


Article

# Hierarchical Classification of Urban ALS Data by Using Geometry and Intensity Information

Xiaoqiang Liu <sup>1,2</sup>, Yanming Chen <sup>1,2,3,\*</sup> , Shuyi Li <sup>1,2</sup>, Liang Cheng <sup>1,2,3</sup> and Manchun Li <sup>1,2,3,\*</sup>

<sup>1</sup> School of Geography and Ocean Science, Nanjing University, Nanjing 210093, China; mg1727062@nju.edu.cn (X.L.); dz1627003@nju.edu.cn (S.L.); lcheng@nju.edu.cn (L.C.)

<sup>2</sup> Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, Nanjing University, Nanjing 210093, China

<sup>3</sup> Collaborative Innovation Center for the South Sea Studies, Nanjing University, Nanjing 210093, China

\* Correspondence: chenyanming@nju.edu.cn (Y.C.); limanchun@nju.edu.cn (M.L.)

Received: 28 August 2019; Accepted: 18 October 2019; Published: 21 October 2019



**Abstract:** Airborne laser scanning (ALS) can acquire both geometry and intensity information of geo-objects, which is important in mapping a large-scale three-dimensional (3D) urban environment. However, the intensity information recorded by ALS will be changed due to the flight height and atmospheric attenuation, which decreases the robustness of the trained supervised classifier. This paper proposes a hierarchical classification method by separately using geometry and intensity information of urban ALS data. The method uses supervised learning for stable geometry information and unsupervised learning for fluctuating intensity information. The experiment results show that the proposed method can utilize the intensity information effectively, based on three aspects, as below. (1) The proposed method improves the accuracy of classification result by using intensity. (2) When the ALS data to be classified are acquired under the same conditions as the training data, the performance of the proposed method is as good as the supervised learning method. (3) When the ALS data to be classified are acquired under different conditions from the training data, the performance of the proposed method is better than the supervised learning method. Therefore, the classification model derived from the proposed method can be transferred to other ALS data whose intensity is inconsistent with the training data. Furthermore, the proposed method can contribute to the hierarchical use of some other ALS information, such as multi-spectral information.

**Keywords:** airborne laser scanning; hierarchical classification; intensity; geometry; supervised learning; unsupervised learning

## 1. Introduction

By the year 2050, 68% of the world population is expected to live in urban areas [1], thus increasing the importance of urban morphology and ecology research, in which two-dimensional land cover products are commonly used [2,3]. However, two-dimensional land cover products cannot represent the vertical differentiation of geo-objects. Thus, airborne laser scanning (ALS), which can directly acquire three-dimensional geometry information of geo-objects, has been introduced into urban morphology and ecology research [4,5], such as change detection [6,7] and carbon storage mapping [8,9]. Unfortunately, a lot of different types of geo-objects, such as buildings, vegetation, and cars, may appear in a small urban local neighborhood [10,11], and it is difficult to automatically extract all geo-objects in the urban environment from raw ALS geometry information. Consequently, many researchers only extracted the ground [12,13], the buildings [14,15], or the powerlines [16,17]. On the other hand, geometry information acquired by ALS was often integrated with passive multi- or hyper-spectral remote sensing image for land cover classification [18–21]. Nevertheless, the automatic fusion of ALS

and passive remote sensing image faces its own challenges, owing to the ALS data and the optical images having different characteristics [22–24].

Besides recording points some ALS devices also record spectral information called “intensity”, which can enable the separation of man-made and natural geo-objects [25,26]. With the development of full-waveform and multi-spectral laser scanning, the intensity information will become an important part of the ALS data [4,27]. As compared to the fusion spectral from passive remote sensing image, the intensity recorded by ALS has several advantages: (1) it is independent of illumination conditions such as shading since LiDAR instruments provide their own light source; (2) ALS allows a vertical differential of the intensity; and, (3) spectral mixing within the measured intensity is minimized [4].

Many researchers utilized the supervised learning method with both geometry and intensity information to classify the ALS data [28–30]. However, the intensity recorded by ALS is fluctuating due to airborne flight height, atmospheric transmittance, detector responsivity, in-situ calibration, time conditions, and others [31], which resulted in the supervised classifier, directly trained using both geometry and intensity information, not being robust, and being hard to transfer to other ALS data. On the other hand, intensity information is more important to distinguish ground-level geo-objects, such as roads and low vegetation [32,33], and it is not elegant to ignore this once this is available.

This paper proposes a method using supervised learning for the stable geometry information and unsupervised learning for the fluctuating intensity information of the ground-level points to reasonably utilize the intensity information recorded by ALS. The proposed method includes three core parts. First, it uses the training data with the geometry feature to train the supervised classifier, and then the trained supervised classifier is used to classify the ALS data. Second, it extracts ground-level points from the classified ALS data and utilizes unsupervised learning with intensity information to reclassify these ground-level points. Finally, it combines the results of the supervised and unsupervised learning by the heuristic rule.

The main contributions of our proposed method are as follows. (1) In the first part, the geometry feature used in the proposed method contains “absolute feature” (e.g., height, normal) varied depending on transformation and rotation, and “relative feature” (e.g., distribution of points), which is invariant to transformation and rotation. This concatenated geometry feature represents the distribution of point neighborhood and the intuition of geo-objects with special direction and height. (2) In the second part, unsupervised learning does not need to be trained and its parameters can be re-estimated in ALS data to be classified. Thus, the classification model that is derived from the proposed method is more robust to fluctuating intensity. (3) In the third part, a heuristic rule is designed to integrate the urban ALS classification results of the supervised and unsupervised learning, which is based on our defined confidence of supervised and unsupervised learning classification results.

After providing a brief summary of related work in Section 2, we clearly explain the proposed method in Section 3. Subsequently, the results are shown in Section 4 and they are discussed in Section 5. Finally, we provide concluding remarks as well as suggestions for future work in Section 6.

## 2. Related Work

ALS data classification belongs to point clouds classification and can be categorized into point-wise and segment-wise method according to the classification primitive. The point-wise method individually classifies each point of the ALS data by using the respective features as the inputs for a standard supervised classifier [10,34]. The segment-wise method first divides the ALS data into segments and then assigns class labels to the segments so that all points within a segment obtain the same class label [35]. The segment-wise method uses more features, like segment size and shape, and therefore obtains a smoother classification result than the point-wise method. However, the performance of the segment-wise method is negatively affected by under- and over-segmentation errors [35]. On the contrary, the point-wise method directly classifies the ALS data point by point without segmenting and extracting features from the segmented object; hence, its result reveals a “pepper and salt noise” behavior, because it ignores the correlation between the labels of the nearby pixels [11,34,36].

Context information, which is the semantical label of a point similar to its nearby points, is usually introduced to smooth the point-wise classification. Schindler gave an overview and comparison of some commonly used filter methods, such as the majority filter, the Gaussian filter, the bilateral filter, and the edge-aware filter for remote sensing image [36], and these methods can be easily adapted to point clouds. Unlike these filter methods, which only handle local information, in the global smoothing method the entire classification of point clouds uses a graph structure, where the contextual information is contained in the adjacent edge of the graph. Conditional random field (CRF) is a commonly used graph model and it adds pairwise potential alongside the unary potential to model the contextual information. Schindler adopted the Potts model (or the contrast-sensitive Potts model) to calculate the pairwise potential and then inferenced the final classification by the approximating method, such as using the  $\alpha$ -expansion graph cut and semiglobal labeling [36]. Niemeyer et al. used the random forest (or linear model) with feature concatenating two neighboring points to learn the pairwise potential [10]. This method of learning the pairwise potential is also called contextual learning method. The contextual learning method considers the smooth in the training time and obtains more contextual information; for instance, it is more probable that cars are situated on a street than on grassland [10]. Secondly, the contextual learning method based on segments can obtain long-range interactions [30,35]. However, it requires the extensive computation of contextual classifiers and the extraction of useful interaction features. The post-processing smoothing technique, such as global smoothing method in [36], is time efficient and flexible enough to be conducted on point clouds without training an additional supervised classifier as compared to the contextual learning method.

Another post-processing smoothing technique is regularization. Landrieu et al. considered the problem of spatially smoothing semantic classification of point clouds from a structured regularization perspective, whose goal is to find an improved labeling with increased spatial smoothness while remaining as close as possible to the initial classification [34]. Under the framework that was proposed in [34], Li et al. utilized optimal graph and probabilistic label relaxation to handle large wrongly labeled regions [28].

The unary potential of CRF and the regularization method both used initial labeling derived with the standard classification. Hence, this study focuses on the standard classification rather than these smooth methods considering contextual information, because the better its result is, the better further processing will perform [37]. A variety of supervised learning methods have been applied for ALS classification, including support vector machine [38,39], boosting [40,41], and random forest [10,42]. In these classification methods, the generation of good features is a vital part of obtaining a good performance.

There have been three major geometry features: covariance-based features, histogram-based features, and deep features. The covariance-based feature is derived from a three-dimensional covariance matrix of a point and its neighbors [43]. Dittrich et al. investigated the accuracy and robustness of the covariance feature [44]. The histogram-based feature accumulates information regarding the spatial interconnection between a point and its neighborhood into a histogram [45–48]. Point feature histograms and its improved versions, denoted as fast point feature histograms (FPFH), were typical histogram-based features. The deep feature is learned from training data by using deep neural networks, which have been rapidly growing in recent years. However, the interpretability has been identified as a potential weakness of deep neural networks [49,50]. The covariance feature is usually used in the ALS classification, whereas the histogram-based feature is more rare. Therefore, this study tests the feasibility of the histogram-based feature in large scale ALS point clouds.

