

基于改进 PointNet++ 模型的苗圃树木点云分类与分割

徐婕, 刘慧*, 沈跃, 杨官学, 周昊, 王思远

江苏大学电气信息工程学院, 江苏 镇江 212013

摘要 激光点云技术可用于苗圃树木生长状态监测与管理,为农业植保机器人提供有效的靶标信息。为了进一步提高树种分类和树冠、树干内部分割的精准性,提出一种基于改进 PointNet++ 的激光点云苗圃树木分类与分割方法。首先,调整 PointNet++ 深度网络邻居点云的相对特征值,同时融合三维点云的低维和高维特征,充分利用各层级点云的特征。然后,将坐标注意力模块与注意力池化融合,进一步增强局部特征提取的能力,提高分类和分割的准确性。最后,针对苗圃常见树木自制了包含 7 类苗圃景观树木点云的数据集并用于实验。实验结果表明,提出的树种识别方法总体精度可达 92.50%,平均类别精度为 94.22%;提出的树冠、树干分割方法的平均交并比为 89.09%。所提方法在分类和分割性能方面均明显优于经典的 PointNet 和 PointNet++,能够为苗圃树木检测识别和农业机器人作业提供更精确的信息。

关键词 遥感; 激光雷达; 深度学习; 树种分类; 苗圃树木树冠和树干分割; PointNet++

中图分类号 TN249

文献标志码 A

DOI: 10.3788/CJL230989

1 引言

近年来,城市绿化工程已经逐渐成为了与城市发展密切相关的热点话题之一^[1]。苗圃用于培育城市绿化工程所需树木,通常都会培育多种不同类型的树木以满足各种城市绿化需求。利用先进的科学技术对苗圃中不同类别的树木进行有效的管理和监测,能够及时发现状态不佳的树木并进行及时有效的养护和处理等操作。其中,树冠和树干是树木的重要组成部分。根据冠层的相关信息,农业机器人可以实现对靶变量的喷雾作业,达到降低成本、减少多余农药对环境造成的影响的目的^[2]。而获取到的树干信息则可以帮助农业机器人实现自主导航等作业^[3]。

随着计算机科学技术的不断发展,与深度学习领域相关的神经网络模型也日渐成熟。目前,运用较广泛的神经网络模型可以分为两种:基于图像和基于点云的模型。其中,激光点云由一组包含三维(3D)坐标信息(x, y, z)的点组成,不受光照条件的影响,并保留了物体的原始几何信息^[4-5],因此更适合用于苗圃、果园等户外场景的检测任务。但是,点云的无序特性造成了不能直接使用基于图像的神经网络模型提取三维点云的特征的问题。为了解决这一问题,早期的研究人员尝试将点云中的点转换为多个二维(2D)图像^[6-7]或体素^[8-9],再用 2D 或 3D 卷积神经网络进行特征提取操作。虽然上述方法可以解决点云中点的无序特性等

问题,但由于额外的数据处理步骤,通常需要更多的内存空间和计算成本。2017 年出现了利用多层感知直接从点云中提取有效特征的神经网络模型 PointNet 以及 PointNet++^[10-11]。PointNet 利用最大池化函数以及设计的联合对准网络,解决了点云的无序特性^[10]。在 PointNet 的基础上,PointNet++ 添加局部区域特征提取模块,通过分层的局部结构捕获越来越高维的点云局部特征^[11]。相比于其他方法,PointNet 以及 PointNet++ 不需要额外的存储空间以及数据转换过程的计算量,因此更加高效。上述两个模型目前被广泛应用于工程领域中^[12-14]。

近年来,国内外很多研究者都对树种分类以及树冠、树干分割等问题进行了相关研究。陈向宇等^[15]结合树木的结构特征、纹理特征以及冠形特征,利用支持向量机实现了树种分类。但该类方法对输入特征的要求较高,若特征选择不当则会导致性能下降。Zhang 等^[16]设置阈值截取由无人机激光雷达和地面激光得到的树干点云,通过绘制不同高度树干的点云截面投影并计算中心坐标来实现树木检测。但该方法主要针对的是林区内的树木。相比之下,苗圃中种植的景观树较矮,部分景观树的树干会被树冠点云遮挡,因此该方法并不适用于苗圃场景中的树木检测。卢华清等^[17]首先构建点云协方差矩阵,根据对应的特征值以及特征向量对树干枝叶进行粗分离处理,再用 Dijkstra 算法计算最低点到其余点的最短路径提取出枝干骨架。该方

收稿日期: 2023-07-04; 修回日期: 2023-08-19; 录用日期: 2023-09-05; 网络首发日期: 2023-09-15

基金项目: 国家自然科学基金(32171908)

通信作者: *amity@ujs.edu.cn

法主要是针对树木的枝叶点云与树干点云进行分割,对采集的数据要求较高,需要先将多个角度采集到的点云进行配准处理,再进行分类、分割作业。但在苗圃场景的实际作业过程中,由于遮挡问题,农业机器人往往只能扫描到树木的单侧点云。为了保证工作的实时性,农业机器人需要根据扫描到的部分点云及时执行树种分类任务并确定树冠、树干等信息,然后再根据获取的信息执行后续的导航和喷药等操作。因此该方法并不适用于农业机器人在苗圃场景下的实时操作。

本文针对农业机器人在苗圃实际应用场景中的需求,对 PointNet++ 进行了改进,并应用其分类和分割分支实现、提高对苗圃中景观树的树种分类和各部分分割任务的识别精度。改进后的模型使用了更深的网络结构以提取更高维度的特征,并充分利用各层级点的特征提高分类和分割的准确性。此外,对划分出的较大的局部区域,根据与中心点的距离调整邻居点的相对特征分布。同时将坐标注意力机制与注意力池化相结合,进一步提高模型提取局部特征信息的能力。采集苗圃中树木的单侧点云并处理制作数据集执行分类和分割实验。结果表明,改进后的模型具有更高的分类和分割精度,为后续在硬件平台的应用提供了基础。

2 方法原理

2.1 特征融合策略

PointNet++ 神经网络模型基于层级框架,可以有效捕获点云的局部特征以完成分类和分割任务。通过降采样算法对点云进行采样,每一层点的数量不断减少,但特征维度却不断增加,因此每个点包含的信息也更丰富。然而,PointNet++ 只使用最后一层集合抽象(SA)层中采样点对应的最高维度的特征,而忽略了点在其他 SA 层对应的低维特征。点云的高维特征更能体现点云的全局信息,而低维特征则可以更好地表征局部特征,体现更多细节信息。因此,充分利用各个 SA 层中点的特征信息有利于提高网络的性能。

本文使用特征融合策略,将点云的低维以及高维特征对应拼接,以提高模型的精度,具体内容如图 1 所示。

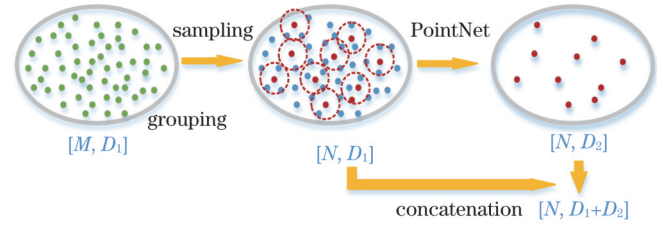


图 1 融合特征策略示意图

Fig. 1 Schematic diagram of feature fusion strategy

图 1 中,绿色、红色和蓝色的点分别代表原有点云中的点、采样点以及采样后剩余的点, M 和 N 分别表示点云中点的个数和采样点的个数。采样点的特征通过 PointNet 层由原来的 D_1 维上升到 D_2 维。本文使用的特征融合策略则是将采样点的特征与点对应之前的特征相结合,从而将特征扩展到 $(D_1 + D_2)$ 维。

