Deeya Kharediya, 2026, 14:3
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science,
Engineering and Technology
An Open Access Journal

Save Vision: Smart Drive Monitor A Real-Time Multi-Indicator Driver Fatigue and Distraction Detection System

Deeya Kharediya [1], Ayushi Gupta [2], Ankita Shrivastava [3], Manish Kr. Suman [4]
[1,2] Student of B.TECH. (CSE)
Lakshmi Narain College of Technology and Science
Bhopal, MP, India
[3,4] Assistant Professor, CSE
Lakshmi Narain College of Technology and Science

Abstract- Driver fatigue, drowsiness, and cognitive distraction are among the most underreported yet dangerous causes of road accidents globally. Unlike impairment from alcohol or mechanical failure, these states are invisible to external observers and develop gradually, leaving little time for corrective action once a critical threshold is crossed. Despite advancements in vehicle safety engineering, a reliable, affordable, and non-intrusive system capable of detecting these states in real time remains largely inaccessible to everyday drivers — particularly in developing nations where road fatality rates continue to rise. Save Vision: Smart Drive Monitor is a camera-based, real-time driver monitoring system developed to bridge this gap. The system continuously captures the driver's facial region through a webcam and applies computer vision techniques to analyse four behaviorally significant risk indicators: prolonged eye closure as a marker of drowsiness, yawning as an early physiological signal of fatigue, head pose deviation as an indicator of distraction,and mobile phone usage as a documented cause of inattention. Detection is performed through facial landmark extraction, Eye Aspect Ratio (EAR) computation, and image-based classification signal of fatigue, head pose deviation as an indicator of distraction,and mobile phone usage as a documented cause of inattention. Detection is performed through facial landmark extraction, Eye Aspect Ratio (EAR) computation, and image-based classification, all processed entirely through software. Upon identifying a risk condition, the system activates a layered alert mechanism — combining an auditory alert through the system speaker and an on-screen visual warning — designed to capture the driver's attention proportionally to the severity of the detected state. The system accepts both live webcam input and pre-recorded video, making it flexible for real-world deployment as well as controlled testing environments. Save Vision is designed with accessibility and practicality at its core, targeting deployment without dependency on expensive hardware or cloud infrastructure. The system establishes a functional foundation for intelligent, preventive road safety — one that can be incrementally extended toward integration with advanced driver assistance systems and fleet-level monitoring platforms. By addressing driver risk at its earliest behavioral signs, Save Vision aims to contribute meaningfully to the reduction of road accidents caused by human inattention.

Keywords: Driver Fatigue Detection, Drowsiness Monitoring, Computer Vision, Eye Aspect Ratio (EAR), Facial Landmark Detection, Driver Monitoring System, Road Safety, Cognitive Distraction Detection, Yawn Detection, Head Pose Estimation, Mobile Phone Usage Detection, Real-Time Alert System, Intelligent Transportation, Advanced Driver Assistance Systems (ADAS), Preventive Safety Technology.

I. INTRODUCTION

Road safety has emerged as one of the most pressing public health challenges of the twenty-first century. Globally, road traffic accidents claim approximately 1.19 million lives each year, making them the leading cause of death among children and young adults between the ages of 5 and 29 [1]. The burden is disproportionately concentrated in low- and middle-income countries, which account for over 90 percent of all road fatalities. India, in particular, faces a severe and worsening crisis. According to the Ministry of Road Transport and Highways (MoRTH), a total of 4,80,583 road accidents were recorded across the country in 2023, resulting in 1,72,890 deaths — the highest figure in nineteen years, translating to approximately 20 fatalities every single hour [2]. Young adults between the ages of 18 and 45 accounted for 66.4 percent of all fatalities, representing a devastating loss of productive human life.

Among the causes identified by MoRTH, human behavioural factors dominate — with distracted driving and driver fatigue consistently cited as major contributors alongside overspeeding [2]. Unlike impairment from alcohol or mechanical failure, fatigue and distraction develop silently and progressively within the driver, leaving little opportunity for corrective action once a dangerous threshold is crossed. Existing vehicle safety technologies such as lane departure warnings and automatic emergency braking are reactive by nature, engaging only after a hazardous situation has already developed. A system capable of detecting unsafe driver states proactively and in real time remains a critical gap, particularly across the vast majority of vehicles on Indian roads that carry no embedded driver monitoring of any kind.