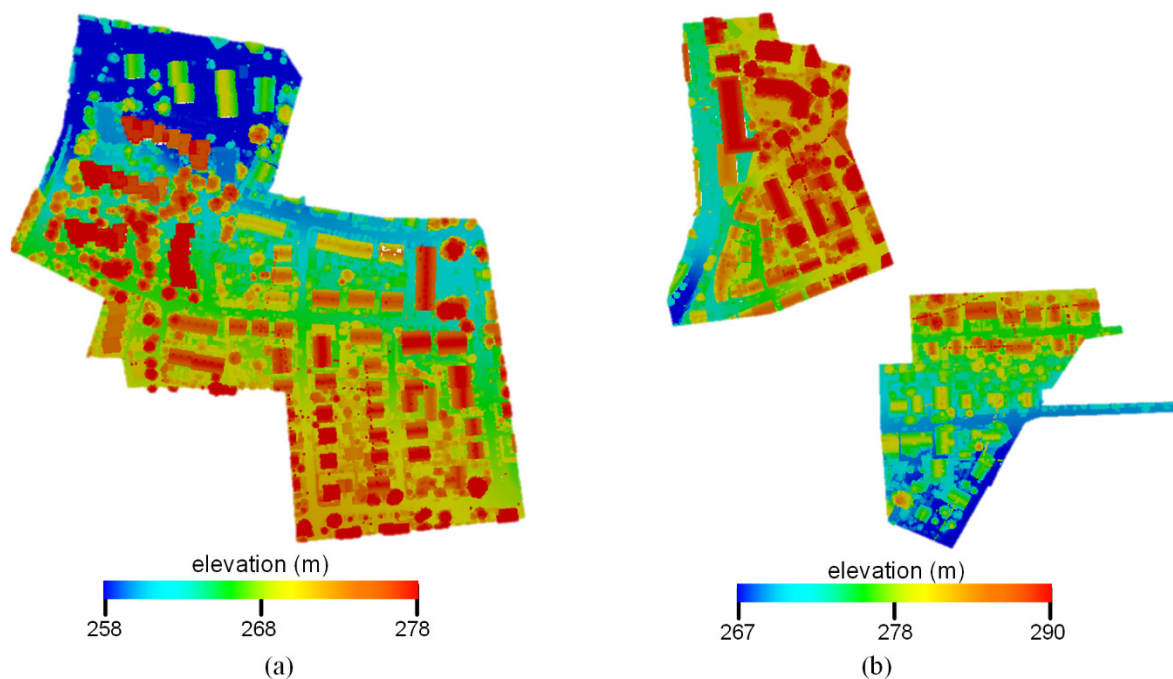
Except for the geometry feature, the intensity was usually used to extract some feature [29,30,51], but it is fluctuating, owing to the system and environmental induced distortions. [52,53] improved the classification accuracy of the airborne LiDAR intensity data by calibrating the intensity. A few factors, such as incidence of angle, range calculated with GPS assistant, and atmospheric effect derived from simulating or insite measurement, should be considered while calibrating the intensity, as reasonably utilizing the intensity recorded by ALS is difficult. Apart from these classifications that are needed to

calibrate intensity, this study attempts to use unsupervised learning to handle intensity, which does not require calibration of intensity and can be transferred to other ALS data whose acquired condition differs from the training data.

### 3. Materials and Methods

#### 3.1. Data Used in This Study

This study used the benchmark data (as shown in Figure 1) provided by the International Society for Photogrammetry and Remote Sensing (ISPRS) [54] to evaluate the performance of the proposed method. This ALS data was acquired by Leica Geosystems using a Leica ALS50 system with a  $45^\circ$  field of view and a mean flying height of 500 m above the ground. The average strip overlap was 30%, and the mean point density was eight points  $\cdot$  m<sup>-2</sup>. Multiple echoes and intensities were recorded. The number of points with multiple echoes was relatively low, owing to the leave-on-condition at the time of the data acquisition. From the scanned data, the ISPRS 3D Semantic Labeling Contest selected two data subsets for the three-dimensional labeling challenge. In total, 9 classes have been defined, namely power line (PL), low vegetation (LV), impervious surface (IS), car (Car), fence/hedge (Fence), roof (Roof), façade (Façade), shrub (Shrub), and tree (Tree). The training and testing areas are from Vaihingen city. The training area consists of a few high-rising buildings surrounded by trees and a purely residential neighborhood with small, detached houses (as shown in Figure 1a). It covers an area of 399 m  $\times$  421 m and contains 753,876 points. Dense, complex buildings and some trees characterize the testing area (as shown in Figure 1b), which covers an area of 389 m  $\times$  419 m and contains 411,722 points.



**Figure 1.** International Society for Photogrammetry and Remote Sensing (ISPRS) benchmark data: (a) training data, (b) testing data.

#### 3.2. Graphical Overview of the Proposed Method

Figure 2 gives a graphical overview of the proposed method, and it consists of three major components:

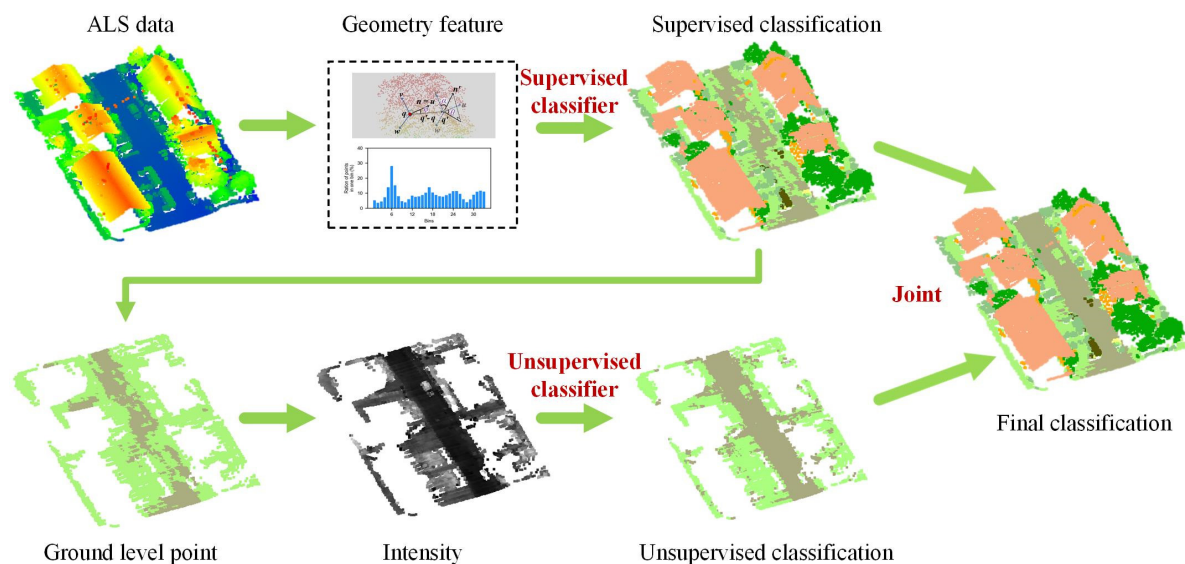
- Supervised learning for geometry information: extract geometry feature from the ALS data and then classify the ALS data by using the trained supervised classifier. The geometry feature in



this study included FPFH, normal, and height. Four common supervised learning methods, namely decision tree (DT), random forest (RF), support vector classification (SVC), and extreme gradient boost (XGBoost) were used alone to test the performance of the proposed method.

- Unsupervised learning for intensity information: after applying the supervised classifier on the ALS data, the ground-level points and the elevated points (such as building and tree) can be split from the ALS data. The ground-level points were reclassified while using an unsupervised learning method based on intensity information. The Gaussian mixture model (GMM) was the selected unsupervised classifier, because the probability distribution of the intensity of some geo-objects is approximately Gaussian distribution [31].
- Join the classification results of the supervised and unsupervised classifier: for the elevated points, the label was the result of the supervised classifier, whereas, for the ground-level points, the label was the selection from the supervised classification result and the unsupervised classification result based on the heuristic rule. Section 3.3 describes this heuristic rule.

The classifiers derived from the supervised and unsupervised learning should be soft, that is, their outputs should be probabilities for all classes. These probabilities help to construct the heuristic rule.



**Figure 2.** Flowchart of the proposed method. The above row shows the supervised learning for geometry information; the bottom row shows unsupervised learning for intensity information, and the final col joint the supervised and unsupervised classification results.

### 3.3. Supervised Learning for Geometry Information

#### 3.3.1. Geometry Feature Description

The feature in this study was the combination of a relative feature (i.e., FPFH ( $fpfh$ )), and two absolute features (i.e., normal ( $n$ ) and height ( $h$ )). Hence, the feature vector was ( $fpfh, n, h$ ), in which  $fpfh$  is a 33-dimensional vector, and  $n$  is a three-dimensional vector. FPFH was proposed in [46]. For each point  $q$  of the ALS data, all of  $q$ 's neighbors enclosed in the sphere with a given radius  $r$  are selected ( $r$ -neighborhood). Given every pair of points,  $q$  and  $q'$ , in the  $r$ -neighborhood, and assuming

$q$  is the point with a smaller angle between its associated normal and the line connecting the point (that is,  $n \cdot (q' - q) \geq n' \cdot (q - q')$ ), a Darboux  $uvw$  frame can be defined while using

$$\begin{cases} u = n \\ v = \frac{(q' - q)}{\|q' - q\|_2} \times u \\ w = u \times v \end{cases} \quad (1)$$

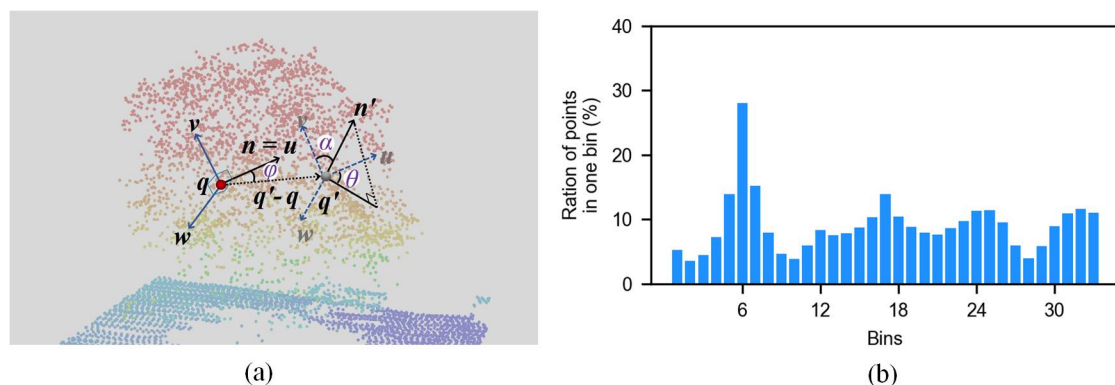
Subsequently, the normal  $n'$  can be represented as an angle tuple  $(\alpha, \theta, \varphi)$  in the  $uvw$  frame (as shown in Figure 3a):

$$\begin{cases} \alpha = v \cdot n' \\ \varphi = u \cdot \frac{(q' - q)}{\|q' - q\|_2} \\ \theta = \arctan(w \cdot n', u \cdot n) \end{cases} \quad (2)$$

A set of the angle tuples for  $q$  and each point in its  $r$ -neighborhood can be obtained. Thus,  $\alpha$  ( $\theta, \varphi$ ) in the angular tuples can be binned into a histogram 11 bins and then these three histograms are concatenated to form a simple point feature histogram (SPFH). FPFH (as shown in Figure 3b) is the weighted sum of SPFH in the  $r$ -neighborhood of  $q$ .

$$\text{FPFH}(q) = \text{SPFH}(q) + \frac{1}{|N(q)|} \sum_{q' \in N(q)} \frac{\text{SPFH}(q')}{\|q' - q\|_2} \quad (3)$$

where  $N(q)$  represents the neighborhood of  $q$  and  $|N(q)|$  represents the number of points in the  $r$ -neighborhood. More details about FPFH can be found in [55].



