Journal of Computer & Robotics 16 (2), Summer and Autumn 2023, 37-48

Computer
& Robotics

Epileptic Seizure Prediction using Multi-Channel Raw EEGs with Convolutional Neural Network

Jamal Nazari [a], Ali Motie Nasrabadi [b,*], Mohammad Bagher Menhaj [c], Somayeh Raiesdana [a]

[a] *Faculty of Electrical, Biomedical and Mechatronics Engineering, Qazvin , Branch, Islamic Azad University, Qazvin, Iran*

[b] *Department of Biomedical Engineering, Faculty of Engineering, Shahed University, Tehran, Iran*

[c] *Department of Electrical Engineering, Amirkabir University of Technology, Tehran, Iran*

Received 07 March 2023; Accepted 18 March 2023

Abstract

Epileptic seizure prediction has been one of the interesting topics among researchers in recent years. Recent evidence suggests that, in many seizures, changes in the preictal signal begin minutes before the ictal begins, raising hopes of predicting the seizure onset before it occurs. Convolutional neural network (ConvNet) is a powerful computational tool with deep learning capacity which is able to detect complex structures in data. In this study, we employed a ConvNet and a set of techniques to make optimal use of the existing data for an end-to-end learning. Multi-channel non-invasive raw EEGs from the CHB-MIT database were used for training of the proposed model. The proposed method resulted in sensitivity of 92.05% and false prediction rate of 0.073/h with the cross-validation approach in distinguishing preictal and ictal. We obtained a 10-minute seizure prediction horizon that is relatively higher than the values obtained in other researches. This longer time period can give the patient more opportunity for preventive actions. Seizure occurrence period was computed nearly 20 minutes which lets the patient wait less for the seizure to occur and this in turn makes him have less anxiety. Furthermore, a feature map visualizing method was employed in the present work to decode the employed deep network and to understand how it learns and what it learns when trying to solve the seizure prediction task. By investigating feature maps of the used ConvNet's middle layer, we observed that the proposed network retains most of the beta and gamma band properties in layers.

Keywords: Seizure Prediction, Epilepsy, ConvNet, EEG

1.Introduction

Epileptic seizures are transient signs or symptoms of abnormal, intense, and synchronous activities of the nervous system caused by electrical discharge of neurons. The percent of the time spent with potential disease related behavioral symptoms is very small and, in some patients, seizures may occur every few months or every few years. Despite the low frequency of clinical symptoms, uncontrollable seizures have profound effects on the patient's life. The effects of seizures are due to their occasional behavioral occurrence and the unpredictability of more than 99.95% when the patients are not having seizures and should be able to live relatively normal, but the stress of seizure may affect their daily lives. It also deprives them of a number of activities, such as driving, swimming, etc., and reduces their quality of life [1].

Therefore, the ability to accurately predict seizures can make significant changes in the lives of people with epilepsy and give them more confidence and

*Corresponding Author. Email: a_m_nasrabadi@yahoo.com

freedom and, it can reduce sudden deaths in patients with epilepsy. These patients could also take the medicine when needed and not constantly. The electroencephalogram (EEG) signal has a higher time resolution than other brain imaging methods and it is mostly used to predict epileptic seizures. EEG signal is basically known as a multivariate time series of a nonlinear and multidimensional system. Noteworthy, the complex and uncertain relationship between timeseries can be realized only with nonlinear functions with a high degree of freedom. Nowadays, with the advances in machine learning, powerful algorithms such as Convolutional Networks have been proposed capable of performing well in natural language processing, object detection and classification. They are also a powerful tool for discovering complex structures in data [2].

One of the hypotheses for predicting epilepsy is that the changes in brainwave patterns occur as the ictal states are reached. There are two perspectives for identifying these changes during the preictal interval. In the first perspective, only the preictal interval is analyzed and it is compared to a predefined threshold level [3]. In the second perspective, the differentiating patterns between the preictal and the interictal intervals are identified and then a binary classification is employed [4]. The commonality between these two perspectives is the extraction of the best features from the EEG signal.

Seizure prediction horizon (SPH) is the interval between the alarm and the onset of the seizure occurrence period (SOP). For the alarm to be true, seizures should begin after SPH and within the SOP. A schematic in Figure 1 shows the intervals of SPH and SOP. Two metrics that are often used to evaluate the predictive systems are false prediction rate (FPR) and sensitivity. FPR is the number of incorrect alarms per hour when they a seizure is predicted but no seizure occurs in the SOP [4]. Sensitivity is defined as the percentage of correctly predicted seizures on the total number of seizures. In clinical use, more SPH is considered to confront and to deal with vital actions and, less SOP is to reduce patient anxiety.

