

doi:10.3969/j.issn.1673-5374.2013.20.003 [http://www.nrronline.org; http://www.sjzsyj.org]

Zhang Z, Zhou Y, Chen ZY, Tian XH, Du SH, Huang RM. Approximate entropy and support vector machines for electroencephalogram signal classification. *Neural Regen Res.* 2013;8(20):1844-1852.

Approximate entropy and support vector machines for electroencephalogram signal classification

Zhen Zhang¹, Yi Zhou¹, Ziyi Chen², Xianghua Tian³, Shouhong Du³, Ruimei Huang¹

1 Department of Biomedical Engineering, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou 510080, Guangdong Province, China

2 Department of Neurology, First Affiliated Hospital of Sun Yat-sen University, Guangzhou 510080, Guangdong Province, China

3 College of Medical Engineering and Technology, Xinjiang Medical University, Urumqi 830011, Xinjiang Uygur Autonomous Region, China

Research Highlights

(1) The characteristic index of the electroencephalogram signal was extracted using a nonlinear dynamics method, and a support vector machine was used for the classification of epileptic electroencephalogram signals. Our findings are more accurate than previous classification studies.

(2) Previous studies used electroencephalogram data from one or two cases in one electroencephalogram database, while we selected electroencephalogram data from four epileptic patients from the electroencephalogram database in two different hospitals, so the results are more representative.

(3) In this study, electroencephalogram data from different cases were regarded as training data and test data for the support vector machine, which was different from previous studies that only used data from the same case, so our results are more meaningful for the prediction of clinical seizures.

(4) We calculated the average classification accuracy rate of three cases as the final results, which are more convincing. Our findings indicate that a nonlinear dynamics index trained classifier can effectively identify epileptic electroencephalogram signals, and has good generalization ability.

Abstract

The automatic detection and identification of electroencephalogram waves play an important role in the prediction, diagnosis and treatment of epileptic seizures. In this study, a nonlinear dynamics index—approximate entropy and a support vector machine that has strong generalization ability were applied to classify electroencephalogram signals at epileptic interictal and ictal periods. Our aim was to verify whether approximate entropy waves can be effectively applied to the automatic real-time detection of epilepsy in the electroencephalogram, and to explore its generalization ability as a classifier trained using a nonlinear dynamics index. Four patients presenting with partial epileptic seizures were included in this study. They were all diagnosed with neocortex localized epilepsy and epileptic foci were clearly observed by electroencephalogram. The electroencephalogram data from the four involved patients were segmented and the characteristic values of each segment, that is, the approximate entropy, were extracted. The support vector machine classifier was constructed with the approximate entropy extracted from one epileptic case, and then electroencephalogram waves of the other three cases were classified, reaching a 93.33% accuracy rate. Our findings suggest that the use of approximate entropy allows the automatic real-time detection of electroencephalogram data in epileptic cases. The combination of approximate entropy and support vector machines shows good generalization ability for the classification of electroencephalogram signals for epilepsy.

Zhen Zhang, Master.

Corresponding author: Yi Zhou, M.D., Master's supervisor, Associate professor, Department of Biomedical Engineering, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou 510080, Guangdong Province, China, zhuyi@mail.sysu.edu.cn.

Received: 2013-03-17

Accepted: 2013-06-18
(N20120831002)

Acknowledgments: We would like to thank all the staff in the Department of Neurology, the First Affiliated Hospital of Sun Yat-sen University in China for providing experimental data and technical support.

Funding: This study was financially supported by the National Natural Science Foundation of China, No. 61263011, 81000554; Program in Sun Yat-sen University supported by Fundamental Research Funds for the Central Universities, No. 11ykpy07; Natural Science Foundation of Guangdong Province, No. S2011010005309; Innovation Fund of Xinjiang Medical University, No. XJC201209.

Author contributions:

Zhang Z was responsible for the study concept and design, analyzed experimental data, and wrote the manuscript. Zhou Y supervised the study and was the fund header. Chen ZY provided and integrated experimental data. Zhou Y, Tian XH, Huang RM and Du SH provided technical or information support, and supervised the study. All authors approved the final version of the paper.

Conflicts of interest: None declared.

Ethical approval: The experiment was approved by Ethics Committee of Zhongshan School of Medicine of Sun Yat-sen University in China.

Author statement: The manuscript is original, has not been submitted to or is not under consideration by another publication, has not been previously published in any language or any form, including electronic, and contains no disclosure of confidential information or authorship/patent application/funding source disputations.

