doi:10.6041/j.issn.1000-1298.2025.05.041

基于改进 YOLO v8 的复杂温室环境黄瓜果实分割方法

夏　天[1]　谢　纯[1]　李琳一[2,3]　陆声链[4]　钱婷婷[2,5]

(1. 上海第二工业大学计算机与信息工程学院，上海 201209；
2. 上海市农业科学院农业科技信息研究所，上海 201403；
3. 上海数字农业工程技术研究中心，上海 201403；
4. 广西师范大学计算机科学与工程学院，桂林 541004；
5. 农业农村部长三角智慧农业技术重点实验室，上海 201403)

摘要：黄瓜果实的检测与分割对于表型分析和黄瓜生长管理至关重要。然而，在复杂温室环境下，果实往往与茎叶相互遮挡，且果实与背景颜色相似，导致传统方法在复杂环境下难以准确识别果实边界并实现高效分割。为此，提出了一种基于改进 YOLO v8 的黄瓜果实分割方法。该方法引入可变形卷积(Deformable convolution network v4，DCNv4)增强模型空间适应性；同时采用 RepNCSPELAN4 模块串联额外的 C2F 模块，细化特征提取与融合；从而提升了模型在复杂温室环境下对黄瓜果实图像的分割性能。实验结果显示，在玻璃温室和塑料连栋大棚两个实验场景中的多个类别上均有出色表现。其中，在玻璃温室场景中的精确率为 96.3%，召回率为 93.1%，平均精度均值 mAP50 为 96.2%，mAP50－95 为 85.3%；在塑料大棚场景中的精确率为 86.8%，召回率为 81.9%，平均精度均值 mAP50 为 90.0%，mAP50－95 为 77.0%。本研究提出的改进方法在处理边界、多重遮挡和多尺度分割方面具有更强的鲁棒性和泛化性，使模型能适应复杂性不同的多样化种植环境而准确分割黄瓜果实。精确的果实图像分割有助于表型参数的获取，为黄瓜果实的表型分析提供了可靠的技术支持，从而促进农业表型机器人的应用。

关键词：黄瓜；果实遮挡；设施温室；图像分割；YOLO v8；可变形卷积

中图分类号：TP391.4　　**文献标识码**：A　　**文章编号**：1000-1298(2025)05-0433-10　　**OSID**：

Improved YOLO v8 Method for Cucumber Fruit Segmentation in Complex Greenhouse Environments

XIA Tian[1]　XIE Chun[1]　LI Linyi[2,3]　LU Shenglian[4]　QIAN Tingting[2,5]

(1. *School of Computer and Information Engineering, Shanghai Polytechnic University, Shanghai* 201209, *China*
2. *Institute of Agricultural Science and Technology Information, Shanghai Academy of Agricultural Sciences, Shanghai* 201403, *China*
3. *Research Center of Shanghai Digital Agricultural Engineering and Technology, Shanghai* 201403, *China*
4. *School of Computer Science and Engineering, Guangxi Normal University, Guilin* 541004, *China*
5. *Key Laboratory of Smart Agricultural Technology (Yangtze River Delta), Ministry of Agriculture and Rural Affairs, Shanghai* 201403, *China*)

Abstract: The detection and segmentation of cucumber fruits are crucial for phenotypic analysis and the management of cucumber growth. However, in complex greenhouse environments, fruits are often occluded by stems and leaves, and their color may be similar to the background, making it difficult for traditional methods to accurately identify fruit boundaries and achieve efficient segmentation. To address this issue, an improved YOLO v8-based method for cucumber fruit segmentation was proposed. This method incorporated deformable convolution network v4 (DCNv4) to enhance the model's spatial adaptability and utilized the RepNCSPELAN4 module in combination with an additional C2F module to refine feature extraction and fusion, thereby improving the model's segmentation performance for cucumber fruit images in complex greenhouse environments. Experimental results showed outstanding

收稿日期：2024－12－20　修回日期：2025－01－25

基金项目：上海市农业科技创新项目(2023－02－08－00－12－F04621)和国家自然科学基金项目(61762013)

作者简介：夏天(1979—)，男，教授，博士，主要从事图像分析与理解、自然语言处理等研究，E-mail：xiatian@sspu.edu.cn

通信作者：钱婷婷(1983—)，女，副研究员，博士，主要从事设施农业模型研究，E-mail：qiantingting@saas.sh.cn

performance across multiple categories in two experimental settings: a glass greenhouse and a plastic greenhouse. Specifically, in the glass greenhouse scenario, the model achieved a precision of 96.3%, recall of 93.1%, mean average precision (mAP50) of 96.2%, and mAP50 - 95 of 85.3%. In the plastic greenhouse scenario, the precision was 86.8%, recall was 81.9%, mAP50 was 90.0%, and mAP50 - 95 was 77.0%. The proposed method demonstrated stronger robustness and generalization in handling boundary issues, multiple occlusions, and multi-scale segmentation, enabling the model to adapt to diverse and complex cultivation environments and accurately segment cucumber fruits. Accurate fruit image segmentation facilitated the acquisition of phenotypic parameters and provides reliable technical support for further phenotypic analysis of cucumber fruits, thereby promoting the application of agricultural phenotyping robots and the intelligent development of agricultural production.

Key words: cucumber; fruit occlusion; greenhouse; image segmentation; YOLO v8; deformable convolution

0 引言

黄瓜是温室内广泛种植的一种蔬菜作物,具有良好的经济效益[1]。在现代农业生产中,黄瓜果实的检测与分割是实现黄瓜表型参数获取和黄瓜生长智能化管理的技术前提[2],所以,研究开发准确、可靠的分割算法具有重要的研究价值和应用前景。然而,在复杂温室环境下,黄瓜果实往往与茎叶存在不同程度的遮挡,加之果实、茎、叶颜色高度相近[3],不易区分,同时,不同温室材质和环境导致的散射、反光等情况,易使图像噪声增强,这些因素大大增加了黄瓜检测与分割任务的难度。

早期研究主要依赖传统图像处理方法[4-5]和机器学习算法进行果实分割[6]。这些方法仅在黄瓜无遮挡时能起作用,难以应对复杂的环境。

随着计算机视觉的发展,深度学习模型在图像分割领域取得了显著进步。LI 等[7]结合残差模块和门控卷积,并引入空间金字塔结构,将 U-Net 高级语义特征和边缘特征融合,实现果园青苹果的高精度语义分割。然而,语义分割主要针对像素级分类,适用于图像前后景及不同材质区域的分割。对于目标数量多且分散独立的果实分割任务,借助目标定位的分割方法更为适宜[8],这种方法也被后续研究所采纳。

