IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 30, 2021, accepted February 3, 2021, date of publication February 9, 2021, date of current version February 23, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3058205

Active Vision-Based Attention Monitoring System for Non-Distracted Driving

LAMIA ALAM[1], (Member, IEEE), MOHAMMED MOSHIUL HOQUE[1], (Senior Member, IEEE), M. ALI AKBER DEWAN[2], (Member, IEEE), NAZMUL SIDDIQUE[3], (Senior Member, IEEE), INAKI RANO[4], AND IQBAL H. SARKER[1], (Member, IEEE)

[1]Department of Computer Science and Engineering (CSE), Chittagong University of Engineering and Technology (CUET), Chattogram 4349, Bangladesh
[2]School of Computing and Information Systems, Faculty of Science and Technology, Athabasca University, Athabasca, AB T9S 3A3, Canada
[3]School of Computing, Engineering, and Intelligent Systems, Ulster University, Londonderry BT47 7JL, U.K.
[4]The Maersk Mc-Kinney Moller Institute, University of Southern Denmark, 5230 Odense, Denmark

Corresponding author: Mohammed Moshiul Hoque (moshiul_240@cuet.ac.bd)

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) under Grant RGPIN-2020-06080.

ABSTRACT Inattentive driving is a key reason of road mishaps causing more deaths than speeding or drunk driving. Research efforts have been made to monitor drivers' attentional states and provide support to drivers. Both invasive and non-invasive methods have been applied to track driver's attentional states, but most of these methods either use exclusive equipment which are costly or use sensors that cause discomfort. In this paper, a vision-based scheme is proposed for monitoring the attentional states of the drivers. The system comprises four major modules-cue extraction and parameter estimation, state of attention estimation, monitoring and decision making, and level of attention estimation. The system estimates the attentional level and classifies the attentional states based on the percentage of eyelid closure over time (PERCLOS), the frequency of yawning and gaze direction. Various experiments were conducted with human participants to assess the performance of the suggested scheme, which demonstrates the system's effectiveness with 92% accuracy.

INDEX TERMS Computer vision, attentional states, attention monitoring, human-computer interaction, driving assistance, gaze direction.

I. INTRODUCTION

Global surveys suggest that drivers' lack of attention is the major cause of road accidents [1]. Every year more than 1.35 million people die and tens of millions more are injured or disabled due to road accidents. Moreover, death rate is three times higher in countries with low income than in high-income countries. It is the responsibility of the driver to keep the attentional state high during driving for the safety of the passengers and the driver him/herself. Attentional state represents the physical, physiological, and behavioral parameters of the driver [2]. Attentional state can be low due to various distracted activities which include using cell phone for texting or talking, eye glance away from the road due to rambling of mind, or sleepiness resulted from a lack of rest and prolonged mental activity or long period of stress or anxiety [3]. It has been found that factors such as fatigue, drowsiness, and distraction impede the drivers' ability to pay attention on the road and surroundings, which result in most road traffic crashes [4]. Monitoring drivers' attention with smart driving assistance system would reduce the risk of road crashes and may help to improve the driving efficiency.

There are various ways to track driver's attentional states through using either invasive or non-invasive methods. In invasive methods, sensors are often used to analyze driver's physiological states and driving performance. Physiological signal includes heart rate (electrocardiogram or ECG/EKG signal), brain activity (electroencephalography or EEG signal), muscle current (electromyography or EMG signal), respiratory rate variability (RRV), eye's cornea-retinal standing potential (Electrooculography or EOG), and skin conductance (electrodermal activity signal). These signals are often collected through electrodes connected to human body. Also, various in-vehicle sensory or external devices such as accelerometer, gyroscope, and magnetometer are used to assess driving performance by acquiring data from steering

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

wheel angle, brake or acceleration statue, lane position changing pattern and so on. Most of these sensors require direct contact with skin which is intrusive for the users and often provides distorted information [5].

On the other hand, the non-invasive methods (e.g., the vision-based systems) do not require any contact with body, and the imaging sensor collects attentional information of the individual from a distance. Vision-based attentional information includes facial features and body movements. Most used non-invasive attentional cues in currently available vision-based systems are eyelid movements (e.g.,eye blink frequency, closure duration) [6]–[11], eye gaze [6]–[8], [12], head movement [7]–[9], [13], [14], facial expressions (e.g., yawning, lip movements etc.) [7], [11], [13], [15], and body movements (like hand movement) [15]–[17]. However, the existing vision-based systems have some limitations: (i) capturing sensors used by the aforementioned systems are either expensive camera(s) [7]–[10], [14], [15], [17] with any additional sensor/hardware [11], [12], [16], [17] or some specialized imaging sensor (e.g., eye tracker [6] and Kinect [13]); (ii) some of the systems used only a single parameter, such as pupil [12], PERCLOS [10] or head pose [14] to estimate the driver's attentional state, making the system unable to adapt to some situations which are common in the real driving scenario (e.g., turning head or wearing sun glass can hide eye) and resulting in incorrect attentional state detection; (iii) some systems detect single inattentional state [8]–[14], [17] or limited only to the level of the same state [6], whereas some other focus on detecting driver's activity [15], [16]; (iv) some systems do not have any alert system to warn the driver of any inattentional state when detected during driving [6]–[8], [10], [11], [14]–[16]; (v) some of the previous works [11], [15], [17] were evaluated in a simulated environment and may not work accurately in the real driving scenario; and (vi) no evidence was provided about the systems mentioned above working in diverse situations (e.g., drivers were having different facial features such as beard, moustache, and hairstyles or wearing accessories (e.g., spectacle, sunglass, and cap) which are common in real driving scenarios.

In this paper, we intend to overcome the above limitations. We propose a vision-based system that extracts driver's attentional cues/features to estimate attentional states and classifies it into attentive, drowsy, fatigued and distracted. The system also alerts the driver in any of inattentional states, such as drowsy, fatigue or distraction. As we found that fatigue, drowsiness, and visual distraction are major causes of inattentiveness which are usually encountered during unsafe driving and there is strong correlation between fatigue, drowsiness, visual distraction and drivers' facial cues [18]. Among those as mentioned earlier non-invasive attentional cues we also found that, the percentage of eyelid closure over time (PERCLOS) [19], frequency of yawning [20] and gaze direction [21] to be most useful indicators for monitoring drivers' attention. Thus, in this work, we have used these parameters to estimate the driver's attentional state. The proposed framework is similar in its concept to our previous framework developed in [7]. However, the main differences between them are in terms of function and data. Our previous framework is developed for a real-time attentional state detection only and assessed with a limited number of tests. This framework is developed for vision-based attention monitoring of drivers in real-time. This work demonstrated how the proposed framework classifies the drivers' attentional state and measures the level of attention by extracting visual cues. Moreover, the proposed system evaluated in the real driving scenario under diverse situations (i.e., drivers having different facial features such as beard, moustache, and hairstyles or wearing accessories) proves its efficiency and accuracy.

A number of approaches have been taken to predict driver attention state for providing traffic safety and to reduce the number of accidents. However, most of the available systems for driver's attention monitoring purpose usually are either expensive or limited to special high-end car models. These systems cannot be affordable for drivers of low income or developing countries. Thus, an attention monitoring system should develop smart driving assistant which maintain a good balance between affordability and functionality. An effective and used friendly system can save people's lives. The major contributions of this work are:

- Propose a vision-based framework that can constantly track the attentional states and level of attention of the driver.
- Develop an awareness system by generating alarm for the driver if inattentional state is detected.
- Evaluate the performance of the proposed framework in real driving scenarios.

The rest of the paper is organized as follows: Section II presents the related work. Section III provides a brief description of the proposed attention monitoring system. Section IV presents a number of experiments with corresponding results and discusses the future works. Finally, in Section V, the paper is concluded with a brief summary.

II. RELATED WORK

Attention is an important activity of the brain that decreases the information flow into brain's sensory system. It enhances the relevant or vital parts of the input stream and discards disruptions [22]. Zivony *et al.* [23] has investigated the spatial attention which endorses the high-level processing and also identifies some boundary conditions of attentional engagement. Their findings suggested that eye blink interrupted attentional engagement, whereas attentional capture (shifting) was unaffected. Benedetto *et al.* [24] also suggested blink duration (BD) and blink rate (BR) as a more profound and trustworthy indicator of driver's visual workload.

In the past decade, detection of attention of driver's has become an active research field. A broad review of different approaches for attention detection has been reported in [5]. These approaches are grouped into five categories such as subjective, physiological, vehicle-based, visual behavioral, and hybrid. Subjective approaches involve detection

of driver's inattention through questionnaire, and feedback are collected as rating scores [25]. However, the subjective approaches are not effective in detecting driver's inattention in a real time, rather these are effective to cross validate the accuracy of other approaches. Physiological approaches depend on some vital information, such as heart rate, brain activity, skin conductance, etc. These approaches typically detect hyper vigilance in a simulated human-machine system based on physiological signals, such as RRV, ECG, EEG, and EOG [26]–[30]. These systems either use wires and electrodes running from the driver to the system causing distraction and annoyance to the drivers [29] or expensive wearable respiratory inductive plethysmography (RIP) band [26], wireless headsets [27], [28] and EEG acquisition equipment [30]. Vehicle-based approaches involve evaluating driving behaviours, such as steering wheel movements, changes in acceleration/speed, and lane position changing and braking patterns over time for detecting inattentiveness [31], [32]. Detecting inattentiveness based on driving behavior is not fairly reliable because the level of errors may vary from person to person. Visual behavior-based approaches involve extraction of visual features of the driver and have been used widely and effectively to detect driver's inattentiveness [33]. For example, Ramírez *et al.* [34] and Takemura *et al.* [35] proposed a method that uses a head mounted sensors. Some recent researches have been conducted for driver's attention detection by combining various types of information, for example, combining driver's physiological signals and visual signal [36], driver's physiological signals and driving contexts [37] or driver's visual cues and driving patterns [38].

