

Article

# Large-Scale ALS Data Semantic Classification Integrating Location-Context-Semantics Cues by Higher-Order CRF

## Wei Han<sup>1</sup>, Ruisheng Wang<sup>2</sup>, Daqing Huang<sup>1,\*</sup> and Cheng Xu<sup>1</sup>

- <sup>1</sup> College of Electronic and Information Engineering, No. 29 Yudao Road, Nanjing University of Aeronautics & Astronautics, Nanjing 210016, China; nuaa\_han@hotmail.com (W.H.); stereom@sina.cn (C.X.)
- <sup>2</sup> Department of Geomatics Engineering, University of Calgary, 2500 University Drive NW,
- Calgary, AB T2N 1N4, Canada; ruiswang@ucalgary.ca \* Correspondence: mxyacc@nuaa.edu.cn

Received: 9 February 2020; Accepted: 17 March 2020; Published: 18 March 2020



Abstract: We designed a location-context-semantics-based conditional random field (LCS-CRF) framework for the semantic classification of airborne laser scanning (ALS) point clouds. For ALS datasets of high spatial resolution but with severe noise pollutions, more contexture and semantics cues, besides location information, can be exploited to surmount the decrease of discrimination of features for classification. This paper mainly focuses on the semantic classification of ALS data using mixed location-context-semantics cues, which are integrated into a higher-order CRF framework by modeling the probabilistic potentials. The location cues modeled by the unary potentials can provide basic information for discriminating the various classes. The pairwise potentials consider the spatial contextual information by establishing the neighboring interactions between points to favor spatial smoothing. The semantics cues are explicitly encoded in the higher-order potentials. The higher-order potential operates at the clusters level with similar geometric and radiometric properties, guaranteeing the classification accuracy based on semantic rules. To demonstrate the performance of our approach, two standard benchmark datasets were utilized. Experiments show that our method achieves superior classification results with an overall accuracy of 83.1% on the Vaihingen Dataset and an overall accuracy of 94.3% on the Graphics and Media Lab (GML) Dataset A compared with other classification algorithms in the literature.

**Keywords:** airborne laser scanning; semantic classification; feature extraction; conditional random fields; higher-order potentials

## 1. Introduction

The semantic classification has been, and still is, of significant interest to the Light Detection and Ranging (LiDAR) processing and machine learning. Airborne laser scanning (ALS) system can acquire both geometric and radiometric information of geo-objects, which has been widely used in semantic classification [1]. An increasing number of applications require the result of semantic classification ranging from object detection to automatic three-dimensional (3D) modeling. Automated urban object extraction from remotely sensed data, especially from ALS point clouds, is a very challenging task due to the complex urban environments and the unorganized point clouds data. We also consider finding different types of objects in a small local neighborhood in this paper, which is obviously difficult for reliable extractions. Compared with a binary decision process, each 3D point in the irregularly distributed point clouds is assigned with a semantic object label in this work. However, due to the obvious defects of ALS point clouds (e.g., noise, inhomogeneity, loss of sharp features and outliers), current methods are not resilient for clutter scenes and heterogeneous ALS point cloud data obtained



from different equipment. Therefore, we incorporate location, spatial contextual, and semantics cues within a higher-order conditional random field (CRF) framework to provide complementary information from varying perspectives, so that it can address the common misjudgment of semantic classes in ALS point clouds, from the perspectives of the accuracy of each class and the overall accuracy.

## 1.1. Related Works for ALS Point Cloud Classification

According to the type of entity used for classification, existing methods can be categorized as point-based and cluster-based (or segment-based) [2,3]. Point-based methods classify each point of the ALS data by using features as the inputs for supervised or unsupervised classifiers [4], while cluster-based methods segment the ALS data into clusters, then class labels are assigned to the clusters in which all points share the same class label [5,6]. We briefly review the aforementioned methods, and demonstrate the rationale for our method in what follows.

Point-based methods generally extract point-wise features locally from the neighborhood defined by a sphere or cylinder. Therefore, such methods usually focus on the selection of discriminative features and effective classifiers. For instance, Reference [7] worked on 3D scene analysis, including geometric features extraction and optimal neighbors selection. Then an optimal eigenentropy-based scale selection method was proposed. Reference [2] combined airborne LiDAR with images to extract more discriminative features. Then, based on these features, the points were classified with AdaBoost. Spectral information from full waveform ALS point clouds can also be used for feature extraction, which exhibit promising results in large-scale urban environments [8]. To take advantage of contextual information, Reference [9] integrated a random forest (RF) classifier into a CRF framework, where the unary and pairwise potentials of CRF were based on probabilities computed by RF. With multispectral images and 3D geometry data, fully connected conditional random field (FCCRF) graph model was employed [10]. Although these point-based methods benefit from using contextual information, their effects have been very limited because they merely consider the coherences between points within a small neighborhood or employ the FCCRF model with great computational cost. Moreover, point-based features are sensitive to noise, such that these methods may not be suitable for complicated scenes and noisy ALS cloud points.

Cluster-based methods choose a point segment in which points share homogeneous properties as the entity to be classified. A variety of methods has been used, which can be roughly classified as exclusive methods, overlapping methods, hierarchical methods and probabilistic methods [11,12]. It has also been proposed to use different entities in the form of voxels, blocks and pillars [13], in the form of planes, smooth surfaces and rough surfaces [6], or in the form of spatial bins, planar segments and local neighborhoods [14]. A robust unsupervised clustering algorithm P2C based on a hierarchical analysis was proposed by Reference [15], which comprises two stages: segmenting the point cloud into non-overlapping patches, and merging the patches into surfaces according to their probability density functions (PDFs). Reference [16] proposed a probability density clustering algorithm to perform hierarchical clustering and a higher-order CRF model was employed based on two different clusters. A natural exponential function was used to obtain hierarchical clusters of ALS point cloud [17], then multilevel point cluster-based features were extracted with latent Dirichlet allocation and sparse coding. Point-homogeneity constraint clustered points with similar geometric and radiometric properties in Reference [18], and markov random field (MRF) model was built based on these clusters. Recently, [19] introduced a super-point graph (SPG) to segment large-scale point clouds. Cluster-based methods uses more semantic features, like cluster size and shape, and therefore obtains a smoother classification result than the point-based method. However, the performance of cluster-based method is negatively affected by under/over-segmentation errors and loss of information. The semantic classification results are strongly connected to the accuracy of cluster-based method.

Specifically, the launch of deep learning in point cloud classification causes significantly promising classification results. PointNet [20] and following works Point-Net++ [21], Recurrent Slice Network (RSNet) [22], Dynamic Graph Convolutional Neural Networks (DGCNN) [23], and Point Convolutional

Neural Networks (PointCNN) [24] further focus on exploring the local context and hierarchical learning architectures. [1] transformed the 3D neighborhood features of a point into 2D images, which were treated as the input of a multi-scale convolutional neural networks (CNN) for training and testing tasks. However, most of the above methods are used in indoor point clouds. Due to the attributes of ALS point cloud, which is generally coarse, noisy, sparse and heterogeneous, the process of model training may be time-consuming and the model may not be suitable for ALS point clouds from different equipment. Then, the present study describes a combination of point-base and cluster-based method in a CRF framework, which takes full advantages of location, context and semantics cues and shows high plausibility in ALS point cloud semantic classification compared with others.

#### 1.2. Contribution

In this paper, an efficient algorithm for the ALS point cloud semantic classification is proposed to improve the overall accuracy and accuracy of each class, especially for classes with small size, which is rarely studied in other papers. Higher-order potentials of the location-context-semantics-based conditional random field (LCS-CRF) play an important role in semantic classification. To increase the efficiency, the Cloth Simulation Filter (CSF) is analyzed and further used to remove ground points with RANdom SAmple Consensus (RANSAC) algorithm for the cluster-based features extraction. Then, a constrained mean-shift clustering method is proposed to obtain clusters which are used to define semantic rules. The higher-order potential and unary potential can be fused based on the class membership probabilities for the inference of LCS-CRF algorithm. The efficiency of the proposed LCS-CRF method is confirmed with two ALS datasets. Compared to the other semantic classification methods, the experimental results confirm that the LCS-CRF algorithm shows a qualitative performance.

The rest of this paper is organized as follows: In Section 2, details of the proposed ALS point cloud semantic classification method are elaborated. Section 3 presents the experimental results produced by the proposed algorithm, and the paper concludes with Section 4.

#### 2. Methodology

It is the goal of this paper to present an efficient CRF-based framework for semantic classification from ALS point cloud data without the use of image data providing spectral information. Firstly, multiple features of ALS point clouds are processed mainly based on their locations which can efficiently improve the results of the point-based classification process. Secondly, a Random Forests (RF) classifier is employed to produce the soft labeling results. Some outliers are found in the initial semantic result, then, a CRF framework is presented to smooth the result with context information between neighboring points. However, we find that it is of low accuracy for the objects with a small size, especially for cars. LCS-CRF is proposed to solve this problem and can achieve higher overall accuracy with a higher-order potential. Cluster-based features are extracted on the cluster obtained by a constrained mean-shift clustering method and semantic rules are defined. Then, based on the common knowledge of semantic rules, we define the higher-order potentials. Finally, the location, context and semantics cues are, respectively, encoded by unary, pairwise and higher-order potentials. Once fused, they can provide complementary information from varying perspectives, to improve the ALS point cloud semantic classification performance. A mean-field approximate inference method is employed to obtain the semantic classification results. Figure 1 shows the flowchart of the proposed method.