2.2 局部特征分布调整策略

在局部特征提取的过程中,主要是根据邻居点与中心点的相对特征提取重要信息。距离中心点越近,则邻居点与中心点的关系越紧密,相对特征也越重要,更能反映出重要的信息。尤其对于划分出的较大的局部区域,距离中心点较远的邻居点可能会影响局部特征信息。因此,本文采用的特征分布调整方法主要是根据邻居点到中心点的距离来调整对应局部区域内的相对特征值。

假设半径为 d 的局部区域的中心点用 p_i 表示,该局部区域共包含 K 个邻居点 $[p_i^1, p_i^2, \dots, p_i^K]$,如图 2 所示。为了获取点的局部特征,PointNet++ 首先将对应的第 j 个邻居点 p_i^j 的坐标转换为以中心点为原点的相对坐标,再使用多层感知机(MLP)提取点云在局部区域内的相对特征 F_i^j :

$$F_i^j = \text{MLP}(p_i^j - p_i). \quad (1)$$

与 PointNet++ 不同,改进后的模型使用特征分布调整(FDA)模块先将邻居点局部特征分布转换为

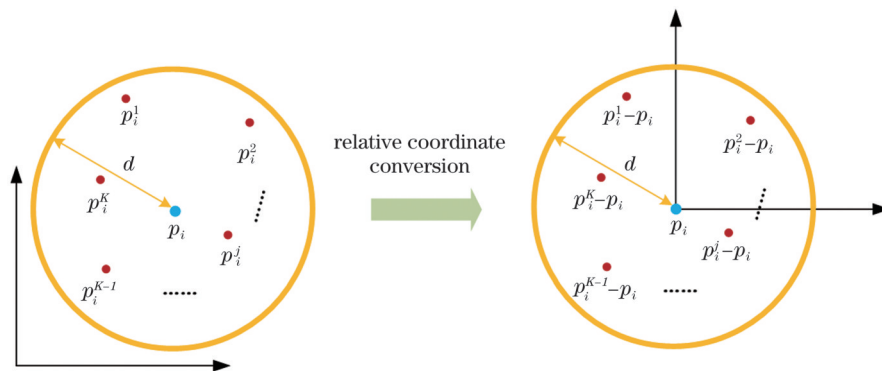


图 2 获取局部区域内点云中点的相对坐标转换

Fig. 2 Acquisition of relative coordinates of neighboring points for each local area

反正切的形式,再提取局部相对特征 F_i^j :

$$F_i^j = \text{MLP} \left\{ \alpha \arctan \left[\beta (\mathbf{p}_i^j - \mathbf{p}_i) \right] \right\}, \quad (2)$$

式中: α 和 β 为调整特征分布形状的参数。

将原特征与修改后的特征进行比较,可以得到如下公式:

$$\begin{cases} \alpha \arctan \left[\beta (\mathbf{p}_i^j - \mathbf{p}_i) \right] > (\mathbf{p}_i^j - \mathbf{p}_i), & (\mathbf{p}_i^j - \mathbf{p}_i) \in (0, l) \\ \alpha \arctan \left[\beta (\mathbf{p}_i^j - \mathbf{p}_i) \right] = (\mathbf{p}_i^j - \mathbf{p}_i), & (\mathbf{p}_i^j - \mathbf{p}_i) = l \\ \alpha \arctan \left[\beta (\mathbf{p}_i^j - \mathbf{p}_i) \right] < (\mathbf{p}_i^j - \mathbf{p}_i), & (\mathbf{p}_i^j - \mathbf{p}_i) \in (l, d) \end{cases}, \quad (3)$$

式中: l 为阈值,当邻居点与中心点的距离为 l 时,原来的相对特征与调整后的相对特征相同。

通过添加 FDA 模块,点云在局部区域的相对特征分布将由原先的线性形式转换为反正切形式。与原来的特征相比,距离中心点更近的邻居点的相对特征值将会增加,而距离中心更远的邻居点的相对特征值将会减少。

2.3 融合坐标注意力机制及注意力池化

注意力机制的主要原理是在输入特征的基础上生成一个权重矩阵,用该矩阵对输入特征进行加权操作。通常情况下,重要的特征对应的权重值较大,而无效特征对应的权重值较小。通过添加不同的权重达到突出重要信息,同时抑制无效特征的目的。由于注意力机

制在神经网络模型中表现出很好的性能,因此被广泛应用于深度学习领域的应用中。

注意力机制种类繁多,包括压缩与激励(SE)注意力机制^[18]、卷积注意力机制模块(CBAM)^[19]、坐标注意力机制(CA)^[20]等。SE注意力模块主要由 squeeze、excitation、reweight 三部分组成,仅仅关注通道信息,忽略了重要的位置信息。CBAM属于混合注意力机制,以串联的方式从空间以及通道两个维度提取特征,但只能捕获局部信息。而CA则是从将通道注意力分解为两个一维的特征编码,沿着两个不同的方向使用并行的池化操作进行特征聚合,可以获取更深远的信息。通过添加SE、CBAM以及CA等注意力模块,能够提升神经网络模型的性能。

平均池化和最大池化是目前最常用的池化方法,它们分别用于求局部区域特征的平均值和最大值,通过特征压缩来加快运算速度。由于可以保持网络的对称性,这两种池化方法在点云局部区域特征提取的过程中,也常被用于局部区域邻居点特征的整合。但这两种池化均不能准确地学习有效特征,可能导致大量有用信息的丢失。在RandLA-Net模型^[21]中提出注意力池化模块,以特征学习的方式获取有用的局部信息。本文用注意力池化替代CA注意力机制中的池化方式,以提升分类和分割结果的准确性,其具体结构如图3所示。

