Research Progress on the Development of Fatigue Driving Detection Based on Deep Learning

Zhibin Guan

School of Software Engineer, Sichuan University, Sichuan, China

zhibin032@gmail.com

Abstract. Inattentiveness, weariness, and drowsiness are prevalent causes of car accidents; however, timely warnings can prevent these dangerous situations. Advances in artificial intelligence (AI) and computer vision have made it feasible to continuously monitor a driver's condition and alert them when their attention starts to wane. AI technology can assess a driver's level of fatigue by analyzing various facial expressions such as yawning, the closing of eyes, and head movements. Additionally, it can collect data on vehicle behavior and the driver's biological signals. The primary metric for evaluating the effectiveness of these systems is detection accuracy, which indicates how reliably the system can identify signs of driver fatigue. This study provides a comprehensive review of recent techniques for detecting fatigue driving, with a strong focus on deep learning methodologies. It highlights the application and reliability of these advanced technologies in real-world scenarios. Furthermore, the study delves into the latest issues and challenges faced in this field, including limitations in current technology and potential obstacles to widespread adoption. Finally, it proposes suggestions and outlines future research directions to enhance the effectiveness and reliability of fatigue detection systems. This expanded analysis aims to contribute to the development of safer driving environments by leveraging cutting-edge AI and computer vision technologies.

Keywords: Fatigue driving, deep learning, vehicle accidents, neural networks, image detection.

1. Introduction

An important field in image processing and human activity recognition is drowsiness detection. A person's reaction time is greatly slowed down during the period between wakefulness and sleep, which is commonly referred to as drowsiness. The main causes of fatal traffic accidents include driver negligence, fatigue, and sleepiness. These elements have a big effect on road safety because they cause drivers to pay far less attention to the road and have less control over their vehicles. Driving while weary or drowsy is the practice of operating a vehicle while experiencing exhaustion or sleepiness.

Fortunately, it is now possible to detect signs of driver fatigue and drowsiness at an early stage and alert them, which can significantly reduce the occurrence of accidents. This is achieved by identifying various symptoms of drowsiness, such as frequent eye closures and prolonged yawning. Depending on the area of detection, these signs of fatigue can be categorized into four main types. These signs can be detected through images or video sequences captured by cameras monitoring the driver's facial expressions.

Furthermore, biological signals can be recorded by sensors attached to the driver, and the vehicle's movement and behavior can be monitored. For the same objective, hybrid approaches are also employed. Based on how they are used, Ramzan et al. categorized the state-of-the-art driver drowsiness detection (DDD) methods in a thorough analysis [1]. There exist three primary categories of these techniques: physiological signal-based, vehicle behavior-based, and behavioral parameter-based. The authors examined the most effective supervised methods for identifying sleepiness and talked about the benefits and drawbacks of the three methods they offered. A review research on the identification of driver weariness and drowsiness was carried out by Sikander et al. [2]. There are five types of DDD techniques: hybrid, vehicle-based, physical, biological, and subjective reporting. A study by Otmani et al. used deep learning methods to measure drowsiness and weariness, including recurrent neural networks (RNN) and convolutional neural networks (CNN) [3].
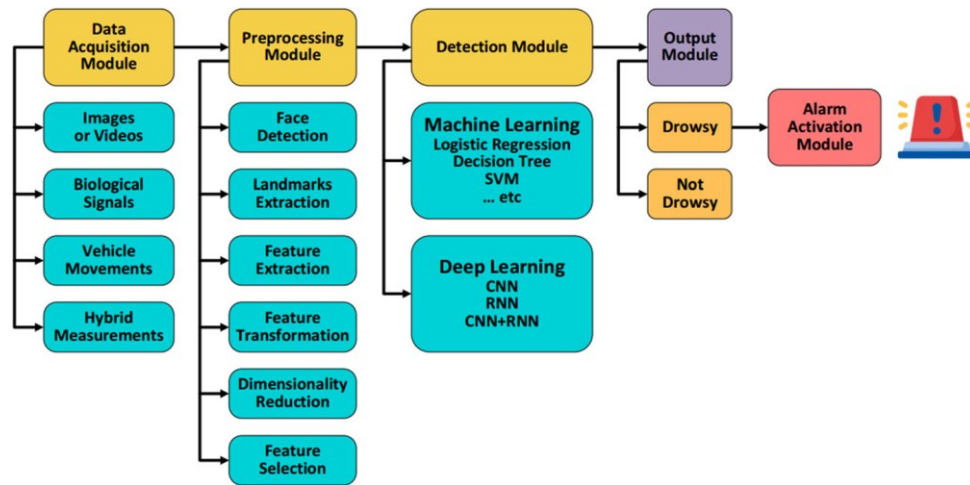
This paper makes significant contributions to the driver fatigue detection systems introduced in recent years, particularly those modern systems, as detailed below:

- Introduces modern deep learning techniques used to detect driver drowsiness for ensuring safe driving.
- Classifies driver fatigue detection systems into four groups based on the tools used to assess drowsiness: eye and facial detection, mouth detection, head detection, and hybrid detection.
- Focuses on the datasets primarily used for detecting drowsiness and introduces their characteristics.
- Compares the applicability and dependability of the four categories for detecting driver weariness.
- explains the drawbacks of drowsiness detection methods, their implications, and the challenges facing the field of detecting driver weariness.

2. Driver Drowsiness Detection (DDD) Systems

When a person is mentally fatigued, they may fall asleep and their attention to tasks decreases. It is crucial to understand that fatigue is not a disease but a state that can be restored through rest and sleep. However, when driving vehicles that require continuous attention, such as cars, buses, or trains, this recovery can become very difficult, potentially leading to serious accidents. Currently, numerous studies are exploring the physiological mechanisms of drowsiness and assessing the degree of fatigue. The literature confirms that there are numerous elements that influence weariness and drowsiness/lethargy in different ways, making it difficult to determine the exact degree of these states. As shown in Figure 1, in order to effectively and reliably depict the driver's condition, fatigue monitoring requires a system that can gather data from various sources and integrate it. Driver fatigue detection systems include several stages:

a. Data Collection: Data is acquired through convenient sensing devices, known as input modules.
b. Preprocessing Module: This module prepares the data by applying landmarks, dimensionality reduction, face detection, feature extraction, feature transformation, and feature selection.
c. Detection or Classification: Relies on machine learning and deep learning techniques for detection or classification.
d. Output Module: This module determines the driver's condition and indicates whether or not they are tired.
e. Alarm Activation Module: Activates an alarm when drowsiness is detected. As shown in Figure 1, this series of processes works in coordination to ensure the driver's safety.

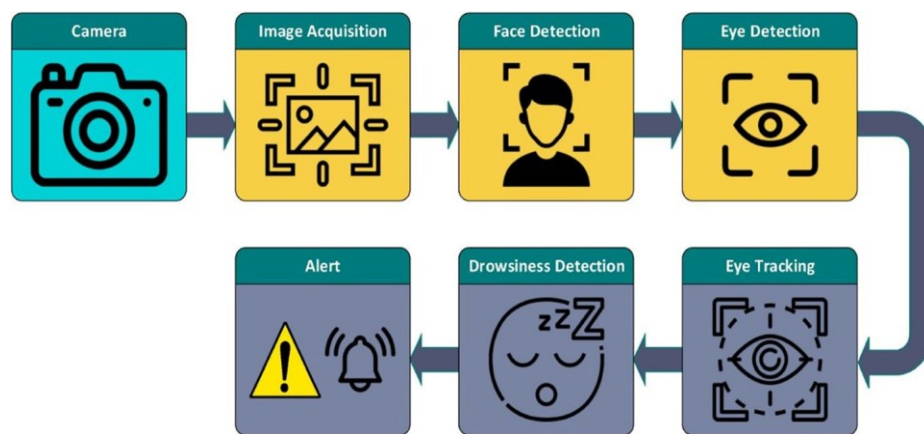**Figure 1.** Data flow and general block diagram in DDD systems [4]

2.1. Eye and Facial Detection

For many computer vision applications, including human-computer interface (HCI), facial emotion recognition (FER), and fatigue and sleepiness monitoring, automatic classification of eye states is essential. These applications are mentioned in Figure 2 and Table 2.

