

基于 EP-YOLO v8 的瓶栽金针菇最优抓取位置定位方法

叶大鹏^{1,2} 景均¹ 吴昊宇³ 李辉煌¹ 谢立敏^{1,2}

(1. 福建农林大学机电工程学院, 福州 350002; 2. 福建省农业信息感知技术重点实验室, 福州 350002;

3. 福建农林大学未来技术学院(海峡联合研究院), 福州 350002)

摘要: 针对工厂化瓶栽金针菇自动切根过程中, 夹持末端因结构设计导致行程固定, 进而影响抓取效果甚至切根质量的问题, 本文基于 YOLO v8 (You only look once) 构建改进的 Enoki-pick_region-YOLO v8 (EP-YOLO v8), 实现瓶栽金针菇整体及最佳受力区域(关键抓取区域)的精准定位与轮廓提取, 保障抓取参数的可靠性。该方法在网络优化基础上, 基于最小欧几里得距离 (Euclidean distance, ED) 构建掩膜关系归属与判断模型, 明确金针菇菇体与关键抓取区域掩膜间父子关系合并优化。通过解析合并前后关键抓取区域的相对位置编码, 确定抓取参数并进行坐标转换, 为建立末端控制映射模型实现末端机械手运动行程的精确控制提供基础。实验结果表明, 本文所提算法的金针菇菇体掩膜识别精确率达 99.3%, 关键抓取区域掩膜识别精确率达 99.6%。同时, 对比发现掩膜质量得到了提高, 获取的参数抓取区域宽度与实际宽度之间的误差仅为 0.7%, 抓取参数基本满足抓取条件, 能有效实现最优抓取位置的精准识别与定位。

关键词: 瓶栽金针菇; 采摘点; 采摘机器人; YOLO v8; 多目标识别

中图分类号: S24; TP391

文献标识码: A

文章编号: 1000-1298(2024)10-0051-11

OSID:



Optimal Grabbing Position Localization Method for Bottle-planted Enoki Based on EP-YOLO v8

YE Dapeng^{1,2} JING Jun¹ WU Haoyu³ LI Huihuang¹ XIE Limin^{1,2}

(1. College of Mechanical and Electrical Engineering, Fujian Agriculture and Forestry University, Fuzhou 350002, China

2. Fujian Key Laboratory of Agricultural Information Sensing Technology, Fuzhou 350002, China

3. School of Future Technology, Haixia Institute of Science and Technology,

Fujian Agriculture and Forestry University, Fuzhou 350002, China)

Abstract: Aiming at the problem that in the automatic root cutting process of factory bottle-planted enokis, the stroke of the clamping end was fixed due to the structural design, which affected the gripping effect and even the quality of root cutting, an improved enoki-pick_region-YOLO v8 based on the you only look once (YOLO v8) was constructed, realized accurate positioning and contour extraction of the whole bottle-planted enoki as well as the optimal stress area (the key picking region). The accurate localization and contour extraction of the whole bottle-planted enoki and the best stress region guaranteed the reliability of the grasping parameters. On the basis of network optimization, a mask attribution and judgment model based on the minimum Euclidean distance (ED) was constructed, the parent-child relationship between the enoki body mask and the key region mask was clarified, and they were merged for optimization. By analyzing the relative position encoding of the key region before and after the merger, the grasping parameters were determined and converted into coordinates, which provided the basis for establishing the end control mapping model to realize the precise control of the end manipulator's motion stroke. The experimental results showed that the algorithm achieved a mask recognition rate of up to 99.3% for the enoki body and 99.6% for the key picking region. At the same time, it was found that the quality of the mask was improved, and the error between the width of the picking area and the actual

收稿日期: 2024-06-11 修回日期: 2024-07-10

基金项目: 福建省农业信息感知技术重点实验室建设项目 (KJG22052A)

作者简介: 叶大鹏 (1971—), 男, 教授, 博士生导师, 主要从事农业生物环境监测与控制等研究, E-mail: ydp@fafu.edu.cn

通信作者: 谢立敏 (1985—), 女, 讲师, 博士, 主要从事非线性系统动力学和机器人运动控制研究, E-mail: lucy_min@163.com

width of the acquired parameters was only 0.7% , and the grasping parameters basically satisfied the conditions of grasping, which effectively realized the accurate identification and localization of the optimal grasping position.

Key words: bottle-planted enokis; picking point; harvesting robot; YOLO v8; multi-target recognition

0 引言

金针菇营养价值高,口感好,在我国广受欢迎^[1-3]。为满足国内市场日益增长的需求,工厂化瓶栽金针菇生产工艺日趋完善。切除根部是确保金针菇卫生和易于烹饪的重要环节。传统的金针菇切根依赖人工,虽然切面平整,但成本较高;现有自动切根设备的末端机械结构是以最小正常生长子实体的尺寸为参数设计,因此一般生长状态下的金针菇在切根时会因夹持过紧而导致切面不平整。此外,部分子实体会因生长不良引起末端夹持状态异常,无法满足自动切根要求^[4-9]。因此,研究一种有效的方案,实现对不同生长形态下金针菇抓取位置的精准定位和识别,进而主动控制机械手爪的张合,对工厂化瓶栽金针菇采收切根的低损化、高效化具有重要意义。

目前,目标检测技术在果蔬识别、精准定位以及智能采摘等领域得到广泛应用^[10-13]。已有研究的研究对象形态轮廓曲线通常较为简单且为单个体。然而,金针菇是具有纵向截面不规则的对象,由于其形态的复杂性,使得目标检测中提取对象的抓取部位和轮廓信息较为困难。

实例分割技术相较于目标检测技术则能获取更精细的轮廓信息和更丰富的语义信息^[14-18]。尽管文献[19-22]的研究证明了实例分割技术在获取精确轮廓和目标点位信息方面的优势,并得到广泛应用,但金针菇的特征使得现有方法无法直接适用于其抓取信息的获取。

本文针对自动采收切根过程中因机械手爪抓取不当引起的问题,提出一种改进的 EP-YOLO v8 (Enoki-pick-region-YOLO v8)模型。该方法首先实现金针菇整体与关键抓取区域的识别与分割。然后构建金针菇整体与关键抓取区域的父子掩膜归属关系判断及合并模型,减弱瓶间相似性和菇体不同区域相似性对抓取位置获取的干扰。最后,依据相对位置关系提取关键抓取参数,为建立映射关系模型,精准调控机械爪,实现智能化、低损化金针菇抓取奠定基础。

1 材料与方法

1.1 瓶栽金针菇智能抓取方法流程

工厂化瓶栽金针菇智能抓取方法流程如图 1 所

示。首先,对采集图像进行人工筛选,剔除低质量图像后扩增;接着,基于改进 YOLO v8 构建识别分割模型,实现金针菇和抓取区域的定位与分割;然后,对分割内容添加标识,依据欧几里得距离 (Euclidean distance, ED) 构建掩膜条件关系判断模型,并基于逻辑算法合并掩膜,微调掩膜质量;最后,获取关键抓取区域参数并进行空间坐标转换,为构建映射关系模型精准控制机械手,实现高效抓取奠定基础。

