Article

Systematic Comparison of Power Line Classification Methods from ALS and MLS Point Cloud Data

Yanjun Wang [1,2] , **Qi Chen** [2,*] , **Lin Liu** [3,4] , **Xiong Li** [5,*], **Arun Kumar Sangaiah** [6] **and Kai Li** [1]

1   National-Local Joint Engineering Laboratory of Geo-Spatial Information Technology, Hunan University of Science and Technology, No.1 Taoyuan Road, Xiangtan 411201, China; wongyanjun@163.com (Y.W.); likai9398@163.com (K.L.)
2   Department of Geography and Environment, University of Hawaii at Mānoa, 2424 Maile Way, Honolulu, HI 96822, USA
3   Department of Geography, University of Cincinnati, Braunstein Hall, 400E, Cincinnati, OH 45221, USA; lin.liu@uc.edu
4   School of Geographic Science, Center of Geo-Informatics for Public Security, Guangzhou University, 230 Guangzhou University City Outer Ring Road, Guangzhou 510006, China
5   School of Computer Science and Engineering, Hunan University of Science and Technology, No. 1 Taoyuan Road, Xiangtan 411201, China
6   School of Computing Science and Engineering, Vellore Institute of Technology (VIT), Vellore-632014, India; arunkumarsangaiah@gmail.com
*   Correspondence: qichen@hawaii.edu (Q.C.); lixiognzhq@163.com (X.L.)

Abstract: Power lines classification is important for electric power management and geographical objects extraction using LiDAR (light detection and ranging) point cloud data. Many supervised classification approaches have been introduced for the extraction of features such as ground, trees, and buildings, and several studies have been conducted to evaluate the framework and performance of such supervised classification methods in power lines applications. However, these studies did not systematically investigate all of the relevant factors affecting the classification results, including the segmentation scale, feature selection, classifier variety, and scene complexity. In this study, we examined these factors systematically using airborne laser scanning and mobile laser scanning point cloud data. Our results indicated that random forest and neural network were highly suitable for power lines classification in forest, suburban, and urban areas in terms of the precision, recall, and quality rates of the classification results. In contrast to some previous studies, random forest yielded the best results, while Naïve Bayes was the worst classifier in most cases. Random forest was the more robust classifier with or without feature selection for various LiDAR point cloud data. Furthermore, the classification accuracies were directly related to the selection of the local neighborhood, classifier, and feature set. Finally, it was suggested that random forest should be considered in most cases for power line classification.

Keywords: laser scanning data; power line classification; random forest; feature selection; classifier

1. Introduction

Power lines are part of the important public infrastructure of cities, and their survey is a crucial task in power supply management and scientific planning [1,2]. Additionally, power lines cover various geographical terrains or different complex scenes, and have long distances. The traditional field-based inspection is labor-intensive and costly. Remote sensing images might obtain spatial information in 2D but in true 3D. In contrast, LiDAR (light detection and ranging) directly provides

high-precision 3D data for the power line corridor, making it a perfect solution for this task by saving a lot of field survey time and labor [3–6]. Airborne laser scanning (ALS) and mobile laser scanning (MLS) are two such LiDAR systems for acquiring accurate 3D data over large areas. However, the data volume of the LiDAR point cloud is usually very large, and the power lines are also close to trees and buildings in different scenes. Therefore, we need to develop highly efficient and rapid methods for power line classification from ALS and MLS in numerous types of scenes, such as forest, suburb, and urban areas.

In general, the classification methods for LiDAR point cloud could be supervised or unsupervised. In contrast to supervised methods, unsupervised ones need more priori knowledge or auxiliary information, such as for example, the direction of the power line corridor, the corridor width, the pole position [7,8], etc. The limitation makes the unsupervised methods not suitable for different types of point cloud, various terrains, and point densities. Therefore, we paid more attention to the systematic comparison of supervised methods for power line classification in this study.

There are several important factors for the supervised classification methods from the LiDAR point cloud, which involve the local neighborhood types, classifiers, and feature sets. By considering the physical characteristics of power lines, we explored the state-of-the-art and possible effects of these parameters as follows:

For the local neighborhood types, the spherical, vertical cylindrical, and k-nearest neighborhoods were commonly used in the classification of ground, tree, and buildings from airborne LiDAR points [9–13], but rarely from power lines. It is relatively unknown how such local neighborhood types work for power line extraction. Based on different types of local neighborhood, the feature extraction for supervised classifiers is a crucial issue. Kim and Sohn [14] and Guo et al. [15] extracted 21 features to characterize the horizontal and vertical properties of power line objects, and used knowledge-based classification methods to separate power lines from their background in two steps by fitting in the XOZ (consisted of X and Y axes) or YOZ (consisted of Y and Z axes) plane. These methods needed to use contextual pylon information, and the accuracy of their point-based classification was 91.04% and 89%, respectively. Guo et al., Jwa and Sohn, and Jwa et al. [16–18] extracted power lines and towers according to the semantic relationship, which is based on the position of towers. The methods were unsuitable for power line classification in complex urban scenes where small electric poles instead of tall towers are ubiquitous. Liang et al. [19] used the fact that the power lines are closely linked to extract power lines from the point cloud, but the method requires airborne LiDAR point cloud of very high density. Ritter and Benger [20] proposed to detect power line candidate points by using the non-linear adjustment of the catenary line, but the method is computationally complex and had large omission errors. Point-level features $\{X, Y, Z, \text{echo number}, \text{intensity}, \cdots\}$ are commonly used to construct the feature vectors [11,16,21–24]. Furthermore, interpretable geometrical and distributional features are extracted from the local neighborhood in many studies [9–11,25]. The determination of local neighborhood types is a critical issue for improving the classification results. The popular classifiers for laser scanning point classification include support vector machines (SVM) [24,26–28], random forests (RF) [9,14,29–32], JointBoost [16], Naïve Bayes (NB) [33,34], and so on. It is difficult to select the classifier for power line classification because of a lack of comparative studies.

In this study, we systematically compared the power line classification methods from ALS and MLS point clouds. We did comparative analysis in three different aspects (neighborhood types, classifiers, and optimized feature sets), so it aims to provide a common and flexible classification framework. Our analysis used various types of ALS and MLS point cloud data, different scenes, and data quality to see whether general findings can be produced.

2. Materials and Methods

We focused on the power line supervised classification methodology from ALS and MLS point cloud, and designed our study framework by considering three elements or parameters: local neighborhoods determination, classifiers selection, and feature sets evaluation. Our methodology

