

PTV3-BLSeg: 面向树木点云精细结构解析的 枝叶分离模型

周书羽, 姚宗琦, 张晓丽

1. 北京林业大学 林学院 林木资源高效生产全国重点实验室, 北京 100083;
2. 北京林业大学 林学院 精准林业北京市重点实验室, 北京 100083;
3. 北京林业大学 森林培育与保护教育部重点实验室, 北京 100083

摘要: 针对树木点云枝叶分离中存在的边界模糊、树干与枝叶点云数量显著不平衡及细粒度结构识别困难等问题, 本文基于 Point Transformer V3 架构提出了一种改进的 PTV3-BLSeg (fusion-enhanced boundary-aware Point Transformer V3 for branch-leaf segmentation) 模型。该模型通过引入融合跳跃连接以增强跨层级特征传递, 优化局部和全局特征的融合效果, 有效提升枝叶交界区域的判别能力。为了进一步加强对枝叶边界的识别能力, 设计边界增强损失函数, 聚焦枝干与叶片的细粒度特征, 减少类别之间的模糊过渡。为提升深层特征学习的稳定性和精度, 采用深层监督机制, 优化枝干特征的学习路径, 推动模型的学习效率与性能提升。在 Smart Tree 数据集上, PTV3-BLSeg 模型的整体精度达到 93.33%、枝干 F1 分数为 75.61%、树叶 F1 分数为 96.09%, 分别较基准模型 (Point Transformer V3) 提升 1.36%、4.14% 和 0.85%; 在内蒙古根河数据集中, PTV3-BLSeg 模型的整体精度达到 83.83%、枝干 F1 分数为 60.77%、树叶 F1 分数为 89.50%, 较基准模型分别提升 5.04%、12.86% 和 9.13%; 在广西横州数据集中, PTV3-BLSeg 模型的整体精度达到 81.87%、枝干 F1 分数为 71.32%、树叶 F1 分数为 86.60%, 分别提升 4.49%、12.73% 和 1.48%, 验证了其良好的泛化能力。总体而言, PTV3-BLSeg 模型在树木点云枝叶分离任务中展现出显著优势, 能够有效缓解树干与枝叶边界不清、类别点云数量严重失衡以及细微结构难以识别等关键难题, 可为森林场景三维模型构建及林木参数精细调查提供可靠的技术支持。

关键词: 枝叶分离, Point Transformer V3, 边界增强, 深层监督, 点云语义分割, 跳跃连接, 特征融合, 三维点云中图分类号: TP79/S757.2/P2

引用格式: 周书羽, 姚宗琦, 张晓丽. 2026. PTV3-BLSeg: 面向树木点云精细结构解析的枝叶分离模型. 遥感学报, 30(6): 1611-1626

Zhou S Y, Yao Z Q and Zhang X L. 2026. PTV3-BLSeg: A branch-leaf segmentation model for fine-scale structural analysis of tree point clouds. National Remote Sensing Bulletin, 30(6): 1611-1626 [DOI: 10.11834/jrs.20265427]

1 引言

近年来, 地基激光雷达 TLS (terrestrial laser scanning) 因其高空间分辨率与非接触式三维扫描能力, 被广泛应用于单木结构提取和冠层建模中 (Li 等, 2014; Yen 等, 2010; 马振宇 等, 2019; 刘鲁霞 等, 2014)。与此同时, 随着对轻便化林业调查设备需求的不断增长, 采用背包和手持等便携式移动激光扫描 MLS (Mobile Laser Scanning) 设备开展数据采集, 进而实现精细化林木三维重建

和表型参数提取也逐渐成为具有广阔应用前景的研究方向。基于 TLS 和 MLS 高密度点云数据的树木枝干与叶片的精细分离, 对于压缩数据量、提高建模效率、提升树干和树木主枝的重建精度与完整性, 进而准确获取叶面积指数 LAI (Leaf Area Index)、非光合作用植被 NPV (Non-Photosynthetic Vegetation)、树木材积与地上生物量 AGB (Above Ground Biomass) 等生态参数具有重要意义; 该过程在精细化三维重建中扮演着“数据净化”和“信息解耦”的双重角色, 是基于地面扫描点云开

收稿日期: 2025-10-10; 预印本: 2026-03-30

基金项目: 国家重点研发计划 (编号: 2023YFD2201700); 国家自然科学基金 (编号: 32171779)

第一作者简介: 周书羽, 研究方向为森林三维重建。E-mail: zhoushuyu@bjfu.edu.cn

通信作者简介: 张晓丽, 研究方向为林草生态遥感。E-mail: zhangxl@bjfu.edu.cn

展森林资源调查、碳循环建模与森林结构重建的基础性任务之一 (Dassot 等, 2011; Liang 等, 2016; Zheng 等, 2016)。然而, 由于观测角度限制, 以及树木结构复杂、枝叶交错密集, 导致枝叶之间空间边界模糊, 尤其在天然林和针叶林中存在严重遮挡, 枝叶点云自动化精细分离面临较大挑战。

目前, 枝叶分离方法主要包括基于规则特征的方法、基于机器学习的方法和基于深度学习的方法3大类。基于规则特征的方法又可以分为基于回波点云强度的方法和基于几何特征的方法两类, 前者利用激光扫描仪获取的辐射信息, 假设枝干与叶片在扫描仪工作波长下具有不同的光学特性 (Tao 等, 2015)。例如, Côté 等 (2009) 利用红外波段反射率差异实现了枝干与叶片的区分, 其依据是枝干在该波段的反射率高于叶片。然而单波段方法对环境光照、传感器特性及林分结构依赖性较强, 难以在枝叶交错密集或遮挡严重的场景中保持稳定精度。为此, Strahler 等 (2008) 与 Mark Danson 等 (2018) 采用双波段激光扫描, 显著提升了分离精度。尽管基于点云回波强度的特征能够有效区分枝叶, 但该类方法的光学特性易受波长、距离及入射角等因素影响, 因而通用性较低。此外, 在茂密林分中, 激光束可能因同时作用于枝干和叶片而产生混合回波, 同时树木不同组织结构间的反射率差异也会引起光谱变化, 二者易相互混淆 (Raumonen 等, 2013)。同时, 该方法通常需要精确的辐射校正, 且不同扫描仪间的辐射校正结果难以直接迁移 (Calders 等, 2017), 限制了其跨设备应用。基于几何特征规则的方法则利用点云的三维空间信息, 通过刻画局部结构差异区分枝干与叶片。此类方法通常首先提取每个点的局部空间特征, 如曲率、线性度、平面度等 (Zhang 等, 2019; Wan 等, 2021), 然后结合阈值或拓扑关系实现分类。局部几何特征的提取依赖邻域点的确定方式, 不同研究采用了多种策略, 包括基于 k 维树 KD-Tree (K-Dimensional Tree) 的高效索引 (Klokov 和 Lempitsky, 2017)、基于 k 近邻 k -NN (k -Nearest Neighbors) 的快速搜索 (Li 等, 2018)。此外, Ferrara 等 (2018) 利用 DBSCAN (Density-Based Spatial Clustering of Applications with Noise) 聚类算法, 结合木质部与

叶片在点云密度上的差异, 实现了无监督的枝叶分离。Wang 等 (2020) 的 LeWoS (Leaf-Wood Separation) 通过点云的几何特征约束图与递归阈值分割实现结构感知的枝叶分离。此类方法在保证可解释性的同时, 具有较好的跨设备适应性, 但其阈值及参数设置对不同树种和扫描密度较为敏感, 仍需根据应用场景进行调整。

部分研究进一步引入机器学习方法以提升枝叶分离效果。典型的分类模型包括随机森林 RF (Random Forest) (Breiman, 2001)、高斯混合模型 GMM (Gaussian Mixture Model) (Wang 等, 2017) 以及支持向量机 SVM (Support Vector Machine)。例如, Zhu 等 (2018) 提出基于自适应半径的近邻搜索策略, 从树木点云中联合提取几何与辐射特征, 并利用 RF 完成枝叶分类; Chen 等 (2018) 通过设计可自适应的搜索半径提取局部几何特征, 并基于 SVM 直接识别枝干点。这类方法无需手动设定复杂的阈值, 能够通过学习样本间的统计差异实现更灵活的分类, 但模型性能仍依赖标注数据的质量和数量, 其在跨树种和多场景条件下的泛化能力仍有待加强。

随着点云深度学习模型的发展, 端到端的枝叶分离方法逐渐成为研究热点。这类方法通过直接在原始点云上学习几何特征与局部拓扑关系, 能够避免传统非监督方法对参数敏感、泛化能力不足的问题。Qi 等 (2017) 提出了 PointNet, 这是首个直接在点云上进行学习的网络, 能够高效提取全局几何特征, 但由于缺乏对局部结构的刻画, 在复杂场景下表现有限。随后, PointNet++ (Qi 等, 2017) 引入层次化采样与局部特征聚合机制, 提升了模型对局部空间关系的建模能力, 但在处理非均匀采样点云时仍存在性能下降问题。为进一步增强局部几何关系的表达, DGCNN (Dynamic Graph CNN for learning on point clouds) (Wang 等, 2019) 采用动态图卷积操作, 通过不断更新邻接关系实现局部拓扑的自适应学习, 但其计算复杂度较高。KPConv (Kernel Point Convolution) (Thomas 等, 2019) 则利用可学习的核点卷积算子, 显著提升了网络对局部空间的感知能力, 但卷积核参数开销较大, 在大规模点云上效率不足。

