聚焦于特定的样本。公式为

$$IoU^{focaler} = \text{clip}\left(\frac{IoU - d}{u - d}, 0, 1\right) \tag{5}$$

其中,d 和 u 为可调节的超参数,用于控制模型对不同难度样本的关注程度。在此基础上,Focaler - DIoU 整合了 DIoU 的几何惩罚项 R_{DIoU},通过计算预测框 b 与真实框 b^{gt} 中心点的欧氏距离 ρ 以及覆盖 2 个框的最小闭合区域对角线长度 c 优化空间位置关系,即

$$R_{\text{DIoU}} = \frac{\rho^2(b, b^{gt})}{c^2} \tag{6}$$

最终,Focaler - DIoU 损失函数 $L_{\text{Focaler-DIoU}}$ 定义为重构 IoU 与几何惩罚项的差值转换

$$IoU_{\text{final}} = IoU^{\text{focaler}} - R_{\text{DIoU}} \tag{7}$$

Focaler - DIoU Loss 最终公式为

$$L_{\text{Focaler-DIoU}} = 1 - IoU^{\text{focaler}} + R_{\text{DIoU}} \tag{8}$$

对于番茄成熟度检测任务,该方法能够有效提升模型对遮挡目标、密集目标及边界模糊目标的定位精度,减少因框偏移或框重叠不足引起的漏检和误检问题。同时,Focaler - DIoU 有助于加快边界框回归收敛,提高模型在复杂环境下的检测稳定性与鲁棒性,Focaler - DIoU 操作流程如图 10 所示。

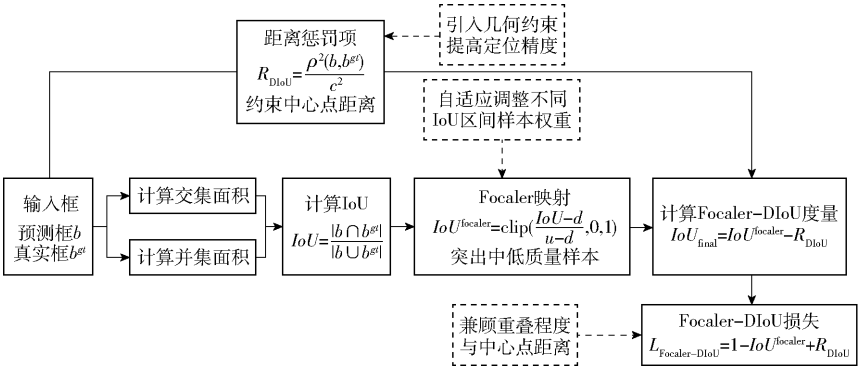


图 10 Focaler - DIoU 操作流程

Fig. 10 Focaler - DIoU process flowchart

3 试验

3.1 试验平台与参数设置

为保证试验结果可重复性,在统一环境下完成模型训练与测试。硬件平台包括 NVIDIA GeForce RTX 4060Ti GPU、13th Intel i5 - 13600KF CPU 和 32 GB 内存,软件环境为 Python 3.10 与 PyTorch 深度学习框架。训练过程中,输入图像尺寸设为 640 像素 × 640 像素,训练轮次为 300,批大小(Batch size)为 16,关闭 Mosaic 数据增强。模型性能从精度和效率 2 个方面进行评价,精度指标采用 mAP@0.5,效率指标采用检测时间和模型内存占用量,同时结合参数量与浮点运算量综合衡量模型轻量化。

3.2 试验结果分析

3.2.1 消融试验

为验证 DySample 上采样模块、LSDA - DHead 检测头、C2f - SHAB 特征提取融合模块以及 Focaler - DIoU 损失函数对模型性能的影响,开展消融试验,结果如表 1 所示。总体来看,各改进模块均对番茄成熟度检测表现出积极作用,但其提升路径和作用侧重点有所不同。基线模型在复杂环境下已具备一定检测能力,但在果实密集分布、局部遮挡及相近成熟度类别区分方面仍存在不足。针对上述问题,从上采样重建、特征提取、检测头设计和损失函数优化 4 个方面进行改进,试验结果表明,各模块均能从不同层面提升模型性能,且联合使用时具有较好的协同增益。

