

基于改进 YOLOv11-Pose 的串番茄识别与 采摘点定位方法^{*}

刘安琪¹ 杨鸿波¹ 潘 航² 陈金龙³ 湛永松³

(1. 北京信息科技大学自动化学院 北京 100192; 2. 长治学院计算机系 长治 046011;

3. 桂林电子科技大学计算机与信息安全学院 桂林 541004)

摘 要: 在温室环境下的串番茄自动化采摘过程中,快速识别串番茄与定位采摘点至关重要。针对温室环境中串番茄自动化采摘中果串快速识别与采摘点精准定位需求,提出了一种基于改进 YOLOv11-Pose 的视觉检测定位方法。该方法首先构建了一个包含串番茄和采摘点同时标注的数据集;其次通过在 YOLOv11-Pose 原模型的不同位置中引入了 GhostC3ECA 与 C3k2_CA 两个模块,分别实现网络的轻量化与采摘点的精准定位,最终实现端到端的串番茄实时检测以及采摘点精准定位;在网络后端,利用面向姿态的非极大值抑制来优化密集果实的检测与关键点的定位效果。实验在玻璃温室采集的串番茄数据集上进行验证,串番茄的识别精度为 96.2%,采摘点的定位精度为 86.5%,同时模型的推理速度为 67.2 fps,能很好的满足实时性处理要求。经现实的温室采摘场景验证,所提出的基于改进 YOLOv11-Pose 的串番茄识别与采摘点定位方法,为自动化采摘机器人提供了一种高效、可靠的视觉检测方案。

关键词: 农业机器人;串番茄采摘;YOLOv11-Pose;关键点检测;计算机视觉;轻量化

中图分类号: TP391.4;TN911.73 **文献标识码:** A **国家标准学科分类代码:** 520.60

Research for identification of cluster tomatoes and positioning of picking points based on an improved YOLOv11-Pose

Liu Anqi¹ Yang Hongbo¹ Pan Hang² Chen Jinlong³ Zhan Yongsong³

(1. College of Automation, Beijing University of Information Science and Technology, Beijing 100192, China;

2. Department of Computer Science, Changzhi University, Changzhi 046011, China; 3. School of Computer Science and Information Security, Guilin University of Electronic Science and Technology, Guilin 541004, China)

Abstract: In the automated harvesting of cluster tomatoes in greenhouse environments, rapid identification of cluster tomatoes and precise localization of picking points are critical. This work addresses the requirements for rapid cluster recognition and accurate picking point localization in automated cluster tomato harvesting within greenhouses by proposing a visual detection and localization method based on an improved YOLOv11-Pose. The method first constructs a dataset containing simultaneous annotations of cluster tomatoes and picking points; then introduces two modules GhostC3ECA and C3k2_CA at different positions within the original YOLOv11-Pose model to achieve network lightweighting and precise picking point localization, respectively. This ultimately enables end-to-end real-time detection of cluster tomatoes and accurate picking point localization. In the network's backend, we utilize pose-aware non-maximum suppression to optimize detection of densely clustered fruits and key points localization. Experiments validated the approach on a glasshouse-collected cluster tomato dataset, achieving 96.2% recognition accuracy and 86.5% picking point localization accuracy. The model operates at 67.2 fps, meeting real-time processing demands. Verified by real-word greenhouse harvesting scenarios, the proposed cluster tomatoes recognition and picking point localization method based on an improved YOLOv11-Pose provides an efficient and reliable visual detection solution for automated harvesting robots.

Keywords: agricultural robot; cluster tomato harvesting; YOLOv11-Pose; key points detection; computer vision; lightweight

0 引 言

近年来,设施农业尤其是温室栽培在全球范围内快速

发展。文献[1]提出温室通过对光照、温度、湿度和营养等要素的精细控制,大幅提高了作物的单位面积产量和产品质量,并使季节性地域性限制得到缓解;与此同时,温室

收稿日期:2025-11-22

^{*} 基金项目:广西省科技计划重点研发项目(AB25069476, AB24010164, AB23026048)、国家自然科学基金京津冀基础研究合作专项(F202405028)、山西省基础研究计划项目(20230302122267)、长治市基础研究计划项目(JC202402)资助

生产也促使农业向集约化、智能化方向转型,推动了对自动化与智能化劳动力的迫切需求。

番茄作为全球最重要的经济作物与蔬菜之一,产量与消费规模位居前列。文献[2]提出,根据联合国粮农组织统计,番茄的全球产量达到数亿吨量级,其产业链在食品安全、加工与贸易中占有重要地位。

由于番茄生产的规模化和市场化特征,降低生产成本、提高采收效率成为提升竞争力的关键。而在番茄全生长周期中,采摘环节长期被认为是劳动最密集、成本最高的环节之一。文献[3]提出人工采摘不仅占用大量劳动力,而且受劳动成本、招聘难度与劳动强度波动等影响显著。因此,用自动化技术替代或辅助人工采摘,能够直接降低人工成本、提升一致性并缓解劳动力短缺问题。

自动化技术在温室管理(如环境控制、灌溉、施肥等)方面已取得显著进展,且由文献[4-10]表明,在大果检测、果实成熟度识别及一般作物采摘上等方面也已取得一定进展。但针对串番茄的端到端智能采摘系统,尤其是实现高鲁棒性的果梗采摘点定位仍存在显著不足,面临诸多挑战。首先,串番茄的群体结构复杂,多个果实通过细小的果梗连成簇状,簇内果实相互近邻、姿态多变,导致视觉感知与机械作用空间都极为受限。其次,叶片、果实之间及果梗与主茎之间的遮挡普遍存在,严重影响基于视觉的检测与定位算法的准确性。再者,果梗纤细且易被损伤,采摘操作要求对果梗采摘点位置与方向实现高精度定位与切割,否则会造成果实或植株的机械损伤,影响商品价值与后续生长。

早期串番茄检测与采摘点定位研究主要依赖颜色阈值、颜色空间变换,形态学处理、边缘/轮廓提取和纹理特征等手工特征对目标进行检测与定位。这类方法在背景简单、光照稳定且果实无遮挡的条件下能取得较高精度,但在温室复杂背景、可变光照、果实密集、叶片与果串之间存在遮挡时鲁棒性显著下降,误检、漏检率较高,难以满足自动化采摘对高可靠感知的要求。

随着深度学习的发展,为串番茄识别与采摘点定位提供了新的思路,主要分为“两阶段”检测方法与“单阶段”检测与定位方法。

“两阶段”检测定位方法首先对番茄进行检测识别,得到果串的候选区域,然后再在候选区域上做精细的果梗检测、分割或关键点回归,从而获取采摘点。如 Rong 等^[11]采用了先在远处对整串进行粗检测果串,再靠近后进行精细定位果梗及切割点的两步策略。Ni 等^[12]先对番茄及果梗所在区域进行识别,然后对采摘点、果实中心点等关键点进行回归,从而确定番茄果实的采摘点。此外,文献[13]总结了一些使用 YOLO 或 Faster R-CNN 算法检测识别果实或果串,然后对候选区域进行专门的果梗分割或关键点回归。这种检测识别精度较高,分为了果实检测识别与果梗分割或关键点回归两个独立网络,便于模块化调试。但在速度与误差累积(检测误差会传到关键点定位中)上存在劣势。

“单阶段”检测定位方法可以实现果串检测与关键点定位的同时输出,可实现更低延迟,减少有检测误差导致的关键点定位误差,非常适合对实时性有强要求的机器人采摘场景。如 Zhang 等^[14]设计了一种 TPMv2 网络,直接以端到端多任务方式估计串番茄的检测框与关键点,为串番茄提供了关于其茎、花梗和果实位置和形态等信息。Li 等^[15]提出了一种自上而下的融合网络 TDFNet,实现了对串番茄的检测识别以及关键点定位,并构建了一种自动番茄采摘系统。“单阶段”端到端的串番茄识别和采摘点定位的方法,将实例检测与关键点输出统一到一个前向网络中,实现了检测与关键点的联合训练与推理,并且兼顾实时性与定位能力,适合兼顾速度与精度要求的自动化采摘场景。

