

面向莲蓬自动采摘的轻量级目标识别与精准定位方法

唐涛^{1,2} 叶秉良^{1,2} 胡淼^{1,2} 丰睿^{1,2} 俞高红^{1,2}

(1. 浙江理工大学机械工程学院, 杭州 310018; 2. 浙江省农业智能感知与机器人全省重点实验室, 杭州 310018)

摘要: 在莲蓬采摘过程中, 实现高效、精准的莲蓬检测与定位对提高采摘效率并降低误采风险具有重要意义, 但现有方法多依赖计算资源密集且结构复杂的深度学习模型, 难以满足现场实时应用需求。为此, 本文提出了一种适用于莲蓬采摘场景的轻量级目标检测与定位方法, 基于轻量级语义分割模型 (Lightweight lotus segmentation network, LLSegNet), 在 DeepLabv3+ 框架中采用 MobileNetV2 作为骨干网络。针对采摘过程莲蓬的多尺度变化、细节捕捉难度及背景干扰等问题, 引入密集空洞空间金字塔池化、条带池化、卷积块注意力及高效通道注意力网络等关键增强策略, 以提升不同尺度下特征提取与表达能力, 同时保持整体模型的轻量化。通过 Ubuntu 20.04 操作系统试验平台, 使用 PyTorch 2.3.1 框架对模型进行训练与评估。结果表明, LLSegNet 模型平均交并比 (mIoU) 和平均像素准确率 (mPA) 分别达到 86.1% 和 92.5%, 模型内存占用量为 15.9 MB, 帧率 (FPS) 为 73.4 f/s, 均优于主流语义分割模型。基于 LLSegNet 模型生成的高质量语义分割结果, 提出了一种结合图像处理与骨架分析的采摘点定位方法, 通过图像预处理、骨架提取、几何分析及法向量扩展映射, 实现了采摘点精准定位, 成功率为 88.5%。研究结果表明, 本文方法不仅提升了莲蓬识别与定位精度, 还实现了模型高效轻量化, 具备在资源受限农业环境中应用与进一步推广的潜力。

关键词: 莲蓬; 目标识别; 语义分割; 轻量级模型; 多尺度特征提取; 采摘点定位

中图分类号: TP391.4; S24

文献标识码: A

文章编号: 1000-1298(2026)08-0013-10

OSID:



Lightweight Target Recognition and Precise Localization Method for Automated Lotus Seedpod Harvesting

TANG Tao^{1,2} YE Bingliang^{1,2} HU Miao^{1,2} FENG Rui^{1,2} YU Gaohong^{1,2}

(1. School of Mechanical Engineering, Zhejiang Sci-Tech University, Hangzhou 310018, China)

2. Zhejiang Provincial Key Laboratory of Agricultural Intelligent Perception and Robotics, Hangzhou 310018, China)

Abstract: During lotus seedpod harvesting, efficient and precise detection and localization are essential for improving harvesting efficiency and minimizing the risk of mispicking. However, existing lotus seedpod recognition methods predominantly rely on computationally intensive and structurally complex deep learning models, rendering them impractical for real-time field applications. To address this limitation, a lightweight object detection and localization method optimized for lotus seedpod harvesting scenarios was proposed. The method was based on the lightweight lotus segmentation network (LLSegNet), a lightweight semantic segmentation model that employed MobileNetV2 as the backbone within the DeepLabv3+ framework. To cope with challenges such as multi-scale variation, difficulty in capturing fine details, and background interference during harvesting, key enhancement strategies, including dense atrous spatial pyramid pooling, strip pooling, convolutional block attention, and an efficient channel attention network. These improvements enhanced multi-scale feature extraction and representation while maintaining the lightweight nature of the overall model. Experiments were carried out on an Ubuntu 20.04 platform utilizing the PyTorch 2.3.1 framework for model training and evaluation. The results demonstrated that the LLSegNet model attained an average intersection over union (mIoU) of 86.1% and an average pixel accuracy (mPA) of 92.5%, with a memory footprint of 15.9 MB and a

收稿日期: 2025-01-16 修回日期: 2025-04-07

基金项目: 国家自然科学基金项目 (32171899)

作者简介: 唐涛 (1995—), 男, 助理研究员, 博士后, 主要从事智能农业装备研究, E-mail: ttzstu@163.com

通信作者: 叶秉良 (1972—), 男, 教授, 博士生导师, 主要从事农业机械设计和机构数值分析与综合研究, E-mail: zist_ybl@zstu.edu.cn

frame rate (FPS) of 73.4 f/s, all of which were superior to mainstream semantic segmentation models. Furthermore, leveraging the high-quality semantic segmentation results produced by the LLSegNet model, a harvesting point localization method that integrated image processing and skeleton analysis was proposed. This method accomplished precise localization of harvesting points through a combination of image preprocessing, skeleton extraction, geometric analysis, and normal vector expansion mapping, achieving a success rate of 88.5%. The findings demonstrated that the proposed method not only improved detection and localization accuracy but also remained computationally efficient, showing strong potential for deployment and further application in resource-constrained agricultural environments.

Key words: lotus seedpod; target recognition; semantic segmentation; lightweight model; multi-scale feature extraction; harvesting point localization

0 引言

中国在全球果蔬产业中占有重要地位,生产与出口规模均居世界前列。莲蓬作为一年生的特色水生经济作物,其种植规模正不断扩大。根据 2022 年统计,中国莲蓬总种植面积约为 $9.0045 \times 10^5 \text{ hm}^2$,产量达到 $2.02 \times 10^5 \text{ t}$,已成为全球主要的莲子生产和出口国之一^[1]。然而,随着社会人口老龄化和农业劳动力减少,莲蓬产业面临劳动力短缺和成本上升问题^[2]。传统的人工采摘方式由于劳动强度大、作业效率低,难以满足规模化种植需求^[3]。此外,莲蓬生长环境复杂,如光照不均、枝叶遮挡及与荷叶颜色差异不明显等因素,增加了采摘定位难度^[4]。莲蓬精准识别与定位对提高采摘成功率至关重要,也是研发莲蓬采摘机器人时需重点解决的关键问题。因此,为提升采摘效率、降低劳动力成本并确保作业质量,迫切需要研发一种适用于莲蓬采摘的农业机器人^[5-8]。

