Chinese Journal of Electronics
Vol.30, No.3, May 2021

Intracranial Epileptic Seizures Detection Based on an Optimized Neural Network Classifier

GONG Chen[1,2,3], LIU Jiahui[1] and NIU Yunyun[1]

(1. *School of Information Engineering, China University of Geosciences in Beijing, Beijing 100083, China*)

(2. *Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China*)

(3. *School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China*)

Abstract — Automatic identification of intracranial electroencephalogram (iEEG) signals has become more and more important in the field of medical diagnostics. In this paper, an optimized neural network classifier is proposed based on an improved feature extraction method for the identification of iEEG epileptic seizures. Four kinds of entropy, Sample entropy, Approximate entropy, Shannon entropy, Log energy entropy are extracted from the database as the feature vectors of Neural network (NN) during the identification process. Four kinds of classification tasks, namely Pre-ictal v Post-ictal (CD), Pre-ictal v Epileptic (CE), Post-ictal v Epileptic (DE), Pre-ictal v Post-ictal v Epileptic (CDE), are used to test the effect of our classification method. The experimental results show that our algorithm achieves higher performance in all tasks than previous algorithms. The effect of hidden layer nodes number is investigated by a constructive approach named growth method. We obtain the optimized number ranges of hidden layer nodes for the binary classification problems CD, CE, DE, and the multitask classification problem CDE, respectively.

Key words — iEEG, Neural network (NN), Entropy, Feature extraction, Mutual range of coefficient, Hidden layer node.

I. Introduction

Epilepsy is one of the most common neurological diseases in the world. It is caused by the abnormal activity of a large number of nerve cells, which is characterized by the duration and uncertainty of seizures[1,2]. In clinical practice, neuroscientists have to analyze large amounts of intracranial electroencephalogram (iEEG) data to detect epileptoid-related activity. With the development of machine learning algorithms, people try to design automatic seizure detection systems to help experts with this time-consuming and tedious process.

An electroencephalogram (EEG) is a signal that conveys information about the brain through electrical communication[3], which is widely used to assess the disorder of neurons and to determine abnormal brain activity. Typically, EEG is done by placing electrodes on the scalp. According to neuroscientists, the brain waves in the scalp are very sensitive, susceptible to noise, and have low spatial resolution[4]. The best way to solve these problems is to transfer electrodes to the cortex and measure electrical activity in the brain, which is iEEG.

There are four states of EEG signal activities under epilepsy conditions: pre-ictal, epileptic, post-ictal, and normal state. The key step of detecting epileptic seizures is recognition of pre-ictal, post-ictal and epileptic states. In this work, we use NN as a classifier and four entropies, SampEn, ApEn, LogEn and ShanEn, as input vectors to detect epileptic seizures. The main contribution of this work can be summarized as follows.

1) We implement four kinds of entropy as an input vector and use NN as a classifier, in order to detect epileptic seizures. The entropy combination of SampEn, ApEn, LogEn, and ShanEn can help to obtain better classification accuracy than previous work.

2) Mutual range of coefficient γ is used to evaluate the classification effect, which reflects the extraction effect of entropy from the perspective of statistics, and has a certain predictive effect on the classification effect.

3) In this paper, we explore the effect of the number of hidden layer nodes on classification performance. The best range of the number is found through analysis, which

Manuscript Received Sept. 13, 2020; Accepted Jan. 5, 2021. This work is supported by the National Natural Science Foundation of China (No.61872325), the Fundamental Research Funds for the Central Universities (No.2652019028), and the China Scholarship Council (No.201806405004).

 DOI:10.1049/cje.2021.03.005

further improves classification accuracy.

II. Related Work and Background

Many scholars have been striving to study the classification of EEG. Fast Fourier transform (FFT) from the recorded signal pretreatment is used to extract EEG, and signals of the different frequency range of average power are regarded as the feature sets[5]. Yoo *et al.* propose a continuous ultra-low-power expandable retractable EEG to continuously detect and record epilepsy with a Support vector machine (SVM) as a classifier[6]. Gigola *et al.* propose an evolution-based method to predict the cumulative energy of wavelet analysis[7]. Five sets of EEG signals are decomposed with Discrete wavelet transform (DWT) into different sub-bands to acquire detail and approximation coefficients, and feed-forward NN, SVM, decision tree and other methods are used to classify EEG[8]. It has been proved that NN can get the best CA. Wavelet packet decomposition (WPD) is used to decompose the wavelet into wavelet coefficients, and the Principal component analysis (PCA) is used to extract characteristic values and the Gaussian mixture model (GMM) as classification[9].

