Mesopotamian Journal of Big Data
Vol. **(2025)**, 2025, **pp**. 61–78
DOI: https://doi.org/10.58496/MJBD/2025/005 ; ISSN: 2958-6453
https://mesopotamian.press/journals/index.php/BigData

Research Article

Coupling a dual tree - complex wavelet transform with K-means to clustering the epileptic seizures in EEG signals

Raid Lafta[1,*], Rasha Hallem Razzaq [2]

[1] *Department of Computer Engineering, Engineering college, Alshatra university, IRAQ.*
[2] *Directorate of Education, Karkh II, Ministry of Education, Baghdad 10011, Iraq.*

ARTICLEINFO

Article History

Received 12 Mar 2025
Revised 18 May 2025
Accepted 18 Jun 2025
Published 29 Jun 2025

Keywords

Epileptic EEG signals
Dual-tree-complex
wavelet transform
Statistical features
K-means clustering
ChaCha20

ABSTRACT

Electroencephalography (EEG) signals are routinely recorded in clinical settings for the diagnosis of epilepsy, that is, brain electrical disorders, by a neurologist. Nevertheless, both the reliability and safety of GB analysis of EEG parameters are not satisfactory. Therefore, identifying the effectiveness of EEG for diagnosis is a considerable challenge for hospitals. This study was conducted to improve the efficiency of detecting epileptic seizures in EEG signals by combining three methodologies: the double tree complex wavelet transform (DT-CWT), K means clustering, and the ChaCha20 encryption algorithm. EEG segments are initially partitioned into regular segments, each of which is further partitioned into smaller clusters by K-means. To analyse the EEG waves to extract frequency information and select six discriminable statistical features, these clusters are examined via the DT-CWT. After feature extraction, epileptic seizures are discerned on the basis of K-means, which can enable very accurate detection of the seizures. The final results are encrypted via the ChaCha20 standard to ensure patient data confidentiality at the send and receipt stages. The findings presented in this study show that the proposed approach has a good clustering accuracy of 99% to help doctors diagnose patients with epilepsy and prescribe the best treatment to cure patients to maintain privacy and prevent data from being seen by unauthorized persons with an overall accuracy of 96.3%. Through the improvement of the accuracy of neurological disease identification, this method opens up the possibility for further progress in the domain of EEG signal analysis.

1. INTRODUCTION

Epilepsy continues to be one of the most common brain diseases affecting millions of people, and rapid diagnosis is needed worldwide [1]. Seizure detection via EEG data is an important tool in clinical applications as it captures dynamic information of the brain [2–5]. Unfortunately, conventional approaches for the interpretation of EEGs face issues of interpatient variability, signal complexity and manual feature extraction, impeding the reproducibility of detection performance [6]. Research has also integrated of machine learning and deep learning models for more accurate classification and automatic seizure detection [7-12].

Although progress has been made in the detection of seizures, several challenges remain. Feature extraction methods such as wavelet-based analysis and spectral decomposition have computational efficiency problems [13-27]. Furthermore, advanced deep-learning techniques (CNNs, Bi-LSTM, and transformers) increase classification accuracy, but such techniques demand a large amount of EEG data and a considerable computational power for real-time applications [28-39]. A further critical issue to overcome is related to data protection because EEGs are stored and transmitted through telehealth systems, which call for sophisticated encryption techniques [40-42].

In the last decade, subband based feature extraction and classification methods have used wavelet transforms, clustering and neural networks. ITD, DWT, PSR achieve good seizure classification performance but computational inefficiency [43-47]. In addition to deep learning feature fusion, the robustness was further enhanced but optimization was necessary for the real-time application [48]. The tunable Q-factor wavelet transform (TQWT) was on the other hand proposed for

*Corresponding author. Email: raidluaibi.lafta@shu.edu.iq.

the detection of seizures and it offered multiresolution analysis, however, at the expense of a heavy computational burden [49–55].

The proposed work presents an integrated framework that combines dual-tree complex wavelet transform (DT-CWT) for improved feature extraction, optimized K-means clustering for effective classification, and ChaCha20 encryption for securing the transport of the EEG input. Through multilevel signal decomposition, clustering enhancements, and lightweight cryptographic methodologies, the presented solution provides a novel, efficient and scalable solution to address the key problems of seizure detection. To overcome these difficulties, the following directions could be adopted in this paper:

- ✓ The feature extraction process improved via dual-tree complex wavelet transform (DT-CWT) to increase the temporal and spectral resolution for the detection of epileptic seizures.
- ✓ Classification accuracy elevation is achieved via advanced K-means clustering for robust and adaptive grouping of EEG features.
- ✓ With the adoption of ChaCha20 encryption, which allows EEG signal processing efficiently and securely in real-time applications, data security and privacy protection can be enhanced.

It proposes a detailed framework built on multiresolution wavelet transformation-based features, and improved clustering techniques as well as advanced cryptography, for the most commonly used application (medical diagnostics), which guarantees both accuracy, and efficiency as well as the application of security.

2. RELATED WORKS

The detection of epileptic seizures has evolved greatly, moving from classical feature extraction methods to machine learning and deep learning approaches. Vanabelle et al. [5] presented an extreme gradient boosting framework whose results significantly surpassed those of classical methods for the seizure classification problem. Similarly, Lafta et al. [6] proposed a multidisease recommendation model for telehealth; however, while it incorporates real-time EEG-based epilepsy seizure detection, scalability was the main issue. Similarly, Zeng et al. [7] studied different decomposition techniques for EEG signals with the use of the ITD, DWT and PSR, and NN to obtain higher detection accuracy.

Malekzadeh et al. [8] used hand-crafted and deep features, which solved some issues related to feature instability, but there were some problems with real-time computational complexity. Deep learning methods have greatly improved the classification of seizures. Wang et al. [9] introduced 1D CNNs for seizure onset detection on long-term EEG recordings Current Limitations Their approach has been able to increase the reliability of detection, although it demands high levels of computational resources for large-scale analysis.

Shoeibi et al. [12] proposed a hybrid of ANFIS classifiers, autoencoder, and fuzzy entropies that resulted in successful seizure detection. However, their approach encounters the high variability in EEG signals, causing the classification performance to be unstable across datasets. Lafta et al. [13] used time series clustering for heart disease and its clustering mechanism was not adaptable for EEG seizure classification which can only be used to detect neurological disorders.

Wavelet transformation methods are commonly used in processing EEG signals. Zhen et al. [14] increased the accuracy of seizure detection via the Revised Tunable Q-Factor Wavelet Transform, but encountered difficulty in real-time EEG processing because of its high computational power.

Ahmed et al. [15] used dual tree complex wavelet transform (DT-CWT) with machine learning; however, they developed a model for not seizures detectable but rather a schizophrenia classifier.

A few studies have investigated the classification of seizures via machine learning. Kavitha et al. [16] used the wavelet domain feature extraction and different classification methods, but worked poorly at generalizing over multiple sets of EEG data. Shukla et al. [17] exploited continuous wavelet transform and deep neural networks, which presented higher accuracy, but they were confronted with the problem of feature instability with the existence of signal artefacts.

Deep learning techniques have significantly improved classification accuracy.

Chashmi [18] applied nonlinear DT-CWT-based feature extraction integrated with deep learning, successfully enhancing seizure detection performance. However, their study did not fully address the computational complexity in real-time scenarios. Al-Salman et al. [19] investigated EEG waveform classification; but did not optimize their approach for handling noise interference, which affects classification reliability.

A comprehensive review of machine learning algorithms applied to published EEG datasets was conducted by Miltiadous et al. [20], who highlighted the strengths and limitations in various classification approaches. However, their findings indicated that clustering-based optimization was still underdeveloped. Pansani & Itikawa [21] integrated convolutional neural networks (CNNs) with continuous wavelet transform, improving feature extraction but facing overfitting issues with limited dataset sizes.

In addition to seizure detection, researchers have explored biometric authentication. Singh & Tiwari [22] proposed a dual multimodal biometric authentication system based on a WOA-ANN and SSA-DBN; however, their focus was outside of seizure detection, and direct EEG application was lacking. Fredes et al. [35] leveraged wavelet-based analysis to refine seizure detection, but the lack of dynamic thresholding limited its robustness across varying patient conditions. Karim et al. [36] improved clustering performance via the enhanced gap statistic in K-means; but struggled with accurate cluster determination for highly variable EEG data.

In the context of security, encryption techniques have been increasingly incorporated into EEG data processing. Fadhil et al. [40] integrated ChaCha20 encryption with the Laplacian of Gaussian filtering, strengthening EEG data confidentiality. Najm et al. [41] conducted a comparative study between the ChaCha20 and Serpentine encryption algorithms, and evaluated their efficiency in lightweight encryption scenarios. Furthermore, Muhammed et al. [42] introduced a hybrid ChaCha20-ECDH encryption model, ensuring secure cloud-based EEG transmissions.

Modern deep learning architectures have further refined seizure classification. Zhao [43] proposed a novel deep neural network tailored for robust seizure detection, whereas Jiwani [44] developed an LSTM-CNN hybrid model, that leverages sequential dependencies for more precise classification. Tuncer et al. [45] expanded on this by implementing a bidirectional LSTM (Bi-LSTM) framework, which demonstrated enhanced temporal feature extraction in EEG signals.

Machine learning applications have also gained prominence in seizure prediction. Tran et al. [46] explored multiple AI-driven techniques for seizure detection, emphasizing model adaptability. Wang et al. [47] combined Symlet wavelet processing, a gradient boosting machine, and a grid search optimizer, which achieved high classification reliability. More recently, Zaid et al. [48] focused on preprocessed and combined EEG datasets, improving deep learning model generalizability.

Several studies have highlighted wavelet-based approaches for optimizing EEG feature extraction. Slimen et al. [50] applied the dual-tree complex wavelet transform (DT-CWT) combined with machine learning classifiers to improve seizure detection rates.

