

基于单视角 RGBD 图像的柑橘果实三维重建与表型检测方法

徐胜勇^{1,2} 易同舟^{1,2} 秦子轶^{1,2} 樊清涛^{1,2} 杨宏磊^{1,2} 李善军^{1,2}

(1. 华中农业大学工学院, 武汉 430070; 2. 农业农村部长江中下游农业装备重点实验室, 武汉 430070)

摘要: 水果表型的测量和分析是植物育种和遗传学研究的一个重要领域。单视角 RGBD 图像的表型检测方法通量高、成本低,但受限于传感器分辨率和视角,通常无法获取果实的表面积和体积等数据。本文提出了一种基于 PFNET 的点云补全网络改进方法,可使用深度相机获取的类球形果实单视角点云进行高精度三维重建并进行表型无损测量。为解决补全网络输入比例不固定的问题,提出了一种自适应几何补全策略将单视角点云补全为近似的半球。在 PFNET 网络框架上增加了第 4 尺度,以充分利用 KINECT 相机获取的稠密点云,有利于复杂形状和细节丰富的结构补全。通过引入四头自注意力模块,能更好地捕捉点云中各点间的相互依赖和空间关系,提升网络特征提取能力。增添了果实点云优化模块,解决原网络生成点云存在局部扩散的问题并提升点云质量,模拟人工测量方式设计了针对性的表型检测方法。实验结果表明,该方法与结构光三维扫描仪获取的柑橘果实点云质量接近,三维重建还原度高。对于横径、纵径、表面积和体积 4 种表型检测的 R^2 均大于 0.96,平均测量精度均超过 93.24%。与 RGBD 图像法相比,单果检测时间增加 17.97 s,但横纵径检测精度大幅提高,且能一次测量 4 项表型参数。与三维扫描仪方法相比,检测精度差值在 4 个百分点以内,但速度超过 48 倍,硬件成本只有后者的 1/10,且易于实现自动化。本文方法在检测精度、运行速度、硬件成本和自动化程度上具有较好的平衡,是一种低成本、综合性能高的三维重建技术,有广泛应用于类球形果实表型无损测量的潜力。

关键词: 球形果实; 表型检测; 点云补全; 深度学习; PFNET; RGBD 图像

中图分类号: TP391.41

文献标识码: A

文章编号: 1000-1298(2025)03-0080-11

OSID:



Three-dimensional Reconstruction and Phenotype Detection of Citrus Fruits Based on Single-view RGBD Images

XU Shengyong^{1,2} YI Tongzhou^{1,2} QIN Ziyi^{1,2} FAN Qingtao^{1,2} YANG Honglei^{1,2} LI Shanjun^{1,2}

(1. School of Engineering, Huazhong Agricultural University, Wuhan 430070, China

2. Key Laboratory of Agricultural Equipment in Mid-lower Yangtze River, Ministry of Agriculture and Rural Affairs, Wuhan 430070, China)

Abstract: The measurement and analysis of fruit phenotypes are crucial aspects of plant breeding and genetics research. Methods for phenotype detection using single-view RGBD images offer high throughput and low cost but are limited by sensor resolution and perspective, often failing to obtain data such as the surface area and volume of fruits. An improved method was proposed based on PFNET that used a depth camera to capture single-view point clouds of spherical-like fruits for high-precision 3D reconstruction and non-invasive phenotype measurements. To address the issue of varying input scales in the completion network, an adaptive geometric completion strategy was introduced to transform single-view point clouds into approximate hemispheres. The addition of a fourth scale to the PFNET framework enhanced the utilization of dense point clouds acquired by KINECT cameras, facilitating the completion of complex shapes and structures rich in detail. By incorporating a four-head self-attention module, the network's ability to capture interdependencies and spatial relationships among points in the point cloud was improved, enhancing feature extraction capabilities. An optimized fruit point cloud module was added to resolve issues with local diffusion in the original network-generated point clouds and to improve their

收稿日期: 2024-12-28 修回日期: 2025-02-10

基金项目: 现代农业(柑橘)产业技术体系岗位科学家项目(CARS-26)

作者简介: 徐胜勇(1980—),男,副教授,博士,主要从事 AI 视觉检测技术研究,E-mail: xsy@mail.hzau.edu.cn

通信作者: 李善军(1977—),男,教授,博士,主要从事柑橘全程机械化研究,E-mail: shanjunlee@mail.hzau.edu.cn

quality. A targeted phenotype detection method, designed to mimic manual measurements, was also proposed. Experimental results showed that this method achieved point cloud quality comparable to that obtained by structured light 3D scanners for citrus fruits, with high fidelity in 3D reconstruction. For the detection of four phenotypes—transverse diameter, longitudinal diameter, surface area, and volume—the R^2 values exceeded 0.96, with average measurement accuracy above 93.24%, approaching that of 3D scanners but at 50 times of the efficiency and one-tenth of the cost. Compared with RGBD image methods, the single-fruit detection time was increased by 17.97 s, but there was a significant improvement in transverse and longitudinal diameter detection accuracy, allowing for the simultaneous measurement of four phenotypic parameters. Compared with the 3D scanner method, the difference in detection accuracy was within 4 percentage points, but the speed was more than 48 times faster, with hardware costs reduced to only one-tenth of the latter's and easier implementation of automation. This method struck a good balance between detection accuracy, operational speed, hardware cost, and level of automation, offering a cost-effective, high-performance 3D reconstruction technology with great potential for non-invasive measurement of spherical-like fruit phenotypes.

Key words: spherical fruit; phenotypic detection; point cloud completion; deep learning; PFNET; RGBD image

0 引言

水果表型的测量和分析是植物育种和遗传学研究的一个重要方面,各种研究采用了不同的方法来表征水果表型并研究潜在的遗传^[1]。果实体积、表面积、横径和纵径^[2-3]等物理特征表现出丰富的遗传多样性,为育种学家和生命科学工作者提供了宝贵的数据资源。随着育种工作者对作物产量目标的提升,所需要采集的表型信息量也相应增加,精确且高通量地获取几何特征具有重要意义。

目前主流的检测方法仍采用人工方式,如使用游标卡尺测量果实直径、排水法测量体积等,使用面积仪测量表面积,需要多人协作且过程繁琐、耗时。基于视觉成像技术的表型分析技术能较好地解决上述问题,其表型特征提取方法有无损、可扩展且易于自动化等优势。基于视觉成像的技术按照成像维度可以分为二维机器视觉和三维立体视觉。早期研究通常依赖于二维视觉图像技术,多用于果实分级和分拣、果实轮廓检测、果实的识别和分割。基于二维图像的表型测量需要进行人工标定且检测精度较低,无法直接进行表面积和体积检测。而随着三维扫描仪技术和深度相机等^[4-5]3D技术的发展,三维立体视觉已成为高通量表型分析的主流方法^[6]。当前,用于果实表型检测的三维立体视觉技术主要包括三维扫描仪和RGBD图像处理等。

三维扫描仪方法是通过扫描仪重建出完整果实点云后进行表型检测,精度最高,经常作为三维重建的基准^[7-8]。基于三维扫描仪的3D方法在实际应用中依然面临着一系列挑战,操作复杂且耗时,数据处理和存储需要较高的计算资源,限制了其普及应用^[9]。