consists of power line candidate filtering, comparison between local neighborhoods, comparison between different classifiers, and comparison between selected feature sets, which are explained in Figure 1 and the following subsections.
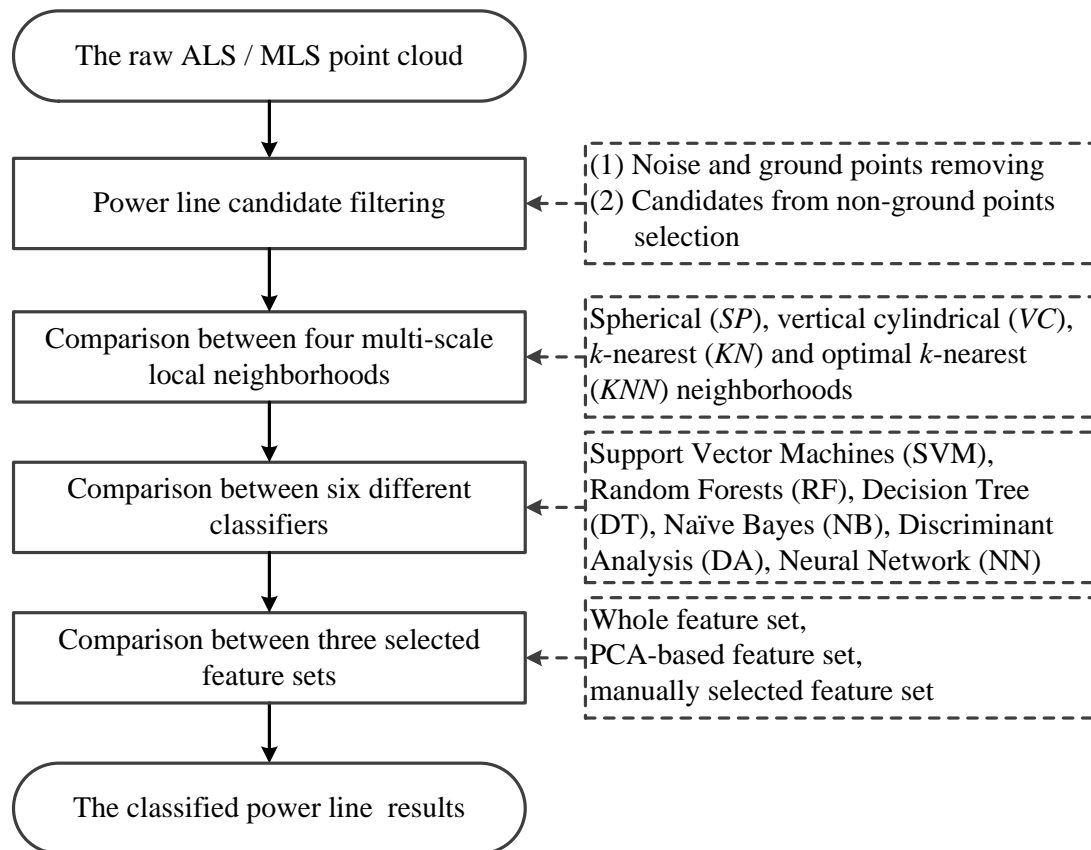

Figure 1. The whole flowchart framework of our study (ALS = airborne laser scanning, MLS = mobile laser scanning, and PCA = principal component analysis).

2.1. Dataset

We have tried to choose our datasets to cover different types of environment (urban versus forest) and data types (ALS versus MLS). However, the number of datasets we can choose is also limited by the LiDAR data available and the large amount of work required for labeling point clouds and creating the ground truth. Our research area covers five test sites (Figure 2) with varying point densities in the range of 3 m^{-2} to 124 points m^{-2} and different geographical topography such as forest, suburb, and urban scenes. In Table 1, we listed the site characteristics. For these datasets, ground truth is available in the form of a manual pointwise labeling of the power line class.

Site I and site II are in an urban area surrounding the campus of University of Hawaii at Manoa in Honolulu, Hawaii. The airborne LiDAR data in these two sites were acquired in the summer of 2013 using an Optech ALTM GEMINI laser system (scan rate: 37 Hz; laser pulse rate: 70,000 Hz; multi-pulse in air mode enabled with up to five echoes) mounted on a twin-engine Piper PA-31 Navajo airplane (aboveground flight height: ~800–1400 m). The dataset areas in site I and II are ~800 × 100 m^2 and ~520 × 360 m^2, respectively. The power lines in these sites are urban distribution lines, and the point density is both ~3.3 points/m^2. Site II is a more complex scene, in which the buildings and trees are closer to power lines.

Site III is a forest area located in Minnesota. The airborne LiDAR data were acquired in May 2013. The dataset area in this site is ~581 × 782 m^2, and the point density is ~1.6 points/m^2.

Site IV and site V are both mobile laser scanning data along primary roads located in Honolulu. The MLS point cloud was acquired in April 2015. The dataset areas in site IV and site V are ~512 × 248 m^2 and ~561 × 1411 m^2, respectively. Their point densities are ~123.7 points/m^2 and ~38.6 points/m^2.

Table 1. Overview of five research sites (ALS = airborne laser scanning, MLS = mobile laser scanning).

	Class	Site I	Site II	Site III	Site IV	Site V
	Data collection date	June 2013	June 2013	May 2013	April 2015	April 2015
Site feature	Data type	ALS	ALS	ALS	MLS	MLS
	Density (points/m^2)	3.3	3.3	1.6	123.7	38.6
	Scene type	Urban	Urban	Forest	Urban	Suburb
	Ground	136,891	279,807	293,425	3,989,964	3,174,792
	Building	48,574	115,785	0	0	0
Labeled numbers	Vegetation	73,523	116,689	426,617	11,576,752	755,585
(points)	Power line	6858	6891	3014	84,668	24,105
	Others (billboard, etc.)	2516	93,452	1733	52,249	12,941
	Total	268,362	612,624	724,789	15,703,633	3,967,423

(**a**)
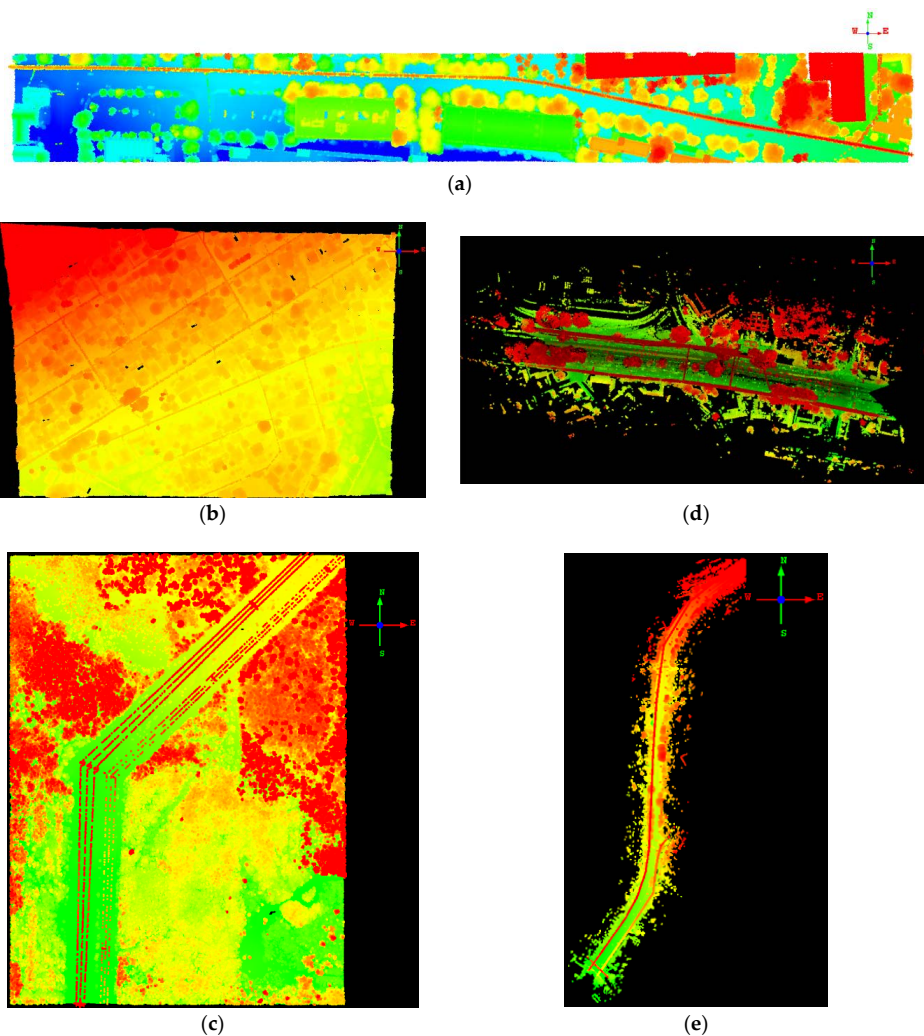

(**b**)

(**d**)

(**c**)

(**e**)

Figure 2. Visualization of the experimental datasets. (**a**) The site I dataset from the large urban ALS (airborne laser scanning) scene around the campus of the University of Hawaii; (**b**) the site II dataset from the large urban ALS scene surrounding the campus of University of Hawaii; (**c**) the site III dataset from the forest ALS scene in Minnesota; (**d**) the site IV dataset of MLS (mobile laser scanning) along primary roads located in Honolulu; (**e**) the site V dataset of MLS along primary roads located in Honolulu.

