ROBUST AND EFFECTIVE AIRBORNE LIDAR POINT CLOUD CLASSIFICATION BASED ON HYBRID FEATURES

L. F. Liao[a] , S. J. Tang[a],*, J. H. Liao[a], W. X. Wang[a], X. M. Li[a], R. Z. Guo[a]

[a] Research Institute for Smart Cities, School of Architecture and Urban Planning & Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Natural Resources, Shenzhen University, Shenzhen, PR China - (2018104003@email,shengjuntang,liaojh@szu.edu.cn,wangwx@szu.edu.cn ,lixming@szu.edu.cn,guorz)@szu.edu.cn

Commission II, WG II/3

KEY WORDS: Point cloud Classification, Supervoxel, Random Forests, Feature fusion, Segmentation

ABSTRACT:

State-of-the-art point cloud classification methods mostly process raw point clouds, using a single point as the basic unit and calculating point cloud features by searching local neighbors via the k-neighborhood method. Such methods tend to be computationally inefficient and have difficulty obtaining accurate feature descriptions due to inappropriate neighborhood selection. In this paper, we propose a robust and effective point cloud classification approach that integrates point cloud supervoxels and their locally convex connected patches into a random forest classifier. We apply a centroid cloud extracted from supervoxels into the proposed classifier, which effectively improves the point cloud feature calculation accuracy and reduces the computational cost. Considering the different types of point cloud feature descriptions, we divide features into three categories (point-based, eigen-based, and grid-based) and accordingly design three distinct feature calculation strategies to improve feature reliability. The proposed method achieves state-of-the-art performance, with average F1-scores of 89.16%, respectively. The successful classification of point clouds with great variation in elevation also demonstrates the reliability of the proposed method in challenging scenes to some extents.

1. INTRODUCTION

With the development of photogrammetry and light detection and ranging (LiDAR) technologies, urban three-dimensional (3D) point clouds can be easily obtained. 3D point cloud data are used in many applications, such as power line inspection (Chen et al., 2017), and urban 3D modeling (Croce et al., 2021, Javernick et al., 2014), unmanned vehicles (Yue et al., 2018). However, the most basic requirement for these applications is the semantic classification of 3D point cloud data, which has been a research focus among photogrammetry and remote sensing communities.

Early classification efforts mainly focused on extracting low-level geometric primitives, such as point features, line features, and surface features, which were used for surface reconstruction or point cloud alignment. In recent years, researchers have developed methods for extracting high-level semantic features for structure model reconstruction from point cloud data through machine learning-and deep learning-based methods (Lafarge and Mallet, 2012, Xiong et al., 2015, Zhou and Neumann, 2010).The core challenges of point cloud data classification are extracting discriminative features from neighborhoods and constructing point cloud classifiers (Hackel et al., 2016, Jie and Zulong, 2014). Accurate classification depends on a combination of robust point cloud features and proper classifiers (Hackel et al., 2016, Wang et al., 2018). Recent works have applied deep learning networks to directly learn per-point features from raw point clouds (Li et al., 2020, Qi et al., 2017a, Qi et al., 2017b). Similar to traditional machine learning, these methods focus on the extraction of higher-order features from point cloud data by building a new neural network. Although remarkable performance has been achieved using these methods, large

training sample sets are required to pre-train the classification models. These semantic tags require manual labeling, which is time-consuming and labor-intensive. Moreover, the training models obtained by such methods are difficult to generalize to other scenarios (Li et al., 2021).To solve the model generalization and incomplete label data problems, many researchers prefer traditional machine learning methods, which require only a small sample dataset to achieve fast and accurate semantic point cloud data classification (Niemeyer et al., 2014, Niemeyer et al., 2016, Zhu et al., 2017). However, original point cloud features are often highly unstable due to the influence of point cloud data accuracy and noise, especially data acquired by tilt photogrammetry. Thus, more researchers are exploiting high-order features and their contextual information for scene classification. As dimensional objects expanding upon the concept of the "superpixel" (Achanta et al., 2012), "supervoxel"(Papon et al., 2013), are generated by partitioning 3D space as point clusters. Supervoxels have been increasingly applied to describe adjoining points related to the same objects (Wu et al., 2016, Zhu et al., 2017). Transferring the original point cloud to the "supervoxel cloud" propagates simple point-based classification to an object-based level. Some point cloud segmentation methods, such as locally convex connected patches (LCCP), recognize points through supervoxel-adjacent relationships. In addition to features, classifiers that can effectively deal with massive data must be considered. Machine learning methods such as random forest (RF) that are capable of handling complex data are gaining attention for this purpose (Breiman, 2001, Ni et al., 2017).

