https://doi.org/10.33472/AFJBS.6.6.2024.2567-2576

AFJBS African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

Detecting Epileptic Seizures Using Machine Learning Algorithms and Discrete Wavelet Transform with EEG Signals

C.Jamunadevi[1], P.Arul[2]

[1]Part time Phd Scholar, Department of Computer Science, Government Arts College (Affliated to Bharathidasan University, Trichy – 24), Tiruchirappalli – 22, India.
[2]Associate Professor, Department of Computer Science, Government Arts College (Affliated to Bharathidasan University, Trichy – 24), Tiruchirappalli – 22, India.

Email: [1]jamu83@gmail.com, [2]phdarul2004@yahoo.co.in

Article Info

Volume 6, Issue 6, June 2024

Received: 09 April 2024

Accepted: 15 May 2024

Published: 10 June 2024

doi: 10.33472/AFJBS.6.6.2024.2567-2576

ABSTRACT:

The neurological problem in the brain causes epileptic seizures, which can have an impact on a patient's health. Electroencephalogram (EEG) signal data is provided, and machine learning (ML) techniques are employed to predict epileptic seizures from the dataset. The major concern is the artifacts and noisy removal. Using a Discrete Wavelet Transform (DWT), the preprocessing method for classifying epileptic seizures is made more computationally fast. The feature extraction cutoff is used and the Daubechies wavelet is used as a scaling function to identify the noisy data from the EEG data set. The proposed method uses Principal Component Analysis with the QR Algorithm to reduce the dimensionality of the dataset, preserving the most significant components of Logistic Regression, as well as the Convolutional Neural Network (CNN) classifier, was compared with the classification relations of EEG signal to identify epileptic seizures. By applying DWT in connection with Convolutional Neural Network obtained novel and reliable classification with accuracy. Detection of epilepsy through the proposed ML algorithm is shown in the experimental results through EEG signals that merit this study with notable parameters like specificity, sensitivity, and accuracy.

Keywords - Machine Learning, EEG Signal, DWT, LR, CNN, and PAC.

1. INTRODUCTION

One of the most prevalent chronic brain conditions affecting people of all ages worldwide is epilepsy seizures. In underdeveloped nations, epilepsy afflicted about 50 million people worldwide.While there is no cure for seizures in the majority of cases, some individuals are accustomed to taking medication to avoid epilepsy. One factor that decides the kind of drug administration and helps neurologists focus on the epilepsy-located zone is the analysis of the epileptogenic zone therapeutic placement. The diagnosis process can be sped up and individuals with epilepsy can receive targeted treatment thanks to automated recognition of seizures from EEG data.

Currently available systems rely on manual identification by domain specialists, however this is mostly because to their limited characteristics that remain rigid even when recognizing large amounts of EEG signal data and patients with exacerbated seizures.

Machine learning focuses on computer programs that can access data and utilize it to learn on their own. Data is classified using both linear and non-linear classifiers, such as artificial neural networks, deep learning, decision trees, logistic regression(LR), support vector machines, k-nearest neighbor, Naive Bayes. Deep learning is one of the several subfields that make up artificial intelligence (DL). Conventional machine learning techniques that relied on feature extraction were used before deep learning (DL) came into being. The competence of individuals who crafted the features by hand limited their efficacy. However, DL offers fully automatic feature extraction and categorization. Numerous medical sectors, particularly the identification of epileptic seizures, have greatly benefited from these procedures.

Despite their advantages, machine learning algorithms have a drawback in that their effectiveness is reliant on the skill of the people creating the characteristics by hand. Deep learning is one of the many diverse subfields that make up artificial intelligence. Deep learning has an advantage over machine learning in that it fully automates the feature extraction and classification procedures. With a focus on EEG and epilepsy models, the publication specified several of the well-known deep learning models, such as hybrid CNN and LR, which detect seizures based on images produced from EEG signals. The input data are derived from the EEG using the intensity image algorithm, which divides the images into two categories: healthy and epileptic.

