Multimedia Tools and Applications
https://doi.org/10.1007/s11042-023-14995-w

Capsule neural network based approach for subject specific and cross-subjects seizure detection from EEG signals

Gopal Chandra Jana[1] · **Keshav Swami**[2] · **Anupam Agrawal**[1]

Received: 25 June 2021 / Revised: 27 January 2023 / Accepted: 22 February 2023

Abstract
The objective of this study is to propose an approach to detect Seizure and Non-Seizure phenomenon from the highly inconsistent and non-linear EEG signals. In the view of performing cross-subject classification over the inconsistency and non-linear characteristics of EEG signals, we have proposed a fine-tuned Capsule Neural Network (CapsNet) based approach to classify the seizure and non-seizure EEG signals through subject specific and cross-subject training and testing. In this experiment, first we have normalized the input data using L2 normalization technique. In the second step, the normalized data have been given to the CapsNet and model level fine-tuning has been carried out. In addition to this, we have performed seizure and non-seizure classification performance evaluation using three more classifiers such as Decision Tree, Logistic Regression, Convolutional Neural Network to compare with the performance of the proposed approach. To estimate the effectiveness of the proposed approach, subject specific and cross-subject training and testing have been performed. In both experiments, we have used multi-channel and single channel EEG datasets. For subject specific experiment, the proposed approach achieved a mean accuracy of 93.50% over the dataset-1 (multi-channel) and an accuracy of 82.61% for dataset-2 (single channel). For cross-subject experiment, the proposed approach achieved a highest mean accuracy of 86.41% over the dataset-1(multi-channel) and a mean accuracy of 48.45% over the dataset-2 (single channel) which shows an advantage of CapsNet in a certain data scenario as described in result section. Overall performance of the proposed approach shown a comparable improvement over the existing approaches.

✉ Gopal Chandra Jana
go.gopal.ch.jana@gmail.com

[1] Interactive Technologies & Multimedia Research Lab, Department of Information Technology, Indian Institute of Information Technology – Allahabad, Prayagraj, UP 211015, India

[2] School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar, Odisha 751024, India

Published online: 11 March 2023

Keywords Electroencephalogram (EEG) · Cross-subject seizure detection · Capsule neural network · Decision tree · Logistic regression · Convolutional neural network

1 Introduction

A seizure occurs when there is an abnormal electrical discharge in the human brain, where Epilepsy is defined as having continuing seizures [13]. More specifically, Epilepsy is a medical condition where some disorder occurs in the neurological level of the brain by some sudden changes of electrical activity in neurons. The result in increasing the electrical activity may result in violent body shakes or sometimes a simple staring spell which may not be noticed. Some physical changes like headache, nausea or some other stomach problems like rising feelings from stomach to throat, tingling in some part of the body occur [19]. According to the affected region of the brain, epilepsy can be classified into two types namely focal and generalized epilepsy. Abnormal EEG of focal epilepsy is in a specific area whereas abnormal EEG in generalized epilepsy is in the entire area. We can localize abnormal regions of the brain by analyzing the EEG data thus seizure detection is important for the treatment of epilepsy [45]. Epilepsy is the most common disorder and affects approximately 1% of the world population [18].

Electroencephalogram (EEG) technique enables us to record neural level activity of our brain [34]. As compare to (ECoG), EEG technique is a non-intrusive method because electrodes are directly placed on the scalp of a subject to acquire the temporal and spatial information of the brain [47]. Generally, EEG test is the recording of brain electrical activity. It can be classified into two types namely scalp and intracranial EEG [9, 47]. Scalp EEG is the recording when electrodes are attached to the scalp and intracranial EEG is the recording of implanting electrodes in the brain during a surgery [9, 47]. Scalp EEG data is commonly used because it is easy to observe than intracranial EEG data although it consists of noise from the scalp and many other external factors that can interfere with seizure detection [9, 47].

A person's brain activity changes when he has an epileptic seizure [15]. This change is known as epileptiform brain activity and can be seen in the EEG recordings [15]. EEG signals have both nonlinear and non-stationary properties that's why they are very complex so in order to perform classification, we should reduce the complexity of signals by normalizing, reducing dimension or channels [20, 32].

In the field of EEG signal analysis, various techniques have been presented for seizure detection using widely known machine learning approaches, some of the approaches using Decision Tree (DT), Logistic Regression (LR), and Convolutional Neural Network (CNN) are mentioned as follows to understand the impact of the present study. In the aspect of the Decision Tree (DT) classification techniques, we have surveyed few papers where DT technique has been considered for seizure detection. In [42] authors have introduced the concepts of asynchronous tree by implementing gradient boosted trees with the help of DT technique to propose an efficient hardware architecture for biomedical applications. Authors have mentioned that the proposed architecture has been tested for automated seizure detection and achieved an average F1 score of 99.23% and 87.86% with an average detection latency of 1.1 s over the random and block-wise splitting of data into train/test sets. In the detection of epileptic seizure, EEG signal features and robust machine learning techniques are very important. With a similar view, in [28] authors have used time and frequency domain feature extraction techniques to extracted the valuable information from EEG data. Authors have considered K-nearest neighbors (KNN), Decision Trees (DT) and Support Vector Machine (SVM) with a linear kernel to compare the seizure detection performance and robustness of these classification techniques. Author mentioned that the DT based classification technique achieved an overall accuracy of 98.5% and SVM achieved the highest performance with an overall accuracy of 99.0%. Time-frequency analysis and feature extraction have been frequently utilized in several seizure detection approaches. As like in [10], authors have used Discrete Wavelet Transform (DWT) to analyze and select the informative EEG sub-bands from the input EEG data. Over the selected sub-bands authors have construct a fuzzy membership function and after that they applied several techniques like associative Petri net (APN), Decision Tree, Support Vector Machine, Neural Network, Bayes net, naive Bayes, and tree augmented naive Bayes for the diagnosis of epilepsy. Authors have claimed that the associative Petri net (APN) achieved diagnosis accuracy rates of 93.8%. Other than DWT based sub-bands selection, some other decomposition based approaches have been frequently used in seizure detection. A Dynamic Mode Decomposition (DMD) based seizure detection approach has been proposed in [43], where the authors mentioned that the Random Undersampling Boosting (RUSBoost) decision trees achieved the highest performance in seizure detection. Other than signal decomposition based approaches, signal transformation domain (time domain to frequency domain and its vice versa) analysis has been frequently used to propose seizure detection approach. In [36] authors have used Fast Fourier Transform (FFT) for extracting time domain features from the input EEG signals and applied decision tree classifier for seizure detection. Authors have claimed that the proposed approach achieved classification accuracies of 98.68% and 98.72% using 5- and 10-fold cross-validation.

In the aspect of Logistic Regression (LR) classification techniques, we have surveyed few papers where LR technique has been considered for seizure detection. In [48], authors have used Logistic Regression model to propose an automated detection approach for postictal generalized EEG Seizure. Similarly, in [8] a technique has been presented for automatic diagnosis of an epileptic seizure where Logistic Regression has been tested with the proposed approach and other four classifiers. Authors have mentioned that first they have applied discrete wavelet transform (DWT) for sub-band selection from EEG signals and after then they have extracted entropy based nonlinear dynamic features from the selected sub-band. Finally, authors have fed the extracted features into six classifiers namely, logistic regression, random forest, K-nearest neighbors, linear discriminant analysis, Naive Bayes classifier, and least square support vector machine (LS-SVM) to execute the classification task. Authors have claimed that the proposed approach with LS-SVM has achieved an accuracy of 99.50%, specificity of 99.40% and a sensitivity of 100.00%, whereas proposed approach with logistic regression achieved an accuracy of 99.00%, specificity of 98.00% and a sensitivity of 100.00% which is also acceptable.

Other than automated detection approach, a simple Epileptic Seizure classification approach has suggested in [30] using time-frequency analysis of the EEG signal. In this experiment author have mentioned that they have used multiscale radial basis functions (MRBF) and a modified particle swarm optimization (MPSO) to enhance the time-frequency data. Finally, an enhanced time-frequency data given into five types of classifiers where Logistic Regression is one of them and performance have been estimated for seizure and non-seizure signal classification. Authors have claimed that the proposed approach with SVM achieved 100% accuracy whereas Logistic Regression achieved 98.00% which is also acceptable performance for seizure detection. Other than with optimization techniques, Logistic Regression has used directly in seizure detection over some feature of EEG signals data. In

[39], a Denoising Sparse Autoencoder (DSA) based seizure detection approach has been proposed, where authors have mentioned that the DSA has been trained over the preprocessed EEG data and after then Logistic Regression model has been used on the top layer of DSA to perform the seizure detection task. Author claimed that the proposed approach achieved an average sensitivity of 100%, specificity of 100%, and recognition of 100% after some post-processed. In the similar fashion, Logistic Regression technique has been directly applied to proposed a seizure detection approach which is described in [46], where authors have used lifting-based discrete wavelet transform (LBDWT) to extract coefficients from the EEG signal. Finally, authors have fed the extracted features into the Logistic Regression and multilayer perceptron neural network (MLPNN) classification model to perform the classification task. Author claimed that the proposed approach with MLPNN has achieved an area under a ROC-AUC curve value of 0.902, specificity of 92.3% and a sensitivity of 92.8% whereas proposed approach with Logistic Regression has achieved an area under a ROC-AUC curve value of 0.853, specificity of 90.3% and a sensitivity of 89.2%.

The above mentioned approaches are based on the feature extraction and uses traditional machine learning approaches. Now a days, several automated feature extraction and processing techniques have been enabled in advanced machine learning techniques with the help of convolution techniques. For example, in [1] an automatic Seizure detection approach has been proposed using Deep Learning, where authors claimed that their approach has minimal involvement in EEG signals pre-processing and it has automatic feature learning capabilities. So, advance machine learning reduces the involvement of traditional features extraction process and has been applied for seizure detection with some data preprocessing techniques.

In the aspect of Convolutional Neural Network (CNN) based classification techniques, we have surveyed few papers where CNN technique has been considered for seizure detection. In [35], an image-based seizure detection approach has been proposed where the authors have mentioned that they have used 3D Convolutional Neural Networks. Authors have claimed that the proposed approach achieved a sensitivity of 85.7%, a false prediction rate of 0.096/h. Instead of image-based techniques, a Multi-View Convolutional Neural Networks has been used in [31] to propose seizure detection approach. Where authors have mentioned that they have considered the time domain and frequency domain features of EEG signals to deliberate two types of view of the input data. After time domain and frequency domain preprocessing authors have fed the data into their proposed Multi-View Convolutional Neural Network framework to detection the occurrence of seizure in epileptic EEG data. Authors have claimed that the proposed approach achieved an average area under the curve (AUCs) value of 0.82 and 0.89 over two subject's EEG data. Another seizure detection approach has been proposed in [17], where authors have used plot EEG images for seizure detection. In this approach authors have mentioned that they have used some preprocessing steps over the EEG data, in this step they have done segmentation process using a time window technique and after then the segmented data have been converted into plot EEG images. Processed plot EEG images have been given into the Convolutional Neural Network (CNN) for seizure and Non-Seizure classification. Authors have claimed that the proposed approach with CNN achieved a 100% of median seizure detection rate by minutes. Similarly, another approach based on spectrographic images of EEG and CNN has suggested in [51] to detect seizure from EEG signals. In this approach, authors have first converted the individual subject's EEG signals data into spectrograms. After then the spectrograms have been segmented into 26,380 total images using a windowing technique. Finally, authors have uses CNN based on VGG-net architecture for seizure and Non-seizure classification. Authors have claimed that the proposed approach achieved a seizure detection sensitivity and specificity of greater than 90%. In the similarly fashion, in [29] a seizure detection approach based on spectrograms and 1D-CNN has been proposed. Authors have claimed that the proposed approach achieved an average accuracy 77.57% along with a 4.74 average positive likelihood ratio and 0.32 average negative likelihood ratio.

In the aspect of a computer-aided diagnosis an approach based on CNN has been proposed in [2], where authors have mentioned that they have uses 13-layer deep convolutional neural network to detect preictal, normal, and seizure classes in epileptic EEG signals. The authors of [2] have claimed that the proposed approach achieved an accuracy of 88.67%, a specificity 90.00%, and a sensitivity of 95.00%. In the view of designing hardware-friendly seizure detection framework an Integer Convolutional Neural Network based seizure detection approach has been presented in [50]. The authors of [50] have claimed that the proposed approach has only 2% drop of accuracy.

