



---

*Research article*

## A sophisticated Drowsiness Detection System via Deep Transfer Learning for real time scenarios

Amina Turki<sup>1</sup>, Omar Kahouli<sup>2,\*</sup>, Saleh Albadran<sup>3</sup>, Mohamed Ksantini<sup>1</sup>, Ali Aloui<sup>2</sup> and Mouldi Ben Amara<sup>2</sup>

<sup>1</sup> Control & Energies Management Laboratory, National Engineering School of Sfax, University of Sfax, Tunisia

<sup>2</sup> Department of Electronics Engineering, Community College, University of Ha'il, 2440, Ha'il, Saudi Arabia

<sup>3</sup> Department of Electrical Engineering, College of Engineering, University of Ha'il, Ha'il 2440, Saudi Arabia

\* **Correspondence:** Email: [a.kahouli@uoh.edu.sa](mailto:a.kahouli@uoh.edu.sa).

**Abstract:** Driver drowsiness is one of the leading causes of road accidents resulting in serious physical injuries, fatalities, and substantial economic losses. A sophisticated Driver Drowsiness Detection (DDD) system can alert the driver in case of abnormal behavior and avoid catastrophes. Several studies have already addressed driver drowsiness through behavioral measures and facial features. In this paper, we propose a hybrid real-time DDD system based on the Eyes Closure Ratio and Mouth Opening Ratio using simple camera and deep learning techniques. This system seeks to model the driver's behavior in order to alert him/her in case of drowsiness states to avoid potential accidents. The main contribution of the proposed approach is to build a reliable system able to avoid false detected drowsiness situations and to alert only the real ones. To this end, our research procedure is divided into two processes. The offline process performs a classification module using pretrained Convolutional Neural Networks (CNNs) to detect the drowsiness of the driver. In the online process, we calculate the percentage of the eyes' closure and yawning frequency of the driver online from real-time video using the Chebyshev distance instead of the classic Euclidean distance. The accurate drowsiness state of the driver is evaluated with the aid of the pretrained CNNs based on an ensemble learning paradigm. In order to improve models' performances, we applied data augmentation techniques for the generated dataset. The accuracies achieved are 97 % for the VGG16 model, 96% for VGG19 model and 98% for ResNet50 model. This system can assess the driver's dynamics with a precision rate of 98%.

**Keywords:** Driver Drowsiness Detection (DDD) system; deep transfer learning; Convolutional Neural Networks (CNNs); the Chebyshev distance

**Mathematics Subject Classification:** 00A06, 68T01, 68T07

---

## 1. Introduction

Road safety is an issue that deserves more attention since it represents one of the great opportunities to save more lives around the world. In accordance with the World Health Organization (WHO) factsheets, road accidents are the eighth reason of mortality globally, while it is fourth in the United States. Each year, roughly 1.3 million people die due to road traffic crashes, that is, one person dies every 25 seconds on average [1]. In addition, road traffic injuries are the first leading cause of death for children and adolescents aged between 5 and 29 years. Besides, about 50 million people suffer from in-juries and as a result many endure disabilities. Nevertheless, all those deaths and injuries are preventable. The critical reason in the crash chain is assigned to the driver. Approximately 94% of all car accidents are caused by human errors [2]. According to the American National Highway Traffic Safety Administration (NHTSA), the major cause of fatal crashes is the driver state related factor: Drowsiness, alcohol and speed [2].

Drowsy driving is the act of driving while tired and feeling fatigued or sleepy. Indeed, driver drowsiness is a hazardous state that occurs unintentionally and can lead to fatalities, but it is difficult to detect. Driver drowsiness is one of the major factors responsible for crash severity, especially for vulnerable road users [3]. The American National Safety Council (NSC) declares that each year, drowsy driving counts about One Hundred Thousand crashes, Seventy Thousand injured persons, and more than One Thousand Five Hundred mortalities [4]. It contributes to an estimated 9.5% of all crashes based on the statistics of the American Automobile Association (AAA) Foundation for Traffic Safety [5].

It is possible to notice the drowsiness state of a driver in its initial phase and keep the driver alert in order to get over a potential crash using DDD systems that monitor and detect the driver's behavior and especially drowsiness [6]. In addition to the obvious need, the integration of these systems will become compulsory for automobile manufacturers in the European Union (EU) soon [7]. Indeed, the European Commission (EC) decided to include DDD as one of the safety mandates in cars sold in the European Union [7]. In addition to that, the European New Car Assessment Program (Euro NCAP) will incorporate drowsiness detection as part of its assessment protocol when evaluating a vehicle's safety ratings in the 2025 Roadmap [6].

Several researchers have recently considered the automatic DDD systems. Three major categories of measures that have been used widely for these systems are [8–10]:

(1) Vehicle measures: Some parameters like deviations from lane position, pressure on the acceleration pedal and the steering wheel shaking are regularly supervised. Any changes in these parameters that surpass a particular threshold is very likely linked to a drowsy driver [11,12].

(2) Behavioral measures: The conduct of the driver especially facial expressions (yawning, eyes' closure, eyes' blinking), head pose, body and hand movements are monitored through a camera or visual sensors and the driver is notified if any of these drowsiness signs are detected [13–15].

(3) Physiological measures based on physiological signals. The evaluation of the cognitive and concentration signs in response to the perceptual stimulus is doable using physiological measurement

signals such as Electro-oculogram (EOG), Electromyogram (EMG), electrodermal activity (EDA), Photoplethysmogram (PPG), Electrocardiogram (ECG) and Electroencephalogram (EEG) signal which is the most used [9,10,16].

In this work, we will develop a sophisticated hybrid DDD system to recognize accurately and efficiently the drowsiness behavior of a driver in order and to alert him depending on both the eyes closure and the mouth opening ratios and using a combination of images issued from car camera and deep learning (DL) techniques. The proposed approach comprises two processes. An offline process is dedicated to the learning module of three CNN networks based on a dataset. We enhanced this dataset using appropriate data augmentation techniques to achieve the classification task with good accuracies. An online process is devoted to the detection module and the prediction module. Indeed, a car cam that records real-time images of the driver is used. The driver's facial landmarks will be limited in a frame using image recognition techniques. We will determine the Eye Closure Ratio (ECR) and the Mouth Opening Ratio (MOR) using the Chebyshev distance for the first time in literature to the best of our knowledge. This distance gives the best result in terms of accuracy compared to other distance measures. The driver drowsiness state is detected based on the ECR or/and the MOR values. Sometimes, and due to environmental factors, some false drowsy states causing false alerts can be detected. This can be intrusive and annoy drivers that may end up not believing them anymore. Thus, in order to evade falsely detected drowsy states, we use the pretrained CNN models to identify the real situation of the driver. The fusion of these predictors, based on ensemble learning methods, will determine whether the driver is tired. Concisely, the novelty of this DDD system lies in; (i) the use of the Chebyshev distance, (ii) the ability to detect the falsely detected drowsy states and (iii) the use of ensemble learning method to accurately recognize the drowsiness state. The proposed DDD system is characterized by its cost-efficiency, ease of use, non-invasiveness, and automatic functionality. These features collectively contribute to its seamless integration into industrial processes, ensuring minimal disruption and maximum efficiency.

Thus, this paper is organized as follows. In Section 2, we will detail concepts in connection with the proposed DDD system including the drowsiness definition, the facial behavioral measures and transfer for DDD systems and some research studies related to that. Section 3 introduces the proposed approach, methodology and materials, it involves the offline process representing the learning module and the online process displaying the detection module. In Section 4, we will detail experimental results showing the performances metrics and some simulations figures related to the proposed system. Section 5 focuses on the proposed method and provides in-depth discussions and comparisons with other existing works. This section highlights the promising nature of our method in terms of both efficiency and accuracy, particularly when compared to other approaches. These findings contribute to establishing the credibility and potential of our proposed method in the field. Finally, in Section 6, the paper ends with a comprehensive and thought-provoking conclusion that summarizes the key findings, implications, and potential avenues for future research.

