Real-Time Driver Fatigue Detection Using NVIDIA Jetson Nano and Deep Learning

Mamona Sadaf*, Sarah Omer, Menahil Ahsan
Department of Electrical Engineering, School of Electrical Engineering and Computer Science (SEECS)
National University of Sciences and Technology (NUST), Islamabad, Pakistan
*Corresponding author: msadaf.bee22seecs@seecs.edu.pk

***Abstract*—Driver fatigue detection using facial analysis has emerged as a critical non-invasive approach to prevent road traffic accidents caused by drowsiness. This paper presents a comprehensive literature review of deep learning and computer vision-based fatigue detection methods, including CNN architectures, Eye Aspect Ratio (EAR), Mouth Aspect Ratio (MAR), and hybrid detection frameworks. The feasibility analysis investigates practical implementation of CNN-based drowsiness detection on the NVIDIA Jetson Nano 4GB platform for real-time automotive deployment. The review identifies key performance metrics, hardware optimization strategies, and challenges that inform the implementation of an efficient embedded fatigue monitoring system. This work demonstrates that optimized deep learning models can achieve real-time performance on resource-constrained edge devices suitable for safety-critical automotive applications.**

***Index Terms*—Driver fatigue detection, NVIDIA Jetson Nano, embedded deep learning, computer vision, Convolutional Neural Network, edge computing, real-time drowsiness detection, Eye Aspect Ratio, Mouth Aspect Ratio, automotive safety**

I. Introduction

Road traffic accidents constitute a major global health concern with driver fatigue identified as a primary contributing factor. The National Highway Traffic Safety Administration estimates that drowsy driving accounts for approximately 20% of annual traffic fatalities [1]. Fatigue impairs cognitive function and reduces reaction time thereby increasing accident risk. The timeliness of detection is critical for preventing accidents before driving performance deteriorates. While physiological signal monitoring remains the gold standard for drowsiness assessment, such methods require intrusive sensor placement limiting practical automotive deployment.

Vision-based detection of facial drowsiness indicators offers a non-invasive, cost-effective, and rapid screening modality. With the advent of compact edge-AI platforms such as the NVIDIA Jetson Nano, real-time facial analysis in vehicles and transportation systems is now feasible. This document reviews the literature on vision-based fatigue detection and examines the feasibility of implementing a CNN-based drowsiness detection pipeline optimized for Jetson Nano deployment. The system targets continuous monitoring of eye closure patterns and yawning frequency to provide early warning alerts to drivers.

II. Literature Review

The literature on vision-based driver fatigue detection spans handcrafted feature methods, deep convolutional models, temporal architectures, landmark-based approaches, and hybrid detection frameworks. Below we present an expanded review with dedicated subsections for the Survey, the State-of-the-Art, the Main Technique, and detailed discussion of supporting works.

A. Survey: Fu et al. (2024)

Fu, Boutros, Lin, Damer (2024) [1]. The comprehensive survey by Fu et al. provides one of the most technically detailed overviews on drowsiness detection across multiple application domains including automotive safety, workplace monitoring, public transportation, aviation, and healthcare. The survey classifies existing approaches into three primary measurement modalities: physiological signal analysis (EEG and ECG), vehicle behavior monitoring, and vision-based facial feature extraction.

Physiological approaches utilizing EEG signals achieve high accuracy by directly measuring brain wave patterns associated with alertness states. EEG-based systems detect drowsiness through spectral power analysis in alpha, beta, theta, and delta frequency bands. However these methods require intrusive electrode placement on the scalp which limits practical deployment in driving scenarios. ECG-based approaches monitor heart rate variability (HRV) as an indicator of autonomic nervous system activity during drowsiness. While less intrusive than EEG, ECG methods still require body-worn sensors and exhibit lower discrimination capability for early-stage drowsiness.

Vehicle-based techniques monitor steering wheel movements, lane departure events, and acceleration patterns to infer driver state. These methods are non-intrusive but detect drowsiness only after driving performance degradation occurs, limiting their effectiveness for accident prevention. The survey identifies vision-based approaches as most suitable for real-world automotive deployment due to their balance of accuracy, non-intrusiveness, and early detection capability.

Vision-based methodologies analyze facial features captured through cameras to identify drowsiness indicators. Key metrics include Eye Aspect Ratio (EAR) measuring eye openness, Percentage of Eye Closure (PERCLOS) tracking blink duration, blink frequency monitoring, yawn detection through Mouth

Aspect Ratio (MAR), and head pose estimation. Traditional computer vision approaches employed handcrafted features including Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP), and geometric facial measurements. These methods achieved early success in controlled settings but exhibited poor generalization across varying illumination, head pose, and occlusion conditions.

