Feature Extraction with Stacked Autoencoders for Epileptic Seizure Detection

Akara Supratak[1], Ling Li[2] and Yike Guo[1*]

***Abstract*— Scalp electroencephalogram (EEG), a recording of the brain's electrical activity, has been used to diagnose and detect epileptic seizures for a long time. However, most researchers have implemented seizure detectors by manually hand-engineering features from observed EEG data, and used them in seizure detection, which might not scale well to new patterns of seizures. In this paper, we investigate the possibility of utilising unsupervised feature learning, the recent development of deep learning, to automatically learn features from raw, unlabelled EEG data that are representative enough to be used in seizure detection. We develop patient-specific seizure detectors by using stacked autoencoders and logistic classifiers. A two-step training consisting of the greedy layer-wise and the global fine-tuning was used to train our detectors. The evaluation was performed by using labelled dataset from the CHB-MIT database, and the results showed that *all* of the test seizures were detected with a mean latency of 3.36 seconds, and a low false detection rate.**

***Index Terms*— epileptic seizures, scalp electroencephalogram, deep learning, unsupervised feature learning, stacked autoencoders**

I. INTRODUCTION

About 50 million people worldwide have epilepsy [1], which is a disorder of the brain characterised by recurrent epileptic seizures. They suffer from repeated, unpredictable seizures, whose effects can vary from disturbances of movement and brief loss of consciousness. These seizures are brief episodes of sign and/or symptoms due to abnormal excessive or synchronous neural activity in the brain [2]. One of the most common approaches to detect and guide therapy for these seizures is through analysis of the *scalp electroencephalogram (EEG)*, which is the recording of the brain's electrical activity.

In order to achieve this, most researchers have developed tools to analyse and detect seizures from patients' brain data by manually hand-engineering features that are representative enough to be used to train supervised learning algorithms [3], [4], [5], [6], [7]. This feature engineering is a way to utilise ingenuity and expert knowledge of human being to create good features for machine learning algorithms [8]. In [5], they utilised Fourier transformation to build patient-specific seizure detectors. They extracted energy of each EEG epoch from selected EEG channels to construct a spectral and spatial feature vector. This vector can be used to detect the presence or absence of some spectral components occurred due to seizures. In [7], they also utilised wavelet decomposition and statistical measurement to extract a useful feature called a combined seizure index (CSI). This CSI is computed from rhythmicity and relative energy of the EEG in the desired EEG channel, and it will be increased significantly during seizure periods. This CSI is also sensitive to the consistency among different EEG channels, which can help reduce false seizure alarms as well as improve the detection performance.

These hand-engineering techniques, however, might not scale well to new patterns of seizure activity, which might be observed in the future. This is due to the fact that EEG data is non-stationary, and the seizure patterns vary across different patients [5]. Thus, hand-engineering new features each time new seizure patterns are observed might be labour-intensive and time-consuming. It would be better if seizure detectors are less dependent on human, and are able to learn features from the observed data by themselves.

One of the promising techniques capable of automatically learn good feature representations from unlabelled data is *Unsupervised Feature Learning* [8], [9], which is the recent development of deep learning. By using multi-layer (or deep) neural networks in combination with special training schemes, meaningful feature representations can be extracted from unlabelled data. Recently, there are some researchers who applied one of the unsupervised feature learning techniques, called deep belief nets (DBNs), to perform multi-class classification and anomaly detection on EEG data [10]. The performance of their system were competitive with the recent hand-engineering approaches.

In this paper, we investigate the possibility of applying *stacked autoencoders* [11], which was trained by using the greedy layer-wise [12] and the global-fine-tuning to learn feature representations from raw, unlabelled EEG data that are informative enough to be used in seizure detection. As the characteristics of EEG vary significantly across different patients [5], implementing seizure detectors for each patient can improve the performance of seizure detection. Therefore we develop *patient-specific* seizure detectors that utilise 1) stacked autoencoders to extract features from the EEG data; and 2) logistic classifiers to perform seizure classification based on the learned features.

