2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 7-9 February, 2019

Real-Time Distraction Detection Based on Driver's Visual Features

Lamia Alam and Mohammed Moshiul Hoque*
Department of Computer Science and Engineering
Chittagong University of Engineering & Technology
Chattogram-4349, Bangladesh
Email: lamiacse09@gmail.com; mmoshiulh@gmail.com*;

***Abstract*— Driver's distraction has been listed as the leading contributing factor to traffic accidents for the past decades. This paper focuses on developing an approach to detect distraction real time by analyzing driver's visual feature from the face region. The proposed approach uses visual features such as movement of eye and head to extract critical information to detect driver attention states and to classify it as either attentive or distracted. Deviation of eye center and head from their standard position for a period of time is considered to be useful cues for detecting lack of attention in this approach. At first face detection is performed after which region of interest (ROI) - eye and head region, are extracted using facial landmarks and lastly, head and eye movements are detected to classify attention state. To evaluate the system performance, we conducted an experiment in a real driving environment with subjects having different characteristics. Our system achieved on average 92% accuracy in detecting attention state for all tested scenarios.**

Keywords—distraction; eye movement; head movement; eye center, yaw angle;

I. INTRODUCTION

In their Global status report on road safety 2015, World Health Organization (WHO) [1], showed that road traffic injuries are emerging as a leading cause of death causing more than 1.2 million people to die globally. In their report, WHO listed Distracted driving as a dangerous and growing impedance to road safety.

There has been a terrible rise in road accidents in Bangladesh over the past few years. According to a leading newspaper of Bangladesh in the past three and a half years, around 25,120 people died in road crashes across the country and 20 people were killed daily. In the last three and a half years 62,482 people were killed in the road according to the report [2]. In WHO's Global Status Report on Road Safety 2015 [1], road traffic death has been estimated 13.6 per 100,000 population in Bangladesh.

There is a strong amount of proof that suggests that driver's distraction is one of the prime factors in road accidents globally [1,3]. Driver's distraction not only caught the attention of international researcher as a prominent risk factor for traffic accident but also researcher's from Bangladesh has also identified it as contributing factor to traffic crashes [4, 5]. All these facts suggest the need of introducing intelligent assistance tools in vehicles which will be able to detect driver's distraction and take countermeasures to reduce traffic crash risks related to driver's distraction.

Distraction can be define as a process of diverting one's attention from a desired area of focus which results in blocking or diminishing the reception of desired information. A more formal definition of distraction is proposed in [6]: "Driver distraction is a diversion of attention away from activities critical for safe driving toward a competing activity". The root of driver distraction are variegated and results in hazards. National Highway Traffic Safety Administration (NHTSA) classifies distractions into the following four categories namely- visual, cognitive, auditory and biomechanical based on driver's in vehicle functionality [7]. Here, we are mainly focusing in visual distraction which can be define as looking away from the roadway focusing on something else. This paper focuses on developing a distraction detection system in real time and it will be helpful to create alertness while the driver is in a distracted state and can be incorporated in driver's attention monitoring system. The proposed system obtain some visual cues such as movement of eye and head which seek a real-time record of driver's to detect visual distraction.

II. RELATED WORK

Driver distraction being one of the most important factors in highway crashes, distraction detection has been a very active research field for the past few decades and a comprehensive number of approaches have been developed. They can be grouped into- physiological, driving behavioral, visual feature based and hybrid approach. This section summarizes some of these very recent works in distraction detection.

Physiological approaches require analysis of crucial cues such as brain signal, skin conductance, heart /pulse rate, etc. A framework was proposed by Wang et al. [8] that is able to detect driver distractions measuring brain activity by analyzing electroencephalographic (EEG) signals. A model was proposed for driver distraction in [9] based on EEG data to estimate the level of distraction by monitoring the effects of the dual tasks in a virtual reality (VR)-based simulation.

Driving-behavior-information-based approaches evaluate the driver's driving behavior i.e. a set of all real-time tractable driving information, such as speed, acceleration, lane changing etc. over time. Craye et al. presented a multi-modal approach based on a driving simulator platform equipped with several sensors, to acquire data such as audio, color video, depth map, heart rate, and steering wheel and pedals positions etc. for driver fatigue and distraction detection [10].