**Figure 3.** Angular representation (a) and an example of Fast point feature histograms (FPFH) (b).

FPFH describes the inertial property of the point distribution and is invariant to the transformation and rotation. However, the façade, roof, tree, and other geo-objects have their own direction, and the designed feature for the ALS data classification should represent this fact. Therefore, the point normal was locally estimated by fitting a plane using the neighborhood points [56]. On the other hand, the raw geometry information that was acquired by the ALS system is under the World Geodetic System 1984 coordinate system. It is necessary to obtain the height based on the ground. In this study, a progressive morphological filter that was accomplished in the point cloud library was used to extract the ground due to its good performance and high efficiency in urban environment [57].

### 3.3.2. A Brief Introduction to Supervised Learning Methods

In this study, four common supervised learning methods, decision tree (DT), random forest (RF), support vector classification (SVC), and extreme gradient boost (XGBoost), are used to process geometry feature. The following is a brief introduction to these methods.

## Decision Tree

Decision tree predicts the label of a target by learning simple decision rules that are represented using a tree structure. The nodes in DT are split using a best feature and a rational value. Gini index or entropy index is used to calculate the best feature and its divided value. The prediction of an unseen point  $\hat{L}_{DT}$  is the fraction of samples of the same class in a leaf that the unseen point falls into:

$$\hat{L}_{DT} = \frac{M}{\text{sum}(M)} \quad (4)$$

where  $M$  is a vector containing the count of samples per class. The predicted class  $\hat{l}_{DT}$  is the one with highest probability:

$$\hat{l}_{DT} = \underset{c \in \mathcal{L}}{\text{argmax}} \hat{L}_{DT}(c) \quad (5)$$

where  $\mathcal{L}$  is the label space.

## Support Vector Classification

For binary classification, the data can be separated by a hyperplane:

$$f(x) = \omega^T \phi(x) + b \quad (6)$$

where  $\omega$  and  $b$  are the parameters in a hyperplane,  $\phi(x)$  maps  $x$  into a higher-dimensional space. Many hyperplanes might classify the data. SVC aims to find the best hyperplane, so that the distance from it to the nearest data point on each side is maximized, which is written as:

$$\begin{aligned} \min_{\omega, b, \zeta} \quad & \frac{1}{2} \omega^T \omega + C \sum_{i=1}^n \zeta_i \\ \text{subject to} \quad & y_i (\omega^T \phi(x_i) + b) \geq 1 - \zeta_i, \\ & \zeta_i \geq 0, \quad i = 1, \dots, n \end{aligned} \quad (7)$$

where  $C > 0$  is regularization parameter,  $\zeta_i$  is slack variables. This object function can be solved using Lagrange multipliers and sequential minimal optimization [58], and  $\omega = \sum_{i=1}^n y_i \alpha_i \phi(x_i)^T$ . For an unseen point, the binary label can be derived while using:

$$\begin{aligned} \hat{l}_{SVC} &= \text{sgn} \left( \sum_{i=1}^n y_i \alpha_i \phi(x_i)^T \phi(x) + b \right) \\ &= \text{sgn} \left( \sum_{i=1}^n y_i \alpha_i \kappa(x_i, x) + b \right) \end{aligned} \quad (8)$$

where  $\kappa(\cdot, \cdot)$  is a kernel that is introduced because of the dimension of  $\phi(x)$  too high to calculate  $\phi(x_i)^T \phi(x)$ . Radial basis function (RBF) is a popular kernel in remote sensing data processing and it is also used in this study.

SVC can be easily adapted to a multi-class task by using the one-vs-one strategy. In addition, SVC does not provide the probability output. Scikit-learn uses Platt scaling to calibrate the probability of SVC in the binary case [59]. In the multiclass case, probabilities are extended as per [60].

## Random Forest

Random forest for classification is an ensemble of unpruned classification decision trees [61,62]. Each decision tree in random forest is built from a sample drawn with the bootstrap sample from all of the training data. When splitting a node during the construction of the decision tree, the split that is the best divided among a random subset of the features rather than among all features. The prediction of an unseen point can be the probability vector  $\hat{L}_{RF}$  of all classes or a single class  $\hat{l}_{RF}$ . The predicted

probability vector  $\hat{L}_{RF}$  of the unseen point is a vote by decision trees  $\hat{L}_{DT}$  in the RF, weighted by their probability estimates:

$$\hat{L}_{RF} = \frac{1}{n} \sum_{i=1}^n \hat{L}_{DTi} \quad (9)$$

where,  $\hat{L}_{DTi}$  is the probability vector of a decision tree output and  $n$  is the number of decision trees in the RF. The predicted class  $\hat{l}_{RF}$  is the one with highest probability:

$$\hat{l}_{RF} = \operatorname{argmax}_{c \in \mathcal{L}} \hat{L}_{RF}(c) \quad (10)$$

where  $\mathcal{L}$  is the label space.

### Extreme Gradient Boost

Extreme gradient boost is tree-boosting ensemble method, its fundamental function predicts a new classification membership after each iteration. XGBoost is built using an additive way that continuously makes new weak classifiers to improve the performance of the previous classifiers. Incorrectly classified samples receive higher weights at the next step, which forces the classifier to focus on their error in the following iterations [63,64]. The final classification of XGBoost combines the improvement of all the previous modeled trees:

$$\hat{L}_{XGB} = \operatorname{softmax} \left( \sum_{i=1}^n f_i(x) \right), f_i \in \mathcal{F} \quad (11)$$

where,  $f$  is a function in the functional space and  $\mathcal{F}$  is the set of all possible DTs. The predicted class  $\hat{l}_{XGB}$  is the one with highest probability:

$$\hat{l}_{XGB} = \operatorname{argmax}_{c \in \mathcal{L}} \hat{L}_{XGB}(c) \quad (12)$$

where  $\mathcal{L}$  is the label space.

### 3.4. Unsupervised Learning for Intensity Information

The GMM is one of the prototype-based clusterings that describe the dataset by a (usually small) set of prototypes [65]. The GMM uses the Gaussian probability model as a prototype to model the dataset and it assumes that the dataset is sampled from the Gaussian mixture distribution

$$p_M(x) = \sum_{k=1}^m \omega_k g(x | \mu_k, \Sigma_k) \quad (13)$$

where,  $x$  is a  $D$ -dimensional continuous-value data (i.e., intensity in this study);  $\omega_k$  are the mixture weights; and,  $m$  is the number of components. Each component is a Gaussian probability density function with mean vector  $\mu_k$  and covariance matrix  $\Sigma_k$

$$g(x | \mu_k, \Sigma_k) = \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma_k|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1} (x-\mu_k)} \quad (14)$$

The mixture weights  $\omega_k$  satisfy the constraint  $\sum_{k=1}^m \omega_k = 1$ .  $\omega_k$ ,  $\mu_k$ , and  $\Sigma_k$  are the three parameters that need to be estimated. The iterative Expectation-Maximization algorithm or the Maximum A Posteriori estimation is often used to estimate these parameters; more details regarding the estimation method and the implementation can be found in [66] and the Scikit-learn package.

Given each point with intensity  $x$ , the output  $\hat{L}'$  of GMM for cluster  $\lambda_k$  is

$$\hat{L}'(\lambda_k|x) = \frac{\omega_k g(x|\mu_k, \Sigma_k)}{\sum_{k=1}^m \omega_k g(x|\mu_k, \Sigma_k)} \quad (15)$$

and the cluster  $\hat{l}'$  is determined by

$$\hat{l}' = \underset{k \in \{1, 2, \dots, m\}}{\operatorname{argmax}} \hat{L}'(\lambda_k|x) \quad (16)$$

It is noted that: (1) the smoothing intensity before feeding to GMM made these parameters  $\omega_k$ ,  $\mu_k$ , and  $\Sigma_k$  more reliable; (2) the result of GMM is the cluster, which has no semantic information. The semantic information can be manually set according to the knowledge or automatically determined by the dominant value of the supervised learning result.

### 3.5. Joint Classification Results of the Supervised and Unsupervised Classifier

The outputs of the supervised and unsupervised learning have different meanings. The output of the supervised learning was the probability vector of both ground-level and elevated geo-objects, whereas the output probability vector of the unsupervised learning was only for ground-level geo-objects. Therefore, directly comparing the probability outputs of supervised and unsupervised learning cannot obtain a rational classification result, which is why this study used the ratios  $r_s$  and  $r_u$  to represent the confidence of supervised and unsupervised learning classification results.

Without any loss of generality, we assume there were two types of ground-level points,  $glo_1$  and  $glo_2$ .  $r_s$  and  $r_u$  were defined as Equations (17) and (18). In order to make  $r_s \geq 1, r_u \geq 1$ , we utilized max and min operation. If  $\hat{L}(glo_1) \geq \hat{L}(glo_2)$ , the probability that the supervised learning classifies an unseen point as  $glo_1$  was  $r_s$  times the probability of classifying the unseen point as  $glo_2$ .  $r_u$  had the same meaning. Thus, we can compare  $r_s$  and  $r_u$  to the joint classification results of the supervised and unsupervised classifier. In addition, the joint coefficient ( $a$  in Equation (19)) was introduced to represent the tradeoff between the supervised classification result and the unsupervised classification result.

$$r_s = \frac{\max(\hat{L}(glo_1), \hat{L}(glo_2))}{\min(\hat{L}(glo_1), \hat{L}(glo_2))} \quad (17)$$

$$r_u = \frac{\max(\hat{L}'(glo_1), \hat{L}'(glo_2))}{\min(\hat{L}'(glo_1), \hat{L}'(glo_2))} \quad (18)$$

$$y = \begin{cases} \hat{l} & ; ar_s \geq r_u \\ \hat{l}' & ; ar_s < r_u \end{cases} \quad (19)$$

A set of combinations of two types would be obtained if there were three or more types of ground-level points. For every combination, the above rule would be used, and then the selection would be determined by voting. The combined result is smoothed using the majority filter to obtain a more homogeneous classification result.

The proposed method was implemented by using python 3.6 in Jupyter Notebook. The code can be found from: [https://github.com/LidarLiu/ALS\\_Classification\\_Hierarchically\\_Geometry\\_Intensity](https://github.com/LidarLiu/ALS_Classification_Hierarchically_Geometry_Intensity).

