Article

PAN: Improved PointNet++ for Pavement Crack Information Extraction

Jiakai Fan [1,2], Weidong Song [2], Jinhe Zhang [1,2], Shangyu Sun [1,2,3,*], Guohui Jia [4] and Guang Jin [5]

1   School of Geomatics, Liaoning Technical University, Fuxin 123000, China; kaikaixiaohuo0428@gmail.com (J.F.); 47221055@stu.lntu.edu.cn (J.Z.)
2   Collaborative Innovation Institute of Geospatial Information Service, Liaoning Technical University, Fuxin 123000, China; lntu_swd@163.com
3   State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China
4   School of Resources and Civil Engineering, Liaoning Institute of Science and Technology, Benxi 117000, China; jiaguohui@lnist.edu.cn
5   Iroadc (Liaoning) Transportation Tech. Co., Ltd., Shenyang 110000, China; jinguang@iroadc.com
*   Correspondence: shangyu_sun@126.com

Abstract: Maintenance and repair of expressways are becoming increasingly important due to the growing frequency of their use. Accurate pavement crack information extraction helps with routine maintenance and reduces the risk of traffic accidents. The traditional 2D crack image detection method has limitations and cannot effectively obtain depth information. Three-dimensional crack extraction from 3D point cloud has become a new solution that can capture pavement crack information more comprehensively and accurately. However, the existing algorithms are not effective in the feature extraction of cracks due to the different and irregular shapes and sizes of pavement cracks and interference from the external environment. To solve this, a new method for detecting pavement cracks in point clouds, namely point attention net (PAN), is herein proposed. It uses a two-branch attention fusion module to focus on space and feature information in the cloud and capture features of crack points at different scales. It also uses the Poly Loss function to solve the imbalance of foreground and background points in pavement point cloud data. Experiments on the LNTU-RDD-LiDAR dataset were carried out to verify the effectiveness of the proposed method. Compared with the traditional method and the latest point cloud segmentation technology, the performance indexes of mIoU, Acc, F1, and Rec achieved significant improvement, reaching 75.4%, 91.5%, 75.4%, and 67.1%, respectively.

Keywords: pavement crack information extraction; 3D point cloud; deep learning; pavement crack 3D point cloud dataset; attention mechanism

1. Introduction

Highway pavement crack information extraction is very important in road quality assurance and driving safety. The early-detection methods relying on manual and sensor have high cost, high detection cost, low efficiency, and strong subjectivity, which makes road crack detection complicated and difficult [1]. The modern machine vision detection method improves the detection efficiency and accuracy to a certain extent. It provides a new solution for the identification and maintenance of highway cracks, which helps to reduce cost, improve efficiency, reduce the impact of road cracks on traffic, and improve the safety and service life of highways.

Traditional methods have been extensively studied for automatic pavement crack detection. Image processing-based techniques primarily include threshold segmentation [2,3], edge detection [4,5], and region growth [6,7]. However, these approaches often rely on prior knowledge or optimal thresholds, limiting their applicability on complex and variable

urban roads, particularly in reliably detecting cracks with weak connections or uneven geometric topologies. Some studies have proposed an information extraction method for calculating crack length [8] and used a data acquisition system combining radar rangefinder and camera [9]. Image segmentation was carried out using gray adaptive threshold algorithm, crack contour was extracted by GaussLaplace algorithm [10], and video-based software system was developed for calculating bridge crack widths. However, traditional methods often have low accuracy and can only approximate crack locations in complex backgrounds or when there is minimal contrast. These methods are also easily affected by the background.

With the rise of deep learning, crack detection has ushered in a new stage of development, and the quality of crack detection has been significantly improved by learning advanced features [11]. The use of deep learning in pavement crack information extraction allows for automated feature extraction and learning without human intervention. Additionally, Transformer models based on the seq2seq architecture have achieved significant success in crack detection, offering not only feature extraction but also multimodal fusion, addressing some of the limitations of CNNs [12]. While two-dimensional detection has made some progress, these image-based methods still face challenges, as pavement images are frequently obscured by lighting, shadows, dirt, and noise, which hinders the accurate capture of terrain and texture details.

With the advancement of 3D data acquisition technology, mobile laser scanning (MLS) systems have become widely used for generating precise 3D coordinate data, enabling efficient and flexible point cloud acquisition on road surfaces [13]. Pavement point clouds exhibit characteristics that include 3D spatial coordinates and intensity information for each point, with crack points typically differing in elevation and intensity from the surrounding pavement points. These features make the pavement point cloud an effective tool for analyzing road conditions and detecting damage. The point cloud's irregularity offers rich information about the pavement surface. The normal vector, indicating the direction at each point, is used for crack segmentation. However, road wear and noise complicate distinguishing elevation and intensity changes between cracks and normal pavement, posing challenges for crack detection.

While current methods have shown promising results, point cloud-based crack-detection techniques face three significant challenges that hinder their practical application. (1) Disorder: LiDAR systems generate 3D point cloud data that are inherently unstructured and unordered, creating difficulties for deep learning algorithms. To address this, many existing methods employ dimension reduction techniques to transform 3D point clouds into 2D images, thereby simplifying the processing. However, this conversion often results in information loss [14]. Moreover, these methods often treat point clouds as discrete, uncorrelated sets, overlooking elevation differences between cracked and non-cracked points. (2) Spatial correlation: Existing point cloud-based crack-detection methods tend to ignore the spatial correlation and adjacency of points in complex pavement structures. Since the pavement is relatively flat, neglecting these correlations can lead to insufficient recognition of small crack details, resulting in incomplete detection. (3) Data dependency: Deep learning models based on point clouds typically require large amounts of manually labeled data and involve training a significant number of parameters. This not only increases the input of manpower, time, and cost but also significantly increases the computing cost with the increase of network complexity [15]. These limitations limit the applicability of current point cloud-based crack-detection methods in different scenarios.

Based on the above problems, the following improvements are proposed:

(1) This paper introduces a PAN network built on a U-Net architecture that processes unordered point sets and directly extracts features, avoiding the information loss often caused by dimensionality reduction in traditional methods. This approach effectively preserves the spatial information and geometric characteristics of point cloud data, enhancing the accuracy and efficiency of point cloud processing;

(2)   In this paper, we present a PC-Parallel module featuring a dual attention module branch parallel structure, which flexibly adapts to pavement point cloud data of varying densities and samples. This design enhances the model's ability to understand both overall and local features, especially in extracting crack information from complex structures, and addresses segmentation result incompleteness. It not only improves the model's robustness across multi-scale and multi-density data but also enhances its capability to capture crack edges and details, thereby increasing the accuracy of point cloud pavement crack segmentation;

(3)   The Poly Loss function is introduced. By adjusting the form of the loss function, the imbalance between crack points and background points can be better balanced, and the sensitivity to edge and detail information can be enhanced. The problem of boundary refinement and class imbalance in point cloud pavement crack segmentation is solved effectively, thus improving segmentation accuracy and model performance;

(4)   A set of large-scale 3D point cloud dataset of pavement cracks suitable for semantic segmentation is established.

2. Related Work

2.1. Methods Based on Traditional Image Processing

Image-based crack-detection algorithms have rapidly advanced. Some studies [2,16] introduced an algorithm for extracting crack thresholds using global or local thresholds. However, selecting suitable thresholds for most crack images is challenging, and the method is highly sensitive to illumination and noise, reducing its stability. Other research [17,18] proposed a block-based segmentation approach, but because detection occurs at the block level, it fails to accurately identify cracks. Different researchers [19,20] developed a filter-based algorithm to detect cracks based on the expected response, but it struggles with complex and varied pavement cracks, and selecting appropriate parameters is often difficult and time-consuming. Another study [6] introduced a seed-based region growth algorithm that uses multidirectional non-minimum suppression and symmetry checking for road crack detection. One study [7] proposed CrackTree, which detects cracks by selecting them in a probability graph and uses a recursive tree edge pruning method to form a minimum spanning tree. This method can detect pavement cracks effectively but does not consider the actual crack width. In the freeway environment, the existing methods often need appropriate parameter presets or prior knowledge to achieve the best performance when dealing with the complex and changeable pavement. This dependence can lead to poor model performance in the absence of human intervention. Pavement cracks often have limited connectivity or irregular geometric topologies, and correctly identifying these cracks remains a major challenge.

2.2. Deep Learning Image Processing Method

In recent years, due to its excellent feature learning ability, many researchers have explored various deep neural networks and demonstrated excellent performance in the efficiency and accuracy of road crack assessment. One study [21] for the first time tried, CNN based on LeNet-5 [22], to detect cracks in local areas of images. This method involves converting the model into a fully convolutional network to obtain image segmentation results. Methods like [23,24] use sliding windows to divide pavement images into smaller blocks, utilizing convolutional neural networks (CNNs) to predict cracks within each block, accurately locating crack positions. However, this approach is limited to block-level detection and does not achieve pixel-level precision. Studies [25,26] have demonstrated that using CNNs for pixel-level crack detection in pavement images allows for precise crack localization and improved accuracy. Despite its precision, the block-based process can be time-consuming, and small blocks often lack sufficient context, which limits real-time performance and efficiency. On the other hand, the application of fully convolutional networks (FCNs) in crack detection has delivered impressive results in both accuracy and speed [27,28]. Building on this success, [29] proposed combining a deep convolutional

neural network (DCNN) with an edge detector to further enhance detection performance. This integration allows for more efficient processing and better edge definition, improving the overall detection capability of pavement cracks. By integrating the DCNN output with the edge detector, this approach effectively reduces the noise rate in crack identification and significantly enhances recognition accuracy. Compared to other deep learning methods, this new approach excels in noise suppression and accuracy. However, the layer number of convolutional networks is deep, resulting in a large increase in model parameters, which makes the training time of the model quite time-consuming. One study [30] utilized the Faster R-CNN two-stage network for high-precision pavement crack information extraction. Compared to the one-stage network, the two-stage approach is more accurate for multi-scale and small-target issues, offering a more reliable solution for accurately locating pavement cracks. However, this increased accuracy comes at the cost of reduced efficiency, resulting in slower detection speeds. Another study [31] proposed a feature pyramid and hierarchical boost network that processes multiple scales simultaneously and retains more details by transmitting information layer by layer. This allows the network to effectively capture features at different scales, adapting better to pavement cracks of varying sizes and shapes, thereby enhancing robustness and accuracy in pavement crack information extraction. However, with the introduction of multi-scale information, it is also affected by some noise and unnecessary information, which reduces the performance. Despite the impressive results of image-based learning methods, their performance heavily depends on external conditions. Two-dimensional pavement images and videos are frequently obscured by light, shadows, stains, and rust, complicating road crack identification. Additionally, 2D image detection techniques are limited in their ability to describe terrain and crack details, reducing the effectiveness of road maintenance assessments.