However, most existing model-driven methods based on supervoxel extraction are prone to include object boundaries in the local neighborhood of voxels, which decreases the homogeneity of supervoxel adjacency and polygonal feature accuracy. Therefore, combining a precise object segmentation util-

---

* Corresponding author. Shengjun Tang, associate researcher, shengjuntang@szu.edu.cn

ity with previous model-driven methods will effectively solve this problem. Object edges can be detected by particular network structures or LCCP(Christoph Stein et al., 2014). Feng et al. (2020) developed a local attention-edge convolutional network that identifies objects by summarizing the features of all neighbors as a weight value learned by the network (Feng et al., 2020) . The LCCP examines the connection between two adjacent supervoxels and determines whether they relate to one object by calculating the included angle of two normal vectors. The former method focuses on whole object segmentation, whereas the latter recognizes as many connected edges as possible. To better exploit supervoxel features and their contextual relationships for point cloud classification, we propose a robust and effective classification approach that integrates point cloud supervoxels and their LCCP relations into an RF classifier to improve the accuracy of feature calculation and reduce computational costs. The proposed method involves three strategies to effectively improve classification accuracy.

(1) Features are divided into three categories based on their description types (point-based, eigen-based, and grid-based), and three unique feature calculation strategies are designed to improve feature reliability.

(2) A centroid point is used to represent supervoxel geometries and every point that belongs to the same cluster shares all properties.

(3) Supervoxel local neighborhoods are segmented by LCCP to avoid the inclusion of object borders.

The rest of this paper is organized into four sections. In Section 2, we presents the framework of the proposed supervoxel-based RF model, providing the feature descriptions and RF model process and algorithm. The statistical and visual results of data training and validation are shown in Section 3, and our research conclusion and remarks are given in Section 4.

2. METHODOLOGY

2.1 Overview of the approach

The approach starts with a voxel-grid-based downsampling algorithm (Rusu and Cousins, 2011) to prevent the point cloud from becoming over-dense without impacting the original structure. Next, a noise-rejection statistical-outlier-removal filter is used to remove dynamic objects and erroneous points from the aerial laser point cloud. The threshold is calculated from the average distance between a single point and its k-neighbors and the same multiplied standard deviation.

The technical route for our approach after data pre-processing is shown in Figure 1.The features are divided into three categories, point-based, eigen-based, and grid-based. First, the original 3D point cloud is transformed into a set of supervoxels by the supervoxel calculation method, in which points located in the same supervoxel generally have similar feature descriptions. At the same time, The original point cloud is also divided using a regular grid to facilitate the extraction of grid-based elevation features in the later stage.Instead of semantic labeling of the raw points, supervoxels are used as the basic unit for semantic classification, and the centroids of the supervoxels are generated from the supervoxel structure. Three kinds of features are calculated: (1) The eigen-based features are first calculated using a principal component analysis algorithm, and the corresponding geometric shape features are generated by deformation and combination with those eigenvalues. Specific-

ally, the adjacency relationship built by voxel cloud connectivity segmentation (VCCS) is used to determine the supervoxel neighborhood ranges. (2) The point-based features, including the local density, point feature histogram, point's normal vectors, and elevation values, are obtained via neighborhood calculation or the point cloud's raw attributes. (3) We introduce a grid-based elevation feature to decrease the influence of uneven topography during point cloud classification. Based on the regularized grid of the point cloud data, the relative elevation of the horizontal location is used as the elevation feature of each supervoxel centroid. Finally, all three feature types are used to train the supervoxel-based RF model, which is used for point cloud classification.
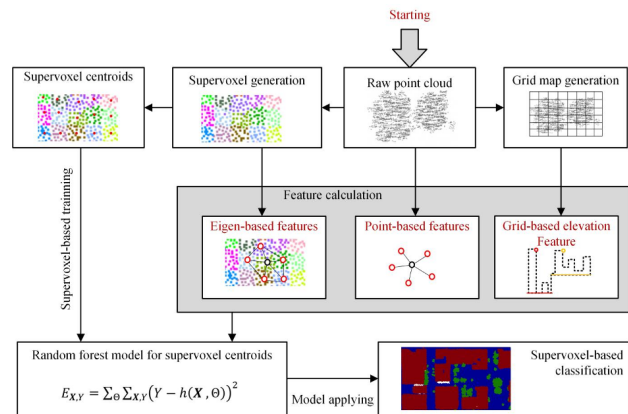

Figure 1. Supervoxel-based random forests framework for point cloud classification

