Real Time Drowsiness Detection System

Amitabh Das[1] Aditya Singh[2] Ashutosh Goyal[3] Aishwarya Gupta[4]
[1,2,3,4]Department of Computer Science and Engineering
[1,2,3,4]ABES Institute of Technology, Ghaziabad-201009, Uttar Pradesh, India

Abstract— With the advancing technology in digital computing world, it has become a lot easier to detect the very minute details of everything. This work is intended to detect the state of drowsiness in drivers in order to minimize the accidents and to improve road safety. The implementation of real time drowsiness detection is done by locating driver's facial gestures along with eye parameters. Our effort is to detect the eye blink of the driver. If the eyes of the driver are in closing state for more than the threshold time, it will show a drowsiness alert and arouse the alarm. The coding is done in python language and Open CV is added for detecting facial features. We aim to develop a real-time algorithm to detect blinking of eyes from a video stream captured from a standard camera. The precise detection of landmarks was enough to estimate the opening of eyes. We propose an algorithm which estimates the landmark positions by extracting a single entity – eye aspect ratio (EAR) – and characterize the opening of eye in each frame. Finally, a Support Vector Machine classifier detects eye blinks as a pattern of EAR values in a short temporal window. The simple algorithm outperforms the state-of-the-art results on two standard datasets.

General Terms: Facial Recognition, Landmark Localization, SVM Classifier
Keywords: Drowsiness, Detection, OpenCV

I. INTRODUCTION

Drowsy means sleepy and having low energy. Drowsiness is the state of feeling sleepiness. It means being unable to keep your eyes open, or feeling tired. Drowsiness is also called excess sleepiness. It is usually accompanied by a lethargic energy and lack of mental alertness. It may be due to some long driving hours by a tired driver. This causes fatigue, so there is a need to detect this in order to avoid accidents. It is not easy to estimate the exact amount of sleep related accidents but research presents that driver fatigue may be a contributing reason in up to 20% in road accidents. These types of accidents are about 50% more expected to result in death or serious hurt. They happen mainly at higher speed impacts and the driver who has fallen asleep cannot apply brakes. It reduces alertness and concentration so that the capacity to perform attention-based activities i.e. driving is compromised. It also affects the speed at which information is processed along with the quality of decision-making. Crashes caused by tired drivers are most likely to happen on long journeys on simultaneous roads, such as motorways, especially after eating or taking an alcoholic drink, after having less sleep than normal, after drinking alcohol, if the driver takes medicines that cause drowsiness and after long working hours or on journeys to home after long shifts, especially night shifts. Tiredness and fatigue often affect a person's driving ability before he/she even notices that he/she is getting tired.

II. EXISTING SYSTEM

By using a non-intrusive OpenCV concept, drowsiness of the driver detected system is developed. Existing systems require a camera installed in front of driver which points straight towards the face of the driver and monitors the driver's eyes in order to identify the drowsiness. For large vehicle such as heavy trucks and buses, this arrangement is not pertinent. Bus has a large front glass window to have a broad view for safe driving. If the camera is placed on the frame which is just about the window, then the camera is unable to detain the anterior view of the face of the driver correctly. If the eyes are closed for five successive frames the system concludes that the driver is declining slumbering and issues a warning signal. Hence existing system is not applicable for large vehicles. In order to conquer the problem of existing system, new detection system is developed in this project work.

A. Proposed System:

This study is intended to develop a closed and opened eye detection algorithm aimed respectively at state of drowsiness and fatigue assessment. The present study is an improved work of the implemented methods tested on real captured video stream using inbuilt webcam. The proposed system performs some steps before determining driver's drowsiness level. Firstly, the face is captured from the video frames. Then, the localization of the eyes and the mouth is performed. Finally, we apply the proposed methods for detecting fatigue and drowsiness.

B. Description of Features

In each video frame, the eye landmarks are detected. The eye aspect ratio (EAR) between height and width of the eye is computed.

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

where p1, . . ., p6 are the 2D landmark locations. The eye aspect ratio is nearly neutral when an eye is open and is about to approach zero while closing of eye. Aspect ratio of the open eye has a small divergence among individuals and it is fully different to a uniform scaling of the image and in-plane rotation of the face. Since eye blinking is performed by both eyes simultaneously, the EAR of both eyes is averaged.

