AlertDriver: a Real-Time Distraction Detection Solution

S Annapoorna, Indu Pagadala, Mahitha Dondeti, Priyanvitha Bethapudi and Nithya Kaandru

May 26, 2024

AlertDriver: A Real-Time Distraction Detection Solution

S Annapoorna[1[0000−1111−2222−3333]], Pagadala Indu[2], Dondeti Mahitha[3], Bethapudi Priyanvitha[4], and Kaandru Nithya[5]

[1] Assistant Professor, BVRIT HYDERABAD College of Engineering for Women, annapoorna.s@bvrithyderabad.edu.in
[2] B.Tech Student,BVRIT HYDERABAD College of Engineering for Women, 20wh1a6623@bvrithyderabad.edu.in
[3] B.Tech Student,BVRIT HYDERABAD College of Engineering for Women, 20wh1a6625@bvrithyderabad.edu.in
[4] B.Tech Student,BVRIT HYDERABAD College of Engineering for Women, 20wh1a6629@bvrithyderabad.edu.in
[5] B.Tech Student,BVRIT HYDERABAD College of Engineering for Women, 21wh5a6604@bvrithyderabad.edu.in

Abstract. In today's modern age, transportation plays a crucial role in our daily routines, making road safety a top priority. Addressing this critical concern, we introduce AlertDriver—a real-time distraction detection system designed to significantly enhance driving safety by promptly identifying and addressing distracted driving behaviors. These distractions can include a range of activities such as using a mobile phone, drinking, or even falling asleep at the wheel.Through comprehensive evaluation using existing datasets, our experiments confirm the effectiveness of AlertDriver in accurately detecting and classifying instances of driver distraction. In essence, AlertDriver represents a proactive approach to improving road safety, harnessing advanced deep learning techniques for real-time distraction detection. With its capacity to swiftly identify and mitigate distracted driving behaviors, this solution holds great potential for reducing road accidents and ultimately saving lives.

Keywords: Distracted driving · Driver safety · Real-time distraction detection · AlertDriver · Driver behavior · Facial expressions · Computer vision · Yawning detection · Distraction behaviors · Alarm triggering.

1 Introduction

In today's era, transportation plays a pivotal role in our daily lives, yet road safety remains a significant concern globally. According to the World Health Organization (WHO), road accidents rank as the eighth highest cause of death worldwide, with factors like drowsy driving accounting for a substantial portion of these incidents. In fact, a study by the Central Road Research Institute highlights that drowsy drivers alone are responsible for 40% of total road accidents. Additionally, distracted driving and sleep-deprived driving contribute significantly to the toll of fatalities and injuries, with over 15 lives lost and more than 1,200 injuries occurring each day due to distracted driving, according to the Centers for Disease Control.To address these critical safety issues, we present AlertDriver—a sophisticated real-time distraction detection solution designed to revolutionize road safety. AlertDriver employs advanced deep learning techniques to promptly identify and mitigate distracted driving behaviors. The methodology begins with the collection and preprocessing of relevant datasets containing video footage of drivers exhibiting various behaviors such as distraction or drowsiness. Data cleaning, annotation, and normalization are performed to prepare the datasets for feature extraction.Key features such as facial expressions, head movements, and eye behaviors are extracted from the video frames using computer vision algorithms. These features serve as input for training a deep learning model, typically a convolutional neural network (CNN). The model is trained to accurately classify different states of driver behavior, including distracted, drowsy, or attentive.The trained model is deployed within the AlertDriver system to perform real-time distraction detection. As video input streams from onboard cameras, the model continuously analyzes the frames, detecting any signs of distraction or drowsiness. When the model identifies a potential instance of distracted driving or drowsiness, AlertDriver generates timely alerts. These alerts can be in the form of visual or auditory signals within the vehicle to prompt the driver to refocus their attention on the road.Through rigorous evaluation using real-world datasets, AlertDriver has demonstrated remarkable efficacy in accurately detecting and classifying instances of driver distraction. This innovative system represents a proactive approach towards enhancing road safety, leveraging technology to prevent accidents before they occur by issuing timely alerts to drivers.AlertDriver embodies a pivotal shift in road safety strategies, providing actionable insights and interventions to mitigate road accidents and ultimately save lives. With AlertDriver, we embark on a journey towards a safer, more secure future on our roads, harnessing the power of technology to protect drivers and passengers alike.

