EEG Recognition of Depressive Mood Tendency and Depression Based on Improved SVM-RFE Algorithm

Jingjie Ma
Portland Institute
University of Posts and Telecommunications
Nanjing, 210023, China

Shuting Zhu
Portland Institute
University of Posts and Telecommunications
Nanjing, 210023, China

Yang Liu
School of Computer Science Software and Cyberspace Security
Nanjing University of Posts and Telecommunications
Nanjing, 210023, China

Ruoyu Du*
School of Communication and Information Engineering
Nanjing University of Posts and Telecommunications
Nanjing, 210023, China
*Corresponding author's email: dury@njupt.edu.cn

Xin Xu*
School of Communication and Information Engineering
Nanjing University of Posts and Telecommunications
Nanjing, 210023, China
*Corresponding author's email: xuxin@njupt.edu.cn

***Abstract*—To better utilize accurate data in measuring depressive mood, this study aims to provide an improved diagnostic reference for confirming and assessing the extent of depression. In this study, a method of electroencephalogram (EEG) signal recognition based on improved Support Vector Machine-Recursive Feature Elimination (SVM-RFE) algorithm is proposed. Firstly, a comprehensive EEG dataset called MO-DEAP (MODMA and DEAP) was constructed, including samples with depressive mood tendency and major depression. Secondly, SMOTE (Synthetic Minority Over-sampling Technique) and ADASYN (Adaptive Synthetic Sampling) were used to process the minority class samples to improve recognition rate. Finally, 11 linear and non-linear features were extracted and used in SVM-RFE. This feature selection method resulted in the identification of an optimal feature subset and improved classification recognition. In addition, the optimal channels were determined by comparing the median and outlier values of six (FP1, FP2, O1, O2, T7, T8) channels between the two groups in MO-DEAP. The analysis revealed that SMOTE can better differentiate between depression patients and individuals with a tendency for depressive mood. The accuracy of DFA and SampEn features in distinguishing depression patients from individuals with a tendency for depressive mood in the training set reached 92.86%. Moreover, SampEn differs significantly across all channels, whereas DFA differs most significantly across FP2, O1, and O2 channels.**

Keywords—EEG signal; Depression; Tendency for depressive mood; SMOTE; SVM-RFE

I. Introduction

As a prevalent mental illness, depression has become the second leading cause of disability and poor health.[1] In fact, depression is not incurable. Diagnosing and treating depression early improves the chances of cure. Depression Tendency is a mental state that does not meet the diagnostic criteria for depression. But if an individual lacks timely diagnosis and treatment, it may seriously affect his or her mental health and quality of life. Therefore, finding effective methods to identify depressive mood tendency and depression can help provide personalized support and treatment options. The diagnosis of most depression still relies on the subjective judgment of clinicians and psychologists, and there is a lack of objective physiological indicators to distinguish individuals with depressive mood tendency from depression patients. EEG signals are generated by the human central nervous system and can exhibit characteristics that objectively reflect emotional states. In various research and diagnostic applications related to brain diseases, EEG signals are widely used due to their low cost and relatively simple recording process. Therefore, studying the characteristics of EEG signals is an important approach to accurately and efficiently distinguish individuals with depressive mood tendencies from depression patients.

There is a certain research basis for distinguishing individuals with depressive mood tendency from depression patients based on EEG signal characteristics. Different channels of EEG signal selection and configuration are used to record and analyze specific EEG signals. Hanshu Cai and others showed in the study that the electrode site located in the prefrontal cortex is related to emotional processes and mental disorders. Therefore, Fp1, Fp2 and Fpz are ideal choices for scalp locations in experiments[2]. The occipital lobe channels O1 and O2 were selected by Mahato[3] to extract EEG features. Additionally, numerous studies related to depression and EEG have utilized both linear and nonlinear features for the identification and classification of EEG signals. Adil O. Khadidos[4] achieved 84.09% accuracy using Detrended Fluctuation Analysis (DFA). Bachmann[5] chose the spectral asymmetry index (SAI) and DFA, and the classification accuracies were 76.5% and 70.6% respectively. K.Kalev[6] selected the Lempel-Ziv complexity (LZC) feature and the classification accuracy in the F7 channel reached 77.27%. Laura Minkowski[7] extracted Higuchi fractal dimensions, correlation dimensions, approximate entropy (ApEn), Lyapunov exponent and DFA. Mahato[8] classified SampEn with an accuracy of 80.60%. Based on the above characteristics of patients with depression, this study will select prefrontal lobe channels FP1, FP2, occipital lobe channels O1 and O2, temporal lobe channels T7 and T8[8] as extraction channels of EEG characteristics. For feature extraction, nonlinear features such as DFA, LZC, SampEn, ApEn, correlation dimensions, along with linear features in the time-frequency domain, are utilized to construct feature matrices for the two groups, aiming to identify the optimal approach to distinguish individuals with depressive mood tendencies from depression patients.

