International Journal of Neural Systems, Vol. 30, No. 4 (2020) 1950024 (15 pages)

DOI: 10.1142/S0129065719500242

Automatic Seizure Detection Based on S-Transform and Deep Convolutional Neural Network

Guoyang Liu, Weidong Zhou* and Minxing Geng
School of Microelectronics, Shandong University
Jinan 250100, P. R. China

Shenzhen Institute of Shandong University
Shenzhen 518057, P. R. China
**wdzhou@sdu.edu.cn*

Received 13 August 2019
Accepted 13 August 2019
Published Online 30 September 2019

Automatic seizure detection is significant for the diagnosis of epilepsy and reducing the massive workload of reviewing continuous EEGs. In this work, a novel approach, combining Stockwell transform (S-transform) with deep Convolutional Neural Networks (CNN), is proposed to detect seizure onsets in long-term intracranial EEG recordings. Primarily, raw EEG data is filtered with wavelet decomposition. Then, S-transform is used to obtain a proper time-frequency representation of each EEG segment. After that, a 15-layer deep CNN using dropout and batch normalization serves as a robust feature extractor and classifier. Finally, smoothing and collar technique are applied to the outputs of CNN to improve the detection accuracy and reduce the false detection rate (FDR). The segment-based and event-based evaluation assessments and receiver operating characteristic (ROC) curves are employed for the performance evaluation on a public EEG database containing 21 patients. A segment-based sensitivity of 97.01% and a specificity of 98.12% are yielded. For the event-based assessment, this method achieves a sensitivity of 95.45% with an FDR of 0.36/h.

Keywords: S-transform; deep learning; Convolutional neural networks (CNN); time-frequency representation; Seizure detection.

1. Introduction

As a chronic neurological disorder, epilepsy is characterized by the recurrent seizures due to abnormal, excessive or hypersynchronous neuronal activities in the brain.[1,2] About 0.6–0.8% of the world's population suffer from epileptic seizures,[3] and for these people, the disease may not only make them lose awareness or consciousness, but also disturb their movement, mood, sensation or mental function.[4] EEG as an effective tool to record activities of the brain has been widely used for the analysis and diagnosis of brain diseases such as epileptic seizures,[5] Alzheimer's disease,[6–8] Attention Deficit Hyperactivity Disorder (ADHD),[9–11] etc. At present, long-term EEG recordings are still inspected by well-trained experts in clinical analysis, which is a time-consuming task. Therefore, an automatic seizure detection system with higher sensitivity and lower false detection rate (FDR) is highly in demand for medical experts in clinical EEG recording analysis.

A variety of seizure detection techniques have been proposed in the past years. In the early 1980s, Gotman *et al.* proposed one of the earliest automatic seizure detection systems,[12] which decomposed the EEG recordings into half-waves and extracted average amplitude, average duration,

*Corresponding author.

background activity, rhythmicity as features for classification. Later, many features based on time, frequency and time-frequency domain were proposed, such as spike rate,[13] mean phase coherence,[14] power spectral density ratio[15,16] and features, from Walsh transformation.[17,18] Meanwhile, a variety of nonlinear dynamics theories including fractal theory.[19,20] and wavelet-chaos methodology.[21,22] have been proposed to analyze the EEG signals. In addition, a classifier with robustness and generalization is needed to classify the features correctly and effectively. Currently, classifiers for classifying EEG features mainly include Support Vector Machine (SVM),[23–25] Bayesian Linear Discriminant Analysis (BLDA)[26] and Artificial Neural Networks (ANN),[27–29] etc.

In previous studies, Short-Time Fourier Transform (STFT) and Continuous Wavelet Transform (CWT) have been widely used in time-frequency analysis of EEG signals.[30,31] S-transform is another effective time-frequency analysis approach proposed by Stockwell *et al.*[32] It is considered as a hybrid transformation between the STFT and the CWT. Unlike the fixed window of STFT, S-transform has an adaptive window like CWT. The S-transform has an advantage in that it provides a wider window at lower frequencies and a narrower window at higher frequencies while retaining the absolute phase of each frequency component of the signal. Due to these advantages of S-transform, it is applied to analyze many kinds of signals such as EEG signals[33–35] and medical imaging.[36,37]

With development of machine learning theory and computing power, deep neural networks have become very effective and powerful tool for classification. In the field of image recognition and classification, Convolutional Neural Network (CNN) is a state-of-the-art method.[38] In recent years, CNN have also been applied to automatic seizure detection. Mirowski *et al.*[39] proposed to use a CNN with five layers to classify a manually selected 2D EEG feature pattern. The method presented by Truong *et al.*[30] inputted the STFT of EEG data into a CNN with 10 layers, and achieved a better performance. On the other hand, due to the strong ability of the CNN to extract features automatically, CNN also acted as an end-to-end classifier in seizure detection.[40,42,43]

In this work, we proposed an automatic seizure detection system based on S-transform spectrogram and deep CNN. The remainder of this paper is organized as follows: Sec. 2 introduces the EEG database used in this study. Section 3 provides a specific description of this method. The proposed method is evaluated based on different criteria and the results are given in Sec. 4. Finally, Sec. 5 discusses the performance of the proposed system and the impact of parameters on system performance.

2. EEG Database

The long-term intracranial EEG (iEEG) database used in this study comes from the Epilepsy Center of the University Hospital of Freiburg, Germany.[44] The database contains EEG recordings of 21 patients suffering from medically intractable focal epilepsy. The EEG data were acquired with intracranial grid-, strip-, and depth-electrodes, using a Neurofile NT digital video system with 128 channels, 256 Hz sampling rate, and a 16-bit AD converter. There are about 725 h of EEG data in total. For each patient, the recordings of three focal and three extra-focal electrode contacts are available.

In this study, only one seizure event is randomly chosen as the training data for each patient. Because empirically the more data is provided, the better performance the deep CNN is achieved, the length of selected normal EEG signals is five times that of the seizure EEG signals. Meanwhile, in order to maintain the balance of training data, each seizure event is overlapped up-sampled with five times. In summary, about 726 h iEEG recordings containing 87 seizure events in all six channels are utilized. 5.5 h recordings containing 21 seizure events are employed as the training set and the rest of 720 h data containing 66 seizure events serve as the test set. The details of the EEG database are summarized in Table 1.

3. Methods

The proposed automatic seizure detection system can be approximately divided into five parts: preprocessing, segmentation, S-transform, deep convolutional neural network and postprocessing. Figure 1 illustrates the architecture of the proposed system and the details of each part are described in the following section.

Table 1. Details of the used database. Seizure types and location: SP represents simple partial, CP represents complex partial, GTC represents generalized tonic-clonic. Electrodes: g represents grid electrode, s represents strip electrode, d represents depth electrode.

