

基于 YOLOv11-VCL 模型的田间辣椒苗 实时检测方法

胡家睿^{1,2} 陆傲鹏^{1,2} 陈蒙^{1,2} 石航³ 韩长杰^{1,2*}

(1. 新疆农业大学机电工程学院, 乌鲁木齐 830052; 2. 新疆智能农机装备工程技术研究中心, 乌鲁木齐 830052;
3. 黑龙江八一农垦大学工程学院, 大庆 163319)

摘要: 为精准监测辣椒苗移栽后的存活状态, 以提升移栽成功率并评估其生长状态, 以 YOLOv11n 为基础模型构建并优化辣椒苗快速检测模型, 针对田间辣椒苗检测中常见的复杂背景与光照条件干扰问题, 在骨干网络的跨阶段局部 v3+2×2 卷积核 (cross-stage partial v3 with kernel 2, C3k2) 模块中引入坐标注意力 (coordinate attention, CA) 机制; 同时加入视觉变换器 (vision transformer, ViT) 注意力机制模块以优化多尺度特征融合的识别效率与精准度, 采用高效交并比 (efficient intersection over union, EIou) 损失函数替代传统定位损失模块, 结合多任务权重自适应策略和数据增强手段优化训练过程, 以提高定位精度。结果表明: 改进的 YOLOv11-VCL 模型对于不同光照、复杂背景条件下的辣椒苗均有较好的检测效果, 精确率、召回率和平均精度均值分别为 95.8%、90.6% 和 95.1%, 明显优于 YOLOv5n、YOLOv8n、YOLOv9n、YOLOv10n 和 YOLOv11n 模型。YOLOv11-VCL 模型的参数量为 3.3 M, 计算量为 7.5 GFLOP, 推理速度为每秒 146 帧, 能够满足精准施药边缘计算设备的运行要求, 有效助力辣椒的自动化田间管理。

关键词: 辣椒苗; 实时检测; 注意力机制; YOLOv11-VCL; 卷积神经网络

A real-time detection approach for pepper seedlings in farmland utilizing YOLOv11-VCL model

Hu Jiarui^{1,2} Lu Aopeng^{1,2} Chen Meng^{1,2} Shi Hang³ Han Changjie^{1,2*}

(1. College of Mechanical and Electrical Engineering, Xinjiang Agricultural University, Urumqi 830052, Xinjiang Uygur Autonomous Region, China; 2. Xinjiang Intelligent Agricultural Machinery Equipment Engineering and Technology Research Center, Urumqi 830052, Xinjiang Uygur Autonomous Region, China; 3. College of Engineering, Heilongjiang Bayi Agricultural University, Daqing 163319, Heilongjiang Province, China)

Abstract: Accurate monitoring of the survival status of pepper seedlings post-transplanting is the key to improving the transplanting success rate and evaluating growth status, and also a prerequisite for realizing the automatic field management of peppers. To this end, this study proposes a fast detection method for the survival status of pepper seedlings named YOLOv11-VCL, which is based on the improved YOLOv11n. Firstly, aiming at the common interference problems of complex background and illumination conditions in field pepper seedling detection, the coordinate attention (CA) mechanism is introduced into the cross-stage partial v3 with kernel 2 (C3k2) module of the backbone network. Meanwhile, to further optimize the recognition efficiency and accuracy of multi-scale feature fusion, the vision trans-

基金项目: 新疆维吾尔自治区重点研发项目(2023 B02021-2), 新疆蔬菜产业技术体系专项(XJARS-07-06), 新疆维吾尔自治区“天山英才”培养计划项目(2024 TSYCCX0035), 新疆维吾尔自治区研究生科研创新项目(XJ2025G118)

* 通信作者 (Author for correspondence), E-mail: hcj_627@163.com

收稿日期: 2025-08-30

former (ViT) attention mechanism module is added to enable the model to detect pepper seedlings more accurately in complex field scenarios. In addition, efficient intersection over union (EIoU) is adopted to replace the traditional localization loss, and the multi-task weight adaptive strategy and data augmentation methods are integrated to optimize the model training so as to improve the localization accuracy. The results showed that the YOLOv11-VCL network model achieved excellent detection performance for pepper seedlings under different illumination and complex background conditions, with the precision, recall and mean average precision (mAP) reaching 95.8%, 90.6% and 95.1%, respectively. The detection results were significantly superior to those of the YOLOv5n, YOLOv8n, YOLOv9n, YOLOv10n and YOLOv11n models, and the missed detection and false detection rates of the improved model were significantly reduced. The improved model has a parameter quantity of 3.3 M, GFLOPs of 7.5 and a frame rate of 146 frames per second (FPS), which can meet the operation requirements of edge computing devices for precision pesticide application, and effectively facilitate the realization of automatic field management.

Key words: pepper seedling; real-time detection; attention mechanism; YOLOv11-VCL; convolutional neural network

中国辣椒年种植面积超过200万 hm^2 ,是重要的经济作物之一。在辣椒苗移栽后,采用高效精准的目标检测技术以实现自动化检测,对后续的移栽成功率和生长状态监测都具有重要意义(戈刚等,2025)。目前,辣椒苗的移栽情况检测仍以人工巡检为主。然而,在实际农业应用中,辣椒移栽环境的多样性会显著增加目标检测难度;且传统的人工特征检测方法受限于光照变化、田间有遮挡物以及植株捕捉不完整等复杂情景;这均会影响对辣椒苗的识别,难以满足田间动态监测需求。因此,开发新的高效检测方法对于及时识别和有效管理移栽过程中辣椒苗的存活情况至关重要。

当前,基于深度学习技术开展的目标检测方法取得了很大进展。YOLO(You Only Look Once)是当前在农业领域应用范围较广的实时目标检测框架(曹亮等,2025;魏冰冰等,2025;周晓彦等,2025),其核心理念在于通过单次前向传播完成对象的分类和定位任务(周秀珊等,2024)。戈刚等(2025)基于YOLOv8n模型进行了改进,以Web应用框架特征(Web application framework feature, WAFF)融合策略优化特征权重,提升特征图表达力,但对模型轻量化关注不足。司超国等(2025)基于YOLOv10s模型提出Chilli-YOLO算法,用GhostConv优化了骨干网络以减少模型参数,并增加二阶通道注意力(second-order channel attention, SOCA)模块来增强特征捕捉,但在复杂的农田情景下该模型的适应性仍有待提升。赵鹏等(2025)引入跨阶段局部 $v_3+2\times 2$ 卷积核(cross-stage partial v_3 with kernel 2, C3k2)_AddBlock_CGLU模块提升了YOLOv11模型的水果