2.2. Power Line Candidate Filtering

The power line candidate filtering is the preprocessing for power line classification from the raw LiDAR point cloud. The power lines are parallel or distributed regularly between two neighboring electric poles or towers with sag [2,15,35,36]. Due to this unique geographical characteristics and large data volume, the candidate filtering aims to collect all of the possible power lines points. The filtering processing includes removing noise and ground points [37–40] and selecting the points that are 4 m above ground [41] as power line candidates.

2.3. Multi-Scale Neighborhood Based Feature Selection

2.3.1. Local Neighborhood Determination

We considered four commonly used neighborhood types for power line classification: spherical, vertical cylindrical, and *k*-nearest neighborhoods, which are defined using different geometrical parameters as follows:

- a spherical neighborhood is formed by all of the 3D points within a sphere around point *P*, which is parameterized with a fixed radius,
- a vertical cylindrical neighborhood is formed by all of the 3D points within a vertical cylindrical whose axis vertically passes through point *P* and whose radius is fixed,
- a *k*-nearest neighborhood is formed by the $k \in N$ nearest neighbors of considered point *P*, the *k* is its parameter,
- an optimal *k*-nearest neighborhood is formed by the optimal *k*-nearest neighbors based on the above-mentioned *k*-nearest neighborhood, the optimal *k* is derived by eigenentropy-based scale selection.

The spherical, vertical cylindrical, and *k*-nearest neighborhoods were used in previous studies for ground, tree, and building classification from airborne LiDAR points [9–12]. These local neighborhood types were involved in this study for classifying power lines.

2.3.2. Feature Extraction

It is important to extract the useful features that distinguish the power line points from the LiDAR point cloud. We chose the multi-scale local neighborhood to characterize a 3D structure for each considered point according to the previous studies [9–12,42]. The multi-scale neighborhood could address the multiple levels of detailed presentation of power lines, and was defined with a series of fixed parameters, which included different radii for spherical, vertical cylindrical neighborhoods, and different *k* values for the *k*-nearest neighborhood. Extracting features from a multi-scale neighborhood implies the concatenation of features from multiple single-scale neighborhoods.

To characterize 3D points in a single-scale neighborhood, their coordinates, geometric features, and distribution features have been proposed and used in the classification of ground, trees, buildings, and cars [9,10,12,24,43]. We followed this computation method and proposed three different feature datasets for power line classification as follows:

For a local neighborhood point set $\mathcal{P}$ of considered point *P*, we first computed the covariance tensor $C = \frac{1}{k} \sum_{i \in \mathcal{P}} (p_i - \overline{p})(p_i - \overline{p})^T$, where $\overline{p} = med_{i \in \mathcal{P}}(p_i)$ is its central point. Then, we acquired the eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ and corresponding normalized eigenvalues e_1, e_2, e_3. Thus, we could form three different feature sets as follows for power line classification.

(1) Feature set *A* includes the whole 26 features that were extracted by the method in Blomley et al. [9], and its partial list of geometric and distributional features can be shown in Table 2.

Table 2. A partial list of geometric features and distributional features.

Feature Class	Formal Definition	Computing Method
Geometric features	Normalized eigenvalues	$e_i = \frac{\lambda_i}{\sum \lambda}\ i \in \{1,2,3\}$
	Linearity	$L_\lambda = \frac{\lambda_1 - \lambda_2}{\lambda_1}$
	Planarity	$P_\lambda = \frac{\lambda_2 - \lambda_3}{\lambda_1}$
	Scattering	$S_\lambda = \frac{\lambda_3}{\lambda_1}$
	Anisotropy	$A_\lambda = \frac{\lambda_1 - \lambda_3}{\lambda_1}$
Distributional features	Omnivariance	$O_\lambda = \sqrt[3]{\lambda_1 \lambda_2 \lambda_3}$
	Sum	$\sum \lambda = \lambda_1 + \lambda_2 + \lambda_3$
	Changing of curvature	$C_\lambda = \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}$
	Radius of local neighborhood	$\gamma_p = dist(p_0 - p_{max})$
	Density of point set $\mathcal{P}$	$D_P = num(p_i)/\frac{4}{3}\pi r_P^3$
	Delta of point set $\mathcal{P}$ in Z axis	$\Delta_Z = max_Z - min_Z$

(2) Feature set *B* is obtained by the optimal feature subset of feature set *A*. The correlation-based feature selection and the principal component analysis (PCA) strategy were used to measure and evaluate the quality of a feature subset. We performed PCA on the extracted feature set *A* and selected the sum of variances of the first few principal components that exceeded 90% as the feature set *B*.

(3) Feature set *Γ* is aggregated by the suitable features in *A* corresponding to the power line physical characteristics. The power lines are usually distributed regularly and linearly between two neighboring electric towers or poles with sag. Considering such characteristics, we manually selected nine optimized core features as set *Γ*, including the linearity, scattering, anisotropy, changing of curvature, density, verticality, eigenvalue entropy, radius, and standard deviation of the *Z* values.

2.4. Classifiers

2.4.1. Support Vector Machines (SVM)

SVM is a non-parametric supervised learning classifier that has become commonly used in remote sensing images and laser scanning data classification [26–28]. The extracted features based on multi-scale neighborhood are the predicators of the SVM classifier. After we tried different kernel functions and compared their corresponding results, we adopted the radial basis function (RBF) kernel, kernel coefficient with four, and automatic scaling of the predicators using a heuristic procedure implemented in Matlab. The RBF kernel was commonly used and validated in many previous SVM applications [26,28,42]. For the RBF SVM, the most important parameters include gamma (the parameter related to the variance of the Gaussian radial basis function) and C (the parameter quantifying how much we penalize the "slack variables" in the objective function). We determined these parameters using the Bayesian optimization method [44] that minimized the fivefold cross-validation classification errors. We also compared this optimization method with another commonly used "grid search" method, and found that the Bayesian method had smaller classification errors and a shorter computation time.

2.4.2. Random Forest (RF)

RF has been increasingly applied in the geographical object extraction and classification from remote sensing images and LiDAR point cloud data [14,30–32,45]. RF employs a random method to establish a forest comprising many mutually independent decision trees. The forest is obtained using the training set, and then each decision tree in the forest makes a judgment about the unlabeled sample

as the category that was voted for most frequently. The number of trees is a key parameter of RF. Usually, the predictive classification error rate of RF decreases initially, and then reaches a minimum before increasing again for increasing number of trees [46]. Moreover, the best performance is usually achieved within the first 250 trees in most datasets. Therefore, we tried different numbers (e.g., 100, 200, 300, and 400) of bagged trees, and investigated the out-of-bag errors over the number of grown trees. We found that an ensemble RF of 300 bagged trees had the smallest error, and thus used it in this study.

2.4.3. Decision Tree (DT)

The use of decision trees (DTs) for remote sensing image classification has increased in recent years. For geographical object classification from LiDAR point cloud data, the most important phase is the construction of an interpretation model (knowledge) for the segmented objects [47–49]. However, it may be difficult to execute in combination with other classifiers, because DTs are similar to a "white box": users are easily able to interpret the links between the response variables of classes and the explanatory features from point cloud data. In this study, a tree is grown by binary recursive partitioning using the response in the specified formula and choosing splits. According to the introduction of DTs [50] and a comparison of parameter optimization methods, we also applied the Bayesian optimization algorithm to minimize the fivefold cross-validation loss for the tree by varying the parameters, including a maximum number of splits and split criterion.

