

基于先验几何特征的深度学习对单木点云组件分割

干瑞霖^{1,2}, 杨健¹, 罗宾汉^{1,3}, 史硕⁴, 杜霖¹,
吴中亮¹, 王思皓¹, 王澳⁴

1. 中国地质大学(武汉) 地理与信息工程学院, 武汉 430074;
2. 长江设计集团有限公司 长江岩土工程有限公司, 武汉 430010;
3. 武汉大学 电子信息学院, 武汉 430072;
4. 武汉大学 测绘遥感信息工程全国重点实验室, 武汉 430079

摘要: 对树木木质和叶部结构的有效区分是准确估算其关键结构参数和精确反演生物量的重要基础。激光雷达为树木结构参数与地上生物量的非破坏性估算提供了新的技术支撑。然而, 当前单木组件分割算法对不同树种的泛化能力有限, 对细小枝干分割能力不足。因此, 本文构建了一个包含713棵单木且树种丰富的大尺度单木组件分割数据集, 同时引入高性能的Point Transformer-V3深度学习网络来获取可靠的分割结果。此外, 通过构建单木点云的先验几何特征(显著性差异指标、主成分指标和垂直度指标), 来优化深度学习网络在单木点云上的分割性能, 并与当前主流的几种深度学习算法进行比较。研究结果表明: 引入构建的先验几何特征的Point Transformer-V3网络进行多树种单木组件分割的OA、mAcc与mIoU分别达到了0.946、0.872和0.806; 即使是深入树冠较小尺度的精细分支也能实现一定提取, 且其在几何形态结构差异较大的不同树种单木上也表现出较高适用性。最后, 通过消融实验验证结果表明在训练时引入单木先验几何特征能较好促进网络分割效果。因此, 本文构建的方法可以进一步促进深度学习在单木组件分割中的应用, 为树木结构参数精确估算提供技术支撑。

关键词: 单木组件, 激光点云, 语义分割, 先验特征, 深度学习, 木叶分离, 枝干提取, 序列化注意力机制
中图分类号: TP701/P2

引用格式: 干瑞霖, 杨健, 罗宾汉, 史硕, 杜霖, 吴中亮, 王思皓, 王澳. 2026. 基于先验几何特征的深度学习对单木点云组件分割. 遥感学报, 30(5): 1326-1340

Gan R L, Yang J, Luo B H, Shi S, Du L, Wu Z L, Wang S H and Wang A. 2026. Deep learning based on prior geometric features for the segmentation of individual tree point cloud components. National Remote Sensing Bulletin, 30(5): 1326-1340 [DOI: 10.11834/jrs.20265201]

1 引言

森林作为地球上最重要的陆地生态系统之一, 约占全球陆地面积31% (Arrizza等, 2024), 是全球碳循环的重要组成部分。研究树木生理生化参数与不同林木结构生物量, 有助于更好了解全球碳动态变化以及冠层与大气之间的相互作用 (Demol等, 2022; Oehmcke等, 2024)。然而, 有效区分木质和叶部结构, 是准确估算树木关键结构参数和精确反演生物量的重要基础。地基激光雷达TLS (Terrestrial Laser Scanning) 的出现, 可无损获取

精细的单木空间结构点云, 提供单木的几何细节信息, 为单木枝叶分割提供技术支持 (Chen等, 2019; Terryn等, 2022)。

目前, 基于TLS点云的单木组件分割算法可分为3大类: 基于辐射特性、基于几何特性以及基于图结构特性的方法 (Arrizza等, 2024; Chen等, 2025)。基于辐射特性的方法主要依赖于叶片和木质部分对激光回波强度LRI (Laser Return Intensity) 的影响实现对TLS点云的叶片和木质部分进行分割 (Li等, 2013)。然而, 这种方法需要使用特定扫描仪器, 在数据采集前还需要对其进行精确的辐射

收稿日期: 2025-05-26; 预印本: 2026-01-23

基金项目: 国家自然科学基金面上项目(编号: 42271388, 42171347); 湖北省自然科学基金(编号: 2024AFA069)

第一作者简介: 干瑞霖, 研究方向为点云林业应用。E-mail: grl_rs@cug.edu.cn

通信作者简介: 杨健, 研究方向为荧光遥感、点云智能解译。E-mail: yangjian@cug.edu.cn

校准 (Calders等, 2017; Danson等, 2014), 且LRI还受距离、激光入射角、植被类型及树种等多种因素的影响 (Kaasalainen等, 2009; Tao等, 2015; Tian和Li, 2022a), 进而限制起在单木组件分离的有效应用。基于几何特性的方法主要利用点云的空间三维坐标信息进行空间特征提取实现木质与叶片组分分割, 主要方法如LeWoS (Wang等, 2020)、TLSeparation (Vicari等, 2019)、基于密度的空间聚类 (Ferrara等, 2018) 以及随机森林等 (Moorthy等, 2020)。这些方法通过分析点云的几何特征对木质点与叶片点进行分割 (Wang等, 2017)。然而, 由于数据集规模较小且几何特征构建不足, 其在细小枝叶分离过程中, 泛化能力仍需进一步提升。基于图结构特性的单木组件分割方法原理是利用点云数据的拓扑特征和几何特性, 通过构建网络图模型实现枝叶分类, 其核心思想是将单木点云表示为一个连通的拓扑网络。Dong等 (2020) 和Strimbu等 (2015) 基于图结构特征分割算法成功从点云中提取出单木冠层; Hui等 (2021) 提出了一种模式点演化分离方法, 通过设置约束条件并利用均值漂移方法构建单木网络图, 通过路径回溯和计算节点访问频率完成木质与叶部分离。Tian等 (2022a) 开发了基于图结构的greedy best-first search (GBS) 算法, 基于最短路径特征进行木叶分离。然而, 这些算法在进行精细枝叶分离过程中, 很难对细支特征进行充分挖掘, 从而限制了其在枝叶分离中的进一步应用。

随着人工智能技术的不断发展, 其可以更好的对点云高维特征进行挖掘并学习, 近年已有研究将深度学习语义分割模型应用于单木点云组件分割任务中 (Wu等, 2020; Xi等, 2023, 2018, 2020; Zhang等, 2024a), 并取得了一定效果。Xi等 (2018) 采用了3D全卷积网络在林地尺度上从TLS数据中过滤木质与非木质成分; Morel等 (2020) 对PointNet++进行改进, 将深度学习与几何特征分析相结合, 以应对点云叶片对细支遮挡的问题; Han和Sánchez-Azofeifa (2022) 引入时间序列分析优化叶木分类, 并使用类激活映射解释深度学习决策过程; Dai等 (2023) 提出了一种新型多方向协作卷积神经网络, 并在不同林地尺度上评估了其在单木组件分割任务中的性能; Zhang等 (2024a) 则将系数卷积与自注意力机制相结合, 构建了高性能的SPTNet模型, 并在细支分离方面取得了较好的

效果。Wu等 (2024) 基于经典的U-Net框架模式提出了Point Transformer-V3深度学习网络, 通过编码器和解码器, 对输入的点云特征进行升维后学习和提取, 再将高维特征还原到原始点云维度, 实现各点类别提取。

当前深度学习模型已经在单木点云的枝叶分离方面开展了大量相关应用研究, 然而树形和尺寸的异质性、树冠内部的遮挡、树干密度, 以及TLS点云的巨大规模和其原始无序性都将影响细小枝干的进一步分割 (Liu等, 2020)。因此, 针对枝叶混合区域进行有效且高效的分离还有待进一步深入探讨, 本文拟通过构建单木先验几何特征 (线性度量、平面度量和离散度量), 用以优化Point Transformer-V3深度学习网络, 对单木组件包括主干 (Trunk)、分支 (Branch) 和叶片 (Foliage) 类别进行分割, 最后, 与当前主流的几种深度学习算法进行实验验证比较, 以期通过构建的点云输入特征优化网络训练, 加强对单木组件各个点云的特征学习, 提升单木组件分割精度。

2 数据来源与研究方法

2.1 数据来源

本研究对目前所能获取的开源带有语义标签的单木尺度和地块尺度的单木组件分割数据集进行了整合, 并对其进行二次语义标注。包含了TLS点云数据 (Anonymous, 2024; Mspace Lab, 2024; Tian和Li, 2022b; Wan等, 2021; Wang等, 2021; Weiser等, 2024; Xi等, 2020), 并引入无人机UAV (Unmanned Aerial Vehicle) 扫描点云数据 (Puliti等, 2023) 来扩充训练样本。最终构建成来自中国、芬兰、德国等15个国家的20余种不同树种的扫描树木数据, 共计713棵具有高质量语义标签的大型单木点云组件分割的Individual Tree Semantic-3D (ITS-3D) 数据集。所有点云数据集均为单木尺度, 统一划分为树干、树枝和树叶3个类别。其中, TLS点云数据的单棵树点云体量在2.0万—1000.0万个, 共计394棵单木。UAV点云数据集, 本研究在原数据集中逐个筛选出了冠层以下部分各级分支完好, 且点密度较高的单木数据, 使其与TLS数据集点云之间的差别尽可能小, 筛选后单棵树点云体量大约在1.6万—30.0万个, 共计319棵单木。

2.2 单木先验几何结构特征构建

相关研究表明基于曲率、密度以及协方差矩阵特征值的特征能够表征单木结构特点的先验几何特征,有助于提升单木组件分割效果(Dai等, 2023; Wan等, 2021; Zhang等, 2024a)。然而,并非所有先验几何特征在不同树种的单木组件上均有较好区分度,加入类别区分度较低的冗余特征不仅无法帮助分类器的训练,还可能会降低分割精度。

