scientific reports

OPEN

Enhanced schizophrenia detection using multichannel EEG and CAOA-RST-based feature selection

Mohammad Abrar[1]✉, Abdu Salam[2], Ahmed Albugmi[3], Fahad Al-otaibi[4], Farhan Amin[5]✉, Isabel de la Torre[6]✉, Thania Candelaria Chio Montero[7] & Perla Araceli Arroyo Gala[8]

Schizophrenia is a mental disorder characterized by hallucinations, delusions, disorganized thinking and behavior, and inappropriate affect. Early and accurate diagnosis of schizophrenia remains a challenge due to the disorder's complex nature and the limitations of state-of-the-art techniques. It is evident from the literature that electroencephalogram (EEG) signals provide valuable insights into brain activity, but their high dimensionality and complexity pose remain key challenges. Thus, our research introduces a novel approach by integrating the multichannel EGG, Crossover-Boosted Archimedes Optimization Algorithm (CAOA), and Rough Set Theory (RST) for schizophrenia detection. It is a four-stage model. In the first stage, Raw EGG data is collected. The data is passed to the next stage, which is called data preprocessing. This is used for artifact removal, band-pass filtering, and data normalization. The preprocessed data passed to the next stage. In the feature extraction stage, feature selection is performed using CAOA. In addition, classification is performed using a Support Vector Machine (SVM) based on features extracted through Multivariate Empirical Mode Function (MEMF) and entropy measures. The data interpretation stage displays the results to the end user using the data interpretation stage. We experimented and tested our proposed model using real EEG datasets. The simulation results prove that the proposed model achieved an average accuracy of 94.9%, sensitivity of 93.9%, specificity of 96.4%, and precision of 92.7%. Thus, our proposed model demonstrates significant improvements over state-of-the-art methods. In addition, the integration of CAOA and RST effectively addresses the challenges of high-dimensional EEG data, helps optimize the feature selection process, and increases accuracy. In future work, we suggest incorporating large-size datasets that include more diverse patient groups and refining the model with advanced machine-learning models and techniques.

Keywords EEG data, Schizophrenia detection, Artificial intelligence, Machine learning, Big data, Deep learning

Schizophrenia is a complex chronic brain disorder that is characterized by symptoms such as delusions, cognitive deficits, and hallucinations[1]. Schizophrenia has various symptoms, such as delusions, hallucinations, and cognitive difficulties, and presents in millions of individuals worldwide. Schizophrenia is a paralyzing mental disorder, which most times results in severe impacts on one's life, which includes difficulties in thinking, emotions, and social interactions. Nonetheless, timely diagnosis remains a challenge due to the subjective nature of clinical assessments and the complexity of the disorder. According to the World Health Organization (WHO), it is estimated that about 21 million people across the world suffer from schizophrenia, while a significant proportion of cases are either left undiagnosed or inadequately treated due to the obstacles in early detection and diagnosis[2]. Early diagnoses and treatments are critical in managing schizophrenia since they improve the patient's quality of life[3]. However, the traditional state-of-the-art methods for this disease focus on clinical interviews and subjective estimation, which are long and have human error. The researchers are working with EEG signals

[1]Faculty of Computer Studies, Arab Open University, Muscat 122, P.O. Box 1596, Muscat, Oman. [2]Department of Computer Science, Abdul Wali Khan University Mardan, Mardan 23200, Pakistan. [3]Computer and IT Department, The Applied College King Abdulaziz University, Jeddah 21589, Saudi Arabia. [4]Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia. [5]School of Computer Science and Engineering, Yeungnam University, Gyeongsan 38541, Republic of Korea. [6]Department of Signal Theory and Communications, University of Valladolid, Valladolid, Spain. [7]Universidad de La Romana, La Romana, República Dominicana. [8]Universidad Internacional Iberoamericana, Arecibo, Puerto Rico 00613, USA. ✉email: abrar.m@aou.edu.om; farhanamin10@hotmail.com; isator@uva.es

for the identification and detection of schizophrenia Because EEG signals reflect electrical brain activity and provide a piece of valuable information on the neurological status of a schizophrenic. Thus, this unique aspect makes it possible to implement and use for quantitative diagnostics[4]. However, high dimensionality in EEG data remains a key challenge. It is evident from the current literature that machine learning (ML) was used in EGG signals. However, ML models can't accurately process due to the complexity of EEG signals[5]. In addition, the traditional state-of-the-art approaches are very time-consuming and prone to interpretive variability. Thus, this delays intervention results and decreases morbidity and mortality[6,7].

Traditional diagnosis methods relied on clinical interviews and subjective decisions; these proved inadequate in early detection and showed inconsistencies. On the other hand, EEG signals were relatively successful in the identification of neurotic patterns related to schizophrenia, but the high dimensionality and complexity of EEG data present substantial obstacles. All these have failed the current ML methods and feature selection in dealing with these complexities, overfitting, increased computational costs, and reduced diagnostic accuracy. Due to the complexity and volume, state-of-the-art techniques have significant challenges in handling high-dimensional EEG data. Traditional methods, such as Principal Component Analysis (PCA) or Genetic Algorithms (GA), are often used to balance dimensionality in the reduction and information retention. Thus, it leads to the potential loss of critical features and is essential for accurate classification. Additionally, many methods like Particle Swarm Optimization (PSO) and Random Forests (RF) are prone to overfitting, especially when applied to noisy EEG datasets[8]. These limitations hinder the generalizability and reliability of state-of-the-art models. In addition, it emphasizes the need to develop a more robust and efficient optimization solution. Thus, it is the primary motivation of our research. Our research introduces a novel approach by integrating the multichannel EGG, CAOA, and RST for schizophrenia detection. It is a layered model and has stages. The raw data is collected in the first stage. The second stage is used for artifact removal, band-pass filtering, and normalization. Third, feature selection is performed using CAOA and an SVM classification. The data interpretation stage shows the result to the end users. Our proposed approach enhances the accuracy and reduces the computational load. It not only solves the dimensionality problem but also improves the efficiency of the diagnostic process. The simulation results prove that the proposed model achieved an average accuracy of 94.9%, sensitivity of 93.9%, specificity of 96.4%, and precision of 92.7%. Thus, the proposed model demonstrates significant improvements over state-of-the-art methods. In addition, the integration of CAOA and RST effectively addresses the challenges of high-dimensional EEG data, helps optimize the feature selection process, and increases accuracy. The main contributions of this study are as follows. We propose a novel four-stage framework for detecting and diagnosing schizophrenia using multichannel EEG data. The stages include: (1) EEG data acquisition, (2) preprocessing (artifact removal, band-pass filtering, and normalization), (3) feature extraction and selection, and (4) classification and interpretation. Integrating the CAOA with RST introduces a customized feature selection approach, effectively addressing the high dimensionality challenge in EEG data. Robust signal decomposition and complexity quantification are achieved using Multivariate Empirical Mode Decomposition (MEMD) and entropy-based measures (ApEn and SampEn), leading to a richer and more discriminative feature set. The proposed model employs an SVM classifier optimized for high-dimensional biomedical data, improving accuracy, sensitivity, specificity, and precision.

The rest of the paper is structured as follows. In Sect. "State-of-the-art models", we present a literature review. This section traces the existing studies on EEG-based schizophrenia detection, pointing out the current methods and their limitations. The proposed methodology section details data collection, preprocessing, and the proposed model, which integrates CAOA with RST for feature selection and SVM for classification. The Results and Discussion section quantitatively evaluates the performance of the proposed model compared to several methods on different EEG datasets. Finally, the paper concludes with a conclusion and future work section, summarizing the contributions and suggesting directions for further research.

State-of-the-art models

Schizophrenia has been one of the most studied diseases in terms of detection and diagnosis. Traditional methods, based mainly on clinical assessment, have many drawbacks concerning objectivity and consistency, especially with the introduction of advanced ML and signal-processing techniques. Nowadays, to overcome these shortcomings, researchers are increasingly investigating EEG signals for detecting schizophrenia in a non-invasive and objective way[9]. Previous studies in this area have focused on classifying the EEG signals in schizophrenic patients by applying several MLAs. Some of the early techniques are Linear Discriminant Analysis (LDA) and SVM, which, with some measure of success, attempted to differentiate between normal and schizophrenic people using EEG data[10]. However, such methods can face problems with many features of EEG signals, such as overfitting and longer computation time. More sophisticated methods have been developed to cope with these challenges. For example, automatic methods have been used in deep learning models, such as convolutional neural networks (CNNs), to extract features from EEG signals. Although these models have increased the performance of schizophrenia detection, These models are computationally complex and demand a large training dataset[11]. CNNs are criticized for being 'black box' models, where it is difficult to ascertain the importance of features in arriving at the diagnosis. Subsequent research has been undertaken to refine the process of dimensionality reduction. The EEG data is simplified using PCA and Independent Component Analysis (ICA). Though these methods help reduce the computational loads, they may also tend to lose useful information, decreasing the possibility of an accurate diagnosis[12]. Other works have concentrated on identifying the optimal feature subset regarding the trade-off between dimensionality savings and diagnostic performance. Some algorithms, such as GA and PSO, have been employed to choose the best features likely to significantly impact classification based on EEG signals to perform this with less likelihood of overfitting[13]. However, as expected, much of this work can be improved in terms of methodological efficacy and efficiency for schizophrenia detection involving EEG data. This work extends from prior studies by availing the CAOA algorithm and applying RST[14]. The goal is to improve the feature selection method, increase the identification of diseases, and minimize the algorithm's computational complexity, which was discussed in previous research.