2.3. Method Based on Traditional 3D Point Cloud Data Processing

Compared to traditional 2D optical images, 3D point clouds provide more precise spatial coordinates and include intensity information. Unlike 2D images, 3D point clouds are unaffected by ambient brightness, offering a more reliable basis for accurately extracting road cracks. Recently, researchers have extensively studied using point clouds for pavement crack detection. By thoroughly analyzing point cloud data, they can detect and locate pavement cracks more accurately and improve detection robustness and accuracy. Some studies [32,33] proposed a deep crack-detection method using 3D pavement data, employing a straightforward threshold strategy to represent crack depth and length. The ITVCrack algorithm proposed by [34] is an automatic crack extraction framework based on iterative tensor voting (ITV). The algorithm first uses the inverse distance weighting (IDW) algorithm to convert pavement points into geo-referenced feature images and then detects candidate cracks using an ITV-based crack extraction framework. Finally, a morphological refinement algorithm is used to distinguish the crack curves of MLS pavement. However, due to the characteristics of the ITV algorithm, ITVCrack has high computational complexity when processing large-scale data, and its generalization performance under different road and environmental conditions is also limited. In another study [35], The Otsu threshold method was used to extract intensity differences in MLS point clouds to identify crack skeletons. Noise was then removed using a spatial density filter, and the Euclidean distance clustering algorithm divided crack points into distinct crack lines. However, this process is affected by changes in point cloud density and local shape, reducing crack line division accuracy. Another study [36] applied the plane triangulation modeling method to detect crack points on a triangular irregular network dataset constructed with the IDW rasterization method. However, the planar triangulation modeling method may lack sufficient adaptability to crack shape changes and multi-scale problems, resulting in missed or false detections in some cases. One study [37] introduced a Gaussian filter to detect the signal-to-noise ratio distribution gradient, efficiently extracted pavement cracks from terrestrial laser scanning (TLS) point cloud, and effectively optimized the accuracy of pavement damage analysis. However, the anti-interference ability against factors such as

noise and occlusion needs to be further improved. In Ref. [38], a random forest classification (RFC) method was applied using LiDAR point clouds captured by drones. Multi-scale and multi-dimensional features are extracted from the intensity and height data of the point clouds and used as input for the RFC method to identify cracks. One study [39] proposed a pavement crack extraction technique by converting MLS point clouds into regular grid structures. This approach introduced a two-dimensional index based on the acquisition time or incident angle of each 3D point. Crack candidate points are detected by combining intensity and height differences. Finally, the MLS data are converted into a "grid map", demonstrating its feasibility and effectiveness for transverse, longitudinal, and oblique cracks.

Compared to traditional 2D optical images, 3D point clouds offer more precise spatial coordinates and include intensity information. Unlike 2D images, 3D point clouds are unaffected by ambient brightness, providing a more reliable basis for accurately extracting pavement cracks. In recent years, researchers have extensively explored using point clouds for pavement crack detection. By thoroughly analyzing point cloud data, researchers can detect and locate pavement cracks with greater precision, enhancing the robustness and accuracy of the detection process [32,33]. A deep crack-detection method using 3D pavement data was proposed, utilizing a straightforward threshold strategy to represent crack depth and length. However, this approach has limitations in successfully extracting complex and fine cracks. The ITVCrack algorithm [34] offers an automatic crack extraction solution based on iterative tensor voting (ITV). It first uses the inverse distance weighting (IDW) algorithm to convert pavement points into geographic reference feature images. Candidate cracks are then identified using the ITV crack extraction framework, and a morphological refinement algorithm is applied to distinguish crack curves in the point cloud pavement. However, due to the characteristics of ITV algorithm, ITVCrack has high computational complexity when processing large-scale data, and its generalization performance under different road and environmental conditions is also limited. In [35], the Otsu threshold method was employed to extract intensity differences in the point cloud for identifying the fracture skeleton. Noise was then removed using a spatial density filter, and fracture points were divided into distinct fracture lines using the Euclidean distance clustering algorithm. However, changes in point cloud density and local shape during fracture line division reduce accuracy. Researchers [36] used the planar triangulation modeling method to detect crack points in the triangular irregular network dataset constructed by the IDW rasterization method. However, the planar triangulation modeling method may not have enough adaptability to the shape change and multi-scale problems of cracks, resulting in missing or false detection in some cases. In [37], a Gaussian filter was introduced to detect the signal-to-noise ratio distribution gradient, and pavement cracks were efficiently extracted from the ground laser scanning (TLS) point cloud to effectively optimize the pavement damage analysis accuracy. However, the anti-interference ability for noise and shielding factors needs to be further improved. In [38], a random forest classification (RFC) method was implemented using LiDAR point clouds captured by drones. Multi-scale and multi-dimensional features are extracted based on the intensity and height information of the point clouds, serving as input for crack extraction via the RFC method. In [39], by converting the point cloud into a regular grid structure, a two-dimensional index was introduced to account for the acquisition time or angle of incidence of each 3D point. Then, the fracture candidate points were detected by the combination of strength difference and height difference. At last, the MLS data were converted losslessly into a "grid graph", and their feasibility and effectiveness on transverse, longitudinal, and oblique cracks were proven. Traditional 3D point cloud processing methods often rely on hand-designed features or simple statistical features, which do not perform well in capturing the details of crack points. Effectively identifying and segmenting these features is challenging, often resulting in the loss or misjudgment of crack point features. In pavement point cloud crack segmentation, capturing local features is essential for detail identification, while combining global features helps understand the overall structure. Traditional methods struggle to

process both local and global features simultaneously, thus impacting the accuracy of pavement point cloud crack segmentation.

2.4. Point Cloud Data Processing Method Based on Deep Learning

2.4.1. Classic Deep Learning Point Cloud Segmentation Algorithm

The tremendous success of deep learning in image processing has led to its expansion into 3D data processing. In point cloud data processing, classic deep learning segmentation algorithms have achieved impressive results. Their core idea is to efficiently segment unstructured and disordered 3D point cloud data using deep learning networks to accurately identify the boundaries and features of various objects or scenes.

PointNet [40], a pioneering model for point cloud segmentation, uses shared MLPs to directly process unordered point sets as input and extract features from them. Building on PointNet, PointNet++ [41] builds on this by introducing a hierarchical feature learning paradigm that recursively captures local geometric structures. By leveraging local point representation and multi-scale information, PointNet++ achieves excellent performance and serves as the foundation for modern point cloud methods [42–44]. An improved PointNet++ [45] was used to classify ejector head defect shapes in seamless steel pipe production, adding a multi-level local feature extraction structure that reduces point count, enhances information acquisition, and provides better stability than PointNet. Advances in natural language processing have led to attention-based methods that excel in exploring point relationships. Models like PCT [46] and Point Transformer [47,48] establish global context in point clouds through self-attention, effectively managing point positions and capturing spatial structures. A Transformer-based point cloud classification network (TransPCNet) [49] was developed to identify sewer defects, using a feature embedding module to map points to a high-dimensional space for feature extraction and closure, learning multi-scale features through a self-attention cascade. To enhance discrimination between similar defect classes, a weighted, smooth cross-entropy loss function was designed to prevent overfitting during classification training. Recently, MLP-type networks have achieved excellent results by simplifying the network structure and strengthening features. PointMLP [50] proposes a geometric affine module to normalize features. RepSurf [51] models the umbrella surface by fitting surface information with triangular planes to provide geometric information. PointNeXt [52] integrates training strategies and model scaling. These methods provide useful exploration and innovation in the development of point cloud processing field.

Currently, research methods are mainly focused on validating the effectiveness of these techniques on relatively more straightforward indoor tasks, while for pavement crack information extraction, there is the challenge that only a few points contain critical information. General point cloud segmentation networks often lack specialized designs for handling sparse feature point clouds, leading to computational redundancy and undue focus on irrelevant points. In extracting pavement crack information, cracks of varying scales require different levels of attention, highlighting the need for further improvement and optimization in current technology. Addressing these issues will enhance the accuracy and robustness of pavement crack information extraction, making it more applicable to real-world complex pavement scenarios.

2.4.2. Road Crack Segmentation Task Based on 3D Point Cloud

In road crack detection, deep learning networks based on point cloud data can generally be categorized into two types: those based on convolutional neural networks (CNN) and those based on graph convolution. These approaches utilize deep learning's powerful representation learning capabilities to significantly enhance the accuracy and efficiency of road crack detection.

Convolutional Neural Network: In [13], a method utilizing an adaptive wavelet neural network (WNN) was proposed for the automatic detection of concrete cracks and other damage types. However, this approach may encounter errors when identifying cracks with fine textures. In [53], CrackNet, an efficient model based on a convolutional neural network (CNN), was proposed for automatic detection of sidewalk cracks on 3D asphalt surfaces. Although it achieves high pixel accuracy, its efficiency is hindered by a static, non-learnable feature generator, which restricts its ability to learn. One study [54] proposed an enhanced architecture based on CrackNet, namely CrackNet II, which abandoned the feature generator in favor of a more complex framework, which enabled the model to detect smaller or finer features while removing more local noise. cracks and maintain faster computing speeds. One study showed [55] CrackNet-V as an efficient deep network that uses smaller filters and introduces a new shallow crack activation unit. Compared to the original CrackNet, CrackNet-V features a deeper network architecture with fewer parameters, enhancing both computational efficiency and accuracy. However, despite these improvements, its performance gains are modest and highly reliant on the data. In [56], CrackNet-R, based on recurrent neural networks, was shown to incorporate a gated recursive multi-layer perceptron (GRMLP) to iteratively update its internal buffer. The GRMLP performs multiple layers of nonlinear transformations using gated units, allowing for deeper abstraction of input and hidden states. Compared to CrackNet, CrackNet-R improves detection speed by four times and significantly enhances detection accuracy.

Graph Convolution: From [57], it can be seen that CrackGCN is an innovative semi-supervised method for extracting 3D pavement crack information. It uses a novel spatial enhancement strategy and graph-based features to identify crack points from MLS data, boosting the effectiveness of GCN. By relying on a small amount of annotated data, Crack-GCN constructs graphs to represent local features, thereby reducing data degradation and minimizing dependence on extensive annotations. RangeSeg [58] is a range-aware instance segmentation framework with a shared encoder backbone and two range-dependent decoders. The heavy decoder focuses on detecting distant and small objects by calculating their distance from the image's top area, enhancing accuracy for small target detection. Meanwhile, the light decoder processes the entire image to reduce computational costs, effectively balancing efficiency and accuracy. SD-GCN [38], a saliency-based extended GCN network, employs two saliency feature spaces and cylinder-based extended graph convolution to detect cracks in moving laser scanning (MLS) point clouds. Both CrackGCN and SD-GCN improve the geometric structure of road point clouds using spatial enhancement strategies. However, they do not account for long-range neighborhoods and multi-scale features, resulting in incomplete crack detection within complex structures. SCL-GCN [15], a hierarchical contrastive learning graph convolution network, was designed for pavement crack extraction from MLS point clouds. It features a novel dual-branch architecture that utilizes multi-scale graphs to expand the effective receptive field for remote contexts while maintaining low computational costs. A graph feature-contrastive learning module guides the dual-branch GCNs, addressing learning biases from imbalanced data and enhancing convergence and performance.

