Lane Detection in Complex Scenes based on Ultra-Fast Structure-aware

Aiwei Guo[1], Muyao Li[2], Dongke Liu[3, *] and Shuyuan Wang[4]

[1]Qingdao University of Science and Technology, Qingdao, China

[2]Camford Royal School, Beijing, China

[3]University of Electronic Science and Technology of China, Chengdu, China

[4]Yangzhou University, Yangzhou, China

* Corresponding Author Email: 26139998L@student.gla.ac.uk

Abstract. With the rapid development of lane detection technology in the field of autonomous driving and intelligent transportation, its robustness and real-time requirements in complex environments are constantly improving. Although the traditional image processing method has a fast-processing speed, it has obvious limitations in complex scenes. Although the segmentation method based on deep learning has advantages in accuracy and adaptability, its high computational cost and large hardware requirements limit its practical application. To solve these problems, this paper adopts the structure-sensing depth lane detection algorithm based on global features. By introducing global features and larger receptive field, the algorithm simplifies the image segmentation process and improves the computational efficiency. We tested TuSimple and CULane datasets, and the results showed that the algorithm significantly improved the detection speed and environmental adaptability while maintaining high accuracy, especially in complex environments such as rainy, foggy and night. This paper further discusses the potential improvement direction of the algorithm in practical application, and lays a foundation for the future lane detection technology.

Keywords: Lane detection, deep learning, Complex Scenes.

1. Introduction

As lane detection technology becomes increasingly sophisticated, it has been widely applied in various fields such as autonomous driving, intelligent traffic management, and intelligent public transportation. With the improvement of processing speed and detection accuracy in various complex scenarios, lane detection technology has gradually entered people's daily lives. At the same time, the requirements for the robustness and real-time performance of lane detection technology in complex environments are also constantly increasing. However, the main lane detection technologies currently suffer from reduced accuracy and slower processing speed under special conditions such as no light at night and road surface water. Therefore, it is necessary to develop more accurate and fast advanced algorithms.

The current main lane detection methods can be roughly divided into two categories: traditional image processing methods and deep learning-based segmentation methods. Traditional image processing method: although the calculation amount is small and the hardware requirement is not high, it has high real-time performance and fast processing speed and can achieve better detection results in simple scenes. However, these methods show obvious limitations in dealing with complex scenes, such as lighting changes, lane lines blurred or blocked, and their accuracy decreases significantly. The segmentation method based on deep learning has significant advantages in detection accuracy and robustness, especially in complex environments showing strong adaptability. These methods can maintain high accuracy in diverse scenarios through a large amount of training data and complex model design. However, they require a large amount of computation, high hardware requirements, long training time, and large storage requirements of the model, which limit their practical application to some extent.

As one of the most representative methods, Qin et al. simplify the image segmentation process by introducing global features and using a larger receptive field to analyze the probability distribution of anchor points, thereby significantly reducing the computational cost and improving the computing speed. Compared with traditional methods, this method can maintain high accuracy while improving detection speed and environmental adaptability, which makes it more practical in diverse driving environments. Additionally, recent work validated the effectiveness of the method by conducting ablation studies using TuSimple and CULane datasets. The experimental results show that the proposed method can greatly improve the operation speed while maintaining high accuracy in both normal and challenging environments.

However, the current research is still insufficient to fully explore the performance of this method under extreme conditions (such as rainstorms, fog, no streetlight at night, etc.). Therefore, based on the existing research, we plan to further explore the robustness and real-time performance of this method in complex environments, and propose corresponding improvement schemes to enhance its reliability in practical applications. Through this comprehensive study, we hope to contribute to the development of more accurate and fast lane detection algorithms suitable for a wide range of pan-driving conditions.