然而,在点云数据中各点可与其邻域内的点构成协方差矩阵,通过对协方差矩阵进行特征分解,计算出其3个非负的特征值: $\lambda_0 \geq \lambda_1 \geq \lambda_2 \geq 0$,则可计算线性度量、平面度量和离散度量3个基本几何结构特征。对分布不均于叶片组分,没有明显方向性,则3个特征值大小相近($\lambda_0 \approx \lambda_1 \approx \lambda_2$),叶片点的离散度相对较高。对于一级主枝(即树干组分),其通常呈较为规则的圆柱体,且直径较大,点密度相对较高,各点邻域通常近似构成平面,故前两个特征值的大小通常较为接近,且远远大于第三特征值($\lambda_0 \approx \lambda_1$, 远大于 λ_2),故树干点的平面度相对较高。对于高级子分支组分,其枝干直径相对较小,呈细长状生长,结构近似线性,其最大特征值远远大于其他两个特征值(λ_0 远大于 $\lambda_1 \approx \lambda_2$),分支点的线性度明显较大。

因此,根据单木组件特点对先验几何特征分析,本研究提取了显著性差异指标 SoD (Significance of Difference)、主成分指标 PCA2 (Principal Component Analysis 2) 以及垂直度 (Verticality) 3个单木先验结构特征作为深度学习网络训练的输入特征。为突出线性度显著性,SoD可充分利用其线性特征,分别细小高级分支(Wan等, 2021)。具体计算公式如下:

$$\text{SoD}(L) = L + (1 - L) \times (L - \max(P, S)) \quad (1)$$

$$L = \frac{\sqrt{\lambda_0} - \sqrt{\lambda_1}}{\sqrt{\lambda_0}}, P = \frac{\sqrt{\lambda_1} - \sqrt{\lambda_2}}{\sqrt{\lambda_0}}, S = \frac{\sqrt{\lambda_2}}{\sqrt{\lambda_0}} \quad (2)$$

式中, $\text{SoD}(L)$ 为线性度的显著性差异指标, L 为线性度, P 为平面度, S 为离散度。 λ_0 、 λ_1 、 λ_2 分别表示主成分分析非负的3个特征值。 SoD 的取值范围为 $[-1, 1]$, 当其值大于0时,代表当前点的线性度处于显著地位,其邻域线性特征明显;反之,则其他两个几何结构度量更显著,其线性特征不明显(Wan等, 2021)。

PCA2指标则根据 λ_1 与其他两个特征值之间的

相对大小关系,对树干、高级分支以及叶片3个组分进行了更进一步的区分。具体计算公式如下:

$$\text{PCA2} = \frac{\lambda_1}{\lambda_0 + \lambda_1 + \lambda_2} \quad (3)$$

式中, PCA2取值范围在 $[0, 0.5]$, 对于具有线性结构的高级分支,其 λ_1 远小于 λ_0 ,故分支组分的PCA2值接近于0,在整个单木结构中最小;而对于枝干组分,其 λ_1 与 λ_0 近似相等且远大于 λ_2 ,故其PCA2值接近于0.5,在整个单木结构中最大;对于3个特征值大小均近似相等的叶片点,其PCA2值则位于0.3左右,处于中间水平。

垂直度 (Verticality) 可以突出单木主干的垂直生长结构,以更好地区分树干与其他组件。具体计算公式如下:

$$\text{Verticality} = 1 - |\mathbf{n}_z| \quad (4)$$

式中, $\mathbf{n}_z$ 为当前点的法向量在 z 方向上的分量,垂直度的取值范围在 $[0, 1]$ 。

根据前面3个特征:显著性差异指标;主成分指标;垂直度指标在单木点云中不同位置的差异性详见图1。可见:单木点云的SoD值在细小的高级分支组件处较高,在叶片组分和第一主枝处则相对较低;在整个单木结构中,具有垂直生长结构的第一主枝其垂直度较大,接近于1,而其他组分的垂直度则较低。

2.3 数据增强

点云数据通常具有较高的稀疏性和不规则性,使深度学习模型容易受训练数据偏差的影响。通过对点云数据进行旋转、缩放、平移等数据增强操作,可以增加样本多样性,帮助模型提升泛化能力,同时缓解不同类别数据分布不均的问题,避免模型对某些类别的分割效果较差。此外,数据增强还能模拟噪声和遮挡等现实情况,提高模型鲁棒性,使其在多树种情况下表现得更加稳定。

本研究中应用的数据增强操作如下:

(1) 随机缩放。选取范围 $[0.9, 1.1]$ 的随机缩放因子 scale , 将其与原始点云数据的坐标 $\mathbf{p}$ 相乘,得到缩放后的坐标 $\mathbf{p}_s$, 其中各坐标轴可独立缩放,即取不同的随机缩放因子。具体计算公式如下:

$$\mathbf{p}_s = \mathbf{p} \times \text{scale} \quad (5)$$

(2) 点云对齐。对原始点云 $\mathbf{p}$ 进行整体平移,使点云重心平移至原点,并在 z 方向上二次平移,使点云在 z 轴上的最小值为零,得到对齐后的点云 $\mathbf{p}_a$ 。

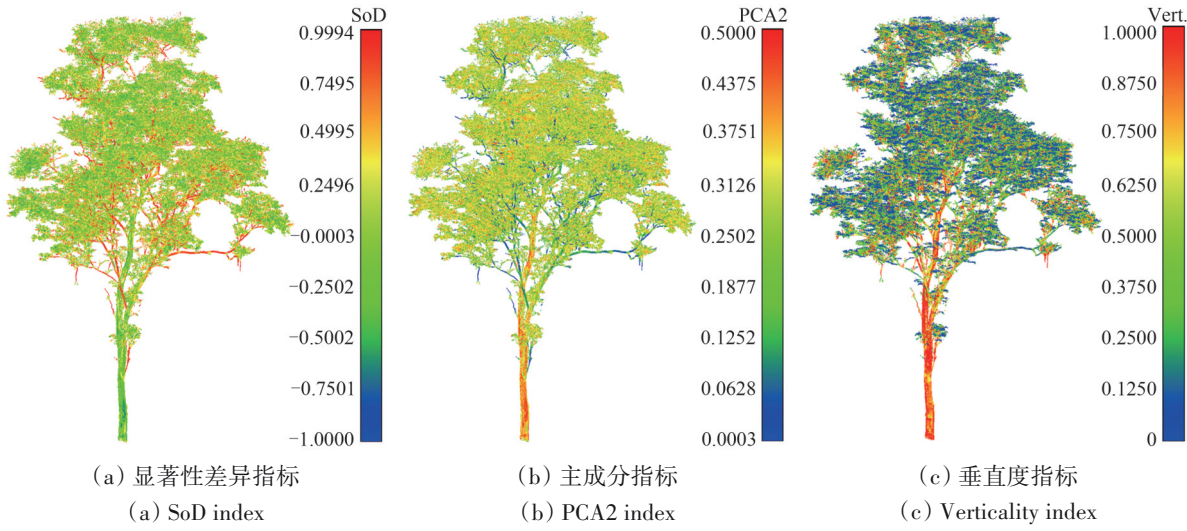


图1 单木先验几何结构特征值示意图

Fig. 1 Schematic diagram of eigenvalues of the prior geometric structure of an individual tree

$$\begin{cases} \mathbf{p}_a = \mathbf{p} - \text{mean}(\mathbf{p}) \\ \mathbf{p}_{a,z} = \mathbf{p}_a \cdot \mathbf{z} - \min(\mathbf{p}_a \cdot \mathbf{z}) \end{cases} \quad (6)$$

(3) 随机旋转。取范围在 $[-\pi, \pi]$ 内的随机旋转角度 θ ，将原始点云 $\mathbf{p}$ 绕 z 轴按 θ 进行旋转，得到旋转后的点云 $\mathbf{p}_r$ 。

$$\mathbf{p}_r = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{p} \quad (7)$$

(4) 随机噪声。随机生成一个与原始点云 $\mathbf{p}$ 的形状相同的噪声张量 noise ，并将其与 $\mathbf{p}$ 相加得到添加噪声后的点云 $\mathbf{p}_j$ 。其满足标准正态分布，并由标准差 σ 进行控制，且通过幅度阈值 c 将噪声最大值限制在 $[-c, c]$ 。

$$\mathbf{p}_j = \mathbf{p} + \text{clip}(N(0, \sigma^2), -c, c) \quad (8)$$

(5) 随机失活。根据预先设置的丢弃率 (ratio)，随机在原始输入数据中丢弃一部分点。

(6) 随机反转。生成一个 $[0, 1]$ 的随机种子，若随机种子 < 0.5 ，则对原始点云 $\mathbf{p}$ 进行坐标反转，反转后的点云坐标为 $\mathbf{p}_f$ 。

$$\mathbf{p}_f = -\mathbf{p} \quad (9)$$

(7) 特征归一化。将点云的原始特征 $\mathbf{f}$ 缩放到 $[-1, 1]$ 的范围内，得到归一化后的特征 $\mathbf{f}_n$ 。

$$\mathbf{f}_n = (\mathbf{f} - \mathbf{f}_{\min}) / (\mathbf{f}_{\max} - \mathbf{f}_{\min}) \quad (10)$$

