

Evolutionary-Optimized Multimodal Driver Risk Assessment Using DE-Enhanced Cross-Attention Fusion of EEG and Telematics

Ch Raga Madhuri*

Department of Computer Science and Engineering, Siddhartha Academy of Higher Education, Deemed to be University, Vijayawada - 520007, Andhra Pradesh, India

E-mail: madhuri.chandra1209@gmail.com

ORCID iD: <https://orcid.org/0000-0002-3687-0783>

*Corresponding author

Dedeepya Manikonda

Department of Computer Science and Engineering, Student of Engineering, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada -520007, Andhra Pradesh, India

E-mail: dedeepyamanikonda@gmail.com

ORCID iD: <https://orcid.org/0009-0006-8562-0509>

Madhurya Chinta

Department of Computer Science and Engineering, Student of Engineering, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada -520007, Andhra Pradesh, India

E-mail: madhurya2303@gmail.com

ORCID iD: <https://orcid.org/0009-0004-4955-3201>

Received: January 20, 2026; Revised: February 16, 2026; Accepted: March 25, 2026; Published: August 8, 2026

Abstract: Road safety depends on both a driver's emotional cognitive state and physical driving behavior, yet most existing systems rely on a single modality, limiting real-world reliability. This paper presents HECANet (Hybrid Evolutionary Cross-Attention Network), a multimodal framework that integrates EEG-based emotional cues and vehicle telematics data for robust driver risk assessment. EEG signals are modeled using a PSO-optimized CNN-LSTM to capture spatiotemporal emotional patterns, while telematics data are analyzed using a GA-optimized XGBoost model to identify safe, distracted, and aggressive driving behaviors. A Differential Evolution-optimized cross-attention fusion layer effectively aligns emotional and behavioral features, enabling interpretable emotion-behavior interaction modeling. The fused representation produces a driver safety score and risk probability, with K-Means clustering used to categorize drivers into Safe, Caution, and Risky groups. Experimental results achieve 94.7% accuracy and a 0.94 macro F1-score, demonstrating that joint emotion-behavior modeling significantly enhances driver risk prediction for intelligent transportation and fleet safety applications.

Index Terms: EEG, CNN-LSTM, Telematics, XGBoost, PSO, Genetic Algorithm, Differential Evolution, Cross-Attention Fusion, Driver Safety, Multimodal Learning, Risk Assessment.

1. Introduction

Road traffic accidents continue to pose a significant global challenge, accounting for approximately 1.3 million deaths annually and ranking among the leading causes of mortality across all age groups. Despite advances in vehicle automation and driver assistance technologies, human factors such as emotional instability, cognitive overload, fatigue, and inattentive behavior remain the dominant contributors to traffic collisions. These internal states often manifest subtly and unpredictably, making them difficult to capture using conventional monitoring systems that rely solely on external vehicle data or visual cues. Addressing this gap requires an integrated understanding of how drivers feel and how they behave, enabling safety systems to detect risk before it becomes observable at the behavioral level, particularly in systems that incorporate driver state awareness into intelligent mobility frameworks [1].

Electroencephalography (EEG) has emerged as a powerful modality for capturing real-time neural activity underlying emotional and cognitive conditions. Emotional drift, such as transitions toward stress, frustration, or drowsiness, directly affects reaction time, decision-making, and situational awareness [2]. At the same time, telematics sensors embedded in modern vehicles generate high-resolution behavioral indicators, including acceleration patterns, gyroscopic motion, braking intensity, and steering behaviour. These signals reflect the outward manifestation of driving risk. While EEG provides insight into internal cognitive dynamics, telematics captures the resulting vehicle actions; individually, each modality offers only a partial view. Together, they can form a comprehensive and context-aware representation of driver safety [3].

Recent multimodal safety research highlights the benefits of integrating physiological and behavioral data, yet several limitations persist. Many approaches rely on simple feature concatenation or shallow models that fail to capture complex temporal and non-linear interactions. While telematics-based methods effectively model driving behavior and provide scalable surrogate safety measures [4], including spatial risk analysis in real-world scenarios [5], they lack insight into drivers' underlying cognitive and emotional states. Additionally, manual hyperparameter tuning and the scarcity of synchronized multimodal datasets limit generalization and real-world deployment of robust driver safety systems.

To address these challenges, the present study introduces HECANet, a Hybrid Evolutionary Cross-Attention Network designed to fuse EEG-based emotional indicators with telematics-based driving behavior for holistic driver risk profiling. HECANet incorporates three levels of innovation. First, EEG signals are processed through a CNN–LSTM architecture that captures both spatial and temporal dynamics of emotional fluctuations, with hyperparameters optimized through Particle Swarm Optimization (PSO) to enhance robustness. Second, driving behaviour is modelled using the XGBoost algorithm, whose parameters are refined using a Genetic Algorithm (GA) to balance model complexity and predictive strength. Third, a Cross-Attention Fusion Mechanism, optimized using Differential Evolution (DE), aligns and integrates emotional and behavioral representations, enabling the system to learn nuanced interdependencies that traditional fusion methods overlook. The unified feature representation produced by HECANet enables the estimation of a driver-specific Safety Score and Risk Probability, which are further stratified using unsupervised clustering to categorize drivers into Safe, Caution, and Risky groups. With a test accuracy of 94.7%, strong macro-level F1 performance, and stable convergence behavior, HECANet demonstrates the effectiveness of optimized multimodal learning for real-time driver safety assessment.

2. Literature Review

Driving safety monitoring has shifted from unimodal sensing to integrated physiological–vehicular frameworks for reliable detection of unsafe driver states. Ju and Li reviewed EEG-based driver state and behavior detection for intelligent vehicles, covering acquisition, preprocessing, feature learning, classification, and deployment challenges such as noise and inter-subject variability [6], while multimodal fusion has been shown to enhance robustness, including EEG–ECG–vision-based fatigue and behavior monitoring [7], and standardized multimodal in-vehicle BCI datasets for cognitive load assessment [8].

