Lane Detection for Autonomous Driving: Comprehensive Reviews, Current Challenges, and Future Predictions

Jiping Bi, Yongchao Song, Yahong Jiang, Lijun Sun, Xuan Wang, *Senior Member, IEEE*, Zhaowei Liu, *Member, IEEE*, Jindong Xu, Siwen Quan, Zhe Dai, and Weiqing Yan, *Senior Member, IEEE*

***Abstract*— Lane detection is crucial for autonomous driving systems (ADS), utilizing sensors like cameras and LiDAR to identify lanes and understand vehicle position, direction, and lane shape. It provides data support for the control system to make informed driving decisions. In this survey, we review recent advancements in lane detection, focusing on both 2D techniques and emerging 3D methods. We begin with an overview of the significance of lane detection in ADS, followed by an analysis of the evolution of 2D techniques over the past decade, covering traditional and deep learning approaches. We also examine recent advancements in 3D lane detection. Additionally, we summarize evaluation metrics and popular datasets in the field. Finally, we discuss current challenges and future directions in lane detection, aiming to provide valuable insights for researchers and developers in this technology.**

***Index Terms*— Computer vision, 2D lane detection, 3D lane detection, autonomous driving.**

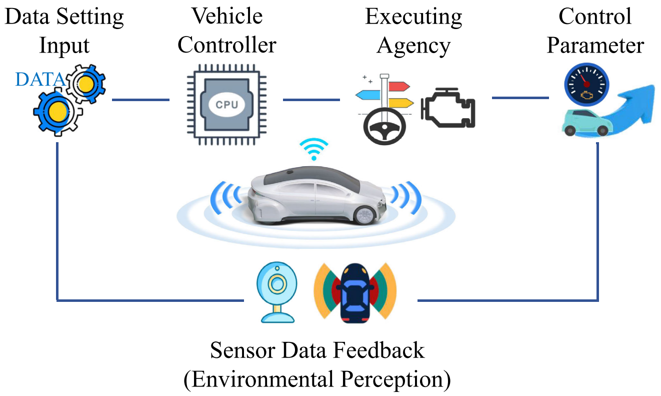


Fig. 1. Diagram of the autonomous driving system with sensing, control, and execution layers. It illustrates that perception is essential; only through accurate environmental sensing can the controller make informed decisions.

I. Introduction

According to the recently published Global Status Report on Road Safety 2023 by the World Health Organization [1], the number of deaths resulting from road traffic accidents reaches 1.19 million per year. In most traffic accidents, human errors are often the primary cause, such as driver fatigue, drunk driving, excessive speed, and inattentive driving. However, with the exponential growth of computer and information technology, autonomous driving technology has emerged as a game-changing [2], efficient method to significantly reduce the frequency of traffic accidents. Autonomous vehicles are equipped with state-of-the-art sensors, which enable them to monitor their surroundings in real-time [3]. Subsequently, the vehicles utilize advanced control algorithms to make precise decisions aimed at avoiding traffic accidents. The successful realization of the above functions requires three essential layers: the sensing layer, control layer, and execution layer [4]. The primary function of the sensing layer is to gather data using a variety of onboard sensors, including cameras and LiDAR, which allows the vehicle to recognize its current state and fluctuations in the surrounding environment. This information is then delivered in real time to the control unit. The control layer processes and analyses the data obtained from the sensors, subsequently transmitting appropriate control signals to the target devices. Finally, after receiving commands from the control center, the execution layer permits the vehicle to perform the appropriate actions as illustrated in Fig. 1.

Autonomous driving technology can be classified as L0 to L5 [5], which encapsulates the full spectrum of driving from purely human operation to fully automated driving. The majority of vehicle companies have successfully reached the L2 level of autonomous driving, and a select few have surpassed that, achieving the L3 level of autonomy [6]. The current generation of vehicles incorporates advanced and highly

Received 29 July 2024; revised 4 November 2024 and 16 December 2024; accepted 29 December 2024. This work was supported in part by the Natural Science Foundation of Shandong Province under Grant ZR2022QF037 and in part by the National Natural Science Foundation of China under Grant 62072391. The Associate Editor for this article was D. F. Wolf. *(Corresponding author: Yongchao Song.)*

Jiping Bi, Yongchao Song, Lijun Sun, Xuan Wang, Zhaowei Liu, Jindong Xu, and Weiqing Yan are with the School of Computer and Control Engineering, Yantai University, Yantai 264005, China (e-mail: bijiping@s.ytu.edu.cn; ycsong@ytu.edu.cn; sunlijun@s.ytu.edu.cn; xuanwang91@ytu.edu.cn; lzw@ytu.edu.cn; xujindong@ytu.edu.cn; wqyan@tju.edu.cn).

Yahong Jiang and Zhe Dai are with the School of Transportation Engineering, Chang'an University, Xi'an 710064, China (e-mail: 2018023002@chd.edu.cn; zhedai@chd.edu.cn).

Siwen Quan is with the School of Electronics and Control Engineering, Chang'an University, Xi'an 710064, China (e-mail: siwenquan@chd.edu.cn).

Digital Object Identifier 10.1109/TITS.2024.3524603

intelligent systems, enabling them to function autonomously without human interference. A prerequisite for this technology is environmental perception [7], i.e., the ability of the vehicle to accurately recognize and understand its surroundings to make appropriate driving decisions. Environmental perception technology utilizes a wide range of sensors, including LiDAR, cameras, and so on. The sensors can swiftly and accurately detect real-time road conditions, traffic, surrounding vehicles, and obstacles. To accurately perceive the road conditions, the vehicle makes driving decisions by analyzing and processing the collected data [8]. Lane detection is a crucial aspect of autonomous driving and environmental perception, analogous to the sensor acquisition in the Internet of Things (IoT) [9]. It allows vehicles to navigate the correct path by identifying and tracking lanes on roads using sensors like cameras [10].

Lanes are prevalent traffic signs in the road scenario, making their detection imperative for vehicle positioning and navigation. It is unique due to their diversity and can be difficult to detect under potential environmental disturbances, thus necessitating a high degree of accuracy and robustness. Lane detection technology is required to effectively handle a multitude of lane variations, including various categories, colors, and continuity, while also being adaptable to a wide array of environmental factors such as shadows, occlusion, wear and tear, weather changes, and lighting conditions to guarantee precise vehicle positioning and reliable navigation. Lane detection plays an integral role in various aspects of vehicle driving systems, including lane departure warning (LDW), lane keeping assist (LKA), lane change assist (LCA), forward collision warning (FCW), adaptive cruise control (ACC), blind spot detection (BSD), and so on [11].

Lane detection techniques can be categorized into 2D and 3D detection methods depending on the detection dimensions. Academic and industrial communities have long been committed to 2D plane lane detection, and the outcomes have remarkable results. These innovative methods encompass traditional computer vision [12], deep learning [13], and multi-sensor fusion [14]. However, the detected lanes have only two-dimensional position information, and they require stringent conditions for practical application. To address these issues, there has been a shift in research focus towards developing 3D lane detection technology [15], [16], [17]. This technology is capable of providing direct detection of lane positions within 3D space, providing more precise 3D lane coordinate information for driverless and assisted driving systems. The precise localization capability contributes significantly to the precise identification of the vehicle's location on the road, hence enabling more precise driving path planning and vehicle control.

The contributions of this paper are as follows:

- We describe the position of lane detection in automated driving and the related status quo.
- We comprehensively review the state-of-the-art lane detection methods of the last decade, elucidating trends in traditional, 2D, and 3D methods.
- This paper highlights and provides a detailed analysis of several popular datasets and the efficiency of lane detection methods applied to them.
- Systematic description of lane detection positioning, current challenges, and future possible directions, with the aim that this will facilitate the community to move further along the journey of automated driving and lane detection.

To the best of our knowledge, the paper is the first comprehensive review of lane detection methods developed in the last decade. The remaining investigations are organized as follows. Section II describes the actual development of autonomous driving and the technical difficulties of the lane detection task. Section III summarizes the 2D lane detection methods of the last decade. Section IV reviews the recent years of 3D lane detection methods. Section V presents the existing dataset, evaluation metrics, and assessment of existing open-source lane detection methods. Subsequently, section VI conclude with a summary of the current challenges and possible directions for the future. Finally, Section VII concludes the paper. The structural framework of this paper is shown in Fig. 2.

II. Background

The field of autonomous driving technology has been extensively explored and developed since the 1970s in various developed nations including the United States, United Kingdom, and Germany. The development of the Driver Assistance System (DAS) in the 1980s marked the emergence of a strategy designed to diminish driver error and boost road safety. The advanced technology is designed to identify and alert drivers to blind spots, thereby creating a more thorough and safe driving experience. The Advanced Driver Assistance System (ADAS) is an advanced version of the original system and its subsequent iterations are based on this foundation [18]. In 2004, the Defense Advanced Research Projects Agency (DARPA) initiated the Autonomous Driving Challenge in the Mojave Desert, thereby catalyzing the technological revolution in autonomous driving [19]. In 2011, the Hong Qi HQ3 vehicle, a joint development of China FAW Group and the National University of Defense Technology (NUDT) conducted a high-speed unmanned driving experiment, covering a remarkable distance of 886 kilometers. In 2014, Google developed a fully autonomous driving vehicle, capable of operating without any human input or supervision. In October 2016, the ceremony of BAIC's autonomous driving and project operation was held in Panjin, Liaoning Province. Since December 2017, four "Alpha" intelligent buses with autonomous driving capabilities have been operating in designated areas of Shenzhen. In February 2018, BYD's fleet of autonomous vehicles, operating on Baidu's Apollo system, accomplished a successful crossing of the Hong Kong-Zhuhai-Macao Bridge in an organized formation. In 2020, Tesla launched a new feature, known as Full Self-Driving (FSD), on top of the pre-existing Autopilot feature. At IoT Expo 2021, Audi demonstrated its advanced V2X (vehicle-to-extraterrestrial) L4 autonomous driving technology on a public road, highlighting its feasibility and future applications. In 2023, the self-driving company Zoox, a subsidiary of Amazon, conducted a successful test of its Robotaxi technology on public roads in Las Vegas. In February 2024, Baidu's Apollo Go vehicle crossed the Yangsigang Yangtze River Bridge and Baishazhou Bridge,

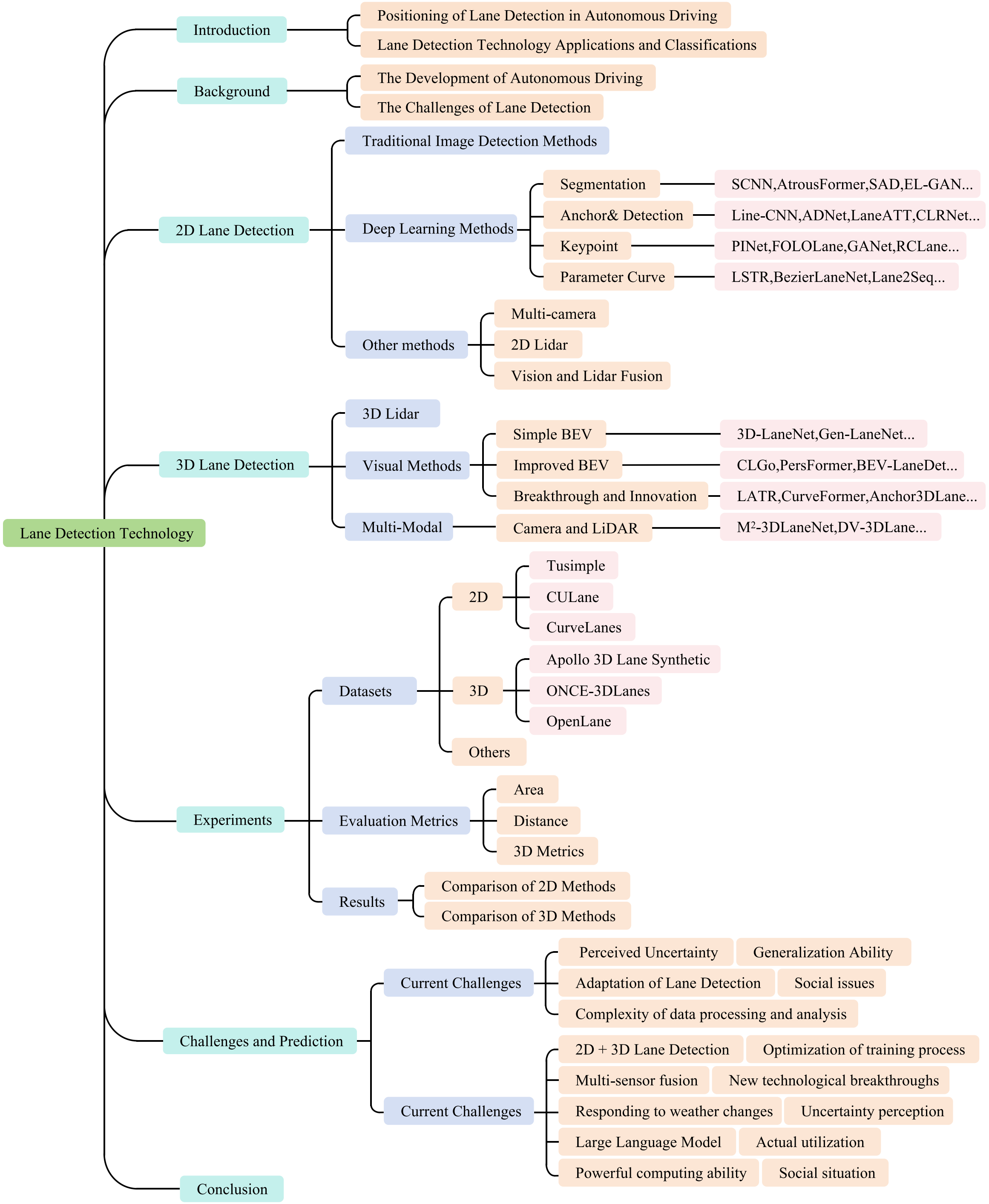

Lane Detection Technology
Introduction
Positioning of Lane Detection in Autonomous Driving
Lane Detection Technology Applications and Classifications
Background
The Development of Autonomous Driving
The Challenges of Lane Detection
2D Lane Detection
Traditional Image Detection Methods
Deep Learning Methods
Segmentation
SCNN,AtrousFormer,SAD,EL-GAN...
Anchor& Detection
Line-CNN,ADNet,LaneATT,CLRNet...
Keypoint
PINet,FOLOLane,GANet,RCLane...
Parameter Curve
LSTR,BezierLaneNet,Lane2Seq...
Other methods
Multi-camera
2D Lidar
Vision and Lidar Fusion
3D Lane Detection
3D Lidar
Visual Methods
Simple BEV
3D-LaneNet,Gen-LaneNet...
Improved BEV
CLGo,PersFormer,BEV-LaneDet...
Breakthrough and Innovation
LATR,CurveFormer,Anchor3DLane...
Multi-Modal
Camera and LiDAR
M^2-3DLaneNet,DV-3DLane...
Experiments
Datasets
2D
Tusimple
CULane
CurveLanes
3D
Apollo 3D Lane Synthetic
ONCE-3DLanes
OpenLane
Others
Evaluation Metrics
Area
Distance
3D Metrics
Results
Comparison of 2D Methods
Comparison of 3D Methods
Challenges and Prediction
Current Challenges
Perceived Uncertainty
Generalization Ability
Adaptation of Lane Detection
Social issues
Complexity of data processing and analysis
Current Challenges
2D + 3D Lane Detection
Optimization of training process
Multi-sensor fusion
New technological breakthroughs
Responding to weather changes
Uncertainty perception
Large Language Model
Actual utilization
Powerful computing ability
Social situation
Conclusion

Fig. 2. Structural framework of this paper.

completing the journey without any driver input. As shown in Fig. 3.

Lane detection is the foundation of ADS, and with the gradual maturation of ADS, lane detection is also evolving. As the name implies, lane detection is designed to precisely determine the position, orientation, and shape of the lane lines. In reality, lane visibility is significantly influenced by numerous factors such as shadows cast by trees and buildings lining the roadsides, occlusions created by vehicles and pedestrians, and various weather conditions like rain, snow, dust, and intense sunlight. As shown in Fig. 4. Therefore lane detection must be robust and keep lane identification even in complex road conditions, different lighting conditions, or encountering occlusions [20].

In the past, when large lane line detection datasets were not available, lane detection relied heavily on manual feature-based methods, which were only capable of achieving the desired performance when lane lines were visible [21]. With the changing and increasingly complex road conditions, manual methods of detection are no longer sufficient due to their inherent limitations and inconsistencies. As opposed to these methods, deep learning methods are capable of automatically

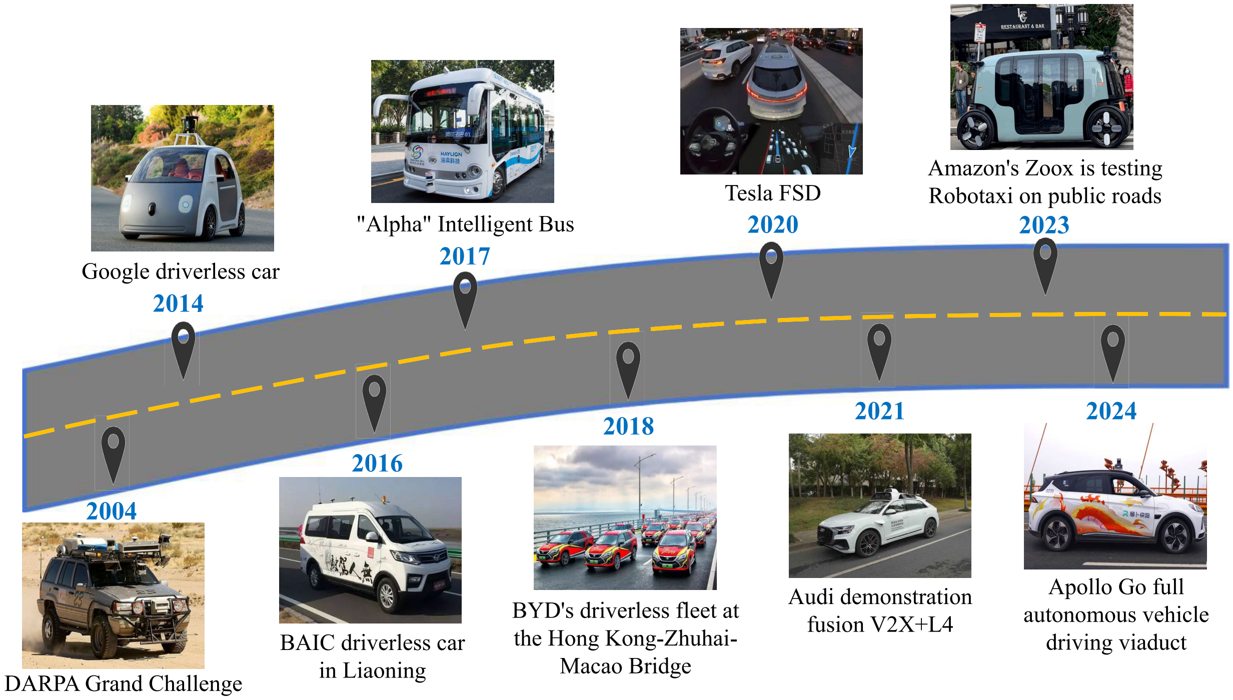

Google driverless car
2014
"Alpha" Intelligent Bus
2017
Tesla FSD
2020
Amazon's Zoox is testing Robotaxi on public roads
2023
2004
DARPA Grand Challenge
2016
BAIC driverless car in Liaoning
2018
BYD's driverless fleet at the Hong Kong-Zhuhai-Macao Bridge
2021
Audi demonstration fusion V2X+L4
2024
Apollo Go full autonomous vehicle driving viaduct

Fig. 3. Practical developments in autonomous driving application technology.

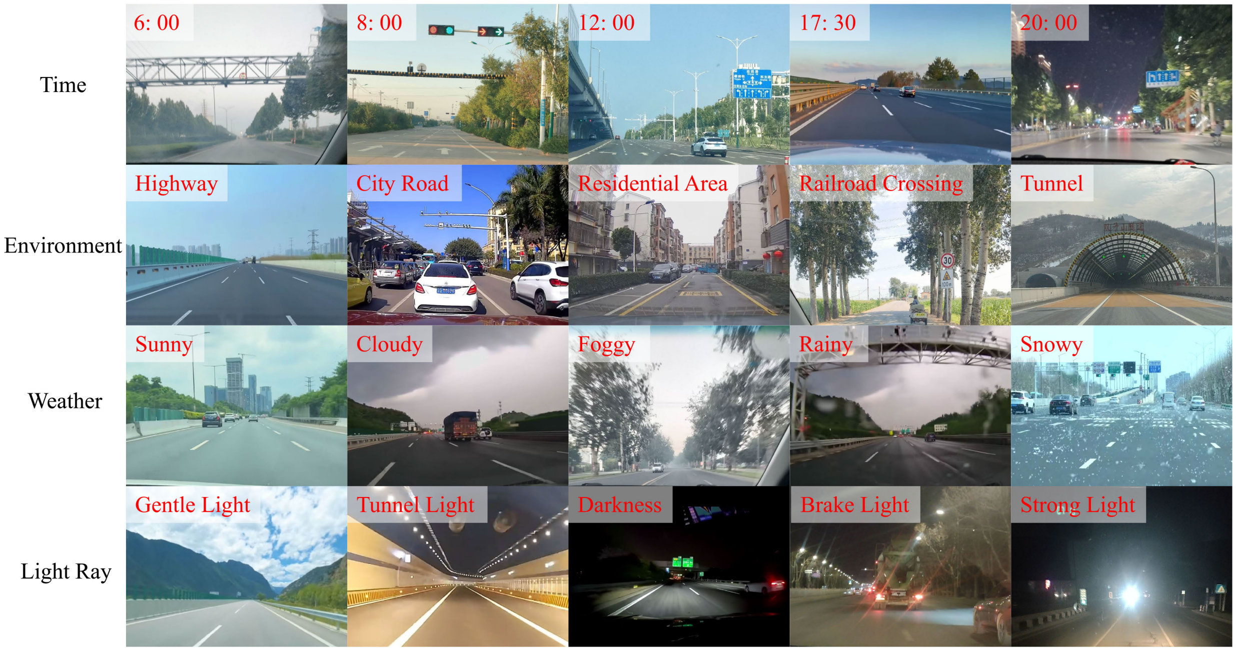


Fig. 4. Various factors affecting lane visibility in real-world environments.

extracting features that align with those required for lane detection. Consequently, they are equipped to yield improved detection results under conditions such as occlusion or missing lane and so on [13]. As shown in Fig. 5, the evolution of lane detection methods is demonstrated. Relevant research results have emerged over time, reflecting the continuous development and widespread interest in this field.

III. Overview of 2D Lane Detection

The development of 2D lane detection technology plays a pivotal role in the deployment of autonomous and assisted driving systems. Researchers have worked tirelessly to continuously enhance the performance of 2D lane detection. The current methods for lane detection can be categorized as traditional image detection methods [22] and deep learning methods [13].

A. Traditional Image Detection Methods

The traditional lane detection process can be described as follows: In the first stage, an image of the front of the vehicle is captured by utilizing a camera or other suitable sensor. The acquired image undergoes a series of pre-processing

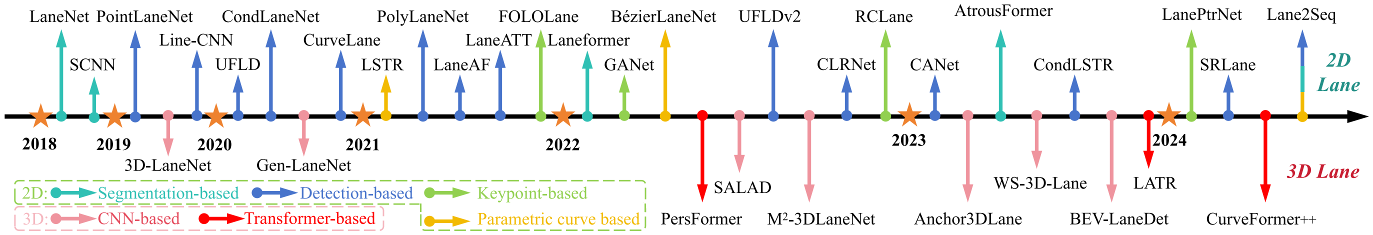


Fig. 5. The different mainstream methods in the field of lane detection in recent years. In the upper part of the axes 2D and the lower part 3D. Each dimension is subdivided into several types, which are labeled in the lower left corner of the figure.

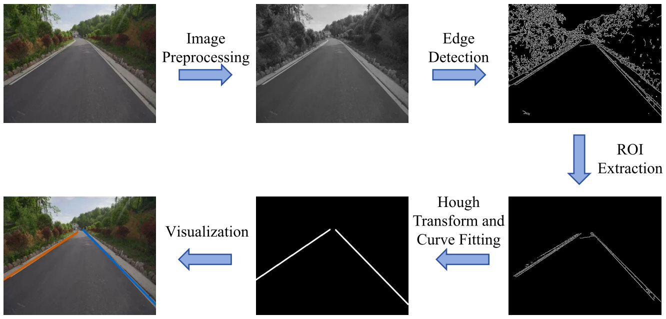


Fig. 6. The general procedure involved in traditional lane detection methods.

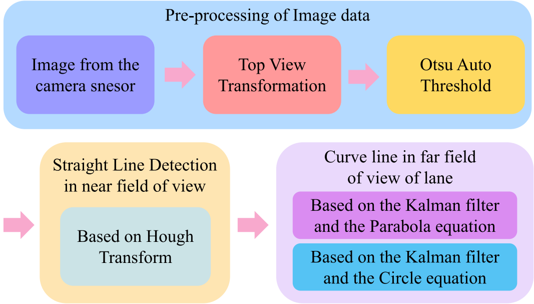

Pre-processing of Image data
Image from the camera snesor
Top View Transformation
Otsu Auto Threshold
Straight Line Detection in near field of view
Based on Hough Transform
Curve line in far field of view of lane
Based on the Kalman filter and the Parabola equation
Based on the Kalman filter and the Circle equation

Fig. 7. Flowchart of lane detection algorithm using Kalman filter [30].

operations prior to subsequent processing. These operations include image noise reduction, color space conversion, and image enhancement, with the aim of enhancing its effectiveness and accuracy. Next, the preprocessed image is used to extract feature information about lane, focusing on characteristics such as color and edge information. The extracted features from the image should undergo further processing to enhance lane visibility, such as binarization, filtering, and edge detection. The common filtering methods are the mean filter [23], median filter [24], bilateral filter [25], and Gaussian filter [26]. Sobel algorithm [27] and Canny algorithm [28] are the most prevalent edge detection techniques. Utilizing the lane model (e.g., straight or curved line) and the extracted features, a specific algorithm (e.g., Hough Transform or Least Squares) is employed to accurately estimate the parameters of the lane, thereby obtaining the precise location and shape of the lane [29]. Finally, the detected lanes are visually represented on the image to aid the driver's visual perception and decision-making process. Refer to Fig. 6 for a visual representation of this process.

Furthermore, as the research continued, traditional lane detection methods were overhauled and innovated based on their original concepts. Like Fig. 7, Dorj et al. [30] proposed a Kalman filter-based algorithm for hypersurface lane detection. The method first transforms the image into a top view with Otsu thresholding [31]. Then, the curved lanes can be efficiently detected by parabolic [32] and circular models [33].

Traditional image processing methods have several limitations in lane line detection, including reliance on complex preprocessing and strong a priori assumptions, leading to poor algorithm portability. Hand-designed features (e.g., colors, edges, or textures) have limited expressiveness, are susceptible to scene noise, and are less robust to environmental changes, often leading to misjudgments and omissions. However, traditional methods have not been eliminated; they have played a significant role in the development of automobiles over the past few decades. Currently, traditional techniques remain effective in certain situations, particularly in environments with good lighting or simple road structures. For instance, classical approaches like the Hough transform continue to deliver reliable results in scenarios with clear features. Additionally, traditional methods offer high real-time performance in embedded systems due to their low computational complexity, one of their key advantages.

With the rapid advancement of autonomous driving technology, the integration of deep learning has enhanced systems' ability to adapt to complex environments and improve accuracy. Deep learning methods overcome the limitations of traditional post-processing methods by automatically learning complex features, allowing detection to adapt to more complex scenarios. Future lane line detection techniques are likely to lean more heavily on deep learning, but leveraging the strengths of traditional methods remains a viable strategy, especially in contexts where real-time performance and computational resources are constrained.

