© June 2020 | IJIRT | Volume 7 Issue 1 | ISSN: 2349-6002

# Drowsiness Detection System using Machine Learning

S. Jansi Rani[1], Anand Rajasekharan[2], Chandrasekhar A R[3]

[1]*Assistant Professor, Department of Information Technology, Sri Ramakrishna Engineering College, Coimbatore, Tamil Nadu, India*

[2,3]*Department of Information Technology, Sri Ramakrishna Engineering College, Coimbatore, Tamil Nadu, India*

***Abstract*- An accident is an incident, occurring suddenly, unexpectedly, and inadvertently under unforeseen circumstances. Each day thousands of road accident victims succumb to injuries and death globally. According to studies, almost one-quarter of all serious motorway accidents are due to drivers handling the vehicle in a state of heavy drowsiness indicating that driver drowsiness tends to cause a greater number of accidents than driving under influence. Long haul drivers who put off taking a break while moving for long-distance often run a high risk of becoming sleepy behind the wheel and fail to identify this early on. By making use of the technologies of computer vision, an interdisciplinary science that deals with how computers can be made more intelligent by developing techniques that help them gather high-level understanding from digital images or videos, a solution can be developed for this problem of detecting and notifying of driver drowsiness. The main idea behind this project is to develop a computer vision system, using OpenCV, which is non- intrusive that can automatically detect driver drowsiness in a real-time video stream and will alert the driver by playing an alarm if they appear to be drowsy. This system would make use of an algorithm that detects the eye landmark points of the driver and based on this can determine if the driver is in a drowsy state or not, and appropriately sound an alarm.**

***Index terms*- drowsiness detection, computer vision, OpenCV, landmark points**

## I.INTRODUCTION

In the last decade, there has been a huge growth in urbanization, and this has led to an exponential increase in the number of vehicles on the road. With this traffic, the number of road accidents has seen a high surge and thus the fatalities occurring in transport have reached a record high. As a result, road safety is one of the most highly focused areas in the field of transport. A slip in this department in the slightest form would lead to catastrophic circumstances. Day by day new innovations is being brought up to circumvent these lapses in safety by integrating upcoming technologies with prevalent methods to solve the issues. This comes as a much-needed aid, especially in a country like India where road accidents claim nearly three lives every minute. Drowsiness occurring due to driving for a long time is one of the most common and dangerous issues related to road safety. According to a study by the Central Road Research Institute (CRRI) [1], about 40% of road accidents occur due to drivers who are exhausted, falling asleep behind the wheel. Many technologies and systems have been proposed to look into this issue but not many have been very successful, some reasons being, the system developed has a detection method is subjective in nature where the driver is made to participate in activities like filling questionnaires, based on which their drowsiness is detected. The more successful detection has been the objective detection method. This method evaluates the driver's behaviour and fatigue in real-time. The Objective method is also further classified depending on whether the system is a contact or non-contact system. A contact system would detect fatigue in the driver by monitoring physiological characteristics such as his heart rate, brain waves, and eye movement, or changes in physical characteristics such as lowering of one's head, sagging posture, etc. To put such a system into use there would be a requirement of attaching some form of hardware like electrodes to the driver which would be in the way of the driver's operation. A non- contact system, however, monitors the physical changes in the driver by making use of non-intrusive systems and computer vision technology that can detect any change in the driver's physical aspects that can help conclude whether the driver is drowsy or not, such as eye blinking, head lowering, etc. In comparison to the contact system, the non- contact system comes out to be more convenient, owing to its low cost of installation and less need for sophisticated technology.

In this paper, we propose a non- contact system that uses a camera facing the driver to capture their face by monitoring the video stream continuously. On successful detection of the driver's face, we extract the facial features of the face, the region of the eye in particular, by making use of the facial landmark detection algorithm. From the eye region data, we compute the eye aspect ratio (E.A.R) [2] of the driver. Using the eye aspect ratio we can determine if the eyes of the driver are closed or not, as the eye aspect ratio remains a constant in the case of the eyes being open but it tends to fall and not increase again if the driver's eyes are closed for a longer period i.e., when they have dozed off behind the wheel, in such a case occurring, an alarm is sounded to wake the driver up.

