Circuits, Systems, and Signal Processing (2023) 42:6780–6797
https://doi.org/10.1007/s00034-023-02423-1

Epilepsy Detection with Multi-channel EEG Signals Utilizing AlexNet

Sohaib Majzoub[1] · Ahmed Fahmy[1] · Fadi Sibai[2] · Maha Diab[1] · Soliman Mahmoud[1]

Received: 5 September 2022 / Revised: 27 May 2023 / Accepted: 29 May 2023 /
Published online: 5 July 2023

Abstract

In this work, we investigate epilepsy seizure detection using machine learning. In the literature, a machine learning model is usually trained to help automate the epileptic detection process, eliminating the need for human intervention. Typically, the dataset is split into training and test sets in a way to maximize the detection accuracy. This requires the training set to include enough EEG samples for every possible patient in order to improve the accuracy numbers during the prediction. However, this might not be easy or practical in real life. A new patient might not have a previous record in the training set, and hence, the prediction for this particular patient might not meet the expected accuracy. The main contribution in this work is to study the impact of the training and test datasets selection from practical point of view on the accuracy and efficacy of the CNN prediction. In this work, a CNN model, namely AlexNet, is trained to detect epileptic states, namely preictal, interictal and ictal, in subjects using electroencephalogram (EEG) signals. The dataset includes the three epileptic zones of

Can be downloaded from https://openneuro.org/datasets/ds003029/versions/1.0.3.

✉ Sohaib Majzoub
smajzoub@sharjah.ac.ae

Ahmed Fahmy
b00077242@aus.edu

Fadi Sibai
Sibai.F@gust.edu.kw

Maha Diab
mdiab@sharjah.ac.ae

Soliman Mahmoud
solimanm@sharjah.ac.ae

1 Department of Electrical Engineering, University of Sharjah, Sharjah 27272, Sharjah, UAE

2 Electrical and Computer Engineering Department, Gulf University for Science and Technology, Mishref, Kuwait

subjects taken from three medical centers, collected by the Fragility Multi-Center Retrospective Study. Furthermore, we propose a framework to utilize a feature extraction technique that exploits the available multiple channels of EEG signals to minimize information loss. As part of the main contribution, three different approaches are proposed to split the EEG sample dataset into the training and test sets. Thus, the prediction performance is evaluated based on the prior knowledge extracted from the particular samples picked for the training set. The results show that the proposed framework achieves an overall accuracy of 94.44% when the training contained samples from each patient. The accuracy is reduced to 92.98% when the training set contained a subset of the patient pool. A binary classification is also performed with up to 98% accuracy for both scenarios.

Keywords Spectrogram · EEG · Machine learning · AlexNet · Ictal · Epilepsy · Seizure detection

1 Introduction

Epilepsy is a chronic neurological disorder affecting the brain, causing patients to suffer from recurrent seizures of involuntary movement. Seizures are burst of sudden electric activity in the brain, which can be monitored through five electroencephalogram (EEG) signals recorded from patients [14]. According to the World Health Organization (WHO), epilepsy affects more than 50 million people worldwide, with 80% of patients living in low and middle-income countries, labeling it as one of the most common neurological diseases globally. Due to recurrent seizures, patients suffer from both physical and psychological problems, such as bruising, and fractures as a result of the seizure as well as anxiety and depression. To minimize the possible physical and psychological consequences of a seizure, researchers have developed methods to detect epileptic individuals and potential seizures in such individuals.

This work aims to detect the onset of a seizure that can help alert both patients and medics, providing them the required time to manage and prevent the seizure, thus improving their quality of life. Depending on the location in the brain in which the electrical disturbance occurs, the intensity and symptoms of an epileptic seizure vary. The stages of an epileptic seizure can be divided into three main stages, namely preictal, ictal (seizure) and interictal. The preictal state is the time before the seizure onset, while the time during which a seizure occurs is the ictal state. The time period between seizures is referred to as the interictal state. Therefore, for early detection of a seizure, it is most important to identify the preictal state which would predict the occurrence of a seizure [28].

Most of machine learning methods in the literature showed very good results, with accuracy percentage from high 80's to high 90's, tested on cleaned and well-studied, but old, datasets such as CHB-MIT, Bonn, Freiburg and others [9, 12, 13, 16, 23]. Examples of the used machine learning methods are support vector machine, KNN, forest trees, random forest, etc. [35, 36]. In this work, AlexNet is used to showcase its performance on a very recent dataset, namely the Fragility Multi-Center Retrospective Study, [27, 29]. Unlike many studies in the literature, multiple centers contributed in the dataset used in this work; this eliminates potential biased samples due to possible center-relevant factors. Furthermore, different approaches were used to split the dataset to study the sensitivity of the results based on the training and test sets selection. The contributions of this work are summarized as follows:

- Building a ML model to predict the different epilepsy time intervals using AlexNet.
- Obtaining up to 98% prediction for new multi-channel EEG signal recordings dataset provided by the Fragility Multi-Center Retrospective Study, [27], after denoising and representing the data in the form of spectrograms.
- Providing prediction performance analysis based on different training and test dataset selection criteria from a practical perspective. The proposed data splitting approaches are sample split, patient–sample split, and patient split. The proposed methods are discussed in Sect. 5.

