

Individual-Tree Segmentation and Extraction based on LiDAR Point Cloud Data

Xiaofeng Liu ^{a,*}, Muhamad Taufik Abdullah ^a, Mas Rina Mustaffa ^a, Nurul Amelina Nasharuddin ^a

^a Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Selangor, 43400 Malaysia

Corresponding author: *gs65728@student.upm.edu.my

Abstract—To extract forest parameters and individual tree information accurately and efficiently from plantations, this study focuses on a plantation of *Pinus tabulaeformis* in Chongli District in China. Utilizing LiDAR point cloud data and ground-measured data from 30 plots, we examined the sensitivity of individual tree segmentation to key parameters by varying the grid values of the point cloud distance discriminant clustering algorithm and adjusting the canopy height resolution (CHR) of the watershed algorithm. The objective was to identify the optimal parameters for both algorithms in terms of tree height extraction precision. In the task of individual tree extraction, the point cloud distance discriminant clustering algorithm outperformed the watershed algorithm. This was evidenced by significantly higher recall, precision, and F1-score. However, in terms of tree height precision, as measured by the coefficient of determination and root mean square error (RMSE), the watershed algorithm proved superior. Specifically, the watershed algorithm achieved a coefficient of determination of 0.88 and an RMSE of 0.93 meters, indicating greater precision in estimating tree parameters. Nonetheless, the optimal parameter settings for the watershed algorithm need to be adjusted based on stand density. Thus, through this study, we found that for individual-tree extraction from LiDAR point cloud data, the initial setting of different grid values and resolutions has a significant impact on segmentation precision. It is essential to design tailored approaches for processing point cloud data under varying environmental conditions to achieve optimal results and precision.

Keywords— Individual tree extraction; LiDAR point cloud data; point cloud distance discriminant clustering algorithm.

Manuscript received 15 Mar. 2024; revised 9 Jun. 2024; accepted 12 Aug. 2024. Date of publication 31 Oct. 2024.
IJASEIT is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.



I. INTRODUCTION

Forests are crucial components of terrestrial ecosystems, playing significant roles in water conservation, soil erosion control, and the global carbon cycle [1], [2]. Therefore, accurately and efficiently obtaining structural parameters such as individual tree count, tree height, canopy width, and diameter at breast height (DBH) is essential for the conservation and sustainable use of forest resources [3], [4]. Traditional forest surveys involve field measurements, including species identification and the measurement of DBH and canopy width. These methods are time-consuming, labor-intensive, and can cause damage to the forest [5]. In recent years, Light Detection and Ranging (LiDAR) has emerged as a new tool for forest surveys by collecting high-density laser point cloud data, facilitating the extraction of forest structural parameters and detailed characterization of vertical structures [6].

Currently, there are two main methods for individual tree segmentation based on LiDAR point cloud data. One method involves using the Canopy Height Model (CHM) to identify the tree crown apex by searching for global maxima and then

applying a segmentation algorithm, such as the watershed segmentation algorithm [7]. Dey et al. [8] compared four algorithms—watershed segmentation, point cloud clustering segmentation, neighborhood growth, and feature point decision tree—and found the watershed algorithm to be the most accurate for individual tree segmentation. Hu et al. [9] employed a marker-controlled watershed algorithm to detect and extract individual fruit tree information, achieving individual tree segmentation and canopy extraction accuracies of 95.03% and 86.39%, respectively, demonstrating the feasibility of the watershed algorithm in handling crown overlap and occlusion issues. Beil et al. [10] found that using CHMs of different resolutions, the optimal results for individual tree information extraction were achieved at a resolution of 0.3 meters, with segmentation and height extraction precision declining and under-segmentation issues increasing as the resolution decreased.

The other method directly uses normalized point cloud data for individual tree segmentation. This approach can utilize the raw point cloud data without constructing additional 3D models, thereby reducing the complexity of data processing

[11]. Key methods include the point cloud distance discriminant clustering algorithm [12] and the k-means clustering algorithm [13]. McTegg et al. [4] designed and improved the k-means algorithm, evaluating the suitability of point cloud clustering methods for individual tree segmentation. Peters et al. [14] proposed a Gaussian clustering algorithm, achieving an individual tree segmentation precision of 89% with normalized point clouds. Milojevic et al. [15] conducted a sensitivity analysis of the point cloud distance discriminant clustering algorithm and found that the best segmentation results were obtained when the distance threshold equaled the average crown radius of the plot.

Applying LiDAR point cloud data for individual tree segmentation allows for highly accurate acquisition of individual tree information, with key parameter values being crucial for the segmentation results. This study focuses on *Pinus tabulaeformis* plantations with varying stand densities in the Chongli District, utilizing airborne LiDAR data and ground-truth measurements to compare the effectiveness of the watershed segmentation algorithm and the point cloud distance discriminant clustering algorithm for individual tree segmentation and height extraction. The primary goal is to explore the optimal parameter values for different stand densities, providing technical support for forest surveys and monitoring, as well as carbon cycle and budget assessments in the region.

The research area is located in Chongli District, Zhangjiakou City, Hebei Province, China, in the northwest of Hebei Province, and the transitional zone between the Inner Mongolia Plateau and North China Plain. The geographical coordinates are between 114°17'–115°34' E and 40°47'–41°17' N, as shown in Figure 1.

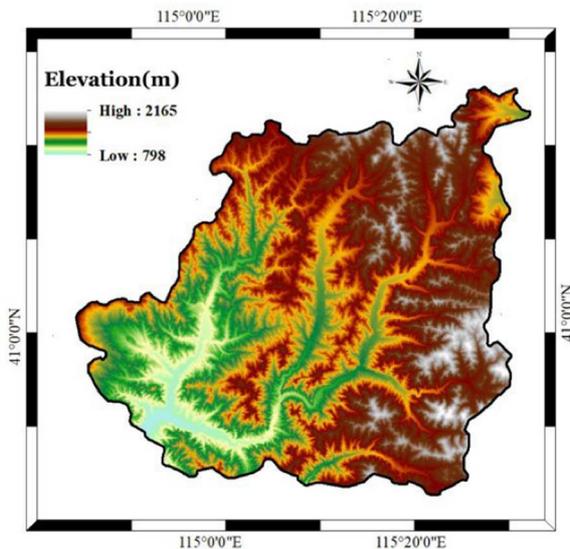


Fig. 1 Chongli District, Zhangjiakou City, Hebei Province, China

The total area of Chongli District is 2334 square kilometers, with 80% being mountainous and a forest coverage rate of 52.38%. The terrain in this area is steep, consisting of Zhongshan, with some low mountains and hills. The elevation ranges from 813 m to 2174 m, with a maximum

height difference of 1361 m. The sample site is located in the Chongli, which covers an area of 39.33 km², and is rich in plant species, with more than 170 common tree species, including *Robinia pseudoacacia*, *Pinus tabulaeformis*, *Platycladus orientalis*, and *Picea abies. principis-rupprechtii*, and so on.

