Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Hybrid Optuna and Q Learning for Adaptive Hyperparameter Optimization in EEG Seizure Classification

Upendra Kumar Swain
Silicon University, odisha
Shreela Dash
Silicon University, odisha
Asit Kumar Das
asitkumar.d@srmap.edu.in

SRM University, Andhra Pradesh

Research Article

Keywords: EEG Seizure Classification, Optuna, Q-learning, Hyperparameter Optimization, GRU–HiBERT–MLP, Explainable AI

Posted Date: November 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11218716/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Hybrid Optuna and Q Learning for Adaptive Hyperparameter Optimization in EEG Seizure Classification

Upendra Kumar Swain[1], Shreela Dash[1], Asit Kumar Das[2]*

[1] Department of Computer Science and Engineering, Silicon University, Bhubaneswar, Odisha, India

[2] Department of Computer Science and Engineering, SRM University-AP, Neerukonda, Mangalagiri, Andhra Pradesh, India

*Corresponding author: Asit Kumar Das | E-mail: asitkumar.d@srmap.edu.in

Abstract

Epileptic seizure classification using electroencephalogram (EEG) signals plays an important role in automated neurological diagnosis and patient monitoring. Although deep learning models have demonstrated promising performance for EEG-based seizure classification, their effectiveness strongly depends on efficient hyperparameter optimization and stable convergence behavior. Motivated by limitations in existing Optuna-optimized GRU–HiBERT–MLP framework, this paper proposes an adaptive hyperparameter optimization that combines Optuna and Q Learning for EEG seizure classification. The proposed framework integrates Bayesian optimization with reinforcement learning–based adaptive exploration to improve hyperparameter tuning efficiency convergence stability and overall classification performance. A GRU–HiBERT–MLP architecture is employed for preictal and ictal EEG classification using wavelet-transformed features extracted from the CHB-MIT scalp EEG dataset. Experimental results demonstrate that the proposed framework achieved a classification accuracy of 93.10% and a ROC–AUC score of 0.9774, outperforming standalone Optuna and Q-learning optimization approaches. Furthermore, convergence analysis demonstrated more efficient hyperparameter search and greater convergence stability, resulting in improved classification performance compared with standalone optimization approaches. SHAP and LIME explainability identify influential features and explain the model predictions for clinical EEG analysis. By integrating adaptive optimization and explainable artificial intelligence, the proposed system serves as an intelligent decision-support framework for epilepsy diagnosis and monitoring.

Keywords: ***EEG Seizure Classification, Optuna, Q-learning, Hyperparameter Optimization, GRU–HiBERT–MLP, Explainable AI***

I. INTRODUCTION AND MOTIVATION

Epilepsy is one of the commonest neurological illnesses globally, where the brain suffers recurrent abnormal electrical activity that could greatly interfere with mental and physiological functioning [1-2]. Electroencephalography (EEG) has been popularly used in the analysis of epileptic activities in the brain because of its non-invasiveness and temporal resolution capability [3-4]. Proper identification of preictal and ictal states of EEG has been significant in early seizure detection and patient observation. Seizure-state identification also contributes to the development of artificial intelligence clinical decision support systems [5]. EEG data are very non-stationary, noisy and complicated, which makes seizure state identification difficult in medical signal processing and intelligent systems [6].

The recent development of deep learning models has greatly boosted the performance of seizure classification methods using EEG signals. Deep architecture, including CNN, LSTM, GRU and transformer among others have been found efficient in learning temporal and spatial features of EEG multichannel data [7-11]. Hybrid approaches that utilize both recurrence and attention have shown great promise in representing complex temporal dynamics of EEG [12-13]. However, despite the success in developing effective deep architectures, their efficiency strongly relies on the proper selection of learning rate, hidden units, attention heads, dropout ratios, and dense layers among others [14].

Hyperparameter optimization has therefore become a critical component in developing robust EEG classification frameworks. Conventional optimization methods such as grid search and random search are computationally expensive and inefficient for high-dimensional search spaces [15]. Bayesian optimization frameworks, particularly Optuna have recently gained considerable attention due to their efficient probabilistic search strategies and automated tuning capability [16-17]. Recently, GRU–Transformer-based EEG seizure classification frameworks utilizing wavelet-transformed EEG features have demonstrated promising performance for preictal and ictal EEG analysis [18]. However, the optimization process primarily relied on Bayesian hyperparameter exploration and exhibited limitations related to adaptive search behavior, convergence stability and exploration–exploitation balancing during optimization [19].

Reinforcement learning techniques, particularly Q-learning, provide an adaptive optimization mechanism capable of learning efficient exploration strategies through reward-driven interactions with the environment [20-21]. Unlike purely probabilistic optimization approaches, Q-learning dynamically updates optimization decisions based on observed performance feedback, enabling improved adaptive exploration of the hyperparameter search space [22]. Integrating reinforcement learning with Bayesian optimization therefore presents a promising direction for improving convergence behavior and optimization efficiency in complex deep learning architectures for EEG analysis [23].

Motivated by these challenges, this paper proposes a Hybrid Optuna–Q Learning Optimization Framework for deep EEG seizure classification. The proposed framework integrates Bayesian optimization with reinforcement learning–based adaptive exploration to improve hyperparameter tuning efficiency, convergence stability and classification performance. A GRU–HiBERT–MLP architecture is employed for preictal and ictal EEG classification using wavelet-transformed features extracted from the CHB-MIT scalp EEG dataset [24]. The hybrid optimization framework adaptively refines critical hyperparameters during training, enabling improved navigation of the optimization search space and better exploration–exploitation balance.

The present work introduces a Hybrid Optuna–Q Learning optimization framework for enhancing optimization stability and adaptive hyperparameter exploration in GRU–HiBERT–MLP-based EEG seizure classification. By integrating reinforcement learning–based adaptive exploration with Bayesian hyperparameter optimization, the proposed framework improves convergence efficiency, optimization stability and classification robustness for preictal and ictal EEG analysis.

The main contributions of this work are summarized as follows:

- A Hybrid Optuna–Q Learning optimization framework is proposed for deep EEG seizure classification by integrating Bayesian hyperparameter optimization with reinforcement

learning–based adaptive exploration to improve convergence stability and optimization efficiency.

- A GRU–HiBERT–MLP classification architecture is employed for preictal and ictal seizure-state classification using wavelet-transformed multichannel EEG features extracted from scalp EEG recordings.
- Comparative experimental analysis demonstrates that the proposed hybrid optimization framework outperforms standalone Optuna and standalone Q-learning optimization approaches in terms of classification accuracy, ROC–AUC performance, convergence behavior and optimization stability.
- SHAP and LIME explainability techniques are incorporated to improve the interpretability, transparency, and reliability of the proposed framework for intelligent clinical EEG seizure analysis.