Preprocessing data through feature selection reduces useless features and maximizes features using classification. Recursive feature elimination based on improved support vector machine (SVM-RFE) is a method that sorts features according to their correlation values to select optimal features. In addition, when classifying data, the number of certain data samples is much smaller than other samples, which will cause the problem of unbalanced data distribution[9]. Therefore, this study aims to employ the SMOTE(Synthetic Minority Over-sampling Technique) and ADASYN (Adaptive Synthetic Sampling) oversampling methods. ADASYN is an extension of SMOTE that can adjust the number of synthetic samples according to the sample density distribution[10]. However, it has the disadvantages of high computational cost and poor processing effect on high-dimensional data sets. Therefore, the above two methods need to be analyzed in specific applications and the oversampling algorithm with better data processing effect should be selected. In this study, SMOTE, which has better classification effect, is combined with SVM-RFE algorithm. This feature selection method has the advantages of faster and more accurate.

To achieve this goal, this article applies an improved SVM-RFE binary classification algorithm and the SMOTE oversampling method, and proposes a method to identify and classify individuals with depressive mood tendencies and depression patients based on EEG signal characteristics.

II. COMPREHENSIVE EEG DATASET MO-DEAP AND EEG PROCESSING

A. Introduction of SVM-RFE Classification Algorithm

The DEAP data set is based on EEG data generated under stimulation induced by music video materials, and records 32 channels of EEG data. This study used the Arousal and Valence data in this data set, that is, the evaluation of Arousal and Valence made by 32 subjects on 40 1-minute music videos they watched.

The MODMA dataset is a publicly available dataset of patients with severe depression, recording 25 minutes of data from 128 EEG channels under stimulation. This data set consists of EEG data from 53 volunteers, including 24 patients with depression and 29 normal subjects. There is no significant difference in gender and age between patients with depression and normal subjects.

B. EEG Signal Preprocessing

In order to eliminate the influence of environmental interference and other factors on the EEG data and retain the effective EEG components, preprocessing operations were first performed on the DEAP and MODMA data sets. In order to remove the EOG noise, ICA independent principal component analysis was performed on the two sets of data, and down-sampling was used to reduce the frequency to 128 Hz, and its band-pass filtering frequency was between 4-45 Hz. The processing process of the two sets of data is shown in Figure 1.

Firstly, preprocess the DEAP dataset. The DEAP data was screened out for the negative emotional data of 32 patients in the range of $1<Arousal<7.1$ and $2<Valence<6.19$. After removing the 3s baseline for each sample, 40 1min data with depressive mood tendency were obtained, and were assigned the label 1.

Secondly, the MODMA data set was preprocessed and the original file was read using the EEGLAB toolbox in MATLAB. In order to be consistent with the number of channels of DEAP, the positions of 32 EEG channels were selected according to the international 10-20 system. Secondly, 24 MDD patients were screened out according to two labels: MDD and HC. In order to control the duration, the events fcue and scue generated by the cue stimulus were used, and the data was processed as 500ms for each event. The number of samples for both events is 160 times.

Finally, in order to process the EEG data matrices of the two data sets in MO-DEAP into the same form, the 320 samples in MODMA were merged into two one-minute data. The EEG data matrix form is 2*32*8064, the numbers represent the number of samples, the number of channels, and the number of sampling points in turn. The data of depression patients are assigned the label 0.

C. ADASYN+SMOTE Balanced Dataset

In this study, the sample size of individuals with depressive mood tendency from the MODMA data set is smaller than the sample size of depression patients from the DEAP data set. In this case, machine learning cannot distinguish between the two groups of individuals well. Currently, SMOTE (Synthetic Minority Over-sampling Technique) and ADASYN (Adaptive Synthetic Sampling) are both methods used to deal with imbalanced data sets. The SMOTE algorithm mainly generates synthetic samples by interpolation in the feature space, while the ADASYN algorithm is based on SMOTE Development, it requires dynamically adjusting the number of generated synthetic samples[9]according to the density of each minority class sample.