2 Related Work

Earlier detection models for epileptic seizures employed machine learning algorithms using feature extraction techniques [35]. The performance of the machine learning models is governed by the extracted features used as input for the classification model, which is a manual preprocessing step mainly based on trial and error. Several machine learning models were developed using different signal preprocessing techniques [2, 4, 7, 8, 11, 14, 19, 21, 25, 33, 34, 37]. The work by [20] focused on comparing different machine learning algorithms for seizure classification, such as generalized relevance learning vector quantization (GRLVQ), support vector machine (SVM), back-propagation, and random forest. The study also considered feature extraction techniques such as the wavelet and principal component analysis (PCA). The results of the comparison reported the best accuracy of 98.66% using the GRLVQ classification model. Another research [38] improved the performance through the use of hybrid models with particle swarm optimization (PSO) and genetic algorithm (GA) to find the optimum parametric values of SVM. The use of PSO with SVM reported an accuracy of 99.38% using extracted features. An epileptic seizure detection machine learning model, proposed by [26], achieved high classification accuracy for six classification problems through feature extraction using discrete wavelet transform (DWT)-based approximate entropy (ApEn). The six classification problems consisted of four binary cases, while the other two had four and five classification classes. The reported accuracy for two of the binary cases was 100%, while the sixth case classifying extracted features of five classes reported 94% accuracy using SVM. Machine learning-based classification of epileptic seizures requires an extra preprocessing step of feature extraction [1]. On the other hand, with deep learning algorithms, features are learned from the input data without the need for an external feature extractor. In the literature, multiple researchers employed deep learning algorithms in developing epileptic seizure detection systems [3, 18, 22, 40]. In [31], a CNN that detects nocturnal frontal lobe epilepsy (NFLE) using EEG signals was presented. The work focused on utilizing CNN to develop patient-specific models classifying EEG signals into binary classes of seizure and no seizure. The model was built using CNN and transfer learning to develop a cross-patient seizure detection model. The CNN architecture was based on deep residual network (ResNet) of the MATLAB deep learning library and resulted in an accuracy above 94% for the three patient-specific models. A more recent work on the classification of epileptic seizures using deep CNN was introduced by [15]. The work proposed a new deep learning-based classification method combining the use of three pretrained CNNs, the Inception ResNet v2, Inception v3, and ResNet 152 followed by two fully connected convolutional layers, a softmax layer, and final classification layer. The proposed classification model named epileptic EEG signal classification (EESC) was used on the CHB-MIT database of children EEG signals. The dataset was preprocessed for denoising and feature extraction. Discrete wavelet transform (DWT) was used for denoising raw EEG signals, and then the power spectral density energy diagrams (PSDED) were created to act as 2D image input for the classification. The classification presented four output classes: interictal, preictal of 30 min duration, preictal of 10 min duration, and seizure. The reported accuracy was greater than 90% for all classes. An autodetection of seizure events using deep neural network (DNN) compared to machine learning techniques was investigated by [1]. Authors used the Bonn University dataset to build a binary classification model for seizure detection. They compared the accuracy of the machine learning techniques, the support vector machine (SVM), and the K-nearest neighbor (KNN), with the DNN using different feature scaling techniques. The proposed DNN reported an accuracy of 97.21%. Using convolutional neural network to detect epileptic seizures, the study by [47] investigated the performance difference between using time and frequency domain signals. The classification model was empirically evaluated by considering two binary problems: interictal versus ictal and interictal versus preictal. In addition, the study considered the ternary problem that involves the three classes, namely the interictal, preictal, and ictal. The proposed model was tested using two epileptic databases and reported accuracies above 90% with better performance using the frequency domain data. In this work, we propose a CNN-based framework for epileptic seizure detection, where multi-channel EEG signals are transformed into two-dimensional power spectral density energy diagrams (PSDED) images. The rest of the paper is organized as follows: Sect. 3 presents an overview of the dataset used in the development process. The data preprocessing, denoising and features extraction are shown in detail in Sect. 4. The CNN architecture employed and its associated layers are explained in the same section. Section 5 demonstrates the empirical evaluation and results of the proposed deep learning framework. Finally, the work is concluded in Sect. 7.

3 Multi-center Multi-channel Epileptic Dataset

As opposed to most of the datasets used in developing deep learning models in prior research studies, the data used in this work are collected by multiple centers, which minimizes potential biases due to possible factors that are associated with a single data source. The dataset was collected by the Fragility Multi-Center Retrospective Study, [10, 17, 27, 30], from multiple epilepsy centers, namely the Cleveland Clinic (CC), the Jackson Memorial Hospital in University of Miami (JMH), the Johns Hopkins Hospital (JHH), the University of Maryland Medical Center (MMC), and the National

Table 1 The EEG dataset with 68 recordings from 3 medical centers used in this work [27]

Subject	Gender	Age	Number of Recordings	Sampling Rate (Hz)	Center
pt1	F	30	4	1000.00	NIH
pt2	F	28	3	1000.00	NIH
pt3	M	45	2	1000.00	NIH
pt6	M	33	3	1000.00	NIH
pt7	F	39	3	1000.00	NIH
pt8	M	25	3	1000.00	NIH
pt10	F	44	3	1000.00	NIH
pt11	M	31	4	1000.00	NIH
pt12	F	43	2	1000.00	NIH
pt13	M	27	4	1000.00	NIH
pt14	F	49	3	1000.00	NIH
pt15	F	59	4	1000.00	NIH
pt16	F	52	3	1000.00	NIH
umf001	F	37	1	999.41	JMH
ummc001	M	23	3	499.71	MMC
ummc002	M	17	3	499.71	MMC
ummc003	M	31	3	249.85	MMC
ummc004	M	38	3	249.85	MMC
ummc005	M	47	2	999.41	MMC
ummc006	M	36	3	249.85	MMC
ummc007	M	54	3	999.41	MMC
ummc008	M	49	3	249.85 (1) – 999.41 (2)	MMC
ummc009	M	36	3	999.41	MMC

Institute of Health (NIH). Each center collected data from different number of patients, each patient has a single or multiple EEG signal recordings, and each recording has one seizure occurrence and is captured using multiple channels.

