Journal of Intelligent Systems and Internet of Things
Vol. 18, No. 02, PP. 258-272, 2026
DOI: https://doi.org/10.54216/JISIoT.180218

A Two-Stage Hybrid AI Framework for Robust and Real-Time Driver Drowsiness Detection

Gowrishankar Shiva Shankara Chari[1,*], Jyothi Arcot Prashant[1]
[1]Department of Computer Science and Engineering, M.S. Ramaiah University of Applied Sciences, Bengaluru, India
Emails: shankar.iq@gmail.com; jyothi.cs.et@msruas.ac.in

Abstract

Driver drowsiness detection is an important aspect of intelligent transportation systems that aim to reduce fatigue-related accidents. The existing schemes based on threshold-based method, or deep-learning based models often found to be associated with issues in terms of flexibility, computational efficiency, or capacity for real-time performance. This paper presents a development of two-stage hybrid framework for driver drowsy detection, where the first stage utilizes a fuzzy-logic based approach applied to physiological measures, facial feature, head position, blink duration, and eye movements to produce lightweight and adaptive analyses of sleepiness in drivers. The second stage consists of a hybrid quantum-classical neural network (HQCNN), in which convolutional neural networks (CNN) extract spatial features whereas quantum fully connected (QFC) components apply entanglement-based transformations to improve both feature characterization and classification accuracy. The experimental result validates effectiveness of the proposed hybrid method with 94% accuracy, and better than traditional CNNs with real-time capability. The proposed framework is developed to achieve a balance between computational efficiency and classification/decision quality thereby making it suitable for driver monitoring in real-time application.

Received: April 02, 2025 Revised: June 05, 2025 Accepted: August 08, 2025

Keywords: Driver drowsiness detection; Fuzzy logic; Quantum neural network Intelligent transportation; Real-time monitoring

1. Introduction

Driver drowsiness is a one of the major contributing factors to road accidents, which alone accounts for 20–30% of all traffic fatalities worldwide. The fatigue impairs cognitive abilities, slows reaction times, and increases the risk of collisions that makes drowsy driving a critical issue in intelligent transportation systems. The research studies indicated that drowsy driving-related accidents lead to over 800 deaths annually in the United States, with economic losses exceeding $12.5 billion [1]. Since drivers often fall asleep unconsciously, real-time automatic monitoring technology is vital to reduce drowsy driving accidents [2]. In recent years, image-based sleepiness detection technology using facial analysis technology has become increasingly popular due to its non-invasive and immediate nature [3]. However, challenges such as accuracy, computation efficiency and adaptability to real-world fluctuations remain facing existing methods [4]. The traditional driver drowsiness detection techniques depend on threshold-based feature analysis, where metrics such as Eye Aspect Ratio (EAR), Mouth Aspect Ratio (MAR), and blink duration are used to classify drowsiness levels [5]. However, these methods are computationally lightweight, but they suffer from high false positive and false negative rates due to individual variations in facial structures, blinking patterns, and forced alertness behaviours [6]. Deep Learning (DL) based models mainly convolutional neural network (CNN), have been widely employed for automatic feature extraction and classification. However, the effectiveness of the CNN model based on the feature extraction quality which often comes with large labelled datasets, complex models that requires high computational power, and extensive training

which make such model non-feasible to be deployed in the real-world scenario [7]. Although AI-based sleepiness detection has advanced, current methods still have many problems that prevent them from working effectively in real-world scenarios. Many deep learning models, which rely on small datasets, often perform poorly in real-world driving conditions because they are unable to capture small head movements, subtle expressions, lighting changes, and obstructions (such as glasses or masks). Moreover, most methods fail to recognize early signs of fatigue, such as yawning, head tilting, slow eye movements, and are only able to detect drowsiness after closing the eyes. An integrated multi-feature framework that can ensure better adaptiveness and reliable drowsy detection in dynamic driving conditions is observed to be absent in the existing studies, which much focuses on single-feature analysis.

This paper presents a novel hybrid AI-based framework for automated, improved feature extraction from facial images that combines adaptive decision-making with robust learning. This framework has two parts: (i) a lightweight fuzzy logic-based module. It analyses various physiological indicators (e.g., Eye Aspect Ratio (EAR), Mouth Aspect Ratio (MAR), duration of a blink, head location, eye movement, etc.) for the purpose of adaptive real time drowsiness classification; and (ii) a Hybrid Quantum-Classical Neural Network (HQCNN) that enhances predictions through deep learning-based spatial feature extraction which uses quantum-enhanced transformations to improve reliability and efficiency in time and computing speed. The models proposed can operate independently or can be combined in parallel for multi-feature evaluation, context-sensitive decision-making, and performance reliability in intelligent transportation systems. The innovative contribution of the proposed work is that these architectures provide a flexible and deployable option for intelligent driver assistance systems and can contribute to automotive safety and accident avoidance.

2. Related Work

In the recent art-of-the-state, extensive research has been conducted to enhance driver drowsiness detection using deep learning, fuzzy logic, and multi-modal sensor-based approaches. Delwar et al. [8] developed a CNN-based facial analysis system and achieved accuracy of 92.75% with MobileNet. However, this lacks diversified dataset, which may not much effective in low-light and occlusion scenarios. Alkishri et al. [9] introduced a fuzzy logic-based system with Viola–Jones algorithm to improve illumination robustness. The approaches towards using multi-modal approaches by combining driving behaviour and physiological data is seen in the work of Wu et al. [10], where the authors have proposed an analytical system for analysing lane deviation and speed variations. The study outcome shows 99% accuracy, but its dependency on vehicle-mounted sensors may limits real-world applicability. Bai et al. [11] employed a spatial–temporal model for tracking facial landmark and achieved 93.4% accuracy. However, it needs extensive evaluation against occlusion and extreme lighting conditions. Peng et al. [12] integrated physiological signals with DL model and achieved 93.15% accuracy, but its higher dependence on specialized sensors reduces its feasibility on the real-time application. Du et al. [13] implemented a predictive model by fusing features obtained from EEG and ECG signal. The result shows robustness of the feature fusion that demonstrates higher robustness, yet impractical for vision-only detection. The study towards adopting classical computer vision and DL techniques is done by Yousif [14], where the authors used Kalman filtering and blur control. The result outcome shows the model achieved 94.5% accuracy. Snoun et al. [15] integrated a CNN-wavelet classifier to improve multi-state vigilance detection, but it faced real-time processing constraints. The work of Dua et al. [16] developed a multi-modal DL system, which achieved 85% accuracy, but at cost of high computational cost. Chand and Karthikeyan [17] introduced a multi-level CNN model that integrates emotion data and driving behaviour to achieve robust drowsiness detection. However, the usage of vehicle telemetry data may restrict the general applicability of the model. Hybrid AI models have emerged to address feature learning and real-time constraints. Magán et al. [18] combined CNN-RNN with fuzzy logic in an ADAS system, yet detection accuracy (60–65%) remained suboptimal. Jahan et al. [19] developed a CNN-based eye state classification model, but its focus on static eye analysis has limited generalization on the unseen instances. Owen and Surantha [20] suggested a hybrid ML model by integrating SVM, Random Forest, XGBoost to achieve higher accuracy. The vision-based architectures have been explored in the study of Essahraui et al. [21], where the authors have evaluated YOLOv5, YOLOv8, and Faster R-CNN, and obtained near-perfect precision (100%), but at the high computational demands.