目前,已有众多目标检测与姿态估计算法的研究^[16-17]。其中,YOLO-Pose 是一种基于“单阶段”目标检测与关键点定位的高效算法,在工业检测、体育分析、疲劳驾驶检测等领域展示了端到端的高效解决优势^[18-21],但是在农业场景中的研究应用尚不充分。

鉴于上述情况,本文提出了一种串番茄的四关键点表示方式,并构建了一个包含串番茄目标框以及检测关键点在内的玻璃温室串番茄数据集;提出了一种基于轻量化改进的 YOLOv11-Pose 网络模型,在 Backbone 层用 GhostConv 替代部分标准 Conv 层进行轻量化处理,插入了轻量化高效通道注意模块(efficient channel attention, ECA),且在 Neck 层引入了坐标注意力机制(coordinate attention, CA),并将以上改进封装为两个复合模块引入原 YOLOv11-Pose 网络模型,从而实现了端到端的串番茄的快速识别与采摘点的精确定位;并在网络后端,利用并优化了面向姿态的非极大值抑制方法(pose-aware non-maximum suppression, Pose-NMS),解决了串番茄在密集场景下传统基于交并比(intersection over union, IoU)的非极大值抑制方法(non-maximum suppression, NMS)容易把不同采摘点错误抑制,从而造成漏检的问题。

本研究聚焦于面向温室串番茄的视觉感知与采摘点定位方法,期望为温室智能采摘系统提供关键技术支撑。

1 方 法

1.1 整体流程设计

本文所提出基于改进 YOLOv11-Pose 构建的端到端的串番茄快速识别与采摘点精准定位方法整体流程如图 1 所示。首先,构建了一个商业化温室的串番茄 RGB 图像数据集,并对串番茄目标框及包含采摘点在内的 4 个关键点进行了标注;然后,对 YOLOv11-Pose 模型进行轻量化改进,进行模型训练,并在网络后端利用并优化 Pose-NMS 降低重复抑制造成的串番茄及关键点的漏检问题;最终在温室环境中实现对串番茄的检测识别与关键点的定位预测实验,以实现提高后续机械臂采摘的成功率的目标。

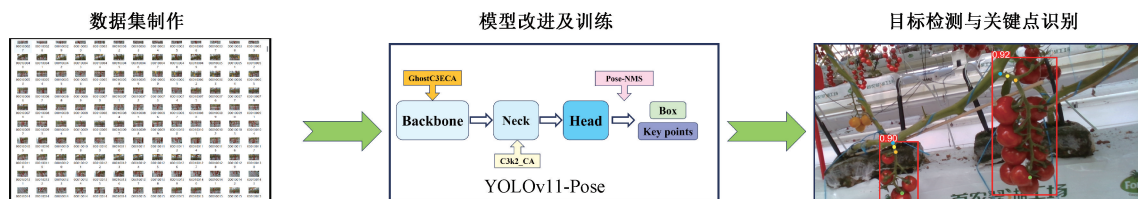


图 1 基于 YOLOv11-Pose 的串番茄识别与采摘点定位方法整体流程

Fig. 1 The overall process of identification and positioning of picking points of cluster tomatoes based on YOLOv11-Pose model

1.2 关键点的定义

由于地球重力的影响,串番茄果实一般垂直向下生长。并且由于串番茄果穗的较为规则拓扑结构,为计算机视觉算法识别串番茄与采摘点提供了基础。为了使关键点具有意义且可以直接映射到串番茄的识别采摘工作,本文提出了一种串番茄四关键点的表示方式,描述了串番茄果串的拓扑与方向性:1)果梗着生节(peduncle node, PN),串番茄果梗与主蔓或侧枝的连接处;2)串番茄集群起始点(start point, SP),串番茄果实集群最上端第 1 个果实的连接点;3)串番茄集群末端点(ending point, EP),串番茄果实集群末端果实的连接点;4)采摘点(cutting point, CP),定位为 PN 与 SP 之间果梗上的串番茄切割点。串番茄关键点选取如图 2 所示。



图 2 串番茄关键点选取

Fig. 2 Selection of cluster tomatoes key points

1.3 YOLOv11-Pose 模型轻量化改进

YOLOv11-Pose 是一种面向实时场景的单阶段检测与关键点回归框架,其设计继承并优化了 YOLO 系列算法在速度与工程可用上的优势。网络由 3 部分构成:骨干(Backbone)用于多尺度特征提取、颈部(Neck)用于多尺度特征融合以及检测头(Head)用于并行输出目标的检测框、类别置信度与关键点信息。YOLOv11-Pose 具有多个版本:YOLOv11n-Pose、YOLOv11s-Pose、YOLOv11m-Pose、YOLOv11l-Pose 以及 YOLOv11x-Pose。其中,YOLOv11n-Pose 具有检测速度快、占用计算资源少以及模型体积小等优势,因此本文选用该模型作为基础模型。与传统的方法相比,YOLOv11-Pose 在一次前向传播中就可以同时完成目标检测与关键点估计,能够在对实时性要求较高的任务应用中提供更少的推理时间。

针对温室环境中,串番茄出现的高度的尺度变化以及果串密集、与枝叶之间的部分遮挡等复杂情况,同时为了

应用于实践作业时的串番茄的检测与采摘点的定位任务,对推理延时与计算资源具有较高要求,本文在保留 YOLOv11-Pose 单阶段检测与定位优势的前提下,对模型进行轻量化改进,且为了使网络模型更为模块化,突出轻量化与注意力模块的协同效果,又便于后续的分析,本文将使用的子结构封装为复合模块并引入 YOLOv11-Pose 网络:GhostC3ECA 模块与 Ck2_CA 模块,以实现在有限算力下保持或提升串番茄目标检测与关键点定位精度与实时性。改进的网络结构如图 3 所示。

1) GhostC3ECA 模块

为了提升在串番茄果串检测以及采摘点定位的速度,本文引入了 GhostC3ECA 模块。GhostC3ECA 模块是在原有 C3k2 模块的基础上,用 GhostConv 替代内部 Bottleneck 模块的多通道 3×3 高计算量卷积以实现模型的轻量化,并在 C3k2 模块末端加入 ECA 注意力机制以恢复通道表达能力,模块结构如图 4 所示。该模块主要部署于 Backbone 的中、深层,提供低成本的高容量表示,同时通过通道注意力机制补偿引入 GhostConv 模块后的特征表示稀疏性,维持对关键点的回归精度,从而实现整个模型的轻量化。

Han 等^[22]提出了 GhostConv,这是一种类似残差的轻量化模块,通过“本征特征+廉价线性变换”策略,以很小的计算量从图像的原始特征中挖掘出所需信息的“幽灵特征图”(ghost features maps),从而大幅降低计算量于参数量,在移动端或嵌入式设备等对高实时、低资源要求等场景,可以兼顾效率与性能的平衡。GhostConv 通过以下两个步骤来实现对图像特征的提取:

(1)先使用少量的标准卷积生成部分核心特征图,这些特征图被称为“主特征图”;

(2)然后对每个本征特征图进行“廉价线性变换”操作(如深度卷积),生成“幽灵特征图”。

最终,GhostConv 将“本征特征图”与“幽灵特征图”进行拼接,得到与传统卷积同等数量的特征图,显著降低了计算成本,加快了计算速度。对于 Backbone 层中的多通道 3×3 高计算量卷积造成的高计算量,使用 GhostConv 替换可以降低浮点运算(FLOPs)与模型延迟。

