Multimedia Tools and Applications (2023) 82:33983–34004
https://doi.org/10.1007/s11042-023-14446-6

Lane detection techniques for self-driving vehicle: comprehensive review

Ashwini Sapkal[1] · **Arti**[1] · **Dishant Pawar**[1] · **Prashant Singh**[1]

Received: 6 August 2021 / Revised: 14 August 2022 / Accepted: 31 January 2023 /
Published online: 10 March 2023

Abstract

According to WHO, 1.35 million people, every year are cut short in road accidents, most of them caused due to human misconduct and ignorance. To improve safety over the roads, road perception and lane detection play a crucial part in avoiding accidents. Lane Detection is a constitution for various Advanced Driver Assisting System (ADAS) like Lane Keeping Assisting System (LKAS) and Lane Departure Warning System (LDWS). It also enables fully assistive and autonomous navigation in self-driving vehicles. Therefore, it has been an effective field of research for the past few decades, but various milestones are yet to be achieved. The problem has encountered various challenging scenarios due to the past limitations of resources and technologies. In this paper, we reviewed the different approaches based on image processing and computer vision that have revolutionized the lane detection problem. This paper also summarizes the different benchmark data sets for lane detection, evaluation criteria. We implemented Lane detection system using Unet and Segnet model and applied it on Tusimple dataset. The Unet performance is better as compared to Segnet model. We also compare the detection performance and running time of various methods, and conclude with some current challenges and future trends for deep learning-based lane marking detection algorithm. Finally, we compare various researcher's approaches with their performances. This paper concluded with the challenges to predict accurate lanes under different scenarios.

Keywords Autonomous driving · Lane detection · Deep learning · Advanced driver assisting system · Lane keeping assisting system · Lane departure warning system

1 Introduction

Advanced driver assisting systems (ADASs) are developed to reduce road accidents by assisting the driver. Lane detection is an essential module of ADAS that provides navigational assistance to the driver. A lane departure warning (LDW) system alerts a driver

Arti, Dishant Pawar and Prashant Singh are contributed equally to this work.

✉ Ashwini Sapkal
asapkal@aitpune.edu.in

Extended author information available on the last page of the article.

if the vehicle diverges from a lane or a narrow road on arterial roads and freeways. The LDW system assists the driver to reduce vehicle crashes that careless drivers cause. As the safety and comfort of passengers is the primary for developing ADAS, this has led to the creation of fully autonomous self-driving vehicles. Many automotive enterprises such as Tesla, Google, Lyft, Uber, etc. have developed their own lane detection models used in assistive autonomous navigation in driver-less cars that have achieved significant importance in research and the real world. Normally, highways are more structured and predictable, with road surfaces well-maintained and lanes well-marked. Contrary, residential or urban driving surrounding features are highly unpredictable with many generic objects, different kinds of lane markings, and complex traffic flow patterns. The comparative consistency and predictability of highways have encouraged some first real-life applications of automation in driving technology. Various automotive enterprises are pursuing highway auto-pilot methods to alleviate driver stress and enable more safety features for vehicles and passengers.

With the recent developments in high computing devices, sensors, computer vision techniques and machine learning theories, real-time lane detection problems are solved more realistically. There are many sensing modalities used for road and lane perception such as monocular vision (i.e. one video camera), stereo, LIDAR, real-time vehicle information secured from car odometry, inertial measurement unit (IMU) alongside global positioning information acquired from global positioning system (GPS) and digital maps. Computer vision is the leading research area in road and lane detection since markings are built for human vision, whereas GPS and LIDAR are prime supplements.

With automaticity, there comes an elevated risk therefore, reliability and robustness are the most essential properties of the system. The high-class automation vehicles should constantly observe their surroundings and must deal with low-accuracy detection problems. Therefore, the assessment of lane detection systems becomes more severe with an increase in automation at each level.

An advanced lane detection model should be robust enough to perform even in critical and diverse conditions such as:

- Poor visibility due to heavy shadows or wet roads.
- A curve or steep roads
- Change in road color or material
- Different slopes
- Continuous or broken lines.
- Environmental conditions such as fog, snow, or heavy rainfall
- Road mark degradation

Due to different environmental conditions and lane variations, gradual degradation of lane marking perfectly accomplished lane detection is very difficult. Therefore, many researchers have been attracted to this field. In the past decade, many sophisticated detection models have been introduced. Some of these were based on conventional geometric modeling techniques and some used machine learning-based approaches. The research efforts in this domain are made since 2005 in DARPA Grand Challenge (2005) and Urban Challenge (2007) and, yet more need to be achieved [11].

The detailed review of existing techniques used for lane detection is described in this paper. This paper covers several insights into recent research toward the goal of enabling secure and robust lane detection. It presents different techniques of lane detection including conventional geometry modelling and deep learning models. It also explores the commonly used benchmark datasets, evaluation metrics and performance evaluation of different techniques. There are many research challenges and limitations that needs to be addressed for accurate lane detection. Some of these challenges are listed in this paper. These challenges opens up the further research in this direction for providing effective driver assistance systems.

This paper specifically discusses the problem of lane detection in order to determine the correct lane position from the markings on the road. The appropriate information of the environment is extracted using a camera mounted on the vehicle. The lane location is then obtained from the image. The pipeline of lane detection model consists of the feature extraction from the road images followed by the classification of each pixel from the image as the lane or not. We have also implemented Lane detection system using Unet and Segnet model and applied it on Tusimple dataset. It is found that the Unet model outperforms as compared to Segnet model when applied on Tusimple Dataset.

The paper is oriented as follows: Section 2 provides the literature survey for the existing lane detection models. The section is subdivided into conventional and machine learning-based models. Section 3 is the discussion portion to analyze and interpret the Literature Survey and research existing gaps that need to be bridged. Finally, we conclude our work in Section 4.

2 Survey on existing models

Researchers faces many problems and challenges in implementing the automatic lane detection system. In Table 1 the critical parameters are listed which makes the lane detection task challenging

2.1 Conventional geometry modelling using lane detection algorithms

This approach particularly uses gradient, texture, boundaries, color, etc., for extracting information about lane markings from the image, followed by line detection and line fitting. Therefore, the pipeline of traditional lane detection models is mainly divided into three stages Image Preprocessing, Feature extraction and model fitting, and detected lane tracking as given in Fig. 1.

Table 1 Factors affecting lane detection

Factors	Its impact on Lane Detection
Lane and road	lane color, inconsistent road texture, Crossway, curve roads, degraded lane markings
Urban Roads	complex painted road surface markings, utility poles, and buildings
Town Roads	road markings, pedestrian sidewalk, and guardrails
Hardware	camera type, camera calibration correctness,camera mounting and positions, different sensor
Traffic	vehicles, shadows, lighting conditions, guardrail
Weather	rain, snow, fog

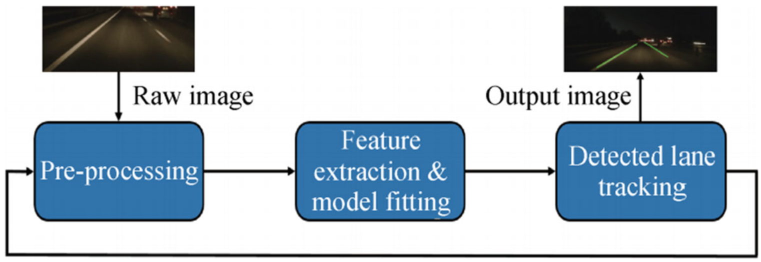


Fig. 1 General architecture of lane detection system

2.1.1 Image pre-processing

The advantage of pre-processing is to enhance image data, and certain features and remove noises that interfere with the accuracy of subsequent recognition of lane lines. In pre-processing the most commonly used methodologies are camera calibration correction, Color space transformation, noise removal, blurring, inverse perspective mapping (IPM) (birds-eye view), region of interest (ROI) selection, etc.