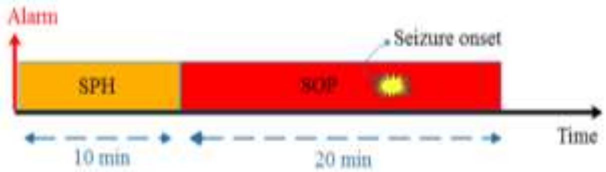

Fig. 1. A schematic shows the seizure occurrence period and the seizure prediction horizon

In the majority of work done to diagnose and predict epileptic seizures, a number of time domain features (such as median, mean, variance, standard deviation, maximum and minimum, etc.) and/or frequency domain features (such as power spectrum density, etc.) and/or time-frequency domain features (such as wavelet transform coefficients, Pseudo-Winger-Will, etc.) and/or chaotic features (such as fractal dimension, the largest Lyapunov exponent, approximate entropy and spectral entropy and correlation dimension, etc.) are extracted. A combination of such features has also been considered by some researchers. In a work [3], researchers predicted epileptic seizures by introducing a similarity index based on symbolic dynamics techniques (statistical behavior of local extremes). They reported sensitivity of 63.75% and FPR of 0.33/h for data of 21 patients from Freiburg database[1] and sensitivity of 96.66% and FPR of 0.33/h for a subset of eight patients. In the last few years, various methods have been proposed to select the most appropriate combination of features and classifiers, including extracting linear features from the EEG signal using autoregressive coefficients [5]. The emergence of dynamical systems theory introduced a number of discriminative nonlinear features including Lempel-Ziv, noise level, correlation entropy complexity, and correlation dimension. Usage of such measures on intracranial EEG resulted in sensitivity of 86.7% and 92.9% with FPR of 0.126/h and 0.096 /h for SOP=30.50 minute and 10-sec forecast horizon [6]. Tailored feature extraction was customized and performed independently for each patient of the CHB-MIT database resulting in a sensitivity of 98% and an FPR of less than 0.05/h [8]. Furthermore, sensitivity of 98.52% and FPR of 0.04/h was obtained with few-shot learning [9]. A research team achieved FPR = 0.11-0.02/h and sensitivity of 99% by using the

[1] https://physionet.org/content/chbmit/1.0.0/

features of statistical moments, zero crossings, Wavelet Transform Coefficients, PSD, graph theory, cross-correlation and using Long short-term memory (LSTM) for CHB-MIT data for 15-120 minutes SOP and zero SPH [13]. In another work, a patient-specific method using the common spatial pattern (CSP) for feature extraction with linear discriminant analysis (LDA) classifier was reported which resulted in 89% sensitivity, FPR = 0.39/h, and 120min SPH for 24 subjects of the CHB-MIT database [10]. In [11], authors presented a patient-specific prediction algorithm using multiple features of spectral power EEG signal and a support vector machine (SVM) for classification. They reported sensitivity of 97.5% and FPR = 0.27/h for 18 patients of the Freiburg dataset. Generally, patient-specific feature-based platforms resulted in high sensitivity and low FPR. Since the best feature combination is extracted independently for each patient, patient-specific systems are more reliable. However, the need for a specialist and being time consuming, as well as the changes required for each new patient are some of the problems against generalizability. Furthermore, given the changing brain dynamics, the selected feature subset may not work well for a new patient in future.

Manual extraction of features is not only time consuming but also imperfect. When faced with a wide range of data, it is challenging to engineer features and achieve high-level features. Generalized networks remove this constraint and allow data features to be extracted and learned without explicit structural information, and in fact create an automated feature extraction path. In [12], authors have used the Recurrent Neural Network (RNN) to learn temporal dependencies between successive samples. In [14], by using resting-state functional magnetic resonance imaging (*rs-fMRI*) and EEG data and LSTM computational tool, 96% sensitivity was achieved. In another study, authors used a convolution neural network (CNN) on Functional near-infrared spectroscopy (fNIRS) and EEG data of 49 patients. They reached 95.24%-100% sensitivity in predicting seizures [15]. In [16], by using CNN and SVM tools on dataset of five dogs and two patients, researchers achieved 0.72% sensitivity in prediction. Another study trained a CNN based on the wavelet coefficients of the EEG signals and achieved a sensitivity of 87.8% and an FPR of 0.147/h on epileptic patients of the CHB-MIT dataset [17]. However, a reliable predictive model to be tested over a larger dataset and be capable of considering variabilities between patients is still required. Such a reliable algorithm can be employed in an assistive alarm system in future.

In this work, we present a ConvNet-based model for predicting epileptic seizures using various techniques in order to use the existing data to optimally. We focus on achieving proper SPH and SOP using multi-channel EEG signals without the need to access hand-crafted features or channel selection.

Deep network architecture allows the reuse of features (mid-level features that are shared between all classes). It also has the potential to create high-level features. Generally, lack of sufficient training samples to efficiently train a ConvNet in predicting epileptic seizures is an issue. In this paper, we intend to use an architecture that makes effective use of existing training instances. We also employ some techniques to achieve higher accuracy with a shallow architecture and fewer parameters. There are different perspectives on applying a convolutional network to EEG data. In some cases, a time-frequency transform like wavelet transform or STFT [4] is applied on multichannel EEG and next the resulting matrix is fed into the network. Another view is to use raw EEGs without any feature extraction which as an end-to-end learning. One of the problems with the deep networks is that they act like a black box, and we have little knowledge of what they are learning. The concept ofvisualizing feature maps helps to figure out which input features are preserved and to recognize the patterns that the filters respond to. Feature maps in the ConvNet are the result of convolution of the filters with input. Previously it was shown that [24] the change of power spectrum in feature map can differentiate interictal. An abnormal increase in the amplitude of gamma-band frequencies before seizures was reported for five epileptic patients in [24]. In this work, we also check to see which frequency band contains more information in ictal/preictal discrimination. For this reason, we take short-time

Fourier transform (STFT) from the middle layer feature maps for each EEG band and then identify the EEG frequency band(s) that have a greater amplitude in the feature maps by evaluating the time-frequency spectrum. This procedure. A statistical analysis is further conducted to show the significance of differences between interictal and ictal periods.