但是 GhostConv 模块的“主特征+廉价线性变换”策略会导致通道信息的稀疏化与表征细粒度的弱化,可能对果梗、果串连接处等关键纹理特征提取产生影响,因此引

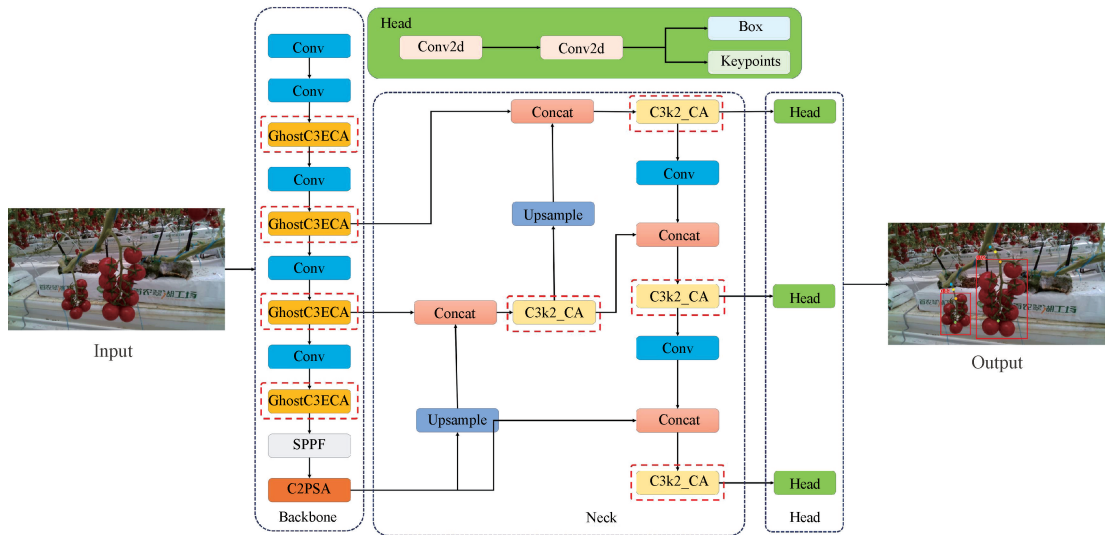


图 3 改进的 YOLOv11-Pose 网络结构
Fig. 3 Structure of the improved YOLOv11-Pose network

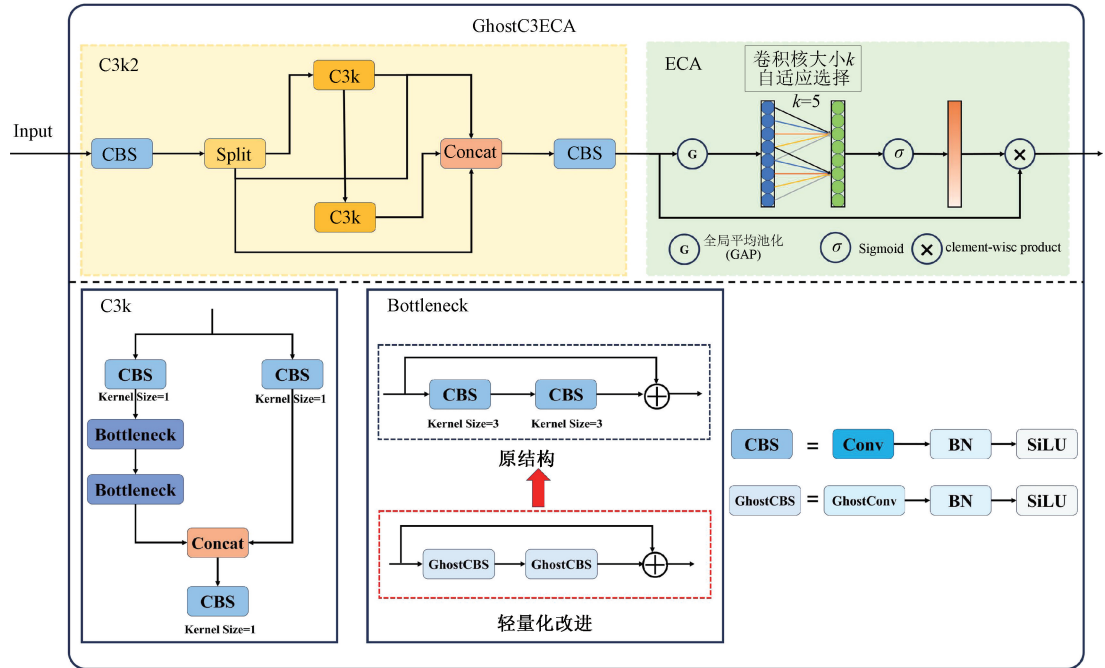


图 4 GhostC3ECA 模块结构
Fig. 4 Structure of GhostC3ECA module

入了 ECA 注意力机制,一种参数极小、延迟低的通道重标定机制,可在通道维对关键特征做放大与抑制,从而补偿引入 GhostConv 模块引入带来的特征信息提取弱化。

Wang 等^[23]提出的 ECA 注意力机制是一种轻量化的通道注意力机制,通过一个一维卷积对每个通道的特征进行卷积,然后使用 Sigmoid 函数来学习通道注意力,从而减少计算复杂度,并在一定程度上提高模型的性能。ECA 中的一维卷积是直接在全局平均池化层(global average pooling, GAP)后直接使用,去除了全连接层,来实现跨通

道间的信息交互,其卷积核的大小 k 计算公式为:

$$k = \left\lfloor \frac{\ln(C)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}} \quad (1)$$

其中, $\lfloor t \rfloor_{\text{odd}}$ 表示最接近于 t 的奇数, C 为特征通道数, γ 与 b 分别为 2 和 1。

Backbone 层的作用是从输入图像中提取目标的特征信息,其特征通道数会随层级的增长导致卷积计算开销占比极高,并且输出通道间常含大量的冗余信息。将 GhostConv 嵌入 Backbone 后,可通过“少量主特征+廉价

线性变换”的策略有效去除冗余并显著降低参数与 FLOPs,从而在保持局部结构与纹理特征信息的同时提升模型的推理效率;而 ECA 注意力机制在模块末端以极小的额外成本对通道间局部交互建模并进行重标定,能够补偿引入 GhostConv 可能带来的判别能力下降,突出对下游检测与关键点回归有用的通道特征。因此, GhostConv + ECA 在 Backbone 中实现了“先降冗余、再增强判别”的协同效应,既满足了边缘部署的计算约束,又保持了必要的表征能力。

2) C3k2_CA 模块

由于串番茄采摘点的定位对于空间位置高度较为敏感,依赖于精确的横纵向的位置信息,因此本文在 Neck 层的 C3k2 模块后附加 CA 注意力机制,以在多尺度融合处提供位置敏感的空间-通道注意力,增强多尺度融合特征的空间感知于方向性位置信息,使得关键点(尤其是采摘点)定位的误差更小,增强关键点回归所需的局部位置信息,从而在串番茄果串密集与部分遮挡场景下提高对串番茄的检测与采摘点定位的准确性。C3k2_CA 模块结构如图 5 所示。

