Facial Expression-Based Drowsiness Detection System for Driver Safety Using Deep Learning Techniques

Amina Turki[1][a], Sirine Ammar[2], Mohamed Karray[3][b] and Mohamed Ksantini[1][c]
[1]*Control & Energies Management Laboratory (CEM-Lab), National Engineering School of Sfax, University of Sfax, Tunisia*
[2]*National School of Electronics and Telecommunications of Sfax, University of Sfax, Tunisia*
[3]*ESME Research Lab, Special School of Mechanics and Electricity (ESME), Ivry Sur Seine, France*

Keywords: Driver Drowsiness Detection (DDD) System, Deep Neural Networks (DNNs), the Chebyshev Distance.

Abstract: Driver drowsiness is a leading cause of road accidents, resulting in severe physical injuries, fatalities, and substantial economic losses. To address this issue, a sophisticated Driver Drowsiness Detection (DDD) system is needed to alert the driver in case of abnormal behaviour and prevent potential catastrophes. The proposed DDD system calculates the Eyes Closure Ratio (ECR) and Mouth Opening Ratio (MOR) using the Chebyshev distance, instead of the classical Euclidean distance, to model the driver's behaviour and to detect drowsiness states. This system uses simple camera and deep transfer learning techniques to detect the driver's drowsiness state and then alert the driver in real time situations. The system achieves 96% for the VGG19 model, and 98% for the ResNet50 model, with a precision rate of 98% in assessing the driver's dynamics.

1 INTRODUCTION

Drowsiness, often underestimated, is a real danger when related to driving. Driver's fatigue and sleepiness becomes a silent threat, contributing significantly to the alarming statistics of road accidents and fatalities. It is not possible to calculate the exact number of sleep related accidents, but research shows that driver fatigue may be a contributory factor in up to 20% of road accidents, and up to one quarter of fatal and serious accidents (ROSPA, 2020). Indeed, the National Highway Traffic Safety Administration (NHTSA, 2017) reported that drowsy driving was involved in an estimated 91,000 crashes, resulting in 795 deaths and 50,000 injuries in the United States in 2017. It is therefore important to detect drowsiness early and accurately.

Preventing drowsiness while driving is a paramount concern, and the integration of Driver Drowsiness Detection (DDD) systems emerges as a crucial solution. These innovative systems represent a proactive and effective approach to preventing the dangers associated with drowsy driving. By leveraging technology to monitor, alert, and respond to signs of fatigue, these systems play a crucial role in safeguarding lives on the road (Ramzan, 2019).

DDD systems can be broadly categorized into several types, each utilizing various measures to monitor and mitigate the risk of drowsy driving.

The most effective type of Driver Drowsiness Detection (DDD) system depends on various factors, including accuracy, real-time responsiveness, and practical implementation. In practice, a combination of technologies often proves to be the most effective approach (Kamti, 2022). Drowsiness detection systems (DDD) based on facial recognition are a promising approach, especially when combined with deep learning (DL) techniques (Aytekin, 2022), (Dua, 2021), (Ahmed, 2023), and (Yu, 2018).

This paper focuses on studying DDD systems based on facial expressions. It proposes a hybrid drowsiness detection system (DDD) that combines eye closure ratio (ECR) and mouth opening ratio (MOR) features extracted from car camera images of

[a] https://orcid.org/0000-0002-4314-3541
[b] https://orcid.org/0000-0001-7293-8696
[c] https://orcid.org/0000-0002-9928-8643

Turki, A., Ammar, S., Karray, M. and Ksantini, M.
Facial Expression-Based Drowsiness Detection System for Driver Safety Using Deep Learning Techniques.
DOI: 10.5220/0012386000003636

In *Proceedings of the 16th International Conference on Agents and Artificial Intelligence (ICAART 2024) - Volume 3*, pages 726-733
ISBN: 978-989-758-680-4; ISSN: 2184-433X

the driver's face using Machine learning (ML) techniques. These features are then used to train classifiers using Deep learning (DL) models to distinguish between drowsy and non-drowsy drivers.