In such a system, the driver's face is the most important area of focus, as it conveys the most information through the driver's facial expressions, thus it is very important for the face detection part to be highly accurate. Previous such non- contact systems made use of Haar feature-based cascade classifiers for the purpose of face detection. Haar feature-based cascade classifiers are very good at detecting lines and edges, which makes it very effective in detecting faces and it is very fast in its operation, but the drawback of using a Haar feature-based system is that the results can tend to be not very accurate. In the case of a driver wearing glasses of any kind, this system would fall short in the task of face detection. Thus, to overcome this problem of inaccuracy and detect the driver's face, we use a Histogram of Oriented Gradient (HOG) feature descriptor, in which the distribution (histograms) of directions of gradients (oriented gradients) are used as features. This descriptor is more accurate than the previously discussed system as it is generated on a dense grid of uniformly spaced cells and makes use of local contrast normalization that is overlapping in nature.

This HOG descriptor is combined with a Linear Support Vector Machine (SVM) into which the feature vector from the HOG algorithm is fed for the purpose of image classification, as a Linear SVM can be used to train highly accurate object classifiers [3]. After successful face detection, we will require proper facial landmark detection to be done, and for this purpose, we make use of the facial landmark detector available in dlib library, which is a general-purpose cross-platform library. The facial landmark detector which is used in dlib library is an implementation of One Millisecond Face Alignment with an Ensemble of Regression Trees [4]. This method works by training a set of marked facial landmarks on an image. These images are manually marked, by using the specific (x, y)-coordinates of the areas around each facial structure. By making use of this trained data, a collection of regression trees are trained to compute the facial landmark positions using the pixel intensities. Through this, we arrive at a facial landmark detector that can successfully detect facial landmarks in real-time with high-quality predictions. By using the coordinates of the position of the eyes, we can compute the Euclidean distance which is used for the eye aspect ratio. Using this eye aspect ratio, we determine if the driver has shown signs of sleepiness and take necessary action in the form of an alarm.

The contribution of this paper is the HOG facial descriptor algorithm working with the Linear SVM classifier for highly accurate image classification and in turn, better face detection, which results in a more reliable system for providing drowsiness alerts to the drivers

## II. MOTIVATION

This section provides an introduction of the existing methods in detecting the drowsiness in drivers and elaborates the challenges.

Visual object tracking is also a vital problem in computer vision. It has a large range of applications in fields like human-computer interaction, behaviour recognition, robotics, and surveillance. Visual object tracking estimates the target location in every frame of the image sequence, given the starting state of the target in the previous frame. Lucas and Kanade [5] suggested that the tracking or locating of the moving target can be realized using the pixel relationship between adjacent frames of the video sequence and displacement changes of the pixels. However, this algorithm can only detect the medium-sized target that shifts between two frames. Using the new methods of the correlation filter in computer vision technology, the Minimum Output Sum of Squared Error (MOSSE) filter proposed by Bolme et al [6], which can produce stable correlation filters to trace the article. Although the MOSSE's computational efficiency is high, its algorithm accuracy is low, and it can only give the grey information of one channel. With the help of the correlation filter, Li and Zhu [7] utilized Histogram of Oriented Gradients (HOG), colour-naming features, and thus the size adaptive scheme to boost object tracking. Danelljan et al [8]. used HOG and the discriminative correlation filter to trace the article. SAMF and DSST solve the matter of deformation or change in scale when the tracking target is rotating. Further, they solve the matter of the tracker's inability to trace objects adaptively and thus the low operation speed. With the successful introduction of the deep-learning algorithm, some attempts join deep learning and thus the correlation filter for tracing the mobile target. Although these algorithms have better precision than the track algorithms supported the correlation filter, their training is time-consuming. Hence, these algorithms cannot track the article in real-time in a very real environment.

The motivation behind facial key-focuses acknowledgment is that getting the vital data about areas of eyebrows, eyes, lips also, nose in the face. With the improvement of profound learning, it is the first run-through for Sun et al. [9] to presented DCNN based on CNN to distinguish human facial key points. This calculation just perceives 5 facial key points, though its speed is quick. To accomplish a higher exactness for the facial key focuses acknowledgment, Zhou et al [10]. utilized FACE++ which streamlines DCNN furthermore, it can perceive 68 facial key points, however, this calculation incorporates an over the top model and the activity of this algorithm is muddled. Wu et al. [11] proposed Changed Convolutional Neural Systems (TCNN) which has a dependency on the Gaussian Blend Model (GMM) to make better the various layers of CNN. In any case, the vigour of TCNN relies upon information exorbitantly. Kowalski et al. [12] presented the Profound Arrangement System (DAN) to perceive the facial key points, which has preferred execution over different calculations. Tragically, DAN needs immense models and count dependent on convoluted capacities. To meet the prerequisite about genuine-time execution, we utilize dlib [13] to perceive facial key points.

