Driver Drowsiness Monitoring System

Amulya H[1], Neha S Patil[2], Preethi TS[3], Priya M Sangram[4]
[1, 2, 3, 4] Dept of Information Science and Engineering
[5]Asst. Professor, Dept of Information Science and Engineering
[1, 2, 3, 4, 5] East West Institute of Technology, Bengaluru.

Abstract- *Drowsy driving being one of the major causes of road accidents and deaths, detection of driver's fatigue and its indication is an active research area. Most of the methods are either vehicle based, or behavioral based, or physiological based which may sometimes be distractive to the driver and the sensors might get expensive. In this study, real time driver's drowsiness monitoring system is developed with sustainable accuracy. In this system with the help of image processing techniques the driver's face is detected in each frame which is recorded using a webcam. Using detected frames facial landmarks such as eye aspect ratio, mouth opening ratio and nose length ratio are computed and depending on these values the drowsiness is monitored.*

Keywords- visual behavior, machine learning, drowsiness detection, eye aspect ratio, mouth opening ratio and nose length ratio, adaptive threshold.

I. INTRODUCTION

Driver drowsiness monitoring is been used as one of the safe technologies to prevent accidents especially in the night times. The driver drowsiness monitoring system based on visual behavior using machine learning which is used to detect and prevent drowsiness has been known as a major challenge in accident prone system. Most of the times the drivers are distracted because of the objects or the events that occurs which takes away the person's attention away from driving.

There are four main factors due to which the driver gets fatigue. These are sleep, work, time of day and physical condition. However, the driver drowsiness and distraction, might have the same effects, i.e., decreased driving performance, longer reaction time, and an increased risk of accidents. Studies have suggested that around 20% of all road accidents are fatigue- related, up to 50% on certain roads and if the driver is alcoholic. To overcome this problem technologies are used for detecting the drowsiness and preventing the road accidents.

Most of the conventional methods are vehicular based (based on vehicles in which it continuously monitors the steering wheel position, lane position and pressure on acceleration pedal.), behavioral based (in which it continuously monitors blinking frequency of eye, eye closure, yawning and head pose.), physiological based (which checks heart rate and brain activity by ECG(electrocardiogram), EOG(electrooculogram), EMG(electromyogram).). in this proposed system we consider behavioral based by using an acquisition of video from the camera that is placed in front of the driver which performs real-time processing of an incoming video stream in order to compute the driver's level of fatigue and estimate the output which activates the alarm system.

II. LITERATURESURVEY

Earlier studies on driver drowsiness monitoring system has been estimated as major causes of road accidents and deaths. Hence, detection of driver's fatigue and its indication has been an active research area. Previous researches have been using sensors such as an infrared camera for pupil detection or voice to detect the fatigue.

Also, the sensors have been used at the front end of the car as a distance sensor to capture the outside event. These have been developed by using the rule-based algorithm and decision tree algorithm. But these approaches or methods have not been driver adaptable. The disadvantages being, the sensors might not work accordingly, false detection or may also be expensive and not all time affordable.

According to the research eye blinking, head tilting and yawning is the major symptom for fatigue also inattentive vehicular movement in the road under fatigue condition can also be another reason for accidents. Previous papers have considered both the behavioral and the vehicular methods for this reason. Two sensors being used, one to capture the driver's face and another sensor for the lane departure. Few papers have used only the behavioral method considering eye aspect ratio, mouth aspect ratio and head bending ratio using raspberry pi and image processing methods. The system should work regardless of the texture and the color of the face and must also be to handle diverse condition such as change in the light, shadows, reflection. But the results lacked the efficiency in detecting the driver's face in the light changes and fault detections.

III. EXISITINGSYSTEM

The guidelines and criteria which is used to estimate the potential ability of the Driver alertness monitoring technologies and devices relate to its functional characteristics and operational properties. This list lacks the technical details in many cases. It is necessary to be considered a functional requirements specification. addressing these general user acceptance and scientific criteria is important to ensure that any proposed device or technology is qualified for its intended purpose of monitoring, unobtrusively and in real time, driver alertness and thereby, in theory helping to reduce motor vehicle crashes related to driver fatigue.

