International Journal of Engineering and Techniques - Volume 8 Issue 4, July 2022

DRIVER DROWSINESS MONITORING SYSTEM USING VISUAL BEHAVIOUR AND MACHINE LEARNING

[1]B.SHIVANI, [2]B.ROHIT GOUD, [3]E.MAHIDHAR REDDY, [4]M.SAI RAM

Dept. of Information Technology, TKR College of Engineering and Technology, Hyderabad, Telangana

bhutam.shivani@gmail.com, bodigarohith1@gmail.com, Erigelamahireddy123@gmail.com

ABSTRACT

Drowsy driving is one of the major causes of road accidents and death. Hence, detection of driver's fatigue and its indication is an active research area. Most of the conventional methods are either vehicle based, or behavioral based or physiological based. Few methods are intrusive and distract the driver, some require expensive sensors and data handling. Therefore, in this study, a low cost, real time driver's drowsiness detection system is developed with acceptable accuracy. In the developed system, a webcam records the video and driver's face is detected in each frame employing image processing techniques. Facial landmarks on the detected face are pointed and subsequently the eye aspect ratio, mouth opening ratio and nose length ratio are computed and depending on their values, drowsiness is detected based on developed adaptive thresholding. Machine learning algorithms have been implemented as well in an offline manner. A sensitivity of 95.58% and specificity of 100% has been achieved in Support Vector Machine based classification.

1. INTRODUCTION

Drowsy driving is one of the major causes of deaths occurring in road accidents. The truck drivers who drive for continuous long hours (especially at night), bus drivers of long-distance route or overnight buses are more susceptible to this problem. Driver drowsiness is an overcast nightmare to passengers in every country. Every year, a large number of injuries and deaths occur due to fatigue related road accidents. Hence, detection of driver's fatigue and its indication is an active area of research due to its immense practical applicability. The basic drowsiness detection system has three blocks/modules: acquisition system, processing system and warning system.

Here, the video of the driver's frontal face is captured in acquisition system and transferred to the processing block where it is processed online to detect drowsiness. If drowsiness is detected, a warning or alarm is sent to the driver from the warning system. Generally, the methods to detect drowsy drivers are classified in three types; vehicle based, behavioral based and physiological based. In vehicle-based method, a number of metrics like steering wheel movement, accelerator or brake pattern, vehicle speed, lateral acceleration, deviations from lane position etc. are monitored continuously. Detection of any abnormal change in these values is considered as driver drowsiness. This is a nonintrusive measurement as the sensors are not attached on the driver. In behavioral based method, the visual behavior of the driver i.e., eye blinking, eye closing, yawn, head bending etc. are analyzed to detect drowsiness. This is also nonintrusive measurement as simple camera is used to detect these features. In physiological based method [8,9], the physiological signals like Electrocardio - gram (ECG), Electrooculogram (EOG), Electroencephalogram (EEG), heartbeat, pulse rate etc. are monitored and from these metrics, drowsiness or fatigue level is detected. This is intrusive measurement as the sensors are attached on the driver which will distract the driver. Depending on the sensors used in the system, system cost as well as size will increase. However, inclusion of more parameters/features will increase the accuracy of the system to a certain extent. These factors motivate us to develop a low-cost, real time driver's drowsiness detection system with acceptable accuracy. Hence, we have proposed a webcam-based system to detect driver's fatigue from the face image only using image processing and machine learning techniques to make the system low-cost as well as portable.

2. LITERATURE SURVEY

Author: W. L. Ou et.al, An intelligent video-based drowsy driver detection system, which is unaffected by various illuminations, is developed in this study. Even if a driver wears glasses, the proposed system detects the drowsy conditions effectively.

By a near-infrared-ray (NIR) camera, the proposed system is divided into two cascaded computational procedures: the driver eyes detection and the drowsy driver detection. The average open/closed eyes detection rates without/with glasses are 94% and 78%, respectively, and the accuracy of the drowsy status detection is up to 91%. By implementing on theFPGA-based embedded platform, the processing speed with the 640×480 format video is up to 16 frames per second (fps) after software optimizations.

