doi:10.6041/j.issn.1000-1298.2026.13.016

西兰花选择性采收机器人成熟度识别与自主采收方法研究

康　朔[1]　李东方[2]　龙思放[1]　张　静[3]　郑成宇[4]　奚　特[1]　王　俊[1]

(1. 浙江大学生物系统工程与食品科学学院, 杭州 310058; 2. 南京农业大学工学院, 南京 211800;
3. 河南农业大学机电工程学院, 郑州 450002; 4. 吉林大学生物与农业工程学院, 长春 130022)

摘要: 西兰花选择性采收的整个过程需要精确的成熟度识别、采摘姿态调整以及快速的作物行导航,但目前缺少一种完成全流程作业的视觉算法。为此,本研究提出了高效采收视觉系统 EH-YOLO,旨在实现复杂农业环境下的准确成熟度识别和自主采摘,提高不同任务之间的协同性。本研究首先在 YOLO v8n 主干网络中加入了高效多尺度注意力(Efficient multi-scale attention)模块,并设计了空间金字塔池化(Soft spatial pyramid pooling-fast)模块,其次采用基于分组混合卷积(Grouped shuffle convolution)的轻量化颈部结构,改进了损失函数,以增强特征提取和融合能力、降低计算复杂度,并提高模型在复杂环境中的性能。最后,通过分析作物的农艺特征,利用西兰花花球的中心点进行植株定位,根据西兰花所处的象限调整采收姿态,并利用聚类算法和最小二乘法提取作物行导航线。模型内存占用量为 3.30 MB,检测速度达到 61.45 f/s,对成熟西兰花的成熟度识别准确率达到 92.84%,在精度和速度方面均超越了现有方法,并通过消融试验验证了改进模块的有效性。此外,与基线模型相比,EH-YOLO 的计算参数量减少了 38.68%,模型内存占用量减少了 46.77%,平均准确率提高了 4.01 百分点,在田间试验中成功进行了全程自主作业,运动和定位成功率为 96.77%,选择性采收成功率为 87.37%。因此,EH-YOLO 是一种实用且高效的西兰花选择性采收机器人视觉系统,对西兰花农田环境具有较好的适用性。

关键词: 选择性采收机器人; 西兰花; 成熟度识别; YOLO

中图分类号: S225.92　　**文献标识码**: A　　**文章编号**: 1000-1298(2026)13-0200-11　　**OSID**:

Maturity Recognition and Autonomous Harvesting Method of Selective Broccoli Harvesting Robot

Kang Shuo[1]　Li Dongfang[2]　Long Sifang[1]　Zhang Jing[3]　Zheng Chengyu[4]　Xi Te[1]　Wang Jun[1]

(1. *College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou* 310058, *China*
2. *College of Engineering, Nanjing Agricultural University, Nanjing* 211800, *China*
3. *College of Mechanical and Electrical Engineering, Henan Agricultural University, Zhengzhou* 450002, *China*
4. *College of Biology and Agricultural Engineering, Jilin University, Changchun* 130022, *China*)

Abstract: The entire process of selective harvesting of broccoli requires precise maturity recognition, harvesting posture analysis, and rapid crop row navigation. However, at present, there is a lack of a visual algorithm suitable for the full-process operation. To address these issues, an efficient harvesting visual system, EH-YOLO, was proposed for the selective harvesting robot of broccoli, aiming to achieve accurate maturity recognition and autonomous harvesting in complex agricultural environments. The efficient multi-scale attention module was incorporated into the YOLO v8n backbone network, and a soft spatial pyramid pooling-fast module was designed. Additionally, a lightweight neck structure based on grouped shuffle convolution was employed, and the loss function was improved to enhance feature extraction and fusion capabilities, reduce computational complexity, and improve model performance in complex environments. Furthermore, by analyzing the crop's agronomic features, the plant location was determined based on the center point of the broccoli head. The harvesting process was adjusted according

收稿日期: 2026-02-05　修回日期: 2026-02-28

基金项目: 浙江省"三农九方"农业科技合作项目(2023SNJF047)

作者简介: 康朔(1998—),男,博士生,主要从事果蔬选择性采收机器人研究,E-mail: kangshuo@sdau.edu.cn

通信作者: 王俊(1965—),男,教授,博士生导师,主要从事农业智能装备研究,E-mail: jwang@zju.edu.cn

to the quadrant in which the broccoli was located, and clustering algorithms, combined with the least squares method, were used to extract the crop row navigation lines. The model size was 3.3 MB, and the detection speed reached 61.45 f/s. The accuracy of maturity recognition for mature broccoli was 92.84%. Additionally, compared with the baseline model, EH - YOLO reduced the computational parameters by 38.68%, decreased the model size by 46.77%, and increased the average accuracy by 4.01 percentage points. In field trials, the system successfully completed the full-process selective harvesting task, with a motion and positioning success rate of 96.77% and a selective harvesting success rate of 87.37%. Therefore, EH - YOLO was a practical and efficient visual system for the selective harvesting robot of broccoli, with strong applicability for commercial broccoli fields.

Key words: selective harvesting robot; broccoli; maturity recognition; YOLO

0　引言

西兰花是一种具有丰富营养价值的重要蔬菜作物[1]，广泛种植于世界各地。根据联合国粮食及农业组织（Food and Agriculture Organization of the United Nations, FAO）统计，全球西兰花和花椰菜的总产量约 2.59×10^9 t，种植面积 1.38×10^6 hm^2。由于西兰花不同步成熟的农艺特性，需要采用选择性收获策略[2]，存在用工成本高、劳动强度大、劳动力短缺等问题[3-4]。限制着西兰花产业的发展，迫切地需要一种选择性收获机器人。西兰花选择性采收机器人的关键技术之一是视觉系统，其核心在于对西兰花的成熟度识别与定位，作物分布的感知与采收姿态的调整，并引导车辆沿作物行行走，对应选择性采收全流程的 3 个重要组成环节[5]。

首先，准确实时的西兰花成熟度识别与定位是关键的上游任务。西兰花种植田内有多种成熟度的花球，成熟度是结合宏观和微观特征的一个综合指标，茎叶遮挡和田间变化的光线对成熟度识别造成影响，检测视野内的杂草及其他背景的干扰也容易对花球成熟度识别过程造成干扰。Blok 等[6]对 Laws 滤波增强后的图像进行阈值分割，可以实现西兰花检测，但存在召回率低、对阈值的依赖性强和局限在实验室环境的问题。Kusumam 等[7]使用 RGB-D 传感器采集三维点云信息，使用视点特征直方图、支持向量机分类器和时间滤波器检测西兰花花球，基于传统机器学习的方法通过提取大小、颜色和纹理特征，构建分类器检测模型，准确率为 89.85%。然而，传统的机器学习方法存在召回率低、依赖特定光照条件、泛化能力差以及实时性能不足等问题，难以满足实际生产需求[8]。

深度学习技术在近几年得到了快速发展[9-13]。针对实际工况下的西兰花成熟度识别研究比较缺乏，复杂农业环境中的成熟度识别仍面临诸多挑战，随着对西兰花选择性采收机器人视觉系统和成熟度识别算法的深入研究，具备高检测精度和快检测速度的轻量级模型是实现部署的关键。

当识别检测到成熟花球及其位置后，还需要控制末端执行器以合适的姿态执行采收动作。迄今为止，西兰花的选择性采收领域还缺乏相关研究[14-16]。当采收完当前区域的所有成熟西兰花后，机器人需要沿着作物行前进至下一工作区域，国内外学者开展了作物行视觉导航的相关研究。Diao 等[17]提出了一种基于改进 YOLO v8 的玉米喷雾机器人导航线提取的方法，旨在解决现有导航线提取算法适应性差和提取质量不佳的问题，这种方法提高了作物识别的鲁棒性和导航精度。然而，特征点提取过程没有充分考虑作物特征，模型相对复杂，导致实时性能不佳。为了在复杂的农田条件下提高作物行检测的准确性，Yang 等[18]采用 YOLO v5 对感兴趣区域内的作物行和背景进行分割，然后利用 FAST 角点检测来定位特征点，并拟合玉米作物行的直线。尽管这些方法都能有效地检测作物行特征点，但在检测框内的图像需要进行预处理才能定位特征点，这增加了特征点定位时间和作物行中心线识别所需的时间[19]。通过观察作物的农艺特性，发现西兰花花球的中心可以有效地定位西兰花植株。利用检测框的中心坐标，可以减少特征点定位时间，还可以实现作物行导航线的提取，从而降低硬件成本和计算能力要求。