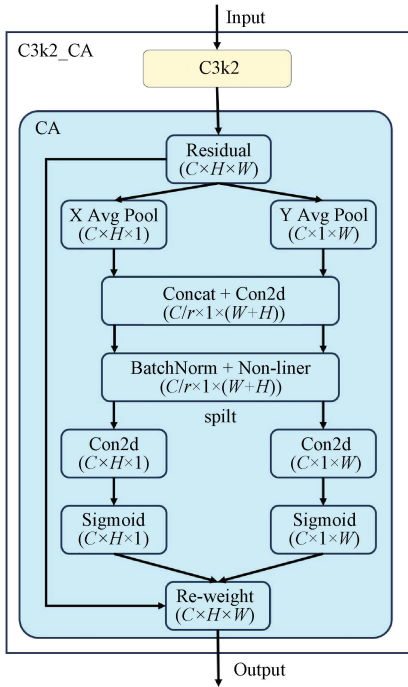


图 5 C3k2_CA 模块结构

Fig. 5 Structure of C3k2_CA module

Hou 等^[24]提出了 CA 坐标注意力机制,这是一种结合通道关系与空间位置信息的注意力机制,旨在在不增加计算成本的前提下增强对特征的表达能力。传统的通道注意力机制,如 Hu 等^[25]提出的压缩与激励网络(squeeze-and-excitation networks, SENet),通过 2D 全局池化提取空间信息,仅关注通道间的关系,而忽略了“特征在哪个位置更为重要”的空间坐标信息。而 CA 的核心思想是将“坐标信

息”嵌入到通道注意力中,可以让每个通道的权重不仅依赖通道自身的特征强度,还与该特征在图像中的水平、垂直位置相关,从而更精准地定位和增强关键区域的特征。

CA 模块主要由两部分组成:

(1)坐标信息嵌入,使用一维池化层对输入特征分别进行水平方向和垂直方向的池化,得到两个包含位置信息的方向特征图;

(2)坐标注意力生成,将水平方向与垂直方向的特征图融合,通过轻量化操作提取通道与位置的关联特征,分别生成对应的水平方向与垂直方向的注意力权重,从而绑定通道与位置信息,最终用生成的坐标注意力权重增强原特征图中的关键区域。

将 CA 部署在 Neck 层的多尺度融合阶段,是因为 Neck 层的核心任务是将不同语义层级的特征对齐并保留空间位置信息以服务检测头和关键点回归头。与仅做通道重标定的注意力机制不同,CA 通过沿水平与垂直方向编码位置信息并将其注入通道注意力权重,能够在保持通道选择性的同时保留并放大空间坐标相关的响应。这对关键点定位尤为重要,即采摘点和果串方向性等几何属性依赖精细的位置信息与跨尺度一致性,而 CA 可以在 Neck 中能增强目标边界、朝向与相对几何关系等特征信息的表达,从而提高多尺度特征融合后关键点回归的稳定性与精度。

3) 模型轻量化改进总结

本文遵循了模块化设计中的“计算减负在前、空间增强在后”的整体思路,在 YOLOv11-Pose 模型中不同位置分别引入了 GhostC3ECA 与 C3k2_CA 两个复合模块: GhostC3ECA 模块在 Backbone 层中侧重于降低计算冗余并恢复通道判别能力提升实时性;C3k2_CA 模块在 Neck 层中增强空间位置信息感知与多尺度特征对齐能力,弥补因轻量化造成的空间细节的丢失,以上改进策略在网络结构上具有明确的理论依据与工程合理性,从而在总体上实现对串番茄的检测精度的提升,同时增强了对关键采摘点的定位精度。

1.4 Pose-NMS

在 YOLO-Pose 的 Ultralytics 系列实现中,目标检测的检测分支与关键点识别的姿态分支共用一个非极大值抑制,即按照检测的置信度对候选框进行排序,再使用 IoU 去排除重叠框,最后将被保留的检测框与对应的关键点一起输出。YOLO-Pose 的“检测框与关键点的锚点绑定设计”,仅使用 NMS 解决检测框的冗余,关键点无需额外分组或者使用 NMS,直接跟随检测框输出。IoU-NMS 的计算公式为:

$$\text{IoU}(B_i, B_j) = \frac{\text{Area}(B_i \cap B_j)}{\text{Area}(B_i \cup B_j)} \quad (2)$$

按照置信度降序遍历候选框,若对于候选框 i 与 j 存在 $\text{IoU}(B_i, B_j) > \tau_{\text{iou}}$, 且 $\text{score}_j < \text{score}_i$, 则移除 j , 实现

单步抑制。

而在温室中串番茄目标检测与采摘点定位任务中,由于果串密集、果实间互相遮挡等问题,往往会导致候选框出现高度重叠以及关键点位置微细差异等情况。传统的基于边界框的 IoU 的 NMS 在此类场景下容易将语义上不同但空间重叠度高的串番茄检测框误抑制,从而导致漏检或关键点丢失等问题,从而降低下游机械臂的采摘成功率。

为此,本文在网络后端利用并优化了面向姿态感知的非极大值抑制 Pose-NMS,其核心思想为在判定是否抑制串番茄检测候选框时,不仅考虑边界框的空间重叠度 IoU,还考虑两候选框对应的关键点(尤其是 CP)的姿态相似度,若两个候选框 IoU 高但关键点的位置相距较远,则保留为两个不同实例;反之则抑制。这样能够在密集的番茄果串区域减少误抑制率,从而提升实例的检测完整性与关键点定位的可靠度。

首先,每张串番茄图像在经过改进的轻量化 YOLOv11-Pose 模型检测后,先用检测框与关键点的置信度进行初筛,过滤掉低质量的预测结果,从而得到每张图像的候选预测集合 C :

$$C = \{c_i\} = \{B_i, s_i^{\text{box}}, (p_{i,k}, s_{i,k}^{\text{kp}}, v_{i,k}), S_i\} \quad (3)$$

其中, B_i 为串番茄的检测候选框, s_i^{box} 为对应候选框的置信度。对于每个关键点 $k \in \{1, 2, 3, 4\}$, $p_{i,k}$ 为每个候选点的坐标, $s_{i,k}^{\text{kp}}$ 为对应关键点的置信度, $v_{i,k}$ 为关键点的可见或有效标志位, S_i 为候选框与关键点的综合置信度,计算公式为:

$$S_i = \alpha s_i^{\text{box}} + (1 - \alpha) \frac{1}{K} \sum_{k \in V_{ij}} s_{i,k}^{\text{kp}} \quad (4)$$

式中: α 为权重系数,用于决定检测框置信度与关键点置信度的相对重要性。

再对候选集合 C 按照综合置信度 S_i 进行降序排序后,得到候选列表。

对于候选列表中的候选 i 与 j , 构建共同可见关键点集合为:

$$V_{ij} = \{k \mid v_{i,k} = 1 \wedge v_{j,k} = 1\} \quad (5)$$

若 $|V_{ij}| = 0$, 即候选 i 与 j 无共同可见点,则使用采摘点 CP 进行判断(在采摘点 CP 可见时),或采用基于 IoU 的策略计算(在采摘点 CP 不可见时)。

对于有公共可见点的候选,用关键点的置信度作为权重,计算置信度加权距离为:

$$D_{ij}^w = \frac{\sum_{k \in V_{ij}} \omega_k \|p_{i,k} - p_{j,k}\|_2}{\sum_{k \in V_{ij}} \omega_k} \quad (6)$$

$$\omega_k = s_{i,k}^{\text{kp}} \cdot s_{j,k}^{\text{kp}} \quad (7)$$

在本文实验中,对串番茄关键点中的采摘点 CP 最为关注,因此在式(6)中,将集合 V_{ij} 定义为了只包含采摘点的集合,即 $V_{ij} = \{CP\}$ 。