2 Literature Survey

In [1], Dr. D. Rosy Salomi Victoria, Dr. D. Glory Ratna Mary, a drowsiness detection model relies on a well-designed Convolutional Neural Network (CNN) architecture, specifically tailored to recognize drowsiness based on eye closure. The initial phase involved creating image datasets containing instances of both open and closed eyes. Seventy-five percent of the dataset was allocated for training the custom CNN, while the remaining 25% served as a test set. The implementation process begins by transforming information videos into frames, with each frame subjected to face and eye detection. The refined CNN is instrumental in automatically learning features that facilitate the classification of eye states (open or closed). An alarm is triggered if the model detects the eyes closing in 15 consecutive frames, serving as a warning for the driver. The proposed CNN exhibits a training accuracy of 97% and a testing accuracy of 67%. For future enhancements, additional facial characteristics may be incorporated to improve detection accuracy. Moreover, there is potential to integrate information on the vehicle's driving patterns obtained through On-Board Diagnostics sensors with the extracted facial features, offering a more comprehensive approach to drowsiness detection.

In [2], Ana-Maria Băias, u, Cătălin Dumitrescu, this paper outlines the utilization of the Haar algorithm for image processing and analysis. We have identified the optimal threshold within the application for achieving the most accurate identification results. By implementing this system within vehicles, it has the potential to significantly reduce the number of road accidents caused by driver fatigue, drowsiness, or inattentiveness. To assess driver drowsiness, the initial step involved identifying key facial regions of interest, including the eyes, nose, mouth, and eyebrows. Subsequently, the focus was on analyzing the eye region. The application calculates the Euclidean distance both vertically and horizontally between the eyes. Based on the Eye Aspect Ratio (EAR) result obtained, it is determined whether the eyes are open or closed. A notable contribution of this study lies in the ability to locate and track the driver's gaze, thereby identifying instances where attention may be diverted while driving. For enhanced accuracy in the future, consideration could be given to incorporating an infrared video camera. This would aid in identifying the driver's eyes during nighttime or in low-light conditions, particularly when the driver is wearing sunglasses.

In [3], Pothuraju Vishesh, Raghavendra S, Santosh Kumar Jankatti, Rekha V, The eye blink detection project successfully achieved its goal of identifying the condition of the eyes in a given image. The primary objective of this project is to contribute to the reduction of road accidents caused by human errors. Various applications for eye blink detection have been developed to address this issue, and the current model presented here is an improved version designed to enhance the detection rate and recognize potential threats more effectively. The approach taken involves retraining the Google MobileNetV2 model to specifically identify eye blinks. The ultimate aim is to enhance road safety for travelers, and the model achieved an impressive accuracy rate of 97%. In future endeavors, the focus will shift towards detecting drowsiness in individuals wearing sunglasses, situations where only one eye is closed, and side views of faces. The overall objective remains making roads safer for everyone.

