World Journal of Advanced Research and Reviews

eISSN: 2581-9615 CODEN (USA): WJARAI

Cross Ref DOI: 10.30574/wjarr

Journal homepage: https://wjarr.com/

(Research Article)

Check for updates

Real-time driver drowsiness and distraction detection using convolutional neural network with multiple behavioral features

Jaira -Salayon- Hernandez *, Fernando -Teston- Pardales Jr., Neña Mae -Sumaylo- Lendio, Ian Exequiel -Sibayan- Manalili, Eufemia -Acol- Garcia and Antonio -Calumba- Tee Jr.

Department of Electronics Engineering, College of Engineering, Pamantasan ng Lungsod ng Maynila, Manila, Philippines.

World Journal of Advanced Research and Reviews, 2024, 23(01), 816–824

Publication history: Received on 23 May 2024; revised on 06 July 2024; accepted on 08 July 2024

Article DOI: https://doi.org/10.30574/wjarr.2024.23.1.1976

Abstract

Road accidents caused by driver drowsiness and distraction represent significant threats to worldwide road safety, with fatalities and injuries at alarming rates in the Philippines. With a significant number of casualties, the need for proactive measures is urgent. Recognizing the human factor as the primary cause of accidents, this study aimed to develop a real-time driver drowsiness and distraction detection system to mitigate risks. Using non-intrusive camera sensors and convolutional neural networks (CNN), the system monitors the driver's behavior, including facial expressions, eye movements, and lane position, to detect signs of drowsiness and distraction. This study meticulously outlines the systematic procedures, employing a quantitative developmental research approach to design and assess the effectiveness of the system. Real-world on-road testing with participants engaged in long-duration driving ensures the authenticity of data collection. The findings highlight the system's promising performance in drowsiness and distraction detection, with high accuracy rates and an effective alert system triggered upon detection of potential risks. The integration of CNN technology underscores the system's potential to significantly enhance road safety, offering immediate benefits for drivers, vehicle manufacturers, and road safety authorities. This research sets a foundation for future advancements in proactive driver safety technologies, emphasizing the critical importance of addressing driver drowsiness and distraction on the roads.

Keywords: Convolutional Neural Network; Behavioral indicators; Alert system; Distraction Detection System; Drowsiness Detection System

1. Introduction

Drowsiness and distraction remain major concerns for road safety. Therefore, it is crucial to identify and detect these issues among drivers in a timely manner to prevent further accidents. Drowsiness and distraction while driving can lead to slower reaction times, reduced situational awareness, and loss of vehicle control [1].

Many researchers are motivated to find a method for lowering the number of traffic accidents. However, inherent intrusiveness in these systems remains a problem [2]. [3] presented a real-time system for monitoring the biological signals of the driver, including respiration, gripping force, and a photoplethysmogram (PPG). Driver monitoring systems (DMS) that employ camera sensors can detect drowsiness and distraction accurately without requiring the driver to wear uncomfortable or intrusive devices but suffers from the issue of illumination. System testing took place in a simulated environment or in controlled settings with adequate lighting [4,5]. The study of [6] implemented a driver monitoring system using Microsoft Kinect.

With the underlying issues, the researchers aim to develop a real-time driver drowsiness and distraction detection system (DDDDS) using Convolutional Neural Network with Eye Aspect Ratio (EAR), Mouth Aspect Ratio (MAR), Eye

* Corresponding author: Jaira Hernandez

Gaze, and Lane Departure to predict signs of driver distraction and drowsiness. In particular, the study aims to: (1) develop a drowsiness detection system that can detect driver drowsiness events by implementing Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR) using the convolutional neural network (CNN); (2) develop a distraction detection system detecting instances of driver distraction through eye gaze analysis using the convolutional neural network; (3) develop a lane departure detection system analyzing lane position of the vehicle; and (4) design a driver safety alert system that can warn the driver in the event of drowsiness and/or distraction detection through audio cues.