B. Deep Learning Methods

Deep learning methods, such as Convolutional Neural Networks (CNNs), offer a more efficient and accurate solution to the lane detection problem, offering several advantages over traditional methods. CNNs excel in handling complex image variations and adapting to different road environments and lane types [34]. The multilayer network structure captures high-level abstract features in the image, leading to superior lane detection accuracy compared to traditional methods. Moreover, deep learning models allow for end-to-end processing from raw images to lane detection results, reducing the need for cumbersome intermediate steps and increasing processing speed. The model is also resistant to noise and interference, making it ideal for real-time applications. Additionally, deep

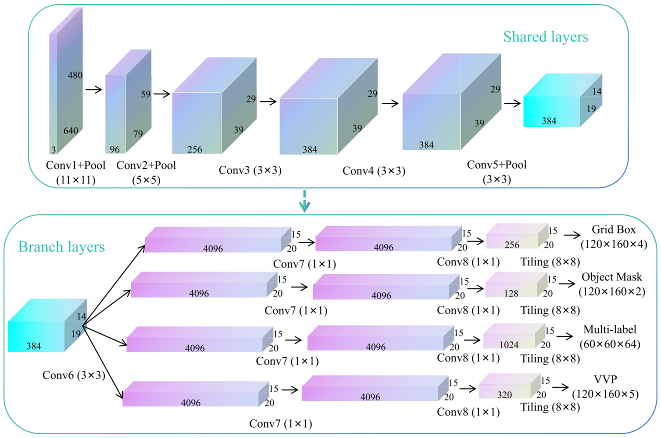


Fig. 8. VPGNet [50] framework structure. It can perform four tasks: lattice regression, object detection, multi-label classification and vanishing point prediction.

learning models can process large-scale datasets and integrate them with other models to enhance the functionality of autonomous driving systems. Based on the representation of lane lines, deep learning methods can be categorized into the following four classes: segmentation-based methods, detection-based methods, keypoint-based methods, parametric regression based methods.

1) Segmentation-Based Methods: Lane detection using semantic segmentation allows accurate identification of lane and non-lane areas by classifying each pixel in the image. The process begins by extracting features from a monocular image to obtain key features and initial lane segmentation results. These results are then refined by a back-end optimization network to generate more accurate lane features. This network classifies each pixel in the image as a lane or non-lane by generating probabilities. Finally, detection results and lane coordinates are generated based on these probability values.

By using pixel-level classification to separate lanes from the background, FCN [51] achieves end-to-end semantic segmentation at the pixel level. It enables the computation of the classification loss between related pixels during backpropagation. U-Net [44], [52] is a neural network architecture using an encoder-decoder structure and upsampling paths for lane segmentation. This method uses channel dimension features to form thicker features for capturing complex lane boundaries. In addition, it maintains feature correlations with minimal training data to maintain accuracy. ENet [53] improves efficiency and accuracy through multi-resolution paths and hopping connections, delivering efficient and accurate lane detection and segmentation in real time. DeepLab [54] lane detection and segmentation accuracy are achieved by the use of null convolution techniques and multi-scale information fusion. As the technology advances, it incorporates CRF and encoder-decoder designs to provide extremely high-resolution images. VPGNet [50] provides a structural guidance mechanism for marking detection by predicting extinction point locations. This model has a multi-task learning framework that can perform predictions not only at the pixel level for lane segmentation, but also at the frame level for detection, at the target level for masks, at the label level for multi-category, and at the location level for extinction points. As shown in Fig. 8. EDANet [55] can skillfully extract the unique characteristics of each layer and summarize the multi-scale information, thereby significantly enhancing the accuracy of lane marking segmentation. In different scenarios, these methods possess unique strengths and limitations, which should be selected and optimized meticulously according to specific requirements.

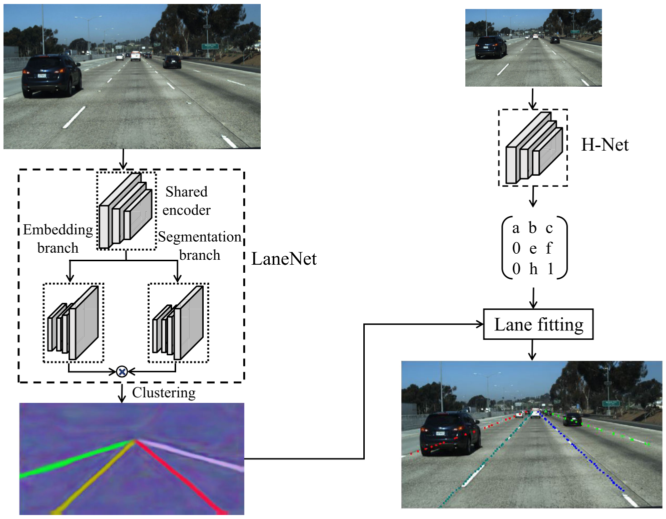


Fig. 9. LaneNet detection system [35]. It has a segmentation and embedding branch. The segmentation branch makes binary lane masks, and the embedding branch clusters and assigns lane pixels to cluster centers with lane IDs. The system first processes an input image to get a lane instance mapping, which is then used to learn perspective transformations. Polynomial fitting then projects the lanes back onto the image.

A brief overview of some of the currently available 2D lane detection methods on segmentation is given in Tab. I, with details given successively later in the paper.

Several new methods have been introduced recently, extending the traditional semantic segmentation techniques mentioned previously. For example, LaneNet [35] architecture utilizes an end-to-end design concept, leveraging a two-branch neural network. This design allows for the successful detection of lane and the execution of pixel-level semantic segmentation with a single forward propagation. As shown in Fig. 9. To efficiently mine the spatial relationships of row and column pixels of linear targets, SCNN [36] proposes a special convolutional approach, shown in Fig. 10. This method extends the traditional deep layer-by-layer convolution to a more comprehensive slice-by-slice convolution in feature mapping. By permitting the information to be aggregated in different dimensions through slices, this strategy ultimately enables the interconnection of pixels within rows and columns of the layer. Wang et al. [56] propose a novel multitask method for lane marking detection that combines CNN-based semantic modeling with handcrafted features for improved localization and introduces a vanishing line prediction for better accuracy in sharp curves and nonflat roads.

RESA [37] expands upon the SCNN model by utilizing the Recursive Feature Shift Aggregator (RESA), which enables feature map slices to continuously shift both vertically and horizontally to collect spatial information between pixels, particularly in regions spanning rows and columns. As shown in Fig. 11. Similarly, PSSNet [38] proposes an

TABLE I

AN OVERVIEW OF SOME TYPICAL 2D LANE DETECTION METHODS ON SEGMENTATION, INCLUDING THE OFFICIAL PUBLICATION DATE, PUBLISHING ORGANIZATION, ADVANTAGES OF THE METHOD, NETWORK ARCHITECTURE, AND OPEN-SOURCE STATUS OF THE PROGRAM. * DENOTES PREPUBLICATION

Methods	Year	Journals / Conferences	Strengths	Architecture	Code
LaneNet [35]	2018	IEEE IV	An embedding vector can be used with a cluster to detect an unlimited number of lane lines. The perspectival transformation learned through HNet enhances the robustness of lane fitting.	CNN	link
SCNN [36]	2018	AAAI	It has the capability to fully explore spatial relationships on image rows and columns. It can make residuals of the transmitted information, which is beneficial for recognizing large objects.	CNN	link
RESA [37]	2021	AAAI	Highly efficient aggregation of information.	CNN	link
PSSNet [38]	2021	BMVC	It proposes a decomposition of spatially separable convolution. PSSNet can be implemented by parallel computing.	CNN	-
LaneAF [39]	2021	IEEE Rob. Autom	Complex problems can be broken down into simpler ones. Decoding algorithms are provided for predicting affinity fields.	CNN	link
Laneformer [40]	2022	AAAI	Combines remote lane points and surrounding objects to induce global contextual relationships.	Transformer	-
AtrousFormer [41]	2023	Pattern Recognition	Global AtrousFormer gathers data globally in an unstructured manner, which is then transferred to the converter via ASPP. It then gets more streamlined by integrating a slicing mechanism.	Transformer	-
ENet-SAD [42]	2019	ICCV	A knowledge distillation-based lane line detection model SAD is proposed, which can enhance the feature representation ability of CNN.	CNN	link
IntRA-KD [43]	2020	CVPR	Interregional affinity maps and attentional map distillations are introduced to efficiently transfer structural knowledge.	KD	link
EL-GAN [44]	2018	ECCV	A new method has been introduced to address complex problems like semantic segmentation by implementing a generative adversarial network architecture and a discriminator trained on both prediction and labeling. The training process is greatly stabilized by embedding loss, which provides more useful gradient feedback than ordinary adversarial loss simulations.	GAN	-
Ripple-GAN [45]	2020	IEEE TITS	It presents RiLLD-Net, a new network that passes feature maps between modules of different spacing, and Ripple-GAN, which is used in combination with RiLLD-Net.	GAN	-
SIM-CycleGAN [46]	2020	IEEE IV	It proposes an efficient data enhancement method under low illumination conditions using the light condition style transfer method. This method solves the scale change problem.	GAN	-
YOLOP [47]	2022	MIR	A single panoramic driving perception network can simultaneously perform traffic target detection, drivable area segmentation, and lane detection. Multi-tasking can be done simultaneously in real-time on an embedded device, the Jetson TX2.	CNN	link
Q-YOLOP [48]	2023	ICMEW	A low-power multitasking model has been developed specifically for traffic scenarios, to overcome the challenges of object detection and semantic segmentation. It can accurately recognizing objects and segmenting lane lines and drivable areas, all while maintaining the lowest computational cost. This makes it ideal for deployment in resource-constrained environments such as mobile devices, IoT devices, and embedded systems.	CNN	-
A-YOLOM [49]	2023	CV*	A new, lightweight model has been created which can handle multiple tasks within a single model. This model includes an adaptive connectivity module for segmenting the neck region and a lightweight and generalized segmentation for the head region.	CNN	link

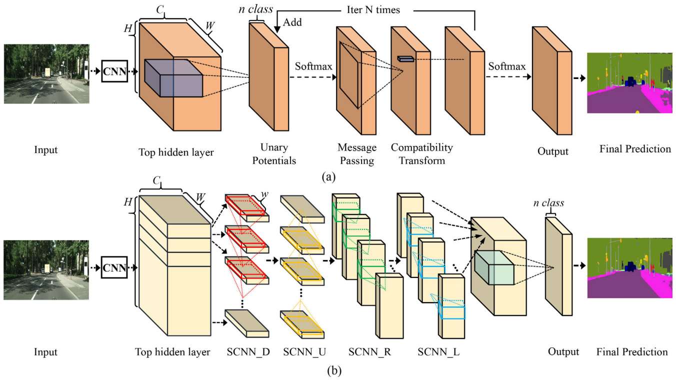


Fig. 10. (a) MRF/CRF-based methods. (b) Spatial CNN [36] realization. Compared to MRF/CRF, SCNN can be applied to the top hidden layer with richer information.

innovative approach called Parallel Space Separation Convolution (PSS-Conv). This method effectively leverages the decomposition of parallel space convolution and merges channel-weighted features for more efficient feature aggregation. LaneAF [39] combines the Affinity Field [57] with traditional binary semantic segmentation, achieving the segmentation of an arbitrary number of lane instances via clustering.

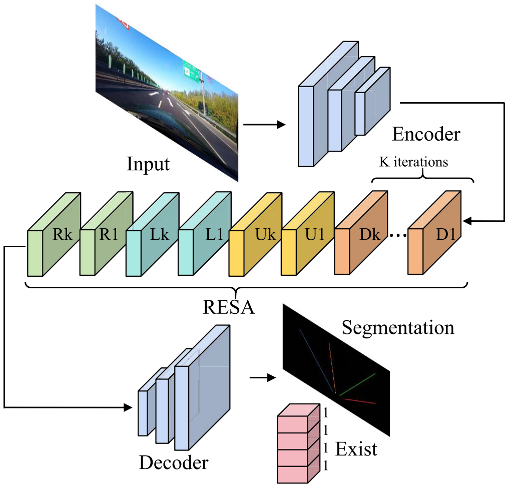


Fig. 11. RESA [37] Architecture Design. It composed by encoder, RESA and decoder. 'Dk', 'Uk', 'Lk', 'Rk' denotes "up-to-down", "down-to-up", "right-to-left", and "left-to-right" respectively at k-thiteration in RESA.

In the YOLO family, YOLOP [47] performs simultaneous traffic target detection, drivable area segmentation, and lane detection with one encoder for feature extraction and three decoders for specific tasks. Q-YOLOP [48] employs an Efficient Layer Aggregation Network (ELAN) as its backbone and

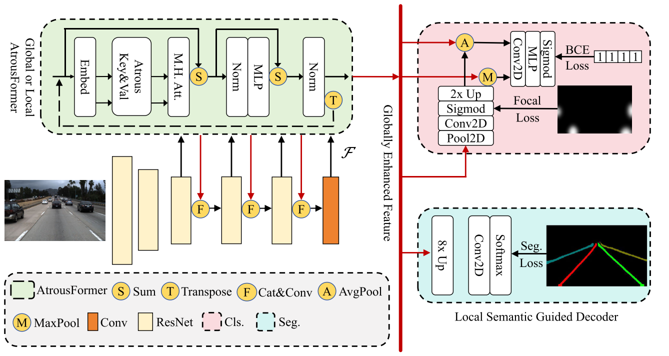


Fig. 12. AtrousFormer [41] architecture overview. In this, the raw image is sequentially passed through Local Augmentation Enhanced Extractor, 1 × 1 Convolution, Global AtrousFormer, and Local Semantic Guided Decoder, and finally the segmentation map and corresponding classification scores are generated.

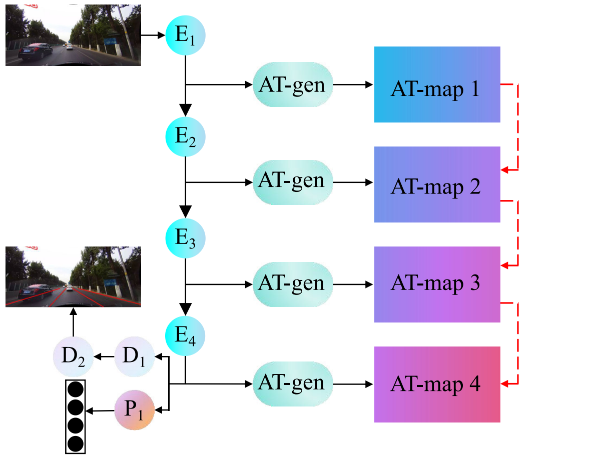


Fig. 13. The process of running the SAD [42] architecture. $E_1 \sim E_4$ comprise the encoder of ENet [53], D_1 and D_2 comprise the decoder of ENet. P_1 is a lane prediction network. AT-gen is the attention generator.

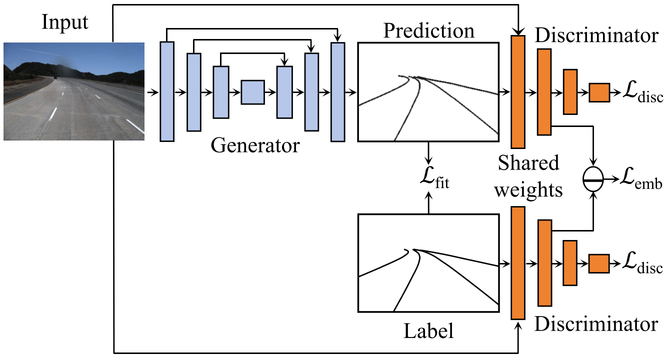


Fig. 14. Overview of the EL-GAN architecture [44].

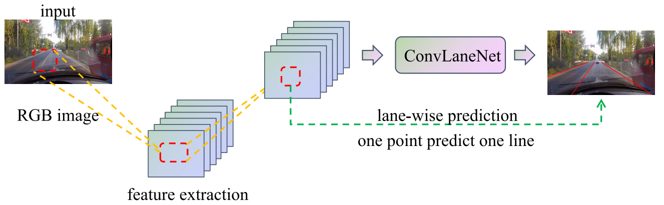


Fig. 15. A brief overview of the PointLaneNet method [60]. It generates coordinates directly from the original input image.

task-specific headers for each task to enable lane segmentation. A-YOLOM [49] introduces an adaptive module specifically for segmentation Neck that can efficiently handle multiple tasks using one model. It eliminates the need for different designs for different scenario tasks and realizes lane segmentation with high accuracy.

Recently, due to the high-performance appeal of the Transformer [58], researchers have also incorporated it into lane detection. Laneformer [40] converts the traditional transformer into a model that can more effectively capture lane shapes and semantic features, thereby minimizing latency. AtrousFormer [41] offers enhanced information-gathering capabilities and improved computational efficiency for the network through its dedicated first-in-first-out approach to information gathering. It guides the decoder through local semantics to accurately characterize the identity and shape of the lanes. To aid this process, predictive Gaussian maps of the starting point of each lane are employed. As shown in Fig. 12.

Furthermore, researchers have explored the potential of Knowledge Distillation (KD) [42] or Generative Adversarial Networks (GAN) [59] to tackle the lane detection issue. SAD [42] proposes a novel, lightweight lane detection method using Self-Attention Distillation, which enables the model to extract multiple attention maps at the encoder level and utilize the top-level attention to regulate the learning of the bottom-level attention. As shown in Fig. 13. Through self-learning, the model is capable of achieving significant improvements without the need for any additional supervision or labeling, thereby enhancing overall segmentation accuracy. IntRA-KD [43] relies on Inter-Region Affinity Distillation to fragment a specific road scene image into numerous distinct regions, each of these regions is symbolized as a node within the graph, and links between nodes are formed depending on their similarity in feature distribution, thereby enabling the shallow network to attain a performance comparable to the deeper network through attentional distillation. EL-GAN [44] employs lane line labels as additional inputs and makes use of GAN to generate segmentation maps that are similar to the true value labels. As shown in Fig. 14. Ripple-GAN [45] incorporates the concepts of feature fusion, Wasserstein generative adversarial training, and multi-objective segmentation, resulting in superior performance, particularly when the lane labeling information is incomplete. SIM-CycleGAN [46] utilizes GAN to create low-light images, which enhances the model's adaptability to low-light environments.

2) Detection-Based Methods: Anchor-based methods generate a set of predefined anchor boxes to efficiently locate lanes within an image. Initially, a variety of anchor frames with different sizes and aspect ratios are created. The anchor point algorithm then analyzes each anchor box using a neural network model to determine if it contains a lane. Finally, the exact location of the lane is further refined based on the anchor points. A brief overview of some of the currently available 2D lane detection methods on detection is given in Tab. II, with details given successively later in the paper.

As shown in Fig. 15, PointLaneNet [60] determines the shape of the lane and the classification of the lanes by creating horizontal straight lines on the input image and determining the shape of the lane and the classification of

TABLE II
AN OVERVIEW OF SOME TYPICAL 2D LANE DETECTION METHODS ON DETECTION, INCLUDING THE OFFICIAL PUBLICATION DATE, PUBLISHING ORGANIZATION, ADVANTAGES OF THE METHOD, NETWORK ARCHITECTURE, AND OPEN-SOURCE STATUS OF THE PROGRAM. * DENOTES PREPUBLICATION

Methods	**Year**	**Journals / Conferences**	**Strengths**	**Architecture**	**Code**
PointLaneNet [60]	2019	IEEE IV	A novel lane detection algorithm is presented that enables the direct extraction of lane line coordinates from the driving environment. The approach boasts a compact network architecture that is well-suited for deployment on the NVIDIA PX2 embedded platform. With the ability to effectively recognize lane lines under challenging conditions such as occlusion, curves, shadows, and traffic congestion, this method demonstrates robust and stable performance.	CNN	-
Line-CNN [61]	2019	IEEE TITS	A novel Line Proposal Unit (LPU) is proposed to locate traffic routes by directly learning global and complete feature representations of traffic routes.	CNN	-
SIIC-Net [62]	2022	CVPR	A new concept called “feature lanes” has been introduced. Feature lanes are used as descriptors for data segmentation to represent structurally distinct lanes in a compact manner in the feature lane space. To detect and regress road lanes in the feature lane space, a new network called SIIC-Net has been developed which is highly efficient.	CNN	link
ADNet [63]	2023	ICCV	To emphasize the importance of anchor flexibility for anchor-based approaches through explicit deconstructive learning of starting point heat maps and their associated directions. Investigates the effectiveness of large kernel mechanisms in lane detection tasks to ensure anchoring quality.	CNN	link
O2SFormer [64]	2023	CV*	Encoding lane anchors as location queries and updating lane anchors layer by layer.	Transformer	link
Sparse Laneformer [65]	2024	CV*	A sparse anchor generation scheme is proposed to generate sparse anchors using position-aware lane queries and angle queries.	Transformer	-
UFLD [66]	2020	ECCV	It can solve problems at very high speeds without visual cues, using only global features for prediction.	CNN	link
ERF-E2E [67]	2020	CVPR	Each lane marker is classified and its vertices are obtained in an end-to-end manner. A new effective level reduction module is designed which does not require any complex post-processing.	CNN	link
CurveLane-NAS [68]	2020	ECCV	The framework facilitates automatically fuse and capture of both long-ranged coherent and precise curve information, establishing a robust computational allocation process.	CNN	link
CondLaneNet [69]	2021	ICCV	The proposed conditional lane detection strategy and row-wise formulation significantly enhanced the capability to discriminate lane instances at the instance level. Additionally, the proposed RIM effectively tackled the intricate problem of lane line detection under complex topologies, such as dense lines and fork lines.	CNN	link
LaneATT [70]	2021	CVPR	A new anchor-based attention mechanism is proposed to aggregate global information with faster training and inference time.	CNN	link
UFLDv2 [71]	2022	IEEE TPAMI	Considering the magnified error problem, propose a new hybrid anchor system designed to effectively diminish localization error.	CNN	link
CANet [72]	2023	IEEE ICASSP	The first mechanism proposes a “curved guide line” to restrict the movement of the lane, which suggests a U-shaped curved guide line to optimize learning by increasing grazing angles.	CNN	-
CLRNet [73]	2022	CVPR	ROIGather can be used to improve lane feature representation by creating relations with all pixels, thereby addressing the lack of visual evidence for the presence of a lane. Furthermore, it is suggested to regress the lane as a holistic unit by adopting the Line IoU loss specifically designed for lane detection tasks.	CNN	link
CondLSTR [74]	2023	ICCV	Using a transformer, dynamic convolution kernels are generated for each lane line in the input image. These convolution kernels allow lane lines to be detected by dynamic convolution, efficiently handling occlusion and lane lines with complex topology.	Transformer	link
CLRerNet [75]	2024	WACV	LaneIoU which is more relevant to the metric by considering local lane angles is proposed.	CNN	link
SRLane [76]	2024	CV*	The advantages of both keypoint-based and suggestion-based methods are utilized. In the “Sketch” phase, the local orientation of the key points can be easily estimated using a fast convolutional layer. In the “Refine” phase, further optimization is performed using a lane segment correlation module.	CNN	link

the lanes by finding the point at which the straight and the lane have a unique intersection. Line-CNN [61] proposes a novel Line Proposal Unit (LPU), LPU can locate accurate traffic curves using line proposals as a reference, enabling the system to learn a global feature representation of the entire traffic route. Su et al. [77] proposed a top-down vanishing point-guided anchoring mechanism for generating dense anchors to efficiently capture various lanes. SIIC-Net [62] introduces the concept of feature lanes, which generates a set of candidate lanes by clustering training in the feature lane space, and then detects road lanes among the candidate lanes. ADNet [63] employs SPGU to generate high-quality anchors, utilizes ALAU to improve lane line feature representation, and utilizes General Lane IoU loss (GLIoU loss) to overcome the shortcomings of Line IoU loss (LIoU loss). As shown in Fig. 16. Additionally, the classical YOLO algorithm can be applied in target detection. For example, Cui et al. [78] proposed an advanced YOLOv3 algorithm for highway lane detection. The algorithm involves refining Anchor parameters with the K-means++ algorithm, extracting features with the Darknet-53 network, and finally, using the upgraded YOLOv3-K-101 network for feature splicing. The novel O2SFormer [64] dynamic anchor-based positional query explores explicit positional prior, employing lane anchors as positional queries and updating these anchors gradually, layer by layer. This strategy enables the detector to resolve semantic conflicts of tags while maintaining end-to-end detection, as shown in Fig. 17. Sparse Laneformer [65] uses a sparse anchor mechanism and no longer relies on dense anchors. That allows the model to be more flexible and not be limited by the density and distribution of the training dataset. It further optimizes the lane prediction results and improves the detection accuracy

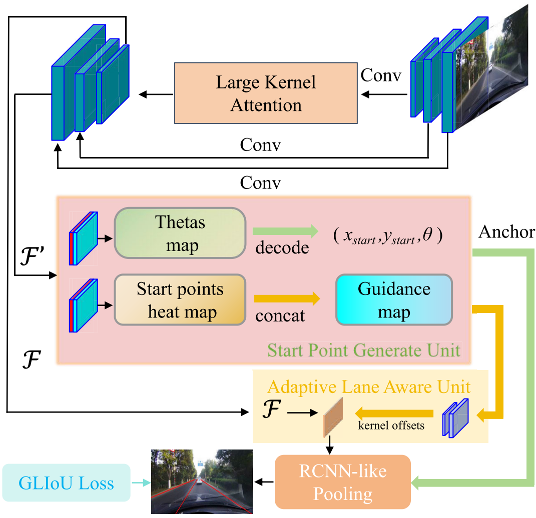


Fig. 16. Overview of ADNet Architecture [63]. Encoder extracts and enhances the lane context with FPN embedded with LKA. Low-level context is transferred to SPGU for the generation of start point guided anchors and guidance map. High-level context is aggregated through ALAU with the aid of the auxiliary guidance map. Pooling optimises the lane lines with the General Lane IoU loss.

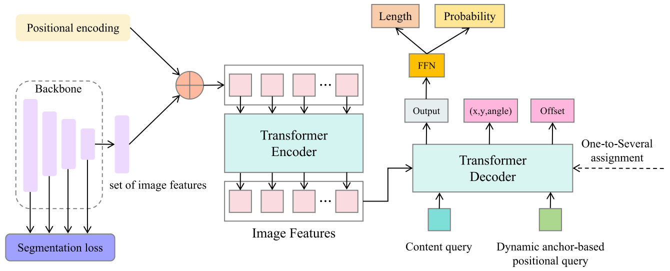


Fig. 17. Structure of O2SFormer [64] pipeline. O2SFormer uses a CNN for 2D feature representation learning and adding positional encoding. The feature representation is passed to the Transformer Encoder where Dynamic anchor-based positional query and content query are taken into account. The output embedding of the Transformer decoder is used by a FFN to predict background and foreground probabilities and lane anchor length. Labels are assigned to lane anchors in each decoder layer using One-to-Several assignment.

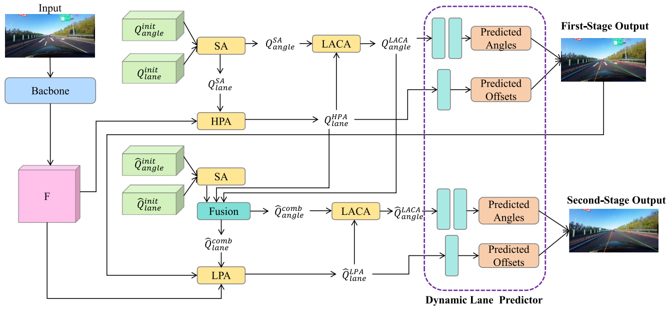


Fig. 18. Overview of Sparse Laneformer [65]. It uses a CNN backbone to extract features from images and then feeds them into a transformer decoder for lane predictions. The decoder includes a two-stage process: an initial stage with attention for learning initial queries and predicting coarse lanes, and a second stage for refining the queries and lanes to generate final predictions.

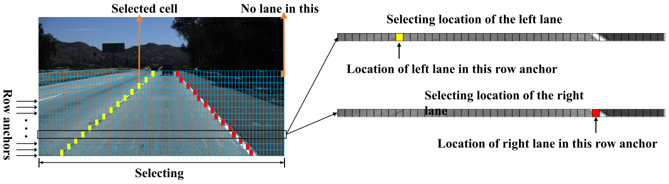

Selected cell
No lane in this
Row anchors
Selecting
Selecting location of the left lane
Location of left lane in this row anchor
Selecting location of the right lane
Location of right lane in this row anchor

Fig. 19. A schematic for selecting on the left and right lane [66].