从单模块结果来看,DySample 能够有效改善高分辨率特征重建质量,增强模型对番茄边缘轮廓、颜色过渡区域及小目标细节的表征能力,从而提升检测精度;C2f - SHAB 在增强检测性能的同时兼顾轻量化,表明其能够在较低复杂度下加强模型对全局上下文和关键区域的建模能力;LSDA - DHead 在显著压缩检测头冗余开销的同时保持了良好的检测性能,体现出共享式细节感知卷积结构在边界特征提取和轻量化设计方面的优势;Focaler - DIoU 主要体现在提升定位质量上,表明其能够通过优化样本回归关注度,增强模型对重叠目标及难检测样本的拟合能力。

从模块组合效果来看,各改进表现出较强的互补性。DySample 与 C2f - SHAB 联合后取得较优检测效果,表明动态特征重建与全局语义建模之间具

表 1　SLD－YOLO v8n 模型消融试验结果

Tab. 1　Ablation test results of SLD－YOLO v8n model

A	B	C	D	参数量	浮点运算量	模型内存占用量/MB	mAP@0.5/%	mAP@0.5:0.95/%
				3.01×10^6	8.18×10^9	5.50	82.47	72.60
√				3.01×10^6	8.12×10^9	5.98	86.33	74.93
	√			2.68×10^6	7.59×10^9	5.42	86.68	75.15
		√		2.32×10^6	6.54×10^9	4.61	84.74	73.85
			√	3.05×10^6	8.18×10^9	5.99	86.19	76.07
√	√			2.73×10^6	7.72×10^9	5.41	87.60	77.34
√		√		2.32×10^6	6.51×10^9	5.09	86.54	76.05
	√	√		2.04×10^6	6.15×10^9	4.53	86.93	75.15
√	√	√		2.09×10^6	6.06×10^9	4.67	87.93	75.58
√	√	√	√	2.09×10^6	6.06×10^9	4.67	89.58	76.71

注：A 表示 DySample，B 表示 C2f－SHAB，C 表示 LSDA－DHead，D 表示 Focaler－DIoU；√表示采用该模块。

有良好的协同作用；C2f－SHAB 与 LSDA－DHead 组合更加突出轻量化优势，表明主干特征优化与检测端冗余压缩能够在较低复杂度下维持较强检测性能；3 个模块联合后，模型在保持较低参数量和浮点运算量的同时进一步提升了整体检测效果，表明多层次协同优化对复杂环境下的番茄成熟度检测是有效的。值得注意的是，B 与 B＋C 组合 mAP@0.5:0.95相同，但在 mAP@0.5 上仍有提升。说明在当前数据集和训练设置下，LSDA－DHead 对较宽松 IoU 阈值下的目标检出和分类判别具有促进作用，但在更严格 IoU 区间内，定位优势尚未完全转化为指标增益。考虑到 mAP@0.5:0.95 为多阈值平均结果，该现象也可能与不同 IoU 区间下的增减抵消有关。

在上述改进基础上进一步引入 Focaler－DIoU 后，模型综合性能达到最优，表明结构优化与损失函数优化之间具有良好的耦合关系。前者提升了特征表达与信息传递能力，后者进一步增强了模型对困难样本和高质量定位的关注程度，从而实现分类精度和边界回归能力的同步提升。综上，本文提出的各改进模块能够从不同层面协同增强模型性能，使其在复杂环境下实现检测精度、定位能力与轻量化部署性能的有效平衡。

3.2.2　不同模型检测性能

由表 2 可知，SLD－YOLO v8n 在检测精度、推理效率与轻量化部署性能之间表现出较优的综合平衡。相较于 YOLO v3s、YOLO v7－tiny 等模型，SLD－YOLO v8n 在参数量、浮点运算量及模型规模方面具有更明显的轻量化优势，表明其在降低资源消耗和提升部署适应性方面具有较强竞争力；同时，该模型在检测时间上也表现更优，说明其能够较好满足番茄成熟度检测任务对实时性的要求。与 YOLO v5n、YOLO v8n、YOLO v10n 及 YOLO 11n 等轻量化模型相比，SLD－YOLO v8n 在保持较低复杂度的同时，进一步提升了召回率与平均精度均值，表明改进后的模型对目标番茄的特征表达能力和检出能力更强，能够有效降低复杂场景下的漏检风险。总体而言，SLD－YOLO v8n 在不显著增加计算负担的前提下，实现了检测精度与检测效率的协同提升，表明本文所提改进策略在增强模型鲁棒性、提高番茄目标识别能力以及满足边缘设备部署需求等方面均具有较好的有效性和应用潜力。

