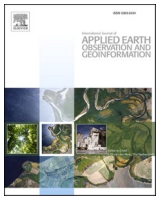
International Journal of Applied Earth Observations and Geoinformation 108 (2022) 102735

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

journal homepage: www.elsevier.com/locate/jag



A graph attention network for road marking classification from mobile LiDAR point clouds

Lina Fang [a], Tongtong Sun [a], Shuang Wang [a], Hongchao Fan [b,*], Jonathan Li [c,*]

[a] Academy of Digital China (Fujian), Fuzhou University, Fuzhou, FJ 350002, China
[b] Department of Civil and Environmental Engineering, Norwegian University of Science and Technology, Trondheim 7491, Norway
[c] Department of Geography and Environmental Management, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

ARTICLE INFO

Keywords:
MLS points clouds
Road marking classification
Graph neural network
Attention mechanism
Deep learning

ABSTRACT

The category of road marking is a crucial element in Mobile laser scanning systems' (MLSs) applications such as intelligent traffic systems, high-definition maps, location and navigation services. Due to the complexity of road scenes, considerable and various categories, occlusion and uneven intensities in MLS point clouds, finely road marking classification is considered as the challenging work. This paper proposes a graph attention network named GAT_SCNet to simultaneously group the road markings into 11 categories from MLS point clouds. Concretely, the proposed GAT_SCNet model constructs serial computable subgraphs and fulfills a multi-head attention mechanism to encode the geometric, topological, and spatial relationships between the node and neighbors to generate the distinguishable descriptor of road marking. To assess the effectiveness and generalization of the GAT_SCNet model, we conduct extensive experiments on five test datasets of about 100 km in total captured by different MLS systems. Three accuracy evaluation metrics: average *Precision*, *Recall*, and F_1 of 11 categories on the test datasets exceed 91%, respectively. Accuracy evaluations and comparative studies show that our method has achieved a new state-of-the-art work on road marking classification, especially on similar linear road markings like stop lines, zebra crossings, and dotted lines.

1. Introduction

Road markings play an important role in improving traffic safety in cities and highways and are considered as the core elements of a variety of applications, including intelligent transportation systems, navigation and positioning services, and smart cities. As a high-tech mapping technology, mobile laser scanning systems (MLSs) have capable of efficiently and accurately acquiring 3D spatial data and reflection intensity data of road scenes. Considering the high *retro*-reflective materials, road markings have a higher intensity value than the surrounding ground, which provides key information for separating road marking from the road surface. But, road marking after segmentation is generally incomplete and discontinuities due to the uneven intensity distribution, which is caused by scanning range, road surface wear, noise, and occlusion. Considering the influence of complexity of road scenes, various categories, levels of incompleteness, finely classifying road marking, particularly similar ones is challenging work.

Recently, some studies employed image-based deep learning frameworks like DBM in Yu et al. (2015), CNN in Wen et al. (2019), and U-net in Ma et al. (2021) on road marking classification tasks and have made some breakthrough achievements in small-size road marking classification. In these studies, large-size road marking like stop lines and solid lines were often separated in advance. Small-size road marking is then translated into fixed-size images, which results in some crucial and detail shape information being lost. Moreover, these deep learning models explore only the shape features associated with individual segments and not incorporate the topology, and semantics information among nearby road marking, which benefit for similar road marking classification like linear road marking.

As non-Euclidian data, graphs are widely utilized to build topology relationships between objects and provide an available representation for road marking. Concretely, graphs allow us to describe various associations between road markings like types, properties, and connections. While, it is challenging to deal with graph learning with varying sizes and topologies, especially in the domain of point clouds. Recent advances in graph neural networks (GNNs) and graph convolution networks (GCNs) extend traditional deep neural networks to graph domain and have achieved some excellent performance on non-

* Corresponding authors.
E-mail addresses: hongchao.fan@ntnu.no (H. Fan), junli@uwaterloo.ca (J. Li).

https://doi.org/10.1016/j.jag.2022.102735
Received 19 January 2022; Received in revised form 25 February 2022; Accepted 27 February 2022

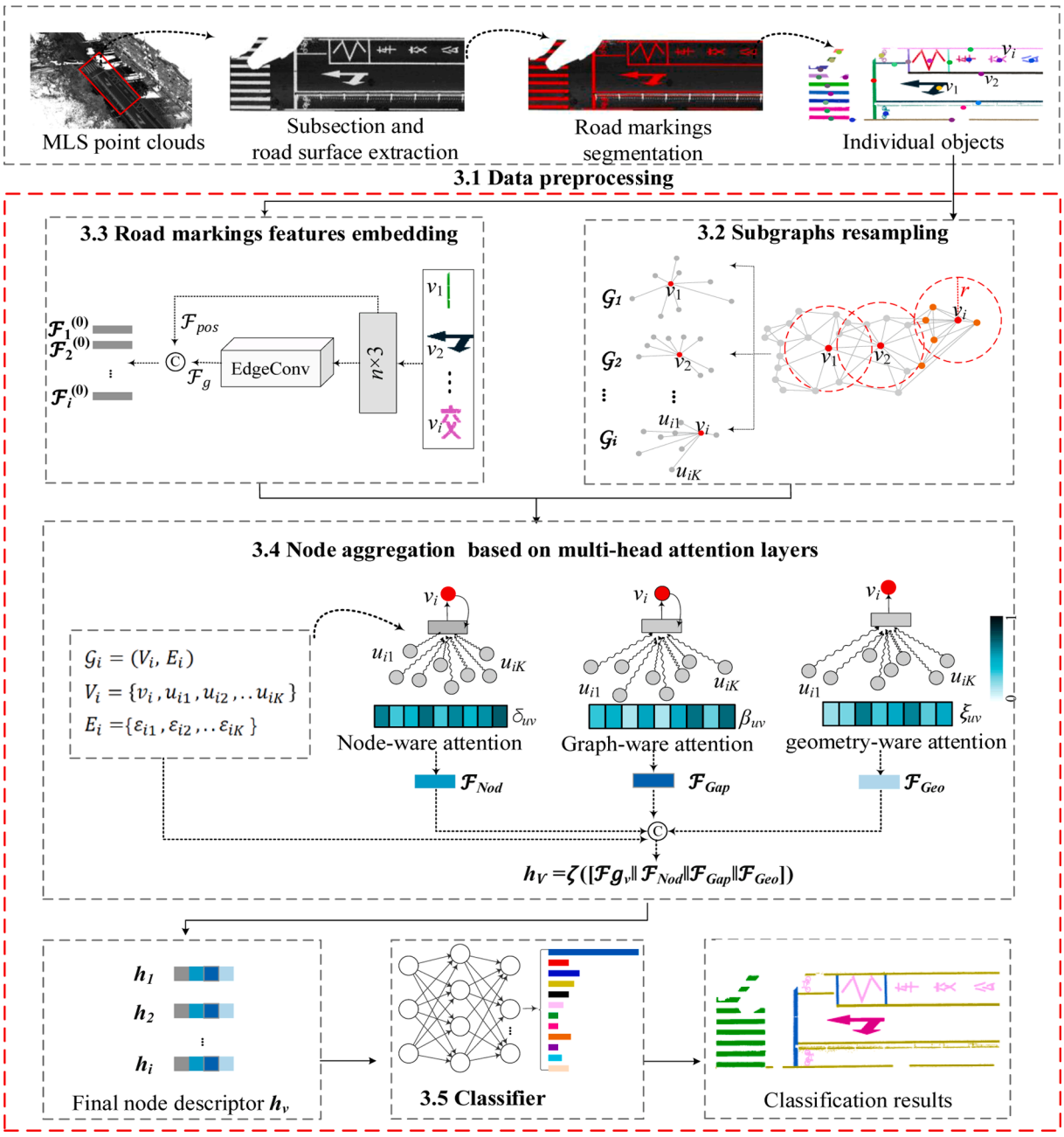


Fig. 1. Overall flowchart of our GAT_SCNet to classify road marking.

structured data like natural language processing (Kipf and Welling, 2016), social networks analysis (Wanda et al., 2021), and point clouds segmentation (Wang et al., 2019b). Inspired by these studies, we proposed a graph attention network on MLS point clouds called GAT_SCNet to address road marking classification issues, which can simultaneously group road marking into 11 categories including large and small size road marking on large scale road scenes. The main contributions of our method are listed as follows:.

- A computational subgraph resampling method suitable for road marking of various sizes.
- A new node aggregation algorithm based on multi-head attention mechanisms for automatically integrating local spatial distribution and semantic information in a unified paradigm.
- A new graph attention network for road marking classification, especially for similar linear road marking.
- Extensive experiments and discussion on the feasibility and generalization of the GAT_SCNet model for large-scale highways and complex urban scenes based on limited computer resources and training samples.

The rest of this paper is structured as follows. Section 2 provides an overview of state-of-the-art work about road marking classification on MLS point clouds, as well as related work of deep learning and graph neural network on point clouds. Section 3 details our GAT_SCNet models. The experimental results together with relevant discussions are presented in Section 4. Section 5 concludes with a summary of the findings.

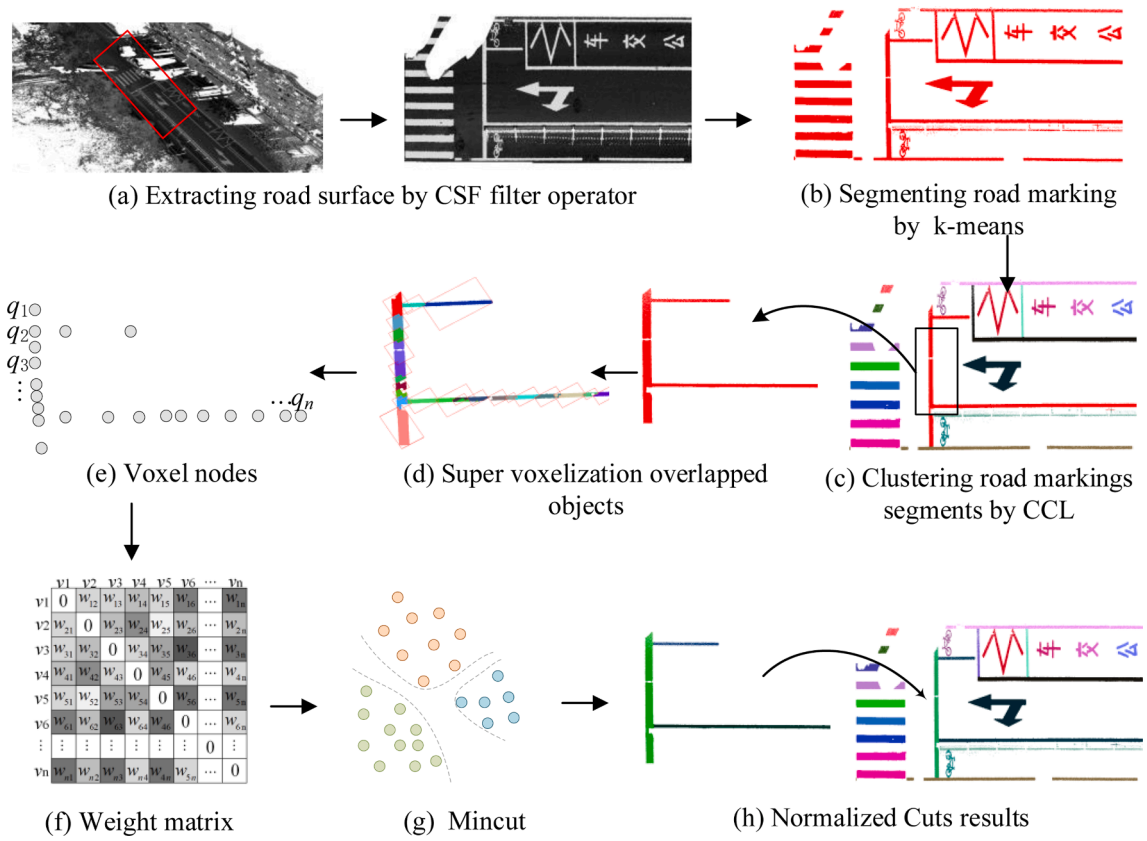


Fig. 2. Illustration of partitioning the overlapped road markings into individual objects based on Ncut algorithm.

