IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 28 July 2023, accepted 12 August 2023, date of publication 18 August 2023, date of current version 24 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3306423

RESEARCH ARTICLE

Road Traffic Marking Extraction Algorithm Based on Fusion of Single Frame Image and Sparse Point Cloud

FEI YU AND ZHAOXIA LU
School of Sport Communication and Information Technology, Shandong Sport University, Jinan 250000, China
Corresponding author: Zhaoxia Lu (zhaoxia_lu@126.com)

This work was supported by the Shandong Social Science Planning and Research Project: Research on the Mechanism and Implementation Path of the Integrated Development of Artificial Intelligence and Shandong Sports Equipment Manufacturing Industry under Grant 21CTYJ24.

ABSTRACT As the boost of modern society, research on the extraction of road traffic markings has become increasingly popular. To improve the regional convolutional neural network, improve the road surface cloud segmentation algorithm based on the radius filtering algorithm and area division method, and combine the two algorithms to improve the sparse point cloud road traffic marking extraction algorithm. Finally, the study will integrate the single frame image and road surface cloud data frame by frame, apply the improved road traffic marking extraction algorithm of sparse point cloud to the road surface cloud with single frame image, and construct the road traffic marking extraction algorithm integrating single frame image and sparse point cloud. The effectiveness of the improved regional convolutional neural network algorithm proposed in the study was verified, and it was found that the average recall rate of the algorithm was 0.841, the average accuracy was 85.4%, and the operation speed was 125.6 seconds. Its performance was superior to other algorithms compared. In addition, the study also compared and analyzed the performance of the fusion road traffic marking extraction algorithm, and found that the average extraction edge length difference of the algorithm's road marking extraction was 0.0315m, and the average relative error between the algorithm and the internal verification points was 0.0493, which is better than the comparison algorithm. Based on the comprehensive experimental results, it was found that the performance of the proposed improved regional convolutional neural network algorithm and the traffic marking extraction algorithm that integrates single frame images and sparse point clouds is superior to the comparison algorithm. Meanwhile, the proposed fusion lane marker extraction algorithm has significantly improved the accuracy and precision compared to traditional lane marker extraction algorithms, and has enormous application potential in the field of road traffic.

INDEX TERMS Single frame image, LiDAR, mask R-CNN, point cloud extraction, road traffic markings.

I. INTRODUCTION

In recent years, with the continuous development of autonomous driving technology, the accurate extraction of road traffic markings has become more and more important. Road markings are not only an important basis for drivers to comply with traffic rules, but also a key information for autonomous vehicles to perform environmental perception and path planning [1], [2]. Therefore, improving the accuracy and robustness of the road traffic marking extraction algorithm is of great significance for the realization of a safe and efficient intelligent transportation system. At present, common road traffic marking extraction methods are mainly based on image processing technologies, such as Canny edge detection, Hough transform, etc. [3]. However, these methods have some problems when dealing with complex scenarios. For example, Canny edge detection is prone to false detection and missing detection when the lighting conditions

The associate editor coordinating the review of this manuscript and approving it for publication was Zhaoqing Pan.

change greatly. However, Hough transform, on the other hand, can lead to mistakes or missing edge detection and straight line detection results when vehicles are moving fast or road signs have complex shapes [4], [5]. In addition, the extraction of road traffic lines is completed by the collection of dense point cloud data by high-thread lidar. However, due to the expensive price of high-thread lidar and the huge amount of point cloud data, the production efficiency of high-precision maps is low and difficult, which cannot meet the market demand. Sparse point cloud extraction technology is a method to extract useful information in sparse point cloud based on point cloud data collected by sensors such as Lidar or depth camera through processing and analysis of point cloud [6]. At present, there are few researches on the fusion of sparse point cloud extraction technology and road extraction algorithm based on image processing technology, so as to fill this gap. This paper fuses single frame image with sparse point cloud extraction technology to construct sparse point cloud road traffic marking extraction algorithm based on single frame image. This algorithm can extract road traffic markings from a single frame image, improve the accuracy and robustness of the road marking extraction algorithm, and control the calculation time and cost at a certain level, which solves the problem of insufficient performance of the traditional road marking extraction algorithm. The theoretical contribution of this algorithm is that it can provide a reference for the research of road marking detection and extraction. At the same time, the algorithm also has a certain contribution in the industrial field, and the algorithm has a positive impact on the development of automatic driving, intelligent transportation system and other fields. This paper will introduce the relevant research on PC algorithms and road traffic extraction algorithms in the first section; In the second section, the construction method of a road traffic marking extraction algorithm that combines single frame images with sparse PCs is introduced; The third section introduces the analysis of the results of the improved algorithm for extracting road traffic markings; The fourth section is the conclusion.

II. RELATED WORKS

As the boost of image technology, the application range of PC technology is becoming wider and wider, and it has been applied in multiple fields. Huang et al. proposed a fast PC ground segmentation method in view of coarse to fine Markov random field to better segment the ground of autonomous vehicle using 3D laser radar, and tested the method in the field. The results show that this segmentation method possesses more excellent accuracy and stronger practical value than the comparison method [7]. The Zou team proposed a 3D PC registration algorithm in terms of local angle statistical histograms to address the 3D PC registration. By comparing this algorithm with other global algorithms, it was found that the presented algorithm has a wider application range and high robustness to noise [8]. To optimize the traditional surfactant separation and recovery system, Alibaba et al. proposed combining PC technology with cloud point reminder technology to construct an improved surfactant separation and recovery system. Then empirical experiments were conducted on the system, and it was found that the system can perform multi-stage separation and recovery of surfactants, effectively separate phenolic compounds, and has practical application value [1]. In response to the issue of environmental interference in the toxicity identification method of metal based nanoparticles, Zhou et al. proposed to integrate the PC extraction method with the cloud point extraction method to construct a species identification model for metal based nanoparticles. Then, the effectiveness of the model was verified, and it was found that the composition and species recognition performance of the metal based nanoparticles were stable under different conditions, and the robustness of the model was stronger than traditional methods [9].

With the rapid development of intelligent vehicles, the demand for images and videos of road conditions by drivers is increasing. Therefore, there are currently several methods used to extract road conditions. To improve the quality of road monitoring, the Heinz team proposed a mobile surveying and mapping system based on kinematics laser scanner. Then, the accuracy and precision of the system were tested, and the outcomes indicated that the physical height accuracy of the system was less than 10 mm, and its estimated standard deviation for cross slope was less than 0.07%. The above results indicate that the system has good road monitoring performance [10]. Ye et al. proposed an extraction method based on moving laser scanning PCs to improve the function of lane feature extraction methods for curved roads. Then, a performance comparison test was conducted on this extraction method, and it was found that its accuracy was 93.76%, which has better accuracy and robustness compared to other comparative algorithms [11]. The Wang team presented a lane detection method in view of semantic discrimination for improving the detection accuracy of lane detection, and incorporated a dual constraint random sample consistency method into this method for improving the accuracy of image extraction. Then empirical analysis was conducted on the detection method, and the results showed that the detection accuracy of the detection method was much higher than the comparison algorithm, and it had reliability in semantic recognition of lane marking [12]. Liu et al. proposed a layered collision avoidance strategy based on road extraction to address the problem of multi-state coupled motion, which makes it difficult to achieve the required accuracy and universal feasibility of different obstacle avoidance behaviors simultaneously in high-speed scenes. This strategy is based on road extraction methods and adopts different collision prevention measures based on different results. Then empirical analysis was conducted on the collision avoidance strategy proposed in the study, and the results showed that the collision avoidance strategy was adaptive to different collision avoidance behaviors and its practicality was better than other methods [13]. The Fang team proposed a transportation facility extraction method based on a mobile laser scanning system to improve the accuracy of transportation facility extraction. Then, when

the extraction method was applied in practice, it was found that the extraction accuracy of the method in the MLS dataset was 96.35%, indicating that the presented extraction method possesses excellent roadside detection performance [14].

The Mask Region-based Convolutional Neural Networks (Mask R-CNN) algorithm may have some difficulties in dealing with occlusion and complex background. When the object is obscured by other objects or background, the algorithm may not be able to segment the object's contour accurately. To achieve fruit and vegetable image segmentation in the supermarket environment, Hameed et al. proposed a fractional-based improved MaskRCNN mask edge improvement method. The empirical analysis of this algorithm shows that compared with traditional image segmentation algorithms, the segmentation accuracy of this algorithm has been greatly improved [15]. In order to accurately calculate the movement trajectory of maintenance personnel in surveillance video, improved MASK-RCNN was mainly used to detect and analyze the dynamics of maintenance personnel based on power grid surveillance video. Tong et al. built a video segmentation and centroid detection path drawing model based on improved MASK-RCNN, and verified the performance of the model. It is found that this model can accurately calculate the trajectory of people with information entering and leaving [16]. In view of the low efficiency and low accuracy of non-folding detection of photovoltaic power stations, Guo proposed an improved segmentation method for defective photovoltaic panels based on improved Mask R-CNN, and conducted an empirical analysis of this method, and found that the accuracy and recall rate of this method have been greatly improved, and it can be applied to industry [17].