Both researchers and drivers found non-intrusive vision-based systems appealing for attentional state detection and monitoring. Recently, Gumaei *et al.* [16] developed a useful camera-based framework for real-time identification of drivers' distraction by using two different deep learning models, a custom deep convolutional neural network (CDCNN) model and a visual geometry group-16 (VGG16)-based fine-tuned model and classified drivers' behaviour in 10 categories. An Interwoven Convolutional Neural Network (Inter-CNN) was proposed by Zhang *et al.* [15] to classify driver behaviors in real-time. This system can classify 9 different behaviors which are the most frequently encountered cases of distracted behavior during driving. In [6], a fatigue detection system was proposed that classifies the behavior into three category such as normal, warning and danger based on eye gaze in real time. Tran *et al.* [17] proposed a rel-time driver distraction detection system that is able to identify 10 types of distractions through multiple cameras and microphone and also alerts through a voice message. Alam and Hoque [7], [8] proposed a system that estimates attentional states of driver based on various visual cues. Shibli *et al.* [9] estimated level of attention and detected fatigue during driving based on assessing eye aspect ratio (EAR) and head pose. Chien *et al.* [12] proposed a system to detect situations when the driver's eyes exhibit distraction for a long duration and generates an alarm. Mandal *et al.* [10] employed a classifier to identify drivers' state based on PERCLOS. Chowdhury *et al.* [13] proposed a framework to estimate driver's attention in terms of facial angle and lip motion. Vicente *et al.* [14] reported a system based on head pose and gaze estimation that detects Eyes off the Road (EoR). Most of these researches investigated either distraction or sleepiness using only one or two visual parameters for detection of drivers' attention. To solve the inconvenience caused by physiological approaches in [11] a vision-based physiological signal measurement system was proposed to estimate driver fatigue. The system uses only one camera to collect physiological information (e.g. remote photoplethysmography (rPPG) signal) to estimate heart rate (HR), pulse rate variability (PRV), and facial features to estimate the percentage of eyelid closure (PERCLOS) and yawning state to measure the fatigue state of the driver. The system was developed and tested in a controlled indoor simulated driving environment with sufficient light to avoid the interference from the external environment and ambient light with a high-resolution camera. Recommended condition for rPPG estimation requires a good lighting condition with high resolution and uncompressed camera. Therefore, when the lighting condition is flawed in a real driving scenario, rPPG based system may not function properly.

Currently available systems are either expensive and limited to special high-end car models or affordable solutions that lack accuracy and robustness. That is what motivated us to focus on implementing a driver's attention monitoring system to bridge the gap between affordability and availability with functionality. In this research, we focused on developing a vision-based system that extracts driver's attentional cues/features and classify them into attentive, drowsy, fatigue and distracted. The system also alerts the driver in the event of any of inattentional state such as drowsy, fatigue or distraction.

III. PROPOSED ATTENTION MONITORING FRAMEWORK

The primary goal of this research is to develop a system that determines the attentional states of the driver during driving. In this work, we considered cues related to eyes, mouth and head region. Fig. 1 demonstrates the schematic illustration of the proposed attention monitoring framework.

Drivers' attention monitoring starts with capturing video input of driver's frontal face for visual cues using a general purpose webcam (Logitech C170) placed at a distance $(0.6m - 0.9m)$ from the driver's face on the vehicle fascia (as shown in Fig. 2). The captured video sequence is sent to the next module for further processing.

Isolating driver's face region throughout the monitoring process is the first important step. So, video sequence is divided into frames, each frame is converted into gray scale and then the face detection and tracking are performed. The Viola-Jones face detection algorithm [39] is employed to detect faces. For each face in the frame, the algorithm returns the positions of the detected face as a rectangle. When more than one face is detected, the largest rectangle

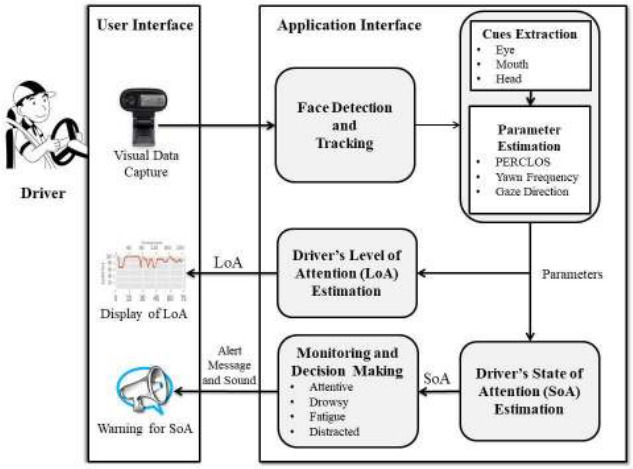


FIGURE 1. Schematic representation of driver's attention monitoring framework.

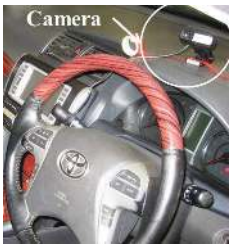


FIGURE 2. Webcam on vehicle fascia for capturing facial cues.

(a) (b)

FIGURE 3. Output of the system, when (a) detecting and (b) tracking the face.

(closest face) is determined using the integral image and is marked as driver's face. As we need to extract cues from the face region to characterize attentional states, analysis of the face region is needed to be done for each frame. But detecting the face for every frame is computationally expensive, so we implemented face tracking algorithm proposed by Danelljan *et al.* [40] once the face is detected. For each frame, a correlation tracker is used to keep record of the tracking quality. Depending on tracking quality, a rectangle image around the face is drawn to indicate that the tracker is following the face. Face detection and tracking starts again if the face is lost or the tracking quality falls. Fig. 3 shows the output from this module. The detected face region is then sent to cues extraction and parameter estimation modules.

A. CUES EXTRACTION AND PARAMETER ESTIMATION

PERCLOS, yawn frequency and gaze direction are visual cue-based parameters which are used to estimate driver's attentional state. These parameters need extraction of

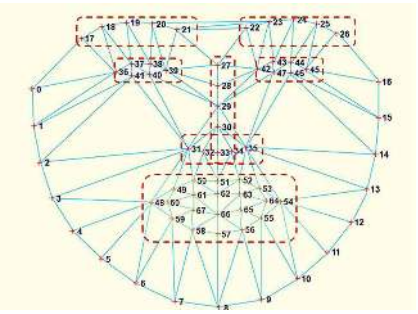


FIGURE 4. Visualization of the 68 facial landmarks.

attentional cues from the face region such as eyes, mouth and head region. Cues indicating drowsiness, fatigue and distraction appear mostly in the eye regions. Thus, cues extracted from eye regions are used to estimate the parameters PERCLOS and eye gaze. Another key cue of fatigue is excessive yawning which is extracted from mouth region to estimate yawn frequency. Usually, normal driving head/face orientation is frontal. Therefore, the deviation of head/face orientation from the standard direction for a substantial time is classified as distraction. The cues related to head position is used to estimate gaze direction. However, the face detector module returns only face region based on which we can infer an approximate orientation. A facial landmark detector proposed by Kazemi and Sullivan [41] is used to extract the additional information, and trained on iBUG300-W data set [42]. Visualization of the 68 facial landmarks from the iBUG 300-W dataset is given in Fig. 4.

For each frame, the face region is found from the face detection and tracking module, facial landmark detector is used to locate salient regions of interest (ROI) of the face to extract cues and the cues are partitioned into three broad classes depending on the parameters PERCLOS, yawn frequency, and gaze direction to estimate driver's attentional states.

1) ESTIMATION OF PERCLOS

It is defined as the percentage of duration of eye being closed over time interval T_1 (60 seconds), excluding the eye blinks and is defined by (1).

$$PERCLOS = \frac{t}{T_1} \times 100\%, \tag{1}$$

where t is the duration of closed-eye state.

Drowsiness and fatigue can also be detected from *PERCLOS* and the period of closed-eyes state. Computation of duration of closed-eye state needs to be determined from the eye state (open or closed). The eye state is also used to detect eye blinks and to measure *PERCLOS* in overlapped time window of one minute continuously. For each frame (f), the eye state is estimated by the measure of eye aspect ratio (*EAR*). The *EAR* is an estimate of the eye opening state [43]. The *EAR* demonstrates a constant value in the case of open-eye state, but quickly changes to 0 when the eye is in closed state. *EAR* shows the correlation between width and

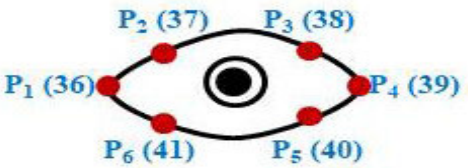


FIGURE 5. Point related to right eye region localized by facial landmark detector.

height of the eye in terms of proportion. A threshold value has been set based on the definition of closed-eye state, i.e., the eyes are 80% or more occluded. Fig. 5 shows the six points of right eye used from Fig. 4 to calculate *EAR* extracted by the landmark detector. After the detection of six (x, y)-coordinates of the right and left eye, points are individually passed to estimate *EAR*. *EAR* for each eye i is calculated using (2).

$$EAR_i = \frac{|P_2 - P_6| + |P_3 - P_5|}{2|P_1 - P_4|}, \tag{2}$$

where $P_1, \ldots, P_6$ are 2-D points rendered in Fig. 5.

The average *EAR* of both eyes is calculated using (3).

$$EAR = \frac{EAR_{Right} + EAR_{Left}}{2}. \tag{3}$$

2) ESTIMATION OF YAWN FREQUENCY

Excessive yawning is associated with the fatigue. Yawning frequency (*YF*) for a time widow of 60 seconds can be estimated by detecting the yawning and incriminating the corresponding counter (*YN*) using the equation defined in (4).

$$YF = YN_{T_2}. \tag{4}$$

where T_2 is the time window. Estimation of yawn frequency need to isolate the mouth region (MR) from the rest of the image (see Fig. 6) and is done using equations (5)-(6).

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x + \frac{w}{4} \\ y + \frac{11h}{16} \end{bmatrix}, \tag{5}$$

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x + \frac{3w}{4} \\ y + h \end{bmatrix}, \tag{6}$$

where (x, y) is initial point of detected face; w and h are the width and height.

After isolating *MR*, the system detects the width of the opening of the mouth due to yawning. The area of the mouth is determined by performing gray scale conversion, histogram equalization and finally an uneven segmentation (S_{MR}) of the unlighted region of *MR* (i.e., inner part) is computed using the threshold (τ), which is set using Otsu's method [44]. The mouth area and the contour of the mouth are measured using the algorithm suggested by Suzuki and Abe [45]. The yawn is presumed to a wide mouth aperture vertically. Thus, the area of the mouth contour tends to extend in successive frames while yawning. Fig. 7 illustrates the detected *MR* by the system.

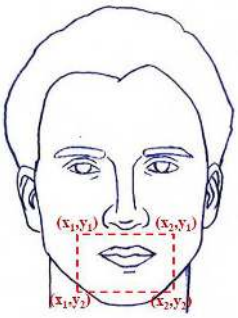


FIGURE 6. Mouth Region (MR) for yawn detection.

(a) (b)

FIGURE 7. Mouth region detected by the proposed system when mouth is (a) closed and (b) open.

Ratio (R_M) of the width (W_M) to the height (H_M) of *MR* describes the rate of increase and it used as an indication of yawning. Therefore, R_M can be defined by (7).

$$R_M = \frac{H_M}{W_M}. \tag{7}$$

R_M is low (i.e., ($\leq Th_Y$)) when mouth is closed and vice-verse. Here, threshold value Th_Y is calculated empirically. Driver is considered to be yawning if a significant number of successive frames (e.g. for 3s) of mouth state is found open. *YN* denotes the number of yawning which is initially 0 and is incremented whenever yawn is detected.