RGBD图像处理是使用深度相机作为采集设备的检测方法。RGBD相机近距离内测量精度高,且成本较低,主要分为多视角和单视角重建与测量^[10]。RGBD多视角重建与测量采用多台相机和电动转盘采集多个视角下的点云信息,将不同视角下点云进行拼接最后进行表型测量,相较于三维扫描仪具有更快的运行速度。多相机情形下需要人为进行标定,或者使用电动转盘,装置成本相较于单视角测量方法有所增加,检测精度不如三维扫描仪^[11-13]。相比RGBD多视角,单视角的重建与测量装置成本更低、速度更快、流程更为繁琐,兼备了测量效率和实用性,已经在果实领域得到了广泛的应用^[14-16]。

基于深度学习的残缺点云补全方法成为研究热点,使得单视角果实的三维重建及表型全面精确测量成为可能。研究者提出了多种深度学习网络,如PCN^[17]、TopNet^[18]、ECG^[19]、MSN^[20]、VRCNet^[21]和PFNET^[22]等。综合来看,PFNET在保留几何特性、细节特征丰富、高效性、高保真度等方面表现出显著优势^[23]。因此,本文基于PFNET网络的优秀特性,提出一种使用RGBD单视角图像的柑橘果实三维重建与表型测量新方法,采用半球和深度学习两阶段补全方法得到果实的完整点云,大幅提高横纵径的检测精度,并能够准确测量表面积和体积。该方法仅使用一台RGBD相机,无需相机或目标旋转即可对类球果实进行高精度三维重建,结构简单、硬件成本低,有集成高通量表型无损检测平台的潜力。

1 材料与方 法

1.1 实验材料

华中农业大学园艺林学学院提供了去除枝叶的

3个品种商品橘,赣南脐橙517个、沃柑97个、丑橘54个,共计668个柑橘,于2023年11月在华中农业大学用农业机器人采集实验数据,包括3类:3D点云补全数据集,用于点云补全网络的开发与测试;单视角RGBD图像,用于补全网络的应用与测试;人工测量果实表型数据,用于验证表型检测方法的性能。

1.2 方法

1.2.1 整体技术流程

图1展示了算法流程,包括RGBD图像采集与点云生成、果实点云半球补全、果实点云网络补全、果实点云优化及表型检测共4个环节。

(1) RGBD图像采集与点云生成

使用KINECT相机采集柑橘果实的单视角RGBD图像数据,对采集的彩色对齐图进行实例分割并编号,按编号使用对应的连通域二值图像对深度图进行掩膜,获取对应编号下的果实深度信息,生成柑橘果实单视角初始点云,并通过统计滤波得到

待补全点云,如图1a所示。

(2) 果实点云半球补全

基于类球形果实的形态特征对上半部分进行补全,从待补全点云中提取出最大横径处的点云,并拟合出圆心与整圆,确定旋转轴。计算点云缺失程度,以确定旋转次数和旋转步长角度。将待补全点云沿旋转轴旋转并与原点云进行拼接,得到拼接后点云,如图1b所示。

(3) 果实点云网络补全

首先采用Reeyee Pro结构光扫描仪获取柑橘完整3D点云并制作点云补全数据集,使用该数据集进行点云补全算法的训练、开发和测试。接着使用训练完毕的 M_1M_2 -PFNET网络对拼接后点云进行补全,并根据现实坐标尺度对网络生成的归一化点云进行等比例缩放,最后进行点云上色得到现实尺度下果实彩色点云,如图1c所示。

(4) 点云优化及表型检测

对现实尺度下点云进行平滑、网格化、平均采样

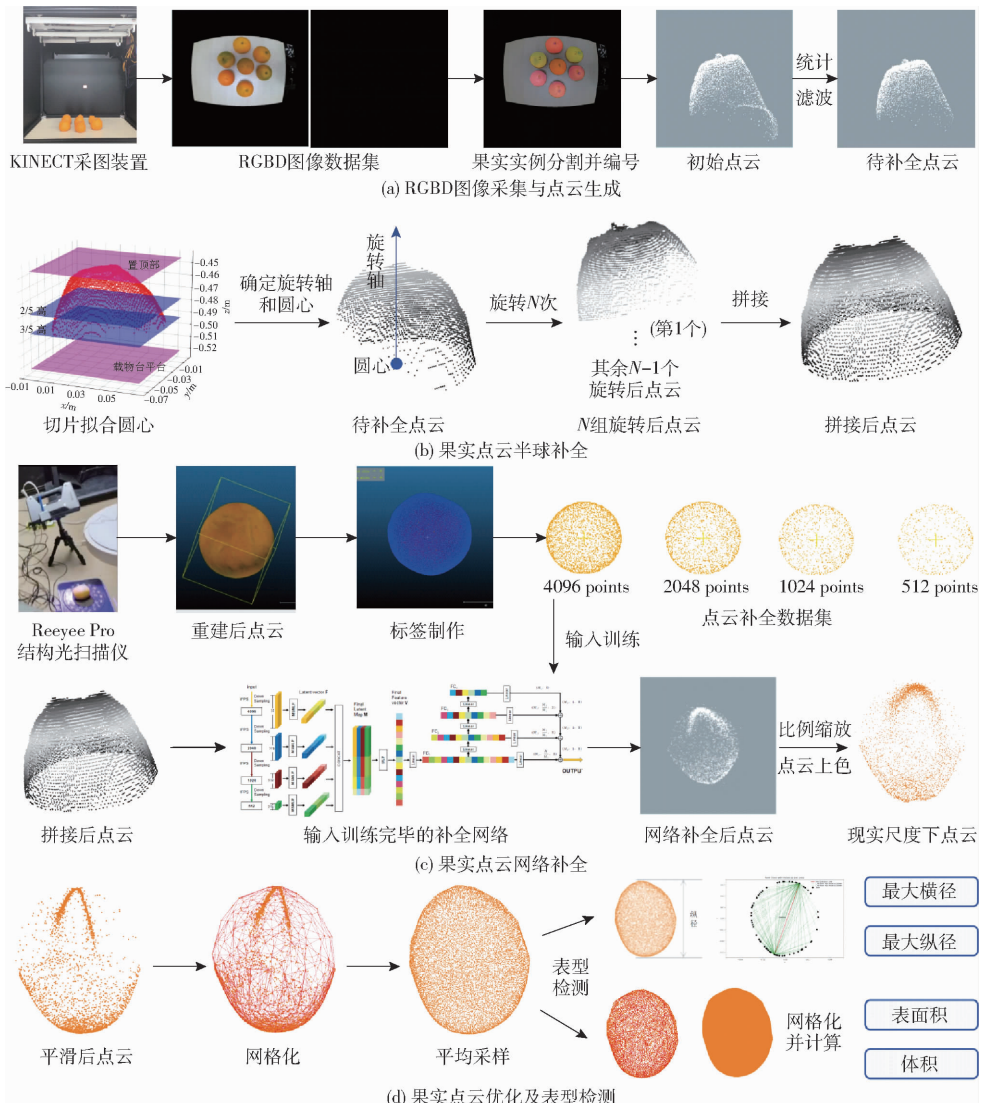


图1 技术流程图

Fig. 1 Technical flowchart

操作,最后使用本文设计的表型检测算法进行表型检测,如图1d所示。

本文算法开发和测试的平台为通用计算机(酷睿11代i9-12900K主频2.5~5.2GHz处理器、64GB内存、NVIDIA GeForce RTX 3090显卡24GB显存)、Windows 10专业版和Pycharm,以及Python 3.6。

