

李 颀, 郭梦媛. 基于深度学习的休眠期苹果树点云语义分割[J]. 江苏农业学报, 2023, 39(5): 1189-1198.  
doi: 10.3969/j.issn.1000-4440.2023.05.011

## 基于深度学习的休眠期苹果树点云语义分割

李 颀<sup>1</sup>, 郭梦媛<sup>2</sup>

(1. 陕西科技大学电子信息与人工智能学院, 陕西 西安 710021; 2. 陕西科技大学电气与控制工程学院, 陕西 西安 710021)

**摘要:** 针对苹果树结构复杂、树干之间相互遮挡、导致国内外大规模机械设备自动剪枝误剪率高等问题, 提出 1 种基于深度学习的休眠期苹果树点云的语义分割。以陕西省礼泉苹果种植基地的休眠期苹果树为研究对象, 为了解决双视角点云配准之间非重叠点对距离过大导致配准误差大的问题, 用 Kinect V2 传感器获取休眠期苹果树点云, 对每株果树采用改进迭代最近点算法 (Iterative closest point, ICP) 进行多视角三维重建, 对于大规模的果树点云, 构建基于超点图的果树分割网络 (Super point graphs network, SPGNet), 对果树点云进行语义分割, 保留果树点云的复杂几何信息。结果表明, 当果树双视角点云的配准误差小于 1 mm 时, 可成功分割休眠期苹果树的树干与分枝, 对分类精度、预测值与真实值的交并比 (*IoU*) 进行评估, 其中树干的分类精度、*IoU* 分别为 94.0%、0.85, 分枝的分类精度、*IoU* 分别为 83.1%、0.75。由此可见, 研究结果可解决机械设备自动剪枝误剪率高的问题, 能在自然光线条件、大规模休眠期苹果树场景下实现对休眠期苹果树树干与分枝的分割, 为大规模自动剪枝提供依据。

**关键词:** 多视角果树三维重建; 超点图; 深度学习; 果树点云语义分割

中图分类号: S661.1 文献标识码: A 文章编号: 1000-4440(2023)05-1189-10

## Semantic segmentation of apple tree point cloud in dormant period based on deep learning

LI Qi<sup>1</sup>, GUO Meng-yuan<sup>2</sup>

(1. School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an 710021, China; 2. School of Electrical and Control Engineering, Shaanxi University of Science and Technology, Xi'an 710021, China)

**Abstract:** Aiming at the problems of complex structure of apple trees, mutual occlusion between the trunks, and false pruning of large-scale mechanical equipment at home and abroad, a semantic segmentation of apple tree point cloud in dormant period based on deep learning was proposed. Taking the dormant apple trees in Liqian apple planting base in Shaanxi province as the research object, in order to solve the problem of large registration error caused by the large distance between non-overlapping point pairs in dual-view point cloud registration, the Kinect V2 sensor was used to obtain the point cloud of dormant apple trees, and the improved iterative closest point algorithm (ICP) was used for multi-view three-dimensional reconstruction of each fruit tree. For large-scale fruit tree point cloud data, a fruit tree segmentation network based on super point graphs (SPGNet) was constructed to perform semantic segmentation on fruit tree point cloud, and the

收稿日期: 2022-09-08

基金项目: 陕西省农业科技创新工程项目 [201806117YF05NC13 (1)]; 陕西省科技厅农业科技攻关项目 (2015NY028); 陕西科技大学博士科研启动基金项目 (BJ13-15)

作者简介: 李 颀 (1973-), 女, 陕西西安人, 博士, 教授, 主要研究方向为农业智能化、信息化、深度学习。(E-mail) liqidq@sust.edu.cn

通讯作者: 郭梦媛, (E-mail) 940516893@qq.com

complex geometric information of fruit tree point cloud was retained. The results showed that when the registration error of the dual-view point cloud of the fruit tree was less than 1 mm, the trunk and branches of the apple tree in the dormant period could be successfully segmented. The classification accuracy and intersection over union of predicted and true values (*IoU*) were evaluated. The classification accuracy and *IoU* of the trunk were 94.0% and 0.85, re-

spectively, and the classification accuracy and *IoU* of the branches were 83.1% and 0.75, respectively. In a word, the research results could solve the problem of high mis-cutting rate in the process of automatic pruning of mechanical equipment, and could realize the segmentation of trunks and branches of dormant apple trees under natural light conditions and large-scale dormant apple tree scenes, and could provide a basis for large-scale automatic pruning.

**Key words:** multi-view fruit tree three-dimensional reconstruction; super-point graph; deep learning; fruit tree point cloud semantic segmentation

2021年,陕西省苹果树的种植面积为 $6.21 \times 10^5$   $\text{hm}^2$ ,陕西省苹果产量为 $1.24 \times 10^7$  t,占全国苹果总产量的26.9%。对休眠期苹果树进行剪枝是提高果树产量的重要措施,传统的苹果树剪枝严重依赖人工操作,不仅成本高昂,而且浪费人力、效率低下<sup>[1]</sup>。虽然目前国内外大规模苹果树自动剪枝的机械设备得到了广泛研究<sup>[2]</sup>,但是机械化剪枝设备错误、遗漏的修剪会导致苹果产量下降。因此,提高休眠期苹果树树干与分枝的分类精度,可以降低机械化设备自动剪枝的误剪率,提高苹果产量并减少劳动力需求<sup>[3]</sup>。

常用于果树枝条语义分割的数据包括二维图像、三维点云等。Amatya等<sup>[4]</sup>使用贝叶斯分类法,将带有叶片的荔枝图像分成4个类别,其中树干识别率达到89.2%。韩文康等<sup>[5]</sup>用Atrus空间金字塔池、门特征金字塔网络来增强网络的特征提取能力,结果显示,苹果和树枝语义分割模型的精确率和召回率分别为87.6%和77.2%。Majeed等<sup>[6]</sup>使用Kinect V2传感器获取苹果幼树的彩色(RGB)和深度图像,使用SegNet对简单的前景二维图像进行分割。