In addition to classification models, researchers have explored self-supervised learning and attention mechanisms. Yuan et al. [55] introduced a hybrid dense net–ViT framework with attention fusion, which significantly improved seizure prediction accuracy. Xiao et al. [56] employed self-supervised learning, highlighting the adaptability of attention-based deep learning models in biomedical signal analysis.

Despite these advancements, existing approaches face challenges in optimizing seizure classification accuracy, reducing computational overhead, and ensuring robust EEG data security. This study builds upon prior works by integrating DT-CWT-based feature extraction, enhanced clustering methods (K-means optimization), and ChaCha20 encryption, addressing critical gaps in feature robustness, classification reliability, and data security.

3. METHODOLOGY

This section elaborates on the methodology of the proposed approach.

3.1 Dataset Description

The University of Bonn EEG database is one of the most popular public databases used in brain signal analysis research, particularly in the field of epileptic seizure detection.

The University of Bonn dataset was used in this research. This dataset serves as a benchmark for most studies published in the literature. Additional information about this dataset can be found in [11]. The international 10–20 system that Bonn University uses was employed to collect the EEG readings. The 10–20 system and the placements of the surface electrodes used to record the EEG signals are shown in Fig. 1. Five sets of recordings have been made, such as A, B, C, D, and E. The surface EEG recordings are the source of sets A and B. While sets C, D, and E, were taken from five epileptic patients, these sets were taken from five healthy volunteers who were awake. Set E shows ictal activity (seizures), whereas sets C and D were recorded while the five epileptic patients experienced seizure-free episodes. All of the sets A, B, C, D, and E, are used in this study to assess the suggested methodology. Five sets of epileptics for the EEG data are shown in Fig. 2.

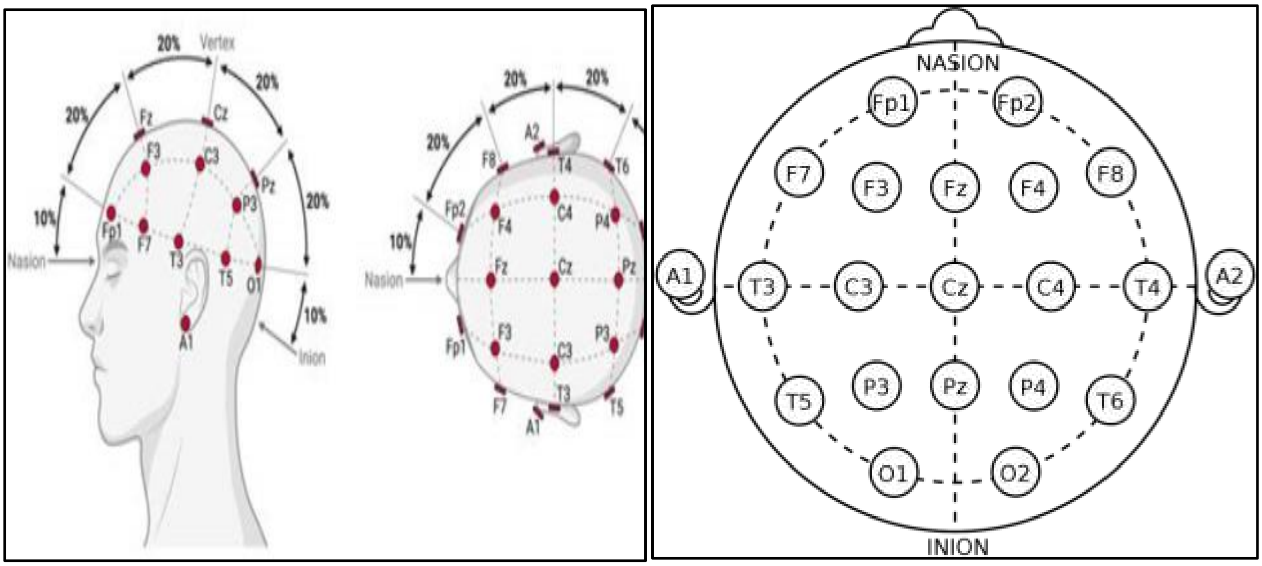

Fig. 1. The 10-20 system of electrode placement for recording an EEG pattern

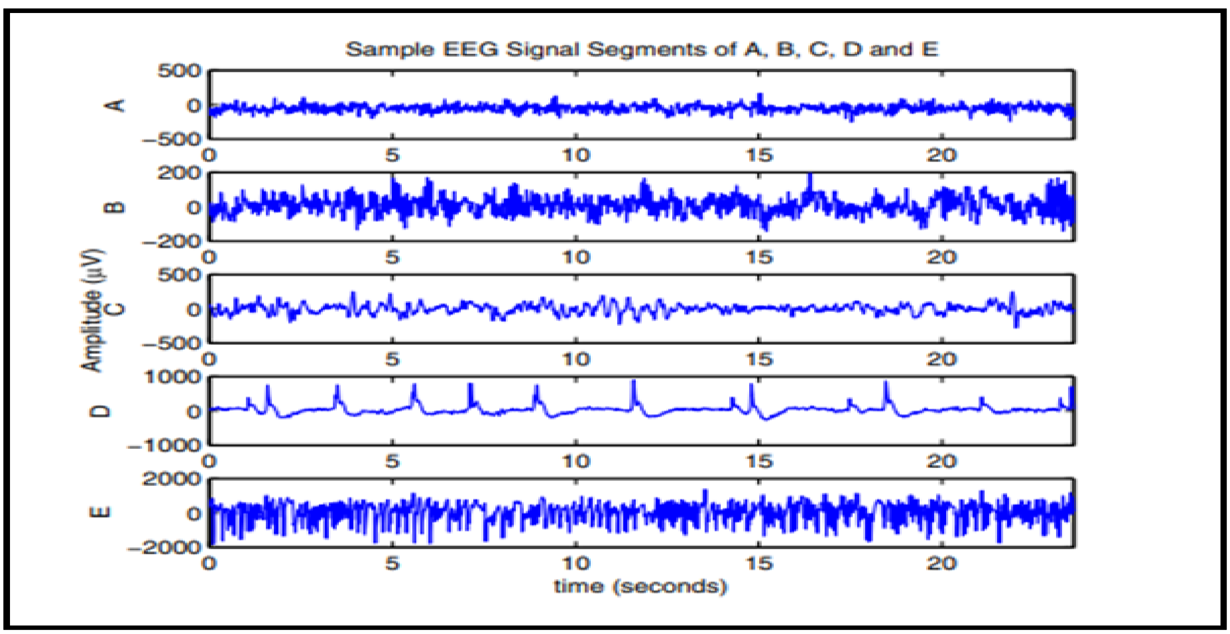

Fig. 2. EEG samples from sets A, B, C, D and E

3.2 Proposed Framework

In this paper, we present a new approach for clustering epileptic seizures in EEG signals. Step 1 is segmentation, as input we have a set of EEG signals composed of five sets (A, B, C, D and E). The EEG time signals are segmented into smaller-time intervals (epochs) for accurate processing and analysis of the relevant features in the epoch. The EEG data are the EEG signals at regular intervals that are ready for analysis. The second phase consists of analysis and segmentation with the use of the DT-CWT (decomposition and features extraction). The input is EEG signals in segmented form. The signals are decomposed and the frequency content is analysed via the DT-CWT. This method can be used to calculate highly useful statistical data such as variance, mean, power, and standard deviation. In this step, the result is a set of statistical features of the EEG signals that serve as criteria to classify the pathologic states. Our next task is tK-means categorization. The input is the statistical properties of the EEG signal. The K-means algorithm is used to cluster the extracted features. The number of clusters is determined automatically by the system to ensure proper classification of epileptic conditions. What you're left with is classified data that tells you how patients suffering from epilepsy can be identified.

The last part of the proposed method is data encryption with ChaCha20 (Encryption). The inputs for this stage are patients' data labeled by the K-means outcome, which includes the chosen EEg signal features as well as the secret encryption key for the ChaCha20 algorithm. The encrypted data is then passed to the ChaCha20 algorithm, which are a very fast stream encryption algorithm used for preserving confidentiality. ChaCha20 uses a 256-bit key and a nonce* for encryption security and diversity. Patient data are transformed into an encrypted version, so that the original data will be unreadable without the correct key. The encoded information is then ready to be safely transmitted or stored protected from unauthorized use. The result is EEG data processed into a very high level of security by the ChaCha20 encryption algorithm. No sensitive patient information is exposed in the process of recording, transmission or storage, and the encrypted file can only be decrypted with the correct key, adding an extra layer of security protection for medical information. The importance of this step.

The ChaCha20 algorithm helps protect patient data from hacking or tampering, especially when they are shared in cloud systems or between different medical institutions. Compared with traditional encryption algorithms such as AES, ChaCha20 offers high processing speed and improved performance, making it a powerful choice for securing medical information.

Figure 3 illustrates the architecture of the proposed approach. The next section contains a set of experiments that were conducted with the datasets.