Camera calibration correction While developing a robust computer vision algorithm, various possible errors need to be addressed. Every camera inherits the lens distortion and occlusions properties during an object's 3D to 2D transformation and the camera image undergoes Radial Distortion and Tangential Distortion.

In Radial Distortion, the straight lines in the image are slightly curved or bent at the corners. In Tangential Distortion, the lens is not parallel configured to the image plane and generates a little extended or tilted image, making the objects look closer or further away.

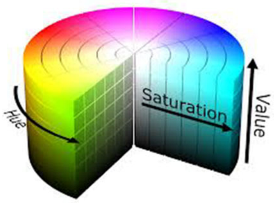
These distortions degrade the quality of data, henceforth the detection. Therefore, camera calibration is an important procedure.

Inverse perspective mapping (IPM) The camera has a perspective view as the human eye due to this property the lane lines seem to merge at a point called the vanishing point. Certain features when perceived from distinct viewpoints, they are identified with better precision. The road curvature lane marking is one such feature. Applying perspective transform would transform the perspective image captured by the camera to a "bird's eye view" of the road, which are mainly focused on the lane markings and present them relatively parallel to each other [1, 2, 7, 20, 46, 47]. This will further make it easier to fit polynomial and measure the curvature of the lane lines. Ju Han Yoo et al. [57] used vanishing point estimation for lane detection; they used inverse perspective transformation rather than IPM. Yingping et al. [26] set a simple trapezoidal transformation from the rectangular region obtained after ROI selection to produce the bird's eye view image (Figs. 2, 3, 4).

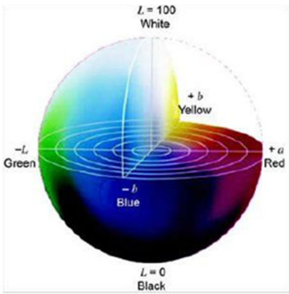
ROI selection Another method [19, 32, 33, 40] to remove noise from the image is by defining Regions of Interest (ROI) on the image plane. ROI selected from the probability of locating the lanes is highest and the remaining region is treated as noise and it is masked. This averts from locating unwanted features from the image. Zhang et al. [60] and Yingmao et al. [26] define the image's lower half as ROI in a simplified way. In [9, 44] the ROI was selected based on the calculated depths.

Fig. 2 Inverse perspective transformation

Color space transformation(CST) The significant intensity edges and heavy shades on the road surfaces are a primary source of clutter. To get rid of this illumination researchers [1, 15, 30] performed various Color Space Transformation (CST) such as HSL(Hue, Saturation and Lightness), LAB (L for lightness and A and B for the color dimensions), YCbCr, YIQ, and others [27, 54]. Illumination invariant images is obtained by combining different colors. This results in the same intensity over a surface in both the illuminated and shadowed areas. Further, to isolate lane marking different threshold masking is applied for both yellow and white color lane lines.



[HSL]



[LAB color space]

2.1.2 Lane feature extraction

In various lane detection systems, image gradient and boundary information are the frequently used image features for image segmentation. This lane feature extraction needs some computation and it finds details depending on the change of image pixel intensities. Well-painted lane markings contrast highly with the road surface, enabling the detection of lanes lines. In literature various authors evaluated the most commonly used feature extractors, such as edge detectors [53], top hat filters [5], steerable filters [37], global threshold [12], local threshold [8] and SLT [51].

The lane features ought not to be intense or degraded by shadow or external influential factors due to the environmental conditions. Therefore a strong and dominant edge detector is required. In contrast to other edge detection algorithms, the Canny edge detection algorithm is widely used to achieve the best feature extraction of the edges from the image [35, 45].

Canny detection The Canny Detector is a multi-stage algorithm for fast real-time edge detection [4, 32, 35, 45]. To detect edge from the given set of thresholds the algorithm detects intense changes in illumination in the image, like a shift from white to black. In

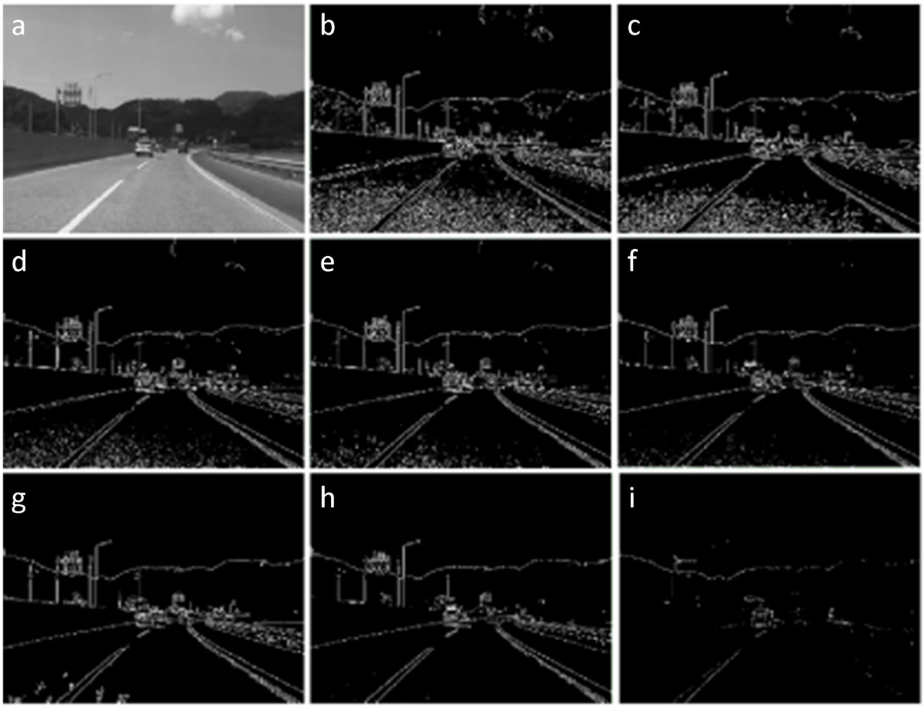


Fig. 3 Canny edge detection with different thresholds

lane feature extraction, the contrast of pixel intensity between the road surface and lane line defines the lane lines. The Gaussian filter is commonly used to remove and mask the noise and then lane features are extracted by Canny edge detection afterward, the lines are recognized from these edges. Figure 3 shows the output of canny detection at different thresholds (th1 and th2)(x is the grey scale image from a source b) th1=40, th2=40; c) th1=90, th2=40; d) th1=90, th2=90; e) th1=140, th2=90; f) th1=140, th24=140; g) th1=40, th2=240; h) th1=240, th2=40; i) th1=450, th2=50. The enigma for this technique is: higher the threshold value lower noises are detected. The crucial point is to apply the appropriate thresholds on the image to locate the correct position of the lane with the lane detection technique.