## 4. Results

Here, precision and recall were used to measure the performance of the proposed method and they are shown in Figures 4 and 5. The precision and recall of the proposed method are higher than supervised learning method without intensity, especially for the low vegetation and the impervious surface, which can also be found in Figures 6 and 7. A comparison of the bottom row and the upper row in Figure 6 (or Figure 7) depicts that: 1) some misclassified low vegetation points were revised to the impervious surface (as shown in lower left inset); 2) some misclassified impervious surface points were



revised to low vegetation (as shown in upper right inset). These two revisions increased the precision and recall of the low vegetation and the impervious surface. For all of the selected supervised learning methods, the precision of low vegetation was greatly increased: DT increased by 4.47%, SVC increased by 7.22%, RF increased by 10.20%, XGBoost increased by 9.85% (as shown in Figure 4); and, the recall of impervious surface was significantly increased: DT increased by 8.97%, SVC increased by 8.2%, RF increased by 13.38%, and XGBoost increased by 12.56% (as shown in Figure 5). These improvements on ground-level points classification also improved the overall performance measured using the weighted average, which reveals that the proposed method can improve the urban ALS classification accuracy by using intensity.

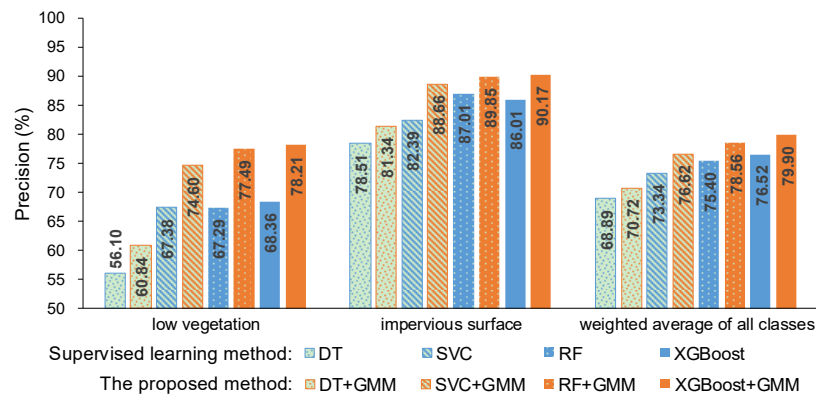


Figure 4. Precision of the supervised learning method without intensity and the proposed method.

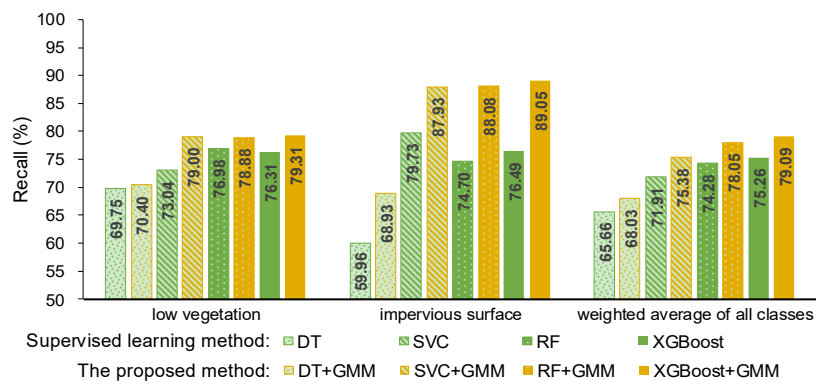
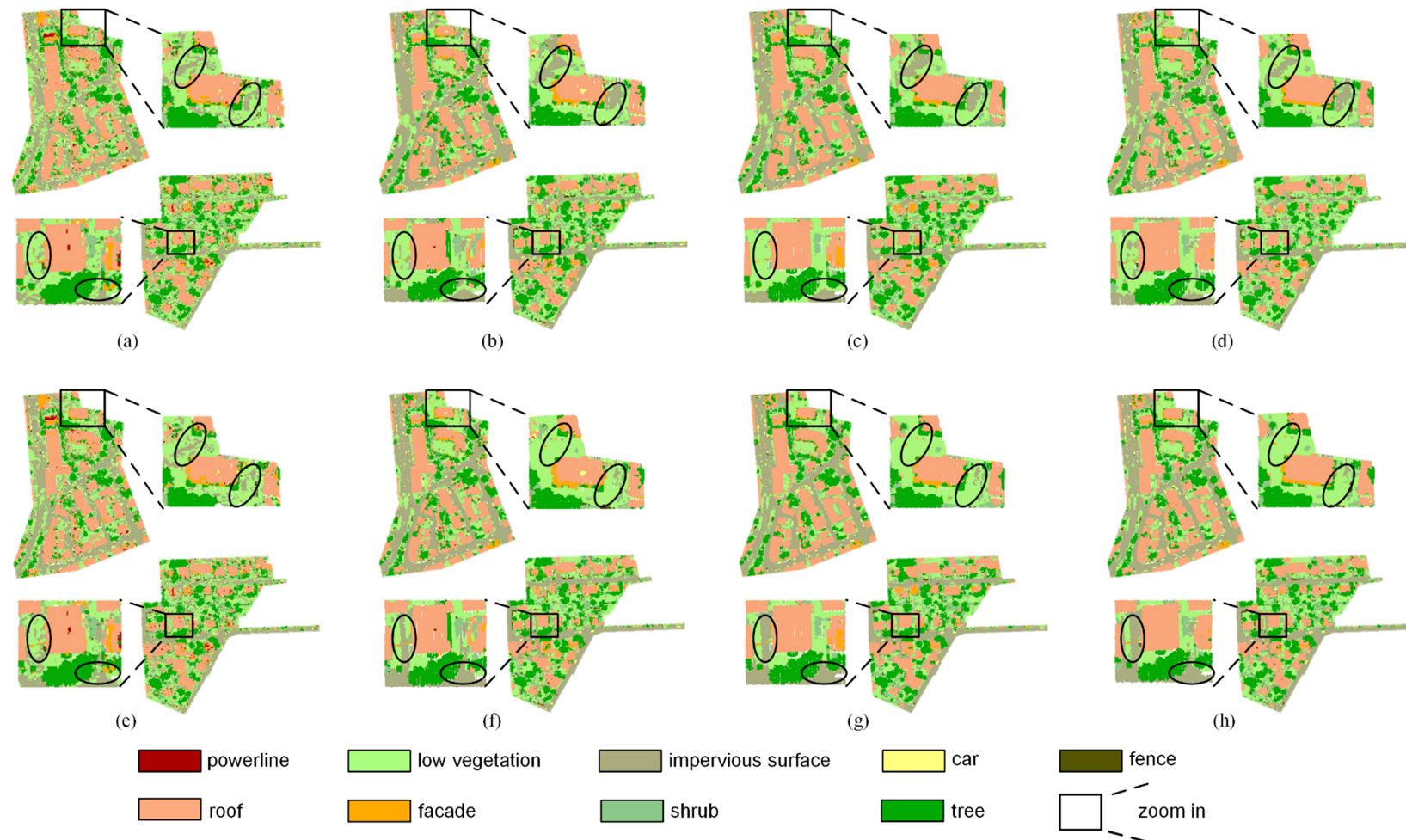
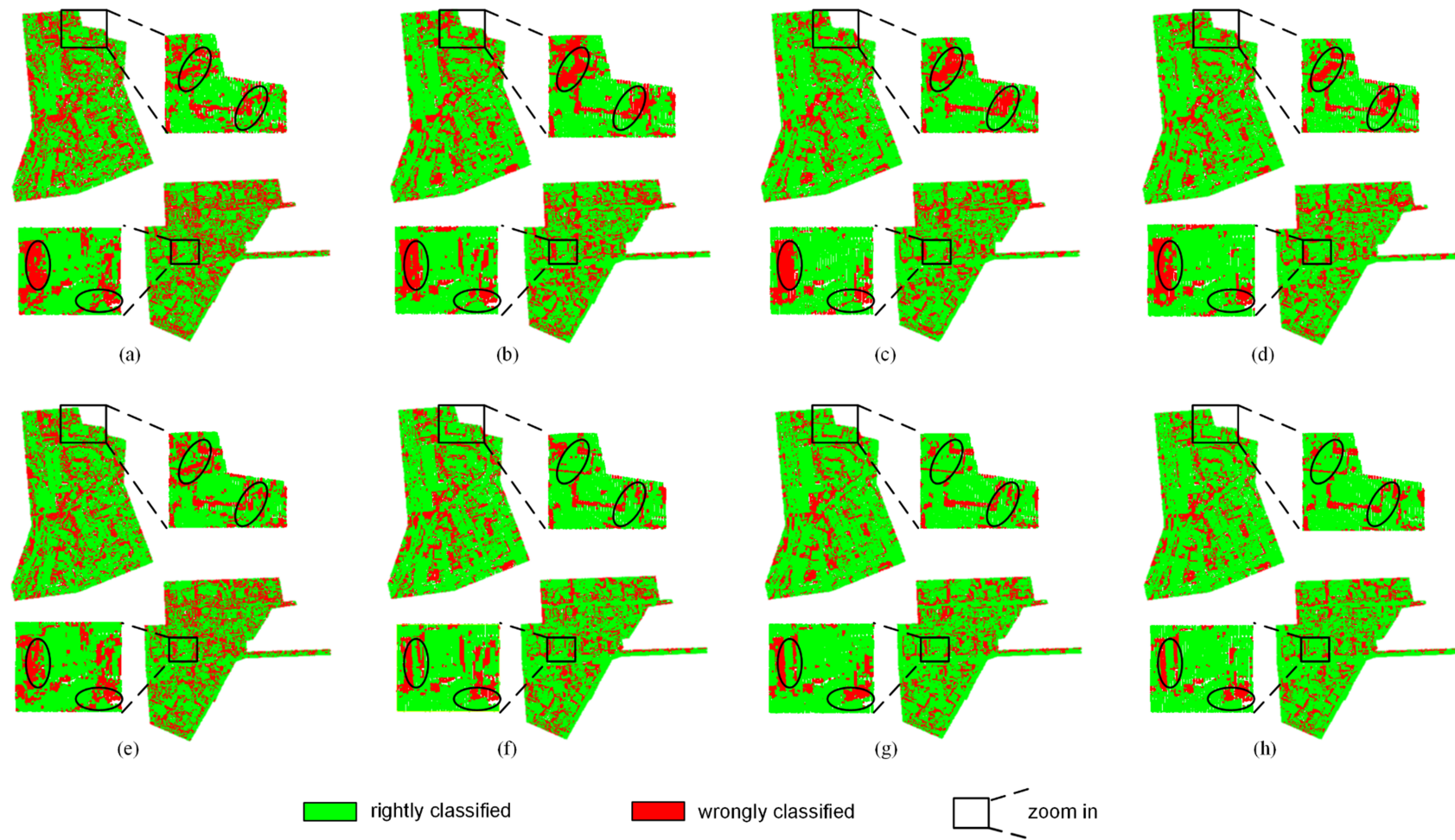


Figure 5. Recall of the supervised learning method without intensity and the proposed method.