2. Material and methods

2.1. System Architecture

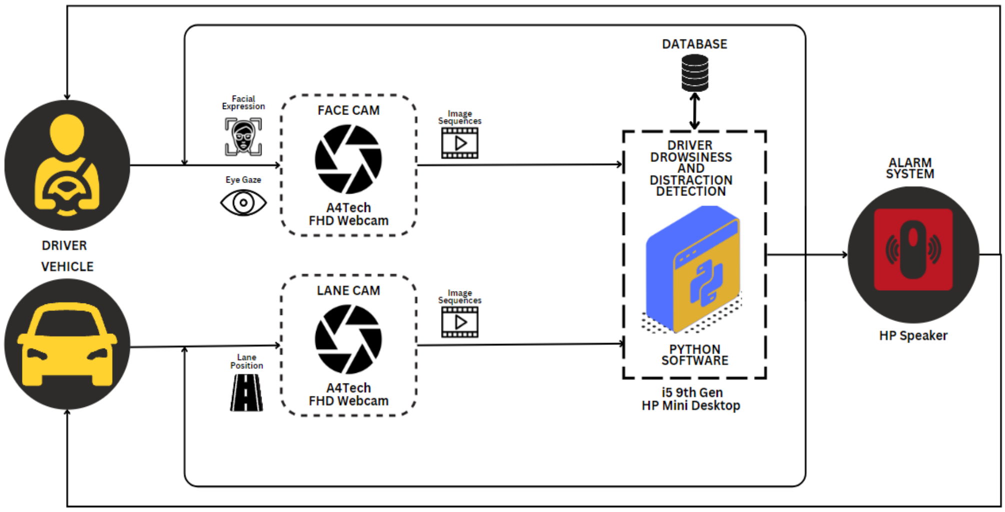

Figure 1 System Architecture for Driver Drowsiness and Distraction Detection System

Figure 1 illustrates the system architecture for detecting driver drowsiness and distraction using several features. The system utilizes two A4Tech FHD web cameras: a face cam that captures the driver's facial expressions and eye gaze direction and a road-facing lane cam. The real-time face cam feed will be processed using Python embedded in the i5 9th Gen HP Mini Desktop, using computer vision and machine learning techniques to detect the driver's facial expressions and eye gaze. The program monitors the behavioral features of the driver, including Eye Aspect Ratio (EAR), Mouth Aspect Ratio (MAR), and eye gaze. Meanwhile, the road cam captures the road environment, and the Python software will also analyze its feed. The software uses lane detection algorithms to detect the lane markings and the road and record the vehicle's lane position. This enables the system to detect misalignment from the standard lane position or driver drifting out from the lane due to drowsiness or distraction. A driver drowsiness and distraction algorithm is developed using behavioral elements taken from the driver's face by the camera focused on the driver and road camera lane detecting data. The algorithm uses CNN, a deep learning system commonly used in image and video processing. The alarm system is triggered if the level of drowsiness or distraction exceeds predetermined algorithm thresholds. The alert mechanism for the alarm system will be the HP Speaker, which will produce an audible alert through the speaker, which will alert the driver.

Figure 2 shows the system flowchart for the drowsiness and distraction detection system, entailing the subsystems of drowsiness detection system, distraction detection system and lane departure detection system integrated to safety alert system. Each subsystem also runs through several processes which monitor and determine instances of drowsiness, distraction, and lane departure, which then the alert system will take an auditory alert if such instances are detected. The researchers trained the models on validated, open-source datasets for drowsiness detection, distraction detection, and lane departure detection.