In order to compare the application effects of the two algorithms in binary classification, this study drew the ROC curve obtained after the binary classification of the two algorithms, as shown in Figure 2:

The AUC value of ADASYN is about 0.5, which has almost no discriminating ability and behaves like random guessing; in the two-classification results after SMOTE processing, the AUC value under the LZC and DFA characteristic curves is between 0.5-1, which has better Discrimination ability. Therefore, this study uses the SMOTE algorithm to create synthetic samples[11] in the feature space of minority class samples.

III. EEG FEATURE EXTRACTION

In this study, the signal was divided into 10 bands, and a total of 11 linear and nonlinear features were extracted.

A. Linear Feature Extraction

1) Wavelet transform. Use Daubechies 5 wavelet to carry out Discrete Wavelet Transform (DWT) [12] to decompose the EEG data in the MO-DEAP data set. The signal is decomposed to level 4, and the approximation coefficient (C) and detail coefficient (L) can be obtained:

$$[C, L] = wavedec(data2, 4, 'db5') \tag{1}$$

The low-frequency energy (Ea) and high-frequency energy (Ed) obtained from wavelet decomposition are:

$$[Ea, Ed] = wenergy(C, L) \tag{2}$$

Since most of the EEG data activity of depressive emotions is concentrated in low-frequency energy areas, each individual with depressive mood tendency and depression patients can extract the Ea values of 10 EEG data segments in the 4-45Hz frequency band.

2) Power Spectral Density. This study used the Welch method[13] to extract the power spectral density, used the Pwelch function to calculate the power spectrum, used the Bandpower function to extract the power information of a specific frequency band, and extracted respectively θ,α,β,γ. The signal power of the four rhythm frequency bands; simply stack the features to obtain a power spectral density matrix of 10×4, that is, each individual with depressive mood tendency and depression patients can extract 40 frequency domain features.

B. Nonlinear Feature Extraction

1) Sample entropy. Sample Entropy (SampEn) is a time series complexity measurement method proposed by Richman[14]. It is calculated as the negative logarithmic conditional probability that two d points with the same sequence remain the same for the next point[15], as shown in formula (3) (4):

$$SE = ln\varphi^m(t) - ln\varphi^{m+1}(t) \tag{3}$$

$$\varphi^m(t) = \frac{1}{n-m+1}\sum_{i=1}^{n-m+1}\left[\frac{1}{n-m}num\{d[x(i), x(j)] < t\}\right] \tag{4}$$

This study used SampEn as a feature to distinguish individuals with depressive mood tendency to depression patients.

2) Detrended fluctuation analysis. Detrended Fluctuation Analysis (DFA) is calculated from the perspective of the time domain, and the relationship α between the root mean square coefficient and each time period is used as the characteristic value. The value of α in [0.5,1] indicates that the sequence is dependent, and the value of α in [0.5,1] indicates that the sequence is correlated. DFA of the oscillation amplitude of the EEG data of the two groups of subjects[15] as shown in equation (5) (6):

$$F(\tau) = \sqrt{\frac{1}{N}\sum_{n=1}^{N}[y(n) - y_\tau(n)]^2} \tag{5}$$

$$log[F(\tau)] = \lambda log(\tau) + log(C) \tag{6}$$

where λ is defined as DFA, and this index is used to characterize the relationship α between the square root coefficient and each time period.

3) Lempel-Ziv complexity. Lempel [16] proposed Lempel-Ziv Complexity (LZC) in 1976. This parameter is characterized by measuring the rate at which new patterns appear in a time series. The degree of disorder in a random signal.

The input is the time series data of two groups of subjects {X(1), X(2),…, X(N)}

$$LZC = \frac{C(N)}{N/log_2 N} \tag{7}$$

As in equation (7), C(N) = 2, N is the number of subjects.