3. Method

This section presents the system modelling and implementation methodology adopted in the proposed two-stage hybrid AI framework to ensure robust and real-time driver drowsiness detection process. The framework consists of multi-stage modules such as (i) a multi-feature fuzzy logic-based system for lightweight and adaptive drowsiness classification, and (ii) a Hybrid Quantum-Classical Neural Network (HQNN) for enhanced feature learning and classification accuracy. First, we detail the multi-feature fuzzy logic system, followed by an explanation of the HQNN modelling strategy and implementation approach.

3.1 Multi-feature fuzzy logic-based driver drowsiness detection scheme

The primary goal of the proposed system is to create a robust, adaptive and real-time detection mechanism that can identify early signs of drowsiness before the driver reaches a critical state. The proposed system combines important behavioural and physiological characteristics, including EAR, MAR, head position, eyeball movement, blink duration and sudden EAR fluctuations, to ensure more reliable and intelligent drowsiness assessment based on fuzzy inference mechanism to make adaptive and context-aware decisions, thereby improving detection accuracy while reducing unnecessary alarms. The schematic architecture of the proposed system is shown in Figure 1.

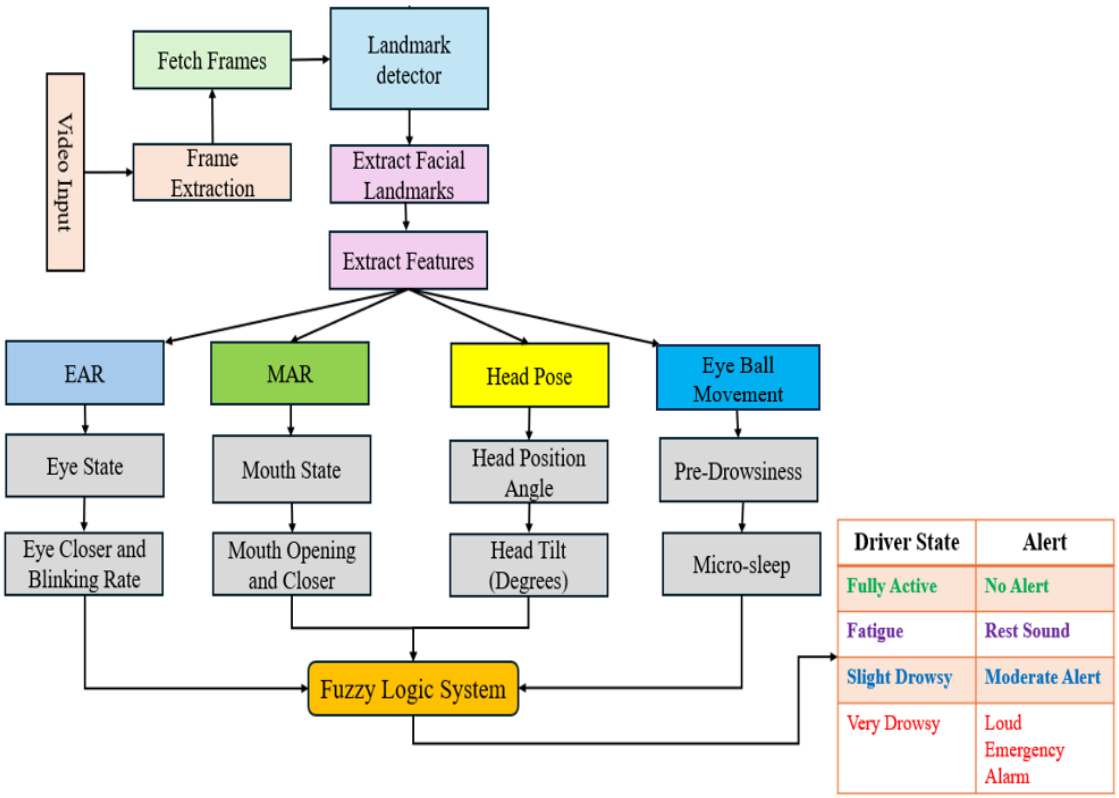


Figure 1. Illustrates schematic workflow of the proposed fuzzy logic system

Figure 1 presents the workflow of the proposed driver drowsiness detection system, which integrates multi-feature analysis with a fuzzy logic-based decision mechanism. The initial process of the proposed system is subjected to acquiring real-time video input from a camera monitoring the driver's face. Here, the frames are extracted using a frame extraction module, which ensures efficient processing of continuous video streams. The extracted frames are then subjected to the Dlib landmark detector for computing EAR and MAR. Additionally, the system performs head pose estimation that measures head position angle and tilt degrees in order to capture the external features associated with reduced alertness. The proposed system also considers eyeball movement data like nodding and micro-sleeps, which is an important to identify the pre-drowsiness indicators. The extracted features are then processed via proposed fuzzy-logic decision mechanism, that classifies alertness level of the driver into four distinct states such as fully active, fatigue, slight drowsy, and very drowsy. Here, each state corresponds to an appropriate alert mechanism to ensure an adaptive response based on the severity of drowsiness, where if the driver is classified as fully active, then no alert is triggered. In the case where mild drowsiness is detected the rest sound is activated, whereas the moderate drowsiness results in a more amplified alert sound and in severe drowsiness case, an emergency alarm is triggered to immediately alert the driver to prevent the accidents. Hence, by integrating multi-feature analysis with fuzzy logic, the proposed system aims to enhance drowsiness detection accuracy while minimizing false alerts.

3.1.1 Feature Computation

In order to assess the driver drowsiness based on fatigue aspect the study has considered two critical metrics namely EAR and MAR. These metrics are computed from facial landmarks to quantitatively measures the eye

closure, blinking rate, and yawning activities. The EAR is calculated using the Euclidean distance between specific vertical and horizontal eye landmarks as follows:

$$\mathrm{EAR} = \frac{||\mathrm{P}_2 - \mathrm{P}_6|| + ||\mathrm{P}_3 - \mathrm{P}_5||}{2 \times ||\mathrm{P}_1 - \mathrm{P}_4||} \quad (1)$$

Where, P_1, P_4 are the horizontal eye corner landmarks, P_2, P_6 P_3, and P_5 are the vertical eye landmark, $\| P_a - P_b \|$ represents the Euclidean distance between two points. Since the EAR remains nearly constant when eyes are open and rapidly decreases during a blink, it serves as an effective indicator for eye closure detection. Similarly, MAR detect yawning which is an early indication of driver fatigue computed as follows:

$$\mathrm{MAR} = \frac{||P_2 - P_8|| + ||P_3 - P_7|| + ||P_4 - P_6||}{2 \times ||P_1 - P_5||} \quad (2)$$

