ISBN: 978-981-18-5852-9
WCSE 2022 Spring Event: 2022 9th International Conference on Industrial Engineering and Applications
doi: 10.18178/wcse.2022.04.166

Classification and Recognition of EEG Signals Based on AdaBoost Algorithm

YanranWang[1], Zhijun Li[1], Rui Zhang[1], Keyong Deng[2]

[1]School of AutomationWuhan University of Technology

[2.]CCCC Second Harbor Consultants Co,Ltd

Abstract. EEG signals directly and objectively reflect the activity of the human body, and are widely used in the field of human movement intention recognition. In order to realize the recognition of the movement intention, firstly, the ERD/ERS phenomenon that occurs when the human body performs a certain side movement or imaginary movement is determined by three methods based on autoregressive model coefficients, based on wavelet packet decomposition coefficient energy and based on co-space mode. Perform feature extraction. After feature extraction, in order to achieve classification, an ensemble learning classification method combining linear discriminant analysis (LDA) and adaptive lifting method is proposed. LAD is used as the weak classifier of AdaBoost to construct a strong classifier for recognition. Compared with the single classification method of the ordinary LDA algorithm, it is proved that the method of using AdaBoosting to strengthen the learning makes the classification accuracy of the motor imagery EEG signal significantly improved. The experimental results show that the use of AdaBoosting-LDA for EEG signal fusion features can achieve a maximum accuracy of 90.7%.

Keywords: EEG signal, Feature extraction, LDA, AdaBoosting

1. Introduction

Since the beginning of the 21st century, my country's population age structure has undergone major changes. Due to the increase in life expectancy and the decline in the birth rate, the proportion of the elderly in my country has reached a high level. Because of the decline in human body functions during the aging process, the elderly need to use auxiliary equipment, such as wearable robots or mechanical exoskeleton, in order to live a better life. There are also many people with disabilities who also need assistive devices in order to live a normal life[1].

In the past human-computer interaction, the signal and instruction are usually the human's intention of movement, that is, the machine judges what kind of action the person is about to complete by judging the action made by the person in the early stage of the movement, and judges the person's movement intention.

But for the elderly or the disabled with weak or lack of mobility, it is obviously not suitable to judge the intention of the movement through the movement trend. In the industrial field, in some extreme operating scenarios, judging intentions based on movement trends may also bring risks[2]. Therefore, current research focuses on EEG signals.

The LDA algorithm is a simple and easy-to-implement classification algorithm. The principle of the algorithm is to assume that there are two types of data. The feature vector of the data is mapped to a straight line or a surface by projection, so that the intra-class distance between the two classification data is minimized. The distance between the classes is the largest to complete the discrimination of the two classifications.

But this is a one-time classification. The classifier does not care about the quality of the classification results, and can only be judged by humans. The ensemble learning method provides a new idea for classification and recognition. It combines multiple weak Classifiers are assembled to form a strong classifier method, which aims to improve the accuracy of classification and recognition. Boosting is a typical integrated learning method, and the AdaBoost algorithm is developed on this basis[3,4].

+ Corresponding author. Tel.: +15621592658; fax:.
E-mail address:447027259@qq.com.

Based on this, this project mainly studies the classification and recognition of EEG signals through the AdaBoosting algorithm. The classification and recognition results are largely related to the selected classification algorithm, so it is necessary to choose a suitable classification algorithm. This article will focus on the LDA classification algorithm, and then will study the method based on AdaBoost ensemble learning, using LDA as the weak classifier of AdaBoost to construct a strong classifier for classification and recognition.

2. Quantification Method of EEG Signal

The brain electrical signal generated by simulating exercise through mental activity is the brain electrical signal of motor imagination.

Studies have shown that when a person moves or imagines a certain side of the body, the amplitude of rhythms and waves in the sensory cortex of the brain on the opposite side will be significantly reduced, while the corresponding rhythm and waveform amplitudes on the ipsilateral side of the brain that are moving It will be significantly enhanced. This is the ERD/ERS phenomenon[5,6], which is an important basis for the classification of brain power in motor imagination.