然后将置信度加权距离进行归一化,以消除尺度的影响,使阈值更具有通用性:

$$\hat{D}_{ij} = \frac{D_{ij}^w}{\sqrt{\omega_i^2 + h_i^2}} \quad (8)$$

其中, ω_i 与 h_i 为候选框的短边 $s_{ij} = \min(\omega_i, h_i)$ 。

Pose-NMS 的软判定(score 衰减)依据为对于得分较低的候选 j 的得分做联合衰减,衰减因子为:

$$\gamma_{ij} = \exp\left(-\frac{\text{IoU}(B_i, B_j)^2}{\sigma_{\text{IoU}}^2} - \frac{D_{ij}^2}{\sigma_{\text{pose}}^2}\right) \quad (9)$$

其中, σ_{IoU} 与 σ_{pose} 为软衰减参数。

然后,更新候选框 j 的得分:

$$S_j \leftarrow S_j \cdot \gamma_{ij} \quad (10)$$

继续下一候选,在任何点将 S_j 低于最终得分阈值 T_{score} 时移除,输出剩余候选作为最终结果。 T_{score} 根据检测精度进行的权衡设定。

上述软衰减策略等价于将 IoU 与归一化姿态距离同时视为“相似性度量”:若两者都高(即非常相似),则衰减显著;若姿态距离大但 IoU 高,则衰减较弱,从而保留近重叠但采摘点不同的实例。

本文在网络后端中采用的 Pose-NMS,在判定抑制时同时考虑边界框重叠度与关键点姿态相似度,并使用软衰减策略联合更新候选得分,在密集子集上显著降低误抑制问题,从而提升串番茄采摘点的精准定位与下游机械臂采摘的成功率。

2 实 验

2.1 实验数据集

实验使用的串番茄数据集来自于北京市首农翠湖工场商业化番茄玻璃温室,使用了 Inter RealSense D435i 相机进行拍摄,包含了不同天气、光照等情况,数据集拍摄场景如图 6 所示。

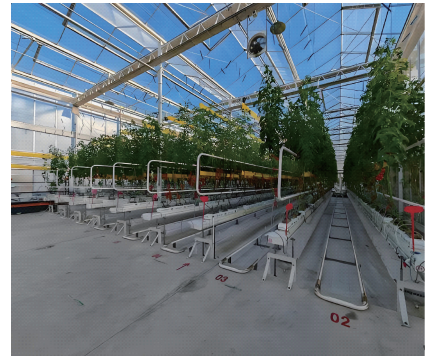


图 6 串番茄数据集的拍摄场景

Fig. 6 Shooting scenarios for the cluster tomatoes dataset

实验数据集由 2 650 张 $1\,280 \times 720$ 的 RGB 图像组成,并使用了 Labelme 标注工具对串番茄实例以及包含采摘点在内的 4 个关键点进行了标注,构建了一个商业化温室

串番茄及关键点的标注数据集。并将整个数据集按照 7:1:2 的比例,将整个数据集划分为训练集、验证集与测试集。且在模型训练期间采用统一的数据增强策略,如将数据集图像进行随机尺度缩放、随机水平翻转等。

2.2 实验环境及参数设置

实验使用的是 Windows11 操作系统,搭载了第 13 代 Intel(R) Core(TM) i5-13450HX 处理器,NIDIA GeForce RTX 4050 独立显卡以及 16 GB 的内存。

模型训练环境为 CUDA 13.0,Python 版本为 3.10,深度学习框架为 PyTorch 2.6.0。训练批次设置为 200,批处理大小为 8,过程中使用了 GPU 进行加速训练。模型训练时使用了 AdamW 优化器,初始学习率为 0.01,采用了 OneCycle 学习率策略,动量为 0.937,权重衰减系数为 0.0005。

2.3 评价指标

在串番茄识别检测中,选择精确率(precision, P)、召回率(recall, R)、平均精度(average-precision, AP)与平均精度均值 mAP₅₀ 作为评价串番茄检测模型的指标,其中 mAP₅₀ 为 IoU 阈值为 0.5 时的平均精度均值。上述评价指标的计算公式为:

$$P = \frac{TP}{TP + FP} \times 100\% \quad (11)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (12)$$

$$AP = \int_0^1 P(R) dR \quad (13)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (14)$$

其中,TP 为模型将正样本正确预测的数量,FP 为模型将负样本错误预测为正样本的数量,FN 为模型将正样本错误预测为负样本的数量,N 为样本中类别的数量。

此外,为评估串番茄检测模型的速度性能,选择了 FLOPs 与 FPS 作为评价指标。FLOPs 为浮点运算次数,用来衡量一个模型的计算复杂度。FPS 为推理速度,表示检测器每秒钟能够处理图片的数量。为评估模型的大小,选择了模型参数量(Parms)作为评价指标。模型中的参数量决定了模型大小以及模型所需的存储空间。模型参数量越小,表明该模型越容易应用于移动终端。

对于串番茄关键点的检测,尤其是采摘点 CP 的定位,选择了目标关键点相似度(object keypoint similarity, OKS)。OKS 可以用来衡量检测到的关键点位置与真实位置之间的匹配程度,OKS 值越接近 1 表示姿态越相似,关键点定位越精准。OKS 计算公式为:

$$OKS = \frac{\sum_i \exp(-\frac{d_i^2}{2s^2 k_i^2}) v_i}{\sum_i v_i} \quad (15)$$

其中, d_i 为第 i 个关键点的预测位置和真实位置之间

的欧氏距离; s 为目标的尺度大小,一般选取目标的边界框面积的平方根; k_i 为与关键点 i 的标准误差,可看作为对每个关键点的重要性赋予权重,用于反映不同关键点的重要性,在本文中采摘点 CP 的重要性最高; v_i 为第 i 个关键点的可见性标识, $v_i \in \{0,1\}$ 。

通过 OKS 计算关键点的相似度后,在计算关键点检测的平均精度:

$$AP_{kpt} = \frac{\sum_m \sum_p \delta(OKS_p > T)}{\sum_m \sum_p v_i} \quad (16)$$

其中, m 为关键点的总数, p 为某关键点在关键点列表中的位置。 $\delta(*)$ 表示若满足条件 $*$, 则 $\delta(*) = 1$, 否则 $\delta(*) = 0$, 在公式中的含义为筛选出满足 OKS 大于阈值 T 的关键点。在本文实验中,将 AP_{kpt} 中的阈值 T 设定为了 0.5,得到了 $AP_{kpt_{50}}$,作为关键点尤其是采摘点 CP 的评价指标。

为衡量串番茄采摘点定位的具体误差,选择将平均定位误差(mean location error, MLE)作为评估指标,在串番茄采摘实验验证时,用于计算预测采摘点坐标与实际采摘点坐标之间的位置偏差,MLE 的计算公式为:

$$MLE = \frac{1}{n} \sum_{i=1}^n \sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + (z_i - \hat{z}_i)^2} \quad (17)$$

其中, (x_i, y_i, z_i) 为实际采摘点坐标, $(\hat{x}_i, \hat{y}_i, \hat{z}_i)$ 为预测的采摘点的坐标。

3 实验结果与分析

为验证本文对 YOLOv11-Pose 的轻量化改进策略以及在网络后端采用 Pose-NMS 方法的有效性,在保证参数与硬件环境一致的情况下进行了对比实验,对比分析本文采用的方案在串番茄检测与采摘点定位时的性能表现。

3.1 消融实验

为了评估在原始 YOLOv11-Pose 网络模型中引入 GhostC3ECA 与 C3k2_CA 两个复合模块后,对串番茄检测与关键点定位的有效性,进行了相关消融实验验证,结果如表 1 所示。

表 1 中的消融实验结果表明,本文提出的在 YOLOv11-Pose 模型中引入 GhostC3ECA 与 C3k2_CA 两个模块后,在串番茄检测与采摘点定位任务中取得了最好的综合性能,相较于原始的 YOLO v11-Pose 模型,本文方法实现了串番茄检测精度与关键点定位精度的同时提升, mAP₅₀ 与 AP_{kpt₅₀} 分别提升了 6.8% 和 5.2%,同时也降低了参数量,提高了检测速度,说明所提模块在精度与速度上均实现了良好的提升。