II. METHOD

The characteristic of the scalp EEG data used in this study is first introduced in Sect. II-A. In Sect. II-B, two main components used in our seizure detectors are briefly

*Corresponding author

[1]A. Supratak and Y. Guo are with Department of Computing, Imperial College London, London SW7 2AZ, UK `{as12212,y.guo} at imperial.ac.uk`

[2]L. Li is with School of Computing, University of Kent, Kent CT2 7NF, UK `c.li at kent.ac.uk`

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

described. Then the details of how we used these two components to build patient-specific seizure detectors is explained in Sect. II-C.

A. Scalp Electroencephalogram

Scalp electroencephalogram (EEG) is the recording of the electrical activity of the brain measured by using electrodes on the scalp. The voltage measured between two electrodes forms an EEG channel.

In this study, an EEG data in the CHB-MIT database [13] was used. This database contained EEG recordings from 23 cases collected from 22 patients who had intractable seizures, and there were 173 events were judged to be clinical seizures by experts. The EEG signals were sampled at 256 Hz, and stored in one- or two-hour long record files. Most of these files contained 23 EEG signals. This EEG data was recorded by positioning scalp electrodes according to the international 10-20 system [14]. The terms *seizure records* and *non-seizure records* are used to refer to record files that contain one or more seizures, and record files that do not contain any seizures respectively.

At the beginning of most seizures, a set of EEG channels start to develop different rhythmic activities, which typically contains a number of frequency components. The patterns of the rhythmic activity also appear differently in different EEG channels across patients. Therefore, in this study, only EEG data from the representative channels were used to train and test the detector for each patient. In particular, Fig. 1 and Fig. 2 show seizures from different patients, and the vertical red dash-lines indicate the onset of Patient A's and B's seizures. In Fig. 1, Patient A's seizure begins at 3801 seconds, and Patient B's seizure begins at 2996 seconds in Fig. 2. From these two figures, it can be seen that the appearances of seizure activities of Patient A and B are most prominent on the different channels. For instance, the representative channels of Patient A are T7-P7(3) and P7-T7(19), while the representative channels of Patient B are FP2-F8(13) and F8-T8(14). Thus, in this example, only EEG data from T7-P7(3) and P7-T7(19) channels is used to train and test the detector for Patient A, and only EEG data from FP2-F8(13) and F8-T8(14) channels is used to train and test the detector for Patient B.

B. Stacked Autoencoders and Logistic Classifiers

1) Stacked Autoencoders: A stacked autoencoder (SAE) [11] is a neural network consisting of multiple layers of sparse autoencoders in which the outputs of each layer is forwarded to the input of each subsequent layer. A sparse autoencoder is a neural network consisting of only one hidden layer. It is an unsupervised learning algorithm capable of extract good feature representations from unlabelled data. By setting the target value of the autoencoder to be equal to the input, the autoencoder tries to learn features that can be used to reconstruct the input.

The important property of the sparse autoencoder is that given a large amount of unlabelled data, the autoencoder can learn good feature representations of the input. Thus,

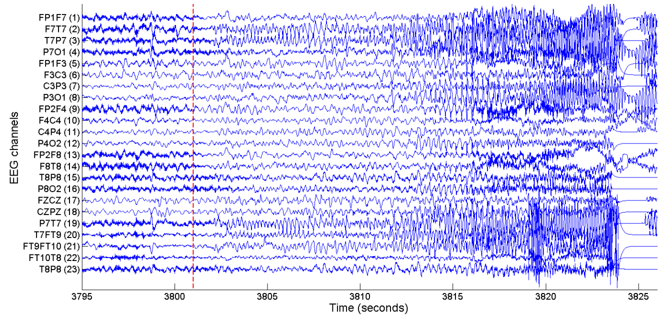

Fig. 1. An example of a seizure pattern in the scalp EEG data of Patient A.

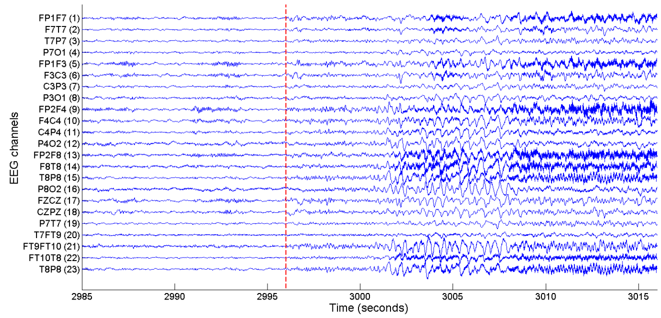

Fig. 2. An example of a seizure pattern in the scalp EEG data of Patient B.

the multi-layer (or stack) of the autoencoders can be used to learn useful feature representations from EEG data, as the subsequent layers can utilise the features learned from the previous layers to produce even more useful features. These feature representations can be then applied in seizure detection.

