ISSN: 2688-8645

Feature Extraction Method of Epileptic EEG Signal based on Wavelet Packet and Improved Fuzzy Entropy

Dehua Zhang[1], Tong Wang[2]

[1]School of Information Science and Engineering, Shenyang University of Technology, Shenyang, China

[2]Department of Electrical and Information Engineering, Dalian Jiaotong University, Dalian, China

Abstract

Epileptic eeg signal has obvious characteristic information, which can be used as an important basis to judge whether epileptic seizure occurs. Because of the low recognition rate of single feature extraction method, a method of eeg feature extraction based on wavelet packet transform and improved fuzzy entropy was proposed. In view of the characteristics of eeg signal with large noise and weak signal, the Wavelet packet Transform (WPT) is used to decompose the EEG signal with multi-resolution and make it into the signal with different characteristics. The original Fuzzy entropy (Fuzzy EN) algorithm was improved to improve its ability of reflecting the degree of irregularity and complexity of time series. Finally, the feature extraction of epileptic EEG signal was completed by combining the wavelet packet transform method.

Keywords

EEG; Wavelet Packet Transform; Improved Fuzzy Entropy; Feature Extraction.

1. Introduction

Epilepsy is a disease caused by the sudden discharge of neurons in the cerebral aneurysm. The usual clinical manifestations are sudden convulsions, the seizure duration is short, and it is accompanied by abnormal phenomena such as consciousness, sensation, and spirit. Severe cases can be life-threatening [1-3]. Since epilepsy is more common in daily life, and the number of patients is increasing year by year, the treatment of epilepsy also has great significance.

EEG activity is generated by neurons in the cerebral cortex and has the characteristics of spontaneous and rhythmic point changes. The sum of the electrical activity of the local neurons recorded from the intracranial or extracranial scalp is called electroencephalogram (EEG). The analysis and discrimination of the EEG can identify and classify the pre, middle and late stages of epilepsy. , is of great significance to the treatment of diseases. The amount of raw EEG signal data is huge, and manual identification has high requirements for doctors. It requires doctors to have rich clinical experience and consumes a lot of time and energy. Therefore, using machine learning methods to identify and classify EEG signals can not only liberate manpower, improve Efficiency, and can well avoid manual misjudgments and missed judgments. To improve the recognition accuracy and efficiency of machine learning, feature extraction of the original EEG signal becomes a necessary process, and the quality of feature extraction will directly affect the accuracy and efficiency of subsequent machine learning recognition.

Scholars at home and abroad have also conducted a lot of research on such problems, and proposed automatic classification and analysis models of various types of epilepsy EEG signals. The main method is to extract features by traditional signal processing methods, such as time-frequency analysis. method, time domain analysis method, frequency domain analysis method, nonlinear analysis method, etc., each type of analysis method includes a variety of methods.

Among them, the time domain analysis method: for the continuous EEG signal, the characteristics of the EEG signal are directly extracted from the time domain for analysis, which can effectively extract the sharp waves and spike waves in the epilepsy signal, and conduct a preliminary diagnosis of the EEG signal. Because EEG signals have obvious rhythm characteristics and waveform characteristics, they can be used for feature analysis, and they are still widely used today. The current popular time domain analysis method is based on the autoregressive model (AR) method, which uses the AR model for secondary feature extraction, and uses random forest for classification, reaching a classification accuracy of 97.352%. However, since the epilepsy EEG signals are dominated by slow waves, the time domain analysis method cannot be used for classification work, and the lack of signal characteristics in the frequency domain will lead to the inability to intuitively reflect the brain situation, resulting in the limitation of time domain analysis.

Frequency domain analysis method: It is based on the assumption that the EEG signal has a stationary assumption, and the various frequency components contained in the EEG signal are transformed from the time domain to the frequency domain, and the frequency information is analyzed and feature extraction. Compared with the epilepsy EEG signal in the interictal period, there will be obvious spike waves, sharp waves and other characteristic waves during the seizure period, and the signal frequency will change significantly. The commonly used frequency domain analysis method is mainly power spectrum analysis, but due to the randomness and non-stationarity of EEG signals, it is greatly limited in frequency domain analysis in China.