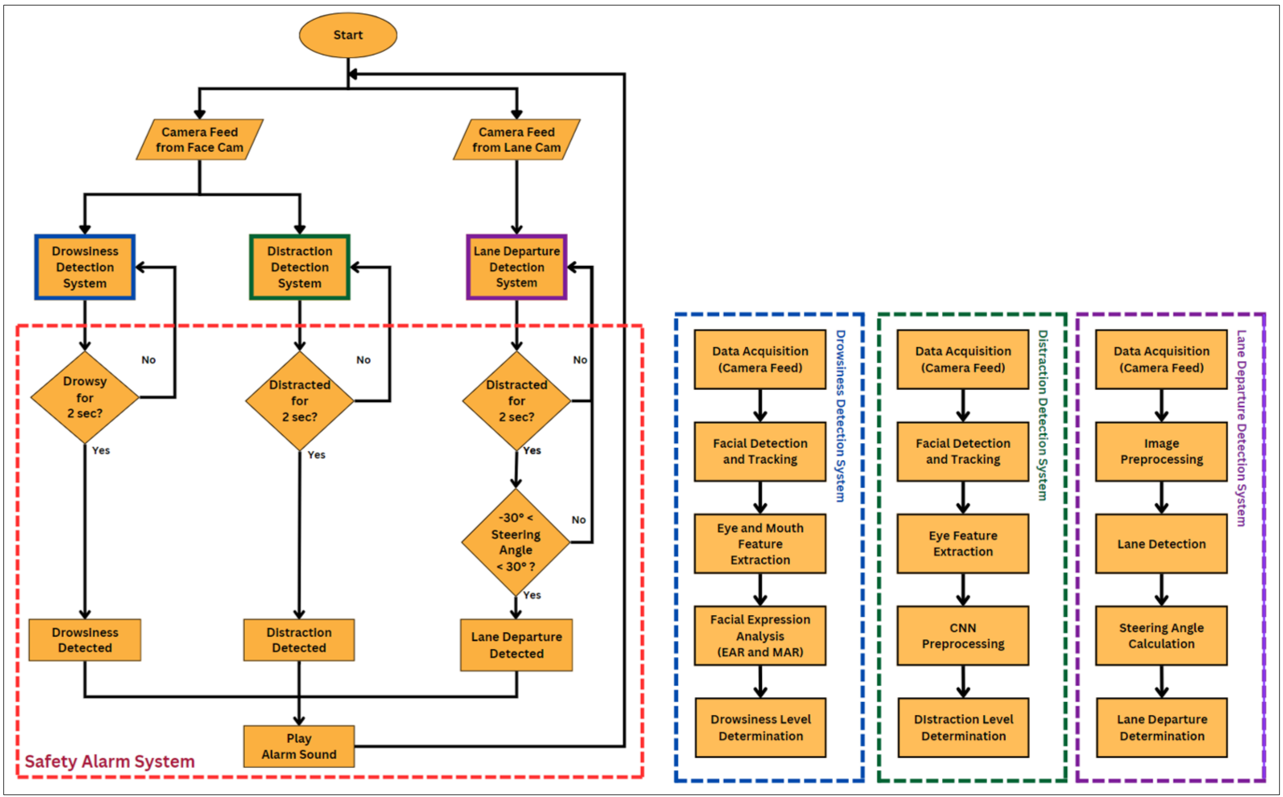

Figure 2 System Flowchart for Driver Drowsiness and Distraction Detection System

2.1.1. Drowsiness Detection System

The drowsiness detection system begins with acquiring a video stream from a webcam camera feed, which records a sequence of image frames where the camera is directed to the driver's face. During the face detection and tracking phase, the driver's face will then be located and tracked within the video stream captured by the camera through algorithms to identify facial landmarks and establish a bounding box around the face. The image frames are preprocessed by converting the color space from BGR to grayscale, and the face detector from the dLib is utilized to detect the face in the grayscale frames. After face detection, the facial landmarks are identified by dLib, and 68 landmark locations are detected for feature extraction. The relevant features, the left eye, right eye, and mouth, will be extracted to calculate drowsiness-related parameters. Facial expression analysis follows, where EAR and MAR are computed based on the extracted eye and mouth features, respectively. EAR and MAR provide insights into eye closure and mouth opening, which are essential indicators of drowsiness. The final step involves drowsiness level determination. Predefined thresholds are employed in this process to ascertain the overall behavioral drowsiness level of the driver.