2. RELATED WORKS

Bhardwaj et al.[1] depict the topography features of the EEG that are obtained from the signal through clinical diagnosis. This enables the electrophysiological alterations to be complemented from the temporal to the spatial dimensions. The topography data, which is thought to be particularly important for epilepsy analysis, is thought to originate from the synchronization patterns of neuronal clusters across different brain areas. One of the main uses of a clinical approach to EEG is the detection of an epileptic neurological disease. Time-consuming has an impact on the efficiency of the laborious process and the performance of neurologists. The neurologists' task of interpreting EEG brain waves is greatly impeded by the manual examination.
Jie Xiang et.al[2] Three key distinctions between seizure and non-seizure attacks are highlighted by the fuzzyen-based method for examining epileptic seizures. D. Selvathi and others [3] This research proposes an effective method that uses wavelet transform and Support Vector Machine (SVM) classifier to detect the presence of seizure in EEG signal.

Park et al.[4] have discussed about how neurological disorders in the brain can cause epileptic seizures and have an impact on a patient's health.Numerous investigations have been conducted using topographic mapping of EEG data. The electroencephalography and brain topography extensive research are developed. Electron brain activity from neuroimaging are represented by means of feature-extracted EEG signals or raw data of EEG signals, which are derived from the amplitude of signals with colored contours. Nevertheless, existing techniques rely mostly on limited traits and are manually recognized by domain specialists.

Zhou M ., et al[5] In this study, the different stages of regions were distinguished for epileptic seizure identification using a convolutional neural network (CNN) based on raw EEG signals rather of a manual feature extraction method.

X. Jiang., et al[6] Dual-tree discrete wavelet transform (DTDWT), which overcomes the disadvantage of directionality and significant aliasing, is the focus of this work.

Subasi, A., et al[7] have discussed about a hybrid model for detecting epileptic seizures which integrate particle swarm optimization (PSO) and genetic algorithms (GA) to find the best support vector machine (SVM) parameters for EEG data categorization.

Wu J, et al [8] suggested a machine learning model to detect the ES using an EEG signal dataset obtained from the Universities of Bonn and CHBMIT. The model used the XGBOOST and CEEMD models. Ramendra Nath Bairagi., et al [9] have talked about automated seizure identification utilizing a band-pass FIR filter, which has a frequency range of 0.5–40 Hz. This filtering method is used initially to remove various sounds and artifacts combined with raw EEG signals.

R Shiva Shankar et al [10] In this dataset, they made use of 11,500 unique values and 179 attributes. The seizure attack was detected using MLP, PCA with RF, QDA, LDA, and PCA with ANN, among other techniques.. A. Zarei, et all [11] The suggested approaches extract various coefficients from the EEG signals using the discrete wavelet transform and OMP algorithms.,

Xin et al.[12] have addressed the use of the Gradient Boost algorithm to lower biasing error and nonlinear aspects of the EEG data for seizure identification. Multiple EEG signal frequency bands are measured using the entropy type of Quality scale (Q), which allows for the manipulation of multi-scale data. In order to modify the Q wavelength, the high level of signals were separated in order to control Q-based entropy (QEn). In this study, k-NN entropy is used to calculate cumulatively from the sub-band. Using classification techniques, the features obtained from the SVM are allocated to identify the stage of an epilepsy seizure that is affected.. Lurong Jiang et all . Ines Assali et all [13] This study presents a method for predicting epileptic seizures using a CNN model in conjunction with spectral and temporal information.This method has been verified using sEEG recordings from the "CHB-MIT" public database. [14]. Mingkan Shen et all [15] This study uses an accurate real-time application based on the CHB-MIT database, which has 41,280 eigenvalues from data from a 120 × 344 matrix graph.

3. PROPOSED METHODOLOGY

The proposed method is to detect epilepsy at the earlier stage with the electroencephalography (EEG) signal. The work is differentiated between several classification states. Clean and preprocess the EEG data to remove artifacts and noise. Filter out noise and remove artifacts using Principal Component Analysis (PCA).

Before applying DWT, preprocess the EEG data to remove artifacts and noise. Common pre-processing steps include filtering where to apply band-pass filters to remove power-line noise and other irrelevant frequency components. The second one is artefact removal where the Principal Component Analysis (PCA) used to remove artifacts such as eye blinks and muscle movements. Using PCA with the QR Algorithm which reduce the dimensionality of the dataset, preserving the most significant components. DWT which decomposes the signal into different frequency components with varying resolutions, useful for capturing time-frequency characteristics of EEG signals. Time-Frequency Analysis with EEG signals are non-stationary and contain transient features that vary over time. DWT provides good localization in both time and frequency domains. The features of the wavelet transform's electro-signals are extracted using a CNN.