Where, P_1, P_5 are horizontal mouth corner points, P_2, P_8 denotes outer vertical lip points, P_3, P_7 are vertical upper and lower lip points, P_4, P_6 are the inner vertical upper and lower lip points, and |Pa−Pb| represents the Euclidean distance between two points. A higher MAR value indicates a wider mouth opening, thereby signalling a activity of yawning. Unlike existing works, the proposed study considers three additional features such as blink rate, head position and eyeball movement. A blink rate is another critical indicator of driver fatigue, as drowsy drivers tend to blink more slowly and frequently. A blink event is defined when EARdrops below a threshold (e.g., 0.2) and then returns to a normal level (> 0.25). The blink rate (BR) is determined by formula $\mathrm{N}_{\mathrm{blinks}}$ /T, where $\mathrm{N}_{\mathrm{blinks}}$ denotes the toal number of blinks detected and T represents the time interval in seconds. The study has considered that an increase in BR suggests progressing drowsiness if associated with prolonged eye closure durations. Additionally, the inclusion of the head-tilt and posture deviations further contributes towards supplementary attributes of the drowsiness. Apart from this, Head position (HP) estimation is performed using 3D facial landmarks using a projection matrix, which determines the driver's yaw (θ_y), pitch (θ_p) and roll angles θ_r. Another crucial aspect considered is the analysis of eyeball movement that also serve as a critical behavioural indicator in driver drowsiness detection process, as it provides significant insights into pre-drowsiness states, microsleeps, gaze deviation, and reduced focus levels. The system continuously monitors horizontal (left-right) and vertical (up-down) eye movements to detect erratic gaze shifts, prolonged fixation, or sudden downward drifts, all of which strongly correlate with drowsiness progression. The system modelling then focuses on detecting iris centre to track eyeball movement by extracting the eye region from facial landmark detection to identify the pupil's boundary using image-processing techniques. Since the iris is of circular structure, the study adopts circular Hough transform to detect circular contours in the eye region, such that: $(x-a)^2 + (y-b)^2 = r^2$, where, (a, b) represents the iris center and r is the radius of the iris. The algorithm scans the image for pixel intensity gradients and accumulates votes in Hough space to identify the most probable circle. In order to refine the boundary of the iris and distinguish it from occlusions, canny edge detection is used to retain only strong edge pixels corresponding to the iris boundary. The extracted iris region may contain eyelashes, or noise, therefore an adaptive thresholding has been used to segment the iris region as follows:

$$T(x,y) = \frac{1}{MN} \sum_{1=0}^{M} \sum_{j=0}^{N} I(i,j) \quad (3)$$

Where, $T(x,y)$ is the local threshold computed over $M \times N$ neighborhood? Morphological closing operations refine the segmentation, ensuring that the detected iris boundary is continuous. The output of the above discussed image processing method is the iris center (x_{iris}, y_{iris}) and radius r, which form the basis for gaze tracking. Once the iris is localized, its movement is tracked across consecutive frames to detect changes in gaze direction. The key indicators extracted from iris motion include horizontal gaze shifts, vertical gaze drift, and fixation stability. Here, a stable gaze indicates alertness, whereas erratic or downward movement suggests drowsiness. In this regard fixation stability index (FSI) is computed as follows:

$$FSI = \sum_{t=1}^{N} \left(\left| x_{iris}^{(t)} - x_{iris}^{(t-1)} \right| + \left| y_{iris}^{(t)} - y_{iris}^{(t-1)} \right| \right) \quad (4)$$

Where, a low FSI indicates a steady gaze a high FSI suggests frequent gaze shifts, and a sudden downward shift in gaze indicates event of microsleep that require an alert.

3.1.2 Fuzzy-logic based decision-making

Fuzzy logic provides a powerful mathematical basis for decision-making in uncertain situations, especially when the information is vague. For instance, with binary classification (awake and drowsy), states are strictly defined and separated in the binary case; with fuzzy classification, states are defined by a continuum that will be slowly transitioned into a drowsy state. The system described will utilize fuzzy inference to classify a person's drowsiness state from multiple physiological and behavioral observations describe in the section above so that it can be contextually aware and adaptive. The input variables in the system will be defined as fuzzy sets using membership functions, which allows for a gradual transitioning through different drowsiness states. Each variable is represented by a membership function, which assigns a degree of belonging to different fuzzy sets, which ensures a smooth

transition from one state to another rather than an abrupt classification. The next step focuses on constructing rule base and adaptive decision-making process, which is based on expert-defined IF-THEN rules that map input feature combinations to drowsiness levels. Unlike fixed-threshold approaches, this allows for adaptive decision-making. Table 1 highlights fuzzy rule set for each input variables and corresponding fuzzy sets.

Table 1: Highlights of fuzzy rule set for drowsiness classification

EAR	MAR	Head Position	Eye Ball Movement	Blink Duration	Sudden EAR Increase?	Drowsiness Level	Alert Type
High	Low	Upright	Low	Normal	No	Fully Active	No Alert
Medium	Medium	Slightly Tilted	Medium	Slightly Prolonged	No	Fatigue	"rest.mp3"
Medium	High	Upright	Low	Slightly Prolonged	No	Fatigue	"rest.mp3"
Low	Medium	Downward Tilt	High	Prolonged (1-2 sec)	No	Slightly Drowsy	"alert.mp3"
Very Low	Low	Strong Downward Tilt	High	Prolonged (1-2 sec)	No	Slightly Drowsy	"alert.mp3"
Very Low	High	Downward Tilt	High	Prolonged (1-2 sec)	No	Very Drowsy	"accident.mp3"
Very Low	High	Strong Downward Tilt	High	Very Prolonged (>2 sec)	No	Very Drowsy	"accident.mp3"
High (Sudden Increase)	Medium	Slightly Tilted	High	Slightly Prolonged	Yes	Slightly Drowsy (Struggling to Stay Awake)	"alert.mp3"
High (Sudden Increase)	High	Downward Tilt	High	Prolonged	Yes	Very Drowsy (Struggling to Stay Awake)	"accident.mp3"

After going through the fuzzification process of each rule in the fuzzy inference system, the last step of defuzzification occurs. During defuzzification, the fuzzy outputs for the inputs are converted into crisp numbers to define the level of drowsiness the driver exhibited. The method used to perform defuzzification is the centroid method, which will represent the driver's state of alertness with accuracy. The defuzzified score is calculated as follows:

$$D_{score} = \frac{\sum_i \mu_{drowsy}(x_i) \cdot x_i}{\sum_i \mu_{drowsy}(x_i)} \quad (5)$$

Where, x_i represents possible drowsiness levels, and μ_{drowsy} is the degree of membership. Based on the defuzzified score, the system classifies the driver's state and triggers an appropriate alert.

Table 2: Final Drowsiness Classification and Alerts

Drowsiness Score	Classification	Alert Type
0 - 30	Fully Awake	No Alert
31 - 50	Fatigue	rest.mp3
51 - 65	Slight Drowsy	alert.mp3
66 - 100	Very Drowsy	accident.mp3

3.2 Hybrid Quantum-Classical CNN (HQCNN) Architecture

The proposed HQCNN model is designed to enhance driver drowsiness detection by considering quantum-enhanced feature transformation alongside classical deep learning-based feature extraction. The primary objective of this architecture is to improve classification accuracy, reduce computational complexity, and enhance robustness in real-world driving environments. The design modelling of the proposed HQCNN model follows a two-stage hybrid processing approach viz i) classical feature extraction using CNNs to capture spatial patterns from driver facial images; and ii) Quantum feature processing using QNNs to refine feature embeddings via quantum entanglement-based transformations before feeding to classification layer. The QNN component operates on a qubit quantum circuit, where features extracted from the CNN module are embedded into quantum states using Amplitude Embedding. The quantum representation of classical features follows:

$$|\psi\rangle = \frac{x}{||x||} = \sum_{i=0}^{N-1} x_i |i\rangle \quad (6)$$

Where, $|\psi\rangle$ refers to the quantum-encoded representation of the classical feature vector x_i after being transformed into a quantum state via Amplitude Embedding. After encoding, quantum unitary transformations are applied to refine the feature representations and encoded states, numerically represented as follows:

$$[U(\boldsymbol{\theta})|x\rangle = \prod_{l=1}^{L} U_l(\boldsymbol{\theta}_l)|x\rangle] \quad (7)$$