II. RELATED WORK

This section reviews prior studies related to EEG-based seizure analysis with emphasis on four closely related research directions: (i) deep learning approaches for EEG seizure classification, (ii) Transformer and hybrid architectures for EEG analysis, (iii) hyperparameter optimization techniques and (iv) reinforcement learning–based optimization strategies. The discussion highlights important limitations in existing approaches and provides the motivation for the proposed Hybrid Optuna–Q Learning optimization framework.

2.1 Deep Learning in Classification of Seizures Using EEG

The use of deep learning algorithms has greatly enhanced seizure classification using EEG through automating feature extraction processes and hierarchical learning from complex neural data [6-11]. Prior attempts to classify seizures using machine learning used hand-crafted EEG features and classifiers, which included SVM, random forests and k-Nearest Neighbor algorithm among others [3]. Even though these algorithms showed some level of efficiency in classifying seizures using EEG data, they largely depended on hand-crafted features and domain-specific processing techniques. More recent studies have focused on the use of deep neural network architectures in automating the seizure detection process and seizure classification. This includes the use of convolutional neural network architecture (CNN), recurrent neural network architecture (RNN), LSTM [25] and Gated Recurrent Unit (GRU) [26] [7-10]. While CNN has shown a high level of efficacy in learning spatial representations, RNN is effective in capturing temporal representations [28]. Among the recurrent neural architectures, Gated Recurrent Unit (GRU) has proven to be more efficient than LSTM when it comes to computation and ability to learn long-term temporal representation. Many studies have shown the effectiveness of GRU architecture in enhancing seizure state classification performance using EEG [12-13].

Despite this progress, however, EEG signals continue to exhibit significant non-stationarity and noise. In addition, the efficiency of using deep learning approaches is heavily influenced by the optimal choice and stability of hyperparameters.

2.2 Transformer and Hybrid Architectures for EEG Analysis

The emergence of transformer architecture models as advanced tools for EEG signal analysis is recent; however, this is attributed to the advantage of the self-attention process and the ability of transformers to handle long temporal dependencies . As much as transformers are widely used in

natural language processing and computer vision, they provide motivation for attention mechanisms in EEG classification models [10].
Hybrid architecture that involve a combination of CNN, GRU, LSTM and Transformer networks in different ways for the purposes of classification and prediction of epileptic seizures have been suggested by several researchers. In particular, combining recurrent learning with attention mechanisms has demonstrated improved capability in modeling complex EEG behavior and seizure-state transitions [27].

Although Transformer and hybrid architectures improve classification performance increasing architectural complexity introduces additional challenges related to hyperparameter optimization, convergence stability and computational efficiency. Consequently, optimization strategies capable of effectively exploring high-dimensional hyperparameter spaces are becoming increasingly important for deep EEG classification systems.

2.3 Hyperparameter Optimization in Deep Learning

Hyperparameter tuning is very important in the enhancement of the model's generalization abilities and performance. Learning rate, hidden layers, attention heads, dropout rates, regularization factors and batch size have a huge impact on how well the model converges and its accuracy in classification tasks [14]. Conventional optimization approaches such as grid search and random search are computationally expensive and inefficient for large-scale deep learning architectures [15].
Bayesian optimization frameworks, particularly Optuna, have recently gained considerable attention for automated hyperparameter tuning in deep learning applications [16]. Optuna employs Tree-structured Parzen Estimators (TPE) and pruning strategies to efficiently explore hyperparameter search spaces while reducing computational overhead [16-17]. Several EEG classification studies have utilized Optuna to optimize deep neural architectures and improve seizure detection performance [16-17].

Recently, Optuna-optimized GRU-HiBERT-MLP frameworks utilizing wavelet-transformed EEG representations have demonstrated promising seizure-state classification performance [18]. Although the framework demonstrated promising classification capability, the Bayesian optimization process exhibited limitations associated with adaptive exploration and convergence stability in the hyperparameter search space [20]. These observations motivate the development of enhanced optimization strategies capable of improving exploration–exploitation balance during hyperparameter tuning.

2.4 Reinforcement Learning–Based Optimization

The concept of reinforcement learning has been identified as an attractive optimization technique that has become increasingly popular in recent times because of its ability to enable adaptation in decision making and dynamic exploration of a domain [20]. Of all reinforcement learning techniques, Q-learning is one that has become very popular owing to its simplicity and ability to learn optimal behavior by interacting with the environment [21].

Recent studies have explored reinforcement learning-based optimization for neural architecture search, adaptive model design, and automated machine learning tasks [29–30]. Unlike purely probabilistic optimization approaches, reinforcement learning dynamically balances exploration and exploitation, potentially improving convergence stability and reducing the risk of local optima [22]. However, standalone reinforcement learning optimization may exhibit slower convergence

behavior and increased exploration overhead when applied to large-scale deep learning architectures.

Integrating reinforcement learning with Bayesian optimization therefore presents a promising direction for improving optimization efficiency, adaptive exploration capability and convergence behavior in complex EEG classification systems.

2.5 Research Gap and Motivation

Although deep learning and Transformer-based architectures have significantly improved EEG seizure classification performance, existing optimization frameworks still face challenges related to convergence stability, computational efficiency and adaptive hyperparameter exploration. Bayesian optimization approaches such as Optuna provide efficient probabilistic search mechanisms but may suffer from premature convergence and insufficient adaptive exploration in highly complex optimization landscapes. Conversely, reinforcement learning techniques provide adaptive exploration capability but may exhibit slower convergence behavior when applied independently.

Motivated by these limitations, this work proposes a Hybrid Optuna-Q Learning Optimization Framework that integrates Bayesian hyperparameter optimization with reinforcement learning–based adaptive exploration. The proposed framework aims to improve convergence stability, optimization efficiency, and classification performance for deep EEG seizure classification using a GRU-HiBERT-MLP architecture.

III. METHODOLOGY

This section presents the proposed Hybrid Optuna-Q Learning optimization framework for deep EEG seizure classification. The overall framework integrates EEG preprocessing, wavelet-based feature extraction, a GRU-HiBERT-MLP classification architecture and a hybrid optimization strategy combining Bayesian optimization with reinforcement learning-based adaptive exploration. Initially, multichannel EEG signals are preprocessed and transformed using Discrete Wavelet Transform (DWT) to obtain informative time–frequency representations. The extracted features are subsequently classified using the GRU-HiBERT-MLP framework, which combines recurrent learning and self-attention mechanisms for effective seizure-state modeling. To improve hyperparameter tuning efficiency and convergence stability, Optuna-based Bayesian optimization is integrated with a Q-learning agent that adaptively guides the hyperparameter search process using reward-driven feedback. The proposed framework aims to enhance exploration–exploitation balance, improve convergence behavior, and achieve robust preictal and ictal EEG classification performance. **Figure 1** illustrates the overall architecture of the proposed optimization-enhanced EEG seizure classification framework.

