

森林结构特征与大模型结合的地基点云单木定位

徐渝杰

重庆交通大学智慧城市学院, 重庆

收稿日期: 2025年2月17日; 录用日期: 2025年3月9日; 发布日期: 2025年3月18日

摘要

森林场景的地面激光扫描数据空间结构复杂, 冠层间郁闭度高, 导致传统森林资源调查方法耗费大量人力物力。随着计算机视觉技术的发展, 深度学习方法在交通、农业、林业等领域已取得显著成果。监督学习作为深度学习的核心范式, 依赖大量标注数据训练模型, 但其对数据质量和数量的要求较高, 在标注数据稀缺的场景中应用受限。相比之下, 基于自监督学习的大模型目标检测方法通过设计预训练任务, 能够从大量未标注数据中提取通用视觉特征, 具备强大的表征能力和泛化性能。这种方法不仅降低了对标注数据的依赖, 还能适应更加复杂、多样化的目标检测场景。因此, 针对现有问题, 本文依据森林空间特有结构提出了一种基于大模型DINO-X Pro的单木定位方法。首先针对森林点云特性设计空间特征, 将其输入网络, 确定森林单木候选框, 再根据候选框反向计算树干点。本文采用森林点云公开数据集进行单木定位实验, 基于本文提出的方法在多样地单木定位中召回率, 精确度, F_1 指数分别达到了84.1%, 88.9%, 86.4%, 与现有方法T-Rex对比均有较大提升, 分别为31.1%, 17.6%, 25.8%; 实验结果表明, 本文方法的单木定位性能好, 可靠性高, 能为森林资源调查提供有效的参考数据。

关键词

单木定位, 地面激光扫描, 深度学习, 特征工程

Single Tree Localization in Ground Laser Scanning Point Clouds Based on the Combination of Forest Structure Characteristics and Large Models

Yujie Xu

School of Smart City, Chongqing Jiaotong University, Chongqing

Received: Feb. 17th, 2025; accepted: Mar. 9th, 2025; published: Mar. 18th, 2025

文章引用: 徐渝杰. 森林结构特征与大模型结合的地基点云单木定位[J]. 图像与信号处理, 2025, 14(2): 109-117.
DOI: 10.12677/jisp.2025.142011

Abstract

The complex spatial structure of ground-based laser scanning data in forest scenes and the high canopy closure make traditional forest resource survey methods labor- and resource-intensive. With the advancement of computer vision technology, deep learning methods have achieved significant success in fields such as transportation, agriculture, and forestry. As a core paradigm of deep learning, supervised learning relies on large amounts of labeled data to train models. However, its high demand for data quality and quantity limits its application in scenarios where labeled data is scarce. In contrast, self-supervised learning-based large-model object detection methods can extract general visual features from vast amounts of unlabeled data by designing pre-training tasks, demonstrating strong representational power and generalization capabilities. This approach not only reduces dependence on labeled data but also enables models to adapt to more complex and diverse object detection scenarios. To address existing challenges, this paper proposes a single-tree localization method based on the large model DINO-X Pro, leveraging the unique spatial structure of forests. First, spatial features tailored to forest point cloud characteristics are designed and input into the network to determine candidate bounding boxes for individual trees. Then, trunk points are back-calculated based on the candidate boxes. This study conducts single-tree localization experiments using publicly available forest point cloud datasets. The proposed method achieves recall, precision, and F₁-score of 84.1%, 88.9%, and 86.4%, respectively, across various sample plots. Compared to the existing T-Rex method, these metrics show significant improvements of 31.1%, 17.6%, and 25.8%, respectively. The experimental results demonstrate that the proposed method offers high single-tree localization performance and reliability, providing valuable reference data for forest resource surveys.

Keywords

Single Tree Positioning, Terrestrial Laser Scanning (TLS), Deep Learning, Feature Engineering

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

森林资源的调查与监测是实现“生态文明”的必要工作，获取精确的森林资源分布信息对森林资源管理、碳储量统计、森林生态价值评估以及生态系统建模等方面具有重要意义[1]-[3]。地面激光扫描(Terrestrial Laser Scanning, TLS)具有高效率、高精度的优势[4]，能有效穿透森林内部并获取到完整的地面及林分的回波信息[5]，因此 TLS 技术已广泛用于森林资源调查[6]。

