

CD-YOLO: 基于改进YOLOv11s的田间胡萝卜苗检测方法

刘浩然^{1,2}, 王 宇¹, 赵学观^{2,4}, 吴华瑞³, 付 豪^{2,4}, 庞树杰⁵, 翟长远^{2,4*}

(1. 辽宁科技大学机械工程与自动化学院, 辽宁鞍山 114051, 中国; 2. 北京市农林科学院智能装备技术研究中心, 北京 100097, 中国; 3. 北京市农林科学院信息技术研究中心, 北京 100097, 中国; 4. 国家农业智能装备工程技术研究中心, 北京 100097, 中国; 5. 中国农业机械化科学研究院集团有限公司, 北京 100083, 中国)

摘 要: [目的/意义] 枝叶遮挡是田间胡萝卜苗精准识别的主要挑战, 对此开发一种轻量、高效的检测模型, 以实现田间遮挡条件下胡萝卜苗的精准鲁棒识别。[方法] 提出了一种基于改进YOLOv11s (You Only Look Once) 的轻量化检测方法CD-YOLO (Carrot Detection-YOLO)。首先, 为降低模型复杂度, 将主干网络中的部分标准卷积替换为深度可分离卷积, 以减少计算量与参数量; 其次, 在C3k2模块中引入高效多尺度注意力机制 (Efficient Multi-Scale Attention, EMA), 构建C3k2-EMA模块, 增强模型对关键特征的动态感知能力, 从而抑制背景与遮挡噪声; 最后, 采用动态检测头, 强化模型对不同尺度特征的融合与感知, 进一步提升检测鲁棒性。[结果和讨论] 结果表明: CD-YOLO模型的计算量和模型大小相较于YOLOv11s分别降低了6.2 GFLOPs和4.8 MB, 单张图像推理速度提升了4.7 ms, 准确率、召回率与平均精度分别提高了3.0、1.5和2.4个百分点; 在应对枝叶遮挡方面, CD-YOLO的漏识率为13.4%, 较YOLOv11s下降了5.7个百分点; 将模型部署于边缘算力设备, 在随机选取的胡萝卜苗图像上进行测试, 漏识率和误识率分别为5.1%和2.7%。[结论] CD-YOLO模型在保持轻量化的同时, 有效提升了田间胡萝卜苗的检测精度与速度, 特别是在应对枝叶遮挡方面表现优异。该研究可为胡萝卜自动化育苗过程中的出苗率统计与精准管理提供可靠的技术支持。

关键词: 胡萝卜苗; 遮挡; 目标检测; YOLOv11s; 轻量化

中图分类号: TP391.4; S24

文献标识码: A

文章编号: SA202511008

引用格式: 刘浩然, 王宇, 赵学观, 吴华瑞, 付豪, 庞树杰, 翟长远. CD-YOLO: 基于改进YOLOv11s的田间胡萝卜苗检测方法[J]. 智慧农业(中英文), 2026, 8(2): 158-174. DOI: 10.12133/j.smartag.SA202511008

LIU Haoran, WANG Yu, ZHAO Xueguan, WU Huarui, FU Hao, PANG Shujie, ZHAI Changyuan. CD-YOLO: A Method for Detecting Carrot Seedlings in Field Based on Improved YOLOv11s[J]. Smart Agriculture, 2026, 8(2): 158-174. DOI: 10.12133/j.smartag.SA202511008 (in Chinese with English abstract)

0 引 言

胡萝卜幼苗状态监测对生长收获起着决定性的作用^[1], 在胡萝卜的生长监测过程中, 出苗状况检测是验证一批胡萝卜田间长势基础的关键^[2]。准确评估出苗情况, 对于早期预测产量、及时发现缺苗断垄问题, 以及指导间苗、补苗等精准农事作业至关重要。传统人工普查方式效率低下且主观性强, 难以实现精准、快速监测^[3]。尤其在复杂田间环境下, 胡萝卜植株与杂草颜色相近^[4], 形态互掩^[5], 容易造成漏识或误识, 因此实现对胡萝卜苗目标的快速、精准识别与定位^[6], 已成为胡萝卜出苗检测

的关键技术挑战。

近年来, 已有许多学者在田间作物与杂草识别领域进行了相关研究。黄友锐等^[7]提出一种基于改进YOLOv8n (You Only Look Once) 的甜菜杂草检测算法, 在主干网络引入感受野坐标注意力卷积, 增强模型对于目标的边缘、纹理、形状等特征信息的提取, 在公开数据集Lincolnbeet上进行测试, 改进后的模型平均精度达到了90.2%。有学者更多关注模型轻量化。曲福恒等^[8]提出一种基于改进DeepLabv3+的轻量化语义分割网络, 将特征提取网络替换为移动端轻量神经网络 (Mobile Net-

收稿日期: 2025-11-09

基金项目: 北京市农林科学院创新能力建设专项 (KJCX20230409); 国家自然科学基金 (32201647); 改革与发展项目 (GGFZ20250205)

作者简介: 刘浩然, 硕士研究生, 研究方向为智能化施药装备研发。E-mail: lhr13646345271@outlook.com

*通信作者: 翟长远, 博士, 研究员, 研究方向为智能化施药装备研发。E-mail: zhaicy@nrcita.org.cn

copyright©2026 by the authors

(登录 www.smartag.net.cn 免费获取电子版全文)

work v2) 降低参数量, 同时在空洞空间金字塔池化模块 (Atrous Spatial Pyramid Pooling, ASPP) 中引入分组逐点卷积, 所提出的网络平均交并比 (Mean Intersection over Union, mIOU) 达到了 86.75%, 模型大小仅有 0.69 MB, 为田间作物与杂草识别分割提出了一种轻量化可行方案。NIU 等^[9] 基于实例分割开发了一套用于生菜田的智能杂草识别与变量喷洒系统。以 YOLOv8-seg 模型为基础, 通过引入深度可分离卷积 (Depthwise Separable Convolution, DWConv)、改进的 C2f-Star 模块, 以及能量函数简单注意力机制 (Simple Attention Mechanism based on Energy Function, SimAM), 使参数量减少 30%, 平均精度达到 90.15%, 处理速度提升至 15.7 ms。针对自然环境下玉米幼苗与杂草识别精度低、实时性和鲁棒性差等问题, 孟庆宽等^[10] 提出一种基于轻量卷积神经网络结合特征层信息融合机制的改进单次多框检测器 (Single Shot MultiBox Detector, SSD) 模型, 在自然环境下玉米检测精度达到了 88.27%, 对比原模型, 检测精度提高了 2.66%。

上述研究多数仅聚焦于田间作物与杂草的分类识别和模型轻量化。实际田间环境中, 目标作物绝大多数会存在不同程度的遮挡问题, 会损失一定的特征信息, 容易出现漏识现象。因此, 也有学者进行了遮挡条件下目标检测方面的研究。张志远等^[11] 基于 YOLOv5 在模型主干网络中添加 Transformer, 颈部网络融合加权双向特征金字塔网络 (Bidirectional Feature Pyramid Network, BiFPN), 在自然环境下樱桃目标检测精度较原始模型提升了 2.9%。为应对果园中严重的枝叶与果实遮挡问题, ZHANG 等^[12] 提出小型遮挡柑橘感知识别 (Small Citrus with Occlusion-Aware REcognition DETection TRansformer, SCORE-DETR) 模型, 通过通道重校准和大卷积核强化空间上下文建模, 并引入小波融合模块优化特征提取能力, 在遮挡率高达 85% 时检测置信度仍达 0.69, 表现出优异的实用性。汤晨等^[13] 改进基于 YOLOv11n 模型来解决密集遮挡环境下百香果识别定位问题, 通过引入部分卷积和改进具有对特征图切片能力的 (A Sliced 3D Spatial and Channel Attention Module, SimAMs 3D) 注意力机制, 增强模型对特征信息的提取和对目标的聚焦程度, 在复杂遮挡条件下的百香果识别精确率可达 93.32%。针对番茄采摘智能机器人识别中广泛存在的因枝叶遮挡而导致识别不准确的问题, 李文峰

等^[14] 采用 GhostNet 主干网络改进 YOLOv4 模型, 该模型在遮挡面积达 65% 条件下识别准确率仍达 93.5%, 具备较高检测精度, 为自然环境下的番茄采摘机器人检测提供了新的方法。王元昊等^[15] 提出一种基于 YOLOv8-SAM 改进算法, 通过集成瓶颈注意力 (Bottleneck Attention Module, BAM) 机制、SAM 分割模型与三维点云轮廓拟合, 实现被遮挡柑橘的精确识别与空间定位, 在遮挡条件下的果实识别平均精度达到 91.1%, 平均果径误差降低了 7.99%。上述研究虽在不同遮挡条件下目标检测取得了一定进展, 但相应地增加了模型的参数量和复杂度, 忽视了模型的处理速度和时效性, 且多数研究仅聚焦于模型优化而没有考虑边缘部署测试, 因此进行边缘端部署时可能由于算力受限而出现效果不佳的情况。

综合上述分析, 田间胡萝卜苗检测的核心挑战可以归纳为“在有限算力下, 对不完整、多尺度目标进行鲁棒感知”的复合优化问题。这就要求模型必须满足 3 个理论需求: 首先, 模型必须在参数量和计算量上实现缩减, 以适应边缘设备的算力约束, 同时还要具备更快的处理速度以满足实时性要求; 其次, 面对枝叶遮挡造成的目标特征碎片化与信息丢失, 模型必须能增强对局部关键特征的感知, 并重建跨尺度的上下文依赖, 以弥补信息缺口; 最后, 田间目标的尺度、空间分布会因遮挡不同而变化, 模型要具备自适应调整特征融合策略的能力, 以实现不完整目标的精准定位。

针对上述问题, 本研究按照“轻量-增强-动态”的优化思路, 提出一种改进 YOLOv11s 的胡萝卜苗识别模型。首先将普通卷积 (Convolution-Batch-Norm-SiLU, CBS) 模块替换为 DWConv 模块, 降低模型参数量, 提高处理速度, 实现模型轻量化; 其次, 在 C3k2 模块中引入高效多尺度注意力 (Efficient Multi-Scale Attention, EMA) 机制, 形成 C3k2_EMA 模块, 利用碎片化特征有效地捕获多尺度空间特征并建立跨维度 (通道和空间) 的依赖关系, 提升模型在遮挡情况下获取特征信息的能力; 最后, 在模型输出之前添加动态检测头 (DynamicHead) 结构, 用于提升目标检测模型在处理不同尺度、不同空间位置时的性能, 降低模型损失值, 提升模型在遮挡条件下的输出判断能力。这 3 个模块的协同组合, “DWConv” 奠定的轻量化基础, 使得 EMA 与 DynamicHead 的引入成为可能。C3k2_EMA 在特征提取阶段完成的抗损性增强, 为

DynamicHead 提供了更高质量、更具判别性的特征输入；DynamicHead 则在此基础上，完成最终的动态融合与决策优化。

1 材料与方法

1.1 数据集制作

1.1.1 数据采集

本研究所用的胡萝卜苗图像采集于北京市昌平区小汤山国家精准农业研究示范基地，图像采集时间为2024年9月，胡萝卜处于幼苗期和叶片生长盛

期，采集设备为罗技C930c摄像头，图像分辨率为1 920×1 080。需要说明的是，当前数据集在生长季节、地理区域和田间管理方式上具有一定单一性。这主要是为了在研究的初始阶段，聚焦于解决“遮挡”与“轻量化”这两个核心算法问题，控制其他环境变量的过度干扰。为确保数据的多样性，使模型训练具有较好的鲁棒性，采集图像时，采取了不同角度、不同拍摄距离、不同形态等方式，共计获取图像1 274张，部分采集图像实例如图1所示。

