doi:10.6041/j.issn.1000-1298.2024.02.024

基于改进 YOLO v5 的复杂环境下桑树枝干识别定位方法

李　丽[1,2]　卢世博[1,2]　任　浩[1,2]　徐　刚[3]　周永忠[4]

(1.西南大学工程技术学院, 重庆 400715; 2.西南大学宜宾研究院, 宜宾 644000;
3.北京市农林科学院智能装备技术研究中心, 北京 100097; 4.重庆祥飞智能科技有限公司, 重庆 401121)

摘要: 为实现复杂自然环境下对桑树嫩叶处枝干的识别检测,改变当前桑叶采摘设备作业过程中依赖人工辅助定位的现状,解决识别目标姿态多样和环境复杂导致的低识别率问题,提出一种基于改进 YOLO v5 模型的桑树枝干识别模型(YOLO v5-mulberry),并结合深度相机构建定位系统。首先,在 YOLO v5 的骨干网络中加入 CBAM(Convolutional block attention module)注意力机制,提高神经网络对桑树枝干的关注度;并增加小目标层使模型可检测 4 像素 ×4 像素的目标,提高了模型检测小目标的性能;同时使用 GIoU 损失函数替换原始网络中的 IoU 损失函数,有效防止了预测框和真实框尺寸较小时无法正确反映预测框及真实框之间位置关系的情况;随后,完成深度图和彩色图的像素对齐,通过坐标系转换获取桑树枝干三维坐标。试验结果表明:YOLO v5-mulberry 检测模型的平均精度均值为 94.2%,较原模型提高 16.9 个百分点,置信度也提高 12.1%;模型室外检测时应检测目标数 53,实际检测目标数为 48,检测率为 90.57%;桑树嫩叶处枝干三维坐标识别定位系统的定位误差为(9.498 5 mm, 11.285 mm,19.11 mm),满足使用要求。该研究可实现桑树嫩叶处枝干的识别与定位,有助于推动桑叶智能化采摘机器人研究。

关键词: 桑叶采摘; 枝干识别定位; YOLO v5; 目标检测; 注意力机制; 坐标转换

中图分类号: S4　　**文献标识码**: A　　**文章编号**: 1000-1298(2024)02-0249-09　　**OSID**:

Mulberry Branch Identification and Location Method Based on Improved YOLO v5 in Complex Environment

LI Li[1,2]　LU Shibo[1,2]　REN Hao[1,2]　XU Gang[3]　ZHOU Yongzhong[4]

(1. *College of Engineering and Technology*, *Southwest University*, *Chongqing* 400715, *China*
2. *Yibin Academy of Southwest University*, *Yibin* 644000, *China*
3. *Intelligent Equipment Research Center*, *Beijing Academy of Agriculture and Forestry Sciences*, *Beijing* 100097, *China*
4. *Chongqing Sunfeel Intelligent Technology Co.*, *Ltd.*, *Chongqing* 401121, *China*)

Abstract: In order to solve the recognition and detection of branches at the young leaves of mulberry trees in complex natural environments, overcome the current situation of relying on manual assisted positioning in the operation process of mulberry leaf harvesting equipment, and improve the problem of low recognition rate caused by diverse target postures and complex environments, a mulberry branch and trunk recognition model was proposed based on the improved YOLO v5 model (YOLO v5-mulberry) and combined it with the depth camera to construct a location system. Firstly, convolutional block attention module (CBAM) attention mechanism was added to the backbone network of YOLO v5 to improve the neural network's attention to the mulberry branches; and a small target layer was added to enable the model to detect 4 pixels × 4 pixels targets, which improved the model's performance in detecting small targets. At the same time, the GIoU loss function was used to replace the IoU loss function in the original network, which effectively prevented the position relationship between the prediction box and the real box from being correctly reflected when the size of the prediction box and the real box was small. Subsequently, the pixel alignment of the depth map and the color map was completed, and the 3D

收稿日期: 2023-07-16　修回日期: 2023-09-03

基金项目: 宜宾市双城协议保障科研经费科技项目(XNDX2022020015)、重庆市杰出青年科学基金项目(2022NSCQ-JQX0030)和中央高校基本科研业务费专项资金项目(SWU-XDJH202302、SWUS23099)

作者简介: 李丽(1982—),女,教授,博士生导师,主要从事智能制造和智能农机装备研究,E-mail: swulili@swu.edu.cn

coordinates of the mulberry tree trunk were obtained through the conversion of the coordinate system. The test results showed that the average accuracy of YOLO v5-mulberry detection model was 94.2%, which was 16.9 percentage points higher than that of the original model, and the confidence level was also 12.1% higher; the number of targets that should be detected by the model outdoor detection was 53, and the number of actually detected targets was 48, and the detection efficiency was 90.57%; the positioning error of the three-dimensional coordinate recognition and location system of the mulberry branch and trunk at the tender leaves was (9.498 5 mm, 11.285 mm, 19.11 mm), which met the requirements for use. The research result can achieve the recognition and positioning of branches and trunks at the tender leaves of mulberry trees, which can help to further promote the research, development and application of intelligent mulberry leaf picking robots.

Key words: mulberry leaf picking; branch identification and location; YOLO v5; target detection; attention mechanism; coordinate transformation

0 引言

2022 年中国蚕茧产值达 357.22 亿元[1],但机械化程度低,桑园劳动力老龄化等问题限制了蚕桑产业的提质增效,在此背景下国家有关部门出台了一系列政策文件鼓励因地制宜发展蚕桑特色产业,对开发蚕桑机械等方向做出了重要指示[2],智能化采桑配套技术的研发对激发蚕桑产业活力至关重要。

当下人工饲料喂养模式不成熟,多批次饲养的大环境下,桑叶的采摘需适应蚕的饲养周期,完成多轮采摘工作,便要求采摘过程中不可损伤桑树枝干,并保留顶端嫩叶,保证桑树继续光合作用发出新叶实现桑叶多批采收[3]。保留嫩叶的采桑要求成为桑叶采收设备的难点,胡迎春教授团队研发的摇杆式桑叶采摘机[4]、往复式桑叶采摘机[5]、螺旋式桑叶采摘机[6]和邸雷等[7]研发的梳刷式电动树叶采摘收集机可实现桑树单枝条采收,西南大学[8]研发的一种桑叶采摘设备实现了单株桑叶的采收,上述设备作业时均需做上下行程,且由人工控制采桑起始点或是结束点,避免伤及枝干嫩叶,这是目前桑叶采摘设备的共性。

