GCN-Based Pavement Crack Detection Using Mobile LiDAR Point Clouds

Huifang Feng, *Student Member, IEEE*, Wen Li, *Student Member, IEEE*, Zhipeng Luo, *Student Member, IEEE*, Yiping Chen, *Senior Member, IEEE*, Sarah Narges Fatholahi, Ming Cheng, *Member, IEEE*, Cheng Wang, *Senior Member, IEEE*, José Marcato Junior, *Member, IEEE*, and Jonathan Li, *Senior Member, IEEE*

***Abstract*—Mobile Laser Scanning (MLS) system can provide high-density and accurate 3D point clouds that enable rapid pavement crack detection for road maintenance tasks. Supervised learning-based algorithms have been proved pretty effective for handling such a large amount of inhomogeneous and unstructured point clouds. However, these algorithms often rely on a lot of annotated data, which is labor-intensive and time-consuming. This paper presents a semi-supervised point-level approach to overcome this challenge. We propose a graph-widen module to construct a reasonable graph structure for point clouds, increasing the detection performance of graph convolutional networks (GCN). The constructed graph characterizes the local features from a small amount of annotated data, avoiding information loss and dramatically reduces the dependence on annotated data. The MLS point clouds acquired by a commercial RIEGL VMX-450 system are used in this study. The experimental results demonstrate that our method outperforms the state-of-the-art point-level methods in terms of recall, F1 score, and efficiency while achieving comparable accuracy.**

***Index Terms*—Pavement crack detection, MLS point clouds, semi-supervised, deep learning, GCN.**

I. Introduction

As THE most common damage for pavements, the crack may arise due to natural and human factors such as climate changes and increased traffic, getting worse depending on its severity. Road distress adversely affects the roadway's regular use and results in traffic accidents and substantial financial losses. At the same time, pavement crack is the essential and standard indicator for evaluating road distress. By quantifying the location, size, shape and depth of pavement cracks, the severity of road distress can be quickly and effectively evaluated, and corresponding maintenance strategies can be formulated in time. Therefore, proper road maintenance and ongoing visual inspections of pavement infrastructures have been the keys to ensure road safety. Fig. 1 shows several examples of pavement cracks.

Fig. 1. Examples of pavement cracks.

Traditional manual inspection methods of pavement cracks based on in-situ measurements are generally time-consuming, labor-intensive, hazardous, and unsteadiness since the measurement results rely on the inspector's subjective experience [1]. Current methods for pavement inspection are often based on image processing and computer vision techniques [2]. Image capturing is safer and more effective than manual methods. However, the accuracy depends on image quality which can be affected by light, shadow, and stains, making crack detection challenging [3]. Compared with images, LiDAR data provides more reliable 3D information independent of adverse conditions suffered by images. LiDAR data provides a more reasonable description of the object's features, helping to achieve more accurate detection results. Moreover, LiDAR data can provide position information, enable point-based methods to locate the detected cracks accurately. Nowadays, supervised learning-based methods have achieved high-accuracy detection results on 3D point clouds [4]–[6]. However, these methods have two significant shortcomings:

(1) Most learning-based methods adopt dimensionality reduction strategies, converting 3D point clouds into 2D images to reduce the processing difficulty, leading to information loss and false detection. Results derived from these methods are biased and not conducive to accurate road evaluation and refined management. These dimension reduction-based methods have two problems: discard the elevation information of point clouds and reduce the resolution of point clouds due to projecting multiple points into a pixel. These methods regard point clouds as a discrete and unrelated point set

Manuscript received October 24, 2020; revised April 28, 2021 and June 22, 2021; accepted July 15, 2021. This work was supported by the National Natural Science Foundation of China under Grant 41871380. The Associate Editor for this article was Z. Duric. *(Corresponding authors: Yiping Chen; Jonathan Li.)*

Huifang Feng, Wen Li, Zhipeng Luo, Yiping Chen, Ming Cheng, and Cheng Wang are with Fujian Key Laboratory of Sensing and Computing for Smart Cities, School of Informatics, Xiamen University, Xiamen 361005, China (e-mail: fenghuifang@stu.xmu.edu.cn; liwen777@stu.xmu.edu.cn; zpluo@stu.xmu.edu.cn; chenyiping@xmu.edu.cn; chm99@xmu.edu.cn; cwang@xmu.edu.cn).

Sarah Narges Fatholahi and Jonathan Li are with the Department of Geography and Environmental Management, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: nfatholahi@uwaterloo.cap; junli@uwaterloo.ca).

José Marcato Junior is with the Faculty of Engineering, Architecture and Urbanism and Geography, Federal University of Mato Grosso do Sul, Campo Grande 79070-900, Brazil (e-mail: jose.marcato@ufms.br).

Digital Object Identifier 10.1109/TITS.2021.3099023

and ignore the elevation difference between crack and non-crack points. Cracks are generally distributed continuously on the road, existing continuous elevation difference between adjacent crack points and surrounding pavement points. Our method uses graph representation to construct point clouds into a graph, utilizing these continuous elevation differences and the adjacency relationship to characterize the local feature of different points.

(2) These data-driven learning-based methods need to train a significant number of parameters, and the computational intensity increases as the network depth grow. Besides, their performance heavily relies on the annotated training data, and their models only work in limited scenarios. To deal with large-scale and more complex point clouds, these methods put forward higher requirements on the quantity, quality, and diversity of annotated training data. However, annotated data acquisition is time-consuming, high-cost, and labor-intensive. These highly data-dependent methods are insufficient to handle the increasing data processing demands in large-scale and complex point cloud scenarios. Semi-supervised learning is a promising approach to alleviate data pressure. However, the existing semi-supervised detection methods mainly focus on 2D image tasks rather than 3D point cloud processing.

To handle the above problems, we propose a semi-supervised learning-based method, leverages GCN and proposes a graph-widen module. The main contributions of our work are as follows:

(1) To the best of our knowledge, this is the first work that combines GCN and point-level pavement crack detection. We represent point clouds with the graph structure, avoiding information loss and characterizing the local features of points through their neighborhood.

(2) The proposed semi-supervised method can learn features from a small amount of annotated data and a large amount of unannotated data, which can significantly reduce the annotated data dependency of our method.

(3) We propose a graph-widen module to expand the graph structure by constructing a reasonable adjacency relationship and new features for point clouds, which enlarges the receptive field of GCN and improves its representation capability to boost the accuracy and efficiency of crack detection.

(4) The experimental results show that our method can achieve satisfactory and even better performance in terms of recall, F1-score, and efficiency than supervised-based methods.

II. Related Work

A. Image-Based Methods

Various traditional image-based approaches, including threshold detection [7]–[9], edge detection [10]–[12], and region growing [13], [14], have been proposed to overcome the limitations of pavement crack detection. In recent years, image-based learning approaches have greatly improved the accuracy of pavement inspection. The models based on convolutional neural networks (CNN) have been widely used to recognize various cracks. Cha *et al.* propose a vision-based method using a deep CNN to detect concrete cracks [15]. The fully convolutional network (FCN) is utilized for the pavement crack detection in [16]. An autonomous surveying scheme [17] is used to collect, analyze and map the image-based distress data in real time. Riid *et al.* automatically detect pavement defects based on a deep CNN model [18]. A feature pyramid and hierarchical boosting network for pavement crack detection is proposed in [19]. A nondestructive testing technique using ground-penetrating radar and network in networks is adopted in [20] to detect pavement distress. Although these image-based learning methods have achieved impressive results in crack detection, the precision is still limited. Their performances heavily dependent on external light conditions, and their high sensitivity to shadow, illumination noise, and oil stains on pavement can cause adverse effects on detection.