Many feature extraction methods and classification algorithms based on supervised learning or unsupervised learning have been applied to detect epileptic seizures[8−16]. Entropy is a numerical measure of signal randomness[17]. Srinibasan *et al.* propose Approximate Entropy (ApEn) as input feature, and Elman, Probabilistic Neural Network (PNN) as classifiers[15]. Sample entropy (SampEn), Spectral entropy (SpEn), and Wavelet entropy are extracted. They use Recursive Elman network (REN) and the radial basis network (RBN) as classifier[18]. Besides, wavelet transform (WT) is used to calculate the Relative wavelet energy (RWE), and an Artificial neural network (ANN) is combined to classify[19]. Aydin *et al.* take ShanEn, Log energy entropy (LogEn), and SampEn as an input of the MLNN architecture. Besides, the LogEn provided the most reliable feature for EEG classification in their experiments [20]. What's more, Raghu proposes the ShanEn, LogEn, SpEn and RenEn as features, and multilayer perceptron neural network (MLP) as a classifier. Mainly, finding which training algorithm and activation functions can make the performance better[10].

The learning ability of NN is largely related to the structure of the network hidden layer, especially the number of nodes in the hidden layer. Generally, the three-layer neural network with a single hidden layer can approximate any continuous function. If the number of hidden layer nodes is too small, the performance of the network will be low[21]. Increasing its number can improve the accuracy of identification and make it easier to observe. However, if there are too many hidden layer nodes, the training time will be longer, resulting in an over-training problem[22]. In the research of[23], a constructive approach is proposed to explore the influence of the number on the results, and the optimal number of hidden layer nodes for iris classification is found with simulated annealing [24].

III. Material and Methods

1. Data basement

The data set is from the short-range EEG database of the epilepsy laboratory at the University of Bonn in Germany. It consists of A, B, C, D, and E data sets. We use three sets, including pre-ictal (set C), post-ictal (set D), and epileptic (set E). Each data set included 100 brain electrical signals of 23.6s over a period of time. All the data are recorded by the signal amplifier of 128 channels. After a 12-bit-analog-to-digital conversion, the data are written continuously onto the disk of a data acquisition computer system at a sampling rate of 173.61Hz with bandpass filter settings at 0.53−40Hz (12 dB/octave). Data on pre-ictal iEEG are acquired from the hippocampus of the opposite hemisphere of the brain of five epileptic patients. To collect iEEG during intermittent epileptic seizures from the same 5 patients by measuring the pathogenic area of epilepsy in the interval between seizures. Also, iEEG during epileptic seizures is collected by measuring the pathogenic areas of epilepsy during epileptic seizures[3].

Our method involves the following steps, see Fig.1. First, four kinds of entropy, SampEn, ApEn, ShanEn, and LogEn, are extracted from signals. Then, a mutual range of coefficient γ is calculated to evaluate features. Next, the value of four kinds of entropy used as the input vector of classifiers to classify four tasks namely CD, CE, DE, CDE. Finally, the effect of the number of hidden nodes is analyzed, in order to obtain the optimal configuration of the Neural Network. Four classification tasks are, 1. Task CD: Pre-ictal v Post-ictal; Task CE: Pre-ictal v Epileptic; Task DE: Post-ictal v Epileptic; Task CDE: Pre-ictal v Post-ictal v Epileptic.

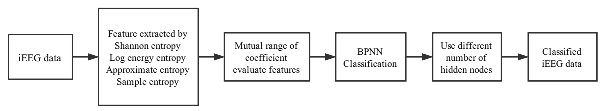

Fig. 1. Block diagram of proposed method

2. Feature extraction

Feature extraction will directly affect the performance of classifiers[25,26]. We use nonlinear feature extraction, because the brain is a chaotic system, whose waves are nonlinear and highly complex. The nonlinear method considers the non-linear system that produces iEEG as definite, and the random iEEG signal produced by the brain is the result of this system. What's more, entropy method is a numerical measure of signal randomness, and good performance has been demonstrated in iEEG experiments on distinguishing between normal and epileptic seizure iEEG with entropy method[20].