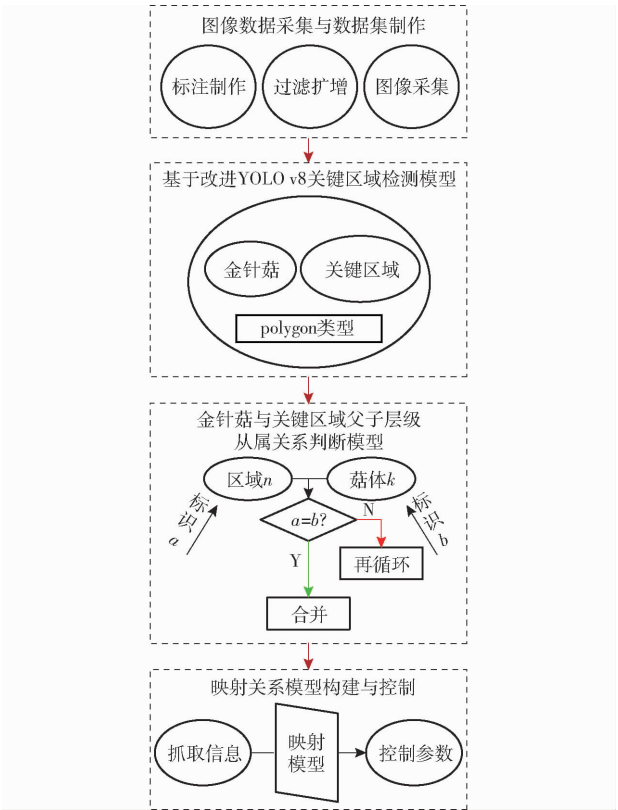


图 1 工厂化瓶栽金针菇智能抓取方法流程图
Fig.1 Flowchart of intelligent grasping method for factory-based bottle-planted enokis

1.2 图像数据采集

通过对福建省漳州市漳浦县万辰生物科技有限公司金针菇采收生产车间的实地采集,共得到瓶栽金针菇图像 1 897 幅,人工剔除低质量图像后,最终得 1 237 幅。图像采集设备为 iPhone13,采用正方形拍摄模式,采样距离为 200 ~ 250 mm,图像分辨率为 3 024 像素 × 3 024 像素。

1.3 数据集制作

1.3.1 最佳受力区域的力学研究与分析

通过游标卡尺测量,金针菇的根部平均高度为 29.81 mm,整株平均高度为 175.77 mm。观察整瓶

发现,瓶口以下的根部明显不适合作为抓取位置,而子实体中上部因间隙较大,也不适合。基于观测结果,最终选取瓶口以上 60 mm 为实验区域,并每隔 20 mm 标注 3 个试抓位置。利用 Universal TA 型质构仪进行压缩实验,以最佳受力区域作为抓取位置。

本研究对 101 株金针菇的 3 个区域(区域①、区域②、区域③)在不同施加载荷和下压速度条件下(15 N,1 mm/s;20 N,1 mm/s;15 N,1.5 mm/s)进行

了压缩实验,结果如图 2 所示。图 2 分别表示在不同外界条件下 3 个区域的压缩表现。图 2d 则展示了 3 个区域对应的多条荷重-位移曲线的拟合结果。由图 2 可以观察到,施加 10 N 的荷载时,区域①的形变量仅为 7.64 mm,明显小于区域②和③的 10.14 mm 和 30.99 mm,这表明区域①具有最佳的刚度。因此,选择区域①作为抓取位置能有效减少抓取过程中的形变。

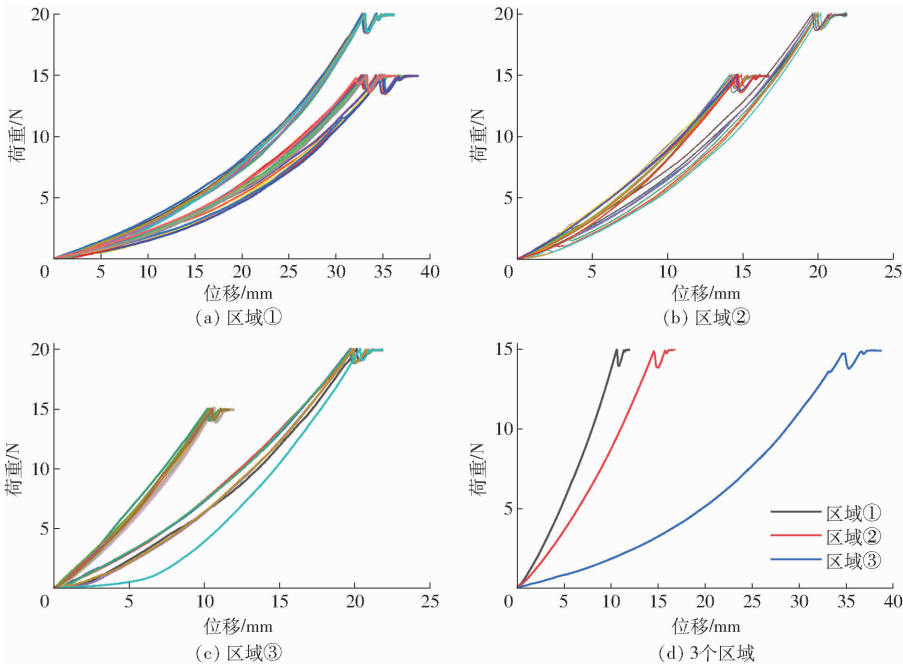


图 2 瓶栽金针菇 3 个区域的压缩实验曲线

Fig.2 Compression test curves of three regions of root of enoki

1.3.2 数据集标注划分与分析

依据 1.3.1 节的结论,使用 Labelme 的多边形模型对图像内容进行标注:金针菇外轮廓标记为“enoki”;抓取区域(区域①)标记为“pick_region”。接着以高斯模糊、镜像翻转、亮度调整等方法扩展数据,模拟不良拍摄情况。共获取 5 740 份样本,按比例 8:1:1 随机划分为训练集、验证集、测试集,其中训练集为 4 592 份,验证集和测试集各为 574 份。

瓶栽金针菇最佳受力区域(抓取区域)位于瓶口上方的菌柄部分。但该区域同菌柄其他部位的颜色、形态特征相似,易干扰其预测位置。如图 3 所示,仅标注“pick_region”时,发现部分个体抓取区域的预测位置出现偏移。因此,为了约束“pick_region”的相对位置避免偏移过大,采用“pick_region”与“enoki”同时标注的方案。

上述方法可构建抓取位置与金针菇整体的相对位置关联性,避免“pick_region”预测位置波动过大。为了具象化这种相对位置关联,统计.json 文件的标签分布,结果如图 4 所示。



图 3 关键位置预测错误结果图

Fig.3 Results of key position prediction errors

由图 4 可知,像素空间中“pick_region”的质心位置分布在“enoki”质心的下方,且同瓶子实体和抓取区域质心间像素距离,在 X 方向上介于

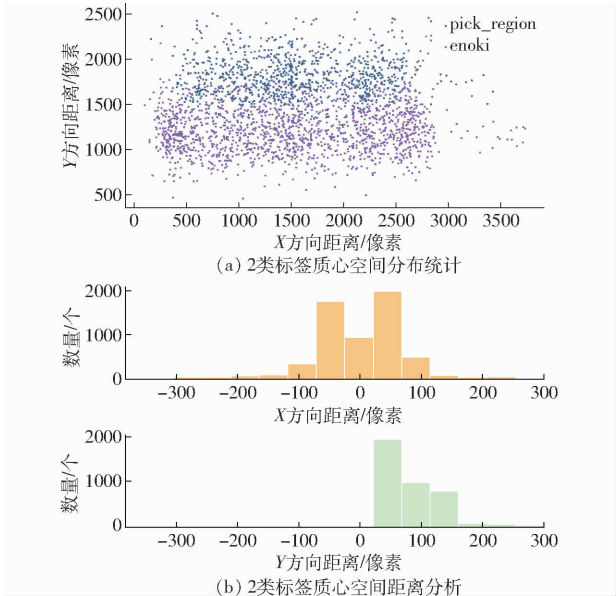


图4 标签信息分析统计

Fig. 4 Analysis and statistics of label information

[-350,250]像素,在Y方向上介于[0,300]像素。该结果明确了相对位置关系,有利于位置收敛,为建立同瓶“enoki-pick_region”关系提供支撑。