ERD/ERS is an energy percentage value, which is used to measure the increase or decrease of power. The calculation method is shown in formula (1).

$$\mathrm{ERD/ERS} = \frac{A-R}{R} \times 100\% \tag{1}$$

Among them, A represents the power value of the specific frequency band of the acquisition point of the same motor imagination type in the experiment, and R is used as a reference baseline, which represents the energy value of the specific frequency band of the acquisition point some time before the motor imagination task occurs[7]. From the expression analysis, it can be seen that when ERD/ERS > 0, it means the energy increase after the motor imaging task, and when ERD/ERS < 0, it means the energy decrease after the motor imaging task. These are all in line with the ERD/ERS phenomenon that occurs when the motor imaging task occurs.

The preprocessing method of EEG signals used in this article is to extract specific frequency bands from the selected EEG data, and the fusion features are used in the final step of classification and recognition.

3. Classification and Recognition of EEG Signals Based on AdaBoost algorithm Fused with LDA

At present, there are many algorithms in the field of EEG signal classification and recognition, such as linear discriminant analysis, artificial neural network, K Nearest Neighbor (KNN) and so on.

ANN:The artificial neural network is an adaptive system. It cascades through the topology structure, and the number of layers of the network structure is selected according to the actual usage. It achieves the purpose of classification by calculating the weight and bias of the hidden layer.

KNN:The K-nearest neighbor method makes classification decisions by calculating the distance between the test set and the training set eigenvalues of the known classification categories, and then finding the K nearest samples of the test set in the training set. KNN needs to calculate the similarity of all samples in the test set and training set in turn, so the increase of data samples will also increase the computational complexity.

LDA:Linear discriminant analysis is a linear learning method, which is widely used in the field of pattern recognition of human electrical signals. Its basic principle is to use the criterion function and the corresponding optimization algorithm to project the training set samples with two types of data from the high-dimensional space to the low-dimensional space in the same direction, so that the same type of data sample points after projection. The distance is the smallest, and the distance between different types of data sample points must be maximized. In this way, a linear discriminant function is generated. Finally, the test set uses this linear discriminant function to achieve the purpose of classification.

After the EEG signal features are extracted, classification and identification are required. Because the LDA algorithm runs fast, the algorithm is simple and easy to implement, and it is suitable for the classification of few samples. It is widely used in the field of pattern recognition of human electrical signals. Therefore, LAD is used as the weak classifier of AdaBoost to build a strong classifier. identify.

3.1.LDA Algorithm Classifiers

The principle of the LDA algorithm is to assume that there are two types of data. The feature vector of the data is mapped to a straight line or a surface by projection, so that the distance between the two categories of data is the smallest and the distance between the categories is the largest, so as to complete the discrimination of the two categories, The schematic diagram of the implementation is shown in Figure1.

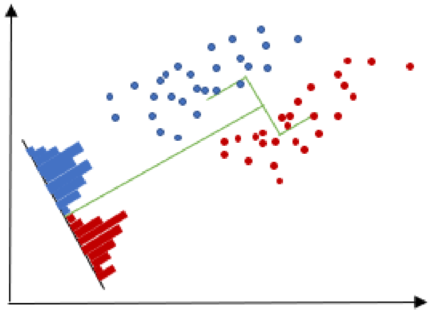

Fig. 1:Principle of LDA algorithm.

The following will study the specific algorithm implementation steps.

Assuming a set of data x is an n -dimensional vector, namely $x=[x_1,x_2,\cdots,x_n]$, the expression of its linear discriminant function can be listed as[8]:

$$g(x)=\omega_1 x_1+\omega_2 x_2+\cdots+\omega_n x_n+\omega_0=W^T x+\omega_0 \tag{2}$$

Assuming that there is a two-classification task, the research and analysis of the decision-making mechanism are carried out by formula:

$$g(x)=g_1(x)-g_2(x) \tag{3}$$

When $g(x)=0$, it means x on the decision plane of the two types of tasks, which can be divided into any one or not, when $g(x)>0$, it is judged as l_1, when $g(x)<0$, it is judged as l_2. To represent the two types of tasks respectively.