在森林 TLS 场景的解析中，单木作为获取有效林分特征的核心组成部分，十分重要，而单木的精确定位则是这一过程中的基础环节[2]。因此，树干检测成为了数据处理链条上的一个关键枢纽。单木定位策略的核心逻辑在于，通过分析森林点云数据的空间分布结构，提取森林信息，精准识别并提取出树木主干，或者识别到树冠中心，进而确定每一株单木的具体位置。在实现单木定位的过程中，存在多种有效的思路和方法。其中一种自上而下的方法的树冠高度模型(Canopy Height Model, CHM)，利用数字表面模型(Digital Surface Model, DSM)与数字高程模型(Digital Elevation Model, DEM)的差异，通过计算得到 CHM。在此基础上，采用动态窗口的局部最大值法，将局部高度最大值点作为树冠的顶点，从而探测到单木的位置[7]。Yang 等人在 CHM 粗分割的基础上，根据三维空间点云无序分布的特点，进行多方向的

空间分布分析, 细化潜在的树顶点位置, 有效地提高了树顶点检测的效率[8]。这类方法能够直观地反映出树冠的空间分布特征, 为单木定位提供了有力的支持。另一种则是通过检测树干, 确定单木位置。Wang 等人[9]提出了一种创新的树干位置检测方法, 该方法融合了平面密度分析与中值 z 法线值的应用, 通过构建两层投影机制, 有效提升了树干位置的检测精度。Chen 等人[10]则基于扫描密度与距离之间的关联, 设计了一种自适应邻域算法提取点云特征, 并结合支持向量机识别树干, 通过聚类分析与圆柱拟合技术, 实现了树干及分支的精细化建模, 确定了单木位置。Liang 等人[11]利用平面度与法向矢量方向作为特征描述符, 从复杂的点云数据中精确提取出树木主干点, 并借助圆柱拟合技术, 成功构建了树干模型, 从而实现了单木位置的准确判定。Chen 等人[12]通过建立系列体素特征, 对不同的特征组合进行评估, 筛选树干提取的最佳特征组合。

近年来, 深度学习技术作为一股新兴力量, 在林业领域的多个分支中得到了广泛应用。特别是在图像分类、目标检测与定位等任务上, 基于卷积神经网络(Convolutional Neural Networks, CNN)的深度学习方法展现出了卓越的性能[13]。自 2017 年起, Charles 等人提出的 PointNet [14]及其改进版本 PointNet++ [15]首次实现了面向点云的深度学习, 此后, 深度学习网络也被成功应用于森林场景的点云分类与单木提取。Xi 等人[16]通过对比不同方法对 TLS 树种数据的分类性能, 发现深度学习相较于传统机器学习, 在语义分割精度上具有显著优势。Wang 等人[17]则利用 Faster R-CNN 网络, 通过对树干点云的横向和正向投影图像进行树干检测, 通过反投影有效计算了单木位置。然而, 将三维点云转化为影响的过程中会空间信息损失不可避免。针对该问题, Chen[18]采用了 PointNet 对体素化后的森林点云进行分类, 并在分类结果的基础上确定了单木的位置, 降低了空间信息损失的影响。而 Windrim 等人[19]从冠层顶部出发, 首先对机载点云进行水平投影, 构建出二维影像, 然后利用 R-CNN 对影像数据进行训练, 实现机载点云的单木提取, 通过二维影像预测框(bounding box)来确定单木的位置。尽管该方法在单木提取上取得了一定的成效, 但监督训练所需的人工成本较高, 且耗时较长。

视觉语言大模型属于一类先进的生成式人工智能模型, 它们具备同时接收并处理视觉与文本信息的卓越能力, 能够进行高效的跨模态联合推理[20][21]。凭借其庞大的参数规模及海量的训练数据支撑, 这类模型能够以多种低成本策略, 包括微调(fine-tuning)、少样本提示(few-shot prompting)以及零样本提示(zero-shot prompting) [22], 灵活应对多种复杂的多模态下游任务, 其中, 医学图像分割领域[23]已经率先引入大模型。同样, 在森林场景中, 大模型的引入能够有效解决森林标注数据缺乏, 或者人工勾画成本大的问题。

基于上述分析, 现有的各类单木定位方法虽然能在如城市等简单场景中实现较好的单木定位效果, 但在森林等复杂场景下[24], 单木定位的精度往往会受到一定限制。同时, 现有的单木定位方法侧重于模型标注, 其需要人工标注, 不能实现快速定位。针对以上问题, 本研究通过分析森林场景特点, 利用几何特征增强对森林点云场景的表征, 再利用大模型实现单木候选框的确定, 再根据矩形候选框反向计算树干中心坐标, 得到单木位置。