近年来,目标识别和定位技术迅猛发展,并在农业领域取得了显著成果^[9-11],为实现莲蓬识别与定位提供了研究基础。传统目标识别算法主要依赖边缘检测、阈值分割、形态学操作等方式,虽具备一定的简单性与可行性,但在自然环境下易受光照、背景等干扰^[12]。为此,研究人员对传统目标识别算法进行了多方面改进。孙建桐等^[13]采用 Canny 边缘检测对串收番茄图像进行分割,获取果实轮廓点,果实识别正确率达 85.1%,并在一定程度上解决了果实粘连分割问题。XING 等^[14]通过自适应非线性函数搜索与最优阈值分割获得红花二值图像,并基于改进的粒子群优化 (PSO) 算法与旋转矩形算法实现红花丝定位,定位精度达 89.75%,平均处理时间 0.14 s。JIN 等^[15]利用阈值分割、过滤和形态学操作等多种图像处理方法分割葡萄区域,再通过 Hough 线检测识别采摘点,在晴天条件下定位成功率为 90.51%。BAI 等^[16]利用 SVM 分类器,并结合形状、纹理、颜色特征,实现了番茄个体识别,通过霍夫圆检测和分

枝-果串几何分析等方法计算采摘点,准确率、召回率及 F1 值均达 100%,识别时间小于 1 s。尽管上述基于传统图像处理的识别与定位方法简单高效,但由于自然环境中的光照和背景变化多样,其鲁棒性尚显不足,难以实现大规模自动化应用^[17]。

随着深度学习与计算机硬件的发展,深度神经网络在众多视觉任务中展现出出色性能,特别是在图像分割领域^[18-19]。这类端到端算法不仅有效提升检测与分割精度,还具有较强鲁棒性^[20]。ZHU 等^[21]提出了一种改进的 EfficientDet 模型,用于果园环境中多品种橄榄果实成熟度检测,通过引入卷积注意力模块和加权双向特征金字塔网络,实现了对不同成熟期橄榄果实的成熟度检测。ZHENG 等^[22]基于 Mask R-CNN 对芒果与茎秆同时分割并测量采摘点,在无遮挡条件下准确率可达 82.4%,F1 值为 95%。周馨墨等^[23]提出了一种基于深度学习与葡萄关键结构多目标识别的采摘点定位方法,采摘点定位成功率为 90.51%。WANG 等^[24]提出了一种改进的 Faster R-CNN 模型,用于复杂环境下多目标番茄成熟度检测,平均精度达 96.14%。上述方法的精度、鲁棒性及适用性相较于传统机器视觉算法有大幅提升,但仍难以满足自动采摘实时作业需求,识别精度和运行效率还有待进一步提升。

虽然深度卷积神经网络在农业图像识别领域已取得广泛应用,但在莲蓬采收方面的研究仍较为有限。因此,本文提出一种轻量级莲蓬分割网络,命名为 LLSegNet (Lightweight lotus segmentation network),用于分割莲蓬和茎秆。该模型基于 DeepLabv3+ 架构,并在此基础上进行针对性改进,以提升在莲蓬采摘任务中的识别精度和鲁棒性,实现采摘点精准定位。

1 材料与方法

1.1 研究方法

提出了一种基于语义分割的莲蓬目标检测与定

位方法,总体流程如图 1 所示,主要包括 3 个阶段:①数据集构建。②莲蓬成熟度识别的轻量级语义分割模型开发。③莲蓬采摘点定位。首先,利用 LabelMe 对采集图像进行标注,构建高质量的训练数据集。然后,设计一种轻量级的莲蓬分割模型 LLSegNet,实现莲蓬及茎秆区域的精确分割,确保识

别过程的高效性和精度,以 MobileNetV2 为骨干网络,引入多种注意力机制和池化模块,提高在复杂环境中分割精度,同时确保模型轻量化。最后,通过对分割结果中茎秆区域提取骨架线,结合法向量扩展映射方法,精准定位采摘点,以满足自动采摘的准确性与实时性要求。

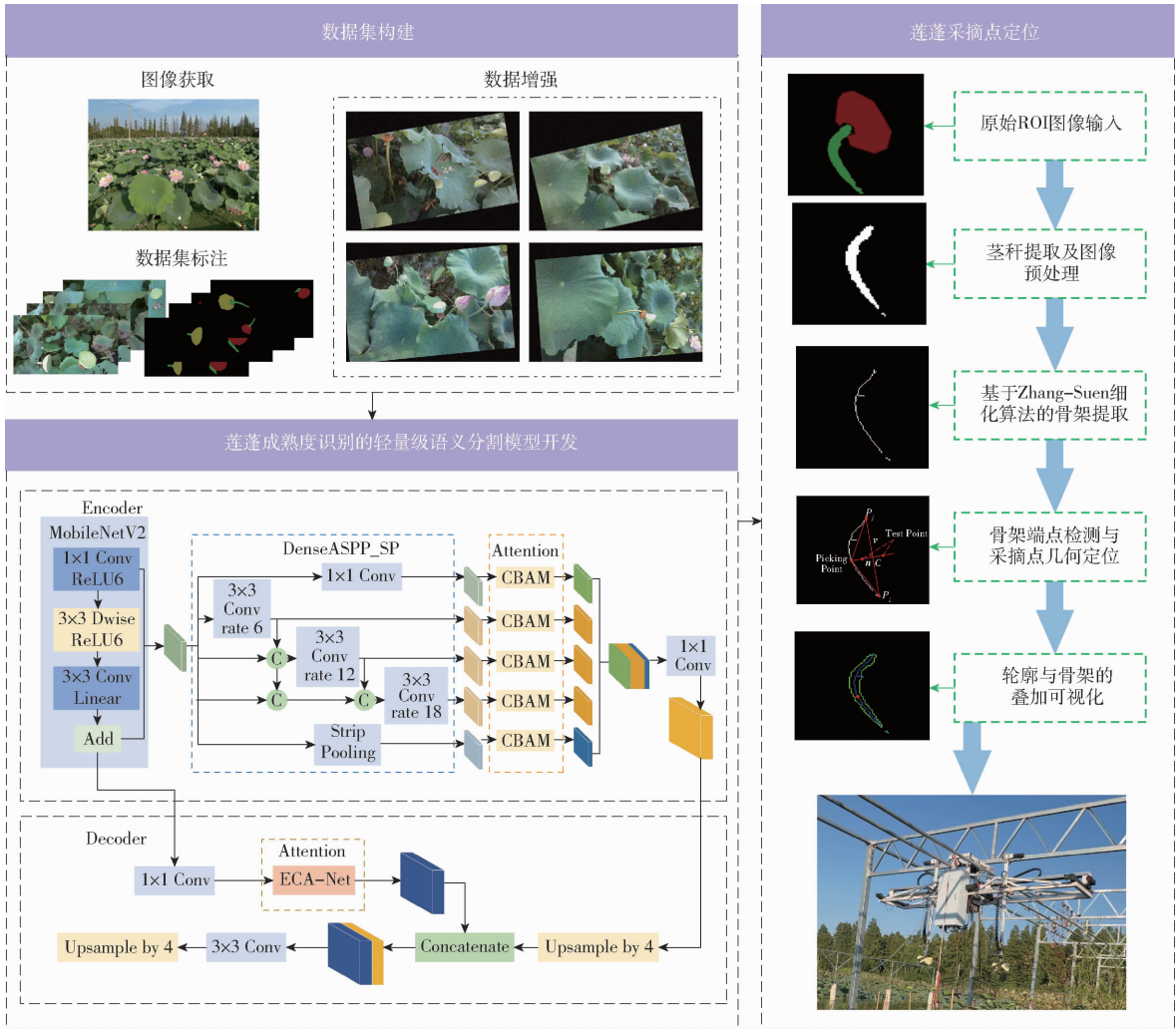


图 1 莲蓬目标识别与采摘点定位方法整体流程

Fig. 1 Overall process of lotus seedpod target recognition and harvesting point localization method

