

DOI: 10.11992/tis.202212026

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20230802.0945.003>

# 基于语义分割视觉伺服的种苗自动夹取系统设计

翟永杰<sup>1</sup>, 王家豪<sup>1</sup>, 张鑫<sup>1</sup>, 胡东阳<sup>1</sup>, 王乾铭<sup>1</sup>, 徐大伟<sup>1,2</sup>, 刘亚军<sup>3</sup>

(1. 华北电力大学自动化系, 河北保定 071003; 2. 中国科学院自动化研究所复杂系统管理与控制国家重点实验室, 北京 100190; 3. 湖北壹鸣生物科技有限公司, 湖北钟祥 431900)

**摘要:** 现代植物组织培养是一项耗时费力的工作, 工作强度大, 工作内容单调, 为了减少劳动力成本并提高产量, 设计了一种基于语义分割视觉伺服的种苗自动夹取系统并进行了测试。首先, 提出了一种基于 DP-BiSeNetV2 语义分割算法的视觉定位方法, 确定了根上合适的夹取点; 之后, 设计开发并测试了一种适合于实际工作环境的夹取装置; 最后, 将视觉定位算法与机器人夹取装置集成, 构建了一个种苗自动夹取系统。在实验环节, 使用蝴蝶兰种苗数据集进行了测试, 在语义分割实验中, DP-BiSeNetV2 模型的平均交并比为 63.51%, 像素准确度为 98.25%; 在夹取实验中, 夹取成功率为 81.7%。实验结果表明, 该自动夹取系统具有很高的潜力, 可以满足植物组织培养生产线的移植需求。

**关键词:** 语义分割; 视觉伺服; 种苗; 夹取点定位; 夹取系统; 机械手; 智能抓取; 深度学习  
**中图分类号:** TP241.2; S126 **文献标志码:** A **文章编号:** 1673-4785(2023)06-1259-09

中文引用格式: 翟永杰, 王家豪, 张鑫, 等. 基于语义分割视觉伺服的种苗自动夹取系统设计 [J]. 智能系统学报, 2023, 18(6): 1259-1267.

英文引用格式: ZHAI Yongjie, WANG Jiahao, ZHANG Xin, et al. Design of automatic picking system for seedlings based on semantic segmentation visual servo[J]. CAAI transactions on intelligent systems, 2023, 18(6): 1259-1267.

## Design of automatic picking system for seedlings based on semantic segmentation visual servo

ZHAI Yongjie<sup>1</sup>, WANG Jiahao<sup>1</sup>, ZHANG Xin<sup>1</sup>, HU Dongyang<sup>1</sup>, WANG Qianming<sup>1</sup>,  
XU Dawei<sup>1,2</sup>, LIU Yajun<sup>3</sup>

(1. Department of Automation, North China Electric Power University, Baoding 071003, China; 2. The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China; 3. Hubei Artisan Biotech Co., Ltd, Zhongxiang 431900, China)

**Abstract:** Modern plant tissue culture is a time-consuming and labor-intensive task with monotonous work. An automatic seedling clamping system based on semantic segmentation visual servo was designed and tested to reduce labor costs and increase production. First, a vision localization method was proposed based on the DP-BiSeNetV2 semantic segmentation algorithm to determine the appropriate clamping point on the root. Further, a clamping device suitable for the actual working environment was designed, developed, and tested. Finally, an automatic seedling clamping system was constructed by integrating the vision localization algorithm with the robot clamping device. In the experimental session, tests were conducted using the Phalaenopsis seedling dataset. In the semantic segmentation experiment, the mIoU and pixel accuracy of the DP-BiSeNetV2 model were 63.51% and 98.25%, respectively. Furthermore, the success rate was 81.7% in the clamping experiment. Experimental results show that the automatic clamping system has a large potential to meet the transplantation requirements of plant tissue culture production lines.

**Keywords:** semantic segmentation; visual servo; seedling; grasping point localization; grasping system; robot arm; intelligent robotic grasping; deep learning

收稿日期: 2022-12-26. 网络出版日期: 2023-08-02.

基金项目: 国家自然科学基金联合基金重点支持项目(U21A20486);  
中国科学院自动化研究所复杂系统管理与控制国家重点实验室开放课题(20220102).

通信作者: 徐大伟. E-mail: [xudawei@ncepu.edu.cn](mailto:xudawei@ncepu.edu.cn).

植物组织培养技术是一种先进的营养繁殖技术, 可在短时间内生产出高质量和无病害的植物<sup>[1]</sup>。然而, 这一过程非常烦琐且劳动强度大<sup>[2]</sup>。

为了减少劳动力成本,提高产量,并通过减少人的接触来降低污染风险,需要开发出自动化种苗夹取系统<sup>[3]</sup>。虽然农业机器人研究起步较早,但是当前还没有成熟的应用在实际现场无菌环境下的自动化种苗夹取系统。

在过去几十年中,世界各地的研究人员对机器人采摘水果蔬菜等进行了大量研究<sup>[4]</sup>, Ringdahl 等<sup>[5]</sup>2019 年研究并设计了一种单臂甜椒采摘机器人,室内环境能够稳定采摘甜椒; Van 等<sup>[6]</sup>研究并设计了一种黄瓜采摘机器人,该机器人搭载一个 7 自由度的机械臂; Wang 等<sup>[7]</sup>设计的番茄采摘机器人由小车搭载机械臂在园区完成了采摘作业; Mu 等<sup>[8]</sup>设计了一种猕猴桃采摘机器人,独特的夹具设计能在多个猕猴桃中准确夹取; Sarabu 等<sup>[9]</sup>设计了一种基于双臂操作的苹果采摘机器人,这些机器人全部使用了现代协作机械臂,根据采摘对象不同设计不同的末端夹具完成采摘作业。在目标的识别和定位方面,李天华等<sup>[10]</sup>提出了一种基于 YOLO v4 与 HSV(hue, saturation, value) 相结合的识别方法,以实现自然环境下成熟期番茄的准确识别。孙红等<sup>[11]</sup>提出使用 SSD-Lite-MobileDet 网络模型识别玉米拔节期冠层, Quan 等<sup>[12]</sup>基于 Faster-RCNN 模型开发了一种玉米幼苗检测机器人,在有大量杂草存在的背景下能有较高的检测精度。Zahid 等<sup>[13]</sup>提出在苹果树修剪机器人研发设计中,采用基于深度学习的图像分割算法为剪枝定位提供数据支持。Philipp 等<sup>[14]</sup>研究了一种田间杂草的语义分割网络,能够对杂草进行分割分类。Shi 等<sup>[15]</sup>研究了一种基于深度学习的表型方法,实现了对植物种苗各组成部分的分割。以上这些研究对本文提出的种苗夹取点定位方法有一定的启发性,另外在夹取机构设计方面,为了尽量避免误伤种苗,并且考虑到夹取之后的移植工作,种苗自动夹取系统需要考虑末端夹取机构的结构、组成材料以及种苗的夹取点和可夹取范围,所以需要设计一种新型夹取机构来完成夹取作业。