EEG-based monitoring of fatigue, workload, and emotion remains central to driver state analysis, where improved feature design and fusion consistently outperform unimodal methods. Lan et al. fused EEG band energy ratios with vehicle motion features using Pearson selection and SVM, achieving 92.37% accuracy [9], while Zhang et al. combined time-, frequency-, and nonlinear-domain EEG features with Logistic Regression–ANOVA–RFE and PCA fusion for robust fatigue recognition using Gaussian SVM [10]. Islam et al. introduced mutual-information-based EEG–vehicular features for workload estimation (MAE = 0.09, 94% accuracy) [11], and recent advances in EEG emotion recognition include enhanced hybrid feature sets [12], contrastive brain-area representation learning [13], simulated accident-related emotion analysis, and efficient all-MLP architectures such as EEG-MLP.

Telematics and connected vehicle data enable scalable driver risk estimation by modeling driving dynamics and behavioral patterns. Mussah and Adu-Gyamfi predicted crash outcomes from connected vehicle trajectories using XGBoost with 0.95 accuracy [14], while Abdelrahman et al. developed supervised driver behavior profiling on SHRP2 data for risk probability estimation [15], complemented by deep learning–XGBoost approaches for real-time assessment [16]. Risk inference from vehicle telemetry has been further supported through interpretable telematics-based insurance frameworks [17], multi-dimensional safety evaluation for electric vehicles using GRU–TCN with transfer learning (MAPE = 2.68%) [18], large-scale fleet safety indices and spatial risk analysis, and systematic reviews highlighting generalization challenges in telematics monitoring [19]. Affective risk has also been linked to aggressive driving through physiological analysis, including emotional arousal prediction from EDA [20], while multimodal affective computing studies consistently demonstrate that fusing EEG with facial expressions, speech, and eye-tracking using attention-based mechanisms improves robustness [21 – 23], motivating multimodal driver monitoring systems.

Foundational studies have established the importance of both physiological and behavioral signals in driver monitoring systems. EEG-based approaches have been widely explored for emotion and cognitive state analysis using benchmark datasets such as DEAP [24], while early works have also demonstrated the effectiveness of EEG signals for detecting driver drowsiness and fatigue [25]. In parallel, driver behavior monitoring systems have utilized vehicle

dynamics and telematics data to assess driver attention and safety [26]. Furthermore, multimodal learning frameworks have highlighted the advantages of integrating heterogeneous data sources to improve system performance and robustness. Despite significant progress in driver monitoring systems, existing approaches exhibit several limitations. EEG-based methods primarily focus on internal cognitive states but lack integration with observable driving behavior, limiting their real-world applicability. In contrast, telematics-based approaches effectively capture external driving patterns but fail to account for underlying emotional and cognitive factors [27].

Although multimodal methods have been explored, most rely on simple feature concatenation or shallow fusion techniques, which do not effectively model complex interdependencies between modalities. Furthermore, many existing works lack systematic optimization strategies, resulting in suboptimal performance and limited generalization. Another critical gap is the insufficient modeling of cross-modal interactions, where the influence of cognitive state on driving behavior is not explicitly learned. Additionally, the lack of synchronized multimodal datasets remains a challenge for practical deployment. To address these limitations, this work proposes an evolutionary-optimized cross-attention fusion framework that enables deep interaction modeling between EEG and telematics data while improving performance and generalization.

3. Methodology

Fig 1 illustrates the proposed multimodal architecture for driver risk assessment, combining EEG-based cognitive state analysis with telematics-based driving behavior modeling. EEG signals collected from multiple scalp electrodes are decomposed into standard frequency bands and preprocessed to remove artifacts and normalize signal variations, enabling effective representation of driver cognitive and emotional states. A CNN–LSTM network is then used to learn spatial and temporal neural patterns, with Particle Swarm Optimization applied to refine model hyperparameters and generate a compact EEG embedding. In parallel, vehicle telematics data capturing driving dynamics are modeled using an XGBoost classifier optimized through a Genetic Algorithm, producing a behavior-aware telematics embedding consistent with established fleet safety analysis approaches [28, 29]. The two modality-specific embeddings are integrated through a cross-attention fusion mechanism to capture interactions between physiological state and driving behavior. The fused features are finally used to compute a driver risk score and cluster driving patterns, supporting interpretable and efficient driver safety assessment[30].

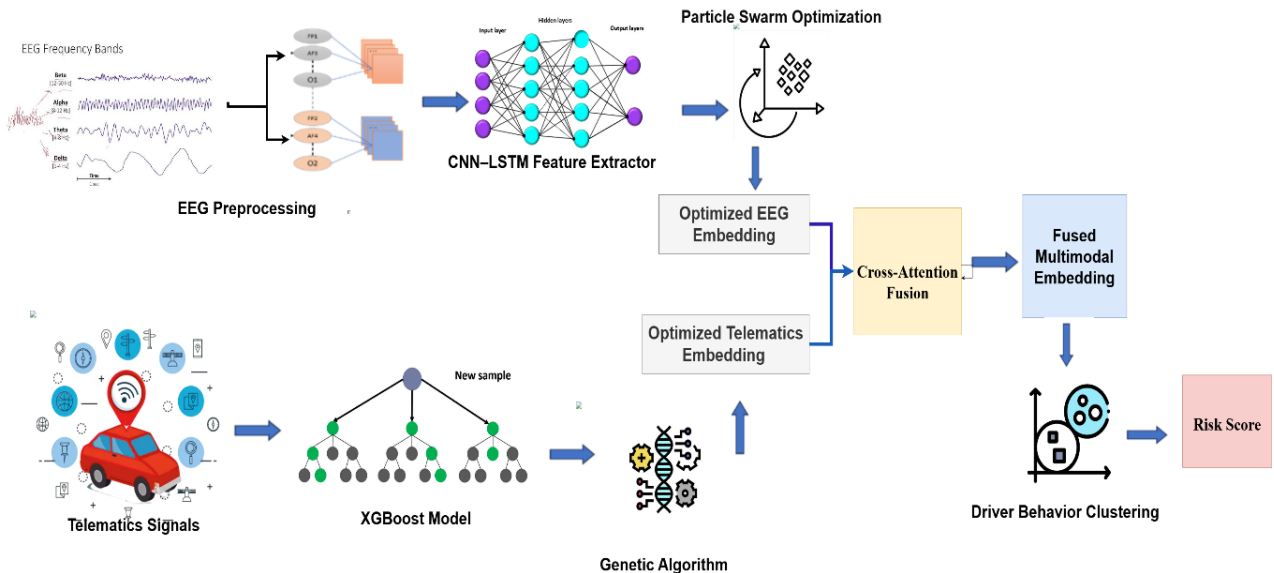


Fig. 1. Multimodal EEG–Telematics Architecture for Driver Risk Assessment.