1.2 莲蓬采摘机器人设计

为实现高效且精准的莲蓬目标检测与定位,设计并搭建了一种莲蓬采摘机器人,由深度相机、上位机、下位机、机械臂、采摘末端和机械本体等部件构成,如图 2 所示。通过深度相机与机械手的协同工作,系统可实时获取并处理莲蓬图像信息。将深度相机安装在机械本体上并与水平面呈 90°,通过 USB 接口与上位机连接,用以获取莲蓬采摘点的三维坐标数据。获取的坐标信息通过串口通信传输至下位机,以驱动机械臂和采摘末端进行精确定位,最终完成对莲蓬的包络及茎秆切割操作。

1.3 图像采集和数据集构建

图像采集于浙江省杭州市建德市大慈岩镇荷田



图 2 莲蓬采摘机器人

Fig. 2 Lotus seedpod picking robot

1. 深度相机 2. 机械臂 3. 采摘末端 4. 控制柜(上位机和下位机) 5. 机械本体

种植基地。拍摄时间为 2023 年 7—8 月,该时段正值莲蓬生长与成熟的关键期。为提高数据集多样性,图像采集涵盖多种天气与光照条件。每日 07:00—18:00 使用搭载在 DJI Mini Air2 型无人机

上的相机进行图像采集。为确保数据集与实际莲蓬采摘机器人的视觉系统一致,使用的无人机相机布置角度与采摘机器人相机安装角度相同,均为与水平面呈 90° 。此外,无人机飞行高度与采摘机器人相机实际安装高度一致。无人机以速度 1 m/s 飞行,以 30 f/s 记录 $1\,920\text{ 像素}\times 1\,080\text{ 像素}$ 视频。试验场地及飞行路径示意图如图 3 所示。



图 3 图像采集过程示意图

Fig.3 Schematic of image acquisition process

为提升自然环境下莲蓬成熟度识别的准确率,首先将无人机以 1 f/s 录制的视频逐帧转换为静态图像,并对其进行筛选与处理,剔除震动、光斑及失焦等可能干扰检测算法的低质量图像。经筛选后共保留 $4\,978$ 幅高质量图像,并统一调整至 $512\text{ 像素}\times 512\text{ 像素}$ 以降低训练计算量。随后,采用 LabelMe 对图像进行标注,如图 4a 所示,将标注文件转换为对应的掩码标签图像,如图 4b 所示,其中红色和黄色分别表示未成熟和成熟莲蓬,绿色表示茎秆。为增强模型性能和泛化能力,同时降低过拟合风险,将图像集按比例 $7:1:2$ 划分为训练集、验证集与测试集。针对训练集,采用数据增强策略,通过裁剪、平移、亮度调节、噪声添加以及旋转等操作的随机组合(图 4c),共生成 $10\,455$ 幅增强图像,最终训练集包含 $13\,940$ 幅图像,而验证集和测试集则保持原始图像数量不变。最终,构建 $15\,433$ 幅图像的完整数据集。

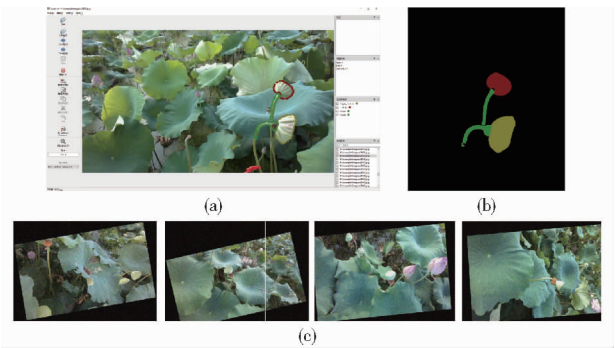


图 4 图像数据集建立过程示意图

Fig.4 Schematics of image dataset creation process

1.4 莲蓬成熟度分割与特征提取方法

为在自然环境下高效且精准地区分莲蓬成熟度,提出新的语义分割模型 LLSegNet,如图 1 所示。模型架构在 DeepLabv3+ 基础上进行优化性改进,以提升其在莲蓬采摘任务的表现。首先,为满足对

检测效率的较高要求,采用 MobileNetV2 作为骨干网络,在充分提取图像特征的同时,有效减少计算量。随后,引入密集空间空洞金字塔池化(Dense atrous spatial pyramid pooling, DenseASPP)和条形池化(Stripe pooling, SP)模块,并在各卷积层后添加卷积块注意力模块(Convolutional block attention module, CBAM),以强化对多尺度与方向性特征的捕捉及在复杂背景中的识别效果。在解码部分,为进一步增强特征表达,在 MobileNetV2 底层特征后增加高效通道注意力网络(Efficient channel attention network, ECA-Net),从而突出关键通道信息并提升莲蓬特征的表达能力。

1.4.1 轻量级骨干网络

由于 DeepLabv3+ 模型的原骨干网络 Xception 参数量较大,在自然环境下应用时会导致总体计算量和推理时间明显增加。为解决这一问题,引入轻量化的 MobileNetV2 网络作为 LLSegNet 模型的骨干网络。MobileNetV2 以深度可分离卷积替代标准卷积,在保证准确率基本不变的情况下,将 3×3 卷积计算量削减约 90% [25]。

此外,MobileNetV2 引入倒置残差结构,进一步提升了网络性能。在倒置残差模块中,输入特征通道首先通过 1×1 卷积进行扩展,然后通过 3×3 深度可分离卷积提取特征信息,最后通过 1×1 卷积恢复特征通道数量,如图 5 所示。