由于果树的二维图像中缺少深度信息,导致树干与分枝的分割精度低。Elfiky等<sup>[7]</sup>提出基于几何特征从两侧对休眠的纺锤形苹果树进行三维重建,

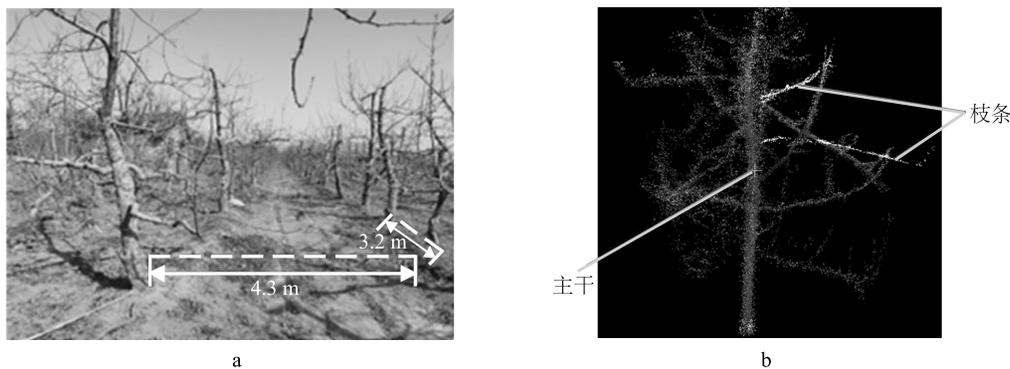
结果显示,基于自适应建模方案的主枝修剪点的检测精度达到96.0%。Medeiros等<sup>[8]</sup>通过分裂合并聚类算法,将休眠苹果树点云分为树干、连接点和分枝,结果显示,平均分割精度为98%,平均直径误差为0.6 cm,该分割使用机器视觉算法手动提取点云特征,由于光照度的变化和背景的复杂,果树枝条分割的鲁棒性通常较差,并且激光雷达的成本较高。

本研究以休眠期苹果树为试验对象,通过改进迭代最近点算法对果树进行多视角三维重建,并将结果输入基于超点图的深度学习网络(Super point graphs network, SPGNet),以期使果树的复杂几何信息得到实质性保留,在大规模休眠期苹果树的场景下实现对休眠期苹果树树干与分枝的分割。

## 1 材料与方法

### 1.1 试验场地

本研究的试验场地位于陕西省咸阳市礼泉苹果种植基地,基地内果树的行距、株距分别为4.3 m、3.2 m,具体场景见图1a。在休眠期,苹果树形成低矮的纺锤形(下文将纺锤形结构的矮化休眠苹果树简称为苹果树),种植基地内的矮化苹果树冠幅为2.1 m,主蔓长度最长可达1.2 m,平均高度约为2.4 m,去除背景后的纺锤形苹果树如图1b所示。



a: 苹果树种植基地场景;b: 去除背景后的纺锤形苹果树。

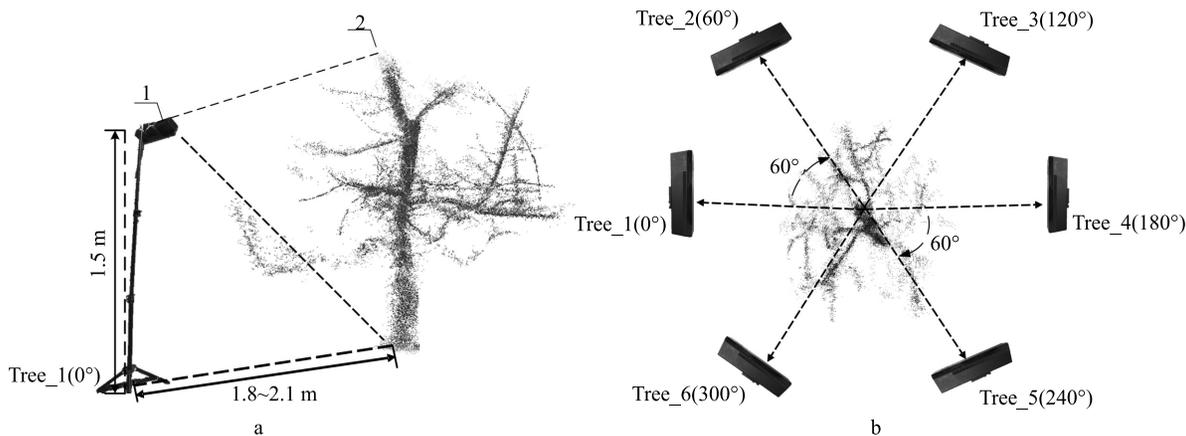
图1 休眠期苹果树点云语义分割的试验场

Fig.1 Test site for semantic segmentation of apple tree point cloud during dormant period

## 1.2 数据采集与点云预处理

1.2.1 数据采集 本研究于2022年2-3月在陕西省咸阳市礼泉苹果种植基地采集数据。由于高精度点云需要高性能的传感器和稳定的成像环境,而以Kinect为代表的基于光子飞行时间(Time-of-Flight, ToF)的RGB-D传感器在点云精度、成本和处理速度3个方面更加平衡,因此本研究使用Kinect V2传感器获取点云。由于果树平均高度约为2.4 m, Kinect V2安装在距离地面150 cm处的支架上,构成数据采集设备,该位置配合Kinect V2的视场角,可将果树分枝完全覆盖, Kinect V2的部分参数见表1。数据采集设备距离苹果树主茎中心1.8~2.1 m, Kinect V2沿顺时针方向绕果树旋转,每旋转60°采集1次数据,旋转360°后数据采集完毕。

图2a描述了获取单株果树三维数据的方法,将其中第1视角(0°)的点云命名为Tree\_1,将其他5个视角的点云分别命名为Tree\_2、Tree\_3、Tree\_4、Tree\_5、Tree\_6,详见图2b。数据采集系统由个人计算机开发和测试,硬件配置为AMD R7-4800H SSD、NVIDIA RTX2060 4G,软件环境选择OpenCV 3.4.13、PCL1.8.1。



a:数据采集方法(1表示数据采集设备,2表示休眠期苹果树);b:多视角数据采集俯视效果(Tree\_1~Tree\_6表示0°~300°视角的拍摄结果)。

图2 休眠期苹果树点云的采集

Fig.2 Data collection of point cloud of apple tree during dormancy period

由于自然光碰到果树枝干边缘时会发生折射现象, Kinect V2传感器无法正常接收信号而产生浮点,从而影响三维重建精度。在本研究中,果树的浮点被统计滤波器(Radius-outlier-removal)消除。当休眠期苹果树的点云集合 $P=\{C_1, C_2, C_3, \dots, C_x, \dots, C_i\}$  ( $x \in N, i \in N$ )时,查询点 $C_i$ 的坐标为 $(x_i, y_i, z_i)$ ,邻域点 $C_n$ 的坐标