C. Experiments

We performed two types of experiments: the accuracy measurement of the landmark detectors, see Sec. 2.3.1, and the experiments the performance evaluators of the eye blink detection algorithm, see Sec 2.3.2.

1) Accuracy of Landmark Detectors

To check the accuracy of tested landmark detectors, we have used the 300-VW dataset [19]. It is a dataset containing 50 videos where each frame is associated with a precise annotation of facial landmarks. The purpose of the following tests is to demonstrate that our landmark detectors are robust and precise in detecting eyes, i.e. the eye corners and outline of the eyelids. Therefore, we have prepared a dataset, a subset of the 300-VW, containing sample images with both open and closed eyes. More precisely, having the ground landmark annotation, we have sorted the frames for each subject by the eye aspect ratio and took 10 frames of the highest ratio (eyes wide open), 10 frames of the lowest ratio (mostly eyes tightly closed) and 10 frames sampled randomly. Thus, we collected 1500 images. Moreover, all the images were later collected (successively 10 times by factor 0.75) in order to check accuracy of tested detectors on small face images.

Two state-of-the-art landmark detectors were tested: Chehra [1] and Intraface [16]. Both run in real-time. Samples from the dataset are shown in Fig. 3. Notice that faces are not always in front of the camera, the expression is not always natural, people are often emotionally speaking or smiling, etc. Sometimes people wear glasses, hair may occasionally partially cover one of the eyes. Both detectors perform generally well, but the Intraface is more robust to very small face images, sometimes at impressive extent.

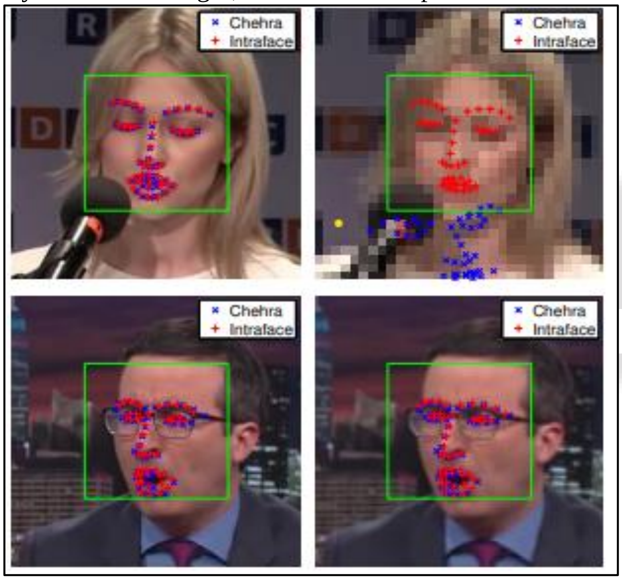


Figure 3: Example images from the 300-VW dataset with landmarks obtained by Chehra [1] and Intraface [16]. Original images (left) with inter-ocular distance (IOD) equal to 63 (top) and 53 (bottom) pixels. Images subsampled (right) to IOD equal to 6.3 (top) and 17 (bottom).

Quantitatively, the accuracy of the landmark detection for a face image is measured by the average relative landmark localization error, defined usually as

$$\epsilon = \frac{100}{\kappa N}\sum_{i=1}^{N} ||x_i - \hat{x}_i||_2,$$

where xi is the ground location of landmark i in the image, xˆi is an estimated landmark location by a detector, N is a number of landmarks and normalization factor κ is the inter-ocular distance (IOD), i.e. Euclidean distance between eye centers in the image. First, we calculated a cumulative graph of the average landmark localization error, see Fig. 4, for a complete set of 49 landmarks and also for a subset of 12 landmarks of the eyes only, since these landmarks are used in the proposed eye blink detector. The results are calculated for all the original images that have average IOD around 80 px, and also for all "small" face images (including subsampled ones) having IOD ≤ 50 px.