2.1.3 Line fitting

A simple lane model cannot illustrate lane shape precisely as the lines differ from straight segment to sharp curve. Whereas a complex lane model requires heavy computational power, thereby, making it inefficient in real-time and also, increasing the error detection rate. Therefore, lower computation and robustness are the essential properties to enable realistic lane detection. Various algorithms such as Least Squares Fitting, Random Sample Consensus (RANSAC), Spline Models, and Hough Transformation are used for line fitting.

Kalman filters The Kalman filter is considered as a solution to various tracking and data prediction tasks [7, 39, 45]. After lane marking detection, Kalman filter is commonly used to predict the lane in the next frame. To predict the future frame, the algorithm makes the noise estimation of the system over time to calculate the parameters of the system (which are unnoticed) and remembers the formerly detected lanes from the former video frame. At every stage, it makes a prediction, draws measurement, and modify itself. To add stability to lane markings, the variant of detected lines are averaged by adding up the measurement errors and previous state by the Kalman filter.

Least squares fitting Least-squares fitting (or least squares estimation) is a method to locate the best fit curve or line from a given set of points [45]. In this technique, instead of the absolute values of the residuals (offsets), the sum of squares of the residuals is used to determine the best-fit curve or line. But the disadvantage of this technique is the outlier's unbalanced effect that influences the line or curve equation. The reason behind this is the residuals squares are used rather than the absolute value of the residuals and the large outliers will influence the curve or line detection more than the points closer. Therefore, this approach cannot guarantee accurate lane findings.

RANSAC RANSAC algorithm is a robust algorithm that has generally been used in numerous computer-vision problems. Researchers have used RANSAC in the Lane Detection model to estimate lane model parameters. The lane features extracted in the previous stage are not distributed evenly and contain too much noise that hinders accurate line prediction. Thus, it is a computationally heavy task to accurately compute the model parameters on feature points with too much noise. RANSAC does not require a training process contrary to other algorithms such as Hough Transformation (HT) or template matching and it can even adjust to the challenging scenarios in model parameters estimation.

Vanishing point Due to the perspective effect of the camera in the process of 3D to 2D transformation images seem to converse to a point called vanishing Point, this property could be exploited to filter out false lane detection. For Lane detection, the intersection point of pair of line segments is computed as the vanishing point for lanes [27, 40, 54]. Liu et al. [34] use maximum intersection points of pairs of line segments detected from Line Segment Detector (LSD) [52] to identify vanishing points of lanes. Yuan et al. [58] estimated vanishing point with a probabilistic framework. The estimated vanishing point in the image space is converted into the approximate line parameters in the parametric space. Ju Ham et al. [57] In order to detect lanes from the input image, firstly the line segment is obtained by the LSD method [52], and for each line segment, the strength is calculated. Then, with the computed line segments strength the vanishing point is obtained using a probabilistic voting procedure.

Spline model Splines are piece wise polynomial functions used to draw curves from the given points. Usually, there is a sequence of the feature points and a curve is designed to smoothly transit (or pass-over) them. Discrete spline models with discrete functions and methodologies were used to identify lane boundaries. To ensure energy-based optimization Wang et al. [53] used the B-Splines model and M Aly et al. [3] used Active Contours (Snakes). Kim [29] uses the Cubic Splines model as the curve comprises the control points. The [18] Catmull Rom spline model is used to find a broader range of lane structure tracking. In most of these models, the curve is represented in parameters from the set of control points. In [26] 100 "seed" is chosen at a closer distance(in Perspective view) as they generally have lower noise points. The resulting spline is evaluated to expose every next point by

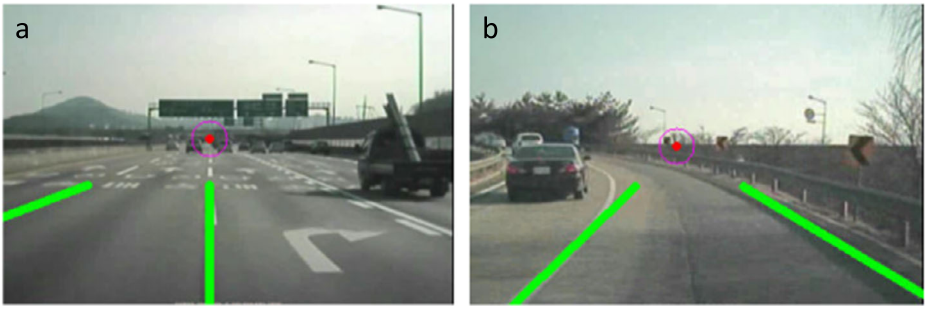


Fig. 4 Hough transformation outputs

a “greedy” search. At each stage evenly spaced control points are generated from the points ranging at 50 pixels from the previous point.

Hough transformation(HT) The HT is commonly used to isolate features of any shape within an image that can be written in a mathematical formula. A typical HT often identifies regular shapes such as lines, ellipses, circles, etc. The advantage of this technique is that it is more tolerant to noise. This technique identifies a line passing through collinear points by detecting the crossing of correspondent curves on the parametric plane in Hough Space. The unique equation for every possible line through the point of an image is counted, and then the dominant lines in an image are computed. Although the points are practically barely collinear and dispersed about the line, curves do not absolutely converge at a point.

Generally, to solve lane detection problems, the image first converted into a binary image using some thresholds then HT is applied to find the dominant lines [28, 45]. This technique is widely used as it also detects false positives in the presence of noise and occlusion.

2.2 Deep learning-based lane detection models

The Deep learning evolution has revolutionized the research on Lane detection. It has gained tremendous popularity as they outrun conventional computationally intensive image-processing-based models, especially in object detection, speech recognition, natural language processing, and image classification. Deep learning-based models do not require additional preprocessing and implicitly learn the features while training the model.

2.2.1 Encoder and decoder model based on CNN

The Convolutional Neural Network (CNN) is one of the most approved and used methods for lane detection. To inhabit more accurate detection, many researchers have integrated CNN and other deep learning techniques [19, 33]. CNN is computationally efficient and excellent in handling extensive unstructured data and has high detection accuracy and automatic feature learning. It was reported that compared to conventional methods CNN, improves the accuracy from 80% to 90% [21]. Therefore, numerous deep learning-based models were proposed. Here we present some of the models.

