Electroencephalography (EEG)-based epileptic seizure prediction using entropy and K-nearest neighbor (KNN)

Sutrisno Ibrahim[a*], Ridha Djemal[a], Abdullah Alsuwailem[a], and Sofien Gannouni[b]

[a]Electrical Engineering Department, King Saud University, Saudi Arabia
[b]Computer Sciences Department, King Saud University, Saudi Arabia
P.O. Box 800, Riyadh 11421, Saudi Arabia

Abstract

Epilepsy is known as a brain disorder characterized by recurrent seizures. The development of a system that is able to predict seizure before its coming has several benefits such as allowing early treatment or even preventing the seizure. In this article, we propose a seizure prediction algorithm based on extracting Shannon entropy from electroencephalography (EEG) signals. The K-nearest neighbor (KNN) method is used to continuously monitor the EEG signals by comparing the current sliding window with normal and pre-seizure baselines to predict the upcoming seizure. Both baselines are continuously updated based on the most recent prediction result using distance-based method. Our proposed algorithm is able to predict correctly 42 from 55 seizures (76 %), tested using up to 570 hours EEG taken from the MIT dataset. With its simplicity and fast processing time, the proposed algorithm is suitable to be implemented in embedded system or mobile application that has limited processing resources.

Keywords: seizure prediction, adaptive learning, EEG, Shannon entropy, KNN.

1. Introduction

Epileptic seizures are the clinical manifestation of transient abnormal excessive or synchronous neuronal activity in the brain [1]. Epilepsy is one of the most common neurological disorders that is affecting more than 65 million people worldwide [2]. In many cases, epileptic seizure is very difficult to be predicted because it has complex clinical features and abrupt interference with daily activity. Seizure prediction along with warning system is important to prevent further injuries of the patient. *Electroencephalography* (EEG) which captures signals from human brain has a great potential to be used to analyze brain activity and condition.

EEG has been used in different application such as in *brain computer interface* (BCI) for home devices control [3] and epilepsy diagnosis [4]. Recently, researchers investigated the use of EEG for seizure detection and prediction (see review papers [5,6] for further details). Seizure detector tries to automatically recognize the seizure while it occurs, while seizure predictor tries to predict the seizure before its occurrence. Although several existing methods have achieved promising results for seizure detection, seizure prediction is still an open and challenging problem. F. Mormann et al. [7], used EEG synchronization measurement to differentiate between normal and pre-seizure state. However, their method was only tested using small number of seizures (14 seizure

events). Recently, Wang et al. [8] and Khalid et al. [9] utilized *largest Lyapunov exponent* (LLE) for seizure prediction. It should be noted that calculating LLE is time consuming [9]. To be implemented in embedded system or mobile application, which has limited processing resource, the algorithm should be simple and fast to be executed.

Another important point that should be considered is the adaptability of the method. EEG is a non-stationary signal that varies between different people as well as varies based on their conditions and mental states. Any proposed method based on EEG signal should be able to tackle this condition. Several artificial intelligence or machine learning techniques have been exploited for EEG signal analysis such as *artificial neural network* (ANN) [10], *support vector machines* (SVM) [11], and *k-nearest neighbor* (KNN) [8]. These techniques should not rely only on pre-training process, but also should have an adaptive learning capability to tackle variation in the EEG signals.

In our previous work [12], we presented the effectiveness of using Shannon entropy for assisting epilepsy diagnosis. In this work, we investigate the use of Shannon entropy for seizure prediction in cooperation with *k-nearest neighbor* (KNN) method. Calculating Shannon entropy is extremely faster than calculating LLE, as much as hundred (s) times faster. The remaining parts of this article are structured as follows. Section 2 presents the description of the proposed methods as well as the EEG data used for the experiment. Experimental results and discussion are provided in Section 3.

* Corresponding author. Tel.: +966-11-4679237.
Email: suibrahim@ksu.edu.sa

Finally, the last section provides conclusion as well as discuss the future work.

2. Data and Methods

2.1. EEG Data

In this study, we used a publicly available EEG dataset provided by a research team from Massachusetts Institute of Technology (MIT) and Boston Children's Hospital [11]. All seizure onsets in this dataset are carefully annotated by medical experts. The data is recorded by 23-channels scalp EEG based on 10-20 international standard acquisition system with 256Hz sampling frequency. Reference [12] provides a more detailed description about the dataset. In total, our testing data consists of 570 hours continuous EEG taken from 10 epileptic patients with 55 seizure onsets as summarized in Table 1.