表 2　不同模型检测性能

Tab. 2　Detection performance among different models

模型	参数量	浮点运算量	模型内存占用量/MB	准确率/%	召回率/%	mAP@0.5/%	检测时间/ms
YOLO v3s	6.15×10^7	1.55×10^{11}	117.01	79.14	80.21	84.03	59.13
YOLO v5n	1.76×10^6	4.12×10^9	3.71	77.23	76.52	80.12	29.14
YOLO v7－tiny	6.21×10^6	1.28×10^{10}	12.12	80.19	79.03	83.07	43.13
YOLO v8n	3.01×10^6	8.18×10^9	5.50	78.52	77.87	82.47	22.10
YOLO v10n	3.00×10^6	8.20×10^9	5.96	78.01	77.07	81.72	17.76
YOLO 11n	2.90×10^6	7.91×10^9	5.30	77.57	78.34	82.06	18.84
SLD－YOLO v8n	2.09×10^6	6.06×10^9	4.67	80.03	86.32	89.58	17.46

3.2.3 改进前后模型性能

如图 11 所示，由改进前后模型的训练曲线可见，SLD－YOLO v8n 在整个训练过程表现出更优的精度演化特征。相较于基线 YOLO v8n，改进模型在中后期保持了更高的 mAP 水平，且曲线波动相对更小，表明其在训练稳定性和特征学习能力方面均有所增强。尤其在更严格的 IoU 评价条件下，SLD－YOLO v8n 的优势更加明显，表明所引入的结构改进不仅提高了目标检出能力，也改善了边界框回归质量。总体来看，改进模型在收敛速度、稳定性和综合检测性能方面均优于基线模型。

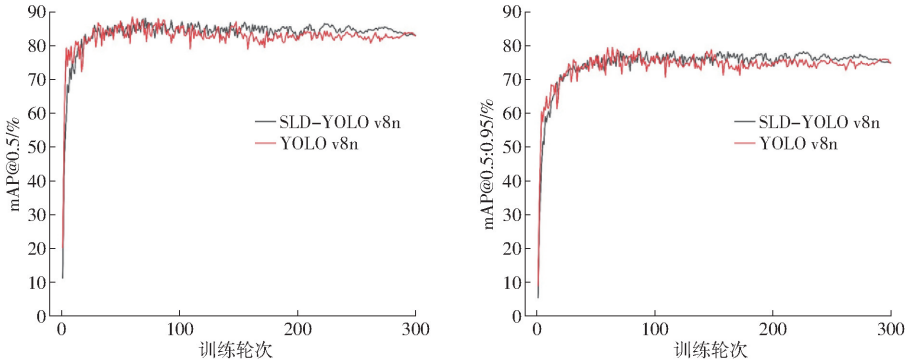
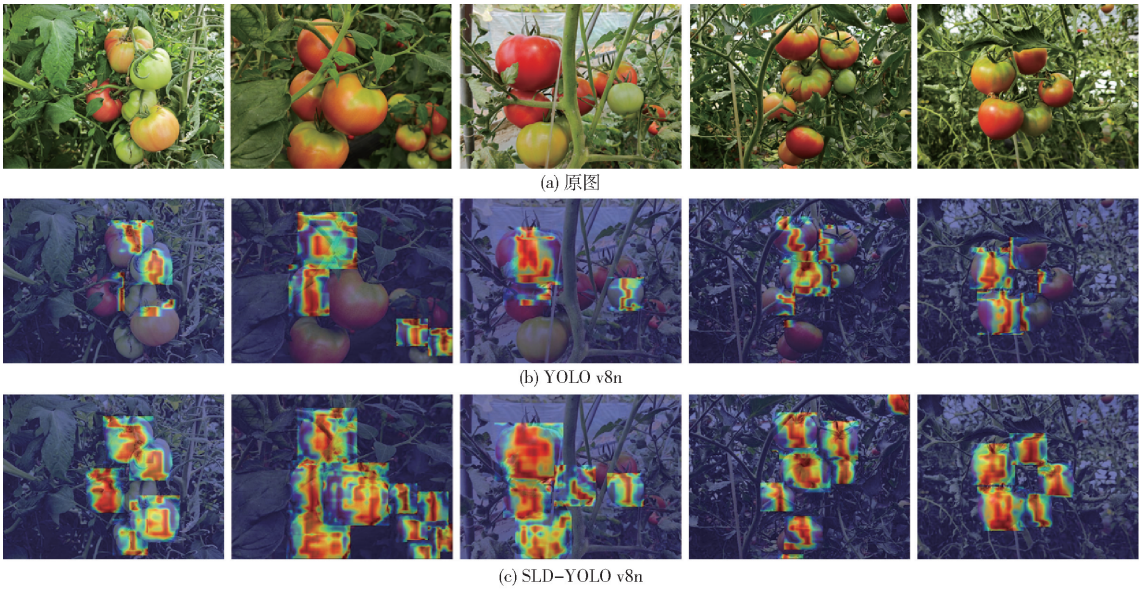