The system first detects the driver's facial landmarks in a frame using image recognition. Then, it calculates the ECR and MOR using the Chebyshev distance, which has been shown to be more accurate than other distance measures. The driver's drowsiness state is then detected by the trained model based on these values. Finally, an ensemble learning methods were used to determine whether the driver is tired. So, the paper is organized as follows: Section 2 discusses concepts related to the proposed DDD system and reviews related research studies. Section 3 introduces the proposed approach, methodology, and materials. Section 4 presents the experimental results and discussions. Finally, section 5 concludes the paper.

2 RELATED WORK

In this work, we focus on the study of DDD systems based on facial expressions measures.

2.1 Facial Expressions' Behavioural Measures for DDD Systems

The features of the driver's physical behaviour represent a good baseline to detect more efficiently the driver's drowsiness. There are many DDD systems which are based on facial expressions. They use many and diverse parameters and methods to conceive their detection procedure.

2.1.1 Eyes' Facial Expressions

The eye state is a relevant method for detecting driver drowsiness (Wilkinson, 2013). Various features like the eye-opening rate, eyelid distance, and PERCLOS are considered top indicators of drowsiness (Wilkinson, 2013). Khan et al. developed a real-time Driver Drowsiness Detection (DDD) system that utilized eyelid closure as a key indicator (Tayab Khan, 2019). The system used surveillance videos to monitor the driver's eyes and classified the eyelids as open or closed based on the curvature of the eyelids. Maior et al. created a sleepiness detection technique using the eyes' movements, calculating the EAR metric to determine whether the eye is open or closed (Marior, 2020). Zandi et al. proposed the use of eye tracking data as a non-intrusive measure for detecting drowsiness, achieving an accuracy of 88.37% to 91.18% with the RF classifier (Zandi, 2019). Hashemi et al. developed a real-time DDD system based on eye closure using deep learning, achieving an accuracy of 98.15% with the FD-NN model (Hashemi, 2020).

2.1.2 Mouth' Facial Expressions

In various studies, the real-time prediction of driver drowsiness has been achieved by analyzing the state of the driver's mouth. Alioua et al. utilized an SVM and the Circular Hough Transform (CHT) to extract features from mouth movements for their DDD system, which proved effective in real-time scenarios across different lighting conditions (Alioua, 2014). The experiment's results indicated that yawning could be detected with an accuracy rate of 81%. Similarly, Xiaoxi et al. developed a DDD system based on CNNs that utilized depth video sequences to detect driver fatigue specifically during nighttime (Xiaoxi, 2017). By employing both spatial and temporal CNNs, the system was able to locate objects and calculate motion vectors, enabling the detection of yawns even when the driver's mouth was covered. The system demonstrated an accuracy of 91.57% in their experiments.

2.1.3 Hybrid Facial Expressions: Eyes and Mouth

In recent studies on Driver Drowsiness Detection (DDD) systems, researchers have explored various approaches to analyze driver behavior. Celecia et al. proposed an economical and accurate DDD system (Celecia, 2020). The system recorded images using a camera with an infrared illuminator and employed a Raspberry Pi 3 Model B for processing. Features from the eyes and mouth were extracted using a cascade of regression tree algorithms. These features were then combined using a Mamdani fuzzy inference system to predict the driver's drowsiness state. The system achieved a high accuracy of 95.5% and remained resilient to various ambient illumination conditions.

Alioua et al. presented a non-intrusive and efficient method for detecting drowsiness (Alioua, 2011). Their approach involved analyzing closed eyelid and open mouth states based on images captured from a webcam. The system used an SVM face detector to identify the face region in each image and applied the Hough transform to locate the mouth and eyes' regions. By assessing the openness of the eye and calculating the mouth opening, the system determined the driver's drowsiness with an accuracy of 94% and an 86% kappa statistic value.

2.2 Deep Learning for DDD Systems

DL is a significant research trend within the Machine Learning (ML) community, known for its remarkable success in various domains. DL networks possess the ability to learn from vast amounts of data, enabling exceptional performance in complex cognitive tasks. Convolutional Neural Networks (CNNs) are a prominent type of DL network. CNNs excel at automatic pattern detection and feature extraction in images, without requiring human guidance. This capability has led to the widespread adoption of CNNs, making them one of the most popular DL networks architectures.