The preprocessing of iEEG signals is relatively simple. In this study, the iEEG data are segmented and processed in the experiment. Each piece of data is divided into 8 segments; and each segment has 512 points (denoted as N). After the segmentation, 800 segments data of set C, set D, and set E are obtained.

1) Shannon entropy

In 1948, Shannon put forward the concept of Shannon entropy (ShanEn) and solved the problem of quantitative measurement of information. ShanEn gives the average percentage of information in the signal and uses non-standardized methods to estimate entropy. ShanEn reflects the degree of disorder (ordering) of a system. The more orderly a system, the lower the information entropy, and vice versa[27]. The non-normalization ShanEn is given by[27].

$$\text{ShanEn} = \sum_{i=1}^{m} -p_i^2 * \log(p_i^2) \tag{1}$$

In the above equation, p_i is the frequency of each data in the signal data segment. The logarithm function is based on two. The value of ShanEn of all training data segments is shown in Fig.2.

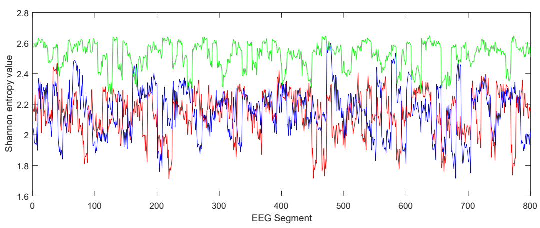

Fig. 2. ShanEn for pre-ictal, post-ictal and epileptic iEEG

2) Log energy entropy

LogEn is a deformation based on ShanEn. Guo *et al.* find that the LogEn provides the most reliable feature for iEEG classification. Therefore, LogEn is considered as a classification feature in this study, which is based on ShanEn with a minute change to Eq.(1) and it is given by[27].

$$\text{LogEn} = \sum_{i=1}^{m} \log(p_i^2) \tag{2}$$

3) Approximate entropy

ApEn is a good criterion for measuring discrete time series. Initially, Pincus *et al.* propose that the ApEn effectively solves short and noisy signals. The ApEn has a strong anti-interference ability and anti-noise ability. Whether it is a random signal, deterministic signal, or the combination of both, the ApEn has strong applicability. The iEEG signal is a chaotic system with strong randomness[15]. The algorithm of ApEn is given below[15,28]:

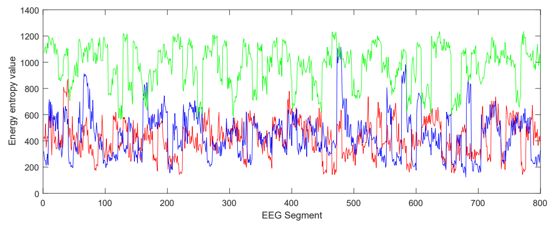

Fig. 3. LogEn for pre-ictal, post-ictal and epileptic

① Data sequence contains N data points: $X = \{x(1), x(2), x(3), \ldots, x(N)\}$.

② $x(i)$ is a subsequence of X. For example, $x(i) = \{x(i), x(i+1), x(i+2), \ldots, x(i+m-1)\}$ and $1 \le i \le N-m$. m represents the number of samples used for the prediction.

③ r represents the noise filter level, which is

$$r = k \times SD \qquad k = 0, 0.1, \ldots, 0.9. \tag{3}$$

SD represtents the standard deviation of the sequence X.

