DOI: http://dx.doi.org/10.26483/ijarcs.v16i2.7195

Volume 16, No. 2, March-April 2025

ISSN No. 0976-5697

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

EEG SEIZURE DETECTION AND ANALYSIS USING MULTI-DIMENSIONAL FEATURE EXTRACTION AND PCA

Vibha Aggarwal
University College, Barnala, Punjab, India

Manjeet Singh Patterh
Department of Electronics and Communication Engineering,
Punjabi University, Patiala, Punjab, India,

Sandeep Gupta*
College of Engineering and Management,
Punjabi University Neighbourhood Campus, Punjab, India,

Lovepreet Singh
University College, Barnala, Punjab, India,

Abstract: Electroencephalography (EEG) is crucial for diagnosing epilepsy by revealing brain electrical patterns. However, various artifacts like eye movements and muscle contractions can distort EEG data, making it difficult to detect seizures accurately. This study proposes an effective methodology to improve seizure detection and analysis by employing advanced feature extraction and principal component analysis (PCA) to reduce artifact impact, ultimately enhancing EEG interpretation and patient outcomes.

Keywords: Electroencephalography, Seizure, PCA, short-time Fourier transform (STFT)

1. INTRODUCTION

Epilepsy is a prevalent neurological disorder that disrupts the normal functions of the brain. It is a chronic and non-communicable condition characterized by sudden seizures, which can cause involuntary movements and affect various parts of the body, either partially or fully. Epilepsy is not a single disease, but rather a collection of syndromes with diverse underlying causes. The human brain continuously generates electrical signals, and an excessive flow of these signals can trigger seizures. The frequency and duration of these seizures may vary. A single seizure does not constitute epilepsy; instead, it is the occurrence of multiple seizures that leads to this condition. Electroencephalograms are complex and irregular in nature, but they provide information about the electrical activities of the brain. Electroencephalography (EEG) can be utilized as a tool for the diagnosis and monitoring of epilepsy.

EEG has been extensively used in the diagnosis and monitoring of neurological disorders, particularly epilepsy [1]. However, the detection and analysis of seizures from EEG signals can be a complex and challenging task due to the inherent variability and non-stationarity of the signals. To address this challenge, researchers have explored various techniques for feature extraction and classification, with the goal of improving the accuracy and reliability of seizure detection. One promising approach is the use of multi-dimensional feature extraction, which can capture the spatial and temporal patterns in the EEG signals [2]. This approach involves representing the EEG data in a higher-dimensional space, which can reveal hidden relationships and patterns that may be obscured in the original, lower-dimensional representation. In addition to multi-dimensional feature extraction, the use of PCA has also been explored for the dimensionality reduction of the extracted features. This technique can help to identify the most relevant and discriminative features, thereby improving the performance of the seizure detection algorithm.

2. METHODOLOGY

For this work, data is taken form CHB-MIT scalp EEG database. Initially, the EEG signals are pre-processed to remove noise and artifacts. The next step is to decompose the EEG signals into multiple segments using a sliding time window approach. For this work, the dataset includes records from 22 individuals, consisting of 5 males aged 3 to 22 and 17 females aged 1.5 to 19 years. For this study, 7 subjects were chosen. Each record lasts one hour and documents the time before, during, and after seizures, capturing both single and multiple seizure occurrences. The data was collected at a sampling frequency of 256 Hz and contains multiple channels of different lengths, with clear labels indicating seizure and non-seizure events.

For both seizure and non-seizure sections, we extract and normalize features such as variance, mean, skewness, kurtosis, Hjorth parameters, power spectral density (PSD), short-time Fourier transform (STFT), and power in specific frequency bands (delta, theta, alpha, beta, and gamma) to achieve zero mean and unit variance for each channel [9]. We then apply PCA to reduce the dimensionality of the feature space, retaining the first three principal components for further analysis.

Recursive feature elimination (RFE) is used to select suitable features for classification, targeting those with a p-value below 0.05. The extracted features are then used to train a support vector machine (SVM) classifier for seizure detection. This approach has been shown to be effective in detecting seizures from EEG signals, with high sensitivity and specificity [1] [2] [4] [5].

The model's performance is evaluated via nested cross-validation, with metrics like F1-score, sensitivity, accuracy,

and precision calculated for each fold. Results are visualized using 3D scatter plots, bar plots, and heat maps.

3. RESULTS

The extracted features that include time domain, frequency domain and statistical features are showing significant differences between single and multiple seizure events and for non-seizure events.