2.4.4. Naive Bayes (NB)

The Bayes network is a powerful probabilistic representation and reasoning tool when dealing with conditions of uncertainty. It has also been widely used as a strategy or single classifier for remote sensing classifications owing to its highly scalable and incremental learning [33,51,52]. Based on the Bayes' theorem, the standard naïve Bayes classifier assumes independence of the predictor variables or features in this study. The problem is to maximize the conditional probability estimation of power line points.

2.4.5. Discriminant Analysis (DA)

Discriminant analysis (DA) is a classical classification method that has been widely applied in remote sensing images classification and other purposes [47]. In this study, we extracted Fisher's linear discriminants based on various features extracted from LiDAR point cloud data.

2.4.6. Neural Network (NN)

Neural network (NN) or artificial neural network is a computing system vaguely inspired by the biological neural network that constitutes animal brains. Such systems "learn" tasks by considering examples, generally without task-specific programming. A NN is based on a collection of connected units or nodes called artificial neurons. Each connection between artificial neurons can transmit a signal from one to another. The artificial neuron that receives the signal can process it, and then signal the artificial neurons that are connected to it. NNs have been widely used on a variety of tasks, including convolutional neural network framework for object-based classification from high-resolution images and ALS point cloud [53–56]. The studies [57,58] discussed the framework and key issues of NN, and the selection rules for the numbers of hidden layers and neurons in the hidden layers. According to these studies, we applied the two-layer feedforward NN with sigmoid output neurons for power line classification, and set the number of hidden neurons to 30 by tuning the network performance.

2.5. Experiments

We considered the following three aspects that play crucial roles for the power line classification from LiDAR points: neighborhood, classifiers, and structural features. Following the procedure described in the previous sections, we did a comparative analysis of power line classification using

four different neighborhood types (spherical, vertical cylindrical, k-nearest, and optimal k-nearest neighborhoods), six different classifiers (SVM, RF, DT, NB, DA, and NN), and three feature sets.

Specifically, we did three kinds of comparative experiments:

(i) Multi-scale neighborhood type experiments based on spherical, vertical cylindrical, and k-nearest and optimal k-nearest neighborhoods, which were denoted as *SP*, *VC*, *KN*, and *OKN*, respectively. Each multi-scale neighborhood resulted from the combination of local neighborhoods, whose radii were 1 m, 3 m, 5 m, 7 m, 9 m, and 11 m. These construction methods of local neighborhoods were commonly used in previous studies [9–12]. We adjusted the parameters and applied the SVM classifier to compare the results of different neighborhoods;

(ii) Classifier experiments including SVM, RF, DT, NB, DA, and NN. We used these six classifiers to train, validate, and test the five datasets. The features in this experiment were the whole features extracted from the datasets based on the better neighborhood type in experiment (i);

(iii) Selected feature set experiments including feature set *A*, *B*, and *Γ*. We adopted the better neighborhood type in experiment (i) and the more suitable classifier in experiment (ii) to compare the results of the different selected feature sets.

In order to compare the point-level classification results of power lines in these different experiments, we applied the commonly used fivefold cross-validation based on the manually labeled points in each data site, respectively. We divided the point clouds into five folds, and used four folds for training, and the remaining fold for testing in each iteration of cross-validation. We considered a variety of measures: (i) precision rate (*PREC*), (ii) recall rate (*REC*), (iii) quality rate (*QUA*) and (iv) processing time (*T*). The *T* contains the whole processing time from power line candidate filtering to classification. The *PREC*, *REC*, and *QUA* are computed as follows:

$$PREC = \frac{(TP)}{(TP) + (FP)} \tag{1}$$

$$REC = \frac{(TP)}{(TP) + (FN)} \tag{2}$$

$$QUA = \frac{(TP)}{(TP) + (FP) + (FN)} \tag{3}$$

where *TP* is the sum of true positives for power lines, *FP* is the sum of false positives for power lines, *FN* is the sum of false negatives for power lines, *PREC* depicts the percentage of *TP* in the power lines classification results, *REC* depicts the percentage of *TP* in the reference data, and *QUA* depicts the quality percentage as an overall measurement. The algorithm proposed was programmed in Matlab (the Mathworks, Inc., Natick, MA, USA). The computer we used has 8 GB RAM and a dual-core 2.20 GHz processor.

3. Results

3.1. Multiple Comparisons between Neighborhood Types

Table 3 summarizes the results of using spherical, vertical cylindrical, *k*-nearest, and optimal *k*-nearest multi-scale neighborhood types for these five datasets.

The vertical cylindrical neighborhood *VC* had the highest mean *PREC*, *REC*, and *QUA* rates. The multi-scale neighborhood *VC* had slightly higher *REC* and *QUA* rates than *SP* in sites I, II, and III, which used ALS point cloud data. In sites IV and V of MLS point cloud data, the *REC* and *QUA* rates of neighborhood *SP* were a little higher than the *VC* ones. Among the five experiment datasets, the *PREC* rates were similar to each other between neighborhoods *VC* and *SP*. The results of neighborhoods *KN* and *OKN* showed the lower mean *PREC*, *REC*, and *QUA* rates. Additionally, we did the paired-sample t-test analysis of the *QUA* rates between any two neighborhoods across five study sites. We found that neighborhoods *VC* and *SP* are statistically significantly better than neighborhoods *KN* and *OKN*.

In contrast, the pair *t*-test between *VC* and *SP* had a *p*-value of 0.1803, which indicated that *VC* is not statistically significantly better than *SP*. Therefore, we chose *VC* and *SP* as the local neighborhood types in the subsequent experiments.

Table 3. Classification performance for different multi-scale neighborhood types (*PREC* = precision rate in %, *REC* = recall rate in %, *QUA* = quality rate in %, *T* = processing time in seconds, *SP* = spherical neighborhood, *VC* = vertical cylindrical neighborhood, *KN* = *k*-nearest neighborhood, and *OKN* = optimal *k*-nearest neighborhood).

Data Site	Neighborhood Type	Performance Measures			
		PREC	*REC*	*QUA*	*T*
Site I	*SP*	98.3	83.7	82.5	1064
	VC	98.2	98.6	96.8	672
	KN	89.0	78.6	71.7	362
	OKN	32.6	15.3	11.6	600
Site II	*SP*	97.7	70.3	69.2	508
	VC	93.4	79.5	75.2	1110
	KN	92.0	38.8	37.5	359
	OKN	73.1	12.5	11.9	220
Site III	*SP*	98.6	49.8	49.5	927
	VC	94.2	55.0	53.2	1001
	KN	36.5	16.6	12.9	607
	OKN	28.3	29.2	16.8	466
Site IV	*SP*	99.4	99.1	98.5	1806
	VC	99.4	98.3	97.7	2694
	KN	61.8	100.0	61.8	1387
	OKN	61.8	100.0	61.8	117
Site V	*SP*	98.7	98.6	97.4	108
	VC	98.5	98.2	96.7	129
	KN	65.1	100.0	65.1	62
	OKN	65.1	100.0	65.1	30
Mean	*SP*	98.5	80.3	79.4	882
	VC	96.7	85.9	83.9	1121
	KN	68.9	66.8	49.8	555
	OKN	52.1	51.4	33.4	287

3.2. Comparisons Between Classifiers

We computed the *QUA* rates of these six classifiers with neighborhood types *VC* and *SP*, and feature sets *A*, *B*, and *Γ* in these five data sites. We also did the paired-sample *t*-test analysis of *QUA* rates between any two classifiers. We found that RF had the highest mean, and at the 5% significance level, it was statistically significantly better than any other classifiers, except for NN. The best results were obtained when using neighborhood type *VC* and feature set *A*, as shown in Table 4.