II. MATERIALS AND METHOD

A. Airborne LiDAR Data

In this study, a DJI M300 RTK Unmanned Aerial Vehicle (UAV) equipped with a Zenith L1 radar lens was selected for data acquisition. In a windless environment, the UAV flew at a speed of 10 m/s, at a flight altitude of 80 m, with a laser bypass overlap rate of 80%, and an average point density of 147 points/m². In order to ensure the quality of the LiDAR point cloud data, the weather was selected to be sunny in May–July 2024, and the wind speed was lower than level 3 for the suitable weather conditions for the operation. The collected LiDAR point cloud data were stored in the standard LiDAR format (.las).

B. Ground Investigation Data

In this study, 30 blocks of oil pine forests with different stand densities, good growth conditions, and complete stand structure were randomly selected from March to July 2024, and the trees within each sample plot were examined on a per-timber basis, and the measurement indexes included stand density, slope, elevation, single-timber height, single-timber diameter at breast height (DBH), and single-timber crown width (as shown in Table I). At the same time, hand-held RTK equipment (QianXun SE Lite network RTK receiver, RTK horizontal precision $\pm (8 + 1 \times 10^{-6}D)$ mm, vertical precision $\pm (15 + 1 \times 10^{-6}D)$ mm) was used to accurately record the geographic coordinates of each sample plot.

C. Individual tree extraction method

The LiDAR point cloud data are filtered and denoised, and the improved progressive dense triangular mesh filtering algorithm (IPTD) [5] is used to classify the ground points, which are processed by two individual tree segmentation methods. (1) Kriging interpolation was used to generate digital elevation model (DEM), digital surface model (DSM), and canopy height model (CHM), in order to optimize the performance of the algorithm and to retain the original height structure characteristics of the canopy height, Gaussian filtering was applied to the canopy height using a 3×3 window to reduce the impact of noise on the algorithm [14], and the local maximum algorithm was used to identify the single tree crown tops in the filtered canopy height. The local maximum algorithm was used to identify the crown tops of single trees in the filtered crown height, and the watershed segmentation algorithm was used to segment the single trees based on the identified crown tops of single trees. (2) For the classified point cloud, according to the ground point normalization process, in order to weaken the interference of topographic relief on the elevation value, the point cloud distance discrimination clustering algorithm is used for individual tree segmentation [15].

TABLE I
BASIC INFORMATION ON SAMPLE PLOTS

Sample Plot No.	Stand density/plant·hm ⁻²	Elevation / (°)	altitude /m	Mean crown width /m	Mean tree height /m	Mean diameter at breast height /cm
1	1 866	38	1 350	(3.01±0.88) a	(8.87±1.30) a	(13.65±3.19) a
2	2 000	0	1 075	(3.54±1.02) a	(9.16±1.83) a	(13.98±3.90) a
3	2 025	30	1 344	(3.21±0.97) a	(8.94±1.89) a	(11.40±4.48) b
4	2 150	18	1 403	(3.39±0.84) a	(9.55±1.85) a	(14.51±3.89) a
5	2 350	38	1 336	(3.33±0.87) a	(9.72±1.52) a	(13.27±3.56) a
6	2 600	25	1 354	(3.13±0.98) b	(7.46±1.87) a	(11.31±4.60) b
7	2 900	26	1 358	(3.21±0.82) a	(10.28±1.20) a	(12.70±3.11) a
8	2 900	30	1 337	(2.97±0.81) a	(10.22±1.71) a	(12.04±4.10) b
9	2 950	40	1 270	(3.24±1.04) b	(7.37±1.33) a	(10.94±3.49) b
10	3 075	25	1 284	(3.00±0.88) a	(8.04±1.58) a	(10.76±3.70) b
11	3 100	0	1 333	(2.81±1.01) b	(6.68±1.15) a	(9.40±4.22) b
12	3 450	30	1 301	(2.84±0.78) a	(10.43±1.80) a	(11.52±3.57) b
13	3 475	24	1 395	(2.93±0.86) a	(9.78±1.67) a	(10.79±3.88) b
14	3 500	16	1 215	(2.91±0.83) a	(7.99±1.32) a	(11.10±3.63) b
15	3 500	16	1 304	(2.93±0.85) a	(8.30±1.54) a	(10.51±3.58) b
16	3 550	33	1 227	(3.59±1.09) a	(8.67±2.26) a	(12.56±3.93) b
17	3 600	27	1 275	(2.86±0.94) b	(8.14±1.68) a	(10.18±4.06) b
18	3 700	31	1 199	(2.79±0.95) b	(9.19±1.24) a	(10.92±3.49) b
19	3 850	23	1 356	(2.61±0.60) a	(6.44±1.45) a	(9.35±2.79) b
20	3 950	21	1 340	(3.06±0.79) a	(10.59±1.73) a	(12.20±3.67) b
21	4 000	29	1 290	(2.67±0.99) b	(7.63±1.68) a	(9.46±4.40) b
22	4 025	28	1 410	(3.11±0.89) a	(9.04±2.12) a	(11.46±3.99) b
23	4 043	10	1 344	(2.74±0.91) b	(9.93±2.00) a	(11.64±3.5) b
24	4 325	25	1 308	(2.66±0.85) b	(10.18±1.95) a	(10.82±3.63) b
25	4 375	36	1 252	(2.65±0.79) a	(6.42±1.14) a	(8.01±3.03) b
26	4 625	25	1 334	(2.76±0.81) a	(9.41±1.60) a	(10.22±3.69) b
27	5 000	20	1 277	(2.67±0.92) b	(8.98±2.00) a	(9.37±4.02) b
28	5 350	25	1 307	(2.45±0.79) b	(8.98±1.91) a	(9.07±3.61) b
29	5 400	24	1 273	(2.38±0.75) b	(6.66±1.12) a	(7.09±3.59) b
30	5 700	0	1 277	(2.66±0.92) b	(7.33±1.36) a	(7.76±3.72) b

Note: Different lowercase letters in the table indicate different coefficients of variation. a is coefficient of variation ≤ 0.3 , b is coefficient of variation > 0.3 .