3) ESTIMATION OF GAZE DIRECTION

Gaze direction (*GD*) is used to detect distraction state of the driver. Both face direction (*FD*) and eye gaze direction (*EGD*) are taken into account to estimate *GD*. Usually, standard *GD* of a driver should be frontal. Deviation from the typical position for long period of time (T_3) is an indication of distraction. *GD* can be calculated by (8).

$$GD = \{FD, EGD\}_{T_3}. \tag{8}$$

Eye centers are detected at first to estimate the *GD*. To simplify the eye center detection, some prepossessing are performed on both eyes individually. First, the region of eyes isolated utilizing the six (x, y)-coordinates (see Fig. 5). Then, by using bi-linear interpolation method, eye regions are resized. To improve the contrast in the image, histogram equalization is performed and skin pixels are eliminated by setting a threshold depending on the highest count and the dimension of the equalized image. Finally, erosion followed by dilation is performed for noise removal. Fig. 8 shows the

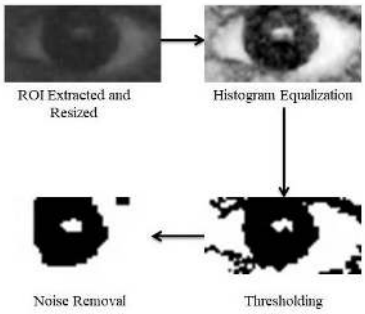


FIGURE 8. Eye center detection processes.

detection process of the right eye. After the pre-processing of the eye image, the visible eyeball area is considered as an ellipse. To identify the outermost border of this ellipse border following an algorithm by Suzuki and Abe [45] is used and then Douglas-Peucker approximation algorithm [46] is utilized to reduce the number of points in the curve. Then, the center $(\bar{x}, \bar{y})$ of the ellipse is estimated using moments [47] and is used to indicate the center of the eye. Once the centers of both eyes are calculated, *EGD* is categorized into right (i.e. $\theta > 8°$), left (i.e. $-8° < \theta$) and front (otherwise) in terms of a threshold value, θ. Here, $\theta = \tan^{-1}(\frac{\triangle x}{\triangle y})$ and $\triangle x = \bar{x_R} - \bar{x_L}, \triangle y = \bar{y_R} - \bar{y_L}$, where, $(\bar{x_L}, \bar{y_L})$ and left and right eye center corresponds to $(\bar{x_R}, \bar{y_R})$.Blue lines which started from center of eye and connected together on the other end represents estimated *EGD* in Fig 9b.

Head orientation plays an important role to indicate driver's attention state. It is connected with the eye gaze direction to determine the person's field of view. Head orientation is estimated by detecting the 15 points marked as red dots in Fig. 9a. At first, the system detected the 15 points (Fig. 9b) and then the standard solution of Perspective-n-Point (PnP) problem [48] is used to determine the head orientation which can be represented as (9).

$$h = (r, t)^T, \tag{9}$$

where h is 3-D head orientation comprises of rotations, $r = (r_x, r_y, r_z)^T$, and translations, $t = (t_x, t_y, t_z)^T$. The 3-D axis is represented as red, green, and blue color in Fig. 9b. The perspective transformation is performed by (10).

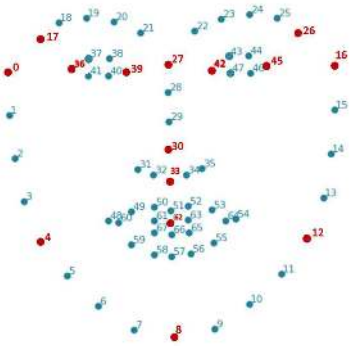
$$s[p, 1]^T = A[R|t]P^T, \tag{10}$$

where s is a scaling factor, A is a camera matrix and $[R|t]$ is joint rotation-translation matrix. Rotations, $r = (r_x, r_y, r_z)$ is determined using the Rodrigues rotation in (11).

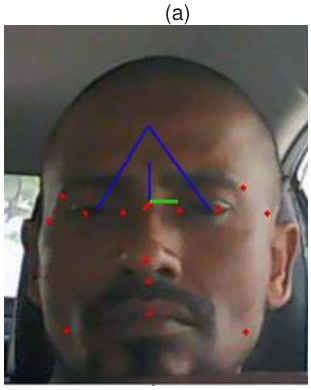
$$R = \cos\theta I + (1 - \cos\theta) r r^T + sin\theta \begin{bmatrix} 0 & -r_z & r_y \\ r_z & 0 & -r_x \\ -r_y & r_x & 0 \end{bmatrix}, \tag{11}$$

where I is a vector in $\mathbb{R}^3$ and $\theta = \|r\|_2$.

The Euler angle yaw (α) is obtained from the vector, r to estimate the face direction(*FD*) to the left ($-90° \leq \alpha < -30°$), right ($30° < \alpha \leq 90°$) and front (otherwise). Our system appears to be operative for a α range



(a)



(b)

FIGURE 9. (a) Points used to calculate the head pose which has marked with red dots and (b) points are detected and estimated GD by the system when EGD and FD are front.

TABLE 1. Parameters used to driver's attentional state monitoring system with notation and observation time.

Parameter	Notation	Details	Time (Sec)
PERCLOS	t	Duration of closed eye state	T_1= 60
Yawn Frequency	Th_Y	0.04	T_2= 60
Gaze Direction	FD EGD	Right	T_3= 2
		Left	
		Front	

of $[-90°, +90°]$ centered at frontal and *FD* detection rate drops significantly when α exceeds the range. Fig. 9b shows an example of gaze estimation of the proposed system.

Table 1 summarizes the values of notations used to estimate the parameters along with the observation time.

B. DRIVER's STATE OF ATTENTION ESTIMATION

Estimated parameters *PERCLOS*, *YF* and *GD* are used to estimate the driver's state of attention (*SoA*). *PERCLOS* is considered to be the most effective visual feature-based measurement of drivers' attentional state in terms of sleepiness detection. Table 2 presents the classification of driver's attentional states which were experimentally found by Jiménez-Pinto and Torres-Torriti [49] on a time window in one minute.

In this work, we characterized the fatigue of the driver's using the parameter *YF*. High *YF* (1-4 yawns per minute) is

TABLE 2. Classification of driver's attentional state in terms of PERCLOS values in 60 seconds windows.

State	$PERCLOS_{MIN}$	$PERCLOS_{MAX}$
Fully awake	0.000	0.048
Fatigue (semi-drowsy)	0.048	0.125
Drowsy	0.125	1.000

the indicator of fatigue [20]. Estimated GD is used to classify driver's attentional states to distracted state. Normally *GD* of driver should be frontal but if *GD* is in the other direction over a time period, it is presumed to be distraction. However, while driving if the driver needs to do some in-vehicle and outside vehicle viewing tasks, which requires him/her to change his gaze direction. So keeping those visually demanding task in mind, we used the *GD* estimated in 2 seconds window according to ISO standard (or more details, see [50]). Based on the above discussion, a procedure to estimate *SoA* is presented in Algorithm 1.

Algorithm 1 Estimate State of Attention (SoA) for Frame, f

Require: $PERCLOS$, YF, GD
  if $PERCLOS \geq 0.125$ **then**
    $SoA \leftarrow Drowsy$
  else if $YF > 1$ or $0.048 \leq PERCLOS < 0.125$ **then**
    $SoA \leftarrow Fatigue$
  else if GD is *Right* or GD is *Left* **then**
    $SoA \leftarrow Distracted$
  else
    $SoA \leftarrow Attentive$
  end if
  Return SoA

C. MONITORING AND DECISION MAKING

The system makes the decision based on the estimated *SoA* for the previous frames. Based on the value of *SoA*, this module generates an alert (sound and message signal) for the driver. The system generates an alarm if the driver is found in any of the inattentive (i.e. drowsy, fatigue, and distracted) SoA. The alarm system trigger a sound and display a message to alert the driver. Fig. 10 shows the output when drowsy, fatigue, and distracted SoA is detected. Alarm deactivates automatically when the driver gets back to the desired SoA (i.e. attentive). Three different types of beep sounds are used for each drowsy, fatigue, and distracted SoA. Fig. 10a illustrates an warning message, ''DROWSINESS ALERT'' is triggered and displayed when the system estimated that the driver is in drowsy state according to Algorithm 1. Similarly warning message for fatigue and distraction detection is displayed (Fig. 10b and Fig. 10c).

D. DRIVER's LEVEL OF ATTENTION ESTIMATION

Parameter values of $PERCLOS_f$, YF_f and GD_f for each frame, f computed from the previous module cue extraction and parameter estimation. These values are scaled down to a range [0,1] for the measured time series using equations (12)-(14). The maximum and minimum values of each parameter is stored and updated from the beginning of monitoring.

$$PERCLOS_{Scale\_f} = \frac{PERCLOS_f - min(PERCLOS)}{max(PERCLOS) - min(PERCLOS)}, \tag{12}$$

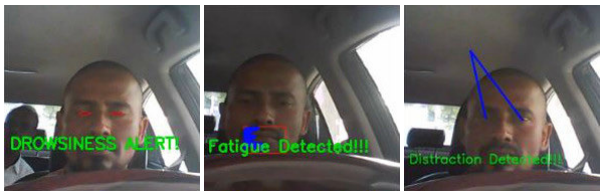
$$YF_{Scale\_f} = \frac{YF_f - min(YF)}{max(YF) - min(YF)}, \tag{13}$$

$$GD_{Scale\_f} = \frac{GD_f - min(GD)}{max(GD) - min(GD)}. \tag{14}$$

Driver's level of attention (*LoA*) is then calculated using redistributed values according to Eq. 15.

$$LoA = \frac{PERCLOS_{Scale\_f} + YF_{Scale\_f} + GD_{Scale\_f}}{3} \times 100\%. \tag{15}$$

A real-time graph (see Fig. 15, 16 and 17) is generated to show the Driver's *LoA* along with live plotting of frame per second (*FPS*), *PERCLOS*, mouth opening ratio, and head and eye position in terms of degree for each frame.



(a) (b) (c)

FIGURE 10. Messages and triggering sound: (a) Drowsy SoA, (b) Fatigue SoA, and (c) Distraction SoA.

IV. EXPERIMENTAL ANALYSIS

To evaluate the proposed framework, three experiments were carried out. Several measures such as false positive rate (FPR), false negetive rate (FNR), accuracy, and processing time have been investigated to assess the performance of the system in real driving scenarios.

A. EXPERIMENTAL SETUP AND IMPLEMENTATION DETAILS

The system is implemented and tested in Python using OpenCV and Dlib libraries on a general-purpose laptop with Intel Core i5-4200U processor. The laptop used 4GB RAM and a 64-bit Windows 10 operating system. A CMOS web-cam (Logitech C170) was connected to the laptop through USB 2.0 port. Detail specification of the used camera is given in Table 3. Considering the video capturing resolution and focal length of the camera, it was mounted perpendicularly in front of the participant' face at a distance of (0.8m) from the car dashboard. During evaluation, the participants were requested to seat in the driving seat in front of the camera.

TABLE 3. Specification of Logitech C170.