*Corresponding author

978-1-5386-9111-3/19/$31.00 ©2019 IEEE

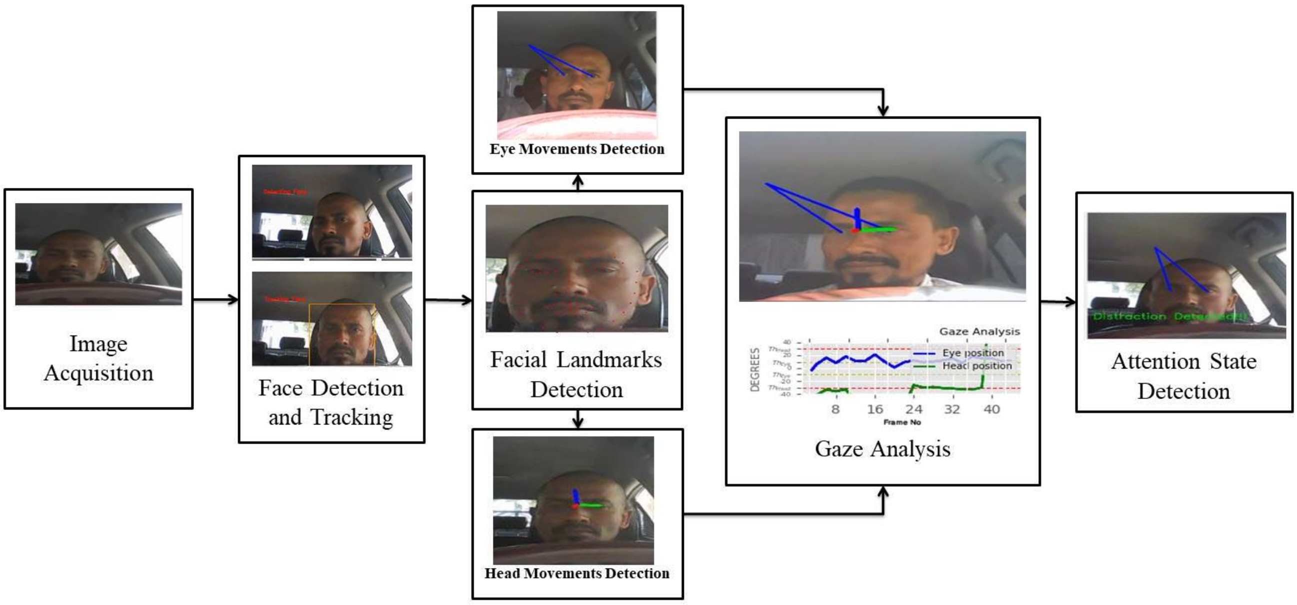


Fig. 1. Schematic diagram of driver's attention monitoring system

Visual-feature-based approach uses visual features from the driver's facial images. Xing et al. proposed a feedforward neural network (FFNN) based system to monitor drivers and identify driving tasks based on 3-D head rotation angles and the upper body (hand and arm at both sides) joint positions using Kinect in a real vehicle [11]. Tores et al. propose a system able to monitor drivers video surveillance using convolutional neural network [12]. A semi-supervised method was proposed by Liu et al. to classify driver's state into the category- attentive and cognitively distracted using eye and head movements [13]. Another similar work was proposed by Vicente to accurately detect Eyes Off the Road (EOR) based on head pose and gaze estimation [14]. Another work based on eye state and head pose was developed by Mbouna et al. that monitors alertness of a vehicle driver continuously[15]. A very recently Alam et al. proposed a driver's attention monitoring system based on visual cues that estimates the driver's attentional status based on a variety of parameters such as, percentage of eyelid closure over time (PERCLOS), yawn frequency and gaze direction [16].

Recently, a very few research works have been done for detecting driver inattention level by combining different approaches. Yao et. al explored the connection between a driver's visual features and driving behaviors of distracted driving and implemented the random forest (RF) method to classify and improve the detection accuracy [17]. Another system was developed recently to detect vigilance level by combining driver's electroencephalogram (EEG) signals and driving contexts [18].

As measures from visual cue have the advantage of being inconspicuous for being able to be recorded using input device from far away, we are interested to develop a system that can detect visual distraction from visual cues such as- eye and head movement.