Drowsiness detection within the drivers may be divided into two types: intrusive approaches and non-intrusive approaches. In intrusive approaches, drivers should wear or be in physical contact with the devices to detect the amount of their fatigue. Warwick et al. [14] proposed a system that used the Bio Harness 3 on the driver's body to gather data and measure their drowsiness using this. However, due to the worth of intrusive approaches and installation, there are some limitations that can't be implemented ubiquitously. So, the other method is to detect the driver drowsiness in a non-intrusive manner, where the measured object doesn't need to contact the motive force. As an example, Omidyeganeh et al. [15] used the camera to get the facial appearance of the drivers to detect the motive force drowsiness, but this method isn't in real-time. Zhang and Hua [16] used fatigue countenance reorganization supported Local Binary Pattern (LBP) features and Support Vector Machines (SVM) to estimate the motive force fatigue, but the complexity of this algorithm is larger than our algorithm. Akrout [17] and Mahdi and Oyini Mbouna et al. [18] used a fusion system for drowsiness detection supported state of eye and position of the head.

In 2016 Manu B.N [19] used a method that detects the facial landmarks using Haar feature-based cascade classifiers. At first, the algorithm must be trained by plenty of images with faces and without faces to coach the classifier to detect human faces more accurately. So along with the Haar feature-based classifiers, cascaded Adaboost classifier is exploited to acknowledge the region of the face then that image is divided into multiple numbers of rectangle areas, in any position and scale within the first image. The haar like features is healthier for real-time face detection. These may be calculated in step with the difference of total of constituent values inside parallelogram space and through this tactic, the Adaboost algorithm can allow all the face samples and it will ignore the non-facial samples of the images.

In 2015, Amna Rahman [20] has created a method to detect the drowsiness using Eye state detection with Eye blinking strategy. During this method, first, the image is converted to greyscale and the corners are discovered using Harris corner detection rule which may detect the corner at each facet and at the down curve of eyelid. After tracing the points then it'll create a straight line between the upper two points and locates the mid-point by calculation of the road, and it connects the mid-point with the lower point. Now for every image, it'll perform the identical procedure and it calculates the space,‘d’, from the mid-point to the lower point to see the eye state. At last, the selection for the attention state is created passionate about separation,’d’ determined. If the distance is zero or is near zero, the attention state is assessed as “closed” otherwise the attention state is identified as “open”. They need also invoked intervals or time to understand that the person is feeling drowsy or not. this is often done by the typical blink duration of an individual is 100-400 milliseconds (i.e. 0.1-0.4 of a second). Different from these methods, we created a straightforward method which makes the project even more accurate and inexpensive.

## III. PROPOSED SYSTEM

The proposed HOG- Linear SVM system aims at successfully identifying drivers who are in a state of drowsiness and immediately wake them up by sounding an alarm. This system requires a camera that is to be used to obtain a live running video stream of the driver behind the wheel. This video stream is analysed for detecting drowsiness on the driver.

### A. Initial Camera Setup:

The first step in the system would be to set up a camera facing the driver so that we can provide successful capturing of the driver’s face for the purpose of further processing. The camera must be set up in such a way that it is not intrusive, i.e., does not get in the way of the driver while on the road and must be placed in a proper manner so that the face captured is clear and so provides accurate results. A Raspberry Pi could be used for integrating the components in such a system, but since Raspberry Pi’s come with a low amount of RAM, the entire load of running all system operations, displaying the GUI, and also the process of handling the compiling for the operation would be too much. If dlib is compiled on a Raspberry Pi, an error message will be displayed. But this can be fixed if we update our Raspberry Pi system while claiming maximum memory and update swap file size. The camera could also be connected to a standard laptop computer. After this hardware setup is done, we can move on to the drowsiness detector algorithm to identify drowsy drivers by monitoring the video stream from the camera.