2. Related work

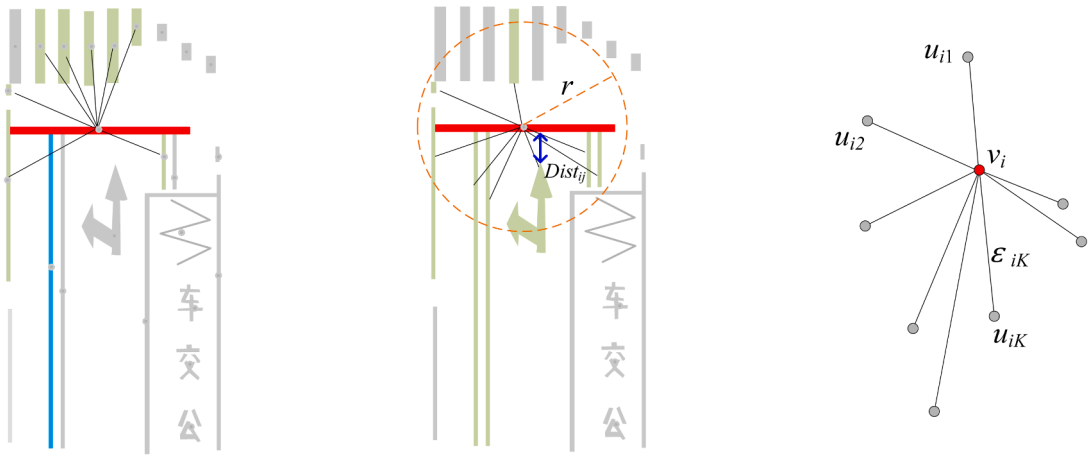
2.1. Road markings classification

As the majority of the road markings are linear or rectangular, some studies employed handcrafted features like size, area of the minimum bounding box of connected components, and orientation as shape signatures to group road marking into large-size and small-size road marking (Yang et al., 2012; Revilloud et al., 2013; Yang and Fang, 2014; Cheng et al., 2017). To further differentiate solid lines and stop lines belonging to large size road marking, most studies used trajectory data or curb lines as a guide (Yang et al., 2013; Revilloud et al., 2013; Wen et al., 2019). To distinguish small-size road marking, Yang et al. (2016) explored a hierarchical classification framework to classify road markings into arrows, rectangular marking, and other objects. Some research exploited semantic information like the repeated pattern or parallel relationship to distinguish zebra crossings, stop lines, and dotted lines (Guan et al., 2014; Kumar et al., 2014; Soilán et al., 2017; Yang et al., 2018). These studies had achieved good preformation on simple or small-scale road scenes, but they required predefined features. Recently, more and more research trends to utilize deep learning architectures to automatically retrieve the distinguishable features of road marking. After clustering the segment results into individual objects, Yu et al. (2015) designed a Deep Boltzmann Machine (DBM) model to extract the advanced features of road marking and successfully grouped small-size road makings into arrows, rectangular-shaped road marking, pedestrian warning markings, and other road marking. Wen et al. (2019) implemented a multi-scale clustering algorithm and stacked four-layer convolution neural networks (CNNs) to automatically capture the high-level features of road marking. Ma et al. (2021) proposed a novel hybrid capsule-based network containing a convolutional capsule network and a fully connected capsule to group connected segments into seven categories including large-size and small-size road marking. These methods all explored image-based deep learning architectures, which translated different sizes of road marking into the fixed-size images and will drop some key shape information. Meanwhile, these methods addressed the road marking segments individually rather than integrating local spatial distribution and semantic information in a unified paradigm, which helps in distinguishing similar road marking.

2.2. Deep learning models on 3D point clouds

As a pioneer, PointNet proposed by Qi et al. (2017a) was directly implemented on point clouds and opened the era of end-to-end deep learning on 3D point clouds. Recently, some research applied convolution-like operation on point clouds to encode the geometric relations between point and its neighborhood points. To preserve the direct relationship between points, the GeoConv layer introduced in Geo-CNN by Lan et al. (2019) was applied to decompose the edge features into different directions and then aggregate the features along each direction. Thomas et al. (2019) designed a flexible and deformable kernel point convolution named KPConv on point clouds to aggregate the features of point clouds.

Inspired by the idea of attention mechanism, some studies attempted to assign different attentional weights to different neighboring points by considering neighbors' geometry and attributes (Yang et al., 2015; Simonovsky and Komodakis., 2017; Qi et al., 2017; Veličković et al., 2018; Wang et al., 2019a; Mi et al., 2021; Zhou et al., 2021). As a pioneering work, Simonovsky and Komodakis (2017) proposed an edge-conditioned convolution (ECC) on graph-structured data, which conditioned weights according to the specific edge labels of nodes. Different from the ECC layer, graph attention network (GAT) by Wang et al. (2019a) automatically assigned different weights to neighbors through an attention mechanism layer, which better aggregated local neighborhood feature information. Wang et al. (2019b) proposed a dynamic graph CNN and implement it on point clouds and designed the EdgeConv



(a) k-NN subgraph resampling using nodes' center distance

(b) k-NN subgraph resampling using nodes' minimum distance in the local area with radius r

(c) Subgraph structure G_i=(V_i, E_i)

Fig. 3. Illustration of resampling subgraphs of road marking.

layer as a geometric convolution module to extract edge features, which indicated the distance between the point and its neighborhood points. Wen et al. (2021) proposed a graph attention convolution (GAC) on point clouds and enhanced the most relevant part of the neighborhood. Similarly, Chen et al. (2021) exploited a GAP Layer to learn the fine-grained feature of point clouds, which embedded a graph attention mechanism within multi-layer perceptron (MLP) layers to highlight the importance of neighbors. Additionally, a multi-head GAP Layer was utilized to specify different importance of neighbors and obtain sufficient contextual attention features, which gained state-of-the-art performance shape classification and part segmentation tasks on ModelNet40 (Vishwanath et al., 2009) and ShapeNet (Chang et al., 2015) datasets.

Overall, deep learning architectures have been achieved breakthrough performance for 3D point cloud segmentation and classification, but graph attention networks for point clouds are still in their infancy. Most graph attention models on point cloud were implemented on public datasets for theoretical analysis and few were implemented on large-scale MLS datasets and real-world environments, which is very important to evaluate the capability and effectiveness of deep learning methods.

3. Method

The target of our GAT_SCNet is to group individual segments into corresponding categories, which considers individual segments as nodes and aggregates topology and semantic relationships among road markings through the graph neural network. To fulfill this task, we stack four main modules in GAT_SCNet (illustrated in Fig. 1): (1) subgraphs sampling, (2) road markings embedding, (3) node aggregating based on multi-head attention layers, (4) classifier with an improved loss function. Before this, we use an unsupervised segmentation approach to separate and cluster road marking points into individual objects.

3.1. Data preprocessing

Considering the efficiency and unavailable large amount of point-based public labeled datasets for road marking segmentation, we choose the unsupervised method instead of the deep learning method for road markings segmentation. To quickly separate the road marking from MLS point clouds, we partition the whole data into a series of subsections and extract road surface based on the CSF filter operator (Zhang et al., 2016). Inspired by the works in Guan et al. (2014), we implement a multi-threshold k-means operator to segment road marking points and then group them into individual components by connected component labeling (CCL). As using CCL may yield one segment containing several road markings, like the stop lines and solid lines shown in Fig. 2(c), we then voxelized the overlapped segments by voxel size s_c and explore the voxel-based Normalized cut (Ncut) algorithm (Fang et al., 2021) to separate the overlapped road marking. In Ncut, overlapped segments were partitioned into two parts by minimizing the dissimilarity within the same group as well as maximizing the discrepancy between two different groups. Hence we calculate the similarity denoted weights w_{ij} between two voxels q_i and q_j according to geometric distribution (see Equation 1).

$$w_{ij} = \begin{cases} \exp\left(-\frac{||q_i - q_j||_2^2}{\sigma_D^2}\right) \cdot \exp\left(-\frac{\arccos^2\left(v_{q_i}.v_{q_j}\right)}{\sigma_A^2}\right), \text{if } ||q_i - q_j||_2 \le d_N 0 \end{cases} \tag{1}$$

where $||q_i - q_j||_2^2$ denoted the horizontal distance between voxel q_i and q_j. $\arccos^2(v_{q_i}.v_{q_j})$ is the intersection angle of two principal directions v_{q_i} and v_{q_j} of voxel q_i and q_j. σ_D and σ_A are the standard deviation of horizontal distances and intersection angles, respectively. d_N is the influence radius of voxel similarity. Then, we implement the optimized solution of Mincut by Shi and Malik (2000) to recursively divide overlapped road markings into individual road marking.

3.2. Subgraphs resampling

Considering computation resources and applicability, it is more challenging for GNNs to predict and train graphs with arbitrary or unseen shapes (Hamilton et al., 2017). Hence, we convert a large graph into a series of subgraphs, which resample the fixed number neighbors of road marking through k-Nearest Neighbors (k-NN) search. Center distances metrics are usually used to search the nearest neighbors. While, this sampling way may accomplish small-size road marking, but cannot cover large-size road marking. As the case in Fig. 3(a), the blue solid line is crucial for red stop line recognition, but it is not within the neighboring domain of the stop line according to their pairwise center distance. To overcome this problem, we adopt a constrained sampling strategy based on minim-distance between the center road marking and its neighbors in the radius r illustrated in Fig. 3(b). In our constrained sampling strategy, if a road segment has points within the radius of the center road marking, we consider it as a candidate neighbor. Then, we

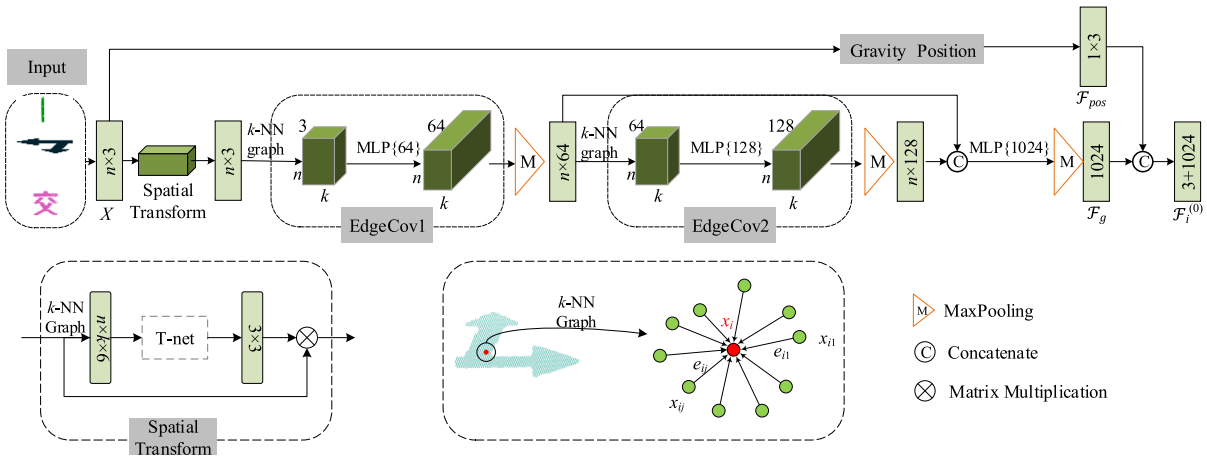


Fig. 4. Initial road marking embedding using DGCNN framework.