III. Proposed Distraction Detection Framework

The key goal of this work is to develop a framework which will be able to detect driver's visual distraction in a real driving environment. We focused on only two cues: eye and head movements in this work despite there are lots of cues involved in detecting the visual distraction. Fig. 1 illustrates the block diagram of our proposed system which consists of seven main modules. Following subsections describe each module in details.

A. *Image Acquisition*

The proposed system starts with acquiring video data of driver's frontal face to extract visual cues that typically characterize a driver's attention state. A simple USB plug and play web camera (Logitech C170) is used for the proposed system. The input device is placed on the car's dashboard in a way that the driver's facial feature can be captured which is sent to the next module.

B. *Face Detection and Tracking*

An RGB frame (F_{RGB}) of a video sequence is taken as input by this module and converted to gray scale using the following equation:

$$F_{Gray} = 0.299 \times R + 0.587 \times G + 0.114 \times B \qquad (1)$$

For face detection, Viola-Jones face detection algorithm [19] is used, which returns the positions of all the detected faces in the frame as rectangles. Then using the integral image we calculated the area of the rectangles, as more than one face can be captured in a frame in the real driving scenario, we considered the "Closest" face (f) as our subject of interest (driver), which has the largest area.

A face is detected in the first frame and is tracked in the consecutive frames using face tracking algorithm [20] unless the face got lost. This mechanism made face region extraction less complex. The face region is then used by the next module to detect landmark points.

C. *Facial Landmarks Detection*

After detecting and tracking face we need to obtain additional information about the face. Based on used detector we can assume rough orientation, but details remain unknown.

Faces have some common stable characteristic such as inner and outer corners of eyes etc., which we need to determine to detect driver's attention state. These prominent regions can be localized using facial landmarks. A wide range of facial landmark detectors is available, all of which are used to localize and label the facial regions. Different facial landmark detectors specify a different number (and location) for these points. The facial landmark detector used by our proposed system is an implementation of the [21] and trained using iBUG300-W dataset [22]. The indexes of the 68 coordinates can be visualized on the image given in Fig. 2.

For each frame (F_{Gray}), the location of the major points of interest of the facial structure of the detected face region (f) by the previous module was obtained using the facial landmark detector. In order to detect eye and head movement, we are particularly interested in point related to eye and head region.

D. Eye Movements Detection

To detect movement of the eye, the center of eyes are detected to determine the rate of deviation of eye centers from the standard position.

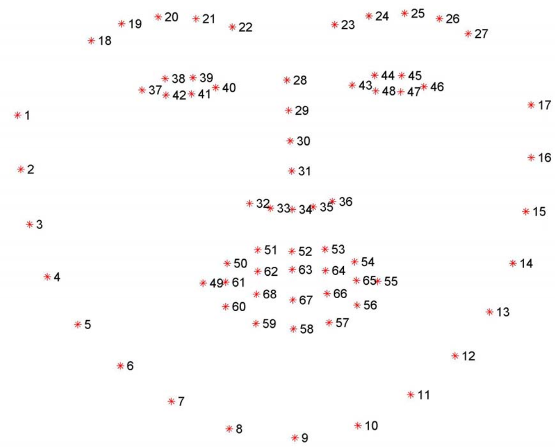
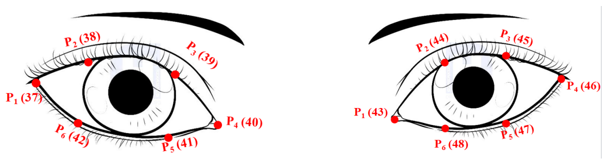


Fig. 2. Visualizing each of the 68 facial coordinate points used by our system

In order to do so at first, the location of the center of the eye has been detected. To facilitate the detection of eye center some preprocessing on both left and right eye image found by six (x, y)-coordinates from the previous modules are performed individually. Fig. 3 shows the points related to eye region which were used to detect eye movements.



(a) Right eye (b) Left eye

Fig. 3. Point used to obtain eye movements

At first, the eye images are resized using the bilinear interpolation method. Then the histogram is calculated improves the contrast in the image. A threshold is determined based on the maximum count and the size of the image size from the histogram. Then pixels with exceeding the threshold value are eliminated and were considered as skin pixels. And lastly, noise removal was done by performing erosion followed my dilation. Fig. 4 shows the preprocessing steps performed to detect eye movements.