Approximate entropy. Approximate Entropy (ApEn) was proposed by Pincus [17] in 1991. It is an algorithm that quantifies the proportion of unpredictability of time series signals. ApEn is defined as follows:

$$ApEn = \frac{1}{N-m}\sum_{i=1}^{N-m}\left[ln\frac{C_i^{m+1}(r)}{lnC_i^{m+1}(r)}\right] \tag{8}$$

4) Correlation dimension. The relevant dimension adopts the most widely used Grassberger and Procaccia algorithm (G-P algorithm for short)[18]. Its algorithm is defined as:

$$D_v(M,\tau) = \frac{\ln C_M^\tau(r_b) - \ln C_M^\tau(r_a)}{\ln(r_b) - \ln(r_a)} \tag{9}$$

As in equation (9), r_a , r_b are used to calculate two different distance thresholds for the fractal dimension, where r_a represents a smaller distance, r_b represents a larger distance.

5) C0 complexity. C0 complexity was proposed by Shen et [19] and others, and is used to quantify the complexity of time series. For a time series x(n) containing N samples, its C0 complexity can be expressed as:

$$C0 = \frac{A1}{A0} = \frac{\sum_{n=0}^{N-1}|x(n) - y(n)|^2}{\sum_{n=0}^{N-1}|x(n)|^2} \tag{10}$$

As in equation (10), A1 and A0 are the irregular components and regular components of x(n) respectively.

C. Summary

Therefore, this study obtained 5 linear features and 6 nonlinear features, for a total of 11 EEG features.

The EEG data of individuals with depressive mood tendency and depression patients are redistributed through the SMOTE non-equilibrium algorithm, and a minority class of depressive samples that retain the characteristic information of the data are synthesized, which improves Single-channel Performance and generalization ability of nonlinear SVM depression recognition model.

IV. NONLINEAR SVM-RFE CLASSIFICATION ALGORITHM

In order to further select the features that distinguish depressed patients from those with depressive mood tendency, this study proposed a feature selection and classification method: Support Vector Machine - Recursive Feature Elimination (RFE) classification algorithm.

A. Introduction of SVM-RFE Classification Algorithm

Currently, there are three main methods for evaluating the importance of variables in SVM: filter, wrapper, and embedded method. Sanz et al [20] proposed a variable selection wrapper method for evaluating a specific subset of variables by training and testing a specific classification model, which is known as SVM-RFE. Due to the experimental selection of more nonlinear EEG features, the algorithm is improved in this study to better capture nonlinear relationships. The method is applied to the nonlinear Radical Basis Function (RBF) kernel to classify depression patients and individuals with depressive mood tendency. The feature selection and classification steps of the improved algorithm are shown in Figure 3:

B. Machine Learning Classification Result.

In this study, we define that the number of subsets is 5 and the selected features of each subset is 2,4,6,8,10. Also, the selected features in each subset is determined by comparing the accuracy. Then, the nonlinear SVM-RFE classification algorithm obtains five subsets with highest accuracy and the output accuracy of each subset is shown below:

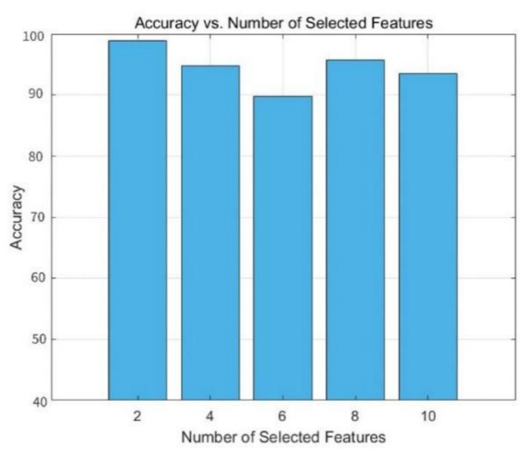

Figure 4. Classification result of SVM-RFE recursive feature elimination

As shown in Figure 4, the highest accuracy is obtained when two features are selected. The output feature subset shows that the two features are DFA and SampEn. Therefore, according to SVM-RFE classification algorithm, the combination of DFA and SampEn is the best features to distinguish between individuals with depressive mood tendency and depression patients. From the feature sort list, it can be obtained that the classification effect of SampEn is more significant compared to DFA.

