International Journal of
AI ELECTRICAL CIVIL AND MECHANICAL ENGINEERING
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0434 www.zesterapublications.com Original Research Paper

INTELLIGENT COMPUTER VISION-BASED DRIVER AND ROAD SAFETY MONITORING SYSTEM USING DEEP LEARNING

Darisavarahalugari Venkata Sushma Charitha [1], Hemalatha T [2], Dr D Vivekananda Reddy[3]

[1]*M.Tech Scholar, Department of CSE, Sri Venkateswara University College of Engineering ,Tirupati.*
E-mail: dv.charitha@gmail.com,ORC-ID: https://orcid.org/0009-0001-0759-3254
[2]*Academic Consultant, Department of CSE, Sri Venkateswara University College of Engineering ,Tirupati.*
E-mail: thirupathuruhemalatha@gmail.com,
[3]*Professor, Department of CSE, Sri Venkateswara University College of Engineering ,Tirupati.*
E-mail: svuvivek@gmail.com, ORC-ID: https://orcid.org/0000-0001-9497-6677

ABSTRACT

Driver and road safety require continuous monitoring of both the driver and the surrounding road environment. This work presents an Intelligent Computer Vision-Based Driver and Road Safety Monitoring System Using Deep Learning that combines different safety-monitoring functions into a single framework. The system uses MediaPipe Face Mesh, facial landmark analysis, deep learning models, YOLOP, YOLO11, and image processing techniques to monitor the driver and road conditions. The driver monitoring module analyzes facial features to identify Normal, Drowsy, Impaired, and Yawning states, along with driver absence and abnormal head movements using eye, mouth, and head-pose analysis. The system also detects seatbelt usage and potholes, while YOLOP and YOLO11 are used for road-scene and road-object detection. CLAHE is used to improve image quality under different lighting conditions. The outputs from these modules are combined using a Unified Processor and Safety Decision Engine to determine the overall safety state and generate an audible alert when an unsafe condition is detected. The system produces annotated output frames and buzzer-synchronized H.264 video, while safety events and session details are stored in a SQLite database. A Streamlit-based analytics dashboard is provided to visualize safety events, session records, detection results, and system outputs. The proposed system provides an integrated approach to driver and road safety monitoring by combining computer vision, deep learning, safety decision-making, real-time alerts, data logging, and analytics in a single framework.

Keywords *— Driver Monitoring, Road Safety, Computer Vision, Deep Learning, Facial Landmark Analysis.*

1. INTRODUCTION:

Vehicle safety depends on several factors, including the condition of the driver, proper use of the seatbelt, and the surrounding road environment. Driver drowsiness, yawning, abnormal head movement, driver absence, improper seatbelt usage, road objects, and potholes can create unsafe driving situations. Recent developments in Artificial Intelligence (AI) [1], Computer Vision, Deep Learning [2], and image processing provide useful methods for monitoring these safety conditions automatically. However, many existing systems mainly focus on a single safety problem, such as drowsiness detection, seatbelt detection, or road-object detection. Monitoring these conditions separately may not provide a complete view of the overall safety condition during vehicle operation. Therefore, an integrated system that can monitor both the driver and road environment is needed.

International Journal of
AI ELECTRICAL CIVIL AND MECHANICAL ENGINEERING
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0434 www.zesterapublications.com Original Research Paper

This work presents an Intelligent Computer Vision-Based Driver and Road Safety Monitoring System Using Deep Learning, which combines multiple safety-monitoring functions within a unified framework. The proposed system can process image, video, and live camera input and analyze the driver as well as the surrounding road environment. The driver-monitoring module uses MediaPipe Face Mesh with 468 facial landmarks for facial analysis and head-pose estimation. Facial features are used to calculate parameters such as Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR). These features, along with temporal analysis and head-pose information, are used to identify Normal, Drowsy, Impaired, and Yawning driver states. The system also detects driver absence and abnormal head movements using pitch, yaw, and roll information. CLAHE (Contrast Limited Adaptive Histogram Equalization) is used in the image-processing stage to improve image contrast under different visual conditions. Pasted markdown

In addition to driver monitoring, the system analyzes important road and vehicle safety conditions. A deep-learning-based model is used for seatbelt detection, while YOLOP is used for road-scene perception and YOLO11 is used for road-object detection. A separate deep-learning-based module is used to identify potholes in the road environment. The information obtained from these different modules is passed to a Unified Processor, which coordinates the results and sends them to the Safety Decision Engine. The Safety Decision Engine evaluates the combined safety information and determines the overall safety state. When an unsafe condition is detected, the system generates a corresponding safety reason and activates an audible buzzer alert. Pasted markdown