2.2 Two-level graphical model generation for feature extraction

Supervoxels are defined as groups of points that contain similar geometric features or attributes, such as location, color, and normal direction. Additionally, adjacency relationships embedded in supervoxels can provide more effective information for neighborhood searching, improving the robustness and accuracy of feature calculation. For this classification method, we use supervoxels, rather than single points, as the basic unit to construct the RF model, and the domain information is constrained via LCCP segmentation. Therefore, a two-level graphical model using supervoxel calculation and LCCP optimization is generated from the raw point cloud. Figure 2 illustrates the two-level graphical model generation process. First, we generate the supervoxel model in two steps, namely, randomly seeding the point cloud and clustering by calculating the feature distances among neighboring points. The supervoxel clustering algorithm estimates the point homogeneity via color, space, and normal dimensions as in Equation 1, 2 and 3.

$$d = i_{space} \times d_{space} + i_{normal} \times d_{normal} \qquad (1)$$

$$d_{space} = \frac{\sqrt{\triangle x^2 + \triangle y^2 + \triangle z^2}}{r_{voxel}}(2)$$

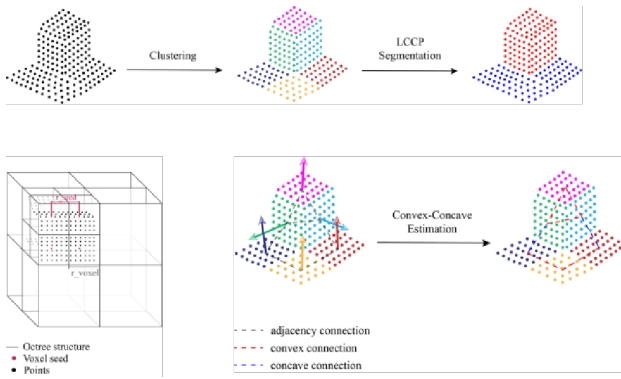$$d_{normal} = \frac{\vec{v_1} \cdot \vec{v_2}}{|\vec{v_1}| \cdot |\vec{v_2}|}(3)$$

Figure 2. Illustration of two-level graphical model generation

where d represents the summarized estimation of homogeneity across all dimensions, d_{space} and d_{normal} represent thte Euclidean distances between seeds and surrounding points and the normal vector directions of a plane approximated using the least-squares method with neighbors within a certain number (set to 15 in the paper balancing the effect and computational costs). The importance weights i_{space} and i_{normal} are set to 0.4 and 0.6 considering optimal homogeneity of a single voxel. r_{voxel} is the size of each supervoxel (set to 1m in this paper according to the point cloud scale), v_1 and v_2 and are normal vectors of pairwise adjacent supervoxels. Then, the adjacency relationship built by VCCS is used to determine the supervoxel neighborhood ranges to calculate local eigen features. To ensure the neighborhood search for eigen feature calculation is accurate, we generate the second-level graphical model using an LCCP algorithm based on the first-level graphical model obtained by supervoxel clustering. The core of the LCCP segmentation is to accurately identify the edges of objects by the angular relationship between supervoxels. Once the adjacency between two supervoxels is confirmed, the relationship properties can be analyzed via normal vector calculation, as shown in Equation 4 and 5.

$$\hat{d} = \frac{\vec{x_1} - \vec{x_2}}{||\vec{x_1} - \vec{x_2}||} \qquad (4)$$

$$\triangle\alpha = \vec{n_1} \cdot \hat{d} - \vec{n_2} \cdot \hat{d} \qquad (5)$$

wherein $\vec{x_1}$ and $\vec{x_2}$ refer to the centroid positions of the two observed supervoxels, and, $\vec{n_1}$ and $\vec{n_2}$ represent their normal vectors. The relationship is considered a convex connection if $\triangle\alpha > 0$, which indicates that the normal vector of the central supervoxel has a minor angle to the straight line defined by $\vec{x_1} - \vec{x_2}$ and a greater angle cosine value. Alternatively, the relationship is considered a concave connection when $\triangle\alpha < 0$ in a similar way.

2.3 Description for hybrid features

Three types of features are used in our proposed RF model, which contains point-based features, Eigen-based features and grid-based elevation features. The descriptions for them are as following.

