# A MULTI-SCALE POINT CLOUDS SEGMENTATION METHOD FOR URBAN SCENE CLASSIFICATION USING REGION GROWING BASED ON MULTI-RESOLUTION SUPERVOXELS WITH ROBUST NEIGHBORHOOD

Junjie Huang[1], Linfu Xie[1,*], Weixi Wang[1], Xiaoming Li[1], Renzhong Guo[1]

[1]  Research Institute for Smart Cities, School of Architecture and Urban Planning, Shenzhen University, Shenzhen, PR China;
huangjunjie2020@email.szu.edu.cn; linfuxie@szu.edu.cn; wangwx@szu.edu.cn; lixming@szu.edu.cn; guorz@szu.edu.cn

KEYWORDS: Point clouds classification, Urban scene, Point clusters, Supervoxels, Region growing

**ABSTRACT:**

Point clouds classification is the basis for 3D spatial information extraction and applications. The point-clusters-based methods are proved to be more efficient and accurate than the point-based methods, however, the precision of the classification is significantly affected by the segmentation errors. The traditional single-scale point clouds segmentation methods cannot segment complex objects well in urban scenes which will result in inaccurate classification. In this paper, a new multi-scale point clouds segmentation method for urban scene point clouds classification is proposed. The proposed method consists of two stages. In the first stage, to ease the segmentation errors caused by density anisotropy and unreasonable neighborhood, a multi-resolution supervoxels segmentation algorithm is proposed to segment the objects into small-scale clusters. Firstly, the point cloud is segmented into initial supervoxels based on geometric and quantitative constraints. Secondly, robust neighboring relationships between supervoxels are obtained based on kd-tree and octree. Furthermore, the resolution of supervoxels in the planar and low-density region is optimized. In the second stage, planar supervoxels are clustered into the large-scale planar point clusters based on the region growing algorithm. Finally, a mix of small-scale and large-scale point clusters is obtained for classification. The performance of the segmentation method in classification is compared with other segmentation methods. Experimental results revealed that the proposed segmentation method can significantly improve the efficiency and accuracy of point clouds classification than other segmentation methods.

## 1. INTRODUCTION

In recent years, with the development of 3D data acquisition technology, large-scale and high-precision 3D point clouds can be easily obtained through lidar technology and photogrammetry. Point clouds have become an ideal data carrier for expressing 3D space, providing spatial information for 3D city modeling. Point clouds classification is the basis for 3D spatial information extraction and application (Che et al., 2019).

Commonly, point clouds classification methods can be divided into point-based methods and point-clusters-based methods (Xu et al., 2012). In point-based classification, the individual point does not have sufficient characteristics to support the classification. There is much research that focuses on the point clouds neighborhood selection (Filin and Pfeifer, 2006), feature extraction, and feature selection (Weinmann et al., 2015; Gupta et al., 2020). However, the point-based methods still suffer from the point clouds density anisotropy, unreasonable neighborhood, and noise. In comparison, point-clusters-based methods are more efficient and more reliable (Vosselman et al., 2017). Point-clusters-based methods first apply segmentation to the point cloud for clustering points with homogeneity. Whether in machine learning classifiers such as support vector machines(SVG) (Zhang et al., 2004), random forest(RF) (Breiman, 2001), or deep learning classifiers such as PointNet (Charles et al., 2017), PCNN (Atzmon et al., 2018), segmentation has many advantages for classification. Firstly, it reduces the number of objects to be classified to improve classification efficiency (Vosselman et al., 2017). Secondly, segmentation avoids expensive point neighborhood selection. It can separate individual point

clusters from the scene. These clusters have information about their properties like size and distinct geometric features. However, point-clusters-based methods require accurate segmentation of initial point clouds. Point clouds segmentation is a hot and difficult area of research in both computer vision and photogrammetry. Currently, there are four major algorithms: model fitting (Schnabel et al., 2007), feature cluster (Filin, 2004), deep learning (Bello et al., 2020), and region growing (Vo et al., 2015). In the urban scene, building facades, roofs, and ground are all pane structures. While region-growing-based methods are widely used in segmenting the point clouds into planar, they are not particularly robust as has been shown in urban scenes (Vo et al., 2015; Li et al., 2019). Mainly because the multiple growth criterion in the algorithm is poorly adaptive to different scale objects in the urban scene. If we set a conservative criterion, it will lead to the problem of under-segmentation in small-scale objects. However, if we set an aggressive criterion, it will fail to extract the continuous plane in the large-scale objects.