B. Traditional Point Cloud-Based Methods

With the development of 3D sensors, many crack detection studies have been conducted using 3D point cloud data [21]–[25]. A planar triangulation modeling method is used in [21] to construct a triangular irregular network dataset and extract cracks based on the inverse distance weighting rasterization method from wall points. The inverse distance weighting interpolation algorithm, the maximum entropy threshold algorithm, and tensor voting are used in [22], [23] to distinguish the crack curves. An Otsu threshold-based method [24] extracts intensity information from MLS point clouds to identify crack skeletons. The maximum gradient of signal-noise ratio distribution for Gaussian filtering is adopted in [25] to identify cracks from the point clouds measurement and optimize the accuracy of the crack analysis efficiently. However, these methods are difficult to extract fine and low-connectivity cracks and are less automated.

C. Data-Driven Point Cloud-Based Methods

With the remarkable achievement of deep learning on image processing, deep learning has been introduced into 3D data processing. CrackNet [4] is a deep learning-based method that detects cracks based on 3D asphalt surface data, which has no pooling layers that downsize the original data and outputs the predicted class scores for all individual pixels and achieves a high-level of pixel-wise accuracy. However, the learning capability is limited by its fixed and non-learnable feature generator. And the processing speed of CrackNet is slow due to the huge number of parameters and large data depth at hidden layers [5]. Several works have been developed to improve CrackNet for enhanced learning capability and faster performance. CrackNet II [6] abandons the feature generator and constructs a deeper architecture. CrackNet-V [5] uses small filter to reduce number of parameters and achieves efficient feature, and proposes a new activation unit for shallow cracks. Both CrackNet II and CrackNet-V achieves better performance in efficiency and accuracy than CrackNet. However, their improvements have limitations, and they heavily dependent on annotated data. Since these CrackNet-like methods are pixel-level methods, they still cannot take full advantage of 3D information and accurately locate the position of detected cracks.

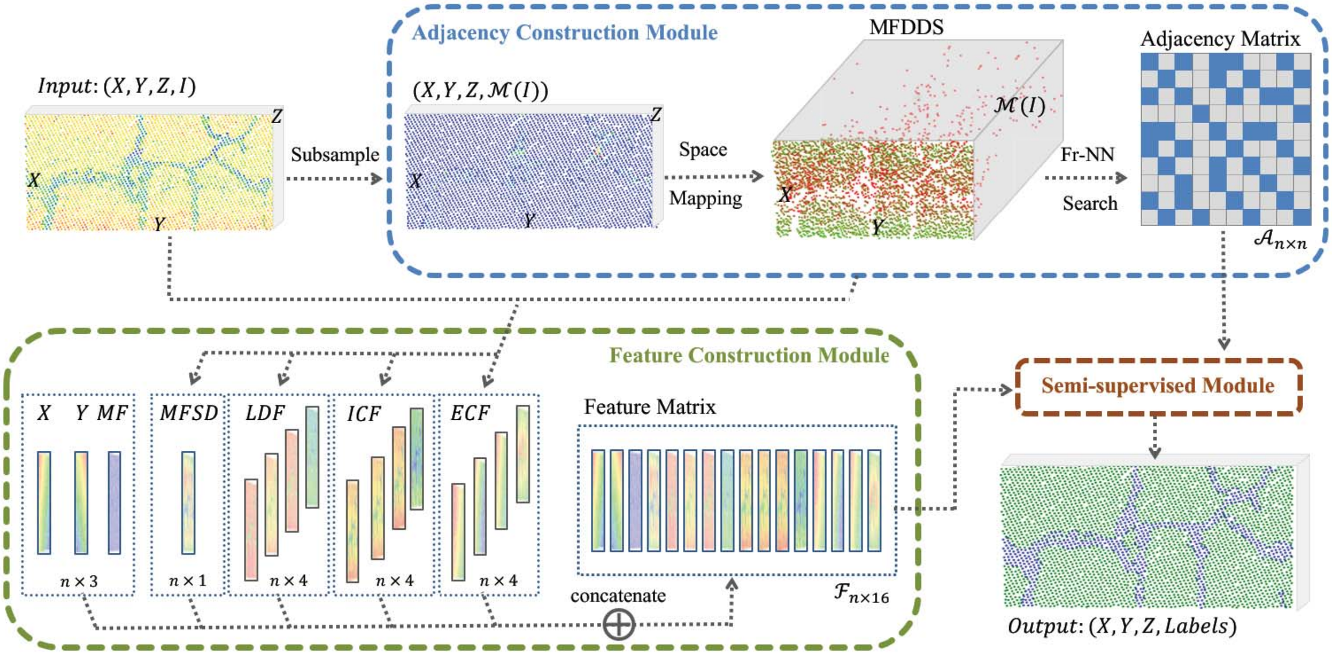


Fig. 2. Workflow of our method. The input point clouds are first subsampled. And then the subsampled point clouds are processed through the graph-widen module, which comprises the adjacency construction module and feature construction module, separated by blue and green dashed boxes. Finally, the constructed adjacency matrix and feature matrix are fed into GCN in the semi-supervised module to train the network and output each point's label.

III. Method

This paper presents a novel semi-supervised approach based on GCN for rapid point-level pavement crack detection. Fig. 2 illustrates the framework of our method. Each point of the input point clouds comprises four channels, i.e., the three-dimensional coordinates (X, Y, Z) and the reflection intensity I. The input point clouds are preprocessed by subsampling and then processed by three modules: adjacency construction, feature construction, and semi-supervised network. The first two form the graph-widen module.

Preprocessing: We subsample the input point clouds to reduce the computational complexity of subsequent processing and change the point arrangement structure, making the structure more general while maintaining the original shape characteristics. A threshold distance sd is used for subsampling, ensuring the distance between any two points in the subsampled point clouds greater than sd. The subsampled point clouds belong to a subset of the original point clouds.

A. Adjacency Construction

1) Space Mapping: The crack points have lower reflection intensity and generally distribute in irregular crossed strip-like shapes. In comparison, the non-crack points have higher reflection intensity (pavement points and road marking points) and scattered distribution (pavement texture points). However, road surfaces are usually located in a planar. The elevation differences among different points are tiny. The intensity differences are disturbed by the pavement texture. As a result, these differences being too weak to serve as classification criteria. Therefore, we propose a space mapping strategy to amplify the reflection intensity distinction and distribution structure difference between crack and non-crack points.

For the subsampled point clouds $P = \{p_1, p_2, \dots, p_n | p_i = (x_i, y_i, z_i, I_i)\}$, where n is the number of points. For $p_i \in P$, $C_i = (x_i, y_i, z_i)$ denotes the three-dimensional coordinates and I_i the reflection intensity. We define the main feature (MF) $\mathcal{M}(I_i)$ of p_i, as follow:

$$\mathcal{M}(I_i) = \csc\left(\frac{1}{1+e^{-\alpha\times I_i}}\right), \tag{1}$$

where $\csc(\cdot)$ indicates the cosecant function, α is a changeable coefficient.

We design MF to make the distances of inter-class points increase much faster than that of intra-class points. We construct a new space, named main feature distance distribution space (MFDDS), based on MF. Its coordinate system is defined by (X, Y, MF). For $p_i \in P$, we perform Eq. (1) to transform p_i into $p_i' = (x_i, y_i, z_i, \mathcal{M}(I_i))$. Then, we switch the positions of z_i and $\mathcal{M}(I_i)$, obtaining the mapping point $q_i = (x_i, y_i, \mathcal{M}(I_i), z_i)$. For $\forall p_i \in P$, we perform the above operation to map P into MFDDS to get $Q = \{q_1, q_2, \dots, q_n | q_i = (x_i, y_i, \mathcal{M}(I_i), z_i)\}$. For $q_i \in Q$, ${C_i}^m = (x_i, y_i, \mathcal{M}(I_i))$ denotes the spatial coordinates, z_i is the elevation feature.