3.1 Inclusion and Exclusion Criteria

According to the Fragility Multi-Center Retrospective Study authors, the CC did not make their records available publicly. On the other hand, the data collected from JMH, MMC, NIH, and JHH can be accessed online. Due to variety of samples' sources, different centers used different wordings during the marking of the seizure onsets on the recordings. The specific clinical events are marked or annotated by clinicians, referring to onset and offset of seizures events for each of the collected data files. The seizure clinicians were not consistent in their description during the labeling of the data. Therefore, these different markings were replaced by regular-expressions such as "start" and "stop" during the streaming of data into the software experimental environment used in this work. The new labels were used then to identify the seizure

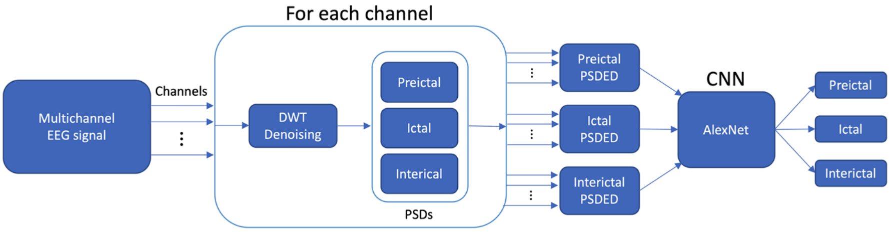


Fig. 1 Proposed CNN-based deep learning framework for seizure detection

zones in patient recordings when cleaning and splitting the data into preictal, ictal, or interictal target images. Recordings with unclear or undetectable onset markers were removed from the dataset. Thus, the useful EEG recordings that were utilized are summarized in Table 1. A total 68 samples were used in this study out of 100 provided samples in the dataset.

4 Methodology

In this work, a CNN-based deep learning framework, namely AlexNet, is used to identify segments of a multi-channel EEG signal as preictal, ictal, or interictal. An overview of the proposed framework is illustrated in Fig. 1.

Each channel of an input EEG signal is denoised to remove potentially existing artifacts. Subsequently, each signal is divided into three periods: preictal, ictal and interictal by using the seizure onset markers, which are identified by clinicians. The obtained signals are then transformed into their power spectral density (PSD) representation. The PSDs of all the channels that corresponds to the preictal, ictal and interictal signals are combined into preictal, ictal and interictal PSDED diagrams, respectively. The PSDEDs are then used to train the AlexNet CNN in order to implement the classification problem, see Sect. 4.2 for details.

4.1 Signal Denoising and Removing of Artifacts

EEG is high in temporal resolution [32], and hence, an EEG signal is easily altered by minor interference, resulting in EEG artifacts while being recorded. Each channel of an EEG signal, therefore, should undergo a denoising procedure [39]. We impose the wavelet threshold denoising method. The DWT denoising is applied using the Daubechies wavelet of order 6 (dB6) [26]. The obtained denoised signals are free of artifacts, while they maintain the characteristics of each of the three zones of interest. An example of an EEG signal and its denoised variation is presented in Fig. 2.

4.2 Features Extraction

EEG signals can be characterized by their change in power in the frequency domain [39]. Such a variation is captured by analyzing the PSD of an EEG signal [5]. Each

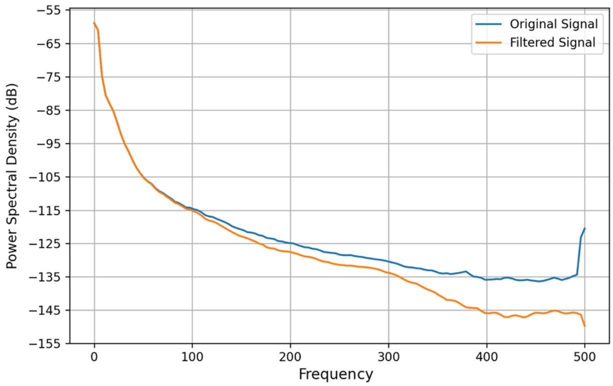


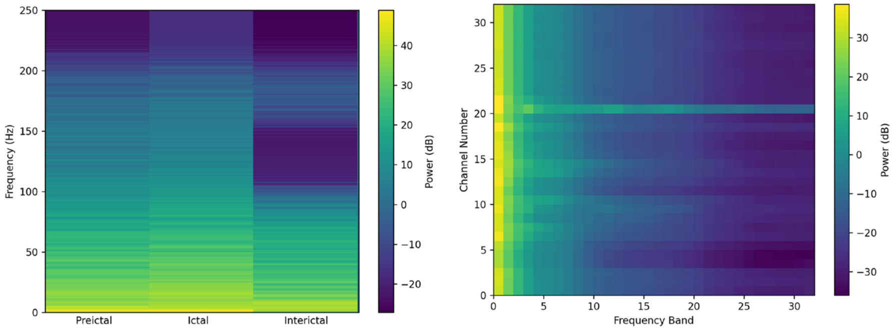
Fig. 2 EEG signal before (blue) and after (orange) denoising (Color figure online)

denoised signal of an EEG channel is divided into three periods: preictal, ictal and interictal. The three periods were marked in the dataset files by the clinicians. An example of heat map representing the PSD of the three periods of the EEG signal is shown in Fig. 3a. Such visible differences can be used as features to characterize each of the three signal periods. It should be noted that the EEG signal files have different sampling rate. For consistency, those sampled at 500Hz and 1000Hz are resampled using 250Hz sample rate for all images during the training and then testing phase.

In this work, a multi-channel signals are considered. The number of available channels varied for each recording. The number of channels was roughly between 70 and 140 channels. Some of these channels were noisy and marked "bad" and other better ones marked as "good". All bad channels were excluded from the study. Furthermore, performing PSD analysis on a single channel would not capture potentially valuable information from the other channel signals. Therefore, the adopted strategy, as in [15], is to choose n channels for every EEG signal S_i that has m_i channels, such that $n < \min_i\{m_i\}$. Then, each channel is divided into n frequency bands in which the power is averaged, constituting an $n \times n$ matrix, referred to as PSDED image. This image has to be a square image as it is a stipulation for the AlexNet model used in the ML for predication. Hence, a PSDED image is an $n \times n$ matrix representing n channels of a recorded multi-channel EEG signal, illustrated in Fig. 3b, that manifests differences between the three periods under consideration. In this work, $n = 32$ is used to construct the images. Thus, there were 32 frequency bands facing 32 channels. The first 32 best channels were considered for each of the recordings [15]. These images are used as the input for the deep learning AlexNet CNN for classification.

