

基于三维点云的群株生菜分割模型与鲜质量预测方法

杨 意¹, 马术涛¹, 黄灿增¹, 朱文鼎¹, 张观康¹, 刘厚诚², 辜 松^{3*}

(1. 华南农业大学电子工程学院(人工智能学院), 广州 510640; 2. 华南农业大学园艺学院, 广州 510640;
3. 华南农业大学工程学院, 广州 510640)

摘 要: 密集场景下群株生菜的有效分割与参数获取是植物工厂生长监测中的关键环节。针对群株生菜中个体生菜鲜质量提取问题, 该研究提出一种利用实例分割模型提取个体生菜点云, 再以深度学习点云算法预测个体鲜质量的方法。该方法以群株生菜为研究对象, 利用深度相机采集群株生菜俯视点云, 将预处理后的点云数据输入实例分割模型 Mask3D 中训练, 实现背景与生菜个体的实例分割, 之后使用鲜质量预测网络预测个体生菜鲜质量。试验结果表明, 该模型实现了个体生菜点云的分割提取, 无多检和漏检的情况。当交并比(intersection over union, IoU)阈值为 0.75 时, 群株生菜点云实例分割的精确度为 0.924, 高于其他实例分割模型; 鲜质量预测网络实现了直接通过深度学习处理点云数据, 预测个体生菜鲜质量的目的, 预测结果的决定系数 R^2 值为 0.90, 均方根误差值为 12.42 g, 优于从点云中提取特征量, 再回归预测鲜质量的传统方法。研究结果表明该研究预测生菜鲜质量的精度较高, 为利用俯视单面点云提取群株生菜中个体生菜表型参数提供了一种思路。

关键词: 作物; 实例分割; 点云处理; 回归预测; 深度学习

doi: 10.11975/j.issn.1002-6819.202408160

中图分类号: S126; S609

文献标志码: A

文章编号: 1002-6819(2025)-01-0192-08

杨意, 马术涛, 黄灿增, 等. 基于三维点云的群株生菜分割模型与鲜质量预测方法[J]. 农业工程学报, 2025, 41(1): 192-199. doi: 10.11975/j.issn.1002-6819.202408160 <http://www.tcsae.org>
YANG Yi, MA Shutao, HUANG Canzeng, et al. Segmenting collective lettuce to predict fresh weight using three-dimensional point cloud[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2025, 41(1): 192-199. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202408160 <http://www.tcsae.org>

0 引 言

实时的品质监测是植物工厂生菜生产中的重要环节, 采集到的表型参数是进行环境调控^[1-2]、产量预测^[3-4]、采收节点确定的重要理论依据。但目前植物工厂生产中, 多采用人工从群株生菜中采集个体样本, 手工测量几何形态参数、去根称量获得鲜质量的监测方式, 采样不全面、效率低。

随着精准农业的发展, 对作物生长状态实时、全面监测的需求日益增长。基于 RGB 图像的植物几何形态参数提取是目前非接触式监测的主要手段^[5-6]。但由于 RGB 图像缺乏深度数据, 因而无法全面捕捉作物的完整信息。为此, 研究人员利用深度(Depth)图像或几何信息作为 RGB 图像的补充开展相关研究, ZHANG 等^[7]使用深度图像补充 RGB 图像, 校正了生菜鲜质量和干质量等参数, 解决单源数据模型精度不佳的问题。LIN 等^[8]利用 RGB 图像、深度图像以及几何特征估计生菜鲜质量, 多源数据融合预测鲜质量的 R^2 为 0.938。朱逢乐等^[9]基

于深度学习融合 RGB 图像和深度图像估算生菜表型参数, 鲜质量预测的 R^2 可提升至 0.982。上述研究均针对单株作物场景, 证明几何信息和深度信息能有效补充 RGB 图像中的缺失信息。

虽然深度图像能够丰富 RGB 图像采集到的作物信息, 但其与 RGB 图像对齐过程较为繁琐。相较之下三维点云数据包含颜色、深度以及详细的几何信息, 是研究植物形态结构更直接有效的数据形式^[10]。SODHI 等^[11]通过多视角图像重建高粱植株, 结合支持向量机与条件随机场分割高粱茎叶, 语义分割精度达 80.2%。MORTENSEN 等^[12]利用田间生菜群株单面点云和高度滤波器分割点云, 通过几何算法计算分割点云的表面积、株高和体积等参数, 并分别回归生菜鲜质量, 其中表面积与两种生菜鲜质量的 R^2 最高, 分别为 0.83 和 0.94。LOU 等^[13]通过 ToF 相机获取单株植物图像, 利用基于颜色区域生长分割方法分割植物与土壤, 以几何对称方式补全点云计算植株面积, 通过提取的参数回归鲜质量。LIU 等^[14]基于多视图立体方法重建植物三维点云, 使用迭代方式实现了油菜点云器官分割。上述研究结果证明三维点云在作物表型参数提取方面的应用潜力, 但传统点云分割算法, 大多需要凭借经验设置阈值参数, 才能得到较理想的分割结果^[15-16]。

随着 PointNet 等^[17-18]网络的提出, 基于深度学习的植物点云分割成为更有潜力的解决方案。GUO 等^[19]通

收稿日期: 2024-08-21 修订日期: 2024-11-29

基金项目: “十四五”国家重点研发计划项目(2021YFD2000703)

作者简介: 杨意, 博士, 副教授, 研究方向为机器视觉作物表型参数提取与现代园艺生产智能装备。Email: yangyih2007@163.com

*通信作者: 辜松, 博士, 教授, 研究方向为现代园艺生产装备。

Email: sgu666@sina.com

过改进 PointNet++模型实现了甘蓝植株点云的茎叶器官语义分割, 分割准确率达 0.95。JIN 等^[20] 基于不同生长阶段的玉米植株激光雷达数据, 提出一种体素卷积神经网络模型, 实现了玉米植株点云的茎叶分类与分割。LI 等^[21] 提出一种双功能深度学习网络, 在烟草等重建点云中实现了叶片实例分割, 分割的平均精确率为 83.30%。上述研究成果证明了深度学习网络从三维点云数据中分割目标作物的优越性。然而, 目前基于作物点云数据集的深度学习模型其目的多为实现作物器官分类和分割^[22], 通常需要耗时的植株重建过程, 且多为语义分割, 难以满足植物工厂快速、批量获取单个作物表型信息的需求。而基于深度学习实例分割模型, 如 Mask3D^[23] 模型可对室内规则点云进行实例分割, 但尚未有形态复杂的作物分割应用。