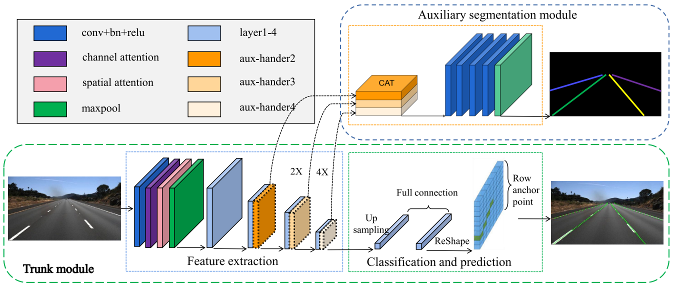


Fig. 20. Unmanned geographic information sensing using hybrid attention mechanism network architecture [79]. It mainly consists three modules: feature extraction module based on mixed-attention mechanism ResNet, auxiliary segmentation module and classification module based on row anchor.

by using not only Horizontal Perception Attention (HPA) and Lane-Angle Cross Attention (LACA) but also the Lane Perception Attention (LPA) mechanism based on deformable cross attention, as shown in Fig. 18.

In addition to the above methods, there exist lane detection methods utilizing a line-by-line search, which means examining the lane pixels in each image line to identify the location and shape of the lane. As shown in Fig. 19. During this process, the algorithm systematically examines each line of the image from the beginning, processing and analyzing each line of pixels in search of the lane's presence and location. Typically, certain predetermined rules, filters, or features are applied to distinguish which pixels may be part of the lane. Following this, the overall lane's contour is gradually approximated based on this accumulated information.

For example, UFLD [66] approaches lane detection as a row selection problem relying on global features, which significantly reduces computation requirements and enhances computational speed. The lightweight version can achieve over 300 frames per second, even in challenging scenarios. The system's performance is further improved by incorporating a large receiver domain for processing global features. Song et al. [79] used a line-direction-based lane line location selection and classification method to detect the presence of a lane at each candidate point based on line anchors, which reduces the high computational complexity associated with the pixel-by-pixel segmentation of traditional semantic segmentation. As shown in Fig. 20. E2E [67] uses a module for efficient horizontal reduction to model each lane marker categorically and obtain its vertices. CurveLane-NAS [68] introduces Neural Architecture Search (NAS) to capture coherent and accurate short-range curve information over long distances. LaneATT [70] extracts features from each anchor using feature mapping. It subsequently combines local and global features to facilitate the use of information from other lanes for accurate lane classification and localization from dense line anchors. CondLaneNet [69] architecture introduces a novel top-down lane detection framework. It first identifies lane instances and then predicts the shape of each instance. A conditional lane detection strategy, integrating conditional convolution and a line-by-line formulation, is proposed to address the lane instance-level recognition problem. UFLDv2 [71] adopts a mixed-anchor approach (row-anchor+column-anchor), which enhances the UFLD row-anchor and efficiently reduces the

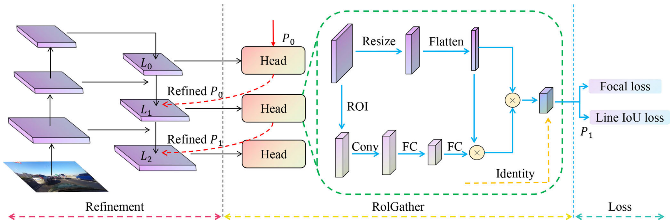


Fig. 21. Overview of the CLRNet framework [73]. Firstly, the network generates feature maps from FPN structure. Subsequently, each lane prior will be refined from high-level features to low-level features. What's more, each head will exploit more contextual information for lane prior features. Finally, classification and regression of lane priors.

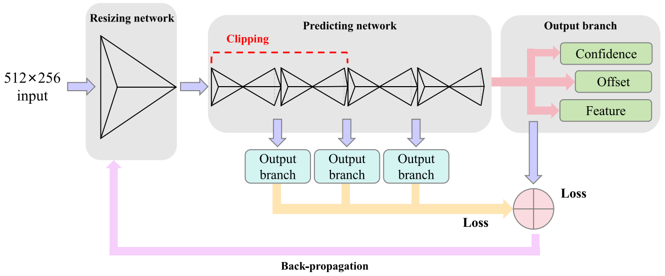


Fig. 22. PINet [80] framework structure. Three-part framework uses resizing network to compress 512 × 256 input data, then feeds compressed data into a predicting network with four hourglass modules for confidence, offset, and embedding predictions. Loss is calculated from each hourglass block output, and clipping modules can adjust resources.

positioning error. However, UFLDv2 does not consider the effect of tilt on losses. To address this issue, CANet [72] uses a heat map form to supervise the construction of a Gaussian distribution centered on the vector line, simulating the modeling of positive sample features with decreasing distance. Specifically, CANet has the capacity to choose a row-wise or column-wise classification, including two types of anchors, all of which are flexible and dynamic to the shape of the instance heatmap.

Moreover, CLRNet [73] uses high-level features to identify lanes and low-level features to adjust the position of the lane. It employs the RoIGather module to gather global semantic information for identification and the proposed Line IoU loss to optimize the lane as a whole. As shown in Fig. 21. On this basis, Honda and Uchida et al. [75] put forward a novel representation, LaneIoU, which incorporates local lane angles into the calculation of the lane. In addition, they created the CLRerNet detector, which enhances the reliability of the confidence scores by optimizing the objective assignment cost and loss function. Chen et al. [76] propose a novel paradigm called "Sketch-and-Refine" that integrates the strengths of both keypoint-based and proposal-based methods. This method, named SRLane, simplifies the lane's local direction and makes it more explicit. It first roughly sketches the lane's shape using local geometric descriptors and then refines it gradually for lane detection.

3) Keypoint-Based Methods: The method estimates lanes by detecting specific key points. Firstly, feature extraction is used to identify important key points from the input image which are invariant and can be recognized across different images. Then, these feature points are matched with corresponding points in different frames to establish inter-frame correspondence. Finally, lanes are reconstructed based on the matched feature points. We summarise several 2D lane detection methods using key points, as shown in Tab. III. Moreover, the specifics are described in the following.

Just as in Fig. 22, PINet [80] boasts triple the confidence, offset information, and embedded features. These features are instrumental for lane key point localization and post-processing key point classification. The confidence and offset information, in particular, contribute to enhancing the accuracy of this process. Reference [81] introduces deep reinforcement learning [82] into lane detection, combining bounding box level convolution neural network lane detector and Deep Q-Learning Localizer (DQLL) to improve the representation of curved lanes [83]. FOLOLane [84] employs two branches, one of which generates a heat map for detecting whether a pixel is a critical point, while the other branch provides offsets for accurately adjusting the position of critical points. The output network completes the local-to-global curve correlation through the correlation algorithm to form multiple complete curves. This method reduces the computational cost while avoiding the problem of noisy and redundant information in the Segmentation strategy. Like Fig. 23, GANet [85] utilizes a Lane-aware Feature Aggregator (LFA) module, a feature aggregator that integrates the global correlation of key points, to model lanes on a full scale. This module is adept at predicting offsets between key points and subsequently aggregating the features of neighboring key points, significantly enhancing the feature representation of the current key point. By integrating local associations with global associations, GANet substantially improves lane detection performance, offering a more comprehensive set of features. RCLane [86] can capture global and local position data for lanes. Specifically, it captures the distance from each point to the two points prior and subsequent, thereby obtaining a localized understanding of the lane. Subsequently, it monitors the endpoints at both lane ends to determine the lane's overall length. LanePtrNet [87] designs a curve-aware centrality as a key measure and proposes the C-FPS algorithm to extract seed points. Further to generate lane clusters, it uses a simple but effective grouping module with cross-instance attentional voting, thus significantly reducing duplicate grouping results.

This approach is usually robust to a certain degree of viewpoint changes, illumination changes, and partial occlusions. At the same time, due to the feature points having invariance, this method can cope with scene changes and image noise to some extent. In practice, the method is usually used in conjunction with other techniques (e.g., motion estimation, image alignment, etc.) to deal with occlusion, breaks, and high density of lanes.

4) Parametric Curve Based Methods: The method models the shape of the lanes and directly outputs a parametric representation of the lanes. Firstly, feature points are identified and extracted from the image, and then a curve-fitting model is used to obtain lane shape parameters. Based on these parameters, the position and shape of the lane in the image, usually one or more curves, are obtained. Finally, the results are post-processed to remove unreasonable curves, connect broken parts, and smooth the curves. We summarize several

TABLE III

An Overview of Some Typical 2D Lane Detection Methods on Key Points and Parametric Curves, Including the Official Publication Date, Publishing Organization, Advantages of the Method, Network Architecture, and Open-Source Status of the Program. * Denotes Prepublication

Modal	Methods	Year	Journals / Conferences	Strengths	Architecture	Code
Key points	PINet [80]	2021	IEEE TITS	It can be tailored according to the computing power of the target system, and the tailored network can be applied directly and has high performance and a low false alarm rate. In particular, PINet performs well under difficult lighting conditions such as nighttime, shadows, and glar	CNN	link
	FOLOLane [84]	2021	CVPR	Focuses on modeling local patterns and enables the prediction of global structures in a bottom-up manner.	CNN	-
	GANet [85]	2022	CVPR	The first method is to regress key points globally. The proposed LFA enhances the correlation between neighboring key points to complement the local information.	CNN	link
	RCLane [86]	2022	ECCV	Proposing a lane relay chain representation to simultaneously model the global geometry and local position information of lanes.	Transformer	link
	LanePtrNet [87]	2024	CV*	Define lane detection as a process of seed point voting and grouping on ordered sets. Introduce centerness as a fundamental aspect of lane detection. Develop a sampling method to predict high-quality candidates.	CNN	-
Parameter curves	PolyLaneNet [88]	2021	ICPR	Output polynomials representing each lane marker in the image, as well as the domain of these polynomials and a confidence score for each lane.	CNN	link
	LSTR [89]	2021	IEEE WACV	End-to-end lane detector with shape model parameters directly outputted. Shape model reflects road and camera state, improving interpretability. Transformer networks for efficient learning global context to help infer occluded parts and capture slender structures.	Transformer	link
	BézierLaneNet [90]	2022	CVPR	Exploit parametric B´ezier curve due to its stability and high freedom in transformations to overcome existing polynomial curve methods optimization difficulties. Propose deformable convolution-based feature flip fusion to utilize lane symmetry properties in driving scenes.	CNN	link
	DecoupleLane [92]	2023	CV*	The lane detection is decomposed into two parts: curve modeling and ground height regression. By using a single 3D lane detection head and using the 2D lane as a projection in perspective space, the 2D and 3D lane detection tasks are unified.	Transformer	-
	Lane2Seq [91]	2024	CV*	A sequence generation-based lane detection method is proposed to convert lane detection into a sequence generation task and unify various lane detection formats. A new multi-format model tuning based on reinforcement learning is proposed to incorporate task-specific knowledge into the lane detection algorithm.	Transformer	-

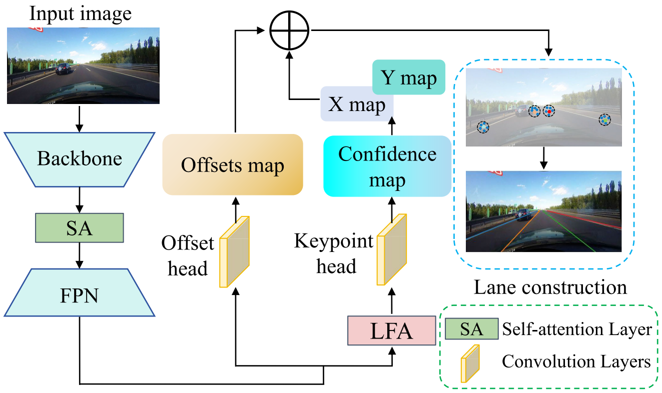


Fig. 23. Overall architectural design of GANet [85]. CNN backbone, SA, and FPN extract multi-scale visual features for the input image. The decoder generates a confidence map and offset map from the key header and offset header, respectively, and clusters key points into groups representing lane line instances. The LFA module is applied to capture local context before keypoint estimation.

methods of 2D lane detection using parametric curves, as listed in Tab. III. The specific parameters for each method are detailed below.

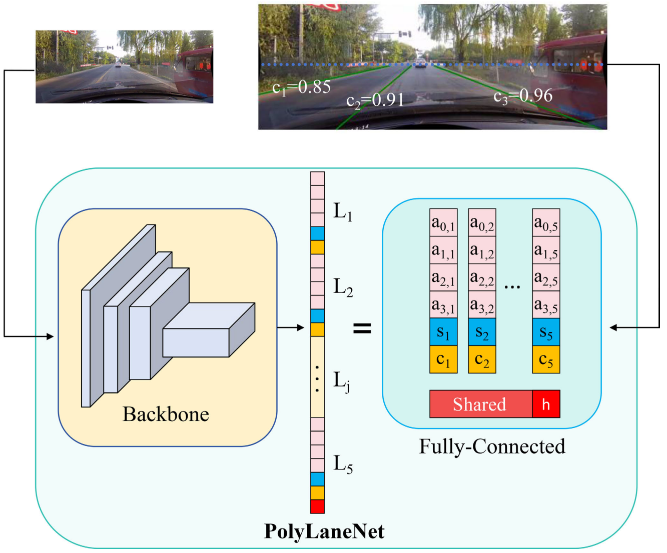

c1=0.85
c2=0.91
c3=0.96
Backbone
L1
L2
Lj
L5
=
a0,1 a1,1 a2,1 a3,1 s1 c1
a0,2 a1,2 a2,2 a3,2 s2 c2
...
a0,5 a1,5 a2,5 a3,5 s5 c5
Shared
h
Fully-Connected
PolyLaneNet

Fig. 24. Overview of PolyLaneNet method [88]. The left to right sequence involves the model inputting an image from a forward-looking camera and outputting information about each lane marking.

As shown in Fig. 24, PolyLaneNet [88] produces a polynomial curve representing each lane in the image and a corresponding confidence score. LSTR [89] adopts a transformer-based network to identify the elongated structure of lanes within a global context by exploiting non-local interactions. Subsequently, the network parameters are directly employed as regression outputs to reflect the road structure and camera pose. BézierLaneNet [90] uses a deep lane detector

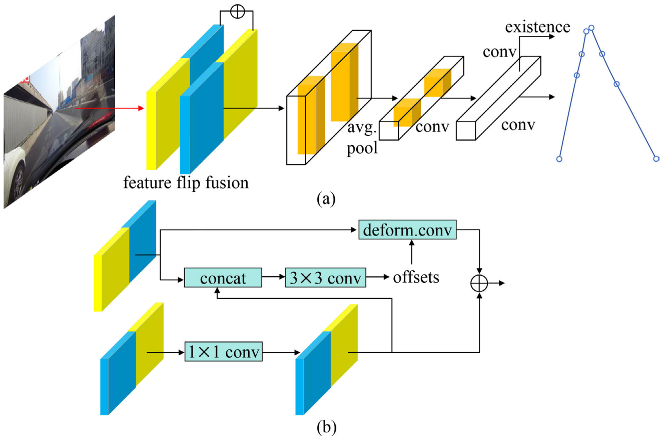


Fig. 25. BézierLaneNet [90] Pipeline. (a) A feature from a typical encoder is fortified through feature flip fusion, subsequently aggregated into 1D feature. Following this, two 1D convolution layers are applied to refine the extracted features. Ultimately, the network predicts Bezier curves through a classification branch and a regression branch. (b) Feature flip fusion. Aligned is achieved by implementing deformable convolution offsets, conditioned on both the flipped and original feature map. This process is most effectively observed when presented in color.

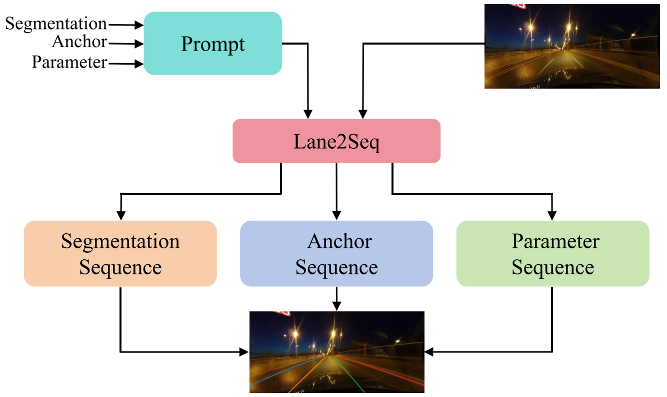


Fig. 26. Lane2Seq [91] interference pipeline. The model senses the images and cues and generates format specific markers. Then it is transformed into the detection format required by Segmentation, Anchor, and Parameter for visualization.

based on Bézier curves that efficiently models the geometry of the lanes and uses parametric Bézier curves to deal with the optimization difficulties of existing polynomial curve methods. The overall model architecture is shown in Fig. 25. Lane2Seq [91] treats lane detection as a sequence generation task, shown in Fig. 26. It employs a simple transformer-based coder-decoder architecture and uses multi-format model tuning based on Reinforcement Learning (MFRL) to integrate task-specific knowledge, thus unifying various lane detection formats.

This approach allows different curve models to be selected depending on the specific problem, for example, a straight-line model may be more appropriate on a highway, whereas a more complex curve model may be required at a curve. Parametric curve-based lane detection methods are usually well adapted to lane shapes in different scenarios and have some mathematical interpretations. Despite its potential benefits, there are certain drawbacks associated with this method. Chief among these are its susceptibility to noise and occlusion, the sensitivity to parameter selection, and the extensive computational requirements for accurate curve fitting.

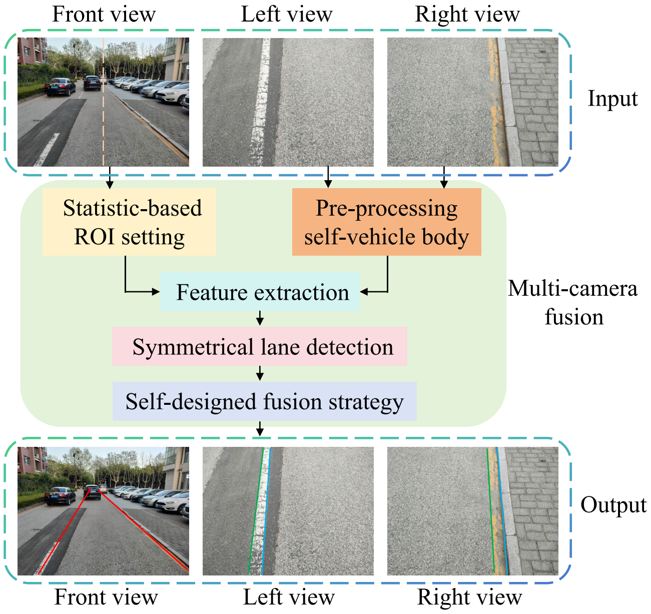

Front view
Left view
Right view
Input
Statistic-based ROI setting
Pre-processing self-vehicle body
Feature extraction
Symmetrical lane detection
Self-designed fusion strategy
Multi-camera fusion
Output
Front view
Left view
Right view

Fig. 27. Multi-camera fusion method architecture for lane detection [93]. The inputs are front/left/right-view images with front-view divided into left/right side. The front view gets ROI setting while side-view undergoes self-vehicle body parts cropping. After feature extraction, lane detection is done for left/right front-/top-view images, and fusion strategy is applied.

C. Other Methods

Researchers also have used multiple cameras to acquire images. Multiple cameras acquire road information from different angles and locations. Compared to single cameras, multi-cameras can provide a more comprehensive view and capture more complex scene features. For instance, Van et al. [94] utilized two cameras, which were discreetly placed under wing mirrors on each side of the vehicle to acquire images. The system then used the EDLines algorithm in real time to detect the line segments. Following this, lane detection was developed by analyzing the angle of these detected line segments. As shown in Fig. 27, Xiong et al. [93] uses a lane detection fusion strategy. It based on vanishing point estimation and specified feature fitting after coordinate transformation using three cameras and a reference vehicle. Zhang et al. [95] proposes a sliding window lane line detection algorithm that combines steering wheel angle data and a binocular camera. The algorithm has two key components: first, it uses the previous steering wheel angle to calculate the radius of the current curve; second, it determines the ratio of pixel distance in the aerial view to real-world distance in the global coordinate system. Yuan et al. [96] multi-stage approach for ground coordinate lane detection and estimation based on vehicle surround view camera module (SVS). Firstly, the raw images are processed using a neural network classifier to generate pixel-level semantic segments. Then, the semantic data labeled lanes are projected to 2D ground coordinates and quantified using a grid representation. Next, outliers are processed and smoothed by spatial and temporal filters. Finally, adaptive polynomial fitting and error statistical analysis are applied to reveal the lane shape and self-vehicle direction information contained in the processed semantic data.

In addition to vision methods, LiDAR is also used for lane detection tasks. 2D LiDAR lane detection is a technique that uses 2D LiDAR sensors to acquire information about the surrounding environment to identify and locate lane lines. By emitting a laser beam and measuring its return time, LiDAR can generate highly accurate point cloud data that reflects the distance and shape of surrounding objects. Han et al. [97] extracts road features as line segments in polar coordinates relative to the LiDAR sensor. These extracted features are then tracked about the vehicle's local coordinates using a nearest-neighbor filter. Wu et al. [98] identifies road boundaries up to 65 m from the LiDAR sensor, constrained by road geometry and the sensor's scan range. To extend this detection range, multiple LiDAR sensors can be placed along the roadside, compensating for point loss at greater distances and allowing for the detection of both road borders. Reference [99] proposed a 2D LiDAR-based road and road edge detection method for recognizing road areas and road edges. Reference [100] proposed a method for recognizing road boundaries and obstacle detection using downward-looking 2D LiDAR, which can detect drivable areas and obstacles on road boundaries, such as curbs, bushes, traffic cones, and vehicles. Unfortunately, it was not able to recognize lanes.

Given the limitations of 2D LiDAR alone, the researchers decided to integrate vision and LiDAR to develop an effective fusion strategy. Li et al. [14] presents a real-time optimal drivable-region and lane detection system that integrates LIDAR and vision data through a multisensory approach, enabling identification of drivable areas via feature-level fusion and an optimal selection strategy for conditional lane detection based on automatic region classification. Caltagirone et al. [119] presents a deep learning approach for road detection by fusing LiDAR point clouds with camera images. Initially, an unstructured and sparse point cloud is projected onto the camera image plane and then upsampled to generate a dense set of 2D images that encode spatial information. Subsequently, multiple fully convolutional networks (FCNs) are trained to carry out the road detection task. Yen?aydin and Schmidt et al. [120] used 2D LiDAR to detect objects on the road, while also utilizing camera data to generate binary bird's-eye view (BEV) [121] images. Subsequently, the BEV is refined to mitigate noise and estimate the position of the object detected by the LiDAR within the BEV, thus performing lane detection. With the maturity of LiDAR technology, 3D LiDAR is beginning to be widely studied in lane detection and has achieved good results, as described in Section IV.

IV. Overview of 3D Lane Detection

Since 2D lane detection methods can only be performed on a planar viewpoint, they face the challenge of viewpoint variations and occlusions, which often lead to false or missed detections. Especially in complex road conditions and situations with many environmental changes, such as inclement weather or road construction zones, these 2D detection technologies may experience performance degradation due to external interference. To address these concerns, in recent years researchers have proposed a variety of 3D lane detection methods that better understand the three-dimensional spatial layout of lanes to improve lane detection accuracy and robustness. It is possible to detect not only the position of the lane in the horizontal direction but also to determine the position of the lane in the vertical direction, thus obtaining the curvature, slope, etc. of the lane. Furthermore, it also works in conjunction with downstream tasks such as path planning and vehicle control, which makes it compatible with additional sensor data such as radar and cameras. In the field of 3D lane detection, the following categories of methods have been developed: 3D LiDAR-based methods, 3D vision-based methods, and multi-modal fusion methods.

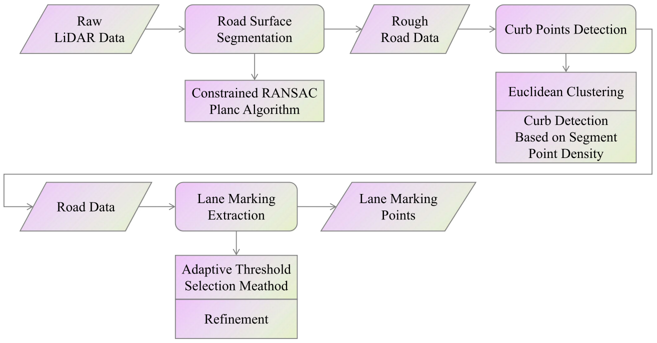


Fig. 28. Workflow of a lane marking detection system using 3D LiDAR [122].

A brief overview of some of the currently available 3D lane detection methods is given in Tab. IV, with details given successively later in the paper.

A. 3D LiDAR Detection

3D Radar, especially 3D LiDAR, can provide extremely accurate and comprehensive three-dimensional spatial data. It can determine the position and shape of surrounding objects by emitting laser pulses and measuring the time of the reflected pulses and provide highly accurate lane positions independent of lighting conditions.