④ $\{x(j)\}$ represents a set of subsequences obtained from $\{x(j)\}$ by varying j from 1 to N. Each sequence $\{x(j)\}$ in the set of $\{x(j)\}$ is compared with $\{x(j)\}$. In this process, $C_i^m(r)$ and $C_i^{m+1}(r)$ are defined as follows,

$$C_i^m(r) = \frac{\sum_{j=1}^{N-m} k_j}{N-m} \tag{4}$$

where

$$k = \begin{cases} 0 & \text{otherwise} \\ 1 & \text{if } |x(i) - x(j)| \le r \text{ for } 1 \le i \le N-m \end{cases} \tag{5}$$

and

$$C_i^{m+1}(r) = \frac{\sum_{j=1}^{N-m} k_j}{N-m}. \tag{6}$$

⑤ $\Phi^m(r)$ and $\Phi^{m+1}(r)$ are defined as follows:

$$\varPhi^m(r) = \frac{\sum_{i=1}^{N-m} \text{In}(C_i^m(r))}{N-m} \tag{7}$$

$$\varPhi^{m+1}(r) = \frac{\sum_{i=1}^{N-m} \text{In}(C_i^{m+1}(r))}{N-m} \tag{8}$$

⑥ ApEn(m, r, N) is calculated using $\varPhi^m(r)$ and

$\Phi^{m+1}(r)$ as follows:

$$\begin{aligned}\mathrm{ApEn}(m,r,N) =&\Phi^m(r)-\Phi^{m+1}(r)\\ =&\frac{1}{N-m}\left[\sum_{i=1}^{N-m}\mathrm{In}\left(\frac{C_i^m(r)}{C_i^{m+1}(r)}\right)\right]\end{aligned} \tag{9}$$

In this study, m equals to 2, and r equals to 0.15SD. m and r are set based on previous studies by Pincus, Srinivasan and others[15,28]. Using the above process, the ApEn of each segments is obtained as shown in Fig.4.

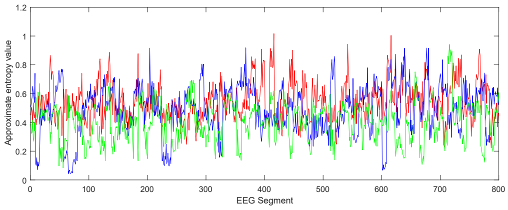

Fig. 4. ApEn for pre-ictal, post-ictal and epileptic iEEG

4) Sample entropy

In 2000, Richman proposed Sample entropy to measure the sequence complexity of nonlinear dynamic systems. By measuring the probability of generating a new pattern in time series, the complexity of EEG signal is obtained, and nonlinear characteristics of EEG signal are described[29]. What's more, SampEn has the advantages of anti-noise and anti-interference at the same time, which avoids the problem of inconsistent statistics caused by the comparison of its own data. The calculation method of SampEn is similar to the process of ApEn, and readers can refer to [29,30] for details.

Usually, m equals to 2 or 3, and r=0.1∼0.25 standard deviation of data segment. In this paper, m=2 and r=0.2 times the standard deviation of the data segment. The sample entropy of the data segment calculated by the above algorithm is shown in the Fig.5.

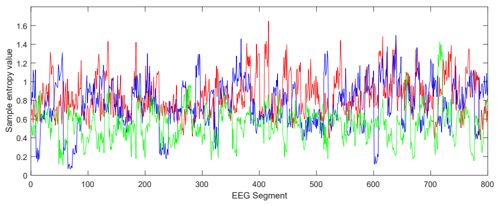

Fig. 5. SampEn for pre-ictal, post-ictal and epileptic iEEG signals

3. Mutual range of coefficient

In order to evaluate the effectiveness of entropy, mutual range of coefficient γ is applied as the evaluation index[31]. The range is defined as the difference between the maximum and minimum values in a segment of data[31]. For example, the ApEn range of dataset C can be defined as[10].

$$\mathrm{Rang}(C)=\max(\mathrm{ApEn})-\min(\mathrm{ApEn}) \tag{10}$$

Similarly, the rang of other three sets of the entropy (ShanEn, SampEn, and LogEn) in dataset C is obtained in the same way. Then the above method is employed to find out the range of data set D and E. γ is defined as the absolute value of the range and the ratio of required data set to their average value[10]. Besides, the following example illustrates the Approximate entropy between dataset C and dataset E.

$$\gamma=\left|\frac{\mathrm{Rang}(C)+\mathrm{Rang}(E)}{\mathrm{Mean}(C)+\mathrm{Mean}(E)}\right| \tag{11}$$

The γ of n data sets defines as follows:

$$\gamma=\left|\frac{\sum_i^{i=n}\mathrm{Rang}(i)}{\sum_i^{i=n}\mathrm{Mean}(i)}\right| \tag{12}$$

where Mean represents the average value of data sets. The smaller the values are, the higher the degree of distinction or proximity overlap and the better frequency band differences between the data sets.