Time-frequency analysis method: In view of the fact that neither time-domain analysis nor frequency-domain analysis can fully reflect the characteristic information, the method of combining time-frequency analysis has become an inevitable development. It has a very good effect in processing non-stationary time-varying signals, and is now the main EEG signal analysis method. Foreign scholars have used discrete wavelet transform to extract features of each frequency sub-band, and used SVM classifier to identify brain diseases, which proves that this method has higher specificity and sensitivity. Ahmadi et al. proposed an epileptic EEG classification model based on wavelet packet transform and SVM.

Nonlinear analysis method and other methods: In view of the nonlinear characteristics of EEG signals, nonlinear dynamics are widely used in the field of EEG signal processing. Common ones include: correlation dimension, Lyapunov exponent, approximate entropy, sample entropy and other nonlinearities Kinetic indicators. Zhang Jianli et al. used wavelet packet decomposition combined with sample entropy for feature extraction, and used Real AdaBoost algorithm and error correction coding for classification, so that the average recognition rate reached 96.78%. In addition, other types of classification methods are also used in EEG signal extraction. Good results have been achieved.

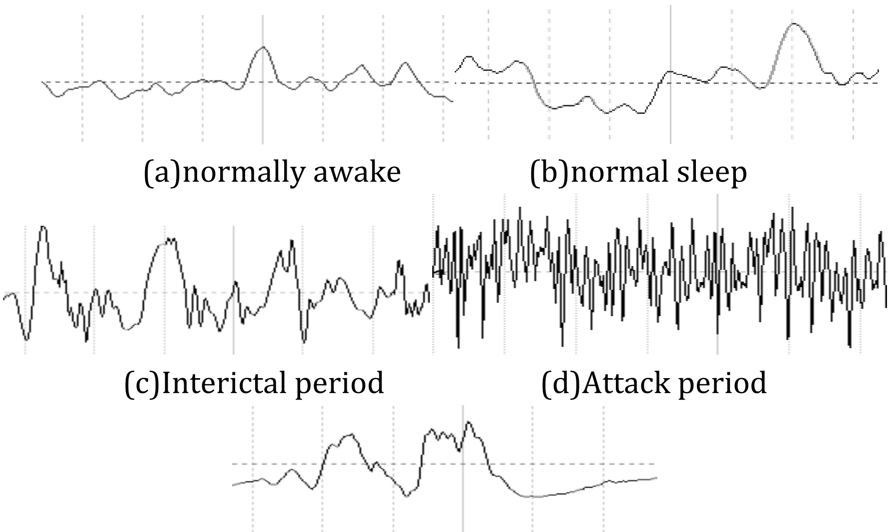
In this paper, the wavelet packet transform method in the time-frequency analysis method is used. First, the EEG signal is decomposed in multiple layers to obtain several groups of different frequency signals with features. to improve the classification accuracy. The fuzzy entropy is similar to the sample entropy and the approximate entropy. It is relatively smooth when the entropy value transitions and changes, and the selection range of its parameters is also large, and it has better robustness to noise. We use the wavelet packet transform to obtain the characteristic waveform in the sub-band, obtain the entropy value by the method of improving the fuzzy entropy, and use the size of the entropy value to reflect the feature extraction result of the signal.

ISSN: 2688-8645

2. Epilepsy EEG

The EEG signal contains a lot of information, including the pathological and physiological aspects of the human body, and its three elements are phase, amplitude, and frequency. The horizontal axis of the signal represents time, and the vertical axis represents the change in potential. EEG has an irreplaceable position in the field of brain research and is one of the main methods for doctors to treat epilepsy. However, the original EEG signal is very weak, and the noise interference is large, accompanied by strong randomness, and has the characteristics of nonlinear and non-stationary. The EEG signal waveform has obvious characteristics during the onset of epilepsy, which can be divided into sharp waves, spike waves, sharp-slow complexes, and spike-slow complexes.