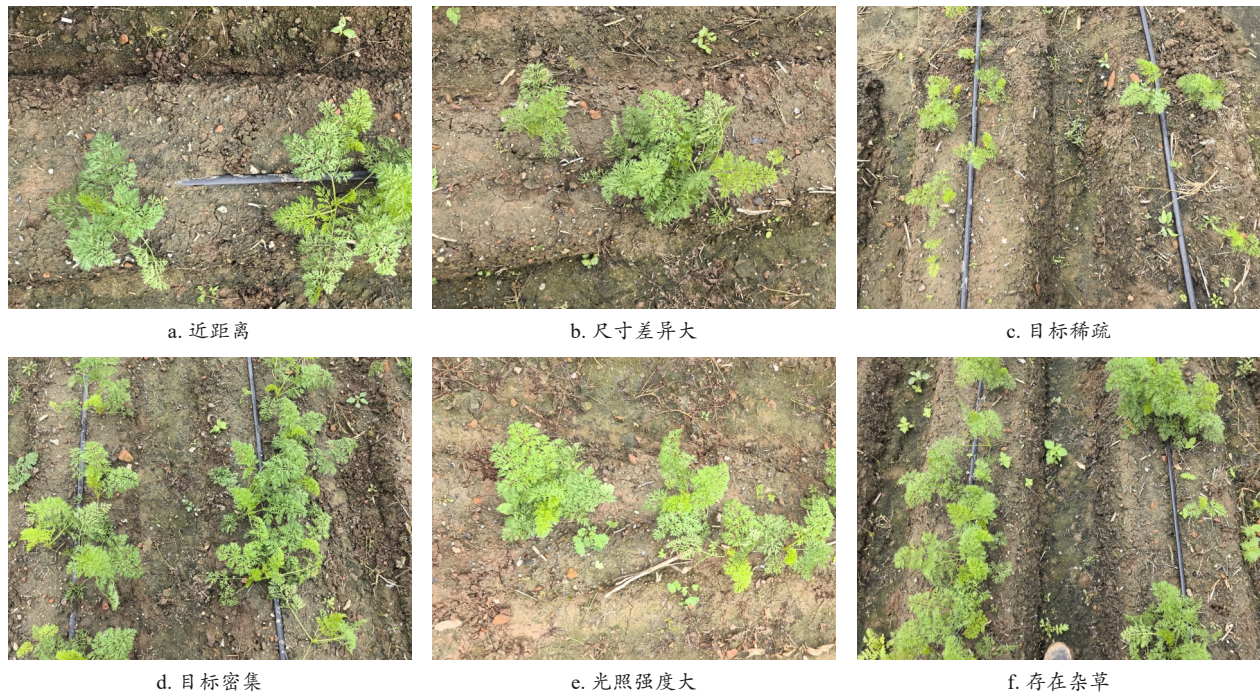


图1 胡萝卜苗采集图像实例

Fig. 1 Carrot seedling image collection examples

1.1.2 数据标注

为了获取精准的胡萝卜苗图像数据集，使用Labelimg软件对所采集到的图像进行数据标注^[16]，得到包含胡萝卜苗目标种类信息和位置信息的标签，保存格式为.txt的标签文件。

1.1.3 遮挡量化标准定义

为对遮挡程度进行量化评估，提出一种基于标注真实框的遮挡程度量化方法，其基本思想是：一个标注框被其他标注框覆盖的区域，可视作其被遮挡的部分。具体如图2所示，对于数据集中每个真实标注框 B_i ，其遮挡程度 O_i 通过计算与其他目标真实标注框 B_j ($j \neq i$) 的交集面积之和占自身面积的比例来衡量，具体计算如公式(1)所示。

$$O_i = \frac{\sum_{i \neq j} \text{Area}(B_i \cap B_j)}{\text{Area}(B_i)} \times 100\% \quad (1)$$

式中： O_i 为当前目标的遮挡程度； B_i 为当前目标标注框； B_j 为相邻目标标注框；Area为标注框所占像素面积； $B_i \cap B_j$ 为相邻两标注框的交集。

使用 O_i 的目的是为目标提供客观的遮挡评分，并依据该评分划分遮挡等级，构建专门的遮挡测试子集，以评估模型的遮挡检测性能。基于 O_i 值，将目标的遮挡程度划分为3个等级：轻度遮挡 (L_1 , $O_i < 10\%$)、中度遮挡 (L_2 , $10\% \leq O_i < 40\%$) 和重度遮挡 (L_3 , $40\% \leq O_i$)。田间胡萝卜苗的遮挡主要表现为苗间互掩，框间重叠面积比能直接反映目标被覆盖的程度，提供了可量化、可重复的评估标准。

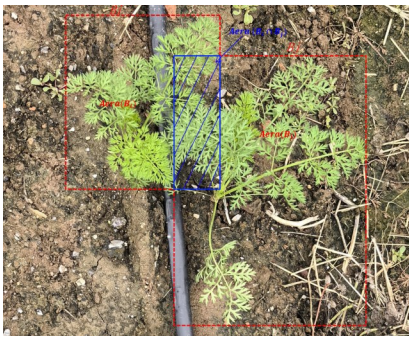


图2 胡萝卜苗遮挡程度量化图

Fig. 2 Quantitative chart of carrot seedling occlusion coverage

1.1.4 数据增强

为提高模型的鲁棒性和泛化能力,同时也为了避免训练过程中出现过拟合现象,本研究采用离线增强^[17]的方式对已标注的胡萝卜苗数据集进行了数据扩充。增强方式主要有图像旋转、图像平移、马赛克增强、添加噪声和亮度增强等方式,通过数据增强,将原始 1 274 张图像扩充至 4 796 张图像,并生成对应的标签文件。通过数据增强丰富了数据集,给模型提供更多可学习的特征,可有效提高模型的鲁棒性和泛化能力。

本研究所构建的数据集共计 4 796 张图像,为了有效验证模型性能,将所有图像在完全打乱后,再按照 8 : 1 : 1 的比例划分为训练集、验证集、测试集。此过程确保了各种情况和增强方式的图像在 3 个子集中近似均匀分布,避免了因人为挑选导致的划分偏差。数据集划分的详细信息如表 1 所示。

表 1 胡萝卜苗数据集分布表

Table 1 Distribution table of carrot seedling dataset

类别	图像数量/张	标签数量/张
训练集	3 837	43 880
验证集	480	6 156
测试集	479	6 867

1.2 胡萝卜苗检测方法

1.2.1 基于 YOLOv11s 的胡萝卜苗检测方法

YOLOv11^[18] 的网络结构沿用了 YOLOv8 的网络结构思想,网络分为 3 部分:主干网络(Backbone)、颈部网络(Neck)、检测头(Head)^[19]。

YOLOv11 的主干网络中,用 C3k2 模块替换了 YOLOv8 主干网络中的 C2f 模块,实现了高层特征与上下文信息的结合,增强了模型对复杂背景的适应性;在快速空间金字塔池化(Spatial Pyramid Pooling-Fast, SPPF)层后面新增了 C2PSA(C2 Position-Sensitive Attention)层,能够聚焦目标的重要

特征,提升模型的特征表达能力。

1.2.2 YOLOv11s 算法改进

针对田间胡萝卜检测中“枝叶遮挡导致特征缺失”与“边缘设备算力受限”问题,本研究的改进策略遵循协同设计的原则“在有限计算的预算下,最大化对不完整、多尺度特征的感知与融合能力”。具体而言:为应对算力受限问题,采用深度可分离卷积替换模型网络中的标准卷积,其目的在于降低模型参数量和计算量,便于部署和达到算力要求。枝叶遮挡导致目标特征信息碎片化、不完整,而 EMA 机制的核心优势在于其高效的多尺度上下文特征信息的捕捉能力,能从碎片化特征中重建更具有鲁棒性的特征,在一定程度上弥补遮挡所造成的特征损失。遮挡和田间环境的复杂性导致目标尺度、空间分布极不规则。DynamicHead 通过尺度感知、空间感知的注意力机制,动态地融合和加权不同层次的特征图,使其能够更好地处理因遮挡而显得尺度模糊或位置不确定的目标。

对于三者的协同关系, DWConv 可降低模型整体计算量,允许嵌入可能提升计算量的 EMA 机制与 DynamicHead 模块; C3k2_EMA 在骨干和颈部网络层强化了关键局部特征的提取,为 DynamicHead 提供了更高质量、更具判别性的多尺度特征输入;而 DynamicHead 则作为检测端的“智能调度器”,进一步融合与优化这些特征,做出最终决策。

鉴于此,本研究基于 YOLOv11s 算法提出 CD-YOLO (Carrot Detection-YOLO) 模型算法,用 DWConv^[20] 替换传统 CBS 卷积操作,减少模型整体参数量和计算量;在部分 C3k2 模块引入 EMA^[21] 机制,构建 C3k2_EMA 模块,强化对关键特征的关注并抑制无关背景噪声;通过“分组-交换”结构捕获多尺度上下文来弥补信息缺失,同时通过跨维度交互精准定位目标的有效部分,聚焦模型对目标的关注度,可实现在遮挡情况下目标部分特征信息缺失情况下的准确推理;添加 DynamicHead^[22] 动态检测头模块,通过尺度感知和空间感知注意力,动态地调整不同通道和空间位置之间的权重信息,增强模型多尺度目标、空间分布不规律条件下的鲁棒性和抗干扰能力。改进之后的识别模型 CD-YOLO 结构图如图 3 所示。

由图 3 可知,本研究对模型做了如下改进。

1) DWConv。YOLOv11 的卷积方式为普通卷积,输入的胡萝卜苗图像由 RGB 三个颜色通道所组成,采用多个 3 维的卷积核进行卷积计算,其

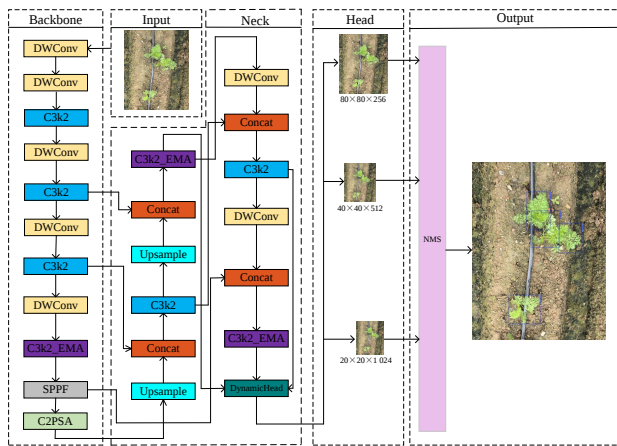


图3 CD-YOLO 结构图

Fig. 3 CD-YOLO model architecture diagram

计算量如公式 (2) 所示。

$$\text{FLOPs} = W \times H \times C \times D^2 \times K \quad (2)$$

式中: FLOPs为浮点运算次数; W 为输入图像的像素宽度; H 为输入图像的像素高度; C 为图像的通道数; D 为卷积核尺寸; K 为卷积核个数。

浮点运算量与设备算力需求成正相关, 算力不足易引发处理延迟、精度下降等问题。由于边缘设备算力通常低于工作站, 故边缘部署对模型的低计算量需求尤为突出。深度可分离卷积 (Depthwise Separable Convolution) 是一种轻量级的卷积操作, 能够减少运算量, 计算量如公式 (3) 所示。

$$\text{FLOPs} = W \times H \times C \times D^2 + D^2 \times C \times K \quad (3)$$

