Machindra K. Gaikwad, 2026, 14:2
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

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

Smart Fatigue Detection Model for Drivers

Shreyas Fatale, Ajinkya Bobade, Divya Jain, Kushal Agrawal, Prof. Mayur Chavan
Dept. of Computer Engineering Pune Institute of Computer Technology Pune, India

Abstract- An abstract of this paper will provide a hybrid driver fatigue system that is supposed to be deployed in real-time on resource-constrained edge devices. The suggested framework integrates the lightweight Convolutional Neural Network (CNN) models with geometrical and temporal feature extraction based on the landmarks to provide a stable fatigue monitoring at the minimal computational complexity. The two parallel CNN models are used to perform eye state, and yawning detection, and geometric fatigue indicators are computed simultaneously through facial landmark Analysis. The percentage of eyelid closure (Eye Closure Percentage (PERCLOS)) is used to determine eye fatigue and monitoring patterns of eyelid closure through time is an efficient measure of drowsiness, even in the process of lip detection in the presence of noise. The way of recognising yawning is the Jaw Drop Angle (JDA), which is a powerful geometric parameters calculated based on the nasal, chin, and jaw positions that is still reliable in spite of imprecision in lips location. A hybrid decision model combines deep learning-based predic-tions with geometric fatigue indicators to enhance the reliability of the system and help to reduce error. The general architecture proves to be practically viable as a cost-efficient approach to the intelligent driver assistance system, specifically in real-time embedded and edge computing applications.

Keywords: fatigue, CNN, computer vision, edge computing.

I. INTRODUCTION

Fatigue-related disabilities play a significant role in road traffic accidents, and thus, the importance of Driver Monitor-ing Systems (DMS) in increasing the safety of transportation is apparent. The traditional methods of monitoring the body such as electroencephalography (EEG) and electrocardiography (ECG) are also very accurate in detection but because they are invasive, expensive and would not work in continuous mode in normal driving situations. By contrast, non-invasive methods Vision-based fatigue detection methods which are based on facial analysis provide an alternative. However, those approaches that rely entirely on deep learning can be characterized by excessive computational complexities and can be vulnerable to occlusions and illumination variations, as well as head rotation.

In order to overcome these constraints, the current study proposes a hybrid driver fatigue recognition system, which concentrates on the eye-state and yawning detection with lightweight convolutional neural network (CNN) models, and real-time landmark-based geometric and temporal analysis. The time-varying eyelid closure patterns are measured as the Eye Closure Percentage (PERCLOS) of assessing eyelid fatigue, which is an indicator of drowsiness. Yawning is identified using the Jaw Drop Angle based on the landmark configurations of the face and thus can be identified reliably without the need to have highly accurate localization of lips. The proposed system, based on a combination of deep learning predictions with geometry-driven fatigue signals and temporal reasoning, allows increasing the reliability of fatigue detection and decreasing the number of false alerts. Its small size facilitates its implementation on edge and embedded computing platforms with resource constraints which supports a practical implementation of real time processes in real world
driving conditions.

II. OBJECTIVES

The current study is aimed at creating an organized scheme of real-time driver fatigue detection that can reveal the signs of drowsiness at the initial stages, therefore, allowing to manage the situation timely and enhance the safety of the road. A hybrid method is employed to balance the efficiency of the system and the amount of calculation required to be performed to augment embedded platforms and employs the deep learning method of visual analysis, complemented by the use of geometric features and localization methods.

The primary objectives of the given research are as follows:

- To design and train lightweight convolutional neural network (CNN) models for visual identification of driver fatigue indicators, particularly focusing on eye-state monitoring and yawning detection.
- To integrate CNN-derived predictions with landmark-based geometric and temporal fatigue features, ensuring reliable and low-latency inference suitable for resource-constrained edge and embedded computing platforms.

A review of the performance of the built fatigue detection system concerning the reliability of its detection, its responsiveness in real time, and its behavior in practical applications in vehicles driving.

By combining predictive models with geometry-based analysis in a resource-efficient design, the system aims to deliver a practical and scalable solution for driver monitoring on edge devices.

