International Journal of Scientific Research and Engineering Development-– Volume 8 Issue 5, Sep-Oct 2025
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Computer Vision in Action: An Advanced Assistance System for Safer Driving

Bhargavi Peddi Reddy[1], Sujeeth Papani[2], Tejash Babu Debbati[3]
[1,2,3]Vasavi College of Engineering, Hyderabad, India
Email: Bhargavi@staff.vce.ac.in[1], sujeethpapani@gmail.com[2], tejashbabu07@gmail.com[3]

Abstract

Road traffic accidents continue to pose a significant threat to human safety, with human error as a primary contributing factor. This study presents a vision-based Advanced Driver Assistance System (ADAS) designed to improve road safety by utilizing cutting-edge computer vision and deep learning methods. The system integrates three vital components: lane detection, using CNN-based U-Net models for precise lane tracking; road segmentation, employing advanced image segmentation for a granular understanding of road scenes; and Forward Collision Warning (FCW), which proactively identifies and alerts drivers to potential collision risks. The proposed ADAS model is thoroughly tested using well-known datasets such as TuSimple and BDD100K, where it delivers impressive results, achieving an accuracy of 85.79%, a mean Intersection over Union (IoU) of 36.67%, a mean Average Precision (mAP) of 51.52%. Despite facing hardware constraints, the system maintains real-time responsiveness and dependable performance. By combining these features, the system effectively tackles key road-related issues, boosts driver awareness, and helps minimize human-induced errors, ultimately promoting safer driving conditions. This work adds valuable insight to the evolving field of AI-driven vehicle safety technologies and underscores the significant role of computer vision in preventing road accidents and creating more secure transportation systems.

Introduction

Road traffic accidents continue to pose a significant threat to human safety, with human error identified as a primary contributing factor. In response, Advanced Driver Assistance Systems (ADAS) have emerged as a critical collection of smart technologies designed to help drivers operate vehicles more safely and comfortably. These systems offer a range of safety functionalities, including speed regulation, automatic emergency braking, and lane deviation alerts, by utilizing a mix of sensors, radar, and cameras to perceive the vehicle's environment. Based on the information collected, the system can provide alerts to the driver or, in certain situations, take direct action to mitigate potential risks.

A significant leap in this field has been the integration of computer vision, which enables systems to analyze real-time visual inputs from cameras and respond appropriately. Vision-based ADAS utilizes computer vision models to interpret the driving environment—including road lanes, signs, nearby vehicles, and pedestrians—with the ultimate goal of improving driver awareness and reducing the risk of accidents. Key features of these systems include Lane Drift Notification (LDN), which monitors the vehicle's position to prevent unintended lane changes; Traffic Sign Recognition (TSR), which detects and displays pertinent traffic sign information ; and Pedestrian Recognition, a safety-oriented feature designed to spot individuals and trigger warnings or automatic braking.

This work presents a comprehensive vision-based ADAS that integrates three core components to enhance road safety: lane detection, road segmentation, and a Forward Collision Warning (FCW) system. The lane detection component employs a Convolutional Neural Network (CNN)-based U-Net model to ensure precise lane tracking. For road segmentation, the system uses advanced image segmentation to achieve a granular understanding of the road scene. Finally, the FCW system proactively identifies and alerts drivers to potential collision risks. The proposed model was evaluated using well-known datasets such as TuSimple and BDD100K, where it achieved an accuracy of 85.79%, a mean Intersection over Union (IoU) of 36.67%, and a mean Average Precision (mAP) of 51.52%. Despite hardware constraints, the system maintains real-time responsiveness and dependable performance.

This paper addresses the research question of how to leverage computer vision and deep learning techniques to assist car drivers in detecting various road elements and distances. By combining these features, the system effectively tackles key road-related issues, boosts driver awareness, and helps minimize human-induced errors. This initiative has the potential to significantly minimize injuries and fatalities on the roads by providing timely and accurate assistance to drivers. Overall, this work adds valuable insight to the evolving field of AI-driven vehicle safety technologies and underscores the significant role of computer vision in creating more secure transportation systems.