All of these can lead to serious classification errors. In this paper, a new supervoxels based on region growing method is proposed, which aims to improve point clouds classification by segmenting the objects in the urban scene to their appropriate scale point clusters. The proposed method is a multi-scale point clouds segmentation algorithm consisting of two stages(see Figure 1). In the first stage, all objects are segmented into small-scale supervoxels. A multi-resolution supervoxels with robust neighborhoods segmentation algorithm is proposed to ease the segmentation errors caused by unreasonable neighborhood and density anisotropy. In the second stage, as the objects at different scales in urban scenes can be distinguished by the planarity, planar supervoxels are
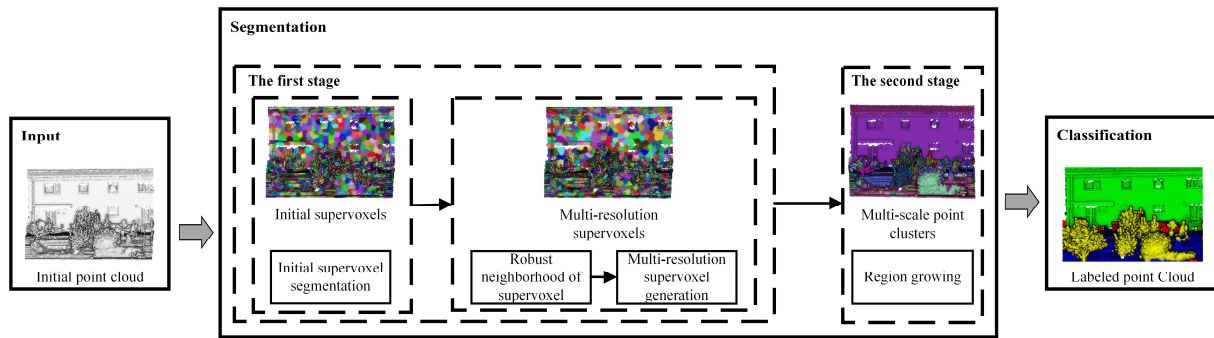
---

\* Corresponding author

**Figure 1**. Workflow of our method.

merged into large-scale planar point clusters by using the region growing.

To verify the effectiveness of the segmentation method proposed in this paper, an RF classifier is used to distinguish the classes of the clusters utilizing five kinds of features.

## 2. METHODOLOGY

### 2.1 Supervoxels-based region growing

In this paper, for different objects in the urban scene, two scales of point clusters are defined, namely, supervoxels and planar point clusters. Supervoxel is compact point clusters that provide a more natural and approximate representation of points. It preserves well the boundaries which mean that the supervoxels of the different scale objects are more distinguishable. Large-scale planar point clusters are obtained by the region growing algorithm when using the supervoxels as the basic unit. Planar point clusters distinguish objects with a planar structure from other objects in the urban scene.

**2.1.1    Initial supervoxels segmentation.** In this paper, the initial supervoxels are segmented by a heuristic algorithm BPSS proposed by (Lin et al.2018). This algorithm constructs an energy function (shown in Equation 1) which consists of a geometric difference term and a clusters number term to abstract the process of supervoxels segmentation. Supervoxels are obtained by using Fusion and Exchange algorithm (Lin et al.2018) to minimize the energy function.

$$minE(Z) = \sum_{i=1}^{N}\sum_{j=1}^{N} z_{ij} D(p_i, p_j) + \lambda|C(Z) - K|$$
$$s.t. z_{ij} = \{0,1\}, \forall i,j; \sum_{i=1}^{N} z_{ij} = 1, \forall j \quad (1)$$

Where $E(Z)$ represents the energy function. $\sum_{i=1}^{N}\sum_{j=1}^{N} z_{ij} D(p_i, p_j)$ is the geometric difference term, and $z_{ij}$ means if the point $p_j$ is represented by $p_i$. $D(p_i, p_j)$ is the geometric difference measure between $p_i$ and $p_j$. $\lambda|C(Z) - K|$ is the cluster number term used to control the supervoxels resolution. $\lambda$ is the regular coefficient used to weigh the geometric difference term and the cluster number term. $C(Z)$ represents the number of supervoxels currently generated. $K$ represents the desired number of supervoxels according to the present resolution.

**2.1.2    Robust neighborhood of supervoxels.** While after using the BPSS algorithm supervoxels cluster points that are homogeneous within the local neighbors, the limited size of individual supervoxel results in its inability to represent the large scale objects. Therefore, robust neighborhood is defined for each supervoxel to gain contextual information. In former research, such as VCCS (Papon et al., 2013) obtain the supervoxels neighborhoods based on the Octree, but fixed resolution voxelization cannot guarantee the effective detection of the density anisotropy scene (Li et al., 2021). In this paper, we first use the kd-tree to obtain supervoxels initial neighborhood, and then dynamically adjust the voxel resolution to optimize the initial neighborhood according to the density of the local region. The illustration of how to define the robust neighborhood for each supervoxel is shown in Figure 2.