成熟度检测精度,却未充分解决模型低延迟问题。这些研究虽提高了模型的检测精度和适应性,但在模型的轻量化、低延迟方面仍有优化空间(邵延华等,2022)。Liang et al.(2023)提出YOLOv5-sbic算法并进行荔枝晚秋幼芽检测,检测效率达79.6%,但准确性有待提高,且对多尺度特征融合及复杂场景适应性的考虑不足。Khanam & Hussain(2024)深入研究了YOLOv11模型,强调其在农作物检测等领域的应用前景,体现出多技术融合趋势,但未针对性解决该模型在小目标检测及定位精度方面的短板。

鉴于人工巡检效率低、识别准确性不足的问题,针对辣椒田的精准检测模型研究正在逐步开展,但难以充分满足移栽后辣椒苗实时监测的实际需求,存在对复杂田间场景适应性弱、轻量化与高精度难以平衡、小目标检测时定位精度不足等问题。本研究以YOLOv11n为基础模型,通过引入视觉变换器(vision transformer, ViT)模块、构建C3k2-坐标注意力(coordinate attention, CA)模块、用高效交并比(efficient intersection over union, EIoU)损失函数替代传统定位损失模块及多任务权重自适应策略的改进,以期提升模型在田间复杂场景下的检测精度,优化模型轻量化与推理速度,以期为辣椒苗自动化田间管理提供技术支撑。

1 材料与方法

1.1 材料

供试图像数据:本试验的数据采集自新疆维吾尔自治区焉耆回族自治县,于2024年4—6月进行采集,采集设备为iPhone 13。图像分辨率为2 340像

素×1 080像素,以jpg格式保存。所有图片均在距辣椒苗1.0~1.5 m处拍摄,包括自上而下、侧面、正面等不同角度,数据集图像包含多种光照和背景条件下的辣椒苗,以确保数据集的多样性。经筛选后共保留了3 192幅图像,其中包含不同光照强度、叶片重叠、非目标物遮挡等复杂环境的图像。

模型运行平台:本研究模型使用PyTorch框架实现,操作系统为Windows 11, GPU为RTX 3090 (24 Gb), Python版本为3.8.18, CUDA版本为11.8, PyTorch版本为2.0.1, torch版本为0.15.2。模型训练参数设置:Epochs为150, Batch为32, Optimizer为ADAM, Imgsz为640, Workers为4, Close_mosaic为10。

1.2 方法

1.2.1 样本预处理和图像增强处理

利用在线矩形区域标注工具LabelImg对采集到的图像进行人工标注,即针对图像中辣椒苗的目标区域,采用矩形框标注方式进行精准框选,统一设置标签名称为“chilli seedling”,确保标注规则一致。标注完成后,通过人工复核与交叉校验剔除错误标注、修正边界框偏差,最终形成高质量标注数据集。同时对图像进行清洗、预处理操作,包括亮度调整、图像平移、增加高斯噪声以及旋转等数据增强手段,以便有效减少过拟合现象,提高模型的泛化能力,以适应不同任务和模型的具体要求。

1.2.2 YOLOv11网络体系的构建及改进途径

YOLOv11n的架构由输入层、骨干网络、连接层以及头部网络构成(魏迎澳等,2025;熊干等,2025)。相较于YOLOv8模型,其将卷积特征融合(convolutional feature fusion, CF2)模块改成C3k2模块,提高了目标定位的准确性,并加强了骨干网络中的多分支信息融合(李成轩等,2025);在分类检测头中增加两个DWConv,极大减少了参数量和计算量。在YOLOv11的5个版本YOLOv11n、YOLOv11s、YOLOv11m、YOLOv11l和YOLOv11x中,本研究采用YOLOv11n模型,其具有体积小、精度高、更适应农田中复杂环境的优势(刘侦龙等,2025)。

在移栽期辣椒苗检测的背景下,对光照变化、植株捕捉不完整以及遮挡物遮挡叶片等挑战,本研究拟对YOLOv11n模型进行一系列优化,其中在骨干网络加入ViT模块,将Transformer架构用于视觉任务,通过图像分块编码,提升视觉任务的性能,能更好地捕捉图像的全局特征,理解图像整体语义信息,有利于把握辣椒苗在场景中的布局、存活率、倒伏等信息。此外对通道特征聚焦精准性进行改进,即在

