Available online at www.HighTechJournal.org

HighTech and Innovation Journal

ISSN: 2723-9535

Vol. 7, No. 1, March, 2026

Towards Safer Roads: A Machine Learning Framework for Driver Fatigue Detection

Deshinta Arrova Dewi [1, 2*], Tri Basuki Kurniawan [3], Misinem [4], Mohd Zaki Zakaria [5], Shahreen Kasim [6]

[1] Faculty of Data Science and Information Technology, INTI International University, Nilai 1800, Negeri Sembilan, Malaysia.

[2] Faculty of Engineering Technology, Shinawatra University, Pathum Thani, Thailand.

[3] Magister of Information Technology, Postgraduate Program, Universitas Bina Darma, Palembang, Indonesia.

[4] Faculty of Vocational, Universitas Bina Darma, Palembang, Indonesia.

[5] Faculty of Computer & Mathematics Sciences, University Technology Mara, Selangor, Malaysia.

[6] Faculty of Computer Sciences and Information Technology, Universiti Tun Hussain Onn Malaysia, Parit Raja 86400, Johor, Malaysia.

Received 07 July 2025; Revised 16 February 2026; Accepted 23 February 2026; Published 01 March 2026

Abstract

The goal of this research is to enhance road safety by delivering an affordable, non-intrusive, and effective drowsiness detection solution suitable for integration into mainstream vehicle systems. Driver drowsiness is a critical factor contributing to road accidents, often resulting in severe injury or death. In Malaysia alone, it is estimated that drowsiness causes between 2,000 and 3,000 traffic accidents annually. Conventional vehicle-based drowsiness detection systems, which rely on steering behavior or lane departure, often fail to detect early physiological signs of fatigue. Recent advancements in artificial intelligence (AI), particularly in computer vision and machine learning, offer new opportunities for developing low-cost, real-time drowsiness monitoring systems. This research proposes a driver drowsiness monitoring system (DDMS) that utilizes deep learning and visual behavior analysis to detect signs of fatigue through real-time monitoring of eye and mouth activity. A customized Convolutional Neural Network (CNN) is developed to classify eye states (open vs. closed), while yawning is detected using facial landmark analysis and Mouth Aspect Ratio (MAR) computations. The system is trained and evaluated using the MRL Eye Dataset, which consists of 4,000 annotated images, with data preprocessing and augmentation applied to enhance robustness. Through systematic experimentation and hyperparameter tuning, the model achieves a peak accuracy of 98%, with equally high precision, recall, and F1 Scores, using the Adam optimizer, a learning rate of 0.001, and 50 training epochs. The system also demonstrated strong real-time performance across varied lighting conditions, though challenges remain in scenarios involving occlusion (e.g., sunglasses) and extreme head positions. The results indicate that the proposed method is not only feasible but highly accurate, marking a significant advancement in proactive traffic accident prevention technologies.

Keywords: Driver Drowsiness Detection; Convolutional Neural Network (CNN); Facial Landmark Analysis; Computer Vision; Time Monitoring System; Public Health; Process Innovation.

1. Introduction

Road traffic accidents are still one of the major causes of death and injury across the globe. According to the WHO, nearly 1.19 million people die each year from road traffic accidents, and a considerable percentage of these accidents involve a fatigued driver [1]. Fatigue can adversely affect a person's cognitive functions, slow reaction times and

* Corresponding author: deshinta.ad@newinti.edu.my

https://doi.org/10.28991/HIJ-2026-07-01-018

reducing situational awareness, thereby increasing the risk of accidents [2]. Given these human and economic costs, the detection of driver fatigue has become a critical concern for the intelligent transportation systems industry. Within the last decade, a wide range of driver fatigue detection systems has been developed, and can be mostly divided into three categories: (I) Physiological signal-based methods. (ii) Behavior-based methods. and (iii) Vehicle Dynamics Based methods. Physiological methods use EEG, EOG, and HRV monitoring to study changes in brain waves and autonomic nervous system responses [3]. Computer vision systems and facial landmark tracking have made it possible to monitor (PERCLO) eye closure rates, yawning rate, and head nodding, which are behavioral methods.

Despite the encouraging outcomes of various approaches, each has drawbacks. Intrusive sensors are often required for physiological signal collection, which can make drivers uncomfortable and hinder their adoption in practical situations. Despite being non-intrusive, behavioral techniques may be sensitive to individual variances, facial occlusions, and changes in illumination. Road conditions, driving style, and vehicle type can all have an impact on vehicle dynamics data, which could lower reliability in a variety of settings [4]. To improve robustness, recent research has started to investigate multimodal techniques that integrate two or more data sources. Nonetheless, three major obstacles still confront many current studies:

- Limited generalizability across drivers and environments: Because demographics, lighting, weather, and road types vary in real-world driving situations, models trained on controlled datasets sometimes perform poorly [5].
- Most datasets are skewed towards either alert or extremely tired states, ignoring the subtle transitional phases where early intervention is feasible. This results in imbalanced datasets and inadequate representation of early fatigue states [6].
- Absence of real-time, lightweight deployment frameworks. The computational cost of many high-performing models retracts their applicability for incorporation into embedded systems within vehicles [4, 7].

To fill these gaps, this study suggests a machine learning architecture for detecting driver weariness that combines contextual driving data, non-intrusive physiological indications, and behavioral cues from computer vision. The proposed system leverages a hybrid feature extraction pipeline combining deep learning–based facial analysis with time-series feature engineering for sensor data. A lightweight ensemble learning model is then deployed to achieve high accuracy while maintaining low computational overhead, enabling real-time detection on embedded hardware.

The contributions of this study are as follows:

- Multimodal, non-intrusive fatigue detection that fuses visual, physiological, and contextual features to improve robustness in real-world conditions.
- Early fatigue state recognition through tailored data augmentation and class balancing techniques, enabling proactive driver alerts.
- Lightweight model architecture optimized for in-vehicle deployment without compromising detection accuracy.

By addressing the limitations of prior work, this framework aims to enhance the reliability, accessibility, and real-world applicability of driver fatigue detection systems, contributing to the broader goal of safer roads and reduced accident rates. As we know that the difficulty in detecting drowsiness poses a significant challenge, as even a few seconds of microsleep—a brief period of unintentional sleep—can have disastrous consequences [4]. Current methods for detecting and alerting drivers of drowsiness have limitations, often relying on vehicle behavior metrics such as lane departure or steering patterns. These methods don't consider the subtle facial expressions and eye movements that show someone is about to drive off the road [5]. The use of these safety systems, which detect drowsiness, is not widespread and is unusual among drivers, as they are often found in expensive automobiles [6]. An AI-powered driver drowsiness detection system offers a promising solution to these shortcomings. A system like this can intelligently analyze facial features and extract meaningful patterns that effectively indicate drowsiness by utilizing machine learning and computer vision techniques. This leads me to develop a machine learning-based driver drowsiness monitoring system to reduce and prevent the frequency of accidents.

Altogether, this project aims to develop an image processing system that detects and alerts drivers who are drowsy, thereby enhancing road safety. The system will focus on identifying signs of drowsiness in the eyes and mouth of drivers with chronic sleepiness, such as those who consistently struggle with sleepiness or have a history of drowsy driving [7]. By utilizing various image processing techniques and machine learning models, the system will detect drowsiness and provide timely warnings to drivers, potentially reducing the number of drowsy driving-related accidents. The significance of this project is to ensure that the road can be used safely for all individuals. By developing an AI-based driver drowsiness monitoring system, it will also help drivers to stay safe and avoid accidents that could be fatal. In addition to reducing the number of accidents caused by drowsy driving, this system can also help to improve the overall safety of our roads.

