IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received July 18, 2021, accepted August 9, 2021, date of publication August 27, 2021, date of current version September 9, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3108561

Drivers Fatigue Level Prediction Using Facial, and Head Behavior Information

HAIDER A. KASSEM, MORSHED CHOWDHURY, AND JEMAL H. ABAWAJY
School of Information Technology, Deakin University, Geelong, VIC 3220, Australia
Corresponding author: Haider A. Kassem (hakas@Deakin.edu.au)

ABSTRACT With driver fatigue continues to cause serious and deadly car and motorcycles accidents, the need for automatically recognizing driver fatigue and alerting the drivers is apparent. Although various approaches that explore physiological and physical factors to classify driver fatigue have been developed, the overall accuracy, recognition speed, distraction in the driving process and the cost of these systems still need to be improved. In this paper, we present a low-cost driver fatigue level prediction framework **(DFLP)** for detecting driver fatigue in its earliest stages. DFLP predicts drive fatigue based on eyes, mouth, and head behavior cues using a non-physical contact sensor input (infrared radiation) (**IR**) camera. DFLP classifies the level of drowsiness and attributes the level of altering accordingly. To validate the proposed fatigue prediction framework, we conducted the experiment using real datasets under night and day illumination conditions. The results of the experiment show that the proposed approach can predict the level of driver's fatigue with 93.3% overall accuracy. The solution proposed in this paper, not only reduces the number of drivers fatigue-related accidents but also addressed an area of sufficient interest for transportation, psychology and public health experts and readers as well as automakers to develop an in-vehicle fatigue prediction system.

INDEX TERMS Drivers fatigue prediction, visual features, fatigue physiological features, fatigue level prediction.

I. INTRODUCTION

Driver fatigue is one of the major contributors to serious and deadly road accidents worldwide. According to the Transport department of New South Wales state in Australia, fatigue has the same level of danger as drink driving [1]. Furthermore, a study conducted by the Adelaide Centre for Sleep Research has shown that drivers who have been awake for 24 hours have an equal driving performance to a person who has a blood alcohol content (**BAC**) [2] of 0.1 g/100ml and is seven times more likely to have an accident [3]. Fatigue can often affect the driving ability of the drivers long before the drivers' even notice, furthermore, there is no standard rule to measure the level of drivers fatigue other than paying attention to fatigue signs [4]. Fatigue associated with accidents is often more severe than others because driver's reaction times are delayed, or the drivers have failed to make any maneuvers to avoid a crash. With unsafe and dangerous driving accounts for the death of more than one million lives and over 50 million serious injuries worldwide each year [5], the need to address this problem is quite obvious.

In order to reduce fatal road crashes caused by fatigued drivers, a variety of systems that actively monitor drivers physiological factors such as heart rate **(HR)** [6] brain waves [7], electroencephalogram **(EEG)** [8] have been developed. Also, approaches that are based on physical factors such as driver position [9], electromyography **(EMG)** [10], and related image data [11] have also been put forward. Other systems that focus on operating parameters such as the strength of the pedal on the brake or accelerator [12] have also been suggested. Although excellent advances in driver fatigue detection have been made, they tend to be intrusive, which creates an obstacle and distraction in the driving process. Also, the overall accuracy is largely dependent on the driving conditions such as time of the day (i.e., daytime or nighttime) and the weather (e.g., clear, cloudy, rainy). The detection speed may also be delayed due to numerous data collection and analysis as well as reconciliation of data in

The associate editor coordinating the review of this manuscript and approving it for publication was Ahmed Farouk.

cases where collected data exhibits differences. Last but not the least, the cost of these systems still needs to be improved. Therefore, there is a need to develop driver fatigue systems with high accuracy, affordability, and less intrusive to reduce fatigue-related road fatalities.

In this paper, we develop an approach for driver's fatigue level prediction framework **(DFLP)** for detecting driver fatigue at an earlier stage. The approach is based on using facial and head behavior information collected using standard computer vision techniques [4]. The proposed approach is low-cost and non-invasive as it uses an infrared camera that will not be attached to the driver's body. Furthermore, the proposed approach does not have the complexity of setting up the physical contact sensors.

The rest of the paper is organized as follows: Section II overviews some related literature research about identifying drivers' fatigue, which is followed by the detailed explanation of the methods used in the DFLP framework in Section III. The different features used to determine the level of driver's fatigue are detailed in Section IV, which is followed by the fatigue classification model in Section V. The performance analysis, experiment setup and results are presented in Section VI. Finally, the conclusion and future directions are discussed in Section VII.

II. RELATED WORK

Driver's fatigue has been an interesting subject of research within the scientific and commercial arenas. According to [13] the review of previous related studies, there are three categories of methods that have been used for the analysis of driver fatigue. The first category involves the use of physiological data like heart rate **(HR)** [6], brain waves [7], electroencephalogram **(EEG)** [8]. The second category involves the use of physical data such as driver position [9], electromyography **(EMG)** [10], and related image data [11]. The third category includes operating parameters such as the strength of the pedal on the brake or accelerator.

A. IDENTIFYING DRIVERS' FATIGUE USING PHYSIOLOGICAL AND PHYSICAL DATA

The work of [13] proposed and designed a Driver Fatigue Prediction Monitoring System, which uses different sensors such as Percent eye-closure over a fixed time window **(PERCLOS)** and Heart rate variability **(HRV)** to detect the driver's level of fatigue. Their experimental results showed that for HRV data, using Back Propagation Neural Networks **(BPNN)** is superior to Long short-term memory **(LSTM)**, because of the true positive rate **(TPR)** and accuracy of BPNN is 24% and 17% more accurate than that of the LSTM, respectively. On the other hand, for PERCLOS data, LSTM is capable of improving the TPR by 14% as compared to the BPNN. The method proposed in their study is capable of detecting fatigue a one-time step ahead, or 3 to 5 minutes prior to drowsiness. In conclusion, the authors proposed to focus on designing a method with a longer prediction time of fatigue and higher accuracy for different ages and genders in their future work. In addition to this, they presented a method of heterogeneous sensor data fusion that can increase the reliability of data and make the input method of the driver fatigue prediction module easier. In another study [14], the authors suggested using physiological and physical information for fatigue recognition. They recorded the time of driving and how the drivers behave while driving as well as the driving performance. artificial neural network **(ANN)** models were used to test the most suitable feature to be used for the purpose of detecting and predicting drivers' fatigue. The results concluded that driver behavior information comes first before the physiological followed by the driving behavior to obtain the best performance. However, the shortcoming of this work is the difficulty to record the head and eyelid movement in a real car, driving behavior and driving performance-based model should be improved.

B. IDENTIFYING DRIVERS' FATIGUE USING PHYSIOLOGICAL DATA

In another study, Hajinoroozi, Zhang [15] introduced deep covariance learning methods with the aim of classifying the mental states of drivers (alert or drowsy). They used EEG epochs data for the experiment. Their experimental results revealed that the performance recorded for all the deep covariance learning methods was better than that of 244 shallow learning methods like spatial-temporal convolutional neural networks **(STCNN)** and Riemannian methods. Among the deep covariance learning methods, the Convolutional Neural Network **(CNN)** model performed best in terms of classification, and it was applied on the spatial sample covariance matrices using EEG epochs. In addition, the CNN model improved the Area under the curve **(AUC)** of the best shallow algorithm (logistic regression + Log-Euclidean Metric) by 12.32% from 70.96% to 86.14%.