Reference [123] utilizes the LiDAR intensity information to reject the point cloud's lane data beyond a certain threshold, thus separating the LiDAR point cloud into asphalt and road markings, which efficiently identifies any road markings (crosswalks, continuous lines, dashed lines). Veronese et al. [124] combined vehicle motion models and 3D LiDAR sensor data to create a detailed visual map. It used infrared reflectance imagery to gain a more complete understanding of the vehicle's surroundings. Subsequently, the image was processed to extract lane markers, allowing the vehicle to accurately recognize the lane position and width. As shown in Fig. 28, Huang et al. [122] used a 3D point cloud generated by 16-line LiDAR to filter curb points and detect lanes on structured roads. Subsequently, they employed the constrained RANSAC algorithm and a curb detection method that relies on the density of road segment points to successfully segment and refine the road data. Zeng et al. [125] proposed a Scatter Hough algorithm for automated lane detection that considers points in the neighborhood of the estimate. For curve fitting, it uses adaptive line segments to fit the curve, extracts candidate points around the straight line and uses the least squares method to fit the candidate points to obtain the curve parameters. It reduces computational complexity,

TABLE IV

OVERVIEW OF REPRESENTATIVE 3D LANE DETECTION METHODS, INCLUDING THE OFFICIAL PUBLICATION DATE, PUBLISHING ORGANIZATION, ADVANTAGES OF THE METHOD, NETWORK ARCHITECTURE, AND OPEN-SOURCE STATUS OF THE PROGRAM. * DENOTES PREPUBLICATION

Modal	Methods	Year	Journals / Conferences	Strengths	Architecture	Code
Monocular	3DLaneNet [15]	2019	ICCV	The first paper to directly predict 3D lane lines from a single front view camera.	CNN	link
	Gen-LaneNet [16]	2020	ECCV	Decoupling image segmentation and geometry encoding to optimize anchor representation in 3D-LaneNet.	CNN	link
	3D LaneNet+ [101]	2020	CV*	Ability to handle more complex topologies. Uncertainty Prediction.	CNN	-
	3DLaneNAS [102]	2022	ICANN	It introduces NAS to enhance the extraction and combination of visual features by automatically optimizing the monocular 3D lane detection method, thus reducing the computational load.	CNN	link
	CLGo [103]	2022	AAAI	A polynomial is used to model the 3D lanes. Geometric constraints are proposed to help estimate the camera pose.	CNN	link
	SALAD [104]	2022	CVPR	A monocular 3D lane line detection network with an external reference free is proposed, divided into two branches, thus regressing the 3D lane coordinates in the image view.	CNN	link
	PersFormer [17]	2022	ECCV	Generation of BEV features using camera parameters as a reference to the relevant front view local area, and simultaneous detection of 2D/3D lanes.	Transformer	link
	GT [106]	2022	CVPR	A task-specific data enhancement method is proposed to address the problem of uneven data distribution for 3D lane detection. The novel loss function exploits the a priori geometric structure of lanes in 3D space to achieve stable reconstruction from local to global.	CNN	-
	STLane3D [106]	2022	BMVC	It combines spatial and temporal information between frames and proposes 3DLane IOULoss.	Transformer	-
	PETRv2 [107]	2023	ICCV	Multi-task learning approach including 3D object detection, BEV segmentation and 3D lane detection	Transformer	link
	GroupLane [108]	2023	CV*	A set of row-level classification heads is proposed to perform 3D lane detection in BEVs.	CNN	-
	CurveFormer [109]	2023	ICRA	The query in the decoder layer is established as a dynamic anchor set, while the curve cross-attention module is applied to compute the query image similarity. The contextual sampling module is incorporated to predict offsets from the combinations of reference features and queries to assist in the learning of sampling offsets.	Transformer	-
	LATR [110]	2023	CVPR	This approach aims to directly identify and detect 3D lanes from the front view without utilizing any 3D surrogate representations. It features a query generator with lane awareness functionality, as well as an innovative dynamic positional embedding technique that bridges the gap between 3D space and 2D images.	Transformer	link
	Anchor3DLane [111]	2023	CVPR	It directly defines anchors in 3D space and regresses 3D lanes directly from FV without introducing BEV. A global optimization method is proposed to exploit the equal-width property of lanes for refinement.	CNN	link
	BEV-LaneDet [20]	2023	CVPR	Uniform parameters of virtual cameras are introduced to solve the problem of inconsistent spatial relationships of different vehicle cameras. A simple but efficient 3D lane representation method is proposed, and the keypoint representation is suitable for representing complex and diverse 3D lane structures. A chip-suitable spatial transformation module, spatial transformation pyramid, is proposed for converting multi-scale forward-looking features into BEV features.	CNN	-
	DecoupleLane [92]	2023	CV*	Comprehensive lane detection using geometric information accurately accounts for fluctuations caused by uneven ground surfaces.	CNN	-
	D-3DLD [112]	2023	ICASSP	The current network architecture includes a lane feature encoder, lane depth network, voxel space mapping, and 3D lane regression. Two existing public datasets - LLAMAS-3DLD and Pandaset-3DLD - were chosen for comparison with our proposed approach, as well as 3D-LaneNet and Gen-LaneNet.	CNN	-
	An Efficient Transformer [113]	2023	CV*	A decomposed cross-attention mechanism is proposed that can learn both lane and BEV representations.	Transformer	-
	WS-3D-Lane [114]	2023	ICRA	It provides a weakly supervised 3D lane detection method with only 2D lane labels.	CNN	link
	LS-3DLane [115]	2024	JIPS	It proposes a new loss function that exploits prior knowledge of the geometric constraints of parallelism in 3D space.	CNN	-
	Curveformer++ [116]	2024	CV*	Information from historical frames is utilized through time fusion.	Transformer	-
Multi-Modal	M2 -3DLaneNet [117]	2022	CV*	It explored the possibility of utilizing LiDARs to detect 3D lanes among modern vision-centric solutions. A multimodal framework for accurate 3D lane detection is proposed that effectively utilizes complementary features from camera images and LiDAR point clouds to detect 3D lanes.	CNN	-
	DV-3DLane-Tiny [118]	2023	ICLR	It proposes BFF strategies to mutually merge features from different modalities and designs UQGs to merge lane-aware queries from dual views to generate a unified query set. A 3D dual-view deformation attention mechanism is introduced to efficiently aggregate dual-view features.	Transformer	link

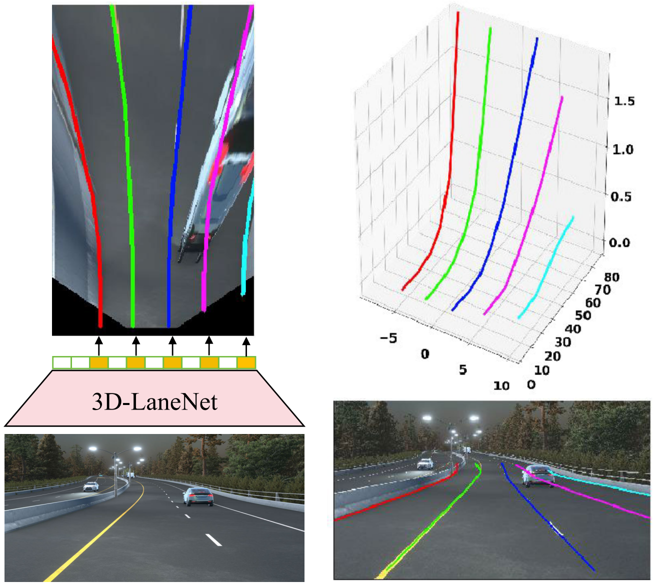


Fig. 29. 3DLaneNet [15] end-to-end approach. Left: Output shown in a top-view perspective. Top-right: Result portrayed in 3D. Bottom-right: Result projected onto the original input image.

enabling better detection of lanes on straight or curved lines given noisy LiDAR data.

However, as the research proceeds, LiDAR alone is becoming less and less of a way to detect lanes. Although it has the advantages of high accuracy, all-weather work, 3D information acquisition, and less susceptibility to occlusion, it also has the disadvantages of higher cost, complex data processing, and dependence on surface materials. Therefore, the current use of LiDAR for lane detection typically works in conjunction with a vision camera.

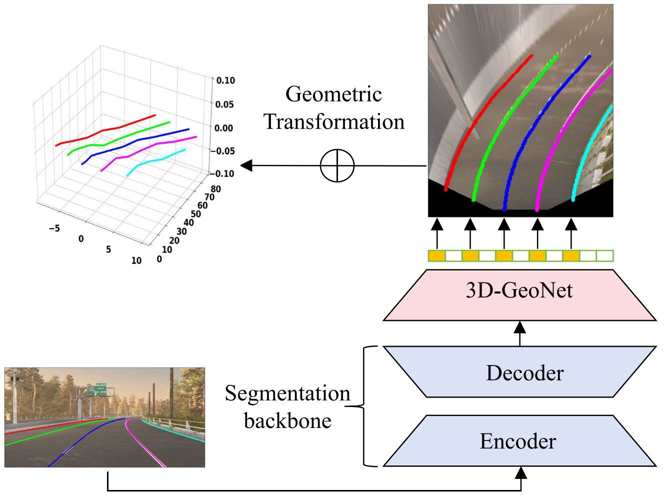


Fig. 30. Overview of Gen-LaneNet [16] method. The segmentation backbone decodes an image into a lane segmentation map, and 3D-GeoNet predicts 3D lane points, specifically represented in top-view 2D coordinates and real heights. At last, the geometric transformation converts the network output to real-world 3D points.

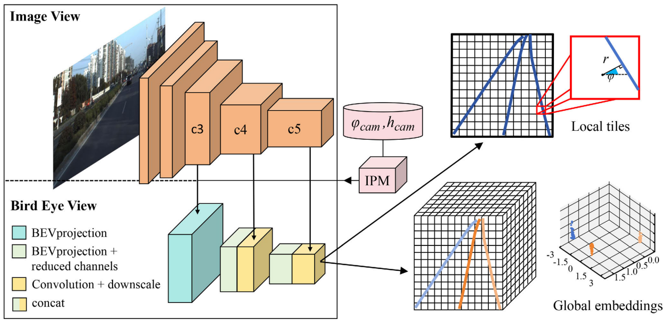


Fig. 31. 3D-LaneNet+ method overview [101] method. The network is comprised of image view and BEV. The last decimated BEV feature map is fed into the lane prediction head, which subsequently produces local lane segments and global embeddings for clustering the segments into entire lane curves.

B. 3D Vision Detection

1) BEV-Based Monocular Vision Inspection: Inspired by recent advancements in monocular 3D target detection methods, researchers have started exploring methods for performing 3D lane prediction directly from monocular images. As shown in Fig. 29, 3D-LaneNet [15] proposes an anchor-based 3D lane representation and projects 2D image features into the Bird-Eye View (BEV) space via Inverse Perspective Mapping (IPM) inside the network. Gen-LaneNet [16] introduces new geometrically guided lane anchor representations in a virtual top-down coordinate system to compute 3D lanes directly from the network output. A scalable two-stage framework is additionally utilized to decouple the learning of image segmentation and 3D lane prediction, as shown in Fig. 30. Compared to 3D-LaneNet, Gen-LaneNet significantly reduces the number of 3D lane labels required to achieve a robust solution in real-world applications. Although this new geometrically guided lane anchor of Gen-LaneNet is more generalizable for unobserved scenarios, it is still limited to long lanes roughly parallel to the direction of self-vehicle travel. 3D-LaneNet+ [101] is an extended 3D lane detection framework that builds on the original 3D-LaneNet to support the detection of more complex lane topologies such as short lanes, vertical lanes, splits, and merges. Just as in Fig. 31, the framework employs an unanchored semi-local representation for capturing geometric features of lane segments, while these are later globally embedded into the complete lane curve to improve detection performance in challenging situations. 3DLaneNAS [102] utilizes Neural Architecture Search (NAS), using multi-objective simulated annealing as the search method. The architecture of the feature extraction and feature fusion modules was finally optimized to improve the accuracy of monocular 3D lane detection in both near and long-distance scenes. In addition, a transfer learning mechanism is integrated to enhance the pace of the search process and improve its accuracy.

In addition to this, Jin et al. [126] introduced the novel attention module of Dual Attention (DA) to lane detection. The attention mechanism employs dual-pathway correlated attention to generate additional features and aggregate information, thus enabling the model to perform robustly and accurately under complex conditions. Efrat et al. [127] proposed a semi-local, BEV, and slice representation to decompose lanes into simple lane segments. It combines learning from parametric modeling of lane segments with deep feature embedding and then clustering the segments into complete lanes. This combination accommodates complex lane topologies, curvatures, and pavement geometries.

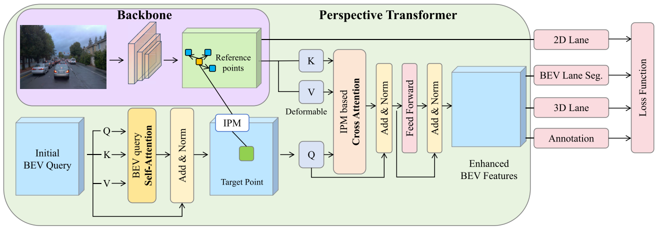


Fig. 32. PersFormer pipeline [17]. The goal is to learn a spatial feature transformation from front view to BEV space to better represent BEV features at target points by considering local context around reference points. PersFormer consists of self-attention and cross-attention modules that interact with BEV queries and reference feature to generate fine-grained BEV features.

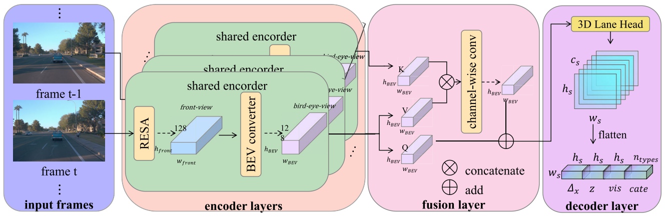


Fig. 33. The architecture of STLane3D [106]. It can be simplified into three main parts: encoder layer, fusion layer, and decoder layer.

2) Improved BEV Detection: However, all of the above 3D detection methods rely on IPM to map image features to BEV features, and IPM causes distortion when encountering uphill and downhill slopes. In response, CLGo [103] better predicts camera pitch angle and mounting height by using a two-stage frame. WS-3D-Lane [114] indirectly supervises the 3D lane heights in the training data by assuming that neighboring lanes have constant widths and equal heights. Meanwhile, it proposes a camera pitch angle self-calibration method to cope with the dynamic changes during data acquisition. PersFormer [17] uses a deformable attention mechanism to iteratively update the BEV features, helping to mitigate the differences introduced by IPM. At the same time using known camera poses to build dense BEVs to look up the query, thus unifying 2D and 3D lane detection under one framework. Nevertheless, the PersFormer method has a high computational resource requirement and a long convergence time. As shown in Fig. 32. To overcome the problems of PersFormer and maintain high performance, Li et al. [128] proposed an approach that combines a perspective converter with MobileNet. The consumption of computational resources is reduced by the spatial feature extraction method provided by MobileNet, while in Perspective Transformer, the spatial feature transformation module further refines high-quality BEV features by generating multi-scale front-end view features. As shown in Fig. 33, STLane3D [106] enables a multi-frame fusion mechanism that harnesses the robust spatiotemporal continuity of successive frames rather than concentrating on a single shot. It first employs a pre-alignment process to align the spatial cues between different frames and then introduces an attention module to fuse the BEV features derived from various temporal stages. As a result, a more comprehensive and continuous comprehension of lane attributes can be achieved.

Furthermore, recent years have seen the development of novel and refined methods. DecoupleLane [92] is a cutting-edge lane detection solution that integrates curve modeling and ground height regression. It employs parametric curves to depict the lanes, preserving their original distribution characteristics and employs ground height regression to address ground height variations due to road conditions and other factors. Moreover, DecoupleLane introduces a novel framework and loss function that unifies 2D and 3D lane detection, enabling the creation of optimized models with or without 3D labels. D-3DLD [112] utilizes voxel mapping with depth awareness to extend rich contextual features from the image domain to 3D space to determine 3D lanes based on voxelized features. Furthermore, an innovative lane representation incorporating uncertainty has been developed, which allows estimating the uncertainty intervals of 3D lane points by applying Laplace loss. BEV-LaneDet [20] proposes a Spatial Transformation Pyramid, a lightweight base on MLP for transforming scale features from front-view to BEV, which solves the problem of invalidating planar assumptions due to factors such as slope. As shown in Fig. 34. Chen et al. [113] propose an efficient transformer for 3D lane detection. It introduces a decomposed cross-attention mechanism that can simultaneously learn lane and BEV representations, which solves the problems of inaccurate view transformation and cumulative errors that may be caused by IPM in traditional methods. GroupLane [108] uses a row classification strategy to represent lanes, divides the feature mapping into groups, and matches each group to a lane instance for detection. During training, predictions are associated with lane labels, losses are computed using one-to-one matching, and no post-processing operations are required for inference. In this way, GroupLane enables end-to-end detection, as shown in Fig. 35. Yao et al. [129] used a sparse point-guided 3D lane detection that consists of two stages: coarse-level lane detection and iterative fine-level sparse point refinement. The coarse lanes are first computed by establishing a dense but efficient correspondence between the front view and the BEV space, and then the 3D lanes are refined layer by layer from low to high resolution by sparse point refinement to improve the efficiency of the information flow and obtain more accurate results.

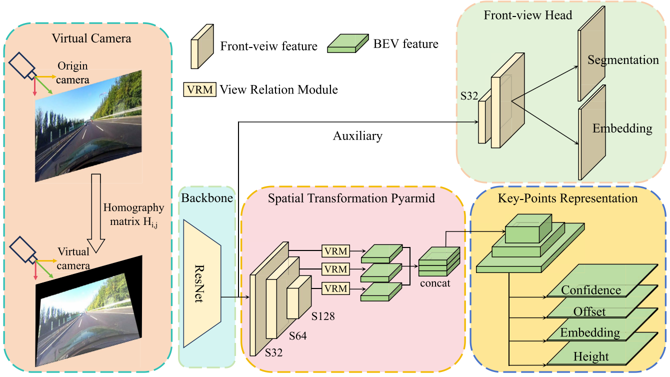


Fig. 34. Schematic of BEV-LaneDet [20] network structure. It consists of five parts: Virtual Camera, Backbone, Spatial Transformation Pyramid, Key-Points Representation, Front-view Head. In which *S32* denotes 32x downsampling of the input image.

PETRv2 [107] is a multi-task learning approach that includes 3D object detection, BEV segmentation, and 3D lane detection. It introduces position embedding transformation to temporal representation learning, which enables pose transformation for temporal alignment through 3D position embedding

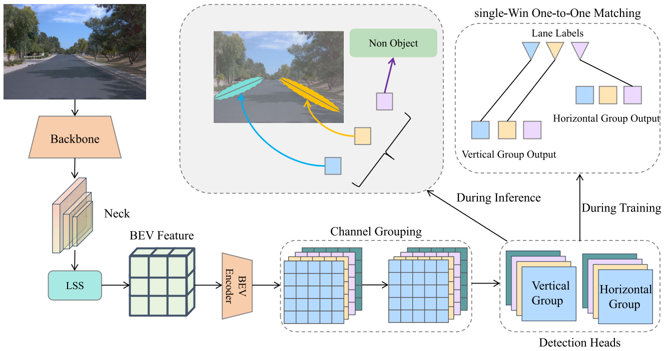


Fig. 35. The overall framework diagram of GroupLane [108]. In this detector, feature maps are split into groups for prediction instance representation. The predictions from detection heads are matched with lane labels using SOM strategy for loss computation. GroupLane generates detection results without post-processing during inference.

(3D PE). A feature-guided position encoder is further proposed to re-weight 3D PE guided by 2D image features. Specifically, the backbone network extracts 2D features from the multiview images and later generates 3D coordinates. To achieve temporal alignment, the 3D coordinates in the PETR of the previous frame t-1 are first transformed by a pose transformation. The 2D image features and 3D coordinates of the two frames are then joined together and injected into the feature-guided position encoder to generate the key and value components of the transformer decoder. The detection, segmentation, and lane queries initialized in different spaces interact with the key and value components in the transformer decoder. The updated queries are further used to predict 3D bounding boxes, BEV segmentation maps, and 3D lanes with task-specific heads.

3) Innovative Detection: The reliance on the flat ground assumption within IPM, combined with the loss of contextual information in the BEV representation, significantly limits its capacity to accurately reconstruct 3D information from the BEV representation. Consequently, a growing number of researchers are currently exploring innovative research breakthroughs that may address these limitations. GT [105] solves the monocular 3D lane detection problem by exploiting the geometric structure under the 2D to 3D lane reconstruction process. It extracts BEV lane information directly from the forward-looking image, which significantly alleviates the confusion of far lane features in previous methods. SALAD [104] can combine the semantic segmentation of 2D lanes and the spatial depth estimation information for 3D reconstruction, to directly obtain the 3D lane positions in the real scene. As shown in Fig. 36.

Huang et al. [111] proposed a BEV-free method called Anchor3DLane to predict 3D lanes directly from FV representations. The 3D lane anchors are projected into the FV elements to extract their features containing good structural and contextual information to make accurate predictions. CurveFormer [109] computes 3D lane parameters directly, avoiding the explicit view conversion step between BEV and front view. It treats the 3D lane detection problem as a curve propagation problem and uses curve queries to represent the 3D lanes. The curve query is expressed by a dynamically ordered set of anchor points and iteratively refined in the Transformer decoder to improve the 3D lane detection results.

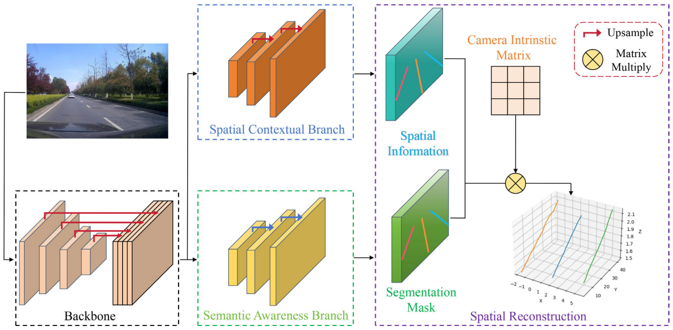


Fig. 36. Schematic diagram of the overall structure of SALAD [104]. The backbone converts an input image into deep features, and two branches decode them to obtain lane spatial info and segmentation mask. 3D reconstruction is then performed, integrating this info to get the 3D lane positions.

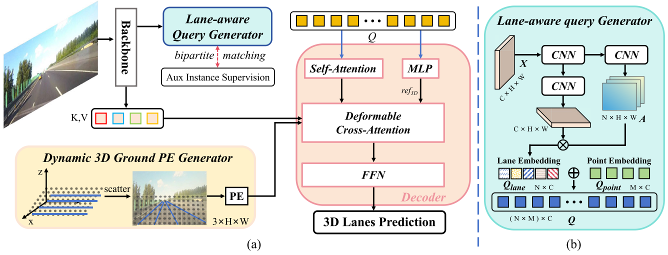


Fig. 37. LATR [110] method overall architecture. LATR is a 3D lane detection framework using a Transformer-based approach. First, the front-view image is processed by the backbone network, as shown in part (a). Then, a Lane-aware Query Generator generates queries using lane-level and point-level embeddings, as illustrated in (b). Dynamic 3D ground positional embeddings are obtained through iterative refinement of a 3D ground plane to capture 3D information.

CurveFormer++ [116] builds on CurveFormerby employing contextual sampling and anchor point constraints to compute curve query-related image features to handle different lane lengths. In addition, it employs a temporal fusion module that combines selected information sparse curve queries and their corresponding anchor point sets to utilize historical lane information. As illustrated in Fig. 37, LATR [110] detects 3D lanes by cross-attention of queries and key-value pairs, using the lane-aware query generator and dynamic 3D ground position embedding constructs.

C. Multi-Modal 3D Lane Detection

In complex traffic environments, monocular 3D detection may be interfered with by light variations, occlusions, and reflections, resulting in inaccurate detection results. Multimodal 3D lane detection can utilize complementary information from multiple sensors to improve the detection of lanes in different environmental conditions, such as multiple cameras, cameras + LiDAR. It can also fuse the data acquired by sensors to provide more sensory information and reduce false detections due to data noise or occlusion from a particular sensor, thus improving the accuracy and robustness of the detection. By combining data from different sensors, real-world lane information can be more accurately reproduced, including lane location, shape, and size.

Bai et al. [130] present a new deep neural network that utilizes both LiDAR and camera sensors and produces very accurate estimates of the complete network architecture directly in 3D space. Specifically, the system takes

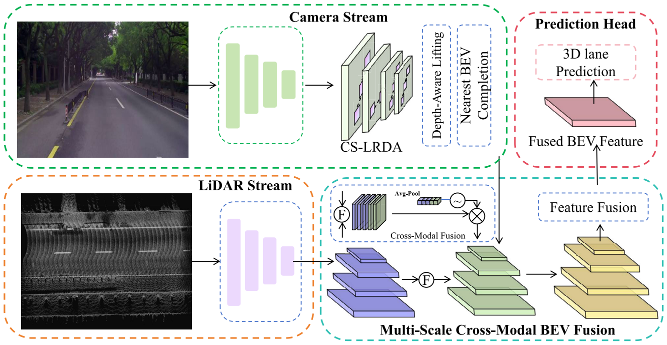


Fig. 38. Overview of M^2-3DLaneNet framework [117]. It's implemented by utilizing both an image and a LiDAR point cloud as inputs, where features from the image are obtained via top-down BEV generation. Subsequently, two BEV features are integrated through bottom-up BEV fusion, which results in the creation of a fused BEV feature. This fused BEV feature is then utilized for the prediction of 3D lanes.

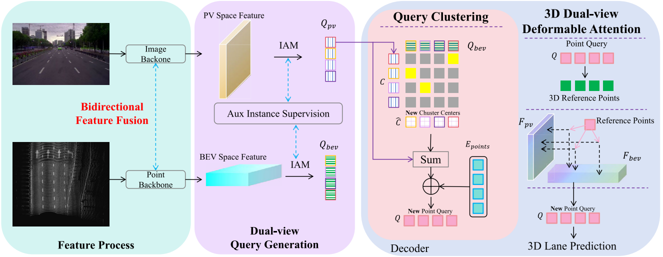


Fig. 39. The overall framework of DV-3DLane [118]. Images and point clouds undergo separate processing in respective backbone. Bidirectional Feature Fusion (BFF) is introduced for fusion of multi-modal features across views. IAM is employed to form lane-aware queries $\mathbf{Q}_{pv}$ and $\mathbf{Q}_{bev}$, which are then aggregated into $\mathbf{C}$ through Dual-view Query Clustering augmented with Epoints to form $\mathbf{Q}$. 3D Dual-view Deformable Attention is introduced to consistently aggregate point features in $\mathbf{Q}$.

input information extracted from LiDAR scans and predicts a dense ground height. Then, it combines the input RGB camera images, which are projected onto the dense ground surface and combined with the LiDAR information to produce lane detection in the top view 3D view. As shown in Fig. 38, M^2-3DLaneNet [117] enhances two-dimensional features into three-dimensional space by fusing geometric information derived from LiDAR data. Consequently, the LiDAR features are further enhanced using a boosted two-dimensional feature set through cross-modal BEV fusion. DV-3DLane [118] achieves accurate 3D lane detection by fusing image and LiDAR data in both PV and BEV views, using innovative feature fusion and query generation methods, and a 3D dual-view deformable attention mechanism. As illustrated in Fig. 39.

With the increasing popularity of autonomous driving, the challenges associated with lane detection have also grown. In the complex road environment, there is an escalating demand for improved lane detection capabilities, and relying solely on traditional sensors or a single perception technology can no longer fulfill these needs. Instead, multimodal fusion sensing is a promising approach that has the potential to meet the ever-evolving requirements. In certain weather conditions, such as when a vehicle is facing glare, driver vision may be disrupted, potentially compromising vehicle safety. To counter the issue, some drivers choose to wear sunglasses to alleviate this effect. However, it should be noted that placing sunglasses over the camera lens may obstruct some light, which may negatively affect the camera's image quality. For these reasons and taking into account actual vehicle configurations, multimodal lane detection can fuse data from a variety of sensors, such as cameras, LiDAR, and millimeter wave radar, to achieve more accurate lane detection. Furthermore, the recent advent of high-precision maps has introduced an element of predefined information, such as road structure and lane markings. When coupled with real-time sensor data, this aids vehicles in more precisely identifying and comprehending their immediate surroundings. Furthermore, the application of such techniques as map matching and lane template matching enhances vehicles' capacity to detect lane boundaries, resulting in enhanced reliability and accuracy.

V. Experiment

A. Datasets

In order to expedite the investigation and assessment of lane detection techniques, researchers have assembled and developed several lane detection datasets. These datasets simulated a variety of scenarios in the real world, including different weather conditions, periods, and road types, to provide rich training and testing samples. The section provides an overview of these datasets, analyzing their characteristics, challenges, and impacts to inform subsequent research.

1) Tusimple Dataset: Tusimple dataset [131] is the only large-scale dataset used to test deep learning methods on the lane detection task before 2018, and many of the lane detection methods are based on it. The basic information of the TuSimple dataset is shown in Tab. V and Tab. VII. Fig. 40 provides six examples of the TuSimple dataset, captured from highway environments under varying weather and congestion conditions.

2) CULane Dataset: The CULane dataset [36] was gathered from cameras mounted on six separate cars, with the road location being Beijing, China. The basic information of the CULane dataset is demonstrated in Tab. V and Tab. VII. The dataset has significant data volume, a wide variety of road scene types, high coverage, and challenging detection difficulty. It has a normal category and eight challenging categories, including congestion, darkness, wireless, and other difficult-to-detect situations, as shown in Fig. 41 (a). Moreover, in its upper left corner, it exhibits the proportion of images corresponding to each scenario type about the overall number of CULane test sets. After the emergence of the dataset, most of the lane detection methods have used it as a target to show their method performance.

3) CurveLanes Dataset: The CurveLanes dataset [68] collects real urban and highway scenarios from several cities in China. Compared with the natural distribution of lanes, CurveLanes has more curves, which account for more than 90% of the dataset. It also has more lanes than CULane and TuSimple, like more than five lanes, so it's more challenging. The CurveLanes dataset provides the essential details in Tab. V and Tab. VII. Some examples of the dataset are shown in Fig. 42.

4) Apollo 3D Synthetic Dataset: Apollo 3D Synthetic dataset [16] is a highly diverse 3D world dataset built through

TABLE V
BASIC INFORMATION OF 2D LANE DATASET

Name	Frame	Train	Validation	Test	Resolution	Year
TuSimple	6408	3268	358	2782	1280×720	2017
CULane	133235	88880	9675	34680	1640×590	2018
CurveLanes	150k	100k	20k	30k	2650×1440	2020

TABLE VI
BASIC INFORMATION OF 3D LANE DATASET

Name	Frame	Train	Validation	Test	Resolution	Year
Apollo 3D Synthetic	10k	-	-	-	1920×1080	2020
ONCE-3DLanes	211k	200k	3k	8k	1920×1020	2022
OpenLane	200K	-	-	-	1920×1280	2022

the Unity game engine. It simulates a variety of visual elements, including highways, cities, homes, and other environments, and renders images with diverse scene structures and visual appearances. The basic information of the Apollo 3D Synthetic dataset is shown in Tab. VI with Tab. VII. In Fig. 43, we show a few examples of the ApolloSim dataset.