C. Statistical Classification Result

Based on the DFA and SampEn found in six (FP1, FP2, O1, O2, T7, T8) channels of depression patients and individuals with depressive mood tendency, t-test was performed on two groups respectively. And the results shows that all the above channels were found to be more statistically significant ($P < 0.05$). And further comparison finds that O1, O2 and FP2 channels, i.e., occipital area and right side of prefrontal lobe were more statistically significant for DFA features. Each channel of SampEn has better statistical significance, which could better differentiate between individuals with depressive mood tendency and depression patients. The significance of the DFA features of O1, O2, and FP2 channels could be represented by box plots, and the statistical results were shown in Table 1:

The left and right box represents the median and outlier DFA values of depression patients and individuals with depressive mood tendency respectively. And this graph shows that the DFA values of depression patients in FP2, O1, O2 channels are significantly lower than those of individuals with depressive mood tendency.

D. Multi-Dataset Validation

To verify the applicability of the SVM-RFE algorithm to other datasets, this study applied these three features to each of the three datasets and obtained the accuracy of classifying the two types of patients for each dataset by single-channel nonlinear SVM classification. The three datasets are further described below.

The three datasets are based on MODMA for major depression patients and the healthy population database DEAP based on EEG data. The present study reorganizes the data for the two sets of data to form three datasets containing depressive mood tendency and depression. Details of the datasets are shown in Table 2:

TABLE 2. DETAILS OF THE SUB DATASET

Dataset Classification	Dataset 1	Dataset 2	Dataset 3
Data Volume	32	20	56

The MODMA and DEAP datasets were used as the original training sets, with 24 depression patients in the MODMA dataset and 32 healthy individuals in the DEAP dataset. The depression stimulated EEG data in MODMA and negative mood data of healthy individuals in DEAP were selected as dataset 1, combined with the stimulated state EEG data of the above individuals. Secondly, in the MODMA dataset, according to the score of PHQ-9 Depression Screening Scale of each subject as well as the criteria for judging the scale, patients with depression were categorized into two groups: 16-19 scores (12 mildly depressed patients), and 20-27 points (8 patients with severe depression), as dataset 2 [21]. Finally, the DEAP data with Valence and Arousal features were classified by k-means algorithm, and the intensity of depressive tendency was differentiated by the high and low arousal, and the above data were categorized as dataset 3. After being processed by SVM-RFE, the SVM classification effect of each dataset is obtained as shown in Table 3:

TABLE 3. CLASSIFICATION RESULTS OF SINGLE CHANNEL NONLINEAR SVM DEPRESSION RECOGNITION MODEL

Feature	Positive and negative mood in DEAP	Minor and major extent in MODMA	DEAP+MODMA
SampEn+DFA	84.49/84.36	94.23/95.86	**92.86/92.75**
SampEn	58.95/59.18	83.78/84.33	80.97/80.99
DFA	56.66/57.17	73.05/72.36	78.39/78.25

Note: a and b of a/b denote the accuracy of the training set and the accuracy of the test set, respectively

As can be seen from Table 3, the feature combination of SampEn and DFA is better able to differentiate between the two categories of individuals with depressive mood tendency in DEAP and depression patients in MODMA in the single-channel results, with an accuracy of 92.86% in the training set. Compared to the SampEn and DFA features alone, the classification accuracy of this feature combination in all three datasets is improved, and the classification accuracy of the feature combination in all three datasets is above 80%. This shows that the SVM-RFE classification algorithm selects the feature combination better and is able to maximize the classification accuracy.

For multi-sample dataset 1, the feature combination of SampEn and DFA improves about 30% compared to a single feature. The classification accuracy of SampEn is also relatively higher than that of DFA features, which indicates that SampEn is able to better differentiate between individuals with depressive mood tendency and depression patients in a single channel.

V. DISCUSSION

This study utilizes the MO-DEAP dataset and processes minority class samples by employing the imbalanced algorithms SMOTE and ADASYN to improve the recognition rate of minority class samples in the SVM classification algorithm. The experimental results demonstrate that the AUC under the LZC and DFA feature curves processed with the SMOTE algorithm ranges from 0.5 to 1. Therefore, SMOTE more effectively distinguishes individuals with depressive mood tendency from depression patients. This is primarily because synthetic samples generated by ADASYN do not fit well with the complex decision boundary in certain nonlinear problems and are prone to noise or outliers.

To identify the optimal features for distinguishing individuals with depressive mood tendency from depression patients, SVM-RFE is employed, determining DFA and SampEn as the best feature subset, and the training set accuracy reaches 92.86%. Regarding channel selection, the DFA and SampEn features on the FP2, O1, and O2 channels show the most significant differences. In practical medical diagnosis, this study can utilize the aforementioned features and channels to accurately and objectively distinguish individuals with depressive mood tendency from depression patients, thereby reducing the misdiagnosis rate of depression.