II. PROBLEM STATEMENT

Current driver monitoring solutions that do exist are either embedded within premium vehicles far beyond the reach of average consumers, or they rely on specialised hardware setups that are costly and impractical for everyday use. Systems that depend on a single detection indicator — such as eyelid tracking alone — are inherently fragile under real-world conditions, failing under poor lighting, partial facial occlusion, or natural physiological variation across different drivers [3]. There is therefore a clear and pressing need for a multi-indicator, camera-based, non-intrusive, and software-driven driver monitoring system capable of functioning reliably using nothing more than a standard webcam and a computing device. Save Vision: Smart Drive Monitor is developed as a direct response to this gap, combining four behavioural detection streams — eye closure, yawning, head pose deviation, and mobile phone usage — into a unified real-time monitoring and alert system that runs entirely in software.

III. LITERATURE REVIEW

Driver monitoring has been a growing area of research, especially as road accidents linked to human inattention continue to rise. Researchers have approached this problem from different angles — some focusing purely on the eyes, others on the face as a whole, and more recent works combining multiple signals together. Going through the existing literature helped us understand what has already been tried, what works well, and more importantly, where the gaps still exist.

One of the most referenced works in this area is by Soukupová and Čech [3], who introduced the concept of Eye Aspect Ratio (EAR) — a simple ratio calculated from six points around the eye that drops significantly when the eye closes. What made this work stand out is that it runs in real time on a standard camera without any special setup. Almost every drowsiness detection system published after 2016 builds on this idea in some way, and our system does too. The simplicity of the EAR formula is what makes it practical for software-based monitoring systems running on everyday computing devices.

A more complete approach was presented in a study [4] that combined EAR with Mouth Aspect Ratio (MAR) and head pose together, testing the system on the NTHU Driver Drowsiness Detection dataset. The results were impressive — up to 99% accuracy — but what was more useful for us was the observation that using multiple facial features together is significantly more reliable than relying on eye closure alone. A driver might have naturally small eyes, or lighting might cause the EAR to behave unexpectedly. Having additional signals like yawning and head movement helps the system stay accurate even when one indicator fails.

On the topic of yawning, Kamboj et al. [5] specifically worked on detecting yawns using the Mouth Aspect Ratio computed from facial landmarks. They also used data augmentation to deal with the fact that yawning datasets are relatively small, and still achieved around

96.69% accuracy. This was useful to us because yawning is often treated as a secondary feature in many systems, but it is actually one of the earliest physical signs of fatigue — often appearing well before eye closure becomes prolonged. Catching it early gives the system a chance to alert the driver sooner.

For distraction detection, Zhao et al. [6] proposed using continuous head pose estimation — tracking the yaw, pitch, and roll angles of the driver's head — as the primary indicator of distraction. They found that head orientation alone carries enough information to classify whether a driver is paying attention to the road. This confirmed our decision to include head pose as one of the four detection streams in Save Vision, since it is computable directly from facial landmarks without any additional sensors or specialised equipment.

A recent paper [7] brought all of these threads together in a unified software-based system using facial mesh and a deep learning classifier, testing on 27 participants in real and simulated driving environments. Their distraction detection hit 100% accuracy, eye closure detection reached 88.89%, but yawning detection sat at 85.19% — and the authors themselves noted that this was because people do not always open their mouths fully when yawning naturally. This is an honest observation that most papers tend to gloss over, and it is something we kept in mind while designing our own yawn detection logic.

What the literature as a whole suggests is that no single indicator is enough on its own. Eye closure, yawning, head pose, and phone detection each capture a different dimension of driver inattention. Systems that combine them consistently outperform those that do not. Save Vision is designed with this lesson at its core.

