A review of vision-based road detection technology for unmanned vehicles

Liu, Chaoyang; Li, Xueyuan; Liu, Qi; Yang, Fan; Li, Zirui; Li, Mengkai

DOI
10.1109/IV55152.2023.10186761

Publication date
2023

Document Version
Final published version

Published in
IV 2023 - IEEE Intelligent Vehicles Symposium, Proceedings

A review of vision-based road detection technology for unmanned vehicles

1st Chaoyang Liu
School of Mechanical Engineering
Beijing Institute of Technology
Beijing, China
3220210263@bit.edu.cn

2nd Xueyuan Li*
School of Mechanical Engineering
Beijing Institute of Technology
Beijing, China
lixueyuan@bit.edu.cn

3rd Qi Liu*
School of Mechanical Engineering
Beijing Institute of Technology
Beijing, China
liuqibeishida@126.com

4th Fan Yang
School of Mechanical Engineering
Beijing Institute of Technology
Beijing, China
yangfanbitdb@163.com

5th Zirui Li
Faculty of Civil Engineering and Geosciences
Delft University of Technology
Delft, The Netherlands
z.li@bit.edu.cn

6th Mengkai Li
School of Mechanical Engineering
Beijing Institute of Technology
Beijing, China
3220200261@bit.edu.cn

2023 IEEE Intelligent Vehicles Symposium (IV) | 979-8-3503-4691-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/IV55152.2023.10186761

***Abstract*—With the development of unmanned vehicle technology, unmanned vehicles have played a huge role in logistics transportation, emergency rescue and disaster relief, etc., so the research on unmanned vehicles is becoming more and more important. Road detection is an important part of environmental perception and an important factor in the realization of assisted driving and unmanned driving technology. High-precision road detection technology can provide important environmental information for efficient planning and reasonable decision-making of unmanned vehicles. Firstly, the technical framework of road detection is given, and the road detection process is introduced in detail. Then, the vision-based road detection algorithm is introduced. Finally, some related data sets in the field of road detection are collected, which provides new ideas and methods for road detection researchers.**

***Index Terms*—road detection, vision, unmanned vehicle, data sets**

I. Introduction

Unmanned vehicle technology has developed rapidly in recent years. It plays an important role in logistics distribution, geographical exploration, troop operations and transportation, and assisted driving. Unmanned vehicle technology will also show broad application prospects in the future.

Unmanned vehicle detection [1] and decision-makingly [2] are developing rapid, and unmanned vehicle road detection is also essential. Road detection can not only provide path planning assistance for unmanned vehicles, but also provide auxiliary information for functions such as lane departure warning, lane keeping, lane change and adaptive cruise of unmanned vehicles.

This paper mainly reviews the lane line detection. In the actual scene of lane line detection, it faces various complex scenarios, such as road surface conditions: lanes have been subjected to rain and snow for a long time, impact and wear of vehicles and pedestrians, etc. The resulting damage and incompleteness, being blocked by vehicles and pedestrians, etc.; environmental factors: changes in lighting, the influence of artificial light sources such as car lights and street lights, shadows of plants, buildings, etc., false edges, fog, rain and snow and other severe weather Noise and occlusion, etc., these influencing factors are challenging problems for lane line detection.This paper also collects the better data sets of lane detection, and makes a detailed description and introduction of them.

The chapters of the article are arranged as follows: Section II introduces the technical framework of road detection and the process of road detection. Section III introduces some representative methods of various methods in recent years are shown, and the advantages and disadvantages of these methods are analyzed and compared. Section IV introduces the data sets related to road detection, and introduces the scene, quantity, and sensor type of the data sets in detail. Finally, the full text is summarized, and the future development direction and research hotspots of road detection are analyzed.

II. Road Detection Technical Framework

Researchers have proposed many methods for road detection, mainly based on vision, lidar and multi-sensor fusion. The basic framework of road detection is shown in Fig. 1. This paper only introduces the road detection method based on vision. Vision-based methods generally input image or video data, and then perform data preprocessing. Region of interest (ROI) extraction is required to obtain the road area, which can reduce the amount of calculation and the impact of other non-road areas on road detection. Some methods also require grayscale image or video data, and use grayscale values to assist in the extraction of road features. Next, operations such as denoising and enhancement need to be performed on the data to make the road features in the data more obvious. After preprocessing, various road detection methods are used for feature extraction and road detection. Vision-based road detection methods are mainly divided into: feature-based, model-based

*Corresponding author: Xueyuan Li and Qi Liu

and learning-based methods. Feature-based methods include color-based, edge-based, vanishing point-based, and feature fusion methods. Model-based methods include methods based on straight line models, curve models, and variable models. Learning-based methods include methods based on traditional machine learning and methods based on deep learning, and methods based on deep learning include methods based on segmentation, curve fitting and point detection. The final road result is output.

III. Road Detection Algorithm Overview

In this part, this paper will introduce and compare the road detection methods in Table I.