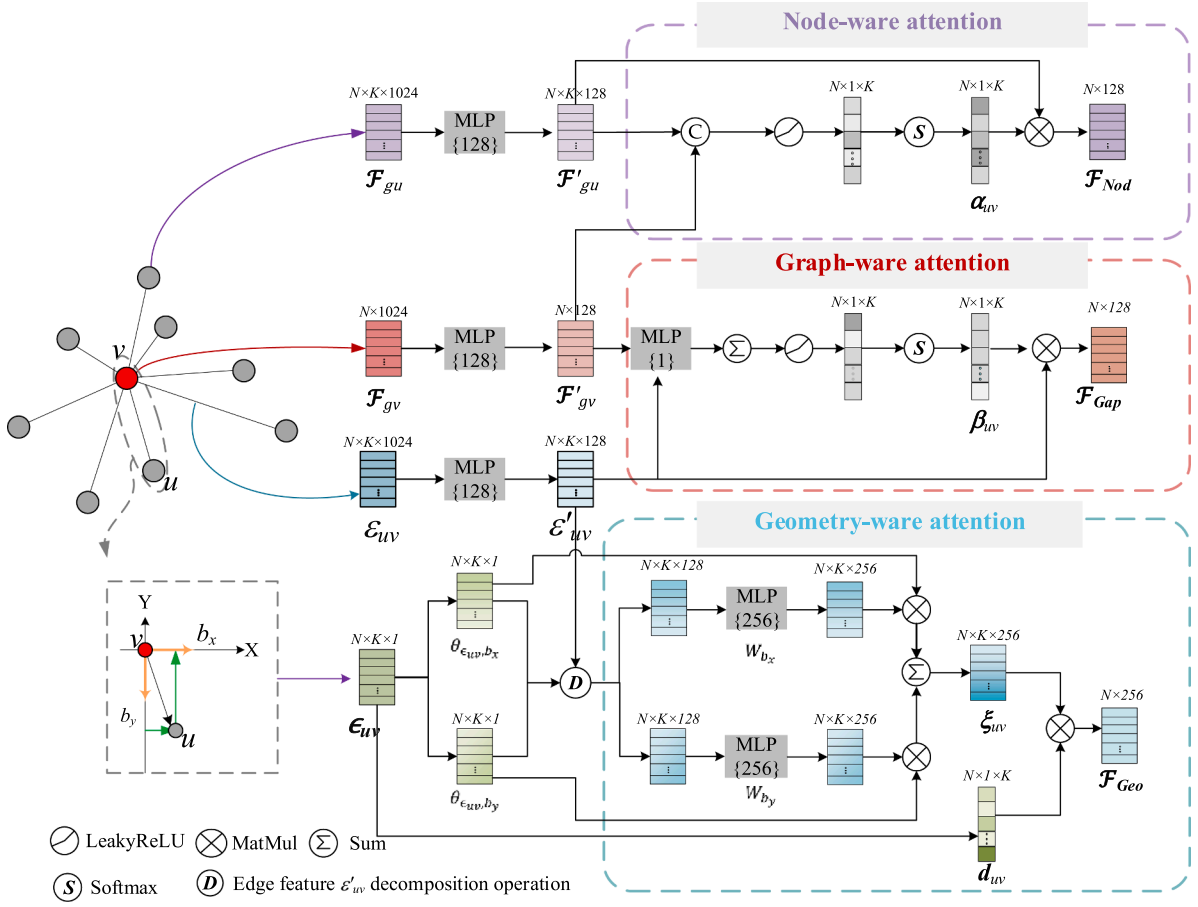


Fig. 5. Node aggregation module of multi-head attention mechanism.

compute the minim-distance $Dist_{ij}$ between the center node v_i and candidate neighbor u_{ij}, $j = 1, 2, \cdots, n_c$, n_c is the number of candidate neighbors. Minim-distance $Dist_{ij}$ is the minimum point-wise distance ($Dist_{ij} = \min\{||\boldsymbol{p} - \boldsymbol{q}||\}$), where $\boldsymbol{p}$ and $\boldsymbol{q}$ denote the points that belong to the center node v_i and candidate neighbor u_{ij}, respectively. Moreover, we sort the minim-distances and select K number of nearest neighbors as the desirable neighbors. In this way, we can accomplish more significant neighbors and convert the large graph into a series of subgraphs. Let $\mathscr{V} = \{v_1, v_2, ..., v_m\}$ be a set of road marking to be classified, the i-th subgraph $\mathscr{G}_i = (V_i, E_i)$ is constructed from the i-th given road marking named center node v_i and its neighbors $\mathscr{N}_v^{\tau}(i)$, where $V_i = \{v_i, u_{i1}, u_{i2}, ..., u_{iK}\}, \forall u_{iK} \in \mathscr{N}_v^{\tau}(i)$ and $E_i = \{\varepsilon_{i1}, \varepsilon_{i2}, ..., \varepsilon_{iK}\}$ denote the nodes and edges in the k-nn subgraph, respectively. This localization processing is easily parallelizable and independent of input data size, which makes it suitable for large-scale road scenes.

3.3. Road marking embedding module

In general, different kinds of road marking vary in shape and size. Therefore, learning the discriminative features associated with road

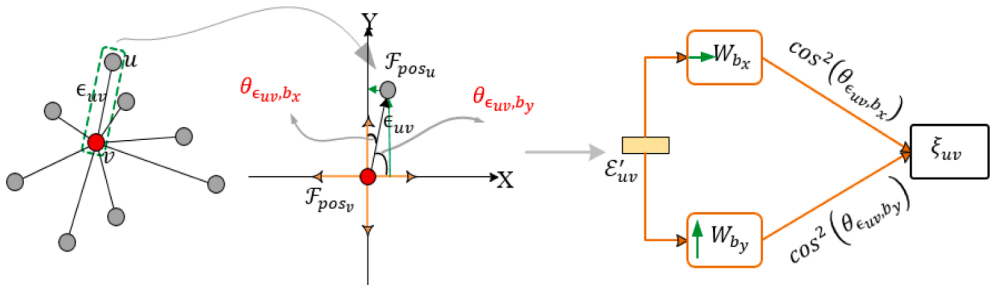


Fig. 6. Combination position vector ϵ_{uv} and edge feature $\mathscr{E}'_{uv}$ to calculate geometric structure feature.ξ_{uv}

marking plays an important role in road marking classification. In the paper, we explore the DGCNN framework (Wang et al., 2019b) to extract high-level features as initial road markings embedding (see in Fig. 4). An F-dimensional point cloud with n points in a road marking could be denoted by $X = \{x_1, x_2, \cdots, x_n\} \subseteq \mathbb{R}^F$. Because the DGCNN network can be irectly operated on point clouds, we take each point in 3D coordinate ($F = 3$) as input.

In DGCNN, the spatial transform block is used to align the input X into a canonical space for maintaining permutation invariance of road marking. Then, EdgeConv layers are implemented to capture the local geometric features named edge features e_{ij} as follows:

$$e_{ij} = \gamma\big((x_i, x_{ij} - x_i), \Theta\big) \tag{2}$$

Where x_i, x_{ij} denote the local geometric features of point i and it's j-th neighboring points j. $\gamma(\bullet)$ is a nonlinear edge function, and Θ is the corresponding coefficient, which is learned by the shared MLP layers with the ReLU activation function. We then use a max-pooling layer as a channel-wise symmetric aggregation operation to yield the output $x_i^{(l+1)}$ of the l-th layer of EdgeConv.

$$x_i^{(l+1)} = \max_{j \in k} \mathrm{ReLU}(e_{ij}^{(l)}) \tag{3}$$

Since different EdgeConv layers capture different local geometric features, we concatenate the outputs of two EdgeConv blocks and use a max-pooling layer to generate the global descriptor $\mathscr{F}_g$ of road marking. To maintain spatial differences, we concatenate the gravity positions of road marking $\mathscr{F}_{pos}$ with global descriptor $\mathscr{F}_g$ as the embedding features $\mathscr{F}^{(0)}$ of each road marking.

$$\mathscr{F}_g = \max_{i \in n}([x_i^{\mathrm{EdgeConv1}} \| x_i^{\mathrm{EdgeConv2}}]) \tag{4}$$

$$\mathscr{F}_i^{(0)} = \big[\mathscr{F}_g(i) \| \mathscr{F}_{pos}(i)\big], i \in [1, m] \tag{5}$$

3.4. Node aggregation based on multi-head attention layers

Because road marking embedding mainly represents the shape characteristic of road marking, they can not describe the local structure and semantic information between road marking. Hence, we propose a shared multi-head graph attention module to encode the local structure and semantic information into the hidden representations $\boldsymbol{h}_v$ of road marking. As illustrated in Fig. 5, the graph attention module consists of three heads: *node-ware attention head*, *graph-ware attention head*, and *geometry-ware attention head*, which dramatically assign different weights to neighbors and explore the crucial information for road marking recognition. In each subgraph, we briefly define node and its neighbors as v, and u ($u \in \mathscr{N}_v^\tau$), respectively.

(1) Node-ware attention head

Generally, the type of road marking is easier predicted by its neighbors with similar shape. The impact of this head is to pay more attention to the neighbor, which is similar to the central node. In more detail, the node-ware attention layer learns the node-coefficients δ_{uv} represent the importance of the neighbor u to the node v according to the similarity of their global features $\mathscr{F}_g$ of road marking, which mainly encodes the shape characteristic of road marking. To gain fine-grained feature, we firstly encode nodes' global feature $\mathscr{F}_g$ into a hyperbolic space with higher representational feature $\mathscr{F}'_g$, and then interpolate center node and neighbors' higher features to learn the node-coefficientsδ_{uv}.

$$\mathscr{F}'_g = h(\mathscr{F}_g, \theta) \tag{6}$$

$$\delta_{uv} = \alpha\Big(\mathscr{F}'_{g_u}, \mathscr{F}'_{g_v}\Big) = \mathrm{LeakyReLU}\Big(\vec{\mathbf{a}}^T\Big[\mathscr{F}'_{g_u} \| \mathscr{F}'_{g_v}\Big]\Big) \tag{7}$$

Where $h(\bullet)$ is a nonlinear function and θ is the corresponding parameters set. LeakyReLU($\bullet$) denotes nonlinear activator leakyReLU in the MLP layer. $\mathbf{a}^T$ is the weight vector parameterizing the node-attention mechanism α. As different subgraphs vary in size and scale, the node-attentional coefficient δ_{uv} are normalized by a softmax function for easily comparable across different neighbors.

$$\alpha_{uv} = \mathrm{softmax}(\delta_{uv}) = \frac{\exp(\delta_{uv})}{\sum_{u \in \mathscr{N}_v} \exp(\delta_{uv})} \tag{8}$$

Then, we combinate the normalized self-attention weights α_{uv} and corresponding neighborhood's features $\mathscr{F}'_{g_u}$ to calculate the node-ware features $\mathscr{F}_{Nod}$ of center node v.

$$\mathscr{F}_{Nod} = f\left(\sum_{u \in \mathscr{N}_v^\tau} \alpha_{uv} \mathscr{F}'_{g_u}\right) \tag{9}$$

Where $f(\bullet)$ denote the nonlinear aggregation function.

(2) Graph-ware attention head

The node-attention mechanism tends to assign more weight to neighbors with similar shape features but drops spatial correlation or topology between the road markings in a subgraph. To exploit the topological relationship between the node and neighbors, we embed a GAP layer introduced in Chen et al., (2021) to capture contextual information of subgraph, dubbed graph features $\mathscr{F}_{Gap}$. In order to highlight different attention weights on neighborhood, we consider the self-geometric information of node denoted by node features $\mathscr{F}_{g_v}$ and local correlations to corresponding neighbors denoted by edge features $\mathscr{E}_{uv}$. In more detail, we combine the center node features $\mathscr{F}_{g_v}$ and edges features $\mathscr{E}_{uv}$ in a subgraph to introduce the normalized graph-ware attention coefficients β_{uv}.

$$\mathscr{E}'_{uv} = h(\mathscr{E}_{uv}, \theta) = h\Big((\mathscr{F}_{g_u} - \mathscr{F}_{g_v}), \theta\Big) \tag{10}$$

$$\beta_{uv} = \frac{\exp\Big(\mathrm{LeakyReLU}\Big(h\Big(\mathscr{F}'_{g_v}, \theta\Big) + h\big(\mathscr{E}'_{uv}, \theta\big)\Big)\Big)}{\sum_{u \in \mathscr{N}_v} \exp\Big(\mathrm{LeakyReLU}\Big(h\Big(\mathscr{F}'_{g_v}, \theta\Big) + h\big(\mathscr{E}'_{uv}, \theta\big)\Big)\Big)} \tag{11}$$

Table 1
Brief information of training datasets.