In [4], Peng Mao, Kunlun Zhang, Da Liang, This paper has introduced a novel driver distraction behavior detection method based on deep learning, aiming to address the escalating safety concerns associated with distracted driving in China. The proposed approach leverages state-of-the-art techniques, including Progressive Calibration Networks (PCN) for precise face detection, Discriminative Scale Space Tracking (DSST) for dynamic face tracking, and the YOLOV3 object detection algorithm for identifying specific distracted behaviors such as smoking and making phone calls in real-time. The research outlines a comprehensive architecture encompassing real-time video data collection, face detection, face tracking, and distraction behavior judgment. Notably, the method has undergone rigorous evaluation against the standalone YOLOV3 approach, showcasing its superior performance in terms of detection accuracy, recall, and speed. This innovative solution exhibits a robust framework capable of effectively enhancing driver safety by accurately identifying distracting behaviours during the driving process. Despite its high recognition rates, the paper acknowledges potential challenges related to false positives, particularly in scenarios where objects resembling smoke may trigger detection errors. The study emphasizes the importance of future research, proposing avenues such as posture estimation for smoking behavior, to refine the method's precision and reduce false detections. In essence, this research represents a significant contribution to the field, offering a promising solution for real-world applications in the ongoing effort to mitigate distracted driving accidents.

In [5], Taner Danisman, Ian Marius Bilasco, Chabane Djeraba, Nacim Ihaddadene, This paper has introduced an innovative method for detecting eye blinks, leveraging the symmetry property and operating independently of head movements within the same frame. Distinguishing itself from systems relying on statistical information from previous frames, the proposed approach demonstrates a commendable frame rate of 110fps on an Intel Xeon 2.9 GHz CPU, accommodating real-time scenarios and allowing for concurrent tasks. While the absence of a common database limits a direct comparison of drowsiness detection results, the paper currently provides detailed insights into eye-blink detection outcomes. Future plans involve extending the evaluation to the Av@Car database for a more comprehensive assessment. Notably, the proposed system exhibits a 94% accuracy in detecting eye blinks, accompanied by a low 1% false positive rate. Real-world experiments emphasize the impact of factors such as the presence of glasses and variations in illumination on the system's performance, influencing critical components like face detection, eye detection, and symmetry calculation.

3 Dataset Description

Our project relies on a publicly available dataset specifically tailored for distracted driving and drowsiness detection. Sourced from a reputable public repository, this dataset provides a comprehensive collection of annotated instances, encompassing distracted driving behaviours like yawning, eye closing, and eye opening, as well as drowsiness-related actions such as phone usage and prolonged inattention.

3.1 Eye Dataset

This dataset comprises images of human eyes categorized into two classes: open eyes and closed eyes. The images, collected from datasets consists of images under diverse lighting conditions. Each image is in JPEG format with varying

Fig. 1. Sample images from eyes dataset

resolutions, predominantly 640x480 pixels, and organized into respective 'Open' and 'Closed' folders. The dataset is split into training (80%), testing (20%) sets, ensuring a balanced distribution of classes. This dataset, useful for eye state classification for drowsiness detection.

3.2 Yawn Dataset

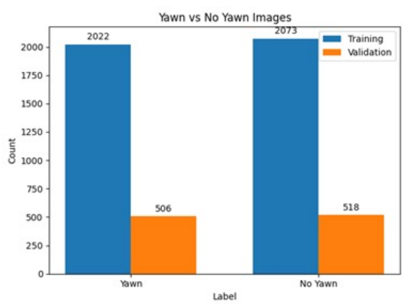

Fig. 2. Distribution of images

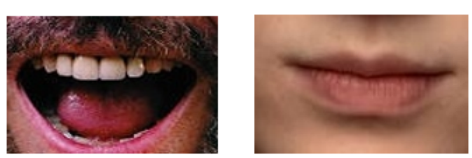

Fig. 3. Sample images from yawn dataset

The dataset comprises two classes: "Yawn" and "No Yawn". Images in the "Yawn" category depict individuals with their mouths open, indicating a yawn, while images in the "No Yawn" category depict individuals with closed mouths or engaging in activities where yawning is unlikely.