IV. COMPARATIVE ANALYSIS

Before presenting our proposed system, it is worth looking at how Save Vision compares to the existing works reviewed in the literature. This helps clarify not just what others have done, but where their approaches fall short and what specific gaps our system is trying to fill.The table below summarises the key differences:

Feature	Soukupova [3]	Gupta [4]	Kamboj [5]	Zhao [6]	Morocho [7]	Save Vision
Eye closure detection	✓	✓	X	X	✓	✓
Yawn detection	X	✓	✓	X	✓	✓
Head pose detection	X	✓	X	✓	✓	✓
Phone detection	X	X	X	X	X	✓
Alert mechanism	X	✓	✓	X	✓	✓
Software only	✓	X	✓	X	✓	✓
Pre-recorded video	X	X	X	X	X	✓

Looking at this honestly, what stands out is that most existing systems focus on one or two detection indicators and leave the rest uncovered. The work by Soukupová and Čech [3] — which is the most widely cited — only detects eye closure and provides no alert mechanism at all. It was designed as a detection method, not a complete safety system. Similarly, Zhao et al. [6] focused purely on head pose without addressing any other fatigue signal, and their setup required GPU-level computing which is not practical for everyday use.

The system by Morocho et al. [7] is the closest to what we are building — it combines multiple detection streams and runs in real time on standard hardware. However it does not include phone detection, which is one of the most commonly documented causes of driver distraction in real-world accidents. Save Vision adds this as a fourth detection stream.
Another distinction worth noting is input flexibility. All the reviewed systems work only with live video. Save Vision is designed to also accept pre-recorded video, which is important for testing and validation when live driving scenarios are not available — something particularly relevant at our current stage of development.

The honest limitation of Save Vision at this stage is that unlike some of the reviewed systems which have been fully tested and report accuracy figures, we are yet to complete implementation and gather performance data.

What the comparative analysis does confirm, however, is that the approach we have chosen — multi-indicator, software-only, webcam-based — is both technically grounded and practically justified.

V. PROPOSED SYSTEM

The basic idea behind Save Vision is simple — if a camera can see the driver's face, it can tell a lot about whether that driver is alert or not. The system we are proposing works on exactly this principle. A webcam continuously captures the driver's face, and the software analyses each frame to check for signs of drowsiness, fatigue, distraction, or phone use. When something concerning is detected, the system immediately alerts the driver through an audio alert played through the speakers and an on-screen warning message. No wearable device or additional hardware is needed — the entire system runs as a software application on a standard laptop or desktop computer, accepting input from either a live webcam or a pre-recorded video file.

To keep things organised, we have divided the system into three parts: the sensing layer which captures the video, the processing layer which analyses it, and the alert layer which responds when a problem is found.

1. Sensing Layer

This is the entry point of the system. Video input is captured either from a live webcam feed or from a pre-recorded driving video, giving the system flexibility across both real-world and testing scenarios. The webcam is expected to be positioned to provide a clear frontal view of the driver's face. Frames from this input are continuously passed into the processing layer for analysis.

2. Processing Layer

This is where most of the work happens. Every frame goes through a fixed pipeline — first the face is located, then key landmark points on the face are identified, and then four separate detection modules analyse those points simultaneously.

2.1 Face Detection

Before anything else can happen, the system needs to find the driver's face within the frame. This step isolates the facial region so that subsequent stages only process the relevant portion of the image, keeping the pipeline efficient and focused.

2.2 Facial Landmark Extraction

Once the face is found, specific points are mapped across it — around the eyes, along the lips, at the nose tip, and along the jawline. These points act as the geometric foundation for all four detection modules that follow.

2.3 Eye Closure Detection

From the landmark points around each eye, the Eye Aspect Ratio (EAR) is calculated — a ratio that stays relatively stable when the eye is open and drops sharply when it closes. If the EAR remains below a set threshold for too many consecutive frames, the system flags a drowsiness event. The EAR method was originally proposed by Soukupová and Čech [3] and has since been widely validated in real-time drowsiness detection research.

2.4 Yawn Detection

A similar approach is applied to the mouth region. The Mouth Aspect Ratio (MAR) is computed from lip landmark points. When a driver yawns, the mouth opens wide and the MAR rises noticeably above its normal range. Since yawning is one of the earliest physical signs of fatigue — often appearing well before prolonged eye closure begins — detecting it allows the system to intervene earlier.

