Journal of Information Systems Engineering and Management
2025, 10(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Ensemble DL Techniques for EEG Epilepsy Classification: Utilizing HRPCS for Enhanced Feature Selection and Improved Accuracy in Seizure Detection and Diagnosis

Jothi. B1* And Dr. D Jayaraj2
1Research scholar, Department of computer and information science, Annamalai University, Annamalainagar, Tamil Nadu, India.
2Assistant Professor, Department of Computer Science,
Government of Arts and Science College. Vadalur, Cuddalore District, India.
*Corresponding Author Email: sriramjairam15@gmail.com

ARTICLE INFO

Received: 10 Oct 2024
Revised: 09 Dec 2024
Accepted: 22 Dec 2024

ABSTRACT

Recurrent seizures are an indicator of epilepsy, a complicated neurological illness that requires prompt and precise diagnosis for successful treatment. Electroencephalogram (EEG) signals are critical in identifying seizure types and patterns, yet the analysis of these signals poses significant challenges. The variability in brain activity, the presence of noise and artifacts, and the complexity of differentiating between epileptic and non-epileptic seizures complicate accurate classification. Conventional methods often fall short, leading to misdiagnoses and inadequate treatment plans. To address these challenges, this study proposes a robust classification model for detecting and classifying EEG epilepsy data. The methodology begins with rigorous data preprocessing, which includes data cleaning and normalization to enhance the signal quality and ensure consistency across the dataset. Next, use dynamic and statistical feature extraction techniques to obtain key EEG signal parameters, which are necessary for accurately differentiating between seizure types. Furthermore, implement a Hybrid Red Piranha Cuckoo Search (HRPCS) algorithm for feature selection, allowing us to identify the most relevant features while reducing dimensionality. Finally, hybrid deep learning techniques HCRNN are utilized, incorporating Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN) for classification and prediction tasks. Blockchain technology is used to secure EEG data, ensuring integrity and patient privacy. After classification, the model also incorporates predictive analytics to forecast potential future seizures, enhancing patient management strategies.

Keywords: Epilepsy; EEG; Seizure; HRPCS; CNN; RNN.

1. INTRODUCTION

Epilepsy is a chronic neurological condition marked by recurrent and unpredictable seizures, affecting millions of individuals globally and significantly challenging patient management and quality of life. Abnormal electrical discharges in the brain induce seizures, which can range in severity from short-term loss of consciousness to severe convulsive bouts including involuntary movements [1,2]. These unpredictable events disrupt daily activities, pose risks of injury, and in severe cases, can lead to disability or death. While advancements in epilepsy treatment have improved seizure management to some extent, the inability to predict seizures accurately remains a critical gap [3]. This limitation exacerbates patient anxiety, restricts the effectiveness of preventive interventions, and impairs overall quality of life [4]. The early prediction of epileptic seizures holds tremendous promise in revolutionizing epilepsy care. By identifying pre-seizure patterns, proactive strategies such as timely medication adjustments, lifestyle changes, or caregiver alerts could be implemented to mitigate risks [5,6]. In addition to lessening the number and intensity of seizures, this strategy may provide useful information for improved management to those who have epilepsy and those who care for them. In addition, seizure prediction offers the potential to improve autonomy, safety, and confidence for patients, allowing them to regain control over their lives [7].

Over the years, research on epileptic seizure prediction has expanded across multiple domains, including neurophysiology, signal processing, and machine learning. The advent of advanced neuroimaging methods and wearable EEG devices has facilitated the collection of large-scale brain activity data, providing unprecedented opportunities to analyze seizure dynamics [8]. These data have paved the way for the development of predictive models that use machine learning algorithms and computational techniques to identify pre-seizure patterns with higher precision [9]. Despite notable progress, developing clinically reliable and robust prediction systems remains a formidable challenge. The complexity and variability of epilepsy contribute significantly to these challenges. Seizures vary greatly in type and frequency, even within the same patient, and EEG signals exhibit high variability due to patient-specific factors and environmental influences [10]. Additionally, the dynamic and intricate nature of brain activity further complicates the prediction process. This research proposes a novel approach using advanced signal processing and machine learning algorithms to extract meaningful features from EEG data, enhancing seizure prediction accuracy and reliability [11]. This research specifically aims to create a data-driven model that incorporates advanced methods for extraction of features, reducing dimensionality, and classifications. The goal is to identify pre-seizure patterns in EEG signals effectively while ensuring scalability and adaptability across diverse patient profiles [12].

The proposed model employs advanced feature extraction techniques to capture relevant information from EEG signals, focusing on key characteristics that distinguish pre-seizure states from normal activity [13]. Dimensionality reduction methods are used to eliminate noise and retain critical features, enhancing computational efficiency. Finally, a robust classification framework is applied to forecast seizures with precision, addressing the inherent variability of EEG data through personalized modeling strategies. By fostering interdisciplinary collaboration and utilizing state-of-the-art technologies, this research aims to advance the field of epileptic seizure prediction. The ultimate objective is to provide individuals living with epilepsy and their caregivers with reliable tools for managing the condition proactively, improving patient outcomes, and enhancing their quality of life [14,15]. Through innovative methodologies, this study strives to bridge the gap in epilepsy care, offering hope for a safer and more independent future for those affected.

The contributions of this paper are manifested below,

- In order to increase the accuracy of seizure diagnosis, this study addresses the difficulties in processing complex and noisy EEG signals by introducing a strong classification model for identifying and categorizing EEG epilepsy data.
- This work incorporates HRPCS algorithm for effective feature selection, enabling the identification of relevant EEG signal parameters while reducing dimensionality for enhanced computational efficiency.
- This work employs hybrid deep learning techniques, HCRNN including CNN and RNN, to achieve accurate classification and prediction of epileptic seizures.
- By integrating blockchain technology, our study protects patient privacy and offers a solid basis for data-driven healthcare solutions by guaranteeing the confidentiality and integrity of EEG data.

This paper's remaining sections are structured as follows. A selection of pertinent works and a problem statement are provided in Part 2. In Section 3, the proposed procedure is presented and explained. Section 5 presents the conclusion, whereas Part 4 presents the findings and discussion.

