

Boundary-Constrained Supervoxel Clustering for Tree Segmentation in Broadleaf Forests

Chaoyong Wu^{1,2}, Shengjun Tang^{1,2}, Weixi Wang^{1,2*}

¹State Key Laboratory of Subtropical Building and Urban Science, China

²School of Architecture and Urban Planning, Shenzhen University, Shenzhen, China
wuchaoyong2023@email.szu.edu.cn, (shengjuntang, wangwx)@szu.edu.cn

Keywords: Airborne LiDAR data, Ultrahigh-resolution RGB imagery, Individual tree segmentation, Supervoxel clustering optimization, Subtropical broadleaf forests.

Abstract

Accurate segmentation of individual trees from Airborne Laser Scanning (ALS) point clouds is essential for urban greening management, ecological conservation, and biodiversity assessment. However, the complex canopy structures of subtropical broadleaf forests, often characterized by multiple peaks, pose significant challenges for existing segmentation algorithms, leading to prevalent over-segmentation. To address this issue, we propose a novel tree segmentation framework that integrates high-resolution RGB imagery with airborne LiDAR point clouds, enhancing the extraction of individual trees in subtropical broadleaf forests. Our method first employs high-resolution imagery to delineate canopy boundaries, which serve as constraints to refine the clustering of supervoxel-segmented point clouds. Furthermore, to mitigate both over- and under-segmentation, an optimization step is introduced based on geometric shape features of tree crowns. Experimental validation conducted in Shenzhen, China, demonstrates the effectiveness of our approach, achieving an average recall of 0.902, precision of 0.890, and F1-score of 0.906 across two study areas. Compared to conventional tree segmentation techniques, our method improves recall, precision, and F1-score by 9.6%, 12.2%, and 13.3%, respectively. These results highlight the advantages of integrating multi-modal remote sensing data for fine-grained tree segmentation in complex forest environments.

1. Introduction

As cities continue to expand, the management and protection of urban greening faces many challenges. Urban greening not only directly affects vegetation coverage, air quality and ecological diversity protection, but is also closely related to the physical and mental health of residents. Especially in the context of increasingly serious air pollution and urban heat island effects, the role of greening is becoming more and more important (Schusler et al., 2018, Ziter et al., 2019). Therefore, accurate and efficient detection and segmentation of single trees is of great significance for the refined management of urban greening, green development, and carbon emission monitoring, and provides technical support for promoting the high-quality and sustainable development of urban greening (O'Neil-Dunne et al., 2014).

Traditional urban greening surveys mainly rely on field ground measurements. However, this method is not only time-consuming and laborious, but also cannot be carried out manually in many places due to environmental conditions (Yang et al., 2022). With the rapid advancement of remote sensing technologies, remote sensing-based urban greening surveys offer a viable alternative. While remote sensing techniques have been widely applied in large-scale vegetation monitoring, the increasing demand for fine-scale urban greening information necessitates more precise tree-level remote sensing approaches.

The application of remote sensing technology in urban greening monitoring has gained traction, particularly for conducting repeatable and rapid assessments. Unmanned Aerial Vehicle (UAV) remote sensing, equipped with high-resolution cameras, captures detailed color information of vegetation, while Light Detection and Ranging (LiDAR) sensors penetrate canopy layers to acquire precise three-dimensional structural attributes of

vegetation (Masek et al., 2015). In recent years, these technologies have demonstrated substantial potential for urban vegetation surveys and monitoring.

Individual tree segmentation methods leveraging remote sensing data can generally be classified into two categories: approaches based on high-resolution imagery and those utilizing LiDAR data. High-resolution imagery captures intricate surface details, facilitating the extraction of spatial variability and texture features that enhance segmentation accuracy. Common techniques for image-based tree segmentation include region growing and watershed segmentation algorithms (Dalponte et al., 2018, Heenkenda et al., 2015). However, these methods often rely on manually set thresholds to control seed point expansion, leading to limited robustness (Ke and Quackenbush, 2011). On the other hand, LiDAR-based approaches capitalize on the technology's ability to acquire accurate vertical structural information, enabling the extraction of detailed forest features. These methods typically include canopy height model (CHM)-based watershed segmentation, region growing (Fang et al., 2016, Zhen et al., 2015), point cloud-based clustering, and deep learning-based point cloud segmentation (Zheng et al., 2024, Henrich et al., 2023, Hao et al., 2021, Latella et al., 2021). However, CHM-based segmentation techniques suffer from several limitations: (a) the smoothing process determines the number of detected trees, leading to inaccuracies; (b) understory trees are often undetectable; and (c) interpolation and smoothing introduce data loss. To address these issues, alternative approaches, such as direct point cloud-based clustering and region-growing algorithms, have been developed. Nonetheless, their segmentation accuracy remains low in broadleaf forests due to the presence of multiple canopy peaks (Yang et al., 2019).

Among segmentation strategies, normalized cut is a top-down

method that partitions data based on similarity, treating points or patches as graph vertices and using similarity weights to construct a graph for segmentation. This approach effectively mitigates under-segmentation issues in forests with distinct canopy peaks, such as coniferous and temperate forests (Lee et al., 2017). However, in complex forest structures, such as subtropical broadleaf forests with multiple canopy peaks, existing methods often result in over-segmentation. Currently, no comprehensive segmentation strategy has been specifically designed to address the challenges posed by these forest types. To bridge this gap, this study proposes a novel individual tree segmentation method tailored for subtropical broadleaf forests by integrating high-resolution imagery with airborne LiDAR point clouds. The method optimizes tree segmentation by refining supervoxel clustering results derived from airborne LiDAR data. The main contributions of this study are as follows:

- **Canopy Boundary-Constrained Supervoxel Clustering:** The proposed method leverages high-resolution imagery to detect canopy boundaries and incorporates these boundaries as constraints to refine the supervoxel clustering of LiDAR point clouds.
- **Clustering Optimization for Improved Segmentation:** To mitigate both under-segmentation and over-segmentation, the clustering results are further refined using geometric shape features of individual trees.

