

# Categories of Point Cloud Segmentation Methods

Subjects: **Others**

Contributor: Xianquan Han , Xijiang Chen , Hui Deng , Peng Wan , Jianzhou Li

Laser point clouds have been widely used in many fields, such as autonomous driving, augmented reality and so on. Point cloud classification and segmentation are key to scene understanding. Therefore, many scholars have systematically studied point cloud classification and segmentation. Machine learning approaches are the most commonly used data analysis methods. In the field of point cloud data analysis, deep learning is used for point cloud classification and segmentation. To improve the accuracy of point cloud classification and segmentation, some researchers have drawn on image methods to handle point clouds.

point cloud

convolutional network

classification

segmentation

## 1. Introduction

There are three main categories of point cloud segmentation methods: projection-based methods, voxel-based methods and methods that directly operate on point clouds. Projection-based methods project the point cloud onto a 2D plane to form an image and use 2D CNNs to extract image features. Voxel-based methods divide the point cloud into blocks, and each block of points represents a voxel. Features are extracted from each voxel point cloud and fused to segment the entire point cloud. Point-cloud-based methods directly perform convolutional operations on point clouds, which can effectively obtain feature information from point cloud data.

## 2. Projection-Based Methods

The key to projection-based methods is constructing multi-view point clouds and then learning the 2D features of the multi-view projections through 2D convolutional operations. Ultimately, the features from the multi-view point clouds are fused. Su et al. <sup>[1]</sup> proposed a novel CNN architecture that combines information from multiple views of a 3D shape into a single and compact shape descriptor, offering even better recognition performance. The feature information of multiple views is integrated into a single, compact 3D shape descriptor by convolution and pooling layers, and the aggregated features are fed into a fully connected layer to obtain classification results. Qi et al. <sup>[2]</sup> improved multi-view CNN performance through a multi-resolution extension with improved data augmentation, which introduced multi-resolution 3D filtering to capture information at multiple scales. Thereby, the performance of the classification model is enhanced. Xu et al. <sup>[3]</sup> designed spatially adaptive convolution based on the structure of the SqueezeSeg model <sup>[4]</sup>. The model has the ability to achieve spatial adaptation and content awareness, which

solves the problem of network performance degradation caused by traditional convolution applied to LiDAR images.

### **3. Voxel-Based Methods**

Point cloud voxelization <sup>[5]</sup> means to voxelize disordered point clouds into regular structures and then use a neural network architecture for feature learning to achieve semantic segmentation of point clouds. Considering the sparsity of point clouds, Graham <sup>[6]</sup> designed a sparse convolutional network and applied it to 3D segmentation <sup>[7]</sup>. Tchapmi et al. <sup>[8]</sup> used the SEGCloud method to subdivide the large point cloud into voxel grids and post-processed them by using trilinear interpolation and conditional random fields. Li et al. <sup>[9]</sup> attempted to sample sparse 3D data and then feed the sampled results into a network for processing to reduce the computational effort. Le et al. <sup>[10]</sup> proposed the PointGrid method, which employs a 3DCNN to learn grid cells containing fixed points to obtain local geometric details. In addition, Hua et al. <sup>[11]</sup> proposed a lightweight, fully convolutional network based on an attention mechanism and a sparse tensor for the semantic segmentation of point clouds.