3.3 Distraction Dataset

Fig. 4. Sample images from distraction dataset

We collected a dataset specifically focused on detecting distracted driving behaviours, comprising four distinct classes: DangerousDriving, Distracted, Drinking, and SafeDriving. Each class represents different driving behaviours and conditions encountered on the road.
Additionally, the dataset has following:
Auto-Orient: Applied to standardize the orientation of the images within the dataset. Isolate Objects: Utilized to isolate relevant objects or regions of interest within the images, focusing on the driver's behaviour and surroundings.
Resize: Images were resized to fit within a standardized resolution of 640x640 pixels, ensuring consistency across the dataset.
Flip: Horizontal flipping was applied as an augmentation technique to introduce variations in the orientation of the images, enhancing the model's ability to generalize across different viewing angles.

4 Methodology

4.1 System architecture

Fig. 5. depicts the flow of the system to be created. The system begins by monitoring the driver's behaviour using a webcam, which captures a continuous video stream. The systematic approach combines computer vision techniques for face and eye detection with deep learning methodologies (specifically CNNs) for real-time drowsiness detection. By incorporating both drowsiness and distraction detection into the system using CNN models and computer vision techniques, the AlertDriver system aims to proactively mitigate risks associated with driver

fatigue and distractions, ultimately enhancing road safety. By continuously monitoring driver behaviour and providing timely alerts, the system aims to enhance road safety by mitigating risks associated with driver fatigue and drowsiness. Ongoing refinement of the CNN model and data collection processes contributes to improving the accuracy and reliability of the drowsiness detection system.

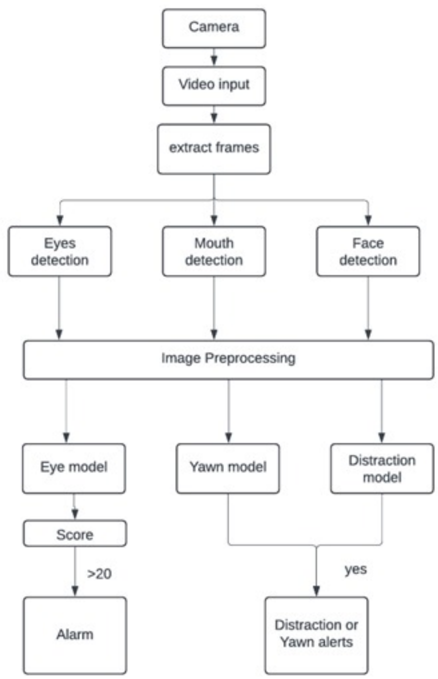

Fig. 5. System architecture

4.2 Implementation

The input video of the driver is taken from the camera placed in front of driver, the video is divided into frames.

Face Detection We utilized the dlib library for face detection and facial landmark localization. The Histogram of Oriented Gradients (HOG) based face detector from dlib was employed to detect faces in each frame of the driver's video feed. Once a face was detected, the pre-trained shape predictor model from dlib was used to localize 68 facial landmarks, such as the eyes, eyebrows, nose, mouth, and jawline. These landmarks, represented as points on the face, provide critical spatial information for subsequent analyses. Visualizations of the detected landmarks were overlaid on the face to illustrate their positions, aiding in further analysis of facial expressions and features.

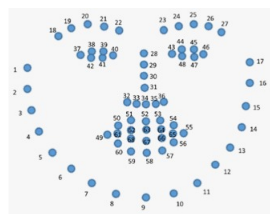

Fig. 6. Face Landmark representation

Fig. 7. Face detection in frame

Drowsiness Detection The mouth is accessed through points [48, 67].
The left eye is accessed with points [42, 47].
The right eye is accessed using points [36, 41].

A. Eye State Detection The face detection algorithm first identifies faces within the input image, then proceeds to detect eyes using a pre-trained eye cascade. For each detected eye, it is cropped, resized, and normalized to prepare for prediction. The model then assesses whether the eye is open or closed. If the prediction suggests closed eyes, a score is incremented; once this score surpasses a predefined threshold, an alarm sound is triggered. Conversely, if the prediction indicates open eyes, the score is decremented. This iterative process enables continuous monitoring of eye states, facilitating timely alerts if drowsiness or fatigue is detected.