A CNN architecture is represented in Figure 1.

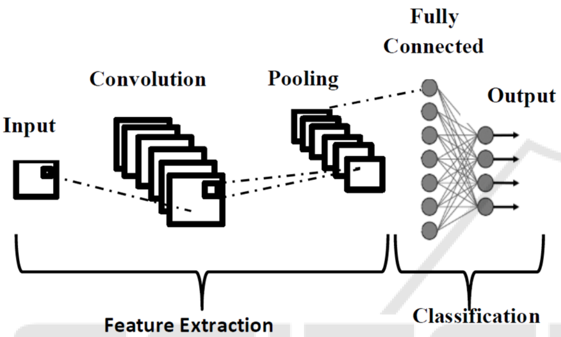

Figure 1: A CNN architecture.

There is a wide range of pre-trained models available for deep learning tasks, such as Inception, VGG family, and ResNet family. Transfer Learning (TL) is a technique that utilizes pre-trained CNN models to solve different tasks within a similar domain (Transfer, 2021). TL saves both resources and time, as it does not require extensive amounts of data or starting the training process from scratch (Ho, 2021). The use of pre-trained structures improves generalization even after fine-tuning to the specific dataset (Kensert, 2019). Several studies have utilized CNNs for drowsy driver detection. The study in (Aytekin, 2022), used a VGG16 model that achieved an accuracy of 91% and an F1-score of over 90% for each class in determining if the driver's eyes are open or closed and if they are yawning. Another study suggested an architecture of four DL models that use RGB videos of drivers as input. It had employed DL models and ensemble processes to detect tiredness, achieving accuracy rates of 85% the with a SoftMax classifier in the output (Dua, 2021). Yu et al. (Yu, 2018) proposed a framework for the DDD based on 3D-deep CNN. The recognition of driver's drowsiness status was done using the condition adaptive representation with an accuracy of 76,2 %.

3 PROPOSED APPROACH

3.1 Description

We present in this section a DDD system that utilizes pretrained CNNs with TL techniques to detect driver drowsiness in various driving scenarios. The proposed approach offers several key contributions:

- Introduction of a novel DL approach that automatically detects and estimates driver drowsiness using camera and deep TL methods.
- Utilization of the Chebyshev distance to analyze the state of the driver's eyes and mouth (open or closed) based on facial landmarks, enabling efficient drowsiness detection.
- Implementation of data augmentation techniques to magnify and enrich the dataset, thereby enhancing the training process.
- Classification of drowsiness states using two pretrained CNN models, resulting in improved performance of the DDD system.
- Utilization of ensemble learning techniques to combine the model outputs and generate the final prediction, ensuring better recognition performance.

3.2 The Learning Procedure

The learning procedure consists of training two CNN models; the VGG19 and the Res-Net50. These models represent the most object identification accuracies (Lee, 2021). They will be used later to decide if the driver is drowsy or not for a real-time detected drowsiness state. The overall procedure is represented by Figure 2.

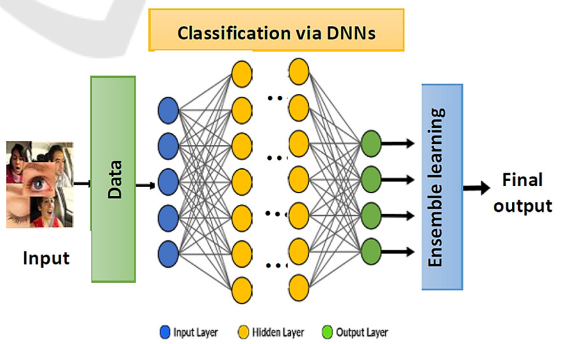

Figure 2: The learning procedure.

