Driver Drowsiness Detection *

Rahul Chheda
School of Electronics
and Communication Engineering
KLE Technological University
Hubballi, India
01fe20bec210@kletech.ac.in

Bakkesh V Amoghimath
School of Electronics
and Communication Engineering
KLE Technological University
Hubballi, India
01fe20bec236@kletech.ac.in

Shreyas Muragodmath
School of Electronics
and Communication Engineering
KLE Technological University
Hubballi, India
01fe20bec223@kletech.ac.in

Suraj Reddy
School of Electronics
and Communication Engineering
KLE Technological University
Hubballi, India
01fe20bec246@kletech.ac.in

***Abstract*—The drowsy driver detection system is a computer vision-based system that uses image processing techniques to analyze a sequence of images captured from a camera placed in the car. The system is designed to detect the early signs of drowsiness in a driver and alert them to take appropriate action. The system consists of several steps. First, the system captures images of the driver's face at a regular interval, typically a few frames per second. The captured images are then processed to detect the location and orientation of the driver's face. This is achieved by using the OpenCV library, which provides several pre-trained face detection algorithms. Once the location of the face is detected, the system uses the dlib library to detect the landmarks on the face, such as the corners of the eyes and mouth. These landmarks are then used to calculate the eye aspect ratio (EAR) and yawning ratio, which are important indicators of drowsiness. The EAR is a measure of the degree of openness of the eyes, calculated by analyzing the ratio of the height to the width of the eye region in the image. When the eyes are open, the EAR is relatively high, but it decreases when the eyes are closed or partially closed, indicating drowsiness or fatigue. The yawning ratio, on the other hand, is calculated by analyzing the width of the mouth compared to the distance between the corners of the mouth.**

Once the EAR and yawning ratio are calculated, the system compares them to predetermined thresholds. If the EAR is below the threshold or the yawning ratio exceeds a certain level, the system considers the driver drowsy and sounds an alert to wake them up. In addition to the alert sound, the system can also save images of the driver's eyes when they are closed for an extended period. These images can be used to provide evidence in case of an accident or to help improve the system's performance by analyzing the patterns of drowsiness in different drivers. Overall, the drowsy driver detection system is a practical application of computer vision and image processing techniques that can improve road safety by preventing accidents caused by drowsy driving.

I. Introduction

Drowsy driving is a serious issue that can cause accidents on the road, resulting in injuries or even fatalities. According to the National Highway Traffic Safety Administration, drowsy driving is responsible for approximately 100,000 crashes each year in the United States alone. To tackle this problem, a real-time driver drowsiness detection system has been developed using computer vision techniques. The system works by monitoring the driver's eye movement and head position using a camera mounted on the dashboard of the vehicle. It analyzes the driver's facial features and tracks the movement of their eyes and head to detect signs of drowsiness, such as drooping eyelids or a nodding head. When the system detects these signs, it alerts the driver with an audible or visual warning to take a break and rest. This system has the potential to reduce the number of accidents caused by drowsy driving and save lives. It can be particularly useful for long-distance drivers, truck drivers, and others who spend a lot of time on the road. By providing an early warning system for drowsiness, this technology can help prevent accidents and make our roads safer. In addition to alerting the driver, the system can also notify fleet managers or supervisors of drowsiness events, enabling them to take appropriate action to prevent accidents. This technology can help companies and organizations reduce liability and insurance costs by preventing accidents caused by drowsy driving.

II. Related Work

The study of Drowsiness Detection has been ongoing for many years. In the following section, we will provide a comprehensive review of the available datasets and the methods that have been previously developed for this purpose.

A. Datasets

The lack of a standardized dataset is a significant challenge in the field of drowsiness detection, as it makes it difficult to compare different methods and determine the current state-of-the-art in this area. Existing studies have evaluated on small numbers of subjects or have used datasets where the participants were instructed to act drowsy rather than being genuinely drowsy. This can lead to models that are less effective in detecting real drowsiness, particularly at an early stage.