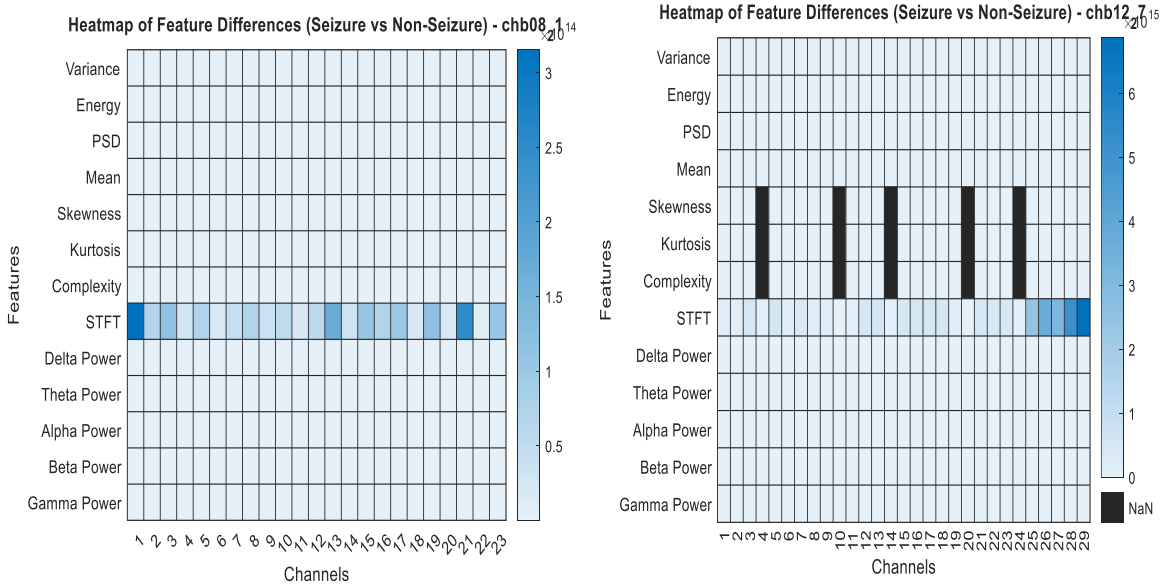

Figure 1: Heatmap of feature difference during single multiple seizure event

The heatmap of feature difference for single and multiple seizure event is showing in figure 1. As the channel numbers 4, 10, 14, 20 and 24 are grounded, so heatmap is showing no feature difference but channel numbers 25-29 are significantly affected particularly channel 29 is most effected during multiple seizure event. While for single seizure event channel number 1 is most effected followed by channel 21.

PCA is applied to reduce the dimensionality while keeping the important information form the extracted features. First three principal components during single seizure event at channel number 8 and multiple seizure event at channel number 12 is shown in figure 2. The projected points in 3D space shows clear distinction between different seizure event types.

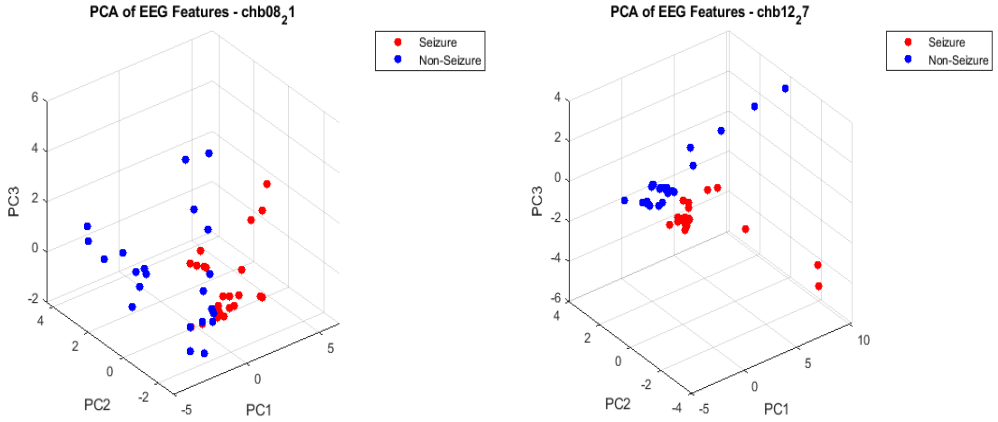

Figure 2: PCA map of seizure and non-seizure event during single and multiple seizure event

RFE was used for feature selection. To highlight the seizure and non-seizure events during multiple seizure event at channel number 12, low p-value (<0.05) features that show high difference are shown in figure 3. Form the figure, it is observed that PSD and variance can effectively notify the seizure event form the EEG signals. Similarly, frequency bands, particularly in the gamma and beta bands, highlights epileptogenic activity.