推动蚕桑产业智能化,需为采桑设备搭建树干识别系统及嫩叶处枝干识别系统,前者用于采桑作业时车辆位置的定位,后者用于采桑作业时执行刀具的定位,本文主要开展采桑作业中嫩叶处枝干识别的研究。任浩等[9]通过 Resnet50 目标检测模型识别顶端 4 ~ 6 片桑叶以下枝干位置,准确率为 87.42%。黄家才等[10]提出的基于 Compact-YOLO v4 算法的茶叶嫩芽移动端识别方法,识别速度及检测准确率有所提高,内存占用量仅为原来的 1/5。XU 等[11]提出一种基于 HSV 模型中 S 分量的甘蔗芽特征提取与识别方法,能有效减少甘蔗芽中心位置的偏差。ZHANG 等[12]通过 RCNN 与拟合的方式对苹果树分支进行检测,其平均召回率与准确率分别达 91.5% 和 85.5%。

由上述文献可看出,除采桑设备需对作业对象开展特定位置识别外,在苹果采摘、烟叶采摘及甘蔗切割领域也有相应的应用和需求,但针对桑树嫩叶处枝干的识别和定位研究较少,且现有成果在识别率及定位精度上有较大提升空间。本文针对实际检测过程中自然环境复杂,嫩叶处枝干遮挡多、体积小等特点,提出一种基于改进 YOLO v5 的桑树嫩叶处枝条识别算法 YOLO v5-mulberry。本文算法中添加 CABM 注意力机制和小目标检测层,融合更改损失函数 3 种改进策略实现桑树嫩叶处枝干识别,并基于 YOLO v5-mulberry 识别算法搭建立体视觉系统,完成深度图与彩色图的像素对齐与坐标转换,获取嫩叶处枝干三维坐标。

1 桑树图像数据获取与预处理

1.1 桑树图像采集

图像采集于重庆祥飞智能科技有限公司合川桑园,主要种植具备无侧枝、叶形大及枝条粗直特点的农桑 14 号,桑园为适应机械化种植收获设有距离适宜的横向间距和纵向间距。拍摄高度充分考虑现有采桑设备尺寸及相机安装位置,于离地 1 m、距树干水平 0.5 m 位置在晴天不同时段和阴天分别采集图像共 1 000 幅,采集的部分图像如图 1 所示。

1.2 图像数据预处理

为增强数据的多样性,提高模型泛化能力,对采集的图像进行镜像旋转、镜像翻转、色彩增强和同态滤波处理,处理后的图像如图 2 所示。

4 种数据扩展方式产生的图像数量见表 1,原图加拓展后共 5 000 幅图像使用 Lambing 工具进行单目标标注,80% 数据集设为训练集,验证集及测试集各占 10%。标注界面如图 3 所示,标注位置在桑树 4 ~ 6 片嫩叶下方枝干处。

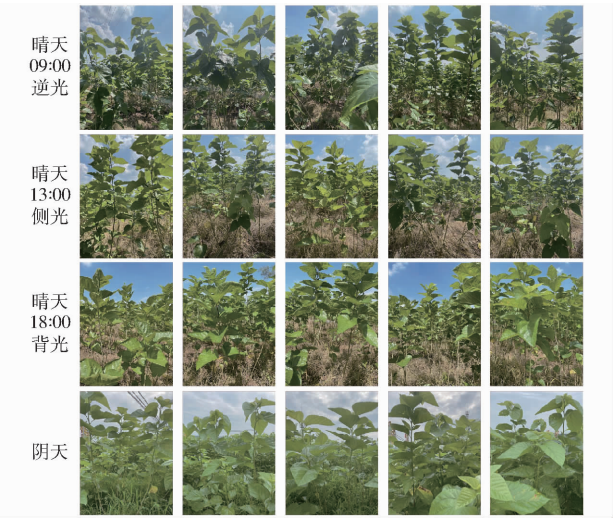


图 1　采集的图像数据集

Fig. 1　Acquired image dataset

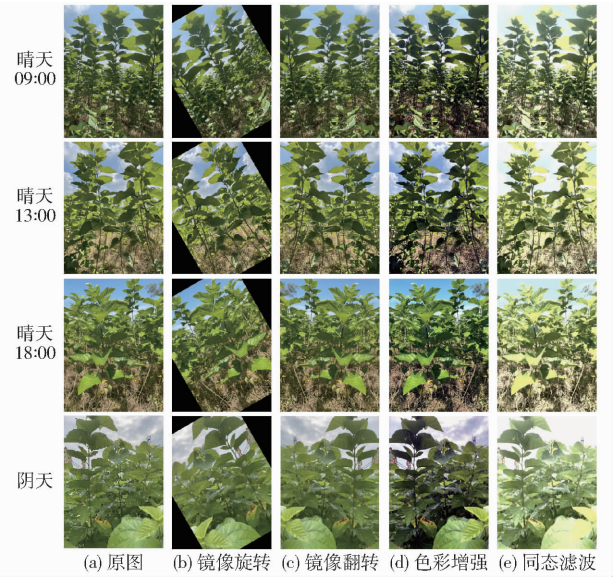


图 2　经过处理的图像数据集

Fig. 2　Processed image dataset

表 1　桑树数据增强后数据集图像数量

Tab. 1　Images generated after data expansion　幅

	原图	镜像旋转	镜像翻转	色彩增强	同态滤波
晴天	700	700	700	700	700
阴天	300	300	300	300	300

2　改进 YOLO v5 的嫩叶处枝干图像目标检测算法

2.1　YOLO v5 网络

目前比较常用的基于深度学习的目标检测网络大致可以分为 one-stage 和 two-stage 两大类[13]。YOLO 系列算法作为一种 one-stage 目标检测算法，检测精度较高，在准确率和识别上达到较好的平衡，适合用于复杂自然环境下的目标检测[14]，其网络结构如图 4 所示[15]。



图 3　标注界面

Fig. 3　Labeling interface

YOLO v5 结构大致可分为骨干网络、颈部网络和检测头网络。骨干网络使用 FOCUS 和 CSP 结构将高分辨率特征图切片拆分为多个低分辨率特征图，即隔列采样、拼接；颈部网络主要对骨干网络所提取到的特征进行区分，融合不同尺度特征[16]，并采用 CSP2_X 结构结合骨干网络中 CSP1_X 结构控制 CSP 结构宽度及深度，在减小计算量时可有效缓解梯度消失程度；检测头部分完成对图中所需特征的检测，得出图像的检测框及目标类别。