Patient	Sex	Seizure type	Electrodes	Number of training seizures	Number of of testing seizures	Training seizure duration (min)	Training interictal duration (min)	Testing EEG length (h)
1	F	SP,CP	g,s	1	3	0.32	1.63	32.93
2	M	SP, CP, GTC	d	1	2	2.46	12.28	29.34
3	M	SP,CP	g,s	1	4	1.98	9.90	31.95
4	F	SP, CP, GTC	d,g,s	1	4	1.55	7.77	33.81
5	F	SP, CP, GTC	g,s	1	4	2.53	12.65	33.61
6	F	CP,GTC	d,g,s	1	2	1.80	9.00	30.20
7	F	SP, CP, GTC	d	1	2	5.32	26.59	30.04
8	F	SP,CP	g,s	1	1	2.98	14.91	27.41
9	M	CP,GTC	g,s	1	4	4.46	22.30	33.45
10	M	SP, CP, GTC	d	1	4	2.55	12.73	34.87
11	F	SP, CP, GTC	g,s	1	3	5.93	29.65	31.42
12	F	SP, CP, GTC	d,g,s	1	3	1.02	5.11	57.49
13	F	SP, CP, GTC	d,s	1	1	4.01	20.04	27.57
14	F	CP,GTC	d,s	1	3	5.99	29.94	30.23
15	M	SP, CP, GTC	d,s	1	3	2.42	12.11	33.72
16	F	SP, CP, GTC	d,s	1	4	2.84	14.22	35.36
17	M	SP, CP, GTC	s	1	4	2.06	10.28	38.62
18	F	SP,CP	s	1	4	0.30	1.50	37.78
19	F	SP, CP, GTC	s	1	3	0.39	1.93	36.77
20	M	SP, CP, GTC	d,s	1	4	2.99	14.93	38.18
21	M	SP,CP	s	1	4	1.14	5.68	35.78
Total				21	66	55.02	275.12	720.53

3.1. *Preprocessing and segmentation*

The raw EEG signals with six channels are first divided into 4-s epochs. Then, discrete wavelet transform (DWT) with Daubechies-4 (Db4) wavelet is applied to each segment to suppress the artifacts and noises arising from eye movement and power line interference, etc. It has been proved in many studies that the Db4 wavelet matches the shape and frequency characteristics of seizures.[45,46] According to Nyquist criteria, the EEG signal with sampling rate of 256 Hz is band limited to 128 Hz. Hence, for each segment, DWT yields six scales representing 64–128 Hz (d1), 32–64 Hz (d2), 16–32 Hz (d3), 8–16 Hz (d4), 4–8 Hz (d5) and approximation coefficients representing 0-4 Hz (a5). Since the frequency of seizure activity is most often ranging from 3–29 Hz,[12] scales 3, 4 and 5 are reconstructed as filtered data. The procedure of filtering raw EEG data using DWT is presented in Fig. 2. An additional preprocessing is segmentation. After filter with wavelet transform, the 4-s epochs are further divided into four 1-s segments without overlapping to capture more sophisticated EEG information in time.

3.2. *S-Transform spectrogram*

S-transform is a combination of STFT and CWT, and is an effective time-frequency analysis method for non-stationary signals. To better control the resolution of the S-transform, the S-transform fine-tuned with parameter p is given as follows:

$$S(\tau, f) = \frac{|f|}{p\sqrt{2\pi}} \int_{-\infty}^{+\infty} x(t) e^{-\frac{(t-\tau)^2 f^2}{2p^2}} e^{-2i\pi f t} dt. \quad (1)$$

As S-transform is directly related to the Fourier transform as STFT, it can be calculated by Fast Fourier Transform (FFT) and Inverse Fast Fourier Transform (IFFT), which can provide a low computation complexity of S-transform. On the other hand, in the view of CWT, the S-transform can be rewritten as follows:

$$S(\tau, f) = e^{-2i\pi f\tau} \sqrt{|f|} W(\tau, a), \quad (2)$$

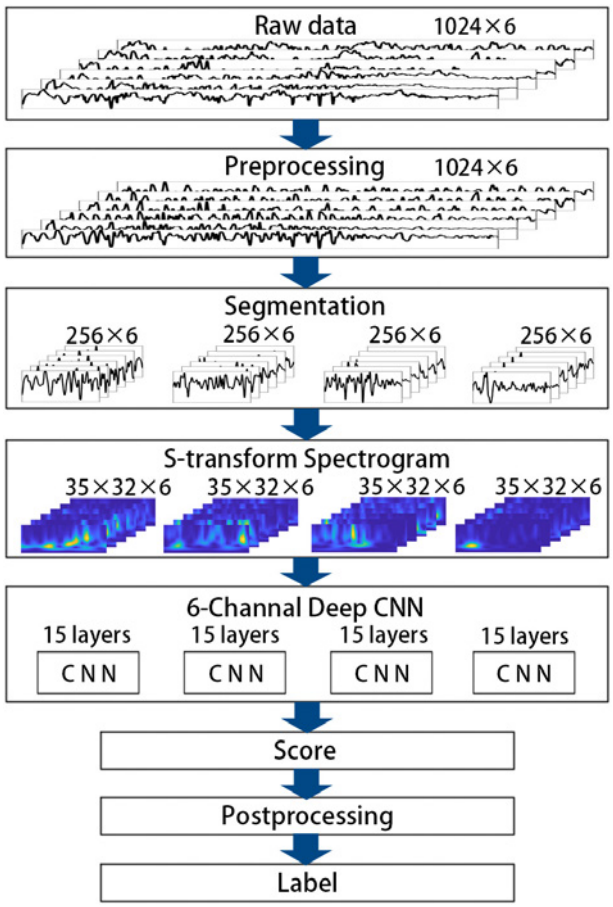

Fig. 1. The architecture of the proposed seizure detection system. Main steps include preprocessing, segmentation, S-transform, deep CNN and postprocessing.

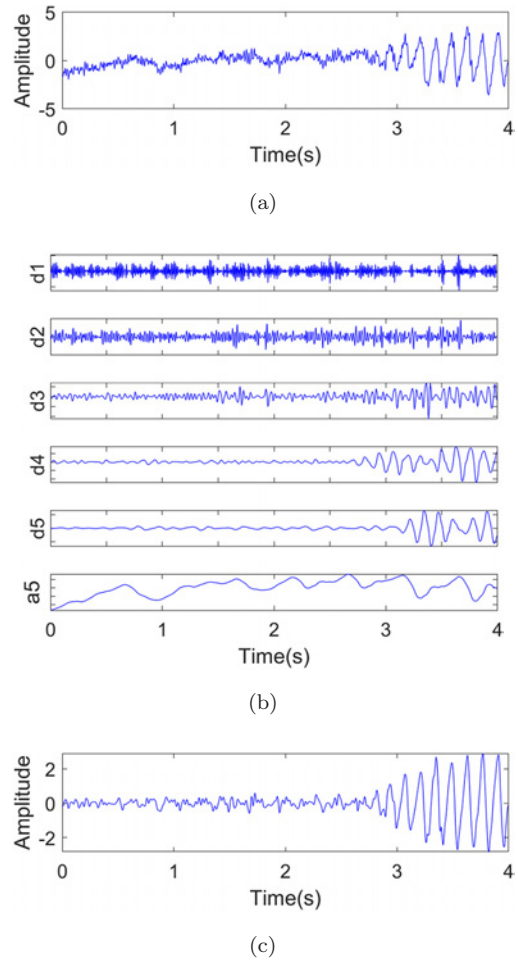

Fig. 2. Filtering EEG signal with DWT. (a) The original 4 s EEG signal, (b) The EEG signal decomposed into five scales using DWT, (c) The filtered EEG signal consisting of d3, d4 and d5 components.

where a represents the scale inversely proportional to frequency, and $W(\tau, a)$ is the CWT of the signal $x(t)$ with a special complex Morlet wavelet satisfying the following equation:

$$\Phi(t) = \frac{1}{p\sqrt{2\pi}} e^{-\frac{t^2}{2p^2}} e^{-2i\pi t}. \tag{3}$$

To obtain the energy distribution of the signal in time-frequency domain, the squared modulus of the S-Transform is taken into account. The S-transform spectrogram is given as:

$$|S(\tau, f)|^2 = S(\tau, f)S^*(\tau, f). \tag{4}$$

The gaussian window of S-transform could provide a higher time resolution for high frequencies and a higher frequency resolution for low frequencies. The seizure EEG signals contain high-frequency components, such as spike waves and sharp waves, while the frequency of normal EEG signals is relatively low, therefore S-transform spectrogram is an effective method to characterize seizure EEG signals. The parameter p in Gaussian window allows us to adjust the energy concentration. After sufficient experiments, the value of p is set to 0.5 to obtain better EEG features.