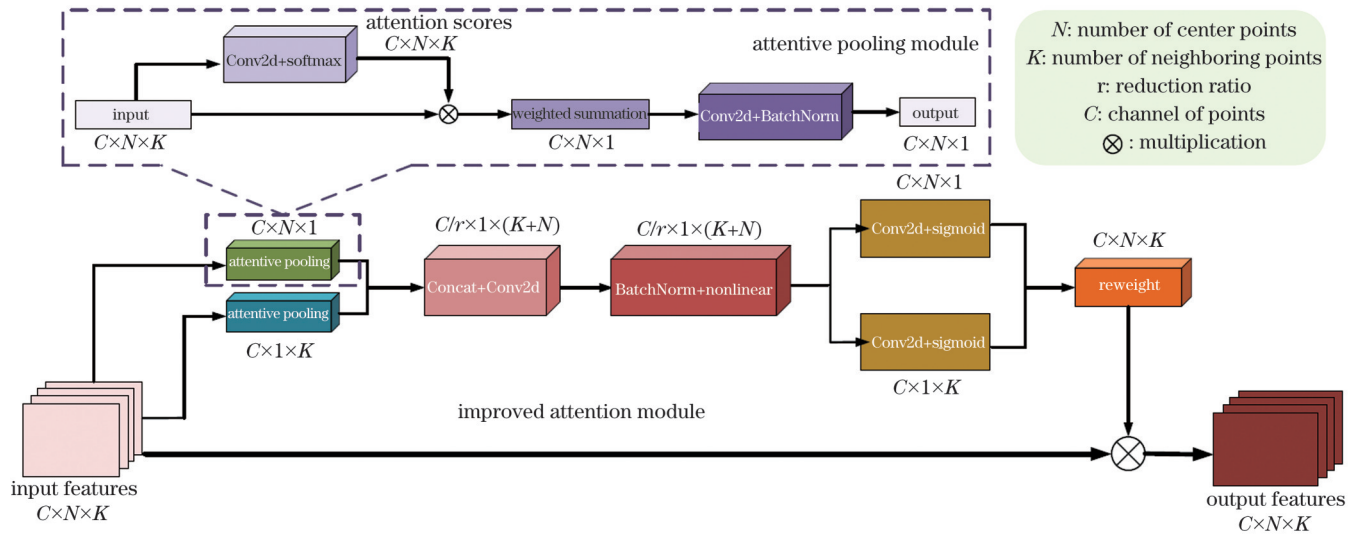


图3 坐标注意力机制与注意力池化融合模型

Fig. 3 Model combining coordinate attention mechanism and attentive pooling

假设给定的输入点云特征为 $F \in \mathbb{R}^{C \times N \times K}$, 其中 C 、 N 、 K 分别表示点云中点的特征维度、中心点个数以及局部区域内邻居点的个数。与只生成单一通道注意力的SE不同,该模型使用两个注意力池化,从中心点和邻居点两个部分分别获取两个特征 F_N 和 F_K , 对应的维度为 $C \times N \times 1$ 和 $C \times K \times 1$ 。在注意力生成阶段,将获取的两个特征进行拼接,并使用卷积函数将原来的 C 维特征压缩为 C/r 维,再使用批量归一化操作以

及swish激活函数得到中间特征 F_m :

$$F_m = \delta \left(\text{Conv} \left([F_N, F_K] \right) \right), \quad (4)$$

式中: $[\cdot, \cdot]$ 为拼接操作; Conv 为卷积变换函数; δ 为swish激活函数。

接着,沿两个方向将中间特征分解为不同的张量 f_N 以及 f_K , 使用不同的卷积将特征转换为 C 维,通过sigmoid激活函数得到两个注意力权重,并与输入特征

执行相乘加权操作,得到该模块的输出特征 F_{out} :

$$F_{out} = F \cdot \left(\sigma \left[\text{Conv}_N(f_N) \right] \right) \cdot \left(\sigma \left[\text{Conv}_K(f_K) \right] \right), \quad (5)$$

式中: Conv_N 与 Conv_K 分别为两个方向上的卷积函数; σ 为 sigmoid 激活函数。

使用注意力池化替代坐标注意力模块中原有的平均池化,如图 3 所示。假设该模块输入的特征为 $F = \{F^1, F^2, \dots, F^K\}$, 对每一个输入特征使用共享的 MLP 以及 softmax 得到对应的注意力分数:

$$s^i = \text{softmax}(\text{MLP}(F^i)). \quad (6)$$

再将得到的注意力分数与输入的特征加权求和,得到局部聚合特征,最后使用卷积函数与批量归一化

操作得到该模块的输出特征 F_{out} :

$$F_{out} = \text{Conv} \left(\sum_{i=1}^K (F^i \cdot s^i) \right). \quad (7)$$

2.4 模型分类分支

由于层级结构,本文提出的网络模型划分出的局部区域逐步扩大,可以获得越来越高维的点云局部特征。模型的输入是一组 $N \times C$ 的点云。在所有实验中,由于只使用输入点云的三维坐标信息 (x, y, z) , 因此输入点云的特征维度 C 设定为 3。模型的分分类分支主要是由多个 SA 层组成的,其结构图分别如图 4 所示。

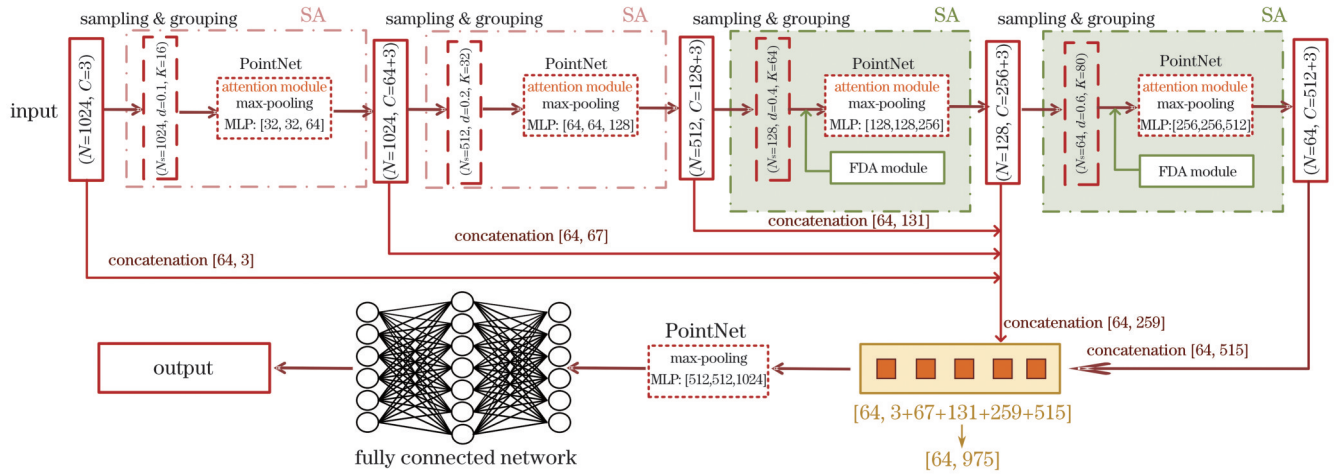


图 4 改进模型的分分类分支

Fig. 4 Classification branch of improved model

SA 由采样层 (sampling)、分组层 (grouping) 和 PointNet 层组成。采样层中使用最远距离采样法从输入的 N 个点中获取 N_s 个采样点。分组层中则以 N_s 个采样点作为中心点,划分出半径为 d 的球形区域作为局部区域,并保证每个局部区域均包含 K 个邻居点。当局部区域内的邻居点的个数少于 K 时,则重复第一个邻居点的坐标,直到达到设定的要求。由于划分出的局部区域较大,在模型的后两个 SA 层的分组层后添加 FDA 模块,用于调整邻居点的相对特征分布,使得越靠近中心点的邻居点对应特征值越大。局部区域中邻居点的坐标转换为以中心点为原点的相对坐标后,再使用 PointNet 层获取每个局部区域中邻居点与中心点之间的相对信息,作为点云的局部特征。两个 FDA 模块中的 α 分别为 $\frac{1}{\pi}$ 以及 $\frac{6}{5\pi}$, β 设定为 5。PointNet 层中添加 2.3 节介绍的注意力模块,用于提升模型提取重要特征的能力。此外,由于特征融合策略,最后一个 SA 中把 64 个采样点在前一层的所有对应的特征进行连接,将特征扩展至 975 维作为点云的全局特征。最后,使用全连接层根据全局特征获得分类预测分数,并根据预测的最高分数确定输入点云的类别。

