Epileptic Seizure Detection Based on Bidirectional Gated Recurrent Unit Network

Yanli Zhang, Shuxin Yao, Rendi Yang, Xiaojia Liu, Wenlong Qiu, Luben Han, Weidong Zhou, and Wei Shang

***Abstract*—Visual inspection of long-term electroencephalography (EEG) is a tedious task for physicians in neurology. Based on bidirectional gated recurrent unit (Bi-GRU) neural network, an automatic seizure detection method is proposed in this paper to facilitate the diagnosis and treatment of epilepsy. Firstly, wavelet transforms are applied to EEG recordings for filtering pre-processing. Then the relative energies of signals in several particular frequency bands are calculated and inputted into Bi-GRU network. Afterwards, the outputs of Bi-GRU network are further processed by moving average filtering, threshold comparison and seizure merging to generate the discriminant results that the tested EEG belong to seizure or not. Evaluated on CHB-MIT scalp EEG database, the proposed seizure detection method obtained an average sensitivity of 93.89% and an average specificity of 98.49%. 124 out of 128 seizures were correctly detected and the achieved average false detection rate was 0.31 per hour on 867.14 h testing data. The results show the superiority of Bi-GRU network in seizure detection and the proposed detection method has a promising potential in the monitoring of long-term EEG.**

***Index Terms*—Epileptic seizure detection, scalp EEG, bidirectional gated recurrent unit, wavelet transform.**

I. Introduction

EPILEPSY is a chronic brain disease and epileptic seizures caused by abnormal discharge of brain nerve cells have complex and diverse clinical manifestations [1].

Manuscript received July 1, 2021; revised November 29, 2021 and December 20, 2021; accepted January 10, 2022. Date of publication January 14, 2022; date of current version January 28, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 61801269 and Grant 61876100, in part by the Shandong Provincial Science and Technology Support Program of Youth Innovation Team in Colleges under Grant 2019KJN041 and Grant 2020KJN005, and in part by the Ph.D. Research Startup Foundation of Shandong Technology and Business University under Grant BS201611. *(Corresponding author: Yanli Zhang.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of the Second Hospital of Shandong University.

Yanli Zhang, Shuxin Yao, Xiaojia Liu, Wenlong Qiu, and Luben Han are with the School of Information and Electronic Engineering, Shandong Technology and Business University, Yantai 264005, China (e-mail: yrmzhang@sdtbu.edu.cn).

Rendi Yang is with the School of Electromechanical and Automotive Engineering, Yantai University, Yantai 264005, China (e-mail: yrndy@126.com).

Weidong Zhou is with the School of Microelectronics, Shandong University, Jinan 250100, China (e-mail: wdzhou@sdu.edu.cn).

Wei Shang is with the Second Hospital of Shandong University, Jinan 250100, China (e-mail: wshang85@aliyun.com).

Digital Object Identifier 10.1109/TNSRE.2022.3143540

In the diagnosis and treatment of patients with epilepsy, electroencephalography (EEG) plays a central role since it shows the electrophysiological activity of cerebral cortex in a convenient and relatively inexpensive way [2], [3]. Long-term EEG recordings are usually analyzed and inspected by neurologists visually, which is a laborious and time-consuming task. It is necessary to study automatic seizure detection methods to provide timely treatment for people with epilepsy and reduce the burden of neurologists.

One of the first seizure detection systems was introduced by Gotman [4], in which three EEG measures including relative amplitude, average duration of half-waves and coefficient of variation were selected to recognize epileptic seizures. Since then, many scholars have improved it and various seizure detection methods have been proposed in the past decades [5]–[8]. Feature extraction from EEG is one of the important processes for most published seizure detection methods. To recognize the seizure EEG, many features has been extracted from time, frequency and time-frequency domain, such as spike rate, power spectral density ratio and signal energy from wavelet transform or Hilbert–Huang transform [9]–[12]. Meanwhile, a variety of nonlinear features, such as largest Lyapunov exponent [13], fractal dimension [14], and entropies [15], [16], also have been employed to study the EEG difference between ictal and interictal phases. At present, some detection systems have utilized the feature extraction ability of deep neural networks [17], however, the extracted features generally have no clear physical meanings.

Another important part of automatic seizure detection methods or systems is classifier. Support vector machine, random forest and various artificial neural networks have exhibited different performance in distinguishing seizure EEG [18]–[25]. However, the robustness and generalization ability of many classifiers are unsatisfactory when the testing data has a different probability distribution with that of the training data. What's more, the training of some classifiers, especially those based on deep neural network, has high demand on the amount of labeled samples and the configuration of hardware [26], [27]. Thus the design of EEG classifier is still one of research hotspots in the field of automatic seizure detection.

In this work, we propose a seizure detection method based on relative energies of time-frequency EEG components and bidirectional gated recurrent unit (Bi-GRU) neural network. As a type of recurrent neural network (RNN), GRU network has effectiveness at capturing long-term dependencies in long sequence, and has superiority in structure and calculation cost

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

compared with other RNN networks such as long short-term memory (LSTM) network [28]–[34].Considering that epileptic seizure has a dynamic evolutionary process and future EEG information may be helpful for the pattern recognition of current EEG, we designed a Bi-GRU network structure in our seizure detection method. By adding a hidden layer with back propagation GRU blocks, the Bi-GRU network analyzes the EEG information in both forward and reverse time directions, which thus can improve the detection performance. The results on public and clinical EEG databases demonstrate that our Bi-GRU-based seizure detection method has the advantages of simple structure, low calculation cost and good ability in mining dependency information in long-term EEG.

The remainder of this paper is organized as follows. The EEG databases used in this work are introduced in Section II. Section III focus on describing the proposed seizure detection method and the used performance evaluation method. Evaluation results are exhibited in Section IV, and Section V discusses on the proposed seizure detection method. Finally, a conclusion of this paper is given in Section VI.

II. Database

The EEG data used in this work partly come from CHB-MIT scalp EEG database that was collected at the Children's Hospital Boston [35], [36]. In the process of EEG data acquisition, International 10-20 system of EEG electrode positions was used, with 256 Hz sampling rate and 16-bit resolution. EEG recordings in CHB-MIT database were collected from 23 epilepsy patients and were grouped into 24 cases, among which case chb21 and chb01 were collected from a same subject and 1.5 years apart. About 9 to 42 continuous EEG files are available for each case, and most files contain one hour of EEG recordings, excepting some files belonging to cases chb04, chb06, chb07, chb09, chb10 and chb23, which are two or four hours long. The EEG files in this database include a total of 198 seizures, and the beginning and end of each seizure has been annotated. The detailed information of CHB-MIT EEG data used in this work is listed in Table I. For each case, we constructed non-overlapping training and testing datasets, which were respectively used to establish and evaluate the seizure detection system. Except that the EEG of one to five seizures were selected as training data, non-seizure EEG were also included in training dataset. The amount of non-seizure EEG in training dataset was equal to or 2-4 times to that of seizure EEG. In total, about 198.14 min EEG recordings were served as training data for 24 cases. In addition, the rest EEG recordings were applied as testing data and the total duration was 867.14 h with 128 seizures included.

The other part of EEG data used in this work were collected from five epilepsy patients in Second Hospital of Shandong University (SH-SDU), Jinan, China. In a total of 58.40 h scalp EEG recordings, there are 41 seizures included, which have been marked by clinical experts. For each patient, a training set and a testing set were constructed following the same method as CHB-MIT database. In total, the duration of training data was 47.54 min and that of testing data was 57.56 h. The number of seizures in training and testing dataset respectively were 12 and 29. Table II gives the details of SH-SDU dataset.