最先被广泛使用的是 Mask R - CNN[9-11],LIU 等[12]在 Mask R - CNN 的基础上,设计逻辑绿色算子来限制锚框的范围,在近色背景中有效分割了黄瓜。LÓPEZ-BARRIOS 等[13]使用 ResNet101 和特征金字塔网络(Feature pyramid networks, FPN)作为 Mask R - CNN 的骨干网络,以增强对象表示和多尺度特征提取能力,使青甜椒在复杂遮挡环境中也能有效分割。尽管这些方法在复杂环境下的果实分割中发挥了重要作用,但它们未能充分考虑果实在生长环境中的不同遮挡情况[14],将所有果实视为一类进行分割,导致果实边缘和遮挡部分的分割效果不佳。

YOLO 系列模型广泛应用于目标检测领域,其后续版本和扩展模型在图像分割领域同样有出色的表现[15-19]。倪纪鹏等[20]在 YOLACT 上引入 Swin - S 注意力机制优化特征提取效果,并采用多尺度特征提取方式克服了模型的局限性,缓解了番茄实例分割中漏检与误检问题。赵文锋等[21]在 YOLOP 骨干网络中置入焦点融合高效卷积模块 FAC,采用 C2F 和 SPPF 模块,优化损失函数,以完成火龙果园中的多任务分割。然而,现有研究往往只针对果实裸露部分进行分割[22],容易将一个果实分离为多个部分。对于黄瓜而言,部分的遮挡并不影响人们对其位置和整体外观的判断,排除遮挡物使黄瓜分段反而不利于表型估算。

因此,本文提出一种能够适应不同遮挡程度的黄瓜图像分割方法。该方法优化 YOLO v8 的特征提取网络,嵌入可变形卷积(Deformable convolution network v4, DCNv4)[23]以增强空间适应性,并串联两个 RepNCSPELAN4 模块和额外的 C2F 模块,提高计算效率的同时细化特征的提取。在复杂性不同的两个场景数据集上对该方法进行验证。

1 材料与方法

1.1 数据

1.1.1 数据采集

图像数据采集于上海市正义园艺有限公司的两个黄瓜种植温室,分别为玻璃温室场景和塑料连栋大棚场景。场景 1 属于玻璃连栋温室(简称玻璃温室),采用无土栽培方式,在种植槽中单行种植,双行吊蔓,种植株距为 35 cm,吊蔓株距为 70 cm,吊蔓行距为 60 cm,相邻种植槽间距为 150 cm。场景 2为塑料连栋大棚(简称塑料大棚),采用传统的土壤栽培方式,双行一垄种植,行内相邻两株黄瓜的株距为 35 cm,相邻行距为 40 cm,相邻垄距为 120 cm。两个温室场景及其种植模式如图 1

所示。

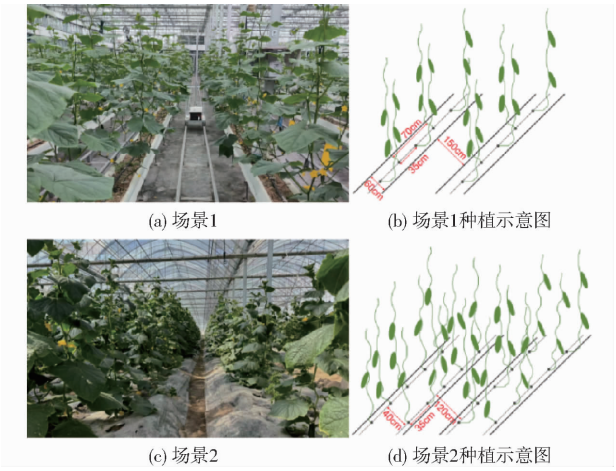

图 1　黄瓜种植场景

Fig. 1　Cucumber planting scenes

图像数据采集时间为 2024 年 5—6 月。采集设备由 Intel RealSense D435i 型深度相机、支架和笔记本式计算机构成，将深度相机固定在支架上，高度范围为 0.1 ~ 2.0 m，使用笔记本式计算机连接及控制，进行图像获取和存储。

考虑到不同光照和环境条件对图像的影响，选择在晴天和阴天的上午、中午和下午，分别对两个温室场景的黄瓜进行拍摄，确保了黄瓜数据的多样性和丰富性。拍摄过程中，相机保持与植株平行，距离植株大约 30 cm，最后从采集的图像中筛选得到 600 幅玻璃温室图像和 212 幅塑料大棚图像。场景 2 内种植密度和冠层环境明显区别于场景 1，遮挡更加严重，且黄瓜与背景不易区分，场景复杂性高。为了探究不同环境复杂性对黄瓜分割性能的影响，本研究将玻璃温室和塑料大棚采集的图像分别构建为两个独立的数据集（数据集 1：玻璃温室场景；数据集 2：塑料大棚场景），用于进一步验证模型在多样化种植环境下的适应性和鲁棒性。

1.1.2　数据分类与标注

模型训练需要输入原始图像及其对应的黄瓜掩膜。为此，采用 Labelme 工具对黄瓜进行像素级别的精确标注，并将得到的 json 文件转换为 txt 格式，以便模型训练。根据田间试验的经验，将黄瓜按照遮挡程度划分为 3 类：①无遮挡（uncover）：指未受任何遮挡的黄瓜。②部分遮挡（Partial cover，PCover）：指中间部分受到遮挡，但顶端裸露且整体位置可以预知的黄瓜。③严重遮挡（Mostly cover，MCover）：指顶端受到遮挡，无法确定整体位置的黄瓜，或者位于图像边缘的黄瓜。

其中，部分遮挡类别的标注包括遮挡物在内，整体仍是黄瓜形状，严重遮挡类别的标注常以不确定的形状出现，如图 2 所示。

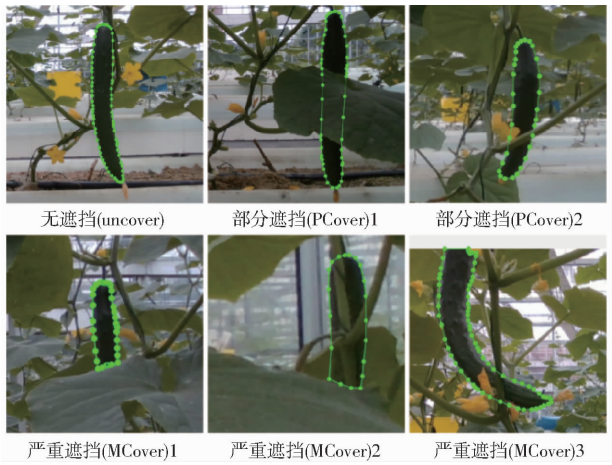

图 2　黄瓜遮挡分类与标注

Fig. 2　Cucumber occlusion classification and annotation

1.1.3　数据增强

在实际应用场景中，可能会遇到空气灰尘较多或镜头未清洁等导致图像模糊的情况，同时黄瓜的生长角度也多种多样。因此，本实验采用 Albumentations 库进行数据增强，以模拟这些复杂情况。例如，使用随机旋转模拟多种黄瓜生长角度，使用亮度调整模拟不同光照条件，以及使用高斯模糊来模拟镜头脏污等。数据的每次增强操作都以几何变换和像素变换的随机组合进行。采用的几何变换包括缩放、旋转、翻转和偏移；像素变换包括高斯模糊、噪声、亮度、对比度和通道像素值的修改。通过数据增强模拟了不同光照条件和各种角度的黄瓜图像。数据集 1 图像为 1 425 幅，数据集 2 图像为 600 幅。图 3 为不同增强处理后的图像示例。