The data recorded by 24 leads are used in this paper, including the whole process data from the normal state to the end of the seizure. After extraction, it can be divided into five categories, namely the wakeful signal and sleep signal in the normal state, the interictal signal, the seizure signal The mid- and seizure-end signals are shown in Fig. 1(a)(b)(c)(d)(e).



(a)normally awake (b)normal sleep

(c)Interictal period (d)Attack period

(e)episode ends

Fig 1. Five types of clinical process signal waveforms

When the functional organization of the human brain is damaged, the EEG will change significantly. Compared with the normal state of EEG signals, there will be certain changes in the interictal and epileptic periods.

3. Wavelet Packet Transform and Improved Fuzzy Entropy

3.1. Principle of Wavelet Packet Transform

Wavelet packet decomposition is a generalization of wavelet decomposition. It has the ability to characterize local characteristics of signals in both time and scale domains, and can be applied to transient and time-frequency characteristics of non-stationary EEG signals. In the process of wavelet decomposition, the original signal is decomposed into a detail part and an approximation part, and the approximation part is decomposed into the next layer of detail part and approximation part again, and this operation is repeated until the set layer is reached. Compared with the wavelet transform, the wavelet packet method is more refined, and can take into account the decomposition of low-frequency and high-frequency parts at the same time, and can adapt to the signal characteristics when selecting the frequency band. The wavelet packet method can retain effective time-frequency information when reconstructing and filtering the EEG signal. In this paper, the wavelet packet method is used to decompose the EEG signal. In order to improve the resolution, it will be decomposed in a binary manner, and its spatial structure is shown in Figure 2 [8~10].

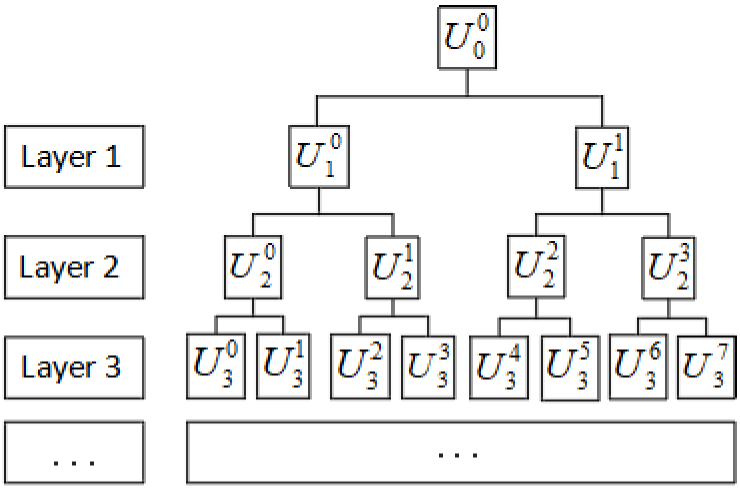


Fig 2. WPT spatial structure

in the text, U_l^τ Expressed as l layer z wavelet packet subspace, in $z=1,2,3\cdots,2^l-1$, the orthonormal basis corresponding to the space is:

$$U_{l,k}^{\tau}(t)=2^{-l/2}u^{\tau}(2^{\tau}t-k) \tag{1}$$

in the formula k is expressed as a translation factor and satisfies the equation:

$$u_{l,0}^{\tau}(t)=\sum_k g_0(k)\cdot u_{l-1,k}^{\tau}(v\text{ is even}) \tag{2}$$

$$u_{l,0}^{\tau}(t)=\sum_k g_1(k)\cdot u_{l-1,k}^{\tau}(v\text{ is odd}) \tag{3}$$

in, $l,k\in Z, z=1,2,3,\cdots,2^l-1$,high pass filter $g_1(k)$ and low pass filter $g_0(k)$ are mutually orthogonal filters, and satisfy:

$$g_1(k)=(-1)^{1-k}g_0(1-k) \tag{4}$$

Use this method to $x(t)$ break down, the original signal can be divided into different wavelet packet subspaces according to different frequency bands, Each frequency band corresponds to the l layer subspace:

$$\{[0,\frac{f_s}{2^{l+1}}];[\frac{f_s}{2^{l+1}},\frac{2f_s}{2^{l+1}}];[\frac{2f_s}{2^{l+1}},\frac{3f_s}{2^{l+1}}];\cdots[\frac{(2l-1)f_s}{2^{l+1}},\frac{f_s}{2}];\} \tag{5}$$

in the formula f_s represents the sample rate of the signal, the first k point wavelet packet coefficients and $l+1$ The reconstruction formula of the layer is:

$$d_{l+1}^{\tau}(k)=\sum_{\varphi}g_0(\varphi-2k)d_l^{2\tau}(\varphi)+\sum_{\varphi}g_1(\varphi-2k)d_l^{2\tau+1}(\varphi) \tag{6}$$

ISSN: 2688-8645

In order to obtain the signal of a specific frequency band, it is necessary to select the corresponding wavelet decomposition coefficient according to the frequency domain characteristics of the extracted signal, and then use the formula (6) to reconstruct. The above operations enable the EEG signal to be further decomposed in both high and low frequency parts, and after the transformation, the amount of information is relatively intact and the loss is small.

3.2. Improved Fuzzy Entropy Principle

Fuzzy entropy is a nonlinear time series analysis method, which is a method of optimizing and improving approximate entropy. The vector similarity is measured under the premise of fuzzy membership function, and the exponential function is used to ensure the validity and continuity of fuzzy entropy [11]. Its advantages are that it can make a smooth transition with parameter changes, and it can still be meaningful in the case of small parameter values. It also has the characteristics of short data set processing and relatively consistent sample entropy characteristics. In order to make the fuzzy entropy have better robustness, it is necessary to choose an appropriate r , m value, Typically, it is recommended to r The value of is set between 0.1 and 0.2, because the threshold cannot be obtained r The real complexity brought about by ignoring the changes in the time series of each time period, deliberately r The value is divided into 10 segments between 0.1 and 0.2, and finally the r The value of is determined at 0.16 times the standard deviation of the first-order difference time series. After calculation, the data tends to be stable, and finally a new series that can reflect the original data is obtained, thus realizing the correlation between time series and complexity. The algorithm is as follows [12~14]:

First put the sequence x_i written in order as m dimensional vector, such as (7):

$$X\left(i\right)=\left[x\left(i\right),x\left(i+1\right),\cdots,x\left(i+m-1\right)\right] \tag{7}$$

in $i=1,2,\cdots,N-m+1$

definition $X\left(i\right)$ and $X\left(j\right)$ The distance between the two, such as (8):

$$d_{ij}[X\left(i\right),X\left(j\right)]=\max\left|x(i+w)-x(j+w)\right| \tag{8}$$

in $w=1,2\cdots,m-1$

definition $X\left(i\right)$ and $X\left(j\right)$ The similarity between the two, such as formula (9):

$$D_{ij}^{m}=\exp\left(-\frac{(d_{ij}^{m})^{n}}{r}\right) \tag{9}$$

in r and n represent the width and gradient of the exponential function boundary, respectively. definition r The values are shown in equations (10)~(12):

$$x(oi)=[x(i+1)-x(i)] \tag{10}$$

in $i=1,2,\cdots,N$

ISSN: 2688-8645

$$x^* = \frac{1}{N}\sum_{i=1}^{N} x(oi) \tag{11}$$

$$r = 0.16\left(\frac{1}{N}\sum_{i=1}^{N}(x(oi)-x^*)^2\right)^{\frac{1}{2}} \tag{12}$$

for each i Calculate the average value, such as formula (12):

$$\mu_m(r) = \frac{1}{N-m}\sum_{i=1}^{N-m}\left(\frac{1}{N-m-1}\sum_{j=1}^{N-m} D_{ij}^m\right) \tag{13}$$

will finally m become $m+1$ Repeat the above operations to obtain fuzzy entropy, as in (14):

$$F(m,r,N) = \ln\frac{\mu_{m+1}(r)}{\mu_m(r)} \tag{14}$$

4. Experimental Verification

The experimental data comes from the original EEG signal of the real patient, which is divided into two groups of experiments without epilepsy-like (control group) and with eclampsia-like (experimental group). frequency bands, as shown in Figure 3:

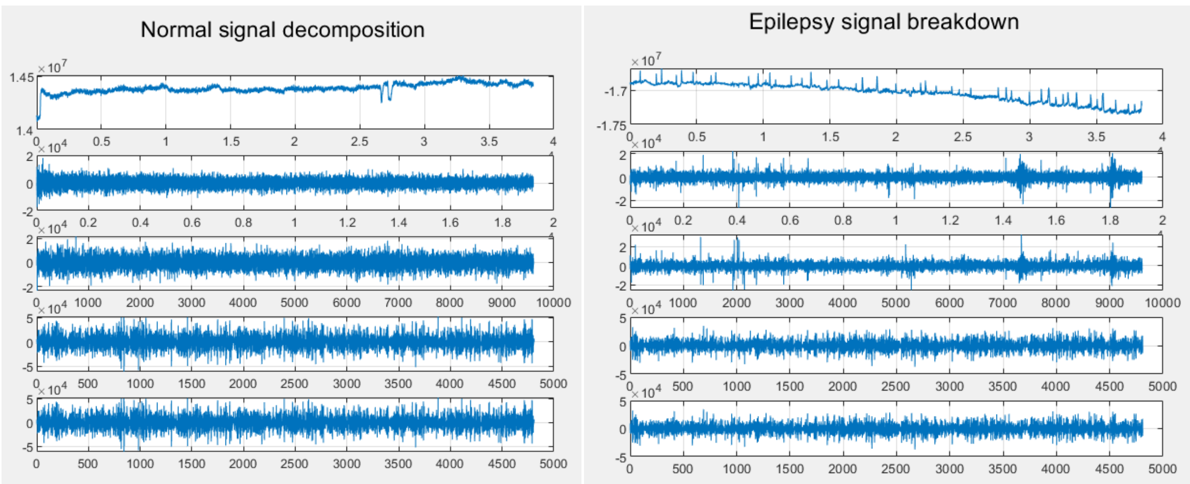


Fig 3. Decomposition of wavelet packet with or without epilepsy

After comparing the two data in the above figure, it is found that the epilepsy-like experimental group has significantly different characteristics than the non-epilepsy-like control group, and it has more characteristic information, which is conducive to the identification and classification of subsequent machine learning. After this operation is completed, the sub-bands with obvious characteristic information in the experimental group are selected. We select the third sub-band as the follow-up experimental object, and then use the method of improving fuzzy entropy to calculate the entropy value of Figure 3. In the algorithm, select = 0.16*std, m=2 to get the fuzzy entropy values [15~16] of the two experiments, as shown in Figure 4:

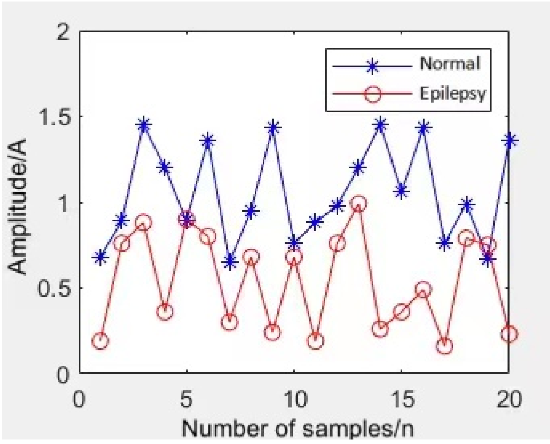


Fig 4. Fuzzy entropy value of EEG signal in two states

Through experiments, it was found that the entropy value of the normal non-eclampsia-like EEG signal sub-bands and the onset eclamptic-like EEG signal sub-bands have a clear degree of distinction, which can fully demonstrate the characteristic difference of the signals between the two. In this way, it is proved that it is feasible to first decompose the wavelet packet, and then use the method of improving the fuzzy entropy to obtain the entropy value of the sub-band, which also lays a solid foundation for the subsequent machine learning of epilepsy EEG signal recognition and classification.