In this work, the frequency range of the S-transform spectrogram is selected from 1 Hz to 35 Hz based on the fact that this range consists of most of the seizure features. Each 1-s EEG segment with time interval of 1/256 s and frequency-sampling interval of 1 Hz yields a real S-transform spectrogram matrix of 35×256. To concentrate the energy and reduce the matrix dimension, we divide the timeline into 32 blocks equally and sum up every block in time dimension. Finally, each 1-s EEG segment is

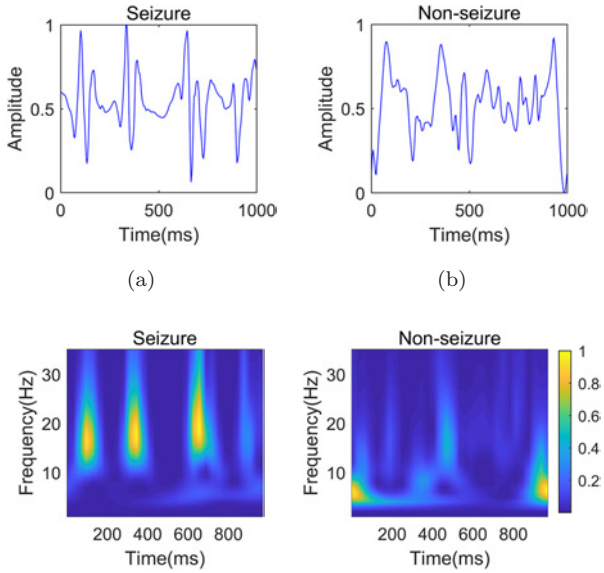

Fig. 3. The typical seizure and non seizure EEG signals and its S-transform spectrograms in 1-s. (a) and (b) represent two types of EEG signal in time domain. (c) and (d) are corresponding to the S-transform spectrogram of (a) and (b), respectively.

transformed into a 35×32 matrix. Figure 3 shows one second of non-seizure and seizure EEG signals and their S-transform spectrogram of single channel EEG signal from Patient 17. Figure 4 illustrates one hundred S-transform spectrograms of non-seizure and seizure EEG signals with first three channels in patient 12, and the data for different channels is represented with different colors. In Figs. 3 and 4, it can be seen obviously that in comparing with the S-transform spectrogram of nonseizure, the S-transform spectrogram of seizure has more ridge-like localized patterns at high frequency.

3.3. *Deep convolutional neural network*

CNN is inspired from the biological structure of a visual cortex, and owing to its powerful capability in extracting feature of image automatically, CNNs have been employed extensively in the field of computer vision.[47] In this work, the multi-channel S-transform spectrogram serves as an image representation of EEG signals. The architecture of the proposed CNN is depicted in Fig. 5. The first convolution layer has $163 \times 3 \times n$ kernels and $1 \times 1 \times 1$ stride, where n varies with the input channel in each layer, and the next three convolution layers have 32,

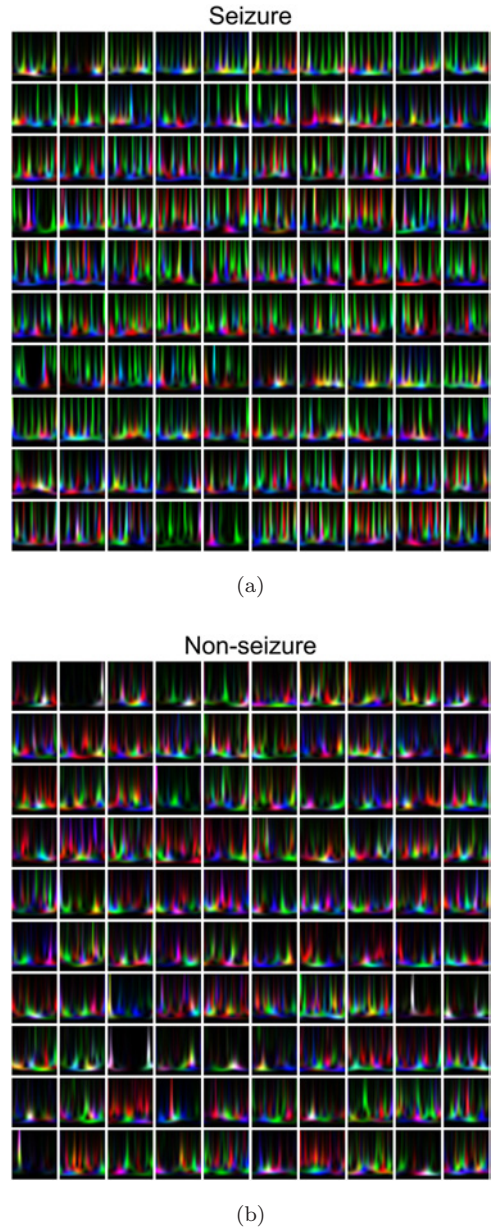

Fig. 4. The image-based representation of the three-channel S-transform spectrogram in RGB mode. (a) and (b) are S-transform spectrograms corresponding to 100 seizure segments and 100 non seizure segments of the first three channels, respectively, which are randomly selected from patient 11.

64 and 64 kernels, respectively. In order to reduce the parameters, max-pooling of size 2×2 acts on the output feature map of No. 1, 2, 4 convolution layers. Following the last max-pooling layer, there are two full-connect layers with the neuron of 120 and 2, respectively. The final output layer uses a soft-max activation function, capable of outputting the possibility

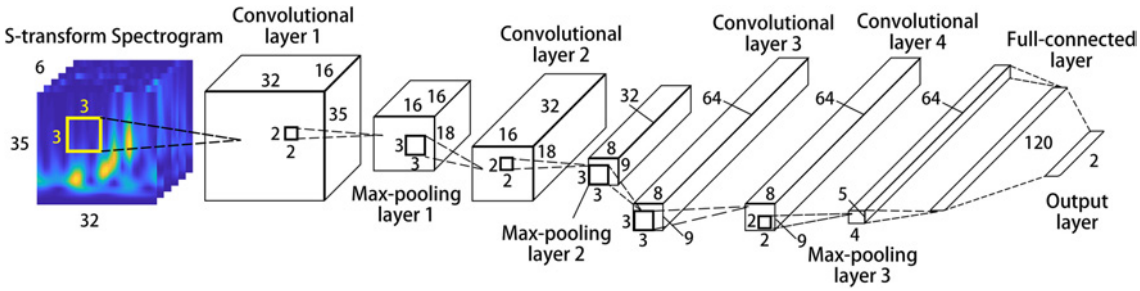

Fig. 5. Architecture of proposed CNN. Feature maps in each layer are illustrated as a cube, the length and height of the cube represent the size of each feature map, and the width represents the number of feature maps. In particular, the width of the last two layers represents the number of neurons.

of each true label. Cross entropy is used as the loss function of CNN model. Let $c = 2$ be the class number, m the number of the training sample and θ the whole parameters of the model, which needs to be optimized, the optimization target as follows:

$$\mathrm{CE}(\theta) = \mathrm{argmin}_\theta \left\{ -\sum_{i=1}^{m}\sum_{j=1}^{c} 1\{y^{(i)} = j\} \times \ln\left(\frac{e^{z_j^{(i)}}}{\sum_{k=1}^{c} e^{z_k^{(i)}}}\right)\right\}, \quad (5)$$

where $z_j^{(i)}$ relateds to θ denotes the total weighted sum of inputs to unit j for sample i, $y^{(i)}$ the true label of the ith sample and $1\{\cdot\}$ the indicator function.

