Circuits, Systems, and Signal Processing (2024) 43:5288–5308
https://doi.org/10.1007/s00034-024-02715-0

Enhanced Epileptic Seizure Detection Through Graph Spectral Analysis of EEG Signals

Ramnivas Sharma[1] · **Hemant Kumar Meena**[1]

Received: 1 November 2023 / Revised: 25 April 2024 / Accepted: 26 April 2024 /
Published online: 18 May 2024

Abstract

Epilepsy is a persistent health condition marked by unusual and highly synchronized electrical activity in the brain cells, resulting in recurring seizures. This paper proposes a novel real-time method to improve the detection of seizures using the spectral features of non-stationary electroencephalogram (EEG) signals. It is observed that the discrete wavelet transform (DWT)-based features do not consider the interrelationship among EEG signal components. This interrelationship has been well captured by the novel representation of EEG in the form of graph signals. Here, the spectral analysis of the graph signals is investigated by the graph-based Fourier transform (GFT). Then, GFT-based features have been selected and fed into different classifiers for analysis. The seizure detection rate in two publicly available EEG-based datasets, the University of Bonn (UB) and the Neurology Sleep Clinic New Delhi (NSC-ND), have been achieved with accuracy of 98.68% and 96.84%, respectively. The accuracy achieved is significantly better than the existing state-of-the-art techniques. This approach demonstrates the impact of utilizing the interrelationship among the EEG components, followed by enhanced feature selection based on GFT for the improved detection of seizures.

Keywords EEG signals · Discrete wavelet transform · Graph Fourier transform · Machine learning algorithms

✉ Ramnivas Sharma
2020ree9534@mnit.ac.in

Hemant Kumar Meena
hmeena.ee@mnit.ac.in

1 Department of Electrical Engineering, Malaviya National Institute of Technology, Jaipur, Rajasthan 302017, India

1 Introduction

Epileptic seizure is an abrupt and abnormal electrical activity in the brain that can alter awareness, sensation, and behavior [43]. It is a common condition, affecting approximately 50 million people worldwide. Detecting epilepsy typically involves monitoring the electrical activity or functioning of the brain using electroencephalography (EEG). A person's medical history, physical examination, and diagnostic procedures like electroencephalography (EEG) and brain imaging are commonly used to diagnose epilepsy [50]. In recent years, advancements in computational neuroscience and signal processing techniques have paved the way for innovative approaches in the field of epileptic seizure detection. Among these, the exploration of graph spectral features in the analysis of EEG signals has garnered significant attention. EEG, a non-invasive method to measure brain electrical activity, provides a wealth of data, yet interpreting its intricate patterns, especially during epileptic seizures, remains a complex challenge. The integration of graph theory principles into the analysis of EEG data allows researchers to represent brain networks as graphs, where nodes represent different brain regions, and edges denote the functional or structural connections between them. Graph spectral features delve into the frequency domain characteristics of these brain networks, capturing nuanced interactions and communication patterns among various brain regions. By leveraging these features, one can identify specific graph patterns associated with epileptic activity, leading to improved seizure detection algorithms. This novel approach holds the potential to enhance our understanding of the underlying neural mechanisms during epileptic seizures, enabling more accurate and timely diagnosis.

2 Related Work

The EEG signal has a μV level magnitude produced by the coordinated neuronal activity of the brain and recorded by electrodes positioned at a particular scalp location [28, 39]. The most common EEG disturbances in epileptic episodes are spike waves and sharp waves [54]. Thus, the proper extraction of feature to identify normal and spike waves in various seizure states become essential. There are various features in the time, frequency, and time-frequency domains, including the short time fourier transform (STFT) [6], discrete wavelet transform (DWT) [32], empirical mode decomposition (EMD), Q-wavelet transformation [22], Hilbert–Huang transform (HHT) [33], and mean amplitude spectrum (MAS) [20] etc. it is possible to combine several features and create algorithms for epilepsy detection using machine learning approaches. These algorithms use various features which are extracted from the EEG signals to identify abnormal patterns that indicate the epilepsy [35]. Common features used for epilepsy detection include spectral mean, standard deviation, kurtosis, and entropy [25]. Earlier research [29, 42] utilized a range of machine learning algorithms such support vector machine (SVM), random forest, naive Bayes, logistic regression (LR)and Convolutions neural network (CNN) to classify the various seizure states. In [14] author presented a disruptive EEG model for seizure detection in real-time applications. In [7] author suggest a paradigm that combines multi-level modular networks, graph theory, and effective brain connection, which correctly distinguish between seizure- and non-seizure-related disorders. In [3] author used fractional linear prediction and SVM in their investigations of two-class classification, and they were able to attain 95.33% accuracy. According to a study cited in [24], the combination of FFT with a k-nearest neighbor (k-NN) model was found to achieve an accuracy of 98.72%. Additionally, a separate study referenced in [15] found that using the Dual-tree complex wavelet with the nearest neighbor (NN) model yielded an accuracy of 95.5%. In contrast, Acharya et al. explored four entropy-based pairs with fuzzy classifiers and reached 98.1 percent accuracy in the classification [9]. In [32] author achieved 98.7% accuracy using the DB4-DWT technique, which is focused on the SVM and artificial neural network models. A method that utilizes a tunable-Q wavelet transform with a multiscale entropy measure classifying seizure activity achieves an accuracy of 98.6% in [1]. As mentioned in [16], the use of FBSE and WMRPE has been shown to accurately predict EEG rhythms with an accuracy of 97.3%. It is also possible to reach 97.7% accuracy in classification using the entropy of FBSE-based eeg rhythms with the empirical wavelet transform (EWT), which was emphasized in [5]. This paper utilizes two publicly accessible datasets to investigate EEG seizure detection. The datasets used include the University of Bonn (UoB) dataset and the Neurology Sleep Clinic New Delhi (NSC-ND) dataset [38, 51], which were employed to identify seizure events. For the study of EEG signals, the researcher can utilize different machine learning models [18, 44] including supervised, unsupervised, deep learning neural networks [2, 45], and graph signal processing [49]. However, conventional techniques for analyzing electroencephalogram (EEG) data usually treat each EEG signal as an individual time series, which can overlook crucial information regarding the intricate interactions among different brain regions. One of the main challenges in automated epilepsy detection is the high degree of variability in EEG signals across individuals and even within the same individual over time.

To overcome these limitation, graph signal processing-based graph fourier transform (GFT) [49] offer a valuable approach by considering the brain's network structure. By representing the brain as a graph, where nodes represent distinct regions of the brain and edges denote their connections, graph signal processing methods enable the examination of interactions between various brain regions during seizures. GFT serves as a mathematical tool utilized for analyzing signals defined on graphs or data with irregular structures. In this paper DWT also used to identify specific features or patterns in the signal that may be related to brain activity. Comparing these techniques for EEG signals, we can see that both GFT and DWT offer unique advantages.

Section 3 of this paper offers a brief introduction to the basics of graph signal processing. In Sect. 4, the methodology has been discussed, including the preprocessing steps and feature extraction techniques such as DWT and GFT. The work of the experiments, which include details about the datasets and the outcomes obtained from the proposed techniques are presented in Sect. 5. The comparative analysis of the results obtained with different models is carried out in Sect. 6. Finally, in Sect. 7, we offer summary of our work.