**Figure 1.** Flowchart of the location-context-semantics-based conditional random field (LCS-CRF) algorithm integrating location, context, and semantics cues. ALS = airborne laser scanning.

#### 2.1. Feature Extraction

#### 2.1.1. Point-Based Feature Extraction

Three types of features are employed in this section, geometric features from the ALS point cloud properties, local shape features from the structure tensor and primitive features from the data source. Since the distinctiveness of point-based features strongly depends on the respective neighborhood encapsulating those 3D points, a data-driven approach is proposed to determine the neighborhood size by selecting the number of nearest neighbors in the local 3D neighborhood of each individual point with eigenentropy-based scale selection [7]. The neighbor size can be determined based on the minimum eigenentropy by varying values of the scale parameter:

$$E_{i,\lambda}(k) = -\sum_{s=1}^{3} \lambda_{i,s}(k) \ln \lambda_{i,s}(k), \qquad (1)$$

$$k_i^* = \underset{k \in \kappa}{\arg\min E_{i,\lambda}(k)},$$
(2)

where  $E_{i,\lambda}(k)$  is the eigenentropy of *i*th point based on the scale parameter *k*, and  $k_i^*$  represents the optimal value for *i*th point. Three eigenvalues ( $\lambda_s, s = 1, 2, 3$ ) can be derived by the symmetric positive semi-definite 3D structure tensor  $T \in \mathbb{R}^{3\times3}$ , which is obtained by the *k* nearest neighbors of each point. In the scope of our work, scales parameters within an interval  $\kappa = [k_{\min}, k_{\max}]$  are considered, with a lower boundary of  $k_{\min} = 10$  neighbors to remain robustness statistically [25–27] and an upper boundary of  $k_{\max} = 100$  to limit the computational effort.

After the recovery of local neighborhoods, we congregate some features which well-suit this semantic classification for ALS point cloud. The features used in our work are shown in Table 1. The point-based feature vector comprises 34 elements.

Table 1. Three types of point-based features used in this work.

Туре	Components
Geometric Features	$H, \Delta H, \sigma_H, r, D, \sigma_D, k1, \sigma_{k1}, k2, \sigma_{k2}, C_g, \sigma_{C_g}, C_m, \sigma_{C_m}, N, \sigma_N, C, \sigma_C, V$
Geometrie i cutures	$r_{2d}, D_{2d}, \sigma(D_{2d})$
Local Shape Features	$L, P, S, O, A, E, \sum_{Es}, \Delta C$
I	$\sum_{Es,2d}$ , $R_{2d}$
Primitive Features	$I, \sigma(I)$

Height H above Digital Terrain Model (DTM) is a discriminating feature to distinguish different classes. The DTM can be generated based on the local topography of the scene [26]. General geometric properties are represented by the radius r of the sphere encompassing k nearest neighbors and the maximum difference  $\Delta H$  within the neighborhood. Density *D*, principle curvatures k1 and k2, Gaussian curvature  $C_g$ , mean curvature  $C_m$ , and verticality V [28] are used to describe the basic properties of ALS data, which has been demonstrated their efficiency by feature important analysis. Normal vector relationships N and curvature C (i.e., normal change rate) are also derived in this work.  $\sigma(\cdot)$  means the variance of above geometric features in a sphere of radius r. With the k nearest neighbors of each point, 3D structure tensor  $T \in \mathbb{R}^{3\times 3}$  can be derived to obtain 8 local shape features: linearity *L*, planarity *P*, scattering S, omnivariance O, anisotropy A, eigenentropy E, sum of eigenvalues  $\sum_{Fs}$ , and change of curvature  $\Delta C$ . Intensity *I* obtained directly by the ALS laser and its variance  $\sigma(I)$  in a sphere of radius r comprise the primitive feature set. In analogy to the 3D case, 2D projection of the 3D points onto XY-plane can reveal complementary information, especially for perfectly vertical structures. Then,  $r_{2d}$  defined by the circle encompassing k nearest neighbors, 2D structure tensor  $T \in \mathbb{R}^{2 \times 2}$  (sum of eigenvalues  $\sum_{Es,2d}$ , ratio of eigenvalues  $R_{2d}$ ), density  $D_{2d}$  [29], and its variance  $\sigma(D_{2d})$  are also derived as the elements of the point-based feature vector.

#### 2.1.2. CSF with RANSAC

To increase the efficiency of LCS-CRF, off-ground points are employed to extract cluster-based features for the higher-order potentials. CSF [30] algorithm can be used to extract off-ground point from LiDAR data, which has been shown superior performance compared with other ground filtering methods.

Two difficulties should be overcome for the ground filtering for ALS point cloud, i.e., insufficient information of small size objects for clustering which will have an obvious effect on the class accuracy, overall accuracy of classification result [17], and misjudgment between ground and classes with lower height (e.g., low vegetation). Then, RANSAC [31] is integrated with CSF to solve these problems, which is able to segment ground and off-ground points simultaneously. Pseudocode of Algorithm 1 for the RANSAC-based CSF algorithm is shown in Appendix A.

Off-ground points set is generated in Algorithm 1 and the result is shown in Figure 2. More information of small size objects (e.g., car) and lower error samples between ground and classes are obtained. Then, clustering is performed on the off-ground points.

#### 2.1.3. Off-Ground Points Clustering

In this section, we first derive an over-segmentation of ALS point cloud by applying the mean-shift algorithm [32,33], a mountain climbing algorithm based on kernel density estimation without the need to initially specify the number of clusters. An adaptive gradient ascent is applied in the iterations of this algorithm, where shift vector value  $||\mathbf{m}||$  will be larger in areas of low point density and lower in areas of high point density [4]. An isotropic Gaussian kernel  $\Gamma$  is adopted, and shift vector value  $||\mathbf{m}||$  of point x can be defined as:

$$\|\mathbf{m}(\mathbf{x})\| = \left\| \frac{\sum\limits_{i \in S_r} \mathbf{x}_i \Gamma\left(\frac{\|\mathbf{x} - \mathbf{x}_i\|^2}{\gamma}\right)}{\sum\limits_{i \in S_r} \Gamma\left(\frac{\|\mathbf{x} - \mathbf{x}_i\|^2}{\gamma}\right)} - \mathbf{x} \right\|,\tag{3}$$

where  $S_r$  represents the set of current point's neighbors within the radius of r, and  $\gamma$  denotes the kernel width selected based on the point distribution for a considered scene.



**Figure 2.** Results contrast between Cloth Simulation Filter (CSF) and Algorithm 1: (**a**) Ground points obtained from CSF algorithm. Details of misjudgment are also shown; (**b**) off-ground points obtained from CSF algorithm. Much information of small size objects is lost; (**c**) ground points obtained from Algorithm 1. Error samples between ground and classes are obviously refined; (**d**) off-ground points obtained from Algorithm 1. Enough information of small size objects, which directly affects the results, can be provided by higher-order potential.

In this work, off-ground ALS data is heterogeneous and it is hard to distinguish different classes closed to each other in distance space (e.g., car and building or building and vegetation). Then, a constrained mean-shift algorithm is proposed, i.e., a post-processing step for the initial over-segmentation performed by mean-shift algorithm and two initial clusters with a low dissimilarity are preferred to be combined into one cluster. Two constraints are used for the dissimilarity discriminate between initial clusters:

• Constraint 1: local connectivity Local connectivity can be measured by the minimum Euclidean distance between points *p*<sub>1</sub> and *p*<sub>2</sub> obtained by

$$d(c_m, c_n) \le th_d \quad \text{s.t.} \ (c_m, c_n) = \arg\min_{(p_1, p_2)} d(p_1, p_2) : p_1 \in c_m, p_2 \in c_n,$$
(4)

where  $d(\cdot)$  is the Euclidean distance between initial cluster  $c_m$  and  $c_n$ , and  $th_d$  is the threshold of the constraint.

Constraint 2: structure correlation

$$\|\log \mathbf{T}_m - \log \mathbf{T}_n\|_F \le th_t,\tag{5}$$

where  $\mathbf{T}_m \in \mathbb{R}^{3\times 3}$  and  $\mathbf{T}_n \in \mathbb{R}^{3\times 3}$  are 3D tensor structures for *m* th and *n* th clusters,  $\log(\cdot)$  the matrix logarithm operator,  $\|\cdot\|_F$  the Frobenius norm [34], and  $th_t$  the threshold of the constraint.

The pseudocode of Algorithm 2, which shows the details of the constrained mean shift algorithm, is presented in Appendix B.

Clusters of different classes exhibit different characteristics, which can be used to extract more discriminative cluster-based features. Clusters derived from mean-shift algorithm, as shown in Figure 3a, are scattered and cluttered, which cannot show the special information for different classes. But, as shown in Figure 3b, more accuracy and discriminative information are provided to perform cluster-based feature extraction.



**Figure 3.** Results contrast between mean-shift algorithm and Algorithm 2: (**a**) Initial clusters achieved with mean-shift algorithm; (**b**) final clusters obtained by a post-processing step with constrained mean-shift algorithm.