2) Logistic Classifiers: In each seizure detector, a binary-class logistic classifier [15] was stacked at the top of the SAE. This enabled the classifier to utilise the feature representations learned from the SAE. Using these learned features in combination with their associated labels from the EEG dataset, the logistic classifier can learn to determine which part of EEG data contains seizure activity.

C. Patient-Specific Seizure Detectors

The seizure detector for each patient was built by using SAEs and logistic classifiers consisting of one input layer, two hidden layers, and one output layer (256-500-500-1). The two hidden layers (i.e., SAEs) were used to learn good feature representations from raw, unlabelled EEG data, which were then used by the output layer (i.e., logistic classifier) to detect the onset of the seizures. The EEG data of each channel was segmented into 1-second epochs before being used in training and seizure detection processes. Due to the fact that the patterns of seizure activity vary across different patients, thus only representative channels were used. The representative channels of different patients might not be the same depending on the appearance of the prominent seizure patterns as discussed in Sect. II-A.

1) Training: The training process was done in two main steps: *pretraining* and *fine-tuning*. In particular, the SAEs with the logistic classifier was first pretrained by using a *greedy layer-wise* approach [12] over raw, unlabelled EEG data to obtain initial model parameters. These parameters were then used as a starting point for the second training step. The second step was to perform a supervised global *fine-tuning* to further improve the performance of the detectors by using labels associated with EEG data. We found that training with only the global fine-tuning step offered poor performance than the two-step training. This is because the pretraining step helps avoid getting stuck in local optima, which can typically occur with random initialisation.

In this study, the *mini-batch* gradient descent, which considers 10 training examples in each step of gradient descent to update model parameters, was used to perform this two-step training. The number of epochs used in the first and the second steps were 10 and 20 respectively. The training set was randomly shuffled at the beginning of each epoch. The regularisation was also used to prevent overfitting problem as well as to prevent hidden layers to learn the identity function (i.e., one-to-one mapping). The learning rate and the regularisation weight were set to 0.05 and 0.003 respectively.

2) Seizure Declaration: Given a set of 1-second epochs of EEG data from the representative channels, our detectors determine whether each 1-second epoch of EEG data in each channel has seizure activity.

Formally, given the i-th 1-second epochs of EEG data $\{x^{(i,1)}, \ldots, x^{(i,k)}\}$ from k channels, where $x^{(i,j)} \in \mathbb{R}^{256}$; and $x^{(i,j)}$ is the i-th 1-second epoch of the j-th channel, our detectors produce an output set $\{u^{(i,1)}, \ldots, u^{(i,k)}\}$ for k channels, where $u^{(i,j)} \in \{0,1\}$; and $u^{(i,j)}$ is the i-th output of the j-th channel. $u^{(i,j)}$ is equal to 1 if the probability that the input $x^{(i,j)}$ having seizure activity is greater than the threshold; otherwise, $u^{(i,j)}$ is equal to 0. In this study, the threshold was set to 0.5.

Our seizure detectors declare each 1-second epoch of EEG data from k representative channels as a seizure when the number of $u^{(i,j)} = 1$ is greater than or equal to a specified threshold (or a *channel threshold*). Different values of channel thresholds affect the performance of our seizure detectors which will be discussed in Sect. III-C. The term *seizure channel* is used to represent the number of representative channels that our seizure detectors determine as having seizure activities.

III. Experimental Evaluation

A. Scalp EEG Dataset

In our experiment, six of 23 cases were randomly chosen to evaluate the performance of our detectors, and only seizure records were used as it is sufficient to demonstrate the performance of our detectors. Thus, our dataset included 44 hours of EEG data with 39 seizures events. Before the EEG data of each channel was used to train and test our detectors, it was divided into 1-second epochs, and pre-processed with the mean normalisation and feature scaling in order to speed up the gradient descent.