3 单木组件分割网络结构与模块

本研究将单木组件划分为主干、分支和叶片3个类别，为更好提取点云局部特征，选取当前最前沿的 Point Transformer-V3 对木质与叶部结构进行

精确分割 (Wu 等, 2024)。该网络总体框架采用经典的 U-Net 框架模式，由编码器 (Encoder) 和解码器 (Decoder) 组成。在编码器中，原始点云及特征张量在第一层首先输入嵌入模块 (Embedding Block) 中，进行特征升维，然后输入若干层序列注意力机制模块 (Serialized Attention Block)，以进行高维特征的学习与提取；而对于第一层外的其他编码层，则由一个序列下采样层 (Serialized Down Pooling) 和若干序列注意力机制模块构成，以获得不同尺度上的点云特征。在解码阶段，每一层由一个序列化上采样层 (Serialized Up Pooling) 和若干注意力模块构成，以将高维特征还原到原始点云维度，最终输出各点的类别得分。网络整体结构及各模块组成如图2所示。

3.1 点云序列化策略

为了充分利用结构化数据的简便性和高效性，本研究将 Point Transformer-V3 网络引入了点云序列化策略，将非结构化点云转换为结构化数据。

空间填充曲线是一种路径，能够穿过高维离散空间中的每一个点，并在一定程度上保留其空间相邻关系。其在数学上可被定义为一个双射函数 $\varphi: \mathbb{Z} \mapsto \mathbb{Z}^n$ ，其中， n 为空间维度，对于点云其取值为 3，且可扩展到更高维度。本网络主要选取 Z-order 和 Hilbert 曲线这两种典型空间填充曲线。标准的空间填充曲线通过沿 x 、 y 、 z 轴依次遍历来处理 3D 空间。然而，为了能够提供关于空间关系的替代视角，可能捕捉到被标准曲线忽视的特殊局部关系。

通过改变遍历顺序, 本研究引入了标准空间填充曲线的重排序变体: Trans Z-order 和 Trans Hilbert

曲线 (Wu 等, 2024)。

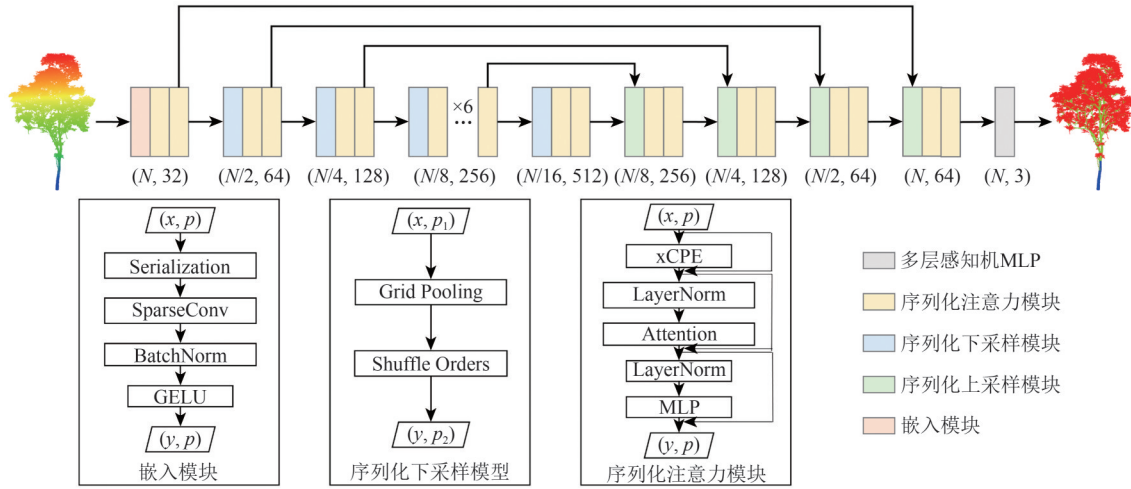


图2 Point Transformer-V3 总体网络框架与各模块组成

Fig. 2 Overall network framework and module composition of Point Transformer-V3

为了利用空间填充曲线的局部保留特性, 网络采用了序列化编码策略, 将点的位置转换为反映其在给定空间填充曲线中顺序的整数序列。空间填充曲线是一种双射函数, 其存在一个逆映射函数 $\varphi^{-1}: \mathbb{Z}^n \mapsto \mathbb{Z}$ 将一个点的位置信息 $p_i \in \mathbb{R}^3$ 转化为序列编码。通过将点的位置投影到格网尺寸为 $g \in \mathbb{R}$ 的离散空间中, 可以得到该点的序列编码为 $\varphi^{-1}(p/g)$ 。这种编码方法也适用于批量点云数据: 为每个点分配一个 64 位整数来记录序列化代码, 其中将后续 k 位分配给由 φ^{-1} 编码后得到的序列位置, 而剩余高位则分配给批次 (Batch) 的索引 $b \in \mathbb{Z}$ 。根据以上序列化编码方式对点进行排序, 即可使得每个批次内点云按照选定空间填充曲线模式有序排列。整个编码过程可以表示为

$$\text{Encode}(p, b, g) = (b \ll k) \mid \varphi^{-1}(p/g) \quad (11)$$

式中, “ $\ll$ ” 表示位运算中的左移位, 而 “ $\mid$ ” 表示位或操作。排序后对点集进行了重新排列, 使其符合由给定空间填充曲线定义的空间序列, 表明数据结构中邻近点很可能在实际三维空间上也很接近。

3.2 序列化注意力机制

为了充分利用点云序列化后的结构化特性, Point Transformer-V3 网络定义了基于序列化点云的补丁注意力机制 (Patch Attention)。该机制将点云中的点分组为不重叠的补丁, 并在每个补丁内执行注意力机制计算, 其有效性依赖于两个主要设计: 补丁分组 (Patch Grouping) 和补丁交互 (Patch

Interaction)。

补丁分组涉及对点集按序列化顺序进行简单分组, 并进行填充, 其过程如图3所示。实际上, 重排序和补丁填充的过程可以集成到一个单一索引操作中。此外, 与本研究中的序列化模式伴生设计的补丁分组策略有效地扩大了注意力机制在 3D 空间中的感受野, 同时保持了可行的空间邻接关系。

来自不同补丁点之间的交互对于网络模型整合输入点云的信息至关重要。因此, 设计合理的补丁交互策略对于克服非重叠结构的局限性具有重要意义, 并在补丁注意力机制功能实现中发挥着关键作用。基于此, 如图4所示, 本网络定义了如下4种交互策略:

(1) 位移扩张 (Shift dilation)。补丁分组通过特定步长跨越序列化点云以进行错位, 从而有效地扩展模型的感受野。

(2) 位移补丁 (Shift patch)。补丁的位置沿序列化点云进行偏移, 类似于图像滤波器中的滑动窗口策略。这种方法最大化了不同补丁之间的交互。

(3) 位移顺序 (Shift order)。点云数据的序列化顺序在每个注意力块之间动态变化。这种策略有助于防止模型对单一模式的过拟合, 并促进了数据特征整合过程中的鲁棒性。

(4) 打乱顺序 (Shuffle order)。在位移顺序策略的基础上引入了随机打乱策略, 以对序列化顺序进行置换。该策略确保每个注意力层的感受野不受限于单一模式, 从而进一步增强了模型的泛化能力。

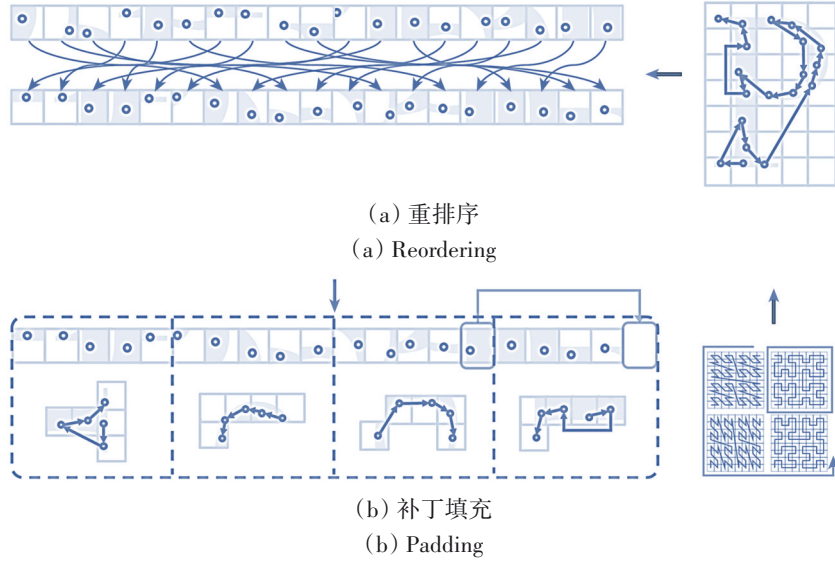


图3 补丁分组实现过程 (Wu 等, 2024)

Fig. 3 Implementation process of Patch Grouping (Wu et al., 2024)