1.4 EP-YOLO v8 瓶栽金针菇分割模型

YOLO 是一种经典的单阶段目标检测算法^[23-27],2021 年受 YOLACT^[28-29] 等影响,YOLO 扩展出了 YOLO-seg,可同时处理目标检测和图像分割任务,其中 YOLO v8 因较高处理效率而受到广泛关注。本文在 YOLO v8 基础上提出一种改进的 EP-YOLO v8 以提高复杂易混淆特征的识别精度,其网络结构如图 5 所示。

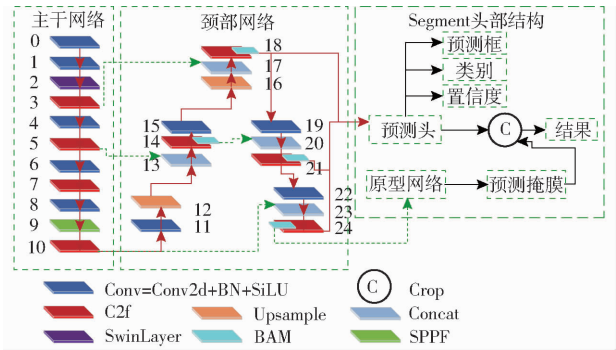


图5 瓶栽金针菇检测网络结构

Fig. 5 Network structure for detecting bottle-planted enokis

首先将 Swin Transformer^[30] 编码结构与传统卷积结合,强化全局特征利用率,提高主干网络特征提取能力。颈部引入瓶颈注意力模块 (Bottleneck attention module, BAM),集中算力处理关键特征。用 SiLU (Sigmoid-weighted linear unit) 激活函数代替整流线性单元 ReLU (Rectified linear unit) 激活函数,提升神经网络的收敛速度与泛化能力。图 5 中, SwinLayer 表示 Swin Transformer 特征提取层;BAM

表示注意力模块;预测头和原型网络表示预测框和掩膜的 2 组并行分支。

1.4.1 Swin Transformer 特征层

由于感受野范围受限,传统的卷积层在处理全局特征时效果不理想。而 Swin Transformer^[30] 则使用类似卷积神经网络中的层次化构建方法,在每一个 Stage 中将特征图下采样成若干个不同尺度的组,对每组特征图进行多头自注意力计算并融合不同尺度特征图;同时引入移位窗口多头自注意力,实现不同窗口间的信息交流,能更好地提取到图像的全局语义特征^[31]。

由于瓶栽金针菇关键抓取区域的位置预测需要借助上下文语义信息,本文通过引入 Swin Transformer 结构提取全局语义信息,将其与成熟的卷积下采样结构进行融合。该方案既能避免过多浮点运算量,也能优化神经网络的多层级特征提取效果,更易利用全局语义特征实现抓取区域的信息获取。其中定义新的特征提取层为 SwinLayer,并集成在 C2f 前 (Layer 2 处),以便在下采样前更好地捕获深层特征,实现更全面的特征表示,如图 5 所示。

1.4.2 基于瓶颈的注意力机制

注意力机制是一种模仿人类注意力机制的计算模型,它可以帮助神经网络在处理输入数据时更加关注重要的部分,从而提高模型性能和效率^[32]。BAM^[33] 基于卷积注意力机制 (Convolutional block attention module, CBAM) 而来,它通过学习输入数据的相关性和重要性来自适应地加强或减弱不同通道的特征响应,提高重要特征的敏感度。

为了缓解 SwinLayer 带来的计算复杂度,在颈部网络插入 BAM 缓解复杂度对网络精度的影响。多次实验结果表明,第 14、18、21、24 层处可最大化提升特征利用效率,同时不会带来浮点运算量的增加。

1.5 基于 ED 的父子层级掩膜归属关系判断及合并

用 YOLO v8-seg 进行瓶栽金针菇关键抓取区域检测任务时,任务窗口会出现大量无用掩膜,这不仅影响了可视化效果,还会影响算法效率。针对该问题,考虑构建父子层级掩膜归属关系判断模型,实现对掩膜归属关系的精准判断;同时基于逻辑算法对掩膜进行合并,以优化掩膜边界。

1.5.1 构建父子层级掩膜归属关系判断模型

定义标签“enoki”(A 类)和“pick_region”(B 类)。并定义 A 类标签为父级对象,B 类标签为子级对象,如图 6 所示。可见每个 A 类对象必然包含一个 B 类对象 (如果检测结果有 B 类)。

为了判断某 B 类对象是否与某 A 类对象有父子关系,提出一种基于最小欧几里得距离的父子层

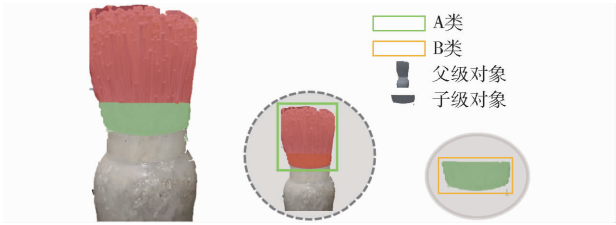


图 6 标签内容定义

Fig. 6 Label content definition



图 7 父子关系判断算法

Fig. 7 Father-son relationship judgment algorithm

级归属关系判断方法,流程如图 7 所示。①读取 det 函数变量“class”,遍历所有“enoki”和“pick_region”标签。②读取标签对象对应的轮廓点集像素坐标。③获取所有轮廓质心坐标。④依据 1.3 节初步筛选符合区域分布条件的质心对。⑤两两计算“pick_region”质心与“enoki”质心间的欧几里得距离,并保留其最小值 $\text{MIN}[\text{ED}]$ 。⑥创建配对空集合 $\{\text{group}\}$,储存获得最小 ED 值的 $\{\text{enoki-pick_region}\}$ 标签对。⑦过滤配对集合 group 元素,避免包含重复元素,最终得到的多组 $\{\text{enoki-pick_region}\}$ 标签对则表示包含父子层级关系的对象。

1.5.2 基于逻辑集合算法的掩膜合并方法

由于 B 类掩膜存在覆盖不全或超出实际范围的问题,无法精确地表示轮廓边界,坐标计算过程中会产生误差累积进而影响精准抓取。因此,提出一种基于逻辑算法的掩膜合并方法,如图 8 所示。

遍历所有 group 标签对,获取掩膜像素值并将其映射到 True-False 阈值空间,通过逻辑或运算合并掩膜,并以 Sobel 算子实行各向梯度运算,细化边缘。该方法既能优化算法效率,还能微调分割轮廓边界,缓解误差传播,提高整体性能。