Research question

How to leverage computer vision and deep learning techniques to assist car drivers to detect lane lines, other vehicles, pedestrians, traffic signs, and the distance from cars in front of them using the road video footage?

Proposed solution, a vision-based ADAS

In recent years, advancements in artificial intelligence and computer vision technologies have opened new possibilities for addressing the challenges of road safety. Specifically, the utilization of machine learning algorithms and computer vision techniques has shown promising potential in developing intelligent systems capable of assisting drivers and reducing the occurrence of accidents. To address these opportunities and challenges, this initiative seeks to create an Advanced Driver Assistance System (ADAS) that harnesses state-of-the-art computer vision techniques. The envisioned ADAS is designed to detect and notify drivers about potential dangers, such as road signs, blind spots in the rear view, and unintended lane changes, thus lowering the likelihood of collisions due to human mistakes. By providing timely and accurate assistance to drivers, this system has the potential to significantly minimize injuries and fatalities on the roads. In the subsequent sections of this dissertation, we will explore the theoretical foundations, methodologies, and experimental results related to the development and implementation of the ADAS. Additionally, we will discuss the implications of this research, including its potential to reshape the automotive industry and pave the way for safer and more efficient transportation systems. Overall, this dissertation serves as a scholarly exploration into the development of an advanced driver assistance system that leverages computer vision and deep learning techniques. By tackling the pressing problem of human mistakes in traffic incidents, this initiative strives to significantly enhance road safety, ultimately protecting lives and reducing the societal and economic consequences of vehicle collisions.

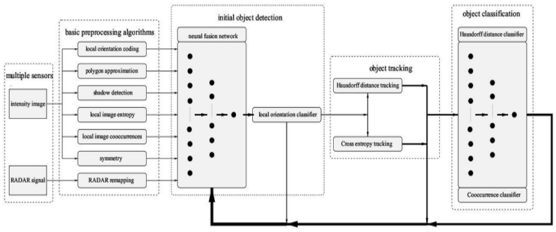

fig1: overall system proposed

1. Vehicle Detection and road segmentation for ADAS

The section explores Vehicle Detection and Road Segmentation as critical components in vision-based Advanced Driver Assistance Systems (ADAS). It begins by emphasizing the importance of object detection and image segmentation in identifying vehicles, pedestrians, lanes, traffic signs, and other road features. Classic image processing methods, such as transforming images into grayscale, removing backgrounds, and computing optical flow, are used to improve image clarity, identify moving objects, and track vehicle movement between frames. Post-processing methods such as morphological operations and blob analysis refine object boundaries and fill gaps for improved accuracy. Alternatively, machine learning-based approaches, particularly Convolutional Neural Networks (CNNs), have emerged as robust solutions, effectively learning patterns and features from extensive datasets. These data-driven models provide advanced capabilities for precise object detection and image segmentation tasks. Collectively, these methodologies play a pivotal role in enhancing situational awareness and driver assistance, underscoring their significance in creating safer and more efficient transportation systems.