A. Methods based on feature extraction

In visual images or videos, roads have their own characteristics, such as color, texture, edge gradient changes, and vanishing points of parallel roads in 2D images. There have been many previous works on road detection based on road feature extraction. The feature-based method does not need to establish a road model, which is low in cost and strong in real-time. However, most of the feature extraction-based methods rely on manual feature extraction, which are less robust and are easily affected by factors such as shadows, occlusion, road wear, rain and snow. It is difficult to adapt to changing and complex traffic scenes.

a) Color-based methods: Lane lines or lanes of unstructured roads have rich color information and are clearly distinguished from the background. Therefore, color features can be used to extract lane information through color segmentation, clustering, and other methods. Some researchers use color-based methods for road detection. Jaesang Lee et al. [3] used RGB color space to segment unstructured roads through Gaussian mixture model. However, because the RGB color space is not intuitive and uneven in color separation, some colors are inconvenient to separate in the RGB space, and are greatly affected by illumination changes, resulting in poor road segmentation. For the problem of color separation in RGB color space, Miguelm Angel Sotelo et al. [4] use HSI color space to segment unstructured roads. This method can reduce the impact of light intensity, but the effect is poor in shadow environment. Xiaolin Li et al. [5] proposed an improved region growing method for unstructured road detection. The hue and saturation in HSV color space are used as criteria in different situations, and then the corresponding growth rules are designed to determine the seeds and determine whether the adjacent pixels belong to the seeds until all the pixels are added to the corresponding region. In order to increase the external constraints to alleviate the road detection under shadow conditions, José M. Álvarez et al. [6] used different components of HLS and LAB color space to detect the entire road surface.

b) Edge-based methods: The gradient change of pixels between the lane and the background is obviously different, so road detection can be performed by extracting edge features. There are mainly edge detection operators such as Canny [7], Roberts [8], Prwitt [9] and Sobel [10] to extract road edges or lane line edges . Yadi Li et al. [11] mainly use Canny edge detection to obtain lane lines. The algorithm alleviates the influence of uneven light to a certain extent. The methods of the edge are easily affected by the false edge caused by shadows, obstacles and other occlusions.

c) Methods based on vanishing points: Hui Kong et al. [12] used vanishing points for unstructured road detection, and improved the voting method when selecting vanishing points. This method first uses the Gabor filter to calculate the main texture direction on each pixel, and discards the hard voting method in the past. A new adaptive soft voting scheme based on variable-sized voting areas solves the shortcomings of the original hard voting method that has large errors in estimating the vanishing point, and this method also speeds up voting. In view of the fact that the texture features extracted by the above traditional methods are vulnerable to problems such as uneven illumination, image resolution, and long follow-up voting time, Yinbo Liu et al. [13] combined the detection of unstructured road vanishing points with convolutional neural networks and heat map regression. This method first overcomes the shortcoming of long prediction time caused by the traditional Resnet [14] module or HRNet [15] as the backbone network in the early heat map regression network, and improves the HRNet backbone feature extraction network. Then uses multi-scale heat map supervised learning to obtain accurate disappearance point, and finally use the strategy of coordinate regression to obtain high-precision coordinates of the vanishing point.

d) Methods based on feature fusion: Considering the limitation of only using a single feature for road detection, many researchers try to integrate different road features for road feature extraction to obtain better robust results. Laksono Kurnianggoro et al. [16] proposed a method based on feature fusion of mixed colors and edges for road detection. This method uses edge detection to detect the first edge of the lane after selecting the ROI, and finds many lane midpoints. Based on these midpoints, a model based on distance and color correlation clustering is applied. Finally, the marked lane line is obtained. The above methods are still greatly affected by lighting and shadows. Zuoquan Li et al. [17] proposed a method of fusing vanishing points and edge extraction to reduce the impact of occlusion, lighting, shadows, etc. on the edge of the lane line.

B. Model-Based Methods

Model-based methods for road detection generally include straight line and curve. Generally speaking, the model-based lane line detection method is not sensitive to image noise and lane line occlusion wear, but the difficulty lies in how to select and establish a model and determine related parameters. High-order curve models will bring a huge amount of calculation. A first-order curve model may not be able to represent road edges or lane lines well.

a) Line-based model: Roads in simple environments and straight roads are suitable for road detection with straight line