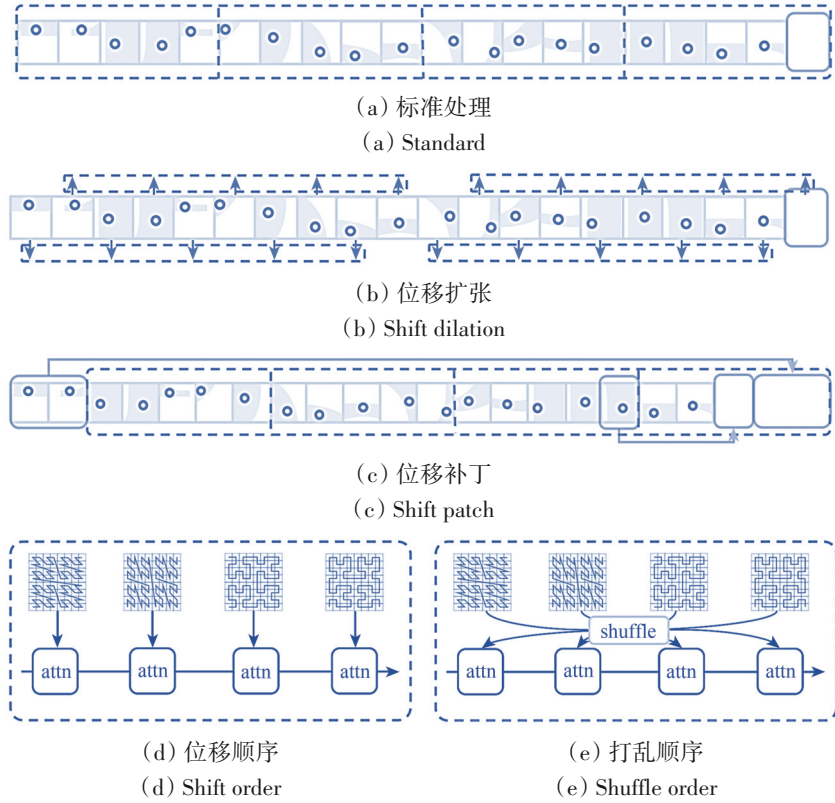


图4 不同补丁交互策略实现方法 (Wu 等, 2024)

Fig. 4 Different patch interaction strategy implementation methods (Wu et al., 2024)

3.3 池化策略

Point Transformer-V3 网络采用格网池化 (Grid Pooling) 策略来进行点云特征的上、下采样。给定一组点集 $\mathbf{M} = (\mathbf{P}, \mathbf{F})$, 通过将三维空间分割成多个不重叠的均匀格网, 可将点集 $\mathbf{M}$ 分成若干子集 $[\mathbf{M}_1, \mathbf{M}_2, \dots, \mathbf{M}_n]$, 再将每个子点集 $\mathbf{M}_i = (\mathbf{P}_i,$

$\mathbf{F}_i)$ 进行如下融合:

$$\begin{aligned} \mathbf{f}'_i &= \text{MaxPool}(\{f_j U \mid f_j \in \mathbf{F}_i\}), \mathbf{p}'_i = \\ &\text{MeanPool}(\{p_j \mid p_j \in \mathbf{P}_i\}) \end{aligned} \quad (12)$$

式中, $(\mathbf{p}'_i, \mathbf{f}'_i)$ 分别为子集 $\mathbf{M}_i$ 经过聚合后得到的池化点位置信息和特征信息, $U \in \mathbb{R}^{c \times c'}$ 为线性投影。通过收集来自 n 个子集的池化点, 可以得到下采样

后的点集。

对于点云张量的上采样,则采用反池化(Up Pooling)操作实现。为了将下采样后的点集 M' 变换回原始点集 M ,可以记录池化过程中点的位置,且只需获取点集 M 中各点的特征信息即可。借助池化阶段中基于网格的分区 $[M_1, M_2, \dots, M_n]$,可以将点特征映射到同一子集中的所有点:

$$f_i^{\text{up}} = f_j', \quad \text{if } (p_i, f_i) \in M_j \quad (13)$$

3.4 位置编码

Point Transformer-V3网络引入条件位置编码CPE(Conditional Positional Encoding)来替代相对低效且复杂的相对位置编码RPE(Relative Positional Encoding),并提出了一种在注意力层之前直接添加一个带有跳跃连接(Skip Connection)的稀疏卷积层(Sparse Convolution)的增强型条件位置编码(xCPE),以提升点云位置信息的提取性能,其结构如图5所示。

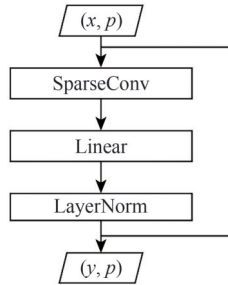


图5 xCPE模块组成结构

Fig. 5 The composition structure of the xCPE module (Wu等, 2024)

3.5 精度评定指标

本研究采用深度学习语义分割研究中广泛应用的各类别交并比IoU (Intersection over Union)、平均交并比mIoU (Mean Intersection over Union),以及总体准确率OA (Overall Accuracy)和平均准确率mAcc (Mean Accuracy)等指标来评价网络的单木组件分割精度,如式(14) — (17)所示。

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \quad (14)$$

$$mAcc = \frac{1}{n} \sum_{i=1}^n \frac{TP_i}{TP_i + FN_i} \quad (15)$$

$$IoU = \frac{TP}{TP + FP + FN} \quad (16)$$

$$mIoU = \frac{1}{n} \sum_{i=1}^n \frac{TP_i}{TP_i + FP_i + FN_i} \quad (17)$$

式中, TP (True Positive) 表示被模型正确预测为正样本的点数; TN (True Negative) 表示被模型正确预测为负样本的点数; FP (False Positive) 表示被模型错误预测为正样本的点数; FN (True Negative) 表示被模型错误预测为负样本的点数; n 为分割的总类别数。

4 实验设计与结果分析

4.1 实验环境及训练超参数设置

本次单木组件分割实验在1块RTX 3060 GPU上进行,其显存为12 GB。将ITS-3D数据集中包含的713棵单木点云,按标准的6:2:2的比例将其按不同子集、不同树种分别均匀的划分为训练集(427棵)、验证集(143棵)和测试集(143棵),分割类别为主干、分支和叶片3类。

在网络训练过程中,本研究使用的超参数设置如下:对原始点云数据进行了格网下采样,格网尺寸为0.04 m;且,对训练集输入数据进行了裁剪以进一步降低数据体量,其数据保留率设为0.6,单个输入点集的最大点数设为81920;由于GPU显存限制,且对单块输入点集保留了较多的点,训练的批大小仅设为2,而验证与测试的批大小默认为1。总共训练轮数为100,其中每轮训练循环次数(loop)为30;使用的优化器为AdamW,学习率为0.0001,权重衰减为0.05;学习率调度器选用OneCycleLR,退火策略设置为余弦退火;损失函数采用CrossEntropy和Lovasz。

对于网络超参数,编码阶段总共设置了5层,各层的注意力模块数量为[2, 2, 2, 6, 2],各层的下采样步长为[2, 2, 2, 2],各层的输出特征维度为[32, 64, 128, 256, 512];解码阶段共设置了4层,各层的注意力模块数量为[2, 2, 2, 2],各层的输出特征维度为[64, 64, 128, 256],其余模型参数使用Point Transformer-V3网络的默认设置(Wu等, 2024)。

4.2 分割结果及分析

利用Point Transformer-V3网络对ITS-3D测试集中的143棵单木样本进行类别预测后,选取了ITS-3D数据集中各子集(前8个为TLS数据子集,第9个为UAV数据子集)的典型数据来进行分割结果的可视化如图6所示,定量评价结果如表1所

示。可见该模型在本数据集的各子集上均呈现出良好的单木组件分割效果。不同子集、不同树种的单木样本之间的总体几何性状存在较大差异，但本研究提出的通过引入构建的先验几何特征用于 Point Transformer-V3 网络较好地应对了这一问题，可以将其推广到大多数形态结构不同的单木数据

上，具有较高的普适性。此外，由图6可见，该网络在木质组分与叶部组分的交界区域均能达到较好的分割效果，准确分离出了木质点与叶片点；在深入树冠区域，传统算法难以识别且被密集叶片点覆盖的高级分支也被 Point Transformer-V3 网络模型较好的提取。

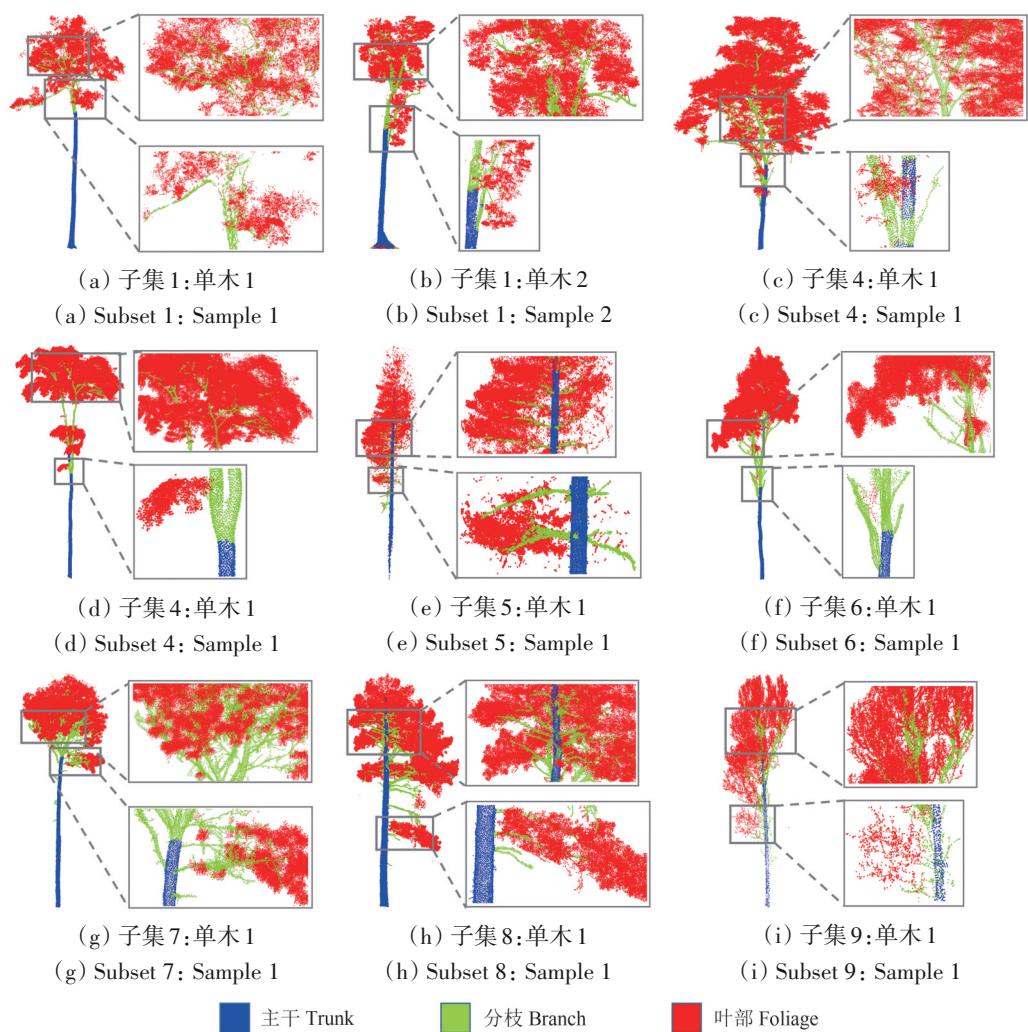


图6 Point Transformer-V3在各子集中的单木组件分割结果
Fig. 6 Individual tree component segmentation result in each subset by using Point Transformer-V3