2.1.2. Distraction Detection System

The distraction detection system starts by acquiring video data from a camera, typically a webcam or a dashboard-mounted camera facing the driver. The video frames are then processed using OpenCV's Haar cascade classifiers. The first Haar cascade classifier detects and tracks the driver's face within each video frame, providing the bounding box coordinates of the detected face region. The system employs another Haar cascade classifier trained to detect eye regions, applying to the detected face region to extract the bounding box coordinates of the driver's eyes. The extracted eye regions then undergo preprocessing steps to prepare the input data for the CNN model. The preprocessing involves extracting the color eye region from the frame, resizing it to match the CNN's input shape, normalizing the pixel values, and converting the eye image to a Numpy array with the appropriate dimensions required by the model. The preprocessed eye images are then fed into the pre-trained CNN model, which has been specifically trained to detect distraction based on eye movements and appearance. The CNN model analyzes the eye images and outputs the probability of distraction for each eye. The system then calculates the average prediction probability across all detected eyes to determine the driver's overall distraction level, with 0.5 as the baseline.

2.1.3. Lane Departure Detection System

The lane departure detection system is designed to perform lane detection and steering angle calculation on a video input, starting by reading frames from a video file using OpenCV's VideoCapture object. Each frame then goes through several image preprocessing steps: color filtering, Sobel edge detection, region of interest selection, and perspective transformation. With the preprocessed image, the lane detection process begins. A sliding window approach is used to detect the lane lines if no previous lane line fit is available, or the detected lane width is outside the expected range. Suppose a previous fit is available and the lane width is within the expected range. In that case, polynomial curves are fitted to the detected lane lines using the previous fit as a starting point, making the detection more efficient. The steering angle is calculated based on the position and curvature of the detected lane lines relative to the center of the image. This angle calculation is performed by determining the angle function. The outputs are the processed video with the detected lane lines and steering angle text overlaid on each frame. The processed frames are written to a new video file using OpenCV's Video Writer object. Additionally, timestamp is written, and the steering angle and the corresponding video time are calculated in a CSV file for further analysis or data logging.

2.1.4. Safety Alert System

The safety alert system will output an auditory response once the driver's distraction or drowsiness and the vehicle's lane departure are detected. Starting from the data acquisition of the driver's facial feature and the road from the camera feed, the data will be processed through the different feature detection algorithms consisting of drowsiness detection, distraction detection, and the lane departure system. For the drowsiness detection, the system should detect the driver to be showing behaviors of being drowsy for at least two seconds before determining that the driver is in the drowsy state; once the driver is deemed drowsy, the alert system will be triggered; if the driver is awake, the system will continue to evaluate the data of the driver's facial features until drowsiness occurs. The same condition applies to distraction detection. When the system detects that the driver is distracted for two seconds, it will interpret it as the driver being distracted, and the alarm system will be triggered; if not within the two-second mark, the driver is considered alert, and the system will continue to evaluate the driver's facial features until distraction becomes apparent. For the lane departure system, before the alert system can be triggered, the steering angle of the vehicle should be above 30 degrees or below -30 degrees while the driver is in the drowsy or distracted phase as it serves as a supplemental detection feature; if not, then the system will continue to monitor the vehicles steering angle and drowsiness or distraction evaluation. The alert system has a feedback loop which means that if the alarm has been triggered, it will continue to create an output sound until the system detects that the driver is not within any of the measures of drowsiness and distraction anymore as it continues to process the data being gathered.