4.3 Classification using CNN

After producing the two-dimensional $n \times n$ images for all recordings in the dataset, the images are then normalized using z-score with mean 0 and standard deviation of 1. The normalized images are fed to the input layer of a convolutional neural network.



(a) Heat map that reflects the PSD of the three periods of an EEG signal for a single channel.

(b) 32 × 32 PSDED image for the seizure (ictal) period of a multi-channel EEG signal.

Fig. 3 (a) a heat map and (b) a PSDED image for single- and multi-channel EEG signals, respectively

4.3.1 AlexNet

AlexNet is a CNN designed by Alex Krizhevsky, [24], in 2012 after it was introduced in the ImageNet Large Scale Visual Recognition Challenge. AlexNet completely revolutionized deep-learning image classification after winning the first prize. AlexNet consists of three max-pooling layers. Generally, a pooling layer encapsulates the features presented in the feature map generated by a convolution layer while reducing the dimensionality of the problem. Therefore, subsequent operations are performed on a reduced feature set instead of the full feature map. Similarly, the max-pooling layers used in AlexNet reduce the dimensions of the feature maps and, thus, minimizes the number of parameters during the learning process, and hence the amount of computations required in the CNN. This also makes the model more robust against the variation in the features' location in the input images. Like traditional neural networks, the two fully connected layers in AlexNet consist of simple neurons that are fed by an activation function, similar to the feed-forward neural networks. However, an AlexNet CNN deploys a dropout strategy to avoid overfitting, where it sets the outgoing edges of a number of arbitrary hidden units to 0 after each training phase. Additionally, an AlexNet includes two layers that apply batch normalization, referred to as normalization layers, to standardize the inputs between layers. This allows stabilizing the learning steps of the model by reducing the number of training epochs and, consequently, the overall computational cost of the neural network.

5 Results and Discussion

The input for the CNN model is the PSDED images extracted from the dataset, such that, 70% were used for training and the remaining 30% were used for the test set. The three data splitting approaches were adapted to simulate different kinds of prior knowledge and its effect on the model prediction accuracy. The exact number of samples in each set, i.e., training and test sets, varied based on the proposed approach. The proposed splitting approaches of the training and test sets are:

- **Sample split**: the training and test sets are picked randomly from all available sample pool regardless of patients. The picked samples of preictal, interictal and ictal do not have to belong to the same patient. This approach is a typical one used in machine learning, but it might not provide the best results. The number of samples in the training and test sets is 48 and 20, respectively.
- **Patient–sample split**: the training and test sets are picked randomly from all available patients. The picked samples of preictal, interictal and ictal have to belong to the same patient. This approach is more consistent with the problem at hand, where each set should have the information of all classes of a given patient. For this scenario, two CNN models were trained, a multi-class prediction model to predict the three classes, namely ictal, preictal and interictal, and a binary prediction model to predict the ictal and non-ictal cases (combining the preictal and interictal classes into one). The number of samples in the training and test sets is 48 and 20, respectively.
- **Patient split**: involves splitting records by patients, such that the network is only exposed to the records of a subset of patients during the training phase. This evaluates the model's performance in real-world scenarios where historic data from multiple centers are mixed together and used to classify future patient records without any previous knowledge about these patients. Two CNN models were trained, a multi-class prediction model to predict the three classes and a binary prediction model to predict the ictal and non-ictal classes. The number of samples in the training and test sets is 49 and 19, respectively.

Table 2 shows the training and test sets for each of the discussed scenarios per patient. It should be noted that in all cases, random permutations are tried while preserving the rules mentioned above. The permutation with the highest accuracy is reported in this work. MATLAB and then Keras open-source library along with TensorFlow were used as the experimental environment in this work. The number of epochs along with different random seeds was selected based on trial and error approach to produce the best results. The best performance for each model is reported and presented in this section. The training and testing times were around 2 min and few seconds, respectively, on iMAC using 3 GHz 6 cores Intel i5, with total memory of 32GB of RAM.

The AlexNet Convolutional Neural Network is used in this work. It is trained using the training sets produced for the multi-class scenarios to classify the three different classes of the epileptic periods: preictal, ictal and interictal, and the two binary classification scenarios to classify seizure and non-seizure classes. The figures of merit and statistical measurements used to assess the quality of the obtained results for each case are classification accuracy, sensitivity, specificity, precision and the F1 score, calculated using the equations below, [26].

Table 2 The training and test EEG dataset split using the three scenarios

Subject	Records	Sample split	Patient-sample split		Patient split	
			Multi-class	Binary	Multi-class	binary
		Training/Test				
pt1	4	2/2	4/0	3/1	4/0	4/0
pt2	3	3/0	2/1	3/0	3/0	3/0
pt3	2	1/1	1/1	2/0	2/0	2/0
pt6	3	2/1	2/1	3/0	3/0	3/0
pt7	3	3/0	3/0	2/1	3/0	3/0
pt8	3	3/0	3/0	2/1	0/3	3/0
pt10	3	3/0	2/1	2/1	3/0	3/0
pt11	4	4/0	2/2	2/2	4/0	4/0
pt12	2	2/0	1/1	2/0	2/0	2/0
pt13	4	3/1	3/1	2/2	4/0	0/4
pt14	3	2/1	1/2	3/0	3/0	3/0
pt15	4	3/1	2/2	3/1	0/4	0/4
pt16	3	1/2	1/2	2/1	0/3	0/3
umf001	1	0/1	1/0	1/0	1/0	1/0
ummc001	3	2/1	2/1	3/0	3/0	3/0
ummc002	3	1/2	2/1	1/2	3/0	3/0
ummc003	3	1/2	3/0	0/3	3/0	3/0
ummc004	3	3/0	3/0	1/2	3/0	0/3
ummc005	2	1/1	1/1	1/1	2/0	0/2
ummc006	3	2/1	3/0	3/0	0/3	3/0
ummc007	3	1/2	2/1	2/1	0/3	3/0
ummc008	3	3/0	1/2	2/1	0/3	0/3
ummc009	3	2/1	3/0	3/0	3/0	3/0
Total	68	48/20	48/20	48/20	49/19	49/19

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{2}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{3}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4}$$

$$\text{F1\_Score} = \frac{2 \times precision \times sensitivity}{precision + sensitivity} \tag{5}$$

where TP, TN, FP and FN are true positives, true negatives, false positives and false negatives, respectively.