2) Adjacency Matrix: For $\forall q_i, q_j \in Q$, we define the main feature distance (MFD) $\mathcal{F}_{MFD}(q_i, q_j)$ as follow:

$$\mathcal{F}_{MFD}(q_i, q_j) = \|{C_i}^m - {C_j}^m\|_2, \tag{2}$$

where $\|\cdot\|_2$ denotes the L_2 norm.

We utilize the nearest neighbors search (NNS) method to identify the neighborhood for each point of Q. For the reason that the k nearest neighbors (k-NN) algorithm always selects a fixed number of points around the center point as its neighbors, ignoring the difference in the distribution structure amplified for different points in the MFDDS construction process. The fr-NN algorithm is used in our method since it treats all the

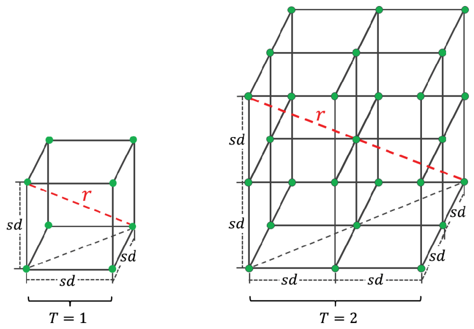


Fig. 3. Illustration of the relationship of the search radius r of fr-NN, radius scale T and the threshold distance sd of subsampling, $r = T \times sd \times \sqrt{3}$.

points in a self-centered sphere with a fixed radius as neighbors of the center point, characterizes the different distribution structures for different points in MFDDS. We construct the adjacency matrix $\mathcal{A}_{n\times n} = (\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n)$ of Q based on MFD. The adjacency vector of q_i is defined as $\mathcal{A}_i = (a_{i1}, a_{i2}, \ldots, a_{in})'$, $a_{ij}\,(j \in [1, n])$ is calculated as follow:

$$a_{ij} = \begin{cases} 1, & \mathcal{F}_{MFD}(q_i, q_j) \leq r \\ 0, & else, \end{cases} \tag{3}$$

where $r = sd \times T \times \sqrt{3}$ is the search radius of fr-NN, and T denotes the radius scale. T also indicates the maximum hop from the center to the points on the sphere's surface. The relationship of r, T and sd is illustrated in Fig. 3. If $a_{ij} = 1$, then q_j is a neighbor of q_i. The neighborhood of q_i is referred to as $\mathcal{N}_{q_i}$. We perform the operation on every point in Q to obtain the adjacency matrix $\mathcal{A}_{n\times n}$.

B. Feature Construction

In this section, we design four new features for the reorganized point clouds in MFDDS. These new features help construct a more compact feature space, increasing the GCN's receptive field. The well-designed features including main feature spatial density (MFSD), elevation context feature (ECF), intensity context feature (ICF), and local distribution feature (LDF). Their details are described below.

1) Main Feature Spatial Density: In road point clouds, the number of crack points is far less than that of non-crack points, and the non-crack points distribute densely while most of the crack points arrange in irregular strip-like shapes. Therefore, in MFDDS, the density of crack points would be much smaller than that of pavement points while bigger than textured pavement points. We design MFSD to describe this character of the reconstructed point clouds. The value of MFSD can be obtained by calculating the number of neighbors of each point. The MFSD of q_i is referred to as $\mathcal{F}_{MFSD}(q_i)$, can be calculated as follow:

$$\mathcal{F}_{MFSD}(q_i) = norm(\sum_{j=1}^{n} a_{ij}, range(0, 1)), \tag{4}$$

where $norm(\cdot, range(0, 1))$ normalizes the input value to a range between 0 and 1.

2) Elevation Context Feature: The elevation differences between crack and non-crack points in the road are very small, easily interfered with by pavement texture points. Therefore, we design ECF to obtain contextual information of the elevation in the neighborhood of each point. Taking advantage of the amplified local distribution structure in MFDDS to strengthen the local feature of points near the crack edge. For $q_i \in Q$, the local elevation set of q_i is called $\mathcal{N}^Z_{q_i}$, which consists of the elevation information of all points in $\mathcal{N}_{q_i}$. The ECF of q_i is referred to as $Z_i{}^{loc}$, describing the local elevation distribution in the neighborhood of q_i, calculated as follow:

$$Z_i{}^{loc} = (max(\mathcal{N}^Z_{q_i}), min(\mathcal{N}^Z_{q_i}), mean(\mathcal{N}^Z_{q_i}), var(\mathcal{N}^Z_{q_i})), \tag{5}$$

where $max(\cdot)$, $min(\cdot)$, $mean(\cdot)$ and $var(\cdot)$ calculate the maximum, minimum, mean and variance of the input, respectively.

For point q_i, the categories of its neighbors can be inferred from the intra-relationship of its ECF. For example, if $max(\mathcal{N}^Z_{q_i}) \approx mean(\mathcal{N}^Z_{q_i}) \gg min(\mathcal{N}^Z_{q_i})$, then it can be inferred that most neighbors of q_i are non-crack points. If $max(\mathcal{N}^Z_{q_i}) \approx mean(\mathcal{N}^Z_{q_i}) \approx min(\mathcal{N}^Z_{q_i})$, then most neighbors of q_i belong to a same class.

3) Intensity Context Feature: In section III-A1, the reflection intensity is transformed into MF to construct MFDDS. Although the transformation amplifies the reflection intensity distinction between crack and non-crack points, information loss is inevitable. To compensate the information loss, we design ICF to make further use of the reflection intensity. The local reflection intensity set of q_i is referred to as $\mathcal{N}^I_{q_i}$, consists of the intensity of all the points in $\mathcal{N}_{q_i}$. We refer to the ICF of q_i as $I_i{}^{loc}$, which can be calculated as follow:

$$I_i{}^{loc} = (max(\mathcal{N}^I_{q_i}), min(\mathcal{N}^I_{q_i}), mean(\mathcal{N}^I_{q_i}), var(\mathcal{N}^I_{q_i})). \tag{6}$$

4) Local Distribution Feature: The MFD set of q_i is referred to as $\mathcal{N}^D_{q_i} = \mathcal{F}_{MFD}(q_i, \mathcal{N}_{q_i})$, consists of the MFD between q_i and all the points in its neighborhood $\mathcal{N}_{q_i}$. We design LDF to utilize the MFD and the connection between each point and its constructed local feature in MFDDS. LDF of q_i is referred to as $D_i{}^{loc}$, which can be calculated as follow:

$$D_i{}^{loc} = (max(\mathcal{N}^D_{q_i}), min(\mathcal{N}^D_{q_i}), mean(\mathcal{N}^D_{q_i}), var(\mathcal{N}^D_{q_i})). \tag{7}$$

Finally, all the features are concatenated in parallel to construct the final feature of q_i, which is referred to as $\mathcal{F}_{q_i} = (x_i, y_i, \mathcal{M}(I_i), \mathcal{F}_{MFSD}(q_i), Z_i{}^{loc}, I_i{}^{loc}, D_i{}^{loc})$. And the feature matrix of Q can be obtained as $\mathcal{F}_{n\times 16} = (\mathcal{F}(q_1), \mathcal{F}(q_2), \ldots, \mathcal{F}(q_n))'$.