2. 单木几何特征构建

现有的语义分割网络在训练和测试时通常需要输入除三维坐标外的单点特征信息, 然而由三维激光扫描仪直接提供的原始点云特征(如 XYZ 坐标、强度值、RGB 值等)不能满足高精度训练的需求[25]。为了提高单木定位的网络性能, 本章节设计出启发式森林特征, 依据森林特有的空间结构进一步提高单木定位的精度, 然后在候选框内计算各单木位置。

2.1. CSF 过滤地面点

由于采用地基激光雷达扫描, 地面点会被大量获取, 但其影响了森林点云的提取[26]。此外, 地面点

占比过高也使得深度学习网络对森林点云的识别效率降低。为实现后续单木分割任务,本文首先利用 CSF (Cloth Simulation Filter)布料滤波算法[27]对森林 TLS 数据进行地面点分离,将原始点云分为地面点集 P_g 与非地面点集 P_o 。

2.2. 森林点云特征

本研究以文献[28]提出的点云三维特征提取算法为基础,计算点云的离散度、各向异性、第一主成分。在此基础上,构建森林场景点云特征图谱。

$$F_1 = [f_c, f_a]^T \quad (1)$$

其中 $f_c = [X, Y, Z]$ 为点云空间坐标, $f_a = [P_a, C_a, S_a]$ 分别为离散度特征、各向异性特征、第一主成分特征的集合,见图 1。

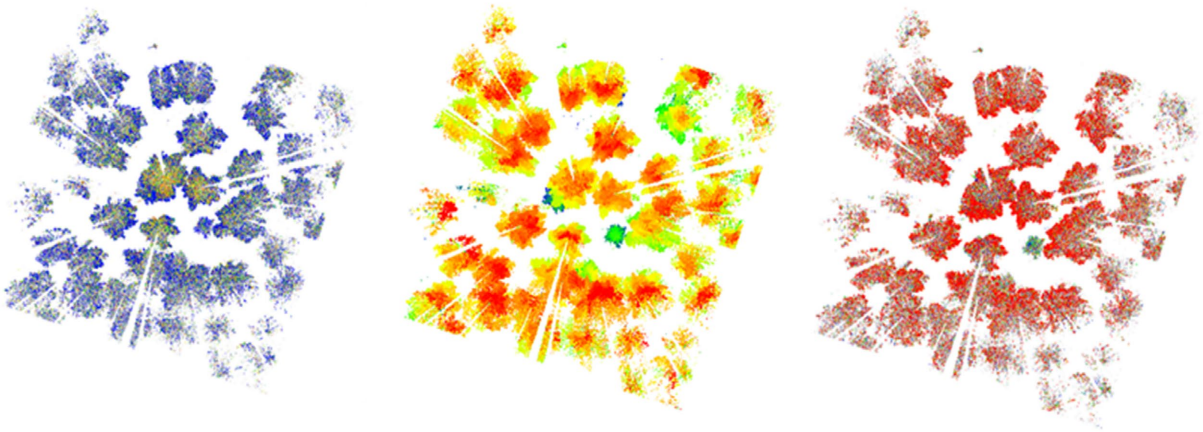


Figure 1. Feature values rendering
图 1. 特征渲染图

2.3. DINO-X 统一视觉模型

DINO-X[29]包括两种模型: DINO-X Pro 以及 DINO-X Edge 模型。前者,为各种场景提供增强的感知能力,后者针对更快的推理速度进行了优化,更适合部署在边缘设备上。

本文采用了 DINO-X Pro 模型,其对文本和视觉提示编码器进行了优化,并引入了自定义提示功能,以提升模型在多模态任务中的性能和灵活性,详见如下图 2。

DINO-X Pro 模型由文本提示编码器,视觉提示编码器,自定义提示包括三部分组成。

文本提示编码器。Grounding DINO [20]及 Grounding DINO 1.5 [21]原用 BERT [30]作文本编码器,其仅在纯文本数据训练,限制多模态感知任务表现。DINO-X Pro 引入预训练 CLIP [22]模型,其在多模态数据预训练,能更好捕捉文本与视觉语义关联,提升开放世界基准测试效率与性能。

视觉提示编码器。DINO-X Pro 中集成的 T-Rex² [31]的视觉提示编码器。通过框和点自定义视觉提示增强目标检测,利用位置嵌入和统一特征空间,结合多尺度注意力层提取特征,以为用户提供的视觉提示为条件,从多尺度特征图中提取视觉提示特征。