Table 3 Prediction results of the proposed splits using AlexNet

Method	Type	Classes	Accuracy	Sensitivity	Specificity	Precision	F1_score
Sample Split	Multi-Class	Interictal	85	76.19	89.74	80	78.05
		Preictal	83.33	77.78	85.71	70	73.68
		Ictal	98.33	95.24	100	100	97.56
		Mean	**88.89**	**83.07**	**91.82**	**83.33**	**83.1**
Patient-Sample Split	Multi-Class	Interictal	95	100	93.02	85	91.89
		Preictal	95	94.74	95.12	90	92.31
		Ictal	93.33	83.33	100	100	90.91
		Mean	**94.44**	**92.69**	**96.05**	**91.67**	**91.7**
	Binary	Non-Ictal	98.25	100	95	97.37	98.67
		Ictal	98.25	95	100	100	97.44
		Mean	**98.25**	**97.5**	**97.5**	**98.68**	**98.05**
Patient Split	Multi-Class	Interictal	94.74	94.44	94.87	89.47	91.89
		Preictal	92.98	85.71	97.22	94.74	90
		Ictal	91.23	88.89	92.31	84.21	86.49
		Mean	**92.98**	**89.68**	**94.8**	**89.47**	**89.46**
	Binary	Non-Ictal	98.33	100	95.24	97.5	98.73
		Ictal	98.33	95.24	100	100	97.56
		Mean	**98.33**	**97.62**	**97.62**	**98.75**	**98.15**

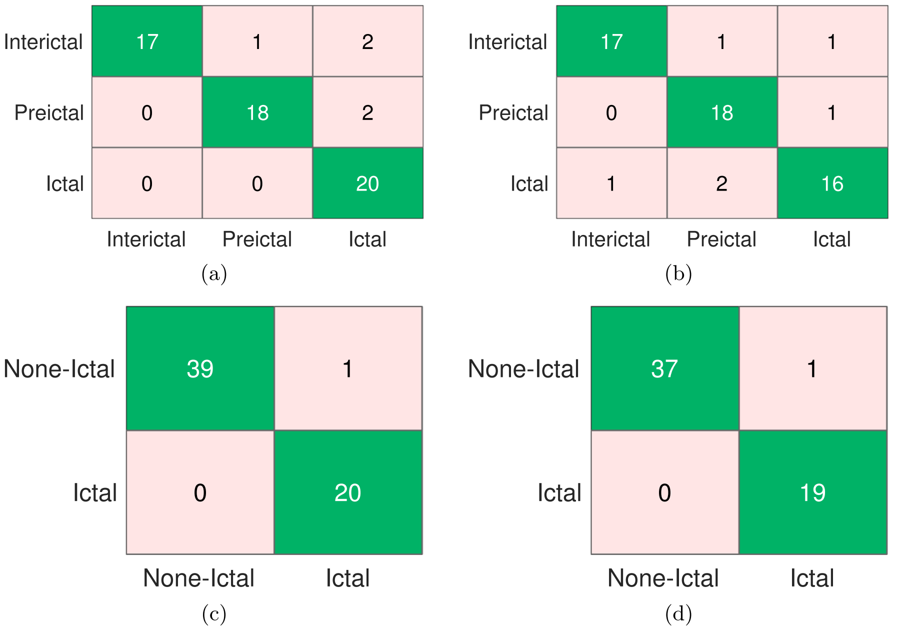


Fig. 4 Confusion matrices of the (a) multi-class patient–sample split, (b) multi-class patient split, (c) binary patient–sample split and (d) multi-class patient split (the y-axis shows the True Class and the x-axis shows the Predicted Class)

Table 3 shows the produced results in this work for each of the scenarios mentioned earlier, namely sample split, patient–sample split and patient split. The confusion matrices of the patient–sample split and patient split are shown in Fig. 4, where the y-axis is the True Class and x-axis shows the Predicted Class. In the context of this work, perhaps the sensitivity is an important measure as it reflects the ratio of the true positives (TPs) to the summation of the false negatives (FNs) and TPs. The model should predict all the positive cases correctly with minimum FN; otherwise, there will be health implications on the patients. Precision is also important for the same reason; however, its significance comes after the sensitivity as it does not include FNs. On the other hand, specificity might not produce serious health implications as TNs and FPs measures are not going to need immediate medical attention. The F1 score is also a good measure because it considers both the precision and sensitivity. Furthermore, the correct prediction of the preictal state is more important compared to other states where a medical procedure has to take place before the seizure. This can optimize the medicine intake for the patient, thus a patient is given the dosage only if a seizure is on the way.

The prediction results per class for the random sample split case are shown in Table 3. As shown, the accuracy is the lowest among all cases, namely 88.89%. It seems that the random picking of samples for the training and test sets regardless of the patient information is not helpful to improve the prediction. The sensitivity, specificity, precision and F1-score are also showing similar trends. The prediction results per class of the patient–sample split scenario is producing the best results as shown in Table 3. This is expected as this scenario helps the model get enough information about most of the patients during the training. Thus, the model is able to produce the highest prediction accuracy, namely 94.44%. All other figures of merit are showing values above 90%. The confusion matrix of this case is shown in Fig. 4a. It is worth noting that the sensitivity of both preictal and ictal is high (100% and 94.74%, respectively), where 5 samples are mislabeled shown in cells (1,2), (1,3) and (2,3), i.e., preictal/ictal false negatives. These cases constitute a small number yet can have health implications.

