

# A detailed synopsis on

“Real Time Drowsiness Detection and Alert System for enhanced safety”

**Submitted in partial fulfillment of the requirements of the award of degree of**

**BACHELOR OF ENGINEERING** **IN**

**COMPUTER SCIENCE & ENGINEERING**

**Submitted By:**

**ABDUL ALEEM [4VV22CS002]** **ATUL K BHARADWAJ [4VV22CS021]**

**CHETHAN T P [4VV22CS036]** **DARSHAN K T [4VV22CS039]**

**Under the guidance of** **Prof. Chaya Kumari H A Assistant professor**

**Department of Computer Science & Engineering** **VVCE, MYSURU**

**2024-2025**

**DEPARTMENT OF COMPUTER SCIENCE** **VIDYAVARDHAKA COLLEGE OF ENGINEERING**

**Abstract**

This project presents the development of a non-intrusive, real-time driver drowsiness detection system designed to mitigate road accidents caused by driver fatigue. Driver fatigue is a leading cause of traffic fatalities globally, and this project aims to provide a practical and accessible technological solution to this critical safety problem. The system leverages computer vision techniques, utilizing a standard webcam to continuously monitor the driver's facial features. The core of the system is built using the OpenCV library for real-time video stream processing and the Dlib library for high- fidelity facial landmark detection. The primary detection mechanism is based on the Eye Aspect Ratio (EAR), a computationally efficient and validated metric that quantifies the level of eye opening from video frames in real-time. The system analyzes the EAR value over a sequence of frames; if the EAR drops below a calibrated threshold for a sustained period, which is a strong indicator of drowsiness or a micro-sleep event, an immediate and distinct audio alarm is triggered to alert the driver. The project will culminate in a standalone, cost-effective proof-of-concept application developed in Python, demonstrating a practical and deployable solution to a critical road safety issue without necessitating specialized or expensive hardware components.

# Introduction

Driver drowsiness is a pervasive and critical issue in transportation safety, contributing significantly to the global burden of road accidents. Statistical analyses consistently identify driver fatigue as a primary factor in a substantial portion of vehicular collisions, with some studies attributing as many as 20% of all road accidents to fatigue-related impairments. This state of reduced alertness, often culminating in "micro-sleeps" i.e, brief, involuntary episodes of sleep lasting a few seconds severely impairs a driver's cognitive functions, reaction time, and decision-making abilities, thereby elevating the risk of catastrophic accidents.

Historically, solutions for monitoring driver alertness have been categorized into two main approaches: intrusive and non-intrusive. Intrusive methods typically rely on physiological sensors to measure biological signals, such as electroencephalography (EEG) for brain activity or electrooculography (EOG) for eye movement. While these methods can achieve high accuracy, their practical application in consumer vehicles is severely limited by their high cost, complexity, and the requirement for drivers to wear cumbersome sensor equipment.

This project is basically a computer vision-based system architected for real- time monitoring of a driver's state of alertness. The system operates by analysing key visual indicators of fatigue, with a primary focus on eye behaviour, which is widely recognized as one of the most reliable visual cues for assessing a driver's cognitive state. The OpenCV library is employed for its robust capabilities in managing and processing real-time video streams, while the Dlib library is utilized for its highly accurate and computationally efficient facial landmark detection model. The core analytical metric of the system is the Eye Aspect Ratio (EAR), a simple yet powerful geometric measure that effectively quantifies eye closure.

The overarching motivation for this project is to contribute meaningfully to the enhancement of transportation safety. By developing an accessible and effective tool to prevent fatigue-related accidents, it aims to demonstrate how established computer vision algorithms can be engineered into a socially impactful application.

# Objectives & Scope

**Objectives:**

* To develop a robust module for real-time face detection and facial landmark localization from a live video stream captured via a standard webcam, utilizing the capabilities of the OpenCV and Dlib libraries.
* To implement and validate the Eye Aspect Ratio (EAR) algorithm, enabling the system to accurately and efficiently classify the driver's eye state (i.e., open or closed) on a frame-by-frame basis.
* To create a standalone Python application that integrates the video capture, landmark detection, EAR calculation, and alerting modules into a single, cohesive, and end-to-end system capable of operating in real-time (e.g., at a processing speed exceeding 20 frames per second) on standard consumer hardware.
* To systematically evaluate the performance and robustness of the developed system using key metrics, including detection accuracy and response time, under a variety of controlled testing conditions to ascertain its real-world viability.