Author: W. B. Hornget.al A: vision-based real-time driver fatigue detection system is proposed for driving safely. The driver's face is located, from color images captured in a car, by using the characteristic of skin colors. Then, edge detection is used to locate the regions of eyes. In addition to being used as the dynamic templates for eye tracking in the next frame, the obtained eyes' images are also used for fatigue detection in order to generate some warning alarms for driving safety. The system is tested on a Pentium III 550 CPU with 128 MB RAM. The experiment results seem quite encouraging and promising. The system can reach 20 frames per second for eye tracking, and the average correct rate for eye location and tracking can achieve 99.1% on four test videos. The correct rate for fatigue detection is l00%, but the average precision rate is 88.9% on the test videos.

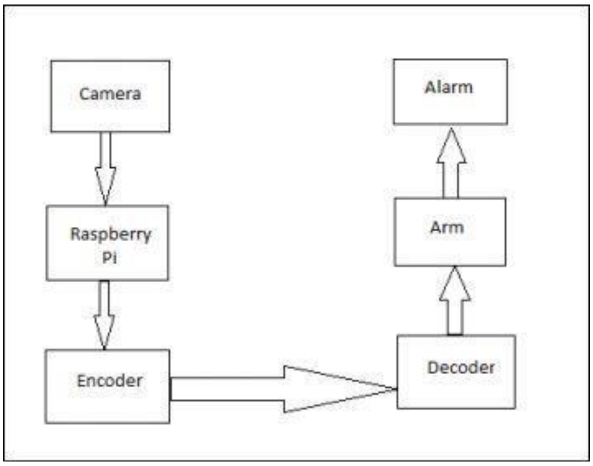
Author: S. Singh and N. P. papa Nikolopoulos: is described a non-intrusive vision-based system for the detection of driver fatigue. The system uses a color video camera that points directly rewards the driver's face and monitors the driver's eyes in order to detect micro-sleeps (short periods of sleep). The system deals with skin-color information in order to search for the face in the input space. After segmenting the pixels with skin like color, we perform blob processing in order to determine the exact position of the face. We reduce the search space by analyzing the horizontal gradient map of the face, consideringthe knowledge that eye regions in the face present a great change in the horizontal intensity gradient. In order to find and track the location of the pupil, we use gray scale model matching. We also use the same pattern recognition technique to determine whether the eye is open or closed. If the eyes remain closed for an abnormal period of time (5-6 sec), the system draws the conclusion that the person is falling asleep and issues a warning signal.

Author: B. Alshaqaqiet.al: Drowsiness and Fatigue of drivers are amongst the significant causes of road accidents. Every year, they increase the amounts of deaths and fatalities injuries globally. In this paper, a module for Advanced Driver Assistance System (ADAS) is presented to reduce the number of accidents due to drivers fatigue and hence increase the transportation safety; this system deals with automatic driver drowsiness detection based on visual information and Artificial Intelligence. Authors propose an algorithm to locate, track, and analyze both the drivers face and eyes to measure PERCLOS, a scientifically supported measure of drowsiness associated with slow eye closure.

Alshaqaqi,et.al. Drowsiness detection has many implications including reducing roads traffic accidents importance. Using image processing techniques is amongst the new and reliable methods in sleepy face. The present pilot study was done to investigate sleepiness and providing images of drivers' face, employing virtual-reality driving simulator. In order to detecting level of sleepiness according to the signal, information related to 25 drivers was recorded with imaging rate of 10 fps, Moreover, on average 3000 frames were analyzed for each driver. The frames were investigated by transforming in grey scale space and based on the Cascade and Viola & Jones techniques and the images characteristics were extracted using Binary and Histogram methods. The MPL neural network was applied for analyzing data.70% of information related to each driver were inserted to the network of which 15% for test and 15% for validation. In the last stage the accuracy of 93% of the outputs were evaluated. The intelligent detection and usage of various criteria in long-term time frame are of the advantages of the present study, comparing to other research. This is helpful in early detection of sleepiness and prevents the irrecoverable losses by alarming. In existing system, the driver drowsiness detection system involves controlling accident due to unconsciousness through Eye blink. Here one eye blink sensor is fixed in vehicle where if driver loses consciousness, then it alerts the driver through buzzer to prevent vehicle from accident. In future we can implement Drowsiness Detection System in aircraft in order to alert. Causes irritation in the eye, May damage retina Highly expensive and distract the driver

3. PROPOSED SYSTEM

A block diagram of the proposed driver drowsiness monitoring system has been depicted in Fig. At first, the video is recorded using a webcam. The camera will be positioned in front of the driver to capture the front face image. From the video, the frames are extracted to obtain 2-D images. Face is detected in the frames using histogram of oriented gradients (HOG) and linear support vector machine (SVM) for object detection [10]. After detecting the face, facial landmarks [11] like positions of eye, nose, and mouth are marked on the images. From the facial landmarks, eye aspect ratio, mouth opening ratio and position of the head are quantified and using these features and machine learning approach, a decision is obtained about the drowsiness of the driver. If drowsiness is detected, an alarm will be sent to the driver to alert him/her.