2.2　YOLO v5 网络改进

本研究使用的桑树嫩叶处枝干检测网络以图 4 中的 YOLO v5 作为基础网络，改进后整体框架如图 5 所示。在 YOLO v5 的骨干网络中加入 CBAM[17]（Convolutional block attention module）注意力机制有效削弱了无效信息，提高了神经网络对图像中桑树枝干的关注度。随后，由于识别区域较小，在与摄像机距离较大时会存在识别目标在图像中仅占不到 8 像素 ×8 像素，由于原始 YOLO v5 模型的特征图尺寸只能达到 80 像素 ×80 像素，很难检测到小于 8 像素 ×8 像素的目标[18]，便增加了小目标层，使得模型可检测 4 像素 ×4 像素的目标，因此模型在检测小目标方面表现出了优越性能。最后使用 GIoU 损失函数替换原始网络中的 IoU 损失函数，有效防止了预测框和真实框尺寸较小时无法正确反映预测框及真实框之间位置关系的情况。

2.2.1　CBAM 注意力机制增加

本研究检测目标存在特殊外形特征及一定空间位置特征，于 YOLO v5 的骨干网络中加入 CBAM 注意力机制可有效提高训练模型检测速度、网络特征提取能力和检测精度，并作为一个多维注意力机制模块加入到关注图像外形特征的通道注意力模块和关注空间位置特征的空间注意力模块中。通道注意力模块（CAM）见图 6，模块通过平均池化和最大池化对输入特征图 F 进行求和提取特征图信息，得到平均池化特征 F^c_{avg} 和最大池化特征 $F^c_{\max}$，提取特征

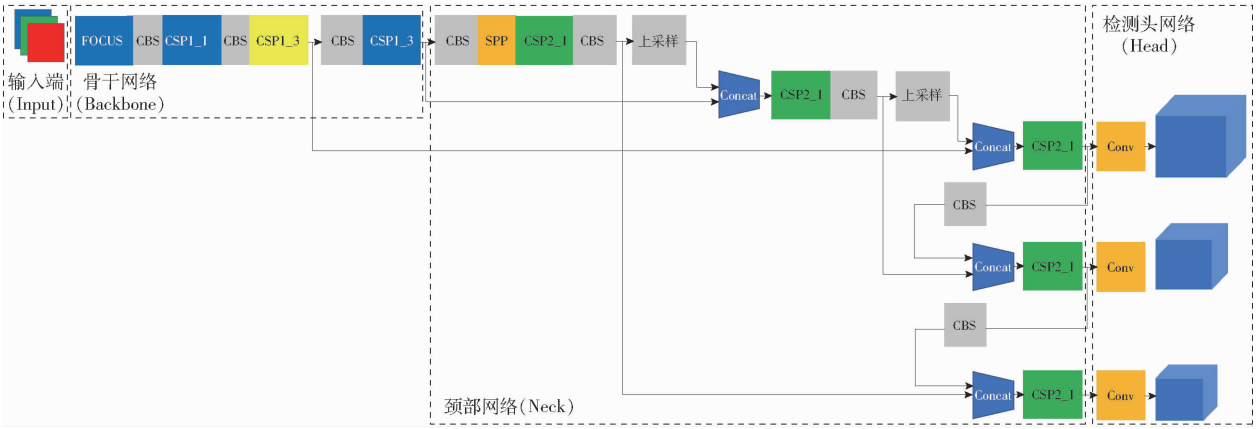


图 4 YOLO v5 网络结构图

Fig. 4 Structure of YOLO v5 network

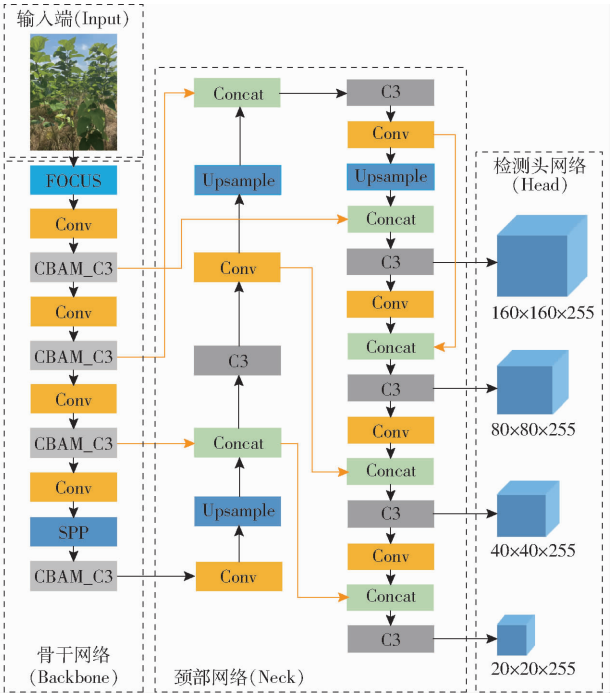


图 5 改进后的 YOLO v5 网络结构图

Fig. 5 Structure diagram of improved YOLO v5 network

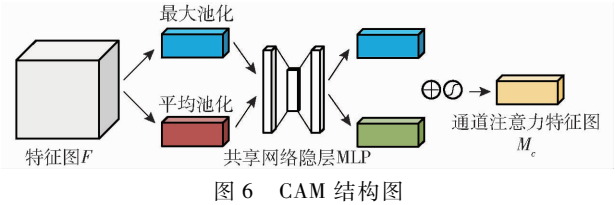


图 6 CAM 结构图

Fig. 6 Structure graph of CAM

图空间信息,随后 2 个特征被传递到共享网络隐层 MLP 进行处理,通过激活函数 σ 得到通道注意力特征图 M_c[19]。多层感知模型中的两层参数用 α_1、α_2 表示,可获得注意力表达式[20]为

$$M_c(F)=\sigma(\alpha_2\alpha_1F^c_{\text{avg}})+\alpha_2\alpha_1F^c_{\max} \tag{1}$$