在农业领域中, 除了需要在复杂环境中准确分割个体作物点云, 还要求能够将点云特征与作物非表观参数(比如鲜质量)等^[24-25] 进行关联。目前研究大多基于从单株场景提取分割点云, 通过计算作物点云的体素体积、高度等^[26] 参数, 再利用上述参数回归鲜质量等非表观参数。在群株场景下分割个体点云, 提取作物表型参数的研究较少。

因此, 针对植物工厂密集种植环境和有限空间限制下, 个体生菜鲜质量的精确提取问题, 本文将基于群株生菜俯视(top-view)单面点云图像, 提出一种结合实例分割模型和深度学习算法的鲜质量预测方法。研究首先创建群株生菜点云数据集; 之后, 使用实例分割模型训练群株生菜点云图像, 获取单株生菜点云; 最后, 对分割得到的单株生菜点云, 通过鲜质量预测网络预测其鲜质量。

1 材料与方法

本研究选取世界范围内广泛种植的生菜为研究对象, 具体选择在广东省良种引进服务公司白云基地(后简称广良)大量种植的奶油生菜品种。试验中, 图像采集时间分别为移植后培育第 14、21、28 天。共采集 108 板群株生菜点云, 每个种植板 15 株, 共获取 1 620 个单株生菜点云。

广良植物工厂内部环境如图 1, 由多层的培养架、光照设备和通风设备等组成, 内部环境复杂, 设备间易相互干涉, 架设多台相机多视角采集图像的难度较高, 因此本研究探讨采用俯视单面点云预测生菜鲜质量的可行性。

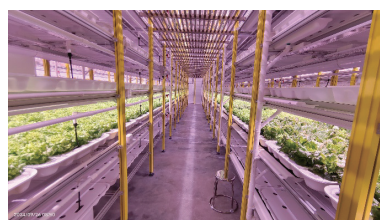


图 1 植物工厂环境图

Fig.1 Plant factory environment Diagram

研究整体流程如图 2 所示, 分为数据获取与预处理、点云标注、点云分割并提取个体生菜点云、鲜质量预测与回归分析等部分。首先利用深度相机获取群株生菜俯视点云图像并进行预处理, 得到仅包含生菜群株和种植板的点云图, 用于创建群株生菜点云数据集, 训练点云实例分割模型。其次采用点云实例分割模型 Mask3D 分割获取个体点云图。最后, 通过鲜质量预测网络预测单株生菜鲜质量。

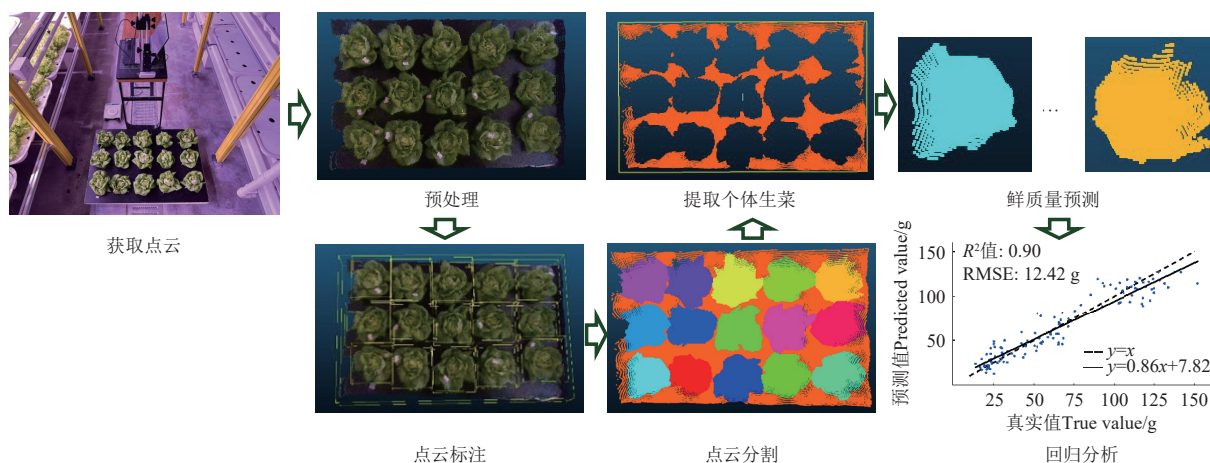


图 2 生菜鲜质量预测流程

Fig.2 Lettuce fresh weight prediction process

1.1 群株生菜数据的获取与处理

Intel RealSense D435i 相机是一款基于结构光原理的深度相机, 运用可见光、红外光和图像处理技术, 实现对物体表面三维感知。本研究中, D435i 处于深度成像模式, 像素分辨率选择 848×480 像素, 捕获群株生菜生成具有深度信息的彩色点云数据。为保证整个种植板处

于视场中, 经测试物距应设置为 1.2 m。

数据预处理的硬件环境为 Intel(R) Core(TM) i5-8300H, 主频 2.30 GHz, 操作系统为 Window 10。在 Pycharm 的编程环境下, 使用点云库 open3d 0.16.0 实现点云数据采集、直通滤波和去噪等数据预处理。原始和预处理后的生菜点云如图 3。

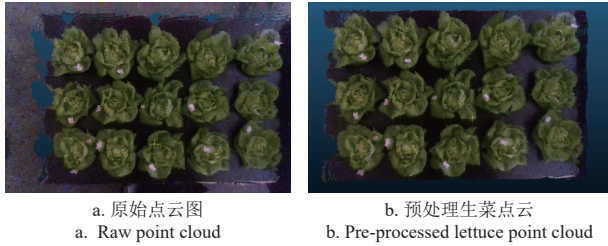


图3 预处理前后的生菜点云图

Fig.3 Point cloud of lettuce before and after preprocessing

1.2 个体生菜点云数据的标注

为了后续模型训练,需对预处理后的群株生菜点云进行人工标注,本文使用三维点云处理软件 Cloudcompare_v2.13 进行点云标注,将点云中各点分类为生菜或背景,同时参考 S3DIS 数据集^[27]的实例标注模式,区分不同生菜实例,结果如图4,其中每个方框内为一个生菜实例对象。