3.2.1 Dataset

The study used the YAWDD dataset (Shabnam, 2014), consisting of 2900 samples of facial features of 322 male and female drivers' videos that were taken in real and varying illumination conditions with different mouth conditions such as normal, talking, singing, and yawning, as well as drivers wear glasses. These samples are mainly used for models and algorithms to classify driver drowsiness. The dataset was divided into four categories: yawn, no-yawn, open eye, and closed eye.

3.2.2 Data Augmentation

Data augmentation techniques are used to increase the quantity of training data DNNs to perform complex tasks with high accuracy. These techniques involve artificially increasing the quantity of data by producing new data points from available data. This was achieved by making small alterations to image data, such as geometric and color transformations, to the original data.

3.2.3 Training

The study focuses on training two CNN models, VGG19 and Res-Net50, to determine if a driver is drowsy in real-time. The models were chosen for their accuracies in object identification and their ability to learn hierarchical representations of visual data. The pre-trained layers of VGG19 and ResNet50 were frozen to preserve their learning features. To adapt to the specific drowsiness state classification task, additional fully connected layers were added to learn high-level features. The models were then compiled for training using the Adam optimizer and sparse categorical cross-entropy loss function. The training process involved many iterations. The performance of each model was evaluated on the validation set, comparing predictions with ground truth labels to measure their accuracy and effectiveness in recognizing different drowsiness states.

3.2.4 Ensemble Learning

This research utilized ensemble learning, a widely recognized and effective machine learning technique, to improve classification performance in drowsiness states. Three distinct ensemble methods were implemented: Ensemble Averaging, Ensemble Stacking, and AdaBoost Ensemble.

- Ensemble Averaging combined predictions from the VGG19 and ResNet50, to derive a final prediction, improving recognition performance. Each model contributed equally to the ensemble's decision, leveraging their strengths and distinctive capabilities.
- Ensemble Stacking introduced a meta-model designed to harness the predictive abilities of individual models, concatenating predictions from both models and feeding them into a densely constructed meta-model. This meta-model aimed to explore higher-order interactions between the models, enhancing performance beyond what each model could achieve independently.
- In AdaBoost Ensemble, individual models were used as base estimators. The meta-model combined the output of these models through weighted averaging, giving more weight to models that performed well and less weight to those with lower accuracy. This process not only enhanced overall performance but also provided a mechanism to adaptively focus on the strengths of specific models.

3.3 The Detection Procedure

To detect driver drowsiness, a basic car camera is installed on the vehicle's roof. The camera captures live video and identifies the driver's face region. Using the Dlib toolkit, the eyes and mouth landmarks are determined. The coordinates of these landmarks are then used to calculate the ECR and the MOR. By analyzing these ratios, the system can identify if the driver's eyes are closed or if they are yawning, indicating a drowsy state.

3.3.1 Identification of Facial Landmarks

The Dlib library (Dlib, 2022) which is an open-source library utilizing C++ language, is used to identify the essential features of the driver's face in the driver video frame by frame. This library provides a facial landmark detector that estimates the positions of 68 face-specific coordinate points, including the eyes, eyebrows, nose, ears, and mouth. The technique for detecting these facial landmarks is based on machine learning algorithms proposed by Viola and Jones (Viola, 2001) and further improved by Kazemi et al. (Kazemi, 2014). The Dlib package offers an efficient solution for real-time facial features detection, enabling accurate identification of the driver's facial landmarks. This face landmarks detector identifies 68 main facial features positions, as shown in Figure 3.

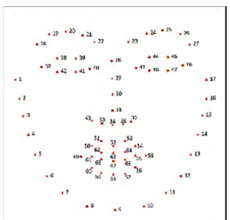

Figure 3: The 68 facial landmark points of human face.

We can detect and access specific facial structures by using the facial landmark index, which identifies sections of the face. Through this method, we can easily extract information from the eye and mouth regions: the right eye: (36, 42), the left eye:(42, 48), and the mouth: (49, 68).

In our study, we utilized a set of 32 facial landmarks, focusing on the left eye, right eye, and mouth regions, to determine the level of eye closure and mouth opening. We employed two distance metrics, namely the Euclidean distance and the Chebyshev distance, to calculate ECR and the MOR. Our findings revealed that the Chebyshev distance outperformed the Euclidean distance, making it the preferred choice for our analysis.