自定义提示。在实际应用中,针对特定场景对模型进行微调[32]的需求非常常见。为此, DINO-X Pro 定义了一系列专门的提示,称为定制提示。通过调优技术覆盖多样场景,如通用对象提示符支持无提示检测,用于支持无提示的开放世界检测,使模型能够检测图像中的任何对象。

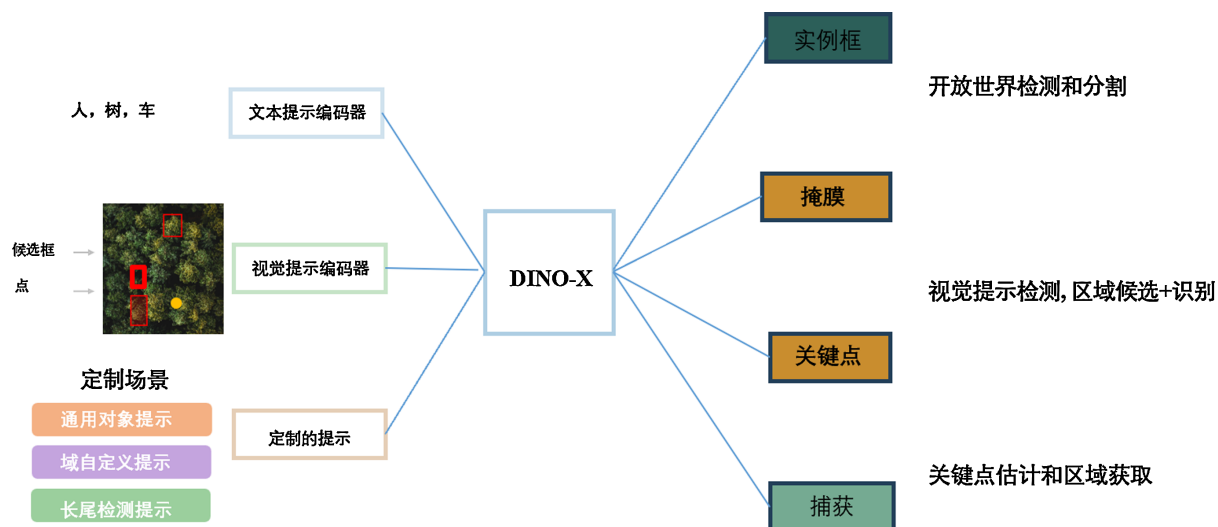


Figure 2. DINO-X network diagram

图 2. DINO-X 网络示意图

通过这些优化, DINO-X Pro 在多模态对齐和目标检测任务中表现出色, 能够更好地适应多样化的实际应用场景。

3. 树干定位

利用网络模型进行预测, 对各样地点云 P_i 按相同方式建立结果图, 通过预测各树干在特征图谱中的位置再进行反向投影来计算树干位置。图 3 为单木点云 3D 候选区构建与反向投影示意图。预测得到的边界框中心表示为 (x, y, w, h) , 其中 (x, y) 表示边界框中心的坐标, w 和 h 分别表示边界框的宽度和高度。基于边界框参数, 计算出单木位置 $P_i^k (x_k, y_k)$ 。