First, a sphere $S(p_c, r)$ is fitted to the supervoxel $s_i$. The definition of the sphere center $p_c$ and the radius $r$ are shown in Equation 2.

$$\begin{cases} P_c = \frac{(\sum_{i=0}^{n} x_i, \sum_{i=0}^{n} y_i, \sum_{i=0}^{n} z_i)}{n} \\ r = \max(dis(P_c, p_i)) \end{cases} s.t. p_i(x_i, y_i, z_i) \in s_i \quad (2)$$

Where $P_c$ represents the centroid point of the supervoxel $s_i$. $dis(P_c, p_i)$ is the distance from $P_c$ to the point $p_i$ in the supervoxel $s_i$. $r$ is the furthest distance.

The spatial indexing of the point cloud is built by the kd-tree. Then, we search the neighboring points of each point in the supervoxel $s_i$ with the search radius of $r$ to constitute the neighboring points set $P_N = \{p_{n_1}, p_{n_2}, p_{n_3}, \dots, p_{n_m}\}$. If the nieghboring point in the $P_N$ satisfies the assignment rule in Equation 3, the *adjancy* relations between the supervoxel $s_{n_i}$ in which the neighboring point is located and the supervoxel $s_i$ is established.

$$(s_i, s_{n_i}) \begin{cases} adjancy, & if(i \neq n_i) \\ equal, & else \end{cases} \quad (3)$$

The initial neighborhood of the supervoxels is obtained by iteratively performing the above operations on each supervoxel. There are both directly connected neighbors and not directly connected neighbors in the initial neighborhoods.

As supervoxels maintain robust adjacency relations in voxelized 3D space (Stein et al., 2014), the supervoxel and its initial neighboring supervoxels are voxelized based on the local density. If neither of the voxels in the neighboring supervoxel is adjacent to the voxels in the target supervoxel, the two supervoxels are not directly connected and remove the
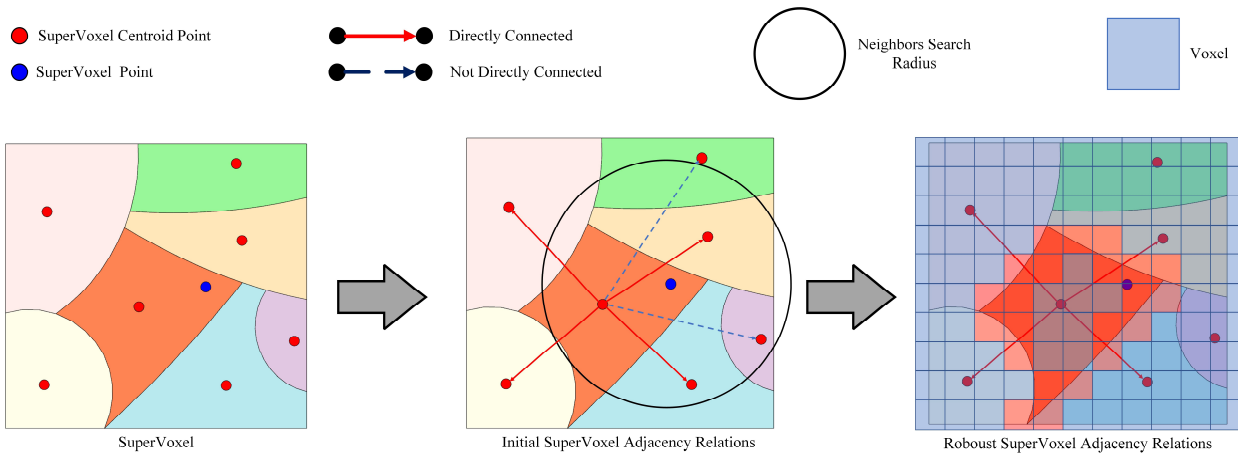
**Figure 2**. Illustration of how to define the robust nieghborhood.

neighboring supervoxel from the target supervoxel neighborhood.

**2.1.3 Multi-resolution supervoxels generation.**
Conventional supervoxels segmentation algorithms including the BPSS method and VCCS method cannot effectively segment low-density point clouds regions and noise effectively, which will result in many small fragments. These small fragments containing few points seriously affect the region growing efficiency. To address the above problems, a multi-resolution suprvoxel generation algorithm is proposed. This algorithm consists of two parts: planar region detection and supervoxel merging. The purpose of using the above algorithm is to merge these small fragments into their neighboring planar supervoxels in the planar region.