1.2.2 RGBD图像采集与点云生成

(1)柑橘果实RGBD图像采集

设计和制作了一个恒光源暗室用于拍摄柑橘的彩色对齐图和深度图。将光源安置在深度相机后方,有效克服KINECT相机在逆光条件下深度信息丢失的问题。柑橘放置于载物台的中心,KINECT相机垂直向下安装,与柑橘保持约0.55m的最优成像距离。得到1024像素×1024像素的深度图和1024像素×1024像素的RGBD对齐图。

(2)柑橘图像分割与编号

如图1a所示,首先使用固定阈值分割方法将7个柑橘从背景中分离,并对每个柑橘进行编号。具体步骤如下:①读取RGBD对齐图像。②将RGB图像转换HSV颜色空间。③在HSV空间中,根据预设的阈值范围创建掩码;色调 H 在0~90之间,这意味着选择的色调范围是从红色开始,覆盖到黄色。饱和度 S 在50~255,这意味着选择的颜色必须有一定的纯度,即不是完全灰色。明度 V 在30~255,这意味着选择的颜色必须有一定的亮度,即不是完全黑色。④将此掩码作用于RGBD对齐图像,生成包含柑橘和噪声的RGBD对齐图像。⑤提取RGB图像中的连通域并剔除小区域噪声。⑥根据连通区域的顶点坐标从上到下从左到右的规则,为每个识别的区域分配唯一的颜色,并在图像上标记数字。⑦使用每个柑橘的连通域二值图像与RGBD对齐图像进行掩模,得到每个柑橘的RGBD分割图像。

(3)单视角点云转换及滤波

本文算法不涉及彩色信息处理,因此使用连通域二值图像与D图像进行掩模,再利用微软SDK提供的函数将深度图直接转换为点云。RGBD传感器得到的点云数据粗糙,存在离群点和悬浮点,因此进行统计滤波,滤除10%的拖影点云,得到柑橘单视角待补全点云。

1.2.3 果实点云半球补全策略

由于柑橘的尺寸、位姿差异,导致单视角RGBD图像转换的点云总是略小于完整点云的一半,一般为40%~45%。而PFNET网络要求待补全点云为固定比例,如果将单视角点云直接输入

会导致补全效果不佳。因此,将PFNET网络补全比例设置为50%,再利用柑橘果实类球体的外形特征,将单视角点云补全到50%再输入,可有效保障补全效果。果实点云半球补全方法和过程如图1b所示。单视角待补全点云中包含有果实最大横径的部分点云,因此在横向切面上搜索最大横径候选点云并进行圆拟合,得到圆心和旋转轴;计算点云缺失程度,以确定旋转次数和旋转步长角度;将待补全点云沿旋转轴旋转,并与原点云拼接,得到半球点云。

1.2.3.1 柑橘点云最大横径拟合圆心计算

图2a所示为柑橘单视角点云,用红色标识。柑橘果实的最大横径处于果实中部,因此在果实中部区域进行最大横径候选点云的搜索。首先,检测柑橘点云最顶部和载物台平面的深度信息,两者相减得到粗略的柑橘纵径;然后,在纵向上截取2/5~3/5高度区域点云,从高到低进行200次切片操作,得到100组候选点云,其中一组点云的3D可视化结果如图2b所示;最后,采用间接平差拟合圆法处理全部候选点云,输出半径最大的拟合圆心坐标,如图2c所示。

1.2.3.2 单视角柑橘点云半球补全

提出了一种依据点云缺失程度进行自适应半球补全方法,过程如下:

(1)对最大横径处点云进行圆弧拟合,首先按照点在圆上的初始角度进行初始排序,如图3a所示。

(2)按照排序计算相邻点连接圆心后的扇形夹角,取夹角最大值的相邻点为圆弧的端点,如图3b所示。将端点连接到上节所述拟合圆心形成一个扇形区域,计算扇形内角 P ,计算式为

$$P = \frac{360a}{A} \quad (1)$$

式中 a ——扇形面积

A ——整圆面积

计算旋转次数 N 和旋转步长角度 S ,计算式为

$$N = \left[\frac{360}{P} \right] \quad (2)$$

$$S = \left[\frac{360}{N+1} \right] \quad (3)$$

式中 $[]$ ——取整数运算符号

(3)以圆心空间坐标为原点,作一条平行于相机坐标系 z 轴的单位向量,以该单位向量为旋转轴。将待补全点云绕旋转轴旋转 N 次,步长角度为 S ,分别得到旋转 S 、 $2S$ 、 \dots 、 NS 角度的 N 组旋转后点云。将旋转前后点云进行拼接,得到拼接后点云,如

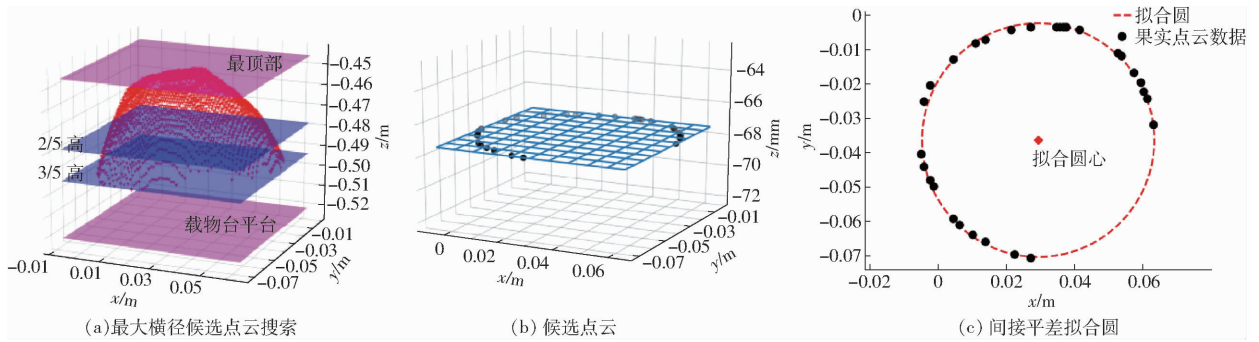


图2 柑橘最大横径点云拟合圆心计算流程图

Fig. 2 Flowcharts of citrus maximum cross-diameter point cloud fitting center calculation