近年来, Transformer 架构被引入点云建模。Point Transformer (Zhao 等, 2021) 通过自注意力

机制实现点间关系建模,提升了全局感知能力。PCT (Point Cloud Transformer) (Guo等, 2021)与Point-BERT (Yu等, 2022)在此基础上进一步优化了特征表达。Point Transformer V2 (Wu等, 2022)在效率和收敛速度上有所提升。Point Transformer V3 (Wu等, 2024)在Point Transformer V2的基础上引入了高效的点间交互机制和改进的注意力算子,显著提升了网络对局部结构与全局上下文的建模能力,并在多个点云基准任务上取得了领先性能。然而,该方法在枝叶分离场景中仍存在以下不足:其一,网络对枝干与叶片交界区域的模糊点判别能力不足;其二,枝干和树叶数量严重不均衡时,模型对小类别(叶片)的响应较弱;其三,训练过程中容易陷入局部最优,限制了整体分割精度。

针对上述枝叶分离任务中的不足,本文以Point Transformer V3为基线,提出了一种改进型PTV3-BLSeg (fusion-enhanced boundary-aware Point Transformer V3 for branch-leaf segmentation)模型,通过引入跳跃连接、边界增强模块与深层监督机制,提升网络对边界点的判别能力、缓解类别不均衡影响,并增强中间层特征的语义引导。同时,本文以传统的LeWoS方法作为基线模型,以综合评估提出的方法在不同树种点云数据上的性能优势。

2 研究区与数据

2.1 Smart Tree公开数据集

本研究在模型训练与评估阶段采用了Smart Tree公开模拟数据集(Dobbs等, 2023)。Smart Tree数据集基于Speed Tree软件仿真生成,该软件广泛应用于数字森林建模,能够高精度控制树木的几何结构和生长形态,从而生成具备真实性和多样性的三维点云数据。Smart Tree数据集中包含6类典型树种:苹果(*Malus domestica*)树、松树(*Pinus ssp.*)、桉树(*Eucalyptus ssp.*)、银杏(*Ginkgo biloba* L.)树、胡桃(*Juglans regia* L.)树与樱桃幼树(Sapling Cherry),每个树种均包含100个单木样本,并提供了逐点的枝干与叶片标签。该数据集覆盖了典型的针叶类与阔叶类树种结构,具备明显的几何特征差异,为训练通用型枝叶分割模

型提供了良好的数据基础。考虑到胡桃树、桉树与樱桃幼树样本中点云密度偏低、结构特征不稳定,可能对模型训练过程产生干扰,故在本研究中未将该3类树种纳入实验流程。最终,选取苹果树、松树和银杏树3类树种作为训练与评估数据来源,3类树种的点云密度分别为23697 pts/m²、15259 pts/m²、18345 pts/m²。实验中将上述数据按7:1:2的比例划分为训练集、验证集与测试集,以保证网络在多树种结构下的充分学习与稳健泛化。

2.2 自建数据集

为进一步验证模型在真实森林场景中的应用效果及其对中国南北方不同树种的泛化能力,本文选取了位于中国北方内蒙古自治区根河样地的实地树木点云数据(所得数据集名称:内蒙古根河数据集)以及位于中国南方广西壮族自治区南宁横州样地的点云数据(所得数据集名称:广西横州数据集),研究区位置如图1所示,树种涵盖白桦(*Betula platyphylla* Suk.)、落叶松(*Larix gmelinii* Rupr.)、桉树(*Eucalyptus ssp.*)、马尾松(*Pinus massoniana* Lamb.)。该数据由手持同时定位与地图构建SLAM(simultaneous localization and mapping)激光扫描设备获取,经预处理后使用CloudCompare软件进行人工精细标注,获得高质量的枝叶分类标签。需特别指出的是,自采数据集仅用于模型测试阶段,未参与任何训练或验证过程,旨在评估模型在高复杂度真实森林场景下的泛化表现。

2.2.1 内蒙古根河数据集

根河市位于内蒙古自治区呼伦贝尔市,地处大兴安岭南麓,地理范围为50°20'N—52°30'N, 120°12'E—122°55'E,平均海拔约714 m。该地区属于季风性亚寒带气候,冬季严寒漫长,夏季凉爽短暂。区域植被类型主要为寒温带针阔混交林,森林覆盖率较高。实验样地内地势相对平坦,郁闭度中等,优势树种为白桦和落叶松。内蒙古根河数据集采用数字绿土H120手持式SLAM设备采集,其中包含白桦35棵,平均点云密度为51678 pts/m²,落叶松35棵,平均点云密度为76933 pts/m²。设备参数如表1所示。

表1 H120手持式SLAM设备参数详情
Table 1 Handheld SLAM device H120 parameter details

参数名称	参数值	参数名称	参数值
电池容量/mAh	5870	测量范围/m	120
电压/V	18	相机类型	360°全景镜头组合
扫描视场角/(°)	280×360	照片分辨率	6080 px×3040 px (2:1)
LiDAR精度/cm	±3	相对精度/cm	≤3
点频(单回波)/(pts/s)	320000	绝对精度/cm	≤5

2.2.2 广西横州数据集

横州市位于中国广西壮族自治区南宁市，地理坐标为22°08′N—23°30′N，108°48′E—109°37′E，属于亚热带季风气候，森林植被主要以常绿阔叶林和常绿针叶林为主，常见树种有马尾松、桉树等，森林覆盖率高。横州地势起伏大，坡度陡，

郁闭度高优势树种为桉树、马尾松。广西横州数据集由千寻公司生产的星光H7手持SLAM设备采集完成，其中桉树35棵，平均点云密度为7050 pts/m²，马尾松35棵，平均点云密度为13193 pts/m²。设备具体参数如表2所示。

表2 H7手持式SLAM设备参数详情
Table 2 Handheld SLAM device H7 parameter details

参数名称	参数值	参数名称	参数值
电池容量/mAh	3450	激光视场角/(°)	水平360，垂直-7—52
卫星系统	GPS+GLO+BDS+GAL+QZSS	照片分辨率	3840×2160 px (2:1)
扫描视场角/(°)	240×143	相对精度/cm	≤2
RTK精度/cm	平面 (0.8+10 ⁻⁶) 高程 (1.5 cm+10 ⁻⁶)	绝对精度/cm	≤5

所有点云数据以.npz格式保存，包含三维坐标及其对应的枝干/叶片标签。在输入网络之前，所有点云样本均进行了归一化与体素下采样处理，以减少内存开销，统一尺度，并确保标签与下采样结果一一对应。本研究下采样参数统一设置为0.02（即将空间划分为边长0.02 m的立方体网格，每个体素只保留一个代表点），点云密度为5000 pts/m²。

3 研究方法

3.1 模型构建

3.1.1 Point Transformer V3模型

Point Transformer V3是一种基于Transformer架构设计的点云深度学习模型，具备建模长距离依赖和多尺度空间结构的能力。其采用编码器—解码器结构，在编码阶段通过多级注意力机制提取点云的全局与局部几何特征，在解码阶段逐层恢复点云空间分辨率。编码器使用多个层级的

Serialized Attention模块，结合点云稀疏特性实现空间感知特征表达；而解码器则配合下采样与上采样模块，实现对结构信息的多尺度重建。Point Transformer V3的网络架构如图1所示。

在本研究中，Point Transformer V3编码阶段包括6个层级，每经过一级编码器，点云数量减少一半，通道数加倍，特征维度由一至六层变化为(N , $C=16$)、($\frac{N}{2}$, $C=32$)、($\frac{N}{4}$, $C=64$)、($\frac{N}{8}$, $C=128$)、($\frac{N}{16}$, $C=256$)、($\frac{N}{32}$, $C=512$)，其中， N 表示单木枝叶点云数量， C 表示点云通道维度。

解码阶段反向恢复空间分辨率，同时维持高维语义信息，最终输出的特征接入分类头进行枝叶二分类预测。

3.1.2 PTV3-BLSeg模型总体框架

PTV3-BLSeg是在Point Transformer V3网络的基础上，结合树木点云中枝干与叶片分布特点而构建的有监督枝叶分离模型。该网络针对类别不

平衡、结构边界模糊等问题, 引入焦点损失 (Focal Loss) 替代常规的二值交叉熵 BCE (Binary Cross Entropy), 并融合3项关键结构设计: 融合跳跃连接、边界增强模块、深层监督机制。

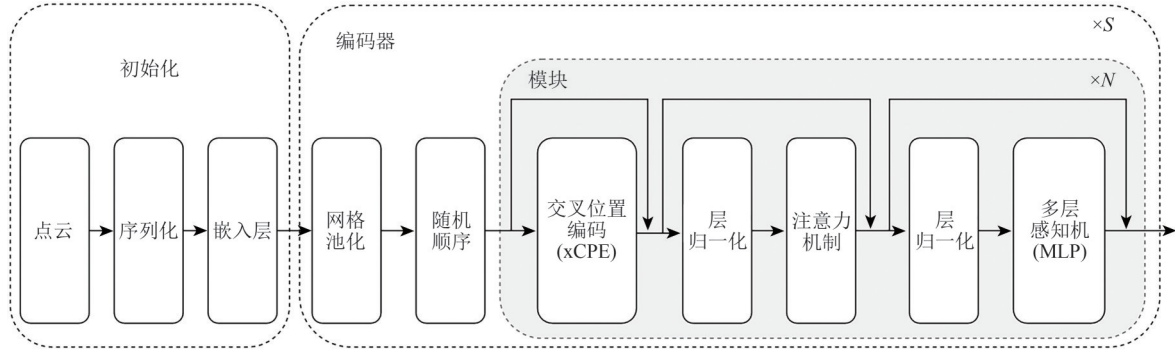


图1 Point Transformer V3网络架构

Fig. 1 Point Transformer V3 network architecture

PTV3-BLSeg输入数据为 $(N, 3)$ 维度的单木点云。在模型训练过程中, 输入数据经由两条分支并行处理。主分支通过PTV3-BLSeg主干网络提取特征, 并在跳跃连接阶段通过特征融合模块 (Fusion Block) 进一步融合特征, 在解码器的最后两层分别获得维度为 $(\frac{N}{2}, C)$ 和 (N, C) 的特征表示。经分类头处理后, 对应输出维度为 $(\frac{N}{2}, 1)$ 和 $(N, 1)$ 的预测标签, 其中前者为辅助分类结