The constant evolution of technology, particularly in the field of artificial intelligence (AI), has created opportunities for groundbreaking applications across various industries. This project explores the domain of artificial intelligence to address a pressing problem: drowsy driving. Given that drowsy driving is a significant contributor to traffic accidents, an AI-powered driver drowsiness detection system is proposed. Leveraging advanced computer vision algorithms, the system analyses facial features, focusing on subtle signs of drowsiness such as eye and mouth movements.

This paper is organized in the following sequence. Section 2 focuses on the fatigue detection from a neuroscience and psychophysiological perspective, with a detailed explanation of visual behavior, facial landmark detection, mouth aspect ratio, and an integrated solution for drowsiness detection. Methodology is discussed in section 3, followed by results discussion in section 4. This paper is concluded at the end of the chapter.

2. Fatigue Detection

Drowsiness represents a transitional state between wakefulness and sleep, characterized by a significant decline in alertness, cognitive performance, and psychomotor functioning [8]. It is a multifactorial condition often precipitated by sleep deprivation, circadian rhythm misalignment, extended periods of monotonous activity, or medical conditions such as sleep apnea and narcolepsy.

On the other hand, fatigue is a more comprehensive concept that includes both mental and physical depletion, and it can happen with or without the body's natural need to sleep. While fatigue results from prolonged cognitive workload, extended driving hours, environmental stressors, or repetitive low-stimulation tasks, it usually takes longer to recover from than drowsiness, which is naturally linked to sleep propensity and can frequently be relieved by short-term rest or napping. Crucially, a driver might be fatigued without visibly being sleepy and still show signs of poor judgment, sluggish reaction times, and a lack of situational awareness. Because drowsiness-focused algorithms may miss non-sleep-related weariness, making them less successful at preventing accidents, distinguishing between these two states is crucial for the design of intelligent detection systems.

According to neuroscientific research, fatigue affects more brain areas than are usually linked to sleep propensity. These include the parietal, basal ganglia, and anterior cingulate cortex, which are in charge of sustained attention, motor coordination, and cognitive control. Prolonged mental and physical exertion reduces activation in attention-regulating networks and disrupts the connection between the frontoparietal network and motor planning areas, according to functional neuroimaging studies using fMRI and EEG. Even when there are no obvious signs of tiredness, these brain changes lead to decreased alertness, delayed decision-making, and less accurate motor performance.

From a psychophysiological perspective, fatigue can be detected from a combination of cardiovascular, muscular, and behavioral indications, including decreased heart rate variability (HRV), elevated electromyographic (EMG) muscle tension, and progressive deterioration in steering stability or pedal control. Multimodal sensor integration combining wearable physiological sensors, driver engagement patterns, and vehicle telemetry helps detect weariness, in contrast to drowsiness, which is frequently detected by ocular metrics alone.

Driver fatigue continues to be a key cause of road traffic accidents worldwide, decreasing cognitive function, slowing reaction times, and diminishing situational awareness. Over the last decade, various detection methods have been developed, ranging from physiological sensing using EEG, ECG, or heart rate variability (HRV) [1-3], to vision-based monitoring of ocular and facial indicators such as PERCLOS, blink duration, and yawning frequency [4-6], and vehicle dynamics analysis including steering wheel movement, lane deviation, and braking patterns [7-8]. While these approaches have shown encouraging outcomes, they have considerable practical limits.

First, relying too heavily on single-modality systems diminishes robustness in real-world circumstances. Physiological approaches, while correct, frequently involve intrusive sensors, which create discomfort and limit consumer adoption [9]. Poor illumination, facial occlusions, and driver accessories such as glasses and masks all have an impact on vision-based techniques. Vehicle telemetry-based detection might be skewed by road conditions, weather, and driving style, making it difficult to assign anomalies to weariness alone [10].

Second, most previous research has been significantly skewed toward drowsiness detection, focusing on sleep-related weariness while ignoring broader non-sleep-related fatigue states generated by prolonged cognitive exertion, monotonous driving, or environmental stresses [11]. This narrow scope limits the ability to detect early fatigue stages, where action could help avert accidents.

To address this, research has focused on the integration of multimodal tiredness detection systems that combine behavioral markers (e.g., face analysis), physiological data (e.g., heart rate variability, EEG), and contextual characteristics (e.g., driving time, time of day) [12]. These technologies, which are frequently embedded in Advanced Driver Assistance Systems (ADAS) or autonomous vehicle platforms, seek to provide robust, adaptive interventions that can either encourage corrective behavior or transition control to automated driving modes when drowsiness is identified [13]. Despite this integration, three persistent challenges remain: (i) limited generalizability due to small, homogeneous datasets and a lack of diverse environmental testing, (ii) under-representation of subtle, early-stage fatigue in datasets, resulting in poor model sensitivity, and (iii) high computational demands of deep learning architectures, which limit real-time deployment in resource-constrained in-vehicle systems.

These shortcomings demand a new approach that provides accurate, multimodal fatigue detection for both drowsy and non-drowsy states, is computationally efficient, and can work dependably under a wide range of real-world settings. The current study addresses this need by offering a lightweight, hybrid machine learning framework that combines behavioral, physiological, and contextual variables into a single architecture, allowing for real-time detection and proactive intervention in ordinary driving scenarios.

While fatigue spans a wide range of physical and mental depletion, this study identifies sleepiness as the predominant expression of fatigue that may be observed in real time [14]. Drowsiness, as a transitional condition between wakefulness and sleep, is frequently the most visible and measurable phase of fatigue, making it an ideal starting point for detection in driver monitoring systems.

This study attempts to develop a reliable baseline for recognizing fatigue states that are significantly associated with sleep propensity by first focusing on drowsiness-related markers such as eye closure rate, blink duration, and head position. Although the long-term goal is to broaden detection capabilities to include non-sleep-related exhaustion, the current framework prioritizes sleepiness due to its more obvious physiological markers, known measuring methodologies, and higher prevalence in documented accident reports.

2.1. Visual Behavior, Drowsiness and Fatigue Detection

Visual behavior encompasses the observable eye and gaze patterns that reflect how individuals process and react to visual stimuli. These behaviors range from basic reflexive responses, such as pupil dilation and saccadic movements, to more sophisticated cognitive actions, including sustained visual attention, environmental scanning, and anticipatory gaze [15]. In the context of driver state monitoring, ocular behavior is an important biometric indication since it immediately indicates a driver's attentiveness, cognitive engagement, and susceptibility to drowsiness and fatigue [16, 17].

Drowsiness, as a state between waking and sleep, is frequently accompanied by specific visual changes such as longer eye closure, reduced blink rate, and impaired gaze stability. These symptoms are typically more prominent and quantifiable than those associated with non-sleep-related fatigue, which may involve a more modest decline in visual-motor coordination over time. Nonetheless, weariness, especially when caused by extended cognitive or physical exertion, can emerge as visual behavior, such as greater gaze dispersion, reduced saccadic velocity, and abnormal fixation patterns, even in the absence of severe sleep demand.