## 2.1.4. Cluster-Based Feature Extraction

In contrast to point-based feature extraction, features of cluster are extracted in this section. Point-based features can describe the details of a single point, whereas whole level information for different classes can be obtained from clusters and used to derive higher-order potentials. Herein, five features are extracted from each cluster:

- Hight  $F_H$  Hight above ground measured by the barycenter of the cluster is used to distinguish the roofs and other classes (e.g., cars, low vegetations), as even the lowest roofs are generally higher than cars or low vegetations.
- Distribution of ground points *F<sub>G</sub>* A circular region centered on the cluster center can be divided in to angular bins. The distribution of ground points can be described by the proportion of bins containing ground points [35]. This feature can be used to classify objects which are adjacent to ground.
- Roughness *F<sub>R</sub>* Roughness can be determined by the variance in distances between the points and the fitting plane computed on its kernel size, namely the scale of a sphere containing nearest points. Smooth surface, such as roofs and facades, can be distinguished by this feature from other classes (e.g., cars, vegetations).
- Compactness *F<sub>C</sub>* Compactness can be measured by the volume of the convex hull divided by the area for each cluster. The number of points in a cluster is defined here as the area. A small compactness will be obtained for erect or small size classes.
- Normal correlation *F<sub>N</sub>* This feature can be measured by the correlation between normal vectors of cluster and the vertical direction of the horizontal plane, which has shown a better performance for regular classes compared with other classes.

All above cluster-based features have been proven effective in distinguishing one or more classes from others. As shown in Figure 4, each feature's capacity is distinguished with different color-coded values.



**Figure 4.** Cluster-based features: (a) height  $F_H$ ; (b) distribution of ground points  $F_G$ ; (c) roughness  $F_R$ ; (d) compactness  $F_C$ ; (e) normal correlation  $F_N$ .

## 2.2. The LCS-CRF Model

To conveniently describe the semantic classification problems, we first establish the notations and definitions used throughout the paper. Consider the input ALS point cloud  $\mathbf{V} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N\}$ , where  $\mathbf{v}_i$  ( $i \in \mathbf{V} = \{1, 2, \dots, N\}$ ) represents a 3D point corresponding to the vertices in a graphical model, and N is the total number of points. A labeled point cloud can be represented by vector  $\mathbf{y} \in \Omega$ , containing the labels  $y_i$  for all points.  $y_i$  takes its value from the label set  $L = \{1, 2, \dots, l\}$ , where l denotes the number of classes. Edges  $e_{ij} \in \mathbf{E}$  are used to model the relations between pairs of adjacent points  $\mathbf{v}_1$  and  $\mathbf{v}_2$ . Then, an undirected graphical model with graph  $G(\mathbf{V}, \mathbf{E})$  consisting of nodes  $\mathbf{V}$  and  $\mathbf{E}$  can be constructed.

## 2.2.1. Pairwise CRF Model

Pairwise CRF model is widely used in semantic classification [13,36,37] to model the spatial interaction in both the labels and observed values, which is of importance in semantic classification. It is a discriminative classification approach, which directly models the posterior probability of the label y conditioned on the observed data x [38,39]. No more than two kinds of cliques are defined in a pairwise CRF. With the Hammersley–Clifford theorem, the CRF model as a Gibbs distribution can be modeled by:

$$P(\mathbf{y}|\mathbf{x}) = \frac{1}{Z(\mathbf{x})} \exp\left\{-\sum_{c \in \mathbf{C}G} \phi c(\mathbf{y}c|\mathbf{x})\right\},\tag{6}$$

where  $Z(\mathbf{x})$  is the partition function,  $\mathbf{C}_G$  the set of all the cliques, and  $\phi c(\mathbf{y}c|\mathbf{x})$  the potential function defined over the clique *c* to model the relationship of the random variables. An assignment of all the random variables (i.e., a labeling) takes values from  $\Omega := L^N$ . Based on the Bayesian maximum a posteriori rule, the most likely labeling  $\mathbf{y}^*$  is inferred based on the given observation, which can be described as:

$$\mathbf{y}^* = \arg \max_{\mathbf{y} \in \Omega} P(\mathbf{y} | \mathbf{x}). \tag{7}$$

The semantic classification problem with pairwise CRF model is therefore equivalent to finding the minimization of the Gibbs energy function  $E(\mathbf{y}|\mathbf{x})$ , which can be described by the sum of the unary and pairwise potentials. As a special case of Equation (6),  $E(\mathbf{y}|\mathbf{x})$  is formulated as:

$$E(\mathbf{y}|\mathbf{x}) = -\log P(\mathbf{y}|\mathbf{x}) - \log Z(\mathbf{x})$$
  
=  $\sum_{i \in \mathbf{V}} \phi_i(y_i, \mathbf{x}) + \sum_{(i,i) \in \mathbf{E}} \phi_{ij}(y_i, y_j, \mathbf{x}) ,$  (8)

where  $\phi_i$  is the unary potential term, a proxy for the initial probability distribution across semantic classes, and  $\phi_{ij}$  is the pairwise potential term to keep smoothness and consistency between predictions.

## 2.2.2. LCS-CRF Model

Compared with pairwise CRF, richer statistics of point cloud can be captured by LCS-CRF. The problem of misclassification among different classes can be efficiently addressed by encoding higher-order semantics information, which can be employed in CRF model to improve the semantic classification performance. In our work, the potential functions are divided in three parts (i.e., unary, pairwise, and higher-order potentials) based on various cliques:

$$\max_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}) \Leftrightarrow \min_{\mathbf{x}} E(\mathbf{y}|\mathbf{x}) = \sum_{i \in \mathbf{V}} \phi_i(y_i, \mathbf{x}) + \sum_{(i,j) \in \mathbf{E}} \phi_{ij}(y_i, y_j, \mathbf{x}) + \sum_{c \in \mathbf{C}} \phi_c(y_c, \mathbf{x}),$$
(9)

where **C** represents the set of higher-order cliques, and  $\phi_c$  are the higher-order potentials defined over cliques.

Then, the mean-field approximate inference algorithm is employed to optimize the energy function to obtain the final labels. Specifically, the location, context, and semantics are congregated in a higher-order CRF model, and the flowchart of the LCS-CRF-based semantic classification implemented in our study is shown in Figure 5.



Figure 5. Flowchart of the LCS-CRF-based semantic classification in this study.

## 2.3. LCS-CRF Energies

2.3.1. Point-based Features for Unary Potentials

The location information of point  $\mathbf{v}_i$  and its optimal neighbors are used to determine the point-based feature vectors, by which the unary potentials  $\phi_i$  linking the point to the class labels determines the most probable label for a single point. The unary potentials  $\phi_i$  can be defined by a discriminative classifier with a probabilistic output [40].

An ensemble learning method, RF classifier is employed to produce the soft labeling results for the unary potentials. RF classifier, constructing a multitude of decision trees during training and integrating the class probabilities of the individual trees at a testing stage, has shown a superior performance based on its robustness, high accuracy, and feasibility for ALS data [9]. In the implementation, each decision tree casts a vote for the most likely class. If the number of votes casts for a class *l* is  $N_l$ , the unary potential is defined by

$$\phi_{i,RF}(y_i = l, \mathbf{x}) = -\ln(N_l/N_t),\tag{10}$$

where  $N_l$  is the total number of decision trees. Based on the point-based features, the location cues are directly used to discriminate the ALS points by the class membership probabilities.

#### 2.3.2. Weighted Potts Model

The pairwise potential  $\phi_{ij}$  incorporates the contextual cues based on the spatial smoothing dependence principle. Based on the prior spatial knowledge, neighboring points are expected to take the same label. The weighted Potts model has been shown to work well for semantic classification in many previous studies [41,42]. Herein, the pairwise potential takes the form of:

$$\phi_{ij}(y_i, y_j, \mathbf{x}) = \mu(y_i, y_j) \left[ w_1 \exp(\frac{\left\| \mathbf{p}_i - \mathbf{p}_j \right\|^2}{2\theta_1^2}) + w_2 \exp(-\frac{\left\| \mathbf{x}_i - \mathbf{x}_j \right\|^2}{2\theta_2^2} - \frac{\left\| \mathbf{p}_i - \mathbf{p}_j \right\|^2}{2\theta_3^2}) \right], \tag{11}$$

where **x** and **p** represent the observed values and 3D coordinates. The label compatibility function  $\mu(\cdot)$ , the weights of the spatial kernel and bilateral kernel  $w_1$  and  $w_2$ , and the parameters of Gaussian kernels  $\theta_1$ ,  $\theta_2$ , and  $\theta_3$  are learned on the training set with the implementation provided in Reference [43].

Based on the spatial relationship, contextual relations between classes can be modeled and weighting factors are defined depending on how likely two classes occur near each other.

#### 2.3.3. Higher-Order Potentials

Higher-order potentials are incorporated in a CRF model to capture richer perception between features and classes with semantics cues. In our work, the higher-order potentials are directly modeled by the cluster-based features with a sigmoid function. The sigmoid function is usually used as the activation function in many classification methods [44–46], which can be seen in Figure 6.



**Figure 6.** The Sigmoid function with scaling parameter  $\lambda$  and translation parameter  $\varepsilon$ .

Before computing the higher-order energy of CRF defined in (9), the cluster-based features are normalized in [0, 1] to balance the perception between features and classes. Furthermore, because some features are only discriminated and beneficial for specified classes, the perception of all of the cluster-based features with regard to the labels on two test datasets, described in Section 3.1, can be summarized in Table 2, respectively. To simplify the description, the perception between a normalized feature f and each label y,  $\mathbf{R}[\bullet]$ , can be modeled by:

$$\mathbf{R}[f,y] = \begin{cases} \frac{1}{1+e^{-\lambda(f-\varepsilon)}} & if " \propto "\\ 1 - \frac{1}{1+e^{-\lambda(f-\varepsilon)}} & if " - \propto "\\ 0 & if "/" \end{cases}$$
(12)

where  $\lambda$  is the scale parameter, and  $\varepsilon$  the translation parameter.

