Journal of Scientific & Industrial Research
Vol. 82, January 2023, pp. 101-108
DOI: 10.56042/jsir.v82i1.70213

EEG Signal Classification Automation using Novel Modified Random Forest Approach

G. Aloy Anuja Mary[1]*, M Purna kishore[2], Sridevi Chitti[3], Ramesh Babu Vallabhaneni[4] & N Renuka[5]

[1]Veltech Rangarajan Dr. Sagunthala R & D Institute of science and Technology, Chennai 600 062, India
[2]Dept of ECE, KKR & KSR Institute of Technology & Science, Vinjanampadu, Guntur 522 017, India
[3]Department of ECE, Kakatiya Institute of technology & science, Warangal 506 371, India
[4]Dept of ECE, NRI Institute of Technology, Pothavarapadu (V) Agiripalli, Vijayawada 520 010, India
[5]Department of ECE, RVR and JC College of Engineering, Guntur 522 019, India

Received 03 May 2022; revised 02 October; accepted 06 October 2022

Digitalization and automation are the two aspects in the medical industry that define compliance with industry 4.0. Automation is essential for speeding up the diagnosis process, while digitalization leads to smart medicine and efficient diagnosis. Epilepsy is one such disease that can use these automation techniques. The automatic monitoring of epilepsy EEG is of great significance in clinical medicine. Aiming at the non-stationary characteristics of EEG signals, the classification of EEG signals is based on the combination of overall empirical mode. It is proposed using the random forest method. The EEG signal data set has an epileptic interval over 200 single-channel signals with a seizure period. A total of 819,400 data are used as samples. First, the overall epileptic EEG signal modal is decomposed into multiple intrinsic modal functions. The effective features are extracted from the first-order intrinsic modal function. Finally, random forest and Least Square SVM (LS-SVM) are considered to classify the EEG signals characteristics. The correct recognition rate of random forest and LS-SVM is compared. The results show that random forest classification method has an ideal classification effect on epilepsy EEG signals during and between seizures. The recognition accuracy is 99% and 60%, which is higher than the accuracy of the LS-SVM. The proposed method improves clinical epilepsy. The efficiency of EEG signals analysis.

Keywords: EEG, Ensemble mode decomposition, Feature extraction and recognition, SVM

Introduction

The entire world has been witnessing the 4th revolution in industrialization. This industrialization is the evolution to fully automated digital smart environment. The cyber-physical systems are the significant part of it. This comprises many diversified innovations and technologies in different sectors. In this paper, the primary emphasis is on the healthcare or medical domain. The entire ecosystem is moving towards healthcare 4.0, with the induction of the standards of Industry 4.0.

Epilepsy is a brain disease that seriously affects human health due to brain damage. It is estimated that 1% of total world's population is affected by epilepsy. Identifying and classifying EEG signals of patients with epilepsy can help doctors determine epilepsy. The specific type of seizures can be used to choose a reasonable treatment plan and help patients avoid danger in time. Therefore, the research on the recognition and classification of epilepsy EEG signals has important theoretical significance and practical value.

In the late 19th century, people began studying electroencephalogram signals (EEG). In 1875, Richard Caton in the United Kingdom recorded the electrical activity of their cerebral cortex through experiments on animals. B19 A term, and the first recorded brain electrical activity from the human scalp. In 1929, Berger, a professor of neurology at the University of Jena, Germany, discovered EEG for the first time, which laid the foundation for EEG research.[1,2] In 1934, Berger's observations were recognized by the human brain. Electrograms are recognized by the scientific community and are widely used in many fields, such as clinical medicine, psychology, pharmacology, brain dynamics, etc.[3]

The analysis of EEG signals mainly includes original signal decomposition, feature extraction, recognition, classification, etc. Signal EEG decomposed into 1/2 waves then features are extracted for EEG epileptic signals processing and is proposed by Gotman *et al*., used wavelet analysis To detect the onset of epileptic