The emergence of Convolutional Neural Networks (CNNs) marked a paradigm shift in drowsiness detection. Deep learning architectures including VGG, ResNet, DenseNet, and Inception families automatically learn discriminative features from raw image data eliminating manual feature engineering. CNNs capture texture and symmetry patterns associated with drowsiness achieving superior accuracy compared to traditional methods. However deployment on resource-constrained embedded platforms requires careful optimization through model compression, quantization, and efficient architecture selection.

Temporal modeling using Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Gated Recurrent Units (GRU) exploits sequential dynamics in video streams. CNN-LSTM hybrid architectures combine spatial feature extraction with temporal pattern recognition enabling detection of transient drowsiness indicators such as slow eyelid closure and micro-expressions. The survey notes that video-based approaches improve sensitivity but require larger computational resources and annotated temporal datasets.

Fu et al. identify critical challenges including dataset scarcity, limited demographic diversity in training data, computational constraints for real-time embedded deployment, illumination and occlusion robustness, and cross-subject generalization. The authors emphasize that publicly available datasets such as NTHU-DDD, YawDD, UTA-RLDD, and NITYMED facilitate algorithm development but often lack diversity in ethnicity, age, lighting conditions, and authentic drowsiness scenarios. Dataset imbalance and limited sample sizes lead to overfitting when training deep models.

The survey advocates for three research priorities: creation of demographically representative public datasets with standardized annotation protocols, development of lightweight architectures optimized for edge deployment with real-time inference capability, and incorporation of explainable AI mechanisms to improve clinical and operational trust. Hardware optimization strategies including pruning, quantization, knowledge distillation, and neural architecture search enable deployment on platforms such as NVIDIA Jetson series, Raspberry Pi, and mobile processors.

Fu et al. conclude that vision-based drowsiness detection represents the most practical approach for automotive deployment. The integration of optimized CNN architectures with efficient facial landmark detection frameworks enables real-time monitoring on embedded GPU platforms. This survey provides the foundational motivation for implementing an optimized fatigue detection system on Jetson Nano addressing the identified challenges of computational efficiency, accuracy, and deployment feasibility.

B. State-of-the-Art: Florez et al. (2024)

Florez, Palomino-Quispe, Alvarez, Coaquira-Castillo, Herrera-Levano (2024), Sensors [2]. Florez et al. present a high-performance real-time driver drowsiness detection system implemented on NVIDIA Jetson Nano utilizing CNN architectures and MAR analysis. The system represents current state-of-the-art in embedded fatigue detection achieving 99.88% accuracy on test data with 14 frames per second throughput on edge hardware.

The methodology introduces an optimized Region of Interest (ROI) extraction technique for eye regions using MediaPipe facial landmark detection. MediaPipe provides 468 facial landmarks enabling precise localization of eye and mouth regions. The authors propose a correction algorithm for ROI extraction ensuring that eye region information is not lost during lateral head movements and inclination. This addresses a critical limitation in existing systems where head pose variation degrades detection accuracy.

The proposed Driver Drowsiness Artificial Intelligence (DD-AI) network comprises three convolutional layers with 50 kernels each using ReLU activation, two max-pooling layers for dimensionality reduction, dropout layers with 0.8 rate for regularization, and fully connected layers with 500 neurons for classification. The architecture is specifically designed for efficient inference on embedded platforms balancing accuracy with computational efficiency.

Performance evaluation compared DD-AI against three transfer learning approaches: InceptionV3, VGG16, and ResNet50V2. Training employed the NITYMED (Night-Time Yawning–Microsleep–Eyeblink–Driver Distraction) dataset containing realistic drowsiness scenarios across diverse subjects. The DD-AI network achieved 99.88% test accuracy outperforming InceptionV3 (98.95%), VGG16 (98.33%), and ResNet50V2 (99.49%). Hardware implementation on Jetson Nano yielded 96.55% accuracy at 14 fps demonstrating feasibility for real-time automotive deployment.

The system integrates Near-Infrared (NIR) camera for robust operation under varying illumination conditions including nighttime driving. NIR illumination ensures consistent facial feature visibility regardless of ambient lighting addressing a critical limitation of RGB-only systems. Drowsiness detection combines eye state classification through CNN with yawn detection using MAR computed from mouth landmarks. Alert mechanisms activate when eye closure exceeds 300 milliseconds or yawning frequency surpasses threshold values.

Key technical contributions include optimized ROI extraction robust to head pose variation, lightweight CNN architecture achieving state-of-the-art accuracy with minimal parameters, real-time implementation on resource-constrained Jetson Nano platform, and integration of NIR imaging for illumination-invariant detection. The work demonstrates that carefully designed lightweight architectures can match or exceed the performance of complex transfer learning models while enabling real-time embedded deployment.