The proposed multimodal driver-safety assessment framework is composed of four synergistic modules: (1) EEG feature extraction using a PSO-optimized CNN–LSTM network, (2) telematics modeling using a GA-tuned XGBoost classifier, (3) Differential Evolution (DE)-optimized cross-attention fusion for effective integration of neurophysiological and behavioral information, and (4) a clustering-based safety scoring module that transforms the fused representation into interpretable driver-state categories. The complete architecture of the system is presented in Fig. 2.

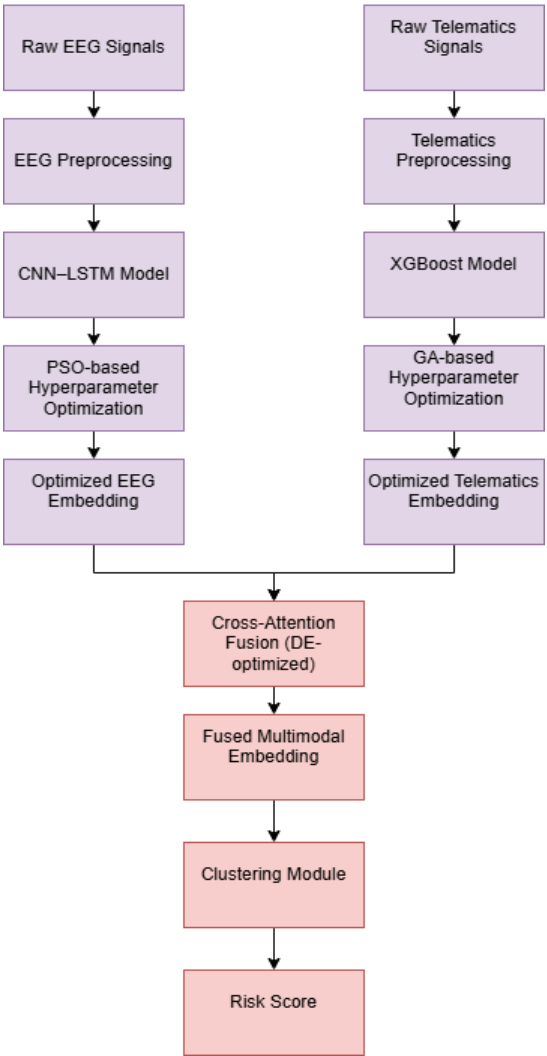


Fig. 2. Proposed multimodal driver-safety assessment pipeline based on optimized EEG and telematics modeling and DE-enhanced cross-attention fusion.

3.1. Datasets

This study utilizes publicly available datasets to support multimodal driver risk assessment using EEG and vehicle telematics data. EEG signals are obtained from the DEAP dataset, which provides 32-channel recordings from multiple subjects along with emotion-related annotations such as arousal and valence, making it suitable for modeling drivers’ internal cognitive and emotional states [24]. Vehicle driving behavior is represented using the Levin Vehicle Telematics dataset, which contains real-world measurements including speed, acceleration, gyroscope signals, and engine parameters collected from multiple vehicles, commonly used for driving behavior and safety analysis. Both datasets are independently preprocessed, segmented into fixed-length time windows, and normalized prior to feature extraction. As the EEG and telematics datasets are collected from different subjects and environments, no direct temporal or subject-level alignment is assumed. Instead, the proposed framework performs representation-level fusion, where modality-specific embeddings are learned independently and integrated using the DE-optimized cross-attention mechanism to capture complementary relationships between cognitive state and driving behavior. This design enables generalized multimodal learning and serves as a preliminary step toward real-world systems with synchronized data.

3.2. System Overview

Since the EEG and telematics data originate from independent datasets, fusion is performed at the feature-representation level rather than through direct temporal alignment. EEG is analyzed using a deep neural architecture capable of capturing both spatial–frequency and temporal patterns, while telematics data are modeled using a powerful tree-based ensemble method. Evolutionary optimization methods (PSO, GA, DE) ensure that each subsystem achieves its best possible performance. The optimized embeddings are fused using a cross-attention mechanism, and the final fused representation is mapped to driver safety levels through unsupervised clustering. This produces both highly accurate predictions and interpretable risk scores.

3.3. Data Acquisition and Preprocessing

3.3.1. EEG Data Preprocessing

EEG signals were preprocessed using a bandpass filter in the range of 0.5–45 Hz implemented using a Butterworth filter to remove low-frequency drift and high-frequency noise. Artifact removal was performed using threshold-based noise suppression techniques to eliminate eye-blink and motion-related interference. The filtered signals were normalized channel-wise using z-score normalization to ensure consistent scaling across all channels. The EEG data were then segmented into fixed-length windows of 2 seconds with 50% overlap to capture temporal dynamics effectively.

$$X^{EEG} \in \mathbb{R}^{T \times C} \quad (1)$$

Each segment forms a matrix of size $T \times C$, where T denotes the number of time samples per window and C represents the number of EEG channels (1). This preprocessing pipeline ensures noise reduction, temporal consistency, and suitability for deep learning-based feature extraction.

3.3.2. Telematics Data Preprocessing

Telematics data collected from vehicle sensors such as speed, acceleration, gyroscope, and steering angle were first resampled to ensure a uniform sampling rate. The data were then segmented into fixed-length time windows for consistent analysis. From each segment, statistical and dynamic features including mean, standard deviation, variance, and entropy were extracted to capture driving behavior patterns. These features provide a compact representation of both steady-state and dynamic driving characteristics. The extracted feature vectors were normalized using z-score standardization to ensure uniform scaling across all variables, resulting in a structured input suitable for the XGBoost model.