In above descriptions, we have mentioned several seizure detection approaches based on Decision Tree (DT), Logistic Regression (LR) which are combined with features extraction techniques to propose seizure detection approach, whereas Convolutional Neural Network (CNN) based approaches are without involvement of traditional features extraction techniques. On the other hand, input EEG signal highly inconsistent and nonlinear in nature which leads massive impact on the classification models. In above literature, inconsistency and nonlinearity of EEG signals aren't been considered as they are used Decision Tree (DT), Logistic Regression (LR) and Convolutional Neural Network (CNN). Instead of DT and LR, CNN has now more popular in seizure detection but CNN has few issues with EEG data which are: CNN takes a lot of data to generalize as it has to learn different filters for each different viewpoints and CNN are Translation Invariant means ConvNets are unable to identify one object 's location with respect to another [22]. In the view of mentioned issues, research community has considering capsule neural network (CapsNet). CapsNet is more robust to change data orientation and size which is one of the most important aspects of seizure detection with inconsistent EEG signals and it has needed much less data and internal representation to classify the occurrence classes [41].

Research community accepts advancement in CapsNets and uses in several areas of EEG signal analysis. We have found few papers in which CapsNets has been used for different types of classification. In the aspect of CapsNets, we found few research papers where CapsNet has been used in Emotion Recognition and Motor Imagery classification through EEG signals. In [26], an approach has been proposed using CapsNet for motor imagery classification. The authors of [26] have mentioned that first they have used short-time Fourier transform (STFT) to convert EEG signal data into 2D images. Finally, 2D images have been fed into the CapsNet for the motor imagery classification. Authors have claimed that the proposed approach achieved an average accuracy of 78.44%. Similarly, in [24] an approach has been proposed using CapsNet for Emotion Recognition. In this experiment author has extracted granger causality feature from the original EEG signals. After then they have made a subset of high relevance features by using sparse group lasso algorithm. Finally, high relevance features set has been fed into the CapsNet to perform emotion classification. In the similar fashion, an approach for emotion recognition has been proposed in [6] where the authors have extracted power spectral density (PSD) from original EEG signal as a frequency domain features. In the next step, PSD of each channel has been categorized into four parts theta, alpha, beta and gamma, after then a multiband feature matrix (MFM) has been generated. Finally, the MFM is given into the CapsNet for emotion recognition. The authors of [6] have claimed that the proposed approach achieved an average recognition accuracy of 68.28%.

After literature review, we found electroencephalogram (EEG) is highly inconsistent in nature, even EEG has recorded from the same person, are not consistent and can be significantly different. In this particular, classification model probably compromised with the performance in seizure detection in epileptic EEG data. To give a better result in terms of accuracy, CNN trains on all the possible combination of input data by using image augmentation technique, where Capsule Neural Network does not need all the combinations of data. CNN uses Max-pooling, which loses spatial information by reducing spatial resolution, Capsule Neural Networks use Dynamic Routing in layers to pass the information from one layer to the next layer Particularly, so CNN based architectures other than CapsNets has compromised classification accuracy when target data are significantly different.

The main motivation and contribution of this work are to use the Capsule Neural Network (CapsNet) to study the subject specific and cross-subject seizure detection performance over the inconsistent behavior of EEG signals. In the experiment first, we normalize the data using L2 normalization, then pass the data into our proposed CapsNet model. In this experiment we have used two datasets from which one is multi-channel (CHB-MIT Scalp EEG dataset [7]) and the other one is a single channel (University of Bonn EEG dataset [14]) EEG dataset. The subject specific and cross-subject seizure detection performance of the proposed approach has been evaluated and compares with traditional (DT and RL) and deep learning (CNN) based classification models. More specifically we highlighted on 1) CapsNet based approach with dynamic routing has been investigated for Seizure detection. 2) Subject specific and cross-subject training and testing have been performed to check the robustness of the proposed approach. 3) Single and Multi-channel both datasets have been considered for this experiment. 4) Also, performances of different classifiers were estimated for seizure detection and compared.

Significance of this work After performing review of the related work, it has been found that a recent work which is based on preictal and inter-ictal EEG signals classification using 1D-CapsNet [49]. But in [49], all EEG channels have not been used for classification and also has not been studied cross-subject training and testing which is our main objective in the study. In our present study, we studied cross-subject as well as subject specific seizure and non-seizure classification experiment by considering all EEG channels of the multi-channel dataset and a single channel dataset separately. So, this study is equally a value addition with respect to the existing related work.

This present study has been presented using five major sections: Section 2 describes Materials and Methods which are used in the present study, Section 3 describes the Proposed CapsNet based Seizure Detection Approach, Section 4 describes the Results and Discussion of the proposed approach, Section 5 shows a comparison with other Schemes and Section 6 Concludes this present study with the aim of extending future scope of this work.

2 Materials and methods

In this section we have described the experimental datasets and technical background of this experiment. ***Subsection 2.1*** describes the experimental datasets, ***subsection 2.2*** describes the input data preprocessing approaches and ***subsection 2.3*** describes the classification techniques which are used to propose the present approach and compare the classification performance.

2.1 Experimental datasets

There are various trial EEG datasets available with different research groups which can be used in seizure detection. In this experiment, we have tested our proposed approach over the two benchmark datasets. Dataset-1 [7] is a multi-channel EEG dataset taken from the CHB-MIT scalp EEG dataset and Dataset-2 [14] is a single channel EEG dataset taken from the University of Bonn EEG dataset. More detailed descriptions about the dataset mentioned in ***subsections 2.1.1*** and ***subsection 2.1.2***.

2.1.1 Dataset-1

Dataset-1 has been taken from CHB-MIT Scalp EEG dataset [7] which is multi-channel and massively appreciated and used by the epileptic seizure analysis research community. As per data description, this EEG recordings were collected from 22 subjects and grouped into 23 cases (chb01, chb02… and so on). This dataset is of 17 females of age 1.5 to 19 years and 5 males of age around 3 to 22 years. The sampling frequency of all the recordings was 256 Hz. The duration and number of seizure events varied from each subject. A total of 25 electrodes were used for each subject in the dataset but they varied from 23 to 28 electrodes and depended on the subject, where 18 electrodes are common to everyone. Detailed information is available on [7], out of 23 cases, we used the data of first five cases to test the propose approach. EEG signal data of 23 channels from the first case (chb01_03) has been plotted and shown in Fig. 1, where ch represent the channel.

2.1.2 Dataset-2

Dataset-2 has been taken from the University of Bonn EEG dataset [3, 14]. A summary of the dataset is given in Table 1. This time series data is under a spectral bandwidth of 0.5 Hz to 85 Hz. This dataset is obtained from five sets (A to E) which are prepared from the EEG recordings of five subjects. For each set, dataset contain 100 single channel EEG recordings. Each recording of the EEG signal is of time duration of 23.6 seconds and sampling rate of

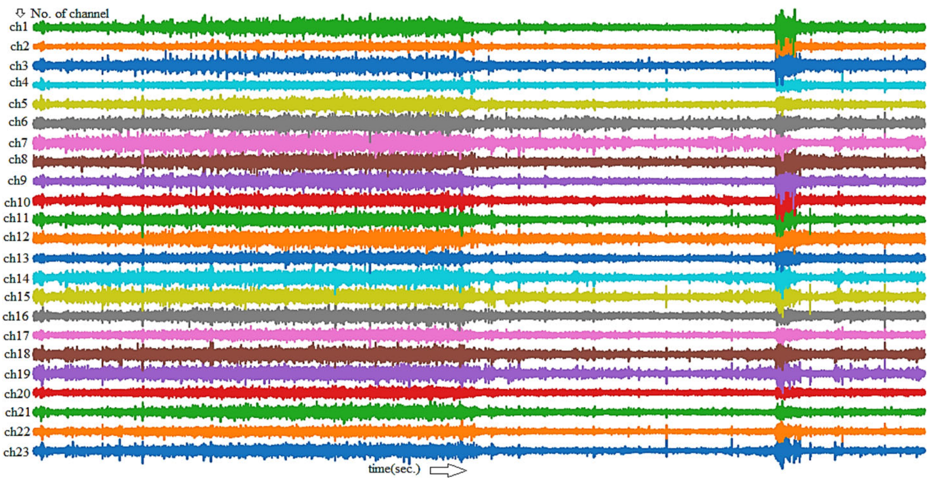

Fig. 1 EEG signals of all channels from Chb01_03

Table 1 Dataset folder description of University of Bonn EEG dataset

Data folder A	Data folder B	Data folder C	Data folder D	Data folder E
Non-Epileptic EEG Data	Non-Epileptic EEG Data	Epileptic EEG Data (Recorded from epileptogenic zone)	Epileptic EEG Data (Recorded with hippocampal formation of the opposite hemisphere of the brain)	Epileptic EEG Data
With Eye Opened	With Eye Closed	Inter-ictal (Non-Seizure) Stage	Inter-ictal (Non-Seizure) Stage	Ictal (Seizure) Stage

173.61 Hz. The normal dataset (Set A and B) has signals of healthy subjects and contains 100 cases. Inter-ictal datasets (Set C and Set D) have EEG signals of five epileptic subjects when they did not have seizure intervals. Particularly Set C EEG data from within the epileptogenic zone and Set D EEG data the hippocampal formation of the opposite hemisphere of the brain [3]. The Ictal dataset (Set E) contains 100 cases with the same number of subjects who had epilepsy with active seizure intervals. Signals with seizure (ictal) and seizure-free (non-ictal) from the Dataset-2 are presented in Fig. 2.

2.2 Input data preprocessing

In this section we have described the experimental data preprocessing strategy that has been followed in this present study for both experimental datasets.

2.2.1 Preprocessing over Dataset-1

Dataset-1: the EEG signals taken from CHB-MIT scalp EEG dataset [7] are in European data format (.edf). Time references for the duration of seizures are available within the dataset summary. Based on the summary, first for each subject all the seizure data have been trimmed out from the files (.edf) and converted into CSV format. Similarly, for each subject, all the non-seizure data have been trimmed out and converted into CSV format. Then label with 0 representing non-seizure data is added to the CSV file. Then corresponding label with 1

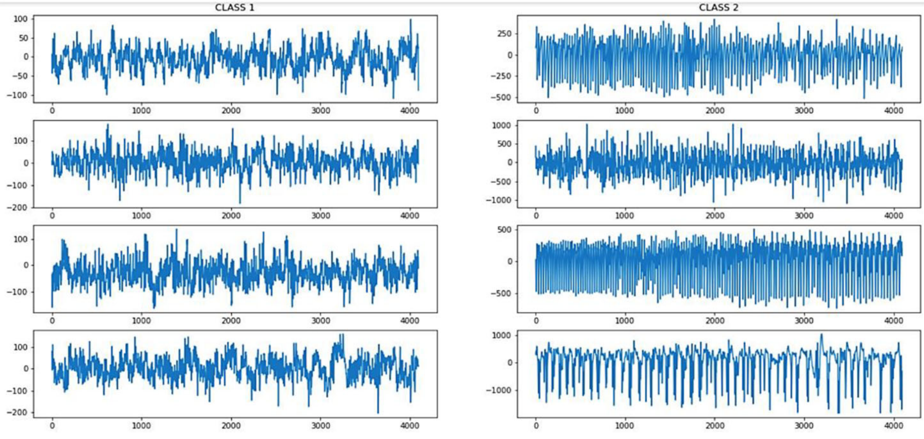

Fig. 2 Seizure (ictal) vs Seizure-free (non-ictal) EEG signals from Dataset-2, signals amplitude is in uV

representing seizure data is added for seizure. The resulting data were normalized using the L2 normalization technique and then split into training and testing data. Normalized data have been fed into the proposed approach. We observed by feeding normalized data into CapsNet that total loss (Eq. 12) of the model has been reduced. EEG signal data distribution of subject-1 of Datatset-1 has been shown in Fig. 3 after normalization process.

2.2.2 Preprocessing over dataset 2

Dataset-2 from the University of Bonn, Germany [14] is also used in this study. We have considered eye-opened non-epileptic data and Ictal epileptic data in this present study for seizure detection. Considered single channel time series data has been divided into 173 segments each of 1 sec as the EEG signal sampling rate is 173.61 Hz (we haven't considered remining fraction), then we normalized the data using L2 normalization. After performing normalization process, we have plotted the EEG signal data from data folder E which has been shown in Fig. 4. As showed in Fig. 4, there are not many changes in the distribution after normalization, as we used Euclidean norm to normalized the data.