表1 Kinect v2传感器相关参数

Table 1 Related parameters of Kinect v2 sensor

特征	参数
测距原理	光子飞行时间(ToF)
RGB相机分辨率(pixel)	1 920×1 080
深度相机分辨率(pixel)	512×424
帧率	30 FPS
有效距离(m)	0.5~7.5
水平视场(°)	70
垂直视场(°)	60
质量(kg)	1.35
外形尺寸(mm)	200×50×50
设备接口	USB 3.0

FPS:每秒传输帧数。

1.2.2 点云预处理 由于数据采集设备获取的果树点云也包括附近的树木和地面,因此根据果树点云在空间坐标系的 $x, y, z$ 坐标,结合树的宽度、高度,设置 $x$ 的阈值范围为 $(-1.4 \text{ m}, 0.9 \text{ m})$ ,  $y$ 的阈值范围为 $(3.1 \text{ m}, 1.0 \text{ m})$ ,可有效过滤点云中的背景。数据采集设备放置在距果树1.8~2.1 m的位置,可以通过将 $z$ 阈值设置为 $(0.9 \text{ m}, -1.3 \text{ m})$ 去除点云中的地面信息。

为 $(x_n, y_n, z_n)$ ,距查询点 $C_i$ 最近的 $K$ 个点 $C_n$ 称为 $K$ 邻域。利用公式(1)、公式(2)求解查询点 $C_i$ 与各邻域点 $C_n$ 的距离( $d_p$ )与均值( $\mu$ )。其中 $\mu$ 服从高斯分布,由公式(3)可以求解出标准差( $\sigma$ )。

$$d_p = \sqrt{(x_i - x_n)^2 + (y_i - y_n)^2 + (z_i - z_n)^2} \quad (1)$$

$$\mu = \frac{\sum_{i=1}^N x_i}{N} \quad (2)$$

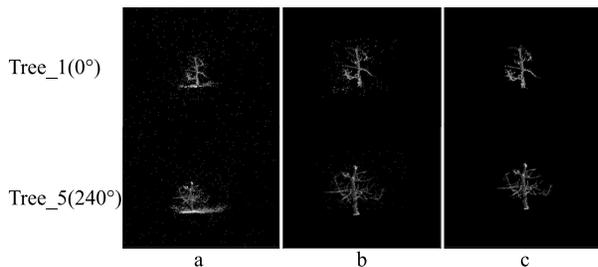
$$\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}} \quad (3)$$

通过公式(4)定义浮点标准距离的范围( $L$ ):

$$L = \mu \pm \sigma \times s \quad (4)$$

式中, $s$  为标准差倍数。

如果某点  $L < \mu$ , 则去除该点。在不丢失果树几何结构信息的情况下, 分别将  $K, S$  设置为 10、2 以去除浮点的点云, 分别精选受光线影响最小的 Tree\_1 和受光线影响最大的 Tree\_5 视角, 如图 3a~图 3c 所示。



a: 多视角采集的苹果树原始点云图片; b: 滤除背景中附近树木和地面后的点云图片; c: 滤除浮点后的点云图片。Tree\_1、Tree\_5 分别表示在 0°、240° 视角采集到的休眠期苹果树点云。

图3 休眠期苹果树的点云预处理结果

Fig.3 Preprocessing of point cloud of apple tree during dormancy period

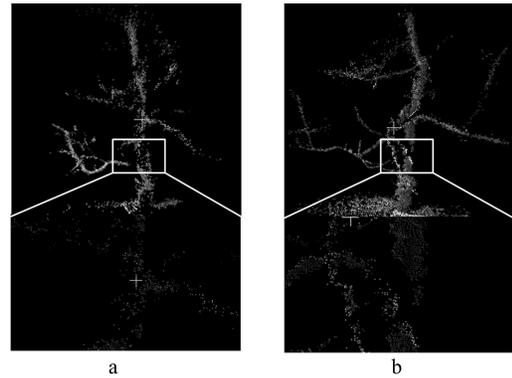
### 1.3 休眠期苹果树的三维重建

由于苹果树枝干相互遮挡, 因此果树点云存在如图 4 所示的破损现象(树干或分枝的点云残缺不全), 导致果树的三维模型无法从单一视角获取。因此, 本研究提出采用改进迭代最近点算法, 给定果树不同角度的  $M$  帧点云  $\{P_i\}_i^M (P_i = \{p_{ij}\}_{j=1}^{N_i}) (M=6)$ 、初始配准参数  $\{R_i^0, t_i^0\}_{i=1}^M$ , 将不同视角的点云转换到同一点云基准帧的坐标系下, 实现休眠期苹果树的三维重建。

传统的迭代最近点算法(Iterative closest point, ICP)是双视角点云配准的基础<sup>[9]</sup>。当给定待匹配点云  $P = \{p_i\}_{i=1}^{N_p}$  和果树精确模型  $Q = \{q_j\}_{j=1}^{N_q}$  时, 双视角点云的配准问题可表示为公式(5)。

$$\begin{cases} \arg \min_{R, t, q_{c(i)} \in Q} \sum_{i=1}^{N_p} \|R p_i + t - q_{c(i)}\|_2^2 \\ \text{s.t. } (R)^T R = I_3, \det(R) = 1 \end{cases} \quad (5)$$

式中,  $p_i, q_{c(i)}$  是待匹配点云、精确模型, 具有如



a: Tree\_1(0°)点云出现破损; b: Tree\_2(60°)点云出现破损。

图4 苹果树点云的破损现象

Fig.4 Damage phenomenon of apple tree point cloud

公式(6)中对应关系的点对:

$$C_k(i) = \arg \min_{j \in \{1, 2, \dots, N_p\}} \|R_{k-1} p_i + t_{k-1} - q_j\|_2 \lim_{x \rightarrow \infty} i = 1, 2, \dots, N_p \quad (6)$$

根据公式(6)的点对关系, 公式(7)不断更新果树点云之间的配准参数( $R, t$ ), 直到前后 2 次迭代所求的配准参数无明显差异时, 即可输出双视角点云配准的最优结果。

$$(R_k, t_k) = \arg \min_{R, t} \sum_{i=1}^{N_p} \|R p_i + t - q_{c_k(i)}\|_2^2 \quad (7)$$