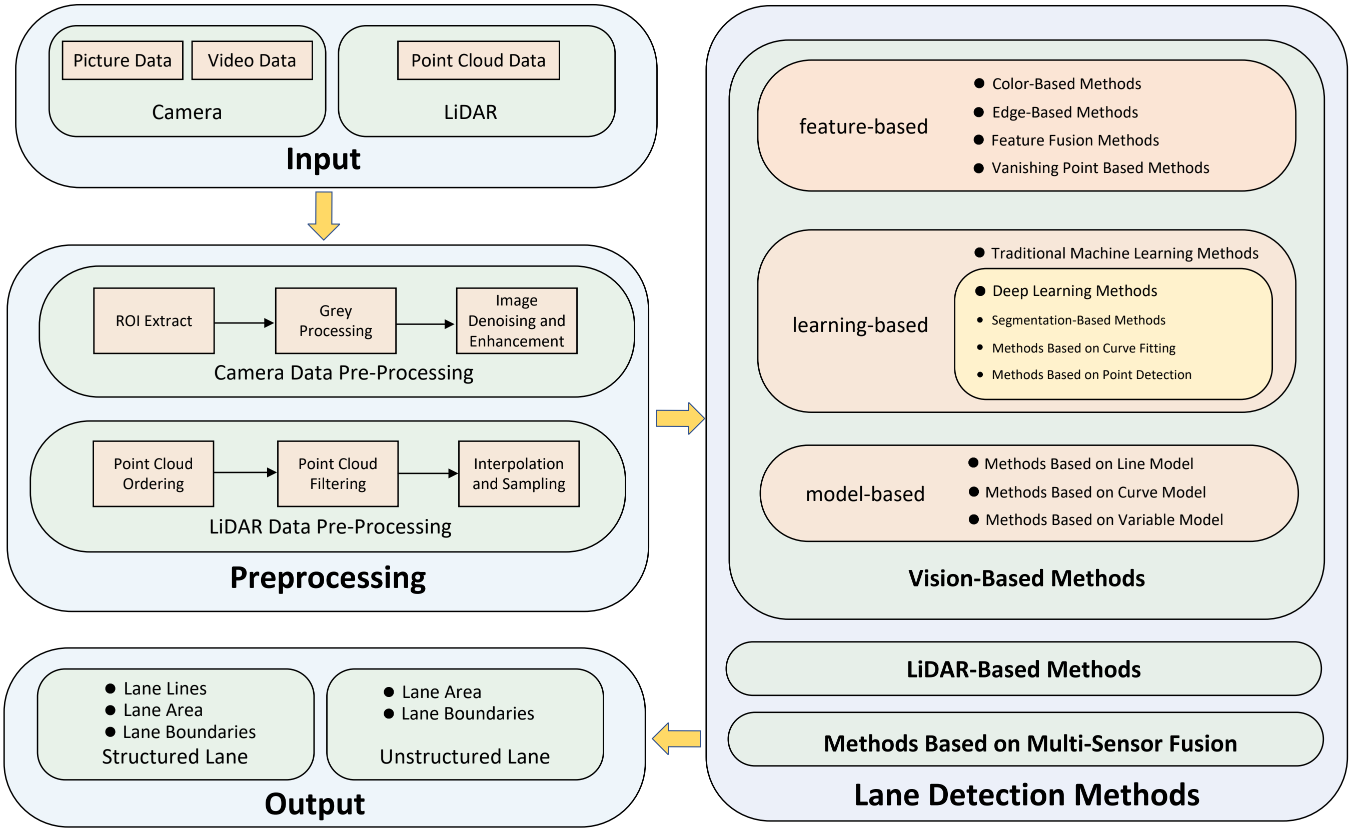

Fig. 1. Road detection flow: four steps of road detection and various methods of road detection

models . Youchun Xu et al. [18] used the Sobel operator to detect the edge, and directly used Hough [19] transform and straight line to fit the lane line. Xinxin Du et al. [20] also directly used Hough transform and straight line to directly fit the lane line. The linear model algorithm has low complexity and strong real-time performance, but its robustness is poor, and it cannot meet the requirements of most scenes such as curves.

b) Curve-based model: The curve model is suitable for most roads, and the curve parameters can be used to fit the road. Christian Roberto Kelber et al. [21] improved the method of parabolic fitting lane line. This method takes advantage of the parallel characteristics of the lane in the world coordinate system. In the camera perspective, the lane lines on both sides are regarded as parabolas, and the vanishing points of the near field and the far field to the road boundary are the same. According to this feature, the fitting parabola is constrained to the road boundary line. The cubic curve has a better fit than the quadratic curve. Mohamed Alyet al. [22] improved the RANSAC straight line fitting method after processing the image with Hough transform, and fitted the RANSAC three Bézier curve to obtain the lane line, but the calculation was more complicated and the efficiency was lower. Jitong Wang et al. [23] also proposed a better curve fitting method. This method is based on the DBSCAN (Density Based Spatial Clustering of Applications with Noise) clustering algorithm and the improved RANSAC method for lane line detection. The improved parabola-based RANSAC algorithm is used to fit the points of each category, which can be fitted as a straight line or curve. This method is more robust and better real-time. The curve model has better accuracy than the linear model, but the high-order curve model has low computational efficiency and is prone to overfitting.

c) Based on variable model: For the problem that the robustness of single-use straight line or curve fitting is not good, some researchers have proposed the method of using a variable model that combines the curve model and the straight line model, which has good robustness. Minchae Lee et al. [24] proposed a robust lane detection algorithm with cascade particle filter and decomposition system model. The cascade particle filter of this method decomposes the lane model into two sub-models: straight line model and curve model. Compared with the traditional particle filter-based lane detection system, it can improve the lane detection accuracy and reduce the calculation time.