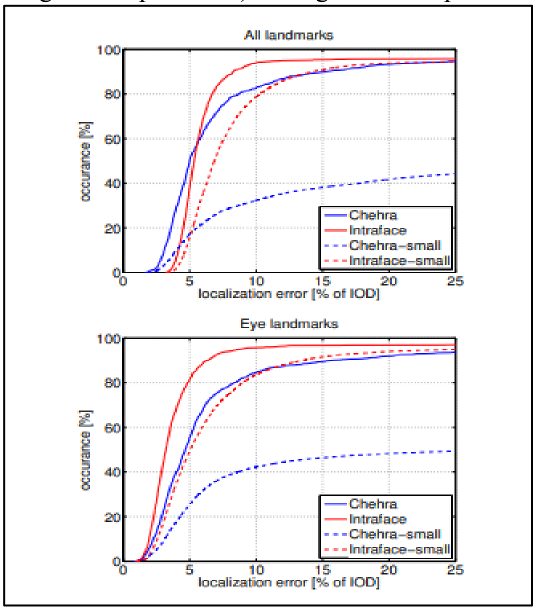


Figure 4: Cumulative histogram of average localization error of all 49 landmarks (top) and 12 landmarks of the eyes (bottom). The histograms are computed for original resolution images (solid lines) and a subset of small images (IOD ≤ 50 px).

The Interface is always more precise than Chehra. As already mentioned, the Intraface is much more robust to small images than Chehra. This behavior is further observed in the following experiment. Taking a set of all 15k images, we measured a mean localization error µ as a function of a face image resolution determined by the IOD. More precisely, µ = 1 |S| P j∈S j, i.e. average error over set of face images S having the IOD in a given range. Results are shown in Fig. 5. Plots have error bars of standard deviation. It is seen that Chehra fails quickly for images with IOD < 20 px. For larger faces, the mean error is comparable, although slightly better for Intraface for the eye landmarks. The last test is directly related to the eye blink detector. We measured accuracy of EAR as a function of the IOD. Mean EAR error is defined as a mean absolute difference between the true and the estimated EAR. We computed the plots for two subsets: closed/closing (average true ratio 0.05 ± 0.05) and open eyes (average true ratio 0.4 ± 0.1). The error is higher for closed eyes. The reason is probably that both detectors are more likely to output open eyes in case of a failure. It is seen that ratio error for IOD < 20 px causes a major confusion between open/close eye states for Chehra, nevertheless for larger faces the ratio is estimated precisely enough to ensure a reliable eye blink detection.

2) Eye Blink Detector Evaluation

We evaluate on two standard databases with ground representations of blinks. The first one is ZJU [11] consisting of 80 short videos of 20 subjects. Each subject has 4 videos: 2 with and 2 without glasses, 3 videos are frontal and 1 is an upward view. The 30fps videos are of size 320 × 240 px. An average video length is 136 frames and contains about 3.6 blinks in average. An average IOD is 57.4 pixels. In this database, subjects do not perform any noticeable facial expressions. They look straight into the camera at close distance, almost not moving, neither smiling nor speaking. A ground blink is defined by its beginning frame, peak frame and ending frame. The second database Eyeblink8 [8] is more challenging. It consists of 8 long videos of 4 subjects that are smiling, rotating head naturally, covering face with hands, yawning, drinking and looking down probably on a keyboard. These videos have length from 5k to 11k frames, also 30fps, with a resolution 640 × 480 pixels and an average IOD 62.9 pixels. They contain about 50 blinks on average per video. Each frame belonging to a blink is annotated by half-open or close state of the eyes.

III. SYSTEM MODULES:

A. *Face Detection*

We extracted the face from the frame to reduce the search region and therefore reduce the computational cost required for the subsequent step. We have used an existing method, based on SVM technique.

B. *Face and Eye Detector Single Image*

This module detects the image captured by webcam and uses it to convert and detect face and eyes of the user in a single image in order to improve precision and accuracy of the detected image.

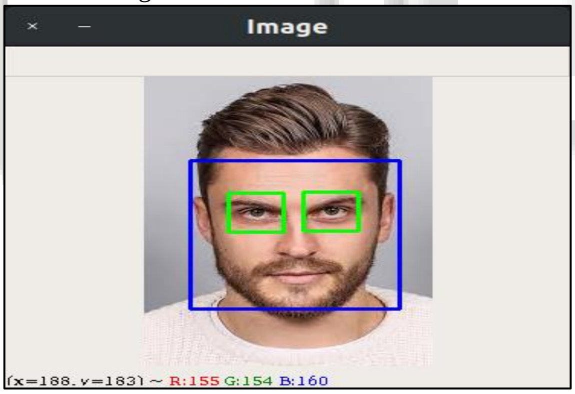


Fig. 1: face and eye detection in single image

C. *Face and Eye Detector Webcam Video*

This module detects the user's face and eye from the webcam feed which is continuously being monitored by the default webcam of system or an external device.