# Scope:

**Our project will cover**

* The system's detection mechanism will focus exclusively on drowsiness detection via eye-blink analysis. The core metric for this analysis will be the Eye Aspect Ratio (EAR), a well-documented and effective method for quantifying eye closure.
* The primary input source for the system will be a standard USB webcam or a built-in laptop camera, ensuring the solution is accessible and does not require specialized hardware.
* The system will be developed and delivered as a standalone desktop application using the Python programming language and its associated scientific computing and computer vision libraries (OpenCV, Dlib, NumPy, mediapipe etc.).
* The system's output will consist of a real-time audio alert, such as a distinct beep or alarm sound, to notify the driver of a detected drowsiness event. For demonstration and debugging purposes, the system will also provide a visual display of the video feed with overlaid facial landmarks and the calculated EAR value.

# Our project will not cover

* **Advanced AI Model Training**: The project will not involve the training of custom deep learning models, such as Convolutional Neural Networks (CNNs) for image classification or Recurrent Neural Networks (LSTMs) for temporal analysis. Instead, it will leverage the highly accurate, pre-trained facial landmark detector provided by the Dlib library (shape\_predictor\_68\_face\_landmarks.dat). This approach ensures project feasibility and allows for a greater focus on system integration and real-time performance optimization.
* **Embedded Systems or Hardware Integration:** There will be no development for or integration with specialized in-vehicle hardware, Controller Area Network (CAN) bus systems, or low-power embedded platforms like Raspberry Pi or NVIDIA Jetson. The project is scoped as a software-based proof-of-concept for standard computer systems.

# Literature Review

The field of driver drowsiness detection has evolved significantly, moving from invasive physiological monitoring to sophisticated, non-intrusive computer vision techniques. A review of the existing literature reveals a clear trajectory of research that informs and justifies the methodology proposed for our project.

# Evolution of Drowsiness Detection Methodologies

Early research into driver fatigue primarily focused on intrusive methods that measure physiological signals directly from the driver's body. These include electroencephalography (EEG), which monitors brainwave patterns, and electrooculography (EOG), which tracks eye movements by measuring the electrical potential between the cornea and retina. They are impractical for widespread real-world application due to their requirement for direct skin contact, cumbersome sensor arrays, and high cost.

This led to the development of non-intrusive, behavior-based methods, which have become the dominant paradigm in modern research. This approach avoids any physical contact with the driver, making it a more acceptable and practical choice for integration into everyday vehicles. Within this domain, the analysis of facial features has emerged as the most promising avenue for research.

**The Cornerstone of Vision-Based Detection: Facial Landmark Analysis** The foundation of modern vision-based drowsiness detection is the ability to accurately and reliably locate key facial features in real-time. This is achieved through facial landmark detection, a technique where algorithms identify a set of pre-defined keypoints on a person's face, such as the corners of the eyes, the tip of the nose, and the contour of the mouth. The Dlib library's 68-point facial landmark predictor is a widely used and highly regarded tool for this purpose.The robustness of these landmark detectors against moderate variations in head orientation, ambient lighting, and facial expressions is a

critical enabler for the development of effective real-world driver monitoring systems.

# The Eye Aspect Ratio (EAR) as a Gold Standard Metric

Among the various metrics derived from facial landmarks, the Eye Aspect Ratio (EAR) has established itself as a gold standard for blink detection and drowsiness analysis. The EAR is a simple yet powerful metric calculated from the coordinates of the six landmarks surrounding each eye. The numerator computes the distance between the vertical eye landmarks, while the denominator computes the distance between the horizontal eye landmarks.

The key advantage of the EAR is that it yields a single scalar value that is largely invariant to the subject's head pose and the scale of the face in the image. When an eye is open, the EAR remains relatively constant at a certain value. When a blink occurs, the EAR value rapidly drops to nearly zero and then returns to its baseline. A state of drowsiness is often characterized by slower blinks and prolonged periods of eye closure, which manifest as a sustained period where the EAR value remains below a certain threshold. Numerous studies have successfully built entire drowsiness detection systems around this metric, demonstrating its validity, computational efficiency, and widespread adoption in the field.

# Beyond Single-Metric Systems: A Spectrum of Complexity

While the EAR provides a robust foundation, the literature also explores more complex systems that integrate multiple fatigue indicators to improve detection accuracy. This reveals a clear engineering trade-off between algorithmic simplicity and detection robustness.