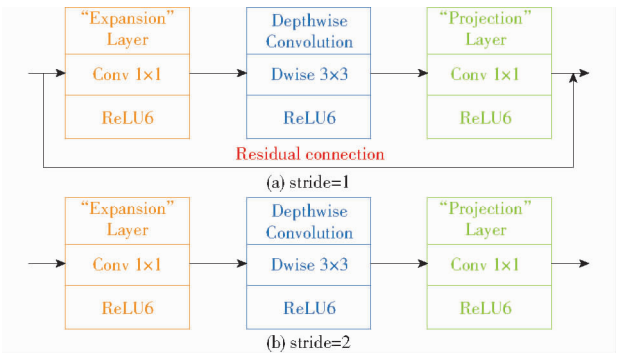


图 5 MobileNetV2 倒置残差结构图

Fig.5 Inverted residual structure of MobileNetV2

为减少特征丢失,MobileNetV2 在通道压缩时采用线性激活,并在扩展层中以线性激活替换常规非线性激活,从而更有效地保留形状、颜色与纹理等关键信息,适用于莲蓬目标的细粒度特征提取。尤其在分割过程,LLSegNet 模型不仅能准确区分莲蓬与茎秆,还可基于成熟度分类实现对不同生长阶段的标注。通过上述改进,MobileNetV2 在资源受限环境下保持了较高效率与性能。其网络架构如表 1 所示,其中 t 为扩展因子, c 为输出通道数, n 为当前层重复次数, s 为步距, k 为分类类别数。

表 1 MobileNet V2 网络架构

Tab. 1 Structure of MobileNetV2 network

输入	算子	<i>t</i>	<i>c</i>	<i>n</i>	<i>s</i>
224 × 224 × 3	Conv2d		32	1	2
112 × 112 × 32	Bottleneck	1	16	1	1
112 × 112 × 16	Bottleneck	6	24	2	2
56 × 56 × 24	Bottleneck	6	32	3	2
28 × 28 × 32	Bottleneck	6	64	4	2
14 × 14 × 64	Bottleneck	6	96	3	1
14 × 14 × 96	Bottleneck	6	160	3	2
7 × 7 × 160	Bottleneck	6	320	1	2
7 × 7 × 320	Conv2d 1 × 1		1 280	1	1
7 × 7 × 1 280	Avgpool 7 × 7			1	
1 × 1 × 1 280	1 × 1		<i>k</i>		

1.4.2 多尺度特征提取网络

为进一步增强网络的多尺度特征提取能力并实现对莲蓬的精准分割与成熟度分类,将 DenseASPP 模块代替原 ASPP 模块,同时结合 SP 条带池化模块,形成 DenseASPP – SP 结构,如图 6 所示。此外,为减小模型内存占用量,在 DenseASPP – SP 模块中将普通卷积替换为深度可分离卷积 (Depthwise separable convolution)。利用空洞卷积在不同感受野上捕捉多尺度特征,并通过密集连接将这些特征融合、传递至后续层,从而在降低模型复杂度的同时充分利用多尺度信息,更有效地处理细节与上下文。

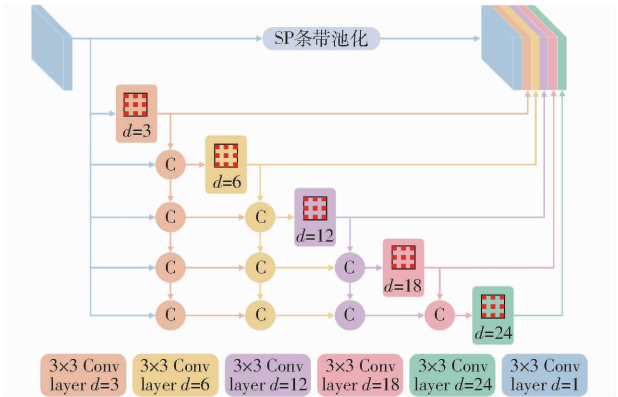


图 6 DenseASPP – SP 模块

Fig. 6 DenseASPP – SP module

在 DenseASPP 模块中,网络并行连接 6 个不同尺度特征层。输入的深度特征首先被传递至输出位置,随后依次执行扩张率 d 分别为 3、6、12、18、24 的深度可分离空洞卷积。每个空洞卷积层的输入由之前所有卷积层的输出与初始特征连接而成,产生多尺度特征图。与原始 ASPP 模块相比, DenseASPP 通过密集连接堆叠所有深度可分离扩张卷积,在减少参数量与计算量的同时,增强了特征表达,对莲蓬分割及成熟度分类尤为关键。

在 DenseASPP 基础上引入的 SP 条带池化模块

可克服全局平均池化在捕捉不同方向图像比例相关性时的不足。SP 使用矩形池化窗口,从水平与垂直 2 个维度获取全局信息,并采用加权平均分别获得列或行特征值,如图 7 所示,其中 H 和 W 分别表示高度和宽度。相较全局平均池化,SP 能同时关注水平与垂直空间维度,从而提升分割精度并为后续成熟度分类提供更丰富的特征支持。

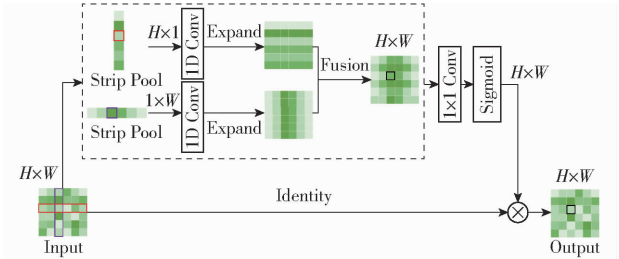


图 7 SP 模块

Fig. 7 SP module

1.4.3 CBAM 注意力模块

在多尺度空洞卷积中,较大的采样扩张率虽然能捕捉更丰富的目标信息,却可能削弱网络对图像边缘特征及目标关联性的表征,尤其在遮挡情况下对莲蓬特征的表示较弱。为此,引入 CBAM 注意力机制,其由通道注意力模块 (Channel attention module, CAM) 和空间注意力模块 (Spatial attention module, SAM) 组成。图 8 为 CBAM 在莲蓬特征图 F 上的处理流程。通过全局池化与维度操作得到一维通道注意力特征权重图 $M_c(F)$,利用池化、通道拼接和 Sigmoid 激活生成空间注意力特征权重图 $M_c(F')$,最后,空间注意力权重图与莲蓬特征图 F' 相乘,得到多维信息集成特征图 F'' 。

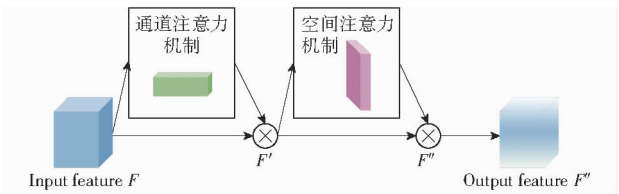


图 8 CBAM 模块

Fig. 8 CBAM module

CAM 结构如图 9 所示。首先,将输入尺寸为 $H \times W \times C$ (C 为通道数) 特征图 F 分别经过平均池化和最大池化聚合特征空间信息,得到 $1 \times 1 \times C$ 特征图。然后,将其输入 2 层共享的多层感知机 (Multi-layer perceptron, MLP) 网络并逐元素相加,再经 Sigmoid 激活生成通道注意力特征权重图 $M_c(F)$ 。最后, $M_c(F)$ 与原特征图 F 相乘,输出通道加权后的特征图 F' 。