D. Individual Tree Height Extraction

Tree-top detection, as the core link in the single-tree segmentation and parameter extraction process, has a decisive impact on the reliability of the final results in terms of its precision. In this study, an adaptive dynamic windowing strategy is applied to find the local maxima of CHM and mark the potential tree tops [16]. These vertices are judged one by one by traversing the whole globe, in which the value of the image height attribute corresponding to each vertex is regarded as the tree height.

E. Precision Evaluation

First, we set the evaluation of individual tree segmentation precision. According to existing studies, individual tree segmentation is divided into correct segmentation, missed segmentation and over-segmentation, correct segmentation is when point cloud data can correctly detect an individual tree, missed segmentation is when point cloud data is not detected as an individual tree, and over-segmentation is when point cloud data divides a tree into multiple individual trees. The precision of individual tree segmentation results was evaluated using three metrics such as recall (r), precision (p) and reconciliation value (F) [17], [18]. Recall rate is the proportion of the number of correctly segmented trees to the number of measured trees, precision rate is the proportion of the number of correctly segmented trees to the number of extracted trees, and reconciliation value is the overall

segmentation precision of the algorithm. The formula is expressed as:

$$r = T_p / (T_p + F_n) \quad (1)$$

$$p = T_p / (T_p + F_p) \quad (2)$$

$$F = \frac{2(r \times p)}{r + p} \quad (3)$$

where: r is the recall rate; p is the precision rate; F is the reconciliation value; T_p is the number of correctly segmented single logs; F_n is the number of missed segmented single logs; and F_p is the number of over-segmented individual trees.

For evaluation of tree height extraction precision, on this research, the coefficient of determination (R^2) and root mean square error (R_{MSE}) were used as the evaluation indexes [19], and the precision of extracting single-tree parameters based on the point cloud distance discriminative clustering algorithm and watershed algorithm was verified by regression analysis between the estimated and measured values of the two methods, which was expressed by the formula:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

$$R_{MSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-1}} \quad (5)$$

where: n is the number of samples; \hat{y}_i is the estimated tree height based on LiDAR point cloud data; y_i is the measured

value of tree height; and \bar{y} is the arithmetic mean of measured tree height.

III. RESULTS AND DISCUSSION

A. Precision of Single Tree Segmentation

As can be seen from Table 2 and Figure 2, the number of correct segmentation of the point cloud distance discriminant clustering algorithm (PCDDCA) is 3,009, the number of missed segmentation is 437, and the number of over-segmentation is 489, which is 149 more than the number of correctly extracted single trees, 149 fewer than the number of missed segmentation, and 65 fewer than the number of over-

segmentation of single trees of the watershed segmentation algorithm (WSA), and the recall, precision, and concordance values of the point cloud distance discriminant clustering algorithm are 0.87, 0.86, 0.87, respectively; the recall, precision, and reconciliation values of the watershed segmentation algorithm are 0.83, 0.84, and 0.83, respectively; the overall segmentation precision of the point cloud distance discriminant clustering algorithm is better than that of the watershed segmentation algorithm. The overall segmentation effect of the point cloud distance discriminant clustering algorithm is relatively clear and significantly better than the overall segmentation effect of the watershed segmentation algorithm.

TABLE II
VERIFICATION OF INDIVIDUAL TREE SEGMENTATION PRECISION

Segmentation method	Correct Segmentation /plant	Missing Segmentation /plant	Over Segmentation/plant	recall rate r	Precision p	F-score F
PCDDCA	3009	437	489	0.87	0.86	0.87
WSA	2860	586	554	0.83	0.84	0.83

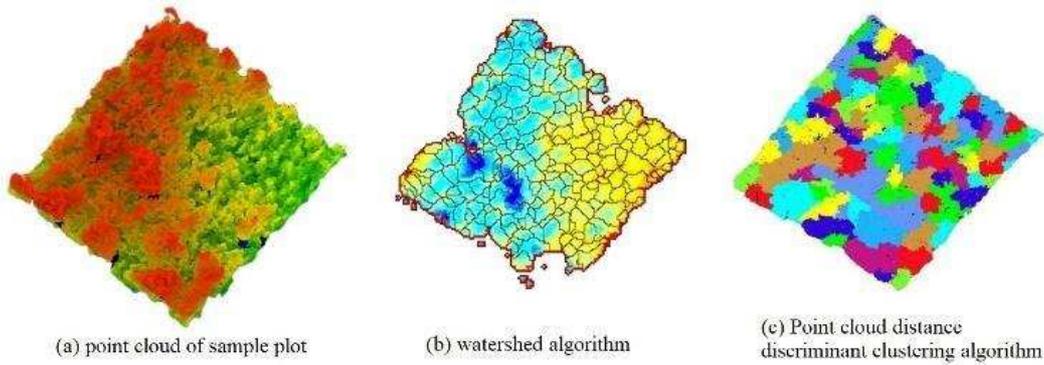


Fig. 2 Segmentation effect diagram of different segmentation methods

B. Sensitivity of Key Parameters

1) Sensitivity of key parameters of the point cloud distance discriminant clustering algorithm:

As can be seen from Figure 3 and Table 3, the key parameter used by LiDAR360 software to identify the location of trees is the grid value, and the recommended value is 1/5 of the crown diameter, and the grid values were set as 1/5 of the measured minimum crown diameter, 1/5 of the average crown diameter and 1/5 of the maximum crown diameter of the sample plot, and sensitivity analysis was performed for the point cloud distance discriminant clustering algorithm. From left to right, the grid value increases sequentially, when the grid value is 1/5 of the minimum crown diameter, compared with the grid value of 1/5 of the average crown diameter, there are 5 more trees correctly segmented, compared with the grid value of 1/5 of the maximum crown diameter, there are 10 more trees correctly segmented. Although the number of over segmentations decreases and the number of missing segmentations increases as the grid value increases, the impact of missing segmentation on segmentation precision is greater than that of over segmentation, so the overall segmentation precision decreases.

In Figure 3, the d, e, and f are local zoom-ins of the same region of Figures a, b, and c, respectively; white dots represent that this tree is correctly segmented, and black dots represent that this tree is over segmented.