综上，现有研究中的算法功能单一，只能实现全流程作业中的一个环节。机器人在田间应用时需要同时处理多种图像数据和多种任务，而且边缘计算设备的算力有限，有必要设计一种易于部署应用的具有高检测精度、快速检测速度、鲁棒性好、可扩展性和共享性高的集成式多功能模型。本文提出一种改进 YOLO v8n 算法完成西兰花成熟度识别及选择性采收全流程的任务。

1　数据集建立与 EH-YOLO 模型构建

1.1　农艺特点与成熟度界定

成熟度识别是一个综合指标，既要结合花球

的宏观特征，又要结合花蕾的微观特征。采收人员在实际采收作业中，往往需要根据西兰花的大小、颜色、花蕾间距、紧实度、市场需求和商品价格等实际情况综合决定是否采收花球。本研究面向实际生产应用，将西兰花成熟度分为 3 类。如图 1 所示，未成熟的西兰花花球发育不完全，花球紧实，表面圆整，颜色深绿，有嫩小且整齐的花蕾颗粒。半成熟的西兰花是介于未成熟和成熟之间的状态，花球已有初步生长，花球紧实，表面圆整，颜色为蓝绿色。成熟西兰花头部完全发育，表面比较圆整，边缘尚未散开，花球相对紧实，颜色为绿色，花蕾整齐且尚未开放[20]。收获过早，花球未充分发育长大、产量不高；收获过迟，则花蕾松散，球面高低不平，花蕾变黄，鲜度和品质下降，失去商品价值。因此，西兰花的分批次成熟特性决定了需要采用选择性收获的策略[21]。



(a) 未成熟 (b) 半成熟 (c) 成熟

图 1 成熟度示意图

Fig. 1 Diagrams of broccoli maturity stages

1.2 图像获取与数据集建立

图 2 是目标检测数据集建立过程示意图，在浙江省宁波市的试验田采集自然收获条件下的西兰花图像。图像采集过程由自主研发的西兰花选择性采收机器人完成，采集视角为俯拍，检测视野中包含一垄两行的多个西兰花花球，这样更加贴近田间西兰花选择性采收机械的实际作业过程，使研究更面向生产应用。图像采集涵盖了不同光照条件，而且图像采集时间与人工采收时间保持一致。当深度相机安装在机械臂末端时，相机距离地面高度为 100 cm，接近弯腰状态下的人体高度；当深度相机安装在移动底盘上时，相机距离地面的高度为 170 cm，与成年人平均身高一致。如图 2 中 5 组图像所示，采集了 5 种具有代表性的西兰花成熟度识别场景，具有不同高度、不同的光线条件（晴天、阴天、多云）、茎叶遮挡和杂草等干扰。试验品种为台绿和浙青，本研究根据成熟过程中的变化将西兰花的成熟度分为 3 类：未成熟、半成熟和成熟。并由经验丰富的农艺师使用 LabelMe 软件对 1 000 幅不同环境下的西兰花成熟图像进行了标注。

1.3 基于视觉信息的成熟度识别

本研究在 YOLO v8n 的基础上改进，提高特征提取能力，采用多种优化策略来实现速度和准确性

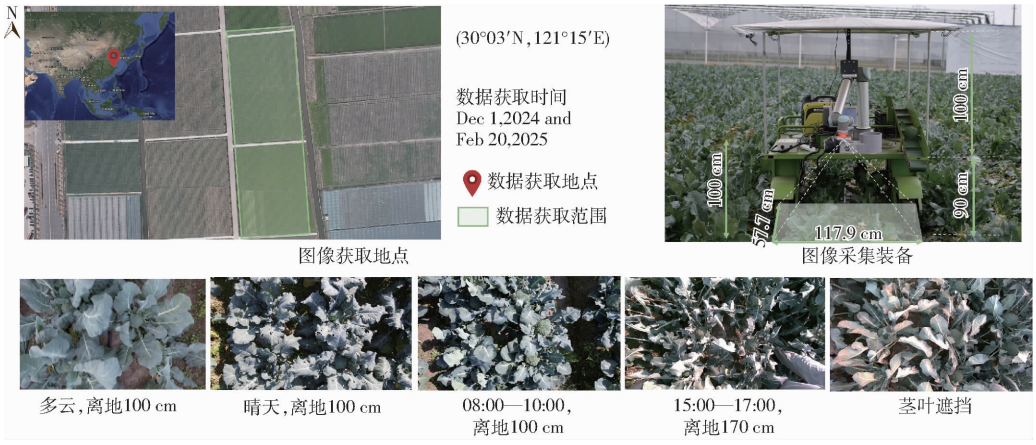


图 2 图像采集过程示意图

Fig. 2 Schematic of image acquisition process

之间的平衡。如图 3 所示，高效收获 YOLO 模型（Efficient harvesting - You only look once，EH - YOLO）分为主干特征提取网络、颈部网络、头部网络和后处理模块。

1.3.1 高效多尺度注意力机制

由于检测视野内存在花球、茎叶、杂草等，这些因素会影响成熟度识别精度。其次，不同成熟度花球之间的颜色、纹理、尺寸存在较小的特征差异，而且茎叶可能会遮挡花球。为增强模型在通道特征捕捉上的能力，并提高模型对不同通道特征的理解能力，更好地区分目标与背景，以及不同成熟度花球之间的特征差异，本文引入高效多尺度注意力（Efficient multi-scale attention，EMA）[22]，结构如图 4 所示，该模块根据通道数对输入特征图进行分组，充分挖掘通道内的特征信息。同时，EMA 模块采用两种不同尺寸的卷积核作为并行子网络，以融合西兰花的多尺度特征信息。两个并行子网络还分别执行全局平均池化，从而有效地编码全局信息并捕捉像素级的成对关系。这一操作建立了通道和空间特征之间的长程依赖，强化了特征图兴趣区域的