由表 1 可知,将 C3k2 模块中的重卷积替换为 GhostConv 后,减少了模型的参数量,提升了模型的检测速度,但是在串番茄识别与关键点定位性能出现下滑,分别

表 1 基于改进 YOLOv11-Pose 的串番茄识别与关键点定位消融实验结果对比

Table 1 Comparison of ablation experiments results for the identification of cluster tomatoes and the positioning of key points based on the improved YOLOv11-Pose model

序号	GhostConv	GhostC3ECA	C3k2_CA	mAP ₅₀ /%	AP _{kpt50} /%	Params/(×10 ⁶)	FPS/fps
1	—	—	—	89.4	81.3	14.2	52.0
2	✓	—	—	87.9	79.0	9.8	78.5
3	—	✓	—	90.6	84.0	10.1	69.8
4	—	—	✓	90.0	85.1	14.4	49.5
5	—	✓	✓	96.2	86.5	10.5	67.2

减少了 1.5%和 2.3%，这一结果符合轻量化算子的常见现象，即在提升计算效率的同时伴随着特征表征能力的降低。在 C3k2 中引入 GhostConv 且在末端加入 ECA 注意力机制后，不仅恢复了检测性能并且超越了原 YOLOv11-Pose 模型，分别提升了 1.2%和 2.7%，且保持了较高的 FPS，这表明 ECA 对于 GhostConv 的引入造成的通道表达损失具有较强的补偿功能；ECA 的轻量通道重标定能够恢复特征通道信息，从而改善关键点检测，且不会显著增加计算量。在 Neck 层引入 C3k2_CA 后，由于 CA 这类空间注意力机制在提升计算精度的同时会带来较大的开销，这一特点可以由表中数据体现出：关键点的定位精度提升了 3.8%，但导致了 FPS 的减少。CA 有益于关键点的定位，但其计算开销不可忽视，因此通过在 Backbone 的中、深层使用 GhostC3ECA 减少计算开销，并在 Neck 层特征融合时使用 C3k2_CA，既能保留 CA 的空间增强效果，又避免了整体计算量过高。

3.2 不同模型对比实验

为了客观清晰地展示本文改进的 YOLOv11-Pose 模

型在串番茄检测与采摘点定位的效果，基于自建的温室串番茄数据集，并选择了 YOLO-Pose 系列模型进行对比实验，有 YOLOv5-Pose、YOLOv8-Pose、YOLOv11-Pose 以及本文改进的 YOLOv11-Pose 模型。

为确保公平的训练条件，对比实验使用同一数据集并确保同样的训练集、验证集与测试集的划分，在相同的软硬件环境、同样的图像增强策略、相同的模型超参数以及相等的训练次数 epoch 进行。不同模型的对比实验结果如表 2 所示。对比实验结果表明，本文改进的 YOLOv11-Pose 串番茄识别与采摘点定位算法的 mAP₅₀ 和 AP_{kpt50} 以及参数量、计算复杂度和 FPS，相较于其他模型，具有更好的效果。其中 mAP₅₀ 与 AP_{kpt50} 相较于 YOLOv5-Pose、YOLOv8-Pose、YOLOv11-Pose 分别高出了 11.5%与 8.5%、8.7%与 4.7%、6.8%与 5.2%。且改进的 YOLOv11-Pose 模型的参数量、模型计算复杂度及 FPS 在这 4 种模型中最小，具有实时性优势。改进的 YOLOv11-Pose 串番茄识别与关键点定位的效果如图 7 所示。

表 2 不同模型性能对比

Table 2 Performance comparison results of different models

模型	mAP ₅₀ /%	AP _{kpt50} /%	Params/(×10 ⁶)	GFLOPs	FPS/fps
YOLOv5-Pose	84.7	78.0	12.6	44.5	47.7
YOLOv8-Pose	87.5	81.8	13.8	20.7	50.6
YOLOv11-Pose	89.4	81.3	14.2	19.0	52.0
改进后的 YOLOv11-Pose	96.2	86.5	10.5	18.2	67.2



图 7 基于改进的 YOLOv11-Pose 串番茄检测与关键点定位效果

Fig. 7 Effect of cluster tomatoes and the positioning of key points based on the improved YOLOv11-Pose model

3.3 Pose-NMS 关键参数选择

为了减少在密集的串番茄果串区域误抑制率,提升串番茄的检测完整性与关键点定位的可靠性,本文在网络后端利用并优化了 Pose-NMS 方法,同时考虑边界框的重叠度与关键点的姿态相似度。方法有两个关键的参数:姿态距离阈值 T_{score} 与得分衰减阈值 $\sigma(\sigma_{IoU}$ 与 $\sigma_{pose})$ 。其中 T_{score} 用于判断姿态距离是否接近, σ 用于 score 衰减函数的尺度参数。 T_{score} 过小,条件要求严格,会导致产生更多的误抑制情况,造成漏检率增加,而过大则会导致检测重复率上升; σ 控制软衰减的强度,其数值越小抑制强度越高,数值越大抑制越温和。

为简单直观地实现上述两个关键参数的选择,本文采用了网格搜索法:首先在取值区间 $[0, 1]$ 采用较大步长进行粗搜索,根据 AP_{kpt50}、漏检率、误检率等相关评价指标,找到指标表现良好且稳定的参数取值区间,经搜索,其中 T_{score} 的稳定区间为 $[0.1, 0.5]$, σ 的稳定区间为 $[0.20, 0.80]$;然后在区间范围区间内,采用较小步长进行细化查找,从而找到合适的参数组合。最终将 T_{score} 取值为 0.2, σ 取值为 0.5,作为本文 Pose-NMS 方法中两个关键参数的取值选择。

3.4 果实密集处理方法对比

在应对使用 YOLOv11-Pose 模型在温室中进行串番

茄检测与采摘点定位作业时,由于果串密集、果实间互相遮挡等导致的候选框高度重叠以及关键点位置微细差异等情况,从而造成串番茄检测框误抑制,导致串番茄漏检或关键点丢失,降低采摘成功率,为此本文在模型的后端引入了面向姿态的 Pose-NMS 的处理方法,在判定抑制时同时考虑边界框重叠度与关键点姿态相似度,从而提升串番茄采摘点的精准定位与下游机械臂采摘的成功率。

为验证本文提出的面向姿态感知方法在实际采摘时,对于果串密集与果串间相互遮挡处理的有效性与可行性,本文进行了专门的对比试验。后处理方法包含传统的 IoU-NMS 与本文提出的 Pose-NMS,测试指标包含 mAP₅₀、AP_{kpt50}、重复率(Duplicate rate)以及漏检率(Miss rate),结果如表 3 所示。在串番茄果串密集或果串间相互遮挡的特殊情形,采用本文提出的面向姿态感知的 Pose-NMS 方法,相较于传统的 IoU-NMS 方法,在串番茄识别的平均检测精度 mAP₅₀ 上提升了 10.6%,在关键点的平均检测精度 AP_{kpt50} 上提升了 8.4%,且显著降低了重复率与漏检率,分别降低了 6.3%、8.5%,相较于传统的 IoU-NMS 方法具有明显的优势。传统方法与本文所提方法在串番茄果实密集处的检测结果对比如图 8 所示。

3.5 串番茄采摘实验验证

为检验本文方法的可行性,在温室环境中进行了实际

表 3 果实密集区域处理方法对比

方法	mAP ₅₀	AP _{kpt50}	Duplicate rate	Miss rate
传统的 IoU-NMS	85.6	78.1	12.4	16.8
本文提出的 Pose-NMS	96.2	86.5	6.1	8.3