### B. Face Detection

The next step involved is detecting the face of the driver that is displayed on the video stream. For the purpose of face detection, we make use of the Histogram of Oriented Gradients (HOG) feature descriptor which uses the distribution of directions of gradients as features and tends to be more accurate and faster in operation than other algorithms. The HOG method suggested by Dalal and Triggs [3] in their seminal 2005 paper, Histogram of Oriented Gradients for Human Detection demonstrated that the Histogram of Oriented Gradients (HOG) image descriptor and a Linear Support Vector Machine (SVM) can be used to train highly accurate object classifiers — or in their particular study, human detectors. In this scenario, we will be tuning it to work for face detection.

The principles followed by the HOG facial descriptor is that the appearance and shape of the local object can often be characterized very well by the local intensity gradients distribution or directions of the edges, even without exact knowledge of the corresponding gradient or the edge positions.

In the practicality, this can be implemented by splitting the image window into tiny spatial regions (“cells”), for every cell accumulating a local 1-D histogram of gradient directions or edge orientations over the pixels of the cell. The combined entries of the histogram will end up forming the representation. For better unchangeability to illumination, shadowing, etc., it is also useful to contrast-normalize the local responses before using them. This can be successfully done by segregating a measure of local histogram “energy” over somewhat larger spatial regions (“blocks”) and using the results from this to normalize all the cells inside the block. The normalized descriptor blocks will be referred to as the Histogram of Oriented Gradient (HOG) descriptors. By covering the detection window with a dense (in fact, overlapping) grid of HOG descriptors and making use of the combined feature vector in a standard Linear SVM based window classifier gives us our human detection chain (see Fig. 1)

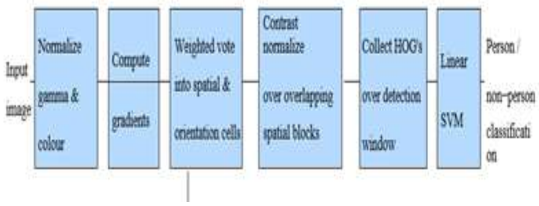


Fig.1. Hog and Linear SVM

C. Face landmark Detection and Extraction:

The next step involved after successful face detection is recognizing the facial landmarks and extraction of the desired facial landmarks. Finding facial landmarks can be done by several methods, but most of the methods work on labeling and localizing the regions such as the right eyebrow, left eyebrow, right eye, left eye, nose, mouth and, jaw. We use the facial landmark detector algorithm which is an implementation of the One Millisecond Face Alignment with an Ensemble of Regression Trees [4]. This detector algorithm is a part of the dlib library. This method works by manually labeling, specific (x,y) coordinates for the regions surrounding each facial structure and using this set of trained facial landmarks on an image. This detector available in dlib library estimates the location of the 68 (x,y) coordinates that are specific to each separate facial structure. The 68 facial landmark coordinates can be visualized in Fig.2.

Fig.2. Visualizing the 68 facial landmark coordinates

We can localize and extract the eye regions by making use of the specific facial indices for the left and right eye regions. The right eye can be accessed by using the coordinates [36,42] and, the left eye can be accessed by using the coordinates [42,48]. These indices are a part of the 68 points iBUG 300-w [21]-[23] dataset on which the facial landmark detector available in dlib library is trained. Irrespective of which dataset is used, if the shape predictor is trained properly on the input training data, the same dlib framework can be used.

D.Eye Aspect Ratio (E.A.R) Computation:

To detect if the driver's eye is closed or not, and to also successfully differentiate between standard eye blinks and eyes being closed during a state of drowsiness, we make use of an algorithm that uses a facial landmark detector. We compute a single, scalar quantity called eye aspect ratio (E.A.R) [2] that reflects whether the eye is closed or not. For each video frame, the landmarks of the eye regions are found, and the Euclidean distance using the height and width of the eye is calculated, which is the eye aspect ratio (E.A.R).

$$EAR = \frac{\|p2 - p6\| + \|p3 - p5\|}{2\|p1 - p4\|}$$

where p1, p2, p3, p4, p5, and p6 are the 2D landmark locations, depicted in Fig. 3.