C. Learning-based methods

Traditional methods based entirely on manual feature extraction are often subject to poor recognition accuracy caused by different scenes, and are susceptible to environmental and other factors, such as rain and snow, lighting conditions, road shape changes, etc. Learning-based methods have changed the previous method of purely manual feature extraction. It can be

TABLE I
COMPARISON OF ROAD DETECTION METHODS

Methods			Advantages	Limitations
Feature -Based	Color-Based	[3]–[6]	• Low complexity • Strong real-time performance • Insensitive to image size and direction	• Susceptible to light, shadow, occlusion, etc.
	Edge-Based	[11]	• Low complexity • Strong real-time performance • Not affected by color features	• Susceptible to image noise, shadows, other edge features
	Based on Vanishing Point	[12], [13]	• Contains global information, • Insensitive to occlusion and shadow	• The lane lines are required to be horizontal and have a constant width
	Based on Feature Fusion	[16], [17]	• Strong robustness • Wide application range • High precision	• High complexity • Poor real-time performance
Mode -Based	Linear Model	[18], [20]	• Low complexity • Strong real-time performance	• Poor robustness and difficult to adapt to complex roads
	Curve Model	[21]–[23]	• Suitable for many environments	• Calculation efficiency is low, prone to overfitting
	Variable Model	[24]	• Good applicability and robustness	• Low calculation efficiency • Poor real-time performance
Learning -Based	Machine Learning	[25], [26]	• Wide application range • High precision	• Easily affected by light, shadow, and occlusion
	Deep Learning	[27]–[36]	• Automatic extraction features • Good robustness • High precision	• Network structure design • Data integration cost • Real-time issues

automatically learned and extracted by the algorithm, which has better robustness in road detection.

a) Traditional machine learning methods: Machine learning provides researchers with a method that can learn and extract features by the machine itself. Shengyan Zhou et al. [25] proposed a road detection machine based on SVM and an effective self-supervised online learning method. Chunzhao Guo et al. [26] used Markov Random Field for road detection, but these machine learning methods still have some inherent shortcomings, some still need manual assistance to extract features, some have poor real-time performance, and are not robust in changing scenes.

b) Deep learning-based methods: The methods based on deep learning generally use convolutional neural networks or the recent use of transformer [37] in vision for road detection . The methods can be divided into three methods: segmentation-based, point-based detection, and curve-based methods. Compared with the method based on machine learning, the method based on deep learning automatically extracts features during the learning process without manual feature extraction. It has higher detection accuracy and robustness, and is suitable for road detection in most road conditions. How to effectively design the network structure, the high cost of network training, the need for a large number of high-quality data sets, the long training and inference time, and how to use the trained model for new scenarios that do not exist in the training set still have great challenges.

Deep learning methods based on segmentation: Segmentation-based deep learning is to perform pixel-level segmentation and extraction of lanes and lane lines through semantic segmentation or instance segmentation. It can directly perform end-to-end lane detection through the network. In the lane line detection methods based on segmentation-based deep learning, Xingang Pan et al. [27] aiming at the problem that the relationship and ability between pixels of the convolutional neural network in the process of image processing in the past has not been fully utilized, so the SCNN(Spatial Convolutional Neural Network) was proposed. The network structure is shown in the figure below. The main part of the network can be replaced by other segmentation networks. The core of the SCNN network lies in the structure in Fig. 2. For the feature layer extracted by the network, it is carried out in four directions. In addition to slicing and convolution, such a structure can make the use of spatial information more complete, and to a certain extent, it can deal with problems such as lane line occlusion. However, the slice convolution of the four directions of the SCNN network results in a large amount of network calculation and slows down the inference time. And in the final predicted lane line predicted by the SCNN network, this network can only predict up to four lane lines.

For the shortcomings of the binary lane segmentation of networks such as SCNN in the past, in order to realize instantiable segmentation of lanes, the LaneNet lane line detection network proposed by Davy Neven et al. [28] designed a multi-task network with branches. In the lane line instance segmentation stage, which consists of a lane segmentation branch and a lane embedding branch to achieve end-to-end

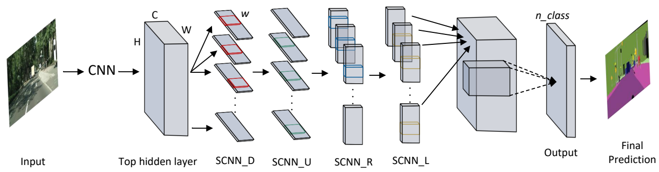

Fig. 2. SCNN network structure.The middle section in the figure is the core part, and the feature map is sliced from four directions: up, down, left, and right. This structure is conducive to identifying long-extended objects such as lane lines [27].

training. The lane segmentation branch outputs the lane line and the background binary map, and the lane embedding branch separates the segmented lane lines into different lane instances to determine which lane line pixels belong to which lane. In the subsequent curve fitting stage, in order to prevent the perspective transformation matrix from being affected by the road slope, a network H-Net is trained before the curve fitting to generate the perspective transformation matrix H to solve this problem.

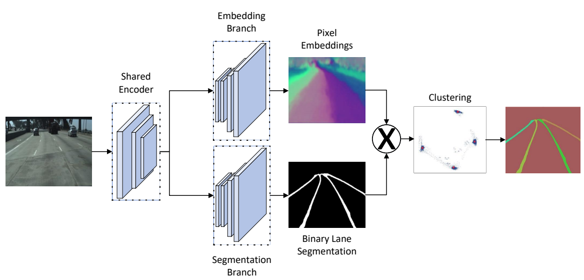

Fig. 3. LaneNet partial network structure. The LaneNet network contains two branches, binary segmentation and embedding. The former extracts the locations where all lane lines exist, and the latter learns a feature vector at each location [28].

