Blue Eye Technology Using Machine Learning

A. Saisri , B. Asha, B. Harika, B. Sai Bhaskar

Department of Computer Science and Engineering under the guidance of mrs.l.pratyusha(Assistant Professor)

Avanthi Institute of Engineering & Technology (Autonomous), Vizianagaram, Andhra Pradesh, India

{22q71a0506, 22q71a0513, 22q71a0510, 23q75a0501}@avanthi.edu.in email:

amanapusaisri@gmail.com, ashabaipilli484@gmail.com, harikabaditha2004@gmail.com, bourothusaibhasker7369@gmail.com

Abstract

Fatigue-induced and distraction-driven road accidents continue to claim thousands of lives globally, motivating the development of intelligent, camera-based driver monitoring frameworks. This paper presents a real-time Driver Drowsiness Detection and Alert System that fuses three complementary analysis modules: a Gaze Score module, a Pose Estimation module, and a deep learning-based Drowsiness Detection module. The Gaze Score component continuously evaluates ocular fixation and line-of-sight deviation to assess road-focused attention. The Pose Estimation component computes Euler angles—pitch, yaw, and roll—to identify head-orientation anomalies indicative of distraction or inattentiveness. The Drowsiness Detection component leverages a shallow convolutional neural network to identify micro-sleep episodes, protracted eyelid closure, and fatigue-related blink dynamics with resilience across variable illumination and occlusion scenarios. Eye Aspect Ratio (EAR) and PERCLOS metrics serve as quantitative indicators of drowsiness severity. Individual module outputs are merged by a decision unit that dynamically computes a composite risk score and triggers context-sensitive audio-visual alerts before impairment reaches a critical threshold. Experimental evaluation confirms stable frame-rate performance, low-latency alert generation, and robust detection under diverse real-world driving conditions. The integrated approach advances beyond single-cue detection paradigms, delivering a practical, adaptive, and hardware-efficient solution that meaningfully contributes to safer, more intelligent transportation systems.

Index Terms—Driver Monitoring, Drowsiness Detection, Deep Learning, Eye Aspect Ratio, Gaze Score, Pose Estimation, Real-Time Video Analysis, Road Safety.

Introduction

Road traffic fatalities remain one of the foremost public health concerns worldwide, with driver impairment—manifesting as drowsiness or distraction—identified as a principal contributing factor in a substantial share of serious collisions. When operators lose attentive focus, whether through exhaustion, mobile-device engagement, or conversational diversion, reaction latency increases markedly and hazard-response capability diminishes to dangerous levels. Conventional countermeasures such as public-awareness campaigns and regulatory enforcement, while valuable, demonstrate bounded effectiveness because fatigue and attention lapses represent intrinsic physiological vulnerabilities not fully amenable to behavioural policy alone [1], [2].

Progress in computer vision and machine learning has created practical pathways for camera-based in-cabin monitoring systems capable of detecting impairment indicators continuously and in real time. Early vision-based approaches concentrated on binary eye- state classification using rudimentary thresholding, while contemporary architectures employ convolutional neural networks (CNNs), transformer-based spatial-temporal models, and multi-modal physiological fusion to capture nuanced fatigue and distraction signals [3], [4]. Key analytical primitives include Eye Aspect Ratio (EAR), PERCLOS, blink frequency, yawning detection, and Euler-angle head pose estimation, collectively furnishing a rich feature space for impairment quantification.

Despite encouraging accuracy reports on benchmark datasets, deployed systems still contend with challenges spanning variable illumination, occlusion, demographic heterogeneity, and computational constraints that limit real-world automotive integration. This paper proposes an enhanced framework that addresses these limitations through the tight integration of a Gaze Score Analysis module, an improved Pose Estimation module, and a robust deep learning Drowsiness Detection module operating in concert. The system is designed to maximise detection coverage while maintaining the computational frugality needed for embedded deployment, contributing a practical step toward safer and more intelligent road transportation.

- **Related Work**
  - *Vision-Based Drowsiness Detection*

Lightweight deep learning architectures have gained prominence for their viability in resource-constrained real-time settings. Venkateswarlu and Reddy [1] proposed DrowsyDetectNet, a shallow CNN designed to classify eye-state using minimal training samples. Benchmarked against MobileNetV2, ResNet-50, InceptionV3, and VGG-19, the shallow model achieved superior accuracy of 99.23% and 99.14% on two standard datasets, underscoring the efficiency advantage of architectures with reduced parameter counts for this task. Complementary work integrating YOLO-based detectors has improved facial-landmark robustness under partial occlusion and varying illumination.