For the 30 sample plots with different stand densities, when the grid value was set to 1/5 of the minimum crown diameter of the sample plot, the segmentation precision was the highest, and the recall, precision, and F values were 0.87, 0.86, and 0.87, respectively; when the grid value was set to 1/5 of the maximum crown diameter of the sample plot, the precision was much higher than the recall because the number of trees that were missing segmented was more than 1/2 of the total number of trees in the sample plot, and the precision of the individual tree segmentation was the lowest, and the recall, precision, and F values were 0.28, 0.88, and 0.42. The recall for segmenting individual trees with different grid values showed significant differences ($P < 0.05$), indicating that the number of missing individual tree segmentation varied greatly, while the precision did not show significant differences, and the F value showed significant differences ($P < 0.05$), so the seriousness of the missing segmentation was the main reason for the impact on the precision of individual tree segmentation with the point cloud distance discriminant clustering algorithm.

TABLE III
EVALUATION OF INDIVIDUAL TREE SEGMENTATION PRECISION FOR DIFFERENT GRID VALUES

Sample Plot No.	Stand density/plant·hm ⁻²	Minimum crown diameter 1/5			Average crown diameter 1/5			Maximum crown diameter 1/5		
		recall	precision	F	recall	precision	F	recall	precision	F
1	1 866	0.86	0.88	0.87	0.52	0.79	0.63	0.43	0.9	0.58
2	2 000	0.93	0.82	0.87	0.59	0.85	0.7	0.41	0.89	0.56
3	2 025	0.86	0.80	0.83	0.57	0.87	0.69	0.35	0.88	0.5
4	2 150	0.84	0.88	0.86	0.59	0.86	0.7	0.36	0.89	0.51
5	2 350	0.93	0.79	0.85	0.61	0.8	0.69	0.35	0.89	0.5
6	2 600	0.84	0.93	0.88	0.6	0.78	0.68	0.36	0.9	0.51
7	2 900	0.83	0.89	0.86	0.62	0.9	0.73	0.48	0.93	0.64
8	2 900	0.89	0.80	0.84	0.59	0.84	0.69	0.3	0.83	0.44
9	2 950	0.88	0.87	0.87	0.49	0.75	0.59	0.2	0.67	0.31
10	3 075	0.89	0.84	0.87	0.64	0.84	0.73	0.3	0.86	0.45
11	3 100	0.84	0.81	0.83	0.61	0.86	0.72	0.39	0.86	0.53
12	3 450	0.92	0.85	0.89	0.52	0.81	0.63	0.27	0.8	0.40
13	3 475	0.80	0.91	0.85	0.57	0.80	0.66	0.23	0.71	0.35
14	3 500	0.88	0.84	0.86	0.54	0.84	0.66	0.26	0.84	0.40
15	3 500	0.86	0.83	0.84	0.58	0.86	0.69	0.23	0.84	0.36
16	3 550	0.79	0.90	0.84	0.58	0.85	0.69	0.25	0.80	0.39
17	3 600	0.90	0.80	0.85	0.50	0.84	0.63	0.19	0.88	0.32
18	3 700	0.86	0.84	0.85	0.57	0.91	0.7	0.38	0.82	0.52
19	3 850	0.89	0.86	0.88	0.49	0.84	0.62	0.24	0.86	0.38
20	3 950	0.78	0.90	0.83	0.56	0.84	0.67	0.27	0.82	0.4
21	4 000	0.90	0.78	0.84	0.40	0.76	0.52	0.33	0.81	0.46
22	4 025	0.88	0.93	0.90	0.55	0.85	0.67	0.24	0.83	0.38
23	4 043	0.89	0.88	0.89	0.56	0.82	0.67	0.32	0.83	0.46
24	4 325	0.94	0.86	0.90	0.62	0.87	0.73	0.27	0.79	0.4
25	4 375	0.91	0.82	0.86	0.56	0.82	0.67	0.24	0.89	0.38
26	4 625	0.89	0.88	0.88	0.55	0.82	0.66	0.31	0.94	0.47
27	5 000	0.92	0.91	0.91	0.55	0.83	0.66	0.24	0.72	0.35
28	5 350	0.84	0.93	0.88	0.47	0.74	0.58	0.24	0.86	0.37
29	5 400	0.87	0.84	0.85	0.57	0.91	0.7	0.33	0.90	0.49
30	5 700	0.88	0.77	0.82	0.47	0.84	0.61	0.26	0.88	0.41
average		0.87 a	0.86 a	0.87 a	0.55 b	0.83 a	b	0.28 c	0.84 a	0.42 c

Note: Different lower-case letters in the table indicate significant differences ($P < 0.05$) between the indicators of individual tree segmentation at different grid values.

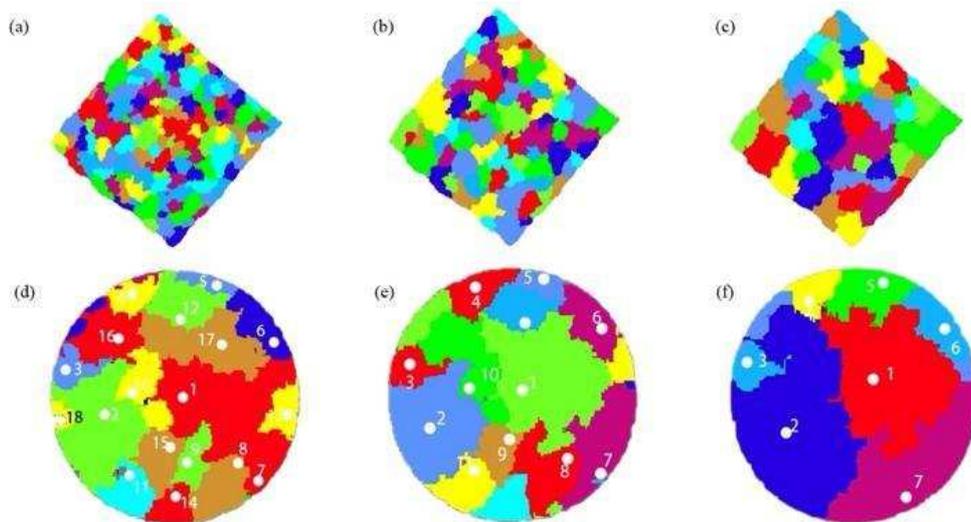


Fig. 3 Performance of segmentation with different grid values

2) Sensitivity of key parameters of the watershed segmentation algorithm:

As shown in Table 4, the optimal canopy resolution of the watershed algorithm for segmenting individual trees varied

for oil pine forests with different densities by selecting 0.2, 0.3, and 0.4 m canopy height resolution. When the density of the stand is 1,866-3,600 plants/hm² (sample plots 1-17), the results of segmenting individual trees with 0.3 m canopy height resolution are optimal, with a recall of 0.79, a precision

of 0.86, and a F value of 0.82. When individual trees are segmented with 0.2 m canopy height resolution, the resolution is too high, and the details of the image are enlarged, and the image is more sensitive to the noise, which results in an over segmentation and a lower precision. The overall segmentation precision is poor, with a recall of 0.83, a precision of 0.46, and a F value of 0.60; when segmenting individual trees with a crown height resolution of 0.4 m, the crown is too smooth, making it difficult to capture more information in the image, resulting in a serious phenomenon of missing segmentation, a reduction in the recall value, and a poor overall segmentation precision, with a recall of 0.49, a precision of 0.82, and a F value of 0.62.