3 Preliminary

3.1 Graph Signal Processing (GSP)

A research area known as graph signal processing [27] works with signals specified on graphs, which are mathematical structures that depict the connections between objects. One possible use for graph signal processing is to enhance the early detection of epilepsy [30]. However, conventional approaches to EEG analysis often examine each EEG signal as a separate time series, which can leave out crucial details about the intricate relationships between various brain regions. By considering the network structure of the brain, graph signal processing can help address this limitation [11, 27]. A field of mathematics known as spectral graph theory examines the properties of a graph by using its adjacency matrix's eigenvalues and eigenvectors [47]. Similarly graph convolutional neural network (GCNN) model [13] based on the GSP technique is used to conduct operations of convolution and pooling on graphs for non-euclidean data.

3.2 Contribution of Paper

- This paper introduces a novel approach utilizing the GFT to represent brain signals in a graph-based format, allowing for the exploration of interconnections among various brain regions. Moreover, the GFT analysis aids in localizing seizure activity by examining the frequency characteristics of the graph signals.
- This research suggests an improved seizure detection with the GFT, based on graph signal processing techniques and gives comparatively more accurate results for seizure classification on both the UB and NSC-ND datasets.
- This paper presents a comparative analysis of the DWT and GFT-based features, employing the Principal Component Analysis (PCA) feature selection technique.

4 Methodology

The diagram in Fig. 1 outlines the structure designed to construct a model for seizure classification. Preprocessing, feature extraction, selection, and result analysis are the four main phases that can be used to analyze EEG signals, as depicted in Fig. 1. In this paper, we incorporated techniques such as DWT and GFT, and employed Principal Component Analysis (PCA) for selecting relevant features.

4.1 Preprocessing

It is the process of preparing data to be analyzed or training a machine learning model [26]. It plays a significant role in model performance and accuracy. In this paper we perform the normalization as a preprocessing step to prepare the data for model implementation.

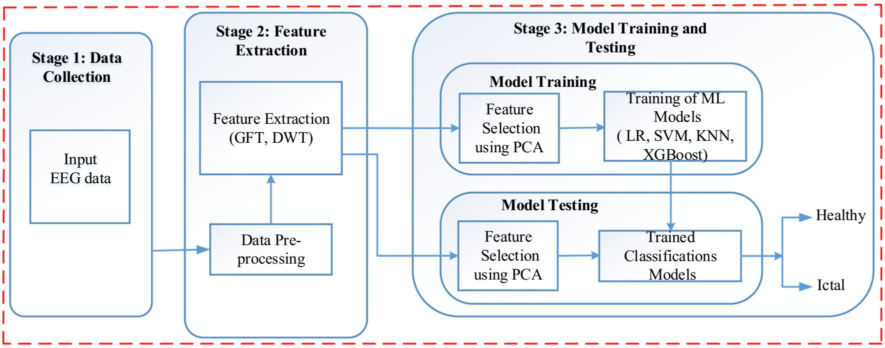


Fig. 1 Block diagram of proposed approach of model implementation

4.2 Feature Extraction

The procedure of choosing raw data and turning it into a set of features suitable for machine learning or other data analysis tasks is referred to as feature extraction [26]. Throughout the process, pertinent data is obtained from the raw data and turned into features that may be used to create models or make predictions. Feature extraction reduces noise and dimensionality and significantly affects the performance of a model. By choosing and transforming the most relevant features from the raw data, the model's accuracy is improved. In our model, we utilize both discrete wavelet transform (DWT) and graph Fourier transform (GFT) as features for both datasets. These features are then used in a machine learning-based model for the purpose of classification, as described in the subsequent discussion.

4.2.1 Discrete Wavelet Transform (DWT)

DWT is a mathematical technique for analyzing signals and images [46]. It separates a signal or image into several wavelets, short waves with varying frequencies and durations. In our paper, we used 4th level DWT, as shown in Fig. 2, which can capture both local and global aspects of a signal. The fine- and coarse-grained features can be captured using the DWT coefficients at various scales [46].

The DWT breaks down a signal into several wavelet coefficients at various scales, each representing the energy in a specific frequency band. During this recursive decomposition, the signal is divided into two frequency sub-bands using a low-pass filter called approximation coefficients, denoted as A and a high-pass filter, also called detail coefficients, designated as D shown in Fig. 2. The different waveforms of approximation coefficients and detail coefficients of 4th level DWT demonstrating the progression of normal and ictal activities, illustrating the amplitude levels using the UB dataset and NSC-ND dataset are presented in Figs. 3 and 4, respectively. To facilitate a comparative analysis of frequencies between normal and ictal instances, the scale of magnitude for ictal instances has been appropriately calibrated. Hence the difference between normal and ictal instances in Fig. 3 has been clearly noticeable at frequency level.