4. Neural network

In this work, the neural network is used as a classifier to detect epileptic seizures. The specific algorithm training process and network performance testing process are listed below.

The activation function of the input layer chooses "prueline", and the activation function of the hidden layer chooses "tansig", which get the best classification effect[10]. The selection of hyperparameters of NN training function: loss function is MSE; the number of epochs (maximum 1000); performance goal (0.001), maximum validation failures (6); the maximum time to train network (infinite). The number of neurons in the hidden layer is obtained according to[23]. Since the number of hidden layers must be greater than 1, the number of hidden layers starts from 1 and increases by 1 after each training. In the same network topology problem studied by[24], the number of hidden layer nodes with optimal classification performance was 11. In our experiment, search conducted within the number of nodes between 1 and 25. For the value of each node number, it runs 10 times and takes the average of classification accuracy. The network is trained with 70% data of the database, and the network performance is tested with 30% data of the database.

IV. Results

In this study, classification accuracy (CA)[15] is used to evaluate the classification performance of neural networks. CA is defined as the ratio of the number of correct data segments classified to the number of test data sets.

1. Features analysis

Figs.2–5 and Fig.6 show entropy values for three different epileptic iEEG conditions. According to previous research results[15], the lower ShanEn is, the more ordered the iEEG signals are. In Fig.2, it can be clearly seen that the entropy value of Shannon pre-ictal and post-ictal are both higher than epileptic iEEG, indicating that the iEEG changes during seizures are more orderly and have less energy. Higher ShanEn obtained for the pre-ictal state, hence it indicates pre-ictal iEEG containing more power in the specified frequency range than the two other states. All figures show that epileptic iEEG produces significantly lower entropy values. In Figs.3–5 and Fig.6, the same analysis can be inferred from LogEn, ApEn and SampEn values. In our experimental result, entropy value present clear discrimination between the other two states and epileptic iEEG signals.

In Table 1, mutual range of coefficient γ between segment of data sets (CD, CE, DE, CDE) is shown.

Table 1. The mutual range of coefficient values between different classification tasks

	CD	CE	DE	CDE
Shannon entropy	1.59	1.36	1.41	1.47
Log energy entropy	1.99	1.00	1.17	1.30
Approximate entropy	2.01	1.75	1.93	1.76
Sample entropy	2.06	1.97	1.84	1.99
sum	7.65	6.25	6.35	6.51

As the proposed study emphasizes more on classification using multi-features, mutual range of coefficient γ of the four entropy combinations of the same classification are added to obtain the sum. Mutual range of coefficient γ of classification task CD, CE, DE and CDE are 7.65, 6.35, 6.25, 6.51. As a result of comparison, we can obtain the size ordering of γ between classified combinations (CD > CDE > DE > CE). Hence, CD data set classification is the most difficult, and CE data set classification is the easiest. Compared with the previous methods, four entropy whose γ is smaller than the result of[10] (where the γ for tasks CD, CE, DE, CDE is 41.38, 9.406, 11.920, 11.977 respectively.), prove that the multi-features is more suitable for extracting features in this problem.

2. The effect of the number of hidden layer nodes

The selection of hidden layer nodes is very complex, which is affected by many factors, including requirements for classification, size of data sets, and so on[24]. The horizontal coordinate represents the number of hidden layer nodes. The vertical coordinate indicates the average CA for four tasks. To avoid the contingency of the experiment, we repeat each experiment ten times, and CA is averaged by 10 experiments with cross validation. The best and worst CA with different numbers of hidden nodes is shown in Table 2.

Tasks CD, CE and DE belong to the binary classification problem. When the number of nodes in the hidden layer is around 5, the classification is better and the complexity of network topology is lower, as shown in Figs.7–9. More precisely, tasks CD, CE, DE can get the best performance when the number of hidden nodes is 6, 7, 5, respectively. The difference between the best and worst classification for task CD, CE, and DE is 13.6%, 7.4% and 7.3%, respectively (see Table 2).