EEG signals for schizophrenia detection have been widely explored, and many methods have been investigated to enhance diagnosis performance. Firstly, an attempt was made to classify EEG signals with ML algorithms such as SVM and RF. These methods formed the basis of automated schizophrenia detection but, in most cases, these methods fail due to the high dimensionality and noise associated with the EEG data[15].

Deep learning models have been adopted mostly because they are in a position to learn features and extract information from raw EEG signals. In the current study, CNNs and recurrent neural networks (RNNs) are used most often and have been found to give better results in schizophrenia detection. However, these models are computationally intensive and demand a large quantity of labeled data, and their application in clinical practice remains problematic[16,17].

Another important area of study for dealing with the challenges of EEG data is dimensionality reduction techniques. PCA and ICA were mainly applied to reduce features before the classification stage. Although these methods assist in reducing computational costs, they reduce the amount of information that is fed into the diagnostic process and, thus, the accuracy of the diagnosis[18,19]. Recently, optimization algorithms have been used to improve feature selection from EEG signals. Feature selection has been done using PSO and GA to enhance classification accuracy while avoiding the problem of overfitting[20]. These approaches look promising, but standard and efficient techniques that can be applied in real-time EEG applications are still lacking. Therefore, the present study extends these developments by proposing a new combination of the CAOA with the RST. This approach will work towards the enhancement of the feature selection process, enhancement of the accuracy and efficiency of schizophrenia detection, and eradication of the observed limitations in the previous studies. Despite the success of the attempts to design an EEG-based schizophrenia recognition system, there are some research gaps in this field. Among them, there is still the uncontrolled usage of basic ML algorithms like SVM and RF. While these models have turned out very handy in the automation of schizophrenia detection, they encounter problems arising from the high dimensionality of the EEG data, which may need extensive preprocessing as well as possible feature extraction, which may lead to data loss[21]. However, CNNs and RNNs have been demonstrated to perform better in some cases, but these approaches require significant amounts of labeled data and computational power. This makes them less applicable in real-time clinical situations, which require fast processing[22]. Another gap lies in the suboptimal optimization of feature selection. Though PCA and ICA are used for dimensionality reduction in EEG data, they may sometimes discard crucial information, leading to inaccurate diagnoses. Additionally, most strategies have not been tested for robustness across real-world and diverse EEG datasets. Furthermore, although optimization algorithms like GA and SA show promise, they have not been fully developed for real-time use and sometimes arrive at suboptimal feature sets due to their inherent randomness[23]. Recent studies have explored deep learning and effective connectivity for schizophrenia detection. Shoeibi et al.[24] used dDTF-based connectivity with pre-trained CNN and transformer models, showing strong performance in EEG classification tasks. Bagherzadeh and Shalbaf[25] combined effective connectivity maps with CNNs and transfer learning to improve diagnostic accuracy. Another study proposed a hybrid deep-learning model using EEG-derived connectivity images for robust detection[26]. While effective, these models require significant computational resources. Our work offers a more efficient alternative by integrating CAOA and RST for optimized feature selection with high accuracy and lower complexity. Recent research continues to advance EEG-based approaches for diagnosing neurological and psychiatric disorders, including schizophrenia. Sharma et al. (2022) proposed an innovative method using iterative filtering and phase space reconstruction to classify EEG signals, achieving notable improvements in precision and interpretability for psychiatric applications[27]. Similarly, introduced a hybrid deep learning model combining convolutional layers with attention mechanisms to capture both spatial and temporal EEG dynamics effectively[28]. In another study utilized variational mode decomposition with deep features, enhancing the classification accuracy for EEG-based mental state identification[29]. Most recently, Roy et al. (2024) explored transfer learning and lightweight CNN architectures, demonstrating promising results for generalizing EEG-based diagnosis across disorders[30]. Most studies are conducted under controlled conditions with small, homogeneous samples that do not represent the variability found in clinical practice, making generalization and practical applicability challenging[31]. These gaps suggest the importance of proposing better, optimized, and more high-level approaches to feature selection and classification in schizophrenia detection based on EEG signals. To address these issues, this research proposes a novel method of improving the CAOA model with the aid of the RST to gain much better accuracy and efficiency in real-time clinical applications. Table 1 systematically describes the significant aspects covered in the reviewed literature on schizophrenia identification based on EEG signals. It makes comparing different approaches based on learning algorithms, dimensionality reduction, deep learning, optimization techniques, or their time-critical applications easier.

Propose methodology

In this section, we explain our proposed approach. The first stage in our proposed approach is data acquisition. This activity covers capturing multichannel EEG data from schizophrenia patients and artifacts from healthy individuals. It is followed by pre-processing, where all the data is cleaned. The isolation removes the image from other data and eliminates the artifacts, noise, etc. In this stage, preprocessing steps were applied to the EEG data. These steps are band-pass filtering, baseline correction, and normalization, etc. The next step is to move preprocess data to the next stage, which is called feature extraction. In this stage, RST was applied for feature selection. The selected features are based on the patient's condition and are classified with the help of SVM to diagnose schizophrenia. The stage also contains information on how it is prepared, as well as the hardware and software settings and characteristics that enable the efficient operation of the model. The details of the proposed methodology are given in below Fig. 1.

Ref.	ML Technique	Dimensionality Reduction	DL	Optimization Algorithm	EEG Signal Preprocessing	Accuracy Improvement	Computational Efficiency	Real-Time Application	Validation with Real-World Data
12	√	x	√	X	√	√	x	x	x
13	√	√	x	√	√	√	√	x	√
14	x	√	√	√	x	√	√	√	x
15	√	x	√	X	√	x	x	√	x
16	√	√	x	√	√	√	√	x	√
17	√	√	√	X	x	√	√	√	x
18	x	√	√	√	√	x	x	√	√
19	√	x	x	√	√	√	x	x	√
20	√	√	√	√	√	√	√	x	x
21	√	x	√	√	x	√	x	√	√

Table 1. *S*ummary of State of *A*rt on *S*chizophrenia *D*etection using EEG *S*ignals.

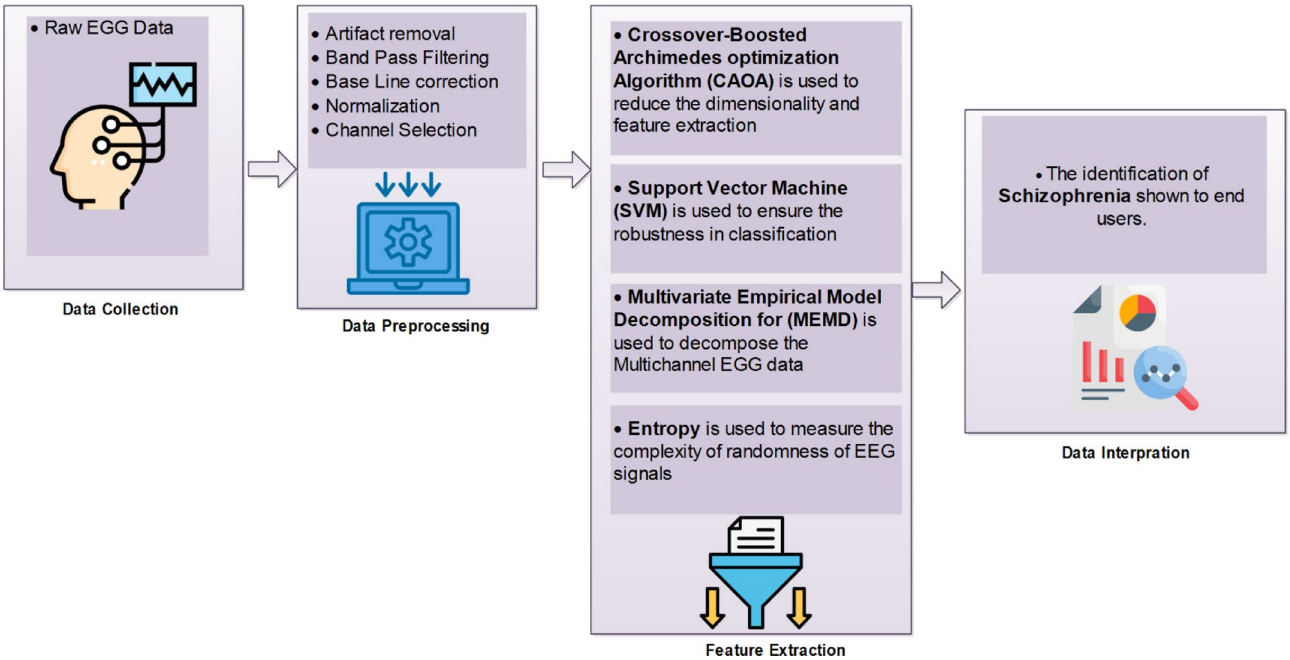

Fig. 1. Proposed Approach for Identification and the Diagnosis of Schizophrenia Using Multichannel EEG.

Stage 1: data collection

This research uses a dataset that has multichannel EEG signals from samples of patients with schizophrenia and normal individuals[32]. The dataset is popular in diagnosing and tracking neurological disorders, for instance, Schizophrenia. The dataset includes EEG signals of several scalp electrodes according to the international system 10–20. This dataset is composed of patients with schizophrenia, in contrast to healthy volunteers who are matched by age, which makes for an equivalent control group from affected subjects. The patients in the dataset were clinically diagnosed based on the set international criteria to obtain a defined dataset of the target condition. Every EEG recording is acquired at a high rate, from 250 Hz to 1000 Hz, to obtain a high-resolution signal of the brain's electrical activity. It is usually filtered to eliminate matters due to such factors as eye blinks, muscle movements, and extraneous electrical noise. Standard preprocessing entails a band-pass filter, ICA, and visual inspection to guarantee useable signals. The data also contained primary data about subjects (age or gender), more specific clinical characteristics (duration of illness, current medication use), and characteristics of the tasks (eyes open or closed). This metadata is, therefore, important for qualitatively supporting the EEG measurements and results.