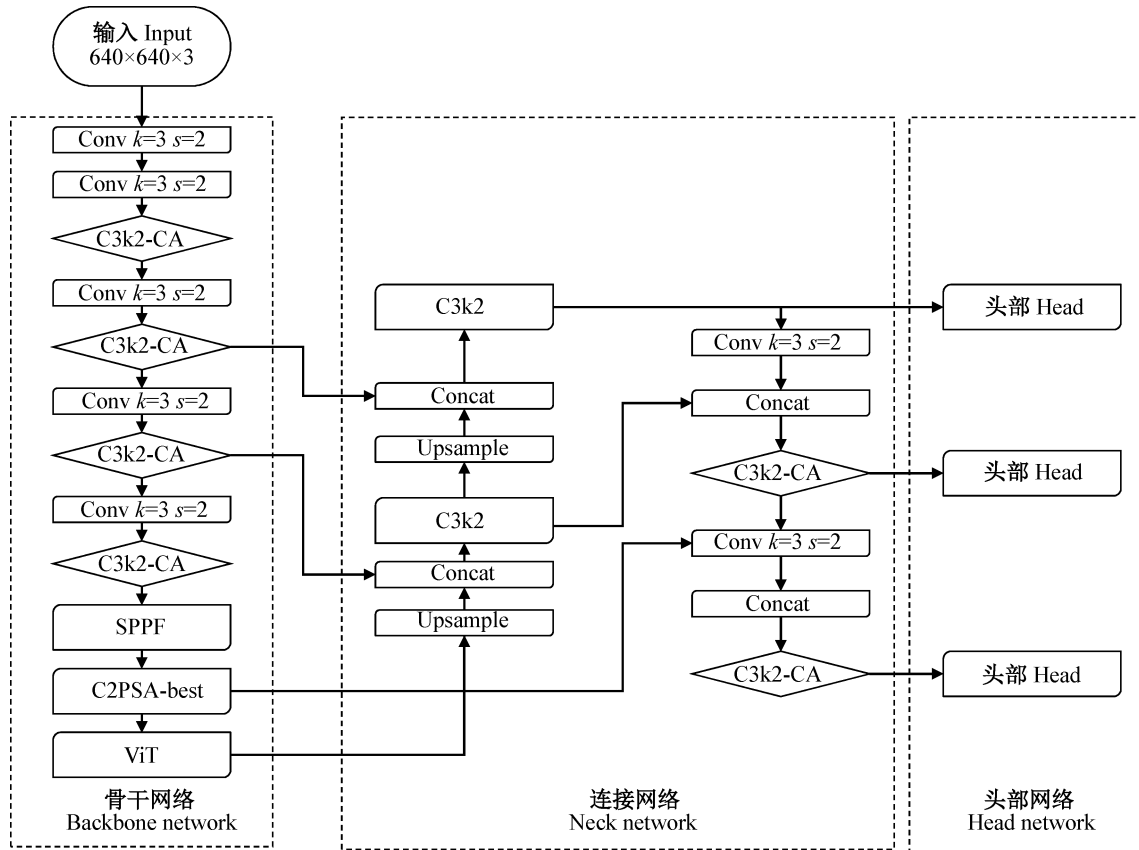
骨干网络及连接网络的C3k2模块引入CA注意力机制,形成C3k2-CA模块(图1),可聚焦关键通道信息,突出与辣椒苗相关的特征,抑制不相关背景的干扰,如土块、地膜等的遮挡,以提升特征提取针对性与有效性。在损失函数优化方面,采用EIoU损失函数替代传统定位损失模块,其综合考量边界框重叠面积、中心点距离及长宽比差异,可精准适配因植株密集、形态不同导致的定位偏差问题。改进后的模型将在辣椒苗检测任务中能更精准定位,对于密集种植的辣椒苗可区分不同个体,避免漏检、误检。

1.2.3 YOLOv11n模型引入ViT模块的改进

在YOLOv11n架构下的辣椒苗识别任务中,当田间辣椒苗出现部分遮挡、形态差异较大或移栽较密集时,模型容易漏检目标或误判边界框位置,导致检测精度不足,在复杂场景下泛化能力较弱(赵静等,2025)。利用YOLOv11n模型进行特征提取时,主要依赖多种卷积操作捕捉局部特征,对于不同光照、密集排列且存在叶片被遮挡的辣椒苗,难以有效获取全局语义信息与长距离依赖关系。为解决该问题,本研究在骨干网络引入ViT模块(图2),该模块是基于Transformer的自注意力机制,可打破卷积操作的局部性限制,直接对图像全局信息进行建模。在辣椒苗识别任务中,ViT模块通过多头自注意力机制,能同时关注不同区域的辣椒苗特征,捕捉叶片、茎秆等部件的长距离关联,精准识别辣椒苗整体形态与空间位置。此外,ViT模板在大规模图像数据集上预训练形成的泛化能力,迁移至辣椒苗检测任务时,可有效提升模型对新环境的适应能力。

1.2.4 YOLOv11n模型引入C3k2-CA模块的改进

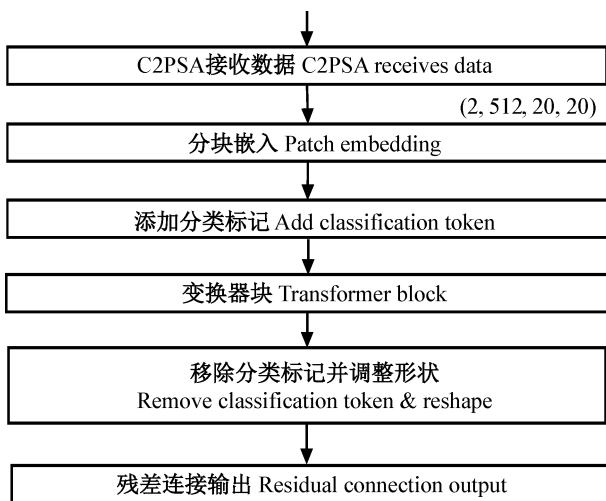
传统C3k2模块作为基于跨阶段局部(cross-stage partial, CSP)结构的核心卷积块,虽通过残差连接增强了特征传播效率,但在空间位置信息提取上存在明显短板(张小艾等,2025)。尤其在田间复杂场景下,面对生长密集、形态各异的辣椒苗,或存在叶片遮挡、背景干扰时,经常因关键位置特征捕捉不足导致小目标辣椒苗漏检、误判或识别不出遮挡、倒伏的辣椒苗。为此,本研究将CA机制嵌入C3k2构建C3k2-CA模块。具体添加于C3k2模块残差分支的特征融合阶段,在3×3卷积与1×1卷积融合前,插入CA模块进行空间-通道维度的注意力加权;CA机制通过水平与垂直方向的全局池化,捕获特征图行列位置依赖关系,经卷积与激活函数处理后生成注意力权重,与原特征图相乘,强化辣椒苗关键区域的特征表达(图3)。



Conv: 卷积层; k : 卷积核尺寸; s : 步幅; C3k2: 跨阶段局部 $v3+2 \times 2$ 卷积核模块; CA: 坐标注意力; SPPF: 快速空间金字塔池化模块; ViT: 视觉变换器模块; Concat: 特征拼接; Upsample: 上采样。Conv: Convolutional layer; k : kernel size; s : stride; C3k2: cross-stage partial $v3$ with kernel 2 module; CA: coordinate attention; SPPF: spatial pyramid pooling-fast module; ViT: vision transformer module; Concat: concatenation.

图1 YOLOv11模型改进网络结构图

Fig. 1 Schematic diagram of the modified YOLOv11 network



C2PSA: 基于C2轻量化残差结构的金字塔分割注意力。C2PSA: C2-based pyramid split attention.