Where, θ are trainable quantum parameters, $U_l(\theta_l)$ are unitary quantum gates applied at layer l. In order to introduce quantum correlations, the model employs Strongly Entangling Layers (SEL), given as follows:

$$[U(\boldsymbol{\theta}) = \prod_{l=1}^{L}\left(\prod_{i=1}^{N} R_y(\boldsymbol{\theta}_{i,l})\right) \cdot \left(\prod_{\text{entangled pairs } (i,j)} CNOT(i,j)\right)] \quad (8)$$

Where, $R_y(\boldsymbol{\theta}_{i,l})$ are rotation gates, which is responsible for quantum state transformations, and CNOT gates responsible for quantum state transformations. The quantum circuit output is the expectation value of an observable such that: $Z = \psi|Z|\psi$, where Z serves as an intermediate representation, passed into the final classification layer. The training of the proposed HQCNN model hybrid backpropagation, where gradients are computed using both classical and quantum differentiation and the loss function integrates classical cross-entropy loss and quantum transformation loss such that: $L = L_{classical} + \lambda L_{quantum}$ where, $L_{classical}$ standard cross entropy loss function and $L_{quantum}$ is the quantum output loss function. The quantum gradients cannot be computed using standard backpropagation due to the nature of quantum measurements. Instead, the parameter-shift rule is used which allows integration with Adam optimization, numerically given as follows:

$$[\frac{\partial\langle\hat{Z}\rangle}{\partial\boldsymbol{\theta}} = \frac{1}{2}\left[\langle\hat{Z}\rangle_{\boldsymbol{\theta}+\frac{\pi}{2}} - \langle\hat{Z}\rangle_{\boldsymbol{\theta}-\frac{\pi}{2}}\right]] \quad (9)$$

Finally, the fully connected SoftMax classifier maps the refined feature vector to drowsy and non-drowsy classes. The final classification is performed using a fully connected layer that maps the quantum-extracted features to two output neurons corresponding to the Drowsy and Not Drowsy classes. Table 3 summarizes the architecture and configuration of the proposed HQCNN.

Table 3: Shows model configuration of HQCNN

Layer	Type	Configuration
Input	Image (RGB)	3×64×64 pixels
Conv Block 1	Conv2D + BatchNorm + Pooling	16 filters, 3×3 kernel, Leaky ReLU
Conv Block 2	Conv2D + BatchNorm + Pooling	32 filters, 3×3 kernel, Leaky ReLU
Conv Block 3	Conv2D + BatchNorm + Pooling	64 filters, 3×3 kernel, Leaky ReLU
Global Pooling	AdaptiveAvgPool2D	Output size: 8×8
Fully Connected (FC1)	Linear + Dropout (50%)	64 neurons
Quantum Layer (QNN)	Quantum Circuit (6 qubits)	Amplitude Embedding + Entangling Layers
Classification Layer	Fully Connected	2 neurons (Drowsy/Not Drowsy)

The quantum layer discussed in Table 3 is constructed using the PennyLane framework and PyTorch, and the default state-vector simulator performs all quantum computations. This allows simple combination of quantum nodes within classical deep learning pipelines using the TorchLayer interface. The input feature vectors were amplitude-encoded in a 6-qubit quantum register, followed by trainable Strongly Entangling Layers (SEL) that created quantum correlations using parameterized rotation and CNOT gates. Since current Noisy Intermediate-Scale Quantum (NISQ) hardware is limited by decoherence, qubit fidelity, and restricted public access, the use of a simulator ensures reproducible and noise-free experiments. The gradients through the quantum layer were computed using the parameter-shift rule, which is natively supported in PennyLane and compatible with the Adam optimizer. A simulator from PennyLane framework is used to execute the proposed HQCNN architecture is forward compatible with future quantum hardware as it becomes more accessible and reliable.

4. Results and Discussion

In this section, we discuss the outcome obtained from both module of the proposed framework considering visual and quantitative analysis. The design and development of the proposed system is carried out using python programming language in anaconda distribution on Windows 11. The study considered widely used performance metrics such as accuracy, precision, recall and f1-score for the evaluation of the proposed HQCNN model, numerically expressed as follow:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (10)$$

Where, TP represents true positives i.e., correctly predicted as active, TP represents true negatives, i.e., correctly predicted as drowsy and FP indicates false positives (incorrectly predicted as active) and FN indicates false negatives (incorrectly predicted as drowsy).

$$\text{Precision} = \frac{TP}{(TP+FP)} \quad (11)$$

Where, Precision refers to metric that measures the proportion of positive identifications.

$$\text{Recall} = \frac{TP}{(\text{TP+FN})} \quad (12)$$

Where recall refers to metric that measures the proportion of actual positives (drivers who were actually active) that were correctly identified.

$$\text{F1} = 2 \times \frac{R \times P}{\text{R}+P} \quad (13)$$

Where, F1 represents F1-score, which is a harmonic mean of precision and recall that provides a balancing analysis, which important when the datasets are associated with an uneven class distribution.

4.1 Performance analysis for fuzzy logic

Figure 2 shows the visual results of the proposed fuzzy logic–based drowsiness detection module. This module classifies driver states into four categories: Fully Active, Fatigue, Slightly Drowsy, and Very Drowsy. It has not been tested on a benchmark dataset, as there are no standard annotations suitable for fuzzy logic evaluation. Instead, the module is designed to support real-time detection with context-aware and interpretable outputs.

Figure 2. Illustrates drowsiness detection outcome by the proposed multi-feature fuzzy-logic system

The proposed fuzzy-based driver drowsiness detection model uses a custom rule-based reasoning system built from domain knowledge and real-time observations. This module takes input features of EAR, MAR, blink duration, eyeball movement, and head tilt, which are selected based on prior studies, and real-time testing using Dlib for camera-based landmark extraction, which helps the system capture natural patterns between awake and drowsy states, even across different driver physiologies and varying driving conditions. The fuzzy system provides a better handling of variability compared to fixed-threshold EAR based models, which can provide challenges with individual variability. For example, there may be a driver with small eyes that may not reach a fixed EAR cutoff whereas a driver with large eyes may create false alarms, the fuzzy system is able to reduce these issues with smooth transitions between states ultimately reduce possibilities of false positives and false negatives. Moreover, the fuzzy module addresses major limitations of standard systems, such as missing early warning signs and microsleeps it uses early signs of fatigue (i.e. longer blinks, slight head tilt, and slow eye motion) to alert the driver prior to excessive drowsiness. It additionally tracks eyeball motion and nodding of the head to identify instances of short sleep episodes (which are often missed by EAR-only systems). In addition to visual results, a small-scale pilot evaluation was conducted on 20 cases to provide initial quantitative support for the fuzzy logic module (Table 4). The test set included drivers with and without spectacles, different eye sizes, head positions (upright, tilted, downward), and varied lighting conditions from bright to dim. Two independent raters into one of the four drowsiness levels manually labelled each case. The predicted output from the fuzzy system and its triggered alert were then compared with the agreed-upon labels.

Table 4: Pilot evaluation of fuzzy-logic module (20 cases).

