Driver Drowsiness Monitoring System using Visual Behaviour and Machine Learning

Ambika Sahitya Batchu*[*], Madhuvani Patnam, Goutham Karnati

Electronics and Communication Engineering
Hyderabad, Telangana, 500039

abisahi2002@gmail.com, madhuvanikashyap@gmail.com, karnatigoutham26@gmail.com

Abstract

One of the leading factors in traffic accidents and fatalities is drowsy driving. Hence, identifying driver weariness and related symptoms is a current research subject. The majority of traditional approaches are either vehicle-, behavior-, or physiological-based. Some approaches necessitate costly sensors and data processing, while others are obtrusive and distracting to the driver. The result of this work is the development of real-time, low-cost technology that can identify drowsy driving. A webcam records the video in the designed system, and utilizing image processing algorithms, the driver's face is recognized in each frame. Following the computation of the eye aspect ratio, mouth opening ratio, drowsiness is determined by adaptive thresholding based on the location of facial landmarks on the identified face. Moreover, offline implementations of machine learning algorithms have been made. Based on support vector machine classification, it was found that the sensitivity was 95.58% and the specificity was 100%.

Keywords: *drowsy driving, detection of drowsiness, eye aspect ratio, mouth opening ratio*

1. Introduction

One of the main factors contributing to fatalities in traffic accidents is drowsy driving. Long-distance bus drivers, nighttime bus drivers, and truck drivers who operate for extended periods of time are particularly prone to this issue. In every country, passengers' most frightening nightmare is a drowsy driver. Many people are hurt or killed in traffic accidents each year as a result of driver drowsiness. Thus, due to its enormous practical relevance, the detection of driver weariness and its indication are significant study areas. Three modules make up the fundamental sleepiness detection system: the acquisition system, the processing system, and the warning system. In this instance, the frontal face of the driver is collected on camera by the acquisition system and sent to the processing block for online analysis to identify tiredness. The alert system will give the driver a warning or alarm if drowsiness is detected.

Drowsy driving detection techniques are often divided into three categories: vehicle-based, behavioral-based, and physiological-based. In the vehicle-based approach, many variables, such as steering wheel movement, brake or accelerator pattern, vehicle speed, lateral acceleration, and lane position deviations, are continuously tracked. Driver drowsiness is deemed to exist when any abnormal change in these parameters is discovered. Due to the sensors' lack of attachment to the driver, this measurement is non-intrusive. With a behavioral-based method, drowsiness is identified by analyzing the driver's visual behaviour, such as eye blinking, eye closing, yawning, head bending, etc. Given that only a basic camera is needed to identify these traits, this assessment is also non-intrusive. The degree of sleepiness or weariness is determined by physiological signals such as heartbeat, pulse rate, Electrocardiogram (ECG), Electooculogram (EOG), and Electroencephalogram (EEG).

This measurement is intrusive since the driver will be distracted by the sensors that are attached to him or her. The cost and size of the system will increase as more sensors are used.

[*] Corresponding author: E-mail: abisahi2002@gmail.com,

Nonetheless, adding more parameters or features could improve the system's accuracy. These elements push us to create a low-cost, accurate real-time driver sleepiness detection system. So, in order to make the system portable and low-cost, we have presented a webcam-based system to identify driver weariness solely from the face image utilizing image processing and machine learning algorithms.

2. The Proposed System

In Fig.A block diagram of the driver drowsiness monitoring system is shown. The footage is initially captured with a webcam. To get a clear shot of the face of the driver, the camera will be placed in front of him or her. To create 2-D images, the video's frames are removed. Histogram of oriented gradients (HOG) and linear support vector machine (SVM) for object detection are used to identify faces in the frames. Face detection is followed by the marking of facial landmarks on the photos, such as the locations of the eyes, nose, and mouth. A determination of the driver's tiredness is made utilising machine learning and variables such as eye aspect ratio, mouth opening ratio, and head posture. These variables are measured from the driver's facial landmarks. In the event of drowsiness, a warning alarm will be sounded to warn the driver. Below, each block's specifics are described.