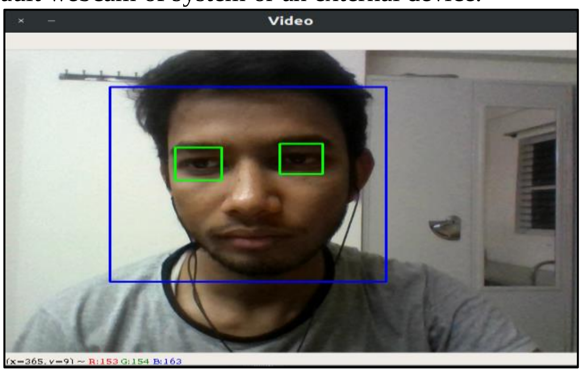


Fig. 2: face and eye detection from webcam feed

D. *Drowsiness Detector*

This script detects if person is drowsy or not using webcam video feed.

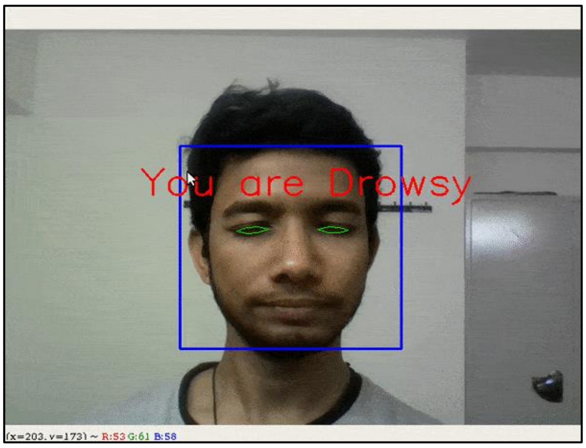


Fig. 3: drowsiness detection using webcam feed

IV. FACIAL PARAMETERIZATION:

Based on various input elements to the face of an ordinary human, the face can be diverted into co-ordinates in order to detect the specified eye positions. The facial landmark detector implemented inside dlib produces 68 (x, y)-coordinates that map to specific facial structures. These 68 point mappings were obtained by training a shape predictor on the labeled iBUG 300-W dataset.

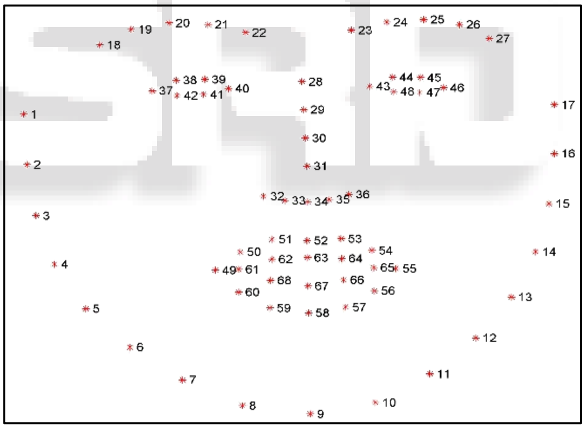


Fig. 4: Visualizing each of the 68 facial coordinate points from the iBUG 300-W dataset (higher resolution).

Examining the image, we found that facial regions can be accessed via simple Python indexing (assuming zero-indexing with Python since the image above is one-indexed):

The right eyebrow through points [17, 22].
The left eyebrow through points [22, 27].
The right eye using [36, 42].
The left eye with [42, 48].
These mappings are encoded inside the FACIAL_LANDMARKS_IDXS dictionary inside face_utils of the imutils library.

V. EYE ASPECT RATIO

We have defined the eye_aspect_ratio function which is used to compute the ratio of distances between the vertical eye landmarks and the distances between the horizontal eye landmarks. The return value of the eye aspect ratio will be approximately constant when the eye is open. The value will then rapidly decrease towards zero during a blink. If the eye is closed, the eye aspect ratio will again remain approximately constant, but will be much smaller than the ratio when the eye is open.