III. LITERATURE SURVEY

Recent advancements in driver fatigue detection have in-creasingly focused on combining deep learning with traditional feature-based approaches to improve robustness and real-time performance.

Savas¸ and Becerikli [1] proposed a multi-task CNN model that jointly analyzes eye and mouth regions to classify fa-tigue levels, demonstrating improved accuracy through shared feature learning. Zhang et al. [2] utilized infrared imaging to achieve reliable eye state detection under low-light conditions, making the system suitable for night driving scenarios.

Physiological signal-based approaches, such as the work by Wang et al. [3], leverage EEG signals and frequency-domain features to achieve high detection accuracy, though they remain intrusive. To address this limitation, Du et al. [4] introduced a multimodal framework combining facial features with heart rate signals using a hybrid CNN-RNN model, significantly improving robustness.

Several works focus on improving visual feature extraction. Zhuang et al. [5] employed segmentation-based CNN models for precise iris and pupil localization, while Khan et al.

[6] proposed transform-domain entropy features for efficient fatigue detection in low-resolution environments. Similarly, Ma et al. [7] introduced a two-stream 3D CNN model utilizing depth information for night-time fatigue detection.

Deep learning-based approaches such as Huang et al. [8] and Xing et al. [15] demonstrate the effectiveness of CNN architectures in capturing spatial facial features for fatigue classification. In contrast, Liu et al. [9] and Coetzer and Hancke [10] explored traditional machine learning techniques based on blink dynamics and eye movement features.

To improve system robustness, fusion-based methods have been proposed. Ji et al. [12] combined multiple fatigue in-dicators including eye closure, yawning, and head pose into a unified score, while Lyu et al. [13] and Gwak et al. [14] incorporated behavioral and physiological features to enhance detection performance.

Overall, recent research trends emphasize hybrid and multi-modal approaches that balance computational efficiency with detection accuracy, making them suitable for real-time deploy-ment in embedded driver monitoring systems.

IV. SYSTEM ARCHITECTURE

The driver fatigue detection system has the hybrid, real-time, and hybrid architecture as its base which refers to a combination of geometric facial analysis and deep learning inference. The system enables the determination of the driver’s cognitive state to be very accurate and dependable since it achieves redundancy. This two different analytical paths is designed to be installed on embedded computing platforms and consideration is given to efficiency and real-time response. The system’s operations flow is represented by five main steps:

Preprocessing and Data Acquisition, (2) Facial Landmark Localization, (3) Multi-Modal Feature Extraction, (4) Hybrid Decision Logic, and (5) Alert Generation

Data Acquisition and Preprocessing

The process starts with gathering the video stream from a camera that is mounted on the vehicle. In order to be consistent with the lighting conditions of the real world drives, each frame is given Contrast Limited Adaptive Histogram Equalization (CLAHE) preprocessing to improve local contrast and the visibility of the facial features.

Facial Landmark Localization

In real time, key facial landmarks are determined by a very accurate impinging algorithm. For every frame, it is set up to identify only one subject to enhance the efficiency of the computation and reduce the number of false positives.

Multi-Modal Feature Extraction

To maintain the strong point of the whole system, the operations are carried on independently in the different parts of the system:

Geometric Feature Analysis

- **Eye Closure Percentage (PERCLOS):** Quantifies eye fatigue by measuring the proportion of time the eyes remain partially or fully closed over a temporal window, computed using vertical eyelid distances normalized with respect to eye width.
- **Jaw Drop Angle (JDA):** Detects yawning behavior by estimating the angular displacement between the nose, chin, and jawline landmarks, providing a robust geometric indicator that is less sensitive to lip detection noise.
- **Head Pose Estimation:** Monitors head orientation and tilt using facial landmarks to identify drowsiness-related postural changes such as nodding and prolonged down-ward head inclination.

Deep Learning Feature Analysis

- **ROI Extraction:** Uses landmarks to locate and crop eye/mouth regions dynamically.
- **CNN Inference:** The pre-trained models, which are optimized for an edge, give the output in the form of probabilistic scores of eye closure and yawning.