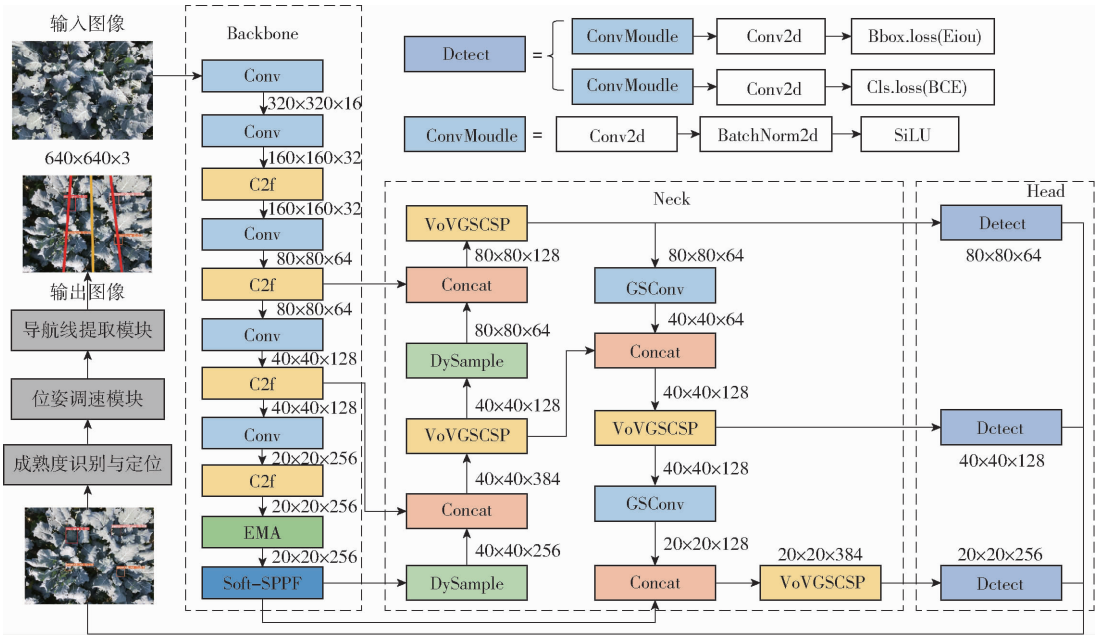


图 3　EH－YOLO 模型结构图

Fig. 3　Diagram of model structure of EH－YOLO

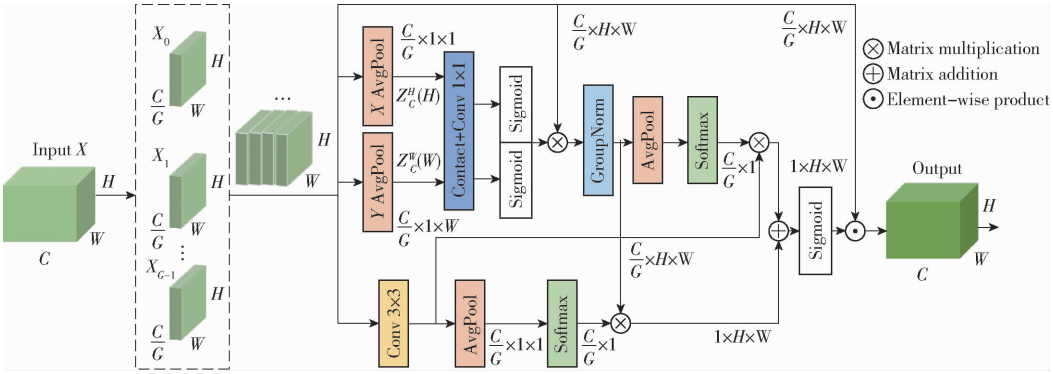


图 4　EMA 注意机制模块图

Fig. 4　EMA attention mechanism module diagram

特征表示。

首先，EMA 模块以主干网络提取的特征图 $X \in \mathbf{R}^{C\times H\times W}$ 为输入，并沿通道维度方向将特征图划分为 G 个组：$X=[X_0, X_1, \cdots, X_i, \cdots, X_{G-1}]$，其中子特征图 $X_i \in \mathbf{R}^{\frac{C}{G}\times H\times W}$，$H$ 为特征图在垂直方向上的维度，W 为水平方向上的维度，C 为通道数量。随后，子特征图被送入两个并行的子网络中，一条对水平方向的特征进行一维全局平均池化，以提取全局上下文信息；另一条路径对垂直方向的特征信息进行一维全局平均池化，以捕捉不同空间维度上的长程依赖性和位置信息，计算公式为

$$z_C^H(H)=\frac{1}{W}\sum_{0\leqslant i\leqslant W} x_C(H,i) \tag{1}$$

$$z_C^W(W)=\frac{1}{H}\sum_{0\leqslant j\leqslant H} x_C(j,W) \tag{2}$$

然后将特征图 $z_C^H(H)$ 和 $z_C^W(W)$ 在垂直方向上进行拼接。使用 1×1 卷积核建立通道和空间位置信息之间的关联，有助于分解表示不同空间方向上的特征图。最终，特征图经过 Sigmoid 激活函数处理，生成注意力权重图。通过矩阵乘法使用水平和垂直维度的注意力权重图与子特征图对其进行更新。中间特征图依次经过组归一化、平均池化和 Softmax 操作。前者实现了全局空间信息和通道信息的融合，从而生成 $\frac{C}{G}\times1\times1$ 的特征图，融合过程计算公式为

$$z_C=\frac{1}{H\times W}\sum_j^H\sum_i^W x_C(i,j) \tag{3}$$

其中，$x_C(i,j)$ 表示第 C 通道在空间维度位置 (i,j) 处的输入特征，z_C 为第 C 通道经过二维全局平均池化后的特征。经过 Softmax 操作后得到的特征图与通过 3×3 卷积核处理过的子特征图进行矩阵乘法，从而生成维度为 $1\times H\times W$ 的空间注意力权重图。另外一个 3×3 子支的输出与经过组归一化处理后的 1×1 分支特征图进行矩阵乘法，得到维度为 $1\times H\times W$ 的

空间注意力权重图,有助于促进空间维度之间信息的交互。最后,将两个空间注意力权重图相加,并通过 Sigmoid 函数和元素积,得到最终的注意力权重图。随后,原始特征图通过最终的注意力权重图加权,网络输出一个强调兴趣区域的特征图。

1.3.2 快速空间金字塔软池化模块

YOLO v8 中使用快速空间金字塔池化(Spatial pyramid pooling-fast, SPPF),通过堆叠不同尺度的池化操作,将特征直接硬性拼接,可能导致边界信息丢失或融合不平滑,而且固定核大小的池化可能导致边界信息丢失。此外,SPPF 在复杂的检测场景(背景混杂、目标边界模糊)下可能不够鲁棒。因此本研究引入了 SoftPool[23],提出了快速空间金字塔软池化模块(Soft spatial pyramid pooling-fast, Soft-SPPF),该模块对多尺度特征的感知更精细,特征融合更平滑自然,减少了信息损失。通过不同尺度的加权池化操作,聚合局部和全局信息,更好地处理背景复杂性和目标模糊性问题,从而提高目标检测的精度。SoftPool 是一种加权池化方法,研究表明[24],相比其他池化方法,它能更有效地保留更多的细节特征,为特征融合提供更有价值的信息。

本研究将 SPPF 中的 Maxpool 层替换为 SoftPool,从而得到 Soft-SPPF 模块,结合了 SPPF 和 SoftPool 的优点,并改善了细节特征,改进了特征融合方式,具备了更灵活和更高效的多尺度特征聚合能力。Soft-SPPF 模块如图 5 所示,首先,输入特征图经过 1×1 卷积操作(包括标准卷积、批归一化和激活函数),接着应用多个 5×5 大小的 MaxPool 层来提取多尺度信息。不同池化层得到的特征与经过初步 1×1 卷积的特征进行拼接。最后,这些拼接后的特征通过另一层 1×1 卷积进行融合。

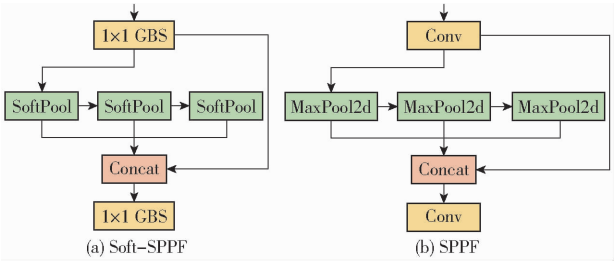


图 5 Soft-SPPF 模块示意图

Fig. 5 Soft-SPPF module diagrams

1.3.3 基于分组混合卷积的轻量化颈部结构

在西兰花选择性采收作业中,检测速度是成熟度识别研究中的另一个关键问题。为了开发出兼具高检测精度和快速检测速度的轻量化模型,并最终实现其在计算能力有限的边缘设备上的部署,本研究使用分组混合卷积(Grouped shuffle convolution, GSConv)模块[25]重构了 YOLO v8 模型的颈部网络,保持模型准确性的同时减少了模型复杂度,GSConv 的结构图如图 6 所示。