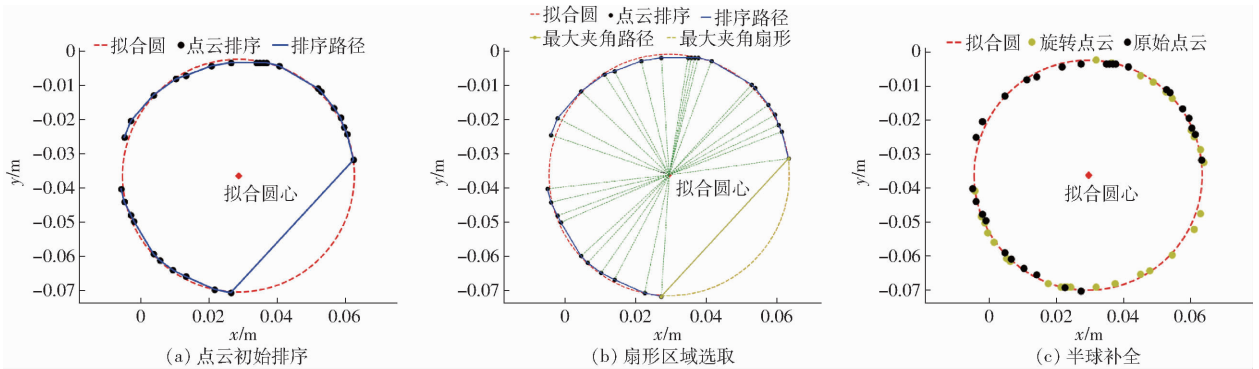


图3 柑橘半球补全示意图

Fig. 3 Schematics of citrus hemisphere completion

图3c所示。

1.2.4 果实点云网络补全

1.2.4.1 柑橘3D点云补全数据集制作

(1) 使用结构光扫描仪获取完整柑橘3D点云

从517个赣南脐橙中随机选取250个果实,依次将柑橘放置在电动旋转托盘上,采用Reeyee Pro结构光扫描仪进行扫描。设置旋转步进值为 20° 。每个柑橘获取18个视角的彩色点云,再使用扫描仪配套软件进行三维重建,得到带有噪声的完整3D点云数据,再经过点云滤波得到高质量的点云,格式为PLY。点云滤波先采用直通滤波删除背景,再使用半径基滤波进一步去除残留的悬浮噪声点。半径基滤波设置阈值为150、半径为0.01 m,此时悬浮点剔除效果最佳。滤波之后,3D点数量从100万左右减少到70万,点云质量大幅提升。

(2) 转换柑橘3D点云格式为TXT文件

提取每一个柑橘点云中的三维空间坐标信息,作为TXT格式文件的前3列,并在第4列全部增添补全类别信息0。结构光扫描仪获取的3D点云PLY格式数据全部转换为TXT格式文件,共250个。

(3) 扩充柑橘3D点云训练集

将上述250个TXT数据集按照3:1:1的比例划分为训练集、验证集和测试集。其中,150组训练集

点云扩充了16倍至2400组。点云的扩充手段包括旋转、翻转、平移以及各向异性缩放。旋转是通过随机选择3个坐标轴之一在 $[-180^\circ, 180^\circ]$ 范围内随机旋转。翻转是相对于3个平面之一进行镜像对称变换。平移是通过在3个坐标轴方向上随机平移一个范围 $[-0.3, 0.3]$ 。各向异性缩放通过对每个维度独立地从 $[0.85, 1.15]$ 范围内随机选取缩放因子来模拟在不同方向上的非均匀伸缩。对原始点云数据使用4种方法的组合,参数设置为随机。

(4) 柑橘3D点云下采样处理

PFNET采用了多尺度分析策略,旨在从宏观到微观不同层面捕捉点云数据的细微特征,需要不同尺度的点云数据作为输入,且最大支持2048个点的输入。本文改进后的PFNET网络可以接受4096个点输入。使用最远距离采样(FPS)将点云数据进行4个层次的下采样处理,从70万个点分别降至4096、2048、1024、512个点。

1.2.4.2 M_1M_2 -PFNET补全网络

(1) M_1M_2 -PFNET网络框架

对PFNET点云补全网络加以改进,设计了引入多头自注意力的特征提取模块(M_1),使得模型能够提取点间依赖关系,丰富了特征表示;增强了整体网络的尺度(M_2),使得补全网络能够进一步充分利用深度相机采集的点云数据。使用柑橘3D点云补全

数据集进行无监督网络的训练。将柑橘点云随机裁剪为两半用于训练,其中一半点云作为输入、另一半点云作为对照,得到能够对半球柑橘点云进行补全的无监督网络。整体的模型框图如图 4 所示。 $M_1 M_2$ -PFNET 无监督学习网络以自我一致性作为目标,通过变换不完整的点云数据生成多个补全版本,并使用自编码器结构重建原始点云数据。通过比较重建的点云与原始点云之间的差异,使用损失函数指导网络的训练。最终输入不完整的半球点云

数据,输出是完整的整球点云数据。编码器部分,将 4 个尺度点云分别输入多头自注意力特征提取模块,得到尺度上的 4 个特征 $F_i(D)$,接着拼接得到潜在特征 $M(D)$,最后使用 MLP 将 $M(D)$ 映射到最终特征向量 $V(D)$ 中。解码器部分,将 $V(D)$ 通过全连接层得到 4 个特征层 FC_i ,在 4 个特征层上预测 4 级点。鉴别器部分,也相应变化为计算 4 个尺度上生成点云预测值与真实值间的差异,即 2.1.1 节中补全损失 L_{com} 。

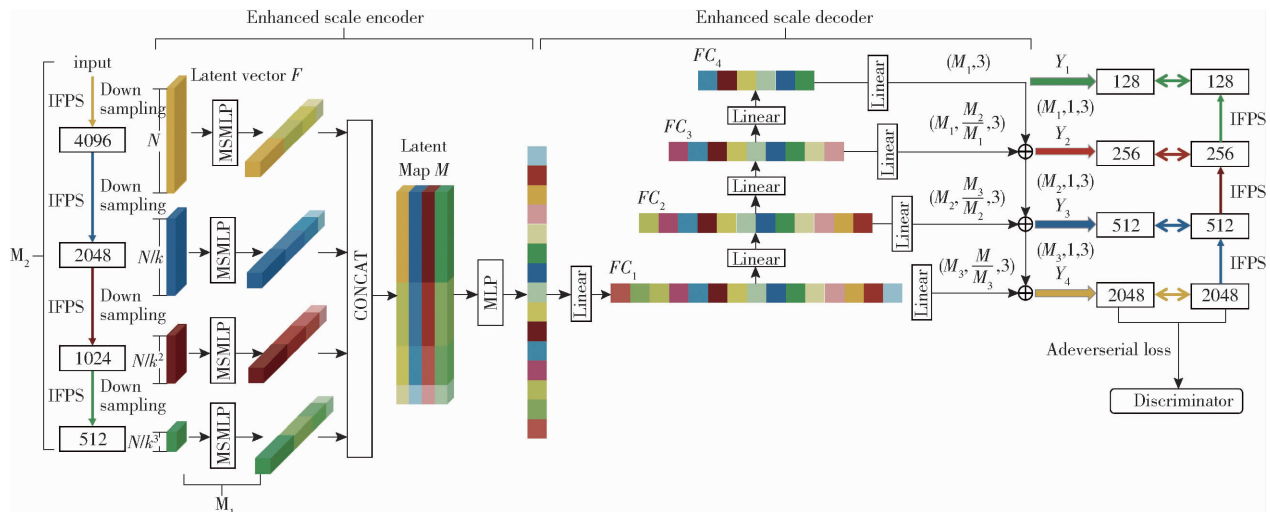


图 4 整体模型框图

Fig. 4 Overall model block diagram

(2) 引入多头自注意力的特征提取模块 (M_1)

PFNET 中的联合多层感知机 (CMLP) 虽在 4 个维度上提取了点的特征信息,但无法直接捕获点与点之间的关系。通过引入自注意力机制,模型能更好地捕捉点云中各点间的相互依赖和空间关系,从而丰富每个点的特征表示。特别是多头自注意力机制 (Multi-head self-attention, MHSA), 它进一步增强了模型对空间关系的理解,使得模型能在细节层面解析并利用点云数据的局部几何结构。这种机制通过自动学习并在不同头之间分配关注点,有效抽取了更丰富的空间特征,这对于主要依赖三维空间坐标信息的点云分析尤为关键。尽管引入自注意力机制增加了模型的复杂度,但通过精细设计的多头自注意力结构,可以充分利用现代 GPU 的矩阵运算能力,实现高效的并行计算。这不仅保证了处理大规模点云数据时的高效率,还提高了模型的灵活性和适应性。