The design of the linear discriminant analysis classifier is to find the optimal W and ω_0 to achieve effective classification.

3.2.Construct a Fisher linear discriminant classifier for EEG signals

Assuming there is a two-category sample data set $X=\{x_1,x_2,\cdots,x_n\}$, the total number of samples is N, and it is divided into two categories, l_1 and l_2, one of which occupies N_1 samples, the other occupies N_2 samples, define the average of each type of sample μ_1, namely the center point of the sample is shown in equation(4).

$$\mu_i=\frac{1}{N_i}\sum_{x_k\in l_i} x_k, i=1,2 \tag{4}$$

After projection, get the center point $\tilde{\mu}_i$.

$$\tilde{\mu}_i=W^T\mu_i \tag{5}$$

In the formula, N_i represents the number of samples of type i, and then the intra-class dispersion matrix of the samples is obtained:

$$s_i=\sum_{x_k\in l_i}(x_k-\mu_i)(x_k-\mu_i)^T, i=1 \tag{6}$$

The overall intra-class dispersion matrix is the sum of the two kinds of dispersion, as shown in formula(7).

$$S_\omega = S_1 + S_2 \tag{7}$$

The dispersion between classes is:

$$s_b = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T \tag{8}$$

The Fisher criterion expects that the difference between the sample means of the two categories takes the maximum value, and the difference between the sample means within each category is expected to take the minimum value, in order to achieve the purpose of linear discriminant analysis, that is, the smaller S_i is, the better, and the larger S_b is, the better .

Then $\tilde{S}_1 = W^T S_1 W$, $\tilde{S}_2 = W^T S_2 W$, suppose the Fisher criterion function is as shown in formula (9).

$$J(W) = \frac{(\tilde{\mu}_1 - \tilde{\mu}_2)^2}{\tilde{S}_1 + \tilde{S}_2} \tag{9}$$

Through the above analysis, the optimal solution can be obtained by maximizing the value of $J(W)$.

Arrange the above formula to get:

$$J(W) = \frac{W^T S_b W}{W^T S_\omega W} \tag{10}$$

Among them, W is the direction vector of the projection, and S_b and S_w represent the dispersion matrix between and within the class, respectively.

Finally, the projection direction vector of the sample is calculated:

$$W^* = S_\omega^{-1}(\mu_1 - \mu_2) \tag{11}$$

In the end, the value of threshold w_0^* has not yet been determined, but the Fisher criterion cannot obtain the solution of the optimal threshold w_0^*. Therefore, approximate estimation methods are generally used, and there are three formula calculation methods:

$$\omega_0 = -\frac{W^T(\mu_1 + \mu_2)}{2} \tag{12}$$

$$\omega_0 = -\frac{W^T(N_1\mu_1 + N_2\mu_2)}{N} \tag{13}$$

$$\omega_0 = -\frac{W^T(\mu_1 + \mu_2)}{2} - \frac{\ln\left(\frac{N_1}{N_2}\right)}{N_1 + N_2 - 2} \tag{14}$$

Here, this article uses the commonly used formula (14) to estimate the threshold w_0.

After constructing the classifier, the Data set III data set can be used for testing.

The characteristics of the data set III data set motion imaging EEG signal extracted in this paper are: the coefficients of the AR model, the energy features related to WPD and the features extracted by the CSP method, as well as the features of the fusion of the WPD and CSP methods. For the feature vector composed of the above Carry out the classification verification of LDA respectively. The Data set III data set has two parts: a test set and a training set, each occupying 140 groups. First, use the training set and the training label to obtain the classified model, and then use the test set to output the corresponding motor imagination EEG category through the training model, and then Compare with the label of the test set to get the accuracy of classification. The classification results are shown in Table 1.