The block diagram of the proposed system has been shown in the above Figure1. The camera captures the image and sends to the Raspberry pi which consists of 32-bit memory card installed with Open CV which helps in image processing. ARM used is the LPC2148 which is the microcontroller. If the signal crosses threshold of 2 sec, it will automatically make the alarm beep and the speed of the vehicle gets reduced. Otherwise, that signal is rejected and next signal is processed.

Working: Driver's face is continuously monitored using a video camera. In order to• detect the drowsiness the first step is to detect the face using the series of frame shots taken by the camera. Then the location of the eyes is detected, and retina of the eye is continuously monitored. The captured image is sent to the Raspberry Pi board for image processing. The raspberry Pi converts the received image to digital signal using Open CV. The digital signal is transmitted from transmitter to the receiver. Both the transmitter and the receiver are paired up. The signal is then passed to the LPC2148, the microcontroller. If the signal crosses the threshold of two seconds, then the alarm beeps and the speed of the vehicle is automatically reduced.

Data Acquisition

The video is recorded using webcam (Sony CMU-BR300) and the frames are extracted and processed in a laptop. After extracting the frames, image processing techniques are applied on these 2D images. Presently, synthetic driver data has been generated. The volunteers are asked to look at the webcam with intermittent eye blinking, eye closing, yawning and head bending. The video is captured for 30 minutes duration.

Face Detection

After extracting the frames, first the human faces are detected. Numerous online face detection algorithms are there. In this study, histogram of oriented gradients (HOG) and linear SVM method [10] is used. In this method, positive samples of fixed window size are taken from the images and HOG descriptors are computed on them. Subsequently, negative samples (samples that do not contain the required object to be detected i.e., human face here) of same size are taken and HOG descriptors are calculated. Usually the number of negative samples is very greater than number of positive samples. After obtaining the features for both the classes, a linear SVM is trained for the classification task. To improve the accuracy of SVM, hard negative mining is used. In this method, after training, the classifier is tested on the labeled data and the false positive sample feature values are used again for training purpose. For the test image, the fixed size window is translated over the image and the classifier computes the output for each window location. Finally, the maximum value output is considered as the detected face and a bounding box is drawn around the face. This non-maximum suppression step removes the redundant and overlapping bounding boxes.

Facial Landmark Detection

Marking After detecting the face, the next task is to find the locations of different facial features like the corners of the eyes and mouth, the tip of the nose and so on. Prior to that, the face images should be normalized in order to reduce the effect of distance from the camera, non-uniform illumination and varying image resolution. Therefore, the face image is resized to a width of 500 pixels and converted to grayscale image. After image normalization, ensemble of regression trees [11] is used to estimate the landmark positions on face from a sparse subset of pixel intensities. In this method, the sum of square error loss is optimized using gradient boosting learning. Different priors are used to find different structures. Using this method, the boundary points of eyes, mouth and the central line of the nose are marked and the number of points for eye, mouth and nose are given in Table I. The facial landmarks are shown in Fig 2. The red points are the detected landmarks for further processing.

Table I: Facial landmark points

Parts	Landmark Points
Mouth	[13-24]
Right eye	[1-6]
Left eye	[7-12]
Nose	[25-28]