由于本研究中采集果树点云的相邻视角为 60°, 而双视角果树点云之间非重叠点对的间隔距离过大会影响配准参数的计算, 从而降低果树三维重建的精度。因此, 本研究引入如公式(8)所示的指数函数  $w_{ij}$ , 以减少果树点云中非重叠点对对配准结果的影响。

$$\begin{cases} w_{ij} = \alpha_{ij} \exp\left(-\frac{d_{ij}^2}{2\sigma^2}\right) \\ \text{s.t. } \begin{cases} \alpha_{ij} = 1, q_{c(ij)} \in P_{ij} \\ \alpha_{ij} = 0.5, q_{c(ij)} \in P \end{cases}, \sigma = \frac{2}{N} \sum_{j=1}^{N_i} d_{ij} \end{cases} \quad (8)$$

式中,  $d_{ij}^2$  为重叠区域内点对  $(p_{ij}, q_{c(ij)})$  之间的欧氏距离,  $\sigma$  为标准差,  $\alpha_{ij}$  为二值变量。当  $q_{c(ij)}$  属于重叠点对时, 取 1.0, 否则取 0.5。

综上所述, 本研究的果树三维重建问题可表示为公式(9):

$$\begin{cases} \arg \min_{R, t, q_{c(ij)} \in Q} \sum_{i=2}^M \sum_{j=1}^{N_i} (w_{ij} \|R_i p_{ij} + t_i - q_{c(ij)}\|_2^2) \\ \text{s.t. } (R_i)^T R_i = I_3, \det(R_i) = 1 \end{cases} \quad (9)$$

式中,  $q_{c(ij)}$  为点  $p_{ij}$  在休眠期苹果树精确模型  $Q$  中的对应点,  $w_{ij}$  为与点对  $[p_{ij}, q_{c(ij)}]$  相关的系数。

根据给定的初始配准参数  $\{R_i^0, t_i^0\}_i^M = 1$ , 不断求解公式(9)的配准参数, 对相邻视角的果树点云进行配准, 不断更新果树的粗糙模型, 将果树三维重建的问题分解为多个双视角果树点云的配准问题。

在本研究中, 每株果树的三维重建都需要采集6个视角的点云, 根据改进的双视角配准算法, 将粗糙模型  $Q_i$  不断修正为果树的精确模型, 具体有4个步骤。

步骤1: 对于前3个连续视角的数据, 将 Tree\_1 定义为基准帧, Tree\_2 根据给定的初始配准参数

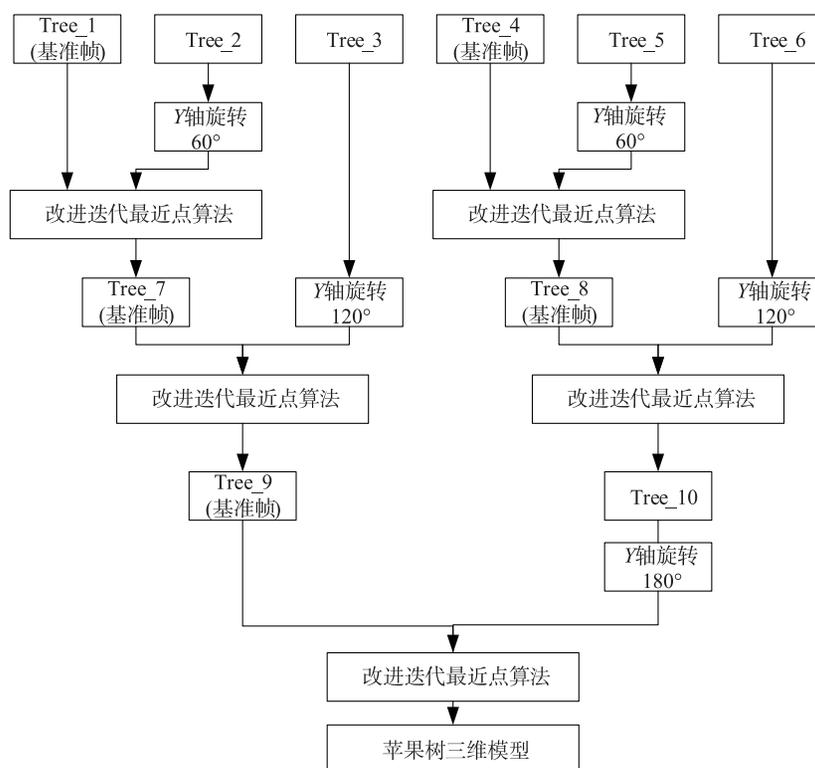
$\{R_i^0, t_i^0\}_i^M = 1$ , 构造出 Tree\_7 粗糙模型, 如公式(10)所示:

$$Q_i = \{R_i^0 P_{ij} + t_i^0\}_{i=1, j=1}^{M, N_j} \quad (10)$$

步骤2: Tree\_7 替换 Tree\_1 成为新的基准帧, 以便进一步计算出 Tree\_8 粗糙模型。

步骤3: 对于后3个连续的视角, 以同样方法构造出 Tree\_10 粗糙模型。

步骤4: Tree\_8 成为新基准帧, 与 Tree\_10 利用双视角配准方法获得果树精确的三维模型, 果树三维重建的流程如图5所示。



Tree\_1~Tree\_10 表示不同视角的休眠期苹果树点云。

图5 休眠期苹果树多视角三维重建算法示意

Fig.5 Schematic diagram of multi-view three-dimensional reconstruction algorithm of apple trees during dormant period

#### 1.4 休眠期苹果树的语义分割

本研究通过改进迭代最近点算法对休眠期苹果树的点云进行三维重建, 提高了果树三维模型的精度。在大规模休眠期苹果树场景下对果树点云进行语义分割, 执行分割任务时不仅可以克服果树点云规模大、无序且不规则的问题, 而且可以保留果树点云的复杂几何信息。

二维图片的语义分割通常使用卷积神经网络(Convolutional neural network, CNN)<sup>[10]</sup>, 但 CNN 需

要输入规则的数据, 因此 CNN 需要先对三维模型进行规则化处理, 常见的处理方法是将三维模型体素化。三维模型体素化的缺点是 CNN 的输入数据会变得稀疏且影响点云的旋转、平移。超点图具有丰富的边缘特征, 并且可编码三维点云中物体各部分之间的上下文关系<sup>[11]</sup>。因此, 本研究拟利用超点图(Super point graph, SPG)对休眠期苹果树进行语义分割。由于受到 PointNet<sup>[12]</sup>、门控图神经网络(Graph neural networks, GNNs)<sup>[13]</sup>的影响, 本研究

拟构建基于超点图的深度学习网络 (Super point graphs network, SPGNet), 实现在大规模休眠期苹果树场景下的语义分割, 休眠期苹果树树干与分枝的分割流程如图 6 所示。