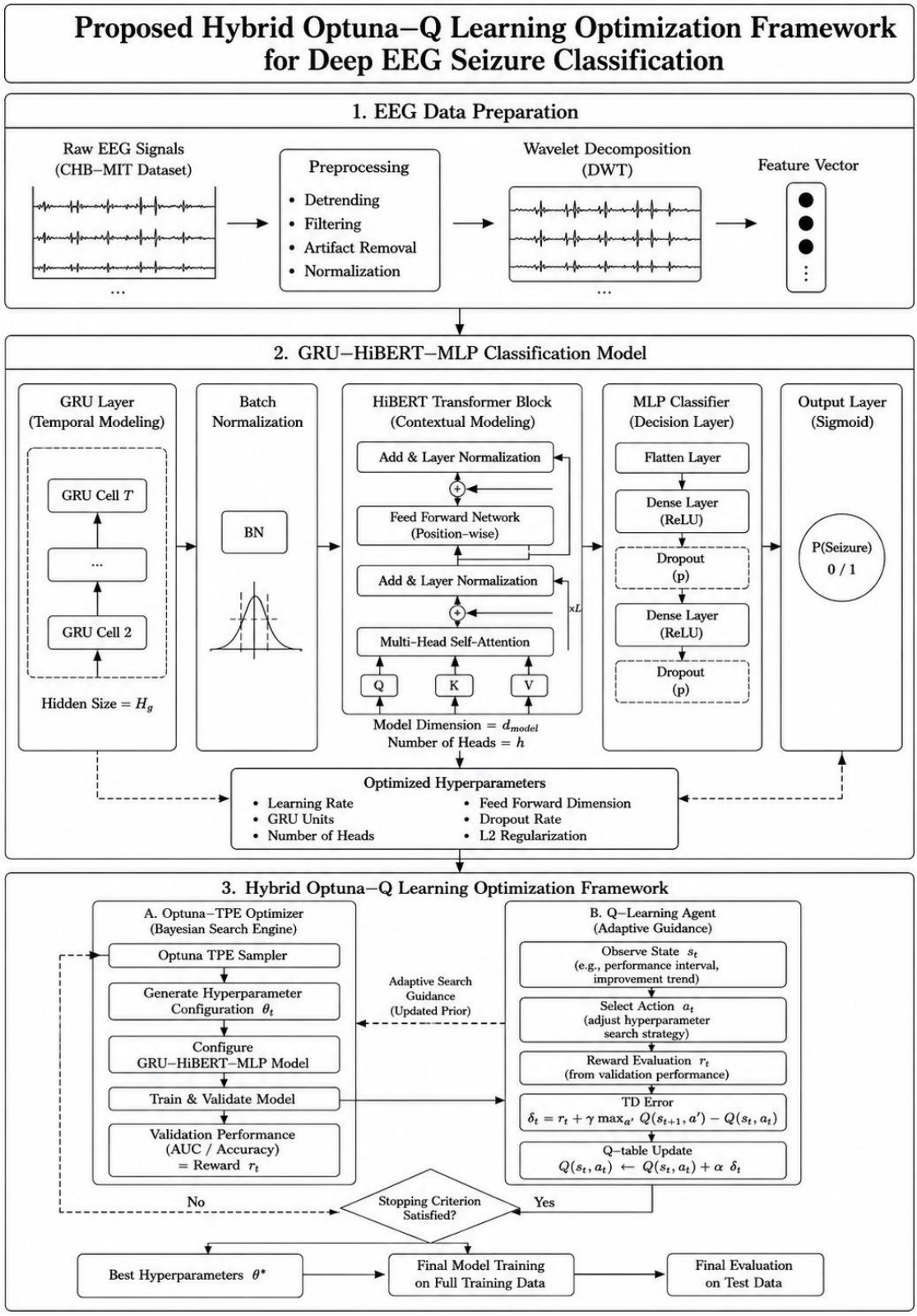


Figure 1. Proposed hybrid Optuna-Q learning framework

3.1 EEG Dataset and Preprocessing

To evaluate the proposed framework, the Preprocessed CHB-MIT Scalp EEG Database obtained from IEEE DataPort was utilized. The dataset derived from the publicly available CHB-MIT EEG database, comprises balanced preictal and ictal EEG recordings collected from 24 pediatric epilepsy patients under clinical monitoring conditions. To ensure consistency across recordings, 23 common EEG channels were retained, and the signals were subsequently transformed into

wavelet-based feature representations. After preprocessing and segmentation, the final dataset used in this study consisted of 1,637 EEG segments for seizure-state classification.

Prior to model training, EEG signals underwent preprocessing and feature extraction to enhance signal representation and reduce the influence of noise and signal variability. Considering that EEG signals have very non-stationary behavior and include many time-frequency features, the Discrete Wavelet Transform (DWT) technique was applied for obtaining effective time-frequency representations from raw EEG data. This study used the Daubechies wavelets (db4) for extracting information due to their efficiency in handling transient behavior of EEG signals. The maximum decomposition level was determined adaptively based on the signal length and wavelet filter characteristics. The resulting approximation and detail coefficients obtained from the wavelet decomposition process were concatenated to construct discriminative feature vectors for seizure-state analysis.

Following wavelet-based feature extraction, StandardScaler normalization was applied to standardize the extracted EEG features by removing the mean and scaling the data to unit variance. The normalization process improves training stability, accelerates convergence behavior, and reduces the influence of feature-scale variations during deep learning optimization. Mathematically, the standardized feature transformation is expressed as

$$X_{scaled} = \frac{X-\mu}{\sigma} \qquad (1)$$

where X represents the original feature vector, μ denotes the meaning of the feature distribution, and σ represents the standard deviation.

After normalization, the processed EEG feature vectors were reshaped into a three-dimensional structure suitable for sequential deep learning analysis and subsequently provided as input to the proposed GRU-HiBERT-MLP classification framework. Finally, the dataset was divided into training and testing subsets using an 80:20 split ratio to evaluate the generalization capability and classification performance of the proposed optimization-enhanced EEG seizure classification framework.

3.2 Proposed GRU-HiBERT-MLP Classification Architecture

The GRU-HiBERT-MLP classification framework employed for preictal and ictal EEG seizure-state classification is illustrated in Figure 2. The framework integrates GRU-based temporal modeling, contextual representation learning, and multilayer nonlinear classification for discriminative EEG feature analysis. Since the primary contribution of the present work focuses on optimization enhancement through the proposed Hybrid Optuna-Q Learning framework, only a concise architectural overview is presented.

Initially, preprocessed wavelet-transformed EEG feature sequences are provided as input to the GRU layer for temporal dependency learning from multichannel EEG signals. The GRU hidden representation is computed as

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \qquad (2)$$

where h_t denotes the hidden state at time step t, z_t represents the update gate, and $\tilde{h}_t$ corresponds to the candidate hidden representation.

Following GRU-based temporal modeling, batch normalization is applied to stabilize feature distributions and improve convergence behavior during training. The normalized feature representations are subsequently forwarded to the HiBERT-inspired Transformer block for contextual EEG modeling using multi-head self-attention mechanisms. The self-attention operation is expressed as

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (3)$$

where Q, K and V denote the query, key and value matrices, respectively and d_k represents the scaling factor used during attention computation.