The Chebyshev distance is particularly advantageous in situations where implementation speed is crucial, as it enables faster computation of pixel distances. This distance metric is commonly used in specialized applications where execution speed is of utmost importance (Potolea, 2010).

$$D\ (x, y) = \max_{i}(|x_i - y_i|) \tag{1}$$

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

3.3.2 Eye Closure Ratio (ECR)

ECR is a scalar value that responds to the estimation of the eye closure state. Each eye is represented by six coordinates, as shown in Figure 4.

ECR value is calculated by using the following equation:

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

3.3.3 Mouth Opening Ratio (MOR):

Yawning is marked by mouth opening as shown in Figure 5. A parameter used to determine whether someone is yawning. Like ECR, MOR is defined as:

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

Figure 4: The facial landmarks related to eyes (p_1-p_6).

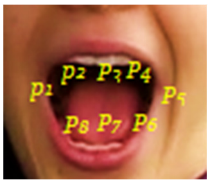

Figure 5: Mouth yawning with facial landmarks (p_1-p_6).

3.3.4 Drowsiness Detection

To detect a drowsy driver, certain conditions need to be met:

1) The driver is considered drowsy if the output of the detector module exceeds a specified drowsiness threshold, typically ranging from 0 to 1. In our case, we have set the threshold at 0.3 after conducting multiple tests.
2) Drowsiness is determined by the ECR, which measures the duration of eye blinks. On average, a blink lasts between 0.1 and 0.4 seconds. If the ECR exceeds this range, indicating prolonged eye closure for more than five seconds, the person is considered drowsy.
3) Drowsiness is also identified by the MOR. When the MOR reaches its maximum value, it indicates yawning, a common sign of drowsiness.
4) If both condition 1 and condition 3 are met simultaneously, with the output exceeding the drowsiness threshold and the MOR indicating yawning, the driver is deemed drowsy.

4 EXPERIMENTAL RESULTS

The DDD system is built based on two DNNs. Furthermore, we tried to achieve the training using the traditional CNN model. A performance comparative analysis of the CNN model with these DNNs models used in the learning module of the DDD system has been performed.

Two DNN models were trained using the YawDD dataset. The dataset was split, with 80% used for training and 20% for testing. Both contain data from the same persons. Data augmentation techniques were applied to the training set. Geometric transformations such as zooming, flipping, and rotation were used to generate new data during the learning step. The generated data was passed through the data augmentation layer before reaching the convolution layers of the DL model.

The trained models were developed in open-source language Python using Collab API with all sup-porting 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, Winodws 10, 64 bits and a Web Camera. The total epochs vary from 43 to 47 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.

Table 1: Performance metrics for models.

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

Table 1 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. According to the achieved results, the CNN model gives the lowest values at all. TL is therefore more suitable to solve the target task. The ResNet50 model is the most efficient CNN model for the drowsiness state classification with a testing accuracy of 98.4%. Figure 6 presents the confusion matrices for the different used CNN models.

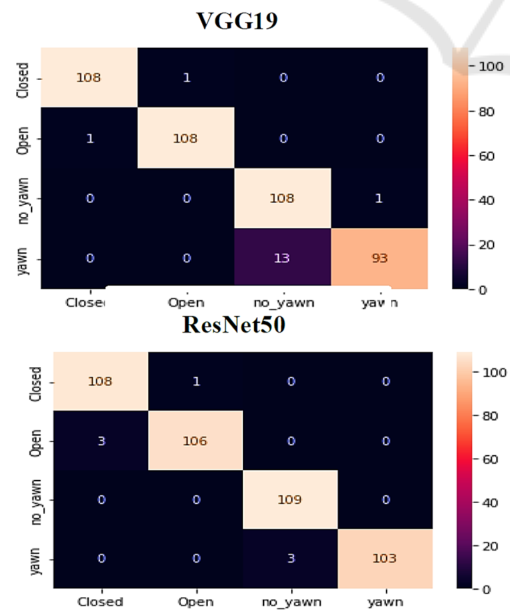

Figure 6: CNN models confusion matrices.