Key Words

neural regeneration; brain injury; epilepsy; electroencephalogram; nonlinear dynamics; approximate entropy; support vector machine; automatic real-time detection; classification; generalization; grants-supported paper; neuroregeneration

INTRODUCTION

Epilepsy is caused by the highly synchronized abnormal discharge of neurons in the brain, it may recur with sudden onset, and can seriously affect the patient's life and work. The electroencephalogram signal contains a large amount of physiological and pathological information about the brain, and is regarded as a key method for the diagnosis and treatment of epilepsy and other brain diseases^[1]. The electroencephalogram signal consists of a group of randomized electrophysiological signals, which have characteristic nonstationary and nonlinear properties. Recently, a nonlinear approach has attracted increasing attention in electroencephalogram signal detection due to the great advantage of the randomness of representative signals^[2]. The nonlinear properties of electroencephalogram signals have become a hot topic.

At present, epileptic patients account for 5% of the total population worldwide, including six million cases in China. The majority of affected cases can be effectively controlled by drug treatment, while the remaining are intractable epilepsy patients because the drug treatment is ineffective. Among them, 21–24% of intractable epilepsy patients can be cured by surgical removal of epileptic loci; however, the symptoms of the remaining 75% of intractable epilepsy patients are unable to be controlled^[3]. Therefore, being able to predict recurrence and sudden seizures is very important for the improvement of the patient's quality of life.

A recent study found that a rule-based seizure prediction method using a nonlinear dynamics index achieved greater than 90% sensitivity for focal neocortical epilepsy^[4]. An effective prediction of an epileptic seizure requires automatic classification of epileptic

electroencephalogram signals^[5]. Therefore, we designed this study to explore the automatic identification and classification of electroencephalogram signals in epilepsy.

Growing evidence has focused on the automatic detection and identification of epilepsy signals^[6]. A variety of classification methods have been proposed^[7], such as decision tree, Gaussian mixture classification, fuzzy classification and support vector machines. In our experiment, we used a support vector machine for classification of electroencephalogram signals, because it requires only a small amount of data and has strong generalization ability^[8-10]. Unfortunately, in previous experiments^[8-10], training data and test data from electroencephalogram signals are often derived from the same cases, which may affect the clinical applicability of the classifiers.

To test the generalization ability of the nonlinear dynamics index and verify the availability of nonlinear dynamics in real-time detection of electroencephalogram epileptic signals, we selected training data and test data from the electroencephalogram signals of different cases. In brief, we first collected a multi-segmental electroencephalogram signal in one epileptic patient and continuous electroencephalogram signals from different cases. Then the signals were divided chronologically, and the nonlinear dynamics index was measured. The electroencephalogram data we collected were scalp data that had high noise disturbance, so we used approximate entropy, which is the best anti-noise nonlinear dynamics index^[11]. After the approximate entropy was calculated, we classified the data in each segment. Our findings indicate that this test method can clearly elucidate the generalization ability of the classifier.