B. Yawn Detection The localized mouth region is cropped, resized, and prepared for analysis. The model is then employed to predict whether the mouth is in a yawning state or not. If a yawn is detected a alert is given.

Distraction Detection The model is trained to classify frames into four distinct categories: Dangerous Driving, Distracted, Drinking, and Safe Driving. It operates directly on the input frames without the need for facial landmarks or other detailed facial analysis. Each frame is processed through the trained model, which assigns it to one of these classes based on the observed behaviour or state depicted in the frame. If any distraction is detected an alert is given.

5 Analysis

5.1 Eyes

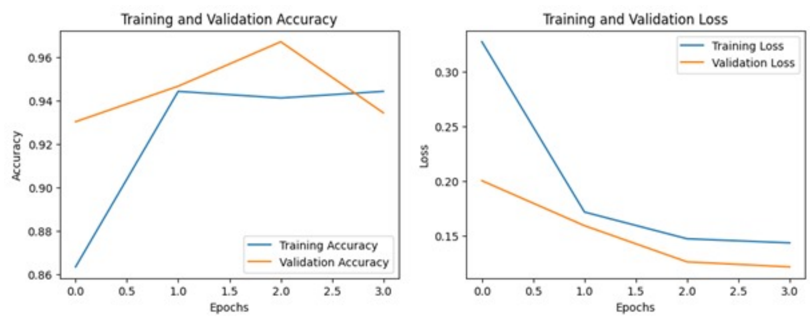

Fig. 8. Training and validation curves of InceptionV3 model on eyes dataset

In our study, we employed an InceptionV3 deep learning architecture to classify drowsiness based on eye images from a specialized dataset. The InceptionV3 model, pretrained on ImageNet, was fine-tuned for our task by adding custom dense layers on top. We utilized transfer learning to leverage the features learned by InceptionV3 on a large dataset to classify drowsiness in our specific domain.The model was trained over four epochs with a batch size of 4, achieving a training accuracy that increased from 86.34% to 94.4%. Our results demonstrate that the InceptionV3 model effectively learned to classify eye images related to drowsiness, indicating strong generalization capability on unseen validation data.

5.2 Eyes

We compared the performance of two deep learning architectures, Convolutional Neural Network (CNN) and Residual Network (ResNet), for the task of yawn de-

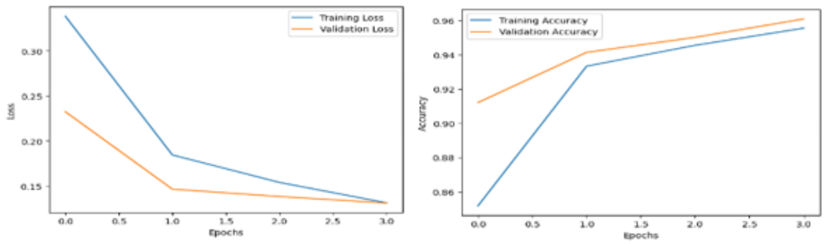

Fig. 9. Training and validation curves of CNN model on yawn dataset

tection. The CNN model out performed resnet model by demonstrated a training accuracy of 95.56% and a validation accuracy of 96.09%, showcasing its effectiveness in distinguishing between yawn and non-yawn.The model demonstrates a consistent decrease in training and validation loss, indicating effective learning and generalization. Moreover, the increasing trend in training and validation accuracy highlights the model's ability to accurately predict yawn instances. The classification report further reveals the model's proficiency in identifying yawns, albeit with some challenges in detecting instances of no yawn. These findings underscore the model's potential for real-world applications.