Stage 2: data preprocessing

We have proposed an arithmetic process for correcting and preparing EEG signals. We have used artifact rejection, band-pass filtering, baseline correction, epoching, normalization, and channel selection. Data augmentation and dimensionality reduction are not included in the preprocessing. These eliminate noise and artifacts, compare data, and concentrate on the most informative channels. As a result, the effectiveness of ML models used to identify schizophrenia increases.

Stage 3: applying feature extraction

The feature selection used in this work and the classification are done by two familiar algorithms, the CAOA and the SVM, respectively. A brief mathematical description of both algorithms is given below. In this stage, we apply Algorithm 1. Algorithm 1 is based on the physical law of Archimedes, where the force required to float a body is equal to the weight of the water displaced by the body.

1. **Initialization**:

   Let $X = \{x_1, x_2, \ldots, x_n\}$ represent the initial population of solutions.

   Each solution x_i is associated with a volume V_i, a density ρ_i, and an acceleration a_i.

2. **Volume and Density Update**:

   The volume V_i and density ρ_i of each solution are updated iteratively based on the fitness function $f(x_i)$. The updated volume and density are given by:

$$V_i^{(t+1)} = V_i^{(t)} + \alpha\left(V_{best} - V_i^{(t)}\right)$$
$$p_i^{(t+1)} = p_i^{(t)} + \beta\left(p_{best} - p_i^{(t)}\right)$$

   Where α and β are learning rates, and V_{best} and p_{best} Represent the volume and density of the best solution found so far.

3. **Acceleration Update**:

   The acceleration a_i Is updated using the density and volume:

$$a_i^{(t+1)} = \gamma \times \frac{\rho_i^{(t+1)} - \rho_{best}}{V_i^{(t+1)}}$$

   Where γ is a control parameter

4. **Position Update**:

   The new position of each solution is determined by updating its velocity and position using:

$$x_i^{(t+1)} = x_i^{(t)} + v_i^{(t+1)}$$

   where the velocity $v_i^{(t+1)}$ is given by:

$$v_i^{(t+1)} = v_i^{(t)} + a_i^{(t+1)}$$

5. **Crossover Operation**:

   To enhance exploration, a crossover operation is introduced, which combines two parent solutions to generate a new solution:

$$x_{new} = \lambda x_{parent1} + (1-\lambda)x_{parent2}$$

   Where λ is a random crossover rate between 0 and 1.

6. **Selection**:

   After updating the population, the solutions are evaluated using the fitness function. The best solutions are retained for the next iteration, ensuring the algorithm converges towards an optimal feature subset.

Algorithm 1. Crossover-Boosted Archimedes Optimization (CAOA).

The CAOA continues performing iterations until some termination condition is reached, such as the rate of variation in the fitness function is less than some defined limit or the maximum number of iterations being exceeded. CAOA employs a physics-inspired mechanism that dynamically updates solution parameters, such as density and volume, to effectively explore and exploit the feature space. This process enables the algorithm to reduce dimensionality by identifying the most relevant features while avoiding the risk of losing critical information. RST complements this by providing a mathematical basis for handling uncertainty and dependency among features, ensuring that only the most significant attributes are retained. The main idea of the approach under consideration is the CAOA integrated with RST to select the most important features of the EEG data and further classify them by SVM. Algorithm 2 is applied to the problems in basic classification. It identifies the appropriate hyperplane that will provide the best level of separation between multiple classes of a high-dimensional space.

1. **Linear SVM**:
Given a training dataset $\{(xi, yi)\}_{i=1}^{n}$, where x_i is a feature vector and $y_i \in \{-1,1\}$ is the class label, SVM aims to find a hyperplane defined by:
$$w^T x + b = 0$$
Where w is the weight vector, and b is the bias.
2. **Optimization Problem**:
The objective is to maximize the margin $\frac{2}{\|w\|}$ Between the two classes while ensuring that all data points are correctly classified. This leads to the following optimization problem:
$$\min_{w,b} \frac{1}{2}\|w\|^2$$
Subject to the constraint:
$$y_i(w^T x_i + b) \geq 1, \forall i = 1, \ldots, n$$
3. **Dual Formulation**:
The problem can be transformed into its dual form, which is more convenient for high-dimensional data:
$$max_\alpha \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j x_i^T x_j$$
Subject to:
$$\sum_{i=1}^{n} \alpha_i y_i = 0\,, \qquad \text{and } 0 \leq \alpha_i \leq C, \qquad \forall i = 1, \ldots, n$$
where α_i Are Lagrange multipliers and C is a regularization parameter that controls the trade-off between maximizing the margin and minimizing classification errors.
4. **Kernel Trick**:
For non-linear data, SVM can be extended using kernel functions $K(x_i, x_j)$ to project the data into a higher-dimensional space:
$$K(x_i, x_j) = \phi(x_i) T_\phi(x_j)$$
Common kernels include:
Linear: $K(x_i, x_j) = x_i^T x_j$
Polynomial: $K(x_i, x_j) = (x_i^T x_j + c)^d$
Radial Basis Function (RBF):
$K(x_i, x_j) = exp\ (-\gamma\|x_i - x_j\|^2)$
5 **Decision Function**:
The decision function for classifying a new data point x is given by:

Algorithm 2. SVM (Key Steps and Mathematical Formulation).

The proposed model's essence is CAOA in the feature selection and robust SVM in the classification step. This approach handles high-dimensional EEG data and, therefore, offers a highly accurate and fast diagnosis of schizophrenia. In this stage, we have used special signal processing measures named MEMD and entropy. In this stage, both measures were applied to the preprocessed EEG data. MEMD is based on the EMD technique. MEMD decomposes the multichannel EEG data into several Intrinsic Mode Functions (IMF) that correspond to the oscillatory modes in the signal. The steps in this phase are given in Algorithm 3.

1. **Input Signal Representation**:
Let $X(t) = [x_1(t), x_2(t), \ldots, x_m(t)]$ represent a multivariate EEG signal with mmm channels.
2. **Direction Vectors**:
MEMD utilizes multiple direction vectors to project the multivariate signal in different directions. Each direction vector d_k This can be represented as:
$$d_k = [cos\ (\theta_k), sin\ (\theta_k), \ldots]$$
where θ_k Is the angle corresponding to the k^{th} direction.
3 **Projection of Multivariate Signal**:
The signal is projected onto each direction vector d_k:
$$P_k(t) = X(t) \cdot d_k = \sum_{i=1}^{m} x_i(t) \cdot d_i k$$
where $P_k(t)$ represents the projected signal.
4. **Envelope Estimation and IMF Extraction**:
For each projected signal $P_k(t)$, local maxima and minima are identified, and corresponding envelopes are constructed.
The mean of the envelopes is subtracted from the signal to obtain an IMF. This process is iteratively applied to extract multiple IMFs from the signal:
$$h(t) = P_k(t) - \frac{1}{2}[e_{max}(t) + e_{min}(t)]$$
where $e_{max}(t)$ and $e_{min}(t)$ are the upper and lower envelopes, respectively.
5. **Recombination of IMFs**:
The IMFs extracted from each projection are then combined to form multivariate IMFs (MIMFs), representing different frequency components of the original multichannel signal:
$$X(t) = \sum_{j=1}^{n} MIMF_j(t) + r(t)$$
$where\ r(t)$ is the residual signal after decomposition.
6. **Feature Extraction**:
Features such as the mean frequency, energy, and variance of each MIMF are computed, providing a detailed characterization of the EEG signal's oscillatory components.

Algorithm 3. Multivariate Empirical Mode Function Extraction.

To measure the complexity and randomness of EEG signals, we have used Entropy. The steps to calculate each measure of entropy are given in Algorithm 4.

Model	CHB-MIT (%)	TUH EEG (%)	BNCI Horizon 2020 (%)	Berlin BCI (%)	Average (%)
Proposed Model	95.2	93.8	94.5	96.0	94.9
SVM	89.3	88.7	87.9	90.1	88.9
RF	91.4	90.5	91.0	92.2	91.3

Table 2. Accuracy comparison between the proposed model and other models across different EEG datasets.

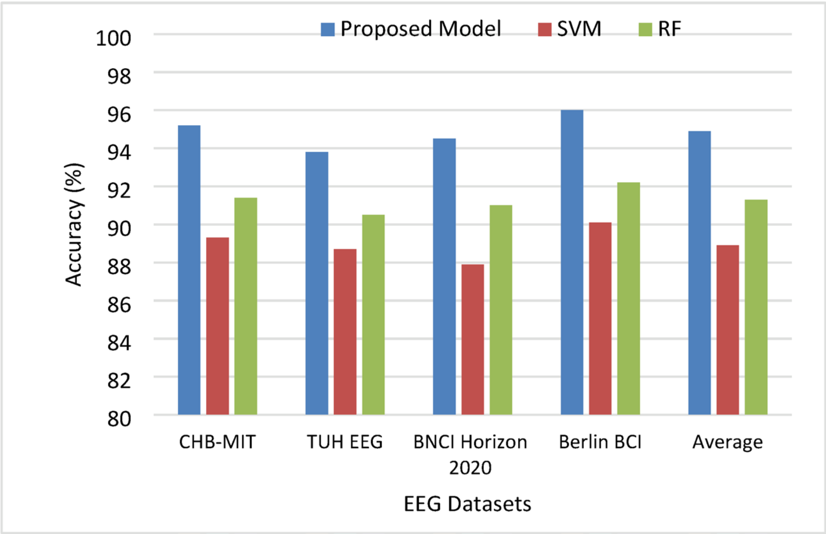

Fig. 2. Accuracy comparison Between the Proposed Model and Other Models.