Because the proposed CNN has approximately 215 thousand learnable parameters and the training dataset is relatively small, several approaches are adopted to improve the robustness of the model, speed up the training process and prevent the model from overfitting. First, every convolution layer is followed by a batch of normalization layer,[48] which can reduce the sensitivity to the CNN initial parameters. Second, we employ dropout of 20% in Nos. 3 and 4 convolution layers. Dropout has proven to be a simple but efficient way to prevent neural networks from overfitting.[49] Third, Adam optimization algorithm is utilized in the training stage of the CNN. Adam algorithm has an adaptive moment, which can accelerate the convergence speed of the model during the training stage.[50] The strong robustness of the method in choosing hyperparameters, such as learning rate, leads us to utilize the default parameters mentioned in the work of Kingma *et al.*[50] Finally, ReLu is chosen as the activation function in CNN to perform a threshold operation to each element of the input, where any value less than zero is set to zero. In this study, the learning rate is set to 0.0001 for all 21 patients, with the max epoch being 150 and the mini-batch size 256.

The data fed into CNN are the spectrogram matrices obtained with S-transform of 1-s EEG

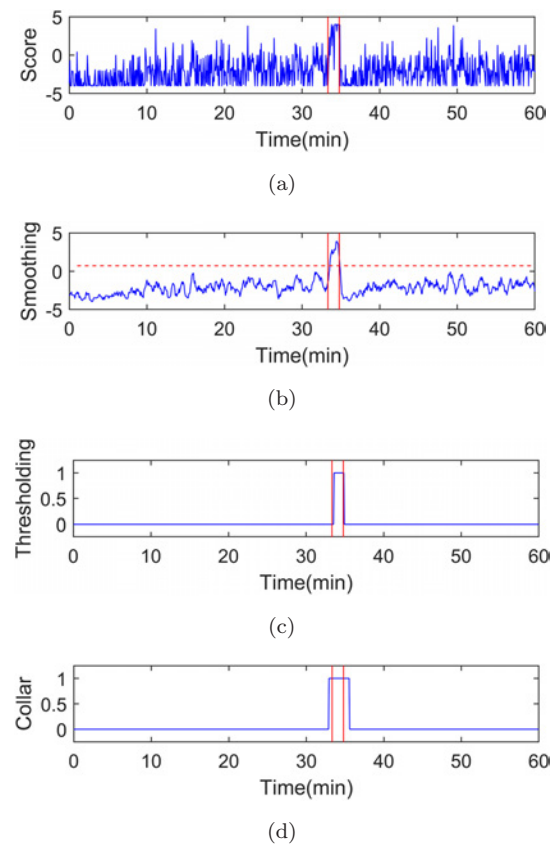

Fig. 6. (Color online) An example of postprocessing 1 h EEG signals. The data between two red vertical lines corresponding to seizure signals. (a) The score obtained from CNN output. (b) The score processed by moving average filter and the horizontal red dotted line represents the threshold. (c) The binary decisions after thresholding. (d) The final label after applying collar technique.

Table 2. The performance of our algorithm on validation set (using 10-fold cross validation for each patient).

Patient	Training set duration (min)	Validation set duration (min)	Single-segment accuracy	Multi-segment accuracy
1	2.92	0.32	93%	100.00%
2	22.10	2.46	94%	98.35%
3	17.81	1.98	95%	99.32%
4	13.98	1.55	99%	99.56%
5	22.78	2.53	80.60%	91.73%
6	16.19	1.80	94%	99.62%
7	47.86	5.32	92%	96.60%
8	26.83	2.98	82%	90.07%
9	40.14	4.46	91.13%	97.44%
10	22.91	2.55	84.68%	95.23%
11	53.37	5.93	98.28%	99.66%
12	9.20	1.02	99%	100.00%
13	36.07	4.01	94%	99.66%
14	53.89	5.99	96%	98.99%
15	21.79	2.42	95.12%	99%
16	25.59	2.84	90.54%	96.93%
17	18.50	2.06	98%	99.67%
18	2.69	0.30	84%	100.00%
19	3.48	0.39	86%	98.11%
20	26.87	2.99	94.19%	98.65%
21	10.24	1.14	90.38%	96.66%
Average	23.58	2.62	92.08%	97.85%

signals, which contains 35 rows, 32 columns and 6 channels. Before input, into CNN, the spectrogram matrix in each channel is normalized to 0-1 to speed up the training. Differential processing is performed on the output of two neurons in the last layer, and the output of single CNN ranges from -1 to 1. We define the score as the sum of four single CNN outputs, ranging from -4 to 4. The higher the score obtained, the greater the chance that this EEG segment belongs to a seizure.

3.4. *Postprocessing*

In order to reduce the isolated false positives and obtain the final detection, postprocessing containing smoothing, thresholding and collar technique, is applied to the output score. A simple way to reduce the FDR effectively is to apply a moving average filter to the CNN output score, and the filter is defined as follows:

$$y_t = \frac{1}{2N+1} \sum_{n=-N}^{N} x_{t+n}, \tag{6}$$

where $2N + 1$ is the length of smoothing, y is the filtered score and x is the CNN score. Then, the filtered score is compared with a fixed threshold Thr to obtain a binary label: 1 represents seizure segment and 0 represents normal segment. Though moving average filter can discard isolated false positives, it may cause the delayed response to seizure detection.

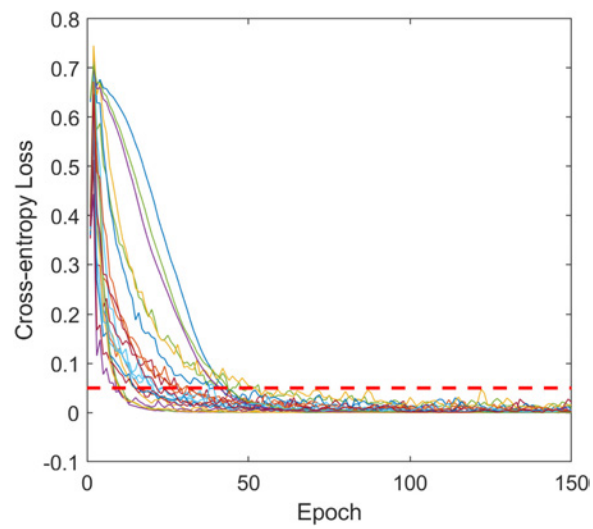

Fig. 7. Curves of the cross-entropy loss function for 21 patients in the training stage.

Table 3. The results of our proposed system on segment-based evaluation criterion.

Patient	Sensitivity	Specificity	Accuracy
1	100.00%	99.92%	99.92%
2	100.00%	99.95%	99.95%
3	100.00%	99.93%	99.93%
4	100.00%	98.95%	98.95%
5	100.00%	99.57%	99.57%
6	100.00%	99.37%	99.37%
7	100.00%	91.73%	91.74%
8	100.00%	99.68%	99.68%
9	100.00%	93.07%	93.09%
10	88.87%	81.68%	81.79%
11	85.07%	99.85%	99.81%
12	100.00%	99.95%	99.95%
13	100.00%	99.96%	99.96%
14	86.61%	99.84%	99.78%
15	100.00%	98.87%	98.87%
16	100.00%	99.45%	99.45%
17	100.00%	99.93%	99.93%
18	100.00%	99.79%	99.79%
19	100.00%	99.20%	99.20%
20	100.00%	99.90%	99.90%
21	76.71%	99.90%	99.85%
Average	97.01%	98.12%	98.12%

To compensate this drawback, the collar technique is applied to the binary label. By utilizing collar, every label, which belongs to seizures, is extended K points to both sides. For each patient, the value of N, Thr and K is adapted independently to obtain an optimal accuracy. An example of postprocessing 1-h CNN output and the detail procedure is shown in Fig. 6.