2.3 Background

Logistic regression Logistic regression [12] is the most-simplest and statistical based machine learning algorithm which is based on sigmoid function and its outcome is related to explanatory variables. Sigmoid function in a S-shaped function that can take a real value and map it in the range of 0 to 1. Logistic regression is easy to implement and easy to train. Generally, seizure detection in epileptic EEG signal can be viewed as a binary classification such as seizure denoted as class 1 and seizure free denoted as class 0. Mathematically, logistic regression can be written with a probability (P) of input m and parameter δ.

$$P(n|m,\delta) = f_\delta(m)^n(1-f_\delta(m))^{1-n} \tag{1}$$

Where, n denotes the label and $n \in (0\ or\ 1)$, and $f_\delta(x) = \frac{e^{\delta^T m}}{1+e^{\delta^T m.}}$

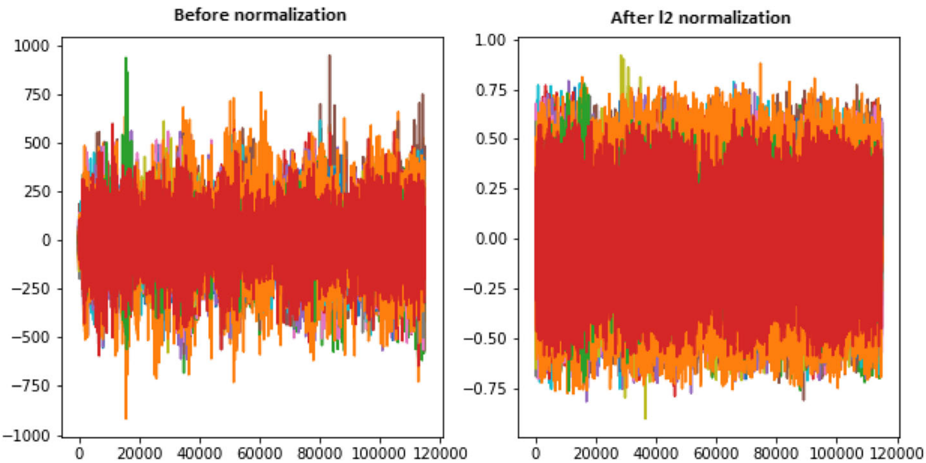

Fig. 3 EEG data distribution before and after normalization of seizure data. Subject-1 EEG data has been considered in this plot; different color indicates signals of the different channels

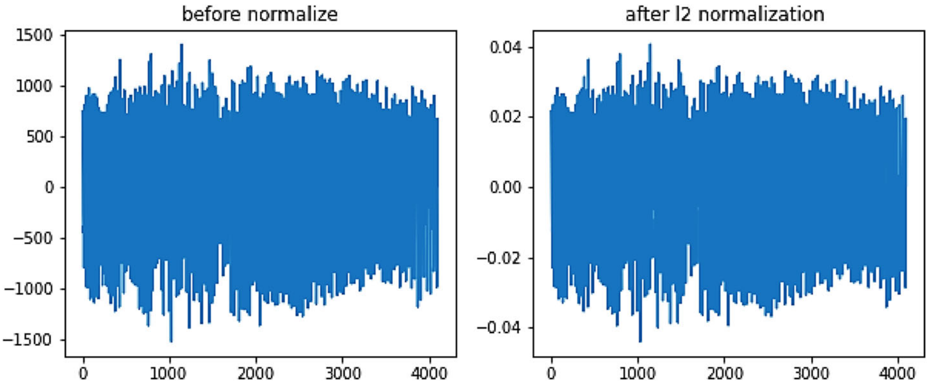

Fig. 4 Sample EEG Data distribution of Dataset-2 before and after performing L2 Normalization

The main purpose of training logistic regression is to generate the optimal parameter δ with the training data by minimizing the **log** likelihood function mentioned below in Eq. 2.

$$L(\delta) = -\frac{1}{s}\sum_{i=1}^{s} y_i \log f_\delta(m_i) + (1-n_i)\log(1-f_\delta(m_i)) \tag{2}$$

With the gradient descent method:

$$\frac{\partial}{\partial\delta}L(\delta) = \sum_{i=1}^{s}(y_i - h_\delta(m_i))m_i^{(j)} \tag{3}$$

At the time of obtaining the optimal δ, the predicted labels P_i could be gained by calculating the posterior probability $P(n|\ m,\ \delta)$ as per Eq. 1. If $P(n|\ m, \delta) > 0.5$ then m$\in$ *class* 1, otherwise $m \in$ *class* 0. This mathematical description of logistic regression has been taken from the paper [39].

Decision tree classifier Decision tree uses predictive modeling approaches in statistics and machine learning [40]. They perform classification without much estimation and they can handle both continuous and categorical values. Decision tree is nothing but a complex and nested if else statement. The whole dataset is fitted into 1 condition and split accordingly. More specifically, Decision tree classifier find the similar pattern from the input dataset and discriminate into different classes [28]. Mathematically, a Decision tree algorithm can be represented using the following equations.

$$\overline{A} = \{A_1, A_2, A_3, \ldots A_m\}^T \tag{4}$$

$$A_i = \left\{a_{i1}, a_{i2}, a_{i3}, \ldots a_{ij}, \ldots a_{in}\right\} \tag{5}$$

$$B = \{B_1, B_2, B_3, \ldots, B_i \ldots B_m\} \tag{6}$$

The main purpose of Decision tree is to predict the observations of $\overline{A}$. From $\overline{A}$ several Decision tree can be constructed with different accuracy level [28]. Here, m denotes the available observations number, n denote the independent variable number, S is the $m - dimension$

vector of the variable forecasted from $\overline{A}$. A_i is the i_{th} component of $n - dimension$ autonomous variables, $a_{i1}, a_{i2}, a_{i3}, \ldots, a_{in}$ are autonomous variable of pattern vector X_i and T is the transpose notation.

Convolutional neural network (CNN) Convolutional neural network is an advance version of a neural network which has convolutional operation instead of matrix multiplication in minimum one of their rows. Basic CNN model can be configured with one input layer, one output layer [21] and some hidden layer which consist of a chain of convolutional layers which convolve with multiplication. For down sampling, CNN uses a pooling function which focuses on important features but leaves the other features. So, feature loss is there in CNN with pooling layers.

Capsule neural network The Capsule Neural Network (CapsNet) [21, 40] is an extended version of CNN, as the cons of CNN is replaced by CapsNet. The pooling function by which feature loss is happening, is replaced by dynamic routing algorithm [41] which preserves all the features. A different activation function called Squash function [41] which is used in CapsNet to extract the non-linear features. Due to these two main reasons (Dynamic routing algorithm, squash function) the extra features such as rotation, position, angle, etc. can be preserved which helps the model in good prediction. In addition to these, CapsNet does not require a large amount of dataset. A summary of different input operation, output, and semantic diagram of input to output process within CapsNet architecture w.r.t a traditional neural network can be constructed from [16, 23, 27, 33, 41, 44] for understating the advancement of CapsNet.

A capsule is nothing but a batch of neurons whose length states the probability and activity vector which represents various properties of data such as rotation, position, direction, thickness etc. [27, 41]. In CapsNet, dynamic routing by agreement algorithm is used instead of max-pooling to achieve a reliable performance [41]. In addition to this, inverse rendering is the method on which CapsNet actually works [27].

In CapsNet, parent capsule's output is fed into the child capsule, and if the actual result of parent capsule is matched with the predicted output of child capsule, then the coupling coefficient increases. Let the output of the capsule i is u_i, now the prediction of the parent capsule j can be calculated using the Eq. 7.

$$\widehat{u_{j|i}} = w_{ij} u_i \tag{7}$$

Where, W_{ij} is a weight matrix. Now, we need to calculate coupling coefficient c_{ij} using the softmax function with the help of Eq. 8.

$$c_{ij} = \frac{\exp\left(p_{ij}\right)}{\sum_k \exp(p_{ik})} \tag{8}$$

Where, p_{ij} represent log prior probability for which capsule i coupled with parent capsule j and at the beginning $p_{ij} = 0$. Therefore, input vector (s_j) of capsule j (parent) can be calculated using the Eq. 9.

$$s_j = \sum_i c_{ij} \cdot \widehat{u_{j|i}} \tag{9}$$

Now, output (v_j) is estimated using a non-linear activation function namely squash function. The relation between output and squash function has mentioned in Eq. 10.

$$v_j = \frac{\left\|s_j\right\|^2}{1+\left\|s_j\right\|^2} \cdot \frac{s_j}{\|s_j\|} \tag{10}$$

After applying the squash function, the entire probability can be expressed in a range of 0 to 1. If capsule output is small, then it shrunk to zeros and if capsule output is large, it shrunk to one. The log probabilities (p_{ij}) are updating through n routing iteration with following agreement.

$$x_{ij} = v_j \cdot \widehat{u_{j|i}} \tag{11}$$

Finally, CapsNet total loss has been calculated using eq. 12,13 and 14 which are mentioned as follows.

$$Total\ Loss = F_{loss} + R_{loss} \tag{12}$$

$$F_{loss} = f_{loss}\max(0, y^{+}-||v_k||)^2 + \lambda(1-f_{loss})\max(0, ||v_k||-y^{-})^2 \tag{13}$$

$$R_{loss} = 0.0005{*}Mean\ Square\ Error\left(Reconstracted_{output}-Original\ Input\right) \tag{14}$$

Where, F_{loss} marginal of each class k, the y^{+} and y^{-} are hyper parameters set to 0.9 and 0.1 respectively. λ is used to decrease the effect of loss on class labels which don't belong to correct class and has value 0.5. If input data matches with its corresponding label, then f_{loss} is 1 else it is 0. On the other hand, R_{loss} is the reconstruction loss where $Reconstracted_{output}$ indicate the out of the decoder.

3 Proposed CapsNet based seizure detection approach

The proposed approach mainly contains three phases such as Phase1: Class level preprocessing, Phases2: Input preprocessing, and Phase3: CapsNet configuration, training and testing over the processed data. Each phase of the proposed approach has been shown in Fig. 5 and described below:

3.1 Class level preprocessing

In this phase we prepare the experimental EEG data into two classes namely seizure and non-seizure. The data preparation has been done for both the datasets (Dataset-1: multi-channel and Dataset-2: single channel) with the help of source data descriptions and the time stamp mentioned in the EEG recordings.

3.2 Input preprocessing (CapsNet input preprocessing)

In this phase data have been normalized and processed to feed into the CapsNet. Details of the individual steps have been mentioned as follows.

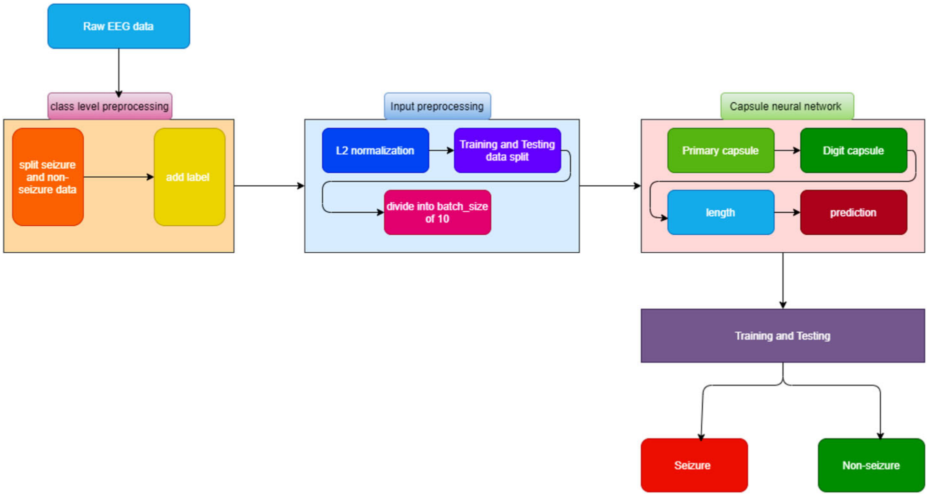

Fig. 5 Illustration diagram of the proposed approach

L2 normalization Input data have been normalized using normalize function [38]. In this work, we have chosen L2 normalization which use Euclidian norm and doesn't take additional input parameter to normalized. Whereas in Z-score normalization, the data is normalized with mean and standard deviation and gives us possible output likely to be gaussian distribution, but gaussian distribution is not required in our case and in Min-Max normalization, it cannot handle outliers, and prone to losing the information, also we have to provide a new min-max value to normalize the data.

Data Split into training and testing For subjects specific training and testing, the data have been divided into training and testing set in the ratio of 70:30. For cross subjects training and testing, 70% of the data is used for training and 50% of data is used for testing. More specifically, in cross subject training and testing session, training have been done over the 70% data of one subject and testing have been done over 50% data of another subject.

Batch size preparation Data have been divided into a batch size of 10 and after then data have been given to the CapsNet model to perform the seizure detection task.