TABLE I
Information of CHB-MIT EEG Data Used in This Work

Case	Gender	Age	Number of Seizures	Training data duration (min)		Testing data duration (h)
				Seizure	Non-seizure	
chb01	F	11	7	2.60	10.40	37.51
chb02	M	11	3	1.33	5.33	34.98
chb03	F	14	7	2.13	8.53	35.86
chb04	M	22	4	3.60	14.40	142.69
chb05	F	7	5	3.93	15.73	37.00
chb06	F	1.5	10	1.07	2.13	56.47
chb07	F	14.5	3	2.40	4.80	62.86
chb08	M	3.5	5	2.67	5.33	14.78
chb09	F	10	4	2.53	10.13	63.50
chb10	M	3	7	1.53	6.13	46.01
chb11	F	12	3	0.33	0.67	33.77
chb12	F	2	40	1.53	3.07	19.25
chb13	F	3	12	1.07	2.13	7.33
chb14	F	9	8	1.33	2.67	23.67
chb15	M	16	20	6.80	13.60	37.39
chb16	F	7	10	0.47	0.47	15.41
chb17	F	12	3	1.47	4.40	18.54
chb18	F	18	6	2.27	6.80	32.95
chb19	F	19	3	1.27	5.07	27.92
chb20	F	6	8	1.40	5.60	26.59
chb21	F	13	4	0.80	3.20	30.93
chb22	F	9	3	1.20	4.80	29.72
chb23	F	6	7	4.07	8.13	18.49
chb24	-	-	16	2.27	4.53	13.53
Total	-	-	198	50.07	148.07	867.14

TABLE II
Information of SH-SDU EEG Data Used in This Work

Patient	Gender	Age	Number of Seizures	Training data duration (min)		Testing data duration (h)
				Seizure	Non-seizure	
1	F	28	19	2.27	9.07	20.38
2	M	61	10	9.73	9.73	15.70
3	M	76	3	4.53	4.53	11.85
4	F	37	6	1.00	4.00	3.68
5	F	38	3	0.53	2.13	5.95
Total	-	-	41	18.07	29.47	57.56

III. Method

The block diagram of the proposed seizure detection method is presented in Fig. 1, which includes EEG pre-processing, feature extraction, classification model based on Bi-GRU neural network and post-processing. Each part of the detection method is described in detail as follows.

A. Pre-Processing

In the pre-processing stage, continuous EEG recordings are firstly segmented without overlapping by a sliding time window with the length of 4 s. Then wavelet transform is performed on the segmented EEG data. As a time-frequency analysis method, discrete wavelet transform decomposes EEG signals into detail and approximate coefficients to characterize the signal components in different time-frequency regions [37].

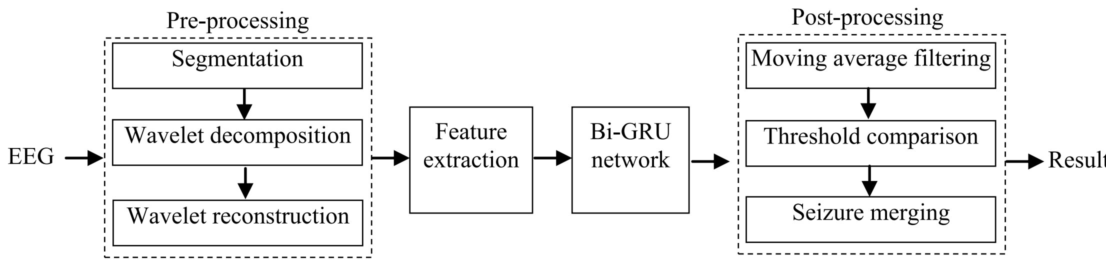

Fig. 1. Block diagram of the proposed seizure detection method.

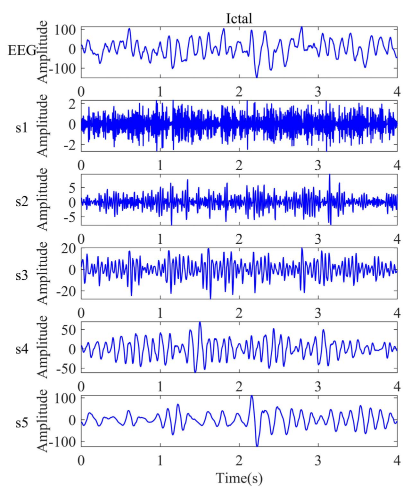

Fig. 2. A segment of ictal EEG and its reconstructed signals.

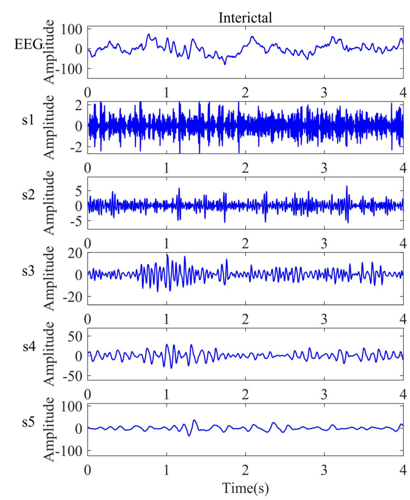

Fig. 3. A segment of interictal EEG and its reconstructed signals.

In this work, Daubechies-4 wavelet is adopted as wavelet function and EEG signals are decomposed in 5 scales. For the EEG with a sampling rate of 256 Hz, the detail coefficients in 5 scales respectively represent the signal components in the frequency band of 64-128 Hz, 32-64 Hz, 16-32 Hz, 8-16 Hz and 4-8 Hz. Obviously, the frequency bands in wavelet scale 3, 4 and 5 are approximately consistent with that of seizure activity, whose is mainly in the range of 3-30 Hz [5]. The detail coefficients in these three scales thus are selected to perform single-scale signal reconstruction and feature extraction, and all the reconstructed signals are equal in length with original EEG.

Fig. 2 and Fig. 3 illustrate two segments of EEG and their reconstructed signals. The signals shown in the top panels of Fig. 2 and Fig. 3 respectively belong to ictal and interictal EEG. After carrying out wavelet decomposition and reconstruction on five scales, the reconstructed signals on scale 1 to 5 are respectively presented as $s1$ to $s5$. As can be seen in Fig. 2 and Fig. 3, the reconstructed signals on scale 3 to 5 have obvious amplitude differences between two classes of EEG.

B. Feature Extraction

Following the above pre-processing, feature extraction is carried out on the reconstructed signals on scale 3, 4 and 5. To facilitate the subsequent establishment of Bi-GRU neural network, each EEG epoch of 4 s are subdivided into eight time slices of 0.5 s. For each EEG channel in j-th time slice, the relative energies of the reconstructed signal on the scale 3, 4 and 5 are respectively calculated as follows.

$$ER_j(k) = E_j(k) / \sum_{i=3}^{5} E_j(i) \tag{1}$$

Here, $1 \leq j \leq 8$, and the value of k is 3, 4 or 5, which indicates the wavelet scale. $E_j(k)$ is the energy of the reconstructed signal s_{jk} on k-th scale in j-th time slice, which is given as:

$$E_j(k) = \sum_{t=1}^{N} \left|s_{jk}(t)\right|^2 \tag{2}$$

where N is the number of sample points of the reconstructed signal s_{jk}, and the value of N is 128 in this work, because the sampling rate is 256 Hz and the length of the signal s_{jk} is 0.5 s.

It can be seen from Eq. (1) that the relative energy values represent the energy proportions of the signal components on different wavelet scales. Relative energy values are in the range 0 to 1, and have different distributions in multi-resolution scales for different EEG patterns and different seizure types. For the part of an EEG recording where a seizure occurs, the relative energy on a particular wavelet scale will rise.