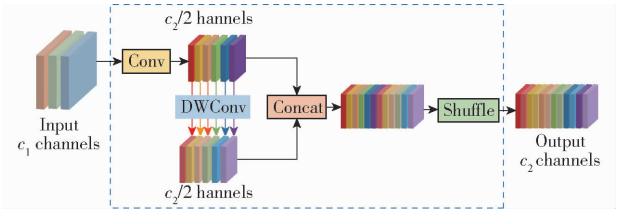


图 6 GSConv 结构图

Fig. 6 Structure diagram of GSConv

首先,通过普通卷积操作对输入特征图进行下采样,然后进行深度可分离卷积(Depthwise convolution, DWConv),并将两次卷积操作的输出进行拼接。最后,执行 Shuffle 操作,将前两次卷积操作中具有紧密语义相关的信息通道重新组合。GSConv 通过 Shuffle 操作,将标准卷积与 DWC 卷积的生成信息进行融合,使深度可分离卷积的输出尽量接近标准卷积,从而更好地平衡模型的精度和速度。

接着,基于 GSConv 引入了新的 GSbottleneck 模块,并使用该 GSbottleneck 模块重构了 YOLO v8 模型的 C2f 模块。所提出的模块命名为 VoVGSCSP,如图 7 所示。通过 GSConv 模块与 VoVGSCSP 轻量化结构的合理结合,能够实现轻量化的同时保证精度不损失,该组合结构被命名为 Slim-neck 网络。本研究中,Slim-neck 被用于重构 YOLO v8 的 neck 网络,一方面提高了 YOLO v8 模型的实时性,另一方面保持了对西兰花花球的成熟度识别精度。

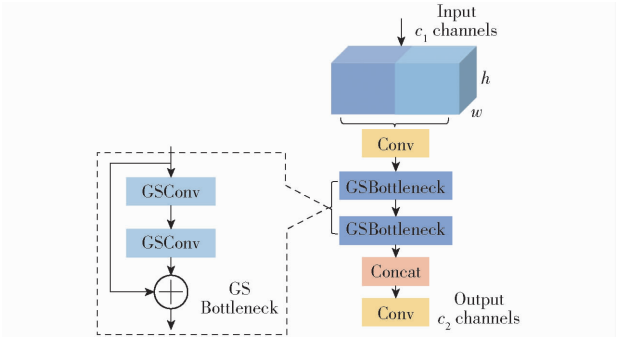


图 7 VoVGSCSP 结构图

Fig. 7 Structure diagram of VoVGSCSP

1.3.4 高效交并比损失函数

西兰花按照标准行间距种植时,沿茎秆往四周茂盛生长的枝叶会部分遮挡周围的花球,所以在西兰花成熟度的实际识别定位中,茎叶遮挡对检测性能有很大影响。因此,本研究引入了高效 IoU 损失(Efficient IoU loss, EIoU)[26]来替代原本的 CIoU 损失(Complete IoU loss, CIoU),提高边界框回归的准

确性和模型的识别定位性能。EIoU 可以在不增加网络复杂度的情况下提高物体检测的定位精度，也不会影响推理速度。EIoU 损失向 IoU 损失中增加了两个惩罚项，第 1 个惩罚项旨在最小化真实框 B^{gt} 的中心点与预测框 B^{p} 中心点之间的欧几里得距离；第 2 个惩罚项则着重于减少真实框和预测框在宽度和高度上的差异。EIoU 损失定义为

$$L_{\mathrm{EIoU}} = 1 - IoU + \frac{d^2(b^{gt}, b^{p})}{(w^{c})^2 + (h^{c})^2} + \frac{d^2(w^{p}, w^{gt})}{(w^{c})^2} + \frac{d^2(h^{p}, h^{gt})}{(h^{c})^2} \tag{4}$$

其中，$d(b^{gt}, b^{p})$ 表示真实框 B^{gt} 和预测框 B^{p} 中心点之间的欧几里得距离；w^{c} 和 h^{c} 表示覆盖 B^{gt} 和 B^{p} 的最小外接框 c 的宽度和高度；$d(w^{p}, w^{gt})$ 和 $d(h^{p}, h^{gt})$ 分别表示真实框和预测框在宽度和高度上的差异，如图 8 所示。

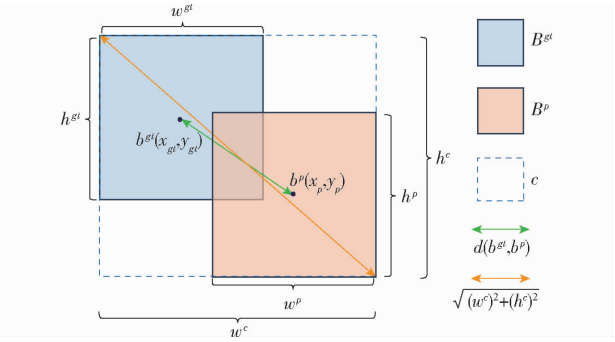


图 8　真值边界框与预测边界框的偏差

Fig. 8　Deviation between ground truth bounding box and predicted bounding box

1.4　基于花球识别信息的自主采收流程设计

1.4.1　末端执行器姿态调整

为了避免在采收过程中末端执行器对附近花球造成压伤和损坏，所以设计了采收姿态的调整。首先，提取每个检测框的中心对于一垄两行种植模式下的西兰花，检测相机俯视的视野如图 9 所示，一般存在 4 个西兰花，标准化种植时的行间距为 45 cm × 45 cm。因为直线往复切割式末端执行器存在一定的体积，其长 × 宽 × 高为 50 cm × 18 cm × 20 cm。

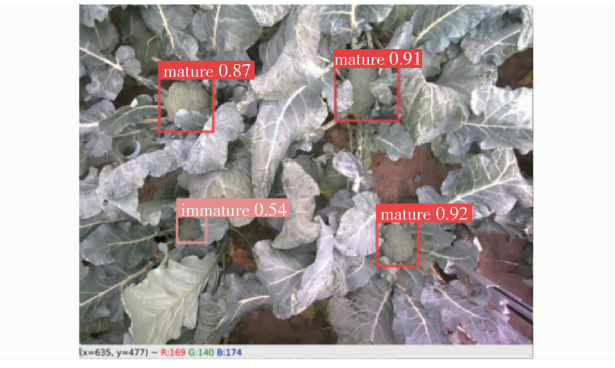


图 9　西兰花的生长和分布

Fig. 9　Growth and distribution of broccoli plants

西兰花植株从地面直立生长，很少会出现倾斜的情况，便于垂直地从上往下收割。值得注意的是，即使根茎部分是倾斜的，花球及球茎部分也会直立生长，这可能是作物本身的生长特性决定的。所以，本研究不考虑直线往复切割式末端执行器滚转角和俯仰角的变化，即 X 轴和 Y 轴不转动。偏航角根据待采收西兰花的情况进行变化，即可以沿 Z 轴转动变化调整末端执行器在水平面的位置。

结合实际检测过程中的西兰花分布特点，将视野划分为 4 个象限，并对处于不同象限的西兰花设置不同的采收姿态，在总控制程序中预设了 orientation(x,y,z,w)。西兰花图像传入 EH－YOLO 网络进行成熟度识别和定位，可以获取所有花球的成熟度类别和坐标(x,y)，然后根据图像宽度、高度判定待采收花球所处象限，实现流程如图 10 所示。

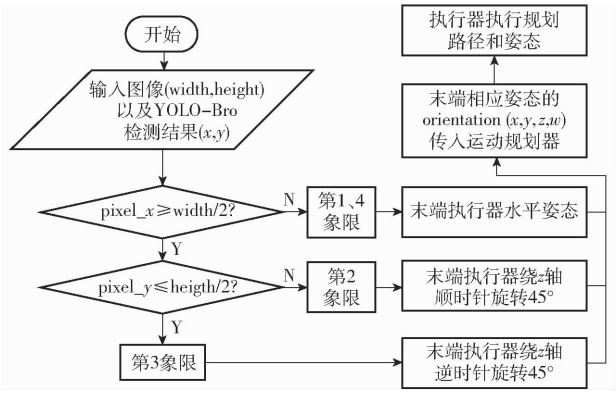


图 10　末端执行器姿态调整流程图

Fig. 10　Implementation of code for end-effector posture adjustment

对于第 1、4 象限的西兰花，其末端姿态为识别位置时的水平姿态；对于第 2 象限的西兰花，其末端姿态为水平绕 z 轴顺时针旋转 45°；对于第 3 象限的西兰花，其末端姿态为水平绕 z 轴逆时针旋转 45°。这样可以很好地避免末端执行器固有体积对西兰花损伤的风险。