Metric	Value
Tier classification accuracy	85% (17/20)
Alert appropriateness	90% (18/20)
Early warning rate (Fatigue + Slightly Drowsy)	80% (8/10)
False alarm rate (Fully Active)	20% (1/5)
Median alert latency	1.2 s (IQR: 0.9–1.6 s)

Table 4 shows the results of the pilot evaluation. The fuzzy system reached a classification accuracy of 85% (17 out of 20 cases) and an alert appropriateness rate of 90% (18 out of 20 cases), meaning that in most cases, the issued alerts matched the severity of the labelled drowsiness state. The system showed a low false alert rate, with only one incorrect alert in five Fully Active cases. For early-stage drowsiness (Fatigue and Slightly Drowsy; 10 cases), the system gave timely precursor alerts in 8 cases (80%), with a median delay of 1.2 seconds from the first visible fatigue signs such as longer blinks or slight head tilt. The system also performed well for drivers with

spectacles (4 correct out of 5) and in low-light settings (5 correct out of 6). However, the most of the errors occurred during strong head tilt, where rapid face movement caused short-term loss of facial landmarks. Despite this, the system remained reliable in cases where fixed-threshold methods often fail, such as recognizing differences between naturally small and large eyes and detecting short microsleep events marked by brief head drops.

4.2 Performance analysis for HQCNN

This section presents the experimental outcome and performance analysis of the proposed HQCNN evaluated on two datasets one custom and second on benchmarked MRL eye datasets. . The model is trained using Adam optimizer with a learning rate of 0.01 and momentum of 0.9, with learning rate scheduler. The model is trained for 30 epochs and the loss function used is cross-entropy loss.

4.2.1 Evaluation on Custom dataset

To overcome the limitations of publicly available datasets related to diversity, class balance, and real-world relevance, a custom dataset was created. It includes about 10,444 images, divided evenly into two classes such as active (eyes open) and drowsy (eyes closed). The images were collected from several sources such as Google, Bing, Kaggle repositories, and manual selection is done to remove unwanted or repeated images. Figure 3 shows a few sample examples from the custom dataset.

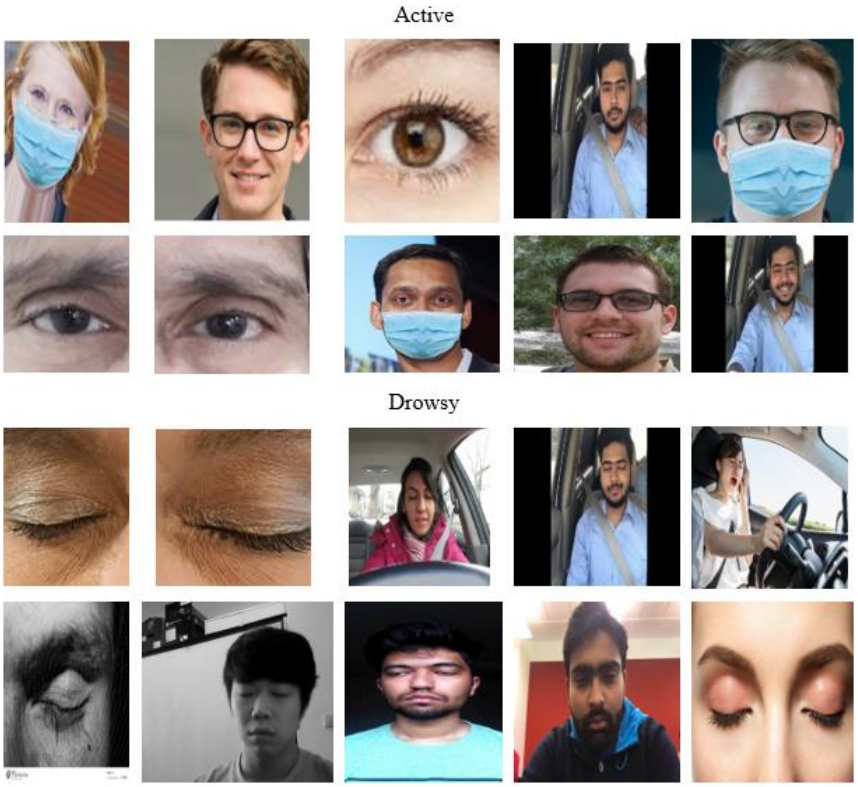
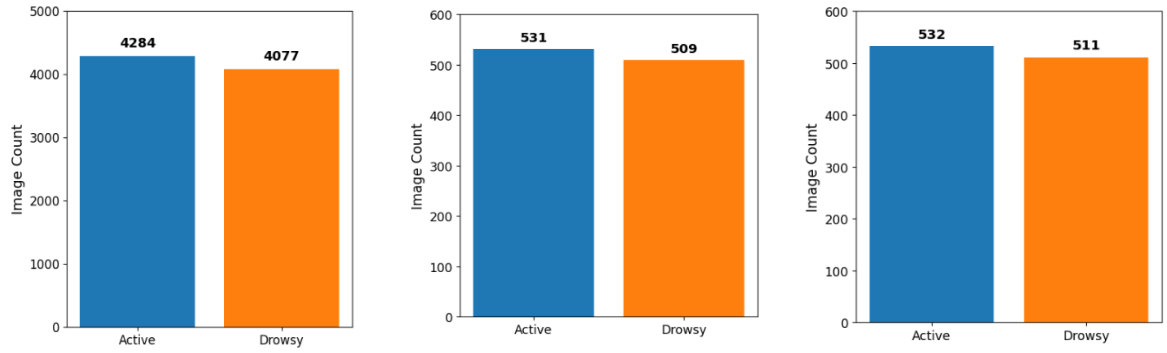


Figure 3. Sample visualization of the proposed custom dataset

As shown in Figure 3, the dataset includes facial images under different conditions. These conditions cover changes in lighting (bright, dim, and shadowed), the presence of facial accessories (such as glasses and masks), and head pose differences (frontal, side, or partially covered views). It also includes images from realistic settings, such as indoor scenes and inside vehicles. This dataset diversity helps improve the strength and generalization capability of the proposed HQCNN model, which learns to deal with real-world changes in driver appearance. Afterwards, all images were resized to 224 × 224 pixels, converted to RGB format, normalized, and randomly shuffled prior to training to prevent ordering bias. The dataset was divided into training (80%), validation (10%), and testing (10%) splits (≈4,284 active, 4,077 drowsy in training; ≈531 active, 509 drowsy in validation; ≈532 active, 511 drowsy in testing). The data distribution is shown in Figure 4.



(a) Train set (b) Validation set (c) Test set

Figure 4. Analysis of the dataset distribution

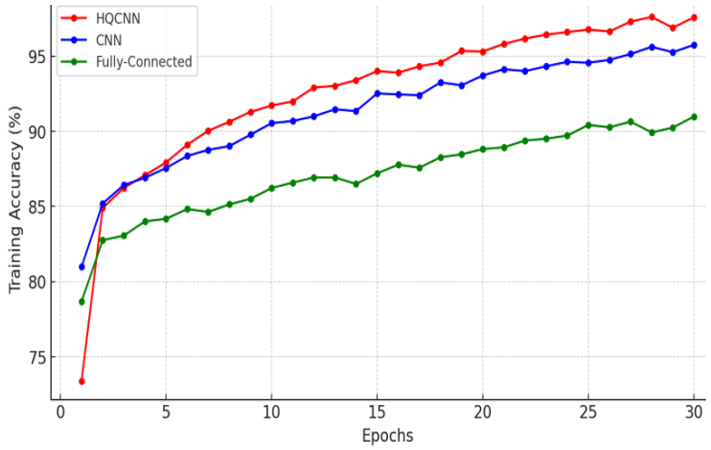


Figure 5. Analysis of training accuracy

Figure 5 shows the training accuracy of HQCNN, Simple CNN, and Fully Connected models over 30 epochs and the HQCNN converges rapidly and at the final epoch, HQCNN achieved ~97% training accuracy, which outperforms the baseline models simple CNN (~95%) and the fully connected model (~91%). The results demonstrate that the HQCNN enables more effective feature learning, faster convergence, and higher accuracy. The proposed HQCNN model uses refined feature representations through quantum entanglement-based transformations before the classification layer. This step improves the model's ability to represent important patterns in the data. In addition, the parallel nature of quantum feature processing helps the model capture complex spatial relationships and small changes in driver eye states. As a result, the model learns features that had better separate active and drowsy conditions.