The remainder of this paper is organized as follows: Section II details the proposed methodology, including canopy boundary-constrained supervoxel clustering and optimization techniques. Section III presents experimental results and evaluates the segmentation performance in the study area. Section IV concludes the paper and discusses future research directions.

2. Method

In this study, UAV-based LiDAR and high-resolution RGB imagery are integrated for individual tree segmentation in subtropical broadleaf forests. The proposed method consists of three key stages: (1) data preprocessing, (2) canopy boundary-constrained supervoxel clustering, and (3) segmentation refinement. The overall framework is illustrated in Figure 1. In Stage 1, the raw airborne LiDAR point cloud undergoes denoising, normalization, and ground filtering to ensure data quality. In Stage 2, supervoxel segmentation is performed using the L0-Cut Pursuit algorithm (Landrieu and Obozinski, 2017), followed by tree crown boundary detection from high-resolution RGB imagery using the Detectree2 (Ball et al., 2023) model. The extracted tree crown boundaries serve as constraints to refine the clustering of supervoxel segments, enhancing segmentation accuracy. In Stage 3, tree crowns with irregular shapes from the initial segmentation are identified and refined through re-segmentation and clustering to mitigate under-segmentation and over-segmentation issues.

2.1 Data pre-processing

First, an progressive morphological filter is applied to separate the raw point cloud into ground and non-ground points. Based on the ground point data, a nearest neighbor interpolation method is used to generate a Digital Terrain Model (DTM). To eliminate the influence of the terrain, the Z-coordinate of each non-ground point is subtracted by the corresponding DTM

value, resulting in a normalized point cloud. For noise points present in the point cloud, surface cleaning is performed using the statistical outlier removal function in CloudCompare. Prior to tree segmentation evaluation, the point cloud for each individual tree is manually separated using CloudCompare, with ground and understory points removed. Additionally, to enhance the visual effect, each tree's point cloud is assigned a random color for easier differentiation and evaluation (Figure 2(b)). For the multi-view high-resolution imagery, we used Smart3D software to generate the corresponding orthophoto (Digital Orthophoto Map, DOM) of the study area, as shown in Figure 2(c).

2.2 Crown Boundary-Constrained Supervoxel Clustering

2.2.1 Crown Boundary Detection For canopy boundary detection in high-resolution imagery, this study employs the Detectree2 model. Detectree2 is based on the Mask R-CNN framework, an extension of Faster R-CNN, incorporating a branch for instance segmentation. The model achieves an F1-score of approximately 0.64 for canopy detection in high-resolution imagery. However, its performance is less effective under certain conditions, particularly when the trees have fewer leaves, low canopy color contrast, or closely spaced crowns. Moreover, due to the accuracy of the DOM imagery and the geometric correction issues when generating orthophotos from multi-view imagery, the detected canopy boundaries often fail to fully overlap with the corresponding tree crown point clouds, as shown in Figure 3.

2.2.2 Point Cloud Supervoxel Segmentation In graph cuts, a point cloud is often represented as a graph consisting of nodes V with attributes and edges E connecting these nodes. Thus, isolating or segmenting the point cloud in a graph cut can be represented by the disruption or cutting of the edges in the graph. Specifically, the graph G is a pair of sets $G = (V, E)$, where V is a set of N vertices and E is the set of edges. Each edge $w_{ij} \in E$ corresponds to a non-negative similarity weight between two vertices $i, j \in V$. The goal of binary graph cut is to partition the graph into two disjoint sets A and B by cutting the edges connecting these two sets, such that $A \cup B = V$ and $A \cap B = \emptyset$. The normalized cut method finds sets A and B by minimizing an energy term. By recursively applying normalized cut, this method can be extended to search for multiple categories, until the process is terminated by a stopping rule.

The L0-Cut Pursuit algorithm is a graph-cut algorithm that transforms the point cloud segmentation problem into a structured optimization problem of the graph. Various algorithms have been proposed for this graph structure optimization problem. The L0-Cut Pursuit algorithm, in particular, minimizes the objective function through graph structure regularization to iteratively cut the 3D point cloud into supervoxels with minimal total variation. Each graph cut is the minimization process of Equation 1, as follows:

$$\min_x \left(\sum_v \|X_{(v)} x_v\|^2 + \lambda \sum_{(u,v)} E_{uv} \varphi(X_u - X_v) \right) \quad (1)$$

where x is the optimization variable, v represents the nodes in the graph, X denotes the point coordinates of each node, E_{uv} represents the weight between two nodes, which is typically defined as the inverse of the distance between the two