蝴蝶兰是最受欢迎的兰花品种之一,但是它们的无性繁殖困难,依赖于植物组织培养技术<sup>[16-18]</sup>。本文以蝴蝶兰种苗为典型种苗对象,分析了蝴蝶兰种苗的形态、颜色特征之后,结合夹取需求,通过深度学习模型训练并提取了蝴蝶兰种苗特征,设计了一种基于语义分割视觉伺服的种苗自动夹取系统,该系统根据种苗自身特性,自动识别出夹取点的位置和姿态,驱动末端夹取

执行机构完成夹取作业。该种苗自动夹取系统拟在保护种苗结构不被破坏的前提下完成夹取工作,以解决人工夹取劳动强度大,劳动内容枯燥的问题。

## 1 基于视觉伺服的种苗自动夹取系统

基于视觉伺服的种苗自动夹取系统的硬件组成包括上位机和机械手两个主要部分,它们之间通过 Wi-Fi 进行通信。上位机通过 USB2.0 端口接收摄像头拍摄的照片,照片发送到伺服控制模块后,经过手眼标定后的坐标转换得到目标的位置和姿态信息,然后由上位机的运动控制模块根据位置和姿态制定夹取计划,通过逆运动学将空间路径转化为机械手的各个关节角度,共计 6 个关节角度和 1 个夹具张开角度发送给机械手,机械手和末端夹具进行夹取动作。图 1 为种苗夹取平台示意图。

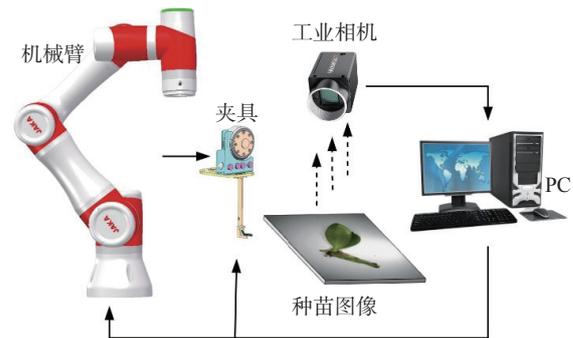


图 1 种苗夹取平台

Fig. 1 Seedling clamping platform

## 2 基于视觉伺服的种苗自动夹取系统

机械手的视觉伺服控制是国内外研究的热点之一,根据摄像机的数量可分为单目、双目和多目。视觉伺服被分为基于图像的视觉伺服<sup>[19]</sup>和基于位置的视觉伺服<sup>[20-21]</sup>。基于位置的视觉伺服相比于基于图像的视觉伺服具有控制器易于实现的优点<sup>[22]</sup>,因此种苗自动夹取系统选用基于位置的视觉伺服控制。本文使用的是海康威视高清单目工业相机。

### 2.1 系统设计

整个控制系统如图 2 所示。首先摄像机拍摄图片,在处理图片中,利用 DP-BiseNetV2 语义分割网络,根据目标颜色和几何特征完成图像分割,再通过分割好的图像颜色特性在 HSV 空间求取夹取点和夹取角度。以目标夹取点和角度作为机械臂的输入,机械臂在得到夹取点和角度信息

后, 判断末端是否超出硬件限制, 机械臂是否在初始位置, 如果二者均无问题, 则机械臂移动到夹取点上方, 并通过调整机械臂关节来改变夹取角度, 完成种苗的夹取。

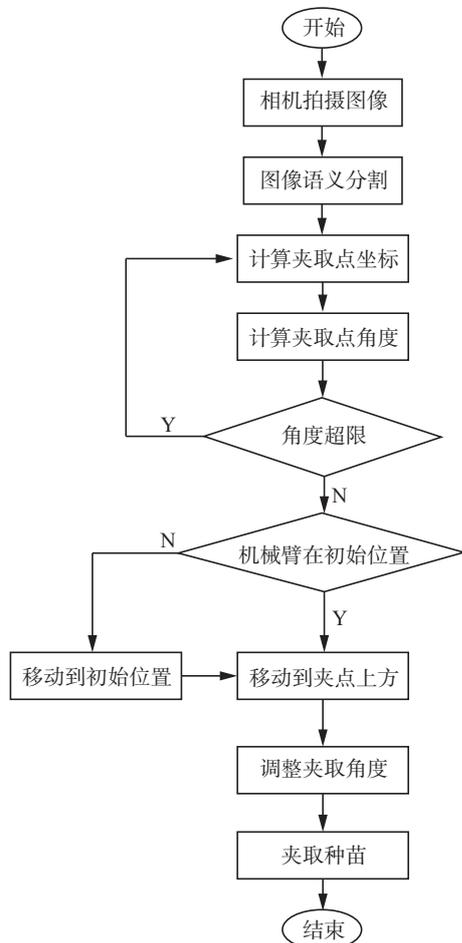


图 2 夹取系统流程  
Fig. 2 Clamping system flow

### 2.2 语义分割视觉定位方法

在进行夹取时, 需要找到合适的夹取点以及夹取角度, 分析了部分种苗器官之间的位置关系, 发现很多种苗的器官之间有固定的相对关系, 且不会随着种苗形状改变而发生位置变化, 于是本文将夹取点选择在主干茎上, 角度通过图像处理的方法得到。由于种苗的形态不是固定的, 传统的图像处理算法局限性太大, 无法识别出每个种苗的姿态, 所以本文中使用时语义分割方法先对种苗进行区域分割, 再利用分割出来的黑色块茎和主茎干的相对关系计算出种苗的位姿。如图 3 所示, 将主干茎的质心点  $S$  作为夹取点,  $G$  是黑色块茎的质心点, 根据两者固定的几何关系, 计算出夹取点的角度信息。分析了部分种苗的根茎关系,