The above research indicates that point cloud extraction method has been widely used in many fields, and there are various methods applied in road extraction, and there are many improvement studies on Mask RCNN. However, there is little research on the fusion of improved Mask RCNN algorithm with point cloud extraction algorithm and its application in road extraction. In order to fill this research gap, the Mask RCNN algorithm and the point cloud extraction algorithm were improved respectively, and the two algorithms were integrated and applied to road traffic marking extraction. The study hopes to improve road safety in this way.

III. CONSTRUCTION OF IMPROVED SPARSE POINT CLOUD ROAD TRAFFIC MARKING EXTRACTION ALGORITHM BASED ON IMAGES

Regional convolutional neural network (Mask Region-based Convolutional Neural Networks, Mask R-CNN) is a common method for dividing road traffic signs, but there are some problems of incomplete dividing road traffic lines and the target area is not obvious [18]. Therefore, the attention module is added to the Mask R-CNN neural network to realize the global sharing of the image and the pixel location relationship, so as to improve the accuracy of the road traffic line marking judgment and the integrity of the extraction line marking. According to whether there are obvious kerbs on both sides of the road, the improved fusion road surface cloud segmentation algorithm is constructed based on the radius filtering algorithm and the area division method, so as to realize the segmentation and extraction of these two road surface clouds. Finally, these two algorithms are combined to build an image-based algorithm for sparse point cloud road traffic marking extraction.

A. IMPROVED MASK R-CNN ALGORITHM BASED ON FUSED ATTENTION MODULE

Road traffic markings are common signs on the current roads in China, mainly divided into indicator markings, prohibition markings, and warning markings [19].The current research on the extraction of road traffic markings mainly focuses on the extraction, detection, and segmentation methods of linear road traffic markings, while ignoring the surface road traffic markings [16]. To separate linear and surface road traffic markings, an improved Mask R-CNN algorithm was constructed by combining the attention module with the Mask R-CNN algorithm. By improving the Mask R-CNN neural network, the attention module is added to the original network architecture, so that it can identify the linear and plane road traffic lines in the more accurate and complete image. As an instance segmentation network based on candidate regions, the overall framework of Mask R-CNN is shown in Figure 1.

Figure 1 illustrates that the overall architecture of Mask R-CNN is constructed based on Fast RCNN, Residual Network Feature Pyramid Network (ResNet FPN), and mask mask prediction branches. Mask R-CNN includes two stages, namely the region feature extraction stage and the processing stage. In the region feature extraction stage, Mask R-CNN uses ResNet as the backbone for extracting input image features, and constructs an FPN through a fully convolutional network. Then, the region of interest is selected through Region Proposal Networks (RPN) to generate candidate boxes for suspected target regions. In the processing stage, the improved algorithm utilizes the classification, bounding box regression, and segmentation mask Mask prediction branch functions of the mask R-CNN to achieve target detection and segmentation. In the average two-value cross-entropy loss entropy loss function of Mask R-CNN, the formula for calculating the classified branch loss is shown in equation (1).

$$L = L_{cls} + L_{reg} + L_{mask} \tag{1}$$

In equation (1), L_{cls} represents the classification loss in the RPN network; L_{reg} is the regularization loss of RPN bounding box regression; L_{mask} represents the mean of mask prediction mask sigmoid function. The formula for calculating the classification loss in the RPN network in the average two value cross entropy loss function is shown in equation (2).

$$L_{cls}(p_i, p_i^*) = -Ib\left[p_i, p_i^* + (1 - p_i)(1 - p_i^*)\right] \tag{2}$$

In equation (2), i represents the index value; p_i represents the true value of probability; p_i^* represents the probability prediction value. The regularization loss calculation formula

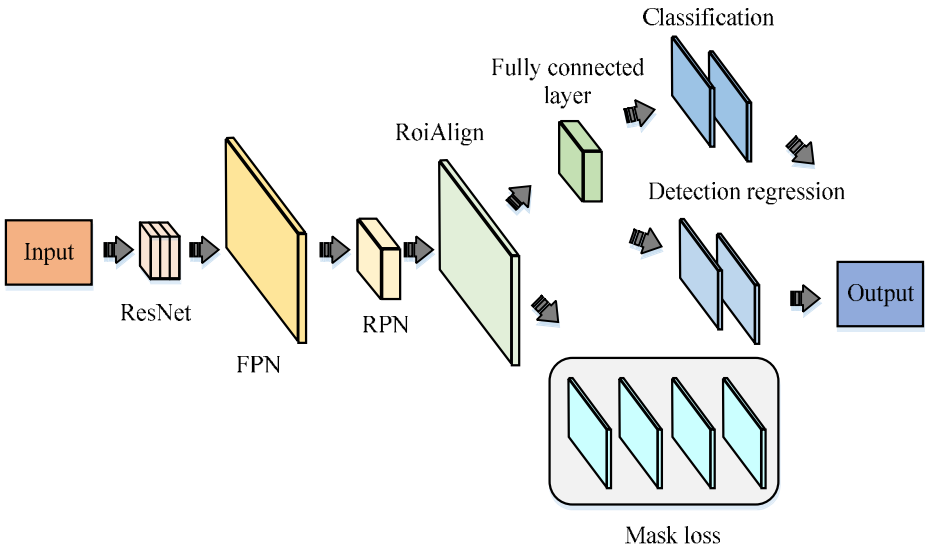


FIGURE 1. Mask R-CNN overall framework.

of RPN boundary box regression in the average two value cross entropy loss function is shown in equation (3).

$$L_{reg}(t_i, t_i^*) = smooth_{L_1}(t_i - t_i^*) \tag{3}$$

In equation (3), i represents the index value; t_i represents the predicted bounding box; t_i^* serves as the true value of the predicted bounding box; $smooth_{L_1}$ is the loss function; L_1 represents the regularization value of RPN bounding box regression calculation. The calculation formula of average two value cross entropy loss function for mask's sigmoid mean is shown in equation (4).

$$L_{mask}(p_i, p_i^*) = \left[p_i lb(p_i^*) + (1 - p_i) lb(1 - p_i^*)\right] \tag{4}$$

In equation (4), p_i represents the true value of the probability; p_i^* represents the probability predicted value; i represents the index value. For enhancing the recognition capability of Mask R-CNN for target information, a fusion attention module was constructed by integrating spatial attention modules and channel attention modules. The attention model added in the research refers to the Convolutional Block Attention Module (CBAM) model structure, which puts the spatial attention model in the front and superposes features several times to finally obtain the fusion features. The overall framework of the integrated attention module constructed in the research is shown in Figure 2.

Figure 2 reveals the fusion attention module. Figure 2 shows that the research fused the spatial attention module and the channel attention module to construct a fused attention model, which was placed in the region feature extraction stage of Mask-RCNN for enhancing the recognition function of target information. Among them, the spatial attention module can guide the Mask-RCNN network to focus on the surrounding environment areas except for the target, thus improving the overall feature extraction performance of the Mask-RCNN network for images. The channel feature module can combine all the channel features and achieve the expression performance of the Mask RCNN network by assigning channel weights. In the spatial attention module, after inputting the feature layer, the module inputs multi-channel feature maps into the pooling layer. The spatial attention module can transform multi-channel feature maps into single channel feature maps through global average pooling and maximum pooling. Then, the transformed feature map is overlaid with features and subjected to convolution operation, and the feature map is activated using the sigmoid function. Finally, the input original multi feature map is multiplied by the activated single feature map, and the output value is the weighted feature. In the channel attention module, the module aggregates the features of the multi-channel feature map, inputs the aggregated features into the shared multi-layer perceptron (MLP) for multi-channel parameter sharing, and then overlays the output features of each channel. Finally, the sigmoid function is used to activate the stacked features, and the activated data is the fused features.