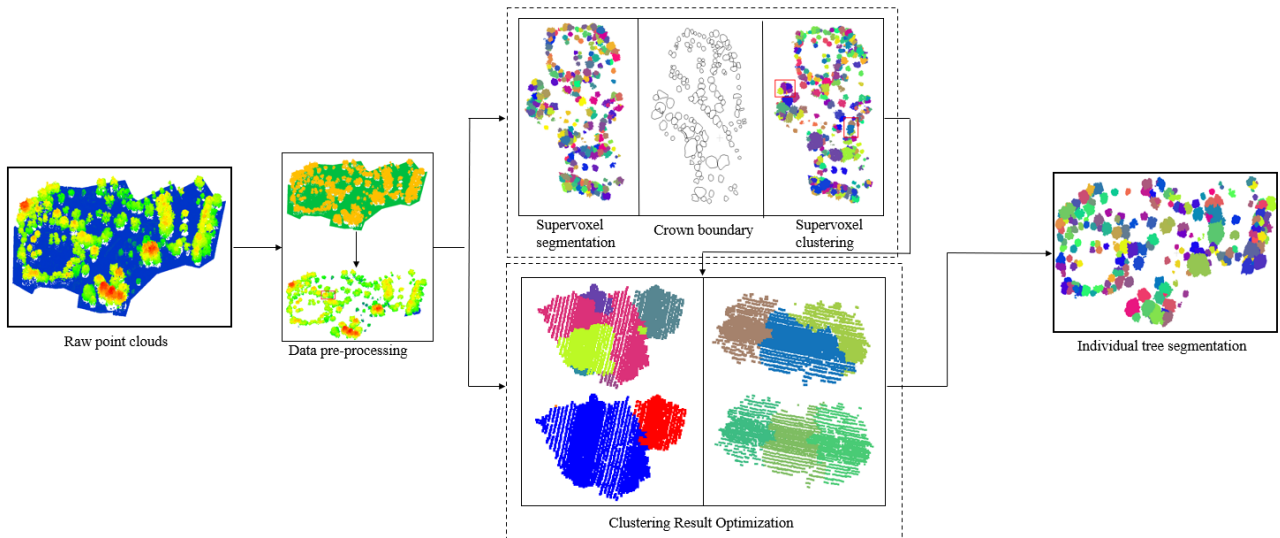


Figure 1. Pipeline of the proposed methodology.

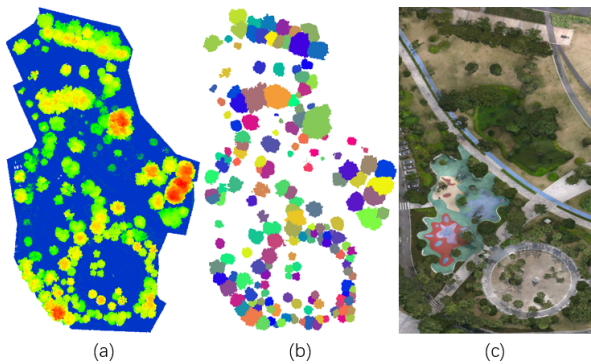


Figure 2. (a) Raw point cloud.(b) Result of manual individual tree segmentation.(c) Digital Orthophoto Map.



Figure 3. Detectree2 canopy boundary detection effect.(a) and (b) There are undetected canopy boundaries.(c) and (d) Detected canopy boundary position offset.(e) The canopy boundary detection is correct.

nodes. $\|X_{(v)}x_v\|^2$ represents the squared Euclidean distance between the node and the centroid.

In the L0-Cut Pursuit algorithm, the first term of the equation is the quadratic fidelity term, which measures the difference

between the current segmentation (i.e., label assignment) and the actual observations. The second term, $\varphi(X_u - X_v)$, represents the edge-based regularization term, which evaluates the differences between adjacent nodes. The regularization parameter λ controls the balance between fidelity and regularization, determining the size of the resulting supervoxels after segmentation. Figure 4 shows the supervoxel segmentation results under different regularization parameters. It can be observed that as the regularization parameter increases, the size of the resulting supervoxels also increases. During supervoxel segmentation of the point cloud data, certain individual trees can be directly clustered, as shown in Figure 4(e). This approach helps compensate for the reduced segmentation accuracy resulting from the limited precision of the Detectree2 model in detecting tree crown boundaries.

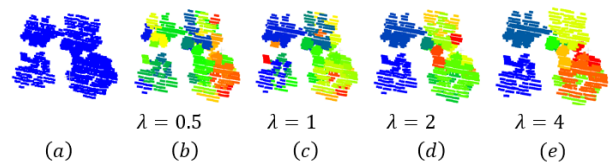


Figure 4. Supervoxel segmentation results for different regularization parameters. (a) Original tree point cloud. (b)-(e) Supervoxel segmentation results for regularization parameters of 0.5, 1, 2 and 4.

2.2.3 Supervoxel clustering If the airborne point cloud is directly constrained for clustering based on the obtained tree crown boundaries, although some crown boundaries can be correctly matched with the corresponding point clouds, most detected crown boundaries exhibit positional deviations from the airborne point cloud. This affects the accuracy of tree segmentation, as shown in Figure 5, where the precision of the segmentation results is relatively poor.

To address the issue of positional deviations between canopy boundaries and the airborne point cloud, this study proposes a canopy boundary-constrained point cloud supervoxel clustering method. Specifically, this method improves the clustering accuracy by merging supervoxels whose voxel centers lie within the same canopy boundary, as shown in Figure 6. This approach

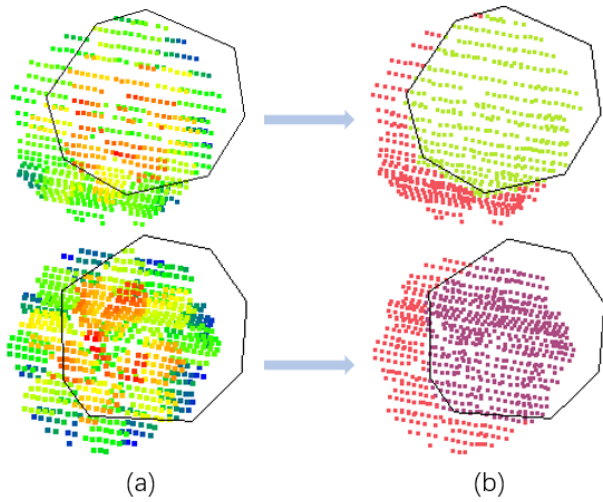


Figure 5. Direct point cloud clustering based on boundary constraints.(a)Raw point cloud overlap with shp.(b)Clustering results

effectively compensates for the decreased clustering performance caused by inaccuracies in the canopy boundary detection, significantly enhancing both the accuracy and stability of the segmentation process.