如图 7 所示,空间注意力模块(SAM)的输入是通道注意力模块上一步的输出特征图。该模块对输入特征映射进行平均池化和最大池化后,得到 F^s_{avg}

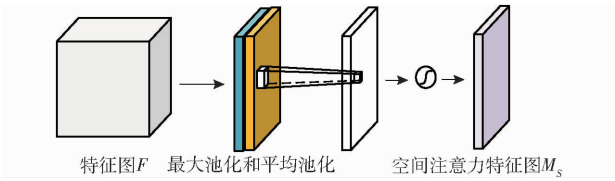


图 7 SAM 结构图

Fig. 7 Structure graph of SAM

和 $F^s_{\max}$ 的聚合通道信息[21],2 个特征图被缝合成一个特征图后用 7×7 的卷积生成一个二维空间特征图,空间注意力表达式为

$$M_s(F)=\sigma(f^{7\times7}([F^s_{\text{avg}};F^s_{\max}])) \tag{2}$$

式中 $M_s(F)$——特征图 F 经过空间注意力的输出权值

$f^{7\times7}$——滤波器大小为 7×7 的卷积运算

$[F^s_{\text{avg}};F^s_{\max}]$——按通道拼接的 2 个 $H\times W\times1$ 描述

CBAM 注意力机制整体流程如图 8 所示,通道注意力机制与空间注意力机制协同学习图像中的关键局部信息,为神经网络特征图中桑树顶端 4~6 片桑叶区域枝干分配较高的权重,提高了神经网络对图像中桑树嫩叶处枝干的关注度,进而提升了网络特征学习和表达能力。

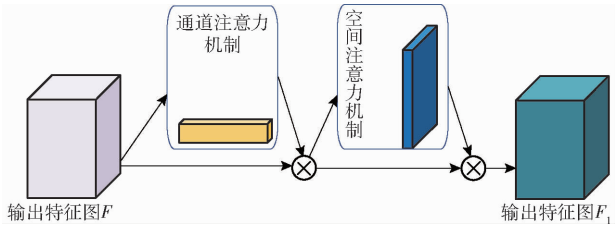


图 8 CBAM 结构图

Fig. 8 Structure graph of CBAM

2.2.2 小目标层增加

原始 YOLO v5 网络包括骨干网络和特征融合,输出 20×20、40×40、80×80 共 3 种尺度的特征图用于检测不同尺寸目标[22]。随桑树枝和摄像机之

间距离的增加，嫩叶处桑树枝在整个图像中仅占不到 8 个像素，原始 YOLO v5 模型很难检测到小于 8 像素 ×8 像素的目标。在原网络中加入一个小目标检测层可使改进后的模型采样获得最大 160 像素 × 160 像素的特征图，从而可以检测到 4 像素 ×4 像素大小的物体。添加小目标层后，采用 COCO 评价方法验证其性能，将 YOLO v5 模型与直接添加小目标层的模型 YOLO v5-SL 进行对比试验，试验结果见表 2。表中，AP_{50} 表示 IoU 阈值为 0.5 时的检测精度；AP_S 表示对小目标（面积小于 32 像素 ×32 像素）的检测精度；AP_M 表示对中目标（面积在 32 像素 × 32 像素 ~96 像素 ×96 像素之间）的检测精度；AP_L 表示对大目标（面积大于 96 像素 ×96 像素）的检测精度。

表 2　小目标层试验结果对比

Tab. 2　Comparison of test results of small target layer

%

模型	AP_{50}	AP_S	AP_M	AP_L
YOLO v5	77.5	19.6	55.4	64.2
YOLO v5-SL	84.3	27.8	56.2	63.5

由表 2 可知，添加了小目标层的改进模型相比于原始模型，AP_S 提高 8.2 个百分点，AP_L 和 AP_M 基本保持不变，这表明添加小目标层的模型在不影响大中目标检测的前提下，能有效提高对小对象的检测性能。

2.2.3　损失函数更改

YOLO v5 目标检测算法使用 IoU 作为边界损失函数判断样本正负，基于此计算置信度损失。IoU 通过计算预测框和真实框的交集部分除以预测框与真实框的并集部分得出结果。当预测框与真实框没有发生交集时，IoU 计算结果为 0，无法正确反映预测框与真实框之间的位置关系，这种情况在预测框和真实框尺寸都较小的情况下极易发生，此时使用 IoU 作为损失函数，训练网络将无法实现优化，针对此情况选择将 IoU 损失函数更换为 GIoU 损失函数[23]。

GIoU[24] 损失函数是 IoU 损失函数的改进形式，它继承了 IoU 尺度不变性的优点，同时 GIoU 损失函数还考虑了预测框与真实框不发生交集的情况。将 GIoU 作为 YOLO v5 目标检测算法度量时，其损失函数为

$$L_{GIOU} = 1 - R_{GIOU} \tag{3}$$

2.3　试验环境

本研究采用的图像采集设备为 Intel RealSense D435i 型相机，通过 USB3.0 接口与计算机连接。相机有效识别范围为 0.105 ~10 m，分辨率为 1 280 像素 ×800 像素，帧率为 90 f/s，像元尺寸为 3 μm × 3 μm。深度学习试验硬件环境为搭载 Intel（R）Xeon（R）Gold 5118 CPU@2.30 GHz 2.29 GHz + NVIDIA Quadro P4000 显卡的工业计算机，软件选择 Windows 10 操作系统、YOLO v5 网络模型、Python 编程语言。

2.4　模型性能评估指标

本文使用精确率（Precision，P）、召回率（Recall，R）、平均精度均值（mAP）和 F1 值 4 个常用指标作为验证模型性能的指标。设定两个边界框之间重叠区域和不重叠区域的交并比（Generalized intersection over union，GIoU）阈值为 0.5，若 GIoU 超过该阈值为正样本，反之为负样本。

2.5　模型性能

2.5.1　消融试验结果

为进一步验证改进方法的有效性，本文进行了一系列消融试验。使用相同的数据集、测试平台及模型参数开展比较分析，评估每种改进方式对最终算法的优化效果。本研究的 3 种改进方法分别为 GIoU、SL 和 CBAM，试验方法如下：①基于原始 YOLO v5 模型，分别加入 3 种改进方法，以评估每种改进方法对最终模型效果的影响。②加入原始模型及完整的 YOLO v5-mulberry 模型进行对比。从表 3 可以看出，对比改进前的 YOLO v5 模型，3 种改进方法均有效提升了网络检测性能，CBAM 注意力机制对改进模型的效果最为显著，精确率和平均精度均值分别提升 6.9、11 个百分点。在经过 3 种方法改进后的 YOLO v5-mulberry 模型中，测试集的 mAP 从 77.3% 提高到 94.2%，精确率、F1 值和召回率分别提高 10、17、22.2 个百分点，损失值也降低到 0.031，说明改进后的模型具有良好的收敛性。

表 3　消融试验结果对比

Tab. 3　Comparison of ablation experiments

序号	GIoU	SL	CBAM	精确率/%	F1 值/%	召回率/%	平均精度均值/%	损失值
1	×	×	×	86.9	77	69.6	77.3	0.059
2	√	×	×	89.6	84	79.9	84.5	0.045
3	×	√	×	90.6	86	81.6	85.9	0.041
4	×	×	√	93.8	88	83.4	88.3	0.037
5	√	×	√	95.1	91	88.1	91.4	0.035
6	√	√	√	96.9	94	91.8	94.2	0.031