5.3 Eyes

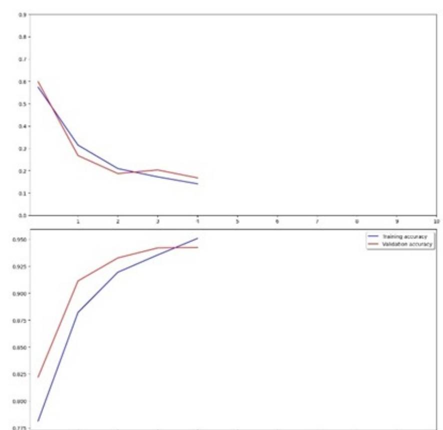

Fig. 10. Training and validation curves of CNN model on Distraction dataset

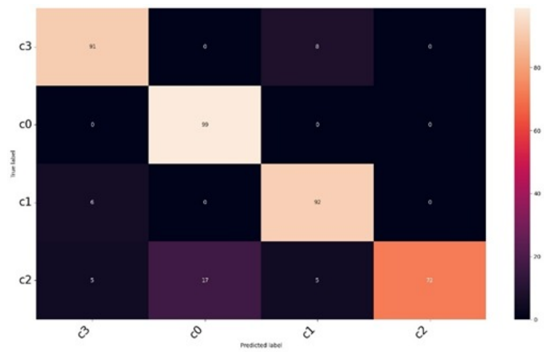

Fig. 11. Confusion matrix

We evaluated the performance of a Convolutional Neural Network (CNN) model for driver distraction classification across four categories: "DANGEROUS DRIVING", "DISTRACTED DRIVING", "DRINKING", and "SAFE DRIVING". The CNN model demonstrated strong performance with a training accuracy of 94.63% and a validation accuracy of 95.72%, showcasing its effectiveness in distinguishing between different distraction types. Figure shows model's consistent decrease in training and validation loss throughout the training process indicates robust learning and generalization ability. Moreover, the increasing trend in both training and validation accuracy highlights the model's capability to accurately predict instances of driver distraction.The classification report revealed the model's proficiency in identifying various distraction types, with notable success in categorizing "DANGEROUS DRIVING", "DISTRACTED DRIVING", and "SAFE DRIVING". These findings underscore the CNN model's potential for real-world applications in driver monitoring systems, where accurate and reliable distraction detection is critical for enhancing road safety. The confusion matrix generated provides valuable insights into the classification performance of the model across different classes. By analysing the distribution of predicted and actual class labels, we can assess the model's strengths and weaknesses, identify misclassifications, and gain actionable feedback for improving model accuracy and reliability. This visual representation is essential for evaluating and fine-tuning machine learning models in classification tasks. Nonetheless, the CNN's high accuracy and effective learning characteristics position it as a promising solution for driver distraction detection tasks.

6 Result

The real-time distraction detection solution we've developed encompasses a range of critical functionalities aimed at enhancing driver alertness and safety. By detecting various behavioral cues and facial attributes, our system provides proactive insights into the driver's state, helping mitigate potential risks associated

