

Article

LWSNet: A Point-Based Segmentation Network for Leaf-Wood Separation of Individual Trees

Tengping Jiang ^{1,2,3,4,†}, Qinyu Zhang ^{1,2,3,†}, Shan Liu ^{1,2,3}, Chong Liang ^{1,2,3}, Lei Dai ⁴, Zequn Zhang ⁵ , Jian Sun ^{1,2,3,*} and Yongjun Wang ^{1,2,3,*} 

¹ Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China; jiangtp_3d@whu.edu.cn (T.J.); w33006@njnu.edu.cn (S.L.); 181302134@njnu.edu.cn (C.L.)

² Key Laboratory of Virtual Geographic Environment, Ministry of Education, Nanjing Normal University, Nanjing 210093, China

³ State Key Laboratory Cultivation Base of Geographical Environment Evolution, Nanjing 210093, China

⁴ State Key Laboratory of Information Engineering in Surveying Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China; dailei@whu.edu.cn

⁵ College of Computer Science and Engineering, Northwest Normal University, Lanzhou 730070, China; 2021222178@nwnu.edu.cn

* Correspondence: 09332@njnu.edu.cn (J.S.); wangyongjun@njnu.edu.cn (Y.W.)

† These authors contributed equally to this work.

Abstract: The accurate leaf-wood separation of individual trees from point clouds is an important yet challenging task. Many existing methods rely on manual features that are time-consuming and labor-intensive to distinguish between leaf and wood points. However, due to the complex interlocking structure of leaves and wood in the canopy, these methods have not yielded satisfactory results. Therefore, this paper proposes an end-to-end **LWSNet** to separate leaf and wood points within the canopy. First, we consider the linear and scattering distribution characteristics of leaf and wood points and calculate local geometric features with distinguishing properties to enrich the original point cloud information. Then, we fuse the local contextual information for feature enhancement and select more representative features through a rearrangement attention mechanism. Finally, we use a residual connection during the decoding stage to improve the robustness of the model and achieve efficient leaf-wood separation. The proposed LWSNet is tested on eight species of trees with different characteristics and sizes. The average F1 score for leaf-wood separation is as high as 97.29%. The results show that this method outperforms the state-of-the-art leaf-wood separation methods in previous studies, and can accurately and robustly separate leaves and wood in trees of different species, sizes, and structures. This study extends the leaf-wood separation of tree point clouds in an end-to-end manner and demonstrates that the deep-learning segmentation algorithm has a great potential for processing tree and plant point clouds with complex morphological traits.



Citation: Jiang, T.; Zhang, Q.; Liu, S.; Liang, C.; Dai, L.; Zhang, Z.; Sun, J.; Wang, Y. LWSNet: A Point-Based Segmentation Network for Leaf-Wood Separation of Individual Trees. *Forests* **2023**, *14*, 1303. <https://doi.org/10.3390/f14071303>

Academic Editor: Eric Casella

Received: 31 May 2023

Revised: 16 June 2023

Accepted: 21 June 2023

Published: 25 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The phenotyping of plant morphological traits is important in plant breeding and intelligent forest management. Trees, with superior environment-protecting functions, are ecologically important to the living conditions of the residents; their spatial structure and corresponding vegetation parameters are the main objectives of forest resource surveys and ecological environment simulations [1]. Trees are composed of photosynthetic materials (leaves) and a non-photosynthetic active material (wood). Because the physiological functions of leaves and woody parts are different, the separation of leaves and wood in individual trees is a prerequisite and basis for many studies (i.e., phenotypic trait

extraction and leaf area index estimation) [2]. Traditionally, optical image is the widely used data for leaf-wood separation task. However, most optical images are devoid of important spatial information, restricting the leaf-wood separation task in several trees with a complex structure [3]. By contrast, many emerging 3D sensors (i.e., laser scanning system, structured light, and light detection and ranging (LiDAR)) can quickly collect high-precision point clouds, characterizing the complete 3D spatial information. In the recent years, laser scanning point clouds have been widely used for high-precision phenotyping in the forest, agriculture, and grass, which shows great potential in the leaf-wood separation of individual tree. However, the point clouds acquired by LiDAR are often large and disordered, and contain significant noise interference, which makes leaf-wood separation based on point clouds more challenging. In the dense vegetation distribution areas, the data occlusion and overlapping problems can easily cause over-segmentation or under-segmentation, restricting lots of traditional feature-engineering methods on the leaf-wood separation from tree point clouds. Therefore, it is important to explore accurate, efficient, and robust leaf-wood separation methods for practical and production applications.

Many studies have been proposed to separate leaves and wood in the terrestrial laser scanning (TLS) of point clouds [4–11]. The existing methods for leaf-wood separation are based on [2] (1) the geometric features of point clouds; (2) backscattered intensity as the primary or complementary data source; and (3) machine-learning methods. However, they are still relatively immature and face challenges in performing leaf-wood separation in TLS point clouds. As shown in Figure 1, TLS-acquired point clouds have an uneven distribution, holes caused by inter-canopy occlusion, different tree sizes and canopy shapes, and complex and diverse internal canopy structures. Most methods rely on laborious manual work or user input of specific data features. These potential drawbacks make the application to different data quality and different tree species datasets less generalizable [8]. In particular, when the internal structure of tree canopies is complex, it is not possible to accurately classify leaves and wood that are close to each other.

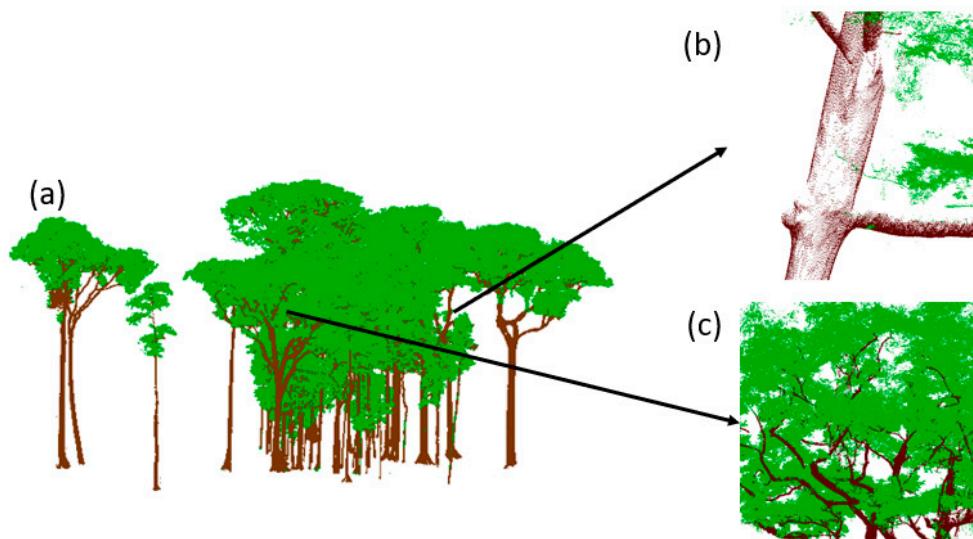


Figure 1. Characteristics of the tree point cloud data: (a) trees of different shapes and heights; (b) incomplete structures; and (c) complex canopy structures.

Recently, benefitting from the advances in convolutional neural network architectures, deep-learning-based methods have shown better performance in object recognition and semantic segmentation. They can extract abstract, high-dimensional features with a high generalization ability from a large number of datasets [12,13]. Qi et al. [12] were the first to propose a network model that acts directly on point clouds to propagate features through an isotropic multilayer perceptron, laying the foundation for a single wood point cloud leaf-wood separation. On this basis, many scholars have investigated leaf-wood separation using deep-learning methods [14–16], which typically use disorderly operations for local

feature aggregation. Despite these early efforts, the contextual information of point clouds is not fully utilized in the leaf-wood separation from point clouds. Specifically, the existing deep-learning networks still lack generalization ability on the leaf-wood separation of different tree species with diverse leaf shapes and canopy structures [1].

In summary, this paper aims to propose a point-based leaf-wood separation model that uses the geometric features of wood and leaf points to enrich the original point cloud information, fuses contextual information by a rearrangement attention mechanism to enhance the features, and adds residual connections to robustly and efficiently distinguish between leaves and wood of different tree species in different environments. The main contributions of this paper are as follows:

- Enriching the original point cloud information by analyzing the structure of leaf and wood points in the canopy and combining local geometric features and point cloud location information;
- Implementing a point-based local context information fusion module, extracting representative local features through a rearrangement attention mechanism, and improving the robustness of the model using a residual connection during the decoding stage to achieve efficient leaf-wood separation;
- Demonstrating that the average F1 score of leaf-wood separation on eight tree species with different features and sizes is as high as 97.29%—it has better performance compared with leaf-wood separation methods in previous studies.

2. Related Work

To separate the leaves and wood in tree point clouds, geometric-feature-based [11] and intensity-feature-based [4] approaches are commonly used. A common geometric-feature-based strategy is a machine-learning classifier that relies on feature descriptors. Knowledgeable and discriminative features are extracted from the spatial arrangement of neighboring points using chunk-based methods [17,18], voxel-based methods [19,20], KD-tree-based methods [21], K-nearest-neighbor search algorithms [22], or spherical neighborhood-search-based methods [23]. Studies have shown that machine-learning methods using geometric-feature-based approaches [7,24,25] can effectively distinguish between leaves and wood. Although methods such as random forest (RF) algorithms [26], Gaussian mixture models (GMM) [27], and support vector machine (SVM) algorithms [28] are independent of forest type and data quality, they require the computation of a large number of feature descriptors. Unsupervised learning methods based on clustering of specific geometric attributes, such as the DBSCAN algorithm [29] and the LeWos model [30], obtain the geometric structures, such as planarity and linearity, of leaves and wood and set debugging thresholds to separate leaf and wood points based on experience.