Lane detection using SegNet Rama Sai et al. [36] present an unique approach using CNN’s SegNet architectural model. As shown in Fig. 5 the encoder consists of convolutional layers, activation function (rectified linear unit/ ReLu), and a pooling layer. It generates

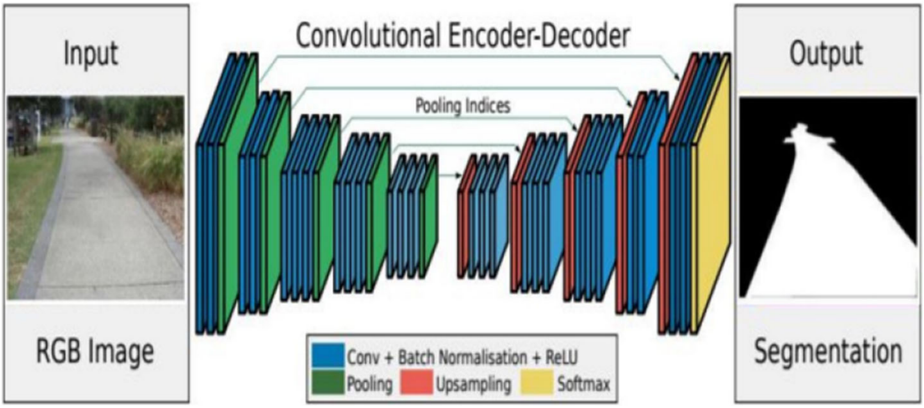


Fig. 5 Illustrating SegNet architecture

low-resolution feature maps from the input image. The output from the encoder is fed to the decoder that consists of deconvolutional layers to compute pixel-wise classification. To enable dynamic navigation, the model's prediction is interfaced with Google APIs such as Google geo-locate, Google maps, and Google street-view APIs. These APIs provide real–time information of the street. Google APIs allow to interface to Google cloud in order to make the applications powerful through the services like computing, storage, networking, etc. Google Cloud APIs are programmatic interfaces to Google Cloud Platform services. They are a key part of Google Cloud Platform, allowing you to easily add the power of everything from computing to networking to storage to machine-learning-based data analysis to your applications. The Encoder-Decoder Seg-net architecture (fully convolutional neural network) and Unet architecture is implemented.

Lane detection using Segnet The encoder-decoder framework model consists of a sequence of nonlinear processing layers (encoders) and a corresponding set of decoders followed by a pixel wise classifier to detect lane in a semantic pixel-wise segmentation manner. The Segnet model contain 13 convolutional layers in the VGG16 network and 5 pooling layers. The model takes single 128*256*3(where 3 is no. of channels) image size and outputs the single channel image which is combined with original image to generate the image of same size. The Encoder consist of convolutional layer with batch normalisation, a ReLU non-linearity and maxpooling layer of 2*2 with stride 2. It takes the input image and construct the feature map. Decoder consist of deconvolutional layers and upsampling layers, the output feature map from encoder is fed to decoder to map the low resolution encoder feature maps to full input resolution feature maps for pixel-wise classification. Specifically, the decoder uses pooling indices computed in the maxpooling step of the corresponding encoder to perform non-linear upsampling. This eliminates the need for learning to upsample. The upsampled maps are sparse and are then convolved with trainable filters to produce dense feature maps.

Lane detection using Unet In Unet, the contraction and expansion part gives the U shaped architecture. The contraction path is actually the encoder part of the model. The VGG network is used as the backbone. There are 10 convolutinal layer with rectifier linear (ReLu) is used. The Kernal size is 2*2, downconvolutional has 3*3 dimension. The Contraction path is convolutional network that consist of repeated application of convolutions. During contraction the saptial information is reduced while feature information has increased. The Expansion part is the decoder part of the model. It combines the feature and spatial information through a sequence of up-convolutions and concatenations with feature maps from the encoder part. The up-convolutions has 2*2 dimension and stride is 2 (Figs. 6 and 7).

The Lane detection system is implemented using Segnet and Unet Architecture. The dataset used in the implementation is Tusimple [49] which is explained in detail further. As shown in Table 2 the Unet performance is better than the Segnet.

2.2.2 GAN model

The generative adversarial network(GAN) [17] model is used for lane detection. The model comprises a generator and a discriminator for semantically separate out lanes from the image. Mohsen et al [17] use an embedding-loss GAN (EL-GAN). The generator produces a semantic segmentation map to predict the lanes from the input image, and the discriminator is used to differentiate produced data from ground truth labels based on weights. The GAN model is primary used to resolve the aforementioned issues with the per-pixel loss. The model avoids soft boundaries generally produced in CNN models, the final predicted lane is thin and accurate.

2.2.3 Dual-view convolutional neutral network (DVCNN)

Bei et al. [22] proposed a DVCNN framework where both perspective-view and the bird eye view images are optimized together for correct lane prediction. The lane detection problem is solved in three folds, initially from perspective-view image all obstacles like vehicles, barriers are excluded to avoid false detection, while in bird eye view generated by

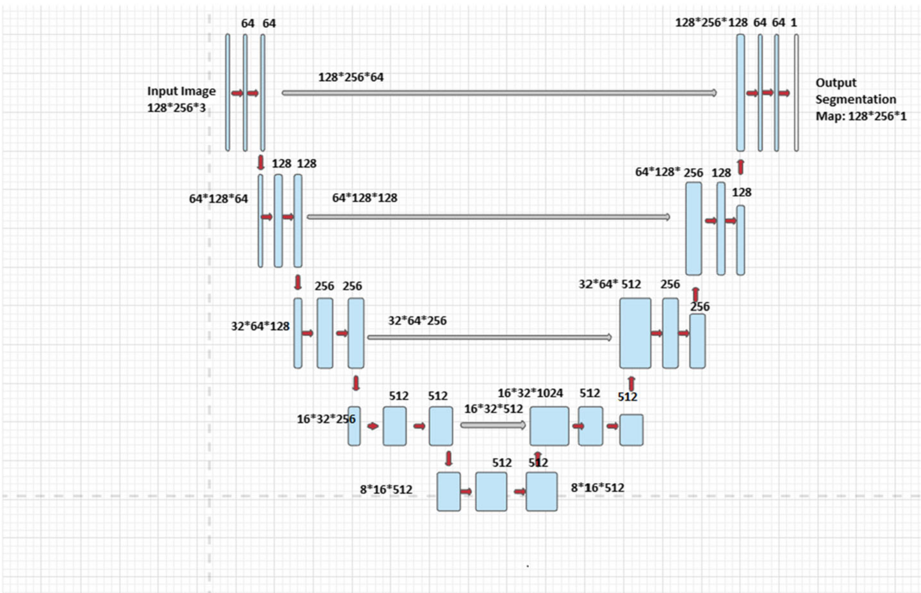


Fig. 6 Illustrating UNet architecture

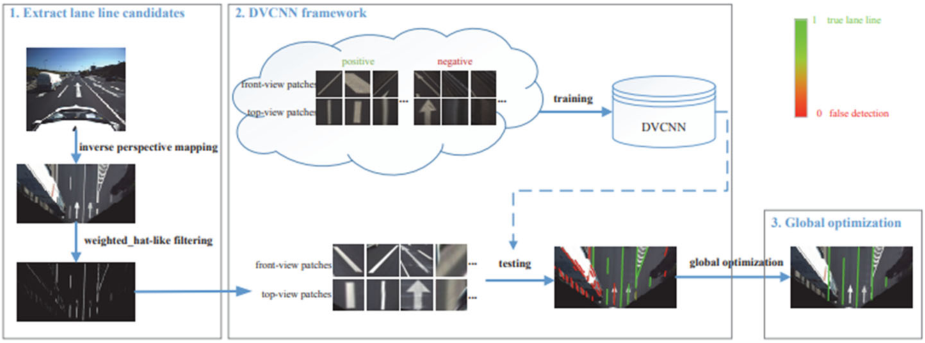


Fig. 7 Methodology for Lane Detection 1) lane line extraction 2) the DVCNN framework 3) the global optimization

perspective transformation mapping from perspective-view image non-club-shaped constructions such as ground arrows, words, etc. are removed. Secondly to extract lane candidate lines a weighted hat-like filter is created. This filter relieves the interference of gradual texture and reduces false detection. Afterward, the front-view and top-view images are the inputs to the DVCNN framework. Thirdly, a global optimization function is constructed to refine the final output using lane probability, lane geometry like length and width. After the optimization, the optimal fusion comprising correct lane lines is detected.