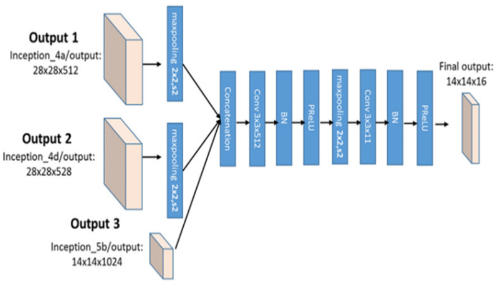

fig2: Fast vehicle detection's prediction module

In fig 2, it can be observed that the model first extracts low-level and high-level features from the given road image using the basenet GoogLeNet. Then the extracted features are fed to the proposed prediction module. In the prediction module, they also applied a couple of ConvNets and some other layers like batch normalization inorder to capture important features obtained from both lower and upper layers. Finally, the model was tested on different datasets and outperformed several popular models in the realm of object detection, such as Faster R-CNN, SSD300, SSD500 and more. When it comes to road segmentation for ADAS, deep learning techniques such as R-CNNs have been widely explored to tackle the problem. Researchers in [23] proposed an improved version of Mask R CNN to perform the task of road segmentation; specifically, the address segmenting pedestrians, cars and buses. The first main block of Mask R-CNNs is Feature Pyramid Networks (FPNs) with the basenet ResNet101, which is responsible for extracting features from the provided image. The authors there modified this component in two ways. Rather than using FPNs, they introduced an innovative approach termed Side Fusion Feature Pyramid Networks (SF-FPNs). They asserted that this method enhances model performance by integrating both highly semantic, low-resolution features and less semantic, high-resolution features, while aiming to minimize any increase in computational time.

2. Road Lane Detection

Lane Detection plays a vital role in vision-driven Advanced Driver Assistance Systems (ADAS), aimed at improving road safety by preventing unintentional lane drifts. This section explores the creation of a lane detection system that harnesses Convolutional Neural Networks (CNNs) and sophisticated image segmentation methods. The system utilizes two key architectures: Fully Convolutional Networks (FCN) and U-Net, both tailored for semantic segmentation tasks. These models proficiently detect and categorize lane characteristics in road visuals, ensuring consistent performance despite challenging scenarios such as shadows or obstacles. The TuSimple dataset, a standard for lane detection studies, is employed for training and assessing the models. The training process employs techniques like data augmentation to tackle diverse roadway scenarios. Model performance is measured using metrics like accuracy, F1-Score, recall, and Intersection over Union (IoU). Findings indicate that U-Net surpasses FCN, delivering superior segmentation quality and resilience, with higher accuracy and IoU values. The system incorporates cutting-edge tools, such as GPU acceleration to expedite training and inference, combined with adaptive thresholding to improve lane marker identification. These advancements support real-time functionality, ensuring the system is ideal for dynamic driving environments.

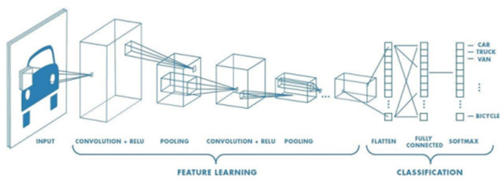

fig3: Overall workflow of CNNs

I. Fully Convolutional Networks(FCNs)

i. This subsection focuses on the application of Fully Convolutional Networks (FCN) for lane detection in Advanced Driver Assistance Systems (ADAS). FCNs are designed to perform end-to-end semantic segmentation by converting road images into lane marking maps. In contrast to conventional CNNs, FCNs maintain spatial dimensions within their design, rendering them well-suited for lane segmentation applications.

ii. The FCN model operates through multiple convolutional layers, enabling detailed feature extraction from road scenes. It replaces fully connected layers with upsampling operations, ensuring the output retains the resolution necessary for precise lane detection. The TuSimple dataset serves as the benchmark for training and testing the model. By leveraging FCN's pixel-wise classification capabilities, lane markings are identified with improved accuracy, even under challenging conditions like varying lighting or occlusions.

iii. Overall, FCN architecture demonstrates its effectiveness in real-time applications, contributing significantly to situational awareness and accident prevention in ADAS systems.

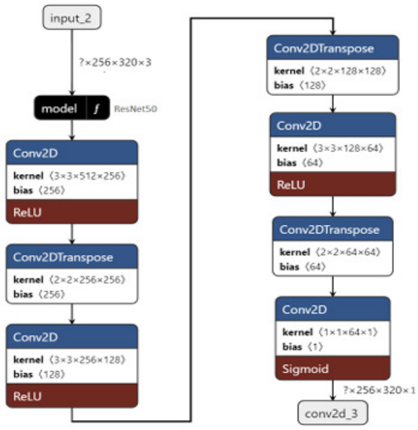

fig 4. The FCN architecture – lane detection

II. U-Net Architecture

i. The U-Net architecture is employed for lane detection tasks to enhance road safety by accurately identifying lane markings. U-Net, originally designed for biomedical image segmentation, and it is effective for segmenting road scenes. The model adopts an encoder-decoder structure, where the encoder extracts high-level features, while the decoder reconstructs spatial information, producing segmentation maps for lane detection.