B. FUSION ROAD SURFACE CLOUD SEGMENTATION ALGORITHM BASED ON RADIUS FILTERING ALGORITHM AND AREA DIVISION METHOD

At present, the main methods for extracting road lane marking information are obtained through 3D laser scanners or LiDARs, and the obtained PC data are processed [20]. At present, there are three main types of research on road PC data extraction, namely PCprojection-based extraction methods, clustering analysis-based extraction methods, and plane feature-based extraction methods [21]. In this study, the improved fusion road surface cloud segmentation algorithm is constructed based on the radius filtering algorithm and the

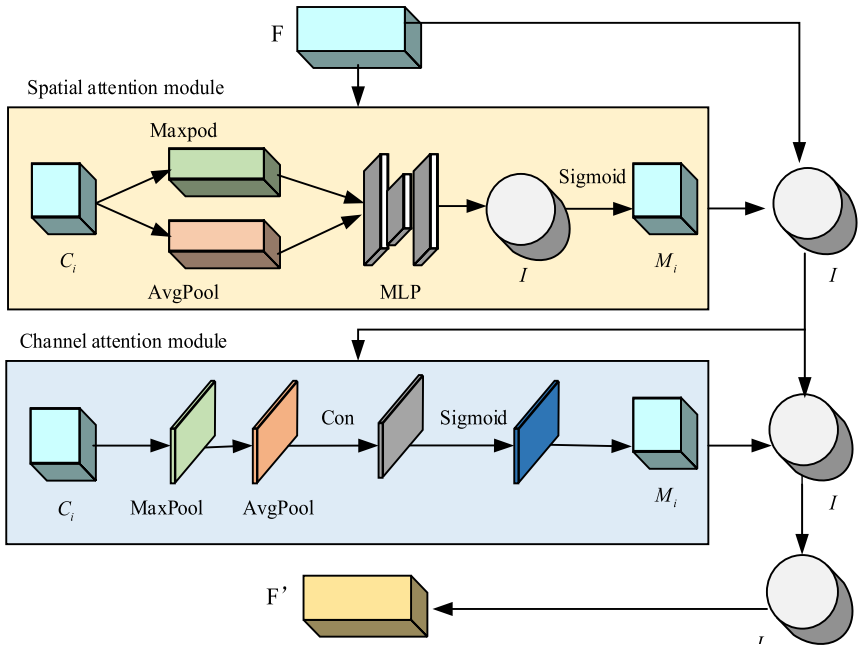


FIGURE 2. Integrated attention module overall architecture.

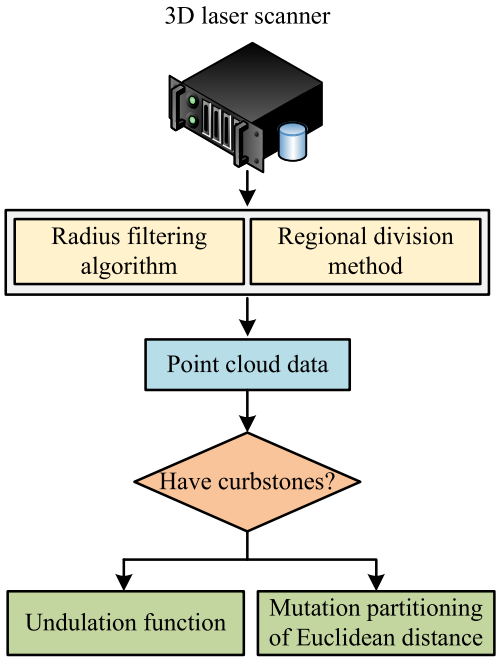


FIGURE 3. The construction steps of the road point cloud extraction algorithm.

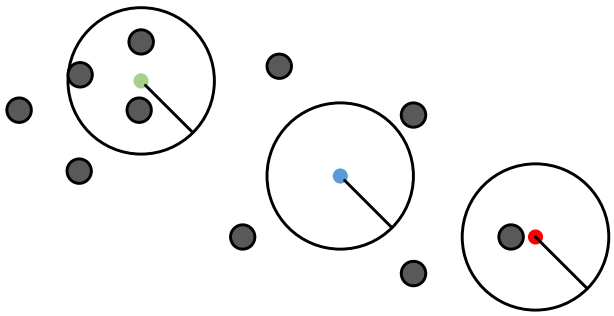


FIGURE 4. Principle of radius filtering algorithm screening.

regional segmentation method. The construction steps of the fusion road surface cloud segmentation algorithm are shown in Figure 3.

As shown in Figure 3, point cloud data is obtained by lidar is studied, but point cloud data obtained by lidar is prone to the problems of huge data and many interference data. In order to accurately extract pavement point clouds, the study first uses radius filtering method and area division method to remove noise points. Then, the study divided the road types into roads with curb and roads without curb. For the road with curb, the pavement fluctuation function is studied and designed, the part with the mutation in the relative height information is selected, and then the point cloud position corresponding to the curb is found to divide the road surface cloud. For the road without curb, the mutation block of Euclidean distance on the pavement point cloud data on the scanning line is used to find the beginning and end of the non-pavement points on the scanning line to divide the road surface cloud. The radius filtering algorithm is an algorithm that achieves image denoising by calculating the average distance between adjacent pixels in an image. The filtering principle of the radius filtering algorithm is shown in Figure 4.

Figure 3 shows that the radius filtering algorithm filters the target PC by setting a threshold in view of the Euclidean

distance between the PCs, and calculating the number of points within the circular range of the set radius. Firstly, define the PC dataset $P(p_1, p_2, p_3, \ldots p_i)$, with a threshold of k and a circle radius of r. Subsequently, take any point p ($p \in P$). Calculate the Euclidean distance from other points q_i ($q_i \in Q$) to point p in the PC dataset. Equation (5) depicts the calculation.

$$d = \sqrt{(x_{q_i} - x_p)2 + (y_{q_i} - y_p)2 + (z_{q_i} - z_p)} \tag{5}$$

In equation (5), x_{q_i} is the abscissa of point q_i; y_{q_i} is the ordinate of point q_i; z_{q_i} is the vertical coordinate of point q_i; x_p is the abscissa of point p; y_p is the ordinate of point p; z_p is the vertical coordinate of point p. Perform neighbor point filtering on point q_i, and the filtering formula is shown in equation (6).

$$\begin{cases} d < r, & q_i \text{ be a neighboring point} \\ d \geq r, & q_i \text{ not a neighboring point} \end{cases} \tag{6}$$

In equation (6), r serves as the radius of the circle; d serves as the Euclidean distance from point q_i ($q_i \in Q$) to point p. Filter and calculate all q_i in the dataset to obtain the number of q_i determined as neighboring points n. Perform noise filtering on n using the filtering formula shown in equation (7).

$$\begin{cases} n < k, & \text{noise point} \\ n \geq k, & \text{target point} \end{cases} \tag{7}$$

It repeats the steps of equations (6) and (7), performing neighbor point filtering and noise filtering judgment on all points p_i in PC dataset $P(p_1, p_2, p_3, \ldots p_i)$ until all noise is removed. The radius filtering algorithm can eliminate outliers and achieve simplification of cloud point data. In addition, there is also a problem of data interference in road extraction cloud point data. To quickly remove such interference from PC, a partition region filtering method is studied, which relies on the coordinate information of PC data to partition regions and process PC, achieving the filtering of PC data in the target area. After the pre-processing of PC data is completed, it is judged and identified whether there are standing curbs on both sides of the road. For roads with curbs, the PC is divided into equal areas, and the PC with obvious fluctuation is selected. Finally, the PC on the road surface is extracted from the elevation line chart. Firstly, the RANSAC algorithm is used to randomly select three points from the original PC to form a plane. If the PC data has the most points falling on this plane, the plane model is output. Subsequently, the coverage area of the PC is evenly separated into blocks, and the road surface relief function is designed using the number of PC in the blocks and the vertical distance between the plane model and the ground points. The formula for calculating the fluctuation function is shown in equation (8).

$$F = \sqrt{\frac{\sum_{i=1}^{n} (d_i - \overline{d})}{2m}} \tag{8}$$

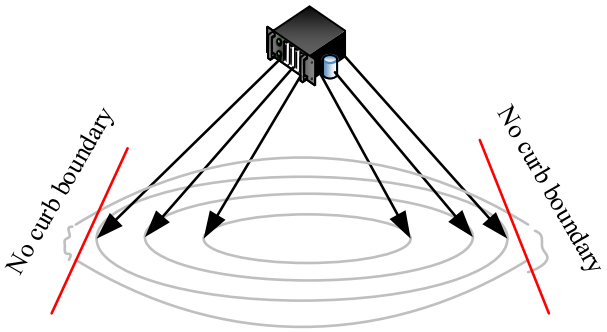


FIGURE 5. Schematic diagram of lidar road scanning line.

In equation (8), d_i represents the projection distance from the i-th point in the uniform block to the planar model; $\overline{d}$ is the average of the vertical projection distance from the midpoint of the uniform block to the plane model; m is the number of midpoints in a uniform block. It uses the fluctuation function to find the PC with large fluctuation, numbers the PC in the PC, and uses the PC number as the abscissa axis and the z coordinate value as the ordinate axis to construct a line graph. According to the line chart, it searches for the mutation location, and then filters, segments and splices through the abscissa to obtain a complete road surface PC. For roads without curbs, this study achieves road surface segmentation and extraction through the characteristics of multi linear LiDAR imaging. The schematic diagram of the LiDAR road scan line is indicated in Figure 5.