*Author for Correspondence
E-mail: draloyanujamary@veltech.edu.in

EEG signals, Gabor *et al*. used artificial neural network technology to detect the onset of EEG.[4–6] Artificial neural networks combined with time-frequency analysis are often used to detect the onset of epilepsy. For example, Ocak *et al*. used discrete wavelet transform to decompose the original signal and extracted approximate entropy for classification to detect epileptic seizures, Oweis *et al*.[7,8] The intrinsic function model is extracted to achieve mode hybrid classification, which is used to classify the actual period of interictal and seizure of the EEG Period.[7,8] In the prediction research based on empirical mode decomposition, Shufang *et al*. proposed a method based on empirical mode decomposition and support vector machine classification to classify EEG signals.[9] The EEG signals are subjected to wavelet packet transform and nonlinear analysis to realize the mental fatigue state.[10] Similarly, the EEG signal denoising method based on the double-density wavelet threshold technique can eliminate the noise mixed in the EEG signal.[11] The Fourier transform can be used to analyze the EEG signal, there were many classical analysis methods of time domain analysis and frequency domain analysis.[12,13] With the continuous development of technology and the emergence of automatic EEG analysis systems, the use of computers to assist in analyzing EEG signals has gradually been favored by scholars. In recent years, in the field of EEG signal analysis, a series of modern analysis methods have appeared successively.[14]

In the 1940s, EEG signals began to be combined with clinical practice, and it has been widely developed worldwide. In addition to EEG signal monitoring, it can detect epilepsy, brain tumors, brain injuries, hypoxia, ischemia, cerebrovascular diseases, and the central nervous system. In addition to providing information on diagnosis, prevention, and treatment of system diseases, as well as metabolism, mental illness, and mental disorders, it can also detect advanced central activities such as brain function, cognition, and thinking.[14,15] However, EEG Linear characteristics are very difficult to analyze. Therefore, the research of EEG signal analysis methods is constantly developing and deepening. In order to improve the recognition rate of EEG signals, this study uses the method of combining empirical mode decomposition and random forest to treat epilepsy. EEG signals are classified for this purpose.

Proposed Method

Experimental Design

The EEMD-RF classification model and the EEMD-LSSVM classification model established in this study are implemented under the Windows7 system, and the development platform adopted is R∗64 3 .3. 0 and MATLAB R2014a.

The experimental data comes from the epilepsy database of the University of Bonn, Germany. The database contains 5 types of EEG signals. Each type of EEG signal contains 100 single-channel with 236s time and 17361 Hz sampling frequency, each single channel signal. Contains 4,096 sample points, among which data sets A and B are the scalp surface EEG signals from 5 healthy volunteers with eyes open and closed, respectively. Data set C is the hippocampal structure of 5 epileptic patients during the seizure interval EEG signals. Data set D is the EEG signal of the epileptic zone during the epileptic period. Data set E is the EEG signal of the epileptic zone during the epileptic period. This study mainly analyzes the D data set and epilepsy of the epileptic period, and E data set during the attack period.

The experiment first realizes EEMD through MATLAB, decomposes the epileptic EEG signal into multiple intrinsic modal functions, calculates the correlation coefficients between the intrinsic modal function of the original EEG to select the effective intrinsic modal function, and uses SA to calculate it. Relevant features. Finally, RF and LSSVM are used to classify epileptic EEG signal features respectively through R. The correct recognition rate of the classification results of the two methods is compared to determine which classification method is effective for the brain in the inter-seizure and the epileptic period. The process of feature extraction is shown in Fig. 1.

Global Decomposition Using Empirical Mode

The overall EEMD method is to add several times to the inherent modal function of the decomposition.