PAN: Our research focuses on innovating the segmentation of pavement point cloud crack scenes. Unlike existing work, we extract 3D pavement crack information directly from MLS data, avoiding the information loss associated with traditional data conversion methods. To more fully leverage the rich information in point clouds, we herein introduce an advanced extended self-attention mechanism module called PC-Parallel.

The PC-Parallel module combines two types of attention modules, operating in parallel to more effectively capture contextual relationships in point cloud data. The spatial attention module targets the local relationships between different feature points, aiding in accurately locating small cracks. In contrast, the channel attention module focuses on long-range contextual information across the channel dimension to better understand overall road conditions. This organic combination significantly improves the model's ability to process pavement crack point cloud data. We constructed a large-scale pavement crack

point cloud dataset, i.e., LNTU-RDD-LiDAR, by ourselves, providing sufficient training and evaluation data for research. Through comprehensive empirical results, the effectiveness of our proposed method on this dataset was verified. Compared with traditional methods, our model directly detects cracks in MLS data, making the results more accurate and practical.

This innovative research not only made significant progress in methods, but also provided valuable resources for research in related fields by constructing a pavement crack point cloud dataset. Our work injects new ideas into the field of pavement crack information extraction and provides strong support for future research.

3. Materials and Methods

3.1. Point Attention Net Model Overview

As a general point cloud processing framework, PointNet++ excels in tasks like point cloud classification, semantic segmentation, and object detection, demonstrating wide applicability. However, in small target segmentation tasks, such as pavement crack extraction, PointNet++ uses a fixed receptive field. Although this field is gradually expanded through multiple Set Abstraction layers, it relies solely on the feature extraction stage of the PointNet layer to input several points into the fully connected layer, resulting in a relatively simple encoding method with low robustness.

To address these issues, this paper proposes a PAN network based on PointNet++, which directly processes unordered point sets as inputs and uses the Set Abstraction module to extract local features at different levels, capturing local structures in the point cloud. Through hierarchical subsampling and aggregation operations, point cloud information is captured at various scales, allowing the network to consider both local structures and global context. This improves the network's understanding of the overall and local structure of the point cloud, enhancing its ability to process point cloud data with multi-scale characteristics. Using symmetric functions to handle the arrangement of input point clouds makes the network insensitive to point arrangements, ensuring consistent outputs across different input configurations. This approach enhances the model's generalization and makes it more adaptable to point clouds with various shapes and structures, boosting its robustness in practical applications. Additionally, the proposed PC-Parallel module enlarges the model's receptive field and strengthens encoding robustness. This module enhances PointNet++'s performance in small target segmentation tasks such as pavement point cloud crack detection, allowing it to better adapt to varying scales and complexities of point cloud data and improving its effectiveness in real-world applications.

The PC-Parallel module enhances the network's capability to capture critical features over long distances, increasing adaptability and robustness. As illustrated in Figure 1, the PC-Parallel module is introduced after the SA block, where PointNet layer features are input. This module combines spatial and channel attention in parallel, leveraging the strengths of both mechanisms to improve point cloud segmentation performance. By modeling the attention matrix, the spatial attention component effectively captures the spatial relationship between any two points, enhancing the ability to identify local crack features and expanding the model's perception of the crack region, which enhances the ability to process crack shapes, filter out irrelevant and noisy points, reduce redundant information in the point cloud, and improve computational efficiency. Simultaneously, the channel attention component focuses on capturing remote context information in the channel dimension, emphasizing feature channels with significant differentiation of crack features while suppressing those with minor contributions. This reduces noise interference and redundant information, improving the model's overall performance.
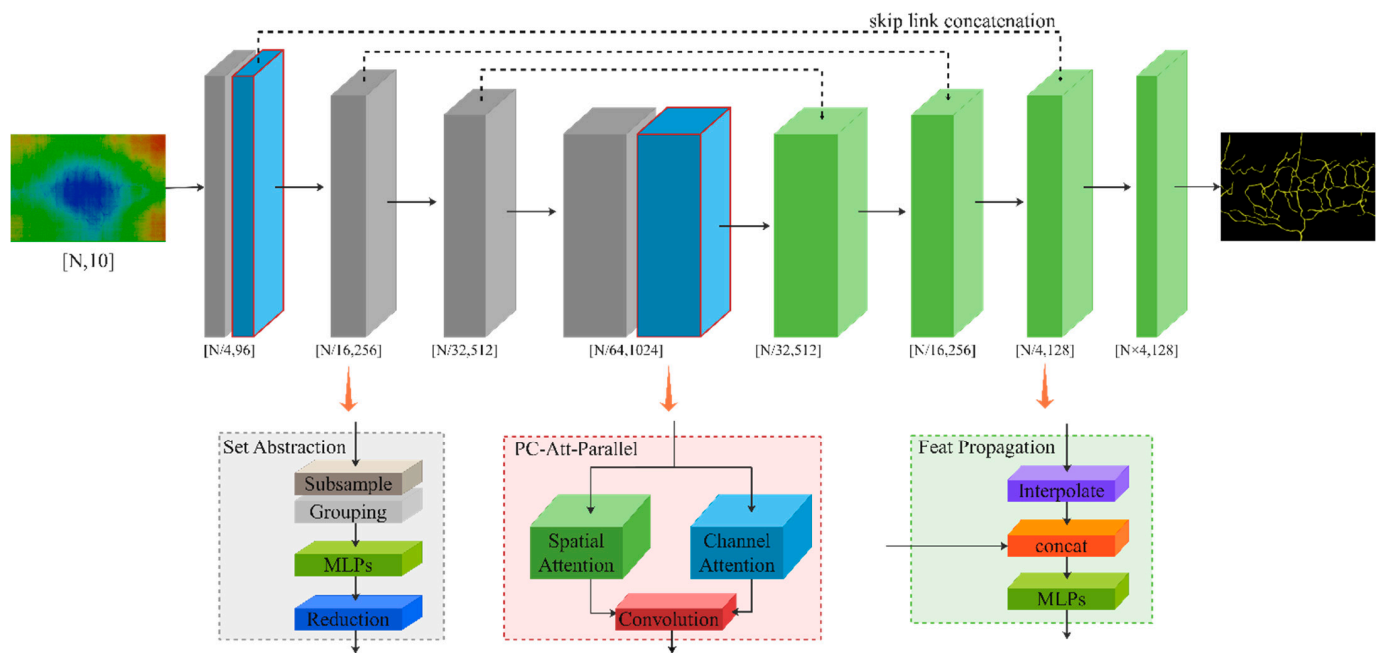
Figure 1. PAN model structure, using PointNet++ as the baseline code and adding PC-Parallel modules in the early and late stages of the network. Adding the PC-Parallel module in the early stage can increase the correlation and encoding between local geometric features; adding PC in the later-stage PC-Parallel module can enable better feature interaction in all aspects of the object and expand the original receptive field.

Finally, aggregating the outputs of the two attention modules enhances the recognition of crack points across spatial and feature dimensions, achieving more effective multi-scale feature fusion. This process captures fracture point characteristics at different sampling levels, allowing the model to better understand the global correlation among fracture points. The combined use of spatial and channel attention enables the model to fully perceive and interpret the information within crack points, thereby improving its understanding of the overall structure and pattern. By obtaining better feature representations, these features can be used more accurately to predict crack points and provide more powerful modeling capabilities for crack point cloud segmentation tasks. Moreover, Poly Loss is introduced to adjust the form of loss function to better balance the imbalance between crack points and background points and significantly improve the identification accuracy of crack areas.

3.2. PC-Parallel Module

The PC-Parallel module consists of two types of attention branches, as illustrated in Figure 2: the spatial attention branch learns the relationship between different feature points, while the channel attention module captures remote context information along the channel dimension. Finally, the outputs from these two attention modules are aggregated to achieve a more effective point-level feature representation.

Spatial Attention Branch: Spatial attention allows the module to selectively focus on local regions around crack points, enhancing local features and improving the perception of crack areas. This is crucial for crack point segmentation, as crack features are typically small and scattered. By incorporating spatial attention, the model becomes more adaptable to point clouds of varying shapes and structures, enabling a deeper understanding and exploitation of spatial relationships between points to better capture local crack details. This enables the model to focus more on areas with important structure. By guiding the model to focus on the key part of the crack point and filtering irrelevant points and noise points, the computational cost of processing redundant information is effectively reduced, and the overall work efficiency is improved. Furthermore, spatial attention helps the model