Feature	Description
Sensor type	VGA sensor
Video capture resolution	1024 x 768
Focus type	Fixed
Focal length	2.3 mm
Maximum frame rate	30 fps
Additional features	Logitech fluid crystal technology

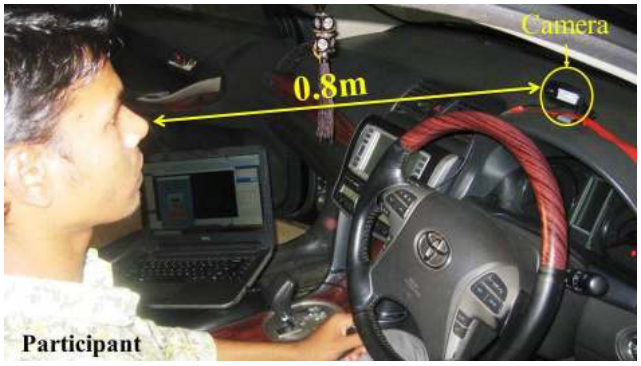


FIGURE 11. Experimental setup.

Fig. 11 illustrates the setting of the experiment. A Toyota Allion of 2011 car was used in the experimental trials.

B. PARTICIPANTS

A total of 10 healthy drivers (mean age = 31.6, SD = 4.45) with different facial features (beard, wearing glasses, and moustache), and hairstyles were participated in the experiments. Seven of them were male and three were female. All the participants were in good health condition and consented before participating. We gave a brief description of the system and explain the whole procedure before conducting the experiments.

C. EXPERIMENT 1: TO VALIDATE THE ESTIMATED PARAMETERS

To investigate the validity of the estimated parameters PERCLOS, YF, and GD, we conducted an experiment with different participants and considering various lighting conditions.

1) EXPERIMENTAL PROCEDURE AND DATA COLLECTION

Three participants (one female and two male) took part in the experiment with distinct facial appearances and hairstyles. During the test, each participant was asked to spend approximately 5 minutes with the system. All of them participated in bared faces and also wore accessories, such as sunglasses, spectacles, and caps. A total of 17 minutes 40 seconds (5 minutes 35 seconds for Participant-1 + 6 minutes for Participant-2 + 6 minutes 5 seconds for Participant-3) long videos with involuntary eye blink and spontaneous yawns were captured by the camera in front of the driver to verify the estimated parameters' correctness under different daylight conditions. Ideally the frame rate of the camera is 30 fps but practically the system cannot capture the video at this rate due to the slow video capturing hardware, the contents of the video, and the computational overload. A total of 10, 805 frames (at an average of 10 fps) were captured and analyzed for this experiment. Table 4 and Table 5 provide the number of frames analyzed separately in each case.

2) EVALUATION MEASURES

Validation of parameters are done based on two criteria: false positive rate (*FPR*) and false negative rate (*FNR*) defined in (16) and (17).

$$FPR = \frac{F_{Pos}}{(F_{Pos} + T_{Neg})} \tag{16}$$

$$FNR = \frac{F_{Neg}}{(F_{Neg} + T_{Pos})} \tag{17}$$

Here,

- T_{Pos} is number of positive instances in a testing sequence which were extracted cues correctly recognized by the system (For example, driver is yawning and system detected is correctly);
- T_{Neg} is number of negative instances in a testing sequence which were correctly recognized by the system (For example, driver's eye are close and system detected them closed);
- F_{Pos} is number of positive instances in a testing sequence which were extracted cues wrongly recognized by the system (For example, driver is yawning but system was unable to detect it);
- F_{Neg} is number of negative instances in a testing sequence which were wrongly recognized by the system (For example, driver's eye are close but system detected them them as open);

3) RESULTS

Accuracy of PERCLOS, YF and GD are dependent on the accuracy of extracting cues such as eye state, yawn frequency, eye gaze direction and face direction respectively and estimating SoA and LoA. At the first stage, techniques to extract the cues were evaluated and then the video sequences were investigated to validate the estimated parameters subjectively.

a: FACE DETECTION, TRACKING AND ROI EXTRACTION

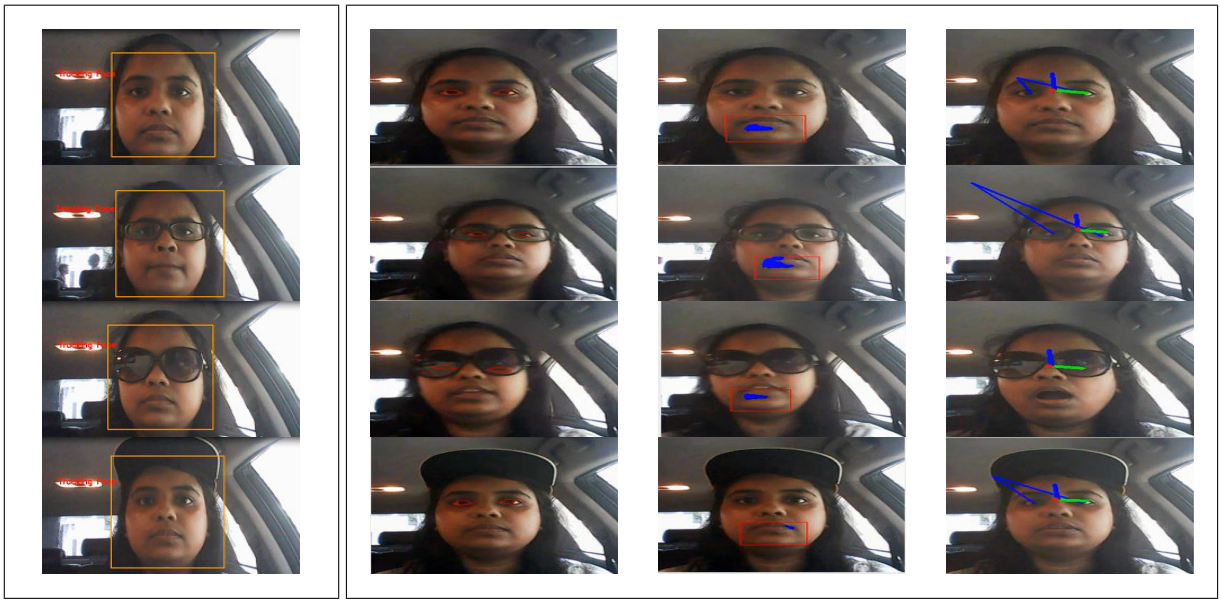
Test was conducted on different participants with different situations to establish the system's (i) face detection and tracking and (ii) region of interest (ROI) (i.e. eye, mouth and head region) extraction capabilities. Fig. 12 shows some sample frames where face of subject with different situation is detected and tracked and ROI is extracted. Observations from this test suggest that algorithm used for detection, tracking and extraction are functioning satisfactory to accommodate participants with different situation.

TABLE 4. Detailed data size of Experiment 1 in different situation.

Situation	No. of frames for each situation			Total no. of frames for each situation
	Participant-1	Participant-2	Participant-3	
Participant without any accessories	1103	1032	1250	3385
Participant with Spectacles	959	825	1065	2849
Participant with Sunglass	678	542	730	1950
Participant wearing Cap	911	989	721	2621
No. of frames for each participants	3651	3388	3766	10805
Duration of captured video	335 sec @ 11 fps (approx.)	360 sec @ 9 fps (approx.)	365 sec @ 10 fps (approx.)	17 min 40 sec

TABLE 5. Detailed data size of Experiment 1 under different lighting condition.

Situation	No. of frames for each situation			Total no. of frames for each situation
	Participant-1	Participant-2	Participant-3	
Broad Day Light	2115	1932	1990	6037
Parking Garage	1536	1456	1776	4768
No. of frames for each participants	3651	3388	3766	10805
Duration of captured video	335 sec @ 11 fps (approx.)	360 sec @ 9 fps (approx.)	365 sec @ 10 fps (approx.)	17 min 40 sec



(a) (b)

FIGURE 12. (a) Face detection and tracking, and (b) ROI extraction of a participant bear faced and wearing different accessories such as spectacles, sun glass and cap.

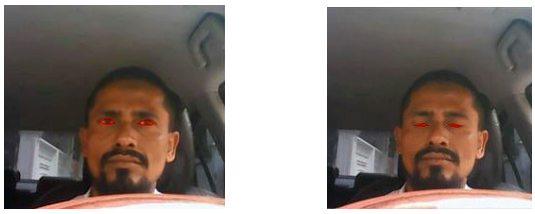
b: CUES EXTRACTION

The correctness of the extracted cues (i.e., eye state, yawn detection, eye gaze direction and face direction) is investigated by *FPR* and *FNR*.

- In case of eye state detection, *FPR* error is occurred when eyes are in open-state but detected as close-state and *FNR* error is occurred when eyes are in close-state but the system detected as in open-state. Table 6 represents the percentages of *FPR* and *FNR*. Results suggest that wearing cap doesn't affect much in detecting eye state but *FNR* of detecting eye state is greater when wearing spectacles than without wearing spectacles. Our system is unable to detect eye state when subject wears sun glass. Eye state detection was not effected by lighting condition. Fig. 13 shows the detection of eye state (both closed and open) by the system.

TABLE 6. Summary of eye state detection.

Situation	FPR (%)	FNR (%)
Subject without spectacles	1.2	3.6
Subject with spectacles	35	18.2
Subject wearing cap	1.8	4.1



(a) (b)

FIGURE 13. Eye state detection by system: (a) Open eyes (b) Closed eyes.

TABLE 7. Summary of yawning detection.

Lighting condition	FPR (%)	FNR(%)
Broad day light	3.5	9.3
Parking garage	7.4	12.5

TABLE 8. Experimental outcomes for eye gaze detection.

Situation	FPR (%)	FNR (%)
Participant without spectacles	3.8	5.1
Participant with spectacles	16.7	12.5
Participant wearing cap	3	5.7

- To calculate the number of yawn, false positive error is occurred when yawn doesn't happen but the system detects as yawn, and false negative error is occurred when yawn occurs but the system did detect it. Summary of the yawn detection under different lighting conditions is shown in Table 7. Results revealed that wearing spectacles, sun glass or cap doesn't affect the yawn detection. But varying lighting condition affected the yawn detection.
- In case of eye gaze direction and face direction detection, false positive occurs when system detected participants' head/eye in standard direction but they are in different direction. On the other hand, false negative occurred when head or eye is in standard position but classified as left or right by the system. Table 8 and Table 9 shows the percentages of *FPR* and *FNR* for eye gaze direction and face direction detection module. Like in eye state detection, the participant wearing spectacle affected the performance of eye gaze detection algorithm whereas face direction detection algorithm is not affected by any of the situation much.

  We also performed an analysis on determining the validity of the detection angle of FD. To detect the head orientation for FD estimation, we used a standard solution of Perspective-n-Point (PnP) problem [48].To verify the range of detection angle we asked a participant to rotate her face left and right to make different head orientations. A total of 2.5 minutes video sequence (i.e. around 1650 frames) was analyzed to observe the estimated Euler angle yaw (α) for different head orientations. The result suggests that implemented head orientation algorithm can estimate α with a range of $-90°$ to $+90°$ efficiently. However, beyond this range, it is unable to detect head so could not estimate α. Fig. 14 shows head orientation at varying yaw angles. Therefore, we set the detection angle for FD in between $-90°$ and $90°$.