IV. PROPOSEDSYSTEM

In the developed system, a webcam is used to capture the visual and the driver's face is detected is each such frame that employs image processing techniques with the help of visual behavior and machine learning. It is known to be profound in the field to help develop a low cost, real time driver's drowsiness detection system with sensitivity and specificity. 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. The proposed system is divided into four the following sub systems:

A. Video Capture Unit

The video capture unit used to record the video in real time of the frame containing the driver face through a camera placed on the car dashboard. the face detection unit receives the sampled frame in which video is sampled with some frequency.

B. Face Detection Unit and Features Extraction

Sampled video frame is received from the video capture unit. These images are the RGB image. Low-light image enhancement and noise elimination is performed for the images in very dim light condition the above process is divided into two subtasks. At first, for denoising the image we apply the super pixel based adaptive noising and secondly, for amplifying the image luminance adaptive contrast enhancement is used. We need to denoise the image before contrast enhancement, to eliminate the noise before its amplification through contrast enhancement.

C. Fatigue Detection on Extracted Features

Sequence of eye and mouth image of the driver is extracted. Fatigue detection analysis on various facial features from this extracted dataset we can perform. Eyes (fast blinking or heavy eyes) are included in these facial features. Fatigue detection on these facial features gives combined result. It is used to give the final result which decides whether the driver is in fatigue or alert state.

D. Alert Unit

The format (r, t) is used in which the modelling of the alert unit is done where {r>=0, t>=0} and r is the resultant value and t is the time. The complete model depends on the output value of the fatigue detection unit. the classification of the image, whether they are fatigued or not fatigued is done by the Support Vector Machine (SVM). Their presentation of the classification output is either a +1 or-1and this number is used by the alert unit in the following way. Specifically, the classified output is the input of a running sum that adds the consecutive output values and has a minimum value up to zero. This unit consists of two threshold level, first one is to detect whether there is no or low fatigue level. The difference between low and high fatigue level is to mark the second threshold value. Keeping the specified threshold value in mind that is if it exceeds implying that fatigue is detected and the alert system goes active and takes action. Hence depending on the fatigue level action is taken that is precisely, for low fatigue level, the alarm rings for 10 seconds and after 10 seconds the system again checks the resultant value and performs the alert process accordingly and for high fatigue level, we can use more effective accident preventive measures .Therefore this system can be used to track/estimate the fatigue level of a driver and also detect the sleep onset with a safe margin.