Table 4 showed that when using neighborhood type *VC* and feature set *A*, the classifiers SVM, DT, and NN had close mean *REC* and *QUA* rates. The classifiers NB and DA performed worse results, while they needed less processing time. For the ALS datasets in the urban areas of sites I and II, the classifiers SVM, RF, and NN achieved higher results of *PREC*, *REC*, and *QUA* rates. The classifiers RF, DT, and NN could be more suitable for the ALS dataset of site III in the forest area.

Table 4. Classification performance for six different classifiers (*PREC* = precision rate in %, *REC* = recall rate in %, *QUA* = quality rate in %, *T* = processing time in seconds, SVM = support vector machines, RF = random forest, DT = decision tree, NB = Naïve Bayes, DA = discriminant analysis, and NN = neural network).

Data Site	Classifier	Performance Measures			
		PREC	*REC*	*QUA*	*T*
Site I	SVM	98.2	98.6	96.8	672
	RF	98.5	95.1	93.8	412
	DT	93.1	91.0	85.3	372
	NB	37.6	96.2	37.1	321
	DA	79.5	77.9	64.9	342
	NN	96.3	96.1	92.6	341
Site II	SVM	93.4	79.5	75.2	1110
	RF	98.5	76.0	75.1	418
	DT	81.9	68.4	59.4	356
	NB	20.7	41.9	16.1	307
	DA	90.1	34.2	33.0	342
	NN	88.2	71.6	65.3	325
Site III	SVM	94.2	55.0	53.2	1001
	RF	98.1	74.2	73.1	1003
	DT	96.3	72.2	70.3	958
	NB	9.7	90.5	9.6	879
	DA	72.5	22.6	20.8	916
	NN	99.2	61.2	60.9	914
Site IV	SVM	99.4	98.3	97.7	2694
	RF	99.6	99.5	99.1	2451
	DT	99.2	98.5	97.7	2376
	NB	93.9	97.4	91.6	2327
	DA	98.6	96.5	95.1	2349
	NN	99.1	97.5	96.7	2345
Site V	SVM	98.5	98.2	96.7	129
	RF	98.9	98.6	97.6	108
	DT	98.6	98.2	96.8	107
	NB	94.1	94.7	89.3	80
	DA	97.0	97.0	94.2	89
	NN	98.7	96.3	95.1	83
Mean	SVM	96.7	85.9	83.9	1121
	RF	98.7	88.7	87.8	878
	DT	93.8	85.7	81.9	834
	NB	51.2	84.1	48.7	783
	DA	87.5	65.7	61.6	808
	NN	96.3	84.5	82.1	802

We found that the *REC* rates for site II were very low, regardless of the classifiers. The reasons are that (1) site II was in an urban scene where the points from trees, buildings, and other objects were close to the ones from power lines; it was even more complex than site I; and (2) the point density of site II was much smaller than those of site IV and site V. The recall rates for site IV were generally low as well, mainly because it had the lowest point density among all of the sites: the points' local neighborhoods in site III contained fewer points, which tended to be misclassified as points from irregular features such as trees.

Among the six classifiers, the classifier NB performed the worst in this study. The omission errors of NB for site I and II were slightly higher than classifiers SVM, RF, and NN. Sites I and II have different geographical scenes. Site I is in a simple urban area, and the power lines are almost distributed linearly and close to some trees. Site II is in a more complex urban area, and the power lines are distributed in

disorder and very close to trees and buildings. These differences result in a low omission error in site I, and a high omission error in site II, for the NB classifier.

For sites IV and V with MLS datasets, these six classifiers all performed very well with precision, recall, and quality rates mostly greater than 94%. In contrast, their performance varied substantially over the other three sites with ALS datasets. This indicates that increasing point density is an effective approach to improve the power line classification accuracy. In contrast, when the point density is relatively low, even the common powerful machine learning methods, such as RF and SVM, could suffer when classifying power lines (especially in sites II and III). Overall, the non-parametric classifiers RF, DT, and SVM had higher rates than the parametric classifiers NB and DA.

3.3. Comparisons between Selected Feature Sets

We computed the QUA rates of these three feature sets with neighborhood types VC and SP, and six classifiers in these five data sites. We also did the paired sample t-test analysis of QUA rates between any two feature sets, and found that feature set A had the highest mean, and at the 5% significance level, was statistically significantly better than the other feature sets. When the neighborhood type VC was used, it had the highest mean QUA rate. Table 5 listed the classification results of these five datasets for all of the feature sets using neighborhood type VC.

Table 5. Classification performance for three different feature sets ($PREC$ = precision rate in %, REC = recall rate in %, QUA = quality rate in %, T = processing time in seconds).

Data Site	Feature Set	Performance Measures			
		$PREC$	REC	QUA	T
Site I	A	98.5	95.1	93.8	412
	B	83.1	38.4	35.6	365
	Γ	98.0	95.7	93.9	202
Site II	A	98.5	76.0	75.1	418
	B	89.7	41.2	39.3	357
	Γ	88.5	71.3	66.7	205
Site III	A	98.1	74.2	73.1	1003
	B	93.1	51.6	49.7	1001
	Γ	85.1	72.5	69.6	415
Site IV	A	99.6	99.5	99.1	2451
	B	95.7	97.9	93.8	2361
	Γ	99.4	99.0	98.5	1061
Site V	A	98.9	98.6	97.6	108
	B	94.8	97.8	92.8	91
	Γ	98.9	98.4	97.4	54
Mean	A	98.7	88.7	87.8	878
	B	91.3	65.4	62.2	835
	Γ	94.0	87.4	85.2	387

The results in Table 5 showed that when the neighborhood type VC was used, the selected feature set Γ had similar mean $PREC$, REC, and QUA rates to the all of those in feature set A. The feature set B, which was obtained by PCA method, had the lowest mean rates. For the mean processing time T, it had dropped dramatically from 878 s in A to 387 s in Γ, because of the smaller number of features in Γ. The T results had not dramatically improved in B compared with A.

For sites IV and V with MLS datasets, the different feature sets all performed relatively well, with the precision, recall, and quality rates being greater than 92%. For the three sites (I, II, and III) with ALS datasets, feature B had the worst performance, with substantially lower quality rates. Relatively, feature set A has the best performance. Compared with feature set A, feature set Γ had a moderate

decrease in precision (from 98.4% to 90.5%) and slight decreases in recall (from 81.8% to 78.8%) and quality (from 80.6% to 76.7%) for these three ALS sites.