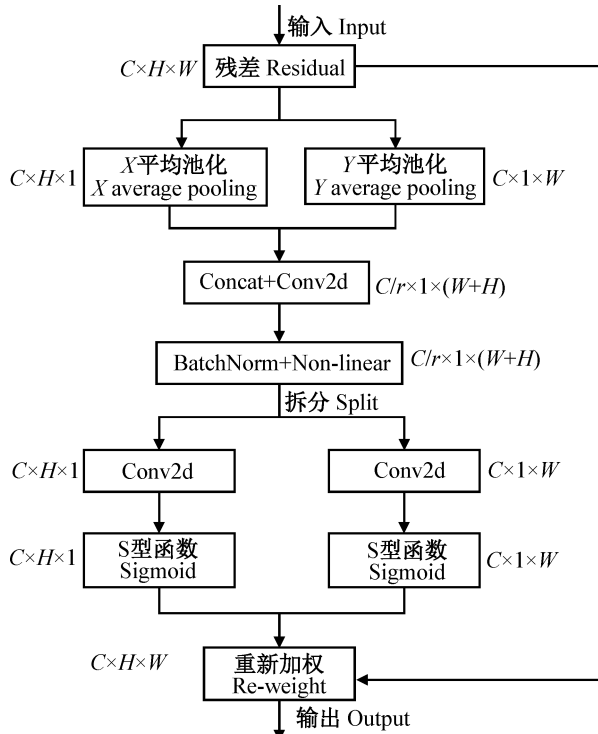
图2 YOLOv11n模型引入ViT模块的改进

Fig. 2 The improvement of the YOLOv11n model by introducing a ViT module

1.2.5 YOLOv11n模型损失函数的改进

在YOLOv11n现有框架下,基于交并比(intersection over union, IoU)损失函数的边框回归方法多聚焦于添加新损失项以加速收敛,却忽视了IoU损失函数自身存在的局限性。尽管理论上IoU损失函数能有效描述辣椒苗边框回归状态,但在实际应用中,面对不同的检测器配置以及复杂多变的检测场景,如辣椒苗处于不同光照、移栽后密度各异、植株远近不一等情况时,IoU损失函数无法自适应调整,泛化性能不足,难以精准适配各类复杂情形。为此,将YOLOv11n中基于C2轻量化残差结构的金字塔分割注意力(C2-based pyramid split attention, C2PSA)模块的IoU损失函数替换为高效IoU (efficient IoU, EIoU)损失函数,其优势在于综合性的考量:不仅继承了传统IoU对边界框重叠面积的度量,确保准确覆盖辣椒苗主体部分,还融入距离IoU (distance IoU, DIoU)对边界框中心点距离的考

量,使预测框在辣椒苗密集或存在遮挡时能更快定位到真实框中心;尤为关键的是EIoU引入长宽比差异,从而精准匹配不同生长阶段下辣椒苗或矮壮或细长的形态特征,避免因长宽比例失配导致的定位偏差,进而显著提升检测的精准度与模型的泛化能力。



C: 特征图通道数; H: 特征图高度; W: 特征图宽度; r: 通道压缩比; Concat: 特征拼接; Conv2d: 二维卷积层。C: No. of channels; H: height of feature map; W: width of feature map; r: reduction ratio; Concat: concatenation; Conv2d: 2D convolutional layer.

图3 C3k2-CA 模块的结构

Fig. 3 Structure of the C3k2-CA module

1.2.6 YOLOv11n 模型多任务权重自适应的改进

为了进一步提升 YOLOv11n 模型的性能,应用了多任务权重自适应的改进策略。该策略能随训练阶段动态调整多任务权重,从而平衡定位损失和分类损失在不同训练时期的重要性。在前 30% 的训练轮次中,辣椒苗的边界框回归准确性是关键。因为只有准确定位辣椒苗的位置,后续分类任务才有可靠的基础。所以在这个阶段应加大定位损失的权重,可以让模型更加关注边界框的精确回归,使得预测的边界框能紧密贴合辣椒苗的实际位置,减少定位误差。而在训练后期,当模型已经具备了较好的定位能力后,分类的准确性就成为提升整体性能的

关键因素。此阶段应提升分类损失的权重,促使模型将更多的精力放在优化类别判别上,从而更准确地识别出不同生长阶段、不同形态的辣椒苗。

$$\alpha(t) = \begin{cases} \alpha_{loc_init} \times \left(1 - \frac{t}{T_{0.3}}\right) + \alpha_{loc_final} \times \frac{t}{T_{0.3}}, & t \leq T_{0.3} \\ \alpha_{cls_init} \times \left(1 - \frac{t - T_{0.3}}{T - T_{0.3}}\right) + \alpha_{cls_final} \times \frac{t - T_{0.3}}{T - T_{0.3}}, & t > T_{0.3} \end{cases}$$

式中, t 表示当前的训练轮次, T 是总训练轮次, $T_{0.3} = 0.3T$ 表示前 30% 训练轮次的界限, α_{loc_init} 和 α_{loc_final} 分别是定位损失在训练前期开始和结束时的权重, α_{cls_init} 和 α_{cls_final} 分别是分类损失在训练后期开始和结束时的权重。

1.2.7 不同模型的检测性能对比

本研究选取 YOLOv5n、YOLOv8n、YOLOv9n、YOLOv10n 和 YOLOv11n 作为基准模型,与本研究改进的 YOLOv11n 模型进行对比试验。所有模型均在相同的试验条件下,使用本研究构建的辣椒数据集进行训练与测试,使用精确率 P 、召回率 R 、平均精度均值(mean average precision, mAP)、 F_1 权重、参数量、计算量及每秒帧数(frames per second, FPS)共 7 项核心评估指标,全方位量化模型的综合性能表现。 $P = T_p / (T_p + F_p) \times 100\%$, $R = T_p / (T_p + F_n) \times 100\%$, 式中, T_p 指样本属于正类并且模型预测结果也为正类的样本数量; F_p 指样本属于负类,但模型却预测为正类的样本数量; F_n 指样本实际为正类,但模型却预测为负类的样本数量。 $mAP = (\sum AP) / n$, AP 表示针对某一类标签所计算出的平均精确率, $AP = \int_0^1 P(R) dR$; n 代表分类任务中类别的总数。本研究中 mAP 仅涉及 mAP@0.5 (IoU 阈值设定为 0.5) 这一指标。 F_1 权重是兼顾精准率与召回率的综合评价指标,通过对二者的调和平均计算平衡权重占比,解决单一指标评价模型性能的片面性问题。在模型复杂度方面,参数量即模型训练时所用参数的总数,是评估模型空间复杂度和规模的关键指标。计算量是指在前向传播过程中,模型所需的浮点运算量,主要用于衡量计算资源的消耗量,通常以每秒十亿次浮点运算(giga floating-point operation, GFLOP)为单位表示。FPS(frames per second)则是衡量模型推理速度的指标,反映单位时间每秒内能够处理的图像帧数。