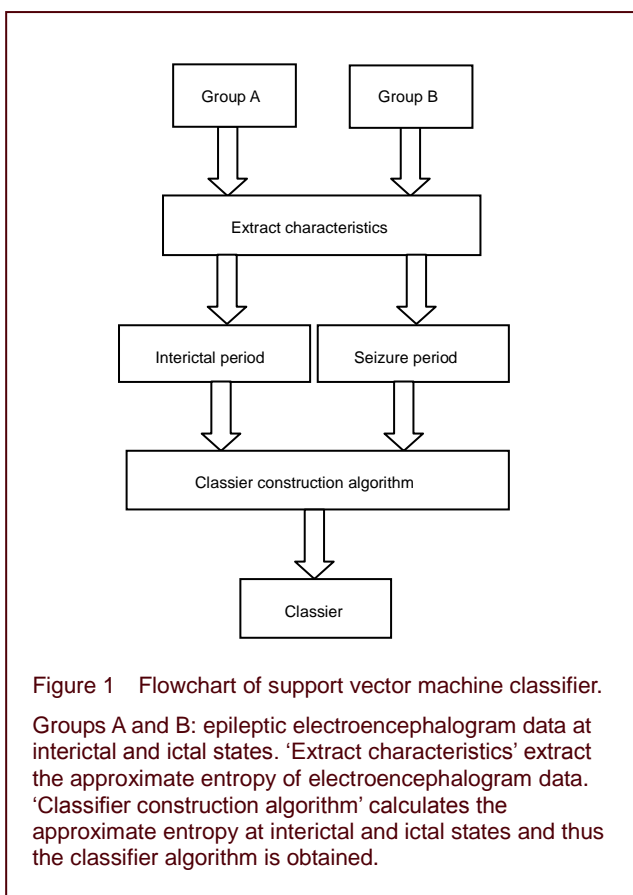
RESULTS

Quantitative analysis of subjects

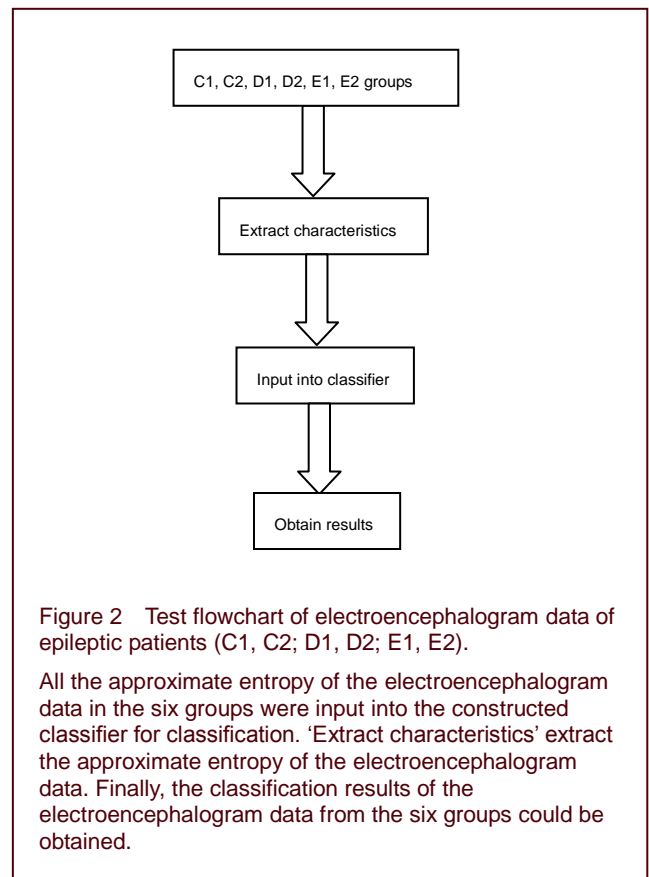
Four patients with partial epileptic seizures (neocortex localized epilepsy with clear loci) were included in this study. Their electroencephalogram data were divided into segments and the characteristic values of each segment were measured, *i.e.*, their approximate entropy. The support vector machine classifier was constructed using the approximate entropy of one epileptic case, and then the electroencephalogram waves of the other three cases were classified. All patients were involved in the result analysis, with no loss.

Construction of classifiers

The electroencephalogram data of patient 1, which were used to construct the support vector machine classifier, were divided into group A and group B. Group A contained 40 segments of electroencephalogram data during the epileptic interictal period, and group B contained 40 segments of electroencephalogram data during the epileptic seizure period. There were significant differences in the approximate entropy between group A and group B ($t = 5.863$, $P < 0.01$; paired sample t -test). The grouping is listed in Figure 1.



After the classifiers were constructed, the electroencephalogram data from the other three patients were chronologically and equally divided into two groups (patient 2: C1, C2; patient 3: D1, D2; patient 4: E1, E2), with 40 segments of data in each group. All of the approximate entropies of the electroencephalogram data in each group were input into the constructed classifier for classification (Figure 2).



As the classification result of the electroencephalogram data in each group was a one-dimensional vector, we can combine electroencephalogram C1, C2, D1, D2, E1, E2 into groups C, D, and E.

Classification results

The characteristics of classifiers can be described using the following index:

Sensitivity = true negativity / (true negativity + false positivity) (1)

Sensitivity means the percentage of successfully verified seizure signals among all test data.

Specificity = true negativity / (true negativity + false positivity) (2)

Specificity means the percentage of successfully excluded seizure signals during the ictal period.

Accuracy = number of correct classification / total number of test samples (3)
 Accuracy means the overall identification rate of the classifiers.

In the group C, according to the clinician's diagnosis report, 96–130 seconds were regarded as the ictal period, and patients presented with apparent spasm in the left upper limb, so we arranged 48–65 segments of data as the ictal data. As shown in Figure 3, the data in segments 49, 51–65, and 75 (17 segments in total) were involved in the ictal data.