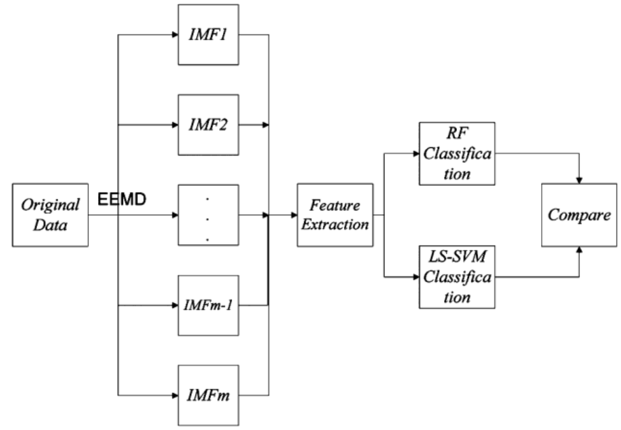

Fig. 1 — EEG ictal detection Process

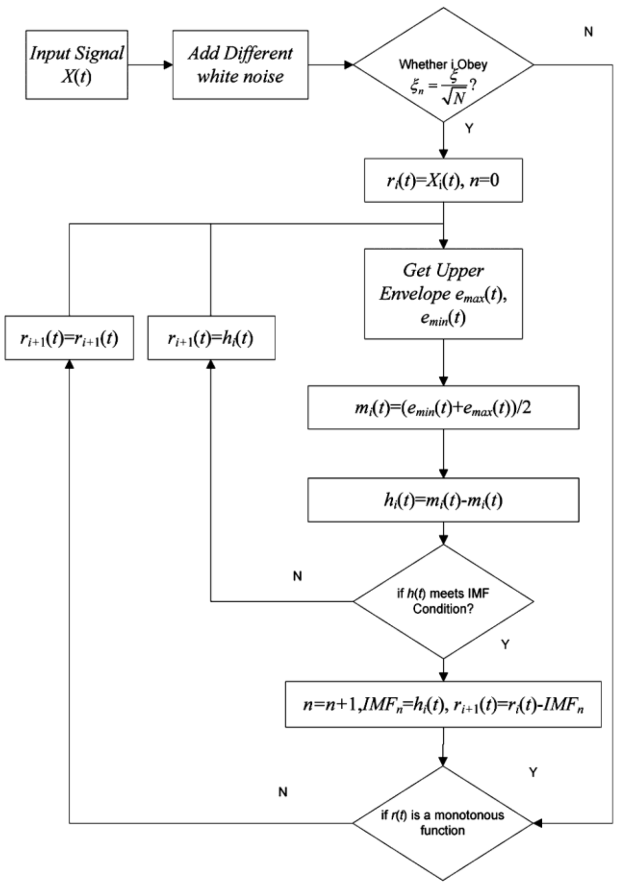

Fig. 2 — EEMD algorithm flow chart

The empirical mode decomposition (EMD) is processed separately, and finally, a global method of averaging.[15,16] This method cancels the added white noise and not only retains signal information of the original sequence but also overcomes it to a large extent. The problem of modal aliasing is solved, and the decomposition is physically unique. The EEMD algorithm flow is shown in Fig. 2.

The EEMD specific steps are below

1) By adding a set of w(t) Gaussian noise to the signal to be analyzed x(t), a signal-to-noise mixture is formed, namely

$$X(t) = x(t) + w(t) \quad \ldots (1)$$

2) Perform EMD decomposition on the signal-to-noise mixture X(t), and decompose it into a combination of IMF components, namely

$$X(t) = \sum_{j=1}^{n} c_j + r_n \quad \ldots (2)$$

3) Add different white noise w(t) to the signal to be analyzed, and repeat the above two steps.

$$X_i(t) = x(t) + w_i(t) \quad \ldots (3)$$

After decomposition, the respective IMF component combinations are obtained, namely $X_i(t) = \sum_{j=1}^{n} c_{ij} + r_{in}$

4) Average the IMF corresponding to all IMF combinations, namely

$$c_j = \frac{1}{N}\sum_{i=1}^{N} c_{ij} \quad \ldots (4)$$

In (4), N indicates the whole number, and final decomposition result is obtained.

$$X(t) = \sum_{j=1}^{n} c_j + r_n \quad \ldots (5)$$

Due to the zero-average characteristic of white noise, the number of times of adding noise is sufficient. After the results of these multiple decompositions are taken average, the noise will eventually be canceled out to achieve the effect of elimination, and the overall average result can be regarded as true. The signal selects 100 single-channels of seizure period and epileptic interval in the Bonn data set, respectively, perform the overall empirical mode decomposition, calculate the correlation coefficients of each order intrinsic mode function and the original EEG, select the appropriate IMF to perform the following analysis

Feature Extraction

The essential features of each order of intrinsic mode functions obtained after the overall empirical mode decomposition are extracted, including the mean, variance, standard deviation, range, coefficient of variation, energy entropy, fluctuation index, and information entropy.