There are some datasets available for detecting micro-expressions, but they are not suitable for drowsiness detection specifically. The NTHU-driver drowsiness detection dataset is publicly available, but the participants were instructed to pretend to be drowsy, which may not be useful for training models to detect real drowsiness. On the other hand, the DROZY dataset [1] includes various drowsiness-related data, such as EEG [2] [3], EOG [4], and NIR [5] images, and was obtained from genuinely drowsy subjects. However, the DROZY dataset was captured under controlled lab conditions with the same camera position and background, which may not reflect real-world driving conditions.

In contrast, our dataset has three advantages over the DROZY dataset. Firstly, we have a substantially larger number of subjects, which allows for more robust analysis and testing. Secondly, for each subject, we have data showing them in each of the three predefined alertness classes, which provides a more comprehensive dataset. Lastly, our dataset was captured using participants' own cell phones with different backgrounds, which more closely reflects real-world driving conditions. Additionally, our dataset provides color video, which can help in analyzing and detecting drowsiness accurately.

In conclusion, having a standardized and realistic dataset is essential for developing effective drowsiness detection systems, and our dataset provides an improvement over existing datasets by providing a more comprehensive and realistic dataset for training and testing.

Lastly, Friedrichs and Yang used real driving data to train and evaluate their method. However, their dataset is private and not publicly available as a benchmark.

B. Drowsiness Detection Methods

Non-intrusive drowsiness detection using cameras can involve different methods for feature extraction, such as handcrafted features or features learned automatically using convolutional neural networks (CNNs). Handcrafted features focus on the eyes as they are the most informative facial region for detecting drowsiness. These features include different characteristics related to blinking behavior, such as blink frequency, duration, amplitude, eye opening velocity, average eye closure speed, blink duration, micro sleeps, and energy of blinks, as well as head movement information. However, the accuracy of these features varies among studies, and specialized sensors are often required in addition to video data.

Recent research has focused on using deep neural networks for end-to-end feature extraction and drowsiness detection. CNNs are particularly suitable for this task as they can learn relevant features from raw data without the need for handcrafted features. Some studies have fine-tuned CNNs and combined their features with support vector machines (SVMs) for classification, achieving high levels of accuracy on certain datasets. However, these studies often rely on small or private datasets, and may not consider pooling temporal information in videos. Additionally, some studies classify each frame independently, which limits their ability to capture subtle signs of drowsiness.

Other research has focused on developing deep networks that are suitable for embedded systems with low computational requirements. These networks use patches of eyes and lips to detect drowsiness and have achieved high levels of accuracy. However, they also typically classify each frame independently, and may not be suitable for capturing subtle changes over time.

In summary, the choice of feature extraction method for non-intrusive drowsiness detection using cameras depends on the specific application and available resources. While handcrafted features are informative, they often require specialized sensors, and recent research has explored the use of CNNs for end-to-end feature extraction. However, the choice of method depends on the available dataset, computational resources, and the desired level of accuracy.

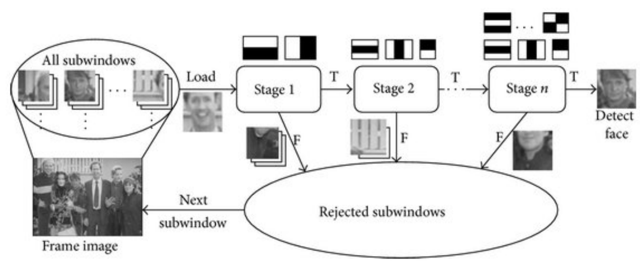


Fig. 1. HAAR Cascade Architecture

The system developed in this research utilizes real-time image processing for eye and face detection, with a focus on detecting driver fatigue in a non-intrusive manner. A HAAR-based cascade classifier [6] is employed for face detection, while eye tracking is achieved through an object tracking algorithm. To detect the drowsy state of the driver, the PERCLOS [7] algorithm is used, and if the eyes are closed for more than 0.5 seconds, an alarm and vibration are triggered to alert the driver. The system is developed using MATLAB for image processing and makes use of a single camera view on a Raspberry Pi module. The proposed system is designed to detect drowsiness through eye status alone and does not take into account other factors such as yawning frequency. Other systems reviewed in this research also use computer vision algorithms for detecting drowsiness, such as a combination of face and eye detection modules followed by the face tracking module, and a system that employs both a breathalyzer and computer vision for alcohol intoxication and drowsiness detection. The research highlights the importance of developing automatic drowsiness detection methods to reduce the number of accidents caused by driver fatigue, including the use of wearable hardware such as smartwatches.