As an example to visualize the classification, we used the vertical cylindrical neighborhood VC, classifier RF (due to its highest mean classification accuracy, as shown in Table 4), and feature set Γ (due to its relatively high quality rate) to classify power lines across the five data sites (Figure 3a–c).
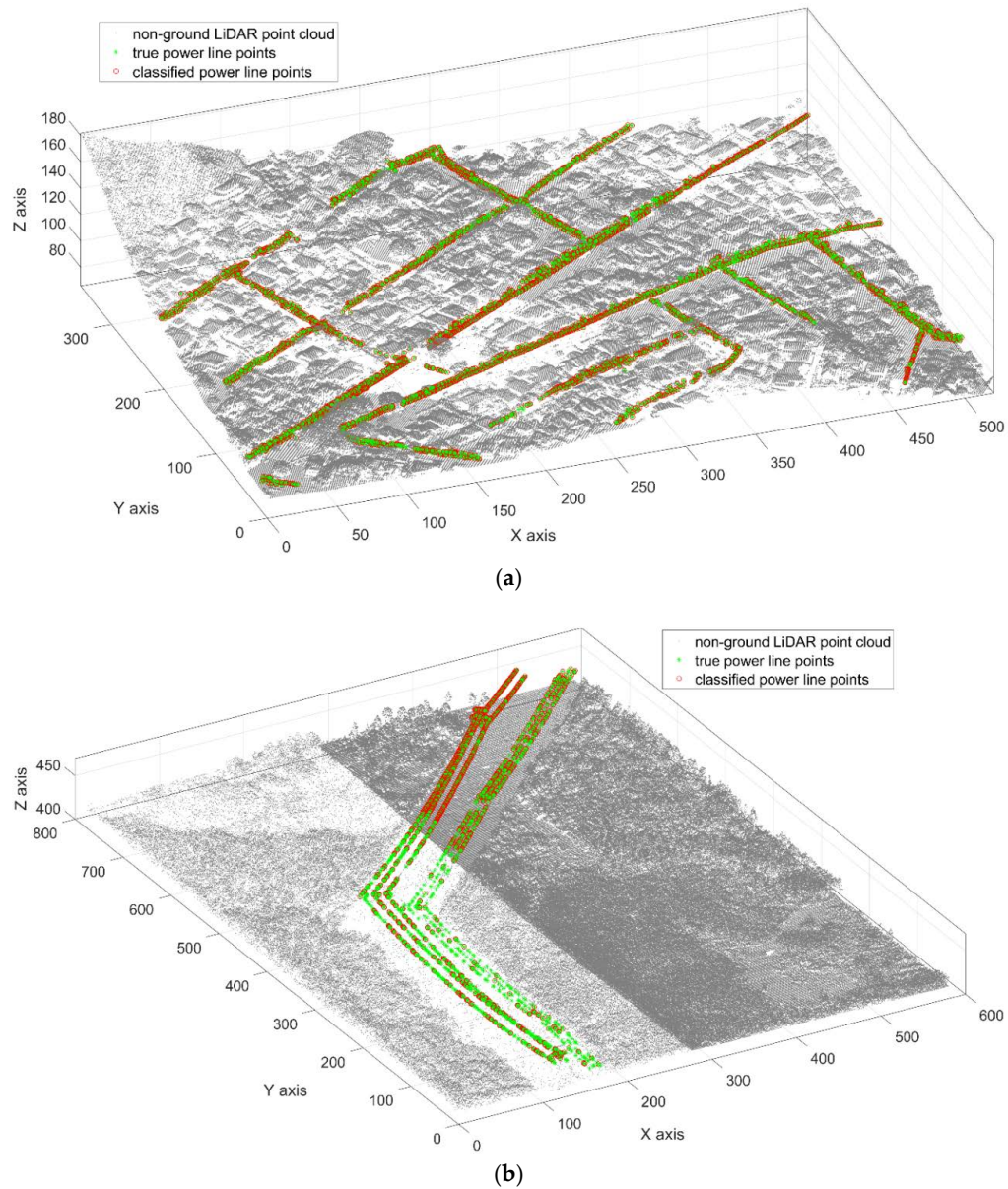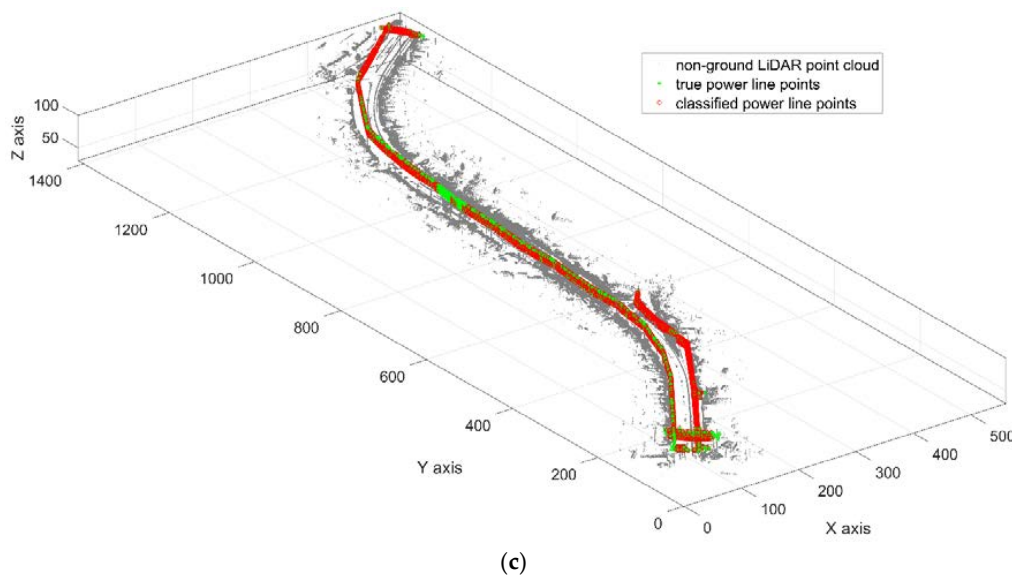

(**a**)

(**b**)

Figure 3. *Cont.*

(**c**)

Figure 3. Visualization of the experiment results of power line classification for the five light detection and ranging (LiDAR) data sites, which involve the non-ground LiDAR point cloud, true power line points, and classified power line points. (**a**) is the power line classification result of data site II, which is a complex urban scene and ALS point cloud; (**b**) is the power line classification result of data site III, which is a forest scene and ALS point cloud; and (**c**) is the power line classification result of data site V, which is a suburban scene and MLS point cloud.

4. Discussion

The aim of this study was to systematically compare the classification methods of power lines from ALS and MLS point cloud data through comparisons between neighborhood types, classifiers, and feature sets. We applied different methodology parameters of proposed classification architecture in the experiment data sites, involving the low density of ALS data and high density of MLS data, as well as forest, suburban, and urban scenes.

4.1. Sensitivity Analysis of Local Neighborhood

The local neighborhood construction of the considered points is an important task for pointwise objects classification from LiDAR point cloud data. Based on the constructed neighborhood, the extraction of crucial features, determination of suitable classifiers, and consequent processing can be realized. The neighborhood types and scales are the two key elements for the local neighborhood. According to the previous studies [9,12,42,59,60], the commonly used neighborhood types are spherical, vertical cylindrical, k-nearest, optimal k-nearest, and slant cylindrical neighborhoods, while the neighborhood scale could be usually divided into single scale and multi-scale types.

Considering the neighborhood types, we found that the four common used ones (spherical, vertical cylindrical, k-nearest, and optimal k-nearest) were widely applied in the classification of different geographical objects (ground, trees, buildings, cars, etc.). The optimal k-nearest type might achieve higher precision results for the classification of special objects. However, it is more time consuming to compute the optimized k for each point, especially when the data volume is very big in large areas or MLS datasets. The experiment in Section 3.1 also revealed that the results of the optimal k-nearest neighborhood type had lower accuracy rates than the others. Therefore, the optimal k-nearest neighborhood type is unacceptable for application in the power line classification. The slant cylindrical neighborhood is more tailored for power line classification, but it needs the power line corridor direction information beforehand. Therefore, in this study, we considered the spherical, vertical cylindrical, k-nearest, and optimal k-nearest neighborhoods in our experiments.

In the aspect of neighborhood scales, it's obvious that the multi-scale neighborhood has more benefits than the single scale neighborhood, because the former one represents more spatial structure characteristics between points. The scale of neighborhoods is commonly identified by the radius parameter, such as 1 m, 3 m, etc. We applied multi-scale neighborhood in this study, which combined multiple single scales of 1 m, 3 m, 5 m, 7 m, 9 m, and 11 m. Obviously, the higher the number of single scales in the multi-scale neighborhood, the more time-consuming the feature extraction and consequent processing.

There should be a balance between the number of scales and the amount of time required. The larger number of single scales covers a wider power line corridor. Therefore, the multi-scale neighborhood could capture more useful physical information of the power lines.