C. Semi-Supervised Network

ChebyNet [26] uses a Chebyshev polynomial to approximate kernels in graph signal processing and the evaluation complexity of the kernel is linear. ChebyNet avoids the Fourier basis and reduces the convolutional kernel size and the computational complexity. The proposed spectral kernels are strictly localized in a ball of K hops from the central vertex. And ChebyNet adopts a multilevel clustering algorithm that produces coarser graphs corresponding to the data domain seen at different resolution. Through the expression of the convolution kernel and the adjacent relationship of the graphs, ChebyNet strictly represents the localize feature of the convolution kernel. The vertices are rearranged as a binary tree structure and turned to be a 1D signals utilizing the coarser graphs. Then the proposed efficient pooling strategy is operated at the 1D signals.

To efficiently achieve point-level pavement crack detection, we utilize a 2-layer ChebyNet to serve as the classifier, which takes the adjacency matrix $\mathcal{A}_{n\times n}$, the feature matrix $\mathcal{F}_{n\times 16}$, and a small number of labels as input. With randomly initialized weights of ChebyNet, we train it in a semi-supervised manner. Our method only inputs part of the labels into ChebyNet and predicts the labels of the remaining nodes. However, the adjacency relations and features of the points with and without a label are known during the training. For $q_a, q_b, q_c \in Q$, q_a and q_b belong to the neighborhood of q_c. In the adjacency construction module, q_a and q_b take part in the adjacency construction of q_c. In the feature construction module, the features of q_a and q_b contribute to constructing the features of q_c. Although the labels of q_a and q_b are unknown, they can still play roles during ChebyNet training through their relationship with q_c.

IV. Experiment and Discussion

In this section, we first introduce the benchmark dataset, implementation details, and parameters setting. Then, we analyze the crack detection results and compare them with state-of-the-art methods. Finally, a series of ablation studies are explained.

A. Data Description

As far as we know, there is no public benchmark or dataset designed for point-level pavement crack detection on point clouds. Therefore, we build our dataset with the road point clouds recorded by the RIEGL VMX-450 scanning system in Qinghai-Tibet Highway in September 2015. The highway starts from Xining, Qinghai Province, China (longitude 36.6171 N, latitude 101.7782 E), ends at Lhasa, Tibet, China (longitude 29.6525 N, latitude 91.1721 E). It is the highest and longest asphalt road in the world. The average elevation is 4,000 meters, the length is up to 1937 km, the mean width of the roadbed is 10 meters, and the slope is less than 7%. Fig. 4 shows the information of the Qinghai-Tibet Highway in Google Earth, and the red line represents the main road.

The RIEGL VMX-450 scanning system integrates two RIEGL VQ-450 laser scanners, four digital cameras, and inertial navigation devices. The scanning system has a high measurement speed of up to 1,100,000 points per second, a high scanning speed of 400 lines per second, and an effective range of up to 800 meters. The maximum scanning precision of the mobile laser scanning system can reach 5mm. Moreover, the average driving speed during travel on Qinghai-Tibet Highway is about 80km/h. These scanning parameters and settings guarantee the millimeter level resolution of the obtained point cloud data.

We segmented 115 pieces of road point clouds from the Qinghai-Tibet Highway point cloud data to build our dataset, of which size is 809 MB in total. Fig. 5 shows the details of road point clouds in our dataset. In the acquired road point clouds, points are densely arranged in scan lines, while almost no points arrange between them. Therefore, it is necessary to conduct subsampling on the original road point clouds to obtain evenly arranged point clouds, facilitating the following operations and reducing the computational complexity.

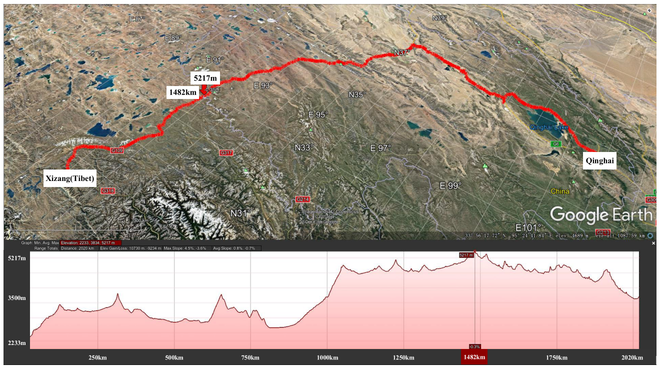
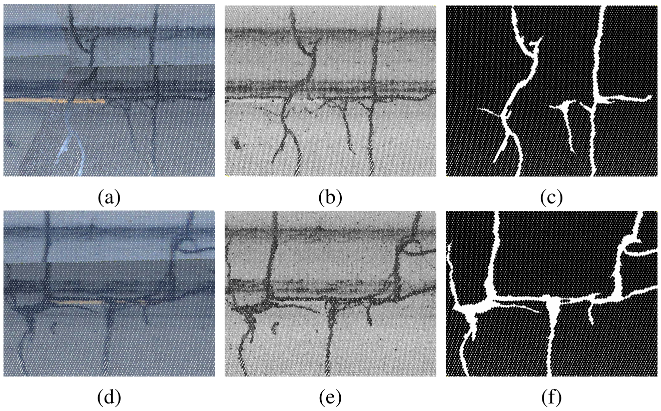


Fig. 4. The detail information of the Qinghai-Tibet Highway in Google Earth. The red line in the map represents the main road of the Qinghai-Tibet Highway, and the line chart below the map illustrates the elevation variance.


(a) (b) (c)

(d) (e) (f)

Fig. 5. Road point clouds. (a) (d): samples with RGB information. (b) (e): samples with reflection intensity, and points with lower intensity are rendered in darker color. (c) (f): the corresponding manually annotated ground truth, where black/white points denote non-crack/crack areas.

B. Ground Truth and Samples

We train and test our method with our dataset to evaluate the performance. To obtain the ground truth, we manually selected the crack points in each road point clouds and annotated them as "crack", the rest were annotated as "non-crack". Based on which we calculate the evaluation metrics. Figs. 5c and 5f show the examples of the manual labelled ground truth of the road point clouds.

For each piece of point clouds in our dataset, the width and length are approximate 10 meters and contain an average of 200,000 points. Since public urban roads are regularly maintained, there are usually few cracks on these roads. It is difficult to ensure regular and comprehensive maintenance due to the long-distance, diverse geographic, and changeable climate environment of the Qinghai-Tibet Highway. Thus, the cracks will inevitably be much more than ordinary roads. We subjectively selected road segments with more cracks since they are urgent to be maintained. We calculated the crack to

TABLE I
PARAMETER SETTINGS OF GCN

Parameters	value
learning rate	0.01
units in layer 1	2
dropout rate	0.5
weight for $L2$ loss	5×10^{-4}
chebyshev polynomial degree	3

the non-crack ratio for each road segment and select segments with a ratio of roughly 2:8 to construct our dataset.

The dataset is divided into two parts: dataset-1 consists of 80 pieces of road point clouds and dataset-2 consists of 35 pieces. Each piece of dataset-2 is sorted in increasing order of the X and Y coordinates, then the sorted points are divided into three parts according to the index order. Specifically, for each point clouds in dataset-2, the part 1 consists of the first 10% of the whole points, the part 2 consist of the last 60% of that, and the part 3 consists of the remained points.

C. Implementation Details

The proposed method is implemented in python 3.7, tensorflow 1.15.2, Ubuntu 16.04, Intel(R) Xeon(R) CPU E5-2698 v4 @ 2.20GHz, Nvidia Tesla V100 and 64GB memory. The following parameters are involved in the proposed method: α, sd, T and r. Below, we describe how to set these parameters.

As described in Section III-A1, α is a coefficient that decides the transformation degree from reflection intensity to main feature, which is set as $\alpha = 1$. The distance between adjacent parallel scanning lines on the road is around 0.1 m, calculated according to the scanning speed and driving speed. Therefore, the threshold distance is set as $sd = 0.05\ m$, ensuring the distance between any two points after subsampling is greater than 0.05 m. Unless otherwise specified, the radius scale is set as $T = 3$. The search radius of fr-NN can be calculated as $r = 0.05m \times 3 \times \sqrt{3}$. The parameters setting of the 2-layer ChebyNet are shown in Table I.