3.3 CapsNet configuration, training and testing

In this phase, we have configured the various entities of the CapsNet model to perform the seizure detection task and to achieve robust detection performance. CapsNet configurations for this experiment have been described as follows.

Primary capsule layer Primary capsule layer contains two convolutional layers. The input to the first layer is a placeholder that will contain input data which we have fed at runtime. Over the processed data, first we have applied a 1D convolutional layer with kernel size three and sixty-four filters which gives the output in the shape of [batch_size,21,64]. After then we have passed the output into another 1D convolutional layer with the same kernel size and filter but with stride of two which gives the output in the shape of [batch_size,10,64]. Here, convolutional layer provides scaler output, but we need in vectors so we have done reshaping it to get eighty maps of eight dimensions and then we have applied squash function to ensure that their length is between 0 and 1.

Digit capsule This is the complex layer inside the CapsNet. In this experiment, we have considered two digit capsule units, to estimate the probability of “0” and “1” where “0” represent non-seizure and “1” represent seizure and both produce output of 16-dimension vectors. ***First Step:*** The first step is to calculate the predicted output vectors. as the first and second layer are fully connected to each other, so for all the pairs of the first and second layer, we find the predicted output vectors. In this computation we have used a transformation matrix W_{11} which has gradually been learned during training period. We multiplied the transformation matrix with the output of the first layer capsule $u_{1\ |\ 1}$ which is the estimated first digit capsule output dependent on the first primary capsule output. Since primary capsule gives the output of 8-dimensional vector and digit capsule generate the output of 16-dimension vector, the transformation matrix must be a 16*8 matrix. Next, we have predicted the second capsule output which depends on the output of the first primary capsule using a different transformation matrix W_{12}. Then we used a second primary capsule to estimate the output of the first digit capsule and do the same for the second digit capsule. Finally, we have 140 predicted output vectors.

Now we have a bunch of output vectors. ***Second step:*** in this we have applied routing by agreement algorithm which is described in Algorithm. 1, it gives the output of the digit capsule. In the routing algorithm, first we have initialized all routing weight to zero.

Algorithm. 1 Routing algorithm [41]

Inputs: $Input\ vectors \rightarrow u_{j|i}, No\ of\ routing\ iteration \rightarrow R_t, Capsules\ Layers \rightarrow C_l$

Outputs: $Digit\ capsule\ output \rightarrow D_j$

CapsNet Routing$(u_{j|i}, R_t, C_l)$
    Routing weight initialization: $w_{ij} \leftarrow 0$, ∀ capsule i in layer C_l and capsule j in layer $(C_l + 1)$
    for R_t iterations do
        $cc_i \leftarrow softmax(w_i)$, ∀ capsule i in layer C_l /* CC ← coupling coefficient, and SoftMax computes Equation. 8 */
        $s_j \leftarrow \sum_i CC_{ij}\widehat{u_{j|i}}$, ∀ capsule j in layer $(C_l + 1)$. /* s_j ← input vector of parent capsule j */
        $D_j \leftarrow squash(s_j)$, ∀ capsule j in layer $(C_l + 1)$ /* Squash compute Equation. 10*/
        $w_{ij} \leftarrow w_{ij} + \widehat{u_{j|i}} \cdot D_j$, ∀ capsule i in layer C_l and capsule j in layer $(C_l + 1)$.
Return D_j

Second, we have applied the softmax function in each primary capsule. Third, with the help of routing weights, we computed the predicted output vector for each of the digit capsule. Fourth, we have applied the squash function over output of the third step so that all the length bound between 0 and 1. This is only the first round of the routing algorithm, but we can run as many rounds as we want. In this proposed approach, we have applied two rounds of this algorithm. The output of this algorithm is the output of digit capsules.

Length of the output vector of a capsule The length of each output vector estimates the probability of seizure and non-seizure class. So, for computing the probability, we have used norm function which is given in Eq. 15.

$$||S_i|| = \sqrt{\sum_i (S_i * S_i) + \in} \tag{15}$$

Where, $\|S_i\|$ is digit capsules output, and here $\in$ is the small number when the output of digit capsule is zero then we don't get the length value zero. But the sum of probability doesn't have to add up to 1 because we have not used a softmax function.

Prediction For calculating the prediction, we have used argmax function which gives the index of the highest probabilities and the index is itself a label. Then we have applied a squeeze function to remove the extra dimensions from of argmax function and now we have a configured capsule network that has been used to estimate seizure and non-seizure class probabilities. Now, we have estimated detection performance of the CapsNet model by comparing predictions with labels. To understanding the layer wise input and output of CapsNet model, a summary of each layer has been shown in Table 2.

Loss For estimating total loss, we have calculated marginal loss and reconstruction loss. For margin loss, we have applied the Eq. 13 described in section 3. For reconstruction loss, we used the output of the digit capsule. First, we apply a mask before sending the output to the decoder because digit capsule's output must be masked out except for the one which relates to the target digit. The shape of the mask is same as the output of digit capsule and it is equal to zero all over except for the target digit location. The input to the decoder is the product of the mask and digit capsule. Decoder is a regular feed forward neural network composed of three fully connected layers. It outputs an array containing 23 values (for Dataset-1) and 173 values (for Dataset-2) from 0 to 1. Now, we have computed the reconstruction loss by multiplying 0.0005 (scale down factor) with the mean square difference between reconstructions and input using Eq. 14. The advantage of the scale down factor is that the reconstruction loss does not dominate the marginal loss during training [41]. After then we have added margin and reconstruction loss to compute the final loss of the CapsNet model. Then we do the training operation by using an Adam optimizer to minimize loss. For the testing, we won't have the labels so masked the output vector using predicted classes rather than the labels. An illustration diagram has been shown in Fig. 6 to understand the complete workflow of CapsNet with loss calculation.

Table 2 Layer wise input and output shape of CapsNet Model

Layer	Input shape	Output shape
conv1d_1 layer	[batch_size,23,1]	[batch_size,21,64]
Conv1d_2 layer	[batch_size,21,64]	[batch_size,10,64]
Squash	[batch_size,10,64]	[batch_size,80,8]
Digit capsule layer	[batch_size,80,8]	[batch_size,1,2,16,1]
Length	[batch_size,1,2,16,1]	[batch_size1,2,1]
Prediction	[batch_size1,2,1]	[batch_size1]

4 Result and discussion

In this experiment we have performed subject specific and cross subject training and testing experiment to estimate the robustness of our proposed seizure detection approach. In subject specific training and testing experiment, we have estimated the performances of Logistic Regression (RL), Decision Tree (DT), CNN and our proposed CapsNet based approach which is described in **subsection** 4.1. In Cross-subject training and testing experiment, we have estimated the performance of our proposed CapsNet based approach which is described in **subsection** 4.2. The performance of all classification models has been estimated using six standard performance evaluation parameters (mathematically described in [25, 37]) namely

$$\text{sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}},$$

$$\text{specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}},$$

$$\text{accuracy} = \left(\frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}}\right),$$

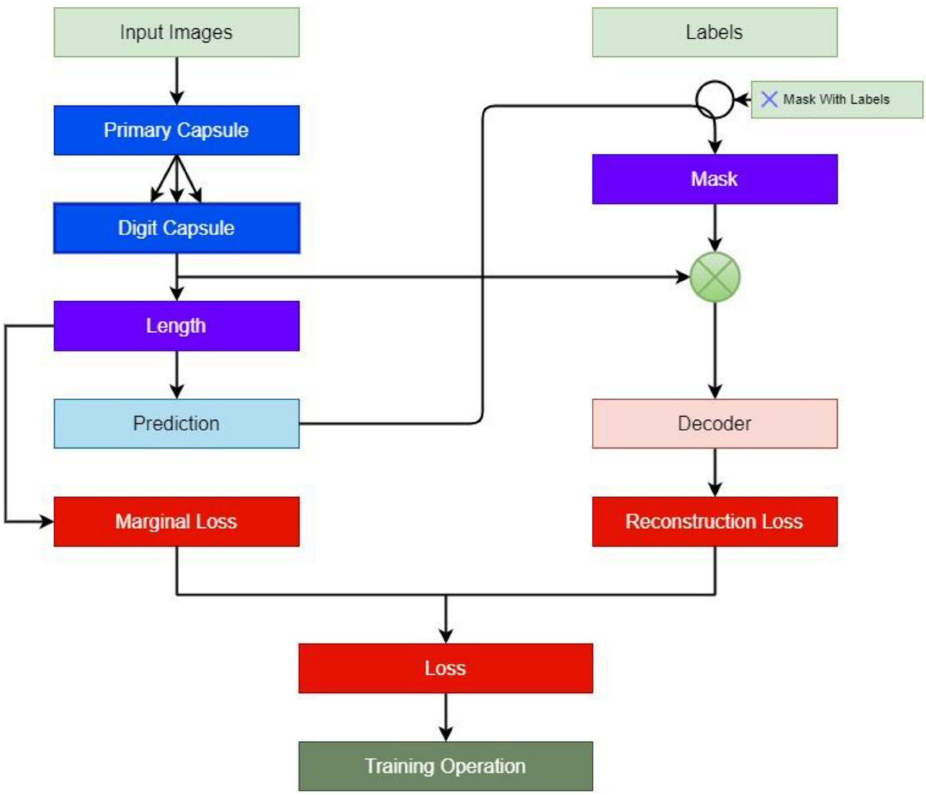

Fig. 6 Illustration diagram of steps involves inside the CapsNet model with loss calculation

$$F1\ Score = \frac{2TP}{(2TP + FP + FN)},$$

$$AUC\ Score = \frac{(TP + TN)}{(TP + TN + FP + FN)}, \text{and}$$

$$\text{False Positive Rate (FPR)} = \frac{FP}{FP + TN}.$$

Also, we have plotted ROC_AUC curve = True Positive Rate against False Positive Rate. Where, True Positive (TP) is the number of truly detected seizures; False Negative (FN) is the number of misclassified detected seizures; True Negative (TN) is the number of truly classified non-seizures; False Positive (FP) is the number of misclassified detected non-seizures [4, 48].

4.1 Results of subject specific experiment

In subject specific training and testing experiment, first we classified seizure and non-seizure EEG signals by considering traditional classification approaches using logistic regression and decision tree. Second, we performed our classification task using CNN technique. In the classification process, we have performed subject specific training and testing, and computed the performances over the Dataset-1 [7] and Dataset-2 [14]. The results have been shown in Tables 3 and 4, and It has been observed that for logistic regression achieved a mean accuracy of 64.292%, a mean sensitivity of 63.886%, a mean specificity of 66.428%, a mean FPR of

Table 3 Subject specific estimated performance for five subjects of Dataset-1

Approaches	Performance Evaluation Parameter	Subject1	Subject2	Subject3	Subject4	Subject5	Mean Performance
Logistic Regression	Accuracy	75.49%	52.80%	60.0%	79.29%	53.88%	64.292%
	sensitivity	72.22%	52.69%	59.81%	75.50%	54.21%	63.886%
	specificity	79.80%	52.91%	60.28%	85.59%	53.56%	66.428%
	FPR	0.2019	0.4708	0.3971	0.1540	0.4643	0.33762
	AUC score	0.7601	0.5280	0.6005	0.8005	0.5388	0.64558
	F1 score	0.7700	0.5294	0.6458	0.8097	0.5364	0.65826
Decision tree	Accuracy	81.48%	80.44%	84.62%	78.04%	83.68%	81.652%
	Sensitivity	82.95%	80.38%	85.71%	78.55%	82.19%	81.956%
	Specificity	80.02%	80.50%	83.45%	78.04%	85.18%	81.438%
	FPR	0.1997	0.1949	0.1654	0.2246	0.1481	0.18654
	AUC score	0.8149	0.8044	0.8458	0.7804	0.8369	0.81648
	F1 score	0.8167	0.8040	0.8528	0.7832	0.8351	0.81836
CNN	Accuracy	89.44%	85.91%	92.61%	89.56%	90.50%	89.604%
	Sensitivity	92.58%	82.84%	92.29%	88.58%	88.69%	88.996%
	Specificity	86.33%	88.90%	92.96%	90.55%	92.02%	90.152%
	FPR	0.1336	0.1102	0.0703	0.0944	0.0797	0.09818
	AUC score	0.8946	0.8590	0.9263	0.8957	0.9035	0.89582
	F1 score	0.8972	0.8544	0.9285	0.8955	0.9023	0.89558
CapsNet	Accuracy	98.17%	96.13%	93.47%	88.75%	91.00%	93.50%
	sensitivity	98.36%	95.78%	95.26%	86.24%	90.35%	93.19%
	specificity	97.98%	96.49%	91.66%	91.72%	91.67%	93.90%
	FPR	0.0201	0.0350	0.0833	0.0827	0.0832	0.05186
	AUC score	0.9817	0.9613	0.9353	0.8872	0.9099	0.93508
	F1 score	0.9818	0.9614	0.9361	0.8925	0.9112	0.9366

