JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Realtime Driver Drowsiness Monitoring System using Visual Behaviour and Machine Learning Techniques

Srinivasarao Udara[1], M.Naresh[2]

[1]Department of Computer Science &Engineering, NIE, Macherla, Andhra Pradesh, India

[2] Department of Computer Science &Engineering, NIE, Macherla, Andhra Pradesh, India

***Corresponding author mail: drsrudara@gmail.com**

Abstract

Fatigue-induced driving is a significant contributor to road accidents and fatalities. Consequently, there is ongoing research into detecting driver fatigue and signaling its onset. Traditional approaches predominantly fall into three categories: vehicle-based, behavioral-based, and physiological-based. Many methods either intrusively disrupt the driver or necessitate costly sensors and complex data processing. Thus, this study focuses on developing a cost-effective, real-time drowsiness detection system with satisfactory accuracy. Within the system architecture, a webcam captures video footage, utilizing image processing techniques to detect the driver's face in each frame. Facial landmarks are then identified, allowing computation of metrics such as eye aspect ratio, mouth opening ratio, and nose length ratio. Drowsiness is determined using adaptive thresholding based on these metrics. Additionally, offline machine learning algorithms have been integrated. The Support Vector Machine based classification achieves a sensitivity of 95.58% and specificity of 100%..

Keywords—Drowsiness detection, visual behaviour, eye aspect ratio, mouth opening ratio, nose length ratio.

1. Introduction

Fatigue-induced driving stands as a significant contributor to fatalities in road accidents. Truck drivers enduring extended shifts, particularly during nighttime and long-haul bus operators face heightened vulnerability to this issue. Driver drowsiness casts a shadow over passengers' safety worldwide, leading to a substantial number of injuries and fatalities annually in fatigue-related road incidents. Therefore, the detection of driver fatigue and its signaling remains a vibrant field of research, owing to its considerable real-world utility. The fundamental drowsiness detection system comprises three modules: acquisition, processing, and warning. In the acquisition phase, video footage of the driver's frontal face is captured and transmitted to the processing module, where real-time analysis identifies signs of drowsiness. Upon detection, the warning system promptly dispatches alerts or alarms to the driver. Typically, methods for identifying drowsy drivers fall into three categories: vehicle-based, behavioral-based, and physiological-based approaches [1-10].

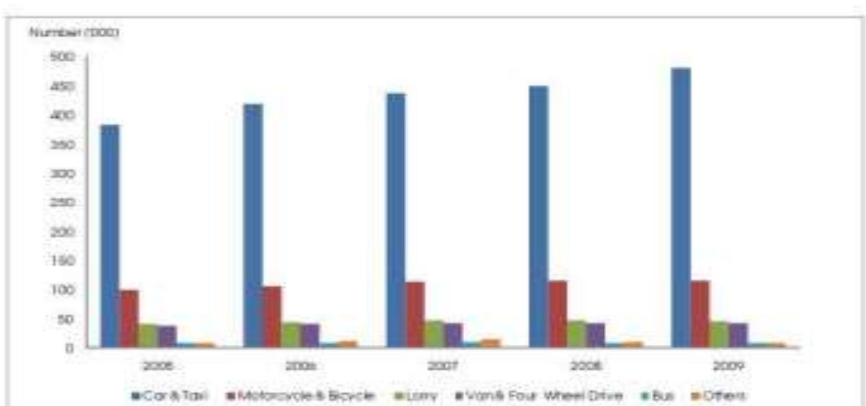
In the vehicle-based approach, various metrics such as steering wheel movements, patterns of accelerator and brake usage, vehicle speed, lateral acceleration, and deviations from lane position are continually monitored. Any abnormal changes in these parameters are indicative of driver drowsiness. This method is non-intrusive, as the sensors are not directly attached to the driver. In the behavioral-based method, the visual behavior of the driver, including eye blinking, eye closure, yawning, and head movement, is analyzed to detect signs of drowsiness.