In the stage of planar region detection, the planarity is calculated firstly for each supervoxel. The planarity are then smoothed based on the supervoxels local context information by using a Gaussian filter. Subsequently, planar and scatter supervoxels are distinguished according to the rule in Equation 4.

$$\begin{cases} Planar, & if(P_\lambda > (maxP_\lambda/k) \\ Scatter, & else \end{cases} s.t. k \in (0,1] \quad (4)$$

Where $P_\lambda$ is the supervoxel planarity after smoothing. $maxP_\lambda$ is largest supervoxel planarity after smoothing. $k$ is the planarity control thresholds to control the planarity of the selected planar supervoxels.

If the proportion of planar supervoxels is greater than 2/3 in the supervoxel neighborhood, the local region of the supervoxel is considered to be a planar region. The neighbor supervoxel that satisfies the assignment rule in Equation 5 will be merged into the supervoxel.

$$\begin{cases} Merge, & if(\frac{r_c}{r_R}\frac{d_R}{d_n} < 1) \\ noMerge, & else \end{cases} \quad (5)$$

Where the $d_R$ is the density of the local region; $d_n$ is the density of the neighbor supervoxel. The lower value of the $\frac{d_R}{d_n}$, the more likely the neighbor supervoxel is a small fragment. Where $r_c$ is the distance from the neighbor supervoxel to the fitting planed based on the supervoxel; $r_R$ is the distance from the neighbor supervoxel to the fitting planed based on the

region. The lower value of the $\frac{r_c}{r_R}$, the more likely the neighbor supervoxel and the supervoxel belong to the same object.

Note the advantages of our planar region detection method over the method proposed in Dong et al., 2018. Our proposed method is more robust because our algorithm need only the planar feature without calculating other local geometric features. Meanwhile, the supervoxel spatial context information is used for planar region detection. The method of Dong et al., 2018 and the planar region detection method in this paper are shown in Figure 3.
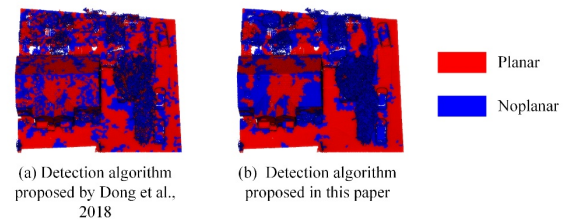


(a) Detection algorithm proposed by Dong et al., 2018

(b) Detection algorithm proposed in this paper

**Figure 3**. Comparison of the planar region detection algorithms.

A comparison of the results before and after the multi-resolution supervoxels generation algorithm is shown in Figure 4.
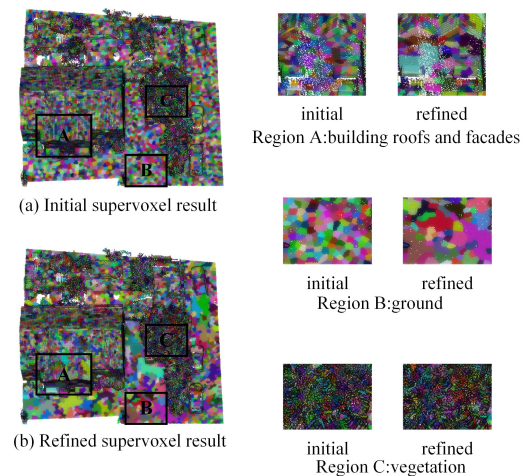


(a) Initial supervoxel result

(b) Refined supervoxel result

Region A: building roofs and facades

Region B: ground

Region C: vegetation

**Figure 4**. Comparison of the results before and after the proposed algorithm.

## 2.2 Region growing for planar structures

The supervoxels-based region growing algorithm groups the neighboring supervoxels with similar planar features into large-scale planar point clusters. The efficiency of the region growing is increased as the not directly connected neighbor supervoxels are removed from the supervoxels robust neighborhood. Meanwhile, the multi-scale point clusters returning from the supervoxels-based region growing also obtain the robust neighborhoods which are derived from the supervoxels robust neighborhoods. As the small scattery supervoxels are merged into the adjacent planarity supervoxels, the planar point clusters retuning from the supervoxels-based region growing are more complete.

The strategies for choosing the seed, obtaining the robust neighborhood, terminating the growing are illustrated in the pseudo-code which is described in Algorithm 1.