2. Related work

Lane detection based on deep learning is an important research area in autonomous driving and advanced driver assistance systems (ADAS), which has achieved significant progress in recent years. LaneNet uses an end-to-end approach to detect lane lines, extracts image features through convolutional neural networks, and uses an embedding layer to distinguish different lane lines. By performing geometric fitting, LaneNet can restore the specific shape of lane lines and maintain high detection accuracy in complex environments. SCNN introduces a spatial attention mechanism to strengthen the connection between neighboring pixels through local connectivity. In addition, SCNN uses multi-scale features to capture lane line information at different scales, further improving the accuracy of lane line detection. LSTR (Lane Segmentation and Tracking by Recurrence) uses past frames to predict the lane line status in the current frame, combines semantic segmentation and lane line tracking based on recurrent neural networks (RNN), and improves the continuity and stability of detection through temporal information. ENet (Encoder-Decoder with Atrous Separable Convolution) is a lightweight semantic segmentation network that uses atrous separable convolution to reduce the number of parameters while maintaining the receptive field. In addition, TransFuser combines the Transformer architecture and convolutional networks, using the powerful representation ability of the Transformer to handle lane detection tasks. These previous methods have utilized advanced neural network architectures (such as convolutional neural networks, Transformers, etc.) and innovative technologies (such as attention mechanisms, multi-task learning, etc.) to improve the accuracy, robustness, and real-time performance of lane detection. Different from existing methods, we focus on exploring the detection effect in various complex scenes, aiming to provide a decision basis for the method selection in real applications.

3. Method

3.1. Ultra-Fast-Lane-Detection Model

As shown in Fig 1, the Ultra-Fast-Lane-Detection model combines the backbone feature extractor based on a deep convolutional network with an additional convolutional layer to enhance the lane analysis capability. The core of the model is a deep convolutional network architecture, which is responsible for extracting image features and loading pre-training weights as needed. If the auxiliary processing function is enabled, the model will add multiple convolution layers to further optimize the intermediate feature map and generate an auxiliary segmentation map to improve the detection accuracy. The model also includes a classification module, which converts the pooled feature map into the lane group classification result through the full connection layer. In the forward propagation process, the input image first extracts feature through the backbone network and generates auxiliary segmentation graphs if auxiliary processing is enabled. Then, the feature graphs output by the backbone network are pooled and classified. Finally, according to the settings of auxiliary processing functions, the model outputs the classification results of the lane group or outputs the classification results and auxiliary segmentation graphs at the same time. This structured approach improves the accuracy and effectiveness of lane detection.

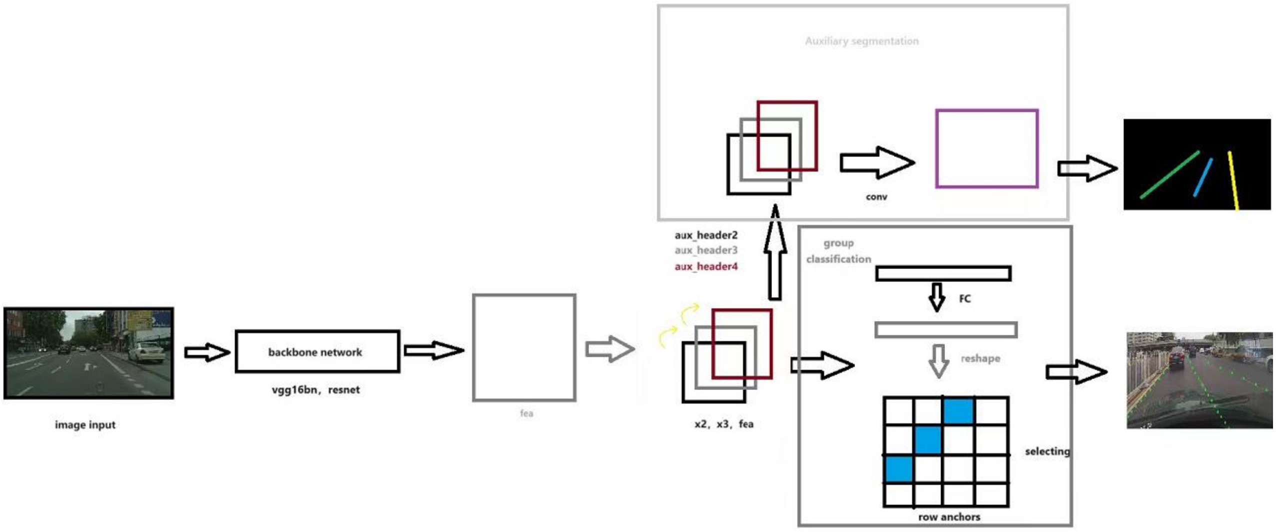

Fig. 1 The overall architecture of the Ultra-Fast-Lane-Detection method. The auxiliary branch is only valid for training, which is shown in the upper gray box, based on segmentation, the tests are based on group classification.