5. Epilogue

During the seizure of epilepsy, the EEG signal will change significantly, which is significantly different from the EEG signal in the normal state, which is also the premise for our study of feature extraction. In this paper, the method of wavelet packet decomposition and improved fuzzy entropy is used to extract the characteristic information of epilepsy EEG signals. After experiments and analysis of clinical experimental data, it is also proved that this method can basically extract the characteristic information of epilepsy EEG signals. This also provides a certain reference value for the follow-up prevention and treatment of epilepsy diseases, and also has certain help for other signal classification and identification tasks.

References

[1] Tran M Q,Elsisi M,Liu M K. Effective feature selection with fuzzy entropy and similarity classifier for chatter vibration diagnosis[J]. Measurement,2021,184.

[2] Zhong Hongyan.Research on Feature Extraction and Classification of EEG Signals[D].Beijing University of Technology,2014.

[3] Lu Xiao J, Zhang J Q,Huang S F,Lu J et al. Detection and classification of epileptic EEG signals by the methods of nonlinear dynamics[J]. Chaos, Solitons and Fractals: the interdisciplinary journal of Nonlinear Science, and Nonequilibrium and Complex Phenomena,2021,151.

[4] Zhang Zhimin. Research and application of entropy measurement and compressed sensing theory in sleep quality assessment [D]. Shandong University, 2020.

[5] Xu Dongping, Chen Feng. Research on EEG information based on improved wavelet transform and fuzzy entropy [J]. Computer Simulation, 2019,36(10):227-232.

[6] Wang M Y,Sheng X Z. Combining empirical wavelet transform and transfer matrix or modal superposition to reconstruct responses of structures subject to typical excitations[J]. Mechanical Systems and Signal Processing,2022,163.

[7] Li Y B,Wang S,Yang Y et al. Multiscale symbolic fuzzy entropy: An entropy denoising method for weak feature extraction of rotating machinery[J]. Mechanical Systems and Signal Processing, 2022, 162.

[8] Meng Qingfang, Zhou Weidong, Chen Yuehui, etc. Feature extraction method of epilepsy EEG signal based on nonlinear prediction effect[J]. Acta Physica Sinica, 2010,59(01):123-130.

[9] Song Huanrong. Extraction and recognition of epilepsy EEG signals[D]. Dalian University of Technology, 2012.

[10] Li K S,Wang Y. A Fuzzy Adaptive Firefly Algorithm for Multilevel Color Image Thresholding Based on Fuzzy Entropy[J]. International Journal of Cognitive Informatics and Natural Intelligence (IJCINI), 2021,15(4).

[11] Akbari H,Sadiq M T,Rehman A U. Classification of normal and depressed EEG signals based on centered correntropy of rhythms in empirical wavelet transform domain.[J]. Health information science and systems,2021,9(1).

[12] Zhang Yuting. Research on EEG Signal Classification Method Based on Wavelet Packet Decomposition [D]. Liaoning Normal University, 2015.

[13] Sun Xiaoqi, Li Xin, Cai Erjuan, etc. Improved fuzzy entropy algorithm and its application in EEG analysis of children with autism [J]. Journal of Automation, 2018, 44(09): 1672-1678.

[14] Cura O K,Akan A. Classification of Epileptic EEG Signals Using Synchrosqueezing Transform and Machine Learning[J]. International Journal of Neural Systems,2021,31(05).

[15] Xu Yonghong, Li Xingxing, Zhao Yong.Classification method of epilepsy EEG signal based on wavelet packet and multivariate multiscale entropy[J].Journal of Biomedical Engineering, 2013, 30 (05): 1073-1078+1090.

[16] Ibrahim A,Chang G L. Selection of optimal wavelet features for epileptic EEG signal classification with LSTM[J]. Neural Computing and Applications,2021(prepublish).

[17] Karabiber C O,Akan A. Analysis of epileptic EEG signals by using dynamic mode decomposition and spectrum[J]. Biocybernetics and Biomedical Engineering,2021,41(1).