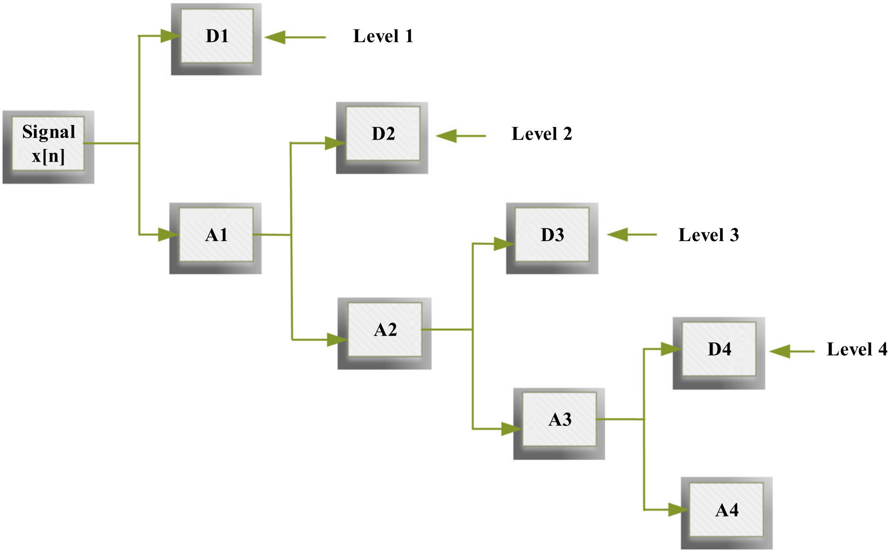


Fig. 2 DWT block diagram of 4th level

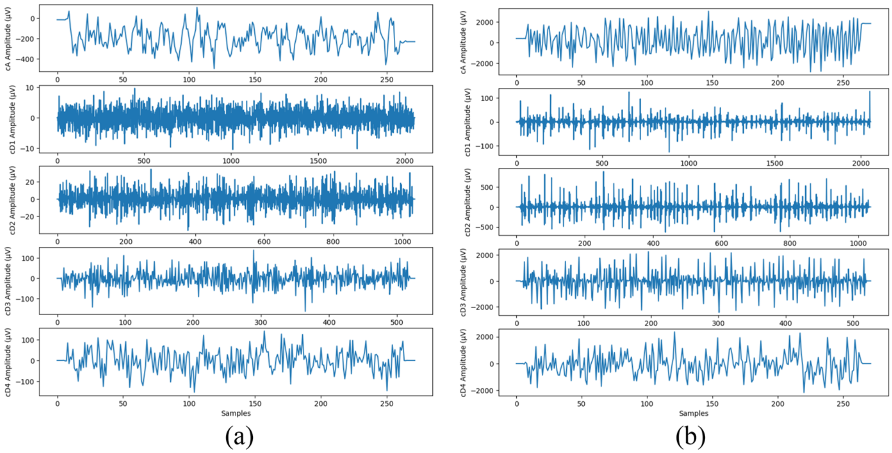
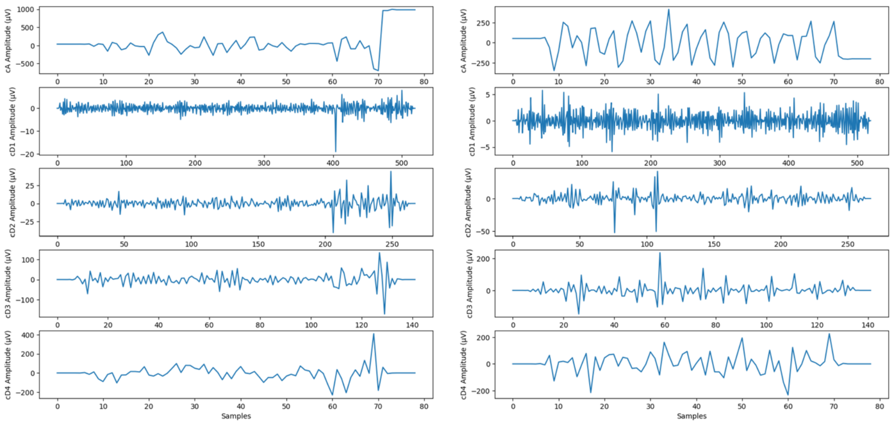


Fig. 3 DWT sub-bands coefficient waveform of UB dataset **a** normal and **b** ictal instances

DWT application includes signal processing and image compression [23, 46]. Nevertheless, one potential drawback of employing DWT as a feature is that it might not be able to fully capture the signal or image's information. The quality of the features retrieved from the DWT coefficients can depend on the chosen wavelet function and decomposition degree. The DWT is also computationally expensive, which can prevent some applications from using it. To avoid some limitations of DWT, we used GFT as another feature for epilepsy detection.



(a) (b)

Fig. 4 DWT sub-bands coefficient waveform of NSC-ND dataset **a** normal and **b** ictal instances

4.2.2 Graph Fourier Transform (GFT)

By extending the Fourier transform to signals specified on graphs, the GFT allows us to separate the signal into its frequency components [47]. GFT is a mathematical tool used to analyze signals on graphs [21]. The graph signal is denoted as $G = \{V, E, W\}$ with N vertices, and is defined as an undirected, weighted graph. Here, $V \rightarrow$ represents the vertex's set, $E \rightarrow$ indicates the set of edges, and $W \rightarrow$ is an $N \times N$ matrix that stores the edge weights. In GSP, when considering all N vertices, a signal $x : V \rightarrow \mathbb{R}$ is expressed on the vertices and can be denoted as a vector $\boldsymbol{x} \in \mathbb{R}^N$. The graph Laplacian matrix $\mathcal{L}$ [47] is defined as

$$\mathcal{L} = D - W \tag{1}$$

where

$D \rightarrow$ diagonal matrix, $W \rightarrow$ adjacency matrix.

The normalized graph Laplacian [47] is represented as

$$\tilde{\mathcal{L}} = I_N - \boldsymbol{D}^{-1/2}\,\boldsymbol{W}\boldsymbol{D}^{-1/2} \tag{2}$$

where I_N an identity matrix of size N and $\boldsymbol{D}$ denotes diagonal matrix. The detail explanation of Laplacian, graph Fourier transform and Chebyshev polynomial is given in [13]. Another representation of GFT with function $x \in \mathbb{R}^N$ on eigenvectors of the graph Laplacian is defined as:

$$\hat{x}(l) = \langle \mathbf{u}_l, x \rangle = \sum_{n=0}^{N-1} \mathbf{u}_l^*(n)x(n) \tag{3}$$

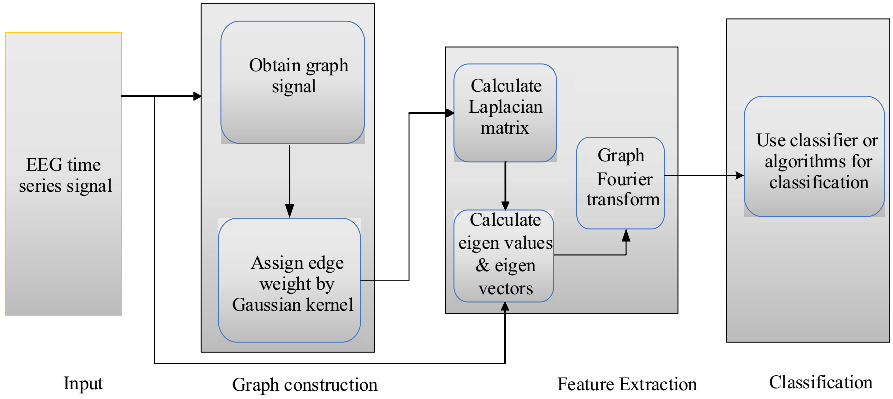


Fig. 5 Flowchart to obtain GFT

The GFT serves as a mathematical tool for analyzing signals on graphs, where a graph represents a collection of nodes (or vertices) connected by edges.

The graph Laplacian, a matrix that depicts the connectivity of the graph, serves as the foundation for the definition of the GFT. Laplacian's eigenvectors represent the frequencies of the graph signal, and each frequency's strength is represented by its eigenvalues [47, 48]. We initially encode the signal as a vector before computing the GFT of the graph signal. Then, we apply the Laplacian to the signal vector and project the result onto the eigenvectors of the Laplacian. With the help of eigenvectors, we evaluate GFT coefficients which represent the frequency domain signal. The steps or basic flowchart to obtain the GFT are shown in Fig. 5.

In the context of feature extraction using DWT and GFT, both techniques serve as powerful tools for capturing distinct characteristics within the data. DWT is particularly effective in analyzing signals with varying frequencies, making it suitable for applications like signal processing and image compression. On the other hand, GFT adeptly captures interrelationships among EEG signal components, making it valuable for scenarios where data is represented as a network or graph. Following the application of these feature extraction techniques, we obtained a dataset consisting of 500×4096 samples. This dataset size poses computational challenges due to its substantial volume.

Further, we used principal component analysis (PCA) to select the relevant features from the feature set. PCA is a dimensionality reduction technique commonly employed in machine learning and data analysis to simplify datasets while preserving essential information. The need for dimensionality reduction often arises when dealing with high-dimensional datasets, as it helps mitigate the curse of dimensionality and enhances model efficiency. By selecting a subset of principal components, PCA allows for a more concise representation of the data, making it easier to visualize and analyze. Figure 6 depicts a visual representation that illustrates the connection between feature extraction, selection, and the implementation of the model.