果, 后者为主分类结果。

另一分支针对枝叶边界点进行加权, 以增强边界区域的损失权重, 从而使网络训练更加聚焦于边界特征。结合上述两级标签预测结果, 模型采用深层监督机制以进一步强化损失约束, 促进网络梯度的有效优化。在推理阶段, 深层监督机制不再启用, 模型仅输出主分类结果。模型结构如图2所示。

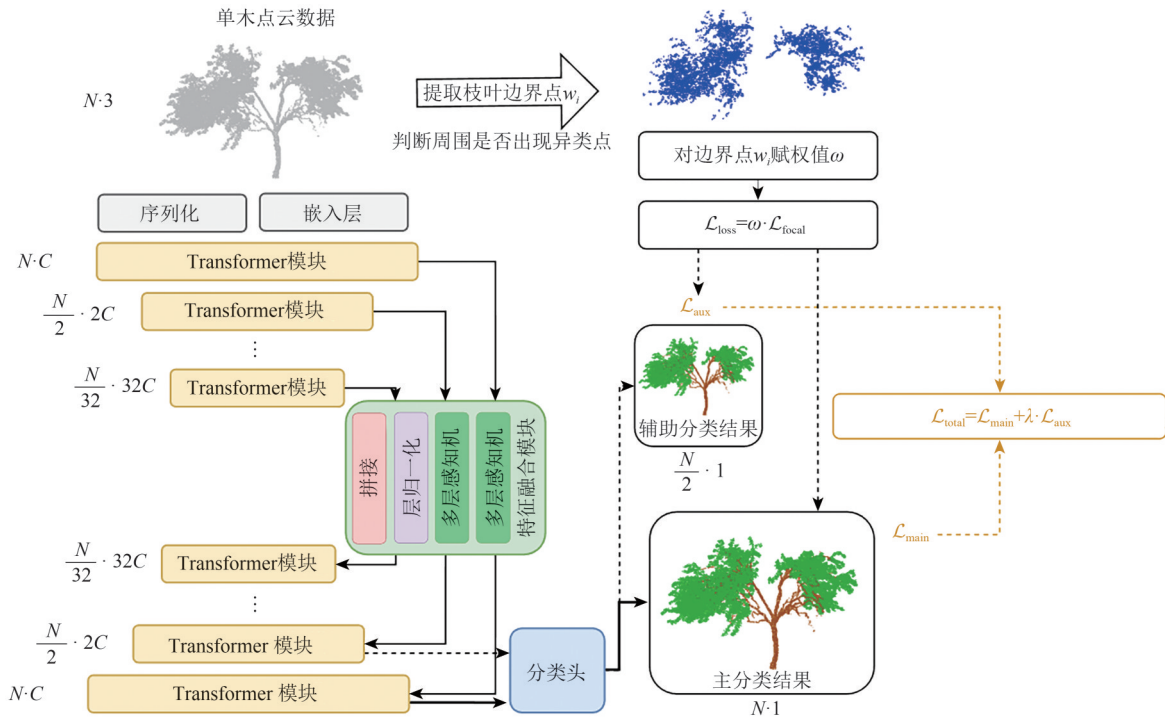


图2 PTV3-BLSeg网络结构示意图(右侧公式见式(6))

Fig. 2 Schematic diagram of PTV3-BLSeg network structure (The formula on the right is shown in Equation (6))

PTV3-BLSeg以Point Transformer V3的编码器—解码器结构为基础,通过5层Transformer编码模块提取多尺度空间特征。在解码过程中,为避免跨层信息融合不足导致的特征退化问题,引入了融合式跳跃连接:将每层解码器上采样后的特征与对应编码器输出进行拼接,并通过多层感知机MLP(Multi-Layer Perceptron)模块进行非线性融合。与常见的跳跃连接不同,本研究在设计时重点考虑了树木点云的结构特性,如枝叶间尺度差异大、叶片呈片状分布而枝干呈线状延展等,因此采用拼接加非线性融合的方式,能够显式保留这些异构几何结构特征,避免简单相加导致的纹理弱化和小尺度叶片特征丢失。这种方式比原始网络中的简单相加更具表达能力,能更好地保留细粒度的空间结构信息。

针对枝干在点云中占比小、边界模糊等难点,PTV3-BLSeg网络引入边界增强模块。具体做法是:通过K近邻方法识别出枝干与叶片的边界点,并在训练过程中对这些边界点赋予更高的损失权重,引导网络关注难分区域,从而提升枝叶交界处的分割准确率。该损失在权重设计上结合了树木点云类别不平衡的特点,对极少量但关键的枝干边界点进行显式强化,使模型在密集叶片背景中仍能学习到稳定的线状枝条结构,这一策略与常规边界损失相比更契合树木点云的结构分布规律。

为提升网络的训练稳定性与中间特征表达能力,PTV3-BLSeg模型进一步引入深层监督机制。在解码器中间层添加辅助监督头,对中间特征直接进行引导,使模型在训练早期便可获得有意义的梯度信号,从而加快收敛并提升整体性能。针对树木枝叶分割中存在的多尺度结构差异大、枝条稀疏细长而叶片密集多变等特点,仅依赖最终输出监督常会导致中间层特征退化,难以充分捕获细枝及局部几何结构。为此,本研究在解码器输出前一层施加显式深层监督,并与边界增强模块协同作用,在解码阶段强化对枝叶局部结构和边界特性的学习,从而显著提升模型对复杂枝叶混合区域的识别与分割能力。

3.1.3 融合跳跃连接

在原始Point Transformer V3的解码阶段,采用的是简单的跳跃连接方式,即直接将上采样特征

与对应编码器输出进行逐元素加和。然而,这种简单操作容易忽略不同层级特征在语义空间与几何尺度上的差异,导致部分细节信息无法有效融合。为增强上下文特征间的信息交互,PTV3-BLSeg引入了融合式跳跃连接机制,即图3中的特征融合模块。

在每个解码阶段,将当前解码器输出特征 $F_d \in \mathbb{R}^{N \times C}$ 与对应编码器输出特征 $F_e \in \mathbb{R}^{N \times C}$ 沿通道维度进行拼接得到 F_{cat} :

$$F_{cat} = (F_d || F_e) \in \mathbb{R}^{N \times 2C} \quad (1)$$

随后通过一个两层感知机进行非线性融合得到融合后的特征 F_{fuse} :

$$F_{fuse} = \text{MLP}(F_{cat}) = W_2(\sigma(\text{LN}(W_1 F_{cat}))) \in \mathbb{R}^{N \times C} \quad (2)$$

式中, W_1 和 W_2 为线性变换参数,LN为LayerNorm, σ 为ReLU激活函数。该机制可有效提升跨层特征对齐与表达能力,增强模型对局部边界结构的辨识能力。

3.1.4 边界增强模块

枝叶分割任务中常面临的一个难点是枝干和树叶的边界分割模糊。为此,PTV3-BLSeg设计了一种边界增强损失策略以提升模型对边界点的识别能力。具体而言,采用基于K近邻的算法检测每个点的局部邻域内是否存在异类标签点,若存在,则将该点标记为边界点。该过程可形式化为

$$\mathcal{B}_{(i)} = \begin{cases} 1, & \exists j \in \mathcal{N}_k(i), j \neq i \\ 0, & \text{其他} \end{cases} \quad (3)$$

式中, $\mathcal{N}_k(i)$ 表示点 i 的 k 个最近邻点集合,点 j 是除点 i 外任一点,如果在 $\mathcal{N}_k(i)$ 中存在任一点 j 与点 i 不属于同一类标签,则 $\mathcal{B}_{(i)}$ 取1,表示点 i 是边界点,反之取0。最终损失函数在二元焦点损失基础上对边界点加权:

$$\mathcal{L}_{focal}(p, y) = -\left(\alpha(1-p)^{\gamma} y \ln(p) + (1-\alpha)p^{\gamma}(1-y) \ln(1-p)\right) \quad (4)$$

式中, $\mathcal{L}_{focal}(p, y)$ 表示预测任意点的损失函数, p 表示预测概率, y 表示真实标签, α 代表类别权重, γ 表示聚焦因子。

$$\mathcal{L} = \sum_i w_i \cdot \mathcal{L}_{focal}(p_i, y_i) \quad (5)$$

式中, $\mathcal{L}$ 为总体损失函数, $w_i = \begin{cases} \omega, & B(i) = 1 \\ 1, & B(i) = 0 \end{cases} (\omega > 1)$,为边界点损失权重。

3.1.5 深层监督

为提高网络的训练稳定性与中间层的特征表达能力, PTV3-BLSeg引入了深层监督机制。该机制在解码器的中间阶段引入一个辅助输出分支, 对中间层特征进行直接监督。设中间层解码特征为 $\mathbf{F}_{\text{aux}} \in \mathbb{R}^{N \times C}$, 其预测输出为: $\hat{y}_{\text{aux}} = \sigma(\mathbf{W}_{\text{aux}} \mathbf{F}_{\text{aux}}) \in \mathbb{R}^{N \times 1}$, 其中 $\mathbf{W}_{\text{aux}}$ 为辅助头线性投影矩阵, σ 为 sigmoid 函数。训练损失 ($\mathcal{L}_{\text{total}}$) 由主损失 ($\mathcal{L}_{\text{main}}$) 和辅助损失 ($\mathcal{L}_{\text{aux}}$) 组成, 二者均为焦点损失, 具体公式如下:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{main}} + \lambda \cdot \mathcal{L}_{\text{aux}} \quad (6)$$

