**EngageWise - Monitoring, Alertness, and Well-Being in the Digital Workspace**

Vignya Durvasula [1] and Shreejit Cheela [2]

Researcher Scholars and Students

Department of Computer Science and Engineering [Artificial Intelligence & Machine Learning]

Vignana Bharathi Institute of Technology, Aushapur, Ghatkesar, Hyderabad-501301.

**Email**: vignyadurvasula@gmail.com [1] and shreejit.jithu2002@gmail.com [2]

**ABSTRACT**

As students approaching our time to work in a continuous, streamlined environment that requires a lot of focus to get things done, we often find ourselves procrastinating and losing focus or attention sitting in front of a computer or a laptop. Such real-life incidents have inspired us to develop an assistant that uses video-feed to report in real-time on how attentive a person sitting in front of the computer is while working or studying. We decided to name this AI assistant EngageWise to reflect its mission of fostering focused and informed engagement during work and study sessions. EngageWise is powered by the concepts Computer Vision and designed to assist individuals in maintaining their attention and productivity, ultimately helping them achieve their goals efficiently in a world full of distractions.

**Keywords:** focus, procrastination, attention, computer, laptop, AI assistant, EngageWise, Computer Vision, productivity, distractions.

**INTRODUCTION**

*You don’t need more time; you need more focus!* is a powerful quote by the author and entrepreneur Jesse Itzler. As students, we sometimes find it hard to focus on our work, be it studying for exams or working on a project on a computer or laptop. We often get distracted by other elements on our screens, toggling between pages and app icons. This behavior wastes time and hampers our attentiveness.

We came up with the idea of building a focus app that analyses how attentive the person is based on their facial movements and gestures and generates a report after the end of the session.

**1.1 Background Study:**

The existing models for an attention or focus detection systems are spread across various domains like attention heatmaps, pupil tracking, Eye Aspect Ratio (EAR), drowsiness detection systems and EEG based attention detection systems.

1. **Attention Heatmaps:** These models involve creating heatmaps of where a person's gaze or attention is concentrated on a screen. While not OpenCV-specific, this technique can be used to provide visualizations of attention patterns of the person.
2. **Pupil Tracking:** Pupil tracking is a common technique for assessing focus and attention. OpenCV can be used to detect and track the position of the pupil in the eye, which can provide insights into a person's focus based on the direction of their gaze.
3. **Eye Aspect Ratio (EAR):** The EAR is a measure of eye openness and can be used to detect blinks and assess alertness. In the proposed model, we use this ratio to detect the blinks and keep a check on how many times a person has blinked during the session.
4. **Deep Learning Models:** Deep learning models, particularly convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have been used to analyse facial expressions, eye movements, and other factors to predict attention and focus.

**1.2 Problem Statement:**

The primary objective of EngageWise is to devise a real-time reporting system that seamlessly integrates aspects from established models. To achieve this, we propose an amalgamation of OpenCV and Machine Learning in our initial prototype. The central challenge we address pertains to the accurate detection of instances where the user's attention is compromised, leading to distractions or loss of focus. Our long-term vision encompasses the development of an interactive application, which may take the form of a mobile or web app. This application will embody the identified functionalities, offering a user-friendly solution for maintaining focus and attentiveness in a variety of contexts.

**1.3 Research Objectives:**

1. **Real-Time Attention Monitoring:** Develop a real-time attention monitoring system that tracks and assesses a person's level of attention and focus while working or studying on a computer.
2. **User-Friendly Assistant:** Create a user-friendly AI assistant that seamlessly integrates into a person's workflow without causing distractions or interruptions.
3. **Early Warning System**: Develop an early warning system that alerts users when their attention starts to wane, helping them take proactive measures to regain focus and productivity.
4. **Reports**: Generate real-time reports and insights on attention patterns, including the duration of focused work, distractions, and periods of reduced attention.
5. **Customizable Settings**: Allow users to customize the system's sensitivity and alert thresholds to match individual preferences and needs.
6. **Privacy and Ethics**: Ensure strict adherence to privacy and ethical guidelines, respecting user data and anonymity while collecting attention-related information.