For each segment of EEG with the length of 4 s, a 3×8 feature matrix can be obtained. To produce the feature matrix for each 4-s EEG epoch, we stack the calculated relative energy values according to the order of channels, for example,

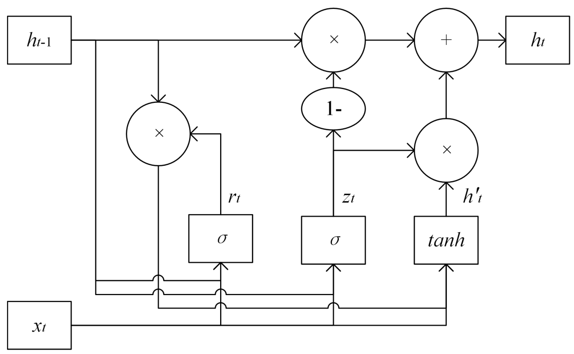

Fig. 4. Internal structure of GRU cell.

the feature matrix will have the dimension of 69×8 if there are 23 EEG channels. The feature matrix of each epoch will be inputted into the following classification model based on bidirectional GRU network.

C. Bidirectional GRU Network

Like LSTM network, GRU was proposed to address the gradient vanishing problem existing in recurrent neural network [28], and it learns the long-term dependencies in long sequence applications through internal gating mechanism. A GRU cell has two gates, that is update gate z_t and reset gate r_t. The activation of the gates in GRU only depends on current input and previous output.

The internal structure of GRU cell is illustrated in Fig. 4, where x_t and h_t respectively are the input vector and the hidden state at time slice t, and h'_t is a candidate of hidden state. At the time slice t, the reset gate r_t determines how much historical information is required to forget and the update gate z_t controls how to update the hidden state using the current EEG information. The detailed formulas are given as follows.

$$r_t = \sigma\left(W_r \cdot \left[h_{t-1}, x_t\right]\right) \tag{3}$$

$$z_t = \sigma\left(W_z \cdot \left[h_{t-1}, x_t\right]\right) \tag{4}$$

$$h'_t = \tanh\left(W_{h'} \cdot \left[r_t * h_{t-1}, x_t\right]\right) \tag{5}$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * h'_t \tag{6}$$

In above formulas, $\sigma(\cdot)$ and $\tanh(\cdot)$ respectively are the sigmoid function and the hyperbolic tangent function. Symbol $\cdot$ and * represent matrix multiplication and Hadamard product respectively, and [] is the concatenation of two vectors. W_r, W_z and $W_{h'}$ are the weight matrixes to be learned by GRU network training.

In GRU recurrent neural network, the information generally propagates in order of time along the GRU cells of a hidden layer. To utilize the dynamic evolutionary mechanism of epileptic seizures and better capture the long-term dependencies in EEG, a neural network model based on bidirectional GRU is constructed in this work and its structure is presented in Fig. 5.

As can be seen in Fig. 5, the proposed neural network model has two hidden layers where the information propagates in forward and backward direction respectively. Each hidden layer has eight GRU cells respectively corresponding to 8 time slices (T1-T8). The input x_t of GRU cell is a vector consisting of the relative energy values extracted from the multichannel EEGs in t-th time slice, and the dimension of hidden state output h_t is about twice to thrice the length of the input. The hidden states of the last GRUs of both directions of forward and backward are then concatenated and transmitted into a dense (fully-connected) layer with one neuron and a sigmoid activation function. Thus, a single output variable is obtained for each EEG epoch of 4 s.

Before the beginning of Bi-GRU network training, the feature samples of seizure EEG are up-sampled by interpolation method to keep the amount of seizure samples equal to that of non-seizure samples, if the duration of seizure EEG in training dataset is less than that of non-seizure EEG. In this work, the multiple of interpolation is determined by the duration ratio of non-seizure EEG to seizure EEG in the training dataset. In the training process of Bi-GRU network, Adam optimization algorithm is selected and binary cross-entropy is employed as loss function. The number of iterations is 300 and batch size is set as 128. As a hyper-parameter, the hidden state dimension of GRU cell is determined by k-fold cross-validation and the value of k is selected as ten in this work.

D. Post-Processing

In the process of seizure detection from EEG recordings, a series of post-processing are needed to carry out on the outputs of the trained Bi-GRU network to obtain the category labels of testing EEG. In detailed, a moving average filter is firstly utilized to the output sequence to reduce the random noises, which can be defined as:

$$y(i) = \frac{1}{2M+1} \sum_{k=-M}^{M} x(i+k) \tag{7}$$

where $x(i)$ is output of Bi-GRU network and $y(i)$ denotes the filtered signal. As can be seen in Eq.(7), $y(i)$ is the average of $2M+1$ points that include the current center point $x(i)$ and its neighborhood points on the left and right sides. $2M+1$ is also known as the order of the moving average filter. In this work, the value of $2M+1$ is in the range of 5 to 11 and different for each patient, which is determined by receiver operating characteristic (ROC) curves.

Then the filtered outputs are compared with a threshold determined in the training stage of Bi-GRU classification model. The threshold is specific for each patient and determined based on the criterion of minimum misclassification for training samples. After comparison with the selected threshold, testing samples are labeled as seizure or non-seizure.

Finally, considered the frequency of epileptic seizures, those seizures in a short time interval will be merged and regarded as belonging to a same seizure event. The time interval used in this work is 1 to 3 min.

Fig. 6 illustrates the post-processing procedures when multichannel EEG sequences in CHB-MIT database are tested by the proposed seizure detection method. The EEG shown in Fig. 6(a) comes from the F4-C4 channel of case chb02, in which the part between two vertical lines is a seizure event marked by clinical expert. Fig. 6 (b) presents the outputs from the trained Bi-GRU network of chb02 and the sequence

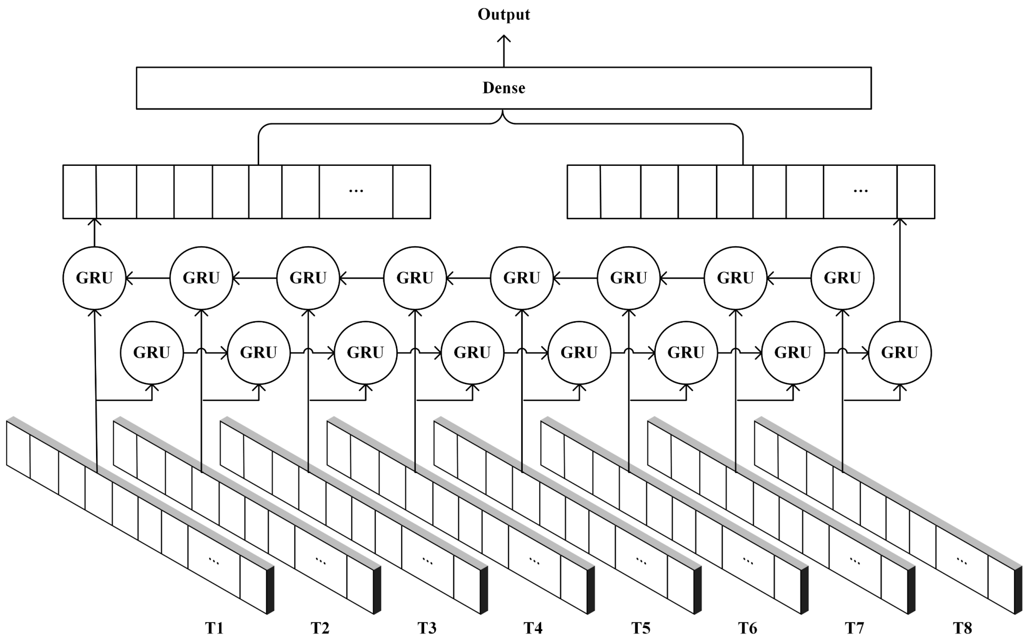

Fig. 5. Structure of the proposed Bi-GRU network.

given in Fig. 6 (c) is the result of moving average filtering. In Fig. 6 (c), the parts with the amplitude greater than the threshold indicated in red dotted line are considered as seizure and are labeled as 1 in Fig. 6(d). As can be seen in Fig. 6 (d), there are two parts labeled as 1 (seizure) and the interval between them is less than 1 min. Then the two seizure parts are merged in Fig. 6 (e) and we are of the opinion that one seizure event is detected in the current EEG recording.