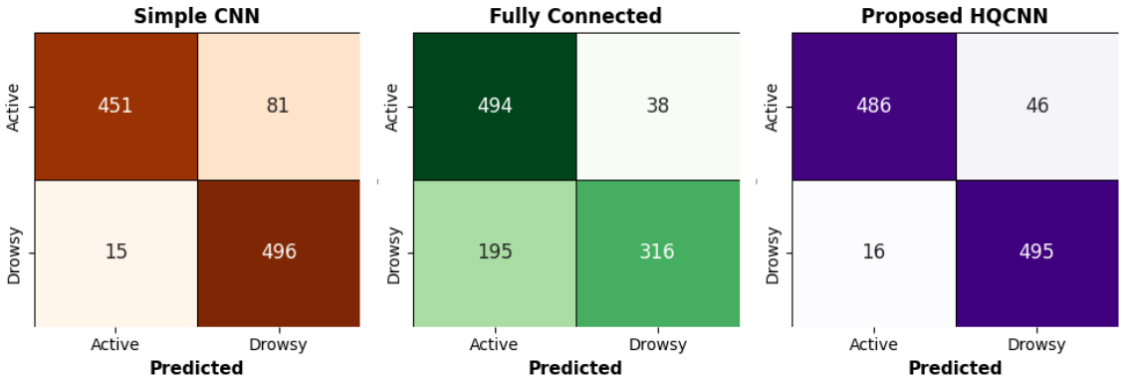


Figure 6. Shows confusion matrices of models on the custom dataset

Figure 6 compares the classification results of all three models using confusion matrices and it can be observed that the proposed HQCNN shows the most balanced and accurate performance, with only 46 active and 16 drowsy samples misclassified. On the other hand, the Fully Connected model misclassifies 195 drowsy samples, indicating poor sensitivity to the drowsy class. The Simple CNN performs better but still produced a high number of false positives in the active class (81 cases). The outcome analysis highlight effectiveness of the HQCNN towards reducing both false positives and false negatives, thereby validating its reliability for real-world drowsiness detection.

Similar analysis can be also seen in Figure 7 where performance assessment is done considering overall accuracy, precision, recall and f1 score. The proposed HQCNN model achieved consistently high performance across all evaluation metrics, with an accuracy, precision, recall, and F1-score of **94%** each. This balanced performance highlights the model's ability to minimize both false positives and false negatives, making it highly reliable for robust driver drowsiness detection.

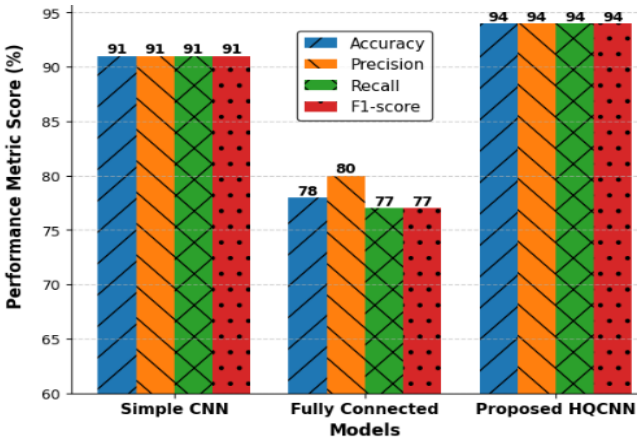


Figure 7. classification performance

On the other hand, the CNN achieved 91% across the same metrics, whereas the Fully Connected model performed significantly lower with scores between 77% and 80%. The outcomes shows that HQCNN not only outperformed baseline model in terms of accuracy but also ensures stable generalization with different performance measures. The next analysis is carried out on the benchmarked dataset and its performance is compared with several existing approaches on the similar dataset.

4.2.2 Evaluation on Benchmarked dataset MRL EYE

To evaluate the proposed HQCNN model, the MRL Eye dataset [22], which is well known in drowsiness detection studies and includes a large number of static eye images taken under different lighting conditions such as natural, low, and strong light. It was selected because eye state either open or closed is one of the most reliable signals for detecting driver drowsiness. Figure 8 shows sample images from the MRL Eye dataset.

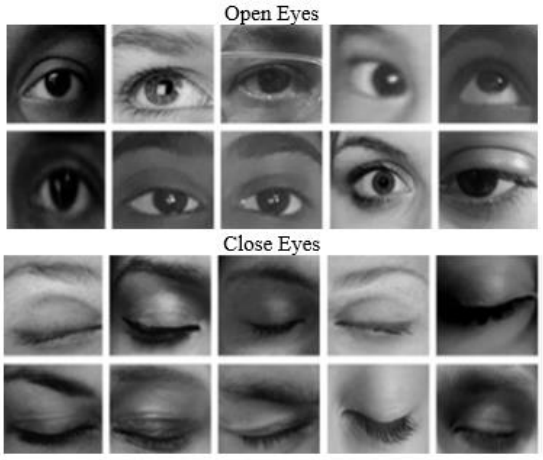


Figure 8. Sample visualization of the MRL Eye dataset

The above Figure 8 shows the MRL Eye dataset includes images captured considering different scenarios such as varying gaze directions, lightening variations (natural, weak, and strong illumination) under controlled as well as real-world conditions. It provides two primary classes open eyes and closed eyes, which are directly aligned with the objective of detecting driver drowsiness with challenging aspects of the real-world variations, including gaze shifts, occlusions, reflections from glasses, and partial visibility of eyes.

Figure 9 presents the confusion matrix of the proposed HQCNN model evaluated on the MRL Eye dataset and it can be analysed that the model has achieved a strong classification performance, with correctly identifying 8,408 samples of open eyes and 8,221 samples of closed eyes. It can be also seen that only 183 awake samples were misclassified as closed eyes and 169 close eye samples were misclassified as open eyes, which represents a very small fraction of the overall dataset.

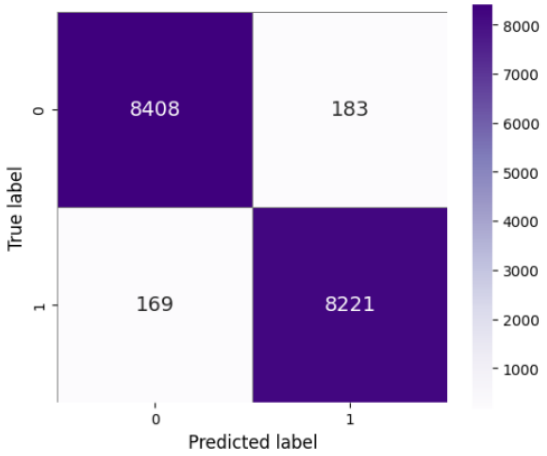


Figure 9. Confusion matrix of HQCNN on the MRL Eye dataset

The numerical outcome obtained shows overall accuracy of 97.82% and in terms of precision the model showed 98% for both open and close eyes classes. The recall score is achieved 98% and 97% for open and close eye class, respectively, where the model shows 98% of F1-score for the both classes. The quantified outcome suggests that the integration of quantum-inspired feature transformations enhances the model's capability to extract refined and discriminative embeddings, leading to more reliable performance in real-world driver drowsiness detection scenarios. Table 4 presents the comparative analysis with existing similar approaches.