1.3 个体生菜点云分割网络

群株生菜点云中包含着多株生菜及种植板,生菜与背景粘连,且生菜个体数量多,若使用传统的区域生长算法分割,难以调整法线阈值等参数值使其适配多个个

体曲面,个体生菜间分割困难。在众多分割模型中,基于深度学习的 Mask3D 模型无需手动调整阈值参数,能够有效处理室内规则点云实例分割任务,故而本研究采用深度学习模型 Mask3D 对群株生菜进行实例分割,通过学习不同生菜点云间的特征实现群株生菜中个体生菜分割,其结构如图5a所示,包括特征主干网络和 Transformer 解码器 2 个模块。该模型以三维整体点云为输入,利用主干网络提取多分辨率层次特征,与实例化查询进行点积,最后返回点云的语义分类和二进制实例掩码^[23]。

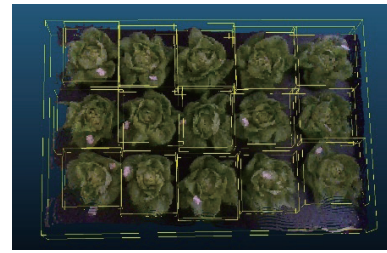
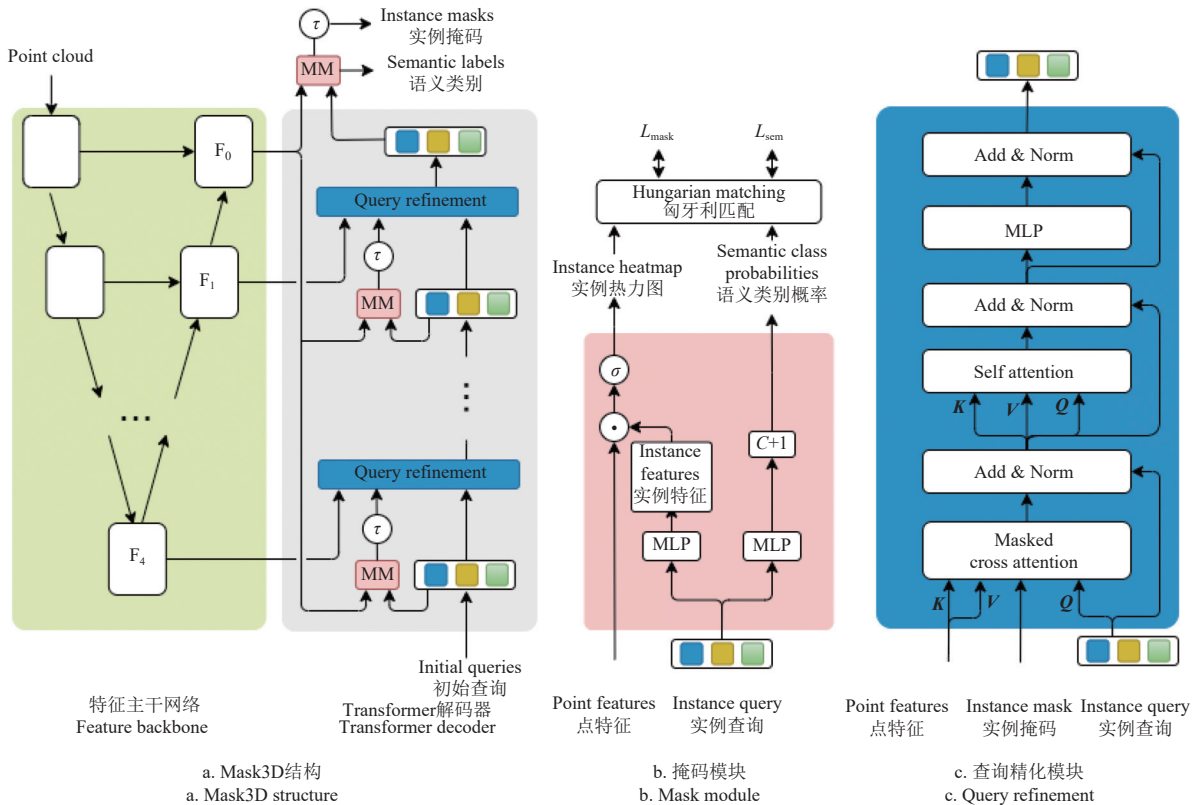


图4 生菜点云标注图

Fig.4 Lettuce point cloud annotation



注: $F_0 \sim F_4$ 代表不同分辨率的特征; τ 代表阈值过滤; MM 代表掩码模块 (图 5b); Query Refinement 代表查询精化模块 (图 5c); L_{mask} 与 L_{sem} 分别代表分割的掩码损失和语义损失, C 代表类别数, σ 和 $\cdot$ 分别代表 sigmoid 函数和点积; Add & Norm 代表残差连接和层归一化, K 、 V 、 Q 分别代表关键向量、值向量、查询向量。

Note: $F_0 \sim F_4$ represent the features with different resolutions; τ represents the threshold filtering; MM represents the mask module (figure 5b); Query Refinement represents the query refinement (figure 5c); L_{mask} and L_{sem} represent the mask loss and semantic loss of segmentation, respectively; C stands for number of classes; MLP stands for multilayer perceptron; σ and $\cdot$ stand for sigmoid function and dot product respectively; Add & Norm stands for residual connection and layer normalization; K , V , and Q represent key vectors, value vectors, and query vectors, respectively.