The contextual EEG representations generated by the Transformer block are subsequently flattened and forwarded to fully connected dense layers for nonlinear seizure-state classification. Dropout regularization is incorporated to reduce overfitting and improve model generalization. Finally, a sigmoid-activated output layer performs binary seizure-state classification between preictal and ictal EEG patterns.

3.3 Proposed Hybrid Optuna-Q Learning Optimization Framework

The performance of deep learning architectures for EEG seizure classification strongly depends on effective hyperparameter optimization and stable convergence behavior during training. Although Bayesian optimization frameworks such as Optuna provide efficient probabilistic hyperparameter exploration, they may suffer from premature convergence and insufficient adaptive exploration when handling complex high-dimensional search spaces. To address these limitations, this work proposes a Hybrid Optuna-Q Learning optimization framework that integrates Bayesian optimization with reinforcement learning-based adaptive exploration to improve convergence stability, optimization efficiency, and classification performance.

Let the hyperparameter search space of the proposed GRU–HiBERT–MLP architecture be represented as

$$\Theta = \{\theta_1, \theta_2, \theta_3, \dots, \theta_n\} \quad (4)$$

where Θ denotes the overall hyperparameter search space including GRU units, attention heads, feed-forward dimensions, dropout rates, learning rates and dense layer configurations. The primary optimization objective is to determine the optimal hyperparameter configuration θ^*that maximizes the validation performance of the EEG classification framework, which can be formulated as

$$\theta^* = argmax_{\theta \in \Theta} f(\theta) \quad (5)$$

Where $f(\theta)$ represents the objective evaluation function corresponding to validation accuracy and ROC-AUC performance. Initially, the Optuna-TPE optimizer probabilistically samples candidate hyperparameter configurations from the search space according to

$$\theta_t \sim p(\theta \mid D_t) \quad (6)$$

where D_t denotes the historical optimization observations obtained from previous trials, and $p(\theta \mid D_t)$ represents the posterior probability distribution estimated using the Tree-structured Parzen Estimator (TPE). The generated hyperparameter configuration θ_t is subsequently used to train the GRU-HiBERT-MLP classification model.

Unlike standalone Bayesian optimization approaches, the proposed framework integrates a reinforcement learning-based Q-learning agent to dynamically guide the hyperparameter search process. In the proposed framework, the optimization state s_t represents the current optimization condition, while the action a_t corresponds to adaptive hyperparameter refinement strategies. The reward r_t is computed based on improvements in validation performance achieved during optimization and is defined as

$$r_t = Acc_t - Acc_{t-1} \tag{7}$$

where Acc_t and Acc_{t-1} denote the current and previous validation accuracies respectively. The Q-learning agent updates the Q-table using temporal difference learning according to

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_t + \gamma maxQ(s_{t+1}, a') - Q(s_t, a_t)] \tag{8}$$

where $Q(s_t, a_t)$ represents the Q-value associated with the state-action pair, α denotes the learning rate, γ represents the discount factor, and r_t corresponds to the obtained validation reward. This update mechanism enables the reinforcement learning agent to progressively learn effective exploration–exploitation strategies during hyperparameter optimization. The adaptive interaction between Optuna and Q-learning can be expressed as

$$\theta_{t+1} = G(\theta_t, Q(s_t, a_t)) \tag{9}$$

where G denotes the adaptive guidance mechanism that refines subsequent hyperparameter exploration based on the learned Q-values. Consequently, the Q-learning agent dynamically guides the Optuna search strategy toward more promising optimization regions while reducing premature convergence and improving convergence stability.

By integrating probabilistic Bayesian optimization with reinforcement learning-based adaptive exploration, the proposed Hybrid Optuna-Q Learning framework enables robust navigation of the hyperparameter search space and contributes to improved EEG seizure-state classification performance for the GRU-HiBERT-MLP architecture.

3.4 Algorithmic Workflow

To provide a systematic representation of the proposed optimization-enhanced EEG classification framework, **Algorithm 1** summarizes the overall workflow of the Hybrid Optuna-Q Learning optimization process. The algorithm integrates Bayesian hyperparameter exploration, reinforcement learning-based adaptive guidance, and GRU-HiBERT-MLP model training for robust EEG seizure-state classification.

Algorithm 1. Hybrid Optuna-Q learning optimization workflow

Input: Preprocessed EEG dataset X, labels Y, hyperparameter search space Θ
Output: Optimized GRU-HiBERT-MLP model

1. *Initialize Optuna-TPE optimizer*
2. *Initialize Q-table Q(s,a)*
3. *Define hyperparameter search space Θ*
4. *Preprocess EEG signals using DWT and normalization*
5. *Split dataset into training and testing subsets*
6. *For each optimization trial t=1→T*
7. *Generate candidate hyperparameter configuration θ_t using Optuna-TPE*
8. *Configure GRU-HiBERT-MLP model using θ_t*
9. *Train model on EEG training dataset*

10. Evaluate validation performance Acc_t and ROC–AUC
11. Compute reward
$r_t = Acc_t - Acc_{t-1}$
12. Observe optimization state s_t
13. Select adaptive action a_t using Q-learning policy
14. Update Q-table using
$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_t + \gamma maxQ(s_{t+1}, a') - Q(s_t, a_t)]$
15. Adapt Optuna search guidance using updated Q-values
16. End For
17. Select optimal hyperparameter configuration θ^*
18. Train final GRU-HiBERT-MLP model using optimized parameters
19. Generate final seizure-state classification results

IV. EXPERIMENTAL RESULTS AND ANALYSIS

4.1 Experimental Setup

The proposed Hybrid Optuna-Q Learning optimization framework was implemented using Python with TensorFlow and Keras deep learning libraries. All experiments were conducted on the Preprocessed CHB-MIT Scalp EEG Database. The dataset consisted of 23-channel EEG recordings segmented into fixed-length windows for seizure-state classification.and it obtained from IEEE DataPort for preictal and ictal seizure-state classification. The extracted wavelet-transformed EEG features were standardized using StandardScaler normalization and reshaped into sequential input representations suitable for GRU-based temporal learning.
The dataset was divided into training and testing subsets using an 80:20 split ratio to evaluate the generalization capability of the proposed framework. The GRU-HiBERT-MLP architecture was trained using the Adam optimizer with binary cross-entropy loss for binary seizure-state classification. These methods have been adopted to ensure stability during training while reducing overfitting. A batch size of 32 and 150 maximum training epochs have been adopted for the model.

For hyperparameter optimization, the proposed framework employed Optuna-based Bayesian optimization integrated with reinforcement learning-based Q-learning guidance. The optimization search space included GRU units, attention heads, feed-forward dimensions, dropout rates, learning rates, and dense layer configurations. Multiple optimization trials and Q-learning episodes were performed to adaptively refine the hyperparameter search process and improve convergence behavior.