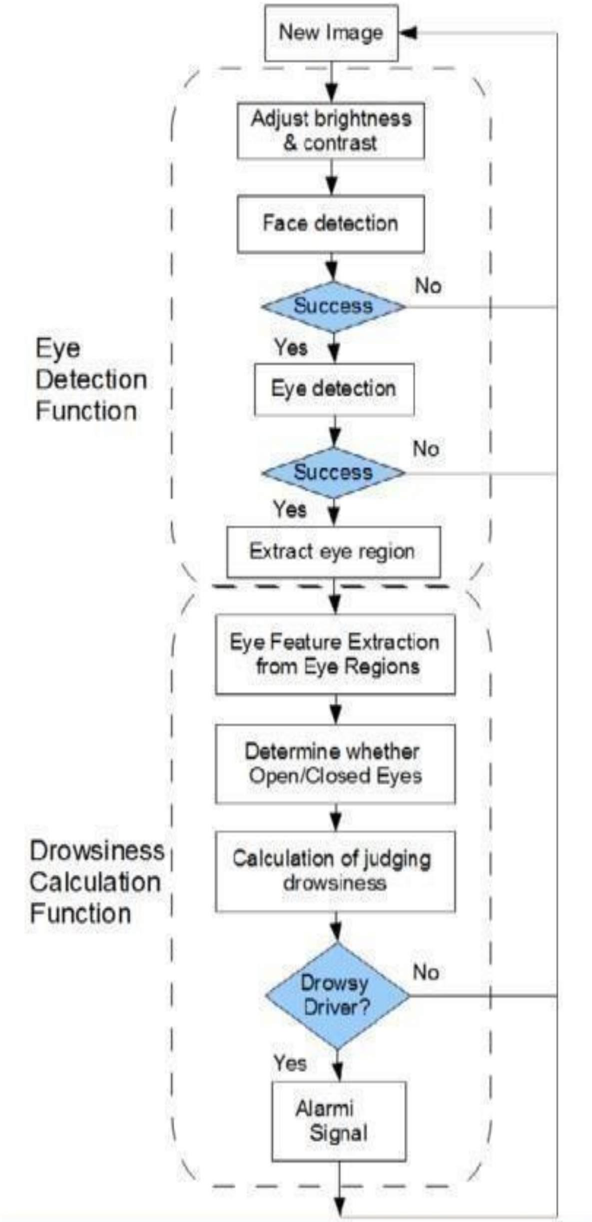


Fig 2. Block Diagram

V. ARCHITECTURE

The system architecture of the proposed system is represented in the figure below. It consists of Video Capture Unit, Face Detection Unit, Head pose estimation system, Facial feature Extraction, closed eye detection, yawning detection, drowsiness detection, Alert Unit. The general flow of drowsiness detection algorithm is given below.

First, we setup a camera that monitors a frame for faces. If a face is detected, we apply facial landmark detection and extract the eye regions. If the eye aspect ratio indicates that the eyes have been closed for a sufficiently long enough amount of time, then sound of an alarm is used to wake the driver up wake up. If the mouth aspect ratio indicates that the yawning for a sufficiently long enough amount of time, then sound an alarm is used to wake up the driver. Estimation of head pose also detects the drowsiness of the driver.

The architecture of the proposed system:

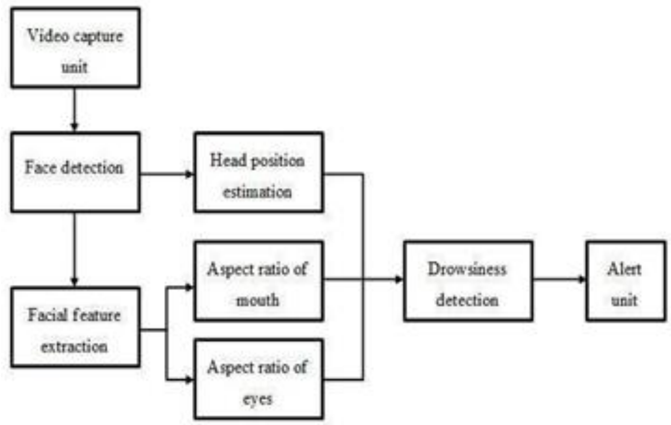


Fig 1. Architecture

VI. IMPLEMENTATION

A. VIDEO CAPTURE

First, we set up a camera in a car that monitors a frame of images. This frame captures the face easily that helps in detection of driver's face and apply facial landmark localization to monitor the eyes, mouth and head length of the person. The web camera is placed on the dashboard of a car to capture the images of the driver's face for video acquisition and an audio alarm system for alerting the fatigued driver and hence in drowsiness detection.

B. FACEDETECTION

The system will start by capturing the video frames one by one. OpenCV provides extensive support for processing live videos. The system detects the face for each frame in the frame image. The system uses Viola-Jones object detector which might be a machine learning approach for visual object detection. This is often achieved by making use of the Haar algorithm for face detection. Haar cascade might be a well-known robust feature-based algorithm which is able to detect the face image efficiently. With the employment of cascade of stages Haar algorithm is prepared to induce obviate the candidates that are non-face. Each of these stages incorporates combination of varied Haar features and each feature is assessed by the employment of Haar feature classifier. The xml file "haarcascade_frontalface_alt.xml" is an inbuilt OpenCV file which can be utilized for searching and detecting the face in individual frames. This file contains variety of features of the face and constructed by employing variety of samples which are positive as well as negative. Post loading the cascaded file, the acquired frame is passed to a position detection function. It then identifies all the possible objects of varied sizes during this particular frame. The sides detector has the potential to detect the objects of all possible sizes. To ease out the tactic, the edge detector is programmed to detect the objects of the specified size i.e. for face region. Next, the output of the sting detector is to store and it's compared with the cascade file to identify the face within the frame. The output of this particular module might be a face detected frame.