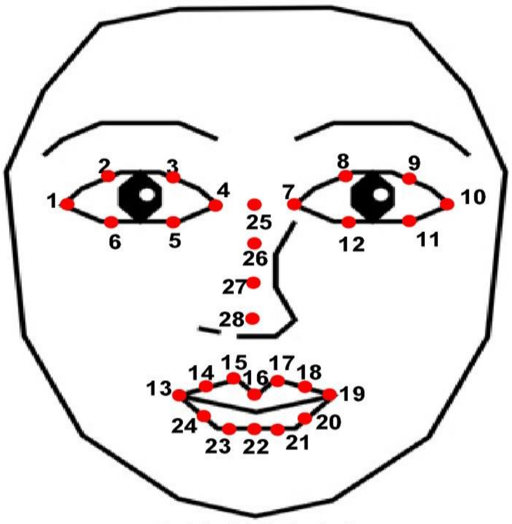


Fig. 2 The facial landmark points

Feature Extraction

After detecting the facial landmarks, the features are computed as described below. Eye aspect ratio (EAR): From the eye corner points, the eye aspect ratio is calculated as the ratio of height and width of the eye as given by where represents point marked as i in facial landmark and is the distance between points marked as i and j. Therefore, when the eyes are fully open, EAR is high value and as the eyes are closed, EAR value goes towards zero. Thus, monotonically decreasing EAR values indicate gradually closing eyes and it's almost zero for completely closed eyes (eye blink). Consequently, EAR values indicate the drowsiness of the driver as eye blinks occur due to drowsiness. Mouth opening ratio (MOR): Mouth opening ratio is a parameter to detect yawning during drowsiness. Similar to EAR, it is calculated as defined, it increases rapidly when mouth opens due to yawning and remains at that high value for a while due to yawn (indicating that the mouth is open) and again decreases rapidly.

Towards zero. As yawn is one of the characteristics of drowsiness, MOR gives a measure regarding driver drowsiness. Head Bending: Due to drowsiness, usually driver's head tilts (forward or backward) with respect to vertical axis. So, from the head bending angle, driver drowsiness can be detected. As the projected length of nose on the camera focal plane is proportional to this bending, it can be used as a measure of head bending. In normal condition, our nose makes an acute angle with respect to focal plane of the camera. This angle increases as the head moves vertically up and decreases on moving down. Therefore, the ratio of nose length to an average nose length while awake is a measure of head bending and if the value is greater or less than a particular range, it indicates head bending as well as drowsiness. From the facial landmarks, the nose length is calculated and it is defined as 28 25 nose length (p p) NLR average nose length − = The average nose length is computed during the setup phase of the experiment as described in the next sub-section.

Classification

After computing all the three features, the next task is to detect drowsiness in the extracted frames. In the beginning, adaptive thresholding is considered for classification. Later, machine learning algorithms are used to classify the data. For computing the threshold values for each feature, it is assumed that initially the driver is in complete awake state. This is called setup phase. In the setup phase, the EAR values for first three hundred (for 10s at 30 fps) frames are recorded. Out of these three hundred initial frames containing face, average of 150 maximum values is considered as the hard threshold for EAR. The higher values are considered so that no eye closing instances will be present. If the test value is less than this threshold, then eye closing (i.e., drowsiness) is detected. As the size of eye can vary from person to person, this initial setup for each person will reduce this effect. Similarly, for calculating threshold of MOR, since the mouth may not be open to its maximum in initial frames (setup phase) so the threshold is taken experimentally from the observations. If the test value is greater than this threshold then yawn (i.e., drowsiness) is detected. Head bending feature is used to find the angle made by head with respect to vertical axis in terms of ratio of projected nose lengths. Normally, NLR has values from 0.9 to 1.1 for normal upright position of head and it increases or decreases when head bends down or up in the state of drowsiness. The average nose length is computed as the average of the nose lengths in the setup phase assuming that no head bending is there. After computing the threshold values, the system is used for testing. The system detects the drowsiness if in a test frame drowsiness is detected for at least one feature. To make this Thresholding more realistic, the decision for each frame depends on the last 75 frames. If at least 70 frames (out of those 75) satisfy drowsiness conditions for at least one feature, then the system gives drowsiness detection indication and the alarm. To make this thresholding adaptive, another single threshold value is computed which initially depends on EAR Threshold value. The average of EAR values is computed as the average of 150 maximum values out of 300 frames in the setup phase. Then offset is determined heuristically, and the threshold is obtained as offset subtracted from the average value. Driver safety is at risk when EAR is below this threshold. This EAR threshold value increases slightly with each yawning and head bending upto a certain limit. As each yawning and head bending is distributed over multiple frames, so yawning and head bending of consecutive frames are considered as single yawn and head bending and added once in the adaptive threshold. In a test frame, if EAR value is less than this adaptive threshold value, then drowsiness is detected, and an alarm is given to the driver. Sometimes it may happen that when the head is too low due to bending, the system is unable to detect the face. In such situation, previous three frames are considered and if head bending was detected in those three frames, drowsiness alarm will be shown. Table II illustrates this calculation for determining the adaptive threshold. Table II: Threshold for the computed parameters EAR from setup phase (average of 150 maximum values out of 300 frames) 0.34 Threshold=EAR- offset 0.34-.045=0.295 At Yawning,(MOR> 0.6)