Eye-gaze behavior is an important measure in evaluating visual behavior, notably in the context of the Quiet Eye notion, which refers to the final visual fixation preceding a motor response. Originally investigated in sports psychology, this notion has application in driving, where continuous and focused gaze patterns are related to attentive driving behavior, but irregular or protracted eye closures frequently indicate fatigue accumulation or microsleep occurrences [16]. Eye-tracking devices use metrics including blink rate, gaze dispersion, saccadic velocity, and eye closure duration to predict the beginning of drowsiness [17], whilst long-term monitoring of fixation stability and saccadic rhythm can provide insights into fatigue-related cognitive slowdown.

In locomotor learning, gaze behavior aids in quantifying how visual inputs influence motor decisions such as steering and braking. Similar ideas apply to driver monitoring systems: decreased saccadic frequency, fixation instability, and greater gaze wandering can indicate a breakdown in visual-motor coordination. This decline is typical of early-stage drowsiness, but it could also suggest cumulative fatigue from long driving times or monotonous road conditions [16, 18]. Visual behavior analysis, which prioritizes drowsiness detection while incorporating fatigue-sensitive measures, serves as a core component for constructing powerful, multimodal driver status monitoring systems capable of managing both immediate sleep danger and long-term performance deterioration [19].

2.2. Facial Landmark Detection for Drowsiness and Fatigue Monitoring

Facial landmark detection is a fundamental approach in the development of advanced driver monitoring systems, particularly for detecting indicators of drowsiness and, to a lesser degree, fatigue. It entails identifying key facial points such as the corners of the eyes, mouth, nose, and jawline that define the geometric structure of the face and allow for examination of minor facial dynamics [20, 21]. The introduction of deep learning-based convolutional neural networks (CNNs) has significantly increased the accuracy and resilience of landmark detection, especially under difficult settings such as head tilts, fluctuating lighting, partial occlusions, or facial accessories like spectacles.

Modern techniques frequently use cascaded CNN architectures, in which many network layers iteratively refine landmark predictions. Cascaded frameworks with over 20 CNN modules, for example, can accurately localize more than 60 facial landmarks, allowing for the extraction of dynamic variables such as the Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR), which are extensively employed as indicators of tiredness and exhaustion [22]. While EAR primarily measures eyelid closure patterns and MAR measures yawning behavior, both measurements can also detect slower, fatigue-related alterations, such as decreased blink frequency, modest facial muscle sagging, or diminished expressiveness, even in the absence of significant sleep desire [23].

During drowsiness, landmark-based assessments frequently reveal noticeable changes in the relative positions and distances surrounding the eyes and mouth, such as prolonged eyelid closure, frequent yawning, and decreased muscular tone [24]. Fatigue, on the other hand, may appear gradually, with symptoms including prolonged micro-expressions of strain, slower recovery of eyelid position following blinks, and decreased facial response to external stimuli. By tracking these landmarks in real time, it is feasible to monitor both the acute onset of drowsiness and the gradual accumulation of weariness.

2.3. Mouth Aspect Ratio (MAR) as a Drowsiness and Fatigue Indicator

The Mouth Aspect Ratio (MAR) is a geometric measure that describes the vertical and horizontal displacement of the mouth. It is computed by dividing the perpendicular distance between the upper and lower lips (measured through midpoints) by the horizontal distance between the mouth corners. This ratio accurately depicts the opening and closing motion of the mouth, which occurs frequently during yawning, a common and visible indicator of tiredness and fatigue. Studies have shown that MAR can be an effective proxy for identifying yawns, which are involuntary reflexes that are frequently prompted by tiredness. Yawns, unlike basic facial expressions, are distinguished by increased mouth openings and prolonged durations, which can be detected with continuous MAR monitoring. A sustained increase in MAR above a specific temporal threshold is commonly interpreted as a yawn occurrence [25].

From a fatigue standpoint, yawning may not necessarily be associated with impending sleep, but it can also result from lengthy mental or physical exertion, boring driving, or diminished stimulation, conditions that degrade performance even in the absence of a strong sleep inclination. In such circumstances, MAR can help identify fatigue-related yawning patterns that occur intermittently over long periods of time rather than in quick succession, as is common in acute sleepiness [26]. Studies have demonstrated that MAR can be a robust proxy for detecting yawns, which are involuntary reflexes often triggered by drowsiness. Unlike simple facial expressions, yawns involve exaggerated mouth openings and prolonged durations, making them distinguishable through continuous MAR monitoring. A sustained increase in MAR over a certain temporal threshold is typically interpreted as a yawn event.

When paired with other indications like Eye Aspect Ratio (EAR), head nodding frequency, and gaze fixation stability, MAR becomes an important part of multimodal drowsiness and fatigue detection frameworks. This combined method improves both sensitivity and specificity, allowing systems to distinguish between fleeting facial expressions and true exhaustion events, whether sleep-related or caused by a prolonged cognitive or physical stress [27].

2.4. Integrated Solution for Drowsiness and Fatigue Detection

A reliable driver state monitoring system combines visual behavior analysis, facial landmark tracking, and aspect ratio calculations into a single, continuous pipeline capable of identifying drowsiness and, to a lesser extent, fatigue. While drowsiness remains the primary emphasis because of its strong and observable physiological markers, using the same framework to recognize fatigue-related symptoms allows for a broader coverage of performance-degrading states that do not manifest immediately after sleep onset.

A typical workflow contains

- Video Input Acquisition: Captured using a near-infrared or RGB camera installed in the car interior to ensure day-night operation.
- Face and Landmark Detection: Using CNN-based detectors, the driver's face is identified and landmark points are computed, regardless of lighting, head pose, or partial occlusion.
- Feature Extraction:
  - EAR for prolonged eye closure and decreased blink rate (drowsiness).
  - MAR for detecting yawns, including those caused by sleep and tiredness.
  - Gaze and head pose metrics are used to estimate attention and detect impaired visual-motor coordination, which is frequent during fatigue buildup.
- State classification involves using threshold-based or machine learning models (e.g., SVM, LSTM, or deep CNNs) to identify alert, drowsy, or fatigued states. This stage can be improved with multimodal fusion by including physiological data (e.g., HRV, EEG) or environmental information (e.g., driving time, time of day).
- When a risk condition is detected, adaptive actions like auditory alarms, steering wheel vibrations, or, in advanced driver assistance systems (ADAS), the change from partial control to automated driving modes is triggered.

3. Research Methodology

3.1. Data Acquisition

It plays a crucial role in this project. Obtaining the correct data is essential for the success of any project, which means gathering reliable and publicly available datasets relevant to the specific aims. Choosing a valuable and good dataset forms the foundation for a strong project. This dataset was retrieved from the Kaggle website (https://www.kaggle.com/datasets/prasadvpatil/mrl-dataset), which contains a total of 4,000 images. These images are evenly split, with 2,000 showing eyes in an open state and 2,000 showing eyes closed. This dataset is highly relevant, as it includes diverse photos that depict various real-world scenarios of driver eye states. Before using the dataset for training machine learning models, it will be carefully processed. By carefully collecting and pre-processing the data, we ensure that the machine learning model has a strong foundation for accurately detecting driver drowsiness based on eye state images.