## **2. Background & related works**

### *2.1. Drowsiness*

Drowsiness is defined as the transition between wakefulness and sleepiness. If a person falls asleep when driving, he will lose control of the car steadily, which often leads to a car crash. Therefore,

drowsy driving, whether it is due to fatigue, medication or a sleep disorder, is a serious issue. In order to obviate these situations, the state of a drowsy driver should be detected accurately and early. DDD systems based on face recognition have a great importance in detecting and alerting the driver to avoid fatal situations, especially, when combined with DL techniques. In our work, we focus on the study of DDD systems based on facial expressions measures.

## 2.2. Facial expressions' behavioral measures for DDD systems

Studying the features of the driver's physical behavior represents a good method to detect more efficiently the driver's drowsiness. DDD systems are generally classified into three types, based on the mouth, head and eye movements.

Table 1 enumerates some of the facial expressions' behavioral measures [17].

**Table 1.** Some of the facial expressions' behavioral measures.

Facial expressions	Description
Blink frequency [18]	The frequency of an eye blinks during a specific period of time.
Maximum closure duration of the eyes [18]	The maximum duration of the eye was closed.
PERcentage of eyelid CLOSure (PERCLOS) [19]	The percent of time in which the eye is 80% closed or more.
Eye Aspect Ratio (EAR) [20–23]	EAR indicates the eye's opening degree. It is equal to zero when the eyes are closed, and it is invariable when the eyes are open.
Yawning frequency [24]	The frequency of the mouth opens for a defined period.
Mouth Aspect Ratio [25]	MAR is calculated based on the height and width of the mouth.

There are many DDD systems which are based on facial expressions. They use many and diverse parameters and methods to conceive their detection procedure.

**Eye state:** the eye state is one of the relevant methods among behavioral measures to detect driver drowsiness. Features like the eye-opening rate, the eyelid distance and PERCLOS are among the top indicators of the driver's drowsiness.

Khan et al. [26] proposed a real-time DDD system that utilized eyelid closure as a key indicator. The technology employed surveillance videos to monitor whether the driver's eyes were open or closed. The system began by identifying the driver's face, followed by locating the eyes and assessing the curvature of the eyelids using an extended Sobel operator. By determining the concavity of the curve, the system classified the eyelids as either open or closed. If the eyes remained closed for a fixed duration, an alarm was triggered. The system was evaluated using three datasets. The first dataset achieved an accuracy of 95%, the second dataset attained 70% accuracy, and the third dataset surpassed 95% accuracy.

A sleepiness detection technique was created by Maior et al. [27] based on the eyes' movements using a basic web camera. The EAR metric was used to calculate the blinking period of time. To determine the EAR value, the ratio between an eye's height and width is computed. When the EAR value is high, the eye is open, and when it is low, the eye is closed. The idea consists of three steps, the EAR computation, the blink classification, and the real-time detection of potential drowsiness state. The EAR were measured and saved at each frame. Unless the blink period is smaller than a specific one,

drowsiness will be indicated. Accuracy attained was 95%.

Zandi et al. [28] proposed the use of eye tracking data as a non-intrusive method for detecting driver drowsiness. In their study, this data was collected from 53 drivers during a simulated driving experiment. First, multichannel EEG signals were recorded to serve as the baseline reference. For the analysis, a binary classification approach was employed, utilizing both a Random Forest (RF) and a Support Vector Machine (SVM) classifier. Different lengths of eye tracking epochs were chosen for feature extraction, and each classifier performance was evaluated for every epoch length. The results showed that the RF classifier achieved an accuracy of 88.37% to 91.18% across all epoch lengths, outperforming the SVM classifier, which achieved an accuracy of 77.12% to 82.62%.

A real-time DDD system based on the eye closure using DL was proposed by Hashemi et al. [29]. For the classification of eye closure, a Fully Designed Neural Network (FD-NN) along with Transfer Learning using the VGG16 and VGG19 models, were developed. To enhance the training data, the authors supplemented the dataset with 4157 new photographs. FD-NN achieved an accuracy of 98.15%. Furthermore, the Transfer Learning approach with the VGG16 model achieved an accuracy of 95.45%, while the VGG19 model achieved an accuracy of 95%.

**Mouth state:** Mouth state is used to predict driver drowsiness in real-time driving situations in numerous studies. Alioua et al. extracted features from mouth movements using an SVM and a method based on the Circular Hough Transform (CHT) for DDD system [30]. These systems were effective in real-time scenarios in various lighting environments. Based on real images, the experiment's findings showed that yawning could be detected with an accuracy rate of 81%.

Xiaoxi et al. [31] performed a DDD system based on CNN using depth video sequences to detect driver fatigue during the nighttime. Indeed, two CNNs were used: A spatial CNN that locates objects, and a temporal CNN that searches for information between adjacent frames. These two CNNs enable the system to calculate motion vectors at each frame, allowing it to detect yawns even when the driver covers his or her mouth with a hand. Experiments display an accuracy of 91.57%.

**Hybrid behavioral features (eye and mouth states):** In recent years, studying DDD systems has been generally based on hybrid behavior analysis. Savasx et al. used a CNN model to detect drowsiness based on both mouth and eye states [32]. The authors showed how to use PERCLOS data and mouth yawning frequency to detect driving drowsiness in real-time. These data were trained with various CNN for the identification of drowsiness factors. The results display an accuracy of 98.81%.

Celecia et al. [33] suggested an economical and accurate DDD system. Images were recorded on camera with an infrared illuminator. The system's process incorporates the obtained features from the eyes and mouths using a Raspberry Pi 3 Model B. The features states were done through a cascade of regression tree algorithms and then they were combined by a Mamdani fuzzy inference system, in order to predict the state of the driver. The system produces a final output that indicates three degrees of drowsiness. The study resulted in a DDD system with 95.5% accuracy resilient to various ambient illumination situations.

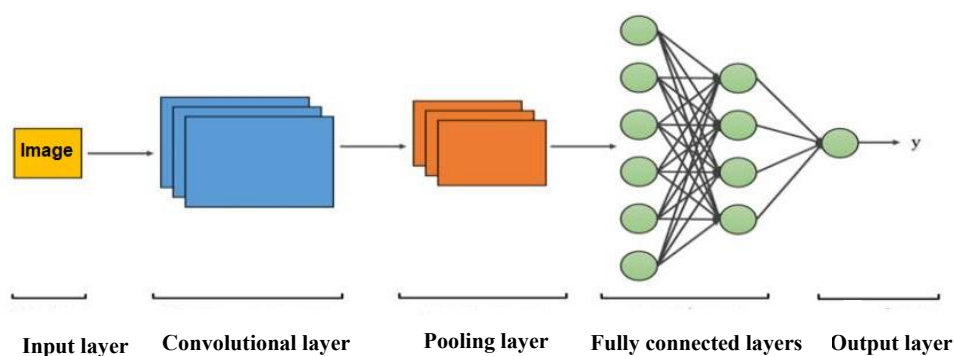
A non-intrusive method that can identify a real state of drowsiness was presented by Alioua et al. [34]. Closed eyelid and open mouth recognition techniques are used to identify drowsiness based on images collected from a webcam. To determine the face region in each image, the system used an SVM face detector. The Hough Transform was employed to locate the mouth and eyes' regions in the face. Based on the extracted features, the system made determinations regarding the driver's drowsiness. The results confirmed the system's robustness, achieving 94% accuracy and 86% kappa statistic value.

In this work, we perform the proposed DDD system according to facial expressions based on both

eyes and mouth states. The ECR and the MOR are calculated utilizing the Chebyshev distance in contrast to the latest works which exclusively used the Euclidean distance. The Chebyshev distance provides better accuracy especially in cases where the operational speed is inherent. The proposed DDD system can recognize a real-time drowsy situation with high performances thanks to the synergy between the offline and the online module of the proposed approach. Indeed, the fusion of DNNs decides the real situation of the driver by overthrowing the falsely detected drowsy states.