Form fig. 4, we can see that the visible eyeball area found from the last step can be considered to be an ellipse. And to detect the boundary of this ellipse on the binary image border following algorithm by Suzuki and Abe [23] and approximation algorithm [24] to present an ellipse shape of vector points were implemented.

The center of this ellipse can be used to indicate the eye center. For pixel intensities $I(x, y)$ moments m_{ij} can be used to calculate the centroid $(\bar{x}, \bar{y})$ [25] of the ellipse:

$$\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}} \qquad (2)$$

where,

$$m_{ij} = \sum_{x,y} I(x, y).x_j.y_i \qquad (3)$$

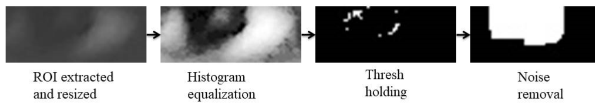


Fig. 4. Preprocessing steps for detecting the eye movemnets

Eye Movement (EM) can be classified into left, front and right using threshold value τ_θ by (4) once center $(\bar{x}, \bar{y})$ of both eyes has been calculated.

$$EM = \begin{cases} Left; if\ \theta > \tau_\theta \\ Right; \text{else } if\ -\tau_\theta < \theta \\ Front; Otherwise \end{cases} \qquad (4)$$

Here, $\theta = \tan^{-1}\left(\frac{\Delta x}{\Delta y}\right)$; whereas $\Delta x = \bar{x}_R - \bar{x}_L, \Delta y = \bar{y}_r - \bar{y}_L$, and $(\bar{x}_L, \bar{y}_L)$ and $(\bar{x}_R, \bar{y}_R)$ are the center of left and right eye respectively.

E. Head Movement Detection

Head movement can be estimated by localizing 15-facial landmarks points marked as blue dots in Fig. 5 found from Facial Landmarks Detection and thus head rotation calculation was performed.

Head rotation was calculated by using the classic solution of Perspective-*n*-Point (PnP) problem [26] - which works as follows:

$$h = (r, t)^T \qquad (5)$$

where, h is 3D head pose and consists of 6 Degree of Freedom (DOF) i.e. r = $(r_x, r_y, r_z)^T$ for rotations, and t = $(t_x, t_y, t_z)^T$ for translations.

Then the head pose is calculated by the perspective transformation as follows:

$$s[p, 1]^T = M[R|t]P^T \qquad (6)$$

Here, s is scaling factor, M stands for camera matrix and the joint rotation-translation matrix is expressed as [R|t].

Matrix R obtained using Rodrigues rotation formula is given in (7) and is used to calculate vector $r = (r_x, r_y, r_z)$.

$$R = \cos\theta I + (1 - \cos\theta) r r^T + \sin\theta \begin{bmatrix} 0 & -r_z & r_y \\ r_z & 0 & -r_x \\ -r_y & r_x & 0 \end{bmatrix} \quad (7)$$

where, I is vector in □^{3} and $\theta = ||r||_2$

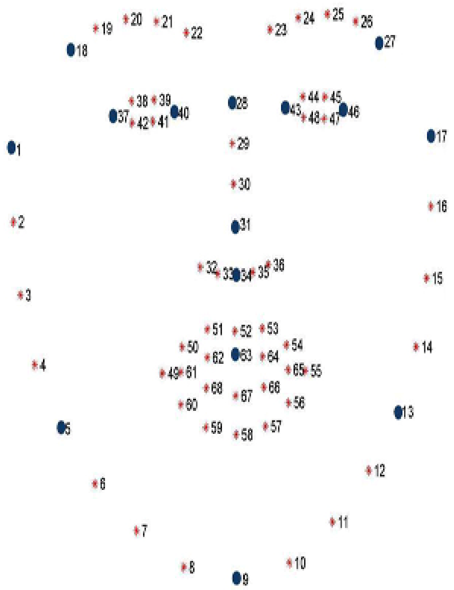


Fig. 5. 15 points marked as blue dots used for estimating head rotation

Then vector r, which represent head rotation in degree is used to obtain the Euler angle yaw(α) and is used to model the head movement (HM) by using threshold values $Th_{\alpha_1}$and Th_{α_2}($Th_{\alpha_1} > Th_{\alpha_2}$) in (8).

$$HM = \begin{cases} Left; & if -Th_{\alpha_1} \leq \alpha < -Th_{\alpha_2} \\ Front; & if -Th_{\alpha_2} \leq \alpha \leq Th_{\alpha_2} \\ Right; & if\ Th_{\alpha_2} < \alpha \leq Th_{\alpha_1} \end{cases} \quad (8)$$

F. *Gaze Analysis*

To detect driver's distraction, we emphasized on analyzing driver's gaze direction. Both eye and head movements were considered to obtain the gaze. These visual cues of driver's facial region are captured, extracted, calculated, stored and updated continuously and are represented graphically.