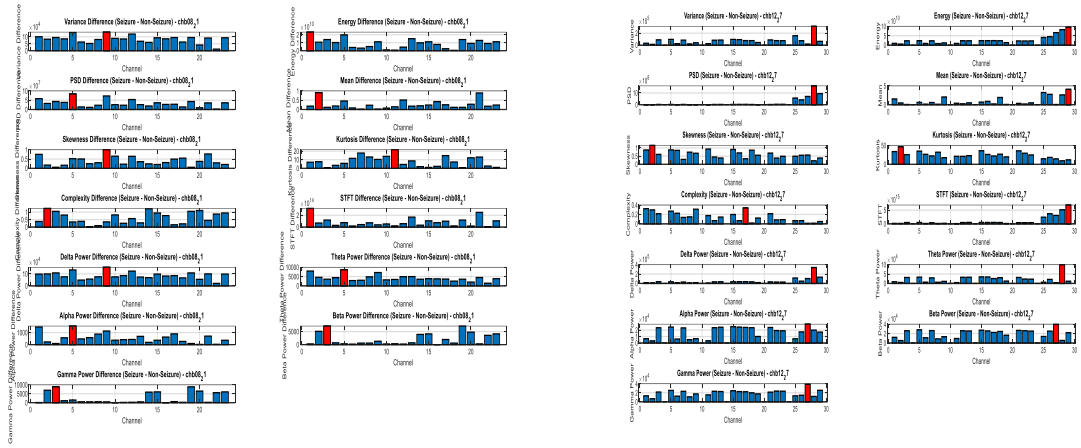

Figure 3: Affected channels vs features for single and multiple seizure event

Furthermore, SVM classifier with 5-fold cross validation show model accuracy and sensitivity 87.88% and 83.33% respectively for multiple seizure event and model accuracy and sensitivity 84.67% and 74.00 % respectively for single seizure event. These high values of accuracy and sensitivity make the model useful in identifying and monitoring the epilepsy patients to provide timely alerts and to reduce the risk during seizures.

4. DISCUSSION

The results show that the proposed method can effectively detect seizures from multi-channel EEG data, with the ability to distinguish between single and multiple seizure occurrences. The use of PCA to reduce the dimensionality of the feature space while retaining important information has been key to the success of the classifier. The feature importance analysis revealed that power spectral density and variance were particularly useful in discriminating between seizure and non-seizure events, corroborating findings from previous studies [4], [6]. Furthermore, the prominence of gamma and beta frequency bands in the selected features suggests that they are reliable indicators of epileptogenic activity, consistent with the literature. [4].

Table 1 and 2 shows the performance metrics of the proposed single and multiple time seizure detection model on a set of EEG recordings. Performance metrics are accuracy, sensitivity, F1 Score and STFT and mathematically these are calculated as given in equation 1-4:

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (1)$$

$$Sensitivity = \frac{TP}{(TP + FN)} \quad (2)$$

$$F1\ score = \frac{2TP}{(2TP + FP + FN)} \quad (3)$$

Where, TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

$$X(m,k) = \sum_{n=0}^{N-1} x(n + mH)\omega(n)e^{-j2\pi kn/N} \quad (4)$$

Where:

- X(m,k) is the STFT result,
- x(n) is the original signal,
- ω (n) is the window function,
- N is the window length,
- H is the hop size,
- k is the frequency bin index,
- m is the time frame index.

The high accuracy and sensitivity values demonstrate the efficacy of the model in real-world applications. During single seizure event, channel 1 and 21 were found to be most affected, while during multiple seizure events, channel 29 was identified as the most discriminative channel. This insight could aid clinicians in identifying the seizure focus by directing their attention to the most affected channels, potentially leading to more targeted treatment and improved patient outcomes.

The cumulative channel analysis highlights the most affected regions during seizure events. The identified brain regions that exhibit the strongest changes during seizures could serve as a starting point for future neurophysiological investigations, potentially revealing novel insights into the underlying mechanisms driving the observed seizure patterns.

The limitations to this study that should be addressed in future work are, dataset could be larger to ensure the model's generalizability. Additionally, the seizure types were not differentiated, and a more fine-grained analysis of different seizure manifestations could lead to further improvements in detection accuracy.

Table 1 Performance parameters during single time seizure

S. No.	Signal	Seizure Time		Difference	Accuracy	Sensitivity	F1 Score	STFT	
		Start	END					Channel (Average)	Difference (Average)
1	chb01_26	1862	1963	101	88.89%	100.00%	0.9	18	2.07839E+14

2	chb04_28	1679	1781	102	91.78%	88.00%	0.91	24	2.35348E+14
3	chb05_17	2451	2571	120	89.11%	100.00%	0.9	6	3.07137E+14
4	chb08_21	2083	2347	264	84.67%	74.00%	0.82	1	3.15292E+14
5	chb11_99	1454	2206	752	68.03%	76.67%	0.69	4	7.11063E+14
6	chb15_22	760	965	205	86.92%	84.64%	0.86	26	3.32714E+11
7	chb17a_04	3025	3140	115	80.15%	82.67%	0.81	19	2.18243E+14
8	chb02_16	130	212	82	86.89%	100.00%	0.88	7	1.06666E+14
9	chb07_13	3285	3381	96	91.33%	96.00%	0.92	4	3.49076E+15
10	chb17b_63	3136	3224	88	65.76%	70.00%	0.64	26	7.65738E+14
11	chb19_30	3159	3240	81	83.79%	90.00%	0.85	8	1.88319E+14
12	chb21_21	2003	2084	81	80.61%	86.67%	0.82	19	2.11259E+14