From a state-of-the-art perspective, Florez et al. excel across multiple dimensions: achieving highest reported accuracy on a challenging nighttime-focused dataset, demonstrating real-world deployment on actual embedded hardware with measured performance metrics, providing complete system integration including camera interface and alert mechanisms, and addressing practical challenges of head pose variation and illumination robustness. This work establishes the benchmark for embedded drowsiness detection systems and serves as the primary technical foundation for our implementation.

C. Main Technique (Implementation Base): Florian et al. (2024)

Florian, Popescu, Hossu (2024), Procedia Computer Science [3]. This work presents the core detection methodology we will implement and optimize for our project. Florian et al. developed a tiredness detection system using Jetson Nano employing computer vision and machine learning techniques for real-time facial analysis.

The system architecture comprises six stages: image acquisition capturing facial video in real-time, preprocessing for noise reduction and normalization, face detection using Haar Cascade classifiers, feature extraction focusing on eye region characteristics, machine learning classification for drowsiness prediction, and graphical user interface for result visualization and system tuning. Face detection employs OpenCV Haar Cascade classifier using the haarcascade frontalface default.xml pre-trained model. Haar Cascades provide computationally efficient face localization suitable for embedded platforms though sensitivity to head pose and illumination variation remains a limitation. Following face detection, the algorithm extracts the eye region as the primary ROI for drowsiness assessment. Feature extraction focuses on eye-related characteristics associated with fatigue including eyelid position, eye openness ratio, and blink patterns. The authors investigate multiple CNN architectures for binary classification of eye states (open versus closed). Training utilizes labeled datasets with data augmentation techniques including rotation, horizontal flipping, and brightness adjustment to improve model robustness. The machine learning pipeline evaluates multiple CNN configurations. Model selection considers accuracy, inference latency, memory footprint, and compatibility with Jetson Nano computational constraints. The final model achieves real-time performance enabling continuous monitoring during driving scenarios. Fine-tuning mechanisms allow accuracy optimization through hyperparameter adjustment and additional training on domain-specific data.

System implementation on Jetson Nano leverages CUDA acceleration for CNN inference and GPU-optimized OpenCV operations for preprocessing. The authors emphasize the importance of efficient algorithm design given the limited 4GB memory and thermal constraints of the Jetson Nano platform. Optimization strategies include model quantization, input resolution reduction, and batch processing where applicable.

The graphical interface displays real-time detection results overlaying drowsiness indicators on the video feed. Alert mechanisms activate upon detecting sustained eye closure or patterns indicative of fatigue. The system provides user-configurable thresholds allowing adaptation to individual driver baseline characteristics and sensitivity preferences.

Key engineering insights include the trade-off between model complexity and inference speed on embedded platforms, the importance of robust preprocessing for handling illumination variation, and the need for careful threshold tuning to minimize false alarms while maintaining high sensitivity. The work demonstrates practical feasibility of deploying CNN-based drowsiness detection on affordable embedded hardware suitable for automotive integration.

This methodology forms the technical foundation for our implementation. We will build upon this framework by incorporating MAR analysis for yawn detection, implementing the optimized ROI extraction from Florez et al., and applying advanced optimization techniques including TensorRT acceleration, FP16 quantization, and model pruning to maximize performance on Jetson Nano.

D. Other Relevant Works

Below we discuss four additional verified references that inform different aspects of our project implementation and provide context for design decisions.

Prasath et al. (2022) [4] implemented machine learning algorithms focusing on eye closure and yawning ratios for drowsiness detection. The work emphasizes the complementary nature of multiple drowsiness indicators. Eye closure patterns alone may not distinguish drowsiness from normal blinking in all cases. Incorporating yawn detection through mouth region analysis improves system sensitivity and reduces false negatives. The study employs traditional machine learning classifiers demonstrating that effective detection does not always require deep learning when appropriate features are engineered. This work informs our decision to implement multi-modal detection combining eye state classification with MAR-based yawn detection.

Rathod et al. (2023) [5] developed RealD3 system utilizing MediaPipe Face Mesh for 468 facial landmarks and YOLO for object detection achieving 94% overall accuracy. The system computes EAR and MAR metrics with PERCLOS classification distinguishing drowsiness from normal blinking based on temporal patterns. Multi-modal feature fusion combining eye state, mouth opening, hand position (hands covering face), and accessory detection (sunglasses) enhances robustness against diverse driving scenarios.

Alert mechanisms activate when EAR falls below 0.25 threshold for sustained duration or MAR exceeds yawning threshold. The work emphasizes importance of temporal filtering to avoid false alarms from brief eye closures during normal blinking. PERCLOS (Percentage of Eye Closure) metric tracking eye closure duration over sliding time windows provides more reliable drowsiness indication than instantaneous measurements. This temporal aggregation approach will be incorporated into our implementation to improve detection reliability.