Saurav S. et al. concentrated on creating and implementing a vision-based real-time eye state identification system for embedded platforms with limited resources in order to overcome these problems [5]. To extract meaningful data from the ocular regions, the system makes use of two lightweight CNNs. Fine-tuned transfer learning (TL) was used to train the CNNs using a small sample eye state database without overfitting.

60-second picture sequences of the face were employed by Magán E. et al. They created two distinct techniques to determine whether the driver is sleepy in order to reduce false positives [6]. While the second method uses DL techniques to extract numerical features that are then fed into a fuzzy logic-based system, the first strategy relies on RNN and CNN.

Dua M. et al. introduced a driver fatigue detection system that utilizes a structure comprising four deep learning models: VGG-FaceNet, AlexNet, FlowImageNet, and ResNet. To identify drowsiness and weariness, they used RGB videos of the driver as input [7]. They took into account four main types of features while applying these techniques: hand and head movements, behavioral traits, and facial expressions.

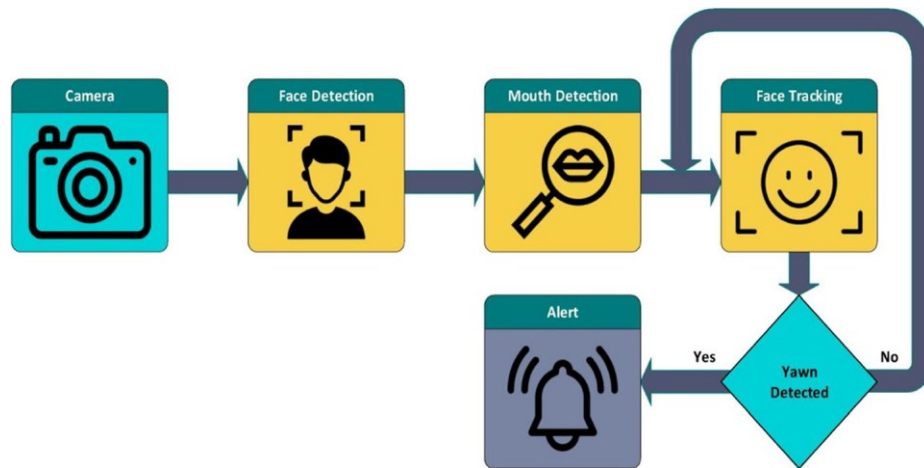**Figure 2.** Procedure for observing blinking of the eyes [4]

2.2. Mouth (Yawning) Detection

The several methods for determining tiredness based on mouth and yawning analysis are shown in Table 2 and Figure 3, along with the benefits and drawbacks of each method.

The goal of the work by Savaş BK et al. was to develop a yawning-based driver detection system that uses a CNN model for classification to detect fatigue [8]. The datasets YawDD, Nthu-DDD, and KouBM-DFD were used to train and evaluate the suggested model.

An enhanced algorithm that can recognize the mouth and face was used in the work by Omidyeganeh M. et al. To determine yawning, the back-projection technique was used to quantify the rate of mouth alterations [9]. Congivue's intelligent embedded Apex camera was employed in this project, fully using the high computational power and low memory of the embedded platform. It could only identify specific alertness levels, though.