Figure 4 shows that based on the uniform distribution and concentric circles of PC data collected by LiDAR, the entire PC is divided into far and near areas and processed separately. It targets the near range area and evenly divides the area around the central axis to obtain n sector-shaped areas. It further divides the sector into m rings, calculates the Euclidean distance from the points on the scan line within the ring area to the central axis, and performsthe roadPC judgment on it. The judgment formula is shown in equation (9).

$$\begin{cases} \Delta d > w_1, & \text{not road point cloud} \\ \Delta d \leq w_1, & \text{be road point cloud} \end{cases} \tag{9}$$

In equation (9), Δd represents the Euclidean distance difference between adjacent points in the annular zone and the central axis; w_1 represents the judgment threshold for the annular zone. For remote areas, voxelization is performed on PCs within the remote area andan elimination judgment is performed. The judgment formula is shown in equation (10).

$$\begin{cases} amount > w_2, \text{eliminate} \\ amount \leq w_2, save \end{cases} \tag{10}$$

In equation (10), *amount* serves as the quantity of points in the remote area; w_2 represents the threshold for determining the quantity of points in the remote area. Finally, the PCs in the near distance and far distance areas are fused and spliced to obtain a complete road surface PC without curbs.

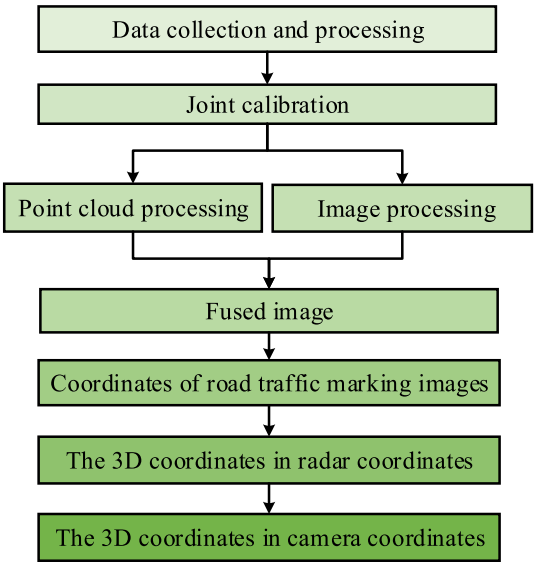


FIGURE 6. Schematic diagram of the algorithm for extracting road traffic markings.

C. CONSTRUCTION AND PRACTICAL APPLICATION OF ROAD TRAFFIC MARKING EXRACTION ALGORITHM FUSED WITH SINGLE FRAME IMAGE AND POINT CLOUD

After the construction of the improved Mask R-CNN algorithm based on the fusion attention module and the fusion road surface cloud segmentation algorithm based on the radius filter algorithm andthe area division method, the two algorithms are studied and integrated to construct the single frame image and point cloud integrating road marking extraction algorithm. Firstly, based on the joint calibration principle of lidar and camera, the single frame image and road surface cloud data are fused frame by frame, and extracts the three-dimensional lidar coordinates of traffic markings are extracted by using the fused image. Then research using improved Mask R-CNN algorithm and fusion road pastry cloud segmentation algorithm processing equipment in the experiment site to obtain image data and point cloud data, then extract the corresponding road pastry cloud through the matrix into the image coordinate system form fusion image, fusion image processing for road traffic marking image coordinates, finally through the road traffic marking in the laser radar coordinate system of three-dimensional coordinates. The schematic diagram of the proposed road traffic marking extraction algorithm is shown in Figure 6.

Figure 6 reveals that the first step is to collect images and cloud point data. A mobile measurement vehicle is used to collect PCs and image data near the south gate of Sichuan University (Jiang'an Campus), and the file is extracted and analyzed for PC data through the ROS system. Then, the obtained PC data and image data are pre-processed, and internal and external verification points are established for subsequent verification research. The second step is joint calibration, in which the joint calibration of the camera and LiDAR is investigated using Hesai's 40-line LiDAR Pandar and Flir monocular cameras. The third step is to use the improved Mask R-CNN proposed in the study to segment and extract road surface PC and images separately. Subsequently, the extracted PC data is transformed into an image coordinate system (CS) through a matrix to construct a fusion image, and the image coordinates of road traffic markings are obtained based on this. Finally, the three-dimensional coordinates of the road traffic markings in radar coordinates are obtained through matrix transformation. The commonly used methods for joint calibration are the checkerboard method and the feature point method. The joint calibration method used in the research is the checkerboard method, and its calibration process is shown in Figure 7.

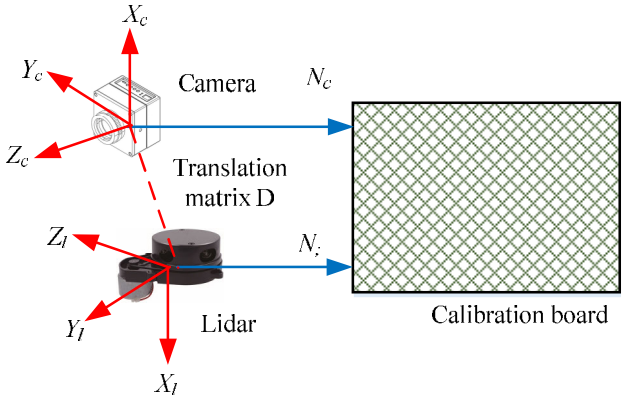


FIGURE 7. Checkerboard calibration process.

As shown in Figure 7, N_c and N_l are the calibration board normals of the origin of the LiDAR camera CS. The normal vector of N_c is obtained, and the calculation is indicated in equation (11).

$$N_c = R_3(R_3^T D) \tag{11}$$

In equation (11), R represents the rotation matrix after camera calibration; R_3 represents the third column vector of the rotation matrix; D represents the translation matrix (TM); T represents the matrix coefficient of the checkerboard calibration board. Subsequently, the normal vector of N_1 is obtained, and its solution formula is shown in equation (12).

$$N_l = [A, B, C]^T \sqrt{D} \tag{12}$$

In equation (12), (A, B, C) represents the plane normal vector that fits the calibration plate plane; D represents the TM; T represents the matrix coefficient of the checkerboard calibration board. Then, the rotation matrix R_{ltoc} and the translation vector D_{ltoc} from the laser radar to the camera CS are solved according to the obtained normal vector. The formula for calculating a random set of D_{ltoc} values is shown in equation (13).

$$\begin{cases} \|N_l\| - \|N_c\| = D_{ltoc} n_l \\ n_l = \dfrac{N_l}{\|N_l\|} \end{cases} \tag{13}$$

In equation (13), $\|N_l\|$ and $\|N_c\|$ represent the range from the origin of the LiDAR and camera CSs to the calibration board plane; n_l represents the unit vector of N_l. Subsequently, the N_c and N_l values of different groups were obtained by adjusting the pose of the LiDAR and camera, and the optimal D_{ltoc} value was solved using the least squares method. The expression of the unit vector n_l in camera coordinates is defined as $R_{ltoc}n_l$, and its calculation formula is shown in equation (14).

$$\begin{cases} R_{ltoc}n_l n_c = 1 \\ n_c = \dfrac{N_c}{\|N_c\|} \end{cases} \tag{14}$$

In equation (14), $\|N_c\|$ represents the range from the origin of the camera CS to the plane of the calibration board; N_c is the calibration board normal of the origin of the LiDAR CS; n_c represents the unit vector of N_c. Subsequently, the optimal R_{ltoc}-value is solved using the least squares method. Perform iterative optimization on the obtained optimal D_{ltoc} and R_{ltoc} values, with the iterative formula shown in equation (15).

$$F = \sum_{i=1}^{n}\sum_{j=1}^{m}\left[\frac{N_{c,j}(P_{ltoc}P_{i,j} + T_{ltoc})}{\|N_{c,i}\|} - \|N_{c,j}\|\right]^2 \tag{15}$$

In equation (15), n is the total number of times the lidar and camera adjust their pose; The number of points scanned and mapped by LiDAR on the calibration board during the i-th pose of m; $N_{c,i}$ represents the vertical vector from the origin of the camera CS to the calibration board; $P_{i,j}$ represents the j-th LiDAR scanning point in the i-th pose; $\|N_{c,j}\|$ represents the vector modulus of $N_{c,j}$. The obtained joint calibration parameters are projected onto the collected image to obtain a fused image, thereby establishing the link between the PC data. Then the road traffic markings in the fusion image are segmented, including color space conversion, ROI binarization, graphic perfection and target area segmentation. After completing the fusion image segmentation, the target 3D LiDAR coordinate information is obtained, and the pixel coordinates of the PC projection are obtained within the segmented area. Then it obtains the LiDAR three-dimensional CS of the road traffic markings through the transformation matrix between pixel and PCCSs.