grasp the global correlations within the point cloud, enhancing its ability to capture overall structures by learning spatial relationships between points. This approach addresses the non-uniformity of point clouds, allowing the model to more accurately process crack points with varying densities and samples, thereby improving its effectiveness in handling pavement crack point cloud data.
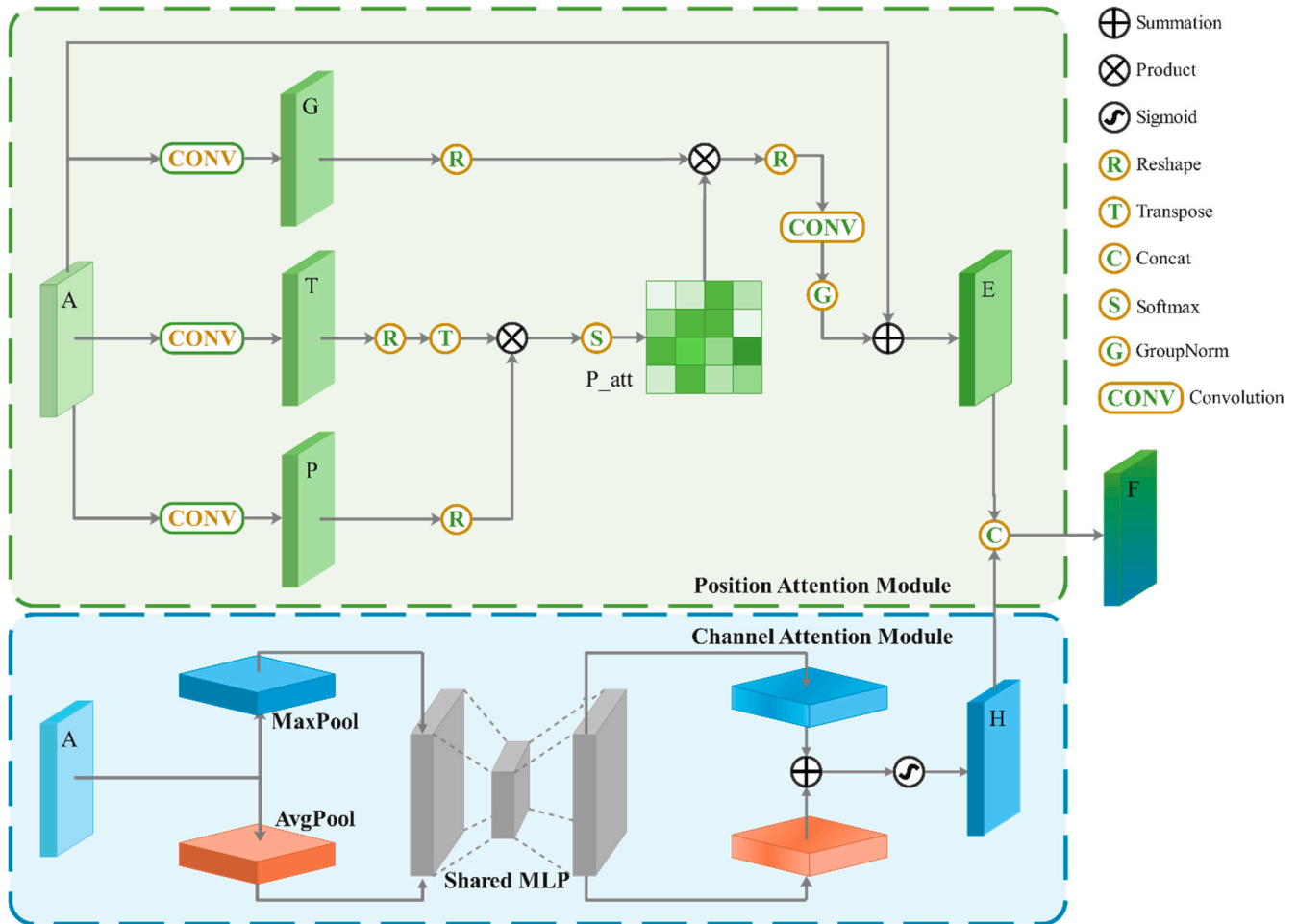

Figure 2. PC-Parallel module structure. The two attention modules work together to improve the point cloud model's understanding of different channel and position combinations and enhance complex structure modeling. Global structure understanding enables more accurate segmentation and better model adaptability, and dynamic weight adjustment promotes feature learning for generalization performance.

We first input local features $A \in \mathbb{R}^{B \times C \times N}$, which are initially processed through two convolutional layers to produce new feature maps T and P, respectively, represented as $T \in \mathbb{R}^{B \times \frac{C}{2} \times N}$, $P \in \mathbb{R}^{B \times \frac{C}{2} \times N}$. Then, we reshape them into $\mathbb{R}^{B \times N}$, where $N = H \times W$. Next, matrix multiplication is performed between the transpose of T and P, followed by a softmax layer to compute the spatial attention map $\text{P\_att} \in \mathbb{R}^{B \times N \times N}$:

$$softmax_{btp} = \frac{e^{\text{P\_att}_{btp}}}{\sum\limits_{l=1}^{N} e^{\text{P\_att}_{btl}}} \tag{1}$$

Among them, P_att_{btp} is a measure of the positional relationship between t and p. b represents the batch dimension, t represents the first spatial dimension, and p represents the second spatial dimension. We first input the feature $A \in \mathbb{R}^{B \times C \times N}$ into another convolu-

tional layer to generate a new feature map $G \in \mathbb{R}^{B \times \frac{C}{2} \times N}$ and then reshape it to $\mathbb{R}^{B \times N}$. Then, we perform matrix multiplication between P_att and G and reshape it to $\mathbb{R}^{B \times N}$. Finally, the results are input into a convolutional layer and are then element-wise summed with feature A after normalization operation to obtain the final output $E \in \mathbb{R}^{B \times C \times N}$:

$$E_{att} = \text{GN}\left(\text{Conv1}d_{\text{out}}\left(\text{softmax}\left(\frac{T^T \cdot P}{\sqrt{c \cdot n}} \right) \cdot G \right) \right) + \text{residual} \tag{2}$$

It can be seen from the formula that the final feature of each position $spatial\_att_E$ is the weighted sum of the features of all positions, taking into account the original features; c is the number of channels, and n is the number of points. Therefore, it has a global context view and selectively aggregates context according to the spatial attention map, helping the model focus on local areas, enhance local features, and reduce the interference of background point clouds. Similar semantic features improve each other, thereby improving intra-class compactness and semantic consistency.

Channel Attention Branch: Channel attention is crucial for dynamically adjusting feature weights across different channels in the point cloud during learning. It emphasizes the importance of various feature channels, enhancing crack point representation by adjusting each channel's weight. Important feature channels are highlighted by emphasizing those with significant differentiation for crack points. Meanwhile, redundant feature channels that contribute little to the crack point cloud segmentation task are suppressed, reducing noise interference and improving the model's overall performance and segmentation accuracy for cracks.

By introducing average pooling and max pooling operations, we integrated the spatial information of the feature map, producing two independent spatial context descriptors named AvgPool and MaxPool, representing average and max pooling features, respectively. These descriptors are then fed into a shared network consisting of a multilayer perceptron (MLP) with hidden layers. To balance parameter count and effectiveness, we set the hidden activation size to $\mathbb{R}^{\frac{C}{r} \times 1 \times 1}$, where r is the decay ratio. After applying the shared network to each descriptor and processing the merged features through the Sigmoid activation function, we obtain the channel attention map H_{att}. This map generation relies on learning a shared network with relatively few parameters, reducing computational load. Channel attention is calculated as follows:

$$H_{att} = \text{sigmoid}(\text{MLP}(\text{AvgPool}(A)) + \text{MLP}(\text{MaxPool}(A))) \tag{3}$$

This channel attention design enables the model to adaptively focus on each channel's information based on task requirements, thereby enhancing the network's sensitivity to point cloud features. This mechanism effectively captures inter-channel relationships in point clouds, offering a more accurate feature representation for segmentation tasks and improving the model's performance and generalization capabilities.

Finally, by integrating the channel attention module and the spatial attention module $F \in \mathbb{R}^{B \times C \times N}$ through matrix concatenation, the model can learn the relationship between the channel and the location more comprehensively so that the model can understand and represent the details and global information of pavement cracks more comprehensively. The synergistic effect of these two attention modules allows the model to more effectively comprehend the combination of different channels at various positions within the point cloud. This capability enables the model to adapt more efficiently to cracks of diverse shapes, sizes, and positions, ensuring stable performance across different types of cracks. Additionally, this synergy enhances the model's ability to accurately locate and segment cracks, significantly improving the precision and accuracy of segmentation tasks. By understanding the spatial and channel relationships within the point cloud, the model can deliver more reliable and consistent results in crack detection and analysis. In summary, this

fusion strategy provides a more comprehensive and flexible feature learning mechanism for point cloud segmentation tasks.

$$F_{att} = E_{att} + H_{att} \tag{4}$$

3.3. Poly Loss Function

In the original PointNet++ framework, NLL Loss (Negative Log Likelihood Loss) is used as a standard loss function and is calculated based on probability distribution. By computing the negative log-likelihood between the predicted class probability distribution of pixels or points and the true label, the model is guided to optimize. However, for fracture point cloud segmentation tasks, NLL Loss overlooks the local structure of fractures and the order of points, leading to a lack of global and local structural information and poor robustness against point cloud rotations or translations.

Given the characteristics of pavement crack point cloud data, there is an imbalance in the distribution of crack sample points and background points. To effectively address this issue in pavement crack segmentation, this study introduced Poly Loss [59]. The core idea of Poly Loss is to enhance the deep learning model's robustness and accuracy by designing specific loss functions tailored to the segmentation task. In the point cloud pavement crack segmentation task, cracks usually occupy a small part of the point cloud, while most of the area is normal pavement, and the data imbalance will cause the original loss function to be unable to effectively identify the crack area. By introducing weighting factors or polynomial terms, Poly Loss effectively addresses the imbalance between crack and background points, significantly enhancing crack region identification accuracy. It makes the model more sensitive to edge and detail information, allowing better preservation and recognition of cracks' fine structure. By adjusting polynomial coefficients, Poly Loss optimizes the model's predictive performance for crack point cloud segmentation, in line with the task requirements for pavement crack point cloud segmentation. We must balance the prediction ability of crack point and background point of the model to improve the overall model performance. Poly Loss is defined as follows:

$$L_{\text{Poly}-1} = -\log(P_t) + \epsilon_1(1 - P_t) \tag{5}$$

Among them, ϵ_1 is an additional hyperparameter used to adjust the first polynomial coefficient, and P_t is the predicted probability of the crack point category. By introducing an additional term $\epsilon_1(1 - P_t)$ into the original cross-entropy loss function $-\log(p_t)$ to adjust the first polynomial coefficient, the classification performance is improved. At the same time, the softmax operation is introduced to effectively deal with the problem of category imbalance and strengthen the learning of minority categories.

In this paper, we employed the Poly Loss function to train the entire network architecture. For the semantic segmentation of pavement cracks in point cloud data, Poly Loss effectively addresses category imbalance by adjusting polynomial coefficients to weight crack and background points. This approach enhances the model's robustness against noise and outliers, significantly boosting performance. It also improves the model's stability and reliability in complex scenarios.

4. Results

4.1. Implementation

This study used a Ubuntu 20.04 operating system, python v3.8.18, pytorch v1.12.1 deep learning framework, and CUDA v11.5.50. The CPU is Intel(R) Xeon(R) Gold 6133 CPU @ 2.50 GHz, and the GPU is NVIDIA GeForce RTX 3090. All models were trained 300 times, the initial learning rate was 0.001, each training epoch updated the learning rate, the optimizer was Adam, and the batch size was 4. The hyperparameters involved in the Point Attention Net model are shown in Table 1. The proposed Point Attention Net model was trained using LNTU-RDD-LiDAR Road-1 and Road-4, verified using LNTU-RDD-LiDAR Road-3, and tested using LNTU-RDD-LiDAR Road-2.

Table 1. Point Attention Net model parameter settings.