The rest of this paper is as follows. In section 2, materials and methods have been presented. In section 3 results are reported. Then, remakes and concluding points are brought in discussion and conclusion sections respectively.

2. Materials and Methods

We employed Boston Children's Hospital (CHB)-MIT dataset [24] in this work. This data contains scalp EEGs from 23 pediatric patients with 844h of continuous EEG recording and 163 seizures. All signals were sampled at 256 samples per second with a 16-bit resolution. The EEG signals were captured with 22 electrodes by the 10-20 recording protocol. In this study, we considered only those patients whose signal was recorded at least 30 minutes before the seizure onset. For this reason, the usable data has decreased to 18 ones. The employed channels included F7-T7, T7-P7, P7-O1, FP1-F3, F3-C3, C3-P3, P3-O1, FZ-CZ, CZ-PZ, FP2-F4, F4-C4, C4-P4, P4-O2, FP2-F8, F8-T8, T8-P8 covering the whole brain. Used the LOOCV perspective to assess the generalizability of the results. Table 1 shows the sensitivity and FPR for SPH=10min and SOP=20min for each patient and their mean.

2.1. Preprocessing

Our method is based on two-class classification to distinguish the preictal from the interictal. Since we use 16 raw EEG channels simultaneously, we need to use a two-dimensional convolution neural network (2DConvNet). But it is important to know how far the preictal interval is from the onset of the seizure because the longer the prognosis horizon, the better for the patient since there is enough time to confront and to deal with vital actions and, the seizure occurrence period should not be long so that the patient have less anxiety. Based on the results obtained to select the most appropriate length of the prediction horizon, about 10 minutes before the onset of the seizure is the best prediction horizon [17].

We considered the signal of 10 to 30 minutes before the onset of the seizure as the preictal interval. The signal was passed through a band-pass filter with a middle frequency of 0.5-100Hz. The 60Hz frequency power line noise was also eliminated via a notch filter. One of the important issues in a classification task is the balance of the dataset which means that we should have an equal number of data points in each class. Due to the fact that the signal length is often different for each patient in the recorded data, we need to use overlap to select data from signals with shorter length and use it for training. It is noteworthy that we should not use overlap to select the test data, and for balancing, we consider a signal with a shorter length as a basis.

Therefore, after normalization, by passing a sliding window with a length of 5 seconds with variable overlap, we select segments with a length of 1280 samples from the preictal and interictal intervals in equal proportions due to the balance of data of the both classes (Figure 2). As a result, each segment is a two-dimensional matrix with a size of 16×1280, which means that each channel with a length of five seconds (1280 samples) is located in a row.

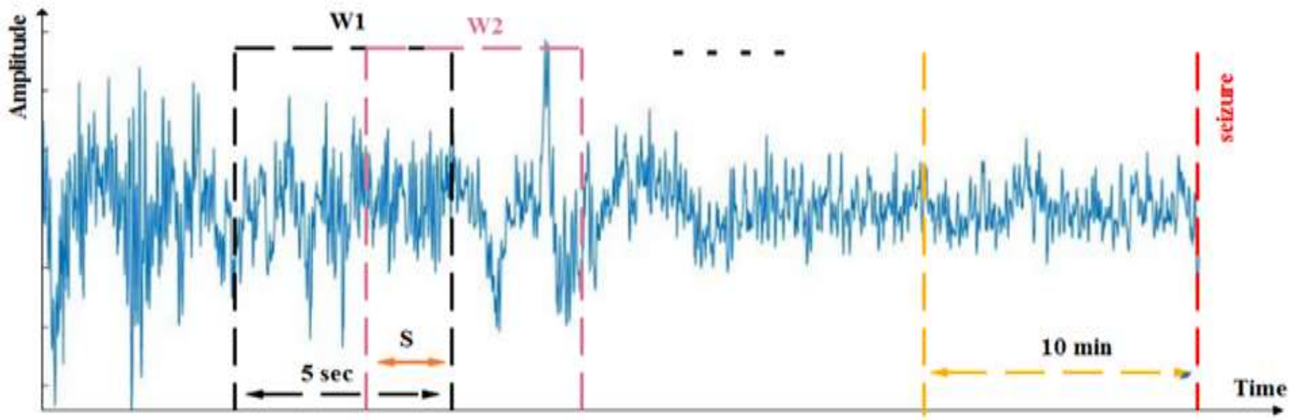

Fig. 2. 5-second sliding window with variable overlap S that 0≤ S ≤0.5.h

2.2. ConvNet Architecture

In this study, we used a ConvNet which is a subtype of feed-forward neural network, inspired from the animal's visual cortex. ConvNets are a powerful representation learning tool that can detect complex structures in data. They require little preprocessing, meaning that the model itself is responsible for learning the features that are extracted manually in traditional algorithms [16].