Liu, *et al.* [16] explained that the brain dynamics of humans are mercurial and time-variant, and as such, extreme fluctuations are experienced in the degree of EEG signals during driving. To this end, researchers have made efforts to improve EEG signals for real-life applications. To achieve this, they presented a real-time brain-computer interfaces **(BCI)** system that uses an online algorithm for the prediction of fatigue. The BCI system was designed with pre-event EEG that allows the evaluation of the mental state of a driver when driving. Furthermore, in their study, the memory capacity for adaptive noise cancellation of the proposed system was increased using recurrent self-evolving fuzzy neural work **(RSEFNN)**. The results of their experiment revealed that the proposed system performed remarkably well in terms of predicting drivers' fatigue while driving.

In [17] Hajinoorzi, *et al.* proposed new deep learning **(DL)** algorithms like channel-wise convolutional neural network **(CCNN)** and CCNN-R (which is a CCNN variation that uses Restricted Boltzmann Machine to replace the convolutional filter) to enable the prediction of a driver's state (alert or drowsy). In the algorithms, raw EEG data characteristics are used together with the CCNN and CCNN-R classifiers to achieve better and robust performance than DNN and CNN alongside other non-DL algorithms that use raw EEG data. The authors who investigated the performance of the proposed deep learning and bagging classifier also compared it with the performance of the CNN, DNN, CCNN-R and CCNN in terms of classification. Raw EEG data and Independent Component Analysis **(ICA)**-transformed EEG features were considered as inputs for different methods. An evaluation of cross-session prediction was carried out for each of the cases. The case is simulated by a cross-session prediction which involves using the data of a session that is not the same as the training epoch for prediction. The performance was evaluated using cross-validation. For the training and validation, a total of seventy sessions were used, out of which ten sessions were used for validation and training of the hyper-parameters, and the other sixty sessions, the classifier is trained using fifty sessions, and then the classifier is tested using ten sessions. The performance is measured using the area under the curve or the Az score. Based on the results obtained, the CNN demonstrated great potentials for use in another EEG-based signal classification.

Similarly, in another work [18], the authors employed the Riemannian measures in machine/deep learning models for the classification of EEG data of the drivers in order to determine the driver's state. The study also involved spatial and temporal covariance matrices calculated from EEG epochs. By using classification and visualization methods, it was perceived that the temporal covariance matrices are more discriminative than spatial covariance matrices. The results of their experiments showed that the use of the temporal covariance matrices with the shallow Riemannian measures produces better performance than that of the Euclidian measures by 3.16%. The authors investigated the deep Riemannian model (SPDNet), for classification by using both temporal and spatial covariance matrices. The experimental results showed that when the SPDNet is used with temporal covariance matrices, it produces better performance than Euclidian measures. The result also showed that the SPDNet improves the classification performance for prediction by 6.02% and 2.86% compared to Euclidian and shallow Riemannian methods, respectively. Their result demonstrated that the novel deep covariance learning model has great potential for robust prediction of driver's state using EEG, and as such efforts should be channelled towards improving the model.

C. IDENTIFYING DRIVERS' FATIGUE USING PHYSICAL DATA

Cheng, *et al.* [19] carried out an investigation on the mental fatigue of a driver, using the camera only data, including Eye Aspect Ratio **(EAR)** and Mouth Aspect Ratio **(MAR)**. They attempted to assess the mental fatigue of a driver using Facial Landmark Sequences. They conducted an experiment where 21 students from the University of China have participated. The experiment involved the collection of features from the 21 participants at intervals of 1 minute. In addition, the authors used four machine-learning algorithms and compared the performance of the algorithm. They proposed an image-based method, and the method was found to have many benefits with a comparatively high level of accuracy at a very low cost, thereby contributing to the generalization and commercialization of the method. The method demonstrated the possibility of having high processing speed, given the development of computing software and hardware. With high processing capacity, the proposed method will be able to meet the real-time requirement in the future. The experimental results showed that the collected feature could be used as indicators for fatigue detection.

Just like in cheng *et al.*, the authors in [20] introduced a new method for the recognition of fatigue in drivers. The proposed method uses the eyes to determine fatigue in drivers. In their work, they combined the convolutional layers with LSTM units to design an end-to-end trainable network. The LSTM units were used due to their ability to learn spatial representations and model temporal dynamics. In their study, infrared videos were captured using an infrared camera with filters that were fixed towards the driver's face. The purpose of doing this was to reduce the effect of illumination. The experimental data used in this study is the image of the eye area obtained by using Multi-task Cascaded Convolutional Neural Networks **(MTCNN)**.

The authors used a residual learning module with SE block and transfer learning [21] with the goal of accelerating convergence and improving the accuracy of classification in their designed network. They were able to achieve video-level prediction of the feature in the time and space domain. The effectiveness of their architecture was proven through comprehensive experiments, which revealed that the performance of the architecture in terms of sleepiness detection is remarkable, without much input pre-processing and manually designed features. Nevertheless, it was observed that the clarity of a driver's eyes is affected by reflective spots when he/she is putting on eyeglasses. It was also observed that the performance of the fatigue detection model is directly affected by the accuracy of eye positioning, and as such must be improved. In conclusion, they proposed to incorporate yawn detection into the proposed model to enhance the performance in their future work.

There are other methods to self-assess fatigue such as the Fatigue Assessment Scale **(FAS)** [22], which consist of a one paper questioner that includes 10 items scale evaluating symptoms of chronic fatigue. However, this method does not provide features to be captured, rather, it contains a self-evaluation of all of the fatigue aspects (physical and mental symptoms).

In summary, the literature review revealed how promising results in terms of driver fatigue detection can be obtained using different features and different methods (see Table 1 for a summary of the related work). However, the most reliable existing method of predicting divers' fatigue is that which uses physiological data or combines physiological and physical data. In this regard, the review of the literature revealed a gap, which is the need for the attachment of sensors on

TABLE 1. A summary of reviewed related work research.