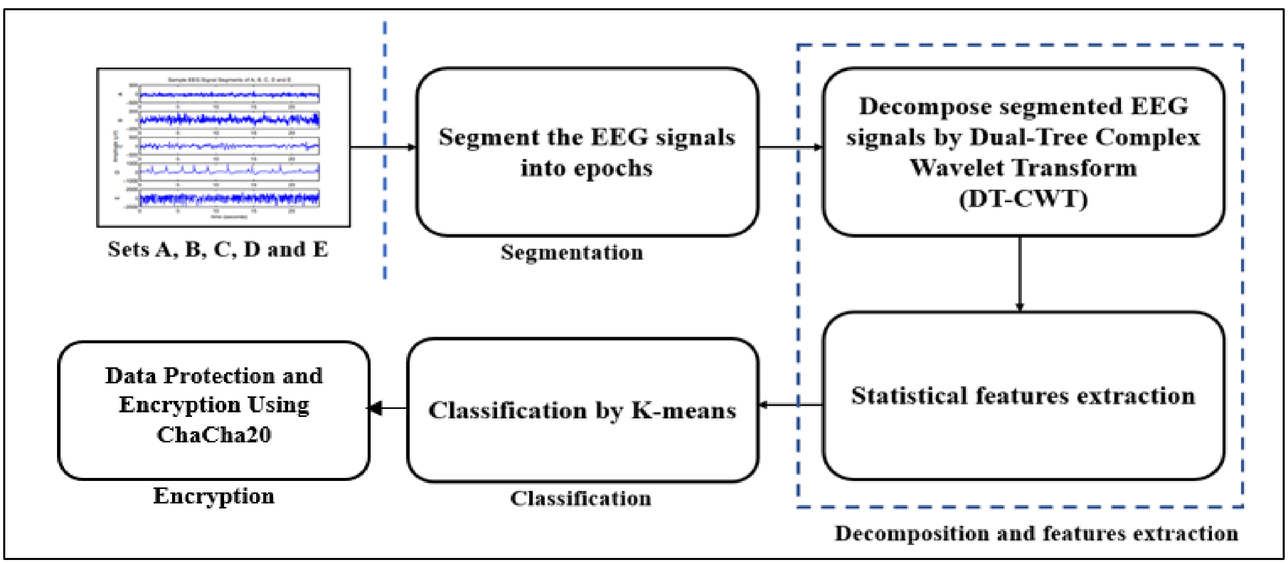

Fig. 3. Architecture of the proposed DT-CWT and K-means clustering method with ChaCha20 encryption.

3.3 EEG segmentation

Finding a more precise and efficient method to divide EEG signals into epochs and remove unnecessary data is crucial since EEG signals are nonstationary and contain redundant data. Each single channel EEG that contains on 4097 data points is partitioned into four segments with 1024 data points after one data point is removed. Each section must be completed within the specified time frame, which is 5.9 seconds. After that, 32 clusters are created from each signal segment. The number of clusters is objectively noted and chosen throughout the training process. Six statistical features are chosen and extracted from every single subband. XMax, XMin, XMean, XSD, XMed, and XRG are the designations for the six extracted features. The brief descriptions of the statistical features that were extracted are displayed in Table 1. While the mean and standard deviation are superior measures for data with a more skewed distribution, min and max are thought to be extremely good measures for time series data that have a symmetric distribution [36,37]. Fig. 4 shows the segmentation process of the EEG signals.

Table 1. Short explanations of the statistical features.

Feature name	Formula	Description
Maximum value	$\mathbf{X_{Max}}$ =Max $[x_n]$	Where X_n=1,2, 3, 4,…, n represents the data and N represents data point number and AM is the mean of samples
Minimum value	$\mathbf{X_{Min}}$ =Minx $[x_n]$	
Mean	$\mathbf{X_{Mean}} = \frac{1}{n}\sum_{1}^{n} x_i$	
Median	$\mathbf{X_{Med}} = (\frac{N+1}{2})^2$	
Standard Deviation	$\mathbf{X_{SD}} = \sqrt{\sum_{n=1}^{N}(x_n - AM)\frac{2}{n-1}}$	
Range	$\mathbf{X_{RG}} = \mathbf{X_{Max}} - \mathbf{X_{Min}}$	

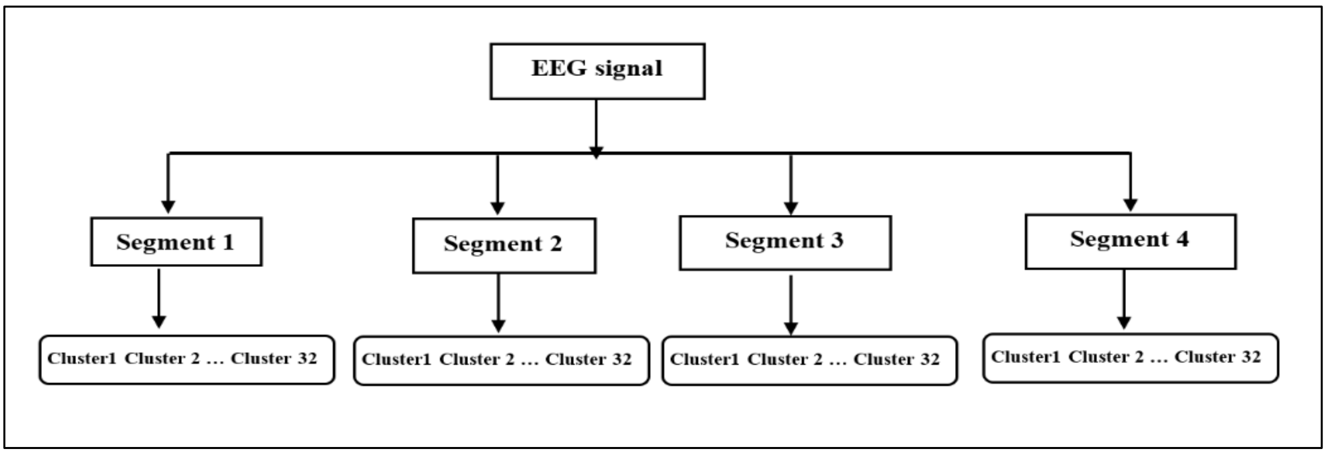

Fig. 4. shows the process of segmentation for each single EEG channel.

3.4 Feature Extraction

Feature extraction, the step in raw EEGs are transformed into the feature set to be later employed in classification and encoding, is shown in Figure 1. It is necessary to refine the estimation to increase accuracy, decrease noise, and lower the computational burden. In this study, feature extraction is performed via the dual-tree complex wavelet transform (DT-CWT), and the K-means clustering algorithm is employed to improve pattern clustering and the ability of the model to differentiate between normal and pathological signals.

To further extract the features from the time frequency space, the DT-CWT decomposes the EEG signal into the frequency channel at multiple scales, and preserves the phase information of the signal, which significantly improves the shift invariance and directional selectivity compared with the traditional wavelet transform. This decomposition leads to a more effective representation of the EEG signals and provides enough details for subtle variations to be distinguished which is important for the separation of epileptic seizures and normal activity within the brain.

After feature extraction, we apply K-means clustering to partition similar patterns so that a classification model for discriminating between seizure and nonseizure states can work efficiently in the model updating steps. The clusters are able to optimize the feature distribution and decrease the intraclass variance, to improve the classification performance. The right number of clusters (K) is found through statistical methodologies (e.g., the K-means gap statistic or elbow method) to provide robust PyG feature clustering.

3.5 Dual Tree Complex Wavelet Transform

Discrete wavelet transformation (DWT) is a powerful tool used in analysing data in time frequency domain. Although the discrete wavelet transformation provides efficient time frequency analysis for nonstationary data, it suffers from some problems related to aliasing, lack of directionality and shift variance. The limitations of the discrete wavelet transformation can be solved by using the DT-CWT which introduces a very powerful time frequency representation for different data. In general, the DT-CWT employs two real DWT trees. Figure 5 Illustrate the structure of the dule tree complex wavelet transformation. As shown in Figure 5, the imaginary part of the tree is represented at the bottom, and the complex wavelet coefficient is represented at the top. DT-CWT is used to decompose the EEG signals into set of subbands i.e., alpha, beta, delta, theta, and gamma for four levels of frequency decomposition. For each tree, five subband can be represented as y1, y2, y3, y4 and z4. Because the DT-CWT has two parts: real and imaginary, ten subbands (five subbands for each tree) can be obtained after four decompositions. (y1,1), (y2,2), (y2,1), (y2,2), (y3,1), (y3,2), (y4,1), (y4,2), (z4,1), and (z4,2) are some possible representations of the subbands. Thus, each pair represents the real or imaginary part; for example, (y1,1) represents the real part, whereas (y1,2) represents the imaginary part. The DT-CWT decomposes the EEG signals (clusters) to extract the frequency information for clustering the epileptic seizures in the EEG signals.

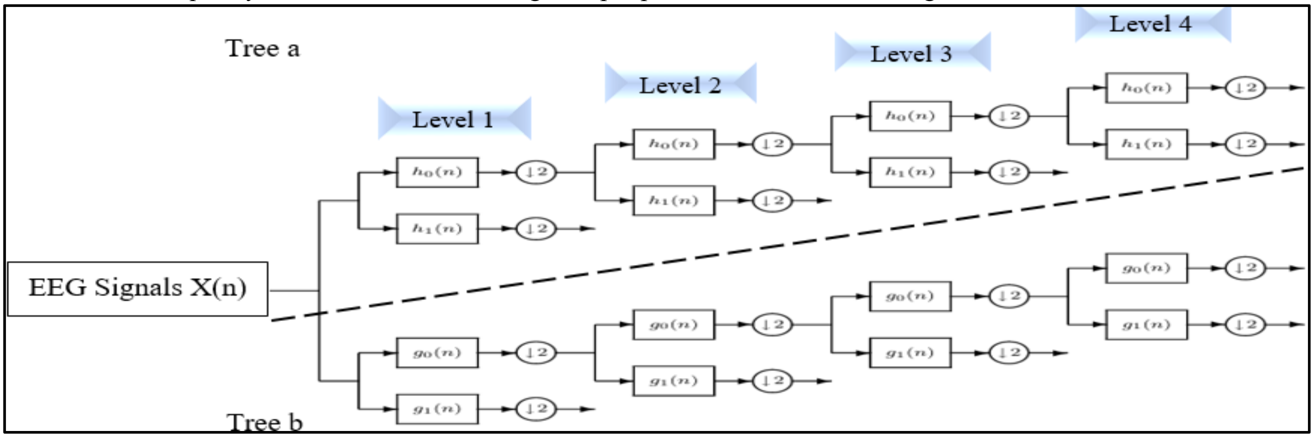

Fig. 5. 1D Dual tree complex wavelet transformation.