式中, λ 控制辅助损失权重。该设计为中间层提供稳定的梯度反馈, 可提升训练效率并增强小目标的表达能力。

3.2 评价指标

为全面评估本文所提出的 PTV3-BLSeg 模型在枝干与叶片精细分割任务中的分类性能, 采用了包括总体精度 OA (Overall Accuracy)、Kappa 系数以及枝干和树叶的 $F1$ 分数在内的多项指标。设真实为枝干且预测为枝干的点数为 T_w , 真实为叶片且预测为叶片的点数为 T_l , 真实为叶片但被误预测为枝干, 以及真实为枝干但被误预测为叶片的点数分别记为 F_w 、 F_l , 总点数为

$$N = T_w + T_l + F_w + F_l \quad (7)$$

OA 用于反映模型对所有点的分类正确率, 定义如下:

$$\text{OA} = \frac{T_w + T_l}{N} \quad (8)$$

$F1$ 分数是精确率 (Precision) 与召回率 (Recall) 的调和平均值, 枝干和叶片的 $F1$ 分数定义如下:

$$P_w = \frac{T_w}{T_w + F_w} \quad (9)$$

$$R_w = \frac{T_w}{T_w + F_l} \quad (10)$$

$$P_l = \frac{T_l}{T_l + F_l} \quad (11)$$

$$R_l = \frac{T_l}{T_l + F_w} \quad (12)$$

式中, P_w 为预测枝干的精确率, R_w 为预测枝干的召回率, P_l 为预测叶片的精确率, R_l 为预测叶片的召回率, 据此, 枝干类别的 $F1$ 分数 ($F1_{\text{wood}}$) 定义为

$$F1_{\text{wood}} = \frac{2P_w R_w}{P_w + R_w} \quad (13)$$

叶片的 $F1$ 分数 ($F1_{\text{leaf}}$) 则定义为

$$F1_{\text{leaf}} = \frac{2P_l R_l}{P_l + R_l} \quad (14)$$

Kappa 系数 (κ) 通过比较预测结果与随机分类结果的一致程度进行调整, 表达式如下:

$$\kappa = \frac{p_o - p_e}{1 - p_e} \quad (15)$$

式中, p_o 代表 OA, $p_e = \frac{(T_w + F_l)(T_w + F_w) + (F_w + T_l)(F_l + T_l)}{N^2}$ 。

3.3 实验环境

模型训练和验证均在 Linux Ubuntu 22.04 操作系统中进行, 其图形处理单元 GPU (Graphics Processing Unit) 为 NVIDIA GTX4090 24 G, 计算统一设备体系结构 CUDA (compute unified device architecture) 版本为 12.4, 所有算法基于开源深度学习框架 PyTorch 实现, 版本为 2.1.1。本研究提出的 PTV3-BLSeg 模型, 其主干网络为 PTV3 网络。为直观展示其改进之处, 在对比实验中, 相同的模型参数保持一致, 模型具体参数如表 3 所示。

表 3 主干网络关键超参数设置

Table 3 Key hyperparameter configuration of the backbone network

参数	参数值
输入通道数	3
下采样步长	2, 2, 2, 2
输出通道数	16, 32, 64, 128, 256
自注意力头数	2, 2, 4, 8, 16
序列块长度	1024, 1024, 1024, 1024
输出通道数	1

训练过程中使用 AdamW 优化器, 学习率初始化为 1×10^{-3} , 权重衰减系数设为 1×10^{-4} , 批大小设为 2, 迭代次数设为 100 次。边界损失超参数 ω 设置为 3, 深层监督超参数 λ 设置为 0.5, 类别权重 α 设为 0.75, 聚焦因子 γ 设为 2。

4 实验结果与分析

4.1 对比实验

为了充分评估 PTV3-BLSeg 在枝叶分离任务的性能, 本研究使用 LeWos 和 Point Transformer V3 深度学习作为对比方法, 在 3 个数据集进行对比试验。

4.1.1 Smart Tree数据集的分割结果与分析

表4描述了这两种方法和本文提出的PTV3-BLSeg在Smart Tree数据集上的精度、Kappa系数,以及枝干F1得分与叶片F1得分上的比较结果。定量分析表明,PTV3-BLSeg在整体性能上全面超越对比方法,其整体精度达93.33%,较Point Transformer V3提升1.36%,较LeWos提升5.86%;

枝干F1分数达到75.61%,较Point Transformer V3提升4.14%,较LeWos大幅提升35.65%;叶片F1分数达96.09%,较Point Transformer V3提升0.85%;Kappa系数达71.89%,较Point Transformer V3提升5.19%。这一系列数据充分验证了PTV3-BLSeg在枝叶分离任务中的综合优势。

表4 不同枝叶分离方法在Smart Tree数据集上的分割性能比较

Table 4 Performance comparison of different branch - leaf separation methods on the Smart Tree Dataset					
树种	模型	总体精度/%	枝干F1分数/%	叶片F1分数/%	Kappa系数/%
苹果树	LeWos	83.94	49.37	91.11	39.14
	Point Transformer V3	89.12	56.24	93.70	50.77
	PTV3-BLSeg(本文)	89.68	59.37	94.01	54.06
松树	LeWos	87.26	51.62	90.37	46.95
	Point Transformer V3	94.88	80.98	97.03	78.02
	PTV3-BLSeg(本文)	96.54	86.54	98.01	84.56
银杏树	LeWos	91.19	57.44	95.00	53.59
	Point Transformer V3	95.81	83.00	97.54	80.60
	PTV3-BLSeg(本文)	96.06	84.28	97.68	82.00
整体	LeWos	87.47	39.96	92.16	36.53
	Point Transformer V3	91.97	71.47	95.24	66.70
	PTV3-BLSeg(本文)	93.33	75.61	96.09	71.89

PTV3-BLSeg对不同树种展现出稳定的泛化能力。对于枝干形态复杂的苹果树,其枝干F1分数达到59.37%,较Point Transformer V3提升3.13%,较LeWos提升10%,显著改善了弯曲枝干的识别精度;对于细枝干密集的松树,枝干F1分数提升尤为显著至86.54%,较Point Transformer V3提升5.56%,较LeWos提升34.92%,证明其对细枝干的精准捕捉能力;而对于枝干相对稀疏的银杏,虽各方法性能差距收窄,但PTV3-BLSeg仍以84.28%的枝干F1分数保持领先,较Point Transformer V3提升1.28%,其82.00%的Kappa系数更较LeWos提升28.41%,体现了其分割结果的高可靠性。

值得注意的是,传统方法LeWos在枝干分割方面表现不佳。其在苹果树、松树、银杏树的枝干F1分数分别仅为49.37%、51.62%、57.44%,整体枝干F1分数更是降至39.96%。相较于LeWos,Point Transformer V3在整体枝干分割性能上有明显提升,其枝干F1分数达到71.47%。然而,在苹果树数据上,其枝干F1分数仅为56.24%,较PTV3-

BLSeg低3.13%;在松树数据上亦低5.56%,显示其在不同树种上的分割能力仍存在一定局限。

图3展示了Smart Tree数据集中3个树种在不同枝叶分割方法下的可视化对比结果。绿色和棕色分别表示叶片与枝干,差异图中红色代表模型误分类成叶片的枝干点,蓝色代表模型误分类成枝干的叶片点。可见,LeWos方法虽然具备一定的结构区分能力,但在枝叶交界区域存在显著误分,尤其在冠层密集或分布不均的情况下表现不稳定。Point Transformer V3在整体分割准确性上较LeWos有明显提升,能够较好保留树木主干及一级枝结构,但仍存在一定程度的边界模糊和叶片误判。相比之下,本文提出的PTV3-BLSeg方法通过引入边界增强损失与辅助监督策略,有效提升了模型对细粒度结构的辨识能力,在不同树种中均实现了更清晰的枝叶分界与更低的误分类率,表现出更强的鲁棒性与泛化能力,验证了所提方法在三维树木点云精细分割任务中的有效性。