SPGNet 输入如图 7a 所示的休眠期苹果树点云后, 将大规模苹果树三维模型的语义分割问题分成几何分割、超点嵌入和上下文分割 3 个不同规模的问题, 如图 7b、图 7c 所示。

1.4.1 基于语义分割的超点图构建 矮化的苹果树结构呈纺锤形, 在大规模苹果树场景下的语义分割目标不是检索如汽车、椅子等简单对象, 而是为了将休眠期苹果树分解为简单的几何部分。SPGNet 将如图 8a 所示的果树三维模型作为输入, 该三维模型遵循全局能量模型且被分割成如图 8b 所示的简单并具有意义的几何块, 又称为超点。最后, 由超点、边构成超边, 使超点之间的空间邻接关系通过超边连接, 如图 8c 所示。

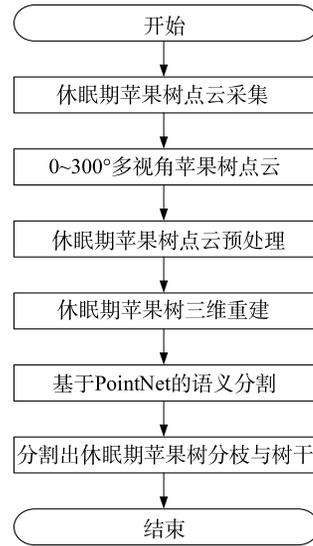
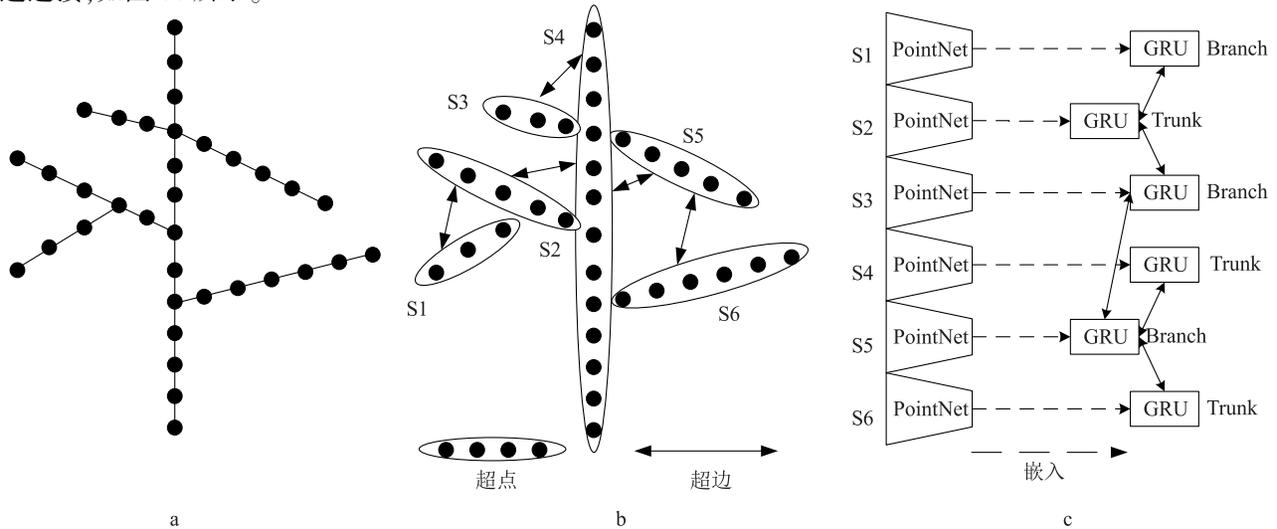
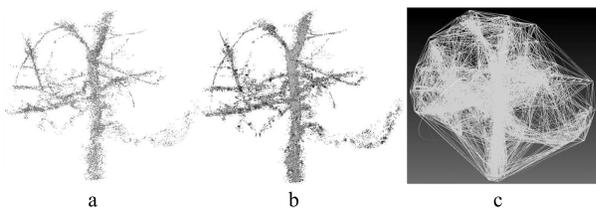


图 6 分割休眠期苹果树树干与分枝的流程  
Fig.6 Process of dividing the trunk and branches of a dormant apple tree



a: 输入的休眠期苹果树三维模型; b: 基于语义分割的超点图 (S1~S6 表示划分出的超点 1~6); c: 超点嵌入和上下文分割 (GRU 表示门控循环单元, Branch 表示休眠期苹果树的分枝, Trunk 表示休眠期苹果树的树干)。

图 7 基于 SPGNet 的休眠期苹果树语义分割算法示意  
Fig.7 Semantic segmentation algorithm of apple tree in dormant period based on SPGNet



a: 果树的三维模型; b: 超点; c: 超点图。

图 8 超点图构建的可视化过程

Fig.8 Visualization of super point graph construction

SPGNet 的全局能量模型是相对于果树点云的近邻点邻接图  $G_n = (V, E_n)$  ( $n = 10$ ) 的定义。  $V$  代表三维模型内的点云的集合,  $E_n$  ( $n = 10$ ) 表示距离三维模型内每个点最近的 10 个点的平均距离。因此, 果树三维模型的语义分割问题可定义为如公式 (11) 所示的最小分割问题。

$$\arg \min_{g \in \mathbb{R}^d} \sum_{i \in Q} \|g_i - f_i\|^2 + \mu \sum_{(i, j) \in E_n} w_{i, j} [g_i - g_j \neq 0] \quad (11)$$

式中:  $\mu$  为正则化系数, 表示几何分割结果的粗

糙度; $w_{i,j}$ 为三维点云之间的边权重; $[?]$ 为艾佛森括号; $f$ 为三维模型的几何特征; $g_i$ 为待求的解,相同的 $g$ 值可视为1个超点。