The binary classification is performed when the three classes and relabeled into two, namely ictal and non-ictal. It should be noted here that merging classes can create imbalanced classes, and this might impact the performance results. Imbalanced classes in a dataset is a typical problem in most realistic datasets. However, the results seem to improve. Figure 4c and 4d shows the binary classification when the interictal and preictal are combined into one class. The accuracy in this case increased up to 98.25% as shown in Table 3. Figure 4c shows the confusion matrix of the binary classification, as shown only one sample of none ictal (i.e., either interictal or preictal) is mislabeled as ictal. In the case of patient split, where the model does not have any information about some patients, the results show less accurate numbers. This is expected as the trained model does not have any information about the new patients. In this case, the accuracy is reduced to 92.98 for the multi-classification case. Similarly, the other figures of merit are reduced by about 2–3% as well. This is still acceptable result as the reduction is not major drop. The results of the same method but with binary classification are showing similar numbers as the previous case. Thus, the accuracy is 98.33% with only one mislabeled sample as shown in Fig. 4d.

Table 4 shows a comparison between this work and previous results reported in the literature. Although the datasets differ, the proposed method is showing competitive numbers. Note that there was no research done on the dataset used in this work.

6 Limitations and Future Directions

As mentioned in the previous section, it is important for a high-accuracy model to be trained on a comprehensive set of samples. Thus, the training dataset should have enough EEG samples from a diverse number of patients in order to reach the 100% accuracy during prediction. This can provide medical physicians with a universal model that can predict correctly for any new patient. This work attempted to showcase that it is possible to predict correctly without a prior knowledge of a particular patient with an accuracy of 92.98% but not up to 100%. It should be noted that the training has to contain samples from the medical center. In order to improve the accuracy, the number of samples has to increase in the training set. This limitation can be addressed through combining different datasets from other resources, given that the number of channels used during data collection is the same. Another future direction is the hardware implementation of the automated system on a portable device. A portable device connected to the cloud can be very help to patients and physicians to reduce the impact of epilepsy on their daily lives. Hardware implementation of AlexNet is done in the

Table 4 Results of the proposed work compared to similar work in the literature

Work	Classifier	Input/feature	Dataset	Classes	Acc	Sen	Spec
*	DNN	Matrix (11500x179)	Bonn University	2	97.21	98.59	91.47
†	EESC	PSDED	CHB-MIT dataset	4	95.0	94.9	98.1
‡	CNN	FFT of raw EEG	Freiburg and CHB-MIT dataset	2	96.7	96.7	96.8
				2	95.4	97.2	97.2
				3	92.3	–	–
		Time		2	91.1	91.1	91
				2	83.8	80.4	87.1
				3	85.1	–	–
This Work	CNN (AlexNet)	Spectrogram	Fragility Multi-Center Retrospective Study	3	94.44	92.69	96.05
				2	98.25	97.5	97.5

* [1]
† [40]
‡ [47]

literature, and it can be either implemented on GPUs or FPGAs. Thus, a new hardware implementation can be realized improving on existing implementation to provide a full portable system. This topic is one of the hot research work as part of the edge AI computing revolutionizing embedded systems [6]. As part of future directions also, other novel types of deep learning with promising hardware implementation features such as high performance and low power are the use of neuromorphic computing. Recent work done in neuromorphic computing and spike-based neural network showed important capabilities and promising potential toward integrating AI and ML accelerators for embedded systems and edge computing [41–46].

7 Conclusion

This work presented the training of a convolutional neural network model to classify three epileptic periods using a dataset of EEG signals from 3 different medical centers. The framework utilizes a feature extraction technique, namely PSDED, that exploits multiple channels of EEG signals to minimize information loss. A high performing CNN, namely AlexNet, is used to produce high prediction accuracy. The paper focused on comparing the performance of the classification model based on different data splitting methods, simulating various kinds of prior knowledge that are used to train the model. Three different splitting approaches are considered, namely the totally random sample split, a patient-assisted patient–sample split and a solely patient split. The proposed framework is evaluated by calculating the accuracy, sensitivity, precision and F1 score per class and for each splitting method. The results show that identifying the three epileptic zones using the proposed framework achieves accuracies of 88.89%, 94.44% and 92.98% for the three scenarios, respectively. Furthermore, a binary classification of the last two scenarios produces up to 98% accuracy. Finally, the given outcomes can be used as a guideline on the best way to split the dataset when machine learning is used for prediction.

References

1. D.K., T., B.G, P., F. Xiong, Auto-detection of epileptic seizure events using deep neural network with different feature scaling techniques. Pattern Recog. Lett. **128**, 544–550 (2019) https://doi.org/10.1016/j.patrec.2019.10.029
2. U.R. Acharya, Y. Hagiwara, H. Adeli, Automated seizure prediction. Epilepsy Behav. **88**, 251–261 (2018). https://doi.org/10.1016/j.yebeh.2018.09.030
3. U.R. Acharya, S.L. Oh, Y. Hagiwara, J.H. Tan, H. Adeli, Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. Comput. Biol. Med. **100**, 270–278 (2018). https://doi.org/10.1016/j.compbiomed.2017.09.017
4. U.R. Acharya, Y. Hagiwara, S.N. Deshpande, S. Suren, J.E.W. Koh, S.L. Oh, N. Arunkumar, E.J. Ciaccio, C.M. Lim, Characterization of focal EEG signals: a review. Futur. Gener. Comput. Syst. **91**, 290–299 (2019). https://doi.org/10.1016/j.future.2018.08.044
5. T. Åkerstedt, M. Gillberg, Sleep duration and the power spectral density of the EEG. Electroencephalogr. Clin. Neurophysiol. **64**(2), 2520 (1986). https://doi.org/10.1016/0013-4694(86)90106-9
6. F. Al-Ali, T.D. Gamage, H.W.T.S. Nanayakkara, F. Mehdipour, S.K. Ray, Novel casestudy and benchmarking of AlexNet for edge AI: from CPU and GPU to FPGA. In: Canadian Conference on Electrical

and Computer Engineering, vol. 2020-Augus (2020). https://doi.org/10.1109/CCECE47787.2020.9255739