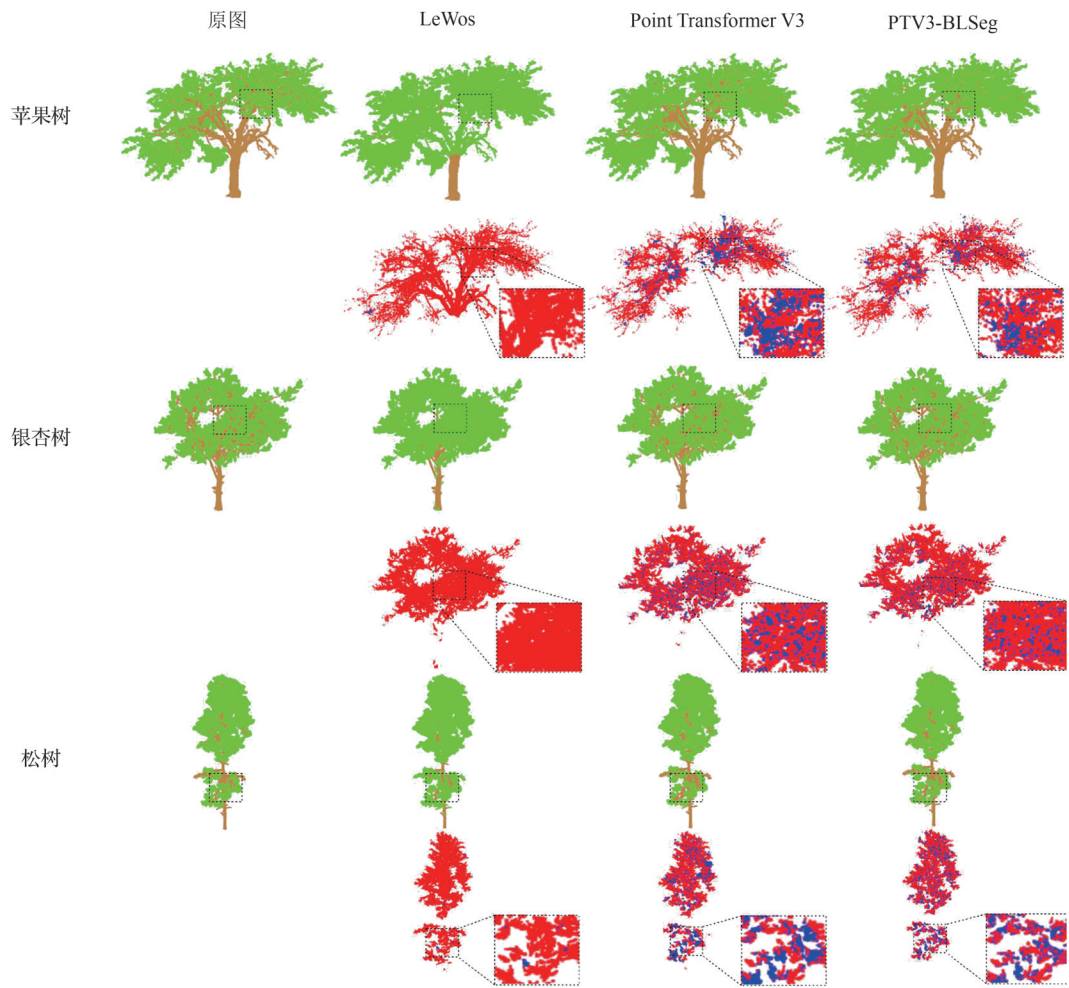


图3 Smart Tree 数据集上不同枝叶分离方法的可视化对比

Fig. 3 Visual comparison of different leaf-branch separation methods on the Smart Tree Dataset

4.1.2 内蒙古根河数据集分割结果与分析

在内蒙古根河数据集上，PTV3-BLSeg 同样展现出卓越的泛化性能。如表 5 所示，该方法在白桦和落叶松两种树种上的分割性能全面领先。总体精度达 83.83%，较 Point Transformer V3 提升 5.04%，较 LeWos 提升 3.25%；枝干 $F1$ 分数达 60.77%，较 Point Transformer V3 大幅提升 12.86%，较 LeWos 提升 13.88%；叶片 $F1$ 分数达 89.50%，较 Point Transformer V3 提升 9.13%；Kappa 系数达 52.66%，较 Point Transformer V3 提升 12.33%。这一结果在特定树种场景中进一步验证了 PTV3-BLSeg 的鲁棒性。

分树种观察发现，PTV3-BLSeg 在复杂树形结构中优势尤为显著。对于枝干表皮特征独特的白桦树，其枝干 $F1$ 分数达 60.85%，较 Point Transformer V3 提升 11.58%，较 LeWos 提升 9.93%；在枝叶密

集交错的落叶松场景，其枝干 $F1$ 分数提升更为突出至 60.65%，较 Point Transformer V3 提升 14.83%，较 LeWos 提升 20.80%。值得关注的是，Point Transformer V3 在该数据集上出现异常表现，其白桦叶片 $F1$ 分数仅为 81.49%，甚至低于传统方法 LeWos 的 88.18%，而 PTV3-BLSeg 以 89.10% 的叶片 $F1$ 分数成功克服此问题，证明其能有效适应不同树种的特性差异。

传统方法 LeWos 在北方树种中仍保持基础分割能力，但枝干识别存在根本缺陷。其落叶松枝干 $F1$ 分数仅 39.85%，表明白桦树皮纹理与落叶松针叶遮挡对基于规则的方法构成严峻挑战。Point Transformer V3 虽在枝干分割上优于 LeWos，但叶片分割性能出现显著波动，其落叶松叶片 $F1$ 分数低至 78.63%，表明标准深度学习模型对区域适应性不足。而 PTV3-BLSeg 在保持枝干分割优势的同时，将叶片 $F1$ 分数稳定提升至 90% 以上。

表 5 不同枝叶分离方法在根河数据集上的分割性能比较

Table 5 Performance comparison of different branch – leaf separation methods on the Genhe Dataset from Inner Mongolia

树种	模型	总体精度/%	枝干 F1 分数/%	叶片 F1 分数/%	Kappa 系数/%
白桦	LeWos	81.38	50.92	88.18	42.78
	Point Transformer V3	78.82	49.27	81.49	39.05
	PTV3-BLSeg(本文)	83.22	60.85	89.10	52.58
落叶松	LeWos	79.44	39.85	87.33	32.48
	Point Transformer V3	78.72	45.82	78.63	42.16
	PTV3-BLSeg(本文)	84.76	60.65	90.15	52.77
整体	LeWos	80.58	46.89	87.83	38.46
	Point Transformer V3	78.79	47.91	80.37	40.33
	PTV3-BLSeg(本文)	83.83	60.77	89.50	52.66

图 4展示了在内蒙古根河数据集中两个树种（白桦与落叶松）上的枝叶分割结果对比。由图可见：LeWos 方法在白桦上存在大量误分类，尤其在主干附近表现出明显的枝叶混淆；Point Transformer V3相较传统方法具有更强的结构表达能力，但在局部复杂区域（图中黑框标注处）仍易出现分界模糊和细枝误判；PTV3-BLSeg 在保持

主结构完整性的基础上，有效抑制了误分类点的出现，尤其在落叶松的冠层边缘与干枝交汇区域，边界识别更为清晰，表现出更好的细粒度分割能力与结构保真性。该结果进一步验证了本文方法在多样性更高的实际数据中仍具备良好的泛化性能与鲁棒性。

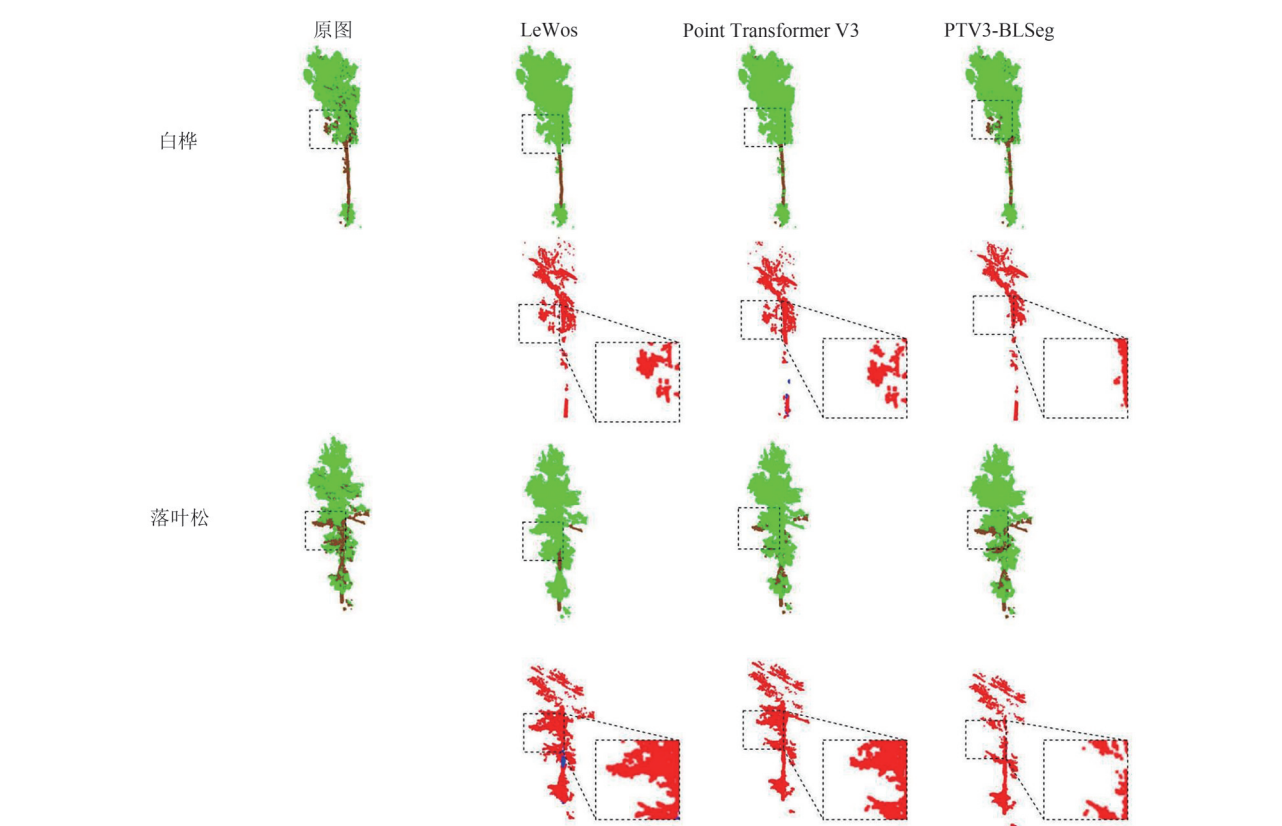


图 4 内蒙古根河数据集上不同枝叶分离方法的可视化对比
Fig. 4 Visual comparison of different leaf–branch separation methods on the Genhe Dataset from Inner Mongolia

4.1.3 广西横州数据集分割结果与分析

在广西横州数据集上，PTV3-BLSeg在枝叶分

割任务中表现最为优异（表6）。整体来看，其总体精度达到81.87%，较Point Transformer V3提升

4.49%，较 LeWos 提升 11.40%；枝干 $F1$ 分数为 71.32%，较 Point Transformer V3 和 LeWos 提升 12.73% 和 20.95%；叶片 $F1$ 分数达到 86.60%，比 Point Transformer V3 提升 1.48%，超越 LeWos 5.32%；Kappa 系数为 59.04%，显示整体分割一致性显著提升。

表 6 不同枝叶分离方法在广西横州数据集上的分割性能比较

Table 6 Performance comparison of different branch – leaf separation methods on the Hengzhou Dataset from Guangxi