2.5 Head Pose Estimation

This module tracks the direction the driver's head is facing by estimating yaw and pitch angles from facial landmarks. A driver paying attention to the road keeps their head roughly centred and level. If the head turns too far sideways or tilts downward beyond a defined angular limit, the system reads this as distraction and raises an alert.

International Journal of Science,
Engineering and Technology
An Open Access Journal

2.6 Phone Detection

The fourth module scans the driver's hand region within the frame for the presence of a mobile phone. To avoid false alarms from brief or accidental detections, the system only raises an alert when a phone is identified consistently across several consecutive frames.

3. Decision Unit

The outputs from all four modules feed into a central decision unit that evaluates the overall driver state moment to moment. It is designed to respond to sustained patterns rather than isolated anomalies — a single dropped EAR value or a brief head turn is not enough to trigger a warning. When a module reports a prolonged abnormal reading, or when multiple modules flag a problem simultaneously, the decision unit escalates the response.

4. Alert Layer

When the decision unit determines that an alert is needed, two things happen. An audio alert is played through the system speakers — a sound that is immediate and difficult to ignore. Simultaneously, an on-screen warning message appears identifying what was detected, whether it was eye closure, yawning, head deviation, or phone use. Together these two channels ensure the alert reaches the driver's attention regardless of what they are focused on at that moment.

VI. METHODOLOGY

Getting the system to actually work the way we described in the proposed system section requires a clear understanding of the technical steps involved at each stage. This section explains the logic and mathematics behind each of the four detection modules — not just what they do, but how they do it.

1. Video Input and Preprocessing

The system accepts video either from a live webcam feed or from a pre-recorded video file. Before any detection begins, each frame is resized and converted from its original colour format to grayscale where needed, reducing the amount of data the system has to process without losing the facial information required for landmark detection. This preprocessing step is small but important — it directly affects how fast the pipeline runs on a standard machine.

2. Face Detection and Landmark Extraction

Every frame first goes through a face detector that locates the driver's face and draws a bounding region around it. Only this region is passed to the landmark extractor, which maps a set of key points across the face — around the eyes, along the lips, at the nose tip, and along the jawline. These points are the foundation of everything that follows. If landmark extraction is inaccurate even slightly, it affects all four detection modules downstream. This is why the choice of landmark detection model matters — it needs to be both accurate and fast enough to keep up with real-time video.

3. Eye Closure Detection using EAR

The Eye Aspect Ratio was introduced by Soukupová and Čech [3] and has become the standard method for real-time eye closure detection. It uses six landmark points around each eye — two horizontal corner points and four vertical points — to compute a single ratio that describes how open the eye is at any given moment.

The formula is:

$$\mathbf{EAR = (||p2 - p6|| + ||p3 - p5||) / (2 \times ||p1 - p4||)}$$

Where p1 to p6 are the six eye landmark points. The numerator adds up the two vertical distances across the eye opening, and the denominator uses the horizontal distance to normalise the ratio so it stays consistent regardless of how close or far the face is from the camera. When the eye is fully open, the EAR stays around a stable value. When the eye closes — whether during a blink or due to drowsiness — the EAR drops sharply toward zero.

The key distinction between a normal blink and drowsiness is duration. A blink lasts a fraction of a second, so the EAR recovers almost immediately. Drowsiness causes the EAR to remain below a set threshold across several

ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

consecutive frames. Research has shown that an EAR threshold of around 0.25 works well across most individuals, though some studies have found values as low as 0.18 to be more accurate in certain conditions [8]. In our system, when the EAR stays below the threshold for a defined number of consecutive frames, a drowsiness alert is triggered.

4. Yawn Detection using MAR

Yawn detection follows a very similar geometric approach. The Mouth Aspect Ratio is calculated from landmark points along the upper and lower lip — specifically the vertical distance between the top and bottom lip divided by the horizontal width of the mouth [4].

MAR = vertical mouth distance / horizontal mouth distance

When the mouth is closed or only slightly open, the MAR stays low. When the driver yawns and the mouth opens wide, the MAR rises noticeably. A threshold is applied in the same way as the EAR — if the MAR exceeds a set value for long enough, the system treats it as a yawn rather than a brief mouth movement such as speaking or swallowing. The challenge with yawn detection, as noted in the literature [7], is that people yawn differently — some open their mouths very wide, others much less so. This is something we are keeping in mind for the threshold calibration stage of the implementation.