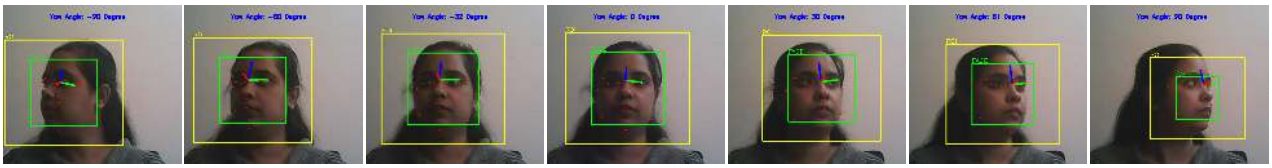
TABLE 9. Experimental results for face direction detection.

Situation	FPR (%)	FNR (%)
Participant without spectacles	6.2	4.8
Participant with spectacles	7.1	4.2
Participant with sunglass	5.8	3.5
Participant wearing cap	6.8	5.1

c: ESTIMATION OF SoA AND LoA

To investigate the the effectiveness of SoA and LoA, we analyzed the video sequences captured by the system and studied State of Attention (SoA) and Level of Attention (LoA) from the video frames. We mainly studied: (1) the validity of the parameters for estimation of SoA and LoA; and (2) the correlation between parameters for estimation of SoA and LoA. For example, Fig. 15, 16 and 17 represent sample snapshots of our monitoring system where fatigue, drowsiness, and distraction can be detected. The analysis is presented in three subsections: (a) SoA and LoA based on PERCLOS; (b) SoA and LoA based on PERCLOS and YF; (c) SoA and LoA based on GD.

1) **SoA and LoA based on PERCLOS:** The objective of this experiment is to estimate the SoA and LoA from PERCLOS. The SoA is classified into attentive, fatigue and drowsiness states according to Algorithm 1. Both fatigue and drowsiness are function of PERCLOS. If the SoA is in any of the non-attentive states (i.e. fatigue, drowsiness) the monitoring system will generate an alarm displaying a warning message and a beep sound. Fig. 15a presents a sample snapshot of the monitoring system showing the analysis of frame 33. It is showing from the top the frame rate (solid blue line), PERCLOS (solid blue line), R_M (solid blue line), FD (solid green line) and EGD (solid blue line), and LoA (solid red line) for over a period of 60 seconds in the subgraphs respectively. In the graph, time in seconds is shown (as secondary axis) on the top of the graph, and frame number is shown (as primary axis) at the bottom of the graph. No yawning (second label from top marked with red circle) is detected during this period and gaze direction (third label from top marked with red circle) is frontal for this frame. PERCLOS is computed using (1) and the SoA is estimated according to Algorithm 1. The LoA is given in percentage (%) and computed using (15).



(a) $-90°$ (b) $-60°$ (c) $-32°$ (d) $0°$ (e) $30°$ (f) $61°$ (g) $90°$

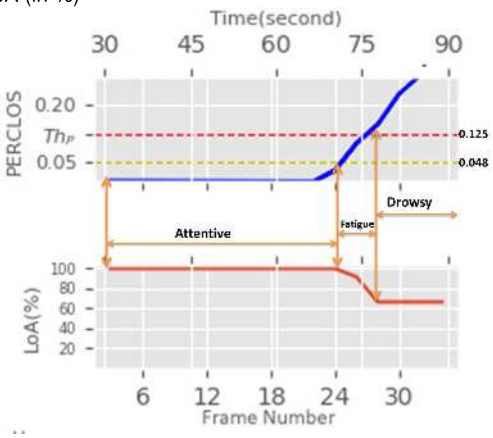
FIGURE 14. Yaw estimation for different head orientation.

a) **Estimation of SoA:** Fig. 15b presents the PERCLOS values (solid blue line) over time (shown on the top) and frames (shown at the bottom). The PERCLOS values define three distinct SoA: attentive when the PERCLOS value is below the threshold value of 0.048 (dotted yellow line), fatigue when the PERCLOS value is equal to or above the threshold value 0.048 and drowsy when the PERCLOS value is equal to or above the threshold value of 0.125 ($Th_p = 0.125$, dotted red line). The driver was in attentive state until frame number 24 (i.e. around 72 second) when PERCLOS value was below 0.048. However, after frame number 24 (i.e. around 72 second 72 second), PERCLOS value has a significant increase and goes over the threshold value ($Th_p = 0.125$) after frame number 27 (i.e. around 76 second), which falls in a drowsy state and resulted in generating an alarm. There is an intermediate fatigue state shown in Fig. 15b, but the PERCLOS value was within this region for a very short time. Therefore, no alarm was generated.

b) **Estimation of LoA:** The lower part of Fig. 15b presents LoA (in %). There is a correlation between PERCLOS value and LoA. The driver has the highest LoA (100%) until frame 24 (i.e. around 72 second) and started falling as the PERCLOS value increased above $Th_p = 0.125$ and LoA reached 66% at frame number 27 (i.e. around 76 second).

2) **SoA and LoA based on PERCLOS and YF:** The objective of this experiment is to estimate the SoA and LoA from PERCLOS and YF. A sample snapshot of our monitoring system for frame number 78 is given in Fig. 16a. It is showing from the top the frame rate (solid blue line), PERCLOS (solid blue line), R_M (solid blue line), FD (solid green line) and EGD (solid blue line), and LoA (solid red line) for over a period of 160 seconds in the subgraphs respectively. In the graph, time in seconds is shown (as secondary axis) on the top of the graph, and frame number is shown (as primary axis) at the bottom of the graph. From this figure, we can see the gaze direction (third label from top marked with red circle) is frontal. PERCLOS value is shown (calculated using eq. 1), R_M value is

(a) Analysis of a frame (frame 33): four insets on the left represent face detection and tracking (top), analysis of eye, head and mouth region (bottom) respectively;five sub-graphs in middle represent Frame Rate, PERCLOS, Yawn frequency, Gaze direction, and Level of Attention from top respectively; four labels on the right represents frame number, YF, GD and LoA (in %)

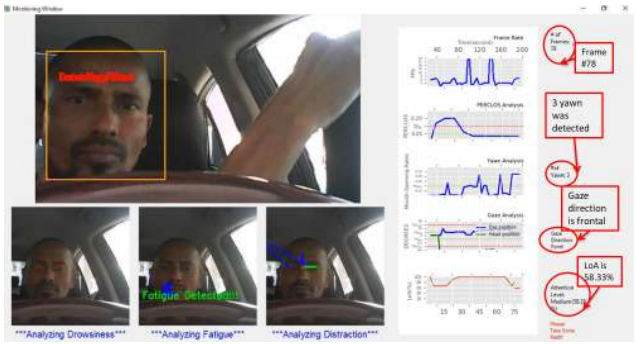


(b) SoA and LoA analysis based on PERCLOS

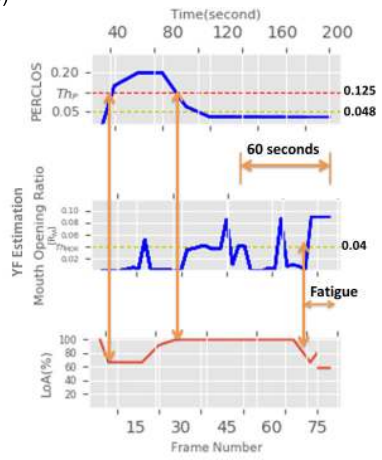
FIGURE 15. Analysis of SoA and LoA based on PERCLOS.

calculated using (7) which is again used in estimating YF according to eq. 4 and SoA is estimated according to Algorithm 1. The LoA is shown in percentage (%) and computed using (15).

a) **Estimation of SoA:** Fig. 16b presents the PERCLOS (top graph) and R_M values (solid blue line) (middle graph) for estimating YF over time (shown on the top) and frames (shown at the bottom). The PERCLOS and YF define two distinct SoA: attentive when the PERCLOS value is below the threshold value of 0.048 (dotted yellow line) and YF is equal to or less than 1, and fatigue when the PERCLOS value is equal to or above the threshold value 0.048 or YF is greater than



(a) Analysis of a frame (frame 78): four insets on the left represent face detection and tracking (top), analysis of eye, head and mouth region (bottom) respectively;five sub-graphs in middle represent Frame Rate, PERCLOS, Yawn frequency, Gaze direction, and Level of Attention from top respectively; four labels on the right represents frame number, YF, GD and LoA (in %)



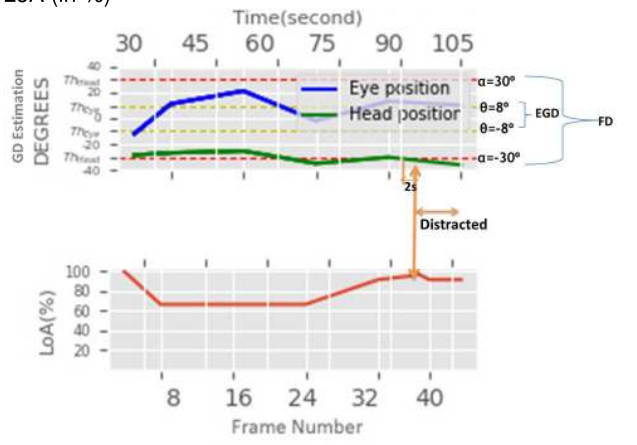
(b) SoA and LoA analysis based on PERCLOS and YF

FIGURE 16. **Analysis of SoA and LoA based on PERCLOS and YF.**

1. We at first analyzed the measured values of PERCLOS. From Fig. 16b we observed that in the first 60s, PERCLOS has significant increase going over the threshold ($Th_p = 0.125$, dotted red line), which resulted in drowsy state. However, to estimate the SoA for this frame we need to observe the PERCLOS measurement for the last 60 seconds i.e. from frame number 45 (at 140 second) (labelled in Fig. 16b). We observed that in the last 60 seconds, PERCLOS value is below 0.048, which represents the attentive state.
YF is another parameter considered to estimate fatigue. Mouth opening ratio (R_M) is used for estimating YF. We can see R_M is showing lots of ups and down throughout the time (160 seconds) under observation. To estimate SoA for current frame, we only need to consider the measurements of the last 60 seconds i.e. from frame number 45 (at 140 second) (labelled in Fig. 16b). We observed that during the last 60 seconds, measurements of R_M is close to the threshold value ($Th_Y = 0.04$) thrice resulting in a YF value of 3 (second label from top marked with red circle in Fig. 16a) for the current frame and identifying SoA as fatigue and generated an alarm.
   b) **Estimation of LoA:** The lower part of Fig. 16b presents LoA (in %). From Fig. 16b we can see that during the first 60 seconds, LoA value (red solid line) is low (66% to 85%) and started to increase when PERCLOS value decreased and reached the threshold value (i.e. under 0.048). Observing the last 60 seconds i.e. from frame number 45 (at 140 second) (labelled in Fig. 16b) we found that LoA started to decrease at frame number 72 (at 180 second) and reached 58.33% at frame number 75 (at 200 second) as three yawn was detected during this time period which is visible from the measurements of R_M.