The proposed system also provides visual and stored evidence of the detected safety conditions. Annotated frames and processed video are generated during system operation, and FFmpeg is used to synchronize the buzzer audio with the processed video to produce an H.264-compatible output video. Safety events, frame-level observations, detection results, head-pose information, session details, and output information are stored in a SQLite database. A Streamlit-based analytics dashboard is used to review session history, safety events, detection results, head-pose information, processed outputs, and system analytics. This allows the system to provide not only real-time monitoring and alerts but also historical information for later analysis. Pasted markdown

The main aim of the proposed system is to bring driver monitoring, seatbelt detection, road perception, pothole detection, safety decision-making, alert generation, video processing, data storage, and analytics into a single framework. By coordinating the outputs of different computer vision and deep-learning modules through the Unified Processor and Safety Decision Engine, the system provides a common safety assessment instead of treating each safety condition independently. The developed framework therefore provides an integrated approach to real-time driver and road safety monitoring using computer vision, deep learning, facial landmark analysis, image enhancement, centralized safety processing, synchronized alerts, persistent data logging, and analytics.

2. LITERATURE REVIEW:

Driver and road safety monitoring has received significant attention with the development of computer vision, machine learning, and deep learning techniques. Existing studies mainly focus on driver monitoring, drowsiness detection, seatbelt detection, road-scene perception, pothole detection, and image enhancement. These studies provide useful methods for identifying different safety conditions during vehicle operation and form the technical basis for developing intelligent vehicle safety-monitoring systems.

Several studies have focused on driver monitoring and drowsiness detection using facial features. Kim et al. [3] proposed a real-time driver monitoring system using facial landmarks for eye-closure detection and head-pose recognition. The system uses facial landmarks to analyze eye closure and head orientation, showing the usefulness of facial landmark-based methods for monitoring driver attention and drowsiness. Phan et al. [4] proposed a driver drowsiness detection and smart alerting system using

International Journal of
AI ELECTRICAL CIVIL AND MECHANICAL ENGINEERING
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0434 www.zesterapublications.com Original Research Paper

deep learning and IoT. The study considered visual indicators such as eye conditions, blinking, yawning, and head movement for identifying drowsy driving conditions and generating alerts. Albasrawi et al. [5] developed a driver drowsiness monitoring system using facial landmark detection and a Convolutional Neural Network for prediction. Their work demonstrated the use of facial information together with deep learning for non-invasive driver monitoring. Keshtkaran et al. [6] investigated the use of facial features from RGB images for assessing alcohol-related driver impairment. These studies show that facial features, eye behavior, mouth movement, and head orientation can provide useful information for identifying different driver conditions.

Seatbelt detection has also been studied using deep-learning-based object detection techniques. Upadhyay et al. [7] proposed a real-time seatbelt detection system using a YOLO-based deep-learning model. The system analyzes images captured inside a vehicle and identifies whether the driver is wearing a seatbelt. Yassin et al. [8] proposed an improved YOLOv5-based method for seatbelt detection in traffic environments and considered image-processing techniques to improve detection under different illumination conditions. These studies demonstrate the usefulness of deep-learning-based object detection for automated seatbelt monitoring. However, the main focus of these approaches is seatbelt recognition rather than combining seatbelt information with driver and road safety conditions.

For road-scene perception, Wu et al. [9] introduced YOLOP, a multi-task deep-learning framework for panoptic driving perception. YOLOP performs traffic-object detection, drivable-area segmentation, and lane detection using a shared feature-extraction architecture. The model was evaluated using the BDD100K driving dataset and demonstrated its ability to support multiple road-perception tasks in real-time driving environments. The study provides an important basis for using deep learning to understand the surrounding road environment.

Pothole detection has been investigated using different deep-learning approaches. Klco et al. [10] studied automated pothole detection using the YOLOv5 neural network. The work demonstrated the use of object detection for identifying road-surface defects automatically. Ruseruka et al. [11] investigated pothole detection and dimension estimation using machine learning and deep learning with in-vehicle technologies. The study showed that vehicle-mounted visual information can be used for road-condition monitoring and pothole identification. Zhong et al. [14] proposed a pothole detection and quantification method using YOLOv8 and point-cloud information. Their approach used object detection to identify potholes and additional information to analyze their geometry and dimensions. These studies show the potential of deep-learning-based methods for identifying road-surface hazards during vehicle operation.

Image quality is also important for computer vision systems operating under different lighting conditions. Chen et al. [12] investigated the use of Contrast Limited Adaptive Histogram Equalization (CLAHE) with YOLO-based models for road-marking recognition at night. The study showed that image enhancement can improve the visual information available for recognition under challenging illumination conditions. This work provides a basis for using CLAHE as an image-processing technique in safety-monitoring systems where input images may have different lighting conditions.

Some studies have considered driver behavior and the surrounding road environment together. The Safe Smart Drive system [13] presented a real-time approach that combines traffic-environment detection and driver-behavior monitoring using machine learning and deep learning. The system includes perception, detection, classification, and alert-generation stages, demonstrating the importance of monitoring both the driver and the surrounding traffic environment. Visconti et al. [15] presented a review of driver-monitoring systems and on-board vehicle technologies in the Internet of Vehicles and intelligent transportation environment. The review covers driver drowsiness, impairment, driving behavior, vehicle conditions, and road-related information and highlights the increasing use of artificial intelligence, computer vision, sensors, and intelligent transportation technologies for vehicle safety.