When the density of the forest stand was 3,700-5,700 plants/hm² (sample plots 18-30), the results were optimal when using 0.2 m canopy height resolution to split single trees, with a recall of 0.88, an precision of 0.84, and a F value of 0.86; when using 0.3 and 0.4 m canopy height resolution to segment individual trees, the precision was relatively high, but the phenomenon of missing segmentation was serious, and the over segmentation phenomenon was worse, the recall was 0.58 and 0.39, the precision was 0.89 and 0.83, and the F values were 0.70 and 0.53, respectively.

As can be seen from Fig. 4, a sample plot with good growth condition and complete forest structure was selected from the sample plots with different stand densities (Sample Plot 6 and Sample Plot 20), and the stand densities of Sample Plot 6 and Sample Plot 20 were 2, 600/hm² and 3, 950/hm² respectively. In Sample Plot 6, the number of missing segmented individual trees at a canopy height resolution of 0.4 m is 4 more than at a resolution of 0.3 m. In this case, missing segmentation is the primary factor affecting segmentation precision; At a canopy height resolution of 0.2 m, there are 7 more over segmented individual trees compared to a resolution of 0.3 m. In this case, over segmentation is the primary factor affecting segmentation precision. The segmentation performance is optimal when the canopy height resolution is 0.3 m for this plot. Sample Plot 20, at canopy height resolutions of 0.3 m and 0.4 m, the number of missing segmented individual trees is 9 and 12 more, respectively, compared to a resolution of 0.2 m; In this area, over segmentation occurs at a canopy height resolution of 0.2 m, but its impact on segmentation precision is less significant than that of missing segmentation. The segmentation performance is optimal at a canopy height resolution of 0.2 m for this sample plot.

TABLE IV
PRECISION OF INDIVIDUAL TREE SEGMENTATION AT DIFFERENT CANOPY HEIGHT RESOLUTIONS

Plot number	Stand density/plant·hm ⁻²	Canopy height resolution 0.2 m×0.2 m			Canopy height resolution 0.3 m×0.3 m			Canopy height resolution 0.4 m×0.4 m		
		recall	precision	F	recall	precision	F	recall	precision	F
1	1 866	0.76	0.39	0.52	0.83	0.74	0.79	0.67	0.78	0.72
2	2 000	0.86	0.42	0.56	0.88	0.74	0.8	0.63	0.71	0.67
3	2 025	0.83	0.35	0.49	0.91	0.66	0.77	0.72	0.77	0.74
4	2 150	0.8	0.42	0.55	0.84	0.73	0.78	0.63	0.73	0.68
5	2 350	0.84	0.52	0.64	0.85	0.9	0.87	0.48	0.85	0.61
6	2 600	0.84	0.51	0.64	0.81	0.88	0.84	0.55	0.83	0.66
7	2 900	0.69	0.47	0.56	0.76	0.73	0.75	0.59	0.74	0.65
8	2 900	0.74	0.41	0.52	0.78	0.76	0.77	0.59	0.86	0.7
9	2 950	0.7	0.48	0.57	0.69	0.92	0.79	0.42	0.88	0.56
10	3 075	0.87	0.51	0.64	0.78	0.92	0.85	0.53	0.84	0.65
11	3 100	0.81	0.48	0.6	0.84	0.84	0.84	0.61	0.95	0.75
12	3 450	0.89	0.73	0.8	0.84	0.92	0.88	0.4	0.89	0.55
13	3 475	0.86	0.79	0.83	0.86	0.92	0.89	0.31	0.84	0.45
14	3 500	0.89	0.68	0.77	0.7	0.94	0.8	0.44	0.95	0.6
15	3 500	0.85	0.58	0.69	0.81	0.93	0.86	0.48	0.8	0.6
16	3 550	0.85	0.67	0.75	0.71	0.93	0.8	0.5	0.84	0.63
17	3 600	0.87	0.58	0.7	0.76	0.92	0.83	0.43	0.74	0.54
18	3 700	0.89	0.8	0.85	0.65	0.83	0.73	0.46	0.74	0.57
19	3 850	0.88	0.76	0.81	0.64	0.9	0.75	0.44	0.89	0.59
20	3 950	0.86	0.79	0.82	0.6	0.87	0.71	0.37	0.74	0.49
21	4 000	0.85	0.71	0.77	0.65	0.87	0.74	0.45	0.86	0.59
22	4 025	0.93	0.9	0.91	0.55	0.86	0.67	0.4	0.84	0.54
23	4 043	0.9	0.78	0.84	0.65	0.87	0.74	0.49	0.85	0.63
24	4 325	0.9	0.91	0.91	0.5	0.88	0.64	0.32	0.86	0.47
25	4 375	0.92	0.78	0.84	0.65	0.92	0.76	0.45	0.81	0.58
26	4 625	0.91	0.81	0.86	0.52	0.91	0.66	0.36	0.87	0.51
27	5 000	0.86	0.91	0.88	0.56	0.91	0.69	0.35	0.8	0.49
28	5 350	0.86	0.93	0.89	0.52	0.91	0.66	0.32	0.84	0.47
29	5 400	0.76	0.85	0.8	0.46	0.83	0.6	0.41	0.88	0.56
30	5 700	0.86	0.75	0.8	0.72	0.89	0.8	0.47	0.82	0.6
Total		0.86a	0.65a	0.74a	0.69b	0.87a	0.77a	0.44c	0.82b	0.58b

Different lowercase letters in the table indicate significant differences between the indicators of single wood segmentation at different CHM resolutions, starting from the highest mean value labelled a, and comparing with the other

groups sequentially, if there is no significant difference then it is recorded as a, and if the difference is significant then it is recorded as b (P<0.05).