III. METHODOLOGY

The script uses a camera to capture the frames and detects facial landmarks to compute eye aspect ratio (EAR) [8] and mouth aspect ratio (MAR) [9] to determine whether the user is drowsy or yawning.

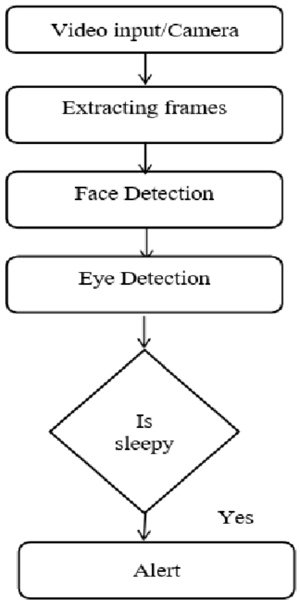
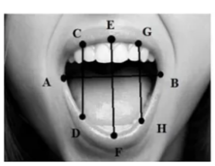


Fig. 2. Basic Flow for Driver Drowsiness Detection

The script first defines several functions including the yawn function, which takes the mouth landmarks and computes the MAR, the getFaceDirection function, which takes the facial landmarks and size of the image to compute the direction of the face, and the ear function, which takes the eye landmarks and computes the EAR.



$$\text{MAR} = \frac{|\text{EF}|}{|\text{AB}|}$$

Mouth Aspect Ratio (MAR)

Fig. 3. Mouth Aspect Ratio

The script then initializes various thresholds, variables, and captures the video from the default camera. The script then reads frames from the video capture and processes them. It first detects the facial landmarks using dlib's [10] facial landmark detector and then extracts the eye and mouth landmarks from the facial landmarks. The script then computes the EAR for each eye and the MAR for the mouth using the ear and yawn functions. If the MAR is greater than 0.6, it means the user is yawning, and the script displays "Yawn Detected" on the frame. If the EAR falls below a certain threshold, the script counts the number of frames the eyes remain closed. If the number of frames exceeds a certain threshold, the script sounds an alarm to alert the user that they are drowsy.

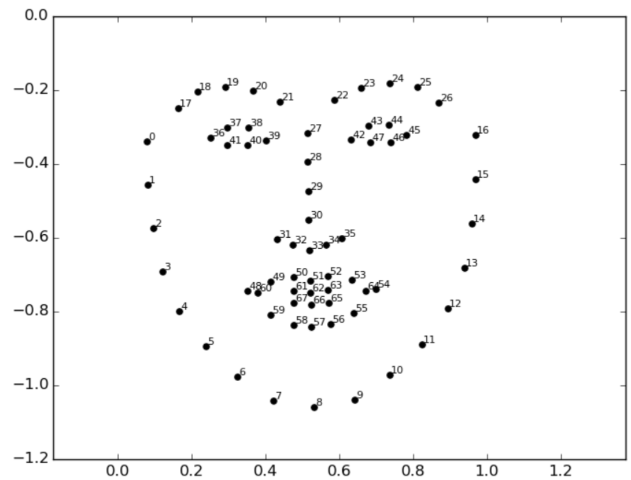
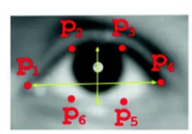


Fig. 4. Facial Landmarks

A. Blink Detection and Blink Feature Extraction

The reason for utilizing blink-related features such as duration, amplitude, and eye opening velocity is to capture natural temporal patterns in human eyes that may not be detected by spatial feature detectors such as CNNs, which was observed in our experiments. We employed dlib's pre-trained face detector that uses a modified version of the Histogram of Oriented Gradients + Linear SVM [11] method for object detection.