2. LITERATURE REVIEW

Emara et al. [16] published three frameworks for processing EEG signals in 2022. The first employed Scale-Invariant Feature Transform (SIFT) for automated seizure identification. In the second, epileptic seizures were predicted using an artificial neural network (ANN) and the Fast Fourier Transform (FFT). The third provided an automated, patient-specific framework for seizure prediction and channel selection, and it employed FFT for feature extraction. Gao et al. [17] introduced a temporal-spatial multi-scale convolutional neural network (CNN) with dilated convolutions for seizure prediction. This framework systematically extracted multi-scale properties in temporal and spatial phases using dilated convolutions to improve receptive fields and prediction accuracy.

Tamanna et al. [18] in 2021 aimed to advance seizure prediction from EEG signals using time-frequency feature extraction and classification techniques. Discrete Wavelet Transform (DWT) was applied to extract features, followed by the use of Support Vector Machine (SVM) and post-processing techniques to predict seizures with high accuracy.

A technique for eliminating motion-related distortions from EEG recordings in epilepsy situations was created in 2020 by Islam et al. [19]. In order to guarantee clean data for additional analysis, this method entailed employing a wearable headset to collect EEG signals and then applying Independent Component Analysis (ICA) to remove artifacts.

In 2021, Zhao et al. [20] developed to enhance quality of life for drug-refractory epilepsy patients, seizure prediction methods strive for energy efficiency and hardware friendliness. Through neural architecture search, a compact model was obtained and evaluated across multiple datasets. Model compression techniques further reduced its size, facilitating low-power operation for wearable and implantable devices.

Five deep learning models for predicting epileptic episodes from intracranial electroencephalogram (iEEG) datasets were proposed by Ouichka et al. [21] in 2022. A Convolutional Neural Network (CNN) was one of these models, along with many fusion techniques such combining two CNNs (2-CNN), three CNNs (3-CNN), and four CNNs (4-CNN). The efficiency of these models in seizure prediction tasks was further demonstrated by the use of transfer learning with ResNet50.

Mahmoodian et al. [22] used the cross-bispectral technique to extract nonlinear multivariate factors in order to study the prediction of epileptic episodes in 2020. The study distinguished between pre-ictal and interictal stages using ten statistical factors. After that, a Support Vector Machine (SVM) classifier was given these characteristics. Ra et al. [23] developed an effective epileptic seizure prediction system in 2021. The method focused on EEG feature extraction and classification by selecting EEG channels using an optimization strategy based on permutation entropy (PE). The classifier used a Genetic Algorithm (GA) in combination with K-Nearest Neighbors (KNN) to assess the CHB-MIT Scalp EEG Database.

Singh and Malhotra [24], in 2022, proposed a two-layer Long Short-Term Memory (LSTM) network model based on spectral features for predicting epileptic seizures. The model utilized spectral power and mean spectrum amplitude features from multiple EEG frequency bands, evaluated over 5 to 50-second segments of EEG data. The two-layer LSTM model demonstrated high accuracy when processing 30-second EEG segments.

In 2022, Xu et al. [25] presented a technique for early seizure prediction that combined Gradient Boosting Decision Trees (GBDT) with nonlinear characteristics of EEG data. In order to extract nonlinear characteristics including approximation entropy, sample entropy, and wavelet entropy, the EEG signals were first denoised using complementary ensemble empirical mode decomposition (CEEMD) and wavelet threshold denoising. A random forest-initialized GBDT classifier successfully differentiated between seizure onset and non-seizure periods.

2.1. Problem Statement

Improving patient treatment and quality of life requires the early and precise prediction of epileptic episodes. People who have epilepsy, a common neurological illness marked by frequent and unexpected seizures, have many difficulties, as do those who care for them. Current seizure prediction methods often lack the precision and reliability needed for timely intervention, increasing risks of injury and complications for patients. Additionally, existing approaches frequently fail to capture the intricate dynamics and variability of EEG signals associated with seizure onset. Addressing these limitations requires the development of advanced predictive models capable of analyzing EEG data in real-time, extracting critical features, and accurately forecasting seizures. By integrating sophisticated signal processing techniques with machine learning algorithms, these models can achieve high sensitivity and specificity while reducing false alarms. Such advancements promise to revolutionize epilepsy management, enabling proactive, personalized care and improving outcomes for individuals living with epilepsy.

3. PROPOSED METHODOLOGY

Epileptic seizure prediction involves utilizing signal processing and DL methods to forecast seizures in epilepsy patients, facilitating timely intervention and enhanced patient care. However, challenges such as diverse seizure patterns among individuals, scarcity of dependable long-term EEG data, the risk of false alarms and the necessity for real-time prediction pose obstacles. Overcoming these challenges requires robust algorithms, comprehensive data collection, and validation procedures to ensure accurate and reliable prediction models.

Addressing these hurdles can significantly improve seizure management and enhance the quality of life for epilepsy patients.

3.1. Data Pre-Processing

In this work, the collected data is pre-processed and carried out using data cleaning and normalization.

3.1.1. Data Cleaning

Data-cleaning corrects errors and inconsistencies, addressing missing values through imputation or deletion. Second, removing duplicates is necessary because duplicate records can inflate the dataset and skew the analysis. Detecting and removing these duplicates ensures that each customer feedback is unique and only contributes once to the analysis. Third, correcting inconsistencies involves standardizing formats, such as dates or categorical values. A date might be entered in different formats by different users, and standardizing it to a common format is essential. Finally, outlier detection and correction are important as outliers are abnormal data points that can skew the results. While some outliers may be valid data, others might result from errors. Outlier detection methods are used to identify these, and corrective measures are taken accordingly.

3.1.2. Normalization

In order to standardize the magnitude of characteristics within a dataset, normalization is an essential step in the preparation of data. Regardless of their initial sizes or units, it guarantees that every variable makes an equal contribution to the study. Converting feature numerical data to a standard scale—typically between 0 and 1 or -1 and 1—is the process. Figure 1 shows the general suggested architecture.

