JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

# DRIVE SAFE EMOTIVE TRANSPORT INITIATIVE USING AI: DROWSINESS, YAWN AND PHONE USAGE DETECTION

**[1]Muhammed Shabin Sadique,[2]Sanu Eldose,[3] Shreyas Shasi A, [4]Surya Dev TK, [5]Dr. Jithendra P R Nayak**

[1]Final year B.E. Student, [2] Final year B.E. Student, [3]Final year B.E. Student, [4]Final year B.E. Student, [5] Associate Professor, CSE
[1]Department of Computer Science & Engineering,
[1]Srinivas Institute of Technology, Mangaluru, India

***Abstract*:** Modern transportation systems face safety risks due to driver fatigue, distraction, and emotional distress. The Drive Safe Emotive Transport Initiative harnesses the power of Artificial Intelligence to monitor and respond to drivers' emotional and cognitive states in real time. Using facial recognition, speech analysis, and physiological data, the system detects signs of stress, drowsiness, and agitation. When triggered, it provides appropriate feedback or interventions such as alerts, calming voice prompts, or auto-slowdown assistance to ensure safer journeys. The AI engine continually learns and adapts to individual driving behaviours, enhancing accuracy over time. Prioritizing both safety and mental well-being, this hands-free, intelligent assistant promotes responsible driving and minimizes road accidents.

**Keywords: Drowsiness detection, Yawn detection, , Facial Recognition, Phone detection, YOLO Technology**

## I. INTRODUCTION

Road safety is a critical concern in modern transportation, yet human factors like fatigue, distraction, and mobile phone usage continue to cause a high number of accidents. Drivers often miss the warning signs of drowsiness or reduced attention, while manual monitoring systems fall short in providing timely intervention. These risks are especially dangerous during long or late-night journeys where constant vigilance is required. Existing safety technologies often lack intelligent, real-time detection of such behaviours.

This paper introduces the *Drive Safe Emotive Transport Initiative*, an AI-powered system designed to monitor and respond to risky driving behaviours. Using computer vision and deep learning, the system detects signs of drowsiness, yawning, and mobile phone usage by analyzing facial features, eye movement, and hand gestures. Once detected, it delivers instant alerts to re-engage driver attention and promote road safety. By combining real-time behavioural analysis with adaptive feedback, this solution aims to reduce accident rates and support safer, more attentive driving for everyone on the road.

## II. LITERATURE REVIEW

Existing research has made notable strides in developing real-time driver monitoring systems aimed at enhancing road safety through AI-based detection of risky behaviors. Studies by Abtahi et al. (2014) and Singh et al. (2019) explored facial landmark tracking and eye aspect ratio (EAR) calculations for detecting drowsiness, while others like Das et al. (2020) implemented CNN models trained on yawning sequences and prolonged eyelid closures. However, many of these systems require infrared (IR) cameras or high-definition facial sensors, which may not be suitable for all vehicles or environments. In addition, phone usage detection, although crucial in distracted driving prevention, has often been treated as an isolated module rather than integrated into a comprehensive system. Solutions reviewed in prior literature typically focused on either drowsiness or distraction separately, limiting their capability for holistic real-time monitoring.

Several Research into multimodal behavioral recognition systems shows promise. For instance, Shaikh et al. (2021) combined video and audio analysis using OpenCV and TensorFlow to recognize drowsy states with moderate accuracy. Others, like Alotaibi et al. (2018), utilized electroencephalogram (EEG) signals in laboratory settings for more accurate fatigue detection. However, such invasive and costly setups are impractical for everyday transport applications. As discussed in Section 2.3 of the source report, yawning detection through mouth aspect ratio (MAR) measurements and temporal frame analysis has become an accepted method, yet issues such as occlusion from sunglasses or variable lighting remain unresolved.