To address this, more advanced systems incorporate multi-modal cue detection. This includes analyzing the Mouth Aspect Ratio (MAR) to detect yawns or using head pose estimation to identify patterns of head nodding, both of which are strong secondary indicators of fatigue. Further along the spectrum are systems that employ deep learning models. The most sophisticated approaches utilize spatio-temporal models, such as Long Short- Term Memory (LSTM) networks, to analyze a sequence of features (e.g., EAR, MAR, and head pose angles) over a time window. By considering the

temporal context of how these fatigue indicators evolve, LSTM-based systems can achieve higher robustness and differentiate between a simple blink and the gradual onset of a micro-sleep.

The following table summarizes key studies that inform the landscape of vision-based drowsiness detection and position the proposed work.

|  |  |  |
| --- | --- | --- |
| **Core Technique / Model and its study by Authors** | **Key Contribution / Finding** | **Limitations Noted** |
| Soukupová & Čech  Eye Aspect Ratio (EAR) | Proposed and validated EAR as a simple, real- time, and robust metric for blink detection using facial landmarks | The EAR value is person-specific; a single threshold is not universally optimal. |
| Rupani et al.  EAR, Dlib, OpenCV | Developed a complete system with an automated alert, demonstrating a practical application pipeline for drowsiness detection. | Relies on a single metric (EAR), potentially missing other fatigue cues like yawning or head nodding. |
| Zhang et al.  Spatio-Temporal LSTM | Used an LSTM network to analyze a sequence of features (eye, mouth, head pose) over time, improving recognition by capturing the temporal context of fatigue  development. | Computationally expensive; requires a more complex architecture and large, annotated training datasets. |

|  |  |  |
| --- | --- | --- |
| Hassaballah et al.  Multi-Scale Facial Landmark Detector (MSFLD) | Proposed an adaptive thresholding mechanism to solve the problem of inter-person variations in eye and mouth aspect ratios, improving robustness across different users. | The proposed detector and adaptive method are more complex to implement than standard, static-threshold approaches. |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

# Methodology

**Research and Analysis**

The initial phase of the project involves in-depth research to inform the selection of algorithms and tools, and to analyze the key parameters that will govern the system's behavior.

* **Algorithm Selection:** The core detection algorithm chosen for this project is the Eye Aspect Ratio (EAR). This decision is based on a comprehensive review of the literature, which consistently highlights the EAR as a highly effective, computationally efficient, and relatively simple metric for real-time blink detection. Its robustness to changes in scale and in-plane head rotation makes it an ideal choice for a proof-of- concept system that must perform reliably under variable conditions.
* **Tool Selection:** The technology stack was selected to leverage a powerful and mature ecosystem for scientific computing and computer vision.
  + **Python:** Chosen as the primary programming language due to its simplicity, extensive library support, and strong community.
  + **OpenCV (Open Source Computer Vision Library):** Selected for its comprehensive set of tools for real-time video capture, image manipulation, and on-screen visualization. It provides the essential backbone for the system's data input and output functionalities.
  + **Dlib:** Chosen specifically for its high-accuracy, pre-trained facial landmark prediction (shape\_predictor\_68\_face\_landmarks.dat).

This model is capable of localizing 68 specific keypoints on a face with remarkable speed and precision, which is fundamental to the EAR calculation.

* **Parameter Analysis:** The success of the system hinges on the careful calibration of two critical parameters. These will be determined empirically through systematic experimentation.
  + **EAR Threshold:** This is the specific EAR value below which the eye is considered closed. As literature suggests this value is person- specific, it will be determined by collecting EAR data from multiple test subjects during normal blinking and simulated prolonged eye closure. An average value that provides a good trade-off between sensitivity and specificity will be chosen as the default static threshold.
  + **Consecutive Frame Count:** This parameter defines the duration of eye closure required to trigger a drowsiness alert. A low value might lead to false alarms from long blinks, while a high value might fail to detect dangerous micro-sleeps. This count will be calibrated experimentally to correspond to a duration of approximately 2-3 seconds of eye closure, a common benchmark for drowsiness.

# Design and Development

* **Stage 1: Video Capture and Pre-processing**

The system will initialize a connection to the default webcam using OpenCV's cv2.VideoCapture() function. Within the main processing loop, it will read frames one by one. Each captured frame will undergo two pre-processing steps: resizing to a fixed width (e.g., 640 pixels) to standardize the input size and reduce computational load, and conversion to grayscale. The grayscale conversion is a critical optimization, as the face detection and landmark prediction algorithms operate on single-channel images, making the process significantly faster and more efficient.