IV. EMPIRICAL ANALYSIS OF ROAD TRAFFIC MARKING EXTRACTION ALGORITHM BASED ON FUSION OF SINGLE FRAME IMAGES AND SPARSE POINT CLOUDS

To verify the effectiveness of the improved Mask R-CNN algorithm, the road surface cloud segmentation algorithm, and the road traffic marking extraction algorithm with single frame image and sparse point cloud. The study conducted empirical experiments separately. The datasets used for the experiments were KITTI dataset, Tusimple dataset and SemanticKITTI dataset. Among them, KITTI data set is a relatively large autonomous driving evaluation data set, which contains real image data of urban, rural and expressway. The Tusimple dataset is the data of the autonomous driving race in the United States. It is collected by a pure vision platform and contains images of a variety of weather conditions, road conditions and time periods. SemanticKITTI The data set contains 150 frames of road frame point cloud data with curb and 150 frames of point cloud data without curb. The experimental environment consists of TensorFlow and Python deep learning framework platforms, as well as an Intel Core i7-7700K 16GB computer.

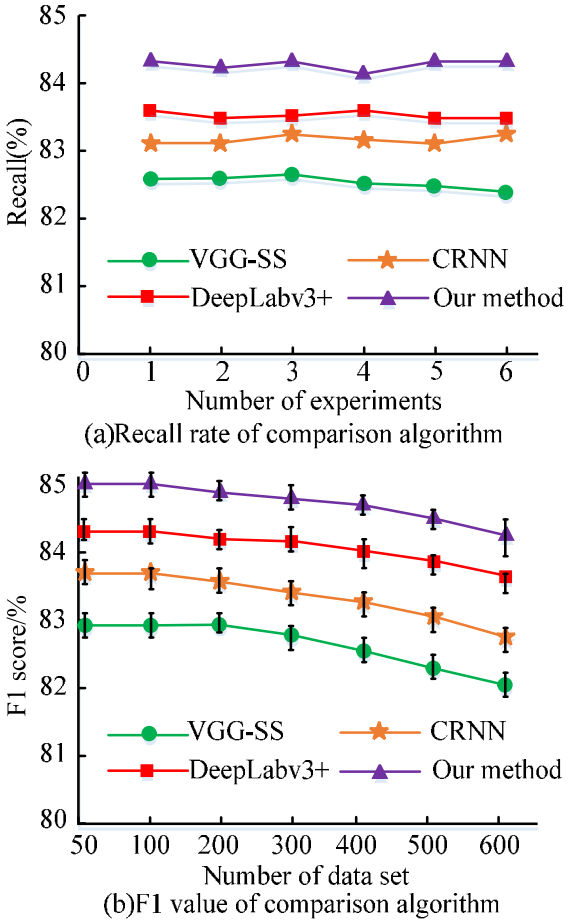


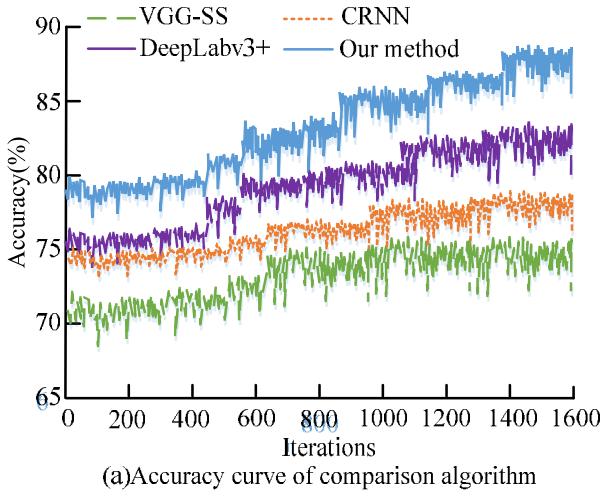
FIGURE 8. Recall rate and F1 score curve of various comparison algorithms.

A. PERFORMANCE COMPARISON EXPERIMENT OF IMPROVED MASK R-CNN ALGORITHM IN ROAD TRAFFIC MARKING EXTRACTION

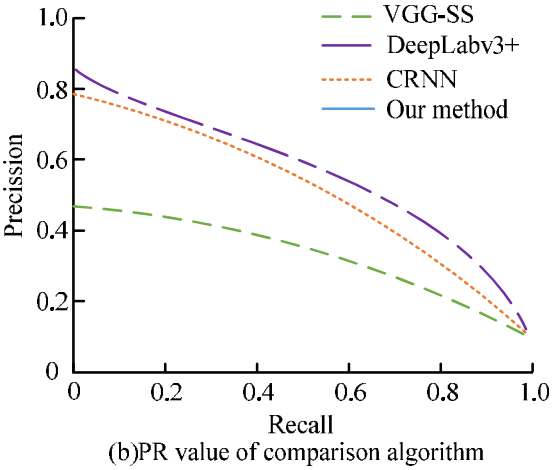
It conducted a performance comparison analysis on the Mask R-CNN algorithm proposed in the study, comparing the neural semantic segmentation network with Visual Geometry Group Sound Source (VGG-SS) algorithm, DeepLabv3+ semantic segmentation network algorithm, and Convolutional Recurrent Neural Network (CRNN) algorithm. The dataset used for this experiment is the Tusimple dataset.The study compares the recall rate, F1 value, accuracy, error, and runtime of each algorithm to perform performance analysis. The recall rate and F1 value of each algorithm are indicated in Figure 8.

TABLE 1. Edge length of road markings extracted from various algorithms and dense point clouds.

Serial number	Extracting edge length from dense point clouds(m)	Our method for extracting edge length(m)	K-means extract edge length(m)	FCN extract edge length(m)
1	0.252	0.258	0.356	0.256
2	0.255	0.128	0.326	0.346
3	0.232	0.215	0.156	0.231
4	0.253	0.238	0.236	0.275
5	0.245	0.232	0.345	0.287
6	0.231	0.31	0.269	0.269
7	0.243	0.201	0.348	0.312
8	0.237	0.234	0.369	0.299
9	0.239	0.212	0.156	0.157
10	0.207	0.213	0.158	0.285
11	0.210	0.263	0.321	0.314
12	0.138	0.135	0.167	0.207
13	0.484	0.421	0.421	0.387
14	0.432	0.423	0.398	0.348
15	0.391	0.345	0.269	0.369
16	0.352	0.351	0.354	0.314
17	0.341	0.312	0.421	0.403
18	0.341	0.391	0.401	0.388
19	0.235	0.254	0.254	0.212
20	0.271	0.249	0.174	0.214

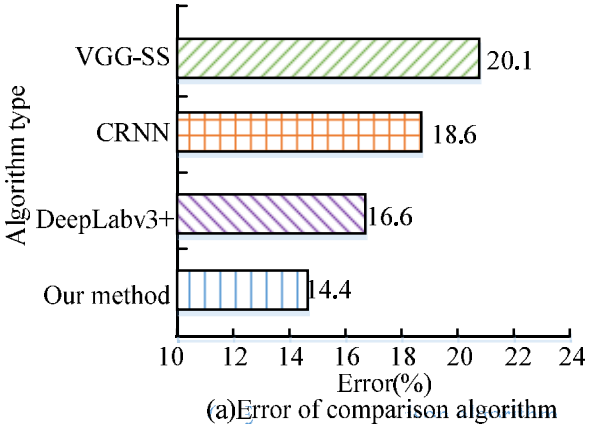


(a)Accuracy curve of comparison algorithm

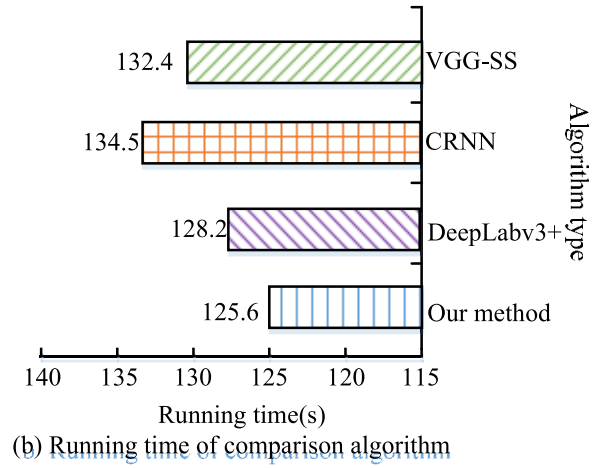


(b)PR value of comparison algorithm

FIGURE 9. Accuracy and PR curve of various comparison algorithms.



(a)Error of comparison algorithm



(b) Running time of comparison algorithm

FIGURE 10. Error and running time of various comparison algorithms.