The growing integration of AI with edge computing also informs this domain. Systems that apply MobileNetV2 or YOLOv5 for real-time object (phone) and pose detection on embedded hardware (e.g., Raspberry Pi or Jetson Nano) are increasingly seen in low-cost safety applications. As cited in Section 2.5 of the source report, real-time inference through optimized deep learning models, aided by Haar Cascades and MediaPipe libraries, enables detection of fine-grained facial expressions and hand positions with relatively low power consumption. These approaches highlight the movement toward lightweight, deployable, and responsive systems capable of operating efficiently in non-laboratory, dynamic driving conditions.

Some existing solutions, like those presented by Roy et al. (2021), employed multi-camera systems and accelerometer data to enhance accuracy, though at the cost of system complexity. In contrast, work by Anitha et al. (2022) demonstrated single-camera solutions with facial feature extraction combined with a heuristic approach for distraction detection. Despite their efficiency, such systems often lack robustness when it comes to distinguishing between deliberate driver actions (e.g., checking rear mirrors) and risky behavior like mobile phone use.

While contribute significantly to road safety through intelligent monitoring systems, the "Drive Safe Emotive Transport Initiative" enhances this foundation by delivering a real-time, unified architecture. It utilizes lightweight CNN models for eye and mouth state detection, paired with advanced object detection techniques for identifying phone usage. The system is designed for cross-platform compatibility and optimized for in-vehicle cameras without specialized hardware. Built with robust error handling and dynamic thresholding, it adapts to individual driver behavior patterns and environmental variability. Moreover, by implementing frame-based behavior fusion and real-time auditory alerts, the system promotes proactive driver intervention, setting itself apart from traditional monitoring setups that merely log events. Ultimately, this integrated AI-based solution addresses key shortcomings of earlier works by providing a holistic, non-intrusive, and scalable driver assistance framework.

## III. PROPOSED SYSTEM: REAL-TIME E-VOICE ARCHITECTURE

The " *Drive Safe Emotive Transport Initiative* adopts a multi-layered architecture to detect unsafe driver behaviors in real time, specifically focusing on drowsiness, yawning, and mobile phone usage. The core components include:

- **Driver Monitoring Device (Camera/Client):** A standard webcam or dashboard-mounted camera captures real-time video of the driver's face and upper body. This acts as the primary sensor input, requiring no additional hardware complexity, and is easily integrable into existing vehicle setups..
- **Frontend:** Built using HTML, Tailwind CSS, and JavaScript, this interface runs inside a lightweight browser or application window. It handles the video stream and communicates detection outcomes. The design is deliberately minimal to avoid visual distraction and prioritizes auditory feedback to alert the driver using text-to-speech output.
- **Backend:** A Flask-based Python server acts as the processing hub. It receives the video feed, performs real-time frame analysis using AI models, and returns detection results to the frontend for alert handling. Detection Modules: *Drowsiness & Yawning Detection:* Powered by Media Pipe and OpenCV, these modules analyze facial landmarks (eye aspect ratio, mouth shape, and blink duration) to detect fatigue-related behaviour . *Phone Usage Detection:* Utilizes YOLOv3 to identify handheld devices near the face or within the field of view, signaling distraction .