Article	Objective	Method	Intrusive & Condition			Dataset			Experiments	
			Intrusive	Day	Night	Type	Features / Variables used	Data Size / No. of participants	Environment	Performance
[13]	Early prediction of fatigue	Back-Propagation Neural Network (BPNN)	✓	✓	×	ECG data (Heart Rate Variability (HRV))	SDNN, RMSSD, HF, and LF extracted from HRV	A 23-year-old from 9:00 am to noon	Simulation	86% Accuracy
[14]	Drowsiness level detection and prediction	Artificial neural network (ANN)	×	✓	×	Car dataset, physiological dataset, behavioural dataset	Physiological, Sensorimotor, and deriving behavior features	21 participants for 110 min	Car simulator	NG
Hajinoroozi, Zhang [15]	Aimed at classifying the mental states of drivers (drowsy or alert).	Introduced deep covariance learning methods.	×	✓	✓	Brain (Physiological data- EGG)	Brain waves	9640 1s EEG epochs	Virtual reality driving	86.14%
Liu, et al. [16]	Aimed at improving EEG signals for real-life applications.	Presented a real-time brain-computer interfaces (BCI) system that uses an online algorithm for the prediction of fatigue. The BCI system was designed with pre-event EEG that evaluates the mental state of a driver when driving.	×	×	✓	Brain (physiological data)	The psychological state was assumed for all the participants.	20 right-handed, healthy young adults participated in the behavioural experiment (mean age ± standard deviation: 23.6 ± 2.9 years old).	Virtual-reality-based environment	75%
Hajinoorzi, et al [17]	Aimed to enable the prediction of a driver's state (alert or drowsy).	Proposed a new deep learning (DL) algorithm like channel-wise convolutional neural network (CCNN) and CCNN-R (which is a CCNN variation that uses Restricted Boltzmann Machine to replace the convolutional filter).	×	✓	✓	EEG data (physiological data characteristics are used together with the CCNN and CCNN-R classifiers.	Brain waves	26 records of approximately three minutes each (13 correspondings to the normal state and 13 more for the relaxed state) of 58 channels for each subject. Approximately 46,000 The experiment involved 44 subjects (mixed gender), non-smoking, aged from 18 to 35 years old	Virtual reality driving scenes	78.8%
[18]	Fatigue state prediction	Deep covariance learning model in which the Riemannian measures were employed in the classification of EEG data of the drivers.	×	NG	NG	EEG data (physiological data). Spatial and temporal covariance matrices were calculated from EEG epochs.	Driver's reaction time	9640 1s EEG of 100 participants	Virtual reality environment	73.82%
Cheng, et al. [19]	They attempted to assess the mental fatigue of a driver using Facial Landmark Sequences	Image-processing.	×	✓	✓	Physical data	Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR)	21 participants' faces	Static driving simulator	83.7% and 85.4%.
[20]	Aimed at recognizing fatigue in drivers.	Introduced a new method for the recognition of fatigue in drivers. An end-to-end trainable network was designed by combining the convolutional layers with LSTM units.	×	×	✓	Physical data	Eyes	Approximately 137,000 data points with timestamp information were collected	NG	83%
[21]	Sleepiness detection.	Proposed a model in which residual learning modules with SE block and transfer learning were used.	×	×	✓	Physical data	Eyes	3364 images for training and 1300 images for testing.	Simulation	85.81%
Our Method	Predicting the level of driver's fatigue	OpenCV DNN	✓	✓	✓	Physical data.	Eye blinking, mouth yawning and head movement	2423 subjects (eyes only) and 200 videos of 36 subjects.	Simulation	93.3%

the driver to enable the use of such physiological data. The attached sensors may serve as an obstacle in a sensitive process like driving, therefore it may not be suitable for all drivers as it may distract some drivers or may even be uncomfortable for drivers to wear.

On the other hand, predicting drivers fatigue studies specifically in currently conducted research suffer greatly from being inconvenient for creating a standard driving situation.

For example, the work of [13], include using BPNN which is known for its efficiency for computing gradient descent. Using BPNN in the scope helps to get less error rate however such a method is also known for taking extensive time for training. Moreover, the biggest limitation in this study is using input variables from attached sensors to drivers, which does not create a comfortable driving environment.

Similarly [14], [18], the EEG data was gathered and evidently, it is a reliable source for accurate results. In contrast, sensors need to be attached to drivers and as we mentioned earlier, it may create an uncomfortable environment.

The work of [17], used raw EEG signals data while considering the EEG data characteristics. The research in a similar sense used data derived from attaching sensors to the driver's which could deem to be uncomfortable, it could create a distraction in a sensitive process like driving.

In conclusion, most existing works on driver fatigue analysis are focused on real-time detection, with less emphasis placed on the driver's reaction time [23]. Therefore, real-time driver fatigue warnings cannot guarantee traffic safety. Instead, early warning or prediction may help save 2%–23% of all crashes [24].

III. DFLP FRAMEWORK STRUCTURE

The framework of the proposed DFLP is shown in Fig. 1 DFLP uses eye, mouth and head orientation which are among the common indicators of fatigue. The DFLP framework consists of many stages including, frame extraction and face detection. DFLP framework uses systematic steps of computer vision techniques such as pre-processing steps to clean the input data, and face detection to crop the face region from the unwanted background. The input image passes through the detector, and then it is subjected to image upscaling, which increases the chances of detecting smaller faces. However, upscaling has a substantial impact on the computation speed. The output is in the form of a list of faces with the (x, y) coordinates of the diagonal corners.

A. PRE-PROCESSING

The pre-processing operations are essential in the preparation of the raw data, as they help in obtaining clear detection features. The model's prediction ability can be affected by many factors such as different scenarios, gender, and skin colour. In terms of the lighting condition of a given scenario, the change of light intensity degrades the prediction by up to 6%. Therefore, it is very important to normalize and equalize the image. In order to minimize the environmental factors like light, skin color, sunglass and other similar elements, the local histogram equalization is generally used for the pre-processing of images. There are many types of histogram equalizations, such as global histogram equalization, local histogram equalization, adaptive histogram equalization amongst many other types. In the proposed model, the main focus is the face region, while the rest of the regions are not considered. In face images, the pixel colors are likely uniform and global histogram equalization is not optimal. Moreover, contrast and edges are more important than individual pixels and color. Therefore, in this work, the use of local histogram equalization is employed to enhance the image contrast and to normalize color distribution. The color transformations are calculated using the following equations:

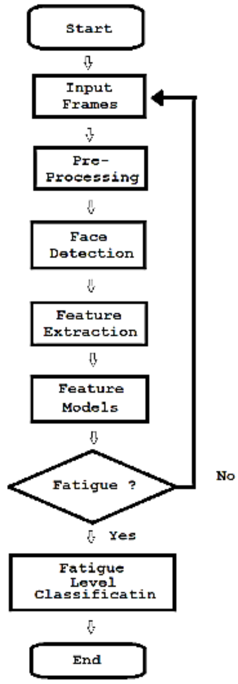


FIGURE 1. The proposed fatigue classification framework.

$$P_i = \frac{n_i}{n} \tag{1}$$

$$Acc_i = \sum_{j=0}^{i} P_j \tag{2}$$

$$T_i(n) = Acc_i^{-1}(n) \times L \tag{3}$$

where P_i : is the probability of pixel of color i in an image, Acc_i, Acc^{-1} are accumulate and its reverse function of P_i for histogram equalization, T_i: pixel and color transformation function, L: is the maximum Pixel Brightness.

The DFLP framework uses the following pre-processing tools: local histogram equalization [25], region of interest **(ROI)** [26], and image equalization [27]. Based on the threshold, there are many results and the tile size of local equalization filter as can be seen in Fig. 2 for threshold 1 and, Fig. 3 for threshold 2. There are 4 categories for each image (upper – original, middle-upper – global equalization, middle down – local equalization, down – adaptive equalization).

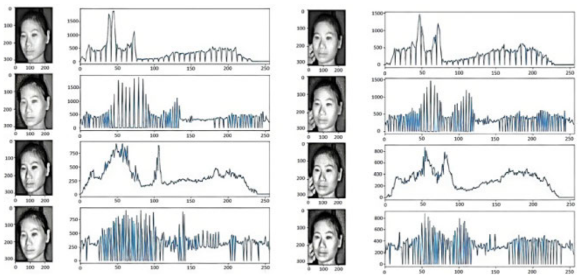


FIGURE 2. Tile size 4, 8, threshold 1.

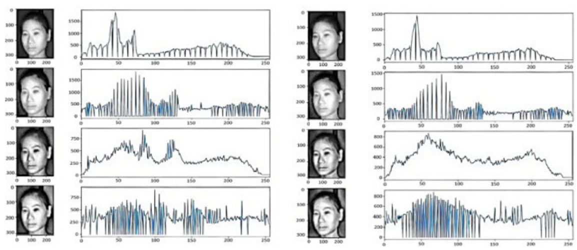


FIGURE 3. Tile size 4,8, threshold 2.