注：“√”表示当前网络使用该结构或方法；“×”表示当前网络未使用该结构或方法。

2.5.2　不同网络试验结果对比

为进一步评价 YOLO v5-mulberry 网络对嫩叶处枝干的检测效果，使用 YOLO v4、YOLO v7、Faster

R－CNN 3 种主流模型在同一环境下对桑树数据集进行训练和测试，结果如表 4 所示。

表 4 与主流模型的对比试验结果

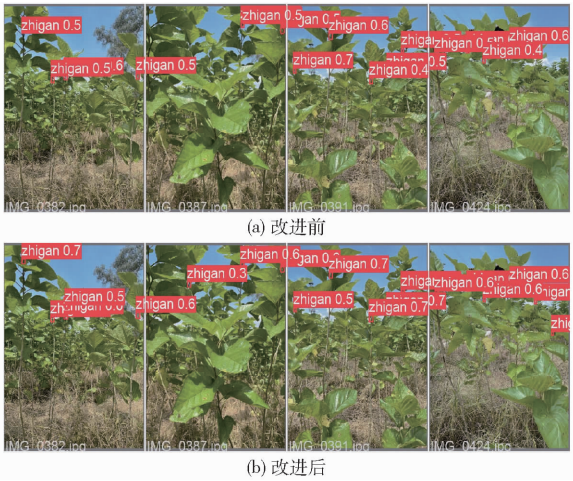
Tab. 4 Comparison test results with mainstream models

模型	精确率/%	召回率/%	平均精度均值/%	检测速度/(f·s^{-1})
Faster R－CNN	80.7	65.1	71.5	5.8
YOLO v4	81.5	71.8	74.6	22.2
YOLO v5	86.9	69.6	77.3	21.9
YOLO v7	90.8	80.4	84.5	28.8
YOLO v5-mulberry	96.9	91.8	94.2	18.6

由表 4 可以看出，本文模型与其他主流检测模型相比，精确率和平均精度均值有明显优势，较 YOLO v7 模型分别提高 6.1、9.7 个百分点。由于本文模型加入了注意力机制及小目标层，综合影响下检测速度由 21.9 f/s 降低至 18.6 f/s，但在采桑作业中，嫩叶处枝干识别定位过程与执行器插入枝条这一行程同时进行，检测时间充分，帧率的略微下降不影响实际应用，因此本文模型在帧率满足现实要求的前提下，保证了较高水平的平均精度均值和召回率，综合性能最优。

2.5.3 改进网络试验结果

图 9 展示了使用同样的验证集图像情况下，YOLO v5 网络改进前后检测结果的对比，在这两组检测结果中可以发现，改进前共检测到 28 处，置信度平均值为 48.3%，改进后检测到 33 处，置信度平均值为 60.4%，改进后的网络模型识别到的桑树嫩叶处枝干数量更多，且置信度也有显著提高。



(a) 改进前

(b) 改进后

图 9 YOLO v5 网络改进前后检测结果对比

Fig. 9 Comparison of detection results of YOLO v5 network before and after improvement

为进一步测试 YOLO v5-mulberry 的性能，在室内及室外分别开展桑树枝检测效果试验，其中室外测试应检测到枝干特定位置 53 处，室内测试应检测到枝干特定位置 58 处，测试结果见表 5。可以看出在室内光照条件较好、背景相对简单的情况下，桑树枝干的检测率更高，在桑园中进行检测时，当光照条件较差、枝干相互交错时算法会对一些枝干漏检，但目前算法的检测率在桑园中仍可达到 90% 以上。

表 5 室外桑园和室内环境下桑树枝干检测数量

Tab. 5 Summary of mulberry stem detection quantity in outdoor mulberry garden and indoor environment

环境	实际目标数	算法检测到目标数	检测率/%
室外桑园	53	48	90.57
室内环境	58	56	96.55
总计	111	104	93.69

注：检测率为算法检测到目标数占实际目标数的百分比。

3 基于双目立体视觉的桑树枝干空间坐标获取

本文以自主研发的桑叶采摘设备容许定位误差作为衡量获取的三维坐标是否可行的标准，该设备采集数据原理为：首先电机控制定位夹爪对从上往下第 4～6 片桑叶处的桑树枝进行夹紧，如图 10a 所示；随后定位爪与采摘爪同步合拢，采摘爪开始向下运动，桑叶被扎于采叶针上，如图 10b 所示；采叶完成后，定位爪与采摘爪松开并复位到原点，丝杆电机正转带动采叶针固定板向上运动，采叶针随之穿过脱叶板，继而使针上叶片随重力落入下方的桑叶打包收集装置，如图 10c 所示。

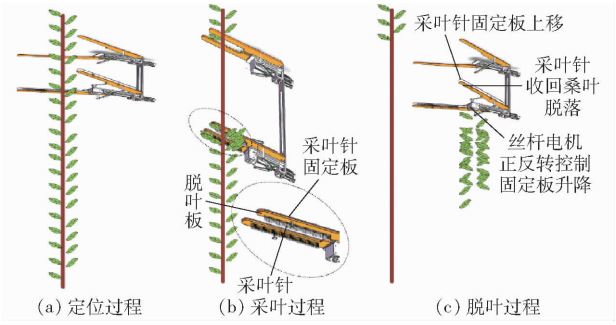


图 10 定位、采叶、脱叶原理图

Fig. 10 Schematic of principle of positioning, leaf picking and leaf shedding

定位爪与采摘爪张开宽度为 150 mm，小于顶端处任意两枝条的横向距离，故 X 轴定位误差不大于 75 mm；测量 50 株桑树第 4 片嫩叶处枝干与顶端距离，最小值为 84.3 mm，大于定位爪与采摘爪之间的距离 40 mm，故 Y 轴定位误差取定位爪与采摘爪的间距，不大于 40 mm；测量 50 株桑树的顶端处任意两枝条的纵向距离最小值为 45 mm，故 Z 轴定位误差不大于 45 mm。综上，最大容许误差为：X 轴 75 mm，Y 轴 40 mm，Z 轴 45 mm。

3.1　相机标定及空间坐标系转换

选用 Intel 公司的 RealSense D435i 系列双目相机开展研究，因相机镜头存在畸变现象，为提高识别检测算法的准确度，需完成标定获取相机内参参数和畸变参数[25]，对相机完成标定及空间坐标系的转换，并将深度图对齐到彩色图，进而建立图像世界与三维世界的准确关系。