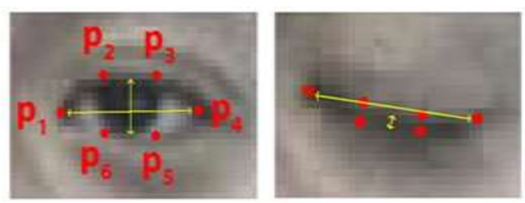


Fig.3. Open and closed eyes with landmarks p(i) automatically detected.

E.Drowsiness Evaluation and counter measures:

After we successfully compute E.A.R, we can use that value to evaluate the driver's state of drowsiness. The E.A.R value remains constant when the eye of the driver is open, but it starts to reduce to a value close to zero when the eye starts to close. E.A.R is invariant with respect to head and body posture. So, using these findings we can classify the eye state as closed when the E.A.R is zero or close to zero, otherwise the state is identified as open.

The final part is making the decision to sound the alarm or not. The average duration of a person's eye blink is 100-400 milliseconds, hence if the driver is in a state of drowsiness, their eye closure time is beyond this interval. In our system, the threshold is set 5 seconds, and if this is crossed the alarm is sounded and an alert regarding this will pop.

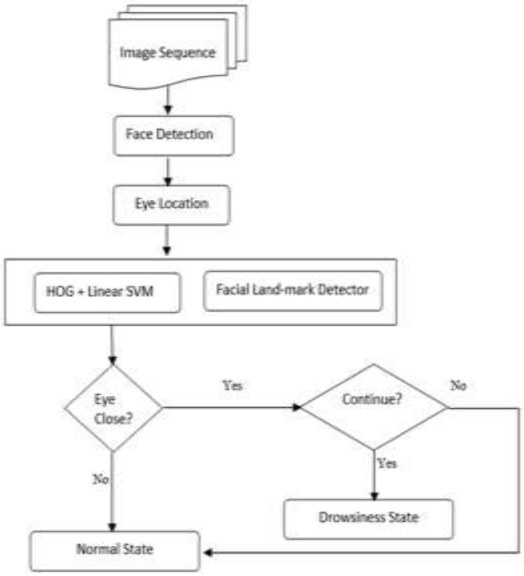


Fig.4. Block diagram of the proposed system for drowsiness detection

IV. RESULT & DISCUSSION

This section shows the results from the proposed system, and the discussions surrounding this.

A. Experimental Setup:

The proposed system works by making use of a system with 8 GB RAM, Intel i5 processor running Windows 10 Operating System. This technique is implemented using OpenCV, which is a cross-platform library containing programming functions specialized in real-time computer vision.

B. Dataset Description:

For this system, we make use of the 68 points iBUG 300-w dataset [21]- [23] on which the facial landmark detector available in dlib library is trained.

C. Experimental Results:

This section shows the sample result of the proposed system that makes use of OpenCV [24] to implement the process of monitoring the live video stream of the driver and detecting if the driver shows signs of drowsiness. First, successful detection of the face takes place followed by detection of the eye as shown in Fig. The eye aspect ratio (E.A.R) is calculated for every frame and as soon as the E.A.R value comes close to zero for more than the threshold time the alert is sounded along with the corresponding message.

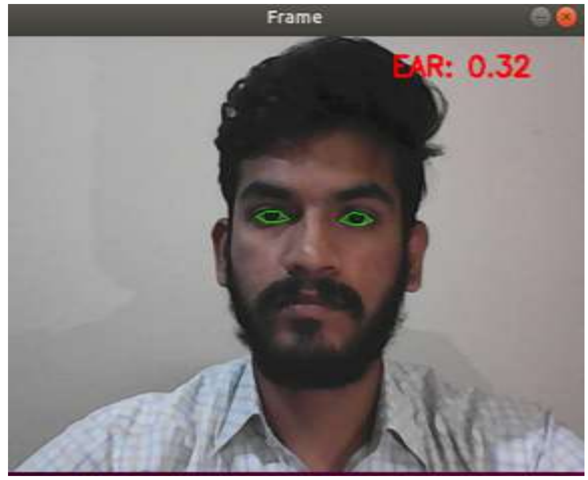


Fig.5. Eye region of the subject is detected, and landmarks are extracted. E.A.R is simultaneously calculated.