3.6 The unsupervised K-means clustering algorithm

K- means is an unsupervised robust algorithm used for large datasets. It works on dividing the data into k mutually exclusive clusters. One of its limitations is that the initial value of k should be more accurate for each spectral data [36]. Let X be n multidimensional data points and the division into K clusters. The Euclidean distance is utilized to obtain a similarity index and the target of clustering is to minimize the summation of the squares of various kinds as follows:

$$c = \sum_{k=1}^{k} \sum_{i=1}^{n} ||(x_i - u_k)||^2 \tag{1}$$

where K denotes the centers of K clusters $\boldsymbol{u_k}$ denotes the Kth center, and $\boldsymbol{x_i}$ denotes to the number of ith points in the dataset. Therefore, the following equation represents the solution of the centroid $\boldsymbol{u_k}$:

$$\begin{aligned} \frac{\partial}{\partial u_k} &= \frac{\partial}{\partial u_k} \sum_{k=1}^{k} \sum_{i=1}^{n} (x_i - u_k)^2 \\ &= \sum_{k=1}^{k} \sum_{i=1}^{n} \frac{\partial}{\partial u_k} (x_i - u_k)^2 \\ &= \sum_{i=1}^{n} 2(x_i - u_k) \end{aligned} \tag{2}$$

The final equation is as follows by setting equation 2 zero:

$$u_k = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{3}$$

The main idea for implementing the k-means algorithm is to randomly extract the sample points (K sample points) from the set of samples and detect them as the center of the initial cluster. Each sample point is divided into cluster depending on the nearest center point. After that, for each cluster, the center point for all samples is selected to be the center point of the cluster. These steps are repeated until the center point of the cluster does not change or until a specific number of iterations is reached.

3.7. ChaCha20 algorithm

ChaCha20 was developed by Daniel Bernstein, and it is a synchronous stream encryption algorithm based on a simple and effective encryption structure that is highly resistant to attacks because it uses a 256-bit key and a 96-bit initialization vector (IV), which makes it a suitable choice for securing sensitive data [40]. In this study, ChaCha20 was used to encrypt EEG data collected from epilepsy the patients. First, the EEG data were collected from the patients, which requires security and protection to maintain privacy. Then, a secret key and a random initialization vector were generated to ensure encryption security. The medical data were encrypted by generating a series of random zeros (Keystream) via the key and the vector (IV). An XOR operation was applied between the original data (Plaintext) and the random zeros (Keystream) to obtain the encrypted medical data (Ciphertext). For the possibility of decryption later, the encrypted data were stored with the IV. Algorithm 1, and Figure 6 illustrate the working steps of the ChaCha20 algorithm for the proposed approach.

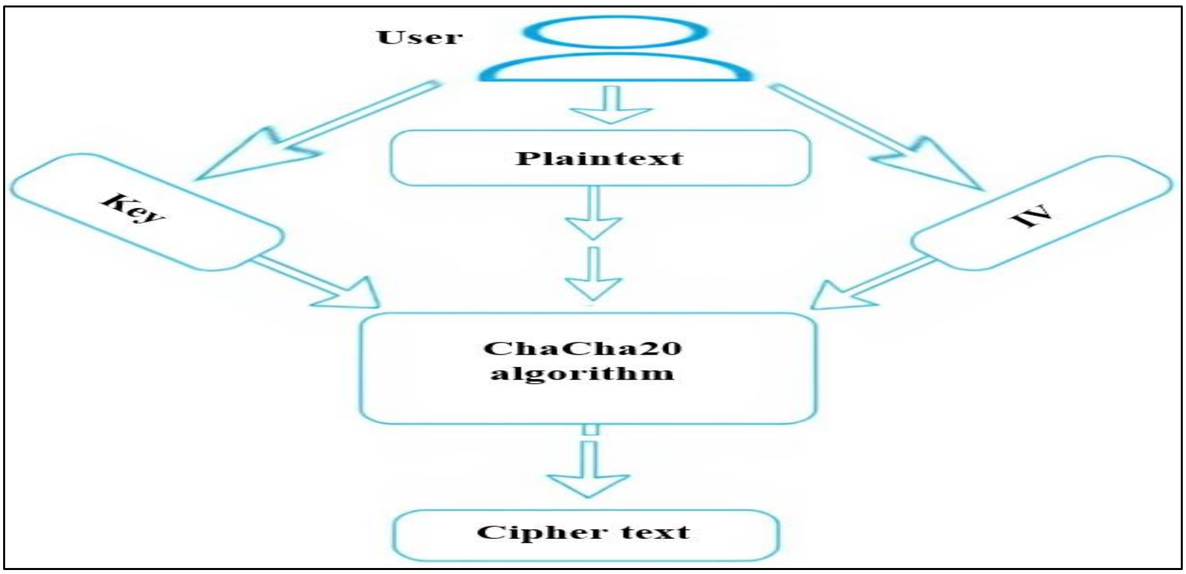

Fig. 6. Flowchart illustrating the encryption process of the results using ChaCha20 for the proposed approach

---

ALGORITHM 1: ChaCha20_Encryption

```
    INPUT: Key: 256-bit secret key. IV:  96-bit initialization vector. Plaintext: Data to be encrypted.
    OUTPUT: Cipher text: Encrypted data. IV: Initialization vector used for encryption
BEGIN
    Step 1: Initialize the state
    state ← Initialize ChaCha20 State (Key, IV)
    Step 2: Prepare to store cipher text
    Cipher text ← []
    Step 3: Process the plaintext in blocks
    FOR each block in Plaintext DO
        Step 3a: Generate keystream for the current block
        keystream ← Generate Key stream(state)
        Step 3b: Encrypt the block using XOR with keystream
        Encrypted _block ← XOR(Plaintext[block], keystream)
        Step 3c: Append the encrypted block to Cipher text
        Append encrypted _block to Cipher text
        Step 3d: Update the state for the next block
        State ← Update State(state)
    END FOR
    Step 4: Return the encrypted data and IV
    RETURN (Cipher text, IV)
END
```

3.8 Performance evaluation

The clustering performance is evaluated via a set of statistical measures as follows:
Accuracy: Accuracy is a common statistical measure used in evaluating the proposed methods. The mathematical formula is as follows:

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \times 100\% \tag{4}$$

Sensitivity: This is a statistical measure utilized to find the rate of actual positive clustering values as follows:

$$Sensitivity = \frac{TP}{TP+FN} \times 100\% \tag{5}$$

Specificity: This is a statistical measure used to find the rate of negative subjects that is correctly defined as follows:

$$Specificity = \frac{TN}{TN+FP} \times 100\% \tag{6}$$

where *TN* (the true negative) represents the genuine nonepileptic subject classified as a correctly nonepileptic subject, whereas *TP* (the true positive) refers to the actual epileptic segment that is correctly classified. Additionally, false negatives (FNs) denote the number of epileptic segments mistakenly classified as nonepileptic, and false positives (FPs) represent the number of epileptic patients misclassified [58,59].

4. EXPERIMENTAL RESULTS

A series of studies were used to test and evaluate the ability of the proposed method to cluster epileptic seizures in EEG signals. These will be described in the subsections of this section of the research.

4.1 Experimental Setup

A series of tests were conducted to assess the suggested method via various statistical feature sets that were taken from the EEG segments. According to the data, there is a positive relationship between the number of retrieved features and the effectiveness of the suggested strategy. Additionally, the approach's performance is assessed using all EEG datasets (A-E). MATLAB software is used for these studies on a desktop computer with an Intel ® Core i7 processor running at 3.40 GHz and 8.00 GB of RAM. The coding experiment was also conducted via Eclipse on a laptop computer equipped with an Intel® core i5 processor and the Ubuntu operating system.

4.2 Feature-based Clustering Results

The approach performance for clustering is evaluated in terms of accuracy, sensitivity, and specificity. To assess the suggested method for every EEG dataset, the retrieved statistical features were first tested independently. The performance of the suggested method, which employs various statistical features for every EEG dataset, is displayed in Figure 7 and Table 2.

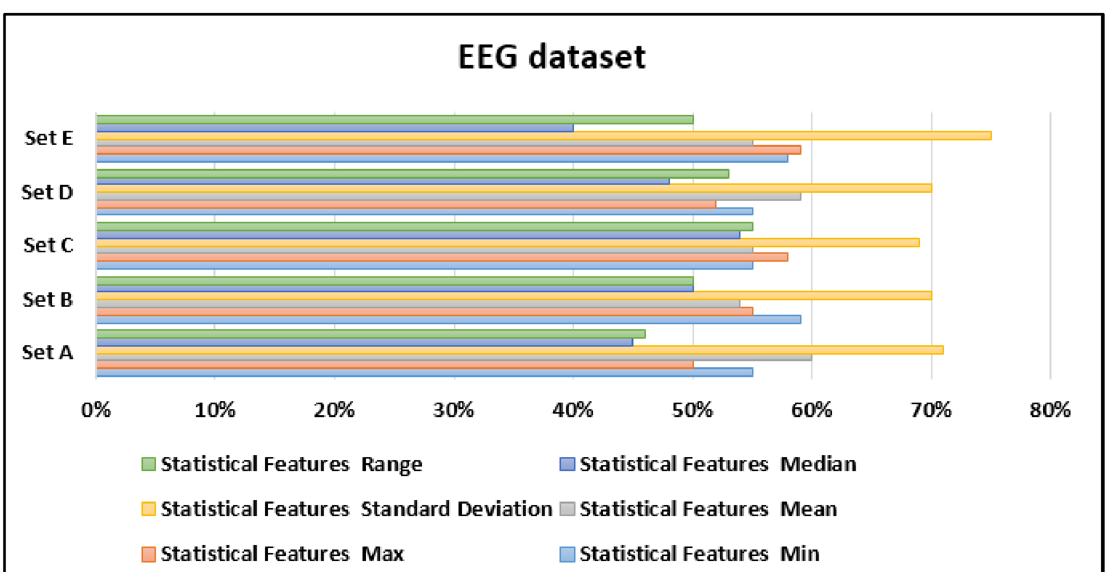

Fig. 7. Variation in Clustering Results Based on Extracted EEG Signal Features

TABLE II. THE PERFORMANCE OF THE APPROACH USING DIFFERENT FEATURES FOR ALL EEG SETS.