Table 4 Subject specific estimated performance over the Dataset-2

Performance Evaluation Parameter	Approaches			
	Logistic Regression	Decision tree	CNN	CapsNet
Accuracy	53.55%	65.02%	94.72%	82.61%
Sensitivity	52.54%	65.94%	94.39%	80.54%
Specificity	54.55%	64.13%	95.03%	84.83%
FPR	0.4544	0.3586	0.0496	0.1516
AUC score	0.5355	0.6504	0.9471	0.8266
F1 score	0.5299	0.6487	0.9460	0.8273

0.33762, a mean AUC score of 0.64558, and a mean F1 score of 0.65826 over the Dataset-1 and an accuracy of 53.55%, a sensitivity of 52.54%, a specificity of 54.55%, a FPR of 0.4544, an AUC score of 0.5355, and a F1 score of 0.5299 over the Dataset-2, whereas decision tree achieved a mean accuracy of 81.652%, a mean sensitivity of 81.956%, a mean specificity of 81.438%, a mean FPR of 0.18654, a mean AUC score of 0.81648, and a mean F1 score of 0.81836 over the Dataset-1 and an accuracy of 65.02%, a sensitivity of 65.94%, a specificity of 64.13%, a FPR of 0.3586, an AUC score of 0.6504, and a F1 score of 0.6487 over the Dataset-

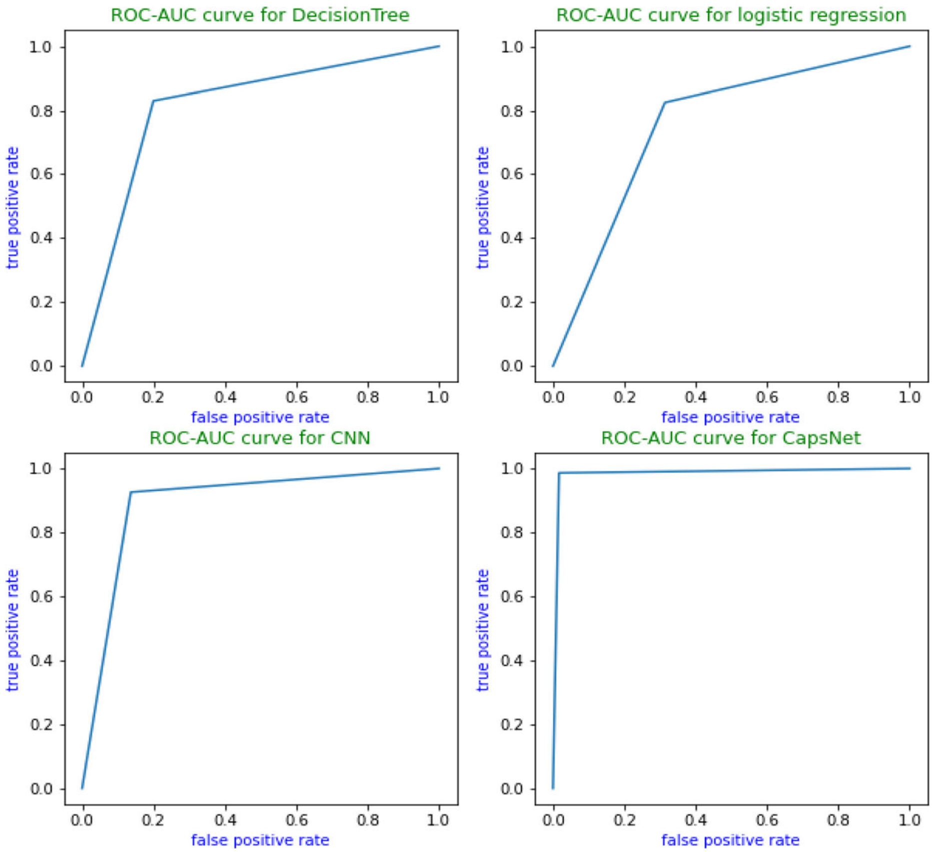

Fig. 7 ROC-AUC curve of RL, DT, CNN, CapsNet over the subject1 from Dataset-1

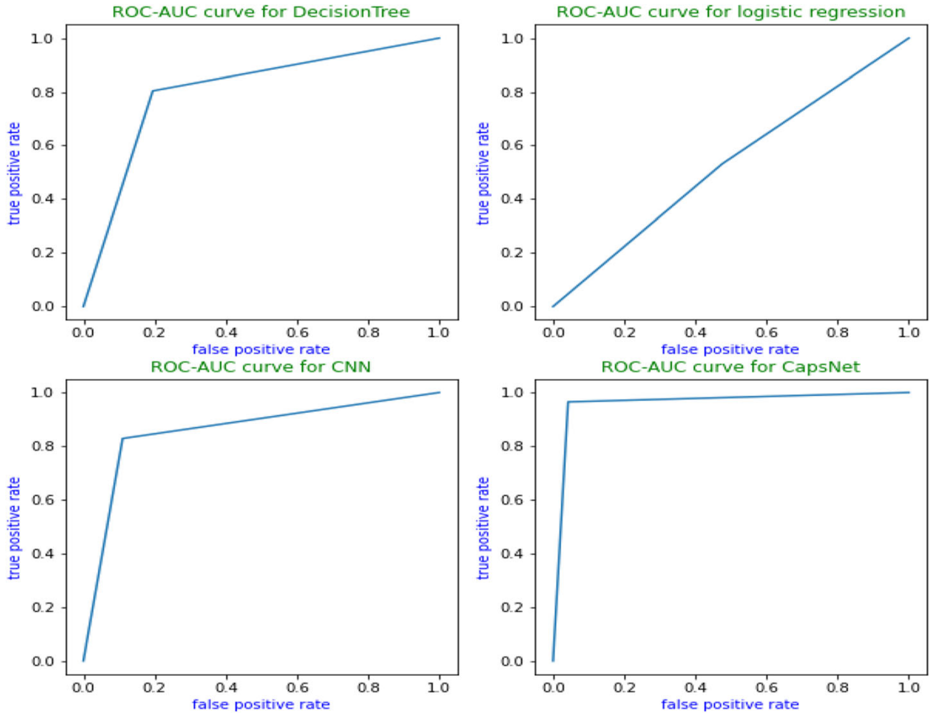

Fig. 8 ROC-AUC curve of RL, DT, CNN, CapsNet over the subject2 from Dataset-1

2. Which tells that the decision tree performs much better than logistic regression. Instead of traditional classification approaches, in this experiment we have tested the performance of CNN as CapsNet belongs to CNN family. The performance of CNN has been recorded also in Tables 3 and 4 over the Dataset-1 and Dataset-2 respectively. From Tables 3 and 4, it has been observed that CNN achieved a mean accuracy of 89.604%, a mean sensitivity of 88.996%, a mean specificity of 90.152%, a mean FPR of 0.09818, a mean AUC score of 0.89582, and a mean F1 score of 0.89558 over the Dataset-1 and an accuracy of 94.72%, a sensitivity of 94.39%, a specificity of 95.03%, a FPR of 0.0496, an AUC score of 0.9471, and a F1 score of 0.9460 over the Dataset-2. Which tells that the CNN performs much better than a decision tree. Finally, we have estimated the performance of our proposed CapsNet based seizure detection approach and it has been recorded in Tables 3 and 4 for the Dataset-1 and Dataset-2 respectively. From Tables 3 and 4, it has been observed that our proposed CapsNet based approach achieved a mean accuracy of 93.50%, a mean sensitivity of 93.19%, a mean specificity of 93.90%, a mean FPR of 0.05186, a mean AUC score of 0.93508, and a mean F1 score of 0.9366 over the Dataset-1 and an accuracy of 82.61%, a sensitivity of 80.54%, a specificity of 84.83%, a FPR of 0.1516, an AUC score of 0.8266, and a F1 score of 0.8273 over the Dataset-2. Which tells that our proposed CapsNet based seizure detection approach performs much better than CNN based approach.

The present study mainly focusses on the seizure detection performance obtained from our proposed CapsNet based approach. Seizure detection performances have been recorded and shown in Tables 3 and 4. It has been observed that our proposed approach achieved better results as compared to other classification techniques which considered in this experiment.

To present a close look of effectiveness of our proposed approach, we have plotted the ROC curve for each considered techniques over the individual subjects. The ROC curves of Logistic Regression, Decision Tree, CNN, CapsNet over the Subject1, Subject2, Subject3, Subject4 and Subject5 for the Dataset-1 have been shown in Figs. 7, 8, 9, 10, and 11 respectively. For Dataset-2, the ROC curves have been shown in Fig. 12. For all ROC curves, it has been observed that our proposed CapsNet based approach has shown consistence performance towards the seizure detection task in compare to other three classifiers.

In subject specific training and testing experiment, Table 3, Table 4, and its related ROC curves (Figs. 7, 8, 9, 10, 11 and 12) suggests that proposed CapsNet based approach can differentiate seizure epochs and non-seizure epochs effectively. And it shows that our proposed approach performs better than CNN performance.

4.2 Results of cross subjects experiment

In the Cross-subject training and testing experiment, to estimate the performance we have only considered our proposed CapsNet based approach as it has better performance in subject-specific training and testing experiments. We have done the cross-subject training and testing experiments to understand the robustness of the proposed approach over the unknown data which haven't been used for training. In this experiment we have considered five subjects from