5) OpenLane Dataset: OpenLane dataset [17] comes from the public perceptual dataset Waymo Open Dataset's valuable content with lane and the closest-in-path object (CIPO) annotations for 1000 road segments. In more detail, OpenLane contains 200,000 frames, over 880,000 instance-level lanes, 14 lane classes, as well as scene labels and closed-path object annotations. The basic information of the dataset is shown in Tab. VI with Tab. VII. Researchers built Openlane-v2 by adding topology based on it, providing 3D annotations of the lanes. In Fig. 44, we show several scenarios for the Openlane dataset.

6) ONCE-3DLanes Dataset: The ONCE-3DLanes dataset [104] is a collection of real-world autonomous driving data in which lane layout annotations are available. Tab. VI and Tab. VII summarize the essential details of the ONCE-3DLanes dataset. The data was sourced from various geographical locations within China, including highways, bridges, tunnels, suburbs, and city centers. It encompasses a broad spectrum of weather conditions and lighting conditions. Furthermore, this dataset includes ample slope scenarios with varying lighting conditions and multiple lanes. Several examples of the dataset are illustrated in Fig. 45.

7) Other Dataset: In addition to the dataset introduction above, a multitude of additional datasets have been released by researchers, including Caltech Lanes [132], VPGNet [50], VIL-100 [133], BDD100K [134], LLAMS [135], SDLane [62], ApolloScape dataset [136], OpenDenseLane [137], and others. Examples of them are shown in Fig. 46. In Tab.VII, we draw a comparison between several of the existing and extensively utilized datasets.

B. Evaluation Metrics

There are two mainstream evaluation methods in lane detection, which are area-based and distance-based, as shown in Fig. 47.

1) Area-Based Evaluation: As displayed in Fig. 47, the sampled points are extrapolated with a suitable curve to stretch the lanes to a consistent width in the original resolution of the image. A true example is recognized when IoU between the predicted value and the true value exceeds a predetermined threshold. When evaluating the performance, the True Positive (TP), False Positive (FP), and False Negative (FN) cases of lane detection in the test set are respectively counted. The Precision, Recall, and F1 values are calculated as the final evaluation metrics, as shown in Eq. 1.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}, \tag{1}$$

where $Precision = \frac{TP}{TP+FP}$ and $Recall = \frac{TP}{TP+FN}$.

This approach is more concerned with the overall similarity of lane detection, globally requiring that the connectivity of all predicted points be as similar as possible to the true value, with no specific requirements for deviations from individual sampling points. In this paper, the CULane dataset [36] was used for this evaluation methodology, where the lane width of the extension is set to 30 pixels and the threshold of IoU is set to 0.5.

2) Distance-Based Evaluation: The distance-based evaluation is determined by the difference in distance between each pair of sampled points and the true value point, and the difference is less than a certain threshold is determined to be the correctly predicted point, such as the points with spacing d2 d4 in Fig. 47. The ratio of the number of correctly predicted points on average per image to the total number of points is defined as the accuracy(Acc) of lane detection. If the percentage of true points in a lane is more than a certain threshold then that lane is considered a TP, else it is an FP or FN. The final metrics used to evaluate the detection performance are Acc, FP, and FN, which are calculated as shown in Eq. 2.

$$\begin{cases} Accuracy = \dfrac{\sum_{clip} C_{clip}}{\sum_{clip} S_{clip}}, \\ FP = \dfrac{F_{pred}}{N_{pred}}, \\ FN = \dfrac{M_{pred}}{N_{gt}}, \end{cases} \tag{2}$$

where C_{clip} represents the number of accurately predicted lane points and S_{clip} denotes the total number of lane points of a clip. F_{pred} denotes the number of incorrectly predicted lanes, N_{pred} denotes the number of all predicted lanes, M_{pred} denotes the number of missed true lanes, and N_{gt} is the number of all true lanes.

The distance-based evaluation approach puts more emphasis on the accuracy of the sampled points, and only points that fall within a smaller interval around the true value are considered to be correct, i.e., the localization has to be more accurate, reflecting the local accuracy of lane detection. In this paper, we adopt this evaluation approach on the Tusimple dataset [131] and follow the official evaluation metrics. If the distance of a lane point is less than a given threshold $t_{pc} = \frac{20}{cos(a_y l)}$, the lane point is considered to be correct, where $a_y l$ denotes the angle of the corresponding ground truth value. Specifically, the distance difference threshold between the predicted correct point and the true point is set to 20 pixels, and the percentage of predicted correct points on the true example of the lane needs to be greater than 85%.

TABLE VII
BASIC INFORMATION OF LANE DETECTION DATASET

Dataset	Year	Traffic Condition	Geographical Locations	Weather	Environment	Difficulty	#Frames	Line Category	#Segments	Types	Average Length	Max #Lanes	Resolution	Inst. Anno	Track. Anno	Link
Caltech Lanes [132]	2008	Light traffic	highway, urban	Day	Real World	Easy	1224/1224	-	4	2D	-	4	640 × 480	✓	×	link
VPGNet [50]	2017	Multi-traffic	urban	Day&night Multi-weather	Real World	Medium	20K/20K	7	-	2D	-	-	1288 × 728	×	-	link
TuSimple [131]	2017	Light traffic	highway	Day	Real World	Easy	6.4K/128K	-	6.4k	2D	1s	5	1280 × 720	✓	×	link
CULane [36]	2018	Multi-traffic	highway, urban, rural	Day&night Multi-weather	Real World	Medium	133K/133K	-	-	2D	-	4	1640 × 590	✓	-	link
Apollo 3D Synthetic [16]	2020	light traffic	highway, urban, residential	Day&night	Synthetic(Unity 3D)	Easy	10.5K/10.5K	-	-	3D	-	6	1920 × 1080	✓	-	link
VIL-100 [133]	2021	Multi-traffic	highway, urban	Day&night Multi-weather	Real World	Medium	10K/10K	10	100	2D	10s	6	640 × 368 ∼ 1920 × 1080	✓	×	link
LLAMAS [135]	2019	Light traffic	highway	Day	Real World	Easy	79K/100K	-	14	2D	-	4	1276 × 717	✓	×	link
ApolloScape [136]	2019	Multi-traffic	highway,urban, residential	Day&night Multi-weather	Real World	Medium	115K/115K	13	235	2D	16s	-	3384 × 2710	×	×	link
BDD100K [134]	2020	Multi-traffic	highway, urban	Day&night Multi-weather	Real World	Medium	100K/120M	11	100k	2D	40s	-	1280 × 720	×	×	link
CurveLanes [68]	2020	Multi-traffic	highway, urban	Day&night	Real World	Medium	150K/150K	-	-	2D	-	9	2650 × 1440	✓	-	link
SDLane [62]	2022	Light traffic	highway, urban	Day	Real World	Medium	43K/43k	-	-	2D	-	7	1208 × 1920	✓	-	link
ONCE-3DLanes [104]	2022	Multi-traffic	highway,bridge,tunnel, suburb,downtown	Day&night Multi-weather	Real World	Medium	211K/211K	-	-	2D/3D	-	8	1920 × 1020	✓	-	link
OpenLane [17]	2022	Multi-traffic	residential,urban,suburbs, highway,parking lot	Day&night Multi-weather	Real World	Hard	200K/200K	14	1k	2D/3D	20s	24	1920 × 1080	✓	✓	link
OpenDenseLane [137]	2022	Light traffic	highway, urban	Day	Real World	Medium	57K/57K	4	1.7k	3D	-	-	1600 × 1025 (RGB + pointcloud)	✓	×	link

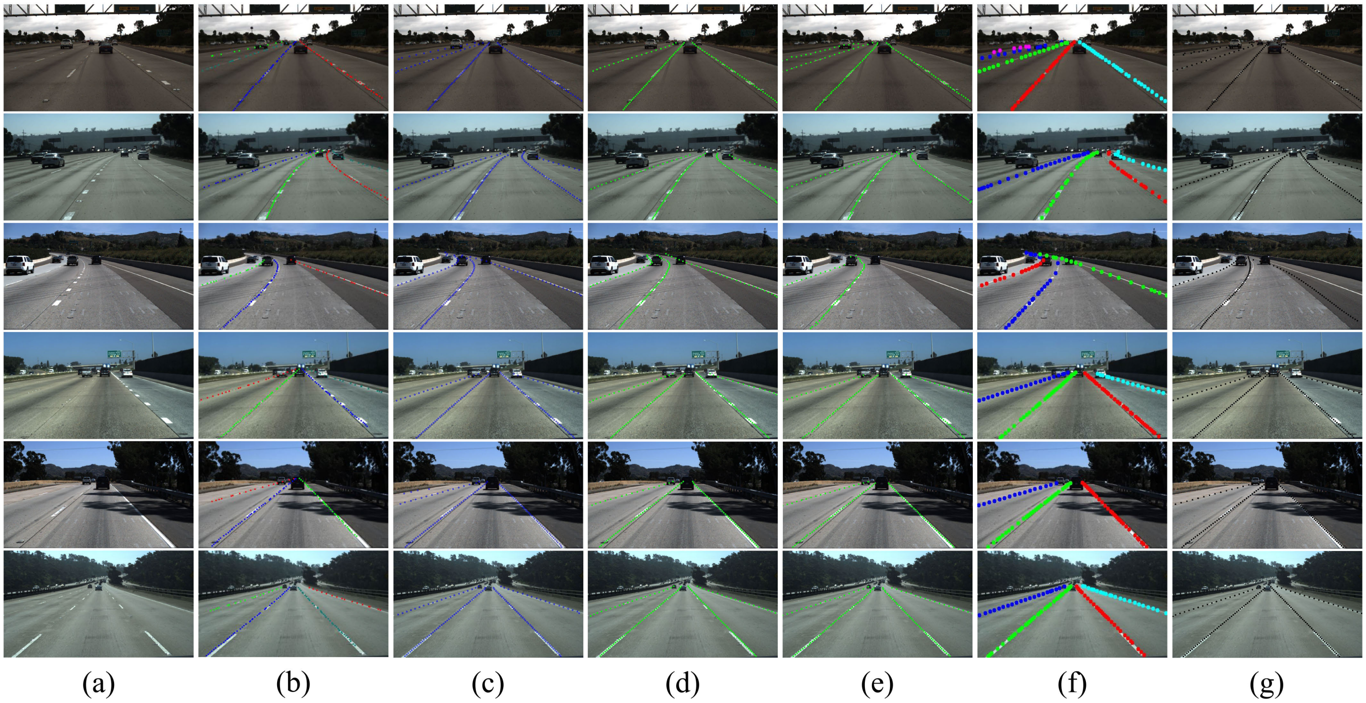


Fig. 40. Visualization results of different methods on the Tusimple dataset. (a) Example of Tusimple dataset [130] with straight and curved lines in different weather and scenarios. (b) Lanenet [35], (c) RESA [37], (d) UFLD [66], (e) UFLDv2 [71], (f) PINet [80], (g) BézierLaneNet [90].

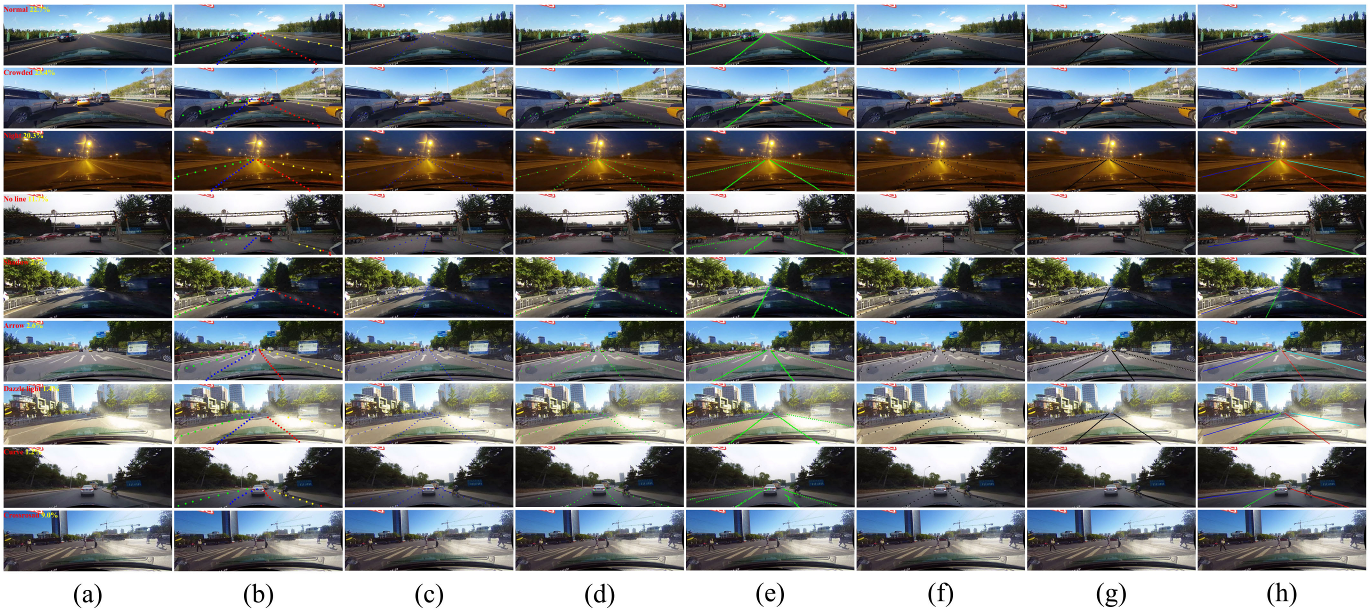


Fig. 41. Visualization results of different methods on the CULane dataset [36]. (a) CULane dataset examples for different scenarios and proportion of each scenario. (b) SCNN [36], (c) RESA [37], (d) UFLD [66], (e) UFLDv2 [71], (f) LSTR [89], (g) BézierLaneNet [90], (h) SRLane [76].

3) 3D Inspection Performance Evaluation: For 3D detection tasks, we adopt the evaluation metrics proposed by Gen-LaneNet [16]. It includes detection accuracy and geometric estimation accuracy. First, detection accuracy was first calculated from standardized average precision (AP) measurements of the precision-recall curve. Following this, the curvilinear distance between the real and detected lanes was calculated as a weighted sum of the Euclidean point-to-point distances. This distance was calculated at predefined y-values placed every 80 cm, extending from 0-80 meters. Performs one-to-one curve matching, selecting pairs of decreasing similarity. If the weighted distance is below some fairly loose threshold (1.5 meters), the match is considered correct. Subsequently, the precision-recall curve is generated by iterating the lane confidence threshold. For the matched detection results, the geometric estimation accuracy was assessed by measuring the error at the same points. As shown in Eq. 3.

$$d^i_{p-t} = \begin{cases} \sqrt{\left(x^i_p v^i_p - x^i_t v^i_t\right)^2 + \left(z^i_p v^i_p - z^i_t v^i_t\right)^2}, & \textit{if } v^i_p = v^i_t, \\ 1.5, & \textit{otherwise}, \end{cases} \tag{3}$$

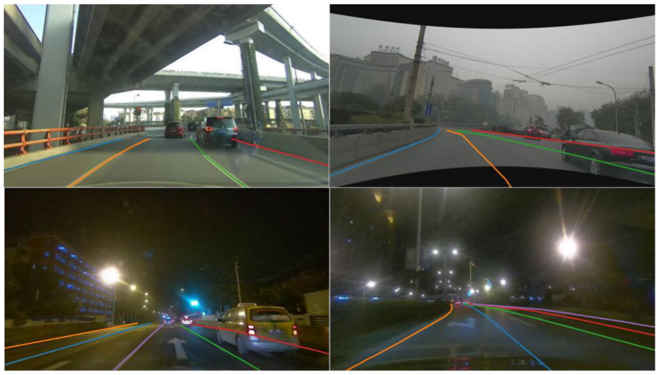


Fig. 42. Example of the CurveLanes dataset [68], which contains complex curved lanes.

where (x_p^i, z_p^i) denotes the 3D position points on a lane derived by model inference, (x_t^i, z_t^i) represents the corresponding point on the label lane, v_p^i and v_t^i represents the corresponding point on the label lane.

For correctly detected lanes, the position error is also evaluated, and the evaluation terms include *Xerror (near)*, *Xerror (far)*, *Zerror (near)*, and *Zerror (far)*. An example of correctly detected lanes and corresponding labels is analyzed, and the error solution formula is shown in Eq. 4.

$$\begin{cases} Xerror = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N}\sqrt{\left(x_i - \hat{x}_i\right)^2}, \\ Zerror = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N}\sqrt{\left(z_i - \hat{z}_i\right)^2}. \end{cases} \tag{4}$$

We investigate F-Score, and error (near and far) of the existing methods on Apollo 3D Lane Synthetic dataset [16], and OpenLane dataset [17] in V-C.

In addition, for ONCE3DLanes dataset [104], SALAD proposes the metric to compute the matching of two lanes in the z-x plane (top view). In this, the lane is represented as $L^k = \left\{(x_i^k, y_i^k, z_i^k)\right\}_{i=1}^n$. To determine whether the predicted lane Lp matches the ground-truth lane Lg, the first matching is first performed in the z-x plane, namely top-view, using the traditional IoU method to determine whether L^p matches L^g. If the IoU is bigger than the IoU threshold, the unilateral chamfer distance (CD) is further used to compute the curve matching error in the camera coordinates. The curve matching error $CD_{p,g}$ between L^p and L^g is calculated as shown in Eq. 5.

$$\begin{cases} CD_{p,g} = \dfrac{1}{m}\displaystyle\sum_{i=1}^{m}\left\| P_{g_i} - \hat{P}_{p_j} \right\|_2, \\ \hat{P}_{p_j} = \displaystyle\min_{P_{p_j}\in L^p}\left\| P_{p_j} - P_{g_i} \right\|_2, \end{cases} \tag{5}$$

where $P_{p_j} = (x_{p_j}, y_{p_j}, z_{p_j})$ and $P_{g_i} = (x_{g_i}, y_{g_i}, z_{g_i})$ are the points of L^p and L^g respectively, and $\hat{P}_{pj}$ is the closest point to a particular point P_{g_i}. m represents the number of points token at an equal distance from the ground-truth lane. If the unilateral chamfer distance is less than the chamfer distance threshold, it is written as τ_{CD}. It is generally accepted that L^p matches L^g and L^p is accepted as a true positive.

C. Results

We selected three 2D datasets, Tusimple [131], CULane [36], and CurveLanes [68], and three 3D datasets, Apollo 3D Lane Synthetic3D [16], OpenLane [17], and ONCE3DLanes [104], for comparison, as they are widely used in lane detection algorithms. In our opinion, nothing is more intuitive than data, therefore we give a large number of data references.

1) Results on Tusimple Dataset: The results of TuSimple [131] are shown in Tab. VIII. We use F1 score, accuracy, false positives, and false negatives to evaluate the model performance. The gap between the different methods on this dataset is smaller because the amount of data is smaller and the scenarios are more homogenous. Fig. 40 shows the visualisation of different methods on this dataset.

2) Results on CULane Dataset: The results of state-of-the-art methods on CULane [36] are shown in Tab. IX It shows the F1 score, as well as the evaluation results for 9 scenes. We adopt the F1 score to measure the performance. Fig. 41 shows the visualisation of different methods on this dataset.

3) Results on CurveLanes Dataset: In addition to the above two datasets, CurveLanes [68] is also widely used. We have chosen F1, Precision, and Recall as the evaluation criteria for CurveLanes. The results of CurveLanes [68] are shown in Tab. IX.

4) Results on Apollo 3D Lane Synthetic Dataset: We compare the test performance of multiple 3D lane detection methods on the Apollo 3D Lane Synthetic dataset [16], including F–Score, AP and X/Z error.As shown in Tab. XI. In addition, we provide Gen-LaneNet and Anchor3DLane visualisation results on this dataset, as shown in Fig. 43.

5) Results on OpenLane Dataset: We provide 3D evaluation results of multiple lane detection algorithms on the OpenLane dataset [17]. In order to thoroughly evaluate the models, we report the F-Score for the entire validation set and for different sets of scenarios, as shown in Tab. XII and XIII, respectively. The visualisation results are shown in Fig. 44.

6) Results on ONCE3DLanes Dataset: In the evaluation of the ONCE-3DLane dataset [104], four metrics were thoroughly analyzed and ensure the accuracy of the model scores. These metrics include the F-Score, Precision, Recall, and CD error, which are all presented in Tab. XIV. The results of the visualisation of this dataset are shown in Fig. 45.

VI. Current Challenges and Future Predictions

Despite impressive breakthroughs in lane detection technology, there still exists various setbacks, which not only hinder its current progress but also wield a considerable impact on its future potential. In this context, we need to delve deeper into the adaptability of lane detection technologies in different environments and find innovative solutions to meet these challenges.

A. Current Challenges

In the real world, the lane detection system is operating at all times, regardless of the environment or weather. The essence of lane detection is to detect information about the position

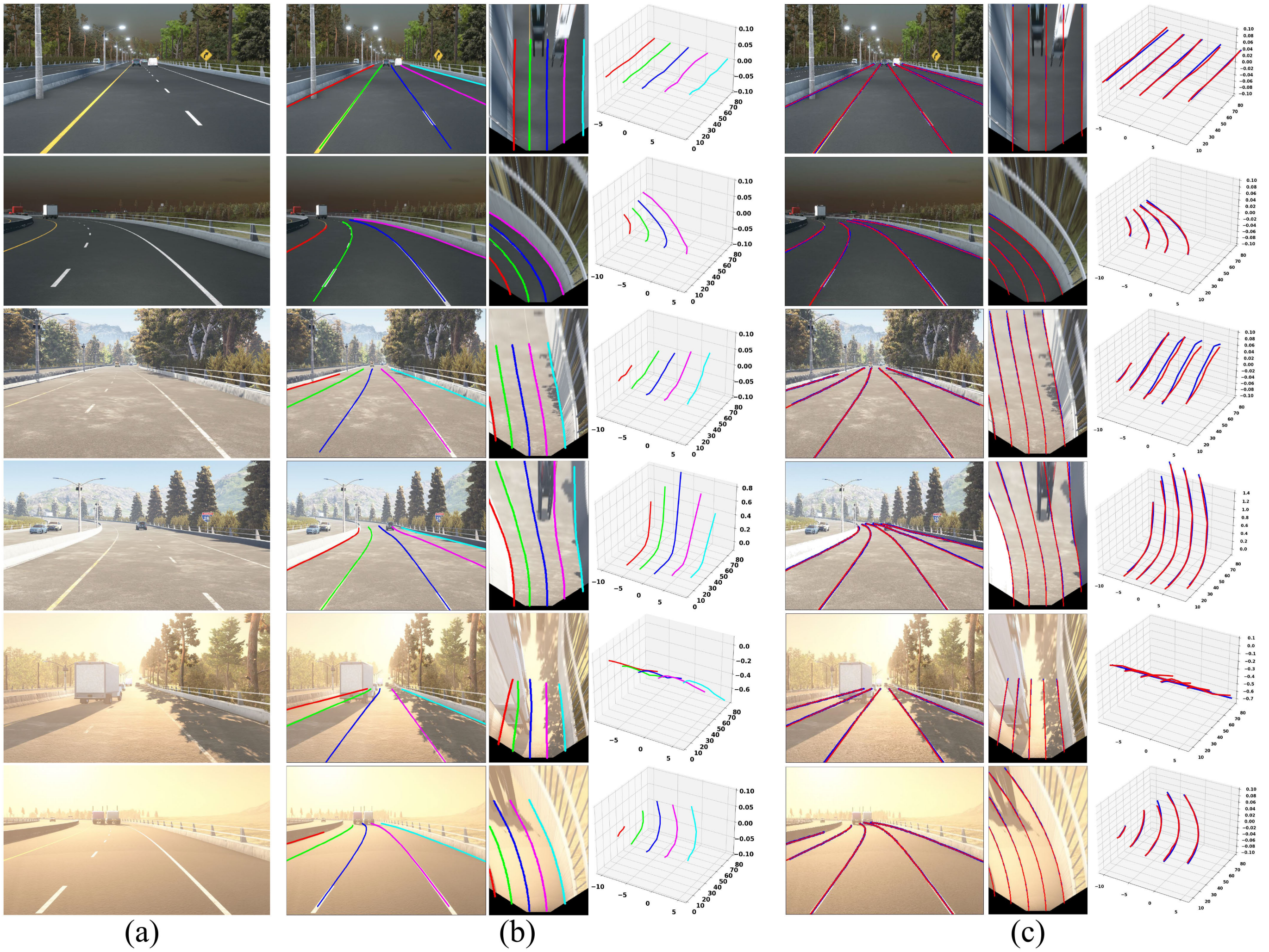


Fig. 43. (a) Several examples from the ApolloSim dataset. The figure shows six different scenarios for daytime and nighttime in sequence, includes both straight and curved lines. (b) Gen-LaneNet [16] visualisation results on the ApolloSim dataset. (c) Visualisation results of Anchor3DLane [111] on the ApolloSim dataset, where red represents the inferred results and blue indicates the dataset annotation.

of the lane relative to the vehicle. With this aim in mind, all strategies and improvements are designed to minimize recognition errors. From the data flow point of view, which means data input, data analysis, and data output, lane detection technology has the following issues.

1) Perceived Uncertainty: The sensors used in the vehicle may have errors that result in a certain degree of bias of the detected lane position, such as LiDAR accuracy, lens aberrations, and so on. Vehicle motion states (acceleration, deceleration, reorientation) can affect the position and shape of the captured lane lines, thus increasing the detection uncertainty. Therefore, the confidence with which deep learning makes predictions about the input data becomes very important. Unfortunately, the prevailing approach still tends to predict scores, and confidence is just now attracting attention. To reduce the error effects of input data, the method of model uncertainty estimation can be introduced to model and quantify the uncertainty in the detection results [138].

2) Generalization Ability: The dataset may not fully cover complex road conditions and environmental changes, such as different regions, weather conditions, and lighting, resulting in the model potentially performing poorly in new environments. There are dynamic scene changes during vehicle travel, such as occlusion by other vehicles, pedestrians, or trees, missing or changing lane lines, and unstructured roads with cracked and potholed pavements. However, the lane detection model fails to adequately consider these scenarios, which may result in low generalization ability. Deep learning models may overfit the training data during the training process, learning about the noise in the dataset and the subtle variations in a particular situation, while ignoring broader patterns and regularities. Additionally, there may be no learned feature representations that are generalizable for the lane detection task. To address the challenges in data processing, researchers can either collect more diverse and real-world data or use data augmentation techniques to simulate different weather conditions, lighting situations, and dynamic scenarios to cover a wide range of road conditions and environmental variations to enhance the robustness of the model. Regularization techniques [139] such as Dropout [140] and L1/L2 [141], [142] regularization can be used to prevent the model from overfitting the training data. In addition, more complex and flexible model architectures can be designed to improve the capture of lane line geometry and features.