EEG dataset	Statistical Features					
	Min	Max	Mean	Standard Deviation	Median	Range
Set A	55%	50%	60%	71%	45%	46%
Set B	59%	55%	54%	70%	50%	50%
Set C	55%	58%	55%	69%	54%	55%
Set D	55%	52%	59%	70%	48%	53%
Set E	58%	59%	55%	75%	40%	50%

In the suggested method, the input vector for training the K-means algorithm was the statistical features that were removed from the subbands. The results indicate that the first four features e.g., Min, Max, Mean and Standard Deviation achieved higher accuracy rates than did the other two features. Therefore, the first four statistical feature sets with high accuracy rates are selected to evaluate the performance of the approach in the next experiments.

4.3 EEG combination results

Most studies in the literature have clustered epileptic EEG signals using unbalanced samples. Equal and unequal samples such as {(AB vs E), (AC vs E), (CD vs E), (ACD vs E), (ABCD vs E), and (AB)(CD) vs E} are used for clustering in this study.

Figures 8-11 illustrate the approach performance when all the clustering cases are used. The results, indicate that, the highest accuracy rate is achieved for clustering set A against set E as presented in Figure 8. The accuracy rate is 99% for clustering set A against set E, whereas it is 96%, 98% and 98% the clustering cases B, C, and D against E respectively.

Another experiment was conducted in this study using equal and unequal samples for the clustering cases. Figure 9. The effectiveness of the proposed approach is shown by the clustering cases of AB and AC against E. The obtained results show that equal and unequal numbers of samples for the clustering cases AB and AC against E reached 97% and 98% respectively.

Furthermore, the clustering cases CD and ABC against E were investigated and recorded in this study. High clustering rates of 97% and 98% are achieved for equal and unequal numbers of samples, respectively. Similarly, the clustering cases ABCD and (AB) (CD) achieved high clustering rates of 96% and 97% respectively.

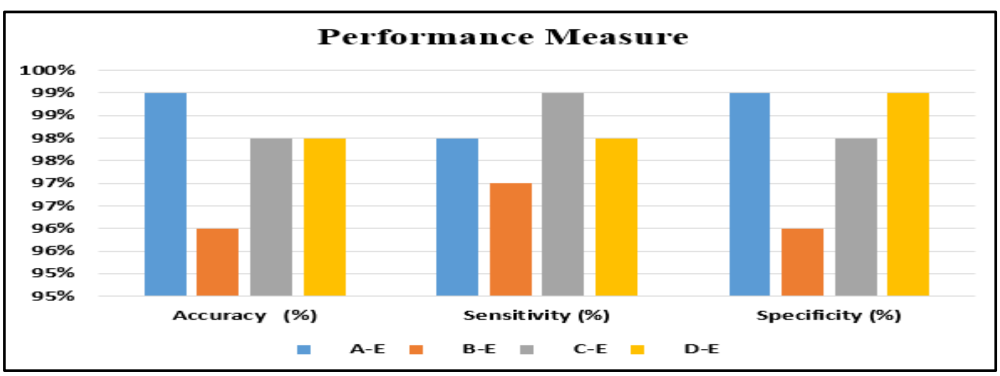

Fig. 8. The clustering accuracy of proposed approach using sets from sets A-D against set E.

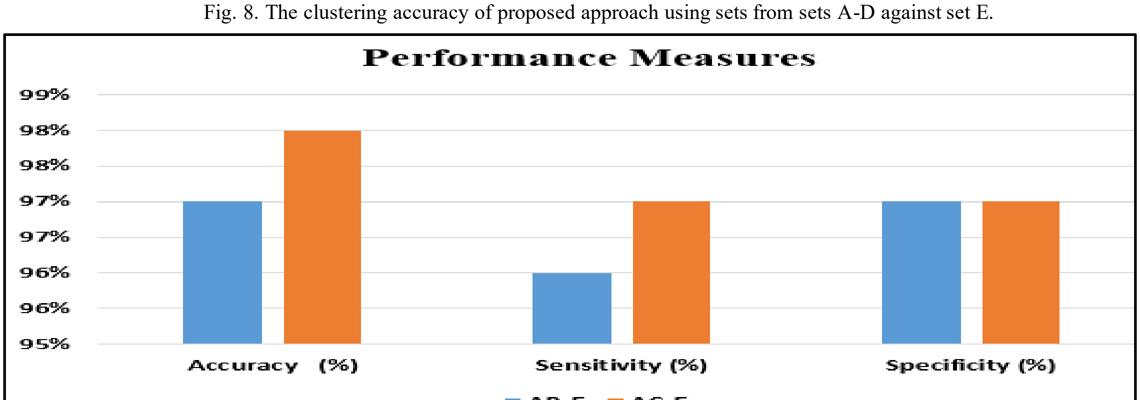

Fig. 9. The clustering accuracy of proposed approach using sets from sets AB and AC against set E.

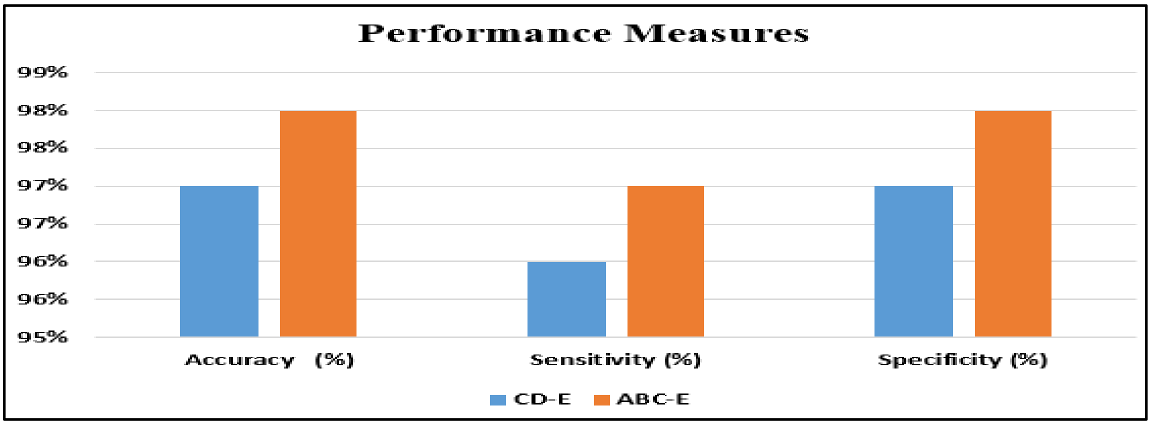

Fig. 10. The clustering accuracy of proposed approach using sets from sets CD and ABC against set E.

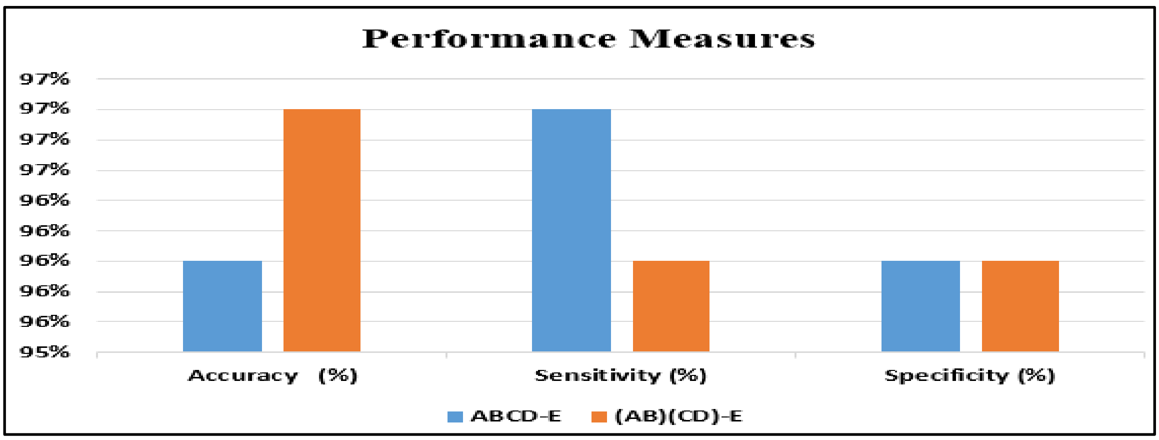

Fig. 11. The clustering accuracy of proposed approach using sets from sets ABCD and (AB)(CD) against set E.

4.4 Comparison with Existing Methods

In this study, a comparative study was conducted with several existing methods to evaluate the performance of the proposed approach. In this comparison, we compared our approach with several methods proposed in the literature that address the problem of epileptic seizure detection in EEG signals via the same dataset. On the basis of the results presented in Tables 3 and 4, wavelet transform techniques were used to find the best wavelet function. On the basis of the feature selection technique, various fixed features were extracted and recorded in this study. Compared with previous techniques, the proposed technique achieves a higher accuracy rate.

Although modern techniques have been used in previous studies, the approach proposed in this study is the most advanced and yields the highest results despite using the same dataset. DT-CWT combined with k-means yielded a higher clustering rate than did the other methods in terms of accuracy, sensitivity, and specificity. Some studies have the use of the same datasets to detect epileptic seizures, and their accuracy rates are not as high as those of our proposed approach.