G. *Attention State Detection*

The data from the previous component is used to detect the attention state of the driver by estimating eye and head movements for past few frames.

For a driver, the nominal eye and head direction are frontal. Deviation of eye or head or both from its standard direction for an extended period of time (T) can be classified as a visual distraction. So, Gaze direction (GD) is computed as a combination of EM and HM for over T time and can be defined as,

$$GD = \{FM, HM\}_T \quad (9)$$

Situations such as wide head rotation, wearing sunglass etc. when eyes are not visible, GD can be obtained using only the HM.

Our system classified driver's attention state using (10).

$$Attention\ State = \begin{cases} Distrcated; GD == Right || GD == Left \\ Attentive; Otherwise \end{cases} \quad (10)$$

IV. Experimental Analysis

The proposed system was tested in terms of robustness, accuracy and detection time on a Windows 10 PC with an Intel Core i5 1.60 GHz processor and 4 GB RAM.

To evaluate the system, we conducted an experiment in a real driving scenario with different subjects (drivers). The major objective of this experiment was to

- Verify the validity of detected eye and head movements
- Measure the accuracy of the system in detecting attention status.
- Report detection time.

A. *Participants*

A total of ten participants both male and female having different facial features, hairstyles and wearing different accessories(such as spectacles, sunglass, and cap) participated in this experiment. The average age of participants is 32.4 years (SD = 4.45).

B. *Experimental Setup.*

Experimental Environment is created by the web camera with the computer via USB and also set the camera in a proper position on the car dashboard in front of the driver. During testing, the participants were asked to be seated directly in front of the camera in the driving seat. For this experiment, Logitech C170 was used. Fig. 6 shows the experimental setup.

Fig. 6. Experimental setup

C. *Evaluation Measures.*

- Verification of the validity of detected eye and head movement are done based on two criteria- false positive rate (FPR) and false negative rate (FNR), using (11) and (12).

$$FPR = \frac{FP}{FP+TN} \quad (11)$$

where FP stands for the total number of false positives and TN for the total number of true negatives.

and $$FNR = \frac{FN}{FN+TP} \qquad (12)$$

where FN stands for the total number of false negative and TP for the total number of true positives.

- Accuracy of detecting attention state was measured using (13), for the video sequence.

$$Accuracy = \frac{C_F}{T_F} \times 100\% \qquad (13)$$

where, C_F is the total number of frames in which attention status was accurately detected and T_F is the total number of frames in the video sequence.

D. *Experimental Procedure*

Participants were asked to ride the with an average speed of 20 kilometers as well as to move their face and eye direction to left and right, speak, blink randomly to evaluate the system performance. A 45 minutes video which contains more than 45000 frames. was captured and analyzed to verify the validity of detected eye and head movements and detected attention state and to measure the attention state detection time of our proposed system.

E. *Results*

The results of the experiment to evaluate the system performance and time complexity are explained as follows:

1) Validification of the detected eye and head movements

In the case of eye and head movement, false positive occurs when the subject's head/ eye is not in frontal direction but the system detects it in the frontal direction. And false negative occurs when the head/ eye is in frontal direction but the system detects it in other (right/left) direction. FPR and FNR for detecting eye and head movements is given in Table I and Table II. Fig. 7 shows some of the sample frames with a subject wearing different accessories from the experiment. Subject wearing spectacle affected the performance of detection of eye movement and wearing sunglasses makes it impossible to detect the eye movement, whereas head movement is not affected by any of the situations much.

TABLE I. EXPERIMENTAL RESULT FOR EYE MOVEMENTS

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

TABLE II. EXPERIMENTAL RESULT FOR HEAD MOVEMENTS

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

2) Accuracy of detecting attention state

The overall performance of the system was determined using equation (13). Table III shows the overall accuracy of the proposed system in detecting attention state. Fig. 8 shows sample frames of a subject with different attention state.