3) **SoA and LoA based on GD:** The objective of this experiment is to estimate the SoA and LoA from GD. Fig. 17a shows a sample snapshot of our monitoring system for frame number 42. It is showing from the top the frame rate (solid blue line), PERCLOS (solid blue line), R_M (solid blue line), FD (solid green line) and EGD (solid blue line), and LoA (solid red line) for over a period of 60 seconds in the subgraphs respectively. In the graph, time in seconds is shown (as secondary axis) on the top of the graph, and frame number is shown (as primary axis) at the bottom of the graph. No yawning (second label from top marked with red circle) is detected during this period. Gaze direction (third label from top marked using red circle) is left for this frame. Here, we have analyzed the measurements FD and EGD. FD and EGD are estimated using yaw angle α of head and direction of eye center measured using angle θ respectively according to the process described in Section III-A3. GD is calculated using FD and EGD according to (8), which can be left, front or right. SoA is calculated using GD value according to Algorithm 1, which can be attentive or distracted. If SoA is distracted an alarm is generated. The LoA is given in percentage (%) and computed using (15).
   a) **Estimation of SoA:** Fig. 17b presents the FD (solid green line) and EGD (solid blue line) over time (shown on the top) and frames (shown at the bottom) which were used to estimate the GD. The GD values define two distinct SoA: attentive when the GD value is front and distracted when the GD value is either left or right. From Fig. 17b we can see that FD and EGD measurements show a lot of variations i.e. crossing threshold values (yellow dotted lines for EGD and red dotted lines for FD) throughout the whole sequence (i.e. 75 seconds) under consideration but but it requires 2 seconds of observation for GD estimation. Therefore, GD estimation requires observation of past few frames of EGD and FD (i.e. from frame number 39, 95 second). We can see

(a) Analysis of a frame (frame 42): four insets on the left represent face detection and tracking (top), analysis of eye, head and mouth region (bottom) respectively;five sub-graphs in middle represent Frame Rate, PERCLOS, Yawn frequency, Gaze direction, and Level of Attention from top respectively; four labels on the right represents frame number, YF, GD and LoA (in %)



(b) SoA and LoA analysis based on GD

FIGURE 17. Analysis of SoA and LoA based on GD.

participant's EGD measurements were between the threshold range ($-8° < \theta < 8°$), but FD measurements exceed the threshold value (i.e. $-30°$), which resulted GD as left, (third label from top marked with red circle in Fig. 17a). The estimated SoA is distracted, which generates an alarm.

b) **Estimation of LoA:** The lower part of Fig. 17b presents LoA (in %). LoA measures (red solid line) in Fig. 17b also showed variation in values though out the whole sequence with respect to other measured parameters. But observing the last few frames i.e. from frame number 39 (95 second) we can see that after reaching the peak, LoA started to decrease again with respect to measurements of GD i.e. FD and LoA reached 83.33%.

We can conclude from the above investigations: (1) estimation of SoA and LoA are accurate and (2) parameters can independently SoA, but estimation of LoA not only depends on the current measurements of parameter but also on previous values of past frames. It is worth mentioning that the correlation between PERCLOS and the two attentional states (fatigue and drowsiness) needs to be further investigated using human subjects with real sleep deprivation.

D. EXPERIMENT 2: TO EVALUATE ATTENTIONAL STATE

The purpose of this experiment is to measure the accuracy of the proposed system in detecting four kinds of attentional status: attentiveness, drowsiness, fatigue, and distraction.

1) EXPERIMENTAL PROCEDURE AND DATA COLLECTION

The experiment was performed in a controlled environment considering the risk of driver's inattentive state for traffic safety. We requested all the participants to pose various expressions which exhibit different attentional state. Each participant spent approximately 3.5 minutes on average, to simulate each state. A total of 140 minutes (10 [participants] $\times$ 3.5 [interaction time] $\times$ 4 [types of status]) long video sequence was captured by the camera placed perpendicular to the driver's face for analyzing. More than 85K frames (For details, see Table 10) were analyzed for this experiment.

2) EVALUATION MEASURES

We measured the accuracy (A) of each attentional states by using (18).

$$A = \frac{D_F}{T_F} \times 100\%, \tag{18}$$

where D_F is total frame number of correctly detected attentional state and T_F is the total number of frames in a testing sequence.

3) RESULTS

The accuracy of different kind of attentional state shown in Table 10. The results revealed that the proposed system is performing quite satisfactory to estimate the states of attention. The average accuracy to classify the attentional states ranges from 91% to 95%. The system demonstrated relatively higher accuracy of 95% for detecting drowsy state than other three state.As the experiment performed in a controlled environment, the results may vary, especially in real sleeping deprivation.

E. EXPERIMENT 3: TO EVALUATE OVERALL PERFORMANCE

This experiment was conducted to investigate the overall performance of the proposed system concerning some real driving scenarios.

1) EXPERIMENTAL PROCEDURE AND DATA COLLECTION

All the participants were requested to drive the car with an average speed of $25km/hr$ and asked them to yawn and blinks randomly. Driving activity was videotaped by the camera placed in front of the driver to estimate the frame processing time in terms of detection, tracking and extraction and to estimate the overall accuracy. Each trial begun with a predefined positioning of the participant. A total of 4 trials

TABLE 10. Accuracy, *A* (%) of different attentional states.

Participant	Drowsiness			Fatigue			Distraction			Attentiveness			Total no. of
no.	D_F	T_F	A	D_F	T_F	A	D_F	T_F	A	D_F	T_F	A	frames
1	2032	2162	94	2078	2165	96	1888	2075	91	2047	2178	94	8580
2	2062	2104	98	1960	2130	92	1897	2132	89	1972	2120	93	8486
3	1901	2136	89	1890	2148	88	1992	2142	93	1890	2100	90	8526
4	2136	2180	98	1948	2164	90	2027	2156	94	1993	2120	94	8620
5	2140	2140	100	2134	2270	94	2192	2260	97	2212	2280	97	8950
6	1910	1990	96	1890	2100	90	1762	2002	88	1835	2017	91	8109
7	2109	2109	100	1967	2115	93	1917	2130	90	1969	2095	94	8449
8	2069	2201	94	1875	2180	86	1933	2172	89	1917	2130	90	8683
9	1945	2210	88	2016	2215	91	1957	2275	86	1954	2220	88	8920
10	1881	2090	90	2003	2131	94	1887	2029	93	2006	2180	92	8430
		Average	95		Average	91		Average	91		Average	92	85753

TABLE 11. Overall accuracy (A) of the system.

Participant no.	No. of trials	D_F	T_F	A(%)
1	4	10829	11520	94
2	3	9828	10800	91
3	5	14418	16200	89
4	3	10260	10800	95
5	5	19404	19800	98
6	4	12096	14400	84
7	3	7776	8640	90
8	4	12528	14400	87
9	5	13392	14400	93
10	4	15048	15840	95
Total	40	125579	136800	92

was conducted with each participants and each trial took 6 minutes and approximately 4 hours in total. Average total distance travelled by each driver is 10 *km*. More than 130K frames were analyzed to detect the number of frames in which the system showed participants' correct attentional state.

2) RESULTS

The overall performance of the system measured in terms of accuracy and the processing time.

- The overall performance of the framework was calculated using (18). Table 11 shows overall accuracy of the system for each participant individually. Analyzing the data of Table 11, we can see that overall accuracy for this system ranges from 84% to 95% and Participant-6, and Participant-8 have lower accuracy compared to other participants. After analyzing the videos sequences of the participants, we have the following observations:
  - Two out of four trials of Participant-6 was conducted in night environment in city areas with lights coming from city street lights. Thus, sometimes the system estimated incorrect attentional state when it was partially or completely dark due to driving under/on fly-over, broken street light, and blackout. As we used a simple web camera our system is unable to extract facial cues and estimate parameters (*PERCLOS*, *YF*, and *GD*) accurately to classify state when it is partially or completely dark outside, which resulted in low overall performance of the system.
  - Participant-8 wore spectacles while driving throughout all the trials. Due to light on spectacle during broad daylight, the system estimated incorrect attentional state, as cues from eye region play a major role in classifying attentional state. Both estimations *PERCLOS* and *GD* depends on cues from the eye region. Thus, the reflection of light on spectacle led to the wrong estimation of *PERCLOS* and *GD* resulting in decreased performance.

  The above discussion shows that the system's overall accuracy depends directly on the parameters' correct estimation. Results also revealed that the system has an average overall accuracy of 92%.
- The processing time (T_P) for each frame is calculated by (19).

$$T_P = T_{FD} + T_C + T_P + T_{LoA} + T_{SoA}, \tag{19}$$

where T_{FD} and T_C denotes the time elapsed by the system to detect face and extract cues, T_P and T_{LoA} indicates the time spent by system to estimate the parameters and driver's level of attention (*LoA*), and T_{SoA} denotes the time spent by the system to classify driver's attentional state (*SoA*). Table 12 shows the average time taken by each task calculated from collected data. From the analysis of the data, it is revealed that the proposed framework took 107 milliseconds (msec) in total to process each frame resulting in a frame rate of 10 fps.

TABLE 12. Average processing time (in msec).

T_{FD}	T_C	T_P	T_{LoA}	T_{SoA}
13	48	30	7	9

TABLE 13. Summary of non-intrusive visual-feature based systems.

Article	Capturing sensor	Used attentional parameters	Estimated output	Alert system
Gumaei et al. [16]	Camera with Raspberry Pi device edge integrated with a SIM card	Driver's behaviour	Distracted behaviour	None
Zhang et al. [15]	Cameras	Body movements Facial expressions	Driver's activity	None
Wang et al. [6]	Eye tracker	Eye gaze PERCLOS Blink frequency	Fatigue behavior into three classes: normal, warning and danger	None
Tran et al. [17]	Cameras Microphone	Body movements	Normal and distracted behaviors	Voice-alert
Alam et al. [7]	Webcam	PERCLOS Yawn frequency Gaze direction	Attention status: drowsy, fatigue, distracted and normal	None
Alam et al. [8]	Webcam	Gaze direction	Attention status: distracted and normal	None
Shibli et al. [9]	Webcam	EAR Face angle Lip motion	Attention class: no attention and attention	Sound
Chien et al. [12]	Infrared camera	Pupil	Drowsiness	Warning
Mandal et al. [10]	Wide-angle cameras	PERCLOS	Attention class: normal and fatigue	None
Chowdhury et al. [13]	Kinect sensor	Face angle Lip motion	Attention class: no attention and attention	Message and sound
Vicente et al. [14]	Webcam	Head pose	Eye off road	None
Tsai et al. [11]	Webcam	rPPg PERCLOS Yawning state	Fatigue state	None
Proposed System	Webcam	PERCLOS Yawn frequency Gaze direction	Attention status: drowsy, fatigue, distracted and normal	Messages and sound based on SoA

F. DISCUSSION

We developed a real-time non-intrusive vision-based driver's attention monitoring system. The evaluation results suggests that our system performs accurately under different conditions with different participants. Since reported non-intrusive vision-based systems had used different parameters and input devices and also estimated outputs are different we are unable to compare our system with them in terms of performance matrices. However, we summarized the existing non-intrusive vision-based techniques [6]–[10], [12]–[15], [17] along with the proposed system in Table 13. The techniques are summarized in terms of capturing sensor, used attentional parameters, estimated output and alert system.