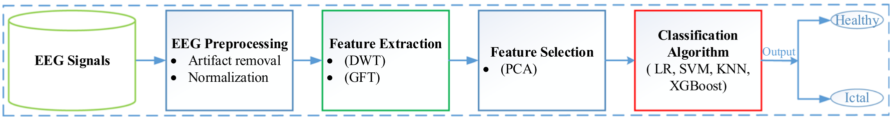


Fig. 6 Main steps of the proposed model for epileptic seizure detection

5 Experiments and Results

The framework for building the model to predict and classify seizure and non-seizure classes for epilepsy detection is illustrated in Fig. 1. In this study, DWT and GFT were used as features, while Principal Component Analysis (PCA) for feature selection. In our study, we utilized Principal Component Analysis (PCA) as a feature selection method to enhance the precision of epileptic activity identification. PCA is a widely recognized technique for reducing the dimensionality of datasets while retaining crucial information. we have expanded our discussion on the criteria employed for feature selection using PCA. Following the implementation of PCA, our dataset now comprises (500×96) sample points. Our selection criteria were driven by the goal of capturing the most significant variance in the data, with a particular focus on features associated with epileptic activities. By emphasizing features that contribute substantially to the variance related to epileptic events, we aim to improve the accuracy and relevance of our identification methodology. Expanding on the significance of feature selection, the impact of PCA-based feature selection on the accuracy of epileptic activity identification can be shown in different tables and comparative analysis section of our paper. For the purpose of detecting epileptic seizures, the chosen features fed into a different model such as support vector machine (SVM), k-nearest neighbor (kNN), extreme gradient boosting (XGBoost), and logistic regression (LR).

5.1 Datasets

The dataset UB, sourced from the University of Bonn (UB) [4], includes file name Z (Set A), O (Set B), N (Set C), F (Set D) and S (Set E). Five sets labeled as (A-E) contains 100 single-channel segments, lasting 23.6 s and sampled at 173.61 Hz. These segments were meticulously chosen from continuous multichannel EEG recordings after a thorough visual examination to identify and eliminate artifacts, such as those caused by muscle activity or eye movements. Moreover, the selected segments had to meet a stationarity criterion. The volunteers were in a relaxed, awake state with eyes open (A) and eyes closed (B) respectively. Sets C, D, and E were derived from our EEG archive focused on presurgical diagnosis. For this study, EEGs from five patients were chosen, all of whom had successfully achieved seizure control following the resection of one of the hippocampal formations, accurately identified as the epileptogenic zone. Sets Z (Set A) and O (Set B), representing EEG data from individuals with open and closed eyes respectively, were acquired from five healthy control subjects utilizing standard 10–20 system caps for surface EEG recordings [17].

Figure 7 displays sample recordings of five distinct EEG signal types obtained from subjects during a single trial. Variations in the waveforms of these five EEG signals

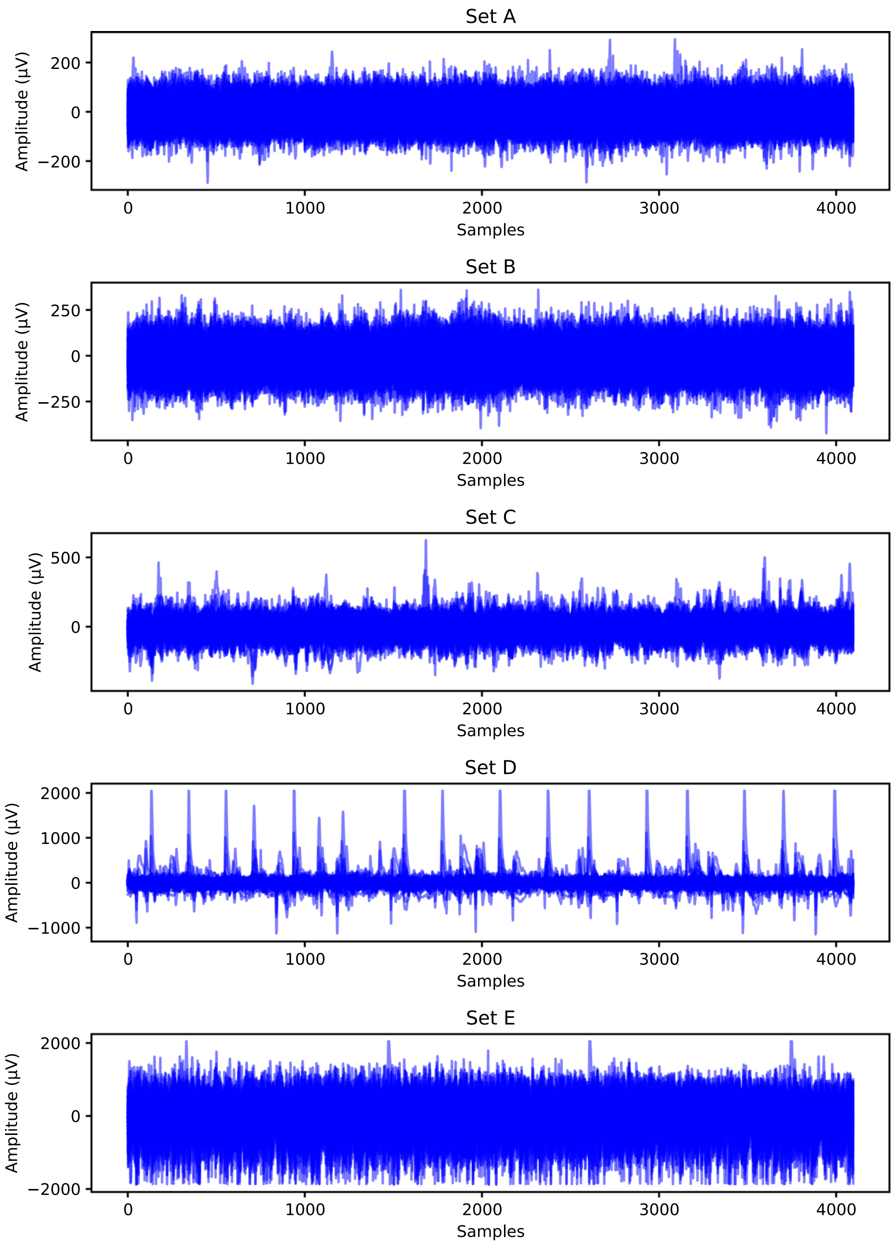


Fig. 7 The sample recording of five different types of EEG signals from UB dataset

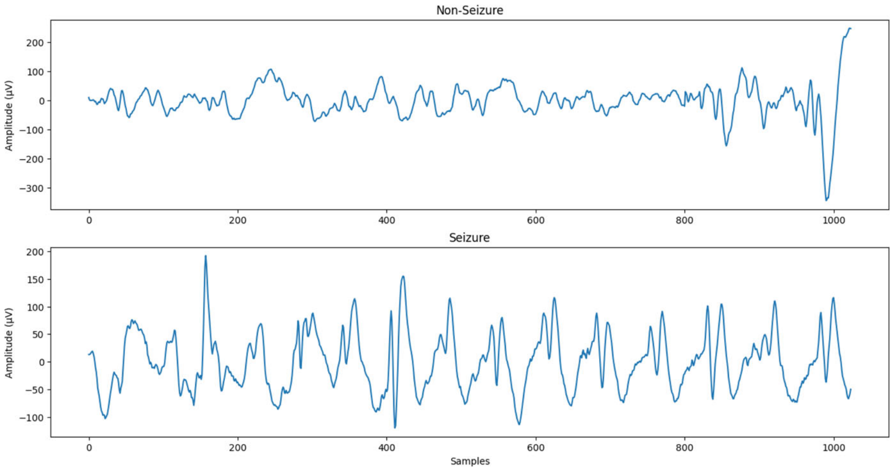


Fig. 8 Original waveform EEG segment data points from NSC-ND dataset. Here x-axis denotes samples and the y-axis denotes amplitude in μV

are evident. Notably, patients with epilepsy exhibit more pronounced fluctuations in their EEG signals compared to those of normal subjects. The epileptic period (Set E) is characterized by distinctive waves, and the fluctuations during this period are notably more severe, with significantly higher amplitudes than EEG signals observed in other periods. The sample of data points visualization of seizure and non-seizure of UB datasets is shown in Fig. 7. Sets N and Sets F were made from intracranial EEG (iEEG) signals from five epileptic patients who were not having seizures. Set S (Set E) is made up of 5 patients with epilepsy who are having active seizures, as shown by iEEG signals [4, 17].