Information	Training Dataset I	Training Dataset II	Training Dataset III	Training Dataset IV
Road scene	Downtown	Suburb	Downtown	Urban
MLS system	Hidarget Hiscan	Trimble	Optech Teledyne	RIEGL VMX-450
Location	Fuzhou China	Beijing China	Toronto Canada	Fuzhou China
Road Length (km)	0.72	0.74	1.00	1.20

Then we take attention coefficients β_{uv} as the graphical masked attention and aggregated the edges features $\mathscr{E}'_{uv}$ to output the graph contextual attention feature $\mathscr{F}_{Gap}$ of center node v:

$$\mathscr{F}_{Gap} = f\left(\sum_{u\in\mathscr{N}_v^r}\beta_{uv}\mathscr{E}'_{uv}\right) \tag{12}$$

(3) Geometry-ware attention head

Different from node-ware and graph-ware attention blocks, Geometry-ware attention block mainly highlights orientation relation-

Fig. 7. Four training datasets and some training samples. For each training dataset, top-to-bottom figures show the training datasets' location in Google Earth, raw MLS point clouds colored by intensity information. To illustrate clearly, some training samples are listed at the bottom.

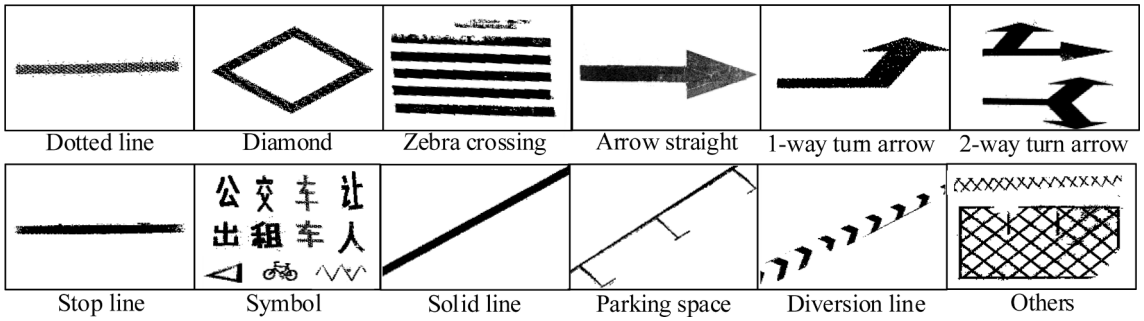


Fig. 8. An illustration of 12 categories of road markings.

Table 2
The optimal parameter configurations.

Notation	Description	Setting
S_c	The size of the voxel in Ncut	0.4 *m*
d_N	The horizontal influence radius of the similarity in Ncut	1 *m*
Mincut	The maximum tolerance allowed in Ncut	0.03
r	The radius of neighborhood area *r* in subgraph sampling	12 *m*
K	the number of the neighbors in a subgraph	8
k	The number of nearest points *k* in EdgeConv layer	20

ships like directionality or arrangement in the subgraph, which is very important to distinguish similar road markings like dotted lines and zebra crossings. Inspired by the Geoconv operator proposed in Monti et al. (2017), we apply the geometry-ware attention block to combinate the position information $\mathscr{T}_{pos}$ and edge features $\mathscr{E}'_{uv}$ on the local coordinate systems shown in Fig. 6. We take the gravity position, the main direction, and the normal direction of center road marking (center node *v*) as the origin, X, and Y direction, respectively. Then, we decompose the edge vector ϵ_{uv} on two orthogonal basis b_x and b_y, and compute corresponding angles $\theta_{\epsilon_{uv},b_i}$. Guiding by direction angle $\theta_{\epsilon_{uv},b_i}$, we further decompose the edge features $\mathscr{E}'_{uv}$ along with the two bases and then aggregate them as the directional feature ξ_{uv}.

$$\epsilon_{uv} = \mathscr{T}_{pos_u} - \mathscr{T}_{pos_v} \tag{13}$$

$$\theta_{\epsilon_{uv},b_i} = \langle \epsilon_{uv}, b_i \rangle b_i \in (b_x, b_y) \tag{14}$$

$$\xi_{uv} = \sum_{b_i \in B} \cos^2(\theta_{\epsilon_{uv},b_i}) W_{b_i} \mathscr{E}'_{uv}, \tag{15}$$

Where ϵ_{uv} is defined by the gravity position of node *v* and neighbors *u*. W_{b_i} denotes the direction-associated weight matrices and is a learned parameter in the geometry-ware attention layers.

Generally, the closer the neighbor node to the center node, the more important. To highlight the proximity from different orientations of road markings, we further aggregate the directional feature ξ_{uv} according to Euclidean distance d_{uv} between the center node *v* and its neighbor *u*, and extract geometry-ware attention feature $\mathscr{T}_{Geo}$ as follows:

$$\mathscr{T}_{Geo} = \frac{\sum_{u \in \mathscr{N}_v^r} d_{uv} \xi_{uv}}{\sum_{u \in \mathscr{N}_v^r} d_{uv}} \tag{16}$$

$$d_{uv} = (r - \|\mathscr{T}_{pos_u} - \mathscr{T}_{pos_v}\|^2) \tag{17}$$

After learning three attention features $\mathscr{T}_{Nod}$, $\mathscr{T}_{Gap}$, and $\mathscr{T}_{Geo}$, we concatenate the three independent attention features and the original global feature $\mathscr{T}_{g_v}$ and produce the nonlinearity aggregation as the final descriptor $\boldsymbol{h}_v$ of node *v*.

$$\boldsymbol{h}_v = \mathrm{ReLU}\left(\left[\mathscr{T}_{g_v} || \mathscr{T}_{Nod} || \mathscr{T}_{Gap} || \mathscr{T}_{Geo}\right]\right) \tag{18}$$

Where ReLU(•) denotes the ReLU activator in the final MLP layer, and || denotes the vector concatenation operation.

3.5. Loss function

Once the final descriptor $\boldsymbol{h}_v$ of road marking has been obtained, we stack an MLP layer and a softmax layer as the classifier to predict its type. In road scenes, the number of dotted lines and zebra crossings accounts for a large proportion, while the number of arrows and text signs is less. It means that the GAT_SCNet model will pay more attention to the classes with more number (e.g., dotted lines) but ignore the classes with fewer numbers (e.g., text), which result in lower accuracy for these

Table 3
Brief information of test datasets.

Information	Test Dataset I	Test Dataset II	Test Dataset III	Test Dataset IV	Test Dataset V
Road scene	Downtown	Downtown	Urban	Highway	Highway
MLS system	Optech Lynx	RIEGL VMX-450	RIEGL VMX-450	StreeMapper 360	Chchav Alpha3D
Location	Hengyang China	Xiamen China	Fuzhou China	Schwyz Switzerland	Hangzhou China
Road Length(km)	1.25	0.44	16.48	23.71	48.32
Dotted line	400	26	1383	3377	5312
Diamond	28	6	28	–	–
Zebra crossing	403	144	931	47	–
Straight arrow	5	4	39	81	67
1-way turn arrow	12	–	29	–	46
2-way turn arrow	2	3	88	5	9
Stop line	10	11	145	–	192
Symbol	–	21	174	1	307
Solid line(m)	3,070.53	1,967.70	16,485.08	100,573.08	211,056.92
Parking space	–	–	72	–	–
Diversion line	–	–	4	30	62

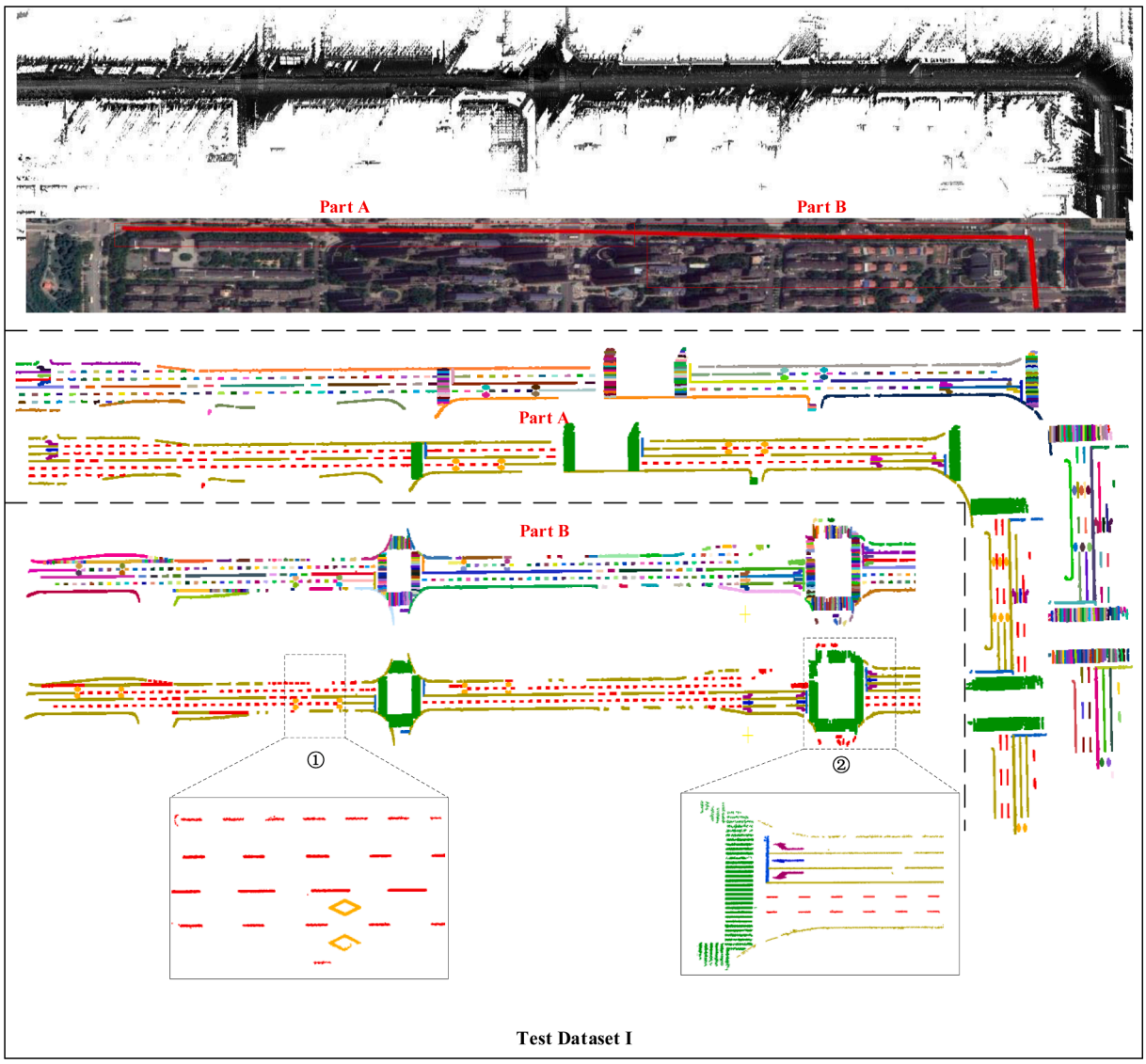


Fig. 9. Road marking classification results of test dataset I denoting downtown. In each part, we presented the test dataset's location in Google Earth, ground points displayed by intensity, the road marking segmentation indicated by different colors, and corresponding classification results colored by category, respectively.

classes. To overcome the influence of the uneven distribution of categories on the performance of road marking recognition, we use the median frequency balance strategy (Eigen and Fergus, 2015) to weight the loss of each class in GAT_SCNet. In the median frequency balance strategy, we pay more attention to the categories with smaller numbers and define the loss function $\mathscr{L}$ as follows:.

$$\mathscr{L} = -\frac{1}{N}\sum_{n=1}^{N}\sum_{c=1}^{C} l_c^{(n)} \log\left(p_c^{(n)}\right) w_c \tag{19}$$

$$w_c = \frac{\text{Median}\left(\left\{\frac{N_c}{N} | c \in C\right\}\right)}{\frac{N_c}{N}} \tag{20}$$

Where N is the number of individual road marking in each Mini-batch. C is the set of all categories. $p_c^{(n)}$ is the prediction of GAT_SCNet, denoting the probability distribution of the n-th objects in subclass c, and $l_c^{(n)}$ is the ground truth, a one-hot probability distribution over the C classes. w_c is the weight of subclass c. N_c denotes the number of subclass c in the total samples N.