When we face high-dimensional inputs, the connection of the neurons to all of the neurons of their previous layer is impractical. Therefore, each neuron is connected to only a small area of the neurons of the previous layer which is called the receptive field. It has been shown that in deep networks the low-level features are learned in the early layers and higher-level concepts are learned as they deepen. It is a logical view that by adding more layers, we will be able to learn more concepts [18]. However, as the network deepens, the number of learnable parameters of the network also increases, and it can be said that increasing capacity does not always improve performance. The second influential factor is the availability of sufficient data for the model. Otherwise, the model only specializes in training data and will not perform well in the test data and in fact "overfitting" will occur. Therefore, there must always be a reasonable proportion between the capacity of the model and the volume of the data. One of the challenges in EEG signal analysis to predict epileptic seizures is the lack of

large datasets. Hence, one of our efforts in this work is to apply some strategies on the adopted architecture in order to use the existing data in an optimal way.

The schematic of a two-layer convolutional network has been shown in figure 3. Input to this network is a row EEG signal with size (n ×16 × 1280) where n is the number of input datasets. There are two layers of convolution each containing a 3×3 kernel and single stride. Afterwards, the convolution output is first flattened and then it is followed by a fully connected layer with two hidden layers and one output layer is used. In every two hidden layers. In this network, the two neurons of the last layers which will determine the classes of "preictal" or "interictal" use the soft-max activation function and the rest of the layers use the Leaky ReLU activation function. More details on layers and training strategies are as follows.

2.2.1. Homogeneous and Symmetrical Filters

Each ConvNet layers consist of a set of learnable filters, which have a small receptive field, but extend through the full depth of the input volume. During the forward pass, each filter is convolved with the input matrix. By computing the dot product between the filter matrix and an input, a feature map is produced. The results have shown that the use of a large number of small filters yields better nonlinear outcomes and higher accuracy [19]. Using homogeneous and symmetric architecture has also been proposed by [20]. In the architecture of the proposed network in this paper, 3*3 filters are used in different layers.

2.2.2. Pooling

Pooling performs downsampling operations in the spatial dimension. The result is a reduction in the number of parameters and consequently a reduction in computational volume and prevention from overfitting. The Pooling layer helps to make the representations of the higher layer immutable in comparison to small displacements in the input and, it also uses inputs of different sizes to do some tasks [21]. However, improper use reduces its performance. It is believed that large feature maps, especially in the primary layers, provide more information on the network than the smaller ones [20]. Networks with the same depth and parameter that use large feature maps are more accurate [2], so the use of fast downsampling is not recommended, especially in the primary layers.

2.2.3. Dropout

Dropout is a new regularization strategy proposed by Hinton and Srivastava [22]. A number of nodes are usually dropped out in each stage and then they are returned in the next stage in order to prevent overfitting. It is also interpreted as a set of several networks that are trained with different subsets of training data. The effect that dropout in the convolution layer has on the robustness of the network is proportional to noise input. It was suggested that dropout is applied to all convolution layers instead of applying solely on the fully-connected (FC) layer. This seems to increase accuracy and generalization of a deep network [20]. We apply this technique to all layers of the proposed network.

2.2.4. Batch Normalization

Batch normalization is a technique for faster training and improves the accuracy. It encompasses normalizing the input to neural networks. It also normalizes the activations in intermediate layers to have zero mean and unit variance [23]. This technique was used in all layers of the proposed network.

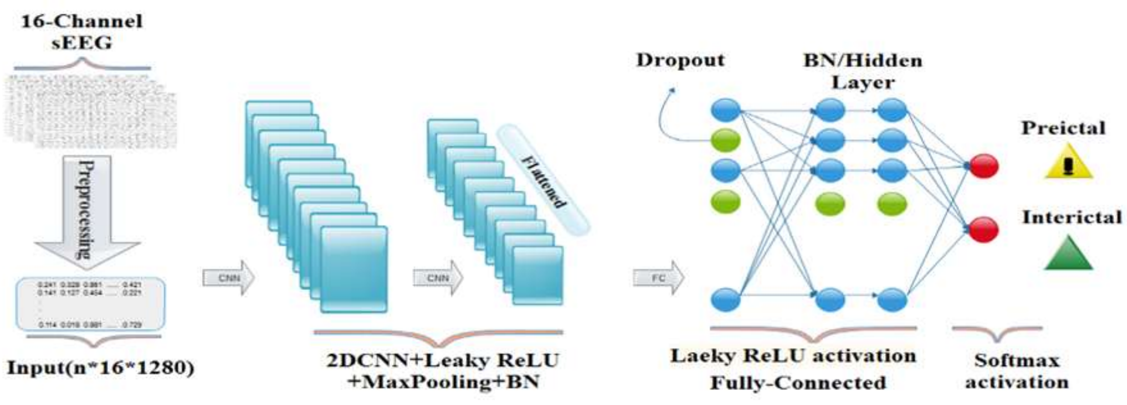

Fig. 3. Two layers of the convolutional layer are applied to the input, which is an EEG signal with size (n * 16 * 1280), n is the number of input datasets. Each convolution layer contains a 3*3 kernel and single stride. Dropout and batch normalization is applied to all layers. A max-pooling layer of 2*2 size is then applied in with two strides. Afterwards, the convolution output is first flattened and then an FC (fully connected) with two hidden layers and one output layer is used. In every two hidden layers, the Dropout and Batch normalization is performed. The two neurons of the last layers use the soft-max activation function and the rest of the layers use the Leaky ReLU activation function.