In feature extraction backbone, two kinds of deep learning models are mainly involved: VGG16bn and Resnet. The VGG16bn model adopts the pre-trained VGG16 feature extractor, and trims and recombines it to simplify its structure and adapt to specific task requirements. The model generates image features through the reduced convolution layer, which provides the basis for the subsequent processing. The ResNet model implements a variety of ResNet architectures, including RESNET-18, RESNET-34, and RESNET-50. The model initializes the pre-trained ResNet network according to the specified network version, and extracts the convolution layer, batch normalization layer, activation function, and four main residual blocks from it. In the forward propagation process, after the initial convolution and batch normalization processing, the input image is extracted by multiple residual blocks, and the feature map of multiple intermediate layers is finally output. These feature maps can be used for further feature processing or fusion. The VGG16 model focuses on the extraction of shallow convolutional features, while the ResNet model efficiently deals with deep features through a residual learning mechanism.

3.2. Data set and experimental setup

To fully evaluate the performance of the Ultra-Fast-Lane-Detection model in harsh environments, we constructed two main datasets: the standard environment data set and the harsh environment data set.

3.2.1. Standard Environment data set

Data sources: Collected from publicly available lane detection datasets, such as the TuSimple data set or CULane data set. These datasets provide high-quality, annotated lane line images, including diverse road scenes under clear daytime conditions.

Data characteristics: The image resolution is usually 1280x720 or 1920x1080 and contains high-contrast lane markings and multiple lane types (live lines, dashed lines).

Annotation information: Provide detailed annotation of lane lines, including pixel coordinates and curvature information of lane lines, as a benchmark for model training and evaluation.

3.2.2. Harsh environment data set

Data source: Generate challenging lane line image data by driving real cars under different harsh environmental conditions. The details include:

Rainy day environment: Capture images containing rain droplets, standing water, and slippery roads. Lane lines in these images may be disturbed or blurred by rain.

Hazy environment: Images are taken in hazy weather conditions with low visibility. Images often suffer from low contrast and high noise.

Night environment: Images taken at night under low light conditions. In the image, the brightness of the lane lines is low and the background noise is high.

Data characteristics: Image resolution is consistent with standard environmental data sets, but image quality and visibility of lane lines are poor due to inclement weather. The lane lines in the image may be partially obscured or blurred due to reflections and shadows.

Annotated information: Provide pixel coordinates of lane lines using an annotated format consistent with standard environmental data sets. Due to the influence of harsh environment, the accuracy of the annotation may be challenged, and the quality of the annotation needs to be strictly reviewed.

3.2.3. Experimental Settings

Training and verification:

Training set: The model is trained using standard environmental data sets to ensure that the model has good basic performance under normal conditions.

Validation set: The model is validated using partial images from a harsh environment dataset to assess its adaptability and robustness in different environments.

Performance evaluation:

Indicators: Precision, Recall and F1 scores were used to evaluate model performance.

Evaluation method: By comparing the detection results of the model in standard environment and harsh environment, the detection accuracy and robustness of the model in different environments are analyzed. According to the test results in harsh environment, the error analysis and error source are discussed in detail.

4. Experiment

We will test the model under different conditions by controlling variables and analyze the detection accuracy and robustness of the model by reference Precision, Recall, F1 score and Frames per Second (FPS). The main control variables are weather conditions (rain, fog, cloudy, sunny), period (daytime, dusk, night), occlusion conditions (no occlusion, slight occlusion, severe occlusion), picture quality (high image quality, low image quality). We first tested the full set and found that the accuracy and recall rate decreased somewhat compared to the provided data set, and then we classified the images in the data set and compared them to the ideal condition (sunny day, good noon light, no occlusion, good driveway, high resolution).

Table 1. Performance for different occlusion and light conditions

Metric	Settings	Greasy weather	Rain	Sun		
				Daytime	Evening	Night
Accuracy	Universal set	0.7715				
	Severe occlusion	0.6781	0.6862	0.8250	0.7725	0.6331
	Light occlusion	0.7415	0.7307	0.8886	0.7866	0.7026
	No shelter	0.7469	0.7680	0.8992	0.8383	0.7525
Recall	Universal set	0.7548				
	Severe occlusion	0.6948	0.6271	0.8420	0.7361	0.5871
	Light occlusion	0.7385	0.6578	0.8889	0.7749	0.6049
	No shelter	0.7494	0.7952	0.9033	0.8415	0.7817
F1 score	Universal set	0.7648				
	Severe occlusion	0.6931	0.6587	0.8280	0.7495	0.6176
	Light occlusion	0.7366	0.7070	0.8880	0.7737	0.6619
	No shelter	0.7476	0.7724	0.8978	0.8404	0.7455

4.1. Performance Analysis for different conditions

4.1.1. Occlusion condition

First of all, we uniformly selected the pictures of noon and sunny days to control the influence of light in different periods and weather. Moreover, the lane clarity and wear degree of these pictures were low, and the resolution was 1080p. The tests were conducted according to no occlusion, slight occlusion and severe occlusion in sequence.