Table 1: Linear Discriminant Classification Result

Extracted features	Accuracy(%)
Coefficients of AR model	67.14
WPD coefficient energy, energy percentage and energy relative deviation	85.00
Features extracted by CSP	86.43
Features of the fusion of WPD and CSP	85.71

It can be observed from Table 1 that the CSP feature extraction method has the highest accuracy rate. The fused feature has a lower accuracy rate than the single CSP method. However, the fused feature has a higher accuracy rate than the single WPD method. It shows that it has a positive effect on one of the feature extraction methods. This also shows that the data fusion feature studied in this paper is not suitable for the LDA classification method alone. It can be seen that the method with the lowest recognition accuracy obtained here is the coefficients of the AR model, and the fusion of the coefficients of the AR model with the features extracted by WPD and CSP will also reduce the accuracy to a certain extent. Therefore, after experimental research, and The coefficients of the AR model are not put into the fusion features for LDA classification and recognition.

3.3.Boosting Classifier Construction

Boosting is a method of assigning a set of weak classifiers to certain weights, and then superimposing the weak classifiers with weights to form a strong classifier.

The implementation method is as follows:

Given an initial training set, classify each sample in the training set, and then assign an initial weight to each class of samples for classification, and use the constructed weak classifier for classification.

After a round of classification, find out the wrong samples, and give higher weight to such samples, so that the strategy is cycled until the result reaches the set number of times.

At this time, the initial weak classifier already has the weights given by multiple iterations, and the strong classifier can be obtained by superimposing these weak classifiers.

The formula is as follows:

For an original training set sample $\{(x_1;y_1),(x_2;y_2),\cdots,(x_n;y_n)\}$, set y_1 to indicate the category label of the i sample x_i, and set a total of T cycles. After the t cycle, the corresponding weight of each sample is set to a_t, and the default The weak classifier is g_t, and for the Boosting algorithm, it is necessary to ensure that the classification estimate of the sample for each iteration of the classifier is greater than one-half, as shown in Equation(15) .

$$P_{(x,y)\in a_i}\left[y=g_t\left(x\right)\right]>\frac{1}{2}+\varepsilon_t,\varepsilon_t\in\left[0,1\right) \tag{15}$$

Finally, a strong classifier can be obtained:

$$G\left(x_i\right)=\operatorname{sgn}\left(\sum_{t=1}^{T}a_i g_t\left(x_i\right)\right) \tag{16}$$

3.4.Introduction to AdaBoost Classification Method

The AdaBoost algorithm is developed on the basis of the Boosting algorithm. The AdaBoost algorithm performs optimization iterations from the data, and finally superimposes the weak classifiers to form a strong classifier.

Next, I will explain how to integrate the LDA algorithm and use the AdaBoost algorithm to build a classifier of EEG signal feature values.

1)There is a two-category training sample data set $\{(x_1;y_1),(x_2;y_2),\cdots,(x_n;y_n)\}$, and the total number of samples is N, $x_i\in R$,where x_i represents the i feature vector of the training data set, and $y_i\in\left(-1,1\right)$ represents the training label corresponding to the i feature vector.

First, the weight distribution of the training set samples needs to be initialized, and the weight distribution of the first training sample is set to D_1. According to the algorithm definition, each type of training sample needs to be assigned the same weight at the beginning, as shown in Equation (17).

$$D_1 = (\omega_{11}, \omega_{12}, \ldots \omega_{1i}, \ldots \omega_{1N}), \omega_{1i} = \frac{1}{N}, i = 1, 2, \ldots, \mathrm{N} \tag{17}$$

Suppose a total of T iteration has been carried out, t represents the t iteration, the weight of the t iteration can be set to D_t, and another classifier with the lowest classification error rate is set to $G_t(x): x \to \{-1, 1\}$, which can be calculated to use this classifier The error rate of classification is shown in Equation (18).

$$e_t = P(G_t(x_i) \neq y_i) = \sum_{i=1}^{N} \omega_{ti} I(G_t(x_i) \neq y_i) \tag{18}$$

2) Next, calculate the coefficient a_t of the weak classifier 2 in the t iteration, which represents the weight of the final classifier, as shown in equation(19).

$$a_t = \frac{1}{2} \log \frac{1 - e_t}{e_t} \tag{19}$$

It can be seen from formula (19) that when $e_t \leq \frac{1}{2}$, $a_t \geq 0$, and when e_t is reduced, a_t will increase, which shows that when there is a minimum error rate in a certain iteration, the weight obtained by the classifier has a maximum value.