图 11 改进前后模型性能

Fig. 11 Model performance before and after improvement

如图 12 所示，SLD－YOLO v8n 对番茄目标区域的关注更加明确，激活区域与果实主体的空间对应关系更紧密。在枝叶交错、果实重叠和尺度变化明显的场景下，基线模型的响应区域相对分散，对部分目标边缘和邻近果实的覆盖不够充分；而改进模型在目标内部和边缘附近均呈现更连续的高响应分布，同时对背景区域的无关激活有所抑制。表明 SLD－YOLO v8n 能够更有效地提取与成熟度判别相关的颜色过渡、轮廓边界及局部纹理信息，从而提升复杂环境下的检测可靠性。



(a) 原图

(b) YOLO v8n

(c) SLD-YOLO v8n

图 12 SLD－YOLO v8n 与 YOLO v8n 热力图

Fig. 12 Heatmaps between SLD－YOLO v8n and YOLO v8n

为进一步验证模型在复杂环境下适应能力，选取果实密集重叠、枝叶遮挡和高亮干扰 3 类典型场景进行试验，如图 13 所示。由图 13 可知，改进模型在复杂场景下均表现出更优的检测稳定性与类别判别能力。在果实密集重叠场景中，番茄目标间相互紧邻甚至局部交叠，基线模型易出现相邻果实边界分离不清和成熟度误判，而改进模型能够较完整地识别各目标，并保持较好的检测框贴合度。表明 C2f－SHAB 模块增强了对局部细节与全局上下文关系的联合建模能力，DySample 改善了密集场景下多尺度特征融合过程中的细节恢复效果，从而有助于区分相互重叠的果实目标。在枝叶遮挡场景中，受枝叶遮挡影响，部分果实仅暴露局部颜色和轮廓信息，基线模型存在漏检或类别混淆现象；相比之

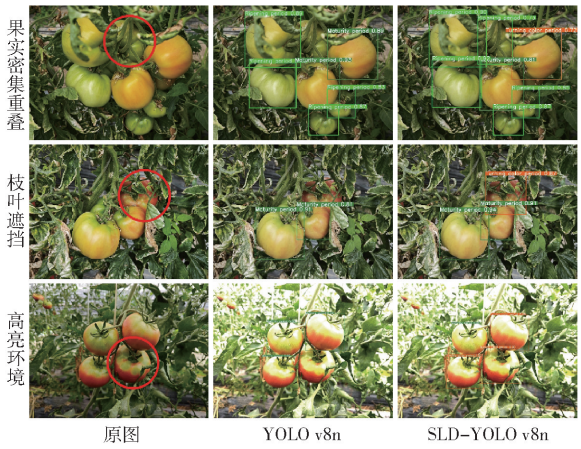


图 13　复杂场景检测效果

Fig. 13　Detection performance in complex scenes

下，改进模型对被遮挡果实仍保持较高的识别准确性，表明 LSDA－DHead 能够增强检测端对边缘轮廓、局部纹理及遮挡区域的特征表达能力，同时 Focaler－DIoU 对遮挡条件下边界框回归的优化作用也有助于提高定位精度。在高亮环境场景中，高亮区域导致果实表面颜色信息失真，易引起成熟度类别混淆，而改进模型仍能较稳定地完成目标定位与成熟度判别，表明所提出方法在复杂光照条件下具有更强的鲁棒性。总体而言，通过对特征提取模块、上采样模块、检测头及损失函数的协同改进，本文模型在果实密集分布、遮挡干扰和高亮变化等复杂场景下均表现出更好的目标检出能力、类别判别能力和边界定位能力。

3.2.4　不同模型检测效果

选取 YOLO v8n、YOLO v10n 和 YOLO 11n 作为对比对象。相较于更早期的检测模型，上述模型在网络结构与轻量化设计上具有更强的可比性，因此更适合用于展示本文方法在当前主流轻量检测框架中的性能差异。