3.1.1　相机标定

采用 Dynamic Calibrator 工具箱及具有固定尺寸的动态校准标定板进行彩色相机及深度相机标定，获取相机内参参数和畸变参数，从而在坐标系转换的矩阵运算中进行精度矫正，相机标定结果见表 6。

表 6　双目相机内部参数及畸变参数

Tab. 6　Internal parameters and distortion parameters of binocular camera

参数		相机标定结果	
		深度相机	彩色相机
焦距	f_x/像素	616.129	908.459
	f_y/像素	616.330	909.059
主点	c_x/像素	602.937	616.345
	c_y/像素	319.643	331.544
旋转矩阵		[0.999 0.006 0.001 0.006 0.999 0.999 0.001 0.001 0.001]	
平移矩阵		[0.015 0.000 2 0.000 7]	
畸变参数 (k_1,k_2,k_3,p_1,p_2)		(0.092 62,0.180 82,0,0.000 65,0.000 36)	

实际成像点 $p'(x',y')$ 和理想成像点 $p(x,y)$ 之间的转换关系表达式为

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = (1 + k_1 r^2 + k_2 r^2 + k_3 r^6) \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 2p_1 xy + p_2(r^2 + 2x^2) \\ 2p_2 xy + p_1(r^2 + 2y^2) \end{bmatrix} \tag{4}$$

式中　k_1、k_2、k_3——径向畸变参数

r——曲率半径

p_1、p_2——切向畸变参数

3.1.2　空间坐标系转换

得到双目相机的内参及畸变参数后，为获得嫩叶枝干的三维空间信息，通过相机的针孔成像原理和矩阵运算的方法进行了像素坐标系、图像坐标系、相机坐标系和世界坐标系的转换[26]：空间坐标系转换示意图如图 11 所示，$Q(X_W,Y_W,Z_W)$ 为在世界坐标系 $O_WX_WY_WZ_W$ 下的嫩叶处枝干坐标点；$O_CX_CY_CZ_C$ 为双目相机坐标系；Oxy 为图像坐标系；uv 为像素坐标系，原点位于图像的左上角（单位为像素）；$q(x,y)$ 为桑树枝干在图像坐标系中的投影点，相机焦距 f 等于 O 与 O_C 间的距离。

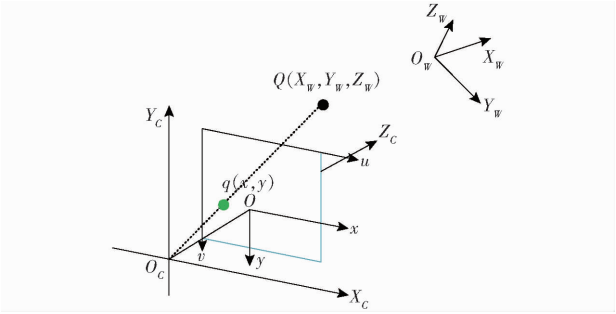


图 11　空间坐标系转换示意图

Fig. 11　Spatial coordinates transform

3.2　深度图像处理

双目相机同时获取彩色图像和深度图像，后者用色彩差异表示红外摄像头的拍摄距离[27]。图 12a、12b 分别是同步采集的深度图和彩色图，由于 2 个摄像头的位置偏差，其对应像素点有位置差异，不能直接使用。为了在目标检测时使用彩色图像数据，通过坐标转换将深度图对齐到彩色图，转换公式为

$$\boldsymbol{T}_{d2c} = \begin{bmatrix} \boldsymbol{R}_{w2c}\boldsymbol{R}_{w2d}^{-1} & \boldsymbol{t}_{w2c} - \boldsymbol{R}_{w2c}\boldsymbol{R}_{w2d}^{-1}\boldsymbol{t}_{w2d} \\ 0 & 1 \end{bmatrix} \tag{5}$$

式中　$\boldsymbol{R}_{w2c}$——世界坐标系转换到彩色坐标系的旋转矩阵

$\boldsymbol{t}_{w2c}$——世界坐标系转换到彩色坐标系的偏移矩阵

$\boldsymbol{R}_{w2d}$——世界坐标系向红外相机的深度坐标系转换的旋转矩阵

$\boldsymbol{t}_{w2d}$——世界坐标系转换到深度相机坐标系的偏移矩阵

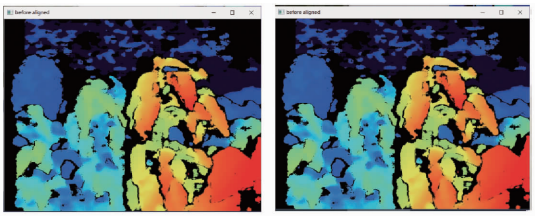
深度图对齐后如图 12c 所示。

3.3　基于双目相机的嫩叶枝干坐标检测

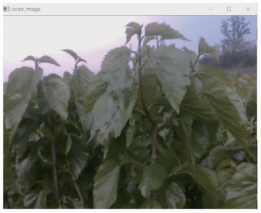
选用 Visual Studio 2019 作为开发环境，主要实现以下功能：①读取双目相机实时采集的彩色图与深度图。②调用 YOLO v5-mulberry 模型对彩色图中桑树枝干位置进行识别。③将深度图对齐到彩色图。④将像素坐标转换为世界坐标系下的真实坐标。

为验证使用 RealSense D435 型相机对嫩叶处枝干三维坐标检测精度是否满足采摘需求，本文使用双目相机及其固定架、吊线锤、便携式计算机、激光测距仪等实验仪器在田间进行三维坐标精度测量试验。

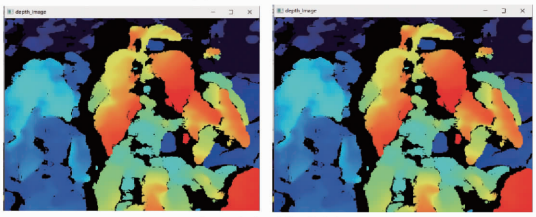
RealSense D435 型相机固定安装，摄像头方向水平向前，使用吊线锤标记相机镜头于地面的投影位置并测量镜头与地面的竖直距离，使相机镜头在多棵桑树上识别有共 20 处嫩叶部位枝干；测量嫩叶处枝干投影位置与相机镜头投影位置的垂直距离与水平距离，计算得桑树枝干三维坐标的实际测量位置，此值与算法识别的三维坐标进行对比并计算误差，算法识别界面见图 13，检测结果见表 7。