### 2.3. Transfer learning for DDD systems

Euclidean distance. Deep Learning (DL), an area of Machine Learning (ML) community, is one of the most important re-search trends thanks to its considerable success in many domains. One of the benefits of DL networks is the ability to learn large amounts of data that allow us to achieve excellent results for many complex cognitive tasks. Among the most popular DL networks, there is a definite type called convolutional neural networks (CNNs). The main advantage of CNNs is their capability to find patterns automatically and to detect the important features among images without any human guidance that made it the most widely used [35,36]. A CNN architecture is represented in Figure 1.



**Figure 1.** A CNN architecture.

A pretrained CNN model is a model that has been trained on a large dataset to learn a specific task, such as image recognition or natural language processing, and then saved for later use. There is a wide range of pre-trained models that are available, such as Inception, Xception, VGG and ResNet families. The choice of the suitable model for one task is based on the domain of the application, the task at hand and the availability of data. Otherwise, deep transfer learning requires a significant correlation between the data of the pre-trained model and the target task so that they are matching [37].

A relevant concept linked to CNNs is transfer learning [37]. This promising technique consists of using a powerful pre-trained CNN model, called Deep Neural Networks (DNNs), to solve a different task of a similar domain, and use it to resolve a new task [38]. Transfer learning means that training does not require high amounts of data and does not need to be started over for every new task. Thus, transfer learning saves both resources and time [39]. The lesser training time leads to better predictive performance and initializing the network with pretrained structures improves the generalization even after considerable fine-tuning to the used dataset [40].

Many studies using CNNs based on the DL techniques have been presented to detect drowsy drivers. The study in [41] analyzed the driver images using a CNN with VGG16 classification algorithm to determine if the eyes of the driver are open or closed and if he is yawning. The VGG16 model attained 91 % of accuracy and an F1-score of over 90 % for each class.

In order to detect tiredness, Dua et al. [42] suggested an architecture of four DL models that use RGB videos of drivers as input. Based on an ensemble process, the output of these models is combined to produce the final output, which is determined by a SoftMax classifier. 85% of accuracy was attained. Yu et al. [43] proposed a framework for the driver drowsiness detection based on 3D-deep convolutional neural network. The recognition of driver's drowsiness status was done using the condition-adaptive representation with an accuracy of 76,2 %. Park et al. [44] proposed a DL network using three deep neural networks to access the driver facial features. For the DDD system, the three networks' outputs are pooled and incorporated into a Softmax classifier. Results from the experiment have a detection accuracy of 73.66%. Abbas [45] designed a HybridFatigue system, which combines both visual and non-visual features to detect driver fatigue. The visual features were measured through PERCLOS to calculate the degree of eye closure. The non-visual features were obtained using ECG sensors, providing physiological data to complement the visual information. To process and analyze these hybrid features effectively, a multi-layer-based transfer learning approach was employed. This approach utilized a combination of CNNs and deep-belief network (DBN) models. This system achieved a detection accuracy of 94.5% when tested on 4250 images.

### 3. Proposed approach

#### 3.1. Description and objectives of our research

In this section, we introduce the considered DDD system to detect driver drowsiness from different driving scenarios based on pretrained CNNs using transfer learning techniques. The DDD system consists of two processes and three modules: an offline process that lies in the learning module and an online process that involves the detection module and the prediction module.

The key contributions of the proposed approach are as follows:

- We propose a novel DL approach that can automatically detect and estimate the driver's drowsiness based on camera and deep transfer learning methods.
- We utilize the Chebyshev distance to check the eyes and the mouth states (open or closed) based on facial landmarks to detect drowsiness.
- We generate data augmentation techniques for the used dataset in order to enlarge and enrich the data towards improving the training procedure.
- We classify the detected drowsiness state by three pretrained CNN models, which effectively improve the performance of the DDD system.
- In order to obtain better recognition performance, we refer to the fusion of classifiers by combining all the models' outputs to find the final prediction based on the ensemble learning theory (the bagging process).
- The proposed system is able to avoid falsely detected drowsiness situations and to alert only the real ones.

### 3.2. Offline process (Learning module)

The offline process consists of training three CNN models; the VGG16, the VGG19 and the Res-Net50. These models represent the most object identification accuracies [46]. They will be used later to decide if the driver is drowsy or not for a real-time detected drowsiness state.

#### 3.2.1. Dataset used

In this study, the YAWDD dataset [47] was used in the offline process for the classification task of driver drowsiness. This data set is composed of 2900 samples with various facial features of both male and female drivers recorded by an in-car camera. It is divided into four categories: Yawn, no-yawn, open eye and closed eye. Figure 2 shows the image samples of the open eye, closed-eye, yawn and no-yawn classes from the dataset. The samples are almost equally distributed on the four classes.



**Figure 2.** Image examples from the dataset.

#### 3.2.2. Training

In this study, the YAWDD considers DL advantages, and we applied deep neural networks with multiple hidden layers for driver drowsiness detection on the presented dataset. Three DNNs were considered for training:

- VGG16 is a pretrained CNN model proposed in 2014. The model attains 92.7% test accuracy in ImageNet that is a dataset of more than 14 million images including 1000 classes [48].
- VGG19 is a pretrained CNN model that is trained using more than one million images from the ImageNet database. The network covers 19 layers and can classify images into 1000 classes [48]. VGG16 and VGG19 are both developed by the Visual Graphics Group (VGG) at the University of Oxford. They are known for their simplicity and uniform architecture. VGG models use small 3x3 convolutional filters with a stride of 1 and max-pooling layers to down sample the spatial dimensions.
- ResNet50 is a widely used ResNet model developed by Microsoft Research. It is a pretrained CNN with 50 deep layers on top of each other to build a network [49]. Like VGG16 and VGG19, ResNet50 is also pretrained on the ImageNet dataset for image classification tasks. ResNet50 introduces skip connections, or shortcuts, to the traditional deep network architecture, which

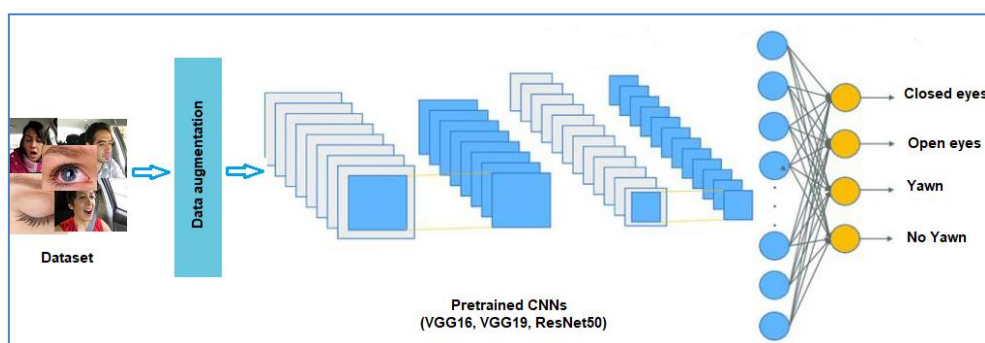


allows the model to learn residual functions. This helps in training deeper networks without the vanishing gradient problem.

The main difference between these two families of model lies in the architecture and approach to learning. While VGG models focus on stacking more layers, ResNet50 emphasizes learning residual functions to address the challenges of training very deep networks.

### 3.2.3. Data augmentation

As they have many tunable parameters, the performance of DNNs depends on the quality and quantity of training data. In fact, these DNNs need a large amount of data in the training phase in order to perform a complex task with high accuracy. However, the dataset used for the training is reduced. Therefore, we refer to data augmentation techniques aiming to artificially increase the quantity of data by producing new data points from available data. This includes making small transformations to the original data by making simple alterations on image data such as geometric transformations, color, and color space transformations. Figure 3 represents an overview of the training process.