To visualize this, consider the following figure from Soukupová and Čech’s 2016 paper, Real-Time Eye Blink Detection using Facial Landmarks.

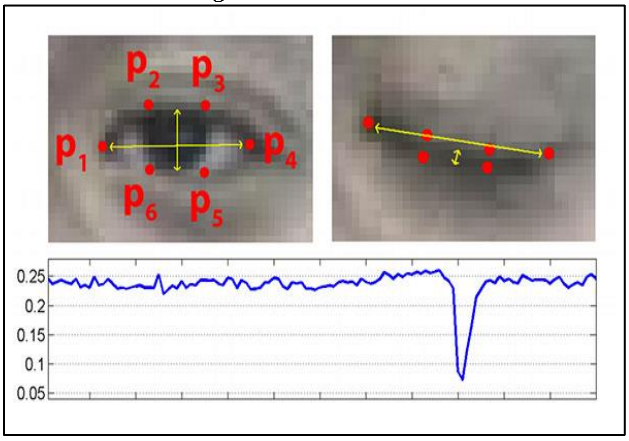


Fig. 5: Top-left: A visualization of eye landmarks when then the eye is open. Top-right: Eye landmarks when the eye is closed. Bottom: Plotting the eye aspect ratio over time. The dip in the eye aspect ratio indicates a blink (Figure 1 of Soukupová and Čech).

On the top-left we have an eye that is fully open with the eye facial landmarks plotted. Then on the top-right we have an eye that is closed. The bottom then plots the eye aspect ratio over time.

As we can see, the eye aspect ratio is constant (indicating the eye is open), then rapidly drops to zero, then increases again, indicating a blink has taken place.

VI. Modulation

A. Face Detection

In this stage, we detect the region containing the face of the driver. The eye_aspect_ratio algorithm is for detection of face in every frame. By face detection we mean locating the face in a frame or in other words finding location of facial characters through a type of technology with the use of webcam installed in the computer.

B. Recognition of Face Region

To detect the region of face after minimizing the background extra portion from the image, labelling method is used. In the labelling method, components are connected in 2-D binary image.

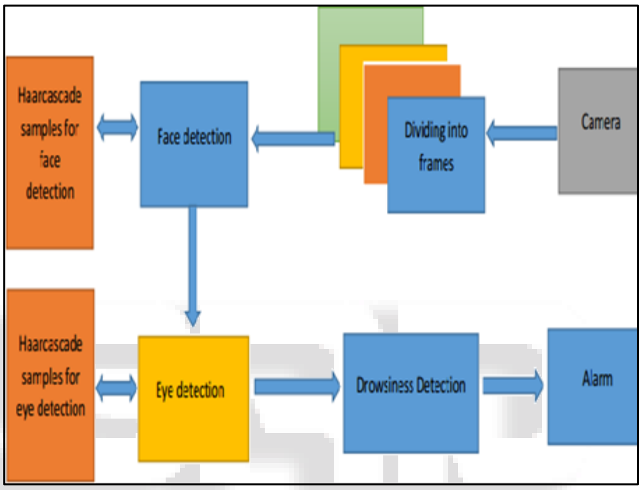
C. Eye Detection

In our method, eye is the decision parameter for finding the state of driver. Though detection of eye may be easier to locate, but it’s really quite complicated. At this point it performs the detection of eye in the required particular region with the use of detection of several features. Generally, Eigen approach is used for this process. It is a time taking process. When eye detection is done then the result is matched with the reference or threshold value for deciding the state of the driver.

D. The Comparison Table

Existing System	Our Added Feature	Accuracy
Facial Detection (up to 78% accuracy).	Eye blink detection with extended facial detection	85% precise system developed.
Eye detection (Showed approximately 80% accurate results).	Addition of Eye Aspect Ratio Formula for calculating accurate landmark points.	Better functioning and 85% accurate.
Detection Technique	Non – Intrusive system and real time detection.	More Reliable and easy to use.

E. Workflow



VII. Conclusion and Scope

We developed a non-intrusive system to localize the eyes and monitor fatigue. We obtained information about the head and eyes position through various self-developed image processing algorithms. During the monitoring, the system was able to decide whether the eyes are opened or closed. When the eyes were closed for two seconds, a warning signal is issued. In addition, during monitoring, the system was able to automatically detect any eye localizing error that might have occurred. In case of this type of error, the system is able to recover and properly localize the eyes. The proposed system was tested on the real driver images. The video image [480 x 640 pixels] of 75 different test persons has been recorded during several day, night and complex background at different places.