For each point clouds in dataset-2, we build a graph structure based on all points and construct the adjacency matrix and feature matrix. ChebyNet is trained with the adjacency matrix, feature matrix, and the labels of points in the corresponding part 1(10%) and predicts points in part 2(60%).

D. Quantitative Assessment Measures

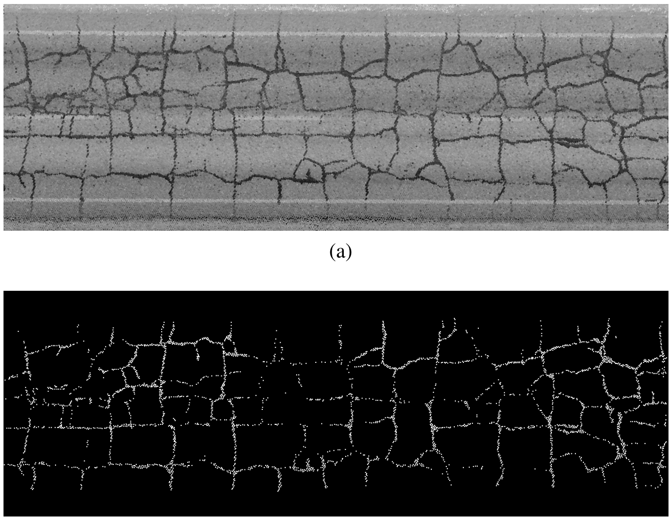
We calculate the evaluation metrics based on each piece of road point clouds, where the precision (PRE), recall (REC), F_1-score (F1), and accuracy (ACC) are defined as follows:

$$PRE = \frac{N_{TP}}{N_{TP} + N_{FP}}, \tag{8}$$

$$REC = \frac{N_{TP}}{N_{TP} + N_{FN}}, \tag{9}$$

$$F_1\text{-}score = 2 \times \frac{PRE \times\ REC}{PRE + REC}, \tag{10}$$

$$ACC = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{TN} + N_{FP} + N_{FN}}, \tag{11}$$


(a)

(b)

Fig. 6. Crack detection results of a 30-meter point cloud scene. (a) Reflection intensity of the road point clouds. (b) Detection result of our method, black points represent the non-crack points and white for the crack points.

where N_{TP} is the number of true positive points; N_{TN} is the number of true negative points; N_{FP} is the number of false positive points; N_{FN} is the number of false negative points.

E. Comparative Experiments

1) Crack Detection Result: As shown in Fig. 6, the experimental results demonstrate the effectiveness of our algorithm. The detection results of a 30-meter road segment with complex cracks are shown in Fig. 6b. The continuous structure of the complex cracks in different types, shapes, widths, and lengths is correctly detected through our method. Which takes advantage of the powerful local representation capability of graph representation and performs satisfactory point-level crack detection result in large road scene.

2) Comparative Analysis: To the best of our knowledge, our method is the first work that focuses on point-level pavement crack detection from point clouds in a semi-supervised manner. To evaluate the feasibility and performance, we compare our method with pixel-level methods [27], [28] and point-level methods [24], [29], [30].

a) Pixel-level methods: We project dataset-1 and dataset-2 into intensity images, and train U-Net and AU-Net with the corresponding intensity images of dataset-1. Test them on the corresponding intensity images of part 2 of dataset-2, and calculate the average pixel-wise assessment measures. During training, all the hyperparameters are set as default.

b) Point-level methods: 3DSkeleton [24] is a traditional point-level method, comprises the Otsu thresholding algorithm, spatial density filter, Euclidean distance clustering and L_1 medial skeleton detection method. We apply the Otsu thresholding algorithm and spatial density filter to extract crack points. The spatial density filter with a local radius $rd = 0.2m$ and a density threshold $ds = 1.2$ are used to remove outliers. PointNet [29] and DGCNN [30] are point-level supervised detection methods. We train them with dataset-1. During training, all the hyperparameters are set as default. For our method, we train it with the part 1 of

TABLE II
COMPARISON RESULTS(%) OF DIFFERENT METHODS

	Method	train:test	training time	PRE	REC	F1	ACC
pixel-level	U-Net [27]	7:3	>6h	**80.9**	73.3	**76.6**	**95.3**
	AU-Net [28]	7:3	>6h	76.5	**77.8**	76.4	95.0
point-level	3DSkeleton [24]	-	-	33.6	71.7	46.7	91.0
	PointNet [29]	7:3	>2h	70.4	63.1	66.1	94.3
	DGCNN [30]	7:3	>4h	**73.0**	67.0	70.1	**94.9**
	Ours	1:6	**<40m**	70.0	**73.9**	**71.9**	93.1

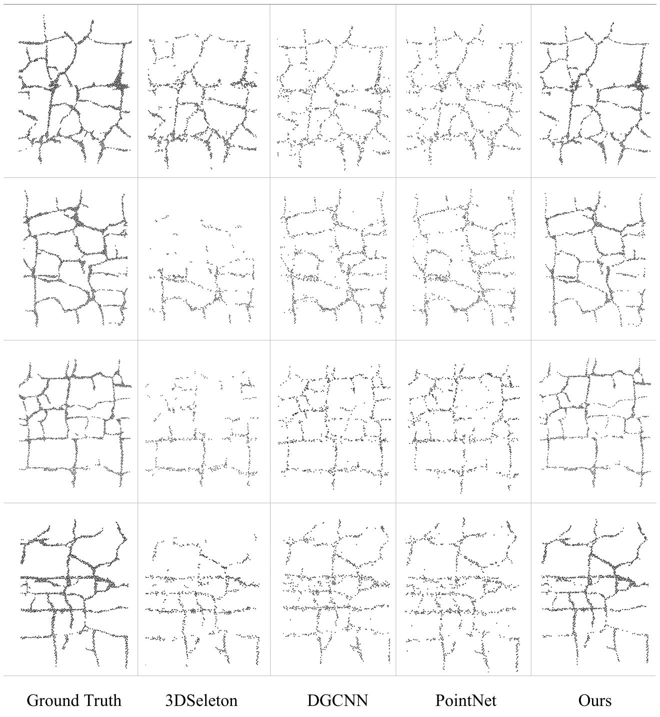
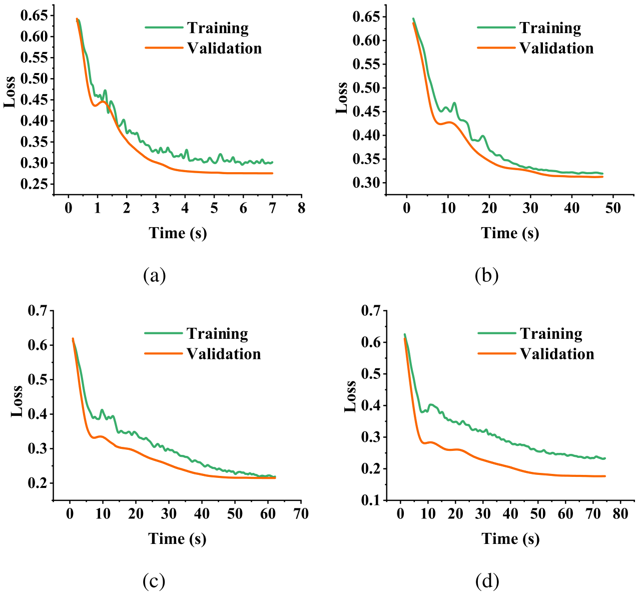


Fig. 7. The point-level crack detection results of comparison methods and ground truth. The first column is the ground truth of the different road point cloud scenes. The last four columns are the detection results of 3DSkeleton [24], DGCNN [30], PointNet [29] and our method, respectively.

dataset-2 from piece to piece. Specifically, for a piece of point clouds D belong to dataset-2, the input adjacency matrix and feature matrix are constructed based on D. The input label belongs to points in part 1 of D. We test all the point-level methods on part 2 of dataset-2 and calculate the average point-wise assessment.