表1 基于 Point Transformer-V3 的 ITS-3D 数据集分割精度
Table 1 Segment accuracy of ITS-3D based on Point Transformer-V3

数据集	总体精度 OA	平均精度 mAcc	平均交并比 mIoU	主干 Trunk	分枝 Branch	叶部 Foliage
子集 1	0.959	0.921	0.865	0.966	0.687	0.943
子集 4	0.954	0.838	0.788	0.957	0.457	0.951
子集 5	0.928	0.872	0.775	0.945	0.473	0.906
子集 6	0.974	0.806	0.746	0.862	0.402	0.975
子集 7	0.884	0.808	0.737	0.936	0.403	0.871
子集 8	0.957	0.907	0.867	0.902	0.741	0.959
子集 9 (UAV)	0.885	0.725	0.609	0.564	0.367	0.897
完整 ITS-3D 数据集	0.946	0.872	0.806	0.932	0.546	0.940

与此同时,为了更好的评估Point Transformer-V3网络的单木组件分割效果,本研究对比了其分割结果与真实标签,结果如图7所示。可见,网络的分割结果与真实值非常接近,在整体上较好的将主干、分支以及叶片进行了区分;而对于局部分割

结果,其分割细节也较为可观,一些细小的分支也能被较好的提取出来。然而,网络在局部叶片覆盖较密的情况下,仍然会存在木质点云被误分为叶片的问题。

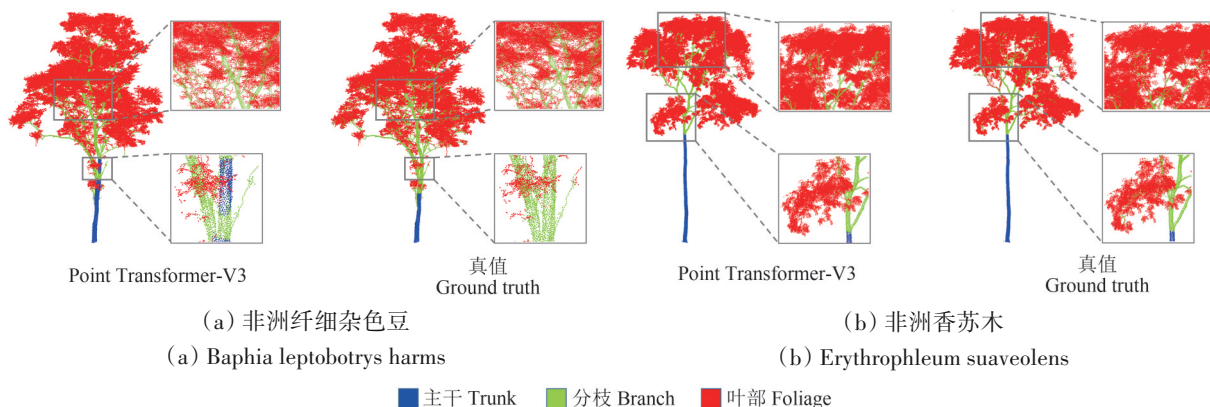


图7 Point Transformer-V3单木组件分割结果与语义标签真值对比

Fig. 7 Comparison of the segmentation results of individual tree components with the true values of semantic labels by using Point Transformer-V3

此外,为了量化的评价Point Transformer-V3网络在ITS-3D数据集上的单木组件分割性能,文中单木组件分割的OA、mAcc和mIoU分别达到了0.946、0.872与0.806,枝干、分支以及叶片类别的IoU指标则分别为0.932、0.546和0.940。实验结果说明该网络对于枝干类别和叶片类别能够有效的分割出来,而对于分支点的分割效果则相对较差,这主要是因为对应的细支在叶片中间,且叶片直接覆盖在细支上,不管是从空间上还是尺度上,都与叶片紧密混合在一起,从而导致提取的特征很难将分枝末端再进一步有效分离。此外,由表1可知在ITS-3D的各子集上也分别取得了较好的分割精度,所有TLS子集的OA、mAcc和mIoU都位于0.928、0.806和0.737以上;UAV的子集9的分割精度则相对较低,其OA、mAcc与mIoU为0.885、0.725和0.609。这可能与UAV点云的点密度相对较低有关,特别是冠层下部以及近地面的枝干点和分支点由于冠层的遮挡导致更加稀疏,从而致使其对这2个类别的分割效果较差。随后,为了进一步探究不同类别分割的精度细节,计算其归一化混淆矩阵(图8)。结果表明Point Transformer-V3网络对于枝干类别和叶片类别都实现了较好的区分,其混淆矩阵对角线上的数值分别到达了0.975和0.974,证明这两类点云被误分为其他类别的情况较少,

分类准确率较高。然而,对于分支类别,其混淆矩阵对角线上的数值仅为0.666,证明分支类别的点云容易被误判为其他类别。当真值标签为分支类别时,其被误判为枝干点的比率仅为0.029,而误判为叶片点的比率却达到了0.305,说明分支点被误判为叶片点的情况较为严重,这也进一步说明了叶片和细支分布在树干上部,两者相互混合在一起,且树冠上层的细小分支容易被大片叶片点包裹,从而在计算先验几何特征过程中,不管从空间还是尺度上都难以有效获取对应位置的特性,进而导致这部分细小分支难以被有效地提取。后续可以考虑采用分步提取算法,设计不同尺度来进行分割。此外,后续为了能够进一步分析构建的先验几何特征以及模型性能,可以探讨各个树种的枝叶分离效果以及模型的鲁棒性。

4.3 不同单木组件分割算法之间的对比实验

为了对比分析Point Transformer-V3网络与传统算法分割性能,分析了GBS、LeWoS算法的分割性能。然而,由于传统算法在进行枝叶分离过程中,无法进一步区分树木的主干与精细分支结构。因此,在与传统算法的对比实验中仅对木质类别与叶片类别两类进行分割性能评估,实验结果如图9和表2所示。可见,在枝叶分割细节上,Point Transformer-

V3模型对细小分支及各分支的末端提取效果更为完整，且各项定量评价指标都优于传统算法。

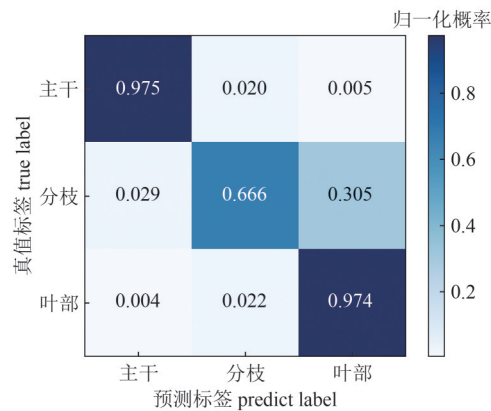


图8 单木组件分割结果混淆矩阵
Fig. 8 Confusion matrix of the individual component segmentation

为了进一步分析 Point Transformer-V3 网络分割性能，选取了近年来主流的点云语义分割网络 PointNet++ (Qi等，2017)，DGCNN (Wang等，2019)，

Point Cloud Transformer (Guo等，2021)，Stratified Transformer (Lai等，2022)，PointNeXt (Qian等，2022)，PointVector (Deng等，2023)，OctFormer (Wang，2023)，SparseUNet (Wu等，2023)，OA-CNNs (Peng等，2024)，Point Cloud Mamba (Zhang等，2024b)，在 ITS-3D 数据集上进行测试。为了保证对比实验的公平性，以上常规网络的训练超参数均与 Point Transformer-V3 网络的超参数保持一致，各对比网络的训练超参数设置如下：下采样网格尺寸 (grid size) 为 0.04 m；球裁剪 (sphere crop) 的数据保留率设为 0.6，单个输入点集的最大点数 (max points) 设为 81920；训练的批大小 (batch size) 设为 2，训练轮数 (epoch) 为 100 轮，其中每轮的训练循环次数 (loop) 为 30 次。对于各网络的结构超参数，本研究都以其原论文提及的性能最优的普适性参数进行设置，以确保各对比网络能获取最好的分割结果。其可视化结果与量化的评价分别如图 10 和表 3 所示。