---

**Algorithm 1: supervoxels-based region growing algorithm**

**Data**: Multi-Resolution supervoxels with robust neighborhood $S = \{s_1, s_2, ..., s_n\}$, residual threshold $r_{th}$, angle threshold $\theta_{th}$;

**Result:** a set of Multi-Scale Point Clusters $C' = \{c'_1, c'_2, ..., c'_m\}$, a neighborhoods set of the Multi-Scale Point Clusters $N' = \{n_{c'_1}, n_{c'_2}, n_{c'_3} ..., n_{c'_m}\}$

**Initialization:** Point Clusters $C'$ , Pending SuperVoxel list $S_L \ll S$ , Seed map $S_m \ll S$ ;

**begin**

    **While** $S_L \neq \emptyset$ **do**

        Set current cluster $c_c$, current seed $s_i$, current nieghbors $N_{s_i}$;

        **if** visited $s_i$ **then**

            **continue**;

        insert $s_i$ into $c_c$; remove $s_i$ from $S_L$;

        **for** each supervoxel $s_j \in N_{s_i}$ **do**

            **if** visited $s_j$ **then**

                insert $c_{s_j}$ into $n_{c_c}$;

            **else**

                set $s_j$ visited; set $N_{s_i} \ll N_{s_j}$;

        **if** $risdual(s_j, c_c,) \leq r_{th}$ && $angle(s_j, c_c,)$

            insert $s_j$ into $c_c$;

            remove $s_j$ from $S_L$;

        insert $c_c$ into $C'$;

    **Return** $C'$;

---

## 2.3 Multi-scale point clusters feature extraction

The individual point contains only 3D spatial coordinate information. As the point clusters and their robust neighborhood are already obtained, it is necessary to fully use them to extract features. It is useful to use the point clusters spatial context information for the local geometric and structural feature extraction. The local 3D shape features such as linearity, planarity, and scattering can describe the geometric and structural features in the local area of the point cluster well. The point clusters at different scales are used to represent different objects in complex urban scenes. The point clusters of roofs and facades are a whole plane with regular orientation, bigger size, and smooth surface. The point clusters of vegetation are discrete with miscellaneous orientation, smaller size, and rough surface. Therefore, using the orientation features, property features, and surface features of the multi-scale point clusters can highly represent the different objects in the urban scene. The height features are also introduced.

In Table 1, we show the definition of each feature introduced in this paper.

## 2.4 Classification

To test the effectiveness of the point clouds segmentation method proposed in this paper in urban scene point clouds classification. A random forest(RF) classifier is trained to distinguish point clusters with different labels. The random forest classifier is chosen because it consists of multiple layers of decision trees that are created from independent random vectors and these are obtained by randomly sampling the feature input vectors. This has a greater advantage when dealing with data with multi-dimensional attributes, not only in terms of faster learning but also in terms of producing more reliable classification results.

## 3. EXPERIMENTS

To prove the effectiveness of the point clouds segmentation method proposed in this paper, e also segmented the point clouds into point clusters using the point-based region growing algorithms and the voxel-based region growing algorithms for classification. The region growing algorithm parameters of the three methods are the same. The classifier and features used for classification are also the same.

## 3.1 Test datasets

Lidar point clouds acquisition via Unmanned Aerial Vehicle(UAV) is becoming a trend. The performance of our method is evaluated on the UAV lidar point cloud of Hessigheim Germany (Koelle et al., 2021) which was acquired by the sensor setup consisting of a RIEGL VUX-1LR scanner and two oblique Sony Alpha 6000 cameras integrated on a RIEGL Ricopter platform. The dataset features a mean point density of about 800 pts/m². Compare to other Airborne Laser Scanning(ALS) point clouds, this point cloud has a higher resolution. This enables the identification of fine-grained structures.

Two test sites with different scenes are selected in the benchmark. They are shown in Figure 5. Area 1 is situated on the main road, with buildings evenly spaced on both sides of the main road and has a topographic drop. Area 2 is a residential neighborhood with compact, small houses with gable floors.

All point clouds in the two areas consist of 21,269,430 points. There are eleven object classes were labeled in the benchmark: low vegetation, impervious surface, vehicle, urban furniture, roof, façade, shrub, tree, soil/gravel, vertical surface, chimney. In this paper, only the 3D coordinates of point clouds are considered, so we group geometrically similar objects into one class. We distinguished the following four object classes: ground(impervious surface, soil/gravel), building(roof, façade, chimney), vegetation(low vegetation, shrub, tree), and, other(vehicle, urban furniture, vertical surface). All experiments are performed using C++ and run on an Intel i7-10700 CPU @ 2.90GHZ and with 64.0GB RAM.