# Stage 2: Face and Landmark Detection

For each grayscale frame, Dlib's built-in frontal face detector (dlib.get\_frontal\_face\_detector()) will be used to identify the region of interest (ROI) containing the driver's face. Once a face is detected, the pre-trained 68-point facial landmark predictor (dlib.shape\_predictor()) will be applied to this ROI. This will return a set of 68 (x, y)-coordinates that map to specific facial features, including the contours of the left and right eyes.

# Stage 3: Feature Extraction (EAR Calculation)

From the 68 landmark points, the coordinates corresponding to the six keypoints for the left eye and the six for the right eye will be extracted. The Euclidean distance between the vertical and horizontal pairs of these keypoints will be calculated using functions from the SciPy library. These distances will then be used to compute the EAR for each eye using the formula previously defined. An average EAR of both eyes will be taken to produce a single, stable value for the current frame.

# Stage 4: Decision Logic and Alerting

A counter variable will be maintained to track the number of consecutive frames where drowsiness is suspected. In each iteration of the loop, the calculated average EAR is compared against the pre- determined EAR threshold.

* + If the EAR falls below the threshold, the counter is incremented.
  + If the EAR is above the threshold, it indicates the eyes are open, and the counter is reset to zero.
  + A final check is performed to see if the counter has exceeded the pre-determined consecutive frame count threshold. If it has, the system concludes that a drowsiness event has occurred and triggers an immediate, audible alarm to alert the driver.

# Expected Outcomes

* A fully functional, real-time drowsiness detection application developed in Python. The application will be capable of processing a live video feed from a standard webcam, analyzing facial cues for signs of fatigue, and issuing timely alerts.
* A robust system that achieves a high level of detection accuracy, benchmarked against figures reported in relevant academic literature. The target is to achieve an accuracy of approximately 90-92% in controlled testing environments.
* A high-performance system demonstrating its real-time viability, characterized by a processing speed of over 20 frames per second (FPS) and a low response time of less than 0.5 seconds from the detection of a drowsiness event to the triggering of an alert.
* A comprehensive final project report detailing the project's background, methodology, implementation, testing results, and conclusions. This will be accompanied by a live demonstration of the working system, showcasing its capabilities in detecting simulated drowsiness events effectively.

# Resources Required

**Hardware Requirements**

* A modern laptop or desktop computer equipped with a multi-core processor (Intel Core i5/i7) to handle real-time video processing.
* A minimum of 8 GB of system RAM to ensure smooth operation without performance bottlenecks.
* A standard USB webcam or an integrated laptop camera with a minimum resolution of 720p to provide a clear video stream for analysis.

# Software Requirements

* **Operating System:** An operating system such as Windows 10/11.

# Core Libraries and Frameworks:

* + **OpenCV-Python:** Library for all computer vision tasks, including video capture, image pre-processing (resizing, color conversion), and visualization (drawing overlays, displaying video).
  + **Dlib:** Essential for its high-performance frontal face detector and its pre-trained 68-point facial landmark prediction model.
  + **NumPy:** A fundamental library for scientific computing in Python, required for efficient numerical operations on the arrays of landmark coordinates.
  + **SciPy:** Utilized for its spatial distance calculation functions (scipy.spatial.distance), which are necessary for implementing the Euclidean distance component of the EAR formula.
  + **imutils:** A collection of convenience functions to make basic image processing tasks with OpenCV easier and more streamlined while working on the iris of the eyes.
* **Programming Language:** Python (version 3.7 or newer).
* **Development Environment:** An Integrated Development Environment (IDE) such as Visual Studio Code or PyCharm, configured for Python development.

# Conclusion

This project sets out to address the critical and persistent issue of driver fatigue, a major contributor to road accidents worldwide. By developing a practical, low-cost, and non-intrusive monitoring system, this project demonstrates a direct application of computer vision technology to enhance public safety. The successful integration of OpenCV for video processing and Dlib for facial landmark detection has resulted in a robust data processing pipeline. The implementation of the Eye Aspect Ratio (EAR) algorithm serves as an effective and computationally efficient core for blink detection, enabling the creation of a functional real-time alert system. The project successfully meets its objective of delivering a tangible proof-of-concept that is both accurate and performant on standard hardware.

However, a critical reflection on the project also highlights its inherent limitations, which in turn illuminate clear and sophisticated pathways for future research and development. A significant future enhancement would be the implementation of an **adaptive thresholding mechanism**. Such a system would perform an initial calibration phase for each new driver, learning their unique baseline EAR and blink characteristics to set a personalized threshold, thereby dramatically improving the system's accuracy and reliability across a diverse user base.