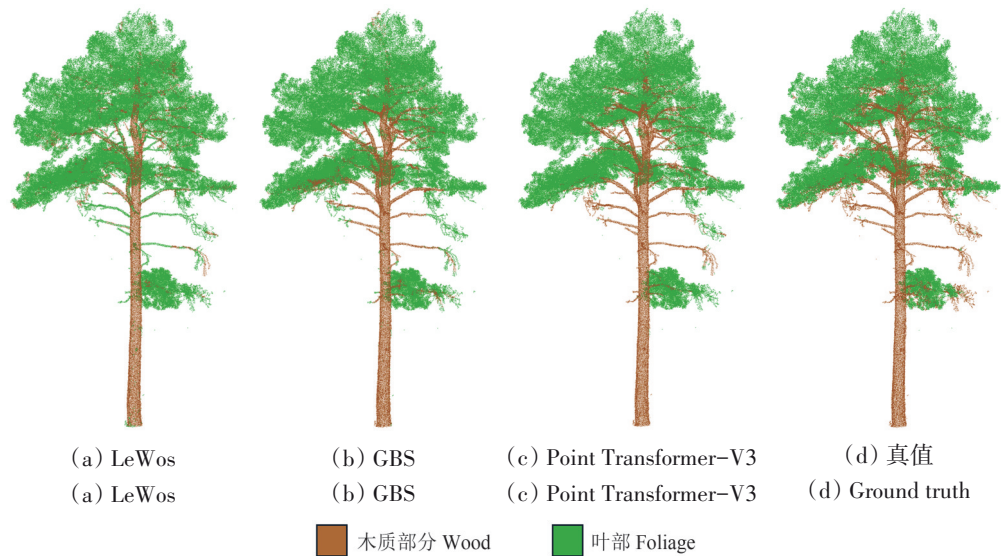


图9 本研究算法与传统单木组件算法分割效果对比
Fig. 9 Comparison of individual-tree component segmentation between the deep learning and the traditional algorithm

表2 本研究算法与传统单木组件分割算法精度对比

Table 2 Comparison of segmentation accuracy of between the deep learning and the traditional algorithm					
对比方法	总体精度 OA	平均精度 mAcc	平均交并比 mIoU	木质部分 Wood	叶部 Foliage
LeWoS	0.873	0.782	0.684	0.514	0.854
GBS	0.871	0.818	0.695	0.542	0.847
Point Transformer V3	0.952	0.922	0.867	0.794	0.940

注：加粗数值表示各对比算法中该指标的最优结果。

根据图 10 可知，Point Transformer-V3 网络较好的提取了冠层上部被叶片覆盖的小尺度高级分

支，且能在一定程度上区分出一级主干与其他高级分支，分割性能较高；Octformer 网络也具有较好

的分割细节,能准确区分木质与叶片组分;此外,SparseUNet与OA-CNNs等网络的整体分割效果也

较好,但其对于主干和高级分支区分能力欠佳;其余网络则均存在不同程度的叶片点过分割现象。

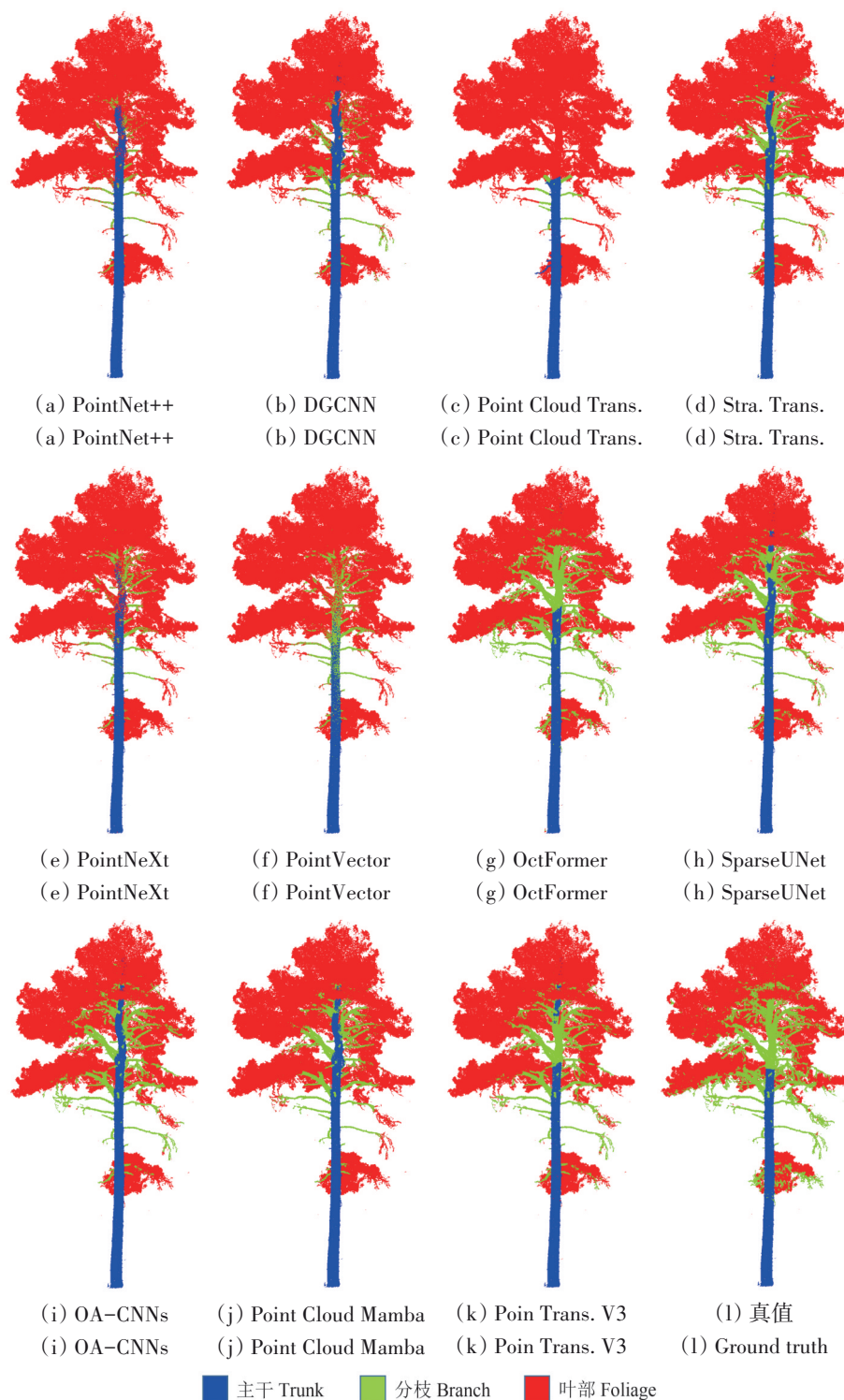


图 10 不同深度学习网络单木组件分割效果对比

Fig. 10 Comparison of individual-tree component segmentation by different deep learning networks

由表3可以发现,在其他对比网络中,OctFormer和SparseUNet也均具有较高的整体分割精度,其mIoU分别达到0.802和0.799,其中OctFormer的

OA指标还达到了与Point Transformer-V3网络相同的水平。而对于各类别的分割精度,研究中使用的Point Transformer-V3网络在主干、分支和叶片类

别的IoU分别为0.932、0.546和0.940，在所有对比网络中均达到第二名的水平，但与最高水平均只存在0.1%—0.2%的差距。对于主干类别，SparseUNet取得了最高的分割精度，其IoU为0.934；而对于分支和叶片类别，则由OctFormer取得了最高分割精度，其IoU分别为0.547和0.941。

表3 不同深度学习网络单木组件分割精度对比

Table 3 Comparison of segmentation accuracy of individual components in different deep learning networks						
对比方法	总体精度 OA	平均精度 mAcc	平均交并比 mIoU	主干 Trunk	分枝 Branch	叶部 Foliage
PointNet++	0.898	0.685	0.628	0.753	0.229	0.903
DGCNN	0.907	0.708	0.650	0.835	0.206	0.909
Point Cloud Transformer	0.818	0.426	0.363	0.273	0.002	0.813
Stratified Transformer	0.925	0.770	0.715	0.899	0.325	0.920
PointNeXt	0.899	0.705	0.641	0.707	0.316	0.900
PointVector	0.886	0.661	0.594	0.635	0.245	0.902
OctFormer	0.946	0.868	0.802	0.918	0.547	0.941
SparseUNet	0.944	0.864	0.799	0.934	0.525	0.937
OA-CNNs	0.942	0.853	0.791	0.932	0.505	0.935
Point Cloud Mamba	0.905	0.694	0.641	0.784	0.234	0.905
Point Transformer V3	0.946	0.872	0.806	0.932	0.546	0.940

注：加粗数值表示各对比算法中该指标的最优结果。

在实验中，注重于提取输入点云局部特征的网络都取得了较好的分割精度，如 Point Transformer-V3、OctFormer等；而对于那些注重全局特征学习的网络，如 Point Cloud Mamba、Point Cloud Transformer等，其分割精度却并不理想。在单木组件分割任务中，分割困难点在于如何区分冠层内部的叶片点和被叶片覆盖的小尺度的高级分支点，这种局部区域的精细分割需要对点云的局部特征进行更好的学习和提取，因此局部特征提取器能表现出更好的性能；而对于全局特征提取器，如 Mamba 的 State space model，其可能无法在该分割任务中发挥其最大的作用。

4.4 消融实验

为了验证在网络训练过程中加入单木先验几何特征参数的有效性，设计了一组减法消融实验，在网络训练过程中保持其他环境不变，仅改变输入特征，包括仅用3维坐标作为输入、以及分别舍去3个特征参数中的一个后，将3维坐标和剩余特征作为输入。各组实验的分割精度如表4所示。

表4表明，当输入所有单木先验几何特征时，分割结果的OA和mIoU分别为94.63%和80.59%，达到了最高水平。当仅输入点云的3维坐标时，相比输入所有特征下降了0.29%和0.42%，分割精度最低；而当仅去除SoD特征时，其OA和mIoU分别