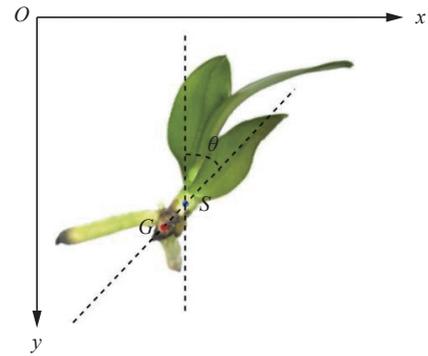


图 3 种苗位姿估计  
Fig. 3 Seedling position estimation

#### 2.2.1 基于改进的 BiSeNetV2 的种苗语义分割

蝴蝶兰种苗图像主要分为 4 部分, 分别是叶片、茎干、根茎、黑色块茎。这几个部分姿态各异, 传统的阈值分割和色差分割算法难以对其进行准确识别和分割。为实现种苗自动夹取系统高效准确夹取, 需要选择识别效率高且准确率较高的语义分割模型<sup>[23-25]</sup>, 本文提出了新型的深度语义分割模型 DP-BiSeNetV2 对蝴蝶兰种苗各组成部分进行识别和分割。DP-BiSeNetV2 的网络结构称为融合位置注意力机制的双边分割网络, 整个网络有两个分支: 语义分支和细节分支。语义分支具有深层和窄通道计数, 允许快速下采样和更多上下文语义信息, 窄通道计数有助于提高速度。细节分支 (DP-Branch) 则相反, 具有浅层和宽通道, 因此可以将注意力集中在局部细节上, 以减少细节损失, 为了降低模型的计算量, 同时提高模型的准确度, 将原始模型 BiSeNetV2 的细节分支中的普通卷积替换成了深度可分离卷积, 并在细节分支中加入了位置注意力机制, 显著提高了准确度。最后, 通过引导聚合层将两个分支的特征合并, 以实现特征的互补融合。这种高效的网络架构在速度和准确性之间提供了极好的平衡。我们将该网络用于语义分割任务, 并调整其数据加载部分以适应数据集的形式。

DP-BiSeNetV2 接收来自工业相机的图片, 并执行像素级语义分割。对图像上目标对象的像素进行分割和着色, 最终输出分割好的图片。图 4 是 DP-BiSeNetV2 的网络结构。图中  $L_1, L_2, L_3$  是深度可分离卷积 (DWConv),  $\varphi$  表示激活函数,  $\otimes$  和  $\oplus$  分别表示元素乘法和元素加法,  $G_1, G_2, G_3, G_4$  表示聚集和扩展层。位置注意力模块中, reshape 表示将特征图  $B, F, D$  的  $C \times H \times W$  格式重塑成  $C \times N$  格式, transpose 表示将矩阵从  $C \times N$  格式转换成  $N \times C$  格式, softmax 表示归一化指数函数。