4. Results

Long-term intracranial EEG data from 21 patients are used for evaluating the proposed system. In training stage, 10-fold cross validation is employed to evaluate the classification performance of CNN, with the result being presented in Table 2. The single-segment accuracy is corresponding to the accuracy of 1-s EEG segment, and the multi-segment accuracy is for 4-s segment. It can be seen that multi-segment decision improves the performance significantly. The change of the cross-entropy loss function in training stage is shown in Fig. 7. The cross-entropy loss values for all 21 patients converge below 0.05 after 50 iterations.

Table 4. The results of our proposed system on event-based evaluation criterion.

Patient	Number of seizures experts marked	Number of seizures Number of seizures detected	Sensitivity	FDR (/h)
1	3	3	100.00%	0.09
2	2	2	100.00%	0.03
3	4	4	100.00%	0.03
4	4	4	100.00%	0.18
5	4	4	100.00%	0.03
6	2	2	100.00%	0.20
7	2	2	100.00%	0.98
8	1	1	100.00%	0.11
9	4	4	100.00%	1.18
10	4	4	100.00%	2.79
11	3	2	66.67%	0.03
12	3	3	100.00%	0.02
13	1	1	100.00%	0.04
14	3	2	66.67%	0.03
15	3	3	100.00%	0.15
16	4	4	100.00%	0.08
17	4	4	100.00%	0.03
18	4	4	100.00%	0.32
19	3	3	100.00%	1.11
20	4	4	100.00%	0.05
21	4	3	75.00%	0.05
Average	66	63	95.45%	0.36

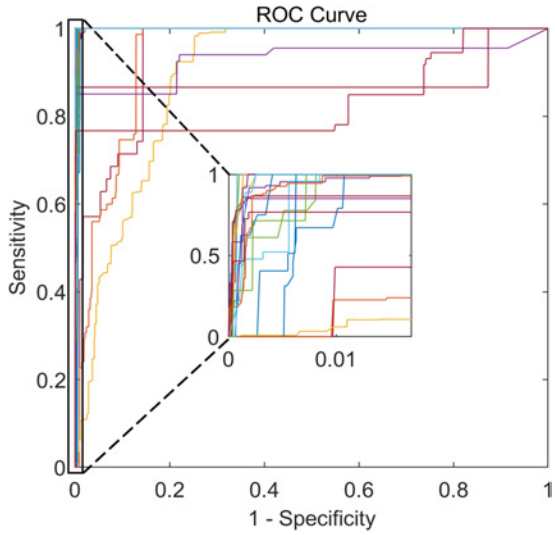

Fig. 8. ROC curves of all 21 patients. Each line with different color represents the ROC curve of a patient.

Table 5. The AUC for each patient.

Patient	AUC	Patient	AUC	Patient	AUC
1	99.82%	8	99.52%	15	99.78%
2	99.94%	9	93.93%	16	99.79%
3	99.94%	10	89.55%	17	99.94%
4	99.79%	11	93.84%	18	99.94%
5	99.68%	12	99.95%	19	99.67%
6	99.69%	13	99.93%	20	99.88%
7	94.05%	14	88.65%	21	84.01%
			Average: 97.21%		

In testing stage, both segment-based criterion and event-based criterion are utilized to evaluate the proposed system. Segment-based criterion contains sensitivity, specificity and accuracy. If we use FP, TP, FN, and TN to represent false positive, true positive, false negative and true negative, respectively, then the sensitivity, specificity and accuracy can be expressed as follows:

$$\text{Sensitive} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{7}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}, \tag{8}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}}. \tag{9}$$

Table 3 lists the detail results of all the patients based on segment-based criterion.

In clinical practice, more attention is placed on the FDR and the event-based method is more suitable for assessing the proposed system. In event-based criterion, sensitivity is defined as the ratio of the number of the detected events to the number of the events experts marked, and FDR as the mean duration of the false positive occurs in 1-h. The results of event-based criterion are presented in Table 4.

For the segment-based evaluation, the mean sensitivity of the 21 patients is 97.01%, while the mean specificity is 98.12%. There are 17 patients with a sensitivity of 100% and a specificity of more than 98%. Due to the missing detection, the sensitivity of the patients 11, 14 and 21 are relatively low. For patients 7 and 9, the difference of S-transform feature among the pre-ictal, post-ictal and ictal signal is relatively blur, but the difference between ictal and inter-ictal is obvious, so the specificity is relatively low. Many disconnected and reconnected events occurred in both normal and abnormal EEG recordings for patient 10, which led to a relatively low score in both sensitivity and specificity.

For the event-based evaluation, the mean sensitivity of 95.45% and the mean FDR of 0.36/h are achieved. All patients except No. 11, 14, 21 have no undetected seizure, and more than 12 patients have an FDR of less than 0.1/h. It is worth mentioning that though the length of training EEG recordings used in patient 1, 18 and 19 is only about half a minute, our system can still achieve a high sensitivity, specificity and low FDR.

To further illustrate the performance of our proposed system, the receiver operating characteristic (ROC) curve is plotted for each patient as shown in Fig. 8 and the area under the curve (AUC) is calculated as listed in Table 5. It can be observed that the mean-value of AUC in 21 patients exceeds 97%, and 15 of them, i.e. more than 2/3 of all patients have an AUC value greater than 99%, which indicates that our system has satisfying classification performance.

5. Discussion

This work presents a novel automatic seizure detection system based on the S-transform spectrogram and deep CNN. The Freiburg EEG database employed in this study is also applied to many other methods for evaluation. Table 6 lists other methods

Table 6. Comparison on the performance of different methods proposed in recent years.

Number	Method	Author	Year	Channels	Number of testing patients	Length of EEG data used (h)	Number of training seizures	Number of testing seizures	Sensitivity (%) (segment-/event-based)	FDR (/h)
1	Lacunarity and BLDA[51]	Zhou *et al.*	2013	3	21	289.14	—	—	96.25%/—	0.13
2	S-transform and singular value decomposition[34]	Xia *et al.*	2015	3	20	156.91	27	55	96.40/96.36	0.16
3	S-transform and boosting algorithm[35]	Yan *et al.*	2015	3	20	582.4	> 26	< 62	94.26/—	0.66
4	Log-Euclidean gaussian kernel-based sparse representation[52]	Yuan *et al.*	2016	3	21	597.95	25	62	95.11/96.77	0.211
5	Improved wavelet neural networks[53]	Geng *et al.*	2016	3	20	225	27	55	96.72/96.36	0.27
6	Improved sparse representation over learned dictionary[54]	Li *et al.*	2016	3	20	508.82	27	54	95.45/94.44	0.23
7	Phase-match error, deviation, fluctuation and LS-SVM[55]	Parvez *et al.*	2017	6	21	509	—	—	95.40/—	0.36
8	Sparse representation-based EMD and BLDA[56]	Yuan *et al.*	2017	3	21	575.26	28	59	93.54/94.92	0.223
9	Wavelet packet transform and imbalanced classification[57]	Yuan *et al.*	2017	3	20	501.4	28	44	95.74/97.73	0.37
10	Dictionary pair learning[58]	Ma *et al.*	2018	3	20	530	26	55	93.39/96.36	0.236
11	S-transform and deep CNN	Our Work	—	6	21	725.11	21	66	97.01/95.45	0.36

assessed on Freiburg EEG database in the part five years. Xia *et al.*[34] and Yan *et al.*[35] also used the S-transform as the feature of the EEG data, but they did not make a comprehensive assessment of all the data in the database and the sensitivity is relatively low. As is shown in Table 6, our method achieves a better segment-based sensitivity of 97.01% with the longest length of EEG data, the least number of training seizures, and the greatest number of testing seizures. Some studies employed a reduced number of testing patients (most of them excluded patient 10) and testing seizures. Hence, the performance measures, even when higher, are not fully comparable.