Hybrid Decision Logic

Integrates geometric and deep learning predictions for in-creased reliability.

Temporal State Buffering: The sliding window used for storing the most recent states is instrumental in smoothing the results and also for transient false positives filtering.

TABLE 1: REFINED SUMMARY OF DRIVER FATIGUE DETECTION APPROACHES

Sr.	Author(s)	Work done	ML Algorithm	Classification Level	Methodology / Features Used
1	Savaş et al. [1]	Multi-task fatigue detection using facial regions	CNN	Multi-level	Joint learning of eye closure and yawning using PERCLOS and FOM on NTHU-DDD dataset
2	Zhang et al. [2]	Infrared-based eye state monitoring	CNN	Binary	Robust eye closure detection under low-light using IR video and blink dynamics
3	Wang et al. [3]	Physiological fatigue detection using EEG	PSD + Entropy	Continuous	Frequency-domain EEG features (alpha, beta, theta) from wearable sensors
4	Du et al. [4]	Multimodal fatigue detection (vision + HR)	CNN + RNN + Fuzzy	Binary	Fusion of facial features and heart rate signals for improved robustness
5	Zhuang et al. [5]	Eye state detection via iris segmentation	CNN + U-Net	Binary	Precise pupil/iris segmentation for accurate eye openness estimation
6	Khan et al. [6]	Low-resolution fatigue detection system	DWT + DCT + Entropy	Multi-class	Transform-domain entropy features for computationally efficient detection

7	Ma et al. [7]	Night-time fatigue detection	3D CNN (Two-stream)	Binary	Temporal depth features from infrared/depth cameras for night driving
8	Huang et al. [8]	Fatigue Detection CNN (FDCN)	CNN	Binary	Dedicated CNN architecture focusing on eye-mouth regions (CEW, ZJU datasets)
9	Liu et al. [9]	Blink-based fatigue estimation	RBF Neural Network	Binary	Blink frequency and duration modeling for fatigue classification
10	Coetzer and Hancke [10]	Comparative study of classifiers	SVM, ANN, AdaBoost	Binary	Evaluation of multiple classifiers using eye shape and motion features
11	Luo et al. [11]	Illumination-robust eye detection	AdaBoost + Tracking	Binary	Adaptive eye tracking under varying lighting conditions
12	Ji et al. [12]	Multi-feature fatigue fusion model	CNN + Fusion	Multi-level	Integration of eye, mouth, and head pose into unified fatigue score
13	Lyu et al. [13]	Fatigue detection using facial posture	Random Forest	Binary	Face orientation, gestures, and posture features from real datasets
14	Gwak et al. [14]	Mild drowsiness detection system	SVM + LR + RF	Multi-class	Combined physiological and behavioral features in simulator environment
15	Xing et al. [15]	CNN-based facial fatigue classification	CNN	Binary	Eye openness and facial region-based classification using deep learning

- **Fused Condition Logic:** Detects tiredness by logical oper-ations, EAR below the limit or eye closure detected by CNN with high confidence, thus enabling the identification of both blinks and microsleep.
- **Drowsiness Determination:** Points to incidents when the duration of the closure or the number of the repeated yawns goes beyond the clinical limits.

Alert Generation

Alerts are prompted by a fatigue detection.

- **Alert Prioritization**
  Ranks according to degree of seriousness (Yawn, Head Tilt, Drowsiness, Severe Drowsiness).
- **Auditory Feedback**

Confirms to the driver with a different pattern of buzzer/speaker or with a voice prompt each fatigue level i.e. easy and quick communication with the driver is achieved.