Figure 8 reveals the recall rate and F1 results for each of the comparison algorithms. Figure 8 (a) reveals the recall rates of each comparison algorithm. Figure 8 (a) reveals that the Mask R-CNN algorithm proposed in the study has the highest overall recall curve, with an average recall rate of 0.841, which is 0.006 higher than DeepLabv3+. Figure 8 (b)

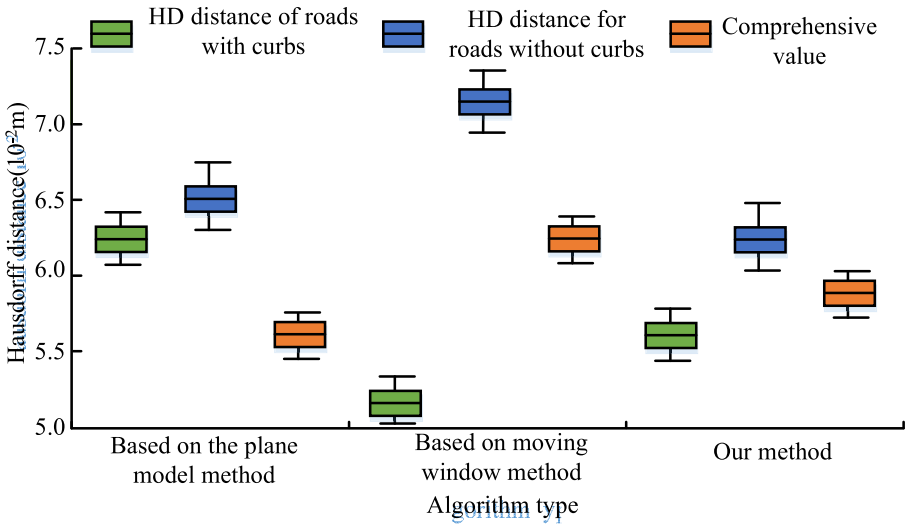


FIGURE 11. The Hausdorff Distance comparison results of various algorithms.

TABLE 2. Edge length of road markings extracted from various algorithms and dense point clouds.

Serial number	The extraction of edge length difference in our method	The extraction of edge length difference in K-means	The extraction of edge length difference in FCN
1	0.006	0.104	0.004
2	0.127	0.071	0.091
3	0.017	0.076	0.001
4	0.015	0.017	0.022
5	0.013	0.1	0.042
6	0.079	0.038	0.038
7	0.042	0.105	0.069
8	0.003	0.132	0.062
9	0.027	0.083	0.082
10	0.006	0.049	0.078
11	0.053	0.111	0.104
12	0.003	0.029	0.069
13	0.063	0.063	0.097
14	0.009	0.034	0.084
15	0.046	0.122	0.022
16	0.001	0.002	0.038
17	0.029	0.08	0.062
18	0.05	0.06	0.047
19	0.019	0.019	0.023
20	0.022	0.097	0.057

shows the F1 values ofeach of the comparison algorithms. Figure 8 (b) shows that the overall F1 value curve of the Mask R-CNN algorithm proposed in the study is the highest, with an average F1 value of 0.847, which is 0.005 higher than DeepLabv3+. The above results indicate that the Mask R-CNN algorithm proposed in the study has better recall performance and F1 value performance than other algorithms. The accuracy and PR curve of each comparison algorithm are shown in Figure 9.

Figure 9 shows the accuracy and PR curve of each comparison algorithm. 9 (a) is the accuracy curve of each comparison algorithm. Figure 9 (a) shows that the accuracy of each comparison algorithm increases with the quantity of iterations, among which the Mask R-CNN algorithm proposed in the study has the highest accuracy, with the highest accuracy of 89.3% and an average accuracy of 85.4%. Figure 9 (b) reveals the PR curves of each comparison algorithm. Figure 9 (b) reveals that the proposed Mask R-CNN algorithm has the largest offline area of the PR curve, with an offline area of 0.73. The above results reveal that the Mask R-CNN algorithm proposed in the study has better accuracy performance. Figure 10 shows the recognition error and running time of each comparative algorithm.

Figure 10 (a) shows the recognition error curves of each comparison algorithm. Figure 10 (a) reveals that the Mask R-CNN algorithm proposed in the study has the lowest error of 14.4%, which is 2.2% lower than DeepLabv3+, 4.2% lower than CRNN, and 5.7% lower than VGG-SS. Figure 10 (b) shows the operation time of each comparison

algorithm. Figure 10 (b) indicates that the Mask R-CNN algorithm proposed in the study has the shortest computation time and the fastest computation speed, at 125.6 seconds, which is 2.6 seconds lower in error compared to DeepLabv3+. Based on the above results, it reveals that the Mask R-CNN algorithm proposed in the study has better error performance and faster computation speed.

B. EMPIRICAL ANALYSIS OF IMPROVED GROUND EXTRACTION ALGORITHM

To understand the effectiveness of the ground PC extraction algorithm proposed in the study, an empirical analysis was conducted on it. The dataset used in the experiment is SemanticKITTI, which contains 150 road frame PC data with curbs and 150 PC data without curbs. In this study, the Hausdorff distance (HD), chamfer distance and extraction error between the PC extracted by the improved extraction algorithm and the original PC are used as evaluation indicators. Then, it is compared with the planar model fitting algorithm and the moving window method to explore the performance and effectiveness of the improved ground extraction algorithm. The HD comparison results of each algorithm are revealed in Figure 11.

Figure 11 shows the HD comparison results of each algorithm. Figure 11 shows that the improved ground extraction algorithm proposed in the study has a lower HD compared to other algorithms. In the PC data of roads with curbs, the HD is 0.056; In the PC data of roads without curbs,the HD is 0.063; A ground cloud point extraction algorithm with an HD of 0.059 is proposed by combining thePC data from roads with and without curbs. The above results indicate that the road surface extraction algorithm proposed in the study has higher extraction accuracy. On the road cloud point dataset with curbstones, the photos are blurred, and the chamfer distance and error performance of each algorithm are compared. The comparison results are illustrated in Figure 12.

Figure 12 (a) shows the extracted cloud point chamfer distances of each cloud point extraction algorithm. Figure 12 (a) shows that the ground extraction algorithm proposed in the study extracts cloud points closer to the original cloud point data, and its cloud point extraction is more accurate, with an average chamfer distance of 0.41 * 45 °. Figure 12 (b) shows the cloud point extraction errors of each cloud point extraction algorithm. Figure 12 (a) shows that the fluctuation amplitude of the error curve of the proposed ground extraction algorithm is smaller than other curves, and its error is more stable. The above results demonstrate that the ground extraction algorithm proposed in the study has better error performance compared to other comparative algorithms, and can better extract ground PC data, which has practical application value.

C. PERFORMANCE CONTRAST EXPERIMENT OF FUSION EXTRACTION OF ROAD TRAFFIC MARKINGS ALGORITHM

After completing the effectiveness verification experiments of the Mask R-CNN algorithm and ground PC extraction

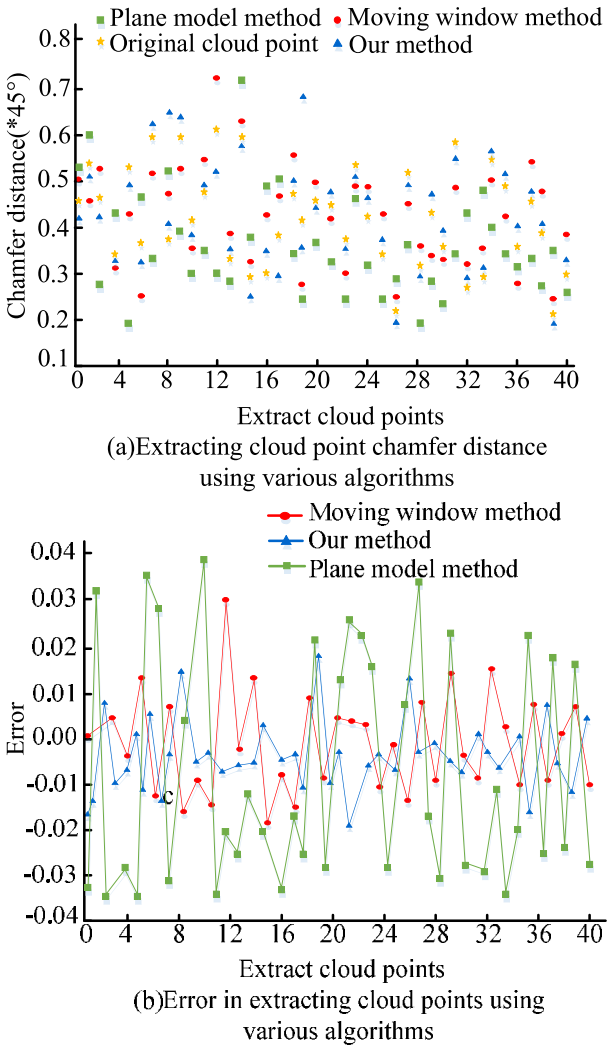


FIGURE 12. Comparison results of chamfering distance and error performance of various algorithms.

algorithm, the empirical analysis of the road traffic marking algorithm proposed in the study, which integrates single frame images and sparse PCs, was conducted.The data set used in the study is point cloud and image data collected near the south gate of Sichuan University (Jiang'an Campus). By comparing the PC data of 200 random internal verification points and 50 random external verification points, the pixel distance and edge length difference of the traffic markings extracted by the fusion algorithm are compared. The comparison results are revealed in Figure 13.