3.2. Data Pre-Processing

Data cleaning can increase accuracy by removing errors and inconsistencies. Filtering helps us identify important clues, such as eye movements, that indicate drowsiness. To enhance the diversity and robustness of the dataset, various pre-processing techniques, including data splitting, data augmentation, and image resizing, are applied. The system can then learn various patterns by dividing the refined dataset into training and testing sets. Data augmentation is a critical technique that artificially expands the dataset by applying various transformations to existing data samples. These transformations include rotating images, shifting images horizontally or vertically, and flipping images horizontally or vertically [28].

3.3. Model Design

To fully develop the DDMS, the design of the system itself is crucial for maintaining a smooth process throughout the project's development. Experiments with the Custom CNN model were conducted to determine the optimal parameter values for the model. The CNN model used in this research project serves as a fundamental component for detecting driver drowsiness through the classification of eye states. CNNs use convolutional layers to extract significant characteristics from images, which makes them particularly well-suited for learning patterns in data. For this research, the CNN architecture was customized to handle the task of eye state classification. Deeper layers were added to extract higher-level features essential for differentiating between open and closed eyes, starting with layers that extract low-level features such as edges and textures.

3.4. Model Development

The following section outlines the step-by-step coding process for model development. This includes importing libraries, loading and pre-processing the dataset, constructing the model architecture, and finally training the model using the prepared dataset.

3.5. Model Evaluation

The evaluation of the driver drowsiness monitoring system's model involved several key steps to ensure its effectiveness and reliability. The dataset was initially divided into training, validation, and test sets using a stratified technique to ensure balanced class distributions. The model was trained using a custom convolutional neural network (CNN) architecture explicitly designed for image classification tasks, such as distinguishing between closed and open eyes. During training, data augmentation techniques such as rotation, zooming, and horizontal flipping were applied to enhance model accuracy. For the evaluation part, the trained model was tested on a separate test set to measure its performance metrics. This includes measuring the accuracy, precision, recall, and F1 score.

3.6. System Development

The developed driver drowsiness detection model was implemented into a functional system to validate its performance and usability. The primary focus was on ensuring the system could process live video feeds and accurately classify drowsiness states in real-time using the embedded model as the classifier engine. The system utilized advanced computer vision techniques, integrating a Convolutional Neural Network (CNN) for eye state analysis and facial landmark detection for yawning detection. The software implementation was carried out using Python, leveraging libraries such as OpenCV and TensorFlow for real-time video processing and model inference, as shown in Figure 1.

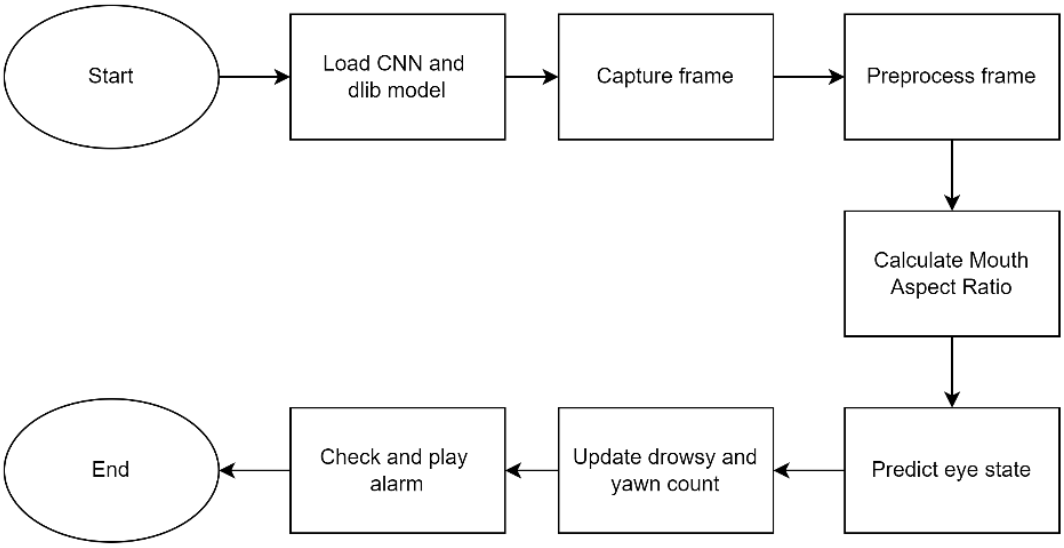

Figure 1. DDMS Flowchart

Based on Figure 1, the drowsiness detection system follows a structured workflow to monitor the driver's state and provide timely alerts effectively. The process begins by loading the constructed CNN and dlib facial landmark detection models, which are essential for extracting the region of interest. The system then captures real-time video frames from the camera, preprocesses them to enhance quality, and calculates the Mouth Aspect Ratio (MAR) for yawning detection.

Next, the CNN model is used to predict the driver's eye state, and the system updates the drowsy and yawn counts based on the predictions. Finally, the system checks these counts against predefined thresholds and plays the appropriate audio alarms if the driver is deemed drowsy or yawning, requiring them to take urgent action. This continuous cycle of data capture, processing, and alert generation ensures the system remains active and contributes to enhanced road safety by reducing the risks associated with drowsy driving.

3.7. Testing Functionalities and Accuracy

The objective of the testing phase is to evaluate the functionality and accuracy of the driver drowsiness monitoring system, ensuring the system performs as intended under various conditions. First, functional testing ensures that all system components, such as eye state detection and yawning detection using facial landmarks, work properly and interact effectively. This involves testing scenarios that mimic real-world driving conditions to assess how well the system performs in detecting drowsiness cues, such as closed eyes and yawning gestures, reliably and promptly.

4. Results and Discussion

4.1. Parameter Setting

Parameter settings in a machine learning model are the specific values assigned to its parameters during training. These settings play a crucial role in determining how the model learns from data and makes predictions, as shown in Table 1.

Table 1. Parameter setting

Parameter Setting	
Parameter	Value / Setting
Learning Rate	0.001
Epoch	15
Optimizer	Adam
Batch Size	32
Training Ratio	7:3
Input Shape	145 × 145
Data Augmentation	Enabled

These initial parameter settings, shown in Table 1, establish the basic configuration for training the machine learning model. These settings define the conditions under which the model was learned and is tested. This setup provides a starting point for further experiments aimed at improving the model's performance and accuracy in predicting outcomes.

The learning rate of 0.001 was chosen because it offers a fair trade-off between convergence speed and model stability. A greater value may result in overshooting the optimal weights, but a significantly lower value may impede the learning process and increase the danger of underfitting. The model was trained over 15 epochs, a value set by preliminary experiments to ensure enough training while minimizing the risk of overfitting. Adam was chosen as the optimizer because of its flexible learning rate capacity and previous competence in image-based deep learning problems.

A batch size of 32 enables effective gradient computing while remaining computationally feasible on existing technology. The 70:30 training ratio ensures that the majority of the data is used for model learning while leaving a significant test set for performance evaluation. The 145 × 145 pixel input form balances visual detail and computational load, making it ideal for real-time or embedded deployments. Data augmentation techniques like as rotation, scaling, flipping, and brightness change were used to artificially expand the training dataset, improve generalization, and increase robustness to real-world perturbations.

These initial parameter values serve as the foundation for training the machine learning model. They specify the settings under which the model was created, tested, and evaluated for this study. By establishing a controlled setup, this configuration allows for systematic experimentation with alternative values in future work—such as varying the learning rate, increasing epochs, or experimenting with different optimizers—to optimize performance, improve robustness, and adapt the model for operational deployment.