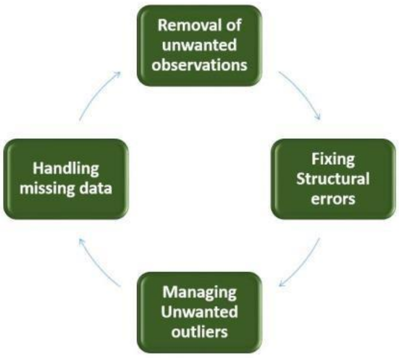
*Threshold=Threshold +0.002 *Max bound* exist At Head Bending, (NLR1.2) *Threshold=Threshold +0.001 *Max bound* exist Apart from using thresholding, the machine learning algorithms are used to detect drowsiness as well. The EAR, MOR and NLR values are stored for the synthetic test data along with actual drowsiness annotation. Prior to classification, statistical analysis of the features has been done. At first, principal component analysis is used to transform the feature space into an independent one. After transforming the feature values, student's t test is used to test whether the features are statistically significant for the two classes. As all the three features are statistically significant at 5% level of significance, all the three features are used for classification using Bayesian classifier, Fisher's linear discriminate analysis and Support vector Machine.

We proposed as an alternative to the user-based neighborhood approach. We first consider the dimensions of the input and output of the neural network. In order to maximize the amount of training data we can feed to the network, we consider a training example to be a user profile (i.e. a row from the user-item matrix R) with one rating withheld. The loss of the network on that training example must be computed with respect to the single withheld rating. The consequence of this is that each individual rating in the training set corresponds to a training example, rather than each user. As we are interested in what is essentially a regression, we choose to use root mean squared error (RMSE) with respect to known ratings as our loss function. Compared to the mean absolute error, root mean squared error more heavily penalizes predictions which are further off. We reason that this is good in the context of recommendersystem because predicting a high rating for an item the user did not enjoy significantly impacts the quality of the recommendations. On the other hand, smaller errors in prediction likely result in recommendations that are still useful—perhaps the regression is not exactly correct, but at least the highest predicted rating are likely to be relevant to the user

Data Processing is a task of converting data from a given form to a much more usable and desired form i.e. making it more meaningful and informative. Using Machine Learning algorithms, mathematical modeling and statistical knowledge, this entire process can be automated. The output of this complete process can be in any desired form like graphs, videos, charts, tables, images and many more, depending on the task we are performing and the requirements of the machine. This might seem to be simple but when it comes to really big organizations like Twitter, Facebook, Administrative bodies like Paliament, UNESCO and health sector organizations, this entire process needs to be performed in a very structured manner.

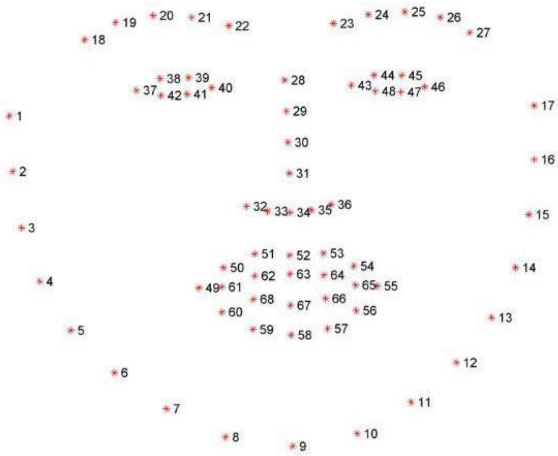
If we have a well-cleaned dataset, we can get desired results even with a very simple algorithm, which can prove very beneficial at times. Obviously, different types of data will require different types of cleaning. However, this systematic approach can always serve as a good starting point.