为94.48%和80.49%，相比输入所有特征下降了0.15%和0.10%；当仅去除PCA2特征时，其OA和mIoU分别为94.45%和80.33%，相比输入所有特征下降了0.18%和0.26%；当仅去除Verticality特征时，其OA和mIoU分别为94.48%和80.36%，相比输入所有特征下降了0.15%和0.23%。因此，结果表明输入构建的3个单木先验几何特征的有效性。

表4 不同输入特征参数情况下的分割精度对比

Table 4 Comparison of segmentation accuracy under different input feature parameters			/%
输入特征	总体精度 OA	平均交并比 mIoU	
仅3维坐标	94.34	80.17	
去除显著性差异指标(SoD)	94.48	80.49	
去除主成分指标(PCA2)	94.45	80.33	
去除垂直度指标(VERTICALITY)	94.48	80.36	
所有特征	94.63	80.59	

注：加粗表示最优精度。

5 结 论

本研究针对单木组件分割算法对不同树种的样本普适性较差，细小枝干分割能力较弱等情况，通过构建先验几何特征优化 Point Transformer-V3 深度学习网络，对单木组件（主干、分支和叶片）开展分割研究，并在 ITS-3D 数据集与传统单木组

件分割算法进行实验验证对比。主要结论如下：(1) 引入构建的先验几何特征的深度学习网络可以深入树冠较小尺度的高级分支也能实现一定提取，且其不同树种几何形态结构差异较大的单木上也表现出较高的适用性；(2) 相比于其他传统算法，Point Transformer-V3网络在单木组件分割应用中具有更高的稳定性能；(3) 通过消融实验表明，在训练Point Transformer-V3网络时引入有效的单木先验特征能更好的促进网络训练实现对于单木组件的分割效果。

参考文献(Reference)

- Anonymous. 2024. Plot-level semantically labelled terrestrial laser scanning point clouds. Zenodo [DOI: 10.5281/zenodo.13268500]
- Arrizza S, Marras S, Ferrara R and Pellizzaro G. 2024. Terrestrial Laser Scanning (TLS) for tree structure studies: a review of methods for wood-leaf classifications from 3D point clouds. *Remote Sensing Applications: Society and Environment*, 36: 101364 [DOI: 10.1016/j.rsase.2024.101364]
- Calders K, Disney M I, Armston J, Burt A, Brede B, Origo N, Muir J and Nightingale J. 2017. Evaluation of the range accuracy and the radiometric calibration of multiple terrestrial laser scanning instruments for data interoperability. *IEEE Transactions on Geoscience and Remote Sensing*, 55(5): 2716-2724 [DOI: 10.1109/TGRS.2017.2652721]
- Chen S L, Liu H Y, Feng Z K, Shen C Y and Chen P P. 2019. Applicability of personal laser scanning in forestry inventory. *PLoS One*, 14(2): e0211392 [DOI: 10.1371/journal.pone.0211392]
- Chen S L, Verbeeck H, Terryn L, Van den Broeck W A J, Vicari M B, Disney M, Origo N, Wang D, Xi Z X, Hopkinson C, Dai W X, Wang M L, Moorthy S M K, Shao J, Ferrara R, MacFarlane D W and Calders K. 2025. The impact of leaf-wood separation algorithms on aboveground biomass estimation from terrestrial laser scanning. *Remote Sensing of Environment*, 318: 114581 [DOI: 10.1016/j.rse.2024.114581]
- Dai W X, Jiang Y H, Zeng W, Chen R B, Xu Y Y, Zhu N N, Xiao W, Dong Z and Guan Q F. 2023. MDC-Net: a multi-directional constrained and prior assisted neural network for wood and leaf separation from terrestrial laser scanning. *International Journal of Digital Earth*, 16(1): 1224-1245 [DOI: 10.1080/17538947.2023.2198261]
- Danson F M, Gaulton R, Armitage R P, Disney M, Gunawan O, Lewis P, Pearson G and Ramirez A F. 2014. Developing a dual-wavelength full-waveform terrestrial laser scanner to characterize forest canopy structure. *Agricultural and Forest Meteorology*, 198-199: 7-14 [DOI: 10.1016/j.agrformet.2014.07.007]
- Demol M, Verbeeck H, Gielen B, Armston J, Burt A, Disney M, Duncanson L, Hackenberg J, K  kenbrink D, Lau A, Ploton P, Sewdinen A, Stovall A, Takoudjou S M, Volkova L, Weston C, Wortel V and Calders K. 2022. Estimating forest above-ground biomass with terrestrial laser scanning: current status and future directions. *Methods in Ecology and Evolution*, 13(8): 1628-1639 [DOI: 10.1111/2041-210X.13906]
- Deng X, Zhang W Y, Ding Q and Zhang X M. 2023. Pointvector: a vector representation in point cloud analysis//*Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver: IEEE: 9455-9465 [DOI: 10.1109/CVPR52729.2023.00912]
- Dong T Y, Zhang X P, Ding Z F and Fan J. 2020. Multi-layered tree crown extraction from LiDAR data using graph-based segmentation. *Computers and Electronics in Agriculture*, 170: 105213 [DOI: 10.1016/j.compag.2020.105213]
- Ferrara R, Virdis S G P, Ventura A, Ghisu T, Duce P and Pellizzaro G. 2018. An automated approach for wood-leaf separation from terrestrial LIDAR point clouds using the density based clustering algorithm DBSCAN. *Agricultural and Forest Meteorology*, 262: 434-444 [DOI: 10.1016/j.agrformet.2018.04.008]
- Guo M H, Cai J X, Liu Z N, Mu T J, Martin R R and Hu S M. 2021. PCT: point cloud transformer. *Computational Visual Media*, 7(2): 187-199 [DOI: 10.1007/s41095-021-0229-5]
- Han T and S  nchez-Azofeifa G A. 2022. A deep learning time series approach for leaf and wood classification from terrestrial LiDAR point clouds. *Remote Sensing*, 14(13): 3157 [DOI: 10.1007/s41095-021-0229-5]
- Hui Z Y, Jin S G, Xia Y P, Wang L Y, Ziggah Y Y and Cheng P G. 2021. Wood and leaf separation from terrestrial LiDAR point clouds based on mode points evolution. *ISPRS Journal of Photogrammetry and Remote Sensing*, 178: 219-239 [DOI: 10.1016/j.isprsjprs.2021.06.012]
- Kaasalainen S, Krooks A, Kukko A and Kaartinen H. 2009. Radiometric calibration of terrestrial laser scanners with external reference targets. *Remote Sensing*, 1(3): 144-158 [DOI: 10.3390/rs1030144]
- Lai X, Liu J H, Jiang L, Wang L W, Zhao H S, Liu S, Qi X J and Jia J Y. 2022. Stratified transformer for 3D point cloud segmentation//*Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans: IEEE: 8490-8499 [DOI: 10.1109/cvpr52688.2022.00831]
- Li Z, Douglas E, Strahler A, Schaaf C, Yang X Y, Wang Z S, Yao T, Zhao F, Saenz E J, Paynter I, Woodcock C E, Chakrabarti S, Cook T, Martel J, Howe G, Jupp D L B, Culvenor D S, Newnham G J and Lovell J L. 2013. Separating leaves from trunks and branches with dual-wavelength terrestrial LiDAR scanning//*2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS*. Melbourne: IEEE: 3383-3386 [DOI: 10.1109/IGARSS.2013.6723554]
- Liu Z C, Zhang Q, Wang P, Li Z and Wang H R. 2020. Automated classification of stems and leaves of potted plants based on point cloud data. *Biosystems Engineering*, 200: 215-230 [DOI: 10.1016/j.biosystemseng.2020.10.006]
- Moorthy S M K, Calders K, Vicari M B and Verbeeck H. 2020. Improved supervised learning-based approach for leaf and wood classification from LiDAR point clouds of forests. *IEEE Transac-*