2.5 模型分割分支

与 PointNet++ 一样,改进模型的分割分支采用了经典的 U-Net 结构,encoder 部分与分类网络相同,在 decoder 部分使用基于插值和连接拼接方式作为该分支主要组成部分。其中,插值是基于 K 个最近邻点的反距离加权平均执行反向插值操作,得到上采样点的特征:

$$f^{(j)}(x) = \frac{\sum_{i=1}^K \omega_i(x) f_i^{(j)}}{\sum_{i=1}^K \omega_i(x)}, \quad (8)$$

式中: $\omega_i(x) = \frac{1}{d_{dis}(x, x_i)^p}$, $d_{dis}(x, x_i)$ 表示点 x 与 x_i 的距离, $p = 2; K = 3; j = 1, \dots, C$ 。

具体而言,对于上采样层中某个点 p_i ,找到上一层中与之距离最近的 K 个点的对应特征进行加权求和,权重为对应距离的反函数。在 decoder 部分采用 skip link concatenation 的方式,将点在 encoder 部分对应的特征进行拼接。最终重复 PointNet 单元 (1×1 卷积网络),直到得到每个点的预测分数,从而确定每一个点的类别。图 5 展示了改进模型的分割分支结构。

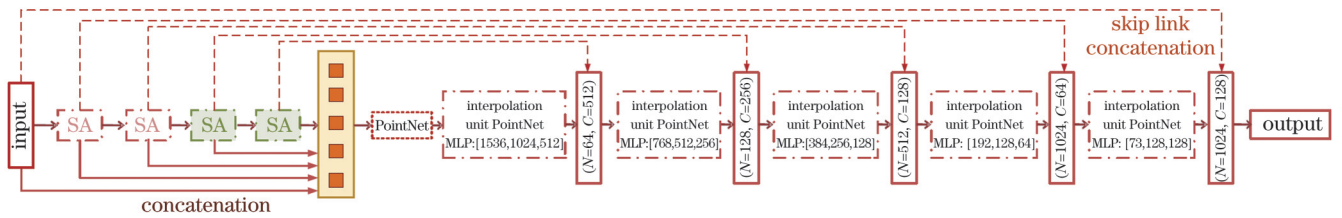


图 5 改进模型的分割分支

Fig. 5 Segmentation branch of improved model

3 实验分析

3.1 点云数据采集与处理

选择江苏镇江宝堰镇南官村杨渡银杏山的林苗圃 (32.12°N, 119.31°E) 作为实验场地, 从 2022 年 7 月到 2023 年 4 月多次采集不同树种的点云并制作数据集。

实验场地如图 6 所示。用 Livox Horizon 激光传感器作为采集工具获取苗圃中树木的点云数据。Livox Horizon 激光传感器具有精确的范围精度和角度精度。传感器积分时间设置为 3000 ms, 以确保采集到足够的点。数据采集过程以及获取到的点云如图 7 所示。激光参数如表 1 所示。

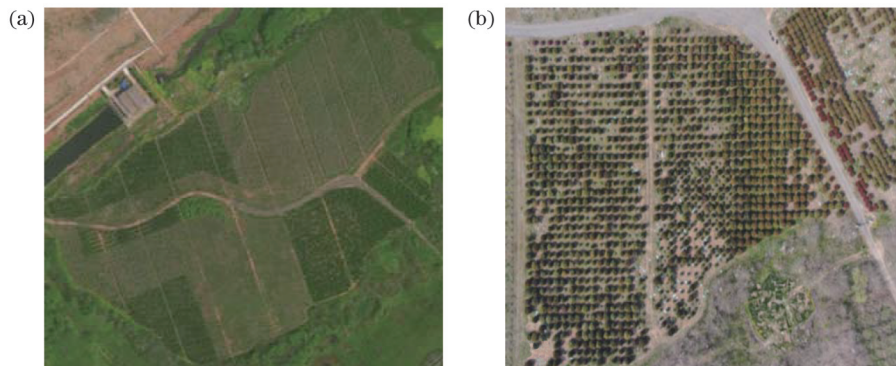


图 6 数据采集的苗圃场景。(a) 苗圃全场景; (b) 苗圃部分场景

Fig. 6 Nursery for data collection. (a) Whole scene of nursery; (b) part of scene of nursery

表 1 激光参数

Table 1 Laser parameters

Parameter	Value
Detection range /m	260
Range error /cm	2
Angle error / (°)	0.05
Field-of-view (FOV) / (°)	81.7×25.1
Data rate / (point·s ⁻¹)	240000

该苗圃中包含多个类别的景观树, 我们采集其中 7 种常见景观树的点云制作数据集: 桂花、海棠、红叶李、鸡爪槭、腊梅、枇杷以及无刺枸骨球。对应的图片如图 8 所示。

为了实现基于点云的树种分类任务以及树冠、树干等各部分割任务, 需要对采集好的点云数据进行处理。除了目标点云, Livox Horizon 激光传感器采集到的数据还包括地面点云、背景点云、噪声点等无用的点。为了获取目标点云, 使用 CloudCompare 软件从原始点云中分割出所需的目标点云。

此外, 分割出的目标点云还需进行标准化处理, 将所有点的坐标转换至 $[-1, 1]$ 之间。可以用下式进行处理:

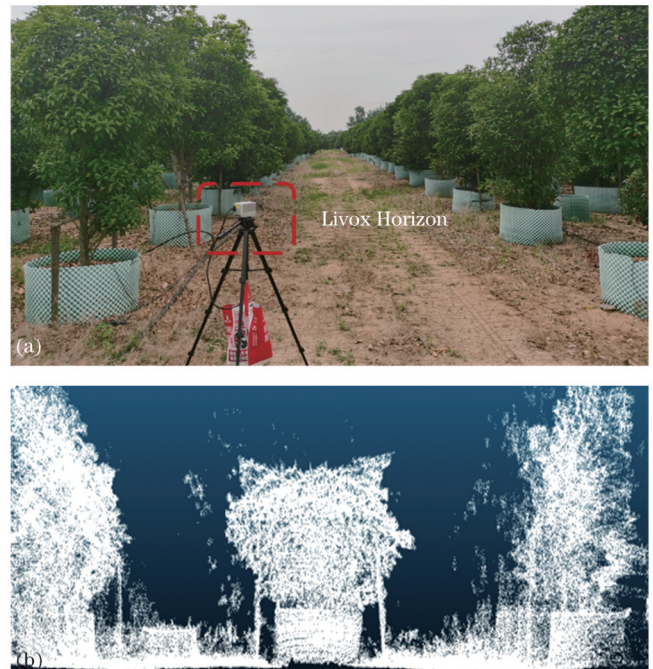


图 7 Livox Horizon 激光传感器以及获取的点云。(a) Livox Horizon 激光传感器; (b) 获取的点云

Fig. 7 Livox Horizon laser sensor and acquired point clouds. (a) Livox Horizon laser sensor; (b) acquired point clouds

