 https://doi.org/10.1051/e3sconf/202339101053

Image-based Driver Alert System for Prevention of Fatigue-related Accidents

G. Vijendar Reddy[1], P. Gopala Krishna[1], L. Sukanya[1], K. Sri Harsha[1], N. Harsha Vardhan[1], V. Arun[1]*

[1]Department of Information Technology, Gokaraju Rangaraju Institute of Engineering and Technology Hyderabad, Telangana, India.

Abstract. The objective of this project is to design a driver unconsciousness detection system using image processing to detect drowsiness and unconsciousness in drivers, thereby preventing accidents resulting from driver fatigue. Driver fatigue is a serious road safety issue, with approximately 20% of all road accidents attributed to this cause. Conventional drowsiness detection systems rely on physiological monitoring, which can be unreliable, expensive, and challenging to implement and maintain. In contrast, the proposed system monitors a sequence of images to identify facial and behavioral patterns indicative of drowsiness or unconsciousness. By detecting facial landmark points and analyzing the duration of eye closure, the system can accurately classify the driver's state and take appropriate measures such as reducing the vehicle's speed and alerting emergency services of the driver's geo-location. The successful implementation of this system holds immense potential for substantially reducing the number of accidents resulting from driver fatigue, thereby mitigating the loss of lives and injuries.

1 INTRODUCTION

Drowsiness is a widely prevalent phenomenon that directly impacts an individual's cognitive functioning and productivity [1]. The symptoms of drowsiness, including fatigue and a pronounced urge to sleep, serve as the primary indicators for its identification. Drowsiness detection has become a critical area of research, particularly in fields like analysis of behaviour, detection of fatigue, and measurement of alertness. The existing techniques for detecting drowsiness can be classified into two major categories: intrusive and non-intrusive. Intrusive techniques such as [2] – [4] employ measuring devices that are directly attached to the human body, whereas non-intrusive techniques such as [5] – [6] do not require any physical contact. While intrusive methods have been proposed in earlier studies, they are often impractical since they require additional hardware. On the other hand, non-intrusive techniques that use high-resolution cameras for image capture face limitations in unconstrained environments. To overcome these limitations, a new non-intrusive technique for detecting drowsiness is proposed. This approach uses a Haar cascade classifier, Histogram of Oriented Gradients, and Support Vector Machines for eye blink classification. The system

*Corresponding author : gurramvijendarreddy@gmail.com

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

detects blinks and computes the PERCLOS value, and if it exceeds 6000ms, it indicates that the person is drowsy. This approach doesn’t require any additional hardware and performs well in regular webcam resolutions, making it immune to variations in lighting conditions. The present study employed the programming language Python to implement the system, and its performance was assessed by means of a comparison between its outputs and those generated by a human expert. The study's results showed that the system's predictions agreed with the expert's opinions 95% of the time. The following research paper includes a detailed exposition of the proposed mechanism aimed at instantly detecting driver fatigue in Section 2. Subsequently, the findings of the experimental evaluations are presented and analyzed in Section 3. Finally, Section 4 furnishes a comprehensive conclusion, encapsulating key insights from the study and outlining potential avenues for future research.

2 DEVELOPED SYSTEM

The system's algorithm is composed of various sequential steps, as described below:

1. Gathering data using a web camera
2. Identifying facial features
3. Detecting and extracting regions of interest
4. Identifying blinks
5. Computing the PERCLOS score
6. Detecting drowsiness

A comprehensive elucidation of each of these procedures is expounded in subsequent sections.

2.1 Gathering data using a web camera

The initial stage of the process entails using a web camera to capture video footage of the test subject. The recorded video is subsequently partitioned into a series of discrete frames or images. Every image is extracted and processed individually as a separate entity.

2.2 Identifying facial features

The recently developed software employs images captured by the computer's camera to identify faces. The Viola and Jones algorithm [6] utilizing Haar-based cascade classifiers is leveraged to achieve this objective.

2.2.1 Haar Features

Haar features constitute a category of image features employed in computer vision and machine learning algorithms to facilitate object detection. The attributes are acquired by performing the Haar wavelet transformation on the image, which is known for its fast edge and contrast change detection capabilities. Through the examination of the response of these features across distinct image locations, an algorithm can ascertain the presence of a target object. Although Haar features have been predominantly employed in facial recognition and detection systems, they can be deployed to detect other objects and patterns in images as well. Overall, the use of Haar features has significantly advanced the development of efficient and robust object detection algorithms [7].