C. FACIAL FEATUREDETECTION

Dlib library is utilized to trace facial landmark and extract the regions. Detecting facial landmarks is accordingly a process of two steps: Localizing the face inside the image. The key facial structures are to be detected on the face ROI. In Step 1 Face detection could also be achieved during variety of the ways. OpenCV’s built-in Haar cascades could be used. A pre-trained HOG + Linear SVM object detectors are applied specifically for the face detection task or deep learning-based algorithms can also be used for face localization. In either of the cases, it does not matter on which the actual algorithm accustomed detect the face within the image. Instead, what is significant is that we obtain the face bounding box along some method (i.e., the (x, y)- coordinates of the face inside the image). In Step 2 to detect the face region we are ready to then apply detecting key facial structures within the face region. There are a variety of facial landmark detectors, nevertheless all of those methods eventually attempt to localize and label the following facial regions: The Mouth, Eyebrow on Right, Eyebrow on Left, The Right Eye, The Left Eye, The Nose and The Jaw. The facial landmark detector that has been included within the 'dlib' library is an application of the Ensemble of Regression Trees paper on One Millisecond Face Alignment.

This method starts on using:

1. A trained set of facial landmarks labelled on an image. These images are manually labelled, identifying the precise (x, y)- coordinates of regions adjoining each facial structure.
2. Priors, the probability of distance between the pairs of input pixels.

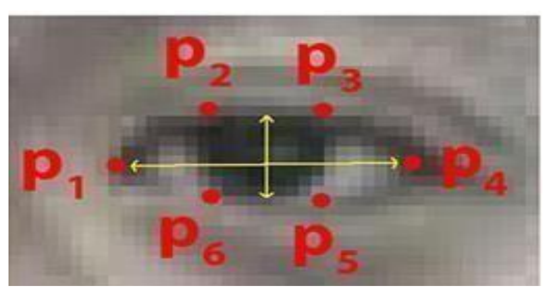
Given the trained datasets, a band of regression trees are instructed so on to reckon the facial landmark positions directly with the pixel intensities assistance themselves. The end result is the facial landmarks detected using the facial landmark detector in real-time with top of the range predictions. The trained facial landmark detector inside the 'dlib' library is beneficial to estimate things of 68 (x, y)-coordinates which map the face facial structures. The 68 coordinates' indexes can certainly be visualized on the image below: These annotations are a part of the 68-point iBUG 300-W dataset in which the dlib facial indication predictor are instructed on. It is important to note that kind of facial landmark detectors exist, which contains the 194-point model which is able to be certainly trained using the HELEN dataset. Nevertheless, of which dataset is utilized, the identical 'dlib' frame work could also be maneuvered to teach as shape predictor on the training data input — this is often useful if the facial landmark detector so custom shape predictors of your own are to be trained. We are ready to see that facial regions are often accessed via simple Python indexing (assuming zero indexing with Python since the image above is one-indexed):

- The mouth is usually accessed through points [48,68].
- The proper eyebrow along the points [17,22].
- The left eyebrow along the points [22,27].
- The proper eye using [36,42].
- The left eye with [42,48].
- The nose using [27,35].
- and so, the jaw via [0,17].

The facial landmark detection could also be applied to localize the important regions of the face, which contains eyes, eyebrows, nose, ears, and mouth. It is shown within the figure given below.

(a) Aspect Ratio of Eyes

This also implies that we can extract specific facial structures by knowing the indexes of the particular face parts. In terms of blink detection, we are only interested in two sets of facial structures eyes. Each of the eye is represented by 6 (x, y)-coordinates, starting from the left-corner of the eye (as if you were looking at the person), and then working clockwise around the remaining part of the region. For every video frame, the eye landmarks are detected according to the coordinates. The eye aspect ratio (EAR) is computed between height and width of the eye.