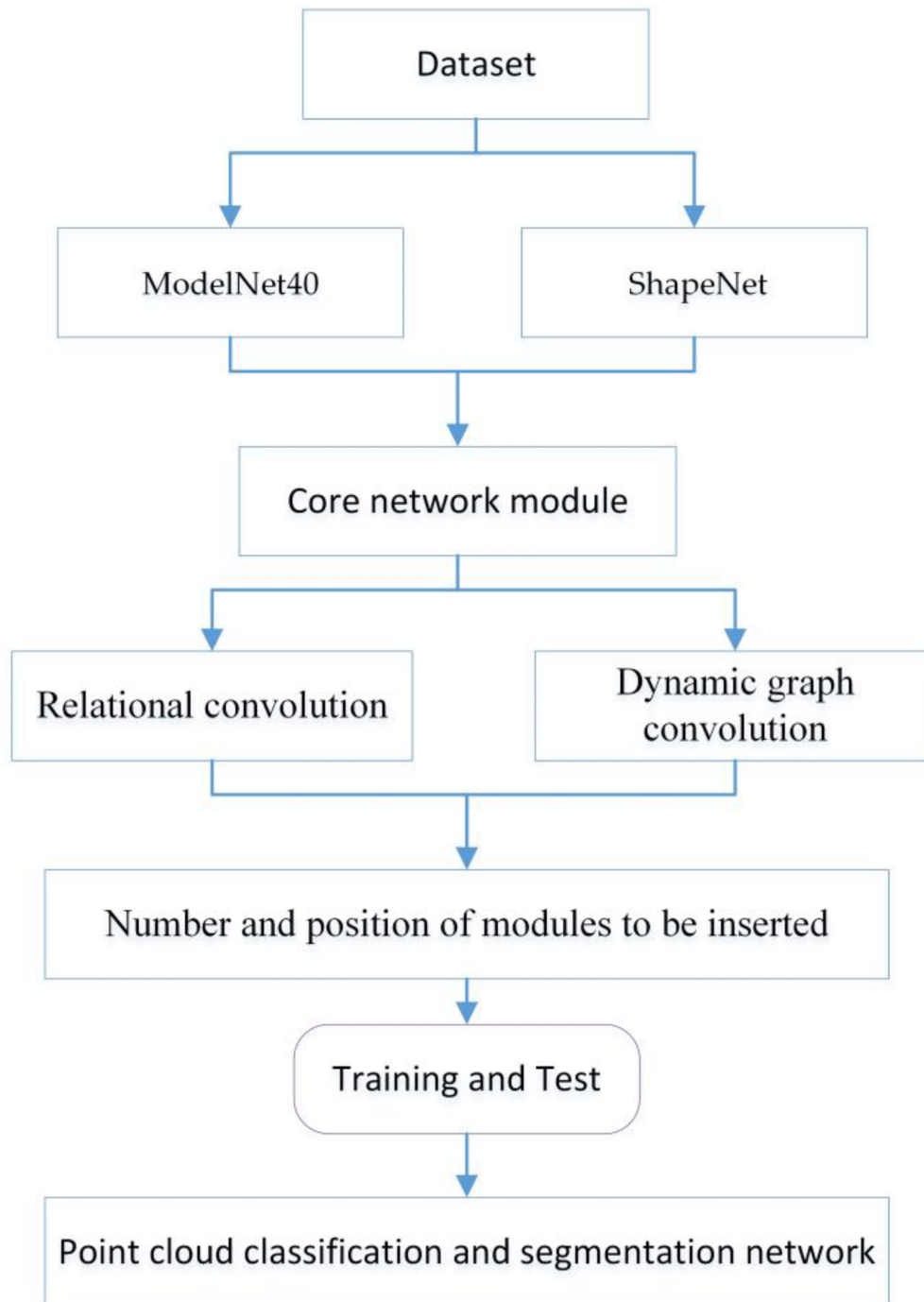
### **4. Point-Cloud-Based Methods**

The above two methods have high computational complexity. In view of this, the network model based on the original point cloud is gradually proposed. PointNet <sup>[12]</sup> was among the earliest approaches to apply deep learning network architectures for processing raw point cloud data. A limitation of this method is that it only extracts global features of point clouds to enable point cloud classification or semantic segmentation without accounting for local features, leading to less-than-ideal network performance for semantic segmentation in large scenes. In order to improve the segmentation performance of point clouds, Charles et al. <sup>[13]</sup> proposed PointNet++ by improving the PointNet framework. PointNet++ aims to address the limitation of the original PointNet that only considers global features of point clouds. The addition of the local region segmentation module allows PointNet++ to learn and utilize local geometric features at different scales, thereby improving its performance on large-scene semantic segmentation. Jiang et al. <sup>[14]</sup> proposed a hierarchical point–edge interaction network for point cloud semantic segmentation. It introduces an edge convolution operation to interact between point features and edge features for capturing local geometric structures. Unbalanced point cloud scenarios will affect the semantic segmentation accuracy of PointNet++. In view of this problem, Deng et al. <sup>[15]</sup> proposed a weighted sampling method based on farthest point sampling (FPS), which can make the sampling process more balanced and reduce the influence of point cloud scene imbalance. Dang et al. <sup>[16]</sup> proposed hierarchical parallel group convolution, which can capture both a single point feature and local geometric features of point clouds. Therefore, the model improves the ability of the network to recognize complex classes. He et al. <sup>[17]</sup> proposed a multi-feature PointNet (MSMF-PointNet) deep learning point cloud model that can extract multiscale, multi-neighborhood features for classification. Liu et al. <sup>[18]</sup> introduces a relation module to learn the relations between points and aggregate neighbor information. Additionally, a shape module is presented to encode absolute geometric features of each point. The two modules are combined to extract comprehensive shape representations. The network can reason about the spatial information features of the point cloud and realize the shape context perception. The global features are obtained through the fully