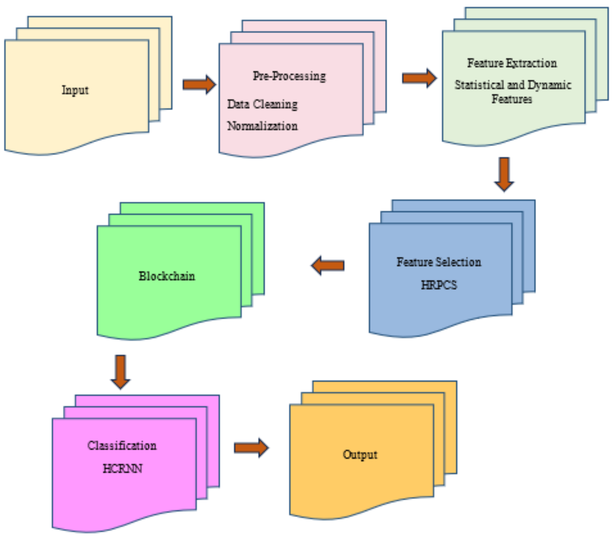


Figure 1: Overall Prosed Methodology

By converting values to a predetermined range, removing the minimum, and dividing by the range, min-max scaling normalizes features. By subtracting the mean and dividing by the standard deviation, Z-score normalization (standardization) centres the data around zero with a standard deviation of 1. Both strategies enhance algorithm performance and convergence by preventing the dominance of features with bigger sizes.

3.2. Feature Extraction

In order to capture important signal properties, statistical and dynamic features are calculated using pre-processed EEG data in the feature extraction phase of this study. The selected features include statistical measures such as Mean, Variance, Skewness, and enhanced Kurtosis, which provide insights into signal distribution and variability. Dynamic features like Lyapunov Exponent, Hjorth Activity, Mobility, and Complexity quantify signal stability and complexity. Additional metrics such as Approximate Entropy, PSD Mean, PSD Variance, Spectral Centroid, and Zero-Crossing Interval Mean reflect signal periodicity, frequency domain properties, and temporal transitions. These features collectively ensure a comprehensive representation of EEG data for accurate analysis and classification.

3.2.1 Statistical Features

- ***Mean***

The mean, or average, is a statistical measure representing the central tendency of a dataset. It's calculated by summing all values and dividing by the total count. While useful for summarizing data, outliers can skew results. Therefore, it's essential to consider other measures like median and mode for a comprehensive analysis. The formula for calculating the mean ($\bar{x}$) of a dataset is given using Eq. (1).

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{1}$$

Where, x_i represents each individual value in the dataset, n is the total number of values in the dataset.

- ***Median***

When a dataset is sorted in either ascending or descending order, the median, a statistical metric, indicates the midway value. It is resilient to outliers since it is unaffected by extreme values. The median is the midway value when the number of observations is odd. The median is the mean of the two middle values if the number of observations is even. Whether the dataset's number of observations (n) is odd or even determines the formula used to get the median:

If n is odd, the median is the value at position $(n + 1)/2$ when the data is arranged in ascending or descending order. If n is even, the median is the average of the values at positions $(n/2) + 1$ when the data is arranged in ascending or descending order.

- ***Skewness***

Skewness is a metric used to quantify how asymmetrically values are distributed within a collection. The following is the formula to determine skewness:

$$Skewness = \frac{3 \times (mean - median)}{standard\ deviation} \tag{2}$$

Eq. (2) involves multiplying the difference between the mean and the median by three and then dividing the result by the standard deviation.

- ***Enhanced Kurtosis***

The fourth-order moment is divided by the population's standard deviation raised to the fourth power to determine kurtosis, a measure of a distribution's tailedness. It reflects how often outliers occur, with excess kurtosis indicating tailedness relative to a normal distribution. The formula for calculating kurtosis (K) is given as per Eq. (3).

$$k = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{\sigma^4} \cdot H(\alpha) \tag{3}$$

Where, n is the number of data points, x_i represents each individual data point, $\bar{x}$ is the sample mean. σ is the standard deviation, Shannon Entropy $H(\alpha)$ assesses the randomness or information content of the dataset. Adding Shannon entropy to kurtosis provides a more subtle understanding of your dataset, especially for tasks requiring detailed pattern recognition or feature differentiation.

3.2.2. Dynamic Features

- ***Lyapunov Exponent***

The Lyapunov Exponent is a measure of the sensitivity of a system to initial conditions, commonly used to quantify chaos and signal stability. In the context of EEG analysis, it captures the rate at which trajectories in the EEG signal space diverge or converge over time. A positive LE suggests chaotic behavior, typical during seizure episodes, while a negative or zero LE implies a stable or predictable state.

- ***Hjorth Activity***

Hjorth Activity is a statistical parameter that evaluates the signal's power or variance, providing insight into its overall energy. High activity levels in EEG signals often indicate abnormal brain activity, such as during seizures. It is particularly useful for identifying regions of heightened neural activity, which could signify a seizure onset.

- ***Hjorth Mobility***

Hjorth Mobility measures the signal's frequency characteristics by computing the ratio of the standard deviation of the first derivative to the original signal. This feature quantifies the rate of change or oscillation in the signal, making it a valuable tool for detecting abrupt changes in brain activity. Increased mobility often corresponds to heightened neural activity or seizure episodes.

- ***Hjorth Complexity***

Hjorth Complexity evaluates the regularity of the signal, capturing its intricacy by analyzing how the signal's mobility changes over time. It provides insights into the structure and organization of the EEG signal, with higher values indicating more complex and unpredictable activity. This feature is crucial for differentiating normal brain activity from pathological conditions, as epileptic seizures often result in a marked increase in signal complexity.

- ***ApproximateEntropy (ApEn)***

One statistical metric for assessing the consistency and prediction of time-series data, like EEG signals, is approximate entropy. It evaluates the likelihood of similar patterns recurring over time. Lower ApEn values suggest more regular signals, often associated with normal brain activity, while higher values indicate irregularity, which is characteristic of seizure episodes.

- ***Power Spectral Density (PSD) Mean***

PSD Mean represents the average power of a signal across its frequency spectrum, calculated from its Power Spectral Density. It provides insights into the distribution of signal energy among different frequency bands. In EEG analysis, PSD Mean can reveal dominant brain wave activity, such as alpha, beta, or gamma waves, and help identify abnormalities like excessive power in specific bands during seizures.

- ***Power Spectral Density (PSD) Variance***

PSD Variance quantifies the variability in power distribution across the frequency spectrum of EEG signals. It helps identify shifts or inconsistencies in brain wave activity, which can be critical for detecting transient anomalies such as seizures. Increased PSD Variance often corresponds to irregular neural activity, making it a valuable feature for seizure classification.