2.2.2 Cascade Classifiers

Cascade classifiers are a popular machine learning technique utilized for object detection in digital images or video. These classifiers comprise several stages, each of which is composed of multiple weak classifiers that have been trained to recognize specific features of the target object. The output of each stage is forwarded to the next stage, where another set of weak classifiers is trained to recognize additional features. This procedure is repeated until the object is detected or rejected. Cascade classifiers are highly efficient since they facilitate the swift rejection of non-object regions in an image, which reduces the computational burden of subsequent stages. The integration of cascade classifiers has enabled the creation of real-time object detection systems that are widely utilized in various applications such as face detection, pedestrian detection, and autonomous vehicles.

2.2.3 Face Detection

In the task of face detection, a cascade classifier is employed to analyze a specific image region and determine the presence or absence of a face. At each stage of the classifier, a predefined set of features is evaluated to assess the existence of a face. The classifier employs a multi-stage approach to evaluate the image region. The first stage serves as an initial test, and if it fails, the image region is immediately dismissed as non-facial. Conversely, if the region passes the first stage, it is then evaluated in the subsequent stages. This cascade strategy streamlines the computational process by allowing the classifier to extract a limited subset of features at each stage instead of evaluating all features simultaneously. The cascade classifier technique is a highly efficient face detection method, as it effectively filters out non-face regions early in the detection process. Due to its efficiency and high detection accuracy, this approach is popularly used in real-time face detection systems.

2.3 Detecting and extracting regions of interest

In the process of detecting faces, geometric characteristics of the human face are utilized to extract the eyes. These properties, as described in [8], are based on a unique geometric ratio of face features such as the eyes, nose, and mouth. The eye region is generally located between 0.2d and 0.6d from the top of the face, where 'd' represents the total length of the face. A more detailed discussion of this process can be found in Section 3 of the implementation description.

Figure 1 - Geometric ratios of human face

The next step after isolating the regions surrounding the eyes during face detection involves detecting the actual eyes within these regions. The detection process relies on a cascade classifier that utilizes Haar features. This classifier categorizes the eye region into those that contain eyes and those that do not by analyzing various eye images. However, it is important to note that this approach may not work if the eyes are obstructed or closed. In such cases, the system utilizes the eye positions identified in the previous video frame to extract eye images for further processing. The system's accuracy in detecting eyes can be evaluated using various metrics, such as precision and recall. By analyzing these metrics, areas for improvement can be identified to optimize the system's performance, especially under challenging conditions such as low lighting or obscured vision. Consequently, such optimization efforts can enhance the reliability and usability of applications dependent on eye-tracking technology. Further details on this process are available in Section 3 of the implementation discussion.

2.4 Identifying blinks.

Upon successful detection of the eyes, the system progresses to the blink detection phase. In order to identify blinks in the eye images, the system employs Histograms of Oriented Gradients [9] as features and Support Vector Machines as binary classifiers. HOG is a widely used technique in computer vision that analyzes the distribution of gradients in local areas of an image to detect object boundaries. SVM, on the other hand, is a machine learning algorithm that maps input features to binary labels, enabling binary classification tasks such as blink detection. By utilizing HOG features and SVM classifiers, the system is able to accurately identify blinks in the eye images, which can have a significant impact on the performance of eye-tracking systems.

2.5 Computing PERCLOS score

In the system being discussed, the PERCLOS (Percentage of Eye Closure over the Length of Observation time in seconds) measure is calculated based on the definition provided in [10]. It determines the time duration for which the eyes were closed and is represented as the fraction of time during a minute in which the eyelids were closed. The PERCLOS value is calculated using the formula shown in the figure below:

$$\frac{\textit{No of frames in which eyes are closed in one min}}{\textit{Total number of frames in one minute}} \times 60$$

Figure 2

2.6 Detecting drowsiness

The method for assessing drowsiness in this system relies on the analysis of several physiological parameters. Once these parameters are collected, they are used to determine the user's level of alertness. The analysis involves monitoring changes in the user's heart rate, blood pressure, and breathing patterns. When a person is drowsy, their heart rate slows down, their blood pressure drops, and their breathing becomes shallow. Based on these patterns, a threshold is established, and if the values fall below that threshold, the user is considered drowsy. The threshold values are calculated based on the data obtained from the healthy population, and this information is used to determine the appropriate levels for drowsiness. The threshold value is set such that it is low enough to avoid missing genuine cases of drowsiness but high enough to avoid false alarms. Therefore, if the user's physiological parameters are below the threshold, it is assumed that the user is drowsy, and appropriate action is taken to prevent any accidents. However, if the values are within the acceptable range, the user is considered to be alert and awake.