In view of the fact that the lane lines in the previous research on the lane line method are often incomplete due to occlusion, shadow, strong light, lane line marking degradation, etc., Qin Zou et al. [29] proposed that the continuous characteristics of the lane lines on the road can be used to extract from the past continuous frames. Infer the position of the lane line when the lane line is incomplete in the current frame. The recurrent neural network has a good advantage in processing continuous frames. Therefore, according to the Conv-LSTM structure proposed by Xingjian Shi et al. [30] when using the long-term short-term memory network to predict rainfall. Combining the convolutional neural network with the recurrent neural network to predict the current occluded or missing lane lines through the past few frames has achieved good results. The overall structure of the network is shown in Fig. 4. This method has achieved good results in dealing with scenes such as occlusion and missing lane lines. But the shortcomings are also obvious. After adding Conv-LSTM, the network structure increases, the training and reasoning time increase, and the requirements for the data set are relatively high. The data set must give the real label of the current frame after a certain number of consecutive frames. The applicability of the data set is not good.

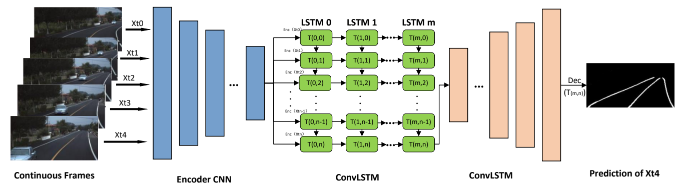

Fig. 4. Schematic diagram of Conv-LSTM network structure. This structure inserts the Conv-LSTM structure in the middle of the segmentation network, enabling the network to predict the occluded lane line in the fifth frame using the consecutive first four frames of images [29].

Inspired by the recent application of Transformer in the field of computer vision, Lang Peng et al. [31] proposed a BEV semantic segmentation method based on transformer for lane detection, named BEVSegFormer.This method uses a shared backbone to encode the image features of any camera. The method also introduces the BEV transformer decoding module to analyze the results of BEV semantic segmentation. the query is reshaped and upsampled to obtain the final lane segmentation results.

Deep learning methods based on curve fitting: The deep learning method based on curve fitting can directly obtain the curve parameters to fit the lane line after extracting the lane line information through the convolutional neural network. The lane line curve fitting obtained by deep learning can reduce the influence of occlusion, loss, shadow and other factors to a certain extent. Lucas Tabelini et al. [32] proposed a convolutional neural network based on PolyLaneNet, as shown in Fig. 5 The network can directly learn the curve parameters of the lane line, and the lane line obtained by the segmentation network needs to be followed. The two-step operation of curve fitting is changed into one step, which reduces the processing cost. The network structure is shown below, including a backbone feature extraction network and a fully connected layer. Through ablation experiments and comparison with SCNN and other networks, this method has certain advantages.

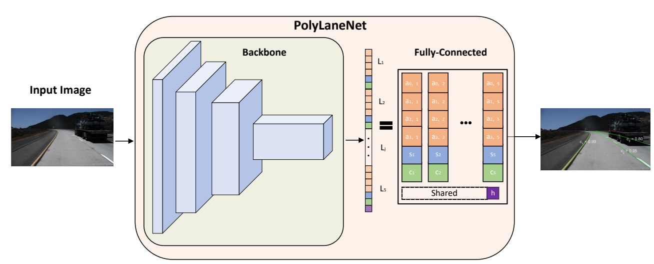

Fig. 5. Schematic diagram of PloyLaneNet network structure. This network takes an input image from a forward-looking camera and outputs a polynomial representing each lane marking in the image, along with a domain lane polynomial and a confidence score for each lane, resulting in lane estimates without post-processing. [32].

Zhengyang Feng et al. [33] also proposed a method for directly fitting lane lines based on convolutional neural networks.

Since the cubic Bézier curve has enough degrees of freedom to parameterize the deformation of lane lines in driving scenes, the computational complexity is low and stable. Therefore, the convolutional neural network is used to obtain the parameters of the cubic Bézier curve to fit the lane line. This method takes advantage of the fact that the lane line is generally symmetrical with respect to the camera, and proposes feature flip fusion. The graph is aggregated with its horizontally inverted version, reinforcing this symmetrical property. This method uses deformable convolution [38] to design feature flip fusion, which is used to align the asymmetry caused by camera rotation, lane change, asymmetry, etc. The overall network structure is shown in the Fig. 6.

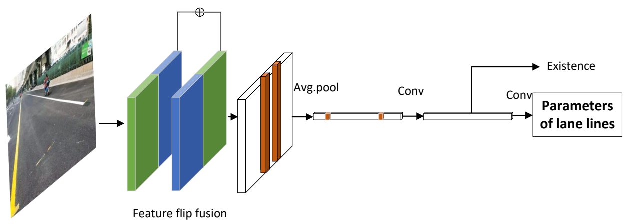

Fig. 6. The detailed structure of feature flip fusion, the feature map before flipping is fused with the reversed feature map after 1x1 convolution, and then 3x3 convolution is performed and then fused by deformable convolution [33].