4.2. Hyper-Parameter Tuning

Proper parameter tuning is crucial for optimizing the performance of the machine learning models. In this research, experiments were conducted on the custom CNN model using various hyperparameters with different values. The objective was to identify the optimal value for each parameter to achieve the best possible performance from the classification model to detect the eye state. These parameters were adjusted with different values to assess their impact on the CNN model and to select the best value for each parameter, as shown in Table 2.

Table 2. Parameter Tuning

Parameter Tuning	
Parameter	Value / Setting
Learning Rate	0.01, 0.001, 0.0001
Epoch	15, 30, 50
Optimizer	Adam, SGD, RMSprop

The first parameter that was experimented with in Table 2 is the learning rate. The learning rate determines how quickly a machine learning model learns from the data during training. By experimenting with different learning rates, we can determine the optimal rate that enables the model to learn effectively without overlearning or underlearning. In this experiment, we use three different values of the learning rate: 0.01, 0.001, and 0.0001. A higher learning rate can speed up learning, but it might cause the model to miss the best solution. Conversely, a lower learning rate ensures more precise learning, but it may take longer.

After experimenting with the learning rate, another critical parameter that is being experimented with is the number of epochs. The epoch parameter defines the number of times the model iterates over the entire training dataset during the learning process. In this research, we tested the model with 15, 30, and 50 epochs to observe how its performance changes with increasing exposure to the data. More epochs can enhance model learning by iteratively adjusting weights, potentially improving accuracy. However, excessive epochs may lead to overfitting, where the model memorizes training data patterns instead of generalizing to new data. This experiment utilizes the optimal learning rate determined in the previous experiment.

Lastly, the experiment focuses on selecting the most effective optimizer for the model. The optimizer plays a crucial role in how the model updates its parameters based on the gradient of the loss function. A total of three different optimizers are used for the experiment. We evaluated Adam, SGD (Stochastic Gradient Descent), and RMSprop optimizers. Each optimizer has its strengths. Adam optimizer combines the advantages of adaptive learning rates and momentum; SGD updates parameters independently for each batch, and RMSprop adapts learning rates based on the moving average of squared gradients. By comparing these optimizers, we can identify which one achieves the best balance of convergence speed and model performance.

4.3. Experiment Results

Each experiment varied in terms of key parameters such as learning rate, epoch settings, and optimizer choices. These experiments aimed to assess how different configurations affected the model's ability to detect eye states accurately. This section presents a detailed evaluation of the model's performance under various conditions, enabling the identification of the most effective setups for real-world deployment. The performance of the model was evaluated using four evaluation metrics: Accuracy, Precision, Recall, and F1-score.

Based on Table 3, the learning rate significantly impacted the model's convergence and prediction performance. A high learning rate of 0.01 resulted in poor performance (accuracy = 0.500, F1-score = 0.335) due to overshooting during weight updates, which caused unstable learning and inefficient convergence. A moderate learning rate of 0.001 resulted in the optimum balance of convergence speed and accuracy, with an F1-score of 0.930 and an accuracy of 0.950. Reducing the learning rate further to 0.0001 resulted in somewhat worse accuracy (0.940) and recall, most likely due to slower learning and possible underfitting. This demonstrates that 0.001 is the best learning rate for this problem, resulting in efficient convergence while avoiding instability.

Table 3. Experimental Results

Parameter	Value/Setting	Precision	Recall	F1-score	Accuracy
Learning Rate	0.01	0.250	0.500	0.335	0.500
	0.001	0.955	0.950	0.930	0.950
	0.0001	0.940	0.935	0.940	0.940
Epoch	15	0.955	0.950	0.930	0.950
	30	0.965	0.965	0.970	0.970
	50	0.980	0.980	0.98	0.980
Optimizer	Adam	0.980	0.980	0.980	0.980
	SGD	0.945	0.940	0.935	0.940
	RMSprop	0.770	0.575	0.480	0.570

Increasing the number of epochs led to consistent performance improvements. At 15 epochs, the model acquired an Accuracy of 0.950 and an F1-score of 0.930, indicating that it learned well in a short amount of time. Extending training to 30 epochs increased all measures (Accuracy = 0.970, F1-score = 0.970), indicating that the model benefited from more passes over the data. The best performance was obtained at 50 epochs, with all metrics reaching or approaching 0.980. This shows that the dataset's size and complexity were ideal for lengthier training with minimal overfitting. However, more increases above 50 epochs would be required to determine whether performance plateaus or declines owing to overfitting.

The choice of optimizer has a significant impact. Adam consistently outperformed others, with the highest Precision, Recall, F1-score, and Accuracy (all 0.980), most likely because of its adaptive learning rate mechanism and ability to deal with sparse gradients efficiently. SGD produced good results (accuracy = 0.940, F1-score = 0.935), but it required more careful tweaking of learning rates and momentum. RMSprop fared much worse (Accuracy = 0.570, F1-score = 0.480), showing that it was less appropriate for the dynamics of this dataset and model design.

The experimental findings show that the ideal configuration for this model is a learning rate of 0.001, 50 epochs, and the Adam optimizer. This combination improves both learning stability and generalization capability, making it ideal for use in real-time eye-state detection applications like sleepiness and fatigue monitoring. These findings emphasize the necessity of hyperparameter tweaking in balancing model accuracy and computing efficiency, particularly in safety-critical applications.

4.4. Experiment Analysis

The experimental investigation focuses on determining the best parameter configuration to maximize the model's performance in detecting driver eye conditions using the Driver Drowsiness Monitoring System (DDMS). Comparing the findings across multiple learning rates, epoch counts, and optimizers reveals that the Adam optimizer, 50 epochs, and a learning rate of 0.001 produced the highest performance. This setup achieved the best accuracy (0.980), as well as similarly good precision, recall, and F1-score values (all 0.980), showing balanced performance across all assessment metrics.

A variety of reasons contribute to this configuration's superiority. The learning rate of 0.001 enabled the model to converge well without overshooting minima or being stranded in shallow local optima. Training for 50 epochs gave enough iterations for the network to fine-tune its feature representations while avoiding overfitting via stable generalization to new data. The Adam optimizer's variable learning rate method improved convergence efficiency and stability, making it especially useful for image-based classification tasks with small fluctuations, such as open versus closed eye states. In practice, high precision is crucial for decreasing false positives (incorrectly recognizing a motorist as drowsy), whereas high recall ensures that true sleepiness situations are not overlooked. The F1-score, which represents the harmonic mean of precision and recall, confirms that the selected configuration achieves a fair trade-off between these competing goals. High accuracy confirms the system's capacity to accurately categorize the vast majority of cases, which is crucial for safety-sensitive applications where detection errors can have serious implications.

In terms of drowsiness detection, these data show that the parameters used allow the model to dependably differentiate between open and closed eye states under varied settings, which is critical for identifying early tiredness cues on time and accurately. The findings further emphasize the importance of appropriate hyperparameter tuning in model performance, implying that even modest configuration changes might result in significant performance loss. As a result, the chosen parameter set not only represents the best configuration for the current dataset and model architecture but also serves as a solid foundation for future improvements, such as the incorporation of additional fatigue-related behavioral and physiological indications.