As seen in Fig. 3, the local equalization is better than other equalizations, and the optimal parameters are tile size 4 and threshold 2.

B. FACE DETECTION

The input for the proposed DFLP framework is video stream or webcam video. The face is detected from every frame of video by using OpenCV which has 2 models (OpenCV Haarcascade and DNN) for face detection.

OpenCV Haarcascade performs the extraction of the frontal face that is bigger than 80 by 80 pixels, and it can even detect partial face images with a confidence value. In a situation whereby, the drivers turned their heads or talk, the OpenCV Haarcascade failed to detect the drivers' faces. This means that the classification performance is affected by this factor. If there were many similar cases during the analysis of the frontal image in training videos, then it would be very difficult to determine whether the drivers are asleep or not.

To overcome this problem, the framework uses OpenCV DNN, which is a very useful tool for detecting frontal and side view of a face in variable face sizes. However, the speed of detection is slow. Driver's drowsiness prediction in real-time is not much important, because fatigue is the feeling of exhaustion and can be determined by observing the behavior of drivers within a time frame of (2∼3s).

C. FEATURE EXTRACTION

The DFLP framework extracts 3 features including, eye blinking, mouth yawning and head movement from detected faces as shown in Fig. 4 using Dlib. Dlib is a technique that is used to locate points on any face, specifically 68 points to locate the mouth corners and eyebrows as well as eyes.

There are always variations in the eye and mouth convex points and center position, and as such, it is very difficult to determine if the eyes and mouth are closed or opened. The eye and mouth size and geometric features vary, depending on gender, age, light condition, weather, wearing glasses and even driver's health.

FIGURE 4. Examples of distracted drivers in different orientations.

Fig. 4 shows that the driver's normal face orientation is slightly right forwarded and left or down in some cases, which means that drivers are distracted or falling asleep. Hence, the right eye images of drivers are less significant than left-eye images because the left-eye images are clear and well detected, but the right eye images are unclear and shadowed in some frames. Hence, the detection of eye blink and its duration based on the left eye is more accurate than the detection based on both eyes.

The eye position varies according to the person's features, gender and age. To enhance left eye detection accuracy, it is necessary to get the right position of the left eye from the face image. Due to some variation in the center and size of the left eye, it is important to use some techniques that can help in minimizing the variation.

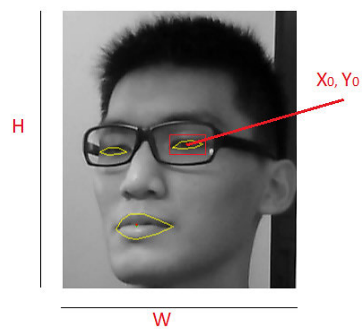


FIGURE 5. An example of defining the eye center ratio.

The figure below (Fig. 5) shows how the eye center position varies depending on the person and scenario. Here, the eye center ratio which is related to the width and height of the face image is defined.

$$Rx = \frac{X_0}{W} \tag{4}$$

$$R_y = \frac{Y_0}{H} \tag{5}$$

As can be seen in Fig. 6, R_y is the ratio of all detected face images. This is calculated by MATLAB based on the Rx, R_y ratio on the face images detected by Dlib from Driver

Drowsiness Detection Dataset by National Tsuing Hua University (**NTHU**) videos dataset [28].

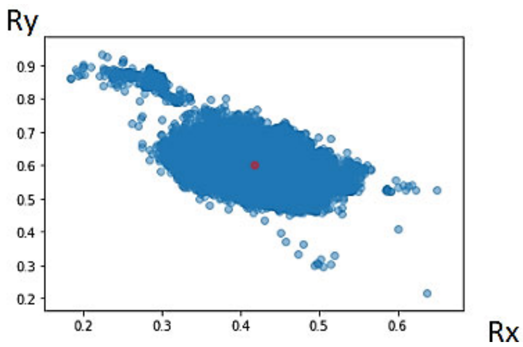


FIGURE 6. RX, RY face images ratio plot by Dlib and Matlab.

Fig. 6 shows that the left eye's ratio is varied 0.3-0.55 for Rx axis and 0.52-0.76 for R_y axis. The mean and standard deviation are as follow.

The mean value of this ratio is:

$$R_{x0} = 0.42(\pm 0.12)$$
$$R_{y0} = 0.58(\pm 0.15)$$

Let us suppose the left eye center is (x_i, y_i), then the eye center, width and height are:

$$Center_{eye}(i,j) = \begin{cases} (R_{x0}, R_{y0}) & \text{if } x_i \text{ not } \in R_{x0}, y_i \text{ not } \in R_{y0} \\ (x_i, y_i) & \text{otherwise} \end{cases} \quad (6)$$

$$WH_{eye}(i,j) = \begin{cases} (0.12W, 0.15H) & \text{if } x_i \text{ not } \in R_{x0}, y_i \text{ not } \in R_{y0} \\ (w, h) & \text{otherwise} \end{cases} \quad (7)$$

where, w, h are the width and height of the face image, while w, h are the width and height of the left eye.

It is also important to detect the mouth region when a driver is in a normal state, talking or yawning.

Fig 4 shows that Dlib does not correctly detect the mouth region when a driver is yawning. Therefore, the mouth region must be detected based on the detected mouth and other features. A look at the images shows that the mouths are centered between the noses and chins. Based on this, the mouth region can be identified. In this case, the mouth region can be refined based on facial landmarks.

Fig. 7 shows the facial landmarks obtained using Dlib. Dlib defines 68 landmarks that identify the face. Some features show the mouth, while others show the mouth, nose, chin, jaw. These features are numbered from 1 to 68. The mouth is in the range of 34 and 6, 12 in height, 7 and 12 in width. Based on this, the mouth region can be defined as follows:

$$Center_{mouth}(i,j) = \left(\frac{dlib_{7x} + dlib_{12x}}{2}, \frac{dlib_{34y} + dlib_{12x}}{2}\right) \quad (8)$$

$$WH_{mouth}(i,j) = (dlib_{7x} - dlib_{12x}, dlib_{34y} - dlib_{12x}) \quad (9)$$

where, $dlib_{34x}$, $dlib_{34y}$ are the x and y values of Dlib landmark 34. We can identify the eye and mouth region based on the features, which is, is mouth: features 49 – 68, while left eye: features 43-48. Therefore, feature 34 implies the nose point 34.

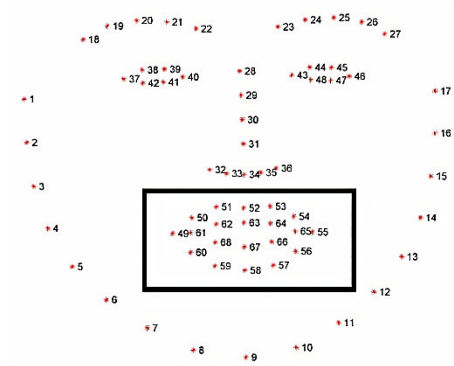


FIGURE 7. The 68 landmarks that identify the face region.

After determining eye and mouth center position, it is possible to extract the eye, mouth and head image approximately based on center position. The input face is not frontal; therefore, it is challenging to get a clear image of the left and right eyes. Moreover, it is impossible to get the images of the mouth and left or right eye. Therefore, it might be a better approach to extract the right eye image when a driver's head orientation is towards the right, while that of the left eye can be extracted when the drivers' head orientation is towards the left.