4. Results and discussion

4.1. Training datasets and samples

For lack of available public MLS samples for road marking classification, we chose four datasets captured by different mainstream MLS systems to build the training corpus. The detailed information about the training datasets is listed in Table 1. All training datasets are common urban scenes and provide a large number of road markings in various shapes, sizes, and incompleteness (see Fig. 7). As seen from the bottom of Fig. 7, we partitioned training datasets into a set of subsections about 150 m in length along road direction and disturbed these subsections through some typical transformation operations including rotation, shaking, marking incompleteness to enlarge the training samples. Moreover, we empirically selected 5500 road sections as the training samples, of which 4000 were randomly selected for training and the rest for verification. In addition, road markings were manually labeled into 12 categories: dotted line, diamond, zebra crossing, arrow straight, 1-way turn arrow, 2-way turn arrow, stop line, symbol (text, triangle, bicycle), solid line, parking space, diversion line and others (shown in Fig. 8).

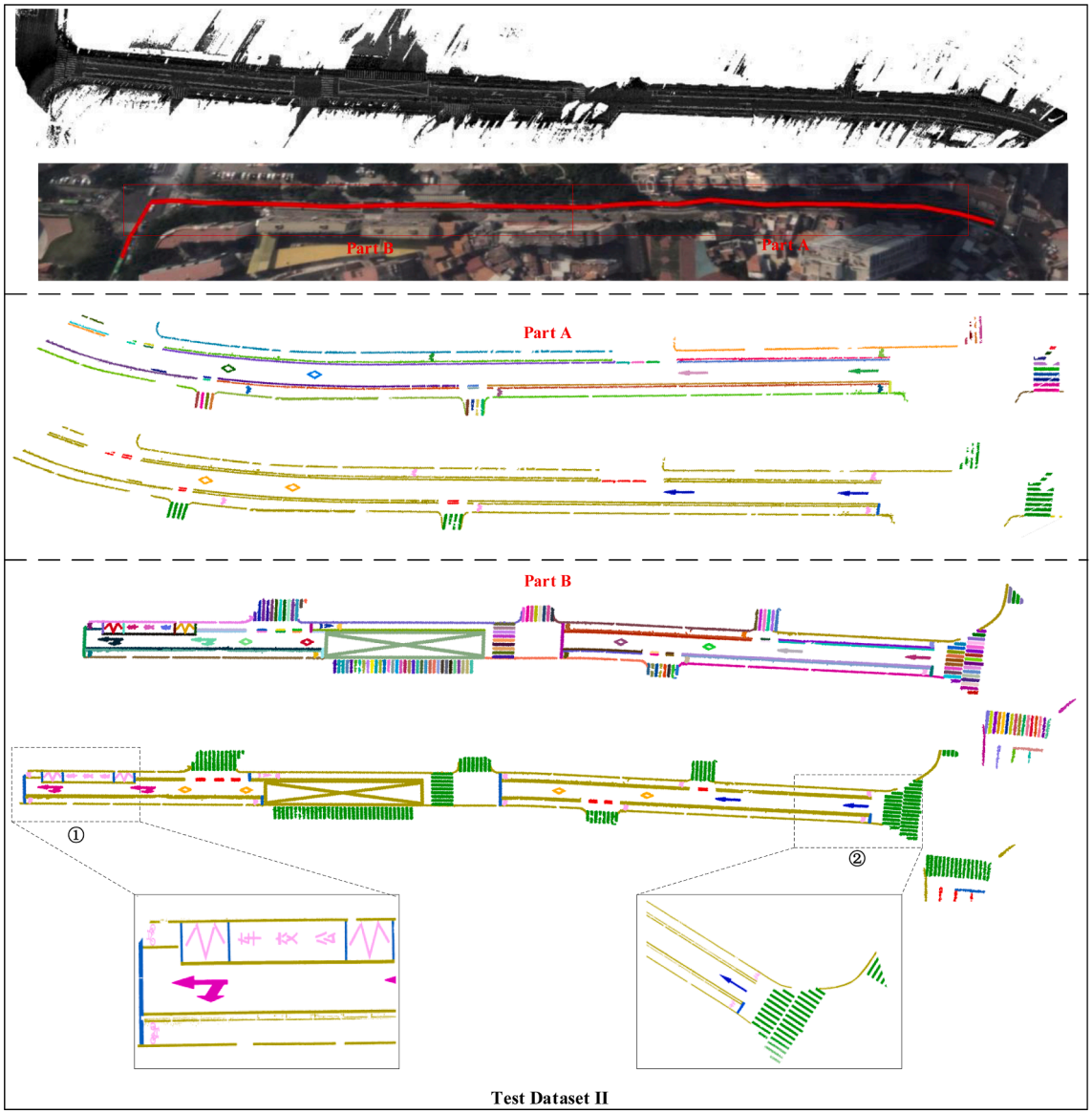


Fig. 10. Road marking classification results of test dataset II denoting downtown. In each part, we presented the test dataset's location in Google Earth, ground points displayed by intensity, the road marking segmentation indicated by different colors, and corresponding classification results colored by category, respectively.

4.2. Parameters configuration and turning

In Section 3.1, the voxel size S_c and the horizontal influence radius of the similarity d_N affected the performance of subdividing the overlapped road marking into individual parts. Considering the size and shape of road marking, we assigned the values of S_c and d_N as $0.4m$ and $1m$, respectively. In Section 3.2, selecting appropriate values of the radius of neighborhood area r, and the number of the neighbors K means a reliable and distinguishable representation of the local relationship of road marking. Considering the shape and semantic rule of road marking, we conducted extensive experiments on training datasets to configure parameters r and K as 12 and 8, respectively. As we used the pre-trained DGCNN model to initial the EdgeConv layers in Section 3.3, the number of nearest points k was set as the default configuration with 20 for all EdgeConv layers (shown in Table 2).

After selecting the optimal hypermeters, we implemented all training datasets to train the GAT_SCNet model until convergence. The GAT_SCNet model was conducted based on the TensorFlow framework. All experiments were run on the environment as NVIDIA GeForce GTX 1060 3 GB, Python 3.5, TensorFlow GPU 1.8.0, CUDA 9.0, and Cudnn 7.0 on window 10.0. During the training stage, we used the random gradient descent method (SGD) and driving quantity (Adam) optimizer with the initial learning rate of 0.001 to train the GAT_SCNet model. Due to the limitation of GPU ability, the mini-batch size, the decay rate, epoch, and the momentum were set to 32, 0.98, 500, and 0.5 respectively.

4.3. Method validation

To validate the performance of the proposed method, we selected five MLS point clouds including three complex urban scenes and two highways in different cities. These datasets were acquired by mainstream MLS systems including Optech Lynx, Riegl VMX-450, StreetMapper 360, Chchav Alpha3D. In addition, test datasets III, IV, and V are large-scale MLS point clouds. The brief information and the ground truth of categories of the test datasets were listed in Table 3.

After removing the off-ground points, we used the unsupervised segmentation operators in Section 3.1 to separate and cluster the road marking from ground point clouds. Then, road markings segments of test datasets are directly fed into the trained GAT_SCNet model to predict the

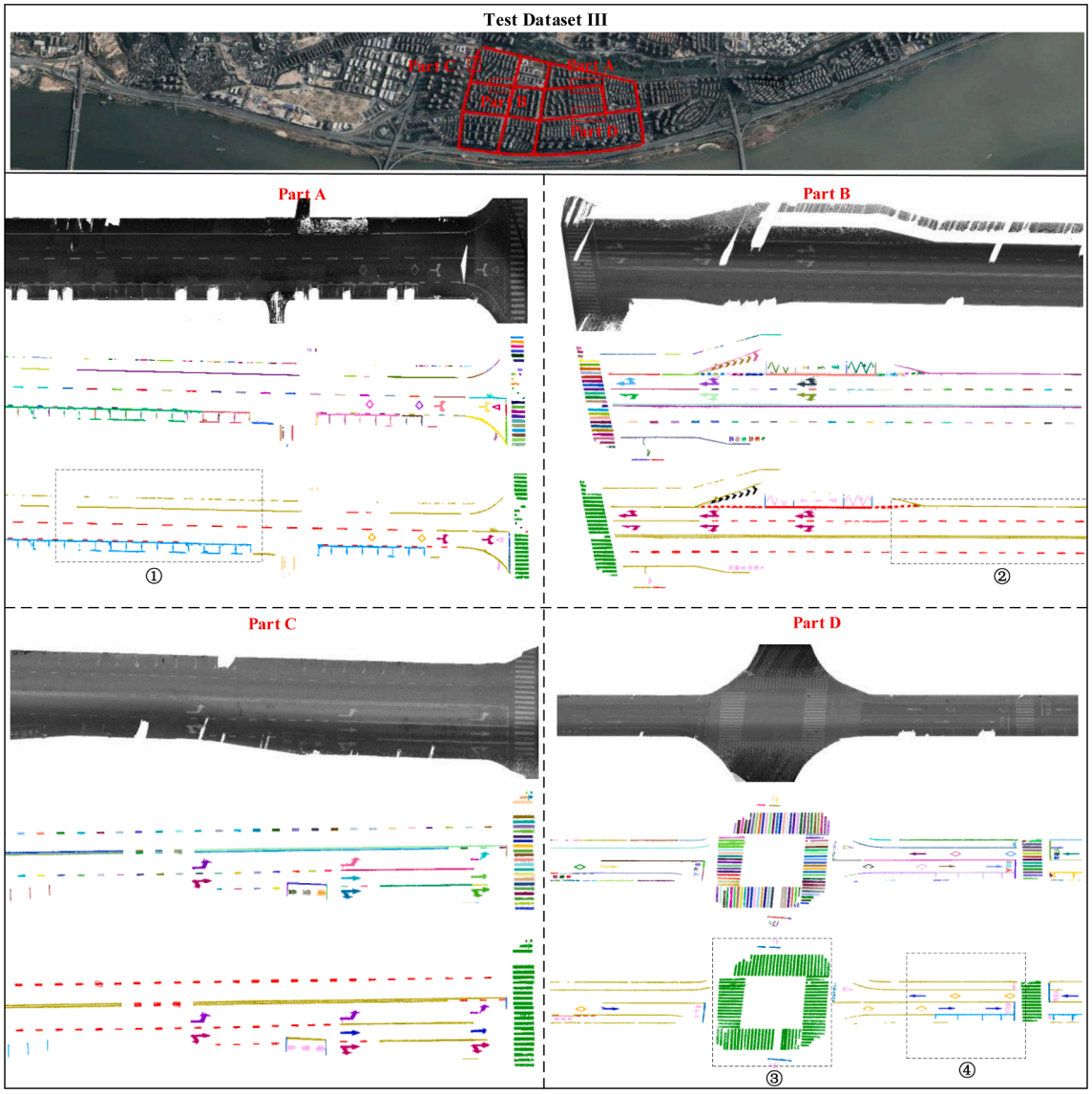


Fig. 11. Road marking classification results of test dataset III denoting urban. The top figure is the location of test dataset III in Google Earth. Part A-D is the selected road scenes. For each selected road scene, we present the ground points displayed by intensity, the road marking segments indicated by different colors, and the corresponding classification results colored by category, respectively.

category. Because some test datasets cover large-scale road scenes, we selected some road sections with road marking segments results and the corresponding road marking classification results of the GAT_SCNet model in Figs. 9-13.

The visualization results in Figs. 9-13 illustrated that GAT_SCNet has successfully grouped the majority of road markings in all test datasets. On the urban scenes, many solid lines and zebra crossings were incomplete for over-segmentation or under-segmentation, especially in test dataset III, the dotted lines can easily cause category confusion. However, our method succeeded to distinguish the dotted lines from the zebra crossings (test datasets I, II, and III) and achieved outstanding results on incomplete ones. In addition, some incomplete road markings caused by wear and occlusion (e.g., the diamond in test dataset I-①, over-segmented solid lines in test dataset III-①) were also correctly recognized, which indicated that the GAT_SCNet model had the capability of capturing the high-level features and overcoming the influence of over-segmentation or missing neighborhood nodes. On urban and highway scenes, solid lines are common and vary in shape, especially on the roundabout of test datasets IV and V. Benefiting from the subgraph structures, the GAT_SCNet model assigned the correct label for most solid lines like the straight lines on test datasets I and III, curve lines on test datasets IV and V. Meanwhile, it overcame the influence of large-scale road marking (e.g., long solid lines) on small-scale road marking (e.g., dotted lines) and simultaneously categorized them. Especially, GAT_SCNet gained exciting results on the unseen road markings and road scenes, such as stop lines with the special layout in test dataset V-Part A, Chinese texts in test datasets V-Part C and V-Part D, and roundabouts with circle lines in test datasets IV and V-Part B, which shows that our proposed GAT_SCNet model has strong generalization ability to different MLS datasets and various road scenes.