This method is also non-intrusive, as it relies on a simple camera to detect these features without directly interfering with the driver. In the physiological-based approach, signals such as Electrocardiogram (ECG), Electrooculogram (EOG), Electroencephalogram (EEG), heartbeat, and pulse rate are monitored to assess the level of drowsiness or fatigue. However, this method is considered intrusive, as it requires attaching sensors to the driver, potentially causing distraction. Additionally, depending on the sensors utilized, the system's cost and size may increase. Nevertheless, integrating additional parameters or features can enhance the system's accuracy to some degree. This impetus drives our endeavor to create a cost-effective, real-time driver's drowsiness detection system with satisfactory precision. Therefore, we propose a webcam-based approach solely utilizing facial images to detect driver fatigue, employing image processing and machine learning techniques to ensure affordability and portability.

Driver drowsiness poses a significant risk in transportation systems, being recognized as a direct or contributing factor to road accidents. Fatigue can substantially impair reaction time, diminish alertness, and compromise a driver's judgment, with women consistently experiencing harassment in our society. Across all cities, there exist areas where harassment of women is a prevalent issue. Studies conducted in urban areas have revealed that 60% of women experience harassment and feel unsafe when leaving their homes. These instances range from verbal remarks to body shaming, representing a concern for our entire society.

The objectives of this study are to propose methodologies for detecting fatigue and drowsiness during driving, to examine the correlation between eye and mouth movements captured in video images from participants in driving simulation experiments as potential indicators of fatigue and drowsiness, to explore the physiological manifestations of fatigue and drowsiness, and to devise a system utilizing eye closure and yawning as indicators for detecting fatigue and drowsiness.

Drowsiness is characterized by a state of almost falling asleep, accompanied by a strong inclination towards sleep. It encompasses two specific interpretations: the typical state preceding falling asleep and a persistent condition where this state occurs regardless of daily rhythms. Sleepiness poses a significant hazard when engaging in tasks demanding sustained focus, such as driving. If a person becomes sufficiently fatigued while driving, they are likely to experience drowsiness, thereby heightening the risk of road accidents.



Source: Department of Road Safety Malaysia

Figure 1: Statistic of Road Accident from 2005 to 2009

Figure 1 illustrates the statistics of road accidents in Malaysia from 2005 to 2009, sourced from MIROS (Malaysia Institute of Road Safety). The data reveals a consistent rise in the number of vehicles involved in accidents each year. Specifically, car and taxi categories account for nearly 400,000 recorded cases of road accidents, with a steady increase observed annually. By 2009, MIROS recorded close to 500,000 incidents of road accidents.

2. Methodology

The proposed driver drowsiness monitoring system is outlined in the block diagram presented in Fig. 2. Initially, video footage is captured using a webcam positioned in front of the driver to capture frontal face images. Frames from the video are then extracted to generate 2-D images. Face detection within the frames is performed using Histogram of Oriented Gradients (HOG) coupled with a linear Support Vector Machine (SVM) for object detection [10]. Once the face is detected, facial landmarks [11], including the positions of the eyes, nose, and mouth, is annotated on the images. Subsequently, measurements such as eye aspect ratio, mouth opening ratio, and head position are computed from these facial landmarks. Leveraging these features and a machine learning approach, a decision regarding the driver's drowsiness is derived.