The performance of the proposed framework was evaluated using classification accuracy, precision, recall, F1-score, ROC-AUC, confusion matrix analysis, and convergence behavior. Comparative experiments were additionally conducted against standalone Optuna and Q-learning optimization strategies to validate the effectiveness of the proposed hybrid optimization framework.

The proposed framework was implemented using Python 3.x and TensorFlow 2.x. Experimental evaluations were conducted on a workstation equipped with an Intel Core i7 processor, 16 GB RAM, and NVIDIA GPU acceleration.

Table 1. Optimal Hyperparameter Configuration Obtained by the Proposed Hybrid Optuna–Q Learning Framework

Hyperparameter	Optimized Value
GRU Units	192
Attention Heads	2
Feed-Forward Dimension (DFF)	256
Dense Units	1024
Dropout Rate	0.14
Learning Rate	1.02×10^{-4}
Weight Decay	1.28×10^{-5}
Mixup Alpha	0.26

The optimal hyperparameter configuration identified by the proposed Hybrid Optuna–Q Learning optimization framework is summarized in **Table 1**. The framework selected 192 GRU units, 2 attention heads, a feed-forward dimension of 256 and 1024 dense units. In addition, the optimization process automatically determined suitable values for dropout regularization, learning rate, weight decay and Mixup augmentation parameters. The optimized configuration was subsequently employed for training and evaluation of the final GRU–HiBERT–MLP seizure classification model.

4.2 Evaluation Metrics

The performance of the proposed Hybrid Optuna-Q Learning optimized GRU-HiBERT-MLP framework was evaluated using several widely adopted classification metrics, including accuracy, precision, recall, F1-score and Receiver Operating Characteristic-Area Under Curve (ROC-AUC). These evaluation metrics provide comprehensive assessment of classification capability, robustness, and discriminative performance for preictal and ictal EEG seizure-state analysis.

Classification accuracy measures the overall proportion of correctly classified EEG samples and is defined as

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

where TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively. Precision evaluates the proportion of correctly predicted positive seizure states among all predicted positive samples and is expressed as

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

Recall, also referred to as sensitivity, measures the ability of the proposed framework to correctly identify seizure-related EEG states and is computed as

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

The F1-score represents the harmonic mean of precision and recall, thereby providing balanced evaluation of classification performance under varying prediction conditions. It is formulated as

$$F1\ score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

In addition to threshold-based metrics, Receiver Operating Characteristic-Area Under Curve (ROC-AUC) analysis was employed to evaluate the discriminative capability of the proposed framework across different classification thresholds. Higher ROC-AUC values indicate improved separability between preictal and ictal EEG states and reflect stronger classification robustness.

Furthermore, convergence behavior, optimization stability and statistical significance analysis were also considered to evaluate the effectiveness of the proposed Hybrid Optuna-Q Learning optimization framework in comparison with standalone Optuna and Q-learning optimization strategies.

4.3 Comparative Performance Analysis and Ablation Study

4.3.1 Training and Validation Performance.

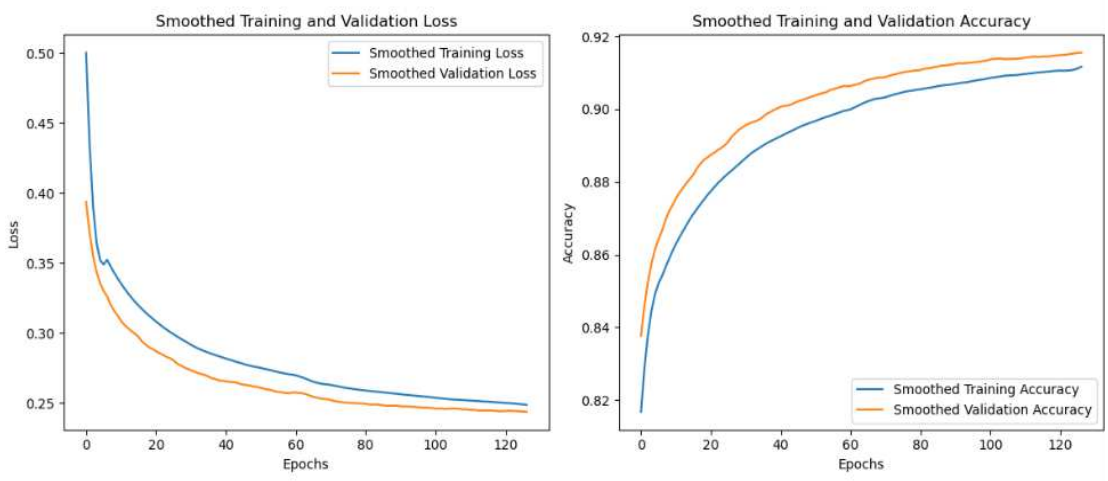


Figure 2. Training vs Validation Loss (Optuna-only)

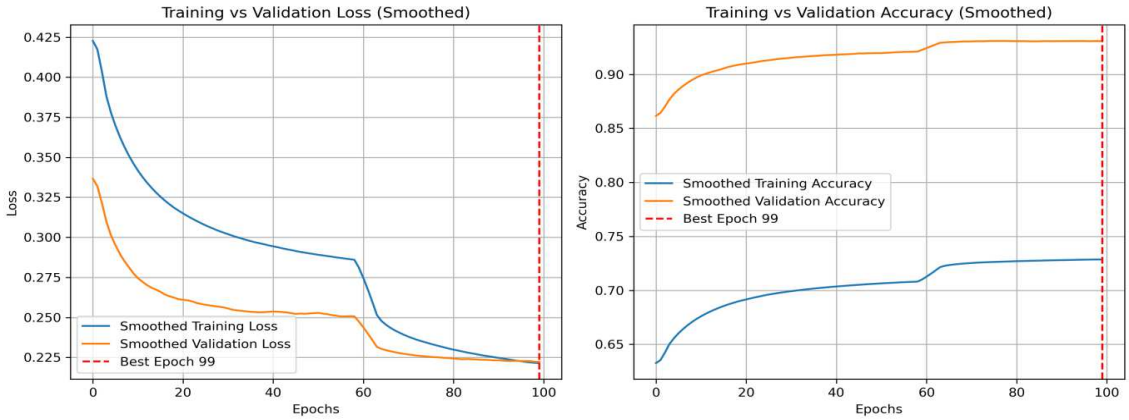


Figure 3. Training vs Validation Loss (Hybrid Optuna–Q Learning)

The training and validation loss curves obtained for Optuna-based hyperparameter optimization are also represented in **Figure 2**. On one hand, the model converges smoothly on the other hand, the validation curves show significant fluctuations in terms of values. This characterizes a very high sensitivity of such curves with respect to the selected hyperparameters. Such fluctuation proves the low generalization stability of a strategy based only on Optuna.

The training and validation performance of the Hybrid Optuna-Q Learning optimized model is displayed in **Figure 3**. This hybrid framework achieves faster convergence and smoother learning curves compared to previous iterations, while significantly narrowing the gap between training and validation losses. Such enhanced stability stems from the reinforcement learning-guided selection of hyperparameters which effectively filters out unstable configurations.