- ***Spectral Centroid***

The Spectral Centroid indicates the "center of mass" of the frequency spectrum, providing a measure of where the majority of the signal energy is concentrated. It is often associated with the perceptual "brightness" of a signal. In EEG analysis, a shift in the spectral centroid can signify abnormal brain activity, such as a transition from normal to preictal or ictal states.

- ***Zero-Crossing Interval Mean (ZCIM)***

Zero-Crossing Interval Mean measures the average duration between successive zero-crossings in the signal, reflecting its oscillatory behavior and periodicity. EEG signals with frequent zero-crossings typically correspond to higher frequency activity, while longer intervals indicate slower waveforms. Variations in ZCIM can highlight transitions in brain activity, such as the onset of a seizure, making it an effective metric for time-domain analysis.

- ***Frequency-Domain Properties***

Frequency-domain properties focus on the distribution of signal energy across various frequency bands. Metrics like PSD Mean, PSD Variance, and Spectral Centroid are used to analyze how EEG signals behave in the frequency domain. These properties help identify dominant neural rhythms, such as delta or gamma waves, and detect anomalies like abnormal power spikes, providing crucial insights into seizure activity and overall brain health.

- ***Temporal Transitions***

Temporal transitions refer to the changes in signal patterns over time, capturing the dynamic nature of neural activity. Metrics such as Approximate Entropy and ZCIM are essential for understanding these transitions, as they reflect how brain states evolve. For instance, the shift from interictal to preictal or ictal states is characterized by temporal changes in EEG signals. Accurately capturing these transitions is critical for seizure prediction and monitoring.

3.3. Feature Selection

In this study, the data extracted undergo feature selection using HRPCS as input.

3.3.1. HRPCS

The HRPCS algorithm integrates the strengths of the Cuckoo Search (CS) and Red Piranha Optimization (RPO) algorithms. CS is inspired by the brood parasitic behavior of cuckoo birds and employs Lévy flights for efficient exploration of the search space. Lévy flights allow the algorithm to take both short steps for local search and long jumps for global exploration, ensuring a balance between exploration and exploitation. CS excels at finding promising regions in the search space with minimal computational effort, making it effective for global optimization tasks. RPO draws inspiration from the cooperative hunting behavior of red piranhas. This swarm-based algorithm focuses on intense local search and dynamic adaptation, refining solutions in regions with higher fitness values. RPO's ability to adjust search intensity and avoid premature convergence makes it effective for local optimization and exploitation. By combining these algorithms, HRPCS leverages the global search efficiency of CS with the local refinement capabilities of RPO. This hybrid approach ensures robust feature selection, allowing the model to focus on the most impactful features while minimizing redundant or irrelevant data.

3.3.1.1. Mathematical Modelling

3.3.1.1.1. Initialization

Functioning as a population-based metaheuristic algorithm, the RPO technique employs red pandas as symbolic representations of individual members. Mathematically, each red panda is depicted as a vector, forming a matrix where rows represent potential solutions and columns hold values for associated problem variables. Initially, red panda coordinates within the search space are randomly initialized using Eq. (4) and Eq. (5). This approach facilitates exploration and exploitation of the solution space to find optimal solutions.

$$Y = \begin{bmatrix} Y_1 \\ . \\ . \\ . \\ Y_i \\ . \\ . \\ Y_M \end{bmatrix}_{M\times n} = \begin{bmatrix} Y_{1,1} \dots Y_{1,j} \dots Y_{1,n} \\ . \\ . \\ . \\ Y_{i,1} \dots Y_{i,j} \dots Y_{i,n} \\ . \\ . \\ Y_{M,1} \dots Y_{M,j} \dots Y_{M,n} \end{bmatrix}_{M\times n} \tag{4}$$

$$y_{i,j} = lob_j + r_{i,j} \cdot \left(upb_j - lob_j\right) \tag{5}$$

The population matrix holding the red panda locations is represented by Y in the RPO technique, where Y_i stands for the ith red panda (possible solution) and $Y_{i,j}$ for its jth dimension (problem variable), random integers $r_{i,j}$ inside the interval [0,1] are used.

The positions of each red panda act as potential solutions, making it possible to assess the objective function associated with each one. A matrix of the form provided by Eq. (6) can be used to represent the final set of evaluated objective function values.

$$f = \begin{bmatrix} f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_i \\ \cdot \\ \cdot \\ f_M \end{bmatrix}_{M\times 1} = \begin{bmatrix} f(Y_1) \\ \cdot \\ \cdot \\ \cdot \\ f(Y_i) \\ \cdot \\ \cdot \\ f(Y_M) \end{bmatrix}_{M\times 1} \tag{6}$$

The value obtained by the ith red panda is indicated by f_i, and f represents the vector of values of the objective function. These values of the objective function are essential for evaluating the caliber of potential solutions. The greatest and lowest values of the objective function are used to identify the best and worst potential solutions, respectively. These potential solutions are modified appropriately during every iteration. Iterative upgrades to potential solutions for the best possible problem-solving are part of the RPO's exploration and exploitation phases.

3.3.1.1.2. Phase 1: Exploration Strategy - Foraging

During the initial phase of RPO, red pandas' positions mimic their foraging behavior in the wild. Leveraging their adeptness in detecting food sources, each red panda evaluates the locations of others with better objective function values as potential feeding grounds. These prospective food positions are identified through comparisons of objective function values, with each red panda randomly selecting one position using Eq. (7). This process simulates the exploration for optimal solutions in the search space.

$$pfs_i = \{Y_k | k \in \{1,2,..,M\} and\ f_k < f_i\} \cup \{Y_{best}\} \tag{7}$$

Based on a comparison with the location of the best candidate solution Y_{best}, the suggested food sources for each red panda $pfs_i$are identified. Approaching these sources causes large positional shifts that improve ability of algorithm to globally search and explore. By determining new locations in relation to the food source (best candidate solution), red pandas' foraging behavior can be replicated. The proposed Eq. (8) to Eq. (10) are used to update the red panda's location.