7. E. Alickovic, J. Kevric, A. Subasi, Performance evaluation of empirical mode decomposition, discrete wavelet transform, and wavelet packed decomposition for automated epileptic seizure detection and prediction. Biomed. Signal Process. Control **39**, 94–102 (2018). https://doi.org/10.1016/J.BSPC.2017.07.022
8. K.M. Almustafa, Classification of epileptic seizure dataset using different machine learning algorithms. Inf. Med. Unlocked **21**, 100444 (2020). https://doi.org/10.1016/j.imu.2020.100444
9. R.G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, C.E. Elger, Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics **64**(6), 8 (2001). https://doi.org/10.1103/PhysRevE.64.061907
10. S. Appelhoff, M. Sanderson, T. Brooks, M. Vliet, R. Quentin, C. Holdgraf, M. Chaumon, E. Mikulan, K. Tavabi, R. Höchenberger, D. Welke, C. Brunner, A. Rockhill, E. Larson, A. Gramfort, M. Jas, MNE-BIDS: organizing electrophysiological data into the BIDS format and facilitating their analysis. J. Open Source Softw. **4**(44), 1896 (2019). https://doi.org/10.21105/joss.01896
11. E. Bou Assi, D.K. Nguyen, S. Rihana, M. Sawan, Towards accurate prediction of epileptic seizures: a review. Biomed. Signal Process. Control **34**, 144–157 (2017). https://doi.org/10.1016/j.bspc.2017.02.001
12. CHB-MIT: CHB-MIT scalp EEG database. (2000). https://physionet.org/content/chbmit/1.0.0/
13. Freiburg: Freiburg seizure prediction project. Freiburg, Germany. https://epilepsy.uni-freiburg.de/freiburg-seizure-prediction-project/eeg-database
14. K. Gadhoumi, J.-M. Lina, F. Mormann, J. Gotman, Seizure prediction for therapeutic devices: a review. J. Neurosci. Methods **260**, 270–282 (2016). https://doi.org/10.1016/j.jneumeth.2015.06.010
15. Y. Gao, B. Gao, Q. Chen, J. Liu, Y. Zhang, Deep convolutional neural network-based epileptic electroencephalogram (EEG) signal classification. Front. Neurol. **11**, 375 (2020). https://doi.org/10.3389/fneur.2020.00375
16. A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.-K. Peng, H.E. Stanley, PhysioBank, physioToolkit and physioNet. Circulation **101**(23), 215 (2000). https://doi.org/10.1161/01.cir.101.23.e215
17. C. Holdgraf, S. Appelhoff, S. Bickel, K. Bouchard, S. D'Ambrosio, O. David, O. Devinsky, B. Dichter, A. Flinker, B. L. Foster, K. J. Gorgolewski, I. Groen, D. Groppe, A. Gunduz, L. Hamilton, C. J. Honey, M. Jas, R. Knight, J.-P. Lachaux, J. C. Lau, C. Lee-Messer, B. N. Lundstrom, K. J. Miller, J.G. Ojemann, R. Oostenveld, N. Petridou, G. Piantoni, A. Pigorini, N. Pouratian, N. F. Ramsey, A. Stolk, N. C. Swann, F. Tadel, B. Voytek, B. A. Wandell, J. Winawer, K. Whitaker, L. Zehl, D. Hermes, BIDS-iEEG: an extension to the brain imaging data structure (BIDS) specification for human intracranial electrophysiology. Sci. Data 6(102), 26 (2019)
18. W. Hu, J. Cao, X. Lai, J. Liu, Mean amplitude spectrum based epileptic state classification for seizure prediction using convolutional neural networks. J. Ambient Intell. Hum. Comput. (2019). https://doi.org/10.1007/s12652-019-01220-6
19. S. Ibrahim, S. Majzoub, Adaptive epileptic seizure prediction based on EEG synchronization. J. Biomim. Biomater. Biomed. Eng. **33**, 52 (2017) https://doi.org/10.4028/www.scientific.net/JBBBE.33.52
20. E.M. Imah, A. Widodo, A comparative study of machine learning algorithms for epileptic seizure classification on EEG signals. In: 2017 International Conference on Advanced Computer Science and Information Systems (ICACSIS), pp. 401–408 (2017). https://doi.org/10.1109/ICACSIS.2017.8355065
21. P. Jahankhani, V. Kodogiannis, K. Revett, EEG Signal classification using wavelet feature extraction and neural networks. In: IEEE John Vincent Atanasoff 2006 International Symposium on Modern Computing (JVA'06), pp. 120–124 (2006). https://doi.org/10.1109/JVA.2006.17
22. H. Khan, L. Marcuse, M. Fields, K. Swann, B. Yener, Focal onset seizure prediction using convolutional networks. IEEE Trans. Biomed. Eng. **65**(9), 2109–2118 (2018). https://doi.org/10.1109/TBME.2017.2785401
23. M.A. Kramer, E.D. Kolaczyk, H.E. Kirsch, Emergent network topology at seizure onset in humans. Epilepsy Res. **79**(2–3), 173–186 (2008). https://doi.org/10.1016/j.eplepsyres.2008.02.002
24. A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks. Commun. ACM **60**(6), 84–90 (2017). https://doi.org/10.1145/3065386