Point-based features description: (1)Local density of points: Measured as the average distance from one point to the nearest k-neighbors. Each measurement constructs a temporary kdtree structure using the fast library for approximate nearest neighbors (FLANN) (Muja and Lowe, 2014) to search for the nearest

centroid points in the input point cloud. The Euclidean distances for pairwise vertices are then calculated, and their average values are used as a feature. (2) Point Feature Histogram (PFH): A descriptor computed by the consistency of the normal vectors of adjacent points (Rusu et al., 2008a, Rusu et al., 2008b). The PFH computation algorithm uses the kdtree to search for available surrounding points and compare their coordinates and normal vectors. An extracted centroid point uses the normal vector of the related supervoxel as a property. (3) The included angle to the horizontal plane of the normal vectors of points: Measured from the coordinates of the 3D normal vectors of points. The formula for calculating the included angle cosine is as follows.

$$c = \frac{n_1 \cdot n_2}{|n_1| \cdot |n_2|} = \frac{z_1}{\sqrt{x_1{}^2 + y_1{}^2 + z_1{}^2}} \qquad (6)$$

where C refers to the cosine value and n_1 and n_2 are the normal vectors of the point and the horizontal plane (defined as (0,0,1)), respectively. The cosine value is applied as a feature instead of an actual degree for feature normalization. (4) Point elevation value: Obtained and transferred from the definite coordinate values of points in the z-axis. For urban scenes with uncertain ground elevation, we propose a grid-based feature optimization strategy to eliminate the effects of terrain undulations. (5) RGB color: Color information can achieve effective judgment of feature types, and this paper uses color features as a basic feature of supervoxels. Considering that this paper uses supervoxels as the basic unit for feature classification experiments, their color features are determined by the average value of points inside the supervoxels.

Eigen-based features description: Different combinations of eigenvalues demonstrate particular shape characteristics (Wang et al., 2018). We apply five typical eigen geometric shape features as inputs to include supervoxel adjacency results in the approach; the computing formulas are shown in Table 1. The adjacency map and LCCP method are used to optimize the neighborhood homogeneity, as described above. (1) Curvature: Describes the extent of the curve for a point group. (2) Linearity: Describes the extent of the line-like shape for a point group. (3) Planarity: Describes the extent of the plane-like shape for a point group. (4) Scattering: Describes the extent of the sphere-like shape for a point group. (5) Anisotropy: Describes the difference between the extents of entropy in respective directions of eigenvectors for a point group.

Feature definition	Computing formula
Curvature	$C_e = \frac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3}$
Linearity	$L_e = \frac{\lambda_1-\lambda_2}{\lambda_1}$
Planarity	$P_e = \frac{\lambda_2-\lambda_3}{\lambda_1}$
Scattering	$S_e = \frac{\lambda_1}{\lambda_3}$
Anisotropy	$A_e = \frac{\lambda_1-\lambda_3}{\lambda_1}$

Table 1. Computing method for eigen value-based shape features. Feature definitions on the left refer to explanation above. Three eigen value symbols are sorted in descending order from 1 to 3 in the formulas

Grid-based elevation feature description: Different elements in urban scenes are mostly vertically distributed; thus, adding z-axis values of vertices as a feature type helps distinguish various objects. However, due to the uncertain elevation of urban terrestrial surfaces, a definite elevation value cannot raise the

discriminability of objects because the homogeneous points are distributed in an unclear range, causing false classification to ground points. Likewise, using the difference between the overall minimum elevation and the value of certain points is not a viable solution. To solve this problem, we apply a grid-based elevation differences system to calculate the elevation feature. As shown in Figure 3(a) and (b), the system projects the entire point cloud onto the plane in two dimensions (x and y), then segregates all points into a certain number of grid squares (a 100×100 grid as the preset minimizing elevation fluctuation in a single square and computational costs). The network range is constrained by maximum and minimum values in the x and y dimensions of all vertices obtained in advance to ensure every point is located in a unique square. The minimum z-values compared in each grid square are used to calculate the difference values, as the fluctuation extent of ground elevation is predictable in minor areas with hardly any steep slopes.
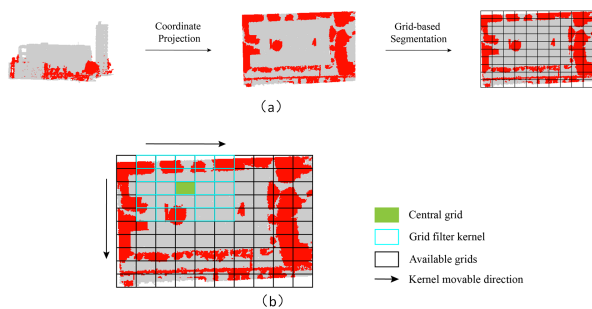

Figure 3. Grid-based elevation computation and filtering. (a) the illustrated point cloud data is on the left and the 2D-projected data with grid meshes is on the right. (b) grid filter examining anomaly of calculated elevation value in grids