In general, wood points are regular with high spatial continuity while leaf points are irregular with low spatial continuity. Wang [31] first decomposed the point cloud into semantically homogeneous superimposed points using the perpendicularity and density features, and then classified the overlapping wood and leaf points using the orientation characteristics and linear features. Wan et al. [32] propose a segmentation classification strategy by using a point cloud local curvature and connection-component-labeling algorithm. Tan et al. [2] remove leaf points from wood points by successively using geometric quantities derived from curvature features, calibration density, and salient features. Zhang et al. [9] use the principal component analysis method and minimization information entropy criterion to achieve the optimal neighborhood scale selection, and then adopt the DBSCAN algorithm to achieve the accurate extraction of tree trunk points. The main drawback of these geometric-data-based methods is their computational intensity and time-consuming nature, particularly when dealing with a large number of TLS point clouds.

Intensity-based methods [4] operate on the assumption that tree components possess distinctive optical properties at the laser scanner's operational wavelength. Tan et al. [33] use a polynomial model to correct for intensity values and then apply a k-means clustering method to obtain two classification categories. However, there are limitations to the

intensity-based approach, which presupposes that leaves and wood have distinct optical properties at the normal operating wavelength of the laser scanner. This wavelength is frequently influenced by factors such as distance, local laser exposure, and angle of incidence and must be calibrated specifically for each instrument. If calibrated, the intensity data may be merged with geometric features, potentially offering advantages over utilizing a single method [26]. Furthermore, in the future, the study of multi-band LiDAR systems [34,35] could be employed to distinguish between leaf and wood points.

In recent years, there are also some scholars based on the graph method for the separation of wood and leaf points. This approach is based on the principle that tree points can be arranged into a connected topological network, and wood–leaf separation is performed by applying the shortest path analysis [10]. Xu et al. [36] design new topological geometric features in the learning process and a new least-cost path model to further separate wood and leaf points. Tian et al. [37] propose an automatic and robust GBS method using only point cloud xyz information based on the point cloud segmentation, cluster recognition, and region growth. However, such methods have poor performance when encountering large number of points and small branches.

Currently, more and more researchers are turning to deep-learning-based point-cloud-processing methods. Hugues et al. [38] define a kernel point-based display convolution kernel to improve the performance of semantic segmentation of point clouds. Wang et al. [39] combine semantic and instance segmentation to extract individual roadside trees from vehicle-mounted mobile laser scanning point clouds. Ao et al. [40] perform automatic segmentation of maize monocots based on convolutional neural networks, while Li et al. [41] develop a plant point cloud segmentation technique, DeepSeg3DMaize, incorporating high-throughput data collection and deep learning. However, these methods have not been tried on wood–leaf separation. Dai et al. [16] propose a directionally constrained and a-priori-assisted neural network for separating wood and leaves from terrestrial laser scanning, but the method does not take the rich contextual information into account. Therefore, this paper considers feature enhancement by fusing contextual information through a rearrangement attention mechanism to improve the performance of wood–leaf separation.

3. Materials and Methods

3.1. Experimental Datasets

Separating leaf and wood from tree point clouds is examined in this study. To validate the effectiveness of the leaf-wood separation algorithm, we selected a dataset of 61 large trees from eastern Cameroon, as presented in the work of Wang et al. [30]. These data were used to calculate the tree biomass and calibrate the anisotropic growth model. The dataset covered a total of 15 different tree species, with trees ranging from 8.7 m to 53.6 m in height (mean value of 33.7 ± 12.4 m) and from 10.8 cm to 186.6 cm in diameter at breast height (mean value of 58.4 ± 41.3 cm). The leaf and wood points of each tree were marked manually, which took 1–15 h and required a significant amount of manual handling time. For this paper, we selected eight of these trees, with the species being Soyauxii, Frake, Scleroxylon, Sapelli, Baphia nitida, Pygeum oblongum, Macrocarpa, and Barteri. One to three trees of the same species were selected to train the network model, and the details of the eight trees are shown in Table 1.

Table 1. Information about the eight species of trees.

	Specie	Height (m)	Diameter at Breast Height (cm)
Tree 1	SO	49.2	50.5
Tree 2	FR	48.3	80.5
Tree 3	SC	49.7	108.5
Tree 4	SA	47.3	153.4
Tree 5	BN	26	51
Tree 6	PO	21.8	17.8
Tree 7	MA	40.6	98.5
Tree 8	BA	29.4	26.7

3.2. Overview of the Proposed Method

The proposed method directly employs point cloud co-ordinates (i.e., xyz) and distinct geometric features (e.g., linear) as inputs to predict the semantic label of each point in an end-to-end manner. Figure 2 illustrates the overall architecture of the proposed LWSNet, which employs a downsampling and upsampling U-Net as the backbone network. During the downsampling stage, the features are gradually mapped to a high-dimensional feature space, and the receptive field is gradually enlarged to acquire higher-level, more abstract features. During the upsampling stage, the original spatial resolution is restored layer by layer, and richer high-dimensional features are extracted through residual connections, enhancing the model's robustness. The proposed LWSNet comprises three key components: (1) local geometric feature extraction; (2) local contextual feature enhancement via a rearrangement attention mechanism; and (3) residual connection optimization during the upsampling stage.

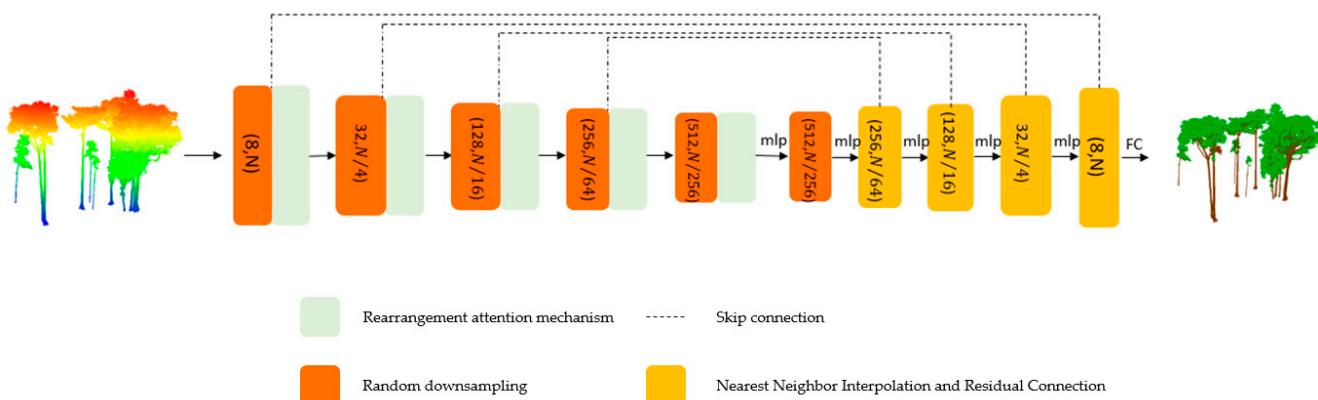


Figure 2. Framework of the proposed LWSNet.

3.3. Local Geometric Feature Extraction

Branches and leaves exhibit distinctly different geometric features in their natural state, with trunk points generally having linear geometric features and leaf points having a discrete distribution. To enhance the distinguishability of leaf and wood points, the proposed LWSNet calculates the local geometric feature vectors of the point cloud by computing the covariance tensor formed by each point and its neighboring points, as shown in Figure 3. In this study, five geometric feature vectors (linearity, planarity, scattering properties, surface variability, and feature entropy) are selected to enrich the original point cloud information. For a given 3D co-ordinate point \hat{p} , the points within a certain radius distance are selected to form its neighborhood point set $S = \{P_1, P_2, \dots, P_k\}$. This neighborhood point set is then used to calculate the covariance tensor C , which is computed using Equation (1):

$$C = \frac{1}{k} \sum_{i=1}^k (p_i - \hat{p})(p_i - \hat{p})^T, \quad (1)$$

where k is the number of 3D points in the neighborhood, p_i is the co-ordinates of the points in the neighborhood, and \hat{p} is the mean value of the co-ordinates of all points in the neighborhood.

After obtaining the symmetric positive definite matrix C , it is decomposed to obtain three eigenvalues $\lambda_1 > \lambda_2 > \lambda_3$ and their corresponding eigenvectors e_1, e_2, e_3 . The three eigenvalues are then normalized so that $\lambda_1 + \lambda_2 + \lambda_3 = 1$. As shown in Table 2, the five geometric features mentioned earlier are then calculated based on these normalized eigenvalues. By using these five geometric features and the 3D co-ordinates of each point as initial features, a training model is built using the dataset that contains marker information.

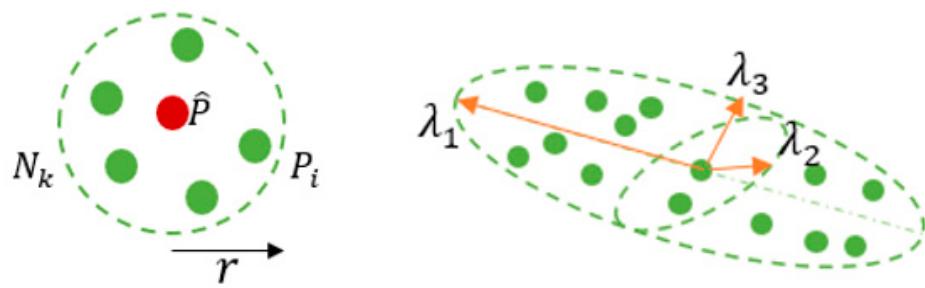


Figure 3. Schematic diagram of the p_i neighborhood and its eigenvalues.