树种	模型	总体精度/%	枝干 $F1$ 分数/%	叶片 $F1$ 分数/%	Kappa 系数/%
桉树	LeWos	71.38	43.26	87.18	40.54
	Point Transformer V3	79.24	52.46	86.33	43.12
	PTV3-BLSeg(本文)	85.02	73.36	89.29	63.10
马尾松	LeWos	69.55	57.48	75.37	46.48
	Point Transformer V3	75.51	64.70	79.28	52.02
	PTV3-BLSeg(本文)	78.72	69.28	83.90	54.97
整体	LeWos	70.47	50.37	81.28	43.51
	Point Transformer V3	77.38	58.59	85.12	47.57
	PTV3-BLSeg(本文)	81.87	71.32	86.60	59.04

对比不同树种的分割结果，PTV3-BLSeg在复杂结构场景中的优势尤为突出。桉树枝干 $F1$ 分数达 73.36%，较 Point Transformer V3 和 LeWos 分别提升 20.90% 和 30.10%，同时叶片 $F1$ 分数保持在 89.29%；马尾松枝干 $F1$ 分数为 69.28%，叶片 $F1$ 分数为 83.90%，均优于其他方法。这表明 PTV3-BLSeg能够有效应对枝叶交错和遮挡复杂的场景。

相比之下，LeWos在枝干分割上存在明显不足，其桉树枝干 $F1$ 分数仅为 43.26%；Point Transformer V3虽有提升，但叶片分割波动较大，局部复杂区域仍易出现分界模糊。PTV3-BLSeg不仅保持枝干分割优势，同时稳定提升叶片分割精度，实现了整体分割的均衡与可靠。

图 5 展示了不同方法在广西横州数据集树木样本上的分割效果。对于桉树样本，在树根及近地面主干区域，LeWos 由于其基于几何连续性与轴向结构建模的特点，因此对粗壮且结构单一的主干部分表现出较好的稳定性。相比之下，Point Transformer V3 及 PTV3-BLSeg 在该区域受限于点云密度不均及局部上下文信息不足，仍存在少量误分现象。

然而，在枝干与叶片交界处及冠层内部等结构复杂区域，LeWos 的几何判别能力明显下降，而 PTV3-BLSeg 通过边界增强与结构约束显著减少了误分类点，在整体结构一致性与细粒度分割方面表现更为稳定。

4.1.4 模型泛化性分析

由表 5 和表 6 可知，PTV3-BLSeg 在内蒙古根河数据集与广西横州数据集上均取得了优于 LeWoS 和 Point Transformer V3 的分割性能，但两地数据仍呈现一定的精度差异。从总体精度来看，PTV3-BLSeg在内蒙古根河数据集上的表现更为突出，整体精度达到 83.83%，Kappa 系数为 52.66%；而在广西横州数据集上总体精度为 81.87%，Kappa 系数为 59.04%。尽管广西横州数据集的 Kappa 系数略高，但内蒙古根河数据集在枝干与叶片的 $F1$ 分数上更加稳定。其主要原因在于两地树种结构特性存在明显差异：内蒙古根河数据集以白桦和落叶松为主，树干结构清晰、枝条层级分明，点云密度较高，有助于模型捕捉枝干与叶片的局部几何差异。

相比之下，广西横州数据集包含桉树和马尾松，具有冠层重叠严重、枝叶尺度细小及树干易被遮挡等特点，加之点云密度偏低，使细枝干与叶片边界处的结构信息不足，从而增加了树干提取的难度。此外，不同树种的冠层结构差异也对模型的泛化能力产生影响。阔叶树（如桉树）枝叶交错密集、叶片纹理细碎、边界过渡模糊；而针叶树（如落叶松）具有相对规则的层状结构，使模型更容易学习其空间层级特征。尽管 PTV3-BLSeg通过边界增强模块与跨层特征融合在一定程度上缓解了这些问题，但在冠层结构高度复杂、叶片大量重叠的树种中，仍可能出现局部细枝干误分或叶片识别不足的情况。

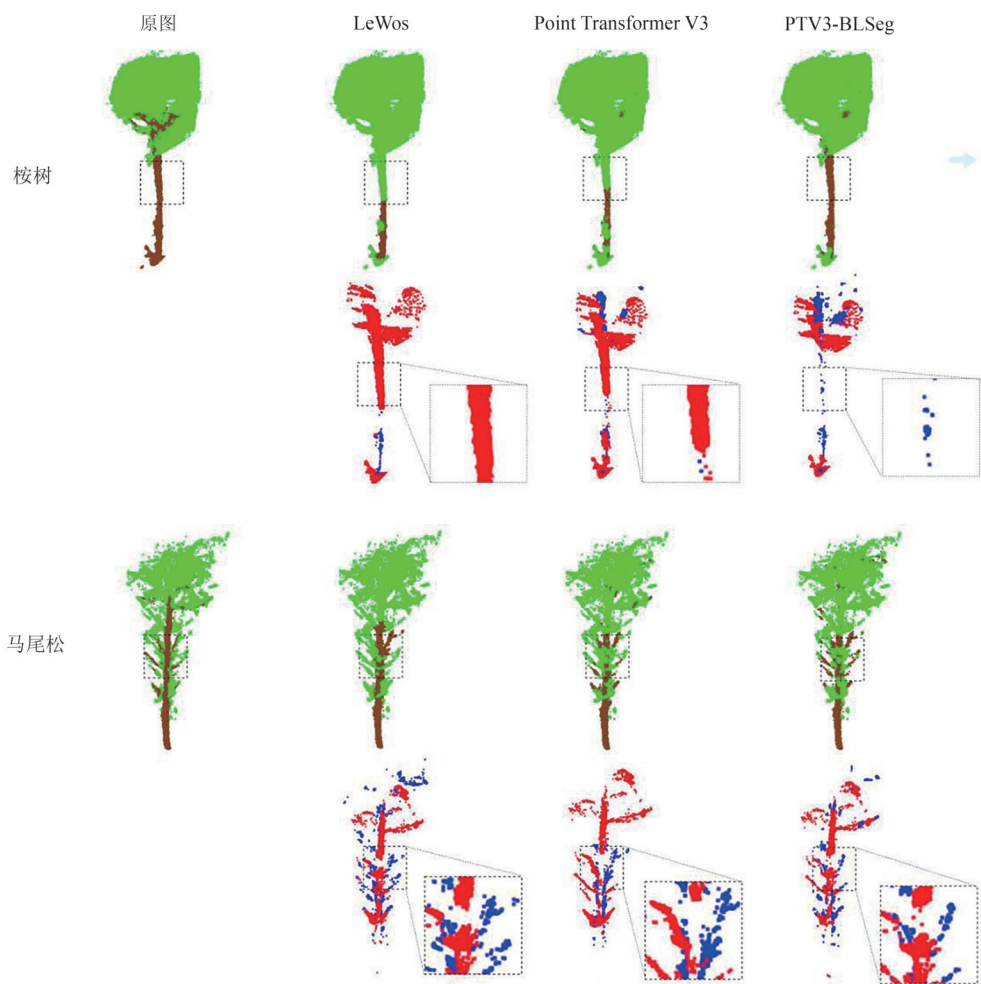


图5 广西横州数据集上不同枝叶分离方法的可视化对比

Fig. 5 Visual comparison of different leaf-branch separation methods on the Hengzhou Dataset from Guangxi

4.2 消融实验与分析

为验证 PTV3-BLSeg 核心模块的有效性，本研究设计了系统的消融实验。如表 7 所示，在 Smart Tree 数据集上逐步引入跳跃连接、边界增强与深层监督，可观察到模型性能的持续提升：当仅添加跳跃连接时，枝干 $F1$ 分数由基础模型 Point transformer V3 的 71.47% 提升至 71.90%，增幅

0.43%；叠加边界增强模块后，枝干 $F1$ 分数进一步升至 72.67%，较单一跳跃连接提升 0.77%；最终引入深层监督的完整模型 PTV3-BLSeg，枝干 $F1$ 分数跃升至 75.61%，较 3 个模块缺失的基础模型提升 4.14%。这种阶梯式增长验证了各模块的协同增效作用。

表 7 Smart Tree 数据集上 PTV3-BLSeg 模型消融实验

Table 7 Ablation experiment of PTV3-BLSeg on the Smart Tree Dataset

模型	改进方式			总体精度/%	枝干 $F1$ 分数/%	叶片 $F1$ 分数/%	Kappa 系数/%
	跳跃连接	边界增强	深层监督				
Point Transformer V3	×	×	×	91.97	71.47	95.24	66.70
	√	×	×	92.55	71.90	95.64	67.73
	√	√	×	92.83	72.67	95.81	68.67
PTV3-BLSeg	√	√	√	93.33	75.61	96.09	71.89

深层监督机制对枝干分割的优化贡献最为显著。对比边界增强模型与完整模型数据，深层监

督的引入使枝干 $F1$ 分数提升 2.94%，同时推动 Kappa 系数提升 3.22%。这表明通过中层特征图施

加的监督信号，有效缓解了枝干特征在深层网络中的衰减问题，尤其改善了细枝干末端的识别精度。边界增强模块则主要提升分割边界清晰度，其添加后叶片 $F1$ 分数提高 0.17%，枝干 $F1$ 分数提高 0.77%，总体精度提升 0.28%，证明该模块能有效区分枝叶粘连区域的过渡点。

跳跃连接的基础性作用同样不可忽视。该模块使模型总体精度提升 0.58%，Kappa 系数提升 1.03%，说明跨层级特征融合增强了空间上下文建模能力。值得注意的是，当仅采用跳跃连接时，枝干 $F1$ 分数增幅有限，而结合边界增强后性能提升扩大，证明几何特征优化与多尺度特征传递存在耦合增益。完整模型 PTV3-BLSeg 最终实现 71.89% 的 Kappa 系数，较基础模型提升 5.19%，标