以模型训练时的RGB三通道彩色图像为例, 分辨率为 640×640 像素, 进行1次卷积所设定的卷积核个数为64, 卷积核尺寸为 3×3 , 则普通卷积的计算量为 7.08×10^8 FLOPs, 而深度可分离卷积的计算量为 1.11×10^7 FLOPs。深度可分离卷积的计算量相较于普通卷积减少了98.43%, 鉴于此, 本研究采用DWConv代替主干网络中的普通卷积, 降低模型整体计算量。

2) 融合注意力的C3k2_EMA。EMA机制是一种高效多尺度注意力机制, 其核心在于通过跨空间学习增强特征表征能力, 提升模型对多尺度信息的捕获效能。通过重组通道维度来保留完整信息, 从而在低计算成本下有效增强了模型捕捉多尺度特征的能力, EMA机制的结构如图4所示。

首先, 输入特征图被送入3个并行的卷积分支中。两个分支采用 1×1 卷积核, 主要承担跨通道信息交互与依赖关系建模的功能: 一个分支用于提取通道间的线性组合特征; 另一分支则通过非线性

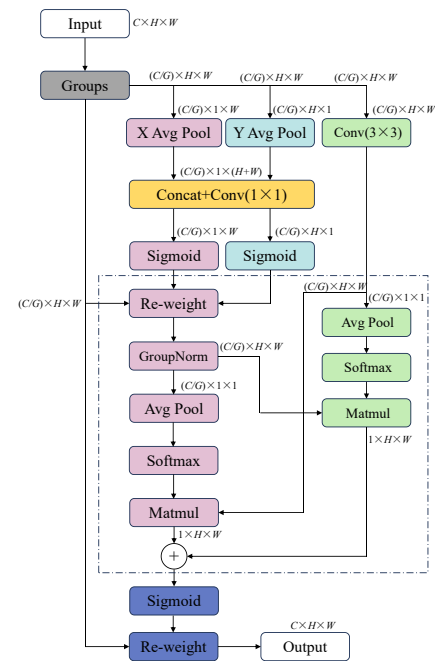


图4 EMA 模块

Fig. 4 EMA module

变换进一步增强通道维度的表征能力。第3个分支采用 3×3 卷积核, 专注于捕获局部空间上下文信息, 通过其较大的感受野增强对邻近语义特征的提取。其次, 3个分支的输出特征在对应位置进行逐元素相加融合, 以整合不同分支所贡献的全局依赖与局部细节信息。融合后的特征经由Sigmoid激活函数归一化, 生成介于0~1的注意力权重图。最后, 该权重图与原始输入特征进行逐点相乘, 从而对关键特征进行强化, 对冗余或噪声特征进行抑制。鉴于此, 本研究在C3k2模块中引入EMA机制, 形成C3k2_EMA模块, 其结构图如图5所示。遮挡问题导致的目标特征不完整, 本质上是局部特征显著性下降和多尺度上下文特征的关联断开。普通卷积或注意力机制难以解决该问题, 而EMA机制通过并行多分支结构与后续的跨纬度交互作用, 能够在单一模块内同时建立强大的通道维与空间维依赖关系。将其嵌入C3k2模块, 相当于在特征提取的关键节点植入“多尺度特征选择器”和“上下文重构器”, 这正是应对特征损失需求的核心。

3) DynamicHead。DynamicHead是针对目标检测任务设计的一种动态统一检测头框架, 其核心优势在于通过结构化注意力机制显式地增强多尺度、多空间和多任务语境下的特征表示能力, 在不引入显著计算开销的前提下显著提升检测性能。

DynamicHead模块通过在尺度、空间、任务维

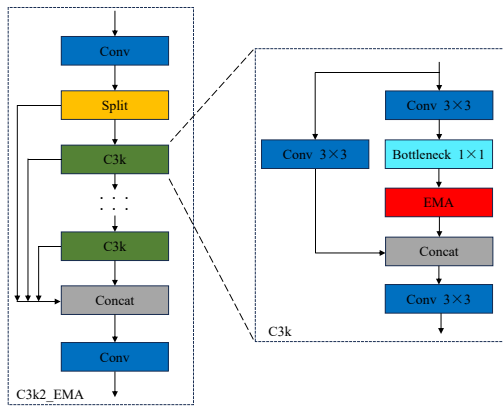


图5 C3k2_EMA 模块

Fig. 5 C3k2_EMA module

度上的协同作用提升检测性能。在尺度上，它融合多尺度特征以增强目标尺寸适应性；在空间上，它利用动态卷积与注意力机制提升模型对复杂场景的鲁棒性；同时，在任务层面，它能动态优化各子任务的特征判别性。引入DynamicHead检测头相当于为模型添加一个“自适应调控器”，使其具备动态感知，能够根据输入特征的具体情况，动态调整推理策略，在提升检测精度的同时保持了良好的泛化能力。

1.3 试验环境及评价指标

1.3.1 试验工作站与嵌入式设备

CD-YOLO 模型训练时所用的平台硬件配置为 i5 13400F CPU, NVIDIA RTX 4060 GPU, 16 GB 内存, 软件配置为 Pycharm 2024.1.4, 环境配置为 Windows 11 系统, 64 位, Python 3.8, Pytorch 1.13.0, CUDA 11.7, 训练过程所采用的参数具体信息如表 2 所示。

表2 CD-YOLO 模型训练参数

Table 2 Training parameters of CD-YOLO model

训练参数	值
训练周期	100
初始学习率	0.001
优化器	Adam
每批次图像数量	24
动量	0.937
优化器权重衰减系数	0.000 5
输入尺寸	640×640

边缘端设备采用 NVIDIA Jetson Orin NX, 配置为 6 核 Arm® Cortex®-A78AE CPU, 1024 核 NVIDIA Ampere GPU, 运行内存 8 GB。搭载系统为 Ubuntu 20.04 系统, 环境配置为 Jetpack 5.1.2, Python 3.8, Pytorch 1.12, CUDA 11.4, Ten-

sorRT 8.5.2。

1.3.2 评价指标

为实现对 CD-YOLO 在胡萝卜苗图像检测性能的精准评估, 采用精确率 (Precision, P)、召回率 (Recall, R)、平均精度 (Mean Average Precision, $mAP_{0.5}$) 关键指标。其中, mAP 在目标检测中常用不同交并比 (Intersection over Union, IoU) 阈值进行计算。本研究主要采用 $mAP_{0.5}$ (即 IoU 阈值为 0.5 时的平均精度), 该指标是评估检测框定位精度的常用标准。各指标计算如公式 (4) ~ 公式 (7) 所示。

$$P = \frac{TP}{TP + FP} \quad (4)$$

$$R = \frac{TP}{TP + FN} \quad (5)$$

$$AP = \int_0^1 P(r) dr \quad (6)$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (7)$$

式中: P 为精确率; R 为召回率; AP 为平均精度; mAP 为平均精度均值; TP 为实际为正样本, 且被模型正确预测为正样本的数量; FP 为实际为负样本, 但被模型错误预测为正样本的数量; $P(r)$ 代表精确率作为召回率的函数, 也就是当召回率为 r 时, 模型对应的精确率的值; N 表示数据集中目标的总类别数。

为了验证改进方式对模型实时性的优化, 本研究采用了计算量、单张图片处理时间和模型大小 3 个计算指标来进行模型评估。计算量指模型在一次前向传播过程中所需的浮点运算次数, 用于衡量模型的计算复杂度, 计算量越大, 需要的算力和时间成本越高; 单张图片处理时间指模型对单个样本从输入到输出推理结果所需的时间, 是模型实际部署时的关键性能指标; 模型大小是训练好的权重文件在终端设备上占用的存储空间。

2 结果与分析

2.1 消融实验

为验证改进方法的可行性, 以 YOLOv11s 为基准模型, 进行消融实验, 以上文所述指标为评价指标, 综合进行结果分析, 试验结果如表 3 所示。

由实验结果可知, 将普通卷积替换为 DWConv 后降低了模型的计算量和尺寸, 模型整体复杂度降低, 同时提高了推理速度。在颈部网络引入 C3k2_EMA 模块, 增强了模型对目标在维度和空间

表3 胡萝卜苗检测研究消融试验结果
Table 3 Results of ablation experiments for carrot seedling detection research

试验编号	改进方式			P/%	R/%	mAP _{0.5} /%	计算量/ GFLOPs	模型大小/MB	单张图片处理时间/ms
	DWConv	C3k2_EMA	DynamicHead						
1	×	×	×	78.2	74.9	81.6	21.5	19.2	14.3
2	√	×	×	79.8	73.6	81.4	14.8	13.8	11.5
3	×	√	×	78.7	77.1	82.9	21.3	19.2	11.7
4	×	×	√	79.9	76.7	83.7	21.5	19.8	12.3
5	√	√	×	79.7	74.5	81.7	14.7	13.9	10.6
6	√	×	√	79.0	74.9	81.8	15.0	14.4	9.8
7	×	√	√	80.2	75.1	83.2	21.6	19.8	11.6
8	√	√	√	81.2	76.4	84.0	15.3	14.4	9.6

注：√表示模型中采用该方式改进；×表示不采用该方式。

进行特征提取的能力，在提高了模型精度的同时降低了复杂度和单张图片处理时间；替换检测头为DynamicHead，使得模型的召回率、平均精度、处理速度等均有所提升。

分析表3数据，试验2使用DWConv模块对模型进行轻量化之后，计算量和模型大小分别减少了31.2%和28.1%，图像处理时间增加了2.8 ms，但平均精度降低了1.2个百分点，其原因在于DWConv模块缺失了通道间的信息交互，不可避免地丢失了部分特征信息，使得模型整体精度有所下降，这印证了轻量化可能带来的特征表达能力减弱问题。试验3引入C3k2_EMA模块之后，对比原YOLOv11s，其计算量减少了0.2 GFLOPs，而精确率、召回率和平均精度分别上升了0.5、2.2和1.3个百分点，单张图片的处理速度提升了2.6 ms，说明C3k2_EMA模块在一定程度降低模型复杂度的同时，还提升了模型的综合性能。试验4替换检测头为DynamicHead之后，虽然模型尺寸增加了0.6 MB，但是计算量保持不变，处理速度由原来的14.3 ms提升至12.3 ms，精确率、召回率和平均精度分别提升了1.7、1.8和2.1个百分点。由试验3和试验4可知，单独引入C3k2_EMA或DynamicHead均能在不同程度上提升检测精度，说明二者均能有效应对遮挡场景下的预测。试验5在使用DWConv轻量化方式之后，引入C3k2_EMA模块，模型平均精度提升1.3个百分点，计算量减少了0.1 GFLOPs，且单张图片的处理速度提升了0.9 ms，这表明EMA有效弥补了DWConv可能带来的特征退化，实现了“轻量化且抗损”的协同目标。试验6采用DWConv轻量化策略和动态检测头，平均精度较试验2提升1.4个百分点，单张图片处理时间降低至9.8 ms。

试验7在试验4的基础上添加了C3k2_EMA模块，召回率和平均精度有小幅下降，处理速度提升0.7 ms。试验8采用3种策略相结合的方式，模型精确率达到了81.2%，召回率为76.4%，平均精度达到了84.0%，较原始YOLOv11s模型分别提升了3.0、1.5和2.4个百分点；改进后模型计算量为15.3 GFLOPs，模型尺寸仅有14.4 MB，较原始YOLOv11s分别减少了6.2 GFLOPs和4.8 MB，下降幅度分别达到28.8%和25.0%，单张图片处理时间降低至9.6 ms，实现了精度与速度的最佳平衡。