**Figure 3.** Description of the learning module.

## 3.3. Online process (Learning module)

### 3.3.1. The detection module

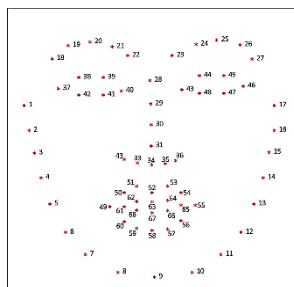
A simple car camera is placed on the top of the vehicle. The face region of the driver is detected from live video. The identification of the eyes and mouth landmarks are obtained by the application of Dlib toolkit. The ECR and the MOR, calculated using the coordinates of landmarks, are presented to identify respectively the state of eyes and the mouth. If the eyes are closed or/and she/he is yawning, a drowsiness state is detected. The detection module mechanism is given in Figure 4.



**Figure 4.** Description of the detection module.

**Identification of facial landmarks:** The driver video is introduced directly to the Dlib [50] library frame by frame. Dlib toolkit is an open-source library using C++ language and including machine learning algorithms and tools used in different applications and in many domains. We used Dlib to identify the driver's face's essential features after extracting the driver's face from the image at each frame. A facial landmark detector with 68 face-specific coordinate points is offered by the Dlib package. Viola and Jones [51] proposed this machine learning-based detection technique.

Kazemi et al. [52] proposed a technique that estimates the positions of facial landmarks on images. This technique can be used for real-time facial features detection including the eyes, eyebrows, nose, ears, and mouth. This detector identifies 68 major facial features positions, as shown in Figure 5.



**Figure 5.** The 68 facial landmark points of human face.

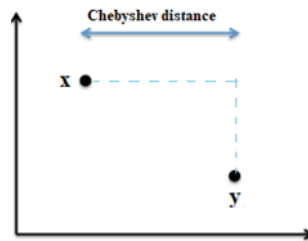
Based on the index of the face sections, we can extract the specific facial structures. Thus, we can detect and access both the eye and mouth regions according to the following facial landmark index:

- The right eye: (36, 42),
- The left eye: (42, 48),
- Mouth: (49, 68).

In this work, we used 32 feature marks characterizing the left eye, right eye, and mouth to calculate eye closure and mouth opening. We calculated the ECR and the MOR, using the Euclidean distance and the Chebyshev distance. The Chebyshev distance performs better than the Euclidean distance. Hence, we refer to the use of the Chebyshev distance.

The Chebyshev distance, contrary to the Euclidean distance, is typically used in very specific use cases which makes it difficult to use as an all-purpose distance. It is generally used when the implementation speed is critical. The main benefit of using the Chebyshev distance over the Euclidean distance is that it takes less time to choose the distances between the pixels. It is the most appropriate distance used in cases where the execution speed is critical [53].

**The Chebyshev distance** (Figure 6) is a metric in which the distance between two vectors is the greatest of their differences. The Chebyshev distance guarantees that the next considered points are potentially placed at the border of the region of point in one dimension, and these points usually discover a new area of the search space [54].



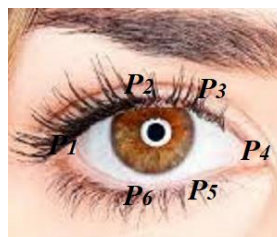
**Figure 6.** Chebyshev distance between two points.

The Chebyshev distance between two points  $x$  and  $y$  or two vectors with standard coordinates  $x_i$  and  $y_i$  is:

$$D(x, y) = \max_i (|x_i - y_i|). \quad (1)$$

**Eye closure ratio (ECR)** is a scalar value that responds to the estimation of the eye closure state. ECR formula is indifferent to the direction and distance of the face and ensures the advantage of identifying faces from a distance [55,56]. Each eye is represented by six coordinates, as shown in Figure 7. ECR value is calculated by using the following equation:

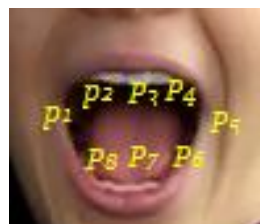
$$ECR = \frac{\max(|p_2 - p_6|) + \max(|p_3 - p_5|)}{2 \max(|p_1 - p_4|)}. \quad (2)$$



**Figure 7.** The facial landmarks related to eyes ( $p_1$ – $p_6$ ).

**Mouth opening ratio (MOR):** Yawning is marked by mouth opening as shown in Figure 8. A parameter used to determine whether or not someone is yawning is the mouth-opening value. Like ECR, MOR is determined as:

$$MOR = \frac{\max(|p_2 - p_8|) + \max(|p_3 - p_7|) + \max(|p_4 - p_6|)}{2 \max(|p_1 - p_5|)}. \quad (3)$$



**Figure 8.** Mouth yawning with facial landmarks ( $p_1$ – $p_6$ ).

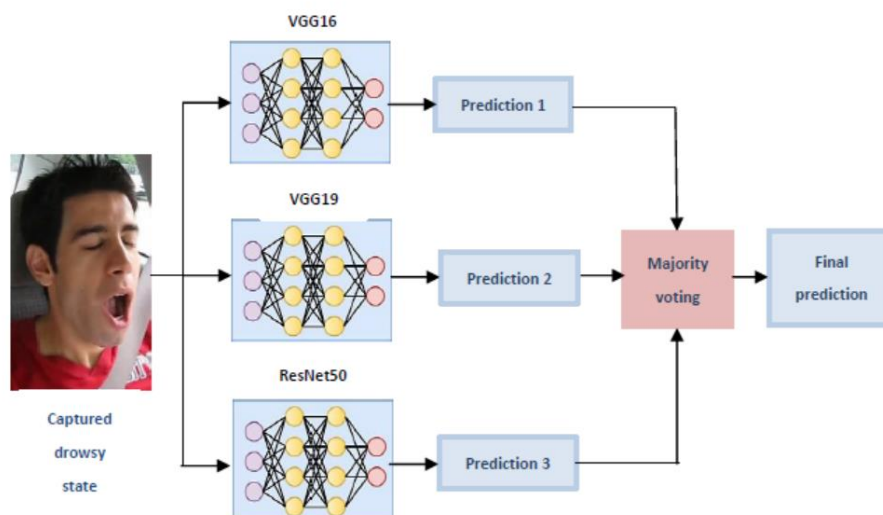
If the eyes of a driver are closed and/or he is yawning, a drowsiness state is detected.

**Drowsiness state:** The following conditions must be crossed to detect a drowsy driver:

- 1) The driver is drowsy when the output of the detector module is greater than a drowsiness threshold. This value ranges generally between 0 and 1. In our case and after several tests, we have chosen a threshold of 0.3.
- 2) The average blink duration of a person is between 0.1 and 0.4 seconds. Thus, a person must be drowsy when his ECR is beyond this interval. Drowsiness is detected if the eyes are closed for over five seconds.
- 3) If the MOR is maximum, the driver is yawning and therefore he is drowsy.
- 4) The driver is obviously drowsy if he meets condition 1 and condition 3 at the same time.

### 3.3.2. The prediction module

Sometimes, some falsely detected drowsiness situations causing false alerts can be recorded. These false alerts can annoy drivers. That is why we perform the prediction module, as shown in Figure 9, to alert only the real drowsiness states.



**Figure 9.** Description of the prediction module.

If a driver is detected drowsy, a captured image of the current frame passes through the three transfer learning CNN models to assess if it is a real drowsiness state or a false one. In order to achieve a performant DDD system, the three models' predictions will be combined to predict the accurate one based on an ensemble learning method. Ensemble learning is a machine learning tool where several models (VGG16, VGG19 and ResNet50) were trained to solve the same problem in order to boost the overall accuracy. For the proposed DDD system, we adopt the bagging fusion strategy as an ensemble learning method. The bagging combines all the predictions of all learners by majority voting and the most-voted class is decided in order to assess the accurate class of the driver (drowsy or not). If the driver is confirmed drowsy, then, the alarm is triggered.

To measure the effectiveness of the proposed approach, we need to evaluate the used CNN models. The evaluation metrics provide convenient measures of performance for the proposed system.

### 3.4. Model evaluation metrics

The performance metrics that have been used to measure the ability of the different used models to detect drowsy subjects are explained below.

#### 3.4.1. Confusion matrix

A confusion matrix is a table arrangement that summarizes the prediction results on a classification task. Each line of the table represents the outcomes in an observed class while each column represents the outcomes in a predicted class, or vice versa.