B. Performance Measurement

The performance of our detector was evaluated by using sensitivity, specificity and latency. *Sensitivity* refers to the percentage of seizures in the test set being identified. *Specificity* refers to the number of times per hour our detector declared the onset of seizures in non-seizure periods. *Latency* refers to the delay between the onset of seizures judged by experts and the onset of seizures declared by our detectors.

To estimate our detector's performance on data from each patient, a *leave-one-record-out* cross-validation scheme [5] was used. Suppose there were N_S seizure records. The evaluation scheme uses $N_S - 1$ seizure records to train the detector. The detector was then used to detect the seizures in the seizure record that was not used in training. This process is repeated N_S times so that each seizure record was tested. For each round, the number of detected seizures and the detection latency was recorded.

C. Results

Fig. 3 shows an example of the scalp EEG data of Patient A (top), and the number of representative channels (bottom) that our seizure detectors determine as having seizure activities (or seizure channels). In this example, the representative channels were the channel 2,3,19,20 and 21. If the channel threshold is set to 1, our detector was able to the detect this seizure with 2-second delay (i.e., at 3803 second). If the channel threshold is set to 2, the onset seizure was detected with 3-second delay (i.e., at 3804).

In our experiment, three channel thresholds: 1, 3 and 5, were used to evaluate the performance of our seizure detectors. Table I shows the summary of the sensitivity and latency evaluated by using these three channel thresholds. The results demonstrated that our seizure detectors were able to detect *all* of 39 test seizures with a mean latency of 3.36 seconds when the channel threshold was set to 1. When the threshold was increased, our detectors started to miss some seizures and became slower in declaring the seizures' onset as most of the prominent seizure activities in the scalp EEG data start to appear after the seizures have begun for a while (see Fig. 3).

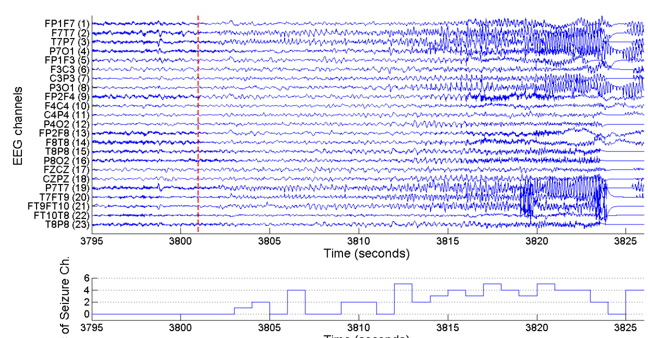

Fig. 3. An example of the scalp EEG data of Patient A (top) associated with the number of seizure channels (bottom) determined by our seizure detector.

TABLE I
SENSITIVITY AND LATENCY OF OUR SEIZURE DETECTORS.

Channel Threshold	Sensitivity (%)	Mean Latency (sec)
1	100%	3.36
3	100%	6.87
5	87.18%	11.18

Fig. 4 shows the number of false detections per hour (i.e., specificity) for each of the six cases with three different channel thresholds. Our detectors produced the lowest false detection rate when the channel threshold was set to 5. This was, as expected, in contrast to the case of sensitivity and latency. When the channel threshold was set to a higher value, our detectors would declared each 1-second epoch of EEG data as having seizure activity only when they were very confident. Even though, the higher threshold made our detectors less sensitive to the noise in the EEG data that has a similar pattern to the seizure activity, some of the seizures, however, might not be detected, or detected with a high latency. Therefore, the channel threshold can be used to specify *tradeoff* between sensitivity, specificity and latency. It is worth mentioning that the false detection rates of Patient 4 were surprisingly high compared with the other five cases. This might be because the non-seizure and seizure activity of Patient 4 were very similar.

Some of the falsely declarations were also observed during the seizures. In particular, our detectors falsely determined some 1-second epochs of EEG data as not having seizure activity as illustrated in Fig. 3. This is because the rhythmic activity during seizures is unstable, and termination of each seizure is less obvious than the onset [2]. Thus, our detectors might misclassify some rhythmic activities during seizures as non-seizure activities. However, this does not affect the performance of our seizure detectors as we are interested only the declaration of the seizures' onset.

The seizure detectors were implemented using MATLAB and evaluated on a common computer (CPU 3.40 GHz and RAM 16.0 GB). The training time varied depending on the amount of training examples, which ranged from 2 to 5 hours. The run time of the detectors was $\sim$10 ms for each 1-second data, which can be used in online seizure detection.