From a fatigue detection standpoint, this parameter configuration ensures that the feature extraction and classification stages are robust enough to generalize beyond binary eye states to more complex multimodal indicators, such as micro-yawning events (via MAR), head pose instability or decreased facial expressiveness. In practice, the same tuning can serve as the foundation for a multimodal, hierarchical classification system: the first stage focuses on detecting drowsiness with high confidence using ocular metrics, while subsequent layers incorporate additional visual, physiological, and contextual indicators to capture non-sleep-related fatigue.

4.5. System Testing

The testing phase of the system will be explained in this section. This step is significant to ensure that the DDMS operates as intended and meets its performance criteria. This phase involves two different types of testing. The first one is functionality testing, and the other one is accuracy testing. We need to conduct functionality testing first before proceeding to accuracy testing to ensure a smooth testing phase.

In this study, a CNN-based framework was built to categorize eye states (open vs. closed) as the key marker for drowsiness, while also providing the groundwork for detecting broader fatigue-related behaviors such as micro-yawning, slow blink recovery, and gaze instability. The model receives real-time video data from a driver-facing camera and uses facial landmark detection to separate key regions of interest, such as the eyes and mouth. Geometric measures like the Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR) are calculated from these regions and used as discriminative features to classify driver attention.

The CNN was trained with settings optimized for stability, speed, and accuracy in visual behavior classification (see Table 1). A learning rate of 0.001 was used to balance convergence speed and model stability, while 50 training epochs provided enough iterations for feature refining without overfitting. The Adam optimizer was chosen for its adjustable learning rate capabilities, which are useful for image classification problems with minor alterations in facial shape. We used a batch size of 32 and an input resolution of 145 × 145 pixels to optimize calculation without sacrificing detail. Data augmentation was used to improve robustness under a variety of illumination situations, head positions, and partial occlusions.

The functionality testing focuses on validating the system's overall functionality, including its user interface, hardware integration, and data processing capabilities. The user interface undergoes usability testing to evaluate how intuitive and simple it is to use, which is crucial for ensuring user acceptance and interaction. Hardware integration tests verify the system's ability to connect with cameras and other peripherals, ensuring stable data transfer and responsive operation. Moreover, it is also essential to test the system's ability to alert the driver and provide feedback. For example, visual alerts, such as on-screen warnings, should be clear, conspicuous, and capable of drawing immediate attention without distracting the driver. Auditory cues, such as alarm sounds, must be loud and distinct enough to be heard.

Next, testing the accuracy of the system. Accuracy testing is crucial for evaluating the system's reliability in accurately and consistently detecting drowsiness. This phase involves benchmarking the system's predictions against known datasets to assess its detection accuracy. Accuracy testing goes beyond mere numerical evaluation; it ensures that the system's detections align with real-world expectations and requirements. It validates not only the algorithmic efficiency of drowsiness detection but also its practical usability and reliability under various operational conditions.

The next test evaluates the system's ability to detect drowsiness and yawning under various scenarios. For this test, we will determine the system's performance across five distinct scenarios.

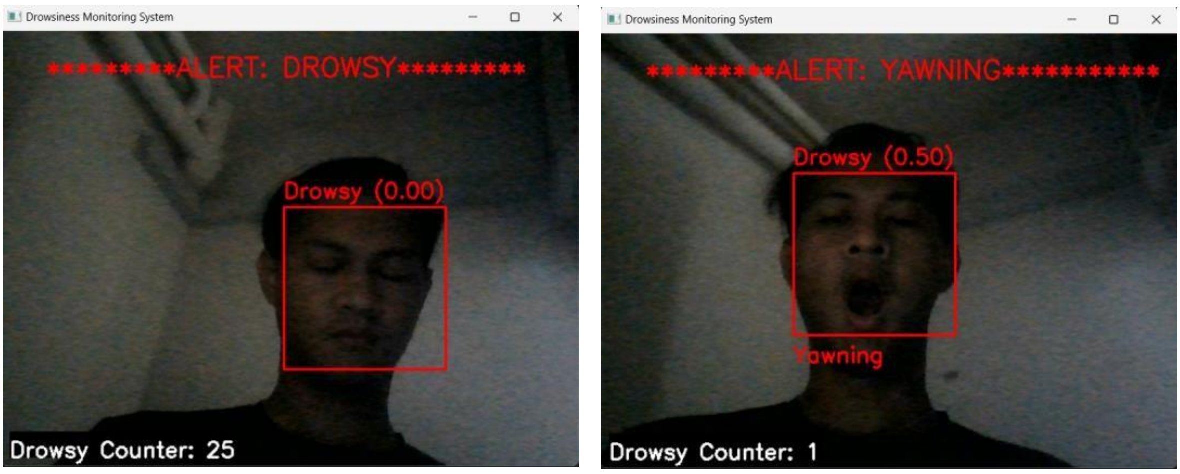

Figure 2. Low light intensity

In this scenario, as shown in Figure 2, we assume the driver operates the vehicle under low-light intensity conditions, without wearing eyeglasses, and maintains a normal head position. Based on the results, we can conclude that the system accurately detects both drowsiness and yawning, even in low-light settings.

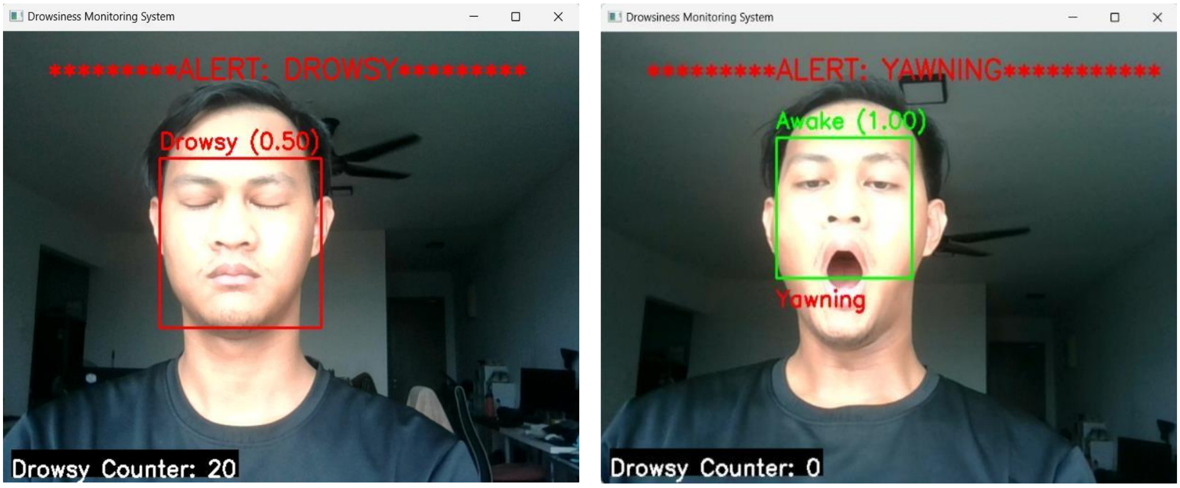

Figure 3. High Light Intensity

In this scenario, as shown in Figure 3, we assume that the driver operates the vehicle under high-light intensity conditions, without wearing eyeglasses, and maintaining a normal head position. Based on the results, we can conclude that the system can detect both drowsiness and yawning correctly even in the high-light settings.