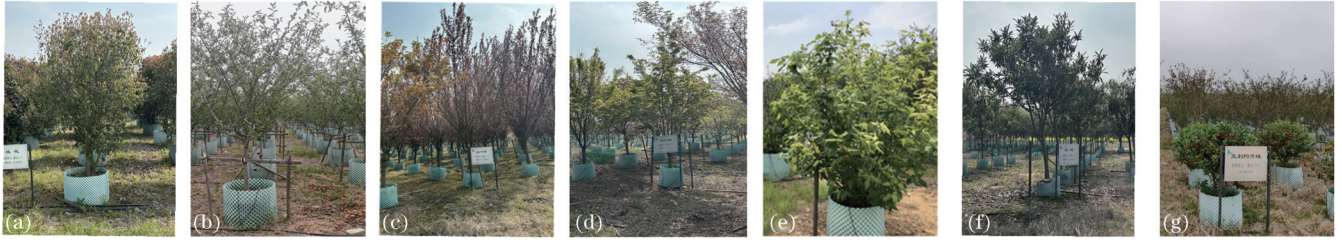


图 8 7 种常见景观树。(a)桂花;(b)海棠;(c)红叶李;(d)鸡爪槭;(e)腊梅;(f)枇杷;(g)无刺枸骨球

Fig. 8 Seven common kinds of landscape trees. (a) *Osmanthus fragrans*; (b) *Malus halliana*; (c) cherry plum; (d) *Acer palmatum*; (e) *Chimonanthus praecox*; (f) loquat tree; (g) Chinese holly

$$\frac{(x - x_{\min}, y - y_{\min}, z - z_{\min})}{\max(x_{\max} - x_{\min}, y_{\max} - y_{\min}, z_{\max} - z_{\min})}, \quad (9)$$

式中: (x, y, z) 表示当前点的坐标; $(x_{\min}, y_{\min}, z_{\min})$ 表示一组点云中三维坐标对应的最小值; $(x_{\max}, y_{\max}, z_{\max})$ 表示一组点云中三维坐标对应的最大值。

共有 1046 棵不同种类的树作为采集样本。为了获取更多的点云数据,从不同的位置进行扫描,最终共获得 1828 组点云数据。根据 7:3 的比例将数据划分为训练集与测试集。每一类树种的数量、获得的点云组数以及对应的训练集和测试集的数量如表 2 所示。

表 2 不同树种采集到的树木数量、得到的点云组数以及训练集和测试集的数量

Table 2 Numbers of samples, point cloud groups, and groups in training and testing sets for different types of trees

Category	Number of trees	Number of point cloud groups	Number of point cloud groups in training set	Number of point cloud groups in testing set
<i>Osmanthus fragrans</i>	120	179	126	53
<i>Malus halliana</i>	144	283	201	82
Cherry plum	164	287	201	86
<i>Acer palmatum</i>	151	290	203	87
<i>Chimonanthus praecox</i>	158	290	203	87
Loquat tree	146	215	151	64
Chinese holly	163	284	199	85

根据实际情况,采集到的点云被进一步细分,每一组点云被分割成 2~4 个部分,并且用不同的标签表示:树冠为 0;树干为 1;种植盆为 2;支撑杆为 3。不同类别的树木分割后的示意图如图 9 所示,为了便于区

分,使用不同颜色的点云对应不同的组成部分(树冠为绿色点云;树干为黄色点云;种植盆为红色点云;支撑杆为褐色点云)。

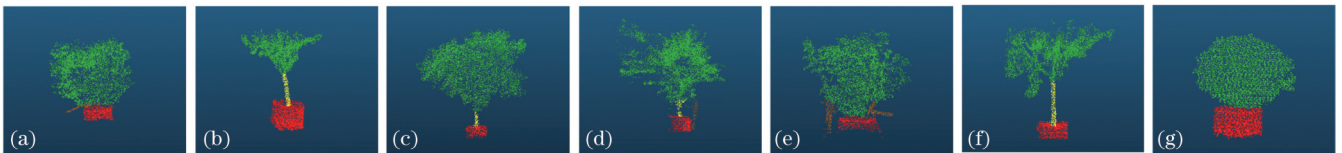


图 9 7 个树种分割后的点云示意图。(a)桂花;(b)海棠;(c)红叶李;(d)鸡爪槭;(e)腊梅;(f)枇杷;(g)无刺枸骨球

Fig. 9 Schematic diagram of segmented point clouds of seven tree species. (a) *Osmanthus fragrans*; (b) *Malus halliana*; (c) cherry plum; (d) *Acer palmatum*; (e) *Chimonanthus praecox*; (f) loquat tree; (g) Chinese holly

3.2 实验环境及参数设置

本文所使用的操作系统为 Ubuntu,对应的软件以及硬件参数如表 3 所示。

3.3 评价指标

在深度学习领域中,评价指标用于评估模型的性能,根据评价指标的值可以判定最优的模型。混淆矩阵主要用于比较分类结果与实际值,并且可以从中了解到每一类预测的情况,以便进行错误分析。混淆矩阵的对角线元素表示模型预测正确的个数。

总体精度(OA, A_o)和平均类精度(mAcc, A_{mAcc})

表 3 实验软硬件参数

Table 3 Software and hardware parameters for experiment

Parameter	Content
CPU	Intel(R) Xeon(R) Gold 6226R
GPU	NVIDIA RTX3090
Cuda	Cuda11.7
Data processing tools	Python3.8, PyCharm2020
Deep learning framework	Pytorch

是用于评价模型分类性能的两个常用指标。OA 表示使用测试集的总体准确率, mAcc 则表示平均类别准

准确性。OA 与 mAcc 的计算公式分别为

$$A_o = \frac{N_{TP}}{W}, \quad (10)$$

$$A_{mAcc} = \frac{1}{m} \sum_{i=1}^m \frac{N_{TP_i}}{W_i}, \quad (11)$$

式中: N_{TP} 表示预测正确结果的数量; W 为测试集中点云的总个数; N_{TP_i} 表示第 i 类点云预测正确个数; W_i 是测试集中第 i 类点云的总个数。

分类任务主要是判定输入点云的类别, 分割任务通常需要确定输入点云中每一个点的类别。交并比 (IoU, R_{IoU}) 以及平均交并比 (mIoU) 用于评价分割模型。IoU 可以用下式表示:

$$R_{IoU} = \frac{N_{TP}}{N_{FP} + N_{TP} + N_{FN}}, \quad (12)$$

式中: N_{TP} 、 N_{FP} 和 N_{FN} 分别表示预测为真阳性、假阳性和假阴性的点数。

mIoU 表示各个部分 IoU 的平均值。

除 mIoU 以外, 精确率 $R_{precision}$ 、召回率 R_{recall} 及 F_1 值都是常用于评价模型分割性能的指标, 分别表示为

$$R_{precision} = \frac{N_{TP}}{N_{FP} + N_{TP}}, \quad (13)$$

$$R_{recall} = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad (14)$$

$$F_1 = \frac{2 \times R_{recall} \times R_{precision}}{R_{recall} + R_{precision}}. \quad (15)$$

本文将使用 OA、mAcc、mIoU、精确率、召回率及 F_1 值作为评价算法性能的指标。评价指标的值越大, 代表算法的性能越好。

3.4 树种分类实验与结果分析

下面使用制作的点云数据进行树种分类实验, 并对实验结果进行分析。点云分类任务的主要目的是确定输入点云的类别。当点云样本数据较少, 又想提高网络模型的泛化性时, 可以使用迁移学习的方法防止模型过拟合现象的发生。因此, 我们先使用公共数据集对网络模型进行预训练, 再调整网络模型结构, 使用自制的数据集得到分类实验结果。