The average, recall, and precision ratios are 92.80% and 95.49% respectively. However, the algorithm fails when the lane lines are over-imposed by obstacles or vehicles.

2.2.4 Cloud point and CNN framework

Bei He et al. [23] suggested lane detection technique using CNN from the point cloud. CNN framework is designed to identify lane lines from point clouds. The road features are extracted using orthogonal projection and a top-view reflectivity image is obtained. The CNN architecture is proposed to distinguish candidate lane markings from reflective images. For accurate detection, the gradual up-sampling is applied to CNN. CNN cannot deploy the global information and domain knowledge, but the lane layout must be evaluated and classified into different groups. To reduce false alarms or detection ground arrows and texts are eliminated by mathematical and geometrical position or length restrictions (Figs. 8 and 9).

2.2.5 Fully connected layer

Alexandru [39] use two laterally-mounted cameras to the input image to fully convolutional deep learning model. They state that the front camera configuration does not furnish an

Table 2 Comparative analysis of lane detection models implemented using Segnet and Unet on tusimple dataset

Model Implemented	Training Accuracy	Validation Accuracy	Testing Accuracy
Segnet	95.81	95.19	95.176
Unet	97.93	97.44	97.527

Fig. 8 The schematic of framework

optimal lane marking sight, as it also records the entire front-facing driving scene, counting all vehicles or clusters. They produced 40,000 artificial image data that provide data augmentation, allows class distribution. The dataset contains images generated in diverse scenarios such as scattered, parallel, and partially interrupted lane markers. A few basic image processing transformations induced variance in the dataset by varying lighting conditions and shadows. Also, the pixel intensities are reduced by constant value in parts of an image to have imitated shadows cast on the lane markings. The network consists of convolutional layers, fully connected layers, and soft-max classification that operates on laterally mounted camera images. The network can predict the lane marking from the image with 97.85 accuracy%

3 Discussion on dataset and evaluation metrics

The Computer vision-based models have dominated the market and will continue due to low-cost hardware and substantial knowledge of image processing and machine learning techniques. Even though its wide range of applications in lane detection, vision-based models suffers from critical, challenging scenarios discussed previously. Many advanced machine learning and deep learning-based models were designed to obtain more reliability and robustness in real-time lane detection. Therefore, robustness is important to assess the system's maturity level and determine the loss ends.

Below Table 3 shows the different methods and its limitation. This table also justifies the significance of using deep learning-based model.

Researchers use various methodologies and assessment criteria to evaluate and benchmark the algorithms. As explained before, a lane detection problem is be broken down

Fig. 9 Sideways cameras focusing on lane lines

Table 3 Analysis of lane detection models

Techniques	Models	Merits	Demerits
Traditional	Geometric modeling with edge detection, line fitting	Works fine with simple lane detection	Does not work for challenging scenarios
Non Deep Learning Based	Uses perspective transformation, camera calibration, ROI	Better for curved roads	Computationally expensive, fail to detect the steep lanes
Deep Learning based	Encoder-decoder CNN Model, SegNet, U-Net	shows better results in challenging situations	Computationally expensive

into preprocessing, lane detection, and tracking. Correspondingly, assessment could be carried out in all 3 phases. Here we have divided the evaluation methods into two categories: Evaluation based on Dataset and evaluation metric.

3.1 Evaluation dataset

The evaluation of the model is broadly categorised in two section based on dataset used is offline or online.

3.1.1 Offline evaluation dataset

The designed lane detection model is generally evaluated offline by image dataset or continuous image frames (videos). There are various public datasets available such as TuSimple, Kitti, Caltech, Camvid (Cambridge-driving Labelled Video Database), Culane, BDD100k, VPGNet, and CurveLanes dataset. The below table describes different dataset used in image segmentation and object detection task

- TuSimple Dataset [49] collected from American Highway contains 7,000 one-second-long video clips of 20 frames. The Training set contains 3626 video clips, 3626 annotated frames, and the Testing set has 2782 video clips and the Testing set has 2782 video clips.
- There are 600 images in Kitti Dataset containing 290 training images and 300 testing images divided into three categories, representing challenging road scenarios. For both training and testing set in each category, there are 100 annotated images. The data is collected from GPS information and Velodyne laser scans. There are 3 categories of dataset urban unmarked (UU), urban marked (UMM), and urban multiple marked lanes (URBAN) [16]
- There are 1225 frames in the Caltech lane dataset, distributed in four clips recorded from the street in Pasadena, CA at distinct daytime. Caltech-lanes Dataset [13]
- CamVid Dataset [10] is road/driving scene video dataset which is captured by the car-mounted camera with a resolution 960×720. The dataset consists of 367 training set, 101 validation set and 233 test sets
- Culane datasetbd6This dataset is collected by mounting cameras on six different vehicles using six different drivers on Beijing roads. The dataset is crested by recording videos for more than 55 hours. Total 133235 frames were extracted. The dataset is

divided into 88880 training set, 9675 validation set and 34860 testing dataset. This dataset contains normal and 8 different challenging categories.

- BDD100k [24] dataset is a diverse driving dataset for Heterogeneous Multitask Learning
- VPGNet [31] was associated with the methods proposed to use the vanishing point to predict the lane marking. In addition to vanishing point labels, various types of lane markings and road signs are marked in detail. This dataset includes varying degrees of rainfall and nighttime images, which is challenging due to severe weather and extreme lighting conditions.
- CurveLanes [2] has more images with curved lines. So it fulfills the requirement of curved scenes.

3.1.2 Online evaluation dataset

To compute detection correctness and confidence the online evaluation techniques integrates roadway and lane configuration stats and combine them with other sensors like GPS or LIDAR. The road and lane geometric configurations are discovered after the camera calibration. The lane configuration restriction such as slope, intercept, the position of the vanishing point, etc. is used as an authentic metric for determining the model robustness. The lane parameters are analyzed to evaluate the correctness of lane detection. If the parameters exceed the constraints, re-estimation is proposed. LLAMAS dataset [6] is unsupervised annotated dataset by creating high definition maps for automated driving including lane markers based on Lidar.

The above Table 4 describes data acquisition details of each dataset.

3.1.3 Evaluation metrics

Recently, Visual Assessment or Simple Detection Rates(SDR) are used as evaluation metrics for lane detection models. The most regular metrics used for performance evaluating of

Table 4 Data acquisition details of datasets

Datasets	Data Acquisition Devices	Image Resolution	Annotation
TuSimple	Color Camera	1280*720	Key point coordinates
Kitti	Grey Scale Camera,Color Camera,Laser Scanner	1242*375	Pixel level,Rectangle Coordinates
Caltech	Color Camera	640*780	Key point coordinates
Camvid	Color Camera	960*720	Pixel level
Culane	Color Camera	1640*590	Key point coordinates
BDD100k	Color Camera	1280*720	Computationally expensive
VPGNet	Color Camera	640*780	Key point coordinates
CurveLanes	Dynamic Color Cameras,Sensor	2560*1440	Key point coordinates
LLAMAS	Color Camera, LiDAR Maps	1276*717	Pixel level

lane detection algorithms are Receiver Operating Characteristic (ROC) curves, Accuracy, Precision, Recall, F-score, and Dice Similarity Coefficient (DSC) [50].

$$Accuracy = (TrP + TrN) \div (m) \tag{1}$$

$$Recall = TrP \div (TrP + FaN) \tag{2}$$

$$Precision = TrP \div ((TrP + FaP) \tag{3}$$

$$F\− measure = 2 * Precision * Recall \div (Precision + Recall) \tag{4}$$

$$DSC = 2 * TrP \div (TrP + FaP + m) \tag{5}$$

Here, TrP is True Positive, TrN is True Negative, FaP is False Positive, FaN is False Negative, n is Total Lanes, m = Total Pixels (Table 5).