2.3. Evaluation Method

Cross-validation is a model evaluation technique that determines the extent to which the results of a used when constructing the model. These data are used to evaluate and measure the performance of the model to predict new data. Thus, we resort to estimating the statistical analysis on a data set can be generalized and be independent of the training data. This method relies on the data that are observed but not model error based on the data that have been set aside for cross-validation to measure the efficiency of the model and its optimality. The leave-one-out cross-validation (LOOCV) technique uses all data, except one, for training and the residuaring data for the method testing. This garlic is repeated N times; N presents the number of data folders. With this work all data will be used to train and test the method. The error rate of the method is equal to the average error rate per iteration.

3. Results

Simulations were done on a core i7 pc with 16 GB RAM and 512 GB SSD. We trained the network in google collaboratory (Google Colab). The results of the implemented network with leave-one-out cross-validation for SOP=20 min and SPH=10 min are tabulated to Table 1 for 18 patients of the introduced dataset. Information for the 18 processed recordings were presented in the first three columns of table 1. Notably, the LOOCV technique was used as described before to provide generalizability for the network. Table 1 shows the sensitivity and FPR for SPH=10min and SOP=20min for each patient separately and in average as well.

Table 1
CHB-MIT 18 cases (Multi-channel raw-EEG) Results of the proposed model

Patient Name	Seizure	Gender	Age(years)	Sensitivity (%)	FPR (/h)
Chb01	7	F	11	94.09	0.10
Chb02	3	M	11	95.00	0.03
Chb03	7	F	14	93.18	0.16
Chb04	4	M	22	95.45	0.11
Chb05	5	F	7	91.36	0.03
Chb06	10	F	1.5	92.72	0.05
Chb07	3	F	14.5	84.34	0.08
Chb09	4	F	10	83.63	0.03
Chb10	7	M	3	95.90	0.06
Chb14	9	F	9	91.98	0.14
Chb15	10	M	16	98.09	0.03
Chb16	7	F	7	88.54	0.03
Chb17	3	F	12	86.73	0.08
Chb18	6	F	18	98.18	0.11
Chb19	3	F	19	99.09	0.07
Chb20	5	F	6	85.90	0.04
Chb21	4	F	13	98.72	0.11
Chb22	3	F	9	84.00	0.07
Avg.				**92.05**	**0.07**

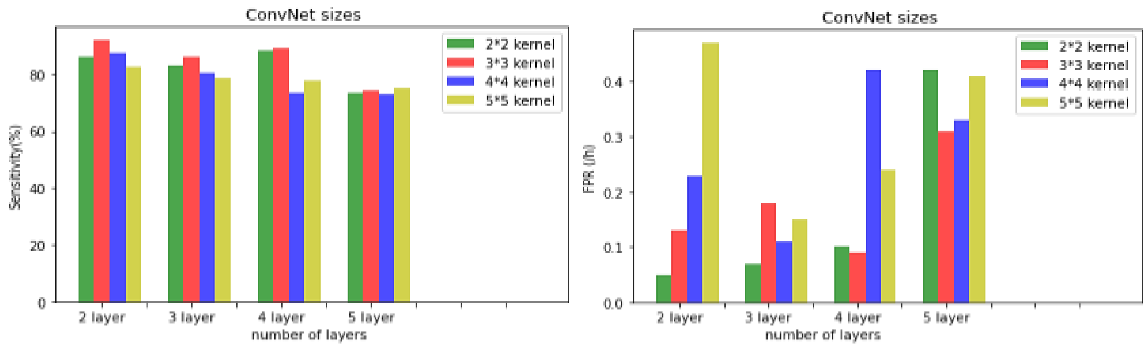

Fig. 4 (a). Average model sensitivity (%) for different number of layers and kernels of different sizes (b). Average model FPR(/h) for the number of different layers and kernels of different sizes .

In order to assess the performance of the proposed method, we performed some evaluation tests. The first one evaluates the performance of the network versus its structure. Figure 4 shows the average model sensitivity and FPR vs the number of layers of the network for different kernel sizes. It shows that by increasing the number of layers and consequently by increasing the capacity of the network, results get worse. The reason is that the number of usable data is little and the network is probably overfitted. Therefore, we do not go beyond the two layers for such a network and the classification task. with the above model and two layers of convolution, the results are acceptable and comparable to existing systems in the field.

In the next test of the proposed model, we calculated time-frequency transforms of the generated feature maps in the network to investigate the amplitude of the frequencies in the spectrum. This amplitude was employed as a distinguishing feature between preictal and interictal samples. A sample Short-time Fourier transform (STFT) from the middle layer feature-map was shown in figure 5. The dimension of the feature-map for an input with dimensions of 16*1280, after passing through the convolution layer and max-pooling of 2*2, changed to 8*640, and then was flattened to dimensions of 1*5120. Next, the STFT transform was applied.