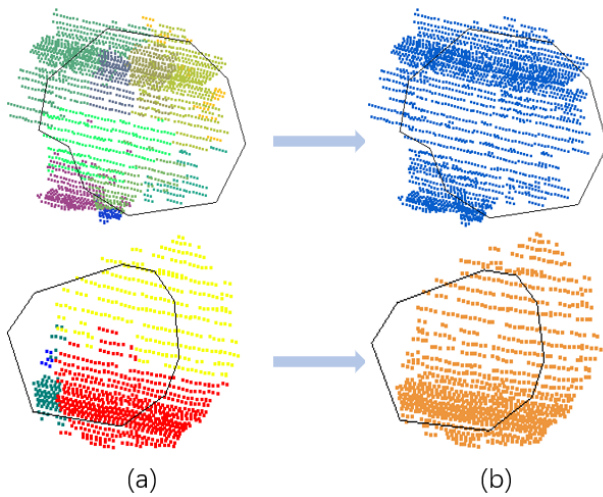


Figure 6. Supervoxels clustering based on boundary constraints.(a) Supervoxel results overlap with shp.(b) Clustering results.

2.3 Supervoxel clustering Result Optimization

Due to the nature of the L0-Cut Pursuit algorithm, which tends to result in under-segmentation for small and tightly packed tree crowns, as well as over-segmentation for large tree crowns. Furthermore, the precision of the Detectree2 model in canopy boundary detection can lead to several issues: a single canopy boundary may encompass multiple tree point clouds, or a canopy boundary may fail to fully cover the corresponding supervoxel centers of individual trees. These issues exacerbate both under-segmentation and over-segmentation problems in the clustering results. To address these under-segmentation and over-segmentation issues in the point cloud, this study proposes a solution based on the shape features of the individual tree segmentation results. By extracting shapes unreasonable

tree point clouds and optimizing the clustering results, the proposed method improves the segmentation accuracy and effectively avoids the occurrence of both under-segmentation and over-segmentation.

Since most tree species in subtropical broadleaf forests have crown projections that are nearly circular in the xy plane, this study proposes a parameter based on the circularity of the crown projection to extract unreasonable crowns. This includes parameters R_1, R_2 and R_3 , with the parameter expressions given in Equation 2:

$$\begin{aligned} r_1 &= \sqrt{\frac{S}{\pi}} \\ R_1 &= \frac{4\pi S}{C^2} \\ R_2 &= \frac{r_1}{r_2} \\ R_3 &= \frac{r_3}{r_4} \end{aligned} \quad (2)$$

Where, S is the crown projection area, C is the perimeter of the crown projection boundary, r_2 is the radius of the circum-circle, r_3 is the shortest distance from the center to the boundary, r_4 is the longest distance from the center to the boundary. Based on the calculations from Equation 2, three thresholds for R_1, R_2 and R_3 were set as 0.6, 0.75, and 0.55, respectively, to extract shapes unreasonable crowns from the clustering results. Figure 7 illustrates the parameter values corresponding to different shapes unreasonable crowns.

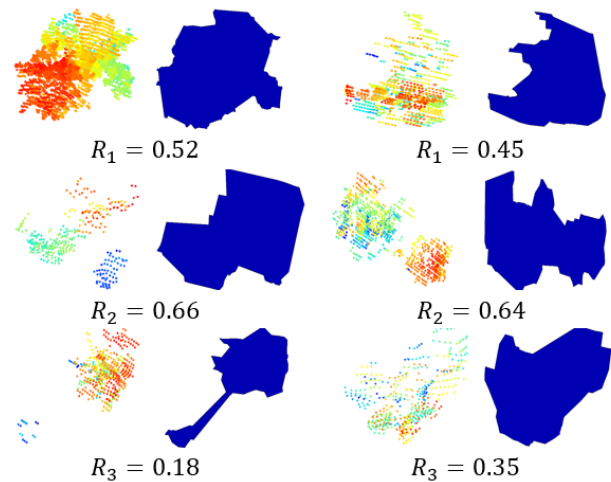


Figure 7. Parameter values corresponding to tree crowns with unreasonable shapes

After extracting shapes unreasonable crowns, this study uses the L0-Cut Pursuit supervoxel segmentation algorithm to further segment these unreasonable crowns. Since the over-segmented point clouds are generated by the L0-Cut Pursuit supervoxel segmentation algorithm, re-segmentation of these over-segmented point clouds does not have a significant impact. Next, for each individual tree segmentation result, unreasonable crowns are extracted again, and each unreasonable crown is merged with its k neighboring crowns. During the merging process, the parameter values of each merged crown are calculated using Equation 2, and the merging result with the highest sum of R_1, R_2 and R_3 parameters is selected. This process is

repeated until the overlap ratio of the point clouds of the extracted unreasonable crowns between two consecutive iterations exceeds 90%.

2.4 Accuracy Evaluation

The accuracy of individual tree segmentation results is evaluated using three measures: Recall, Precision, and F-score (Qin et al., 2022). Specifically, if a tree exists in the ground reference and is correctly segmented, it is considered a true positive (TP). If the overlap between the predicted tree and the actual tree is greater than 50%, it is deemed correctly segmented. If a tree exists in the ground reference but is not segmented, it is considered a false negative (FN). If a tree does not exist in the ground reference but is segmented, it is considered a false positive (FP). TP, FP, and FN represent correct segmentation, over-segmentation, and under-segmentation, respectively. Precision represents the ratio of true positives to all detected trees, while Recall represents the ratio of true positives to all ground reference trees. F-score is the overall accuracy considering both omission and false detection.