3.4. Optimized EEG Modeling Using CNN–LSTM With PSO

3.4.1. CNN–LSTM Architecture

The EEG stream was modeled using a hybrid convolutional and recurrent neural network architecture. Convolutional layers were used to extract localized temporal–frequency features across channels, capturing brainwave patterns related to driver alertness and fatigue. The extracted CNN features were then passed into an LSTM layer to learn temporal transitions. The final hidden state of the LSTM served as the EEG embedding, represented as:

$$E_{EEG} = f_{CNN-LSTM}(X^{EEG}) \quad (2)$$

This equation represents the EEG embedding extracted from the input EEG window (Eq. (2)). The optimized CNN–LSTM model $f_{CNN-LSTM}$ converts raw EEG into a meaningful feature vector. The CNN component consists of two convolutional layers with ReLU activation, followed by max-pooling layers. The LSTM layer contains 64 hidden units to capture temporal dependencies, followed by a fully connected layer to generate the EEG embedding. Dropout is applied to reduce overfitting. For the telematics model, XGBoost is configured with a maximum depth of 6, learning rate of 0.1, and 100 estimators, with parameters optimized using GA. The model was trained using the Adam optimizer with a learning rate of 0.001 and a batch size of 32 for 50 epochs. Early stopping based on validation loss was employed to prevent overfitting and ensure optimal model convergence.

3.4.2. PSO Hyperparameter Optimization

Manually tuning deep neural networks is inefficient and often produces suboptimal results. Therefore, Particle Swarm Optimization (PSO) was used to identify the best set of hyperparameters for the CNN–LSTM model. Each PSO particle encoded a combination of hyperparameters such as the number of convolutional filters, kernel sizes, number of LSTM units, dropout rates, and learning rates. The fitness of each particle was evaluated using the validation F1-score:

$$Fitness = 1 - F1_{validation} \quad (3)$$

PSO minimizes this fitness value, so higher validation F1-scores produce lower fitness (Eq. (3)). Through iterative updates guided by personal and global best solutions, PSO converged to a configuration that significantly improved EEG-based driver-state classification.

3.5. Optimized Telematics Modeling Using XGBoost With GA

3.5.1. XGBoost Model

Telematics features were modeled using XGBoost, a high-performance gradient-boosted decision tree algorithm that excels at structured data modeling. XGBoost constructs multiple boosted decision trees while minimizing a regularized loss function to balance accuracy and model complexity. The intermediate output from the penultimate boosting layer was used as the telematics embedding:

$$E_{TEL} = f_{XGB}(X^{TEL}) \quad (4)$$

The boosting layer is used to generate the telematics embedding, as shown in (Eq. (4)). An optimized XGBoost model processes the input telematics features through an ensemble of decision trees, capturing complex non-linear relationships in the data. For each input sample, the leaf node indices from all trees are extracted and combined to form a fixed-length embedding vector. This representation encodes important feature interactions and serves as the telematics embedding for further fusion.

3.5.2. GA-Based Hyperparameter Optimization

A Genetic Algorithm (GA) was employed to identify the optimal XGBoost hyperparameters, including maximum tree depth, number of estimators, learning rate, subsample ratio, and column sampling rate. GA evolves populations of candidate solutions through crossover, mutation, and elitist selection. The fitness of each candidate was determined using validation accuracy, and the best-performing chromosome produced a telematics model with enhanced predictive capability.

3.6. Cross-Attention Fusion Optimized by Differential Evolution

3.6.1. Cross-Attention Mechanism

After individual optimization, the EEG and telematics embeddings were integrated using a cross-attention mechanism. Cross-attention allows each modality to highlight the most relevant patterns in the other modality. Given the two embeddings, the fused representation was computed as:

$$E_{FUSED} = \text{Attn}(E_{EEG}, E_{TEL}) + \text{Attn}(E_{TEL}, E_{EEG}) \quad (5)$$

where $\text{Attn}(E_{EEG}, E_{TEL})$ represents EEG attending to telematics features, and $\text{Attn}(E_{TEL}, E_{EEG})$ represents telematics attending to EEG, enabling complementary feature integration (5).

To provide a formal definition, cross-attention is computed using query (Q), key (K), and value (V) projections. Given embeddings E_{EEG} and E_{TEL} , the fused representation is defined as:

$$E_{FUSED} = \frac{\text{softmax}(Q_{EEG} K_{TEL}^T)}{\sqrt{d}} \cdot V_{TEL} + \frac{\text{softmax}(Q_{TEL} K_{EEG}^T)}{\sqrt{d}} \cdot V_{EEG} \quad (6)$$

where Q, K, and V represent learned linear projections of the embeddings, and d denotes the embedding dimension. This formulation enables each modality to attend to the most informative features of the other, resulting in a richer and more discriminative fused representation (Eq. (6)).

3.6.2. DE Optimization of Fusion Parameters

Differential Evolution (DE) was used to tune the fusion-related parameters such as attention dimension, number of attention heads, fusion weights, and regularization coefficients. DE explores a population of parameter vectors using mutation and crossover strategies, gradually converging toward the configuration that maximizes model performance. The DE-optimized fusion module produced the most discriminative multimodal representation. The use of Differential Evolution is motivated by its effectiveness in optimizing continuous parameter spaces and avoiding local optima. Compared to manual tuning or grid search, DE provides a more efficient and robust optimization strategy for the cross-attention fusion parameters, contributing to improved overall model performance.

3.7. Clustering-Based Safety Scoring

To generate interpretable safety levels, the fused multimodal embedding was further processed using a clustering algorithm such as K-Means or Gaussian Mixture Models. The fused features:

$$E_{Fused} \in \mathbb{R}^D \quad (7)$$

where it denotes the fused feature vector in a D-dimensional real-valued space, representing the combined multimodal embedding obtained from EEG and telematics data (7). The fused feature vector is partitioned into three clusters corresponding to driver risk levels are Safe, Mild Risk, and High Risk Cluster. quality is assessed using silhouette and Davies–Bouldin indices, and each cluster is subsequently mapped to a final safety score. This clustering-based formulation enables not only risk prediction but also an interpretable and continuous understanding of driver state progression.