The second EEG dataset, which we referred to as dataset 2, was obtained from the Institute for Neurology and Sleep in Hauz Khas, New Delhi (NSC-ND) [37, 52]. Ten epileptic patients' EEG recordings were gathered. Gold-plated scalp EEG electrodes were positioned using the 10–20 electrode placement technique during acquisition. EEG time series signals contains fifty MAT files per downloaded folder. Each MAT file contains 1024 samples from a single EEG time series with a duration of 5.12 s. The representation of data points seizure and nonseizure of dataset 2 (NSC-ND) where the y-axis denotes amplitude in μV is shown in Fig. 8. The sampling rate is 200 hertz (Hz). In dataset 2, there are three classes: preictal, interictal, and ictal. There are 1024 sample points in each time series signal.

5.2 Evaluation Indicators

In this study, the prediction effect of the proposed model was evaluated using accuracy, precision, recall, and F1-score [19]. The number of positive samples that were really anticipated to be positive samples is shown by TP in this instance. TN stands for true negative and represents the quantity of projected negative samples. False positives (FP), also called type 1 errors, are the number of positive samples that are expected from negative samples. False negatives (FN) show how many positive samples are mistakenly labeled as negative samples (type 2 error).

The ration of correctly classified instances to all instances is called accuracy defined as:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{4}$$

True positive rate (TPR), or sensitivity, is the proportion of cases accurately categorized as positive to all positive instances.

$$TPR = Sensitivity = \frac{TP}{TP + FN} \tag{5}$$

The ratio of instances mistakenly identified as positive to all negative instances is known as the false positive rate (FPR) or specificity:

$$FPR = specificity = \frac{TN}{FP + TN} \tag{6}$$

Precision is the proportion of correctly classified occurrences to all instances of positive classification:

$$precision = \frac{TP}{TP + FP} \tag{7}$$

Recall is the relationship between instances that were correctly classified or true positives, and cases that were misclassified, or false negatives.

$$recall = \frac{TP}{TP + FN} \tag{8}$$

Balancing indicators leads to a major increase in the complexity of the model selection process since no single criterion can be used to rank candidate models and select the best one. Two complementing indications have been combined into one in an effort to make this task simpler. The F-measure, is defined by the given below equation.

$$F - measure = 2 \times \frac{precision \times recall}{preecision + recall} \tag{9}$$

5.3 Results

In this paper, the XGBoost, SVM, LR, and kNN are taken into consideration to predict epilepsy, and their parameters are illustrated in Table 1. The specified model parameters in Table 1 represent configurations for four different machine learning algorithms: LR, SVM with an RBF Kernel, KNN, and XGBoost. For LR, a regularization parameter (C) of 0.01 suggests strong regularization, and the ’newton-cg’

Table 1 Different parameters of classifiers

Model	Algorithm	Model parameters
Logistic regression	Logistic regression	C = 0.01, solver = 'newton-cg', penalty = '12'
SVM	SVC	kernel = 'rbf', C = 100, gamma = 0.001
K-Nearest neighbors	KNeighbors classifier	n_neighbors = 5, weights = 'distance', metric: 'manhattan'
XGBoost	XGB classifier	reg_lambda = 1, reg_alpha = 0.5, n_estimators = 100, max_depth = 5 learning_rate = 0.1

Table 2 Different machine learning model implementation without any feature

Dataset	Features	Classifiers	Accuracy (%)
UB dataset	Without any features	Logistic regression	84
		XGBoost	90.01
		SVM	90.22
		kNN	82
NSC-ND dataset	Not included	Logistic Regression	77.56
		XGBoost	80.36
		SVM	83
		kNN	66.09

solver is chosen for optimization. SVM employs an RBF kernel with a relatively high regularization parameter (C=100) and a small gamma (0.001) for controls the width of the RBF kernel. KNN utilizes 5 neighbors with distance-weighted voting and a Manhattan distance metric. XGBoost incorporates L1 and L2 regularization terms (reg_alpha and reg_lambda), 100 boosting rounds (n_estimators), limits tree complexity with max_depth=5, and uses a moderate learning rate of 0.1. These parameter choices align with common practices for each algorithm, emphasizing regularization, appropriate kernel selection, and sensible values for hyperparameters, while further fine-tuning may be necessary based on specific dataset characteristics.

In epilepsy detection, we predict the occurrence of seizures in individuals by employing advanced techniques such as SVM, XGBoost [10] and LR [31]. A multivariate linear regression model serves as the foundation for the sigmoid function that constitutes logistic regression. In higher dimensional space, aggregation is visible in comparable samples. kNN [41] frequently utilize the Euclidean distance to calculate distances. In this article, we show how different machine learning models are used on the two datasets, with or without features. The model implementation results without including any feature are shown in Table 2.

The study utilized two datasets and employed the 4-level discrete wavelet transform and graph Fourier transform as inputs for the machine learning algorithms kNN, XGBoost, LR, SVM. The results, showcased in Tables 3 and 4, demonstrate the performance of each model in various seizure states. We have included a table that outlines the trials of different features used in the proposed approach, which is presented in Table 5. The sensitivity of the results was assessed, and the SVM model, utilizing

Table 3 Experimental result with discrete wavelet transform (DWT) as a Feature

Dataset	Features	Classifiers	Accuracy (%) with DWT Feature	Accuracy (%) using DWT with PCA
UB dataset	DWT, PCA	kNN	85.98	91
		Logistic regression	88.62	88
		XGBoost	92.45	92.21
		SVM	96.62	94
NSC-ND dataset	DWT, PCA	kNN	80	90
		Logistic regression	87	76.22
		XGBoost	90	83.33
		SVM	92.50	90

Table 4 Experimental results using GFT as Feature

Dataset	Features	Classifiers	Accuracy (%) with GFT Feature	Accuracy (%) using GFT with PCA
UB dataset	GFT, PCA	kNN	83.22	92.86
		Logistic regression	90	91.20
		XGBoost	95.33	97.03
		SVM	97.45	**98.68**
NSC-ND dataset	GFT, PCA	kNN	81.36	87.12
		Logistic regression	80	89.86
		XGBoost	90	96.63
		SVM	93	**96.84**

Bold values indicate the highest accuracy of the proposed method among the other included methods

Table 5 Operational procedure at each trial

Cases/ Trials	without feature	with DWT	with GFT	GFT with PCA
Trial 1	✓	–	–	–
Trial 2	–	✓	–	–
Trial 3	–	–	✓	–
Trial 4	–	–	–	✓

features from the graph Fourier transform, demonstrated superior performance in comparison to other models, as outlined in Tables 6 and 7. The Figs. 9 and 10 illustrate the comparative accuracy of various implemented models on both UB and NSC-ND datasets respectively.