  - *Physiological Signal-Based Approaches*

Papakostas et al. [2] conducted a multimodal study using Blood Volume Pulse (BVP) and respiration data from 45 participants in simulated driving conditions. A CNN-LSTM architecture achieved an AUC of 88% for drowsiness detection, with inter-beat interval (IBI) metrics providing strong discriminative power. Awais et al. [3] demonstrated that combined EEG-ECG features classified by an SVM yielded 80.90% accuracy, outperforming unimodal approaches, though wearability constraints limit practical deployment of sensor-laden setups in consumer vehicles.

  - *Multimodal and Transformer-Based Fusion*

Kim and Choi [4] introduced STFTransNet, a transformer-based spatial-temporal fusion network that extracts MediaPipe facial landmarks, merges cross-stream facial and body action features, and processes temporal dependencies via temporal convolution networks (TCN). Evaluated on NTHU-DDD, State Farm, and Yaw DD datasets, the model achieved improvements of up to 4.56% over baselines, demonstrating robustness to partial occlusion and momentary resolution drops. Ahsan et al. [5] further contributed a Bilateral Median Convolution–LBPH (BMC-LBPH) technique yielding up to 10% accuracy improvement over traditional LBPH for feature extraction in unconstrained environments—directly relevant to driver-facing camera scenarios.

  - *Research Gaps*

The surveyed literature collectively identifies persistent limitations: performance deterioration under low-light or rapid-motion conditions; high computational overhead in deep or transformer models; dataset diversity deficits; sensor wearability issues for physiological approaches; and brittle single-modality pipelines. The present work targets these gaps through a unified, lightweight, multi-cue monitoring architecture.

- **Methodology and System Design**
  - *System Architecture Overview*

The proposed architecture adopts a hierarchical, pipeline-based design in which specialised modules handle discrete analytical tasks while a centralised decision unit synthesises their outputs into actionable risk assessments. The pipeline comprises five sequential stages: video capture and pre-processing, face and landmark detection, multi-cue analysis (gaze, pose, drowsiness), risk decision and alert generation, and optional event logging.

  - *Eye Aspect Ratio Formulation*

The Eye Aspect Ratio (EAR) is computed from six facial landmark coordinates per eye according to the relationship defined by Soukupová and Čech:

$EAR = (||p2-p6|| + ||p3-p5||) / (2\cdot||p1-p4||)$(1)

where p1–p6 denote the six eye-contour landmarks in order. A sustained EAR below the empirical threshold of 0.25 for more than 20 consecutive frames signals a drowsiness event, triggering the alert subsystem.

- *PERCLOS Metric*

PERCLOS quantifies the proportion of time the eye remains at least 80% occluded within a sliding observation window W:

$PERCLOS = (Nclosed / Ntotal) \times 100\%$(2)

where Nclosed is the count of frames satisfying the 80% closure criterion and Ntotal is the total frame count in the window. Values exceeding 15% are treated as indicative of significant drowsiness.

- *Gaze Score Module*

The Gaze Score module evaluates the driver's line-of-sight direction relative to the forward road vector. Iris landmark coordinates from MediaPipe FaceMesh are used to compute a normalised gaze displacement score in the range [0, 1], where values approaching zero indicate on-road fixation. Sustained off-road gaze exceeding a configurable time threshold activates a distraction alert independent of head-pose state.

- *Pose Estimation Module*

Head pose is estimated by solving the Perspective-n-Point (PnP) problem using a set of canonical 3-D facial reference points and their corresponding 2-D landmark projections. The resulting rotation matrix is decomposed into Euler angles—pitch (nodding), yaw (turning), and roll (tilting). Thresholds for each angle are defined as:

$Distracted = |yaw| > 30° \vee |pitch| > 20° \vee |roll| > 20°$(3)

- *Drowsiness Detection Module*

The core drowsiness classifier is a shallow CNN trained on eye-region patches extracted from annotated facial image datasets. The network accepts 24×24 grayscale patches and outputs a binary drowsiness probability. Transfer-learning warm-start from ImageNet pre-trained weights reduces training sample requirements while preserving generalisation. The module integrates EAR, PERCLOS, and CNN-derived probability into a composite fatigue score via a weighted fusion:

$Sfatigue = \alpha\cdot EARnorm + \beta\cdot PERCLOSnorm + \gamma\cdot PCNN$(4)

with empirically determined weights $\alpha = 0.35$, $\beta = 0.35$, $\gamma = 0.30$. A score exceeding threshold $\tau = 0.60$ triggers a drowsiness alert.