The development a non-intrusive system that can detect human fatigue and alert quickly is the idea of the project. Drivers who do not regularly stop on long distances are highly vulnerable to drowsiness, which they often do not recognize early enough. According to the research conducted by the expert, approximately one-quarter of all grave motorway accidents are attributed to sleep-paced drivers in need of rest. They can be prevented by taking the effort to get enough sleep before driving, drink coffee or energy drink, or have a rest when the signs of drowsiness occur. Complex approaches, such as EEG and ECG, are used to detect drowsiness popularly. This method is very accurate for measuring, but contact metering needs to be used, and driver fatigue and drowsiness monitor are subject to many restrictions. Therefore, driving in real-time is not comfortable. This document provides a method by measuring the eye closing rate and yawning for the detection of drowsiness signs among drivers.

This system monitors the driver's eyes with a camera, and we can detect driver fatigue symptoms early enough by developing an algorithm that avoids sleep. This project will therefore assist in detecting driver fatigue in advance and alarms and pop-up warning output. In addition, a warning is manually disabled instead of automatically.

A driving simulation system is driven by a participant in the video, and a webcam takes place in front of the driving system is designed for the identification of the face of the video image. The aim to use the face area so that eyes and mouth can be detected in the face area. Once the face is found, the eyes and the simulator. The video is recorded with the webcam to see the change from awake to sleepy.

The eye and mouth detection parameters are generated inside the face image. The video will be transformed into frames per second. The eyes and mouth can be located from there. Once the eyes have been located, the eyes are opened or closed by measuring the intensity changes in the eye area. If the eyes are found to be closed for 4 consecutive frames, the driver is confirmed to be drowsy.

ISBN - 978-93-95038-50-8

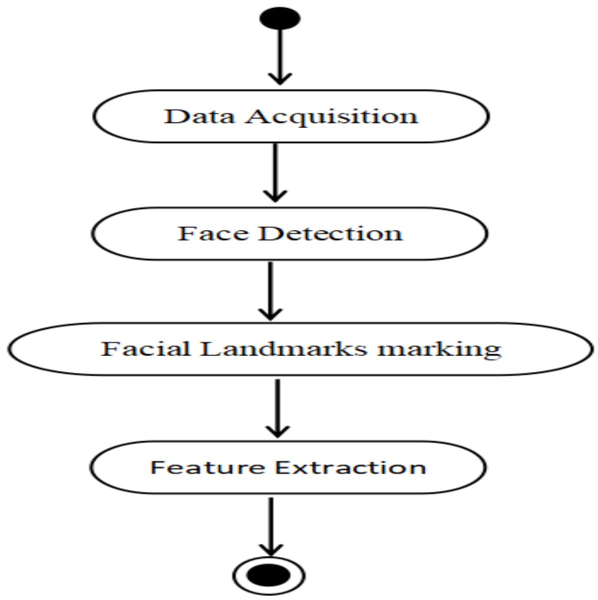

Figure 1: The block diagram of detection of drowsiness

A. DATA ACQUISITION

Sony CMU-BR300 camera is used to record the video, and a laptop is used to extract and process the individual frames. These 2D images are subjected to image processing methods after the frames have been extracted. At this time, artificial driver data has been produced. In order to participate in this experiment, participants are instructed to gaze at the webcam while yawning, sighing, and blinking occasionally. The video is recorded for thirty minutes.

B. FACE DETECTION

Faces are initially found after frames have been extracted. There are many face detection algorithms available online. In this study, the linear SVM algorithm is used in conjunction with the histogram of oriented gradients (HOG). This method computes HOG descriptors on positive samples of a fixed window size obtained from the images. Thereafter, identical-sized negative samples of the desired object to be detected (in this case, a human face) are taken, and HOG descriptors are computed. Often, there are far more negative samples than positive samples. A linear SVM is trained for the classification job after the characteristics of both classes are obtained. Intensive data mining is used to increase the SVM's accuracy. The classifier is evaluated on the labelled data in this technique after training, and the false positive sample feature values are used once more for training. The classifier computes the output for each window point for the test picture after translating the fixed-size window over the image. Ultimately, the recognized face is identified as the largest value output, and a bounding box is formed around the face. This non-maximum suppression step eliminates bounding boxes that are unnecessary and overlap each other.