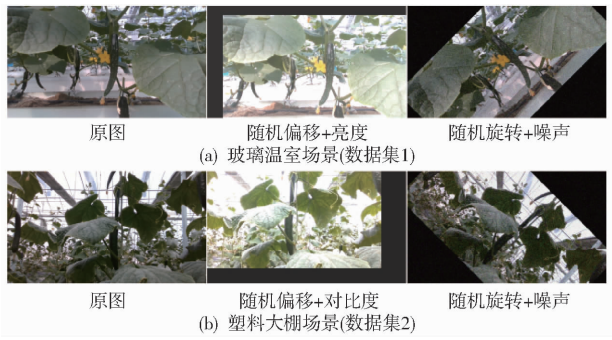

图 3　数据增强图像示例

Fig. 3　Example of augmented images

两个数据集均按照 8∶2 的比例划分为训练集和测试集。表 1 为两个数据集的训练集和测试集中 3 个类别果实标签的数量。

1.2　用于黄瓜果实分割的改进 YOLO v8 模型

1.2.1　基准模型 YOLO v8n

YOLO v8[24] 遵循 YOLO 架构，由骨干网络、颈部网络和检测头组成。其骨干网络使用 C2F 模块

表 1 各类别标签数量

Tab. 1 Number of labels for each category

数据集		遮挡程度			总计
		无遮挡	部分遮挡	严重遮挡	
数据集 1	训练集	989	603	955	2 547
	测试集	97	72	110	279
	总计	1 086	675	1 065	2 826
数据集 2	训练集	382	616	535	1 533
	测试集	35	59	74	168
	总计	417	675	609	1 701

取代 YOLO v7[25] 的 C3 模块，以增强梯度流。检测头采用无锚点的解耦设计，引入任务对齐分配器和分布焦点损失来优化损失计算。检测头的 3 个分支，分别用于分类、定位和分割任务。本文选用参数量最小的 YOLO v8n-seg 作为基准模型。

1.2.2 可变形卷积模块

为了增强模型的空间适应性，更好捕捉图像中各区域的空间关系，本研究引入 DCNv4[23] 构造新的卷积模块。如图 4 所示，输入指定通道数（c_1）为 3 的黄瓜特征图，首先用 1×1 卷积将输入通道转化为所需的输出通道数 16（c_2），扩展特征维度并保持空间分辨率。应用 DCNv4 动态调整卷积核的采样位置，对黄瓜果实的分割尤其重要。每个黄瓜的弯曲程度不同，形状存在差异，且严重遮挡的黄瓜边界不规则，可变形卷积能够有效捕捉空间变换和不规则模式，并灵活地适应这些变化，从而准确定位和分割黄瓜果实的边界。可变形卷积的输出经过批次归一化层（Batch normalization，BN），规范化特征分布以稳定训练并加速收敛，随后应用 SiLU 激活函数引入非线性映射，帮助模型捕捉数据中的复杂关系。最后，经过处理的特征图以特定的输出通道数传递到后续层进行下一步特征提取。

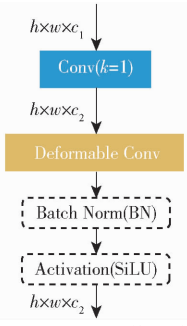

图 4 可变形卷积模块结构图

Fig. 4 Structure diagram of deformable convolution module

1.2.3 RepNCSPELAN4 模块

由于黄瓜生长环境中存在茎、叶、花和果等不同遮挡物，部分遮挡和严重遮挡这两种类别的特征相比无遮挡类别更加复杂。为了使模型更好应对多样性特征，在骨干网络中使用了 RepNCSPELAN4 模块，RepNCSPELAN4 模块的整体结构如图 5 所示。该模块通过并行路径处理输入的特征图，然后将它们连接并进一步转换。具有指定通道数（c_1）的输入特征图首先通过第 1 层卷积层，将通道维度转换为 c_3，随后在通道维度将特征图分成两部分，分别与后续特征图进行连接，并对第 2 部分进行处理。第 2 部分先经过一个 RepNCSP 模块，将通道维度从 $c_3/2$ 减少到 c_4，再经一个保留通道数量的卷积层，得到的特征图进入下一步处理，同时也作为连接的一部分。此时通道数为 c_4 的特征图依次经过 RepNCSP 模块和卷积层 Conv，随后进入最后的连接。Concat 连接了 4 部分的特征图，按照先后顺序，沿着通道维度进行连接，最终获得通道总数为 c_3+2c_4 的组合特征图，经卷积将其转换为原始空间维度（h，w）和指定的输出通道数 c_2。该模块设计通过并行路径处理输入来增强特征提取，使网络能够捕获不同的特征并提高其表示能力。这使细节特征不易丢失，模型能从不同处理阶段的特征图中获取关键特征信息，有效地区分果实与背景之间的细微差异。

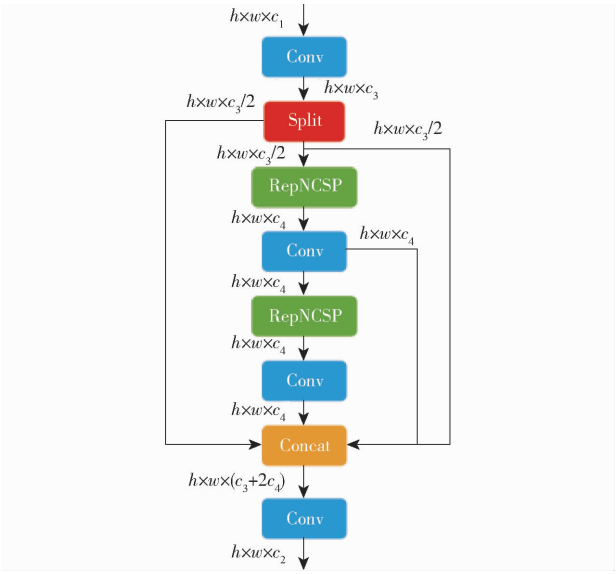

图 5 RepNCSPELAN4 模块结构图

Fig. 5 Structure diagram of RepNCSPELAN4 module

其中，RepNCSP 模块是基于 CSP 模块的结构改进而来的，它在网络中引入跨阶段部分连接来提高模型的上下文提取能力，有效结合多尺度和瓶颈特征，增强特征表示，使模型不受遮挡物的影响，准确分类并分割部分遮挡类别的黄瓜。如图 6 所示，模块分为 2 个分支，在主要分支上，先进行 1×1 卷积，接着经过若干个 RepNBottleneck 模块处理；另一个分支保存输入的特征信息，仅通过 1×1 的卷积提高特征图通道数，方便与主要分支相连接，以进行特征融合。