图5 Mask3D 结构说明^[23]Fig.5 Mask3D structural illustrations^[23]

特征主干网络是基于 MinkowskiEngine^[28] 的 U-Net (图 5a), 其具有对称的编码器和解码器结构。编码器

将输入的群株生菜点云体素化, 将连续的点云数据转换为离散的体素网格, 使其适用于卷积操作。同时, 编码

器使用稀疏卷积提取生菜点云特征，并同时生菜点云进行降采样，逐层提取不同分辨率下的特征。解码器则将提取的点特征逐步还原到原始分辨率，并与提取到的生菜点云特征进行融合。

Transformer 解码器是由图 5b 的掩码模块和图 5c 的查询精化堆叠组成。图 5b 中的掩码模块接收经过细化的实例查询和点特征，通过点积计算生成代表每个点与实例查询相似度的热力图。基于点积计算结果，生成该生菜实例在点云位置中的二进制编码。图 5c 中的查询精化模块作用是优化实例查询，通过交叉细化生菜点云特征，并利用自注意力机制在实例级别进行交互，逐步细化实例查询。

Mask3D 处理点云的核心是实例查询（图 5b），每个实例均表示场景中的一个生菜对象，并预测相应的点级别实例掩码。为此，实例查询通过 Transformer 解码器进行迭代精化，允许实例查询对多分辨率特征（ $F_0 \sim F_4$ ）进行交叉优化，对其他实例查询进行自精化。重复此过程，得到最终的精化实例查询集合。最后，掩码模块结合点特征消耗精化后的实例查询，根据点特征与实例查询之间的结果返回每个实例的类别（如生菜类或背景类）和其在整体点云中的位置。

1.4 鲜质量预测网络

鲜质量预测（fresh weight prediction, FWP）网络的功能是通过分割后的个体生菜点云预测个体生菜鲜质量。该网络的输入是 Mask3D 分割后单株生菜的点云数据。由于点云具有无序性且缺乏固定结构，传统深度学习模型（如卷积神经网络）通常对输入数据的顺序和位置敏感，难以直接处理无序点云数据。而 PointNet 通过使用对称函数保证了模型对于输入不同排序具有置换不变性，使得该深度学习模型拥有处理无序点云的能力。因此，本研究使用基于 PointNet 构建的回归网络预测生菜鲜质量，结构如图 6，包括数据准备模块、特征提取模块以及由多层感知机（multilayer perception, MLP）组成的回归模块。

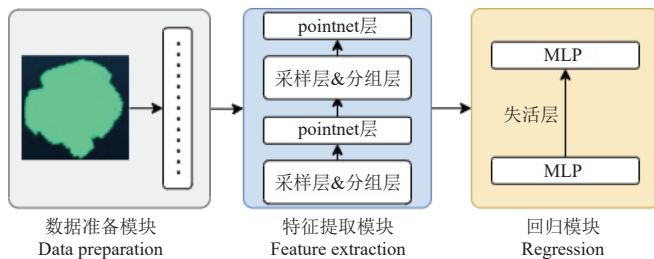


图 6 鲜质量预测网络结构图

Fig.6 Fresh weight prediction (FWP) network structure diagram

数据准备模块通过最远点采样，将个体生菜点云采样为满足模型输入要求的 1 024 点。特征提取模块由采样层、分组层和 pointnet 层堆叠组成。采样层通过设置不同的比例参数选取不同的点作为中心点，采用最远点采样策略，使中心点覆盖整体轮廓，从而有效捕获数据的全局特征。分组层以采样层选择的中心点为聚类中心，

设置不同的半径值，构建局部区域集，捕捉局部结构信息。pointnet 层通过多层感知机提取分组层构造的局部区域特征，在各维度使用最大池化操作得到全局特征。

回归模块由两层感知机组成，每个 MLP 包括 1 024 个输入神经元、512 个隐藏神经元以及一个输出神经元组成。用于将 pointnet 层提取到的特征与鲜质量关联，预测生菜鲜质量。

1.5 模型训练

模型在硬件环境为 Intel Core i9-12900KF CPU，NVIDIA GeForce RTX 3 090 Ti GPU，软件环境为 Ubuntu 22.04 LTS 的电脑上训练，同时基于 CUDA 11.3 和 cuDNN 8.2 加速。训练分为分割阶段和预测阶段。分割阶段利用 Mask3D 分割个体生菜。Mask3D 模型训练的批量大小、初始学习率和迭代次数分别设为 4、0.001 和 1 000 次，使用 Adam W 优化器优化网络。预测阶段，利用 FWP 网络建立生菜点云与其鲜质量的对应关系。初始学习率、批次大小和迭代次数分别设为 0.001、8 和 500，训练采用 Adam 优化器。

分割阶段训练集、测试集分别设为 86 板、22 板群株生菜。分割阶段测试集的 22 板（每板包含 15 株生菜）共计 330 株单株生菜点云，作为预测阶段的数据集，训练集和测试集各为 250 株和 80 株。

2 结果与分析

2.1 总体评价指标

为评价分割模型的有效性，本文采用平均精确度 A_p 衡量模型在不同类别下的分割精度，即

$$P = \frac{TP}{TP + FP} \times 100\% \quad (1)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (2)$$

$$A_p = \frac{1}{n} \sum_{k=1}^n (P(k) \cdot \Delta R(k)) \quad (3)$$

式中 n 是真实实例的数量， $P(k)$ 是在前 k 个检测结果中正确检测到实例的比例， $\Delta R(k)$ 是在第 $(k-1)$ 个和第 k 个检测结果之间的召回率变化。真正例（true positive, TP）、假正例（false positive, FP）和假负例（false negative, FN）分别对应正确检测、误检和漏检。计算 A_p 时，只有当分割掩码与真实点云的交并比（intersection over union, IoU）值大于阈值，才会被纳入计算范围。

为评价鲜质量预测网络的性能，采用决定系数 R^2 和均方根误差 RMSE 作为评价指标，计算式如下：

$$R^2 = 1 - \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{\sum_{i=1}^m (y_i - \bar{y})^2} \quad (4)$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (5)$$