V.EVALUATION METRICS

Detection Accuracy

The identification power of the CNN-based and geometric classifiers is a result of their performance evaluated with standard classification metrics extracted from the confusion matrix (True Positives, False Positives, True Negatives, and False Negatives). The used metrics are:

- **Precision:** Displays the percentage of the correctness of the identification of the fatigue events in the total positive predictions.

$$\text{Precision} = \frac{TP}{TP + FP}$$

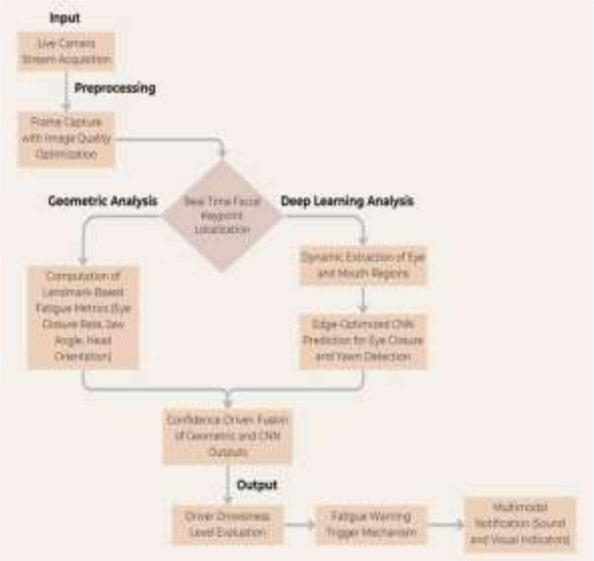

Fig.1. Block diagram of the proposed system.

- **Recall (Sensitivity):** Reflects the total number of fatigue cases that the system has correctly identified out of the total real ones, which is a vital point for driver safety.

$$\text{Recall} = \frac{TP}{TP + FN}$$

F1-Score: Is a single metric that combines both precision and recall by calculating their weighted harmonic mean. This metric is appropriate for imbalanced datasets.

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Real-Time Performance

This section of the paper takes into account the wearable embedded system's capabilities and its response time to show results when live camera footage is available.

Frames Per Second (FPS): It denotes the number of frames the system is able to handle within one second. The higher the FPS, the smoother the flow of images and thus the better the real-time performance will be.

End-to-End Latency: The total time that elapsed from the instance the frame was taken until the alert was generated. The smaller the time span, the quicker the driver gets the signal and thus the better the timing for intervention.

Resource Utilization: It is the portion of computer re-sources that are in use for a particular activity. The perfor-mance of edge devices with limited capabilities like Raspberry Pi can be maintained by controlling CPU and memory con-sumption which is part of resource utilization and is being monitored through system logs.

System Robustness

Robustness metrics are the ones that guarantee the normal-ized functioning of the system, even in hard situations in the real-life environment. The key metrics are:

Environmental Resilience: The assessment of a system under different light conditions, with the presence of the additional elements (e.g. spectacles), and partial occlu-sions in order to verify the system's generalization and stability.

False Alarm Rate (FAR): The figure that shows the percentage of those frames which were actually non-fatigued but mistakenly labeled as fatigue events.

$$\text{FAR} = \frac{FP}{FP + TN}$$

Missed Detection Rate (MDR): The figure that shows the percentage of the parts of the total real fatigue that was unnoticed.

$$\text{MDR} = \frac{FN}{TP + FN}$$

VI. PROPOSED METHODOLOGY

The proposed fatigue detection framework features a dual, pathway hybrid architecture that is optimized for deployment on real, time edge hardware. It combines CNN, based classifi-cation with landmark, driven geometric analysis, and a hybrid logic engine enhanced by an advanced reasoning capability.

CNN Model Architecture

Two parallel 5, layer CNNs (256 neurons/layer) to obtain high offline accuracy and are implemented as lightweight TFLite models:

- **Eye State Model:** Takes 128×128 grayscale eye images as input and produces open/closed probability.
- **Yawn Detection Model:** Is given 96×96 RGB mouth images to analyze and return yawn probability scores.

Geometric Feature Analysis

- **Eye Closure Percentage (PERCLOS):** Quantifies eye fatigue by measuring the proportion of time the eyes remain partially or fully closed over a temporal window, computed using vertical eyelid distances normalized with respect to eye width.
- **Jaw Drop Angle (JDA):** Detects yawning behavior by estimating the angular displacement between the nose, chin, and jawline landmarks, providing a robust geometric

indicator that is less sensitive to lip detection noise.