In addition, we quantitatively evaluated the road marking classification results on test datasets with three accuracy scores *Precision* $\left(P = \frac{TPs}{TPs+FPs}\right)$, *Recall* $\left(R = \frac{TPs}{TPs+FNs}\right)$, and F_1 $\left(F_1 = \frac{2\times P\times R}{P+R}\right)$ and corresponding average scores of each category. As the scope of our proposed method is to finely group the road marking, we only calculate the types of road marking. Hence, we define *TPs*, *FPs* and *FNs* as the numbers of correctly classified road marking, unclassified road marking, and incorrectly classified road marking, respectively, and calculate accuracy scores for

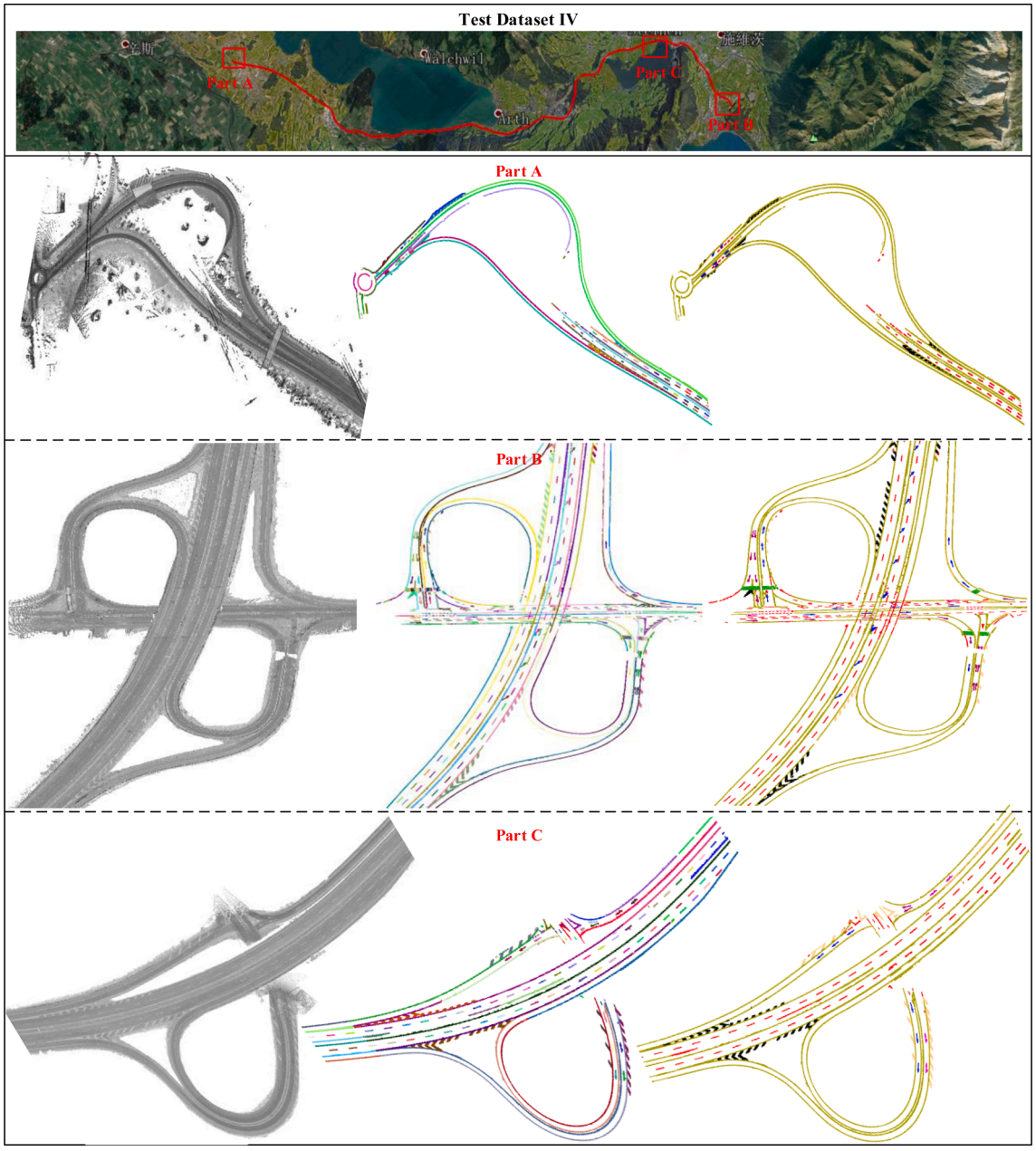


Fig. 12. Road marking classification results of test dataset IV denoting highway. The top figure is the location of test dataset IV in Google Earth. Part A-C are selected roundabouts. For each selected road scene, we present the ground points displayed by intensity, the road marking segments indicated by different colors, and the corresponding classification results colored by category, respectively.

each test dataset (see Table 4). We manually labeled the ground truth listed in Table 3 based on the results of road markings segmentation. Overall, the GAT_SCNet model gained F_1 over 91% for all categories and 100% for 1-way turn arrow, 2-way turn arrow, and diversion line on all test datasets. In terms of linear road marking, the GAT_SCNet model also gains high F_1, 93.42% for dotted line, 95.64% for zebra crossing, 91.09% for stop line, and 98.23% for solid line. Although test dataset III are complex and large, our GAT_SCNet model achieves good results with F_1 over 94%.

Although most road markings were correctly grouped, little road marking are still misclassified due to incompleteness or unfavorable significant neighbors, such as the misclassified stop line shown in Fig. 14 (a) and misclassified solid lines shown in Fig. 14(b). Compared with other road marking, the accuracy scores of stop lines were relatively lower and gained an average F_1 of 91.09%. The main reason is the number of stop lines in test datasets I and II were small. Hence, little misclassified stop lines will significantly drop accuracy scores. The lower *Precision* of dotted lines in test dataset II suffers for the same reason. In test dataset V, some overlapped stop lines were misclassified as solid lines in the deceleration zones illustrated in Fig. 14(c), resulting in a lower *Recall*. On highway scenes, our method gained low *Precision* for the dotted lines of 91.36% and 89.14%, respectively. This is because some solid lines in Fig. 14(d) were partitioned into small sections due to occlusion and grouped into dotted lines for similar shape and spatial arrangement. Overall, the proposed method has been achieved very promising performance for road marking classification from MLS point clouds, despite a few road markings being misclassified.

4.4. Comparative study

Because of unavailable open datasets and execution codes for road marking classification contests, the comparison between our GAT_SCNet

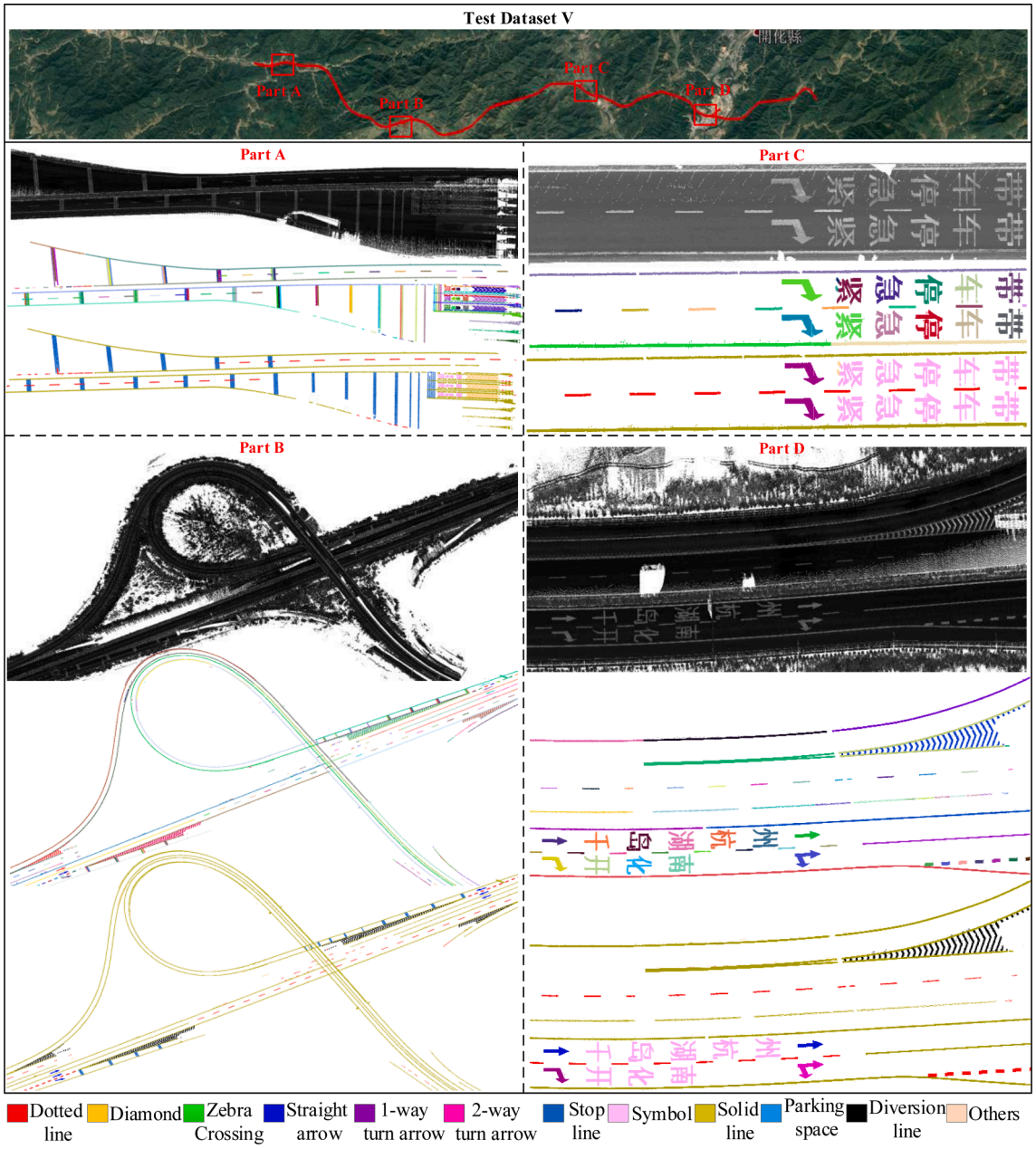


Fig. 13. Road marking classification results of test dataset V denoting highway. The top figure is the location of test dataset IV in Google Earth. Part A-C is selected high scenes. For each part, we presented ground points displayed by intensity, the road marking segments indicated by different colors, and the corresponding classification results colored by category, respectively.

model with other state-of-the-art methods is challenging. Hence, we compared our models with three popular deep learning models on point clouds including PointNet by Qi et al. (2017a), and two graph deep learning models on point clouds: DGCNN by Wang et al. (2019b) and GAPNet by Chen et al. (2021). Additionally, we analyzed the impact of three attention mechanisms and constructed serial ablation experiments with 7 ablation models through ablating different head attention layers. According to the optimal parameters configuration, we reconstruct and re-implement these comparison and ablation models on the same training samples and test datasets. Qualitative and quantitative comparison results on five test datasets with average F_1 were listed in Table 5. Some visual compared performances on typical scenes from the test datasets were illustrated in Fig. 15.

Based on quantitative illustrated in Table 5 and Fig. 15, we can see that our method outperformed all the compared methods and gained significant results. According to our findings, incorporating local graph structures and multi-head attention layers into deep learning frameworks contributes to improved performance of road marking classification.

(1) the impact of the graph structure

We found that local graph structures embedded in a deep learning framework can significantly improve the classification accuracy of linear road marking. In comparison, 7 ablation models (Model^{1}_{noGAT} ~ $\text{Model}^{7}_{Geo+Gap}$) and our GAT_SCNet successfully categorized the majority of linear road marking using neighborhood information encoded in the subgraphs. Since PointNet, DGCNN, and GAPNet were developed for 3D shape recognition, they gained general performance on all categories and could not distinguish similar linear road marking like in Scenes A, C, E, and F in Fig. 15. Even unstacked the graph attention mechanism

Table 4
Accuracy of road marking classification results.