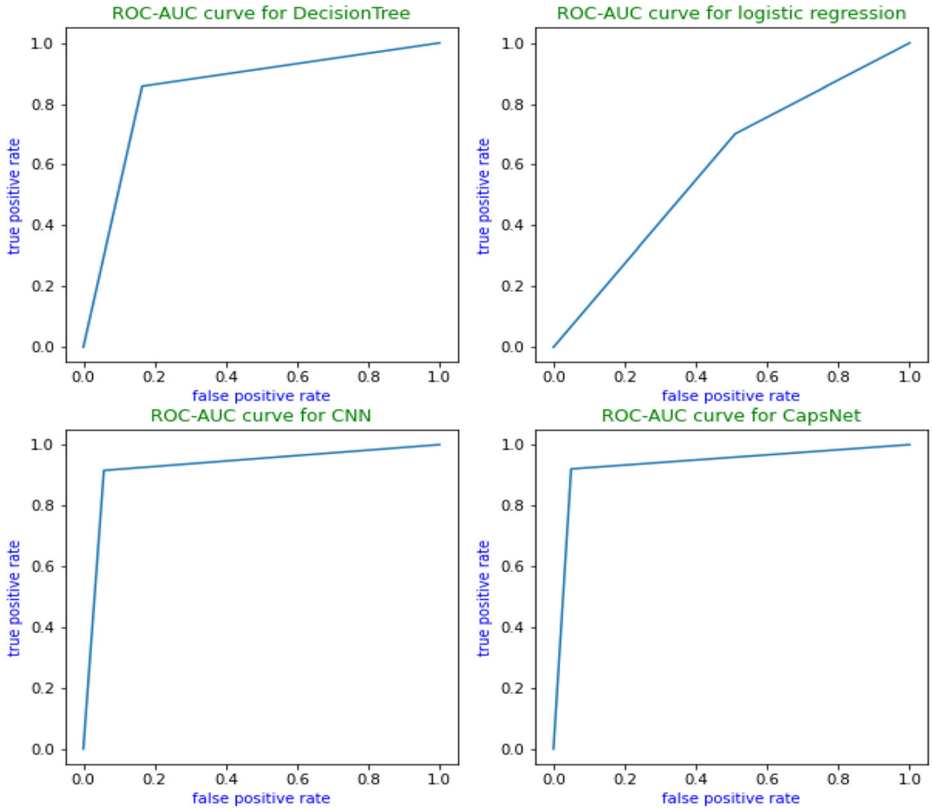

Fig. 9 ROC-AUC curve of RL, DT, CNN, CapsNet over the subject3 from Dataset-1

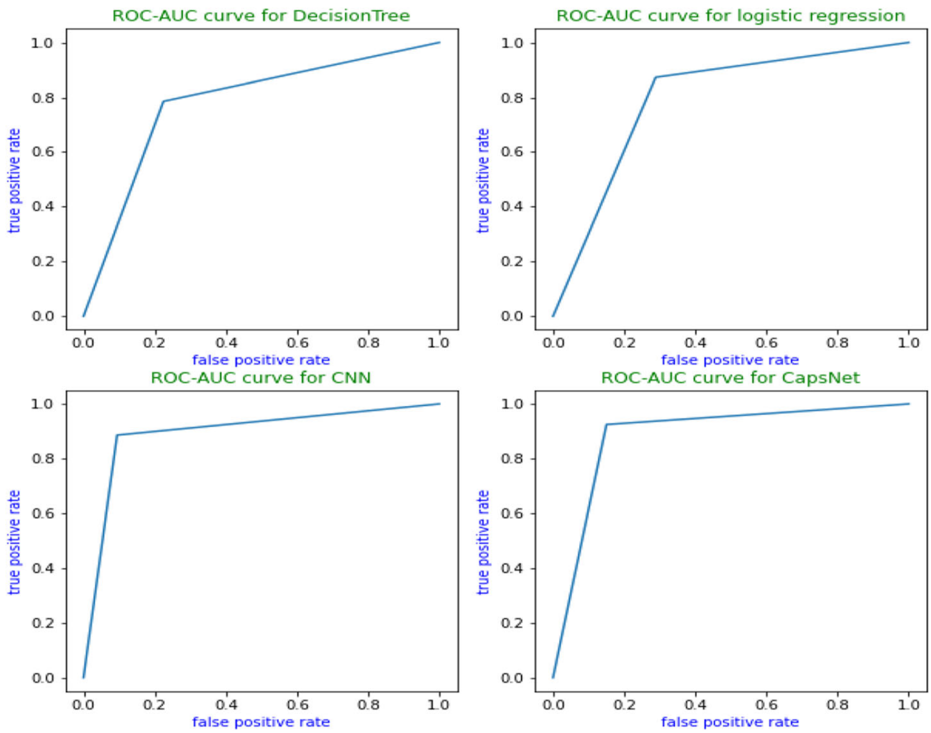

Fig. 10 ROC-AUC curve of RL, DT, CNN, CapsNet over the subject4 from Dataset-1

Dataset-1 and five data folders from Dataset-2. For estimating performance of the proposed CapsNet approach, we have considered five cases for Dataset-1 and two cases for Dataset-2. For each case in Dataset-1, we have trained the CapsNet over one subject EEG data and testing has been performed over the remaining four subjects individually. For Dataset-2, data folder A (Non-seizure) and E (Seizure) have been used for training, for testing folder (B and C), and (B and D) have been used. The performance of the proposed approach has been estimated and recorded in Tables 5 and 6 for Dataset-1 and Dataset-2 accordingly. From Table 5, it has been observed that our proposed approach achieved a mean accuracy of 85.21%, a mean sensitivity of 84.66%, a mean specificity of 93.14%, a mean FPR of 0.0684, a mean AUC score of 0.8525, and a mean F1 score of 0.8742 in Case-1; a mean accuracy of 86.41%, a mean sensitivity of 88.85%, a mean specificity of 91.40%, a mean FPR of 0.0857, a mean AUC score of 0.8653, and a mean F1 score of 0.8832 in Case-2; a mean accuracy of 72.37%, a mean sensitivity of 67.59%, a mean specificity of 98.95%, a mean FPR of 0.0103, a mean AUC score of 0.7232, and a mean F1 score of 0.7959 in Case-3; a mean accuracy of 76.52%, a mean sensitivity of 75.50%, a mean specificity of 95.98%, a mean FPR of 0.0400, a mean AUC score of 0.7898, and a mean F1 score of 0.8360 in Case-4; a mean accuracy of 56.07%, a mean sensitivity of 55.63%, a mean specificity of 56.98%, a mean FPR of 0.4297, a mean AUC score of 0.5600, and a mean F1 score of 0.6063 in Case-5 over the Dataset-1. Which shows that our proposed approach performed a consistent seizure detection over the Dataset-1. Also, we have estimated the performance of our proposed approach over the Dataset-2 and has been recorded in Table 6. From Table 6, it has been observed that our proposed approach achieved a mean accuracy of 48.45%, a mean sensitivity of 18.66%, a mean specificity of 78.25%, a mean

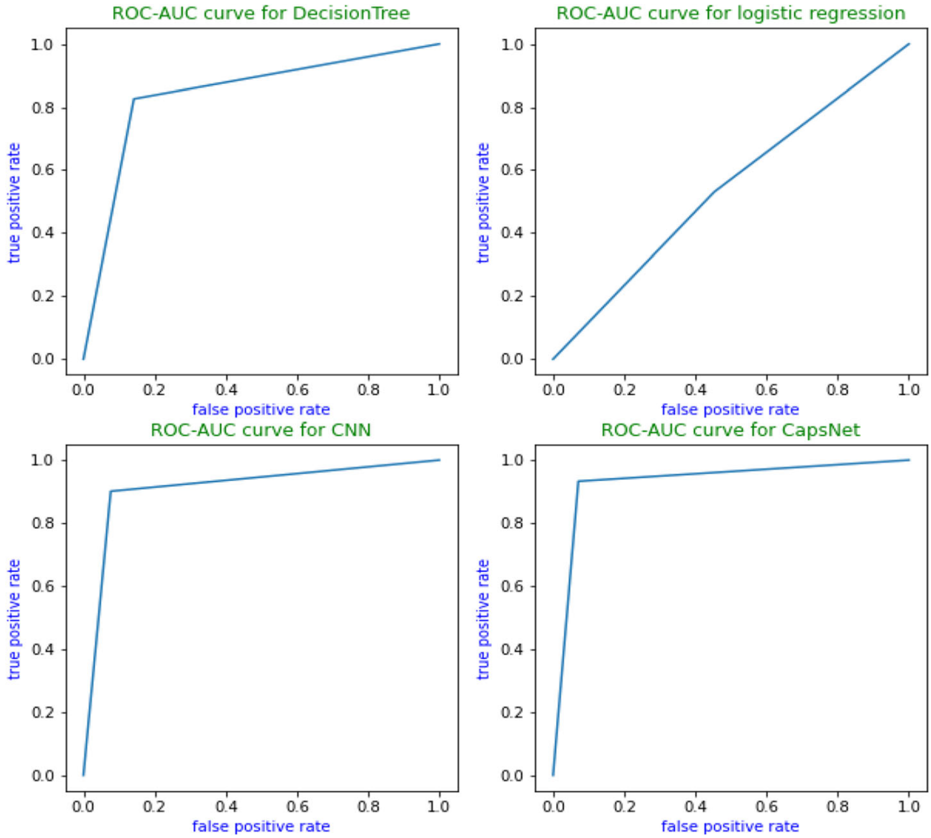

Fig. 11 ROC-AUC curve of RL, DT, CNN, CapsNet over the subject5 from Dataset-1

FPR of 0.2174, a mean AUC score of 0.4845, and a mean F1 score of 0.2653 over the Dataset-2. The performance of the proposed approach is not up to the marks because, in cross-subject training and testing experiment testing data has not sufficient seizure onset data. In subject-specific experiment we have already considered data folder A (having non-seizure data) and E (having Ictal or seizure onset data), so in cross-subject experiment remaining cases we can form is with data folder B (having non-epileptic data), C (having Inter-ictal data) and D (having Inter-ictal). For Dataset-2, in cross-subject experiment testing data don't exactly have seizure onset data but having inter-ictal EEG data. Due this type of data scenario our proposed approach has not shown high accuracy, it is indicating that either the proposed approach is underfitting or it is not blindly giving the high performance. This is an advantage of CapsNet. So, the performance of the proposed approach is appreciable for seizure onset detection.

Like in subject-specific experiment, here also we have plotted the ROC curve of the proposed approach over the individual subjects to present a close look of the effectiveness of our proposed approach. The ROC curves of CapsNet over the Subject1, Subject2, Subject3, Subject4 and Subject5 from Dataset-1 have been shown in Figs. 13, 14, 15, 16, and 17 respectively. For Dataset-2, the ROC curves have been shown in Fig. 18. For all ROC curves, it has been observed that our proposed CapsNet based approach has achieved a consistent performance in seizure detection task in compare to other three classifiers.

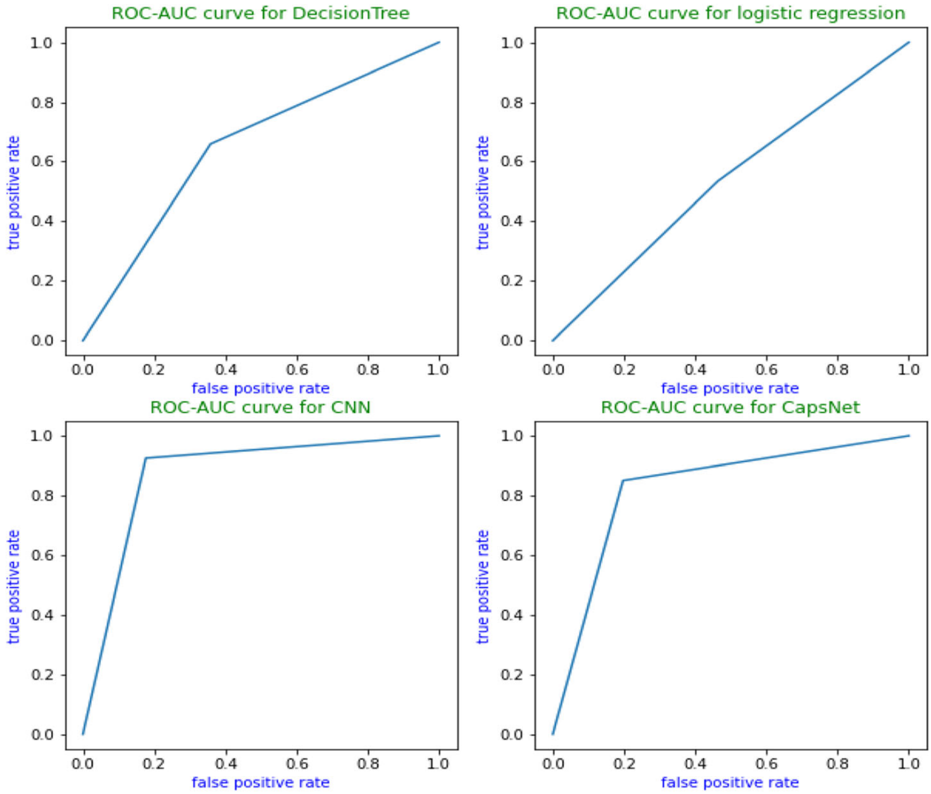

Fig. 12 ROC-AUC curve of RL, DT, CNN, CapsNet over the Dataset-2

So, the estimated results of our proposed CapsNet based seizure detection approach have shown better performance in compare to CNN, RL and DT. The result illustrated that the shortcoming in the CNN based seizure detection approach can be overcome with the help of CapsNet techniques, whereas CapsNet is at state-of-the-art phase for the detection of seizure. So, our proposed CapsNet based seizure detection has an impact in the diagnosis of epileptic seizure. In this present study, all experimental phases have been executed on Google Colab [5, 11] with Python7.3.