An active frequency range is defined, such as the frequency range from A to B in Figure 5, to specify the frequency components that are employed in computation and classification. Figure 6 shows the mean and variance of the feature map transformation that were calculated for each EEG frequency bands including delta (0-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (14-34 Hz), the gamma (34-44 Hz), and for each class including interictal and ictal.

The analysis of feature maps of the network in figure 6 shows that the trained ConvNet model makes use of the beta and gamma spectra mostly. Also, it can be inferred that features were most prominent in the beta band for interictal and in the gamma band for preictal.

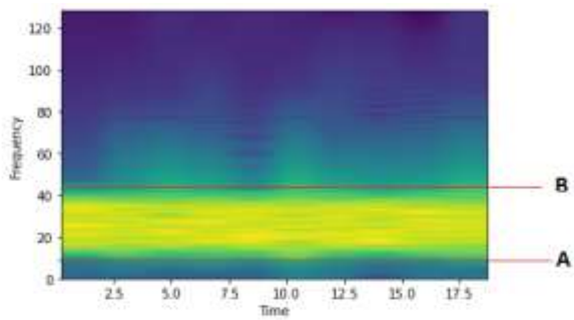

Fig.5. Short-time Fourier transform (STFT) from the middle layer feature-map.

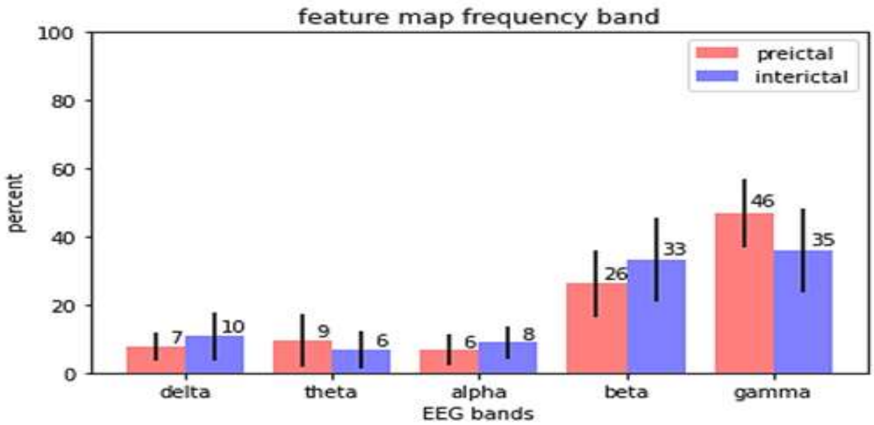

Fig. 6. The average and variance of feature-maps for each of the Delta, Theta, Alpha, Beta and Gamma frequency bands. For interictal input (blue) and preictal (red).

Finally, the proposed method was compared with some state-of-the-art. The results have been reported in Table 2 showing a relatively good sensitivity and FPR for the selected prediction horizon.

However, we can see that the combination method presented in [25] has yielded a relatively high sensitivity of 99.72% and FPR=0.004. Also references [14] and [26] has higher sensitivity (96% and 98.6%, respectively) compared to our method. Noteworthy, the SPH has been set to zero in all of them while we have selected a 10min SPH in our implementation. Furthermore. References [8] and [11] had higher sensitivity (97.5% and 98% respectively) as well although using a patient-specific method that needs to be performed independently for each patient and not generalizable.

Table 2
Comparison of the results of the state-of-the-art.

Authors	Feature	Classifier	Database	Sensitivity(%)	FPR	SOP(min)	SPH(sec)
Niknazar et al. 2016[3]	symbolic dynamic	Threshold	FB	63.75	0.33	45	0
Cho et al. 2017[27]	Phase locking value	SVM	23Chb	82.44	-	5	0
Miri et al. 2011[28]	Zero-crossing rate, astatistical index	Threshold	6FB	86.6	0.06	20±5	0
Aarabi et al. 2017[6]	Bayesian inversion of power spectral density	RBD	21FB	87.07	0.20	30	10
Tsiouris et al.2018[13]	WTC,PSD	LSTM	Chb	90	0.11	15-120	0
Truong et al. 2018[4]	STFT	CNN	13FB 13Chb	81.2 81.4	0.06 0.16	30	300
Alotaiby et al.2017[10]	CSP	LDA	24Chb	89	0.39	120	0
Zisheng et al.2016[26]	Power spectral density	SVM	17Chb	98.68	0.05	50	0
Khan et al.2018[17]	Wavelet transform	CNN	23Chb	87.8	0.14	10	0
Hisham 2019 [25]	Raw EEG	DCAE+ Bi-LSTM	22Chb	99.72	0.004	60	0
Qin et al. 2020[29]	STFT	CNN+ELM	Chb	95.85	0.045	-	-
This work	Raw EEG	ConvNet	18Chb	92.02	0.073	20	600

4. Discussion

Today generalized neural networks and especially ConvNet have shown to be powerful computational tools for predicting epileptic seizures. One of the main advantages of such methods is that they eliminate hand-crafted feature extraction procedures and do not need a feature engineering step which has always been controversial. Furthermore, they have better generalization. However, the use of such networks is almost new and more work is still needed to be done to certify their extensive use.