connected layer and then classified. Hu et al. [19] proposed an efficient and lightweight RandLA-Net network. The model used a local feature aggregation module to expand the range of k-nearest neighbor points to reduce information loss. The random sampling is used to reduce the storage cost and improve computational efficiency. Landrieu et al. [20] proposed a method for large-scale point cloud semantic segmentation using superpoint graphs. The approach first oversegments the point cloud into a set of superpoints. It then constructs a graph to represent relationships between superpoints. Graph convolutional networks are applied to perform semantic segmentation on the superpoint graph. It can greatly reduce the number of points in a point cloud, and the network can be applied in the field of large-scale point cloud datasets.

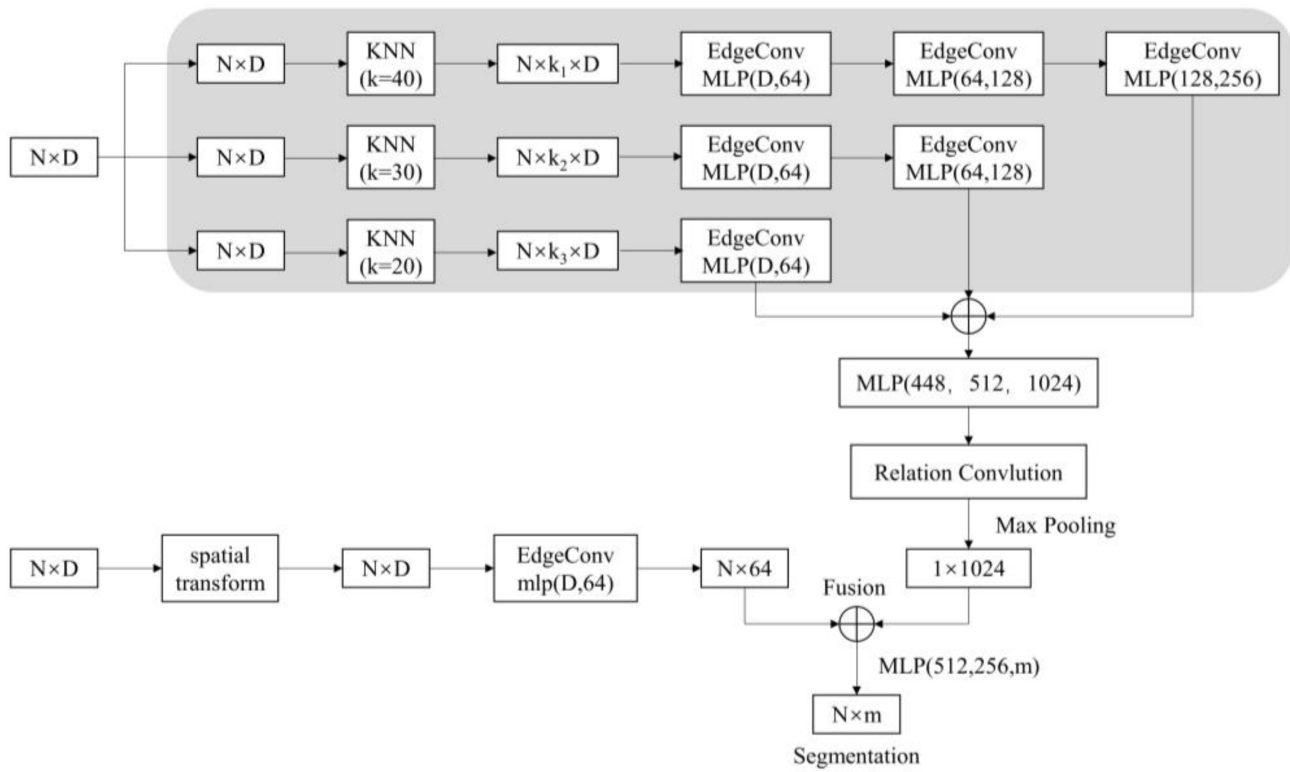
## **| 5. Proposed Network Architecture**

The flowchart of the proposed method is described in **Figure 1**. First, the input point cloud data are locally partitioned using three different scales of k-nearest neighbors (KNN), and a directed graph is constructed for the extraction of features between points.