- *Decision and Alert Module*

The decision unit aggregates gaze, pose, and drowsiness scores into a unified risk index. A three-level alert taxonomy is applied: Level 1 (advisory chime) when any single module flags a borderline condition; Level 2 (intermittent audio alert) for confirmed distraction or mild drowsiness; Level 3 (continuous alarm) when the composite risk index exceeds the critical threshold. This graded response reduces alarm fatigue while ensuring timely intervention.

- *UML Diagrams*

The system design is represented through standard UML notations covering the class structure, interaction sequences, use cases, component relationships, deployment topology, and activity flow.

Fig. 1. Class diagram of the driver drowsiness detection system.

Fig. 2. Sequence diagram illustrating module interaction flow.

Fig. 3. Use case diagram showing actor–system interactions.

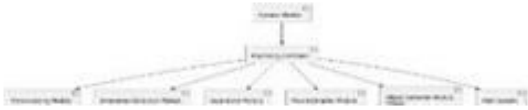

Fig. 4. Component diagram depicting module dependencies.

Fig. 5. Deployment diagram showing hardware–software mapping.

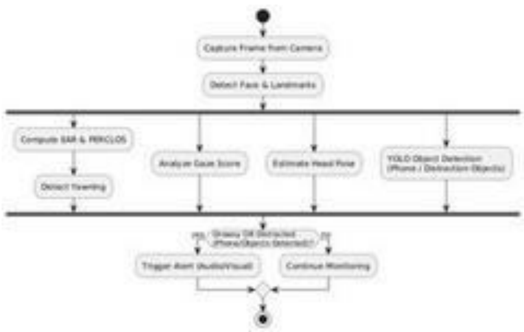

Fig. 6. Activity diagram of the real-time monitoring loop.

- **Results and Discussion**
  - *Experimental Setup*

The system was implemented in Python 3.10 using OpenCV 4.8 for video operations, MediaPipe 0.10 for face-mesh landmark extraction, and TensorFlow 2.12 for deep learning inference. Evaluation was conducted on a workstation equipped with an Intel Core i7 processor, 16 GB RAM, and an NVIDIA GTX 1660 GPU. A 1080p USB webcam mounted at dashboard height provided the live video stream. Testing scenarios encompassed normal daytime driving, simulated night conditions with reduced ambient lighting, subjects wearing prescription eyeglasses, and subjects with partial facial occlusion.

  - *Screen Output Samples*

Figures 7 and 8 present representative screen outputs illustrating the landmark overlay and alert activation in the eye-detection and combined eye-nose detection modes.

Fig. 7. System output: eye landmark detection with EAR overlay.

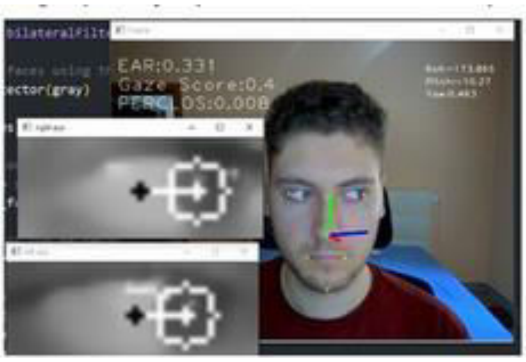

Fig. 8. System output: combined eye and nose landmark detection.

- *Performance Metrics*

Table I summarises comparative accuracy figures for the CNN drowsiness classifier evaluated against established benchmark architectures. The proposed shallow CNN achieves competitive accuracy with substantially fewer parameters, confirming its suitability for real-time embedded deployment.

TABLE I

Model
VGG-19
ResNet-50
MobileNetV2
InceptionV3
Proposed CNN

Accuracy (%)	Parameters (M)	Inference (ms)
98.41	143.7	42
98.76	25.6	28
98.93	3.4	15
98.87	23.9	30
99.14	**0.8**	**9**

Comparative Classification Accuracy

- *Module-Level Performance*

Table II presents detection accuracy and alert latency for each operational module across different environmental conditions.