Secondly, the fence, the car, and the façade were underestimated (as shown in Table 1), whereas four dominant geo-objects, namely the low vegetation, the impervious surface, the roof, and the tree, had a higher F1 score that is the harmonic mean of precision and recall, which reveals that the imbalance in the environment in the real world influences the classification. The confusion matrix (Table 2) and the visualization of the FPFH of all the geo-objects (Figure 8) were used to analyze the influence. Each geo-object has its own histogram, which helps the supervised learning method to distinguish different geo-objects; however, some geo-object's FPFH might be confused with others. The powerline is located above the roof and its point density is low, so the powerline points were typically misclassified as the roof. Car points and fence points were mainly misclassified as the shrub, whereas the shrub points misclassified as car or fence were rare, which might be because their height and FPFH are similar and the number of shrub points is much larger than car and fence points (as shown in Table 2). Some roof points and façade points were confused because the edge of the building (i.e., the interaction of roof and façade) is hard to distinguish.



**Figure 6.** Classification result of testing data. Upper row is the result of supervised learning methods without intensity: (a) decision tree (DT), (b) support vector classification (SVC), (c) random forest (RF), (d) extreme gradient boost (XGBoost). Bottom row is the result of the proposed method which joint the result of supervised and unsupervised learning methods: (e) DT + Gaussian mixture model (GMM), (f) SVC + GMM, (g) RF + GMM, and (h) XGBoost + GMM.



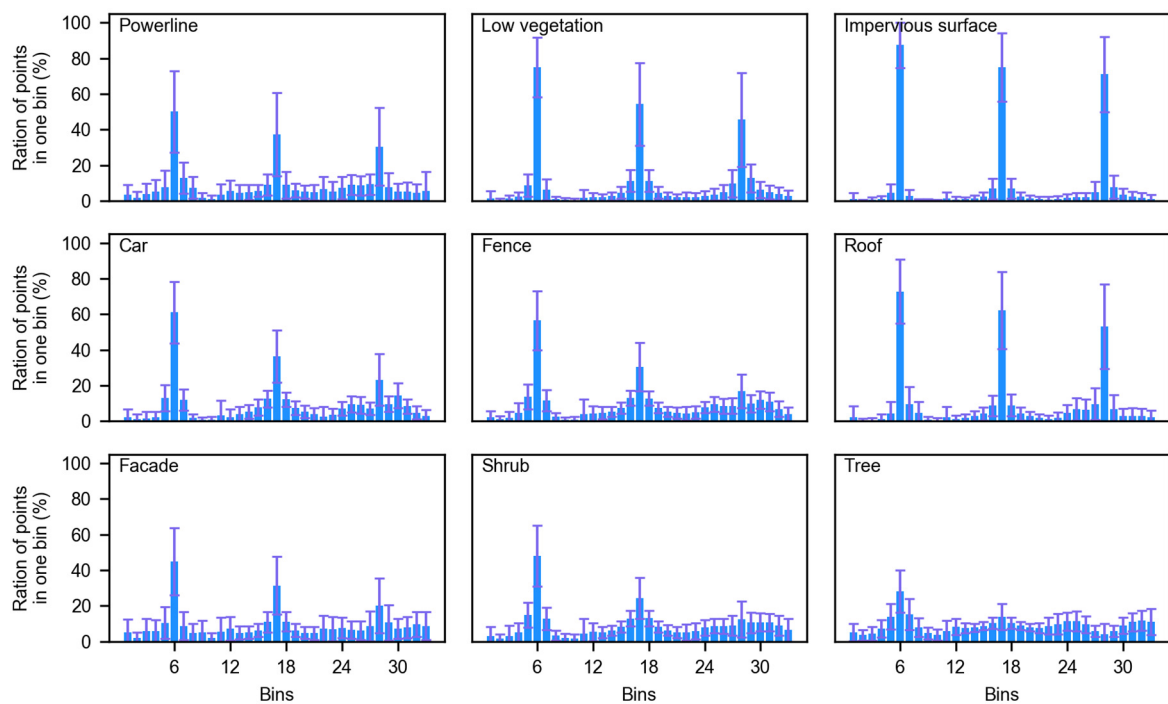
**Figure 7.** Classification error of testing data. Upper row is the classification error of supervised learning methods without intensity: (a) DT, (b) SVC, (c) RF, (d) XGBoost. Bottom row is the classification error of the proposed method: (e) DT + GMM, (f) SVC + GMM, (g) RF + GMM, and (h) XGBoost + GMM.

**Table 1.** F1 score of the proposed method (%).

Geo-objects→ Method↓	PL	LV	IS	Car	Fence	Roof	Façade	Shrub	Tree	WA
DT + GMM	1.83	65.27	74.62	33.26	14.43	81.84	28.67	34.76	73.23	68.89
SVM + GMM	5.46	76.74	88.29	40.97	19.19	83.15	35.12	34.81	72.16	75.57
RF + GMM	N/A	78.18	88.96	44.76	10.00	86.56	40.18	39.71	76.18	77.80
XGBoost + GMM	7.44	78.76	89.60	53.52	17.93	88.30	45.51	40.10	76.51	79.01

**Table 2.** Confusion matrix between the reference and the predict of the proposed method (RF+GMM as an example). UA is user accuracy and PA is producer accuracy.

Reference→ Predicted↓	PL	LV	IS	Car	Fence	Roof	Façade	Shrub	Tree	Total	UA
PL	0	0	0	0	0	0	0	0	0	0	-
LV	3	77,853	11,558	569	1442	1732	992	5396	929	100,474	77%
IS	0	8811	89,820	417	95	158	244	288	123	99,956	90%
Car	0	17	6	1125	24	13	15	81	38	1319	85%
Fence	0	85	8	342	434	117	5	203	63	1257	35%
Roof	383	3605	218	167	623	90,211	1453	1096	1637	99,393	91%
Façade	84	195	39	5	321	5700	4646	230	678	11,898	39%
Shrub	8	6430	231	1064	3298	2330	1208	10,821	4301	29,691	36%
Tree	122	1694	106	19	1185	8787	2661	6703	46,457	67,734	69%
Total	600	98,690	101,986	3708	7422	10,9048	11,224	24,818	54,226	411,722	
PA	0%	79%	88%	30%	6%	83%	41%	44%	86%		78%

**Figure 8.** Fast point feature histogram of all geo-objects included in the training data.

## 5. Discussion

### 5.1. Compare the Proposed Method with the Supervised Learning Method Considering Intensity

#### 5.1.1. Comparison Conditions Setting

One way of utilizing the intensity is to concatenate it with the geometry feature. This study compared the proposed method with the supervised learning method by using the feature ( $fpfh, n, h, i$ ), in which  $i$  is the intensity of a point. The intensity information was considered under certain conditions;



one condition was the same as that of the training data and another was different from the training data. The ISPRS benchmark testing data is acquired under the same conditions as the training data. The following method was used to simulate the intensity of the testing data under different conditions.

A simplified form of the ALS range equation can be obtained under the assumptions of an extended target (i.e., one that intercepts the entire laser beam) and Lambertian reflectance [31]:

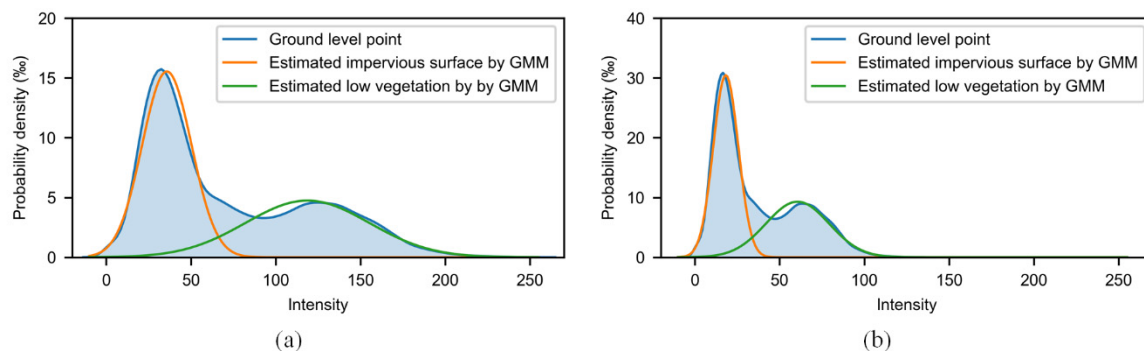
$$P_r = \frac{\cos \alpha}{4R^2} P_t D_r^2 \eta_{atm} \eta_{sys} \rho \quad (20)$$

where  $P_r$  is the received optical power (which is directly proportional to the intensity  $i$ ),  $\alpha$  is the angle of incidence,  $R$  is the range between the scanner and the target,  $P_t$  is the transmitted power,  $D_r$  is the received aperture diameter,  $\eta_{atm}$  is the atmospheric transmission factor,  $\eta_{sys}$  is the system transmission factor, and  $\rho$  is the target reflectance at the laser wavelength.

The ISPRS benchmark data was acquired at a height of 500 m. This study modified the intensity by simulating its acquisition above a height of 700 m, for which the effect of  $\alpha$  on the intensity is negligible when compared to the range. Thus, this study simply modified intensity by using

$$i_{modified} = \frac{R_{original}}{R_{modified}} \times i_{original} \approx \left( \frac{500 - h}{700 - h} \right)^2 \times i_{original} \quad (21)$$

Another reason to use this simplified equation was that the emission angle of the laser is not recorded in this data. Figure 9 shows the original and modified intensity distribution, in which the result of the GMM is also shown.



**Figure 9.** Intensity distribution of ground-level point in the supervised classification result, and GMM result for original intensity (a) and modified intensity (b).

The study has also introduced a consistency coefficient to compare the proposed method and supervised learning methods with intensity (i.e., DT\_i, SVC\_i, RF\_i, and XGBoost\_i). The consistency coefficient was defined as the ratio of the number of points for which the supervised learning method considering intensity and the proposed method have the same label to the total number of points.