4.3.2 ROC Analysis and Classification Performance

The ROC curve of Optuna-only optimization strategy, presented in Figure 4(a), provides an AUC of nearly 0.971. Although this suggests a good level of classification power, the fairly gradual trajectory towards the top-left corner suggests that there exists a particular shortcoming on the part of this strategy, according to which this method performs inadequately on different complex variations of the EEG classes.

The ROC Curve of the proposed Hybrid Optuna-Q Learning Framework is depicted as Figure 4(b**)** below. Here, it can be noticed that there is an improved AUC of 0.9774. This is due to the increased class separability as well as the reduced number of false positive as well as false negative values. The main reason for this is the reinforcement learning-based optimization process that allows it to search for better values of hyper-parameters.

The Precision Recall (PR) curve of the optimized hybrid model, as shown in Figure 4(c), indicates a high value of PR-AUC with a high precision of 0.9784. This clearly confirms the efficacy of the developed framework, especially within the medical setup where the need for a lower false alarm rate is a pressing requirement for valid seizure detection. Overall, the results of the comparative study clearly confirm the superiority of the hybrid Optuna and Q-learning model over the Optuna model regarding the robustness of the model's classification and discriminatory power. This improvement in the ROC & PR curve clearly confirms the utility of the inclusion of reinforcement learning within the optimization process and asserts the proposed model as a highly efficacious technique for EEG signal applications.

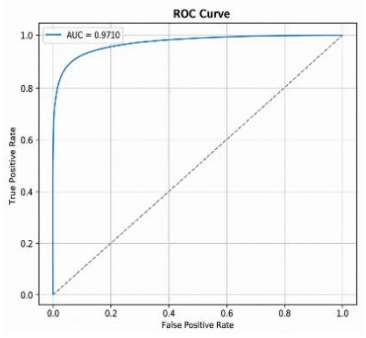


Figure 4(a). ROC curve obtained using the Optuna-only optimization framework

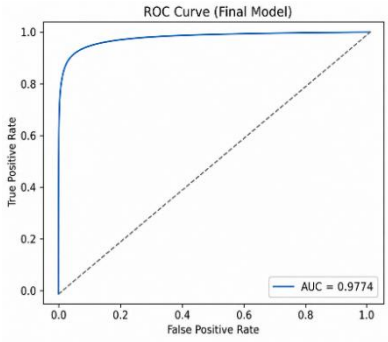


Figure 4(b). ROC curve obtained using the proposed Hybrid Optuna-Q Learning optimization framework.

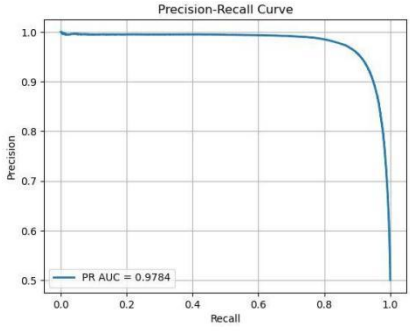


Figure 4(c). Precision-Recall (PR) curve of the proposed Hybrid Optuna-Q Learning framework.

4.3.3 Confusion Matrix Analysis

The confusion matrix for the Hybrid Optuna-Q Learning optimized model, as evident from Figure 5, indicates equitable classification accuracy for both classes: preictal and ictal. It can be noted that the optimized model has resulted in a dramatic decrease in the number of false negatives, which forms an essential aspect for being safe within the medical field. The equitable classification capability illustrates the efficiency of the proposed approach for its intended application.

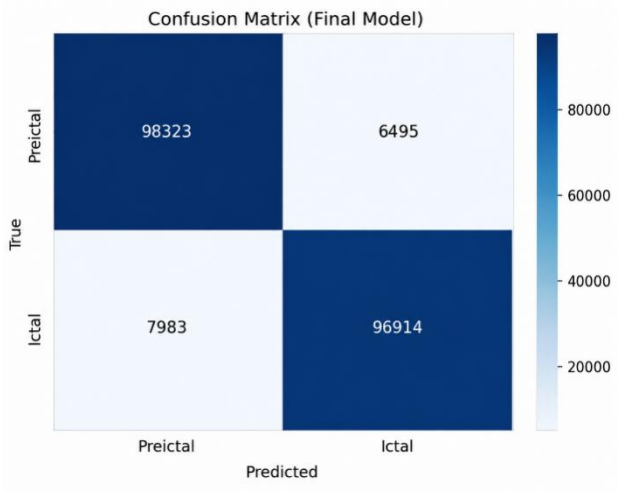


Figure 5. Confusion Matrix of Hybrid Optimization

The hybrid model exhibits better classification balance and lower misclassification when compared to the baseline method, especially when it comes to reducing false negatives. This is crucial for accurate identification of seizure activity, which is necessary for dependable seizure detection. Overall, the suggested framework's robustness and dependability are confirmed by the confusion matrix.

4.3.4 Ablation Study

The experimental results demonstrate in **Table 2**. that the proposed Hybrid Optuna + Q-learning optimization strategy achieved the best overall performance, obtaining an accuracy of 93.10% and ROC-AUC of 0.9774 while also exhibiting faster convergence and higher stability across trials compared to standalone Optuna and Q-learning approaches.

Table 2.
Comparative Performance Analysis of Different Hyperparameter Optimization Strategies for EEG-based Preictal vs. Ictal Classification.

Metric	Optuna-only	Q-learning Only	Hybrid Optuna + Q-learning
Accuracy (%)	92.00	91.51	**93.10**
Precision	0.9200	0.9347	**0.9300**
Recall	0.9200	0.8926	**0.9300**
F1-score	0.9200	0.9132	**0.9300**
ROC–AUC	0.9710	0.970	**0.9774**
Convergence Speed	Moderate	Slower	**Faster**

Stability Across Trials	Moderate	Moderate	**High**
Optimization Strategy	Bayesian Optimization (TPE)	Reinforcement Learning	**Hybrid TPE + Q-learning**

4.4 Convergence Analysis

To further investigate the optimization behavior and convergence stability of the proposed framework, a convergence comparison was conducted among standalone Optuna optimization, standalone Q-learning optimization, and the proposed Hybrid Optuna–Q Learning optimization strategy

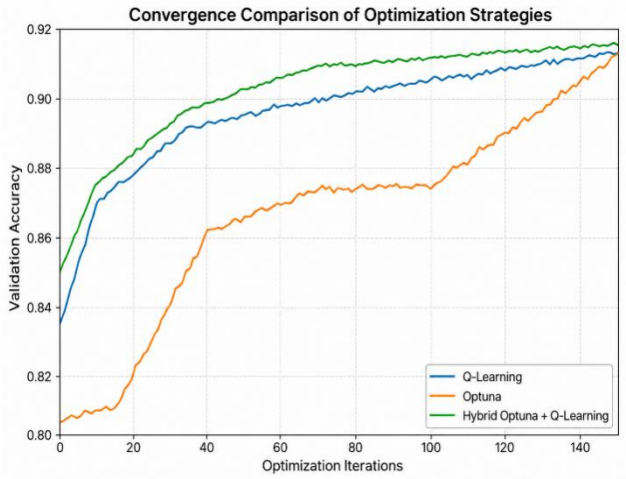


Figure 6. Convergence comparison of standalone Optuna optimization, standalone Q-learning optimization, and the proposed Hybrid Optuna–Q Learning framework in terms of validation accuracy during optimization iterations.