5 Comparison with existing related schemes

Several research papers have been published on seizure detection with some machine learning techniques as described in the introduction section. Our proposed approach based on CapsNet technique which is a special configuration of CNN architecture. As we have mentioned earlier that the CapsNet itself is at state-of-the-art phase for seizure detection, so we have not found related published articles where CapsNet technique has been applied directly for cross-subject seizure detection. However, we have found few articles where CapsNet has been used for Motor Imagery classification and Emotion Recognition over the EEG data. On this view, we have shown two comparison tables (Tables 7 and 8), where Table 7 shows a comparison with

Table 5 Estimated performance of proposed approach in Cross-subject experiment over the Dataset-1

Cases	Training Vs. Testing		Performance of proposed approach					
	Training Subject	Testing Subject	Accuracy	sensitivity	specificity	FPR	AUC score	F1 score
Case-1	Subject1	Subject2	87.08%	81.79%	94.44%	0.0555	0.8712	0.8804
		Subject3	91.21%	98.60%	85.20%	0.1479	0.9153	0.9092
		Subject4	95.38%	97.71%	93.66%	0.0633	0.9540	0.9536
		Subject5	67.17%	60.55%	99.29%	0.0070	0.6695	0.7536
Mean Performance			85.21%	84.66%	93.14%	0.0684	0.8525	0.8742
Case-2	Subject2	Subject1	95.58%	98.31%	93.60%	0.0639	0.9581	0.9569
		Subject3	88.23%	98.85%	80.70%	0.1925	0.8868	0.8742
		Subject4	95.24%	98.10%	92.63%	0.0736	0.9528	0.9517
		Subject5	66.60%	60.16%	98.68%	0.0131	0.6638	0.7500
Mean Performance			86.41%	88.85%	91.40%	0.0857	0.8653	0.8832
Case-3	Subject3	Subject1	79.39%	70.87%	99.13%	0.0086	0.7949	0.8277
		Subject2	64.44%	58.45%	98.31%	0.0168	0.6454	0.7364
		Subject4	94.08%	90.07%	99.15%	0.0084	0.9400	0.9444
		Subject5	51.60%	50.97%	99.22%	0.0077	0.5127	0.6752
Mean Performance			72.37%	67.59%	98.95%	0.0103	0.7232	0.7959
Case-4	Subject4	Subject1	80.94%	84.04%	98.20%	0.0179	0.8998	0.9069
		Subject2	75.37%	67.43%	96.24%	0.0375	0.7542	0.7987
		Subject3	93.80%	96.69%	90.99%	0.0900	0.9392	0.9389
		Subject5	56.90%	53.87%	98.52%	0.0147	0.5661	0.6997
Mean Performance			76.52%	75.50%	95.98%	0.0400	0.7898	0.8360
Case-5	Subject5	Subject1	55.15%	53.65%	58.25%	0.4174	0.5524	0.6169
		Subject2	54.81%	53.78%	56.49%	0.4350	0.5483	0.5966
		Subject3	59.10%	59.99%	57.98%	0.4201	0.5882	0.6245
		Subject4	55.22%	55.13%	55.22%	0.4465	0.5511	0.5874
Mean Performance			56.07%	55.63%	56.98%	0.4297	0.5600	0.6063

Table 6 Estimated performance of proposed approach in cross-subject experiment over the Dataset-2

Training	Testing	Accuracy	Sensitivity	Specificity	FPR	AUC score	F1_score
Dataset A and E	Dataset B and C	47.33%	16.42%	78.25%	0.2174	0.4733	0.2377
	Dataset B and D	49.57%	20.90%	78.25%	0.2174	0.4957	0.2930
Mean Performance		48.45%	18.66%	78.25%	0.2174	0.4845	0.2653

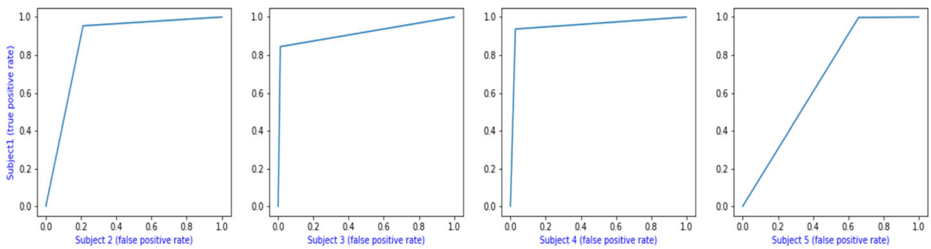

Fig. 13 ROC curve of CapsNet over the case-1 of Dataset-1

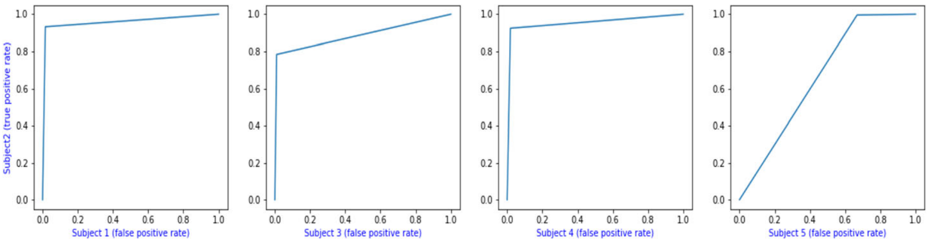

Fig. 14 ROC curve of CapsNet over the case-2 of Dataset-1

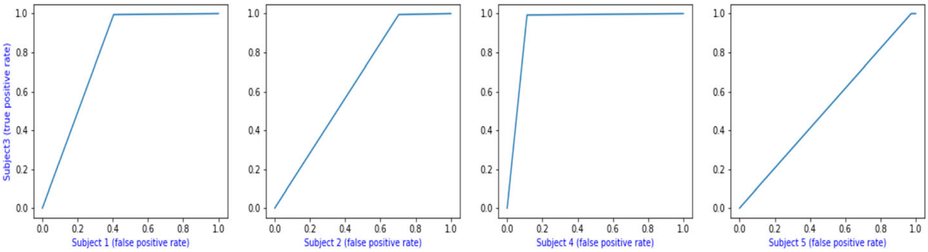

Fig. 15 ROC curve of CapsNet over the case-3 of Dataset-1

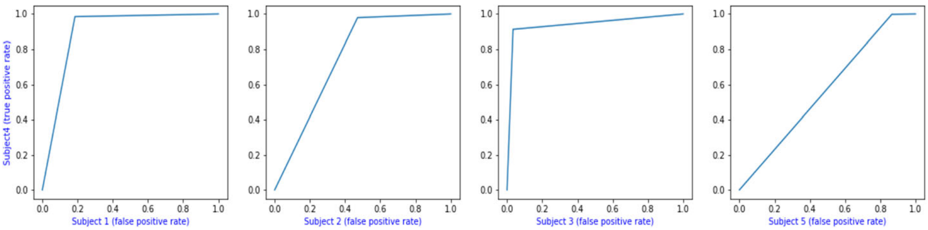

Fig. 16 ROC curve of CapsNet over the case-4 of Dataset-1

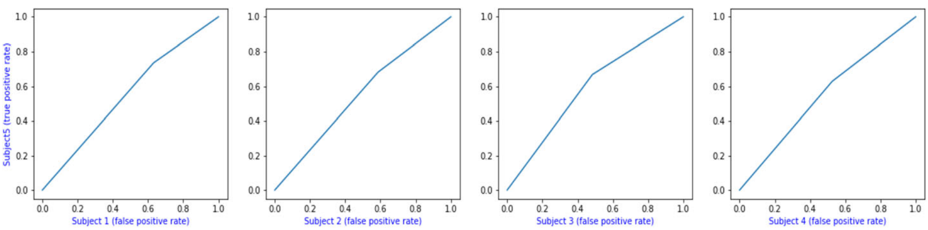

Fig. 17 ROC curve of CapsNet over the case-5 of Dataset-1

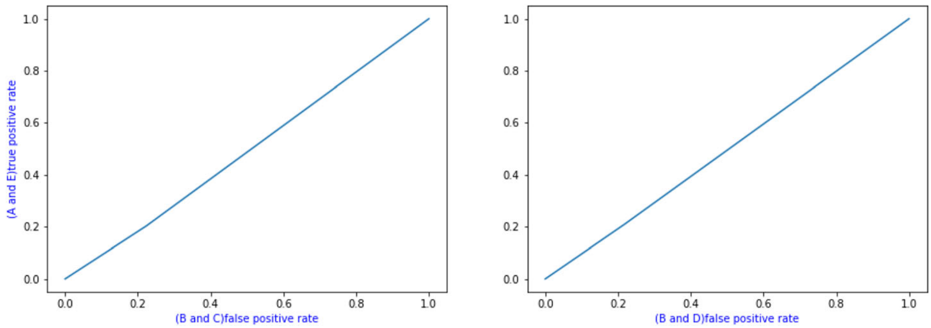

Fig. 18 ROC curve of CapsNet over the of Dataset-2

CNN based approach against our proposed approach. Table 8, shows a comparison against all existing approach based on CapsNet with our proposed CapsNet based seizure detection approach. In our experiment, Tables 7 and 8 shows that the CapsNet based approach can be one of the approaches for seizure detection using EEG signals, also it has advantages for cross-subject seizure detection.

6 Conclusion and future work

In this present study, a Capsule Neural Network (CapsNet) based seizure detection approach has been proposed for the classification of seizure and non-seizure over the epileptic EEG signals. We have performed subject specific and cross-subject training and testing over the two datasets. Dataset-1 [7] is multi-channel EEG dataset and Dataset-2 [14] is a single channel EEG dataset. The EEG datasets have been arranged as per the seizure and non-seizure time stamp. After that L2 normalization technique has been used to normalize the raw EEG data. After normalization, the data have been divided in batch size 10 then fed into the configured CapsNet to perform the seizure detection task. It has been observed that CapsNet perform better as compared to CNN and other traditional classification algorithms by giving a highest average accuracy of 93.50% and also overcomes all issues related to CNN over the seizure detection. Thus, this could be a capable approach for seizure and non-seizure EEG signals classification over the epileptic EEG signals and it can be used effectively for seizure detection.

Limitations of this work is belonging to the generalization of the proposed approach for achieving more robust performance of cross-subject seizure detection. Future work can be focused on the generalization of the proposed approach, which may help us to increase the more robustness of the model. Also, we can use graph neural network based approaches, or by introducing attention based mechanism in capsule neural network for cross-subject seizure detection. During the generalization of the proposed approach, we can use the whole dataset to verify the performance instead of the five subjects EEG data from the Dataset-1 which is used in this experiment. Our proposed approach can be tested and useable for Alzheimer’s, Autism, Schizophrenia, and related disease detection over EEG signals.

Table 7 Comparison with various CNN based seizure detection approaches against our proposed approach

Suggested by	Preprocessing Techniques used	Classification Method	Experimental Dataset	Estimated Performance
A. R. Ozcan et al. [35]	Multi-color image generation from selected features Spectral band power, statistical moment and Hjorth parameters	3D Convolutional Neural Networks (3D CNN) techniques	CHB-MIT scalp EEG dataset	Sensitivity: 85.7%, False prediction rate (FPR): 0.096/h
C. Liu et al. [31]	PCA for data segmentation and FFT for time domain conversion	Multi-View Convolutional Neural Network	Two subjects of CHB-MIT scalp EEG dataset and Intracranial EEG dataset by a Kaggle contest	AUC value: 0.82 and 0.89 on two subjects respectively
Ali Emami et al. [17]	Filter technique, Segmentation, Finally, segmented data have been converted into Plot EEG	Convolutional Neural Network	Authors in-house EEG data	Median true positive rate 74%/sec and median of detected seizure rate 100%/min
Peter Z. Yan et al. [51]	EEG spectrograms processing	Convolutional Neural Networks	CHB-MIT scalp EEG dataset and NYP-WC EEG data	Grater that 90% of seizure detection sensitivity and specificity on CHB-MIT scalp Dataset and grater that 90% sensitivity and 75–80% specificity on the NYP-WC dataset
Gopal Chandra Jana et al. [29]	EEG spectrograms processing	1D- Convolutional Neural Networks	Eight subjects from CHB-MIT scalp EEG dataset	Highest accuracy 88.00%, Highest Sensitivity 86.52%, Highest Specificity 91.11%
U. Rajendra Acharya et al. [2]	Z-score normalization	Deep Convolutional Neural Network (13-layer)	University Bonn EEG dataset	Accuracy: 88.67%, specificity: 90.00%, and Sensitivity: 95.00%
N. D. Truong et al. [50]	FFT	Integer Convolutional Neural Network	CHB-MIT scalp EEG dataset, UPenn and Mayo Clinic's seizure detection data set and Dataset by a Kaggel contest	Only 2% drop of accuracy
This work	L2-normalization	Capsule Neural Network (CapsNet) With Subject specific and Cross subject training and Testing	Five subjects from CHB-MIT scalp EEG dataset, University of Bonn EEG dataset	For Subject Specific: over Dataste-1, a mean accuracy of 93.50%, sensitivity of 93.19%, specificity of 93.90%, FPR of 0.05186. AUC score of 0.93508, F1 score of 0.9366. For Dataset-2: a mean accuracy of 82.61%, sensitivity of 80.54%,

Table 7 (continued)

Suggested by	Preprocessing Techniques used	Classification Method	Experimental Dataset	Estimated Performance
				specificity of 84.83%, FPR of 0.1516. AUC score of 0.8266, F1 score of 0.8273. For Cross subject experiment: Results have been mentioned in Tables 5 and 6.