其中 y_i 为生菜点云预测的鲜质量值， $\bar{y}$ 为所有预测鲜质量

值均值, y_i' 为生菜真实鲜质量值, m 为样本数量。

2.2 分割结果与分析

2.2.1 Mask3D 模型结果与分析

表 1 为 Mask3D 模型在不同 IoU 阈值 (A_{p25} 、 A_{p50} 、 A_{p75}) 时生菜与背景两类点云的分割精确度。点云图像中背景部分多为平面或近似平面, 识别难度较低, 因此背景类别的分割精确度即使在 A_{p75} 时也能达到 0.995。但生菜几何结构复杂、叶片层叠多, 所以模型在生菜类别的分割精确度有所下降, 在 A_{p75} 时生菜的分割精确度降至 0.924, 但背景和生菜的平均分割精度也能达到了 0.960 以上, 综合数据表现, 说明该方法分割生菜点云的可行性较高。

表 1 不同 IoU 阈值下分割精确度

Table 1 Segmentation accuracy under different IoU thresholds

实例对象 Instance object	A_{p75}	A_{p50}	A_{p25}
背景 Background	0.995	1.000	1.000
生菜 Lettuce	0.924	0.936	1.000
均值 Averages	0.960	0.968	1.000

注: A_{p25} 、 A_{p50} 、 A_{p75} 指交并比为 0.25、0.50、0.75 时, 模型的分割精确度。
Note: A_{p25} , A_{p50} , and A_{p75} refer to the segmentation accuracy of the model when the IoU is 0.25, 0.50, and 0.75.

在掩码分割结果的可视化分析中, 采用 open3d 对生菜个体及背景着色, 结果如图 7。对比图 7b 和 7c, 图 7c 的个体生菜分割存在少量的孔洞, 但各生菜实例无多检或漏检。观察发现孔洞主要出现在生菜与背景的交界处, 形成的原因: 1) 标注过程中生菜与背景的边界较难确定, 导致边界判定不一致, 使得 Mask3D 模型训练时出现阈值较低的区域。2) 模型设定了得分阈值, 用以滤掉得分较低的边缘区域, 导致出现孔洞。

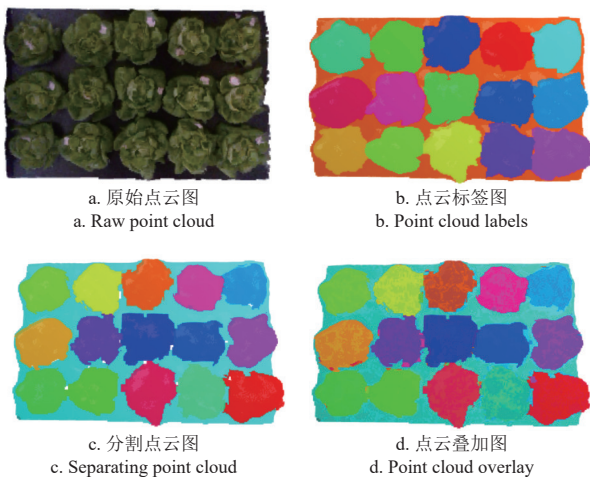


图 7 生菜分割结果的可视化

Fig.7 Visualization of lettuce segmentation results

2.2.2 与不同分割模型的对比试验

为评估 Mask3D 的点云分割效果, 本研究选择了 4 个流行的点云实例分割模型 Pointcept^[29]、Jsnet^[30]、Asis^[31] 和 Jsis3d^[32] 与之对比。表 2 为 Mask3D 与上述模型在 IoU 为 0.75 阈值时的分割结果。结果显示, 对于个体生菜分割, Mask3D 以 0.924 的分割精度领先于其他模

型, 说明相比于 Pointcept 等其他模型, Mask3D 区分不同生菜实例个体的能力强; 对于背景分割结果, Mask3D、Pointcept 和 Jsnet 模型对整体背景的分割精度均达到 0.990 以上, 但 Pointcept 和 Jsnet 对不同生菜实例的分割能力不足, 出现个体生菜实例边界误判, 边缘识别错误或漏检、多检的情况。对生菜和背景两类目标的分割精确度均值, Mask3D 分割模型以 0.960 领先于表 2 中 Pointcept 等其他四个模型, 证实了其处理生菜点云分割任务的均衡性。

表 2 不同分割模型的分割精确度 A_{p75} 对比

Table 2 Segmentation accuracy (A_{p75}) for different segmentation models

实例对象 Instance object	Mask3D	Pointcept	Jsnet	Asis	Jsis3d
背景 Background	0.995	0.992	1.000	0.733	0.623
生菜 Lettuce	0.924	0.832	0.776	0.717	0.806
均值 Averages	0.960	0.912	0.888	0.725	0.715

综上, Mask3D 模型无需依靠经验调整参数去适应不同的分割平面, 虽然分割后在生菜与背景交界处存在少量孔洞, 但不同生菜实例都被几乎完整的分割, 无漏检和多检, 较好地完成了预定的个体生菜分割目标, 说明使用 Mask3D 模型实现群株生菜中个体生菜实例分割具有较高的可行性。

2.3 鲜质量预测结果与分析

2.3.1 FWP 网络结果与分析

图 8 是以 Mask3D 分割点云、手动分割点云作为 FWP 网络输入, 预测得到的鲜质量与真实鲜质量间的关系图。结果表明, 分割点云预测鲜质量与鲜重真实值相关性为 0.90 (图 8a)。与同采用单面点云提取群株生菜参数的文献 [12] 相比, 此相关性处于其 2 种生菜预测结果 0.83 和 0.94 之间, 但文献 [12] 采用单变量输入, 方法的输入变量单一、抗风险能力不强, 且其需要提取参数和预测等多个步骤。与手动分割点云 (图 8b) 相比, Mask3D 分割点云在 R^2 值略低于手动分割点云, 但根据实测, 由于点云图的复杂结构, 手动分割单板生菜点云约需 1 h, Mask3D 在测试集上分割每板平均用时约为 963 ms。考虑时间消耗, Mask3D 在效率和成本方面具有明显的优势。综合相关性数据和方法的自动化流程, 表明本研究方法预测生菜鲜质量具有较高可行性, 更适合应用在植物工厂等数据量较大的场景。