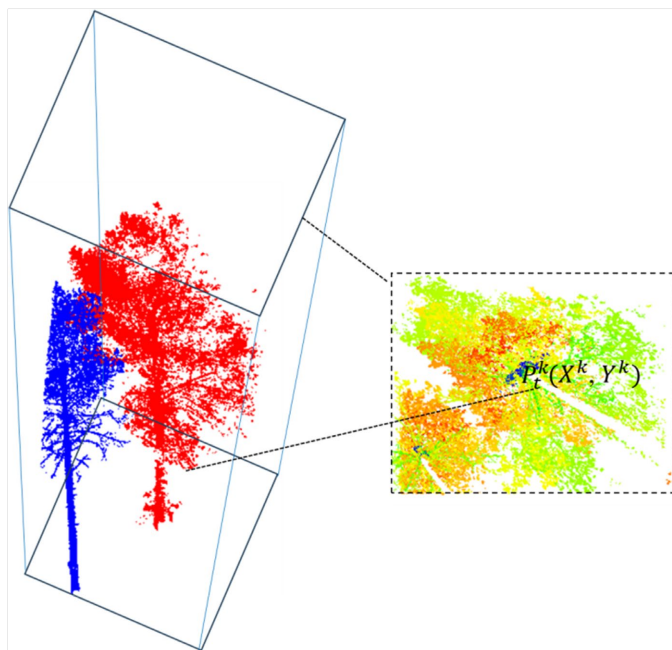


Figure 3. Construction and back-projection of trunk points bounding box

图 3. 树干点候选框构建与反向投影

4. 实验与结果分析

4.1. 实验数据

为验证本研究的可靠性, 本实验采用了 TLS 森林公共数据集[33]。数据集采集于芬兰湖地区的森林地带(61.19°N, 25.11°E), 样地内主要包含欧洲赤松、挪威云杉、银杉等植被。该数据集共包含 4 组样地, 以植被密度、林下植被数量等因素分为简单、中等两个难度等级, 图 4(a)、图(b)分别为简单、中等样地。每个难度等级包含两组样地, 每组样地提供单测站及多测站 TLS 数据, 以样地 1~4 表示。样地 1~2 为简单、3~4 为中等。由于公共数据集并未提供可供参考的语义信息及实例信息, 本研究使用 RiSCAN PRO 逐点标记语义及实例信息。使用样地 1~4 约 70%点云数据作为训练数据集, 剩余数据作为测试、验证数据集。

本实验使用 Tensorflow 与 PyTorch 框架作为深度学习网络训练平台。编程环境为 Python3.8, 操作系统为 Ubuntu 18.04。GPU 为 GeForce RTX 4090, CPU 为 Intel(R) Core(TM) i7-12700F, 内存大小为 64 G。

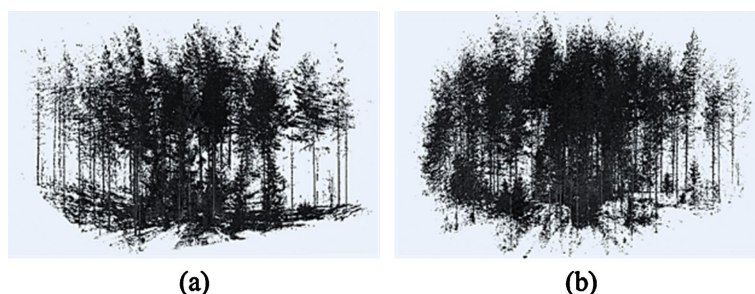


Figure 4. Forest point cloud intensity images for various plots
图 4. 各样地森林点云强度影像

4.2. 单木定位结果

本文对比了 T-Rex 的树干识别与定位精度。对比实验均采用相同网络配置参数与硬件条件进行训练。采用召回率(Recall)、精确度(Precision)、F₁ 指数等精度评估指标综合评判模型性能。

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3)$$

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

图 5 为单木定位结果。表 1 为本文树干检测各类别检出数量表, 其中 TP、FP、FN 分别为树干正确检出数、错误检出数、漏检数。表 2 为本文与 T-Rex 树干检测精度对比。在各种等级难度样地中, 本文方法与对比方法精度评估差值均较大, 其中召回率, 精确度, F₁ 指数均有较大提升, 分别为 31.1%, 17.6%, 25.8%。因此, 本文方法在不同场景下的树干提取精度较对比方法有一定优势。

6. 结论

本研究提出了一种基于地面激光扫描(TLS)数据的单木定位方法, 通过深度学习模型实现对森林场景中点云数据的精准分析。该方法以森林特有的三维空间结构为基础, 将点云的空间特征作为输入 DINO-X Pro 网络中, 显著提升了模型对树冠的感知能力, 从而精确计算出单木位置。实验结果表明, 与现有