The integration of YOLO object detection enables recognition of contextual factors affecting facial analysis such as sunglasses occlusion or hands covering face. These scenarios require system adaptation or alert suppression to prevent false positives. RealD3 demonstrates the value of comprehensive scene understanding beyond isolated facial feature analysis. We will explore similar contextual awareness mechanisms in our implementation.

Fouad (2023) [6] demonstrated EEG-based drowsiness detection using machine learning algorithms including Linear Discriminant Analysis (LDA), Support Vector Machines (SVM-RBF and SVM-Linear), K-Nearest Neighbors (KNN), and Random Forest achieving 100% accuracy across twelve subjects. The study utilized EEG signals from only three electrodes (Fp1, Fp2, Fpz) demonstrating that effective drowsiness detection does not require full-scalp electrode arrays.

Feature extraction focused on spectral power in delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), and beta (13-30 Hz) frequency bands. Drowsiness correlates with increased theta and alpha power combined with decreased beta power. The study employed comprehensive preprocessing including bandpass filtering, artifact removal, and epoch segmentation with one-second windows.

While physiological signal-based methods achieve exceptional accuracy, the intrusive nature of electrode placement limits automotive applicability as discussed in the survey [1]. Drivers are unlikely to tolerate wearing EEG caps during routine driving. However this work establishes performance benchmarks that vision-based systems should approach. The 100% accuracy achieved with EEG provides an upper bound for drowsiness detection performance and motivates continued optimization of vision-based approaches.

The study also highlights the importance of subject-independent validation. Cross-subject generalization remains challenging as individual differences in baseline EEG patterns, facial morphology, and drowsiness manifestation require robust model training. Our implementation will employ cross-validation strategies and diverse training data to ensure generalization across different drivers.

Fu et al. (2024) - Survey [1] provides additional context beyond the detailed discussion in Section II-A. The survey identifies hardware optimization as a critical enabler for embedded deployment. Model compression techniques including pruning remove redundant parameters reducing memory footprint and computational requirements. Quantization converts floating-point weights to lower precision (FP16 or INT8) accelerating inference with minimal accuracy degradation. Knowledge distillation transfers knowledge from complex teacher models to compact student models suitable for edge deployment.

Neural architecture search (NAS) automates discovery of efficient architectures optimized for specific hardware constraints. Platforms such as NVIDIA Jetson benefit from GPU-accelerated inference libraries including TensorRT providing automatic kernel optimization, layer fusion, and precision calibration. The survey emphasizes that deployment success requires co-design of algorithms and hardware optimization strategies rather than treating them as independent concerns.

The survey also discusses emerging trends including federated learning for privacy-preserving model training across distributed vehicles, self-supervised learning reducing reliance on labeled data, and multimodal fusion combining vision with audio analysis of speech patterns or vehicle sensor data. These directions represent promising avenues for future enhancement beyond our current implementation scope.

References

[1] B. Fu, F. Boutros, C.-T. Lin, and N. Damer, "A survey on drowsiness detection – modern applications and methods," 2024. [Online]. Available: https://arxiv.org/abs/2408.12990

[2] R. Florez, F. Palomino-Quispe, A. B. Alvarez, R. J. Coaquira-Castillo, and J. C. Herrera-Levano, "A real-time embedded system for driver drowsiness detection based on visual analysis of the eyes and mouth using convolutional neural network and mouth aspect ratio," *Sensors*, vol. 24, no. 19, 2024. [Online]. Available: https://www.mdpi.com/1424-8220/24/19/6261

[3] N. Florian, D. Popescu, and A. Hossu, "Real-time tiredness detection system using nvidia jetson nano and opencv," *Procedia Computer Science*, vol. 242, pp. 536–543, 2024, 11th International Conference on Information Technology and Quantitative Management (ITQM 2024). [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050924018209

[4] P. Nithiyanandam, J. Sreemathy, and V. Pandi, "Driver drowsiness detection using machine learning algorithm," 03 2022, pp. 01–05.

[5] S. Rathod, T. Mali, Y. Jogani, N. Faldu, V. Odedra, and P. K. Barik, "Reald3: A real-time driver drowsiness detection scheme using machine learning," in *2023 IEEE Wireless Antenna and Microwave Symposium (WAMS)*, 2023, pp. 1–5.

[6] I. A. Fouad, "A robust and efficient eeg-based drowsiness detection system using different machine learning algorithms," *Ain Shams Engineering Journal*, vol. 14, no. 3, p. 101895, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2090447922002064