There are four potential values to get in a classification task that can be defined as follows [57]:

- True Positive (TP): The predicted model outcome was true, and the observed value was true.
- False Positive (FP): The predicted model outcome was true, but the observed was false.
- False Negative (FN): The model expected a false outcome, while the observed value was true.
- True Negative (TN): The model predicted outcome and the observed outcome were false.

Confusion matrices can be used to calculate performance metrics for classification models. The most commonly used metrics to judge model performance are accuracy, precision, recall, and F1 score.

#### 3.4.2. Accuracy

Accuracy gives the percentage of the total number of the correctly predicted values. It should be as high as possible. The accuracy calculating formula is:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}. \quad (4)$$

#### 3.4.3. Precision

The precision is the rate of correctly defined positive values. Its formula can be written as:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \quad (5)$$

This metric allows us to calculate the rate of the positive predictions which are actually positive.

#### 3.4.4. Recall

Recall (or sensitivity) is the measure of true positive over the count of actual positive outcomes. It represents the percentage of observed positive cases that are correctly identified. The formula for recall can be expressed as:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (6)$$

Using this formula, we can evaluate our model if it is able to identify the actual true result.

### 3.4.5. F1 Score

The F1 score is the Harmonic Mean of precision and recall. This metric highlights these two factors. The formula for the F1 score can be expressed as:

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

### 3.4.6. AUC-ROC curve

The ROC (Receiver Operating Characteristics) curve is a significant evaluation metric to measure any classification model's performance in DL. This curve plots two parameters:

- True Positive Rate (TPR), which is the recall rate.
- False Positive Rate (FPR) is defined as follows:

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \quad (8)$$

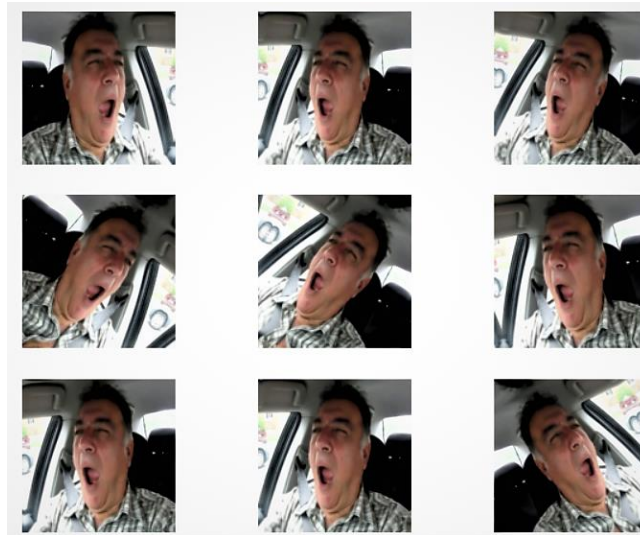
TPR and FPR represent the y-axis and the x-axis of the ROC curve. The region under the curve is known as the Area Under the Curve (AUC). The larger the area, the better the performance.

## 4. Experimental results

As we mentioned before, our DDD system is built based on DNN. Moreover, we tried to achieve the training using the traditional CNN model. A performance comparative analysis of the CNN model with the three DNN models used in the learning module of the DDD system has been performed based on the previous detailed evaluation metrics.

To evaluate the proposed DDD system, we trained the previously described models using the YAWDD dataset [44]. A total of 80% of the dataset was used for training, and only 20% was used for testing. The dataset used in the training phase was different from the one used in the testing phase. Furthermore, data augmentation techniques are used to the training set in order to overcome image transformations and noise. We referred to geometric transformations as data augmentation tools which are: Zooming, flipping and rotation in order to generate new data. The generation of new data was done during the learning step. All the data moved using forward backward propagation. At each iteration, the image passed through the data augmentation layer before reaching the convolution layers of the DL model. Finally, the model improved the accuracy and minimized the error at each iteration.

The data augmentation layer randomly generated a combination of the real image and other different augmented images. It may generate the original image as it is, one of the random modifications, or a combination of 2 or more of them, that guarantees a minimum of multiplication of data by 5 during 20 passages (20 epochs). This action has increased the number of images of the training dataset four times higher. Increasing the diversity of the training set had made an imperative improvement in the DNN testing performances. Figure 10 shows an example of an augmented image.



**Figure 10.** An example of an augmented image.

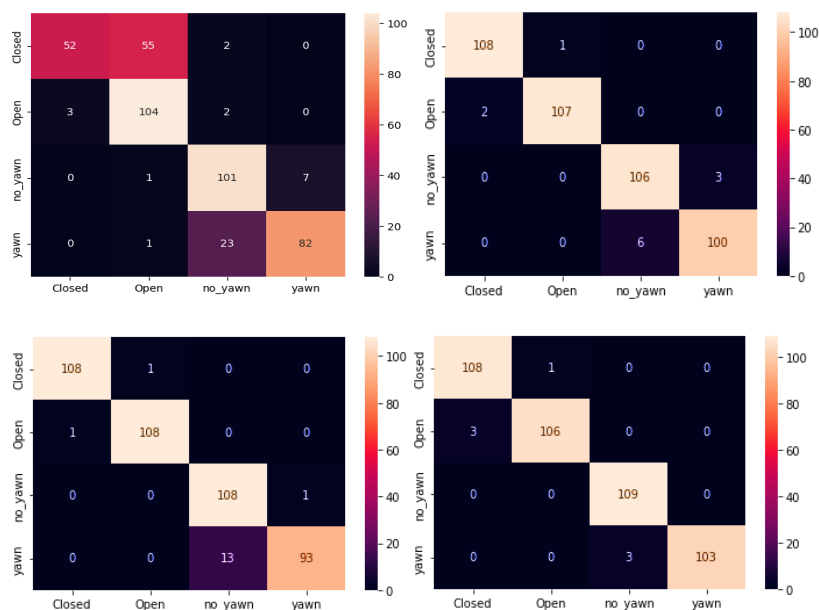
Table 2 lists the number of layers for the various models that are used.

**Table 2.** Number of layers for various used models.

Model	CNN	VGG16	VGG19	ResNet50
Number of Layers	10	16	19	50

These models were developed in open-source language Python using Collab API with all supporting libraries related to computer vision and deep-learning architectures as OpenCV, Keras, and Tensorflow tools on a PC with the following configuration: Intel® Core (TM) 10th generation CPU, 8 Go of RAM, Windows 10, 64 bits and a Web Camera. The total epochs vary from 8 to 50 according to the model. The time processing is therefore different for each model. It increases unless the number of layer increases. However, on average, the DDD system took 0.22 seconds to train a single image for each model.

Figure 11 presents the confusion matrices for the different used models. This figure gives an overview of the attained accuracies.



**Figure 11.** CNN models' confusion matrices.

According to the confusion matrices, the system is able to eliminate all the false positive values. This can consolidate the challenges raised in this work to conceive a reliable system. Table 3 depicts the performance metrics for the CNN model with the different DNN models.