2.2. Implementation and Testing

2.2.1. Prototype Development

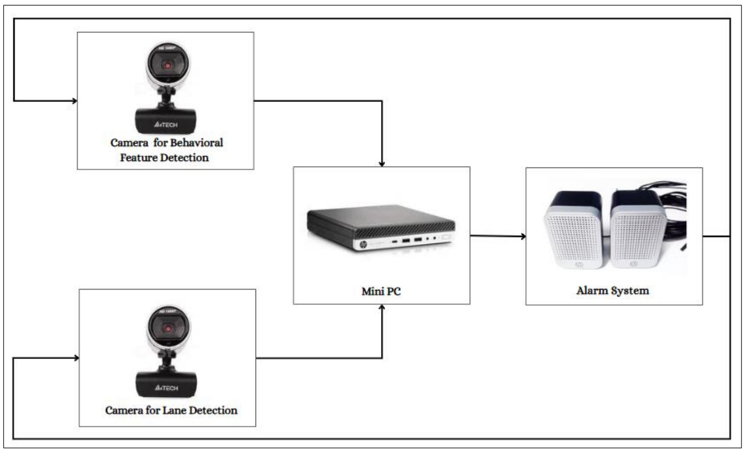

Figure 3 Block Diagram for Drowsiness Detection System

The whole process for prototype development consists of local server consists of the mini-PC that houses the Python program that communicates with the behavioral feature detection and the alarm system. The Behavioral Feature Detection system utilizes an A4Tech PK-910 HD webcam connected to the HP ProDesk 400 G5 Mini Desktop to capture images and video frames for daylight and night vision. The system leverages the HP ProDesk 400 G5 Mini Desktop's computational capabilities to process the captured lane data in real-time. The alarm system is a speaker that will communicate with the server once drowsiness and distraction are detected.

2.2.2. System Testing

To validate the real-time driver drowsiness and distraction detection system, rigorous testing will be conducted on a carefully selected road section. This allows controlling variables like traffic, road type, and weather conditions. Testing will occur at various times to assess performance under different lighting and noise conditions. The system uses computer vision models trained on labeled datasets to detect drowsiness, distraction, and lane departure. The testing has two main parts: algorithm testing and real-time on-road testing with participants who are regular long-distance drivers. For algorithm testing, driving data simulating all possible conditions is collected. The algorithms are run on this data, and performance metrics like accuracy, false positives/negatives, and response time are analyzed. For real-time testing, participants drive on the road while observers monitor the system's output and record detected instances of drowsiness and distraction. Convenience sampling is used for real-time testing by recruiting available volunteers with varying driving experiences. The testing concludes with a report evaluating the system's performance and recommendations for improvement. This comprehensive testing approach under simulated and real-world conditions ensures the system's effectiveness before deployment.

2.2.3. Performance Metrics

The observed results of the system testing will be classified into four interpretations: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FP). These then will be used to evaluate the performance of the driver monitoring system to calculate the different performance metrics: accuracy, precision, and recall. These metrics can be used to assess the system's overall performance and identify areas where it can be improved.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

Accuracy is the most widely used metric for evaluating the performance of classification models. It is simply the proportion of correct predictions made by the model. Precision is the proportion of positive predictions that are correct. It is the number of times the model correctly predicts that an instance belongs to the positive class divided by the total number of times the model predicts that an instance belongs to the positive class. Recall is the proportion of positive cases correctly identified. It is the number of times the model correctly predicts that an instance belongs to the positive class divided by the total number of instances that belong to the positive class.

3. Results and discussion

3.1. Results for Objective 1: Drowsiness Detection System

In this study, the drivers were asked to act the two drowsy states "drowsy" and "non drowsy" to provide data for the accuracy of the system. This involves participants intentionally showing drowsiness, such as by closing their eyes partially or completely and opening their mouth. The result of these controlled scenarios will contribute to improving the system's ability to differentiate drowsy and non-drowsy in real-time situations

Table 1 Drowsiness Detection System Test Results

Test Scenario	True Positive	False Positive	True Negative	False Negative
Drowsy: closed/drowsy eyes, yawning	105	-	-	20
Non-drowsy: eyes open, mouth closed	-	25	100	-

Table 1 shows the performance of the drowsiness detection system predicting two states of driver drowsiness: drowsy state (closed/drowsy eyes, yawning) and non-drowsy state (eyes open, mouth closed). The system can identify drowsiness properly, resulting in the system detecting 105 correct drowsy predictions and 100 non-drowsy predictions. However, there was a notable amount of misclassification (25 false positives and 20 false negatives), which suggests that the system can miscalculate in detecting driver drowsiness.