In the study by Knapik M. et al., a yawning detection fatigue system based on thermal imaging was described. This device can function without disturbing the driver during the day or night because it makes use of thermal imaging [10]. Facial alignment is first accomplished by identifying the corners of the eyes, and yawning is then detected using the suggested thermal model.
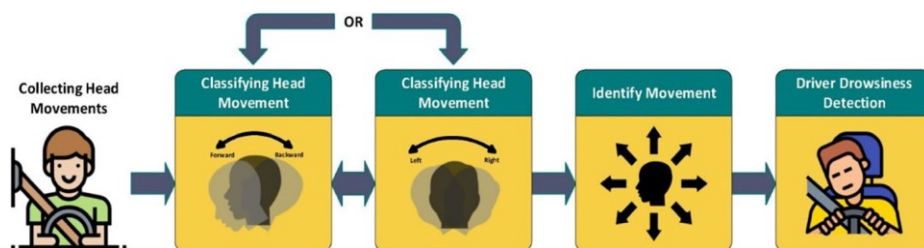

Figure 3. Procedure for observing mouth and yawn [4]

2.3. Head Detection

Numerous techniques have been researched in recent years to detect driver drowsiness. The primary method used by these tactics to identify fatigue is driving behavior observation. Disturbances in a driver's driving style are frequently used to detect driver weariness. Numerous authors cited in Table 2 use gaze and head position variables to detect driver weariness, as illustrated in Figure 4.

Even when the driver wears sunglasses that cover their eyes, the study by Wijnands JS et al. provides an example explaining how to balance real-time inference requirements and good prediction accuracy using deeply separated 3D convolutions and early spatial and temporal feature fusion [11].

Figure 4. Head pose and gaze observation procedure [4]

2.4. Hybrid Eye, Mouth, and Head Detection
Table 2 presents driver fatigue detection systems that rely on and incorporate hybrid architectures with parameters from eye, mouth, and head movements.

De Lima Medeiros PA et al. used volunteers' blinking as a human-computer interaction interface and developed an advanced computer vision detector capable of processing information acquired by a standard camera in real-time [12]. Blink detection was carried out using the following processes as an improved technique for eye state recognition: face recognition, modeling, region of interest (ROI) extraction, and eye state categorization. Moving average filter, ROI evaluator, and rotation compensator were other components of this method. Moving average filter, ROI evaluator, and rotation compensator were other components of this method. Two other datasets were also generated: the Autonomous Blinking Dataset (ABD) and the YouTube Eye State Classification (YEC) dataset, which was produced by taking facial photos out of the AVSpeech collection. CNN and SVM models were trained on the YEC dataset, and experiments were carried out on multiple datasets, such as CEW, ZJU, Talking Face, Eyeblink, and ABD, to assess the algorithms' performance.

Moujahid A. et al. introduced a facial monitoring method that relies on compact facial texture descriptors, which can capture the most apparent drowsiness features [13]. This compactness was further improved by applying a feature selection procedure to the initial recovery features and using a multi-scale pyramid facial representation with fundamental sections comprising both local and global information. The model was separated into four stages: (i) facial description utilizing a multi-level multi-scale hierarchy for feature extraction; (ii) pyramid multi-level (PML) facial modeling; (iv) feature subset selection and classification. In the study by Ed-Doughmi Y. et al., RNNs were used to evaluate a sequence of driver facial images and predict the driver's fatigue status. The model was designed and evaluated using a dataset to detect driver drowsiness. A multi-layer model based on a recurrent neural network architecture that relies on 3D convolutional networks was constructed [14].

3. Results Analysis

3.1. Analysis of Commonly Used Datasets
Table 1 lists the most commonly used datasets in driver fatigue detection systems, providing descriptions for each. Various datasets have been utilized in multiple driver fatigue detection systems.