Variation Coefficient

EEG signal amplitudes are analyzed, and its features that are often used are simple statistics such as mean, and variance. The coefficient of variation measures amplitude changes in EEG signals. For EEG between epileptic seizures, the signals with regular amplitude changes such as signals, the value of the coefficient of variation is relatively small, which is defined as

$$CV = \frac{\sigma^2}{\mu^2} \quad \ldots (6)$$

Here, μ is given as

$$\mu = \frac{1}{l}\sum_{j=1}^{l} \left|IMF_j\right|$$

$$\sigma = \sqrt{\frac{1}{l}\sum_{j=1}^{l}\left(IMF_j - \frac{1}{l}\sum_{j=1}^{l} IMF_j\right)^2}$$

In the formula, l is intrinsic length.

Volatility index

The index of volatility is to measure intensity of signal fluctuation during the period of epilepsy is usually more intense than the fluctuation of the interval of the seizure, and is defined as

$$F_i = \frac{1}{l}\sum_{j=1}^{l}\left|IMF_{j+1} - IMF_j\right| \quad \dots (7)$$

In the formula, l is intrinsic length.

Energy Entropy

To facilitate energy entropy, feature extraction characterizes the difference in the characteristics of different intrinsic modal functions. The formula for the energy of the intrinsic modal function $c_i(t)$ is shown below, t_1 and t_2 are the signal start and end time respectively[16,17]

$$E_1 = \int_{t_1}^{t_2} c_i^2(t)dt \quad \dots (8)$$

It is defined as

$$H = -\sum_{i=1}^{n} P_i log P_i \quad \dots (9)$$

In the formula, $P_i = \frac{E_i}{E}$ is the ratio of energy.

Information Entropy

The information entropy of the epileptic seizure signal is usually lower than that of the interracial period, which is defined as

$$H(x) = E\left[\log_2\left(2, \frac{1}{p(x_i)}\right)\right] = -\sum p(x_i)\log_2\left(2, p(x_i)\right) (i = 1, \dots, n) \quad \dots (10)$$

In the formula, x represents a random variable.

The interictal period and the interictal period IMF1 ~ IMF5 in the Bonn data set are selected to calculate the relevant features like mean, variance, standard deviation, range, etc.

Random Forest

Random forest was first proposed in 2001.[17,18] It is a non-parametric statistical method. Its basic idea is to use the bootstrap method to extract multiple sub-samples from the original sample. Sub-samples are used to model the decision tree, and the prediction results of multiple decision trees are combined using the voting method or the average method to determine the final prediction result.

Compared with NN, SVM, decision trees, and other methods, random forest has better noise tolerance and higher prediction accuracy and is not prone to overfitting problems. Random forest is composed of several tree classifiers$\{h(x, \theta_k), k = 1, \dots$. The final output of the random forest, the specific steps[19] are as below

Step 1 Use the Bagging method to form individual training sets, that is, each individual training set is to extract n samples from original training signal dataset with replacement.

Step 2 For each training set, use the following process to generate A classification regression tree without pruning.

1) Suppose there are a total of M primitive attributes, given a positive integer $mtry$, which satisfies $mtry \leq M$, at each internal node, randomly select mtry attributes from the M primitive attributes as the candidate attributes of the split node. In the process of generating the entire forest $Mtry$ remains unchanged.

2) Select the best split method from $mtry$ candidate attributes to split the node.

3) Let each tree grow fully without pruning.

4) Repeat 1) and 2). Until generate $ntree$ classification regression tree ($ntree$ is large enough).

5) When classifying samples of unknown categories, the output is

$$c = \arg\max_c \left(\frac{1}{ntree}\sum_{k=1}^{ntree} I(h(x, \theta_k) = c)\right) \quad \dots (11)$$

Random forest introduces random selection attributes on the basis of Bagging, which reduces the correlation between trees to a greater extent. At the same time, the establishment of a single non-pruned classification regression tree can obtain lower deviations, thereby ensuring the performance of random forests classification performance.