3. Experiments and Results

3.1 Dataset Descriptions

The study area is located in Shenzhen Bay Park, Shenzhen, China, with geographic coordinates of 113°56' 26" E, 22°30' 43" N. This area belongs to the subtropical broadleaf forest ecosystem and is characterized by high tree species diversity. The dominant species include coconut trees, flame trees, fan palms, and mangroves. Figure 8 illustrates the study area, which is divided into two sub-areas. Area 1 features a relatively regular arrangement of trees, with small differences in canopy size and fewer species. Area 2, on the other hand, has a denser arrangement of trees, with larger differences in canopy size and greater species diversity.

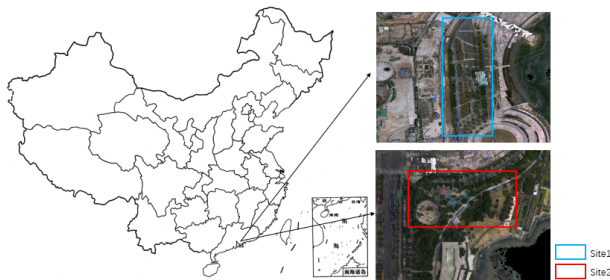


Figure 8. The location of the study area in ShenZhenWan Park, Shenzhen City of southern China.

The unmanned aerial vehicle (UAV) LiDAR data used in this study were acquired in 2024 using the RIEGL VQ-1560i-DW airborne LiDAR scanner. This device is a lightweight laser scanner specifically designed for UAVs, offering high spatial resolution and adaptability. During the data acquisition process, the UAV flew at a constant speed of 190 km/h at an altitude of 450 meters, with the resulting airborne point cloud data density approximately 45 points/m². Simultaneously, high-resolution imagery was captured by the UAV using the S4500_X_40 camera in 2024, providing multi-view images with a resolution of 2 cm.

3.2 Evaluation of Tree Segmentation

Figure 9 shows the results of individual tree segmentation for the two study areas. It can be observed that most of the trees in both regions are correctly segmented. In Region 1, where the tree arrangement is relatively simple, the segmentation results (Figure 9(a)) show that nearly all trees are accurately segmented, with no noticeable over-segmentation or under-segmentation. The boundaries between adjacent trees are clearly visible, and the individual tree contours and detailed features are well preserved. In Region 2, where the tree arrangement is more complex, the segmentation results (Figure 9(b)) indicate that some closely spaced trees and larger trees with wider crowns exhibit over-segmentation. However, the majority of trees are still accurately segmented.

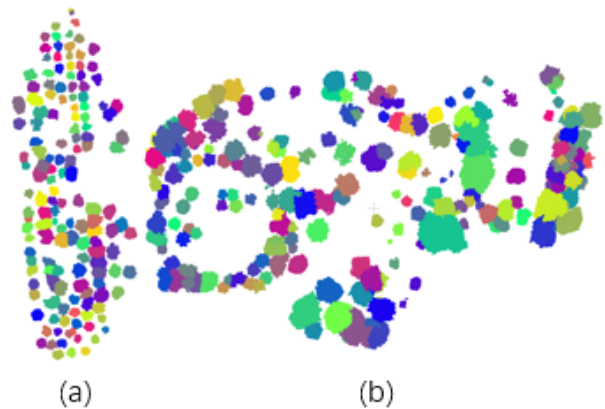


Figure 9. Individual tree segmentation results.(a) Study area 1 individual tree segmentation results.(b) Study area 2 individual tree segmentation results.

To further assess the accuracy of the tree segmentation results, this study conducted a precision evaluation for both study areas, with the results shown in Table 1. In Area 1, the precision, recall, and F1-score of tree segmentation were 0.927, 0.957, and 0.942, respectively, demonstrating high segmentation accuracy. In contrast, the precision, recall, and F1-score for tree segmentation in Area 2 were 0.852, 0.847, and 0.849, representing a 9.3% decrease in the overall F1-score compared to Area 1. Specifically, in terms of precision, Area 2 showed a 7.5% reduction in precision compared to Area 1. This difference is mainly attributed to the denser arrangement of tree crowns in Area 2, leading to more under-segmentation issues. Regarding recall, Area 2 experienced an 11% decrease compared to Area 1. This difference is primarily due to the misclassification of small, tightly clustered trees as a single large tree during segmentation.

Method	Site	Prec	Rec	F1
Proposed	Site 1	0.927	0.957	0.942
	Site 2	0.852	0.847	0.849

Table 1. The accuracy assessment results of individual tree segmentation.

3.3 Comparative Analysis

To evaluate the effectiveness of the proposed method, this study compared it with several classic traditional tree segmentation methods, including Li 2012 (Li et al., 2012), PyCrown (Zörner et al., 2018), Watershed (Zhao et al., 2014), Ams3d (Ferraz et al., 2016), Silva (Silva et al., 2016), and MCRC (Lee et al.,

2017). Figure 10 illustrates the partial segmentation results of six methods in Study Area 1 and Study Area 2. As shown in Table 2 and Table 3, the proposed method achieves the following precision, recall, and F1 scores in the two study areas: Area 1: 0.866, 0.795, 0.829; Area 2: 0.927, 0.957, 0.942. Overall, regardless of whether the trees in Area 1 are more regularly spaced with fewer species or the trees in Area 2 are more densely arranged with higher species diversity, the proposed method outperforms the other methods in terms of tree segmentation accuracy.