TABLE III. COMPARISON OF DIFFERENT APPROACHES ON THE UNIVERSITY OF BONN EEG DATASET INCLUDING THE PROPOSED METHOD

Reference	Approach	Average Accuracy	Average Sensitivity	Average Specificity	Additional Metrics
[43]	1D CNN Deep Neural Network	97.63-99.52%	-	-	96.73-98.06% (3-class)
[44]	LSTM-CNN Hybrid Model	98.50%	97.80%	98.90%	End-to-end classification
[45]	Bi-LSTM Network Architecture	96-99%	-	-	Mean success: 97.78%
[46]	Discrete Wavelet Transform + ML	98.40%	-	-	75% dimensionality reduction
[47]	Symlet Wavelet + Gradient Boosting	93-99.66%	-	-	Three-class: 93.9%
[48]	Pre-processed Combined EEG + FFT	95-98%	-	-	Original + FFT optimal
Proposed DT-CWT + K-means	**Coupling DT-CWT with k-means**	**99%**	**98%**	**98%**	**ChaCha20 encryption**

TABLE IV. EXTENDED COMPARISON OF STATE-OF-THE-ART EEG CLASSIFICATION METHODS INCLUDING THE PROPOSED DT-CWT + K-MEANS APPROACH

Reference	Approach	Average Accuracy	Average Sensitivity	Average Specificity	Additional Metrics
[49]	DWT + SVM + RUSBoosted Ensemble	97.00%	96.67%	-	FPR: 3.24%
[50]	Dual-Tree Complex Wavelet Transform + Machine Learning	98.50%	97.80%	96.90%	Multiple classifiers tested
[51]	Wavelet Transform + SVM (Long-term iEEG)	-	94.46%	95.26%	FDR: 0.58/h
[52]	ResBiLSTM (ResNet + BiLSTM)	98.88-100%	-	-	Binary/Ternary classification
[53]	Meta-sampling + Ensemble Classifier	92.52%	92.58%	92.51%	Scalp EEG dataset
[46]	Machine Learning with DWT + Binary PSO	98.40%	-	-	75% dimensionality reduction
[54]	Transformer-based Multi-Channel EEG	97.50%	97.50%	-	Event-based: FDR 0.06/h
[55]	DenseNet-ViT with Attention Fusion	96.80%	95.20%	97.40%	Seizure prediction task
[56]	Self-supervised Transformer (SLAM)	97.07%	-	-	Patient-specific performance
[57]	Multi-channel Vision Transformer (MViT)	94.50%	93.80%	95.10%	Scalogram-based input
[58]	Tunable-Q Wavelet + CNN (Real-time)	97.57%	98.90%	-	FPR: 2.13%, Delay: 10.46s
[59]	Wavelet-based Alpha/Beta Band Analysis	96.30%	94.80%	97.20%	Computationally efficient
Proposed DT-CWT + K-means	**Coupling DT-CWT with k-means**	**99%**	**98%**	**98%**	**ChaCha20 encryption**

4.5 ChaCha20 algorithm performance analysis

The ChaCha20 algorithm, which uses a 256-bit secret key and a 96-bit IV, has proven to be highly effective and capable of securing data. It has also proven to be highly fast in encryption, with hash values ranging from 5.0E+08 to 2.0E+09 during repeated encryption operations ranging from 1-96. Figure 12 illustrates this. This level of security makes it difficult for attackers to decrypt sensitive information, which is important in healthcare applications.

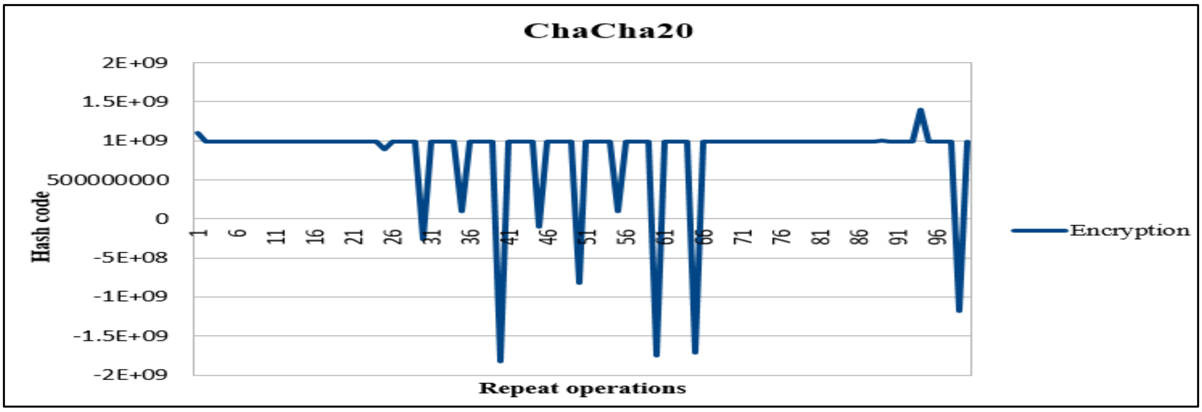

Fig. 12. ChaCha20 algorithm (Hash code).

The performance evaluation of the ChaCha20 algorithm across 100 iterative trials demonstrated a high level of consistency in both speed and efficiency. The number of encrypts remained within a narrow margin—ranging from 6.8 ms to 7.2 ms—indicating temporal stability regardless of repeated execution. On average, ChaCha20 achieved a processing time of approximately 7.0 ms, which underscores its suitability for real-time applications where latency must be minimized. In parallel, the efficiency rates fluctuated mildly between 94.4% and 96.3%, with a mean value of 95.6%, reflecting the algorithm's ability to maintain reliable throughput while conserving computational resources. Figure 13 illustrates this. These results highlight ChaCha20's balanced design, which enables secure encryption without imposing significant delays, making it a viable candidate for medical data transmission systems such as portable EEG devices.

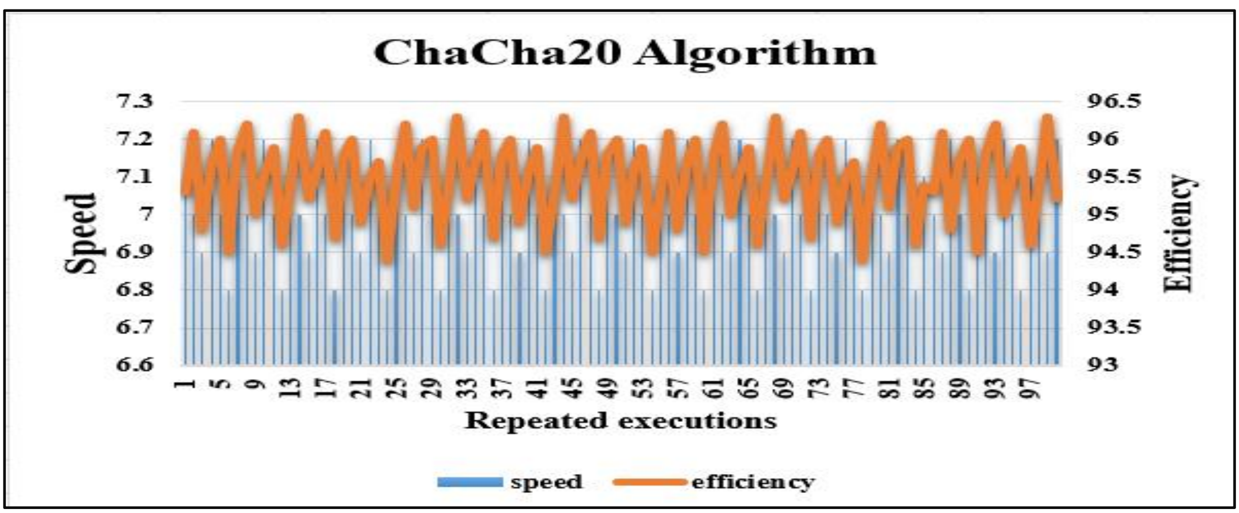

Fig 13: Performance Metrics of ChaCha20 – Speed and Efficiency Across Repeated Executions

Its consistent performance under repeated stress reinforces its practicality in time-sensitive environments that demand both security and responsiveness.

4.6 Discussion

The findings indicate that the introduced models DT-CWT, K-means and ChaCha20 achieve excellent performance for epilepsy classification reaching an accuracy of 99% for the ten scenarios. Results: Compared with existing techniques, the proposed approach shows better performance in the analysis of EEG and detection of epileptic seizures with superior accuracy. Second, the data were secured via encryption with ChaCha20 in our application to secure medical information and keep patient information secure during submission.

- **Strengths of the methodology**
  - ✓ *High accuracy in classification:* The model showed good performance in the identification and classification of epileptic cases and contributed to good support for physician decisions.

- ✓ *Improved data security:* The method usesChaCha20 encryption for the confidentiality of patient data and hence is well suited for medical studies with high security needs.
- ✓ *Computational efficiency:* DT-CWT allows for efficient analysis of neural data, enhancing model performance while maintaining efficiency in computational time. Scalability: The introduced model can be extended to diagnose other neurological diseases via engineered ensemble learning methods such as AdaBoostand, which paves the way for enhancing brain signal analysis in the future.

- **Limitations and challenges**
  - ✓ *Model sensitivity to K in K-means:* For the K-means method, the number of clusters to form K as such, must be known a priori-Selecting the K in K-means achieves the best classification accuracy.
  - ✓ *Compability of ChaCha20 in bulk applications:* Even though ChaCha20 is efficient in encryption, other algorithms such as Speck, may need to be tested to determine whether it can become more efficient with a large volume of data.
  - ✓ *Machine learning potential for improvement:* Model performance can be further improved via deep learning (DL) to increase classification accuracy and the ability of the model to support more complex data.

5. LIMITATIONS

However, the performance of the proposed model in DT--CWT and K-means clustering as well as the classification of EEG signals has several limitations, which could potentially influence the accuracy of the findings as well as the practical applicability of the obtained results. The most significant limitations are as follows:

- **Data and Sample Limitations:**
  - ✓ Sample size: The dataset used in the study is the University of Bonn EEG database, which has few samples in the category and may prohibit generalizing the results.
  - ✓ Environmental noise factors: We do not consider external interferences that could occur during real-time EEG recording (e.g. electrical noise and environmental artefacts).
- **Methodological and Technical Considerations:**
  - ✓ Wavelet Transform Effect: Even though the DT-CWT enhances the feature extraction, for classification, extracting the most significant features relies on the number of decomposition levels.
  - ✓ Computational complexity: The computational cost might be a bottleneck when this method is applied in real-time monitoring systems of EEGs because of the necessity of high resolution analysis of signals.
- **Practical Issues and Challenges for Implementation:**
  - ✓ ChaCha20 Encryption Issue: Even though the encryption technique is proven to protect patient privacy, further confirmation to ensure successful integration between EEG signal processing and the encryption algorithm is needed.
  - ✓ Clinical Implementation: For the model to be incorporated successfully into healthcare systems, the groups of research-design-validation development teams collaborate during its deployment.