- tions on Geoscience and Remote Sensing, 58(5): 3057-3070 [DOI: 10.1109/TGRS.2019.2947198]
- Morel J, Bac A and Kanai T. 2020. Segmentation of unbalanced and inhomogeneous point clouds and its application to 3D scanned trees. *The Visual Computer*, 36(10): 2419-2431 [DOI: 10.1007/s00371-020-01966-7]
- Liang X L, Qi H W, Deng X J, Chen J C, Cai S S, Zhang Q J, Wang Y S, Kukko A, Hyyppä J.. 2025. ForestSemantic: a dataset for semantic learning of forest from close-range sensing. *Geo-Spatial Information Science*, 28(1), 185-211 [DOI: 10.1080/10095020.2024.2313325]
- Oehmcke S, Li L, Trepekli K, Revenga J C, Nord-Larsen T, Gieseke F and Igel C. 2024. Deep point cloud regression for above-ground forest biomass estimation from airborne LiDAR. *Remote Sensing of Environment*, 302: 113968 [DOI: 10.1016/j.rse.2023.113968]
- Peng B H, Wu X Y, Jiang L, Chen Y K, Zhao H S, Tian Z T and Jia J Y. 2024. OA-CNNs: omni-adaptive sparse CNNs for 3D semantic segmentation//*Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE: 21305-21315 [DOI: 10.1109/CVPR52733.2024.02013]
- Puliti S, Pearce G, Surový P, Wallace L, Hollaus M, Wielgosz M and Astrup R. 2023. FOR-instance: a UAV laser scanning benchmark dataset for semantic and instance segmentation of individual trees. *Zenodo* [DOI: 10.5281/zenodo.8287792]
- Qi C R, Yi L, Su H and Guibas L J. 2017. PointNet++: deep hierarchical feature learning on point sets in a metric space//*Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach: Curran Associates Inc.: 5105-5114
- Qian G C, Li Y C, Peng H W, Mai J J, Al Kader Hammoud H A, Elhoseiny M and Ghanem B. 2022. PointNeXt: revisiting PointNet++ with improved training and scaling strategies//*Proceedings of the 36th International Conference on Neural Information Processing Systems*. New Orleans: Curran Associates Inc.: 23192-23204
- Strimbu V F and Strimbu B M. 2015. A graph-based segmentation algorithm for tree crown extraction using airborne LiDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 104: 30-43 [DOI: 10.1016/j.isprsjprs.2015.01.018]
- Tao S L, Guo Q H, Xu S J, Su Y W, Li Y M and Wu F F. 2015. A geometric method for wood-leaf separation using terrestrial and simulated lidar data. *Photogrammetric Engineering and Remote Sensing*, 81(10): 767-776 [DOI: 10.14358/PERS.81.10.767]
- Terryn L, Calders K, Bartholomeus H, Bartolo R E, Brede B, D'Hont B, Disney M, Herold M, Lau A, Shenkin A, Whiteside T G, Wilkes P and Verbeeck H. 2022. Quantifying tropical forest structure through terrestrial and UAV laser scanning fusion in Australian rainforests. *Remote Sensing of Environment*, 271: 112912 [DOI: 10.1016/j.rse.2022.112912]
- Tian Z L and Li S H. 2022a. Graph-based leaf - wood separation method for individual trees using terrestrial lidar point clouds. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 5705111 [DOI: 10.1109/TGRS.2022.3218603]
- Tian Z L and Li S H. 2022b. Graph-based Leaf-Wood Separation Method for Individual Trees Using Terrestrial Lidar Point Clouds: Labeled validation data. *Zenodo* [DOI: 10.5281/zenodo.6831378]
- Vicari M B, Disney M, Wilkes P, Burt A, Calders K and Woodgate W. 2019. Leaf and wood classification framework for terrestrial LiDAR point clouds. *Methods in Ecology and Evolution*, 10(5): 680-694 [DOI: 10.1111/2041-210X.13144]
- Wan P, Zhang W M and Jin S N. 2021. Plot-level wood-leaf separation for terrestrial laser scanning point clouds. *Dryad* [DOI: 10.5061/dryad.rfj6q5799]
- Wang D, Hollaus M and Pfeifer N. 2017. Feasibility of machine learning methods for separating wood and leaf points from terrestrial laser scanning data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-2/W4: 157-164 [DOI: 10.5194/isprs-annals-iv-2-w4-157-2017]
- Wang D, Takoudjou S M and Casella E. 2020. LeWoS: a universal leaf-wood classification method to facilitate the 3D modelling of large tropical trees using terrestrial LiDAR. *Methods in Ecology and Evolution*, 11(3): 376-389 [DOI: 10.1111/2041-210X.13342]
- Wang D, Takoudjou S M and Casella E. 2021. LeWoS: a universal leaf-wood classification method to facilitate the 3D modelling of large tropical trees using terrestrial LiDAR. *Dryad* [DOI: 10.5061/dryad.np5hqbzp6]
- Wang P S. 2023. OctFormer: octree-based transformers for 3D point clouds. *ACM Transactions on Graphics*, 42(4): 155 [DOI: 10.1145/3592131]
- Wang Y, Sun Y B, Liu Z W, Sarma S E, Bronstein M M and Solomon J M. 2019. Dynamic graph cnn for learning on point clouds. *ACM Transactions on Graphics*, 38(5): 146 [DOI: 10.1145/3326362]
- Weiser H, Ulrich V, Winiwarter L, Esmoris A M and Höfle B. 2024. Manually labeled terrestrial laser scanning point clouds of individual trees for leaf-wood separation. V1 edn. *heiDATA* [DOI: 10.11588/DATA/UUMEDI]
- Wu B X, Zheng and Chen Y. 2020. An improved convolution neural network-based model for classifying foliage and woody components from terrestrial laser scanning data. *Remote Sensing*, 12(6): 1010 [DOI: 10.3390/rs12061010]
- Wu X Y, Jiang L, Wang P S, Liu Z J, Liu X H, Qiao Y, Ouyang W L, He T and Zhao H S. 2024. Point Transformer V3: simpler, faster, stronger//*Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE: 4840-4851 [DOI: 10.1109/cvpr52733.2024.00463]
- Wu X Y, Wen X, Liu X H and Zhao H S. 2023. Masked scene contrast: a scalable framework for unsupervised 3D representation learning//*Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver: IEEE: 9415-9424 [DOI: 10.1109/cvpr52729.2023.00908]
- Xi Z X, Chasmer L and Hopkinson C. 2023. Delineating and reconstructing 3D forest fuel components and volumes with terrestrial laser scanning. *Remote Sensing*, 15(19): 4778 [DOI: 10.3390/rs15194778]
- Xi Z X, Hopkinson C and Chasmer L. 2018. Filtering stems and branches from terrestrial laser scanning point clouds using deep 3-

- D fully convolutional networks. *Remote Sensing*, 10(8): 1215 [DOI: 10.3390/rs10081215]
- Xi Z X, Hopkinson C, Rood S B and Peddle D R. 2020. See the forest and the trees: effective machine and deep learning algorithms for wood filtering and tree species classification from terrestrial laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 168: 1-16 [DOI: 10.1016/j.isprsjprs.2020.08.001]
- Zhang S, Chen Y P, Wang B, Pan D, Zhang W M and Li A G. 2024a. SPTNet: sparse convolution and transformer network for woody and foliage components separation from point clouds. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5702718 [DOI: 10.1109/TGRS.2024.3376454]
- Zhang T, Yuan H B, Qi L, Zhang J N, Zhou Q Y, Ji S P, Yan S C and Li X T. 2024b. Point cloud mamba: point cloud learning via state space model. *arXiv preprint arXiv: 2403.00762* [DOI: 10.48550/arXiv.2403.00762]

Deep learning based on prior geometric features for the segmentation of individual tree point cloud components

GAN Ruilin^{1,2}, YANG Jian¹, LUO Binhuan^{1,3}, SHI Shuo⁴, DU Lin¹,
WU Zhongliang¹, WANG Sihao¹, WANG Ao⁴

1. School of Geography and Information Engineering, China University of Geosciences, Wuhan 430074, China;

2. Changjiang Geotechnical Engineering Co., Ltd., CISPDR Corporation, Wuhan, 430010, China;

3. Electronic Information School, Wuhan University, Wuhan 430072, China;

4. State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

Abstract: Forest aboveground biomass (AGB) is a crucial component of the global carbon cycle, playing a key role in monitoring global carbon dynamics and effectively mitigating climate change. The segmentation of individual tree components to distinguish wood and leaf structures effectively is an important foundation for accurately estimating the key structural parameters of trees and precisely inferring AGB. With advancements in technology, the advent of terrestrial LiDAR has provided a novel nondestructive method for estimating tree structural parameters and AGB. However, current algorithms for individual tree component segmentation that use point clouds exhibit limited universality across different tree species, and their capability to segment fine branches remains relatively constrained.

Therefore, this study constructs a large individual tree component segmentation dataset, namely, ITS-3D, which contains 713 tree samples to address the issue of insufficient high-quality training samples for individual tree component segmentation. In addition, the state-of-the-art point transformer-V3 deep learning network is employed on the ITS-3D dataset for the segmentation of individual tree components, including the main trunk, branches, and leaf categories. Furthermore, this study optimizes the segmentation performance of the point transformer-V3 network for individual tree point clouds by constructing individual tree prior geometric features (SoD index, PCA2 index, and verticality index). Subsequently, a comparative study will be conducted with several current major deep learning algorithms. Finally, this study also validates the effectiveness of the introduced prior geometric features through ablation experiments.

Experimental results demonstrate that the segmentation performance achieved an Overall Accuracy (OA) of 0.946, mean accuracy of 0.872, and mean Intersection over Union (mIoU) of 0.806, reflecting high segmentation precision. Even smaller, higher-order branches within the tree crown are successfully extracted. Moreover, the model's OA and mIoU values for various TLS subsets of ITS-3D exceed 0.928 and 0.737, respectively, showing high generalization capability across tree species with large geometric structural differences. Furthermore, compared with other mainstream deep learning networks, the point transformer-V3 network adopted in this study attains the highest values for segmentation accuracy metrics. In addition, in the performance comparison with traditional leaf-wood separation algorithms, such as LeWoS, the point transformer-V3 network increases OA and mIoU by 7.9% and 17.2%, respectively. This result fully demonstrates the excellent performance and generalization capability of high-performance deep learning techniques in individual tree component segmentation across multiple tree species. Finally, when all the prior geometric features of trees are incorporated as input in the ablation experiments, the OA and mIoU of the segmentation results reach their peak values of 94.63% and 80.59%, respectively. This finding represents an increase of 0.29% and 0.42% compared with the scenario where only the 3D coordinates of point clouds are used as input, indicating that introducing effective individual tree prior features during training can better promote the segmentation effect of a deep learning network.

Through this research, the application of state-of-the-art deep learning technology to individual tree component segmentation can be improved further, providing technical support for the precise estimation of tree structural parameters.

Key words: individual tree component, LiDAR point clouds, semantic segmentation, prior features, deep learning, leaf-wood separation, branch extraction, serialized attention mechanism

Supported by National Natural Science Foundation of China (No. 42271388, 42171347); Natural Science Foundation of Hubei Province (No. 2024AFA069)