The ROC curves corresponding to the used CNN models are presented in Figure 7.

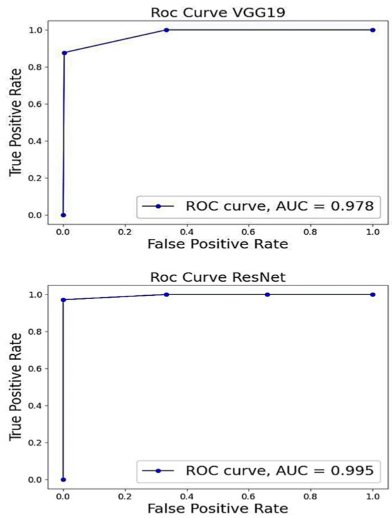

Figure 7: The ROC curves related to the DNNs.

These figures confirms that all models are good classifiers.

To ensure a high-performance DDD system, an ensemble learning approach based on three ensembles methods is implemented to combine the outputs of the models and accurately determine the driver's state. If the driver is confirmed as drowsy, an alarm is triggered. Each ensemble method was rigorously evaluated to assess their effectiveness in improving recognition of drowsiness states using metrics such as accuracy, precision, and confusion matrices. Table 1 depicts the performance metrics for the obtained models. Table 2 shows the performance metrics for the used ensemble methods.

Table 2: Performance metrics for ensemble methods.

Ensemble method	Ensemble Averaging	Ensemble Stacking	AdaBoost Ensemble
Accuracy	0.89	0.94	0.98
Precision	0.92	0.95	0.98
Recall	0.89	0.94	0.98
F1 Score	0.88	0.94	0.98

Table 2 provides a comprehensive overview of performance metrics for ensemble methods, with ensemble Averaging achieves a precision of 0.92, indicating 92% correct positive predictions. It identifies 89% of all actual positive cases with a recall of 0.89. The F1-Score of 0.88 balances precision and recall, indicating a well-balanced model. Ensemble Stacking performs even better with a precision of 0.95, indicating a high proportion of correct positive predictions. It also has a recall of 0.94, indicating strong performance in identifying positive cases. The F1-Score of 0.94 signifies a well-balanced model, accurately classifying 94% of the data. AdaBoost Ensemble outperforms the others with a precision of 0.98, indicating extremely accurate positive predictions and a recall of 0.98, identifying almost all positive cases.

The experiments conducted in this study show that combining car cameras with DL technology is highly beneficial for drowsiness detection. DL algorithms can effectively capture and analyze various drowsiness characteristics from the images captured by the car camera, enhancing the accuracy and effectiveness of the drowsiness detection system. Additionally, the experiments demonstrate that using ensemble learning approaches can greatly improve the performance of the DDD system. Ensemble learning techniques enhance the robustness and reliability of the system, making it more effective in detecting and preventing drowsy driving incidents.

5 COMPARISONS

Numerous DDD systems have been suggested in the literature, employing a wide range of methods and techniques to formulate their detection procedures. Among these, the behavioral parameter-based techniques, also known as image-based systems, have gained significant popularity. These systems focus particularly on facial expressions such as eye closure, eye blinking, and yawning. To conduct a comparative analysis of the proposed DDD system with these techniques, we assessed the performance metrics of the aforementioned DDD systems mentioned in the paper. Table 3 reviews the DDD systems mentioned in this paper with the proposed one.

Table 3: Performance metrics for ensemble methods.

Facial expressions	Reference	Accuracy
Based on eye state	(Tayab Khan, 2019)	95% for the first data set 70% for the second data set 95% for the third data set
	(Marior, 2020)	95%
	(Zandi, 2019)	88.37% to 91.18% with the RF classifier
	(Hashemi, 2020)	98.15% with the FD-NN model
Based on mouth state	(Alioua, 2014)	81%
	(Xiaoxi, 2017)	91.57%
Based on eye and mouth states	(Celecia, 2020)	95.5%
	(Alioua, 2011)	94%
	The proposed approach	**98%**

According to table 3, the best accuracy is assigned to our DDD system proposed in this paper.