After the t iteration of training, it is necessary to calculate the weights required for the next iteration, that is, the weights need to be updated, and the weights are distributed:

$$D_{t+1} = (\omega_{t+1,1}, \omega_{t+1,2}, \ldots, \omega_{t+1,i}, \ldots \omega_{t+1,N}) \tag{20}$$

$$\omega_{t+1,i} = \frac{\omega_{t,i}}{Z_T} \exp(-a_t y_i G_t(x_i)), i = 1, 2, \ldots, \mathrm{N} \tag{21}$$

Among them, Z_t represents the normalization factor, so that D_{t+1} is a probability distribution, and the calculation method of Z_t is shown in Equation (22).

$$Z_t = \sum_{i=1}^{N} \omega_{t,i} \exp(-a_t y_i G_t(x_i)) \tag{22}$$

3) The weight of the iteration corresponding to the misclassified sample will increase, and the weight of the iteration corresponding to the correctly classified sample will decrease. This is also the principle that AdaBoost pays more attention to the more difficult samples.

Finally, the weak classifiers are linearly combined to form a strong classifier, as shown in Equation (23).

$$G(x) = \mathrm{sgn}\left(\sum_{t=1}^{T} a_t G_t(x)\right) \tag{23}$$

3.5.Analysis of Results

The AdaBoost-LDA method is used to classify and identify the extracted feature vectors, and the number of loop iterations is set to 50. The obtained recognition accuracy is shown in the figure.

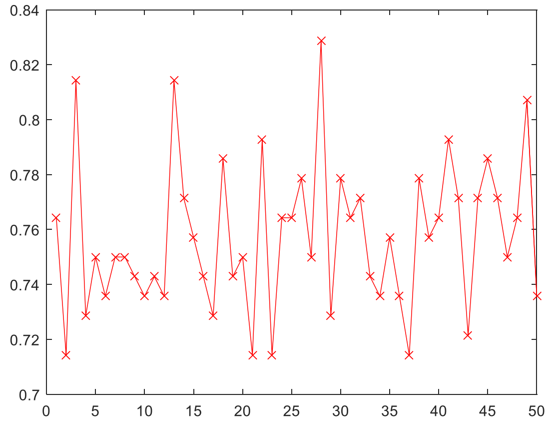

Fig. 2: AdaBoost-LDA classification and recognition based on AR model coefficient features.