In the RF classification process, the mean, variance, standard deviation, range of IMF1~IMF5 are extracted. The feature vector of the training data is sent to an RF classification to obtain the best RF parameters. Finally, the trained RF classifier is used to distinguish the EEG signal categories in the test set.

Least Square Support Vector Machine

The basic idea of the core support-vector of the theorem 2 of the 0 theorem is SVM). The e_i and inequality constraints are changed to the second normal form and equality constraints of the selection error.[20,21] Compared with the SVM algorithm, the LSSVM algorithm simplifies the calculation

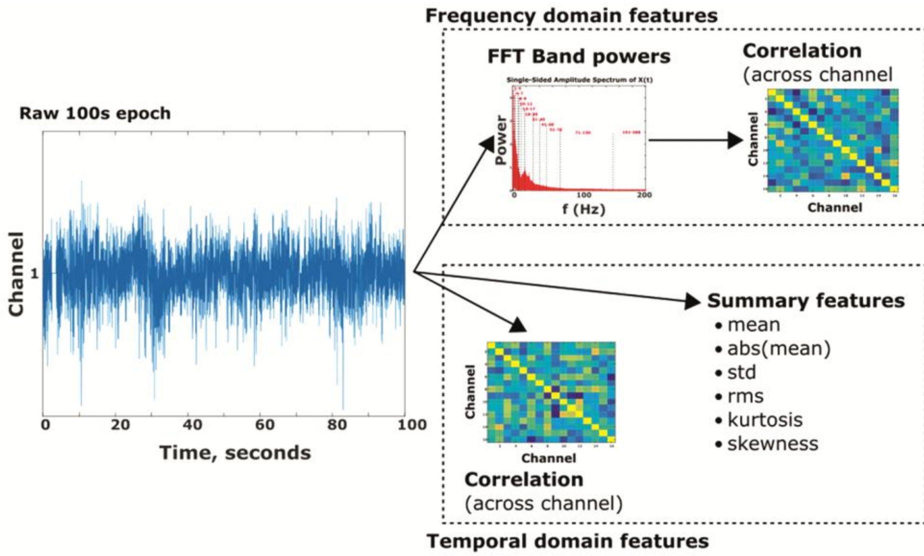

Fig. 3 — Sample of preictal and interictal EEG

complexity, reduces the prediction error, and ensures the accuracy of the calculation results.

Assume that the training sample set $(x_i, y_i)(i = 1, 2, \ldots, N$ as dimensional feature space mapping. Construct the optimal decision function in the high-dimensional space, and the optimization model according to the principle of structural risk minimization.

$$\min J(W,\xi) = \min \frac{1}{2} W^T W + \frac{C}{2} \sum_{i=1}^{M} \xi_{i,2} \quad \ldots (12)$$

The Constraints are

$$y_i = W^T g(x_i) + b + \xi_i (i = 1, 2, \ldots, M)$$

Since W is generally infinite dimensional, it is difficult to calculate the optimization model directly. This planning problem is usually transformed into its dual space. According to the Hilbert-Schmidt principle, introducing a kernel function$K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$. Convert the inner product calculation problem in the transformation space. Finally, the LSSVM regression function is obtained as

$$y(x) = \sum_{i=1}^{M} a_i K(x, x_i) + b \quad \ldots (13)$$

In the LSSVM classification process, the first 75% feature vectors of the interictal and interictal data sets are selected as training samples, and the last 25% of the feature vectors are used as test samples. To find the best parameters to obtain the trained LSSVM. The classifier sends the test sample to the classifier and performs multiple experiments to obtain the average classification result.

Results and Discussion

Epilepsy EEG Signal Decomposition

A sample of EEG during the inter-seizure period is shown in Fig. 3. One single-channel 200 single-channel signals in Bonn data set D is selected, which contains 4,097 sample points. It can be seen that the signal amplitude is large, and it has regularity. The EEG signal samples correspond to epileptic seizures. One single-channel over 200 of Bonn data set E is selected, containing 4,097 sample points. The signal amplitude can be seen as small and unstable. It is observed there is a significant gap between the EEG signal during the epileptic period and the EEG signal during the epileptic period.