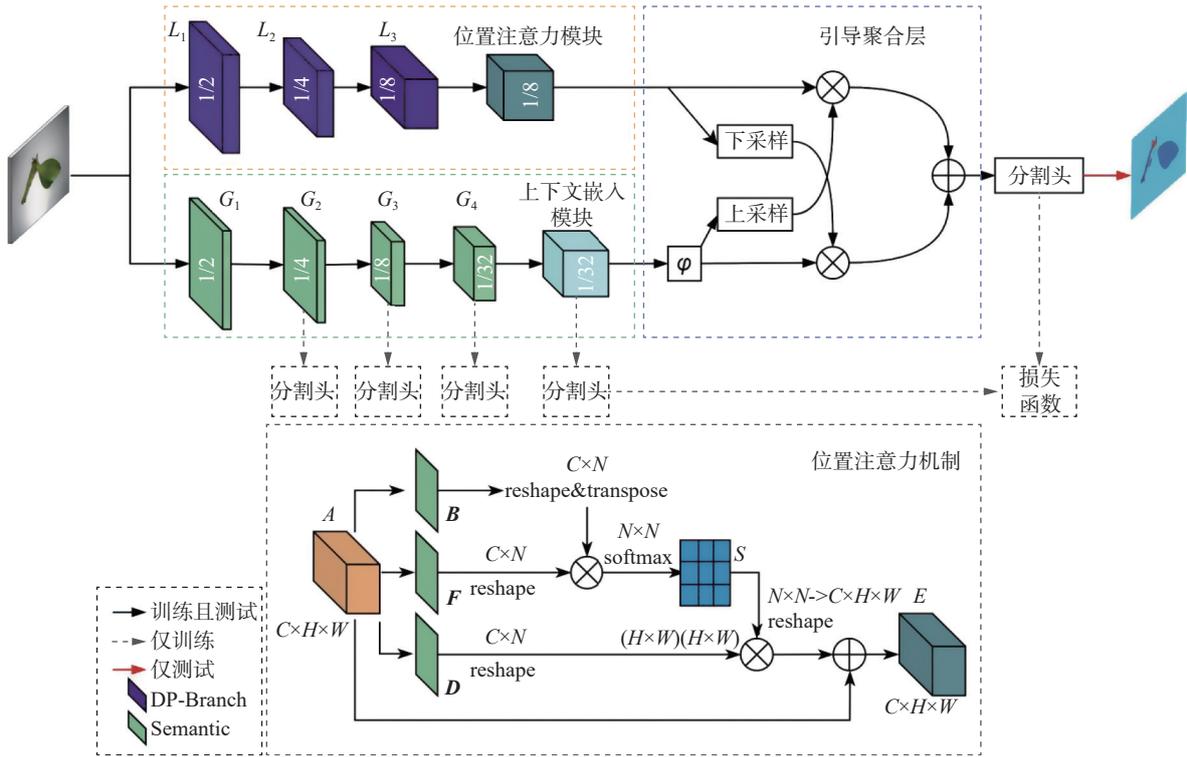


图 4 DP-BiseNetV2 网络结构

Fig. 4 DP-BiseNetV2 network structure

2.2.2 求取夹取点坐标及斜率

原图和分割结果对比如图 5。

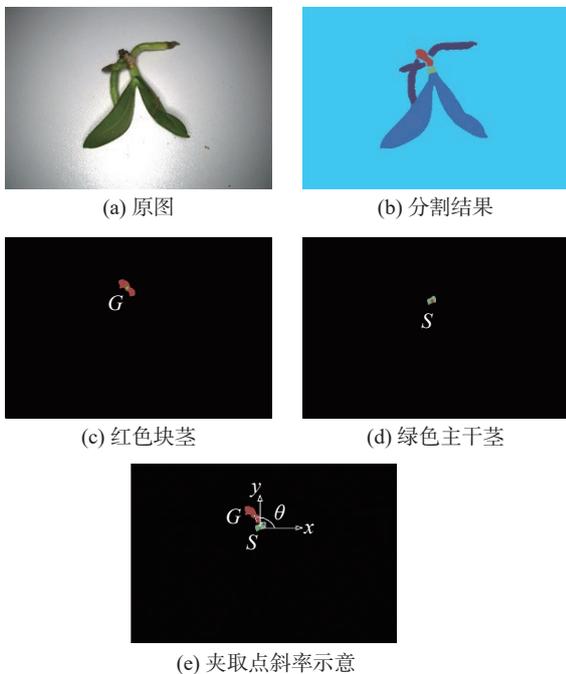


图 5 原图和分割结果对比

Fig. 5 Comparison of original image and segmentation result

DP-BiseNetV2 分割好的结果如图 5(b) 所示, 蓝色部分代表背景, 绿色部分代表主干茎, 红色部分代表黑色块茎, 深蓝色和紫色分别代表叶片和根茎。上文介绍到, 需要获取的是夹取点和夹

取点末端夹具的角度, 所以只关注语义分割结果中蒙版的红色部分和绿色部分。基于这一点, 选择将分割结果转化为 HSV 格式, HSV 颜色空间能通过调整 hue(色调)、saturation(饱和度)和 value(亮度)3 个通道的参数选择提取图片中固定的颜色, 可以把红色部分(黑色块茎)和绿色部分(主干茎)单独分离出来, 如图 5(c)、5(d) 所示。

如图 5 所示, 将这两部分分离出来之后, 根据图像矩的定义求两部分的几何中心, 由于灰度图密度分布均匀, 所以几何中心就是质心, 根据如下公式:

图像的零阶矩:

$$M_{00} = \sum_i \sum_j V(i, j) \tag{1}$$

图像的一阶矩:

$$M_{10} = \sum_i \sum_j i \cdot V(i, j) \tag{2}$$

$$M_{01} = \sum_i \sum_j j \cdot V(i, j) \tag{3}$$

图像的质心为

$$x_c = \frac{M_{10}}{M_{00}}, \quad y_c = \frac{M_{01}}{M_{00}} \tag{4}$$

其中:  $V(i, j)$  是点  $(i, j)$  处的像素值, 图像的零阶矩表示像素值之和, 一阶矩表示横坐标和对应像素值的乘积和纵坐标和对应像素值的乘积。

如此求出两部分各自的质心坐标。将绿色主干茎的质心  $S = (x_s, y_s)$  选为夹取点, 红色根茎的质

心为  $G = (x_G, y_G)$ , 在世界坐标系中将  $S$  点和  $G$  点连接起来得到直线  $SG$ , 计算  $SG$  的斜率  $k_{SG}$ :

$$k_{SG} = \frac{y_S - y_G}{x_S - x_G} \quad (5)$$

得到了夹取斜率后, 根据斜率和角度的关系, 即可求出夹取点角度  $\theta$ 。至此, 已经获得了夹取点坐标和夹取角度信息, 记为  $(x_S, y_S, \theta)$ 。斜率和角度的关系为

$$\theta = \arctan k_{SG} (0^\circ \leq \theta \leq 180^\circ) \quad (6)$$

如图 5(e) 是夹取点角度示意图。

### 2.3 夹取方案设计

作为整个系统的核心部分, 夹取策略应根据目标位姿与当前位姿之间的偏差, 使机械手安全、准确、快速地完成夹取动作。夹取机械臂选择 JAKA-Zu3, 该机械臂的本体结构如图 6 所示, 主要由 6 个关节和 2 个铝制臂杆组成, 基座用于机器人本体和底座连接, 工具端用于机械臂与工具连接。另外, 根据实验平台的可操作空间和摄像头的识别范围, 设计了一种小电机绳驱长杆形夹具来完成夹取任务。



图 6 机械臂本体结构  
Fig. 6 Robotic arm body structure

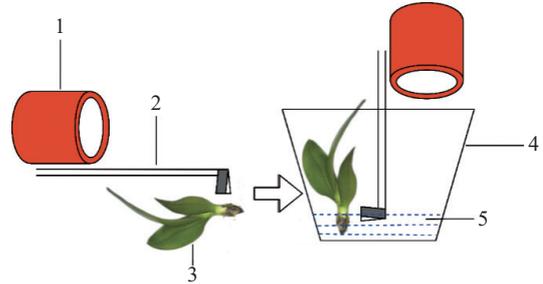
#### 2.3.1 末端夹具

机械臂的执行末端机构是进行种苗夹取和移动的部件, 安装部位在机械臂第 6 个自由度的末端。根据作业要求, 机械臂需要利用该夹具进行取苗操作, 并且需要在取苗之后, 插入到培养基所在的敞口瓶中, 考虑到这一过程种苗的姿态需要旋转  $90^\circ$  (如图 7 所示), 再以工业相机的准确识别范围和机械臂的工作空间为参考, 将夹具的机械结构末端设计成外形为“L”型, 便于在工作时夹苗。