Table 2. Geometric features based on eigenvalues.

Eigenvector	Formula
Linearity	$V_1 = (\lambda_1 - \lambda_2)/\lambda_1$
Planarity	$V_2 = (\lambda_2 - \lambda_3)/\lambda_1$
Scattering properties	$V_3 = \lambda_3/\lambda_1$
Surface variability	$V_4 = \lambda_3$
Characteristic entropy	$V_5 = - \sum_{i=1}^3 \lambda_i \times \ln(\lambda_i)$

3.4. Local Contextual Feature Enhancement via a Rearrangement Attention Mechanism

Given the input point set $X_{in} \in R^{N \times 3}$ and the corresponding local geometric features $F_{in} \in R^{N \times 5}$, the local high-dimensional feature representation space of the modeled point set is $X \in R^{L \times 3}$. The local context features of the point cloud are represented as $f \in R^{L \times D}$; this part considers local feature augmentation for a point location, while the local feature aggregation method is applicable to the whole point set. p denotes the point location and f denotes the corresponding feature. The process of contextual feature enhancement is shown in Figure 4.

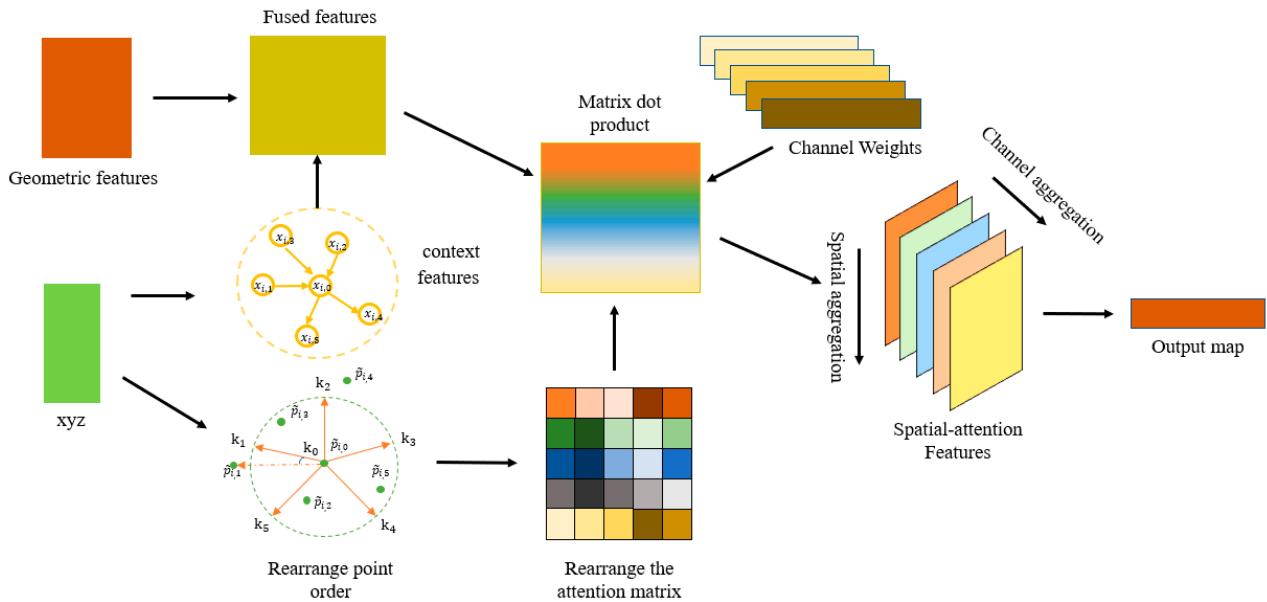


Figure 4. A module for local contextual feature enhancement via a rearrangement attention. This module mainly includes local contextual feature fusion, and learning of rearrangement attention matrix.

3.4.1. Local Contextual Feature Fusion

Before each feature fusion, the neighbors of each point in each downsampling point set are first grouped as contextual feature aggregation units $N(p)$ using the K-nearest-neighbor (KNN) algorithm. K points continued to be sampled on top of the previous layer in an

attempt to slightly expand the sensory field. It is important to note that the searched neighborhood points are queried in the point set X_{in} before downsampling. The local contextual features of the fused point cloud include the original spatial location, the relative position between points in the neighborhood, and the geometric relative features. For each point p , the original spatial information constituted by its neighborhood point set is expressed as Equation (2):

$$\mathcal{L}_{ij} = p_i \oplus p_{i,j}, \quad (2)$$

where \oplus is the concatenation operation, $p_{i,j}$ are the xyz co-ordinates of the points in the neighborhood, and p_i is the xyz co-ordinates of the current point. The relative spatial position is encoded as Equation (3):

$$\mathcal{L}_{rj} = (p_i - p_{i,j}) \oplus \| p_i - p_{i,j} \|, \quad (3)$$

where $\| \cdot \|$ calculates the Euclidean distance from the neighboring points to the center point. The geometric relative characteristics are expressed as Equation (4):

$$f_{i,j} = f_i^{in} - f_j^{in}, \quad (4)$$

where $f_{i,j}$ is the relative geometric feature between the local neighborhood points computed above or the high-dimensional relative feature obtained from the intermediate layer. To encode the contextual features of point p for the final concatenation as Equation (5):

$$x_{ij} = MLP(\mathcal{L}_{ij} \oplus \mathcal{L}_{rj} \oplus f_{i,j}), \quad (5)$$

where $x_{ij} \in \mathbb{R}^{D_{in}}$ and D_{in} is the dimensionality of the feature. For each point, the construction $X_i = \{x_{i,0}, x_{i,1}, \dots, x_{i,K-1}\} \in \mathbb{R}^{D_{in} \times K}$ is used for the convolution operation.

3.4.2. Rearrangement Attention Mechanism

Most deep neural networks for point cloud segmentation typically adopt weight-sharing multilayer perceptron to encode point features directly. However, when it comes to tree canopy points, which have a scattered distribution, it becomes difficult to differentiate between leaf and wood points through weight sharing alone. Therefore, anisotropic convolution is better suited for learning the differentiated high-dimensional features between leaf and wood points. This approach is inspired by traditional convolution on Euclidean structured data. The local neighborhood points are rearranged and assigned corresponding weights based on the relationship between the magnitude of the angle between each point's local neighborhood and a fixed set of kernel points. Representative features of each point are then extracted by applying the generated attention matrix, similar to traditional convolutional neural networks.

For each point $\mathcal{P} = \{p_1, \dots, p_N\}$ in a point cloud, a feature $f \in \mathbb{R}^D$ is associated with it, where D and N represent the dimensionality of the feature and the number of points, respectively. The Euclidean space co-ordinates of each point are represented by a three-dimensional vector $p_i = [x_i, y_i, z_i]^\top$. The feature of each point can contain information such as color, normal vector, or high-dimensional embedded semantic features learned through deep neural networks.

The local neighborhood information N_i of the point cloud is obtained through nearest-neighbor search, where $N_i = \{p_i, p_{i,1}, \dots, p_{i,K-1}\}$ represents the nearest K -neighbor points of the i -th point. A fixed set of kernel points is first generated by mapping the Fibonacci lattice [20] onto the sphere using equal area projection, resulting in uniformly distributed kernel points on the surface of the sphere, denoted as $K = \{k_0, k_1, \dots, k_{L-1}\}$. Here, $k_0 = [0, 0, 0]^\top$ represents the position at the origin and L denotes the number of kernel points. The relative position information between the i -th point and its neighboring points is obtained as $\tilde{p}_{i,j} = p_i - p_{i,j}$, where $p_{i,j} \in N_i$, $\tilde{p}_{i,0} = [0, 0, 0]^\top$ represents the position at the origin. The rearrangement attention matrix is generated by encoding the dot product

between the local neighborhood co-ordinates and the kernel points using the Softmax function, with the expression of Equation (6):

$$\mathcal{H}_i = f(\tilde{P}_i K^\top), \quad (6)$$

where $\tilde{P}_i \in \mathbb{R}^{K \times 3}$, $K \in \mathbb{R}^{L \times 3}$, $\mathcal{H}_i \in \mathbb{R}^{K \times L}$, and $f(\cdot)$ is the Softmax function. For each kernel point, the dot product ensures that neighboring locations with smaller angles to the kernel point are assigned larger weights. The Softmax function allows the generated attention weight matrix to be expressed as probabilities. When the local neighborhood co-ordinate points have smaller angles to the kernel point, they are assigned higher probability values and vice versa. Figure 4 illustrates this phenomenon, where $\tilde{p}_{i,1}$ and $\tilde{p}_{i,3}$ have relatively higher probability values when they are closer to the kernel point k_1 , and the darker the color of the attention matrix, the higher the probability value. Since $k_0 = \tilde{p}_{i,0} = [0, 0, 0]^\top$, the dot product of these two vectors is zero. Therefore, it is necessary to set $\mathcal{H}[0, 0] = 1$ so that the point itself is selected as the first point in the resampled convolution neighborhood.