International Journal of
AI ELECTRICAL CIVIL AND MECHANICAL ENGINEERING
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0434 www.zesterapublications.com Original Research Paper

Overall, the reviewed studies demonstrate the use of facial landmark analysis, deep learning, object detection, road-scene perception, image enhancement, and intelligent monitoring techniques for different vehicle and road safety applications. Driver-monitoring research provides methods for analyzing eye behavior, facial features, yawning, drowsiness, impairment, and head movement [3]–[6]. Seatbelt detection studies demonstrate the application of YOLO-based models for monitoring seatbelt compliance [7], [8]. Road-perception and pothole studies provide methods for detecting road objects, understanding road scenes, and identifying road-surface hazards [9]–[11], [14]. Image enhancement using CLAHE has also been investigated to improve visual information under difficult lighting conditions [12]. Studies such as SafeSmartDrive and the Internet of Vehicles review further show the importance of considering both driver behavior and the surrounding road environment in intelligent vehicle safety systems [13], [15]. These studies provide the technical foundation for the different computer vision and deep-learning components used in the proposed driver and road safety monitoring system.

3. SYSTEM METHODOLOGY:

The proposed Intelligent Computer Vision-Based Driver and Road Safety Monitoring System Using Deep Learning is implemented as an integrated computer vision and deep-learning framework for monitoring driver behaviour, vehicle safety conditions, and the surrounding road environment. The system accepts image, video, and webcam inputs and processes the acquired frames through image preprocessing, driver monitoring, seatbelt detection, road-scene perception, road-object detection, pothole detection, unified processing, and safety decision-making. The outputs from these modules are combined by the Unified Processor, and the Safety Decision Engine determines the overall safety condition and generates appropriate alerts. The system also provides synchronized output video, persistent database storage, and a Streamlit-based analytics interface.

3.1 Input Acquisition and Image Preprocessing:

The system supports three types of input: image, video, and webcam. Image inputs are processed individually, whereas video and webcam inputs are processed frame by frame. The input frames undergo basic preprocessing operations such as resizing, noise reduction, and color conversion before being passed to the monitoring modules. Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied to improve image contrast and enhance relevant visual information under varying illumination conditions. The enhanced frames are then used as input to the driver and road monitoring components.

The image-processing flow is:

Input → Preprocessing → CLAHE Enhancement → Enhanced Frame

3.2 Driver Monitoring Module:

The Driver Monitoring Module is implemented using MediaPipe Face Mesh, which provides 468 facial landmarks for facial analysis. These landmarks are used to obtain information from the driver's eyes, mouth, and facial orientation.

Two important geometric features are calculated from the facial landmarks:

- **Eye Aspect Ratio (EAR):** used to analyze eye closure and eye-related behaviour.
- **Mouth Aspect Ratio (MAR):** used to analyze mouth opening and yawning-related behaviour.

The module also estimates the driver's head orientation using pitch, yaw, and roll. These facial and head-pose features are used to identify the driver states Normal, Drowsy, Impaired, and Yawning. In addition, the module performs driver absence detection and identifies abnormal head movement based on head-pose information. Thus, the driver-monitoring module combines facial landmarks, eye behaviour, mouth movement, and head orientation to characterize the driver's condition.

3.3 Seatbelt Detection Module:

International Journal of
AI ELECTRICAL CIVIL AND MECHANICAL ENGINEERING
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0434 www.zesterapublications.com Original Research Paper

The Seatbelt Detection Module uses a deep-learning-based object detection model to identify whether the driver is wearing a seatbelt. The module analyses the driver region and produces the corresponding seatbelt detection result. The detected seatbelt condition is forwarded to the Unified Processor so that seatbelt compliance can be considered together with driver and road safety conditions during the final safety evaluation.

3.4 Road-Scene Perception Using YOLOP:

The YOLOP (You Only Look Once for Panoptic Driving Perception) model is used for road-scene perception. It provides information related to the driving environment, including drivable-area perception and lane-related information, along with traffic-scene understanding. The processed frame is passed to YOLOP to obtain the required road-scene observations, which are subsequently forwarded to the Unified Processor.

3.5 Road-Object Detection Using YOLO11:

YOLO11 is used for detecting relevant objects present in the surrounding road environment. The model processes the enhanced input frames and identifies road objects such as vehicles and other relevant objects present in the scene. The detected objects are used by the safety-processing stage to determine whether a road-related safety condition exists. The resulting road-object information is passed to the Unified Processor for further evaluation.