图 7 为 RepNBottleneck 和 RepConvN 模块结构

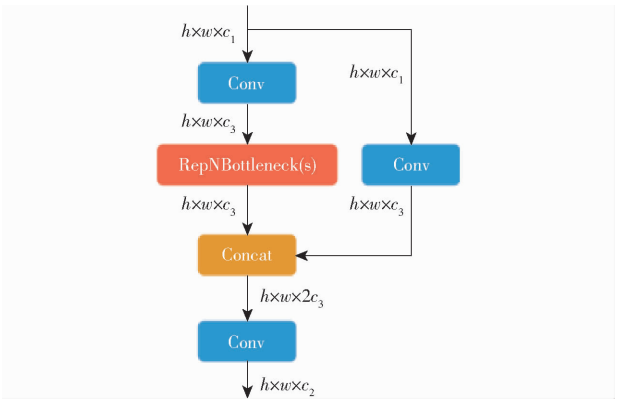

图 6　RepNCSP 模块结构图

Fig. 6　Structure diagram of RepNCSP module

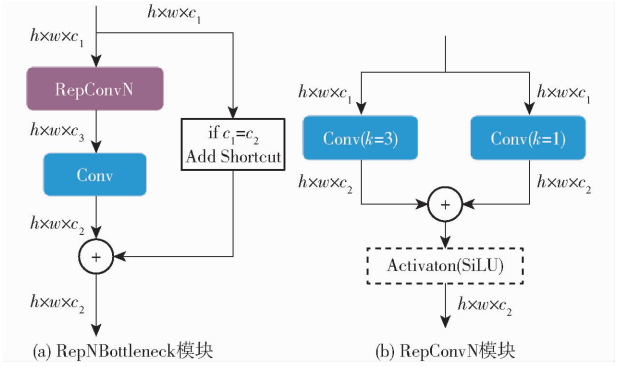

图 7　RepNBottleneck 模块和 RepConvN 模块结构图

Fig. 7　Structure diagrams of RepNBottleneck module and RepConvN module

图。RepNBottleneck 结合 RepConvN 和跳跃连接实现高效的梯度传播。输入的特征图通过一个 RepConvN 层，利用重参数化卷积操作提升表示能力，之后使用 3×3 卷积将通道维度从 c_1 降低至 c_3，这种设置有助于提高特征提取网络的健壮性。当设定的输入和输出的通道维度相同时，在上述操作之后进行跳跃连接，将原始输入特征直接与处理后的特征相加，使模块能够保留输入的关键信息，并在反向传播过程改善梯度流。RepConvN 模块是将输入特征图并行通过两个卷积层，3×3 的卷积用于捕捉空间特征，1×1 卷积用以细化特征映射，同时保留通道维度。随后将两者的输出张量逐元素相加，获得的特征图将应用 SiLU 激活函数进行非线性变换。这种设计使模型在训练过程中保持计算效率和稳定性。

1.2.4　改进 YOLO v8 模型整体结构

本文优化了 YOLO v8 的特征提取网络，使用可变形卷积（DCNv4）替换原第 1 层卷积以增强空间适应性，并串联 2 个 RepNCSPELAN4 模块和额外的 C2F 模块，提高计算效率的同时细化特征的提取。

图 8 为改进后的 YOLO v8 网络结构图，主要分为：骨干网络、颈部网络和检测头。骨干网络负责从输入图像（640 像素，640 像素，3）中提取层次特征，从 DCNv4 模块开始，该模块通过可变形卷积增强空间适应性。允许网络动态采样输入特征映射，有效地捕获几何变换，尤其对不规则和弯曲形状的黄瓜。接下来，卷积层和两个 RepNCSPELAN4 模块进一步细化提取的特征，增强模型对重叠黄瓜图像特征的感知与融合。RepNCSPELAN4 模块计算效率高，用于多尺度特征表示，将瓶颈结构与轻量级重参数化技术相结合，以最小的计算成本确保丰富的特征提取。额外的卷积层和 C2F 模块，以及 SPPF（Spatial pyramid pooling-fast），可以在多个尺度上池化特征，以捕获全局和局部上下文，有效关联部分遮挡黄瓜的“果肉”部分与“遮挡物”部分，还原部分遮挡类别黄瓜的整体位置。颈部融合来自骨干不同层的特征，使用上采样，串联额外的 C2F 模块来聚合不同分辨率的特征。最后，Head 以多种分辨率输出特征图，允许模型有效地处理不同大小的对象。

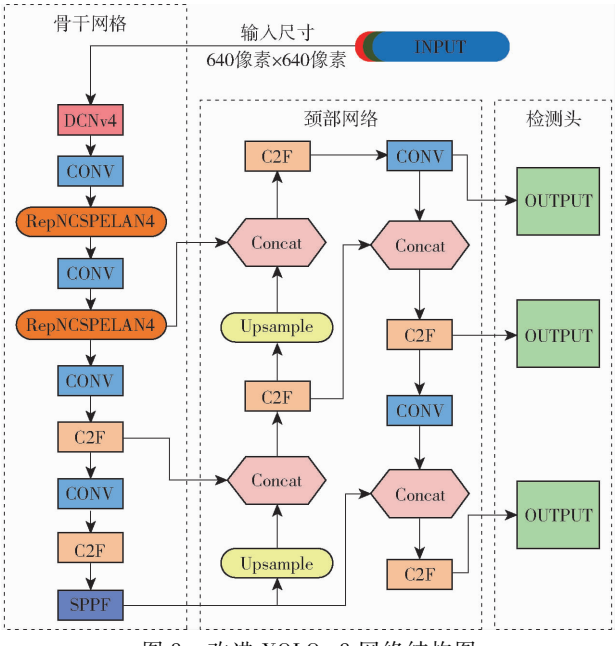

图 8　改进 YOLO v8 网络结构图

Fig. 8　Structure diagram of improved YOLO v8 network

DCNv4 和 RepNCSPELAN4 模块在改进骨干网络内的特征提取过程中发挥了关键作用。DCNv4 通过提供采样的灵活性来增强网络处理空间变换的能力。同时，RepNCSPELAN4 模块确保了高效的多尺度特征提取，使体系结构在保留多样特征表示的同时最大限度地减少计算开销。改进措施显著增强了模型的性能，进一步增强了模型在下游任务中捕获复杂和多样化特征的能力。

2　实验

2.1　实验环境

实验平台搭载了 Intel（R）Gold 6226R CPU 中

央处理器(主频为 2.90 GHz),配备 NVIDIA 的 GeForce RTX3090 GPU 高性能显卡(显存为24 GB),足以支撑深度学习模型的训练,确保实验稳定、快速进行。操作系统为 Ubuntu 20.04.5 LTS,采用 Pytorch 1.12 深度学习框架,并使用 Python 3.8 作为编程语言。