Table 2 Drowsiness Detection System Performance

Accuracy	Precision	Recall
0.82	0.81	0.84

Table 2 displays the accuracy, precision and recall for a drowsiness detection system that employs Mouth Aspect Ratio and Eye Aspect Ratio models. The accuracy of this system is 82%, which implies that the classification was correct for 82% of instances it evaluated as either drowsy or not drowsy. The precision of the system is at 0.81, meaning that out of all the predicted positives cases by the system, 81% were positive cases of drowsiness. The recall for this system stands at .84. Overall, these metrics suggest that EAR- and MAR-based drowsiness detection perform well with acceptable levels of accuracy, precision and recall.

3.2. Results for Objective 2: Distraction Detection System

In this study, the performance of the distraction detection system was evaluated by testing the system regarding distraction detection and eye gaze analysis. The drivers are asked to simulate the two distraction states, which are "distracted" and "alert". The distracted state is characterized by the driver gazing off the forward roadway, looking up, at the console/dashboard, and side mirrors. The alert state is characterized by the driver looking at the road ahead, eyes on the forward roadway.

Table 3 Distraction Detection System Test Results

Test Scenario	True Positive	False Positive	True Negative	False Negative
Distracted: eyes off the forward roadway	109	-	-	16
Alert: eyes on the forward roadway	-	5	120	-

Table 3 illustrates the prediction performance of the distraction detection system under two states of driver distraction: distracted state (eyes off the forward roadway) and alert state (eyes on the forward roadway). As the table shows, the distraction detection system produced accurate predictions of 109 true positives and 120 true negatives and yielded prediction errors of 5 false positives and 16 false negatives.

Table 4 Distraction Detection System Performance

Accuracy	Precision	Recall
0.91	0.95	0.87

The distraction detection system performs well in correctly detecting driver distraction instances, as depicted in Table 4. This table reports an accuracy of 0.91, meaning that 91% of the cases are correctly classified as distracted or not distracted driving scenarios. The high accuracy indicates that the system is reliable enough to detect the driver's distraction state. Furthermore, the system has a precision of 0.95, meaning that for all situations flagged as distracted driving, it is correct 95% of the time. That great precision value ascertains that alerts or interventions triggered by this system are primarily justified, hence avoiding unnecessary distraction or false alarms. Meanwhile, the system exhibits a recall of 0.87, which indicates an acceptable rate of correctly identifying a driver being distracted.

3.3. Results for Objective 3: Lane Departure Detection System

The angle vs time graph of the daytime vehicle lane detection testing on Figure 8 analyzes the vehicle's steering angle based on the position and curvature of the detected lane lines relative to the picture center over time. By analyzing the patterns of the plots of the angle values over time, one can see the responsiveness of the lane detection algorithm. At the start, the pattern can be seen as a straight plot which means that the vehicle is still centered. After 3 seconds, the graph remains consistent at the positive level which tells that the vehicle is constantly stays in that angle to the right, while the continuous up and down slope indicates that the algorithm is detecting that the vehicle is turning from left to right and vice versa as the angle increases and decrease consistently. Sudden spikes in the graph such as in 3.1 seconds, 5.3 seconds, and 17.8 seconds which may indicate that the vehicle suddenly turned or there may be errors such as shadows during daytime that interfere with the detection of the algorithm.