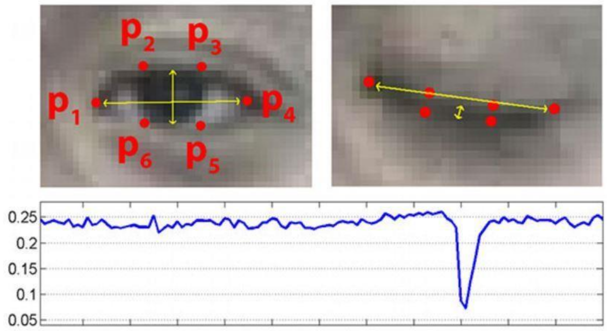
Steps involved in Data Cleaning

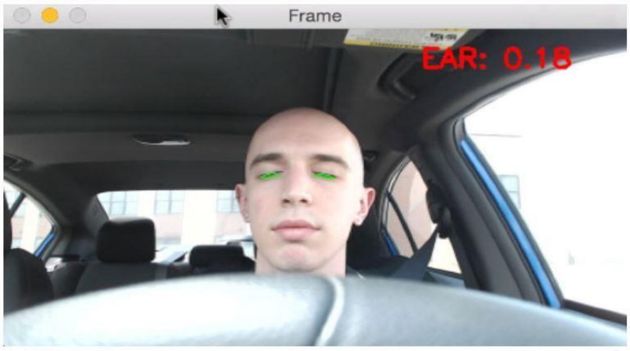


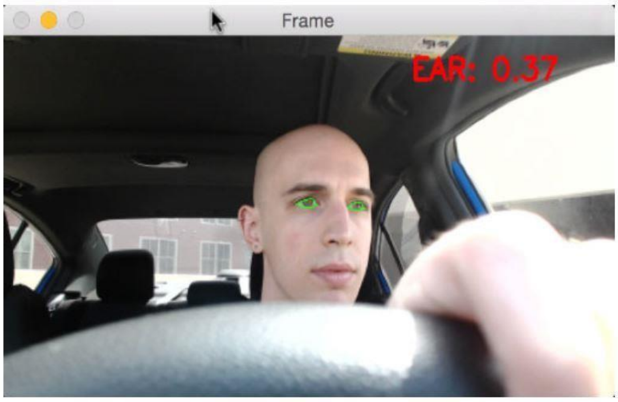
4. RESULTS AND DISCUSSION

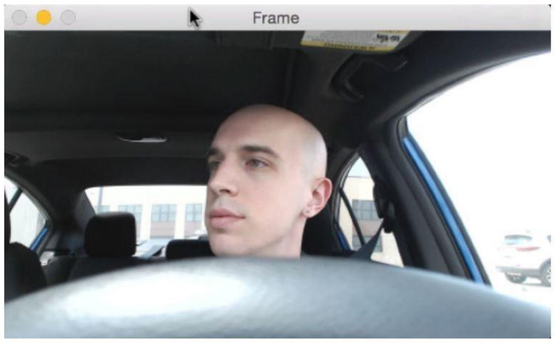
The actual purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner.











Test Cases:

Test Id	Test case Description	Test Data	Expected Result	Actual Result	Pass/Fail
1	Input video	Video testing	Video loaded	Successfully loaded	Pass

2	Face detection	Testing faces	Faces tested	Successfully Tested	Pass
3	Facial land marking	Eyes detect in face	Eyes detected	Successfully detected	Pass
4	Calculate the EAR	Calculate the Eye Aspect Ratio(EAR)	Calculated Eye Aspect Ratio	Success fully detected	Pass
5	Drowsy or not Drowsy	Eye close or not close tested	Drowsy or not drowsy detected	Successfully detected	Pass
6	Alarm	Alarm ringing based on the eye closed	Alarm ringing	Successfully alarm ringing	Pass

The proposed system has been developed and tested with the generated data. The webcam is connected with the laptop for further processing and classification of the video streaming in an online manner. Subsequently, the feature values are stored for statistical analysis and classification as well. One frame from the normal or awake state is shown in Fig. 3. The feature values for these frame Different drowsy conditions are displayed in Fig. 4. Figure 4(a) illustrates an example of drowsiness alert due to yawn and Fig. 4(b) illustrates an example of same due to eye closing. An example of head bending and detecting it as drowsiness alert is shown in Fig. 4(c). Figure 4(d) depicts the condition when the head is too low due to bending and drowsiness is detected as described in previous section. Table III illustrates sample values of the parameters for different states. Table III: Sample values of different parameters for different states State EAR MOR NLR Normal 0.35 0.34 1.003 Yawning 0.22 0.77 0.76 Eye Closed 0.15 0.419 0.876 Head Bending 0.15 0.577 0.66 the developed system can also detect drowsiness with persons wearing spectacles as depicted.