**1.4 Significance of Study:**

The existing models are only directed towards the attention of students during classroom sessions or eLearning. We aim to develop a system that can be used by every individual working on a computer. EngageWise can be used by students studying for their semester examinations, working professionals completing their tasks, teachers grading the students’ assignments, divergently, gamers who want to track their level of immersion and attention during gaming sessions and so on.

**LITERATURE SURVEY**

With EngageWise, we aim to develop a personal assistant that alerts the user when there are signs of drowsiness and decreased attention levels. One can seldom find an exact implementation of EngageWise; the probability is nearly negligible. However, it is a compilation of many computer vision tasks put in one place.

We started this journey considering the driver drowsiness or alert systems as our predecessors, as the principle behind EngageWise and these existing systems overlap a little.

**2.1 Previous studies**

Our idea to determine if a person is attentive or feeling drowsy based on the non-verbal behavior (facial, gestures) is inspired from this paper. The author of the paper [1] had a few objectives which aligned with our interests. The author carries out their research in a classroom, where a teacher and students are the actors, to determine the attention of the group of students during a lecture considering many non-verbal gestures like head pose estimation, gaze detection (to determine where the students are looking) and facial expressions. The base for determining the above-mentioned criteria is facial landmark detection, which is used in this paper and incorporated in our research.

The authors of the paper [2] conducted research to measure students' attentiveness in an eLearning environment based on facial detection and eye states. While this research is similar to EngageWise, we addressed some of its limitations. The authors of the paper considered only the state of the eye [open, closed] to assess students' visual attention. We extended this principle to detect if the person has their eyes closed or open, therefore incrementing the blink count.

The authors of [3] aimed to solve the attention deficit problems of students in online learning. AttenQ was designed as a tool to help teachers assess the students’ attention level so that the teacher can decide on what innovations they should come up with to keep the students attentive. The proposed solution uses computer vision and feature detection methods to monitor students' activities and provide teachers with information about their attention span. The system takes into account factors such as face detection, eye position estimation, head pose estimation, and drowsiness detection to determine whether a student is attentive or not.

The paper [4] discusses how a robot can calculate the level of visual focus of human attention. The authors of the paper proposed a system that can determine the level of visual focus of attention by analyzing eye movements and facial expressions. The system defines four levels of attention: high, medium, low, and no attention. In this paper, the scale of high-level attention is defined as ten, and the minimum scale of low-level attention is defined as zero. The level of attention is calculated based on the count of eyeball moves and time.

The authors of [5] have proposed a real time system using real time image processing, face/eye detection techniques, eye blink rates and yawning, which is designed a non-intrusive real time monitoring system for detecting driver’s drowsiness. This study forms a stepping-stone to calculate the blink and yawn detection in our research.

Our proposed model incorporates blink frequency and yawn count to determine if the person is drowsy or not. This idea is inspired from the paper [6] which uses EAR and PERCLOS to detect the blink and yawn counts to analyze if the driver is feeling drowsy. The methodology and libraries used in this research are incorporated in EngageWise with a few changes.

In the paper [7], the authors developed a system for monitoring driver fatigue and attention using various sensors and algorithms. The system aims to detect signs of fatigue, such as changes in mouth shape, eye closing time, and eye position, and alert drivers to prevent accidents caused by drowsy driving. We leveraged changes in mouth shape, eye position and closing time in EngageWise to determine if the person is drowsy or active.

[8] Although EngageWise is nowhere related to driver drowsiness detection systems that detect and alert a driver if he is drowsy or sleepy in a moving truck, EngageWise follows a same system to detect if a person in front of the system is not active during the session. Both the systems have a common functionality. The camera system used here monitors facial features and head position for signs of drowsiness, such as yawning and sudden head nods. Additionally, eye blinking rate and eye closure duration are measured to detect drowsiness.