**Table 3.** Performance metrics for various models.

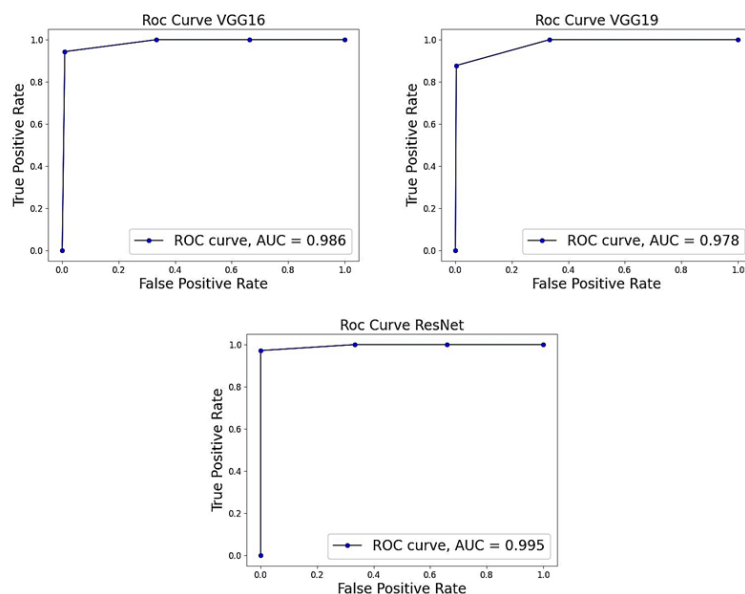
Metric/Model	CNN	VGG16	VGG19	ResNet50
Accuracy	0.8900	0.9722	0.9630	0.9838
Precision	0.8247	0.9724	0.9658	0.9842
Recall	0.7829	0.9720	0.9624	0.9837
F1 Score	0.7740	0.9722	0.9641	0.9839
Time processing(s)	800	128	688	752
Epochs	50	8	43	47

Table 3 reveals that the ResNet50 model achieved the highest values for all metrics. The time processing is as higher as the number of layers increased and it is relative to all hardware and software materials. On other hand, the CNN model gives the lowest values at all. Transfer learning is therefore more suitable to solve the target task.

According to the achieved results, the ResNet50 model is the most efficient DNN model for the drowsiness state classification with a testing accuracy of 98.4%. Furthermore, VGG16 and VGG19 also achieved competitive results, as shown in the table.

The ROC curves corresponding to the DNN models used in our DDD system are presented in Figure 12.





**Figure 12.** The ROC curves related to the DNN models.

The area under the ROC curves of all models confirms that all DNN models are good classifiers. Indeed, all the models present good AUC values, so they have a great ability to recognize the drowsiness state. The ResNet50 model is the most performant model to recognize the drowsiness state.

## 5. Comparison and discussion

Several DDD systems were proposed in the literature. They used many and diverse methods and techniques to conceive their detection procedure. The most popular DDD system is the behavioral parameter-based techniques or also known as imaged based system. This system continuously monitors the drivers' physical behavior especially the facial expressions like eye closure ratio, eye blinking and yawning.

The comparative analysis of the proposed DDD system with such techniques can be performed by comparing the performance's metrics of the mentioned DDD systems in the paper or through other DDD systems that had used the same dataset utilized by our system (the YAWDD dataset).

Table 4 summarizes the DDD systems mentioned in the literature review of this paper with the one proposed.

**Table 4.** DDD systems mentioned in the paper.

<b>Reference</b>	<b>Facial expressions</b>	<b>Methodology</b>	<b>Accuracy</b>
[26]	Eyelid closure	Surveillance videos to check if the driver's eyes were closed. The eyes were located to identify curvature of the eyelids using an extended Sobel operator. Three datasets were used by the system.	95% for the first data set 70% for the second data set 95% for the third data set
[27]	Eye movements	Three ML algorithms were used to identify eye behavior using a simple web camera.	95%
[28]	Eye tracking	Drowsiness detection is based on eye-tracking signals' features. Different lengths of eye tracking epochs were chosen for feature extraction, and the performance of each classifier was evaluated for every epoch length.	RF: 88.37% to 91.18% SVM: 77.12% to 82.62%
[29]	Eye closure	A real-time system based on eye closure using DL. A Fully Designed Neural Network (FD-NN) along with Transfer Learning models (VGG16 and VGG19), were developed to classify the drowsiness state.	FD-NN: 98.15% TL-VGG16: 95.45% TL-VGG19: 95%
[30]	Mouth movements	Features of mouth movements were extracted using an SVM and a method based on the Circular Hough Transform (CHT) for DDD system	81%
[31]	Mouth (yawning)	DDD system based on CNN using depth video sequences to detect driver fatigue during the nighttime.	91.57%
[33]	Eye and mouth states	The eyes and mouths features were set using a Raspberry Pi 3 Model B. A Mamdani fuzzy inference system is used to estimate the driver state.	95.5%
[34]	Eye and mouth states	The Hough transform was employed to locate the mouth and eyes' regions in the face. Then, it is used to assess whether the extracted eye was open and to calculate the mouth opening.	94%
[41]	Eye and mouth states	CNN with VGG16 classification algorithm to detect eye and mouth state.	91%
[42]	Eye and mouth states	Four DL models are used. The output of these models is combined to produce the final output, which is determined by a SoftMax classifier.	85%
[43]	Facial features and scenes conditions	The drowsiness detection used a 3D-deep convolutional neural network. The recognition of driver's drowsiness status was done using the condition-adaptive representation.	76.2
[44]	Facial features	The three networks' outputs are pooled and incorporated into a Softmax classifier.	73.66%
[45]	Visual and non-visual features	A combination of CNNs and deep-belief network (DBN) models was used to detect driver fatigue from both visual and non-visual features.	94.5%
<b>Proposed approach</b>	<b>Eye and mouth state</b>	<b>A classification module includes three DNN to detect the drowsiness state and a prediction module to assess the real drowsy state based on an ensemble learning method.</b>	<b>98%</b>

According to Table 4, the best accuracy is assigned to our DDD system proposed in this paper.

Table 5 makes a comparison in terms of accuracies between DDD systems that used the same dataset utilized in this paper.

**Table 5.** DDD systems that used the YAWDD dataset.

Reference	Methodology	Accuracy
[58]	The Dlib toolkit is used to extract facial feature points from the driver face detected using the improved YOLOv3-tiny CNN. Used the SVM classifier to decide the drowsiness state.	94.32%
[59]	Using CNNs and kernelized correlation filters to detect driver face and to extract face features. Drowsiness is assessed based on eyes and mouth states. The driver is alerted if detected drowsy.	92%
[60]	Using a two-stream spatial-temporal graph convolutional network for videos to detect driver drowsiness.	93.4%
<b>Proposed approach</b>	<b>A classification module includes three DNN to detect the drowsiness state and a prediction module to assess the real drowsy state based on an ensemble learning method.</b>	<b>98%</b>

According to Table 5, all the systems used cameras or images to detect the fatigue and the drowsiness of the driver, and they show valuable accuracies. Besides, the proposed approach overshoots the other approaches that used the YAWDD dataset in terms of accuracy.

To conclude, experiments demonstrate that using car camera and DL technology simultaneously is promoting drowsiness detection. Indeed, the valuable information issued from the car camera can be boosted by DL technologies which capture various observable fatigue and drowsiness characteristics. The experiments also reveal that ensemble learning approaches can significantly increase the DDD system performance that will effectively promote robustness and reliability. Liu et al. [61] confirmed the same idea in their review paper.

The main advantage of the proposed DDD system is that it can be easily industrialized because it is cost-efficient, easy to use, non-invasive system and automatic. However, this system is very dependent on the quality of the image processing which makes the detection of the driver's state highly related to the quality of the image. Moreover, some conditions may affect the DDD systems like wearing sunglasses, sudden changes in lighting, and camera distance from the face. All these factors can reduce the accuracy or even lead to false detection. The major challenge for the DDD system is guaranteeing a high-quality camera and carefully considering environmental factors of the car during system development and testing.

Successfully, our DDD system is an advanced and intelligent system, comparable to other cutting-edge technologies like the Traffic Sign Recognition System (TSRS) [62]. By integrating the DDD system into the Advanced Driver Assistance Systems (ADAS) and/or Automated Driving Systems (ADS) in smart vehicles, we aim to provide valuable assistance to drivers in order to improve their efficiency and safety.