Below Table 6 shows the performance comparison of different deep learning based techniques.

3.2 Current limitation & challenges

Significant advancements are being made in road perception and lane finding. It is practically used in ADAS systems such as LDWS and LKAS. Most of these models used camera-vision based techniques due to low-cost device that gives extensive knowledge of surrounding to enable lane detection. Still, the vision-based models suffer from illumination variation, shadow, bad weather, etc. Due to the inevitable limitations of the camera, building a high-reliability vision-based model requires lots of development efforts. Therefore, various milestones are to be achieved.

The complexity of the ADAS system varies from simple (LWS) to the hardest problem (autonomous navigation), which led to different approaches by researchers. The most important properties of a system are accuracy and robustness which determine if the model

Table 5 Performance comparison of different technique

Techniques	Accuracy	FaP	FaN	F1-Score
Ripple-GAN [59]	97.82	0.0048	0.0289	97.67
Keep Eye on Lane [48]	95.63	0.0353	0.0292	96.77
Salmnet [55]	96.91	0.0263	0.0252	93.44
EndToEnd-LaneMarker [56]	96.22	0.0308	0.0376	92.28
EL-GAN [17]	96.39	0.0412	0.0336	90.56
PointLaneNet [14]	96.34	0.0467	0.0518	89.61
Lightweight Lane [25]	96.64	0.0602	0.0205	87.56
Spatial as Deep [41]	96.53	0.0617	0.0180	87.27
Agnostic lane detection [24]	96.29	0.0722	0.0218	85.52
Towards end-to-end lane detection [38]	96.40	0.0780	0.0244	84.71
FastDraw [42]	95.20	0.0760	0.0450	84.39
Cascaded CNN [43]	95.24	0.197	0.0620	78.35

Table 6 Synopsis for various lane detection models

Ref	Preprocessing	Lane detection	Tracking	Evaluation	Comments
Wu et al [54]	YIQ color space transform, vanishing point detection	Fan-scanning line detection	None	Visual Examination and Correct Detection Rate	The highway lane departure warning and front collision system is built with straight lane model
Niu et al [39]	Temporal blur, IPM, adaptive threshold	RANSAC	Kalman filter	Camera Quantitative analysis and visual assessment	ALD 2.0 is used for labeling video ground truth systematically
Jung and Kelber [28]	Edge detection, Gaussian filter	Gradient Weighted Hough transformation	a lane boundary model	Camera Frame images and visual assessment	proposed linear-parabolic model requires low computational power and memory requirements and showed better predictions in noise, shadows, lack of lane painting and change of illumination conditions.
Wang et al [53]	Vanishing point detection and canny detection	Control point detection	None	Camera Frame images and visual assessment	The B-snake model is powerful in shadows and lighting discrepancy
Aly [2]	IPM, Gaussian kernel filter	RANSAC	Hough transform	Evaluated on Caltech dataset and Quantitative analysis	This method is robust on curve lanes and shadows but affected by road painting and crossways
Borkar et al [7]	IPM, temporal blur	RANSAC	Kalman filter	Visual assessment and detection metrics	The model is created mainly for night vision.
Haloi and Jayagopi [20]	IPM, steerable filter	RANSAC	None	Evaluated on public KITTI dataset	optical flow was employed to built LDWS and the model is robust to shadows
Gurghian et al [19]	ROI, artificial image generation	CNN	None	Pixel level distance evaluation	end-to-end lane recognition method is applied in real time
Rose et al [45]	Dynamic thresholds, Canny detector	Least Squares fitting , Hough Transform	Kalman filter	Real time assessment for spatial and slope criterion and MAE with ground truth position	A reliable lane detection model and lateral offset measurement using sensors like camera and Lidar

Table 6 (continued)

Ref	Preprocessing	Lane detection	Tracking	Evaluation	Comments
Sehestedt et al [46]	IPM	Markov model	Clustered Particle Filters	Tested on dataset from testbed vehicle CRUISE and Evaluation using Matlab	Robust in challenging illumination scenario and reliable for urban straight and curved roads
Li et al [33]	ROI, IPM	CNN, RNN	None	Quantitative analysis with receiver operating characteristic curve	The model uses long-short-term memory(LSTM) that record lane's geometrical structures for a times-pan in the video sequences
Assidiq et al [4]	Canny edge detector	Hough transform	Hyperbola Fitting	Visual assessment and correct detection rate	Robust in shadows and lighting discrepancy situations, Suitable for color coated curved or linear roads
Liu et al [35]	Gaussian filter	Canny Edge detection	Hough transform	Evaluation using Performance Metric	Computationally expensive and poor at curve lane detection
Guo et al [18]	IPM	Cascade lane feature detector	Catmull Rom splines	Analysis using 6121 images sequences	the model use weighted graph particle filter. Robust in various lighting and weather conditions.
Ozgunalp and Dahnoun [40]	Symmetrical Local Threshold	Hough accumulator	ROI,vanishing point	Camera Frame images and visual assessment	Robust in shadows and night Suitable for both straight and curved roads
Li et al [32]	Camera Calibration, Noise Cancellation filer	ROI, Canny edge detector	Kalman filter	Evaluated using manually labeled frames with multiple metrics	Suitable simple straight lane detection but has inadequate performance in challenging scenarios like heavy traffic, complex road textures and poor illumination
Mamidala et al [36]	Reduce Resolution	SegNet	None	Performance Evaluation using multiple metrics	dynamic data is collected from Google street-view api that gives quality information about roads and lane.
He et al [22]	IPM and weighted hat-like filter	DVCNN framework	Google Optimization	Visual and Quantitative Evaluation	Fails when the lanes in the image is occluded by the vehicle.

can be used in real-time. Various lane detection models implemented using machine learning techniques have achieved significantly better processing and lane prediction results. But generally, the detection fails as the camera moves out of the region of interest (ROI) and in various challenging scenarios. Also, there is a considerable research gap in the road perception issues: perception of multi-lane and stochastic typologies of road and lane, this is known as middle-complexity. The analysis carried out on full autonomy does not answer the middle-complexity problem. In fact, as the full autonomous problem is very difficult to solve, ad-hoc solutions are created to avoid extensive onboard understanding and focus on finite problem aspects only. Therefore, the biggest challenge to future models in the next decade is extending the scope of lane perception and support secure and reliable lane prediction under challenging scenarios.