The authors of [9] have conducted in-depth research in the field of driver drowsiness detection systems that have been developed till date. This is a systematic review paper that discusses the best possible techniques, measures, tools, and classification methods for detecting driver drowsiness. The paper is divided into seven main sections, which include a detailed analysis and evaluation of selected papers, a review of drowsiness detection techniques, a comparative study of drowsiness detection techniques, and a discussion of classification methods used for drowsiness detection. The paper concludes with a summary of the findings and suggestions for future research. The study aims to provide a comprehensive understanding of the state-of-the-art research in the field of drowsiness detection systems.

Our idea to use Eye Aspect Ratio to determine if the person’s eyes are open or closed is inspired from this research [10]. The driver's images are processed to detect the drowsiness of the driver by performing live monitoring of Eye Aspect Ratio (EAR) using pre-trained Neural Network based dlib functions. The EAR technique calculates the ratio of the eyes using Python interpreter in order to recognize drowsiness. The system can identify the occurrence of microsleep easily and alert the driver through an alarm for any positive drowsy condition. The developed system occupied with the Pi camera, Raspberry Pi 4, and GPS module are used to detect and analyze continuously the state of eye closure in real-time. The system can recognize whether the driver is drowsy or not, with the initial, wearing spectacles, dim light, and microsleep condition experimental conducted successfully give 90% of accuracy.

**Summary:**

To recapitulate, EngageWise is developed as a personal assistant for detecting drowsiness and decreased attention levels. Drawing inspiration from various research papers, the system combines elements of facial behavior analysis and non-verbal cues to assess attentiveness. Notable influences include research on student attention assessment in classrooms, attentiveness in eLearning, and driver drowsiness detection techniques. EngageWise's methodology incorporates features like blink frequency and yawn count, similar to the techniques employed in these papers. While its primary focus is on personal attention monitoring, EngageWise shares common functionalities with driver drowsiness detection systems, such as assessing facial features and head position for signs of drowsiness.