## 6. Conclusions

Drivers' behavior, especially drowsiness, is the major cause of road accidents worldwide. Detecting and modeling the drowsiness state by means of DDD systems makes it possible to alert the driver about a dangerous situation. However, these systems have not yet been operated for many reasons especially inaccessibility and lack of performance. Therefore, it is necessary to build a reliable drowsiness detection system that can correctly and effectively detect real-time drivers' behavior. Based on the eye closure and mouth opening, we can perform a functional DDD system. In this study, we have proposed a hybrid DDD system to detect a drowsy driver in a real-time state. The working process has been divided into an offline process dedicated to the learning module and an online process for the detection and the prediction modules. For the offline training, we have applied data augmentation techniques for the used database to enhance the data. Besides, the pre-trained CNN models used in the learning module have achieved good performances for the classification of the driver's state and the recognition of a drowsy driver in the prediction module. In the online process, we have used transfer learning techniques to evaluate the drowsiness state. The considered system can eliminate false drowsiness alerts and to identify only the real ones. This hybrid approach increases the system performances which are consequently reflected by the obtained results. In the proposed study, one camera could be insufficient in case of fast head rotation, which leads to the vanishing of an eye from the captured image. This can be solved by using extra cameras in the car. Moreover, our approach studies only the features of a driver's face. It does not extract the features from limb movements. Future works should further consider other body parts as indicators.

### Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

### Acknowledgments

This research has been funded by Deputy for Research & Innovation, Ministry of Education through Initiative of Institutional Funding at University of Ha'il – Saudi Arabia through project number IFP-22 122

### Conflict of interest

All the authors declare that there are no potential conflicts of interest and approval of the submission.

### References

1. World Health Organization, Road traffic injuries, 2022. Available from: <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>.
2. S. Singh, Critical reasons for crashes investigated in the national motor vehicle crash causation survey, National Highway Traffic Safety Administration (NHTSA), 2015.

3. M. M. R. Komol, M. M. Hasan, M. Elhenawy, S. Yasmin, M. Masoud, A. Rakotonirainy, Crash severity analysis of vulnerable road users using machine learning, *PLoS ONE*, **16** (2021), e0255828. <https://doi.org/10.1371/journal.pone.0255828>
4. National Safety Council, Drivers are falling asleep behind the wheel. 2022. Available from: <https://www.nsc.org/road/safety-topics/fatigued-driver?>
5. J. M. Owens, T. A. Dingus, F. Guo, Y. Fang, M. Perez, J. McClafferty, et al., *Prevalence of drowsy driving crashes: Estimates from a large-scale naturalistic driving study*, AAA foundation for traffic safety, Washington, 2018.
6. Euro NCAP 2025 Roadmap. Available from: <https://cdn.euroncap.com/media/30700/euroncap-roadmap-2025-v4.pdf>.
7. Official Journal of the European Union, Document 32019R2144—Regulation (EU) 2019/2144 of the European Parliament and of the Council. 2019. Available from: <https://eur-lex.europa.eu/eli/reg/2019/2144/oj>.
8. M. Ramzan, H. U. Khan, S. M Awan. A. Ismail, M. Ilyas, A. Mahmood, A survey on state-of-the-art drowsiness detection techniques, *IEEE Access*, **7** (2019), 61904–61919. <https://doi.org/10.1109/ACCESS.2019.2914373>
9. C. N. Watling, M. M. Hasan, G. S. Larue, Sensitivity and specificity of the driver sleepiness detection methods using physiological signals: A systematic review, *Accident Anal. Prev.*, **150** (2021), 105900. <https://doi.org/10.1016/j.aap.2020.105900>
10. M. M. Hasan, C. N. Watling, G. S. Larue, Physiological signal-based drowsiness detection using machine learning: Singular and hybrid signal approaches, *J. Safety Res.*, **80** (2022), 215–225. <https://doi.org/10.1016/j.jsr.2021.12.001>
11. C. C. Liu, S. G. Hosking, M. G. Lenné, Predicting driver drowsiness using vehicle measures: Recent insights and future challenges, *J. Safety Res.*, **40** (2009), 239–245. <https://doi.org/10.1016/j.jsr.2009.04.005>
12. P. M. Forsman, B. J. Vila, R. A. Short, C. G. Mott, H. P. A. Van Dongen, Efficient driver drowsiness detection at moderate levels of drowsiness, *Accident Anal. Prev.*, **50** (2013), 341–350. <https://doi.org/10.1016/j.aap.2012.05.005>
13. X. Zhang, X. Wang, X. Yang, C. Xu, X. Zhu, J. Wei, Driver drowsiness detection using mixed-effect ordered logit model considering time cumulative effect, *Anal. Methods Accid. Res.*, **26** (2020), 100114. <https://doi.org/10.1016/j.amar.2020.100114>
14. E. Ouabida, A. Essadike, A. Bouzid, Optical correlator based algorithm for driver drowsiness detection, *Optik*, **204** (2020), 164102. <https://doi.org/10.1016/j.ijleo.2019.164102>
15. Y. Sun, P. Yan, Z. Li, J. Zou, D. Hong, Driver fatigue detection system based on colored and infrared eye features fusion, *Comput. Mater. Con.*, **63** (2020), 1563–1574. <https://doi.org/10.32604/cmc.2020.09763>
16. M. K. Kamti, R. Iqbal, Evolution of driver fatigue detection techniques—A review from 2007 to 2021, *Transport. Res. Rec.*, **2676** (2022), 485–507. <https://doi.org/10.1177/03611981221096118>
17. Y. Albadawi, M. Takruri, M. Awad, A review of recent developments in driver drowsiness detection systems, *Sensors*, **22** (2022), 2069. <https://doi.org/10.3390/s22052069>
18. A. A. Bamidele, K. Kamardin, N. S. N. A. Aziz, S. M. Sam, I. S. Ahmed, A. Azizan, et al., Non-intrusive driver drowsiness detection based on face and eye tracking. *Int. J. Adv. Comput. Sci. Appl.*, **10** (2019), 549–569. <https://doi.org/10.14569/IJACSA.2019.0100775>

19. S. T. Lin, Y. Y. Tan, P. Y. Chua, L. K. Tey, C. H. Ang, Perclos threshold for drowsiness detection during real driving. *J. Vision*, **12** (2012), 546. <https://doi.org/10.1167/12.9.546>
20. A. Rosebrock, Eyeblink detection with OpenCV, Python, and Dlib, PyImageSearch, 2017. Available from: <https://pyimagesearch.com/2017/04/24/eye-blink-detection-opencv-python-dlib/>.
21. V. Pradeep, Namratha, T. Nisha, Shravya, M. Vshker, A review on eye aspect ratio technique, *IJAR SCT*, **3** (2023), 98–100. <https://doi.org/10.48175/IJAR SCT-7843>
22. R. C. Chen, C. W. Chang, C. Dewi, Determining the driver's mental state using detecting the eyes, *2022 IET International Conference on Engineering Technologies and Applications (IET-ICETA)*, 2022. <https://doi.org/10.1109/IET-ICETA56553.2022.9971619>
23. S. Thiha, J. Rajasekera, Efficient online engagement analytics algorithm toolkit that can run on edge, *Algorithms*, **16** (2023), 86. <https://doi.org/10.3390/a16020086>
24. A. Moujahid, F. Dornaika, I. Arganda-Carreras, J. Reta, Efficient and compact face descriptor for driver drowsiness detection, *Expert Syst. Appl.*, **168** (2021), 114334. <https://doi.org/10.1016/j.eswa.2020.114334>
25. A. C. Huang, C. Yuan, S. H. Meng, T. J. Huang, Design of fatigue driving behavior detection based on circle hough transform, *Big Data*, **11** (2023), 1–17. <https://doi.org/10.1089/big.2021.0166>
26. M. T. Khan, H. Anwar, F. Ullah, A. Ur Rehman, R. Ullah, A. Iqbal, et al., Smart real-time video surveillance platform for drowsiness detection based on eyelid closure. *Wirel. Commun. Mob. Com.*, **2019** (2019), 2036818. <https://doi.org/10.1155/2019/2036818>
27. C. B. S. Maior, M. J. das Chagas Moura, J. M. M. Santana, I. D. Lins, Real-time classification for autonomous drowsiness detection using eye aspect ratio, *Expert Syst. Appl.*, **158** (2020), 113505. <https://doi.org/10.1016/j.eswa.2020.113505>
28. A. S. Zandi, A. Quddus, L. Prest, F. J. E. Comeau, Non-intrusive detection of drowsy driving based on eye tracking data, *Transport. Res. Rec.*, **2673** (2019), 247–257. <https://doi.org/10.1177/0361198119847985>
29. M. Hashemi, A. Mirrashid, A. B. Shirazi, Driver safety development: Real-time driver drowsiness detection system based on convolutional neural network, *SN Comput. Sci.*, **1** (2020), 289. <https://doi.org/10.1007/s42979-020-00306-9>
30. N. Alioua, A. Amine, M. Rziza, Driver's fatigue detection based on yawning extraction, *Int. J. Veh. Technol.*, **2014** (2014), 678786. <https://doi.org/10.1155/2014/678786>
31. X. Ma, L. P. Chau, K. H. Yap, Depth video-based two-stream convolutional neural networks for driver fatigue detection, *2017 International Conference on Orange Technologies (ICOT)*, 2017. <https://doi.org/10.1109/ICOT.2017.8336111>
32. B. K. Savasx, Y. Becerikli, Real time driver fatigue detection system based on multi-task ConNN, *IEEE Access*, **8** (2020), 12491–12498. <https://doi.org/10.1109/access.2020.2963960>
33. A. Celecia, K. Figueiredo, M. Vellasco, R. González, A portable fuzzy driver drowsiness estimation system, *Sensors*, **20** (2020), 4093. <https://doi.org/10.3390/s20154093>
34. N. Alioua, A. Amine, M. Rziza, D. Aboutajdine, Driver's fatigue and drowsiness detection to reduce traffic accidents on road, In: *Computer analysis of images and patterns*, Berlin, Heidelberg: Springer, 2011. [https://doi.org/10.1007/978-3-642-23678-5\\_47](https://doi.org/10.1007/978-3-642-23678-5_47)
35. L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, et al. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions, *J Big Data*, **8** (2021), 53. <https://doi.org/10.1186/s40537-021-00444-8>