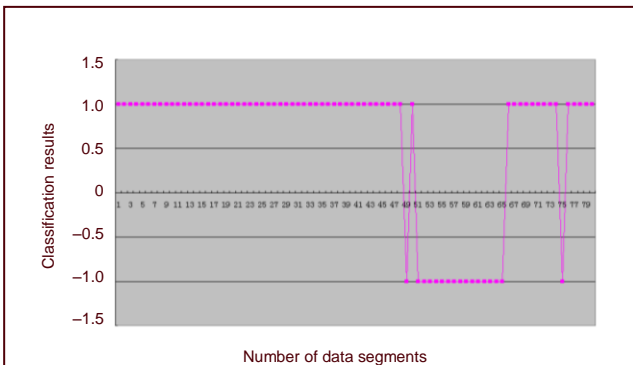


Figure 3 Electroencephalogram data classification results of patient 2 (group C).

According to the support vector machine classifier algorithm, the Y axis represents classification results, 1 represents the epileptic interictal period and -1 represents the epileptic ictal period. The X axis represents electroencephalogram data at 1–80 segments for 0–160 seconds.

Each segment of data represents 2 seconds. Segments 49, 51–65, 75 (total 17 segments) can be considered as ictal data.

In the group D, according to the clinician's diagnosis report, the data in segments 44–63 were the ictal data (Figure 4).

In the group E, according to the clinician's diagnosis report, the data in segments 53–72 were the ictal data (Figure 5).

Based on the clinician's diagnosis report, the classification sensitivity, specificity and accuracy rate were very high in the groups C, D, and E. The accuracy rate of electroencephalogram data classification in the three groups reached 93.33% (Table 1).

DISCUSSION

In this study, we tried to extract an electroencephalogram

signal index using a nonlinear dynamics method and used a support vector machine in the classification of epileptic electroencephalogram signals.

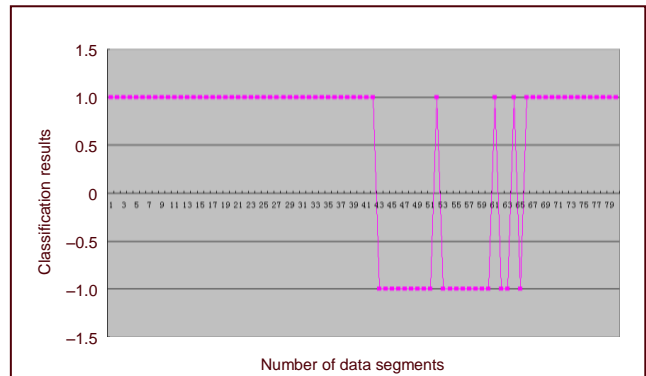


Figure 4 Electroencephalogram data classification results of patient 3 (group D).

According to the support vector machine classifier algorithm, the Y axis represents classification results, 1 represents the epileptic interictal period and -1 represents the epileptic ictal period. The X axis represents the electroencephalogram data at 1–80 segments for 0–160 seconds.

Each segment of data represents 2 seconds. Segments 44–63 can be considered as ictal data.

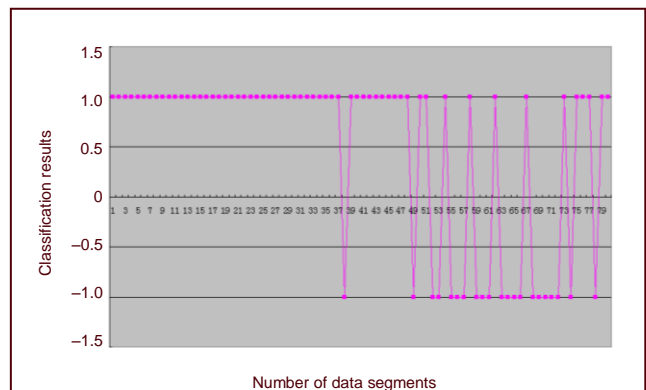


Figure 5 Electroencephalogram data classification results of patient 4 (group E).

According to the support vector machine classifier algorithm, the Y axis represents classification results, 1 represents the epileptic interictal period and -1 represents the epileptic ictal period. The X axis represents the electroencephalogram data at 1–80 segments for 0–160 seconds.

Each segment of data represents 2 seconds. Segments 53–72 can be considered as ictal data.