Table 1. Summary of the EEG data

Patient number	Length of EEG	Number of seizures
1	40	7
2	35	3
3	37	7
4	158	4
5	37	5
6	66	10
7	66	3
8	19	5
9	70	4
10	47	7
Total	570 Hours	55

2.2. Seizure prediction method

Researchers found strong evidence for the existence of pre-seizure state prior to seizure onsets that can be used for seizure prediction [5]. Different states of an EEG signal (normal, pre-seizure, seizure onset and post-seizure) are illustrated in Fig. 1. One may predict the upcoming seizure onset within pre-defined prediction horizon (H) by detecting the pre-seizure states. In the literature, researcher used different prediction horizon starting from several minutes up to couple of hours [8]. In this study, we defined the prediction horizon equal to one hour (60 minutes).
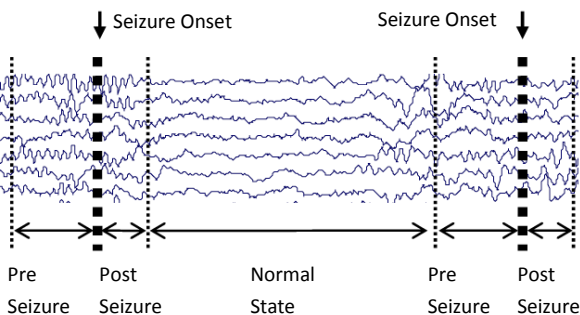

Fig. 1. EEG signals with different seizure events.

Fig. 2 shows an overview of our proposed method. Initially, a sliding window with a length 10 minutes is applied

to a continuous EEG signals. This sliding window is then divided into smaller 60 epochs, each has 10 seconds duration. With sampling frequency equal to 256 Hz, every epoch has 2560 samples for each of its EEG channel.
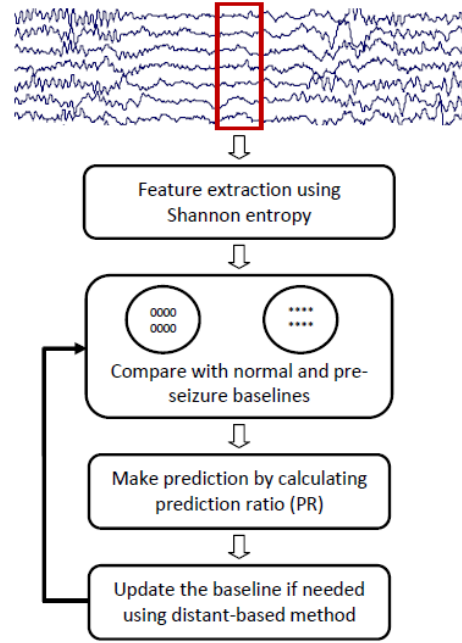

Fig. 2. Overview of the proposed method.

For feature extraction, we employ the Shannon entropy that measures the distribution of the data in the EEG signals. The different distribution and complexity in the brain signal may reveal functionality and condition of brain. These motivate us to do further investigation on using the entropy as a tool to predict the seizure. Let us consider discrete single channel EEG data (S):

$$S = [s_1, s_2, ..., s_{2500}] \quad (1)$$

The Shannon entropy (SE) can be calculated as follows [13]:

$$SE = -\sum_{i=1}^{k} p_i \log_2 p_i \quad (2)$$

where j is the number of distinct values in the discrete data (S) and p_i represents the probability or normalized frequency for these distinct values. We calculate the EEG Shannon entropy from the moving windows as well as from the two pre-defined baselines: normal and pre-seizure baseline. The size of the baseline is the same as the moving window that is 60 EEG epochs.

The basic principle of our seizure predictor is comparing entropy value of all epochs in the moving window with the normal and pre-seizure baselines using the KNN method. Intuitively, this quantifies the degree of similarity of the windowed EEG epoch with two baselines. Let us define WSE_c^n as the entropy value of an epoch in the current moving window with n represents the epoch number and c the channel number. Note that the EEG data has 23 channels with total 60 epochs in the moving window as well as in the two baselines. Accordingly, NSE_c^n and PSE_c^n represent the entropy values in the normal and pre-seizure baseline, respectively. First, we

calculate the distance of the current moving window under consideration with the normal baseline [9].