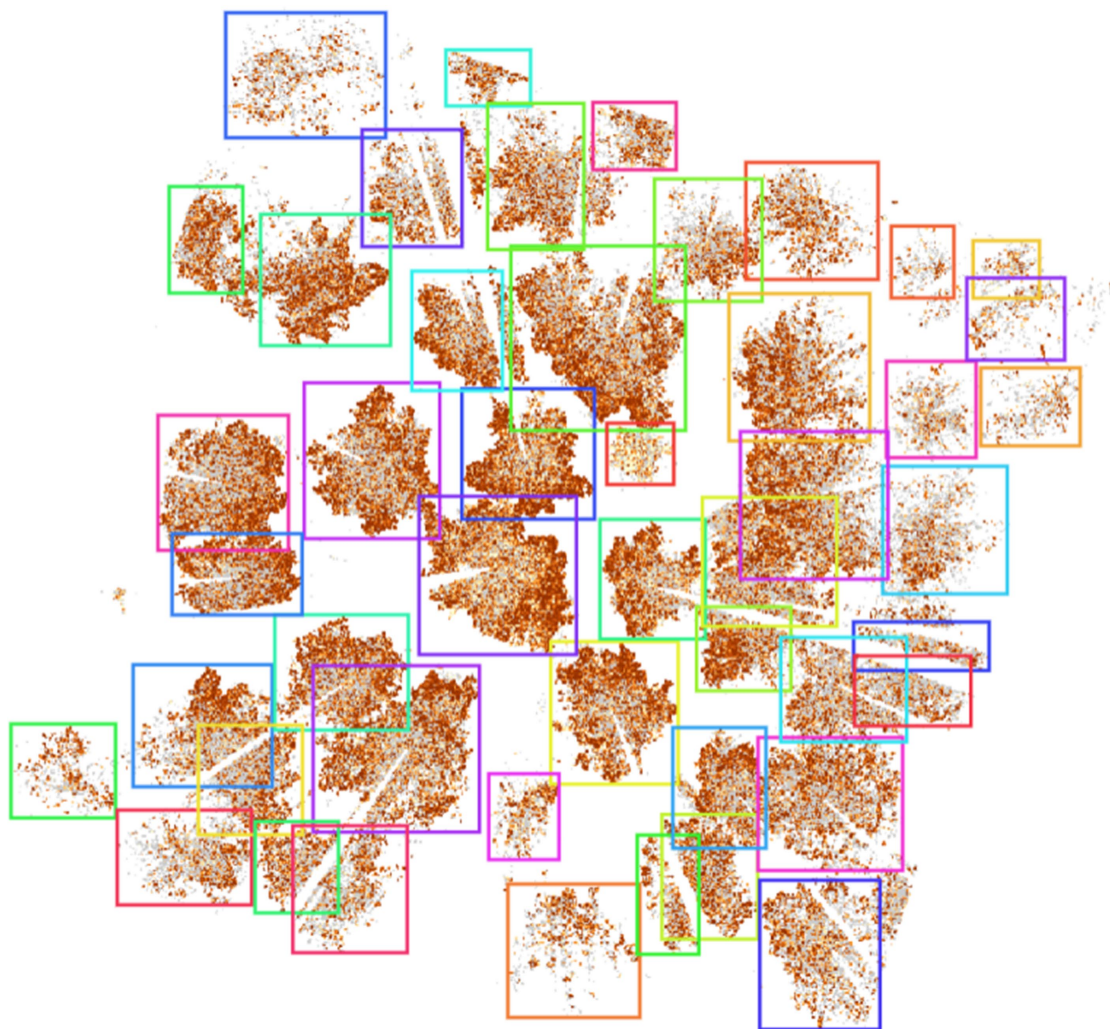


Figure 5. Neighbor tree points segmentation based on tree trunk localization

图 5. 基于单木定位结果

Table 1. Counts for each category of tree trunk detection

表 1. 树干检测各类别检出数量对比表

样地号	1	2	3	4
参考个数	46	48	81	79
TP	42	41	65	63
FP	5	4	10	8
FN	4	7	16	16
样地号	1	2	3	4
参考个数	46	48	81	79
TP	30	31	36	30
FP	10	8	20	15
FN	16	17	45	49

Table 2. Comparison of tree trunk detection accuracy
表 2. 树干检测精度对比

精度指标\样地号		1	2	3	4	均值
DINO-X Pro	召回率/%	91.3%	85.4%	80.2%	79.7%	84.1%
	精确度/%	89.3%	91.1%	86.6%	88.7%	88.9%
	F ₁ 指数	90.3%	88.1%	83.3%	84.0%	86.4%
T-Rex	召回率/%	65.2%	64.5%	44.4%	37.9%	53.0%
	精确度/%	75.0%	79.4%	64.2%	66.6%	71.3%
	F ₁ 指数	69.7%	71.3%	52.8%	48.8%	60.6%

方法相比，本方法在单木定位精度和鲁棒性方面均表现出显著优势，尤其在复杂森林样地中展现了良好的适应性，为森林资源调查提供了高精度的数据支持。未来研究将进一步优化特征工程，探索更具创新性的点云特征提取方法，如融合点云的语义信息，以进一步提升定位效率与模型性能。

基金项目

重庆市研究生导师团队建设项目(JDDSTD2022002)；重庆交通大学研究生科研创新项目资助(2024S0159)。

参考文献

[1] 曹林, 余光辉, 代劲松, 等. 激光雷达技术估测森林生物量的研究现状及展望[J]. 南京林业大学学报, 2013, 37(3): 163-169.