3.6 Pothole Detection Module:

A separate deep-learning-based pothole detection component is implemented to identify potholes in the road scene. The detected pothole information, including the detection result and associated information, is forwarded to the Unified Processor. This allows road-surface hazards to be considered together with driver-related and other road-related safety conditions during safety evaluation.

3.7 Unified Processor:

The Unified Processor acts as the central coordination component of the proposed system. It receives the outputs generated by the Driver Monitoring Module, Seatbelt Detection Module, YOLOP-based road-scene perception, YOLO11-based road-object detection, and pothole detection module.

The processor organizes these outputs into a common processing flow and combines the observations for centralized safety evaluation. This integration allows multiple safety conditions to be considered together rather than treating each detection independently. The major inputs to the Unified Processor include:

Driver State + Driver Presence + Head Movement + Seatbelt Status + Road Objects + Pothole Information

The combined observations are then passed to the Safety Decision Engine.

3.8 Safety Decision Engine:

The Safety Decision Engine evaluates the outputs received from the Unified Processor using predefined safety rules. The decision process considers driver state, driver presence, abnormal head movement, seatbelt condition, detected road objects, and pothole information.

Based on the combined conditions, the engine determines the overall safety state and identifies the corresponding safety reasons. When an unsafe condition is detected, the system activates the buzzer and displays the relevant safety information on the processed frame. The decision flow is:

Driver Observations + Seatbelt Status + Road Observations + Pothole Information → Unified Processor → Safety Decision Engine → Overall Safety State

3.9 Alert Generation and Buzzer System:

The system provides an audible buzzer alert when an unsafe condition is identified. The alert mechanism is controlled by the safety state generated by the Safety Decision Engine. At the same time, the processed frame is annotated with the relevant driver, road, and safety information.

International Journal of
AI ELECTRICAL CIVIL AND MECHANICAL ENGINEERING
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0434 www.zesterapublications.com Original Research Paper

The generated buzzer events are associated with the corresponding processed frames, allowing the alert information to be synchronized with the monitoring output.

3.10 Output Video Generation and FFmpeg Integration:

The processed frames containing the detected conditions, safety information, and visual annotations are combined to generate the output video. Since the buzzer is generated during system execution, FFmpeg is used to integrate the corresponding audio with the processed video.

This produces a synchronized audio-video output containing both the visual monitoring information and the corresponding safety alerts. The final output is generated in a browser-compatible H.264 video format. The output-generation flow is:

Safety Decision → Visual Annotation + Buzzer → Processed Video → FFmpeg Audio Integration → Final Output Video

3.11 Database and Event Logging:

The system uses SQLite for persistent storage of monitoring and session information. The database maintains both session-level and event-level records generated during system operation.

The stored information includes:

- Driver-state observations
- Frame-level measurements
- Head-pose information
- Seatbelt detections
- Road-monitoring observations
- Pothole detections
- Safety events
- Output-frame information
- Session information

This database structure allows the results of completed monitoring sessions to be retained and reviewed later. It also provides the stored data required for session-based analytics and reporting.

3.12 Streamlit Analytics and Visualization:

A Streamlit-based analytics dashboard is developed to provide an interactive interface for reviewing the stored monitoring information. The dashboard provides access to session history, safety events, detection results, event records, and graphical analysis.

The analytics component allows previously completed monitoring sessions to be examined after processing rather than relying only on real-time output. This supports analysis of driver states, head-pose behaviour, seatbelt detections, road observations, pothole detections, and safety events.

3.13 Datasets and Model Development:

Task-specific datasets are used for developing the different detection and monitoring components. The Driver Drowsiness Dataset (DDD) and the manually curated driver-state dataset support driver-state analysis. The Seatbelt Detection Dataset is used for seatbelt detection, the IIT Madras Pothole Dataset is used for pothole detection, and BDD100K is used for road-scene and road-object perception.