C. FACIAL LANDMARKS MARKING

The next objective is to locate various facial features, such as the corners of the lips and eyes, the tip of the nose, and so on. This is after the face has been detected. Prior to that, the facial photos should be normalized to lessen the impact of camera distance, uneven lighting, and changing image resolution. The face image is therefore scaled down to 500 pixels in width and made into a grayscale image. An ensemble of regression trees is used to estimate landmark positions on the face from a sparse selection of pixel intensities after picture normalization. The sum of square error loss is improved through gradient boosting learning in this method. To find various structures, several priors are employed. Using this technique, the boundary points for the eyes, mouth, and the central line of the nose are marked. Table I lists the number of points for each boundary point. Figure B displays facial landmarks. The detected landmarks needing additional processing are indicated by red pointers.

Table 1 Points of Facial Landmarks

Facial Landmarks	
Parts	***Points***
Right eye	[37-42]
Left eye	[43-48]
Mouth	[49-68]

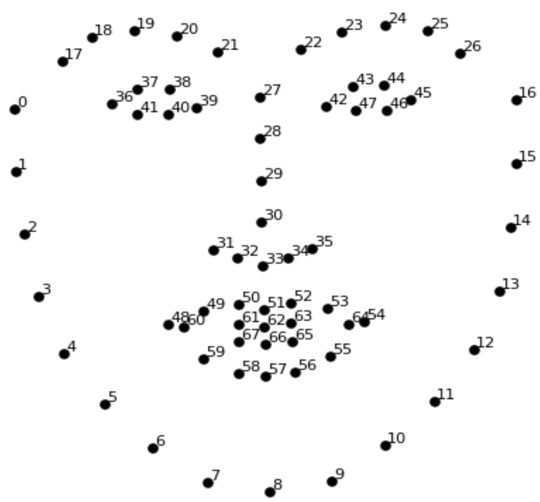

Figure 2 : Facial Landmark Points

D. FEATURE EXTRACTION

The characteristics are computed as follows after identifying the facial landmarks.

Eye aspect proportion (EAR): The ratio of the eye's height to its width, as determined by the eye corner points, is known as the eye aspect ratio and is given as below

$$EAR = \frac{||(p_2 - p_6)|| + ||(p_3 - p_5)||}{2||(p_1 - p_4)||} \qquad 1$$

Equation 1 : Eye Aspect Ratio

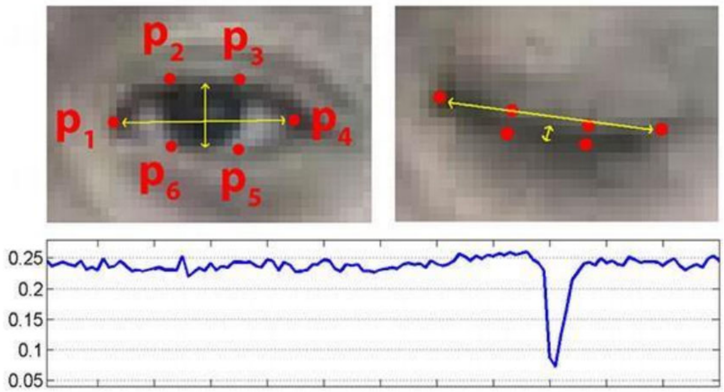

Figure 3: Top-left: A visualization of eye landmarks when then the eye is open.

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.