ConvNet networks can automatically select the best features and as they apply a hierarchical strategy on data in multiple layers, they can pick a range of features from low-level to high-level which well corresponds to a binary classification task.

We used raw EEGs as input in this work and as it was mentioned there is no longer need for a hand-crafted feature extraction step. Noteworthy this is very vital for a real-time task such as seizure prediction paradigm. Our results show that using Leaky ReLU as an activation function and prior batch normalization as well as homogeneous, small size and large number of filters, especially in the first layers, and the use of dropouts in all layers can enhance performance. On the other hand, besides of prioritizing metrics of performance such as sensitivity and FPR, we focused on setting a suitable alarm period in this work. It is obvious that in a predictive system, a patient should receive an alarm some time before seizure onset to have enough time for preventive actions. By setting reasonable SOP and SHP intervals and training the network based on this requirement, we can guarantee any patient using this system that he will not have seizures occur for the upcoming 10 minutes after receiving the alarm. This waiting time is almost enough for the patient to take preventive actions and, on the other hand, after this period of time the patient will wait for a seizure occurrence in a maximum of 20 minutes. In comparing ConvNet with other methods with the same SOP and SPH in Table 2, our results showed superiority. Furthermore, investigating the ConvNet's middle layer feature map revealed some facts about the network performance such as knowing which input features are maintained in the layers and which EEG signal frequency band is more discriminative for preictal and interictal classification.

5. Conclusion

Prognosis of the occurrence time of seizure onset can make a variety of treatments possible. For example, instead of continuous medication, causing neurological complications, treatment can be limited to times when seizure occurrence is probable. For example, patients who are taking persistent antiepileptic drugs can take seizure preventive drugs such as episodic ones. In this paper, a new perspective was proposed to predict epileptic seizures using “the raw multichannel EEG signals” based on generalized neural networks. In this proposed model there is no need for hand-crafted feature extraction. Even, unlike most of the previous works using the short-time Fourier transform (STFT) of the EEG signals as the network input (which is indeed a type of feature extraction) we employed the raw unprocessed EEG signals as input. We have presented a relatively high seizure prediction horizon -around 10 minutes- for our model which is a reasonable time for the patient to take preventive actions. Also, the seizure occurrence period has determined 20 minutes in this work. This makes the patient wait less for a seizure to occur and will have less anxiety. To have a better training procedure, we followed a leave-one-out cross-validation approach for each subject. The results in this task with the LOOCV method showed high sensitivity with a mean of 92.05% and a low FPR of 0.073/h for 18 EEG recordings from the CHB-MIT database. Finally, by investigating feature maps of the trained network and extracting a common frequency range in the middle layer feature maps using an STFT transformation, we were able to characterize which input properties are preserved in the network layers and which frequency band of the EEG signal spectrum contains more information for this classification with these networks. This information is helpful for future investigations since a successfully learned-network, based on a dataset from interictal and preictal EEG time series, has been computationally decoded, and some undercover facts related to its learning and classification performance are available. In future work, we intend to examine the EEG channels separately and find the connectivity between the brain regions.

Reference

[1] Freestone DR, Karoly PJ, Cook MJ. “A forward-looking review of seizure prediction”, Curr Opin Neurol;2017,30:167–173. DOI:10.1097/WCO.0000000000000429.

[2] Brock A, De S, Simonyan K, Smith SL. “High-Performance Large-Scale Image Recognition Without Normalization”, arXiv; 2021,2102.06171. DOI:arxiv.org/abs/2102.06171.

[3] Niknazar H, Nasrabadi AM. “Epileptic Seizure Prediction Using a New Similarity Index for Chaotic Signals”, Int J Bifurcat Chaos; 2016, 26:165-186. DOI:10.1142/S0218127416501868.

[4] Truong ND, Nguyen AD, Kuhlmann M, Bonyadi MR, Yang J, Ippolito S, Kavehei O. “Convolutional Neural Networks for Seizure Prediction Using Intracranial and Scalp Electroencephalogram”, Neural Networks; 2018,105:104-111. DOI:10.1016/j.neunet.2018.04.018.

[5] Chisci L, Mavino A, Perferi G, Sciandrone M, Anile C, Colicchio G, Fuggetta F. “Real-Time Epileptic Seizure Prediction Using AR Models and Support Vector Machines”, IEEE Trans Biomed Eng; 2010, 57: 1124-1132. DOI:10.1109/TBME.2009.2038990

[6] Aarabi A, He B. “Seizure prediction in patients with focal hippocampal epilepsy”, Clin Neurophysiol; 2017, 128:1299-1307. DOI:10.1016/j.clinph.2017.04.026.