如图 14 所示，各模型在复杂温室环境下对番茄目标的识别能力存在明显差异。总体而言，YOLO v8 前的模型虽能完成部分显著目标的检测，但在果实密集分布、枝叶遮挡、小目标识别以及相近成熟度类别区分方面，普遍存在漏检、误检和定位不准等问题，主要与其特征提取能力有限、多尺度信息融合不足以及对复杂背景干扰抑制能力较弱有关。相比之下，SLD－YOLO v8n 的综合检测效果最优，在多组样本中均能较完整地识别出不同成熟阶段番茄目标，检测框覆盖更准确，类别判别也更稳定，尤其在枝叶遮挡、果实重叠和尺度差异较大场景下仍可保持较好的检测一致性。主要原因为，SLD－YOLO v8n 通过 DySample 增强了上采样阶段的细节恢复能力，使小目标和边缘信息在多尺度融合过程中得到更充分保留；C2f－SHAB 强化了局部细节与全局上下文的联合建模能力，有助于在复杂背景中突出与成熟度相关的关键区域；LSDA－DHead 进一步提高了检测头对轮廓边界、颜色过渡和遮挡区域的敏感性；Focaler－DIoU 改善了密集目标和边界模糊目标的回归质量，因此整体表现出更高的检出率和更



图 14　不同模型检测效果

Fig. 14　Detection results among different models

准确的定位能力。

从基线模型 YOLO v8n 检测结果来看,其对多数显著番茄目标已具备较好检测能力,但在局部遮挡较强、果实彼此靠近或成熟度差异较小的区域,仍出现一定程度漏检和类别混淆,部分检测框对目标边界的贴合程度也不够理想。表明 YOLO v8n 虽具有较好的轻量化优势和基础检测性能,但在复杂场景下对细粒度特征和长程依赖关系的建模仍存在不足。YOLO v10n 在部分样本中的检测结果相较 YOLO v8n 并未表现出明显优势,对密集分布果实和背景交错区域的识别仍不够充分,部分小尺度或被遮挡目标未能被有效检出,说明其在复杂农业场景下对弱显著目标和局部细节信息的利用仍受限制。YOLO 11n 在整体检测稳定性上较 YOLO v10n 有所改善,对部分显著目标的分类和定位更加平稳,但在多目标重叠、枝叶干扰较强以及成熟度相邻类别判别方面仍存在一定偏差,表明虽增强了特征表达能力,但对番茄成熟度检测任务中细粒度差异和复杂空间关系的适配性仍不及本文模型。

综合而言,SLD - YOLO v8n 在复杂温室环境下能够更有效地兼顾目标检出、类别判别与边界定位,表现出更强的鲁棒性和适用性,验证了所提出改进策略在番茄成熟度检测任务中的有效性。

4 结论

(1)提出了一种面向复杂温室环境的轻量化番茄成熟度检测模型 SLD - YOLO v8n。以 YOLO v8n 为基础,通过引入 C2f - SHAB、DySample、LSDA - DHead 和 Focaler - DIoU,从特征表达、细节恢复、检测输出和边界框回归等方面完成协同优化。

(2)不同模型性能测试试验及热力图可视化结果表明,SLD - YOLO v8n 在复杂背景下对番茄目标区域的响应更加集中、完整,对重叠目标、局部遮挡目标及小尺度目标具有更强的识别与定位能力。与 YOLO v8n、YOLO v10n 和 YOLO 11n 相比,本文模型在目标检出、类别判别和边界贴合程度方面均表现出更好的稳定性和鲁棒性。

(3)与 YOLO v8n 相比,SLD - YOLO v8n 在检测精度与轻量化方面均表现更优:参数量由 3.01×10^6 降至 2.09×10^6,浮点运算量由 8.18×10^9 降至 6.06×10^9,模型内存占用量由 5.50 MB 降至 4.67 MB,检测时间缩短至 17.46 ms;同时,精确率达到 80.03%,召回率达 86.32%,mAP@0.5 提升至 89.58%,表明本文模型在降低复杂度的同时,实现了检测精度和推理效率的协同提升。

参 考 文 献

[1] Wang A C, Xu Y Z, Hu D, et al. Tomato yield estimation using an improved lightweight YOLO11n network and an optimized region tracking-counting method[J]. Agriculture, 2025, 15: 1353.