综合看来，DWConv、C3k2_EMA与DynamicHead的协同组合，构成了一个从“特征提取轻量化”到“特征感知增强化”，再到“决策输出自适应”的完整优化闭环。DWConv的引入显著降低了模型的计算复杂度和尺寸，为后续引用相关模块腾出了算力预算，使整体模型满足边缘部署的实时性要求。嵌入在特征提取网络中的C3k2_EMA模块充当了“特征增强器”，通过高效的多尺度注意力机制，在特征提取阶段主动聚焦于目标的关键局部特征，并重建因遮挡而断裂的上下文关联，从而为后续特征融合提供了更鲁棒、信息更完整的特征图。DynamicHead作为“动态决策器”，接收并融合这些增强后的多尺度特征，通过尺度感知与空间感知机制，自适应地调整对不同位置和大小目标的检测权重，从而准确定位在复杂遮挡场景下形态各异的目标。3个模块分别从算力效率、特征完整性、表示适应性3方面入手，解决对应的子问题，轻量化是基础，特征增强是核心，动态决策是保障，该设计使得CD-YOLO在有限算力下仍能对遮挡目标保持高敏感性与高定位精度，实现了速度、精度与鲁棒性的协同提升。

为进一步体现 C3k2_EMA 模块对遮挡目标的特征提取能力的提升与上下文特征关联的增强, 将模型中间层特征图可视化, 特征图是一系列卷积核对输入图像进行卷积操作并进行特征提取的结果, 通过特征图可以反映出模型特定层所学习到的特征, 例如边界、纹理、颜色等, 特征图中黄色区域表示高激活值, 意味着模型在该区域检测到了与当前层

任务相关的显著特征, 蓝色区域表示低激活值, 表明该区域的特征与当前层任务关联性较低。在 Backbone 中, 改进的 C3k2 模块位于第 9 个特征提取层, 特征图尺寸为 20×20 像素, 分别提取 C3k2 模块和引入 EMA 机制之后的 C3k2_EMA 模块的特征图, 对比结果如图 6 所示。

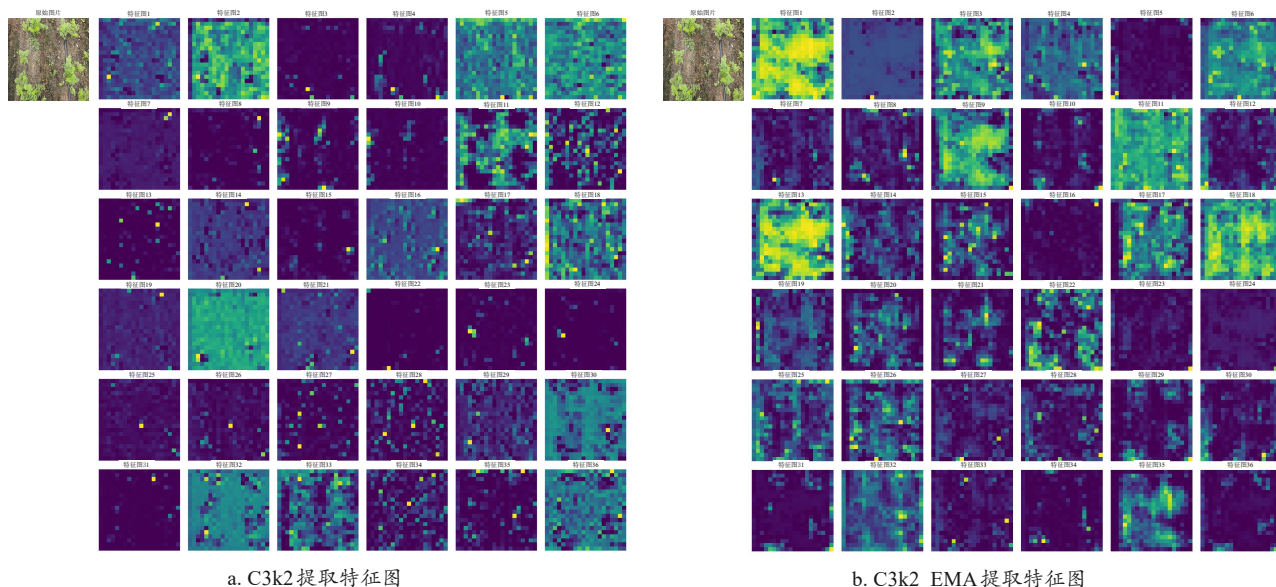


图 6 C3k2 与 C3k2_EMA 模块特征图对比

Fig. 6 Comparison of feature maps between C3k2 and C3k2_EMA modules

特征提取方面, 对比图 6a 和图 6b, 图 6b 的整体亮度更大, 特征激活程度更高, 说明在特征图提取上 C3k2_EMA 模块获取到的信息更全面, 更能提取到重要特征。面对遮挡导致的特征碎片化问题, C3k2 模块的特征整合能力有限, 如图 6a 中特征图 7、8、13、27 等所示, 其提取的叶片特征呈现离散分布, 难以形成完整表征。而 C3k2_EMA 通过其多尺度特征融合机制, 能够有效聚合这些碎片化特征, 在特征图 15、17、22、26 中可见, 离散的特征片段被整合为连贯的特征区域, 为后续增强遮挡目标碎片化特征上下文联系提供了基础。综合对比来看, 改进后的 C3k2_EMA 模块, 在相同深度下,

所捕捉到的特征信息更加全面完整, 能充分融合碎片化特征, 激活区域更加集中, 能有效地聚焦目标关键区域, 有利于在遮挡条件下保持更高激活强度和更丰富细节, 提高检测精度。

2.2 轻量化主干试验

本试验以改进 CD-YOLO 为基础网络, 对比该主干网络与其他主流轻量化特征提取主干网络对模型的影响。选择 MobileNetv3^[23]、ShuffleNetv2^[24]、EfficientVit^[25] 特征提取网络对模型进行主干替换, 训练策略相同, 对比不同主干的训练效果, 试验结果如表 4 所示。

表 4 不同轻量化主干网络检测胡萝卜苗效果对比

Table 4 Comparison of different lightweight backbone network effects for carrot seedling detection

主干网络	P/%	R/%	mAP _{0.5} /%	计算量/GFLOPs	参数量	模型大小/MB	单张图片处理时间/ms
MobileNetv3	76.2	73.0	79.9	13.2	7.21×10^6	16.9	14.6
ShuffleNetv2	73.0	72.3	77.8	10.4	5.34×10^6	11.0	12.2
EfficientVit	76.0	74.5	79.7	14.6	7.39×10^6	15.7	11.4
CD-YOLO	81.2	76.4	84.0	15.3	7.06×10^6	14.4	9.6

由表中结果可知, 本研究所提出的轻量化模型 CD-YOLO, 相较于替换 MobileNetv3、Shuffle-Netv2、EfficientVit 主干网络方式, 在精确率、召回率、平均精度和处理速度上都具有优势。平均精度相较于其他主干网络分别提升了 4.1、6.2、4.3 个百分点, 单张图片处理速度则分别加快了 5.0、2.6、1.8 ms。

2.3 主流模型对比

为全面地评估 CD-YOLO 模型在胡萝卜苗图像检测任务中的优势, 本研究将 CD-YOLO 与 SSD^[26]、更快基于区域的卷积神经网络 (Faster Reigon-Based Convolutional Neural Network, Faster-RCNN)^[27]、YOLOv5s^[28]、YOLOv8s^[29]、YOLOv11s 等主流目标检测模型和 YOLOv8s-P2、DWG-YOLOv8^[30]、HAD-YOLO 等针对遮挡的检测模型进行了对比试验, 试验结果如表 5 所示。

表5 胡萝卜苗检测研究不同模型对比效果
Table 5 Comparison results of different models for carrot seedling detection research

模型	P/%	R/%	mAP _{0.5} /%	计算量/ GFLOPs	模型大 小/MB	单张图片处 理时间/ms
SSD	63.5	81.4	72.2	63.5	90.5	36.6
Faster R-CNN	69.1	84.5	77.7	142.6	108.1	113.3
YOLOv5s	80.2	73.9	80.5	16.2	15.1	14.2
YOLOv8s	79.1	74.8	81.6	49.1	22.5	9.8
YOLOv11s	78.2	74.9	81.6	21.5	19.2	14.3
YOLOv8s-P2	79.3	74.4	81.5	55.3	24.8	10.7
DWG-YOLOv8	78.2	73.2	82.0	17.4	14.7	10.5
HAD-YOLO	77.9	74.3	81.1	35.4	12.5	19.3
CD-YOLO	81.2	76.4	84.0	15.3	14.4	9.6

由表 5 可知, CD-YOLO 相较于其他模型在精确率、召回率、平均精度等方面均具有一定优势, 平均精度对比 SSD、Faster-RCNN、YOLOv5s、YOLOv8s、YOLOv11s、YOLOv8s-P2、DWG-YOLOv8、HAD-YOLO 模型在胡萝卜苗数据集上分别领先 11.8、6.3、3.5、2.4、2.4、2.5、2.0、2.9 个百分点; 同时 CD-YOLO 在计算量上均小于其他模型, 模型大小也小于绝大多数模型, 其中模型大小较 SSD、Faster-RCNN、YOLOv5s、YOLOv8s、YOLOv11s、YOLOv8s-P2 和 DWG-YOLOv8 分别减少了 84.1%、86.7%、4.6%、36.0%、25.0%、41.9%、2.1%, 处理速度快于其他模型, 单张图片处理速度较其他模型分别提升了 27.0、103.7、4.6、

0.2、4.7、1.1、0.9、9.7 ms。从试验结果来看, 无论是主流目标检测模型, 还是遮挡检测模型, 在含有遮挡, 以及多种田间条件的验证集上的综合表现都不及 CD-YOLO。由此可知, 改进后的 CD-YOLO 模型兼具精度和速度优势, 在遮挡和复杂多变的田间环境下表现更好, 更适用于田间环境下的胡萝卜苗检测任务。

2.4 测试集表现

为初步探究模型对田间常见变化的适应性, 本研究对整体测试集进行了测试, 其中包含不同土壤状态、光照强度差异等情况。部分模型在测试集不同情况下的表现如图 7 所示。

通过不同模型在测试集上的效果对比发现, 在土壤潮湿条件下, CD-YOLO 对胡萝卜苗的预测结果较为准确, 而 YOLOv11s 存在着因目标显示不全导致的漏检, DWG-YOLOv8 存在着漏检和密集遮挡目标的重复检测问题。在干燥土壤条件下, CD-YOLO 能将图片中的目标很好地检测出来, YOLOv11s 和 DWG-YOLOv8 只能完成单体目标的准确检测, 对遮挡较为严重目标仍存在漏检和框选不全的问题。在不同光照条件下, 光照强度变化导致成像中损失一定纹理特征信息, 而 CD-YOLO 由于引入 EMA 模块, 增强了多尺度特征提取, 对亮度变化、纹理缺失的目标具有更好的感知能力, 因此 CD-YOLO 在不同光照条件下的目标检测都表现出优异的效果, 反观 YOLOv11s 和 DWG-YOLOv8, 因其缺少上下文信息的联系重构, 导致在特征信息不全的基础上无法学习到一系列特征的上下文联系, 进而出现漏识现象。在小目标的预测结果上, CD-YOLO 与 DWG-YOLOv8 均能实现较好的预测结果, 而 YOLOv11s 对尺度较小且特征不明显的目标无法完成预测, 导致漏识现象发生。综合来看, CD-YOLO 在不同土壤条件、不同光照条件下检测效果最好, 显示出对单一基地内主要干扰因素的有效应对能力, 同时也表明模型具备了一定的场景内泛化能力。然而, 对于跨季节、跨地域等引起的更具根本性的特征分布变化, 其泛化性能仍需通过更广泛的数据进行验证, 后续研究将通过采集不同时期、不同地域, 以及更复杂的田间变化条件的数据, 来进一步提升模型的泛化性和鲁棒性。