**Table 2.** Perception of the features for each class. Respectively, the symbols " $\propto$ " and " $-\alpha$ ", indicate the feature value that tends to be large or small in the corresponding class.

Label\Feature	$F_H$	$F_G$	$F_R$	F <sub>C</sub>	$F_N$
Low vegetation	$-\infty$	-∞	-∝	œ	œ
Ground	$-\infty$	œ	$-\infty$	œ	œ
Car	$-\infty$	œ	$\infty$	œ	$\infty$
Fence	$-\infty$	œ	$\infty$	$-\infty$	$\infty$
Roof	œ	$-\infty$	$-\infty$	œ	$\infty$
Facade	œ	œ	$-\infty$	$-\infty$	$-\infty$
Shrub	œ	œ	$\infty$	œ	$\infty$
Tree	cc	$-\infty$	œ	œ	œ
		(a)			
Label\Feature	$F_H$	$F_G$	$F_R$	$F_C$	$F_N$
Ground	$-\infty$	œ	$-\infty$	$-\infty$	œ
Building	$\propto$	$-\infty$	$-\infty$	$-\infty$	œ
Car	$-\infty$	œ	$\propto$	$-\infty$	$\propto$
Low vegetation	$-\infty$	œ	œ	œ	œ
High vegetation	œ	$-\infty$	œ	œ	$-\infty$
		(b)			

Specifically, some semantic rules are defined to adjust the higher-order potentials. Discriminative thresholds  $\tau_H$  and  $\tau_G$  for  $F_H$  and  $F_G$ , respectively, can be used to classify buildings and vehicles. Buildings and facades have a lower value in  $F_R$ , which must be smaller than a threshold  $\tau_R$ . The values of  $\tau_H$ ,  $\tau_G$ , and  $\tau_R$  are semantically defined based on common knowledge, and are generally suitable in all scenes. Then, the higher-order potentials are defined as:

$$\phi_c(y_c = l, \mathbf{x}) = -\ln(S(y_c, \mathbf{x}))$$

$$S(y_c = l, \mathbf{x}) = \begin{cases} 0, & \text{if against rules} \\ \sum_{f_c \in \mathbf{N}_F} \mathbf{R}[f_c, l], & \text{otherwise} \end{cases},$$
(13)

where  $\mathbf{N}_F = [f_H, f_G, f_R, f_C, f_N]$  is the normalized set of cluster-based features. We consider that off-ground points in a cluster share the same higher-order potential. To reduce the complexity of inference, the higher-order potentials can be rewritten by class membership probabilities and turned into unary potentials [43]. The integrated unary potentials can be written as:

$$\phi_i(y_i = l) = \begin{cases} -\ln(N_l/N_t), & \text{if } i \in G\\ -\ln[\zeta(\frac{N_l}{N_t}) + (1-\zeta)\frac{S(y_i=l)}{\sum\limits_{y \in L} S(y)}], & \text{otherwise} \end{cases},$$
(14)

where  $\zeta$  is a free parameter from 0 to 1, to compromise the location cues and semantics cues.

### 2.4. Evaluation Metrics

For evaluation, we compare the derived semantic labeling to the ground truth on a per-point basis. The confusion matrix and five commonly used measures are employed. The evaluation metrics are represented by overall accuracy (OA), Kappa coefficient (KA), recall (R), precision (P), and  $F_1$ -score. Generally, the number of examples per class is inhomogeneous in the test data, and then OA and KA are used to reflect the overall performance and the degree of consistency. Meanwhile, R represents a measure of completeness or quantity, and P represents a measure of exactness or quality. The  $F_1$ -score is a compound metric which combines P and R with equal weights. Appendix C describes the formulas in detail.

## 3. Experimental Analysis

To evaluate the performance of the proposed LCS-CRF algorithm, experiments with two ALS data sets were performed on a Windows 10 64-bit, Intel Core i7-4790k 4.00GHz processor with 32 GB of RAM, using Python language.

## 3.1. Study Areas

Two labeled benchmark datasets, Vaihingen Dataset (Figure 7) and GML Dataset A (Figure 8), are employed to evaluate our methodology for ALS data of different characteristics.



Figure 7. Vaihingen Dataset.

The Vaihingen Dataset is provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) and was acquired with a Leica ALS50 system over Vaihingen, Germany, with an average point density of 4 points/m<sup>2</sup>. In the scope of the ISPRS Benchmark on 3D Semantic Labeling, a reference labeling was performed with respect to nine semantic classes, (namely, power line, low vegetation, impervious surfaces, car, fence/hedge, roof, facade, shrub, and tree). Thereby, each point in the data set is labeled accordingly [9]. For this dataset, containing about 1.166 M points in total, a split into a training scene (about 754 k points) and a test scene (about 412 k points) is provided. For each point, its XYZ-coordinates and intensity value are provided.

The GML Dataset A is provided by the Graphics & Media Lab, Moscow State University, and publicly available sources. This dataset has been acquired with an ALTM 2050 system (Optech Inc.) and contains about 2.077M labeled 3D points, whereby the reference labeling has been performed with respect to five semantic classes (namely, ground, building, car, tree, and low vegetation). For this dataset, a split into a training scene and a test scene is provided. For each point, its XYZ-coordinates are provided without intensity value.



Figure 8. GML Dataset A.

#### 3.2. Qualitative Comparison

In this section, we mainly focus on the analysis of three stages, i.e., ground points filtering, off-ground points clustering, and LCS-CRF performing.

To visually compare our proposed Algorithm 1 with the CSF method, some small parts with meaningful information are selected from Vaihingen Dataset and GML Dataset A, as shown in Figure 9. In Figure 9, each group (Figure 9a–h) presents the comparison of filtering results for off-ground point with CSF method and our proposed Algorithm 1. We can observe that some confusing object information, especially for small size objective, can be extracted from ground points set, which can be obtained from CSF method. Not only our method can extract off-ground points from ground point sets, but it can also enhance the reliability of higher-order potentials by eliminate the misclassification between off-ground and ground points. Yet, it has two shortcomings: (1) a fraction of ground points are filtered as off-ground points, which cause a coarse cluster-based classification result; and (2) different parameters should be explored for ALS data diversity. To overcome these shortcomings, we further consider the ground as one of the objectives classified in the calculation of higher-order potentials. Besides, sensitivity analysis for parameters is shown in Section 3.4.1.

Compared with point-based features, the cluster-based features can provide new attributes, upon which semantics cues can be effectively employed. We define five cluster-based features for the derivation of higher-order potentials, which relate closely to the clustering results of off-ground points. Figure 10 presents the clustering results for the test data from Vaihingen Datasets and GML Dataset A, based on the off-ground points, which are extracted with Algorithm 1. As shown in Figure 10, class *roof* (green in Figure 10a)/*building* (blue in Figure 10c), which is far from ground with smooth surface; class *car* (cyan in Figure 10a and reseda in Figure 10c), which has a high correlation with ground; and class *tree* (yellow in Figure 10a)/*high vegetation* (orange in Figure 10c), which has a roughness surface, tend to be aggregated into single cluster, and we can make the utmost of semantics cues on these clusters. Due to the similarity of attributes for some different classes, mis-clusters, which means multiple classes contained in a cluster, also exist in the clustering results. Then, we employ the clustering result to define the higher-order potential in the LCS-CRF model, rather than the final semantic classification

result. In the LCS-CRF model, we integrate the point-based features and cluster-based features, which show different attributes for each point and complement mutually.



**Figure 9.** Comparison of CSF method with our off-ground points filtering method: (**a**–**d**) More *façade* points, *car* points, etc., can be integrated into off-ground points for Vaihingen Dataset; (**e**–**h**) more *car* points, *vegetation* points, etc., can be integrated into off-ground points for GML Dataset A.

To better evaluate the effectiveness of the LCS-CRF model, the qualitative results for three classification algorithms (i.e., RF, CRF, and LCS-CRF) of the two test datasets are, respectively, shown in Figures 11 and 12. To learn the RF models, 400 trees are sufficient in our work. One thousand training samples for each class are randomly chosen from the reference ground-truth data of Vaihingen Dataset and GML Dataset A. The performance of RF in the case of limited training samples can be shown in Figure 11a,b and Figure 12a,b. The soft labeling results for each class, produced by RF, are considered as the unary term of CRF and LCS-CRF.



**Figure 10.** Filtered and clustering results of off-ground points from two ALS datasets: (**a**) off-ground points of Vaihingen Dataset obtained with Algorithm 1; (**b**) clusters of (**a**) obtained with Algorithm 2; (**c**) off-ground points of GML Dataset A obtained with Algorithm 1; (**d**) clusters of (**c**) obtained with Algorithm 2.

(c)

(d)

As can be seen in Figure 11a,b, RF results in a discontinuous shape with lots of discrete points, due to the lack of consideration for the spatial contextual information. By considering the contextual information to alleviate the effect of noise, CRF can deliver a smoother classification map. Although the classification performance of a CRF model can be promoted dramatically by combining contextual information compared with RF method, their classification performance in keeping useful details are different. Due to the similarity between point-based features of different classes, e.g., ground and low vegetation, tree, and roof, etc., misclassified points are aggregated together, as shown in Figure 11d, which always directly affects the accurate interpretation of the various classes. It is a challenging task to accurately discriminate similar classes. However, on the whole, our proposed LCS-CRF model can achieve the semantic classification result with fewer misclassified regions and less salt and pepper classification noise by employing location-contextual-semantics cues. As shown in Figure 11e–f, the proposed model shows a competitive visual performance and can preserve useful detail information.