36. Z. Li, F. Liu, W. Yang, S. Peng, J. Zhou, A survey of convolutional neural networks: Analysis, applications, and prospects, *IEEE T. Neur. Net. Lear. Syst.*, **33** (2022), 6999–7019. <https://doi.org/10.1109/TNNLS.2021.3084827>
37. F. Chollet, Transfer learning & fine-tuning, 2020. Available from: [https://keras.io/guides/transfer\\_learning/](https://keras.io/guides/transfer_learning/)
38. E. Magán, M. P. Sesmero, J. M. Alonso-Weber, A. Sanchis, Driver drowsiness detection by applying deep learning techniques to sequences of images, *Appl. Sci.*, **12** (2022), 1145. <https://doi.org/10.3390/app12031145>
39. N. Ho, Y. C. Kim, Evaluation of transfer learning in deep convolutional neural network models for cardiac short axis slice classification, *Sci Rep.*, **11** (2021), 1839. <https://doi.org/10.1038/s41598-021-81525-9>
40. A. Kensert, P. J. Harrison, O. Spjuth, Transfer learning with deep convolutional neural networks for classifying cellular morphological changes, *SLAS Discov.*, **24** (2019), 466–475. <https://doi.org/10.1177/2472555218818756>
41. A. Aytekin, V. Mençik, Detection of driver dynamics with VGG16 model, *Appl. Comput. Syst.*, **27** (2022), 83–88. <https://doi.org/10.2478/acss-2022-0009>
42. M. Dua, Shakshi, R. Singla, S. Raj, A. Jangra, Deep CNN models-based ensemble approach to driver drowsiness detection, *Neural Comput. Applic.*, **33** (2021), 3155–3168. <https://doi.org/10.1007/s00521-020-05209-7>
43. J. Yu, S. Park, S. Lee, M. Jeon, Driver drowsiness detection using condition-adaptive representation learning framework, *IEEE T. Intell. Transp. Syst.*, **20** (2018), 4206–4218. <https://doi.org/10.1109/TITS.2018.2883823>
44. P. Sanghyuk, P. Fei, S. Kang, C. D. Yoo, Driver drowsiness detection system based on feature representation learning using various deep networks. In: *Computer Vision—ACCV 2016 Workshops*. Springer, Cham, 2017. [https://doi.org/10.1007/978-3-319-54526-4\\_12](https://doi.org/10.1007/978-3-319-54526-4_12)
45. Q. Abbas, HybridFatigue: A real-time driver drowsiness detection using hybrid features and transfer learning, *Int. J. Adv. Comput. Sci. Appl.*, **11** (2020), 585–593. <https://doi.org/10.14569/IJACSA.2020.0110173>
46. D. Lee, Which deep learning model can best explain object representations of within-category exemplars? *J Vision*, **21** (2021), 12. <https://doi.org/10.1167/jov.21.10.12>
47. A. Shabnam, M. Omidyeganeh, S. Shirmohammadi, B. Hariri, YawDD: A yawning detection dataset, *MMSys '14: Proceedings of the 5th ACM Multimedia Systems Conference*, 2014, 24–28. <https://doi.org/10.1145/2557642.2563678>
48. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition. *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*.
49. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016. <https://doi.org/10.1109/CVPR.2016.90>
50. Dlib C++ toolkit, Available from: <http://dlib.net/>.
51. P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, 2001, <https://doi.org/10.1109/CVPR.2001.990517>

52. V. Kazemi, J. Sullivan, One millisecond face alignment with an ensemble of regression trees, *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014. <https://doi.org/10.1109/CVPR.2014.241>
53. R. Potolea, S. Cacoveanu, C. Lemnaru, Meta-learning framework for prediction strategy evaluation, *International Conference on Enterprise Information Systems*, Berlin, Heidelberg: Springer, 2011. [https://doi.org/10.1007/978-3-642-19802-1\\_20](https://doi.org/10.1007/978-3-642-19802-1_20)
54. R. Dillmann, J. Beyerer, J.D. Hanebeck, T. Schultz, Advances in artificial intelligence, *Proceedings of the Springer 33rd Annual German Conference on AI. Karlsruhe*, 2010.
55. C. Dewi, R. C. Chen, X. Jiang, H. Yu, Adjusting eye aspect ratio for strong eye blink detection based on facial landmarks, *PeerJ Comput. Sci.*, **8** (2022), e943. <https://doi.org/10.7717/peerj-cs.943>
56. C. Dewi, R. C. Chen, C. W. Chang, S. H. Wu, X. Jiang, H. Yu, Eye aspect ratio for real-time drowsiness detection to improve driver safety, *Electronics*, **11** (2022), 3183. <https://doi.org/10.3390/electronics11193183>
57. N. Kadri, A. Ellouze, M. Ksantini, S. H. Turki, New LSTM deep learning algorithm for driving behavior classification, *Cybernet. Syst.*, **54** (2023), 387–405. <https://doi.org/10.1080/01969722.2022.2059133>
58. F. You, Y. Gong, H. Tu, J. Liang, H. Wang, A fatigue driving detection algorithm based on facial motion information entropy, *J. Adv. Transport*, **2020** (2020), 8851485. <https://doi.org/10.1155/2020/8851485>
59. W. Deng, R. Wu, Real-time driver-drowsiness detection system using facial features, *IEEE Access*, **7** (2019), 118727–118738. <https://doi.org/10.1109/ACCESS.2019.2936663>
60. J. Bai, W. Yu, Z. Xiao, V. Havyarimana, A. C. Regan, H. Jiang, et al., Two-stream spatial-temporal graph convolutional networks for driver drowsiness detection, *IEEE T. Cybernetics*, **52** (2022), 13821–13833. <https://doi.org/10.1109/TCYB.2021.3110813>
61. F. Liu, D. Chen, J. Zhou, F. Xu, A review of driver fatigue detection and its advances on the use of RGB-D camera and deep learning, *Eng. Appl. Artif. Intel.*, **116** (2022), 105399. <https://doi.org/10.1016/j.engappai.2022.105399>
62. N. Triki, M. Karray, M. Ksantini, A real-time traffic sign recognition method using a new attention-based deep convolutional neural network for smart vehicles, *Appl. Sci.*, **13** (2023), 4793. <https://doi.org/10.3390/app13084793>



AIMS Press

© 2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)