[2] Malik M H, Zhang T, Li H, et al. Mature tomato fruit detection algorithm based on improved HSV and watershed algorithm [J]. IFAC PapersOnLine, 2018, 51(17): 431 - 436.

[3] 荣佳诚, 熊梓茗, 元嘉策, 等. 复合机械臂构型番茄采摘机器人设计与试验[J]. 农业机械学报, 2026, 57(5): 61 - 70,94.
Rong Jiacheng, Xiong Ziming, Yuan Jiace, et al. Design and experiment of composite-manipulator-type tomato harvesting robot [J]. Transactions of the Chinese Society for Agricultural Machinery, 2026, 57(5): 61 - 70,94. (in Chinese)

[4] Badgujar C M, Poulose A, Gan H. Agricultural object detection with you only look once (YOLO) algorithm: a bibliometric and systematic literature review[J]. Computers and Electronics in Agriculture, 2024, 223: 109090.

[5] Gu Z C, Ma X D, Guan H, et al. Tomato fruit detection and phenotype calculation method based on the improved RTDETR model[J]. Computers and Electronics in Agriculture, 2024, 227: 109524.

[6] Begum N, Hazarika M K. Maturity detection of tomatoes using transfer learning[J]. Measurement: Food, 2022, 7: 100038.

[7] 樊湘鹏, 张岩琪, 周硕, 等. 基于改进 YOLOv8s 与 RGB - D 信息融合的番茄采摘机器人识别与定位方法[J]. 农业工程学报, 2025, 41(15): 106 - 116.
Fan Xiangpeng, Zhang Yanqi, Zhou Shou, et al. Recognition and position method of tomato picking robot based on improved YOLOv8s and RGB - D information fusion[J]. Transactions of the CSAE, 2025, 41(15): 106 - 116. (in Chinese)

[8] Chen W B, Liu M C, Zhao C J, et al. MTD - YOLO: multi-task deep convolutional neural network for cherry tomato fruit bunch maturity detection[J]. Computers and Electronics in Agriculture, 2024, 216: 108533.

[9] Wu Q, Huang H M, Song D K, et al. YOLO - PGC: a tomato maturity detection algorithm based on improved YOLOv11[J]. Applied Sciences, 2025, 15: 5000.

[10] Qin J H, Chen Z L, Zhang Y C, et al. YOLO - CT: a method based on improved YOLOv8n - Pose for detecting multi-species mature cherry tomatoes and locating picking points in complex environments[J]. Measurement, 2025, 254: 117954.

[11] 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)

[12] Lawal M O. Tomato detection based on modified YOLOv3 framework[J]. Scientific Reports, 2021, 11:1447.

[13] Gao X, Ding J Y, Zhang R H, et al. YOLOv8n－CA: improved YOLOv8n model for tomato fruit recognition at different stages of ripeness[J]. Agronomy, 2025, 15:188.
[14] Appe S N, Arulselvi G, Balaji G N. CAM－YOLO: tomato detection and classification based on improved YOLOv5 using combining attention mechanism[J]. PeerJ Computer Science, 2023, 9: e1463.
[15] 任晶秋, 万恩晗, 单蜜, 等. 基于改进 YOLO v8n 网络的番茄成熟度实时检测算法[J]. 农业机械学报, 2025, 56(3): 374－382,450.
Ren Jingqiu, Wan Enhan, Shan Mi, et al. Improved YOLO v8n network for real-time detection of tomato maturity[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(3): 374－382,450. (in Chinese)
[16] 丁杰源, 奚小波, 张宝峰, 等. 基于 YOLOv5s－CBS 的番茄成熟度识别方法[J]. 扬州大学学报(自然科学版), 2025, 28(2): 29－37.
Ding Jieyuan, Xi Xiaobo, Zhang Baofeng, et al. Tomato maturity recognition method based on YOLOv5s－CBS[J]. Jouranl of Yanzghou University (Natural Science Edition), 2025, 28(2): 29－37. (in Chinese)
[17] Wang Q, Hua Y, Lou Q D, et al. SWMD－YOLO: a lightweight model for tomato detection in greenhouse environments[J]. Agronomy, 2025, 15: 1593.
[18] Hao F Q, Zhang Z Y, Ma D X, et al. GSBF－YOLO: a lightweight model for tomato ripeness detection in natural environments[J]. Journal of Real-Time Image Processing, 2025, 22:47.
[19] 奚小波, 丁杰源, 翁小祥, 等. 基于轻量化 YOLO v5s－MCA 的番茄成熟度检测方法[J]. 农业机械学报, 2025, 56(3): 383－391,436.
Xi Xiaobo, Ding Jieyuan, Weng Xiaoxiang, et al. Tomato maturity detection method based on lightweight YOLO v5s－MCA [J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(3): 383－391,436. (in Chinese)
[20] Koirala A, Walsh, K B, Wang Z, et al. Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of 'MangoYOLO'[J]. Precision Agriculture, 2019, 20(6): 1107－1135.
[21] Parico A I B, Ahamed T. Real time pear fruit detection and counting using YOLOv4 models and deep SORT[J]. Sensors, 2021, 21: 4803.
[22] Yun S, Ro Y. Shvit: single-head vision transformer with memory efficient macro design[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 5756－5767.
[23] Ha C K, Nguyen H, Van V D. YOLO－SR: an optimized convolutional architecture for robust ship detection in SAR imagery [J]. Intelligent Systems with Applications, 2025, 26: 200538.
[24] Chen Z, He Z, Lu Z M. DEA－Net: single image dehazing based on detail-enhanced convolution and content-guided attention [J]. IEEE Transactions on Image Processing, 2024, 33: 1002－1015.
[25] Juan T, Dm C, Ja R, et al. A comprehensive survey of loss functions and metrics in deep learning[J]. Artificial Intelligence Review, 2025, 58(7): 195.