SAM 结构如图 10 所示。首先,将 CAM 输出的特征图 F' 作为输入特征,特征图尺寸为 $H \times W \times C$,经过基于通道的最大池化和平均池化,分别生成

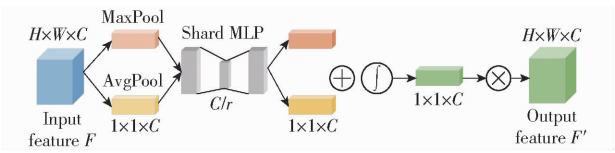


图 9 通道注意力模块

Fig. 9 CAM module

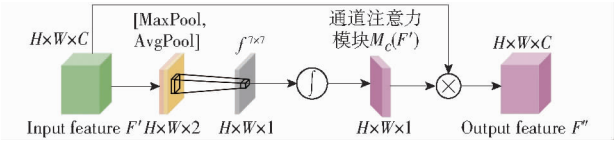


图 10 空间注意力模块

Fig. 10 SAM module

2 个尺寸为 $H \times W \times 1$ 的特征图。随后,将这 2 个特征图沿通道维度拼接,经过 7×7 卷积操作,降维至 $H \times W \times 1$,并通过 Sigmoid 激活函数生成空间注意力特征权重图 $M_c(F')$ 。最后,将空间注意力特征权重图 $M_c(F')$ 与输入特征图 F' 相乘,得到最终经过空间注意力机制加权后的特征图 F'' 。

1.4.4 ECA-Net 注意力模块

为增强网络的通道注意力表达能力并优化计算效率,在 MobileNetV2 浅层特征中嵌入 ECA-Net^[26]。如图 11 所示,作为 SE-Net(Squeeze and excitation networks)的改进,ECA-Net 通过将全连接层替换为一维卷积层,显著降低了计算量和参数量,同时避免了降维操作带来的信息损失,从而实现更高效、灵活的通道注意力机制。

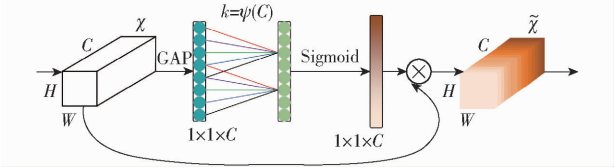


图 11 有效通道注意力模块

Fig. 11 ECA-Net module

MobileNetV2 虽因轻量化与高效性而适合资源受限环境,但在细粒度信息和局部特征捕捉方面仍存在局限。考虑到采用了预训练的骨干网络,若直接在深层集成 ECA-Net,可能会破坏原有结构和预训练权重,从而影响模型性能。因此,选择将 ECA-Net 嵌入浅层特征中。这一策略不仅保持了预训练网络的完整性,还在不增加计算负担的前提下,显著增强了浅层特征的表达能力,提升了模型在分割与分类任务中的精度。

1.5 采摘点定位策略

为实现莲蓬采摘点的精准定位并克服茎秆弯曲带来的偏移问题,提出基于图像处理与骨架分析的定位方法,包括图像预处理、骨架提取、几何分析以及基于法向量扩展映射。该方法可将采摘点准确映

射至茎秆骨架上,从而提高定位精度并保持较高鲁棒性。

利用本文提出的 LLSegNet 语义分割模型获取莲蓬与茎秆的感兴趣区域(RoI of interest, ROI)并对 ROI 进行灰度化、二值化和形态学操作,生成清晰的二值化区域图。采用 Zhang-Suen 细化算法提取茎秆骨架^[27],将其简化为单像素宽度的连通路径,为几何分析提供便利。为进一步适应不同茎秆形态及姿态,提出法向量扩展映射方法,计算骨架两端点的中点并作为初步采摘点;再通过计算法向量扩展路径,进一步修正初步计算的采摘点,确保精度。该方法在茎秆采摘点定位的灵活性上表现出色,能够适应不同茎秆形态,操作过程如图 12 所示。

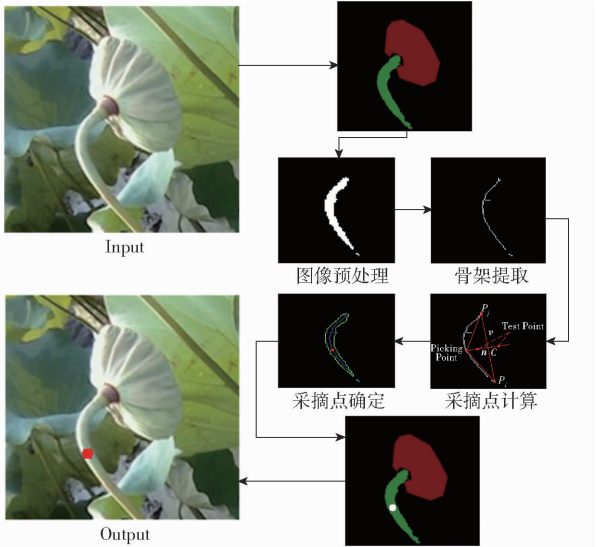


图 12 莲蓬采摘点定位算法流程图

Fig. 12 Algorithm flowchart for lotus seedpod harvesting point localization

具体步骤如下:

(1)端点确定:计算骨架上所有像素点对的欧几里得距离,选出最大距离对应的点对作为骨架端点。令骨架点集合为 $\{P_1, P_2, \dots, P_n\}$,其中每个点 P_i 的二维坐标为 (x_i, y_i) ,计算式为

$$d(P_i, P_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (1)$$

式中 $d(P_i, P_j)$ ——点 P_i, P_j 距离

(2)中心点计算:计算 2 个端点的中点,作为骨架中心点,并将其作为初始采摘点。中点 C_1 坐标为

$$C_1 = \left(\frac{x_i + x_j}{2}, \frac{y_i + y_j}{2} \right) \quad (2)$$

(3)方向向量与法向量生成:计算 2 个端点之间方向向量 $\mathbf{v}$,并根据该向量生成其法向量 $\mathbf{n}$ (即垂直向量)。方向向量 $\mathbf{v}$ 计算式为

$$\mathbf{v} = (x_j - x_i, y_j - y_i) \quad (3)$$

根据方向向量 $\mathbf{v}$ 生成法向量 $\mathbf{n}$,公式为

$$\boldsymbol{n} = (-(y_j - y_i), x_j - x_i)$$

(4)