Task CDE belongs to a multi-classification problem. For the classification of CDE data sets, best result is obtained when the number of hidden layer nodes is 14 as shown in Fig.10. The difference between the best and the worst classification is 7.6% (see Table 2).

Table 2. The best and worst CA with different numbers of hidden layer nodes

Classification	The best CA	The worst CA	Difference
CD	66.3%	52.7%	13.6%
CE	99.5%	92.1%	7.4%
DE	98.5%	91.2%	7.3%
CDE	84.6%	77.0%	7.6%

For the binary classification problem CD, CE and DE, it is recommended to set the number of nodes in the hidden layer between 5 and 8. In this case, the CA of task CD is between 63.4% and 66.3%; the CA of task CE is between 98.5% and 99.5%; the CA of task DE is between 96.8% and 98.5%. For multi-classification task CDE, it is recommended to set the number of nodes between 10 and 15. The CA of task CDE is between 80.9% and 84.6%.

3. Analysis of time consumption

In this section, we analyze the time consumption in our research. The experimental results are presented in Table 3. The top figure is the time consumption for the classifier using one or four kinds of entropy as input. The bottom figure is the time consumption for calculating different entropy. In the part of the classification, we find that inputting in one entropy or four entropy has almost no effect on time consumption for neural network classifier. In the part of feature extraction, we record the time consumption for calculating different entropy with all date (the size of data is 800). For ShanEn and LogEn, the computation time can be negligible. The computation of ApEn and SampEn cost relatively long time. For each data, we just need about 0.01s for calculating ApEn and SampEn. Therefore, compared with the method only using one entropy as inputs, our algorithm consumes less time, but gets better performance. Besides, because the size of our data size is small, four kinds of different types of entropy in our research do not add much computational cost. What's more, failure in medical research is costly. Even with some additional computational costs, the increased accuracy is worth it.

Compared with previous algorithms, our algorithm shows obvious superiority when dealing with all the four tasks CD, CE, DE and CDE, see Table 4. Obviously, our method obtained the best results for all four tasks.

Table 3. The time consumption in our research

	one entropy	four entropy
CD	20.7s	21.6s
CE	18.9s	19.0s
DE	19.1s	19.4s
CDE	28.3s	29.4s

	ShanEn	LogEn	ApEn	SampEn
C	0.393 s	0.394 s	8.83 s	7.74 s
D	0.422 s	0.421 s	8.86 s	7.85 s
E	0.517 s	0.505 s	9.20 s	8.02 s

Table 4. Comparison of proposed study with other techniques

Researcher	Year	Method	Task	CA(%)
N. Nicolaou[13] J. Georgiou	2012	Permutation entropy-SVM	CE	88.00
			DE	79.94
Y. Kumar *et al.* [8]	2012	DWT based on Relative wavelet energy Wavelet entropy-SVM	CE	97.50
			DE	97.50
Y. Kumar *et al.* [14]	2014	DWT based on ApEn Feed forward BP network	CE	98.00
			DE	94.00
A. Sharmila [32] P. Geethanjali	2016	DWT- Naive Bayes classifer k-NN classifiers	CE	99.62
			DE	95.62
S. Raghu *et al.* [10]	2017	Shannon, Log energy Spectral and Renyi entropy using MLP	CE	97.68
			DE	94.56
			CDE	84.58
			CD	57.80
Present reporting	2020	Shannon, Log energy Approximate entropy and Sample entropy using NN	CE	99.5
			DE	98.5
			CD	66.3
			CDE	84.6

V. Conclusions

In this study, an optimized neural network classifier is proposed to detect intracranial epileptic seizures. Four kinds of entropy, namely, SampEn, ApEn, ShanEn, LogEn are extracted to train neural networks for four classification tasks, namely, CD, CE, DE, and CDE. Mutual range of coefficient γ is used for proving the validity of the entropy method in feature extraction. Compared with previous research[10], the multi-features in our work are more suitable for extracting features in epileptic iEEG signals detecting. Then, the growth method is used to explore the effect of the number of hidden layer nodes on classification performance. We obtain the number of hidden layer nodes which can achieve better performance for different tasks.