4.2. Effects of Different Classifiers

We compared six common classifiers for the power line classification from LiDAR point cloud data in this study. These classifiers contain SVM, RF, DT, NB, DA, and NN, which are popular in the classification of geographical objects from remote sensing images or laser scanning point cloud [16,26,27,33,42]. However, it is difficult to know whether these classifiers are suitable for power line classification or different types of point cloud data. Therefore, we applied six classifiers and five different data sites in our experiments, which could be more creditable in various situations.

By exploiting the experiment results, we found that the RF and NN classifiers could be better options for power line classification from ALS and MLS point cloud data. The accuracy was similar to each other for these two classifiers. Generally, the classifiers of RF, DT, and NN were more suitable for ALS data, while those of RF, DT, and SVM were better for MLS data. Moreover, the six compared classifiers performed consistently better for MLS point clouds (sites IV and V) than ALS ones (sites I, II, and III). The classification results of these six classifiers had slight differences for MLS point clouds, but had noticeable differences for the ALS point clouds. The main difference between ALS and MLS data is the point density. The point density of ALS was usually <10 points/m^2, while that of MLS could be up to 150 points/m^2. The local neighborhoods of points in MLS, which include more points, could provide more detailed spatial structure information. Therefore, it was revealed that: the higher the point density in the raw LiDAR point cloud, the better the classification result that could be achieved.

4.3. Differences between Selected Feature Sets

The selection of feature set in the power line classification is crucial to the processing time and performance of the proposed methodology. The previous studies [9,10,12,13] proposed up to 26 features for the classification of different objects from airborne LiDAR point cloud data. We concentrated these features as the initial whole feature set, as they were validated in the applications of building, tree, and car classification. Based on these features, we obtained two other feature sets by applying the PCA method and manual selection. According to the physical distribution characteristics of power lines, we selected some core features that could better represent the contextual information.

The experiments revealed that the selected core feature set could achieve accuracy that was close to the whole feature set for MLS datasets and had a slightly lower accuracy for the ALS datasets. Meanwhile, it decreased the processing time dramatically from 878 s to 387 s for these data sites on average. Such core features are useful when computation speed is critical (e.g., processing large-area data with limited computer resources) and for users who can slightly sacrifice accuracy.

5. Conclusions

In this study, we systematically compared the power line classification methods for ALS and MLS point cloud data. Based on the various parameters of the model, we specially focused on the result comparison of different local neighborhood types, classifiers, and selected feature sets. Through comparison analysis, we provided a common simple and validated framework of power

line classification method for different types of point cloud, variable geographical scenes, and point density. We found that the classification method composed with the multi-scale vertical cylindrical neighborhood, the RF classifier, and selected core feature sets could be an optimal solution that can balance both classification accuracy and processing time. The high point density of MLS could achieve higher classification accuracy than ALS. On the other hand, the methods in this study almost belonged to supervised classification algorithms. In the future work, we should develop an efficient unsupervised power line classification method from LiDAR point cloud in complex scenes.

Author Contributions: Y.W. and Q.C. together designed the research and methods and wrote the code. Y.W. conducted the analysis and wrote the manuscript. Q.C. provided the LiDAR point cloud and assisted the manuscript writing. L.L. assisted refining the research design and manuscript writing. X.L. and A.K.S. assisted processing the LiDAR data and methods design. K.L. assisted processing LiDAR data and results interpretation.

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

References

1. Ahmad, J.; Malik, A.S.; Xia, L.; Ashikin, N. Vegetation encroachment monitoring for transmission lines right-of-ways: A survey. *Electr. Power Syst. Res.* **2013**, *95*, 339–352. [CrossRef]
2. Matikainen, L.; Lehtomäki, M.; Ahokas, E.; Hyyppä, J.; Karjalainen, M.; Jaakkola, A.; Kukko, A.; Heinonen, T. Remote sensing methods for power line corridor surveys. *ISPRS J. Photogramm. Remote Sens.* **2016**, *119*, 10–31. [CrossRef]
3. Glennie, C.L.; Carter, W.E.; Shrestha, R.L.; Dietrich, W.E. Geodetic imaging with airborne lidar: The earth's surface revealed. *Rep. Prog. Phys. Phys. Soc.* **2013**, *76*, 8. [CrossRef] [PubMed]
4. McManamon, P.F. Review of ladar: A historic, yet emerging, sensor technology with rich phenomenology. *Opt. Eng.* **2012**, *51*, 060901. [CrossRef]
5. Zhang, Y.; Yuan, X.; Fang, Y.; Chen, S. Uav low altitude photogrammetry for power line inspection. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 14. [CrossRef]
6. Kwoczyńska, B.; Dobek, J. Elaboration of the 3D model and survey of the power lines using data from airborne laser scanning. *J. Ecol. Eng.* **2016**, *17*, 65–74. [CrossRef]
7. Popovic, D.; Pajic, V.; Jovanovic, D.; Sabo, F.; Radovic, J. Semi-automatic classification of power lines by using airborne lidar. In *FIG Working Week 2017, Surveying the World of Tomorrow—From Digitalisation to Augmented Reality*; International Federation of Surveyors FIG: Helsinki, Finland, 2017.
8. Cheng, L.; Tong, L.; Wang, Y.; Li, M. Extraction of urban power lines from vehicle-borne lidar data. *Remote Sens.* **2014**, *6*, 3302–3320. [CrossRef]
9. Blomley, R.; Jutzi, B.; Weinmann, M. Classification of airborne laser scanning data using geometric multi-scale features and different neighbourhood types. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *III-3*, 169–176. [CrossRef]
10. 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]
11. Weinmann, M.; Schmidt, A.; Mallet, C.; Hinz, S.; Rottensteiner, F.; Jutzi, B. Contextual classification of point cloud data by exploiting individual 3D neigbourhoods. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *II-3/W4*, 271–278. [CrossRef]
12. 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]
13. Yang, B.; Huang, R.; Li, J.; Tian, M.; Dai, W.; Zhong, R. Automated reconstruction of building lods from airborne lidar point clouds using an improved morphological scale space. *Remote Sens.* **2017**, *9*, 14. [CrossRef]

14. Kim, H.B.; Sohn, G. Point-based classification of power line corridor scene using random forests. *Photogramm. Eng. Remote Sens.* **2013**, *79*, 821–833. [CrossRef]

15. Guo, B.; Li, Q.; Huang, X.; Wang, C. An improved method for power-line reconstruction from point cloud data. *Remote Sens.* **2016**, *8*, 36. [CrossRef]

16. 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]

17. Jwa, Y.; Sohn, G. A piecewise catenary curve model growing for 3D power line reconstruction. *Photogramm. Eng. Remote Sens.* **2012**, *78*, 1227–1240. [CrossRef]

18. Jwa, Y.; Sohn, G.; Kim, H.B. Automatic 3d powerline reconstruction using airborne lidar data. In Proceedings of the IAPRS Laser Scanning 2009, Paris, France, 1–2 September 2009; Bretar, F., Pierrot-Deseilligny, M., Vosselman, G., Eds.; IAPRS: Paris, France, 2009; Volume XXXVIII, pp. 105–110.

19. Liang, J.; Zhang, J.; Deng, K.; Liu, Z. A New Power-Line Extraction Method Based On Airborne Lidar Point Cloud Data. In Proceedings of the International Symposium on Image and Data Fusion, Tengchong, China, 9–11 August 2011; IEEE: Tengchong, China, 2011; pp. 1–4.

20. Ritter, M.; Benger, W. Reconstructing Power Cables from Lidar Data Using Eigenvector Streamlines of the Point Distribution Tensor Field. In Proceedings of the WSCG 2012—20th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, Plzen, Czech Republic, 25–28 June 2012.