ii. This approach leverages skip connections to combine features from different layers, ensuring the preservation of fine details in the lane markings. The TuSimple dataset, a benchmark in lane detection, is used for training the model, and extensive data augmentation techniques are applied to handle diverse road conditions. Performance is evaluated using metrics like accuracy, IoU, and F1-Score, showcasing the robustness of U-Net in handling complex scenarios such as shadows, occlusions, and varying lighting.

iii. The U-Net model delivers precise lane segmentation results, outperforming traditional methods and other architectures, thus contributing significantly to situational awareness and accident prevention in ADAS systems.

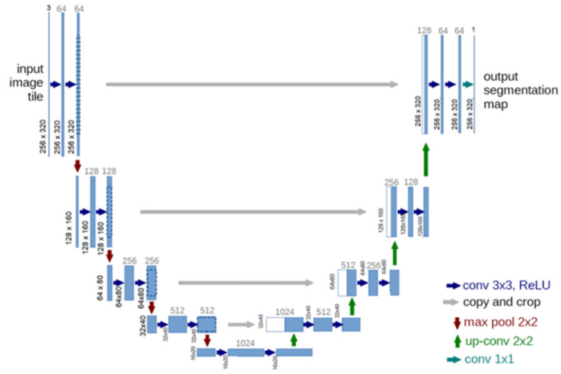

fig 5. U-Net architecture

III. Model Evaluation Metrics

To assess the effectiveness of our road lane detection model, we employ a range of evaluation metrics designed specifically for this task. These include accuracy, precision, recall, and F1 score. Below, we explain each metric in the context of road lane detection:

Accuracy reflects the overall correctness of the model's lane predictions, measuring the percentage of pixels correctly classified as either lane or non-lane. It is calculated by taking the sum of true positives (correctly identified lane pixels) and true negatives (correctly identified non-lane pixels) and dividing by the total number of pixels (true positives + true negatives + false positives + false negatives).

Accuracy = (TP+TN) /(TP+TN+FP+FN)

Recall, also referred to as Sensitivity, also known as sensitivity or true positive rate, evaluates the model's ability to correctly identify lane pixels out of all actual lane pixels present in the image. It is computed by dividing the true positives (correctly identified lane pixels) by the sum of true positives and false negatives (missed lane pixels).

Recall = (TP) / (TP+FN)

Precision assesses the accuracy of the model's predictions when identifying a pixel as part of a lane. It is determined by dividing the true positives (correctly identified lane pixels) by the sum of true positives and false positives (non-lane pixels incorrectly classified as lane pixels).

Precision = (TP) / (TP+FP)

The F1 score is a balanced measure that combines both precision and recall. It provides an overall assessment of the model's performance by considering their trade-off. The F1 score is calculated using the harmonic mean of precision and recall.

F1 score = 2×(Precision × Recall) /(Precision+Recall)

Intersection over Union (IoU) is a metric used to evaluate the performance of a segmentation model in accurately localizing objects or regions of interest. It measures the overlap between the predicted segmentation mask and the ground truth mask. IoU is calculated by dividing the area of intersection between the predicted and ground truth regions by the area of their union. It provides a measure of how well the model's predictions align with the ground truth annotations. IoU values range from 0 to 1, where 0 indicates no overlap between the predicted and ground truth regions, and 1 indicates a perfect match.

IoU = (Area of Intersection) /(Area of Union)

By evaluating our road lane detection model using these metrics, we can gain insights into its accuracy, ability to correctly identify lane pixels, and the balance between precision and recall. These metrics allow us to assess the model's performance and make informed decisions about its effectiveness for road lane detection tasks.