- **Head Pose Estimation:** Monitors head orientation and tilt using facial landmarks to identify drowsiness-related postural changes such as nodding and prolonged down-ward head inclination.

Hybrid Logic Engine

Multi, modal fusion is used to achieve dependable inference:

- **Eye Closure:** A trigger occurs when EAR is below threshold or CNN very confidently predicts closed.
- **Yawn Detection:** Is done by the CNN scores that are sustained for a long period of time or the prolonged MAR getting higher over a number of frames.
- **Temporal Filtering:** The 15, frame sliding buffer is used to get rid of noise that is transient, therefore, a persistent state is required to get a warning signal.
- **Cool down Mechanism:** Is used for the prevention of redundant yawn triggers during long events.

This architecture compromises computational efficiency with detection robustness, thus making it appropriate for embedded driver monitoring systems.

VII. DATASET AND TRAINING

The deep learning models were trained and validated on a wide range of publicly available and custom, built datasets. Details about the data sources, preprocessing pipeline, and training setup are provided in this section.

Dataset Sources

The models were trained on extensive datasets to ensure robustness and high accuracy:

- **Eye State Dataset:** The eye state classification system utilized the MRL Eye Dataset for its training, which is a large, scale, publicly available repository of eye images.
- **Yawn Dataset:** The yawn detection model utilized a large, publicly available yawn dataset.
- **Dataset Scale:** Each of the datasets had close to 100,000 images to make the datasets quite diverse for training purposes.
- **Custom Data Supplement:** Besides, a normal webcam was used to capture more data that can cause changes of light, facial expressions, and head poses. In this way, the public datasets were augmented for real, world applicability.

Preprocessing Steps

A standardized preprocessing pipeline was applied before training the models:

- **Image Resizing:** Eye images were resized to 128 × 128 pixels and mouth images to 96 × 96 pixels to match their respective model input layers.
- **Color Space Conversion:** Eye images were converted to grayscale to emphasize textural features. Mouth images were retained in RGB format.
- **Pixel Normalization:** In order to make the training more stable and the convergence faster, the pixels of all images were normalized to the continuous range [0, 1] by scaling each pixel through division by 255.0.

Training Details

The deep learning models were compiled and trained with standard configurations:

- **Loss Function:** Binary Cross-Entropy was used for both eye state and yawn classification tasks.
- **Optimizer:** To maintain an adaptive learning rate and achieve efficient convergence, the Adam optimizer was employed.
- **Training Procedure:** Each model was trained for 32 epochs using a 90:10 split for training and validation data.
- **Performance Metrics:** During the training, the model was evaluated using four performance metrics Accuracy, Precision, Recall, and F1, Score that provided a detailed assessment not only of the general correctness of the classification but also of the balance between sensitivity and specificity.

Data Augmentation

In order to improve model generalization and also prevent overfitting, the below augmentation methods have been uti-lized

- Horizontal Flipping
- Random brightness Adjustment
- Extra noise added

VIII. EXPERIMENTAL RESULTS

This section presents both quantitative and qualitative eval-uation of the proposed driver fatigue detection system. The ex-periments assess the individual performance of CNN models, the effectiveness of the hybrid framework, and overall system behavior.

Quantitative Comparison of CNN Models

The CNN models for eye state and yawning detection were evaluated using a dedicated test dataset. The eye detection model achieved an accuracy of 92.34%, correctly classifying 41,333 closed-eye and 42,152 open-eye instances, with min-imal misclassifications. This demonstrates strong generaliza-tion and robustness. Similarly, the yawning detection model achieved an accuracy of 91.84%, correctly identifying 2,572 non-yawn and 2,385 yawn instances. Low false positives and false negatives indicate reliable differentiation between normal mouth movements and yawning.

Comparative Analysis: Geometric vs. Hybrid Model

A comparison was conducted between a Geometric-Only model and the proposed Hybrid Model. The geometric ap-proach, based on landmark thresholds, is computationally efficient but sensitive to variations such as lighting, head pose, and occlusions, leading to false positives. The hybrid model combines geometric features with CNN predictions for improved reliability. Eye fatigue is measured using PERCLOS (temporal eye closure), while yawning is detected using jaw drop angle. Temporal buffering ensures consistency by filtering out transient errors.