In contrast to wavelet techniques, the performance of the GFT features for seizure detection is the main emphasis of this work. By using GFT in EEG analysis, we can focus on specific frequency bands instead of all frequency bands. This reduces

Table 6 Overall Results of various stages using UB dataset with SVM

Cases/ Trials	Precision	Recall	F1 score	Accuracy (%)
Trial 1	0.90	0.768	0.80	90.2
Trial 2	0.95	0.91	0.93	96.62
Trial 3	0.951	0.953	0.94	97.45
Trial 4	0.98	0.914	0.945	**98.68**

Bold value indicates the highest accuracy of the proposed method among the other included methods

Table 7 Overall Results of various stages using NSC-ND dataset with SVM

Cases/ Trials	Precision	Recall	F1 score	Accuracy (%)
Trial 1	0.89	0.81	0.82	83
Trial 2	0.94	0.93	0.93	92.50
Trial 3	0.95	0.90	0.92	93
Trial 4	0.94	0.98	0.96	**96.84**

Bold value indicates the highest accuracy of the proposed method among the other included methods

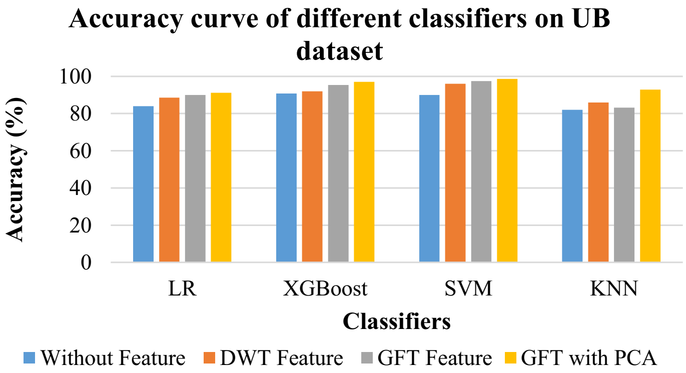


Fig. 9 Comparison of accuracy of the different implemented models on UB dataset

the amount of work the computer has to do during the calculation. This can give us additional insights from the perspective of clinic detection.

6 Comparative Analysis

Different tables present experimental classifiers, each revealing distinct outcomes. Table 2 indicates that machine learning classifiers like kNN, LR, XGBoost, and SVM perform inferior as compared to Table 3 when features from both datasets are overlooked. In contrast, Table 3, utilizing a discrete wavelet transform-based feature, outperforms Table 2. Comparing Table 4, where the graph Fourier transform is used as a feature, to Table 3, all models exhibit enhanced accuracy. Further, we utilized Principal Component Analysis (PCA) to reduce dimensionality and select features

Table 8 Comparison of the proposed method with existing state-of-the-art

Ref.	Dataset	Method	Accuracy (%)
[46]	University of Bonn (UoB)	DWT+ SVM RUSBoosted tree Ensemble	97
[23]	UoB (Z, F, S)	DWT+ SVM	96.59
[8]	UoB	VMD- bandwidths spectral features-RF	98.7
[55]	UoB (ZO-NF-S)	CNN	96.97
[16]	UoB (Z, F, S)	Morlet wavelet + Least squares	97.3
[5]	UoB	EWT + Ensemble classifiers	97.7
[53]	UoB (ZO-NF-S)	Bi-LSTM	88
[34]	UoB	Stein kernel based SR	97.99
	NSC-ND	SPD matrices	97.50
[36]	UoB	LCADC	100
	NSC-ND	ASA	97
[40]	UoB	VMD, HT	100
	NSC-ND	EMRVFLN	99.74
[12]	UoB (S-ZNOF)	Intersecting ED	99.4
	NSC-ND (A-B)	Intersecting ED	94
Proposed Method	UoB	Feature: GFT	
		Logistic regression	91.20
		KNN	92.86
		XGBoost	97.03
		SVM	**98.68**
	NSC-ND	Feature: GFT	
		Logistic regression	89.86
		KNN	87.12
		XGBoost	96.63
		SVM	**96.84**

Bold values indicate the highest accuracy of the proposed method among the other included methods

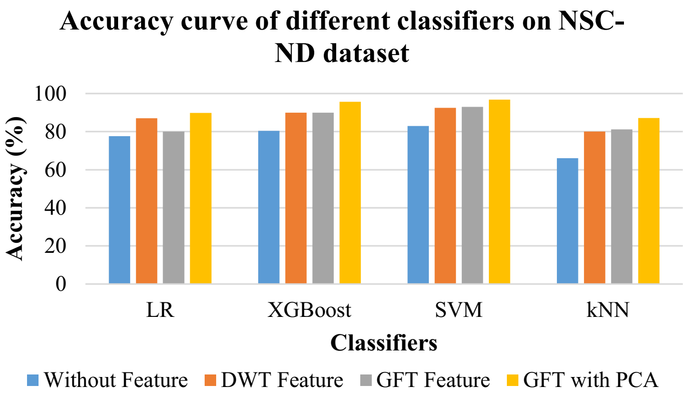


Fig. 10 Accuracy curve of the different implemented model on NSC-ND dataset

from the Graph Fourier Transform, as demonstrated in Table 4. This approach provides an improvement in accuracy. Table 5 outlines four specific trials: trial 1, where the model operates without feature extraction (as seen in Table 2); trial 2, where only the DWT-based feature is considered, implemented on both datasets and reflected in Table 3.

In Trial 3, detailed in Table 5, we exclusively utilized the graph Fourier transform (GFT) as a feature, drawing upon graph signal processing techniques. In this trial, a range of machine learning models are applied, and the corresponding results can be found in Table 4. Trial 4, is also documented in Table 5, incorporated solely the graph Fourier transform as an input feature, combined with the application of PCA. This particular trial demonstrated significantly improved accuracy, as evident in Table 4. Throughout these experiments, we specifically focused on evaluating the performance of SVM classifiers, which consistently showed superior accuracy in comparison to other models across both datasets. The outcomes of these evaluations have been presented in Tables 6 and 7. Initially, concerning the vital recall index, a noticeable pattern emerged as the dataset was divided. For the SVM model, the recall initially increased, reaching 0.95 and 0.91 in trial 3 and trial 4 on the UB datasets, as depicted in Table 6. In comparison, KNN and LR exhibited lower accuracy, precision, and F1-score than SVM and XGBoost. In the case of the SVM model, all four indexes showed improvement. The recall increased to 0.98 in trial 4, which is more than that of trial 3 on the NSC-ND dataset as shown in Table 7. In trial 4, the SVM model achieved an accuracy of 98.68%, precision of 0.98, and F1-score of 0.945 for UB datasets (Table 6), and 96.84%, 0.94, and 0.96 for the NSC-ND dataset (Table 7). These results highlight the effectiveness of our proposed approach, utilizing SVM, XGBoost, kNN, and LR, with accuracy rates of 98.68%, 97.03%, 92.86%, and 91.20% on the UB dataset and 96.84%, 96.63%, 87.12%, and 89.86% on the NSC-ND epilepsy dataset, respectively as shown in Table 4.