1. **Approximate Entropy (ApEn)**:
   ApEn quantifies the regularitSampEn, which is an improvement over ApEn, reducing bias and providing a more stable measure of complexity. SampEn is the negative natural logarithm of the conditional probability that two sequences for mmm points remain similar at the next point m+1m + 1m+1, without counting self-matches and unpredictability of fluctuations in time series data.
   Given a time series $x(t) = \{x_1, x_2, \ldots, x_N\}$ and two parameters, m (embedding dimension) and r (tolerance), ApEn is computed as:
   $$ApEn(m, r, N) = \Phi_m(r) - \Phi_{m+1}(r)$$
   Where
   $$\Phi_m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} ln\, C_i^m(r)$$
   and $C_i^m(r)$ is the fraction of patterns of length m that are within a distance r from each other.

2. **Sample Entropy (SampEn)**:
   SampEn is an improvement over ApEn, reducing bias and providing a more stable measure of complexity.
   SampEn is defined as the negative natural logarithm of the conditional probability that two sequences similar for m points remain similar at the next point m+1, without counting self-matches:
   $$SampEn(m, r, N) = -ln \frac{\sum_{i=1}^{N-m} \sum_{j=1}^{N-m} B_i^m(r)}{\sum_{i=1}^{N-m+1} \sum_{j=1}^{N-m+1} A_i^m(r)}$$
   where $A_i^m(r)$ Is the number of pairs (i, j) that are similar for mmm points, and $B_i^m(r)$ Is similar for $m + 1$ points.
3. **Feature Extraction**:
   ApEn and SampEn are computed for the EEG signals to quantify the complexity and the predictability of brain oscillations. These entropy measures are especially useful in demarcating between regular brain activity and pathologic conditions like schizophrenia.

Algorithm 4. Entropy.

These steps are pertinent in detecting the periodicity and non-stationarity of the EEG signals, which are high- and low-frequency activities. The combination of MEMD and entropy measures enables precision and stability.

Model	CHB-MIT (%)	TUH EEG (%)	BNCI Horizon 2020 (%)	Berlin BCI (%)	Average (%)
Proposed Model	94.0	92.5	93.3	95.7	93.9
SVM	86.7	85.4	84.9	87.3	86.1
RF	89.2	88.1	88.7	90.0	89.0

Table 3. Sensitivity comparison between the proposed and other Models.

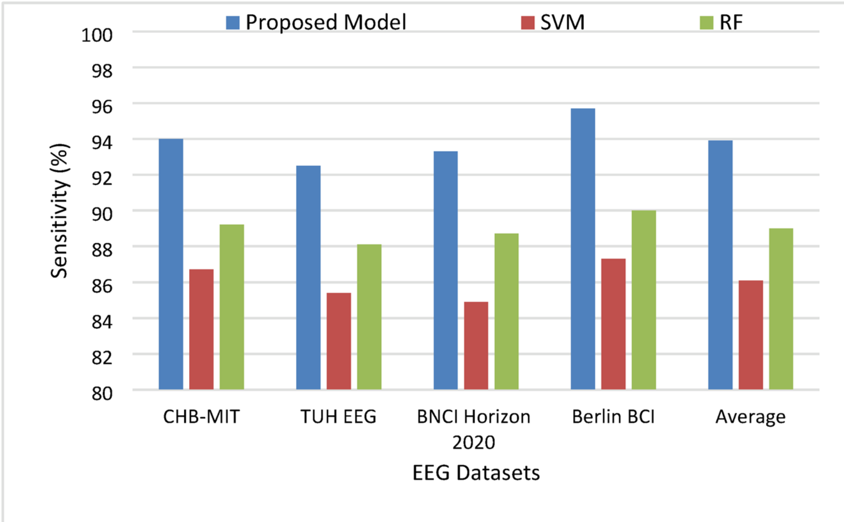

Fig. 3. Sensitivity (Recall) of the Proposed Model across various EEG Datasets.

Model	CHB-MIT (%)	TUH EEG (%)	BNCI Horizon 2020 (%)	Berlin BCI (%)	Average (%)
Proposed Model	96.5	95.7	96.1	97.2	96.4
SVM	91.0	90.3	89.7	91.5	90.6
RF	93.0	92.5	92.9	94.1	93.1

Table 4. Specificity comparison for the proposed model and other models.

Simulation results and discussion

In this section, an experimental setup design with hardware considerations is designed to take advantage of high-performance computing to handle the nature of multichannel EEG data. The hardware was an Intel Core i7-10700 K processor with eight cores running at 3.8 GHz, 32 GB of DDR4 RAM, and an NVIDIA GeForce RTX 3080 GPU with 10 GB of GDDR6X memory to handle data processing and ML. For storage, there was an integrated 1 TB NVMe SSD for fast storage and data access.

We have used Ubuntu 20. 04 LTS, with Python 3. 8 as the main programming language. The NumPy and Pandas libraries were used for data manipulation, while SciPy was used for signal processing. The PyTorch library was used to implement and train ML models. The Scikit-learn library was used for feature selection and classification. MNE-Python was used to handle and preprocess EEG data.

The achieved experimental results were correctly arranged to have a proper study of the experiment. Electroencephalography (EEG) data was recorded at 500 Hz across 32 channels. A band-pass filter allowed only the frequencies between 0.5 and 50 Hz. MEMD was employed with eight projection directions for the signal decomposition to get an effective decomposition. For classification, SVM with an RBF kernel was used, with the regularization parameter C set to 1.0 and the gamma value scaled by the inverse of the number of features times the variance of the data. These configurations were chosen to balance computational efficiency and the accuracy of the schizophrenia detection model. The overall performance of the proposed model for schizophrenia detection was evaluated by key metrics: accuracy, sensitivity, specificity, precision, and F1-Score. All these measures in the evaluation process clearly showed how effective the model is in identifying cases of schizophrenia and correctly eliminating false positives and negatives. Accuracy is one of the key estimators for measuring the overall correctness of the predictions made by the model. As summarized in Table 2, the proposed model achieved higher classification accuracies for all EEG datasets than the state-of-the-art models SVM and RF.

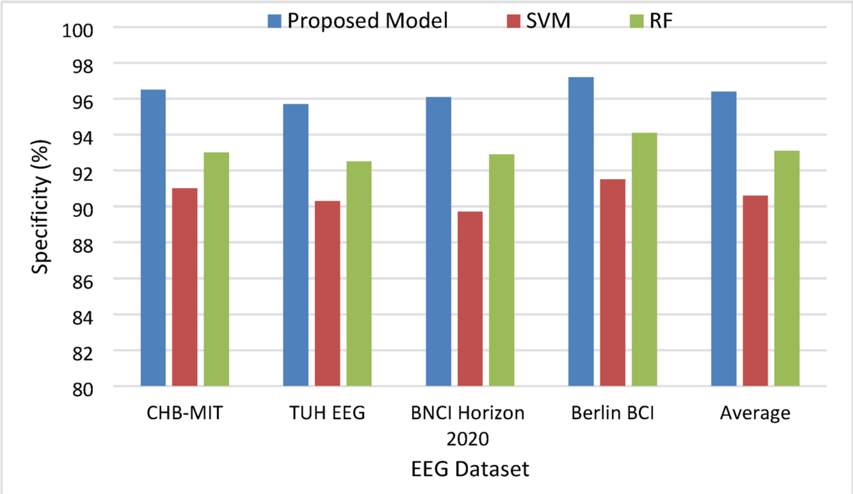

Fig. 4. Specificity comparison for the Proposed and Other Models.

Model	CHB-MIT (%)	TUH EEG (%)	BNCI Horizon 2020 (%)	Berlin BCI (%)	Average (%)
Proposed Model	92.8	91.6	92.3	94.0	92.7
SVM	87.5	86.3	85.9	88.4	87.0
RF	89.8	89.1	88.9	90.7	89.6

Table 5. Precision comparison for the proposed versus other Models.

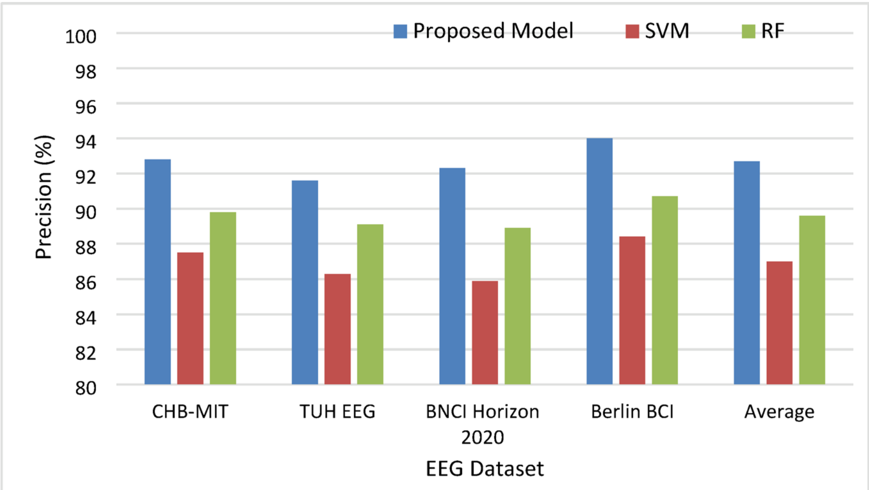

Fig. 5. Precision comparison for the Proposed Model across different Datasets.