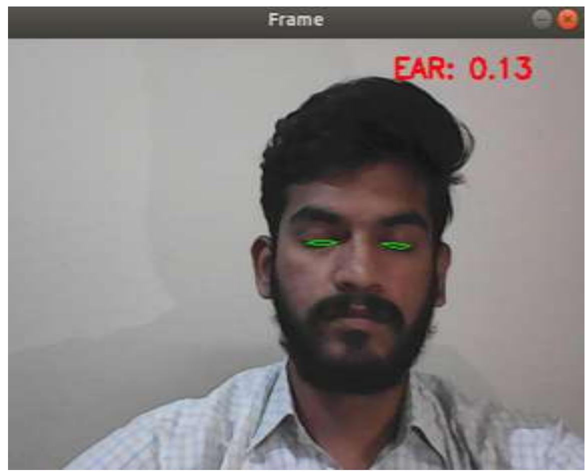


Fig.6. Subject's eyes are closed, and the E.A.R value is decreasing to almost zero.

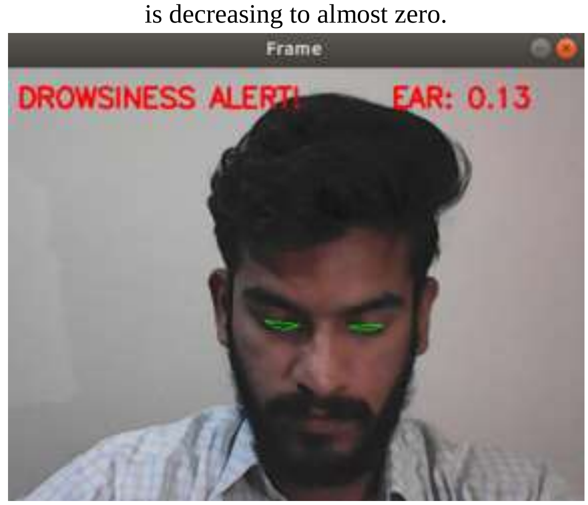


Fig.7. Alarm is sounded and alert is displayed when the subject's eyes have remained closed for more than the threshold time.

## V. CONCLUSION

We successfully developed a sleepiness detection system to look out for the fatigue within the drivers in real-time that helps avoid accidents due to sleepiness. The system has a face and eye blinking detection formula based upon the Histogram of Oriented Gradients (HOG) image descriptor and a Linear Support Vector Machine (SVM) facial detectors. These are precise enough to reliably estimate the positive photos of the face and level of eye openness. Moreover, the results show that it's having a higher accuracy even within the low light conditions.

## REFERENCES


[1] CISR- Central Road Research Institute, "https://www.crridom.gov.in/"

[2] Tereza Soukupova and Jan Cech, Center for Machine Perception, Department of Cybernetics Faculty of Electrical Engineering, Czech Technical University in Prague, "Real-Time Eye Blink Detection using Facial Landmarks", 21st Computer Vision Winter Workshop Luka Cehovin, Rok Mandeljc, Vitomir Struc (eds.) Rimske Toplice, Slovenia, February 3–5, 2016.

[3] Navneet Dalal and Bill Triggs, INRIA Rhone-Alps, ˆ 655 avenue de l'Europe, Montbonnot 38334, France, "Histograms of Oriented Gradients for Human Detection".

[4] Vahid Kazemi, Josephine Sullivan, "One millisecond face alignment with an ensemble of regression trees", IEEE Conference on Computer Vision and Pattern Recognition 2014.

[5] Bruce D. Lucas and Takeo Kanade, "An iterative image registration technique with an application to stereo vision," in Proceedings of the 7th International Joint Conference on Artificial Intelligence- Volume 2, San Francisco, CA, USA, 1981, IJCAI'81, pp. 674–679, Morgan Kaufmann Publishers Inc.

[6] D. S. Bolme, J. R. Beveridge, B. A. Draper, and Y. M. Lui, "Visual object tracking using adaptive correlation filters," in Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition, June 2010, pp. 2544–2550.

[7] Yang Li and Jianke Zhu, "A scale adaptive kernel correlation filter tracker with feature integration," in Computer Vision - ECCV 2014 Workshops, Lourdes Agapito, Michael M. Bronstein, and Carsten Rother, Eds., Cham, 2015, pp. 254–265, Springer International Publishing.