GFT is particularly well-suited for analyzing brain connectivity, while DWT is more widely used for feature extraction and signal noise reduction. In terms of computational complexity, DWT is generally faster and more efficient, while GFT can be more computationally demanding due to its use of graph theory concepts. The effectiveness of the suggested method and comparative analysis of previous studies in epilepsy detection using same datasplit such as UB and NSC-ND datasets is summarised in Table 8. The graphical representation of the comparative accuracy of different implemented models on both UB and NSC-ND datasets is depicted in Fig. 9 and Fig. 10, respectively. This result highlights the superiority of our technique in classifying seizure-free, inter-ictal, and healthy EEG data. Our proposed method has undergone assessment using a public dataset, resulting in enhanced accuracy. This underscores the broad applicability of our graph-based approach, with a key strength residing in the extraction of interrelationships among EEG components-based features. This approach minimizes data volume, distinguishing it from deep learning techniques. However, relying solely on two public datasets may not suffice for direct application in practical settings. The inclusion of diverse datasets would enable us to refine our model and parameters effectively, subsequently augmenting the applicability of our model in predicting actual results. The primary challenge lies in obtaining access to EEG datasets from various locations, necessitating dependence on publicly available datasets.

7 Conclusion

This paper suggested an approach for detecting epilepsy seizures based on EEG data. It assessed its effectiveness using Discrete Wavelet Transform (DWT) and Graph Fourier Transform (GFT) as features with machine learning classifiers such as LR, XGBoost, kNN, and SVM. The utilization of DWT and GFT as features provided valuable insights into the underlying graph spectral patterns associated with epileptic seizures, leading to improved detection capabilities. This study validates the potential of graph spectral features in EEG analysis and paves the way for further advancements in epileptic seizure detection. The findings underscore the importance of leveraging graph spectral techniques for more accurate and efficient diagnosis, significantly contributing to the ongoing efforts in epilepsy research. The results demonstrate that our proposed approach, which combines SVM, XGboost, kNN, and LR, achieves an accuracy of 98.68%, 97.03%, 92.86%, and 91.20% on the UB dataset and 96.84%, 96.63%, 87.12%, and 89.86% on the NSC-ND Epilepsy dataset respectively. This result shows the efficiency of our method in classifying seizure-free, interictal, and healthy EEG data. In future studies, integrating advanced signal processing techniques beyond DWT and GFT, along with machine learning algorithms for complex graph spectral features, could enhance the accuracy of seizure detection models. Exploring real-time applications and diverse datasets further would significantly improve the practicality and effectiveness of EEG-based epileptic seizure detection systems, leading to better diagnosis and patient care. Our paper explores these aspects in depth, emphasizing the technical integration of the user-centered design and collaborative problem-solving strategies employed to enhance its clinical utility.

Author Contributions Ramnivas Sharma: Original draft, Software, Review and editing of the paper, Formal analysis, and Results obtained.

Funding No Funding is involved.

Data Availability Publicly available data has been referred to in this paper.

Declarations

Conflict of interest No competing financial or interpersonal conflicts.

Ethical Approval Not involve any studies with animals or humans.

Consent to Participants Human participants are not involved.

References

1. U.R. Acharya, F. Molinari, S.V. Sree, S. Chattopadhyay, K.-H. Ng, J.S. Suri, Automated diagnosis of epileptic EEG using entropies. Biomed. Signal Process. Control **7**(4), 401–408 (2012)
2. U.R. Acharya, S.L. Oh, Y. Hagiwara, J.H. Tan, H. Adeli, Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. Comput. Biol. Med. **100**, 270–278 (2018)
3. B. Akbarian, A. Erfanian, A framework for seizure detection using effective connectivity, graph theory, and multi-level modular network. Biomed. Signal Process. Control **59**, 101878 (2020)
4. R.G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, C.E. Elger, Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. Phys. Rev. E **64**(6), 061907 (2001)
5. A. Anuragi, D.S. Sisodia, R.B. Pachori, Automated fbse-ewt based learning framework for detection of epileptic seizures using time-segmented EEG signals. Comput. Biol. Med. **136**, 104708 (2021)
6. S.M. Beeraka, A. Kumar, M. Sameer, S. Ghosh, B. Gupta, Accuracy enhancement of epileptic seizure detection: a deep learning approach with hardware realization of stft. Circuits Syst. Signal Process. **41**, 461–484 (2022)
7. W. Bomela, S. Wang, C.-A. Chou, J.-S. Li, Real-time inference and detection of disruptive EEG networks for epileptic seizures. Sci. Rep. **10**(1), 8653 (2020)
8. M. Chakraborty, D. Mitra et al., Epilepsy seizure detection using kurtosis based vmd's parameters selection and bandwidth features. Biomed. Signal Process. Control **64**, 102255 (2021)
9. G. Chen, Automatic EEG seizure detection using dual-tree complex wavelet-Fourier features. Expert Syst. Appl. **41**(5), 2391–2394 (2014)
10. T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system, in *proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, (San Francisco, California, 2016) pp. 785–794
11. M. Cheung, J. Shi, O. Wright, L.Y. Jiang, X. Liu, J.M. Moura, Graph signal processing and deep learning: convolution, pooling, and topology. IEEE Signal Process. Mag. **37**(6), 139–149 (2020)
12. N. Darjani, H. Omranpour, Phase space elliptic density feature for epileptic EEG signals classification using metaheuristic optimization method. Knowl.-Based Syst. **205**, 106276 (2020)
13. M. Defferrard, X. Bresson, P. Vandergheynst, Convolutional neural networks on graphs with fast localized spectral filtering. Adv. Neural. Inf. Process. Syst. **29** (2016)
14. N. Fatma, P. Singh, and M. K. Siddiqui. An in-depth examination of machine learning approaches for detecting epileptic seizures, in *2024 International Conference on Automation and Computation (AUTOCOM)*, (IEEE, 2024) pp. 507–511
15. P. Ghaderyan, A. Abbasi, M.H. Sedaaghi, An efficient seizure prediction method using knn-based undersampling and linear frequency measures. J. Neurosci. Methods **232**, 134–142 (2014)
16. V. Gupta, R.B. Pachori, Epileptic seizure identification using entropy of fbse based EEG rhythms. Biomed. Signal Process. Control **53**, 101569 (2019)
17. U. Herwig, P. Satrapi, C. Schönfeldt-Lecuona, Using the international 10–20 EEG system for positioning of transcranial magnetic stimulation. Brain Topogr. **16**, 95–99 (2003)
18. M.-P. Hosseini, A. Hosseini, K. Ahi, A review on machine learning for EEG signal processing in bioengineering. IEEE Rev. Biomed. Eng. **14**, 204–218 (2020)