The performance of our method is analyzed according to the assessment measures type. The quantitative results are shown in Table II. Pixel-wise measures could obtain a higher assessment value than point-wise due to the information loss caused by dimensionality reduction. Considering that our approach costs much less training data and the training time than U-Net and AU-Net, our approach has achieved satisfactory results than these pixel-level supervised learning-based methods. Our method achieves a recall of 73.9%, compared with 71.7% of 3DSkeleton, 63.1% of PointNet and 67.0% of DGCNN. As for F1, our method achieves a score of 71.9%, which is 25.2% higher than 3DSkeleton, 5.8% than PointNet and 1.8% than DGCNN. The results show that our method achieves significant performance improvement from the traditional and supervised learning-based point-level approaches in recall and F1, mainly due to the powerful local feature representation capability generated by the constructed adjacency relationship and strictly designed features. And the visualization results in Fig. 7 show that our method has great advantages when the cracks have obviously local structure and can detect more complete and continuous cracks. However, our method causes a small number of failure cases when dealing with line-shaped cracks with narrow widths. And when the proportion of crack points are too small in the point clouds, due to the data imbalance, the network can easily converge to a state that all points are regarded as non-crack points while still achieve a good loss.



(a) (b) (c) (d)

Fig. 8. Training/validation loss on four different scenes.

TABLE III
RESULTS(%) OF DIFFERENT NNS METHODS AND TRAIN/TEST SPLIT

Features	NNS	train:test	PRE	REC	F1
X, Y, Z	-	-	-	-	-
X, Y, Z, I	-	-	-	-	-
X, Y, MF, MFSD	k-NN	-	-	-	-
	fr-NN	1:7	83.6	53.2	65.1
		1:6	83.0	63.4	71.9
		1:4	**84.0**	50.9	63.4
X, Y, MF, MFSD, ICF	k-NN	-	-	0	0
	fr-NN	1:6	80.9	72.6	76.6
		1:5	79.3	67.8	73.1
		1:4	82.5	67.6	74.3
X, Y, MF, MFSD, ICF, LDF, ECF	k-NN	-	-	-	-
	fr-NN	1:6	78.9	**75.5**	**77.2**

3) Efficience Analysis: When achieving the above-mentioned performance on the same test data, the training data and training time consumed by our method is much less than PointNet and DGCNN. Specifically, our method only consumes 30% of the time consumed by

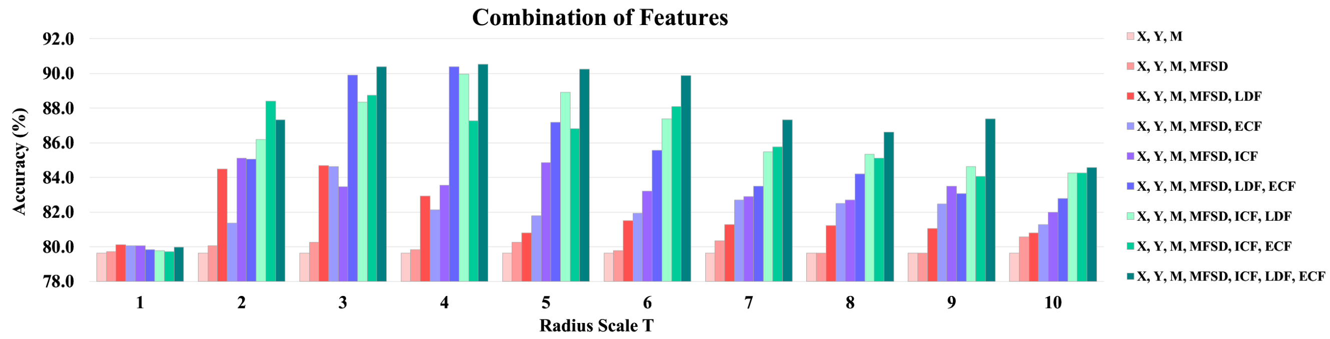


Fig. 9. The comparative experimental results of different combinations of features. y-axis: accuracy; x-axis: the value of radius scale T. The crack detection accuracy generally increases with the addition of features.

TABLE IV
ACCURACY(%) OF DIFFERENT RADIUS SCALE VS. FEATURE COMBINATION

Feature Combination	Coordinates	NNS	T=1	T=2	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
X, Y, Z	(X,Y,Z)	fr-NN	79.6	79.6	79.6	79.6	79.6	79.6	79.6	79.6	79.6	79.6
X, Y, Z, I	(X,Y,Z)	fr-NN	79.6	79.6	79.6	79.6	79.6	79.6	79.6	79.6	79.6	79.6
X, Y, MF	(X,Y,MF)	fr-NN	79.6	79.6	79.6	79.6	79.6	79.6	79.6	79.6	79.6	79.6
X, Y, MF, Z	(X,Y,MF)	fr-NN	79.6	79.6	79.6	79.6	79.6	79.6	79.6	79.6	79.6	79.6
X, Y, MF, MFSD	(X,Y,MF)	fr-NN	79.7	80.1	80.2	79.8	80.2	79.8	80.3	79.6	79.6	80.6
X, Y, MF, MFSD, ECF	(X,Y,MF)	fr-NN	**80.1**	81.4	84.6	82.1	81.8	81.9	82.7	82.6	82.5	81.3
X, Y, MF, MFSD, ICF	(X,Y,MF)	fr-NN	**80.1**	85.1	83.5	83.5	84.9	83.2	82.9	82.7	83.5	82.0
X, Y, MF, MFSD, LDF, ECF	(X,Y,MF)	fr-NN	79.8	85.1	89.9	90.4	87.2	85.6	83.5	84.2	83.1	82.8
X, Y, MF, MFSD, LDF, ICF	(X,Y,MF)	fr-NN	79.8	86.2	88.4	90.0	88.9	87.4	85.5	85.3	84.6	84.3
X, Y, MF, MFSD, ECF, ICF	(X,Y,MF)	fr-NN	79.7	**88.4**	88.7	87.2	86.8	88.1	85.8	85.1	84.1	84.3
X, Y, MF, MFSD, LDF, ECF, ICF	(X,Y,MF)	fr-NN	80.0	87.3	**90.4**	**90.5**	**90.2**	**89.9**	**87.3**	**86.6**	**87.4**	**84.6**

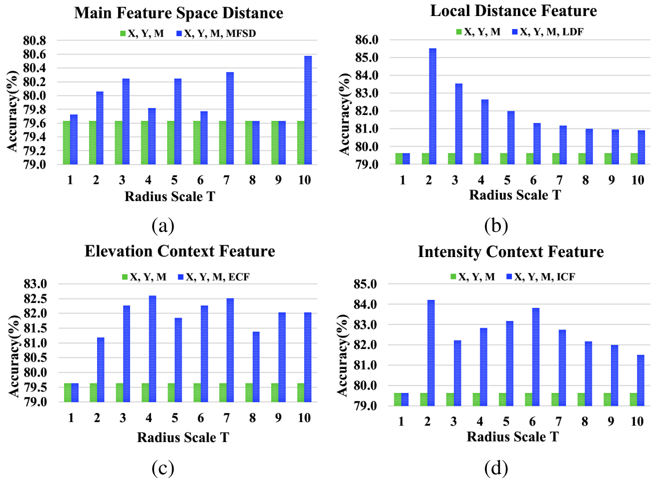
PointNet and 17% of the time consumed by DGCNN. As a semi-supervised method, our method can be trained with much less data while achieving satisfactory or even better performance than supervised methods. The demands of training data of PointNet and DGCNN is more than 2.3 times of test data, while the training data of our semi-supervised method is less than 0.17 times the same test data. These results validate the high efficiency and low data dependency of our method. Moreover, as shown in Fig. 8, in the early stage of training, the training loss dropped rapidly and then stabilized near 0.3 and 0.2, indicating that the model tends to converge. Furthermore, the validation loss has the same trend as training loss. These results show that the proposed method can converge and has excellent performance on the validation set.