在本次实验的预训练中, 使用了经典的点云分类数据集 ModelNet40。该数据集中共包含常见的 40 个类别物体的点云, 共有 12311 组点云数据。其中, 9843 组点云作为测试集数据, 2468 组点云作为训练集数据。每组点云包含 10000 个点, 在实验过程中, 网络模型将每次从数据中采样 1024 个点进行训练以及测试。使用该数据集对 PointNet, PointNet++ 以及改进网络模型进行训练与测试。本文提出的模型中添加了融合注意力池化的 CA 模块, 为了验证添加改进的注意力模块对模型性能的影响, 分别使用 SE、CBAM 以及 CA 进行训练测试。预训练完成后, 调整网络模型中全连接层最后一层输出节点的个数, 将预训练好的模型中的 40 个输出节点调整成 7 个, 网络模型其他参数

保持不变。更换数据集后, 再次训练网络模型, 对模型的参数不断进行迭代。用训练好的网络模型进行测试, 测试结果如表 4 所示。

表 4 使用自制数据集的模型分类测试结果

Table 4 Testing results of classification on self-made dataset

unit: %		
Model	OA	mAcc
PointNet	81.25	80.00
PointNet++	89.71	89.44
Ours (SE)	90.33	92.39
Ours (CBAM)	90.83	92.28
Ours (CA)	91.33	93.13
Ours	92.50	94.22

根据表 4, 本文提出的模型的性能优于 PointNet 以及 PointNet++; 改进后模型预测的总体精度为 92.50%, 比 PointNet 高 11.25 个百分点, 比 PointNet++ 高 2.79 个百分点; 改进后模型的平均类精度为 94.22%, 比 PointNet 高 14.22 个百分点, 比 PointNet++ 高 4.78 个百分点。另外, 与使用 SE、CBAM 以及不包含注意力池化的 CA 模型相比, 本文提出的改进模型可以获得更高的分类精度。

为了进一步对预测结果进行分析, 我们使用混淆矩阵展示三个模型对每一类树种的预测结果, 如图 10 所示。图 10 中每一列数据的总和表示该类树种在测试集中的个数。对角线元素颜色越深, 则表示该类预测越准确。与其他两个模型相比, 改进后模型除了对腊梅以及枇杷树预测准确性略低, 其他树种预测准确性均优于其他两种模型。比较其他树种分类, 三种模型对鸡爪槭、桂花以及枇杷这三类分类的准确性较低: 桂花误判为腊梅, 鸡爪槭误判为枇杷, 枇杷误判为鸡爪槭。造成分类准确率低的原因可能是由于桂花与腊梅、鸡爪槭与枇杷外形相似。自制的点云数据集仅保留了物体的坐标信息, 没有保留颜色信息, 因此, 只依靠点云信息很难实现准确分类。

3.5 树冠树干等分割实验与结果分析

与判断输入点云类别的分类模型相比, 分割模型需要确定点云中每一个点的类别。由于本文所述分割实验的主要目的是分割出树的树冠、树干、种植盆以及支撑杆, 因此在分割实验中, 不再根据树的种类进行划分。将第 2 节中处理好的点云数据按照 7:3 的比例划分训练集以及测试集。分割实验中, 每次将从原有的点云中采样 2048 个点, 并经过模型特征提取、分类, 最终确定采样点中每一个点的类别。用训练好的 PointNet, PointNet++ 以及包含各种注意力模块的网络模型进行测试, 测试结果如表 5 所示。

根据表 5 所示, 在分割任务中, 本文提出的模型得到的 mIoU、精确率、召回率以及 F_1 值分别为 89.09%、90.09%、95.44% 以及 92.59%, 比 PointNet++ 分割得

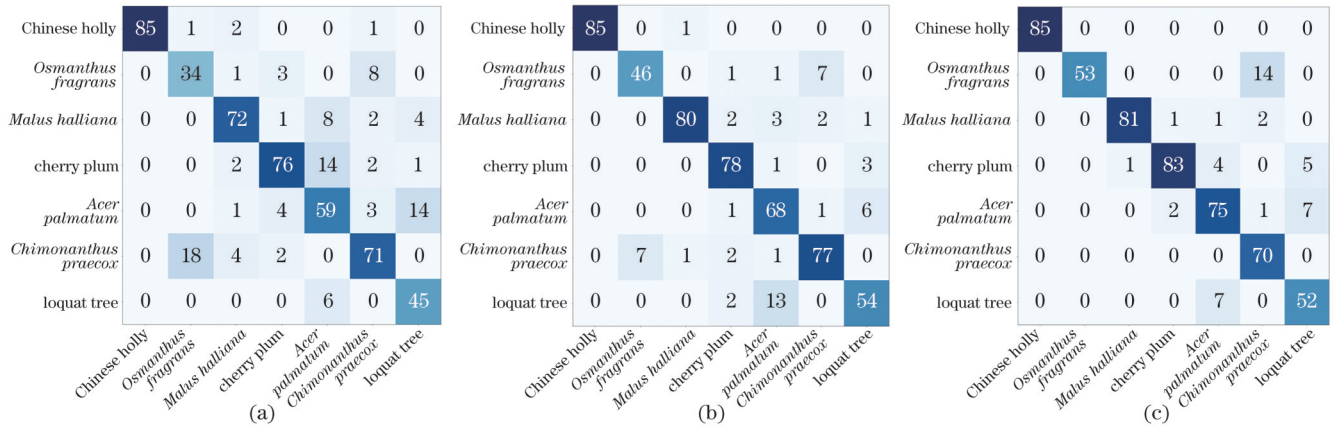


图 10 混淆矩阵。(a)PointNet; (b)PointNet++; (c)ours

Fig. 10 Confusion matrixes. (a) PointNet; (b) PointNet++; (c) ours

表 5 使用自制数据集的模型分割测试结果

Table 5 Testing results of segmentation on self-made dataset

Model	mIoU	Precision	Recall	F_1 score
PointNet	76.95	70.71	76.01	73.01
PointNet++	83.85	85.35	92.40	88.47
Ours (SE)	87.08	89.32	94.86	91.45
Ours (CBAM)	87.61	88.56	94.29	91.17
Ours (CA)	89.89	88.29	94.51	91.12
Ours	89.09	90.09	95.44	92.59

unit: %

到的结果分别高出 5.24、4.74、3.04、4.12 个百分点。与包含不同注意力模型的模型相比,尽管本文提出的模型 mIoU 略低于直接使用 CA 的模型,但其他指标均为最高。

为了进一步对分割后的结果进行说明,将 PointNet, PointNet++ 以及改进模型分割出的部分结果进行可视化操作,并进行比较,如图 11 所示。

图 11 中白框标出了明显预测错误的点云。与 PointNet 相比,PointNet++ 的预测准确度较高,但仍然在细节部分分割错误,如将支撑杆预测为树冠或种植盆等。推测导致这一结果的主要原因是 PointNet++