$$Y_i^{p1}: y_{i,j}^{p1} = y_{i,j} + r.\left(sfs_{i,j} - Is.y_{i,j}\right) \cdot S \tag{8}$$

$$S = \frac{u}{|v|^{\frac{1}{\beta}}} \tag{9}$$

Integrating the step size (S) from CS into RPO can enhance its performance by improving exploration and exploitation balance. The step size in CS controls the magnitude of solution perturbations, promoting better exploration of the solution space. By adding this concept to RPO, it can prevent premature convergence by making larger moves in the early stages and refining the search later. This dynamic step adjustment helps RPO avoid local optima and improves convergence. The step size can be implemented using a scaling factor during position updates, inspired by Lévy flights in CS.

$$Y_i = \begin{cases} Y_i^{p1}, f_i^{p1} < f_i \\ Y_i, else \end{cases} \tag{10}$$

The new location of the ith red panda as ascertained from the RPO's first phase is represented by Y_i^{p1}. Objective function is denoted by f_i^{p1}, and its position in the jth dimension is indicated by $y_{i,j}^{p1}$. For the ith red panda, sfs_i denotes the preferred food source, and $sfs_{i,j}$ denotes its location in the jth dimension. Is is a randomly chosen number from the set {1, 2}, and the variable r is a random value between 0 and 1.

3.3.1.1.3. Phase 2: Proficiency in ascending and perching on trees (exploitation)

During the second phase of RPO, red pandas' tree-climbing behavior guides their positioning. These animals typically rest on trees for extended periods and move to nearby trees for food after ground foraging. In promising regions, this behavior results in minor positioning tweaks that improve the RPO algorithm's exploitation and local search capabilities. Mathematically, this behavior entails computing new positions for each red panda and updating previous positions if the objective function improves, as described in Eq. (11) and Eq. (12). This process mimics the iterative refinement of solutions as red pandas navigate the search space in pursuit of optimal solutions.

$$Y_{i,j}^{p2} = y_{i,j} + \frac{lob_j + r_{i,j} \cdot (upb_j - lob_j)}{t} \cdot S, i = 1,2, \dots, M, j = 1,2, \dots, n, t = 1,2, \dots, T \quad (11)$$

$$Y_i = \begin{cases} Y_i^{p2}, f_i^{p2} < f_i \\ Y_i, else \end{cases} \quad (12)$$

The ith red panda's modified position, obtained from the second phase of RPO, is represented by Y_i^{p2}. Objective function is shown by f_i^{p2}, and its position in the jth dimension is indicated by $Y_{i,j}^{p2}$. A random number between 0 and 1 represents the variable r. The symbol t denotes the algorithm's iteration counter, whereas T stands for the maximum iterations. Algorithm 1 illustrates how HRPCS integrates the CS into the RPO, adjusting red panda positions during both exploration and exploitation phases.

Algorithm 1: HRPCS
Initialize population of red pandas Y with random positions in the search space
Initialize the step size S using Lévy flight
Evaluate the objective function for each red panda
For each iteration t = 1 to T do:
Phase 1: Exploration Strategy (Foraging)
For each red panda
Find the best solution and potential food sources using Eq. (7)
Update position of red panda based on food sources
Select new position
Update position using Eq. (8)
Calculate the step size (S) using the CS step size using Eq. (9)
If the new position improves the objective function
Update the position to the new one Eq. (10)
Phase 2: Exploitation Strategy (Tree Climbing)
For each red panda
For each dimension
Update position using Eq. (11)
Evaluate the new objective function
Update the position to the new one using Eq. (12)
Update the best solution found so far
Update if a better solution is found in this iteration
Return (best solution found)

3.4. Data Security-Blockchain

To ensure the security, integrity, and effective utilization of sensitive EEG data, blockchain technology is seamlessly integrated into the proposed model. This combination secures data storage, manages access rights, and streamlines the reuse of key features for classification tasks. A key feature of blockchain integration is the generation of cryptographic keys for secure data access. When EEG data is uploaded to the blockchain, public-private key pairs are generated using cryptographic algorithms. The public key is used to encrypt data, ensuring only authorized users with the corresponding private key can decrypt and access it. For healthcare providers and researchers, smart contracts validate access requests, allowing only authenticated entities to retrieve specific EEG data or derived features. This mechanism safeguards against unauthorized access, ensuring compliance with data protection regulations.

The EEG classification process involves selecting critical features from pre-processed data to optimize classification accuracy. Features such as statistical and dynamic features are extracted during the feature extraction stage. These features are prioritized based on their relevance using the HRPCS algorithm, which reduces dimensionality by identifying the most discriminative attributes. Selected features are then securely stored on the blockchain. Each feature vector is hashed using a cryptographic hash function and appended to the blockchain as a transaction. The hash ensures data integrity, while the blockchain's distributed nature provides fault tolerance and immutability. Alongside the features, metadata such as patient IDs (anonymized) and timestamps are stored for future reference.

Storing features on the blockchain enables their secure reuse for subsequent classification or research tasks. When the EEG classification model is retrained or updated, the blockchain is queried to retrieve the stored features. Smart contracts streamline this process, ensuring that only authorized entities can access the data. Additionally, the retrieved features can be combined with new EEG data for longitudinal studies, predictive analysis, or personalized treatment plans. By leveraging blockchain, this approach not only enhances the reliability of data storage and access but also empowers healthcare providers with a transparent and tamper-proof system. This integration ensures robust security, data privacy, and efficient resource utilization, making it a cornerstone of proactive and personalized epilepsy management solutions.

3.5. Classification

In this study, the selected data is classified using HCRNN for enhanced accuracy.

3.5.1. HCRNN

CNN are a subset of deep learning models that are skilled in identifying and deriving meaningful patterns from unprocessed input, such as images. Convolutional, pooling, activation, completely connected, and an output layer are among them. While pooling layers down sample features, convolutional layers use filters to extract features. Fully connected layers produce predictions, and activation functions incorporate non-linearities. CNN are trained to minimize a selected loss function by optimizing their parameters using methods like gradient descent and backpropagation. CNN are able to adjust and enhance their recognition and classification skills of objects in images through this iterative process. Three different sorts of layers commonly make up ANN architecture: input, hidden, and output layers. There are neurons, or nodes, in every layer, and information is carried by the connections between these neurons. A neural network built for sequential data is called an RNN.