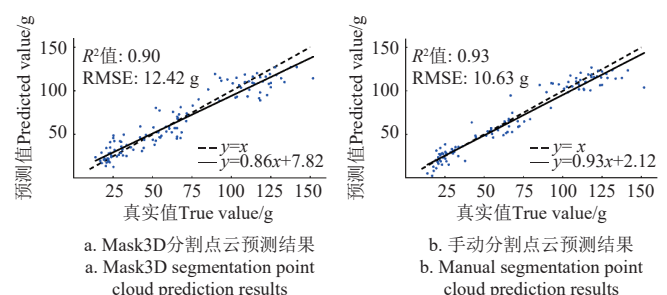


图 8 FWP 网络生菜鲜质量预测分析

Fig.8 FWP network fresh weight predictive analysis

2.3.2 与传统机器学习方法结果对比

为对比 FWP 网络的预测效果, 本论文借鉴了相关文献计算生菜点云预测因子的方式^[12,33], 即通过 open3d 计算 Maske3D 分割后个体生菜点云的冠幅、投影面积、高度和体积, 并将其作为生菜鲜质量预测因子, 使用 5 种典型的机器学习算法预测其鲜质量, 计算结果的 R^2 和 RMSE 值。

为分析上述所选冠幅等 4 因子与生菜鲜质量的相关性, 绘制相关性热力图 (图 9)。其中鲜质量与冠幅、投影面积、体积相关性均达到了 0.88 以上, 而株高与鲜质量的相关性为 0.60。株高相关性低的原因: 1) 分割后生菜点云与种植板分离, 株高计算缺乏参考点, 使其与鲜质量的相关性降低; 2) 多数作物具有矮壮特性, 其株高与鲜质量的相关性较冠幅和投影面积等参数低。相关性结果表明, 当点云存在部分残缺时, 部分提取的特征量会失真 (株高), 但其他特征量 (如冠幅等) 与鲜质量仍保持较强相关性。

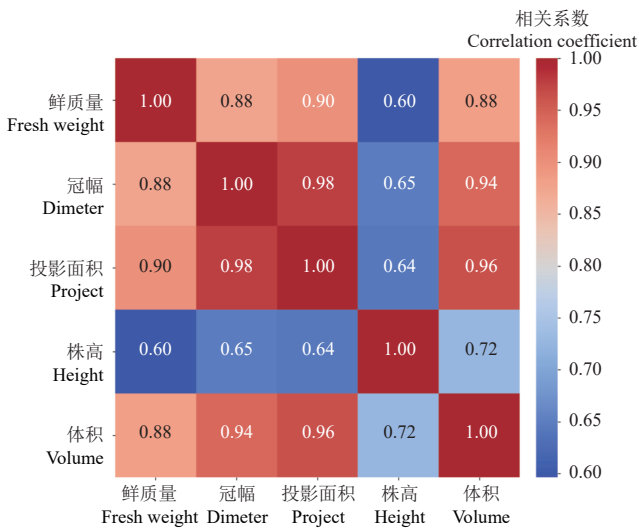


图9 生菜鲜质量与预测因子的相关性热力图

Fig.9 Heat map of correlation between lettuce fresh weight and predictors

图 10 为各算法在测试集上预测生菜鲜质量的交叉验证散点图, 对比其他 5 种方法, FWP 网络的 R^2 值最高, 分别比线性 (套索回归和线性回归) 和非线性 (多项式回归、随机森林回归和决策树回归) 拟合算法的最大值高出 0.12 和 0.07。FWP 网络的 RMSE 最低, 分别比线性和非线性拟合算法的最小值低 4.53、2.64 g, 且 FWP 网络的预测结果数据的离散程度更低。综合上述结果, 证明 FWP 网络在准确度和误差值方面均优于其他 5 种典型算法。

综上所述, FWP 网络通过深度学习方式直接处理点云, 具有自动学习点云特征的能力, 自动化程度明显优于传统从点云中提取特征量, 再回归预测鲜质量的方式; FWP 网络的预测不涉及不同特征的计算结果和参数的调整, 预测生菜鲜质量的精度与误差值也均优传统方法。

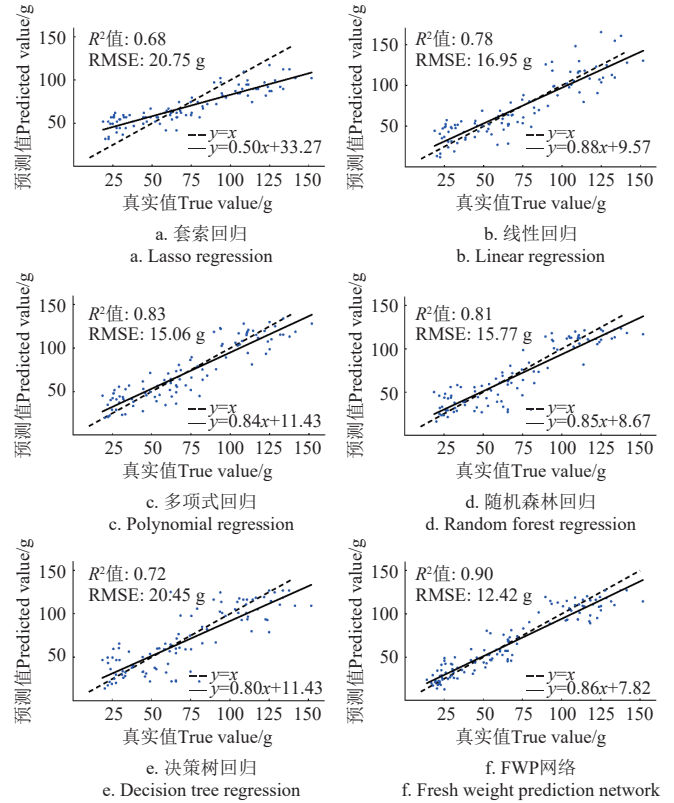


图 10 生菜鲜质量的回归分析图

Fig.10 Regression analysis plot for lettuce fresh weight

3 结 论

本文针对密集场景中作物有效分割和个体参数获取的问题, 基于 Top-view 单面点云, 提出了一种结合 Mask3D 实例分割模型和鲜质量预测 (fresh weight prediction) 网络的个体生菜鲜质量预测方法, 为利用 Top-view 单面点云图像评估群株作物鲜质量提供一种参考方法。主要结论如下:

1) Mask3D 分割模型实现了生菜点云与背景点云的语义分割和不同生菜个体间的实例分割。试验结果表明, 在 IoU 值为 0.75 时, 群株生菜的实例分割精确度 0.924, 高于同训练环境下实例分割模型。

2) FWP 网络使用深度学习方式直接处理个体生菜点云数据, 实现了生菜点云与其鲜质量的直接关联。生菜鲜质量预测结果的决定系数 R^2 为 0.90, 均方根误差为 12.42 g, 优于传统利用提取特征量预测鲜质量的方式。

本研究实现了利用 Top-view 单面群株点云预测密植环境下的个体生菜鲜质量的目标, 为设施蔬菜的无损监测提供新的思路。未来将使用半监督方式辅助人工标注, 提高边缘标注的准确性。也将通过采集更多叶菜品种的点云图像, 构建多类型叶菜数据集, 验证该方法在其他叶菜鲜质量预测上的可靠性。

[参 考 文 献]

- [1] HE R, ZHANG Y, SONG S, et al. UV-A and FR irradiation improves growth and nutritional properties of lettuce grown in an artificial light plant factory[J]. Food Chemistry, 2021, 345:

- 128727.
- [2] GANG M S, KIM H J, KIM D W. Estimation of greenhouse lettuce growth indices based on a two-stage CNN using RGB-D images[J]. *Sensors*, 2022; 22(15): 5499.
 - [3] 李毅念, 杜世伟, 姚敏, 等. 基于小麦群体图像的田间麦穗计数及产量预测方法[J]. *农业工程学报*, 2018, 34(21): 185-194.
LI Yinian, DU Shiwei, YAO Min, et al. Method for wheatear counting and yield predicting based on image of wheatear population in field[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2018, 34(21): 185-194. (in Chinese with English abstract)
 - [4] BROWN T B, CHENG R, SIRAULT X R, et al. TraitCapture: Genomic and environment modelling of plant phenomic data[J]. *Current Opinion in Plant Biology*, 2014, 18: 73-79.
 - [5] DU J, LU X, FAN J, et al. Image-based high-throughput detection and phenotype evaluation method for multiple lettuce varieties[J]. *Frontiers in Plant Science*, 2020, 11: 563386.
 - [6] 崔庭源, 杨其长, 张义, 等. 基于迁移学习和卷积神经网络的生菜鲜重估测[J]. *中国农业大学学报*, 2022, 27(11): 197-206.
CUI Tingyuan, YANG Qichang, ZHANG Yi, et al. Lettuce fresh weight estimation based on transfer learning and convolutional neural network[J]. *Journal of China Agricultural University*, 2022, 27(11): 197-206. (in Chinese with English abstract)
 - [7] ZHANG Q, ZHANG X, WU Y, et al. TMSCNet: A three-stage multi-branch self-correcting trait estimation network for RGB and depth images of lettuce[J]. *Frontiers in Plant Science*, 2022, 13: 982562.
 - [8] LIN Z, FU R, REN G, et al. Automatic monitoring of lettuce fresh weight by multi-modal fusion based deep learning[J]. *Frontiers in Plant Science*, 2022, 13: 980581.
 - [9] 朱逢乐, 严霜, 孙霖, 等. 基于深度学习多源数据融合的生菜表型参数估算方法[J]. *农业工程学报*, 2022, 38(9): 195-204.
ZHU Fengle, YAN Shuang, SUN Lin, et al. Estimation method of lettuce phenotypic parameters using deep learning multi-source data fusion[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2022, 38(9): 195-204. (in Chinese with English abstract)
 - [10] MIAO T, WEN W, LI Y, et al. Label3DMAize: Toolkit for 3D point cloud data annotation of maize shoots[J]. *Gigascience*, 2021, 10(5): giab031.
 - [11] SODHI P, VIJAYARANGAN S, WETTERGREEN D. In-field segmentation and identification of plant structures using 3D imaging[C]//2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Vancouver, BC, Canada: IEEE, 2017: 5180-5187.
 - [12] MORTENSEN A K, BENDER A, WHELAN B, et al. Segmentation of lettuce in coloured 3D point clouds for fresh weight estimation[J]. *Computers and Electronics in Agriculture*, 2018, 154: 373-381.
 - [13] LOU M, LU J, WANG L, et al. Growth parameter acquisition and geometric point cloud completion of lettuce[J]. *Frontiers in Plant Science*, 2022, 13: 947690.
 - [14] LIU J, LIU Y, DOONAN J. Point cloud based iterative segmentation technique for 3D plant phenotyping[C]//2018 IEEE International Conference on Information and Automation (ICIA). Wuyishan, China: IEEE, 2018: 1072-1077.
 - [15] YAN J, TAN F, LI C, et al. Stem-leaf segmentation and phenotypic trait extraction of individual plant using a precise and efficient point cloud segmentation network[J]. *Computers and Electronics in Agriculture*, 2024, 220: 108839.
 - [16] GE L, ZOU K, ZHOU H, et al. Three dimensional apple tree organs classification and yield estimation algorithm based on multi-features fusion and support vector machine[J]. *Information Processing in Agriculture*, 2022, 9(3): 431-442.
 - [17] CHARLES R Q, SU H, KAICHUN M, GUIBAS L J. PointNet: Deep learning on point sets for 3D classification and segmentation[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017: 77-85.
 - [18] QI C R, YI L, SU H, et al. Pointnet++: Deep hierarchical feature learning on point sets in a metric space[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). Long Beach, California, USA: NIPS, 2017: 5105-5114.
 - [19] GUO R, XIE J, ZHU J, et al. Improved 3D point cloud segmentation for accurate phenotypic analysis of cabbage plants using deep learning and clustering algorithms[J]. *Computers and Electronics in Agriculture*, 2023, 211: 108014.
 - [20] JIN S, SU Y, GAO S, et al. Separating the structural components of maize for field phenotyping using terrestrial LiDAR data and deep convolutional neural networks[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(4): 2644-2658.
 - [21] LI D, SHI G, LI J, et al. PlantNet: A dual-function point cloud segmentation network for multiple plant species[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2022, 184: 243-263.
 - [22] WANG L, ZHENG L, WANG M. 3D Point cloud instance segmentation of lettuce based on PartNet[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). New Orleans, LA, USA: IEEE, 2022: 1646-1654.
 - [23] SCHULT J, ENGELMANN F, HERMANS A, et al. Mask3D: Mask transformer for 3D semantic instance segmentation[C]//2023 IEEE International Conference on Robotics and Automation (ICRA). London, UK: IEEE, 2023: 8216-8223.
 - [24] OHASHI Y, ISHIGAMI Y, GOTO E. Monitoring the growth and yield of fruit vegetables in a greenhouse using a three-dimensional scanner[J]. *Sensors*, 2020, 20(18): 5270.
 - [25] DUBE N, BRYANT B, SARI-SARRAF H, et al. Cotton boll distribution and yield estimation using three-dimensional point cloud data[J]. *Agronomy Journal*, 2020, 112(6): 4976-4989.
 - [26] LI J, WANG Y, ZHENG L, et al. Towards end-to-end deep RNN based networks to precisely regress of the lettuce plant height by single perspective sparse 3D point cloud[J]. *Expert Systems with Applications*, 2023, 229: 120497.
 - [27] ARMENI I, SENER O, ZAMIR A R, et al. 3D Semantic