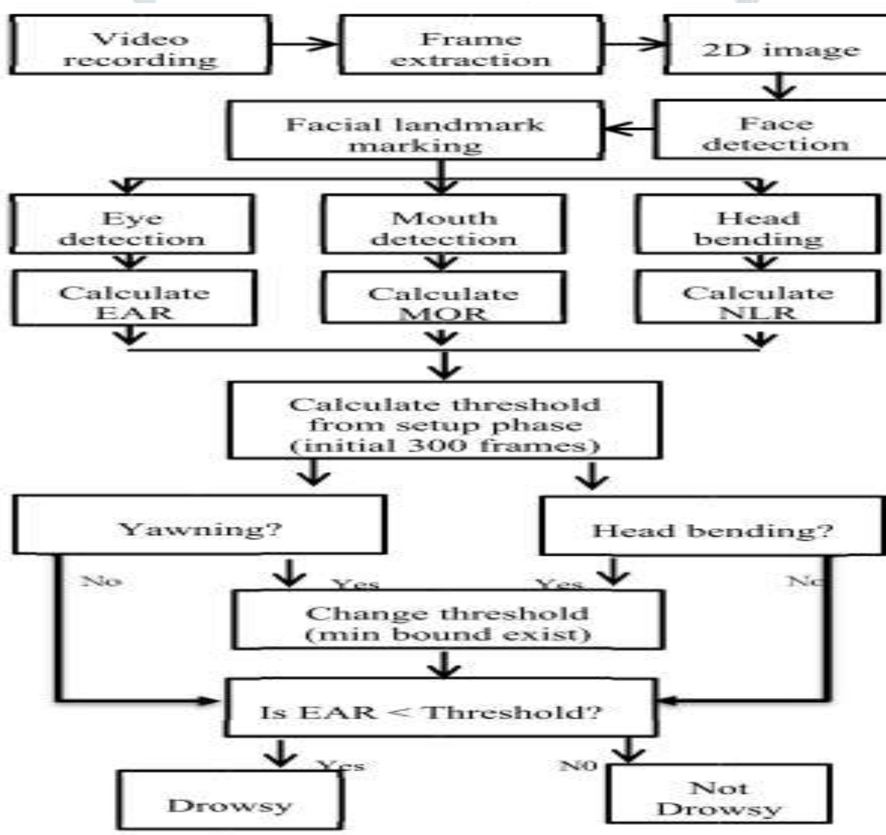


Figure 2: The block diagram of the proposed drowsiness detection system

A. Data Acquisition

The video footage is captured using a Sony CMU-BR300 webcam, and the frames are subsequently extracted and processed on a laptop. Following frame extraction, image processing techniques are applied to these 2D images. Currently, synthetic driver data has been generated for experimentation purposes. Volunteers are instructed to gaze at the webcam while intermittently performing actions such as eye blinking, eye closing, yawning, and head bending. The video is recorded for duration of 30 minutes.

B. Face Detection

After frame extraction, the initial step involves detecting human faces. Various online face detection algorithms are available, but for this study, the histogram of oriented gradients (HOG) coupled with a linear SVM method [10] is employed. In this approach, positive samples comprising 118,728 fixed window sizes are selected from the images, and HOG descriptors are computed on them. Subsequently, negative samples—those lacking the object of interest for detection, in this case, human faces—are taken at the same size, and HOG descriptors are computed. Typically, the number of negative samples far exceeds the number of positive samples. Following the acquisition of features for both classes, a linear SVM is trained for the classification task. To enhance the SVM's accuracy, hard negative mining is implemented. In this technique, post-training, the classifier is evaluated on labeled data, and the feature values of false positive samples are reintroduced for training purposes. During testing, a fixed-size window is traversed across the image, and the classifier computes the output for each window position. Ultimately, the maximum output value is designated as the detected face, and a bounding box is delineated around it. A non-maximum suppression step is then executed to eliminate redundant and overlapping bounding boxes.

C. Facial Landmark marking

Following face detection, the subsequent objective is to identify the locations of various facial features such as the corners of the eyes and mouth, and the tip of the nose. Before this, normalization of the face images is imperative to mitigate the effects of camera distance, non-uniform illumination, and varying image resolution. Hence, the face image is resized to a width of 500 pixels and converted to grayscale. Following image normalization, an ensemble of regression trees [11] is employed to estimate the landmark positions on the face from a sparse subset of pixel intensities. This method optimizes the sum of square error loss through gradient boosting learning. Different priors are utilized to identify different facial structures. Through this approach, the boundary points of the eyes, mouth, and the central line of the nose are delineated, with the corresponding numbers of points for each feature outlined in Table I. The facial landmarks, as shown in Fig 2, depict the detected landmarks marked in red for further processing.

D. Feature Extraction

After detecting the facial landmarks, the features are computed as described below.

Eye aspect ratio (EAR): The eye aspect ratio (EAR) is computed from the eye corner points, representing the ratio of the height and width of the eye. This ratio is calculated using the formula $EAR = (\|p2 - p6\| + \|p3 - p5\|) / (2 * \|p1 - p4\|)$, where $\|pi - pj\|$ denotes the Euclidean distance between points marked as i and j in the facial landmark. Consequently, when the eyes are fully open, EAR attains a high value, gradually decreasing as the eyes close. A monotonically decreasing trend in EAR values indicates gradually closing eyes, with values approaching zero indicating complete closure (eye blink). Hence, EAR values serve as indicators of driver drowsiness, as eye blinks occur due to drowsiness.