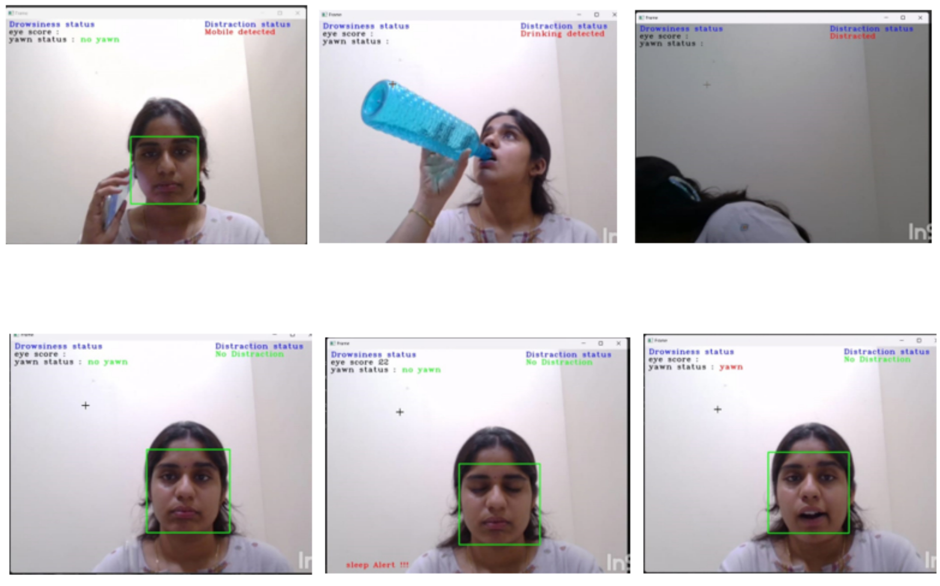

Fig. 12. Results

with driver distraction. Firstly, our solution includes the capability to detect signs of drowsiness, such as eyelid drooping, indicating driver fatigue. By identifying these subtle cues, our system can promptly alert drivers to take necessary breaks or corrective actions to prevent accidents caused by drowsy driving. Furthermore, the detection of yawning behaviour plays a pivotal role in our system. Yawning is often associated with fatigue and reduced alertness, serving as an important indicator of the driver's cognitive state. Our solution promptly recognizes yawning instances and alerts drivers to re-engage with the driving task or take rest as needed. Moreover, the detection of behaviours like drinking while driving is crucial for maintaining focus and control on the road. By identifying these activities, our system encourages safe driving practices and discourages distractions that can compromise driver attention and reaction times. Another key aspect of our solution is the monitoring of mobile phone usage while driving. Detecting instances of mobile phone handling provides essential feedback to drivers, promoting adherence to hands-free policies and minimizing the risks associated with distracted driving. Lastly, our system includes face alignment monitoring to ensure that the driver's face remains centred within the field of view. This feature helps maintain proper attention towards the road and reduces the likelihood of visual distractions caused by off-centre positioning. The integration of these advanced detection capabilities within our real-time distraction detection solution empowers drivers with actionable insights to maintain optimal focus and attentiveness while driving. By leveraging technologies, we aim to enhance road safety and mitigate potential accidents caused by driver distractions in diverse driving environments.

7 Conclusion & Future Scope

In conclusion, the development of our real-time distraction detection solution marks a pivotal advancement in promoting driver alertness and enhancing road safety. Leveraging state-of-the-art technologies in computer vision and machine learning, our system effectively identifies various driver behaviors and facial attributes indicative of distraction and fatigue. By promptly alerting drivers to potential risks and providing proactive interventions, our solution aims to significantly reduce the likelihood of accidents caused by driver distraction.

Looking ahead, there are several promising avenues for future research and development in this field. Firstly, continuous refinement and optimization of our detection algorithms will be crucial to enhance accuracy and robustness in identifying subtle behavioral cues and facial attributes associated with distraction and fatigue. Additionally, the integration of our distraction detection system with existing Advanced Driver Assistance Systems (ADAS) holds immense potential to provide comprehensive driver assistance and improve overall vehicle safety.Exploring the use of multi-modal sensors, such as eye-tracking and heart rate monitors, could further complement visual detection methods and offer a more comprehensive understanding of the driver's state. Furthermore, extensive validation and real-world testing of the system across diverse driving scenarios and conditions will be essential to ensure its reliability and effectiveness in practical applications.Collaborations with automotive manufacturers and industry stakeholders will play a vital role in integrating our solution into next-generation vehicles and contributing to the development of safer driving environments. By fostering ongoing innovation and collaboration, we aim to advance this technology and contribute significantly to the realization of safer and more secure transportation systems for drivers and passengers alike.