#### 2.3.2 夹取姿态分析

图 8 是末端夹具示意图, 夹取姿态如图 9 所示, 由于种苗所处平台是恒定水平的, 在坐标系  $O_P$  中,  $Z_P$  恒垂直于水平面, 所以夹取姿态的选择中, 末端坐标系  $O_J$  只有  $Z_J$  轴是在旋转的, 结合上文讲到  $Z$  轴的高度固定, 所以机械臂的末端位姿  $(x, y, z, r_x, r_y, r_z)$  中只有  $x, y, r_z$  是变化量, 其余的全

部是固定值。这 3 个未知量在 2.2.2 节已经求取, 即夹取点坐标和夹取角度。



1. 机械臂末端, 2. 夹具, 3. 种苗, 4. 敞口瓶, 5. 培养基

图 7 夹具工作示意图

Fig. 7 Fixture working diagram

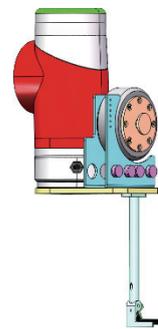


图 8 末端夹具  
Fig. 8 End clamps

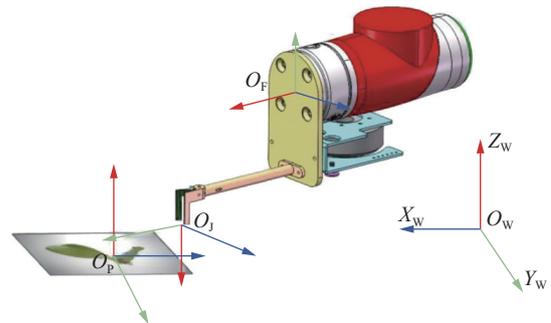


图 9 夹取时末端姿态  
Fig. 9 End posture during clamping

## 3 实验结果与分析

区别于大多数农业机器人在户外场景试验, 本实验面对的是组织培养蝴蝶兰种苗的移植, 这需要无菌环境, 因此, 需要在室内环境搭建工作台。室内搭建的工作台背景比较简单, 与图像采集时使用的场景保持一致。

### 3.1 语义分割实验

#### 3.1.1 评价指标

夹取点识别主要由两部分组成: 语义分割和图像处理。分割结果如图 5(b) 所示, 由于在图像处理时的色彩是不变的, 所以在图像处理部分使用图像一阶矩求质心。由于色彩的不变性, 这一步的输出结果是完全固定的。

在语义分割部分,选择 DP-BiseNetV2 为分割模型,采用帧率 (frame rate, FR)、均交并比 (mean Intersection over Union, mIoU)、像素准确度 (pixel accuracy, pixAcc) 和模型大小 (Size) 作为模型的评价指标。mIoU 是语义分割的标准度量,计算所有类别交集和并集之比的平均值。mIoU 的公式为

$$M_{IoU} = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{i=0}^k p_{ji} - p_{ii}} \quad (7)$$

式中:  $k$  是类别数,  $k+1$  表示加上了背景类,  $i$  表示真实值,  $j$  表示预测值,  $p_{ij}$  表示将类别  $i$  预测为  $j$  的像素的数目,  $p_{ji}$  将类别  $j$  预测为  $i$  的像素的数目,  $p_{ii}$  表示实际类别为  $i$ 、预测类别也为  $i$  的像素的数目。

像素准确度用预测对的像素的数量除以总的像素数量来反映准确度,计算公式为

$$P_{ixAcc} = \frac{\sum_{i=0}^k p_{ii}}{\sum_{i=0}^k \sum_{j=0}^k p_{ij}} \quad (8)$$

帧率是图像处理领域中的定义,是指画面每秒传输帧数,通俗来讲就是指动画或视频的画面数。FR 是测量用于保存、显示动态视频的信息数量。每秒传输帧数越多,所显示的动作就会越流畅。放在本实验中,是为了看模型处理一张图片需要多长时间,由于本系统要求尽量提高工作效率,需要识别速度尽量快,所以 FR 是一个很关键的指标。

### 3.1.2 种苗数据集获取和标注

种苗的数据集构建使用工业相机分别对 A 类蝴蝶兰种苗、B 类蝴蝶兰种苗、A 类模型苗和 B 类模型苗等 4 类种苗进行拍摄,得到原始数据集 (图 10(a)), 然后对原始图像用 labelme 软件进行人工标注,标注完成后得到每幅图像的 JSON 数据 (图 10(b)), 然后把数据集格式改成 Cityscapes 数据集格式,获得对应的掩膜图片 (图 10(c)), 图中红色标签对应的是种苗的叶,绿色部分是主干茎,黄色部分是根茎,蓝色部分是黑块茎,黑色是背景色。将标注完的 2000 张 A 类蝴蝶兰种苗图片按照 8:2 的比例随机分为训练集和验证集,用于模型的有监督训练。将 500 张 B 类蝴蝶兰种苗和 500 张 A、B 两类模型苗按照同样方式进行训练。



图 10 数据集构建过程  
Fig. 10 Dataset construction process

### 3.1.3 种苗分割效果分析

为了测试语义分割模型对种苗的分割效果,实验中使用了 A 类蝴蝶兰种苗、B 类蝴蝶兰种苗、A 类模型苗和 B 类模型苗等 4 种不同种类的种苗进行训练和语义分割操作。DP-BiseNetV2 的分割结果如图 11 所示,图 11(a) 是 A 类蝴蝶兰种苗的分割结果,图 11(b) 是 B 类蝴蝶兰种苗的分割结果,图 11(c)、图 11(d) 分别是 A 类模型苗和 B 类模型苗的分割结果。

在分割效果图中,可以看出模型 DP-BiseNetV2 能够准确地分割出 4 类种苗的各组成部分,能够将种苗的黑色块茎和主干茎分割出来(即图 11 中的红色和绿色部分)。