2.2 训练设置

在模型训练过程中,初始学习率设置为 0.01,批大小设置为 16,动量为 0.937,以确保模型权重稳定、有效更新;选择 SGD 优化器并设定权重衰减系数为 0.000 5,使模型在面对未知数据时有更好的泛化能力。输入图像在进入模型时,尺寸被统一调整为 640 像素×640 像素,设定 1 000 次迭代,每训练完一次迭代,模型会进行一次验证,指导模型有效学习。

2.3 评价指标

为了客观地评价模型在黄瓜分割任务中的有效性,本实验使用精确率(Precision,P)、召回率(Recall,R)、平均精度均值 mAP50 和 mAP50-95 对模型的性能进行评价。

另外,采用浮点运算量(GFLOPs)和参数量(Parameters)来衡量模型运算复杂度与资源需求,采用图像处理帧率(FPS)来衡量模型推理速度。

2.4 模型损失

模型在训练和验证过程中,损失函数和关键性能指标的变化趋势如图 9 所示。边界框损失值(box_loss)的下降表明模型对目标边界定位的逐步改进;分割损失值(seg_loss)随着迭代次数的增加迅速下降,大约在 100 次迭代之后,损失下降速度开始放缓,最终趋近 0,表明模型的分割性能逐渐稳定;分类损失值(cls_loss)的减少反映了模型区分不同类别能力的增强;分布式焦点损失值(dfl_loss)的降低进一步表明了模型识别难以分类样本能力的提升。同时,验证阶段的损失趋势与训练阶段一致,表明模型具有良好的泛化能力。

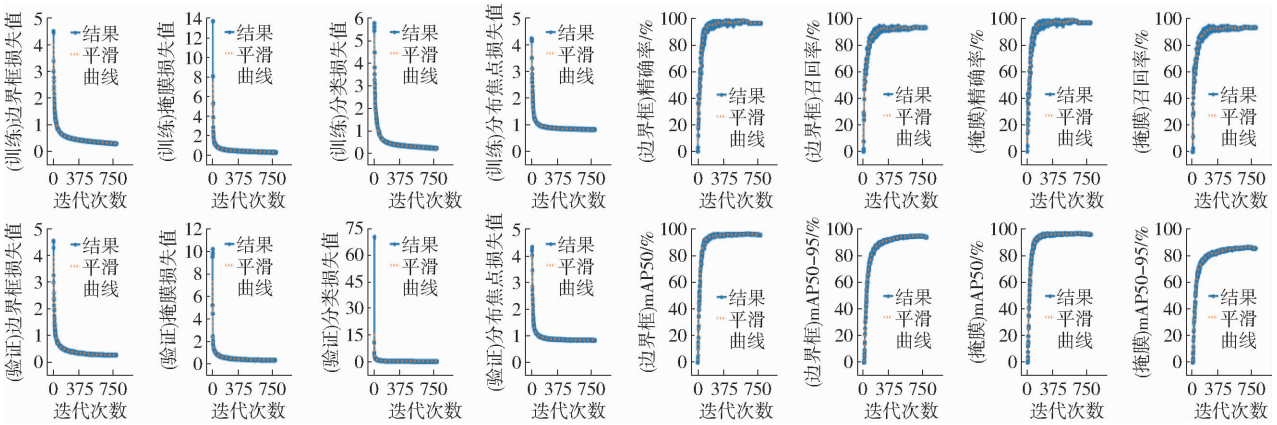

图 9 损失值与性能曲线

Fig. 9 Curves of loss values and performance metrics

随着迭代次数的增加,边界框(Box)和掩膜(Mask)上的精确率、召回率和不同阈值下的平均精度均值,都呈快速上升趋势,后期稳定在较高水平,进一步验证了模型的鲁棒性和通用性。所有子图中,平滑曲线与原始点高度一致,没有剧烈振荡或不正常的波动,这体现模型整体的可靠性。

2.5 模型在两个数据集上的实验结果对比

2.5.1 不同类别分割性能

改进后的模型在不同遮挡程度的黄瓜果实分割任务上表现出色,实验结果如表 2 所示。在数据集 1 中,所有类别平均精确率为 96.3%,召回率为 93.1%,平均精度均值 mAP50 为 96.2%,mAP50-95 为 85.3%。在不同类别中,无遮挡类别的分割性能最佳,平均精度均值 mAP50 和 mAP50-59 分别达到 97.4% 和 86.7%;部分遮挡类别次之,其 mAP50 为 96.8%,mAP50-59 为 85.9%;相比之下,严重遮挡类别表现较弱,mAP50 为 94.4%,mAP50-59 仅为 83.3%。这种差异表明,无遮挡和部分遮挡类别的特征较为明显,易于模型识别,而严重遮挡类别由于其复杂性,包括多种遮挡情况和特征混合,导致黄瓜主体和边界的分割难度显著增加。

在数据集 2 中,黄瓜与背景区分度大幅降低,遮挡情况复杂,所有类别平均精确率下降至 86.8%,召回率降至 81.9%,mAP50 为 90.0%,mAP50-95 为 77.0%,相比数据集 1,分别下降 9.5、11.8、6.2、8.3 个百分点。尽管各项指标都有所下降,但结果表明,在塑料大棚更复杂的栽培环境下,模型仍展现出良好的泛化能力。不同于数据集 1,在单个类别的分割上,部分遮挡类别的分割性能表现最佳,其平均精度均值 mAP50 为 90.9%,mAP50-59 为 78.4%,无遮挡类别性能略低于部分遮挡类别。这种差异主要是数据集中各类别样本数量的不平衡所致,如表 1 所示,数据集 2 中无遮挡果实数量不足 25%。

表 2　分割结果对比

Tab. 2　Comparison of segmentation results　%

数据集	类别	精确率	召回率	mAP50	mAP50－95
数据集 1	无遮挡	97.2	96.9	97.4	86.7
	部分遮挡	95.7	93.1	96.8	85.9
	严重遮挡	96.1	89.4	94.4	83.3
	所有类别	96.3	93.1	96.2	85.3
数据集 2	无遮挡	74.8	84.8	90.0	77.8
	部分遮挡	92.3	84.7	90.9	78.4
	严重遮挡	93.4	76.2	89.1	74.9
	所有类别	86.8	81.9	90.0	77.0

2.5.2　类别预测结果分析

两种场景数据集中黄瓜类别的预测结果如图 10、11 所示。对于数据集 1 中的黄瓜，预测效果最好的是无遮挡类（uncover），预测精度达到 97%；部分遮挡（PCover）的预测精度为 93%，被错误分类为严重遮挡和背景的比例均为 3%，有 1% 被预测为无遮挡；严重遮挡（MCover）的预测精度为 90%，被错误分类为无遮挡（uncover）和部分遮挡（PCover）的比例均为 2%，6% 被预测为背景（background）。数据集 2 内的黄瓜类别预测，无遮挡（uncover）的准确率最高为 89%，有 11% 错误分类；部分遮挡（PCover）为 86%；严重遮挡（MCover）的分类精度下降到 77%。