As the point cloud is scattered and disordered, directly applying isotropic filters to the disordered neighborhood points would reduce expressiveness of the proposed LWSNet. To this end, corresponding weights are assigned based on the magnitude of the angle between the local neighborhood points and the kernel points, generating an attention matrix. The attention matrix is then rearranged to resample the neighborhood of each point and select more representative features, as described in Equation (7):

$$\tilde{X}_i = X_i \mathcal{H}_i, \quad (7)$$

where $\tilde{X}_i \in \mathbb{R}^{D_{in} \times L}$. The weights between the central point and its neighboring points are assigned corresponding weights based on the rearrangement of the magnitude of the angle with the kernel points. Based on this, a weight-sharing MLP (multilayer perceptron) filter can be applied to each point in the point cloud, and the operation is the same as the conventional convolution, expressed as Equation (8):

$$y_i = g(\text{vec}(\tilde{X}_i)^\top W + b), \quad (8)$$

where $W \in \mathbb{R}^{(D_{in} \cdot L) \times D_{out}}$, containing D_{out} anisotropic filters. The bias term is denoted by $b \in \mathbb{R}^{D_{out}}$. The output feature point $y_i \in \mathbb{R}^{D_{out}}$ corresponds to the input feature point $x_i \in \mathbb{R}^{D_{in}}$. The $\text{vec}(\cdot)$ function converts a matrix into a column vector. The function $g(\cdot)$ is an activation function that introduces nonlinearity (e.g., ReLU).

3.5. Residual Connection Optimization during the Upsampling Stage

Although the anisotropic convolution-based contextual feature enhancement module extracts rich, high-level semantic features from the point cloud during the downsampling stage, it ignores the geometric detail information of the input point cloud, which hinders the distinction between leaf and wood points that are close to each other. To address this issue, the original spatial resolution of the point cloud is recovered through nearest-neighbor interpolation on the high-dimensional semantic feature information. While inverse distance weighting and trilinear interpolation upsampling methods are commonly used in deep neural networks, they are computationally inefficient and memory-intensive. To improve the model's efficiency, this paper employs the simple nearest-neighbor interpolation sampling method. However, this method tends to select duplicate points, leading to redundancy of information and reduced model accuracy. To mitigate this issue, the rich high-dimensional semantic features generated through forward propagation are connected to the original abstract features using residual connections during the upsampling stage, compensating for the information redundancy problem caused by nearest-neighbor sampling and improving the robustness of the model.

The flow of residual connections is shown in Equation (9). Based on the semantic feature information $F_{out} \in R^{1 \times 2d}$ output by the local contextual feature fusion module and the rearrangement attention module, the input information is upsampled using nearest-neighbor interpolation. The upsampling point cloud is then non-linearly operated upon using a simple multilayer perceptron (MLP) with ReLU and sigmoid activation functions. Finally, the result is connected with the input features for residuals to output the final feature $F_{enhance} \in R^{k \times 2d}$:

$$F_{enhance} = \text{Sigmoid}\{\text{MLP}[\text{Relu}(\text{MLP}(F_{out}))] + F_{out}\}, \quad (9)$$

3.6. Implementation and Evaluation Metrics

The programming language used for the proposed LWSNet is Python 3.6, and the deep neural network is built using the Tensorflow framework. The operating system used was Ubuntu 16.04, with a RAM of 32G and a GPU Quadro GP100 graphics card model. The network training settings were as follows: batch size of 4, epoch set to 50, initial learning rate set to 0.001, optimization of network parameters performed using Adam's algorithm, and momentum set to 0.9. The evaluation metrics used in this paper are *Precision* (Equation (10)), *Recall* (Equation (11)), and *F1 score* (Equation (12)), which are defined as follows:

$$P = \frac{TP}{TP + FP} \times 100\%, \quad (10)$$

$$R = \frac{TP}{TP + FN} \times 100\%, \quad (11)$$

$$F1 = \frac{2PR}{P + R} \times 100\%, \quad (12)$$

where TP (true positive case) refers to the number of actual positive samples that were correctly predicted as positive, while FP (false positive case) refers to the number of actual negative samples that were incorrectly predicted as positive. FN (false negative case) refers to the number of actual positive samples that were incorrectly predicted as negative, while TN (true negative case) refers to the number of actual negative samples that were correctly predicted as negative. The precision rate indicates the percentage of correctly predicted positive samples to all predicted positive samples. The recall rate, on the other hand, indicates the percentage of correctly predicted positive samples to the actual positive samples. The F1 score is the harmonic mean of the precision rate and recall rate. A higher F1 score indicates a more robust classification model.

4. Results

4.1. Results and Analysis of Leaf-Wood Separation

To demonstrate the effectiveness of the proposed LWSNet, we visually inspect the leaf-wood separation results using several representative tree point clouds. The displayed eight tree point cloud samples in Figure 5 are chosen with diversified spatial structures and under different growth environments to show the exceptional leaf-wood separation ability and accuracy of our LWSNet. In the Figure 5, the first row represents the ground-truths, the middle row shows the leaf-wood separation results of the LeWos [30], and the bottom row is the leaf-wood separation results of the proposed method. Visually, the leaf-wood separation results of the proposed method are in good agreement with the ground-truths, showing that the proposed method is better at detecting the main trunk of trees.

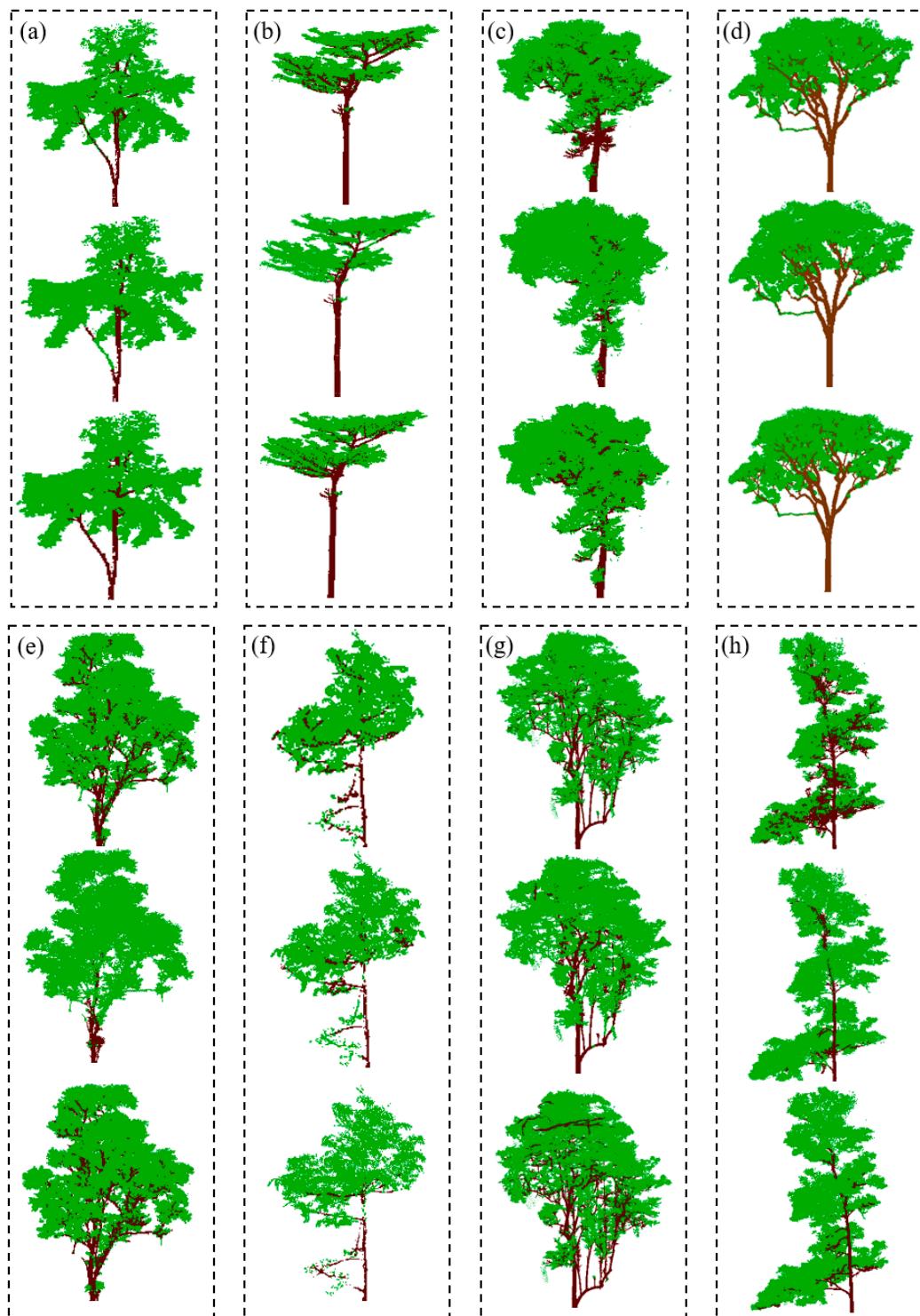


Figure 5. Leaf-wood separation results of the eight trees (green: leaves and brown: wood): (a) Tree 1; (b) Tree 2; (c) Tree 3; (d) Tree 4; (e) Tree 5; (f) Tree 6; (g) Tree 7; and (h) Tree 8. (Top) Manual separation; (middle) LeWos algorithm; and (bottom) proposed method.

By comparing various results, we can find that the proposed LWSNet has a high sensitivity to detecting small branches and leaves. The main reason is some branch and trunk points are often misclassified as leaf points in the complex and diverse internal structure of the canopy. Specially, the proposed method compares favorably with the LeWos method, which mainly distinguishes leaf points and wood points based on linear structural features. In order to provide a more obvious comparison, the wood point

separation results are further visualized in this paper. As shown in Figure 6, both the LeWos and the proposed method show good results for tree trunk differentiation, mainly because both the LeWos and the proposed method introduce linear geometric features, which are better suited for data with more obvious linear features such as rootstocks.