Table 8 Comparison with existing CapsNet based classification approaches against our proposed approach

Suggested by	Approach intended for	Preprocessing and Classification approach used	Cross Subject training and Test Experiment	Dataset used	accuracy
Kwon-WooHa et.al [26]	Motor Imagery classification	STFT has been used for converting EEG signal into 2D images and finally CapsNet has been used for classification	No	BCI competition IV 2b	Accuracy: 78.44%
Jinliang GUO [24]	Emotion Recognition	Granger causality feature has been extracted and feature screening has performed using sparse group lasso algorithm, finally CapsNet has been used for classification	No	DEAP dataset	Accuracy: 96%
Hao Chao et.al [6]	Emotion Recognition	Multiband feature matrix (MFM) preparation, and finally CapsNet has been used for classification	No	DEAP dataset	Accuracy: 68%
Suat Toraman [49]	Preictal and inter-ictal EEG signals Classification	Channel selection (two channel considered) and segmentation	**No**	CHB-MIT scalp EEG dataset	**For subject Specific:** Best Accuracy: 97.74% **For Cross subject:** Not performed/Mentioned.
This work	Seizure Detection	L2 Normalization, and finally CapsNet with dynamic routing has been used for classification. (All EEG Channels has been considered)	**Yes**	Five subjects from CHB-MIT scalp EEG dataset, University of Bonn EEG dataset	**For subject Specific:** a *mean accuracy* of 93.50% over Dataset-1, Best accuracy 98.17% for Subject-1 of Dataset-1 and an accuracy of 82.61% for Dataset-2. **For Cross subject:** Highest mean accuracy of 86.41% over the Dataset-1, Best accuracy 95.58% for Subject-2 Vs Subject-1 of Dataset-1 and a mean accuracy of 48.45% over the Dataset-2

Acknowledgements This work was carried out at Interactive Technologies & Multimedia Research Lab (ITMR Lab) supported by the Department of Information Technology, Indian Institute of Information Technology Allahabad (https://www.iiita.ac.in/), UP, India. The authors are grateful for this support.

Data availability Experimental datasets used in this proposed approach are publicly available is at https://physionet.org/content/chbmit/1.0.0/ (CHB-MIT scalp EEG dataset) and https://repositori.upf.edu/handle/10230/42894 (University of Bonn, EEG time series dataset).

Declaration

Conflict of interest Authors declare that there is not any conflict of interest in this present study.

References

1. Abdelhameed A, Bayoumi M (2021) A deep learning approach for automatic seizure detection in children with epilepsy. Front Comput Neurosci 15:650050. https://doi.org/10.3389/fncom.2021.650050
2. Acharya UR, Oh SL, Hagiwara Y, Tan JH, Adeli H (2018) Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. Comput Biol Med 100:270–278. https://doi.org/10.1016/j.compbiomed.2017.09.017
3. Andrzejak RG, Lehnertz K, Mormann F, Rieke C, David P, Elger CE (2001) 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 64:061907. https://doi.org/10.1103/PhysRevE.64.061907
4. Bhattacharyya A, Pachori RB (2017) A multivariate approach for patient-specific EEG seizure detection using empirical wavelet transform. IEEE Trans Biomed Eng 64:2003–2015. https://doi.org/10.1109/TBME.2017.2650259
5. Bisong E (2019) Google Colaboratory, building machine learning and deep learning models on Google cloud platform. Apress, Berkeley, CA
6. Chao H, Dong L, Liu Y, Lu B (2019) Emotion recognition from multiband EEG signals using CapsNet. Sensors 19:2212. https://doi.org/10.3390/s19092212
7. CHB-MIT Scalp EEG dataset, https://physionet.org/content/chbmit/1.0.0/. Accessed: 27 Jan 2023
8. Chen S, Zhang X, Chen L, Yang Z (2019) Automatic diagnosis of epileptic seizure in electroencephalography signals using nonlinear dynamics features. IEEE Access 7:61046–61056. https://doi.org/10.1109/ACCESS.2019.2915610
9. Chernecky CC, Berger BJ (2013) laboratory tests and diagnostic procedures. 6th ed, Elsevier, ISBN 9781455706945
10. Chiang H-S, Chen M-Y, Huang Y-J (2019) Wavelet-based EEG processing for epilepsy detection using fuzzy entropy and associative petri net. IEEE Access 7:103255–103262. https://doi.org/10.1109/ACCESS.2019.2929266
11. Colaboratory: Frequently Asked Questions, https://research.google.com/colaboratory/faq.html. Accessed: 27 Jan 2023
12. Cole TJ (1991) Applied logistic regression. D. W. Hosmer and S. Lemeshow, Wiley, New York, 1989. No. of pages: xiii + 307. Price: £36.00. Stat Med 10:1162–1163. https://doi.org/10.1002/sim.4780100718
13. Constantino TM, What's the Difference Between a Seizure and Epilepsy? In: intermountainhealthcare.org. https://intermountainhealthcare.org/blogs/topics/live-well/2017/12/whats-the-difference-between-a-seizure-and-epilepsy/. Accessed 27 Jan 2023
14. Department of Epileptology, University of Bonn, EEG time series download page, https://repositori.upf.edu/handle/10230/42894. Accessed 27 Jan 2023
15. EEG (Electroencephalogram)-Epilepsy Society. https://epilepsysociety.org.uk/about-epilepsy/diagnosing-epilepsy/electroencephalogram-eeg. Accessed 27 Jan 2023
16. Eldor T, Capsule Neural Networks – Part 2, https://towardsdatascience.com/capsule-neural-networks-part-2-what-is-a-capsule-846d5418929f. Accessed 27 Jan 2023
17. Emami A, Kunii N, Matsuo T, Shinozaki T, Kawai K, Takahashi H (2019) Seizure detection by convolutional neural network-based analysis of scalp electroencephalography plot images. NeuroImage: Clin 22:101684. https://doi.org/10.1016/j.nicl.2019.101684
18. Engel J, Pedley TA (2008) Epilepsy: a comprehensive textbook, 2nd edn. Wilkins, Wolters Kluwer Health/Lippincott Williams &

19. Epilepsy, factsheet - World Health Organization, (9th February 2022), https://www.who.int/en/news-room/fact-sheets/detail/epilepsy. Accessed 27 Jan 2023
20. Gajic D, Djurovic Z, Gligorijevic J, di Gennaro S, Savic-Gajic I (2015) Detection of epileptiform activity in EEG signals based on time-frequency and non-linear analysis Front Comput Neurosci 9. https://doi.org/10.3389/fncom.2015.00038
21. Goodfellow I, Bengio Y, Courville A (2016) Convolution neural network, deep learning 321–359. MIT press
22. Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, Liu T, Wang X, Wang G, Cai J, Chen T (2018) Recent advances in convolutional neural networks. Pattern Recogn 77:354–377. https://doi.org/10.1016/j.patcog.2017.10.013
23. Xifeng Guo, A Keras implementation of CapsNet, Dynamic Routing Between Capsules, https://github.com/XifengGuo/CapsNet-Keras. Accessed 27 Jan 2023
24. Guo J, Fang F, Wang W, Ren F (2018) EEG emotion recognition based on granger causality and CapsNet neural network. 5th IEEE international conference on cloud computing and intelligence systems (CCIS) 47–52, IEEE, Nanjing, China. https://doi.org/10.1109/CCIS.2018.8691230
25. Guptha NS, Balamurugan V, Megharaj G, Sattar KNA, Rose JD (2022) Cross lingual handwritten character recognition using long short term memory network with aid of elephant herding optimization algorithm. Pattern Recogn Lett 159:16–22. https://doi.org/10.1016/j.patrec.2022.04.038
26. Ha K-W, Jeong J-W (2019) Motor imagery EEG classification using capsule networks. Sensors 19:2854. https://doi.org/10.3390/s19132854
27. Ha K-W, Jeong J-W (2019) Decoding two-class motor imagery EEG with capsule networks. IEEE international conference on big data and smart computing (BigComp) 1–4, EEE, Kyoto, Japan. https://doi.org/10.1109/BIGCOMP.2019.8678917
28. Hussain L (2018) Detecting epileptic seizure with different feature extracting strategies using robust machine learning classification techniques by applying advance parameter optimization approach. Cogn Neurodyn 12:271–294. https://doi.org/10.1007/s11571-018-9477-1
29. Jana GC, Sharma R, Agrawal A (2020) A 1D-CNN-spectrogram based approach for seizure detection from EEG signal. Procedia Comput Sci 167:403–412. https://doi.org/10.1016/j.procs.2020.03.248
30. Li Y, Wang X-D, Luo M-L, Li K, Yang XF, Guo Q (2018) Epileptic seizure classification of EEGs using time–frequency analysis based multiscale radial basis functions. IEEE J Biomed Health Inform 22:386–397. https://doi.org/10.1109/JBHI.2017.2654479
31. Liu C-L, Xiao B, Hsaio W-H, Tseng VS (2019) Epileptic seizure prediction with multi-view convolutional neural networks. IEEE Access 7:170352–170361. https://doi.org/10.1109/ACCESS.2019.2955285
32. Mei Z, Zhao X, Chen H, Chen W (2018) Bio-signal complexity analysis in epileptic seizure monitoring: a topic review. Sensors 18:1720. https://doi.org/10.3390/s18061720
33. Naturomics, A Tensorflow implementation of CapsNet (Capsules Net), Dynamic Routing Between Capsules, https://github.com/naturomics/CapsNet-Tensorflow. Accessed 27 Jan 2023
34. Niedermeyer E, Lopes da Silva FH (2005) Electroencephalography: basic principles, clinical applications, and related fields. 5th ed, Lippincott Williams & Wilkins
35. Ozcan AR, Erturk S (2019) Seizure prediction in scalp EEG using 3D convolutional neural networks with an image-based approach. IEEE Trans Neural Syst Rehabil Eng 27:2284–2293. https://doi.org/10.1109/TNSRE.2019.2943707
36. Polat K, Güneş S (2007) Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform. Appl Math Comput 187:1017–1026. https://doi.org/10.1016/j.amc.2006.09.022
37. Praveena HD, Guptha NS, Kazemzadeh A, Parameshachari BD, Hemalatha KL (2022) Effective CBMIR system using hybrid features-based independent condensed nearest neighbor model. J Healthcare Eng 2022: 1–9. https://doi.org/10.1155/2022/3297316
38. Preprocessing data: Normalization, https://scikit-learn.org/stable/modules/preprocessing.html#preprocessing-normalization. Accessed 27 Jan 2023
39. Qiu Y, Zhou W, Yu N, Du P (2018) Denoising sparse autoencoder-based ictal EEG classification. IEEE Trans Neural Syst Rehabil Eng 26:1717–1726. https://doi.org/10.1109/TNSRE.2018.2864306
40. Rokach L, Maimon OZ (2008) Data mining with decision trees Theroy and applications. World Scientific, ISBN 978-9812771711.
41. Sabour S et al. (2017) Dynamic Routing Between Capsules, NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing. PP. 3859–3869
42. Shoaran M, Haghi BA, Taghavi M, Farivar M, Emami-Neyestanak A (2018) Energy-efficient classification for resource-constrained biomedical applications. IEEE J Emerg Sel Topics Circuits Syst 8:693–707. https://doi.org/10.1109/JETCAS.2018.2844733

43. Solaija MSJ, Saleem S, Khurshid K, Hassan SA, Kamboh AM (2018) Dynamic mode decomposition based epileptic seizure detection from scalp EEG. IEEE Access 6:38683–38692. https://doi.org/10.1109/ACCESS.2018.2853125
44. Srihari S, Capsule Networks, https://cedar.buffalo.edu/~srihari/CSE676/9.12%20CapsuleNets.pdf. Accessed 27 Jan 2023
45. Subasi A (2005) Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients. Expert Syst Appl 28:701–711. https://doi.org/10.1016/j.eswa.2004.12.027
46. Subasi A, Erçelebi E (2005) Classification of EEG signals using neural network and logistic regression. Comput Methods Prog Biomed 78:87–99. https://doi.org/10.1016/j.cmpb.2004.10.009
47. Tatum WO (2013) Handbook of EEG interpretation. Demos Medical Publishing pp155–190. ISBN 9781617051807
48. Theeranaew W, Ryvlin P, Surges R, Thijs R, Schuele S, Lhatoo S, Loparo KA, McDonald J, Zonjy B, Kaffashi F, Moseley BD, Friedman D, So E, Tao J, Nei M (2018) Automated detection of postictal generalized EEG suppression. IEEE Trans Biomed Eng 65:371–377. https://doi.org/10.1109/TBME.2017.2771468
49. Toraman S (2021) Automatic recognition of preictal and interictal EEG signals using 1D-capsule networks. Comput Electr Eng 91:107033. https://doi.org/10.1016/j.compeleceng.2021.107033
50. Truong ND, Nguyen AD, Kuhlmann L, Bonyadi MR, Yang J, Ippolito S, Kavehei O (2018) Integer convolutional neural network for seizure detection. IEEE J Emerg Sel Topics Circuits Syst 8:849–857. https://doi.org/10.1109/JETCAS.2018.2842761
51. Yan PZ, Wang F, Kwok N, Allen BB, Keros S, Grinspan Z (2019) Automated spectrographic seizure detection using convolutional neural networks. Seizure 71:124–131. https://doi.org/10.1016/j.seizure.2019.07.009