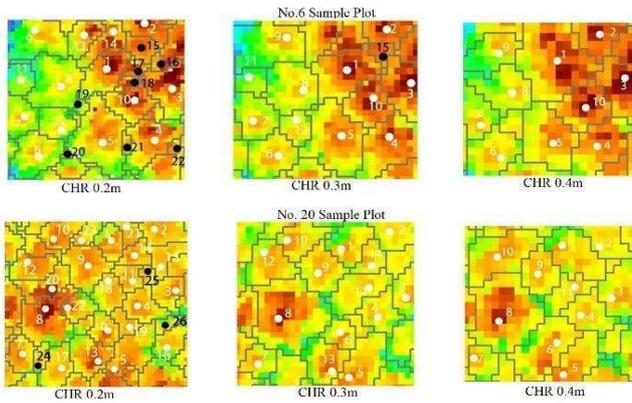


Fig. 4 Sensitivity analysis of the Watershed Algorithm

White dots represent trees that are correctly segmented, and black dots represent trees that are over-segmented. In order to verify the precision of the point cloud distance discriminant clustering algorithm and the watershed segmentation algorithm in extracting tree heights, the measured values were linearly fitted to the estimated values 1 (point cloud distance discriminant clustering algorithm) and 2 (watershed segmentation algorithm), respectively. Since it was difficult to directly correspond the location of the segmented greasewood with the actual data of the trees, the estimation of tree height was judged based on the sample level statistics of the mean tree height of each sample site.

As shown in Figure 5, the coefficient of determination (R^2) in the point cloud distance discriminant clustering algorithm and the watershed segmentation algorithm were 0.82 and 0.88, the root mean square error was 1.46 m and 0.93 m, and the correlation coefficients were 0.94 and 0.91, respectively ($P < 0.05$). The correlation coefficients were 0.94 and 0.91 ($P < 0.05$). This indicates that the height values extracted by splitting a single tree in the watershed algorithm are closer to the measured values.

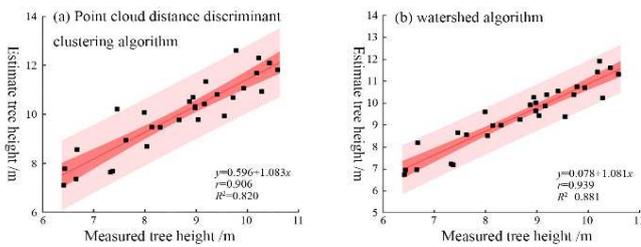


Fig. 4 Extraction precision of tree height by different algorithms

IV. CONCLUSION

Taking airborne LiDAR point cloud data as the data source, the improved progressive encrypted triangular mesh filtering algorithm (IPTD) was used to classify the point cloud, and the point cloud distance discriminant clustering algorithm and watershed algorithm were applied to extract the structural parameters of the single trees for precision evaluation, and the following conclusions were drawn: the overall segmentation precision of the point cloud distance discriminant clustering algorithm and watershed segmentation algorithm for the man-made oleander forests was higher, with a reconciliation value of 0.87 and 0.83, respectively. For oil pine forests with different stand densities, the point cloud distance discriminant clustering algorithm had the highest segmentation precision

when the optimal grid value for segmenting single trees was 1/5 of the minimum crown diameter of the sample site, with a recall of 0.87, a precision of 0.86, and a concordance value of 0.87. The optimal resolution of crown height for the watershed algorithm for segmenting single trees varied, and the canopy height resolution varied when the density of the stand was between 1,866-3,600 trees/hm², and the canopy height resolution varied from 1,000 to 1,600 plants/hm², with the canopy height of 1,000 to 1,600 plants/hm². When the stand density is 1,866-3,600 trees/hm² and the canopy height resolution is 0.3 m, the extraction of single trees is the best, with a recall of 0.79, a precision of 0.86, and a reconciliation value of 0.82. When the stand density is 3,700-5,700 trees/hm² and the canopy height resolution is 0.2m, the extraction of single trees is the best, with a recall of 0.88, a precision of 0.84, and a reconciliation value of 0.86. The tree height resolution of the watershed segmentation algorithm is different, and the optimum tree height resolution for the split of single trees differs. The coefficient of determination (R^2) of the watershed segmentation algorithm was 0.88, and the root mean square error (RMSE) was 0.93 m. The tree height extraction precision was higher than that of the point cloud distance discriminant clustering algorithm (the coefficient of determination (R^2) was 0.82, and the RMSE was 1.46 m. The tree height extraction precision was higher than that of the point cloud distance discriminant clustering algorithm.

The grid value is a key parameter of point cloud distance discriminant clustering algorithm for single-tree segmentation, Gu et al [20]. When applying the ground-based radar point cloud data to model the digital elevation of glacier surface, they found that a lower grid value would result in more voids in the digital elevation, and a higher grid value would result in a smoother surface of digital elevation, which would be difficult to describe the real details of the glacier surface, so choosing the appropriate grid value is crucial for segmentation effect. In this study, when the grid value was set to 1/5 of the minimum crown diameter of the sample plots in 30 sample plots with different stand densities, the segmentation precision was the highest ($r=0.87$, $p=0.86$, $F=0.87$), and as the grid size increases, the number of over segmentations decreases while the number of missing segmentations increases. The impact of missing segmentations on segmentation precision is greater than that of over segmentations, resulting in a gradual decrease in the number of correctly segmented instances [21], [22]. Consequently, the recall rate declines, leading to an overall reduction in segmentation precision [23], [24]. In this study, the grid value is the key parameter because of the different versions of the software used (the essence of the two key parameters of the average value of the crown radius of the sample site and the grid value is to discriminate the distance of the point cloud for the purpose), and the optimal value of the grid for the minimum value of the crown diameter of the sample site 1/5 of the segmentation effect is the best [25], because of the large density of the oil pine stands in the study area (1,866-5,700 plants/hm²), and the density of the sample site in the previous study is the average of 478 plants/hm². The average stand density of the previous study was 478 plants/hm², so the canopy overlap was serious, which affected the point cloud distance discriminant clustering algorithm.