To verify the robustness of our method, another high-resolution ALS data of a different sensor is used to assess the performance of proposed method. Similarly, the semantic classification results of GML Dataset A obtained by three methods, i.e., RF, CRF, and LCS-CRF, are shown in Figure 12. Similar to the above test, CRF can deliver smoother results than RF and an improvement in the classification accuracy. Compared with RF model, CRF tends to greatly reduce the classification noise based on context cues. Then, some potentially useful details may also be eliminated. In this experiment, there is a slight difference of point-based features between the class car and low vegetation, which are easily confused. As shown in Figure 12c,d, an obvious misclassification has been presented. Most *low vegetation* points are classified as *car*, which limited the accuracies of the low vegetation and car. With the proposed LCS-CRF model, not only the location and context information are considered, but also

the semantics to alleviate the misclassification effectively are fused. The visual results in Figure 12e, *f*, show an improvement for the *car* and *low vegetation* classification.



**Figure 11.** The semantic classification results and classification errors for the Vaihingen Dataset: (**a**,**b**) RF; (**c**,**d**) CRF; (**e**,**f**) LCS-CRF.



**Figure 12.** The semantic classification results and classification errors for the GML Dataset A: (**a**,**b**) RF; (**c**,**d**) CRF; (**e**,**f**) LCS-CRF.

It is observed that our proposed method outperforms RF and CRF. An improvement in the quantitative metrics will be analyzed in the next section, in which the quantitative performances of Vaihingen Dataset and GML Dataset A are also reported.

## 3.3. Quantitative Comparison

In this section, the corresponding quantitative performances of Vaihingen Dataset and GML Dataset A are reported and analyzed. In accordance with Figure 11e,f and Figure 12e,f, our method can correctly label most of the test data. It can achieve a high OA of 83.1% and KA of 78.5% on the Vaihingen Dataset with eight categories of objects and a high OA of 94.3% and KA of 89.3% on the GML Dataset A with five categories of objects.

We classify semantic classification methods for the Vaihingen Dataset into two categories: traditional machine learning-based and deep learning-based. We compare our method with the result provided in Reference [26] and the submitted results with published papers provided by the ISPRS Semantic Labeling Benchmark. Reference [5,47,48] adopted the traditional machine learning classifiers to classify ALS point clouds, while Reference [49–52] leveraged deep learning for the semantic classification. For the sake of clarity and readability, the results achieved by each research group and our model (namely LCS-CRF) are listed for comparison in Table 3.

Methods	P (%)							OA	$\overline{F_{1}}$	
memous	low_veg	imp_sur	car	fence	roof	facade	shrub	tree	- Un	1
MSF [26]	67.5	82.7	35.7	14.1	86.3	39.9	32.2	69.9	68.1	52.6
HM_1 [47]	83.8	89.1	51.4	36.6	91.6	61.9	38.6	77.9	80.5	66.4
UM [48]	78.6	88.0	89.6	28.8	93.6	66.5	38.8	71.8	80.8	59.0
LUH [5]	83.0	91.8	86.4	49.5	97.3	52.4	34.1	87.4	81.6	68.4
RF	83.0	88.2	15.9	11.9	91.0	23.4	26.9	66.9	71.2	51.7
CRF	83.8	89.5	73.6	18.4	92.0	34.4	30.8	74.2	78.3	59.3
LCS-CRF	84.9	89.3	92.2	29.6	91.9	45.7	44.9	76.5	83.1	60.8
BIJ_W [49]	77.1	88.5	61.6	55.7	92.5	85.5.	39.2	80.1	81.5	60.3
RIT_1 [50]	88.0	89.6	70.1	66.5	95.2	51.4	33.4	86.0	81.6	63.3
Whu Y4 [51]	80.6	90.4	71.0	73.0	93.1	62.4	55.2	81.9	84.9	69.2
NANJ2 [52]	90.0	89.2	83.4	50.5	95.7	47.6	45.4	88.3	85.2	69.3

**Table 3.** Scores per class for each method and corresponding overall accuracy (OA) and  $\overline{F_1}$  (%).

MSF: the method based on multi-scale features.

We perform experiments on another ALS dataset, i.e., GML Dataset A, to verify the effectiveness of our method. The LCS-CRF model ranks first in terms of the OA and  $\overline{F_1}$  compared with other methods listed in the Table 4.

Methods		OA	$\overline{F_1}$				
memous	Ground	Building	Car	Tree	Low Vegetation	UII	• 1
MSF [26]	97.5	47.2	17.2	98.7	10.8	90.5	58.5
AMN [53]	74.8	7.9	32.6	98.8	88.7	-	-
RF+LBP	95.3	66.8	13.3	97.9	14.6	86.0	57.3
RF+α-exp	94.0	69.9	14.0	98.1	17.3	87.8	57.6
RF	98.5	22.8	7.3	98.4	6.8	84.4	48.3
CRF	96.6	42.8	10.8	97.1	19.9	92.4	57.1
LCS-CRF	96.8	59.6	33.3	97.2	30.8	94.3	70.2

**Table 4.** Scores per class for each method and corresponding OA and  $\overline{F_1}$  (%).

AMN: Associative Markov Networks; LBP: Loopy Belief Propagation.

#### 3.4. Sensitivity Analysis for Parameters

In our experiments, the LCS-CRF model obtained a good classification performance. However, there are so many parameters in the LCS-CRF model to be determined, which play an important role in the classification. These parameters distribute in three parts, i.e., Algorithm 1, Algorithm 2, and higher-order potentials.

#### 3.4.1. Parameters for Algorithm 1

The implementation of CSF requires three essential parameters, including the *GR* to determine the number of particles, *CT* to select the distances between points and the simulated terrain, and *MI* to

end the simulation process. To study the sensitivity of *GR* and *CT* for the CSF algorithm, *MI* is set to be 200, which is enough for our scene. *GR* varies from 0.2 to 1.2 and 0.2 to 1.0 for the test data of Vaihingen Dataset and GML Dataset A, respectively, with a step of 0.2. *CT* selected from 0.3 to 1.3 and 0.4 to 2.4 for the test data of Vaihingen Dataset and GML Dataset A, respectively, with a step of 0.4. Sensitivity analysis for parameters is presented in Figure 13. As can be observed, better results can be obtained, which are considered as the initial input of Algorithm 1, with *GR* equal to 0.6 and *CT* equal to 0.5 for Vaihingen Dataset, and *GR* equal to 0.4 and *CT* equal to 1.2 for GML Dataset A.



**Figure 13.** (a) Parameters analysis of CSF algorithm for Vaihingen Dataset; (b) Parameters analysis of CSF algorithm for GML Dataset A.

Although the ground points filtering accuracy can be 95.1% and 96.0% for Vaihingen Dataset and GML Dataset A, more detailed off-ground object information, especially objects with small size and low height, are essential for our scene to improve the semantic classification results. Then, we employ RANSAC for the ground points obtained by CSF to enrich the off-ground information with enough filtering accuracy. The property of RANSAC for each point is mostly determined by two thresholds, the maximum distance to distinctive initial inliers among current point's neighbors and the minimum inlier ratio to determine whether the current point is an element of ground points set on the premise that the current point belongs to initial inliers.

To find an appropriate value of maximum distance and minimum inlier ratio, we test the procedure with maximum distance varying from 0.1 to 0.4 with a step of 0.05 and 0.1 to 0.8 with a step of 0.1 for the test data of Vaihingen Dataset and GML Dataset A, respectively. It is worth noting that 0.4 and 0.8 are not the cut-off values of maximum distance, only representing the variation tendency of OA for ground/off-ground points. The minimum inlier ratio varies from 0.5 to 0.8 with a step of 0.05 for the two test datasets. For evaluating the filtering results, we utilize OA for ground/off-ground points to analyze the OA. Analysis for these two parameters are shown in Figure 14. We can observe that the OA for ground/off-ground points converges to a certain value, due to the higher maximum distance and the lower minimum inlier ratio with a slight influence on the OA. To make the results more reliable, the observed results, which can show the details directly, parts of them as shown in Figure 9, are also considered to determine the values of these two parameters.



**Figure 14.** (a) Thresholds analysis for RANdom SAmple Consensus (RANSAC) algorithm on Vaihingen Dataset; (b) thresholds analysis for RANSAC algorithm on GML Dataset A.

In order to obtain more details of off-ground object information and keep the OA, the parameters are utilized to perform Algorithm 1, which are listed in Table 5. These parameters are determined based on the experimental results (as shown in Figures 9 and 14) and the properties of the input ALS point cloud.

Table 5. The Parameters setting of Algorithm 1 for Vaihingen Dataset and GML Dataset A.

Parameters	GR	СТ	MI	Max Distance	Min Ratio
Vaihingen	0.6	0.5	200	0.6	0.7
GML A	0.4	1.2	200	0.9	0.8

#### 3.4.2. Parameters for Algorithm 2

Algorithm 2 is proposed to produce clusters of off-ground points, which can be used to extract discriminative cluster-based features. In the first step, two parameters, r and  $\gamma$ , are selected for the mean-shift algorithm, which are based on the prior knowledge about the expected point distribution for the scene we consider. Then, parameters k,  $th_d$ , and  $th_t$ , which were described in Section 2.1.3, are determined for the post-processing step.