1.6 关键抓取点评估算法

在抓取空间对称性较高的物体时,常以质心为

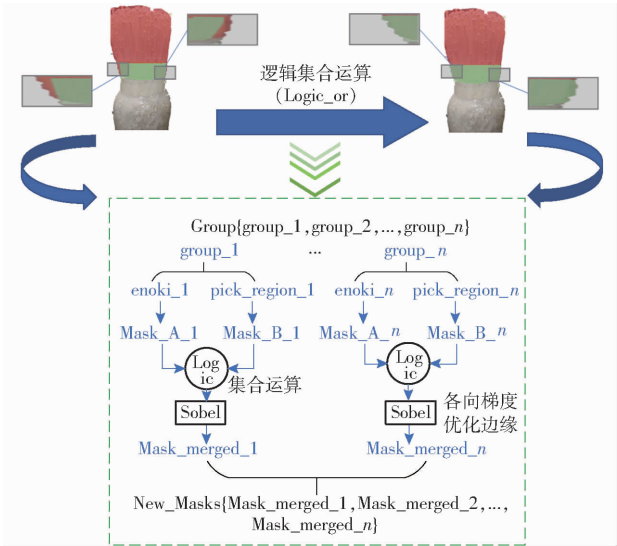


图 8 掩膜合并算法

Fig. 8 Mask merging algorithm

目标点。然而,对于形态不完全对称的物体,如瓶栽金针菇等,其质心并不一定表示最佳抓取位置。因此,本文提出一种关键抓取点评估算法。该算法在质心基础上,结合抓取区域的具体形态特征,确定最终抓取点,如图 9 所示。

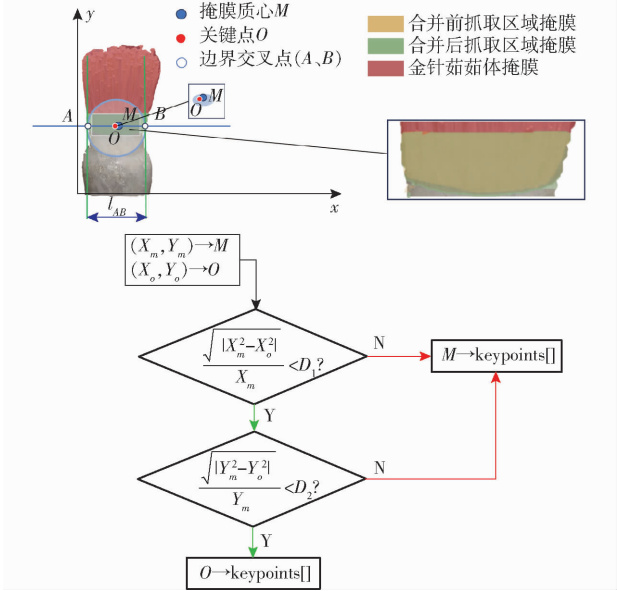


图 9 关键点获取算法示意图

Fig. 9 Schematic of key point acquisition algorithm

具体步骤如下:

- (1) 获取合并后的抓取区域掩膜质心 M 在像素空间内坐标。
- (2) 绘制过点 M 的水平直线,交掩膜边界于点 $A、B$ 。
- (3) 以 AB 长度为直径绘制圆 C ,记其圆心为点 O 。
- (4) 判断点 O 与点 M 的位置关系。若点 O 位于点 M 的范围椭圆内,则选择点 O 作为抓取点;否

则,仍然选取点 M 作为抓取点。椭圆范围由 D_1 、 D_2 确定,其中 D_1 表示允许点 O 的横坐标在点 M 左右波动 3% (经验值),而 D_2 表示允许点 O 的纵坐标在点 M 上下波动 1%, D_1 、 D_2 分别设置为 0.05 和 0.02。

(5) 输出关键抓取点“keypoints”和区域宽度 l_{AB} ,经坐标转换后映射为控制参数。其中“keypoints”可用于更新执行末端初始对齐位置, l_{AB} 可用于末端行程控制。

1.7 实验平台与通信方式

Tensor RT 加速处理后,EP-YOLO v8 模型被转换成.engine 格式,以便在实验平台上部署并控制末端抓手。该平台包括转盘、移动滑台、抓取执行末端、切刀和深度相机等部件,如图 10 所示。深度相机的基础坐标系位置已固定,以避免过多的运算量。

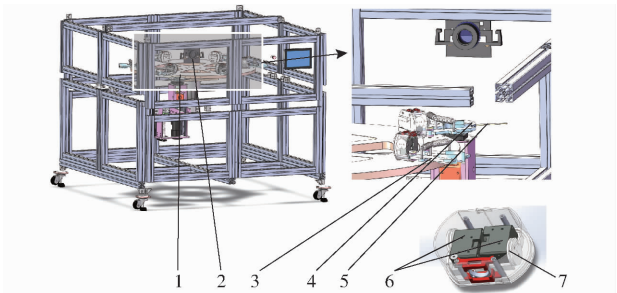


图 10 瓶栽金针菇自动抓取切根实验平台
Fig. 10 Automatic platform for bottle-planted enoki harvesting and root cutting experiment

1. 抓手转盘 2. 视觉输入端 3. 末端抓手 4. 三叶刀片转盘
5. 切根刀片 6. 舵机 7. 牵引轮

在此过程中,视觉系统捕获的图像将传输到边缘计算设备上,以获取金针菇抓取区域的宽度信息。经过坐标转换后,实际的宽度信息将通过 IIC 通信传输到 STM32 单片机上,通过 {Width-PWM} 映射关系转换为对应的 PWM 脉冲信号控制舵机的转动量,以精确控制末端抓手的行程量,从而实现对不同形态金针菇的抓取。在通信协议和数据处理方面,本文采用了定义数据帧的方式。例如,本文串口传输的完整数据形式为 [0xa5 (帧头),width (宽度信息),0xc5 (帧尾)],该方式有助于确保传输信息的质量和同步性。

2 实验

2.1 模型训练与环境

模型训练在 Windows 11 操作系统上进行,PyTorch 框架被集成到 Anaconda 环境中。程序采用 Python 3.7 编写,并使用 CUDA 11.7。硬件设置包含 12th Gen Intel (R) Core (TM) i9-12900H

2.50 GHz,主频为 2.50 GHz,其内部存储为 16 GB,显卡为 NVIDIA GeForce RTX3060。初始学习率设置为 0.01,共计 120 个轮次,其中前 3 个轮次为热身,动量因子为 0.937,批处理尺寸为 -1 (自动取值)。训练结果如图 11 所示。

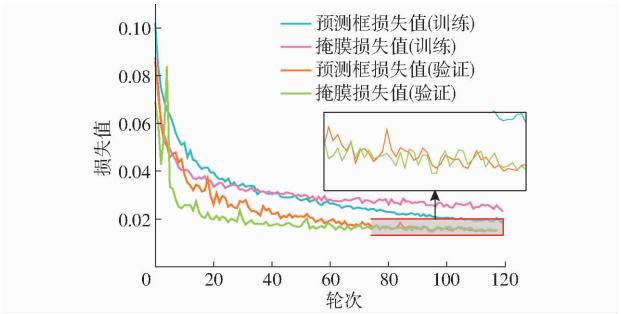


图 11 训练/验证损失曲线

Fig. 11 Training/validation loss curves

由图 11 可知,EP-YOLO v8 在训练阶段预测框预测损失值为 0.023,掩膜预测损失值为 0.02;验证阶段预测框预测损失值为 0.015,掩膜预测损失值为 0.0153,均低于训练损失,表现出良好的性能。