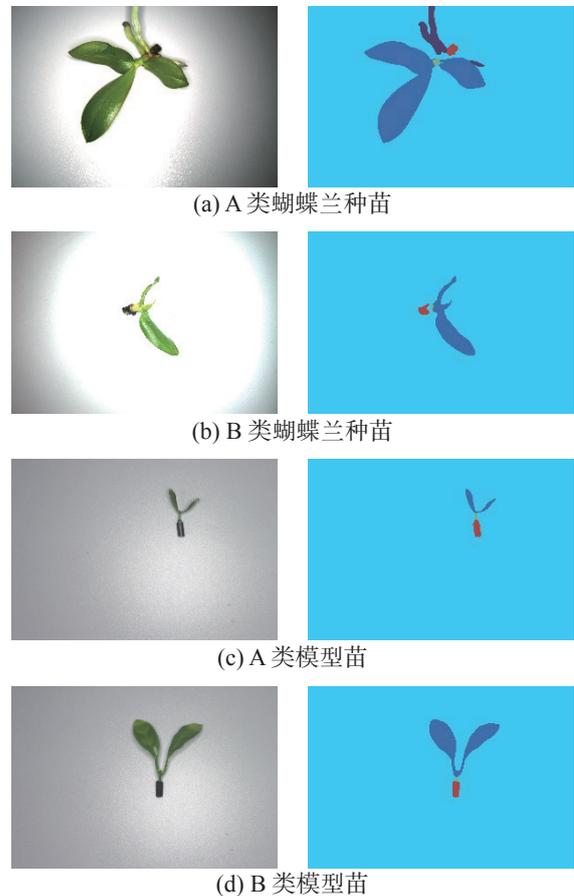


图 11 不同种苗的分割结果  
Fig. 11 Segmentation results of different seedlings

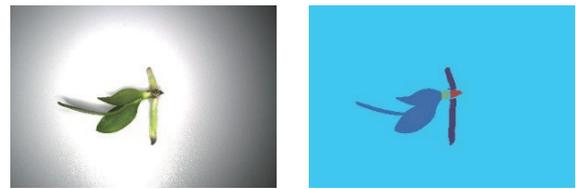
### 3.1.4 A 类蝴蝶兰种苗分割结果与分析

为了验证模型的先进性和优越性,本文以数据集图像和数量最多的 A 类蝴蝶兰种苗作为整体实验对象,对比了 BiseNet<sup>[26]</sup>、BiseNetV2<sup>[27]</sup>、DeeplabV3+<sup>[28]</sup>、DFANet<sup>[29]</sup>、DABNet<sup>[30]</sup>、Fast-SCNN<sup>[31]</sup> 语义分割模型。设置 mIoU、帧率、像素准确度率 (pixAcc) 和模型大小 (Size) 为评价指标。对比实验中,保持模型所设置初始参数一致,输入图像分辨率大小为 3072×2048,检测结果如表 1 所示。

表 1 不同模型分割效果对比  
Table 1 Results of seedling cutting test

模型	mIoU/%	FR/(f/s)	pixAcc/%	Size/MB
BiseNet	58.62	15.92	98.25	53.23
BiseNetV2	61.35	19.62	97.85	13.38
DeeplabV3+	61.79	3.85	<b>98.32</b>	164.21
DFANet	57.21	13.01	98.06	8.14
DABNet	57.01	16.70	98.08	<b>3.01</b>
Fast-SCNN	58.36	<b>36.41</b>	98.14	4.83
DP-BiseNetV2	<b>63.51</b>	22.65	98.28	11.63

原始种苗的输入图像分辨率是  $3072 \times 2048$ , 在这样的高分辨率情况下, 能同时保证准确度和处理图片的速度是目前语义分割研究的热门问题。种苗数据集的背景为纯色, 分割部位与背景的颜色对比明显, 所以 pixAcc 的结果均能达到较高指标, DeeplabV3+ 不属于轻量级语义分割模型, 整体参数量和计算量都要比其余几个模型大, 模型大小为 164.21 MB, 虽然 pixAcc 达到了较高的 98.32%, 但是该模型参数冗余量较大, 影响推理速度。而本文提出的基于 BiseNetV2 框架的 DP-BiseNetV2 的模型大小仅为 11.63 MB, 小于 BiseNetV2, 并且 mIoU 达到了 63.51%, pixAcc 达到了 98.28%, 两项指标均领先于其他几个模型, 在处理图片的速度上虽然不及 Fast-SCNN, 但是 Fast-SCNN 的准确率太低, 无法胜任分割任务。权衡几种模型的优缺点, 本文提出的 DP-BiseNetV2 模型在保证分割精度的同时模型更轻量, 推理速度较快, 适合用于种苗的分割任务。DP-BiseNetV2 的检测结果如图 12 所示。



(d) 种苗 4

图 12 A 类蝴蝶兰种苗的分割结果

Fig. 12 Results of segmentation of class A phalaenopsis seedlings

### 3.2 种苗夹取实验

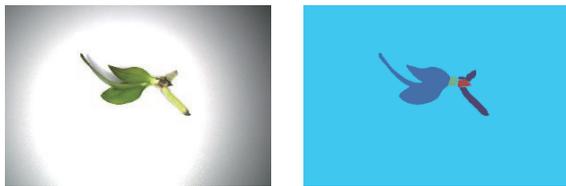
为了验证方法的实际夹取效果, 选取 60 株 A 类蝴蝶兰种苗进行实验。表 2 给出了当使用视觉系统引导夹具到指定的夹取位置, 然后夹具夹取种苗所获得的结果。如果满足以下条件, 则判定夹取操作成功: 1) 指定夹取点位于植株的主干茎上, 2) 夹取并抬起植株成功, 3) 横移 10 cm 种苗不脱落。

表 2 种苗夹取实验结果

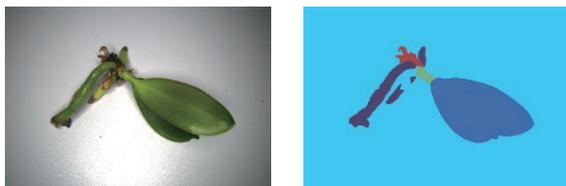
Table 2 Results of seedling clamping experiments