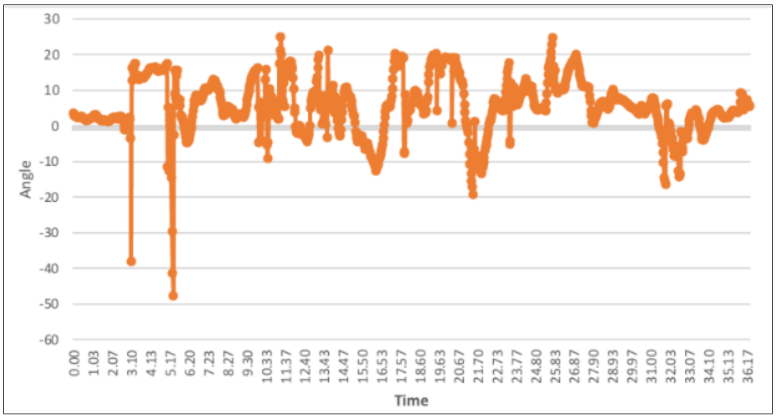

Figure 4 Angle VS Time Graph of the Lane Detection During Daytime

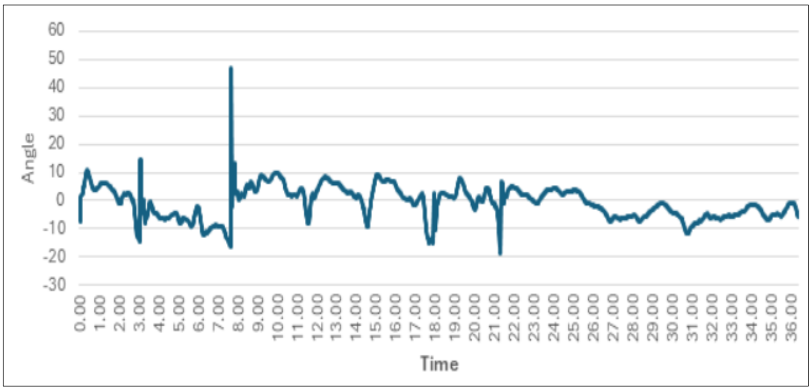

Figure 5 Angle VS Time Graph of the Lane Detection During Nighttime

Figure 9 shows the measurements of the vehicle's steering angle as the time changes during the conducted experiment at nighttime. By analyzing the patterns of the plots of the angle values over time, one can see the responsiveness of the lane detection algorithm during nighttime. However, during nighttime, several factors affect the lane detection algorithm such as the lighting from the bumpers of the front cars and the lighting of the streetlights which can result to sudden spikes such as seen in the 7.67 second mark. Sudden spikes like this may also indicate sudden changes in lanes by the vehicle such as those recorded at the 21 second mark. Overall, the plotting shows that the lane detection algorithm consistently detects the vehicle lane as most of the angle measured gradually increases and decreases creating a slope in the graph stating that the vehicle is detected as turning left or turning right continuously.

3.4. Results for Objective 4: Safety Alert System

The system monitors the driver's behavioral features such as Eye Aspect Ratio (EAR), Mouth Aspect Ratio (MAR) and Eye Gaze. Also, the lane angle is an additional for more accurate results. By providing timely audio cues and warnings, the system will enable drivers to acknowledge the alarm, reducing the risk of accidents caused by drowsiness and distraction while driving.