The internal decision logic applied within each detection module on a per-frame basis is illustrated in Figure 1.

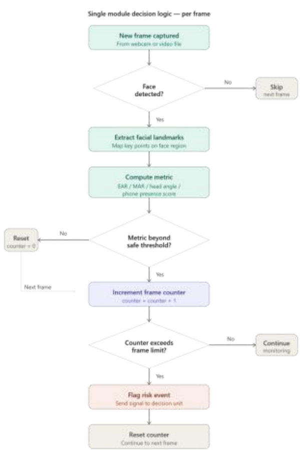

Figure 1: Per-frame decision logic within a single detection module

5. Head Pose Estimation

Head pose estimation works differently from the EAR and MAR approaches — instead of a simple ratio, it involves estimating the three-dimensional orientation of the driver's head from a two-dimensional image. This is done by matching the detected 2D facial landmarks to a known 3D reference model of the human face and then solving for the rotation that best explains the difference between them [6].

The three angles estimated are:

- **Yaw** — left and right rotation of the head
- **Pitch** — up and down tilt of the head
- **Roll** — sideways tilt of the head

For driver distraction, yaw and pitch are the most relevant. A driver looking straight at the road maintains a near-zero yaw and a level pitch. If the head turns sideways beyond a defined angular threshold, or tilts downward significantly — which often happens when a driver looks at their phone or at something inside the car — the system flags a distraction event. The Perspective-n-Point algorithm is the standard method used to solve this 3D orientation problem from 2D landmark coordinates, and it works well when the face is reasonably visible and not too far from the camera [6].

6. Phone Detection

Phone detection is approached differently from the landmark-based modules. Rather than computing geometric ratios, the system scans the region around the driver's hands within the frame and classifies whether a mobile phone is present. This is done through image-based object classification. To keep the system from raising false alarms every time the driver momentarily moves their hand, phone detection is only confirmed when the object is consistently identified across multiple consecutive frames. A single frame detection is ignored — the alert only fires when the pattern persists.

ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

7. Alert Logic

Once the four modules have been evaluated for each frame, the decision unit looks at their combined output and decides whether to alert. The system is designed to be tolerant of momentary anomalies — a single low EAR frame, a brief head turn, or one frame where a phone-like object appears is not enough to trigger anything. It is the persistence of these signals across time that the system responds to. When an alert is warranted, an audio signal plays through the system speakers and an on-screen warning appears simultaneously, telling the driver specifically what was detected.

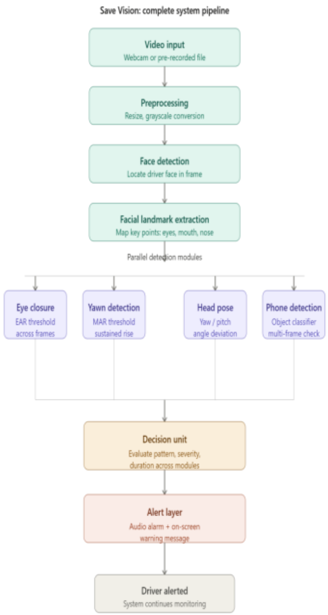

Figure 2: Save Vision — complete system detection pipeline

VII. APPLICATIONS AND BENEFITS

The problem Save Vision addresses is not limited to one type of driver or road. Fatigue and distraction affect every category of road user, and the system's software-based design makes it deployable across multiple real-world contexts without significant cost or infrastructure.

1. Personal Vehicles

Everyday commuters represent the largest group exposed to drowsy and distracted driving. According to the WHO, drivers using a mobile phone while driving are approximately four times more likely to be involved in a crash [10]. Yet for most personal vehicle owners in India, no monitoring technology exists in their vehicles at all. Save Vision requires nothing beyond a webcam and a computing device, making it accessible without any vehicle modification or significant expenditure.