According to the results obtained from our experiment, we believe that the SAE trained with the two-step training is capable of learning good features from raw, unlabelled EEG data that are representative enough to be used in seizure detection. The best choice of the channel threshold depends on the seizure patterns in each patient.

IV. CONCLUSIONS

We investigate the possibility of applying the recent development of unsupervised feature learning to extract meaningful features from raw, unlabelled EEG data such that they can be used in seizure detection. We utilise the SAEs and logistic classifiers to build patient-specific seizure detectors. The results of our experiment confirmed that our detectors were able to detect seizures by using features learned from the EEG data.

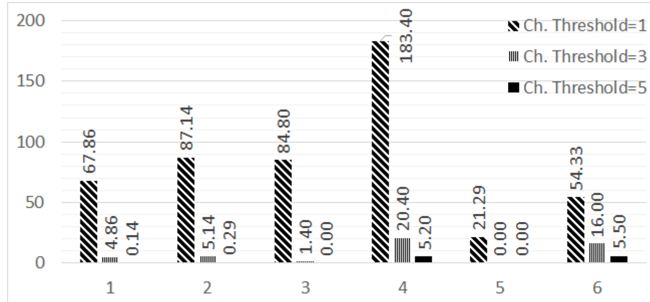

Fig. 4. False detection rates per hour for each of the six cases with three different channel thresholds.

REFERENCES

[1] W. H. Organization, "Epilepsy." [Online]. Available: http://www.who.int/mediacentre/factsheets/fs999/en/

[2] R. S. Fisher, W. van Emde Boas, W. Blume, C. Elger, P. Genton, P. Lee, and J. Engel, J., "Epileptic seizures and epilepsy: Definitions proposed by the international league against epilepsy (ilae) and the international bureau for epilepsy (ibe)," *Epilepsia*, vol. 46, no. 4, pp. 470–472, 2005.

[3] S. A. Asha, C. Sudalaimani, P. Devanand, T. E. Thomas, and S. Sudhamony, "Automated seizure detection from multichannel eeg signals using support vector machine and artificial neural networks," in *Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), 2013 International Multi-Conference on*, Conference Proceedings, pp. 558–563.

[4] Y. U. Khan and J. Gotman, "Wavelet based automatic seizure detection in intracerebral electroencephalogram," *Clinical Neurophysiology*, vol. 114, no. 5, pp. 898–908, 2003.

[5] A. H. Shoeb and J. V. Guttag, "Application of machine learning to epileptic seizure detection," in *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, Conference Proceedings, pp. 975–982.

[6] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Epileptic seizure detection in eegs using time-frequency analysis," *IEEE Trans Inf Technol Biomed*, vol. 13, no. 5, pp. 703–10, 2009.

[7] A. S. Zandi, M. Javidan, G. A. Dumont, and R. Tafreshi, "Automated real-time epileptic seizure detection in scalp eeg recordings using an algorithm based on wavelet packet transform," *Biomedical Engineering, IEEE Transactions on*, vol. 57, no. 7, pp. 1639–1651, 2010.

[8] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, no. 8, pp. 1798–1828, 2013.

[9] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[10] D. F. Wulsin, J. R. Gupta, R. Mani, J. A. Blanco, and B. Litt, "Modeling electroencephalography waveforms with semi-supervised deep belief nets: fast classification and anomaly measurement," *J Neural Eng*, vol. 8, no. 3, p. 036015, 2011.

[11] Y. Bengio, "Learning deep architectures for ai," *Foundations and Trends in Machine Learning*, vol. 2, no. 1, pp. 1–127, 2009.

[12] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy layer-wise training of deep networks," *Advances in neural information processing systems*, vol. 19, pp. 153–160, 2007.

[13] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[14] R. W. Homan, J. Herman, and P. Purdy, "Cerebral location of international 1020 system electrode placement," *Electroencephalography and clinical neurophysiology*, vol. 66, no. 4, pp. 376–382, 1987.

[15] S. Dreiseitl and L. Ohno-Machado, "Logistic regression and artificial neural network classification models: a methodology review," *Journal of Biomedical Informatics*, vol. 35, no. 56, pp. 352–359, 2002.