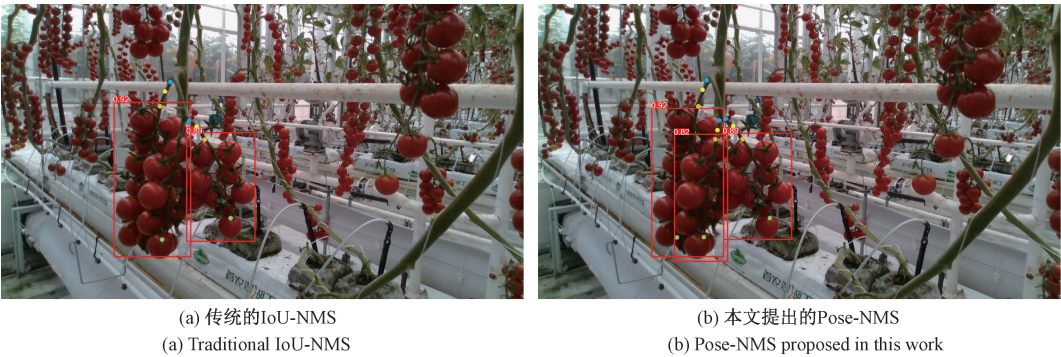


图 8 果串密集区域的检测结果

Fig. 8 Detection results for cluster tomatoes that are densely packed

的串番茄采摘作业,测试实际采摘的成功率,如图 9、10 所示。主要的实验验证过程为利用本文的串番茄的识别与采摘点的定位算法,定位到要进行采摘的串番茄以及对应的采摘点,结合 Inter RealSense D435i 相机得到的深度信息,将采摘点的在相机中的 2D 坐标转化为实际场景中的 3D 坐标,并利用法澳意威机械臂进行采摘验证。

采摘实验验证选择在温室中的果串密集串区域,共进行了 3 组串番茄采摘实验,同时记录了每组采摘实验的平均采摘成功率与采摘点平均定位误差 MLE,如表 4 所示。在串番茄采摘实验验证完成后,对未能成功采摘的案例进行了分析总结。

1)在第 1 组和第 2 组实验中,存在果串被番茄枝叶、茎

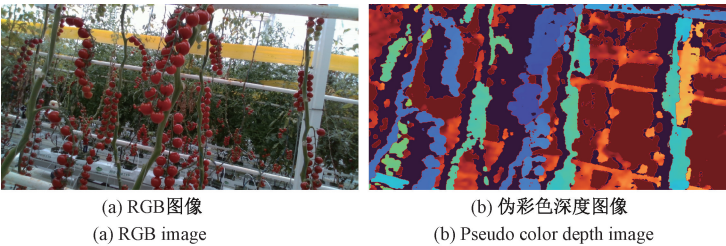


图 9 RealSense D435i 相机同时得到的 RGB 图与伪深度图

Fig. 9 The RGB image and pseudo-depth map captured simultaneously by the RealSense D435i camera

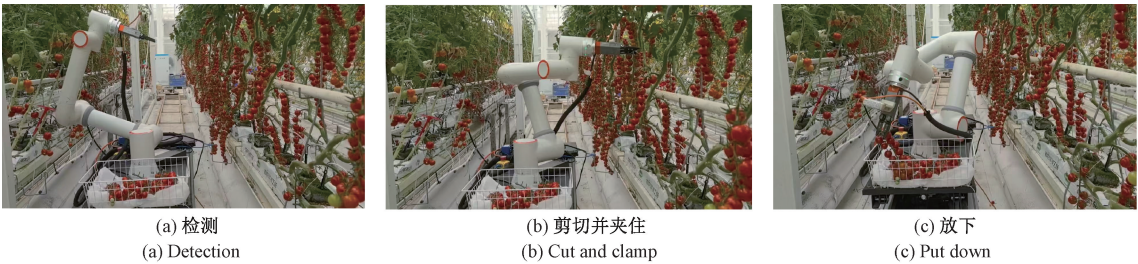


图 10 串番茄采摘实验验证

Fig. 10 Experimental verification of cluster tomatoes harvesting

干等严重遮挡的情况,无法检测到串番茄或无法精准定位采摘点,导致采摘失败;

2)在第 3 组实验中,由于存在对几株串番茄的采摘点定位误差较大,在采摘测试时,机械臂末端到达的位置错误或到达的位置并不适合采摘作业,从而导致采摘失败;

3)在所有测试实验中,均存在机械臂的路径规划不合适的情况,从而会导致机械臂与串番茄植株的茎干发生碰撞,或者与采摘点所在果梗发生碰撞,使得机械臂末端无法到达采摘点位置,从而不能正常的串番茄采摘作业。

表 4 串番茄采摘实验记录

Table 4 Cluster tomatoes picking experiment record

序号	总采摘 株数	成功采摘 株数	采摘 成功率/%	MLE/ mm
1	18	15	83.3	8.2
2	15	13	86.7	6.5
3	16	12	75.0	12.2

上述 3 种串番茄采摘失败的原因也是目前困扰自动化采摘技术发展的主要原因,需要更进一步的研究推进。

4 结 论

为实现在温室中串番茄的快速识别与采摘点的精准定位,本文设计了一种包含采摘点的四关键点串番茄表示方式,并提出了基于 YOLOv11-Pose 的端到端方法。通过在中深层引入了 GhostConv 实现轻量化、在模块末端加入了 ECA 以补偿通道表示损失,并在多尺度融合处采用位置敏感的 CA 注意力机制,并将其封装为 GhostC3ECA 与

C3k2_CA 两个复合模块,并引入原网络。实验证明,该模块化改造在保证实时性的前提下可提升整体性能,在自建温室数据集上的检测结果显示,本文方案在串番茄检测与关键点定位上优于原模型,在 mAP_{50} 与 AP_{kpt50} 分别提升了 6.8%和 5.2%。

本文通过在网络后端引入面向姿态的 Pose-NMS 后处理方法,能显著降低串番茄果串密集及果串相互遮挡场景下的重复率与漏检率,从而提高采摘点定位的准确率以及下游机械臂作业的采摘成功率。尽管面向姿态感知的 NMS 策略已有研究,本文提出的后处理方式有两点不同:将姿态相似度判据专门适配于串番茄的 CP 的定位任务,设计了采摘点优先的权重策略;将关键点置信度作为姿态距离的权重,并与检测框置信度联合做软衰减。经实验验证,相较于传统的 IoU-NMS 方法,本文方法在串番茄平均检测精度 mAP_{50} 与关键点的平均检测精度分别提升了 10.6%、8.4%。

尽管本文在网络的轻量化改造与后处理方面取得了较好结果,但仍存在较多限制:当前验证只是基于单一温室中的规模有限的数据集,对于其他温室环境、不同品种的串番茄的并不适用,因此模型的跨温室能力、不同串番茄品种检测的泛化性有待提高;对于串番茄严重遮挡时,检测与关键点定位效果并不理想,有待改进;以及在采摘时机机械臂路径的规划问题也是影响采摘成功率的重要因素之一。对于未来的研究,将进一步扩充数据集,涵盖各种复杂情景,包括不同的天气、光照,番茄的品种以及不同的温室环境,提高泛化性,同时不断改善模型,提升对于严重遮挡场景中的检测与定位能力,并对采摘时的机械臂的运动路径进行合理规划,减少与植株茎干、果梗等的碰撞,