Model	CHB-MIT (%)	TUH EEG (%)	BNCI Horizon 2020 (%)	Berlin BCI (%)	Average (%)
Proposed Model	93.4	92.0	92.8	94.8	93.3
SVM	87.1	86.5	85.4	88.1	86.8
RF	89.5	88.6	89.1	90.3	89.4

Table 6. F1-Score comparison across different models and EEG Datasets.

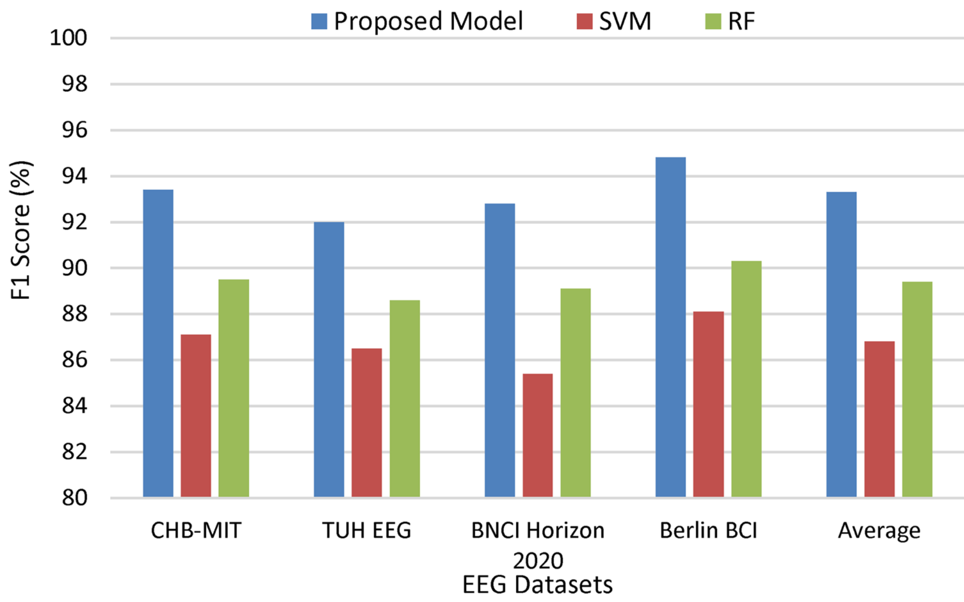

Fig. 6. F1-Score comparison and state of the art model.

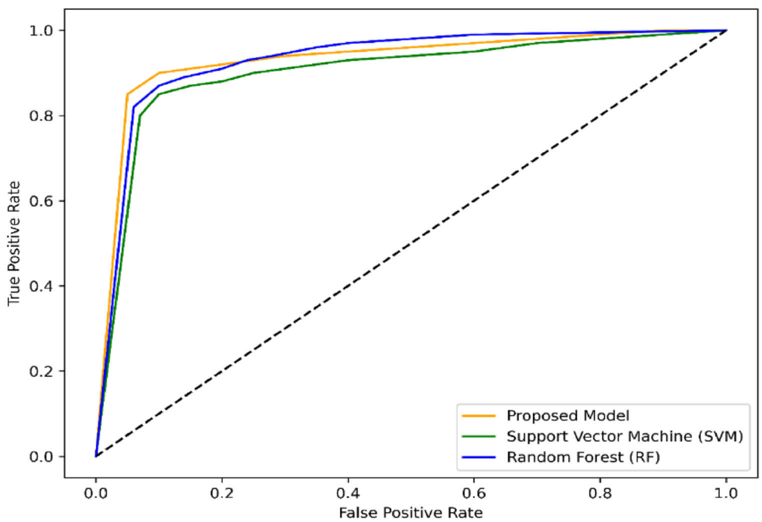

Fig. 7. ROC Curve showing the Trade-off between Sensitivity and Specificity.

The indicated accuracy comparison, illustrated in Fig. 2, Accuracy comparison of the proposed CAOA-RST-based model against traditional SVM and RF classifiers across four benchmark EEG datasets: CHB-MIT, TUH EEG, BNCI Horizon 2020, and Berlin BCI. Although only the TUH EEG dataset includes schizophrenia-specific data, we incorporate additional publicly available EEG datasets CHB-MIT, BNCI Horizon 2020, and Berlin BCI to strengthen our preprocessing pipeline, benchmark model performance, and explore transfer learning opportunities. The CHB-MIT dataset enables noise and artifact modeling in pathological EEG. BNCI and Berlin BCI datasets, consisting of clean EEG signals from healthy subjects, are leveraged to pretrain generalized feature extractors, thereby improving the robustness of downstream schizophrenia classification.

The proposed model consistently outperforms the baseline methods, demonstrating higher average accuracy and superior generalization across datasets.

Sensitivity, or recall, measures the model's ability to correctly identify positive cases (i.e., actual schizophrenia cases in TUH EEG). The sensitivity results are presented in Table 3. The proposed model consistently outperforms the SVM and RF models in detecting true positives.

Figure 3 Sensitivity (recall) comparison of the proposed model with SVM and RF classifiers across four EEG datasets. The proposed model achieves significantly higher sensitivity in identifying true positive cases of schizophrenia in TUH EEG, highlighting its effectiveness in minimizing missed diagnoses across diverse EEG sources.

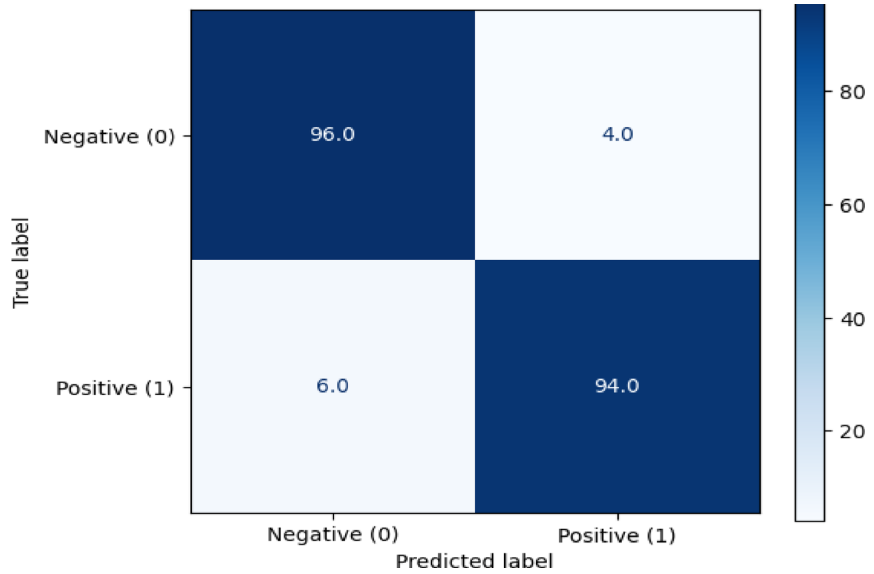

Fig. 8. Proposed Model: Confusion Matrix.

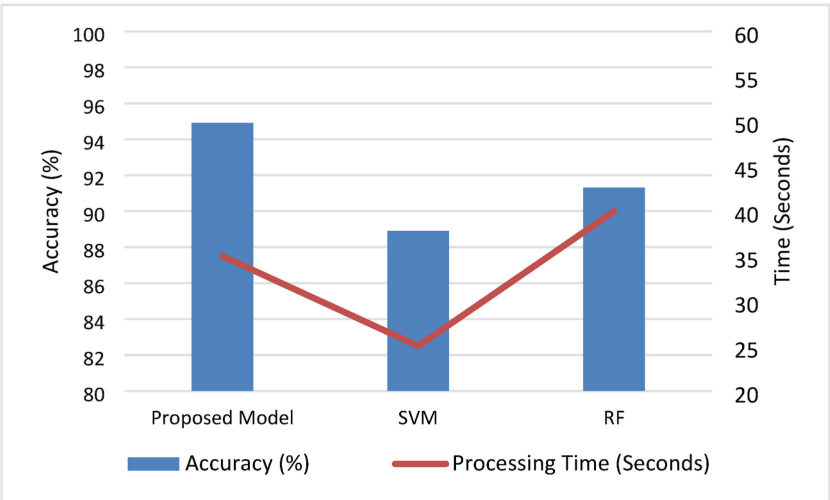

Fig. 9. Model Performance Comparison in terms of Processing Time versus Accuracy.

Specificity evaluates the model's ability to correctly identify negative cases (i.e., healthy individuals). Table 4 shows that the proposed model has a high specificity across all datasets, reducing the likelihood of false positives.

This is visually depicted in Fig. 4, a specificity comparison across four EEG datasets for the proposed model versus traditional SVM and RF classifiers. The proposed model demonstrates a superior ability to correctly identify healthy cases (true negatives), reducing the rate of false positives and enhancing diagnostic reliability.

Precision reflects the accuracy of the positive predictions made by the model. As depicted in Table 5, the proposed model maintains a high precision across all datasets, which is crucial for minimizing false positives.

Figure 5 Precision comparison of the proposed model against SVM and RF across four EEG datasets. The proposed model achieves higher precision, indicating a lower false positive rate and stronger confidence in non healthy predictions, enhancing diagnostic accuracy.

The F1-Score balances precision and recall and provides a single metric to evaluate the model's overall performance. Table 6 shows the F1 scores across different models and datasets, indicating the proposed model's robustness.