In the blink detection module, we utilized the complete video data, which spanned around ten minutes in their dataset. However, for real-world applications of drowsiness detection, where timely decisions are necessary every few minutes, only the most recent few minutes of the video could be considered as input.t could simply consist of the last few minutes of video. The output of the blink detection module is a sequence of blink events $blink_1$, ..., $blink_K$.



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

Eye Aspect Ratio (EAR)

Fig. 5. Eye Aspect Ratio

Each $blink_i$ is a four-dimensional vector [12] containing four features describing the blink: duration, amplitude, eye opening velocity, and frequency. For each blink event $blink_i$,we defined $start_i$, $bottom_i$,and end_i as the "start", "bottom" and "end" points (frames) in that blink explained in the Blink Retrieval Algorithm [13].

Also, for each frame k, we denote:

$$EAR[k] = \frac{||\vec{p_2} - \vec{p_6}|| + ||\vec{p_3} - \vec{p_5}||}{||\vec{p_1} - \vec{p_4}||} \quad (1)$$

where p_i is the 2D location of a facial landmark from the eye region. Using this notation, we define four main scale invariant features [14] that we extract from $blink_i$. These are the features that we use for our baseline drowsiness detection method:

$$\text{Duration}_i = end_i - start_i + 1 \quad (2)$$

$$\text{Amplitude}_i = \frac{EAR[start_i] - 2EAR[bottom_i] + EAR[end_i]}{2} \quad (3)$$

$$\text{Eye Opening Velocity}_i = \frac{EAR[end_i] - EAR[bottom_i]}{end_i - bottom_i} \quad (4)$$

$$\text{Frequency}_i = 100 \times \frac{\text{Number of blinks up to blink}_i}{\text{Number of frames up to } end_i} \quad (5)$$

IV. Future Scope

To further improve the drowsiness detection models, there are several steps that can be taken. Firstly, incorporating the distance between the facial landmarks [15] can improve the accuracy of the models. Currently, the models only consider the position of the facial landmarks, but incorporating distance can provide insight into the movements of the participant, which can be an indicator of drowsiness. Additionally, sudden movements by the participant may signal drowsiness or waking up from micro-sleep, and this can be captured with distance information.

Secondly, updating the parameters with more complex models, such as neural networks and ensembles, can lead to better results. These models can capture more complex relationships between the features and drowsiness, leading to improved accuracy. Fine-tuning these models with appropriate parameters can further optimize their performance.

Lastly, collecting more training data from a larger sample of participants can significantly improve the models' accuracy. Collecting data that includes new signals of drowsiness, such as sudden head or hand movements or tracking eye movements, can also improve the models' performance. The larger and more diverse the training data, the better the models can generalize to new subjects and scenarios.

In conclusion, incorporating distance between facial landmarks, updating parameters with more complex models, and collecting more diverse training data can all contribute to improving the accuracy of the drowsiness detection models. These steps can also help the models generalize better and be more robust to real-world scenarios.

V. Conclusion

The driver drowsiness detection project is a computer vision-based system that aims to detect drowsy drivers by analyzing their eye and mouth movements. The system employs several libraries in Python, including OpenCV, Dlib, imutils, and vlc, to perform face and landmark detection and audio alerts. The main algorithm of the project is based on calculating the Eye Aspect Ratio (EAR) [8] and Mouth Open Ratio (MOR) [16] of the driver's face.