4. Results and Discussion

This section presents a comprehensive evaluation of the proposed multimodal driver-safety assessment framework. We analyse model behaviour across multiple dimensions, including learning stability, convergence characteristics, class-wise discrimination capability, and overall predictive accuracy. The evaluation incorporates training and validation accuracy curves, a detailed confusion matrix, and class-specific performance metrics to rigorously assess the effectiveness and reliability of the optimized multimodal architecture. To ensure reliable evaluation, the datasets are divided into training 70%, validation 15%, and test 15% sets. The training set is used for model learning, while the validation set is used during PSO and GA optimization for hyperparameter tuning. The test set is strictly held out for final performance evaluation.

To prevent data leakage, preprocessing steps such as normalization and feature extraction are performed independently within each split. The data partitioning is completed prior to model training to ensure no information from the test set influences the learning process. In addition to the fixed train–validation–test split, cross-validation was performed to further assess model robustness. A 5-fold cross-validation strategy was applied, where the dataset was partitioned into five subsets, and the model was trained and evaluated across multiple folds. The results showed consistent performance across folds, confirming the stability and generalization capability of the proposed framework.

4.1. Model Performance

4.1.1. Baseline Model Performance (Without Optimization)

This subsection reports the classification performance of baseline models trained using default hyperparameters, without applying any optimization techniques. These results serve as a reference for evaluating the effectiveness of optimization and multimodal fusion.

Table 1. Baseline accuracy of evaluated models.

Model configuration	Accuracy (%)
EEG-based model	82.0
Telematics-based model	85.0
EEG + Telematics (Fusion)	88.0

The EEG-based model achieves an accuracy of 82%, as shown in Table 1, which reflects the inherent variability and noise in physiological signals. The telematics-based model performs better, achieving 85%, as vehicle dynamics directly capture observable driving behavior. Combining EEG and telematics features improves accuracy to 88%, demonstrating the benefit of multimodal fusion even without optimization. Traditional machine learning models such as Support Vector Machine (SVM) and Random Forest were also evaluated as baselines. However, they achieved lower performance compared to the proposed model, highlighting the effectiveness of deep learning and multimodal fusion in capturing complex nonlinear relationships.

4.1.2. Optimized Model Performance

In this subsection, modality-specific optimization techniques are applied to enhance model performance. PSO is used for the EEG-based model, GA for the telematics-based model, and DE for the fused multimodal model.

Table 2. Accuracy of models after optimization.

Model configuration	Optimization method	Accuracy (%)
EEG-based model	PSO	86.0
Telematics-based model	GA	89.0
EEG + Telematics (Proposed)	DE	94.7

After optimization in Table 2, all models show improved performance. The optimized EEG and telematics models achieve accuracies of 86% and 89%, respectively. The proposed DE-optimized fusion model attains the highest accuracy of 94%, confirming that evolutionary optimization combined with multimodal fusion significantly enhances driver risk assessment performance.

Table 3. Comparison of Optimization Strategies.

Optimization Method	Accuracy (%)
No Optimization	88.0
Grid Search	91.2
Proposed (PSO + GA + DE)	94.7

The effectiveness of the proposed evolutionary optimization strategy was further evaluated by comparing it with conventional tuning approaches such as grid search. As shown in Table 3, grid search improves performance over baseline models; however, the proposed combination of PSO, GA, and DE achieves superior accuracy. This improvement is attributed to the ability of evolutionary algorithms to efficiently explore complex and high-dimensional hyperparameter spaces, avoiding local optima and producing more robust configurations.

4.2. Training and Validation Performance

Fig. 3 shows the training and validation accuracy curves for the optimized HECANet model over 50 epochs. The model exhibits rapid learning in the initial phase, with training accuracy rising from approximately 72% to above 88% by Epoch 5. Validation accuracy follows a similar trend, increasing from about 70% to 87% over the same range, indicating strong early generalization.



Fig. 3. Training and validation accuracy curves for the optimized HECANet model over epochs.

4.3. Confusion Matrix Analysis

Fig. 4 illustrates the confusion matrix for the optimized multimodal driver-state classifier across three categories: Safe, Mild Risk, and High Risk. The model demonstrates strong class-wise performance, correctly identifying the majority of Safe and High Risk instances, with only limited confusion between these two classes. Notably, the Mild Risk category is classified without any misclassification, indicating a clear separation of intermediate-risk states from both low- and high-risk driving conditions. Overall, the confusion patterns confirm the model’s reliability in distinguishing distinct driver risk levels while maintaining stable and balanced classification behavior.

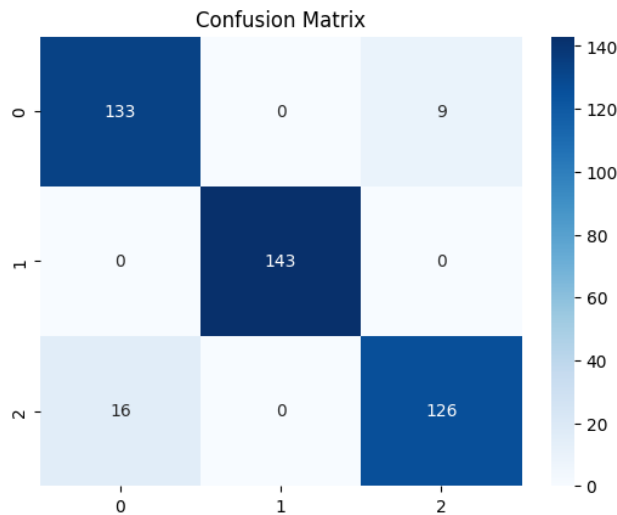


Fig. 4. Confusion matrix showing the classification performance of the optimized multimodal driver-state model across three classes: Safe (0), Mild Risk (1), and High Risk (2).