VI. CONCLUSION

Due to the significant work and study pressures faced by modern individuals and the lack of timely emotional counseling, an increasing number of college students and individuals in society exhibit a tendency towards depressive mood. Addressing this social phenomenon, the goal of this study is to identify the optimal features that can differentiate between depression patients and individuals with a tendency towards depressive mood by analyzing the MO-DEAP dataset. The objective is to develop personalized treatment plans for these two groups, addressing their specific challenges and ultimately reducing the prevalence of depression.

Due to the unique characteristics of the population with depression and time constraints, this study did not establish its own experimental dataset. Therefore, the finding of this study have certain limitations. In future research, it will be advantageous to investigate features extracted from different frequency bands and to employ various classifiers to distinguish between different subjects. Additionally, it is critical to classify the severity of depression in patients, as varying levels of depression necessitate different treatment plans.

REFERENCE

[1] Bairy G, Niranjan U.C., and Puthankattil S.D. (2016). Automated classification of depression EEG signals using wavelet entropies and energies. Journal of Mechanics in Medicine and Biology, 16(03), 1650035.

[2] Cai H, Han J, Chen Y, Sha X, Wang Z, Hu B, Yang J, Feng L, Ding Z, Chen Y and Gutknecht J. (2018). A pervasive approach to EEG-based depression detection. Complexity, 2018, 1–13.

[3] Mahato S and Paul S. (2018). Detection of major depressive disorder using linear and non-linear features from EEG signals. Microsystem Technologies, 25(3), 1065–1076.

[4] Khadidos A.O., Alyoubi K.H., Mahato S and Mohanty S.N. (2023). Computer aided detection of major depressive disorder (MDD) using electroencephalogram signals. IEEE Access, 11, 41133–41141.

[5] Bachmann M, Lass and Hinrikus H (2017). Single channel EEG analysis for detection of depression. Biomedical Signal Processing and Control, 31, 391-7.

[6] Kalev K, Bachmann M, Orgo L, Lass J and Hinrikus H (2015). Lempel-Ziv and multiscale Lempel-Ziv complexity in depression. Engineering in Medicine & Biology Society.

[7] Minkowski L, Valter K and Gurve D (2021). Feature extraction to identify depression and anxiety based on EEG. 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC).

[8] Mahato S, Goyal N, Ram D and Paul S (2020). Detection of depression and scaling of severity using six channel EEG data. Journal of Medical Systems, 44(7).

[9] Jiang H, Jiang R, Wang X and Wang H (2020). Unbalanced Data Classification Algorithm Based on Combination of ADASYN and SMOTE. School of Computer and Information Security, Guilin University of Electronic Technology, Guilin Guangxi 541000, China.

[10] Alhudhaif A (2021). A novel multi-class imbalanced EEG signals classification based on the adaptive synthetic sampling (ADASYN) approach. PeerJ. Computer Science, 7, e523.

[11] Wang H (2019). Research and application of imbalanced classification problems. M.S. thesis, Suzhou University, Suzhou, China, June 2019.

[12] Zhang G, Yu M, Chen G, Han Y, Zhang D, Zhao G and Liu Y (2019). A review of EEG feature research for emotion recognition. Science China: Information Sciences, 49(9), 1097-1118.

[13] Tian X (2022). Research on wearable information perception methods for fatigue driving. Student Thesis Joint Comparison Database, Apr. 26, 2022.

[14] Richman J and Moorman J (2000). Physiological time-series analysis using approximate entropy and sample entropy, Am J Physiol Heart Circ Physiol, 278(6), H2039-49.

[15] Sun J (2014). Abnormal Dynamics of EEG Oscillations in Schizophrenia Patients on Multiple Time Scales. IEEE Transactions on Biomedical Engineering, 61(6), 1756-1764.

[16] Lempel A and Ziv, J(1976). On the complexity of finite sequences. IEEE Transactions on Information Theory, 22(1), 75-81.

[17] Pincus S. M. (1991). Approximate entropy as a measure of system complexity. Proceedings of the National Academy of Sciences, 88(6), 2297-2301.