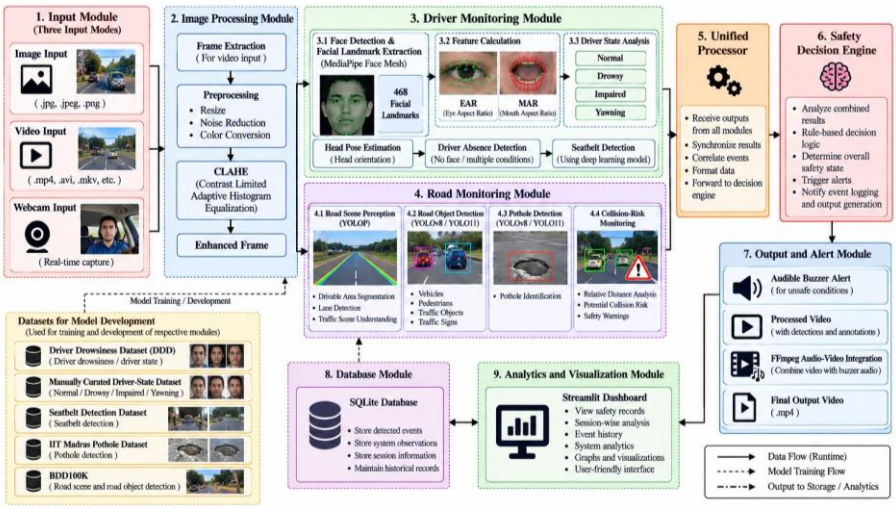
Table 1: The Seatbelt Detection Dataset

Dataset	Associated Component	Purpose
Driver Drowsiness Dataset (DDD)	Driver Monitoring	Drowsiness and driver-state analysis
Manually Curated Driver-State Dataset	Driver Monitoring	Normal, Drowsy, Impaired, and Yawning recognition
Seatbelt Detection Dataset	Seatbelt Monitoring	Seatbelt detection

International Journal of
AI ELECTRICAL CIVIL AND MECHANICAL ENGINEERING
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0434 www.zesterapublications.com Original Research Paper

IIT Madras Pothole Dataset	Pothole Detection	Pothole detection
BDD100K	Road Monitoring	Road-scene and road-object perception

The general model-development flow is:
Relevant Dataset → Model Development/Training → Developed Model → Monitoring Component → Unified Processor → Safety Decision Engine
Thus, the proposed methodology integrates computer vision, facial landmark analysis, deep-learning-based detection, road-scene perception, safety decision logic, real-time alerts, synchronized video processing, persistent database storage, and interactive analytics into a single framework for driver and road safety monitoring.



"Fig 1: Proposed Architecture"

4. RESULTS AND ANALYSIS:

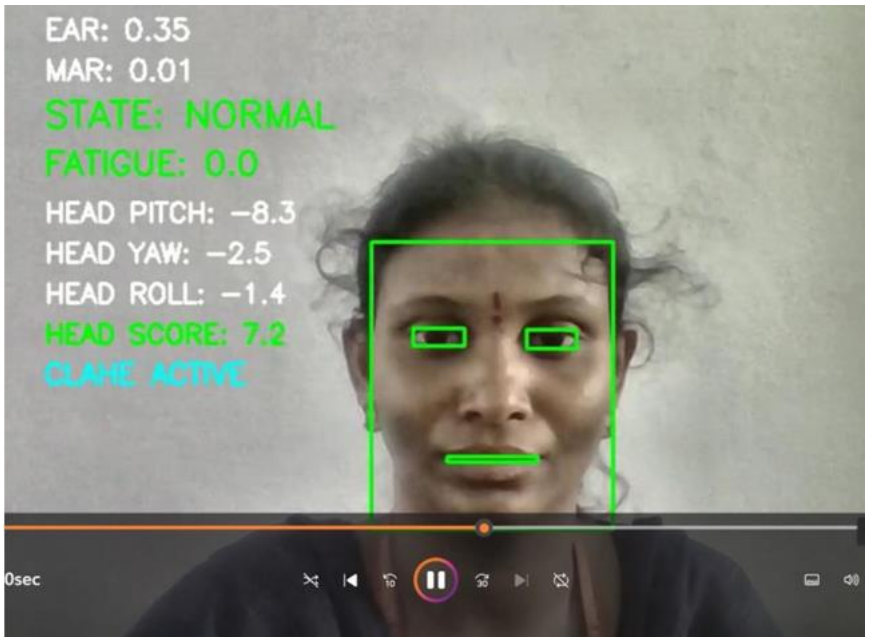
The proposed Intelligent Computer Vision-Based Driver and Road Safety Monitoring System Using Deep Learning was evaluated using image, video, and monitoring inputs to verify the functional operation and integration of the developed components. The evaluation covered driver monitoring, image enhancement, seatbelt detection, road-object and road-scene monitoring, pothole detection, safety decision-making, alert generation, video processing, database logging, and session-level analytics. A total of 30 functional and system-level test cases were evaluated, and all were recorded as Pass.

The driver-monitoring module successfully identified the tested Normal, Drowsy, Yawning, and Impaired conditions, along with driver absence and abnormal head movement. EAR, MAR, and head-pose parameters such as pitch, yaw, and roll were also obtained. The results show that multiple facial and head-related observations could be used together for driver-condition monitoring. CLAHE was successfully incorporated into the image-processing pipeline. The seatbelt module detected the compliant condition with a recorded confidence of 0.55, while non-detection generated the corresponding safety response.

Road-monitoring tests successfully detected objects such as cars and buses and generated corresponding safety conditions. Pothole testing produced zero detections for a pothole-free scene, while other tests detected one pothole at 44.0% confidence and three potholes at 64.9% confidence. These confidence values represent individual detection outputs and are not overall model-accuracy measures.

The outputs of the individual modules were integrated through the Unified Processor and Safety Decision Engine. Fig. 2 shows the annotated unified processed output, demonstrating the combined presentation of driver, seatbelt, road, safety, and alert information.