Table 1. Description of datasets

Ref/Year	Dataset Name	Description
[15], 2016	NTHUDDD public dataset	This dataset includes 36 people from different ethnic backgrounds who participated in a range of driving simulation scenarios wearing and not wearing glasses or sunglasses. These situations include everyday driving, yawning, blinking slowly, falling asleep, laughing fits, and more, and they are performed in both day and nighttime scenarios. • There are five distinct scenarios and eighteen participants in the training dataset. Over the course of about a minute, the slow blink rate, nodding signals, and yawning of each participant are recorded. • Each of the two most important scenarios is demonstrated by a 1.5-minute sequence: one shows a mix of non-drowsiness-related behaviors (such as talking, laughing, and looking in both directions), and the other a combination of drowsiness-related indicators (like nodding, yawning, and slow blinking).
[16], 2020	ZJU gallery	NeuralBody introduced a multi-view dataset known as LightStage. A multi-camera system with more than 20 synchronized cameras was used to record a variety of dynamic human movies, which is how this dataset was created. Driving data was evaluated in both daytime and nighttime scenarios.
[17], 2020	YawDD	An in-car camera was used to record a number of movies featuring actual drivers with a variety of facial features, including men, women, people of different ethnic backgrounds, and even robots. Talking, singing, yawning, and other actions were among the activities. The videos were split up into the following two sets: • The camera was installed beneath the front mirror of the vehicle in the first batch. Three different scenarios are covered in this collection of 322 videos: normal driving (no talking), talking or singing, and yawning while driving. There are three or four videos in each scenario.
[18], 2015	CelebA dataset	There are 200,000 celebrity photos in the CelebA collection, and each one has 40 attributes added to it. With 202,599 face photos, 40 binary feature labels per image, 5 landmark locations, and 10,177 identities, this extensive collection of facial attributes—also referred to as CelebFaces—offers a broad range of attribute annotations.

3.2. Evaluation Metrics

Various metrics have been adopted to evaluate the system's ability to recognize drowsy and fatigued participants, including precision, accuracy, F1 score, and sensitivity. Equation (1) provides the accuracy evaluation metric. Accuracy is the most widely used metric for evaluating driver fatigue detection systems. It provides a good indicator of the system's ability to distinguish between true positives (TP) and true negatives (TN). True positives, true negatives, false positives, and false negatives are denoted by the letters TP, TN, FP, and FN, in that order.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{1}$$

3.3. Experimental Results Analysis

The work on driver drowsiness detection (DDD) that combines deep learning (DL) and machine learning (ML) is shown in Table 2. covering most algorithms based on eye, facial, mouth, and head measurements.

The table clearly indicates the source, year, measurement area, algorithm classification, implementation environment, accuracy, and dataset for each algorithm.

Following data analysis with the NTHUDDD public dataset shown in Table 2, the following conclusions can be drawn:

The accuracy of deep learning is superior to that of traditional machine learning. This is because deep learning has more parameters, allowing it to learn from more data and features, thus enhancing its learning performance significantly compared to machine learning.

Measuring mouth features plays a crucial role in improving the model's accuracy. Yawning is a prominent and critical feature of driver fatigue, serving as an important criterion for determining whether a driver is fatigued. Therefore, mouth measurement holds significant weight in DDD systems.

Table 2. Image and video-based DDD systems

Ref, Year	Image and video parameters	Classification methods	Accuracy	Datasets
[5], 2022	Eye and Face	Dual CNN Ensemble (DCNNE)	CEW:97.56% ZJU:97.99% MRL:98.98%	CEW, ZJU, MRL
[7], 2021	Eye and Face	Deep-CNN-based ensemble	85%	NTHU-DDD video dataset
[4], 2019	Mouth and Eye	Multiple CNN-kernelized correlation filters method	92%	CelebA dataset, YawDD dataset
[6], 2022	Face	RNN and CNN	60%	UTA-RLDD dataset
[8], 2021	Mouth	ConNN model	99.35%	YawDD, Nthu-DDD and KouBM-DFD datasets
[9], 2016	Mouth	SVM Viola-Jones	75%	Self-prepared dataset
[10], 2019	Mouth	Cold and hot voxels	Cold voxels: 71% Hot voxels: 87%	Self-prepared dataset
[19], 2019	Eye and mouth	Hybrid structure of CNN and LSTM	84.45%	NTHUDDD public dataset
[11], 2020	Facial features, and head movements	3D CNN	73.9%	NTHUDDD public dataset
[12], 2022	Eye, head, and mouth	CNN and SVM	97.44%	new datasets were created: YEC and ABD
[13], 2021	Eye, head, and mouth	SVM	79.84%	NTHUDDD public dataset
[14], 2020	Eye, head, and mouth	Multi-layer model based on 3D convolutional networks	97.3%	NTHUDDD public dataset