3 RESULTS

3.1 Implementation

The paper describes a system that employs a regular CMOS web camera to capture video frames. The camera has a resolution of 1280x720 pixels and captures 30 frames per second. Each frame of the video is treated as a separate image and is processed independently. Figure 3 illustrates a single video frame extracted from the recorded video using the web camera.
The initial stage of the process involves the extraction of facial features from the captured image. The process of feature extraction uses a complex set of algorithms that analyze the image for specific patterns and features. These algorithms are based on the principles of computer vision and pattern recognition, which have been developed over several decades. The algorithms are designed to identify specific patterns in the image, such as edges, corners, and curves, which are then used to identify the different facial features, such as the eyes, nose, and mouth. Once these features have been extracted, they are used to create a virtual representation of the face, which is then compared to a database of known faces. This database is constantly updated with new faces and features, making the system highly accurate and efficient. The feature extraction process is critical to the success of subsequent stages, such as face recognition and tracking, as it provides the necessary information for the system to identify and track individuals accurately.

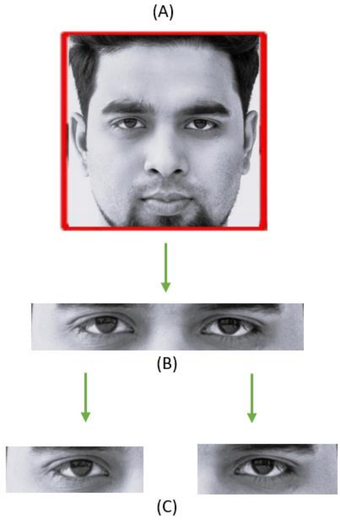

Figure 3 system to identfy

The system described in this passage computes facial geometry properties for human face images captured from video frames, as detailed in Section 2.3. The system extracts the eye regions and geometric ratios, as shown in Fig 3(a) and Fig 3(b) respectively. To analyze the eye regions further, the system divides the extracted region equally in half by cutting it at the center, as illustrated in Fig 3(c). The resulting halves of the eye region are then transmitted to the subsequent stages of the system for further analysis. The process of detecting eyes from the eye region involves the utilization of Haar cascade classifiers, as described in Section 2.4. The classifiers are similar to those previously used for detecting faces. The left and right eyes that are extracted are shown in Figure 4. Following this, the eye images are forwarded to the next stage for classification.

After detecting the eyes using Haar-based cascade classifiers, the system employs an eye blink detection algorithm, as presented in Figure 4. The algorithm uses a combination of Fourier Transform-based features and a Random Forest classifier for the detection of eye blinks. In this approach, 256 Fourier Transform-based features are extracted from each eye image and then fed to the Random Forest classifier for open or closed-eye classification. The PERCLOS value is computed by dividing the total number of frames where the eyes are closed by the total number of frames within a minute. The resulting value is then compared to a threshold of 5 seconds. If the PERCLOS value exceeds the threshold, the user is classified as drowsy, while if it is below the threshold, the user is considered to be awake. This methodology provides a precise and dependable technique for the real-time detection of drowsiness, which is crucial in ensuring the safety of drivers and operators in various industries.

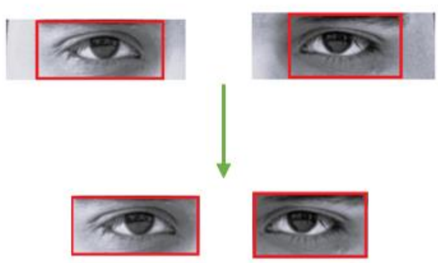

Figure 4- Eyes detected using Haar based cascade classifier applied on the eye region

3.2 Validation

In order to assess the efficiency of the proposed method, a total of twelve test videos were employed. These videos varied in their duration and frame rates and were captured in a standard room with regular lighting conditions. The algorithm was employed on the test videos, and the PERCLOS values obtained were utilized to classify the state of the individual in the video as drowsy or non-drowsy.

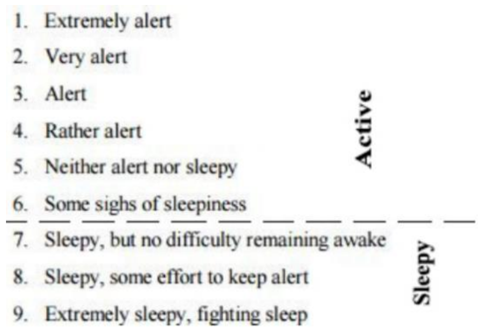

Figure 5 steps

The precision of the system was evaluated by comparing its predictions against the ratings obtained by a human rater using the Karolinska Sleepiness Scale. The KSS scale was simplified into two categories: "Active" and "Drowsy," which are illustrated in Figure 5.

The human rater evaluates the test videos by observing the subject and assigning a rating based on the Karolinska Sleepiness Scale (KSS) [11] . The KSS is presented as a visual scale, which is divided into two categories: "Active" and "Drowsy." [12] The graphical representation of the KSS scale used by the human rater is shown in Figure 5. The results of the human rater's evaluation are recorded in Table 1.