1.2.8 改进 YOLOv11n 模型的消融试验

为系统评估 ViT 模块、C3k2-CA 机制、EIoU 损失函数和多任务权重自适应策略对田间辣椒检测性

能的贡献,本研究在采集并构建的辣椒数据集上开展了消融试验。即以原始YOLOv11n模型为基线,依次叠加上述改进模块,通过控制变量法逐步验证各模块对模型性能的影响。模型性能评价方法同1.2.7。

1.2.9 改进YOLOv11n模型的实际检测效果评价

将本研究构建的辣椒图像数据集按训练集:验证集:测试集=8:1:1的比例随机划分,其中训练集2553幅图像,验证集319幅图像,测试集含320幅图像。为验证YOLOv11n模型的改进效果,本研究选取320幅不参与训练的测试集图像,通过对比其与原始YOLOv11n模型的置信度,评判最终模型性能。为进一步证明本研究所改进的YOLOv11模型的优势,对其训练损失曲线和验证损失曲线进行分析,验证改进模型在田间辣椒苗识别任务中的实用性与优越性。

2 结果与分析

2.1 不同模型的检测性能对比结果

本研究以YOLOv11模型为基础模型,成功通过引入ViT模块、C3k2-CA机制、EIoU损失函数和多任务权重自适应策略进行了改进,改进后的模型命

名为YOLOv11-VCL。与YOLOv5n、YOLOv8n、YOLOv9n、YOLOv10n和YOLOv11n模型相比,YOLOv11-VCL模型在目标检测的关键指标上展现出明显优势,其中,精确率提升了2.3~9.4百分点,召回率提升了0.9~9.2百分点, F_1 权重提升了0.8~7.4百分点,mAP@0.5增加了0.9~4.0百分点,均明显高于上述5个传统模型(表1),充分体现了YOLOv11-VCL模型在田间辣椒苗检测任务中降低漏检与误检方面的优越性。在计算资源方面,YOLOv11-VCL模型的参数量为3.3M,计算量为7.5GFLOP,推理速度为146FPS,使得训练时辣椒苗候选框的生成速度大幅加快,在模型性能、鲁棒性和泛化能力上表现优异。

2.2 ViT模块对YOLOv11-VCL模型性能的影响

原始YOLOv11n模型在加入ViT模块后,小目标检测的mAP@0.5较原始模型提升2.5百分点,达到92.9%;同时,在密集种植场景下,ViT模块凭借全局特征提取优势与其他模块形成协同作用,进一步强化目标区分度,使模型的精确率提升了约0.7百分点(表2)。此外,ViT模块通过优化特征提取效率还提升了模型的推理速度,使模型能更快速输出辣椒苗检测结果,提升了实际应用效率。

表1 YOLOv11-VCL模型与其他相关YOLOv模型的性能对比结果

Table 1 Performance comparison results of YOLOv11-VCL model and other related YOLOv models

模型 Model	精确率 Precision/%	召回率 Recall/%	mAP@0.5/%	F_1 权重 F_1 score/%	计算量 Computational complexity/GFLOP	参数量 No. of parameters/M	推理速度 Reasoning speed/FPS
YOLOv5n	86.4	81.4	91.1	89.5	15.4	3.7	139
YOLOv8n	92.3	89.3	93.5	95.9	8.6	3.3	146
YOLOv9n	92.1	88.4	93.1	95.8	11.4	3.5	142
YOLOv10n	93.2	89.5	93.9	93.3	9.6	3.8	137
YOLOv11n	93.5	89.7	94.2	96.1	8.2	3.5	141
YOLOv11-VCL	95.8	90.6	95.1	96.9	7.5	3.3	146

GFLOP:每秒十亿次浮点运算;FPS:每秒帧数。GFLOP: Giga floating-point operation; FPS: frames per second.

2.3 C3k2-CA模块对YOLOv11-VCL模型性能的影响

C3k2-CA模块的引入使模型召回率从基线的85.4%提升至86.2%,提升幅度为0.8百分点,能有效召回更多田间真实存在的辣椒苗目标;在密集种植的复杂场景下,C3k2-CA模块与ViT模块协同作用,可使模型的精确度提升约1.3百分点(表2),解决了因植株密集遮挡导致的目标漏检问题。C3k2-CA模块的融合还优化了模型的训练效率,原始YOLOv11n模型需经过150轮训练才能完成收敛,

引入该模块后仅需约120轮即可达到收敛状态,大幅节省了模型训练的时间成本与计算资源消耗,加速了模型的迭代优化。

2.4 EIoU对YOLOv11-VCL模型性能的影响

本研究通过引入EIoU损失函数优化定位损失,结合多任务权重自适应策略,使模型在辣椒苗检测任务中训练效果明显提升,在引入ViT模块将模型的mAP@0.5提升至92.9%的基础上,经EIoU损失函数与多任务权重自适应策略的协同优化,模型的

mAP@0.5 进一步提升至 93.5%，较原始模型提升了 3.1 百分点(表 2)。

2.5 YOLOv11-VCL 模型的消融试验结果

原始 YOLOv11n 模型仅通过嵌入 ViT 模块,其小目标检测的 mAP@0.5 便较基线模型提升 2.5 百分点,达到 92.9%;在此基础上进一步引入 EIoU 定位损失优化策略后, mAP@0.5 进一步提升至 93.5%;

最终构建的 YOLOv11-VCL 模型在复杂田间场景的测试验证中, mAP@0.5 相较于原始模型提升 4.7 百分点,达到 95.1%,且检测精度大幅提升,精确率达到 95.8%,召回率达到 90.6%(表 2)。表明该模型不仅可精准识别辣椒苗,有效降低误检情况,还能对幼苗、小型植株苗及被遮挡辣椒苗实现可靠识别,减少目标漏检现象。

表 2 改进 YOLOv11-VCL 模型的消融试验性能比较

Table 2 Performance comparison of ablation experiments on the improved YOLOv11-VCL model

模型 Model	精确度/% Precision/%	召回率/% Recall/%	mAP@0.5/%	计算量 Computational complexity/GFLOP	参数量 No. of parameters/M
YOLOv11n	92.5	85.4	90.4	7.1	3.6
YOLOv11n+C3k2-CA	93.1	86.2	91.6	7.4	3.7
YOLOv11n+ViT	93.2	86.6	92.9	7.6	3.7
YOLOv11n+EIoU	94.2	87.4	94.1	8.4	3.9
YOLOv11n+C3k2-CA+ViT	93.8	86.9	92.4	7.7	3.8
YOLOv11n+EIoU+C3k2-CA	92.8	85.5	91.5	7.2	3.6
YOLOv11n+ViT+EIoU	94.5	88.9	93.5	8.1	3.9
YOLOv11-VCL	95.8	90.6	95.1	7.5	3.3

C3k2: 跨阶段局部 v3+2×2 卷积核模块; ViT: 视觉变换器模块; CA: 坐标注意力机制; EIoU: 高效交并比损失函数; GFLOP: 每秒十亿次浮点运算; FPS: 每秒帧数。C3k2: Cross-stage partial v3 with kernel 2 module; ViT: vision transformer module; CA: coordinate attention mechanism; EIoU: efficient intersection over union loss function; GFLOP: giga floating-point operation; FPS: frames per second.