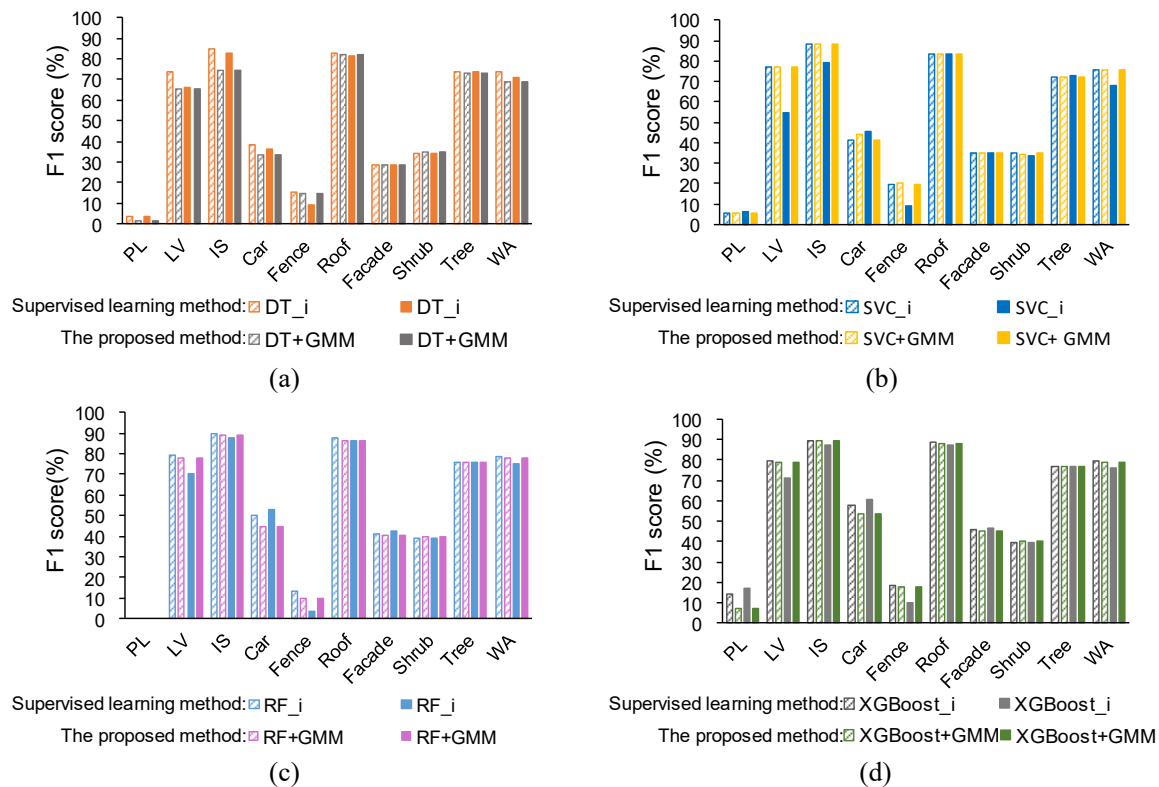
### 5.1.2. Comparison under the Same Conditions

The comparison result can be seen from Figure 10, where the metrics are the F1 score. The supervised learning methods considering intensity and the proposed method can both obtain better performances than the supervised learning methods without intensity and this improved performance was mainly reflected on the low vegetation and the impervious surface. There was no significant improvement on elevated geo-objects, such as roof, tree, and shrub. This improvement in ground-level geo-objects and the absence of elevated geo-objects indicated that the intensity information is more important for distinguishing ground-level geo-objects.

In addition to DT\_i, the performance of the proposed method is similar to the supervised learning method with intensity (as shown in Figure 10). The consistency coefficient between the proposed



method and SVC\_i, RF\_i, XGBoost\_i were 91.30%, 93.07%, 94.24%, respectively, which indicates that the classification result of the proposed method is consistent with supervised learning method with intensity. Although DT\_i obtained better performance than the proposed method, the performance of DT\_i did not exceed RF and XGBoost, which means that DT cannot make full use of the geometry feature and its classification ability is slightly worse. This phenomenon is also shown in Figure 7, where (a) has more red areas than (b), (c), and (d)



**Figure 10.** Comparison between supervised learning methods considering intensity and the proposed method under two conditions: same condition was filled with colored slashes, and different condition was filled with solid colors. WA is an abbreviation for weighted average of all classes. The supervised method in (a), (b), (c), (d) is DT, SVC, RF and XGboost, respectively.

### 5.1.3. Comparison under Different Conditions

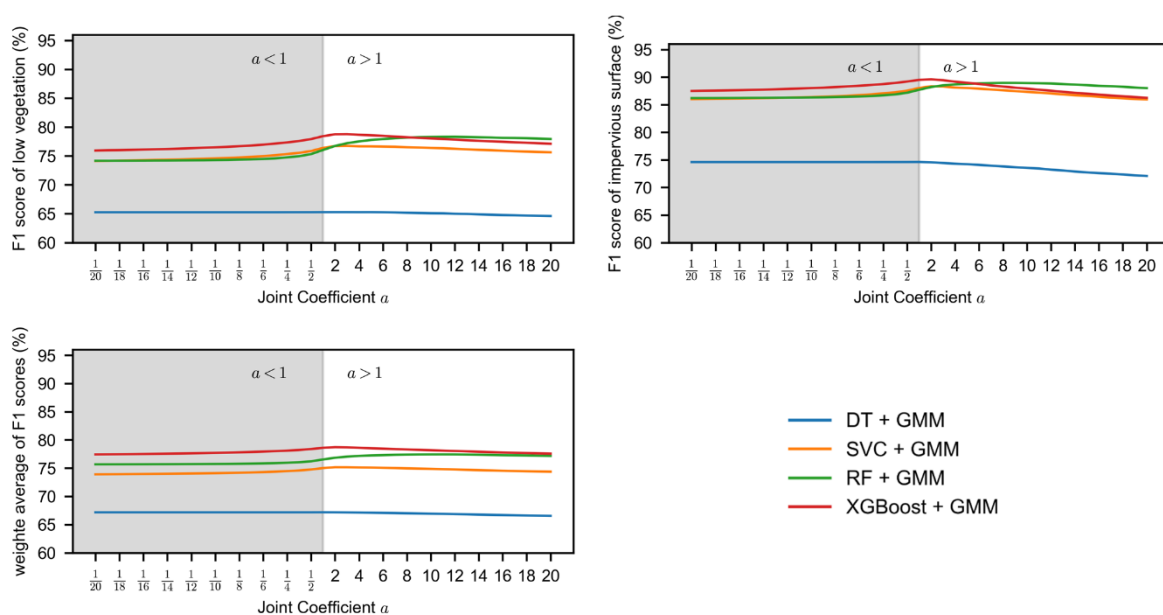
Figure 10 also compares the supervised learning methods while considering intensity and the proposed methods under two conditions. It can be concluded that supervised learning method with intensity is sensitive to intensity whereas the proposed method is more robust. For supervised learning method, after modifying the intensity, the F1 score of low vegetation was significantly reduced, which implies that supervised learning, which directly uses the intensity, has some problem. Apart from low vegetation, the impervious surface was slightly decreased because the average intensity of the impervious surface is lower than low vegetation (i.e., the mean of the orange line is smaller than the green line in Figure 9a). The modified intensity was decreased overall. Therefore, the modified intensity of low vegetation is more similar to the original intensity of the impervious surface, which results in a higher recall of the impervious surface and, correspondingly, a higher F1 score.

This study only controlled one factor (i.e., flight height), and some uncertainty appeared in the supervised learning method with intensity. The F1 score of the car significantly increased whereas the F1 score of the fence reduced. Therefore, the intensity of ALS should be carefully used instead of using it directly in supervised learning. In addition, the F1 score of the roof, the tree, the façade, and the

shrub were a little varied after modifying intensity, which also evaluated the basic of our method: geometry feature is enough for elevated geo-objects.

### 5.2. The Effect of Joint Parameter on the Performance of the Proposed Method

The joint coefficient,  $a$ , was the key parameter that was used to trade off the result of the supervised learning method and unsupervised learning method. If the joint result was more dependent on unsupervised classification (i.e.,  $a < 1$  in Figure 11), the performance will decrease. On the contrary, if the supervised classification was considered to be more reliable (i.e.,  $a > 1$  in Figure 11), the performance would be improved. However, if  $a$  was very large, it would ignore the result of the unsupervised classification, resulting in the performance decrease. In addition, the value of  $a$  is changed as the supervised learning methods changes: 1 for DT + GMM, 3 for SVC + GMM, 8 for RF + GMM, and 2 for XGBoost + GMM. Other important parameters can be found in the Appendix A.



**Figure 11.** Effect of joint coefficient  $a$  on the performance of the proposed method.

### 5.3. The Limitation of the Proposed Method

The classification of façades, fences, and shrubs is relatively poor because of two facts in the ALS data: (1) the geo-objects in the real world are imbalanced (that can be seen from Table 2); (2) the volume density of the fences ( $0.22 \text{ points} \cdot \text{m}^{-3}$ ), the façades ( $0.26 \text{ points} \cdot \text{m}^{-3}$ ), and the shrubs ( $1.29 \text{ points} \cdot \text{m}^{-3}$ ) are lower than the trees ( $2.26 \text{ points} \cdot \text{m}^{-3}$ ), the roofs ( $4.03 \text{ points} \cdot \text{m}^{-3}$ ), the low vegetations ( $4.05 \text{ points} \cdot \text{m}^{-3}$ ), and the impervious surfaces ( $5.87 \text{ points} \cdot \text{m}^{-3}$ ). In future work, the rebalance machine learning method should be considered and the optimal or multi-scale radius should be used in the ALS classification, as [43,48]. This study focuses on the intensity information of ALS, so the majority filter is used for small wrongly labeled regions. A graph-structured regularization framework will be considered for large wrongly labeled regions [28,34].

In addition, the type of land cover should be clearly defined to obtain a more generalized ALS classification model and transfer it to other ALS data. This study simply uses the benchmark label. Our further studies will rely on other more complex environments than those found in benchmark data.

## 6. Conclusions

This paper has proposed a novel method that uses the ALS information hierarchically to reasonably utilize fluctuating ALS intensity in a classification method and improve the generalization ability of the classifier. After applying a supervised classifier on the ALS data with the geometry feature,

the intensity information is used to reclassify ground-level points by an unsupervised classifier. The final classification result is the integration of the supervised and unsupervised classification results. The proposed method has higher classification accuracy than supervised learning without intensity: in the ISPRS benchmark data, a 5.68 % improvement for low vegetation and a 7.77% improvement for impervious surfaces were achieved.

Unlike supervised learning using intensity, the proposed method handled intensity through unsupervised learning, whose parameters could be re-estimated in the ALS data to be classified, which resulted in the classifier derived from the proposed method being adaptive to the fluctuating intensity and transferrable to other ALS data. Although this study did not prove that a supervised classifier trained both using geometry and intensity information is not useful in principle, it points to the conclusion that one should be cautious about directly using intensity in supervised learning for ALS data classification on a larger scale.

Furthermore, this study utilized the geometry information and intensity information of ALS hierarchically without calibrating intensity, which paves a novel way to use the ALS information and it can exhibit great potential for ALS data processing. However, the type of land cover should be clearly defined, and the imbalance in the real world should be considered in future work to obtain a more generalized classifier and apply it to a three-dimensional urban environment.