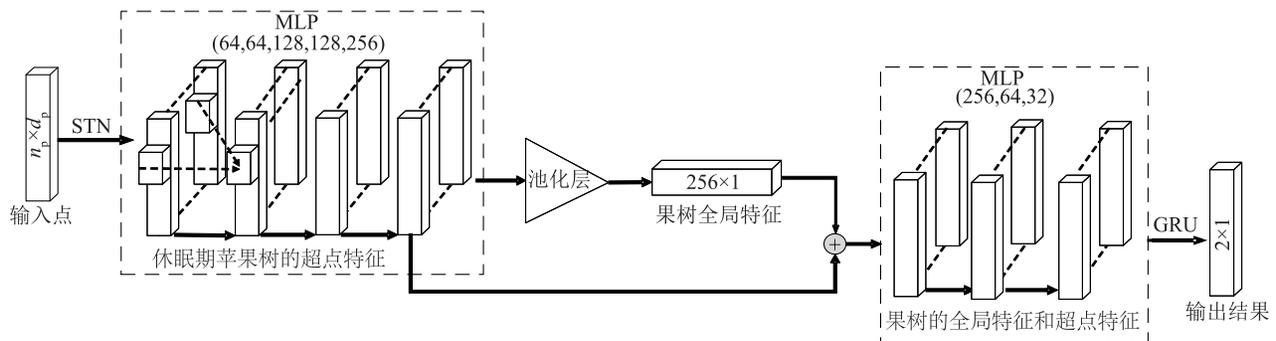
根据公式(11)的结果 $S=\{S_1,\dots,S_k\}$ ,将果树语义分割的结果定义为 $k$ 个超点。本研究提出采用有向属性图 $G=(S,E,F)$ ( $G$ 表示果树的超点图; $S$ 表示超点的集合; $E$ 表示连接超点的所有边的集合,又称超边; $F$ 表示编码的超点图边界的权重)来表示整个果树的三维模型。如果2个相邻的超点( $S,T$ )在 $E$ 中至少有1条边,则该超边的表达见公式(12)。(S,T)从超边的偏移量中获得的空间特征如公式(13)所示。

$$E=\{(S,T)\in S^2\mid\exists(i,j)\in E\cap(S\times T)\} \quad (12)$$

$$\delta(S,T)=\{(p_i-p_j)\mid(i,j)\in E\cap(S\times T)\} \quad (13)$$

1.4.2 超点嵌入与上下文分割 果树点云的超点嵌入是将果树三维模型的超点( $S_i$ )嵌入到维度大小为 $d_p$ 的特征向量中,计算承载特定语义信息的特征

描述符;上下文分割的目的是根据边缘特征传输的超边信息细化嵌入的原始超点,获取每个点的语义标签,即苹果树的2个类(树干和分枝)。SPGNet的输入维度为 $s\times n_p\times d_p$ ( $s$ 代表果树三维模型被分割的超点数, $n_p$ 表示超点包含的三维点数, $d_p$ 代表每个采样点的原始特征维度)。SPGNet首先通过空间变换网络(Spatial transform networks, STN)将每个超点对齐并重新缩放为独立的单位球体以作为SPGNet的输入,保证三维模型的平移、旋转不变性;使用权重共享多层感知器(Multilayer perceptron, MLP)提取超点特征,然后使用最大池化层聚集该特征。连接全局特征和超点特征后,MLP再次应用于超点嵌入的输出。最后,为了实现果树三维模型的语义分割,将MLP生成的32维特征嵌入到门控循环单元(Gated recurrent unit, GRU)<sup>[14]</sup>,得到的休眠期苹果树点云的分割结果见图9。



MLP表示多层感知器,STN表示空间变换网络,GRU表示门控循环单元, $n_p$ 表示超点包含的三维点数, $d_p$ 表示每个采样点的原始特征维度。

图9 SPGNet网络结构

Fig.9 SPGNet network structure

1.4.3 模型训练和超参数 本研究对休眠期苹果树的语义分割网络进行了端到端的训练,使用120株果树三维模型进行训练和测试,每株果树有31 832~100 521个点,随机选择84株果树模型(70%)进行训练,其余36株果树模型(30%)用于测试。同时,通过Cloud Compare v2.11.3软件手动将果树的点云标记为树干(Trunk)、树枝(Branch)2类,分别用灰色、浅黑色标注。为了提高计算性能,使用0.006的体素对果树点云进行二次采样。本研究使用的SPGNet的超参数见表2。

SPGNet试验在显卡为NVIDIA GeForce RTX2080 Ti、操作系统为Ubuntu 18.04、运行内存为128 G、算力为7.5的服务器上运行,使用的软件工具包括CUDA 10.0、CUDNN 9.1和Python 3.7,试验在Pytorch 1.10

中实现。

表2 SPGNet的超参数

Table 2 Hyperparameters of SPGNet

训练参数	参数值
BASE LEARNING RATE	0.001
LR_DECAY	0.7
LR_STEPS	[275, 320]
EPOCHS	200
BATCH_SIZE	4
MOMENTUM	0.9
OPTIM	ADAM
RE_STRENGTH	0.005
VOXEL_WIDTH	0.006
PTN_NPTS	[40, 128]

ADAM:自适应矩估计优化器。

1.4.4 性能评价指标 为了评估休眠期果树三维点云配准的准确性,根据公式(14)计算平均距离误差(MDE):

$$MDE = \frac{\sum_{i=1}^n \sqrt{(x_{si} - x_{ti})^2 + (y_{si} - y_{ti})^2 + (z_{si} - z_{ti})^2}}{n} \quad (14)$$

式中, $n$ 为点云的数量; $(x_{si}, y_{si}, z_{si})$ 为源点云的坐标; $(x_{ti}, y_{ti}, z_{ti})$ 为目标点云的坐标。

此外,本研究使用整体精度(OA)、类精度(ClassAccuracy)、预测值与真实值的交并比(IoU)、平均精度(mAcc)和平均联合交集度(mIoU)来评估SPGNet在测试数据集中评估分割果树树干和树枝的性能,计算方法见公式(15)~公式(19)。

$$OA = \frac{\sum_{i=1}^N p_{ii}}{\sum_{j=1}^N \sum_{k=1}^N p_{jk}} \times 100\% \quad (15)$$

$$ClassAccuracy = \frac{TP}{(TP+FN)} \times 100\% \quad (16)$$

$$IoU = \frac{TP}{(TP+FP+FN)} \quad (17)$$

$$mIoU = \frac{\sum_{i=1}^n IoU}{N} \times 100\% \quad (18)$$

$$mAcc = \frac{\sum_{i=1}^n ClassAccuracy}{N} \times 100\% \quad (19)$$

式中: $p_{jk}$ 表示类 $k$ 预测为类 $j$ 的点云总数; $TP$ 、 $FP$ 、 $FN$ 分别表示正确分割的果树分枝类别(真阳性)、错误分割的果树分枝类别(假阳性)、漏检果树分枝类别(假阴性); $N$ 表示类别数( $N=2$ )。