E. Performance Evaluation

To evaluate the performance of the proposed seizure detection method, segment-based and event-based assessment indexes are both applied in this work. In the segment-based evaluation, detection sensitivity, specificity and accuracy rate are used as evaluation criteria, which can be respectively calculated by the following three formulas.

$$Sensitivity = \frac{TP}{TP + FN} \times 100\% \tag{8}$$

$$Specificity = \frac{TN}{TN + FP} \times 100\% \tag{9}$$

$$Accuracy\ rate = \frac{TP + TN}{TP + FN + FP + TN} \times 100\% \tag{10}$$

In above formulas, TP (true positive) and TN (true negative) respectively refer to the number of seizure and non-seizure segments which are correctly recognized by our detection system. FP (false positive) refers to the number of non-seizure EEG segments incorrectly judged as seizure by the detection system, and FN (false negative) is that of incorrectly labeled seizure segments.

The method of event-based assessment uses sensitivity and false detection rate (FDR) as evaluation criteria. Event-based sensitivity is applied to describe the percentage of seizure events correctly detected by the proposed system. For each patient, FDR is defined as the average number of false detections per hour during the non-seizure period.

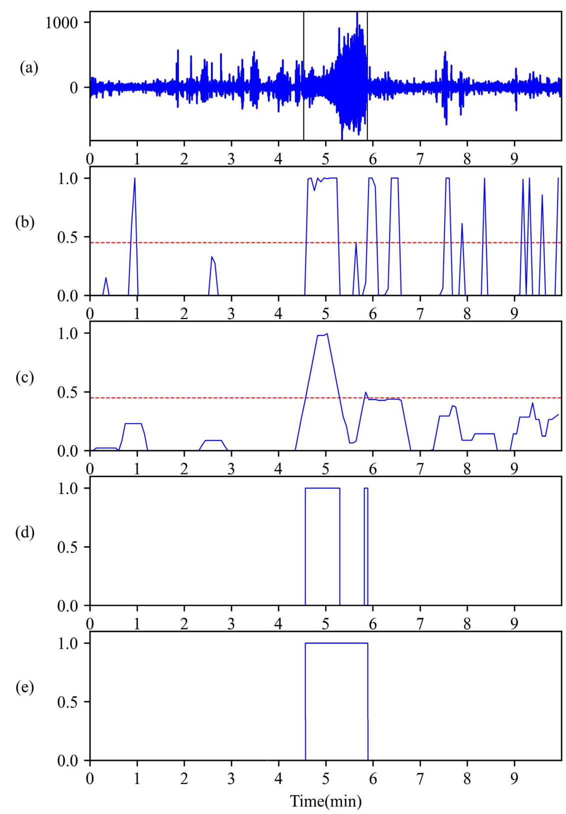

Fig. 6. A graphic illustration of post-processing procedures. (a) An EEG sequence coming from one of channels of case chb02 in CHB-MIT database, and the part between two vertical lines is a seizure event marked by expert. (b) The output of the trained Bi-GRU network. (c) The result after moving average filtering. (d) The binary decision after threshold comparison. (e) The final detection result after merging the seizures having the interval less than 1 min.

TABLE III
RESULTS OF SEGMENT-BASED PERFORMANCE ASSESSMENT ON CHB-MIT EEG DATA

Case	Sensitivity (%)	Specificity (%)	Accuracy rate (%)
chb01	100.00	99.54	99.54
chb02	90.91	99.34	99.33
chb03	100.00	99.69	99.69
chb04	97.50	98.42	98.42
chb05	95.00	99.95	99.94
chb06	100.00	99.99	99.99
chb07	97.78	99.78	99.77
chb08	98.41	94.53	94.59
chb09	96.77	99.46	99.46
chb10	94.25	99.87	99.86
chb11	100.00	99.52	99.52
chb12	94.62	97.50	97.47
chb13	91.21	84.83	84.92
chb14	90.00	98.83	98.82
chb15	95.47	98.57	98.53
chb16	50.00	99.62	99.60
chb17	100.00	99.81	99.81
chb18	84.09	99.85	99.83
chb19	89.74	99.23	99.22
chb20	100.00	99.97	99.97
chb21	91.89	98.96	98.95
chb22	100.00	99.43	99.43
chb23	100.00	99.08	99.09
chb24	95.65	98.07	98.05
Average	93.89	98.49	98.49

TABLE IV
RESULTS OF EVENT-BASED PERFORMANCE ASSESSMENT ON CHB-MIT EEG DATA

Case	Number of seizures in testing set	Number of detected seizures	Sensitivity (%)	FDR (/h)
chb01	4	4	100.00	0.05
chb02	2	2	100.00	0.06
chb03	5	5	100.00	0.17
chb04	2	2	100.00	0.78
chb05	3	3	100.00	0.08
chb06	6	6	100.00	0.00
chb07	2	2	100.00	0.40
chb08	4	4	100.00	1.37
chb09	2	2	100.00	0.43
chb10	5	5	100.00	0.00
chb11	2	2	100.00	0.00
chb12	25	25	100.00	0.58
chb13	8	8	100.00	0.14
chb14	5	5	100.00	0.76
chb15	15	15	100.00	0.57
chb16	4	2	50.00	0.06
chb17	2	2	100.00	0.00
chb18	4	3	75.00	0.09
chb19	2	2	100.00	0.18
chb20	6	6	100.00	0.04
chb21	3	2	66.67	0.68
chb22	2	2	100.00	0.30
chb23	3	3	100.00	0.65
chb24	12	12	100.00	0.15
Average	-	-	95.49	0.31

IV. RESULTS

The proposed epileptic seizure detection method was implemented on Matlab 2016a and TensorFlow 2.1.0 with Python 3.7.7, which were running on a notebook computer with a 2.9 GHz AMD Ryzen 7 4800H processor. The running time in training stage of a Bi-GRU network is about 14 s and the testing time taken for one-hour multi-channel EEG are about 41.8 s, among which the time for EEG pre-processing and feature extraction is 41.2s.

The performance of the proposed seizure detection method was evaluated on the testing data of all cases in CHB-MIT database. The results of segmented-based and event-based performance assessment are listed in Table III and Table IV respectively. It can be noticed from Table III that a sensitivity of 93.89% and a specificity of 98.49% were achieved averaging on 24 cases in the segment-based performance evaluation. Among all the cases, eight cases have the sensitivity of 100%, and sixteen cases have the specificity and the accuracy rate higher than 99%.

On the other hand, 124 out of 128 seizures in the testing datasets were successfully detected and an average sensitivity of 95.49 was achieved for 24 cases, which can be seen in Table IV. In addition, there were 14 cases having the false detection rate (FDR) less than 0.2 times per hour and high FDR mainly occurred on the testing data of case 4, 8, 12, 14, 15, 21 and 23. In total, an average false detection rate of 0.31 times per hour was obtained.