Mouth opening ratio (MOR): The mouth opening ratio (MOR) serves as a parameter for detecting yawning during drowsiness. Like the EAR, it is calculated as $MOR = (\|p62 - p66\| + \|p63 - p65\|) / (2 * \|p61 - p64\|)$, where $\|pi - pj\|$ denotes the Euclidean distance between points marked as i and j in the facial landmark. This ratio increases rapidly when the mouth opens due to yawning, maintaining a high value for a period (indicating an open mouth), before decreasing rapidly towards zero. As yawning is a characteristic of drowsiness, MOR provides a measure of driver drowsiness.

Head bending, often observed due to drowsiness, involves the driver's head tilting (forward or backward) concerning the vertical axis. By analyzing the head bending angle, one can detect driver drowsiness. Since the projected length of the nose on the camera focal plane correlates with this bending, it serves as a measure of head bending. Under normal conditions, our nose forms an acute angle with respect to the camera's focal plane. This angle increases as the head moves upward and decreases as it moves downward. Therefore, the ratio of the nose length to an average nose length while awake acts as an indicator of head bending and, consequently, drowsiness. Using the facial landmarks, the nose length is computed as NLR = $\|p28 - p25\|$ / average nose length. The average nose length is determined during the setup phase of the experiment, as detailed in the subsequent subsection.

E. Classification

Following the computation of all three features, the subsequent step is to identify drowsiness within the extracted frames. Initially, adaptive thresholding is employed for classification. Subsequently, machine learning algorithms are employed to classify the data. To compute the threshold values for each feature, it is presumed that initially, the driver is in a fully awake state.

This phase is referred to as the setup phase. During this phase, the EAR values for the first three hundred frames (corresponding to 10 seconds at 30 frames per second) are recorded. Among these initial three hundred frames containing a face, the average of the top 150 maximum values is established as the hard threshold for EAR. Higher values are chosen to ensure that no instances of eye closure are overlooked. If the test value falls below this threshold, indicating eye closure (i.e., drowsiness), it is detected. Since the size of the eye can vary from person to person, this initial setup for each individual minimizes this variability. Similarly, to determine the threshold for MOR, since the mouth may not be fully open in the initial frames (setup phase), the threshold is determined experimentally based on observations. If the test value exceeds this threshold, indicating a yawn (i.e., drowsiness), it is detected. The head bending feature is utilized to determine the angle formed by the head with respect to the vertical axis, expressed as the ratio of projected nose lengths. Typically, the NLR ranges from 0.9 to 1.1 for a normal upright head position, fluctuating as the head bends up or down during drowsiness. The average nose length is computed as the mean of the nose lengths during the setup phase, assuming no head bending occurs. Following the computation of threshold values, the system enters the testing phase. Drowsiness is detected by the system if at least one feature exhibits signs of drowsiness in a test frame. To enhance the realism of this thresholding, the decision for each frame depends on the last 75 frames. If at least 70 frames (out of those 75) satisfy the drowsiness conditions for at least one feature, the system indicates drowsiness detection and triggers an alarm.

In order to ensure the adaptability of this thresholding process, an additional single threshold value is calculated. Initially, it is based on the EAR average derived from the top 150 maximum values out of 300 frames during the setup phase. Subsequently, an offset is determined heuristically, and the threshold is obtained by subtracting the offset from the average value. When the EAR falls below this threshold, driver safety is compromised. This EAR threshold value experiences a slight increase with each instance of yawning and head bending, up to a certain limit. As each occurrence of yawning and head bending spans multiple frames, consecutive instances of yawning and head bending are treated as a single occurrence and are included once in the adaptive threshold calculation. In a test frame, if the EAR value falls below this adaptive threshold, drowsiness is detected and an alarm is triggered for the driver. Occasionally, when the head is excessively lowered due to bending, the system may fail to detect the face. In such instances, the system checks the previous three frames, and if head bending was detected in those frames, a drowsiness alarm is activated. Table II provides an illustration of the calculation for determining the adaptive threshold value. Additionally, machine learning algorithms are employed to detect drowsiness. The EAR, MOR, and NLR values, along with the actual drowsiness annotations, are stored for the synthetic test data. Prior to classification, statistical analysis of the features is conducted.