There are small variations of eye and mouth region detected by Dlib, and this degrades the accuracy of prediction. To get the right region of the eye and mouth, the moving average should be used in obtaining center and width of the eye and mouth, because the sequence frame does not change much in eye and mouth position. The transformation equation is given below:

$$ROI_{eye}(i,j) = Center_{eye}(i_0, j_0) + WH_{eye}(i_0, j_0) \quad (10)$$

$$Center_{eye}(i_0, j_0) = \frac{1}{m}\sum Center_{eye}(i,j) \quad (11)$$

$$WH_{eye}(i_0, j_0) = \frac{1}{m}\sum WH_{eye}(i,j) \quad (12)$$

$$ROI_{mouth}(i,j) = Center_{mouth}(i_0, j_0) + WH_{mouth}(i_0, j_0) \quad (13)$$

$$Center_{mouth}(i_0, j_0) = \frac{1}{m}\sum Center_{mouth}(i,j) \quad (14)$$

$$WH_{mouth}(i_0, j_0) = \frac{1}{m}\sum WH_{mouth}(i,j) \quad (15)$$

where, ROI_{eye} is the extracted eye region of interest. ROI_{mouth} present the extracted mouth region of interest. $Center_{eye}(i_0, j_0)$ is center point of the eye, $WH_{eye}(i_0,j_0)$ is width, height points of the eye.

IV. FEATURE DETECTION

The following section will explain in detail the proposed DFLP model. DFLP model based on three features mouth, eyes and head information that defines the level of drivers' fatigue, as explained in features models below:

A. EYES BLINKING CNN MODEL

The use of CNN model is employed for eye blinking. There are lists of mouth talk and yawn (ground truth), and head movement that can be used in the training video and annotation text file, but there is none for eye blinking. Hence, using only such output data to train eye blink is a difficult task. Therefore, in this step, the model is pre-trained using the eye open and close image (**CEW**) dataset [29].

The eyes model consists of 9 layers modified from LeNet-5 CNN layer as can be seen in TABLE 2.

TABLE 2. Eye model layers that contain 9 CNN layers.

CNN Layers	No. pixels	Purpose
ZeroPadding2D	24,24,1	Extends the image and computes interpolated spectra of the image
Convolution2D	16	Convolution layer
MaxPooling2d		Find the max point in a tile size of convolution image
Flatten		Flatten 2D parameters to 1D parameters
Dense	512	Densely connected layer
Dropout		Randomly parameters selection layer
Dense	512	Densely connected layer
Dropout		Randomly parameters selection layer
Dense	2	Output layer

There are many layers ranging from ZeroPadding2d layers (as input layer) to Dense layer (as output layer). The input of this model is a gray scale eye image with the size of 24×24 pixels, whereas the output of this model is a 1D vector of (eye close, eye open). ZeroPadding2D layer is used for input 2D array extension for getting more features. The eye images of the open and closed eyes are similar, but it is very difficult to differentiate between closed eyes and open eyes. Thus, the feature parameters can be extended by using Zero Padding 2D. The Convolution2D layer is used for feature extraction based on different criteria. MaxPooling2d layer is used in reducing the feature parameters and minimizing the effect of noise and relevant pixels or regions. This layer contains several heat maps that are very vital for the classification of the image. The dense layer is used to find the feature parameters and connected feature relations, while the Dropout layer randomly selects the parameters and dropouts so that the model does not encounter the problem of overfitting.

B. MOUTH AND HEAD MOVEMENT CNN MODELS

CNN models for mouth yawn detection and head movement are similar to eye blink detection. The difference between eye and mouth models is that there is no ZeroPadding2D layer because the yawn, talk and closed mouth are quite different. The input of this model is also a 2D gray image of 24x 24pixels, and the output is a 1D vector of (close, yawn, talk) for mouth model, and (normal, falling sleep, distraction) as can be seen in TABLE 3.

At this step, it is necessary to identify the driver's state based on the outputs of CNN models. For example, the output of the eye blink CNN model shows if the driver's eye is opened or closed but does not show the fatigue level. Therefore, there is a need to identify the driver's state based on observation of these values. Based on this model, the eye blinking ratio can be identified. In general, normal eye blinking duration is less than 300 $\sim$ 500ms, and fatigued 500 $\sim$ 1000ms and falling sleep are longer than 1s. Thus, the blinking ratio can be defined as the moving average of eye blink over time. Sometimes, the blinking ratio can be 1 or 2 sec, which includes the whole eye blinking process). Here, there is a need to classify all cases including falling asleep, therefore the time duration is about 1 $\sim$ 1.5second.

TABLE 3. Mouth and head movement CNN models that contains 9 layers.

CNN Layers	No. pixels	Purpose
Convolution2D	24,24,1	Convolution layer
MaxPooling2d	16	Find the max point in a tile size of convolution image
Flatten		Flatten 2D parameters to 1D parameters
Dense	512	Densely connected layer
Dropout		Randomly parameters selection layer
Dense	512	Densely connected layer
Dropout		Randomly parameters selection layer
Dense	3	Output layer

The moving average parameters that were used in this work are divided into three different thresholds: Eye blinking model is 300-400ms, normal blinking time, and blink period: 3 $\sim$ 4s. The Moving average size is 50 frames. The weighted class value is: 1 for class 0 (open), 10 for class 1(closed), and the threshold: > 300/3000 * (1+11) = 0.11. As for the mouth, the yawn duration is > 5 ∼ 6s, and 1.5s in some cases (Hence 2s < optimal < 6s). The Moving average size is 80 frames, while the Weighted class value is 1 for class 0 (closed), 10 for class 2 (talking), and 30 for class 1 (yawning). The threshold is 5 < talking < 15, yawning > 15. For the head, the nodding notification time is > 1s, while the Moving average size is 30, and the Weighted class value is 1 for class 0 (normal), 10 for class 2(turn left or right), and 30 for class 1 (falling asleep). The threshold value is 5 < turning < 15, falling asleep > 15.

$$Avg_{eye} = \frac{1}{n}\sum_{i=0}^{n} Eye(i) \tag{16}$$

where $Eye(i)$ is the status of the eye at the frame i, And the fatigue level is like the following:

$$P_{eye} = \begin{cases} normal & if\,Avg_{eye} < 0.3 \\ early\ fatigue & else\ if\,Avg_{eye} < 0.1 \\ fatigued & otherwise \end{cases} \tag{17}$$

We can use the same method to identify the mouth yawning and head movement.

$$Avg_{mouth} = \frac{1}{m}\sum_{i=0}^{m} Mouth(i) \tag{18}$$

$$Avg_{head} = \frac{1}{k}\sum_{i=0}^{k} Head(i) \tag{19}$$

And the fatigue level is identified as the following:

$$P_{mouth} = \begin{cases} closed & if\ Avg_{mouth} < 0.1 \\ talking & else\ if\ Avg_{mouth} < 0.5 \\ yawning & otherwise \end{cases} \tag{20}$$

$$P_{head} = \begin{cases} normal & if\ Avg_{mouth} < 0.1 \\ distraction & else\ if\ Avg_{mouth} < 0.7 \\ nodding & otherwise \end{cases} \tag{21}$$

In the first case (eye blink), if the time duration is set at 1.5 sec, the frame size must be n = 1.5 / (1/30) = 45 frames provided that the frame speed is 30fps. k and m are the frame size for mouth yawning and head movement, then specific time duration is cited in the next section (result and comparison/ Moving average parameters), and Avg is the moving average, mean value; any notation can be found for that variable and parameters.