Table V shows the results of the performance evaluation on SH-SDU EEG data. The sensitivities of the first two patients are less than 80%, and the specificity values of patient 1 and 4 failed to achieve 90%. Averaged on five patients, the sensitivity, specificity and accuracy rate respectively are 87.08%, 88.65% and 88.55%. In addition, 2 out of total 29 seizures are missed detection and the averaged false detection rate is more than 1 per hour. The total results are not as good as those of CHB-MIT database. One of the reasons is that there are more noises and artifacts in SH-SDU EEG dataset according to the clinical manifestations of some patients. What's more, high seizure frequency and short duration both increase the detection difficulty, for example, the number of seizures contained in about 20 h of testing data are up to 16 for patient 1.

TABLE V
RESULTS OF SEGMENT-BASED PERFORMANCE ASSESSMENT ON SH-SDU EEG DATA

Patient	Sensitivity (%)	Specificity (%)	Accuracy rate (%)
1	77.78	81.40	81.37
2	76.37	95.70	95.15
3	93.02	90.01	90.02
4	94.12	82.60	82.66
5	94.12	93.54	93.54
Average	87.08	88.65	88.55

V. DISCUSSION

A. Characteristics of the Seizure Detection Method

In this paper, a seizure detection method based on bidirectional GRU recurrent network is proposed, which was tested on the scalp EEG datasets. In the proposed seizure detection method, wavelet transform is performed as a preprocessing of multichannel EEG. In different decomposition scales, wavelet transform has different time-frequency resolution, which overcomes the shortcoming of Short-Time Fourier Transform (STFT) having fixed-size time window and makes it suitable for the analysis of singular signals and nonstationary signals. After carrying out a wavelet decomposition on original EEG, the wavelet coefficients on scale 3, 4 and 5 are applied to signal reconstruction, which can function as a filtering and is helpful for the classification between seizure and non-seizure EEG, because seizure activity commonly focuses on the frequency band of the three scales and the reconstructed signals of seizure EEG in these scales have obvious amplitude and energy difference with those of non-seizure EEG. However, in the propagation and termination phases of epileptic seizures, there are still existing some high-amplitude/low frequency activities that are not in the frequency range of scale 3 to 5. Neglecting these low frequency activities in the seizure detection will produce some false negative results.

In the subsequent feature extraction section, relative energies of the reconstructed signals on wavelet scale 3, 4 and 5 are calculated respectively. Generally, the EEG shows a rhythmic behavior when a seizure occurs, most of its energy thus exhibit in limited scales of the multi-resolution framework, which is different with interictal EEG that often spread across most of the wavelet scales [37]. There are higher relative energy values existing in a particular wavelet scale, which indicates the presence of rhythmic activity and can be used for seizure detection. The values of relative energy calculated from the reconstructed signals on scale 3, 4 and 5 were combined into feature matrixes and inputted into a Bi-GRU neural network.

In the proposed seizure detection method, Bi-GRU network is adopted as classification model, which includes two hidden layers and a dense output layer. Excepting the merits that small calculation cost, fast training speed and needing small amount of training samples, Bi-GRU network is designed in this work because it propagates the information in positive and negative time directions by two hidden layers and can thoroughly capture the long-term dependencies in EEG sequences. The seizure detection based on Bi-GRU network can achieve a better detection results than that based on GRU when they were tested on a same testing dataset. Table VI gives a comparison on the detection results between GRU and Bi-GRU based seizure detection methods. As can be seen in Table VI, the segment-based sensitivity of 93.89 and event-based sensitivity of 95.49 obtained by Bi-GRU based seizure detection method are both higher than those of GRU-based method, which means that Bi-GRU has a better detection performance for seizure activity. Meanwhile, the higher specificity and the lower FDR also indicate that Bi-GRU based detection method has a more accurate classification ability for EEG patterns.

On the other hand, the performance of Bi-GRU based detection method is more stable when system parameters are changed. The dimension of hidden state in GRU cells is a parameter influencing the performance of recurrent neural network and generally without definite selection method. To demonstrate the performance stability of Bi-GRU based

TABLE VI
COMPARISON OF GRU AND BI-GRU BASED DETECTION METHODS

Method	Sensitivity (%)	Specificity (%)	Event-based Sensitivity (%)	FDR (/h)
GRU	88.16	97.38	92.96	0.57
Bi-GRU	93.89	98.49	95.49	0.31

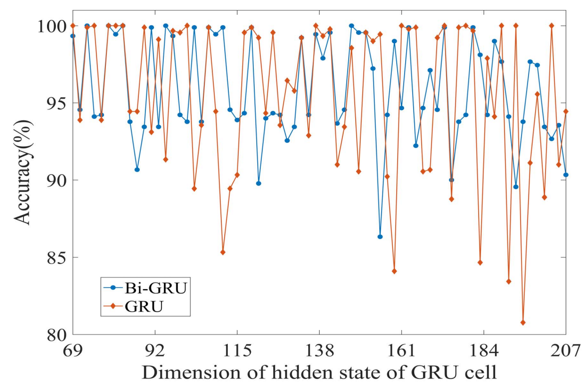

Fig. 7. Comparison of classification accuracy rates between Bi-GRU and GRU based detection systems. (Blue dots present the accuracies of Bi-GRU based detection system corresponding to different hidden state dimensions, and red diamond points are those of GRU based detection system.)

detection system when the hidden state dimension is changed, we tested on an interictal EEG file from case chb13 in CHB-MIT database. The dimension is changed in the range of 69-207, with step length of 2. Here, the lower limit equal to the input vector dimension of each GRU cell in the network of case chb13, and the upper limit is set as three times to the lower limit. Fig. 7 shows the accuracy rates corresponding to different hidden state dimensions for GRU network and Bi-GRU network. In Fig. 7, blue dots present the accuracy rate values of Bi-GRU based detection system and the red diamond points are those of GRU based system. Obviously, the distribution of accuracy rate for the case with Bi-GRU is more concentrated and stable when the hidden state dimension is changed.

On the whole, the time-frequency energy analysis and Bi-GRU network based classification model make the proposed method perform well in seizure detection from scalp EEG. Moreover, a series of post-processing strategies haven been applied to the outputs of Bi-GRU networks, which can further improve the detection performance. In the post-processing procedures, moving average filtering is employed to reduce the false detections caused by random noises and sporadic fluctuations. However, the moving average filter is a finite impulse response (FIR) filter and will produce a time-delay influencing the segment-based evaluation results. It is even possible that those seizures having short duration will be missed detection because of the delay of moving average filtering. It is obvious that a reasonable filter order is required in order to balance its influence on the detection results. In this work, the filter order was determined by plotting ROC curve.

Because the dynamics progress of epileptic seizures vary greatly among different patients and EEG recordings have high heterogeneity in different patients and seizures, majority of the current seizure detection methods are patient-specific and generally fail to produce satisfactory detection results for some patients in different datasets [38]. The seizure detection method proposed in this work also is patient-specific, that is, the established detection systems have patient-dependent training dataset and parameters such as GRU output dimension and threshold etc. In order to further analyze the performance of the proposed seizure detection method, a case-independent verification was also performed in this work. From the point of clinic application, the training and testing data for a detection system were selected from different cases. Using the EEG of the first five cases in CHB-MIT database, we built a training dataset with the total duration of 20.8 min. The trained detection system was tested on the EEG data coming from case chb06 to chb24, more specifically, we selected four EEG files (at least two files containing seizures) from each case to be used as testing data. On the testing dataset that has the total duration of 75.8 h and contains 42 seizures, the case-independent detection system obtained a sensitivity of 58.25%, a specificity of 79.97% and an accuracy rate of 79.79%. Obviously, the results of case-independent verification needed to be further improved from the point of clinic application.