Table 2 Performance parameters during multiple time seizure

S. No.	Signal	Seizure Time												Difference	Accuracy	Sensitivity	F1 Score	STFT	
		Start	END	Start	END	Start	END	Start	END	Start	END	Start	END					Channel (Average)	Difference (Average)
1	chb12_23	253	333	425	522	630	670							217	78.48%	79.33%	0.78	17	5.58536E+12
2	chb12_27	916	951	1097	1124	1728	1753	1921	1963	2388	2440	2621	2669	229	87.88%	83.33%	0.87	29	6.86124E+15
3	chb12_29	107	146	554	592	1163	1199	1401	1447	1884	1921	3557	3584	223	87.88%	90.00%	0.88	3	2.39439E+15
4	chb12_33	2185	2206	2427	2450									192	77.27%	83.33%	0.79	4	8.03719E+13
5	chb12_42	699	750	945	973	1170	1199	1676	1701	2213	2236			156	73.73%	80.00%	0.74	3	1.04681E+13
6	chb13_62	851	916	1626	1691	2664	2721							187	71.36%	82.67%	0.74	26	7.22977E+14
7	chb15_40	834	894	2378	2497	3362	3425							242	60.50%	57.14%	0.55	38	2.75354E+13

5. CONCLUSION

In this work, we proposed a comprehensive framework for seizure detection and analysis using multi-channel EEG data. The key components of this framework include robust feature extraction from time, frequency, and statistical domains dimensionality reduction using PCA feature selection via recursive elimination SVM-based classification with parameter optimization, and interpretability through feature importance analysis and interactive visualization.

The results demonstrate the effectiveness of the proposed approach, achieving high accuracy and sensitivity in distinguishing between single and multiple seizure occurrences. The interpretability provided by the feature importance analysis and interactive dashboard offers valuable insights that can inform clinical decision-making.

Future work will focus on expanding the dataset to include a broader range of seizure patterns that may increase the model's adequacy.

REFERENCES

[1] Z. Xiong et al., "A Study on Seizure Detection of EEG Signals Represented in 2D," Jul. 29, 2021, Multidisciplinary Digital Publishing Institute. doi: 10.3390/s21155145.

[2] X. Wei, L. Zhou, Z. Chen, L. Zhang, and Y. Zhou, "Automatic seizure detection using three-dimensional CNN based on multi-channel EEG," Dec. 01, 2018, BioMed Central. doi: 10.1186/s12911-018-0693-8.

[3] W. A. W. Azlan and Y. F. Low, "Feature extraction of electroencephalogram (EEG) signal - A review." p. 801, Dec. 01, 2014. doi: 10.1109/iecbes.2014.7047620.

[4] A. S. M. Murugavel and S. Ramakrishnan, "Optimal feature selection using PSO with SVM for epileptic EEG classification," Jan. 01, 2014, Inderscience Publishers. doi: 10.1504/ijbet.2014.066229.

[5] L. Chen, Y. Zhao, P. Ye, J. Zhang, and J. Zou, "Detecting driving stress in physiological signals based on multimodal feature analysis and kernel classifiers," Mar. 01, 2017, Elsevier BV. doi: 10.1016/j.eswa.2017.01.040.

[6] A. Emami, N. Kunii, T. Matsuo, T. Shinozaki, K. Kawai, and H. Takahashi, "Seizure detection by convolutional neural network-based analysis of scalp electroencephalography plot images," Jan. 01, 2019, Elsevier BV. doi: 10.1016/j.nicl.2019.101684.

[7] S. Chatterjee, S. Pratiher, and R. Bose, "Multifractal detrended fluctuation analysis based novel feature extraction technique for automated detection of focal and non-focal electroencephalogram signals," Aug. 02, 2017, Institution of Engineering and Technology. doi: 10.1049/iet-smt.2017.0117.

[8] S. T. Ahi, H. Kambara, and Y. Koike, "A comparison of dimensionality reduction techniques for the P300 response," Jan. 01, 2009. doi: 10.1145/1592700.1592732.

[9] A. Subasi, "EEG signal classification using wavelet feature extraction and a mixture of expert model", Expert Systems with Applications, May 01, 2007, 32(4), 1084-1093. doi: 10.1016/j.eswa.2006.02.005