2. Commercial Fleet and Long-Haul Transport

Trucks account for approximately 12.3% of road accidents and 15.8% of road fatalities in India [2]. A study of Indian truck drivers found that around 23% regularly battle sleep deprivation and over 53% report chronic fatigue-related health issues [12]. Save Vision can be deployed across entire fleets affordably, alerting drivers in real time while providing fleet managers with detection logs to identify high-risk drivers and routes.

3. Public Transport

Fatigue is estimated to be a contributing factor in up to 20% of road collisions, and is disproportionately linked to high-severity crashes since a sleeping driver cannot brake or swerve [13]. Bus operators managing overnight routes face this risk acutely. Save Vision operates passively throughout an entire shift without requiring any driver interaction, making compliance automatic.

4. Key Benefits

Beyond specific contexts, Save Vision offers four broad advantages. First, cost accessibility — a software-only approach removes the hardware barrier that keeps existing solutions out of reach for most Indian drivers. Second, non-intrusiveness — the driver changes nothing about their behaviour; the system observes silently. Third, flexibility — support for both live and pre-recorded video input makes the system useful in operational and research environments alike. Fourth, robustness — four independent detection streams ensure the system remains functional even when individual modules face limitations, reducing missed detections at critical moments.

VIII. CHALLENGES

Every system has limitations, and being honest about them is an important part of any research work. Save Vision is no exception. While the multi-indicator approach makes the system more robust than single-modality alternatives, several real-world challenges remain that affect its reliability and scope.

1. Lighting Conditions

The most immediate challenge is lighting. Camera-based facial landmark detection performs well under adequate and consistent lighting but degrades significantly in low-light environments, harsh shadows, or direct sunlight glare. Since Save Vision relies entirely on visual input, any condition that reduces image quality directly affects detection accuracy. Research has consistently identified variable lighting as one of the primary failure conditions for computer vision systems operating in real-world environments [4]. Night driving — which is also the highest-risk period for fatigue-related accidents — is therefore the condition where the system is most needed but also most challenged.

2. Facial Occlusion

Drivers wearing sunglasses, face masks, or hats that cast shadows over the eye region create significant difficulties for eye-closure and landmark-based detection. The EAR method specifically depends on clear visibility of the eyelid contour — even partial occlusion of the eye region can cause the computed ratio to behave unpredictably, leading to either missed drowsiness events or false positives [3]. This is a known limitation of landmark-based approaches and one that has not been fully resolved in the existing literature.

3. Individual Variation

Different people have naturally different facial geometries, eye shapes, and resting expressions. A fixed EAR threshold that works well for one driver may be too sensitive or too lenient for another. Similarly, not all drivers yawn in the same way — some open their mouths very wide, others much less so — which makes a fixed MAR threshold unreliable across diverse users [7]. Personalised threshold calibration per driver would address this, but adds complexity to the deployment process.

4. Camera Placement and Stability

The system assumes a stable, consistent frontal view of the driver's face. In practice, camera placement inside a vehicle is not always ideal — vibrations from the road, varying seat positions across different drivers, and suboptimal mounting angles can all affect the quality and consistency of the captured frames. A frame where the driver's face appears at an extreme angle or partially outside the camera's field of view will either fail face detection entirely or produce unreliable landmark coordinates.

5. Computational Load

Running four detection modules simultaneously in real time places a meaningful demand on the processing device. On a standard laptop this is manageable, but performance can degrade on older or lower-specification machines, increasing processing latency and reducing the effective frame rate. Since timely detection is central to the system's safety function, any significant latency introduced by hardware limitations directly reduces its effectiveness.

6. Phone Detection Reliability

Among the four modules, phone detection is the most susceptible to false positives and false negatives. Objects that resemble a phone in shape or size — a wallet, a book, or a dark-coloured object held in a similar position — can trigger incorrect detections. Conversely, phones held at unusual angles or partially hidden may go undetected. While the multi-frame confirmation approach reduces false alerts, it does not eliminate them entirely.

IX. FUTURE WORK

The current version of Save Vision establishes a working foundation for software-based, multi-indicator driver monitoring. However, several directions exist where the system can be meaningfully extended and improved in future iterations.