| Feature | Definition | Category |
|---|---|---|
| Linearity | $\dfrac{\lambda_1 - \lambda_2}{\lambda_1}$ | Local 3D shape features |
| Planarity | $\dfrac{\lambda_2 - \lambda_3}{\lambda_1}$ | |
| Scattering | $\dfrac{\lambda_3}{\lambda_1}$ | |
| Anisotropy | $\dfrac{\lambda_1 - \lambda_3}{\lambda_1}$ | |
| Eigenentropy | $-\sum_{i=1}^{3} \lambda_i ln(\lambda_i)$ | |
| Normal | $N(n_x, n_y, n_z)$ | Oritention features |
| Verticality | $1 - n_z$ | |
| Cluster Size | $size(C)$ | property feature |
| Distance to plane | $\dfrac{1}{n}\sum_{i=1}^{n} dis(p_i, P)$ | Surface features |
| Dispersion | $1 - \dfrac{num(Grid_{occupied})}{num(Grid)}$ | |
| Elevation | $\dfrac{1}{n}\sum_{i=1}^{n} dis(p_i, G)$ | Height features |
| Height bleow | $\dfrac{1}{n}\sum_{i=1}^{n} (E_i - E_{min})$ | |
| Height above | $\dfrac{1}{n}\sum_{i=1}^{n} (E_{max} - E_i)$ | |
| Vertical range | $E_{max} - E_{min}$ | |

**Table 1.** List of features.



(a) Area 1                    (b) Area 2

**Figure 5.** Test sites of Hessigheim benchmark.

### 3.2 Accuracy evaluation metrics

The classification results are quantitatively evaluated. The number of True Positives (TP), the number of False Positives (FP), and the number of False Negatives (FN) are calculated for each class $i$. We use precision (Equation 6), recall (Equation 7), $F_1$-score (Equation 8), Intersection over UNIO (IoU) (Equation 9) for each class.

We also show mean accuracy, mean $F_1$-score, and mean IoU over the whole test areas.

$$P_i = \frac{TP_i}{TP_i + FP_i} \quad (6)$$

$$R_i = \frac{TP_i}{TP_i + FN_i} \quad (7)$$

$$F1_i = \frac{2 \cdot P_i \cdot R_i}{P_i + R_i} \quad (8)$$

$$IoU_i = \frac{TP_i}{TP_i + FP_i + FN_i} \quad (9)$$

### 3.3 Experimental result

Quantitative classification results of three methods in the two test datasets are shown in Table 2. Our mean recall is 0.822 for Area 1 and 0.833 for Area 2 while the mean precision is 0.933 and 0.929. Meanwhile, the classification of objects (Ground, Building) with planarity has achieved a good result. Most evaluation values of different classes are higher than the other methods. We attribute the result primarily to the reasons that we segmented the different classes of objects in the urban scene into appropriate scale point clusters, making the feature

more responsive to their characteristics. The overall view of the segmentation and classification results is shown in Figure 6. In Figure 7, we can find the classification errors due to the segmentation errors in the other two methods. Due to the innovative multi-resolution supervoxels with robust neighborhood segmentation algorithm proposed in this paper, these segmentation errors are successfully avoided.

| Area 1 | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Method | Point-based | | | | Voxel-based | | | | Our method | | | |
| **Class** | **Pre** | **Rec** | **F1** | **IoU** | **Pre** | **Rec** | **F1** | **IoU** | **Pre** | **Rec** | **F1** | **IoU** |
| Ground | 0.916 | 0.909 | 0.912 | 0.838 | 0.967 | 0.791 | 0.870 | 0.771 | 0.968 | 0.906 | 0.936 | 0.880 |
| Building | 0.919 | 0.841 | 0.878 | 0.783 | 0.877 | 0.925 | 0.901 | 0.820 | 0.953 | 0.924 | 0.939 | 0.885 |
| Vegetation | 0.834 | 0.958 | 0.892 | 0.805 | 0.881 | 0.970 | 0.923 | 0.858 | 0.865 | 0.981 | 0.919 | 0.851 |
| Other | 0.644 | 0.171 | 0.271 | 0.156 | 0.222 | 0.453 | 0.298 | 0.175 | 0.627 | 0.477 | 0.542 | 0.371 |
| MEAN | 0.886 | 0.719 | 0.738 | 0.646 | 0.872 | 0.784 | 0.748 | 0.656 | 0.922 | 0.822 | 0.834 | 0.747 |
| Area 2 | | | | | | | | | | | | |
| Method | Point-based | | | | Voxel-based | | | | Our method | | | |
| **Class** | **Pre** | **Rec** | **F1** | **IoU** | **Pre** | **Rec** | **F1** | **IoU** | **Pre** | **Rec** | **F1** | **IoU** |
| Ground | 0.929 | 0.905 | 0.917 | 0.847 | 0.944 | 0.793 | 0.826 | 0.703 | 0.943 | 0.959 | 0.950 | 0.906 |
| Building | 0.945 | 0.929 | 0.937 | 0.882 | 0.956 | 0.932 | 0.944 | 0.894 | 0.967 | 0.958 | 0.962 | 0.928 |
| Vegetation | 0.571 | 0.803 | 0.667 | 0.501 | 0.442 | 0.893 | 0.592 | 0.420 | 0.739 | 0.870 | 0.799 | 0.665 |
| Other | 0.584 | 0.277 | 0.376 | 0.231 | 0.569 | 0.483 | 0.523 | 0.354 | 0.768 | 0.427 | 0.549 | 0.378 |
| MEAN | 0.884 | 0.728 | 0.724 | 0.615 | 0.844 | 0.775 | 0.721 | 0.593 | 0.929 | 0.803 | 0.815 | 0.719 |