4.4. Class-Wise Performance Metrics

Beyond raw counts, Table 4 reports precision, recall, and F1-score for each class, as well as the macro average across all classes.

Table 4. Class-Wise Precision, Recall, and F1-Score.

Class	Precision (%)	Recall (%)	F1-Score (%)
Safe (0)	89.5	93.6	91.5
Mild Risk (1)	100.0	100.0	100.0
High Risk (2)	93.8	88.7	91.1
Macro Average	94.4	94.1	94.2

The proposed model demonstrates strong and well-balanced classification performance across all driver risk categories. The Safe and High Risk classes achieve F1-scores above 91%, indicating reliable discrimination between low- and high-risk driving behaviors, while the macro-average F1-score of 94.2% confirms consistent predictive capability without class bias. Notably, the Mild Risk category achieves perfect precision, recall, and F1-score, suggesting clear separability of intermediate-risk patterns in the learned multimodal feature space. However, this result should be interpreted cautiously, as it may also reflect dataset-specific characteristics due to the use of independently collected EEG and telematics datasets. Overall, the confusion matrix shows strong diagonal dominance with minimal misclassification, and statistical evaluation across multiple runs indicates stable performance and good generalization.

4.5. Statistical Validation and Robustness Analysis

To evaluate the robustness of the proposed model, multiple independent runs were conducted using different random initializations. The performance metrics are reported as mean and standard deviation. The results demonstrate that the model achieves stable and consistent performance across runs, with low variance indicating reliable generalization.

Table 5. Statistical Performance Across Multiple Runs.

Metric	Mean (%)	Standard Deviation (%)
Accuracy	94.7	± 0.8
Precision	94.4	± 0.7
Recall	94.1	± 0.9
F1-Score	94.2	± 0.7

As shown in Table 5, the low standard deviation confirms that the model performance is stable and not sensitive to random initialization.

4.6. Comparison With Existing Models

The comparison in Table 6 highlights the effectiveness of the proposed HECANet framework relative to existing driver monitoring approaches. Prior studies largely focus on either EEG–vehicle motion fusion using conventional classifiers, mutual-information-based EEG–vehicular feature fusion, or telematics-only crash prediction using ensemble and XGBoost models, achieving accuracies between 92.37% and 95.0%. Emotion-centric approaches based on physiological signals such as EDA show comparatively lower performance due to limited behavioral context. In contrast, HECANet integrates EEG and telematics through evolutionary optimization and cross-attention fusion, achieving 94.7% accuracy while jointly modeling drivers’ internal cognitive–emotional states and external driving behavior. This balanced multimodal design enables competitive performance while providing richer interpretability and broader applicability for real-world driver risk assessment.

Table 6. Comparison With Other Models.

Related Work	Task	Dataset	Accuracy (%)
Lan et al. [1]	Driving fatigue detection using EEG + vehicle motion fusion (SVM)	Experimental driving data (EEG + steering angle, acceleration, speed)	92.37%
Islam et al. [2]	Mental workload classification using mutual-information EEG + vehicular fusion features	Real driving experiment (EEG + vehicular signals)	94.0%
Mussah & Adu-Gyamfi [3]	Real-time crash outcome prediction using connected vehicle trajectory data (ensemble/XGBoost)	Connected vehicle high-resolution trajectory dataset	95.0%
Nahid et al. [10]	Emotional arousal prediction using neural networks from EDA features	20 participants (18–30), 84 driving sessions (EDA + telematics)	65.0%
Our Proposed Work	HECANet	EEG + Telematics	94.7%

4.7. Overall System Performance and Discussion

The overall results highlight the effectiveness of the proposed multimodal, evolutionarily optimized framework:

- The high overall accuracy ($\approx 95\%$) and macro F1-score ($\approx 94\%$) indicate that the combination of EEG and telematics data provides a rich, complementary view of driver state.
- The PSO-optimized CNN–LSTM efficiently models neurophysiological patterns, while the GA-optimized XGBoost captures complex driving behavior from telematics.
- The DE-optimized cross-attention fusion further enhances performance by allowing each modality to selectively attend to the most relevant features in the other, leading to the near-perfect recognition of Mild Risk.

The model demonstrates a small generalization gap between training ($\approx 96\%$), validation ($\approx 94\%$), and test accuracy (94.7%), indicating that overfitting is effectively controlled. This is further supported by the use of dropout in the CNN–LSTM model and validation-based optimization strategies.

5. Conclusion

This study presented an optimized multimodal driver safety assessment framework that integrates EEG signals and vehicle telematics data through evolutionary learning and attention-based fusion. The use of PSO-tuned CNN–LSTM models for EEG analysis, GA-optimized XGBoost for telematics modeling, and DE-enhanced cross-attention fusion enabled effective learning of complex neuro-behavioral patterns associated with different driver risk states. Experimental results demonstrate high accuracy, stable convergence, and reliable class-wise performance, particularly in identifying Mild and High-Risk driving conditions. The inclusion of clustering-based safety scoring further enhances interpretability by enabling transparent monitoring of driver state transitions. Beyond private vehicle applications, the proposed framework has strong potential for deployment in on-demand mobility services, large-scale ride-sharing platforms, and fleet-based transportation systems, where continuous driver monitoring is critical for improving safety and service reliability. Despite the strong performance achieved by the proposed framework, several limitations should be acknowledged. The reliance on independently collected EEG and telematics datasets may introduce dataset-specific bias and limit the generalizability of the model to real-world scenarios. In addition, multimodal inputs are inherently sensitive to noise, and the integration of multiple learning components increases computational complexity, which may affect real-time deployment.

Furthermore, the absence of extensive statistical significance testing may constrain the robustness evaluation of the proposed approach. Future work will focus on validating the framework using synchronized, real-world multimodal datasets, incorporating additional sensing modalities, and performing comprehensive statistical analysis. Efforts will also be directed toward developing lightweight and adaptive models to enable efficient real-time deployment in embedded vehicular systems.