The proposed DDD system offers several advantages that make it suitable for industrialization. However, the accuracy of driver state detection in this system heavily relies on the quality of image processing. Various factors such as wearing sunglasses, sudden changes in lighting, and the distance between the camera and the driver's face can affect the system's performance, potentially leading to reduced accuracy or false detections. Despite these challenges, our DDD system is highly advanced and comparable to other state-of-the-art technologies like the Traffic Sign Recognition System (TSRS) (Triki, 2023). The DDD system can be integrated into Advanced Driver Assistance Systems (ADAS) and/or Automated Driving Systems (ADS) in smart vehicles.

6 CONCLUSIONS

The major cause of road accidents worldwide is drivers' behavior, particularly drowsiness. To address this issue, DDD systems have been developed to detect and model the drowsiness state, allowing for timely alerts to drivers in dangerous situations. However, these systems face challenges such as inaccessibility and lack of performance. Therefore, there is a need to build a reliable drowsiness detection system that can accurately and effectively detect drivers' behavior in real-time. By analysing eye closure and mouth opening, we have proposed a functional DDD system to detect a drowsy driver in a real-time state. The working process has been divided into learning process and detection process. For the training, we have applied data augmentation techniques for the used database to enhance the training data. Additionally, the DNN models utilized for learning displayed promising results for classifying the driver's state and identifying drowsiness. Moreover, ensemble learning techniques were employed to assess the drowsiness state.

The proposed DDD system is cost-efficient, easy to use, non-invasive, and automatic, which makes it suitable for industrial applications. However, to ensure a high-quality camera and account for environmental factors during system development and testing, careful consideration is necessary.

REFERENCES

Ahmed, M.I.B.; Alabdulkarem, H.; Alomair, F.; Aldossary, D.; Alahmari, M.; Alhumaidan, M.; Alrassan, S.; Rahman, A.; Youldash, M.; Zaman, G. (2023). A Deep-Learning Approach to Driver Drowsiness Detection. Safety, 9, 65. https://doi.org/10.3390/safety9030065

Alioua, N., Amine, A., Rziza, M., Aboutajdine, D. (2011). Driver's fatigue and drowsiness detection to reduce traffic accidents on road. *In Proceedings of the International Conference on Computer Analysis of Images and Patterns,* Seville, Spain, 29–31 August 2011.

Alioua, N., Amine, A., Rziza, M. (2014). Driver's Fatigue Detection Based on Yawning Extraction. Int. J. Veh. Technol. https://doi.org/10.1155/2014/678786

Aytekin, A., Mençik, V. (2022). Detection of Driver Dynamics with VGG16 Model. Appl. Comput. Inform. 27, 83-88. https://doi.org/10.2478/acss-2022-0009

Celecia, A., Figueiredo, K., Vellasco, M., González, R. (2020). A portable fuzzy driver drowsiness estimation system. Sensors, 20, 4093. https://doi.org/10.3390/s20154093

Dlib C++ toolkit. Available online: http://dlib.net/ (accessed on 08 Mai 2022).

Dua, M., Shakshi, Singla, R., et al. (2021). Deep CNN models-based ensemble approach to driver drowsiness detection. Neural Comput & Applic. 33, 3155–3168. https://doi.org/10.1007/s00521-020-05209-7

Hashemi, M., Mirrashid, A., Shirazi, A.B. (2020). Driver Safety Development: Real-Time Driver Drowsiness Detection System Based on Convolutional Neural Network. SN Comput. Sci. 1, 1–10.