[7] Li S, Zhou W, Yuan Q, Liu Y. “Seizure Prediction Using Spike Rate of Intracranial EEG”, IEEE Trans Neural Syst Rehabil Eng; 2013, 21:880–886. DOI:10.1109/TNSRE.2013.2282153

[8] Zhang Z, Parhi K. “Low-Complexity Seizure Prediction From iEEG/sEEG Using Spectral Power and Ratios of Spectral Power”, IEEE Trans Biomed Circuits Syst; 2016, 10:693-706. DOI:10.1109/TBCAS.2015.2477264

[9] Nazari J, Motie Nasrabadi A, Menhaj MB, Raiesdana S. “Epilepsy Seizure Prediction with Few-Shot Learning Method”, Brain Informatics; 2022, 9:21. DOI:10.1186/s40708-022-00170-8

[10] Alotaiby TN, Alshebili SA, Alrshoud SR. “Epileptic Seizure Prediction Using CSP and LDA for Scalp EEG Signals”, Comput Intell Neurosci; 2017, 2017:323-334. DOI:10.1155/2017/1240323

[11] Park Y, Luo L, Parhi K, Netoff T. “Seizure prediction with spectral power of EEG using cost-sensitive support vector machines”, Epilepsia; 2011, 52:1761-1770. DOI:10.1111/j.1528-1167.2011.03138.x

[12] Prasad SC, Prasad P. “Deep recurrent neural networks for time series prediction”, arXiv; 2014, 2:1407.5949. DOI:arxiv.org/abs/1407.5949

[13] Tsiouris KM, Pezoulas VC, Zervakis M, Konitsiotis S, Koutsouris DD, Fotiadis DI. “A Long Short-Term Memory Deep Learning Network for the Prediction of Epileptic Seizures Using EEG Signals”, Comput Biol Med; 2018, 99:24-37. DOI:10.1016/j.compbiomed.2018.05.019

[14] Hosseini MP, Tran TX, Pompili D, Elisevich K, Zadeh HS. “Multimodal Data Analysis of Epileptic EEG and rs-fMRI via Deep Learning and Edge Computing”, Artif Intell Med; 2020, 104:813-823. DOI:10.1016/j.artmed.2020.101813

[15] Rosas RR, Guevara E, Peng K, Nguyen DK, Lesage F, Pouliot P, Lima WE. Prediction of epileptic seizures with convolutional neural networks and functional near-infrared spectroscopy signals. Comput Biol Med. 2019; 111:103355. DOI:10.1016/j.compbiomed.2019.103355

[16] Liang J, Lu R, Zhang C, Wang F. “Predicting Seizures from Electroencephalography Recordings: A Knowledge Transfer Strategy”, IEEE International Conference on Healthcare Informatics, Chicago; 2016. DOI:10.1109/ICHI.2016.27

[17] khan h, marcuse l, Fields M, Swann K, Yener B. “Focal onset seizure prediction using convolutional”, IEEE Trans Biomed Eng; 2018, 65:2109-2118. DOI:10.1109/TBME.2017.278540.

[18] Krizhevesky A, Sutskever I, Honton GE. “Imagenet classification with deep convolutional neural networks”, NIPS; 2012, 60:1097-1105. DOI:10.1145/3065386.

[19] Christian S, et al. “Going deeper with convolutions”, CVPR; 2014,1-9. DOI:arxiv.org/abs/1409.4842

[20] Hasanpour SH, Rouhani M, Fayyaz M, Sabokrou M, Adeli E. “Towards Principled Design of Deep Convolutional Networks: Introducing SimpNet”, arXiv; 2018. DOI:arxiv.org/abs/1802.06205v1

[21] Yann L, Yoshua B and Geoffrey H. “Deep learning”, Nature; 2015, 436-444.DOI:10.1038/nature14539

[22] Srivastava N, Hinton G, Krizhevsky A. “Dropout: A Simple Way to Prevent Neural Networks from Overfitting”, Machin Learning Research; 2014, 15:1929-1958. DOI:mendeley.com/catalogue/e8d99ee4-3f9a-3e60-be30-02827562c697

[23] Bjorck J, Gomes C, Selman B, Weinberger KQ.

"Understanding Batch Normalization", arXiv; 2018, 4:1806.02375. DOI:arxiv.org/abs/1806.02375v4

[24] Shoeb A. "Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment","PhD Thesis, Massachusetts Institute of Technology; 2009. DOI:10.13026/C2K01R

[25] Hisham D, Magdy BA. "Efficient Epileptic Seizure Prediction Based on Deep Learning", IEEE Trans Biomed Circuits Syst; 2019, 13: 804-813. DOI:10.1109/TBCAS.2019.2929053

[26] Zisheng Z and Keshab P. "Low-Complexity Seizure Prediction From iEEG/sEEG Using Spectral Power and Ratios of Spectral Power", IEEE Trans Biomed Circuits Syst; 2016, 693-706. DOI:10.1109/TBCAS.2015.2477264

[27] Cho D, Min B, Kim J, Lee B. "EEG-Based Prediction of Epileptic Seizures Using Phase SynchronizationElicited from Noise-Assisted Multivariate Empirical Mode Decomposition", IEEE Trans Neural Syst Rehabil Eng; 2017, 25:1309 – 1318. DOI:10.1109/TNSRE.2016.2618937

[28] Miri M and Nasrabadi AM. "A new seizure prediction method based on return map", Proceedings of the Iranian Conference on BioMedical Engineering, tehran; 2011. DOI:10.1109/ICBME.2011.6168565

[29] Qin Y, Zheng H, Chen W, Qin Q, Han C. "Patient-specific Seizure Prediction with Scalp EEG Using Convolutional Neural Network and Extreme Learning Machine", 39th Chinese Control Conference (CCC) Shenyang; 2020. DOI:10.23919/CCC50068.2020.9189578.