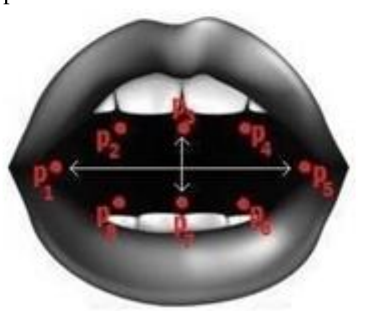
$$\text{EAR} = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|}$$

Fig 6.1 2D Landmark of Eye

Where p1-p6 are the 2D landmark locations. The Eye Aspect Ratio is mostly constant when an eye is open and is zero when the eye is closed. It is partially person and head pose insensitive. among individuals Aspect ratio of the open eye has as mall variance and it is fully invariant to a uniform scaling of the image and in-plane rotation of the face, the EAR of both eyes is averaged, since eye blinking is performed by both eyes synchronously. The numerator in this equation computes the distance between the vertical eye landmarks and the denominator computes the distance between horizontal eye landmarks, weighting the denominator appropriately since there is only two sets of vertical points and one set of horizontal points. the eye aspect ratio is constant, then rapidly drops close to zero, then again increases, indicating a single blink has occurred.

(b) Aspect Ratio of Mouth

This also implies that the specific facial structures can be extracted by knowing the indexes of the particular face features. In terms of yawning detection, we are only interested in facial structures of mouth. Mouth is represented by 8 (x, y)-coordinates, starting at the left-corner of the inner lips and then working clockwise around the remainder of the region: For every video frame, the mouth landmarks are detected. The mouth aspect ratio (MAR) between height and width of the mouth is computed.



$$\text{MAR} = \frac{\|p_3 - p_7\|}{\|p_1 - p_5\|}$$

Fig 6.2 2D Landmark of Mouth

Where p1- p8 are the 2D landmark locations, The Mouth Aspect Ratio is mostly constant when a mouth is open and is varied while opening of mouth. It is partially person and head pose insensitive. Aspect ratio of the closed mouth has a small variance among individuals and it is fully invariant to a uniform scaling of the image and in- plane rotation of the face. The numerator of this equation computes the distance between the vertical eye landmarks while the denominator computes the distance between horizontal mouth landmarks, weighting the denominator.

(c) Head Pose Estimation

The next step is to estimate driver's head pose/head bend. Driver's face image from head detection and tracking has been adjusted for identifying the head tilting. A method to estimate the head pose is tracking the head using the motion of the head or the motion of features on the face. A well-known method to compute the motion is to use an optical flow algorithm. A method for tracking a rigid head on video uses optical flow. They used a technique called motion regularization for tracking process together with an ellipsoid model as a base. The main method is to find the rigid motion of the head model that accounts best for the optical flow. the first step is to calculate the optical flow of each point and then use a gradient descent technique to find the accurate motion of the head. The tracker in their test is stable over a large number of frames and also when using sequences with a low frame rate and the noisy images. The proposed an optical flow-based method to constraint the motions of a deform able model. They combined optical flow information and edge information to prevent drifting caused by the optical flow.

(d) Alert Unit

The modelling of the alert unit is done when the driver is drowsy. The representation of the classification output is either a 1 or 0 and this number is used by the alert unit. Whenever the resultant value goes above a specified threshold, it implies that fatigue is detected and the alert system goes active and takes action according to the detected fatigue level.

Table 1. Test conditions

EAR (eye aspect ratio)	Eye threshold= 0.3 Eye consecutive frames= 48	If EAR > 0.3 If EAR < 0.3	Drowsiness not detected. Drowsiness detected.
MOR (mouth aspect ratio)	Mouth threshold= 0.6 Mouth consecutive frames= 30	If MOR > 0.6 If MOR < 0.6	Drowsiness detected. Drowsiness not detected.
Head length ratio	Max head movement= 20 Movement threshold = 50 Gesture threshold = 175	If x_movement > gesture threshold if y_movement > gesture_threshold	Drowsiness not detected. Drowsiness detected.

VII. RESULTS

Eye aspect ratio is used as the indicator of drowsiness in this eye closing part and Mouth aspect ratio is used as the indicator of drowsiness in this yawning part. We extract the video data to its frames and the frames are input to the part eye and mouth region. Calculating and comparing the Eye aspect ratio and Mouth aspect ratio with the threshold value and detecting the drowsiness of driver. If the Driver is drowsy the alarm will sound and EAR and MAR values and the alert will show in display. Also analyzing the head pose of the driver will detect the driver drowsiness.