2.6 YOLOv11-VCL 模型的实际检测效果

对原始 YOLOv11n 模型和改进的 YOLOv11-VCL 模型在不同场景下的检测结果进行分析,并通过随机抽取具有代表性的图片进行展示(图 4)。YOLOv11n 和 YOLOv11-VCL 模型在近距离无干扰数据集上均能准确识别辣椒苗,但 YOLOv11-VCL 模型检测置信度更高(图 4-A);在高光处理场景中, YOLOv11n 模型出现误检,而 YOLOv11-VCL 模型减少了误检,精准识别出辣椒苗(图 4-B);在暗处理场景中, YOLOv11n 模型发生漏检, YOLOv11-VCL 模型的精确率更高(图 4-C);在边缘残缺图片检测中,目标处于图片边缘时导致植株不全, YOLOv11n 模型因缺乏处理不完整特征的机制导致误检和定位不准,而 YOLOv11-VCL 模型则基于全局语义准确推断残缺辣椒苗特征(图 4-D);在多植株小目标检测场景下, YOLOv11n 模型因小目标混淆出现较多的漏检和误检,而 YOLOv11-VCL 模型提升了检测精度(图 4-E);在遮挡处理场景中, YOLOv11n 模型对被遮挡辣椒苗检测困难,而 YOLOv11-VCL 模型增强了对遮挡场景下辣椒苗的检测能力(图 4-F)。这表明 YOLOv11-VCL 模型在多种复杂场景下的检

测性能均明显优于原始 YOLOv11n 模型。

为进一步证明本研究所改进的 YOLOv11-VCL 模型的优势,对其训练损失曲线和验证损失曲线进行分析,模型训练过程中的边界框回归损失以及分类损失均随迭代次数增加呈现快速下降后趋于平稳的趋势,表明模型在训练集上对辣椒苗的边界框定位与分类判别能力逐步收敛(图 5-A);而验证损失曲线与训练损失曲线的下降趋势一致且数值接近(图 5-B),说明模型未出现过拟合,对辣椒苗样本具备良好的泛化能力。同时也分析了模型对验证集的 mAP 指标, mAP@0.5 最终接近 95%(图 6),体现了常规判定标准下模型对辣椒苗检测的精度,反映出模型在不同 IoU 严苛程度下的鲁棒性均良好。

3 讨论

本研究提出的 YOLOv11-VCL 模型,针对移栽期辣椒苗检测中光照变化、植株捕捉不完整及遮挡叶片等核心难题,构建了全局特征强化、局部特征聚焦与损失函数优化的三维改进体系。通过在 YOLOv11n 架构中引入 ViT 模块、改进 C3k2-CA 模块及优化损失函数策略,实现了检测精度与实际适

配性的双重提升。这一改进思路与现有研究共识一致,即通过针对性的模块优化与机制创新,可有效突

破传统农业目标检测模型在复杂场景下的性能瓶颈 (Xue & Wang, 2025; Shi et al., 2025)。



图4 YOLOv11n与YOLOv11-VCL模型在不同场景下对辣椒幼苗的检测

Fig. 4 Detection of pepper seedlings by YOLOv11n and YOLOv11-VCL models under different scenarios