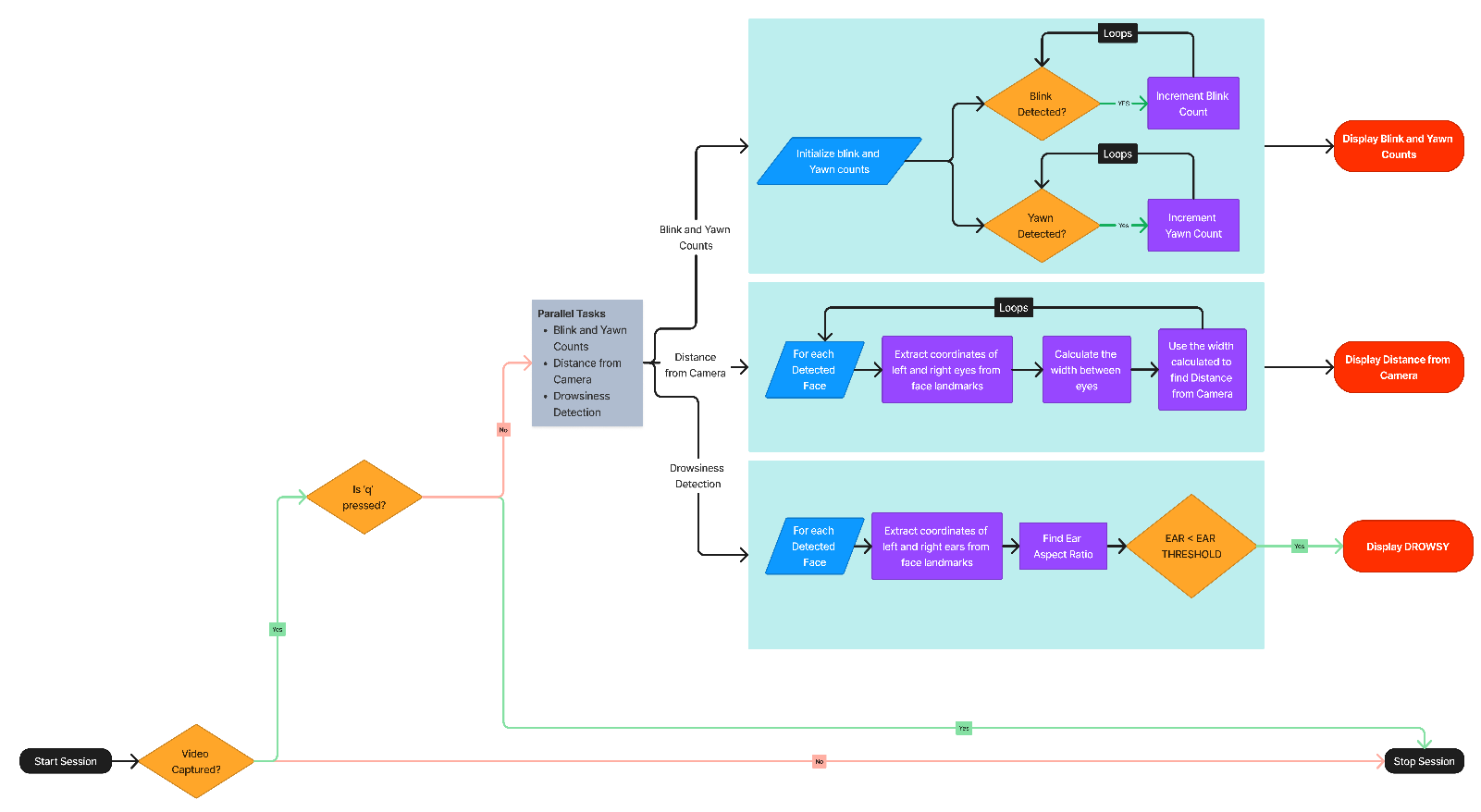
**METHODOLOGY**

**3.1 Design and Architecture**

The success of real-time computer vision applications, such as facial landmark detection and distance estimation, often hinges on the capabilities of the underlying hardware. For the Python-based facial analysis code snippet under consideration, the hardware and software requirements are notably influenced by the need for efficient image processing.

Our model captures the real-time video through the user’s webcam, calculates how far the user is from the camera, keeps a track of the blink frequency and yawn count, therefore display an alert that the user is drowsy if they seem to doze or show sleepiness.

The flowchart below clearly explains all the events starting from the user initiating the session to ending the session.



**3.2 Implementation**

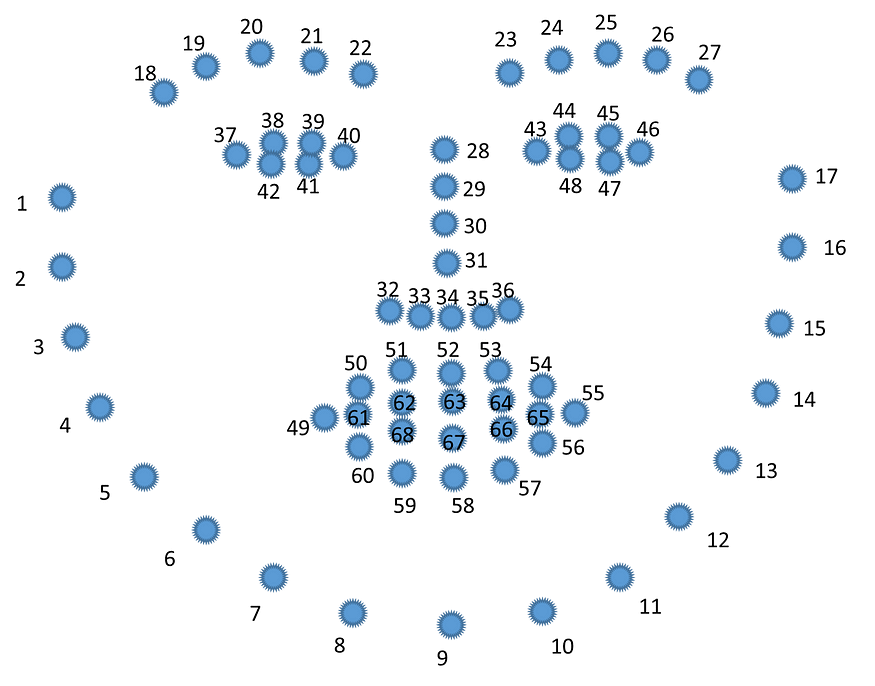
In this section, we present the implementation details of EngageWise using Computer Vision and Facial Landmarks. The system leverages various computer vision techniques and libraries to perform real-time detection of blink and yawn events