## 2 结果与分析

### 2.1 果树的双视角配准误差分析

为了选择双视角果树点云的配准视角,本研究随机选择自然光下10株结构不同的果树,针对2种不同的双视角配准方案进行配准误差分析。图10显示了以 $0^\circ$ 视角点云为源点云,利用公式(14)对果树源点云与目标点云的配准误差分析结果,即果树点云在视角差分别为 $60^\circ$ 、 $180^\circ$ 时前后2帧点云平均距离误差(MDE)的对比结果。相比于视角差为 $60^\circ$ 的休眠期苹果树点云配准结果,视角差为 $180^\circ$ 时存在更多的非重叠点对。在进行双视角配准时, $0\sim 180^\circ$ 的MDE最大为3.80 mm,而 $0\sim 60^\circ$ 的MDE最大为0.28 mm,因此本研究中双视角配准的视角差选择 $60^\circ$ 。

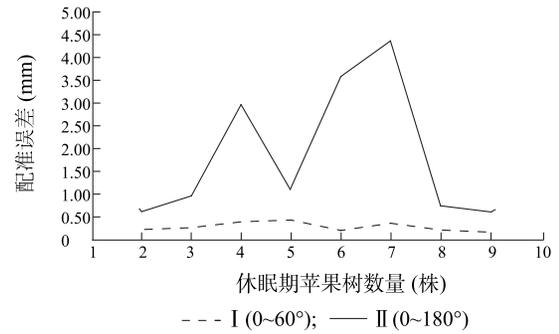


图10 不同双视角配准方案误差的比较

Fig. 10 Error comparison of different dual-view registration schemes

### 2.2 果树点云分割结果的分析

本研究采用SPGNet模型,在休眠期苹果树点云的训练集与测试集内进行性能测试。首先评估模型在训练集上是否出现过拟合现象。图11a、图11b分别显示了果树点云在训练过程中的学习准确率、最小批量损失曲线。本研究发现,模型的最大学习准确率为97%,最小批量损失为0.07,进行第200次训练时的学习准确率与最小批量损失趋于稳定。因此可见,本研究模型在果树训练集上得到了适当训练。

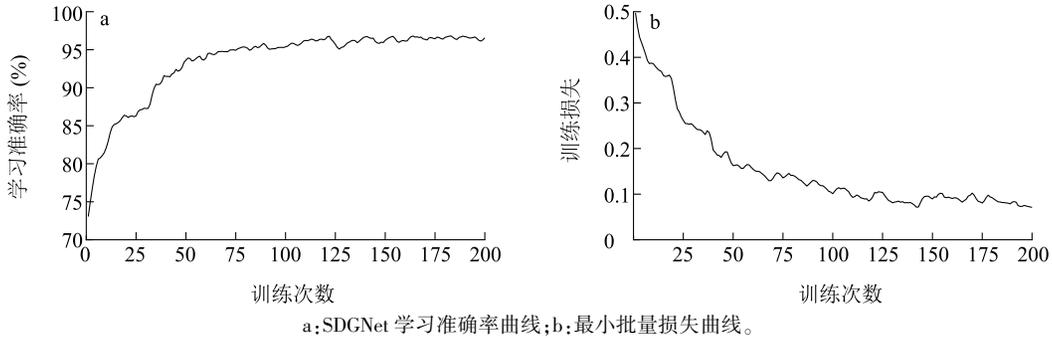
此外,本研究使用OA、mAcc、mIoU等指标,在由36株果树三维模型构成的测试集上评估本模型的分割结果。如图12a所示,SPGNet在测试集上的OA、mAcc、mIoU分别为90.1%、89%、81%。如图12b所示,果树树干的分类精度、IoU分别为94.0%、0.85;果树分枝的分类精度、IoU分别为83.1%、0.75,虽然果树分枝的分割精度低于果树树干,但是该结果获得了休眠期果树的预期分割性能。

### 2.3 本研究模型与目前先进点云分割网络性能的对比

为了分析本研究模型点云分割的性能,本研究将所提方法与目前先进的点云分割网络Point CNN<sup>[15]</sup>、DGCNN<sup>[16]</sup>在休眠期苹果树数据集上进行比较试验。为了保证比较试验的可靠性,2种比较模型均按照本研究模型的训练方式进行训练。由表3可知,本研究模型的mAcc、OA均排在首位,分别比Point CNN的mAcc、OA高1.9个百分点、1.9个百分点。经分析,Point CNN面对输入顺序不同的点云,定义了1个 $x$ 变换矩阵,保证了点云输入顺序不影响点云输出结果,从而利于Point CNN模型保留苹果树完整的点云特征并能够自适应地学习特征的权重,相比之下,DGCNN使用基于动态图的方法来学

习点云之间的关系,但对点云的局部和全局结构不够敏感,因此 Point CNN 的特征学习方法略优于

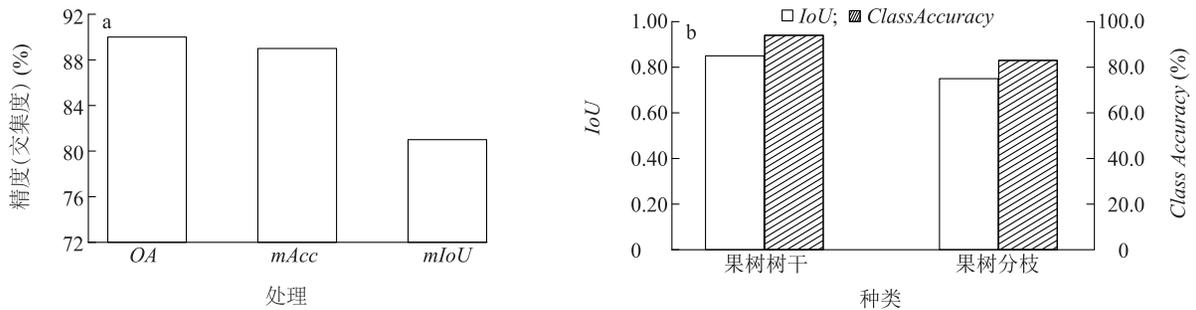
DGCNN。



a:SDGNet 学习准确率曲线;b:最小批量损失曲线。

图 11 SPGNet 训练模型

Fig.11 SPGNet training model analysis



a:SPGNet 整体测试性能分析(OA 表示整体精度,mAcc 表示平均精度,mIoU 表示平均联合交集度);b:树干与分枝分割性能分析(IoU 表示预测值与真实值的交并比,ClassAccuracy 表示树干类与分支类的精度)。