(a) 对齐前深度图



(b) 同一时刻彩色图



(c) 对齐后深度图

图 12 桑树同一时刻彩色图与深度图

Fig. 12 Color map and depth map of mulberry at the same time

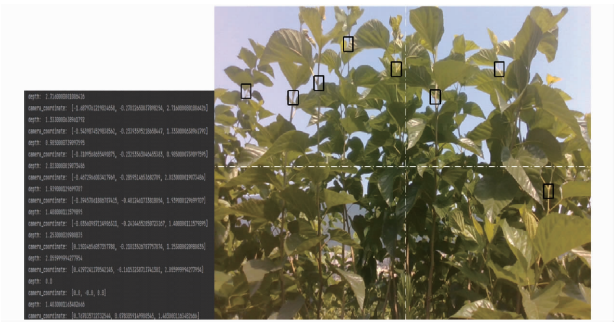


图 13 三维坐标算法识别界面

Fig. 13 3D coordinate algorithm recognition interface

由表 7 可知,X 轴最大误差为 18 mm,平均误差为 9.498 5 mm;Y 轴最大误差为 34.7 mm,平均误差为 11.285 mm;Z 轴最大误差为 42 mm,平均误差为 19.11 mm,误差在允许范围内,符合使用要求。

4 结论

(1)因为检测目标的特殊性,通过增加 CBAM 注意力机制、增加小目标层、损失函数替换为 GIoU 来提高原模型的性能。改进 YOLO v5-Mulberry 模型精确率和平均精度均值达到 96.9% 和 94.2%,较 YOLO v5 模型提高 10、16.9 个百分点,与 Faster R-CNN、YOLO v4 和 YOLO v7 模型相比有较大优势。并于室外和室内开展模型检测,识别率分别为 90.57% 和 96.55%,验证了本文改进模型的可行性。

(2)对相机进行标定,完成坐标系转换与深度图对齐后,于桑园开展了嫩叶处枝干三维坐标检测,受模型识别精度、测量精度和相机精度等因素干扰,得到桑树嫩叶处枝干三维坐标识别定位系统的定位误差为(9.498 5 mm,11.285 mm,19.11 mm),通过改进 YOLO v5 模型的识别定位系统提高了桑树嫩叶处枝干目标的识别精度,定位精度满足需求,为采桑设备智能化研究提供了参考。

表 7 桑树枝干三维坐标检测结果

Tab. 7 Results of mulberry stem three-dimensional coordinate detection cm

序号	桑树枝干坐标		
	真实值	检测值	误差
1	(4.1,-21.4,138.7)	(4.38,-28.87,142.20)	(0.28,7.47,3.50)
2	(-66.3,-1.7,44.3)	(-68.09,-1.05,41.10)	(1.79,0.65,3.20)
3	(18.0,21.1,92.1)	(16.81,20.08,93.80)	(1.19,1.02,1.70)
4	(23.4,56.9,59.2)	(23.14,57.43,59.23)	(0.26,0.53,0.03)
5	(17.6,-14.4,64.9)	(18.7,-14.86,64.32)	(1.10,0.46,0.58)
6	(29.4,32.6,66.9)	(29.86,33.94,66.73)	(0.46,1.34,0.27)
7	(-3.0,10.5,43.7)	(-3.35,9.44,42.90)	(0.35,1.06,0.80)
8	(45.3,-9.8,70.3)	(45.78,-9.30,72.34)	(0.48,0.50,2.04)
9	(-56.2,-21.9,135.9)	(-54.40,-23.39,133.30)	(1.80,1.49,2.60)
10	(-32.1,-22.1,99.5)	(-31.09,-23.25,98.50)	(1.01,1.15,1.00)
11	(-46.9,-29.4,204.3)	(-46.72,-28.60,202.30)	(0.18,0.80,2.00)
12	(-38.6,-49.5,191.4)	(-39.46,-48.12,193.90)	(0.86,1.48,2.50)
13	(-2.7,-26.3,145.4)	(-3.56,-24.24,140.80)	(0.86,2.06,4.20)
14	(16.5,-22.8,127.2)	(15.02,-21.02,125.30)	(1.48,1.78,1.90)
15	(44.3,-16.4,203.1)	(42.97,-16.15,205.60)	(1.33,0.25,2.50)
16	(78.4,8.4,142.9)	(76.70,7.83,140.30)	(1.70,0.57,2.60)
17	(-12.3,-11.1,61.3)	(-13.87,-11.69,59.50)	(1.57,0.59,1.80)
18	(16.5,-8.5,82.9)	(17.23,-9.83,81.60)	(0.73,1.33,1.30)
19	(35.1,-4.7,73.9)	(33.83,-6.11,72.70)	(1.27,1.41,1.20)
20	(-34.2,-19.5,71.5)	(-33.93,-18.87,69.00)	(0.27,0.63,2.50)

参 考 文 献

[1] 李建琴, 梁巧, 顾国达. 新冠肺炎疫情对我国蚕丝业的影响及发展对策[J]. 中国畜牧杂志, 2023, 59(3): 336-341.

[2] 六部门印发《蚕桑丝绸产业高质量发展行动计划(2021—2025 年)》[J]. 北方蚕业, 2020, 41(4): 60.

[3] 赵永财, 张晋, 李法德, 等. 桂桑优 12 和农桑 14 号草本式栽培桑枝条力学特性研究[J]. 蚕业科学, 2022, 48(6): 489-495.
ZHAO Yongcai, ZHANG Jin, LI Fade, et al. Study on the mechanical characteristics of mulberry branches of Guisangyou 12 and Nongsang 14 in herbaceous cultivation[J]. Acta Sericologica Sinica, 2022, 48(6): 489-495. (in Chinese)

[4] 胡迎春, 齐勇落, 胡裔志, 等. 摇杆式桑叶采摘机采摘装置的有限元分析[J]. 江苏农业科学, 2019, 47(22): 255-257, 268.
HU Yingchun, QI Yongluo, HU Yizhi, et al. Finite element analysis of picking device of cocker type mulberry picking machine [J]. Jiangsu Agricultural Sciences, 2019, 47(22): 255-257, 268. (in Chinese)