Compared with previous classification results^[1-2, 7-10], our findings are more accurate and convincing for the following reasons: First, previous studies used electroencephalogram data from one database and included only a few cases^[1-2, 7-10], while our study collected the electroencephalogram data of four patients from two different hospitals.

Table 1 Comparison of the sensitivity, specificity and accuracy of electroencephalogram data classification results for three epileptic patients

Group	Total data segments (segment)	Data at seizure (segment)	True positivity (segment)	False positivity (segment)	True negativity (segment)	False negativity (segment)	Sensitivity (%)	Specificity (%)	Accuracy (%)
C	80	18	15	1	62	2	88.24	98.41	96.25
D	80	20	18	2	58	2	90.00	96.67	95.00
E	80	20	17	4	54	5	77.27	93.10	88.75

The sensitivity, specificity and accuracy of electroencephalogram data classification results for three epileptic patients (groups C, D, and E) are high. The average classification accuracy rate reached 93.33%.

Second, previous studies also used training data and test data from one database^[1-2, 7-10], while our study collected training data and test data for the support vector machine from different databases and different cases. Finally, we calculated the average classification accuracy rate of three patients as the classification results, which are more convincing. Our experimental findings suggest that classifiers constructed with nonlinear dynamics indexes can effectively identify epileptic electroencephalogram signals and have good generalization ability. This method is simple and easy to use, and may benefit real-time detection and identification systems for epilepsy, thus assisting clinical diagnosis and treatment.

Currently, neuroscience has become a hot spot in life science research. The characteristics of the nervous system may be quantitatively described through research into neurophysiological signals and the use of nonlinear dynamics methods, thus furthering our understanding of epilepsy.

In this study, we used nonlinear dynamics and a support vector machine to classify epileptic electroencephalogram signals, and have obtained a high accuracy rate compared with previous studies^[1-2, 7-10]. Furthermore, the classifiers constructed with a nonlinear dynamics index showed good generalization ability. However, our experiments are theoretical research and are not yet fit to be used clinically, thus practical applications deserve further exploration.

SUBJECTS AND METHODS

Design

A computer-simulation, neuro-electrophysiological experiment.

Time and setting

Experiments were performed at the Medical Information Processing Laboratory, Department of Biomedical Engi-

neering, Zhongshan School of Medicine, Sun Yat-sen University, China from July to October in 2012.

Subjects

Electroencephalogram data from partial epileptic patients were collected from the First Affiliated Hospital of Sun Yat-sen University and the Second Affiliated Hospital of Guangzhou Medical College, China.

Inclusion criteria

(1) Neocortical localized epilepsy; (2) clear epileptic loci; (3) onset within 1 minute, with clear ictal and interictal periods.

Exclusion criteria

(1) Limbic system localized epilepsy; (2) unclear epileptic loci; (3) onset later than 1 minute, ictal and interictal periods were elusive.

Electroencephalogram data

Electroencephalogram data were collected from the electroencephalogram room, First Affiliated Hospital of Sun Yat-sen University (patient 1) and the electroencephalogram room, Department of Neurology, Second Affiliated Hospital of Guangzhou Medical College (patients 2-4), China, at 200 Hz and 256 Hz sampling frequency respectively.

Patient 1: Temporal lobe epilepsy. Electroencephalogram data were collected *via* the F7 lead (left anterior temple) for 1 200 seconds continuously, and divided into ictal data (40 segments; group A) and ictal data (40 segments; group B) according to clinician's reports. Each group contained 2 000 points of data.

Patient 2: Temporal lobe epilepsy. Electroencephalogram data were collected *via* the F8 lead (right anterior temple) for 160 seconds continuously, and chronologically divided into 80 segments (group C). Each segment contained 514 points of data, with each point representing 2 seconds. According to clinical perfor-

mance, 88–126 seconds were regarded as the ictal period, and the patient presented with left upper limb convulsions.

Patient 3: Temporal lobe epilepsy. Electroencephalogram data were collected via the T4 lead (right middle temple) for 160 seconds continuously, and chronologically divided into 80 segments (group D). Each segment contained 514 points of data, with each point representing 2 seconds. According to clinical performance, 88–126 seconds were regarded as the ictal period, and the patient presented with left upper limb convulsions and a left deflection of the head and the eyes.

Patient 4: Frontal lobe epilepsy. Electroencephalogram data were collected via the Fp1 lead (left anterior temple) for 160 seconds continuously, and chronologically divided into 80 segments (group E). Each segment contained 514 points of data, and each point represented 2 seconds. According to clinical performance, 106–144 seconds were regarded as the ictal period, and the patient presented with left upper limb stiffness and a right deflection of the head and the eyes.