Parameter	Value
Learning rate	0.001
Batch size	4
Optimizer	Adam
Momentum parameter	0.9
Training cycle	300

4.2. Quantitative Assessment Measures

To comprehensively evaluate the effectiveness of construction disaster detection, this study used IoU, $mIoU$, P_{re}, R_{ec}, and F_1 as the main evaluation metrics. Here, P_{re} calculates the percentage of correctly predicted pavement cracks to assess model effectiveness, as shown in Equation (6). R_{ec} calculates the ratio of correctly identified crack points among all crack points to evaluate detection completeness, as shown in Equation (7). F_1 is the comprehensive evaluation index of P_{re} and R_{ec}, as shown in Formula (8). TP represents the number of point clouds in the fracture area correctly identified; FP represents the number of point clouds incorrectly identified as fractures, and FN represents the number of fracture point clouds not identified.

$$P_{re} = \frac{TP}{TP + FP} \tag{6}$$

$$R_{ec} = \frac{TP}{TP + FN} \tag{7}$$

$$F_1 = \frac{2 \times P_{re} \times R_{ec}}{P_{re} + R_{ec}} \tag{8}$$

$mIoU$ is an indicator used to evaluate the performance of deep learning point cloud segmentation tasks. Its calculation method is shown in Equation (10), where N represents the number of categories. Among them, IoU is only used as an evaluation index for each crack category segmentation, such as Equation (9).

$$IoU_i = \frac{TP_i}{TP_i + FP_i + FN_i} \tag{9}$$

$$mIoU = \frac{1}{N} \sum_{i=1}^{N} \frac{TP_i}{TP_i + FP_i + FN_i} \tag{10}$$

4.3. LNTU-RDD-LiDAR Dataset

In similar studies, most pavement crack information extraction methods focus on a single crack type, leading to incomplete crack type coverage. To address this, we developed a pavement crack point cloud dataset named LNTU-RDD-LiDAR for the segmentation task of pavement cracks on point cloud data. Provincial roads served as the collection source, where varying weather conditions and terrain changes resulted in a wider variety of surface cracks. This diversity makes regular, timely, and comprehensive road maintenance more challenging.

Collect: In the data acquisition stage, a laser WPL7-808-10W was used as a point cloud scanner, and combined with an airborne laser measurement system, the road point cloud was recorded to construct the original point cloud dataset. The airborne laser measurement system consists of two WPL7-808-10W lasers, two 3D AT-C2 cameras, and an inertial navigation unit. The clear distance between the bottom of the laser sensor and the road surface was 1900 mm. The point cloud measurement covers a range of 4 m to 2.1 m, the ranging accuracy reaches 0.5 mm, the absolute accuracy is 10 mm, and the elevation error between lanes is only 3 mm. The selection of these scanning parameters and settings was designed to ensure that the obtained point cloud data reached an accuracy level of 0.01 mm in resolution. The entire experimental dataset contains 188 pavement point cloud segments,

each of which contains an average of 20 million points. Each segment has a width of 3.8 m and a length of about 16 m. Table 2 shows the size and high resolution of the dataset that provides detailed and comprehensive pavement information for this study, which lays a solid foundation for subsequent crack detection and analysis.

Table 2. Parameters of 3D laser road acquisition equipment.

Hyperparameter	Range
Point cloud resolution	0.01 mm
Clear distance between instruments and ground	1900 mm
Point cloud measurement coverage	4 m–2.1 m
Point cloud ranging accuracy	0.5 mm
Point cloud absolute accuracy	1 cm
Elevation error between lanes	3 mm
Use environment	Free from ambient light all day long
Detection speed	40 km/h

Mark: After collecting and processing the pavement point cloud data through the airborne laser measurement system, a pavement crack point cloud labeling task was carried out. The point cloud labeling tool CloudCompare_v2.13 is used in this paper. CloudCompare_v2.13 is a powerful open-source point cloud data processing software with rich point cloud processing functions, including import, export, and editing. Its intuitive 3D visual interface makes visual analysis and interactive operation more convenient. CloudCompare_v2.13 can efficiently process and annotate point cloud data, which provides a good data foundation for subsequent tasks.

Table 3 shows the detailed requirements of the labeling task, including the accurate location, shape, and size of the cracks. The whole labeling process covered 188 crack point cloud pavement segments, among which the ratio of crack point to non-crack point was about 0.6:9.4. Then, the LNTU-RDD-LiDAR experimental data were divided into four subsets: Road-1, Road-2, Road-3, and Road-4, among which Road-1 contains 51 pavement segments. Road-2 contains the latter 17 road segments, Road-3 contains 59 road segments, and Road-4 contains 61 road segments. At the same time, the background points were also marked to consider the overall road surface information. This provides a high-quality labeled dataset for semantic segmentation of pavement crack point cloud-type data, as illustrated in Figure 3. Through multi-person verification and quality control, the accuracy and consistency of the annotation results were ensured, which provides a reliable basis for the training and evaluation of the model.

Table 3. LNTU_RDD_PCD dataset description table.

	Hyperparameter	Range
	Point cloud dimension	3D
Dataset size	Dataset size	713 GB
	Number of datasets	188
	Position coordinate	x, y, z
	Color	R, G, B
Point cloud attribute	Intensity	Intensity
	Category	Label
	Normal vector	Nx, Ny, Nz
Labeling information	Class tag	Pavement point, crack point
	Annotation method	Manual labeling
	Number of training sets	127
	Number of verification sets	21
Dataset partitioning	Number of test sets	20
	Partition method	Random sampling

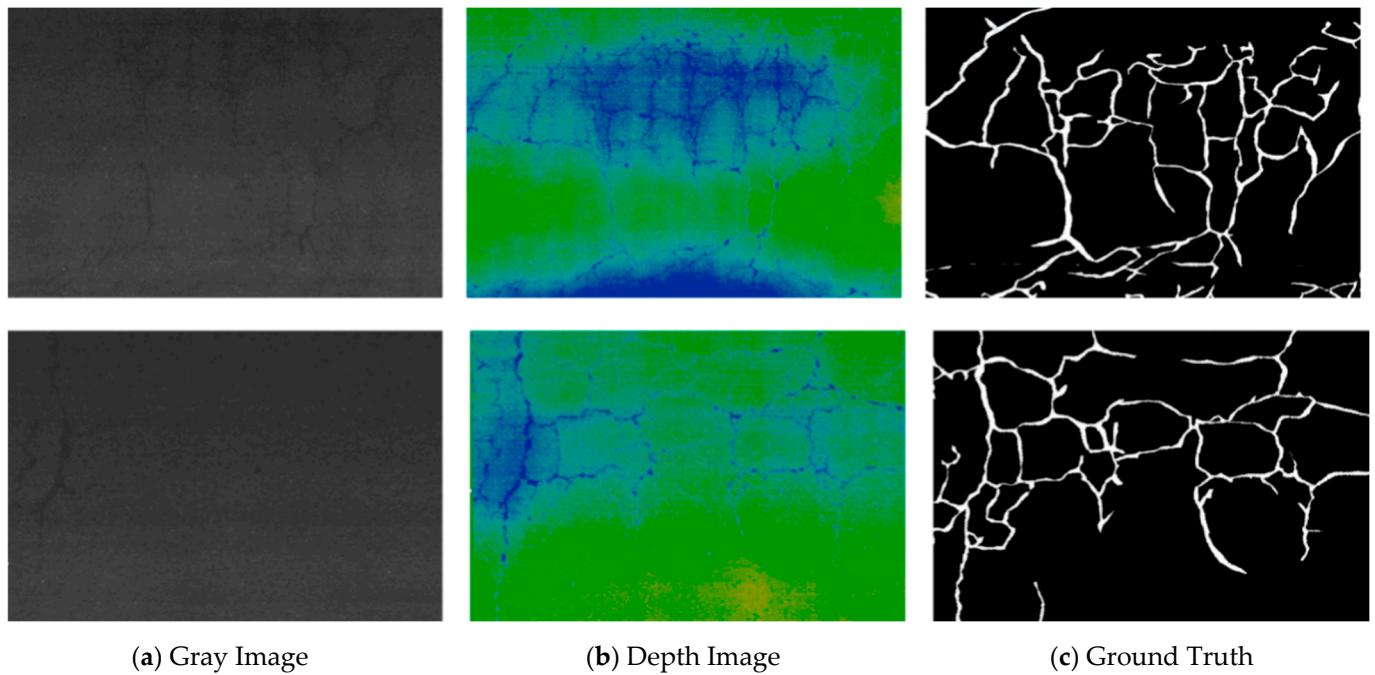(**a**) Gray Image   (**b**) Depth Image   (**c**) Ground Truth

Figure 3. LNTU-RDD-LiDAR pavement crack point cloud dataset. (**a**) Generate a gray value image based on the z value of the point cloud. (**b**) Generate a depth value image based on the z value of the point cloud. (**c**) Corresponds to the true value of the manually labeled road crack.

4.4. Comparative Experiment

To verify the effectiveness of the proposed Point Attention Net model for pavement crack point cloud data segmentation, this study tested the LNTU-RDD-LiDAR point cloud dataset and compared the model with existing methods, including PointNet [40], Pointnet++ [41], Point Transformer [47], and PointMLP [50]. PointNet, a pioneer of point-by-point classification, uses a symmetric function invariant to permutations but struggles to learn local features in complex road scenarios. Point Transformer introduces a self-attention mechanism to capture the spatial structure of point clouds and establish a global context. PointMLP incorporates a lightweight geometric affine module to enhance performance.

From Table 4, it is clear that the PAN model improves in mIoU, R_{ec}, F1, and Acc compared to the PointNet, PointNet++, Point Transformer, and PointMLP models. mIoU improved by 0.9% relative to the Point Transformer model, Acc improved by 0.1% relative to the PointMLP model, and F1 improved by 1.3% relative to the Point Transformer model. As illustrated in Figure 4, experimental results show that, compared with most traditional deep learning algorithms, the Point Attention Net model mainly benefits from two-branch attention fusion, which enhances the ability of description feature coding.

Table 4. Comparison of semantic segmentation results between PAN and PointNet, PointNet++, PointTransformer, and PointMLP on the LNTU-RDD-LiDAR dataset.