[2] 易静, 马开森, 向建平, 等. 点云切片结合聚类算法的 TLS 单木探测方法研究[J]. 南京林业大学学报(自然科学版), 2024, 48(4): 113-122.

[3] 刘开红, 汪琴. LiDAR 与传统测绘在森林资源补充调查中的应用对比分析[J]. 江西测绘, 2024(4): 16-19.

[4] Kuma, H., Fukuoka, H. and Komatsu, M. (2020) A Quantitative Analysis of Mine Mills by 3D Laser Scanner. *Materials Science Forum*, **983**, 73-80. <https://doi.org/10.4028/www.scientific.net/msf.983.73>

[5] Xu, D., Wang, H., Xu, W., Luan, Z. and Xu, X. (2021) Lidar Applications to Estimate Forest Biomass at Individual Tree Scale: Opportunities, Challenges and Future Perspectives. *Forests*, **12**, Article 550. <https://doi.org/10.3390/fl2050550>

[6] Chen, M., Xiao, L., Jin, Z., Pan, J., Mu, F. and Tang, F. (2023) Registration of Terrestrial Laser Scanning Data in Forest Areas Using Smartphone Positioning and Orientation Data. *Remote Sensing Letters*, **14**, 381-391. <https://doi.org/10.1080/2150704x.2023.2206974>

[7] 李响, 甄贞, 赵颖慧. 基于局域最大值法单木位置探测的适宜模型研究[J]. 北京林业大学学报, 2015, 37(3): 27-33.

[8] Yang, J., Kang, Z., Cheng, S., Yang, Z. and Akwensi, P.H. (2020) An Individual Tree Segmentation Method Based on Watershed Algorithm and Three-Dimensional Spatial Distribution Analysis from Airborne Lidar Point Clouds. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **13**, 1055-1067. <https://doi.org/10.1109/jstars.2020.2979369>

[9] Wang, D., Hollaus, M., Puttonen, E. and Pfeifer, N. (2016) Automatic and Self-Adaptive Stem Reconstruction in Landslide-Affected Forests. *Remote Sensing*, **8**, Article 974. <https://doi.org/10.3390/rs8120974>

[10] Chen, M., Wan, Y., Wang, M. and Xu, J. (2018) Automatic Stem Detection in Terrestrial Laser Scanning Data with Distance-Adaptive Search Radius. *IEEE Transactions on Geoscience and Remote Sensing*, **56**, 2968-2979. <https://doi.org/10.1109/tgrs.2017.2787782>

[11] Liang, X., Litkey, P., Hyypä, J., Kaartinen, H., Vastaranta, M. and Holopainen, M. (2012) Automatic Stem Mapping Using Single-Scan Terrestrial Laser Scanning. *IEEE Transactions on Geoscience and Remote Sensing*, **50**, 661-670. <https://doi.org/10.1109/tgrs.2011.2161613>

[12] Chen, M., Liu, X., Pan, J., Mu, F. and Zhao, L. (2023) Stem Detection from Terrestrial Laser Scanning Data with Features