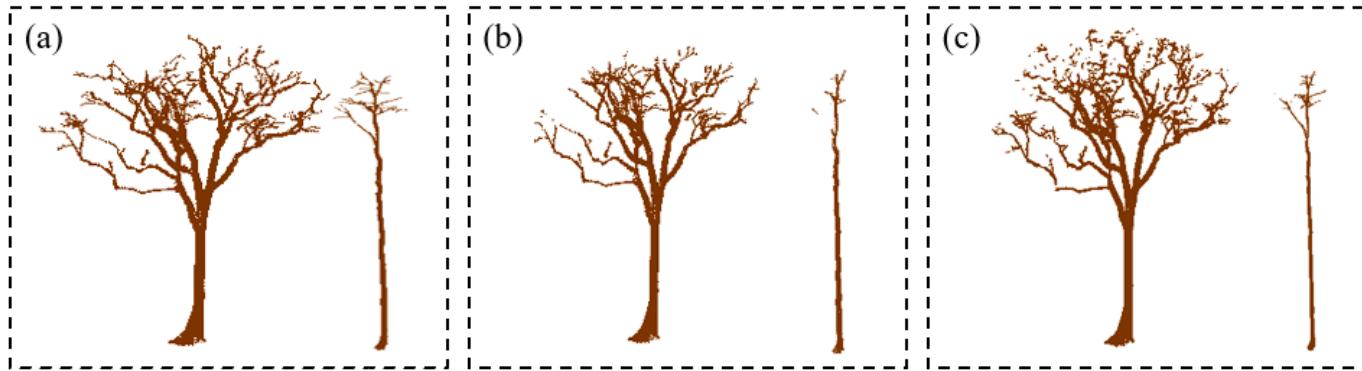


Figure 6. Wood point separation results: (a) ground-truth, (b) the result of LeWos, and (c) the result of the proposed method.

When distinguishing the locations of tree canopies, it can be found that the accuracy of the proposed method in distinguishing leaf and wood points within a canopy is better than that of the LeWos method, whether it is distinguishing leaf and wood points in complex or simple canopies. LeWos assumes that wood points are linearly distributed at different scales, while leaves are usually flat or dispersed, but in the face of a complex structured canopy, the branch points within it are more likely to be misclassified into leaf points, leading to a decrease in overall classification accuracy. In contrast, the proposed method introduces several geometric features in the input stage to further enrich the feature information and learn more robust contextual features through the network, which is more helpful in distinguishing leaf and wood points in more complex canopies. Figure 7 shows the heat map of the classification accuracy of the proposed method for eight different trees' leaves and wood, and the numbers on the diagonal line are the segmentation accuracy for leaf and wood points. It can be found that the proposed LWSNet has similar accuracy for the leaf-wood separation of eight trees, which shows LWSNet has better robustness. From Figure 7h, we can find that the classification accuracy of leaf points of the eighth tree is lower, and more leaf points are misclassified as wood points. On the one hand, it is because Tree 8 has a complex canopy structure and multiple canopies are stacked together, which is challenging. On the other hand, Tree 8 has only one tree in the training set, and the training sample is not rich enough to make the model fit better. From the classification results, we can find that there is always a misclassification of leaf and wood points inside the canopy, which is the challenge that most algorithms face in distinguishing wood and leaf points inside the canopy.

The above visualization compares the classification results between the deep-learning methods proposed in this paper and traditional clustering methods based on prior knowledge. To provide further evidence, representative classification methods in this field were validated based on this dataset in this paper. Table 3 presents the accuracy classification results of feature-engineering-based machine-learning methods [42], the unsupervised learning LeWos algorithm [30], the supervised deep-learning KPconv network [38], and the methods proposed in this paper.

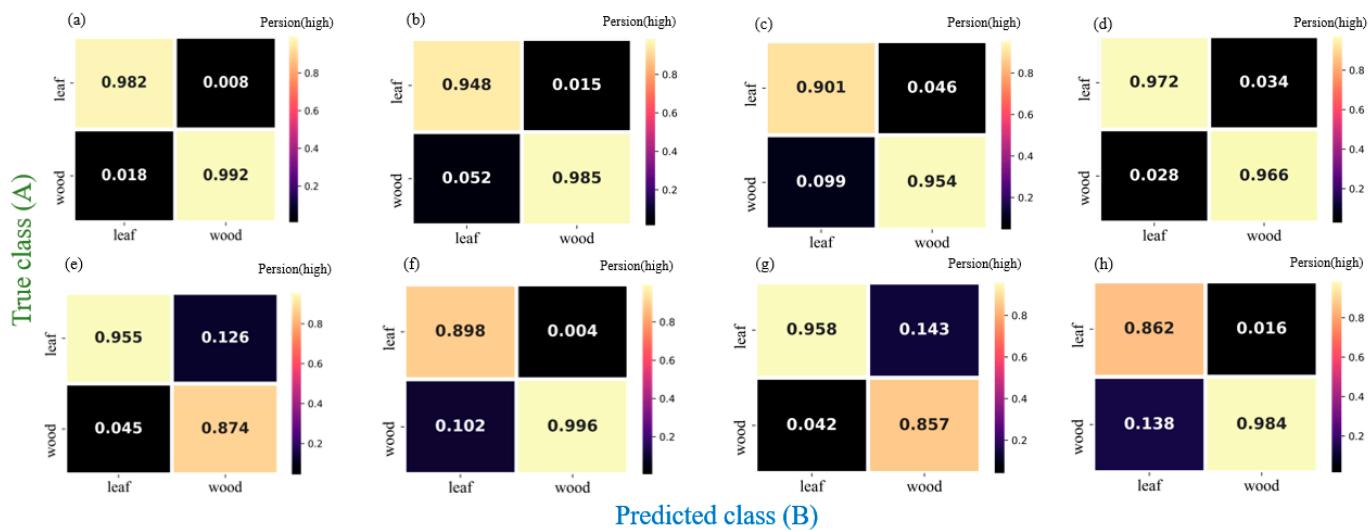


Figure 7. Leaf-wood classification confusion matrix of eight trees: (a) Tree 1; (b) Tree 2; (c) Tree 3; (d) Tree 4; (e) Tree 5; (f) Tree 6; (g) Tree 7; and (h) Tree 8.

Table 3. Quantitative evaluation of leaf-wood separation of eight tree species.

Method		Precision (%)	Recall (%)	F1 Score (%)	Mean F1 Score (%)
Random forest [28]	Leaf	94.38	66.52	77.98	77.12
	Wood	80.52	88.04	76.25	
LeWos [30]	Leaf	90.69	92.00	90.85	86.98
	Wood	86.30	85.25	83.10	
KPconv [38]	Leaf	94.48	98.62	96.98	95.12
	Wood	97.12	88.04	93.25	
Ours	Leaf	97.57	98.95	98.25	97.29
	Wood	97.77	94.94	96.33	

The leaf and wood point mean F1 score of this paper's method reached 97.29%, which are the highest accuracy, recall, and F1 score among these four methods. Compared with the random forest algorithm, which requires the design of a large number of manual features, the leaf point F1 score is improved by 20.27 points and the wood point F1 score accuracy is improved by 20.08 points. The KPconv algorithm, which has better performance in point cloud classification, has slightly lower classification accuracy than the method proposed in this paper. The KPconv training is set to 200 iterations, while the method proposed in this paper is set to only 50 iterations. This method also outperforms KPconv in terms of time and efficiency. An analysis of the classification accuracy in the table shows that the wood points have lower accuracy than the leaf points due to the complex canopy structure and the quality of the scanned data. Vicari et al. [34] argue that data quality is one of the main prerequisites for the successful separation of leaf and wood, and for tall trees with complex canopies, this challenge is even more severe. Data quality depends on a variety of factors, such as laser scanner characteristics, scan settings, alignment accuracy, and degree of occlusion. To overcome these challenges, Vicari et al. [34] suggested the use of an optimal scanning domain and high-resolution TLS. Another strategy proposed in Paynter et al. [43] is the application of dynamic adjustment to control the TLS data quality.

The accuracy of the method in this paper is comparable to or better than the state-of-the-art leaf and wood separation methods in previous studies. Moorthy et al. [42] proposed a leaf-wood classification method combining geometric features defined by radially bounded nearest neighbors on multiple spatial scales in a machine-learning model with an average overall accuracy of 94%. Ma et al. [7] proposed an improved salient feature-based leaf and wood separation method, which achieved an overall accuracy of 95%. Wang et al. [30] proposed an unsupervised geometry-based dynamic segmentation method for separating leaf and wood points with an average overall accuracy of 88%. Zhu et al. [26] combined geometric and radiometric features and used an adaptive radius nearest-neighbor search

algorithm for leaf and wood material identification with an overall accuracy of 84%, where the relatively low accuracy may be because the intensity data were not corrected for the effects of incidence angle and distance. Since the applicable scenarios and data sources are different from the methods in this paper, the methods in this paper are compared with the commonly used random forest classifier, LeWos method, and KPconv deep neural network on the test dataset. The results show that the accuracy of the method is comparable to existing methods for branch and leaf classification, providing a new alternative method for separating leaves and wood of single trees from a different and new perspective.

4.2. Ablation Analysis

To further analyze the effectiveness of the proposed module for deep neural networks, ablation experiments were conducted on the test dataset of this paper:

- The local context feature fusion module (LCFM), which enables each point to obtain its local geometry explicitly, was removed. After removing this module, the simple local point features were directly passed to the subsequent rearrangement attention module.
- The rearrangement attention module (RAM), which assigns attention weights to the disorderly arrangement of the point cloud, was removed. After removing this module, more representative information about each point was not selected.
- The residual connection optimization module (RCOM), which compensates for the lack of spatial information caused by nearest-neighbor upsampling, was removed. After removing this module, the features of the downsampled layer were directly interpolated by the nearest neighbor to recover the original spatial resolution.

Table 4 compares the Precision, Recall, and F1 score of all ablation networks. It can be seen that:

- Removing the fusion of local contextual features had the greatest impact on the overall performance of the network. The fusion of features effectively resolved the misclassification of stem and leaf points within the complex canopy.
- Removing the rearrangement attention module had the second-highest impact on the performance because it could not effectively retain representative features.
- Removing the residual connection optimization module reduced the performance because it could not effectively resolve the problem of information redundancy due to sampling.