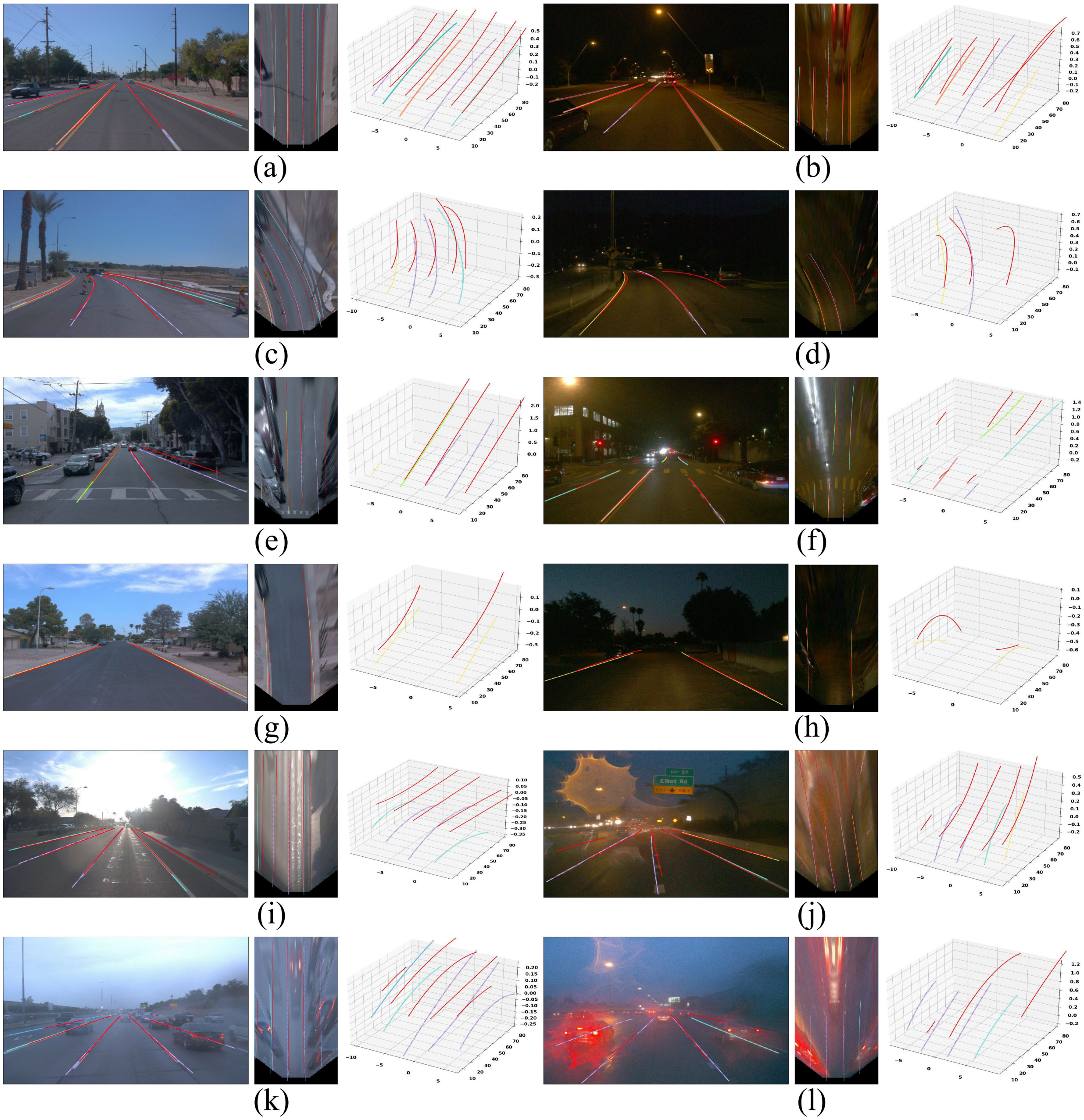


Fig. 44. Anchor3DLane [111] results on the Openlane dataset, with the red line indicating the predicted lane lines and the other color line indicating the ground truth lane lines labeled on the dataset. Figures a and b depict straight-line scenes during the day and night, respectively. Figures c and d display curved scenes during these times. Figures e and f showcase residential neighborhoods during the day and night. Figures g and h show rural paths during the day and night. Figures i and j depict exposure scenes during the day and night. Figures k and l showcase foggy scenes during the day and night.

We think lane detection in AI is in a DIKW model, i.e., data, information, knowledge, and wisdom are incorporated into a kind of pyramid-shaped hierarchy, with each layer endowed with some qualities over the next. Raw observations and measurements yield data and analyzing relationships between data yields information. The application of information in action produces knowledge. Wisdom is concerned with the future; it implies implication and lagging influence. Lane detection performance based on a certain dataset is effective, achieving better results on the same dataset. However, it's critical to highlight that the trained model's results on other datasets, testing its generalization ability, are substandard. We aim to establish a comprehensive, universally applicable model, suitable for all situations, but current scientific and technological progress in this area is challenging to translate into practical applications.

3) Adaptation of Lane Detection: In addition to lane detection, there are other detection tasks in automated driving, such as vehicle detection, pedestrian detection, and traffic marking detection, so lane detection must leave actionable space for other tasks. Recent multitask learning [47] proposes a solution idea to merge various detection tasks but multitask learning still has a lot of development space. To effectively detect lane lanes while leveraging features relevant to other tasks, it is paramount to design suitable feature extractors that enable efficient feature sharing and isolation. There may be competition

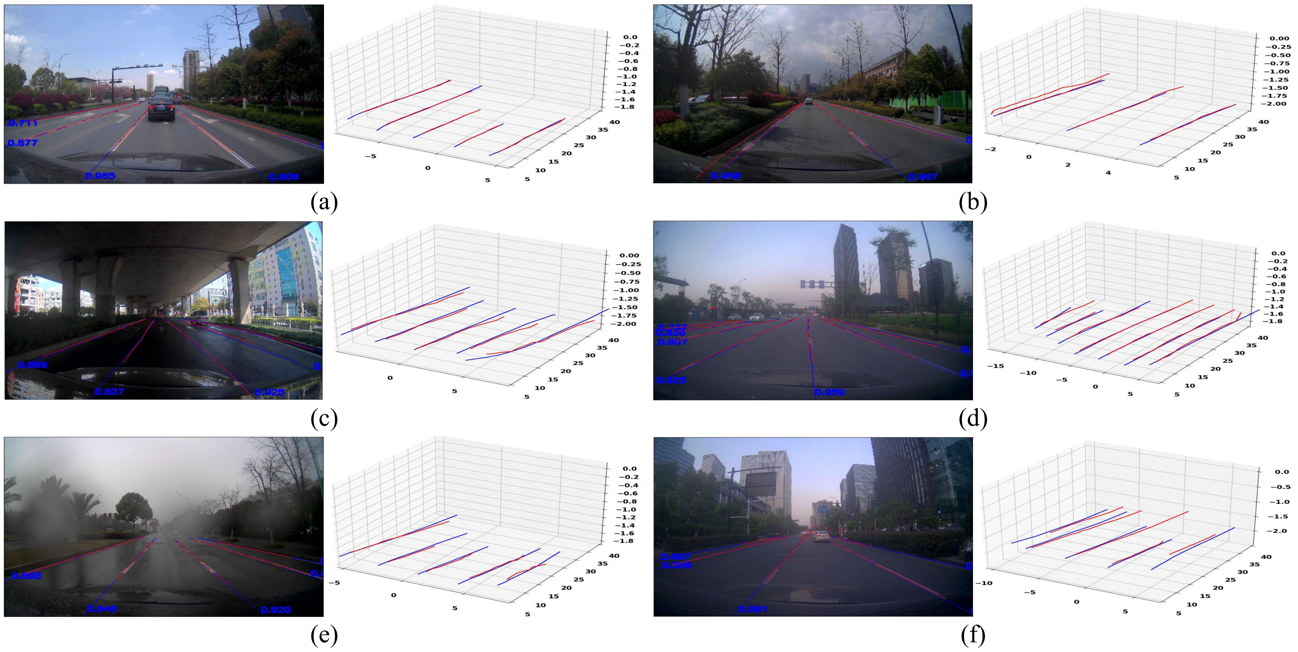
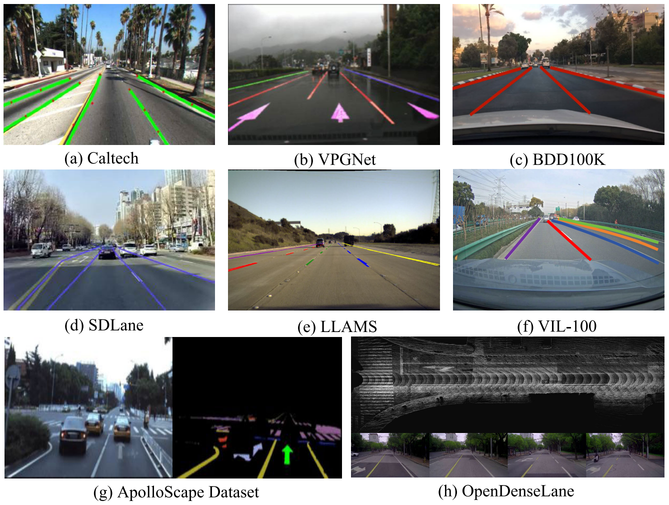


Fig. 45. Scenes under different weather conditions in the ONCE-3DLanees dataset, including (a) city roads, (b) residential areas, (c) under the viaduct, (d) multiple lanes, (e) foggy day, and (f) cloudy day scenes. The figure shows the lane representations and their 3D position coordinates in the original image after the Anchor3DLane [111] test.



(a) Caltech (b) VPGNet (c) BDD100K (d) SDLane (e) LLAMS (f) VIL-100 (g) ApolloScape Dataset (h) OpenDenseLane

Fig. 46. Example of partial lane dataset.

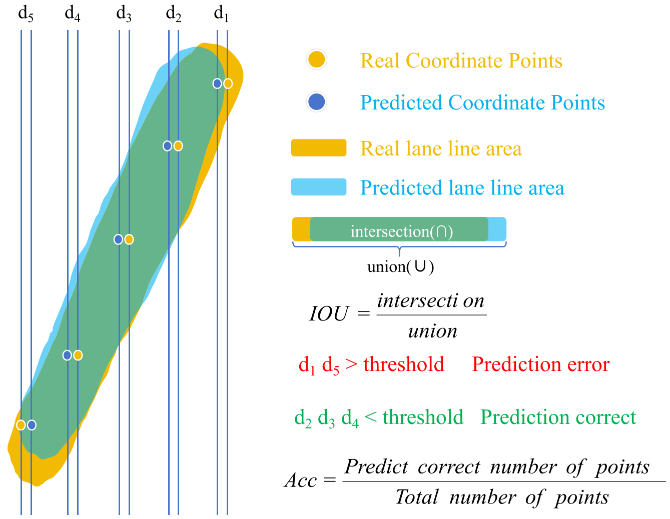

d_5
d_4
d_3
d_2
d_1
Real Coordinate Points
Predicted Coordinate Points
Real lane line area
Predicted lane line area
intersection(∩)
union(∪)
$IOU = \frac{intersecti\ on}{union}$
d_1 d_5 > threshold Prediction error
d_2 d_3 d_4 < threshold Prediction correct
$Acc = \frac{Predict\ correct\ number\ of\ points}{Total\ number\ of\ points}$

Fig. 47. Schematic representation of area-based and distance-based assessment approaches.

and cooperation between different tasks. For example, lane detection tasks may compete with vehicle detection tasks, thus computational resource allocation is an emerging issue. On the other hand, they may also improve each other's performance by sharing features, which requires a reasonable trade-off when designing the network architecture. In addition, there may be an imbalance in the distribution of data across different tasks, which can cause the model to have an excessive performance advantage for some tasks and poor performance for others during training. Therefore, appropriate strategies such as weighted loss functions or resampling techniques are needed to improve the model's performance.

4) Complexity of Data Processing and Analysis: In this era where data is king, lane detection systems require large amounts of data to train and optimize algorithms to ensure accuracy and stability in different environments and conditions. Therefore, extensive preprocessing and labeling are required to ensure the quality and consistency of the data. Secondly, the lane detection algorithm should have strong data analysis capabilities. It should be able to extract effective features from large amounts of data and can recognize and judge them accurately. As introduced in 6.1.3. In addition, lane detection needs to take into account the real-time requirements. It can identify and track lanes in a timely and accurate manner even at high speeds or in complex traffic situations. As a result, there is also a higher demand for efficiency and speed in data processing and analysis. On the one hand, relying on existing theories, parallel computing, GPU acceleration, and other techniques are used to speed up processing and improve system performance. Another possible approach is to develop

TABLE VIII

PERFORMANCE OF DIFFERENT METHODS ON TUSIMPLE

Methods	**Backbone**	**Year**	**F1(%)↑**	**Acc(%)↑**	**FP(%)↓**	**FN(%)↓**
Segmentation-based						
LaneNet [35]	ENet	2018	94.8	96.38	7.8	2.44
SCNN [36]	VGG16	2018	95.97	96.53	6.17	1.8
RESA [37]	ResNet34	2021	96.93	96.82	3.63	2.48
LaneAF [39]	DLA-34	2021	96.49	95.62	2.8	4.18
Laneformer [40]	ResNet18	2022	-	96.54	4.35	2.36
Laneformer [40]	ResNet34	2022	-	96.56	5.39	3.37
AtrousFormer [41]	XR34	2023		96.59	2.83	3.26
AtrousFormer [41]	XR34	2023	-	96.71	2.82	3.24
ENet-SAD [42]	ENet	2019	95.92	96.64	6.02	2.05
EL-GAN [44]	EL-GAN	2018	96.26	94.9	4.12	3.36
Ripple-GAN [45]	RiLLD-Net	2020	97.67	97.28	0.48	2.89
Lane2Seq(segmentation) [91]	ViT-Base	2024	97.95	96.85	2.01	2.03
Anchor&Detection based						
PointLaneNet [60]	MobileNet-v2	2019	95.07	96.34	4.67	5.18
Line-CNN [61]	ResNet122	2019	-	96.87	4.42	1.97
SIIC-Net [62]	ResNet18	2022	-	95.62	3.2	3.99
ADNet [63]	ResNet18	2023	96.9	96.23	2.91	3.29
ADNet [63]	ResNet34	2023	97.31	96.6	2.83	2.53
Sparse Laneformer [65]	ResNet34	2024	96.81	95.69	-	-
UFLD [66]	ResNet34	2020	88.02	95.86	18.91	3.75
ERF-E2E [67]	ERFNet	2020	-	96.02	3.21	4.28
CondLaneNet [69]	ResNet18	2020	97.01	95.48	2.18	3.8
CondLaneNet [69]	ResNet34	2020	96.98	95.37	2.2	3.82
LaneATT [70]	ResNet18	2021	96.71	95.57	3.56	3.01
LaneATT [70]	ResNet34	2021	96.77	95.63	3.53	2.92
UFLDv2 [71]	ResNet34	2022	96.13	95.53	3.06	4.66
CLRNet [73]	ResNet18	2022	97.89	96.84	2.28	1.92
CLRNet [73]	ResNet34	2022	97.82	96.87	2.27	2.08
CANet-S [72]	ResNet-18	2023	97.51	96.56	2.29	2.68
CANet-M [72]	ResNet-34	2023	97.44	96.66	2.32	2.79
CANet-L [72]	ResNet-101	2023	97.77	96.76	1.92	2.53
CondLSTR [74]	ResNet-18	2023	97.71	96.06	1.79	2.82
CondLSTR [74]	ResNet-34	2023	97.64	96.06	1.84	2.92
CondLSTR [74]	ResNet-101	2023	97.94	96.02	1.55	2.56
SRLane [76]	ResNet18	2024	97.66	96.85	2.8	1.85
Lane2Seq(anchor) [91]	ViT-Base	2024	97.86	96.72	2.21	2.05
Keypoint-based						
PINet (4H) [80]	PINet	2021	96.75	-	3.1	2.5
FOLOLane [84]	ERFNet	2021	96.59	96.92	4.47	2.28
GANet-S [85]	ResNet18	2022	97.71	95.95	1.97	2.62
GANet-M [85]	ResNet34	2022	97.68	95.87	1.99	2.64
GANet-L [85]	ResNet101	2022	97.45	96.44	2.63	2.47
RCLane-S [86]	SegFormer-B0	2022	97.52	96.49	2.21	2.57
RCLane-M [86]	SegFormer-B1	2022	97.61	96.51	2.24	2.36
RCLane-L [86]	SegFormer-B2	2022	97.64	96.58	2.28	2.27
LanePtrNet [87]	HRNet-18	2024	-	95.8	4.54	3.01
Parameter based						
PolyLaneNet [88]	EfficientNetB0	2021	90.62	93.36	9.42	9.33
LSTR [89]	ResNet18	2021	96.85	96.18	2.91	3.38
BézierLaneNet [90]	ResNet18	2022	-	95.41	5.3	4.6
BézierLaneNet [90]	ResNet34	2022	-	95.65	5.1	3.9
DecoupleLane [92]	DLA-34	2023	97.93	97.01	2.03	3.31
Lane2Seq(parameter) [91]	ViT-Base	2024	96.59	96	2.23	3.54

new computational acceleration algorithms, as well as to push for a new generation of more powerful processors.

5) Social Issues. a) Resource Sharing: In the automotive industry, various vehicle companies have accumulated a huge amount of data resources. However, they are cautious about sharing data publicly. Vehicle companies have invested a lot of resources in data collection, processing, and analysis, which may contain their core technologies and trade secrets. The vehicle industry is highly competitive, with companies vying for market share and technological advantage. Public data sharing may give competitors access to favorable information, intensifying competition in the marketplace and reducing the firm's competitive advantage. In addition, data protection regulations such as the General Data Protection Regulation (GDPR) have strict requirements on the process and protection of personal data. Vehicle companies need to ensure the legality, security, and privacy protection of their data so as not to violate laws and regulations and bear legal liabilities and business risks.

6) b) Formulation of Laws, Regulations and Standards: With the development of autonomous driving technology, relevant laws, regulations, and standards need to be established to regulate lane detection system design, use, and testing. This involves issues such as road traffic regulations, safety standards, and allocation of responsibilities.

7) c) Road Infrastructure Adaptation: Autonomous vehicles need to rely on road marks for localization and navigation. Therefore road infrastructure must be suitable to ensure that it can meet the needs of autonomous driving systems. It involves work such as updating road markings and installing road sensors.

8) d) Human-Computer Inflation (HCI) and Human Intervention (HI): In the automatic driving process, HI may be required for intervention or decision-making, such as in complex traffic situations and emergency situations. Therefore, appropriate HCI interfaces and decision-making mechanisms are necessary for ensuring effective communication between humans and the autonomous driving system.

9) e) Technology Popularization and Acceptance: The popularization of autonomous driving technology needs to face public acceptance and recognition, including aspects such as trust in the technology, safety concerns, and cost considerations.

B. Predictions of Future Directions

1) 2D + 3D Lane Detection: The 2D lane detection results provide important support for the acquisition of 3D coordinates. By extracting features from 2D images, we can obtain information about the position and shape of the lanes in the 2D plane. This information lays the foundation for subsequent 3D detection. After acquiring the 2D lane data, this 2D information can be mapped to the 3D space by combining the internal and external parameters of the camera. For example, using known camera parameters and image depth information, the system can convert 2D coordinates to 3D coordinates, thereby accurately determining the position of an object in 3D space. This mapping not only helps to enhance scene understanding but also supports path planning and decision-making.

2) Multi-sensor Fusion: As of now, the mainstream sensors equipped in autonomous vehicles are cameras, LiDAR, millimeter wave mines, inertial measurement units (IMUs), and GPS. These sensors play an essential role in L2 as well as L3. As the technology iterates, new sensors will gradually make their way onto the scene, just like event cameras [143], depth cameras, multi-threaded 3D LiDAR, and other sensors with higher accuracy. Combining data from different sensors can compensate for the limitations of a single sensor, thus improving the accuracy, reliability, and applicability of lane detection.

TABLE IX
PERFORMANCE OF DIFFERENT METHODS ON CULANE

Methods	Backbone	Year	Normal↑	Crowd↑	Dazzle↑	Shadow↑	No line↑	Arrow↑	Curve↑	Cross↓	Night↑	Total↑
Segmentation-based												
SCNN [36]	VGG16	2018	90.60	69.70	58.50	66.90	43.40	84.10	64.40	1990	66.10	71.60
RESA [37]	ResNet34	2021	91.90	72.40	66.50	72.00	46.30	88.10	68.60	1896	69.80	74.50
PSSNet [38]	PSSNet(N=1)	2021	92.30	73.70	68.10	72.80	49.40	88.50	70.80	1747	71.20	75.76
LaneAF [39]	DLA-34	2021	91.80	75.61	71.78	79.12	51.38	86.88	72.70	1360	73.03	77.41
Laneformer [40]	ResNet18	2022	88.60	69.02	64.07	65.02	45.00	81.55	60.46	25	64.76	71.71
Laneformer [40]	ResNet34	2022	90.74	72.31	69.12	71.57	47.37	85.07	65.90	26	67.77	74.70
AtrousFormer [41]	XR18	2023	92.72	75.56	68.16	73.67	50.07	88.82	70.32	1169	73.49	77.63
AtrousFormer [41]	XR34	2023	92.83	75.96	69.48	77.86	50.15	88.66	71.14	1054	73.74	78.08
ENet-SAD [42]	ENet	2019	90.10	68.80	60.20	65.90	41.60	84.00	65.70	1998	66.00	70.80
IntRA-KD [43]	ERFNet	2020	-	-	-	-	-	-	-	-	-	72.40
SIM-CycleGAN [46]	SIM-CycleGAN	2020	91.80	71.80	66.40	76.20	46.10	87.80	67.10	2346	69.40	73.90
Lane2Seq(segmentation) [91]	ViT-Base	2024	93.39	77.27	73.45	79.69	53.91	90.53	73.37	1129	74.96	79.64
Anchor&Detection based												
PointLaneNet [60]	MobileNet_v2	2019	88.00	68.10	61.50	63.30	44.00	80.90	65.20	1640	63.20	70.20
SIIC-Net [62]	ResNet18	2022	91.50	74.80	69.70	72.30	51.10	87.70	62.00	1507	71.40	76.50
ADNet [63]	ResNet18	2023	91.92	75.81	69.39	76.21	51.75	87.71	68.84	1133	72.33	77.56
ADNet [63]	ResNet34	2023	92.90	77.45	71.71	79.11	52.89	89.90	70.64	1499	74.78	78.94
O2SFormer[64]	ResNet34	2023	92.50	75.25	70.93	77.72	50.97	87.63	68.10	2749	72.88	77.03
Sparse Laneformer[65]	ResNet34	2024	-	-	-	-	-	-	-	-	-	77.77
UFLD [66]	ResNet34	2020	90.70	70.20	59.50	69.30	44.40	85.70	69.50	2037	66.70	72.30
ERF-E2E [67]	ERFNet	2020	91.00	73.10	64.50	74.10	46.60	85.80	71.90	2022	67.90	74.00
CurveLane-NAS [68]	CurveLanes-S	2020	88.30	68.60	63.20	68.00	47.90	82.50	66.00	2817	66.20	71.40
CurveLane-NAS [68]	CurveLanes-M	2020	90.20	70.50	65.90	69.30	48.80	85.70	67.50	2359	68.20	73.50
CurveLane-NAS [68]	CurveLanes-L	2020	90.70	72.30	67.70	70.10	49.40	85.80	68.40	1746	68.90	74.80
CondLaneNet [69]	ResNet18	2020	92.87	75.79	70.72	80.01	52.39	89.37	72.40	1364	73.23	78.14
CondLaneNet [69]	ResNet34	2020	93.38	77.14	71.17	79.93	51.85	89.89	73.88	1387	73.92	78.74
LaneATT [70]	ResNet18	2021	91.11	72.96	65.72	70.91	48.35	85.49	63.37	1170	68.95	75.09
LaneATT [70]	ResNet34	2021	92.14	75.03	66.47	78.15	49.39	88.38	67.72	1330	70.72	76.68
UFLDv2 [71]	ResNet34	2022	92.50	74.90	65.70	75.30	49.00	88.50	70.20	1864	70.60	75.90
CANet-S [72]	ResNet18	2023	93.07	76.59	70.51	77.82	52.24	89.39	72.48	1213	72.91	78.46
CANet-M [72]	ResNet34	2023	93.58	77.88	73.11	75.06	51.68	90.09	75.54	1176	73.92	79.16
CANet-L [72]	ResNet101	2023	93.60	78.74	70.07	79.35	52.88	90.18	76.69	1196	74.91	79.86
CLRNet [73]	ResNet-18	2022	93.30	78.33	73.71	79.66	53.14	90.25	71.56	1321	75.11	79.58
CLRNet [73]	ResNet34	2022	93.49	78.06	74.57	79.92	54.01	90.59	72.77	1216	75.02	79.73
CLRNet [73]	DLA34	2022	93.73	79.59	75.30	82.51	54.58	90.62	74.13	1155	75.37	80.47
CondLSTR [74]	ResNet18	2023	94.11	79.17	73.55	80.39	54.41	90.37	75.89	1214	75.39	80.36
CondLSTR [74]	ResNet34	2023	94.12	79.72	77.02	82.51	53.76	90.59	76.65	1370	75.57	80.55
CondLSTR [74]	ResNet101	2023	94.17	79.90	75.43	80.99	55.00	90.97	76.87	1047	75.11	80.77
CLRerNet [75]	ResNet34	2024	93.93	79.51	73.88	83.16	55.55	90.87	74.45	1088	76.02	80.76
CLRerNet [75]	DLA-34	2024	94.02	80.20	74.41	83.71	56.27	90.39	74.67	1161	76.53	81.12
SRLane [76]	ResNet18	2024	93.50	77.80	71.60	78.80	52.10	90.20	74.70	1365	74.70	78.90
Lane2Seq (anchor) [91]	ViT-Base	2024	93.11	77.43	73.25	79.46	53.74	90.02	72.44	1173	75.12	79.27
Keypoint-based												
PINet (4H) [80]	PINet	2021	90.30	72.30	66.30	68.40	49.80	83.70	65.60	1427	67.70	74.40
FOLOLane [84]	ERFNet	2021	92.70	77.80	75.20	79.30	52.10	89.00	69.40	1569	74.50	78.80
GANet-S [85]	ResNet18	2022	93.24	77.16	71.24	77.88	53.59	89.62	75.92	1240	72.75	78.79
GANet-M [85]	ResNet34	2022	93.73	77.92	71.64	79.49	52.63	90.37	76.32	1368	73.67	79.39
GANet-M [85]	ResNet101	2022	93.67	78.66	71.82	78.32	53.38	89.86	77.37	1352	73.85	79.63
RCLane-S [86]	SegFormer-B0	2022	93.41	77.93	73.32	80.31	53.84	89.04	75.66	1298	74.33	79.52
RCLane-M [86]	SegFormer-B1	2022	93.59	78.77	72.44	84.37	52.77	90.31	78.39	907	73.96	80.03
RCLane-L [86]	SegFormer-B2	2022	94.01	79.13	72.92	81.16	53.94	90.51	79.66	931	75.10	80.50
LanePtrNet [87]	HRNet-18	2024	92.02	79.88	67.77	81.95	58.86	87.59	66.61	1574	75.35	79.82
Parameter based												
LSTR [89]	ResNet18	2021	86.78	67.34	56.63	59.82	40.10	78.66	56.64	1166	59.92	68.72
BézierLaneNet [90]	ResNet18	2022	90.22	71.55	68.70	70.91	45.30	84.09	62.49	996	58.98	73.67
BézierLaneNet [90]	ResNet34	2022	91.59	73.20	69.90	76.74	48.05	87.16	69.20	62.45	888.00	75.57
DecoupleLane [92]	DLA-34	2023	93.85	80.32	76.31	82.34	55.40	90.48	75.12	1035	75.40	80.82
Lane2Seq (parameter) [91]	ViT-Base	2024	93.03	76.42	72.17	78.32	52.89	89.67	72.67	1319	73.98	78.39

TABLE X
PERFORMANCE OF DIFFERENT METHODS ON CURVELANES

Methods	Backbone	Year	F1 (%) ↑	Pre (%) ↑	Recall (%) ↑
SCNN [36]	VGG16	2018	65.02	76.13	56.74
ENet-SAD [42]	ENet	2019	50.31	63.60	41.60
PointLaneNet [60]	MobileNet_v2	2019	78.47	86.33	72.91
CurveLane-NAS[68]	CurveLanes-S	2020	81.12	93.58	71.59
CurveLane-NAS [68]	CurveLanes-M	2020	81.80	93.49	72.71
CurveLane-NAS [68]	CurveLanes-L	2020	82.29	91.11	75.03
CondLaneNet-S [69]	ResNet18	2020	85.09	87.75	82.58
CondLaneNet-M [69]	ResNet34	2020	85.92	88.29	83.68
CondLaneNet-L [69]	ResNet101	2020	86.10	88.98	83.41
UFLDv2 [71]	ResNet34	2022	80.45	81.49	79.44
UFLDv2 [71]	ResNet34	2022	81.34	81.93	80.76
RCLane-S [86]	SegFormer-B0	2022	90.47	93.33	87.78
RCLane-M [86]	SegFormer-B1	2022	90.96	93.47	88.58
RCLane-L [86]	SegFormer-B2	2022	91.43	93.96	89.03
CLRNet [73]	DLA34	2022	86.10	91.40	81.39
CANet-S [72]	ResNet18	2023	86.57	91.37	82.25
CANet-M [72]	ResNet34	2023	87.19	91.53	83.25
CANet-L [72]	ResNet101	2023	87.87	91.69	84.36
CondLSTR [74]	ResNet-18	2023	87.99	90.90	85.27
CondLSTR [74]	ResNet-34	2023	88.23	91.24	85.41
CondLSTR [74]	ResNet-101	2023	88.47	91.32	85.80
CLRerNet [75]	DLA34	2024	86.47	91.66	81.83

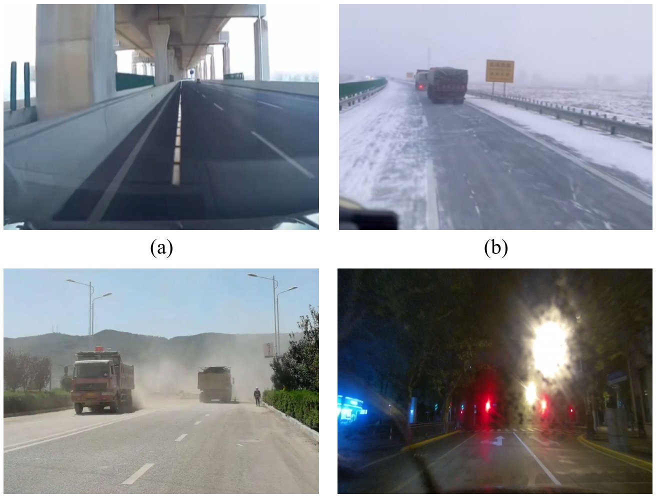

(a) (b) (c) (d)

Fig. 48. Interference with lane detection in different environments. (a) Light projected onto the roadway. (b) Snow on the road surface. (c) Dust on the road. (d) Various lights at night.