RNN, in contrast to feedforward neural networks, have connections that form directed cycles, which allow them to handle sequential input and store hidden state information. Because the order of the input pieces matters, they are especially useful for jobs involving time series data or natural language processing. Recurrent units, such Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), are used by RNNs to update hidden states over time and manage long-term dependencies. They use methods such as Backpropagation through Time (BPTT) to update parameters according to the data's sequential structure. Furthermore, RNNs are used in security frameworks to continually monitor incoming data and identify anomalies or departures from predicted patterns that could point to system faults or security risks. The hidden state hs_t at time t in an RNN is computed using Eq. (13).

$$hs_t = \sigma(w_{hsi} i_t + w_{hshs} hs_{t-1} + bi_{hs}) \tag{13}$$

The output $o_t$at time t is computed based on the hidden state and expressed as per Eq. (14).

$$o_t = softmax(w_{ohs} hs_t + bi_o) \tag{14}$$

The RNN processes sequences by iterating through time steps, updating the hidden state at each step based on the current input and the previous hidden state. Algorithm 2 illustrates how HCRNN integrates the CNN into the RNN.

Algorithm 2: HCRNN for Classification

Input: EEG data, Labels, Epochs, Learning rate, Batch size
Output: Trained HCRNN model, Predictions
1. Initialize Model
  a. Define CNN layers (convolution, pooling, activation)
  b. Define RNN layers
  c. Add output layer with Softmax for classification
2. Preprocess Data
  a. Normalize input data
  b. Split data into training and testing
  c. Create batches of size
3. Train Model
  FOR epoch = 1 to E
    FOR each batch
      a. Forward Pass
        i. Extract features using CNN layers
        ii. Process features sequentially with RNN layers
        iii. Generate output using Softmax
      b. Compute loss between predictions and labels
      c. Backward Pass
        i. Update model weights using gradient descent
    END FOR
  END FOR
4. Test Model
  a. Pass through the trained model
  b. Predict labels for test data
5. Evaluate Model
6. Output
  a. Trained model and predictions

After classification, the model leverages predictive analytics to forecast potential future seizures, providing valuable insights for proactive patient management. By analyzing patterns in preictal and interictal EEG data, the system identifies trends and markers indicative of imminent seizures. This forecasting capability allows healthcare providers to implement timely interventions, such as medication adjustments or alert systems, to mitigate seizure risks. The integration of predictive analytics also supports personalized treatment plans, enhancing the quality of care for epilepsy patients. This forward-looking approach not only improves patient safety but also empowers individuals to better manage their condition with early warnings and actionable insights.

4. RESULT AND DISCUSSION

4.1. Experimental Setup

The experiments were conducted using Python. A high-performance computing environment was utilized to handle the computational demands of training the HCRNN model on a large-scale dataset. The model was trained for epochs and the learning rate was dynamically adjusted based on validation loss.

4.2. Dataset Collection

The American Epilepsy Society Seizure Prediction Challenge dataset [26] is designed to advance seizure prediction methods using intracranial EEG (iEEG) data. The dataset contains iEEG recordings for human and canine subjects, categorized into training and testing sets. Training data includes ten-minute clips labelled as Preictal (pre-seizure) or Interictal (non-seizure). Preictal data represents one hour prior to a seizure onset with a five-minute seizure horizon, ensuring sufficient warning time for intervention. Interictal segments are collected from non-seizure periods, far from any seizure activity, minimizing contamination. Each clip is stored in .mat format with fields like EEG signal matrix (electrode × time), duration, sampling frequency, electrode names, and sequence index. This comprehensive structure aids in feature extraction and modeling. The dataset also includes random testing data for evaluation. With over 113 GB of data, this resource supports the development of accurate seizure prediction models using advanced machine learning techniques.

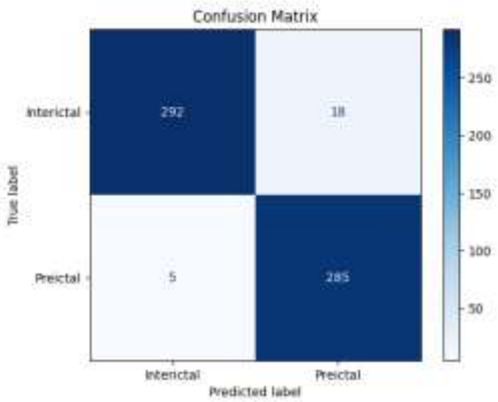


Figure 2: Confusion Matrix (Prediction)

The confusion matrix in fig. 2 represents the predicted results of the model's classification of EEG data into preictal (seizure-prone) and interictal (non-seizure) states. It provides a detailed comparison between the actual and predicted outcomes. Key observations include:

- True Positives (TP): High values indicate the model effectively identifies seizure-prone periods.
- True Negatives (TN): A strong ability to classify non-seizure periods accurately.
- False Positives (FP): Minimal false alarms, showcasing the model's robustness.
- False Negatives (FN): A low rate, demonstrating reliable sensitivity in detecting preictal states.

The matrix reflects the model's excellent overall accuracy, supported by high precision and recall values.

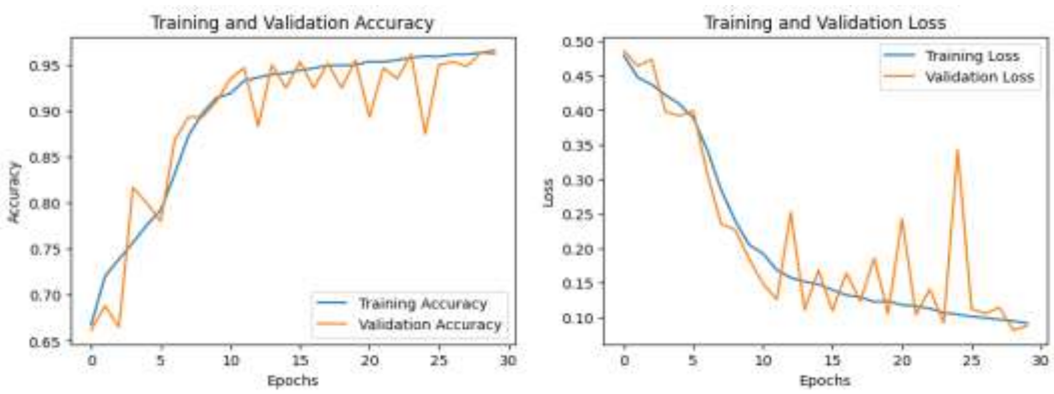
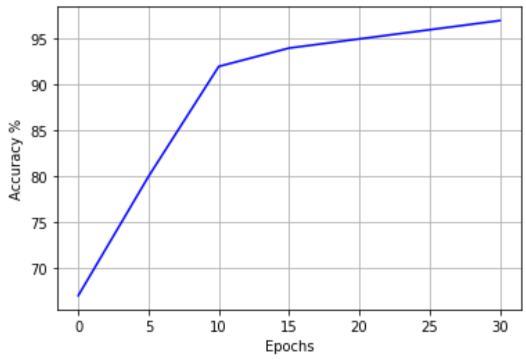
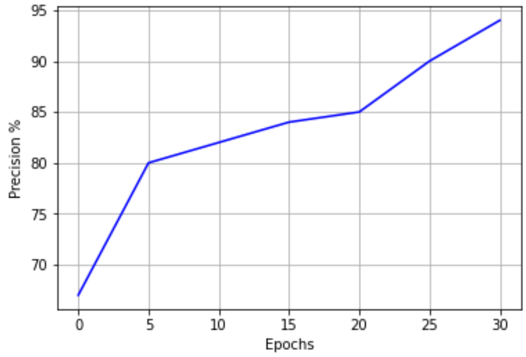