结果	夹取成功	夹取失败		
		坐标错误	夹取脱落	角度超限
种苗数目	49	4	5	2
成功/失败率/%	81.7	6.7	8.3	3.3

种苗夹取实验现场如图 13 所示, 图 13(a) 的机械臂和夹具属于待机状态, 经过视觉系统检测和坐标转换之后, 将夹取点坐标传输给机械臂控制系统, 驱动机械臂到达目标位置, 并驱使夹具夹取种苗。图 13(b) 是机械臂和夹具已经移动到了夹取点正上方, 图 13(c) 是夹具已经成功夹取种苗。



(a) 种苗 1



(b) 种苗 2



(c) 种苗 3



(a) 待机状态 (b) 夹取点上方 (c) 夹取成功

图 13 种苗夹取实验现场

Fig. 13 Seedling clamping experiment site

夹取统计结果如表 2 所示, 其中完成夹取操作条件的种苗有 49 株, 成功率为 81.7%。

不成功夹取统计中, 其中有 4 株种苗是因为识别过程中叶片遮挡住黑色块茎, 导致没能分割出黑色块茎和主干茎, 进行夹取点检测时检测不到夹取位置以致于角度求取失败, 机械臂无法调整夹取角度, 导致夹取失败; 有 5 株种苗因为主干茎比起正常种苗要幼小, 为了保护种苗, 末端夹具在投入使用前要调整好夹取力度, 由于种苗的幼小使得夹取力度不够大, 导致在移动过程中脱落; 2 株种苗由于摆放角度超出机械臂的旋转范

围,导致机械臂角度超限,作业失败,机械臂之所以角度超限,是因为在设计夹具之初,要考虑到夹取之后的栽种工作,以致于在夹取时调整角度并不是只调整机械臂第 6 个自由度,而是调整至少 3 个自由度,于是在夹取种苗过程中机械臂有可能会出奇异点,从而导致夹取作业停止。

实验结果表明,正常条件下,该种苗自动夹取系统能够适应大部分种苗的夹取作业,夹取失败的案例是因为个别苗本身生长结构的偏差造成,在后续的改进中,可以在视觉系统中针对这类苗进行单独分类,进行单独处理。

#### 4 结束语

1) 本文设计了一种基于语义分割视觉伺服的种苗自动夹取系统,提出了轻量化语义分割模型 DP-BiseNetV2 作为种苗识别检测器,基于种苗自身结构设计了由分割模板在 HSV 空间分离引导的夹取点定位算法,通过手眼标定建立视觉采集系统和夹取执行系统间映射关系,配合机械臂夹取控制策略可以完成对蝴蝶兰种苗的自动化夹取作业。针对种苗的夹取需求设计了一种夹具,该夹具能在保证夹取末端稳定、有效完成夹取作业。

2) 对比 BiseNet、BiseNetV2、DeeplabV3+、DABNet、Fast-SCNN 等 5 种常见的语义分割算法,轻量级语义分割算法 DP-BiseNetV2 的 mIoU 达到 63.51%,帧率达到 22.65 f/s,像素准确率达到 98.28%,模型大小为 11.63 MB。

3) 实验证明该种苗自动夹取系统能在保证种苗结构不被破坏的前提下有效完成夹取任务。在 60 株种苗夹取试验中系统夹取合格率高于 81.7%。

针对种苗各部位间可能存在的遮挡或重叠情况,需要控制机械臂先将种苗调整到系统可以操作的姿态,再进行视觉系统识别和夹取,目前单机械臂难以完成如此复杂的夹取任务。下一步研究中,拟在视觉系统中采用识别复杂种苗的单一分支任务,单独对复杂结构种苗进行特殊夹取。

#### 参考文献:

- [1] DEBNATH M, MALIK C, BISEN P. Micropropagation: a tool for the production of high quality plant-based medicines[J]. *Current pharmaceutical biotechnology*, 2006, 7(1): 33–49.
- [2] LEE T J, ZOBAYED S, FIRMANI F, et al. A novel automated transplanting system for plant tissue culture[J]. *Biosystems engineering*, 2019, 181: 63–72.
- [3] 翟永杰, 胡东阳, 苑朝, 等. 基于视觉伺服的蝴蝶兰种苗切割系统设计与试验[J]. *农业工程学报*, 2022, 38(6): 148–156.
- [4] ZHAI Yongjie, HU Dongyang, YUAN Chao, et al. Design and experiments of phalaenopsis seedling cutting system using visual servo[J]. *Transactions of the Chinese society of agricultural engineering*, 2022, 38(6): 148–156.
- [5] NING Zhengtong, LUO Lufeng, DING Xinming, et al. Recognition of sweet peppers and planning the robotic picking sequence in high-density orchards[J]. *Computers and electronics in agriculture*, 2022, 196: 106878.
- [6] RINGDAHL O, KURTSEER P, EDAN Y. Evaluation of approach strategies for harvesting robots: case study of sweet pepper harvesting[J]. *Journal of intelligent & robotic systems*, 2019, 95(1): 149–164.
- [7] VAN H E J, HEMMING J, VAN TUIJL B A J, et al. Collision-free motion planning for a cucumber picking robot[J]. *Biosystems engineering*, 2003, 86(2): 135–144.
- [8] WANG Guohua, YU Yabo, FENG Qingchun. Design of end-effector for tomato robotic harvesting[J]. *IFAC-Papers OnLine*, 2016, 49(16): 190–193.
- [9] MU Longtao, LIU Yadong, CUI Yongjie, et al. Design of end-effector for kiwifruit harvesting robot experiment[C]// 2017 ASABE Annual International Meeting. Washington: American Society of Agricultural and Biological Engineers, 2017.
- [10] SARABU H, AHLIN K, HU Aiping. Leveraging deep learning and RGB-D cameras for cooperative apple-picking robot arms[C]// 2019 ASABE Annual International Meeting. Boston: American Society of Agricultural and Biological Engineers, 2019.
- [11] 李天华, 孙萌, 丁小明, 等. 基于 YOLO v4+HSV 的成熟期番茄识别方法[J]. *农业工程学报*, 2021, 37(21): 183–190.
- [12] LI Tianhua, SUN Meng, DING Xiaoming, et al. Tomato recognition method at the ripening stage based on YOLO v4 and HSV[J]. *Transactions of the Chinese society of agricultural engineering*, 2021, 37(21): 183–190.
- [13] 孙红, 乔金博, 李松, 等. 基于深度学习的玉米拔节期冠层识别[J]. *农业工程学报*, 2021, 37(21): 53–61.
- [14] SUN Hong, QIAO Jinbo, LI Song, et al. Recognition of the maize canopy at the jointing stage based on deep learning[J]. *Transactions of the Chinese society of agricultural engineering*, 2021, 37(21): 53–61.
- [15] QUAN Longzhe, FENG Huaigu, LYU Yingjie, et al. Maize seedling detection under different growth stages and complex field environments based on an improved Faster R-CNN[J]. *Biosystems engineering*, 2019, 184: 1–23.
- [16] ZAHID A, MAHMUD M S, HE Long, et al. Technological advancements towards developing a robotic pruner for apple trees: a review[J]. *Computers and electronics in agriculture*, 2021, 189: 106383.
- [17] LOTTES P, BEHLEY J, CHEBROLU N, et al. Robust joint stem detection and crop-weed classification using image sequences for plant-specific treatment in precision farming[J]. *Journal of field robotics*, 2020, 37(1): 20–34.