2.4 Super-voxel based random forests classifier model

Supervoxel based RF classifier model The RF algorithm relies on features extracted from the original point cloud to generate decision trees from randomly selected point inputs for classification. All above-computed features are sent into the RF construction system with a segmented training set of points. As an algorithm with a random feature selection strategy, the system arbitrarily draws a subset from the original training dataset and grows a new decision tree from the extracted set, which allows the RF method to efficiently handle large-scale datasets (Breiman, 2001). Two thresholds limit the growth of the forest: the max depth and the total number of decision trees (set to 25 and 10, respectively, in this paper, balancing efficiency and result quality). The growth of trees ceases when the preset thresholds are reached, and the output forest classifies the validation set using corresponding features to verify accuracy. The algorithm applies the mean-squared generalization error to evaluate the classification correctness (Breiman, 2001), as follows.

$$E_{X,Y} = \sum_{\theta} \sum_{X,Y} (Y - h(X,\theta))^2 \qquad (7)$$

where X refers to the random feature vector and Y refers to the corresponding label. h means a single tree inside the forest, appearing in tandem with one X.

3. EXPERIMENTAL RESULTS

Data preparation: The performance of the proposed method is verified using ISPRS benchmark dataset. The ISPRS benchmark datasets collected in Toronto, Canada provided by ISPRS benchmark (Rottensteiner et al., 2012) is used. We select three indices considered effective in previous approaches, the overall accuracy (OA), the mean intersection over union (mIoU), and the F1-score, the values of which are compared with other methods using the same datasets from ISPRS benchmarks and computed as follows. True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN) values are extracted from the confusion matrix of the classification result, p and r are the precision and recall percentages, respectively.

$$OA = \frac{TP}{TP + TN + FP + FN} \qquad (8)$$

$$mIOU = \frac{TP}{TP + 2 \times (TN + FP + FN)} \qquad (9)$$

$$p = \frac{TP}{TP + FP} \qquad (10)$$

$$r = \frac{TP}{TP + FN} \qquad (11)$$

$$F_1 = \frac{2 \times p \times r}{p + r} \qquad (12)$$

Experimental results of Toronto sites: The Toronto dataset is divided into two regions including Area 1 and Area 2. The classification results are visualized in Figure 4. because few instances have been tested on the features of the Toronto sites, comparisons with five similar approaches are given in Table 2. There are plentiful high buildings and vegetation in miniature in the Toronto scenes, plus relatively sparse point cloud data and uncompleted construction facades, which increase the difficulty of distinguishing categories. Moreover, as a dataset generated by an airborne laser scanner, it lacks the RGB band information in the original point cloud data as properties, which disables three effective features in the proposed classifier. Nonetheless, the proposed method performs well in these situations with the advantages of optimized elevation and precisely segmented neighborhood eigen features, even without the integration of multispectral oblique images. As Figure 5 illustrates, the grid-based elevation feature corrects parts of the false building point labels recognized as misclassification in previous approaches (Zhu et al., 2017).
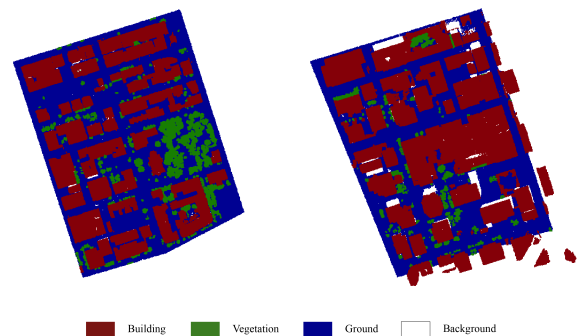

Figure 4. Classification results of two areas in Toronto sites

Methods	Area 1			Area 2		
	OA (%)	mIoU (%)	F1-score (%)	OA (%)	mIoU (%)	F1-score (%)
MAR_2	94.3	89.2	88.9	94.0	88.7	88.4
MSR	95.5	91.4	91.2	94.8	90.1	89.7
ITCM	81.3	68.5	66.1	83.0	70.9	67.9
ITCR	84.2	72.7	69.2	85.4	74.5	72.4
TUM	82.6	70.4	68.1	83.1	71.1	68.9
Ours	93.2	87.4	92.6	93.1	87.0	85.8