1. Adaptive Thresholding

One of the most impactful improvements would be moving from fixed EAR and MAR thresholds to adaptive, driver-specific calibration. In this approach, the system spends the first few minutes of each session observing the driver's baseline facial metrics — their natural blink rate, resting eye openness, and typical head position — and sets thresholds relative to that individual rather than using a universal value. This would directly address the individual variation challenge identified in the previous section and is likely to reduce both false positives and missed detections significantly.

2. Night-Time and Low-Light Performance

Improving performance under poor lighting conditions is a critical next step. One approach is integrating an infrared camera module, which captures facial detail independent of ambient light. Infrared-based facial landmark detection has been used effectively in several driver monitoring studies and would extend Save Vision's operational reliability into the nighttime driving scenarios where fatigue risk is highest [6].

3. Incorporating Additional Fatigue Indicators

Future versions of the system could incorporate additional behavioural signals beyond the current four. Head nodding — where the driver's head slowly drops forward and snaps back — is a well-documented indicator of microsleep and could be detected through temporal analysis of pitch angle changes across consecutive frames. Blink rate over time is another signal; research has shown that blink frequency decreases progressively as fatigue deepens [8], making it a useful supplementary metric alongside the binary open/closed assessment provided by EAR alone.

4. Driver Emotion and Stress Detection

Emotional state is increasingly recognised as a factor in driving safety — a highly stressed or agitated driver is also an impaired one. Future work could extend Save Vision to incorporate facial action unit analysis for basic emotion recognition, building on the approach taken by Morocho et al. [7] who demonstrated real-time emotion detection alongside fatigue monitoring using a MobileNetV2-based classifier. This would give the system a more holistic picture of driver wellness beyond just physical drowsiness.

5. Mobile and Edge Deployment

Currently the system runs on a standard laptop or desktop environment. A natural progression would be porting it to a mobile application or a lightweight edge computing device that can be mounted permanently inside the vehicle without depending on a separate computer. This would make the system significantly more practical for everyday deployment, particularly in commercial fleet contexts where a dedicated computing device per vehicle is not feasible.

6. Dataset Collection and Benchmarking

Since Save Vision is currently in the development stage, a dedicated evaluation dataset collected under Indian driving conditions — varying road types, cabin lighting, and driver demographics — would provide a more relevant and reliable benchmark than internationally collected datasets. Future work should prioritise building such a dataset and using it to formally evaluate and publish the system's detection accuracy, response latency, and false positive rates.

X. CONCLUSION

Road accidents caused by driver fatigue and distraction continue to claim an enormous number of lives every year, both in India and globally. What makes this particularly difficult is that unlike mechanical failure or external road hazards, these are internal, invisible states that develop gradually within the driver — often without the driver themselves realising how impaired they have become. The need for a system that monitors this in real time, without burdening the driver with wearable devices or complex setup, is clear.

Save Vision: Smart Drive Monitor is developed as a response to this need. By combining four independent detection streams — eye closure through EAR computation, yawning through MAR analysis, head pose deviation, and mobile phone usage detection — into a unified software pipeline, the system monitors the driver's behavioural state continuously and alerts them through audio and on-screen warnings the moment a risk condition is identified. The entire system runs on a standard computing device with a webcam, requiring no specialised hardware and no vehicle modification, which makes it accessible to a far wider range of users than existing commercial solutions.

The work reviewed in the literature confirmed that the technical approach we have adopted is well grounded. EAR-based drowsiness detection, MAR-based yawn detection, and head-pose-based distraction monitoring are each individually validated across multiple published studies. What Save Vision adds is the combination of all four into a single, coherent, software-only system that also includes phone detection — a stream absent from most comparable work in the literature.
Honest acknowledgement is also important here. The system has real limitations — lighting sensitivity, occlusion from sunglasses, individual variation in facial geometry, and the computational demands of running four modules simultaneously are all challenges that require further work. The future directions outlined in the previous section — adaptive thresholding, infrared support, edge deployment, and formal benchmarking under Indian conditions — represent the natural next steps for turning this foundation into a fully reliable safety tool.