F. Ablation Experiments

We conducted a series of ablation studies to investigate the importance of different components of our method. We perform some experiments to compare the effectiveness of different NNS methods and the influence of different training and test data split. The experimental results are shown in Table III. The first two rows show that when (X, Y, Z) serve as the coordinates to perform NNS, neither k-NN nor fr-NN can accurately segment crack points from the road point clouds. When (X, Y, MF) serve as the coordinate to calculate the neighbors for each point, the k-NN still cannot conduct any accurate prediction, while the fr-NN achieves the best recall with the train/test split equal to 1:6. The precision of detection improving with the increase in the proportion of training data, but the recall and F1-measure have dropped to a certain extent. Note that, in the other experiments, our method uniformly the train/test split to 1: 6.

We conduct a series of experiments to investigate the importance of the designed features in Section III-B. Fr-NN is used to conduct the adjacency matrix for all experiments. The experimental results are shown in Table IV. For the first two rows, (X, Y, Z) represents coordinates, and (X, Y, MF) for the others. The value of radius scale T varies from 1 to 10 to change the radius r, leading to a different neighborhood for each point and resulting in a different adjacency matrix.

As shown in the first row of Table IV, the accuracy remains unchanged. Since all points are predicted as non-crack points due to the serious data imbalance of crack and non-crack points. While in other rows, accuracy gradually increased with the additional input of features. Our method achieves the best average performance with all features, which demonstrates that each designed feature has its contribution. The visualization results shown in Figs. 9 and 10 also verify the above observation. Fig. 9 shows that the accuracy increased with the addition of features. For $T \in [1, 10]$, all the best accuracy occurs with all features. Overall, our method achieves best performance when $T \in [3, 4]$ and inputting all features. This is because the neighbor points located in the range when $T \in [3, 4]$ contribute most to the local feature representation of each point. And Fig. 10 shows that each designed feature can significantly increase the detection accuracy.



(a) (b) (c) (d)

Fig. 10. The comparative experimental results that illustrate the effect and necessity of the designed features. y-axis: accuracy; x-axis: the value of radius scale T. (a), (b), (c), and (d) show the effectiveness of MFSD, LDF, ECF, and ICF, respectively.

V. Conclusion

This paper presents a novel semi-supervised 3D pavement crack detection algorithm, increasing the detection reliability and efficiency. Our method is successfully applied to the 3D MLS dataset. Evaluations demonstrate that our method achieves a satisfactory performance and illustrate our method's high efficiency and low data dependency. Our method achieves 73.9% in recall and 71.9% in F1 in the Qinghai-Tibet Highway dataset. Compared with the state-of-the-art point-level methods, our method outperforms recall, F1 score, and efficiency while achieving comparable accuracy, which benefits from the reasonably constructed adjacency relationship and the strictly designed features based on the cracks' characteristics. Consequently, we provide a feasible and promising semi-supervised solution, which boosts the performance of pavement cracks detection from point clouds. Future extensions of this work will be oriented to explore the internal connections and mutual influences of different features and propose a loss function with adaptive weight to relieve the sample imbalance problem of object detection in large-scale point cloud scenes.

References

[1] T. B. J. Coenen and A. Golroo, "A review on automated pavement distress detection methods," *Cogent Eng.*, vol. 4, no. 1, Jan. 2017, Art. no. 1374822.

[2] E. Salari and G. Bao, "Pavement distress detection and classification using feature mapping," in *Proc. IEEE Int. Conf. Electro/Inf. Technol.*, May 2010, pp. 1–5.

[3] X. Chen and J. Li, "A feasibility study on use of generic mobile laser scanning system for detecting asphalt pavement cracks," in *Proc. ISPRS Arch.*, vol. 41, Jul. 2016, pp. 1–5.

[4] A. Zhang *et al.*, "Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network," *J. Comput.-Aided Civil Infrastruct. Eng.*, vol. 32, no. 10, pp. 805–819, 2017b.

[5] Y. Fei *et al.*, "Pixel-level cracking detection on 3D asphalt pavement images through deep-learning- based CrackNet-V," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 1, pp. 273–284, Jan. 2020.

[6] A. Zhang *et al.*, "Deep learning–based fully automated pavement crack detection on 3D asphalt surfaces with an improved CrackNet," *J. Comput. Civil Eng.*, vol. 32, no. 5, Sep. 2018, Art. no. 04018041.

[7] J. Tang and Y. Gu, "Automatic crack detection and segmentation using a hybrid algorithm for road distress analysis," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Oct. 2013, pp. 3026–3030.

[8] N. Katakam, "Pavement crack detection system through localized thresholing," Ph.D. dissertation, College Eng., Univ. Toledo, Toledo, OH, USA, 2009.

[9] H. Oliveira and P. L. Correia, "Automatic road crack segmentation using entropy and image dynamic thresholding," in *Proc. EUSIPCO*, 2009, pp. 622–626.

[10] B. Tremblais and B. Augereau, "A fast multi-scale edge detection algorithm," *Pattern Recognit. Lett.*, vol. 25, no. 6, pp. 603–618, 2004.

[11] R. Salim Lim, H. M. La, Z. Shan, and W. Sheng, "Developing a crack inspection robot for bridge maintenance," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2011, pp. 6288–6293.

[12] R. S. Lim, H. M. La, and W. Sheng, "A robotic crack inspection and mapping system for bridge deck maintenance," *IEEE Trans. Autom. Sci. Eng.*, vol. 11, no. 2, pp. 367–378, Feb. 2014.

[13] M. Gavilán *et al.*, "Adaptive road crack detection system by pavement classification," *Sensors*, vol. 11, no. 10, pp. 9628–9657, 2011.

[14] Q. Zou, Y. Cao, Q. Li, Q. Mao, and S. Wang, "CrackTree: Automatic crack detection from pavement images," *Pattern Recognit. Lett.*, vol. 33, no. 3, pp. 227–238, 2012.

[15] Y.-J. Cha, W. Choi, and O. Büyüköztürk, "Deep learning-based crack damage detection using convolutional neural networks," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 32, no. 5, pp. 361–378, May 2017.

[16] W. M. Jung, F. Naveed, B. Hu, J. Wang, and N. Li, "Exploitation of deep learning in the automatic detection of cracks on paved roads," *Geomatica*, vol. 73, no. 2, pp. 29–44, Jun. 2019.

[17] M.-M. Naddaf-Sh, S. Hosseini, J. Zhang, N. A. Brake, and H. Zargarzadeh, "Real-time road crack mapping using an optimized convolutional neural network," *Complexity*, vol. 2019, pp. 1–17, Sep. 2019.

[18] A. Riid, R. Lõuk, R. Pihlak, A. Tepljakov, and K. Vassiljeva, "Pavement distress detection with deep learning using the orthoframes acquired by a mobile mapping system," *Appl. Sci.*, vol. 9, no. 22, p. 4829, Nov. 2019.