志着分割结果可靠性的本质提升。

图 6 展示了 Smart Tree 数据集中一例松树样本的消融实验可视化对比结果。由图 7 可观察到：在标准 Point Transformer V3 基础上引入融合跳跃连接后，枝叶过渡区域的误分类点呈现更加分散且细碎的形态，说明跳跃连接在一定程度上增强了局部结构特征的表达，但仍存在边界模糊的问题。当进一步加入边界增强模块时，模型在枝叶边界处的误差分布更加细腻，表明焦点损失函数有效提升了对过渡区域的辨识能力。在此基础上加入深层监督后，整体分割质量达到最佳，误分类点最少，边界形态最为清晰，体现出深层监督对于提高模型稳定性与边界感知能力的重要性。

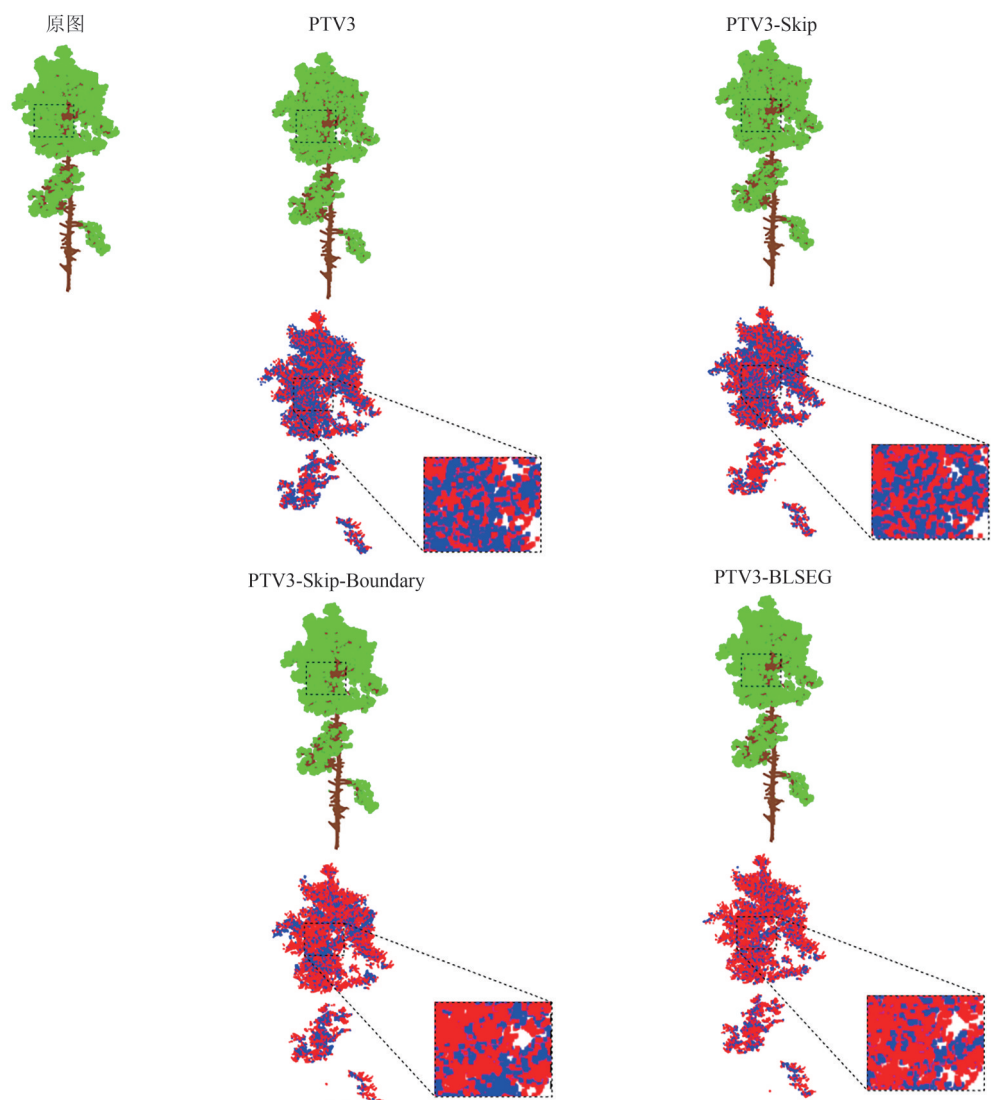


图 6 Smart Tree 数据集中一例松树样本的消融实验可视化对比
Fig. 6 Ablation study visualization comparison on a Pine sample from the Smart Tree Dataset

综合消融实验结果, PTV3-BLSeg的技术突破源于三重创新设计的有机整合: 跳跃连接保障了多尺度特征的完整性传递, 边界增强模块强化了枝叶交界区域的判别能力, 深层监督机制则针对性地优化了枝干特征学习路径。这种结构化改进方案成功解决了点云枝叶分割中的三大核心挑战——细粒度特征保留、边界模糊抑制以及枝干语义信息衰减。

5 结 论

本文针对森林点云中枝干与叶片空间交错、类别比例失衡及边界模糊等问题, 以 Point Transformer V3为基础, 引入融合式跳跃连接、边界增强模块和深层监督机制, 提出了改进的 PTV3-BLSeg 模型, 有效提升了深度学习枝叶分离算法的多尺度特征传递、复杂区域判别和深层语义表达能力。实验结果表明, 该方法在 Smart Tree 数据集、内蒙古根河数据集、广西横州数据集上均取得较基线方法更高的总体精度和更好的枝干识别能力, 在不同树种、不同采样环境下表现出良好的稳定性和泛化性。

与现有基于深度学习的枝叶分割方法相比, PTV3-BLSeg 针对枝干识别薄弱、边界区域误分率高、训练过程易陷局部最优等瓶颈问题提出了有针对性的改进, 并通过消融实验验证了3项创新设计的协同增效作用。其中, 融合跳跃连接缓解了跨层特征丢失导致的局部细节不足; 边界增强模块有效提升了枝叶交界点的判别能力; 深层监督机制增强了中间层梯度信号, 提高了细枝干和局部结构的识别精度。

尽管如此, 模型仍存在一定局限。一是对极端稠密或稀疏点云适应性不足, 当点云异常稠密时, 体素下采样会在局部区域集中删除大量细枝干与叶片的原始结构信息, 导致特征表达出现“过度平滑”, 从而在复杂冠层或细小枝条区域产生结构细节丢失; 而在极度稀疏点云中, 局部几何纹理不足、邻域特征不稳定, 使模型难以充分捕获枝干与叶片的细粒度差异, 因此更易出现边界模糊和枝干误分。换言之, 模型的性能在点云密度接近其表征能力上限或下限时均会下降。二是训练过程对硬件资源要求较高, 不利于大范围快速部署。三是未充分引入树木形态学和生物学先验, 对特殊树形或受损个体适应性有限。这些

问题主要源于深度网络对数据分布和特征稳定性的依赖。

未来研究可从两方面展开: 其一, 探索网络轻量化和模型压缩技术, 实现林下设备的快速推理; 其二, 开发跨尺度、多源数据联合建模策略, 提高模型在不同传感器、采样密度和多林型环境下的适用性与稳定性。

综上, PTV3-BLSeg有效提升了复杂森林环境中枝叶精细分割的准确性和鲁棒性, 为后续森林三维信息智能提取与生态参数高精度反演提供了可行路径与方法参考。

参考文献(References)