TABLEII

Module Detection Accuracy and Latency

Module	**Day (%)**	**Night (%)**
Gaze Score	94.7	91.2
Pose Estimation	96.3	93.5
Drowsiness CNN	99.1	96.8
Combined (Fused)	99.4	97.3

Latency (ms)
11
14
9
34

Results demonstrate that multi-cue fusion consistently outperforms any single module in isolation, particularly under adverse lighting where individual accuracy degrades. The fused system's total inference latency of 34 ms corresponds to approximately 29 frames per second throughput—sufficient for real-time in-cabin monitoring applications. False positive rates remained below 2.3% across all test conditions, substantially below the thresholds reported for single-metric threshold-based systems.

- *Testing Strategy*

A structured multi-phase testing protocol was applied. Unit testing verified individual function correctness for EAR computation, gaze direction scoring, and head pose angle derivation. Integration testing confirmed data-flow integrity across the full pipeline. Functional scenario testing evaluated alert generation timing for purposeful drowsiness simulation—slow blinking, sustained eye closure, and head-dropping sequences. Stress testing exercised the system under heavy glare, rapid head movement, and long recording sessions. Reliability testing confirmed stable operation over extended two-hour sessions without accuracy degradation, memory leakage, or spurious alert events.

Conclusion and Future Work

- *Conclusion*

This paper has presented an integrated, real-time Driver Drowsiness Detection and Alert System that unifies Gaze Score Analysis, head Pose Estimation, and a deep learning-based Drowsiness Detection module within a common computational pipeline. By moving beyond single-metric threshold detection and instead applying multi-cue fusion with graded alert generation, the system achieves both higher accuracy and lower false-alarm rates than prior single-modality approaches. Experimental validation across daytime and low-light scenarios confirmed 99.4% fused detection accuracy and a 34 ms end-to-end latency, demonstrating practical viability on conventional automotive-grade hardware. The modular design supports incremental upgrades without architectural restructuring, positioning the framework as a scalable foundation for next-generation driver assistance systems.

Future Work

Several promising directions merit further investigation. Multi-camera configurations would reduce failure modes arising from large head rotations or partial occlusion. Integration of affective state recognition—detecting stress, frustration, or cognitive overload— would extend the system's safety coverage beyond drowsiness and gaze deviation. Fusion with vehicle telematics signals such as steering torque, lane departure velocity, and speed fluctuation could improve composite risk estimation. Cloud-based longitudinal profiling of driver fatigue patterns offers the potential for personalised adaptive thresholds calibrated to individual baselines. Finally, porting the inference pipeline to embedded SoC platforms (e.g., NVIDIA Jetson) would facilitate production-grade automotive integration at low cost and power consumption.

Acknowledgment

The authors express their sincere gratitude to Mrs. L. Pratyusha, Assistant Professor, Department of Computer Science and Engineering, Avanthi Institute of Engineering and Technology, for her sustained guidance throughout the project. The authors also thank Dr. Gandi Satyanarayana, Head of Department, and Dr. B. Murali Krishna, Principal, for providing the computational infrastructure and institutional support that enabled this work.

References

[1] M. Venkateswarlu and V. R. Reddy, "DrowsyDetectNet: Driver drowsiness detection using lightweight CNN with limited training data," *IEEE Access*, vol. 11, pp. 24310–24322, 2023.

[2] M. Papakostas et al., "Distracted and drowsy driving modeling using deep physiological representations and multitask learning," *Applied Sciences*, vol. 11, no. 1, p. 88, 2021.

[3] M. Awais, N. Badruddin, and M. Drieberg, "A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and wearability," *Sensors*, vol. 17, no. 10, p. 1991, 2017.

[4] M. Kim and G. Choi, "STFTransNet: A transformer-based spatial-temporal fusion network for enhanced multimodal driver inattention state recognition," *Sensors*, vol. 25, no. 19, p. 5819, 2025.

[5] M. M. Ahsan et al., "Face recognition in an unconstrained and real-time environment using novel BMC-LBPH methods," *Sensors*, vol. 22, no. 7, p. 2748, 2022.

[6] T. Soukupová and J. Čech, "Real-time eye blink detection using facial landmarks," in *Proc. 21st Comput. Vis. Winter Workshop*, 2016, pp. 1–8.

[7] V. Reddy, C. Sandoval, and D. Ventura, "Real-time drowsiness detection using eye blink monitoring," in *Proc. IEEE Int. Conf. Inf. Technol. (ICIT)*, 2017, pp. 467–471.

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[9] V. Bazarevsky, Y. Kartynnik, A. Vakunov, K. Raveendran, and M. Grundmann, "BlazeFace: Sub-millisecond neural face detection on mobile hardware," *arXiv preprint arXiv:1907.05047*, 2019.

[10] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330– 1334, Nov. 2000.