1.4.2　作物行导航中心线提取

(1)提取每个检测框的中心点：EH－YOLO 预测的每个矩形框都有一个中心点坐标(c_x, c_y)，可以根据检测框的左上角和右下角坐标来计算中心点坐标。图 11 红色点代表检测到的西兰花花球中心顶点。

(2)中心点聚类：根据空间分布，使用 K-means 聚类算法将相邻的中心点分为一类，得到一个中心点坐标集合 $C = \{(c_{xi}, c_{yi}) \mid i = 1, 2, \cdots, n\}$。这样做的目的是将属于同一导航路径的西兰花分为一簇，避免其他行的西兰花干扰路径的生成。

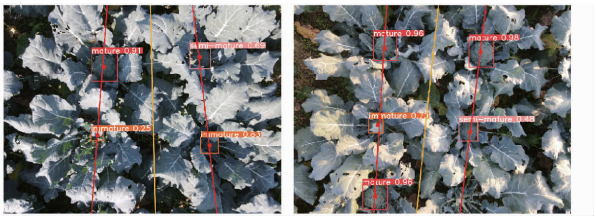


图 11 导航线提取示意图

Fig. 11 Diagram of navigation line extraction

(3)拟合导航路径:对聚类后的中心点坐标使用最小二乘法进行多项式拟合,以获得平滑的导航路径。图 11 中红色线是通过这些点拟合得到的导航辅助线。作物行的两条边界线可作为独立的检测区域,根据农艺学原理,两条边界线之间的中线可作为导航路径。因此,检测到的两条作物行边界线的平分线即为机器人可用的导航中心线,如图 11 所示,黄色线是导航路径。

1.4.3 基于 PID 控制算法的机器人路径跟踪

路径跟踪控制是采收机器人自动导航中的重要环节,其目的是确保机器人能够沿识别的作物行中心线行驶。通过调整控制输入,路径跟踪控制可以精确地执行任务,并保证系统的稳定性与精度。路径跟踪控制分为横向和纵向两个方面,横向控制用于调整电动底盘的转向,确保其沿着规划路径行驶;纵向控制则负责调节底盘的速度和加速度,由于机器人的前进速度较低且西兰花的种植行曲率较小,纵向控制无需过多调整。比例-积分-微分控制(Proportional - integral - derivative,PID)是一种广泛应用的经典反馈控制策略,常用于路径跟踪中的角度调节。PID 控制由比例、积分和微分组成,三者协同作用以减小系统误差,提升稳定性与控制精度。

使用 PID 算法进行西兰花选择性采收机器人的转向控制时,以电动底盘与所识别出的作物行预瞄点间的横向偏差作为输入,以电动底盘的转向角作为输出。转向角 $\delta(t)$ 的计算公式为

$$\delta(t) = k_p e(t) + k_i \sum_{i=0}^{k} e(t) + k_d(e(t) - e(t-1)) \quad (5)$$

其中,k_p、k_i、k_d 分别为 PID 的比例系数、积分系数及微分系数,$e(t)$ 为 t 时间步所对应的横向偏差。如图 12 所示,由于相机的识别位置位于电动底盘前方中心位置处,因此横向偏差使用图像中作物行与图像中点间的横向偏差替代。

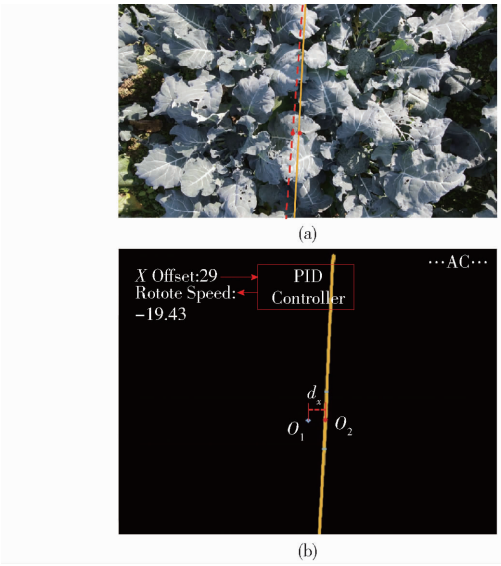


图 12 作物行中心线横向偏差的计算方法

Fig. 12 Calculation method for lateral deviation of centerline of crop rows

PID 参数整定是确保控制器性能的关键步骤,目标是找到合适的比例系数 k_p、积分系数 k_i 和微分系数 k_d,以实现系统的稳定性和快速响应。在试验过程中,采用手动试错法对 PID 参数进行调整,最终得到的参数值为 $k_p = 1.2$,$k_i = 0.8$,$k_d = 0.09$。

视觉导航系统和控制流程如图 13 所示。其中,信息获取模块负责图像信息采集及转角信息采集;导航线提取模块负责从图像中完成准确而稳定的导航线提取及跟踪;控制信号计算模块负责导航偏差的解算、视觉导航路径向空间导航路径的映射及控制信号的输出,通过 PID 算法实现对转向角度的闭环控制;车机通讯模块负责将控制命令传达至底盘电机。

1.5 评价指标

本研究采用 COCO(Common objects in context)数据集的评估指标,包括精确率(Precision)和召回率(Recall)。平均精度(Average precision, AP)反映了测试模型的整体性能,其中 AP_{50} 和 AP_{75} 提供

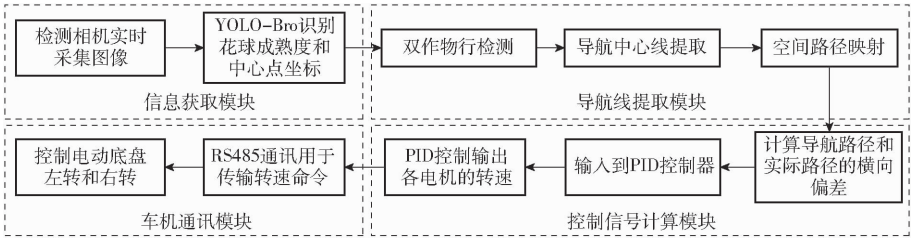


图 13 视觉导航系统流程图

Fig. 13 Flowchart of visual navigation system

了在特定掩码 IoU 阈值下对预测的更详细评估，从而全面评估算法的整体有效性。平均精度均值（mean Average Precision，mAP）表示所有类别的平均 AP。

除了模型的检测准确率指标外，模型内存占用量和帧率也常被用于评估模型的性能，反映了其计算复杂度和实时能力。收获作业中的关键是识别并收获成熟的花球。在田间试验中，提出了补充评估指标：成熟度识别准确率（Maturity recognition accuracy，MRA），机器人在摄像头视野范围内识别出的成熟西兰花球的数量与实际的成熟西兰花球数量的百分比。第 2 个指标是运动与定位成功率（Motion and position success rate，MPSR），末端执行器定位准确到达西兰花顶部的数量与机器人在摄像头视野范围内识别出的成熟西兰花数量的百分比。