[8] M. Danelljan, G. HÃd'ger, F. S. Khan, and M. Felsberg, "Discriminative scale space tracking," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 8, pp. 1561–1575, Aug. 2017.

[9] Y. Sun, X. Wang, and X. Tang, "Deep convolutional network cascade for facial point detection," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, June 2013, pp. 3476–3483.

[10] E. Zhou, H. Fan, Z. Cao, Y. Jiang, and Q. Yin, "Extensive facial landmark localization with coarse-to-fine convolutional network cascade," in Proc. IEEE Int. Conf. Computer Vision Workshops, Dec. 2013, pp. 386–391.

[11] Y. Wu, T. Hassner, K. Kim, G. Medioni, and P. Natarajan, "Facial landmark detection with tweaked convolutional neural networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 12, pp. 3067–3074, Dec. 2018.

[12] M. Kowalski, J. Naruniec, and T. Trzcinski, "Deep alignment network: A convolutional neural network for robust face alignment," in Proc. IEEE Conf. Computer Vision and Pattern Recognition Workshops (CVPRW), July 2017, pp. 2034–2043.

[13] Davis E. King, "Dlib-ml: A machine learning toolkit," J. Mach. Learn. Res., vol. 10, pp. 1755–1758, Dec. 2009.

[14] B. Warwick, N. Symons, X. Chen, and K. Xiong, "Detecting driver drowsiness using wireless wearables," in Proc. IEEE 12th Int. Conf. Mobile Ad Hoc and Sensor Systems, Oct. 2015, pp. 585–588.

[15] M. Omidyeganeh, A. Javadtalab, and S. Shirmohammadi, "Intelligent driver drowsiness detection through fusion of yawning and eye closure," in Proc. IEEE Int. Conf. Virtual Environments Human-Computer Interfaces and Measurement Systems, Sept. 2011, pp. 1–6.

[16] Yan Zhang and Caijian Hua, "Driver fatigue recognition based on facial expression analysis using local binary patterns," Optik, vol. 126, pp. 4501– 4505, 2015.

[17] Belhassen Akrout and Walid Mahdi, "Spatio-temporal features for the automatic control of driver drowsiness state and lack of concentration," Machine Vision and Applications, vol. 26, no. 1, pp. 1–13, Jan 2015.

[18] R. Oyini Mbouna, S. G. Kong, and M. Chun, "Visual analysis of eye state and head pose for driver alertness monitoring," IEEE Transactions on Intelligent Transportation Systems, vol. 14, no. 3, pp. 1462–1469, Sept. 2013.

[19] B.N, "Facial Features Monitoring for Real Time Drowsiness Detection" Manu, 2016, 12th International Conference on Innovations in Information Technology (IIT) [Pg. 78-81] https://ieeexplore.ieee.org/document/7880030

[20] Amna Rahman, Department of Software Engineering, Fatima Jinnah Women University, "Real Time Drowsiness Detection using Eye Blink Monitoring" 2015 National Software Engineering Conference (NSEC 2015)https:// ieeexplore.ieee.org/document/7396336

[21] Sagonas, E. Antonakos, G, Tzimiropoulos, S. Zafeiriou, M. Pantic, "300 faces In-the-wild challenge: Database and results", Image and Vision Computing (IMAVIS), Special Issue on Facial Landmark Localisation "In-The-Wild". 2016.

[22] Sagonas, G. Tzimiropoulos, S. Zafeiriou, M. Pantic, "300 Faces in-the-Wild Challenge: The first facial landmark localization Challenge", Proceedings of IEEE Int'l Conf. on Computer Vision (ICCV-W), 300 Faces in-the-Wild Challenge (300-W). Sydney, Australia, December 2013.

[23] Sagonas, G. Tzimiropoulos, S. Zafeiriou, M. Pantic, "A semi-automatic methodology for facial landmark annotation", Proceedings of IEEE Int'l Conf. Computer Vision and Pattern Recognition (CVPR-W), 5th Workshop on Analysis and Modeling of Faces and Gestures (AMFG 2013). Oregon, USA, June 2013.

[24] I. Culjak, D. Abram, T. Pribanic, H. Dzapo and M. Cifrek, "A brief introduction to OpenCV," 2012 Proceedings of the 35th International Convention MIPRO, Opatija, 2012, pp. 1725-1730.