Initially, principal component analysis [12] is applied to reconfigure the feature space into an independent one. Subsequently, the transformed feature values undergo testing via Student's t-test to ascertain their statistical significance for the two classes. Since all three features demonstrate statistical significance at the 5% level, they are employed for classification using Bayesian classifier [12], Fisher's linear discriminant analysis [12], and Support Vector Machine [12].

Algorithm steps:

Detecting driver drowsiness typically involves analyzing various factors such as eye movement, facial expressions, steering behavior, and vehicle dynamics. Here's a high-level algorithm for driver drowsiness detection:

1. Capture Input:
   - Use a camera to capture the driver's face and eyes.
   - Utilize sensors to monitor steering wheel movements, vehicle speed, and lane position.
2. Preprocessing:
   - Apply image processing techniques to detect and track the driver's face and eyes.
   - Extract relevant features such as eye closure, eye movement, head position, and facial expressions.
   - Process sensor data to analyze steering patterns, vehicle speed, and lane deviation.
3. Feature Extraction:
   - Extract features from the preprocessed data. For example:
     - Eye aspect ratio (EAR): Ratio of eye width to eye height to detect blink rate and eye closure.
     - Mouth aspect ratio (MAR): Ratio of mouth width to mouth height to detect yawning.
     - Head pose angles: Detect head nods or tilts.
     - Steering wheel angle changes: Analyze steering behavior.
4. Feature Fusion:
   - Combine features from different sources (e.g., facial features and steering behavior) to improve accuracy and robustness.
5. Classification:
   - Train a machine learning model (e.g., SVM, Random Forest, CNN) using labeled data to classify the driver's state into drowsy or alert.
   - Use the extracted features as input to the classifier.
6. Decision Making:
   - Set a threshold for the classifier output to determine if the driver is drowsy or alert.
   - If the classifier predicts drowsiness, trigger a warning system to alert the driver. This could be auditory alerts, visual warnings, or haptic feedback.
   - Depending on the severity of drowsiness, take appropriate actions like playing loud music, suggesting a break, or even activating automated safety features like lane-keeping assistance.
7. Continuous Monitoring:
   - Continuously monitor the driver's state in real-time.
   - Update the model if necessary to adapt to changing conditions or to improve accuracy.
8. Feedback Loop:
   - Incorporate feedback from the driver's responses to the warnings (e.g., whether they responded by becoming more alert or ignored the warnings) to improve the system's performance over time.
9. Optimization:
   - Optimize the algorithm for real-time performance and efficiency, considering computational resources and response time requirements.
10. Validation and Testing:
   - Validate the algorithm using real-world driving data, including various driving conditions and scenarios.

- Test the system under different lighting conditions, weather conditions, and with different drivers to ensure reliability and effectiveness.

This algorithm provides a general framework for driver drowsiness detection. Depending on the specific requirements and constraints of the application, you may need to customize and fine-tune each step accordingly.

3. Results and discussions

In this paper, we employ a webcam and the SVM (Support Vector Machine) algorithm to monitor the visual behavior of a driver and detect drowsiness. The application utilizes the built-in webcam to capture images of the driver, and then employs the OPENCV SVM algorithm to extract facial features from these images. Subsequently, it checks whether the driver in the image exhibits signs of blinking eyes for a consecutive 20 frames or yawning. If either of these behaviors is detected, the application alerts the driver with drowsiness messages. We utilize a pre-trained SVM model for drowsiness detection and continuously monitor or predict the distance between the eyes and mouth using the Euclidean distance function. If this distance approaches a drowsiness threshold, the application promptly alerts the driver.

To implement above concept we are using following modules

Video Recording: Using this module we will connect application to webcam using OPENCV built-in function called Video Capture.

Frame Extraction: Using this module we will grab frames from webcam and then extract each picture frame by frame and convert image into 2 dimensional array.

Face Detection & Facial Landmark Detection: Using SVM algorithm we will detect faces from images and then extract facial expression from the frames.