B. Analysis of Missed Detections and False Detections

When the seizure detection method proposed in this work were evaluated on CHB-MIT dataset, there were 4 seizures not be detected. The missed seizures respectively came from case chb16, chb18 and chb21, and the missed detections may be caused by low amplitude and unobvious seizure activity, for example the seizures of case chb18 and chb21. In addition, short duration of seizures and the moving average filtering in post-processing also were the reasons leading to missed detection. The EEG shown in Fig. 8 come from three of channels of case chb16 and contain a missed seizure that is the part in shade of grey. This seizure event only lasts 6 s. Similarly, another missed seizure of case chb16 is illustrated in Fig. 9, where the EEG in the beginning and ending parts of the seizure event present low amplitudes and have little difference with non-seizure EEG, although there is obvious seizure activity in a short time that is about second 14 to 21.

On the other hand, some interictal EEG were incorrectly labeled as seizures by the detection system, especially the EEG files from case chb08 which has the average FDR up to 1.37 per hour. The false detections in interictal EEG were mainly caused by large amplitude epileptiform discharges and artifacts contained in EEG recordings. As an example, Fig. 10 illustrates one of false detections of chb08, that is, the part in shade of grey. It can be seen that there are obvious large amplitude rhythmic activities existing in the channel of T7-P7 and P7-O1.

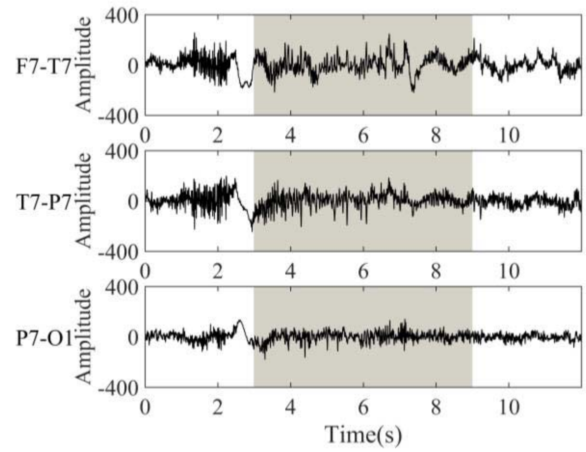

Fig. 8. A missed seizure from case chb16.

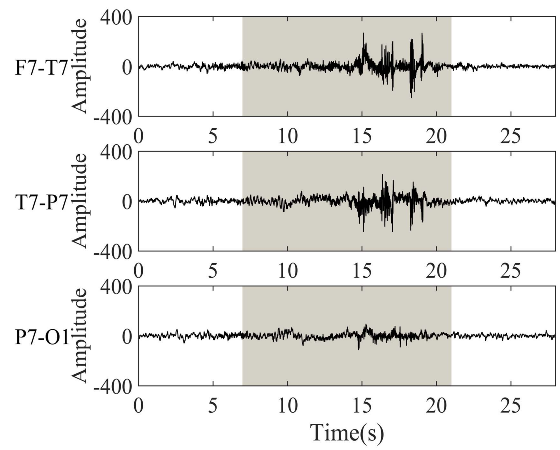

Fig. 9. Another missed seizure from case chb16.

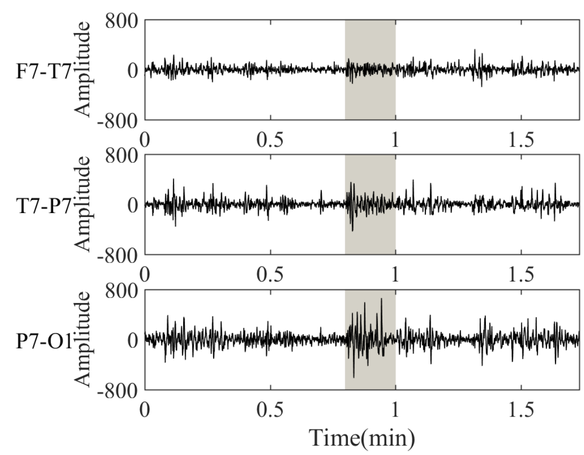

Fig. 10. Example of a false detection.

C. Comparison With Other Seizure Detection Methods

Table VII lists several exiting seizure detection methods for comparison, all of which were evaluated on CHB-MIT EEG database. In the work of Kiranyaz et al, a patient-specific classification system was proposed based on the collective network of binary classifiers, which obtained an average sensitivity of 89.01% along with an average specificity of 94.71 over a testing dataset of 146 h from 21 cases [39]. Zabihi *et al.* investigated the dynamics of EEG signals and proposed a seizure detection approach based on linear discriminate

TABLE VII
PERFORMANCE COMPARISON OF SEVERAL SEIZURE DETECTION METHODS

Authors	Method	Number of Cases	Duration of testing data (h)	Sensitivity (%)	Specificity (%)	Accuracy rate(%)	False detection rate (/h)
Kiranyaz et al.[39]	Collective network of binary classifiers	21	146	89.01	94.71	-	-
Zabihi et al.[40]	Linear discriminate analysis +Naive Bayesian	23	-	89.10	94.80	94.69	-
Samiee et al.[41]	Sparse rational decomposition + Local Gabor binary patterns	23	163	70.4	99.1	83.53	0.35
Liang et al.[42]	CNN+LSTM	24	-	84.00	99.00	99.00	0.2
Hu et al.[43]	Local mean decomposition + Bi-LSTM	24	874.92	93.61	91.85	-	-
Tsiouris et al.[44]	EEG features + LSTM network	24	980	99.38	99.60	-	-
Yang et al.[38]	STFT spectral images + Dual self-attention residual network	13	268.6	89.25	92.67	92.07	-
Wang et al.[45]	Stacked 1D-CNN	24	518.6	88.14	99.62	99.54	0.2
Sahani et al.[46]	Reduced deep convolutional stack autoencoder + Improved kernel random vector functional link network	24	-	100	99.96	99.96	0.0391
Peng et al.[47]	Stein-kernel based sparse representation	20	-	97.85	98.57	98.21	-
Shoka et al.[48]	Channel selection+Ensemble classifier	23	-	100	77.5	89.02	-
This work	Bi-GRU network	24	867.14	93.89	98.49	98.49	0.31

analysis and naive Bayesian classifiers, which achieved the average sensitivity of 89.10% and specificity of 94.80% when the training rate was set as 50% [40]. To address the problem of off-line detection of epileptic seizures, Samiee *et al.* proposed an EEG feature extraction method based on sparse rational decomposition and local Gabor binary patterns, which was assessed on 163 h of EEG recordings and obtained the overall sensitivity of 70.4%, specificity of 99.1%, and false alarms rate of 0.35 per hour [41]. Obviously, the detection method proposed in this work outperformed the above methods and yielded better results on a greater testing dataset.

In Table VII, the detection methods of Liang, Hu and Tsiouris all adopted LSTM recurrent network [42]–[44]. Liang *et al.* constructed a long-term recurrent convolutional network as epilepsy detector by combining CNN and LSTM, and a sensitivity of 84% with a specificity of 99% were achieved when the detector was tested on CHB-MIT database [42]. Hu *et al.* applied bidirectional LSTM network to seizure detection after a local mean decomposition of EEG. Tested on the EEG recordings of 24 cases, the detection method of Hu *et al.* achieved an average sensitivity of 93.61% with the specificity of 91.85% [43]. With the time domain, frequency domain and graph theoretic EEG features as inputs, a two-layer LSTM network structure was designed by Tsiouris *et al.* to address the binary classification problem between preictal and interictal EEG [44]. Compared with these LSTM-based methods, the detection method proposed in this work shows a comparable performance, as can be seen in Table VII. Bi-GRU network has similar properties in capturing the long-term dependencies in EEG as LSTM networks, but with a simpler internal structure and fast training speed.