V. CLASSIFICATION MODEL

After the models of eye, mouth and head have been trained, there is a need to classify the driver's state. The process of the classification involves classifying the driver's state into normal and fatigued. When the drivers are not sleeping or distracted, they are classified as normal, but when the drivers are asleep, feel exhausted or are distracted from driving, they are classified as fatigued. Therefore, a logical model is built for the classification of the fatigue level based on eye blinking, mouth yawn and head movement.

$$Fatigue = f1(head, f2(eye, mouth)) \tag{22}$$

where, $f1$ is the function that classifies fatigue and normal state of drivers where $f2$ is the function that drivers are falling asleep, distracted from driving, or in a normal state.

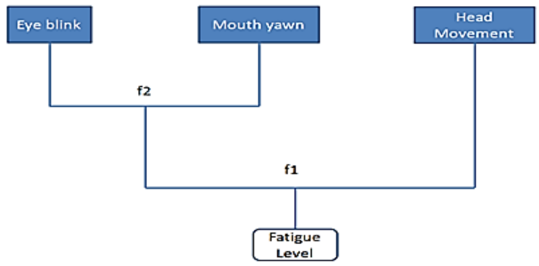


FIGURE 8. The logical model that contains f1, f2 function that classifies the fatigue.

The logical model in Fig.8 is presented in the following equations:

$$\begin{aligned} P(Fatigue) &= P(head\ and\ f2) = P(f2) \times P(head\,|\,f2) \\ &= P(f2)\,P((normal, nodding, distracting)\ |\,f2) \\ &= P(f2) \times (P(normal\,|\,f2) + P(nodding\,|\,f2) \\ &\quad \times P(distracting\,|\,f2)) \end{aligned} \tag{23}$$

$$\begin{aligned} P(f2) &= P(eye, mouth) = P(eye) \times P(mouth\,|\,eye) \\ &= P(eye) \times P((closed, talking, yawning)\ |\,eye) \\ &= P(eye) \times (P(closed\,|\,eye) + P(talking\,|\,eye) \\ &\quad + P(yawing\,|\,eye))) \end{aligned} \tag{24}$$

$$\begin{aligned} P(eye) &= P(eyeblinking, eyeblinkingrate, \\ &\quad blinkingspeed) \end{aligned} \tag{25}$$

where $P(eye)$, means all status of the eye, that is, eye state while blinking, eye state while talking, and eye state while yawing.

Based on $P(Fatigue)$, we can identify the driver's fatigue level. $P(Fatigue)$ is the fatigue level calculated by the equation $f1(head, f2(eye, mouth))$- meaning 0 – normal, not sleeping, 0.5 –early fatigued, 1.0 falling asleep.

VI. EXPERIMENTS

A. EXPERIMENTAL SETUP

We have used Intel Core i5-6200U CPU, 2.3GHz with 8 GB of RAM in our experiment. We proposed a framework for driver fatigue level prediction.

B. DATASET

We used publicly available datasets for our experiment. The main dataset was from National Tsuing Hua University which was used in many recent publications including [30], [31]. NTHU dataset consists of 36 subjects from different ethnicities including men and women, and the videos of the subjects were recorded under night and day illumination conditions. Moreover, the subjects were recorded in different simulated driving scenarios such as barefaced, night-bare face, sunglasses, glasses, night-glasses as shown in Fig 9.

In every scenario, the subjects were recorded in two states (sleepy/non-sleepy), and a total of 200 videos were obtained.

Furthermore, the videos were converted to images to serve as input in the proposed DFLP framework. A video with a duration of 1.5-minute was produced; the video contains a combination of drowsiness-based symptoms such as slow blinking, nodding, and yawning and non- drowsiness-based actions such as being distracted, laughing or talking.

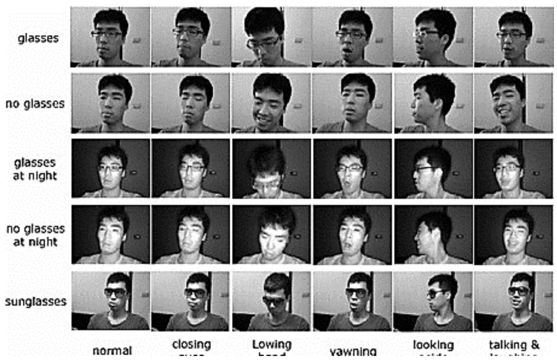


FIGURE 9. Different driving scenarios of NTHU dataset [32].

The second public available dataset that we used in our experiment was closed eyes in the wild, which was used for driver drowsiness recognition by many researchers including [33]. CEW dataset was used to train our eyes blinking CNN model. The dataset contains unconstrained scenarios of real-world applications, hence, it contains a variety of different individuals with different ages, gender, background as well as wearing different accessories, such as glasses and hats. The dataset has images of different quality including blurry, disguised, and occlusion under various lighting conditions. More so, the dataset contains different view angles and zooming levels. CEW includes 24×24 pixels centered eye patches extracted from 100×100 pixels face images. The total number of images includes 2423 subjects, out of which 1192 of them were closed eyes images and the remaining 1231 opened eyes images. Fig. 10 illustrates an example of CEW dataset.

FIGURE 10. An example of CEW centered eye patches dataset.

As can be seen in TABLE 4, OpenCV DNN extracts more faces than Haarcascade, Therefore, the classification would be greater than OpenCV Haarcascade.

C. RESULTS AND DISCUSSION

Predicting drivers' fatigue is more important than just detecting fatigue, due to the fact that prediction could give a bigger window for drivers to act upon a sudden hazard.

TABLE 4. The total extracted number of OPENCV DNN frames in comparison with OPENCV HAARCASCAD.

Algorithm / Data	OpenCV DNN	OpenCV Haarcascade
S1	3330/3330 frames	1967/3330 frames
S2	1945/2167 frames	868/2167 frames
S3	11268/12704 frames	10017/12704 frames

[13] have presented fatigue classification method that into two classifications: drowsy, and alert.

The data was obtained from an HRV in the form of EEG epochs. Their study concluded that deep covariance learning methods reported better performance than shallow learning methods and that was through using a CNN model. The study has successfully improved upon the accuracy from 70.96% to 86.14%.

Similarly, the work of [14], has also been used to improve the quality of the EEG data that was constructed from the driver's brain waves reading signals. A real-time brain-computer interfaces system was developed using an online algorithm with pre-event EEG. One of the main achievements in this study was to create a noise cancellation system that can improve the prediction of driver's fatigue. The drawback of this study is similar to [13] in a sense of using attached sensors data to drivers.

In the work [17], the characteristics of the EEG data was considered to develop DL algorithms, CCNN and CCNN-R. Raw data were used to achieve robust performance.

In addition, the work of [18] also included using EEG data. The research achieved a performance increase of 3.16% compared to Euclidian measures in classifying driver's fatigue when using temporal covariance matrices.

This experiment focused on measuring the accuracy of eye blinking, mouth yawning, and head movement CNN models. In the experiment, Video 1 includes drivers with glasses (S1), Video 2 includes drivers with sunglasses (S2), and Video 3 include drivers with a mixture of glasses and sunglasses (S3).

TABLE 5. CNN model for eye blink, mouth yawn, and head movement training and validation accuracy results.