4. Discussion

In-depth literature research has revealed various strategies for detecting driver fatigue and mitigating potential issues. With advancements in technology and ongoing research in the field of artificial intelligence, the performance of these systems has significantly improved, successfully addressing many

challenges. Considering factors such as invasiveness, interference, cost, ease of use, and accuracy, the practicality of these systems highlights their capability to accurately identify actual levels of drowsiness.

4.1. Image-Based Measurement Methods

Image-based measurement methods are considered practical because they are non-invasive, non-interfering, cost-effective, and do not require sensor setup each time the system is used. Common tools include webcams, smartphone cameras, and thermal cameras. Cameras that gather data are positioned at a specific distance from the driver in order to prevent obscuring their view.

4.2. Relevant Parameters and Accuracy

These systems are dependent on factors associated with weariness, like eye closures, head motions, and yawning. Their accuracy typically ranges from 80% to 99%. However, as mentioned earlier, these systems are influenced by various parameters. Typically, controlled conditions are used for their construction and evaluation, or video records of drivers detecting weariness are used. Therefore, the response to fatigue and its symptoms is mainly simulated, with drivers being prompted to mimic certain symptoms during the data collection phase. In simulated test environments or when using actual databases, the accuracy of image-based systems can exceed 90%. When tiredness manifests visually, it is easy to identify.

4.3. Advantages of Hybrid Systems

Overall, using an appropriate fatigue detection system is the most important step in minimizing car accidents caused by exhaustion. Furthermore, hybrid systems are seen to be the best option for detecting tiredness because of their outstanding track record of dependability.

5. Conclusion

This work offers a thorough analysis and current assessment of systems for detecting driver fatigue, with an emphasis on research findings from 2012 to 2024. Given that this field directly impacts human safety, extensive research is currently being conducted. Given the direct impact on human safety, extensive research is currently being conducted in this field. Driver carelessness or inattention is the primary cause of most traffic accidents. In this review, a comprehensive classification of driver fatigue detection systems is provided, discussing four fundamental methods based on different features for identifying fatigue and drowsiness. These four methods are: eye or facial detection, mouth detection, head detection, and hybrid detection systems. This study presents a comprehensive review of all the systems that are involved, along with the datasets that are utilized, the deep learning models that are implemented, the parameters that are used, and the accuracy of each system's detection. Additionally, typical datasets used for detecting driver fatigue are introduced. Finally, the paper proposes new research directions and potential solutions to address the current issues. Through continuous innovation and improvement, driver fatigue detection systems will be able to more effectively prevent fatigue-related road accidents, ensuring the safety of human lives.

References

[1] 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, 61904-61919. https://doi.org/10.1109/ACCESS. 2019.2917046

[2] Sikander, G., & Anwar, S. (2018). Driver fatigue detection systems: a review. IEEE Transactions on Intelligent Transportation Systems, 20(6), 2339-2352. https://doi.org/10.1109/TITS.2018.2878884

[3] Otmani, S., Pebayle, T., Roge, J., & Muzet, A. (2005). Effect of driving duration and partial sleep deprivation on subsequent alertness and performance of car drivers. Physiology & Behavior, 84(5), 715-724. https://doi.org/10.1016/j.physbeh.2005.02.021

[4] Pandey, N. N., & Muppalaneni, N. B. (2022). A survey on visual and non-visual features in Driver's drowsiness detection. Multimedia Tools and Applications, 1-41. https://doi.org/10.1007/s11042-021-11606-1

[5] Saurav, S., Gidde, P., Saini, R., & Singh, S. (2022). Real-time eye state recognition using dual convolutional neural network ensemble. Journal of Real-Time Image Processing, 1-16. https://doi.org/10.1007/s11554-021-01120-4