In the following, a number of key observations are highlighted:

- Capturing sensors used by the reported systems were expensive. Chowdhury *et al.* [13] used a Kinect sensor to estimate driver's attention level and in [16] a fixed specialized camera was used with Raspberry Pi which integrated with a SIM card. Multiple cameras were used by [14], [15], [17] and Chien *et al.* [12] used an IR camera to capture video sequences. Wang *et al.* [6] was prepared using a eye tracker. Similar to us [7]–[9], [11] also used a normal camera, but camera used by Tsai *et al.* [11] and Shibli *et al.* [9] was expensive than the proposed system.
- Although Alam and Hoque [7], [8] used the similar parameters to estimate the attention state but they did not develop any alarm system to alert the driver in the lower level of attention during driving. Also, in [8] only distraction was detected and in [7] the system was evaluated only for three participants. Moreover, their system's overall accuracy is lower than the proposed system. Attentional parameters used by the systems [6], [10], [11], [14] were partially similar to us but estimated outputs of the systems are different, for example systems described in [6], [11] and [10] focused on detecting fatigue behavior only, whereas [14] concentrated on gaze direction/eye of road. Methods in [15] and [17] analyzed driver's body movements

and facial expressions to classify driver's activity. Shibli *et al.* [9] and Chowdhury *et al.* [13] used entirely different parameters such as face angle and lip motion to estimate the attention level.

- The proposed system generates different sound as the alert and message for different inattentional state. The systems developed in [9] and [13] also display warning message and sound similar to the proposed system. System reported in [17] has a voice alert system, but [12] has a alert system which generates warning but type of warning is absent. The systems such as in [7], [8], [10], [11], [14]–[16] have not included any alert system.

Based on the above observations, it is evident that the proposed attention monitoring framework is simple and affordable compared to existing non-intrusive vision-based techniques.

The current version of the proposed system can extract driver's attentional cues/features and classify them into attentive, drowsy, fatigued and distracted, but numerous promising dimensions to be included for further improvements. Following issues can be addressed to make the system more functional:

- Distraction can be categorized into visual (e.g., looking away from roadway focusing on something else), cognitive (e.g., rambling of mind), auditory (e.g., focusing on the ringing cell phone or loud music) and biomedical (e.g., using a cell phone and adjusting audio device) distraction. In this work, we selected visual distraction as we mainly focused on developing a vision-based system and used facial cues to classify driver's attentional state. However, the remaining types of distractions are also major causes of unsafe driving. Therefore, a module may develop to detect various types of distraction, including additional sensors mounted on the car for necessary data accumulation. We expect with the addition of this module will improve the comprehensive functionality and robustness of the attention monitoring system.
- Impairment due to alcohol is another vital factor that causes traffic crashes along with inattentive driving. Thus, alcohol detection module can be incorporated, which will generate an alarm, when the system detects the driver is under the influence of alcohol by analyzing visual characteristics.
- To ensure safe driving and to deal with traffic accidents, a vehicle tracking module may incorporate and establish a connection to the server for effective monitoring and notify traffic incident.

In addition to the above, the proposed system should improve to deal with the driving at night in the highway with no street lights, driving under real sleeping deprivation and different driving positions. A statistical model can be utilized to improve the robustness and accuracy of the current implementation.

V. CONCLUSION

A vision-based framework for determining drivers' attention states is presented in this paper. The framework is able to detect inattentiveness of drivers in day or night environments when dim light is available (i.e., in city areas). Evaluation shows that the proposed framework is functioning well with 92% accuracy in real-time driving scenarios. We plan to add more detection modules such as cognitive, auditory and biomedical distraction detection, alcohol detection, and vehicle tracking and monitoring in future. Additionally, the proposed system may extend to function under more challenging and practical driving scenarios. The proposed system may be installed in vehicles and make a substantial impact to reduce the road crashes and save human lives.

REFERENCES

[1] *The Global Status Report on Road Safety 2018*, World Health Organization (WHO), Geneva, Switzerland, Dec. 2018.

[2] P. T. Gimeno, G. P. Cerezuela, and M. C. Monta nés, "On the concept and measurement of driver drowsiness, fatigue and inattention: Implications for countermeasures," *Int. J. Veh. Des.*, vol. 42, nos. 1–2, pp. 67–68, 2006.

[3] M. A. Regan, C. Hallett, and C. P. Gordon, "Driver distraction and driver inattention: Definition, relationship and taxonomy," *Accident Anal. Prevention*, vol. 43, no. 5, pp. 1771–1781, Sep. 2011.

[4] M. Gastaldi, R. Rossi, and G. Gecchele, "Effects of driver task-related fatigue on driving performance," *Procedia - Social Behav. Sci.*, vol. 111, pp. 955–964, Feb. 2014.

[5] Y. Dong, Z. Hu, K. Uchimura, and N. Murayama, "Driver inattention monitoring system for intelligent vehicles: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 596–614, Jun. 2011.

[6] Y. Wang, R. Huang, and L. Guo, "Eye gaze pattern analysis for fatigue detection based on GP-BCNN with ESM," *Pattern Recognit. Lett.*, vol. 123, pp. 61–74, May 2019.

[7] L. Alam and M. M. Hoque, "Vision-based driver's attention monitoring system for smart vehicles," in *Intelligent Computing & Optimization*, P. Vasant, I. Zelinka, and G.-W. Weber, Eds. Cham, Switzerland: Springer, 2019, pp. 196–209.

[8] L. Alam and M. M. Hoque, "Real-time distraction detection based on Driver's visual features," in *Proc. Int. Conf. Electr., Comput. Commun. Eng. (ECCE)*, Feb. 2019, pp. 1–6.

[9] A. M. Shibli, M. M. Hoque, and L. Alam, "Developing a vision-based driving assistance system," in *Emerging Technologies in Data Mining and Information Security*, A. Abraham, P. Dutta, J. K. Mandal, A. Bhattacharya, and S. Dutta, Eds. Singapore: Springer, 2019, pp. 799–812.

[10] B. Mandal, L. Li, G. S. Wang, and J. Lin, "Towards detection of bus driver fatigue based on robust visual analysis of eye state," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 3, pp. 545–557, Mar. 2017.

[11] Y.-C. Tsai, P.-W. Lai, P.-W. Huang, T.-M. Lin, and B.-F. Wu, "Vision-based instant measurement system for driver fatigue monitoring," *IEEE Access*, vol. 8, pp. 67342–67353, 2020.

[12] J.-C. Chien, Y.-S. Chen, and J.-D. Lee, "Improving night time driving safety using vision-based classification techniques," *Sensors*, vol. 17, no. 10, p. 2199, Sep. 2017.

[13] P. Chowdhury, L. Alam, and M. M. Hoque, "Designing an empirical framework to estimate the driver's attention," in *Proc. 5th Int. Conf. Informat., Electron. Vis. (ICIEV)*, May 2016, pp. 513–518.

[14] F. Vicente, Z. Huang, X. Xiong, F. De la Torre, W. Zhang, and D. Levi, "Driver gaze tracking and eyes off the road detection system," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 4, pp. 2014–2027, Aug. 2015.

[15] C. Zhang, R. Li, W. Kim, D. Yoon, and P. Patras, "Driver behavior recognition via interwoven deep convolutional neural nets with multi-stream inputs," *IEEE Access*, vol. 8, pp. 191138–191151, 2020.

[16] A. Gumaei, M. Al-Rakhami, M. M. Hassan, A. Alamri, M. Alhussein, M. A. Razzaque, and G. Fortino, "A deep learning-based driver distraction identification framework over edge cloud," *Neural Comput. Appl.*, pp. 1–16, Sep. 2020, doi: 10.1007/s00521-020-05328-1.