All the Declarations and Statements

Author Contributions Statement

Ch Raga Madhuri – Conceptualization, Methodology, and Supervision: Proposed the research idea, designed the overall system architecture, and supervised the project execution.

Dedeepya Manikonda – Data Curation, Software Implementation, and Model Development: Performed data acquisition, preprocessing, implementation of models, and conducted model training and validation.

Madhurya Chinta – Formal Analysis, Visualization, and Writing: Conducted performance evaluation, prepared visualizations and analysis, and contributed to manuscript writing, review, and editing.

All authors have read and agreed to the published version of the manuscript.

Conflict of Interest Statement

The authors declare that there is no conflict of interest regarding the publication of this paper.

Funding Declaration

This research received no external funding.

Data Availability Statement

The datasets used in this study are publicly available. The EEG data were obtained from the DEAP dataset (<https://www.eecs.qmul.ac.uk/mmv/datasets/deap/>), and the vehicle telematics data were obtained from the Levin Vehicle Telematics dataset (<https://www.kaggle.com/datasets/yunlevin/levin-vehicle-telematics>). Processed data and implementation details are available from the corresponding author upon reasonable request.

Ethical Declarations

The authors should declare the following: Statements of ethical approval for studies involving human subjects and/or animals.

Acknowledgments

The authors would like to thank the reviewers and domain experts for their valuable feedback and suggestions, which helped improve the quality and reliability of this research work.

Declaration of Generative AI in Scholarly Writing

During the preparation of this manuscript, AI-assisted tools were used only for language improvement, grammar correction, and formatting purposes. All technical content, methodology, results, and conclusions were developed and verified by the authors. The authors take full responsibility for the content of this publication.

Abbreviations

None.

Appendix A/B/C..., with appendix title

None.

References

- [1] M. Jang and K. Oh, "Development of an integrated longitudinal control algorithm for autonomous mobility with EEG-based driver status classification and safety index," *Electronics*, vol. 13, no. 7, p. 1374, 2024, doi: 10.3390/electronics13071374.
- [2] Chen, J., Lin, X., Ma, W., Wang, Y., & Tang, W. (2024). EEG-based emotion recognition for road accidents in a simulated driving environment. *Biomedical signal processing and control*, 87, 105411, 2024, doi: 10.1016/j.bspc.2023.105411.
- [3] Liu, D., Yang, L., Ni, P., Wang, Q., Sun, H., Zhang, Q., & Tang, C. (2023, December). EEG-MLP: An all-MLP architecture for EEG emotion recognition. In *2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 2655-2662). IEEE, 2023, doi: 10.1109/BIBM58861.2023.10385434.
- [4] P. Alrassy, A. W. Smyth, and J. Jang, "Driver behavior indices from large-scale fleet telematics data as surrogate safety measures," *Accident Analysis & Prevention*, vol. 179, p. 106879, 2023, doi: 10.1016/j.aap.2022.106879.
- [5] Nikolaou, D., Ziakopoulos, A., Kontaxi, A., Theofilatos, A., & Yannis, G. (2025). Spatial analysis of telematics-based surrogate safety measures. *Journal of Safety Research*, 92, 98-108, 2025, doi: 10.1016/j.jsr.2024.09.012.
- [6] J. Ju and H. Li, "A survey of EEG-based driver state and behavior detection for intelligent vehicles," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 6, no. 3, pp. 420–434, 2024, doi: 10.1109/TBIOM.2024.3400866
- [7] L. Wang, F. Song, T. H. Zhou, J. Hao, and K. H. Ryu, "EEG and ECG-based multi-sensor fusion computing for real-time fatigue