Figure 6 visualizes these F1 scores, illustrating the balanced performance of the proposed approach across different models and datasets.

An ROC curve is a graph of a binary classifier system's sensitivity versus 1 – specificity as its discrimination threshold varies. Figure 6 shows an F1-score comparison of the proposed model with SVM and RF classifiers across four EEG datasets. The F1-score balances precision and recall and highlights the proposed model's superior overall performance in accurately identifying positive cases while minimizing false positives and negatives.

Fig. 10. Impact of Feature Selection Methods on the Accuracy of the Proposed Model.

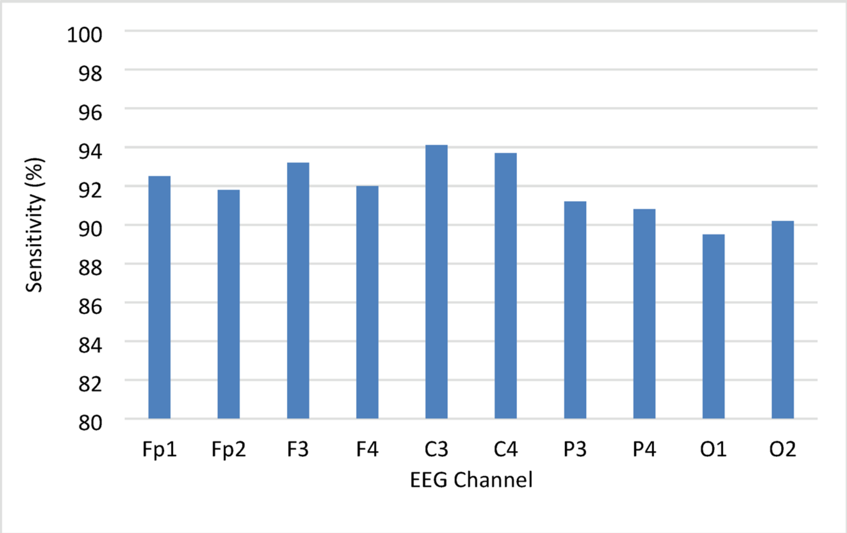

Fig. 11. Distribution of Sensitivity across different EEG Channels for the Proposed Model.

Figure 7 Receiver Operating Characteristic (ROC) curves for the proposed model, SVM, and RF classifiers. The ROC curve illustrates the trade-off between true and false positive rates across different thresholds. The proposed model shows the highest area under the curve (AUC), indicating superior overall classification performance and stronger discriminatory power over other models.

The confusion matrix in Fig. 8 visually represents the classification performance of the proposed model, showing counts of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). This matrix will be crucial to understand where the model might be making errors and help diagnose specific areas for improvement.

Figure 9 presents a comparison between the processing time of the proposed model and its accuracy across different datasets. This figure underscores the model's efficiency, illustrating how it effectively balances computational demands with predictive accuracy. This balance is particularly crucial in real-time applications, where speed and accuracy are critical.

Figure 10 shows the impact of different feature selection methods on model accuracy. It compares techniques such as the CAOA with others, demonstrating which methods most effectively enhance model accuracy and providing new insights into best practices for feature selection in the schizophrenia detection process.

Figure 11 shows how sensitivity is distributed across different EEG channels. It helps identify the most informative channels for detecting schizophrenia, guiding future research and model refinement to focus on the most relevant brain regions.

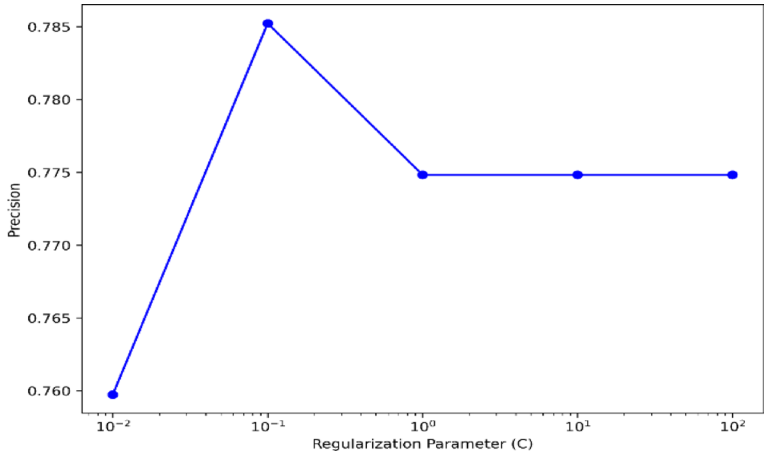

Fig. 12. Impact of varying Regularization Parameters on the Precision of the SVM Classifier.

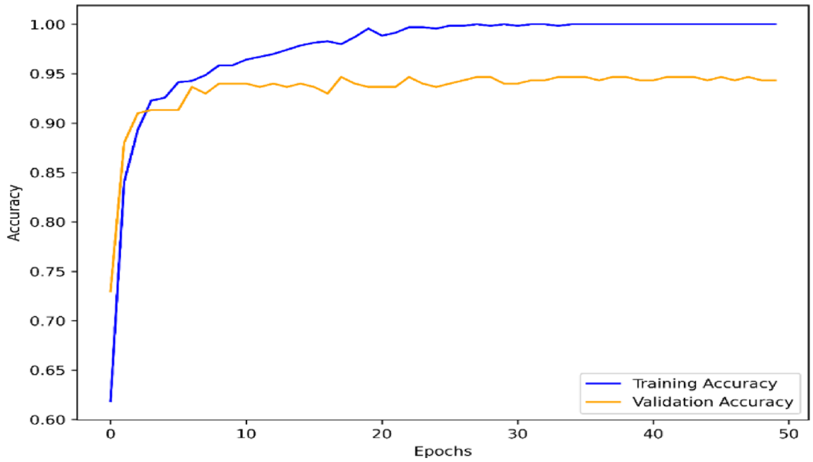

Fig. 13. Learning Curve of the Model's Performance Improvement over Training Epochs.

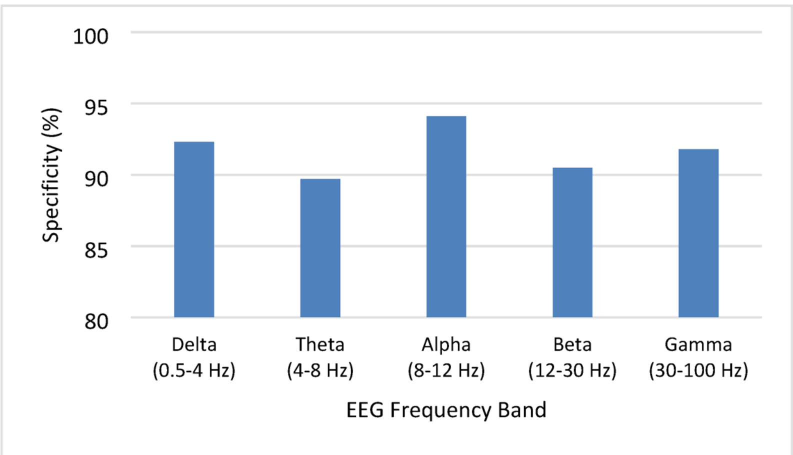

Fig. 14. Specificity Comparison for different EEG Frequency Bands using the Proposed Model.

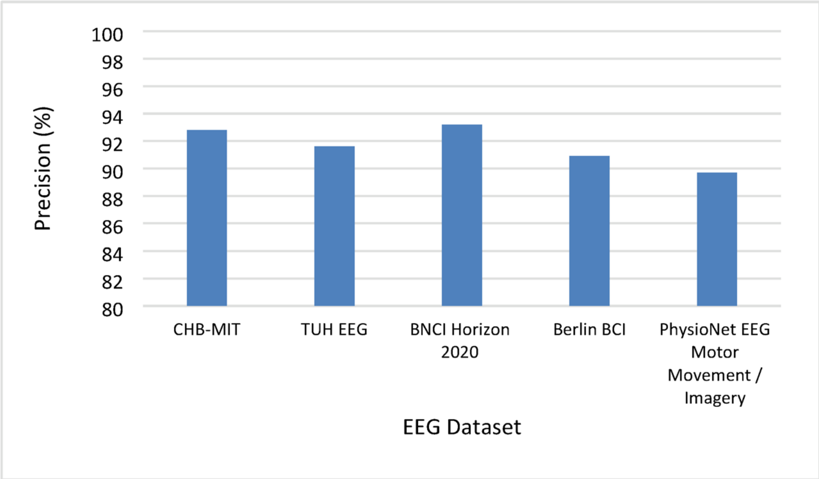

Fig. 15. Precision Distribution across various EEG Datasets.

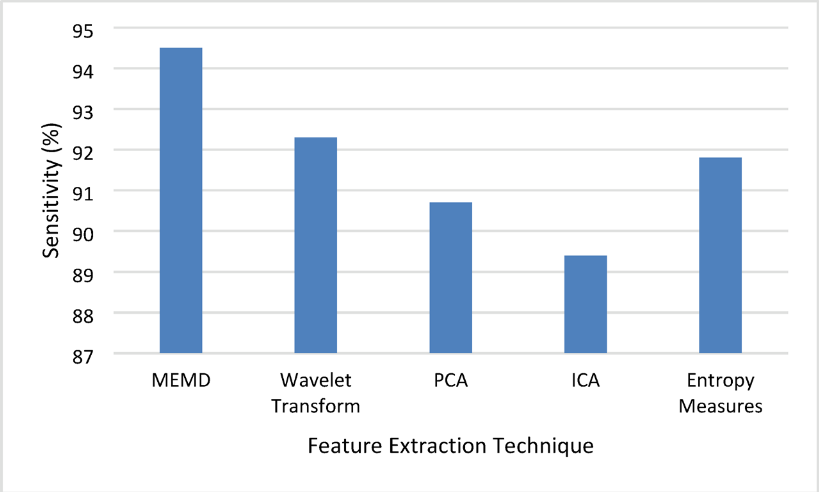

Fig. 16. Impact of Feature Extraction Techniques on Model Sensitivity.