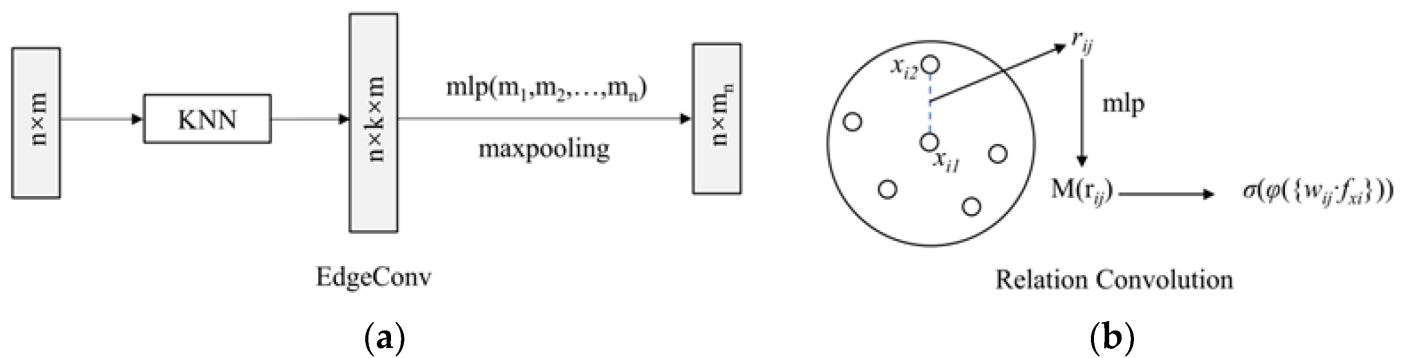
**Figure 1.** Flowchart of the proposed method.

Simultaneously, researchers utilize EdgeConv (edge convolution) layers to extract and aggregate local features of the point cloud. The regular network is then extended to the irregular structures by using relation convolution to further extract the features of the points. Finally, global features are output through maximum pooling. The point cloud deep learning network proposed in this research is constructed, as shown in **Figure 2**. The input of the network is the point set  $P = \{p_1, p_2, \dots, p_n\} \subseteq \mathbb{R}^F$ ,  $N$  is the number of sampled points and  $D$  is the feature dimension of each point of input.



**Figure 2.** Proposed network architecture.

When  $D = 3$ , the input points have only three dimensions, which means that the input point cloud data are  $(x, y, z)$ . The  $N \times D$  dimensional point clouds are input to the spatial transform networks (STN). The spatial transformation matrix learned by this module can align the coordinates of the input point cloud data. The aligned data are manipulated using a layer of the dynamic graph convolutional network module to obtain local neighborhood features, which are mapped to 64, 128 and 1024 dimensions in sequence by Multi-Layer Perceptron (MLP). The local features of the point cloud are extracted and aggregated by using EdgeConv, as shown in **Figure 3a**. Deeper feature extraction is then performed using a relation-shape convolution network module, as shown in **Figure 3b**. It utilizes relation-shape convolution to capture local geometric structures by modeling pairwise spatial relations between points. Global features are obtained through max pooling of the local features.



**Figure 3.** Structure of edge and relation convolution. (a) Edge Convolution Module. (b) Relation Convolution Module.

Point cloud classification is achieved by applying max pooling and an MLP to the global features. Point cloud segmentation is realized according to the combination of global and local features. It concatenates global and local features and outputs the score per point.