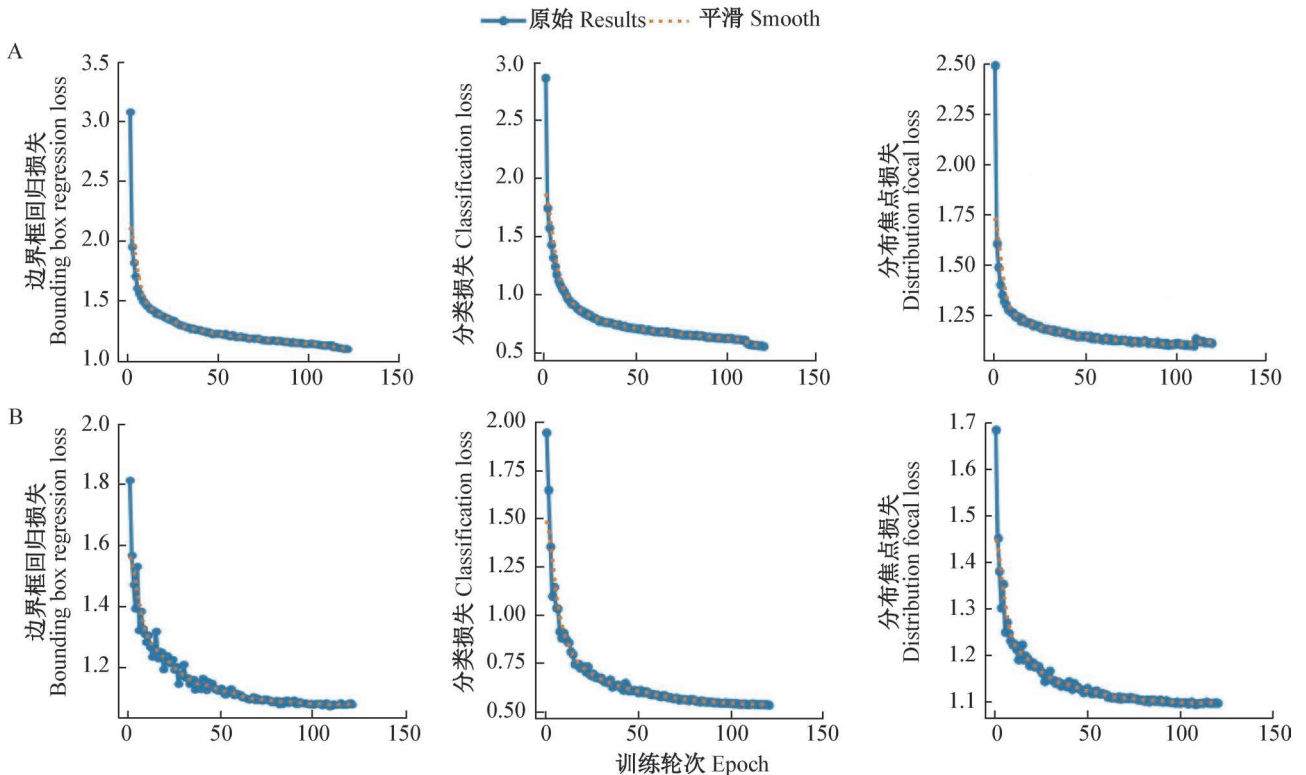


图5 YOLOv11-VCL模型实际效果验证中的训练损失曲线(A)与验证损失曲线(B)

Fig. 5 Training (A) and validation (B) loss curves of YOLOv11-VCL model in practical effect verification

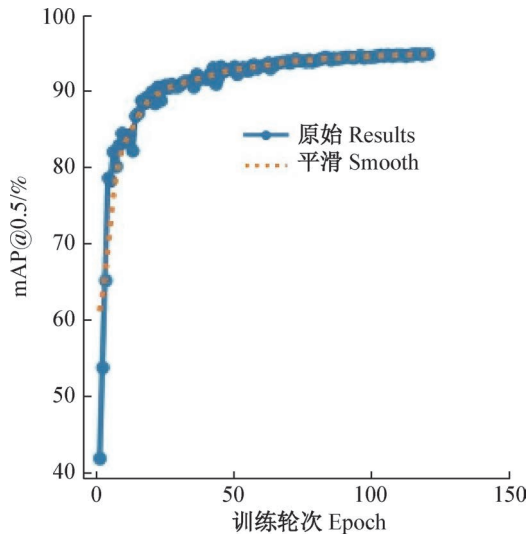


图6 YOLOv11-VCL模型实际效果检测中验证集的mAP曲线

Fig. 6 mAP curve of the YOLOv11-VCL model on the validation set during practical effect detection

与蒋平等(2025)现有农业检测研究相比,本研究在数据层面专门构建了移栽期辣椒苗数据集,覆盖多种光照、种植密度及遮挡场景,并通过亮度调整、图像平移、高斯噪声添加及旋转等数据增强手段扩充样本多样性,相较于通用农业数据集更贴合移栽期检测的实际需求。这与Shi et al.(2025)强调的模型需适配实际农业生产场景的核心诉求高度契合。本研究改进的YOLOv11-VCL模型针对YOLOv11n现有架构的不足,引入ViT模块以强化全局特征捕捉,解决了局部卷积难以建模复杂场景中辣椒苗长距离依赖关系的难题(Xue & Wang, 2025);改进C3k2模块并融入CA机制形成C3k2-CA模块,精准聚焦关键通道信息,提升特征提取的针对性并抑制背景干扰;创新性地采用EIou替代传统IoU损失模块,综合考量边界框多维度差异,同时结合多任务权重自适应策略,依训练阶段动态调整定位损失与分类损失权重,有效提升了模型对复杂场景的适应能力,实现检测精度与泛化性能的双重优化。这些改进使YOLOv11-VCL模型在保证高精度的同时,具备更优的轻量化与低延迟特性,更契合实际农业生产中对辣椒苗检测的高效性与准确性需求(Shi et al., 2025)。

本研究结果表明,YOLOv11-VCL模型在准确率与模型体量上均表现出明显优势:精确率达95.8%,召回率达90.6%,mAP@0.5为95.1%,计算量为7.5 GFLOP,参数量为3.3 M;与YOLOv11n模型

相比,改进后的模型在复杂场景下的mAP@0.5提升4.7个百分点,精确度提升3.3个百分点,且检测速度快至6 ms,突出了其在检测任务中的准确性与高效性。但受田间环境复杂性限制,当前研究仅局限于无病虫害侵染的辣椒苗。未来工作可侧重于扩充辣椒苗更多生长状态的数据集,以便进一步开展目标识别与实时检测研究。本研究证实了深度学习在农业检测领域的应用潜力,未来将进一步探索多作物多场景的适应性优化,推动计算机视觉技术与智慧农业的深度融合。

参考文献 (References)

- Cao L, Xiao W, Li XL. 2025. Review on YOLO algorithm and its application in crop identification and pest detection. *Journal of Zhongkai University of Agriculture and Engineering*, 38(6): 60-71 (in Chinese) [曹亮, 肖伟, 李湘丽. 2025. YOLO算法及其在农作物识别及病虫害检测应用综述. *仲恺农业工程学院学报*, 38(6): 60-71]
- Ge G, Yang J, Liu Y, Hu YX, Liu HH. 2025. Detecting tomatoes in complex agricultural environments using improved YOLOv8n. *Transactions of the Chinese Society of Agricultural Engineering*, 41(9): 143-153 (in Chinese) [戈刚, 杨珺, 刘毅, 胡逸轩, 刘昊辉. 2025. 改进YOLOv8n模型对复杂农业场景下西红柿目标的检测. *农业工程学报*, 41(9): 143-153]
- Jiang P, Chen M, Han CJ, Liang J, Li GX. 2025. Design and test of plug seedling transplanter seedling replenishment system based on machine vision. *Journal of Agricultural Mechanization Research*, <https://link.cnki.net/urlid/23.1233.S.20251210.1335.006> (in Chinese) [蒋平, 陈蒙, 韩长杰, 聂诗良, 梁佳, 李光新. 2025. 基于机器视觉的穴盘苗移栽机补苗系统设计与试验. *农机化研究*, <https://link.cnki.net/urlid/23.1233.S.20251210.1335.006>]
- Khanam R, Hussain M. 2024. YOLOv11: an overview of the key architectural enhancements. arXiv: 2410.17725
- Li CX, Zhao CJ, Zhang C, Huang WQ, Li JQ, He X, Wang QY. 2025. Multi-surface defect detection method for Dangshan pears based on AIC-YOLOv11n model. *Transactions of the Chinese Society of Agricultural Engineering*, 41(14): 320-328 (in Chinese) [李成轩, 赵春江, 张驰, 黄文倩, 李佳琪, 何鑫, 王庆艳. 2025. 基于AIC-YOLOv11n模型的砀山梨多表面缺陷检测方法. *农业工程学报*, 41(14): 320-328]
- Liang JT, Chen X, Liang CJ, Long T, Tang XY, Shi ZM, Zhou M, Zhao J, Lan YB, Long YB. 2023. A detection approach for late-autumn shoots of litchi based on unmanned aerial vehicle (UAV) remote sensing. *Computers and Electronics in Agriculture*, 204: 107535
- Liu ZL, Wang J, Mai RG. 2025. Detecting underwater sea cucumber in marine ranching using improved YOLOv10s. *Transactions of the Chinese Society of Agricultural Engineering*, 41(10): 186-194 (in Chinese) [刘侦龙, 王骥, 麦仁贵. 2025. 基于改进YOLOv10s的海洋牧场水下海参检测方法. *农业工程学报*, 41(10): 186-194]