Sub No:	Time (Sec)	Resolution	Fps	PERCLOS(s)	Prediction	Human rater
1	35	1280x720	30	26.71	Drowsy	9
2	31	1280x720	25	15.11	Drowsy	7
3	100	1280x720	15	18.71	Drowsy	7
4	68	1280x720	10	1.09	Active	1
5	79	1280x720	15	3.05	Active	5
6	65	1280x720	15	17.48	Drowsy	9
7	**46**	**1280x720**	**30**	**9.09**	**Drowsy**	**6**
8	47	1280x720	30	27.72	Drowsy	8
9	40	1280x720	30	2.74	Active	4
10	36	1280x720	30	39.37	Drowsy	9
11	38	1280x720	30	27.41	Drowsy	8
12	37	1280x720	30	4.02	Active	2
13	41	1280x720	20	11.15	Drowsy	7
14	49	1280x720	20	5.03	Active	4
15	67	1280x720	15	29.51	Drowsy	9
16	78	1280x720	30	37.48	Drowsy	8
17	32	1280x720	25	8.17	Active	1
18	63	1280x720	30	20.07	Drowsy	7
19	83	1280x720	25	17.10	Drowsy	8
20	92	1280x720	30	31.32	Active	5

Table 1 - Comparison of predictions made by the developed and system and human rater

According to the data in Table 1, it can be inferred that the drowsiness detection system described in this research was successful in detecting drowsiness in individuals. The system's forecasts were in agreement with those of the human rater in 19 out of the 20 test cases. The only instance where the system's prediction did not align with the human rater's assessment is indicated in the table. The overall accuracy rate of the system is 95%.

4 CONCLUSION

This research paper focuses on the detection of drowsiness through web camera imagery. The proposed methodology entails utilizing a Haar-based cascade classifier to track the eyes and a HOG-SVM combination for eye blink detection. The PERCLOS value is computed and compared against a predetermined threshold of 6 seconds to ascertain the individual's level of alertness. The system's efficacy was evaluated and found to be comparable to that of a human evaluator. This non-invasive approach is suitable for implementation on desktop computers and mobile devices, without the need for specialized hardware, and performs optimally under normal lighting and resolution conditions. This technique presents potential applications in a variety of domains, including monitoring driver attentiveness, detecting vitality, measuring concentration, and assessing alertness.

References

1. Sandberg, D., Kecklund, G., Åkerstedt, T., & Lendahl, M. (2011). The Characteristics of Sleepiness During Real Driving at Night: A Study of Driving Performance, Physiology and Subjective Experience. Sleep.
2. Pal, N.R., Nasrabadi, N.M., & Principe, J.C. (2008). EEG-based Subject-and Session-Independent Drowsiness Detection: An Unsupervised Approach. EURASIP Journal on Advances in Signal Processing.
3. Lin, C.T., Ko, L.W., Chiou, J.C., & Duann, J.R. (2005). EEG-based Drowsiness Estimation for Safety Driving Using Independent Component Analysis. IEEE Transactions on Circuits and Systems I: Regular Papers.
4. Sahayadhas, A., Sundaraj, K., & Murugappan, M. (2013). Drowsiness Detection During Different Times of Day Using Multiple Features. Australasian Physical & Engineering Sciences in Medicine.
5. Kurian, D., Joseph, P.L.J., Radhakrishnan, K., & Balakrishnan, A. (2014). Drowsiness Detection Using Photoplethysmography Signal. In Fourth International Conference on Advances in Computing and Communications. IEEE.
6. You, C. W., Huang, J. C., Lin, J. J., Tsai, Y. C., & Lee, S. S. (2012). CarSafe: A Driver Safety Application for Detecting Hazardous Driving Behavior Using Dual-Cameras on Smartphones. Proceedings of the 2012 ACM Conference on Ubiquitous Computing. ACM.
7. Freund, Y., & Schapire, R. E. (1999). A brief overview of boosting. Journal of Japanese Society for Artificial Intelligence.
8. Oguz, Ö. (1996). Thumb rule of Leonardo da Vinci to determine the proportion of the face in younger adults: A study on surgical and radiologic anatomy. Surgical and Radiologic Anatomy.
9. Dalal, N. and Triggs, B. (2005). "Histograms of Oriented Gradients for Human Detection". Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2005. IEEE.

10. Wierwille, W. W. (1999). Historical Perspective on Slow Eyelid Closure: Whence PERCLOS? In Proceedings of Technical Conference on Ocular Measures of Driver Alertness.
11. Johns, M.W. (2009). Defining excessive daytime sleepiness.
12. Trutschel, U., Berger, A., Adame, B., & Krajewski, J. (2011). PERCLOS: An Alertness Measure of the Past. In Proceedings of the Sixth International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design.