Detection: Using this module we will detect eyes and mouth from the face

Calculate: Using this module we will calculate distance with Euclidean Distance formula to check whether given face distance closer to eye blinks or yawning, if eyes blink for 20 frames continuously and mouth open as yawn then it will alert driver.

OpenCV is an artificial intelligence API available in python to perform various operation on images such as image recognition, face detection, and convert images to gray or coloured imagesetc. This API written in C++ languages and then make C++ functions available to call from python using native language programming. Steps involved in face detection using OpenCV.

Face Detection Using OpenCV

This seems complex at first but it is very easy. Let me walk you through the entire process and you will feel the same.

Step 1: Considering our prerequisites, we will require an image, to begin with. Later we need to create a cascade classifier which will eventually give us the features of the face.

Step 2: This step involves making use of OpenCV which will read the image and the features file. So at this point, there are NumPy arrays at the primary data points.

All we need to do is to search for the row and column values of the face NumPy N dimensional array. This is the array with the face rectangle coordinates.

Step 3: This final step involves displaying the image with the rectangular face box.

SVM Description

Machine learning involves predicting and classifying data and to do so we employ various machine learning algorithms according to the dataset. SVM or Support Vector Machine is a linear model for classification and regression problems. It can solve linear and non-linear problems and work well for many practical problems. The idea of SVM is simple: The algorithm creates a line or a hyperplane which separates the data into classes. In machine learning, the radial basis function kernel, or RBF kernel, is a popular kernel function used in various kernelized learning algorithms. In particular, it is commonly used in support vector machine classification. As a simple example, for a classification task with only two features (like the image above), you can think of a hyperplane as a line that linearly separates and classifies a set of data.

Intuitively, the further from the hyperplane our data points lie, the more confident we are that they have been correctly classified. We therefore want our data points to be as far away from the hyperplane as possible, while still being on the correct side of it.

So when new testing data is added, whatever side of the hyperplane it lands will decide the class that we assign to it.

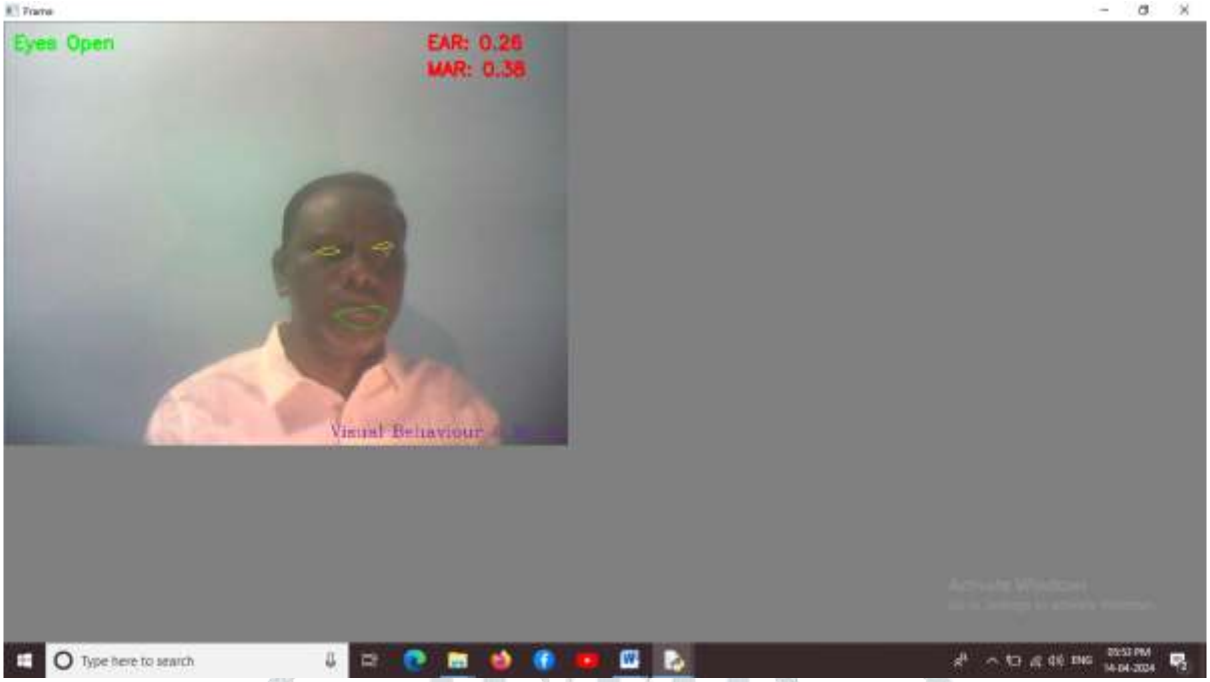
The distance between the hyperplane and the nearest data point from either set is known as the margin. The goal is to choose a hyperplane with the greatest possible margin between the hyperplane and any point within the training set, giving a greater chance of new data being classified correctly.