All of the electroencephalogram data were collected from the scalp, processed and stored in the First Affiliated Hospital of Sun Yat-sen University, China and the Second Affiliated Hospital of Guangzhou Medical College, China. Prior to the experiments, all four epileptic patients were informed of the experimental scheme and risk, and gave their informed consent^[12]. The electroencephalogram data from the four epileptic patients are summarized in Table 2.

Table 2 General information of the four partial epileptic patients

Patient	1	2	3	4
Sex	F	F	M	F
Age (year)	16	25	15	13
Sample frequency (Hz)	200	256	256	256
Origin	Temporal	Temporal	Temporal	Frontal
Ictal segments (segment)	40	18	20	20
Interictal segments (segment)	40	62	60	60
Segment duration (second)	10	2	2	2

F: Female; M: male.

Methods

Extraction of nonlinear characteristics value

Introduction of approximate entropy: a common nonlinear kinetic parameter, which includes correlation dimension, Lyapunov exponent, and approximate entropy.

Among these, approximate entropy is a rule for the complexity of a measure series and statistical quantization. It can measure the probability of new patterns when the dimension increases from m to $m + 1$ and the incidence of new information in a time sequence. Complex signals tend to have high approximate entropy; on the contrary, regular signals tend to have low approximate entropy. Approximate entropy can reflect the overall characteristics of the signals. Compared with correlation dimension and Lyapunov exponent, approximate entropy requires a small amount of data (500–1 000 points) and has anti-noise ability. Generally, approximate entropy is calculated by its definition as follows:

For a given N -point time series $\{u(i)\}$, approximate entropy can be obtained by the following steps (m is the dimension of the preset pattern, and r is the preset similar tolerance):

(1) The sequence $\{u(i)\}$ composes the m -dimensional vector $X(i)$, namely:

$$X(i)=[u(i), u(i + 1) \cdots u(i + m - 1)], i = 1 - N - m + 1 \quad (1)$$

(2) For every i value, the distance between vector $X(i)$ and another vector $X(j)$ is calculated:

$$d[X(i), X(j)] = \max |u(i + k) - u(j + k)|, k = 0 - m - 1 \quad (2)$$

(3) Given the threshold value r ($r > 0$), for each i value, the number of $d[X(i), X(j)] < r$ is analyzed and the ratio of the number to the total number of vectors $N - m + 1$ is calculated, denoted as $C_i^m(r)$, namely:

$$[C_i^m(r) = \{d[X(i), X(j)] < r \text{ number}\} / (N - m + 1) \quad (3)$$

(4) $C_i^m(r)$ logarithm is obtained, and the average i value is calculated, denoted as $\Phi^m(r)$, namely:

$$\Phi^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \ln C_i^m(r) \quad (4)$$

(5) For $m + 1$, repeat steps 1–4, to obtain $\Phi^{m+1}(r)$

(6) Obtain the final results as follows:

$$ApEn(m, r) = \Phi^m(r) - \Phi^{m+1}(r) \quad (5)$$

Usually $m = 2$, $r = 0.1 - 0.25 \text{ SD}(u)$ (SD: standard deviation), and the obtained approximate entropy has reasonable statistical properties, which was consistent with previous studies^[13], so we set $m = 2$, $r = 0.25 \text{ SD}(u)$ for approximate entropy.

A fast approximate entropy algorithm: Approximate entropy was extracted using a fast algorithm, which overcomes the shortcomings of common algorithms, such as low calculation efficiency, slow speed and being not conducive to real-time operation. This algorithm was proposed by Hong *et al*^[14]. using the theory of binary distance matrix; the speed was 5 times the definition calculation.