In the below figure 1 illustrated the proposed model of detecting seizure attack. The EEG dataset is pre-processed and the noise removed data is futher processed for dimensionality reduction where PCA and QR algorithm where used. Followed with DWT type of Daubechies wavelet used for remove unwanted data like noisy data and feature extracted with the classification of CNN and LR used and finally the epilepsy seizure accuracy detecting the stages either normal or abnormal

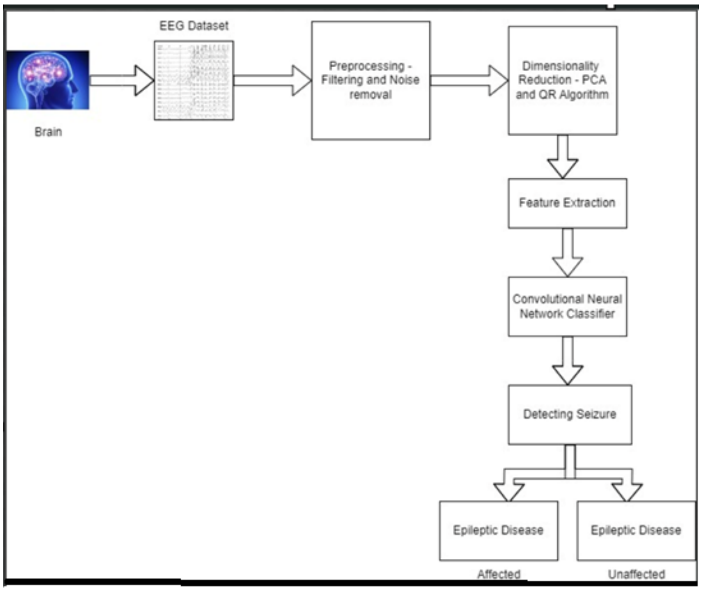

Figure.1. Proposed Method Flow Diagram

CNN's architecture learns from models. The EEG signals instinctively collect the data and perform end-to-end way classification, in contrast to the traditional approach where features are extracted and a subset of selected retrieved features is then passed to classify the classification.

3.1 EEG Dataset

The EEG dataset is one of the best methods used for the detecting of Epilepsy Seizure. The data set is taken from open repository where each of the 5 folders in the original dataset from the reference has one hundred files, each represents a particular subject or individual. Every file holds a 23.6-second recording of brain activity. Data points with the sum of 4097 are sampled from the related time-series. The result of the EEG recording at a particular instant in time is represented by each data point. There are 500 people in total, and each of them has 4097 data points for 23.5 seconds.

Every 4097 data points were split and scrambled into 23 chunks, each of which has 178 data points for a single second. The values of the EEG recordings at each distinct time are represented by the data points. Thus, we currently have 23 x 500 = 11500 informational bits (row), with 178 pieces each information 1 second(column), the last column represents the label y {1,2,3,4,5}.

3.2 Data Pre-Processing

The EEG signal is processed using the time-frequency domain using the aforementioned procedure once more. The data pre-processing includes the first EEG signal filtering and artifact reduction amplification. Additionally, the raw dataset consists of binary data with values of 0 and 1. Machine learning techniques will be used to preprocess the data and turn it into a machine-readable process. Certain illustrative modulations between various frequency bands are used to quantify brain activity. Moreover, a recent prediction suggests a high correlation between the reduction of complications and EEG signals.