Description

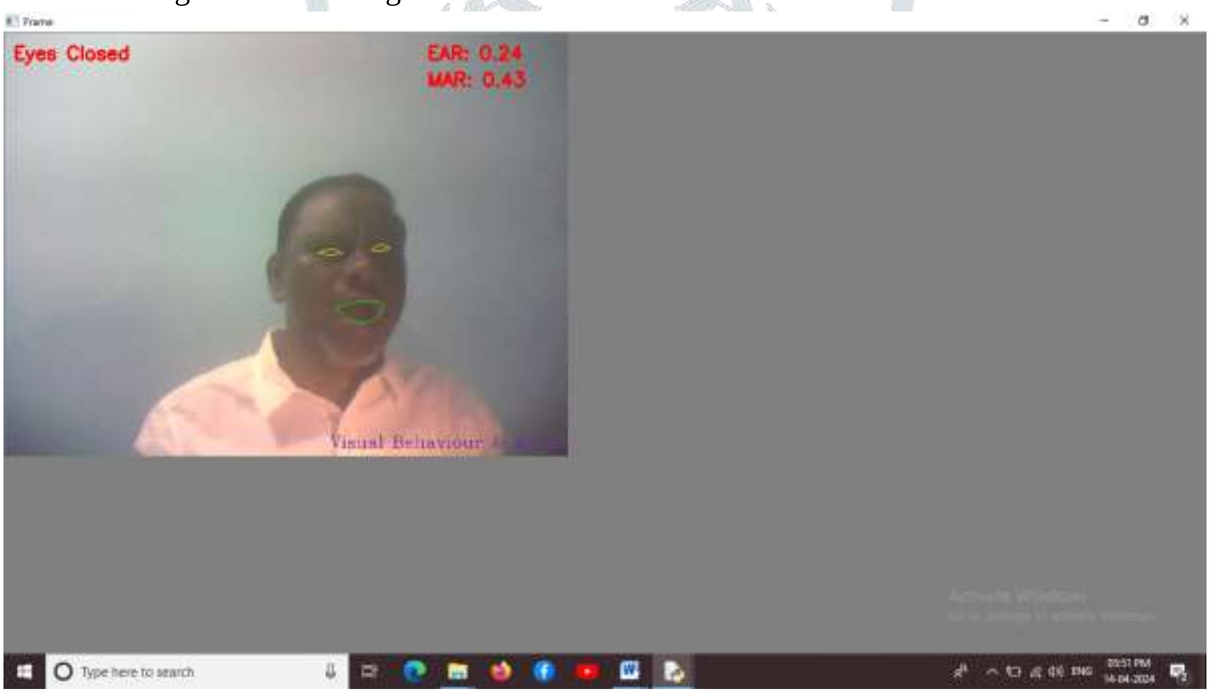
Drowsy driving is one of the major causes of road accidents and death. Hence, detection of driver's fatigue and its indication is an active research area. Most of the conventional methods are either vehicle based, or behavioural based or physiological based. Few methods are intrusive and distract the 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.