**3.2.1 Important Terminology**

* **Haar Cascade Classifier:** Although it's not explicitly mentioned in the code, the cv2.VideoCapture and dlib.get\_frontal\_face\_detector functions often use Haar Cascade Classifiers for face detection. Haar Cascades are a machine learning object detection method used to identify objects in images or video.
* **Facial Landmark Detection (dlib):** The code utilizes dlib's facial landmark detection model. While not named in the code, dlib uses a combination of shape predictors and trained models to locate specific facial landmarks (such as eyes, nose, mouth, etc.) on detected faces.
* **Eye Aspect Ratio (EAR):** The EAR is a blink detection technique. It calculates the ratio of distances between specific points on the eye, typically using the landmarks of the left and right eyes. If this ratio falls below a threshold, a blink is detected.
* **Mouth Aspect Ratio (MAR):** The MAR is used for yawn detection. It computes the ratio of the vertical distance between the upper and lower lip to the horizontal distance between the corners of the mouth. A yawn is detected if this ratio exceeds a threshold.
* **Gaussian Blur:** cv2.GaussianBlur is applied to the video frames. Gaussian blurring is a technique for reducing noise and detail in an image, which can improve the accuracy of facial feature detection.

**3.2.2 Facial Landmark Detection**

The system relies on facial landmarks to track the movements of eyes and the mouth. The facial landmarks are detected using the dlib library, which provides a pre-trained shape predictor model ("shape\_predictor\_68\_face\_landmarks.dat"). These landmarks are used to calculate various metrics related to eye and mouth movement.

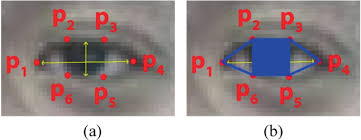


The key landmarks used in the system include:

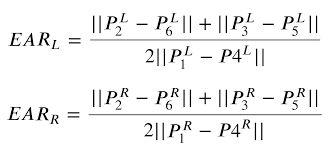
* Mouth landmarks (Indexes 48 to 68)
* Right eyebrow landmarks (Indexes 17 to 22)
* Left eyebrow landmarks (Indexes 22 to 27)
* Right eye landmarks (Indexes 36 to 42)
* Left eye landmarks (Indexes 42 to 48)
* Nose landmarks (Indexes 27 to 35)
* Jaw landmarks (Indexes 0 to 17)

**3.2.3 Blink Detection**

Blink detection is performed by measuring the Eye Aspect Ratio (EAR). The EAR is calculated using the Euclidean distances between key points of the left and right eyes. If the EAR falls below a predefined threshold (EAR\_THRESHOLD), it is considered as a blink event. The system keeps track of the blink count.

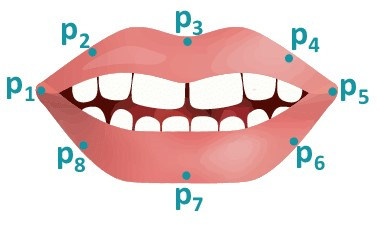


The formula for calculating the EAR is given below



**3.2.4 Yawn Detection**

Yawn detection is based on the Mouth Aspect Ratio (MAR). The MAR is calculated by measuring the ratio of the vertical distance between the upper and lower lip to the horizontal distance between the corners of the mouth. If the MAR exceeds a predefined threshold (MAR\_THRESHOLD), it is detected as a yawn event. The system maintains a count of yawn events.

**3.2.5 Distance Estimation**

The system also incorporates distance estimation functionality using face detection. It utilizes the dlib face detector to identify faces in the video stream. The width of the detected face is used to estimate the distance from the camera, considering the known width of an average human face.

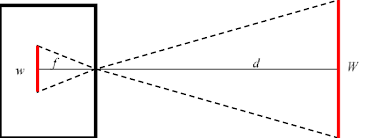


Figure 4.5

**3.2.6 User Interaction**

The system provides a user interface for capturing images by pressing the 'c' key. Captured images are saved with incremental filenames (e.g., "image1.png"). The 'q' key is used to exit the program.