1) For the N -point sequence, the $N \times N$ binary distance matrix D was calculated, and the element in the D (i^{th} row and j^{th} column) is denoted by d_{ij} :

$$d_{ij} = \begin{cases} 1 & |u(i)-u(j)| < r \\ 0 & |u(i)-u(j)| \geq r \end{cases} \quad i = 1, 2, \dots, N \quad (6)$$

2) C_i^2 and C_i^3 were calculated using the elements in matrix D :

$$C_i^2(r) = \sum_{j=1}^{N-1} d_{ij} \cap d_{(i+1)(j+1)} \quad (7)$$

$$C_i^3(r) = \sum_{j=1}^{N-2} d_{ij} \cap d_{(i+1)(j+1)} \cap d_{(i+2)(j+2)}$$

3) $C_i^2(r)$ logarithm is obtained and the average i value is calculated:

$$\Phi^2(r) = \frac{1}{N-1} \sum_{i=1}^{N-1} \ln C_i^2(r) \quad (8)$$

$C_i^3(r)$ logarithm is obtained and the average i value is calculated:

$$\Phi^3(r) = \frac{1}{N-2} \sum_{i=1}^{N-2} \ln C_i^3(r) \quad (9)$$

4) Approximate entropy estimate:

$$ApEn(2, r) = \Phi^2(r) - \Phi^3(r) \quad (10)$$

This fast approximate entropy algorithm omits the vector construction process and directly obtains the difference of data in the time series, thus improving the calculation speed of approximate entropy. During seizures, a large group of neurons in the human brain synchronously discharge, and several brain functions are inhibited, therefore, compared with normal brain electrical activity, the complexity will be reduced. From the interictal period to the ictal period, approximate entropy reduces to varying degrees and then begins to increase after the end of the seizure^[15-16].

Support vector machine

Definition of support vector machine: It is the optimal design rule of linear classifier proposed by Vapnik *et al*^[17-19]. Its basic principle is to find the optimal hyperplane in the feature space, and to separate two types of samples correctly. The interval between different types of sample sets that are close to the hyperplane should be the maximum, so as to achieve a greater generalization capability; when the optimal classification hyperplane cannot completely separate two types of sample, it is suggested to introduce a relaxation factor ξ , allowing the presence of misclassified samples, thus balancing the empirical risk and the promotion.

A support vector machine is based on structural risk minimization, while others are based on empirical risk minimization. Compared with other classification methods, a support vector machine will keep the sample characteristics and consider the adaptability of new input data; therefore, it has a stronger learning ability for small sample sizes^[20].

Data points are expressed as an n -dimensional vector x , and the present study discusses a binary classification algorithm, with $y = 1$ or -1 to represent different levels. The classifier aims to find a hyperplane in the n -dimensional space, using the following equation:

$$w^T x + b = 0 \quad (11)$$

Through the hyperplane, two types of data can be separated. If the data point on one side of the hyperplane is -1 ,

then that on the other side of the hyperplane is all 1:

$$f(x) = w^T x \quad (12)$$

Obviously, if $f(x) > 0$, then x is a point located on the hyperplane. As for all $f(x) < 0$ points, $y = -1$, while $f(x) > 0$ corresponds to $y = 1$. During the classification process, we introduced the data points x into $f(x)$. If the result was less than 0, we defined it as -1 , and if greater than 0, we defined it as 1 ^[21]. After the optimal classification plane in the character space was chosen, the binary classification results could be tested with the test data. For the new data x , $f(x)$ can be defined as:

$$f(x) = \sum_{i=1}^N y_i \lambda_i K(x_i, x) + b \quad i = 1, 2, \dots, N \quad (13)$$

Here, $\lambda_i \geq 0$ is the Lagrangian factor, meeting

$$y_i (w \cdot \varphi(x_i) + b) = 1. \quad K(x_i, x) \text{ is the kernel function,}$$

which receives two low-dimensional space vectors and mappings to high-dimensional space, calculating the vector inner product through a transformation and constructing a linear discriminant function in high-dimensional space. Many problematic nonlinear classification problems in low-dimensional space can be solved through the kernel function conversion into a high-dimensional space.

Influencing factors: There are two factors contributing to the properties of the support vector machine classifier: the kernel function and the error penalty parameter C ^[22-23]. Currently, the selection criteria for the kernel function remain elusive, and radial basis functions have achieved good outcomes in electroencephalogram signal analysis^[24].

The radial basis functions can map nonlinear time series low-dimensional space vector sets into high-dimensional space, so the kernel function in the support vector machine we used was a radial basis function. The error penalty parameter C is applied to optimize the generalization ability of the machine, which may be affected by the misclassified sample proportion and the algorithm complexity. In our experiments, we tried $C = 50$, $C = 100$, and C is close to infinity for numerical calculation. When $C = 50$, there were two misclassified samples; when $C =$

100, there were also two misclassified samples; and when C was approaching infinity, only one sample was misclassified, so we set the error penalty parameter C as approaching infinity.