3) Detection time of the system

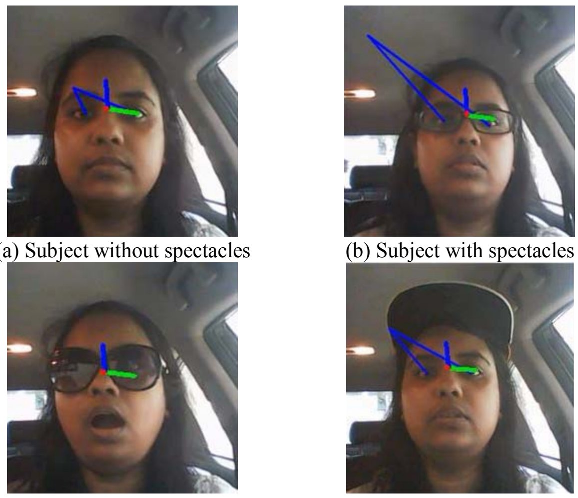
In order to evaluate the time complexity, we observed the time taken by the system to detect the face, facial landmarks, eye and head movement and attention state for a single frame for each participant in milliseconds (ms). Analyzing the data we can say that our system on average takes 14 ms to detect face, 15 ms to detect facial landmarks, 22 ms to detect eye and head movements and 5 ms to classify attention state. In another word, we can say that it takes 56 ms i.e. 0.056 seconds(s) to perform the operations on each frame.

F. *Discussion*

Results from this experiment suggest that our system is quite accurate in detecting eye and head movements and indicate that overall accuracy of our system is 92% and the system can run at about 18 frames per second (fps).

V. CONCLUSION

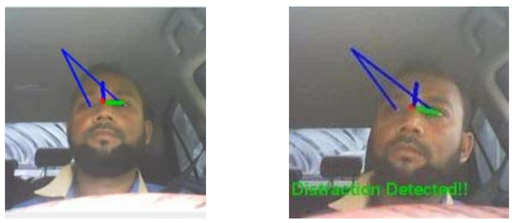
This paper proposes a system that detects distraction in real-time using the video stream captured from a web camera placed on the dashboard. The system is able to detect distraction under a variety of facial characteristics (such as facial features, hairstyles, and accessories) of the driver. The system achieved accuracy of around 92 % for the entire situation. Our experiments showed that our eye and head movement detection algorithm is robust. The inclusion of warning system to warn the driver while distraction is detected is kept as future work.



(a) Subject without spectacles (b) Subject with spectacles

(c) Subject with sunglass (d) Subject wearing cap

Fig. 7. Sample frames form the experiment while valdating detection of eye and head movements



(a) Attentive (b) Distracted

Fig. 8. Sample frames form the experiment while determining of detecting attention state

TABLE III. OVERALL ACCURACY OF THE SYSTEM IN DETERMINING ATTENTION STATE- ATTENTIVE AND DISTRACTED

Participant No.	**Tf**	**Cf**	**Accuracy**
1	4875	4388	90
2	4932	4697	95
3	4842	4692	97
4	4156	3727	90
5	4860	4392	90
6	4902	4462	91
7	4730	4117	87
8	4172	3933	94
9	4275	3957	93
10	4529	4287	95
Total	46273	***Average***	92

REFERENCES

[1] "Global status report on road safety 2015," World Health Organization (WHO), Geneva 27, Switzerland, Rep., 2015.

[2] S. Rahman. (2018, August 04) Over 25,000 killed on roads in 3 years. Prothom Alo. Dhaka, Bangladesh. [Online]. Available: https://en.prothomalo.com/bangladesh/news/181012/25-120-killed-on-roads-in-3-years [Accessed15 Oct. 2018].

[3] "Undercounted is underinvested: How incomplete crash reports impact efforts to save lives," National Safety Council (NSC), Itasca, IL 60143, USA, Rep., 2017.

[4] R. R. Wahi, and M. T. Rahman, "An exploratory study of driver distraction in bangladesh: A contributing factor to crashes," in *Proc. 4th Global Engineering, Science and Technology Conf.*, Dhaka, Bangladesh, 2013, pp. 725-730.