[19] F. Yang, L. Zhang, S. Yu, D. V. Prokhorov, X. Mei, and H. Ling, "Feature pyramid and hierarchical boosting network for pavement crack detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 4, pp. 1525–1535, Apr. 2020.

[20] Z. Tong, D. Yuan, J. Gao, Y. Wei, and H. Dou, "Pavement-distress detection using ground-penetrating radar and network in networks," *Construct. Building Mater.*, vol. 233, Feb. 2020, Art. no. 117352.

[21] H. Jiang, Q. Li, Q. Jiao, X. Wang, and L. Wu, "Extraction of wall cracks on earthquake-damaged buildings based on TLS point clouds," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 9, pp. 3088–3096, Sep. 2018.

[22] H. Guan *et al.*, "Iterative tensor voting for pavement crack extraction using mobile laser scanning data," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1527–1537, Mar. 2015.

[23] H. Guan, J. Li, Y. Yu, M. Chapman, and C. Wang, "Automated road information extraction from mobile laser scanning data," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 1, pp. 194–205, Feb. 2015.

[24] Y. Yu, J. Li, H. Guan, and C. Wang, "3D crack skeleton extraction from mobile LiDAR point clouds," in *Proc. IEEE Geosci. Remote Sens. Symp.*, Jul. 2014, pp. 914–917.

[25] X. Xu and H. Yang, "Intelligent crack extraction and analysis for tunnel structures with terrestrial laser scanning measurement," *Adv. Mech. Eng.*, vol. 11, no. 9, Sep. 2019, Art. no. 168781401987265.

[26] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," in *Proc. NeuralIPS*, 2016, pp. 3844–3852.

[27] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. MICCAI*, 2015, pp. 234–241.

[28] O. Oktay *et al.*, "Attention U-Net: Learning where to look for the pancreas," 2018, *arXiv:1804.03999*. [Online]. Available: http://arxiv.org/abs/1804.03999

[29] R. Q. Charles, H. Su, M. Kaichun, and L. J. Guibas, "PointNet: Deep learning on point sets for 3D classification and segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 652–660.

[30] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon, "Dynamic graph CNN for learning on point clouds," *ACM Trans. Graph.*, vol. 38, no. 5, pp. 1–12, 2019.

Huifang Feng (Student Member, IEEE) received the M.Sc. degree in mathematics from Jimei University, China, in 2017. She is currently pursuing the Ph.D. degree in information and communications engineering with Fujian Key Laboratory of Sensing and Computing for Smart Cities, School of Informatics, Xiamen University, China. Her research interests include environment perception for point cloud processing, scene understanding, machine, and deep learning.

Wen Li (Student Member, IEEE) received the B.Eng. degree in communications engineering from Shandong Agricultural University, Tai'an, China, in 2018. He is currently pursuing the M.Eng. degree in electronic and communications engineering with Fujian Key Laboratory Sensing and Computing for Smart Cities, School of Informatics, Xiamen University, China. His research interests include point cloud processing, object detection, 3D vision, machine, and deep learning.

Zhipeng Luo (Student Member, IEEE) received the Ph.D. degree in information and communications engineering from Xiamen University, China, in 2020. His current research interests include environment perception for autonomous driving, mobile laser scanning for road inventory, point cloud analytics, 3D vision, machine, and deep learning. He has published several research papers in referred journals and conferences, including *ISPRS Journal of Photogrammetry and Remote Sensing*, IEEE Transactions on Intelligent Transportation Systems, and IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.

Yiping Chen (Senior Member, IEEE) received the Ph.D. degree in information and communications engineering from the National University of Defense Technology, Changsha, China, in 2011. She is a Senior Engineer with Fujian Key Laboratory of Sensing and Computing for Smart Cities, School of Informatics, Xiamen University, China. From 2007 to 2011, she was an Assistant Researcher with Chinese University of Hong Kong, China. Her current research interests include image processing, mobile laser scanning data analysis, 3D point cloud computer vision, and autonomous driving. She has published more than 60 papers in referred journals, including IEEE Transactions on Intelligent Transportation Systems, IEEE Transactions on Geoscience and Remote Sensing, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, and conferences, including CVPR, IGARSS, and ISPRS. She was a receipient of the 2020 Best Reviewer of the IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.

Sarah Narges Fatholahi received the M.Sc. degree in geomatics engineering from the University of Tehran, Iran, in 2015. She is currently pursuing the Ph.D. degree in geospatial data science with the Geospatial Sensing and Data Intelligence Laboratory, Department of Geography and Environmental Management, University of Waterloo, Canada. Her research interests include urban remote sensing, 3D mapping using LiDAR, spatial machine learning, and point cloud analytics. She has published research papers in the *International Journal of Remote Sensing*, *Survey Review*, *Geomatica*, and in ISPRS Annuals.

Ming Cheng (Member, IEEE) received the Ph.D. degree in biomedical engineering from Tsinghua University, Beijing, China, in 2004. He is currently a Professor with Fujian Key Laboratory of Sensing and Computing for Smart Cities and Xiamen Key Laboratory of Geospatial Sensing and Computing, School of Information Science and Engineering, Xiamen University, Xiamen, China. His research interests include remote sensing image processing, point cloud processing, computer vision, and machine learning.

Cheng Wang (Senior Member, IEEE) received the Ph.D. degree in signal and information processing from the National University of Defense Technology, Changsha, China, in 2002. He is currently a Professor with the School of Informatics and the Executive Director with Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University, China. He has coauthored more than 150 papers in referred journals and top conferences including IEEE Transactions on Geoscience and Remote Sensing, *Pattern Recognition*, IEEE Transactions on Intelligent Transportation Systems, IEEE Conference on Computer Vision and Pattern Recognition, Association for the Advancement of Artificial Intelligence (AAAI), and *ISPRS Journal of Photogrammetry and Remote Sensing*. His current research interests include point cloud understanding, multi-sensor fusion, mobile mapping, and geospatial big data. He is a fellow of the Institution of Engineering and Technology and an Associate Editor of IEEE Geoscience and Remote Sensing Letters. He is also the Chair of the Working Group I/6 on Multi-Sensor Integration and Fusion of the ISPRS.

José Marcato Junior (Member, IEEE) received the Ph.D. degree in cartographic science from Sao Paulo State University, Brazil. He is currently a Professor with the Faculty of Engineering, Architecture and Urbanism and Geography, Federal University of Mato Grosso do Sul, Campo Grande, MS, Brazil. His current research interests include UAV photogrammetry and deep neural networks for object detection, classification, and detection. He has published more than 30 in refereed journals and over 70 in conferences, including papers published in *ISPRS Journal of Photogrammetry and Remote Sensing*, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Transactions on Intelligent Transportation Systems, and *Remote Sensing*.

Jonathan Li (Senior Member, IEEE) received the Ph.D. degree in geomatics engineering from the University of Cape Town, South Africa, in 2000. He is currently a Professor with the Department of Geography and Environmental Management and the Department of Systems Design Engineering, University of Waterloo, Canada. His main research interests include image and point cloud analytics, mobile mapping, and AI-empowered information extraction from LiDAR point clouds and earth observation images. He has coauthored more than 400 publications, including over 240 refereed journal articles . He is currently serving as an Associate Editor for IEEE Transactions on Intelligent Transportation Systems, IEEE Transactions on Geoscience and Remote Sensing, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, and *Canadian Journal of Remote Sensing*. He was a recipient of the Outstanding Achievement in Mobile Mapping Technology Award in 2019 for his pioneering contributions in developing and promoting mobile mapping technology and the ISPRS Samuel Gamble Award in 2020 for his significant contributions to the development, organization or professional activities of the photogrammetry, remote sensing, and spatial information sciences at national or international level.