3.3 Discrete Wavelet Transform

In our investigation, DWT was crucial in eliminating noisy data from EEG signals. In order to stop electrode signals from fluctuating in the temporal and frequency domains, the wavelet is passed. The scaling function is the Daubechies wavelet, and for orthogonal scaling, multi-resolution analysis is used after it is broken down into a translated version. the Daubechies function ψ(t) is scaled into a wavelet called (ψ_(a,b) (t)).

$$\frac{1}{\sqrt{a}}\psi\frac{(t-b)}{(a)}$$

where parameters were transformed, scaled by a and b, respectively, with a, b ϵ R, and a≠0.Daubechies wavelet was used to acquire the DWT signal form's parameters a and b.

$$d_{j,i} = \int_{-\alpha}^{+\alpha} s(t)^{\frac{k}{2}}\psi(t-k)dt = \big(s(t), \psi_{j,i}(t)\big)$$

The Daubechies Wavelet (dj) is identified by a location given by 'k' and a co-efficient level of 'j'. The feature vectors of the EEG classification in seizure detection are constructed using the Daubechies wavelet. It is used to threshold denoising after the DWT kind of wavelet. Time and frequency are evaluated using signal characteristics, and non-linear analysis is done with the help of Daubechies Wavelet.

Decomposition Equations:

For a signal x[n], the decomposition at level j is given by:

$aj[k]=\sum_n h[n-2k]aj-1[n]$

$dj[k]=\sum ng[n-2k]aj-1[n]$

where:

- $a_j[k]$ - approximation coefficients at level j.
- dj[k] - the detail coefficients at level j.
- h[n] - low-pass and high-pass filter coefficients, respectively.
- a0[n]=x[n] is the original signal.

3.4 CNN for feature extraction

Due to CNN's proficiency with input models for local dependencies and their capacity to share weight when training neural networks with a large number of trained parameters, they have been notably effective. EEG signals are used by both to change the wavelet characteristics. (Fig 2).

The weighted sums of the values are obtained in the surrounding pixels and output values during the convolution process itself. Convolutional kernel filters are the name given to the matrix weight. Using CNN classification techniques, kernel filters are applied to the entire convoluted input. Below the equation are the CNN feature extractions:

$$b = \sum_{n=0}^{N-1} x_n C_{k-n}$$

where the EEG signal data filter itself, the number of "x" elements, and the output vector's b, x, C, and N denoted, respectively. On the other hand, the windowing effect of the applied C kernel filter allows the features to be identified in the image.

The EEG signal[16] in Figure 2 is analyzed using the time domain on the x-axis and microwaves on the y-axis, covering both positive and negative 1s. The epileptic seizure lobs' normal and interictal stages are shown by the wave's spikes. Patients experiencing epileptic seizures are in the middle stage, known as the ictal stage, and early detection of this stage may improve the patient's prognosis.

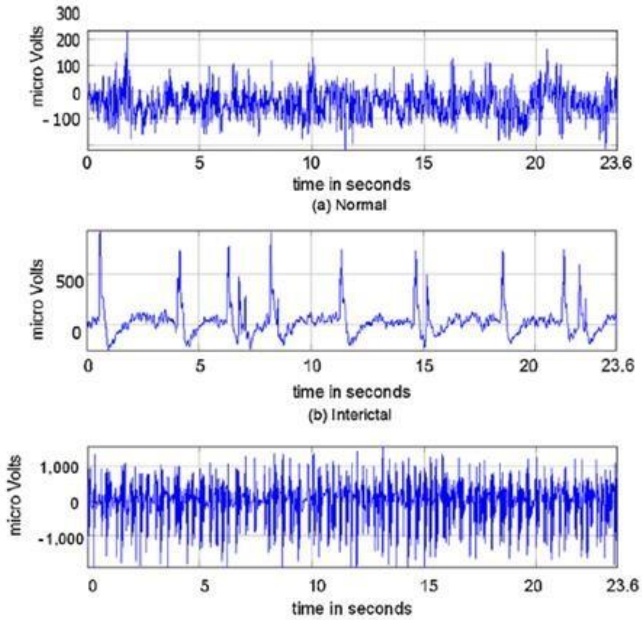

Fig.2. EEG Signal

The feature learns short temporal term connections of EEG signals and convolutional kernel filters while simultaneously looking for a relationship between the closed frequency bandwidth. The CNN method is used to pass the EEG, extract its features using a wavelet, and classify the data.

3.5 Logistic Regression (LR)

Categorization of the machine learning process that determines a dependent categorical variable's probability. The dependent variable in logistic regression (LR) is a binary variable that represents a coded value for the data. Each subset of this data set's subset k is divided into a test set and a training set, with the remaining set's k-1 acting as the latter. This helps to avoid the problem of overfitting procedures. After the holdout technique is run k times and the average error of the k trail is finished. The discrete wavelet type of the Daubechies wavelet is used to classify the raw data from the noisy binary dataset—a subset of EEG signals—into de-noising data. For an appropriate classification of epilepsy seizures from EEG signals, the noisy data are eliminated.