Table 2. Quantitative comparison of the proposed method and previous related methods testing on Toronto sites.
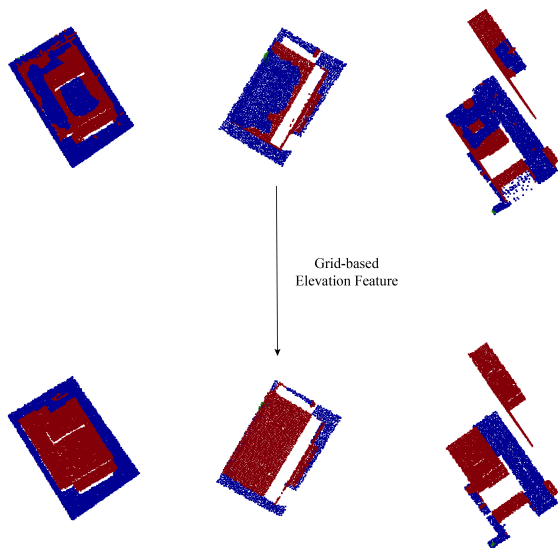
Figure 5. Corrections of grid-based elevation filter for constructions lacking corresponding ground points

Building   Vegetation   Ground   Background   Area of misclassification
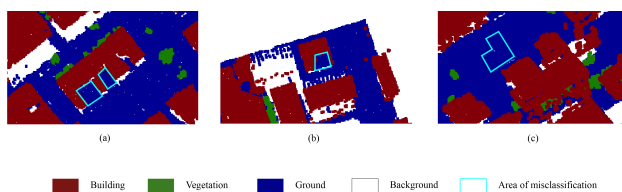
Figure 6. Misclassification cases which recognized roof as ground points in Toronto sites

The F1 scores shown in Table 2 indicate the performance of the proposed method overall six approaches. Even though compared to three indices, our method is one of the best in both Area 1 and 2. The majority of objects including constructions and vegetation clusters in vast scale are distinguished correctly viewed from Figure 4. However, similar shape features lead to some confusion. The most obvious misclassification part is a gym-like construction located at the right-bottom of Area 1, with a maximum height lower than most the vegetations, which pretends to be misclassified as ground. There are still parallel examples illustrated in Figure 6, in which we believe that it contains mainly roof points with few corresponding façade points relating to the same buildings, hence this is fault brought up by the dataset itself.

Experimental results of urban scenes in Shenzhen: The selected dataset included four urban regions, one for the training set and three for independent validation [marked as (a), (b), (c)]. The training area was 350 m × 200 m, and the validation areas were approximately 400 m × 300 m. The entire dataset was downsampled to a resolution of 0.3 m. The classification results are illustrated in Figure 7. Most vegetation points and ground points were accurately classified, and explicit outlines of buildings were visible in the resulting figure. In most scenes, vegetation was distinguished from adjacent buildings. Moreover, the centroid-based classification method enabled low computation costs, even though each validation area contained more than four million points after the downsampling process. This demonstrates that the proposed classifier successfully handles large datasets. The point-based classification method in CGAL library (Fabri and Pion, 2009) was used for comparison purpose. The quantitative performance evaluations of our proposed method and the pointbased method are shown in Table 3. As expected, the super voxel-based method proposed in this paper achieved better classification accuracy in all three regions compared to the traditional point cloud-based methods. Specifically, the proposed method achieved 3.6, 5.8, and 4.4 percent, respectively, in the OA, mIoU, and F1_score in Area (a). Similar results were found in the other two regions.

The average performance of the proposed method was higher for the Shenzhen dataset than the Vaihingen and Toronto datasets. The mostly rectangular rooftop shapes and integrated facade structures prevented building points from being recognized as vegetation, whereas the uncertainty of object consistency in the Vaihingen set led to false classification. Compared with the Toronto sites, which were comparably generated except without color information, most elevated vegetation points and buildings with low height and more detailed facades were successfully distinguished using RGB color features in the Shenzhen dataset. However, some exceptional situations in the dataset affected the overall accuracy of the classification results. As shown in Figure 8a, the neighborhood information of partial rooftop points that were similar to roads, such as rises at the edge or street light posts, reduced the contextual consistency of the local region and affected the classification. Additionally, due to the intricate and uncertain shape appearances in modern urban scenes, a single training area provided limited polygonal examples. Parts of buildings with minor scale or unusual contours that were not provided in the training region were misclassified as ground pieces in the validation sets [Figure 8b], which reduced the overall classification accuracy.

Benefiting from supervoxel extraction processing, the point cloud of Shenzhen University can be rapidly aggregated into supervoxel structures, which effectively reduced the point cloud density and complexity. In turn, with supervoxels as the basic unit, the classification method proposed in this paper achieved point cloud classification with high efficiency, and the overall computation costs were about 1.5 h. Moreover, the utilization of LCCP object homogeneity segmentation in supervoxel-based neighborhoods contributed to the considerable classification precision with complete object surfaces consisting of point arrays, which advanced the object-based theory.