Resolution is an important parameter in the process of generating DEM, DSM, and CHM. Ao et al [26] found that by analysis of the precision of CHM single-tree canopy extraction at different resolutions, when data collection was carried out in forests with differences in scenes and scales, the spatial resolution could be adjusted to improve the recognition precision [10], and the segmentation effect was superior to that of resolutions of 0.1, 0.5, and 1.0 m when the resolution was 0.3 m × 0.3 m. This study found that for different densities of stands, the CHM resolution could be adjusted to make it more adaptable to the density of stands in the study area. In this study, it was found that for stands with different densities, by adjusting the CHM resolution to make it more adaptable to the density of the stands in the study area, when the density of the stand was ≤3,600 plants/hm² (sample plots 1-17) and the canopy height resolution was 0.3 m, the best effect was achieved in the extraction of single logs, and when the density was ≥3,700 plants/hm² (sample plots 18-30), the crown height resolution was 0.2 m. The best effect was achieved in the extraction of single logs. The best results were obtained when the canopy height resolution was 0.2 m. The resolution varied with stand density because the resolution would directly affect how well the interpolated canopy height model (CHM) matched the real situation of the stand. When the density was ≤3,600 plants/hm² and the canopy height resolution was 0.2 m, many pits and bumps did not match the actual situation of the sample plot due to the high resolution, which led to an increase in the number of over-segmentation; when the canopy height resolution was 0.4 m, the resolution was low, the number of missing segmentation increased, and the overall segmentation precision declined; when the density was ≥3,700 plants/hm², the canopy height resolution was 0.3, 0.4 m, and 0.4 m. The resolution varied with the stand density because the interpolated values would directly affect the agreement between the canopy height model (CHM) and the real situation of the stand. 0.3, 0.4 m, the low resolution causes the generated CHM to be too smooth and cannot accurately capture the details, resulting in an increase in omissions. Peng et al. suggested that the choice of spatial resolution is different for different scenes, and the resolution is not as small as the segmentation effect [9], but close to 1/10 of the canopy diameter when the segmentation effect is the best. In this study, the canopy diameter of sample plots 1-17 was closer to 3 m, and the canopy diameter of sample plots 18-30 was closer to 2 m, which verified that the appropriate canopy height resolution for stand density could improve the segmentation precision of single trees.

According to the point cloud distance clustering discriminant algorithm and the watershed algorithm, the height of single trees was extracted from 30 sample plots, and the watershed segmentation algorithm had a higher precision of extracting the height of oil pine trees in the study area. The precision of the point cloud clustering algorithm is related to the density of the point cloud of the single tree, due to the high density of the oil pine forest in the study area, the point cloud cannot completely penetrate the forest, so it is more difficult to obtain the information of the low trees, which ultimately led to the estimated tree height values are higher than the actual values. As the top of the oleander canopy is more obvious, the combination of identifying local maxima through adaptive windowing and suitable CHM resolution improves

the tree height extraction precision. The heights of individual trees extracted by the watershed segmentation algorithm ranged from 6.75 to 11.92 m, which was consistent with the manually measured heights of *Pinus sylvestris* samples in the similar area, as shown by Wang [27] et al. In this paper, the coefficient of determination (R^2) of the average tree height precision of the ground-measured tree height combined with the LiDAR point cloud data was 0.88, and the root mean square error was 0.93 m. The results were better than those of Wang et al [27]. who estimated the height of single trees in natural forests ($R^2=0.77$), and were slightly lower than those of Reckziegel et al. [28] who fused the data from the airborne radar and the backpack Lidar data and estimated the heights of single trees in the four artificial eucalyptus forests ($R^2=0.77$), which were lower than those of Liu et al. [29] who estimated the height of single trees in four artificial eucalyptus forests by combining the data from the airborne radar and the backpack Lidar data. ($R^2=0.895$). The variability in estimation precision mainly originated from the forest structure, geographical features, and point cloud data collection, classification, and density of the study area [30], [31]. Therefore, more data are needed to verify whether the key parameters for single tree partitioning of planted oil pine forests in Chongli District are applicable to other planted forests and other regions according to the appropriate parameters set for different stand densities.

Through this study, we found that for individual-tree extraction from LiDAR point cloud data, the initial setting of different grid values and resolutions has a significant impact on segmentation precision. It is essential to design tailored approaches for processing point cloud data under varying environmental conditions to achieve optimal results and precision.

REFERENCES

- [1] C. Beil, R. Ruhdorfer, T. Coduro, and T. H. Kolbe, "Detailed Streetspace Modelling for Multiple Applications: Discussions on the Proposed CityGML 3.0 Transportation Model," *ISPRS International Journal of Geo-Information*, vol. 9, no. 10, p. 603, Oct. 2020, doi:10.3390/ijgi9100603.
- [2] F. T. Kurdi, "Efficiency of Terrestrial Laser Scanning in Survey Works: Assessment, Modelling, and Monitoring," *International Journal of Environmental Sciences & Natural Resources*, vol. 32, no. 2, May 2023, doi: 10.19080/ijesnr.2023.32.556334.
- [3] G. Schwoch, T. J. Lieb, and H. Lenz, "Flight Testing Drone Contingencies during Runway Inspection in U-space Shared Airspace," *2023 Integrated Communication, Navigation and Surveillance Conference (ICNS)*, pp. 1–7, Apr. 2023, doi:10.1109/icns58246.2023.10124322.
- [4] S. J. McTegg, F. Tarsha Kurdi, S. Simmons, and Z. Gharineiat, "Comparative Approach of Unmanned Aerial Vehicle Restrictions in Controlled Airspaces," *Remote Sensing*, vol. 14, no. 4, p. 822, Feb. 2022, doi: 10.3390/rs14040822.
- [5] X. Zhou and W. Li, "A Geographic Object-Based Approach for Land Classification Using LiDAR Elevation and Intensity," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 5, pp. 669–673, May 2017, doi: 10.1109/lgrs.2017.2669994.
- [6] D. Marinelli, C. Paris, and L. Bruzzone, "An Approach Based on Deep Learning for Tree Species Classification in LiDAR Data Acquired in Mixed Forest," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022, doi: 10.1109/lgrs.2022.3181680.
- [7] Z. Gharineiat, F. Tarsha Kurdi, and G. Campbell, "Review of Automatic Processing of Topography and Surface Feature Identification LiDAR Data Using Machine Learning Techniques," *Remote Sensing*, vol. 14, no. 19, p. 4685, Sep. 2022, doi:10.3390/rs14194685.