2　结果与讨论

2.1　模型性能比较

本研究选取经典先进的 YOLO v5s、YOLO v6n 和 YOLO v8n 算法作为对比，图 14 是各目标检测模型的成熟度识别效果比较图。如图 14a 所示，在一般情况下，4 个算法都有较好的表现，不存在漏检和错误检测情况。但是当处理花球差异不明显，光线变化和茎叶遮挡等复杂问题时，算法的识别精度受到影响下降。如图 14b 所示，YOLO v5s 和 YOLO v6n 没有识别到右上角的西兰花，而且除了 EH－YOLO 准确预测视野中左侧半成熟西兰花之外，其余 3 个算法均错误预测为未成熟西兰花，这可能是 EMA 机制和 Soft－SPFF 改进模块提高了网络的特征提取能力，可以更好地区分不同成熟度花球之间的细微差异。

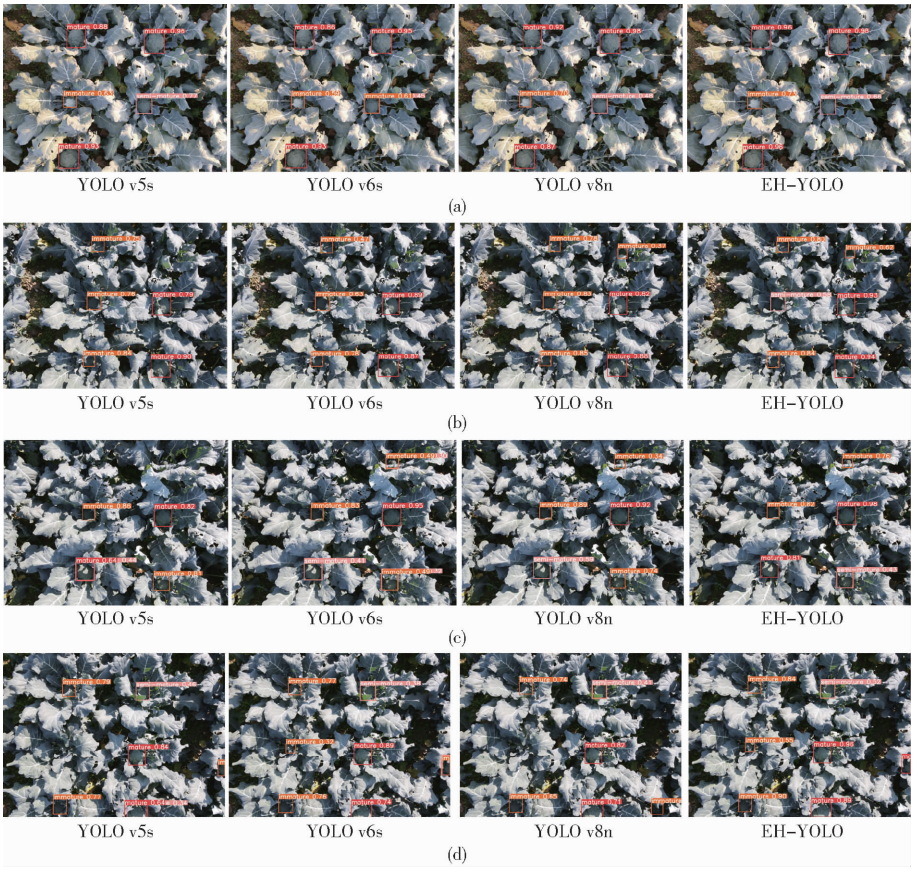


图 14　不同目标检测模型的成熟度识别性能比较

Fig. 14　Comparison of maturity recognition performance across different object detection models

田间存在阴天弱光、晴天强光及云影变化等情况，容易造成花球表面高光或阴影并引发误判。图 14c 中左下角花球受阴影影响时，部分基线模型将成熟误判为半成熟或产生漏检，而 EH－YOLO 仍能正确识别。Soft－SPPF 通过软加权池化在多尺度聚合时更好地保留细节纹理，配合 EMA 对有效区域的增强，有助于在光照不均匀条件下维持稳定特征表示。茎叶遮挡会导致目标可见区域缩小、边界模糊，易产生漏检或类别混淆，在图 14d 的典型遮挡场景中，YOLO v5s 和 YOLO v8n 出现漏检，而 EH－YOLO 仍能稳定检测目标，体现出更强的鲁棒性。其原因在于：EMA 模块增强了对花球关键纹理与颜色差异的关注能力，抑制了叶片等背景干扰；同时，EIoU 损失对边界框中心点与宽高差异的显式

约束使遮挡样本的定位收敛更稳定,从而降低漏检风险。

与其余算法相比,EH - YOLO 具有更高的成熟度识别准确率、查全率,而且识别花球成熟度的置信度是最高的,表明本研究所提出模型的先进性和有效性。需要指出的是,漏检情况会影响导航线的提取,进而影响对行行走效果,存在损伤植株的风险。错误检测情况会影响西兰花选择性采收机器人选择作业对象,进而影响经济效益。

如表 1 所示,所提出模型的精确率、召回率和平均精度分别比基线提高 4.3、5.1、4.01 个百分点,均超过了轻量化模型 YOLO v5s、YOLO v6n。而且 EH -YOLO 在保持模型轻量化特性的同时,对于成熟花球的识别准确率为 92.84%,分别比其他模型高 5.17、10.12、4.71 个百分点。EH - YOLO 在保持较高精度的同时实现了 61.45 f/s 的推理速度,模型内存占用量约 3.30 MB。并在端到端流程中共享检测结果用于姿态调整与作物行导航:基于检测框中心点即可完成植株定位,避免了传统方法中的图像预处理与特征点提取步骤,从而进一步降低计算开销。

表 1 不同目标检测模型的性能比较

Tab. 1 Performance comparison of different object detection models

模型	精确率/%	召回率/%	帧率/(f·s^{-1})	mAP_{50}/%	mAP/%		
					未成熟	半成熟	成熟
YOLO v5s	70.5	75.7	11.26	79.80	84.93	66.92	87.67
YOLO v6n	64.4	71.1	29.22	76.56	81.55	65.41	82.72
YOLO v8n	68.8	78.3	41.58	80.72	86.70	67.33	88.13
EH - YOLO	73.1	83.4	61.45	84.73	89.28	72.08	92.84

2.2 改进模块的消融试验

开展消融试验以评估改进模块的有效性及成熟度识别精度,每个试验在训练和测试阶段均采用相同的超参数设置,如表 2 所示,反映了各改进模块对鲁棒性与轻量化的不同贡献。从 mAP_{50} 指标来看,Soft - SPPF 提高了 1.11 个百分点,说明软池化的多尺度聚合更利于保留花球边缘与纹理细节;EMA 提高了 2.23 个百分点,表明在背景复杂、类别差异细微的成熟度识别任务中,跨空间与通道的注意力建模能够有效抑制茎叶和杂草干扰并强化花球特征;GSConv 主要贡献在于显著压缩模型尺寸并保持精度,适合资源受限的边缘部署场景。

进一步地,各模块组合呈现协同效应。当使用 Soft - SPPF 和 EMA 来集成和提取特征图时,mAP_{50} 相较于基线分别提升了 3.19 个百分点。这表明二者使得模型在复杂场景下对多尺度细节与全局上下文同时增强,使不同成熟度之间的细微差异更易被区分。当使用 Soft - SPPF 和 GSConv 组合时,识别准确度提升了 3.20 个百分点,可以动态调整采样位置,更准确地捕捉目标的细节特征,聚焦于最重要的特征或区域。EMA 与 GSConv 的组合也使 mAP_{50} 提升了 2.41 个百分点,增强了模型在通道特征捕捉上的能力,从而更好地区分目标与背景。这表明各模块之间存在有益的协同效应,每个模块在跨模态特征融合过程中均作为有效的融合机制,共同提升了模型的目标检测性能。

表 2 EH - YOLO 模型的消融试验结果

Tab. 2 Ablation study results of EH - YOLO model

Soft - SPPF	EMA	GSConv	精确率/%	mAP_{50}/%	模型内存占用量/MB
			68.80	80.72	6.20
√			69.95	81.83	5.11
	√		70.16	82.95	6.37
		√	69.73	81.34	4.32
√	√		71.84	83.91	5.28
	√	√	71.01	83.13	4.49
√		√	72.16	83.92	3.23
√	√	√	73.10	84.73	3.30