- driving recognition based on feedback mechanism," *Sensors*, vol. 23, no. 20, p. 8386, 2023, doi: 10.3390/s23208386.
- [8] Angkan, P., Behinaein, B., Mahmud, Z., Bhatti, A., Rodenburg, D., Hungler, P., & Etemad, A. (2024). Multimodal brain–computer interface for in-vehicle driver cognitive load measurement: Dataset and baselines. *IEEE Transactions on Intelligent Transportation Systems*, 25(6), 5949–5964, 2024, doi: 10.1109/TITS.2023.3345846.
- [9] Z. Lan, J. Zhao, P. Liu, C. Zhang, N. Lyu, and L. Guo, "Driving fatigue detection based on fusion of EEG and vehicle motion information," *Biomedical Signal Processing and Control*, vol. 92, p. 106031, 2024, doi: 10.1016/j.bspc.2024.106031.
- [10] Y. Zhang, H. Guo, Y. Zhou, C. Xu, and Y. Liao, "Recognising drivers' mental fatigue based on EEG multi-dimensional feature selection and fusion," *Biomedical Signal Processing and Control*, vol. 79, p. 104237, 2023, doi: 10.1016/j.bspc.2022.104237.
- [11] M. R. Islam, S. Barua, M. U. Ahmed, S. Begum, P. Aricò G. Borghini, and G. Di Flumeri, "A novel mutual information based feature set for drivers' mental workload evaluation using machine learning," *Brain Sciences*, vol. 10, no. 8, p. 551, 2020, doi: 10.3390/brainsci10080551.
- [12] D. Kulkarni and V. V. Dixit, "Hybrid classification model for emotion detection using EEG with improved feature set," *Biomedical Signal Processing and Control*, vol. 100, p. 106893, 2025, doi: 10.1016/j.bspc.2024.106893.
- [13] S. Dai et al., "Contrastive learning of EEG representation of brain areas for emotion recognition," *IEEE Transactions on Instrumentation and Measurement*, 2025, doi: 10.1109/TIM.2025.3533618.
- [14] A. R. Mussah and Y. Adu-Gyamfi, "Machine learning framework for real-time assessment of traffic safety utilizing connected vehicle data," *Sustainability*, vol. 14, no. 22, p. 15348, 2022, doi: 10.3390/su142215348.
- [15] A. E. Abdelrahman, H. S. Hassanein, and N. Abu-Ali, "Robust data-driven framework for driver behavior profiling using supervised machine learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 4, pp. 3336–3350, 2020, doi: 10.1109/TITS.2020.3035700.
- [16] L. Shi, C. Qian, and F. Guo, "Real-time driving risk assessment using deep learning with XGBoost," *Accident Analysis & Prevention*, vol. 178, p. 106836, 2022, doi: 10.1016/j.aap.2022.106836.
- [17] H. J. Li, X. G. Luo, Z. L. Zhang, W. Jiang, and S. W. Huang, "Driving risk prevention in usage-based insurance services based on interpretable machine learning and telematics data," *Decision Support Systems*, vol. 172, p. 113985, 2023, doi: 10.1016/j.dss.2023.113985.
- [18] J. Hong, F. Liang, H. Zhang, Y. Chen, R. Li, K. Li, and J. Yang, "Data-driven multi-dimension driving safety evaluation for real-world electric vehicles," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 7, pp. 9721–9733, 2024, doi: 10.1109/TVT.2024.3379677.
- [19] J. Boylan, D. Meyer, and W. S. Chen, "A systematic review of the use of in-vehicle telematics in monitoring driving behaviours," *Accident Analysis & Prevention*, vol. 199, p. 107519, 2024, doi: 10.1016/j.aap.2024.107519.
- [20] M. S. H. Nahid, T. Z. Tila, and T. S. Seecharan, "Detecting emotional arousal and aggressive driving using neural networks: A pilot study involving young drivers in Duluth," *Sensors*, vol. 24, no. 22, p. 7109, 2024, doi: 10.3390/s24227109.
- [21] Pan, J., Fang, W., Zhang, Z., Chen, B., Zhang, Z., & Wang, S. (2023). Multimodal emotion recognition based on facial expressions, speech, and EEG. *IEEE Open Journal of Engineering in Medicine and Biology*, 5, 396–403, 2023, doi: 10.1109/OJEMB.2023.3240280.
- [22] Bhatlawande, S., Shilaskar, S., Pramanik, S., & Sole, S. (2024). Multimodal emotion recognition based on the fusion of vision, EEG, ECG, and EMG signals. *International journal of electrical and computer engineering systems*, 15(1), 41–58, 2024, doi: 10.32985/ijeces.15.1.5.
- [23] Wang, S., Qu, J., Zhang, Y., & Zhang, Y. (2023). Multimodal emotion recognition from EEG signals and facial expressions. *IEEE Access*, 11, 33061–33068, 2023, doi: 10.1109/ACCESS.2023.3263670.
- [24] Koelstra, S., Muhl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., ... & Patras, I. (2011). Deap: A database for emotion analysis; using physiological signals. *IEEE transactions on affective computing*, 3(1), 18–31, 2011, doi: <https://doi.org/10.1109/T-AFFC.2011.15>.
- [25] Lin, C. T., Chang, C. J., Lin, B. S., Hung, S. H., Chao, C. F., & Wang, I. J. (2010). A real-time wireless brain–computer interface system for drowsiness detection. *IEEE transactions on biomedical circuits and systems*, 4(4), 214–222, 2010, doi: 10.1109/TBCAS.2010.2046415.
- [26] Vasudevan, K., & Das, A. P. (2017, July). Driver drowsiness monitoring by learning vehicle telemetry data. In 2017 10th International Conference on Human System Interactions (HSI) (pp. 270–276). IEEE, 2017, pp. 270–276, doi: 10.1109/HSI.2017.8005054.
- [27] Baltrušaitis, T., Ahuja, C., & Morency, L. P. (2018). Multimodal machine learning: A survey and taxonomy. *IEEE transactions on pattern analysis and machine intelligence*, 41(2), 423–443, 2018, doi: 10.1109/TPAMI.2018.2798607.
- [28] Zhou, X., Lin, D., Jia, Z., Xiao, J., Liu, C., Zhai, L., & Liu, Y. (2023, July). An EEG channel selection framework for driver drowsiness detection via interpretability guidance. In 2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) (pp. 1–5). IEEE, 2023, doi: 10.1109/EMBC40787.2023.10341126.
- [29] M. Winlaw, S. H. Steiner, R. J. MacKay, and A. R. Hilal, "Using telematics data to find risky driver behaviour," *Accident Analysis & Prevention*, vol. 131, pp. 131–136, 2019, doi: 10.1016/j.aap.2019.05.017.
- [30] Dong, Y., Hu, Z., Uchimura, K., & Murayama, N. (2010). Driver inattention monitoring system for intelligent vehicles: A review. *IEEE transactions on intelligent transportation systems*, 12(2), 596–614, 2010, doi: 10.1109/TITS.2010.2092770.

Authors' Profiles



Ch Raga Madhuri is an Assistant Professor in the Department of Computer Science and Engineering at Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, having held this position since April 2015. Alongside her teaching responsibilities, she is currently pursuing a Ph.D. degree from JNTUK, Kakinada. With over 10 years of teaching experience, she has gained extensive knowledge and expertise in artificial intelligence, machine learning, data analytics, and IoT fields. She has a distinguished research record, having published 25 papers in many national and international conferences, with proceedings indexed by IEEE and Scopus. In addition, she has published research papers in Scopus and UGC Indexed journals. She also holds three patents in her name.



Dedeepya Manikonda is a student in the Department of Computer Science and Engineering at Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, India, currently pursuing her Bachelor of Technology (B.Tech). Her research interests include Machine Learning and Artificial Intelligence. She has published a Scopus-indexed research paper in the field of crop price forecasting using deep learning techniques.



Madhurya Chinta is a student at Velagapudi Ramakrishna Siddhartha Engineering College, Kanuru, Vijayawada, India, pursuing her Bachelor of Technology in Computer Science and Engineering. Her research interests lie in machine learning and artificial intelligence.

How to cite this paper:

Ch Raga Madhuri, Dedeepya Manikonda, Madhurya Chinta, " Evolutionary-Optimized Multimodal Driver Risk Assessment Using DE-Enhanced Cross-Attention Fusion of EEG and Telematics", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.18, No.4, pp. 184-196, 2026. DOI:10.5815/ijigsp.2026.04.10