Road safety is ultimately a shared responsibility. Technology alone cannot eliminate accidents, but a system that catches a driver nodding off at 2 AM on a highway, or notices that their head has turned away from the road for too long, and responds in time — that system saves lives. Save Vision is built toward that goal, and we believe the foundation laid in this work is a meaningful step in that direction.

REFERENCES

1. World Health Organization, Mobile Phone Use: A Growing Problem of Driver Distraction, Geneva: WHO, 2011. [Online]. Available: https://www.who.int/publications/i/item/mobile-phone-use-a-growing-problem-of-driver-distraction
2. A. Khan, S. Phulre, and V. Richariya, "Integrated FLAG based Energy Efficient Routing with Drop Minimization in MANET," International Journal of Computer Applications, vol. 86, no. 17, 2014.
3. S. Phulre, P. Gautam, and S. K. Mishra, "Implementation of trusted multitier method for intrusion detection in mobile ad hoc networks with DSR algorithm," in 2014 Science and Information Conference, London, UK, 2014, pp. 666-673.
4. N. Sahu, S. Phulre, and V. Richhariya, "A Review: Enhancement in the Robustness and Capacity of Watermarking Technique using DWT, DCT and LSB," International Journal of Computer Applications, vol. 118, no. 1, pp. 1-6, 2015.
5. T. Soukupová and J. Čech, "Real-Time Eye Blink Detection using Facial Landmarks," in 21st Computer Vision Winter Workshop, Rimske Toplice, Slovenia, Feb. 2016. [Online]. Available: https://cmp.felk.cvut.cz/~cechj/EAR/
6. SaveLIFE Foundation, Distracted Driving in India: A Study on Mobile Phone Usage, Pattern and Behaviour, New Delhi: SaveLIFE Foundation, 2017.
7. Castrol India, Indian Truck Driver Health and Lifestyle Study, Mumbai: Castrol India, 2018. Cited in: Quartz India.
8. A. Chourey, S. Phulre, and S. Mishra, "A survey paper on employee attrition prediction using machine learning techniques," Journal of Interdisciplinary Cycle Research, vol. 11, no. 12, pp. 199-202, 2019.
9. Y. Zhao et al., "Driver Distraction Detection Method Based on Continuous Head Pose

Estimation," Computational Intelligence and Neuroscience, 2020. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1155/2020/9606908

10. A. K. Phulre, M. Kamble, and S. Phulre, "Content Management Systems hacking probabilities for Admin Access with Google Dorking and database code injection for web content security," in 2nd International Conference on Data, Engineering and Applications (IDEA), Bhopal, India, 2020, pp. 1-5.
11. P. S. Tomar, S. Phulre, and M. Shrivastava, "Impact of Varying Physical Links on QoS of Virtual Topology Management in WDM Networks," in 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT), Bhopal, India, 2021, pp. 74-80.
12. C. Dewi, R. C. Chen, and H. Liu, "Eye Aspect Ratio for Real-Time Drowsiness Detection to Improve Driver Safety," Electronics, vol. 11, no. 19, p. 3183, 2022.
13. S. Zhang and M. Abdel-Aty, "Drivers' Visual Distraction Detection Using Facial Landmarks and Head Pose," Transportation Research Record, 2022.
14. World Health Organization, Global Status Report on Road Safety 2023, Geneva: World Health Organization, 2023.
15. A. Gupta et al., "Real-Time Machine Learning-Based Driver Drowsiness Detection Using Visual Features," Journal of Imaging, vol. 9, no. 5, p. 91, 2023.
16. S. Kamboj et al., "Detection of Drowsiness among Drivers Using Novel Deep Convolutional Neural Network Model," PMC/NCBI, 2023. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10650052/
17. Ministry of Road Transport and Highways, Government of India, Road Accidents in India 2023, New Delhi: MoRTH, 2024.
18. Royal Society for the Prevention of Accidents (RoSPA), Driver Fatigue and Road Accidents: Road Safety Factsheet, Birmingham: RoSPA, March 2024.
19. D. Morocho-Cayamcela et al., "Driver Monitoring System Using Computer Vision for Real-Time Detection of Fatigue, Distraction and Emotion via Facial Landmarks and Deep Learning," Sensors, vol. 26, no. 3, p. 889, 2026.