Dataset	Type Accuracy	Dotted line	Diamond	Zebra crossing	Straight arrow	1-way turn arrow	2-way turn arrow	Stop line	Symbol	Solid line (m)	Parking space	Diversion line
Test	*TPs*	390	28	394	5	12	2	7	–	2,967.17	–	–
Dataset I	*FPs*	3	0	5	0	0	0	0	–	103.58	–	–
	FNs	10	0	9	0	0	0	3	–	103.36	–	–
	P(%)	99.24	**100.00**	98.75	**100.00**	**100.00**	**100.00**	**100.00**	–	96.63	–	–
	R(%)	97.50	**100.00**	97.77	**100.00**	**100.00**	**100.00**	70.00	–	96.63	–	–
	F_1(%)	98.36	**100.00**	98.25	**100.00**	**100.00**	**100.00**	82.35	–	96.63	–	–
Test	*TPs*	24	6	138	4	–	3	11	21	1,936.61	–	–
Dataset	*FPs*	6	0	13	0	–	0	1	0	0	–	–
II	*FNs*	2	0	6	0	–	0	0	0	31.09	–	–
	P(%)	80.00	**100.00**	91.39	**100.00**	–	**100.00**	91.67	**100.00**	**100.00**	–	–
	R(%)	92.31	**100.00**	95.83	**100.00**	–	**100.00**	100.00	**100.00**	98.42	–	–
	F_1(%)	85.71	**100.00**	93.56	**100.00**	–	**100.00**	95.65	**100.00**	96.77	–	–
Test	*TPs*	1296	28	828	39	29	88	145	174	16449.91	72	4
Dataset	*FPs*	5	2	3	2	0	0	18	10	162.5	8	0
III	*FNs*	90	0	3	0	0	0	0	0	35.16	0	0
	P(%)	99.62	93.33	99.64	95.12	**100.00**	**100.00**	88.98	94.57	99.02	90.00	**100.00**
	R(%)	93.51	**100.00**	99.64	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	99.19	**100.00**	**100.00**
	F_1(%)	96.55	96.46	99.64	97.50	**100.00**	**100.00**	94.16	97.21	99.40	94.74	**100.00**
Test	*TPs*	3,184	–	41	81	–	5	–	1	96,859.41	–	30
Dataset	*FPs*	301	–	2	0	–	0	–	0	848.26	–	0
IV	*FNs*	193	–	6	0	–	0	–	0	3,713.67	–	0
	P(%)	91.36	–	95.35	**100.00**	–	**100.00**	–	**100.00**	99.13	–	**100.00**
	R(%)	94.28	–	87.23	**100.00**	–	**100.00**	–	**100.00**	96.31	–	**100.00**
	F_1(%)	92.80	–	91.11	**100.00**	–	**100.00**	–	**100.00**	97.70	–	**100.00**
Test	*TPs*	5,359	–	–	67	46	9	171	277	201,729.46	–	62
Dataset	*FPs*	653	–	–	0	0	0	0	2	4,321.71	–	0
V	*FNs*	58	–	–	0	0	0	29	13	9,327.46	–	0
	P(%)	89.14	–	–	**100.00**	**100.00**	**100.00**	**100.00**	99.28	97.90	–	**100.00**
	R(%)	98.93	–	–	**100.00**	**100.00**	**100.00**	85.50	95.52	95.58	–	**100.00**
	F_1(%)	93.78	–	–	**100.00**	**100.00**	**100.00**	92.18	97.36	96.73	–	**100.00**
All	*Ave_P*(%)	91.87	97.78	96.28	99.02	**100.00**	**100.00**	95.16	98.46	98.65	90.00	**100.00**
Datasets	*Ave_R*(%)	95.31	**100.00**	95.12	**100.00**	**100.00**	**100.00**	88.88	98.88	97.52	**100.00**	**100.00**
	Ave_F$_1$(%)	93.42	98.85	95.64	99.50	**100.00**	**100.00**	91.09	98.64	98.23	94.74	**100.00**

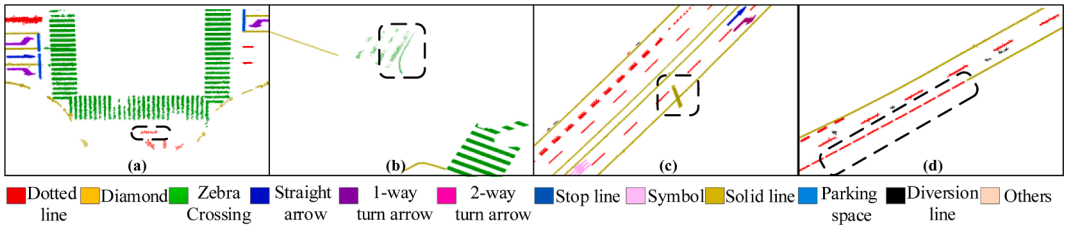


Fig. 14. Some classical error classification. (a) misclassified stop line, (b) misclassified solid line, (c) misclassified stop line, (d) misclassified solid lines.

Table 5
Quantitative comparisons (in %) between our method and comparison models on five test datasets.

Category	F_1									
Method	Dotted line	Diamond	Zebra crossing	Straight arrow	1-way turn arrow	2-way turn arrow	Stop line	Symbol	Solid line	Parking space	Diversion line
PointNet	60.61	86.79	66.81	51.28	62.80	84.75	37.44	79.42	92.55	49.62	9.09
DGCNN	56.20	86.06	53.20	60.20	78.10	92.16	50.61	85.97	93.69	60.00	25.00
GAPNet	70.03	86.41	58.75	42.24	72.46	90.13	61.29	85.51	92.03	51.79	33.33
$Model^1_{noGAT}$	85.12	67.54	75.02	88.51	84.75	74.34	88.24	83.22	92.15	70.59	36.36
$Model^2_{Nod}$	88.74	73.97	92.17	82.08	86.37	83.77	80.78	88.08	93.83	73.85	60.05
$Model^3_{Gap}$	89.46	75.65	93.96	84.19	86.68	85.68	82.83	88.81	93.94	75.39	100.00
$Model^4_{Geo}$	90.91	91.99	94.63	96.13	98.36	96.00	89.50	98.74	97.01	99.31	100.00
$Model^5_{Nod+Gap}$	88.08	73.64	92.78	85.12	87.18	79.21	82.02	86.62	95.78	70.94	95.43
$Model^6_{Nod+Geo}$	92.33	92.07	95.82	97.11	99.61	96.16	88.01	99.07	95.06	99.31	94.88
$Model^7_{Gap+Geo}$	93.06	96.04	95.72	98.43	100.00	98.98	89.42	88.97	97.07	99.31	100.00
GAT_SCNet (Ours)	**93.42**	**98.85**	**95.64**	**99.50**	**100.00**	**100.00**	**91.09**	**98.64**	**98.23**	**94.74**	**100.00**

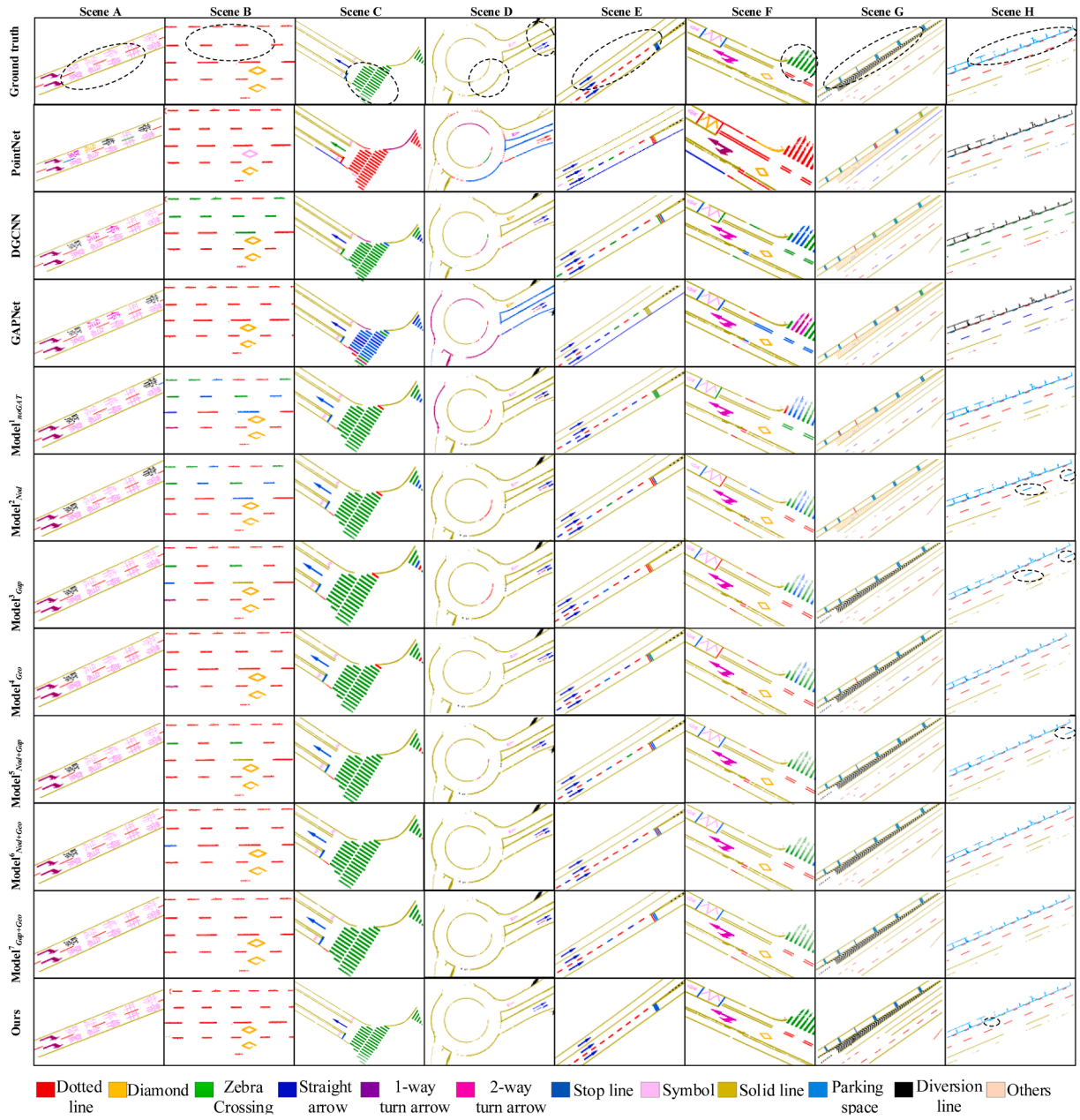


Fig. 15. Visual compared performances on some typical scenes from the test datasets.

modules and relying on the sole node embedding features, Model$^1_{noGAT}$ outperformed the above three compared methods and significantly raised the average F_1 of dotted line, zebra crossing, straight-arrow, stop line and parking space from 70%, 67%, 61%, 62%, 61% to 85.12%, 75.02%, 88.51%, 88,24%, 70.59% respectively. Compared with PointNet, DGCNN had better results on road marking in most categories, GAPNet tends to classify similar shape road marking into the same category such as dotted lines seen in Scene B. These studies also confirm the reliability of integrating EdgeConv layers and GAP layers into our GAT_SCNet model.

(2) the impact of the multi-head attention mechanisms

Compared with Model$^1_{noGAT}$ without attention mechanism modules, other ablation models stacking attention mechanism layers gained more promising road marking classification. In contrast, Model$^4_{Gap}$ with geometry-ware attention layers outperformed Model$^2_{Nod}$ and Model$^3_{Geo}$ stacking node-ware and graph-ware attention layer, respectively. The reason may be that the geometry-ware attention module integrates more prominent information including global features and direction features of nodes and pays more attention to the spatial arrangement of road marking in subgraphs, which is more beneficial for improving the road marking recognition results. In addition, we observed that Model$^4_{Gap}$ were more conducive to recognizing solid lines especially curve lines (shown in Fig. 15(D)). Generally, more attention heads, higher road marking classification accuracy. Compared with only stacking graph-ware attention module in Model$^3_{Gap}$, Model$^7_{Gap+Geo}$ integrating geometry-ware attention mechanism improved F_1 over 21% for diamond, 13% for three arrows. Combining three head attention mechanisms, our methods gained the best competitive performance on road marking classification, especially on linear road marking like dotted line, straight arrow, zebra crossing, stop line, and solid line. For some road scenes (see Scenes A, D, and E in Fig. 15), our method outperformed all compared methods and correctly group the incompleteness arrows, stop lines, and

texts into the corresponding categories. Contrary to expectations, our methods and $Model^{5}_{Nod+Gap}$ gained lower F_1 of parking spaces than $Model^{4}_{Geo}$. The main reasons is some dotted lines were misclassified as parking spaces for the season of similar shape and spatial distribution.