$$N_c^n = abs(NSE_c^n - WSE_c^n) \qquad (3)$$

with $n = 1, 2, ..., 60$ and $c = 1, 2, …, 23$. In Eq. (3), N_c^n represents the distance of the current moving window with the normal baseline. Taking the summing for all channels, we obtain the distance for each epoch:

$$N^n = \sum_{c=1}^{23} (N_c^n) \qquad (4)$$

Only the first three epochs with minimum distance value are considered for calculating the final distance value:

$$N = \sum_3 \min(N^n) \qquad (5)$$

We may consider this as *k-nearest neighbor* (KNN) with k equals 3. Accordingly, the distance of the moving window with the pre-seizure baseline is obtained with same procedure.

$$P_c^n = abs(PSE_c^n - WSE_c^n)$$
$$n = 1, 2, ..., 60 \ \& \ c = 1, 2, …, 23 \qquad (6)$$

$$P^n = \sum_{c=1}^{23} (P_c^n) \qquad (7)$$

$$P = \sum_3 \min(P^n) \qquad (8)$$

Finally, the prediction ratio (R) is the ratio between the two distances (P and N):

$$R = \frac{P}{N} \qquad (9)$$

The prediction ratio (R) is then compared to a pre-defined threshold value (T). Based on several experiments, we set the threshold equal 0.99. If the prediction ratio falls lower than this threshold ($R<T$), an alarm is generated as a warning that there might be an upcoming seizure within the prediction horizon (H); otherwise it is a normal state.

2.3. Updating the baselines

The user may provide a feedback to the system regarding the prediction output to improve the performance of the system over the time. Four possible outcomes of the seizure prediction are as follows.

- *True Positive* (TP): an alarm is generated and seizure occurs during H
- *True Negative* (TN): no alarm and no seizure occur during H
- *False Positive* (FP): an alarm is generated but no seizure occurs during H
- *False Negative* (FN): no alarm but seizure occurs during H.

If the prediction is true, no update is required, whether it is true positive (TP) or true negative (TN). In case of false positive (FP), the normal baseline is updated using the current

moving window, while in the case of false negative (FN), the pre-seizure baseline is updated. There are several possible methods to update the baseline such as total replacement, random replacement, score-based method and distance-based method [8]. In this study, we use distance-based updating method with 75% replacement. With the baseline size equal to 60, that means 45 'bad' epochs in the baselines are replaced with the new ones from the current moving window. Baseline epoch is considered as 'bad' if it has a big or far distance from the moving window in the case of false prediction. The distance is calculated using the same formulas as previously discussed [see Eqs. (3)-(8)]. However, it is case we use the maximum distance value instead of the minimum one, that mean instead of using minimum function in the Eqs (5) we used maximum one.

3. Results and Discussion

Long term 570 hours scalp EEG recording taken from 10 patients in MIT dataset is used to evaluate the proposed seizure prediction algorithm. The normal and pre-seizure baselines are initialized randomly from the data preceding the first seizure event almost for all patients. Exception for patient number #3, the baselines are initialized from the data preceding the third seizure events because the first and second seizure events happen close with the beginning of the EEG recording. The prediction performance is assessed in terms of sensitivity and specificity as formulated in Eqs. (10)-(11) below. The sensitivity is defined as the number of correctly predicted seizure (TP) over the total seizure events [8]:

$$Sensitivity = \frac{TP}{Total\,seizure} \qquad (10)$$

The specificity is the estimation of the portion of the normal period that is not considered to be false awaiting time:

$$Specificity = \left(1 - \frac{false\,waiting\,period}{normal\,period}\right) \times 100\% \qquad (11)$$

The normal period is duration between post-seizure and the pre-seizure in the upcoming seizure onset (see Fig. 1) and the false waiting period is the number of false alarm multiplied by the prediction horizon (H). In this paper, the post-seizure period is set to be 20 minutes, while the pre-seizure as well as the prediction horizon are set to be 60 minutes (1 hour).

Fig. 3 shows the extracted Shannon entropy values in the first EEG channel for patient number #01. For this patient, the total duration of EEG recording is around 40 hours with seven seizures onsets or events. By observing the pattern, we note that there are always increasing entropy values prior to seizure onset. Fig. 4 shows the prediction ratio (R) for the same patient. Once the prediction ratio falls under the threshold, an alarm is generated indicating that there will be a seizure onset in the near future. All seizures are correctly predicted for this patient. However, there are also several false predictions.