To build a secure and reliable model, various functional blocks and diverse conditions, assumptions need to be recognized. Additional validation tasks must be performed, as many failure cases are infrequent and tough to recognize. Hence, significant research and development efforts are required to achieve satisfying results. One of the solutions to overcome camera limitations is combining other modalities. Considering the scenario following are few ways that may be helpful:

3.2.1 Fusion or integration technique

If Single vision-based techniques gives less performance due to camera limitations, fusion of other perception models is helpful. To enhance the lane prediction, multiple backup models are integrated, this is known as algorithm level integration. The algorithms are parallel or serially integrated and weighted. The system should be able to infer the corner cases where the algorithm fails to weigh them accordingly or transit among these.

3.2.2 Public benchmark evaluation

All the researchers follow different evaluation metrics to assess the system performance in terms of accuracy and robustness as there is no standard benchmark to evaluate the model and it becomes very difficult to compare across publications. A public video benchmark solves this problem.

4 Conclusion

In this review, various conventional and recent lane detection techniques based on deep learning are discussed. We compared various methodologies to give an overview of preprocessing, detection, and tracking techniques used by various researchers. Many evaluation metrics and data-sets used for Lane detection systems are also presented here. The Lane detection system is implemented using Unet and Segnet architecture using Tusimple dataset. As compared to the Segnet, Unet performs better and gives an accuracy 97.52%. Based on in depth survey on deep learning based techniques used for lane detection system, we conclude that the performance of the GAN model and Encoder Decoder based deep learning techniques is quite better as compared to the other deep learning based techniques. Finally, we discuss current gaps or limitations in lane detection and the future challenges, furthermore we provide approaches to bridge these gaps.

Editorial Policies for:

Springer journals and proceedings: https://www.springer.com/gp/editorial-policies

Nature Portfolio journals: https://www.nature.com/nature-research/editorial-policies
Scientific Reports: https://www.nature.com/srep/journal-policies/editorial-policies
BMC journals: https://www.biomedcentral.com/getpublished/editorial-policies

Acknowledgements We thank Dr Swati Shinde, Professor, Pimpari Chinchwad College of Engineering, Pune for her assistance in the revision of this paper. Her comments and suggestions has greatly improved the manuscript.

Author Contributions All authors have equally contributed to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript.

Funding Not applicable

Data Availability Not applicable

Code Availability Not applicable

Declarations

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Conflict of Interests We certify that there is no actual or potential conflict of interest in relation to this article.

Ethics approval The submitted work is original and not have been published elsewhere in any form or language

Consent to participate Not applicable

Consent for Publication Not applicable

References

1. Alvarez JM, Salzmann M, Barnes N (2014) Large-scale semantic co-labeling of image sets. In: IEEE winter conference on applications of computer vision. IEEE, pp 501–508
2. Aly M (2008) Real time detection of lane markers in urban streets. In: 2008 IEEE intelligent vehicles symposium. IEEE, pp 7–12
3. Aly M (2008) Real time detection of lane markers in urban streets. In: 2008 IEEE intelligent vehicles symposium. IEEE, pp 7–12
4. Assidiq AA, Khalifa OO, Islam MR, Khan S (2008) Real time lane detection for autonomous vehicles. In: 2008 International conference on computer and communication engineering. IEEE, pp 82–88
5. Aubert D, Kluge KC, Thorpe CE (1991) Autonomous navigation of structured city roads. In: Mobile Robots V. SPIE, vol 1388, pp 141–151
6. Behrendt K, Soussan R (2019) Unsupervised labeled lane markers using maps. In: Proceedings of the IEEE/CVF international conference on computer vision workshops, pp 0–0
7. Borkar A, Hayes M, Smith MT (2009) Robust lane detection and tracking with ransac and kalman filter. In: 2009 16th IEEE international conference on image processing (ICIP). IEEE, pp 3261–3264
8. Broggi A, Cappalunga A, Caraffi C, Cattani S, Ghidoni S, Grisleri P, Porta PP, Posterli M, Zani P (2010) Terramax vision at the urban challenge 2007. IEEE Trans Intell Transp Syst 11(1):194–205
9. Broggi A, Cattani S (2006) An agent based evolutionary approach to path detection for off-road vehicle guidance. Pattern Recogn Lett 27(11):1164–1173
10. Brostow GJ, Fauqueur J, Cipolla R (2009) Semantic object classes in video: a high-definition ground truth database. Pattern Recogn Lett 30(2):88–97
11. Buehler M, Iagnemma K, Singh S (2007) The 2005 DARPA grand challenge: the great robot race. Springer, vol 36
12. Burrow M, Evdorides H, Snaith M (2003) Segmentation algorithms for road marking digital image analysis. In: Proceedings of the institution of civil engineers-transport. Thomas telford ltd, vol 156, pp 17–28

13. Caltech-lanes Dataset (2022) https://www.vision.caltech.edu/malaa/datasets/caltech-lanes/. Accessed April 2021
14. Chen Z, Liu Q, Lian C (2019) Pointlanenet: efficient end-to-end cnns for accurate real-time lane detection. In: 2019 IEEE intelligent vehicles symposium (IV). IEEE, pp 2563–2568
15. Cheng H-Y, Jeng B-S, Tseng P-T, Fan K-C (2006) Lane detection with moving vehicles in the traffic scenes. IEEE Trans Intell Trans Syst 7(4):571–582
16. Fritsch J, Kuehnl T, Geiger A (2013) A new performance measure and evaluation benchmark for road detection algorithms. In: 16th International IEEE conference on intelligent transportation systems (ITSC 2013). IEEE, pp 1693–1700
17. Ghafoorian M, Nugteren C, Baka N, Booij O, Hofmann M (2018) Elgan: embedding loss driven generative adversarial networks for lane detection. In: Proceedings of the European conference on computer vision (ECCV) workshops, pp 0–0
18. Guo C, Mita S, McAllester D (2010) Lane detection and tracking in challenging environments based on a weighted graph and integrated cues. In: 2010 IEEE/RSJ international conference on intelligent robots and systems. IEEE, pp 5543–5550
19. Gurghian A, Koduri T, Bailur SV, Carey KJ, Murali VN (2016) Deeplanes: End-to-end lane position estimation using deep neural networksa. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 38–45
20. Haloi M, Jayagopi DB (2015) A robust lane detection and departure warning system. In: 2015 IEEE intelligent vehicles symposium (IV). IEEE, pp 126–131
21. He B, Ai R, Yan Y, Lang X (2016) Lane marking detection based on convolution neural network from point clouds. In: 2016 IEEE 19th international conference on intelligent transportation systems (ITSC). IEEE, pp 2475–2480
22. He B, Ai R, Yan Y, Lang X (2016) Accurate and robust lane detection based on dual-view convolutional neutral network. In: 2016 IEEE intelligent vehicles symposium (IV). IEEE, pp 1041–1046
23. He B, Ai R, Yan Y, Lang X (2016) Lane marking detection based on convolution neural network from point clouds. In: 2016 IEEE 19th international conference on intelligent transportation systems (ITSC). IEEE, pp 2475–2480
24. Hou Y (2019) Agnostic lane detection. arXiv:1905.03704
25. Hou Y, Ma Z, Liu C, Loy CC (2019) Learning lightweight lane detection cnns by self attention distillation. In: 2019 IEEE/CVF international conference on computer vision (ICCV), pp 1013–1021. https://doi.org/10.1109/ICCV.2019.00110
26. Huang Y, Li Y, Hu X, Ci W (2018) Lane detection based on inverse perspective transformation and kalman filter. KSII Trans Internet Inf Syst (TIIS) 12(2):643–661
27. Huang AS, Moore D, Antone M, Olson E, Teller S (2009) Finding multiple lanes in urban road networks with vision and lidar. Auton Robot 26(2):103–122
28. Jung CR, Kelber CR (2004) A robust linear-parabolic model for lane following. In: Proceedings. 17th Brazilian symposium on computer graphics and image processing. IEEE, pp 72–79
29. Kim Z (2008) Robust lane detection and tracking in challenging scenarios. IEEE Trans Intell Trans Syst 9(1):16–26
30. Kuderer M, Kretzschmar H, Burgard W (2013) Teaching mobile robots to cooperatively navigate in populated environments. In: 2013 IEEE/RSJ Int Conf Intell Robots Syst. IEEE, pp 3138–3143
31. Lee S, Kim J, Shin Yoon J, Shin S, Bailo O, Kim N, Lee T-H, Seok Hong H, Han S-H, So Kweon I (2017) Vpgnet: vanishing point guided network for lane and road marking detection and recognition. In: Proceedings of the IEEE international conference on computer vision, pp 1947–1955
32. Li Y, Iqbal A, Gans NR (2014) Multiple lane boundary detection using a combination of low-level image features. In: 17th International IEEE conference on intelligent transportation systems (ITSC). IEEE, pp 1682–1687
33. Li J, Mei X, Prokhorov D, Tao D (2016) Deep neural network for structural prediction and lane detection in traffic scene. IEEE Trans Neural Netw Learn Syst 28(3):690–703
34. Liu W, Li S, Huang X (2014) Extraction of lane markings using orientation and vanishing point constraints in structured road scenes. Int J Comput Math 91(11):2359–2373
35. Liu G, Wörgötter F, Markelić I (2010) Combining statistical hough transform and particle filter for robust lane detection and tracking. In: 2010 IEEE intelligent vehicles symposium. IEEE, pp 993–997
36. Mamidala RS, Uthkota U, Shankar MB, Antony AJ, Narasimhadhan A (2019) Dynamic 0approach for lane detection using google street view and cnn. In: TENCON 2019-2019 IEEE region 10 conference (TENCON). IEEE, pp 2454–2459