As observed in **Figure 6**, the standalone Optuna optimization framework demonstrates moderate convergence behavior during the initial optimization stages due to probabilistic Bayesian hyperparameter exploration. Although the Optuna-based approach gradually improves validation accuracy over successive optimization iterations, its convergence stability remains comparatively moderate because of fluctuations in the hyperparameter search process. Similarly, the standalone Q-learning framework exhibits relatively stable adaptive exploration capability; however, the convergence process remains comparatively slower during later optimization stages.

In contrast, the proposed Hybrid Optuna-Q Learning framework achieves faster and more stable convergence behavior throughout the optimization process. The integration of reinforcement learning–based adaptive guidance with Bayesian hyperparameter exploration enables the proposed framework to effectively balance exploration and exploitation during optimization. Thus, the hybrid approach tends to converge towards the best hyperparameters faster and with greater optimization stability at each iteration.

Additionally, the hybrid approach has shown higher validation accuracies than any independent optimization method, validating the significance of using reinforcement learning for hyperparameter tuning. The convergence analysis therefore validates that the proposed Hybrid Optuna–Q Learning optimization strategy enhances convergence stability, accelerates optimization efficiency, and contributes to improved EEG seizure-state classification performance for preictal and ictal EEG analysis.

4.5 Explainability Analysis

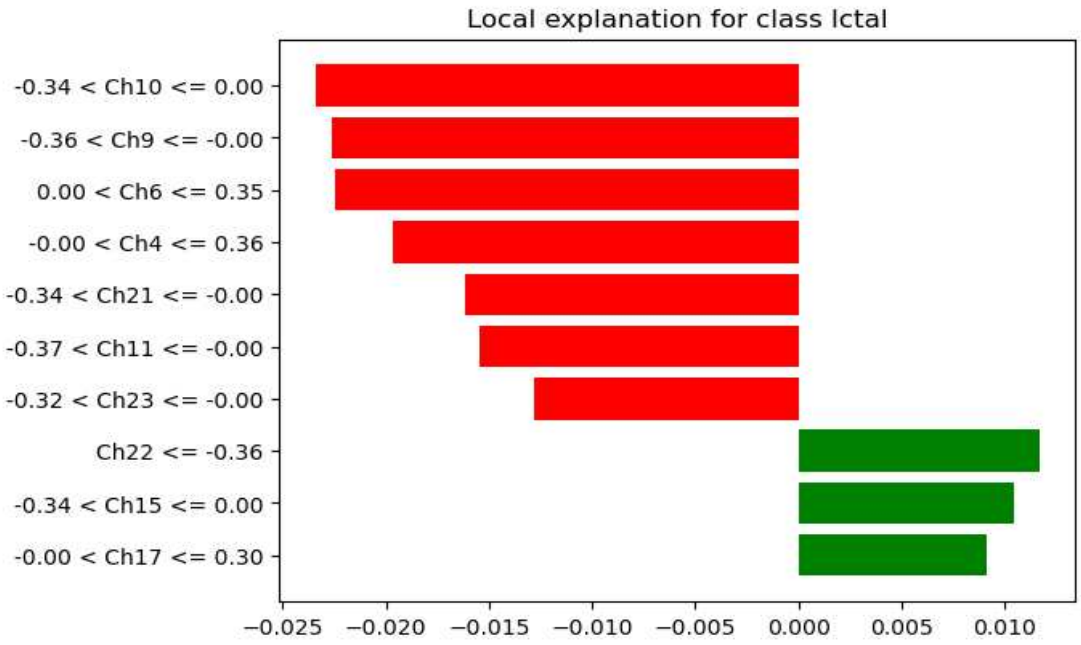


Figure 7. LIME-based explanation

.

The LIME explanation for a typical ictal EEG sample is shown in **Figure 7**. The most significant EEG characteristics influencing the model's choice are highlighted in the visualization. Positive and negative contributions show how each channel influences the prediction result, proving that the suggested model depends on significant signal properties rather than arbitrary patterns.

4.6 Discussion

The experimental results demonstrate that the proposed Hybrid Optuna-Q Learning optimization framework significantly improves the performance of the GRU-HiBERT-MLP architecture for EEG seizure-state classification. Comparative analysis against standalone Optuna and standalone Q-learning optimization approaches confirms that the proposed hybrid framework achieves superior classification accuracy, improved ROC-AUC performance and enhanced convergence stability. The observed improvements indicate that the integration of Bayesian hyperparameter exploration with reinforcement learning-based adaptive guidance provides a more effective optimization strategy for complex EEG classification models.

One of the major observations from the experimental analysis is the improved convergence behavior achieved by the proposed hybrid optimization framework. Bayesian optimization algorithms like Optuna perform effective exploration in the search space of hyperparameters; nonetheless, there may be cases of instability and early convergence in the optimization problem that is high-dimensional. On the other hand, reinforcement learning techniques, which include Q-learning, can provide adaptive exploration as well, although there will be additional optimization iterations in order to ensure stability in convergence.

Convergence analysis also shows that the proposed approach provides a much more smooth and faster rate of convergence than the individual optimization methods. In this context, the reinforcement learning model automatically tunes the process of hyperparameter tuning in Optuna according to its validation rewards. Consequently, the proposed framework reduces unstable

optimization behavior while maintaining improved validation performance across optimization iterations. These findings validate the effectiveness of reinforcement learning-guided Bayesian optimization for deep EEG classification tasks.

From a clinical point of view the better classification ability of the proposed approach is of relevance for automated seizure state detection and intelligent EEG monitoring systems. It is essential for accurate detection of preictal and ictal EEG states due to the significance of this distinction in seizure prediction and intervention. The use of the SHAP and LIME interpretability approaches makes the proposed model more understandable due to the identification of EEG features responsible for making classification decisions.

However, there are some limitations associated with this approach despite its good performance. Firstly, this research has mainly been performed on the Preprocessed CHB-MIT Scalp EEG Database, and more validation can be done on large EEG datasets from multiple centers to enhance the generalization ability of the framework. Another possible research direction in the future is to consider the lightweight training and real-time implementation of the framework along with using reinforcement learning approaches.

V. LIMITATION AND FUTURE WORK

Although the proposed Hybrid Optuna-Q Learning framework achieved promising performance for EEG seizure-state classification several limitations remain. The current study was evaluated primarily on the Preprocessed CHB-MIT Scalp EEG Database, and additional validation using larger multi-center EEG datasets would further strengthen the generalization capability and clinical applicability of the framework. Furthermore, the integration of Bayesian optimization and reinforcement learning introduces increased computational complexity and training overhead during hyperparameter optimization.