The test results show that the algorithm can maintain high speed and recognition accuracy for all kinds of occlusion cases. Even in the case of severe occlusion, it can still identify the location of the lane more accurately.

4.1.2. Time Period

We selected the images with no shelter and clear lanes at noon and a resolution of 1080p, and classified them according to different periods, including good lighting during the day, streetlights not turned on at dusk but with sun residual light, and no streetlights at night.

As can be seen from the Fig 2, different lighting conditions in different time periods have a certain impact on the algorithm's lane recognition. When dusk brightness decreases, the accuracy decreases to a certain extent, while the recognition accuracy is even lower under extreme conditions without streetlights at night. However, when there are streetlights at night, the recognition accuracy and speed can reach the level of good daylight.

4.1.3. Weather conditions

The pictures without shelter at noon and with a clear resolution of 1080p were selected, and classified according to three road conditions: sunny, rainy and foggy.

As can be seen from the figure, weather conditions also have a certain impact on lane recognition, especially in rainy and foggy days. On cloudy days, because the road can still maintain sufficient lighting conditions, the detection speed and accuracy can be maintained, while due to the reflection of the ground in rainy days and the low visibility in foggy days, both the detection speed and detection accuracy have declined significantly.

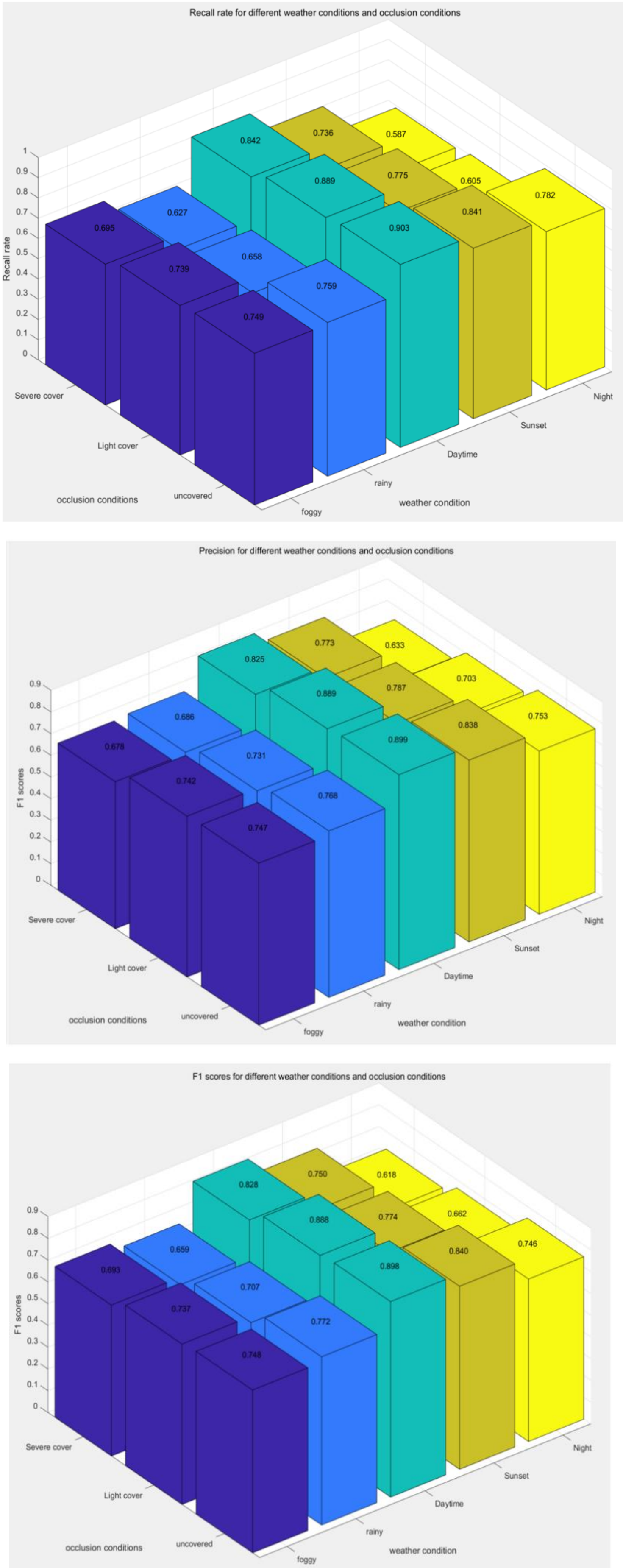

Fig. 2 Bar chart of accuracy, recall and f1 scores under different conditions

4.2. Improvement Suggestions

For the optimization of lane recognition speed, in the case of consistent hardware or algorithm conditions, we put forward some more practical operations for the vehicle recognition system:

4.2.1. Model compression

Quantization: Converting floating-point weights and activation values in the model to low-precision representations (such as 8-bit integers), which can significantly reduce computation and memory usage. Pruning: Removing unimportant weights from the network, reducing computation while maintaining performance.

4.2.2. Optimization model architecture

Lightweight network: The use of lightweight network architectures such as MobileNet, EfficientNet or ShuffleNet is an important optimization direction for in-vehicle systems. These networks are designed to provide efficient performance in situations with limited computing resources and are particularly suitable for real-time lane detection tasks. Network Scaling: Adjusting the complexity of the network, such as reducing the number of convolutional layers or the number of feature maps per layer, can run faster models on on-board devices.

4.2.3. Image preprocessing optimization

Resolution adjustment: In an onboard system, reducing the resolution of the input image to a level sufficient to maintain the accuracy of lane detection can reduce the computational burden. For example, the resolution can be adjusted according to real-time requirements to balance processing speed and detection accuracy. Area of interest (ROI): When processing images, focusing on areas where lane lines are likely to appear (such as the middle of the road) and ignoring other areas can significantly improve efficiency. This is especially important for real-time processing, as on-board systems typically need to handle video streams with high frame rates.