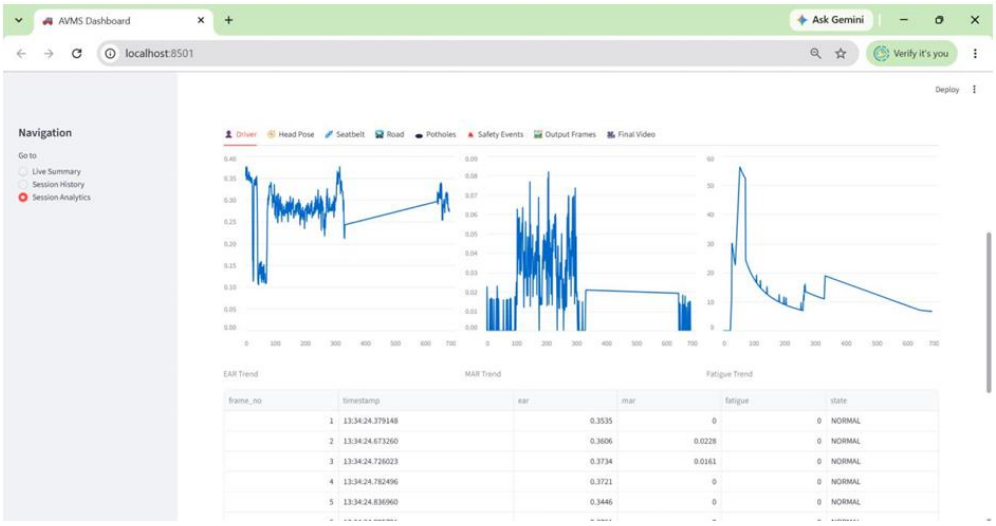
International Journal of
AI ELECTRICAL CIVIL AND MECHANICAL ENGINEERING
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0434 www.zesterapublications.com Original Research Paper



"Fig. 2. Annotated Unified Processed Video Output"

During system-level testing, 202 frames were processed, with 27 frames requiring buzzer activation, and the final safety state was NORMAL. An extended 97.35-second video containing 1,947 frames was also processed successfully. The buzzer audio was integrated with the processed video using FFmpeg to generate the final synchronized output. These results demonstrate the coordinated operation of the monitoring, decision, alert, and video-processing stages. The system also provided session-level analysis through SQLite and the Streamlit dashboard.

Fig. 3 presents the driver-monitoring trend analysis, which provides a temporal view of driver-related observations and allows changes in driver behaviour to be examined across a session.

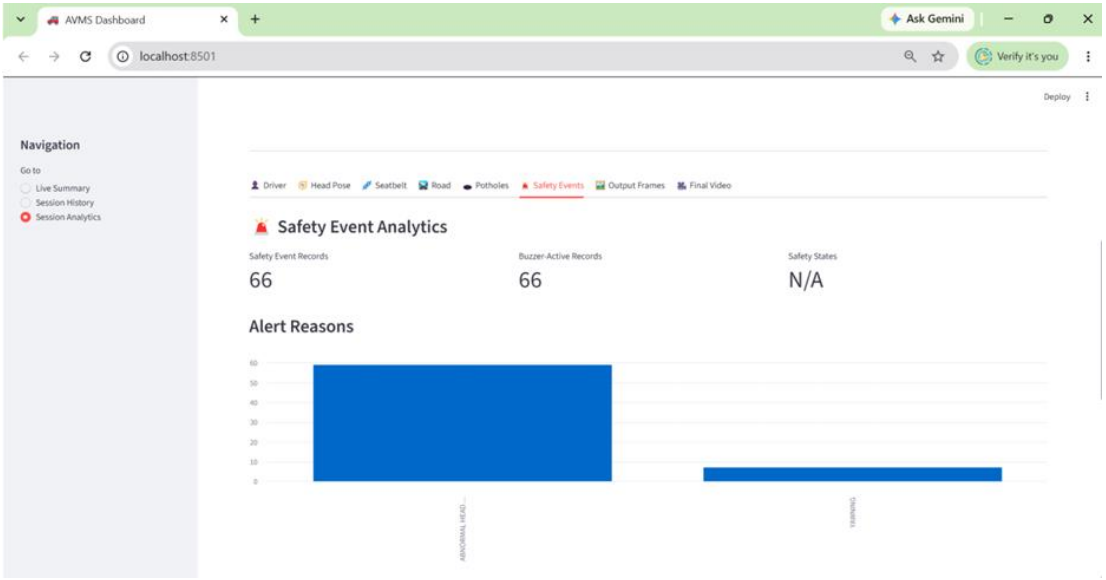


"Fig. 3. Driver Monitoring Analytics and Trend Analysis"

In the evaluated dashboard session, 284 total frames, 93 buzzer events, and 50 abnormal head-pose frames were recorded. The system also retained 67 monitoring sessions, enabling previous results to be reviewed.