[18] Sanz H, Valim C, Vegas E, Oller J.M. and Reverter F (2018). SVM-RFE: selection and visualization of the most relevant features through non-linear kernels. BMC Bioinformatics, 19(1).

[19] Shen E, Cai Z and Gu F (2005). Mathematical foundation of a new complexity measure. Applied Mathematics and Mechanics-English Edition, 26(9), 1188-1196.

[20] Sanz H, Valim C, Vegas E, Oller J.M. and Reverter F(2018). SVM-RFE: selection and visualization of the most relevant features through non-linear kernels. BMC Bioinformatics, 19(1).

[21] Yang Z, Luo T, Li J and Fan Y (2022). Research on Multi-Classification Algorithm for Depression Based on EEG and Common Spatial Patterns. Beijing Biomedical Engineering, Apr.

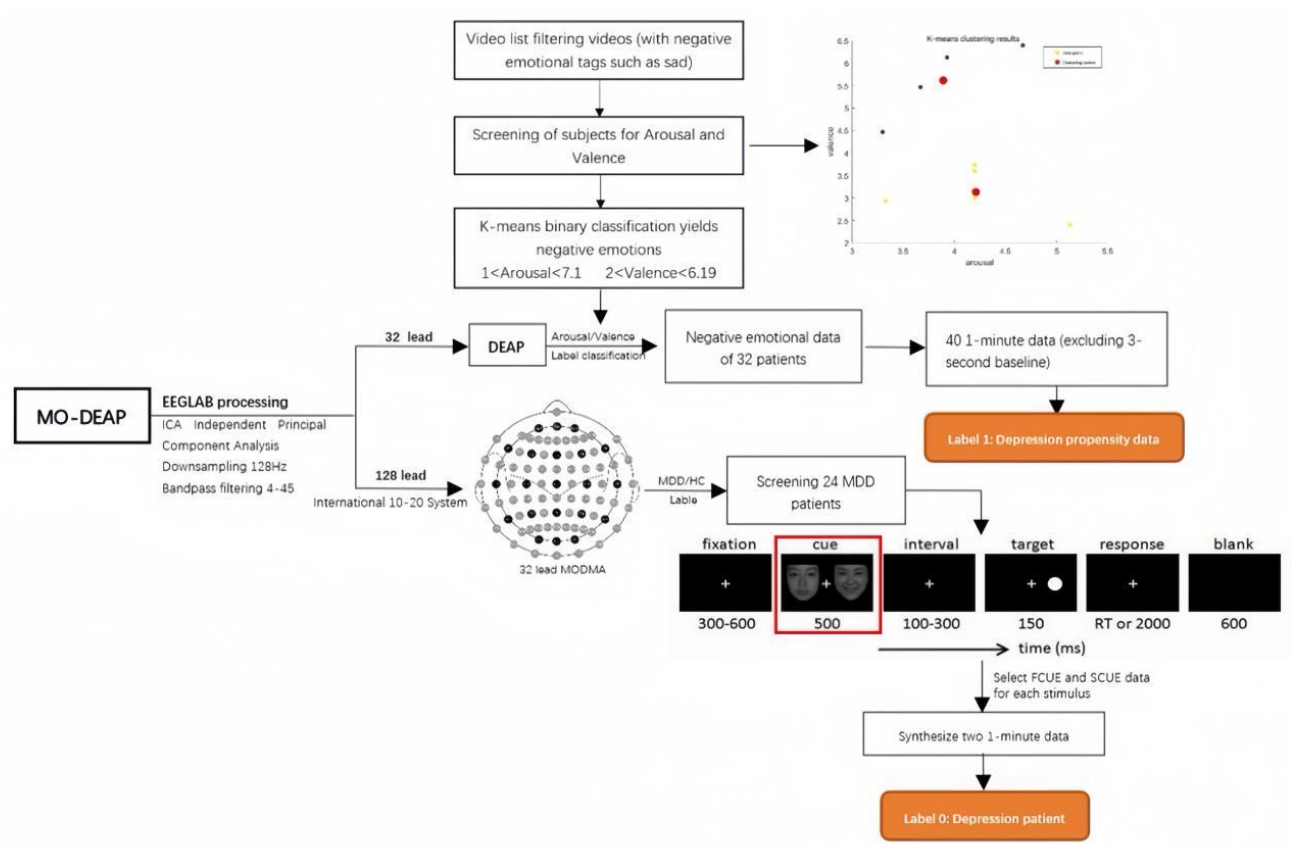

Figure 1. MO-DEAP preprocessing process

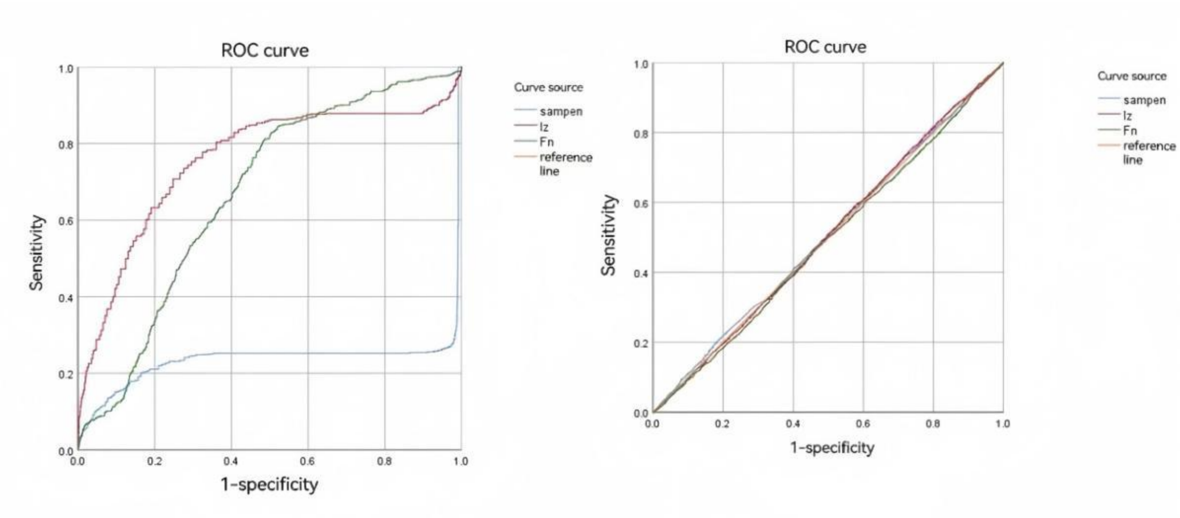

Figure 2. ROC curve after (a) SMOTE treatment (b) ADASYN treatment

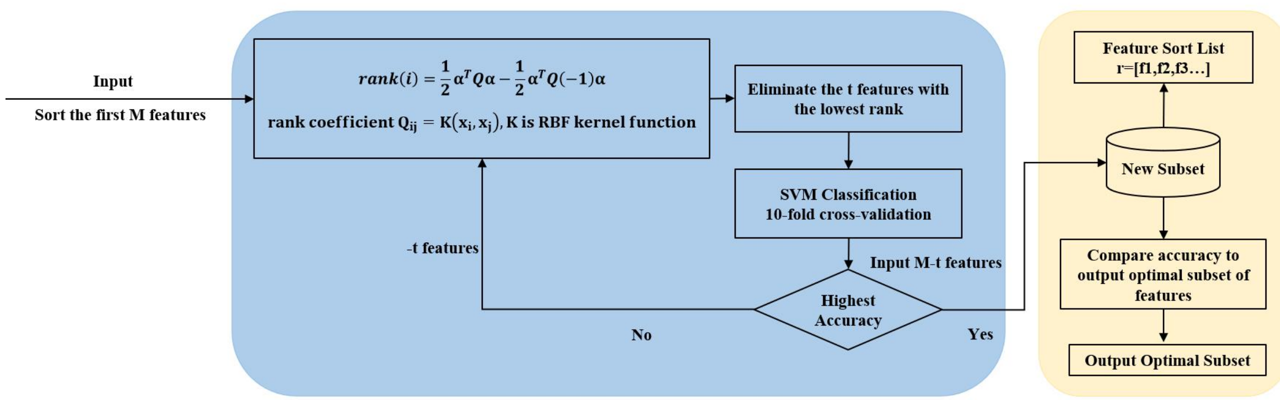

Figure 3. Basic procedure of nonlinear SVM-RFE classification algorithm

TABLE 1 THE DIFFERENCE OF DFA IN FP2, O1, O2 BETWEEN DEPRESSED GROUP AND DEPRESSION-PRONE GROUP

Channel FP2	Channel O1	Channel O2