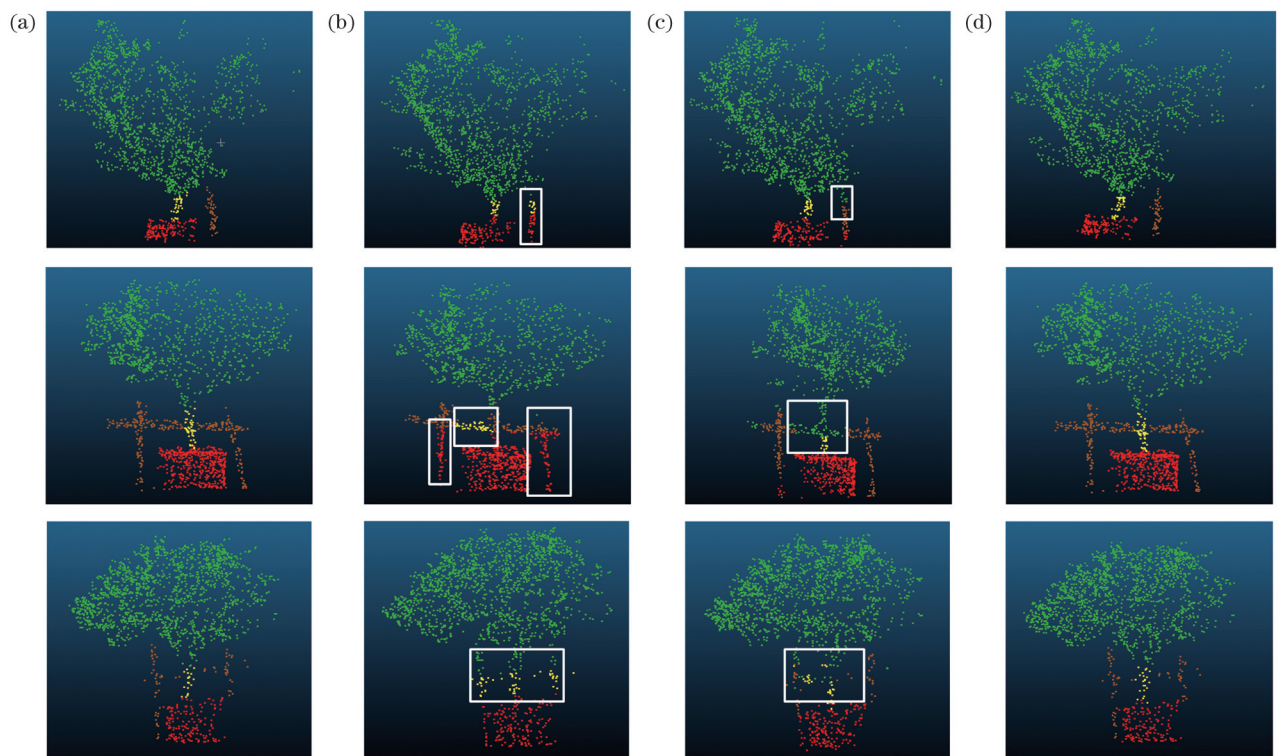


图 11 PointNet, PointNet++ 和改进模型分割结果可视化示例(白框标出明显预测错误的点)。(a)基准真值; (b)PointNet 分割的结果; (c)PointNet++ 分割的结果; (d)本文模型分割的结果

Fig. 11 Examples of visualization of segmentation results using PointNet, PointNet++, and proposed model (white boxes denote wrong predicted points). (a) Ground truth; (b) segmentation results of PointNet; (c) segmentation results of PointNet++; (d) segmentation results of proposed model

忽略了包含更多细节信息的点云低维特征。但由于改进的模型同时结合了低维以及高维的点云特征,并且提高了在局部区域特征提取的能力,因此可以将支撑杆与树干、树冠较为准确地分割出来。

4 结 论

本文提出了一种基于 PointNet++ 改进的神经网络模型,可实现树种分类和树木各部分的分割任务,适用于农业机器人在苗圃场景中的实际应用。和经典的点云神经网络模型 PointNet++ 不同,所提出的模型充分利用点云的高维以及低维特征以提高性能。此外,在构建局部区域时,将处于较大局部区域内邻居点的相对特征调整为反正切的形式分布,提高更靠近中心点的邻居点对应的特征值。使用坐标注意力机制与注意力池化相结合的模块,加强对点云中关键局部特征的提取。

使用自制苗圃点云数据集进行了分类、分割实验。实验结果表明,与 PointNet 和 PointNet++ 相比,本文提出的模型在分类和分割任务中均实现了更优性能。后续工作中,拟将本文提出的模型运用于硬件平台,用以获取苗圃中各类树木信息,从而帮助实现机器人路径规划、自主对靶、变量喷雾等操作。

参 考 文 献

- [1] 王雁. 我国城市绿化理念探索的回顾与发展趋势[J]. 浙江林学院学报, 2003, 20(3): 315-320.
Wang Y. Review and expectation of the development and ideas of China city greening[J]. Journal of Zhejiang Forestry College, 2003, 20(3): 315-320.
- [2] 姜红花, 刘理民, 柳平增, 等. 面向精准喷雾的果树冠层体积在线计算方法[J]. 农业机械学报, 2019, 50(7): 120-129.
Jiang H H, Liu L M, Liu P Z, et al. Online calculation method of fruit trees canopy volume for precision spray[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(7): 120-129.
- [3] 沈跃, 庄珍珠, 刘慧, 等. 基于 RealSense 深度相机的多特征树干快速识别方法[J]. 农业机械学报, 2022, 53(4): 304-312.
Shen Y, Zhuang Z Z, Liu H, et al. Fast recognition method of multi-feature trunk based on RealSense depth camera[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(4): 304-312.
- [4] Zamanakos G, Tsochatzidis L, Amanatiadis A, et al. A comprehensive survey of lidar-based 3D object detection methods with deep learning for autonomous driving[J]. Computers & Graphics, 2021, 99: 153-181.
- [5] Guo Y L, Wang H Y, Hu Q Y, et al. Deep learning for 3D point clouds: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(12): 4338-4364.
- [6] Su H, Maji S, Kalogerakis E, et al. Multi-view convolutional neural networks for 3D shape recognition[C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2016: 945-953.
- [7] Feng Y F, Zhang Z Z, Zhao X B, et al. GVCNN: group-view convolutional neural networks for 3D shape recognition[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 264-272.
- [8] Maturana D, Scherer S. VoxNet: a 3D convolutional neural network for real-time object recognition[C]//2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), September 28-October 2, 2015, Hamburg, Germany. New York: IEEE Press, 2015: 922-928.
- [9] 张满囤, 燕明晓, 马英石, 等. 基于八叉树结构的三维体素模型检索[J]. 计算机学报, 2021, 44(2): 334-346.
Zhang M D, Yan M X, Ma Y S, et al. 3D voxel model retrieval based on octree structure[J]. Chinese Journal of Computers, 2021, 44(2): 334-346.
- [10] Charles R Q, Hao S, Mo K C, et al. PointNet: deep learning on point sets for 3D classification and segmentation[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 77-85.
- [11] Charles R Q, Li Y, Hao S, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space[C]//Advances in Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA. [S.l.: s.n.], 2017: 5099-5108.
- [12] 易倩, 钟浩宇, 刘龙, 等. 基于 ROI-RSICP 算法的车轮廓动态检测[J]. 中国激光, 2020, 47(11): 1104006.
Yi Q, Zhong H Y, Liu L, et al. Dynamic inspection of wheel profile based on ROI-RSICP algorithm[J]. Chinese Journal of Lasers, 2020, 47(11): 1104006.
- [13] 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 L, Wang H Y, Ji C M, et al. Estimation algorithm of apple phenotypic parameters based on PointNet and transfer learning[J]. Journal of Nanjing Agricultural University, 2021, 44(6): 1209-1216.
- [14] Hu H, Yu J C, Yin L, et al. An improved PointNet++ point cloud segmentation model applied to automatic measurement method of pig body size[J]. Computers and Electronics in Agriculture, 2023, 205: 107560.
- [15] 陈向宇, 云挺, 薛联凤, 等. 基于激光雷达点云数据的树种分类[J]. 激光与光电子学进展, 2019, 56(12): 122801.
Chen X Y, Yun T, Xue L F, et al. Classification of tree species based on lidar point cloud data[J]. Laser & Optoelectronics Progress, 2019, 56(12): 122801.
- [16] Zhang Y P, Tan Y L, Onda Y, et al. A tree detection method based on trunk point cloud section in dense plantation forest using drone lidar data[J]. Forest Ecosystems, 2023, 10: 100088.
- [17] 卢华清, 伍吉仓, 张子健. 面向地基激光点云的树木枝叶分离[J]. 中国激光, 2022, 49(23): 2310001.
Lu H Q, Wu J C, Zhang Z J. Tree branch and leaf separation using terrestrial laser point clouds[J]. Chinese Journal of Lasers, 2022, 49(23): 2310001.
- [18] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7132-7141.
- [19] Zhao Z X, Chen K X, Yamane S. CBAM-Unet: easier to find the target with the attention module "CBAM"[C]//2021 IEEE 10th Global Conference on Consumer Electronics (GCCE), October 12-15, 2021, Kyoto, Japan. New York: IEEE Press, 2021: 655-657.
- [20] Hou Q B, Zhou D Q, Feng J S. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 13708-13717.
- [21] Hu Q Y, Yang B, Xie L H, et al. RandLA-net: efficient semantic segmentation of large-scale point clouds[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 11105-11114.