IV. Results

No	Models	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	IoU (%)
1	FCN	0.9783897	0.76614946	0.72483397	0.74491928	0.593523
2	U-Net	0.9790163	0.7661074	0.7456274	0.7557287	0.607366

Table 1: The lane detection models' performance metrics

Based on Table 1, the FCN model achieves a high accuracy of 97.84%, indicating its ability to correctly classify pixels as lane or non-lane. It demonstrates a precision of 76.61%, meaning it accurately identifies lane pixels among all the pixels classified as lane. The model achieves a recall of 72.48%, indicating its capability to detect actual lane pixels. The F1-Score, which balances precision and recall, stands at 74.49%, reflecting a solid overall performance. The IoU for the FCN model reaches 59.35%, showing the degree of overlap between predicted and actual lane boundaries. Comparatively, the U-Net model achieves a slightly higher accuracy of 97.90% than the FCN model. Both models share identical precision at 76.61%, correctly pinpointing lane pixels. However, the U-Net model outperforms the FCN model in recall, reaching 74.56%, which reflects its superior capability to identify lane pixels. Additionally, it records a higher F1-Score of 75.57%, showcasing a stronger balance between precision and recall compared to the FCN model. The U-Net model also achieves a greater IoU of 60.74%, indicating closer alignment with the actual lane boundaries. Both models perform well in road lane detection, with high accuracy and precision. The U-Net model shows improved performance in terms of recall, F1-Score, and IoU compared to the FCN model. These results suggest that the U-Net model has a better ability to accurately detect lane pixels and align with the ground truth lane boundaries. However, the differences between the models' performance are relatively small, and further analysis may be needed to determine the most suitable model for specific application requirements.

3. Road Segmentation

The road segmentation module is a vital element of the Advanced Driver Assistance System (ADAS), aimed at boosting road safety and driver awareness by precisely dividing the road scene into key components, such as vehicles, pedestrians, traffic signs, and other relevant objects. This approach utilizes advanced computer vision techniques to analyze visual data and provide immediate insights. Recognized for its diversity and superior quality, this dataset offers annotated road scene images that encompass a broad spectrum of driving situations and environments, serving as an outstanding tool.

The chosen architecture for this task is inspired by the U-Net framework, which excels in image segmentation due to its encoder-decoder structure. The encoder captures spatial details from input images, while the decoder reconstructs these details into segmented outputs. By connecting matching layers from the encoder to the decoder, skip connections help retain fine-grained information, which is crucial for precisely identifying road features. The adaptable structure and strong performance of the U-Net architecture make it highly effective for accurately labeling different components within road scenes.

This system empowers ADAS to provide substantial driver support, enhancing road safety by accurately detecting critical aspects of the driving environment.

V. Results

We evaluated the performance of our semantic segmentation model by assessing its ability to accurately segment 20 different objects/categories in an image. The evaluation was conducted using three commonly used metrics: overall pixel accuracy, intersection over Union (IoU), and the average Precision. For each object/category, we calculated the pixel accuracy, the IoU score, and average Precision to measure the model's performance in accurately classifying and segmenting the corresponding pixels.

Table 2 showcases the results obtained for each category. It presents a comprehensive summary of the model's performance.

We evaluated the model's performance across various classes using three main metrics: overall pixel accuracy, Mean Intersection over Union (IoU), and Mean Average Precision (mAP). The model achieved 85.79% in pixel accuracy, a mean IoU of 36.67%, and a mAP of 51.52%. Nonetheless, limitations in memory forced us to reduce the image resolution significantly, which hindered the model's ability to effectively learn segmentation patterns for specific categories like trains and radar. As a result, this constraint negatively influenced the evaluation outcomes.

Category ID	Categories	IoU	Average Precision
10	sky	0.9242	0.9659
0	road	0.8667	0.9177
13	car	0.8136	0.8906
8	vegetation	0.7845	0.8640
2	building	0.7197	0.8139
19	unknown	0.6369	0.7845
1	sidewalk	0.4694	0.6567
9	terrain	0.3557	0.5386
15	bus	0.2802	0.7941
6	traffic light	0.2622	0.4997
5	pole	0.2553	0.5259
11	person	0.2486	0.4522
14	truck	0.2116	0.3839
7	traffic sign	0.1631	0.3694
4	fence	0.1386	0.2874
3	wall	0.1048	0.2679
17	motorcycle	0.0528	0.173
18	bicycle	0.0354	0.0724
12	rider	0.0114	0.0472
16	train	0.0000	0.0000