The present work focuses on binary classification of preictal and ictal EEG states. Future research may extend the proposed framework toward multi-class seizure analysis, patient-independent learning, and real-time seizure prediction systems. Additionally, advanced reinforcement learning strategies and lightweight optimization techniques may be explored to improve optimization efficiency and facilitate deployment in real-time clinical EEG monitoring environments. Further investigation into clinically interpretable EEG biomarkers and explainable AI frameworks may also enhance the transparency and reliability of automated seizure analysis systems.

VI. CONCLUSION

This paper presented a Hybrid Optuna–Q Learning optimization framework for GRU-HiBERT-MLP based EEG seizure-state classification using wavelet-transformed EEG features. The proposed framework integrated Bayesian hyperparameter optimization with reinforcement learning-based adaptive exploration to improve convergence stability and classification performance for preictal and ictal EEG analysis.

Experimental results on the Preprocessed CHB-MIT Scalp EEG Database demonstrated that the proposed hybrid framework outperformed standalone Optuna and standalone Q-learning optimization approaches in terms of classification accuracy, ROC-AUC performance and convergence behavior. The integration of SHAP and LIME explainability techniques additionally improved model interpretability and transparency for EEG seizure-state analysis.

The proposed Hybrid Optuna-Q Learning framework provides an effective optimization strategy for deep EEG classification and demonstrates the potential of combining Bayesian optimization with reinforcement learning for intelligent biomedical signal analysis applications.

DECLARATIONS

Funding

The authors received no funding for this work.

Author Contributions

Upendra Kumar Swain contributed to the conceptualization, methodology, software implementation, experimentation, data analysis, visualization, and preparation of the original manuscript. Shreela Dash supervised the research, provided methodological guidance, and contributed to the review and revision of the manuscript. Asit Kumar Das provided academic guidance and contributed to the review and revision of the manuscript. All authors read and approved the final manuscript.

Ethics and Consent to Participate

Not applicable. This study involved secondary analysis of the publicly available Preprocessed CHB-MIT Scalp EEG Database obtained from IEEE DataPort and did not involve direct recruitment of participants or collection of new human-subject data by the authors.

Consent to Publish

Not applicable.

Competing Interests

The authors declare that they have no competing interests.

Data Availability

The EEG dataset used in this study is the publicly available Preprocessed CHB-MIT Scalp EEG Database, obtained, from IEEE DataPort. The dataset is available at https://ieee-dataport.org/open-access/preprocessed-chb-mit-scalp-eeg-database (DOI: 10.21227/awcw-mn88).

REFERENCES

[1] World Health Organization, “Epilepsy”, *WHO Fact Sheet*, 2023.

[2] R. S. Fisher et al., “ILAE official report: A practical clinical definition of epilepsy”, *Epilepsia*, vol. 55, no. 4, pp. 475-482, 2014.

[3] A. Subasi, “EEG signal classification using wavelet feature extraction and machine learning methods”, *Expert Systems with Applications*, vol. 32, no. 4, pp. 1084-1093, 2007.

[4] H. Adeli, Z. Zhou, and N. Dadmehr, “Analysis of EEG records in an epileptic patient using wavelet transform”, *Journal of Neuroscience Methods*, vol. 123, no. 1, pp. 69-87, 2003.

[5] U. R. Acharya et al., “Automated diagnosis of epilepsy using entropies”, *Biomedical Signal Processing and Control*, vol. 7, no. 4, pp. 401-408, 2012.

[6] S. Roy et al., “Deep learning-based electroencephalography analysis: A systematic review”, *Journal of Neural Engineering*, vol. 16, no. 5, 2019.

[7] V. J. Lawhern et al., “EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces”, *Journal of Neural Engineering*, vol. 15, no. 5, 2018.

[8] A. Graves et al., “Speech recognition with deep recurrent neural networks”, *IEEE ICASSP*, 2013.

[9] K. Cho et al., “Learning phrase representations using RNN encoder-decoder for statistical machine translation”, *arXiv preprint arXiv:1406.1078*, 2014.

[10] A. Vaswani et al., “Attention is all you need”, *Proceedings of NeurIPS*, 2017.

[11] M. Schirrmeister et al., “Deep learning with convolutional neural networks for EEG decoding and visualization”, *Human Brain Mapping*, vol. 38, no. 11, pp. 5391-5420, 2017.

[12] H. Bashivan et al., “Learning representations from EEG with deep recurrent-convolutional neural networks”, *arXiv preprint arXiv:1511.06448*, 2015.

[13] S. Raghu et al., “EEG-based multi-class seizure classification using deep learning techniques”, *Biomedical Signal Processing and Control*, vol. 70, 2021.

[14] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.

[15] J. Bergstra and Y. Bengio, “Random search for hyper-parameter optimization”, *Journal of Machine Learning Research*, vol. 13, pp. 281-305, 2012.

[16] T. Akiba et al., “Optuna: A next-generation hyperparameter optimization framework”, *Proceedings of ACM SIGKDD*, 2019.

[17] F. Hutter, L. Kotthoff, and J. Vanschoren, *Automated Machine Learning*, Springer, 2019.

[18] U. K. Swain and S. Dash, “Optimized Hybrid GRU-HiBERT-MLP Model with Wavelet Features for Seizure State Classification from EEG Signals”, *Proceedings of the 2025 IEEE 4th International Conference for Advancement in Technology (ICONAT)*, 2025.

[19] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization”, *arXiv preprint arXiv:1412.6980*, 2014.

[20] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, MIT Press, 2018.

[21] C. J. C. H. Watkins and P. Dayan, “Q-learning”, *Machine Learning*, vol. 8, no. 3-4, pp. 279–292, 1992.

[22] V. Mnih et al., “Human-level control through deep reinforcement learning”, *Nature*, vol. 518, pp. 529–533, 2015.

[23] K. Arulkumaran et al., “Deep reinforcement learning: A brief survey”, *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 26–38, 2017.

[24] A. Shoeb, “Application of machine learning to epileptic seizure onset detection and treatment”, Ph.D. dissertation, MIT, 2009.

[25] S. Hochreiter and J. Schmidhuber, “Long short-term memory”, *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[26] J. Chung et al., “Empirical evaluation of gated recurrent neural networks on sequence modeling”, *arXiv preprint arXiv:1412.3555*, 2014.

[27] D. Bahdanau et al., “Neural machine translation by jointly learning to align and translate”, *arXiv preprint arXiv:1409.0473*, 2014.

[28] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning”, *Nature*, vol. 521, pp. 436-444, 2015.

[29] B. Zoph and Q. V. Le, “Neural architecture search with reinforcement learning”, *arXiv preprint arXiv:1611.01578*, 2016.

[30] H. Pham et al., “Efficient neural architecture search via parameter sharing”, *Proceedings of ICML*, 2018.