2.2 评价指标

YOLO 模型数据集以 COCO 数据集格式制成,因此本文选择平均精度均值 (Mean average precision, mAP) 作为评价指标,分别对检测精度和分割精度进行评价。mAP 越高,模型检测与分割效果越好。实验中设置不同阈值的交并比 (Intersection over union, IoU) 计算 mAP。本文选取 mAPv、mAP50 作为模型的评价指标。其中,mAPv (mAP@50-95) 表示 IoU 在 0.5~0.95 区间且增量为 0.05 的所有取值下的 mAP 的均值,mAP50 表示 IoU 为 0.5 时的 mAP。

2.3 消融实验

开展 YOLO v8-seg、YOLO v8-seg + SwinLayer 和 YOLO v8-seg + SwinLayer + BAM 模型的消融对比实验,探究不同功能模块对瓶栽金针菇检测网络性能的具体优化效果。可视化结果如图 12 所示,网络性能显著提升。消融实验各阶段的训练指标如表 1 所示。

以所有类开展分析,不难发现,相较于 YOLO v8,引入 SwinLayer 特征提取结构后,网络预测框平均检测精度指标 mAPv 和 mAP50 均提高 1.1%;掩膜平均检测精度指标 mAPv 和 mAP50 分别提高 0.8% 和 0.7%。该结果表明 SwinLayer 能有效提高模型全局特征的利用率,优化性能。引入 BAM 注意力机制后,预测框平均检测精度指标 mAPv 和 mAP50 分别提高 0.6% 和 0.3%;掩膜平均检测精度

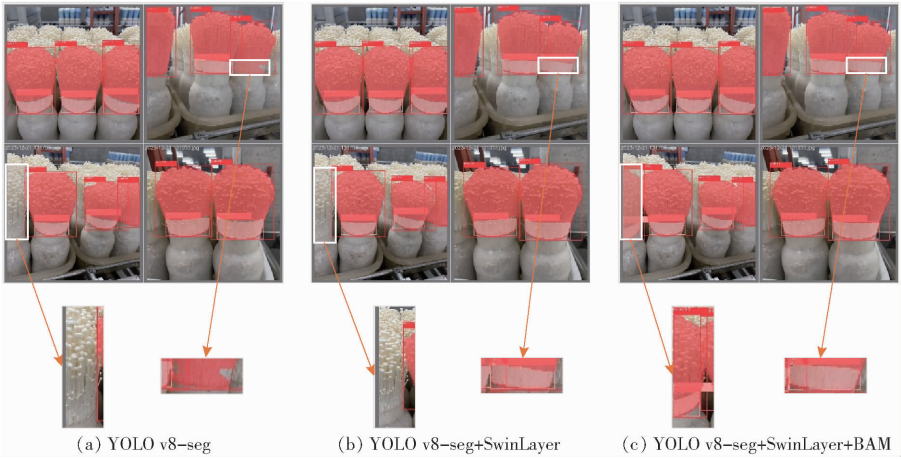


图 12 消融实验结果可视化

Fig. 12 Visualizations of ablation experiment results

表 1 消融实验结果

Tab. 1 Ablation experiment results

Swin Layer	BAM	所有类				金针菇				关键区域			
		预测框		掩膜		预测框		掩膜		预测框		掩膜	
		mAPv	mAP50	mAPv	mAP50	mAPv	mAP50	mAPv	mAP50	mAPv	mAP50	mAPv	mAP50
×	×	84.3	98.3	82.7	98.2	87.6	98.5	83.2	98.4	79.2	98.7	76.5	99.0
✓	×	85.2	99.2	83.4	98.9	87.9	99.0	85.9	99.1	80.3	99.1	78.9	99.4
×	✓	84.8	98.6	83.1	98.5	87.9	98.9	84.0	98.9	79.7	98.8	77.3	99.2
✓	✓	86.8	99.7	85.1	99.4	88.7	99.2	86.1	99.3	80.8	99.1	79.9	99.6

注：表中“✓”表示模块已添加，“×”表示模块未添加。

指标 mAPv 和 mAP50 分别提高 0.5% 和 0.3%。这证明 BAM 能够专注于提升重要特征,提升模型表现。同时集成上述模块,预测框平均检测精度指标 mAPv 和 mAP50 分别提高 3.0% 和 1.4% ;掩膜平均检测精度指标 mAPv 和 mAP50 分别提高 2.9% 和 1.2%。上述结果验证了本方法的可行性,优化后的网络能更精确地分割掩膜以提高定位精度,有助于后续研究工作的开展。

2.4 模型间横向实验对比

为进一步评估 EP-YOLO v8 的检测效果,将其与几类主流检测分割模型 (YOLO v5-seg、YOLO v7-seg、YOLO v8-seg、Mask R-CNN、YOLACT、YOLACT++、Swin Transformer、Sota Model^[19]) 展开横向对比,实验可视化结果如图 13 所示。

不难观察到,EP-YOLO v8 模型在检测分割能力和掩膜质量上明显优于其他同类检测和分割模

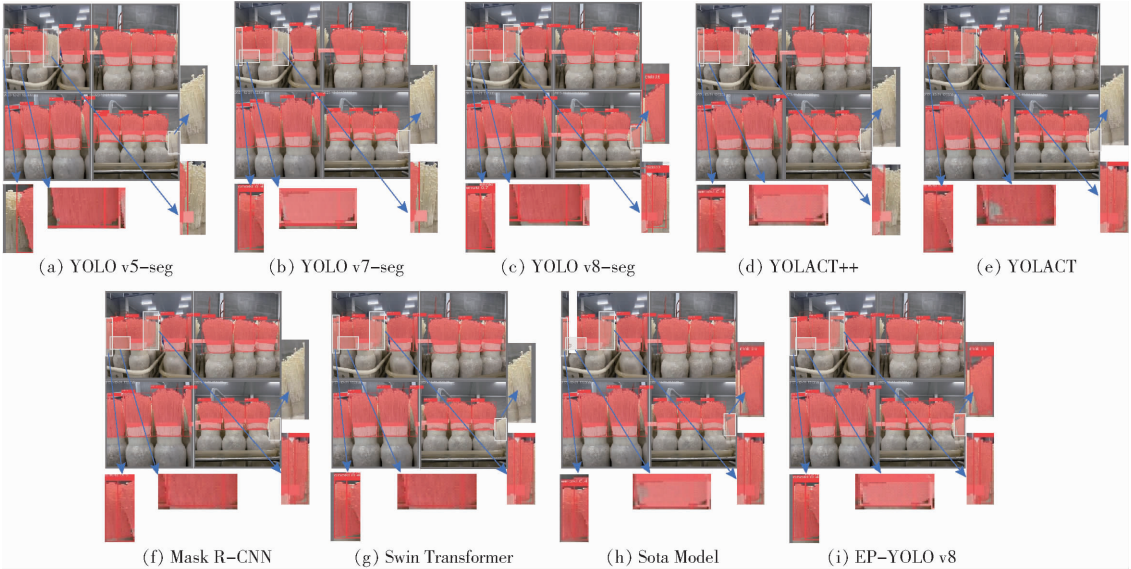


图 13 模型横向结果可视化

Fig. 13 Visualization of model horizontal result

型,表现出了良好的性能。

分析表 2 可知,在所有类中,EP-YOLO v8 展现出了出色的预测框和掩膜预测能力。其预测框和掩膜的 mAPv 均为最高,分别达 86.8% 和 85.1%。相较于 YOLO v8-seg,其 mAPv 分别提高 3.0% 和 2.7%;相较于 YOLO v5-seg,则分别提高 3.8% 和 3.7%。结合图 13 的可视化结果分析,EP-YOLO v8 与 Sota Model 均具有较强的掩膜分割能力,但在掩膜质量和细节方面,EP-YOLO v8 更胜一筹。