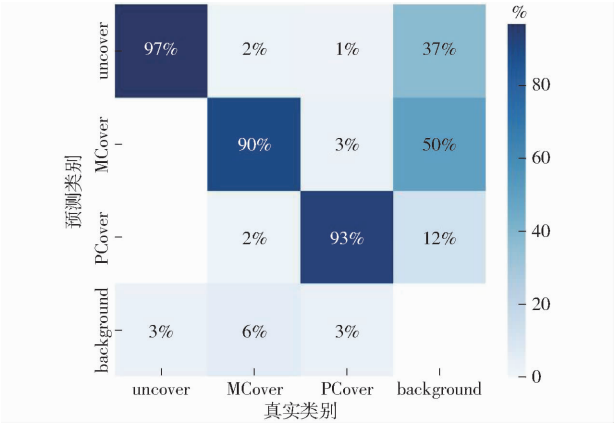

图 10　数据集 1 混淆矩阵

Fig. 10　Confusion matrix of dataset 1

其中，无遮挡（uncover）的特征区分度高，所以有更高的预测精度，分别比部分遮挡（PCover）和严重遮挡（MCover）高 4、7 个百分点。部分遮挡（PCover）错误分类的主要原因是遮挡物情况多样；而严重遮挡（MCover）因为既涵盖无遮挡的特征，又有部分遮挡的特征，情况复杂，所以分类误差相对较大。另外，被预测为背景（background）的实例均是像素面积很小且处于边缘的黄瓜，模型不易识别。由于黄瓜叶柄颜色和形状与幼期果实相近，所以处于背景中的部分叶柄也会被错误识别为黄瓜，两个数据集的背景（background）误识为黄瓜的实例中，分别有 50% 和 57% 被识别为严重遮挡（MCover），这也从侧面反映了严重遮挡类别（MCover）的复杂性。

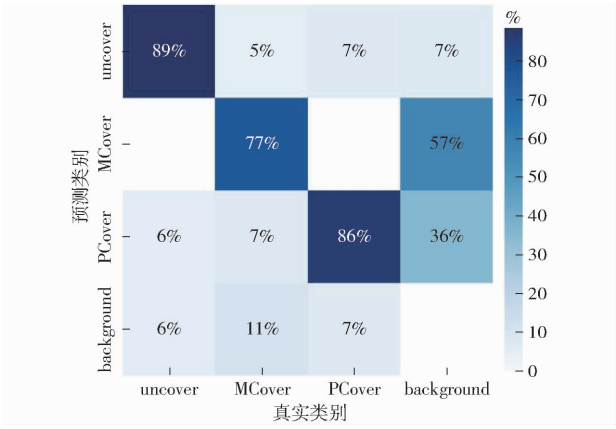

图 11　数据集 2 混淆矩阵

Fig. 11　Confusion matrix of dataset 2

基于分割所得的掩膜，可以计算获得黄瓜的表型参数，无遮挡和部分遮挡的果实掩膜呈现了黄瓜的完整形态，表型参数的获取比较容易，但对于严重遮挡的果实，仅使用目标检测和分割的方法难以实现表型参数的计算，给黄瓜生长智慧管控带来了困难。所以，对黄瓜不同遮挡程度类别的预测也十分必要。如何尽可能减少遮挡的影响，并利用遮挡信息辅助农业表型机器人分析，这是未来的研究方向。

2.5.3　分割结果对比

图 12 为模型在数据集 1、2 上的分割结果，包括不同遮挡程度的果实。其中，A－H 来源于玻璃温室场景（数据集 1），I－P 来源于塑料大棚场景（数据集 2）。

从图 12D、H、L、P 可以看到，改进 YOLO v8 拥有更精细的分割结果，分割的掩膜与果实更贴合，对于边界的分割能力优于原始模型，且黄瓜边界框的置信度均高于原始模型预测结果。更重要的是，改进的 YOLO v8 能清楚分割多个重叠的黄瓜，避免了“一物多识”的问题，如图 12K、L 所示，原始模型将重叠严重的两个黄瓜识别为多个，而改进 YOLO v8 可以准确分割并区分其类别。另外，对于严重遮挡类别和处于边缘的小黄瓜，改进 YOLO v8 有较好的识别能力，如图 12 O、P 所示，原始模型将叶柄错误识别为严重遮挡类别的黄瓜果实，而改进后 YOLO v8 能将其正确视为背景，改善了原始模型存在的误检问题。从整体上看，改进后的 YOLO v8 模型对细节特征的感知效果大大提升，不仅显著提高了边界分割能力，还在处理遮挡、多重重叠和多尺度目标分割上展现了卓越的能力，并且在玻璃温室和