Herein, the performance of Algorithm 2 is mainly evaluated based on the intuitive result, and an experiment example has been shown in Figure 3. Then, we only provide the configuration of these parameters for Vaihingen Dataset and GML Dataset A, which is shown in Table 6.

Table 6. The Parameter configuration of Algorithm 2 for Vaihingen Dataset and GML Dataset A.

Parameters	r	γ	k	$th_d$	$th_t$
Vaihingen	2	1	20	1.4	0.8
GML A	3	1	30	1.4	0.9

## 3.4.3. Parameters for Higher-Order Potentials

In the LCS-CRF model, the higher-order potentials are derived with semantics cues based on a Sigmoid function. Two parameters are utilized to determine the formulation of Sigmoid function, and they are, respectively, denoted as  $\lambda$  and  $\varepsilon$ . Parameter  $\lambda$  mainly controls the scaling of Sigmoid function, while  $\varepsilon$  controls the translation. In this section, we also normalize the cluster-based features into [0,1], and then parameter  $\varepsilon$  is set as 0.5 to consist with the distribution of cluster-based features. The expression of Sigmoid function with different values of parameter  $\lambda$  is shown in Figure 15. The

datum line is represented by a red straight line, which is treated as a reference to Sigmoid function. It means that the values of cluster-based features are directly used for the calculation of higher-order potentials. Different curves in the figure represent the projection values of cluster-based features through Sigmoid function with different  $\lambda$ . We employ Sigmoid function to enhance the discrimination of the cluster-based features to obtain a better classification result. However, there have been a few misjudgments in terms of cluster-based features, which are utilized to obtain the higher-order potential based on the regulations described in Section 2.3.3. Then, the corresponding analysis for parameter  $\lambda$  is given to test its effect in the LCS-CRF algorithm.



**Figure 15.** Sigmoid functions with different  $\lambda$  in the interval from 0 to 1.

In order to study the sensitivity of the parameter  $\lambda$  for our method, other parameters are set to be constants. Experiments are conducted to analyze the effect of the parameter  $\lambda$ , which is varied from 2 to 12 with a step of 2 for Vaihingen Dataset and GML Dataset A. The sensitivity analysis for the parameter  $\lambda$  is presented in Figure 16. To make them more concise, we also compute the variation tendency of the OA under different settings of parameter  $\lambda$ , as shown in Figure 16a,b. The parameter  $\lambda$  shows obvious impact on the OA compared with employing the datum function, and the relative importance of the higher-order potential is increased as parameter  $\lambda$  increases.



**Figure 16.** (a) Parameter  $\lambda$  analysis in Vaihingen Dataset; (b) Parameters  $\lambda$  analysis in GML Dataset A.

We can observe that, the OA first increases as parameter  $\lambda$  increases since the semantic rules are properly utilized with Sigmoid functions to enhance the discrimination of cluster-based features. Then, the OA no longer increases at a certain value of parameter  $\lambda$  (i.e., around 6 for Vaihingen Dataset and around 8 for GML Dataset A), and even shows a slight decreasing trend, since the large varying degrees of cluster-based features can lead to the accumulation of noise from cluster-based features and cause misjudgments of clusters. The red dotted lines in Figure 16, serving as a reference, represents the classification results based on the higher-order potentials derived by datum function.

Another parameter,  $\zeta$ , is also analyzed with Vaihingen Dataset and GML Dataset A, which mainly controls the effect of the higher-order potentials in the classification. As shown in the Figure 17, parameter  $\zeta$  is selected from 0 to 1 with a step of 0.1, while other parameters are set to be constant values. The OA gradually increases in the beginning with the increase in parameter  $\zeta$ , in which the semantic rules dominate the tendency compared with location information in the unary potential. After parameter  $\zeta$  reaches up to a certain value (i.e., around 0.6 for Vaihingen Dataset and around 0.7 for GML Dataset A), the OA also shows a slight decreasing trend, since the unary potential become dominant with the increase in parameter  $\zeta$ . When  $\zeta$  equals to 1, the overall accuracies for Vaihingen Dataset and GML Dataset A reach 0.783 and 0.924, respectively, where the classification result is obtained by the CRF model. It is found that an obvious improvement of the classification results was shown in both test datasets by integrating higher-order potentials, compared with the results directly derived by CRF model.



**Figure 17.** (**a**) The trends of the OA with parameter  $\zeta$  in the Vaihingen Dataset; (**b**) the trends of the OA with parameter  $\zeta$  in the GML Dataset A.

#### 4. Discussion

From Table 3, we can observe that the OA of LCS-CRF model performs the best among all of the traditional machine learning based method. As far as the eight specific classes are concerned, our method ranks first in the *imp\_sur*, *car*, and *shrub* classes within the traditional machine learning-based methods, and its *P* surpass previous highest results with absolute advantages (+1.1%, +2.6%, and +6.1%). The RF model is mainly based on the point-based features, which are derived by the location cues of points, to perform semantic classification for ALS data. The CRF model integrates the location and contextual cues and shows a smoother result compared with the RF model (as shown in Figure 11). Obviously, the LCS-CRF model shows a superior result by incorporating location, context, and semantics cues into a higher CRF model. Especially for the *car* class, a great improvement of *P* is obtained by adding semantics cues. The class *low-veg*, with a higher *P*, mainly benefits from Algorithm 1. The OA of the LCS-CRF model ranks first among the traditional machine learning-based methods and third among the deep learning-based methods, with minor disadvantages (1.8% and 2.1% lower than the second and the first OA, respectively). Though some deep learning-based methods perform better than our method, the LCS-CRF model can also satisfy the general demand with less training costs.

In Table 4, the *P* of *car* class with LCS-CRF model surpass the results of RF and CRF model with +26% and +22.5%, which means that semantics cues play an important role in the semantic classification.

We perform the methods RF+LBP and  $RF+\alpha$ -exp by adding a regulation framework to smooth the semantic results derived by RF model. Though significant improvements are shown in *building*, *car*, and *low vegetation* classes compared with RF model, the OA of methods RF+LBP and  $RF+\alpha$ -exp are still less than 90%. The *P* of *car* class for our method is superior to others, and plausible results are shown in *ground*, *building*, *tree*, and *low vegetation* classes, which validate our proposed method.

In comparison to other approaches, our method shows several strengths. We compare the results achieved with our methodology to the ones obtained by recent approaches. Similarly, Reference [5] proposed a hierarchical higher-order CRF framework, in which, spatial and context were integrated via a two-layer CRF. The Robust  $P^n$  Potts model was utilized to build the higher-order potential in their first layer CRF. Their framework iterated and mutually propagated context to improve the classification results. The results, with their framework on the Vaihingen Dataset, have been described in Table 3 (LUH), which showed outstanding performance in  $F_1$  and revealed a rather high quality of the results in several classes. In contrast, our methodology extra integrates semantic cue in a higher-order CRF, which is a one-layer CRF with neither iteration nor propagation of context, and shows obvious increases in class car and OA by 5.8% and 1.5%, respectively. Currently, the only approach delivering semantic classification results of higher quality (with OA = 85.2% and  $F_1 = 69.3\%$ ) for the Vaihingen Dataset is the one presented by Reference [52] that leverages deep learning for the semantic labeling of ALS point clouds. Yet, a multi-convolutional neural network (MCNN) was trained to automatically learn deep features of each point from the generated contextual images across multiple scales, which was time-consuming in training process and had relatively high requirements to hardware, while the proposed LCR-CRF framework only employs explicit point-based and cluster-based features. Comparable results can be observed in Table 3 with P in classes *imp\_sur* (+0.1%), *car* (+8.8%), *façade* (-1.9%), and shrub (-0.5), and with the OA (-2.1%). Compared with [49] and [50], which also adopted deep learning for the semantic classification, the OA is, respectively, raised by 1.6% and 1.5% in our framework and P in several classes shows better performance, especially in class car. Due to the consideration of multi-scale neighborhoods, Reference [26] obtained an improved performance on the GML Dataset A by exploring contextual information across different scales in the, respectively, extracted features, while we obtain the optimal neighbors with the algorithm proposed in Reference [7] and integrate meaningful semantic cues. As shown in Table 4, our method increases the OA by 3.8% and the  $F_1$  by 11.7%, and three of the five classes' *P* are improved. The methods *RF*+*LBP* and *RF*+ $\alpha$ -*exp*, which was performed based on the methodology proposed in Reference [25], constructed graph models and employed structured regularization for spatially smoothing semantic labeling of point clouds. In our method, not only spatial information is utilized, but also context and semantic cues are integrated in a posterior probability model. In contrast with these two methods, our method better addresses some hard-to-retrieve classes, such as classes car and low vegetation, and increases OA by 8.3% and 6.5%, as observed in Table 4.

Experiment results suggest that the LCS-CRF model shows superior performance on the semantic classification for ALS data. However, there are still some misclassification in the results. For the Vaihingen Dataset, classes *fence* and *facade* are at a disadvantage due to their attributes, including the small cardinal number, sparsity, and similar characteristics with some other classes. A close-up visual inspection shows that the class *fence* is often classified as class *low\_veg* or *shrub*, which causes adverse effects on the OA and  $\overline{F_1}$ . For the GML Dataset A, classes *building* and *car* produce lower precisions compared with classes *ground* and *tree*. Based on the visual inspection of test data, class *building* with small height shows similar attributes to classes *ground* and *car*, due to its planarity and clustering. Class *low vegetation* with smaller clusters is easily classified as *car*, which is very sensitive to the *P* of class *car* due to the extremely small size of class *car* compared with the whole test dataset.