Deep learning methods based on point detection: The lane line detection method based on point detection extracts the lane line by clustering the feature points on the lane line through the deep learning network. Dongkwon Jin et al. [34] proposed an anchor-based lane line detection method based on the anchor-based target detection method. In order to solve the problem of false detection and false detection caused by the complex structure of the lane, the concept of feature lane is proposed to describe the lanes of different structures. In order to obtain the feature lane, the singular value decomposition is used to obtain the optimal rank of the lane matrix of all lanes contained in the training set. Then, a set of lane candidate sets are generated by clustering training lanes in feature lanes. Finally, the author develops an anchor-based detection network SIIC-Net to determine the most lane set.

The above methods all perform lane line detection at the 2D level. In order to solve the problem of inaccurate road local estimation caused by uphill, downhill, and road bumps in the 2D lane line detection scene, Li Chen et al. [35] proposed a 3D Lane line detection method, in which PersFormer (Perspective Transformer), an end-to-end monocular 3D lane line detector, is proposed, which includes a transformer-based spatial feature conversion module. This method combines the spatial features of the front view and BEV map transformation is modeled as a learning process, which is more robust than inverse perspective transformation. Fan Yan et al. [36] also proposed a no-external-parameter, no-anchor-point 3D lane line detection method,named SALAD. This method returns lane 3D coordinates in the image without converting the feature map to BEV, which can solve the problem of 3D road structure perception error caused by the continuous change of camera external parameters due to road roughness. This method can directly perform three-dimensional lane detection on monocular images without the need to manually make anchor points and the supervision of external parameters. The two branches of semantic perception branch and spatial context branch decode the features to obtain the spatial information and segmentation mask of the lane line, and then integrate these information for 3D reconstruction, and finally obtain the 3D lane line position of the real scene.

IV. Road Detection Dataset

This review collects twelve excellent road detection datasets, as shown in Table II.

The Caltech Lanes dataset is four segments taken on the streets of Pasadena, California, containing segments from different city streets. A total of 1224 frame images were marked including 4172 lane boundaries.

The KITTI dataset was jointly established by the Karlsruhe Institute of Technology in Germany and the Toyota American Institute of Technology and collected using a variety of sensors. A total of more than 200,000 images of labeled objects are composed. The Road part is the lane line detection dataset.

The VPGNet dataset is a dataset proposed by the Korea Robotics and Computer Laboratory in the ICCV paper published in 2017 to detect lane lines through vanishing points in VPGNet [39]. The image resolution is 1288X728.

The CUlanes dataset is a large-scale challenging dataset for lane line detection research. It is collected by cameras installed on six vehicles, collecting a total of 133,235 frames, and dividing the dataset into 88,880 pictures as a training set. , a validation set of 9,675 images and a test set of 34,680 images, including eight difficult-to-detect conditions such as crowding, night, wireless, and shadow.

Tusimple is a company engaged in autonomous driving. In June 2018, Tusimple held a part of the public dataset in a competition for lane line detection using camera image data. The dataset is collected under good and moderate weather conditions and different traffic conditions during the day, which contains highway scenes with 2 lanes, 3 lanes and 4 lanes up to 5 lanes. The training set contains 3626 frames, and the test set contains 2782 frames.

The BDD100K dataset is a public driving data set released by the University of Berkeley AI Laboratory (BAIR). There are a total of 100,000 pictures of 1280x720 size and are labeled. The lane detection part is a training set of 70,000, a test set of 20,000, and a verification set. 10000. And it also distinguishes between solid lines, dashed lines, double lines and single lines;

The Apolloscapes dataset is collected under different traffic conditions in different cities by equipping a medium-sized SUV with a high-resolution camera and a Riegl acquisition system. High-quality pixel-level annotations of more than 110,000 frames are included.

The CurveLanes dataset is a public data set in the open source curve lane detection work of Huawei Noah's Ark

TABLE II
ROAD DETECTION DATA SETS

Data Sets	Data Format	Sensor	Scenes	Marking Methods
Caltech-Lanes	Image	Camera	Different city streets with/without curves and shadows	Spline curve
KITTI-ROAD	Image Video Point Cloud	Camera, 3D laser scanner GPS/IMU	Including various traffic scenarios such as rural areas, expressways, and city centers	Semantic segmentation
VPGNet	Image	Camera	Severe weather such as no rain, rainfall, heavy rainfall, and low lighting at night	Grid unit
CULane	Image	Camera	Beijing's cities, villages, and highways	Semantic segmentation
Tusimple	Image	Camera	mostly highway	Broken line
BDD100K	Video Image	Camera	The four regions of the United States are sunny, cloudy, rainy, day and night	Instance segmentation
Apolloscapes	Image Video Point Cloud	Camera Lidar	Several cities in China, diverse weather	Semantic segmentation
Curve Lanes	Image	Camera	Most of them are curved lanes in the city	Instance segmentation
RUGD	Image	Camera, Lidar GPS/IMU	Contains no identifiable geometric edges and vanishing points, and semantic boundaries are highly irregular	Semantic segmentation
VIL100	Image Video	Camera	10 scenes such as normal, crowded, occlusion, haze, dark, strong reflection	Instance segmentation
Open Lane	Image Video	Camera	14 lane categories, diverse scenarios	Instance segmentation
ONCE-3DLanes	Image Point Cloud	Camera Lidar	Various weather, lighting conditions, various geographical locations	Instance segmentation

Laboratory and Sun Yat-sen University. Each image has more curved lanes. The dataset consists of 150,000 images with 680,000 labels.