21. 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]

22. Stal, C.; Briese, C.; De Maeyer, P.; Dorninger, P.; Nuttens, T.; Pfeifer, N.; De Wulf, A. Classification of airborne laser scanning point clouds based on binomial logistic regression analysis. *Int. J. Remote Sens.* **2014**, *35*, 3219–3236. [CrossRef]

23. Zhou, G.; Zhou, X. Seamless fusion of lidar and aerial imagery for building extraction. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 7393–7407. [CrossRef]

24. Ramiya, A.M.; Nidamanuri, R.R.; Krishnan, R. Object-oriented semantic labelling of spectral–spatial lidar point cloud for urban land cover classification and buildings detection. *Geocarto Int.* **2016**, *31*, 121–139. [CrossRef]

25. Zhang, Z.; Zhang, L.; Tong, X.; Mathiopoulos, P.T.; Guo, B.; Huang, X.; Wang, Z.; Wang, Y. A multilevel point-cluster-based discriminative feature for als point cloud classification. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 3309–3321. [CrossRef]

26. Zhang, J.; Lin, X.; Ning, X. Svm-based classification of segmented airborne lidar point clouds in urban areas. *Remote Sens.* **2013**, *5*, 3749–3775. [CrossRef]

27. Xu, S.; Vosselman, G.; Oude, E.S. Multiple-entity based classification of airborne laser scanning data in urban areas. *ISPRS J. Photogramm. Remote Sens.* **2014**, *88*, 1–15. [CrossRef]

28. Dalponte, M.; Ene, L.T.; Marconcini, M.; Gobakken, T.; Næsset, E. Semi-supervised svm for individual tree crown species classification. *ISPRS J. Photogramm. Remote Sens.* **2015**, *110*, 77–87. [CrossRef]

29. Chehata, N.; Guo, L.; Mallet, C. Airborne lidar feature selection for urban classification using random forests. In Proceedings of the IAPRS Laser Scanning 2009, Paris, France, 1–2 September 2009; Bretar, F., Pierrot-Deseilligny, M., Vosselman, G., Eds.; IAPRS: Paris, France, 2009; Volume XXXVIII, pp. 207–212.

30. Belgiu, M.; Drăguţ, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [CrossRef]

31. 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]

32. Ni, H.; Lin, X.; Zhang, J. Classification of als point cloud with improved point cloud segmentation and random forests. *Remote Sens.* **2017**, *9*, 288. [CrossRef]

33. Kang, Z.; Yang, J.; Zhong, R. A bayesian-network-based classification method integrating airborne lidar data with optical images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 1651–1661. [CrossRef]

34. Xu, Y.; Yao, W.; Hoegner, L.; Stilla, U. Segmentation of building roofs from airborne lidar point clouds using robust voxel-based region growing. *Remote Sens. Lett.* **2017**, *8*, 1062–1071. [CrossRef]

35. Guan, H.; Yu, Y.; Li, J.; Ji, Z.; Zhang, Q. Extraction of power-transmission lines from vehicle-borne lidar data. *Int. J. Remote Sens.* **2016**, *37*, 229–247. [CrossRef]

36. Xu, K.; Zhang, X.; Chen, Z.; Wu, W.; Li, T. Risk assessment for wildfire occurrence in high-voltage power line corridors by using remote-sensing techniques: A case study in hubei province, china. *Int. J. Remote Sens.* **2016**, *37*, 4818–4837. [CrossRef]

37. Mongus, D.; Lukač, N.; Žalik, B. Ground and building extraction from lidar data based on differential morphological profiles and locally fitted surfaces. *ISPRS J. Photogramm. Remote Sens.* **2014**, *93*, 145–156. [CrossRef]

38. Yan, L.; Liu, H.; Tan, J.; Li, Z.; Chen, C. A multi-constraint combined method for ground surface point filtering from mobile lidar point clouds. *Remote Sens.* **2017**, *9*, 958. [CrossRef]

39. Meng, X.; Currit, N.; Zhao, K. Ground filtering algorithms for airborne lidar data: A review of critical issues. *Remote Sens.* **2010**, *2*, 833–860. [CrossRef]

40. Chen, Q. Improvement of the edge-based morphological (em) method for lidar data filtering. *Int. J. Remote Sens.* **2009**, *30*, 1069–1074. [CrossRef]

41. Zhu, L.; Hyyppä, J. Fully-automated power line extraction from airborne laser scanning point clouds in forest areas. *Remote Sens.* **2014**, *6*, 11267–11282. [CrossRef]

42. Wang, Y.; Chen, Q.; Liu, L.; Zheng, D.; Li, C.; Li, K. Supervised classification of power lines from airborne lidar data in urban areas. *Remote Sens.* **2017**, *9*, 771. [CrossRef]

43. Hackel, T.; Wegner, J.D.; Schindler, K. Fast semantic segmentation of 3d point clouds with strongly varying density. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *III-3*, 177–184. [CrossRef]

44. Snoek, J.; Larochelle, H.; Adams, R.P. Practical bayesian optimization of machine learning algorithms. In *Advances in Neural Information Processing Systems 2012, Proceedings of the Neural Information Processing Systems Conference, Stateline, NV, USA, 3–8 December 2012*; NIPS 2012: Stateline, NV, USA, 2012; pp. 2951–2959.

45. Sun, X.; Lin, X.; Shen, S.; Hu, Z. 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]

46. Probst, P.; Boulesteix, A.-L. To tune or not to tune the number of trees in random forest? *arXiv* **2017**.

47. Li, M.; Ma, L.; Blaschke, T.; Cheng, L.; Tiede, D. A systematic comparison of different object-based classification techniques using high spatial resolution imagery in agricultural environments. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *49*, 87–98. [CrossRef]

48. Reinartz, P.; Samadzadegan, F.; Abdi, G. Deep learning decision fusion for the classification of urban remote sensing data. *J. Appl. Remote Sens.* **2018**, *12*, 016038.

49. Bigdeli, B.; Pahlavani, P. High resolution multisensor fusion of sar, optical and lidar data based on crisp vs. Fuzzy and feature vs. Decision ensemble systems. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 126–136. [CrossRef]

50. Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. *Classification and Regression Trees*; CRC Press: Boca Raton, FL, USA, 1984.

51. Thompson, D.R.; Hochberg, E.J.; Asner, G.P.; Green, R.O.; Knapp, D.E.; Gao, B.-C.; Garcia, R.; Gierach, M.; Lee, Z.; Maritorena, S.; et al. Airborne mapping of benthic reflectance spectra with bayesian linear mixtures. *Remote Sens. Environ.* **2017**, *200*, 18–30. [CrossRef]

52. Alonso-Montesinos, J.; Martínez-Durbán, M.; del Sagrado, J.; del Águila, I.M.; Batlles, F.J. The application of bayesian network classifiers to cloud classification in satellite images. *Renew. Energy* **2016**, *97*, 155–161. [CrossRef]

53. Zhang, X.; Wang, Q.; Chen, G.; Dai, F.; Zhu, K.; Gong, Y.; Xie, Y. An object-based supervised classification framework for very-high-resolution remote sensing images using convolutional neural networks. *Remote Sens. Lett.* **2018**, *9*, 373–382. [CrossRef]

54. 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]

55. 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]

56. Zhao, W.; Du, S.; Emery, W.J. Object-based convolutional neural network for high-resolution imagery classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 3386–3396. [CrossRef]

57. Hornik, K. Approximation capabilities of multilayer feedforward networks. *Neural Netw.* **1991**, *4*, 251–257. [CrossRef]

58. Hinton, G.E.; Osindero, S.; Teh, Y.-W. A fast learning algorithm for deep belief nets. *Neural Comput.* **2006**, *18*, 1527–1554. [CrossRef] [PubMed]

59. Blomley, R.; Weinmann, M. Using multi-scale features for the 3d semantic labeling of airborne laser scanning data. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *IV-2/W4*, 43–50. [CrossRef]

60. 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]