Algorithm 1 Epileptic seizure detection

1. Step 1: Pre-Processing
   Band-pass filters to remove power-line noise
2. Step 2: Principal Component Analysis (PCA) used to remove artifacts such as eye blinks and muscle movements.
3. Step 3 : QR Algorithm which reduce the dimensionality of the dataset.

4. Step 4: Discrete Wavelet Transform decomposes the signal into different frequency components with varying resolutions.
5. Step 5: CNN feature extractions
6. Step 6: Detection of Seizure
7. End

The chance that an event will occur and how it relates to the explanatory set of factors for classification is provided by statistical approaches given as

$$Logit(F_1) = l_n\left(\frac{F_1}{1 - F_1}\right)$$

$$= \mu 0 + \mu 1 * p1 + \cdots + \mu m p2$$

$$= \mu 0 + \sum_{i=1}^{m} \mu 1 \; X \; pi$$

Here, "µ0" represents the intercept, while the explanatory variables p1, p2,... pm are represented by the coefficients µ1, µ2, µ3... εm. Rather than altering the logarithm as a variable, LR computes the logarithm's change of the variables. The correlation between nonlinear regressions is known as the independent variable. Probability of occurrence of occurrences The independent variables' nonlinear function is provided as

P(Y=1∣X)=1+e−(β0+β1X1+β2X2+…+βnXn)1

Here, the probability values were compelled by logistic regression to fall between 0 and 1, and the LR was determined using a maximum of estimate approach with a coefficient in variables. The diagnosis process can be sped up and individuals with epilepsy can receive targeted treatment thanks to automated recognition of seizures from EEG data. Presented here is a novel application of machine learning (ML) technology for the automatic classification of EEG signals.

4. RESULT AND DISCUSSION

Data set are typically divided using training and testing. The model gains knowledge about circulated data from the training set, which contains output. In order to detect epileptic seizures in this work, 25% the testing mode makes use of the testing group's data collection, while the remaining 75% are used as a training group. The Db wavelet is utilized for decomposition and the EEG signal is used for well-being. The performance of the proposed Convolution Neural Network and LR wavelet function classification system is displayed in Table 1.

4.1. Sensitivity

Electroencephalogram (EEG) researchers have classified the total number of true positives based on seizure segmentation. The suggested approach for expert identification is used to represent the True positive for the purpose of detecting the seizure segment.

$$SN = \frac{TP}{TP + TN} \; X \; 100$$

4.2. Specificity

The free classification of all epilepsy seizures, including many genuine negatives, is represented by EEG specialists. The label of seizure for both algorithms that the EEG experts proposed was reflected by the True Negative segmentation.

$$SN = \frac{TN}{TN + FP} \; X \; 100$$

4.3 Recognition Accuracy

The total number of correctly segmented and identified cases of epilepsy that are accurately detected.

$$AC = \frac{True\ Positive + True\ Negative}{True\ P + True\ N + \ False\ P + False\ N}\ X\ 100$$

The total number of correctly identified events of a true positive and true negative is represented by the variables TN and TP, respectively. The total number of erroneously positive and negative events is represented by the FN and FP, respectively. The index performance of Daubechies wavelet transformations were quite comparable to one another in a lot of instances.

Using the Daubechies wavelet, the statistical parameters are monitored to attain the highest number of accuracy. A comparison of the findings in Table 1 indicates that the estimated EEG signals are more accurate. When compared to the current seizure detection approach, the suggested method showed an accuracy of 96.52%.

Features including CNN, testing and training sets, logistic regression, and feature extraction are to blame for the increased sensitivity.

The overall epilepsy work performance is indicated in the below table in terms of not detecting and detecting the seizure attacks.

The accuracy of seizure revealing in the classification is depending upon the feature classification and correctly segmented out of the signals. The accuracy level is used to classify using CNN's performance level and the LR technique.