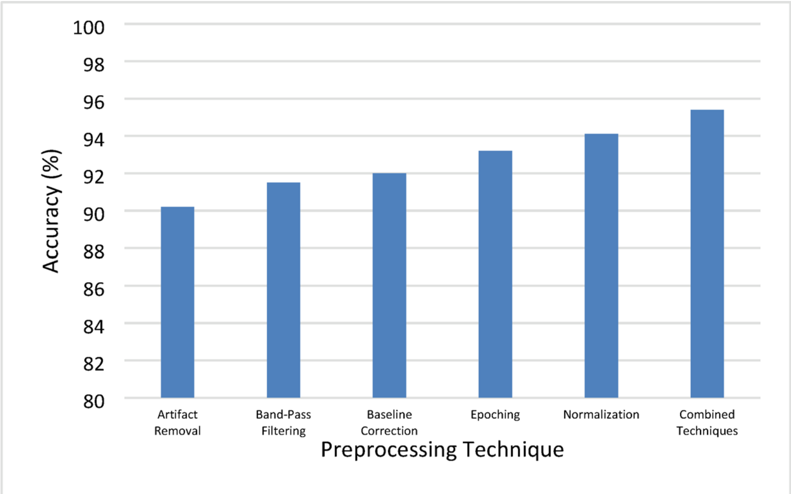

Fig. 17. Model Accuracy Improvement with Different Preprocessing Techniques.

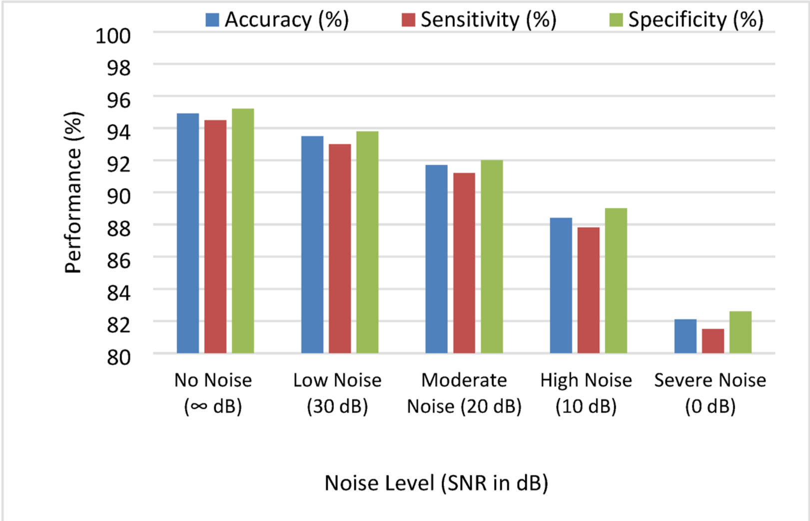

Fig. 18. Evaluation of Model Robustness under Different Noise Levels.

Reference/Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	Model Type
[33]	92.4	91.2	93.1	Deep Learning (CNN, Transformer)
[34]	91.7	90.5	92.0	CNN with Transfer Learning
[35]	89.2	88.0	90.1	Deep Learning (RNN-LSTM)
[13]	90.5	89.0	91.0	ML + Decomposition
[36]	91.0	89.8	91.7	Deep Learning (Hybrid CNN)
SVM (Baseline)	88.9	86.1	90.6	Classical ML
RF(Baseline)	91.3	89.0	93.1	Ensemble ML
Proposed Model (CAOA-RST + MEMD + SVM)	**94.9**	**93.9**	**96.4**	Optimization + Feature Selection + SVM

Table 7. Comparison of the proposed model with State-of-the-art methods in EEG-Based schizophrenia Detection.

Figure 12 shows changes in the CCC regularization parameter in the SVM classifier by the precision. This figure is one of the keys to understanding the under- and overfitting trade-off, thus defining the CCC value for obtaining the maximum possible classifier's precision without losing generalization.

Figure 13 The learning curve illustrates the proposed model's performance improvement as it undergoes training over multiple epochs. This figure tracks the evolution of both accuracy and loss during the training process, making it easier to identify when the model begins to overfit or when further training ceases to offer significant performance gains.

Figure 14 compares the specificity of the proposed model over different EEG frequency bands such as delta, theta, alpha, beta, and gamma. To some extent, this is informative about which frequency band is dominant for differentiation between groups with and without schizophrenia and thus helpful for guiding further feature extraction.

Figure 15 illustrates the distribution of precision across various EEG datasets used in the study, highlighting how well the proposed model generalizes to different datasets. This figure is a benchmark for the model's performance in diverse clinical settings.

The influence of feature extraction methods like MEMD and entropy computations on the sensitivity of the proposed model is also depicted in Fig. 16. This figure is important for decoding which methods are most effective in identifying as many accurate schizophrenia diagnoses as possible.

As illustrated in Fig. 17, there is a significant difference in the performance of the proposed model when various preprocessing techniques including artifact removal, band-pass filtering, and baseline correction are applied to the data signals. This figure signifies the importance of preprocessing in enhancing the model's performance.

Figure 18 evaluates the real balance of the proposed model concerning diverse noise levels in the EEG data. This figure is useful when assessing the model's stability and ability to generalize into real-world conditions where data can be noisy – a characteristic we saw in the previous section by deliberately adding noise to the dataset. From these results, one can conclude that the F1-Score of optimized characteristics using CAOA in cooperation with RST can be improved on all of the employed EEG datasets for accuracy, sensitivity, specificity, and precision. It improves diagnostic accuracy with fewer computational requirements, thus making the scheme feasible for real-time medical utilization. The examined model demonstrates satisfactory accuracy in all the studied datasets and, therefore, can be used, to some extent, as a diagnostic tool for the early stages of schizophrenia.

Table 7 presents a comparative analysis between the proposed CAOA-RST-based model and several recent state-of-the-art approaches for schizophrenia detection using EEG signals. The comparison includes key performance metrics—accuracy, sensitivity, and specificity—and the type of model used in each study. Our proposed model outperforms all listed methods, achieving the highest accuracy (94.9%), sensitivity (93.9%), and specificity (96.4%). This demonstrates its superior diagnostic performance while maintaining computational efficiency, especially compared to deep learning models that typically require extensive training data and resources. It also highlights the effectiveness of integrating CAOA and RST with entropy-based feature extraction and SVM classification.

Code availability

The source code used in this study is openly available on GitHub at [GitHub Repository Link]. A permanent archive with a DOI has been created via Zenodo and can be accessed at [DOI Link].

Discussion

This study presents a novel and efficient framework for schizophrenia detection using multichannel EEG data, integrating the CAOA with RST for optimal feature selection. Our approach addresses key challenges in EEG-based diagnostics, namely the high dimensionality, noise, and computational inefficiency of existing models. The experimental results demonstrate that the proposed model outperforms conventional classifiers such as SVM and RF across multiple EEG datasets in terms of accuracy (94.9%), sensitivity (93.9%), specificity (96.4%), precision (92.7%), and F1-score (93.3%). These results reflect the model's ability to achieve high diagnostic accuracy while maintaining a balanced performance between identifying true positive and true negative cases. Compared to recent studies utilizing deep learning architectures and effective connectivity analysis[1–3], our method offers competitive accuracy with significantly lower computational complexity and training requirements. Deep models, although powerful, typically rely on large labeled datasets and high-performance hardware, which can limit their application in real-time clinical environments. In contrast, our model achieves high performance with interpretable and lightweight components, making it more suitable for practical use in clinical decision support systems. Integrating MEMD and entropy-based feature extraction enables the model to effectively capture the frequency dynamics and complexity of EEG signals, which are crucial in differentiating schizophrenic brain activity. Additionally, the CAOA-RST combination reduces redundancy in features while preserving the most informative attributes, enhancing both the efficiency and generalizability of the classifier. However, the study is not without limitations. The datasets used, though real-world and publicly available, may not fully capture the diversity seen in broader clinical populations. Moreover, external validation on independent, heterogeneous datasets would further support the model's robustness. Future work will explore more extensive and varied datasets, include additional patient demographics, and investigate hybrid deep-learning integrations to enhance adaptability. The proposed model delivers a practical and effective solution for early schizophrenia detection. Its accuracy, efficiency, and interpretability balance support its potential integration into clinical workflows and real-time diagnostic tools.