The overall empirical mode decomposition of the epileptic EEG signal is performed to obtain order for each intrinsic modal. By calculating the correlation coefficients of the intrinsic modal functions of each order, IMF1 ~ IMF5 are selected to extract the features and calculate them. Seizure interval and epileptic seizures, the first five intrinsic mode functions of a certain period of EEG are shown in Fig. 4. After the overall empirical mode decomposition, the intrinsic mode function is given in Fig. 4(a) for the inter-seizure EEG signal. Choose the Bonn data set one of the 100 single-channel signals in D. The inter-seizure EEG signal after overall empirical mode decomposition has 8 intrinsic mode functions and 1 trend item.

Similarly, Fig. 4(b) is the intrinsic modal function of the epileptic seizure EEG signal after the overall empirical mode decomposition. One of the 100 signals in Bonn dataset E is selected as the signal sample.

From Fig. 4 (a) and Fig. 4(b), it is observed that there are differences in the frequency and amplitude of EEG signal between epileptic seizures and the EEG signal during seizures. The modal function calculates the correlation coefficient.

Feature Extraction of Epilepsy EEG Signal

Feature extraction is performed on 200 single-channel signals of epilepsy EEG signal between seizures and seizures. After the overall empirical mode decomposition, the correlation coefficients are calculated and the first five intrinsic modal functions selected are extracted. The main selections are the mean, variance, and standard deviation; Simple statistics such as range are used as features for the next classification.

Both Table 1 and Table 2 are the characteristic result samples of the epilepsy EEG signal interval and the epilepsy period of the Bonn data set. Take a single channel signal as an example. After the overall empirical mode decomposition, the intrinsic modal function and the original brain Correlation coefficients of electrical signals, IMF1 ~ IMF5 are selected to calculate the relevant features.

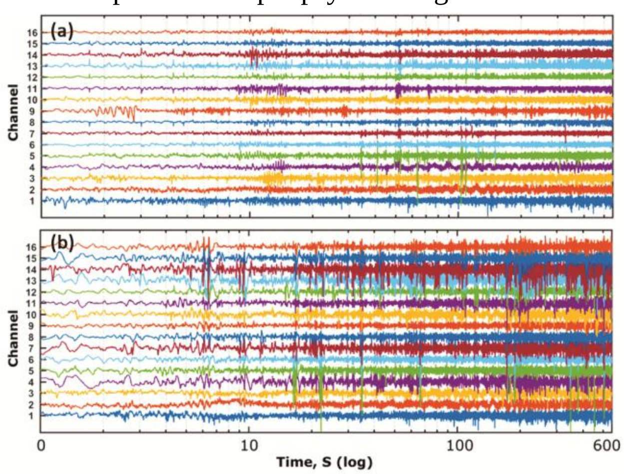

Fig. 4 — Top 16 IMFs of (a) interictal and (b) preictal EEGs on Bonn data sets

Recognition and Classification of Epilepsy EEG Signals

In this study, two methods of RF and LSSVM were selected to classify the intrinsic modal functions, and the results were compared. The results of classifying all the characteristics of intrinsic modal functions and their combinations were summarized in Table 3, which showed that the use of RF to classify IMF1 ~ IMF5 combined features have the best classification effect, and the highest correct recognition rate is 99.60%.

The experimental results show that the characteristics of IMF1~IMF5 selected by the EEMD-RF method have the best classification effort on interictal and EEG, and the correct recognition rate of 99% is 60%. The EEG classification effect during the interictal period is the best, with a correct recognition rate of 96% and 100%. After comparison, it is found that the EEMD-RF method has better classification accuracy for EEG than the EEMD-LSSVM method.