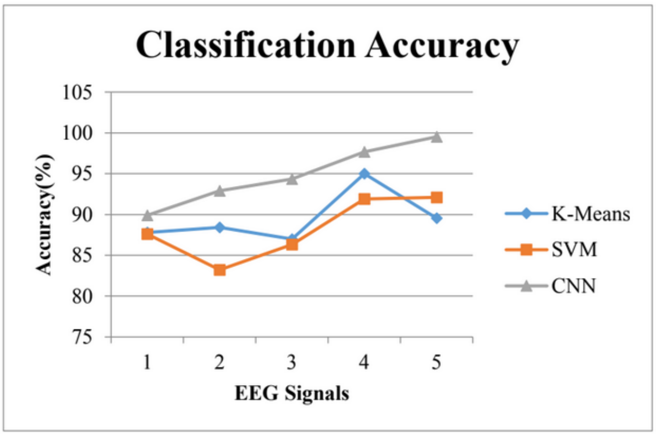

Fig.3. Accurate Epilepsy Detection

Figure 3 illustrates the accuracy level of both the suggested and current procedures. The high degree of accuracy feature classification for epilepsy detections is utilized to compare with the current SVM and K-means methods, as shown in the comparison table above. 96.52% accuracy in identifying epilepsy seizures using EEG signals that belonged to CNN and LR classification methods.

The performance comparison of the proposed identification of epilepsy seizure through advanced machine learning methods has been summarized in the table.2.

Illustrations for the comparison tables between the existing ML methods and our designed algorithm is shown table 2. According to the comparison table above, the suggested method outperformed RNN and ANN classifier techniques in terms of sensitivity level (91%), specificity level (89%), and accuracy level (97.2%).

Table.2. Performance analysis of proposed and existing machine learning algorithm

Parameter	CNN	RNN	ANN
Sensitivity	90%	71%	80%

Specificity	88%	78%	75%
Accuracy	96.5%	60%	72%

Therefore, in comparison to other approaches now in use, the suggested classification method for classification, as shown in figure 4.4, can yield a highest percentage of sensitivity, specificity, and accuracy. Convolution neural networks, or CNNs, are compared to classification and regression trees, or CART, and artificial neural networks, or ANNs, in order to evaluate these machine learning techniques' statistical features.

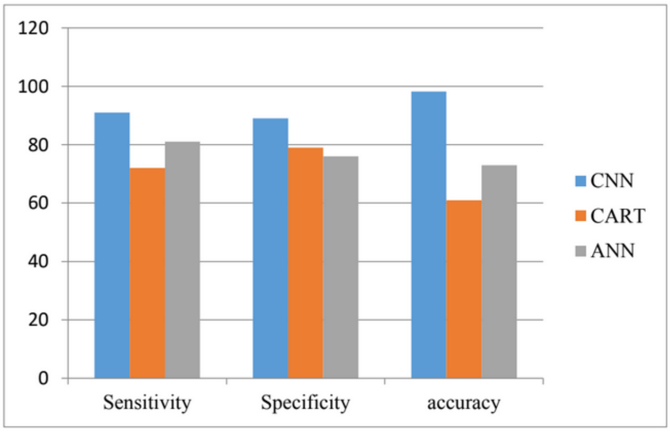

Fig.4. Comparison of Stastical Parameters

In a manner similar to this, the classifier algorithms' total percentage of correctly classified test data records acts as a proxy for the test dataset's classification accuracy. As a substitute for the accuracy metric, specificity and sensitivity are used to evaluate the classifier's performance.

4.4 Accurate Classification of Epilepsy Seizure

By merging image processing and machine learning approaches, it is possible to examine the prediction accuracy of both the presented and current methodologies by looking at how epilepsy is recognized by picture mining. By integrating image processing and machine learning approaches, epilepsy can be identified by picture mining, which allows for an analysis of the prediction accuracy of both the existing and proposed methodologies.