Overall, the proposed seizure prediction method successfully predicts 42 seizures from a total of 55 seizure events (76% sensitivity) for all 10 epileptic patients. Fig. 5 shows in detail the number of predicted seizures and total seizures for each patient. The average specificity is equal to

62%. Table 2 show the comparison with other methods. The proposed method achieved better performance than the previous method by Wang et al. that employed largest Lyapunov exponent (LLE) [8]. Their method achieved 73% sensitivity. It should be noted also that they used intracranial EEG that has a better signal quality compared to scalp EEG. Scalp EEG is more practically to be used in normal life because it is noninvasive.
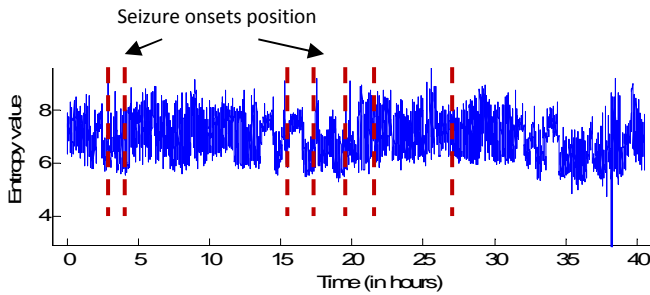

Fig. 3. Entropy values in the first EEG channel for patient #01.
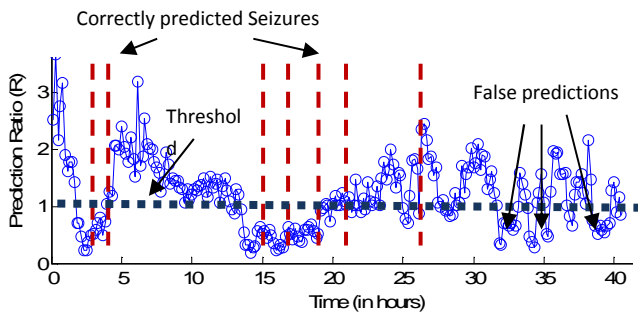

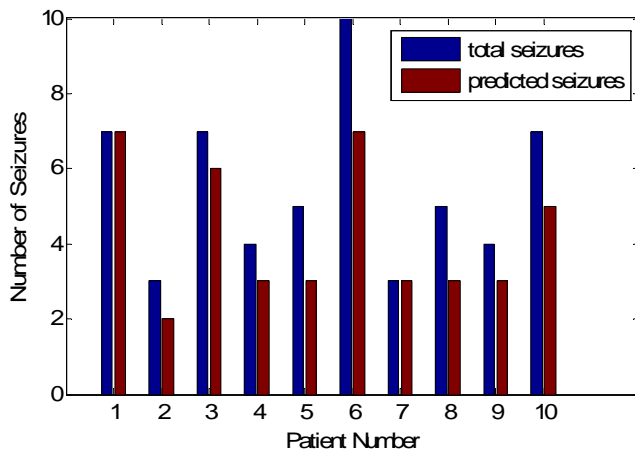Fig. 4. Prediction ratio evolution for patient #01.

Fig. 5. Correctly predicted seizures for all subjects.

Our proposed method based on calculating Shannon entropy is much simpler and faster compared to other methods such as using phase synchronization [7] or largest Lyapunov exponent (LLE) [8,9]. Although the work in reference [9] achieved better performance using LLE (91% sensitivity), this algorithm is time consuming. It takes more than 28 second to extract feature from a 23-channel EEG epoch with duration 10 second (2560 samples), while our method only takes 0.045 second, that means more than 600 times faster. Both

experiments were carried out using the same environment (PC with Intel i7 CPU @2.67GHz, Windows 7, MATLAB R2013a). For the embedded system application which has limited processing resource, our algorithm is more promising to be used.

Table 2. Comparison with pervious methods

Publication	Method	EEG Data	Performance
Wang *et al.* 2013 [8]	Largest Lyapunov exponent (LLE)	Intracranial EEG, 10 subjects, total 2051 hours with 155 seizures.	Sensitivity: 73% Specificity: 67% Very slow
Khalid *et al.* 2015 [9]	Largest Lyapunov exponent (LLE)	MIT scalp EEG, 10 subjects, total 391 hours with 47 seizures	Sensitivity: 91% Specificity: 89% Very slow
This work	Shannon entropy	MIT scalp EEG, 10 subjects, total 570 hours with 55 seizures	Sensitivity: 76% Specificity: 63% Much faster and simpler

Adaptive learning is very important to improve the system performance over the time. In our work (also in [8]), the adaptive learning capability is achieved by updating the baseline with a distance based method; it only replaces the 'bad' samples, while in [9] all samples (whether it is good or bad) are replaced. Always replacing entire samples in the baseline with the most current sliding window is not effective for long-term performance improvement because it would delete also the good baseline samples. Good baseline sample is able to differentiate between normal and pre-seizure states in the EEG signal more clearly.