Traditional seizure detection is commonly achieved by manually selecting features, for example, spike rate, mean phase coherence, power spectral density ratio, etc. and its performance is vulnerable to the influence of those handcrafted features. In this work, the EEG features are automatically extracted by CNN with no need for any manual selection. Meanwhile, in the previous studies, channel fusion (or multi-channel decision) is a common postprocessing procedure for reducing FDR caused by single channel artifact, which needs hand-designed fusion rules.[53,58] This work, channel fusion can be achieved automatically in CNN without any manual intervention.

The parameters of S-transform and hyperparameters in deep CNN can directly influence the final results, and several experiments are implemented to compare the performance with different parameters using 10-fold cross validation. The parameter p in Eq. (3) is closely related to the feature of S-transform. Along with the increase of the p value, the width of the Gaussian window of the S-transform increases, which leads to the decrease of the frequency-domain resolution of the high-frequency EEG signal and the increase of the time-domain resolution of the low-frequency EEG signal. Hence, the choice of p is the key to obtain a good EEG feature. Here, we use the training data of Patient 9 to do several relevant experiments.

In general, the number of layers of deep CNN is positively related to the complexity of the model and the number of parameters. Too many parameters will increase the computational burden, and the model tends to overfit. Meanwhile, few parameters will leads to lower performance and under-fitting of the model. In this work, CNNs with 7-, 10-, 15- and 19-layer CNNs, which contain 1, 2, 4 and 6 convolution layers, respectively, are taken into consideration. Figures 9 and 10 show the classification results with different parameters of the S-transform and different parameters of the CNN model, where Fig. 9 illustrates the accuracies with different parameter

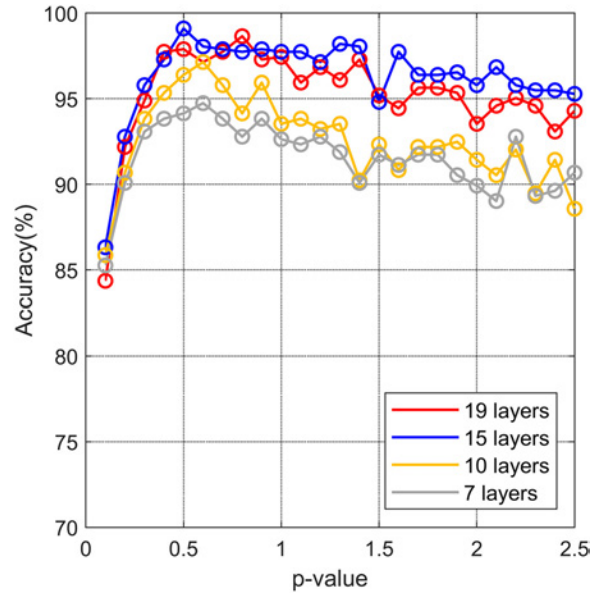

Fig. 9. Comparison of multi-segment accuracies under different number of CNN layers and p-value of S-transform spectrogram.

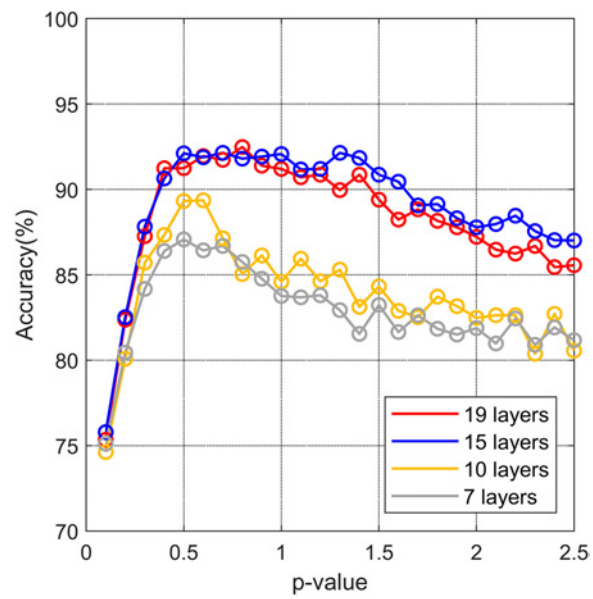

Fig. 10. Comparison of single-segment accuracies under different number of CNN layers and p-value of S-transform spectrogram.

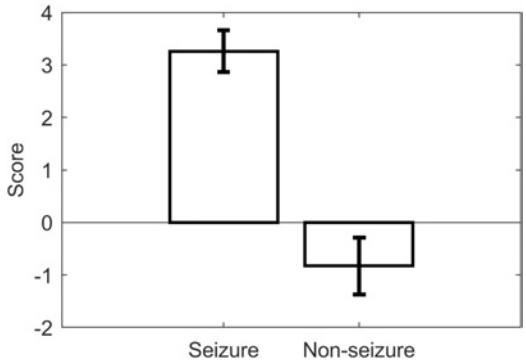

Fig. 11. Mean and standard deviation of the seizure and non seizure scores for 21 patients.

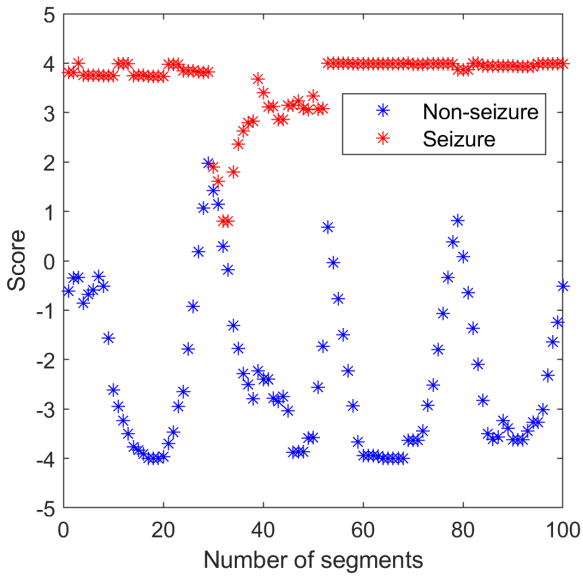

Fig. 12. 100 seizure and non seizure scores randomly selected from 21 patients.

combinations, and Fig. 10 the single-segment decision result. It is obvious that the performance of the 15- and 19-layer CNN is better than that of the 10- and 7-layer CNN in the same S-transform parameters. In the 15-layer CNN and the 19-layer CNN, the performance of the former architecture is slightly better than the latter, especially when the value of p is greater than 1.5. The multi-segment accuracy of the 15-layer CNN reaches 99.01% when the parameter p of the S-transform is set to 0.5. Hence, we choose the S-transform spectrogram with the parameter $p = 0.5$ as the time-frequency feature and the 15-layer CNN as the classifier in this work.

Figure 11 shows the mean and standard deviation of the scores for 21 patients with seizures and non seizures. Besides, we randomly pick 100 original scores of the seizure and non-seizure segments in 21 patients shown in Fig. 12, which demonstrates the excellent classification capability and the great robustness of our detection system. Although the proposed system achieves a high sensitivity in both segment- and event-based criterion, FDR still needs to be reduced. Figure 13 depicts an example of true positive detection and false positive detection with 6-channel raw EEG data.