Yang *et al.* proposed a dual self-attention residual network (RDANet) to recognize the preictal state from long-term EEG recordings, and the RDANet obtained an averaged accuracy of 92.07% on the EEG datasets of 13 selected patients [38]. In the work of Wang *et al.*, a stacked one-dimensional convolutional neural network (1D-CNN) model was proposed for seizure detection, which achieved 88.14% sensitivity and 99.62% specificity on about 518.6 h of EEG dataset [45]. Sahani *et al.* combined a reduced deep convolutional stack autoencoder with an improved kernel random vector functional link network to recognize the epileptic seizure from EEG recordings [46]. Compared with these deep neural networks based detection methods, our Bi-GRU network based detection method has comparable detection performance, but with a simpler architecture and less computational cost.

Using the Stein kernel-based sparse representation, Peng *et al.* proposed a seizure detection method that achieved the sensitivity and specificity of 97.85% and 98.57% respectively [47]. The automated seizure diagnosis system of Shoka *et al.* presented a 100% sensitivity based on an ensemble classifier, but the specificity was only 77.5% [48]. In addition, both the detection methods of Peng and Shoka needed channel selection by computing the variance of training samples. Unlike them, the seizure detection method proposed in this work is insensitive to the number of EEG electrodes, which can be different for different epilepsy patients.

D. Limitations and Future Work

Although the proposed seizure detection method achieved remarkable results on CHB-MIT EEG database, there are limitations in the current work, which need to be further studied. In one aspect, the number of GRU cells of each hidden layer in Bi-GRU neural network is set as 8 in the current method and the time slice corresponding to each GRU cell is 0.5 s. The number of GRU cells and the length of time slices are the parameters influencing how much information participates in the recognition of EEG patterns, thus further comparative experiments should be carried out on the selection of these parameters to better mine the long-term dependencies in EEG sequences. In another aspect, the performance of our seizure detection method are not good enough in the patient-independent verification experiment and in the evaluation to the clinical SH-SDU EEG dataset. As part of our future work, great efforts will be made to research adaptive adjustment method for system parameters to fit the EEG characteristics of different epilepsy patients.

VI. Conclusion

With the purpose of decreasing the workload of EEG monitoring units in the analysis of large EEG recordings and therapeutic evaluation, an automatic seizure detection method is proposed in this paper by combining the relative energy of particular wavelet frequency bands with Bi-GRU neural network. Benefitting from the considerable ability of Bi-GRU network in mining the long-term dependencies containing in EEG recordings in positive and negative time directions, the proposed detection method achieved an average sensitivity of 93.89% and an average specificity of 98.49% on CHB-MIT scalp EEG database. Our detection method have obvious advantages in the requirement for hardware equipment and the amount of calculation, because it needs to extract fewer EEG features and applies a simple structure for Bi-GRU network. What's more, the proposed detection method is insensitive to the number of EEG electrodes and can be applied to establish an effective detection system for specific epilepsy patient. However, improving the detection performance in patient-independent application and on those practical clinical EEG data with more noises and artifacts is an issue worth researching in our future work.

Acknowledgment

The authors would like to thank Guoyang Liu in Shandong University, Zhen Liu and Shang Wang in the Second Hospital of Shandong University, for collecting the EEG data used in this study.

References

[1] S. J. M. Smith, "EEG in the diagnosis, classification, and management of patients with epilepsy," *J. Neurol. Neurosurg. Psychiatry*, vol. 76, no. 2, pp. 2–7, 2005.

[2] S. Noachtar and J. Rémi, "The role of EEG in epilepsy: A critical review," *Epilepsy Behav.*, vol. 15, no. 1, pp. 22–33, Jun. 2009.

[3] Y. Zhang, R. Yang, and W. Zhou, "Roughness-length-based characteristic analysis of intracranial EEG and epileptic seizure prediction," *Int. J. Neural Syst.*, vol. 30, no. 12, Nov. 2020, Art. no. 2050072.

[4] J. Gotman, "Automatic recognition of epileptic seizures in the EEG," *Electroencephalogr. Clin. Neurophysiol.*, vol. 54, no. 5, pp. 530–540, 1982.

[5] S. Grewal and J. Gotman, "An automatic warning system for epileptic seizures recorded on intracerebral EEGs," *Clin. Neurophysiol.*, vol. 116, no. 10, pp. 2460–2472, Oct. 2005.

[6] A. Aarabi, R. Fazel-Rezai, and Y. Aghakhani, "A fuzzy rule-based system for epileptic seizure detection in intracranial EEG," *Clin. Neurophysiol.*, vol. 120, no. 9, pp. 1648–1657, 2009.

[7] Y. Zhang, W. Zhou, and S. Yuan, "Multifractal analysis and relevance vector machine-based automatic seizure detection in intracranial EEG," *Int. J. Neural Syst.*, vol. 25, no. 6, Sep. 2015, Art. no. 1550020.

[8] A. Zarei and B. M. Asl, "Automatic seizure detection using orthogonal matching pursuit, discrete wavelet transform, and entropy based features of EEG signals," *Comput. Biol. Med.*, vol. 131, Apr. 2021, Art. no. 104250.

[9] O. Faust, U. R. Acharya, L. C. Min, and B. H. Sputh, "Automatic identification of epileptic and background EEG signals using frequency domain parameters," *Int. J. Neural Syst.*, vol. 20, no. 2, pp. 159–176, 2010.

[10] F. Riaz, A. Hassan, S. Rehman, I. K. Niazi, and K. Dremstrup, "EMD-based temporal and spectral features for the classification of EEG signals using supervised learning," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 1, pp. 28–35, Jan. 2016.

[11] M. K. Kiymik, İ. Güler, A. Dizibüyük, and M. Akin, "Comparison of STFT and wavelet transform methods in determining epileptic seizure activity in EEG signals for real-time application," *Comput. Biol. Med.*, vol. 35, no. 7, pp. 603–616, Oct. 2005.

[12] U. R. Acharya, S. V. Sree, P. C. A. Ang, R. Yanti, and J. S. Suri, "Application of non-linear and wavelet based features for the automated identification of epileptic EEG signals," *Int. J. Neural Syst.*, vol. 22, no. 2, Apr. 2012, Art. no. 1250002.

[13] N. F. Güler, E. D. Übey, and İ. Güler, "Recurrent neural networks employing Lyapunov exponents for EEG signals classification," *Expert Syst. Appl.*, vol. 29, no. 3, pp. 506–514, 2005.

[14] Y. Zhang, W. Zhou, S. Yuan, and Q. Yuan, "Seizure detection method based on fractal dimension and gradient boosting," *Epilepsy Behav.*, vol. 43, pp. 30–38, Feb. 2015.

[15] V. Srinivasan, C. Eswaran, and N. Sriraam, "Approximate entropy-based epileptic EEG detection using artificial neural networks," *IEEE Trans. Inf. Technol. Biomed.*, vol. 11, no. 3, pp. 288–295, May 2007.

[16] U. R. Acharya, F. Molinari, S. V. Sree, S. Chattopadhyay, K.-H. Ng, and J. S. Suri, "Automated diagnosis of epileptic EEG using entropies," *Biomed. Signal Process.*, vol. 7, no. 4, pp. 401–408, Jul. 2012.

[17] J. Craley, E. Johnson, C. Jouny, and A. Venkataraman, "Automated inter-patient seizure detection using multichannel convolutional and recurrent neural networks," *Biomed. Signal Process. Control*, vol. 64, Feb. 2021, Art. no. 102360.