表 2 模型间横向结果对比

Tab.2 Horizontal comparison of results among models

模型	单幅图像 检测时 间/ms	所有类			
		预测框		掩膜	
		mAPv/%	mAP50/%	mAPv/%	mAP50/%
YOLO v5-seg	8.4	83.6	97.4	82.0	97.1
YOLO v7-seg	10.6	83.9	97.6	82.6	97.4
YOLO v8-seg	9.0	84.3	98.3	82.9	98.2
Mask R-CNN	10.3	83.9	97.3	81.2	97.4
YOLACT	12.8	83.9	97.8	82.5	97.3
Swin Transformer	15.6	84.3	98.2	82.8	98.1
YOLACT++	13.1	84.2	98.0	82.5	97.9
Sota Model	12.4	86.5	99.3	84.9	98.9
EP-YOLO v8	11.7	86.8	99.7	85.1	99.4

上述结果进一步验证了 EP-YOLO v8 在工厂化瓶栽金针菇识别分割任务中的良好适应性,为后续合并与采收点定位工作提供了坚实基础。

2.5 算法效果评估

2.5.1 与传统视觉方法的对比研究

工业零部件因标准化,传统视觉检测方法在准确率和时效性方面通常优于基于深度学习的方法。然而,对于农产品(如瓶栽金针菇)等,传统检测方法往往面临多重挑战。以检测时间和识别效果为指标,全面比较这 2 种检测算法的优缺点。在此过程中,传统算法通过对 1.3.2 节中金针菇标注区域的 RGB 颜色通道进行分析,确定上下阈值分别为 141、121、98 和 255、205、190。

由表 3 可知,无论在 PC 端还是在边缘设备上,传统视觉识别算法在实时性方面均优于 EP-YOLO v8。如处理单幅图像时,PC 端传统视觉算法的检测时间比 EP-YOLO v8 快 75.2%;而在边缘设备上,传统视觉算法的检测时间也比 EP-YOLO v8 快 74.4%。尽管 EP-YOLO v8 的检测速度相对较慢,但其在特征提取和表征方面表现更优秀。特别是基于颜色特征的图像处理方法使得传统视觉算法难以准确识别关键区域,容易误识别瓶口、瓶身等部分(图 14),从而影响参数抓取的准确性。相比之下,EP-YOLO v8 能够更精确地识别和分割目标,获取

更精确的形态特征参数,从而优化抓取效果。

表 3 EP-YOLO v8 与传统视觉检测时效性对比
Tab.3 Comparison of real-time performance between EP-YOLO v8 and traditional method ms

检测类别	PC 端	边缘设备
传统方法单幅图像检测时间	2.9	4.2
EP-YOLO v8 单幅图像检测时间	11.7	16.4
传统方法视频流单帧图像平均检测时间	3.1	7.1
EP-YOLO v8 视频流单帧图像平均检测时间	12.4	19.3

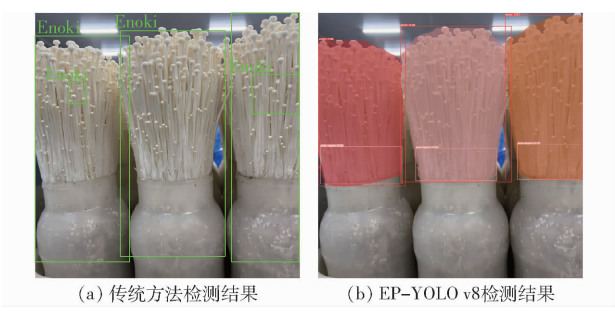


图 14 传统算法与 EP-YOLO v8 的识别效果对比

Fig.14 Comparison of recognition performance between traditional algorithms and EP-YOLO v8

2.5.2 合并效果评估分析

为评估掩膜合并效果,分别在普通视角(A1)、广视角(A2)、阴影普通视角(A3)和阴影广视角(A4)的拍摄条件下对 100 个样本进行合并效果评估。根据以下标准判断合并效果:

(1)合并的对象是否跨越了父级掩膜的边界:合并后的对象是否能够正确识别并包含在一个连续的父级掩膜内,确保合并的对象不会错误地延伸到相邻的非目标区域。

(2)对于分割后的样本,检查所有具有父子关系的掩膜是否被完全合并:要求能够准确识别并合并具有明确层级关系的掩膜。

(3)合并的对象是否同时包含了父级掩膜“enoki”以及子级掩膜“pick_region”的预测框:合并后的对象必须同时涵盖父子层级中所有相关的预测框,确保信息的完整性和准确性。

(4)未满足合并条件的对象是否被正确地从掩膜中移除:保证合并结果的纯净度,通过去除那些不符合合并条件的对象,提高最终合并掩膜的准确度和可用性。

(5)合并后掩膜的几何形状和位置准确性:合并后的掩膜是否保持了物体的自然形状,并且其在图像中的位置是否准确,避免了合并过程中可能出现的形变或位移。

(6)对变化场景的适应能力:在多变的场景下(包括在不同光照、遮挡情况以及不同尺寸和

形态的对象),合并算法是否仍然能够准确地执行。

实验结果如图 15 和表 4 所示。图 15 的第 1 行(B1)为采集的原始样本图像,第 2 行(B2)为 EP-YOLO v8 分割后的结果,第 3 行(B3)为合并失败的样例(包括合并失败、漏检等问题),第 4 行(B4)为微调后成功合并的结果。结合表 4 中的数据可以看出,微调后每份样本中成功合并对象的数量显著提升。同时,统计得到 4 类样本的合并成功率分别提高至 96.4%、95.8%、96.0% 和 96.5%,较调整前有一定改善,证明了合并算法的有效性。

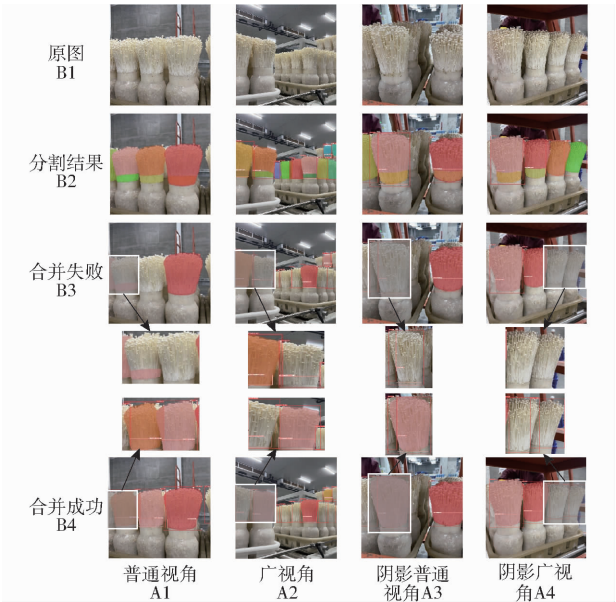


图 15 合并效果评估

Fig. 15 Evaluation of merging effects

表 4 合并结果评估

Tab. 4 Evaluation of merging results

样本种类	统计数量/ 份	调整前合并 成功率/%	调整后合并成 功率/%
普通视角	100	95.4	96.4
广视角	100	94.8	95.8
阴影普通视角	100	95.1	96.0
阴影广视角	100	95.7	96.5