Parameters	Training	Validation
Eye blink	99.9%	98.7%
Mouth yawn	99.8%	97.6%
Head movement	99.9%	97.3%

The overall mathematical Model accuracy was gotten by comparing the ground truth and the predicted list from our mathematic model.

TABLE 6. Test accuracy results of eyes blinking, mouth yawn, and head movement CNN models.

Parameters	S1	S2	S3
Eye blink	96.7%	86.3%	93.7%
Mouth yawn	94.3%	95.1%	94.7%
Head movement	96.5%	89.3%	93.7%

TABLE 7. Our overall mathematical model accuracy score.

Data	Test accuracy
S1	96.5%
S2	89.3%
S3	94.1%
Overall	93.3%

In this paper, we proposed DFLP framework that can identify the level of driver's fatigue through facial and dynamic indications. We use a systematic methodology of pre-processing, face detection and feature extraction to get the maximum accuracy score. We used NTHU and CEW datasets in our experiment. Drivers face images in each driving scenario were captured. For each face extracted image, EARs, MARs and Head Pose were determined. Several CNN algorithms were used to measure eye blink, mouth, and head movement. We used the subjective annotation to assess the level of fatigue and to play as the ground truth of the driver's fatigue level. The general block diagram of the proposed methodology was shown in Fig. 8.

There are three findings of this study. First, we demonstrate the fatigue level prediction assessed from subjective rates can be predicted by eye, mouth and head movement feature with high accuracy. The eyes, mouth and head features can be predicted by CNN models using facial landmarks as input features. In our study, the prediction accuracy for each feature was more than 97%. Thus, there exists a great potential to apply inexpensive nonintrusive in-vehicle cameras for the purpose of driver's fatigue monitoring and warning.

Second, our results demonstrate that the features extraction can be improved by a combination of global, local and adaptive histogram equalizations. It was confirmed that the various illumination conditions and accessories such as sunglasses create obstacles to obtaining eye features. However, the feature extraction of EARs was greatly improved by the use of CNN (over 86%). Furthermore, it is shown that the moving average applied to the feature sequence stabilizes the features greatly.

Third, we suggested a logical model for fatigue level prediction based on eye blink, mouth and head movement. The various states of the driver, talking, yawning, turning, and falling asleep are determined and a final prediction is made based on the driver's state.

VII. CONCLUSION AND FUTURE WORK

This study shows using facial and head behavior images as an only input source is effective for predicting the level of fatigue. The advantage of this work is that it predicts and measures the level of fatigue instead of detecting the driver's fatigue while using low-cost equipment for the input feed. Moreover, the input sensors are detached from the driver unlike using attached sensors that hinder the driving process and may play as an obstacle for drivers. Furthermore, since developed countries will use automotive cars, many developing countries have many trucks and other vehicles operated by humans as an example, therefore we can use this work to assist driver's safety.

Our proposed framework can also be used as a baseline for an engineered device that has an input sensor, and a processing unit, alternatively a device that has the ability to be connected online to a server in the cloud for online processing. this device can be in a small size to be placed on the dash of the car in Infront of the driver.

However, the limitation of this work is that it works in simulated driving scenarios, hence, under uncontrolled conditions, this framework may fail to capture the facial landmark of drivers. Future work will tend to tackle this obstacle potentially by using more advanced machine vision techniques specifically (facial expression embedding and sequence embedding) that can be used to calibrate the eye and mouth parameters.

Another suggestion to tackle such limitation can be by using more facial and behavioral features classifications, such as drivers nodding behavior and drivers frown expressions that occur during fatigue.

In addition, an alert can be added as an output to our DFLP to advise drivers of their fatigue level. Such an alert can be useful to take the necessary adjustments such as advising the drivers to drive with caution or to not drive based on the level of their fatigue.

REFERENCES

[1] H. Abbood, W. Al-Nuaimy, A. Al-Ataby, S. A. Salem, and H. S. AlZubi, "Prediction of driver fatigue: Approaches and open challenges," in *Proc. 14th UK Workshop Comput. Intell. (UKCI)*, Sep. 2014, pp. 1–6.

[2] O. Akiva, A. Atsmon, and D. Atsmon, "Continuous identity monitoring for classifying driving data for driving performance," U.S. Patent 10 229 46, Mar. 12, 2019.

[3] A. Girit, "Drowsy driver detection using image processing," M.S. thesis, Graduate School Natural Appl. Sci., Middle East Tech. Univ., Ankara, Turkey, 2014.

[4] N. Alioua, A. Amine, A. Rogozan, A. Bensrhair, and M. Rziza, "Driver head pose estimation using efficient descriptor fusion," *EURASIP J. Image Video Process.*, vol. 2016, p. 1–14, Jan. 2016.

[5] C. Yan, "Vision-based driver behaviour analysis," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Univ. Liverpool, Liverpool, U.K., 2016.

[6] N. Munla, M. Khalil, A. Shahin, and A. Mourad, "Driver stress level detection using HRV analysis," in *Proc. Int. Conf. Adv. Biomed. Eng. (ICABME)*, Beirut, Lebanon, Sep. 2015, pp. 61–64, doi: 10.1109/ICABME.2015.7323251.

[7] Giovanni, T. Suprihadi, and K. Karyono, "DROWTION: Driver drowsiness detection software using MINDWAVE," in *Proc. Int. Conf. Ind. Autom., Inf. Commun. Technol.*, Bali, Indonesia, Aug. 2014, pp. 141–144, doi: 10.1109/IAICT.2014.6922096.

[8] Z. Gao, X. Wang, Y. Yang, C. Mu, Q. Cai, W. Dang, and S. Zuo, "EEG-based spatio–temporal convolutional neural network for driver fatigue evaluation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 9, pp. 2755–2763, Sep. 2019, doi: 10.1109/TNNLS.2018.2886414.

[9] A. Mittal, K. Kumar, S. Dhamija, and M. Kaur, "Head movement-based driver drowsiness detection: A review of state-of-art techniques," in *Proc. IEEE Int. Conf. Eng. Technol. (ICETECH)*, Coimbatore, India, Mar. 2016, pp. 903–908, doi: 10.1109/ICETECH.2016.7569378.

[10] V. Balasubramanian and R. Bhardwaj, "Grip and electrophysiological sensor-based estimation of muscle fatigue while holding steering wheel in different positions," *IEEE Sensors J.*, vol. 19, no. 5, pp. 1951–1960, Mar. 2018.

[11] I. R. Tayibnapis, D.-Y. Koo, M.-K. Choi, and S. Kwon, "A novel driver fatigue monitoring using optical imaging of face on safe driving system," in *Proc. Int. Conf. Control, Electron., Renew. Energy Commun. (ICCEREC)*, Bandung, Indonesia, Sep. 2016, pp. 115–120.

[12] A. Guettas, S. Ayad, and O. Kazar, "Driver state monitoring system: A review," *Proc. 4th Int. Conf. Big Data Internet Things, Rabat, Morocco (BDIoT)*, Tangier-Tetuan, Morocco, Oct. 2019, pp. 1–7.

[13] D. Utomo, T.-H. Yang, D. T. Thanh, and P.-A. Hsiung, "Driver fatigue prediction using different sensor data with deep learning," in *Proc. IEEE Int. Conf. Ind. Cyber Phys. Syst. (ICPS)*, Taipei, Taiwan, May 2019, pp. 242–247.