[18] R. Meier, H. Dittrich, A. Schulze-Bonhage, and A. Aertsen, "Detecting epileptic seizures in long-term human EEG: A new approach to automatic online and real-time detection and classification of polymorphic seizure patterns," *J. Clin. Neurophysio.*, vol. 25, no. 3, pp. 119–131, 2008.

[19] A. Temko, E. Thomas, W. Marnane, G. Lightbody, and G. Boylan, "EEG-based neonatal seizure detection with support vector machines," *Clin. Neurophysiol.*, vol. 122, no. 3, pp. 467–473, 2011.

[20] C. Donos, M. Dümpelmann, and A. Schulze-Bonhage, "Early seizure detection algorithm based on intracranial EEG and random forest classification," *Int. J. Neural Syst.*, vol. 25, no. 5, Aug. 2015, Art. no. 1550023.

[21] O. Kocadagli and R. Langari, "Classification of EEG signals for epileptic seizures using hybrid artificial neural networks based wavelet transforms and fuzzy relations," *Expert Syst. Appl.*, vol. 88, pp. 419–434, Dec. 2017.

[22] Z. Wei, J. Zou, J. Zhang, and J. Xu, "Automatic epileptic EEG detection using convolutional neural network with improvements in time-domain," *Biomed. Signal Process. Control*, vol. 53, Aug. 2019, Art. no. 101551.

[23] M. A. Tanveer, M. J. Khan, H. Sajid, and N. Naseer, "Convolutional neural networks ensemble model for neonatal seizure detection," *J. Neurosci. Methods*, vol. 358, Jul. 2021, Art. no. 109197.

[24] A. Narin, "Detection of focal and non-focal epileptic seizure using continuous wavelet transform-based scalogram images and pre-trained deep neural networks," *IRBM*, to be published, doi: 10.1016/j.irbm.2020.11.002.

[25] M. Geng, W. Zhou, G. Liu, C. Li, and Y. Zhang, "Epileptic seizure detection based on stockwell transform and bidirectional long short-term memory," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 3, pp. 573–580, Mar. 2020.

[26] M. Rashed-Al-Mahfuz, M. A. Moni, S. Uddin, S. A. Alyami, M. A. Summers, and V. Eapen, "A deep convolutional neural network method to detect seizures and characteristic frequencies using epileptic electroencephalogram (EEG) data," *IEEE J. Transl. Eng. Health Med.*, vol. 9, 2021, Art. no. 2000112.

[27] J. Prabin Jose, M. Sundaram, and G. Jaffino, "Adaptive rag-bull rider: A modified self-adaptive optimization algorithm for epileptic seizure detection with deep stacked autoencoder using electroencephalogram," *Biomed. Signal Process. Control*, vol. 64, Feb. 2021, Art. no. 102322.

[28] K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in *Proc. EMNLP*, Doha, Qatar, 2014, pp. 1724–1734.

[29] G. Huang, X. Li, B. Zhang, and J. Ren, "PM2.5 concentration forecasting at surface monitoring sites using GRU neural network based on empirical mode decomposition," *Sci. Total Environ.*, vol. 768, May 2021, Art. no. 144516.

[30] J. Yu, X. Zhang, L. Xu, J. Dong, and L. Zhangzhong, "A hybrid CNN-GRU model for predicting soil moisture in maize root zone," *Agricult. Water Manage.*, vol. 245, Feb. 2021, Art. no. 106649.

[31] G. Ayzel and M. Heistermann, "The effect of calibration data length on the performance of a conceptual hydrological model versus LSTM and GRU: A case study for six basins from the CAMELS dataset," *Comput. Geosci.*, vol. 149, Apr. 2021, Art. no. 104708.

[32] S. Zhang *et al.*, "Modeling pedestrians' near-accident events at signalized intersections using gated recurrent unit (GRU)," *Accident Anal. Prevention*, vol. 148, Dec. 2020, Art. no. 105844.

[33] Y. Ma and H. Li, "GRU-auto-encoder neural network based methods for diagnosing abnormal operating conditions of steam drums in coal gasification plants," *Comput. Chem. Eng.*, vol. 143, Dec. 2020, Art. no. 107097.

[34] B. L. Priya, S. Jayalakshmy, J. K. Pragatheeswaran, D. Saraswathi, and N. Poonguzhali, "Scattering convolutional network based predictive model for cognitive activity of brain using empirical wavelet decomposition," *Biomed. Signal Process. Control*, vol. 66, Apr. 2021, Art. no. 102501.

[35] A. Shoeb, "Application of machine learning to epileptic seizure onset detection and treatment," Ph.D. dissertation, Dept. Harvard-MIT Division Health Sci. Technol., Massachusetts Inst. Technol., Cambridge, MA, USA, 2009.

[36] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. E215–E220, Jun. 2000.

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

[38] X. Yang, J. Zhao, Q. Sun, J. Lu, and X. Ma, "An effective dual self-attention residual network for seizure prediction," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1604–1613, 2021.

[39] S. Kiranyaz, T. Ince, M. Zabihi, and D. Ince, "Automated patient-specific classification of long-term electroencephalography," *J. Biomed. Informat.*, vol. 49, pp. 16–31, Jun. 2014.

[40] M. Zabihi *et al.*, "Analysis of high-dimensional phase space via Poincaré section for patient-specific seizure detection," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 3, pp. 386–398, Mar. 2016.

[41] K. Samiee, P. Kovács, and M. Gabbouj, "Epileptic seizure detection in long-term EEG records using sparse rational decomposition and local Gabor binary patterns feature extraction," *Knowl.-Based Syst.*, vol. 118, pp. 228–240, Feb. 2017.

[42] W. Liang, H. Pei, Q. Cai, and Y. Wang, "Scalp EEG epileptogenic zone recognition and localization based on long-term recurrent convolutional network," *Neurocomputing*, vol. 396, pp. 569–576, Jul. 2020.

[43] X. Hu, S. Yuan, F. Xu, Y. Leng, K. Yuan, and Q. Yuan, "Scalp EEG classification using deep bi-LSTM network for seizure detection," *Comput. Biol. Med.*, vol. 124, Sep. 2020, Art. no. 103919.

[44] K. M. Tsiouris, V. C. Pezoulas, M. Zervakis, S. Konitsiotis, D. D. Koutsouris, and D. I. Fotiadis, "A long short-term memory deep learning network for the prediction of epileptic seizures using EEG signals," *Comput. Biol. Med.*, vol. 99, pp. 24–37, Aug. 2018.

[45] X. Wang, X. Wang, W. Liu, Z. Chang, T. Kärkkäinen, and F. Cong, "One dimensional convolutional neural networks for seizure onset detection using long-term scalp and intracranial EEG," *Neurocomputing*, vol. 459, pp. 212–222, Oct. 2021.

[46] M. Sahani, S. K. Rout, and P. K. Dash, "Epileptic seizure recognition using reduced deep convolutional stack autoencoder and improved kernel RVFLN from EEG signals," *IEEE Trans. Biomed. Circuits Syst.*, vol. 15, no. 3, pp. 595–605, Jun. 2021.

[47] H. Peng *et al.*, "Automatic epileptic seizure detection via stein kernel-based sparse representation," *Comput. Biol. Med.*, vol. 132, May 2021, Art. no. 104338.

[48] A. A. Ein Shoka, M. H. Alkinani, A. S. El-Sherbeny, A. El-Sayed, and M. M. Dessouky, "Automated seizure diagnosis system based on feature extraction and channel selection using EEG signals," *Brain Informat.*, vol. 8, no. 1, Feb. 2021.