- Breiman L. 2001. Random forests. *Machine Learning*, 45(1): 5-32 [DOI: 10.1023/A:1010933404324]
- 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]
- Charles R Q, Su H, Mo K C and Guibas L J. 2017. PointNet: Deep learning on point sets for 3D classification and segmentation//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE: 77-85 [DOI: 10.1109/cvpr.2017.16]
- Chen M L, Wan Y C, Wang M W and Xu J Z. 2018. Automatic stem detection in terrestrial laser scanning data with distance-adaptive search radius. *IEEE Transactions on Geoscience and Remote Sensing*, 56(5): 2968-2979 [DOI: 10.1109/TGRS.2017.2787782]
- Côté J F, Widlowski J L, Fournier R A and Verstraete M M. 2009. The structural and radiative consistency of three-dimensional tree reconstructions from terrestrial LiDAR. *Remote Sensing of Environment*, 113(5): 1067-1081 [DOI: 10.1016/j.rse.2009.01.017]
- Dassot M, Constant T and Fournier M. 2011. The use of terrestrial LiDAR technology in forest science: application fields, benefits and challenges. *Annals of Forest Science*, 68(5): 959-974 [DOI: 10.1007/s13595-011-0102-2]
- Dobbs H, Batchelor O, Green R and Atlas J. 2023. Smart-tree: neural medial axis approximation of point clouds for 3D tree skeletonization//11th Iberian Conference on Pattern Recognition and Image Analysis. Alicante: Springer: 351-362 [DOI: 10.1007/978-3-031-36616-1_28]
- 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]
- Klokov R and Lempitsky V. 2017. Escape from cells: deep Kd-networks for the recognition of 3D point cloud models//Proceedings of the IEEE International Conference on Computer Vision. Venice: IEEE: 863-872 [DOI: 10.1109/ICCV.2017.99]
- Li W, Niu Z, Gao S, Huang N and Chen H Y. 2014. Correlating the horizontal and vertical distribution of LiDAR point clouds with components of biomass in a *Picea crassifolia* forest. *Forests*, 5(8): 1910-1930 [DOI: 10.3390/f5081910]
- Li Y Y, Bu R, Sun M C, Wu W, Di X H and Chen B Q. 2018. PointCNN: convolution on X-transformed points//Proceedings of the 32nd International Conference on Neural Information Processing Systems. Montréal: Curran Associates Inc.: 828-838
- Liang X L, Kankare V, Hyypä J, et al. 2016. Terrestrial laser scanning in forest inventories. *ISPRS Journal of Photogrammetry and Remote Sensing*, 115: 63-77 [DOI: 10.1016/j.isprsjprs.2016.01.006]
- Liu L X, Pang Y, Li Z Y, Xu G C, Li D and Zheng G. 2014. Retrieving structural parameters of individual tree through terrestrial laser scanning data. *Journal of Remote Sensing (in Chinese)*, 18(2): 365-377 (刘鲁霞, 庞勇, 李增元, 徐光彩, 李丹, 郑光. 2014. 用地基激光雷达提取单木结构参数——以白皮松为例. *遥感学报*, 18(2): 365-377) [DOI: 10.11834/jrs.20143091]
- Ma Z Y, Pang Y, Li Z Y, Lu H, Liu L X and Chen B W. 2019. Fine classification of near-ground point cloud based on terrestrial laser scanning and detection of forest fallen wood. *Journal of Remote Sensing (in Chinese)*, 23(4): 743-755 (马振宇, 庞勇, 李增元, 卢昊, 刘鲁霞, 陈博伟. 2019. 地基激光雷达森林近地面点云精细分类与倒木提取. *遥感学报*, 23(4): 743-755) [DOI: 10.11834/jrs.20197383]
- Mark Danson F, Sasse F and Schofield L A. 2018. Spectral and spatial information from a novel dual-wavelength full-waveform terrestrial laser scanner for forest ecology. *Interface Focus*, 8(2): 20170049 [DOI: 10.1098/rsfs.2017.0049]
- 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
- Raunonen P, Kaasalainen M, Åkerblom M, Kaasalainen S, Kaartinen H, Vastaranta M, Holopainen M, Disney M and Lewis P. 2013. Fast automatic precision tree models from terrestrial laser scanner data. *Remote Sensing*, 5(2): 491-520 [DOI: 10.3390/rs5020491]
- Strahler A H, Jupp D L B, Woodcock C E, Schaaf C B, Yao T, Zhao F, Yang X Y, Lovell J, Culvenor D, Newnham G, Ni-Miester W and Boykin-Morris W. 2008. Retrieval of forest structural parameters using a ground-based LiDAR instrument (Echidna®). *Canadian Journal of Remote Sensing*, 34(S2): S426-S440 [DOI: 10.5589/m08-046]
- Tao S L, Guo Q H, Su Y J, Xu S W, Li Y M and Wu F F. 2015. A geometric method for wood-leaf separation using terrestrial and simulated LiDAR data. *Photogrammetric Engineering & Remote Sensing*, 81(10): 767-776 [DOI: 10.14358/pers.81.10.767]
- Thomas H, Qi C R, Deschaud J E, Marcotegui B, Goulette F and Guibas L. 2019. KPConv: flexible and deformable convolution for point clouds//Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul: IEEE: 6410-6419 [DOI: 10.1109/iccv.2019.00651]
- Wan P, Shao J, Jin S N, Wang T J, Yang S M, Yan G J and Zhang W M. 2021. A novel and efficient method for wood-leaf separation from terrestrial laser scanning point clouds at the forest plot level. *Methods in Ecology and Evolution*, 12(12): 2473-2486 [DOI: 10.1111/2041-210X.13715]
- 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, Momo Takoudjou S 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 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[J]. *ACM Transactions on Graphics (TOG)*, 38(5): 146 [DOI: 10.1145/3326362]
- 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 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE: 4840-4851 [DOI: 10.1109/cvpr52733.2024.00463]
- Wu X Y, Lao Y X, Jiang L, Liu X H and Zhao H S. 2022. Point transformer V2: grouped vector attention and partition-based pooling//Proceedings of the 36th International Conference on Neural Information Processing Systems. New Orleans: Curran Associates Inc.: 2415
- Yen T M, Ji Y J and Lee J S. 2010. Estimating biomass production and carbon storage for a fast-growing makino bamboo (*Phyllostachys makinoi*) plant based on the diameter distribution model. *Forest Ecology and Management*, 260(3): 339-344 [DOI: 10.1016/j.foreco.2010.04.021]
- Yu X M, Tang L L, Rao Y M, Huang T J, Zhou J and Lu J W. 2022. Point-BERT: pre-training 3D point cloud transformers with masked point modeling//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE: 19291-19300 [DOI: 10.1109/cvpr52688.2022.01871]
- Zhang W M, Wan P, Wang T J, Cai S S, Chen Y M, Jin X L and Yan G J. 2019. A novel approach for the detection of standing tree stems from plot-level terrestrial laser scanning data. *Remote Sensing*, 11(2): 211 [DOI: 10.3390/rs11020211]
- Zhao H S, Jiang L, Jia J Y, et al. 2021. Point transformer//Proceedings of the IEEE/CVF International Conference on Computer Vision. Montreal: IEEE: 16239-16248 [DOI: 10.1109/ICCV48922.2021.01595]

Zheng G, Ma L X, He W, Eitel J U H, Moskal L M and Zhang Z Y. 2016. Assessing the contribution of woody materials to forest angular gap fraction and effective leaf area index using terrestrial laser scanning data. *IEEE Transactions on Geoscience and Remote Sensing*, 54(3): 1475-1487 [DOI: 10.1109/tgrs.2015.2481492]

Zhu X, Skidmore A K, Darvishzadeh R, Niemann K O, Liu J, Shi Y F and Wang T J. 2018. Foliar and woody materials discriminated using terrestrial LiDAR in a mixed natural forest. *International Journal of Applied Earth Observation and Geoinformation*, 64: 43-50 [DOI: 10.1016/j.jag.2017.09.004]

PTV3-BLSeg: A branch-leaf segmentation model for fine-scale structural analysis of tree point clouds

ZHOU Shuyu, YAO Zongqi, ZHANG Xiaoli

1. State Key Laboratory of Efficient Production of Forest Resources, College of Forestry, Beijing Forestry University, Beijing 100083, China;

2. Beijing Key Laboratory of Precision Forestry, College of Forestry, Beijing Forestry University, Beijing 100083, China;

3. Key Laboratory of Forest Cultivation and Protection, Ministry of Education, Beijing Forestry University, Beijing 100083, China

Abstract: Tree point cloud segmentation plays a crucial role in forest structural analysis, 3D forest reconstruction, and the accurate estimation of tree parameters such as diameter at breast height, crown structure, and biomass. However, automatic branch-leaf separation from tree point clouds remains a challenging task. Major difficulties include the ambiguity of boundaries between trunks, branches, and foliage, the severe class imbalance caused by the significantly larger number of foliage points compared with that of trunk points, and the complex fine-scale geometric structures present in tree crowns. These challenges often lead to misclassification of trunk points as foliage and inaccurate representation of structural details, which ultimately limits the reliability of downstream forestry applications. Therefore, developing a robust and accurate segmentation method capable of effectively handling boundary ambiguity, class imbalance, and fine-grained structural representation is essential for improving the quality of tree point cloud analysis.

To address these issues, this study proposes Point Transformer V3 (PTV3)-BLSeg, an improved tree point cloud segmentation model based on the PTV3 architecture. The proposed model introduces three key improvements. First, fusion skip connections are incorporated into the decoder to strengthen cross-level feature propagation and enhance the interaction between shallow geometric features and deep semantic features, thereby improving the model's ability to capture local structural details and global contextual information. Second, a boundary-enhanced loss function is designed to emphasize points located near trunk-leaf boundaries by assigning them higher training weights. This strategy helps the model focus on subtle structural transitions and improves discrimination between adjacent classes. Third, a deep supervision mechanism is introduced by adding auxiliary supervision to intermediate layers, which guides the feature learning process at multiple depths, stabilizes network training, and improves the learning efficiency of trunk-related features that are typically underrepresented in the dataset. The proposed model was evaluated using three different tree point cloud datasets representing diverse forest environments: the Smart Tree Dataset, the Genhe Dataset from Inner Mongolia, and the Hengzhou Dataset from Guangxi.

Experimental results demonstrated that the proposed PTV3-BLSeg model consistently outperformed the baseline PTV3 model across all evaluation datasets. On the Smart Tree Dataset, PTV3-BLSeg achieved an overall accuracy of 93.33%, with trunk and leaf *F1*-scores of 75.61% and 96.09%, representing improvements of 1.36%, 4.14%, and 0.85%, respectively, compared with the baseline. On the Genhe dataset, the model achieved an overall accuracy of 83.83%, with trunk and leaf *F1*-scores of 60.77% and 89.50%, improving by 5.04%, 12.86%, and 9.13%, respectively. On the Hengzhou Dataset, the model obtained an overall accuracy of 81.87%, with trunk and leaf *F1*-scores of 71.32% and 86.60%, which corresponded to improvements of 4.49%, 12.73%, and 1.48%, respectively. These results indicate that the proposed model significantly enhances trunk recognition performance while maintaining high accuracy in leaf segmentation.

Overall, PTV3-BLSeg effectively addresses key challenges in tree point cloud segmentation, including boundary ambiguity, class imbalance, and the difficulty of recognizing fine-scale structures. By integrating fusion skip connections, a boundary-enhanced loss function, and a deep supervision strategy, this model improves segmentation accuracy and training stability. The consistent performance improvements observed across multiple datasets demonstrate the robustness and generalization capability of the proposed approach. This study provides an effective and reliable solution for high-precision tree point cloud segmentation and offers valuable technical support for advanced forestry applications such as 3D forest inventory, tree structure analysis, and ecological monitoring.

Key words: branch-leaf separation, Point Transformer V3, boundary enhancement, deep supervision, point cloud semantic segmentation, skip connection, feature fusion, 3D point clouds

Supported by National Key Research and Development Program of China (No. 2023YFD2201700); National Natural Science Foundation of China (No. 32171779)