[14] X. Zhang, X. Zhao, H. Du, and J. Rong, "A study on the effects of fatigue driving and drunk driving on drivers' physical characteristics," *Traffic Injury Prevention*, vol. 15, no. 8, pp. 801–808, Jul. 2014, doi: 10.1080/15389588.2014.881996.

[15] M. Hajinoroozi, J. M. Zhang, and Y. Huang, "Driver's fatigue prediction by deep covariance learning from EEG," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Banff, AB, Canada, Oct. 2017, pp. 240–245.

[16] Y.-T. Liu, S.-L. Wu, K.-P. Chou, Y.-Y. Lin, J. Lu, G. Zhang, W.-C. Lin, and C.-T. Lin, "Driving fatigue prediction with pre-event electroencephalography (EEG) via a recurrent fuzzy neural network," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE)*, Vancouver, BC, Canada, Jul. 2016, pp. 2488–2494.

[17] M. Hajinoroozi, Z. Mao, and Y. Huang, "Prediction of driver's drowsy and alert states from eeg signals with deep learning," in *Proc. IEEE CAMSAP*, Cancun, Mexico, Dec. 2015, pp. 493–496.

[18] M. Hajinoroozi, J. Zhang, and Y. Huang, "Prediction of fatigue-related driver performance from EEG data by deep Riemannian model," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jeju, South Korea, Jul. 2017, pp. 4167–4170.

[19] Q. Cheng, W. Wang, X. Jiang, S. Hou, and Y. Qin, "Assessment of driver mental fatigue using facial landmarks," *IEEE Access*, vol. 7, pp. 150423–150434, 2019.

[20] Z. Xiao, Z. Hu, L. Geng, F. Zhang, J. Wu, and Y. Li, "Fatigue driving recognition network: Fatigue driving recognition via convolutional neural network and long short-term memory units," *IET Intell. Transp. Syst.*, vol. 13, no. 9, pp. 1410–1416, Sep. 2019.

[21] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Learning and transferring mid-level image representations using convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Columbus, OH, USA, Jun. 2014, pp. 1717–1724.

[22] H. J. Michielsen, J. De Vries, and G. L. Van Heck, "Psychometric qualities of a brief self-rated fatigue measure," *J. Psychosomatic Res.*, vol. 54, no. 4, pp. 345–352, Apr. 2003.

[23] Y.-Y. Lin and P.-A. Hsiung, "An early warning system for predicting driver fatigue," in *Proc. IEEE Int. Conf. Consum. Electron.-Taiwan (ICCE-TW)*, Taiwan, Taipei, Jun. 2017, pp. 283–284.

[24] T. Kundinger, C. Mayr, and A. Riener, "Towards a reliable ground truth for drowsiness: A complexity analysis on the example of driver fatigue," *Proc. ACM Hum.-Comput. Interact. J.*, vol. 4, Mar. 2020, Art. no. 78, doi: 10.1145/3394980.

[25] H. Wang, F. You, X. Chu, X. Li, and X. Sun, "Research on customer marketing acceptance for future automatic driving—A case study in China city," *IEEE Access*, vol. 7, pp. 20938–20949, Feb. 2019.

[26] Y. Wang and Z. Pan, "Image contrast enhancement using adjacent-blocks-based modification for local histogram equalization," *Infr. Phys. Technol.*, vol. 86, pp. 59–65, Nov. 2017.

[27] R. Brinkmann, *The Art and Science of Digital Compositing*. Burlington, MA, USA: Morgan Kaufmann, 1999.

[28] Y. Wang, Q. Chen, and B. Zhang, "Image enhancement based on equal area dualistic sub-image histogram equalization method," *IEEE Trans. Consum. Electron.*, vol. 45, no. 1, pp. 68–75, Feb. 1999.

[29] C. H. Weng, Y. H. Lai, and S. H. Lai, "Driver drowsiness detection via a hierarchical temporal deep belief network," in *Proc. Asian Conf. Comput. Vis. (ACCV)*, Taipei, Taiwan, Nov. 2016, pp. 117–133.

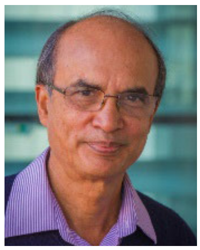
[30] W. Liu, J. Qian, Z. Yao, X. Jiao, and J. Pan, "Convolutional two-stream network using multi-facial feature fusion for driver fatigue detection," *Future Internet*, vol. 11, no. 5, p. 115, May 2019, doi: 10.3390/fi11050115.

[31] B. K. Savas and Y. Becerikli, "Real time driver fatigue detection system based on multi-task ConNN," *IEEE Access*, vol. 8, pp. 12491–12498, 2020.

[32] J. Lyu, H. Zhang, and Z. Yuan, "Joint shape and local appearance features for real-time driver drowsiness detection," in *Proc. Asian Conf./Workshop Comput. Vis.*, Taipei, Taiwan, Nov. 2016, pp. 178–194, 2016.

[33] L. Zhao, Z. Wang, G. Zhang, Y. Qi, and X. Wang, "Eye state recognition based on deep integrated neural network and transfer learning," *Multimedia Tools Appl.*, vol. 77, no. 15, pp. 19415–19438, Nov. 2017.

HAIDER A. KASSEM received the B.Sc. degree in software engineering from the University of Baghdad, Iraq, in 2009, and the M.Sc. degree in information technology from Osmania University, India, in 2013. He is currently pursuing the Ph.D. degree in information technology with Deakin University, Australia. He was a Lecturer with the Laser and Optical Engineering Department, Al-Kut University College, Kut, Iraq, and the Mathematics and Computer Science Department, Wasit University, Kut. He has published and contributed to several research papers. His research interests include machine learning, computer vision, and digital image processing.



MORSHED CHOWDHURY received the Ph.D. degree from Monash University, Australia, in 1999. He is currently an Academic Staff Member with the School of Information Technology, Deakin University, Australia. Prior to joining Deakin University, he was an Academic Staff with Gippsland School of Computing and Information Technology, Monash University. He has more than 12 years of industry experience in Bangladesh and Australia. As an International Atomic Energy Agency (IAEA) Fellow, he has visited a number of international laboratory/centers, such as Bhaba Atomic Research Centre, India; Brookhaven National Laboratory, NY, USA; and International Centre for Theoretical Physics (ICTP), Italy. He has published more than 165 research papers, including a number of journal articles, conference papers, and book chapters. His current research interests include security of the Internet of Things, wireless network security, health data analytics, image processing, and documentation security. He has organized a number of international conferences and has been serving as a member of the technical and program committee for several international conferences, since 2001. He has also acted as a reviewer of many journal articles.

JEMAL H. ABAWAJY is currently a Full Professor with the Faculty of Science, Engineering and Built Environment, School of Information Technology, Deakin University, Australia. He is also the Director of the Parallel and Distributing Computing Laboratory. He is the author/coauthor of five books, more than 250 papers in conferences, book chapters, and journals, such as IEEE Transactions on Computers and IEEE Transactions on Fuzzy Systems. He is a Senior Member of IEEE Computer Society, IEEE Technical Committee on Scalable Computing (TCSC), IEEE Technical Committee on Dependable Computing and Fault Tolerance, and IEEE Communication Society. He has served on the editorial board for numerous international journals and currently serving as an Associate Editor for the *International Journal of Big Data Intelligence* and *International Journal of Parallel, Emergent and Distributed Systems*. He has also guest edited many special issues. He also edited ten conference volumes.