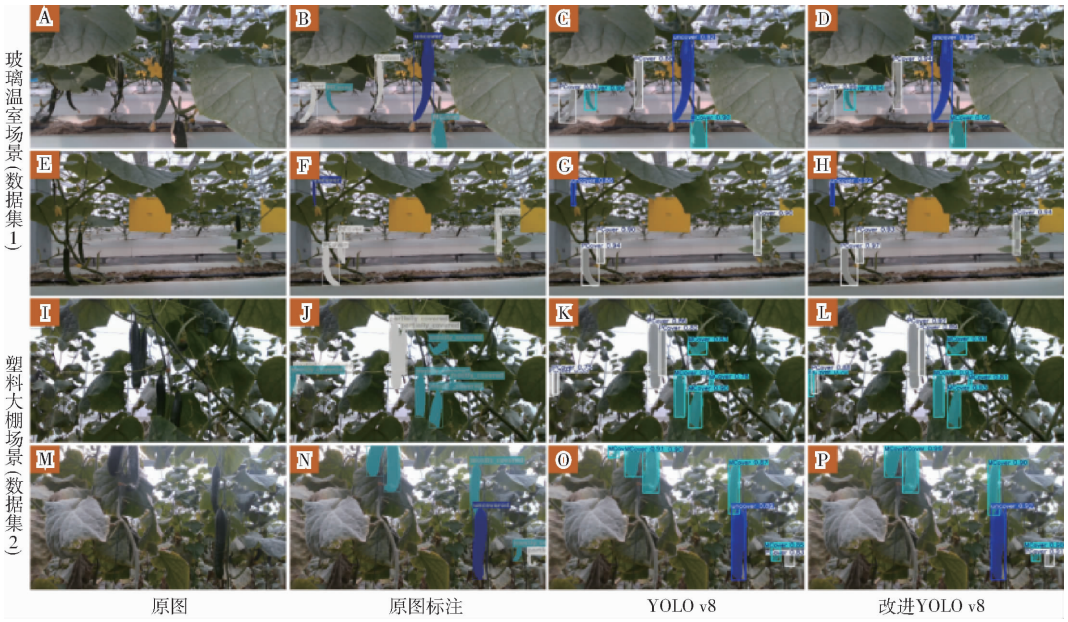

图 12 模型分割结果示例

Fig. 12 Examples of model segmentation results

塑料大棚两种不同复杂场景中,表现出更强的鲁棒性与适应性。

2.6 消融实验

为了明确所提的改进方法对模型分割性能的影响,以 YOLO v8 为基准模型并构造了 5 种变体,通过 6 组消融实验来量化、分析改进方法对模型整体性能的贡献。实验在数据集 1 上进行。实验结果如表 3 所示,基准模型 YOLO v8 精确率为 94.4%,召回率为 91.8%,mAP50 和 mAP50 - 95 分别为 95.8% 和 78.8%。引入可变形卷积模块(DCNv4)后,精确率提升了 1.1 个百分点,mAP50 提升了 0.3 个百分点,而召回率和 mAP50 - 95 略有下降。单独添加 RepNCSPELAN4 模块,使模型精确率提升了 1.3 个百分点,mAP50 提升了 0.3 个百分点,但对召回率和 mAP50 - 95 的影响不大。实验同时加入 2 个 RepNCSPELAN4 模块考察对模型性能的影响,结果显示,精确率和召回率分别提升至 95.9% 和 92.7%,mAP50 和 mAP50 - 95 分别提升至 96.0% 和 79.3%,相比基准模型,各项指标均有所提升,RepNCSPELAN4 模块使模型能有效捕捉特征。在此基础上,对两个模块协同作用进行实验,在置换了 DCNv4 的基础上,加入一个 RepNCSPELAN4 模块,精确率达到 96.1%,mAP50 - 95 显著提升至 82.4%,这表明,模型精度有效提升,尤其是对黄瓜边界等精细分割上。当 DCNv4 模块和 2 个 RepNCSPELAN4 模块同时引入,模型达到了最佳性能,精确率和召回率分别为 96.3% 和 93.1%,相比提升了 1.9、1.3 个百分点,较好地改善了模型误检和漏检的问题;另外,平均精度均值提升十分明显,mAP50 和 mAP50 - 95 分别达到了 96.2% 和 85.3%,mAP50 - 95 相比提升了 6.5 个百分点。实验结果表明,本文所提的改进方法能够增强模型对复杂场景的适应能力,在不同遮挡程度的黄瓜分割任务上,比原始模型有更优秀的表现。

表 3 消融实验结果对比

Tab. 3 Comparison of ablation experiment results

%

模型	精确率	召回率	mAP50	mAP50 - 95
YOLO v8	94.4	91.8	95.8	78.8
+ DCNv4	95.5	91.7	96.1	78.3
+ RepNCSPELAN4 × 1	95.7	91.7	96.1	78.7
+ RepNCSPELAN4 × 2	95.9	92.7	96.0	79.3
+ DCNv4 + RepNCSPELAN4 × 1	96.1	91.4	95.7	82.4
+ DCNv4 + RepNCSPELAN4 × 2	96.3	93.1	96.2	85.3

2.7 不同模型实验结果对比

为了进一步评价改进模型的有效性,在相同实验环境下,使用原始 YOLO v8、YOLO v7、YOLACT 和 Mask R - CNN,在数据集 1 上进行训练和验证。由表 4 可知,改进后的 YOLO v8 与其他模型相比,在精确率、召回率和两个不同阈值的平均精度均值上都有较大提高,与原始 YOLO v8 相比分别提升了 1.9、1.3、0.4、6.5 个百分点,与 Mask R - CNN 相较分别提高了 17.6、7.6、5.1、6.6 个百分点,与 YOLACT,以及同系列的 YOLO v7 相比也有显著提升,其中,mAP50 - 95 分别提高了 46.2、16.3 个百分点。改进后的 YOLO v8 分割精度显著提高,能够更好地应对分割任务中的复杂挑战。

表 4　不同模型实验结果

Tab. 4　Experimental results of different models

模型	精确率/%	召回率/%	mAP50/%	mAP50－95/%	浮点运算量	参数量	帧率/(f·s^{-1})
Mask R－CNN	78.7	85.5	91.1	78.7	2.550×10^{12}	4.398×10^{7}	11.5
YOLACT	58.9	60.7	58.9	39.1	8.487×10^{11}	5.373×10^{7}	30.5
YOLO v7	94.4	92.1	95.1	69.0	1.419×10^{12}	3.785×10^{7}	107.5
YOLO v8	94.4	91.8	95.8	78.8	1.200×10^{11}	3.259×10^{6}	285.7
改进的 YOLO v8	96.3	93.1	96.2	85.3	9.010×10^{11}	3.259×10^{6}	138.9

此外，尽管改进的 YOLO v8 的浮点运算量有所增加，从原始的 1.200×10^{11} 增长至 9.010×10^{11}，但相较于 Mask R－CNN 和 YOLO v7，改进的 YOLO v8 仍然保持了较低的计算复杂度。同时，其参数量保持在 3.259×10^{6}，与原始 YOLO v8 一致，仅为 Mask R－CNN 的 7.4%，显著降低了模型存储和部署的资源需求。改进的 YOLO v8 对图像的推理速度是 138.9 f/s，虽然低于原始 YOLO v8 的 285.7 f/s，但分别为 Mask R－CNN、YOLACT 和 YOLO v7 帧率的 12.1、4.6、1.3 倍，符合实际应用的需求。总体而言，改进的 YOLO v8 模型不仅在精确率、召回率和分割性能上均显著优于原始模型及其他对比模型，还保持了轻量化的设计和较高的推理速度，在需要高精度和高效分割的任务中，改进的 YOLO v8 模型具有更大的优势。

3　结论

(1)针对黄瓜分割任务中存在的分割精度低、未区分遮挡程度、果实被分段等问题提出了一种基于 YOLO v8 的改进模型：在特征提取网络中引入可变形卷积(DCNv4)以增强模型空间适应性，使模型能有效地捕获几何变换，灵活处理不规则目标；采用 RepNCSPELAN4 模块串联额外的 C2F 模块，在网络中引入跨阶段部分连接来提高模型的上下文提取能力，使模型能有效结合多尺度和瓶颈特征，在保留多样的特征表示同时最大限度地减少了计算开销。

(2)根据场景复杂性的不同，自建了两个数据集，并结合几何变换和像素变换进行数据增强，其中涵盖不同光照条件和各种角度的黄瓜图像，确保数据集的多样性和丰富性。根据遮挡程度进行分类标注，还原黄瓜完整形态避免分段，更好地为图像分割和表型分析提供图像数据基础。

(3)在两个数据集上开展丰富的实验，实验结果表明改进 YOLO v8 能有效捕捉图像的细节特征，在边界分割、遮挡和重叠处理以及多尺度目标的分割任务上表现良好。在玻璃温室场景中，模型的精确率为 96.3%，召回率为 93.1%，平均精度均值 mAP50 为 96.2%，mAP50－95 为 85.3%，相较原始提升了 6.5 个百分点。在塑料大棚场景中，模型精确率为 86.8%，召回率为 81.9%，平均精度均值 mAP50 为 90.0%，mAP50－95 为 77.0%。无论是整体上，还是单个分类上的分割，两个数据集上的优秀表现都表明改进 YOLO v8 具有较强的泛化能力，能适应多样化的种植环境。此外，与 YOLO v7、YOLACT 和 Mask R－CNN 等模型对比，改进的 YOLO v8 不仅拥有更高精度，还能保持较高的帧率(138.9 f/s)，较低的计算复杂度和较小的参数量。综合来看，该模型兼具泛化性、鲁棒性和实时性，能在多种复杂环境中准确、高效完成黄瓜果实的分割，为表型机器人的应用及农业智能化管理提供新思路。

参　考　文　献

[1] WEI Xue, DING Haifeng, JIN Tao, et al. CucumberAI: cucumber fruit morphology identification system based on artificial intelligence[J]. Plant Phenomics, 2024, 6: 0193.