RUGD dataset, highly irregular terrain with semantic boundaries, mainly unstructured terrain passable areas. The dataset contains pixel-level labels for 24 different visual categories with a resolution of 1376x1110. The complete large-scale dataset exceeds 37,000 images.

The VIL100 dataset is a relatively new video instance lane line detection dataset. Its purpose is to use continuous video frames to solve problems such as lane line occlusion and damage that cannot be solved by a single frame. The data set contains 100 videos with 10,000 frames of images.

OpenLane is a dataset proposed in the 3D lane line detection method PersFormer. OpenLane contains 200,000 frames, more than 880,000 instance-level lanes, and 14 lane categories.

The ONCE-3DLanes dataset is a dataset proposed by Fudan University and Huawei Noah's Ark Laboratory. The data set contains various scenes. The data set is used to solve 3D data. dataset is from 211,000 2D lane markings of road scenes, automatically generate high-quality 3D lane marking positions.

CONCLUSION AND FUTURE OUTLOOK

In this paper, we make two main contributions to the field of road detection. First of all, this article conducts a comprehensive review of road detection methods, open source platforms, and data sets in recent years, which can help readers have a broad and comprehensive understanding of the current lane detection field. Secondly, we have made a more detailed interpretation of the network structure of some of these methods, which can help researchers analyze and learn from it in their own network structure.

From the interpretation and analysis of the full text, it can be seen that lane line detection has many difficulties that are currently being studied and will continue to be studied in the future. There are mainly three points.

A. Road detection in harsh environments

The lane line or the boundary line of the drivable area of the road is easily blocked by other vehicles, shadows, etc., and even they are damaged, which can easily lead to false detection or missed detection.

B. Design of Multiple Types of Road Linearity Detectors

There are many types of actual road lines, and different types of lane lines have different constraints on the current lane driving behavior, such as solid and virtual, yellow and white, single and double lane lines.

C. Road detection under high-speed operating conditions of unmanned vehicle

When the camera is running at high speed in the unmanned vehicle, the motion blur of the image will be generated. The accuracy of the laser radar will be lower when the unmanned vehicle is running at high speed. At high speed, the real-time requirements of the road detection algorithm are high. How to efficiently detect the lane.

REFERENCES

[1] Q. Liu, Z. Li, S. Yuan, Y. Zhu, and X. Li, "Review on vehicle detection technology for unmanned ground vehicles," *Sensors*, vol. 21, no. 4, p. 1354, 2021.

[2] Q. Liu, X. Li, S. Yuan, and Z. Li, "Decision-making technology for autonomous vehicles: Learning-based methods, applications and future outlook," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pp. 30–37, IEEE, 2021.

[3] J. Lee and C. D. Crane, "Road following in an unstructured desert environment based on the em (expectation-maximization) algorithm," in *2006 SICE-ICASE International Joint Conference*, pp. 2969–2974, IEEE, 2006.

[4] M. A. Sotelo, F. J. Rodriguez, L. Magdalena, L. M. Bergasa, and L. Boquete, "A color vision-based lane tracking system for autonomous driving on unmarked roads," *Autonomous Robots*, vol. 16, no. 1, pp. 95–116, 2004.

[5] L. Xiaolin, J. Yufeng, G. Yan, F. Xiaoxue, and L. Weixing, "Unstructured road detection based on region growing," in *2018 Chinese Control And Decision Conference (CCDC)*, pp. 3451–3456, IEEE, 2018.

[6] J. M. Á. Alvarez and A. M. Lopez, "Road detection based on illuminant invariance," *IEEE transactions on intelligent transportation systems*, vol. 12, no. 1, pp. 184–193, 2010.

[7] J. Canny, "A computational approach to edge detection," *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 679–698, 1986.

[8] R. Ouda and P. Hart, "Pattern classification and scene analysiswiley," *New York*, 1973.

[9] J. M. Prewitt *et al.*, "Object enhancement and extraction," *Picture processing and Psychopictorics*, vol. 10, no. 1, pp. 15–19, 1970.

[10] W. Gao, X. Zhang, L. Yang, and H. Liu, "An improved sobel edge detection," in *2010 3rd International conference on computer science and information technology*, vol. 5, pp. 67–71, IEEE, 2010.

[11] Y. Li, L. Chen, H. Huang, X. Li, W. Xu, L. Zheng, and J. Huang, "Nighttime lane markings recognition based on canny detection and hough transform," in *2016 IEEE International Conference on Real-time Computing and Robotics (RCAR)*, pp. 411–415, IEEE, 2016.

[12] H. Kong, J.-Y. Audibert, and J. Ponce, "Vanishing point detection for road detection," in *2009 ieee conference on computer vision and pattern recognition*, pp. 96–103, IEEE, 2009.

[13] Y.-B. Liu, M. Zeng, and Q.-H. Meng, "Unstructured road vanishing point detection using convolutional neural networks and heatmap regression," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–8, 2020.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[15] K. Sun, B. Xiao, D. Liu, and J. Wang, "Deep high-resolution representation learning for human pose estimation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5693–5703, 2019.

[16] D. Cáceres Hernández, L. Kurnianggoro, A. Filonenko, and K. H. Jo, "Real-time lane region detection using a combination of geometrical and image features," *Sensors*, vol. 16, no. 11, p. 1935, 2016.