- [15] SHI Weinan, VAN DE ZEDDE R, JIANG Huanyu, et al. Plant-part segmentation using deep learning and multi-view vision[J]. *Biosystems engineering*, 2019, 187: 81–95.
- [16] GUO W J, LIN Yuzu, LEE N. Photosynthetic light requirements and effects of low irradiance and daylength on phalaenopsis amabilis[J]. *Journal of the American society for horticultural science*, 2012, 137(6): 465–472.
- [17] LIN Mingju, HSU B D. Photosynthetic plasticity of Phalaenopsis in response to different light environments[J]. *Journal of plant physiology*, 2004, 161(11): 1259–1268.
- [18] LIU Y C, TSENG K M, CHEN C C, et al. Warm-night temperature delays spike emergence and alters carbon pool metabolism in the stem and leaves of Phalaenopsis aphroide[J]. *Scientia horticulturae*, 2013, 161: 198–203.
- [19] XU De, LU Jinyan, WANG Peng, et al. Partially decoupled image-based visual servoing using different sensitive features[J]. *IEEE transactions on systems, man, and cybernetics: systems*, 2017, 47(8): 2233–2243.
- [20] VICENTE P, JAMONE L, BERNARDINO A. Towards markerless visual servoing of grasping tasks for humanoid robots[C]//2017 IEEE International Conference on Robotics and Automation. Singapore: IEEE, 2017: 3811–3816.
- [21] CHANG W C. Robotic assembly of smartphone back shells with eye-in-hand visual servoing[J]. *Robotics and computer-integrated manufacturing*, 2018, 50: 102–113.
- [22] 杨月全, 秦瑞康, 李福东, 等. 机器人视觉伺服控制研究进展与挑战 [J]. *郑州大学学报 (理学版)*, 2018, 50(2): 41–48.
- YANG Yuequan, QIN Ruikang, LI Fudong, et al. The development and challenges of studies on robot visual servo control[J]. *Journal of Zhengzhou university (natural science edition)*, 2018, 50(2): 41–48.
- [23] YU Changqian, WANG Jingbo, PENG Chao, et al. BiSeNet: bilateral segmentation network for real-time semantic segmentation[C]//European Conference on Computer Vision. Cham: Springer, 2018: 334–349.
- [24] 李国进, 黄晓洁, 李修华, 等. 采用轻量级网络 MobileNetV2 的酿酒葡萄检测模型 [J]. *农业工程学报*, 2021, 37(17): 168–176, 317.
- LI Guojin, HUANG Xiaojie, LI Xiuhua, et al. Detection model for wine grapes using MobileNetV2 lightweight network[J]. *Transactions of the Chinese society of agricultural engineering*, 2021, 37(17): 168–176, 317.
- [25] 宁政通, 罗陆锋, 廖嘉欣, 等. 基于深度学习的葡萄果梗识别与最优采摘定位 [J]. *农业工程学报*, 2021, 37(9): 222–229.
- NING Zhengtong, LUO Lufeng, LIAO Jiaxin, et al. Recognition and the optimal picking point location of grape stems based on deep learning[J]. *Transactions of the Chinese society of agricultural engineering*, 2021, 37(9): 222–229.
- [26] 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 Rendu, LIU Xiaoyang, et al. Apple positioning based on YOLO deep convolutional neural network for picking robot in complex background[J]. *Transactions of the Chinese society of agricultural engineering*, 2019, 35(3): 164–173.
- [27] YU Changqian, GAO Changxin, WANG Jingbo, et al. BiSeNet V2: bilateral network with guided aggregation for real-time semantic segmentation[J]. *International journal of computer vision*, 2021, 129(11): 3051–3068.
- [28] CHEN L C, ZHU Yukun, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//European Conference on Computer Vision. Cham: Springer, 2018: 833–851.
- [29] LI Hanchao, XIONG Pengfei, FAN Haoqiang, et al. DFANet: deep feature aggregation for real-time semantic segmentation[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2020: 9514–9523.
- [30] LI Gen, YUN I, KIM J, et al. DABNet: depth-wise asymmetric bottleneck for real-time semantic segmentation[EB/OL]. (2019–07–26)[2022–12–26]. <https://arxiv.org/abs/1907.11357>.
- [31] POUDEL R P K, LIWICKI S, CIPOLLA R. Fast-SCNN: fast semantic segmentation network[EB/OL]. (2019–02–12)[2022–12–26]. <https://arxiv.org/abs/1902.04502>.

### 作者简介:



翟永杰, 教授, 博士, 主要研究方向为电力视觉。主持国家自然科学基金面上项目 2 项、河北省自然科学基金项目 2 项、横向科研项目 20 余项。编著 1 部, 参编教材 1 部、著作 3 部。发表学术论文 30 余篇。



王家豪, 硕士研究生, 主要研究方向为电力视觉与智能机器人。



徐大伟, 讲师, 博士, 主要研究方向为机器人。参与国家重点研发计划项目 1 项、国家自然科学基金面上项目 1 项, 主持横向科研项目 2 项。