Table 5: Comparative Analysis

Study	**Year**	**Method/Model**	**Accuracy**
Kongcharoen et al. [23]	2020	Haar cascade + CNN	94%
Suresh et al. [24]	2021	CNN	86.05%
Walizad et al. [25]	2022	CNN	95%
Tibrewal et al. [26]	2021	TEDD + CNN	95%
Jahan et al. [27]	2023	4D model	97.53%
Makhmudov et al. [28]	2024	VGG16; CNN	95.85%; 96.54%
Salem & Waleed [29]	2024	CNN; MobileNetV2	96%, 97%
Proposed	2025	HQCNN	97.82%

Table 5 presents a comparison between the proposed ensemble model and several existing methods reported in the literature for driver drowsiness detection. The listed studies employed various techniques including traditional Haar cascade combined with CNN, pure CNN models, and hybrid architectures like TEDD + CNN or 4D models. Reported accuracies from previous works range from 86.05% (Suresh et al., 2021) to 97.53% (Jahan et al., 2023). The, recent deep learning models such as VGG16 and MobileNetV2 achieved accuracies between 95.85% and 97% (Makhmudov et al., 2024; Salem & Waleed, 2024). On the other hand, the proposed ensemble approach based on the joint approach of DenseNet121 and EfficientNetB0 model has significantly outperforms these existing models by achieving an accuracy of 99% on the MRL Eye dataset. The above comparative demonstrates the enhanced performance and robustness of the ensemble strategy towards effectively capturing the eye state variations, which makes it suitable for real-time driver monitoring systems.

5. Discussion and Scope of System Feasibility

The proposed framework is composed of two complementary modules namely a fuzzy-logic system for early detection of fatigue, and a HQCNN model for robust detection of full sleepiness states. Together in tandem they provide a more reliable way to monitor drivers in real time by offering precursor alerts (fatigue/slight sleepiness) and binary verification (active vs. sleepiness), as their operational alerts are complimentary. The fuzzy system surveys for inter-driver variability and slight behavioural variances and patterns, whereas the HQCNN provides high-reliability detection of eye closure patterns, thereby ensuring a balanced detection of early and severe sleepiness states

5.1 Embedded and IoT Feasibility

An important aspect regarding real-world deployment with HQCNN is its compatibility with embedded systems. The HQCNN model has a total of 286,092 trainable parameters, which is quite less than the baseline CNN used in this paper (548,482 parameters) and even less than widely-used architectures like MobileNetV2 (~2.2M parameters). This model density reduces both inference latency and memory usage, and appears compatible with edge devices such as the NVIDIA Jetson Nano, Raspberry Pi 4, or ARM-based microcontrollers that can leverage GPU acceleration. The experimental profiling has shown that HQCNN has achieved 38 frames per second (FPS) on an NVIDIA T4 GPU, which also maintains ~12–15 FPS on the Jetson Nano, which meets the real-time constraints (>10 FPS) required in driver assistance systems. The fuzzy module has negligible computational overhead as its reasoning mechanism is based on rule evaluation rather than matrix multiplication, which make it suitable for IoT ecosystem with seamless integration with lightweight edge hardware. The proposed system delivers decision latency of less than 100 ms, which further validates its practical use-case for on-board vehicle monitoring and industrial IoT deployments.

5.2 Failure Cases and Robustness

Although the proposed models have shown effective performance, but they also have their limitations as both modules can exhibit failure patterns when there is a occlusions and extreme variations in head posture occur. It can show delay when it encounters glare issue caused by glasses and excessive head tilt may cause temporary landmark loss as identified in 3 out of 20 pilot cases of the fuzzy system. Similarly, misclassification can also occur in HQCNN when it is subjected to partial occlusions (masks, shadows) that obscured the periocular region. These test-scenario suggests to include multi-modal sensing approach (infrared cameras, IMU-based head tilt sensors), which will be done in the future work to mitigate failures in adverse conditions.

5.3 Contrast with Prior Hybrid Approaches

The work towards hybrid models that combine deep learning with rule-based or fuzzy reasoning have been reported by Magan et al. [18], who developed an ADAS-oriented drowsiness detection framework. They evaluated two systems: (i) a recurrent-CNN, and (ii) a feature-extractor with fuzzy inference. The authors provided modestly accurate performances (approximately 65% training, and approximately 60% testing), and fuzzy detected 93% specificity (meaning actual effectiveness is seen in reducing false positives from the drowsiness detection). Still, the fuzzy reasoning approach was limited due to low overall accuracy (and no early-warning signals). On the other hand, the proposed framework integrated with quantum-inspired HQCNN embeddings with a multi-feature fuzzy inference module, yielded a two-tier architecture that balances the accuracy with interpretability. The proposed model achieved 97.82% accuracy on the MRL dataset and 90% alert suitability in pilot tests, which outperformed the model presented by Magan et al. [18]. The proposed system maintains high specificity with introducing early-warning interpretability, which is, absents in prior hybrids and make our approach it as a scalable advancement for real-time ADAS applications.

5.4 Broader Applicability and Scalability

In addition to driver drowsiness monitoring, the proposed framework can also be applied to other IoT-based safety monitoring applications, such as industrial machinery operations, railway driver alerting systems, and construction

site fatigue monitoring. Since both HQCNN and fuzzy reasoning are lightweight and interpretable, this framework can be generalized to detect other human behavioural conditions (e.g., distraction, fatigue, or abnormal motion) when it is customized with task-specific features. The possible scope of the proposed system is not only limited to transportation safety but also it can be applicable to workplace safety and industrial IoT deployments.

6. Conclusion

This paper has presented a hybrid AI-enabled framework for real-time detection of driver drowsiness that combines multi-feature fuzzy logic-based classification and an HQNN module to improve accuracy and robustness. The fuzzy logic model incorporates multiple physiological indicators to enable adaptive drowsiness detection and the HQNN utilizes quantum-enhanced transformations to better learn features. The experimental results demonstrated that the effectiveness of the proposed approach in terms of higher classification accuracy and low computational complexity. The proposed fuzzy logic approach offers early drowsiness detection process with minimal false alarms. Our HQNN model we have explored its efficiency against overfitting issue where there less observation sample but with dynamic lightning effective, varying facial expressions such as size difference in eye and mouth for different individuals. The proposed dual-stage system addresses key challenges in real-time driver monitoring, and offers a scalable and computationally efficient solution for intelligent transportation systems (ITS). To further enhance real-world applicability, future work will explore multi-model and feature fusion approach and adaptive rule- with neural network for improved decision-making in advanced driver assistance systems (ADAS).

Funding: "This research received no external funding"

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] U.S. Department of Transportation, "Drowsy driving and traffic safety: Research findings and strategies for prevention," NHTSA, 2021. [Online]. Available: https://www.nhtsa.gov. [Accessed: Feb. 17, 2025].