Point Clouds Classification and Segmentation for Nursery Trees Based on Improved PointNet++ Model

Xu Jie, Liu Hui*, Shen Yue, Yang Guanxue, Zhou Hao, Wang Siyuan

School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, Jiangsu, China

Abstract

Objective Nurseries span large areas, so that managing and protecting landscape trees manually requires significant resources. Substituting agricultural robots for human forces can effectively address issues of low productivity and labor shortages. The perception systems are essential elements of agricultural robots. Agricultural robots can obtain essential information for tree statistical analysis, autonomous navigation, and target spray operations by perceiving tree species, trunks, crowns, and others. Therefore, it is of great significance to design a neural network model which can accurately obtain the information of tree species, crowns and trunks. Point clouds are more suitable for outdoor scenes due to their greater robustness to different lighting conditions. While there are many studies related to tree species classification and crown or trunk segmentation, most of them focus on trees located in forestry areas or require prior point cloud registration operations. Compared with trees in forest areas, landscape trees planted in nurseries are generally smaller in size. Moreover, during the process of work, robots are only able to capture partial point clouds of trees. In order to ensure the real-time performances, robots need to obtain relevant information on tree species, trunks, and crowns from partially scanned point clouds. Hence, based on PointNet++, we present an improvement model in this paper. We collect partial point clouds of trees from nurseries to train the model for tree species classification, as well as segmentation tasks of tree crowns, trunks and other parts. Experimental results demonstrate its superior classification and segmentation performances, which can prepare for the subsequent implementation of the application on the hardware platform.

Methods The proposed neural network model adopts the hierarchical structure. Multiple set abstraction (SA) layers are used to extract the local features of the point clouds. Each SA layer consists of a sampling layer, a grouping layer, and a PointNet layer. The sampling layer employs the iterative farthest point sampling to select sampling points. Then, taking the sampling points as the center points, spherical regions with a fixed radius d are constructed as the local areas in the grouping layer. Each local area contains K neighboring points. As the number of layers increases, the radius d of the spherical region expands continuously. For larger local areas, the feature distribution adjustment module is employed to convert the relative features of neighboring points from linear to arc tangent, increasing the relative features of those closer to the central point while reducing the influence of distant neighboring points in each local area. Multilayer perceptrons (MLPs) are employed for extracting local features in the PointNet layer. To enhance the model's capacity to capture important information, we integrate the coordinate attention (CA) module with attention pooling for extracting local features. Furthermore, in the final SA layer, the model concatenates the low-dimensional and high-dimensional features of sampled points, and employs fully connected networks to predict the category of the input point clouds. The segmentation branch utilizes the U-Net structure, and employs an interpolation method based on inverse distance weighted average with K nearest neighbors, combining with skip links across levels for feature propagation. Finally, the category of each point is obtained to realize segmentation.

Results and Discussions For classification and segmentation experiments, we collect seven kinds of point clouds of common landscape trees in the nursery using the Livox Horizon laser. During the data collection process, the parameters of the laser are shown in Table 1, and more detailed information of the collected point clouds is shown in Table 2. To perform segmentation tasks, each group of collected point clouds is further processed by dividing into 2–4 parts represented with different labels (Fig. 9). To demonstrate better classification and segmentation performances, the experimental results of the proposed model are compared with classic PointNet and PointNet++ using the self-made dataset. Besides, we conduct experiments by replacing the improved attention module in our proposed model with squeeze-and-excitation (SE) attention module, convolutional block attention module (CBAM), and CA attention module. For the classification experiments, the models are first pre-trained using the public dataset ModelNet40 before being trained with the self-made dataset. And the results are shown in Table 4. The overall accuracy (OA) and mean class accuracy (mAcc) of the proposed model are 92.50% and 94.22%, respectively, which are higher than the corresponding metrics of PointNet and PointNet++. Compared with other kinds of attention modules, the attention model utilized in the proposed model achieves the highest classification accuracy. The same models are also trained for the segmentation experiments, and the testing results are presented in Table 5. The evaluation indicator values including average intersection over union (mIoU), precision, recall and F_1 score of the proposed model are 89.09%, 90.09%, 95.44% and 92.59%, respectively, which are higher than those of PointNet++. Furthermore, our proposed model outperforms PointNet++ in terms of accurately capturing fine details (Fig. 11).

Conclusions In this paper, we present an improved neural network model based on point clouds, which is designed to classify tree species in the nursery and segment tree crowns, trunks, and other parts. Different from the classic neural network model

PointNet++ , the proposed model combines the high-dimensional and low-dimensional features of the points to improve the performance. And the relative features of the neighboring points are adjusted to the form of arc tangent distribution , which can improve the corresponding values of the neighboring points closer to the central point in the local areas. Furthermore, the CA attention module and attention pooling module are combined to improve the ability of capturing the important features. To enable the practical application of agricultural robots, we collect and process point cloud data of landscape trees in the nursery to train the models. The experimental results demonstrate that our proposed network outperforms PointNet and PointNet++ in terms of both classification and segmentation tasks. The model proposed in this paper provides the theoretical foundation for future applications of agricultural robots in the nursery, such as tree statistical analysis, autonomous navigation, and target spray operations.

Key words remote sensing; lidar; deep learning; tree species classification; crown and trunk segmentation of nursery trees; PointNet++