注:√表示使用该模型。

2.3 田间试验结果

西兰花选择性采收机器人的田间试验于 2024 年 12 月在浙江省宁波市康绿蔬菜合作社进行,种植的西兰花品种为“浙青 80”和“台绿”。试验地块随机选取 5 行长势一致的西兰花,各选取 30 m 作为试验区,田间设备的配置如表 3 所示。

表 3 西兰花选择性收获机器人软硬件配置

Tab. 3 Hardware and software configuration of broccoli selective harvesting robot

配置	模块	型号
硬件配置	CPU	13th Intel(R) Core(TM) CPU i9 - 13900H@2.6 GHz, 16 GB
	GPU	NVIDIA GeForce RTX4070, 16 GB
	移动平台	Electric high-gap chassis
	深度相机	Intel RealSense D455
	末端执行器	Linear reciprocating cutting end effector
	机械臂	UR10
	系统	Ubuntu 18.04
软件配置	深度学习框架	Pytorch 1.12.1, Python 3.8.6
	GPU 并行计算框架	Cuda 11.8
	机器人操作系统	ROSmelodic, Gazebo 9, RViz 1.13

2.3.1 成熟度识别效果

采收作业中的图像处理工作是在西兰花选择性

采收机器人的中央处理器上进行，田间试验的西兰花成熟度识别与定位效果如图 15 所示，EH－YOLO 可以检测深度相机视野内的全部西兰花，没有漏检的情况，EH－YOLO 对花球成熟度识别准确率高，而且置信度高。漏检和错误检测会影响导航线的提取和作业对象的选取，进而影响对行行走效果和作业效果，影响经济效益。EH－YOLO 花球成熟度识别准确率高，而且置信度高。对于阴天弱光、晴天强光、多云变化光线以及茎叶遮挡造成的光线不均匀等复杂环境都有很好的适应性，而且预测锚框准确覆盖西兰花花球的所在区域，为后续的准确运动定位提供了保障。经过统计，在浙青和台绿两个品种的 103 个真实成熟西兰花中，成功识别 93 个，成熟类花球的识别准确率为 90.29%。

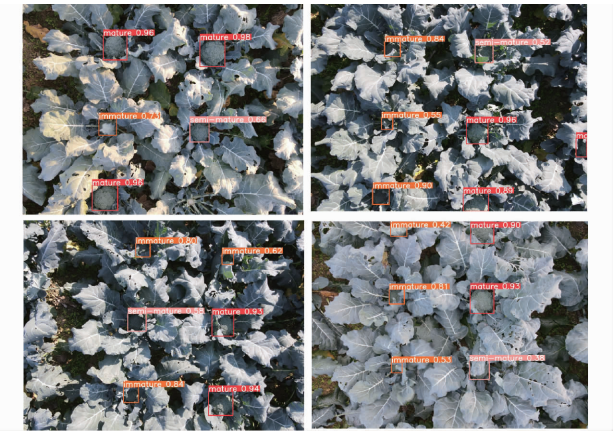


图 15　成熟度识别与定位效果

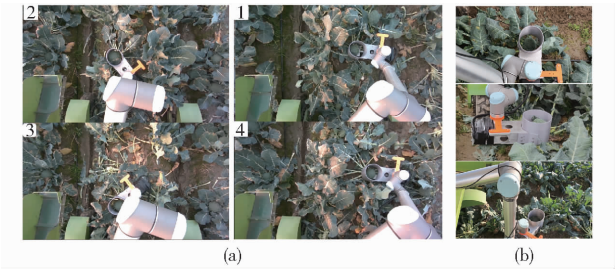
Fig. 15　Maturity recognition and localization effect

2.3.2　末端姿态调整效果

在成熟度识别信息的基础上可以实现末端姿态调整和作物行导航等功能，进而实现自主采收。如图 16a 所示，西兰花选择性采收机器人的末端执行器可以根据西兰花所处象限，及时调整为合适的采收姿态，从而避免对周围未采收西兰花的影响，可以很好地适应标准化种植的西兰花农田。如图 16b 所示，基于机器人操作系统（Robot operating system，ROS）可以控制机械臂从收集点准确到达切割点，经测量与统计，套筒中心和花球中心在平面上的误差为 ±1 cm，仿形套筒可以包络各个尺寸的成熟西兰花。

2.3.3　对行行走效果

如图 17 所示，当当前视野范围内的成熟西兰花采收完成后，选择性采收机器人可在 EH－YOLO 视觉系统的引导下沿作物行自主前进。机器人能够根据西兰花在行内的空间分布特征，动态调整行驶方向，实现连续稳定的对行行走。经过测量与统计，采收机器人实际行驶轨迹与作物行中心线之间的横向偏差控制在 3 cm 以内，满足西兰花种植行距条件下的安全作业要求，验证了视觉系统在行走引导任务中的有效性。



(a)　(b)

图 16　选择性采收机器人采收姿态调整与定位

Fig. 16　Selective harvesting robot's harvesting posture adjustment and location



图 17　选择性采收机器人对行行走效果

Fig. 17　Field navigation effect of selective harvesting robots

田间采收试验的数据统计如表 4 所示，部署了 EH－YOLO 视觉系统的机器人选择性采收成功率为 87.37%，识别准确率为 90.29%，运动定位成功率为 96.77%。田间试验中西兰花选择性采收机器人的采收效果如图 18 所示，可以完成对各种成熟西兰花花球的识别、定位、切割与收集。作业过程中系统运行平稳，证实了西兰花选择性采收机器人视觉系统的有效性和实用性，具有较高的可行性及商业价值。

表 4　田间试验结果

Tab. 4　Field experiment results of selective harvesting robots

品种	真实成熟花球数量	成熟度识别		运动定位	
		数量	准确率/%	数量	成功率/%
浙青	48	43	89.58	42	97.67
台绿	55	50	90.91	48	96.00
总计	103	93	90.29	90	96.77



图 18　西兰花选择性采收机器人的采收效果

Fig. 18　Harvesting effect of selective broccoli harvesting robot

3 结论

(1)基于 YOLO v8 设计一种高效的采收视觉模型 EH - YOLO,其具有检测精度高、检测速度快、共享性高的特点,采用模块化的集成控制策略,避免了单独使用成熟度识别、姿态调整和导航功能所带来的计算负担。提高了成熟度识别精度和不同任务间的关联性和协同性,进而在复杂农田环境中实现选择性采收全流程作业。

(2)通过多种针对性的改进模块提高成熟度识别的准确率,实现模型的轻量化,提高边界框定位的准确性,为后续步骤打下了基础。在对比试验中,模型检测速度为 61.45 f/s,对成熟花球的识别准确率为 92.84%,相较于其他模型具有更高的检测速度和精度。结合作物特点设计了导航线提取和末端姿态调整方法,基于花球圆心的定位方法无需进行之前的图像预处理,大大减少了特征点的定位时间。

(3)本研究为选择性采收机器人的视觉系统设计提供了一种新范式,强调图像信息的共享和多功能协同,提高彼此之间关联,降低硬件成本。经过田间试验,本文所提出的视觉算法能够满足西兰花选择性采收机器人在复杂农田环境中成熟度识别、末端姿态调整以及视觉导航的实时性和准确性要求,选择性采收成功率为 87.37%。

参 考 文 献

[1] Li H, Xia Y, Liu H Y, et al. Nutritional values, beneficial effects, and food applications of broccoli (*Brassica oleracea* var. italica Plenck)[J]. Trends in Food Science & Technology, 2022, 119: 288 - 308.

[2] Kusumam K, Krajnik T, Pearson S, et al. 3D-vision based detection, localization, and sizing of broccoli heads in the field[J]. Journal of Field Robotics, 2017, 34(8): 1505 - 1518.