2.5 遮挡效果检测

基于前面定义的遮挡量化标准, 对测试集中所有 479 张图像进行了遮挡量化处理, 筛选出所有包

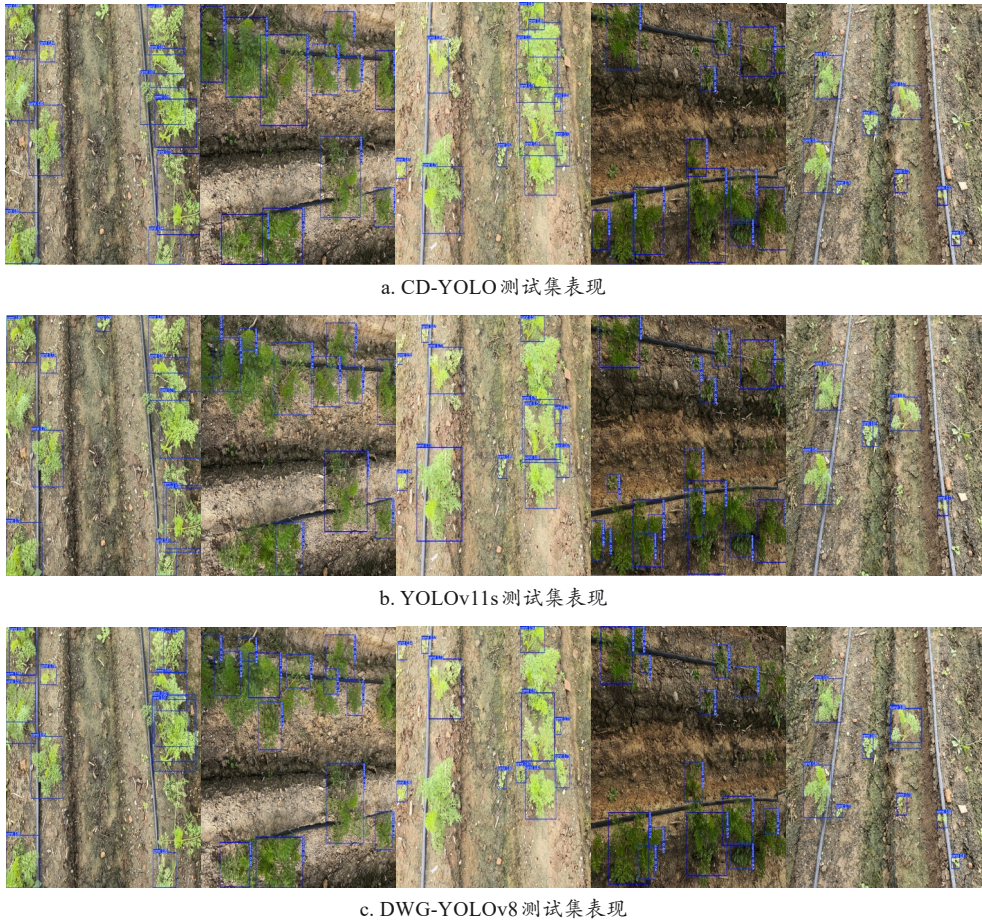


图7 不同模型在胡萝卜苗测试集上的效果对比

Fig.7 Comparative results of different models on the carrot seedling test set

含轻度及以上遮挡的目标图像，构成“遮挡测试子集”，共得到 122 张图像，约占整体测试集的 25.5%，共包含 1 652 个胡萝卜苗目标，将该子集作为评估模型抗遮挡能力的核心数据集。剩余图像构成“常规测试子集”，用于评估模型在相对理想条件下的综合性能。

为验证本研究所提模型在遮挡条件下的识别效果，本研究对遮挡测试子集和常规测试子集分别进行了测试，实验结果如表 6 所示。

由表 6 可知，在遮挡子集上，CD-YOLO 的 P 、 R 和 $mAP_{0.5}$ 分别为 79.6%、75.1% 和 80.6%，相较于 YOLOv11s 在遮挡子集上分别提升了 3.7、4.0 和 5.1 个百分点，而在常规子集上分别提升了 1.1、1.8 和 1.8 个百分点，在整体测试集上的精确率、召回率和平均精度则分别提升了 0.8、1.5 和 3.0 个百分点。经对比发现，在遮挡子集上各项指标的提升效果大于在常规子集和整体测试集上的表现。由此可知，本研究针对遮挡所采取的模型优化方式得到验证。对遮挡子集中的图片进行结果统计，统计 YO-

表 6 YOLOv11s 与 CD-YOLO 在不同测试子集上效果对比

Table6 Comparative results of YOLOv11s and CD-YOLO performance across different test subsets

测试子集	模型	图像数量	$P/\%$	$R/\%$	$mAP_{0.5}/\%$
常规子集	YOLOv11s	357	78.7	76.3	77.6
	CD-YOLO	357	79.8	78.1	79.4
遮挡子集	YOLOv11s	122	75.9	71.1	75.5
	CD-YOLO	122	79.6	75.1	80.6
整体测试集	YOLOv11s	479	78.9	74.7	76.9
	CD-YOLO	479	79.7	76.2	79.9

LOv11s 和 CD-YOLO 的识别准确数与漏识数，测试结果如表 7 所示。

改进前 YOLOv11s 的漏识数为 315，正确识别数为 1 337，漏识率和识别率分别为 19.1% 和 80.9%。改进后 CD-YOLO 漏识数为 221，正确识别数为 1 431，漏识率和识别率分别为 13.4% 和 86.6%，相比原模型漏识率下降了 5.7 个百分点。由此可知，改进后 CD-YOLO 模型在遮挡子集数据上

表7 YOLOv11s与CD-YOLO遮挡条件下胡萝卜苗检测效果对比

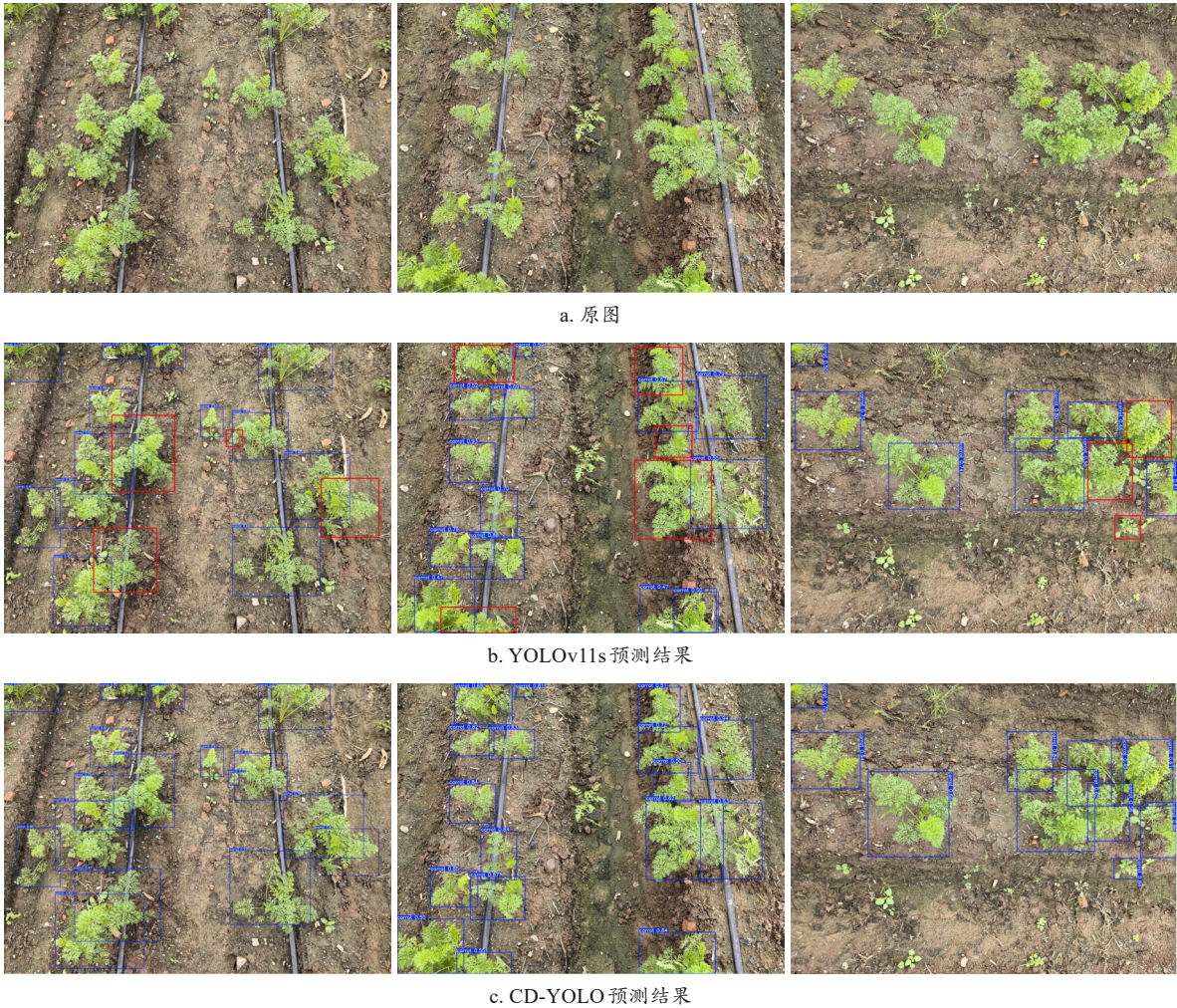
Table 7 Comparison of carrot seedling detection performance between YOLOv11s and CD-YOLO under occlusion conditions

模型	目标总数/个	正确识别数/个	漏识数/个	识别率/%	漏识率/%
YOLOv11s	1 652	1 337	315	80.9	19.1
CD-YOLO	1 652	1 431	221	86.6	13.4

漏识率和识别率都领先于原始YOLOv11s，在遮挡条件下表现出更好的准确性和鲁棒性。

模型部分识别效果对比如图8所示。由图8b可知，YOLOv11s在遮挡条件下易出现漏识现象，特别是多目标形态互掩下，漏识现象较为严重，因为YOLOv11s预测时所提取的特征信息缺失局部关键特征，无法突出被掩盖目标的关键部分，且不能

分离混合特征，导致无法实现对混合特征目标的预测。观察图8c发现，CD-YOLO模型对遮挡条件下的胡萝卜苗可以实现较好的检测效果，通过引入C3k2_EMA模块与DynamicHead结构，模型的特征提取与多尺度信息融合能力增强，有效提升了遮挡目标的识别效果。C3k2_EMA模块有效扩大了感受野同时强化了对关键特征的关注，即使在目标部分可见的情况下仍能提取具有判别力的局部特征。同时，DynamicHead通过自适应权重融合不同尺度的特征图，增强了模型对多尺度上下文信息的整合能力，从而在目标被遮挡时能够综合利用其可见部分与周围上下文信息进行推理。二者的协同作用使CD-YOLO模型在遮挡条件下仍能准确捕捉目标的本质特征，因此在检测性能上优于YOLOv11s模型。



注:图中蓝色框为检测框,红色框为漏检目标。

图8 YOLOv11s与CD-YOLO在遮挡条件下胡萝卜苗检测效果对比

Fig. 8 Comparison of carrot seedling detection performance between YOLOv11s and CD-YOLO under occlusion conditions