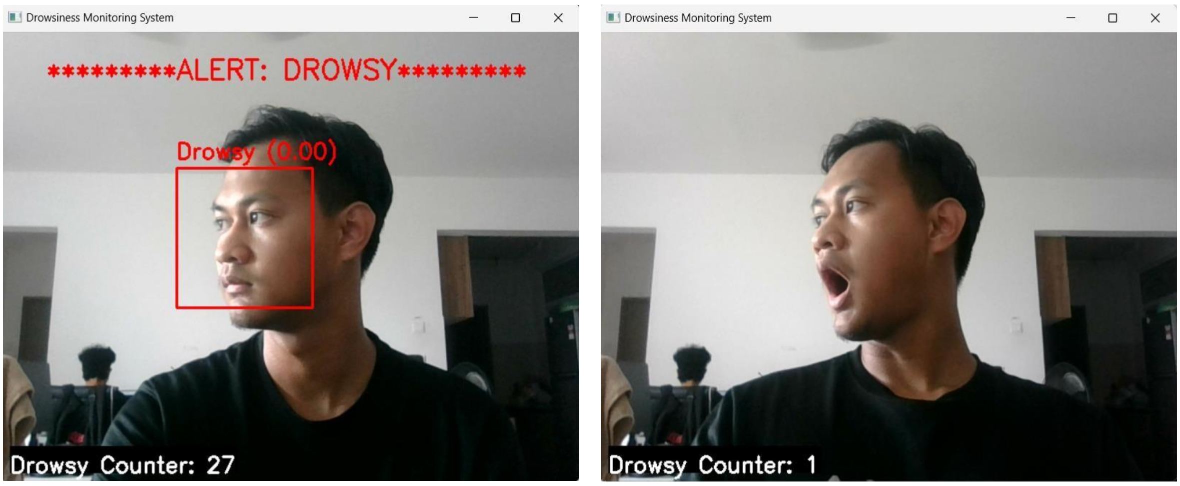

Figure 4. Head position tilt

In this scenario, as shown in Figure 4, we assume that the driver operates the vehicle under normal light intensity conditions, without wearing eyeglasses, but with their head turned to the side. The results indicate that even though the driver's eyes are wide open, the system still detects drowsiness. This means that the system did not accurately predict the drowsiness. The system is also unable to detect yawning events in this head position accurately.

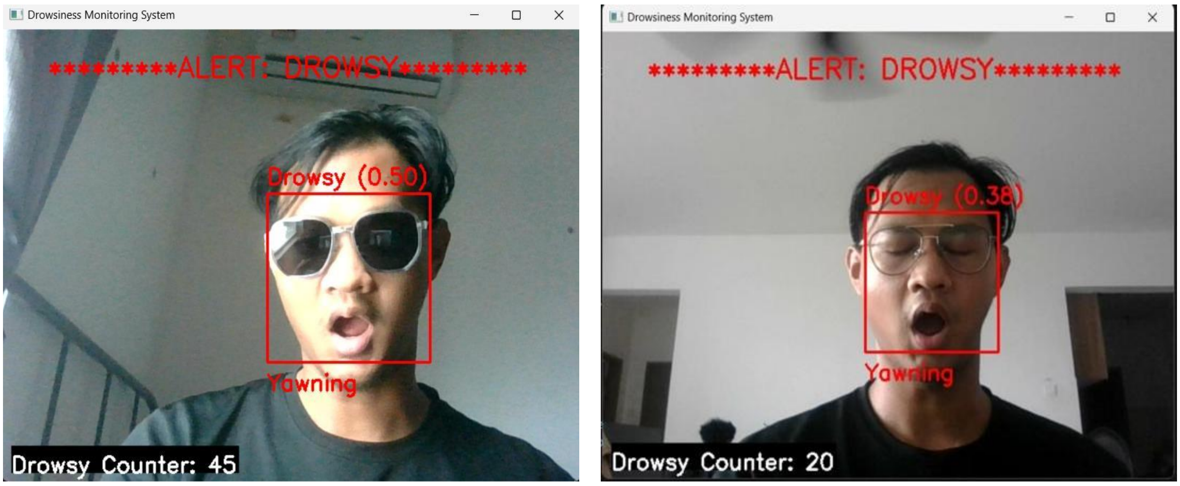

Figure 5. With eyeglasses and sunglasses

In this scenario, as shown in Figure 5, we assume the driver operates the vehicle under normal lighting conditions, wearing sunglasses, and with a normal head position. The results indicate that the system struggles to detect drowsiness effectively because sunglasses block the eye detection process.

5. Conclusions

This study successfully created a real-time Driver Drowsiness and Fatigue Monitoring System (DDFMS) using advanced computer vision techniques, notably Convolutional Neural Networks (CNNs) paired with facial landmark detection. The system was created to improve road safety by detecting early signs of drowsiness and fatigue through continuous monitoring of the driver's eye and mouth movements, with a focus on detecting prolonged eye closures, yawning, and other visual-motor cues that indicate performance decline. The study achieved three primary objectives:

- Identifying appropriate computer vision approaches for the accurate and non-intrusive identification of drowsiness and fatigue-related visual behaviors.
- Design and optimization of a CNN-based classification model with a curated eye state dataset to provide high detection accuracy for early intervention.
- Implementation of a fully functional real-time system capable of sending out timely alerts to lessen the risk of drowsiness and fatigue-related accidents.

The best configuration is using the Adam optimizer, a learning rate of 0.001, and 50 training epochs—achieved an accuracy of 98%, with equally strong precision, recall, and F1-score values. These findings verify the system's ability to identify eye state changes as a basic indicator of drowsiness, while also laying the groundwork for future research that will include additional indicators of non-sleep-related exhaustion. By combining CNN with facial landmark analysis and real-time alarm systems, the system displayed great accuracy and responsiveness in a variety of environments. However, limitations were found in circumstances of facial occlusion (e.g., sunglasses) and severe head positions, which affected identification reliability. Future enhancements should focus on:

- Expanding the dataset to include more diverse causes, contexts, and fatigue manifestations.
- Improving detection methods for occlusions and big head pose fluctuations.
- Using multimodal inputs (such as EEG, heart rate variability, and steering behavior) to detect non-drowsy exhaustion states that affect driver performance without causing sleepiness.

6. Declarations

6.1. Author Contributions

Conceptualization, D.A.D. and T.B.K.; methodology, D.A.D.; software, T.B.K.; validation, D.A.D., M.Z.Z., and S.K.; formal analysis, D.A.D.; investigation, M.; resources, S.K.; data curation, T.B.K.; writing—original draft preparation, D.A.D.; writing—review and editing, D.A.D., M.Z.Z., and S.K.; visualization, T.B.K.; supervision, D.A.D.; project administration, D.A.D. and M.; funding acquisition, D.A.D. All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

This research was funded by INTI International University, Malaysia, under the Internal Research Grant Scheme INTI IU Research Grant 2025: INTI-FDSIT-01-38-2025.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

7. References

[1] Liu, F., Chen, D., Zhou, J., & Xu, F. (2022). A review of driver fatigue detection and its advances on the use of RGB-D camera and deep learning. Engineering Applications of Artificial Intelligence, 116, 105399. doi:10.1016/j.engappai.2022.105399.