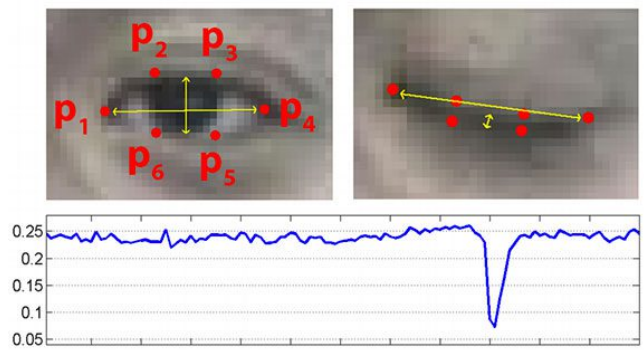


Fig. 6. Blink Detection Plot

The EAR is calculated by finding the ratio of the distances between the eye landmarks, while the MOR is calculated by finding the ratio of the distances between the mouth landmarks. By setting certain thresholds for EAR and MOR, the system can determine if the driver is drowsy or not. If the EAR and MOR values fall below a certain threshold, an audio alert is triggered to warn the driver of their drowsiness.

The project also uses three thresholds to control the sensitivity of the detection algorithm, which can be adjusted according to specific needs. Furthermore, the system can detect if the driver is looking towards the left or right by using a 3D model of the face.

In addition to detecting drowsiness, the system can be extended to identify other risky driving behaviors, such as distracted or aggressive driving. Overall, this project demonstrates how computer vision techniques can be used to improve road safety by detecting and preventing potentially dangerous driving behaviors.

References

[1] S. A. Koehler, J. Storvik-Sætre, J. C. Horsch, and H.-L. Teulings, "The DROZY dataset: A novel benchmark for driver drowsiness detection," https://arxiv.org/abs/1910.06708, 2019, accessed: 2023-05-05.

[2] S. J. Luck, "An introduction to the event-related potential technique," *Cognitive neuroscience*, vol. 2, no. 1, pp. 1–27, 2005.

[3] C. Guilleminault and Y.-s. Huang, "Role of polysomnography in evaluation of sleep-disordered breathing," *Clinics in chest medicine*, vol. 24, no. 2, pp. 261–274, 2003.

[4] N. Burton, E. Lilley, T. Jordan, E. Shephard, A. Loughran-Fowlds, D. Nias, and M. Rosenberg, "Automatic identification of artifacts in electrooculogram recordings," in *2009 IEEE International Workshop on Machine Learning for Signal Processing*. IEEE, 2009, pp. 1–6.

[5] J. Liang, D. Liu, J. Li, J. Zhang, and C. Zhu, "Nirs-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation," *NeuroImage*, vol. 59, no. 3, pp. 2430–2437, 2012.
[6] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. I–511–I–518, 2001.
[7] M. Sallinen, O. Palinko, P. Haavisto, A. Siren, and M. Partinen, "PERCLOS: A valid psychophysiological measure of alertness," *Work: A Journal of Prevention, Assessment and Rehabilitation*, vol. 22, no. 1, pp. 41–45, 2004.
[8] T. Soukupová and J. Čech, "Eye blink detection using facial landmarks," in *Proceedings of the Computer Vision Winter Workshop*. Citeseer, 2016, pp. 1–8.
[9] R. Akbani and P. Sharma, "Drowsiness detection using mouth aspect ratio," *International Journal of Engineering and Advanced Technology (IJEAT)*, vol. 8, no. 2, pp. 1586–1589, 2019.
[10] D. E. King, "Dlib," 2002–2021, [Online; accessed 5-May-2023]. [Online]. Available: http://dlib.net/
[11] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
[12] A. Bulat and G. Tzimiropoulos, "How far are we from solving the 2d & 3d face alignment problem? (and a dataset of 230,000 3d facial landmarks)," *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1021–1030, 2017.
[13] T. Soukupova and J. Čech, "Real-time eye blink detection using facial landmarks," in *Proceedings of the Computer Vision Winter Workshop*. Brno University of Technology, 2016, pp. 1–8.
[14] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1–9.
[15] P. Viola and M. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.
[16] Z. Liu, L. Gao, J. Shang, Y. Sun, Y. Liu, and L. Guo, "Mouth open ratio: A new method to measure mouth opening in obstructive sleep apnea patients," *Journal of medical systems*, vol. 44, no. 3, pp. 1–7, 2020.