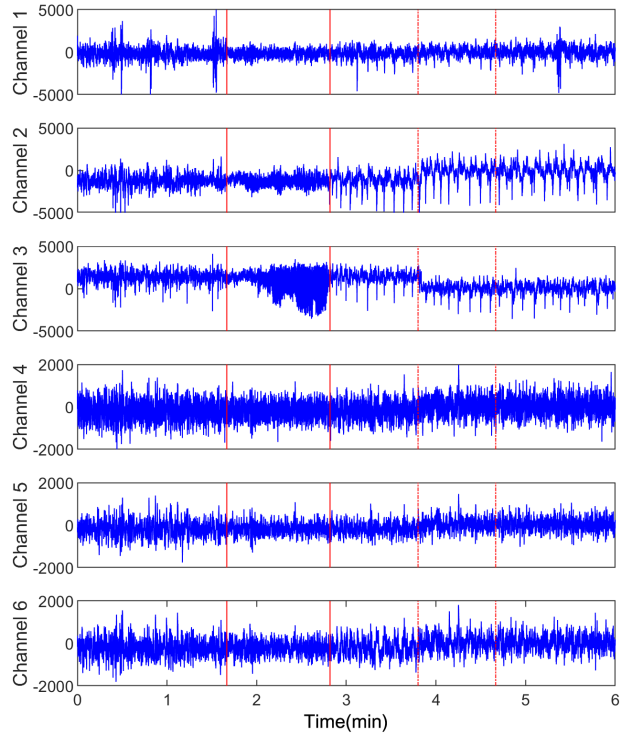

Fig. 13. An example of true positive detection and false positive detection of patient 21. The seizure event marked by experts is between two vertical lines and it is successfully detected by our system with sensitivity of 100%. The event marked between two dot dash lines is a false positive detection.

All experiments are carried out in MATLAB R2018a, running in a workstation with an Intel E3-1230v3 CPU, a NVDIA GTX 1060 6 GB GPU and 16 GB memory. To make full use of the complex computing ability of CPU and the parallel computing ability of GPU, the preprocessing including normalization and DWT are executed on CPU, while the deep CNN training and prediction are executed on GPU. It takes only about 0.05-s to predict the 4-s original EEG recordings with our proposed system, which can completely meet the needs of real-time

detection. For offline training, benefit from the parallelism of the GPU, our algorithm can be more effectively achieved. Taking Patient 3 as an example, it takes 12.8-s for the training dataset containing 40-min EEG recordings.

6. Conclusion

Usually, analysis of EEG depends on the visual inspection of neurologists. However, massive amounts of EEG data from patients make the task of reviewing EEGs very tedious and time-consuming. Therefore, the automatic seizure detection technology is of great worth for assisting the diagnosis of epilepsy. In this work, we propose a novel seizure detection system based on S-transform and deep CNN. In the proposed framework, each multi-channel EEG segment is filtered by DWT with Db4 wavelet. Then, their corresponding multi-channel time-frequency representation is calculated with fine-tuned S-transform and fed into a well-trained deep CNN for feature extraction and classification. Finally, the postprocessing including soothing, threshold and collar technique is applied to improve the final detection. Our system was assessed on Freiburg EEG database with approximately 725-h EEG recordings from all 21 patients. A sensitivity of 97.01% and specificity of 98.12% are achieved under the segment-based evaluation criteria, and 95% sensitivity and 0.36/h FDR are yielded under the event-based evaluation criteria. In our future work, we will further evaluate the performance of our algorithm with scalp EEG data.

Acknowledgments

The support of the Research Funds of Science and Technology Innovation Committee of Shenzhen Municipality (No. JCYJ20180305164357463), and the Development Program of Science and Technology of Shandong (No. 2014GSF118171) is gratefully acknowledged.

References

1. S. Ghosh-Dastidar, H. Adeli and N. Dadmehr, Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection, *IEEE Trans. Biomed. Eng.* **54**(9) (2007) 1545–1551.
2. P. Sharma, Y. U. Khan, O. Farooq, M. Tripathi and H. Adeli, A wavelet-statistical features approach for nonconvulsive seizure detection, *Clin. EEG Neurosci.* **45**(4) (2014) 274–284.
3. F. Mormann, R. G. Andrzejak, C. E. Elger and K. Lehnertz, Seizure prediction: The long and winding road, *Brain.* **130**(2) (2006) 314–333.
4. E. Wyllie, G. D. Cascino, B. E. Gidal and H. P. Goodkin, *Wyllie's Treatment of Epilepsy*: *Principles and Practice* (Lippincott Williams & Wilkins, 2012).
5. H. Adeli, and S. Ghosh-Dastidar, *Automated EEG-Based Diagnosis of Neurological Disorders*: *Inventing the Future of Neurology*, (CRC Press, Florida, 2010) p. 72.
6. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, Alzheimer's disease: models of computation and analysis of EEGs, *Clin. EEG and Neurosci.* **36**(3) (2005) 131–140.
7. M. Ahmadlou, H. Adeli and A. Adeli, New diagnostic EEG markers of the Alzheimer's disease using visibility graph, *J. Neural Transm.* **117**(9) (2010) 1099–1109.
8. Z. Sankari, H. Adeli and A. Adeli, Intrahemispheric, interhemispheric, and distal EEG coherence in Alzheimer's disease, *Clin. Neurophysiol.* **122**(5) (2011) 897–906.
9. M. Ahmadlou and H. Adeli, Functional community analysis of brain: A new approach for EEG-based investigation of the brain pathology, *Neuroimage* **58**(2) (2011) 401–408.
10. M. Ahmadlou and H. Adeli, Wavelet-synchronization methodology: A new approach for EEG-based diagnosis of ADHD, *Clin. EEG Neurosci.* **41**(1) (2010) 1–10.
11. M. Ahmadlou and H. Adeli, Fuzzy synchronization likelihood with application to attention-deficit/hyperactivity disorder, *Clin. EEG Neurosci.* **42**(1) (2011) 6–13.
12. J. Gotman, Automatic recognition of epileptic seizures in the EEG, *Electroencephalogr. clin. Neurophysiol.* **54**(5) (1982) 530–540.
13. S. Li, W. Zhou, Q. Yuan and Y. Liu, Seizure prediction using spike rate of intracranial EEG, *IEEE Trans. Neural Syst. Rehabil. Eng.* **21**(6) (2013) 880–886.
14. Y. Zheng, G. Wang, K. Li, G. Bao and J. Wang, Epileptic seizure prediction using phase synchronization based on bivariate empirical mode decomposition, *Clin. Neurophysiol.* **125**(6) (2014) 1104–1111.
15. A. Aarabi and B. He, Seizure prediction in patients with focal hippocampal epilepsy, *Clin Neurophysiol.* **128**(7) (2017) 1299–1307.
16. Z. Zhang and K. K. Parhi, Low-complexity seizure prediction from iEEG/sEEG using spectral power and ratios of spectral power, *IEEE Trans. Biomed. Circuits and Syst.* **10**(3) (2016) 693–706.