为保证方向一致性并简化计算过程,对法向量 $\boldsymbol{n}$ 进行归一化处理,即

$$\hat{\boldsymbol{n}} = \frac{\boldsymbol{n}}{\|\boldsymbol{n}\|}$$

(5)

(4)测试点生成:沿归一化垂直向量的正负方向,按预设步长 d_1 逐步移动,生成 2 个测试点 TestPoint1 和 TestPoint2,计算式为

$$d_{\text{TestPoint}_1} = C_1 + d_1 \hat{\boldsymbol{n}}$$

(6)

$$d_{\text{TestPoint}_2} = C_1 - d_1 \hat{\boldsymbol{n}}$$

(7)

(5)最近点映射:对每个测试点与骨架中所有点计算欧几里得距离,选择最近的骨架点作为最终采摘点。最终采摘点计算式为

$$P = \underset{P_k \in \{P_1, P_2, \dots, P_n\}}{\operatorname{argmin}} d(d_{\text{TestPoint}}, P_k)$$

(8)

2 结果与分析

2.1 试验平台

在 x86_64 架构的 Ubuntu 20.04 环境下开展,硬件包含 NVIDIA GeForce RTX 3090 显卡和 Intel® Xeon® Silver 4310 CPU@ 2.10 GHz 处理器。采用 PyTorch 2.3.1 深度学习框架,结合 CUDA 12.1 进行加速,并使用 Python 3.9 作为编程语言。

2.2 模型训练及参数设置

在训练前,将输入图像尺寸统一调整为 512 像素 $\times$ 512 像素。训练阶段参数如下:迭代次数 (Epochs) 为 100 次,批次大小 (Batch size) 为 16,初始学习率为 7×10^{-3} ,学习率衰减至 7×10^{-5} ,优化算法采用随机梯度下降 (SGD)。每 5 个迭代次数进行一次验证并保存当前模型。

2.3 评估指标

为准确评估模型性能,采用平均交并比 (Mean intersection over union, mIoU) 和平均像素准确率 (Mean pixel accuracy, mPA) 作为模型语义分割精度

的主要指标。

2.4 模型性能评估与验证

2.4.1 消融试验

为验证本文提出的 3 种增强策略 (DenseASPP-SP、CBAM 及 ECA-Net) 对 DeepLabv3+ 模型性能的影响,采用控制变量法设计 9 种不同模型配置,并通过消融试验对其性能进行系统评估,试验结果见表 2。结果表明,基线模型 1 (仅使用 DeepLabv3+) mIoU 为 84.3%,模型内存占用量为 209 MB。引入 MobileNetV2 后 (模型 2) 将模型内存占用量减小至 22.4 MB,且 mIoU 小幅提升至 84.7%,表明 MobileNetV2 在减小参数量的同时保持了较强的特征学习能力。进一步引入 DenseASPP-SP (模型 3),将模型内存占用量压缩至 15.7 MB,并将 mIoU 提升至 85.2%,表明该多尺度模块有助于捕捉不同空间尺度的关键信息。引入 CBAM (模型 4) 和引入 ECA-Net (模型 5),mIoU 分别提升至 85.5% 和 85.8%,mPA 分别达到 91.8% 和 92.1%,表明注意力机制能突出重要特征并增强表达效果。此外,多模块组合 (模型 6 ~ 模型 9) 进一步提升了 mIoU 和 mPA,其中模型 9 同时引入 DenseASPP-SP、CBAM 和 ECA-Net,mIoU 和 mPA 分别提升至 86.1% 和 92.5%,且模型内存占用量为 15.9 MB。由此可见,这 3 种增强策略的协同作用不仅显著提升了模型分割性能,还有效控制了模型体积,验证了各模块在整体模型中的有效性。综合消融试验结果显示,所提出的增强策略不仅有效提高了分割精度与分类准确率,也显著降低了模型存储需求,为实时应用的快速高精度目标定位与分割提供了坚实基础。

2.4.2 不同模型性能对比

为进一步验证改进模型性能,将其与主流语义分割模型进行比较分析,结果如表 3 所示。试验结果表明,本文模型在 mIoU、mPA、模型内存占用量和

表 2 消融试验结果
Tab. 2 Ablation test results

模型	DeepLabv3 +	MobileNetV2	DenseASPP - SP	CBAM	ECA - Net	mIoU/%	mPA/%	模型内存占用量/MB
1	√					84.3	91.0	209.0
2	√	√				84.7	90.9	22.4
3	√	√	√			85.2	91.5	15.7
4	√	√		√		85.5	91.8	22.6
5	√	√			√	85.8	92.1	22.4
6	√	√	√	√		85.6	91.9	15.9
7	√	√	√		√	85.7	92.0	15.7
8	√	√		√	√	86.0	92.3	22.6
9	√	√	√	√	√	86.1	92.5	15.9

注:√表示采用策略。

帧率等指标上显著优于其他 3 种模型。本文模型 mIoU 和 mPA 分别为 86.1% 和 92.5%，较 DeepLabv3 + 模型提高 1.8、1.5 个百分点，较 PSPNet 模型提高 1.1、1.4 个百分点，较 U - Net 模型提高 0.9、1.0 个百分点。此外，本文模型内存占用量为 15.9 MB，相较于 DeepLabv3 + 、PSPNet 和 U - Net，分别减小约 92%、91% 和 90%。同时，本文模型检测速度为 73.4 f/s，显著高于 DeepLabv3 + 、PSPNet 和 U - Net。

表 3 不同模型性能

Tab.3 Performance of different models

模型	mIoU/ %	mPA/%	模型内存占 用量/MB	帧率/ (f·s ⁻¹)
DeepLabv3 +	84.3	91.0	209.0	28.1
PSPNet	85.0	91.1	178.0	49.8
U - Net	85.2	91.5	167.0	61.7
本文模型	86.1	92.5	15.9	73.4

综合分析表明,在维持或提升分割精度的同时,本文算法实现了较高轻量化与计算效率,尤其适合在资源受限设备中部署,可有效满足实际应用场景对于高效、便携的需求。

不同模型分割性能如图 13 所示。DeepLabv3 + 和 U - Net 在分类过程存在一定误差,将同一莲蓬同时标记为成熟和不成熟,导致分类结果不准确。PSPNet 在检测莲蓬边缘时存在不完整现象,影响了分割精度。相比之下,本文模型在分割性能上表现优越,不仅能实现更平滑的莲蓬和茎秆边缘分割,还显著减少了漏分和误分。因此,本文模型在整体分割效果上明显优于其他 3 种模型,提供了最为精确和可靠的分割结果。