Although this study employs advanced techniques such as Dual-Tree Complex Wavelet Transform (DT-CWT) for detecting abnormal patterns in EEG signals and incorporates ChaCha20 encryption to ensure data security, certain limitations remain. Specifically, the proposed framework does not directly address data poisoning threats within federated learning environments, nor has it been validated in broader cybersecurity threat [60,61].

6. CONCLUSION AND FUTURE WORK

Although remarkable progress has been made in the field of medical technology, fast, reliable, and automatic methods for clustering epilepsy cases correctly and uniformly are lacking. This paper proposes an elaborate approach for aggregating EEG cases of epilepsy; and ensures data security through encryption. The proposed approach can achieve good accuracy in disease state classification by applying the dual-tree complex wavelet transform (DT-CWT) to obtain the statistical features and then the K-means classifier. The ChaCha20 algorithm was used to encrypt the output to enhance the privacy safety of medical records, and to ensure the confidentiality of patient information.

The experimental results show that the proposed scheme achieves high performance for both encryption and clustering, with 99% clustering accuracy under ten cases. The proposed method outperforms several other methods in the literature in terms of both accuracy and consistency in epileptic seizure classification.

Further work is intended to apply the proposed approach to EEG-based signal classification combined with some other adaptive clustering methods such as AdaBoost and boosting, increasing the precision of EEG signal classification. To enhance the universality of the model in medical diagnosis, we will check whether it can be extended to patients with other types of neurologic diseases. To show how well the patient data are secured by the Speck algorithm, it can also be evaluated as an alternative to the ChaCha20 algorithm. Finally, we discuss the adaptability of the approach to the telehealth context, enabling a detailed and accurate investigation of medical time series data.

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

None.

Acknowledgment

None.

References

[1] S. Raghu, N. Sriraam, A.S. Hegde, P. L. Kubben, (2019)," A novel approach for classification of epileptic seizures using matrix determinant," *Expert Systems with Applications*, 127, 323–341. https://doi.org/10.1016/j.eswa.2019.03.021.

[2] C. Yuan, H. Yang, (2019), "Research on K-value selection method of K-means clustering algorithm", *MDPI, 2(2)*, 226-235. https://doi.org/10.3390/j2020016.

[3] H. Al-Hadeethi, S. Abdulla, M. Diykh, R. C. Deo, J. H. Green, (2020), "Adaptive boost LS-SVM classification approach for time-series signal classification in epileptic seizure diagnosis applications", *Expert Systems with Applications*, 161, 113676. https://doi.org/10.1016/j.eswa.2020.113676.

[4] E. Vafaei, and M. Hosseini, (2025), "Transformers in EEG Analysis: A Review of Architectures and Applications in Motor Imagery, Seizure, and Emotion Classification", *Sensors 2025, 25(5), 1293*, https://doi.org/10.3390/s25051293.

[5] P. Vanabelle, P. De Handschutter, R. El Tahry, M. Benjelloun, M. Boukhebouze, (2020)," Epileptic seizure detection using EEG signals and extreme gradient boosting", *Journal of biomedical research*, 34(3), 228.

[6] R. Lafta, J. Zhang, X. Tao, X. Zhu, H. Li, L. Ghang, R. Deo, (2020)," A general extensibale learning approach for multi-disease recommendations in a telehealth environment", *Pattern Recognition Letters*,132, 106-114. https://doi.org/10.1016/j.patrec.2018.11.006.

[7] W. Zeng, Mengqing. L, Chengzhi. Y, Qinghui. W, Fenglin. L, (2020)," Identification of epileptic seizures in EEG signals using time-scale decomposition (ITD), discrete wavelet transform (DWT), phase space reconstruction (PSR) and neural networks", https://doi.org/10.1007/s10462-019-09755-y.

[8] A. Malekzadeh, A. Zare, M. Yaghoobi, H. R. Kobravi, R. Alizadehsani, (2021)," Epileptic seizures detection in EEG signals using fusion handcrafted and deep learning features," *Sensors*, 21(22), 7710.

[9] X. Wang, X. Wanga, W. Liu, Zheng. C, T. Kärkkäinen, F. Cong, (2021), "One dimensional convolutional neural networks for seizure onset detection using long-term scalp and intracranial EEG" *Neuro computing*, DOI: 10.1016/j.neucom.2021.06.048

[10] A. Malekzadeh, A. Zare, M. Yaghoobi, R. Alizadehsani, (2021)," Automatic diagnosis of epileptic seizures in EEG signals using fractal dimension features and convolutional Autoencoder method", *Big Data and Cognitive Computing*, 5(4), 78.

[11] I. Ahmad, X. Wang, M. Zhu,(2022) "EEG-Based Epileptic Seizure Detection via Machine/Deep Learning Approaches: A Systematic Review", *Computational Intelligence and Neuroscience*, https://doi.org/10.1155/2022/6486570.

[12] A. Shoeibi, N. Ghassemi, M. Khodatars, P. Moridian, R. Alizadehsani, A. Zare, J. M. Gorriz, (2022), "Detection of epileptic seizures on EEG signals using ANFIS classifier, autoencoders and fuzzy entropies", *Biomedical Signal Processing and Control*, 73, 103417.

[13] R. L. Lafta, M. S. AL-Musaylh, Q. M. Shallal, (2022)," Clustering similar time series data for the prediction the patients with heart disease", *Indonesian Journal of Electrical Engineering and Computer Science*, 26(2), 947-954. DOI: 10.11591/ijeecs.v26.i2.pp947-954.

[14] L. Zhen, B. Zhu, M. Hu, Z. Deng, J. Zhang, (2022), "Revised Tunable Q-Factor Wavelet Transform for EEG-Based Epileptic Seizure Detection" *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, DOI 10.1109/TNSRE.2023.3257306

[15] F. A. Ahmed, A. S. M. A. Shamsul, M. R. Tamjid, A. C. Sabbir, (2022), “Detection of Schizophrenia from EEG Signals using Dual Tree Complex Wavelet Transform and Machine Learning Algorithms” *Bangladesh Journal of Medical Physics*, DOI:10.3329/bjmp.v15i1.63559

[16] K.V. N. Kavitha, A. Sharmila, A. I. Lucky, O. Stephen, K. S. Senthamil, T. A. Ahamed, A. Musah, (2022),” On the Use of Wavelet Domain and Machine Learning for the Analysis of Epileptic Seizure Detection from EEG Signals” *Journal of Healthcare Engineering*, https://doi.org/10.1155/2022/8928021.

[17] R. Shukla, B. Kumar, G. Gaurav, G. Singh, A. K. Sahani, (2022),” Epileptic Seizure Detection Using Continuous Wavelet Transform and Deep Neural Networks”, *SPRINGER NATURE*, https://doi.org/10.1007/978-3-030-98886-9_23.

[18] A. J. Chashmi, (2023),” Presenting a technique for identifying and diagnosing epileptic seizures using nonlinear feature extraction based on DT-CWT coefficients of brain EEG signals with a deep learning approach”, *Future Generation of Communication and Internet of Things (FGCIOT),* 2(4): 28-35, 2023.

[19] W.Al-Salman, A. B. Al-Khafaji, and M. Al-Zubaidie, (2024) "Classification Manner Utilizing Electroencephalography Signals to Investigate Waveforms," *Karbala International Journal of Modern Science:* Vol. 10: Iss. 4, Article 8. https://doi.org/10.33640/2405-609X.3377.

[20] A. Miltiadous, K. D. Tzimourta, N. Giannakeas, M. G. Tsipouras, E. Glavas, K. Kalafatakis, and A. T. Tzallas,” Machine Learning Algorithms for Epilepsy Detection Based on Published EEG Databases: A Systematic Review” *IEEE Access*, DOI: 10.1109/ACCESS.2022.3232563.

[21] A. Pansani, E .Itikawa, (2023),” Epileptic seizure detection with Convolutional Neural Networks and the Continuous Wavelet Transform”, https://www.researchgate.net/publication/368837292. https://doi.org/10.5753/sibgrapi.est.2022.

[22] S. P. Singh, S. Tiwari, (2023), “A Dual Multimodal Biometric Authentication System Based on WOA-ANN and SSA-DBN Techniques”*MDPI*, https://doi.org/10.3390/sci5010010.

[23] Y. Wang, M. Chen, X. Xi, and H. Yang, (2023),” Bathymetry Inversion Using Attention-Based Band Optimization Model for Hyperspectral or Multispectral Satellite Imagery”, *MDPI,* https://doi.org/10.3390/w15183205.

[24] G. Jahnavi, A. D. Chowdary, A. Jagadeesh, K. B. Raju, (2023),” Fusion of CT and MRI Medical Images Using Dual Tree Complex Wavelet Transform” *International Journal for Research in Applied Science*, https://doi.org/10.22214/ijraset.2023.50366.

[25] M. Tajmirriahi, H. Rabbani, (2024),” A Review of EEG-based Localization of Epileptic Seizure Foci: Common Points with Multimodal Fusion of Brain Data”, *Journal of Medical Signals & Sensors,* DOI: 10.4103/jmss.jmss_11_24.

[26] J. Nie, H. Shu, and F. Wu, (2024), “An epilepsy classification based on FFT and fully convolutional neural network nested LSTM”, *Front. Neurosci*. 18:1436619, DOI: 10.3389/fnins.2024.1436619

[27] S. A. Holguin-Garcia, E. Guevara-Navarro, A. E. Daza-Chica, M. A. Patiño-Claro, H. B. Arteaga-Arteaga, G. A. Ruz, R. Tabares-Soto, and M. A. Bravo-Ortiz, (2024), “A comparative study of CNN-capsule-net, CNN-transformer encoder, and Traditional machine learning algorithms to classify epileptic seizure”, *BMC Medical Informatics and Decision Making*, https://doi.org/10.1186/s12911-024-02460-z.