	Area (a)			Area (b)			Area (c)		
	OA (%)	mIoU (%)	F1-Score (%)	OA (%)	mIoU (%)	F1-Score (%)	OA (%)	mIoU (%)	F1-Score (%)
Our method	94.0	88.7	90.1	93.5	87.8	91.8	93.5	87.8	91.7
Point based	90.6	82.9	85.6	87.6	78.0	84.2	86.1	75.6	79.8

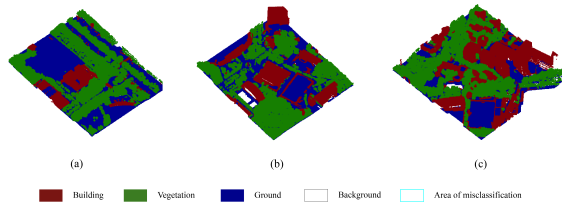Table 3. Quantitative evaluation of the supervoxel-based results and point-based results of the proposed method on the Shenzhen airborne LiDAR dataset.

Figure 7. Classification results of airborne LiDAR-generated Shenzhen sites. Three selected sites have been marked as (a), (b) and (c).
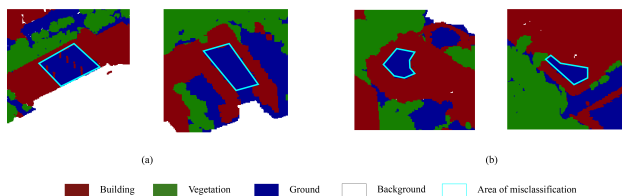

Figure 8. Misclassification cases in the Shenzhen dataset. (a) Faults due to edge interruption. (b) Faults due to untrained object shapes.

4. CONCLUSION

In this paper, we proposed a robust and effective airborne LiDAR point cloud classification method that integrates hybrid features, including point-based features, eigen-based features, and elevation-based features, into a supervoxel RF model. Three main innovations are applied to effectively improve the classification accuracy of the proposed model. . (1) Rather than single points, we use supervoxels as the basic entity to construct the RF model and constrain the domain information via LCCP segmentation. (2) A two-level graphical model involving supervoxel calculation and LCCP optimization is generated from the raw point cloud, which significantly improves the reliability and accuracy of neighborhood searching. (3) The features are divided into three categories based on feature descriptions (point-based, eigen-based, and grid-based), and three unique feature calculation strategies are accordingly designed to improve feature reliability. We conduct two experiments using ALS data collected from Toronto site which has been provided by ISPRS benchmark and real scene data collected from Shenzhen, China, respectively. We compare the quantitative analysis of ALS datasets with other existed methods maintaining high performance, and the classification results demonstrate the robustness and effectiveness of the proposed method. However, the proposed method still has some limitations on scene generalizability. The algorithm may fail to recognize roof components when lacking facade information, which is caused by a loss of the connection relationship between supervoxels. In the future, we would like to integrate external constraints into the classification process to prevent the influence of over-segmentation.

REFERENCES

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S., 2012. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE transactions on pattern analysis and machine intelligence*, 34(11), 2274–2282.

Breiman, L., 2001. Random forests. *Machine learning*, 45(1), 5–32.

Chen, H., Wang, C., Chen, T., Zhao, X., 2017. Feature selecting based on fourier series fitting. *2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, IEEE, 241–244.

Christoph Stein, S., Schoeler, M., Papon, J., Worgotter, F., 2014. Object partitioning using local convexity. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 304–311.

Croce, V., Caroti, G., De Luca, L., Jacquot, K., Piemonte, A., Véron, P., 2021. From the Semantic Point Cloud to Heritage-Building Information Modeling: A Semiautomatic Approach Exploiting Machine Learning. *Remote Sensing*, 13(3), 461.

Fabri, A., Pion, S., 2009. Cgal: The computational geometry algorithms library. *Proceedings of the 17th ACM SIGSPATIAL international conference on advances in geographic information systems*, 538–539.

Feng, M., Zhang, L., Lin, X., Gilani, S. Z., Mian, A., 2020. Point attention network for semantic segmentation of 3D point clouds. *Pattern Recognition*, 107, 107446.

Hackel, T., Wegner, J. D., Schindler, K., 2016. Fast semantic segmentation of 3D point clouds with strongly varying density. *ISPRS annals of the photogrammetry, remote sensing and spatial information sciences*, 3, 177–184.