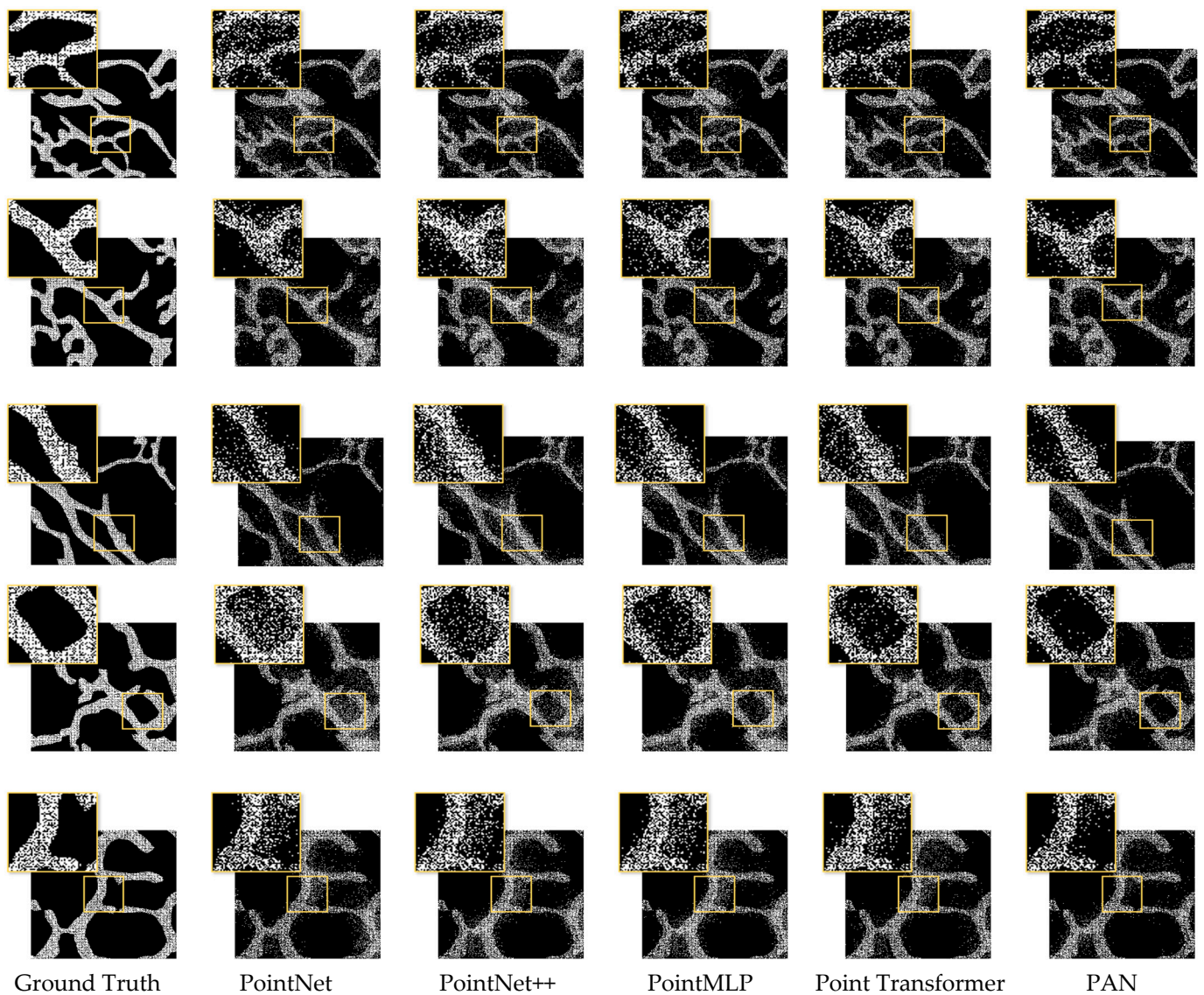| Model	R_{ec} ↑	F1 ↑	Acc ↑	mIoU ↑	Param.
PointNet	48.1	64.9	89.1	68.4	3.5 M
PointNet++	59.4	70.4	90.4	69.2	1.41 M
PointMLP	61.3	73.6	91.4	74.3	12.6 M
Point Transformer	62.2	74.1	91.5	74.5	7.8 M
PAN	**67.1**	**75.4**	**91.5**	**75.4**	**1.76 M**

Figure 4. Visual comparison results and corresponding detailed observations of pavement crack information extraction between PAN and PointNet, PointNet++, PointTransformer, and PointMLP methods on the LNTU-RDD-LiDAR dataset.

4.5. Ablation Experiments

PC-Parallel: Table 5 shows the results of the gradual improvement of pavement crack information extraction performance with the increase of attention blocks. Compared with the original Pointnet ++ baseline network, the introduction of location attention module and channel attention module can significantly improve the segmentation effect of pavement crack point cloud. In this case, mIoU reached 75.4%, R_{ec} reached 67.1%, F1 reached 75.4%, and Acc reached 91.5%.

Table 5. Ablation experiments of PC-Att-Parallel module test on LNTU-RDD-LiDAR dataset.

Model	R_{ec} ↑	F1 ↑	Acc ↑	mIoU ↑
BaseNet	48.1	64.9	89.1	68.4
+Position Attention	56.2	70.7	90.9	72.2
+Channel Attention	57.9	70.8	90.7	72.1
+PC-Parallel	**67.1**	**75.4**	**91.5**	**75.4**

As shown in Figure 5, incorporating the PC-Parallel module effectively addresses issues of incomplete segmentation and imprecise boundaries. The location attention module learns spatial interdependencies between features, while the channel attention module captures dependencies between channels. Both attention blocks positively impact pavement crack extraction performance. Introducing these modules and fusing their features significantly enhances performance. Consequently, the two-branch attention block demonstrates remarkable effectiveness in the crack segmentation task of pavement point cloud data.


| Ground Truth | BaseNet | +Position Attention | +Channel Attention | +PC-Parallel |
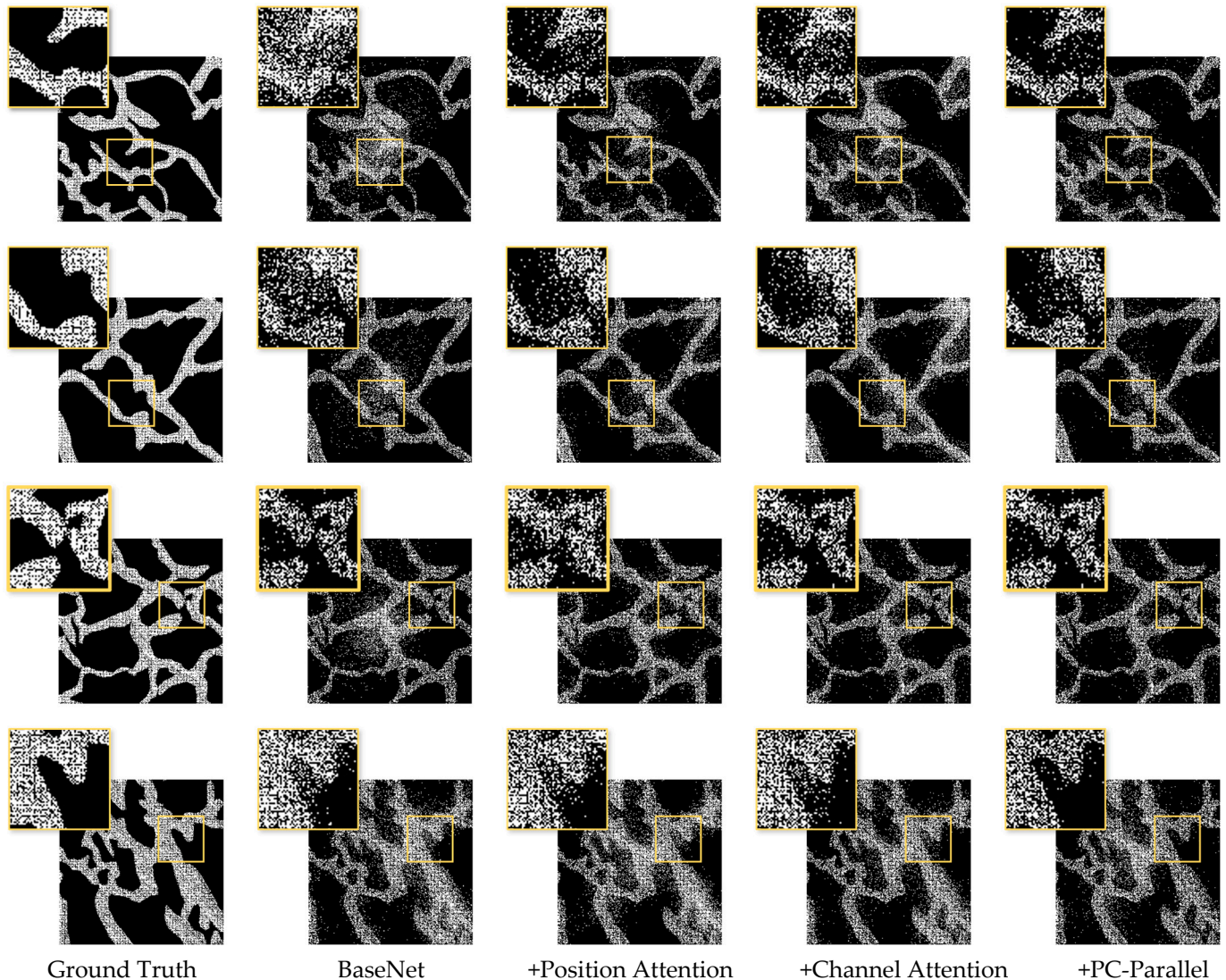
Figure 5. The PC-Parallel module on LNTU-RDD-LiDAR dataset tested the results of visual comparison and detailed observation of the ablation experimental method.

Poly Loss: Table 6 shows the results of the gradual improvement of pavement crack information extraction performance with the increase of loss function. Compared to the original PointNet++ baseline code, the introduction of Poly Loss to handle class imbalance and focus on difficult-to-classify samples effectively improves the performance and robustness of the model. In this case, mIoU reached 74.0%, R_{ec} reached 63.3%, F1 reached 73.5%, and Acc reached 91.1%. Therefore, Poly Loss shows remarkable effectiveness in the pavement point cloud data crack segmentation task. As one can see from Figure 6, replacing the Loss function in the original PointNet++ baseline code with Poly Loss effectively solves the class imbalance problem. The loss function highlights the features of foreground points and background points so that the model can better distinguish the slit points and other points.

Table 6. Ablation experiments of Poly Loss test on LNTU-RDD-LiDAR dataset.