[2] BAI Yuhao, GUO Yunxiang, ZHANG Qian, et al. Multi-network fusion algorithm with transfer learning for green cucumber segmentation and recognition under complex natural environment[J]. Computers and Electronics in Agriculture, 2022, 194: 106789.

[3] WANG Ning, QIAN Tingting, YANG Juan, et al. An enhanced YOLO v5 model for greenhouse cucumber fruit recognition based on color space features[J]. Agriculture, 2022, 12(10): 1556.

[4] 刘振宇，丁宇祺. 自然环境中被遮挡果实的识别方法研究[J]. 计算机应用研究，2020，37(增刊 2)：333－335，339.
LIU Zhenyu, DING Yuqi. Research on recognition method of occluded fruit in natural environment[J]. Application Research of Computers, 2020, 37(Supp. 2): 333－335, 339. (in Chinese)

[5] LI Dahua, ZHAO Hui, ZHAO Xiangfei, et al. Cucumber detection based on texture and color in greenhouse[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2017, 31(8): 1754016.

[6] MAO Shihan, LI Yuhan, MA You, et al. Automatic cucumber recognition algorithm for harvesting robots in the natural environment using deep learning and multi-feature fusion[J]. Computers and Electronics in Agriculture, 2020, 170: 105254.

[7] LI Qianwen, JIA Weikuan, SUN Meili, et al. A novel green apple segmentation algorithm based on ensemble U - Net under complex orchard environment[J]. Computers and Electronics in Agriculture, 2021, 180: 105900.

[8] CHARISIS C, ARGYROPOULOS D. Deep learning-based instance segmentation architectures in agriculture: a review of the scopes and challenges[J]. Smart Agricultural Technology, 2024, 8: 100448.

[9] HE K, GKIOXARI G, DOLLAR P, et al. Mask R - CNN[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961 - 2969.

[10] GIMÉNEZ-GALLEGO J, MARTINEZ-DEL-RINCON J, GONZÁLEZ-TERUEL J D, et al. On-tree fruit image segmentation comparing Mask R - CNN and vision transformer models. Application in a novel algorithm for pixel-based fruit size estimation [J]. Computers and Electronics in Agriculture, 2024, 222: 109077.

[11] 龙洁花,赵春江,林森,等. 改进 Mask R - CNN 的温室环境下不同成熟度番茄果实分割方法[J]. 农业工程学报,2021,37(18):100 - 108.
LONG Jiehua, ZHAO Chunjiang, LIN Sen, et al. Segmentation method of the tomato fruits with different maturities under greenhouse environment based on improved Mask R - CNN[J]. Transactions of the CSAE, 2021, 37(18): 100 - 108. (in Chinese)

[12] LIU Xiaoyang, ZHAO Dean, JIA Weikuan, et al. Cucumber fruits detection in greenhouses based on instance segmentation [J]. IEEE Access, 2019, 7: 139635 - 139642.

[13] LÓPEZ-BARRIOS J D, ESCOBEDO CABELLO J A, GÓMEZ-ESPINOSA A, et al. Green sweet pepper fruit and peduncle detection using Mask R - CNN in greenhouses[J]. Applied Sciences, 2023, 13(10): 6296.

[14] 肖伸平,赵倩颖,曾甲元,等. 基于 YOLO - DCL 的复杂环境油茶果遮挡检测与计数研究[J]. 农业机械学报,2024,55(10):318 - 326,480.
XIAO Shenping, ZHAO Qianying, ZENG Jiayuan, et al. *Camellia oleifera* fruits occlusion detection and counting in complex environments based on improved YOLO - DCL[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(10): 318 - 326, 480. (in Chinese)

[15] 翟肇裕,张梓涵,徐焕良,等. YOLO 算法在动植物表型研究中应用综述[J]. 农业机械学报,2024,55(11):1 - 20.
ZHAI Zhaoyu, ZHANG Zihan, XU Huanliang, et al. Review of applying YOLO family algorithms to analyze animal and plant phenotype[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 1 - 20. (in Chinese)

[16] MA Jie, ZHAO Yanke, FAN Wanpeng, et al. An improved YOLO v8 model for lotus seedpod instance segmentation in the lotus pond environment[J]. Agronomy, 2024, 14(6): 1325.

[17] WANG Aichen, QIAN Weihao, LI Ao, et al. NVW - YOLO v8s: an improved YOLO v8s network for real-time detection and segmentation of tomato fruits at different ripeness stages[J]. Computers and Electronics in Agriculture, 2024, 219: 108833.

[18] YUE Xiang, QI Kai, NA Xinyi, et al. Improved YOLO v8-seg network for instance segmentation of healthy and diseased tomato plants in the growth stage[J]. Agriculture, 2023, 13(8): 1643.

[19] CHEN Jincheng, JI Chao, ZHANG Jing, et al. A method for multi-target segmentation of bud-stage apple trees based on improved YOLO v8[J]. Computers and Electronics in Agriculture, 2024, 220: 108876.

[20] 倪纪鹏,朱立成,董力中,等. 基于 SwinS - YOLACT 的番茄采摘机器人实时实例分割算法研究[J]. 农业机械学报,2024,55(10):18 - 30.
NI Jipeng, ZHU Licheng, DONG Lizhong, et al. Real-time instance segmentation algorithm for tomato picking robot based on SwinS - YOLACT[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(10): 18 - 30. (in Chinese)

[21] 赵文锋,黄袁爵,钟敏悦,等. 基于 OrchardYOLOP 的火龙果园多任务视觉感知方法[J]. 农业机械学报,2024,55(11):160 - 170.
ZHAO Wenfeng, HUANG Yuanjue, ZHONG Minyue, et al. Multi-task visual perception method in dragon orchards based on OrchardYOLOP[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 160 - 170. (in Chinese)

[22] WU Tao, MIAO Zhonghua, HUANG Wenlei, et al. SGW - YOLO v8n: an improved YOLO v8n-based model for apple detection and segmentation in complex orchard environments[J]. Agriculture, 2024, 14(11): 1958.

[23] XIONG Yuwen, LI Zhiqi, CHEN Yuntao, et al. Efficient deformable convnets: rethinking dynamic and sparse operator for vision applications[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 5652 - 5661.

[24] VARGHESE R, SAMBATH M. YOLO v8: a novel object detection algorithm with enhanced performance and robustness[C]// 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS). IEEE, 2024: 1 - 6.

[25] 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, 2023: 7464 - 7475.