5. Conclusion

In this paper, we propose a novel graph deep learning named GAT_SCNet for road marking classification on MLS point clouds, which aims to category large-size and small-size road marking in a unique framework. Unlike previous studies, we construct a computable subgraph structure for each road marking and address the road marking classification as node classification. To solve the problem of learning graphs with different shapes, sizes, and node relationships, a multi-head attention mechanism has been explored in the GAT_SCNet model, which dramatically aggregates geometric features represented by node embedding and spatial distribution represented by graph features into distinguishable features of road marking. To assess the effectiveness and generalization of the GAT_SCNet model, we chose four training datasets and five test datasets from different MLS systems and cities. Quantitative and visual results show that our method has achieved very encouraging performance on fine road marking classification, with the average *Precision*, *Recall*, and F_1 of 11 categories exceeding 91%. Especially, the GAT_SCNet model properly identifies two types of arrows and diversion lines with F_1 of 100%. For linear road marking, the GAT_SCNet model also gains very high F_1 score, 93.42% for dotted line, 95.64% for zebra crossing, 91.09% for a stop line, and 98.23 % for solid line in the test datasets. Comparative studies indicated that the proposed GAT_SCNet model outperforms the state-of-the-art methods by utilizing the graph structure and multi-head attention mechanism. Results of comparative experiments also illustrate that the distinctiveness of the subgraph suffers from the quantity and quality of neighboring road marking. Therefore, in future research, we will focus on a "point-to-point" deep learning model for road markings segmentation to overcome the influence of the uneven intensity data.

CRediT authorship contribution statement

Lina Fang: Validation, Formal analysis, Investigation, Resources, Writing – original draft, Visualization. **Tongtong Sun:** Conceptualization, Software, Methodology, Writing – original draft, Validation. **Shuang Wang:** Conceptualization, Software, Methodology, Writing – original draft, Writing – review & editing, Funding acquisition. **Hongchao Fan:** Conceptualization, Methodology, Supervision. **Jonathan Li:** Validation, Resources, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Our study is jointly supported by the National Natural Science Foundation of China (NSFC) project (No. 42071446), and the Fujian Foreign Cooperation Project Foundation (No. 2020I0007). Thanks to Optech, StreetMapper, and Trimble Inc. for providing the datasets.

References

Chen, C., Fragonara, L.Z., Tsourdos, A., 2021. GAPointNet: Graph attention based point neural network for exploiting local feature of point cloud. Neurocomputing. 438, 122–132. https://doi.org/10.1016/j.neucom.20 21.01.095.

Cheng, M., Zhang, H., Wang, C., Li, J., 2017. Extraction and Classification of Road Markings Using Mobile Laser Scanning Point Clouds. IEEE J. Selected Topics Applied Earth Observations Remote Sensing. 10 (3), 1182–1196.

Chang, A.X., Funkhouser, T., Guibas, L., Hanrahan, P., Huang, Q., Li, Z., Savarese, S., Savva, M., Song, S., Su, H.C., 2015. ShapeNet: An Information-Rich 3D Model Repository. Science arXiv preprint arXiv:1512.03012.

Eigen, D., Fergus, R., 2015. Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture. Proceedings of the IEEE international conference on computer vision 2650–2658.

Fang, L., Shen, G., Luo, H., Chen, C., and Zhao, Z., 2021. Automatic Extraction of Roadside Traffic Facilities From Mobile Laser Scanning Point Clouds Based on Deep Belief Network, IEEE Transactions on Intelligent Transportation Systems. 22, 1964-1980. https://doi.org/09/TITS.2020.3017629.

Guan, H., Li, J., Yu, Y., Wang, C., Chapman, M., Yang, B., 2014. Using mobile laser scanning data for automated extraction of road markings. ISPRS J. Photogrammetry Remote Sensing. 87, 93–107. https://doi.org/10.1016/j.isprsjprs.2013.11.005.

Hamilton, W. L., Ying, R., and Leskovec, J., 2017. Inductive Representation Learning On Large Graphs, Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, California, USA, 1025–1035. https://dl.acm.org/doi/abs/10.5555/32944771.3294869.

Kipf, T., Welling, M., 2016. Semi-Supervised Classification with Graph Convolutional Networks. Paper presented at the ICLP 2016: International Conference on Learning Representations. arXiv preprint arXiv:1609.02907.

Kumar, P., McElhinney, C.P., Lewis, P., McCarthy, T., 2014. Automated road markings extraction from mobile laser scanning data. International Journal of Applied Earth Observation and Geoinformation. 32, 125–137. https://doi.org/10.1016/j.jag.2014.03.023.

Lan, S., Yu, R., Yu, G., and Davis, L. S., 2019. Modeling local geometric structure of 3D point clouds using Geo-CNN, In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 998-1008. https://doi.org/10.1109/CVPR.2019.00109.

Simonovsky, M., and Komodakis N., 2017. Dynamic Edge-Conditioned Filters in Convolutional Neural Networks on Graphs, In: IEEE Conference on Computer Vision and Pattern Recognition. 29-38.https://doi.or g/10.1109/cvpr.2017.11.

Soilán, M., Riveiro, B., Martínez-Sánchez, J., Arias, P., 2017. Segmentation and classification of road markings using MLS data. ISPRS J. Photogrammetry Remote Sensing 123, 94–103.

Ma, L., Li, Y., Li, J., Yu, Y., Junior, J.M., Goncalves, W.N., Chapman, M.A., 2021. Capsule-Based Networks for Road Marking Extraction and Classification From Mobile LiDAR Point Clouds. IEEE Trans. Intelligent Transportation Systems. 22 (4), 1981–1995. https://doi.org/10.1109/TITS.2020.2990120.

Monti, F., Boscaini, D., Masci, J., Rodola, E., Svoboda, J., Bronstein, M.M., 2017. Geometric deep learning on graphs and manifolds using mixture model cnns. Proc. IEEE Conference on Computer Vision and Pattern Recognition 5115–5124. https://doi.org/10.1109/CVPR.2017.576.

Mi, X., Yang, B., Dong, Z., Liu, C., Zong, Z., Yuan, Z., 2021. A two-stage approach for road marking extraction and modeling using MLS point clouds. ISPRS J. Photogrammetry Remote Sensing. 180, 255–268. https://doi.org/10.1016/j.isprsjprs.2021.07.012.

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 77–85. https://doi.org/10.1109/CVPR.2017.16.

Qi, C. R., Yi, L., Su, H., and Guibas, L. J., 2017. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. Paper presented at the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA. arXiv preprint arXiv:1706.02413, 2017.

Revilloud, M., Gruyer, D., and Pollard, E., 2013. An improved approach for robust road marking detection and tracking applied to multi-lane estimation, IEEE Intelligent Vehicles Symposium (IV). 783-790.https://dio.org /10.1109/IVS.2013.6629562.

Shi, J., Malik, J., 2000. Normalized cuts and image segmentation. IEEE Trans Pattern Anal Mach Intell 2000, 22(8), 888-905. https://doi.org/10.1109/34.868688.

Thomas, H., Qi, C.R., Deschaud, J., Marcotegui, B., Goulette, F., Guibas, L., 2019. KPConv: Flexible and Deformable Convolution for Point Clouds, Computer Vision and Pattern Recognition. Paper presented at the IEEE/CVF International Conference on Computer Vision (ICCV) 2019, 6411–6420.

Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., and Bengio, Y., 2018. Graph attention networks. International Conference on Learning Representations (ICLR) 2018. arXiv preprint arXiv:1710.10903.

Vishwanath, K. V., Gupta, D., Vahdat, A., and Yocum, K., 2009. ModelNet: Towards a datacenter emulation environment, IEEE Ninth International Conference on Peer-to-Peer Computing. 81-82. https://doi.org/10.11 /09/p2p.2009.5284497.

Wanda, P., Jie, H. J., and Mining., 2021. DeepFriend: finding abnormal nodes in online social networks using dynamic deep learning, Social Network Analysis. 11, 1-12. https://doi.org/10.1007/s13278-021-00742-2.

Wang, L., Huang, Y., Hou, Y., Zhang, S., and Shan, J., 2019a. Graph Attention Convolution for Point Cloud Semantic Segmentation, In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 10288-10297. https://doi.org/10.1109/CVPR.2019.01054.

Wen, C., Sun, X., Li, J., Wang, C., Guo, Y., Habib, A., 2019. A deep learning framework for road marking extraction, classification and completion from mobile laser scanning point clouds. ISPRS J. Photogrammetry Remote Sensing. 147, 178–192. https://doi.org/10.1016/j.isprsjprs.2018.10.007.

Wang, Y., Sun, Y., Liu, Z., Sarma, S.E., Bronstein, M.M., Solomon, J.M., 2019b. Dynamic Graph CNN for Learning on Point Clouds. ACM Trans. Graphics (TOG). 38 (5), 1–12. https://doi.org/10.1145/3326362.

Wen, C., Li, X., Yao, X., Peng, L., Chi, T., 2021. Airborne LiDAR point cloud classification with global-local graph attention convolution neural network. ISPRS J. Photogrammetry Remote Sensing. 173, 181–194. https://doi.org/10.1016/j.isprsjprs.2021.01.007.

Yang, B., Fang, L., Li, Q., and Li, J. , 2012. Automated Extraction of Road Markings from Mobile Lidar Point Clouds, Photogrammetry Engineering And Remote Sensing. 78, 331-338. https://doi.org/10.14358/PER S.7 8.4.331.

Yang, B., Fang, L., 2014. Automated Extraction of 3-D Railway Tracks from Mobile Laser Scanning Point Clouds. IEEE J. Selected Topics Applied Earth Observations Remote Sensing 7 (12), 4750–4761. https://doi.org/10.1109/JSTARS.2014.2312378.

Yang, B., Fang, L., Li, J., 2013. Semi-automated extraction and delineation of 3D roads of street scene from mobile laser scanning point clouds. ISPRS J. Photogrammetry Remote Sensing. 79, 80–93. https://doi.org/10.1016/j.isprsjprs.2013.01.016.

Yu, Y., Li, J., Guan, H., Jia, F., Wang, C., 2015. Learning Hierarchical Features for Automated Extraction of Road Markings From 3-D Mobile LiDAR Point Clouds. IEEE J. Selected Topics Applied Earth Observations Remote Sensing. 8 (2), 709–726. https://doi.org/10.1109/JSTARS.2014.2347276.

Yang, Z., Yang, D., Dyer, C., He, X., Hovy, E., 2016. Hierarchical Attention Networks for Document Classification. Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies.

Yang, M., Wan, Y., Liu, X., Xu, J., Wei, Z., Chen, M., Sheng, P., 2018. Laser data based automatic recognition and maintenance of road markings from MLS system. Optics & Laser Technology 107, 192–203.

Yang, B., Fang, L., Li, Q., Li, J., 2012b. Automated Extraction of Road Markings from Mobile Lidar Point Clouds. Automated Extraction of Road Markings from Mobile Lidar Point Clouds. 78 (4), 331–338.

Zhang, W., Qi, J., Wan, P., Wang, H., Xie, D., Wang, X., Yan, G., 2016. An Easy-to-Use Airborne LiDAR Data Filtering Method Based on Cloth Simulation. Remote Sensing. 8 (6), 501. https://doi.org/10.3390/rs8060501.

Zhou, Y., Huang, R., Jiang, T., Dong, Z., Yang, B., 2021. Highway alignments extraction and 3D modeling from airborne laser scanning point clouds. International J. Applied Earth Observation Geoinformation. 102, 102429. https://doi.org/10.1016/j.jag.2021.102429.