[5] N. Sharmeen, and M. R. Islam, "Road accidents: Contemporary scenario and policy issues in bangladesh," *J. of Bangladesh Institute of Planners*, vol. 4, pp. 45-55, 2011.

[6] D. Lee, K. L. Young, and M. A. Regan, "Defining driver distraction," in *Driver Distraction: Theory, E ects, and Mitigation*. CRC Press, Taylor & Francis Group, Boca Raton, FL, USA, 2008, p. 31.

[7] R. G. T. A. Ranney, E. Mazzae, and M. J. Goodman, "Nhtsa driver distraction research: Past, present, and future," Transportation Research Center Inc. East Liberty, OH, Rep., 2000.

[8] S. Wang, Y. Zhang, C. Wu, F. Darvas, and W. A. Chaovalitwongse, "Online Prediction of Driver Distraction Based on Brain Activity Patterns," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 1, pp. 136-150, Feb. 2015.

[9] Y.-K. Wang, S.-A. Chen, and C.-T. Lin, "An EEG-based brain–computer interface for dual task driving detection," *Neurocomputing*, vol. 129, pp 85-93, 2014.

[10] C. Craye, A. Rashwan, M. S. Kamel, and F. Karray, "A multi-modal driver fatigue and distraction assessment system," *Int. J. Intell. Transp. Syst. Res.*, vol.14, no. 3, pp. 173-194, 2016.

[11] Y. Xing et al., "Identification and analysis of driver postures for in-vehicle driving activities and secondary tasks recognition," *IEEE Trans. Comput. Social Syst.*, vol. 5, no. 1, pp. 95-108, March 2018.

[12] R. Torres, O. Ohashi, E. Carvalho, and G. Pessin, "A deep learning approach to detect distracted drivers using a mobile phone," in *Proc. Int. Conf. on Artificial Neural Networks (ICANN 2017)* in Artificial Neural Networks and Machine Learning, LNCS 10614, Springer, Cham, 2017.

[13] T. Liu, Y. Yang, G. Huang, Y. K. Yeo, and Z. Lin, "Driver distraction detection using semi-supervised machine learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 4, pp. 1108-1120, April 2016.

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

[15] R. Oyini Mbouna, S. G. Kong, and M. Chun, "Visual analysis of eye state and head pose for driver alertness monitoring," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 3, pp. 1462-1469, Sept. 2013.

[16] L. Alam, and M. M. Hoque, "Vision-based driver's attention monitoring system for smart vehicles," in *Proc. 1st Int. Conf. on Intelligent Computing & Optimization (ICO 2018)* in Advances in Intelligent Systems and Computing, vol. 866, Springer, Cham, 2018.

[17] 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," *Transportation Research Record: J. Trans. Res. Board*, Sep. 2018.

[18] 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, March 2018.

[19] P. Viola, and M. J. Jones, "Robust real-time face detection," *Int. J. Computer Vision*, vol. 57, pp. 137-154, May 2004.

[20] M. Danelljan, G. Häger, F. Khan, and M. Felsberg, "Accurate scale estimation for robust visual tracking," in *Proc British Machine Vision Conf. (BMVC 2014)*, Uk, September 2014.

[21] V. Kazemi, and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proc 2014 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2014)*, Ohio, USA, pp. 1867-1874.

[22] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "A semiautomatic methodology for facial landmark annotation," in *Proc. 2013 IEEE Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW 2013)*, Oregon, USA, pp. 896-903.

[23] S. Suzuki and K. be, "Topological structural analysis of digitized binary images by border following," *Computer Vision, Graphics, and Image Processing*, vol. 30, no. 1, pp. 32 -46, 1985.

[24] S. . Wu and M. R. G. Marquez, "A non-self-intersection douglas-peucker algorithm," in *Proc. 16th Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI 2003)*, Sao Carlos, Brazil, Oct 2003, pp. 60-66.

[25] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Inf. Theory*, vol. 8, no. 2, pp. 179-187, February 1962.

[26] F. Chang, A. T. Tran, T. Hassner, I. Masi, R. Nevatia, and G. Medioni, "Faceposenet: Making a case for landmark-free face alignment," in *Proc. 2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*, Oct 2017, Venice, Italy, pp. 1599-1608.