According to the previous research, ApEn and SpEn can effectively solve short and noisy signals. Besides, ApEn evaluation system generally requires no more than 1000 points, which is suitable for our question. As the performance of classification method varies with different kinds of entropy, we will investigate the effects of some new entropy, such as reverse dispersion entropy, on classification task in future research.

References

[1] Fisher. R. S, Boas. W. V. E, Blume. W, *et al.*, "Epileptic seizures and epilepsy: Definitions proposed by the International league against epilepsy (ILAE) and the International Bureau for Epilepsy (IBE)", *Epilepsia*, Vol.46, No.4, pp.470–472, 2005.

[2] Duncan. J. S, Sander. J. W, Sisodiya. S. M, *et al.*, "Adult epilepsy", *The Lancet*, Vol.367, No.9516, pp.1087–1100, 2006.

[3] Andrzejak. R. G, Lehnertz. K, Mormann. F, *et al.*, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state", *Physical Review E*, Vol.64, No.6, pp.061907, 2001.

[4] Janjarasjitt. S, "Spectral exponent characteristics of intracranial EEGs for epileptic seizure classification", *IRBM*, Vol.36, No.1, pp.33–39, 2015.

[5] Theeranaew. W, McDonald. J, Zonjy. B, *et al.*, "Automated detection of postictal generalized EEG suppression", *IEEE Transactions on Biomedical Engineering*, Vol.65, No.2, pp.371–377, 2017.

[6] Yoo. J, Yan. L, El-Damak. D, *et al.*, "An 8-channel scalable EEG acquisition SoC with patient-specific seizure classification and recording processor", *IEEE Journal of Solid-state Circuits*, Vol.48, No.1, pp.214–228, 2012.

[7] Gigola. S, Ortiz. F, D'attellis. C. E, *et al.*, "Prediction of epileptic seizures using accumulated energy in a multiresolution framework", *Journal of Neuroscience Methods*, Vol.138, No.1-2, pp.107–111, 2004.

[8] Kumar. Y, Dewal. M. L and Anand. R. S, "Relative wavelet energy and wavelet entropy based epileptic brain signals classification", *Biomedical Engineering Letters*, Vol.2, No.3, pp.147–157, 2012.

[9] Acharya. U. R, Sree. S. V, Alvin. A. P. C, *et al.*, "Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework", *Expert Systems with Applications*, Vol.39, No.10, pp.9072–9078, 2012.

[10] Raghu. S and Sriraam. N, "Optimal configuration of multilayer perceptron neural network classifier for recognition of intracranial epileptic seizures", *Expert Systems With Applications*, Vol.89, pp.1087–1100, 2006.

[11] Gao. J. F, Hui. S. I, Bin. Y. U, *et al.*, "Lie detection analysis based on the sample entropy of eeg", *Acta Electronica Sinica*, Vol.45, No.8, pp.1836–1841, 2017. (in Chinese)

[12] Huang. J. R, Fan. S. Z, Abbod. M, *et al.*, "Application of multivariate empirical mode decomposition and sample entropy in EEG signals via artificial neural networks for interpreting depth of anesthesia", *Entropy*, Vol.15, No.9, pp.3325–3339, 2013.

[13] Nicolaou. N. and Georgiou. J., "Detection of epileptic electroencephalogram based on permutation entropy and support vector machines", *Expert Systems with Applications*, Vol.39, No.1, pp.202–209, 2012.

[14] Kumar. Y., Dewal. M. L., and. R. S., "Epileptic seizures detection in EEG using DWT-based ApEn and artificial neural network", *Signal, Image and Video Processing*, Vol.8, No.7, pp.1323–1334, 2014.