Figure 3: Accuracy vs Loss

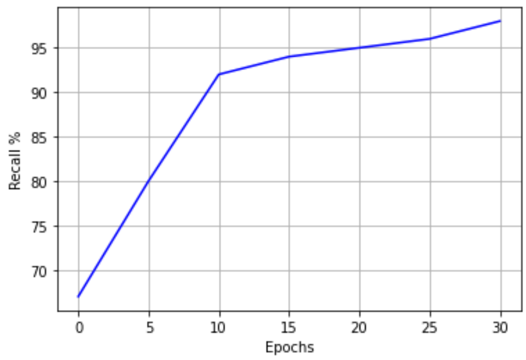
The model effectively learns to distinguish seizure-prone from non-seizure situations without overfitting, as evidenced by the training accuracy's steady improvement with epochs. The model's generalizability to unknown data is confirmed by the validation accuracy, which closely resembles the training curve. Figure 3 illustrates how the training and validation loss curves steadily decline across epochs until stabilizing at low values. This pattern indicates that the optimization algorithm is learning and convergent. These indicators' convergence demonstrates the model's dependability and effectiveness. As seen in fig. 4, the suggested model performs exceptionally well, exhibiting excellent accuracy, precision, recall, and F1-score.



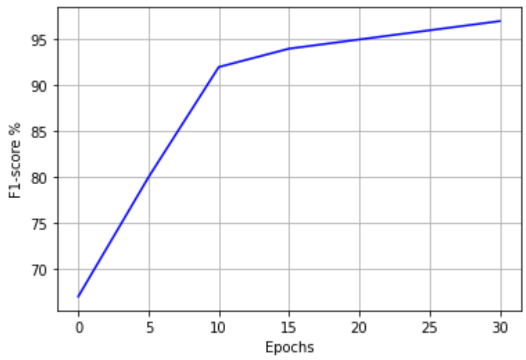
(a)



(b)



(c)



(d)

Figure 4: Performance Metrics Varying Epochs

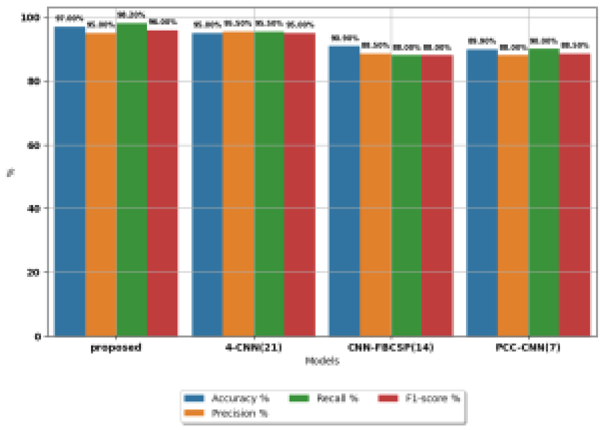


Figure 5: Comparison of Proposed and Existing Models

A comparison of the suggested and current models is shown in Fig. 5, which emphasizes gains in F1 scores, recall, accuracy, and precision. The suggested model performs better than conventional methods, proving its usefulness in tasks involving seizure categorization and prediction.

5. CONCLUSION

This study developed a reliable classification technique for identifying and categorizing data related to EEG epilepsy. To maintain consistency throughout the dataset and improve signal quality, the technique started with thorough data preparation, which included data cleaning and normalization. To get important EEG signal parameters like mean, median, skewness, and improved kurtosis that were necessary for precisely distinguishing between seizure types, statistical feature extraction approaches were used. In order to choose the most pertinent features while lowering dimensionality, the HRPCS algorithm was used. For tasks involving classification and prediction, hybrid deep learning methods were applied, such as an HCRNN that combines CNN and RNN. Blockchain technology was utilized to ensure the security and integrity of the EEG data in order to safeguard patient privacy. By employing predictive analytics to foresee potential future seizures following categorization, the method enhanced patient care practices.

DECLARATION OF INTERESTS:

Funding: On Behalf of all authors the corresponding author states that they did not receive any funds for this project.

Conflicts Of Interest: The authors declare that we have no conflict of interest.

Competing Interests: The authors declare that we have no competing interest.

Data Availability Statement: All the data is collected from the simulation reports of the software and tools used by the authors. Authors are working on implementing the same using real world data with appropriate permissions.

Ethics Approval: No ethics approval is required.

Consent To Participate: Not Applicable

Consent For Publication: Not Applicable

Human And Animal Ethics: Not Applicable.

Code Availability: Not Applicable.

Author's Contributions:

Author 1: She Performed the Analysis the overall concept, writing and editing.

Author 2: He participated in the methodology, Conceptualization, Data collection and writing the study.

References

[1] Savadkoohi, M., Oladunni, T. and Thompson, L., 2020. A machine learning approach to epileptic seizure prediction using Electroencephalogram (EEG) Signal. Biocybernetics and Biomedical Engineering, 40(3), pp.1328-1341.