[2] M. A. Rahman, S. Das, and X. Sun, "Understanding the drowsy driving crash patterns from correspondence regression analysis," *J. Saf. Res.*, vol. 84, pp. 167–181, 2023, doi: 10.1016/j.jsr.2022.10.017.

[3] Y. Albadawi, M. Takruri, and M. Awad, "A review of recent developments in driver drowsiness detection systems," *Sensors (Basel)*, vol. 22, no. 5, p. 2069, 2022, doi: 10.3390/s22052069.

[4] A. Quddus, A. S. Zandi, L. Prest, and F. J. E. Comeau, "Using long short term memory and convolutional neural networks for driver drowsiness detection," *Accid. Anal. Prev.*, vol. 156, p. 106107, 2021, doi: 10.1016/j.aap.2021.106107.

[5] G. S. Shankara Chari and J. A. Prashant, "Scaling effectivity in manifold methodologies to detect driver's fatigueness and drowsiness state," *IAES Int. J. Artif. Intell. (IJ-AI)*, vol. 13, no. 2, pp. 1227–1240, 2024, doi: 10.11591/ijai.v13.i2.pp1227-1240.

[6] M. Kamboj, K. Kadian, V. Dwivedi, A. Wary, and S. Ojha, "Advanced detection techniques for driver drowsiness: A comprehensive review of machine learning, deep learning, and physiological approaches," *Multimed. Tools Appl.*, 2024, doi: 10.1007/s11042-024-19738-z.

[7] K. Ahmad, P. P. Em, and N. A. A. Aziz, "Machine learning approaches for detecting driver drowsiness: A critical review," *Int. J. Membr. Sci. Technol.*, vol. 10, no. 1, pp. 329–346, 2023.

[8] T. S. Delwar et al., "AI- and deep learning-powered driver drowsiness detection method using facial analysis," *Appl. Sci. (Basel)*, vol. 15, no. 3, p. 1102, 2025, doi: 10.3390/app15031102.

[9] W. AlKishri, A. Abualkishik, and M. Al-Bahri, "Enhanced image processing and fuzzy logic approach for optimizing driver drowsiness detection," *Appl. Comput. Intell. Soft Comput.*, vol. 2022, pp. 1–14, 2022, doi: 10.1155/2022/9551203.

[10] B.-F. Wu, Y.-H. Chen, C.-H. Yeh, and Y.-F. Li, "Reasoning-based framework for driving safety monitoring using driving event recognition," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 3, pp. 1231–1241, 2013, doi: 10.1109/TITS.2013.2257759.

[11] J. Bai et al., "Two-stream spatial-temporal graph convolutional networks for driver drowsiness detection," *IEEE Trans. Cybern.*, vol. 52, no. 12, pp. 13821–13833, 2022, doi: 10.1109/TCYB.2021.3110813.

[12] Y. Peng et al., “A multi-source fusion approach for driver fatigue detection using physiological signals and facial image,” *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 11, pp. 1–11, 2024, doi: 10.1109/TITS.2024.3420409.

[13] G. Du, S. Long, C. Li, Z. Wang, and P. X. Liu, “A product fuzzy convolutional network for detecting driving fatigue,” *IEEE Trans. Cybern.*, vol. 53, no. 7, pp. 4175–4188, 2023, doi: 10.1109/TCYB.2021.3123842.

[14] J. Yousif, “Fuzzy logic approach for detecting drivers’ drowsiness based on image processing and video streaming,” *TechRxiv*, 2021, doi: 10.36227/techrxiv.16777927.v1.

[15] A. Snoun, T. Bouchrika, I. Teyeb, and O. Jemai, “A multimodal hypovigilance detection system based on fuzzy logic and transfer learning,” *J. Signal Process. Syst.*, vol. 94, no. 12, pp. 1411–1427, 2022, doi: 10.1007/s11265-022-01813-z.

[16] M. Dua, Shakshi, R. Singla, S. Raj, and A. Jangra, “Deep CNN models-based ensemble approach to driver drowsiness detection,” *Neural Comput. Appl.*, vol. 33, no. 8, pp. 3155–3168, 2021, doi: 10.1007/s00521-020-05209-7.

[17] H. V. Chand and J. Karthikeyan, “CNN based driver drowsiness detection system using emotion analysis,” *Intell. Autom. Soft Comput.*, vol. 31, no. 2, pp. 717–728, 2022, doi: 10.32604/iasc.2022.020008.

[18] E. Magán, M. P. Sesmero, J. M. Alonso-Weber, and A. Sanchis, “Driver drowsiness detection by applying deep learning techniques to sequences of images,” *Appl. Sci. (Basel)*, vol. 12, no. 3, p. 1145, 2022, doi: 10.3390/app12031145.

[19] I. Jahan et al., “4D: A real-time driver drowsiness detector using deep learning,” *Electronics (Basel)*, vol. 12, no. 1, p. 235, 2023, doi: 10.3390/electronics12010235.

[20] V. Owen and N. Surantha, “Computer vision-based drowsiness detection using handcrafted feature extraction for edge computing devices,” *Appl. Sci. (Basel)*, vol. 15, no. 2, p. 638, 2025, doi: 10.3390/app15020638.

[21] S. Essahraui et al., “Real-time driver drowsiness detection using facial analysis and machine learning techniques,” *Sensors (Basel)*, vol. 25, no. 3, 2025, doi: 10.3390/s25030812.

[22] R. Fusek, “Pupil localization using geodesic distance,” in *Lecture Notes in Computer Science*, vol. 11241, pp. 433–444, 2018.

[23] W. Kongcharoen, S. Nuchitprasitchai, Y. Nilsiam, and J. M. Pearce, “Real-time eye state detection system for driver drowsiness using convolutional neural network,” in *Proc. 2020 17th Int. Conf. Electr. Eng./Electron., Comput., Telecommun. Inf. Technol. (ECTI-CON)*, Phuket, Thailand, Jun. 24–27, 2020, pp. 551–554.

[24] Y. Suresh, R. Khandelwal, M. Nikitha, M. Fayaz, and V. Soudhri, “Driver drowsiness detection using deep learning,” in *Proc. 2021 2nd Int. Conf. Smart Electron. Commun. (ICOSEC)*, Trichy, India, Oct. 20–22, 2021, pp. 1526–1531.

[25] M. E. Walizad, M. Hurroo, and D. Sethia, “Driver drowsiness detection system using convolutional neural network,” in *Proc. 2022 6th Int. Conf. Trends Electron. Informat. (ICOEI)*, Tirunelveli, India, Apr. 28–30, 2022, pp. 1073–1080.

[26] M. Tibrewal, A. Srivastava, and R. Kayalvizhi, “A deep learning approach to detect driver drowsiness,” *Int. J. Eng. Res. Technol.*, vol. 10, pp. 183–189, 2021.

[27] I. Jahan, K. M. A. Uddin, S. A. Murad, M. S. U. Miah, T. Z. Khan, M. Masud, S. Aljahdali, and A. K. Bairagi, “4D: A real-time driver drowsiness detector using deep learning,” *Electronics*, vol. 12, p. 235, 2023.

[28] F. Makhmudov, D. Turimov, M. Xamidov, F. Nazarov, and Y.-I. Cho, “Real-time fatigue detection algorithms using machine learning for yawning and eye state,” *Sensors*, vol. 24, p. 7810, 2024.

[29] D. Salem and M. Waleed, “Drowsiness detection in real-time via convolutional neural networks and transfer learning,” *J. Eng. Appl. Sci.*, vol. 71, p. 122, 2024.