EEG signal classification method is based on the combination of EMD and SVM.[9] An improved method based on this is proposed for the EEG signal classification combining EEMD and LSSVM to deliver better classification results.[16–19] This confirms that the signal decomposition method based on EMD is compared with machine learning. The combination of EEG classification methods significantly affects the classification of EEG signals. The work presented

Table 1 — Sample of Interictal EEG features

IMF	Mean	Variance	Standard Deviation	Extremum	Coefficient of variance	Volatility Index	Energy entropy	Information entropy
IMF_1	11.39	39197.87	197.98	1843.50	1738.81	0.01	3.31	27.26
IMF_2	7.49	62194.69	249.39	1959.35	3329.42	−0.05	2.82	27.92
IMF_3	−4.39	60439.04	245.84	2139.20	−5595.23	0.28	2.01	27.88
IMF_4	−0.05	427.79	20.68	318.31	−43423.76	−0.02	6.65	20.74
IMF_5	0.36	1288.87	35.90	537.91	9983.24	0.01	7.19	22.33

Table 2 — Sample of ictal EEG features

IMF	Mean	Variance	Standard Deviation	Extremum	Coefficient of variance	Volatility Index	Energy entropy	Information entropy
IMF_1	0.01	116.86	10.81	151.05	177642.53	0.01	4.67	18.87
IMF_2	0.02	311.75	17.66	257.08	73778.05	−0.02	3.68	20.28
IMF_3	0.17	128.33	11.33	152.53	6620.47	0.00	4.83	19.00
IMF_4	−0.13	268.94	16.40	217.49	−12415.83	0.00	3.82	20.07
IMF_5	0.06	165.76	12.87	257.63	20448.32	0.00	4.38	19.37

Table 3 — Comparison of EEG results classification

Intrinsic mode function	Number of Samples	RF Recognition rate/%	LSSVM recognition rate
IMF_1-IMF_2	200	99.50	96.00
IMF_1-IMF_3	300	99.33	92.00
IMF_1-IMF_4	400	99.25	90.00
IMF_1-IMF_5	500	99.60	88.00

by Li *et al.*[16] selected the overall EMD and LS-SVM classification method when studying the problem of epilepsy EEG signals. The electrical signal is decomposed and classified, which proves that the overall empirical mode decomposition method can effectively decompose epileptic EEG signals and retain meaningful information in the research of epileptic EEG signal classification.

EEMD has great advantages in time-frequency analysis, especially in processing nonlinear and non-stationary signals. It can maintain the benefits of EMD and solve the end effect and modal aliasing problems in EMD. In addition to EEG signal decomposition, it can also be widely used. It is applied to other fields, such as fault diagnosis, weather forecasting, etc. RF can resist noise and interference well compared with other current classification algorithms. As a supervised learning method, it is not prone to overfitting. For imbalances in terms of classification data, it can balance errors. RF has become an important analysis method since the algorithm was proposed. It can be widely used in other fields, including genetic data, nuclear magnetic resonance spectroscopy, land classification, etc. On this basis, the study proposed a new classification method for epilepsy EEG signals based on EEMD and RF.

Approximate entropy can be used as a feature to predict in chaotic time series analysis while energy entropy is used a feature in the application of gear fault diagnosis and identification for effective classification.[20,21] This showed that it can be obtained from different data the characteristics of the optimal classification are not fixed. This study selects multiple characteristics for combination and classification, and the optimal classification result is selected.

Through an empirical analysis of epilepsy EEG data from the University of Bonn, Germany, combining the adaptive decomposition characteristics of EEMD with the advantages of RF anti-interference and avoiding overfitting. The correct recognition rate of the classification results of the EEMD-RF algorithm is compared with the classification results of the EEMD-LSSVM algorithm. The results show that the EEMD-RF method is significantly better than the EEMD-LSSVM method in classifying the EEG signals between seizures and seizures. This method can be applied to analyzing and processing biomedical signals and many other fields, such as financial data, mechanical fault diagnosis, environmental early warning, etc. It provides feasible ideas and models for studying classification, identification, early warning, and forecasting in different fields.

Conclusions

There is a large amount of physiological and pathological information in EEG signals. The recognition and classification of EEG signals play a vital role in the detection and adjuvant treatment of epilepsy diseases. The analysis and display of EEG signals in this study are related to the EEMD-LSSVM method. Compared with the classification results, the EEMD decomposition of the epileptic EEG signals is performed, and the energy entropy and information entropy of IMF1~IMF5 are selected as features, and the characteristics of the epileptic seizure inter-seizure and epileptic seizure EEG signals are identified and classified by the realization of RF classification. The accuracy of the results is higher, showing a correct recognition rate of 99.60%. The EEMD-RF method has a better effect on classifying EEG signals between seizures and seizures. The method proposed in this study has the advantage of automatic monitoring of epilepsy. Important theoretical significance and practical value, it can prompt and warn patients in time and provide scientific decision-making support for medical staff.