References

1. Victoria DR, Mary DG. Driver drowsiness monitoring using convolutional neural networks. In2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS) 2021 Feb 19 (pp. 1055 1059). IEEE.
2. Băias,u, Ana-Maria, and Cătălin Dumitrescu. "Contributions to driver fatigue detection based on eye-tracking." Int J Circuit Syst Signal Process 15 (2021): 1-7.
3. Vishesh, Pothuraju, et al. "Eye blink detection using CNN to detect drowsiness level in drivers for road safety." Indonesian Journal of Electrical Engineering and Computer Science 22.1 (2021):222-231.
4. Peng Mao, Kunlun Zhang, Da Liang ,"Driver Distraction Behaviour Detection Method Based on Deep Learning", Peng Mao et al 2020 IOP Conf. Ser.: Mater. Sci. Eng. 782 022012.

5. Danisman, Taner, et al. "Drowsy driver detection system using eye blink patterns." 2010 International Conference on Machine and Web Intelligence. IEEE, 2010.
6. Arif, Mubeen, et al. "A Real-Time Driver Drowsiness Detection and Warning System Based on an Eye Blinking Rate." Intelligent Technologies and Applications: Second International Conference, INTAP 2019, Bahawalpur, Pakistan, November 6–8, 2019, Revised Selected Papers 2. Springer Singapore, 2020.
7. Hashemi, Maryam, Alireza Mirrashid, and Aliasghar Beheshti Shirazi. "Driver safety development: Real-time driver drowsiness detection system based on convolutional neural network." SN Computer Science 1.5 (2020): 289.
8. Pandey, Nageshwar Nath, and Naresh Babu Muppalaneni. "Real-time drowsiness identification based on eye state analysis." 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS). IEEE, 2021.
9. Dewi, Christine, et al. "Eye aspect ratio for real-time drowsiness detection to improve driver safety." Electronics 11.19 (2022): 3183.
10. Phan, Anh-Cang, et al. "An efficient approach for detecting driver drowsiness based on deep learning." Applied Sciences 11.18 (2021): 8441.
11. J. Anitha, G. Mani and K. Venkata Rao, "Driver drowsiness detection using viola jones algorithm", Smart Intelligent Computing and Applications, pp. 583-592, 2020.
12. Kolpe, P., Kadam, P., Mashayak, U.: Drowsiness detection and warning system using python. In: International Conference on Communication and Information Processing (ICCIP). (2020)
13. Suhaiman, A.A., May, Z., Rahman, N.A.A.: Development of an intelligent drowsiness detection system for drivers using image processing technique. In: IEEE Student Conference on Research and Development. (2020)
14. Maior, C. B. S. et al. (2020) ‘Real-time classification for autonomous drowsiness detection using eye aspect ratio’, Expert Systems with Applications. doi: 10.1016/j.eswa.2020.113505.
15. Miah, Abdullah Arafat, Mohiuddin Ahmad, and Khatuna Zannat Mim. "Drowsiness detection using eye-blink pattern and mean eye landmarks’ distance." Proceedings of International Joint Conference on Computational Intelligence: IJCCI 2018. Springer Singapore, 2020.
16. Rajasekaran, Rajkumar, et al. "Enhancing Driver Safety: Real-Time Drowsiness Detection through Eye Aspect Ratio and CNN-Based Eye State Analysis." 2024 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE). IEEE, 2024.
17. Florez, Ruben, et al. "A cnn-based approach for driver drowsiness detection by real-time eye state identification." Applied Sciences 13.13 (2023): 7849.
18. Al-Razgan, Muna S., Issema Alruwaly, and Yasser A. Ali. "Eye-Blink Event Detection Using a Neural-Network-Trained Frame Segment for Woman Drivers in Saudi Arabia." Electronics 12.12 (2023): 2699.
19. Ahmed, M.; Laskar, R. Eye center localization using gradient and intensity information under uncontrolled environment. Multimed. Tools Appl. 2022, 81, 7145–7168.
20. Bitkina, O.V.; Park, J.; Kim, H.K. The ability of eye-tracking metrics to classify and predict the perceived driving workload. Int. J. Ind. Ergon. 2021, 86, 103193.