Table 2: Performance Measurements of the Model on the Validation Set

4. Forward Collision Warning (FCW) System

Forward Collision Warning (FCW) System, which is an important component of the Advanced Driver Assistance System (ADAS). This section introduces the methodology used to analyze potential collision risks by leveraging segmented road information. The system identifies hazards within the Region of Interest (RoI) and provides real-time warnings to the driver. It highlights how the FCW system enhances safety by proactively alerting drivers to potential collisions, thereby reducing the likelihood of accidents.

I. Results

The performance of the FCW system widely relies on the image segmentation that is used as the backbone of the system. Therefore, based on the performance of the road segmentation discussed above, the system achieved a good performance at detecting the presence of extra objects within the designated RoI. The illustrations in Figure 7 demonstrate some outcomes of the system.

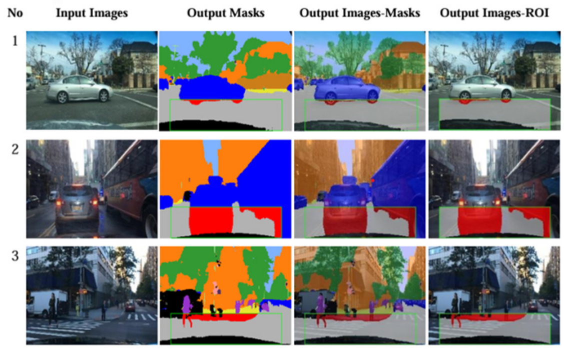

fig 6. Outcomes of the FCW system

5. Combining All Three Components

The vision-based ADAS integrates three key components: the road lane detection model, road segmentation model, and a simple Forward Collision Warning (FCW) system. The workflow begins with processing captured road images using the lane detection model to identify lane markings and generate a mask. The road segmentation model then processes this mask, assigning specific pixel values to classify different road components into separate categories. Finally, the FCW system utilizes the segmented mask to identify potential hazards within the Region of Interest (RoI), marking them red to alert the driver. Combined, these elements improve situational awareness, helping drivers better understand their environment.

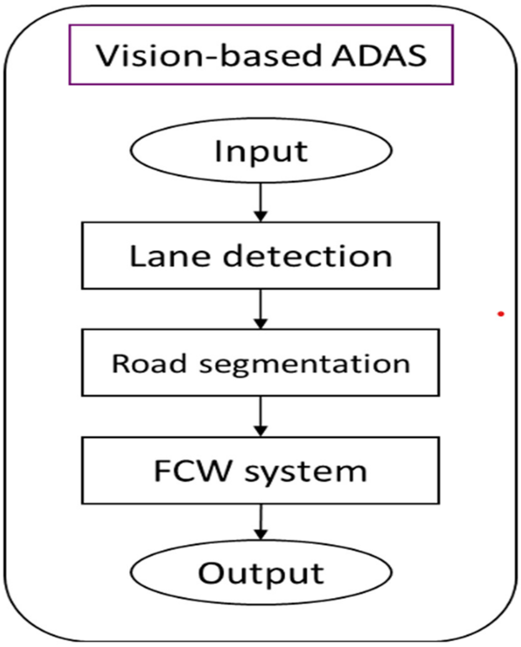

fig 7. ADAS as a Whole

6. Conclusion

The combination of lane detection, road segmentation, and Forward Collision Warning (FCW) systems creates a powerful, vision-based Advanced Driver Assistance System (ADAS). Leveraging deep learning and computer vision, it significantly improves road safety and offers real-time assistance to drivers.

The lane detection component uses Convolutional Neural Networks (CNNs) to recognize lane markings, helping to prevent unintended lane changes. The U-Net architecture outperforms FCN in key metrics such as accuracy, recall, F1-Score, and IoU. For road segmentation, the U-Net model with skip connections classifies various road elements, including vehicles, pedestrians, and traffic signs. Despite the challenge of downscaling images, it achieves an overall accuracy of 85.79%, a mean IoU of 36.67%, and a mAP of 51.52%.