[5] 胡迎春，闫鑫，庄锦芳，等. 往复式桑叶采摘机设计及采摘效益分析[J]. 农机化研究，2016，38(4)：76－79.
HU Yingchun，YAN Xin，ZHUANG Jinfang，et al. Design of reciprocating machine of picking mulberry leaves and analysis of its efficiency and benefit[J]. Journal of Agricultural Mechanization Research，2016，38(4)：76－79.（in Chinese）
[6] 胡迎春，华南，牟向伟，等. 螺旋式桑叶采摘机 PLC 控制系统的设计与仿真[J]. 机械研究与应用，2016，29(3)：56－59.
HU Yingchun，HUA Nan，MOU Xiangwei，et al. Design and simulation of PLC control system of spiral mulberry leaves plucking machine[J]. Mechanical Research & Application，2016，29(3)：56－59.（in Chinese）
[7] 邸雷，杨自栋. 梳刷式喜树叶采摘收集机设计[J]. 农业装备与车辆工程，2020，58(8)：28－31.
DI Lei，YANG Zidong. Design of brush-type *Camptotheca acuminata* leaf picking and collecting machine[J]. Agricultural Equipment & Vehicle Engineering，2020，58(8)：28－31.（in Chinese）
[8] 西南大学. 一种桑叶采摘设备:CN116210455A[P]. 2023－06－06.
[9] 任浩，李丽，卢世博，等. 基于深度学习的复杂自然环境下桑树枝干识别方法[J]. 中国农机化学报，2023，44(2)：182－188.
REN Hao，LI Li，LU Shibo，et al. Identification method of mulberry tree branches in complex natural environments based on deep learning[J]. Journal of Chinese Agricultural Mechanization，2023，44(2)：182－188.（in Chinese）
[10] 黄家才，唐安，陈光明，等. 基于 Compact-YOLO v4 的茶叶嫩芽移动端识别方法[J]. 农业机械学报，2023，54(3)：282－290.
HUANG Jiacai，TANG An，CHEN Guangming，et al. Mobile recognition solution of tea buds based on Compact-YOLO v4 algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery，2023，54(3)：282－290.（in Chinese）
[11] XU H Z，LI Q，LIU J D，et al. Feature extraction and identification of sugarcane bud based on S component in HSV model [J]. Recent Adv. Electr. El.，2023，16(1)：78－89.
[12] ZHANG J，HE L，KARKEE M，et al. Branch detection for apple trees trained in fruiting wall architecture using depth features and regions-convolutional neural network (R－CNN)[J]. Computers and Electronics in Agriculture，2018，155：386－393.
[13] ZHANG X H，WANG H P，XU C G，et al. A lightweight feature optimizing network for ship detection in SAR image[J]. IEEE Access，2019，7：141662－141678.
[14] 赵德安，吴任迪，刘晓洋，等. 基于 YOLO 深度卷积神经网络的复杂背景下机器人采摘苹果定位[J]. 农业工程学报，2019，35(3)：164－173.
ZHAO Dean，WU Rendi，LIU Xiaoyang，et al. Apple positioning based on YOLO deep convolutional neural network for picking robot in complex background[J]. Transactions of the CSAE，2019，35(3)：164－173.（in Chinese）
[15] ZHOU J C，JIANG P，ZOU A R，et al. Ship target detection algorithm based on improved YOLO v5[J]. J. Mar. Sci. Eng.，2021，9：908.
[16] ZHU L L，GENG X，LI Z，et al. Improving YOLO v5 with attention mechanism for detecting boulders from planetary images [J]. Remote Sens.，2021，13(18)：3776.
[17] WOO S，PARK J，LEE J Y，et al. CBAM：convolutional block attention module[C] // Proceedings of the European Conference on Computer Vision(ECCV)，2018：3－19.
[18] 韩俊，袁小平，王准，等. 基于 YOLO v5s 的无人机密集小目标检测算法[J]. 浙江大学学报(工学版)，2023，57(6)：1224－1233.
HAN Jun，YUAN Xiaoping，WANG Zhun，et al. UAV dense small target detection algorithm based on YOLO v5s[J]. Journal of Zhejiang University(Engineering Science)，2023，57(6)：1224－1233.（in Chinese）
[19] WANG D C，CHEN X N，JIANG M Y，et al. ADS-Net:an attention-based deeply supervised network for remote sensing image change detection[J]. Int. J. Appl. Earth Obs.，2021，101：102348.
[20] DAI J，ZHAO X，LI L P，et al. GCD－YOLOv5：an armored target recognition algorithm in complex environments based on array LiDAR[J]. IEEE Photonics J.，2022，14(4)：1－11.
[21] 姚艳清，程塨，谢星星，等. 多分辨率特征融合的光学遥感图像目标检测[J]. 遥感学报，2021，25(5)：1124－1137.
YAO Yanqing，CHENG Gong，XIE Xingxing，et al. Optical remote sensing image object detection based on multi-resolution feature fusion[J]. National Remote Sensing Bulletin，2021，25(5)：1124－1137.（in Chinese）
[22] ZHANG B，SUN C F，FANG S Q，et al. Workshop safety helmet wearing detection model based on SCM－YOLO[J]. Sensors，2022，22(17)：6702.
[23] 储震，张小玲，殷高方，等. 基于改进 YOLO v3 的浮游藻类检测算法[J]. 激光与光电子学进展，2023，60(2)：257－264.
CHU Zhen，ZHANG Xiaoling，YIN Gaofang，et al. Detection algorithm of planktonic algae based on improved YOYO v3[J]. Laser & Optoelectronics Progress，2023，60(2)：257－264.（in Chinese）
[24] CHEN T K，ZHOU Y，WANG S F，et al. GIoU-CLOCs：import generalized intersection over union-involved object-detection tasks based on LiDAR and camera[J]. J. Russ. Laser Res.，2023，44(1)：100－109.
[25] HAN F F，BIAN Y X，LIU B，et al. Research on calibration of a binocular stereo-vision imaging system based on the artificial neural network[J]. J. Opt. Soc. Am. A.，2023，40(2)：337－354.
[26] 黄玲涛，王彬，倪涛，等. 基于 Kinect 的机器人抓取系统研究[J]. 农业机械学报，2019，50(1)：390－399.
HUANG Lingtao，WANG Bin，NI Tao，et al. Research on robotic grasping system based on Kinect camera[J]. Transactions of the Chinese Society for Agricultural Machinery，2019，50(1)：390－399.（in Chinese）
[27] 徐胜勇，卢昆，潘礼礼，等. 基于 RGB－D 相机的油菜分枝三维重构与角果识别定位[J]. 农业机械学报，2019，50(2)：21－27.
XU Shengyong，LU Kun，PAN Lili，et al. 3D reconstruction of rape branch and pod recognition based on RGB－D camera [J]. Transactions of the Chinese Society for Agricultural Machinery，2019，50(2)：21－27.（in Chinese）