Furthermore, the current system operates on a single modality - eye closure. While effective, this singular focus means it may miss other important fatigue indicators. This reflects the fundamental trade-off between algorithmic simplicity and detection robustness. Future work could evolve Vigil-Eye into a multi-modal system by integrating additional detection modules. This would involve incorporating the Mouth Aspect Ratio (MAR) to detect yawning and implementing head pose estimation to identify patterns of head nodding, both of which are strong complementary indicators of drowsiness. Fusing these multiple data streams would create a more comprehensive and context-aware assessment of the driver's state, reducing the likelihood of false positives and negatives.

Looking further ahead, a more advanced avenue of research would be to move beyond frame-by-frame analysis and incorporate spatio-temporal analysis. By employing models such as Long Short-Term Memory (LSTM) networks, a future version of the system could analyze sequences of features over time. This would allow the model to learn the temporal dynamics of fatigue—how eye closure, yawning, and head posture evolve and correlate over several seconds enabling it to distinguish between an innocuous long blink and the dangerous onset of a micro-sleep with even greater certainty.

In its current form, our project theme stands as a robust proof-of-concept that successfully bridges the gap between established computer vision algorithms and a pressing real-world problem. It underscores the potential for accessible technology to create socially impactful applications and provides a solid foundation upon which more sophisticated and intelligent driver safety systems where we can ensure that the suitable system can be built forward

# References

1. Bajaj, S., et al. (2023). A Real-Time Driver Drowsiness Detection Using OpenCV, DLib. In: *ICT Analysis and Applications*. Springer Nature Singapore.
2. Chen, L., et al. (2023). Real-Time Driver Drowsiness Detection System. *International Journal of Research and Publication Reviews*, 6(2).
3. Guo, R., et al. (2022). Fatigue Driving Recognition Method Based on Multi-Scale Facial Landmark Detector. *Sensors*, 22(24), 9898.
4. Jain, A., et al. (2023). Driver Drowsiness Detection System Using OpenCV and Keras. *International Journal of Progressive Research in Engineering Management and Science*, 3(5).
5. Kumbhar, S., et al. (2022). Driver Drowsiness Detection System Using Python. *International Research Journal of Modernization in Engineering Technology and Science*, 4(4).
6. Maior, C. B. S., et al. (2018). A Model for Drowsiness Detection Based on Eye Features in Real-Time Video Streams. *Proceedings of the 14th International Conference on Probabilistic Safety Assessment and Management (PSAM 14)*.
7. Mandal, B., et al. (2021). A Novel Approach for Driver Drowsiness Detection Using Eye Aspect Ratio and Dynamic Time Warping. *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*.
8. Mane, S., et al. (2022). Driver Drowsiness Detection Using Eye Aspect Ratio. *International Research Journal of Modernization in Engineering Technology and Science*, 4(5).
9. Mehta, S., et al. (2019). Real-time driver drowsiness detection system using eye aspect ratio and eye closure ratio. *2019 International Conference on Intelligent Computing and Control*

*Systems (ICCS)*.

1. Patel, M., et al. (2019). Driver Drowsiness Detection System using Eye-Blink and Yawn Detection. *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)*.
2. Rahman, M. M., et al. (2018). Fatigue Detection Using Facial Landmarks. *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*.
3. Reddy, B., et al. (2017). Driver Drowsiness Detection System using Computer Vision.

*International Journal of Engineering and Technology*, 9(4), 3233-3238.

1. Singh, S., et al. (2020). Driver Drowsiness Detection using Dlib. *International Journal of Engineering Research & Technology*, 9(5).
2. Sriram, B., et al. (2023). Driver Drowsiness Detection Using Python. *International Journal of Novel Research and Development*, 8(6).
3. Vamshi, K., et al. (2023). Driver Drowsiness Detection System Using Machine Learning.

*Tuijin Jishu/Journal of Propulsion Technology*, 44(3).

1. Zhang, C., Lu, X., & Huang, Z. (2020). A Driver Fatigue Recognition Algorithm Based on Spatio-Temporal Feature Sequence. *arXiv preprint arXiv:2003.08134*.
2. Zhang, W., et al. (2021). Driver Fatigue Detection Based on Facial Multi-Feature Fusion.

*Journal of Advanced Transportation*, 2021.

1. Zhang, Y., et al. (2023). A Driver Fatigue Detection Method Fusing ShuffleNet V2K16 and Dlib. *Applied Sciences*, 13(8), 5074.