2.6 热力图可视化

为了更加直观地观察到改进方式对模型识别效果的提升, 绘制热力图^[31]。热力图用于表示模型预测过程中对不同位置的关注程度, 能直观地反映

出模型对图像的感兴趣区域。热力图的颜色代表了模型特征提取情况, 红色区域代表对模型预测很重要, 蓝色则表示该区域不重要, 改进前后的热力图如图9所示。

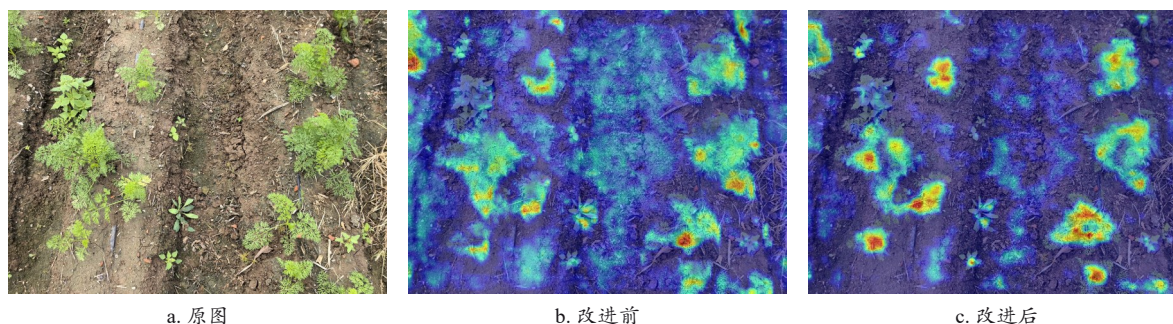


图9 YOLOv11s 与 CD-YOLO 模型热力图可视化对比

Fig. 9 Visual comparison of heatmaps between YOLOv11s and CD-YOLO models

对比热力图, 原始 YOLOv11s 的目标区域较背景区域颜色更亮, 但背景中的部分区域也有一定亮度, 说明模型对于目标特征提取的聚焦能力不足, 且对于遮挡的目标赋予相同的注意力, 易导致模型将多个遮挡重叠目标预测为一个目标。引入 C3k2_EMA 模块并添加 DynamicHead 动态检测头之后, 图片中胡萝卜目标区域颜色更亮, 覆盖程度更高, 背景区域亮度变低, 对于相互遮挡的目标也分别赋予较高程度的关注。C3k2_EMA 模块增强了模型的特征提取能力, 对目标特征信息的提取也更加完善, 而动态检测头 DynamicHead 使得模型对于通道信息和空间信息的关注度提升, 对于正确目标的检测感知增强, 同时也抑制了非目标区域对于模型整体预测能力的影响。热力图可视化显示, 改进后的模型在目标区域激活更强, 且目标的覆盖程度更高, 对背景信息的抑制更明显, 表明 C3k2_EMA 与 DynamicHead 能有效提升模型对目标的注意力集中度, 进而提升检测精度。

2.7 边缘端部署与检测效果

实际作业时, 模型往往部署于边缘端算力设备, 与工作站相比, 边缘算力设备往往面临着算力不足的问题。为了验证 CD-YOLO 在边缘端的检测情况, 同时也为了加速推理速度, 本研究采用 TensorRT 库^[32]在边缘端进行推理加速。将训练好的 *.pt 模型文件转换为 *.onnx 格式, 并在边缘端优化编译为 *.engine 文件。执行检测任务时, 加载 *.engine 文件就可实现推理加速操作。

为准确测试模型在边缘设备上的实时推理性

能, 本研究所有帧率与单张图像处理时间的测试均在统一的测试条件下进行, 模型输入像素为 640×640 , 采用批处理大小为 1, 以模拟实际单张图像流式处理的场景, 使用相同的 predict.py 测试脚本, 分别在工作站和 Jetson Orin NX 上运行。单张图片处理时间包含了数据预处理、前向推理, 以及后处理的全流程耗时。帧率 (Frames Per Second, FPS) 由该流程的单次平均耗时计算得出。模型部署在工作站和 Jetson Orin NX 的帧率情况如表 8 所示。

表 8 YOLOv11s 与 CD-YOLO 模型在不同设备上检测帧率对比

Table 8 Frame rate comparison between YOLOv11s and CD-YOLO models on different devices

模型	工作站	边缘算力设备	单位: 帧/s
			TensorRT
YOLOv11s	89.6	10.5	29.7
CD-YOLO	104.2	12.9	32.5

由表 8 可知, 在 TensorRT 加速之前, 由于边缘端设备算力有限, 处理速度无法与工作站相比, 实时检测帧率较低, 只有 12.9 帧/s, 加速之后帧率提升到 32.5 帧/s, 检测速度提升 2.5 倍, 单张图片处理速度达到 30.7 ms。

将 YOLOv11s 和 CD-YOLO 模型部署于 Jetson Orin NX 之后, 随机选择胡萝卜苗测试集图像进行测试, 统计模型的正确识别数、误识数、漏识数, 结果如表 9 所示。

表 9 展示了模型改进前后在田间胡萝卜苗图片数据上的识别情况。在目标总数为 492 的图片数据中, 改进前 YOLOv11s 的漏识数为 63, 误识数为

表9 YOLOv11s与CD-YOLO边缘端上胡萝卜苗检测效果对比

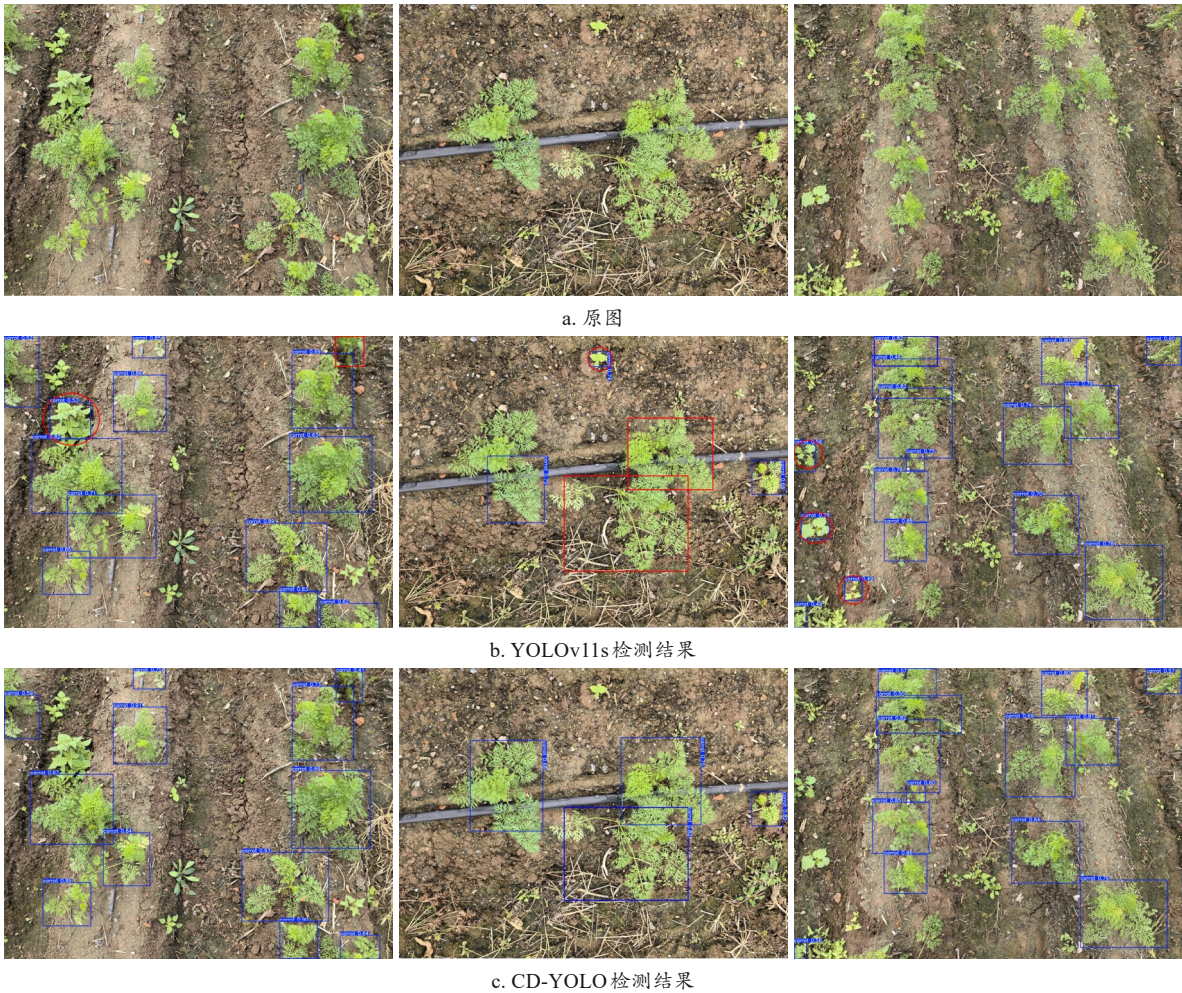
Table 9 Comparison of carrot seedling detection performance between YOLOv11s and CD-YOLO on edge devices

模型	目标总数/个	识别数/个	正确识别数/个	漏识数/个	误识数/个	漏识率/%	误识率/%
YOLOv11s	492	453	429	63	24	12.8	5.3
CD-YOLO	492	480	467	25	13	5.1	2.7

24, 漏识率和误识率分别为12.8%和5.3%。改进后CD-YOLO漏识数为25, 误识数仅有13, 漏识率和误识率分别为5.1%和2.7%, 相比原模型下降了7.7个和2.6个百分点。由此可知, 改进后的模型不论是在漏检率还是误识率上都优于原始YOLOv11s,

在真实田间环境条件下, CD-YOLO表现出了更好的鲁棒性和准确性。

图10为部分YOLOv11s与CD-YOLO部署于边缘端NVIDIA Jetson Orin NX的检测结果。



注:图中蓝色框为检测框,红色框为漏检目标,红色圈为误检目标。

图10 YOLOv11s与CD-YOLO边缘端上胡萝卜苗检测效果对比

Fig. 10 Comparison of carrot seedling detection performance between YOLOv11s and CD-YOLO on edge devices

由图10可知, YOLOv11s存在漏检、误检的情况, 对杂草和遮挡条件下的胡萝卜苗检测未能表现出很好的准确性和鲁棒性。而改进后的CD-YOLO对胡萝卜苗检测有着较好的准确率, 在杂草和遮挡条件下也有着较好表现。YOLOv11s的检测结果如图10b。图10b左边图像中存在1处误检、1处

漏检, 由于图中杂草的多株簇拥, 在轮廓、颜色上与胡萝卜苗具有较高相似性, 且YOLOv11s低层次与高层次特征信息融合时会存在一定程度的丢失, 因此缺失了低层次的纹理信息, 导致形态近似的杂草被误识为胡萝卜苗; 观察图中的漏检目标可知, 该目标未能完整出现在图像中, 仅有部分茎秆信

息, 缺少基本的轮廓和纹理特征信息, 重要特征信息的缺失导致该目标无法被基线模型识别。图 10b 中间图像中存在 1 处误检、2 处漏检, 基线模型 YOLOv11s 注意力倾向于关注图像中的“显著性区域”(如绿色区域), 而非具有区分性的“关键局部”(如叶片纹理), 且图中杂草尺寸较小, 所含有的绿色通道信息权重大于细节纹理信息权重, 因此模型易将小目标杂草检测为小目标胡萝卜苗; 图 10b 中间图像中的 2 处漏检目标为粘连目标, 像素区域高度重叠, 模型难以区分单个目标的“空间边界”, 锚框回归会因空间信息的重叠而产生歧义, 导致预测框无法准确包围单个目标, 甚至完全漏检。图 10b 右边图像中存在 3 处误检, 3 个误检目标均为小尺寸杂草, 存在的纹理特征信息较少, 颜色特征信息权重较大, 模型对其预测时易检测为小尺度胡萝卜目标。特别注意到, 图 10c 中尺度更小的杂草未被错误检测为胡萝卜苗目标, 原因在于多尺度特征融合机制对目标尺度存在“感知阈值”, 更小尺度的杂草在像素面积、语义信息(纹理复杂度等)上低于模型对胡萝卜苗目标的最小尺度预期, 在候选框生成阶段, 其特征信息会因尺度不匹配而被抑制, 因此不会激活模型的目标判定逻辑。进一步观察, 多数误检目标的置信度较其他胡萝卜目标偏低, 这从一定程度上说明了杂草与胡萝卜苗