**Author Contributions:** Conceptualization, X.L., Y.C. and M.L.; Data curation, X.L. and Y.C.; Formal analysis, X.L. and Y.C.; Funding acquisition, Y.C., L.C. and M.L.; Investigation, Y.C. and L.C.; Methodology, X.L. and S.L.; Project administration, Y.C.; Resources, L.C.; Software, X.L.; Supervision, Y.C. and M.L.; Validation, S.L.; Visualization, X.L., Y.C. and S.L.; Writing – original draft, X.L. and S.L.; Writing – review & editing, X.L., Y.C., L.C. and M.L.

**Funding:** This work was supported by the National Key R&D Program of China (2017YFB0504205), and the National Natural Science Foundation of China (41501456, 41622109).

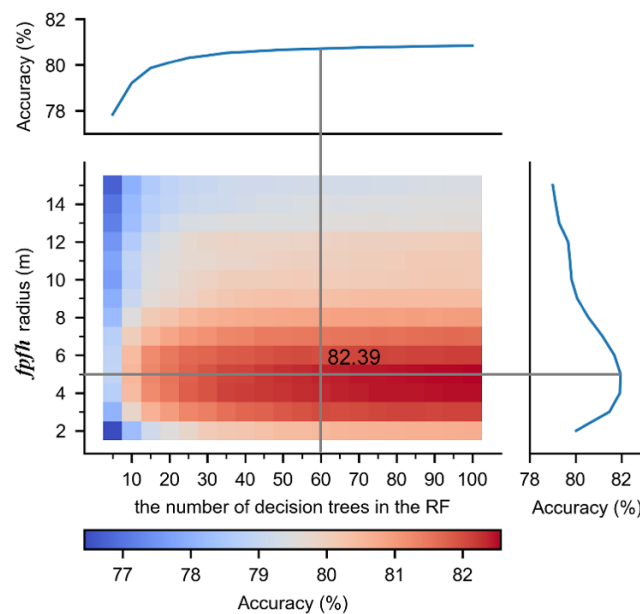
**Acknowledgments:** We thank German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) for providing The Vaihingen data set [54]: <http://www.ifp.uni-stuttgart.de/dgpf/DKEPAllg.html>.

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

## Appendix A.

### *Appendix A.1. The Parameters in RF and FPFH Radius*

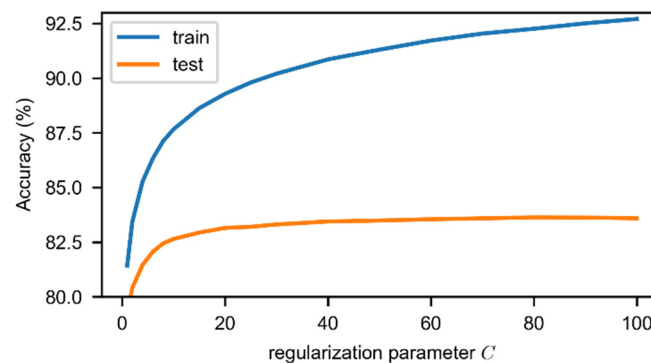
We used a 5-fold cross validation method to evaluate the effect of the FPFH radius and the number of decision trees in RF, and the result was shown in Figure A1. The accuracy of RF grows with the number of decision trees and when it exceeds 60, the accuracy of the classifier becomes more stable and cannot significantly improve. Hence, this study set the number of decision trees to 60. On the other hand, the overall accuracy of RF to the FPFH radius is more complex; it first increases and then decreases as the FPFH radius increases. This is caused by the fact that a larger radius contains more geo-objects, so the distinction between different geo-objects becomes smaller. On the other hand, a smaller radius contains fewer points resulting in the neighborhood not containing enough points to represent the geo-object. Hence, this study set the FPFH radius to 5 m.



**Figure A1.** Effect of FPFH radius and the number of decision trees on the accuracy of RF in the proposed method.

#### Appendix A.2. The Parameter in SVC

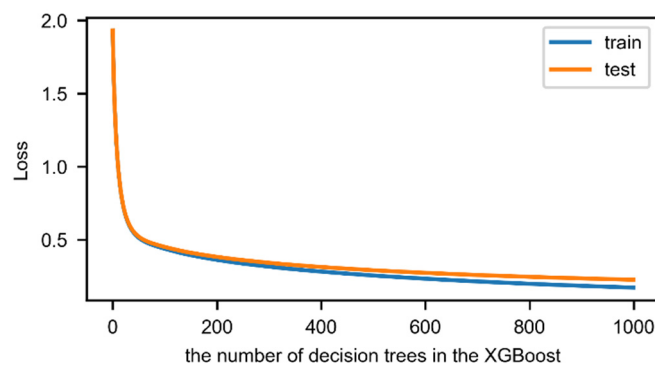
We used a 5-fold cross validation method to evaluate the effect of regularization parameter  $C$  in the SVC, and the result was shown in Figure A2. The accuracy of SVC grows with the  $C$  and when it exceeds 40, the test accuracy of the classifier becomes more stable and cannot significantly improve. Hence, this study set the regularization parameter to 40.



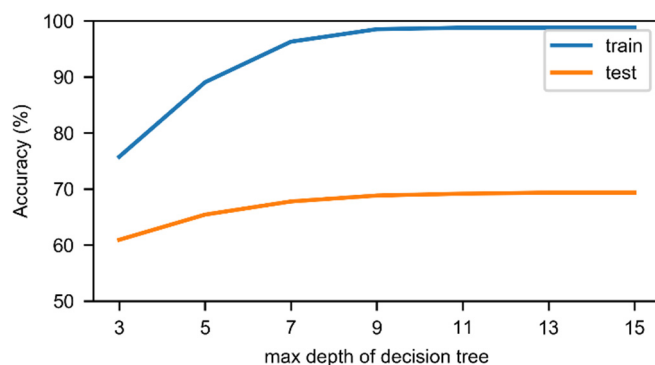
**Figure A2.** Effect of regularization parameter  $C$  on the accuracy of SVC in the proposed method.

#### Appendix A.3. The Parameters in XGBoost

We used a 5-fold cross validation method to evaluate the effect of the number of decision trees and its max depth in the XGBoost, and the result was shown in Figures A3 and A4. The loss of XGBoost decrease with the number of decision trees and when it exceeds 500, the loss of the classifier becomes more stable and cannot significantly reduce. Hence, this study set the number of decision trees to 500. Accordingly, the max depth of decision trees is set to 11.



**Figure A3.** Effect of the number of decision trees on the loss of XGBoost in the proposed method.



**Figure A4.** Effect of max depth of decision trees on the accuracy of XGBoost in the proposed method.

## References

1. Nations, U. *World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420)*; United Nations, Department of Economic and Social Affairs: New York, NY, USA, 2018.
2. Walde, I.; Hese, S.; Berger, C.; Schmullius, C. From land cover-graphs to urban structure types. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 584–609. [\[CrossRef\]](#)
3. Zhu, Z.; Zhou, Y.; Seto, K.C.; Stokes, E.C.; Deng, C.; Pickett, S.T.A.; Taubenböck, H. Understanding an urbanizing planet: Strategic directions for remote sensing. *Remote Sens. Environ.* **2019**, *228*, 164–182. [\[CrossRef\]](#)
4. Eitel, J.U.H.; Höfle, B.; Vierling, L.A.; Abellán, A.; Asner, G.P.; Deems, J.S.; Glennie, C.L.; Joerg, P.C.; LeWinter, A.L.; Magney, T.S.; et al. Beyond 3-D: The new spectrum of lidar applications for earth and ecological sciences. *Remote Sens. Environ.* **2016**, *186*, 372–392. [\[CrossRef\]](#)
5. Yan, W.Y.; Shaker, A.; El-Ashrawy, N. Urban land cover classification using airborne LiDAR data: A review. *Remote Sens. Environ.* **2015**, *158*, 295–310. [\[CrossRef\]](#)
6. Tran, T.; Ressel, C.; Pfeifer, N. Integrated change detection and classification in urban areas based on airborne laser scanning point clouds. *Sensors* **2018**, *18*, 448. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Okyay, U.; Telling, J.; Glennie, C.L.; Dietrich, W.E. Airborne lidar change detection: An overview of Earth sciences applications. *Earth Sci. Rev.* **2019**. [\[CrossRef\]](#)
8. Godwin, C.; Chen, G.; Singh, K.K. The impact of urban residential development patterns on forest carbon density: An integration of LiDAR, aerial photography and field mensuration. *Landsc. Urban Plan.* **2015**, *136*, 97–109. [\[CrossRef\]](#)
9. Raciti, S.M.; Hutyrá, L.R.; Newell, J.D. Mapping carbon storage in urban trees with multi-source remote sensing data: Relationships between biomass, land use, and demographics in Boston neighborhoods. *Sci. Total Environ.* **2014**, *500*, 72–83. [\[CrossRef\]](#)
10. Niemeyer, J.; Rottensteiner, F.; Soergel, U. Contextual classification of lidar data and building object detection in urban areas. *ISPRS J. Photogramm. Remote Sens.* **2014**, *87*, 152–165. [\[CrossRef\]](#)