- Selected via Stem-Based Evaluation. *Forests*, **14**, Article 2035. <https://doi.org/10.3390/f14102035>
- [13] Fan, X., Yang, X., Ye, Q. and Yang, Y. (2018) A Discriminative Dynamic Framework for Facial Expression Recognition in Video Sequences. *Journal of Visual Communication and Image Representation*, **56**, 182-187. <https://doi.org/10.1016/j.jvcir.2018.09.011>
 - [14] Qi, C.R., Su, H., Mo, K., et al. (2017) PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 77-85.
 - [15] Qi, C.R., Yi, L., Su, H., et al. (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, 4-9 December 2017, 5105-5114.
 - [16] Xi, Z., 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. <https://doi.org/10.1016/j.isprsjprs.2020.08.001>
 - [17] Wang, J., Chen, X., Cao, L., An, F., Chen, B., Xue, L., et al. (2019) Individual Rubber Tree Segmentation Based on Ground-Based Lidar Data and Faster R-CNN of Deep Learning. *Forests*, **10**, Article 793. <https://doi.org/10.3390/f10090793>
 - [18] Chen, X., Jiang, K., Zhu, Y., Wang, X. and Yun, T. (2021) Individual Tree Crown Segmentation Directly from Uav-Borne Lidar Data Using the PointNet of Deep Learning. *Forests*, **12**, Article 131. <https://doi.org/10.3390/f12020131>
 - [19] Windrim, L. and Bryson, M. (2020) Detection, Segmentation, and Model Fitting of Individual Tree Stems from Airborne Laser Scanning of Forests Using Deep Learning. *Remote Sensing*, **12**, Article 1469. <https://doi.org/10.3390/rs12091469>
 - [20] Liu, S., Zeng, Z., Ren, T., Li, F., Zhang, H., Yang, J., et al. (2024) Grounding DINO: Marrying DINO with Grounded Pre-Training for Open-Set Object Detection. In: Leonardis, A., Ricci, E., Roth, S., Russakovsky, O., Sattler, T. and Varol, G., Eds., *Computer Vision—ECCV 2024*, Springer, 38-55. https://doi.org/10.1007/978-3-031-72970-6_3
 - [21] Ren, T.H., Jiang, Q., Liu, S.L., Zeng, Z.Y., Liu, W.L., Gao, H., Huang, H.J., Ma, Z.Y., Jiang, X.K., Chen, Y.H., Xiong, Y.D., Zhang, H., Li, F., Tang, P.J., Yu, K. and Zhang, L. (2024) Grounding DINO 1. 5: Advance the “Edge” of Open-Set Object Detection. arXiv: 2405.10300.
 - [22] Xu, H., Xie, S.N., Ellen Tan, X.Q., et al. (2023) Emystifying CLIP Data. arXiv: 2309.16671.
 - [23] 孙兴, 蔡肖红, 李明, 等. 视觉大模型 SAM 在医学图像分割中的应用综述[J]. 计算机工程与应用, 2024, 60(17): 1-16.
 - [24] Chang, L., Fan, H., Zhu, N. and Dong, Z. (2022) A Two-Stage Approach for Individual Tree Segmentation from TLS Point Clouds. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **15**, 8682-8693. <https://doi.org/10.1109/jstars.2022.3212445>
 - [25] 肖龙. 基于深度学习的森林地面点云分类与单木分割研究[D]: [硕士学位论文]. 重庆: 重庆交通大学, 2024.
 - [26] Xiao, L., Chen, M., Zhang, X. and Liu, X. (2022). Classification of Forest Point Cloud Considering Relative Elevation and Change of Curvature Using Randla-Net. 2022 *IEEE International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC)*, Chongqing, 5-7 August 2022, 333-337. <https://doi.org/10.1109/sdpc55702.2022.9915816>
 - [27] Zhang, W., Qi, J., Wan, P., Wang, H., Xie, D., Wang, X., et al. (2016) An Easy-To-Use Airborne Lidar Data Filtering Method Based on Cloth Simulation. *Remote Sensing*, **8**, Article 501. <https://doi.org/10.3390/rs8060501>
 - [28] Weinmann, M., Jutzi, B., Hinz, S. and Mallet, C. (2015) Semantic Point Cloud Interpretation Based on Optimal Neighborhoods, Relevant Features and Efficient Classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing*, **105**, 286-304. <https://doi.org/10.1016/j.isprsjprs.2015.01.016>
 - [29] Ren, T.H., et al. (2024) DINO-X: A Unified Vision Model for Open-World Object Detection and Understanding. arXiv: 2411.14347.
 - [30] Devlin, J., Chang, M.W., Lee, K. and Toutanova, K. (2019) BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. *Proceeding of North American Chapter of the Association for Computational Linguistics*, Minneapolis, 2-7 June 2019, 4171-4186.
 - [31] Zhang, H., Li, F., Liu, S., Zhang, L., Su, H., Zhu, J., Ni, L. M. and Shum, H. (2022) DINO: DETR with Improved DeNoising Anchor Boxes for End-to-End Object Detection. arXiv: 2203.03605.
 - [32] Jiang, Q., Li, F., Zeng, Z., Ren, T., Liu, S. and Zhang, L. (2024) T-Rex2: Towards Generic Object Detection via Text-Visual Prompt Synergy. In: Leonardis, A., Ricci, E., Roth, S., Russakovsky, O., Sattler, T. and Varol, G., Eds., *Computer Vision—ECCV 2024*, Springer, 38-57. https://doi.org/10.1007/978-3-031-73414-4_3
 - [33] Liang, X., Hyypä, J., Kaartinen, H., Lehtomäki, M., Pyörälä, J., Pfeifer, N., et al. (2018) International Benchmarking of Terrestrial Laser Scanning Approaches for Forest Inventories. *ISPRS Journal of Photogrammetry and Remote Sensing*, **144**, 137-179. <https://doi.org/10.1016/j.isprsjprs.2018.06.021>