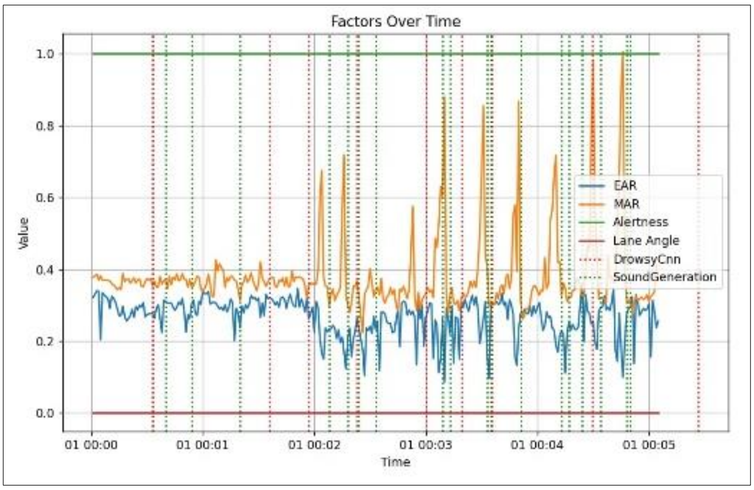

Figure 6 Real-Time Driver Drowsiness and Distraction Detection

The visual representation is shown on how the system worked during the real-time testing of a participant in real-life conditions of driving. In this graph, it shows the trend on where the system has detected the driver is drowsy, distracted, or not and its effectiveness to alert the driver during these moments. The sudden high spikes of the MAR and the low spikes of the EAR in the period of 01 00:02 until 01 00:05 indicates that during this timeframe, the second participant has been detected as drowsy as it reached the threshold indicated in the system. Consequently, after these spikes in the MAR and EAR metrics happened, the graph shows that the system has generated a sound to alert the second participant right after as the sound generation can only be triggered after the two second period of drowsiness or distraction, which means that the system has effectively functioned during times of driver drowsiness or distraction.

Table 5 Safety Alert System Test Results and Performance

Test Scenario	Tally
Correctly activated alarms	214
False activated alarms	36
Overall System Accuracy	85.60 %

The table presents the result of the accuracy test for the real-time driver drowsiness and distraction detection system. The system correctly identified and triggered the alarm in 214 instances that means that the system accurately detects drowsiness and distraction in the driver. However, the system also falsely activated the alarms for 36 instances, meaning it incorrectly detected drowsiness and distraction and triggered the alarm even though the driver was not drowsy or not distracted. The overall accuracy of the system resulted to 85.6%, which implies a good performance on the alerting capability of the proposed system.

4. Conclusion

A real-time driver drowsiness and distraction detection system is successfully developed in this study, effective in detecting instances of driver drowsiness and distraction using CNN with multiple behavioral features as inputs. Improving the hardware components of the system in terms of size and performance, using of additional features or combining the other features of detection, increasing the dataset, testing the device in various settings, and including more participants are all possible areas for future research initiatives.

Compliance with ethical standards

Acknowledgements

The researchers would like to extend their gratitude to their advisers, Engr. Eufemia A. Garcia and Engr. Antonio C. Tee, for the never-ending guidance and valuable insights throughout the thesis journey. Moreover, the authors also like to thank the panel board, the chair Dr. Carlos C. Sison and members Engr. Fernando Victor V. De Vera and Prof. Mark

Christopher R. Blanco, who also gave the researchers inspiration and guidance in this research endeavor through scrutiny and engaging discourse.

Disclosure of Conflict of interest

Authors have all agreed that no conflict of interest exists in the creation of the study.

Statement of informed consent

Informed consent was obtained from each participant involved in the study through a confidentiality agreement, establishing protocols on confidentiality, recording consent, and non-disclosure of identifiable information.

References

[1] Kirley BB, Robison KL, Goodwin AH, Harmon KJ. O'Brien NP, West A, Harrell SS, Thomas L, Brookshire K. Countermeasures that work: A highway safety countermeasure guide for State Highway Safety Offices. 11th ed. Washington: National Highway Traffic Safety Administration; 2023.

[2] Perkins EM, Situala C, Burke M, Marzbanrad F. Challenges of Driver Drowsiness Prediction: The Remaining Steps to Implementation. IEEE Transactions on Intelligent Vehicles. 2021 Sept; 8(1): 1319-38.

[3] Ju JH, Park YJ, Park J, Lee BG, Lee J, Lee JY. Real-Time Driver's Biological Signal Monitoring System. Sensors and Materials. 2015;27(1):51-59.

[4] Alioua N, Aouatif A, Rogozan A, Bensrhair A, Rziza M. Driver head pose estimation using efficient descriptor fusion. EURASIP Journal on Image and Video Processing. 2016; 2016(1):1-14.

[5] Altameem A, Kumar A, Poonia RC, Kumar S, Saudagar AKJ. Early Identification and Detection of Driver Drowsiness by Hybrid Machine Learning. IEEE Access. 2021; 9(1):162805-19.

[6] Haggag O. Automated Drowsiness Detection while Driving using Depth Camera. Proceedings of Umeå's 23rd Student Conference in Computing Science. 2019; 2019(1): 11-22.