Table 4. Network performance of the network at different settings.

	Precision (%)	Recall (%)	F1 Score (%)
Remove LCFM	92.21	91.63	91.92
Remove RAM	93.11	93.10	93.10
Remove RCOM	96.64	96.32	96.60
The full framework	97.64	97.63	97.63

From this ablation study, it can be seen that the modules complement each other to enhance each other, achieving the best performance of the model.

5. Discussion

5.1. The Number of Downsampled Points

The number of downsampled points may affect the leaf-wood separation performance of the proposed LWSNet. Following the work of [1], the limited number of downsampled points is insufficient to represent structural details, so complicated tree point clouds usually have unsatisfactory leaf-wood separation results. However, increasing the number of downsampled points costs more computational time without a great performance improvement. To avoid this situation, we increase the number of downsampled points of each tree point cloud while satisfying the limit of the computation capacity. Simply put, we explore the

influence of the number of downsampled points on the leaf-wood separation performance of the proposed LWSNet.

According to our observations, the convergence ability of a network relies on the number of downsampled points. To evaluate the impact of the number of downsampled points on the leaf-wood separation performance, the number of downsampled points should ensure that a tree point cloud is included without too many background points. Accordingly, we perform some comparison experiments with the number of downsampled points empirically defined as 512, 1024, 1536, 2048, 2560, 3072, 3584, 4096, 4608, and 5120, respectively. The quantitative evaluation results obtained using the different number of downsampled points are provided in Table 5. We can find that the leaf-wood separation performance increases with the number of downsampled points. When the number of downsampled points is less than 2048, the evaluation indicators of the leaf-wood separation task are low since insufficient input points may result in difficulty in the local semantic connectivity learning. When the number of downsampled points increases from 2048 to 4096, the leaf-wood separation performance increases significantly. When the number of downsampled points is 4096, all evaluation indicators are above 97%. Within the range 4096–5120, the leaf-wood separation performance growth is slow. Although the larger number of downsampled points brings in better leaf-wood separation results, we must consider the trade-off between accuracy and efficiency for leaf-wood separation in tree point clouds. Therefore, 4096 downsampled points not only obtain satisfactory leaf-wood separation results for LWSNet, but also consume a lower processing time.

Table 5. Quantitative evaluation results obtained by different numbers of downsampled points.

	Precision (%)	Recall (%)	F1 Score (%)
512	94.87	94.93	94.90
1024	95.31	95.09	95.22
1536	96.24	96.12	96.20
2048	97.08	97.15	97.10
2560	97.40	97.42	97.41
3072	97.46	97.51	97.48
3584	97.60	97.58	97.59
4096	97.64	97.63	97.63
4608	97.65	97.65	97.65
5120	97.67	97.66	97.66

5.2. Generalization Capability on Different Types of Plants

In order to prove that the proposed method is robust enough to be applied to different types of plants, generalization experiments were conducted on two different types plant point cloud datasets. We used our LWSNet trained on the tree point clouds to predict labels for the Pheno4D dataset [44] directly without retraining. The Pheno4D dataset has two crops (i.e., maize and tomato) scanned by LiDAR, including about 260 million manual annotated 3D points. We visualize the qualitative segmentation results of the proposed LWSNet on the Pheno4D dataset in Figure 8. As can be seen, the proposed LWSNet successfully distinguishes the exact junction between the leaf and the stem. Note the tomatoes have more complex canopy structures (e.g., more leaves) compared with the maize in the Pheno4D dataset. Specifically, leaf-stem separation on tomato crops is regarded as a challenging task because they have large variances in the spatial structure and the degree of the leaf curvature. However, the proposed LWSNet has satisfactory leaf-stem separation results, which shows that our segmentation method can be used to process crop point clouds. Meanwhile, the numerical comparisons are also provided in Table 6. KPconv [38] and Eff-3DPSeg [45] are selected for comparison with the proposed LWSNet, and we used recommendations from their original configuration selection. It can be seen that our LWSNet has the best leaf-stem separation performance across all methods compared in Table 6. The Precision, Recall, F1-score, and Intersection over Union (IoU) (leaf

and stem) of LWSNet are all followed by Eff-3DPSeg [45] with gaps of around 1%, while the KPconv [38] is slightly inferior to Eff-3DPSeg [45] by about 0.1%/1.4%, 0.5%/1.1%, 0.3%/2.1%, and 0.6%/2.4%, respectively. The proposed LWSNet has balanced advantages on both maize and tomato plants, showing excellent adaptability in different plant species.

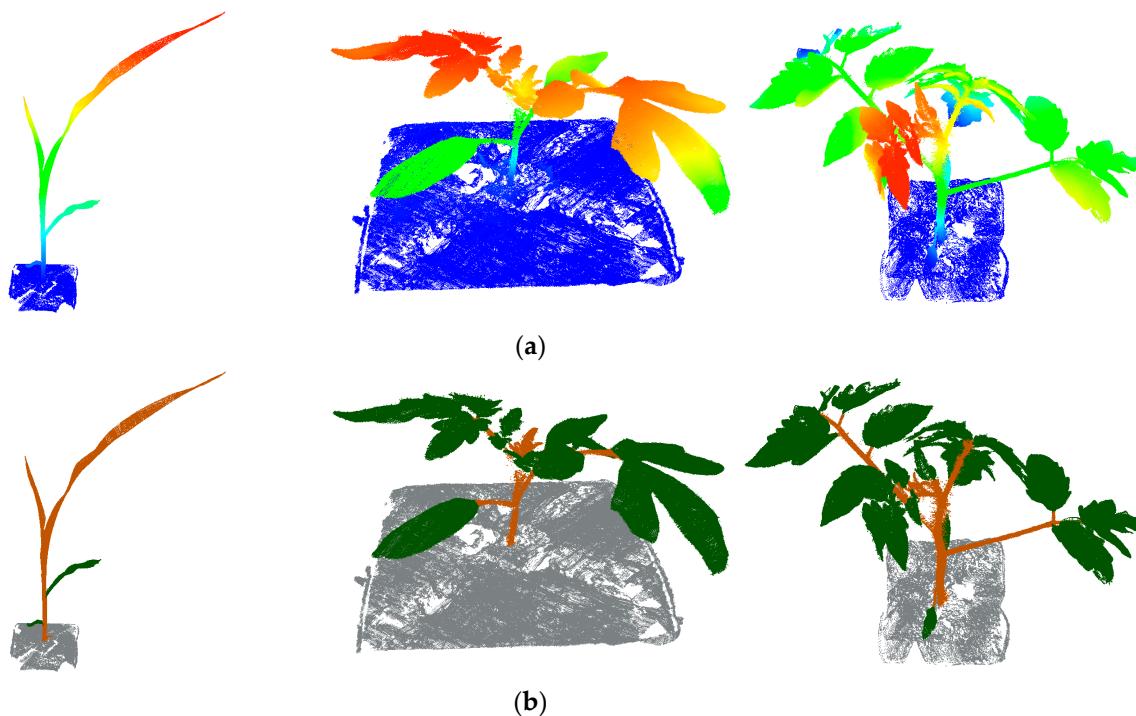


Figure 8. Example of leaf-stem separation on Eff-3DPSeg [45] dataset: (a) raw point clouds colored by elevation; and (b) leaf-stem separation results, where gray, brown, and green points represent points from ground, stems, and leaves, respectively. Note that the first point cloud is maize, and the last two point clouds are tomatoes.

Table 6. Quantitative evaluation of leaf-stem separation on Eff-3DPSeg [45] dataset.

Method		Precision (%)	Recall (%)	F1 Score (%)	IoU (%)
KPconv [38]	Leaf	98.7	97.0	97.5	94.9
	Stem	86.2	91.3	88.2	80.1
Eff-3DPSeg [45]	Leaf	99.4	99.2	99.3	98.6
	Stem	94.7	95.9	95.3	91.0
LWSNet	Leaf	99.5	99.7	99.6	99.2
	Stem	96.1	97.0	97.4	93.4

Moreover, the above tree point clouds and Pheno4D dataset were collected by the laser scanner, which is expensive, and the data acquisition process is time-consuming. Thus, we hope to check out whether the pre-trained LWSNet can work directly on the leaf-stem separation task of crop point clouds captured from the multi-view stereo (MVS) technique. In this generalization experiment, two ornamental plants—*Codiaeum variegatum* and *Hibiscus rosa-sinensis Linn.*—are constructed by the MVS technique, which is very different from LiDAR point clouds [46]. A total of 45 and 50 phone images with a resolution of 4032×3024 are collected for a *Codiaeum variegatum* and a *Hibiscus rosa-sinensis Linn.*, respectively. Then, MVS is applied to obtain two plant point clouds shown in Figure 9a. Figure 9b shows the leaf-stem separation results on *Codiaeum variegatum* and *Hibiscus rosa-sinensis Linn.* by the pre-trained LWSNet, respectively. As can be perceived, the proposed LWSNet can also generate satisfactory leaf-stem separation results for crop MVS point clouds, showing LWSNet is also good at recognizing different crop species with varied

shapes and orientations. The quantitative comparison results for two MVS point clouds are listed in Table 7; all quantitative measures are larger than 90%, which again demonstrates that the proposed LWSNet is versatile enough to produce good leaf-stem separation results for different plant species of point clouds.