2.5.3 抓取参数获取评估分析

如 1.6 节中所述,末端行为调控主要依托于 2 个相关联的参数:关键点(keypoint)和区域宽度 l_{AB} 。聚焦于区域宽度,探讨算法的可靠性。图 16 可视化了某样本获取关键点的结果。

图 16a 为掩膜合并结果,图 16b 为关键点获取情况,不难发现有 3 组成功合并的掩膜,其中质心及关键点信息如表 5 所示。

表 5 观察到“enoki”质心与合并后的质心 M 仍满足 1.3.2 节的相对位置关系,初步满足抓取条件;同时点 O 在点 M 基础上横向波动不超过 3%,实现

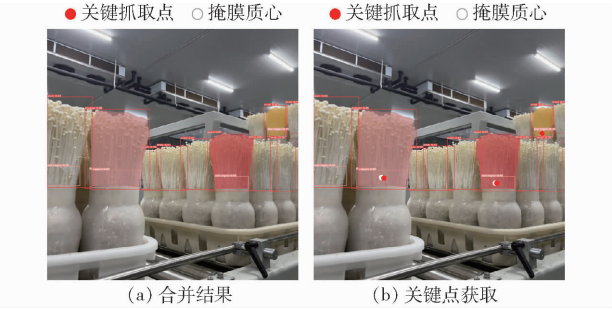


图 16 关键点获取结果对比

Fig. 16 Comparison of key point acquisition results

表 5 质心及关键点坐标信息

Tab. 5 Center of mass and key points

点类别	coordinates			像素
	组别 1	组别 2	组别 3	
“enoki”质心	(468,322)	(180,287)	(582,215)	
“pick_region”质心	(470,393)	(170,382)	(583,239)	
合并后关键区域质心 M	(470,384)	(171,374)	(587,232)	
关键点 O	(469,384)	(172,374)	(584,232)	

了依据个体形态二次优化抓取点的目的,确保了“一瓶一点”的抓取方案。

为进一步评估算法的可行性,以单瓶金针菇的关键区域抓取宽 l_{AB} 与实际宽度 b 进行对比验证,如图 17 所示。



图 17 瓶栽金针菇抓取宽度分析

Fig. 17 Analysis of enoki grasp-width

图 17 中,“enoki”质心坐标为(196 像素,191 像素),“pick_region”质心为(204 像素,348 像素),关键点“keypoints”质心为(200 像素,314 像素),抓取区域的像素宽度 $2l_{AB}$ 为 200 像素。由于已知拍摄距离和相机参数,转换坐标系后计算得到宽度为 86.2 mm,而人工测量得到的实际宽度 b 为 87.0 mm,二者误差为 0.9%,具有一定可靠性。

对 30 份样本进行测试,结果如图 18 所示,可观察到获取的抓取宽度与测量宽度相近,平均误差仅为 0.7%,能满足不同生长状态下瓶栽金针菇的识别检测和抓取需求。

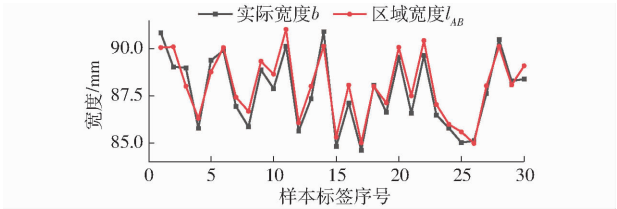


图 18 样本宽度对比
Fig. 18 Comparison of sample widths

3 结 论

(1)针对工厂化瓶栽金针菇自动切根设备抓取末端因交互性弱引起的切根质量问题,提出了一种优化方法。

(2)在 YOLO v8 基础上引入 Swin Transformer 特征提取层 SwinLayer 和 BAM 注意力模块,构建 EP - YOLO v8 模型。在自建的瓶栽金针菇数据集上测试发现,EP - YOLO v8 相较于 YOLO v8 - seg,其预测框和掩膜准确度指标 mAPv 分别提高 3.0% 和 2.7%,有效地提升了模型性能。与其他检测模型相

比,EP - YOLO v8 也表现优异。因此改进的 EP - YOLO v8 能实现工厂环境下瓶栽金针菇抓取位置的高质量分割。

(3)提出了父子层级掩膜关系归属判断及其合并算法,并根据合并结果进行二次微调。结果表明该方法在工厂环境常见的 4 种检测视角下,掩膜合并成功率分别达到 96.4%、95.8%、96.0% 和 96.5%,这说明算法能有效识别并判断两类掩膜合并,降低了同瓶子实体区域特征混淆、邻瓶关键区域关系相似、模型识别分割结果不准确等干扰因素对抓取参数准确性的消极影响。

(4)依据质心提出了一种关键抓取参数优化算法。相较于传统的质心定位方法,该算法能够更有效地缓解因边界非对称导致的质心偏移问题,并获取不同大小金针菇个体的抓取区域宽度。实验结果表明,该算法获取的抓取参数基本满足抓取条件,且抓取区宽与实际宽度间误差仅为 0.7%,能够有针对性地完成识别分割,为后续构建精准的金针菇抓取行程控制映射模型提供理论和方法基础。

参 考 文 献

[1] 陆欢,王瑞娟,刘建雨,等. 不同品种金针菇的营养分析与评价[J]. 食品与机械,2021,37(6):69-75,96.
LU Huan, WANG Ruijuan, LIU Jianyu, et al. Analysis and evaluation of nutrient components of different strains of *Flammulina filiformis*[J]. Food and Machinery, 2021, 37(6): 69-75, 96. (in Chinese)

[2] ZHANG Y, WANG D, CHEN Y, et al. Healthy function and high valued utilization of edible fungi[J]. Food Science and Human Wellness, 2021, 10(4): 408-420.

[3] ZHONG Y, DONG S, CUI Y, et al. Recent advances in postharvest irradiation preservation technology of edible fungi: a review [J]. Foods, 2022, 12(1): 103.

[4] LI C, XU S. Edible mushroom industry in China: current state and perspectives[J]. Applied Microbiology and Biotechnology, 2022, 106(11): 3949-3955.

[5] 高文硕. 瓶栽金针菇机械化采摘技术试验研究[D]. 北京: 中国农业科学院, 2021.
GAO Wenshuo. Research on mechanized picking technology of bottled flammulina [D]. Beijing: Chinese Academy of Agricultural Sciences, 2021. (in Chinese)

[6] 赵俊. 瓶栽金针菇自动采收关键技术研究及试验[D]. 福州: 福建农林大学, 2022.
ZHAO Jun. Research and experiments on key technologies for automatic harvesting of enoki mushrooms in bottles [D]. Fuzhou: Fujian University of Agriculture and Forestry, 2022. (in Chinese)

[7] 郑义姚. 基于深度学习的金针菇分类系统设计与应用[D]. 厦门: 华侨大学, 2021.
ZHENG Yiyao. Design and application of *Flammulina velutipes* classification system based on deep learning [D]. Xiamen: Huaqiao University, 2021. (in Chinese)

[8] 张俊. 面向工厂化褐菇种植的智能蘑菇采摘机器人设计[D]. 南京: 南京农业大学, 2021.
ZHANG Jun. Design of intelligent picking mushroom robot for factory-made brown mushroom [D]. Nanjing: Nanjing Agricultural University, 2021. (in Chinese)

[9] 易山林. 工厂化种植双孢菇自动采摘系统关键技术研究[D]. 扬州: 扬州大学, 2023.
Yi Shanlin. Research on key technology of automatic picking system for industrial planting of *Agaricus bisporus* [D]. Yangzhou: Yangzhou University, 2023. (in Chinese)

[10] REN Z, FANG F, YAN N, et al. State of the art in defect detection based on machine vision[J]. International Journal of Precision Engineering and Manufacturing-Green Technology, 2022, 9(2): 661-691.