四头自注意力机制能够根据输入数据的特性动态调整特征权重,优化聚焦于不同特征维度,以提取最相关的信息供后续处理使用。具体流程如图 5 所示。

① 输入点拼接成输入矩阵 D

$$D = \begin{bmatrix} x_1 & x_2 & \cdots & x_N \\ y_1 & y_2 & \cdots & y_N \\ z_1 & z_2 & \cdots & z_N \end{bmatrix} \quad (4)$$

其中, N 代表点云的数量,每个点包含空间 3 维坐标信息特征。

② 通过全连接层 (Fully connected layer, FC) 将输入点矩阵 D 逐步映射至 128、256、512、1 024 维,分别得到 h_{head_1} 、 h_{head_2} 、 h_{head_3} 和 h_{head_4} 等 4 个维度的计算结果。

③ 拼接 4 次自注意力的计算结果得到特征 $F(D)$ 。

$$F(D) = [h_{head_1} \quad h_{head_2} \quad h_{head_3} \quad h_{head_4}] \quad (5)$$

4 次自注意力计算依次挖掘了 128、256、512、1 024 共 4 个维度上每个点之间的关联信息,拼接 4 次自注意力计算结果,输出包含 4 个不同维度点间相关性的特征 $F(D)$ 。最终不仅提取点自身的空间坐标,还获取了点云中各点间的相互依赖和空间关系信息,进一步提升了网络的性能。

(3) 扩展 PFNET 网络尺度 (M_2)

PFNET 需要不同尺度的点云数据作为输入。KINECT 相机捕获的柑橘点云数据包含 4 500 ~

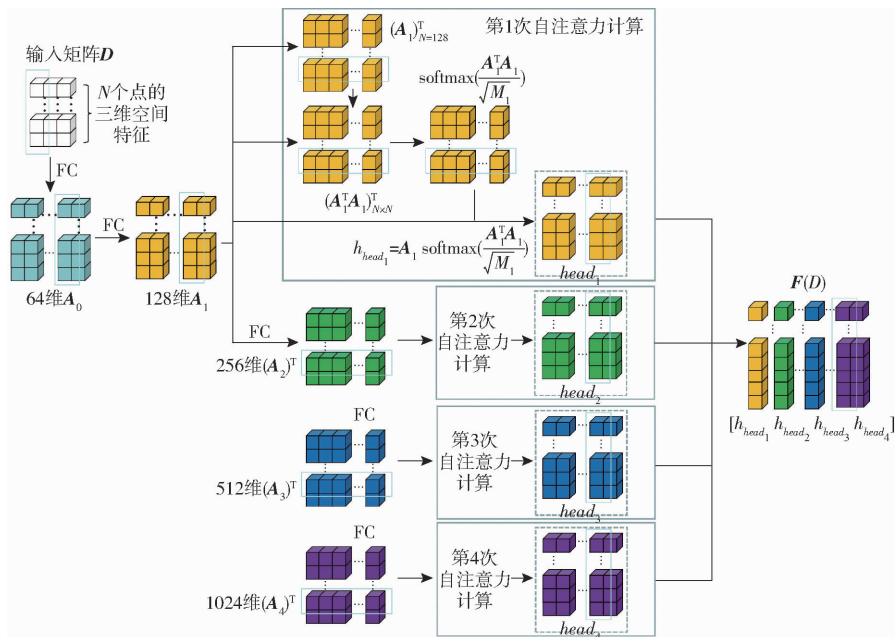


图5 多头自注意力计算过程

Fig. 5 Multi-head self-attention calculation process

7 500 个有效点,而 PFNET 仅支持 2 048 个点的输入,导致无法充分利用这些数据。因此,在原有架构中增加了第 4 尺度,即一个能够接受 4 096 个点的输入层,并对编码器和解码器结构进行了对应的调整和优化,以确保兼容性和性能提升。新增尺度增加了网络深度,有助于学习更深层次的特征,能更好地捕捉局部和全局特征,提升模型表达能力和拟合复杂模式的能力,有利于复杂形状和细节丰富的结构补全。PFNET 采用多分辨率编码器(MRE),多尺度的方法提升了高低层次点云语义信息提取的效率,能够更好地捕捉和融合不同层次的特征信息。本文增添了一个 4 096 尺度层,需对 4 个尺度的点云分别进行特征提取并拼接后映射。具体细节如图 4 中编码器部分所示。对输入的 512、1 024、2 048、4 096 共 4 个尺度待补全柑橘点云分别进行 M_1 模块特征提取流程,得到 $F_1(D)$ 、 $F_2(D)$ 、 $F_3(D)$ 和 $F_4(D)$ 4 个特征,拼接 4 个特征 $F_i(D)$,得到大小为 $1\ 920 \times 4$ 的潜在特征 $M(D)$,最后使用 MLP 将这些特征映射到最终特征向量 $V(D)$ 中。

PFNET 采用点金字塔解码器(PPD),可将全连接解码器与基于折叠的解码器相结合,以加强预测形状的局部几何信息。增强多尺度 PFNET 的编码器尺度已改为 4 层,为了进一步增强模型的生成能力,解码器也应相应扩充为 4 层结构。增强尺度的点金字塔解码器有助于生成更完整、连续且结构合理的点云,具体细节如图 6 所示。解码器将最终特征向量 $V(D)$ 作为输入,目的是输出大小为 $M \times 3$ ($M = 2\ 048$) 的缺失区域点云。具体流程如下:①将 $V(D)$ 依次通过全连接层,得到 4 个特征层 FC_i ($i =$

1, 2, 3, 4), 每个特征层负责生成不同分辨率的点云。②256 维特征层 FC_4 预测生成大小为 $M_1 \times 3$ ($M_1 = 128$) 的一级中心点 Y_1 。③512 维特征层 FC_3 基于 Y_1 预测生成大小为 M_2/M_1 的二级中心点 Y_2 。将特征层 FC_3 预测生成的大小为 $M_1 \times M_2/M_1 \times 3$ 的特征向量和一级中心点 Y_1 扩展后大小为 $M_1 \times 1 \times 3$ 的特征向量进行合并操作。④1 024 维特征层 FC_2 基于 Y_2 预测生成大小为 M_3/M_2 的三级中心点 Y_3 。具体同步步骤③中方法,如图 6 所示进行第 2 次合并操作。⑤最终进行第 3 次合并操作,得到增强尺度 PPD 的最终预测结果四级最终点 Y_4 ,即补全网络最终输出的点云信息。

1.2.5 柑橘果实点云优化与表型检测方法

(1) 柑橘果实点云优化

网络补全后的点云稀疏且不均匀,视觉效果不佳且影响果实表型测量精度,因此对点云进行优化。首先,采用均值滤波器对点云数据进行平滑处理,去除表面噪声;然后,将平滑后点云中的点连接成面,再由这些面构成整个表面的网格表示,即 mesh 化,得到果实的三维网格模型;最后,对三维网格模型进行均匀采样,得到稠密且均匀的点云。