**3.2.6 Real-time Visualization**

The processed video frames are displayed in real-time using OpenCV's 'imshow' function. Various textual information, such as blink count, yawn count, and drowsiness status, is overlaid on the video feed to provide real-time feedback to the user.

**3.2.7 Execution and Termination**

The system continuously captures frames from the camera feed, processes them for blink and yawn detection, and displays the results. The program, if running on the terminal, can be terminated by pressing the 'q' key else, if running on website after flask integration should be terminated using “Stop Session” or CTRL+C. Upon termination, the camera is released, and all OpenCV windows are closed.

**RESULTS**

**4.1 Active/Open Screen**



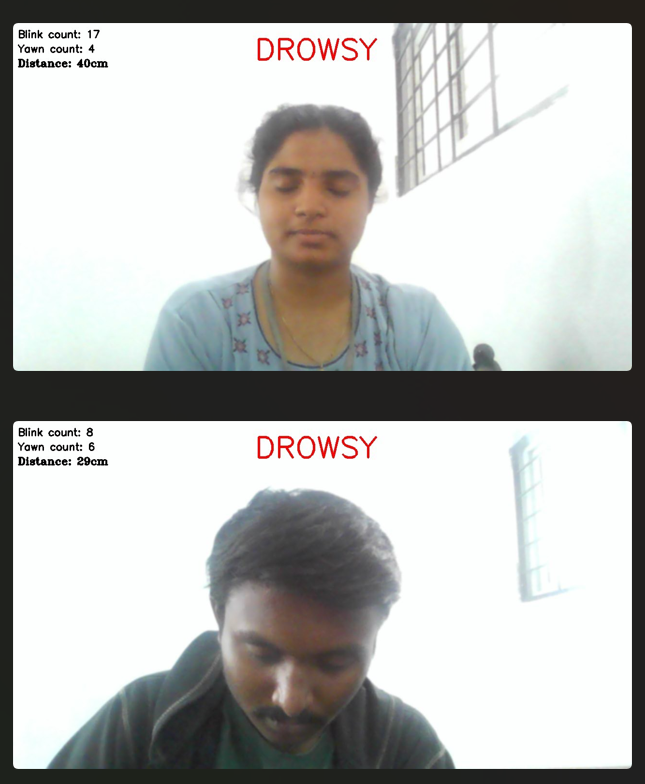
**4.2 Blink Detection**



**4.3 Yawn Detection**



**4.4 Drowsiness Detection**

****

**CONCLUSION**

EngageWise is a sophisticated focus detection application designed to assess the attention span of individuals, including students preparing for exams, working professionals, and teachers. Leveraging advanced computer vision techniques and facial feature analysis, the application provides real-time insights into the user's level of attentiveness.

The system employs a combination of facial landmark detection, blink, and yawn analysis to gauge the user's focus during study or work sessions. By monitoring the Eye Aspect Ratio (EAR) for blink detection and the Mouth Aspect Ratio (MAR) for yawn detection, EngageWise can effectively identify signs of drowsiness or reduced attention.

**5.1 Future Enhancements:**

EngageWise is not entirely accurate and opens the possibility of enhancements in many possible ways. We definitely look forward toenhance the model by fine tuning the parameters, adding a few functionalities to make it more engaging, incorporate user feedback mechanism and prioritize security and privacy of the users.

**REFERENCES**

[1] Raca, Mirko. (2015). Camera-based Estimation of Student's Attention in Class.

[2] Qingshan Deng and Zhili Wu 2018 IOP Conf. Ser.: Earth Environ. Sci. 199032042

[3] Pooja Koshti(2022) AttenQ- Attention Span Detection Tool for Online Learning

[4] Chakraborty, Partha & Yousuf, Mohammad & Zahid, Zahidur & Faruqui, Nuruzzaman. (2020). How Can a Robot Calculate the Level of Visual Focus of Human’s Attention. 10.1007/978-981-15-3607-6\_27.