17. M. Adjouadi *et al.*, Interictal spike detection using the Walsh transform, *IEEE Trans. Biomed. Eng.* **51**(5) (2004) 868–872.
18. M. T. Tito *et al.*, A new algorithm for seizure detection using orthogonal transformations, *WSEAS Trans. Signal Process* **3** (2007) 155–162.
19. M. Ahmadlou, H. Adeli and A. Adeli, Improved visibility graph fractality with application for the diagnosis of autism spectrum disorder, *Physica A*: *Stat. Mech. Appl.* **391**(20) (2012) 4720–4726.
20. M. Ahmadlou, H. Adeli and A. Adeli, Fractality analysis of frontal brain in major depressive disorder, *Int. J. Psychophysiol.* **85**(2) (2012) 206–211.
21. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, A spatio-temporal wavelet-chaos methodology for EEG-based diagnosis of Alzheimer's disease, *Neurosci. Lett.* **444**(2) (2008) 190–194.
22. M. Ahmadlou, H. Adeli and A. Adeli, Fractality and a wavelet-chaos-methodology for EEG-based diagnosis of Alzheimer disease, *Alzheimer Disease Associated Disorders.* **25**(1) (2011) 85–92.
23. A. Eftekhar, W. Juffali, J. El-Imad, T. G. Constandinou and C. Toumazou, Ngram-derived pattern recognition for the detection and prediction of epileptic seizures, *PloS One.* **9**(6) (2014) e96235.
24. R. S. S. Kumari and J. P. Jose, Seizure detection in EEG using time frequency analysis and SVM, in (*ICETECT*), *2011 Int. Conf. Emerging Trends in Electrical and Computer Technology* (Nagercoil, India, 2011), pp. 626–630.
25. Y. Park, L. Luo, K. K. Parhi and T. Netoff, Seizure prediction with spectral power of EEG using cost-sensitive support vector machines, *Epilepsia.* **52**(10) (2011) 1761–1770.
26. S. Yuan, W. Zhou and L. Chen, Epileptic Seizure Prediction Using Diffusion Distance and Bayesian Linear Discriminate Analysis on Intracranial EEG, *Int. J. Neural Syst.* **28**(1) (2018) 1750043.
27. L. Guo, D. Rivero and A. Pazos, Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks, *J. Neurosci. Meth.* **193**(1) (2010) 156–163.
28. S. Ghosh-Dastidar and H. Adeli, Improved spiking neural networks for EEG classification and epilepsy and seizure detection, *Int. Comput.-Aided Eng.* **14**(3) (2007) 187–212.
29. S. Ghosh-Dastidar, H. Adeli and N. Dadmehr, Principal component analysis-enhanced cosine radial basis function neural network for robust epilepsy and seizure detection, *IEEE Trans. Biomed. Eng.* **55**(2) (2008) 512–518.
30. N. D. Truong *et al.*, Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram, *Neural Netw.* **105** (2018) 104–111.
31. O. Faust, U. R. Acharya, H. Adeli and A. Adeli, Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis, *Seizure.* **26**(2015) 56–64.
32. R. G. Stockwell, L. Mansinha and R. Lowe, Localization of the complex spectrum: The S transform, *IEEE Trans. Signal Process.* **44**(4) (1996) 998–1001.
33. C. R. Pinnegar, H. Khosravani and P. Federico, Time–frequency phase analysis of ictal EEG recordings with the S-transform, *IEEE Trans. Biomed. Eng.* **56**(11) (2009) 2583–2593.
34. Y. Xia, W. Zhou, C. Li, Q. Yuan and S. Geng, Seizure detection approach using S-transform and singular value decomposition, *Epilepsy Behav.* **52** (2015) 187–193.
35. A. Yan *et al.*, Automatic seizure detection using Stockwell transform and boosting algorithm for long-term EEG, *Epilepsy Behav.* **45** (2015) 8–14.
36. S. Assous, A. Humeau, M. Tartas, P. Abraham and J.-P. L'Huillier, S-transform applied to laser Doppler flowmetry reactive hyperemia signals, *IEEE Trans. Biomed. Eng.* **53**(6) (2006) 1032–1037.
37. S. Gonzalez Andino, R. Grave de Peralta Menendez, C. Lantz, O. Blank, C. Michel and T. Landis, Non-stationary distributed source approximation: An alternative to improve localization procedures, *Human brain Mapping.* **14**(2) (2001) 81–95.
38. A. Krizhevsky, I. Sutskever and G. E. Hinton, Imagenet classification with deep convolutional neural networks, in *Advances in Neural Information Processing Systems* (Lake Tahoe, USA, 2012), pp. 1097–1105.
39. P. Mirowski, D. Madhavan, Y. LeCun and R. Kuzniecky, Classification of patterns of EEG synchronization for seizure prediction, *Clin. Neurophysiol.* **120**(11) (2009) 1927–1940.
40. A. O'Shea, G. Lightbody, G. Boylan and A. Temko, Neonatal seizure detection using convolutional neural networks, in *2017 IEEE 27th Int. Workshop on Machine Learning for Signal Processing* (*MLSP*) (Tokyo, Japan, 2017), pp. 1–6.
41. U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan and H. Adeli, Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals, *Comput. Biol. Med.* **100** (2017) 270–278.
42. A. Antoniades *et al.*, Detection of interictal discharges with convolutional neural networks using discrete ordered multichannel intracranial EEG, *IEEE Trans. Neural Syst. Rehabil. Eng.* **25**(12) (2017) 2285–2294.
43. A. Antoniades, L. Spyrou, C. C. Took and S. Sanei. Deep learning for epileptic intracranial EEG data, in *2016 IEEE 26th Int. Workshop Machine Learning for Signal Processing* (*MLSP*) (Vietri sul Mare, Italy, 2016), pp. 1–6.
44. Freiburg Seizure Prediction Project, Freiburg, Germany. Available from: http://epilepsy.uni-freburg.de/freiburg-seizure-prediction-project/eeg-database.

45. Y. Khan and J. Gotman, Wavelet based automatic seizure detection in intracerebral electroencephalogram, *Clin. Neurophysiol.* **114**(5) (2003) 898–908.
46. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, A wavelet-chaos methodology for analysis of EEGs and EEG subbands to detect seizure and epilepsy, *IEEE Trans. Biomed. Eng.* **54**(2) (2007) 205–211.
47. T. N. Sainath, A.-R. Mohamed, B. Kingsbury and B. Ramabhadran, Deep convolutional neural networks for LVCSR, in *2013 IEEE Int. conf. Acoustics, Speech and Signal Processing* (*ICASSP*) (2013), pp. 8614–8618.
48. S. Ioffe and C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, arXiv:1502.03167.
49. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, Dropout: A simple way to prevent neural networks from overfitting, *J. Mach. Learn. Res.* **15**(1) (2014) 1929–1958.
50. D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, arXiv:1412.6980.
51. W. Zhou, Y. Liu, Q. Yuan and X. Li, Epileptic seizure detection using lacunarity and Bayesian linear discriminant analysis in intracranial EEG, *IEEE Trans. Biomed. Eng.* **60**(12) (2013) 3375–3381.
52. S. Yuan, W. Zhou, Q. Wu and Y. Zhang, Epileptic Seizure Detection with Log-Euclidean Gaussian Kernel-Based Sparse Representation, *Int. J. Neural Syst.* **26**(3) (2016) 1650011.
53. D. Geng, W. Zhou, Y. Zhang and S. Geng, Epileptic seizure detection based on improved wavelet neural networks in long-term intracranial EEG, *Biocybern. Biomed. Eng.* **36**(2) (2016) 375–384.
54. J. Li, W. Zhou, S. Yuan, Y. Zhang, C. Li and Q. Wu, An Improved Sparse Representation over Learned Dictionary Method for Seizure Detection, *Int. J. Neural. Syst.* **26**(1) (2016) 1550035.
55. M. Z. Parvez and M. Paul, Seizure prediction using undulated global and local features, *IEEE Trans. Biomed. Eng.* **64**(1) (2017) 208–217.
56. S. Yuan, W. Zhou, J. Li and Q. Wu, Sparse representation-based EMD and BLDA for automatic seizure detection, *Med. Biol. Eng. Comput.* **55**(8) (2017) 1227–1238.
57. Q. Yuan *et al.*, Epileptic seizure detection based on imbalanced classification and wavelet packet transform, *Seizure.* **50** (2017) 99–108.
58. X. Ma, N. Yu and W. Zhou, Using Dictionary Pair Learning for Seizure Detection, *Int. J. Neural Syst.* **29** (2018) 1850005.