(2) 表面积和体积测量

将最终优化点云 mesh 化,得到三角网格模型,再使用标准的几何算法来计算表面积和体积。使用海伦公式计算每个三角形的面积,将所有单个三角形的面积相加,即可得到整个 mesh 模型的表面积。使用积分法将 mesh 模型划分成子模型,通过几何计算得到体积,然后累加得到总体积。

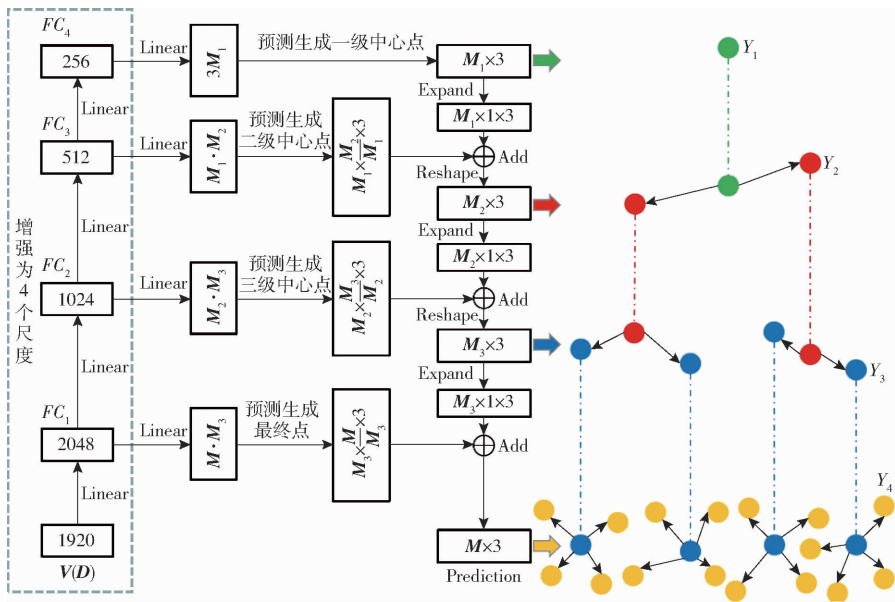


图 6 增强尺度点金字塔解码器

Fig. 6 Enhanced scale point pyramid decoder

(3) 横径和纵径测量

柑橘通常呈现球形或椭球形,如图 7 所示,采用模拟人工测量的方式,得到果实的横径、纵径。

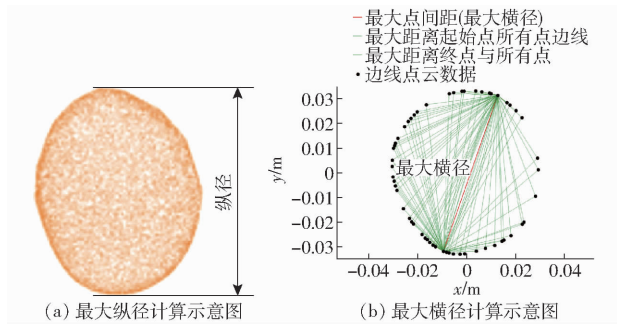


图 7 纵横径测量方法示意图

Fig. 7 Schematics of transverse and longitudinal diameter measurement method

具体步骤如下:①模仿人工进行纵径测量。人工测量是用游标卡尺卡住果蒂和果实底部度数测量。而由于柑橘果实是人工将果蒂朝上置于平台上,与人工测量时标准一致,因此计算顶部点云与底部点云的深度差值,即可直接获得果实的纵径。②模仿人工进行横径测量。人工测量时是将果蒂正对自己,使用卡尺测量果实横径,多次测量取最大值。并且绝大多数柑橘最大横径位于果实中部。基于此,模仿人手卡住柑橘果实,采用 1.2.3 节方法得到横切点云后,计算每个点之间的距离,最后取点云间距最大值作为最大横径。

2 实验结果与分析

2.1 M_1M_2 -PFNET 网络模型验证实验

2.1.1 损失函数和评价指标

倒角距离 ζ_{cd} 常用来表示缺失点云的补全效果,

描述了预测结果与真实点云之间的差异。

$$(S_1, S_2) =$$

$$\frac{1}{2} \left(\frac{1}{|S_1|} \sum_{x \in S_1} \min_{y \in S_2} \|x - y\| + \frac{1}{|S_2|} \sum_{y \in S_2} \min_{x \in S_1} \|x - y\| \right) \quad (6)$$

式中 S_1 ——预测点云 S_2 ——真实点云

由于解码器阶段变为 4 个尺度,在预测生成的 4 个阶段分别计算损失函数,计算训练过程中完整的损失函数 L_{com} 。

$$L_{com} = d_{cd_1}(Y_1, Y_{gt1}) + d_{cd_2}(Y_2, Y_{gt2}) + d_{cd_3}(Y_3, Y_{gt3}) + d_{cd_4}(Y_4, Y_{gt4}) \quad (7)$$

式中 Y_1, Y_2, Y_3, Y_4 ——4 个阶段的预测点云输出

$Y_{gt1}, Y_{gt2}, Y_{gt3}, Y_{gt4}$ ——从真实点云通过 IFPS

采样到 $N/16, N/8, N/4$

和 N 分辨率的真实完整点云

整点云

$d_{cd_1}, d_{cd_2}, d_{cd_3}, d_{cd_4}$ ——4 个阶段的倒角距离

2.1.2 M_1M_2 -PFNET 网络在柑橘果实数据集上的验证

为了评估 M_1M_2 -PFNET 对柑橘果实进行补全的能力,在 1.2.1 节中柑橘 3D 点云基准数据集上进行实验,在得到缺失柑橘果实点云后,进行一次归一化处理,与完整柑橘果实点云构成点云对。表 1 中给出不同网络点云补全的评估结果。倒角距离越小,证明补全的点云与原来的真实点云位置越接近,补全效果越好。 M_1 -PFNET 是仅改进特征提取模块的网络, M_2 -PFNET 是仅增强尺度的改进网络。对比原 PFNET,2 种改进网络均取得了更小的倒角距离,证明了改进的有效性。与其他补全网络相比,

M_1M_2 -PFNET 具有最小的倒角距离,展示了 M_1M_2 -PFNET 网络优越的性能。图 8b 展示了一组 M_1M_2 -PFNET 对缺失 50% 柑橘点云的补全效果,其表面存在少量噪声并且不均匀,这是因为原补全网络 PFNET 本身在补全时存在局部细节扩散的问题。但在经过优化之后,扩散的点云和噪声点均被滤除,最终在外形相似度和均匀度上均接近基准点云,图 8c 为优化后的效果。图 8e 是图 8c 最终优化点云与图 8d 所示的激光扫描仪获取的基准点云进行比较的重叠图。计算最终优化点云和真实完整点云间的倒角距离为 1.896,小于 M_1M_2 -PFNET 网络倒角距离,具有更佳的点云质量。