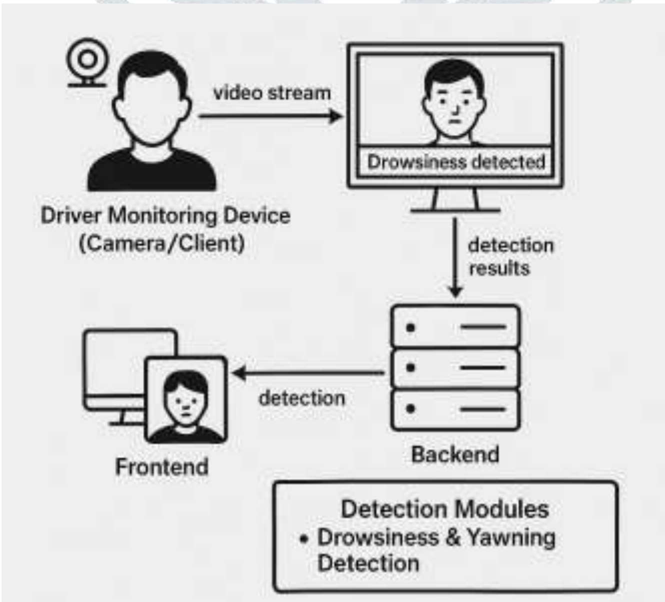


**Fig.1: Architecture of a Voice-Controlled Email Application**

Figure 1 illustrates depicts the high-level system architecture for detecting driver fatigue and distraction. At the top, the User Layer represents the driver, monitored via a camera that continuously feeds video to the system. This input flows into the Presentation Layer (Frontend), responsible for displaying status and delivering real-time auditory feedback. The captured video is forwarded to the Application Layer (Backend Server) using asynchronous HTTP or WebSocket protocols.

The backend, developed with Flask and Python, performs real-time frame processing using MediaPipe for facial analysis and YOLOv3 for object detection. These AI models operate on extracted features like eye state, mouth openness, and object presence to identify drowsy states, yawning, or phone use..

The final layer, the Detection and Alert Layer, uses pyttsx3 to deliver audible warnings such as "Please stay alert" or "Avoid phone use while driving," allowing the system to respond without distracting the driver. This modular design ensures each layer is independently testable and upgradable, improving system maintainability and deployment flexibility in real-world driving environments.

## IV. IMPLEMENTATION DETAILS AND TECHNOLOGIES

The implementation of the *Drive Safe Emotive Transport Initiative* relies on a strategic integration of computer vision, deep learning, and real-time monitoring technologies to detect and respond to unsafe driver behaviors. At the core of the detection mechanism is the use of MediaPipe, a lightweight, real-time framework for facial landmark detection, which allows continuous tracking of eye aspect ratio (EAR), mouth opening width, and head movement to recognize signs of drowsiness and yawning. OpenCV-Python is utilized to process video frames from the camera feed, extract key facial features, and support gesture analysis.The system employs YOLOv3 (You Only Look Once), a robust object detection algorithm, to identify mobile phone usage by detecting the presence of handheld devices near the driver's face or in their line of sight. YOLOv3 is pretrained on large image datasets and fine-tuned to improve accuracy in detecting phone-related gestures in various lighting and angle conditions. The logic for integrating detection models is written in Python, forming the backend engine for behavioral analysis and alert generation.

The frontend interface is developed using HTML, Tailwind CSS, and JavaScript, designed to be minimal and distraction-free. The client-side logic runs within a lightweight Flask application that initiates and handles the video feed, processes AI detection outputs, and interacts with the alert system. When a risk behavior such as drowsiness, yawning, or mobile phone use is detected, the system triggers appropriate responses using pyttsx3, a Python-based Text-to-Speech (TTS) library, which vocally alerts the driver in real time to refocus their attention on the road.

Implementation efforts focused on fine-tuning detection thresholds for EAR, mouth aspect ratio, and object classification confidence levels to ensure low false positives while maintaining sensitivity to genuine risks. Additional attention was paid to error handling scenarios, such as temporary frame drops, occluded facial features, and camera access issues, enabling the system to remain stable and responsive during extended operation.

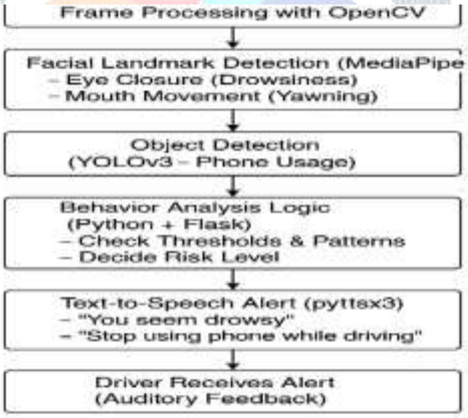


**Fig.2: Voice-Driven Email Composition Process Flow**

Figure 2 outlines the process by which the system identifies and responds to driving risks. It begins with a live camera feed, continuously analyzed by the Python-based backend using Media Pipe and YOLOv3 for facial and object detection. If a behavioural cue such as eye closure, yawning, or mobile phone usage is detected, the event is flagged and processed. The Flask server sends the appropriate alert message to the Text-to-Speech engine, which then delivers an audible warning to the driver. This feedback loop ensures real-time communication between detection modules and the driver, creating an adaptive, safety-focused monitoring system without requiring driver input.This cohesive interaction between video analysis, deep learning models, and real-time alerts showcases the effective application of AI in minimizing accident risks through behavioral monitoring and proactive intervention.