37. McCall JC, Trivedi MM (2004) An integrated, robust approach to lane marking detection and lane tracking. In: IEEE Intelligent vehicles symposium. IEEE, 2004, pp 533–537
38. Neven D, Brabandere B, Georgoulis S, Proesmans M, Van Gool L (2018) Towards end-to-end lane detection: an instance segmentation approach, pp 286–291. https://doi.org/10.1109/IVS.2018.8500547
39. Niu J, Lu J, Xu M, Lv P, Zhao X (2016) Robust lane detection using two-stage feature extraction with curve fitting. Pattern Recogn 59:225–233
40. Ozgunalp U, Dahnoun N (2014) Robust lane detection & tracking based on novel feature extraction and lane categorization. In: 2014 Ieee international conference on acoustics, speech and signal processing (icassp). IEEE, pp 8129–8133
41. Pan X, Shi J, Luo P, Wang X, Tang X (2018) Spatial as deep: spatial cnn for traffic scene understanding. In: AAAI
42. Philion J (2019) Fastdraw: addressing the long tail of lane detection by adapting a sequential prediction network, pp 11574–11583. https://doi.org/10.1109/CVPR.2019.01185
43. Pizzati F, Allodi M, Barrera A, García F (2019) Lane detection and classification using cascaded cnns
44. Pomerleau D (1995) Ralph: rapidly adapting lateral position handler. In: Proceedings of the intelligent vehicles' 95. Symposium. IEEE, pp 506–511
45. Rose C, Britt J, Allen J, Bevly D (2014) An integrated vehicle navigation system utilizing lane-detection and lateral position estimation systems in difficult environments for gps. IEEE Trans Intell Transp Syst 15(6):2615–2629
46. Sehestedt S, Kodagoda S, Alempijevic A, Dissanayake G (2007) Robust lane detection in urban environments. In: 2007 IEEE/RSJ international conference on intelligent robots and systems. IEEE, pp 123–128
47. Sivaraman S, Trivedi MM (2013) Integrated lane and vehicle detection, localization, and tracking: a synergistic approach. IEEE Trans Intell Transp Syst 14(2):906–917
48. Tabelini L, Berriel R, Paixao TM, Badue C, De Souza AF, Oliveira-Santos T (2021) Keep your eyes on the lane: real-time attention-guided lane detection. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 294–302
49. Tusimple Dataset (2017). https://github.com/TuSimple/tusimple-benchmark. Accessed April 2021
50. Veit T, Tarel J-P, Nicolle P, Charbonnier P (2008) Evaluation of road marking feature extraction. In: 2008 11th International IEEE conference on intelligent transportation systems. IEEE, pp 174–181
51. Veit T, Tarel J-P, Nicolle P, Charbonnier P (2008) Evaluation of road marking feature extraction. In: 2008 11th International IEEE conference on intelligent transportation systems. IEEE, pp 174–181
52. Von Gioi RG, Jakubowicz J, Morel J-M, Randall G (2008) Lsd: a fast line segment detector with a false detection control. IEEE Trans Pattern Anal Mach Intell 32(4):722–732
53. Wang Y, Teoh EK, Shen D (2004) Lane detection and tracking using b-snake. Image Vis Comput 22(4):269–280
54. Wu C-F, Lin C-J, Lee C-Y (2011) Applying a functional neurofuzzy network to real-time lane detection and front-vehicle distance measurement. IEEE Trans Syst Man Cybern Part C (Appl Rev) 42(4):577–589
55. Xu X, Yu T, Hu X, Ng WW, Heng P-A (2020) Salmnet: a structure-aware lane marking detection network. IEEE Trans Intell Transp Syst 22(8):4986–4997
56. Yoo S, Lee HS, Myeong H, Yun S, Park H, Cho J, Kim DH (2020) End-to-end lane marker detection via row-wise classification. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, pp 1006–1007
57. Yoo JH, Lee S-W, Park S-K, Kim DH (2017) A robust lane detection method based on vanishing point estimation using the relevance of line segments. IEEE Trans Intell Transp Syst 18(12):3254–3266
58. Yuan J, Tang S, Pan X, Zhang H (2014) A robust vanishing point estimation method for lane detection. In: Proceedings of the 33rd chinese control conference. IEEE, pp 4887–4892
59. Zhang Y, Lu Z, Ma D, Xue J-H, Liao Q (2020) Ripple-gan: lane line detection with ripple lane line detection network and wasserstein gan. IEEE Trans Intell Transp Syst 22(3):1532–1542
60. Zhang G, Zheng N, Cui C, Yan Y, Yuan Z (2009) An efficient road detection method in noisy urban environment. In: 2009 IEEE intelligent vehicles symposium. IEEE, pp 556–561

Affiliations

Ashwini Sapkal[1] · **Arti**[1] · **Dishant Pawar**[1] · **Prashant Singh**[1]

Arti
artisaradhna12@gmail.com

Dishant Pawar
dishantpawar_17343@aitpune.edu.in

Prashant Singh
prashantsingh_17459@aitpune.edu.in

1 Department of Information Technology, Pune, Army Institute of Technology, Dighi, Pune, 411015, Maharashtra, India