In terms of precision, the proposed method’s precision is second only to the watershed algorithm. Compared to other methods excluding the watershed algorithm, it achieves improvements of 4.4% to 17.1% in Area 1 and 5.7% to 34.3% in Area 2. This indicates that most of the tree crowns segmented by the proposed method are correct. On the other hand, the lower precision of Ams3d is likely due to its use of a mean-shift algorithm, which was developed for coniferous forests and often detects multiple tree apex points for a single tree in broadleaf forests, resulting in significant over-segmentation. In terms of recall, the proposed method’s recall is second only to the MCRC algorithm. Compared to other methods excluding the MCRC algorithm, it achieves improvements of 1.1% to 22.7% in Area 1 and 6.5% to 64.7% in Area 2. This indicates that the method is better at identifying the true tree crowns and reducing omissions. The Watershed algorithm has the lowest recall, suggesting a significant under-segmentation problem that fails to capture most of the tree crowns. The overall F1-score, which balances precision and recall, is the highest for the proposed method. Compared to other tree segmentation methods, the F1-score improves by 4.0%-10.6% in Area 1 and 7.8%-52.1% in Area 2, demonstrating the best balance between precision and recall.

Method	Metrics		
	Pre	Rec	F1
Li2012	0.807	0.920	0.860
PyCrown	0.883	0.918	0.900
Watershed	0.990	0.730	0.836
Ams3d	0.756	0.946	0.844
Silva	0.879	0.914	0.896
MCRC	0.851	0.961	0.902
Proposed	0.927	0.957	0.942

Table 2. Segmentation accuracy of each method in area 1.

Method	Metrics		
	Pre	Rec	F1
Li2012	0.693	0.782	0.735
PyCrown	0.795	0.718	0.755
Watershed	0.882	0.200	0.328
Ams3d	0.509	0.760	0.610
Silva	0.768	0.662	0.712
MCRC	0.702	0.855	0.771
Proposed	0.852	0.847	0.849

Table 3. Segmentation accuracy of each method in area 2.

3.4 Module evaluation

3.4.1 Comparison of clustering effects before and after point cloud supervoxel segmentation To verify the impact of supervoxel segmentation on the final tree segmentation results when performing point cloud clustering based on tree crown boundaries, this study compares the precision of two tree segmentation methods. The first method clusters the point cloud data directly based on the detected tree crown boundaries,

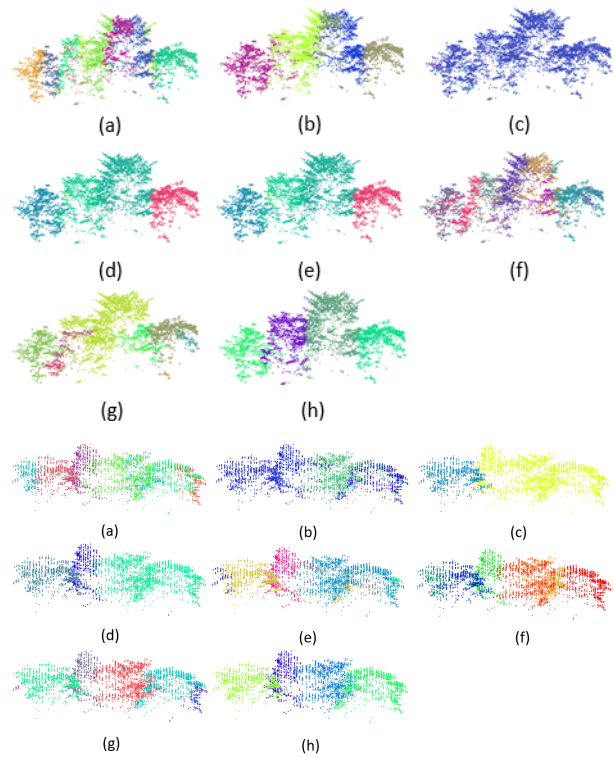


Figure 10. Three-dimensional visualization of partial segmentation results in Study Area 1 and Study Area 2. (a) Li2012 segmentation results. (b) PyCrown segmentation results. (c) Watershed segmentation results. (d) Ams3d segmentation results. (e) Silva segmentation results. (f) MCRC segmentation results. (g) Proposed method segmentation results. (h) Groundtruth.

while the second method first performs supervoxel segmentation on the point cloud and then clusters the supervoxel segmentation results based on the tree crown boundaries. The tree segmentation results and their precision for both methods are shown in Figure 11 and Table 4.

Site	Method	Prec	Rec	F1
Site1	Before clustering	0.924	0.688	0.788
	After clustering	0.870	0.901	0.885
Site2	Before clustering	0.750	0.614	0.675
	After clustering	0.664	0.824	0.735

Table 4. The accuracy assessment results of individual tree segmentation.