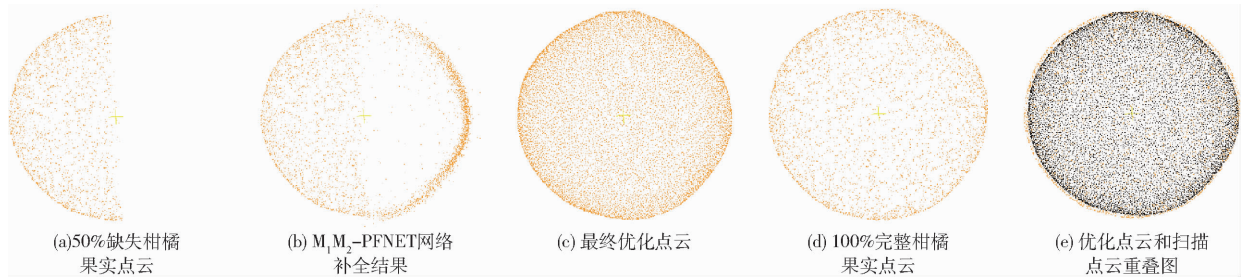


图 8 一组柑橘果实 3D 点云的补全效果对比

Fig. 8 Comparisons of completion effect of a set of 3D point clouds of citrus fruits

2.2 柑橘果实表型测量性能分析

2.2.1 人工测量值对比分析

使用游标卡尺测量柑橘的最大横径和最大纵径;使用排水法测量柑橘的体积;将柑橘皮剥离再撕成小块放入扫描仪中压平,使用 Epson Expression 12000XL 型扫描仪(爱普生公司,分辨率 0.01 mm^2)扫描得到柑橘的表面积。使用本文算法测量和分析了 668 个柑橘果实的表型,其中 517 个赣南脐橙、54 个丑橘、97 个沃柑。3 个品种柑橘的外形尺寸统计数据如表 2 所示。计算决定系数(R^2)和均方误差(MSE)来比较结果,并计算平均误差和最大误差来验证鲁棒性。使用 KINECT 相机获取了 668 个果实的单视角点云并使用本文方法进行三维重建和表型检测,统计结果如图 9 所示,4 种表型测量平均 R^2 均大于 0.96、平均测量精度均超过 93.24%。

表 2 果实尺寸统计数据

Tab.2 Fruit size statistical data mm

柑橘品种	个数	果径最小值	果径最大值	果径平均值	标准差
赣南脐橙	517	52.83	93.19	73.28	10.91
丑橘	54	67.33	91.34	81.26	6.18
沃柑	97	53.81	86.03	63.89	8.87

实验结果表明,本文方法能够有效克服果实位姿和拍摄视角导致的点云缺失问题,实现高精度的表型测量,满足实际检测的精度需求。

表 1 在柑橘点云数据集上的点云补全结果

Tab.1 Point cloud completion results on citrus point cloud dataset

补全网络	输入点云数	倒角距离
PCN	4 096	2.247
MSN	4 096	2.014
TopNet	4 096	2.148
VRCNet	4 096	2.189
ECG	4 096	2.225
PFNET	4 096	2.213
M_1 -PFNET	4 096	2.035
M_2 -PFNET	4 096	2.141
M_1M_2 -PFNET	4 096	1.973
M_1M_2 -PFNET + Optimization	4 096	1.896

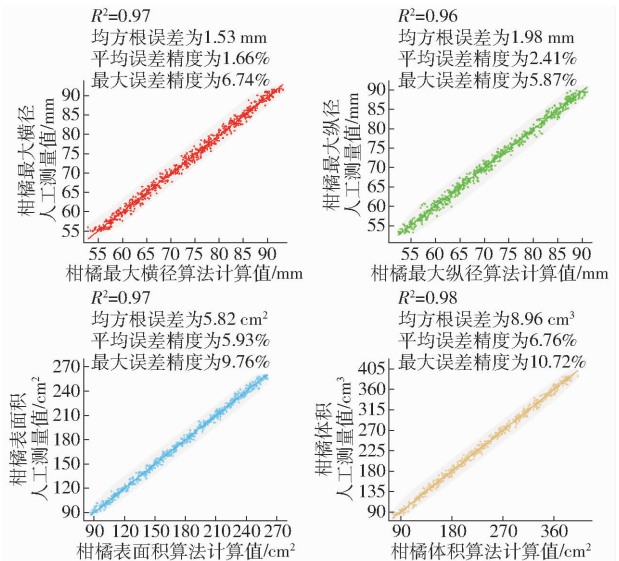


图 9 果实表型测量结果相关性分析

Fig.9 Correlation analysis after fruit phenotypic measurements

为了验证本文方法的泛化性能,分别统计了 3 个品种柑橘表型检测结果,如表 3 所示。本文使用

表 3 不同品种柑橘测量精度

Tab.3 Measurement accuracy of different citrus varieties %

柑橘品种	横径	纵径	表面积	体积
赣南脐橙	98.45	97.64	94.12	93.26
丑橘	97.81	97.06	93.82	93.03
沃柑	98.05	97.62	93.94	93.25

的训练数据集为赣南脐橙。赣南脐橙的检测精度最高,而与赣南脐橙形态更相近的沃柑品种精度次之,形态差异最大的丑橘品种检测精度最差。3 个品种柑橘均取得了很好的检测精度,测量精度差值小于 0.7%,证明了本文方法的泛化性能较好。

2.2.2 3 种典型果实表型检测方法对比

与 RGBD 图像和三维扫描仪两种表型测量方法进行了对比实验。RGBD 图像法是在 RGB 机器视觉基础上结合了 D 图像的深度信息进行的测量,无需提前标定,只需要一幅图像即可完成测量,对于横径、纵径的测量速度快但无法测量表面积和体积。三维扫描仪法在对柑橘果进行完整的三维重建后,再使用本文的表型计算方法进行检测。激光扫描仪

的数据采集速度和点云融合速度都较为费时,人工劳动强度大。本文方法是在 RGBD 方法的基础上,对单视角点云进行补全,重建出完整的柑橘果实点云后进行表型检测。表 4 展示了 3 种测量方法的结果。本文方法与 RGBD 图像法相比,单果检测时间增加 17.97 s,但横纵径检测精度大幅提高,且能一次测量 4 项表型参数。与三维扫描仪方法相比,检测精度差值在 4 个百分点以内,但速度超过 48 倍,硬件成本只有后者的 1/10,通量大于 180 个/h,且易于实现自动化。本文方法在检测精度、运行速度、硬件成本和自动化程度上具有较好的平衡,是一种低成本、综合性能高的三维重建技术,有广泛应用于果实表型无损测量的潜力。

表 4 3 种典型果实表型检测方法性能对比

Tab. 4 Performance comparison of three typical fruit phenotype detection methods

测量方法	测量精度/%				平均耗时/s	硬件成本/元	自动化程度
	横径	纵径	表面积	体积			
RGBD 图像方法	81.56	96.53			1.56	2×10^4	全自动
三维扫描仪方法	99.17	98.96	96.98	97.12	981.60	2.5×10^5	需人工
本文方法	98.34	97.59	94.07	93.24	19.53	2×10^4	全自动

2.3 讨论

影响 M_1M_2 -PFNET 网络性能的因素主要包括:本文使用的数据集为赣南脐橙,如果对于特定品种制作数据集可以有效提升补全性能。PFNET 网络经过改进之后能接受 4 096 个点的输入规模,与 KINECT 相机获取的单视角点云规模匹配。如果输入点云规模进一步提高,网络复杂度和计算资源需求大幅提升而补全性能提升有限。PFNET 具有保持原有几何结构的同时进行补全的特性,但要求输入点云为固定的比例。本文方法利用类球体性质对单视角点云进行了半球补全以满足 PFNET 的需求,当果实外形偏离类球体时重建精度会下降。