Statistical analysis

All data are derived from raw electroencephalogram images. The data of patient 1 were divided into groups A and B, and the approximate entropy of the two groups was compared using a paired samples t -test. A $P < 0.01$ value was considered significant.

REFERENCES

- [1] Li SF, Zhou WD, Cai DM, et al. EEG signal classification based on EMD and SVM. *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi*. 2011;28(5):891-894.
- [2] Yuan Q, Zhou WD, Liu YX. Epileptic seizure detection with linear and nonlinear features. *Epilepsy Behav*. 2012;24(4):415-421.
- [3] Iasemidis LD. Epileptic seizure prediction and control. *IEEE Trans Biomed Eng*. 2003;50(5):549-558.
- [4] Aarabi A, He B. A rule-based seizure prediction method for focal neocortical epilepsy. *Clin Neurophysiol*. 2012;123(6):1111-1122.
- [5] Mirowski P, Madhavan D, Lecun Y, et al. Classification of patterns of EEG synchronization for seizure prediction. *Clin Neurophysiol*. 2009;120(11):1927-1940.
- [6] Yuan Q, Zhou W, Li S. Epileptic EEG classification based on extreme learning machine and nonlinear features. *Epilepsy Res*. 2011;96(1-2):29-38.
- [7] Gadhoumi K, Lina JM, Gotman J. Discriminating preictal and interictal states in patients with temporal lobe epilepsy using wavelet analysis of intracerebral EEG. *Clin Neurophysiol*. 2012;123(10):1906-1916.
- [8] Zhao JL, Zhou WD, Liu K, et al. Application of SVM and wavelet method in classification of EEG signals. *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi*. 2011;28(2):277-279.
- [9] Liu YX, Zhou WD, Yuan Q, et al. Automatic seizure detection using wavelet transform and SVM in long-term intracranial EEG. *IEEE Trans Neural Syst Rehabil Eng*. 2012;20(6):749-755.
- [10] Ocak H. Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. *Expert Syst Appl*. 2009;36(2):2027-2036.
- [11] Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol*. 2000;278(6):2039-2049.
- [12] State Council of the People's Republic of China. Administrative Regulations on Medical Institution. 1994-09-01.
- [13] Zhou Y, Huang RM, Chen ZY, et al. Application of approximate entropy on dynamic characteristics of epileptic absence seizure. *Neural Regen Res*. 2012;7(8):572-577.

- [14] Hong B, Tang YQ, Yang FS, et al. Primary application of appropriate entropy(ApEn), Cross-ApEn's property and fast algorithm in EEG and it's recognition reseach. *Signal Process*. 1999;15(2):100-108.
- [15] Srinivasan V, Eswaran C, Sriraam N. Approximate entropy-based epileptic EEG detection using artificial neural networks. *IEEE Trans Inf Technol Biomed*. 2007;11(3): 288-295.
- [16] Guo L, Rivero D, Pazos A. Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks. *J Neurosci Methods*. 2010; 193(1):156-163.
- [17] Vapnik VN, Chapelle O. Bounds on error expectation for support vector machines. *Neural Computation*. 2000;12(9): 2013-2036.
- [18] Drucker H, Wu DH, Vapnik VN. Support vector machines for spam categorization. *IEEE Trans Neural Netw*. 1999; 10(5):1048-1054.
- [19] Chapelle O, Haffner P, Vapnik VN, et al. Support vector machines for histogram-based image classification. *IEEE Trans Neural Netw*. 1999;10(5):1055-1064.
- [20] Vapnik VN. An overview of statistical learning theory. *IEEE Trans Neural Netw*. 1999;10(5):988-999.
- [21] Saitta L, Kripakaran P, Raphael B, et al. Feature selection using stochastic search: an application to system identification. *J Comput Civil Eng*. 2010;24(1):3-10.
- [22] Feng GH. Parameter optimizing for support vector machine classification. *Jisuanji Gongcheng yu Yingyong*. 2011;47(3):123-124.
- [23] Fu YY, Ren D. Choice of kernel function and other parameters in support vector machine. *Keji Chuangxin Daobao*. 2010;(9):6-7.
- [24] Zavar M, Rahati S, Akbarzadeh MR. Evolutionary model selection in a wavelet-based support vector machine for automated seizure detection. *Expert Syst Appl*. 2011;38(9):10751-10758.

(Reviewed by Jackson C, Norman C, Chi ZF, Chen ZY)
(Edited by Wang J, Yang Y, Li CH, Song LP, Liu WJ, Zhao M)