5. CONCLUSION

This study came to new conclusions about how to detect epileptic seizures using EEG signals. A brain EEG measures electrical activity. This research examines how the brain recognises seizures.When analysing seizure data, wavelet signals from EEG signals describe electrodes spatially, Electroencephalography is used to categorise automatically detected seizures using logistic regression and CNN. achieving a promising outcome with high classification accuracy rates and concentrating on EEG signal spectrum representation to enhance categorization. Experts' exposure to EEG data is dependent on their work with epileptic episodes. These studies demonstrate how logistic regression (LR) and convolutional neural networks (CNNs) can function with the EEG signal to provide more accurate classification than current techniques. Categorizing these two approaches to EEG readings and achieves an overall classification accuracy of 96.52% in detecting epileptic episodes. Experimental results provide EEG signals with excellent specificity, sensitivity, and accuracy.

6. REFERENCES

1. Bhardwaj, A., Gupta, A., Jain, P., Rani, A., & Yadav, J. Classification of human emotions from eeg signals using svm and lda classifiers. In 2015 2nd international conference on signal processing and integrated networks (spin), 2015.
2. Jie Xiang, Conggai Li et.al., “The detection of epileptic seizure signals based on fuzzy entropy”, Journal of Neuroscience Methods, Volume 243, 30, pages 18-25, 2015.
3. Selvathi, D., and V. Krishna meera. "Realization of epileptic seizure detection in EEG signal using wavelet transform and svm classifier." 2017 international conference on signal processing and communication (icspc). IEEE, 2017
4. Park, Chulkyun, et al. "Epileptic seizure detection for multi-channel eeg with deep convolutional neural network." 2018 international conference on electronics, information, and communication (iceic). IEEE, 2018.
5. Zhou M, Tian C, Cao R, Wang B, Niu Y, Hu T, Guo H, Xiang J. Epileptic Seizure Detection Based on EEG Signals and CNN. Front Neuroinform., 2018.
6. X. Jiang, K. Xu, R. Zhang, H. Ren, W. Chen., “Redundancy removed dual-tree discrete wavelet transform to construct compact representations for automated seizure detection”, Applied Sciences., 9(23), 5215., 2019..
7. Subasi, A.; Kevric, J.; Abdullah Canbaz, M. Epileptic seizure detection using hybrid machine learning methods. Neural Comput. Appl., 31, 317–325., 2019.
8. Wu, J., Zhou, T., & Li, T.,” Detecting epileptic seizures in EEG signals with complementary ensemble empirical mode decomposition and extreme gradient boosting. Entropy”, 22(2), 140., 2020..
9. Ramendra Nath Bairagi., et al., “Epileptic seizure identification in EEG signals using DWT, ANN and sequential window algorithm”, Soft Computing Letters, Volume 3, 100026, 2021. .
10. R Shiva Shankar et al.,”Detection of Epilepsy based on EEG Signals using PCA with ANN Model”, J. Phys.: Conf. Ser. 2070 012145., 2021. .
11. A. Zarei, B.M. Asl., Automatic seizure detection using orthogonal matching pursuit, discrete wavelet transform, and entropy based features of EEG signals Comput. Biol. Med., 131, Article 104250., 2021. .
12. Xin Xu,Makoun Lin,Tingting Xu, “Epilepsy seizures prediction based on nonlinear features of eeg signal and gradient boosting decision tree”. International journal of environmental research and public health ; 19(18): 11326, 2022.
13. Ines Assali, Ahmed Ghazi Blaiech, Asma Ben Abdallah, Khaled Ben Khalifa, Marcel Carrère, Mohamed Hédi Bedoui, CNN-based classification of epileptic states for seizure prediction using combined temporal and spectral features,Biomedical Signal Processing and Control,Volume 82,2023,.
14. Lurong Jiang, Jiawang He, Hangyi Pan, Duanpo Wu, Tiejia Jiang, Junbiao Liu, Seizure detection algorithm based on improved functional brain network structure feature extraction, Biomedical Signal Processing and Control, Volume 79, Part 1, 2023.
15. Mingkan Shen, Fuwen Yang, Peng Wen, Bo Song, Yan Li., “A real-time epilepsy seizure detection approach based on EEG using short-time Fourier transform and Google-Net convolutional neural network”., Heliyon, Volume 10, Issue 11, 2024.
16. P. Ashok, S. Praveen Chakkravarthy, U. Arunkumar and P. Parthasarathy, "LR-CNN based Classification for Improved EEG Epilepsy Seizure Detection using Discrete Wavelet Transform Technique," *2023 7th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, Coimbatore, India, 2023