**Table 2.** Quantitative analysis of classification results**.**



(a) Point-based in Area 1

(b) Voxel-based in Area 1

(c) Our method in Area 1

(d) Point-based in Area 2

(e) Voxel-based in Area 2

(f) Our method in Area 2

**Classification Legend**
Ground
Building
Vegetation
Other
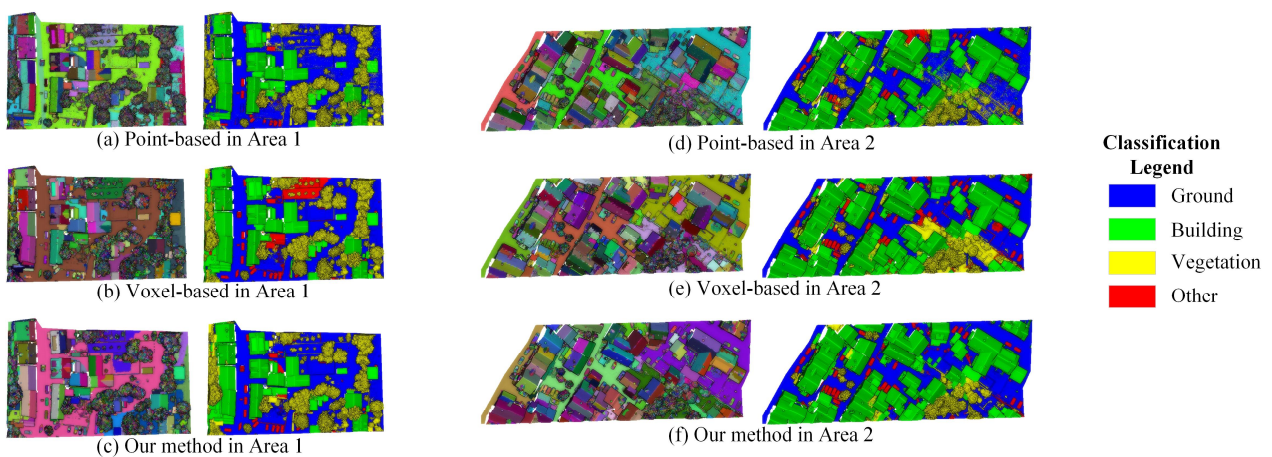
**Figure 6.** Segmentation results (left) and Classification results (right) in Area 1 and Area 2.
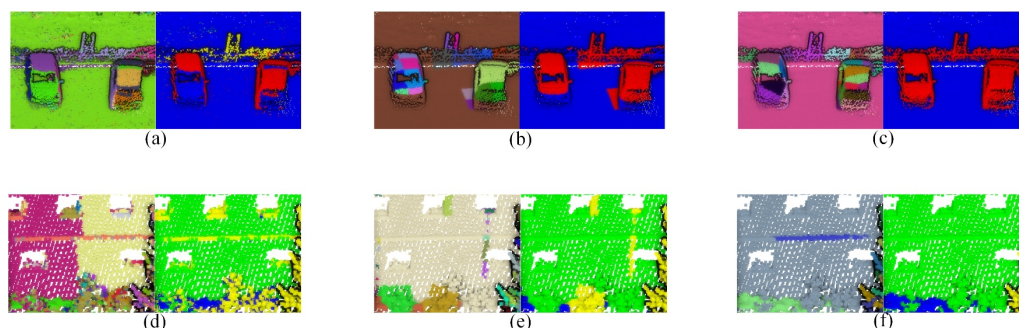


(a)　(b)　(c)

(d)　(e)　(f)

**Figure 7.** Comparison of segmentation details and classification results. (a) (d) are the results of point-based methods. (b) (e) are the results of voxel-based methods. (c) (f) are the results of voxel-based methods.

## 4. CONCLUSION

We have proposed an efficient segmentation method for point-clusters-based classification in urban scenes and evaluated the performance on several datasets. A multi-resolution supervoxels with robust neighborhood based region growing algorithm is proposed in this paper. As a contribution, different classes of urban objects can be segmented into appropriate scales point clusters by our method. This makes it possible to extract rich features in point clusters which is meaningful for classification. Through experiments on the different datasets, we find that the problem of inhomogeneous density and lack of topological information of point clouds leading to segmentation errors and further classification errors is solved by using our method. This is mainly due to our innovative acquisition of multi-resolution supervoxels with the robust neighborhood to obtain more spatial contextual information. However, the method proposed in this paper remains to be improved. For segmentation, the initial supervoxel resolution needs to be set manually, and there are too many parameters required in the region growing process. This can result in some complex planar structures such as building façades with window frames not being segmented in their entirety. We think that A CRF can be applied to segmentation to solve this problem. For classification, the feature extraction of point clusters still needs to be investigated.