图 12 果树语义分割网络的性能

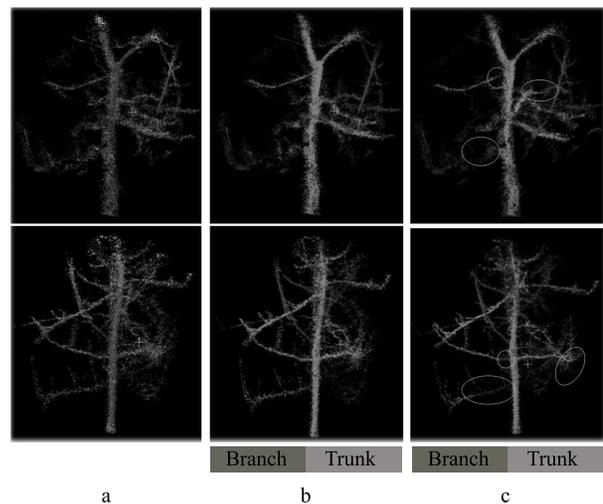
Fig.12 Performance analysis of fruit tree semantic segmentation network

本研究得出的休眠期苹果树树干与分枝的分割结果如图 13c 所示。由于存在人工标注错误及分枝形状类似于树干的情况,因此在树干和分枝之间的连接处会出现漏检、树干被错误分割为分枝的现象。虽然对果树的复杂结构存在一些分割错误,但是最终的果树分枝分割准确率为 83%,并识别出果树分枝的破损。在大规模休眠期苹果树场景下,将来可通过提取苹果树的骨架点、使深度网络在复杂的果树点云中学习更多特征等方法进一步提高果树树干与分枝的分割精度。

表 3 本研究模型 (SPGNet) 与目前先进点云分割网络性能的对比  
Table 3 Performance comparison between this research model (SPGNet) and the current state-of-the-art point cloud segmentation network

模型方法	mAcc (%)	OA (%)	类精度 (%)	
			Trunk	Branch
DGCNN	81.8	87.4	83.9	79.8
Point CNN	87.1	88.2	92.1	82.0
SPGNet	89.0	90.1	94.0	83.1

mAcc 表示平均精度,OA 表示整体精度,Trunk 表示树干类,Branch 表示分支类。



a:果树的原始点云;b:标记的分类结果;c:SPGNet 分割结果[浅黑色和灰色分别代表树枝(Branch)和树干(Trunk),圆圈表示错误案例]。

图 13 果树语义分割结果

Fig.13 Analysis of fruit tree semantic segmentation results

### 3 结论

本研究针对大规模机械设备自动剪枝时误剪率高、苹果树相互遮挡导致双视角点云配准误差大的问题,采用视角差为 $60^\circ$ 的果树点云,使用改进最近迭代点算法进行多视角三维重建,配准误差为 $0.28\text{ mm}$ ,可有效保留果树的真实特征。本研究提出的基于超点图的深度学习网络,实现了 $OA$ 为 $90.1\%$ 、 $mAcc$ 为 $89\%$ 、 $mIoU$ 为 $81\%$ 的果树树干与分枝的分割,其中树干的分类精度最高为 $94.0\%$ 。与目前先进点云语义分割网络性能对比发现,本研究模型的整体精度比Point CNN高 $1.9$ 个百分点,比DGCNN高 $2.7$ 个百分点,实现了自然光线下大规模休眠期苹果树场景的果树树干与分枝的分割。当休眠期苹果树的分类精度较高时,可降低现有机械设备自动剪枝时的误剪率。为了进一步推动机械设备在大规模休眠期苹果树场景下自动剪枝效率的提高,下一步可侧重于研究休眠期苹果树分枝的长度,并根据分枝长度匹配自动剪枝的最佳方案。

#### 参考文献:

- [1] KARKEE M, ADHIKARI B, AMATYA S, et al. Identification of pruning branches in tall spindle apple trees for automated pruning [J]. Computers & Electronics in Agriculture, 2014, 103: 127-135.
- [2] MIKA A, BULER Z, TREDER W. Mechanical pruning of apple trees as an alternative to manual pruning [J]. Ogronictwo, 2016, 15(1): 113-121.
- [3] HE L, SCHUPP J. Sensing and automation in pruning of apple trees: a review [J]. Agronomy, 2018, 8(10): 211.
- [4] AMATYA S, KARKEE M, GONGAL A, et al. Detection of cherry tree branches with full foliage in planar architecture for automated sweet-cherry harvesting [J]. Biosystems Engineering, 2016, 146(4): 3-15.
- [5] HAN W K, CHEN C. Fruit detection and segmentation for apple harvesting using visual sensor in orchards [J]. Sensors, 2019, 19(20): 4599.
- [6] MAJEED Y, ZHANG J, ZHANG X, et al. Deep learning based segmentation for automated training of apple trees on trellis wires [J]. Computers and Electronics in Agriculture, 2020, 170: 105277.
- [7] ELFIKY N M, AKBAR S A, SUN J, et al. Automation of dormant pruning in specialty crop production: an adaptive framework for automatic reconstruction and modeling of apple trees [C]. Piscataway, NJ: IEEE Computer Society, 2015: 65-73.
- [8] MEDEIROS H, KIM D, SUN J, et al. Modeling dormant fruit trees for agricultural automation [J]. Journal of Field Robotics, 2017, 34(7): 1203-1224.
- [9] 徐思雨, 祝继华, 田智强, 等. 逐步求精的多视角点云配准方法 [J]. 自动化学报, 2019, 45(8): 1486-1494.
- [10] 张蕊, 孟晓曼, 曾志远, 等. 图卷积神经网络在点云语义分割中的研究综述 [J]. 计算机工程与应用, 2022, 58(24): 29-46.
- [11] LANDRIEU L, SIMONOVSKY M. Large-scale point cloud semantic segmentation with superpoint graphs [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4558-4567.
- [12] GUINARD S, LANDRIEU L. Weakly supervised segmentation-aided classification of urban scenes from 3d LiDAR point clouds [C] // International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2017: 151-157.
- [13] LI Y, TARLOW D, BROCKSCHMIDT M, et al. Gated graph sequence neural networks [C]. Ithaca, NY: Openreview. net, 2016.
- [14] CHO K, VAN MERRIËNBOER B, BOUGARES F, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation [C]. Stroudsburg: ACL, 2014.
- [15] LI Y, BU R, SUN M, et al. Pointcnn: convolution on x-transformed points [C]. Montreal Canada: Neural Information Processing System. Foundation, 2018.
- [16] WANG Y, SUN Y B, LIU Z W, et al. Dynamic graph CNN for learning on point clouds [J]. ACM Transactions on Graphics, 2018, 38(5): 1-12.

(责任编辑:徐艳)