In future, this prototype can be extended to raise alarm before sleeping by calculating the heart beat measure without physical disturbance i.e., non-intrusive method using modified ECG methods. Usually in ECG method key points of body (For example chest, head, wrist etc.,) are sticked with wire. In the extended method, sticking wire may be avoided. This will lead us to a way to find out the optimum level of drowsiness. Further, this prototype will be extended to monitor the reflect ray from eye using nano camera. If the reflection ray is absent, then eye is closed otherwise eye is opened. We believe that this will create a better opportunity to detect drowsiness.

References

[1] Baskan S. Bulut M. and Atalay V., "Projection based method for segmentation of human face and its evaluation", Pattern Recognition, Vol. 23, pp. 623–1629, 2002.
[2] Belhumeur P. and Kriegman D., "What is the set of images of an object under all possible lighting conditions", Int. J. of Computer Vision, Vol. 28, pp. 245- 260, 1998.
[3] Campadelli P. and Lanzarotti R., "Localization of facial features and fiducial points" in Proc. Int. Conf. Visualization, Imaging and Image Processing, pp. 491–495, 2002.
[4] D'Orazio T., Leo M., Cicirelli G. and Distante A., "An algorithm for real time eye detection in face images," in Proc. 17th Int. Conf. on Pattern Recognition, Vol. 3, pp. 278–281, 2004.
[5] Fletcher L., Apostoloff N., Chen J. and Zelinsky A., "Computer vision for vehicle monitoring and control", Proceedings of the 2001 Australian Conference on Robotics and Automation, Sydney, pp. 67-72, 14-15 November 2001.
[6] Fletcher L., Apostoloff N., Petersson L. and Zelinsky A., "Vision in and out of Vehicles", In Broggi, A. (Ed.), Intelligent Transportation Systems. IEEE Computer Society, pp. 12-17, 2003.
[7] Fletcher L., Petersson L. and Zelinsky A., "Driver assistance systems based on vision in and out of Vehicles", Proceedings of the IEEE Intelligent Vehicles Symposium (IV2003), Columbus, Ohio, USA, 2003.
[8] Gu H., Ji Q. and Zhu Z., "Active facial tracking for fatigue detection", Proceedings of the Sixth IEEE Workshop on Applications of Computer Vision, Orlando, 3-4, pp. 137-142, December 2002.
[9] Guo Can Feng and Pong C. Yuen, "Multi-cues eye detection on gray intensity Image", Pattern Recognition. Vol. 34, pp. 1033-1046, 2001.
[10] Haro A., Flickner M. and Essa I., "Detecting and tracking eyes by using their Physiological properties, dynamics and appearance", in Proc IEEE Conf. Computer Vision and Pattern Recognition", Vol. 1, pp. 163–168, 2000.
[11] Ito T., Mita S., Kozuka K., Nakano T., and Yamamoto S., "Driver blinks measurement by the motion picture processing and its application to drowsiness detection", Proceedings of the 5th IEEE International Conference on Intelligent Transportation Systems, Singapore, pp. 168 – 173, 3-6 September 2002.
[12] Jeffrey Huang and Harry Wechsler, "Eye Detection Using Optimal Wavelet Packets and Radial Basis Functions", International Journal of Pattern Recognition and Artificia1 Intelligence, Vol. 13, No 7, 1999.
[13] Kawaguchi T. and Rizon M., "Iris detection using intensity and edge information", Pattern Recognition, Vol.36, pp. 549–562, 2003.
[14] Lal S.K.L., Craig A., Boord P., Kirkup L. and Nguyen H., "Development of an algorithm for an EEG-based driver fatigue countermeasure", Journal of Safety Research, Vol. 34, pp. 321-328, 2003.
[15] Moorthi M., Arthanari M. and Sivakumar M., "Preprocessing of Video Image with Unconstrained Background for Drowsy Driver Detection", International Journal of Computer Science and Network Security (IJCSNS), USA, ISSN 1947-5500 Vol. 8 No.2, pp. 145-151, May 2010.