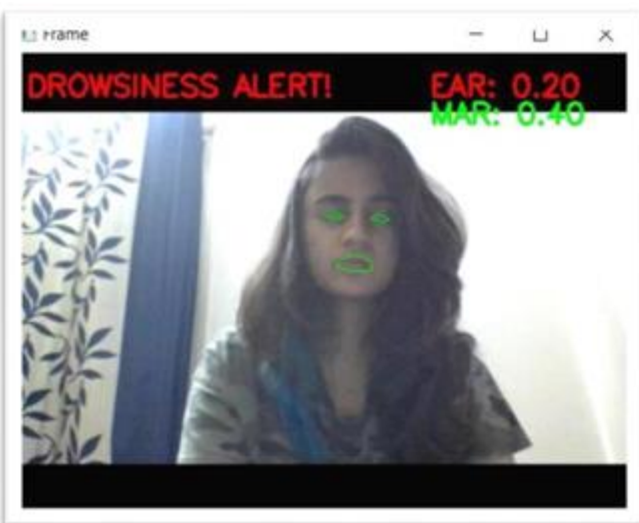


Fig 7.1 Detecting Drowsiness through eyes

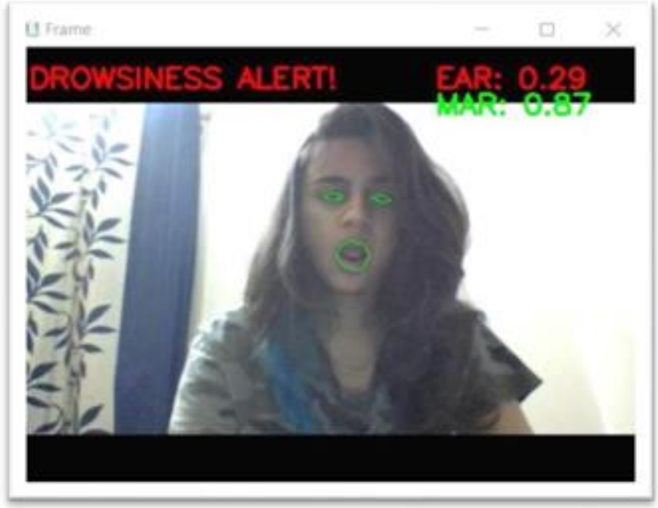


Fig 7.2 Detecting Drowsiness through mouth

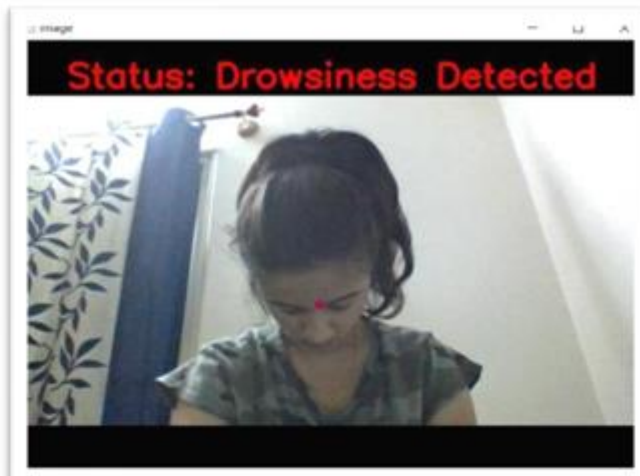


Fig 7.3 Detecting Drowsiness through head length

Fig 7.4 Detecting Drowsiness also with spectacles

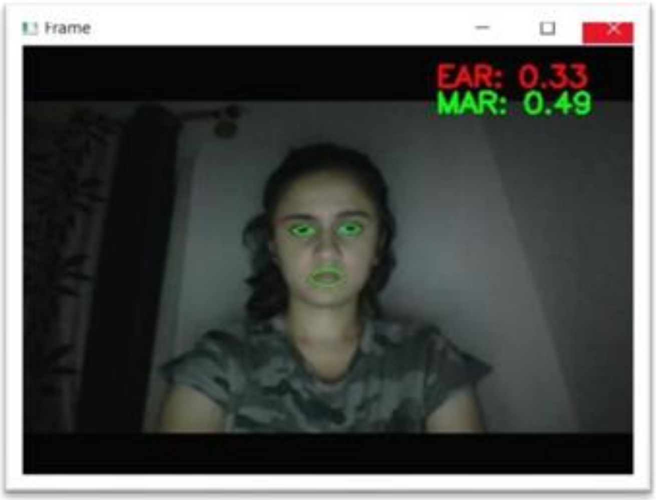


Fig 7.5 Detecting face in low light condition

VIII. CONCLUSION

This system works only on the required part of face image i.e. eyes and mouth, rejects the rest. This step decreases the unnecessary features in the feature set. Eye and mouth detection are less accurate compared to the face detection, and hence the use of face detection to get the eye and mouth image part of the driver. Estimating the head position of the driver will detect the drowsiness of driver. It makes the system optimized in the context of time and accuracy. The model which has divided alert unit into three units is an efficient way to alert the driver. It works in such a way that the driver is not subjected to sudden attacks, which may lead to an accident.

IX. ACKNOWLEDGEMENT

Any achievement, be it scholastic or otherwise does not depend solely on the individual efforts but on the guidance, encouragement and cooperation of intellectuals, elders and friends. A number of personalities, in their own capacities have helped us in carrying out this project.

We would like to take this opportunity to thank them all. First and foremost we would like to thank Mr.Sanju D J, Asst. Professor, EWIT, for his moral support and valuable suggestions and expert advice. towards completing our project work. and constantly guiding me to organize the project in a systematic manner.

We thank our Parents, and all the Faculty members of Department of Information science & Engineering for their constant support and encouragement. Last, but not the least, we would like to thank our peers and friends who provided us with valuable suggestions to improve our project.

REFERENCES

[1] Stan, Ovidiu, Liviu Miclea, and Ana Centea. "EyeGaze Tracking Method Driven by Raspberry PI Applicable in Automotive Traffic Safety." Artificial Intelligence, Modelling and Simulation (AIMS), 2014 2nd International Conference on. IEEE, 2014.