[2] Rahman, A., Hriday, M. B. H., & Khan, R. (2022). Computer vision-based approach to detect fatigue driving and face mask for edge computing device. Heliyon, 8(10), e11204. doi:10.1016/j.heliyon.2022.e11204.

[3] Civik, E., & Yuzgec, U. (2023). Real-time driver fatigue detection system with deep learning on a low-cost embedded system. Microprocessors and Microsystems, 99, 104851. doi:10.1016/j.micpro.2023.104851.

[4] Yu, L., Yang, X., Wei, H., Liu, J., & Li, B. (2024). Driver fatigue detection using PPG signal, facial features, head postures with an LSTM model. Heliyon, 10(21), e39479. doi:10.1016/j.heliyon.2024.e39479.

[5] AL-Quraishi, M. S., Azhar Ali, S. S., AL-Qurishi, M., Tang, T. B., & Elferik, S. (2024). Technologies for detecting and monitoring drivers' states: A systematic review. Heliyon, 10(20), e39592. doi:10.1016/j.heliyon.2024.e39592.

[6] Hasan, M. M., Watling, C. N., & Larue, G. S. (2024). Validation and interpretation of a multimodal drowsiness detection system using explainable machine learning. Computer Methods and Programs in Biomedicine, 243, 107925. doi:10.1016/j.cmpb.2023.107925.

[7] Rad, R., & Jamzad, M. (2005). Real time classification and tracking of multiple vehicles in highways. Pattern Recognition Letters, 26(10), 1597-1607. doi:10.1016/j.patrec.2005.01.010.

[8] Watling, C. N., Mahmudul Hasan, M., & Larue, G. S. (2021). Sensitivity and specificity of the driver sleepiness detection methods using physiological signals: A systematic review. Accident Analysis and Prevention, 150, 105900. doi:10.1016/j.aap.2020.105900.

[9] Kumar, V., Pham, H., Pandey, P. K., & Goel, A. (2021). Driving to safety: real-time danger spot and drowsiness monitoring system. Soft Computing, 25(23), 14479–14497. doi:10.1007/s00500-021-06381-1.

[10] Radzuan, N. Q., Hassan, M. H. A., Abu Kassim, K. A., Ab. Rashid, A. A., Mohd Razelan, I. S., & Othman, N. A. (2021). The Analysis of Road Traffic Fatality Pattern for Selangor, Malaysia Case Study. Mekatronika, 3(1), 79–88. doi:10.15282/mekatronika.v3i1.7155.

[11] El-Nabi, S. A., El-Shafai, W., El-Rabaie, E. S. M., Ramadan, K. F., Abd El-Samie, F. E., & Mohsen, S. (2024). Machine learning and deep learning techniques for driver fatigue and drowsiness detection: a review. Multimedia Tools and Applications, 83(3), 9441–9477. doi:10.1007/s11042-023-15054-0.

[12] Fu, B., Boutros, F., Lin, C. T., & Damer, N. (2024). A Survey on Drowsiness Detection: Modern Applications and Methods. IEEE Transactions on Intelligent Vehicles, 9(11), 7279–7300. doi:10.1109/TIV.2024.3395889.

[13] Wang, S., Zhong, L., Fu, Y., Chen, L., Ren, J., & Zhang, Y. (2024). UFace: Your Smartphone Can "Hear" Your Facial Expression! Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 8(1), 27. doi:10.1145/3643546.

[14] Nair, A., Patil, V., Nair, R., Shetty, A., & Cherian, M. (2024). A review on recent driver safety systems and its emerging solutions. International Journal of Computers and Applications, 46(3), 137–151. doi:10.1080/1206212X.2023.2293348.

[15] Albadawi, Y., Takruri, M., & Awad, M. (2022). A Review of Recent Developments in Driver Drowsiness Detection Systems. Sensors, 22(5). doi:10.3390/s22052069.

[16] Bitkina, O. V., Park, J., & Kim, H. K. (2021). The ability of eye-tracking metrics to classify and predict the perceived driving workload. International Journal of Industrial Ergonomics, 86, 103193. doi:10.1016/j.ergon.2021.103193.

[17] Joliot, M., Cremona, S., Tzourio, C., & Etard, O. (2024). Modulate the impact of the drowsiness on the resting state functional connectivity. Scientific Reports, 14(1), 1–15. doi:10.1038/s41598-024-59476-8.

[18] Andrillon, T., Taillard, J., & Strauss, M. (2024). Sleepiness and the transition from wakefulness to sleep. Neurophysiologie Clinique, 54(2), 102954. doi:10.1016/j.neucli.2024.102954.

[19] Chang, R. C. H., Wang, C. Y., Chen, W. T., & Chiu, C. Di. (2022). Drowsiness Detection System Based on PERCLOS and Facial Physiological Signal. Sensors, 22(14), 5380. doi:10.3390/s22145380.

[20] Jones, A. W. (2022). Driving Under the Influence of Alcohol. Handbook of Forensic Medicine, 1387–1408. Portico. doi:10.1002/9781119648628.ch59.

[21] Freitas, A., Almeida, R., Gonçalves, H., Conceição, G., & Freitas, A. (2024). Monitoring fatigue and drowsiness in motor vehicle occupants using electrocardiogram and heart rate − A systematic review. Transportation Research Part F: Traffic Psychology and Behaviour, 103, 586–607. doi:10.1016/j.trf.2024.05.008.

[22] Antony, M. M., & Whenish, R. (2021). Advanced Driver Assistance Systems (ADAS). Automotive Embedded Systems, 165–181. doi:10.1007/978-3-030-59897-6_9.

[23] Mahanama, B. (n.d.). Eye Movement and Pupil Measures: A Review. Frontiers in Computer Science, 3, 733531. doi:10.3389/FCOMP.2021.733531/XML/NLM.

[24] Niu, J., & Ma, C. (2022). Is it Good or Bad to Provide Driver Fatigue Warning During Take-Over in Highly Automated Driving? Transportation Research Record, 2676(2), 762–774. doi:10.1177/03611981211046920.

[25] Shoshina, I. I., Kovalenko, S. D., Kuznetsov, V. V., Brak, I. V., & Kashevnik, A. M. (2024). Literature Review on Detection of Fatigue State Based on Eye Movement Monitoring. Human Physiology, 50(3), 260–275. doi:10.1134/S0362119724700737.

[26] Sri Mounika, T. V. N. S. R., Phanindra, P. H., Sai Charan, N. V. V. N., Kranthi Kumar Reddy, Y., & Govindu, S. (2022). Driver Drowsiness Detection Using Eye Aspect Ratio (EAR), Mouth Aspect Ratio (MAR), and Driver Distraction Using Head Pose Estimation. ICT Systems and Sustainability, 619–627. doi:10.1007/978-981-16-5987-4_63.

[27] Karthickmanoj, R., Aasha Nandhini, S., Sasilatha, T., & Lakshmi, D. (2023). Smart Cyber-Physical System-Based Plant Disease Detection for Agriculture. Contemporary Developments in Agricultural Cyber-Physical Systems, 204–222. doi:10.4018/978-1-6684-7879-0.ch011.

[28] Badashah, S. J., Alam, A., Jawarneh, M., Moharekar, T. T., Hariram, V., Poornima, G., & Jain, A. (2025). Cancer Classification and Detection Using Machine Learning Techniques. Natural Language Processing for Software Engineering, Wiley, 95-111. doi:10.1002/9781394272464.ch6.