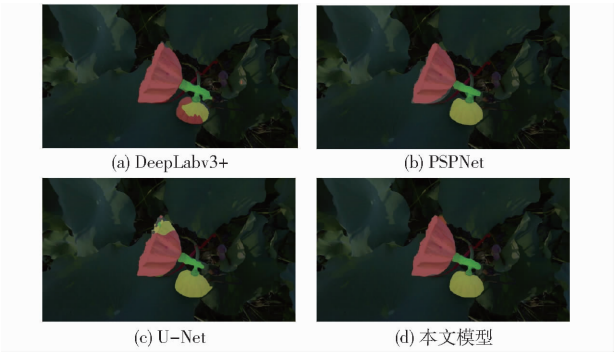


图 13 不同模型检测效果

Fig.13 Detection results for different models

2.4.3 采摘点定位性能评估

为验证采摘点定位方法的可靠性,从测试数据集中随机抽取 100 幅图像开展定位试验,共包含 148 根莲蓬茎秆,覆盖多样环境与莲蓬姿态。利用 LLSegNet 语义分割模型获取莲蓬与茎秆的 ROI,对

图像进行灰度化、二值化及形态学处理,得到更清晰的二值化图像。在此基础上,通过骨架提取简化茎秆结构,并采用法向量扩展映射对采摘点进行精确定位。图 14 为定位效果示例,试验结果表明,131 个样本被成功定位,成功率为 88.5%。然而,17 个样本未能成功定位采摘点,主要原因在于茎秆未能准确识别(图 14e)。该情况通常发生在茎秆被严重遮挡或形态不规则时,导致分割过程茎秆的起始和终止位置未能被识别,从而无法进一步确定采摘点,影响了茎秆的精确分割,并导致定位准确性降低。试验结果表明,所提出的采摘点定位方法具有较强鲁棒性,能够有效应对复杂环境条件下的莲蓬采摘任务。

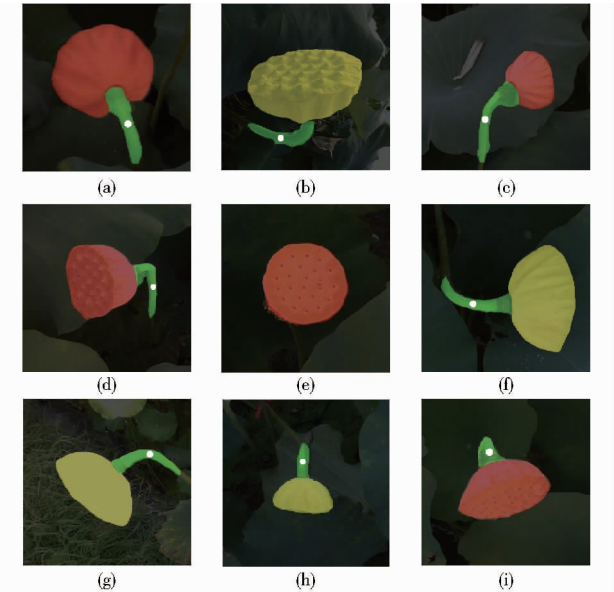


图 14 不同角度莲蓬采摘点定位效果

Fig.14 Lotus seedpod harvesting point localization results at different angles

此外,本研究前期在田间环境中采用与真实莲蓬相似的仿真莲蓬进行初步验证,如图 15 所示。结果表明所提出的采摘点定位方法具有较高的准确性和鲁棒性。然而,由于实际莲蓬生长周期长及试验条件限制,目前尚未开展大规模采摘平台试验。未来将利用现有平台通过多个地点和不同生长条件进一步开展验证试验,以全面评估该方法在真实田间环境中的性能和鲁棒性。



图 15 田间试验

Fig.15 Field experiment

3 结束语

提出基于 DeepLabv3 + 架构的轻量级分割模型 LSegNet,融合多种增强策略,在提升多尺度特征提取与表达能力的同时保持较低计算复杂度。试验结果表明,LSegNet 模型在多个关键性能指标上表现优异,mIoU 和 mPA 分别为 86.1% 和 92.5%,模型内存占用量仅为 15.9 MB,帧率为 73.4 f/s,相较主流模型具备更

高精度与更优效率,适合在资源受限的移动设备或农业自动化环境中部署。此外,结合图像处理与骨架分析的采摘点定位方法在测试中定位成功率达 88.5%,显示出在复杂农业场景中的可行性与应用潜力。所提出的轻量级目标识别与精准定位方法有效提升了莲蓬识别与分割精度,并实现了高效轻量化部署,为资源受限环境下的农业自动化设备提供了技术支撑,也为莲蓬自动采摘系统研发奠定了基础。

参 考 文 献

[1] TANG T, WANG X, MA Z, et al. Lightweight detection method for lotus seedpod in natural environment[J]. International Journal of Agricultural and Biological Engineering, 2023, 16(6): 197–206.

[2] REDMOND S R, CORNELIA W, IBRAHIM H A, et al. Research and development in agricultural robotics: a perspective of digital farming[J]. International Journal of Agricultural and Biological Engineering, 2018, 11(4): 1–11.

[3] 赵春江, 范贝贝, 李瑾, 等. 农业机器人技术进步、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 and trends[J]. Smart Agriculture, 2023, 5(4): 1–15. (in Chinese)

[4] TANG Y C, CHEN M, WANG C, et al. Recognition and localization methods for vision-based fruit picking robots: a review[J]. Frontiers in Plant Science, 2020, 11: 510.

[5] 刘继展, 江应星. 农业采摘机器人产业化进程分析与多臂高速化技术走向[J]. 农业机械学报, 2024, 55(10): 1–17.
LIU Jizhan, JIANG Yingxing. Industrialization trends and multi-arm technology direction of harvesting robots[J]. Transaction of the Chinese Society for Agricultural Machinery, 2024, 55(10): 1–17. (in Chinese)

[6] MONTOYA-CAVERO L E, DE LEÓN TORRES R D, GÓMEZ-ESPINOSA A, et al. Vision systems for harvesting robots: produce detection and localization[J]. Computers and Electronics in Agriculture, 2021, 192: 106562.

[7] 刘继展. 温室采摘机器人技术研究进展分析[J]. 农业机械学报, 2017, 48(12): 1–18.
LIU Jizhan. Research progress analysis of robotic harvesting technologies in greenhouse[J]. Transaction of the Chinese Society for Agricultural Machinery, 2017, 48(12): 1–18. (in Chinese)

[8] MA J, LU A E, CHEN C, et al. YOLOv5-lotus an efficient object detection method for lotus seedpod in a natural environment[J]. Computers and Electronics in Agriculture, 2023, 206: 107635.