4. Conclusion and Future Direction

An electroencephalography (EEG)-based epileptic seizure prediction has been presented in this article. The proposed method achieved promising results tested using 570 hours EEG recording taken from Massachusetts Institute of Technology (MIT) scalp EEG dataset. The proposed method successfully predicted 42 out of 55 seizures. With its simplicity and fast processing time, the proposed method can be easily implemented in any embedded system such as in FPGA board or mobile devices.

A possible future work includes testing the proposed method using a larger EEG dataset. More robust feature extraction method will be investigated to improve the prediction accuracy. To tackle the low specificity of the proposed method, several smoothing approaches will be investigated. Furthermore, the proposed seizure prediction method should also be integrated with a suitable application such as mobile devices to perform more effective seizure monitoring.

References

1. N. Tetzlaff and V. Senger, *The seizure prediction problem in epilepsy: Cellular nonlinear networks*, IEEE Circuits Syst. Magaz. 12 (2012) 8-20.

2. D.J. Thurman *et al*; ILAE Commission on, Epidemiology, *Standards for epidemiologic studies and surveillance of epilepsy*, Epilepsia 52 (2011) 2-26.

3. A. Palumbo, F. Amato, B. Calabrese, et al., *An embedded system for eeg acquisition and processing for brain computer interface applications*, in Wearable and Autonomous Biomedical Devices and Systems for Smart Environment. Heidelberg, Germany: Springer Berlin Heidelberg, 2010.

4. S. Noachtar, J. Rémi, *The role of EEG in epilepsy: A critical review*, Epilepsy & Behavior 15 (2009) 22-33.

5. S. Ramgopal *et al*, *Seizure detection, seizure prediction, and closed-loop warning systems in epilepsy*, Epilepsy & Behavior 37 (2014) 291–307.

6. T. N. Alotaiby et al, *EEG seizure detection and prediction algorithms: a survey*, EURASIP J. Adv. Sign. Proc. 2014 (2014) 183.

7. F. Mormann, R. Andrzejak, T. Kreuz, C. Rieke, P. David, C. Elger, and K. Lehnertz, *Automated detection of a preseizure state based on a decrease in synchronization in intracranial electroencephalogram recordings from epilepsy patients*, Physical Rev. E 67 (2003) 021912.

8. S. Wang, W. Chaovalitwongse, and S. Wong, *Online seizure prediction using adaptive learning approach*, IEEE Trans. Knowl. Data Eng. 25 (2013) 2854-2866.

9. M. I. Khalid, S. A. Aldosari, S. A. Alshebeili, T. Alotaiby, and F. E Abd El-Samie, *Online adaptive seizure prediction algorithm for scalp EEG*, *2015 International Conference on Information and Communication Technology Research (ICTRC), Abu Dhabi, UAE, 2015, pp. 44-47.*

10. A. J. Gabor, R. R. Leach and F. U. Dowla, *Automated seizure detection using a self-organizing neural network.*, Electroencephalogr Clin Neurophysiol. 99 (1996) 257-66.

11. T. N. Alotaiby, F. .E. Abd El-Samie, S. A. Alshebeili, K. H. Aljibreen and E. Alkhanen, *Seizure detection with common spatial pattern and support vector machines, 2015 International Conference on Information and Communication Technology Research (ICTRC), Abu Dhabi, UAE, 2015, pp. 152–155.*

12. S. Ibrahim, K. AlSharabi, R. Djemal, and A. Alsuwailem, *An adaptive learning approach for EEG-based computer aided diagnosis of epilepsy*, *International Seminar on Intelligent Technology and Its Application (ISITIA), Lombok, Indonesia, 2016, pp. 55-60.*

13. A. Goldberger, et al., *PhysioBank, physioToolkit, and physioNet: components of a new research resource for complex physiologic signals circulation*, 101 (23):e215-e220. Link to database: http://physionet.org/pn6/chbmit/ [last visited on Sept. 5, 2016].

14. A. Shoeb, *Application of machine learning to epileptic, seizure onset detection and treatment*, Ph.D. Thesis, MIT, USA, 2009.

15. C. E. Shannon, *A mathematical theory of communication*, Bell Syst. Technical J. 27 (1948) 379–423.