(上接第 283 页)

[25] 崔岩. 基于深度学习的玉米植株识别及行间中心线提取方法研究[D]. 大庆:黑龙江八一农垦大学, 2025.
Cui Yan. Deep learning-based corn plant identification and interline centerline extraction method study [D]. Daqing: Heilongjiang Bayi Agricultural University, 2025. (in Chinese)
[26] 江庆, 安东, 韩华宇, 等. 基于激光雷达与 RGB 相机融合的玉米作物行检测算法研究[J]. 农业机械学报, 2024, 55(10): 263－274.
Jiang Qing, An Dong, Han Huayu, et al. Maize crop row detection algorithm based on fusion of LiDAR and RGB camera[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(10): 263－274. (in Chinese)
[27] 傅灯斌, 江茜, 齐龙, 等. 基于区域生长顺序聚类－RANSAC 的水稻苗带中心线检测[J]. 农业工程学报, 2023, 39(7): 47－57.
Fu Dengbin, Jiang Qian, Qi Long, et al. Detection of the centerline of rice seedling belts based on region growth sequential clustering－RANSAC[J]. Transactions of the CSAE, 2023, 39(7): 47－57. (in Chinese)
[28] 刘宇玎, 邵明玺. 基于河马优化算法的成行作物导航路径提取方法[J]. 中国农机化学报, 2025, 46(10): 120－127, 145.
Liu Yuding, Shao Mingxi. Navigation path extraction method for row-planted crops based on the Hippopotamus optimization algorithm[J]. Journal of Chinese Agricultural Mechanization, 2025, 46(10): 120－127, 145. (in Chinese)
[29] Xie E, Wang W, Yu Z, et al. SegFormer: simple and efficient design for semantic segmentation with transformers[PP/OL]. V1. arXiv (2021－10－28)[2026－01－08]. https://doi.org/10.48550/arXiv.2105.15203.
[30] Howard A, Sandler M, Chu G, et al. Searching for MobileNetV3[PP/OL]. V1. arXiv (2019－11－20)[2026－01－08]. https://doi.org/10.48550/arXiv.1905.02244.
[31] Hu J, Shen L, Albanie S, et al. Squeeze-and-Excitation networks[PP/OL]. V1. arXiv (2019－05－16)[2026－01－08]. https://doi.org/10.48550/arXiv.1709.01507.
[32] Ronneberger O, Fischer P, Brox T. U－Net: convolutional networks for biomedical image segmentation[PP/OL]. V1. arXiv (2015－05－18)[2026－01－08]. https://doi.org/10.48550/arXiv.1505.04597.
[33] Xiong Y, Li Z, Chen Y, et al. Efficient Deformable ConvNets: rethinking dynamic and sparse operator for vision applications [PP/OL]. V1. arXiv (2024－01－11)[2026－01－08]. https://doi.org/10.48550/arXiv.2401.06197.