[5] Kumar, N., & Barwar, D.N. (2014). Detection of Eye Blinking and Yawning for Monitoring Driver’ s Drowsiness in Real Time.

[6] Maganti Manasa, Vikas B, K. Subhadra (2019). Drowsiness detection using Eye-Blink frequency and Yawn count for Driver Alert.

[7] Ling Gan, Bing Cui and Weixing Wang, "Driver Fatigue Detection Based on Eye Tracking," 2006 6th World Congress on Intelligent Control and Automation, Dalian, China, 2006, pp. 5341-5344, doi: 10.1109/WCICA.2006.1714090.

[8] Vandna and Rekha Saini. “Driver Drowsiness Detection System and Techniques: A Review.” (2014).

[9] M. Ramzan, H. U. Khan, S. M. Awan, A. Ismail, M. Ilyas and A. Mahmood, "A Survey on State-of-the-Art Drowsiness Detection Techniques," in IEEE Access, vol. 7, pp. 61904-61919, 2019, doi: 10.1109/ACCESS.2019.2914373.

[10] S. Sathasivam, A. K. Mahamad, S. Saon, A. Sidek, M. M. Som and H. A. Ameen, "Drowsiness Detection System using Eye Aspect Ratio Technique," 2020 IEEE Student Conference on Research and Development (SCOReD), Batu Pahat, Malaysia, 2020, pp. 448-452, doi: 10.1109/SCOReD50371.2020.9251035.

[11] B. Alshaqaqi, A. S. Baquhaizel, M. E. Amine Ouis, M. Boumehed, A. Ouamri and M. Keche, "Driver drowsiness detection system," 2013 8th International Workshop on Systems, Signal Processing and their Applications (WoSSPA), Algiers, Algeria, 2013, pp. 151-155, doi: 10.1109/WoSSPA.2013.6602353.

[12] L. Thulasimani; Poojeevan P; Prithashasni S P. "Real Time Driver Drowsiness Detection Using Opencv and Facial Landmarks". Int. J. of Aquatic Science, 12, 2, 2021, 4297-4314.

[13] R. Vishnu et al 2021 J. Phys.: Conf. Ser. 1916 012172 "Driver drowsiness detection system with opencv and keras"

DOI 10.1088/1742-6596/1916/1/012172

[14] Y. Suryawanshi and S. Agrawal, "Driver Drowsiness Detection System based on LBP and Haar Algorithm," 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2020, pp. 778-783, doi: 10.1109/I-SMAC49090.2020.9243347.

[15] P. P. Patel, C. L. Pavesha, S. S. Sabat and S. S. More, "Deep Learning based Driver Drowsiness Detection," 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2022, pp. 380-386, doi: 10.1109/ICAAIC53929.2022.9793253.

[16] Mohanty, Archit, and Saurabh Bilgaiyan. "Drowsiness detection system using KNN and OpenCV." Machine Learning and Information Processing: Proceedings of ICMLIP 2020. Springer Singapore, 2021.

[17] Neelanjan. (n.d.). GitHub - neelanjan00/Driver-Drowsiness-Detection: https://github.com/neelanjan00/Driver-Drowsiness-Detection

[18] Misbah. (n.d.). GitHub - misbah4064/drowsinessDetector: OpenCV Drowsiness Detector using Python and Dlib. GitHub. https://github.com/misbah4064/drowsinessDetector

[19] Steflayanto. (n.d.). GitHub - steflayanto/nwHacks-2020-Back-End. GitHub. https://github.com/steflayanto/nwHacks-2020-Back-End

[20] Zhermin. (n.d.). GitHub - zhermin/ripplecreate-attention-model: Attention Indexing Computer Vision POC for ripplecreate. GitHub. https://github.com/zhermin/ripplecreate-attention-model.

**AUTHORS PROFILE:**

[1]: Vignya Durvasula: Final Year Student in Department of CSE [AI & ML], VBIT.

[2]: Shreejit Cheela: Final Year Student in Department of CSE [AI & ML], VBIT.

**OUR IMPLEMENTATION:**

<Paste github collaborative public link>