[11] 王昱,姚兴智,李斌,等. 基于改进 YOLO v7 - tiny 的甜椒畸形果识别算法[J]. 农业机械学报, 2023, 54(11): 236-246.
WANG Yu, YAO Xingzhi, LI Bin, et al. Maldormed sweet pepper fruit identification algorithm based on improved YOLO v7 - tiny[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(11): 236-246. (in Chinese)

[12] 赵明岩,吴顺海,李一欣,等. 基于改进 YOLOv5s 的黑皮鸡枞菌检测方法[J]. 农业工程学报, 2023, 39(12): 265-274.
ZHAO Mingyan, WU Shunhai, LI Yixin, et al. Improved YOLOv5s - based detection method for *Termitomyces albuminosus* [J]. Transactions of the CSAE, 2023, 39(12): 265-274. (in Chinese)

[13] 王磊磊,王斌,李东晓,等. 基于改进 YOLOv5 的菇房平菇目标检测与分类研究[J]. 农业工程学报, 2023, 39(17):

163 – 171.

WANG Leilei, WANG Bin, LI Dongxiao, et al. Object detection and classification of *Pleurotus ostreatus* using improved YOLOv5[J]. Transactions of the CSAE, 2023, 39(17): 163 – 171. (in Chinese)

[14] 张博洋. 工业视觉中目标检测与实例分割技术的研究与应用[D]. 济南: 山东大学, 2023.

ZHANG Boyang. Research and application of object detection and instance segmentation technology in industrial vision[D]. Ji'nan: Shandong University, 2023. (in Chinese)

[15] 夏晨斐. 密集场景下的精细行人实例分割方法研究[D]. 武汉: 华中科技大学, 2022.

XIA Chenfei. Research on fine pedestrian instance segmentation in crowded scenes[D]. Wuhan: Huazhong University of Science and Technology, 2022. (in Chinese)

[16] 卢伟, 邹明萱, 施浩楠, 等. 基于 YOLO v5-TL 的褐菇采摘视觉识别-测量-定位技术[J]. 农业机械学报, 2022, 53(11): 341 – 348.

LU Wei, ZOU Mingxuan, SHI Haonan, et al. Technology of visual identification measuring location for brown mushroom picking based on YOLO v5-TL[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(11): 341 – 348. (in Chinese)

[17] 王玲, 徐伟, 等. 基于 SR300 深度相机的褐蘑菇原位测量技术[J]. 农业机械学报, 2018, 49(12): 13 – 19, 108.

WANG Ling, XU Wei, et al. Portabella mushrooms measurement in situ based on SR300 depth camera[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(12): 13 – 19, 108. (in Chinese)

[18] 陈燕, 陆嘉豪, 胡小春, 等. 基于 VGG-UNet 的食用菌菌丝体表型参数自动测量方法[J]. 农业机械学报, 2023, 55(1): 233 – 240.

CHEN Yan, LU Jiahao, HU Xiaochun, et al. Automated measurement method of phenotypic parameters of edible mushroom mycelium based on VGG-UNet[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 55(1): 233 – 240. (in Chinese)

[19] 周馨翌, 吴烽云, 邹湘军, 等. 基于多目标识别的葡萄果串采摘点定位方法[J]. 农业工程学报, 2023, 39(22): 166 – 177.

ZHOU Xinzhaoh, WU Fengyun, ZOU Xiangjun, et al. Grape cluster picking point location method based on multi-object recognition[J]. Transactions of the CSAE, 2023, 39(22): 166 – 177. (in Chinese)

[20] 王映龙, 肖欢, 殷华, 等. 基于 RGBD 相机的黑皮鸡枞菌子实体形态视觉测量[J]. 农业工程学报, 2022, 38(20): 140 – 148.

WANG Yinglong, XIAO Huan, YIN Hua, et al. Measurement of morphology of *Oudemansiella raphanipies* based on RGBD camera[J]. Transactions of the CSAE, 2022, 38(20): 140 – 148. (in Chinese)

[21] 周华茂, 王婧, 殷华, 等. 基于改进 Mask R-CNN 模型的秀珍菇表型参数自动测量方法[J]. 智慧农业(中英文), 2023, 5(4): 117 – 126.

ZHOU Huamao, WANG Jing, YIN Hua, et al. Phenotypic analysis geesteranus based on improved Mask R-CNN[J]. Smart Agriculture, 2023, 5(4): 117 – 126. (in Chinese)

[22] 杜文圣, 王春颖, 朱衍俊, 等. 采用改进 Mask R-CNN 算法定位鲜食葡萄疏花夹持点[J]. 农业工程学报, 2022, 38(1): 169 – 177.

DU Wensheng, WANG Chunying, ZHU Yanjun, et al. Fruit stem clamping points location for table grape thinning using improved Mask R-CNN[J]. Transactions of the CSAE, 2022, 38(1): 169 – 177. (in Chinese)

[23] FU L, YANG Z, WU F, et al. YOLO-Banana: a lightweight neural network for rapid detection of banana bunches and stalks in the natural environment[J]. Agronomy, 2022, 12(2): 391.

[24] CHENG T, SONG L, GE Y, et al. Yolo-world: real-time open-vocabulary object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16901 – 16911.

[25] SU P, HAN H, LIU M, et al. MOD-YOLO: rethinking the YOLO architecture at the level of feature information and applying it to crack detection[J]. Expert Systems with Applications, 2024, 237: 121346.

[26] MIDHUN P. MATHEW, THERESE Y M. Leaf-based disease detection in bell pepper plant using YOLO v5[J]. Signal, Image and Video Processing, 2022, 16(3): 841 – 847.

[27] GAI R, CHEN N, YUAN H. A detection algorithm for cherry fruits based on the improved YOLO-v4 model[J]. Neural Computing and Applications, 2023, 35(19): 13895 – 13906.

[28] 肖珂, 梁聪哲, 夏伟光. 基于改进 YOLACT 的果树叶墙区域实时检测方法[J]. 农业机械学报, 2023, 54(4): 276 – 284.

XIAO Ke, LIANG Congzhe, XIA Weiguang. Real-time detection dethod of fruit leaf wall area based on improved YOLACT [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(4): 276 – 284. (in Chinese)

[29] 朱德利, 余茂生, 梁明飞. 基于 SwinT-YOLACT 的玉米果穗实时实例分割[J]. 农业工程学报, 2023, 39(14): 164 – 172.

ZHU Deli, YU Maosheng, LIANG Mingfei. Real time instance segmentation of maize ears using SwinT-YOLACT[J]. Transactions of the CSAE, 2023, 39(14): 164 – 172. (in Chinese)

[30] LIU Z, LIN Y, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 10012 – 10022.

[31] PARK J, WOO S, LEE J Y, et al. A simple and light-weight attention module for convolutional neural networks[J]. International Journal of Computer Vision, 2020, 128(4): 783 – 798.

[32] NIU Z, ZHONG G, YU H. A review on the attention mechanism of deep learning[J]. Neurocomputing, 2021, 452: 48 – 62.

[33] LIU Y, SHAO Z, TENG Y, et al. NAM: normalization-based attention module[J]. arXiv Preprint, arXiv:2111.12419, 2021.