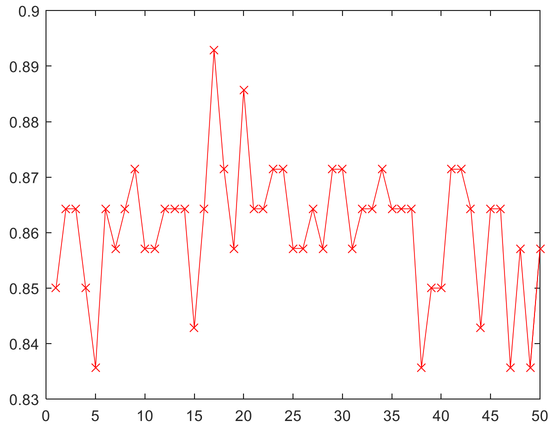

Fig. 3: AdaBoost-LDA Classification and Recognition Based on WPD

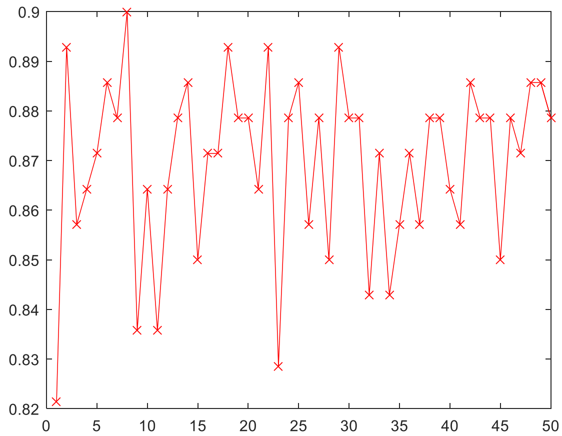

Fig. 4: AdaBoost-LDA Classification and Recognition Based on CSP

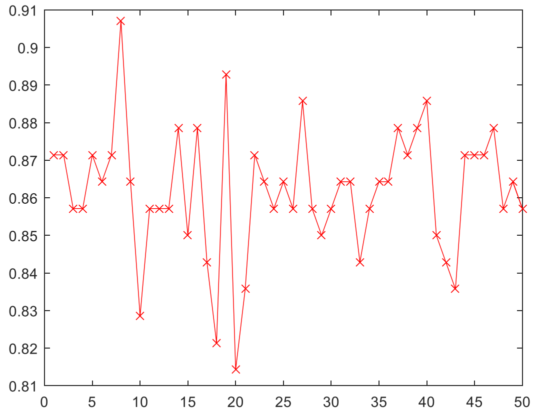

Fig. 5: AdaBoost-LDA Classification and Recognition Based on WPD-CSP

Table 1 ：CLASSIFICATION RESULTS OF DIFFERENT CHARACTERISTICS AND CLASSIFICATION METHODS

Classi-fication \ feature / Accuracy (%）	AR	WPD	CSP	WPD-CSP
LDA	67.14	85.00	86.43	85.71
AdaBoost-LDA	82.86	89.29	90.00	90.71

4. Conclusions

Introduce the ensemble learning AdaBoost algorithm, and use LDA to construct a strong classifier for classification and recognition. It is found that comparing the single LDA classification methods respectively, the recognition accuracy has improved to varying degrees. For the coefficient features of AR models that are difficult to classify by LDA, the AdaBoost-LDA method can significantly improve the accuracy. Therefore, the AdaBoost-LDA method can significantly improve the accuracy of features that are difficult to classify by a single classification method. For WPD-CSP fusion features, AdaBoost's ensemble learning method also has obvious advantages. The use of AdaBoost-LDA fusion features has higher accuracy than single feature recognition, and the combination of WPD-CSP feature fusion and AdaBoost-LDA classification method It has the highest recognition accuracy rate, which can reach 90.71%.

5. Acknowledgements

This work was supported by the National Key R&D Program of China under Grant NO.2017YFB130050 2.

6. References

[1] XingguoLong.ResearchonElectroencephalogram-basedControlMethodologiesforExoskeletonRobot[D]. University of Chinese Academy of Sciences (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences),2020.

[2] Ding Qi-Chuan, Xiong An-Bin, Zhao Xin-Gang, Han Jian-Da. A Review on Researches and Applications of sEMG-based Motion Intent Recognition Methods [J]. Acta Automatica Sinica,2016,42(01):13-25.

[3] Bandarabadi M,Dourado A,Teixeira C A,et al. Seizure prediction with bipolar spectral power features using Adaboost and SVM classifiers[C]//Engineering in Medicine & Biology Society. IEEE, 2013:6305-6308.

[4] Donos Cristian,Dümpelmann Matthias,Schulze-Bonhage Andreas. Early Seizure Detection Algorithm Based on Intracranial EEG and Random Forest Classification.[J]. International journal of neural systems,2015,25(5):155-159.

[5] Gert Pfurtscheller,Christa Neuper. Future prospects of ERD/ERS in the context of brain–computer interface (BCI) developments[J]. Progress in Brain Research,2006,159:433-437.

[6] Neuper C,Schlögl A,Pfurtscheller G. Enhancement of left-right sensorimotor EEG differences during feedback-

regulated motor imagery.[J]. Journal of clinical neurophysiology:official publication of the American Electroencephalographic Society,1999,16(4):373-382.

[7] G. Pfurtscheller,F.H. Lopes da Silva. Event-related EEG/MEG synchronization and desynchronization: basic principles[J]. Clinical Neurophysiology,1999,110(11):1842-1857.

[8] R.Duda,P.Hart,D.Stork.Pattern Classification[D]. New York:wiley 2001.