---

## References

1. Su, H.; Maji, S.; Kalogerakis, E.; Learned-Miller, E. Multi-view convolutional neural networks for 3d shape recognition. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 945–953.
2. Qi, C.R.; Su, H.; Nießner, M.; Dai, A.; Yan, M.; Guibas, L.J. Volumetric and multi-view cnns for object classification on 3d data. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 5648–5656.
3. Xu, C.; Wu, B.; Wang, Z.; Zhan, W.; Vajda, P.; Keutzer, K.; Tomizuka, M. Squeezesegv3: Spatially-adaptive convolution for efficient point-cloud segmentation. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; Springer: Cham, Switzerland, 2020; pp. 1–19.
4. Wu, B.; Wan, A.; Yue, X.; Keutzer, K. Squeezeseg: Convolutional neural nets with recurrent crf for real-time road-object segmentation from 3d lidar point cloud. In Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, QLD, Australia, 21–25 May 2018; pp. 1887–1893.
5. Lee, M.-y.; Lee, S.-h.; Jung, K.-d.; Lee, S.-h.; Kwon, S.-c. A Novel Preprocessing Method for Dynamic Point-Cloud Compression. *Appl. Sci.* 2021, 11, 5941.
6. Graham, B. Spatially-sparse convolutional neural networks. *arXiv* 2014, arXiv:1409.6070.
7. Verdoja, F.; Thomas, D.; Sugimoto, A. Fast 3D point cloud segmentation using supervoxels with geometry and color for 3D scene understanding. In Proceedings of the 2017 IEEE International Conference on Multimedia and Expo (ICME), Hong Kong, China, 10–14 July 2017; pp. 1285–1290.
8. Tchapmi, L.; Choy, C.; Armeni, I.; Gwak, J.; Savarese, S. Segcloud: Semantic segmentation of 3d point clouds. In Proceedings of the 2017 International Conference on 3D Vision (3DV), Qingdao, China, 10–12 October 2017; pp. 537–547.
9. Li, Y.; Pirk, S.; Su, H.; Qi, C.R.; Guibas, L.J. Fpnn: Field probing neural networks for 3d data. *Adv. Neural Inf. Process. Syst.* 2016, 29, 1–10.
10. Le, T.; Duan, Y. Pointgrid: A deep network for 3d shape understanding. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22

June 2018; pp. 9204–9214.

11. Wang, F.; Yang, Y.; Wu, Z.; Zhou, J.; Zhang, W. Real-Time Semantic Segmentation of Point Clouds Based on an Attention Mechanism and a Sparse Tensor. *Appl. Sci.* 2023, 13, 3256.
12. Qi, C.R.; Su, H.; Mo, K.; Guibas, L.J. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, 21–26 July 2017; pp. 652–660.
13. Qi, C.R.; Yi, L.; Su, H.; Guibas, L.J. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *Adv. Neural Inf. Process. Syst.* 2017, 30, 1–14.
14. Jiang, L.; Zhao, H.; Liu, S.; Shen, X.; Fu, C.W.; Jia, J. Hierarchical point-edge interaction network for point cloud semantic segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 10433–10441.
15. Deng, C.; Peng, Z.; Chen, Z.; Chen, R. Point Cloud Deep Learning Network Based on Balanced Sampling and Hybrid Pooling. *Sensors* 2023, 23, 981.
16. Dang, J.; Yang, J. HPGCNN: Hierarchical Parallel Group Convolutional Neural Networks for Point Clouds Processing. In *Proceedings of the Asian Conference on Computer Vision*, Kyoto, Japan, 30 November–4 December 2020.
17. He, P.; Ma, Z.; Fei, M.; Liu, W.; Guo, G.; Wang, M. A Multiscale Multi-Feature Deep Learning Model for Airborne Point-Cloud Semantic Segmentation. *Appl. Sci.* 2022, 12, 11801.
18. Liu, Y.; Fan, B.; Xiang, S.; Pan, C. Relation-shape convolutional neural network for point cloud analysis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, 15–20 June 2019; pp. 8895–8904.
19. Hu, Q.; Yang, B.; Xie, L.; Rosa, S.; Guo, Y.; Wang, Z.; Trigoni, N.; Markham, A. Randla-net: Efficient semantic segmentation of large-scale point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, 13–19 June 2020; pp. 11108–11117.
20. Landrieu, L.; Simonovsky, M. Large-scale point cloud semantic segmentation with superpoint graphs. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 18–22 June 2018; pp. 4558–4567.

---

Retrieved from <https://encyclopedia.pub/entry/history/show/113761>