Figure 13 (a) shows the pixel distance distribution of the internal verification points of the fused road traffic marking extraction algorithm. Figure 13 (a) shows that the average pixel difference between the proposed fusion traffic marking extraction algorithm and the internal verification points is 0.394 pixels. Figure 13 (b) shows the comparison results of the edge length differencethe lane markings at the external verification points of the fusion lane marking extraction algorithm. Figure 13 (b) shows that the average edge length

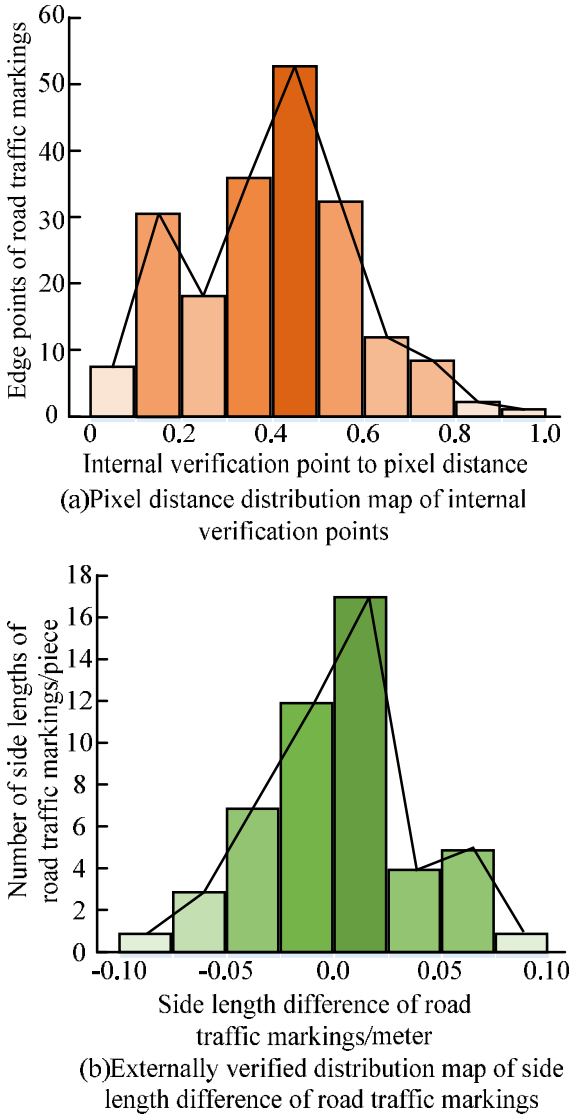


FIGURE 13. Pixel distance and edge length difference of traffic markings extracted by fusion algorithm.

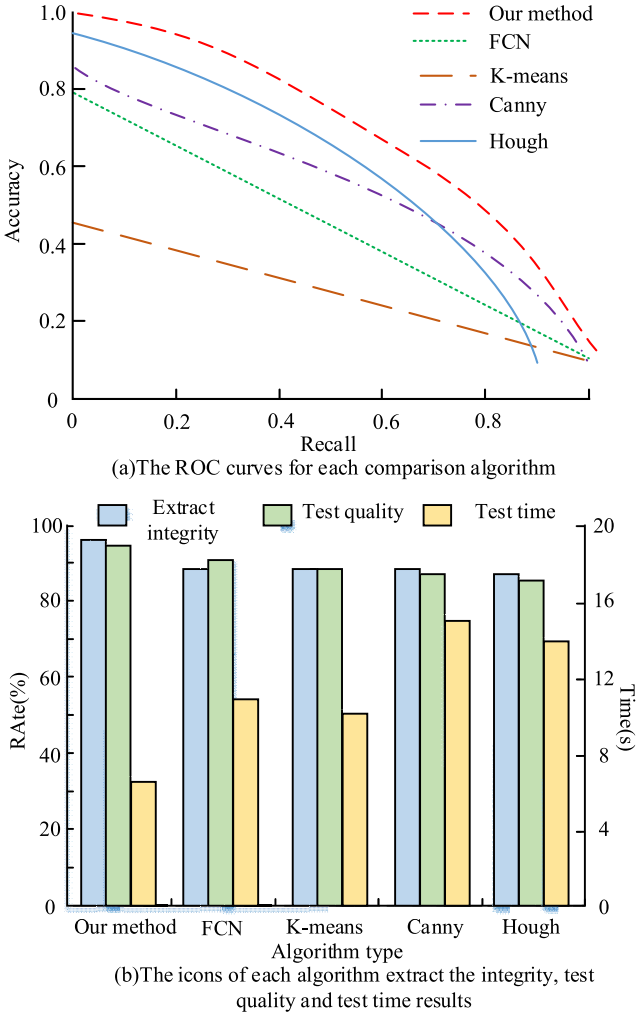


FIGURE 14. Comparison results of the operation speed and performance of each algorithm.

difference of the proposed fusion road traffic marking extraction algorithm is-0.0025m, and the average relative error of edge length is 0.0493. The above results show that the proposed fusion road traffic marking extraction algorithm can accurately and effectively extract road traffic markings, which has practical application value. To further validate the extraction performance of the proposed fusion road traffic marking extraction algorithm, a performance comparison analysis was conducted on the algorithm. The comparison algorithms were k-means clustering algorithm (K-means) and Full Convolutional Networks (FCN) road marking extraction algorithms. Performance comparison experiments were conducted with dense PC extracted road traffic marking edges and their differences. The edge lengths of road markings extracted by different algorithms and dense PCs are shown in Table 1.

Table 1 shows that the average side length of the road markings extracted from dense PCs is 0.27945m; The average extracted edge length of the fusion algorithm proposed in the study is 0.26925m; The average extracted edge length of the K-means extraction algorithm is 0.29495m; The average extraction edge length of the FCN extraction algorithm is 0.29365m. The above results demonstrate that the road traffic marking algorithm proposed in the study, which integrates single frame images and sparse PCs, extracts the average edge length more closely than the road marking algorithm extracted from dense PCs. The difference in edge length between each algorithm and dense PC extraction of road markings is shown in Table 2.

Table 2 shows that the average extracted edge length difference of the fusion algorithm proposed in the study is 0.0315m; The average extracted edge length difference of the K-means extraction algorithm is 0.0696m; The average extraction edge length difference of the FCN extraction algorithm is 0.0546. The above results demonstrate that the proposed road traffic marking algorithm that integrates single

frame images and sparse PCs has a smaller average edge length difference compared to other algorithms. It can extract the edges of road markings efficiently and accurately.In order to further explore the advantages of the fusion road line extraction algorithm, the research on the basis of FCN and K-means extraction algorithm, continue to study the algorithm and Canny edge detection [23] and Hough transformation [24] performance comparison, comparison for the ROC curve offline area, icon extraction integrity, detection quality and detection time. The comparison results of each algorithm are shown in Figure Figure 14.

Figure 14 (a) shows the ROC curve of each algorithm. From Figure 14 (A), the proposed algorithm has the largest ROC line area of 0.87, which is higher than other comparison algorithms and has the best performance. Figure 14 (b) shows the comparison results of the icon extraction integrity, icon detection quality and chart detection time of each algorithm. From Figure 14 (b), the proposed algorithm not only has the best icon extraction integrity and icon extraction quality, but also takes the shortest detection time of 5.7s. These results show that the proposed icon extraction algorithm is the best algorithm with the fastest calculation speed.

V. CONCLUSION

With the increasing demand for road marking extraction, traditional road marking extraction methods can no longer meet the needs of today's society. In response to the above issues, the study combines the improved Mask R-CNN algorithm with the improved cloud point extraction algorithm to construct a road traffic marking extraction algorithm that integrates single frame images and PCs. The performance analysis of the improved Mask R-CNN algorithm proposed in the study found that the average recall rate of the algorithm was 0.841, the average F1 value was 0.847, the average accuracy was 85.4%, the error was 14.4%, and the operation time was 125.6 seconds. This indicates that the algorithm performs better than other comparative algorithms. In addition, empirical analysis was conducted on the proposed road traffic marking extraction algorithm that integrates single frame images and PCs. The analysis found that the average pixel difference between the algorithm and the internal verification points was 0.394 pixels, the average extracted edge length of traffic markings was 0.26925m, the difference from the average edge length was -0.0025m, and the average relative error was 0.0493. Its performance is superior to the comparison algorithm. Based on the above results, it can be found that the research proposes an improved Mask R-CNN algorithm and a road traffic marking extraction algorithm that integrates single frame images and PCs, which is superior to comparative algorithms in terms of computational speed, accuracy, and error, and has practical application value. The subsequent research direction of the experiment is to conduct empirical analysis on the performance of algorithms in more complex signal environments. The subsequent research aims to verify the accuracy of the absolute coordinates of the fusion road traffic marking extraction algorithm.