Fig. 4 presents the safety-event analytics and alert-reason distribution. A total of 93 safety events were recorded in the evaluated session. The analysis shows how observations from driver, seatbelt, road, and pothole monitoring were converted into safety events through the centralized decision process.

International Journal of
AI ELECTRICAL CIVIL AND MECHANICAL ENGINEERING
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0434 www.zesterapublications.com Original Research Paper



"Fig 4: Safety Event Analytics and Alert Reason Distribution"

Overall, the results demonstrate the functional operation and integration of the proposed system under the tested conditions. The system combines driver and road monitoring with centralized safety decisions, audible alerts, synchronized video output, persistent data storage, and session-level analytics. The evaluation is primarily functional and system-level; therefore, the reported confidence values and test-case results should not be interpreted as overall machine-learning accuracy. Quantitative measures such as accuracy, precision, recall, F1-score, mAP, and FPS are not reported because corresponding ground-truth-based measurements were not established for all implemented components.

5. CONCLUSION:

The project "Intelligent Computer Vision-Based Driver and Road Safety Monitoring System Using Deep Learning" presents an integrated computer vision-based framework for real-time monitoring of driver behaviour, vehicle safety conditions, and the surrounding road environment. The system combines driver-state monitoring, head-movement analysis, seatbelt detection, road-object detection, and pothole detection within a unified processing pipeline. The outputs are integrated through a Safety Decision Engine to determine the overall safety state and generate corresponding alerts.

The driver-monitoring component uses facial landmark analysis to obtain Eye Aspect Ratio (EAR), Mouth Aspect Ratio (MAR), fatigue indicators, and head-pose information. These features support the detection of normal driving, yawning, drowsiness, driver absence, and abnormal head movement. The road and vehicle safety components complement this analysis by monitoring seatbelt compliance, detecting relevant road objects, and identifying potholes. The system outputs are demonstrated through annotated frames, trend graphs, state distributions, detection records, and safety-event analytics.

A key feature of the system is the Unified AVMS Processor, which integrates the individual monitoring components into a common processing framework. The system provides synchronized buzzer alerts for detected safety conditions and generates annotated H.264-compatible video output. An interactive Streamlit dashboard enables analysis of session history, driver states, head-pose trends, seatbelt statistics, road-monitoring results, pothole information, safety events, saved output frames, and generated video outputs.

The system also incorporates SQLite-based persistent data management for storing structured information related to monitoring sessions, driver-monitoring frames, head pose, seatbelt detections, road observations, pothole detections, safety events, output frames, and session outputs. The stored records support historical analysis and session-based reporting.

International Journal of
AI ELECTRICAL CIVIL AND MECHANICAL ENGINEERING
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0434 www.zesterapublications.com Original Research Paper

Overall, the implementation demonstrates the feasibility of integrating deep-learning-based object detection, facial landmark analysis, road-scene perception, safety decision logic, real-time alerts, video processing, database storage, and interactive analytics within a unified vehicle safety-monitoring framework. The developed system provides a foundation for intelligent driver and road safety assistance and can be further extended through additional datasets, improved detection models, broader real-world testing, and integration of additional vehicle safety sensors.

REFERENCES:

[1] G. Viswanath, "Artificial Intelligence-driven Frameworks for Fostering Active Participation and Learning in Language Classrooms", in International Journal of Interpreting Enigma Engineers (IJIEE), 2024, Vol.1, No.3, pp.23-32.

[2] G. Viswanath, "Multiple Cancer Types Classified Using CTMRI Images Based On Learning Without Forgetting Powered Deep Learning Models", in International Journal of HRM and Organizational Behavior, 2024, Vol.12, pp 243-253.

[3] D. Kim, H. Park, T. Kim, W. Kim, and J. Paik, "Real-time driver monitoring system with facial landmark-based eye closure detection and head pose recognition", in Scientific Reports, vol. 13, Article 18264, 2023. doi: 10.1038/s41598-023-44955-1.

[4] A.-C. Phan, T.-N. Trieu, and T.-C. Phan, "Driver drowsiness detection and smart alerting using deep learning and IoT", in Internet of Things, vol. 22, Article 100705, 2023. doi: 10.1016/j.iot.2023.100705.

[5] R. Albasrawi, F. F. Fadhil, and M. T. Ghazal, "Driver drowsiness monitoring system based on facial landmark detection with convolutional neural network for prediction", in Bulletin of Electrical Engineering and Informatics, vol. 11, no. 5, pp. 2637–2644, 2022. doi: 10.11591/eei.v11i5.3966.

[6] E. Keshtkaran, B. von Berg, G. Regan, D. Suter, and S. Z. Gilani, "Estimating blood alcohol level through facial features for driver impairment assessment," in Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2024. doi: 10.1109/WACV57701.2024.00448.

[7] A. M. Upadhyay, B. Sutrave, and A. Singh, "Real time seatbelt detection using YOLO deep learning model," in 2023 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), 2023, pp. 1–6. doi: 10.1109/SCEECS57921.2023.10063114.