As shown in Section 3.3, parameters in three parts, i.e., Algorithm 1, Algorithm 2, and higher-order potentials, are analyzed. Most parameter values are tested in a general interval based on the attributes of point clouds and common experience. Based on the hardware described in Section 3, it takes about 1.5 h to calibrate the parameters in the first and second parts both on the Vaihingen Dataset and GML

Dataset A. The decision of parameters in the third part need a heavier time cost due to the large-scale ALS point clouds, and the time for each inference on the LCR-CRF model is about 1.2 h. Then, parallel computing is utilized to speed up the process to a great extent. Once the parameters are determined, automatic interpretation can be performed on large-scale ALS point clouds. In addition, it takes only about 0.5 h to train a CRF model on the Vaihingen Dataset in our work, while the training time in a deep learning framework takes about three to six days [54].

## 5. Conclusions

In this paper, we presented an LCS-CRF model for ALS data semantic classification. The main novelty of this framework consists of the integration of location, context, and semantics cues from irregularly distributed ALS points to semantically labeled point clouds in a higher-order CRF framework. The method processes in three main stages, i.e., (i) feature extractior; (ii) off-ground points extraction and clustering; and (iii) classification. A total of 34 point-based features from their locations and 5 cluster-based features from off-ground points' clusters are extracted to form the feature space. To effectively employ the semantics cues, off-ground points extraction and clustering are performed for the cluster-based feature extraction. Based on the location and semantics cues, the unary potentials and higher-order potentials can be derived by the RF classifier and the sigmoid function. Then, the context information between neighbor points is integrated in a higher-order CRF as a pairwise potential to smooth the classification. Therefore, the location, context, and semantics cues are, respectively, formulated in unary, pairwise, and higher-order potentials within the probabilistic LCS-CRF model to alleviate the misclassification. The experiments with two ALS point cloud data sets confirm the competitive semantic classification performance of the proposed method in both the qualitative and quantitative evaluations.

However, parameters with different values are sensitive to the classification results. In our future work, further improvements aim at preserving more potentially useful details to improve the results with fewer parameters. We also intend to investigate the potential of deep learning adapted to the ALS point cloud data.

**Author Contributions:** All the authors contributed extensively to the work presented in this paper. W.H. conceived the original idea and implemented all the experiments for the study; R.W. contributed to the article's organization and provide suggestions that improved the quality of the paper; D.H. revised the manuscript and guided the overall study; C.X. supplied experimental facilities and analyzed the experiments results. All authors have read and agreed to the published version of the manuscript.

**Funding:** We thank the ISPRS Working group III/4 and the Graphics & Media Lab, Moscow State University for providing the datasets for experiment. This work was supported by Funding of Jiangsu Innovation Program for Graduate Education, the Fundamental Research Funds for the Central Universities (Grant No. KYLX15\_0287), National Natural Science Foundation of China (61601222), Natural Science Foundation of Jiangsu Province (BK20160789), China Postdoctoral Science Foundation (2018M632303).and National Key Research and Development Project (Grant No.2017YFC0822404).

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

## Appendix A

Parameters Max Iterations (*MI*), Classification Threshold (*CT*), and Grid Resolution (*GR*) are utilized for the CSF algorithm which have been analyzed to obtain a better initial result in Section 3. The number of neighbors for a query point is determined by parameter k. th1 and th2 are used to determine the inliers and calculate the inlier proportion thresholds of the RANSAC algorithm and inlier proportion.

#### Algorithm 1. CSF with RANSAC

**Input**: ALS point cloud set  $P = \{p_1, p_2, \dots, p_n\}$ . **Parameters**: MI, CT, GR, k, th1, th2 1: Ground points set  $\{G\} \leftarrow \phi$ , Off-ground points set  $\{NG\} \leftarrow \phi$ 2: Derive the initial ground and off-ground points with CSF {G} & {NG}  $\leftarrow$  (P, MI, CT, GR) 3: **For** i = 0 to size {*G*} **do** 4: Find k neighbors of *i*-th query point  $p_i \leftarrow (i, k, P)$ 5: Inliers are determined with RANSAC algorithm  $\{IP_i\} \leftarrow th1$ 6: *Calculate the proportion of*  $\{IP_i\}$  *to all neighbors ir*<sub>i</sub>  $\leftarrow$  (*th*1, *k*) 7: If  $p_i \in \{IP_i\} \& ir_i > th2$ 8: Keep  $p_i \in \{G\}$ 9: Else then 10:  $\{NG\} \leftarrow \{NG\} \cup p_i$ 11: End For Output: {*G*}, {*NG*}

## Appendix B

Parameters *r* and  $\gamma$ , scales for neighbor size and gaussian kernel, respectively, are utilized for the mean-shift algorithm which have been analyzed to obtain a better initial result in Section 3. The number of neighbors for a query cluster is determined by parameter *k*.  $th_d$  and  $th_t$  are the constraints for local connectivity and structure correlation.

Algorithm 2. Constrained mean-shift algorithm

**Input**: ALS off-ground point set {NG} **Parameters**: r,  $\gamma$ , k,  $th_d$ ,  $th_t$ 1: Derive the initial clusters and cluster centers with mean-shift algorithm  $\{C\} \& \{C_{cen}\} \leftarrow (NG, r, \gamma)$ 2: While true do For j = 1 to size  $\{C\}$  do 3: 4: If  $C_i \in \{C\}$  then 5: Find neighbors for each cluster  $N_i \leftarrow (C_{cen}, k)$ 6: *Compare*  $C_i$  *with cluster*  $n_i \in N_i$ If local connectivity  $< th_d and$  structure correlation  $< th_t$  then 7: 8:  $C_i \leftarrow C_i \cup n_i, \{C\} \leftarrow \{C\} \setminus n_i$ 9: End If End If 10: 11: **End For** No merging happened. 12: 13: End While **Output**: *Final cluster set* {*C*}

## Appendix C

The OA, KA, *R*, *P*, and *F*<sub>1</sub>-score can be computed by the confusion matrix as follows:

$$OA = \frac{1}{N} \sum_{i=1}^{m} x_{ii}$$
(A1)

$$KA = \frac{N\sum_{i=1}^{m} x_{ii} - \sum_{i=1}^{m} [(x_{ii} + \sum_{j \neq i}^{m} x_{ji}) \times (x_{ii} + \sum_{j \neq i}^{m} x_{ij})]}{N^2 - \sum_{i=1}^{m} [(x_{ii} + \sum_{j \neq i}^{m} x_{ji}) \times (x_{ii} + \sum_{j \neq i}^{m} x_{ij})]}$$
(A2)

$$P = \frac{x_{ii}}{x_{ii} + \sum_{j \neq i}^{m} x_{ji}}$$
(A3)

$$R = \frac{x_{ii}}{x_{ii} + \sum_{j \neq i}^{m} x_{ij}}$$
(A4)

$$F_1 = 2 \cdot \frac{P \cdot R}{P + R} \tag{A5}$$

where  $x_{ij}$  represents the element of confusion matrix on *i*th row and *j*th column.

## References

- 1. Liu, X.Q.; Chen, Y.M.; Li, S.Y.; Cheng, L.; Li, M.C. Hierarchical Classification of Urban ALS Data by Using Geometry and Intensity Information. *Sensors* **2019**, *19*, 583. [CrossRef] [PubMed]
- Lodha, S.K.; Fitzpatrick, D.M.; Helmbold, D.P. Aerial LiDAR data classification using AdaBoost. In Proceedings of the 6th International Conference on 3-D Digital Imaging and Modeling (3DIM'07), Montreal, QC, Canada, 21–23 August 2007.
- 3. Zhou, Y.; Yu, Y.; Lu, G.L.; Du, S. Super-segments based classification of 3D urban street scenes. *Int. J. Adv. Rob. Syst.* **2012**, *9*, 248. [CrossRef]
- Weinmann, M.; Hinz, S.; Weinmann, M. A Hybrid Semantic Point Cloud Classification-Segmentation Framework Based on Geometric Features and Semantic Rules. *PFG J. Photogramm. Remote Sens. Geoinf. Sci.* 2017, 85, 183–194. [CrossRef]
- 5. Niemeyer, J.; Rottensteiner, F.; Sörgel, U. Hierarchical higher order crf for the classification of airborne lidar point clouds in urban areas. *Int. Arch. Photogramm. Remote Sens.* **2016**, *41*, 655–662. [CrossRef]
- 6. Ni, H.; Lin, X.G.; Zhang, J.X. Classification of ALS Point Cloud with Improved Point Cloud Segmentation and Random Forests. *Remote Sens.* 2017, *9*, 288. [CrossRef]
- Weinmann, M.; Jutzi, B.; Hinz, S.; Mallet, C. Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. *ISPRS-J. Photogramm. Remote Sens.* 2015, 105, 286–304. [CrossRef]
- 8. Hackel, T.; Wegner, J.D.; Schindler, K. Fast semantic segmentation of 3D points clouds with strongly varying density. *ISPRS Ann. Photogramm., Remote Sens. Spat. Inform. Sci.* **2016**, *3*, 177–184. [CrossRef]
- 9. 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]
- Sun, X.F.; Lin, X.G.; Shen, S.H. High-Resolution Remote Sensing Data Classification over Urban Areas Using Random Forest Ensemble and Fully Connected Conditional Random Field. *ISPRS Int. J. Geo Inf.* 2017, 6, 245. [CrossRef]
- 11. Nguyen, A.; Le, B. 3D point cloud segmentation: A survey. In Proceedings of the 6th International Conference on Robotics, Automation and Mechatronics (RAM), Manila, Philippines, 12–15 November 2013.
- 12. Sanatan, M. Unsupervised Learning and Data Clustering. Towards Data Science. 2017. Available online: https://towardsdatascience.com/unsupervised-learning-and-data-clustering-eeecb78b422a (accessed on 15 October 2019).
- 13. Hu, H.; Munoz, D.; Bagnell, J.A.; Hebert, M. Efficient 3-D scene analysis from streaming data. In Proceedings of the IEEE International Conference on Robotics and Automation, Karlsruhe, Germany, 6–10 May 2013.
- 14. Gevaert, C.M.; Persello, C.; Vosselman, G. Optimizing multiple kernel learning for the classification of UAV data. *Remote Sens.* **2016**, *8*, 1025. [CrossRef]
- 15. Poullis, C. A Framework for Automatic Modeling from Point Cloud Data. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 2563–2575. [CrossRef] [PubMed]
- 16. Li, Y.; Chen, D.; Du, X.; Xia, S.; Wang, Y.; Xu, S.; Yang, Q. Higher-Order Conditional Random Fields-Based 3D Semantic Labeling of Airborne Laser-Scanning Point Clouds. *Remote Sens.* **2019**, *11*, 1248. [CrossRef]
- 17. Zhang, Z.X.; Zhang, L.Q.; Tong, X.H. A Multilevel Point-Cluster-Based Discriminative Feature for ALS Point Cloud Classification. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 3309–3321. [CrossRef]