According to the results in Figure 11 and Table 4, when comparing direct clustering based on detected tree crown boundaries with first performing supervoxel segmentation and then clustering based on tree crown boundaries, the first method shows higher precision by 5.4% and 8.6% in the two study areas, respectively, while its recall is 21.3% and 21% lower. This difference is primarily due to the denser tree arrangement in certain areas, where some tree crown boundaries were not detected or where a single tree crown boundary encompassed multiple trees. On the other hand, in simpler areas, tree crown boundaries could be more accurately detected, leading to a higher precision for the first method, but a sharp decline in recall. However, in terms of the overall evaluation metric, F1-score, the second method outperforms the first in both study areas, with improve-

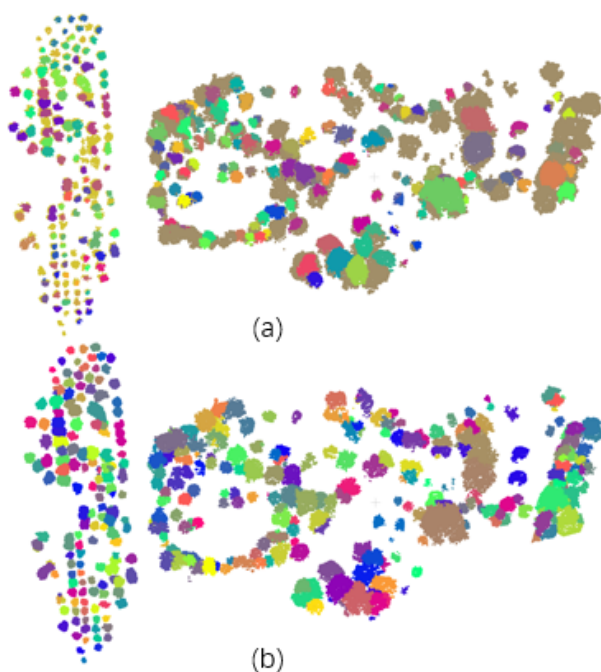


Figure 11. Individual tree segmentation result.(a) Direct point cloud clustering based on canopy boundaries.(b) Point cloud supervoxel segmentation results are clustered based on canopy boundaries (without clustering optimization).

ments of 9.7% and 6%, respectively.

3.4.2 Comparison of clustering results before and after optimization To verify the effectiveness of clustering optimization on the tree segmentation results, this study compares the segmentation precision of two methods. The first method clusters the supervoxel segmentation results based on detected tree crown boundaries without performing clustering optimization, while the second method clusters the supervoxel segmentation results based on tree crown boundaries and subsequently applies clustering optimization. The tree segmentation results and their precision for both methods are shown in Figure 12 and Table 5.

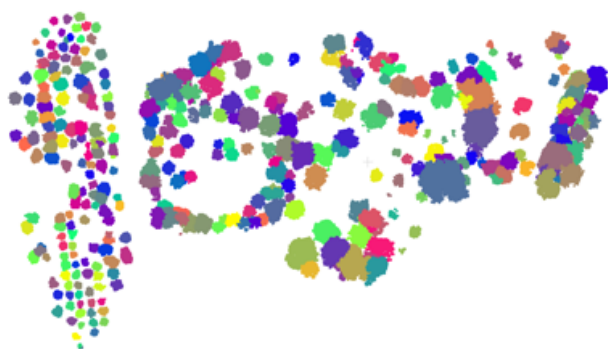


Figure 12. Individual tree segmentation results for clustering optimization.

The results from Figure 12, Figure 11 (b) and Table 5 reveal that, compared to individual tree segmentation results without clustering optimization, the segmentation accuracy after clustering optimization shows significant improvements in both study areas. Specifically, in Region 1 and Region 2, precision

Site	Method	Prec	Rec	F1
Site1	Before optimization	0.870	0.901	0.885
	After optimization	0.927	0.957	0.942
Site2	Before optimization	0.664	0.824	0.735
	After optimization	0.852	0.847	0.849

Table 5. The accuracy assessment results of individual tree segmentation.

increased by 5.7% and 18.8%, recall increased by 5.6% and 2.3%, and the overall evaluation metric, F1-score, improved by 5.7% and 11.4%, respectively. The improvement in precision in Area 1 is primarily due to the re-segmentation of densely packed trees that were previously under-segmented. In Area 2, the precision gain results from the re-segmentation of closely spaced small trees and the effective clustering of over-segmented large tree crowns.

4. Conclusion

This study presents a novel individual tree segmentation method that integrates airborne LiDAR point clouds with high-resolution imagery. The proposed approach was evaluated in two study areas with varying complexity and benchmarked against several traditional single-tree segmentation algorithms. Experimental results demonstrate the effectiveness of our method in subtropical broadleaf forests, achieving recall rates of 0.957 and 0.847, precision rates of 0.927 and 0.852, and F1-scores of 0.942 and 0.849. By leveraging the complementary strengths of LiDAR-derived vertical structural features and high-resolution RGB imagery-based texture features, our method significantly mitigates both under-segmentation and over-segmentation issues, leading to more accurate segmentation outcomes. Compared to traditional tree segmentation techniques, our approach achieves an average improvement of 9.6%, 12.2%, and 13.3% in recall, precision, and F1-score, respectively.

Despite these advancements, certain limitations remain. The accuracy of tree crown boundary extraction from high-resolution imagery can still be improved, and performance decreases in densely packed forest regions where tree crowns overlap. Future research will focus on refining boundary detection techniques and optimizing segmentation strategies to further enhance accuracy and robustness across diverse forest environments.

Acknowledgements

This work was supported in part by the National Key Research and Development Program of China (Project No. 2022YFB3903700), National Natural Science Foundation of China (Project Nos. 42471442, 42171265, 42371431), Research Project of Shenzhen S and T Innovation Committee (Project No. KJZD20230923115508017), Natural Science Foundation of Guangdong Province (Project No. 2024A1515030061) and Research project of State Key Laboratory of Subtropical Building and Urban Science (Project No. 2023ZB18).