[28] W. Shi, Y. Cao, F. Chen, W. Tong, L. Zhang, and J. Wan, (2024), “Multi perspective characterization of seizure prediction based on microstate analysis”, *Front. Neurosci.,* https://doi.org/10.3389/fnins.2024.1474782.

[29] S. A. Patel, (2024), “The Efficiency Impact of Gershgorin Circle Theorem for Feature The Efficiency Impact Of Gershgorin Circle Theorem For Feature Extraction In Computer Vision And Biomedical Signal Analysis”, *All ETDs from UAB*. 3870. https://digitalcommons.library.uab.edu/etd-collection/3870.

[30] H. Torkey, S. Hashish, S. Souissi, E. Hemdan, and A. Sayed, (2025), “Seizure Detection in Medical IoT: Hybrid CNN-LSTM-GRU Model with Data Balancing and XAI Integration”, *Algorithms 2025, 18, 77*. https://doi.org/10.3390/a18020077.

[31] H. Partamian, S. Jahromi, L. Corona, M. S. Perry, E. Tamilia, J. R. Madsen, J. Bolton, S. S. D. Stone, P. L. Pearl, & C. Papadelis, (2025), “Machine learning on interictal intracranial EEG predicts surgical outcome in drug resistant epilepsy”, *npj | digital medicine*, https://doi.org/10.1038/s41746-025-01531-3.

[32] M. A. El-Shorbagy, A. Bouaouda, L. Abualigah, and F. A. Hashim, (2025), “Atom Search Optimization: a comprehensive review of its variants, applications, and future directions”, *PeerJComput.Sci*.11:e2722, DOI10.7717/peerj-cs.2722.

[33] L. Zougagh, H. Bouyghf, M. Nahid, I. Sabiri, (2024),” A New Approach for Epileptic Seizure Detection from EEG and ECG Signals Using Wavelet Decomposition”, https://www.researchgate.net/publication/378067301, DOI: 10.1007/978-3-031-52388-5_33

[34] Q. Sun, L. Yuanjian, L. Shuangde, W. Chaodong, (2024)," Automatic Epileptic Seizure Detection Using PSO-Based Feature Selection and Multilevel Spectral Analysis for EEG Signals", *Journal of Sensors*, https://doi.org/10.1155/2022/6585800.
[35] R. Lafta, H. Alshaheen, F. H. Enad, (2024), "Classification of epileptic based on fast Fourier transform and ensemble model", *AIP Conf. Proc. 3051, 040003,* https://doi.org/10.1063/5.0191753.
[36] I. Karim. K, H. Binti. D, N. B. Zainuddin, R. Sokkalingam, M. Farooq, Muzammil. E, Gohar. A, Mudasar. Z, (2024)," Determining the optimal number of clusters by Enhanced Gap Statistic in K-mean algorithm" *Egyptian Informatics Journal*, https://doi.org/10.1016/j.eij.2024.100504
[37] S. Das, S. A. Mumu, M. A. H. Akhand, and M. A. S. Kamal, (2024)," Epileptic Seizure Detection Using Integrated Decomposed Features from EEG Signal", *Journal of Computer Science*, https://doi.org/10.3844/jcssp.2024.1270.1280.
[38] S. Yadav, S. Sharma, (2024)," Study Of Existing Methods & Techniques of K-Means Clustering" *Educational Administration Theory and Practice journal*, https://doi.org/10.53555/kuey.v30i4.1755.
[39] N. Huang, Z. Xi, Y. Jaio, Y. Zhang, Z. Jiao, X. Li, (2024)," Multi-modal feature fusion with multi-head self-attention for epileptic EEG signals", *Mathematical Biosciences and Engineering*, https://doi.org/10.3934/mbe.2024304.
[40] F. A. Fadhil, F. T. A. Alhilo, M. T. Abdulhadi, (2024), " Enhancing data security using Laplacian of Gaussian and Chacha20 encryption algorithm", *De Gruyter.* https://doi.org/10.1515/jisys-2024-0191.
[41] H. Najm, M. S. Mahdi, W. R. Abdulhussien, (2024), " Light weight Image Encryption Using Chacha20 and Serpent Algorithm", *Journal of Internet Services and Information Security,* https://doi.org/10.58346/JISIS.2024.I4.027.
[42] R. K. Muhammed, Z. N. Rashid, S. J. Saydah, (2024), "A Hybrid Approach to Cloud Data Security Using ChaCha20 and ECDH for Secure Encryption and Key Exchange", *Kurdistan Journal of Applied Research (KJAR),* https://doi.org/10.24017/science.2025.1.5.
[43] W. Zhao, (2020), "A Novel Deep Neural Network for Robust Detection of Seizures Using EEG Signals", *Computational and Mathematical Methods in Medicine*, 2020, 9689821.
[44] N. Jiwani, (2022) "A LSTM-CNN Model for Epileptic Seizures Detection using EEG Signal". *International Conference Proceedings*,
[45] E. Tuncer, E. D. Bolat (2021), "Classification of epileptic seizures from electroencephalogram (EEG) data using bidirectional short-term memory (Bi-LSTM) network architecture", *Biomedical Signal Processing and Control*, 73, 103046. https://doi.org/10.1016/j.bspc.2021.103462.
[46] L. V. Tran, H. M. Tran, T. M. Le, T. T. M. Huynh, H. T. Tran, and S. V. T. Dao, (2022), "Application of Machine Learning in Epileptic Seizure Detection", *Diagnostics*, 12(11), 2879. https://doi.org/10.3390/diagnostics12112879.
[47] X. Wang, G. Gong, N. Li (2019), "Automated Recognition of Epileptic EEG States Using a Combination of Symlet Wavelet Processing, Gradient Boosting Machine, and Grid Search Optimizer", *Sensors*, 19(2), 219. https://doi.org/10.3390/s19020219.
[48] Y. Zaid, M. Sah, C. Direkoglu, (2023), "Pre-processed and combined EEG data for epileptic seizure classification using deep learning", *Biomedical Signal Processing and Control*, 84, 104738. https://doi.org/10.1016/j.bspc.2023.104738.
[49] S. Y. Shah, H. Larijani, R. M. Gibson, and D. Liarokapis, (2024), "Epileptic Seizure Classification Based on Random Neural Networks Using Discrete Wavelet Transform for Electroencephalogram Signal Decomposition", *Appl.Sci*, https://doi.org/10.3390/app14020599.
[50] I. B. Slimen, L. Boubchir, Z. Mbarki1, H. Seddik, (2020). "EEG epileptic seizure detection and classification based on dual-tree complex wavelet transform and machine learning algorithms", *Journal of Biomedical Research*, 34(3), 151-161.
[51] Y. Liu, W. Zhou, Q. Yuan, and S. Chen, (2012), "Automatic seizure detection using wavelet transform and SVM in long-term intracranial EEG", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 20(6), 749-755.
[52] W. Zhao, W. -F Wang, L. M. Patnaik, B. -C Zhang, S.-J. Weng, S.-X. Xiao, D.-Z. Wei, and H.-F. Zhou, (2024), "Residual and bidirectional LSTM for epileptic seizure detection", *Frontiers in Computational Neuroscience, 18, 1415967*. https://doi.org/10.3389/fncom.2024.1415967.
[53] M. Liu, J. Liu, M. Xu, Y. Liu, J. Li, W. Nie, and Q. Yuan, (2025), "Combining meta and ensemble learning to classify EEG for seizure detection (Scalp EEG)", Scientific Reports, 15, 10755. https://doi.org/10.1038/s41598-025-88270-3.
[54] Y. Zhu, M. D. Wang, (2023), "Automated Seizure Detection using Transformer Models on Multi-Channel EEGs", *Research Gate*. https://doi.org/10.13140/RG.2.2.15851.98080.
[55] S. Yuan, K. Yan, S. Wang, J–X. Liu, and J. Wang, (2024), "EEG-Based Seizure Prediction Using Hybrid DenseNet–ViT Network with Attention Fusion", *Bioengineering, 11(8), 823*. https://doi.org/10.3390/brainsci14080839.

[56] T. Xiao, Z. Wang, Y. Zhang, Hongbin, S. Wang, H. Feng, Y. Zhao (2024), "Self-supervised Learning with Attention Mechanism for EEG-based seizure detection". *Biomedical Signal Processing and Control*, 89, 105789. https://doi.org/10.1016/j.bspc.2023.105464.

[57] H. Daoud, & M. A. Bayoumi, (2022), "Multi-Channel Vision Transformer for Epileptic Seizure Prediction", *Bioengineering*, 9(7), 285. https://doi.org/10.3390/bioengineering9070285

[58] M. Shen, P. Wen, B. Song, Y. Li, (2022). "An EEG based real-time epilepsy seizure detection approach using discrete wavelet transform and machine learning methods", *ELSEVIER,* https://doi.org/10.1016/j.bspc.2022.103820.

[59] S. U. Fredes, A. D. Firoozabadi, P. Adasme, P. P. Játiva, and C. Azurdia-Meza, (2024), "Enhanced Epileptic Seizure Detection through Wavelet-Based Analysis of EEG Signal Processing". *Applied Sciences, 14(13), 5783*. https://doi.org/10.3390/app14135783.

[60] M. A. Ameedeen, R. A. Hamid, T. H. H. Aldhyani, Laith A. M. Al-Nassr, S. O. Olatunji, P. Subramanian, (2024), "A Framework for Automated Big Data Analytics in Cybersecurity Threat Detection", *MJBD*, https://doi.org/10.58496/MJBD/2024/012.

[61] M. Aljanabi, H. Ahmad, (2023), "Navigating the Void: Uncovering Research Gaps in the Detection of Data Poisoning Attacks in Federated Learning-Based Big Data Processing: A Systematic Literature Review", MJBD, https://doi.org/10.58496/MJBD/2023/019.