[9] TANG Y, QIU J, ZHANG Y, et al. Optimization strategies of fruit detection to overcome the challenge of unstructured background in field orchard environment: a review[J]. Precision Agriculture, 2023, 24(4): 1183–1219.

[10] 项荣, 应义斌, 蒋焕煜. 田间环境下果蔬采摘快速识别与定位方法研究进展[J]. 农业机械学报, 2013, 44(11): 208–223.
XIANG Rong, YING Yibin, JIANG Huanyu. Development of real-time recognition and localization methods for fruits and vegetables in field[J]. Transaction of the Chinese Society for Agricultural Machinery, 2013, 44(11): 208–223. (in Chinese)

[11] 郑太雄, 江明哲, 冯明驰. 基于视觉的采摘机器人目标识别与定位方法研究综述[J]. 仪器仪表学报, 2021, 42(9): 28–51.
ZHENG Taixiong, JIANG Mingzhe, FENG Mingchi. Vision based target recognition and location for picking robot: a review[J]. Chinese Journal of Scientific Instrument, 2021, 42(9): 28–51. (in Chinese)

[12] 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]. Transaction of the Chinese Society for Agricultural Machinery, 2020, 51(9): 1–17. (in Chinese)

[13] 孙建桐, 孙意凡, 赵然, 等. 基于几何形态学与迭代随机圆的番茄识别方法[J]. 农业机械学报, 2019, 50(增刊): 22–26,61.
SUN Jiantong, SUN Yifan, ZHAO Ran, et al. Tomato recognition method based on iterative random circle and geometric morphology[J]. Transaction of the Chinese Society for Agricultural Machinery, 2019, 50(Supp.): 22–26,61. (in Chinese)

[14] XING Z, ZHANG Z, SHI R, et al. Filament-necking localization method via combining improved PSO with rotated rectangle algorithm for safflower-picking robots[J]. Computers and Electronics in Agriculture, 2023, 215: 108464.

[15] JIN Y, YU C, YIN J, et al. Detection method for table grape ears and stems based on a far-close-range combined vision system and hand-eye-coordinated picking test[J]. Computers and Electronics in Agriculture, 2022, 202: 107364.

[16] BAI Y, MAO S, ZHOU J, et al. Clustered tomato detection and picking point location using machine learning-aided image analysis for automatic robotic harvesting[J]. Precision Agriculture, 2023, 24(2): 727–743.

[17] RONG Q, HU C, HU X, et al. Picking point recognition for ripe tomatoes using semantic segmentation and morphological

processing[J]. Computers and Electronics in Agriculture, 2023, 210: 107923.

[18] KOIRALA A, WALSH K B, WANG Z L, et al. Deep learning-method overview and review of use for fruit detection and yield estimation[J]. Computers and Electronics in Agriculture, 2019, 162: 219 – 234.

[19] MINAEE S, BOYKOV Y Y, PORIKLI F, et al. Image segmentation using deep learning: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(7): 3523 – 3542.

[20] 孙红, 李松, 李民赞, 等. 农业信息成像感知与深度学习应用研究进展[J]. 农业机械学报, 2020, 51(5): 1 – 17.
SUN Hong, LI Song, LI Minzan, et al. Research progress of image sensing and deep learning in agriculture[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(5): 1 – 17. (in Chinese)

[21] ZHU X, CHEN F, ZHANG X, et al. Detection the maturity of multi-cultivar olive fruit in orchard environments based on Olive-EfficientDet[J]. Scientia Horticulturae, 2024, 324: 112607.

[22] ZHENG C, CHEN P, PANG J, et al. A mango picking vision algorithm on instance segmentation and key point detection from RGB images in an open orchard[J]. Biosystems Engineering, 2021, 206(6): 32 – 54.

[23] 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 Xinzhaoy, WU Fengyun, ZOU Xiangjun, et al. Method for locating picking points of grape clusters using multi-object recognition[J]. Transactions of the CSAE, 2023, 39(22): 166 – 177. (in Chinese)

[24] WANG C, WANG C, WANG L, et al. A lightweight cherry tomato maturity real-time detection algorithm based on improved YOLOV5n[J]. Agronomy, 2023, 13(8): 2106.

[25] MA D M, LI P H, HUANG X Y, et al. Efficient semantic segmentation based on improved DeepLabV3[J]. Computer Engineering & Science, 2022, 44(4): 737 – 745.

[26] WANG Q, WU B, ZHU P, et al. ECA – Net: efficient channel attention for deep convolutional neural networks[C] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020.

[27] ZHANG T Y, SUEN C Y. A fast parallel algorithm for thinning digital patterns[J]. Communications of the ACM, 1984, 27(3): 236 – 239.

~~~~~

(上接第 12 页)

[14] 劳彩莲, 李鹏, 冯宇. 基于改进 A\* 与 DWA 算法融合的温室机器人路径规划[J]. 农业机械学报, 2021, 52(1): 14 – 22.  
LAO Cailian, LI Peng, FENG Yu. Path planning of greenhouse robot based on fusion of improved A\* algorithm and dynamic window approach[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(1): 14 – 22. (in Chinese)

[15] 熊俊涛, 李中行, 陈淑绵, 等. 基于深度强化学习的虚拟机器人采摘路径避障规划[J]. 农业机械学报, 2020, 51(增刊 2): 1 – 10.  
XIONG Juntao, LI Zhongxing, CHEN Shumian, et al. Obstacle avoidance planning of virtual robot picking path based on deep reinforcement learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(Supp. 2): 1 – 10. (in Chinese)

[16] WANG Y, LU C, WU P, et al. Path planning for unmanned surface vehicle based on improved Q – Learning algorithm[J]. Ocean Engineering, 2024, 292: 116510.

[17] MAOUDJ A, HENTOUT A. Optimal path planning approach based on Q-Learning algorithm for mobile robots[J]. Applied Soft Computing, 2020, 97: 106796.

[18] 李潇宇, 张君华, 郭晓光, 等. 基于强化学习的机器人底盘能量管理与路径规划优化算法[J]. 农业工程学报, 2024, 40(21): 175 – 183.  
LI Xiaoyu, ZHANG Junhua, GUO Xiaoguang, et al. Reinforcement learning-based optimization algorithm for energy management and path planning of robot chassis[J]. Transactions of the CSAE, 2024, 40(21): 175 – 183. (in Chinese)

[19] ZHOU Q, LIAN Y, WU J, et al. An optimized Q – Learning algorithm for mobile robot local path planning[J]. Knowledge-Based Systems, 2024, 286: 111400.

[20] LOW E S, ONG P, LOW C Y, et al. Modified Q-Learning with distance metric and virtual target on path planning of mobile robot[J]. Expert Systems with Applications, 2022, 199: 117191.