References

Ball, J. G., Hickman, S. H., Jackson, T. D., Koay, X. J., Hirst, J., Jay, W., Archer, M., Aubry-Kientz, M., Vincent, G., Coomes, D. A., 2023. Accurate delineation of individual tree crowns in

- tropical forests from aerial RGB imagery using Mask R-CNN. *Remote Sensing in Ecology and Conservation*, 9(5), 641–655.
- Dalponte, M., Frizzera, L., Ørka, H. O., Gobakken, T., Næsset, E., Gianelle, D., 2018. Predicting stem diameters and above-ground biomass of individual trees using remote sensing data. *Ecological Indicators*, 85, 367–376.
- Fang, F., Im, J., Lee, J., Kim, K., 2016. An improved tree crown delineation method based on live crown ratios from airborne LiDAR data. *GIScience & Remote Sensing*, 53(3), 402–419.
- Ferraz, A., Saatchi, S., Mallet, C., Meyer, V., 2016. Lidar detection of individual tree size in tropical forests. *Remote Sensing of Environment*, 183, 318–333.
- Hao, Y., Widagdo, F. R. A., Liu, X., Liu, Y., Dong, L., Li, F., 2021. A hierarchical region-merging algorithm for 3-D segmentation of individual trees using UAV-LiDAR point clouds. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–16.
- Heenkenda, M. K., Joyce, K. E., Maier, S. W., 2015. Mangrove tree crown delineation from high-resolution imagery. *Photogrammetric Engineering & Remote Sensing*, 81(6), 471–479.
- Henrich, J., van Delden, J., Seidel, D., Kneib, T., Ecker, A. S., 2023. Treelearn: A comprehensive deep learning method for segmenting individual trees from forest point clouds. *CoRR*.
- Ke, Y., Quackenbush, L. J., 2011. A review of methods for automatic individual tree-crown detection and delineation from passive remote sensing. *International Journal of Remote Sensing*, 32(17), 4725–4747.
- Landrieu, L., Obozinski, G., 2017. Cut pursuit: Fast algorithms to learn piecewise constant functions on general weighted graphs. *SIAM Journal on Imaging Sciences*, 10(4), 1724–1766.
- Latella, M., Sola, F., Camporeale, C., 2021. A density-based algorithm for the detection of individual trees from LiDAR data. *Remote Sensing*, 13(2), 322.
- Lee, J., Coomes, D., Schonlieb, C.-B., Cai, X., Lellmann, J., Dalponte, M., Malhi, Y., Butt, N., Morecroft, M., 2017. A graph cut approach to 3D tree delineation, using integrated airborne LiDAR and hyperspectral imagery. *arXiv preprint arXiv:1701.06715*.
- Li, W., Guo, Q., Jakubowski, M. K., Kelly, M., 2012. A new method for segmenting individual trees from the lidar point cloud. *Photogrammetric Engineering & Remote Sensing*, 78(1), 75–84.
- Masek, J. G., Hayes, D. J., Hughes, M. J., Healey, S. P., Turner, D. P., 2015. The role of remote sensing in process-scaling studies of managed forest ecosystems. *Forest Ecology and Management*, 355, 109–123.
- O’Neil-Dunne, J., MacFaden, S., Royar, A., 2014. A versatile, production-oriented approach to high-resolution tree-canopy mapping in urban and suburban landscapes using GEOBIA and data fusion. *Remote sensing*, 6(12), 12837–12865.
- Qin, H., Zhou, W., Yao, Y., Wang, W., 2022. Individual tree segmentation and tree species classification in subtropical broadleaf forests using UAV-based LiDAR, hyperspectral, and ultrahigh-resolution RGB data. *Remote Sensing of Environment*, 280, 113143.
- Schusler, T., Weiss, L., Treering, D., Balderama, E., 2018. Research note: Examining the association between tree canopy, parks and crime in Chicago. *Landscape and Urban Planning*, 170, 309–313.
- Silva, C. A., Hudak, A. T., Vierling, L. A., Loudermilk, E. L., O’Brien, J. J., Hiers, J. K., Jack, S. B., Gonzalez-Benecke, C., Lee, H., Falkowski, M. J. et al., 2016. Imputation of individual longleaf pine (*Pinus palustris* Mill.) tree attributes from field and LiDAR data. *Canadian journal of remote sensing*, 42(5), 554–573.
- Yang, M., Zhou, X., Liu, Z., Li, P., Tang, J., Xie, B., Peng, C., 2022. A review of general methods for quantifying and estimating urban trees and biomass. *Forests*, 13(4), 616.
- Yang, Q., Su, Y., Jin, S., Kelly, M., Hu, T., Ma, Q., Li, Y., Song, S., Zhang, J., Xu, G. et al., 2019. The influence of vegetation characteristics on individual tree segmentation methods with airborne LiDAR data. *Remote Sensing*, 11(23), 2880.
- Zhao, D., Pang, Y., Li, Z., Liu, L., 2014. Isolating individual trees in a closed coniferous forest using small footprint lidar data. *International journal of remote sensing*, 35(20), 7199–7218.
- Zhen, Z., Quackenbush, L. J., Stehman, S. V., Zhang, L., 2015. Agent-based region growing for individual tree crown delineation from airborne laser scanning (ALS) data. *International Journal of Remote Sensing*, 36(7), 1965–1993.
- Zheng, J., Yuan, S., Li, W., Fu, H., Yu, L., Huang, J., 2024. A Review of Individual Tree Crown Detection and Delineation From Optical Remote Sensing Images: Current progress and future. *IEEE Geoscience and Remote Sensing Magazine*.
- Ziter, C. D., Pedersen, E. J., Kucharik, C. J., Turner, M. G., 2019. Scale-dependent interactions between tree canopy cover and impervious surfaces reduce daytime urban heat during summer. *Proceedings of the National Academy of Sciences*, 116(15), 7575–7580.
- Zörner, J., Dymond, J., Shepherd, J., Jolly, B., 2018. PyCrown-Fast raster-based individual tree segmentation for LiDAR data. *Landcare Research Ltd.: Lincoln, New Zealand*, 8.