Method	$R_{ec} \uparrow$	F1 $\uparrow$	Acc $\uparrow$	mIoU $\uparrow$
BaseNet	48.1	64.9	89.1	68.4
+Poly Loss	**63.3**	**73.5**	**91.1**	**74.0**
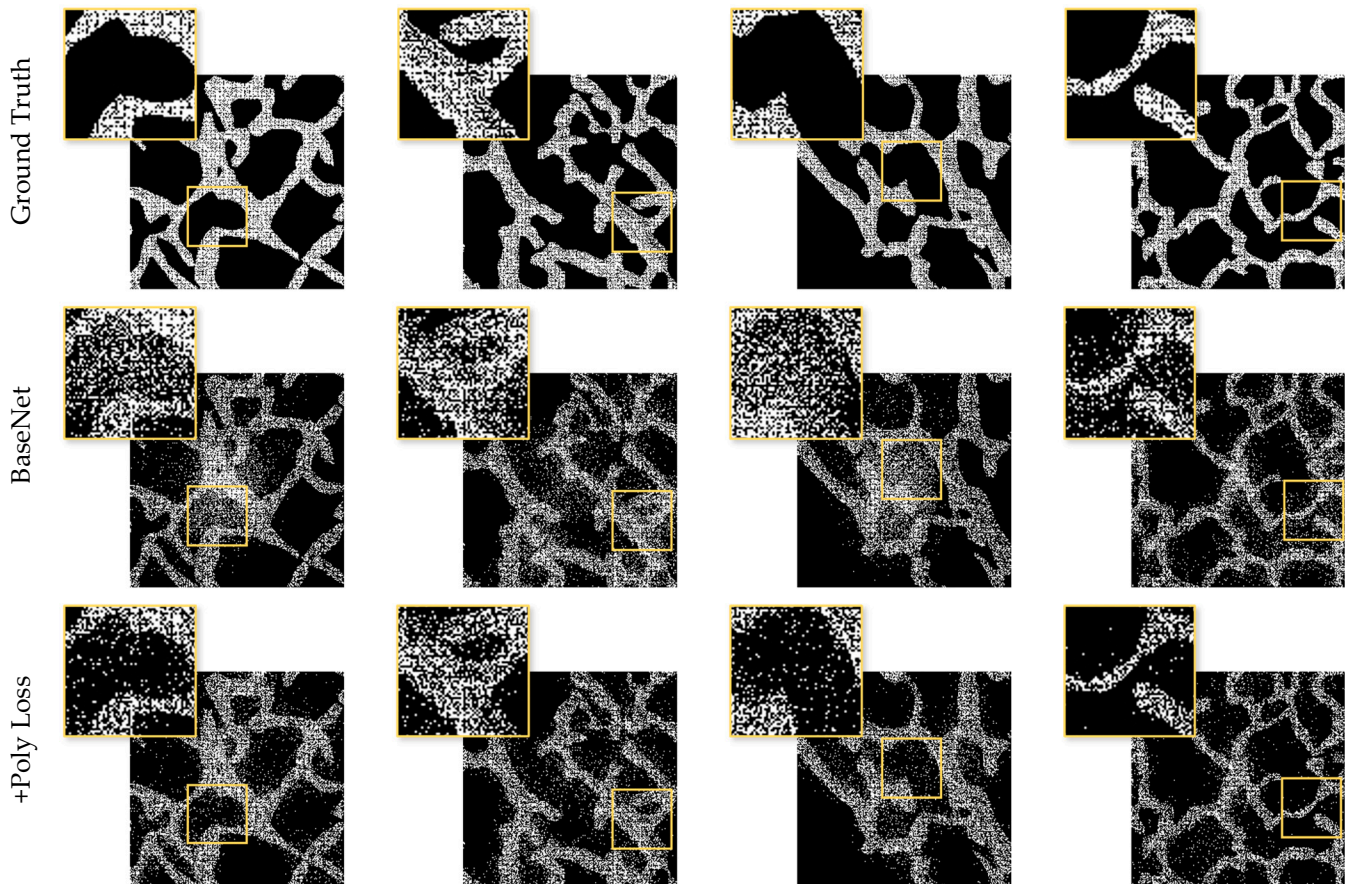
Figure 6. Visual comparison and detailed observation of ablation experiments with Poly Loss test on LNTU-RDD-LiDAR dataset.

From Figure 6, replacing the Loss function in the original PointNet++ baseline code with Poly Loss effectively solves the class imbalance problem. The loss function highlights the features of foreground points and background points so that the model can better distinguish the slit points and other points.

5. Discussion

In this paper, the 3D laser point cloud pavement crack data have better robustness than a 2D image under varying illumination conditions and in a low-intensity contrast environment. It can effectively deal with various kinds of rust and oil covering the road surface. Experiments on the proposed model show that the indexes of mIoU, Acc, F1, and R_{ec} are significantly improved, which confirms the effectiveness of the proposed method compared with traditional methods. The introduction of the Poly Loss function helps to better capture the edges of cracks and details.

However, the dataset in this paper includes a large number of 3D point cloud data of pavement cracks collected under clear-weather conditions. However, these data do not fully cover all environmental conditions. On rainy and snowy days, water and snow can mask pavement cracks, making it difficult for laser scans to capture precise crack features.

In addition, point cloud data may introduce additional noise, reducing the detection effect and recognition ability of the model.

To overcome these limitations and further explore the application of 3D point cloud datasets in pavement crack segmentation tasks, future research will focus on the following aspects:

1. Expanding datasets and multi-source data integration: We plan to collect more samples under various environmental conditions to enhance the generalization ability of the model. By combining other sensor data such as RGB images and thermal imaging, the deficiency of point cloud data can be supplemented to provide richer environmental information. In addition, we will develop automated data-labeling tools to automatically generate annotated data using deep learning and machine learning algorithms. This will help speed up labeling, reduce reliance on manual labeling, and improve the consistency and accuracy of labeling;

2. Technology integration and practical application: We will study how to seamlessly integrate the technology into the existing monitoring system and apply it to the actual road maintenance scenario. This includes addressing compatibility issues and adding new features without affecting existing system functionality. To ensure the effectiveness of the technology in practical scenarios, we plan to promote the application in cooperation with experts in relevant fields. This will help us identify and address potential implementation issues and optimize the technology based on actual needs. At the same time, we will evaluate implementation and maintenance costs to design solutions that are both efficient and cost-effective to support a wide range of applications.

Through these efforts, we hope to significantly improve the usefulness of the dataset and the applicability of the technology, providing a solid foundation for future road monitoring and maintenance.

6. Conclusions

This paper introduces the Point Attention Net (PAN) network for extracting 3D pavement crack information, aiming to overcome the limitations of previous point cloud-based deep learning approaches. The PAN network incorporates a novel PC-Parallel module that was specifically designed to learn spatial interdependencies and channel dependencies of features separately. This design significantly enhanced the performance of point cloud pavement crack segmentation by allowing the network to more effectively capture and process the intricate details of crack features. Additionally, boundary refinement and class imbalance issues in point cloud pavement crack segmentation were addressed by introducing the Poly Loss function. The test results on the LNTU-RDD-LiDAR dataset show that the proposed method has excellent performance on mIoU, R_{ec}, F1, and Acc, reaching 75.4%, 75.4%, 67.1%, and 86.8%, respectively. In comparison to existing point cloud segmentation methods, the proposed approach demonstrates superior performance, as the proposed method improved the mIoU and F1 indexes by 1.1% and 1.3%, respectively. The experimental results demonstrate that the proposed method significantly enhances point cloud pavement crack segmentation.

In future studies, we will explore more effective class-balancing strategies to enhance the model's performance and generalization on unbalanced datasets such as highway pavement crack point cloud data. Additionally, we aim to increase the number of samples and scene categories, expanding the model's training set to encompass a broader range of practical situations, thereby improving its generalization ability when encountering unknown data.

Author Contributions: Conceptualization, J.F. and S.S.; data curation, J.F.; funding acquisition, W.S. and S.S.; investigation, J.F. and J.Z.; methodology, J.F.; project administration, W.S.; funding acquisition, W.S.; resources, G.J. (Guohui Jia) and G.J. (Guang Jin); software, J.F.; visualization, J.F.; writing—original draft, J.F.; writing—review and editing, S.S., J.Z. and G.J. (Guohui Jia). All authors have read and agreed to the published version of the manuscript.

References

1. Liu, Z.; Cao, Y.; Wang, Y.; Wang, W. Computer vision-based concrete crack detection using U-net fully convolutional networks. *Autom. Constr.* **2019**, *104*, 129–139. [CrossRef]
2. Oliveira, H.; Correia, P.L. Automatic road crack segmentation using entropy and image dynamic thresholding. In Proceedings of the 2009 17th European Signal Processing Conference, Glasgow, UK, 24–28 August 2009; pp. 622–626.
3. Peng, L.; Chao, W.; Shuangmiao, L.; Baocai, F. Research on crack detection method of airport runway based on twice-threshold segmentation. In Proceedings of the 2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC), Qinhuangdao, China, 18–20 September 2015; pp. 1716–1720.
4. Lim, R.S.; La, H.M.; Shan, Z.; Sheng, W. Developing a crack inspection robot for bridge maintenance. In Proceedings of the 2011 IEEE International Conference on Robotics and Automation, Shanghai, China, 9–13 May 2011; pp. 6288–6293.
5. Lim, R.S.; La, H.M.; Sheng, W. A robotic crack inspection and mapping system for bridge deck maintenance. *IEEE Trans. Autom. Sci. Eng.* **2014**, *11*, 367–378. [CrossRef]
6. Gavilán, M.; Balcones, D.; Marcos, O.; Llorca, D.F.; Sotelo, M.A.; Parra, I.; Ocaña, M.; Aliseda, P.; Yarza, P.; Amírola, A. Adaptive road crack detection system by pavement classification. *Sensors* **2011**, *11*, 9628–9657. [CrossRef] [PubMed]
7. Zou, Q.; Cao, Y.; Li, Q.; Mao, Q.; Wang, S. CrackTree: Automatic crack detection from pavement images. *Pattern Recognit. Lett.* **2012**, *33*, 227–238. [CrossRef]
8. Tian, F.; Zhao, Y.; Che, X.; Zhao, Y.; Xin, D. Concrete crack identification and image mosaic based on image processing. *Appl. Sci.* **2019**, *9*, 4826. [CrossRef]
9. Cao, X.; Li, T.; Bai, J.; Wei, Z. Identification and Classification of Surface Cracks on Concrete Members Based on Image Processing. *Trait. Du Signal* **2020**, *37*, 519–525. [CrossRef]
10. Kang, Y.; Yu, A.; Zeng, W. Construction of Concrete Surface Crack Recognition Model Based on Digital Image Processing Technology. *J. Phys. Conf. Ser.* **2021**, *2074*, 012067. [CrossRef]
11. Tong, Z.; Yuan, D.; Gao, J.; Wei, Y.; Dou, H. Pavement-distress detection using ground-penetrating radar and network in networks. *Constr. Build. Mater.* **2020**, *233*, 117352. [CrossRef]
12. Zhang, J.; Pu, J.; Xue, J.; Yang, M.; Xu, X.; Wang, X.; Wang, F.-Y. HiVeGPT: Human-machine-augmented intelligent vehicles with generative pre-trained transformer. *IEEE Trans. Intell. Veh.* **2023**, *8*, 2027–2033. [CrossRef]
13. Turkan, Y.; Hong, J.; Laflamme, S.; Puri, N. Adaptive wavelet neural network for terrestrial laser scanner-based crack detection. *Autom. Constr.* **2018**, *94*, 191–202. [CrossRef]
14. Wang, L.; Zhang, X.; Song, Z.; Bi, J.; Zhang, G.; Wei, H.; Tang, L.; Yang, L.; Li, J.; Jia, C. Multi-modal 3D Object Detection in Autonomous Driving: A Survey and Taxonomy. *IEEE Trans. Intell. Veh.* **2023**, *8*, 3781–3798. [CrossRef]
15. Feng, H.; Ma, L.; Yu, Y.; Chen, Y.; Li, J. SCL-GCN: Stratified Contrastive Learning Graph Convolution Network for pavement crack detection from mobile LiDAR point clouds. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *118*, 103248. [CrossRef]
16. Cheng, H.; Shi, X.; Glazier, C. Real-time image thresholding based on sample space reduction and interpolation approach. *J. Comput. Civ. Eng.* **2003**, *17*, 264–272. [CrossRef]
17. Huang, Y.; Xu, B. Automatic inspection of pavement cracking distress. *J. Electron. Imaging* **2006**, *15*, 013017. [CrossRef]
18. Ying, L.; Salari, E. Beamlet transform-based technique for pavement crack detection and classification. *Comput.-Aided Civ. Infrastruct. Eng.* **2010**, *25*, 572–580. [CrossRef]
19. Zhang, A.; Li, Q.; Wang, K.C.; Qiu, S. Matched filtering algorithm for pavement cracking detection. *Transp. Res. Rec.* **2013**, *2367*, 30–42. [CrossRef]
20. Zalama, E.; Gómez-García-Bermejo, J.; Medina, R.; Llamas, J. Road crack detection using visual features extracted by Gabor filters. *Comput.-Aided Civ. Infrastruct. Eng.* **2014**, *29*, 342–358. [CrossRef]
21. Zhang, L.; Yang, F.; Zhang, Y.D.; Zhu, Y.J. Road crack detection using deep convolutional neural network. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016; pp. 3708–3712.
22. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [CrossRef]
23. Cha, Y.J.; Choi, W.; Büyüköztürk, O. Deep learning-based crack damage detection using convolutional neural networks. *Comput.-Aided Civ. Infrastruct. Eng.* **2017**, *32*, 361–378. [CrossRef]
24. An, Y.-K.; Jang, K.; Kim, B.; Cho, S. Deep learning-based concrete crack detection using hybrid images. In Proceedings of the Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2018, Denver, CO, USA, 5–8 March 2018; pp. 273–284.