[15] Srinivasan. V., Eswaran. C., Sriraam. N., *et al.*, "Approximate entropy-based epileptic EEG detection using artificial neural networks", *IEEE Transactions on Information Technology in*

Biomedicine, Vol.11, No.3, pp.288–295, 2007.
[16] YIN Yi and SHANG Pengjian, “Multivariate multiscale sample entropy of traffic time series”, *Nonlinear Dynamics*, Vol.86, No.1, pp.479–488, 2016.
[17] Kannathal. N, Min. L. C, *et al.*, “Entropies for detection of epilepsy in EEG”, *Computer methods and programs in biomedicine*, Vol.80, No.3, pp.187–194, 2005.
[18] Kumar. S. P, Sriraam. N, Benakop. P. G, *et al.*, “Entropies based detection of epileptic seizures with artificial neural network classifiers”, *Expert Systems with Applications*, Vol.37, No.4, pp.3284–3291, 2010.
[19] Ling. G, Rivero. D, Seoane. J. A, *et al.*, “Classification of EEG signals using relative wavelet energy and artificial neural networks”, *Proceedings of the First ACM/SIGEVO Summit on Genetic and Evolutionary Computation*, Shanghai, China, pp.177–184, 2009.
[20] Aydın. S, Saraoğlu, *et al.*, “Log energy entropy-based EEG classification with multilayer neural networks in seizure”, *Annals of Biomedical Engineering*, Vol.37, No.12, pp.2626, 2009.
[21] Curteanu. S and Cartwright. H, “Neural networks applied in chemistry. I. Determination of the optimal topology of multilayer perceptron neural networks”, *Journal of Chemometrics*, Vol.25, No.10, pp.527–549, 2011.
[22] Sun. J, “Learning algorithm and hidden node selection scheme for local coupled feedforward neural network classifier”, *Neurocomputing*, Vol.79,pp.158–163, 2012.
[23] Islam. M. M, Yao. X and Murase. K, “A constructive algorithm for training cooperative neural network ensembles”, *IEEE Transactions on Neural Networks*, Vol.14, No.4, pp.820–834, 2003.
[24] ZHANG Shirui and LI Xinke, “An estimation algorithm of BP neural network hidden layer nodes based on simulated annealing”, *Journal of Hefei University of Technology (Natural Science)*, No.11, pp.10, 2017. (in Chinese).
[25] ZHANG Xuejun, HUO Yan and WAN Dongsheng, “Improved EMD based on piecewise cubic hermite interpolation and mirror extension”, *Chinese Journal of Electronics*, Vol.29, No.5, pp.899–905, 2020.
[26] WANG Danyang and SHAO Fangming, “Research of neural network structural optimization based on information entropy”, *Chinese Journal of Electronics*, Vol.29, No.4, pp.632–638, 2020.
[27] Shannon. C. E, “A mathematical theory of communication”, *Bell System Technical Journal*, Vol.27, No.3, pp.379–423, 1948.
[28] Pincu. S, “Approximate entropy (ApEn) as a complexity measure”, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, Vol.5, No.1, pp.110–117, 1995.
[29] Richman. J. S and Moorman. J. R, “Physiological time-series analysis using approximate entropy and sample entropy”, *American Journal of Physiology-Heart and Circulatory Physiology*, Vol.278, No.6, pp.H2039–H2049, 2000.
[30] Yao. W, Hu. H, Wang. J, *et al.*, “Multiscale ApEn and SampEn in quantifying nonlinear complexity of depressed MEG”, *Chinese Journal of Electronics*, Vol.28, No.4, pp.817–821, 2019.
[31] Woodbury. George, “An introduction to statistics”, *Cengage Learning*, 2001.
[32] Sharmila. A and Geethanjali. P, “DWT based detection of epileptic seizure from EEG signals using naive Bayes and k-NN classifiers”, *IEEE Access*, Vol.4, pp.7716–7727, 2016.

GONG Chen was born in 1998. He received the B.E. degree in computer science and technology at China University of Geosciences (Beijing) in 2020. He now is a master student at Institute of Automation, Chinese Academy of Sciences. His research interests include machine learning and reinforcement learning. (Email: ChenG_abc@outlook.com)

LIU Jiahui was born in 1998. He received the B.E. degree in computer science and technology at China University of Geosciences (Beijing) in 2020. His research interests include machine learning and deep learning. (Email: JHLiu_2018@outlook.com)

NIU Yunyun (corresponding author) was born in 1983. She received the Ph.D. degree in Huazhong University of Science and Technology. She is an associate professor in School of Information Engineering, China University of Geosciences (Beijing). Her research interests include intelligent algorithms and machine learning. (Email: yniu@cugb.edu,cn)