- 18. Zhu, Q.; Li, Y.; Hu, H. Robust point cloud classification based on multi-level semantic relationships for urban scenes. *ISPRS J. Photogramm. Remote Sens.* **2017**, *129*, 86–102. [CrossRef]
- Landrieu, L.; Boussaha, M. Point Cloud Over segmentation with Graph-Structured Deep Metric Learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–20 June 2019.
- Qi, C.R.; 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, Hawaii, HI, USA, 21–26 July 2017; pp. 2–8.
- 21. Qi, C.R.; Yi, L.; Su, H.; Guibas, L.J. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *arXiv* **2017**, arXiv:1706.02413v1.
- Huang, Q.G.; Wang, W.Y.; Neumann, U. Recurrent slice networks for 3d segmentation of point clouds. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018.
- 23. Wang, Y.; Sun, Y.B.; Liu, Z.W.; Sarma, S.E.; Bronstein, M.M.; Solomon, J.M. Dynamic graph cnn for learning on point clouds. *ACM Trans. Graph.* **2019**, *38*, 1–12. [CrossRef]
- 24. Li, Y.Y.; Bu, R.; Sun, M.C.; Wu, W.; Di, X.H.; Chen, B.Q. PointCNN: Convolution On X-Transformed Points. *arXiv* 2018, arXiv:1801.07791v5.
- Landrieu, L.; Raguet, 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]
- 26. Blomley, R.; Weinmann, M. Using multi-scale features for the 3d semantic labeling of airborne laser scanning data. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* **2017**, *4*, 43–50. [CrossRef]
- 27. Weinmann, M.; Jutzi, B.; Mallet, C. Geometric features and their relevance for 3d point cloud classification. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* **2017**, *4*, 157–164. [CrossRef]
- 28. Demantké, J.; Vallet, B.; Paparoditis, N. Streamed vertical rectangle detection in terrestrial laser scans for facade database production. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* 2012, 1, 99–104. [CrossRef]
- 29. Lari, Z.; Habib, A. Alternative methodologies for estimation of local point density index: Moving towards adaptive lidar data processing. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* 2012, 39, 127–132. [CrossRef]
- 30. Zhang, W.M.; Qi, J.B.; Wan, P.; Wang, H.T.; Xie, D.H.; Wang, X.Y.; Yan, G.J. An Easy-to-Use Airborne LiDAR Data Filtering Method Based on Cloth Simulation. *Remote Sens.* **2016**, *8*, 501. [CrossRef]
- 31. Albano, R. Investigation on Roof Segmentation for 3D Building Reconstruction from Aerial LIDAR Point Clouds. *Appl. Sci.* **2019**, *9*, 4674. [CrossRef]
- 32. Fukunaga, K.; Hostetler, L. The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Trans. Inf. Theory* **1975**, *21*, 32–40. [CrossRef]
- 33. Cheng, Y. Mean shift, mode seeking, and clustering. *IEEE Trans. Pattern Anal. Mach. Intell.* **1995**, *17*, 790–799. [CrossRef]
- 34. Arsigny, V.; Fillard, P.; Pennec, X.; Ayache, N. Log-euclidean metrics for fast and simple calculus on diffusion tensors. *Magn. Reson. Med.* **2006**, *56*, 411–421. [CrossRef]
- 35. Criminisi, A.; Shotton, J. Decision Forests for Computer Vision and Medical Image Analysis; Springer: London, UK, 2013. [CrossRef]
- 36. Zheng, S.; Cheng, M.M.; Warrell, J.; Sturgess, P.; Vineet, V.; Rother, C.; Torr, P.H. Dense semantic image segmentation with objects and attributes. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014.
- 37. Liu, F.; Lin, G.; Shen, C. CRF learning with CNN features for image segmentation. *Pattern Recognit.* **2015**, *48*, 2983–2992. [CrossRef]
- 38. Sánchez-Lopera, J.; Lerma, J.L. Classification of LiDAR bare-earth points, buildings, vegetation, and small objects based on region growing and angular classifier. *Int. J. Remote Sens.* **2014**, *35*, 6955–6972. [CrossRef]
- Lafferty, J.D.; McCallum, A.; Pereira, F.C.N. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In Proceeding of the 9th International Conference on Machine Learning, Sydney, Australia, 8–12 July 2002.
- 40. Kumar, S.; Hebert, M. Discriminative random fields. Int. J. Comput. Vis. 2006, 68, 179–201. [CrossRef]

- Wolf, D.; Prankl, J.; Markus, V. Fast semantic segmentation of 3D point clouds using a dense CRF with learned parameters. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Seattle, WA, USA, 26–30 May 2015.
- 42. Zhao, J.; Zhong, Y.F.; Shu, H.; Zhang, L.P. High-Resolution Image Classification Integrating Spectral-Spatial-Location Cues by Conditional Random Fields. *IEEE Trans. Image Process.* **2016**, *25*, 4033–4045. [CrossRef] [PubMed]
- 43. Krähenbühl, P.; Koltun, V. Parameter Learning and Convergent Inference for Dense Random Fields. In Proceeding of the International Conference on Machine Learning, Atlanta, GA, USA, 16–21 June 2013.
- 44. Ratnagiri, M.V.; Rabiner, L.; Juang, B.H. Multi-Class Classification Using a New Sigmoid Loss Function for Minimum Classification Error (MCE). In Proceedings of the 9th International Conference on Machine Learning & Applications, Fairfax, VA, USA, 11–13 December 2010.
- 45. Angelo, M.F.; Carneiro, P.C.; Granado, T.C. *Influence of Contrast Enhancement to Breast Density Classification by Using Sigmoid Function*; Springer: Berlin/Heidelberg, Germany, 2015.
- 46. Hu, Y.C.; Tsai, J.F. Evaluating classification performances of single-layer perceptron with a Choquet fuzzy integral-based neuron. *Expert Syst. Appl.* **2009**, *36*, 1793–1800. [CrossRef]
- Steinsiek, M.; Polewski, P.; Yao, W.; Krzystek, P. Semantic analysis on ALS data in urban areas using Conditional Random Fields. In Proceeding of the Tagungsband 37 Wissenschaftlich-Technischen Jahrestagung DGPF, 7–10 March 2017; Volume 26, pp. 521–531.
- 48. Horvat, D.; Zalik, B.; Mongus, D. Context-dependent detection of non-linearly distributed points for vegetation classification in airborne LiDAR. *ISPRS J. Photogramm. Remote Sens.* **2016**, *116*, 1–14. [CrossRef]
- 49. Wang, Z.; Zhang, L.Q.; Zhang, L.; Li, R.; Zheng, Y.; Zhu, Z. A deep neural network with spatial pooling (DNNSP) for 3-D point cloud classification. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 4594–4604. [CrossRef]
- 50. Mohammed, Y.; Kelbe, D.J.; Ientilucci, E.J.; Carl, S. A multi-scale fully convolutional network for semantic labeling of 3d point clouds. *ISPRS J. Photogramm. Remote Sens.* **2018**, *143*, 191–204.
- 51. Yang, Z.; Tan, B.; Pei, H.; Jiang, W. Segmentation and multiscale convolutional neural network-based classification of airborne laser scanner data. *Sensors* **2018**, *18*, 3347. [CrossRef]
- 52. 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]
- Shapovalov, R.; Velizhev, A.; Barinova, O. Non-Associative Markov Networks for Point Cloud Classification. In Proceeding of the ISPRS Technical Commission III Symposium on Photogrammetry Computer Vision and Image Analysis, Paris, France, 1–3 September 2010.
- 54. Zhao, C.; Guo, H.; Lu, J.; Yu, D.; Li, D.; Chen, X. ALS Point Cloud Classification with Small Training Data Set Based on Transfer Learning. *IEEE Geosci. Remote Sens. Lett.* **2019**, 1–5. [CrossRef]



© 2020 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/).