25. L. Kuhlmann, K. Lehnertz, M.P. Richardson, B. Schelter, H.P. Zaveri, Seizure prediction—ready for a new era. Nat. Rev. Neurol. **14**(10), 618–630 (2018). https://doi.org/10.1038/s41582-018-0055-2
26. Y. Kumar, M.L. Dewal, R.S. Anand, Epileptic seizures detection in EEG using DWT-based ApEn and artificial neural network. Signal Image Video Process. **8**(7), 1323–1334 (2014). https://doi.org/10.1007/s11760-012-0362-9
27. A. Li, C. Huynh, Z. Fitzgerald, I. Cajigas, D. Brusko, J. Jagid, A.O. Claudio, A.M. Kanner, J. Hopp, S. Chen, J. Haagensen, E. Johnson, W. Anderson, N. Crone, S. Inati, K.A. Zaghloul, J. Bulacio, J. Gonzalez-Martinez, S.V. Sarma, Neural fragility as an EEG marker of the seizure onset zone. Nat. Neurosci. **24**(10), 1465–1474 (2021). https://doi.org/10.1038/s41593-021-00901-w
28. M. Moshinsky, Characterization of focal EEG signals: A review — Science direct (1959). https://www.sciencedirect.com/science/article/pii/S0167739X18318818
29. S. Opałka, B. Stasiak, D. Szajerman, A. Wojciechowski, Multi-channel convolutional neural networks architecture feeding for effective EEG mental tasks classification. Sensors (Switzerland) **18**(10), (2018) https://doi.org/10.3390/s18103451
30. C.R. Pernet, S. Appelhoff, K.J. Gorgolewski, G. Flandin, C. Phillips, A. Delorme, R. Oostenveld, EEG-BIDS, an extension to the brain imaging data structure for electroencephalography (2019). https://doi.org/10.1038/s41597-019-0104-8
31. F. Pisano, G. Sias, A. Fanni, B. Cannas, A. Dourado, B. Pisano, C.A. Teixeira, Convolutional neural network for seizure detection of nocturnal frontal lobe epilepsy. Complexity **2020**, 4825767 (2020). https://doi.org/10.1155/2020/4825767
32. T. Radüntz, J. Scouten, O. Hochmuth, B. Meffert, Automated EEG artifact elimination by applying machine learning algorithms to ICA-based features. J. Neural Eng. **14**, 46004 (2017). https://doi.org/10.1088/1741-2552/aa69d1
33. M. Sharma, R.B. Pachori, U. Rajendra Acharya, A new approach to characterize epileptic seizures using analytic time-frequency flexible wavelet transform and fractal dimension. Pattern Recognit. Lett. **94**, 172–179 (2017). https://doi.org/10.1016/j.patrec.2017.03.023
34. M. Sharma, R.B. Pachori, U. Rajendra Acharya, A new approach to characterize epileptic seizures using analytic time-frequency flexible wavelet transform and fractal dimension. Pattern Recognit. Lett. **94**, 172–179 (2017). https://doi.org/10.1016/J.PATREC.2017.03.023
35. A. Shoeibi, M. Khodatars, N. Ghassemi, M. Jafari, P. Moridian, R. Alizadehsani, M. Panahiazar, F. Khozeimeh, A. Zare, H. Hosseini-Nejad, A. Khosravi, A.F. Atiya, D. Aminshahidi, S. Hussain, M. Rouhani, S. Nahavandi, U.R. Acharya, Epileptic seizures detection using deep learning techniques: a review. Int. J. Environ. Res. Pub. Health 2021, Vol. 18, Page 5780 18(11), 5780 (2021) https://doi.org/10.3390/IJERPH18115780, arXiv:2007.01276
36. M.K. Siddiqui, R. Morales-Menendez, X. Huang, N. Hussain, A review of epileptic seizure detection using machine learning classifiers. Brain Inf. **7**(1), 1–18 (2020). https://doi.org/10.1186/s40708-020-00105-1
37. Y. Song, J. Crowcroft, J. Zhang, Automatic epileptic seizure detection in EEGs based on optimized sample entropy and extreme learning machine. J. Neurosci. Methods **210**(2), 132–146 (2012). https://doi.org/10.1016/j.jneumeth.2012.07.003
38. A. Subasi, J. Kevric, M. Abdullah Canbaz, Epileptic seizure detection using hybrid machine learning methods. Neural Comput. Appl. **31**(1), 317–325 (2019). https://doi.org/10.1007/s00521-017-3003-y
39. P. Thodoroff, J. Pineau, A. Lim, Learning robust features using deep learning for automatic seizure detection. (2016) CoRR **abs/1608.0**
40. N.D. Truong, A.D. Nguyen, L. Kuhlmann, M.R. Bonyadi, J. Yang, S. Ippolito, O. Kavehei, Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram. Neural Netw. **105**, 104–111 (2018). https://doi.org/10.1016/j.neunet.2018.04.018
41. S. Yang, B. Deng, J. Wang, H. Li, M. Lu, Y. Che, X. Wei, K.A. Loparo, Scalable digital neuromorphic architecture for large-scale biophysically meaningful neural network with multi-compartment neurons. IEEE Trans. Neural Netw. Learn. Syst. **31**(1), 148–162 (2020). https://doi.org/10.1109/TNNLS.2019.2899936
42. S. Yang, J. Tan, B. Chen, Robust spike-based continual meta-learning improved by restricted minimum error entropy criterion. Entropy **24**(4), 24040455 (2022). https://doi.org/10.3390/e24040455
43. S. Yang, B. Linares-Barranco, B. Chen, Heterogeneous ensemble-based spike-driven few-shot online learning. Front. Neurosci. **16**, 850932 (2022). https://doi.org/10.3389/fnins.2022.850932

44. S. Yang, T. Gao, J. Wang, B. Deng, M.R. Azghadi, T. Lei, B. Linares-Barranco, SAM: a unified self-adaptive multicompartmental spiking neuron model for learning With working memory. Front. Neurosci. **16**, 850945 (2022). https://doi.org/10.3389/fnins.2022.850945
45. S. Yang, J. Wang, B. Deng, M.R. Azghadi, B. Linares-Barranco, Neuromorphic context-dependent learning framework with fault-tolerant spike routing. IEEE Trans. Neural Netw. Learn. Syst. **33**(12), 7126–7140 (2022). https://doi.org/10.1109/TNNLS.2021.3084250
46. S. Yang, J. Wang, X. Hao, H. Li, X. Wei, B. Deng, K.A. Loparo, BiCoSS: toward large-scale cognition brain with multigranular neuromorphic architecture. IEEE Trans. Neural Netw. Learn. Syst. **33**(7), 2801–2815 (2022). https://doi.org/10.1109/TNNLS.2020.3045492
47. M. Zhou, C. Tian, R. Cao, B. Wang, Y. Niu, T. Hu, H. Guo, J. Xiang, Epileptic seizure detection based on EEG signals and CNN. Front. Neuroinform. **12**, 95 (2018). https://doi.org/10.3389/fninf.2018.00095