REFERENCES

[1] A. Alibade, G. Batra, E. Bozinou, and S. Lalas, "Optimization of the extraction of antioxidants from winery wastes using cloud point extraction and a surfactant of natural origin (lecithin)," *Chem. Papers*, vol. 74, pp. 4517–4524, Dec. 2020.

[2] L. Ma, Y. Li, J. Li, W. Tan, Y. Yu, and M. A. Chapman, "Multi-scale point-wise convolutional neural networks for 3D object segmentation from LiDAR point clouds in large-scale environments," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 2, pp. 821–836, Feb. 2021.

[3] S. Chen, H. Ma, C. Jiang, B. Zhou, W. Xue, Z. Xiao, and Q. Li, "NDT-LOAM: A real-time LiDAR odometry and mapping with weighted NDT and LFA," *IEEE Sensors J.*, vol. 22, no. 4, pp. 3660–3671, Feb. 2022.

[4] Q. Shi, J. Wu, Z. Ni, X. Lv, F. Ye, Q. Hou, and X. Chen, "A blast furnace burden surface deeplearning detection system based on radar spectrum restructured by entropy weight," *IEEE Sensors J.*, vol. 21, no. 6, pp. 7928–7939, Mar. 2021.

[5] X. Gao and D. Hu, "MVDCANet: An end-to-end self-attention-based multiview-dualchannel 3D object detection," *IEEE Sensors J.*, vol. 21, no. 24, pp. 27789–27800, Dec. 2021.

[6] Y. Guo, Z. Mustafaoglu, and D. Koundal, "Spam detection using bidirectional transformers and machine learning classifier algorithms," *J. Comput. Cognit. Eng.*, vol. 2, pp. 5–9, Apr. 2022.

[7] W. Huang, H. Liang, L. Lin, Z. Wang, S. Wang, B. Yu, and R. Niu, "A fast point cloud ground segmentation approach based on coarse-to-fine Markov random field," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 7841–7854, Jul. 2022.

[8] X. Zou, H. He, Y. Wu, Y. Chen, and M. Xu, "Automatic 3D point cloud registration algorithm based on triangle similarity ratio consistency," *IET Image Process.*, vol. 14, no. 14, pp. 3314–3323, 2022.

[9] X. Zhou, W. Jiang, J. Wang, S. He, J. Li, and B. Yan, "Speciation analysis of Ag_2S and ZnS nanoparticles at the ng/L level in environmental waters by cloud point extraction coupled with LC-ICPMS," *Anal. Chem.*, vol. 92, no. 7, pp. 4765–4770, 2020.

[10] E. Heinz, C. Eling, L. Klingbeil, and H. Kuhlmann, "On the applicability of a scan-based mobile mapping system for monitoring the planarity and subsidence of road surfaces—Pilot study on the A44n motorway in Germany," *J. Appl. Geodesy*, vol. 14, no. 1, pp. 39–54, 2019.

[11] C. Ye, H. Zhao, L. Ma, H. Jiang, H. Li, R. Wang, M. A. Chapman, J. M. Junior, and J. Li, "Robust lane extraction from MLS point clouds towards HD maps especially in curve road," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 2, pp. 1505–1518, Feb. 2022.

[12] Y. Zhong, J. Zhang, Y. Li, T. Geng, and M. Wang, "Robust multi-lane detection method based on semantic discrimination," *IET Intell. Transp. Syst.*, vol. 14, no. 9, pp. 1142–1152, Sep. 2020.

[13] Z. Liu, J. Chen, F. Lan, and H. Xia, "Methodology of hierarchical collision avoidance for high-speed self-driving vehicle based on motion-decoupled extraction of scenarios," *IET Intell. Transp. Syst.*, vol. 14, no. 3, pp. 172–181, Mar. 2020.

[14] L. Fang, G. Shen, H. Luo, C. Chen, and Z. Zhao, "Automatic extraction of roadside traffic facilities from mobile laser scanning point clouds based on deep belief network," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 4, pp. 1964–1980, Apr. 2021.

[15] Z. U. Rehman, M. A. Khan, F. Ahmed, R. Damaševičius, S. R. Naqvi, W. Nisar, and K. Javed, "Recognizing apple leaf diseases using a novel parallel real-time processing framework based on MASK RCNN and transfer learning: An application for smart agriculture," *IET Image Process.*, vol. 15, no. 10, pp. 2157–2168, Mar. 2021, doi: 10.1049/ipr2.12183.

[16] T. Chen, Y. Jiang, W. Jian, L. Qiu, H. Liu, and Z. Xiao, "Maintenance personnel detection and analysis using mask-RCNN optimization on power grid monitoring video," *Neural Process. Lett.*, vol. 51, no. 2, pp. 1599–1610, 2019, doi: 10.1007/s11063-019-10159-w.

[17] S. Guo, Z. Wang, Y. Lou, X. Li, and H. Lin, "Detection method of photovoltaic panel defect based on improved mask R-CNN," *J. Internet Technol.*, vol. 23, no. 2, pp. 397–406, Mar. 2022.

[18] Y. Wu, H. Xu, X. Gong, and Y.-C. Yeo, "A ladder transmission line model for the extraction of ultralow specific contact resistivity—Part I: Theoretical design and simulation study," *IEEE Trans. Electron Devices*, vol. 67, no. 7, pp. 2682–2689, Jul. 2020.

[19] Z. Gao, G. Yang, E. Li, Z. Liang, and R. Guo, "Efficient parallel branch network with multi-scale feature fusion for real-time overhead power line segmentation," *IEEE Sensors J.*, vol. 21, no. 10, pp. 12220–12227, May 2021.

[20] H. Guan, H. E. Shizhong, L. I. Qiu, Z. Yang, C. Qin, and H. E. Wei, "A review of convolutional neural networks in equipment wear particle recognition," *Tribology*, vol. 42, no. 2, pp. 426–445, 2022.
[21] V. Manee, W. Zhu, and J. A. Romagnoli, "A deep learning image-based sensor for real-time crystal size distribution characterization," *Ind. Eng. Chem. Res.*, vol. 58, no. 51, pp. 23175–23186, Dec. 2019.
[22] R. Godse and S. Bhat, "Mathematical morphology-based feature-extraction technique for detection and classification of faults on power transmission line," *IEEE Access*, vol. 8, pp. 38459–38471, 2020, doi: 10.1109/ACCESS.2020.2975431.
[23] M. Kalbasi and H. Nikmehr, "Noise-robust, reconfigurable Canny edge detection and its hardware realization," *IEEE Access*, vol. 8, pp. 39934–39945, 2020, doi: 10.1109/ACCESS.2020.2976860.
[24] M. Marzougui, A. Alasiry, Y. Kortli, and J. Baili, "A lane tracking method based on progressive probabilistic Hough transform," *IEEE Access*, vol. 8, pp. 84893–84905, 2020, doi: 10.1109/ACCESS.2020.2991930.

FEI YU was born in Shandong, China, in 1980. She received the bachelor's degree in computer science and application from the Qilu University of Technology, in 2002, and the Master of Engineering degree in computer science and technology from Shandong University, in 2005.

Since 2005, she has been teaching with Shandong Sport University. Her main courses include discrete mathematics, operating systems, artificial intelligence, and python language foundation. She has participated in writing three textbooks and published multiple papers in domestic and foreign journals and conferences. Her research interests include image processing, pattern recognition, machine learning, and sports big data analysis.

ZHAOXIA LU was born in Shandong, China, in 1974. She received the Ph.D. degree in computer software and theory from Shandong University.

Currently, she is the Vice Dean of the School of Sports Media and Information Technology, Shandong Sport University, the Director of the Shandong Provincial Staff Physical Health Management Guidance Center, the Master Supervisor of Shandong Sport University, and the Director of the Shandong Sports Science Society. Her research interests include sports information processing and analysis, research on the mechanisms, methods, applications of information technology in promoting sports health, and intelligent sports information processing. There is a certain research foundation in the construction of a scientific fitness guidance network information platform and the analysis of physical health big data. She has presided over or participated in the completion of ten scientific research projects at or above the provincial and ministerial level, published 28 academic papers, of which five were included in SCI/EI. In recent three years, the main courses are programming language C, data structure, and computer professional English.