25. Fan, Z.; Wu, Y.; Lu, J.; Li, W. Automatic pavement crack detection based on structured prediction with the convolutional neural network. *arXiv* 2018, arXiv:1802.02208.

26. Yang, X.; Li, H.; Yu, Y.; Luo, X.; Huang, T.; Yang, X. Automatic pixel-level crack detection and measurement using fully convolutional network. *Comput.-Aided Civ. Infrastruct. Eng.* **2018**, *33*, 1090–1109. [CrossRef]

27. Huang, H.-W.; Li, Q.-T.; Zhang, D.-M. Deep learning based image recognition for crack and leakage defects of metro shield tunnel. *Tunn. Undergr. Space Technol.* **2018**, *77*, 166–176. [CrossRef]

28. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.

29. Dorafshan, S.; Thomas, R.J.; Maguire, M. Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete. *Constr. Build. Mater.* **2018**, *186*, 1031–1045. [CrossRef]

30. Song, L.; Wang, X. Faster region convolutional neural network for automated pavement distress detection. *Road Mater. Pavement Des.* **2021**, *22*, 23–41. [CrossRef]

31. Yang, F.; Zhang, L.; Yu, S.; Prokhorov, D.; Mei, X.; Ling, H. Feature pyramid and hierarchical boosting network for pavement crack detection. *IEEE Trans. Intell. Transp. Syst.* **2019**, *21*, 1525–1535. [CrossRef]

32. Jahanshahi, M.R.; Jazizadeh, F.; Masri, S.F.; Becerik-Gerber, B. Unsupervised approach for autonomous pavement-defect detection and quantification using an inexpensive depth sensor. *J. Comput. Civ. Eng.* **2013**, *27*, 743–754. [CrossRef]

33. Ouyang, W.; Xu, B. Pavement cracking measurements using 3D laser-scan images. *Meas. Sci. Technol.* **2013**, *24*, 105204. [CrossRef]

34. Guan, H.; Li, J.; Yu, Y.; Chapman, M.; Wang, H.; Wang, C.; Zhai, R. Iterative tensor voting for pavement crack extraction using mobile laser scanning data. *IEEE Trans. Geosci. Remote Sens.* **2014**, *53*, 1527–1537. [CrossRef]

35. Yu, Y.; Li, J.; Guan, H.; Wang, C. 3D crack skeleton extraction from mobile LiDAR point clouds. In Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada, 13–18 July 2014; pp. 914–917.

36. Jiang, H.; Li, Q.; Jiao, Q.; Wang, X.; Wu, L. Extraction of wall cracks on earthquake-damaged buildings based on TLS point clouds. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 3088–3096. [CrossRef]

37. Xu, X.; Yang, H. Intelligent crack extraction and analysis for tunnel structures with terrestrial laser scanning measurement. *Adv. Mech. Eng.* **2019**, *11*, 1687814019872650. [CrossRef]

38. Ma, L.; Li, J. SD-GCN: Saliency-based dilated graph convolution network for pavement crack extraction from 3D point clouds. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *111*, 102836. [CrossRef]

39. Zhong, M.; Sui, L.; Wang, Z.; Hu, D. Pavement Crack Detection from Mobile Laser Scanning Point Clouds Using a Time Grid. *Sensors* **2020**, *20*, 4198. [CrossRef]

40. Qi, C.R.; Su, H.; Mo, K.; Guibas, L.J. Pointnet: Deep learning on point sets for 3d classification and segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 652–660.

41. Qi, C.R.; Yi, L.; Su, H.; Guibas, L.J. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *Adv. Neural Inf. Process. Syst.* **2017**, *30*.

42. Wang, Y.; Sun, Y.; Liu, Z.; Sarma, S.E.; Bronstein, M.M.; Solomon, J.M. Dynamic graph cnn for learning on point clouds. *ACM Trans. Graph. (Tog)* **2019**, *38*, 146. [CrossRef]

43. Fan, S.; Dong, Q.; Zhu, F.; Lv, Y.; Ye, P.; Wang, F.-Y. SCF-Net: Learning spatial contextual features for large-scale point cloud segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 14504–14513.

44. Xu, M.; Ding, R.; Zhao, H.; Qi, X. Paconv: Position adaptive convolution with dynamic kernel assembling on point clouds. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 3173–3182.

45. Yu, H.; Huang, H.; Zheng, J.; Zhao, T.; Zhou, X. Non-contact on-line inspection method for surface defects of cross-rolling piercing plugs for seamless steel tubes. *China Mech. Eng.* **2022**, *33*, 1717.

46. Guo, M.-H.; Cai, J.-X.; Liu, Z.-N.; Mu, T.-J.; Martin, R.R.; Hu, S.-M. Pct: Point cloud transformer. *Comput. Vis. Media* **2021**, *7*, 187–199. [CrossRef]

47. Zhao, H.; Jiang, L.; Jia, J.; Torr, P.H.; Koltun, V. Point transformer. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021; pp. 16259–16268.

48. Engel, N.; Belagiannis, V.; Dietmayer, K. Point transformer. *IEEE Access* **2021**, *9*, 134826–134840. [CrossRef]

49. Zhou, Y.; Ji, A.; Zhang, L. Sewer defect detection from 3D point clouds using a transformer-based deep learning model. *Autom. Constr.* **2022**, *136*, 104163. [CrossRef]

50. Ma, X.; Qin, C.; You, H.; Ran, H.; Fu, Y. Rethinking network design and local geometry in point cloud: A simple residual MLP framework. *arXiv* 2022, arXiv:2202.07123.

51. Ran, H.; Liu, J.; Wang, C. Surface representation for point clouds. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 18942–18952.

52. Qian, G.; Li, Y.; Peng, H.; Mai, J.; Hammoud, H.; Elhoseiny, M.; Ghanem, B. Pointnext: Revisiting pointnet++ with improved training and scaling strategies. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 23192–23204.

53. Zhang, A.; Wang, K.C.; Li, B.; Yang, E.; Dai, X.; Peng, Y.; Fei, Y.; Liu, Y.; Li, J.Q.; Chen, C. Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network. *Comput.-Aided Civ. Infrastruct. Eng.* **2017**, *32*, 805–819. [CrossRef]

54. Zhang, A.; Wang, K.C.; Fei, Y.; Liu, Y.; Tao, S.; Chen, C.; Li, J.Q.; Li, B. Deep learning–based fully automated pavement crack detection on 3D asphalt surfaces with an improved CrackNet. *J. Comput. Civ. Eng.* **2018**, *32*, 04018041. [CrossRef]
55. Fei, Y.; Wang, K.C.; Zhang, A.; Chen, C.; Li, J.Q.; Liu, Y.; Yang, G.; Li, B. Pixel-level cracking detection on 3D asphalt pavement images through deep-learning-based CrackNet-V. *IEEE Trans. Intell. Transp. Syst.* **2019**, *21*, 273–284. [CrossRef]
56. Zhang, A.; Wang, K.C.; Fei, Y.; Liu, Y.; Chen, C.; Yang, G.; Li, J.Q.; Yang, E.; Qiu, S. Automated pixel-level pavement crack detection on 3D asphalt surfaces with a recurrent neural network. *Comput.-Aided Civ. Infrastruct. Eng.* **2019**, *34*, 213–229. [CrossRef]
57. Feng, H.; Li, W.; Luo, Z.; Chen, Y.; Fatholahi, S.N.; Cheng, M.; Wang, C.; Junior, J.M.; Li, J. GCN-based pavement crack detection using mobile LiDAR point clouds. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 11052–11061. [CrossRef]
58. Chen, T.-H.; Chang, T.S. RangeSeg: Range-aware real time segmentation of 3D LiDAR point clouds. *IEEE Trans. Intell. Veh.* **2021**, *7*, 93–101. [CrossRef]
59. Leng, Z.; Tan, M.; Liu, C.; Cubuk, E.D.; Shi, X.; Cheng, S.; Anguelov, D. Polyloss: A polynomial expansion perspective of classification loss functions. *arXiv* **2022**, arXiv:2204.12511.