- [8] E. K. Dey, M. Awrangjeb, F. Tarsha Kurdi, and B. Stantic, "Machine learning-based segmentation of aerial LiDAR point cloud data on building roof," *European Journal of Remote Sensing*, vol. 56, no. 1, May 2023, doi: 10.1080/22797254.2023.2210745.
- [9] C. Hu, Z. Pan, and T. Zhong, "Leaf and wood separation of poplar seedlings combining locally convex connected patches and K-means++ clustering from terrestrial laser scanning data," *Journal of Applied Remote Sensing*, vol. 14, no. 01, p. 1, Feb. 2020, doi:10.1117/1.jrs.14.018502.
- [10] Y. Wang, Y. Lin, H. Cai, and S. Li, "Hierarchical Fine Extraction Method of Street Tree Information from Mobile LiDAR Point Cloud Data," *Applied Sciences*, vol. 13, no. 1, p. 276, Dec. 2022, doi:10.3390/app13010276.
- [11] T. Kutzner, K. Chaturvedi, and T. H. Kolbe, "CityGML 3.0: New Functions Open Up New Applications," *PGF – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, vol. 88, no. 1, pp. 43–61, Feb. 2020, doi: 10.1007/s41064-020-00095-z.
- [12] J. P. Richa, J.-E. Deschaud, F. Goulette, and N. Dalmaso, "AdaSplats: Adaptive Splatting of Point Clouds for Accurate 3D Modeling and Real-Time High-Fidelity LiDAR Simulation," *Remote Sensing*, vol. 14, no. 24, p. 6262, Dec. 2022, doi: 10.3390/rs14246262.
- [13] Z. Zhang et al., "Vectorized rooftop area data for 90 cities in China," *Scientific Data*, vol. 9, no. 1, Mar. 2022, doi: 10.1038/s41597-022-01168-x.
- [14] R. Peters, B. Dukai, S. Vitalis, J. van Liempt, and J. Stoter, "Automated 3D Reconstruction of LoD2 and LoD1 Models for All 10 Million Buildings of the Netherlands," *Photogrammetric Engineering & Remote Sensing*, vol. 88, no. 3, pp. 165–170, Mar. 2022, doi:10.14358/pers.21-00032r2.
- [15] N. Milojevic-Dupont et al., "EUBUCCO v0.1: European building stock characteristics in a common and open database for 200+ million individual buildings," *Scientific Data*, vol. 10, no. 1, Mar. 2023, doi:10.1038/s41597-023-02040-2.
- [16] E. Lewandowicz, F. Tarsha Kurdi, and Z. Gharineiat, "3D LoD2 and LoD3 Modeling of Buildings with Ornamental Towers and Turrets Based on LiDAR Data," *Remote Sensing*, vol. 14, no. 19, p. 4687, Sep. 2022, doi: 10.3390/rs14194687.
- [17] E. Hyypää et al., "Accurate derivation of stem curve and volume using backpack mobile laser scanning," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 161, pp. 246–262, Mar. 2020, doi:10.1016/j.isprsjprs.2020.01.018.
- [18] L. Comesaña-Cebral, J. Martínez-Sánchez, H. Lorenzo, and P. Arias, "Individual Tree Segmentation Method Based on Mobile Backpack LiDAR Point Clouds," *Sensors*, vol. 21, no. 18, p. 6007, Sep. 2021, doi: 10.3390/s21186007.
- [19] L. Liu et al., "Single Tree Segmentation and Diameter at Breast Height Estimation With Mobile LiDAR," *IEEE Access*, vol. 9, pp. 24314–24325, 2021, doi: 10.1109/access.2021.3056877.
- [20] Z. Gu and F. Pei, "A multi-layer K-means based algorithm for individual tree identification in forest point clouds," *Forest Resources Management*, vol. 1, pp. 124–131, 2022.
- [21] A. Adhikari, C. R. Montes, and A. Peduzzi, "A Comparison of Modeling Methods for Predicting Forest Attributes Using Lidar Metrics," *Remote Sensing*, vol. 15, no. 5, p. 1284, Feb. 2023, doi:10.3390/rs15051284.
- [22] L. Lombard, R. Ismail, and N. K. Poona, "Modelling forest species using LiDAR-derived metrics of forest canopy gaps," *South African Journal of Geomatics*, vol. 9, no. 1, pp. 31–43, Sep. 2022, doi:10.4314/sajg.v9i1.3.
- [23] G. Fan, H. Liang, Y. Zhao, and Y. Li, "Automatic reconstruction of three-dimensional root system architecture based on ground penetrating radar," *Computers and Electronics in Agriculture*, vol. 197, p. 106969, Jun. 2022, doi: 10.1016/j.compag.2022.106969.
- [24] S. Cai and Y. Xing, "Backpack Lidar Filtering Low Intensity Point Clouds to Extract Forest Tree Breast Diameter," *Forest Engineering*, vol. 37, pp. 12–19, 2021.
- [25] X. Xu, F. Iuricich, K. Calders, J. Armston, and L. De Floriani, "Topology-based individual tree segmentation for automated processing of terrestrial laser scanning point clouds," *International Journal of Applied Earth Observation and Geoinformation*, vol. 116, p. 103145, Feb. 2023, doi: 10.1016/j.jag.2022.103145.
- [26] Z. Ao et al., "Automatic segmentation of stem and leaf components and individual maize plants in field terrestrial LiDAR data using convolutional neural networks," *The Crop Journal*, vol. 10, no. 5, pp. 1239–1250, Oct. 2022, doi: 10.1016/j.cj.2021.10.010.
- [27] F. Tarsha Kurdi, Z. Gharineiat, G. Campbell, M. Awrangjeb, and E. K. Dey, "Automatic Filtering of Lidar Building Point Cloud in Case of Trees Associated to Building Roof," *Remote Sensing*, vol. 14, no. 2, p. 430, Jan. 2022, doi: 10.3390/rs14020430.
- [28] R. Bohn Reckziegel, E. Larysch, J. P. Sheppard, H.-P. Kahle, and C. Morhart, "Modelling and Comparing Shading Effects of 3D Tree Structures with Virtual Leaves," *Remote Sensing*, vol. 13, no. 3, p. 532, Feb. 2021, doi: 10.3390/rs13030532.
- [29] X. Liu et al., "A novel entropy-based method to quantify forest canopy structural complexity from multiplatform lidar point clouds," *Remote Sensing of Environment*, vol. 282, p. 113280, Dec. 2022, doi:10.1016/j.rse.2022.113280.
- [30] Q. Shu, T. Rötzer, A. Detter, and F. Ludwig, "Tree Information Modeling: A Data Exchange Platform for Tree Design and Management," *Forests*, vol. 13, no. 11, p. 1955, Nov. 2022, doi:10.3390/f13111955.
- [31] S. A. Ahmadi, A. Ghorbanian, F. Golparvar, A. Mohammadzadeh, and S. Jamali, "Individual tree detection from unmanned aerial vehicle (UAV) derived point cloud data in a mixed broadleaf forest using hierarchical graph approach," *European Journal of Remote Sensing*, vol. 55, no. 1, pp. 520–539, Oct. 2022, doi:10.1080/22797254.2022.2129095.