以期望为作物的自动化采摘提供更加高效、可靠的技术。

参考文献

- [1] Simpson A, Harvey R, Fox C. A review of adaptable technologies for robotic urban horticulture[J]. *Frontiers in Sustainable Food Systems*, 2025,9:91605107.
- [2] Food and Agriculture Organization of the United Nations. Agricultural production statistics 2010-2023 [EB/OL]. [2025-11-20]. <http://www.fao.org/about/zh/>.
- [3] Tan Y R, Liu X, Zhang J M, et al. A review of research on fruit and vegetable picking robots based on deep learning[J]. *Sensors(Basel)*, 2025, 25(12):3677.
- [4] 石国照, 张富贵, 苟园旻, 等. 水果采摘机器人目标识别与采摘点定位研究进展[J]. *中国农机化学报*, 2025, 46(5): 115-122.
- Shi Guozhao, Zhang Fugui, Gou Yuanmin, et al. Research progress on target recognition and picking point localization of fruit picking robot[J]. *Journal of Chinese Agricultural Mechanization*, 2025, 46(5): 115-122.
- [5] 樊湘鹏, 张岩琪, 周硕, 等. 基于改进 YOLOv8s 与 RGB-D 信息融合的番茄采摘机器人识别与定位方法[J]. *农业工程学报*, 2025, 41(15): 106-114.
- Fan Xiangpeng, Zhang Yanqi, Zhou Shuo, et al. Recognition and position method of tomato picking robot based on improved YOLOv8s and RGB-D information fusion[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2025, 41(15): 106-114.
- [6] Zu L L, Zhao Y P, Liu J Q, et al. Detection and segmentation of mature green tomatoes based on Mask R-CNN with automatic image acquisition approach[J]. *Sensors*, 2021, 21(23): 7842.
- [7] Mu Y, Chen T S, Ninomiya S, et al. Intact detection of highly occluded immature tomatoes on plants using deep learning techniques[J]. *Sensors(Basel)*, 2020, 20(10): 2984.
- [8] 刘坤, 吉宏亚, 黄程菲, 等. 基于改进 YOLOv5s 的番茄成熟度识别技术研究[J]. *中国农机化学报*, 2025, 46(5): 79-84.
- Liu Kun, Ji Hongya, Huang Chengfei, et al. Research on tomato maturity recognition technology based on improved YOLOv5s[J]. *Journal of Chinese Agricultural Mechanization*, 2025,46(5): 79-84.
- [9] 马智杰, 李长勇. 基于机器视觉的鲜食葡萄采摘点和抓取点定位方法[J]. *农机化研究*, 2025, 47(7): 191-196.
- Ma Zhijie, Li Changyong. Table grape picking point and grasping point positioning method based on machine vision [J]. *Journal of Agricultural Mechanization Research*, 2025, 47(7): 191-196.
- [10] 庄煜, 徐琨林, 李美琪, 等. 基于改进 YOLO v11n 的轻量化葡萄穗梗采摘点定位方法与试验[J]. *农业机械学报*, 2025, 56(9): 618-625.
- Zhuang Yu, Xu Kunlin, Li Meiqi, et al. Lightweight grape cluster stalk picking point localization method and test based on improved YOLO v11n [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2025, 56(9): 618-625.
- [11] Rong J C, Dai G L, Wang P B. A peduncle detection method of tomato for autonomous harvesting[J]. *Complex & Intelligent Systems*, 2022, 8(4): 2957-2968.
- [12] Ni J P, Zhu L C, Dong L Z, et al. TomatoPoseNet: An efficient keypoint-based 6D pose estimation model for non-destructive tomato harvesting[J]. *Agronomy*, 2024, 14(12): 3027.
- [13] Paul A, Machavaram R. Advanced segmentation models for automated capsicum peduncle detection in night-time greenhouse environments[J]. *Systems Science & Control Engineering*, 2024, 12(1): 2437162.
- [14] Zhang F, Gao J, Song C Y, et al. TPMv2: An end-to-end tomato pose method based on 3D key points detection[J]. *Computers and Electronics in Agriculture*, 2023, 210: 107878.
- [15] Li X X, Han Y H, Ma N, et al. Autonomous tomato harvesting with top-down fusion network for limited data[J]. *IEEE Transactions on Robotics*, 2025, 41: 3610-3627.
- [16] 刘韵婷, 刘欣然, 肖培宇, 等. 基于改进 YOLOv8 室内老人跌倒检测算法研究[J]. *电子测量与仪器学报*, 2026,40(2): 164-174.
- Liu Yunting, Liu Xinran, Xiao Peiyu, et al. Research on improved YOLOv8 algorithm for indoor elderly fall detection[J]. *Journal of Electronic Measurement and Instrumentation*, 2026,40(2): 164-174.
- [17] 李帆雅, 张泽辉, 陈博洋, 等. 面向复杂工业场景的人体姿态估计性能增强方法[J]. *仪器仪表学报*, 2025, 46(8): 255-260.
- Li Fanya, Zhang Zehui, Chen Boyang, et al. Human pose estimation performance enhancement method for complex industrial scenes [J]. *Chinese Journal of Scientific Instrument*, 2025, 46(8): 255-260.
- [18] Maji D, Nagori S, Mathew M, et al. YOLO-Pose: Enhancing YOLO for multi person pose estimation using object keypoint similarity loss[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops(CVPRW), 2022: 2637-2640.

- [19] 成彬, 张博豪, 雷华. 改进 YOLOv8-Pose 的钢筋焊接节点识别[J]. 计算机技术与发展, 2025, 35(2): 174-182.
Cheng Bin, Zhang Bohao, Lei Hua. Improvement of YOLOv8-Pose for identification of steel bar welding joints[J]. Computer Technology and Development, 2025, 35(2): 174-182.
- [20] 罗智杰, 王泽宇, 岑飘, 等. 基于改进 YOLOv8pose 的校园体测运动姿势识别研究[J]. 电子测量技术, 2024, 47(19): 24-28.
Luo Zhijie, Wang Zeyu, Cen Piao, et al. Research on human motion pose recognition algorithm based on improved YOLOv8pose[J]. Electronic Measurement Technology, 2024, 47(19): 24-28.
- [21] 朱周华, 侯智杰, 田成源, 等. 基于改进 YOLOv8-pose 的分心驾驶检测与识别[J]. 电子测量技术, 2024, 47(15): 135-138.
Zhu Zhouhua, Hou Zhijie, Tian Chengyuan, et al. Distracting driving detection and identification based on an improved YOLOv8-pose [J]. Electronic Measurement Technology, 2024, 47(15): 135-138.
- [22] Han K, Wang Y H, Tian Q, et al. GhostNet: More Features from Cheap Operations [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2020: 1580-1582.
- [23] Wang Q L, Wu B G, Zhu P F, et al. ECA-Net: Efficient channel attention for deep convolutional

neural networks [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 11534-11538.

- [24] Hou Q B, Zhou D Q, Feng J S. Coordinate attention for efficient mobile network design[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2021: 13713-13715.
- [25] Hu J, Shen L, Sun G, et al. Squeeze-and-excitation networks[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018: 7132-7135.

作者简介

刘安琪, 硕士研究生, 主要研究方向为模式识别与图像处理。

E-mail: 2023020375@bistu.edu.cn

杨鸿波(通信作者), 教授, 硕士生导师, 主要研究方向为模式识别与图像处理。

E-mail: anonbo@bistu.edu.cn

潘航, 博士, 讲师, 主要研究方向为人工智能与人机交互。

E-mail: panhang@czc.edu.cn

陈金龙, 高级实验师, 硕士生导师, 主要研究方向为三维室内定位、深度学习、手势识别、人机交互、机器视觉等。

E-mail: jinlong.chen@guet.edu.cn

湛永松, 教授, 主要研究方向为数字媒体技术、计算机图形图像技术、计算机游戏及 GIS 开发应用等。

E-mail: yszhan@guet.edu.cn