[2] Usman, S.M., Khalid, S. and Bashir, Z., 2021. Epileptic seizure prediction using scalp electroencephalogram signals. Biocybernetics and Biomedical Engineering, 41(1), pp.211-220.

[3] Das, K., Daschakladar, D., Roy, P.P., Chatterjee, A. and Saha, S.P., 2020. Epileptic seizure prediction by the detection of seizure waveform from the pre-ictal phase of EEG signal. Biomedical Signal Processing and Control, 57, p.101720.

[4] Dissanayake, T., Fernando, T., Denman, S., Sridharan, S. and Fookes, C., 2021. Deep learning for patient-independent epileptic seizure prediction using scalp EEG signals. IEEE Sensors Journal, 21(7), pp.9377-9388.

[5] Rasheed, K., Qadir, J., O'Brien, T.J., Kuhlmann, L. and Razi, A., 2021. A generative model to synthesize EEG data for epileptic seizure prediction. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 29, pp.2322-2332.

[6] Aslam, M.H., Usman, S.M., Khalid, S., Anwar, A., Alroobaea, R., Hussain, S., Almotiri, J., Ullah, S.S. and Yasin, A., 2022. Classification of EEG signals for prediction of epileptic seizures. Applied Sciences, 12(14), p.7251.

[7] Zhang, S., Chen, D., Ranjan, R., Ke, H., Tang, Y. and Zomaya, A.Y., 2021. A lightweight solution to epileptic seizure prediction based on EEG synchronization measurement. The Journal of Supercomputing, 77, pp.3914-3932.

[8] Dissanayake, T., Fernando, T., Denman, S., Sridharan, S. and Fookes, C., 2021. Geometric deep learning for subject independent epileptic seizure prediction using scalp eeg signals. IEEE Journal of Biomedical and Health Informatics, 26(2), pp.527-538.

[9] Wang, G., Wang, D., Du, C., Li, K., Zhang, J., Liu, Z., Tao, Y., Wang, M., Cao, Z. and Yan, X., 2020. Seizure prediction using directed transfer function and convolution neural network on intracranial EEG. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 28(12), pp.2711-2720.

[10] Affes, A., Mdhaffar, A., Triki, C., Jmaiel, M. and Freisleben, B., 2022. Personalized attention-based EEG channel selection for epileptic seizure prediction. Expert Systems with Applications, 206, p.117733.

[11] Prathaban, B.P. and Balasubramanian, R., 2021. Dynamic learning framework for epileptic seizure prediction using sparsity based EEG Reconstruction with Optimized CNN classifier. Expert Systems with Applications, 170, p.114533.

[12] Alshebeili, S.A., Sedik, A., Abd El-Rahiem, B., Alotaiby, T.N., El Banby, G.M., El-Khobby, H.A., Ali, M.A., Khalaf, A.A. and Abd El-Samie, F.E., 2020. Inspection of EEG signals for efficient seizure prediction. Applied Acoustics, 166, p.107327.

[13] Zambrana-Vinaroz, D., Vicente-Samper, J.M., Manrique-Cordoba, J. and Sabater-Navarro, J.M., 2022. Wearable epileptic seizure Prediction System based on machine learning techniques using ECG, PPG and EEG signals. Sensors, 22(23), p.9372.

[14] Jemal, I., Mezghani, N., Abou-Abbas, L. and Mitiche, A., 2022. An interpretable deep learning classifier for epileptic seizure prediction using EEG data. IEEE Access, 10, pp.60141-60150.

[15] Cheng, C., You, B., Liu, Y. and Dai, Y., 2021. Patient-specific method of sleep electroencephalography using wavelet packet transform and Bi-LSTM for epileptic seizure prediction. Biomedical Signal Processing and Control, 70, p.102963.

[16] Emara, H.M., Elwekeil, M., Taha, T.E., El-Fishawy, A.S., El-Rabaie, E.S.M., El-Shafai, W., El Banby, G.M., Alotaiby, T., Alshebeili, S.A. and Abd El-Samie, F.E., 2022. Efficient frameworks for EEG epileptic seizure detection and prediction. Annals of Data Science, 9(2), pp.393-428.

[17] Gao, Y., Chen, X., Liu, A., Liang, D., Wu, L., Qian, R., Xie, H. and Zhang, Y., 2022. Pediatric seizure prediction in scalp EEG using a multi-scale neural network with dilated convolutions. IEEE journal of translational engineering in health and medicine, 10, pp.1-9.

[18] Tamanna, T., Rahman, M.A., Sultana, S., Haque, M.H. and Parvez, M.Z., 2021. Predicting seizure onset based on time-frequency analysis of EEG signals. Chaos, Solitons & Fractals, 145, p.110796.

[19] Islam, M.S., El-Hajj, A.M., Alawieh, H., Dawy, Z., Abbas, N. and El-Imad, J., 2020. EEG mobility artifact removal for ambulatory epileptic seizure prediction applications. Biomedical Signal Processing and Control, 55, p.101638.

[20] Zhao, S., Yang, J. and Sawan, M., 2021. Energy-efficient neural network for epileptic seizure prediction. IEEE Transactions on Biomedical Engineering, 69(1), pp.401-411.

[21] Ouichka, O., Echtioui, A. and Hamam, H., 2022. Deep learning models for predicting epileptic seizures using iEEG signals. Electronics, 11(4), p.605.

[22] Mahmoodian, N., Haddadnia, J., Illanes, A., Boese, A. and Friebe, M., 2020. Seizure prediction with cross-higher-order spectral analysis of EEG signals. Signal, Image and Video Processing, 14, pp.821-828.

[23] Ra, J.S., Li, T. and Li, Y., 2021. A novel permutation entropy-based EEG channel selection for improving epileptic seizure prediction. Sensors, 21(23), p.7972.

[24] Singh, K. and Malhotra, J., 2022. Two-layer LSTM network-based prediction of epileptic seizures using EEG spectral features. Complex & Intelligent Systems, 8(3), pp.2405-2418.

[25] Xu, X., Lin, M. and Xu, T., 2022. Epilepsy seizures prediction based on nonlinear features of EEG signal and gradient boosting decision tree. International Journal of Environmental Research and Public Health, 19(18), p.11326.

[26] Dataset taken from: "https://www.kaggle.com/c/seizure-prediction/data?select=Patient_1", dated 1/1112024.