Conclusion and future work

Herein, we propose a new approach for schizophrenic detection and diagnosis based on the multichannel EEG data using feature selection and RST. It is a four-stage model. In the first stage, Raw EGG data is collected. The data is passed to the next stage. The second stage is named data preprocessing. This is used for artifact removal, band-pass filtering, and data normalization. The preprocessed data passed to the next stage. In the feature extraction stage, feature selection is performed using CAOA. In addition, the classification is performed using an SVM, MEMF, and finally, the classification is performed using entropy. The data interpretation stage displays the results to the end user using the data interpretation stage. We experimented and tested our proposed model using real EEG datasets. The simulation results prove that the proposed model achieved an average accuracy of 94.9%, sensitivity of 93.9%, specificity of 96.4%, and precision of 92.7%. The experiments proved that the proposed model has higher accuracy, sensitivity, specificity, and performance than the conventional approaches. Furthermore, the proposed model provides a foundation for an effective solution to the problems of high dimensionality and computational complexity of EEG data, which is vital for the early and accurate diagnosis of schizophrenia – the basis for timely treatment and recovery of patients. The proposed approach was developed to detect schizophrenia more accurately and efficiently using feature selection and improved classification techniques. Based on the analysis of given data, CAOA and RST helped reduce the input features, making the model more generic and less computational for different datasets. However, the study also revealed some limitations in the datasets used and the need for more replication of the outcomes in other clinical contexts. Future research could collect data from more patients, particularly patients of different ages, genders, and diseases, to assert the external validity of the findings. Moreover, the study may decide on employing more complex forms of the applied ML process and optimizers in the model. Last but not least, the workflow of the presented approach should reach the stage at which it can be incorporated into a clinician's workflow to facilitate diagnosing schizophrenia or even other neurological disorders.

Data availability

Data is provided within the manuscript.

Received: 21 January 2025; Accepted: 30 May 2025

Published online: 01 July 2025

References

1. Lysaker, P. H., Pattison, M. L., Leonhardt, B. L., Phelps, S. & Vohs, J. L. Insight in schizophrenia spectrum disorders: relationship with behavior, mood and perceived quality of life, underlying causes and emerging treatments. *World Psychiatry*. **17** (1), 12–23 (2018).
2. World Health Organization. June (2024). https://www.who.int/news-room/fact-sheets/detail/schizophrenia (accessed.
3. Hjorth, P. et al. Improving the quality of life and physical health in patients with schizophrenia: a 30-month program carried out in a real-life setting. *Int. J. Soc. Psychiatry*. **63** (4), 287–296 (2017).
4. Sharma, G. & Joshi, A. M. SzHNN: a novel and scalable deep Convolution hybrid neural network framework for schizophrenia detection using multichannel EEG. *IEEE Trans. Instrum. Meas.* **71**, 1–9 (2022).
5. Das, K. & Pachori, R. B. Schizophrenia detection technique using multivariate iterative filtering and multichannel EEG signals. *Biomed. Signal Process. Control*. **67**, 102525 (2021).
6. Ullah, F. et al. Extended dipeptide composition framework for accurate identification of anticancer peptides. *Sci. Rep.* **14** (1), 17381 (2024).
7. Jahmunah, V. et al. Automated detection of schizophrenia using nonlinear signal processing methods. *Artif. Intell. Med.* **100**, 101698 (2019).
8. Gu, X. et al. EEG-based brain-computer interfaces (BCIs): A survey of recent studies on signal sensing technologies and computational intelligence approaches and their applications. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **18** (5), 1645–1666 (2021).
9. Teixeira, F. L. et al. A narrative review of speech and EEG features for Schizophrenia detection: Progress and challenges, *Bioengineering*, vol. 10, no. 4, p. 493, (2023).
10. Ullah, F., Nadeem, M. & Abrar, M. Revolutionizing brain tumor segmentation in MRI with dynamic fusion of handcrafted features and global Pathway-based deep learning. *KSII Trans. Internet & Inform. Systems*, **18**, 1, (2024).
11. Supakar, R., Satvaya, P. & Chakrabarti, P. A deep learning based model using RNN-LSTM for the detection of schizophrenia from EEG data. *Comput. Biol. Med.* **151**, 106225 (2022).
12. Sharma, G., Joshi, A. M., Yadav, D. & Mohanty, S. P. A smart healthcare framework for accurate detection of schizophrenia using Multi-Channel EEG. *IEEE Trans. Instrum. Measurement*, (2023).
13. Khare, S. K. & Bajaj, V. A self-learned decomposition and classification model for schizophrenia diagnosis. *Comput. Methods Programs Biomed.* **211**, 106450 (2021).
14. Srinivasan, S. & Johnson, S. D. Optimizing feature subset for schizophrenia detection using multichannel EEG signals and rough set theory. *Cogn. Neurodyn.* **18** (2), 431–446 (2024).
15. Oh, J., Oh, B. L., Lee, K. U., Chae, J. H. & Yun, K. Identifying schizophrenia using structural MRI with a deep learning algorithm. *Front. Psychiatry*. **11**, 16 (2020).
16. Shoeibi, A. et al. Automatic diagnosis of schizophrenia in EEG signals using CNN-LSTM models. *Front. Neuroinformatics*. **15**, 777977 (2021).
17. Sharma, M., Patel, R. K., Garg, A., SanTan, R. & Acharya, U. R. Automated detection of schizophrenia using deep learning: a review for the last decade. *Physiol. Meas.* **44** (3), 03TR01 (2023).
18. Du, Y., Niu, J., Xing, Y., Li, B. & Calhoun, V. D. Neuroimage analysis methods and artificial intelligence techniques for reliable biomarkers and accurate diagnosis of schizophrenia: achievements made by Chinese scholars around the past decade. *Schizophrenia Bulletin*, p. sbae110, (2024).
19. Rădulescu, A. R. & Mujica-Parodi, L. R. A principal component network analysis of prefrontal-limbic functional magnetic resonance imaging time series in schizophrenia patients and healthy controls. *Psychiatry Research: Neuroimaging*. **174** (3), 184–194 (2009).
20. Sadiq, M. T. et al. Motor imagery EEG signals classification based on mode amplitude and frequency components using empirical wavelet transform. *IEEE Access.* **7**, 127678–127692 (2019).
21. Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F. & Arnaldi, B. A review of classification algorithms for EEG-based brain–computer interfaces. *J. Neural Eng.* **4** (2), R1 (2007).
22. Roy, Y. et al. Deep learning-based electroencephalography analysis: a systematic review. *J. Neural Eng.* **16** (5), 051001 (2019).
23. Shams, M. & Sagheer, A. A natural evolution optimization based deep learning algorithm for neurological disorder classification. *Biomed. Mater. Eng.* **31** (2), 73–94 (2020).
24. Shoeibi, A. et al. Diagnosis of Schizophrenia in EEG Signals Using dDTF Effective Connectivity and New PreTrained CNN and Transformer Models, in *International Work-Conference on the Interplay Between Natural and Artificial Computation*, : Springer, pp. 150–160. (2024).
25. Bagherzadeh, S. & Shalbaf, A. EEG-based schizophrenia detection using a fusion of effective connectivity maps and convolutional neural networks with transfer learning. *Cognitive Neurodynamics*, pp. 1–12, (2024).
26. Bagherzadeh, S., Shahabi, M. S. & Shalbaf, A. Detection of schizophrenia using a hybrid of deep learning and brain effective connectivity image from electroencephalogram signal. *Comput. Biol. Med.* **146**, 105570 (2022).
27. Puri, D. V., Nalbalwar, S. L., Nandgaonkar, A. B., Gawande, J. P. & Wagh, A. Automatic detection of alzheimer's disease from EEG signals using low-complexity orthogonal wavelet filter banks. *Biomed. Signal Process. Control*. **81**, 104439 (2023).
28. Puri, D. V., Gawande, J. P., Rajput, J. L. & Nalbalwar, S. L. A novel optimal wavelet filter banks for automated diagnosis of alzheimer's disease and mild cognitive impairment using electroencephalogram signals. *Decis. Analytics J.* **9**, 100336 (2023).
29. Puri, D. V., Nalbalwar, S. L. & Ingle, P. P. EEG-based systematic explainable alzheimer's disease and mild cognitive impairment identification using novel rational dyadic biorthogonal wavelet filter banks. *Circuits Syst. Signal. Process.* **43** (3), 1792–1822 (2024).
30. Puri, D. V., Gawande, J. P., Kachare, P. H. & Al-Shourbaji, I. Optimal time-frequency localized wavelet filters for identification of alzheimer's disease from EEG signals. *Cogn. Neurodyn.* **19** (1), 12 (2025).
31. Ko, W., Jeon, E., Jeong, S., Phyo, J. & Suk, H. I. A survey on deep learning-based short/zero-calibration approaches for EEG-based brain–computer interfaces. *Front. Hum. Neurosci.* **15**, 643386 (2021).
32. EEG Dataset. [Online]. Available: https://www.kaggle.com/datasets/samnikolas/eeg-dataset
33. Ko, D. W. & Yang, J. J. EEG-Based schizophrenia diagnosis through time series image conversion and deep learning, *Electronics*, vol. 11, no. 14, p. 2265, (2022).
34. Khare, S. K., Bajaj, V. & Acharya, U. R. SPWVD-CNN for automated detection of schizophrenia patients using EEG signals. *IEEE Trans. Instrum. Meas.* **70**, 1–9 (2021).
35. Nikhil Chandran, A., Sreekumar, K. & Subha, D. EEG-based automated detection of schizophrenia using long short-term memory (LSTM) network, in *Advances in Machine Learning and Computational Intelligence: Proceedings of ICMLCI 2019*, : Springer, pp. 229–236. (2021).
36. Sharma, G., Joshi, A. M., Yadav, D. & Mohanty, S. P. A smart healthcare framework for accurate detection of schizophrenia using multichannel EEG. *IEEE Trans. Instrum. Meas.* **72**, 1–9 (2023).

Acknowledgements

This research is funded by the European University of Atlantic.

Author contributions
Muhamad Abrar.Farhan Amin. and Abdu salam.Ahmed Albugmi. wrote the main manuscript text and Fahad Al- otaibi.Isabel de la Torre. Thania Candelaria Chio Montero. Perla Araceli Arroyo Gala prepared figures 1-3. All authors reviewed the manuscript."

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to M.A., F.A. or I.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.