For example, the camera may not be able to sense when there is strong light or light interference, as shown in Fig. 48 (a). At this time, the high-precision LiDAR will detect the road coating, road asphalt, and other factors via the reflected echoes, so as to detect the lanes. In addition, the shape and position of lane lines in an image may change, making lane line detection difficult. Camera vibrations and changes in viewing angle can also affect accuracy and stability. A multi-sensor fusion approach can effectively address these difficulties. Combining camera, LiDAR, and other sensor data improves the accuracy and stability of lane line position. In this way, the fusion design of these sensors will enable their deployment for advanced levels of autonomous driving (L4 and L5), and they are likely to serve as the data source for future lane detection systems. But when fusing data from different sensors, new challenges arise, one of which is the need to design applicable data fusion algorithms. It is a topic worthy of our research.

3) Responding to Weather Changes: Current lane detection performance has progressed significantly in well-lit, clear lane marking conditions. However, there is still room for improvement in accuracy in harsh environments, especially in low-light conditions, such as cloudy days and nights, as shown in Tab. IX XIII. On cloudy days or at night, lack of light makes lane marks unclear, making it difficult for the algorithm to recognize and track lanes [46]. In addition, background clutter information (reflections from street and vehicle lights) may interfere with detection and analysis. We can use special sensors, such as low-light enhanced cameras, infrared cameras, or LiDAR, to obtain additional information and improve system robustness and stability through optimization algorithms and augmented learning.

In addition, most of the current research on severe weather focuses on cloudy days, glare, rain, and fog [17], but there is a lack of research on snow and dust, just like in Fig. 48. In winter, autonomous driving encounters situations rarely found in lane datasets, such as snowy landscapes. In the lane dataset, snowy scenarios are relatively rare, but in winter, snow poses a significant obstacle that must be factored in when developing autonomous driving technology. Ice on the road and snow on the road has a significant impact on the decision control of the automated driving system. Sandy and dusty weather also have a major interference in the detection interference of various sensors. Therefore, how lane detection can cope with these weather conditions will become a future research direction. Combining the existing theories [144], [145], [146], [147], we believe that image snow removal and image denoising can be used in snowy and other unfavorable weather lane detection to enable the preprocessing of camera data acquisition.

4) Large Language Model (LLM): The launch of Chat Generative Pre-trained Transformer (ChatGPT) [148] shows the fascination of large language models in the field of artificial intelligence (AI). Transformer [58] is one of the core architectures for large-scale language modeling, and after Persformer [17] introduced the Transformer into lane detection, lane detection performance was greatly improved. And as a result, we have a new idea. Going to develop a dedicated LLM that focuses on lane detection tasks. Inputting data captured by vehicle sensors and cameras into LLM utilizes its powerful language understanding and contextual inference capabilities to interpret lane markings and road conditions. LLM can analyze various visual features in an image and generate detailed descriptions of lane position, type, and state based on context, thereby providing more accurate and richer information to the vehicle's automated driving system. In addition, LLM can combine information such as vehicle sensors and GPS data to provide a deeper understanding and inference of the vehicle's surroundings. Combining language models with traditional computer vision techniques can improve the ability of lane detection systems to understand complex traffic scenes.

We think that a grand language model in autonomous driving can accept image sequences as input and output scene

TABLE XI
PERFORMANCE OF DIFFERENT METHODS ON APOLLO 3D LANE SYNTHETIC

Scene	Methods	Backbone	Year	F1(%)↑	AP(%)↑	X error (m)↓		Z error (m)↓	
						near	*far*	*near*	*far*
Balanced Scene	3DLaneNet [15]	VGG-16	2019	86.4	89.3	0.068	0.477	0.015	0.202
	Gen-LaneNet [16]	ERFNet	2020	88.1	90.1	0.061	0.496	0.012	0.214
	CLGo [103]	ResNet18	2022	91.9	94.2	0.061	0.361	0.029	0.250
	PersFormer [17]	EfficientNet	2022	92.9	-	0.054	0.356	0.010	0.234
	GT [105]	ERFNet	2022	91.9	93.8	0.049	0.387	0.008	0.213
	CurveFormer [109]	EfficientNet	2023	95.8	97.3	0.078	0.326	0.018	0.219
	LATR [110]	ResNet50	2023	96.8	97.9	0.022	0.253	0.007	0.202
	Anchor3DLane [111]	ResNet18	2023	95.6	97.2	0.052	0.306	0.015	0.223
	BEV-LaneDet [20]	ResNet34	2023	96.9	-	0.016	0.242	0.020	0.216
	WS-3D-Lane [114]	3DLaneNet	2023	93.5	95.7	0.027	0.321	0.006	0.215
	LS-3DLane [115]	EfficientNet-b0	2024	89.4	89.6	0.083	0.480	0.022	0.271
Rarely Observed	3DLaneNet [15]	VGG-16	2019	72.0	74.6	0.166	0.855	0.039	0.521
	Gen-LaneNet [16]	ERFNet	2020	78.0	79.0	0.139	0.903	0.030	0.539
	CLGo [103]	ResNet18	2022	86.1	88.3	0.147	0.735	0.071	0.609
	PersFormer [17]	EfficientNet	2022	87.5	-	0.107	0.782	0.024	0.602
	GT [105]	ERFNet	2022	83.7	85.2	0.126	0.903	0.023	0.625
	CurveFormer [109]	EfficientNet	2023	95.6	97.1	0.182	0.737	0.039	0.561
	LATR [110]	ResNet50	2023	96.1	97.3	0.050	0.600	0.015	0.532
	Anchor3DLane [111]	ResNet18	2023	94.4	96.9	0.094	0.693	0.027	0.579
	BEV-LaneDet [20]	ResNet34	2023	97.6	-	0.031	0.594	0.040	0.556
	LS-3DLane [115]	EfficientNet-b0	2024	78.1	79.9	0.173	0.952	0.047	0.705
Visual Variations	3DLaneNet [15]	VGG-16	2019	72.5	74.9	0.115	0.601	0.032	0.230
	Gen-LaneNet [16]	ERFNet	2020	85.3	87.2	0.074	0.538	0.015	0.232
	3DLaneNAS [102]	VGG-16	2022	92.1	92.4	0.037	0.358	0.005	0.192
	CLGo [103]	ResNet18	2022	87.3	89.2	0.084	0.464	0.045	0.312
	PersFormer [17]	EfficientNet	2022	89.6	-	0.074	0.430	0.015	0.266
	GT [105]	ERFNet	2022	89.9	92.1	0.060	0.446	0.011	0.235
	CurveFormer [109]	EfficientNet	2023	90.8	93.0	0.125	0.410	0.028	0.254
	LATR [110]	ResNet50	2023	95.1	96.6	0.045	0.315	0.016	0.228
	Anchor3DLane [111]	ResNet18	2023	91.4	93.6	0.068	0.367	0.020	0.232
	BEV-LaneDet [20]	ResNet34	2023	95.0	-	0.027	0.320	0.031	0.256
	LS-3DLane [115]	EfficientNet-b0	2024	87.3	89.0	0.085	0.529	0.032	0.330

TABLE XII
PERFORMANCE OF DIFFERENT METHODS ON OPENLANE

Modal	Methods	Backbone	Year	F1(%)↑	Categor Accuracy↑	X error(m)↓		Z error(m)↓	
						near	*far*	*near*	*far*
Monocular	3DLaneNet [15]	VGG-16	2019	44.1	-	0.479	0.572	0.367	0.443
	Gen-LaneNet [16]	ERFNet	2020	32.3	-	0.591	0.684	0.411	0.521
	PersFormer [17]	EfficientNet	2022	50.5	92.3	0.485	0.553	0.364	0.431
	STLane3D [106]	ResNext-50	2022	50.6	-	-	0.500	-	0.178
	PETRv2 [107]	VoVNetV2	2023	61.2	-	0.400	0.573	0.265	0.413
	CurveFormer [109]	EfficientNet	2023	50.5	-	0.340	0.772	0.207	0.651
	LATR [110]	ResNet50	2023	61.9	92.0	0.219	0.259	0.075	0.104
	Anchor3DLane [111]	ResNet18	2023	53.1	90.0	0.300	0.311	0.103	0.139
	BEV-LaneDet [20]	ResNet34	2023	58.4	-	0.309	0.659	0.244	0.631
	An Efficient Transformer [113]	EfficientNet	2023	63.8	91.5	0.245	0.304	0.104	0.129
	GroupLane [108]	ConvNext-B	2023	64.1	92.8	0.320	0.441	0.233	0.402
	CurveFormer++ [116]	EfficientNet	2024	52.7	88.1	0.337	0.801	0.198	0.676
Multi-Modal	M^2-3DLaneNet [117]	3D pillar-based+ EfficientNet B7	2022	55.5	-	0.431	0.487	0.327	0.401
	DV-3DLane-Tiny [118]	ResNet18	2023	63.4	91.6	0.137	0.159	0.034	0.063

description results. The extensive prior knowledge from pre-training a large visual language model (VLM) can be utilized to improve perception and comprehension performance [149]. In this, vision-to-text (V2T) is key. We can integrate visual

TABLE XIII
PERFORMANCE OF DIFFERENT METHODS ON OPENLANE

Modal	Methods	Backbone	Year	All	Up&Down	Curve	Extreme Weather	Night	Intersection	Merge&Split
Monocular	3DLaneNet [15]	VGG-16	2019	44.1	40.8	46.5	47.5	41.5	32.1	41.7
	Gen-LaneNet [16]	ERFNet	2020	32.3	25.4	33.5	28.1	18.7	21.4	31.0
	PersFormer [17]	EfficientNet	2022	50.5	42.4	55.6	48.6	46.6	40.0	50.7
	STLane3D [106]	ResNext-50	2022	50.6	41.3	47.4	54.0	51.3	42.5	47.9
	CurveFormer[109]	EfficientNet	2023	50.5	45.2	56.6	49.7	49.1	42.9	45.4
	LATR [110]	ResNet50	2023	61.9	55.2	68.2	57.1	55.4	52.3	61.5
	Anchor3DLane [111]	ResNet18	2023	53.1	45.5	56.2	51.9	47.2	44.2	50.5
	BEV-LaneDet [20]	ResNet34	2023	58.4	48.7	63.1	53.4	53.4	50.3	53.7
	DecoupleLane [92]	DLA-34	2023	51.2	43.5	57.3	-	48.9	43.5	-
	An Efficient Transformer [113]	EfficientNet	2023	63.8	57.6	73.2	57.3	59.7	57.0	64.9
	GroupLane [108]	ConvNext-B	2023	64.1	-	-	-	-	-	-
	CurveFormer++ [116]	EfficientNet	2024	52.7	48.3	59.4	50.6	48.4	45.0	48.1
Multi-Modal	M^2-3DLaneNet [117]	3D pillar-based+ EfficientNet B7	2022	55.5	53.4	60.7	56.2	51.6	43.8	51.4
	DV-3DLane-Tiny [118]	ResNet18	2023	63.4	-	-	-	-	-	-

TABLE XIV
PERFORMANCE OF DIFFERENT METHODS ON ONCE3DLANES

Methods	Backbone	Year	F1(%)↑	Precision(%)↑	Recall(%)↑	CD Error(m)↓
3DLaneNet [15]	VGG-16	2019	44.73	61.46	35.16	0.127
Gen-LaneNet [16]	ERFNet	2020	45.59	63.95	35.42	0.121
SALAD [104]	Segformer	2022	64.07	75.90	55.42	0.098
PersFormer [17]	EfficientNet	2022	74.33	80.30	69.18	0.074
STLane3D [106]	ResNext-50	2022	74.05	76.63	71.64	0.085
LATR [110]	ResNet50	2023	80.59	86.12	75.73	0.052
Anchor3DLane [111]	ResNet18	2023	74.44	80.50	69.23	0.064
DecoupleLane [92]	DLA-34	2023	75.07	81.19	69.26	0.062
An Efficient Transformer [113]	EfficientNet	2023	80.84	84.50	77.48	0.056
WS-3D-Lane sup [114]	3D-LaneNet	2023	77.02	84.51	70.75	0.058
GroupLane [108]	ConvNext-B	2023	79.42	82.41	76.54	0.054
CurveFormer++ [116]	EfficientNet	2024	77.22	82.22	72.79	0.081

embeddings and linguistic embeddings using various fusion methods to jointly optimize the feature representation used for the target task [150]. Alternatively, by comparative learning on large numbers of image-text pairs, image features relevant to language are captured [151]. Specifically, for lane detection, the image captured by the camera with the associated lane markings is used as the learning object. Enhance the lane perception in various traffic environments through extensive prior knowledge of VLM. Then, search for the association between traffic scene images and lane markings to capture representations related to "lane language".

5) Powerful Computing Ability: Advanced deep learning algorithms and complex data fusion techniques require execution in real-time environments to achieve accurate perception and understanding of the environment. Therefore, lane detection systems require sufficient computational resources to support the real-time execution of these algorithms. In addition, lane detection systems usually require cooperation with high-precision maps to achieve more accurate lane sensing and localization. Constructing and updating high-precision maps involves extensive data processing and computation work, which also requires the support of powerful computing resources. The computational speed improvement has two directions. On the one hand, computational strategies can be optimized, which requires good models and algorithms to reduce the algorithmic complexity, including space complexity and time complexity. On the other hand, developing more powerful computing processors and accelerating hardware updates and iterations that lead to powerful hardware computing capabilities. It requires a concerted effort by companies, research institutes, and other researchers. One point we need to make clear is that these algorithms are serving autonomous driving. Unless all the data are sent to the cloud via 4G/5G or even newer communication technology, it will only work in the vehicle [152]. While the vehicle resources are limited, its primary task is traveling, not computing. Therefore, how to run these algorithms on an on-board computer [153], [154] has become a worthy direction to explore.

6) Optimization of Training Process: The training data needs to be accurate and diverse. Filtering is required before training to minimize noise [138], [155]. Therefore, we can learn from parameter adjustment in control theory. The training hyperparameters can also be adaptively adjusted so as to reach the target state as soon as possible. More specifically, the model training situation is monitored in time during the training process, including the change of the loss function, the improvement of the model performance, and so on. And timely adjust the training strategy and hyperparameters to accelerate the model convergence and improve the training efficiency.

7) Uncertainty Perception: In order to solve the problem of 6.2.1, the output of the lane detection model can be changed from deterministic to probabilistic, which means that the probability distribution of each pixel or lane is outputted instead of a simple binary classification result. It can better reflect the model's uncertainty about detection results. Develop new methods for estimating uncertainty in lane detection models to simulate and estimate uncertainty in detection results.

In addition, when the vehicle faces unknown, randomly occurring scenarios or complex lane structures, ADS needs to have more decision-making capabilities to ensure that it is able to make informed control maneuvers. In this case, the lane detection model not only accurately recognizes lanes but also understands and responds to complex traffic environments. So active learning [156] and incremental learning [157] for roads is very important. In road scenarios, active learning can select informative or challenging scenarios to help the model better adapt to different road conditions and changes. For example, the system can collect data on rainy days, at night, or in congested traffic situations to improve the model's ability to understand and generalize to these complex scenarios. Incremental learning updates the model by introducing new data after the model has already been trained, rather than retraining the entire model. It can be used to quickly adapt to new road signs, lane line configurations, or changes in traffic rules. To build a human-machine interaction system that incorporates feedback and decision-making from the human driver before fully automated driving is achieved. Enable the autonomous driving system to adjust the output of the lane detection model according to the driver's intention and feedback.

8) New Technological Breakthroughs: After John McCarthy, the father of Artificial Intelligence, introduced Artificial Intelligence at the Dartmouth Conference in 1956, AI gradually evolved. But later, as it is difficult to improve the efficiency of the attribution method, it enters into a dilemma. Expert systems and neural networks, including RNN [158], CNN [34], and GNN [159], later emerged, but overall it was not a hot state of affairs. Until the introduction of AlexNet [160] in 2012 ushered in a historic breakthrough in AI algorithms and kicked off the AI boom. OpenAI releases ChatGPT [148] in 2022, once again producing a breakthrough in the development of artificial intelligence. Hough [161], CNN [34], and Transformer [17] have been successively applied to lane detection. In this decade CNN has driven lane detection forward rapidly, and in the future Transformer and other new theories will drive it forward as well.

Lane detection has also evolved without a mathematical foundation. Early lane detection is realized by parabola, likelihood function [161], and so on. Later, as the amount of data increases and the complexity increases, people also through mathematical principles to reduce the computational complexity [162]. For example, FDA [163] uses a combination of Fourier series to estimate the gradient of a symbolic function in the frequency domain to train a BNN, Wave-MLP [164] introduces complex numbers into neural networks by treating each token as a wave with an amplitude and phase, and mamba [165] will apply state space to deep learning. In addition, the limitation of the development of artificial intelligence is the computational power. From the CPU and GPU update to look at the development of artificial intelligence, it can be seen that when the computational power is improved, AI becomes more "intelligent". Improving the arithmetic power of lane detection computational devices is equally effective in increasing detection results. And the above described are not single but integrated and mutually reinforcing. We can safely anticipate that any significant breakthroughs in the creation of new AI models, applications of mathematical theories, or the introduction of novel computational tools will significantly impact the field as a whole. We believe that in this era of rapid technological development, "new AlexNet" and "new ChatGPT" will appear shortly. Of course, there are many more breakthrough points, such as new materials, new sensing devices, and new communication technologies.

9) Actual Utilization: With the development of autonomous driving technology, lane detection systems need to be more automated and intelligent. In practical use, the first condition is stable, accurate, and fast. In detail, it means that it is running stably, with a high level of accuracy and fast operational computing power. Future lane detection systems may be integrated with other sensors and modules to enable more advanced driver assistance functions such as automatic lane changing and automatic parking. There are many instances where various sensors work together, running multiple tasks simultaneously. Like Xiaomi intelligent driving high with the chip is Nvidia double Orin-X, combined arithmetic 508TOPS. The low version uses Orin-N, single chip arithmetic 84TOPS [166]. Only a small portion of these computational resources are used by lane detection. As a result, future lane detection systems need to be lower cost and more efficient.

On real-world roads, there are various types of lane lines, such as white solid dashed lines, double yellow lines, etc., and these markings are crucial for vehicle movement. However, most of the current lane detection only detects the location of the lane lines and does not give a reasonable explanation of the type of lane lines, which is precisely an indispensable part of autonomous driving. Moreover, it's vital to contemplate the potential future applications of this data within control systems, the intended format for the data transmission structure, and to assess scenarios such as dashed lines turning into solid lines at intersections, or a predicted growth in the number of lanes. All of these elements are topics for subsequent research.

Furthermore, UniAD [167] proposes an integrated framework for establishing a complete perception-making process. It integrates full-stack driving tasks into a single network, enabling more efficient exchange of information and coordinating perceptual prediction decisions to further enhance path planning. It makes for a revolutionary moment in autonomous driving: End-to-end Autonomous Driving! For example, the Tesla FSD v12 directly reduced the amount of code from over 300,000 lines to 2,000 lines. This may mean that lane detection will be moving in that direction.

10) Social Situation: Demand for lane detection technology as a means of enhancing driving safety and comfort is set to increase further in the future, attracting increased attention and significant market interest. Many countries and regions have established relevant traffic regulations and standards that require vehicles to have certain driver assistance systems, such as Lane Keeping Assist (LKA). Through regulatory requirements, market demand, safety needs, and technological advances, lane detection will continue to move forward, increasing its value and impact in the application of autonomous driving.

VII. Conclusion

In this survey, we offer an in-depth exploration of lane detection technologies employed in autonomous driving technology. Initially, we outline the progression of autonomous driving technology, including the requisite lane detection task and its associated challenges. Next, we provide an in-depth study of lane detection algorithms (both 2D and 3D) from the past decade, which focus mainly on visual detection. Furthermore, we present a comprehensive review of lane detection datasets, highlighting six prominent datasets from which existing methods are evaluated, thereby providing readers with insights into the optimal performance and potential optimization strategies. Finally, we present the current challenges that lane detection technology faces today and forecast possible future directions of advancement. In summary, lane detection not only allows for precise positioning of cars within a lane, but also provides an important basis for subsequent lane departure warning and trajectory planning. We hope that this review can benefit the community and serve as an insightful guide for future lane detection research.

References

[1] *Global Status Report on Road Safety 2023: Summary*, WHO, Geneva, Switzerland, 2023.

[2] J. Ayoub, F. Zhou, S. Bao, and X. J. Yang, "From manual driving to automated driving: A review of 10 years of AutoUI," in *Proc. 11th Int. Conf. Automot. User Interfaces Interact. Veh. Appl.*, Sep. 2019, pp. 70–90.

[3] E. Marti, M. A. de Miguel, F. Garcia, and J. Perez, "A review of sensor technologies for perception in automated driving," *IEEE Intell. Transp. Syst. Mag.*, vol. 11, no. 4, pp. 94–108, Apr. 2019.

[4] S. Behere and M. Törngren, "A functional architecture for autonomous driving," in *Proc. 1st Int. Workshop Automot. Softw. Archit. (WASA)*, May 2015, pp. 3–10.

[5] L. Chen et al., "Milestones in autonomous driving and intelligent vehicles—Part I: Control, computing system design, communication, HD map, testing, and human behaviors," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 53, no. 9, pp. 5831–5847, Sep. 2023.

[6] X. Wang, K. Li, and A. Chehri, "Multi-sensor fusion technology for 3D object detection in autonomous driving: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 2, pp. 1148–1165, Feb. 2024.

[7] H. Zhu, K.-V. Yuen, L. Mihaylova, and H. Leung, "Overview of environment perception for intelligent vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 10, pp. 2584–2601, Oct. 2017.

[8] W. Schwarting, J. Alonso-Mora, and D. Rus, "Planning and decision-making for autonomous vehicles," *Annu. Rev. Control Robot. Auton. Syst.*, vol. 1, no. 1, pp. 187–210, 2018.

[9] Y. Song, J. Bi, and X. Wang, "Design and implementation of intelligent monitoring system for agricultural environment in IoT," *Internet Things*, vol. 25, Apr. 2024, Art. no. 101029.

[10] C. K. Chandni, V. V. S. Variyar, and K. Guruvayurappan, "Vision based closed loop pid controller design and implementation for autonomous car," in *Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI)*, Sep. 2017, pp. 1928–1933.

[11] S. P. Narote, P. N. Bhujbal, A. S. Narote, and D. M. Dhane, "A review of recent advances in lane detection and departure warning system," *Pattern Recognit.*, vol. 73, pp. 216–234, Jan. 2018.

[12] A. B. Hillel, R. Lerner, D. Levi, and G. Raz, "Recent progress in road and lane detection: A survey," *Mach. Vis. Appl.*, vol. 25, no. 3, pp. 727–745, 2014.

[13] Y. Zhang, Z. Lu, X. Zhang, J.-H. Xue, and Q. Liao, "Deep learning in lane marking detection: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 5976–5992, Jul. 2022.

[14] Q. Li, L. Chen, M. Li, S.-L. Shaw, and A. Nüchter, "A sensor-fusion drivable-region and lane-detection system for autonomous vehicle navigation in challenging road scenarios," *IEEE Trans. Veh. Technol.*, vol. 63, no. 2, pp. 540–555, Feb. 2013.

[15] N. Garnett, R. Cohen, T. Pe'er, R. Lahav, and D. Levi, "3D-LaneNet: End-to-end 3D multiple lane detection," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, Sep. 2019, pp. 2921–2930.

[16] Y. Guo et al., "Gen-LaneNet: A generalized and scalable approach for 3D lane detection," in *Proc. 16th Eur. Conf. Comput. Vis.*, Glasgow, U.K. Cham, Switzerland: Springer, Aug. 2020, pp. 666–681.

[17] L. Chen et al., "PersFormer: 3D lane detection via perspective transformer and the OpenLane benchmark," in *Proc. Eur. Conf. Comput. Vis.*, Jul. 2022, pp. 550–567.

[18] R. Okuda, Y. Kajiwara, and K. Terashima, "A survey of technical trend of ADAS and autonomous driving," in *Proc. Tech. Papers Int. Symp. VLSI Design, Autom. Test*, Apr. 2014, pp. 1–4.

[19] J. Kumagai, "Tech expeditions: Sand trap," *IEEE Spectr.*, vol. 41, no. 6, pp. 44–50, Jun. 2004.

[20] R. Wang, J. Qin, K. Li, Y. Li, D. Cao, and J. Xu, "BEV-LaneDet: An efficient 3D lane detection based on virtual camera via key-points," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 1002–1011.

[21] Q. Huang and J. Liu, "Practical limitations of lane detection algorithm based on Hough transform in challenging scenarios," *Int. J. Adv. Robotic Syst.*, vol. 18, no. 2, Mar. 2021, Art. no. 17298814211008752.

[22] H. Gajjar, S. Sanyal, and M. Shah, "A comprehensive study on lane detecting autonomous car using computer vision," *Expert Syst. Appl.*, vol. 233, Dec. 2023, Art. no. 120929.

[23] G. Gupta, "Algorithm for image processing using improved median filter and comparison of mean, median and improved median filter," *Int. J. Soft Comput. Eng.*, vol. 1, no. 5, pp. 304–311, 2011.

[24] H. Hwang and R. A. Haddad, "Adaptive median filters: New algorithms and results," *IEEE Trans. Image Process.*, vol. 4, no. 4, pp. 499–502, Apr. 1995.

[25] M. Elad, "On the origin of the bilateral filter and ways to improve it," *IEEE Trans. Image Process.*, vol. 11, no. 10, pp. 1141–1151, Oct. 2002.

[26] G. Deng and L. W. Cahill, "An adaptive Gaussian filter for noise reduction and edge detection," in *Proc. IEEE Conf. Rec. Nucl. Sci. Symp. Med. Imag. Conf.*, Oct. 1993, pp. 1615–1619.

[27] W. Gao, X. Zhang, L. Yang, and H. Liu, "An improved Sobel edge detection," in *Proc. 3rd Int. Conf. Comput. Sci. Inf. Technol.*, vol. 5, Jul. 2010, pp. 67–71.

[28] P. Bao, L. Zhang, and X. Wu, "Canny edge detection enhancement by scale multiplication," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 9, pp. 1485–1490, Sep. 2005.

[29] 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.

[30] B. Dorj, S. Hossain, and D.-J. Lee, "Highly curved lane detection algorithms based on Kalman filter," *Appl. Sci.*, vol. 10, no. 7, p. 2372, Mar. 2020.

[31] X. Yuan, J.-F. Martínez, M. Eckert, and L. López-Santidrián, "An improved Otsu threshold segmentation method for underwater simultaneous localization and mapping-based navigation," *Sensors*, vol. 16, no. 7, p. 1148, Jul. 2016.

[32] C. Rosito Jung and C. R. Kelber, "A lane departure warning system based on a linear-parabolic lane model," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2004, pp. 891–895.

[33] L. Luo, D. Xu, Z. Zhang, J. Zhang, and W. Qu, "A fast and robust circle detection method using perpendicular bisector of chords," in *Proc. 25th Chin. Control Decis. Conf. (CCDC)*, May 2013, pp. 2856–2860.