Ho, N., Kim, YC. (2021). Evaluation of transfer learning in deep convolutional neural network models for cardiac short axis slice classification. Sci Rep. 11, 1839. https://doi.org/10.1038/s41598-021-81525-9 https://doi.org/10.1007/s42979-020-00306-9

Kamti, M. K.; Iqbal, R. (2022). Evolution of Driver Fatigue Detection Techniques-A Review From 2007 to 2021. Transp. Res. Rec., 2676, 485–507. https://doi.org/10.1177/03611981221096118

Kazemi, V., Sullivan, J. (2014). One millisecond face alignment with an ensemble of regression trees. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23-28 June 2014. https://doi.org/10.1109/CVPR.2014.241

Kensert, A., Harrison, P.J., Spjuth, O. (2019). Transfer Learning with Deep Convolutional Neural Networks for Classifying Cellular Morphological Changes. SLAS Discov. 24, 466-475. https://doi.org/10.1177/2472555218818756

Lee, D. (2021). Which deep learning model can best explain object representations of within-category exemplars? J Vis. 1;21(10):12. https://doi.org/10.1167/jov.21.10.12

Marior, C.B.S., das Chagas Moura, M.J., Santana, J.M.M., Lins, I.D. (2020). Real-time classification for autonomous drowsiness de-tection using eye aspect ratio. Expert Syst. Appl. 158, 113505. https://doi.org/10.1016/j.eswa.2020.113505

NHTSA. (2017). "Traffic safety facts 2015."

Potolea, R., Cacoveanu, S., Lemnaru, C. (2010). Meta-learning Framework for Prediction Strategy Evaluation. In Proceedings of the International Conference on Enterprise Information Systems, Funchal-Madeira, Portugal, 8–12 June 2010.

Ramzan, M., Khan, H.U., Awan, S.M., Ismail, A., Ilyas, M., Mahmood, A. (2019). A Survey on State-of-the-Art Drowsiness Detection Techniques. IEEE Access. 7. https://doi.org/61904-61919

ROSPA: The Royal Society for the Prevention of Accidents (2020), *Driver Fatigue and Road Accidents Factsheet.*

Shabnam, A., Mona, O., Shervin, S., Behnoosh, H. (2014). YawDD: A yawning detection dataset. In Proceedings of the 5th ACM Multimedia Systems Conference, Singapore, 19 March 2014. https://doi.org/10.1145/2557642.2563678

Tayab Khan, M., Anwar, H., Ullah, F., Ur Rehman, A., Ullah, R., Iqbal, A., Lee, B.H., Kwak, K.S. (2019). Smart real-time video surveillance platform for drowsiness detection based on eyelid closure. Wirel. Commun. Mob. Comput. 1–9. https://doi.org/10.1155/2019/2036818

Transfer Learning & Fine-Tuning. Available online: https://keras.io/guides/transfer_learning/ (accessed on 20 August 2021).

Triki, N., Karray, M., Ksantini, M. (2023). A Real-Time Traffic Sign Recognition Method Using a New Attention-Based Deep Convolutional Neural Network for Smart Vehicles. Appl. Sci. 13, 4793. https://doi.org/10.3390/app13084793

Viola, P., Jones, M. (2011). Rapid object detection using a boosted cascade of simple features. In Proceedings of the IEEE Computer Society Conference. Kauai, HI, USA, 8-14 December 2001.

Wilkinson, VE., Jackson, ML., Westlake, J, Stevens, B, Barnes, M, Swann, P, Rajaratnam, S.M, Howard ME. (2013). The accuracy of eyelid movement parameters for drowsiness detection. J Clin Sleep Med. 15; 9(12):1315-24. https://doi.org/10.5664/jcsm.3278

Xiaoxi, M., Chau, L.P., Yap, K.H. (2017). Depth video-based two-stream convolutional neural networks for driver fatigue detection. In Proceedings of the 2017 International Conference on Orange Technologies (ICOT), Singapore, 8–10 December 2017.

Yu, J., Park, S., Lee, S., Jeon, M. (2018). Driver drowsiness detection using condition-adaptive representation learning framework. *IEEE Trans. Intell. Transp. Syst.* 20,4206–4218. https://doi.org/10.48550/arXiv.1910.09722

Zandi, A.S., Quddus, A., Prest, L., Comeau, F.J. (2019). Non-intrusive detection of drowsy driving based on eye tracking data. *Transp. Res. Rec.* 2673, 247–257. https://doi.org/10.1177/0361198119847985 .