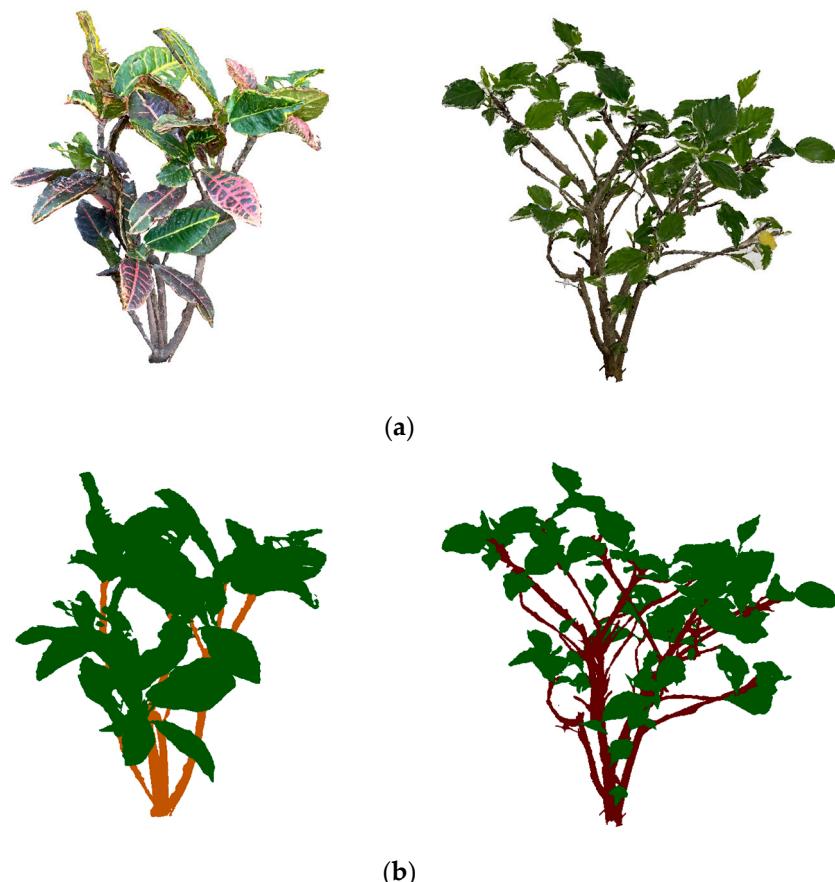


Figure 9. Example of leaf-stem separation on plant MVS point clouds: (a) raw point clouds colored by RGB; and (b) leaf-stem separation results, where brown and green points represent points from stems and leaves. Note that the left point cloud is *Codiaeum variegatum*, and the right point cloud is *Hibiscus rosa-sinensis Linn.*

Table 7. Quantitative evaluation of leaf-stem separation on two plant MVS point clouds.

		Precision (%)	Recall (%)	F1 Score (%)	IoU (%)
<i>Codiaeum variegatum</i>	Leaf	95.1	95.4	95.2	94.1
	Stem	92.3	92.8	92.5	90.7
<i>Hibiscus rosa-sinensis Linn.</i>	Leaf	94.3	95.2	94.6	93.9
	Stem	91.7	92.2	91.9	90.2

5.3. Future Improvements

Due to the massive nature of point cloud data, the available hardware devices have limitations in terms of computational memory storage and capacity [47]. Therefore, a random sampling approach is often used to reduce the number of point clouds. However, this approach can result in the loss of accurate contextual information, despite the presence of specific modules designed to compensate for this. As technology continues to evolve, the development of robust and efficient sampling methods will lead to significant improvements in the application of deep learning to point clouds. Additionally, different tree species with varying heights and interlaced canopies present a significant challenge

to leaf-wood separation. Although this paper has achieved some results in the algorithm, it still struggles to distinguish between severe overlapping canopies, and when there is too little distance between leafy trees. Therefore, further research that incorporates ecological theory is necessary to address these complex cases of trees and integrate them into deep-learning processing.

6. Conclusions

Existing algorithms based on the traditional feature descriptors still have their limitations in the face of challenges in the leaf-wood separation task. Therefore, we propose a novel deep-learning method, named LWSNet, to make full use of the geometric features of point clouds and enhance the expressiveness of local contextual features. Meanwhile, the robustness of the proposed LWSNet is enhanced by nearest-neighbor interpolation and residual connections, thus improving the classification accuracy of leaf-wood separation. The proposed LWSNet is validated on eight trees of different species and sizes, achieving the highest average F1 score of 97.29% and outperforming three existing methods (i.e., random forest [28], LeWos [30], and KPconv [38]). The ablation experiments show that the geometric features have the greatest impact on the leaf-wood separation performance of tree point clouds. Our study further demonstrates the potential of the rearrangement attention mechanism in leaf-wood separation, and future work should focus on its application in tree model reconstruction.

Although the proposed LWSNet achieves reasonable leaf-wood separation performance in experimental sets, it still has difficulty in distinguishing between the cases of severely overlapping canopies and leaves with distances that are too small between them. To better solve this problem, further research is necessary to combine ecological theory and fully exploit the potential features of leaf and wood point clouds. Furthermore, we will enrich the plant point cloud datasets by using different 3D sensors and reconstruction techniques to add more new species to the dataset to form a more robust leaf-wood separation network.

Author Contributions: Conceptualization, T.J., Q.Z. and C.L.; methodology, T.J., Q.Z. and C.L.; software, T.J. and C.L.; validation, T.J., Q.Z., S.L. and C.L.; formal analysis, T.J., Q.Z. and Z.Z.; investigation, Q.Z., L.D. and Z.Z.; resources, T.J., C.L. and Y.W.; data curation, T.J., C.L., L.D. and Y.W.; writing—original draft preparation, T.J., Q.Z., S.L., C.L. and Y.W.; writing—review and editing, T.J., Q.Z., S.L., C.L., J.S. and Y.W.; visualization, T.J., Q.Z., C.L. and Z.Z.; supervision, T.J., J.S. and Y.W.; project administration, T.J., J.S. and Y.W.; funding acquisition, T.J., J.S. and Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 42171446 and 41771439.

Data Availability Statement: Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Acknowledgments: The authors acknowledge all reviewers for their valuable comments.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Li, D.; Shi, G.; Li, J.; Chen, Y.; Zhang, Y.; Xiang, S.; Jin, S. PlantNet: A dual-function point cloud segmentation network for multiple plant species. *ISPRS J. Photogramm. Remote Sens.* **2021**, *184*, 243–263. [[CrossRef](#)]
2. Tan, K.; Ke, T.; Tao, P.; Liu, K.; Duan, Y.; Zhang, W.; Wu, S. Discriminating Forest Leaf and Wood Components in TLS Point Clouds at Single-Scan Level Using Derived Geometric Quantities. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5701517. [[CrossRef](#)]
3. Jiang, T.; Wang, Y.; Liu, S.; Cong, Y.; Dai, L.; Sun, J. Local and global structure for urban ALS point cloud semantic segmentation with ground-aware attention. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5702615. [[CrossRef](#)]
4. Zhang, W.; Wan, P.; Wang, T.; Cai, S.; Chen, Y.; Jin, X.; Yan, G. A Novel Approach for the Detection of Standing Tree Stems from Plot-Level Terrestrial Laser Scanning Data. *Remote Sens.* **2019**, *11*, 211. [[CrossRef](#)]