[34] J. Tang, S. Li, and P. Liu, "A review of lane detection methods based on deep learning," *Pattern Recognit.*, vol. 111, Mar. 2021, Art. no. 107623.

[35] D. Neven, B. D. Brabandere, S. Georgoulis, M. Proesmans, and L. V. Gool, "Towards end-to-end lane detection: An instance segmentation approach," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2018, pp. 286–291.

[36] X. Pan, J. Shi, P. Luo, X. Wang, and X. Tang, "Spatial as deep: Spatial CNN for traffic scene understanding," in *Proc. AAAI Conf. Artif. Intell.*, Apr. 2018, vol. 32, no. 1, pp. 1–8.

[37] T. Zheng et al., "RESA: Recurrent feature-shift aggregator for lane detection," in *Proc. AAAI Conf. Artif. Intell.*, 2021, pp. 3547–3554.

[38] X. Shen, Z. Lu, Y. Zhang, and J. Xue, "Lane line detection based on parallel spatial separation convolution," in *Proc. BMVC*, 2021, pp. 1–13.

[39] H. Abualsaud, S. Liu, D. B. Lu, K. Situ, A. Rangesh, and M. M. Trivedi, "LaneAF: Robust multi-lane detection with affinity fields," *IEEE Robot. Autom. Lett.*, vol. 6, no. 4, pp. 7477–7484, Oct. 2021.

[40] J. Han et al., "Laneformer: Object-aware row-column transformers for lane detection," in *Proc. AAAI Conf. Artif. Intell.*, 2022, pp. 799–807.

[41] J. Yang, L. Zhang, and H. Lu, "Lane detection with versatile AtrousFormer and local semantic guidance," *Pattern Recognit.*, vol. 133, Jan. 2023, Art. no. 109053.

[42] Y. Hou, Z. Ma, C. Liu, and C. C. Loy, "Learning lightweight lane detection CNNs by self attention distillation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 1013–1021.

[43] Y. Hou, Z. Ma, C. Liu, T.-W. Hui, and C. C. Loy, "Inter-region affinity distillation for road marking segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 12483–12492.

[44] M. Ghafoorian, C. Nugteren, N. Baka, O. Booij, and M. Hofmann, "EL-GAN: Embedding loss driven generative adversarial networks for lane detection," in *Proc. Eur. Conf. Comput. Vis. (ECCV) Workshops*, Jan. 2019, pp. 256–272.

[45] Y. Zhang, Z. Lu, D. Ma, J.-H. Xue, and Q. Liao, "Ripple-GAN: Lane line detection with ripple lane line detection network and Wasserstein GAN," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 3, pp. 1532–1542, Mar. 2021.

[46] T. Liu, Z. Chen, Y. Yang, Z. Wu, and H. Li, "Lane detection in low-light conditions using an efficient data enhancement: Light conditions style transfer," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Oct. 2020, pp. 1394–1399.

[47] D. Wu et al., "YOLOP: You only look once for panoptic driving perception," *Mach. Intell. Res.*, vol. 19, no. 6, pp. 550–562, Dec. 2022.

[48] C. Chang et al., "Q-YOLOP: Quantization-aware you only look once for panoptic driving perception," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)*, Jul. 2023, pp. 52–56.

[49] J. Wang, Q. M. Jonathan Wu, and N. Zhang, "You only look at once for real-time and generic multi-task," 2023, *arXiv:2310.01641*.

[50] S. Lee et al., "VPGNet: Vanishing point guided network for lane and road marking detection and recognition," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 1947–1955.

[51] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440.

[52] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. 18th Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, vol. 9351, Munich, Germany. Cham, Switzerland: Springer, 2015, pp. 234–241.

[53] A. Paszke, A. Chaurasia, S. Kim, and E. Culurciello, "ENet: A deep neural network architecture for real-time semantic segmentation," 2016, *arXiv:1606.02147*.

[54] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, Apr. 2017.

[55] S.-Y. Lo, H.-M. Hang, S.-W. Chan, and J.-J. Lin, "Multi-class lane semantic segmentation using efficient convolutional networks," in *Proc. IEEE 21st Int. Workshop Multimedia Signal Process. (MMSP)*, Sep. 2019, pp. 1–6.

[56] Q. Wang, T. Han, Z. Qin, J. Gao, and X. Li, "Multitask attention network for lane detection and fitting," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 3, pp. 1066–1078, Mar. 2022.

[57] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2D pose estimation using part affinity fields," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 7291–7299.

[58] A. Vaswani et al., "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, Jun. 2017, pp. 5998–6008.

[59] X. Wang, L. Sun, A. Chehri, and Y. Song, "A review of GAN-based super-resolution reconstruction for optical remote sensing images," *Remote Sens.*, vol. 15, no. 20, p. 5062, Oct. 2023.

[60] Z. Chen, Q. Liu, and C. Lian, "PointLaneNet: Efficient end-to-end CNNs for accurate real-time lane detection," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2019, pp. 2563–2568.

[61] X. Li, J. Li, X. Hu, and J. Yang, "Line-CNN: End-to-end traffic line detection with line proposal unit," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 1, pp. 248–258, Jan. 2020.

[62] D. Jin, W. Park, S.-G. Jeong, H. Kwon, and C.-S. Kim, "Eigenlanes: Data-driven lane descriptors for structurally diverse lanes," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 17163–17171.

[63] L. Xiao, X. Li, S. Yang, and W. Yang, "ADNet: Lane shape prediction via anchor decomposition," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2023, pp. 6381–6390.

[64] K. Zhou and R. Zhou, "End-to-end lane detection with one-to-several transformer," 2023, *arXiv:2305.00675*.

[65] J. Liu et al., "Sparse laneformer," 2024, *arXiv:2404.07821*.

[66] Z. Qin, W. Huanyu, and X. Li, "Ultra fast structure-aware deep lane detection," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Glasgow, U.K. Cham, Switzerland: Springer, 2020, pp. 276–291.

[67] S. Yoo et al., "End-to-end lane marker detection via row-wise classification," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2020, pp. 1006–1007.

[68] H. Xu, S. Wang, X. Cai, W. Zhang, X. Liang, and Z. Li, "CurveLane-NAS: Unifying lane-sensitive architecture search and adaptive point blending," in *Proc. 16th Eur. Conf. Comput. Vis. (ECCV)*, Glasgow, U.K. Cham, Switzerland: Springer, 2020, pp. 689–704.

[69] L. Liu, X. Chen, S. Zhu, and P. Tan, "CondLaneNet: A top-to-down lane detection framework based on conditional convolution," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 3773–3782.

[70] L. Tabelini, R. Berriel, T. M. Paix ao, C. Badue, A. F. D. Souza, and T. Oliveira-Santos, "Keep your eyes on the lane: Real-time attention-guided lane detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 294–302.

[71] Z. Qin, P. Zhang, and X. Li, "Ultra fast deep lane detection with hybrid anchor driven ordinal classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 5, pp. 2555–2568, May 2024.

[72] Z. Yang et al., "CANet: Curved guide line network with adaptive decoder for lane detection," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2023, pp. 1–5.

[73] T. Zheng et al., "CLRNet: Cross layer refinement network for lane detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 898–907.

[74] Z. Chen, Y. Liu, M. Gong, B. Du, G. Qian, and K. Smith-Miles, "Generating dynamic kernels via transformers for lane detection," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2023, pp. 6812–6821.

[75] H. Honda and Y. Uchida, "CLRerNet: Improving confidence of lane detection with LaneIoU," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, Jan. 2024, pp. 1176–1185.

[76] C. Chen, J. Liu, C. Zhou, J. Tang, and G. Wu, "Sketch and refine: Towards fast and accurate lane detection," 2024, *arXiv:2401.14729*.

[77] J. Su, C. Chen, K. Zhang, J. Luo, X. Wei, and X. Wei, "Structure guided lane detection," 2021, *arXiv:2105.05403*.

[78] W. Cui, Y. Wang, S. Kang, J. Xie, Q. Wang, and M. V. Ivanovich, "Road lane line detection method based on improved yolov3 algorithm," *Acta Autom. Sinica*, vol. 48, no. 6, pp. 1560–1568, 2022.

[79] Y. Song et al., "A novel lane line detection algorithm for driverless geographic information perception using mixed-attention mechanism ResNet and row anchor classification," *ISPRS Int. J. Geo-Inf.*, vol. 12, no. 3, p. 132, Mar. 2023.

[80] Y. Ko, Y. Lee, S. Azam, F. Munir, M. Jeon, and W. Pedrycz, "Key points estimation and point instance segmentation approach for lane detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 8949–8958, Jul. 2022.

[81] Z. Zhao, Q. Wang, and X. Li, "Deep reinforcement learning based lane detection and localization," *Neurocomputing*, vol. 413, pp. 328–338, Nov. 2020.

[82] V. Mnih et al., "Playing Atari with deep reinforcement learning," 2013, *arXiv:1312.5602*.

[83] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Mach. Learn.*, vol. 8, pp. 279–292, May 1992.

[84] Z. Qu, H. Jin, Y. Zhou, Z. Yang, and W. Zhang, "Focus on local: Detecting lane marker from bottom up via key point," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 14122–14130.

[85] J. Wang et al., "A keypoint-based global association network for lane detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 1382–1391.

[86] S. Xu et al., "RCLane: Relay chain prediction for lane detection," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, Jan. 2022, pp. 461–477.

[87] J. Cao, X. Zhu, and C. Qian, "LanePtrNet: Revisiting lane detection as point voting and grouping on curves," 2024, *arXiv:2403.05155*.

[88] L. Tabelini, R. Berriel, T. M. Paix ao, C. Badue, A. F. De Souza, and T. Oliveira-Santos, "PolyLaneNet: Lane estimation via deep polynomial regression," in *Proc. 25th Int. Conf. Pattern Recognit. (ICPR)*, Jan. 2021, pp. 6150–6156.

[89] R. Liu, Z. Yuan, T. Liu, and Z. Xiong, "End-to-end lane shape prediction with transformers," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2021, pp. 3694–3702.

[90] Z. Feng, S. Guo, X. Tan, K. Xu, M. Wang, and L. Ma, "Rethinking efficient lane detection via curve modeling," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 17062–17070.

[91] K. Zhou, "Lane2Seq: Towards unified lane detection via sequence generation," 2024, *arXiv:2402.17172*.

[92] W. Han and J. Shen, "Decoupling the curve modeling and pavement regression for lane detection," 2023, *arXiv:2309.10533*.

[93] H. Xiong et al., "Fast and robust approaches for lane detection using multi-camera fusion in complex scenes," *IET Intell. Transp. Syst.*, vol. 14, no. 12, pp. 1582–1593, Dec. 2020.

[94] Q. N. Van, M. Yoon, W. Che, D. Yun, H. Kim, and K. Boo, "A study on real time integrated lane detection and vehicle tracking method with side-mirror cameras," in *Proc. IEEE 14th Int. Workshop Adv. Motion Control (AMC)*, Apr. 2016, pp. 346–352.

[95] Q. Zhang, J. Liu, and X. Jiang, "Lane detection algorithm in curves based on multi-sensor fusion," *Sensors*, vol. 23, no. 12, p. 5751, Jun. 2023.

[96] T. Yuan, W. Cao, S. Zhang, K. Yang, M. Schoen, and B. Duraisamy, "Lane detection and estimation from surround view camera sensing systems," in *Proc. IEEE SENSORS*, Oct. 2023, pp. 1–4.

[97] J. Han, D. Kim, M. Lee, and M. Sunwoo, "Road boundary detection and tracking for structured and unstructured roads using a 2D LiDAR sensor," *Int. J. Automot. Technol.*, vol. 15, no. 4, pp. 611–623, Jun. 2014.

[98] J. Wu, H. Xu, and J. Zhao, "Automatic lane identification using the roadside LiDAR sensors," *IEEE Intell. Transp. Syst. Mag.*, vol. 12, no. 1, pp. 25–34, Spring. 2020.

[99] W. Zhang, "LiDAR-based road and road-edge detection," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2010, pp. 845–848.

[100] J. Han, D. Kim, M. Lee, and M. Sunwoo, "Enhanced road boundary and obstacle detection using a downward-looking LiDAR sensor," in *Proc. IEEE Trans. Veh. Technol.*, Jan. 2012, vol. 61, no. 3, pp. 971–985.

[101] N. Efrat, M. Bluvstein, S. Oron, D. Levi, N. Garnett, and B. E. Shlomo, "3D-LaneNet+: Anchor free lane detection using a semi-local representation," 2020, *arXiv:2011.01535*.

[102] A. Zoljodi, M. Loni, S. Abadijou, M. Alibeigi, and M. Daneshtalab, "3DLaneNAS: Neural architecture search for accurate and light-weight 3D lane detection," in *Proc. Int. Conf. Artif. Neural Netw.* Cham, Switzerland: Springer, Jan. 2022, pp. 404–415.

[103] R. Liu, D. Chen, T. Liu, Z. Xiong, and Z. Yuan, "Learning to predict 3D lane shape and camera pose from a single image via geometry constraints," in *Proc. AAAI Conf. Artif. Intell.*, Jun. 2022, vol. 36, no. 2, pp. 1765–1772.

[104] F. Yan et al., "ONCE-3DLanes: Building monocular 3D lane detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Apr. 2022, pp. 17143–17152.

[105] C. Li, J. Shi, Y. Wang, and G. Cheng, "Reconstruct from top view: A 3D lane detection approach based on geometry structure prior," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2022, pp. 4369–4378.

[106] Y. Wang, Q. Guo, P. Lin, G. Cheng, and J. Wu, "Spatio-temporal fusion-based monocular 3D lane detection," in *Proc. BMVC*, 2022, p. 314.

[107] Y. Liu et al., "PETRv2: A unified framework for 3D perception from multi-camera images," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2023, pp. 3239–3249.

[108] Z. Li et al., "GroupLane: End-to-end 3D lane detection with channel-wise grouping," 2023, *arXiv:2307.09472*.

[109] Y. Bai, Z. Chen, Z. Fu, L. Peng, P. Liang, and E. Cheng, "CurveFormer: 3D lane detection by curve propagation with curve queries and attention," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2023, pp. 7062–7068.

[110] Y. Luo et al., "LATR: 3D lane detection from monocular images with transformer," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2023, pp. 7907–7918.

[111] S. Huang et al., "Anchor3DLane: Learning to regress 3D anchors for monocular 3D lane detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 17451–17460.

[112] N. Kim, M. Byeon, D. Ji, and D. Oh, "D-3DLD: Depth-aware voxel space mapping for monocular 3D lane detection with uncertainty," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2023, pp. 1–5.

[113] Z. Chen, K. Smith-Miles, B. Du, G. Qian, and M. Gong, "An efficient transformer for simultaneous learning of BEV and lane representations in 3D lane detection," 2023, *arXiv:2306.04927*.

[114] J. Ai, W. Ding, J. Zhao, and J. Zhong, "WS-3D-lane: Weakly supervised 3D lane detection with 2D lane labels," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2023, pp. 5595–5601.

[115] M. Hassoubah and G. Sistu, "Data driven 3D-lane detection using parallelism loss function," *J. Image Graph.*, vol. 12, no. 1, pp. 16–22, 2024.

[116] Y. Bai, Z. Chen, P. Liang, and E. Cheng, "CurveFormer++: 3D lane detection by curve propagation with temporal curve queries and attention," 2024, *arXiv:2402.06423*.

[117] Y. Luo et al., "M^2-3DLanenet: Exploring multi-modal 3D lane detection," 2022, *arXiv:2209.05996*.

[118] Y. Luo, S. Cui, and Z. Li, "Dv-3DLane: End-to-end multi-modal 3D lane detection with dual-view representation," in *Proc. 12th Int. Conf. Learn. Represent.*, 2023, pp. 1–13.

[119] L. Caltagirone, M. Bellone, L. Svensson, and M. Wahde, "LiDAR–camera fusion for road detection using fully convolutional neural networks," *Robot. Auto. Syst.*, vol. 111, pp. 125–131, Jan. 2019.

[120] Y. Yeniaydin and K. W. Schmidt, "Sensor fusion of a camera and 2D LiDAR for lane detection," in *Proc. 27th Signal Process. Commun. Appl. Conf. (SIU)*, Apr. 2019, pp. 1–4.

[121] H. Li et al., "Delving into the devils of bird's-eye-view perception: A review, evaluation and recipe," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 4, pp. 2151–2170, Apr. 2024.

[122] J. Huang, P. K. Choudhury, S. Yin, and L. Zhu, "Real-time road curb and lane detection for autonomous driving using LiDAR point clouds," *IEEE Access*, vol. 9, pp. 144940–144951, 2021.

[123] A. Hata and D. Wolf, "Road marking detection using LiDAR reflective intensity data and its application to vehicle localization," in *Proc. 17th Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Qingdao, China, Oct. 2014, pp. 584–589.

[124] L. D. P. Veronese, A. Ismail, V. Narayan, and M. Schulze, "An accurate and computational efficient system for detecting and classifying ego and sides lanes using LiDAR," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2018, pp. 1476–1483.

[125] H. Zeng et al., "ScatterHough: Automatic lane detection from noisy LiDAR data," *Sensors*, vol. 22, no. 14, p. 5424, Jul. 2022.

[126] Y. Jin, X. Ren, F. Chen, and W. Zhang, "Robust monocular 3D lane detection with dual attention," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2021, pp. 3348–3352.

[127] N. Efrat, M. Bluvstein, N. Garnett, D. Levi, S. Oron, and B. E. Shlomo, "Semi-local 3D lane detection and uncertainty estimation," 2020, *arXiv:2003.05257*.

[128] M. Li, P. M. Chu, and K. Cho, "Perspective transformer and MobileNets-based 3D lane detection from single 2D image," *Mathematics*, vol. 10, no. 19, p. 3697, Oct. 2022.

[129] C. Yao, L. Yu, Y. Wu, and Y. Jia, "Sparse point guided 3D lane detection," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2023, pp. 8329–8338.

[130] M. Bai, G. Mattyus, N. Homayounfar, S. Wang, S. K. Lakshmikanth, and R. Urtasun, "Deep multi-sensor lane detection," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 3102–3109.

[131] (2024). *Tusimple Dataset*. Accessed: Apr. 15, 2024. [Online]. Available: https://github.com/TuSimple/tusimple-benchmark

[132] M. Aly, "Real time detection of lane markers in urban streets," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2008, pp. 7–12.

[133] Y. Zhang et al., "VIL-100: A new dataset and a baseline model for video instance lane detection," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 15661–15670.

[134] F. Yu et al., "BDD100K: A diverse driving dataset for heterogeneous multitask learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 2636–2645.

[135] K. Behrendt and R. Soussan, "Unsupervised labeled lane markers using maps," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019, pp. 832–839.

[136] X. Huang, P. Wang, X. Cheng, D. Zhou, Q. Geng, and R. Yang, "The ApolloScape open dataset for autonomous driving and its application," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 10, pp. 2702–2719, Oct. 2020.

[137] X. Chen, W. Liao, B. Liu, J. Yan, and T. He, "Opendenselane: A dense LiDAR-based dataset for HD map construction," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2022, pp. 1–6.

[138] A. Loquercio, M. Segu, and D. Scaramuzza, “A general framework for uncertainty estimation in deep learning,” *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 3153–3160, Apr. 2020.

[139] F. Kamalov and H. H. Leung, “Deep learning regularization in imbalanced data,” in *Proc. Int. Conf. Commun., Comput., Cybersecurity, Informat. (CCCI)*, Nov. 2020, pp. 1–5.

[140] L. Wu et al., “R-Drop: Regularized dropout for neural networks,” in *Proc. Neural Inf. Process. Syst.*, vol. 34, 2021, pp. 10890–10905.

[141] H. Wang, C. Qin, Y. Zhang, and Y. Fu, “Neural pruning via growing regularization,” 2020, *arXiv:2012.09243*.

[142] C. A. Mehdi, J. Nour-Eddine, and E. Mohamed, “Check for updates regularization in CNN: A mathematical study for l1, l2 and dropout regularizers,” in *Proc. Int. Conf. Adv. Intell. Syst. Sustain. Develop.*, vol. 1. Cham, Switzerland: Springer, 2023, p. 442.

[143] L. Gao et al., “VECtor: A versatile event-centric benchmark for multisensor SLAM,” *IEEE Robot. Autom. Lett.*, vol. 7, no. 3, pp. 8217–8224, Jul. 2022.

[144] W.-T. Chen et al., “ALL snow removed: Single image desnowing algorithm using hierarchical dual-tree complex wavelet representation and contradict channel loss,” in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 4176–4185.

[145] C. Wang, Z. Zheng, R. Quan, Y. Sun, and Y. Yang, “Context-aware pretraining for efficient blind image decomposition,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 18186–18195.

[146] Y. Zhu et al., “Learning weather-general and weather-specific features for image restoration under multiple adverse weather conditions,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 21747–21758.

[147] X. Chen, H. Li, M. Li, and J. Pan, “Learning a sparse transformer network for effective image deraining,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 5896–5905.

[148] T. Wu et al., “A brief overview of ChatGPT: The history, status quo and potential future development,” *IEEE/CAA J. Autom. Sinica*, vol. 10, no. 5, pp. 1122–1136, May 2023.

[149] F.-L. Chen et al., “VLP: A survey on vision-language pre-training,” *Mach. Intell. Res.*, vol. 20, no. 1, pp. 38–56, Feb. 2023.

[150] W. Kim, B. Son, and I. Kim, “ViLT: Vision-and-language transformer without convolution or region supervision,” in *Proc. Int. Conf. Mach. Learn.*, 2021, pp. 5583–5594.

[151] A. Radford et al., “Learning transferable visual models from natural language supervision,” in *Proc. Int. Conf. Mach. Learn.*, vol. 139, 2021, pp. 8748–8763.

[152] A. Arooj, M. S. Farooq, A. Akram, R. Iqbal, A. Sharma, and G. Dhiman, “Big data processing and analysis in Internet of Vehicles: Architecture, taxonomy, and open research challenges,” *Arch. Comput. Methods Eng.*, vol. 29, no. 2, pp. 793–829, 2022.

[153] S. Costache, V. Gulisano, and M. Papatriantafilou, “Understanding the data-processing challenges in intelligent vehicular systems,” in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2016, pp. 611–618.

[154] Z. Hu and H. Tang, “Design and implementation of intelligent vehicle control system based on Internet of Things and intelligent transportation,” *Sci. Program.*, vol. 2022, pp. 1–11, Jan. 2022.

[155] F. Ma, Y. Liu, S. Wang, J. Wu, W. Qi, and M. Liu, “Self-supervised drivable area segmentation using LiDAR’s depth information for autonomous driving,” in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2023, pp. 41–48.

[156] D. Yuan et al., “Active learning for deep visual tracking,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 10, pp. 13284–13296, Oct. 2023.

[157] Y. Yuan, Z. Xiong, and Q. Wang, “An incremental framework for video-based traffic sign detection, tracking, and recognition,” *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 7, pp. 1918–1929, Jul. 2017.

[158] J. Li, X. Mei, D. Prokhorov, and D. Tao, “Deep neural network for structural prediction and lane detection in traffic scene,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 3, pp. 690–703, Mar. 2017.

[159] H. Sun, Z. Liu, S. Wang, and H. Wang, “Adaptive attention-based graph representation learning to detect phishing accounts on the Ethereum blockchain,” *IEEE Trans. Netw. Sci. Eng.*, vol. 11, no. 3, pp. 2963–2975, May 2024.

[160] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 25, 2012, pp. 1–11.

[161] Y. Wang, D. Shen, and E. K. Teoh, “Lane detection using spline model,” *Pattern Recognit. Lett.*, vol. 21, no. 8, pp. 677–689, Jul. 2000.

[162] Y. Matviychuk, N. Kryvinska, N. Shakhovska, and A. P. Maranda, “New principles of finding and removing elements of mathematical model for reducing computational and time complexity,” *Int. J. Grid Utility Comput.*, vol. 14, no. 4, pp. 400–410, 2023.

[163] Y. Xu, K. Han, C. Xu, Y. Tang, C. Xu, and Y. Wang, “Learning frequency domain approximation for binary neural networks,” in *Proc. Adv. Neural Inf. Process. Syst.*, Jan. 2021, pp. 25553–25565.

[164] Y. Tang et al., “An image patch is a wave: Phase-aware vision MLP,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 10935–10944.

[165] R. Xu, S. Yang, Y. Wang, Y. Cai, B. Du, and H. Chen, “Visual mamba: A survey and new outlooks,” 2024, *arXiv:2404.18861*.

[166] (2024). *Xiaomi Pilot*. Accessed: Apr. 22, 2024. [Online]. Available: https://www.xiaomiev.com/pilot

[167] Y. Hu et al., “Planning-oriented autonomous driving,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2023, pp. 17853–17862.

Jiping Bi was born in Weifang, Shandong, China, in 2001. He received the B.S. degree from the School of Computer and Control Engineering, Yantai University, in 2023, where he is currently pursuing the master’s degree. His current research interests include image processing, traffic target detection, and automatic control.

Yongchao Song was born in Weihai, Shandong, China, in 1990. He received the B.S. and Ph.D. degrees from the School of Electronic and Control Engineering, Chang’an University, Xi’an, China, in 2015 and 2020, respectively. He is currently an Associate Professor with the School of Computer and Control Engineering, Yantai University. His current research interests include remote sensing information processing and deep learning.

Yahong Jiang was born in Baoding, Hebei, China, in 1992. She is currently pursuing the Ph.D. degree in transportation planning and management with the School of Transportation Engineering, Chang’an University. Her current research interests include transportation policy evaluation and self-driving electric vehicle promotion.

Lijun Sun was born in Yantai, Shandong, China, in 2001. She received the B.S. degree from the School of Computer and Control Engineering, Yantai University, in 2023, where she is currently pursuing the master’s degree. Her current research interests include deep learning, remote sensing image super resolution, and artificial intelligence.

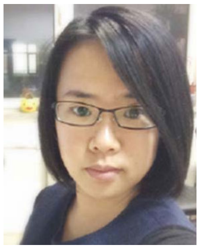
Xuan Wang (Senior Member, IEEE) was born in Weihai, Shandong, China, in 1991. She received the B.S. and Ph.D. degrees from the Traffic Information Engineering and Control, Chang’an University, China, in 2013 and 2018, respectively. She is currently an Associate Professor with the School of Computer Science and Control Engineering, Yantai University. Her research interests include intelligent traffic control, artificial intelligence, and computer vision.

Zhaowei Liu (Member, IEEE) received the B.S., M.S., and Ph.D. degrees from the College of Control Science and Engineering, Shandong University, Jinan, China, in 2002, 2005, and 2018, respectively. He is currently a Professor with the School of Computer Science and Control Engineering, Yantai University, Yantai, China. His research interests include graph learning and machine learning.

Jindong Xu received the B.S. degree in communication engineering from Shandong University, Jinan, China, in 2003, and the M.S. and Ph.D. degrees from the College of Information Science and Technology, Beijing Normal University, Beijing, China, in 2008 and 2014, respectively. From 2015 to 2017, he was a Visiting Scholar with the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing. He is currently a Professor with the School of Computer and Control Engineering, Yantai University, Yantai, China. His research interests include image processing, pattern recognition, and multi-view fusion.

Siwen Quan received the B.S. degree from Chang'an University, Xi'an, China, in 2015, and the Ph.D. degree from the Huazhong University of Science and Technology, Wuhan, China, in 2019. She is currently an Associate Professor with the School of Electronics and Control Engineering, Chang'an University. Her research interests include local geometric shape description, 3-D object recognition, and image fusion.

Zhe Dai was born in Xi'an, Shannxi, China, in 1993. He received the Ph.D. degree from the Traffic Information Engineering and Control, Chang'an University, China, in 2020. He is currently a Lecturer and Master's Tutor with Chang'an University. His research interests include intelligent transportation systems, research directions: 1) multi-source fusion traffic information perception; 2) transportation safety warning and control; and 3) multi-modal transportation network modeling.



Weiqing Yan (Senior Member, IEEE) received the Ph.D. degree in information and communication engineering from Tianjin University, Tianjin, China, in 2017. From 2015 to 2016, she was a Visiting Ph.D. Student with the Visual Spatial Perceived Laboratory, University of California at Berkeley, Berkeley, CA, USA. From 2022 to 2023, she was a Research Fellow with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. She is currently an Associate Professor with the School of Computer and Control Engineering, Yantai University, Yantai, China. Her research interests include 3D vision, multi-view clustering, and pattern recognition.