Mouth opening ratio (MOR): MOR is a parameter used to identify yawning while drowsy. It is calculated similarly to EAR as,

$$MOR = \frac{||(p_2 - p_8)|| + ||(p_3 - p_7)|| + ||(p_4 - p_6)||}{3||(p_1 - p_5)||} \qquad 2$$

Equation 2 : Mouth Opening Ratio

E. Classification

The following challenge is to identify sleepiness in the extracted frames after computing all three features. First, adaptive threshold is taken into account for categorization. Later, to categorize the data, machine learning methods are employed. It is assumed that the driver is fully awake at the start of computing the threshold values for each feature. The setup process is at this point. The EAR values for the first 24 frames are recorded during the setup process. The hard threshold for EAR is the average of 20 maximum values out of these 24 initial face-containing frames. The higher numbers are taken into account to ensure that there are no instances of eye closure. If the test result falls below this cutoff, drowsiness (or eye closure) is identified. This initial configuration for each person will lessen this effect because eye size might differ from person to person. Similar to this, the threshold for MOR is determined empirically from the observations because the mouth might not be fully open in the initial frames (setup phase). Yawn (i.e., drowsiness) is recognized if the test value exceeds this threshold. By measuring the angle the head makes with regard to the vertical axis in terms of the ratio of projected nose lengths, the head bending feature may be used. NLR typically ranges from 0.9 to 1.1 for the normal upright position of the head and it rises or falls when the head bends down or up while drowsy. Assuming there is no head bending, the average nose length is calculated as the average of the nose lengths throughout setup. The system is tested after the threshold values have been calculated. If at least one feature in a test frame indicates drowsiness, the system will recognise it.

This threshold is based on the previous 75 frames, which makes it more realistic. The system will indicate drowsiness detection and sound an alarm if at least 70 frames satisfy the sleepiness criterion for at least one characteristic. Another single threshold value that initially depends on the EAR threshold value is computed in order to make this threshold adaptive. The average of the 150 maximum values from among the 300 frames in the setup phase is used to calculate EAR values. The threshold is then obtained by subtracting the offset from the average value, which is determined heuristically. When EAR falls below this level, driver safety is put in jeopardy. With each yawn and head bend, this EAR threshold value rises slightly till a certain point. Because each yawn and head bending occurs over a number of frames, the yawning and head bending of succeeding frames are treated as a single event and are only added once to the adaptive threshold. If the EAR value in a test frame is below this adaptive threshold value, drowsiness is identified and the driver is alerted. The device may occasionally fail to detect the face when the head is too low as a result of bending. In this case, the previous three frames are taken into account, and if there was any head bending in those three frames, a sleepiness alarm will be displayed.

3. Results and Discussions

With the generated data, the suggested system has been created and tested. For additional processing and categorization of the internet video streaming, the webcam is connected to the laptop. The feature values are then saved for later statistical analysis and classification.

Figure 4 displays one frame from the usual or waking condition.

For this frame, the feature values are

EAR = 0.35 MOR = 0.38

Figure 4: Normal or awake state with facial landmarks

Figure 5: Drowsiness detected by the system due to (a) eye closing, (b) yawn, (c) head bending.

4. Conclusions

A low-cost, real-time driver drowsiness monitoring system based on visual behaviour and machine learning has been proposed in this work. Here, streaming video from a webcam is used to compute aspects of visual behaviour such eye aspect ratio, mouth opening ratio. A real-time method for detecting driver intoxication has been created using adaptive thresholding. With the generated synthetic data, the built system performs precisely. The feature values are then saved, and machine learning methods are applied to classify the data.

References

Journal Article

1. 1.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.
2. Book/Monograph
3. Richard O. Duda, Peter E. Hart, David G. Stork, “Pattern Classification”, Wiley student edition.
4. A chapter from a Book

5. 3.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,” 343IEEE Transactions on Industrial Informatics, vol. 12, no. 4, pp. 1438- 1452, Aug. 2016.
6. Conference Paper
7. 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.
8. V. Kazemi and J. Sullivan; "One millisecond face alignment with an ensemble of regression trees", IEEE Conf. on Computer Vision and Pattern Recognition, 23-28 June, 2014, Columbus, OH, USA.
9. 6.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
10. international Conf. on Consumer Electronics - Taiwan, 2015.