19. M. Hossin, M.N. Sulaiman, A review on evaluation metrics for data classification evaluations. Int. J. Data Min. Knowl. Manag. Process **5**(2), 1 (2015)
20. W. Hu, J. Cao, X. Lai, J. Liu, Mean amplitude spectrum based epileptic state classification for seizure prediction using convolutional neural networks. J. Ambient Intell. Hum. Comput. **14**(11), 15485–15495 (2019)
21. S. Jang, S.-E. Moon, and J.-S. Lee. EEG-based video identification using graph signal modeling and graph convolutional neural network, in *2018 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, (IEEE, 2018), pp. 3066–3070
22. K. Jindal, R. Upadhyay, H.S. Singh, Application of tunable-q wavelet transform based nonlinear features in epileptic seizure detection. Analog Integr. Circ. Sig. Process **100**, 437–452 (2019)
23. J. Jing, X. Pang, Z. Pan, F. Fan, Z. Meng, Classification and identification of epileptic EEG signals based on signal enhancement. Biomed. Signal Process. Control **71**, 103248 (2022)
24. V. Joshi, R.B. Pachori, A. Vijesh, Classification of ictal and seizure-free EEG signals using fractional linear prediction. Biomed. Signal Process. Control **9**, 1–5 (2014)
25. A.A. Khalil, M.I. El Sayeid, F.E. Ibrahim, A.A. Khalaf, E. Gemeay, H. Kasem, S.E.A. Khamis, G.M. El-Banby, W. El-Shafai, E.-S.M. El-Rabaie et al., Efficient frameworks for statistical seizure detection and prediction. J. Supercomput. **79**(16), 17824–17858 (2023)
26. A. Khosla, P. Khandnor, T. Chand, A comparative analysis of signal processing and classification methods for different applications based on EEG signals. Biocybern. Biomed. Eng. **40**(2), 649–690 (2020)
27. R. Li, X. Yuan, M. Radfar, P. Marendy, W. Ni, T.J. O'Brien, P.M. Casillas-Espinosa, Graph signal processing, graph neural network and graph learning on biological data: a systematic review. IEEE Rev. Biomed. Eng. **16**, 109–135 (2021)
28. Y. Liang, C. Chen, F. Li, D. Yao, P. Xu, L. Yu et al., Altered functional connectivity after epileptic seizure revealed by scalp EEG. Neural Plast. (2020). https://doi.org/10.1155/2020/8851415
29. J. Mateo-Sotos, A. Torres, J. Santos, O. Quevedo, C. Basar, A machine learning-based method to identify bipolar disorder patients. Circuits Syst. Signal Process. **41**(4), 2244–2265 (2022)
30. P. Mathur and V. K. Chakka. Graph signal processing of EEG signals for detection of epilepsy, in *2020 7th International Conference on Signal Processing and Integrated Networks (SPIN)*, (IEEE, 2020), pp. 839–843
31. M. Nourelahi, A. Zamani, A. Talei, S. Tahmasebi, A model to predict breast cancer survivability using logistic regression. Middle East J. Cancer **10**(2), 132–138 (2019)
32. M. Omidvar, A. Zahedi, H. Bakhshi, EEG signal processing for epilepsy seizure detection using 5-level db4 discrete wavelet transform, ga-based feature selection and ann/svm classifiers. J. Ambient Intell. Hum. Comput. **12**(11), 10395–10403 (2021)
33. R.J. Oweis, E.W. Abdulhay, Seizure classification in EEG signals utilizing Hilbert–Huang transform. Biomed. Eng. Online **10**, 1–15 (2011)
34. H. Peng, C. Lei, S. Zheng, C. Zhao, C. Wu, J. Sun, B. Hu, Automatic epileptic seizure detection via stein kernel-based sparse representation. Comput. Biol. Med. **132**, 104338 (2021)
35. S. Poorani, P. Balasubramanie, Seizure detection based on EEG signals using asymmetrical back propagation neural network method. Circuits Syst. Signal Process. **40**(9), 4614–4632 (2021)
36. S.M. Qaisar, S.F. Hussain, Effective epileptic seizure detection by using level-crossing EEG sampling sub-bands statistical features selection and machine learning for mobile healthcare. Comput. Methods Programs Biomed. **203**, 106034 (2021)
37. X. Qin, D. Xu, X. Dong, X. Cui, S. Zhang, EEG signal classification based on improved variational mode decomposition and deep forest. Biomed. Signal Process. Control **83**, 104644 (2023)
38. K. Rasheed, A. Qayyum, J. Qadir, S. Sivathamboo, P. Kwan, L. Kuhlmann, T. O'Brien, A. Razi, Machine learning for predicting epileptic seizures using EEG signals: a review. IEEE Rev. Biomed. Eng. **14**, 139–155 (2020)
39. E.A. Refaee, S. Shamsudheen, A computing system that integrates deep learning and the internet of things for effective disease diagnosis in smart health care systems. J. Supercomput. **78**(7), 9285–9306 (2022)
40. S.K. Rout, P.K. Biswal, An efficient error-minimized random vector functional link network for epileptic seizure classification using vmd. Biomed. Signal Process. Control **57**, 101787 (2020)
41. S. Rukmawan, F. Aszhari, Z. Rustam, J. Pandelaki, Cerebral infarction classification using the k-nearest neighbor and naive Bayes classifier. J. Phys. Conf. Ser. **1752**, 012045 (2021)

42. J. Saeedi, K. Faez, M.H. Moradi, Hybrid fractal-wavelet method for multi-channel EEG signal compression. Circuits Syst. Signal Process. **33**, 2583–2604 (2014)
43. N. Sairamya, S.T. George, D.N. Ponraj, M. Subathra, Detection of epileptic eeg signal using improved local pattern transformation methods. Circuits Syst. Signal Process. **37**, 5554–5575 (2018)
44. M. Savadkoohi, T. Oladunni, L. Thompson, A machine learning approach to epileptic seizure prediction using electroencephalogram EEG signal. Biocybern. Biomed. Eng. **40**(3), 1328–1341 (2020)
45. G. Sharma, A. Parashar, A.M. Joshi, DepHNN: a novel hybrid neural network for electroencephalogram EEG-based screening of depression. Biomed. Signal Process. Control **66**, 102393 (2021)
46. M. Shen, P. Wen, B. Song, Y. Li, An EEG based real-time epilepsy seizure detection approach using discrete wavelet transform and machine learning methods. Biomed. Signal Process. Control **77**, 103820 (2022)
47. D.I. Shuman, S.K. Narang, P. Frossard, A. Ortega, P. Vandergheynst, The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. IEEE Signal Process. Mag. **30**(3), 83–98 (2013)
48. T. Song, W. Zheng, P. Song, Z. Cui, EEG emotion recognition using dynamical graph convolutional neural networks. IEEE Trans. Affect. Comput. **11**(3), 532–541 (2018)
49. L. Stanković, M. Daković, E. Sejdić, Introduction to graph signal processing, in *Vertex-frequency analysis of graph signals*. ed. by L. Stanković, M. Daković, E. Sejdić (Springer, Berlin, 2019), pp.3–108
50. P. Swami, T.K. Gandhi, B.K. Panigrahi, M. Tripathi, S. Anand, A novel robust diagnostic model to detect seizures in electroencephalography. Expert Syst. Appl. **56**, 116–130 (2016)
51. P. Swami, T.K. Gandhi, B.K. Panigrahi, M. Tripathi, S. Anand, A novel robust diagnostic model to detect seizures in electroencephalography. Expert Syst. Appl. **56**, 116–130 (2016)
52. P. Swami, B. Panigrahi, S. Nara, M. Bhatia, T. Gandhi. EEG epilepsy datasets. (2016c). https://doi.org/10.13140/RG,2(14280.32006)
53. E. Tuncer, E.D. Bolat, Classification of epileptic seizures from electroencephalogram EEG data using bidirectional short-term memory (bi-lstm) network architecture. Biomed. Signal Process. Control **73**, 103462 (2022)
54. J. Wu, T. Zhou, T. Li, Detecting epileptic seizures in EEG signals with complementary ensemble empirical mode decomposition and extreme gradient boosting. Entropy **22**(2), 140 (2020)
55. W. Zhao, W. Zhao, W. Wang, X. Jiang, X. Zhang, Y. Peng, B. Zhang, G. Zhang, A novel deep neural network for robust detection of seizures using EEG signals. Comput. Math. Methods Med. (2020). https://doi.org/10.1155/2020/9689821