11. Kang, Z.; Yang, J. A probabilistic graphical model for the classification of mobile LiDAR point clouds. *ISPRS J. Photogramm. Remote Sens.* **2018**, *143*, 108–123. [[CrossRef](#)]
12. Sithole, G.; Vosselman, G. Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. *ISPRS J. Photogramm. Remote Sens.* **2004**, *59*, 85–101. [[CrossRef](#)]
13. Liu, X.; Chen, Y.; Cheng, L.; Yao, M.; Deng, S.; Li, M.; Cai, D. Airborne laser scanning point clouds filtering method based on the construction of virtual ground seed points. *J. Appl. Remote Sens.* **2017**, *11*, 16032. [[CrossRef](#)]
14. Awrangjeb, M.; Fraser, C. Automatic segmentation of raw LiDAR data for extraction of building roofs. *Remote Sens.* **2014**, *6*, 3716–3751. [[CrossRef](#)]
15. Sun, S.; Salvaggio, C. Aerial 3D building detection and modeling from airborne LiDAR point clouds. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 1440–1449. [[CrossRef](#)]
16. Jwa, Y.; Sohn, G.; Kim, H.B. Automatic 3d powerline reconstruction using airborne lidar data. *Int. Arch. Photogramm. Remote Sens.* **2009**, *38*, W8.
17. Sohn, G.; Jwa, Y.; Kim, H.B. Automatic powerline scene classification and reconstruction using airborne lidar data. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2012**, *13*, 28. [[CrossRef](#)]
18. Xu, Z.; Guan, K.; Casler, N.; Peng, B.; Wang, S. A 3D convolutional neural network method for land cover classification using LiDAR and multi-temporal Landsat imagery. *ISPRS J. Photogramm. Remote Sens.* **2018**, *144*, 423–434. [[CrossRef](#)]
19. Cook, B.D.; Bolstad, P.V.; Næsset, E.; Anderson, R.S.; Garrigues, S.; Morisette, J.T.; Nickeson, J.; Davis, K.J. Using LiDAR and quickbird data to model plant production and quantify uncertainties associated with wetland detection and land cover generalizations. *Remote Sens. Environ.* **2009**, *113*, 2366–2379. [[CrossRef](#)]
20. Buján, S.; González-Ferreiro, E.; Reyes-Bueno, F.; Barreiro-Fernández, L.; Crecente, R.; Miranda, D. Land Use Classification from Lidar Data and Ortho-Images in a Rural Area. *Photogramm. Rec.* **2012**, *27*, 401–422. [[CrossRef](#)]
21. Cao, Y.; Wei, H.; Zhao, H.; Li, N. An effective approach for land-cover classification from airborne lidar fused with co-registered data. *Int. J. Remote Sens.* **2012**, *33*, 5927–5953. [[CrossRef](#)]
22. Murphy, J.M.; Le Moigne, J.; Harding, D.J. Automatic image registration of multimodal remotely sensed data with global shearlet features. *IEEE Trans. Geosci. Remote Sens.* **2015**, *54*, 1685–1704. [[CrossRef](#)] [[PubMed](#)]
23. Mastin, A.; Kepner, J.; Fisher, J. Automatic registration of LIDAR and optical images of urban scenes. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009.
24. Yang, B.; Chen, C. Automatic registration of UAV-borne sequent images and LiDAR data. *ISPRS J. Photogramm. Remote Sens.* **2015**, *101*, 262–274. [[CrossRef](#)]
25. Song, J.-H.; Han, S.-H.; Yu, K.Y.; Kim, Y.-I. Assessing the possibility of land-cover classification using lidar intensity data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2002**, *34*, 259–262.
26. Charaniya, A.P.; Manduchi, R.; Lodha, S.K. Supervised Parametric Classification of Aerial LiDAR Data. In Proceedings of the Computer Vision and Pattern Recognition Workshop, 2004. CVPRW'04, Washington, DC, USA, 27 June–2 July 2004; p. 30.
27. Matikainen, L.; Karila, K.; Hyypä, J.; Litkey, P.; Puttonen, E.; Ahokas, E. Object-based analysis of multispectral airborne laser scanner data for land cover classification and map updating. *ISPRS J. Photogramm. Remote Sens.* **2017**, *128*, 298–313. [[CrossRef](#)]
28. Li, N.; Liu, C.; Pfeifer, N. Improving LiDAR classification accuracy by contextual label smoothing in post-processing. *ISPRS J. Photogramm. Remote Sens.* **2019**, *148*, 13–31. [[CrossRef](#)]
29. Yang, Z.; Jiang, W.; Xu, B.; Zhu, Q.; Jiang, S.; Huang, W. A convolutional neural network-based 3D semantic labeling method for ALS point clouds. *Remote Sens.* **2017**, *9*, 936. [[CrossRef](#)]
30. Niemeyer, J.; Rottensteiner, F.; Sörgel, U.; Heipke, C. Hierarchical higher order crf for the classification of airborne lidar point clouds in urban areas. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. Arch.* **2016**, *41*, 655–662. [[CrossRef](#)]
31. Kashani, A.; Olsen, M.; Parrish, C.; Wilson, N. A review of LiDAR radiometric processing: From ad hoc intensity correction to rigorous radiometric calibration. *Sensors* **2015**, *15*, 28099–28128. [[CrossRef](#)]
32. Yu, Y.; Li, J.; Guan, H.; Jia, F.; Wang, C. Learning hierarchical features for automated extraction of road markings from 3-D mobile LiDAR point clouds. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *8*, 709–726. [[CrossRef](#)]

33. Hu, X.; Li, Y.; Shan, J.; Zhang, J.; Zhang, Y. Road Centerline Extraction in Complex Urban Scenes From LiDAR Data Based on Multiple Features. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 7448–7456.
34. Landrieu, L.; Raguét, H.; Vallet, B.; Mallet, C.; Weinmann, M. A structured regularization framework for spatially smoothing semantic labelings of 3D point clouds. *ISPRS J. Photogramm. Remote Sens.* **2017**, *132*, 102–118. [[CrossRef](#)]
35. Vosselman, G.; Coenen, M.; Rottensteiner, F. Contextual segment-based classification of airborne laser scanner data. *ISPRS J. Photogramm. Remote Sens.* **2017**, *128*, 354–371. [[CrossRef](#)]
36. Schindler, K. An overview and comparison of smooth labeling methods for land-cover classification. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 4534–4545. [[CrossRef](#)]
37. Thomas, H.; Goulette, F.; Deschaud, J.-E.; Marcotegui, B. Semantic Classification of 3D Point Clouds with Multiscale Spherical Neighborhoods. In Proceedings of the 2018 International Conference on 3D Vision (3DV), Verona, Italy, 5–8 September 2018; pp. 390–398.
38. Mallet, C.; Bretar, F.; Roux, M.; Soergel, U.; Heipke, C. Relevance assessment of full-waveform lidar data for urban area classification. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, S71–S84. [[CrossRef](#)]
39. Secord, J.; Zakhor, A. Tree detection in urban regions using aerial lidar and image data. *IEEE Geosci. Remote Sens. Lett.* **2007**, *4*, 196–200. [[CrossRef](#)]
40. Lodha, S.K.; Fitzpatrick, D.M.; Helmbold, D.P. Aerial lidar data classification using adaboost. In Proceedings of the Sixth International Conference on 3-D Digital Imaging and Modeling (3DIM 2007), Montreal, QC, Canada, 21–23 August 2007; pp. 435–442.
41. Guo, B.; Huang, X.; Zhang, F.; Sohn, G. Classification of airborne laser scanning data using JointBoost. *ISPRS J. Photogramm. Remote Sens.* **2015**, *100*, 71–83. [[CrossRef](#)]
42. Guo, L.; Chehata, N.; Mallet, C.; Boukir, S. Relevance of airborne lidar and multispectral image data for urban scene classification using Random Forests. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 56–66. [[CrossRef](#)]
43. Weinmann, M.; Urban, S.; Hinz, S.; Jutzi, B.; Mallet, C. Distinctive 2D and 3D features for automated large-scale scene analysis in urban areas. *Comput. Graph.* **2015**, *49*, 47–57. [[CrossRef](#)]
44. Dittrich, A.; Weinmann, M.; Hinz, S. Analytical and numerical investigations on the accuracy and robustness of geometric features extracted from 3D point cloud data. *ISPRS J. Photogramm. Remote Sens.* **2017**, *126*, 195–208. [[CrossRef](#)]
45. Osada, R.; Funkhouser, T.; Chazelle, B.; Dobkin, D. Shape distributions. *ACM Trans. Graph.* **2002**, *21*, 807–832. [[CrossRef](#)]
46. Rusu, R.B.; Blodow, N.; Beetz, M. Fast point feature histograms (FPFH) for 3D registration. In Proceedings of the 2009 IEEE International Conference on Robotics and Automation, Kobe, Japan, 12–17 May 2009; pp. 3212–3217.
47. Rusu, R.B.; Marton, Z.C.; Blodow, N.; Beetz, M. Persistent point feature histograms for 3D point clouds. In Proceedings of the 10th International Conference Intelligent Autonomous Systems (IAS-10), Baden-Baden, Germany, 23–25 July 2008; pp. 119–128.
48. Blomley, R.; Weinmann, M.; Leitloff, J.; Jutzi, B. Shape distribution features for point cloud analysis—a geometric histogram approach on multiple scales. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2014**, *2*, 9. [[CrossRef](#)]
49. Zhu, X.X.; Tuia, D.; Mou, L.; Xia, G.-S.; Zhang, L.; Xu, F.; Fraundorfer, F. Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 8–36. [[CrossRef](#)]
50. Reichstein, M.; Camps-Valls, G.; Stevens, B.; Jung, M.; Denzler, J.; Carvalhais, N. Deep learning and process understanding for data-driven Earth system science. *Nature* **2019**, *566*, 195. [[CrossRef](#)] [[PubMed](#)]
51. Zhao, R.; Pang, M.; Wang, J. Classifying airborne LiDAR point clouds via deep features learned by a multi-scale convolutional neural network. *Int. J. Geogr. Inf. Sci.* **2018**, *32*, 960–979. [[CrossRef](#)]
52. Yan, W.Y.; Shaker, A. Radiometric correction and normalization of airborne LiDAR intensity data for improving land-cover classification. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 7658–7673.
53. Yan, W.Y.; Shaker, A.; Habib, A.; Kersting, A.P. Improving classification accuracy of airborne LiDAR intensity data by geometric calibration and radiometric correction. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 35–44. [[CrossRef](#)]
54. Cramer, M. The DGPF-test on digital airborne camera evaluation—Overview and test design. *Photogramm. Fernerkund. Geoinf.* **2010**, *2010*, 73–82. [[CrossRef](#)]

55. fpfh\_estimation @ pointclouds.org. Available online: [http://pointclouds.org/documentation/tutorials/fpfh\\_estimation.php](http://pointclouds.org/documentation/tutorials/fpfh_estimation.php) (accessed on 23 August 2019).
56. Girardeau-Montaut, D. CloudCompare Version 2.6. 1-User Manual. 2015. Available online: <http://www.danielgm.net/cc/doc/qCC/CloudCompare%20v2.6.1%20%20User%20manual.pdf> (accessed on 20 August 2019).
57. Zhang, K.; Chen, S.-C.; Whitman, D.; Shyu, M.-L.; Yan, J.; Zhang, C. A progressive morphological filter for removing nonground measurements from airborne LIDAR data. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 872–882. [[CrossRef](#)]
58. Chang, C.-C.; Lin, C.-J. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 27. [[CrossRef](#)]
59. svm @ scikit-learn.org. Available online: <https://scikit-learn.org/stable/modules/svm.html#svm-classification> (accessed on 23 August 2019).
60. Wu, T.-F.; Lin, C.-J.; Weng, R.C. Probability estimates for multi-class classification by pairwise coupling. *J. Mach. Learn. Res.* **2004**, *5*, 975–1005.
61. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
62. Chehata, N.; Guo, L.; Mallet, C. Airborne lidar feature selection for urban classification using random forests. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2009**, *38*, W8.
63. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
64. Georganos, S.; Grippa, T.; Vanhuyse, S.; Lennert, M.; Shimoni, M.; Wolff, E. Very high resolution object-based land use-land cover urban classification using extreme gradient boosting. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 607–611. [[CrossRef](#)]
65. Borgelt, C. *Prototype-Based Classification and Clustering*. Magdeburg, Germany. 2006. Available online: <https://opendata.uni-halle.de/bitstream/1981185920/10725/1/chrborgelt1.pdf> (accessed on 10 October 2019).
66. Reynolds, D. Gaussian mixture models. *Encycl. Biom.* **2015**, 827–832. [[CrossRef](#)]



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).