存在一定的隐形差异, 也从一定程度上反映了模型对不同样本的认知不确定性。

CD-YOLO 通过 C3k2_EMA 模块和 DynamicHead 检测头的联合作用, 在存在杂草和遮挡条件下的胡萝卜苗识别中有更好的表现。在杂草条件下, C3k2_EMA 模块可对胡萝卜特征赋予更高的权重, 同时对杂草的“干扰特征”进行抑制。DynamicHead 检测头可针对每个目标与非目标的形态特异性动态生成匹配的分类子网络, 在各种场景下保持分类准确率。对于存在一定程度的遮挡黏连, C3k2_EMA 模块可捕捉目标不同区域的碎片化与整体化联系的特征信息, 进而获取更多的特征。DynamicHead 检测头可根据目标的实时特征自适应调整检测头的结构, 使模型能更精准地捕捉单个目标的“空间边界”, 避免因黏连导致的边界框融合或漏检。综合对比 YOLOv11s 和 CD-YOLO 在边缘端的识别结果, CD-YOLO 表现更佳, 更能适应胡萝卜苗检测的任务。

2.8 失败案例分析

尽管 CD-YOLO 模型在一般检测场景中展现出较好的识别精度和鲁棒性, 但在某些极端物理条件下仍面临检测失败的挑战。如图 11 所示, 当测试样本出现严重运动模糊时, 该模型也会发生漏检现象。

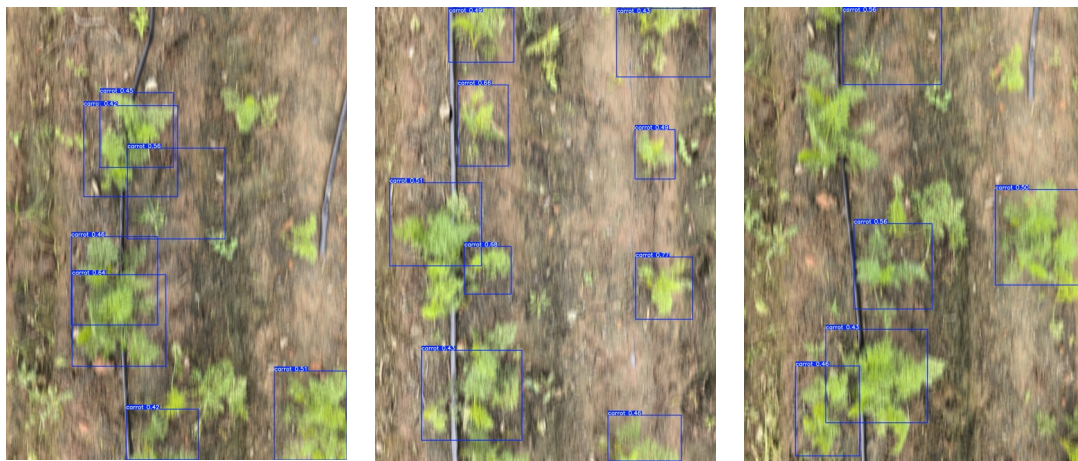


图 11 强烈运动模糊下的 CD-YOLO 对胡萝卜苗检测效果

Fig. 11 CD-YOLO performance in carrot seedling detection under severe motion blur

对模型出现漏识现象进行分析, 首先, 从模糊成因来看, 本研究采集图像时所采用的摄像头最高帧率仅有 30 帧/s, 单帧的曝光时间最大为 33.3 ms, 在速度为 0.6 m/s 时, 相机与场景之间的最小相对位移可达 2 cm, 该物理位移在图像传感器平面上直接

表现为像素级的线性拖影, 导致目标的高频空间特征缺失, 从而造成捕获图像自身所含信息缺失。其次, 运动模糊会导致成像质量整体退化, 主要表现为空间域中边缘与纹理结构的高频信息损失, 以及图像对比度的显著下降。在通道层面, 模糊导致不

同颜色与亮度通道间的区分度减弱,使得模型在提取通道注意力时难以聚焦于有效目标区域;在空间维度上,由于目标与背景的梯度差异减小,模型在构建特征金字塔时,深层特征图中包含的语义信息与浅层特征中的几何细节均受到抑制,从而削弱了多尺度特征融合的有效性。最后,从目标特性角度分析,运动模糊会导致图像中胡萝卜苗的轮廓边缘产生弥散效应,叶片区域的纹理细节在频域中高频成分衰减,导致模型在卷积过程中所捕获关键特征减少,进而出现由于特征信息不足而导致的漏识现象。正是由于这种图片质量下降、目标特征信息的缺失,使得模型在特征提取过程中难以准确区分目标与背景,最终导致漏识率的上升。

为提升本研究在高速运动场景下的胡萝卜苗检测效果,后续研究可从成像硬件优化、图像传输与预处理、模糊自适应模型3个层面展开系统性优化:采用帧率更高的摄像头,从图像获取上减少运动模糊;图像输入模型前,加入基于“编码-解码”的去模糊操作,保证成像质量;为模型设计自适应特征提取网络,引入频域分析分支与增强多尺度特征融合能力,提升模型对模糊图像的固有鲁棒性和特征提取能力。

3 结 论

1) 制作胡萝卜苗期数据集,对YOLOv11s深度学习模型进行了轻量化改进,获取了目标检测算法CD-YOLO,在验证集上的预测准确率为81.2%,召回率为76.4%,平均精度为84.0%,相较于YOLOv11s在验证集上的表现分别提升了3.0、1.5、2.4个百分点,单张图片处理时间9.6 ms,计算量15.3 GFLOPs,模型大小仅14.4 MB。

2) 深度可分离卷积DWConv替换普通卷积CBS,使CD-YOLO模型大小减少28.1%,计算量下降31.2%,相对于MobileNetV3、ShuffleNetV2、EfficientViT等轻量化主干网络,mAP分别提高4.1、6.2、4.3个百分点。在C3k2模块中引入EMA机制形成C3k2_EMA模块,替换检测头为DynamicHead,遮挡黏连条件下的检测漏识率为13.4%,较原YOLOv11s下降5.7个百分点。

3) 模型边缘端检测帧率达到了32.5帧/s,在随机选择的胡萝卜苗测试图片中,CD-YOLO的漏识率和误识率为5.1%和2.7%,较YOLOv11s分别下降了7.7和2.6个百分点。对比YOLOv11s和CD-YOLO在边缘端设备上的表现,CD-YOLO的准确

性和处理速度优于YOLOv11s,综合表现更佳,更能满足田间作业的准确性和实时性要求。

数据可用性声明: 本研究所使用的数据集可通过合理请求向作者获取。

利益冲突声明: 本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献:

- [1] 刘星,黄建新,欧承刚,等.胡萝卜根色及其色素组分的遗传和育种研究进展[J].植物遗传资源学报,2022,23(5): 1241-1248.
LIU X, HUANG J X, OU C G, et al. Current advances on inheritance and breeding of carrot root color and its pigment components[J]. Journal of Plant Genetic Resources, 2022, 23(5): 1241-1248.
- [2] 赵童,米月花,籍镭钊,等.胡萝卜收割机的结构优化设计[J].工程机械,2025,56(6): 154-156, I0008.
ZHAO T, MI Y H, JI L F, et al. Structural optimization design of carrot harvester[J]. Construction Machinery and Equipment, 2025, 56(6): 154-156, I0008.
- [3] 张清蓉,王国栋,赵正伟,等.基于自动控制技术的胡萝卜种植收割一体机设计[J].南方农机,2024,55(21): 46-50.
ZHANG Q R, WANG G D, ZHAO Z W, et al. Design of carrot planting and harvesting integrated machine based on automatic control technology[J]. South Agricultural Machinery, 2024, 55(21): 46-50.
- [4] 倪建功,李娟,邓立苗,等.基于知识蒸馏的胡萝卜外观品质等级智能检测[J].农业工程学报,2020,36(18): 181-187.
NI J G, LI J, DENG L M, et al. Intelligent detection of carrot appearance quality grade based on knowledge distillation[J]. Transactions of the Chinese Society of Agricultural Engineering, 2020, 36(18): 181-187.
- [5] XIE W J, WEI S, ZHENG Z H, et al. Recognition of defective carrots based on deep learning and transfer learning[J]. Food and Bioprocess Technology, 2021, 14(7): 1361-1374.
- [6] 王春桃,梁伟健,郭庆文,等.农业害虫智能视觉检测研究综述[J].中国农机化学报,2023,44(7): 207-213.
WANG C T, LIANG W J, GUO Q W, et al. Summary of research on intelligent vision detection of agricultural pests[J]. Journal of Chinese Agricultural Mechanization, 2023, 44(7): 207-213.
- [7] 黄友锐,王小桥,韩涛,等.基于改进YOLOv8n的甜菜杂草检测算法研究[J].江苏农业科学,2024,52(24): 196-204.
HUANG Y R, WANG X J, HAN T, et al. A detection method for sugar beets and weeds based on improved YOLOv8n algorithm[J]. Jiangsu Agricultural Sciences, 2024, 52(24): 196-204.
- [8] 曲福恒,李金状,杨勇,等.基于改进DeepLabv3+的轻量化作物杂草识别方法[J].石河子大学学报(自然科学版),2024,42(1): 117-125.
QU F H, LI J Z, YANG Y, et al. Lightweight crop and weed recognition method based on improved DeepLabv3+[J]. Journal of Shihezi University (Natural Science), 2024, 42(1): 117-125.
- [9] NIU L T, SU W H, ZHANG H Y, et al. Development of intelligent equipment for weed identification and variable spraying in lettuce fields based on instance segmentation framework[J]. Engineering Applications of Artificial Intelligence, 2025, 159: 111634.
- [10] 孟庆宽,张漫,杨晓霞,等.基于轻量卷积结合特征信息融合的玉米幼苗与杂草识别[J].农业机械学报,2020,51(12): 238-245, 303.
MENG Q K, ZHANG M, YANG X X, et al. Recognition of maize seedling and weed based on light weight convolution and feature fusion[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(12): 238-245, 303.
- [11] 张志远,罗铭毅,郭树欣,等.基于改进YOLOv5的自然环境下樱桃果实识别方法[J].农业机械学报,2022,53(S1): 232-240.
ZHANG Z Y, LUO M Y, GUO S X, et al. Cherry fruit detection