- Shao YH, Zhang D, Chu HY, Zhang XQ, Rao YB. 2022. A review of YOLO object detection based on deep learning. *Journal of Electronics & Information Technology*, 44(10): 3697-3708 (in Chinese) [邵延华, 张铎, 楚红雨, 张晓强, 饶云波. 2022. 基于深度学习的YOLO目标检测综述. *电子与信息学报*, 44(10): 3697-3708]
- Shi H, Liu CX, Wu M, Zhang H, Song H, Sun H, Li YF, Hu J. 2025. Real-time detection of Chinese cabbage seedlings in the field based on YOLO11-CGB. *Frontiers in Plant Science*, 16: 1558378
- Si CG, Liu MC, Wu HR, Miao YS, Zhao CJ. 2025. Chilli-YOLO: an intelligent maturity detection algorithm for field-grown chilli based on improved YOLOv10. *Smart Agriculture*, 7(2): 160-171 (in Chinese) [司超国, 刘梦晨, 吴华瑞, 缪祎晟, 赵春江. 2025. Chilli-YOLO: 基于改进YOLOv10的露地辣椒成熟度智能检测算法. *智慧农业*, 7(2): 160-171]
- Wei BB, Zhang Z, Yan K, Zhang GJ, Wang P, Huang QY, Niu CY. 2025. Intelligent monitoring of tephritid fruit fly pests using an improved YOLOv5s model. *Journal of Plant Protection*, 52(5): 1163-1171 (in Chinese) [魏冰冰, 张志, 闫科, 张桂健, 王鹏, 黄求应, 牛长纓. 2025. 基于改进YOLOv5s模型的实蝇类害虫智能监测. *植物保护学报*, 52(5): 1163-1171]
- Wei YA, Xu J, Li YH, Hu HT. 2025. Anomaly detection of substation equipment and operational behavior based on the improved YOLOv11n model. *Journal of Hubei Minzu University (Natural Sciences Edition)*, 43(1): 41-46 (in Chinese) [魏迎澳, 徐建, 李英豪, 胡浩特. 2025. 基于改进YOLOv11n模型的变电站设备及生产行为异常检测. *湖北民族大学学报(自然科学版)*, 43(1): 41-46]
- Xiong G, Chen CF, Zhang S. 2025. QMDF-YOLO11: rice pests detection algorithm in complex scenarios. *Computer Engineering and Applications*, 61(13): 113-123 (in Chinese) [熊干, 陈慈发, 张上. 2025. QMDF-YOLO11: 复杂场景下水稻害虫检测算法. *计算机工程与应用*, 61(13): 113-123]
- Xue RZ, Wang LQ. 2025. Research on lightweight citrus leaf pest and disease detection based on PEW-YOLO. *Processes*, 13(5): 1365
- Zhang XA, Qiao HB, Xi L, Ma CW, Xu X, Hui XH, Tang Q, Zhang H. 2025. Detecting small objects for field crop pests using improved YOLOv8. *Transactions of the Chinese Society of Agricultural Engineering*, 41(13): 225-233 (in Chinese) [张小艾, 乔红波, 席磊, 马超伟, 许鑫, 惠向晖, 唐琦, 张浩. 2025. 基于改进YOLOv8的大田农作物害虫小目标检测方法. *农业工程学报*, 41(13): 225-233]
- Zhao J, Li JQ, Yang L, Cui WH, Han GT, Zhou Q, Lu LQ, Lan YB. 2025. Real-time detecting cotton seedlings under a lightweight film using PConv-CGLU with a heavy parameter detection head. *Transactions of the Chinese Society of Agricultural Engineering*, 41(18): 151-162 (in Chinese) [赵静, 李京谦, 杨蕾, 崔文豪, 韩国涛, 周琦, 鲁力群, 兰玉彬. 2025. 基于PConv-CGLU与重参数检测头的轻量化膜下棉苗实时检测算法. *农业工程学报*, 41(18): 151-162]
- Zhao P, Qiang GL, Lu B, Gao Y, Zhang QX. 2025. Fruit ripeness detection based on improved YOLOv11. *Modern Information Technology*, (8): 34-40 (in Chinese) [赵鹏, 强光磊, 卢波, 高扬, 张仟祥. 2025. 基于改进YOLOv11的水果成熟度检测. *现代信息科技*, (8): 34-40]
- Zhou XS, Wen LT, Jie BF, Zheng HF, Wu QQ, Li KE, Liang JN, Li YJ, Wen JY, Jiang LY. 2024. Real-time detection algorithm of expanded feed image on the water surface based on improved YOLOv11. *Smart Agriculture*, 6(6): 155-167 (in Chinese) [周秀珊, 文露婷, 介百飞, 郑海锋, 吴其琦, 李克讷, 梁军能, 黎一键, 文家燕, 江林源. 2024. 改进YOLOv11的水面膨化饲料颗粒图像实时检测算法. *智慧农业*, 6(6): 155-167]
- Zhou XY, Ju X, Yu QD, Ma H, Xian XQ, Zhang GF. 2025. Dual-dimensional recognition of tomato leafminer *Tuta absoluta* based on a multi-scale fusion model and its application. *Journal of Plant Protection*, 52(4): 972-984 (in Chinese) [周晓彦, 鞠醒, 于启烜, 马浩, 洗晓青, 张桂芬. 2025. 基于多尺度融合模型的双维度番茄潜叶蛾识别方法及应用. *植物保护学报*, 52(4): 972-984]

(责任编辑:李美娟)