Hybrid Decision Logic: Fatigue is detected using a logical OR fusion of geometric and CNN outputs. This allows early detection of fatigue cues while maintaining robustness against noise. As a result, the hybrid model achieves higher preci-sion, reduced false alarms, and improved overall performance compared to the geometric-only approach.

IX. CONCLUSION

This paper presents a robust hybrid driver fatigue detection system that combines landmark-based geometric and temporal analysis with lightweight convolutional neural network (CNN) architectures for efficient and accurate real-time monitoring. By fusing CNN-based visual classification with geometric fatigue indicators derived from facial landmarks, our approach achieves an effective balance between detection accuracy and computational efficiency, making it suitable for deployment on resource-constrained edge and embedded platforms.

Extensive experimental evaluation demonstrates that the system achieves an overall classification accuracy of approximately 92% while maintaining stable real-time performance. These results position the proposed hybrid framework as a promising solution for next-generation intelligent driver assistance systems, with strong potential to enhance road safety through reliable and non-intrusive fatigue detection.

REFERENCES

1. Savas¸, M., Becerikli, Y.: A Multi-task Convolutional Neural Network for Driver Fatigue Detection. IEEE Access 8, 124351–124362 (2020)
2. Zhang, X., et al.: Driver Drowsiness Detection Using Eye Closure and Yawning Features With Infrared Videos. Sensors 20(8), 1–17 (2020)
3. Wang, Z., Liu, J., Zhao, L.: EEG-Based Driver Fatigue Detection Using PSD and Entropy Measures. Biomedical Signal Processing and Control 55, 101678 (2020)
4. Du, Y., Yu, X., Liu, X.: Multimodal Recurrent Neural Network for Driver Fatigue Detection. IEEE Transactions on Intelligent Transportation Sys-tems 21(5), 1912–1923 (2020)
5. Zhuang, Y., et al.: Lightweight Pupil and Iris Localization for Eye State Detection Using CNN with Segmentation. Journal of Ambient Intelligence and Humanized Computing 12, 837–849 (2021)

6. Khan, M., et al.: Transform-Domain Entropy Features for Driver Fatigue Detection in Low-Resolution Videos. Pattern Recognition Letters 140, 48–56 (2020)
7. Ma, L., et al.: Real-Time Driver Fatigue Detection Using Depth Video Streams and Two-Stream 3D CNNs. Neurocomputing 438, 172–181 (2021)
8. Huang, C., et al.: FDCN: A CNN-Based Fatigue Detection Model Using Eye and Mouth Region Features. IEEE Access 9, 53582–53593 (2021)
9. Liu, Q., et al.: Eye Blink Detection and Fatigue Assessment Using Radial Basis Function Neural Networks. Applied Sciences 10(5), 1–13 (2020)
10. Coetzer, M., Hancke, G.: Classifier Performance Evaluation for Driver Drowsiness Detection Using Eye Features. IEEE Sensors Letters 5(3), 1–4 (2021)
11. Luo, H., et al.: Robust Eye Detection in Low-Light Conditions Using AdaBoost and Tracking Algorithms. Image and Vision Computing 89, 1–10 (2019)
12. Ji, Q., et al.: Multimodal Fusion for Drowsiness Detection Using CNN and Heuristic Score Aggregation. IEEE Transactions on Multimedia 21(5), 1250–1262 (2019)
13. Lyu, Y., et al.: Driver Fatigue Detection Using Random Forest on Face Posture and Gesture Features. Expert Systems with Applications 139, 112851 (2020)
14. Gwak, Y., et al.: Mild Drowsiness Detection Using Physiological and Behavioral Features. IEEE Access 7, 165485–165499 (2019)
15. Xing, T., et al.: A CNN-Based Fatigue Detection Framework Using Eye Openness Features. Multimedia Tools and Applications 79, 24201–24218 (2020)