References

1 Gordon K L, McComas A J & John W O, *Clinical Electroencephalography*, Butterworth-Heinemann, 2013.

2 Hans B, Über das elektrenkephalogramm des menschen *DMW*, **60(51)** (1934) 1947–1949.

3 Michal T, Fundamentals of EEG measurement, *Meas Sci Rev*, **2(2)** (2002) 1–11.

4 Jones G, Bouwer M S & Gotman J, EEG slow waves and memory performance during the intracarotid amobarbital test, *Epilepsia*, **35(1)** (1994) 61–69.

5 Saab R, McKeown M J, Myers L J & Gharbieh R A, A wavelet based approach for the detection of coupling in EEG signals, *Proc 2nd Int IEEE EMBS Conf Neural Eng*, 2005, 616−620, doi: 10.1109/CNE.2005.1419700.

6 Gabor A J & Masud S, Automated interictal EEG spike detection using artificial neural networks, *Electroencephalogr Clin Neurophysiol*, **83(5)** (1992) 271–280.

7 Ocak H, Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy, *Expert Syst Appl*, **36(2)** (2009) 2027–2036.

8 Oweis R J & Enas W A, Seizure classification in EEG signals utilizing Hilbert-Huang transform, *Biomed Eng* **10(1)** (2011) 1–15.

9 Li S, Zhou W, Yuan, Geng S & Cai D, Feature extraction and recognition of ictal EEG using EMD and SVM, *Comput Biol Med*, **43(7)** (2013) 807–816.

10 Wang Y, Liu Z, Zhou Q & Chen X, Wavelet packet entropy analysis of resting state electroencephalogram in sleep deprived mental fatigue state, *Int Conf Human-Computer Interaction*, Springer, Cham, (2019)

11 Chen S, Zhang X, Chen L & Yang Z, Automatic diagnosis of epileptic seizure in electroencephalography signals using nonlinear dynamics features, *IEEE Access* **7** (2019) 61046–61056.

12 Celka P, Boualem B & Paul C, Preprocessing and time-frequency analysis of newborn EEG seizures, *IEEE Eng Med Biol Mag*, **20(5)** (2001) 30–39.

13 Anderson C W, Erik A S & Sanyogita S, Multivariate autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks, *IEEE. Trans Biomed Eng,* **45(3)** (1998) 277–286.

14 Jandó G, Ralph M S, Zsolt H & Buzsáki, G, Pattern recognition of the electroencephalogram by artificial neural networks, *Electroencephalogr Clin Neurophysiol*, **86(2)** (1993) 100–109.

15 Fisch B J & Spehlmann R, EEG primer: basic principles of digital and analog EEG, *Elsevier Health Sci J*, (1999) 23–27.

16 Ying L I & Zhaocheng L V, Recognition of epileptic EEG signals based on EEMD and LS-SVM, *J Hunan Univ Nat Sci* (2014).

17 Lu P, Yuan D, Lou Y & Liu C, Single-trial identification of motor imagery EEG based on HHT and SVM, *Proc 2013 Chinese Intell Automat Conf* (Springer, Berlin, Heidelberg) (2013).

18 Mahesh P, Random Forest classifier for remote sensing classification, *Int J Remote Sens*, **26(1)** (2005) 217–222.

19 Farrikh A, R Arya R M & Fanani A Z, Diabetic retinopathy grade classification based on fractal analysis and random forest, *2019 Int Seminar Appl Technol Info Commun (iSemantic)* IEEE, (2019).

20 Zhang Y, Liu K, Qin L & An X, Deterministic and probabilistic interval prediction for short-term wind power generation based on variational mode decomposition and machine learning methods, *Energy Convers Manag*, **112** (2016) 208–219.

21 Zhang C, Jian J C & Xun G, A gear fault diagnosis method based on EMD energy entropy and SVM, *Shock Vib*, **29(10)** (2010) 216–220.