The developed algorithm has been tested on INVEDRIFAC dataset. It is a video and image database of faces of in vehicle automotive drivers. The developed algorithm has been tested on 6 different driver videos. The performance of the same on this data is with acceptable accuracy. Also, the videos have different illumination conditions which indicate that the algorithm can perform well even at low illumination condition. Subsequently statistical analysis and classification of the features into two classes have been explored as well. As the features are correlated, principal component analysis has been used to transform the feature space into an independent one. The independent features are statistically significant at 5% level of significance. Bayesian classifier, Fisher's linear discriminate analysis (FLDA) and Support Vector Machine (SVM) with linear kernel have been used for classification. Presently, this has been done in offline manner on the stored data. Two-third data is used for training and one-third data is used for testing the algorithms. The classifier results are given in Table IV. Sensitivity is calculated as the ratio of correctly classifying drowsy states out of all actual drowsy states and specificity is computed as the ratio of correctly classifying awake states out of all actual awake states. Overall accuracy is computed as the correctly classified states out of all the frames. It is evident that the overall accuracy of FLDA and SVM is better than that of Bayesian whereas Bayesian gives best sensitivity of 97%. However, the specificity of Bayesian is quite low (56%). This may be due to the error in approximating the probability distributions. Due to low specificity, the alarm may ring when actually drowsiness is not there. This can be disturbing to the driver.

CONCLUSION

In this paper, a low-cost, real-time driver drowsiness monitoring system has been proposed based on visual behavior and machine learning. Here, visual behavior features like eye aspect ratio, mouth opening ratio and nose length ratio are computed from the streaming video, captured by a webcam. An adaptive thresholding technique has been developed to detect driver drowsiness in real time. The developed system works accurately with the generated synthetic data. Subsequently, the feature values are stored and machine learning algorithms have been used for classification. Bayesian classifier, FLDA and SVM have been explored here. It has been observed that FLDA and SVM outperform Bayesian classifier. The sensitivity of FLDA and SVM is 0.896 and 0.956 respectively whereas the specificity is 1 for both. As FLDA and SVM give better accuracy, work will be carried out to implement them in the developed system to do the classification (i.e., drowsiness detection) online. Also, the system will be implemented in hardware to make it portable for car system and pilot study on drivers will be carried out to validate the developed system

REFERENCES

1. *L. Ou, M. H. Shih, C. W. Chang, X. H. Yu, C. P. Fan, "Intelligent Video-Based Drowsy Driver Detection System under Various Illuminations and Embedded Software Implementation", 2015 international Conf. on Consumer Electronics - Taiwan, 2015.*
2. *W. B. Horng, C. Y. Chen, Y. Chang, C. H. Fan, "Driver Fatigue Detection based on Eye Tracking and Dynamic*

Template Matching", IEEE International Conference on Networking,, Sensing and Control, Taipei, Taiwan, March 21-23, 2004.

3. *S. Singh, N. P. papanikolopoulos, "Monitoring Driver Fatigue using Facial Analysis Techniques", IEEE Conference on Intelligent Transportation System, pp 314-318.*
4. *B. Alshaqaqi, A. S. Baquhaizel, M. E. A. Ouis, M. Bouumehed, A. Ouamri, M. Keche, "Driver Drowsiness Detection System", IEEE International Workshop on Systems, Signal Processing and their Applications, 2013.*
5. *M. Karchani, A. Mazloumi, G. N. Saraji, A. Nahvi, K. S. Haghighi, B. M. Abadi, A. R. Foroshani, A. Niknezhad, "The Steps of Proposed Drowsiness Detection System Design based on Image Processing in Simulator Driving", International Research Journal of Applied and Basic Sciences, vol. 9(6), pp 878-887, 2015.*
6. *R. Ahmad, and J. N. Borole, "Drowsy Driver Identification Using Eye Blink Detection," IJISET - International Journal of Computer Science and Information Technologies,vol. 6, no. 1, pp. 270-274, Jan. 2015.*
7. *A. Abas, J. Mellor, and X. Chen, "Non-intrusive drowsiness detection by employing Support Vector Machine," 2014 20th International Conference on Automation and Computing (ICAC), Bedfordshire, UK, 2014, pp. 188- 193.*
8. *A. Sengupta, A. Dasgupta, A. Chaudhuri, A. George, A. Routray, R. Guha; "A Multimodal System for Assessing Alertness Levels Due to Cognitive Loading", IEEE Trans. on Neural Systems and Rehabilitation Engg., vol. 25 (7), pp 1037-1046, 2017.*