- parsing of large-scale indoor spaces[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 1534-1543.
- [28] CHOY C, GWAK J, SAVARESE S. 4D Spatio-temporal ConvNets: Minkowski convolutional neural networks[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA: IEEE, 2019: 3070-3079.
- [29] JIANG L, ZHAO H, SHI S, et al. PointGroup: Dual-set point grouping for 3D instance segmentation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Seattle, WA, USA: 2020: 4866-4875.
- [30] ZHAO L, TAO W. JSNet: Joint instance and semantic segmentation of 3D point clouds[C]//Proceedings of AAAI Conference on Artificial Intelligence. New York, NY, USA: Association for the Advancement of Artificial Intelligence, 2020: 12951-12958.
- [31] WANG X, LIU S, SHEN X, et al. Associatively segmenting instances and semantics in point clouds[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA: IEEE, 2019: 4091-4100.
- [32] PHAM Q H, NGUYEN D T, HUA B S, et al. JSIS3D: Joint semantic-instance segmentation of 3D point clouds with multi-task pointwise networks and multi-value conditional random fields[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA: IEEE, 2019: 8819-8828.
- [33] ZHANG Y, LI M, LI G, et al. Multi-phenotypic parameters extraction and biomass estimation for lettuce based on point clouds[J]. *Measurement*, 2022, 204: 112094.

Segmenting collective lettuce to predict fresh weight using three-dimensional point cloud

YANG Yi¹, MA Shutao¹, HUANG Canzeng¹, ZHU Wending¹, ZHANG Guankang¹,
LIU Houcheng², GU Song^{3*}

(1. College of Electronic Engineering (College of Artificial Intelligence), South China Agricultural University, Guangzhou 510640, China;

2. College of Horticulture, South China Agricultural University, Guangzhou 510640, China; 3. College of Engineering, South China Agricultural University, Guangzhou 510640, China)

Abstract: Acquiring the individual parameters of collective lettuce under dense scenarios can greatly contribute to environmental regulation, yield prediction, and harvest timing determination in the growth monitoring center of the plant factory. Traditional monitoring can often involve the manual measurement of geometric parameters and root removal for fresh weight determination, leading to the less comprehensive and efficient. Fortunately, non-destructive monitoring can be expected to extract the crop phenotypic parameters using machine vision and machine learning at present. However, most existing machine learning exhibited certain limitations in the application of lettuce point clouds. For instance, the majority of application scenarios have been focused on the organ segmentation of individual crops. It is still lacking in the individual segmentation of collective plants. Additionally, the extraction of crop phenotypes from point clouds can often rely mainly on manually predefined feature quantities. There is a high demand to fully explore the effective phenotypic information within the lettuce point clouds. In this study, instance segmentation was proposed to process the point cloud data of collective crops. Subsequently, deep learning of point clouds was employed to predict the fresh weight of individual crops. The collective lettuce was also taken as the research object. A depth camera was also utilized to collect the single-plane point clouds of the collective lettuce. After point cloud preprocessing, the data was then input into the instance segmentation model (Mask3D) for training. A feature backbone network was employed to extract the features from the point clouds of the collective lettuce. A Transformer decoder was utilized to process the instance queries. Point cloud features were integrated with the instance queries through the mask module. A mask was then generated for each instance. The background and lettuce point clouds were segmented to distinguish the individual lettuce. Finally, the fresh weight prediction (FWP) network was employed to predict the fresh weight of individually segmented lettuce. The feature extraction network (PointNet) was utilized to extract the features from the segmented point clouds of individual lettuce. A multilayer perceptron was also employed to regressively predict the fresh weight of the lettuce. Experimental results indicated that the segmentation and extraction of individual lettuce point clouds were successfully achieved without over- or under-detection on the point cloud dataset. When the Intersection over Union (IoU) threshold was set to 0.75, the average precision (A_p) of instance segmentation was 0.924 for collective lettuce point clouds, superior to the instance segmentation models, such as Jsnet. Furthermore, the direct prediction of lettuce fresh weight reduced the errors associated with the manual feature extraction during processing using deep learning point cloud. The coefficient of determination (R^2) and the root mean squared error (RMSE) were 0.90 and 12.42 g, respectively, indicating the superior accuracy of the fresh weight prediction network. Traditional and manual feature extraction from point cloud parameters with machine learning was achieved in the maximum R^2 of 0.83 and the minimum RMSE of 15.06 g. Therefore, the deep learning instance segmentation and point cloud regression can be expected to estimate the fresh weight of collective lettuce, indicating exceptional better performance. The finding can also provide significant importance for the growth monitoring, yield estimation, and harvest timing determination of facility vegetables.

Keywords: crop; instance segmentation; point cloud processing; regression prediction; deep learning