5. Hu, T.; Wei, D.; Su, Y.; Wang, X.; Zhang, J.; Sun, X.; Liu, Y.; Guo, Q. Quantifying the shape of urban street trees and evaluating its influence on their aesthetic functions based on mobile LiDAR data. *ISPRS J. Photogramm. Remote Sens.* **2022**, *184*, 203–214. [[CrossRef](#)]
6. Dai, W.; Yang, B.; Dong, Z.; Shaker, A. A new method for 3D individual tree extraction using multispectral airborne LiDAR point clouds. *ISPRS J. Photogramm. Remote Sens.* **2018**, *144*, 400–411. [[CrossRef](#)]
7. Ma, L.; Zheng, G.; Eitel, J.U.; Moskal, L.M.; He, W.; Huang, H. Improved Salient Feature-Based Approach for Automatically Separating Photosynthetic and Nonphotosynthetic Components Within Terrestrial Lidar Point Cloud Data of Forest Canopies. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 679–696. [[CrossRef](#)]
8. Vicari, M.B.; Disney, M.I.; Wilkes, P.; Burt, A.; Calders, K.; Woodgate, W. Leaf and wood classification framework for terrestrial LiDAR point clouds. *Methods Ecol. Evol.* **2019**, *10*, 680–694. [[CrossRef](#)]
9. Zhang, J.; Wang, J.; Dong, P.; Ma, W.; Liu, Y.; Liu, Q.; Zhang, Z. Tree stem extraction from TLS point-cloud data of natural forests based on geometric features and DBSCAN. *Geocarto Int.* **2022**, *37*, 10392–10406. [[CrossRef](#)]
10. Hui, Z.; Jin, S.; Xia, Y.; Wang, L.; Ziggah, Y.Y.; Cheng, P. Wood and leaf separation from terrestrial LiDAR point clouds based on mode points evolution. *ISPRS J. Photogramm. Remote Sens.* **2021**, *178*, 219–239. [[CrossRef](#)]
11. Wang, D.; Brunner, J.; Ma, Z.; Lu, H.; Hollaus, M.; Pang, Y.; Pfeifer, N. Separating Tree Photosynthetic and Non-Photosynthetic Components from Point Cloud Data Using Dynamic Segment Merging. *Forests* **2018**, *9*, 252. [[CrossRef](#)]
12. Qi, C.; Su, H.; Mo, K.; Guibas, L.J. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 77–85.
13. Wang, Y.; Jiang, T.; Liu, J.; Li, X.; Liang, C. Hierarchical instance recognition of individual roadside trees in environmentally complex urban areas from UAV laser scanning point clouds. *ISPRS Int. J. GeoInf.* **2020**, *9*, 595. [[CrossRef](#)]
14. Windrim, L.; Bryson, M. Detection, Segmentation, and Model Fitting of Individual Tree Stems from Airborne Laser Scanning of Forests Using Deep Learning. *Remote Sens.* **2020**, *12*, 1469. [[CrossRef](#)]
15. Krisanski, S.; Taskhiri, M.S.; Aracil, S.G.; Herries, D.; Turner, P. Sensor Agnostic Semantic Segmentation of Structurally Diverse and Complex Forest Point Clouds Using Deep Learning. *Remote Sens.* **2021**, *13*, 1413. [[CrossRef](#)]
16. Dai, W.; Jiang, Y.; Zeng, W.; Chen, R.; Xu, Y.; Zhu, N.; Xiao, W.; Dong, Z.; Guan, Q. MDC-Net: A multi-directional constrained and prior assisted neural network for wood and leaf separation from terrestrial laser scanning. *Int. J. Digit. Earth* **2023**, *16*, 1224–1245. [[CrossRef](#)]
17. Koma, Z.; Rutzinger, M.; Bremer, M. Automated Segmentation of Leaves from Deciduous Trees in Terrestrial Laser Scanning Point Clouds. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 1456–1460. [[CrossRef](#)]
18. Jin, S.; Su, Y.; Gao, S.; Wu, F.; Ma, Q.; Xu, K.; Ma, Q.; Hu, T.; Liu, J.; Pang, S.; et al. Separating the Structural Components of Maize for Field Phenotyping Using Terrestrial LiDAR Data and Deep Convolutional Neural Networks. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 2644–2658. [[CrossRef](#)]
19. Ayrey, E.; Hayes, D.J. The Use of Three-Dimensional Convolutional Neural Networks to Interpret LiDAR for Forest Inventory. *Remote Sens.* **2018**, *10*, 649. [[CrossRef](#)]
20. Wang, Y.; Jiang, T.; Yu, M.; Tao, S.; Sun, J.; Liu, S. Semantic-Based Building Extraction from LiDAR Point Clouds Using Contexts and Optimization in Complex Environment. *Sensors* **2020**, *20*, 3386. [[CrossRef](#)]
21. Klokov, R.; Lempitsky, V.S. Escape from Cells: Deep Kd-Networks for the Recognition of 3D Point Cloud Models. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 863–872.
22. Li, Y.; Bu, R.; Sun, M.; Wu, W.; Di, X.; Chen, B. PointCNN: Convolution on X-Transformed Points. In Proceedings of the Advances in Neural Information Processing Systems, Montréal, Canada, 3–8 December 2018; pp. 820–830.
23. Qi, C.R.; Yi, L.; Su, H.; Guibas, L.J. PointNet++: Deep hierarchical feature learning on point sets in a metric space. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 5099–5108.
24. Eitel, J.U.; Vierling, L.A.; Long, D.S. Simultaneous measurements of plant structure and chlorophyll content in broadleaf saplings with a terrestrial laser scanner. *Remote Sens. Environ.* **2010**, *114*, 2229–2237. [[CrossRef](#)]
25. Li, D.; Li, J.; Xiang, X.; Pan, A. PSegNet: Simultaneous Semantic and Instance Segmentation for Point Clouds of Plants. *Plant Phenomics* **2022**, *2022*, 9787643. [[CrossRef](#)]
26. Zhu, X.; Skidmore, A.K.; Darvishzadeh, R.; Niemann, K.O.; Liu, J.; Shi, Y.; Wang, T. Foliar and woody materials discriminated using terrestrial LiDAR in a mixed natural forest. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *64*, 43–50. [[CrossRef](#)]
27. Wang, D.; Hollaus, M.; Pfeifer, N. Feasibility of machine learning methods for separating wood and leaf points from terrestrial laser scanning data. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *4*, 157–164. [[CrossRef](#)]
28. Yun, T.; An, F.; Li, W.; Sun, Y.; Cao, L.; Xue, L. A Novel Approach for Retrieving Tree Leaf Area from Ground-Based LiDAR. *Remote Sens.* **2016**, *8*, 942. [[CrossRef](#)]
29. Ferrara, R.; Virdis, S.G.; Ventura, A.; Ghisu, T.; Duce, P.; Pellizzaro, G. An automated approach for wood-leaf separation from terrestrial LIDAR point clouds using the density based clustering algorithm DBSCAN. *Agric. For. Meteorol.* **2018**, *262*, 434–444. [[CrossRef](#)]
30. Wang, D.; Momo Takoudjou, S.; Casella, E. LeWoS: A universal leaf-wood classification method to facilitate the 3D modelling of large tropical trees using terrestrial LiDAR. *Methods Ecol. Evol.* **2020**, *11*, 376–389. [[CrossRef](#)]
31. Wang, D. Unsupervised semantic and instance segmentation of forest point clouds. *ISPRS J. Photogramm. Remote Sens.* **2020**, *165*, 86–97. [[CrossRef](#)]

32. Wan, P.; Shao, J.; Jin, S.; Wang, T.; Yang, S.; Yan, G.; Zhang, W. A novel and efficient method for wood–leaf separation from terrestrial laser scanning point clouds at the forest plot level. *Methods Ecol. Evol.* **2021**, *12*, 2473–2486. [[CrossRef](#)]
33. Tan, K.; Zhang, W.; Dong, Z.; Cheng, X.; Cheng, X. Leaf and Wood Separation for Individual Trees Using the Intensity and Density Data of Terrestrial Laser Scanners. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 7038–7050. [[CrossRef](#)]
34. Disney, M.I.; Boni Vicari, M.; Burt, A.; Calders, K.; Lewis, S.L.; Raumonen, P.; Wilkes, P. Weighing trees with lasers: Advances, challenges and opportunities. *Interface Focus* **2018**, *8*, 20170048. [[CrossRef](#)]
35. Calders, K.; Disney, M.I.; Armston, J.; Burt, A.; Brede, B.; Origio, N.; Muir, J.; Nightingale, J.M. Evaluation of the Range Accuracy and the Radiometric Calibration of Multiple Terrestrial Laser Scanning Instruments for Data Interoperability. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 2716–2724. [[CrossRef](#)]
36. Xu, S.; Zhou, K.; Sun, Y.; Yun, T. Separation of Wood and Foliage for Trees from Ground Point Clouds Using a Novel Least-Cost Path Model. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 6414–6425. [[CrossRef](#)]
37. Tian, Z.; Li, S. Graph-Based Leaf-Wood Separation Method for Individual Trees Using Terrestrial Lidar Point Clouds. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5705111. [[CrossRef](#)]
38. Thomas, H.; Qi, C.; Deschaud, J.; Marcotegui, B.; Goulette, F.; Guibas, L.J. KPConv: Flexible and Deformable Convolution for Point Clouds. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea, 27 October 2019–2 November 2019; pp. 6411–6420.
39. Wang, P.; Tang, Y.; Liao, Z.; Yan, Y.; Dai, L.; Liu, S.; Jiang, T. Road-Side Individual Tree Segmentation from Urban MLS Point Clouds Using Metric Learning. *Remote Sens.* **2023**, *15*, 1992. [[CrossRef](#)]
40. Ao, Z.; Wu, F.; Hu, S.; Sun, Y.; Su, Y.; Guo, Q.; Xin, Q. Automatic segmentation of stem and leaf components and individual maize plants in field terrestrial LiDAR data using convolutional neural networks. *Crop J.* **2022**, *10*, 1239–1250. [[CrossRef](#)]
41. Li, Y.; Wen, W.; Miao, T.; Wu, S.; Yu, Z.; Wang, X.; Guo, X.; Zhao, C. Automatic organ-level point cloud segmentation of maize shoots by integrating high-throughput data acquisition and deep learning. *Comput. Electron. Agric.* **2022**, *193*, 106702. [[CrossRef](#)]
42. Krishna Moorthy, S.M.; Calders, K.; Vicari, M.B.; Verbeeck, H. Improved Supervised Learning-Based Approach for Leaf and Wood Classification from LiDAR Point Clouds of Forests. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 3057–3070. [[CrossRef](#)]
43. Paynter, I.; Genest, D.C.; Saenz, E.J.; Peri, F.; Li, Z.; Strahler, A.H.; Schaaf, C. Quality Assessment of Terrestrial Laser Scanner Ecosystem Observations Using Pulse Trajectories. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 6324–6333. [[CrossRef](#)]
44. Schunck, D.; Magistri, F.; Rosu, R.A.; Cornelissen, A.; Chebrolu, N.; Paulus, S.; Léon, J.; Behnke, S.; Stachniss, C.; Kuhlmann, H.; et al. Pheno4D: A spatio-temporal dataset of maize and tomato plant point clouds for phenotyping and advanced plant analysis. *PLoS ONE* **2021**, *16*, 0256340. [[CrossRef](#)]
45. Luo, L.; Jiang, X.; Yang, Y.; Samy, E.R.; Lefsrud, M.G.; Hoyos-Villegas, V.; Sun, S. Eff-3DPSeg: 3D organ-level plant shoot segmentation using annotation-efficient point clouds. *arXiv* **2022**, arXiv:2212.10263.
46. Jiang, T.; Sun, J.; Liu, S.; Zhang, X.; Wu, Q.; Wang, Y. Hierarchical semantic segmentation of urban scene point clouds via group proposal and graph attention network. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *105*, 102626. [[CrossRef](#)]
47. Du, R.; Ma, Z.; Xie, P.; He, Y.; Cen, H. PST: Plant segmentation transformer for 3D point clouds of rapeseed plants at the podding stage. *ISPRS J. Photogramm. Remote Sens.* **2023**, *195*, 380–392. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.