[3] 刘成良, 贡亮, 苑进, 等. 农业机器人关键技术研究现状与发展趋势[J]. 农业机械学报, 2022, 53(7): 1 - 22, 55.
Liu Chengliang, Gong Liang, Yuan Jin, et al. Current status and development trends of agricultural robots[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7): 1 - 22, 55. (in Chinese)

[4] 张静, 王俊, 杜冬冬, 等. 履带自走式单行大白菜收获机设计与试验[J]. 农业机械学报, 2022, 53(12): 134 - 146.
Zhang Jing, Wang Jun, Du Dongdong, et al. Design and experiment of crawler self-propelled single-row harvester for Chinese cabbage[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(12): 134 - 146. (in Chinese)

[5] 赵春江, 范贝贝, 李瑾, 等. 农业机器人技术进展、挑战与趋势[J]. 智慧农业(中英文), 2023, 5(4): 1 - 15.
Zhao Chunjiang, Fan Beibei, Li Jin, et al. Agricultural robots: technology progress, challenges andtrends[J]. Smart Agriculture, 2023, 5(4): 1 - 15. (in Chinese)

[6] Blok P M, Barth R, Van Den Berg W. Machine vision for a selective broccoli harvesting robot[J]. IFAC - PapersOnLine, 2016, 49(16): 66 - 71.

[7] Kusumam K, Krajnik T, Pearson S, et al. Can you pick a broccoli? 3D-vision based detection and localisation of broccoli heads in the field[C]// 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). New York: IEEE, 2016: 646 - 651.

[8] Blok P M, Evert F K, Tielen A P M, et al. The effect of data augmentation and network simplification on the image-based detection of broccoli heads with Mask R - CNN[J]. Journal of Field Robotics, 2021, 38(1): 85 - 104.

[9] Blok P M, Van Henten E J, Van Evert F K, et al. Image-based size estimation of broccoli heads under varying degrees of occlusion[J]. Biosystems Engineering, 2021, 208: 213 - 233.

[10] Blok P M, Kootstra G, Elghor H E, et al. Active learning with MaskAL reduces annotation effort for training Mask R - CNN on a broccoli dataset with visually similar classes[J]. Computers and Electronics in Agriculture, 2022, 197: 106917.

[11] Garcia-Manso A, Gallardo-Caballero R, Garcia-Orellana C J, et al. Towards selective and automatic harvesting of broccoli for agri-food industry[J]. Computers and Electronics in Agriculture, 2021, 188: 106263.

[12] Kang S, Li D, Li B, et al. Maturity identification and category determination method of broccoli based on semantic segmentation models[J]. Computers and Electronics in Agriculture, 2024, 217: 108633.

[13] Kang S, Fan J, Ye Y, et al. Maturity recognition and localisation of broccoli under occlusion based on RGB - D instance segmentation network[J]. Biosystems Engineering, 2025, 250: 270 - 284.

[14] Zhang J, Xie J, Zhang F, et al. Greenhouse tomato detection and pose classification algorithm based on improved YOLOv5[J]. Computers and Electronics in Agriculture, 2024, 216: 108519.

[15] Meng Z, Du X, Sapkota R, et al. YOLOv10 - pose and YOLOv9 - pose: real-time strawberry stalk pose detection models[J]. Computers in Industry, 2025, 165: 104231.

[16] 苑进. 选择性收获机器人技术研究进展与分析[J]. 农业机械学报, 2020, 51(9): 1 - 17.
Yuan Jin. Research progress analysis of robotics selective harvesting technologies[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(9): 1 - 17. (in Chinese)

[17] Diao Z, Guo P, Zhang B, et al. Navigation line extraction algorithm for corn spraying robot based on improved YOLOv8s network[J]. Computers and Electronics in Agriculture, 2023, 212: 108049.

(下转第 303 页)

[20] Cao J. White button mushroom RGB - D images[EB/OL]. Mendeley Data (2020 - 11 - 07). https://doi.org/10.17632/d3x2w9p439.2.

[21] 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)

[22] 刘雨青,张昕程,万家城,等. 基于双目立体视觉的冷冻黄鳍金枪鱼体尺与体质量测量方法[J]. 农业机械学报, 2025, 56(10): 36 - 44.
Liu Yuqing, Zhang Xincheng, Wan Jiacheng, et al. Binocular stereo vision-based measurement method for body size and body weight measurement of frozen yellowfin tuna[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(10): 36 - 44. (in Chinese)

[23] 王伟,张彦斐,宫金良,等. 基于自适应升温模拟退火算法的农业机器人全区域覆盖策略[J]. 华南农业大学学报, 2021, 42(6): 126 - 132.
Wang Wei, Zhang Yanfei, Gong Jinliang, et al. Whole area coverage strategy of agricultural robot based on adaptive heating simulated annealing algorithm[J]. Journal of South China Agricultural University, 2021, 42(6): 126 - 132. (in Chinese)

[24] 马全坤,张彦斐,宫金良. 基于记忆模拟退火和 A* 算法的农业机器人遍历路径规划[J]. 华南农业大学学报, 2020, 41(4): 127 - 132.
Ma Quankun, Zhang Yanfei, Gong Jinliang. Traversal path planning of agricultural robot based on memory simulated annealing and A* algorithm[J]. Journal of South China Agricultural University, 2020, 41(4): 127 - 132. (in Chinese)

[25] 李洪利,刘欣悦,杜博军,等. 面向自由空间光通信波前校正的改进模拟退火算法[J]. 中国光学(中英文),2025, 18(4): 784 - 793.
Li Hongli, Liu Xinyue, Du Bojun, et al. Improved simulated annealing algorithm for wavefront correction in free-space optical communication[J]. Chinese Optics, 2025, 18(4): 784 - 793. (in Chinese)

[26] 李宏伟,英昌杰,管官. 基于改进粒子群-模拟退火算法的集装箱船结构优化方法[J/OL]. 中国舰船研究. [2025 - 11 - 19]. https://doi.org/10.19693/j.issn.1673 - 3185.04540.
Li Hongwei, Ying Changjie, Guan Guan. A structural optimization method for container ships based on an improved particle swarm-simulated annealing algorithm[J/OL]. Chinese Journal of Ship Research. [2025 - 11 - 19]. https://doi.org/10.19693/j.issn.1673 - 3185.04540. (in Chinese)

(上接第 210 页)

[18] Yang Y, Zhou Y, Yue X, et al. Real-time detection of crop rows in maize fields based on autonomous extraction of ROI[J]. Expert Systems with Applications, 2023, 213: 118826.

[19] Guo P, Diao Z, Zhao C, et al. Navigation line extraction algorithm for corn spraying robot based on YOLOv8s - CornNet[J]. Journal of Field Robotics, 2024, 41(6): 1887 - 1899.

[20] 高旭, 檀国印, 朱长志, 等. 青花菜品种资源重要农艺性状的鉴定与评价[J]. 浙江农业科学, 2022, 63(5): 1098 - 1105.

[21] 张俊萍. 稀特蔬菜生产新技术[M]. 北京:中国农业出版社, 2017.

[22] Ouyang D, He S, Zhang G, et al. Efficient multi-scale attention module with cross-spatial learning[C]// ICASSP 2023—2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2023: 1 - 5.

[23] Stergiou A, Poppe R, Kalliatakis G. Refining activation downsampling with SoftPool[C]// 2021 IEEE/CVF International Conference on Computer Vision (ICCV). New York: IEEE, 2021: 10337 - 10346.

[24] Wang X, Zhao Q, Jiang P, et al. LDS - YOLO: a lightweight small object detection method for dead trees from shelter forest [J]. Computers and Electronics in Agriculture, 2022, 198: 107035.

[25] Li H, Li J, Wei H, et al. Slim-neck by GSConv: a lightweight-design for real-time detector architectures[J]. Journal of Real-Time Image Processing, 2024, 21(3): 62.

[26] Zhang Y F, Ren W, Zhang Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146 - 157.