[2] SS, Kulkarni, and A. D. Harale. " application of raspberry pi based embedded system for real time protection against road accidents due to driver’s drowsiness and/or drunk and drive cases." international journal of engineering sciences &research technology, 5(9):September,2016.

[3] Anjali, K. U., et al. "Real-time nonintrusive monitoring and Detection of eye blinking in view of accident prevention due to drowsiness." Circuit, Power and Computing Technologies (ICCPCT), 2016 International Conference on. IEEE, 2016. Stan, Ovidiu, Liviu Miclea, and Ana Centea. "Eye Gaze Tracking Method Driven by Raspberry PI Applicable in Automotive Traffic Safety." Artificial Intelligence, Modelling and Simulation (AIMS), 2014 2nd International Conference on IEEE,2014.

[4] Turan, Garima, and Sheifali Gupta. "Road Accidents PreREvention system using Driver’s Drowsiness Detection." International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2(2013).

[5] Ruian Liu,et.a;, "Design of face detection and tracking system," Image and Signal Processing (CISP), 2010 3rd International Congress on , vol.4, no., pp.1840,1844, 16-18 Oct.2010.

[6] Goal, P, et.al., "Hybrid Approach of Haar Cascade Classifiers and Geometrical Properties of Facial Features Applied to Illumination Invariant Gender Classification System," Computing Sciences (ICCS), 2012 International Conference on, vol., no., pp.132,136, 14-15 Sept.2012.

[7] Parris, J., et.al, "Face and eye detection on hard datasets," Biometrics (IJCB), 2011 International Joint Conference on, vol., no., pp.1,10, 11-13Oct.2011.

[8] Qiang Ji, Zhiwei Zhu and Peilin Lan ―IEEE transactions on Vehicular Technology Real Time Nonintrusive Monitoring and prediction of driver fatigue. 53, no. 4, July2004.

[9] Wei-niin Huang & Robert Mariani, ―Face Detection and precise Eyes Location ―, Proceeding of the International Conference on Patten Recognition (ICPP‖OO), Vol.4,2000 Gonzalez, Rafel C. and Woods, Richard E. ―Digital Image Processing‖, Prentice Hall: Upper Saddle River, N.J.,2002.