The FCW system processes segmentation data to detect potential hazards within the Region of Interest (RoI) and provides timely alerts to the driver. This integration of various components boosts situational awareness, helping drivers better understand their surroundings and make informed decisions. Through the continued enhancement of these technologies, ADAS contributes to safer and more secure roadways for everyone.

7. References

[1] “Road traffic injuries”, WHO. Accessed: May 02, 2023. [Online]. Available: WHO Fact Sheet.
[2] S. Chen, M. Kuhn, K. Prettner, and D. E. Bloom, “The global macroeconomic burden of road injuries: estimates and projections for 166 countries”, vol. 3, no. 9, Sep. 2019, doi: 10.1016/S2542-5196(19)30170-6.
[3] 3N. S. Aminuddin et al., “A new approach to highway lane detection by using hough transform technique”, vol. 16, no. 2, Nov. 2017, doi: 10.32890/JICT2017.16.2.8231.
[4] W. Fan et al., “Median Filtered Image Quality Enhancement and Anti-Forensics via Variational Deconvolution”, vol. 10, no. 5, Jan. 2015, doi: 10.1109/TIFS.2015.2398362.
[5] S. Hiremath and B. Shreenidhi, “Dynamic Hough transform for robust lane detection and navigation in real time”, vol. 3, no. 1, Apr. 2023, doi: 10.54646/bijscit.2022.28.
[6] S. Kumar et al., “An efficient approach for highway lane detection based on the Hough transform and Kalman filter”, vol. 7, no. 5, Oct. 2022, doi: 10.1007/s41062-022-00887-9.
[7] A. Istiningrum et al., “Lane Detection With Conditions of Rain and Night Illumination Using Hough Transform”, Aug. 2022, doi: 10.1109/ICOIACT55506.2022.9972068.
[8] R. Li and Y. Dong, “Robust Lane Detection through Self Pre-training with Masked Sequential Autoencoders and Fine-tuning with Customized PolyLoss”, vol. abs/2305.17271, May 2023, doi: 10.48550/arXiv.2305.17271.
[9] P. H. Phan et al., “Robust Autonomous Driving Control using Auto-Encoder and End-to-End Deep Learning under Rainy Conditions”, Jul. 2023, doi: 10.1145/3591569.3591616.
[10] Q. Zou et al., “Robust Lane Detection From Continuous Driving Scenes Using Deep Neural Networks”, vol. 69, no. 1, Apr. 2020, doi: 10.1109/TVT.2019.2949603.
[11] V. Badrinarayanan et al., “SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation”, Oct. 2016, doi: 10.48550/arXiv.1511.00561.
[12] O. Ronneberger et al., “U-Net: Convolutional Networks for Biomedical Image Segmentation”, Jan. 2015, doi: 10.48550/arXiv.1505.04597.
[13] TuSimple. (2021). TuSimple Lane Detection Challenge Dataset [Dataset]. Retrieved from Kaggle: TuSimple Dataset.

[14] F. Hasan et al., "Multiple Vehicle Detection and Segmentation in Malaysia Traffic Flow", in IOP Conference Series Materials Science and Engineering, IOP, Mar. 2018, doi: 10.1088/1757-899X/318/1/012041.

[15] U. Handmann et al., "Computer vision for driver assistance systems", in Proceedings of SPIE, International Society for Optics and Photonics, doi: 10.1117/12.317463.

[16] J. Pyo et al., "Front collision warning based on vehicle detection using CNN", ISOCC2016, Oct. 2016, doi: 10.1109/ISOCC.2016.7799842.

[17] J. C. Hayward, "Near-miss determination through use of a scale of danger", no. 384, Jan. 1972.

[18] Tarmizi and A. A. Aziz, "Vehicle Detection Using Convolutional Neural Network for Autonomous Vehicles", in 2018 International Conference on Intelligent and Advanced System (ICIAS), Aug. 2018, doi: 10.1109/ICIAS.2018.8540563