[17] Z.-Q. Li, H.-M. Ma, and Z.-Y. Liu, "Road lane detection with gabor filters," in *2016 International conference on information system and artificial intelligence (ISAI)*, pp. 436–440, IEEE, 2016.

[18] Y. Xu, R. Wang, K. Li, and Y. Zhao, "A linear model based road identification algorithm," *journal of image and graphics*, vol. 9, no. 7, pp. 858–864, 2004.

[19] D. H. Ballard, "Generalizing the hough transform to detect arbitrary shapes," *Pattern recognition*, vol. 13, no. 2, pp. 111–122, 1981.

[20] X. Du and K. K. Tan, "Vision-based approach towards lane line detection and vehicle localization," *Machine Vision and Applications*, vol. 27, pp. 175–191, 2016.

[21] C. R. Jung and C. R. Kelber, "An improved linear-parabolic model for lane following and curve detection," in *XVIII Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI'05)*, pp. 131–138, IEEE, 2005.

[22] M. Aly, "Real time detection of lane markers in urban streets," in *2008 IEEE intelligent vehicles symposium*, pp. 7–12, IEEE, 2008.

[23] J. Wang, W. Hong, and L. Gong, "Lane detection algorithm based on density clustering and ransac," in *2018 Chinese Control And Decision Conference (CCDC)*, pp. 919–924, IEEE, 2018.

[24] M. Lee, C. Jang, and M. Sunwoo, "Probabilistic lane detection and lane tracking for autonomous vehicles using a cascade particle filter," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 229, no. 12, pp. 1656–1671, 2015.

[25] S. Zhou, J. Gong, G. Xiong, H. Chen, and K. Iagnemma, "Road detection using support vector machine based on online learning and evaluation," in *2010 IEEE intelligent vehicles symposium*, pp. 256–261, IEEE, 2010.

[26] C. Guo, S. Mita, and D. McAllester, "Robust road detection and tracking in challenging scenarios based on markov random fields with unsupervised learning," *IEEE Transactions on intelligent transportation systems*, vol. 13, no. 3, pp. 1338–1354, 2012.

[27] X. Pan, J. Shi, P. Luo, X. Wang, and X. Tang, "Spatial as deep: Spatial cnn for traffic scene understanding," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, 2018.

[28] D. Neven, B. De Brabandere, S. Georgoulis, M. Proesmans, and L. Van Gool, "Towards end-to-end lane detection: an instance segmentation approach," in *2018 IEEE intelligent vehicles symposium (IV)*, pp. 286–291, IEEE, 2018.

[29] Q. Zou, H. Jiang, Q. Dai, Y. Yue, L. Chen, and Q. Wang, "Robust lane detection from continuous driving scenes using deep neural networks," *IEEE transactions on vehicular technology*, vol. 69, no. 1, pp. 41–54, 2019.

[30] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-c. Woo, "Convolutional lstm network: A machine learning approach for precipitation nowcasting," *Advances in neural information processing systems*, vol. 28, 2015.

[31] L. Peng, Z. Chen, Z. Fu, P. Liang, and E. Cheng, "Bevsegformer: Bird's eye view semantic segmentation from arbitrary camera rigs," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 5935–5943, 2023.

[32] L. Tabelini, R. Berriel, T. M. Paixao, C. Badue, A. F. De Souza, and T. Oliveira-Santos, "Polylanenet: Lane estimation via deep polynomial regression," in *2020 25th International Conference on Pattern Recognition (ICPR)*, pp. 6150–6156, IEEE, 2021.

[33] Z. Feng, S. Guo, X. Tan, K. Xu, M. Wang, and L. Ma, "Rethinking efficient lane detection via curve modeling," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 17062–17070, 2022.

[34] D. Jin, W. Park, S.-G. Jeong, H. Kwon, and C.-S. Kim, "Eigenlanes: Data-driven lane descriptors for structurally diverse lanes," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 17163–17171, 2022.

[35] L. Chen, C. Sima, Y. Li, Z. Zheng, J. Xu, X. Geng, H. Li, C. He, J. Shi, Y. Qiao, *et al.*, "Persformer: 3d lane detection via perspective transformer and the openlane benchmark," in *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXXVIII*, pp. 550–567, Springer, 2022.

[36] F. Yan, M. Nie, X. Cai, J. Han, H. Xu, Z. Yang, C. Ye, Y. Fu, M. B. Mi, and L. Zhang, "Once-3dlanes: Building monocular 3d lane detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 17143–17152, 2022.

[37] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.

[38] J. Dai, H. Qi, Y. Xiong, Y. Li, G. Zhang, H. Hu, and Y. Wei, "Deformable convolutional networks," in *Proceedings of the IEEE international conference on computer vision*, pp. 764–773, 2017.

[39] S. Lee, J. Kim, J. Shin Yoon, S. Shin, O. Bailo, N. Kim, T.-H. Lee, H. Seok Hong, S.-H. Han, and I. So Kweon, "Vpgnet: Vanishing point guided network for lane and road marking detection and recognition," in *Proceedings of the IEEE international conference on computer vision*, pp. 1947–1955, 2017.