Screen shots

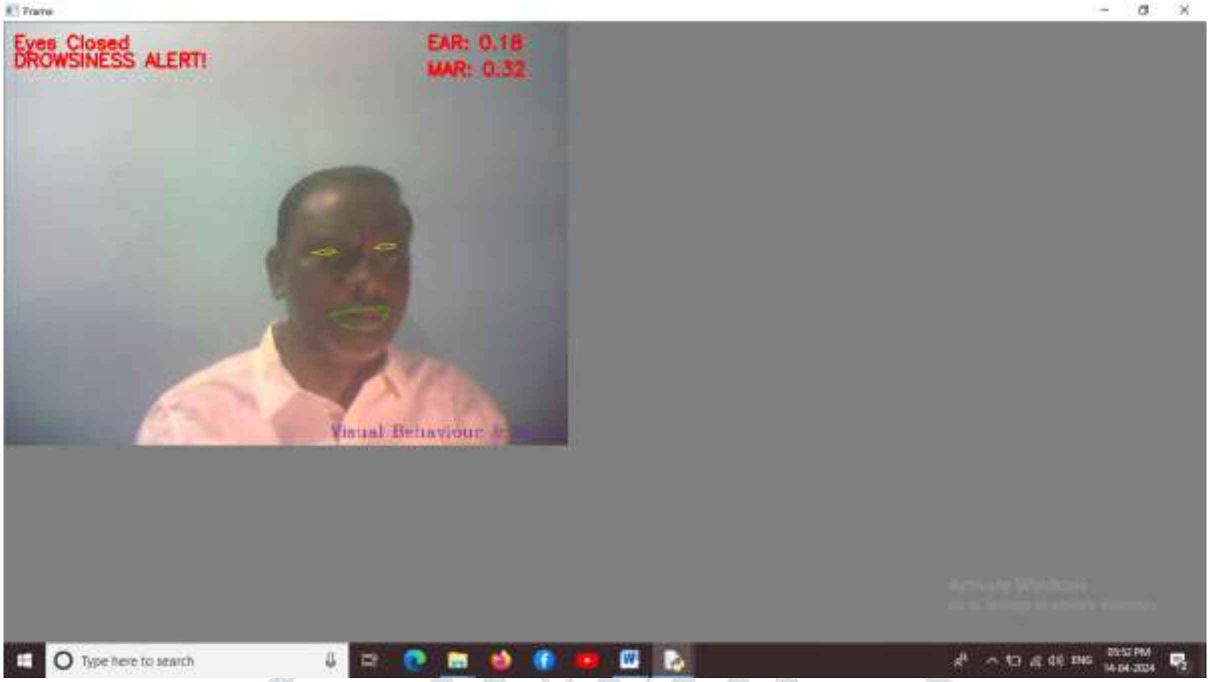
In above screen click on 'Start Behaviour Monitoring Using Webcam' button to connect application with webcam, after clicking button will get below screen with webcam streaming



In above screen we can see web cam stream then application monitor all frames to see person eyes are open or not, if closed then will get below message



Similarly if mouth starts yawn then also will get alert message



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] W.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”,IEEEInternationalConferenceonNetworking,,Sensing and Control, Taipei, Taiwan, March 21-23,2004.

[3] S.Singh,N.P.papanikolopoulos,“MonitoringDriverFatigueusing Facial Analysis Techniques”, IEEE Conference on Intelligent Transportation System, pp314-318.

[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.188193.

[8] A.Sengupta,A.Dasgupta,A.Chaudhuri,A.George,A.Routray,R. Guha;"AMultimodalSystemforAssessingAlertnessLevelsDueto Cognitive Loading", IEEE Trans. on Neural Systems and Rehabilitation Engg., vol. 25 (7), pp 1037-1046,2017.

[9] K. T. Chui, K. F. Tsang, H. R. Chi, B. W. K. Ling, and C. K. Wu, “An accurate ECG based transportation safety drowsiness detection scheme,”118732IEEETransactionsonIndustrialInformatics,vol. 12, no. 4, pp. 14381452, Aug.2016.

[10] N. Dalal and B. Triggs, “Histograms of Oriented Gradients for Human Detection”, IEEE conf. on CVPR,2005.

[11] V.KazemiandJ.Sullivan;"Onemillisecondfacealignmentwithan ensemble of regression trees", IEEE Conf. on Computer Vision and Pattern Recognition, 23-28 June, 2014, Columbus, OH,USA.

[12] Richard O. Duda, Peter E. Hart, David G. Stork, “Pattern Classification”, Wiley studentedition.

[13] Dataset:https://sites.google.com/site/invedrifac/