- method in natural scene based on improved YOLOv5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53 (S1): 232-240.
- [12] ZHANG B Y, ZHANG F K, AN S, et al. SCORE-DETR: An efficient Transformer-based network for small and occluded citrus detection[J]. Computers and Electronics in Agriculture, 2025, 238: 110843.
- [13] 汤晨, 刘振青, 邵阳, 等. 基于改进 YOLOv11n 的密集遮挡环境百香果识别方法[J]. 农业机械学报, 2026, 57 (5): 167-176.
TANG C, LIU Z Q, SHAO Y, et al. Passion fruit recognition method in densely occluded environments based on improved YOLOv11n[J]. Transactions of the Chinese Society for Agricultural Machinery, 2026, 57 (5): 167-176.
- [14] 李文峰, 胡世康, 杨琳琳, 等. 基于轻量化 YOLOv4 对不同遮挡程度成熟番茄果实的识别[J]. 云南农业大学学报(自然科学版), 2024(4): 184-189.
LI W F, HU S K, YANG L L, et al. Recognition of mature tomato fruits with different occlusion degrees based on lightweight YOLOv4[J]. Journal of Yunnan Agricultural University (Natural Science), 2024(4): 184-189.
- [15] 王元昊, 姜欢欢, 罗红品, 等. 基于改进 YOLOv8 算法对被遮挡柑橘的识别与定位优化[J]. 西南大学学报(自然科学版), 2025, 47(2): 171-183.
WANG Y H, LOU H H, LUO H P, et al. Recognition and location optimization of shaded Citrus based on improved YOLOv8 algorithm [J]. Journal of Southwest University (Natural Science), 2025, 47(2): 171-183.
- [16] 李会, 郭家文, 黄世醒, 等. 基于改进 YOLOv7 的甘蔗幼苗检测方法试验研究[J]. 农机化研究, 2025, 47(9): 146-154.
LI H, GUO J W, HUANG S X, et al. Experiment on sugarcane seedling detection method based on improved YOLOv7[J]. Journal of Agricultural Mechanization Research, 2025, 47(9): 146-154.
- [17] 郑健林, 黄世醒, 郑丁科, 等. 基于改进 YOLOv5 的机收蔗含杂率检测方法试验研究[J]. 农机化研究, 2026, 48(2): 217-224.
ZHENG J L, HUANG S X, ZHENG D K, et al. Experimental study on impurity content detection method of machine-harvested sugarcane based on improved YOLOv5[J]. Journal of Agricultural Mechanization Research, 2026, 48(2): 217-224.
- [18] 牛子昂, 裴正军. 基于改进 YOLOv11-Pose 的玉米植株骨架及表型参数提取方法[J]. 智慧农业(中英文), 2025(2): 95-105.
NIU Z A, QIU Z J. Extraction method of maize plant skeleton and phenotypic parameters based on improved YOLOv11-Pose [J]. Smart Agriculture, 2025(2): 95-105.
- [19] 谭泗桥, 陈涵, 朱磊, 等. 基于改进 YOLOv8m 的稻田害虫识别方法[J]. 农业工程学报, 2025, 41(2): 185-195.
TAN S Q, CHEN H, ZHU L, et al. Identification method of rice pests based on improved YOLOv8m[J]. Transactions of the Chinese Society of Agricultural Engineering, 2025, 41(2): 185-195.
- [20] HOWARD A G, ZHU M L, CHEN B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications[EB/OL]. arXiv: 1704.04861, 2017.
- [21] 李亚, 蒋晨, 王海瑞, 等. 基于 EDW-YOLOv8 的棉花叶片病害检测[J]. 华中农业大学学报, 2025, 44(5): 189-197.
LI Y, JIANG C, WANG H R, et al. Cotton leaf disease detection based on EDW-YOLOv8[J]. Journal of Huazhong Agricultural University, 2025, 44(5): 189-197.
- [22] DENG L, MIAO Z H, ZHAO X G, et al. HAD-YOLO: An accurate and effective weed detection model based on improved YOLOv5 network [J]. Agronomy, 2025, 15(1): 57.
- [23] DENG J L, LIANG Q, HE J J, et al. Flavor grading of zanthoxylum based on computer vision-multi-chromatography fusion [J]. Journal of Food Composition and Analysis, 2025, 148: 108323.
- [24] 刘坤, 吉宏亚, 黄程非, 等. 基于改进 YOLOv5s 的番茄成熟度识别技术研究[J]. 中国农机化学报, 2025, 46(5): 79-85.
LIU K, JI H Y, HUANG C F, et al. Research on tomato maturity recognition technology based on improved YOLOv5s[J]. Journal of Chinese Agricultural Mechanization, 2025, 46(5): 79-85.
- [25] 曹玉莹, 刘银川, 高新悦, 等. LightTassel-YOLO: 一种基于无人机遥感的玉米穗实时检测方法(英文)[J]. 智慧农业(中英文), 2025, 7(6): 96-110.
CAO Y Y, LIU Y C, GAO X Y, et al. LightTassel-YOLO: A real-time detection method for maize tassels based on UAV remote sensing[J]. Smart Agriculture, 2025, 7(6): 96-110.
- [26] 李大华, 孔舒, 李栋, 等. 基于改进 SSD 模型的柑橘叶片病害轻量化检测模型[J]. 浙江农业学报, 2024, 36(3): 662-670.
LI D H, KONG S, LI D, et al. Lightweight detection model of citrus leaf diseases based on improved SSD model[J]. Acta Agriculturae Zhejiangensis, 2024, 36(3): 662-670.
- [27] REN S Q, HE K M, 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, 2017, 39(6): 1137-1149.
- [28] WEN F, WU H, ZHANG X X, et al. Accurate recognition and segmentation of northern corn leaf blight in drone RGB Images: A CycleGAN-augmented YOLOv5-Mobile-Seg lightweight network approach[J]. Computers and Electronics in Agriculture, 2025, 236: 110433.
- [29] JIA X F, HUA Z L, SHI H T, et al. A soybean pod accuracy detection and counting model based on improved YOLOv8[J]. Agriculture, 2025, 15(6): 617.
- [30] LIU H R, WANG Y, ZHAI C Y, et al. DWG-YOLOv8: A lightweight recognition method for broccoli in multi-scene field environments based on improved YOLOv8s[J]. Agronomy, 2025, 15 (10): 2361.
- [31] 李茂, 肖洋帆, 宗望远, 等. 基于改进 YOLOv8 模型的轻量化板栗果实识别方法[J]. 农业工程学报, 2024, 40(1): 201-209.
LI M, XIAO Y Y, ZONG W Y, et al. Detecting chestnuts using improved lightweight YOLOv8[J]. Transactions of the Chinese Society of Agricultural Engineering, 2024, 40(1): 201-209.
- [32] DHASARATHAN C, GNANASEKARAN S, PATTANAYAK A, et al. Tensor RT optimized driver drowsiness detection system using edge device[J]. Ain Shams Engineering Journal, 2025, 16 (10): 103620.

CD-YOLO: A Method for Detecting Carrot Seedlings in Field Based on Improved YOLOv11s

LIU Haoran^{1,2}, WANG Yu¹, ZHAO Xueguan^{2,4}, WU Huarui³, FU Hao^{2,4},
PANG Shujie⁵, ZHAI Changyuan^{2,4*}

(1. School of Mechanical Engineering and Automation, University of Science and Technology Liaoning, Anshan 114051, China; 2. Intelligent Equipment Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China; 3. Information Technology Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China; 4. National Engineering Research Center of Intelligent Equipment for Agriculture (NERCIEA), Beijing 100097, China; 5. Chinese Academy of Agricultural Mechanization Sciences Group Co., Ltd., Beijing 100083, China)

Abstract: [Objective] In field environments under natural conditions, leaf occlusion and mutual plant shading pose significant challenges to the accurate identification of carrot seedlings. Furthermore, practical agricultural applications often rely on edge devices with limited computational power, necessitating a detection model that combines lightweight design, high accuracy, and robust anti-occlusion capability. The purpose of this research is to develop a robust recognition method for carrot seedlings suitable for complex field conditions, thereby enhancing the accuracy and efficiency of seedling emergence statistics in automated seedling raising processes and providing reliable technical support for precise farm management. **[Methods]** The CD-YOLO (Carrot Detection-You Only Look Once), a lightweight detection model was proposed based on an improved YOLOv11s. First, to reduce model complexity, several standard convolutions in the backbone network were replaced with depthwise separable convolutions (DWConv), thereby decreasing floating-point operations (FLOPs) and the number of parameters, establishing a lightweight foundation for edge deployment. Secondly, the efficient multi scale attention (EMA) mechanism was embedded into the critical feature extraction module C3k2, constructing a C3k2_EMA module. This module enhanced dynamic perception of local key features and reconstructed cross-scale contextual dependencies broken by occlusion through its parallel multi-branch structure, effectively suppressing background and occlusion noise. Finally, the DynamicHead detection head was introduced. Leveraging its scale-aware and spatial-aware mechanisms, it achieved a dynamic fusion of multi-level features and adaptive weight adjustment, further improving the model's decision-making robustness in complex scenes. To comprehensively evaluate model performance, a carrot seedling dataset covering various field scenarios was independently constructed. Through offline data augmentation, the original 1 274 images were expanded to 4 796, which were then split into training, validation, and test sets in an 8:1:1 ratio. Meanwhile, to systematically quantify the model's anti-occlusion performance, an occlusion severity assessment criterion based on the overlapping area of bounding boxes was proposed. Targets were categorized into three occlusion levels: mild, moderate, and severe. Based on this, a dedicated "Occlusion Test Subset" was separated from the main test set, providing an objective and reproducible benchmark for evaluating the model's anti-occlusion capability. **[Results and Discussions]** Experimental results on the custom dataset demonstrated that CD-YOLO comprehensively improved detection performance while maintaining its lightweight characteristics. Compared to the baseline model YOLOv11s, CD-YOLO reduced computational load by 6.2 GFLOPs (a 28.8% decrease), decreased model size by 4.8 MB (a 25.0% reduction), improved single-image inference speed by 4.7 ms, reaching 9.6 ms. Concurrently, precision, recall, and mean average precision (mAP_{0.5}) increased by 3.0, 1.5, and 2.4 percentage points, respectively, ultimately reaching 81.2%, 76.4%, and 84.0%. In comparisons with other lightweight backbone networks like MobileNetv3 and ShuffleNetv2, CD-YOLO consistently outperformed them on the accuracy-speed comprehensive metric, validating the effectiveness of its improvement strategies. In occlusion performance tests, the missed detection rate of CD-YOLO on the occlusion test subset was 13.4%, a 5.7 percentage points decrease compared to YOLOv11s. Its mAP_{0.5} on the occlusion subset reached 80.6%, a 5.1 percentage points improvement over the baseline, whereas the improvement on the regular subset was 1.8 percentage points, proving the model's enhanced efficacy in occlusion scenarios. After deploying the model on an NVIDIA Jetson Orin NX edge device and accelerating it with TensorRT, the inference frame rate increased to 32.5 f/s. On random test images, CD-YOLO achieved missed detection and false detection rates of 5.1% and 2.7%, respectively, representing decreases of 7.7% and 2.6% compared to YOLOv11s, demonstrating promising practical application potential. Ablation studies and feature map visualizations further indicated that DWConv, C3k2_EMA, and DynamicHead formed a synergistic optimization loop: DWConv achieved computational compression, freeing up computational budget for subsequent modules; C3k2_EMA enhanced local perception and contextual reconstruction of occluded targets during the feature extraction stage; and DynamicHead performed dynamic fusion of multi-scale features at the decision-making end. Together, they ensured high-precision detection of incomplete targets under limited computational resources. **[Conclusions]** Through the synergistic design of "lightweighting, feature enhancement, and dynamic fusion", the CD-YOLO model achieved an excellent balance between computational efficiency, detection accuracy, and anti-occlusion capability. The model not only significantly reduced reliance on the computational power of edge devices but also effectively improved robustness and adaptability in complex field environments through structured attention and dynamic fusion mechanisms.

Keywords: carrot seedlings; occlusion; object detection; YOLOv11s; lightweight

Foundation items: Beijing Academy of Agriculture and Forestry Sciences Innovation Capacity Building Project (KJCX20230409); National Natural Science Foundation of China (32201647); Reform and Development Project (GGFZ20250205)

First author: LIU Haoran, E-mail: lhr13646345271@outlook.com

***Corresponding author:** ZHAI Changyuan, E-mail: zhaicy@nercita.org.cn