[6] Magán, E., Sesmero, M. P., Alonso-Weber, J. M., & Sanchis, A. (2022). Driver drowsiness detection by applying deep learning techniques to sequences of images. Applied Sciences, 12(3), 1145. https://doi. org/10.3390/app12031145

[7] Dua, M., Singla, R., Raj, S., & Jangra, A. (2021). Deep CNN models-based ensemble approach to driver drowsiness detection. Neural Computing and Applications, 33(8), 3155-3168. https://doi.org/10.1007/s00521-020-05448-8

[8] Savaş, B. K., & Becerikli, Y. (2021). A deep learning approach to driver fatigue detection via mouth state analyses and yawning detection. https://doi.org/10.1109/ACCESS.2021.3095074

[9] Omidyeganeh, M., Shirmohammadi, S., Abtahi, S., Khurshid, A., Farhan, M., Scharcanski, J., et al. (2016). Yawning detection using embedded smart cameras. IEEE Transactions on Instrumentation and Measurement, 65(3), 570-582. https://doi.org/10.1109/TIM. 2015.2504018

[10] Knapik, M., & Cyganek, B. (2019). Driver's fatigue recognition based on yawn detection in thermal images. Neurocomputing, 338, 274-292. https://doi.org/10.1016/j.neucom. 2019.02.048

[11] Wijnands, J. S., Thompson, J., Nice, K. A., Aschwanden, G. D., & Stevenson, M. (2020). Real-time monitoring of driver drowsiness on mobile platforms using 3D neural networks. Neural Computing and Applications, 32(13), 9731-9743. https://doi.org/10.1007/s00521-019-04194-3

[12] de Lima Medeiros, P. A., da Silva, G. V. S., dos Santos Fernandes, F. R., Sánchez-Gendriz, I., Lins, H. W. C., da Silva Barros, D. M., et al. (2022). Efficient machine learning approach for volunteer eye-blink detection in real-time using webcam. Expert Systems with Applications, 188, 116073. https://doi.org/10.1016/j.eswa.2021.116073

[13] Moujahid, A., Dornaika, F., Arganda-Carreras, I., & Reta, J. (2021). Efficient and compact face descriptor for driver drowsiness detection. Expert Systems with Applications, 168, 114334. https://doi.org/ 10.1016/j.eswa.2020.114334

[14] Ed-Doughmi, Y., Idrissi, N., & Hbali, Y. (2020). Real-time system for driver fatigue detection based on a recurrent neuronal network. Journal of Imaging, 6(3), 8. https://doi.org/10.3390/jimaging6030008

[15] Weng, C. H., Lai, Y. H., & Lai, S. H. (2016). Driver drowsiness detection via a hierarchical temporal deep belief network. In Asian Conference on Computer Vision (pp. 117-133). Springer, Cham. https://doi.org/ 10.1007/978-3-319-54187-7_8

[16] Zhou, K., Wang, K., & Yang, K. (2020). A robust monocular depth estimation framework based on light-weight erf-pspnet for day-night driving scenes. Journal of Physics: Conference Series, 1518(1), 012051. https://doi.org/10.1088/1742-6596/1518/1/012051

[17] Abtahi, S., Omidyeganeh, M., Shirmohammadi, S., & Hariri, B. (2014). YawDD: a yawning detection dataset. In Proceedings of the 5th ACM Multimedia Systems Conference (pp. 24-28). https://doi.org/10. 1145/2557642.2563671

[18] Liu, Z., Luo, P., Wang, X., & Tang, X. (2015). Deep learning face attributes in the wild. In Proceedings of the IEEE International Conference on Computer Vision (pp. 3730-3738). https://doi.org/10.1109/ICCV. 2015.425

[19] Guo, J. M., & Markoni, H. (2019). Driver drowsiness detection using hybrid convolutional neural network and long short-term memory. Multimedia Tools and Applications, 78(20), 29059-29087. https://doi.org/10.1 007/s11042-018-7016-1