[17] D. Tran, H. M. Do, W. Sheng, H. Bai, and G. Chowdhary, "Real-time detection of distracted driving based on deep learning," *IET Intell. Transp. Syst.*, vol. 12, no. 10, pp. 1210–1219, Dec. 2018.
[18] S. Kaplan, M. A. Guvensan, A. G. Yavuz, and Y. Karalurt, "Driver behavior analysis for safe driving: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 6, pp. 3017–3032, Dec. 2015.
[19] D. S. Bowman, W. A. Schaudt, and R. J. Hanowski, *Handbook of Intelligent Vehicles*. London, U.K.: Springer, 2012, ch. Advances in Drowsy Driver Assistance Systems Through Data Fusion, pp. 895–912.
[20] S. Gupta and S. Mittal, "Yawning and its physiological significance," *Int. J. Appl. Basic Med. Res.*, vol. 3, no. 1, pp. 11–15, 2013.
[21] J. Grippenkoven and S. Dietsch, "Gaze direction and driving behavior of drivers at level crossings," *J. Transp. Saf. Secur.*, vol. 8, no. sup1, pp. 4–18, Jun. 2016.
[22] J. G. Taylor and N. F. Fragopanagos, "The interaction of attention and emotion," *Neural Netw.*, vol. 18, no. 4, pp. 353–369, May 2005.
[23] A. Zivony, A. S. Allon, R. Luria, and D. Lamy, "Dissociating between the N2pc and attentional shifting: An attentional blink study," *Neuropsychologia*, vol. 121, pp. 153–163, Dec. 2018.
[24] S. Benedetto, M. Pedrotti, L. Minin, T. Baccino, A. Re, and R. Montanari, "Driver workload and eye blink duration," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 14, no. 3, pp. 199–208, May 2011.
[25] E. De Valck and R. Cluydts, "Slow-release caffeine as a countermeasure to driver sleepiness induced by partial sleep deprivation," *J. Sleep Res.*, vol. 10, no. 3, pp. 203–209, Sep. 2001.
[26] F. Guede-Fernandez, M. Fernandez-Chimeno, J. Ramos-Castro, and M. A. Garcia-Gonzalez, "Driver drowsiness detection based on respiratory signal analysis," *IEEE Access*, vol. 7, pp. 81826–81838, 2019.
[27] J. Moon, Y. Kwon, J. Park, and W. C. Yoon, "Detecting user attention to video segments using interval EEG features," *Expert Syst. Appl.*, vol. 115, pp. 578–592, Jan. 2019.
[28] G. Li and W.-Y. Chung, "Combined EEG-gyroscope-tDCS brain machine interface system for early management of driver drowsiness," *IEEE Trans. Human-Mach. Syst.*, vol. 48, no. 1, pp. 50–62, Feb. 2018.
[29] A. Sahayadhas, K. Sundaraj, M. Murugappan, and R. Palaniappan, "Physiological signal based detection of driver hypovigilance using higher order spectra," *Expert Syst. Appl.*, vol. 42, no. 22, pp. 8669–8677, Dec. 2015.
[30] J. Zhang, Z. Yin, and R. Wang, "Recognition of mental workload levels under complex human–machine collaboration by using physiological features and adaptive support vector machines," *IEEE Trans. Human-Mach. Syst.*, vol. 45, no. 2, pp. 200–214, Apr. 2015.
[31] A. Alamri, A. Gumaei, M. Al-Rakhami, M. M. Hassan, M. Alhussein, and G. Fortino, "An effective bio-signal-based driver behavior monitoring system using a generalized deep learning approach," *IEEE Access*, vol. 8, pp. 135037–135049, 2020.
[32] C. C. Liu, S. G. Hosking, and M. G. Lenné, "Predicting driver drowsiness using vehicle measures: Recent insights and future challenges," *J. Saf. Res.*, vol. 40, no. 4, pp. 239–245, Aug. 2009.
[33] A. Fernández, R. Usamentiaga, J. Carús, and R. Casado, "Driver distraction using visual-based sensors and algorithms," *Sensors*, vol. 16, no. 11, p. 1805, Oct. 2016.
[34] J. M. Ramirez, M. D. Rodriguez, A. G. Andrade, L. A. Castro, J. Beltran, and J. S. Armenta, "Inferring drivers' visual focus attention through head-mounted inertial sensors," *IEEE Access*, vol. 7, pp. 185422–185432, 2019.
[35] K. Takemura, K. Takahashi, J. Takamatsu, and T. Ogasawara, "Estimating 3-D point-of-regard in a real environment using a head-mounted eye-tracking system," *IEEE Trans. Human-Mach. Syst.*, vol. 44, no. 4, pp. 531–536, Aug. 2014.
[36] M. Gjoreski, M. Z. Gams, M. Lustrek, P. Genc, J.-U. Garbas, and T. Hassan, "Machine learning and end-to-end deep learning for monitoring driver distractions from physiological and visual signals," *IEEE Access*, vol. 8, pp. 70590–70603, 2020.
[37] Z. Guo, Y. Pan, G. Zhao, S. Cao, and J. Zhang, "Detection of driver vigilance level using EEG signals and driving contexts," *IEEE Trans. Rel.*, vol. 67, no. 1, pp. 370–380, Mar. 2018.
[38] Y. Yao, X. Zhao, H. Du, Y. Zhang, and J. Rong, "Classification of distracted driving based on visual features and behavior data using a random forest method," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2672, no. 45, pp. 210–221, Dec. 2018.
[39] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, May 2004.
[40] M. Danelljan, G. Hager, F. S. Khan, and M. Felsberg, "Discriminative scale space tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 8, pp. 1561–1575, Aug. 2017.
[41] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1867–1874.
[42] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "A semi-automatic methodology for facial landmark annotation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2013, pp. 896–903.
[43] T. Soukupova and J. Cech, "Real-time eye blink detection using facial landmarks," in *Proc. 21st Comput. Vis. Winter Workshop*, Rimske Toplice, Slovenia, Feb. 2016, pp. 1–8.
[44] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.
[45] S. Suzuki and K. Be, "Topological structural analysis of digitized binary images by border following," *Comput. Vis., Graph., Image Process.*, vol. 30, no. 1, pp. 32–46, Apr. 1985.
[46] S.-T. Wu, A. C. G. D. Silva, and M. R. G. Márquez, "The Douglas-Peucker algorithm: Sufficiency conditions for non-self-intersections," *J. Brazilian Comput. Soc.*, vol. 9, no. 3, pp. 67–84, Apr. 2004.
[47] M.-K. Hu, "Visual pattern recognition by moment invariants," *IEEE Trans. Inf. Theory*, vol. IT-8, no. 2, pp. 179–187, Feb. 1962.
[48] F.-J. Chang, A. T. Tran, T. Hassner, I. Masi, R. Nevatia, and G. Medioni, "FacePoseNet: Making a case for landmark-free face alignment," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2017, pp. 1599–1608.
[49] J. Jiménez-Pinto and M. Torres-Torriti, "Optical flow and driver's kinematics analysis for state of alert sensing," *Sensors*, vol. 13, no. 4, pp. 4225–4257, Mar. 2013.
[50] *Road Vehicles—Measurement and Analysis of Driver Visual Behaviour With Respect to Transport Information and Control Systems*, Standard ISO 15007:2020, 2020.

LAMIA ALAM (Member, IEEE) received the B.Sc. and M.Sc. degree in computer science and engineering from the Chittagong University of Engineering and Technology (CUET), Bangladesh, in 2014 and 2018, respectively. She is currently working as an Assistant Professor with the Department of Computer Science and Engineering (CSE), CUET. Her research interests include human–computer interaction (HCI), computer vision (CV), and machine learning (ML). She is an Associate Member of the Institution of Engineers, Bangladesh (IEB).

MOHAMMED MOSHIUL HOQUE (Senior Member, IEEE) received the Ph.D. degree from the Department of Information and Computer Sciences, Saitama University, Japan, in 2012. He is currently a Distinguish Professor with the Department of Computer Science and Engineering (CSE), CUET, where he is also the former Head. He is also serving as the Director for Students' Welfare, CUET, where he is also the Director for Sheikh Kamal IT Business Incubator. He served as the TPC Chair, TPC Co-Chair, the Publication Chair, and TPC members in several international conferences. He was the Award Coordinator from 2016 to 2017), Conference Coordinator from 2017 to 2018 and the Vice-chair (Technical) from 2018 to 2020 of IEEE Bangladesh Section. Moreover, He was served as the Vice-Chair (Activity) from 2018 to 19, the Award Coordinator from 2017 to 18 of IEEE Computer Society Bangladesh Chapter and Educational Activity Coordinator from 2017 to 2020), IEEE Robotic and Automation Society, Bangladesh Chapter, respectively. He is also the Founding Director of the CUET Natural Language Processing Laboratory and the Fab Laboratory CUET. He published more than 125 publications in several International Journals, and Conferences. His research interests include human robot/computer interaction, artificial intelligence, machine learning, and natural language processing. He is a Senior Member of the IEEE Computer Society, IEEE Robotics and Automation Society, IEEE Women in Engineering, IEEE Signal Processing Society, USA, and a Fellow of Institute of Engineers, Bangladesh.

M. ALI AKBER DEWAN (Member, IEEE) received the B.Sc. degree in computer science and engineering from Khulna University, Bangladesh, in 2003, and the Ph.D. degree in computer engineering from Kyung Hee University, South Korea, in 2009. From 2003 to 2008, he was a Lecturer with the Department of Computer Science and Engineering, Chittagong University of Engineering and Technology, Bangladesh, where he was an Assistant Professor, in 2009. From 2009 to 2012, he was a Postdoctoral Researcher with Concordia University, Montreal, QC, Canada. From 2012 to 2014, he was a Research Associate with the École de Technologie Supérieure, Montreal. He is currently an Assistant Professor with the School of Computing and Information Systems, Athabasca University, Athabasca, AB, Canada. He has published more than 50 articles in high-impact journals and conference proceedings. His research interests include artificial intelligence, affective computing, computer vision, data mining, information visualization, machine learning, biometric recognition, medical image analysis, and health informatics. He has served as an Editorial Board Member, the Chair/ Co-Chair, and a TPC member in several prestigious journals and conferences. He received the Dean's Award and the Excellent Research Achievement Award for his academic performance and research achievements during his Ph.D. studies in South Korea.

NAZMUL SIDDIQUE (Senior Member, IEEE) received the Dipl.-Ing. degree in cybernetics from TU Dresden, Germany, the M.Sc. degree in computer science from BUET, Bangladesh, and the Ph.D. degree in intelligent control from the Department of Automatic Control and Systems Engineering, University of Sheffield, U.K. He is currently with the School of Computing, Engineering and Intelligent Systems, Ulster University. His research interests include robotics, cybernetics, computational intelligence, nature-inspired computing, stochastic systems, and vehicular communication. He has published more than 170 research articles in the broad area of computational intelligence, vehicular communication, robotics and cybernetics. He authored or coauthored five books published by John Wiley, Springer, and Taylor & Francis. He guest edited eight special issues of reputed journals on *Cybernetic Intelligence*, *Computational Intelligence*, *Neural Networks*, and *Robotics*. He has been involved in organizing many national and international conferences and co-edited seven conference proceedings. He is a Fellow of the Higher Education Academy and a member of different committees of IEEE SMCS. He is on the Editorial Board of the *Scientific Research* (Nature), *Journal of Behavioral Robotics*, *Engineering Letters*, *International Journal of Machine Learning and Cybernetics*, *International Journal of Applied Pattern Recognition*, *International Journal of Advances in Robotics Research*, and also on the Editorial Advisory Board of the *International Journal of Neural Systems*.

INAKI RANO received the M.Sc. degree in physics and the Ph.D. degree in computer sciences from the University of the Basque Country, Spain. From 1997 to 2004, he was a Research Assistant with the San Sebastian Technology Park and a member of the Robotics and Autonomous Systems Group, University of the Basque Country working in control architectures for mobile robots. In 2005, he joined the Computer Science and Systems Engineering Department, University of Zaragoza, Spain, where he developed several biologically inspired navigation strategies for mobile robots. In 2008, he was on Sabbatical Leave with the University of Essex, U.K., working on the application of systems identification to the generation of bio-inspired controllers. In 2011, he joined the Institute for Neural Computation of the Ruhr-Universitaet Bochum, Germany. He was Lecturer on Cognitive Robotics with the School of Computing and Intelligent Systems, Ulster University, from 2013 to 2018. He is currently an Assistant Professor with SDU Biorobotics, University of Southern Denmark. His current research interests include the applications of dynamical systems and control theory to biorobotics.

IQBAL H. SARKER (Member, IEEE) received the Ph.D. degree from the Department of Computer Science and Software Engineering, Swinburne University of Technology, Melbourne, VIC, Australia, in 2018. He is currently working as a Faculty Member of the Department of Computer Science and Engineering, Chittagong University of Engineering and Technology. His professional and research interests include data science, machine learning, AI-driven computing, NLP, cybersecurity analytics, behavioral analytics, the IoT-smart city technologies, and healthcare analytics. He has published a number of peer-reviewed Journals and Conferences in top venues, such as Journals [*Journal of Network and Computer Applications* (Elsevier, USA), *Internet of Things* (Elsevier), *Journal of Big Data* (Springer Nature, U.K.), *Mobile Network and Applications* (Springer, The Netherlands), *Sensors* (Switzerland), *The Computer Journal* (Oxford University Press, U.K.), IEEE TRANSACTIONS ON ARTIFICIAL INTELLIGENCE, IEEE ACCESS (USA), and so on] and Conferences [IEEE DSAA (Canada), IEEE Percom (Greece), ACM Ubicomp (USA and Germany), ACM Mobiquitous (Australia), PAKDD (Springer, Australia), ADMA (Springer, China), and so on]. He is one of the Research Founders of the International AIQT foundation, Switzerland, and a member of ACM and IEEE.