## ACKNOWLEDGEMENTS

## REFERENCES

Che, E., Jung, J. and Olsen, M., 2019. Object Recognition, Segmentation, and Classification of Mobile Laser Scanning Point Clouds: A State of the Art Review. Sensors, 19: {}.

Xu, S., Oude Elberink, S. and Vosselman, G., 2012. Entities and features for classification of airborne laser scanning data in urban areas. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, I-4: 257-262.

Filin, S. and Pfeifer, N., 2006. Segmentation of airborne laser scanning data using a slope adaptive neighborhood. ISPRS JOURNAL OF PHOTOGRAMMETRY AND REMOTE SENSING, 60(2): 71-80.

Weinmann, M., Urban, S., Hinz, S., Jutzi, B. and Mallet, C., 2015. Distinctive 2D and 3D Features for Automated Large-Scale Scene Analysis in Urban Areas. Computers & Graphics: {}.

Gupta, A., Watson, S., and Yin, H., 2020. 3D Point Cloud Feature Explanations Using Gradient-Based Methods, pp. 1-8.

Vosselman, G., Coenen, M. and Rottensteiner, F., 2017. Contextual segment-based classification of airborne laser scanner data. ISPRS JOURNAL OF PHOTOGRAMMETRY AND REMOTE SENSING, 128: 354-371.

Zhang, L., Zhou, W.D. and Jiao, L.C., 2004. Wavelet support vector machine. IEEE TRANSACTIONS ON SYSTEMS MAN AND CYBERNETICS PART B-CYBERNETICS, 34(1): 34-39.

Breiman, L., 2001. Random forests. MACHINE LEARNING, 45(1): 5-32.

Charles, R., Su, H., Mo, K. and Guibas, L., 2017. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation, pp. 77-85.

Atzmon, M., Maron, H. and Lipman, Y., 2018. Point Convolutional Neural Networks by Extension Operators. ACM TRANSACTIONS ON GRAPHICS, 37(4).

Schnabel, R., Wahl, R. and Klein, R., 2007. Efficient RANSAC for Point-Cloud Shape Detection. Computer graphics forum, 26(2): 214-226.

Filin, S., 2004. Surface classification from airborne laser scanning data. COMPUTERS \& GEOSCIENCES, 30(9-10): 1033-1041.

Bello, S.A., Yu, S., Wang, C., Adam, J.M. and Li, J., 2020. Review: Deep Learning on 3D Point Clouds. REMOTE SENSING, 12(11).

Vo, A., Truong-Hong, L., Laefer, D.F. and Bertolotto, M., 2015. Octree-based region growing for point cloud segmentation. ISPRS Journal of Photogrammetry and Remote Sensing, 104: 88-100.

Li, P., Wang, R., Wang, Y. and Gao, G., 2019. Automated Method of Extracting Urban Roads Based on Region Growing from Mobile Laser Scanning Data. SENSORS, 19(23).

Lin, Y., Wang, C., Zhai, D., Li, W. and Li, J., 2018. Toward better boundary preserved supervoxel segmentation for 3D point clouds. ISPRS Journal of Photogrammetry and Remote Sensing, 143(SEP.): 39-47.

Papon, J., Abramov, A., Schoeler, M. and Wörgötter, F., 2013. Voxel Cloud Connectivity Segmentation - Supervoxels for Point Clouds, pp. 2027-2034.

Li, H., Liu, Y., Men, C. and Fang, Y., 2021. A novel 3D point cloud segmentation algorithm based on multi-resolution supervoxel and MGS. International Journal of Remote Sensing, 42: 8492-8525.

Stein, S., Schoeler, M., Papon, J. and Wörgötter, F., 2014. Object Partitioning Using Local Convexity, pp. {}.

Dong, Z., Yang, B., Hu, P. and Scherer, S., 2018. An efficient global energy optimization approach for robust 3D plane segmentation of point clouds. ISPRS JOURNAL OF PHOTOGRAMMETRY AND REMOTE SENSING, 137: 112-133.

Koelle, M. et al., 2021. The Hessigheim 3D (H3D) benchmark on semantic segmentation of high-resolution 3D point clouds and textured meshes from UAV LiDAR and Multi-View-

Stereo. ISPRS Open Journal of Photogrammetry and Remote
Sensing, 1: 100001.