Javernick, L., Brasington, J., Caruso, B., 2014. Modeling the topography of shallow braided rivers using Structure-from-Motion photogrammetry. *Geomorphology*, 213, 166–182.

Jie, S., Zulong, L., 2014. Airborne LiDAR feature selection for urban classification using random forests. *Geomatics and Information Science of Wuhan University*, 39(11), 1310–1313.

Lafarge, F., Mallet, C., 2012. Creating large-scale city models from 3D-point clouds: a robust approach with hybrid representation. *International journal of computer vision*, 99(1), 69–85.

Li, X., Wang, L., Wang, M., Wen, C., Fang, Y., 2020. DANCE-NET: Density-aware convolution networks with context encoding for airborne LiDAR point cloud classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 166, 128–139.

Li, X., Wen, C., Cao, Q., Du, Y., Fang, Y., 2021. A novel semi-supervised method for airborne LiDAR point cloud classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 180, 117–129.

Muja, M., Lowe, D. G., 2014. Scalable nearest neighbor algorithms for high dimensional data. *IEEE transactions on pattern analysis and machine intelligence*, 36(11), 2227–2240.

Ni, H., Lin, X., Zhang, J., 2017. Classification of ALS point cloud with improved point cloud segmentation and random forests. *Remote Sensing*, 9(3), 288.

Niemeyer, J., Rottensteiner, F., Soergel, U., 2014. Contextual classification of lidar data and building object detection in urban areas. *ISPRS journal of photogrammetry and remote sensing*, 87, 152–165.

Niemeyer, J., Rottensteiner, F., Sörgel, U., Heipke, C., 2016. Hierarchical higher order crf for the classification of airborne lidar point clouds in urban areas. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences-ISPRS Archives 41 (2016)*, 41, 655–662.

Papon, J., Abramov, A., Schoeler, M., Worgotter, F., 2013. Voxel cloud connectivity segmentation-supervoxels for point clouds. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2027–2034.

Qi, C. R., Su, H., Mo, K., Guibas, L. J., 2017a. Pointnet: Deep learning on point sets for 3d classification and segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 652–660.

Qi, C. R., Yi, L., Su, H., Guibas, L. J., 2017b. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *arXiv preprint arXiv:1706.02413*.

Rottensteiner, F., Sohn, G., Jung, J., Gerke, M., Baillard, C., Benitez, S., Breitkopf, U., 2012. The ISPRS benchmark on urban object classification and 3D building reconstruction. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences I-3 (2012), Nr. 1*, 1(1), 293–298.

Rusu, R. B., Blodow, N., Marton, Z. C., Beetz, M., 2008a. Aligning point cloud views using persistent feature histograms. *2008 IEEE/RSJ international conference on intelligent robots and systems*, IEEE, 3384–3391.

Rusu, R. B., Marton, Z. C., Blodow, N., Beetz, M., 2008b. Learning informative point classes for the acquisition of object model maps. *2008 10th International Conference on Control, Automation, Robotics and Vision*, IEEE, 643–650.

Wang, Y., Chen, Q., Liu, L., Li, X., Sangaiah, A. K., Li, K., 2018. Systematic comparison of power line classification methods from ALS and MLS point cloud data. *Remote Sensing*, 10(8), 1222.

Wu, F., Wen, C., Guo, Y., Wang, J., Yu, Y., Wang, C., Li, J., 2016. Rapid localization and extraction of street light poles in mobile LiDAR point clouds: A supervoxel-based approach. *IEEE Transactions on Intelligent Transportation Systems*, 18(2), 292–305.

Xiong, B., Jancosek, M., Elberink, S. O., Vosselman, G., 2015. Flexible building primitives for 3D building modeling. *ISPRS Journal of Photogrammetry and Remote Sensing*, 101, 275–290.

Yue, X., Wu, B., Seshia, S. A., Keutzer, K., Sangiovanni-Vincentelli, A. L., 2018. A lidar point cloud generator: from a virtual world to autonomous driving. *Proceedings of the 2018 ACM on International Conference on Multimedia Retrieval*, 458–464.

Zhou, Q.-Y., Neumann, U., 2010. 2.5 d dual contouring: A robust approach to creating building models from aerial lidar point clouds. *European conference on computer vision*, Springer, 115–128.

Zhu, Q., Li, Y., Hu, H., Wu, B., 2017. Robust point cloud classification based on multi-level semantic relationships for urban scenes. *ISPRS journal of photogrammetry and remote sensing*, 129, 86–102.