## V. RESULTS AND DISCUSSION

The "*Drive Safe Emotive Transport Initiative"* successfully validated its core capabilities, offering a real-time, AI-driven approach to driver behaviour monitoring. During testing, each detection module delivered reliable performance. The Drowsiness Detection feature utilized facial landmark tracking to monitor eye closure duration and blink frequency, issuing alerts when prolonged eye closures were identified—effectively warning drivers before attention levels dropped critically. Similarly, the Yawning Detection system monitored mouth aspect ratios and jaw movements, triggering auditory alerts upon detecting repeated or extended yawns, indicating potential fatigue. For distracted driving, the Mobile Phone Usage Detection module employed object recognition algorithms to analyze hand positioning and identify the presence of mobile devices near the driver's face or within the hand. When usage was detected during motion, the system promptly activated voice warnings to discourage the behaviour. These safety prompts were delivered through a built-in text-to-speech system, providing clear and timely feedback while minimizing additional distractions

Throughout testing, the system maintained responsive feedback and accurate detections in real-time driving scenarios, demonstrating strong performance in controlled conditions. The user interface, designed with minimal visual elements, supported non-intrusive operation, allowing drivers to focus on the road. Integration of Media Pipe and YOLOv3 technologies ensured precise landmark identification and object tracking. These results affirm the solution's potential for enhancing driver awareness and preventing accidents caused by fatigue or distraction.

## VI. CONCLUSION

The "Drive Safe Emotive Transport Initiative using AI" offers a robust and intelligent approach to enhancing road safety by detecting critical driver behaviors such as drowsiness, yawning, and mobile phone usage in real-time. Leveraging computer vision, deep learning, and real-time monitoring, the system provides a proactive solution to reduce road accidents caused by human error and inattention. Its ability to function effectively in diverse lighting and environmental conditions makes it highly practical for real-world deployment. With a strong emphasis on safety, the initiative underscores the transformative power of AI in transportation. Future developments may include emotion detection for stress levels, integration with vehicle control systems for automated alerts or intervention, multi-modal sensor fusion for increased accuracy, and cloud-based analytics for fleet safety monitoring. This initiative stands as a compelling example of how AI can be harnessed to create emotionally aware, responsive, and intelligent transport systems that prioritize human life and safety on the roads.

## VII. REFERENCES

[1] A. D. Rana, A. Deshpande, A. Bhor, and D. Patil, "Real-Time Driver Drowsiness Detection System Using Eye Aspect Ratio and Yawn Detection," *2022 6th International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, 2022, pp. 1355-1360, doi: 10.1109/ICCMC53470.2022.9754102.Jain, V., et al. (2021). "Voice based E-mail for the Visually Impaired."

[2] M. Abtahi and M. Sarrafzadeh, "A Vision-Based System for Monitoring Driver's Drowsiness and Distraction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 4, pp. 2017–2029, 2015, doi: 10.1109/TITS.2014.2374591."

[3] M. A. Khan, A. Azam, and H. Abbas, "AI-Based Smartphone Usage Detection While Driving Using Deep Learning," *Journal of Advanced Transportation*, vol. 2020, Article ID 8896352, 2020, doi: 10.1155/2020/8896352."

[4] T. R. Doshi and K. R. Yadav, "Real Time Driver Alertness Monitoring System Using CNN," *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, Coimbatore, India, 2021, pp. 710–715, doi: 10.1109/ICAIS50930.2021.9395894."

S. Jaiswal, R. Garg, and R. Ranjan, "Real-Time Drowsiness Detection Using Eye Blink Monitoring," *2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, Greater Noida, India, 2021, pp. 927–932, doi: 10.1109/ICACCCN52233.2021.9538829.