在果实表型检测中,误差主要来源有:由于 KINECT 相机自身硬件性能的限制,获取的深度数据总是存在误差,在当前系统的拍摄距离下存在 1 mm 的系统误差,影响横径、纵径的测量精度;部分柑橘果实的形状不规则,果蒂偏离了中轴区域 2~3 mm,对纵径测量结果影响较大;补全网络自身性能限制,导致补全的半球点云不完全与真实值一

致,影响表面积、体积的测量精度。

3 结论

(1)为了对类球体果实单视角点云进行补全,对 PFNET 进行了改进,增添了多头自注意力特征提取模块用于提取点间关系,而且增强了尺度使得网络具有更丰富的特征层次,使得传统 PFNET 可以适应 4 096 个点的输入规模,并能够捕捉更多的细节信息,补全的点云经过优化之后倒角距离达到 1.896,实现了高质量的单视角三维重建。

(2)提出了一种基于 RGBD 单视角图像的果实表型检测方法,在精度、通量、成本和自动化等方面实现了较好的平衡,对于 3 个品种柑橘的果实横径、纵径、表面积、体积的测量准确率均大于 93.03%,通量大于 180 个/h、硬件成本低于 2 万元。相比三维扫描仪方法,表型检测精度略低但成本、通量具有巨大优势,且易于实现自动化;相比传统 RGBD 图像方法,通量略差但精度大幅提升,且可以准确测量表面积和体积。

参 考 文 献

- [1] BELO A P M, MORGADO C M A, DE SOUZA E R B, et al. Comparative and organic analysis and characterization of varieties of tangerines[J]. *Scientia Horticulturae*, 2018, 240: 102 - 108.
- [2] 李沛, 何治霖, 谈月霞, 等. 基于重测序数据与表型性状的宽皮柑橘遗传多样性分析与优异种质筛选[J]. *中国农业科学*, 2024, 57(23): 4761 - 4795.
LI Pei, HE Zhilin, TAN Yuexia, et al. Genetic diversity analysis and excellent germplasm screening of wide skinned citrus based on resequencing data and phenotypic traits[J]. *Scientia Agricultura Sinica*, 2024, 57(23): 4761 - 4795. (in Chinese)

- [3] ZHANG Ying, CAO Yufen, HUO Hongliang, et al. An assessment of the genetic diversity of pear (*Pyrus L.*) germplasm resources based on the fruit phenotypic traits[J]. *Journal of Integrative Agriculture*, 2022, 21(8): 2275–2290.
- [4] FORERO M G, MURCIA H F, MENDEZ D, et al. LiDAR platform for acquisition of 3D plant phenotyping database[J]. *Plants*, 2022, 11(17): 2199.
- [5] TENG X, ZHOU G, WU Y, et al. Three-dimensional reconstruction method of rapeseed plants in the whole growth period using RGB–D camera[J]. *Sensors*, 2021, 21(14): 4628.
- [6] SUN D, ROBBINS K, MORALES N, et al. Advances in optical phenotyping of cereal crops[J]. *Trends in Plant Science*, 2022, 27(2): 191–208.
- [7] 凡超, 吴秋萍, 张国辉, 等. 基于三维点云数据的芒果果形鉴定平台搭建及试验[J]. *农业工程学报*, 2024, 40(22): 290–296.
FAN Chao, WU Qiuping, ZHANG Guohui, et al. Construction and experiment of mango fruit shape identification platform based on three-dimensional point cloud data[J]. *Transactions of the CSAE*, 2024, 40(22): 290–296. (in Chinese)
- [8] 张煜恒, 周宏平, 张超. 基于激光点云的苹果果实定位与体积测量方法[J]. *应用激光*, 2023, 43(8): 102–111.
ZHANG Yuheng, ZHOU Hongping, ZHANG Chao. Apple fruit localization and volume measurement method based on laser point cloud[J]. *Applied Laser*, 2023, 43(8): 102–111. (in Chinese)
- [9] SAHA K K, TSOULIAS N, WELTZIEN C, et al. Estimation of vegetative growth in strawberry plants using mobile LiDAR laser scanner[J]. *Horticulturae*, 2022, 8(2): 90.
- [10] 徐胜勇, 李磊, 童辉, 等. 基于 RGB–D 相机的黄瓜苗 3D 表型高通量测量系统研究[J]. *农业机械学报*, 2023, 54(7): 204–213, 281.
XU Shengyong, LI Lei, TONG Hui, et al. Research on high-throughput measurement system for 3D phenotype of cucumber seedlings based on RGB–D camera[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(7): 204–213, 281. (in Chinese)
- [11] LIU J, XU X, LIU Y, et al. Quantitative potato tuber phenotyping by 3D imaging[J]. *Biosystems Engineering*, 2021, 210: 48–59.
- [12] YAMAMOTO S, KARKEE M, KOBAYASHI Y, et al. 3D reconstruction of apple fruits using consumer-grade RGB-depth sensor[J]. *Engineering in Agriculture Environment*, 2018, 11(4): 159–168.
- [13] WANG Y, CHEN Y. Fruit morphological measurement based on three-dimensional reconstruction[J]. *Agronomy*, 2020, 10(4): 455.
- [14] 麦春艳, 郑立华, 孙红, 等. 基于 RGB–D 相机的果树三维重构与果实识别定位[J]. *农业机械学报*, 2015, 46(增刊): 35–40.
MAI Chunyan, ZHENG Lihua, SUN Hong, et al. Three dimensional reconstruction and fruit recognition and localization of fruit trees based on RGB–D camera[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2015, 46(Supp.): 35–40. (in Chinese)
- [15] WANG Z, WALSH K B, VERMA B, et al. On-tree mango fruit size estimation using RGB–D images[J]. *Sensors*, 2017, 17(12): 2738.
- [16] 陈龙, 王浩云, 季呈明, 等. 基于 Pointnet 和迁移学习的苹果表型参数估算研究[J]. *南京农业大学学报*, 2021, 44(6): 1209–1216.
CHEN Long, WANG Haoyun, JI Chengming, et al. Research on apple phenotypic parameter estimation based on Pointnet and transfer learning[J]. *Journal of Nanjing Agricultural University*, 2021, 44(6): 1209–1216. (in Chinese)
- [17] YUAN W, KHOT T, HELD D, et al. PCN: point completion network [C] // 2018 International Conference on 3D Vision (3DV), 2018.
- [18] TCHAPMI L P, KOSARAJU V, REZATOFIGHI H, et al. Topnet: structural point cloud decoder [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019.
- [19] PAN L, CEN X, CAI Z, et al. ECG: edge-aware point cloud completion with graph convolution[J]. *IEEE Robotics*, 2020, 5(3): 4392–4398.
- [20] LIU M, SHENG L, YANG S, et al. Morphing and sampling network for dense point cloud completion [C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2020.
- [21] PAN L, CHEN X, CAI Z, et al. Variational relational point completion network [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021.
- [22] HUANG Z, YU Y, XU J, et al. Pf-net: point fractal network for 3D point cloud completion [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020.
- [23] 陈建文, 赵丽丽, 任蓝草, 等. 深度学习点云质量增强方法综述[J]. *中国图象图形学报*, 2023, 28(11): 3295–3319.
CHEN Jianwen, ZHAO Lili, REN Lancuo, et al. A review of deep learning point cloud quality enhancement methods[J]. *Chinese Journal of Image and Graphics*, 2023, 28(11): 3295–3319. (in Chinese)