4.2.4. Noise removal

In bad weather conditions, image noise and blur are common problems. Noise removal algorithms (such as Gaussian filter, mean filter, bilateral filter) are used to improve the image quality, so as to improve the accuracy of lane detection. Image enhancement techniques: Apply image enhancement techniques (such as contrast stretching, histogram equalization) to improve low-contrast images and make lane lines more visible. Image repair: Apply image repair techniques (such as deep learning-based image repair models) to damaged or blocked areas to fill in the lane lines that are blocked by weather conditions such as rain and snow or specifically for the situation, such as weather condition specific models: Train lane recognition models for specific adverse weather conditions (such as rain, snow, fog). Annotated data collected under severe weather conditions were used for training to improve the model's adaptability to these conditions

5. Discussion

In this study, we comprehensively evaluated the performance of the lane detection model under different test conditions, and analyzed the influence of weather conditions, period, occlusion conditions and picture quality on the model performance through the method of control variables. We evaluated the detection accuracy and robustness of the model using Precision, Recall and F1 scores. Here is a discussion of the main results:

5.1. Precision

The accuracy of the model shows obvious difference in different weather conditions. The accuracy of the model is relatively high in sunny conditions, whether it is day, evening or night, especially in the case of no occlusion (the highest accuracy is 0.8992), which indicates that the model can effectively detect lane lines under ideal conditions. However, when faced with bad weather (such as fog or rainy days), the accuracy decreases significantly. For example, in the case of severe occlusion, the accuracy dropped to 0.6781 on foggy days, while it remained at 0.8250 on sunny days. At the same time, visible occlusion also significantly affects the performance of the model, especially severe occlusion will significantly reduce the accuracy, which may be due to the occlusion of the visible part of the lane line, resulting in detection difficulties.

5.2. Recall Rate

The trend in recall rates is similar to accuracy, but slightly different. Under all test conditions, the recall rate of the model showed a relatively stable trend. Under sunny conditions, especially in the case of no shelter, the recall rate is higher (up to 0.9033), indicating that the model can identify the lane lines better. However, recall rates are also affected by bad weather and shelter conditions. For example, recall rates were reduced in foggy and severely obscured conditions (0.6948 for foggy severely obscured conditions), indicating increased omissions in the model's identification of lane lines.