[8] W. Yassin, F. Abdollah, A. Ismail, and P. Ragam, "Seatbelt detection in traffic system using an improved YOLOv5", in Journal of Advanced Computing Technology and Application, vol. 5, no. 2, pp. 1–17, 2023. doi: 10.54554/jacta.2023.05.02.001.

[9] D. Wu, M.-W. Liao, W.-T. Zhang, X.-G. Wang, X. Bai, W.-Q. Cheng, and W.-Y. Liu, "YOLOP: You only look once for panoptic driving perception", in Machine Intelligence Research, vol. 19, no. 6, pp. 550–562, 2022. doi: 10.1007/s11633-022-1339-y.

[10] P. Klco, D. Koniar, L. Hargas, and M. Paskala, "Automated detection of potholes using YOLOv5 neural network", in Transportation Research Procedia, vol. 74, pp. 1150–1155, 2023. doi: 10.1016/j.trpro.2023.11.255.

[11] C. Ruseruka, J. Mwakalonge, G. Comert, S. Siuhi, F. Ngeni, and Q. Anderson, "Augmenting roadway safety with machine learning and deep learning: Pothole detection and dimension estimation using in-vehicle technologies", in Machine Learning with Applications, vol. 16, Article 100547, 2024. doi: 10.1016/j.mlwa.2024.100547.

[12] R.-C. Chen, C. Dewi, Y.-C. Zhuang, and J.-K. Chen, "Contrast limited adaptive histogram equalization for recognizing road marking at night based on YOLO models", in IEEE Access, vol. 11, pp. 92926–92942, 2023. doi: 10.1109/ACCESS.2023.3309410.

[13] S. Bouhsissin, N. Sael, F. Benabbou, A. Soultana, and A. Jannani, "SafeSmartDrive: Real-time traffic environment detection and driver behavior monitoring with machine and deep learning", in IEEE Access, vol. 12, pp. 169499–169517, 2024. doi: 10.1109/ACCESS.2024.3498596.

Received: 14-09-2026 | Accepted: 19-10-2026 | Published: 26-10-2026

[14] J. Zhong, D. Kong, Y. Wei, and B. Pan, "YOLOv8 and point cloud fusion for enhanced road pothole detection and quantification", in Scientific Reports, vol. 15, Article 11260, 2025. doi: 10.1038/s41598-025-94993-0.
[15] P. Visconti, G. Rausa, C. Del-Valle-Soto, R. Velázquez, D. Cafagna, and R. De Fazio, "Innovative driver monitoring systems and on-board-vehicle devices in a smart-road scenario based on the Internet of Vehicle paradigm: A literature and commercial solutions overview", in Sensors, vol. 25, no. 2, Article 562, 2025. doi: 10.3390/s25020562.
[16] F. Yu, H. Chen, X. Wang, W. Xian, Y. Chen, F. Liu, V. Madhavan, and T. Darrell, "BDD100K: A diverse driving dataset for heterogeneous multitask learning", in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020. doi: 10.1109/CVPR42600.2020.00271.
[17] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M. G. Yong, J. Lee, W.-T. Chang, W. Hua, M. Georg, and M. Grundmann, "MediaPipe: A framework for building perception pipelines," arXiv preprint arXiv:1906.08172, 2019.
[18] Y. Kartynnik, A. Ablavatski, I. Grishchenko, and M. Grundmann, "Real-time facial surface geometry from monocular video on mobile GPUs," in Proceedings of the IEEE/CVF Workshop on Computer Vision for Augmented and Virtual Reality, 2019.
[19] T. Soukupová and J. Čech, "Real-time eye blink detection using facial landmarks," in Proceedings of the 21st Computer Vision Winter Workshop, 2016, pp. 1–8.
[20] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," Computer Vision, Graphics, and Image Processing, vol. 39, no. 3, pp. 355–368, 1987. doi: 10.1016/S0734-189X(87)80186-X.
[21] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779–788. doi: 10.1109/CVPR.2016.91.
[22] G. Jocher and J. Qiu, "Ultralytics YOLO11," version 11.0.0, Ultralytics, 2024. Available: https://github.com/ultralytics/ultralytics
[23] G. Bradski, "The OpenCV library," Dr. Dobb's Journal of Software Tools, 2000.
[24] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "PyTorch: An imperative style, high-performance deep learning library," in Advances in Neural Information Processing Systems (NeurIPS), vol. 32, 2019.
[25] FFmpeg Developers, "FFmpeg: A complete, cross-platform solution to record, convert and stream audio and video," FFmpeg Documentation. Available: https://ffmpeg.org/documentation.html
[26] Streamlit, "Streamlit documentation," Streamlit Documentation. Available: https://docs.streamlit.io/
[27] SQLite Consortium, "SQLite documentation," SQLite Documentation. Available: https://www.sqlite.org/docs.html