5.3. F1 Score

The F1 score takes accuracy and recall into account, providing a comprehensive assessment of the model's performance. We observed that F1 scores fluctuated significantly under different test conditions. Especially in the case of bad weather and severe occlusion, the F1 score is significantly lower (for example, the F1 score of severe occlusions in foggy days is 0.6931), which further reflects that the detection ability of the model under these conditions is greatly challenged. Under ideal conditions (sunny and unshaded), the F1 score reached a high level (up to 0.8978), indicating that the model performed well under good test conditions.

6. Conclusion

This study shows that the lane detection model performs well under ideal conditions, while the detection accuracy and recall rate of the model decrease under bad weather, different lighting conditions and occlusion conditions. These results suggest that future research can address these challenges, for example by enhancing the diversity of the dataset, improving the robustness of the model, or incorporating other sensor information to improve the model's performance in complex environments. In addition, given the importance of detection speed in real-time applications, subsequent work should further evaluate the real-time processing capability of the model.

Authors Contribution

All the authers contributed equally and their names were listed in alphabetical order.

References

[1] Qin, Z., Wang, H., & Li, X. (2020). Ultra Fast Structure-aware Deep Lane Detection. arXiv:2004.11757v4 [cs.CV].

[2] M. Galinski, V. Danylov, P. Lehoczký, R. Bencel and L. Soltes, "Ultra-Fast Visible Lane Distance Estimation Using a Single Camera," in IEEE Access, vol. 12, pp. 79251-79263, 2024, doi: 10.1109/ACCESS.2024.3409475.

[3] Z. Qin, P. Zhang and X. Li, "Ultra Fast Deep Lane Detection With Hybrid Anchor Driven Ordinal Classification," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 46, no. 5, pp. 2555-2568, May 2024, doi: 10.1109/TPAMI.2022.3182097.

[4] 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.

[5] Y. Hou, Z. Ma, C. Liu, and C. C. Loy, “Learning lightweight lane detection cnns by self attention distillation,” in Int. Conf. Comput. Vis., 2019, pp. 1013–1021.

[6] Z. Tian, C. Shen, H. Chen, and T. He, “Fcos: A simple and strong anchor-free object detector,” IEEE Trans. Pattern Anal. Mach. Intell.,vol. 1, no. 1, pp. 1–12, 2020.

[7] N. bin Abdul Razak, M. Z. bin Mazlan, J. B. Johari, S. A. Bin Che Abdullah and N. K. Mun, "A Lane Detection Using Image Processing Technique for Two-Lane Road," 2022 IEEE 10th Conference on Systems, Process & Control (ICSPC), Malacca, Malaysia, 2022, pp. 214-219, doi: 10.1109/ICSPC55597.2022.10001801.

[8] S. Sultana and B. Ahmed, "Lane Detection and Tracking under rainy weather challenges," 2021 IEEE Region 10 Symposium (TENSYMP), Jeju, Korea, Republic of, 2021, pp. 1-6, doi: 10.1109/TENSYMP52854.2021.9550931.

[9] M. S. Saved, M. Elsharkawy, B. Kobasy and H. Eltabakh, "Enhancing Lane Detection for Autonomous Vehicles Using Image Processing and Real-Time Analysis," 2023 11th International Japan-Africa Conference on Electronics, Communications, and Computations (JAC-ECC), Alexandria, Egypt, 2023, pp. 38-41, doi: 10.1109/JAC-ECC61002.2023.10479625.

[10] Yoichiro Iwasaki and Yuji Kurogi, "Real-time robust vehicle detection through the same algorithm both day and night," 2007 International Conference on Wavelet Analysis and Pattern Recognition, Beijing, China, 2007, pp. 1008-1014, doi: 10.1109/ICWAPR.2007.4421579.

[11] S. Sultana and B. Ahmed, "Lane Detection and Tracking under rainy weather challenges," 2021 IEEE Region 10 Symposium (TENSYMP), Jeju, Korea, Republic of, 2021, pp. 1-6, doi: 10.1109/TENSYMP52854.2021.9550931.

[12] S. Sultana, B. Ahmed, M. Paul, M. R. Islam and S. Ahmad, "Vision-Based Robust Lane Detection and Tracking in Challenging Conditions," in IEEE Access, vol. 11, pp. 67938-67955, 2023, doi: 10.1109/ACCESS.2023.3292128.