

RESEARCH

Open Access



# A novel EEG-based major depressive disorder detection framework with two-stage feature selection

Yujie Li<sup>1</sup>, Yingshan Shen<sup>1</sup>, Xiaomao Fan<sup>2\*</sup>, Xingxian Huang<sup>3</sup>, Haibo Yu<sup>3</sup>, Gansen Zhao<sup>1</sup> and Wenjun Ma<sup>1</sup>

## Abstract

**Background:** Major depressive disorder (MDD) is a common mental illness, characterized by persistent depression, sadness, despair, etc., troubling people's daily life and work seriously.

**Methods:** In this work, we present a novel automatic MDD detection framework based on EEG signals. First of all, we derive highly MDD-correlated features, calculating the ratio of extracted features from EEG signals at frequency bands between  $\beta$  and  $\alpha$ . Then, a two-stage feature selection method named PAR is presented with the sequential combination of Pearson correlation coefficient (PCC) and recursive feature elimination (RFE), where the advantages lie in minimizing the feature searching space. Finally, we employ widely used machine learning methods of support vector machine (SVM), logistic regression (LR), and linear regression (LNR) for MDD detection with the merit of feature interpretability.

**Results:** Experiment results show that our proposed MDD detection framework achieves competitive results. The accuracy and  $F_1$  score are up to 0.9895 and 0.9846, respectively. Meanwhile, the regression determination coefficient  $R^2$  for MDD severity assessment is up to 0.9479. Compared with existing MDD detection methods with the best accuracy of 0.9840 and  $F_1$  score of 0.97, our proposed framework achieves the state-of-the-art MDD detection performance.

**Conclusions:** Development of this MDD detection framework can be potentially deployed into a medical system to aid physicians to screen out MDD patients.

**Keywords:** Depression detection, Two-stage feature selection, EEG

## Introduction

Major depressive disorder (MDD) is a debilitating disease characterized by at least one discrete depressive episode lasting no less than 2 weeks, which involves clear-cut changes in mood, interests and pleasure, as well as changes in cognition and vegetative symptoms [1]. It is one of the three major diseases throughout the world,

and its prevalence is still on the rise [2]. MDD may affect a patient's life severely when it attacks. Patients suffering from MDD are often accompanied by symptoms such as anhedonia, appetite and physical changes, poor sleep quality, slow thinking, loss of willpower activity, easy fatigue, excessive self-blame, and so on. In clinical practice, the most used methods for MDD diagnosis are mainly dependent on professional depression rating scales like the 17-item Hamilton rating scale (HAM-D-17) [3]. However, the rating scale assessment conducted requires a trained physician available, which is difficult for community and township hospitals to provide such

\*Correspondence: xmfan@scnu.edu.cn

<sup>2</sup> College of Big Data and Internet, Shenzhen Technology University, Shenzhen, China

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

kind of MDD detection services. Herewith, developing an automatic MDD detection method has great significance.

In recent years, EEG signals are widely being employed in brain function research and clinical applications, such as Brain-Computer Interfacing (BCI) [4], emotion classification [5], epilepsy [6, 7], and dementia [8]. Not limited to these, researchers also utilize EEG signals as a tool to build an MDD detection model. EEG signals are formed by the summation of postsynaptic potentials occurring simultaneously in many neurons in the brain, divided into  $\delta$ ,  $\theta$ ,  $\alpha$ ,  $\beta$ , and  $\gamma$  bands [9]. In addition, it contains rich human physiological information, which can be used to discover the features connected to depressive disorders. Although the existing methods [10–13] have been achieved promising results in MDD detection, these methods have not sufficiently been solved the following challenges: (1) How to find highly MDD-correlated features with medical interpretations. (2) How to reduce the feature searching space on high-dimensional data.

Many researchers attempted to generate MDD-related features from EEG signals to solve the first challenge. Specifically, Hosseinifard et al. [9] extracted features directly from EEG band power and nonlinear features of detrended fluctuation analysis (DFA), Higuchi fractal, correlation dimension and Lyapunov exponent from EEG signals utilizing signal processing. Acharya et al. [14] presented an MDD-related feature named depression diagnosis index (DDI) through the combination of different nonlinear features, including fractal dimension, largest Lyapunov exponent, and sample entropy. Mahato et al. [15] derived EEG-based linear features of band power, interhemispheric asymmetry, and nonlinear features of relative wavelet energy (RWE) and wavelet entropy (WE) and fused these features together for depression detection. In general, most of the MDD-related features presented in their work were derived through signal processing or mathematical analysis from EEG signals. However, these features were not proven to be directly correlated to MDD in prior medical research and MDD detection performance. Instead, their correlation needs to be deduced backward according to the classification results of MDD detection models. Therefore, there is no sufficient evidence to show that these features are highly correlated with MDD, and how to introduce the prior medical research knowledge when extracting the highly correlated features of MDD remains to be solved.

For the second challenge, Mohammadi et al. [16] employed the designed feature selection method of genetic algorithm (GA) on the features extracted by linear discriminant analysis (LDA), aiming to achieve the optimal MDD detection result. However, GA is a randomized searching method, which is unstable and still has the potential risk of falling into the optimal local

solution. Wajid et al. [17] utilized a rank-based feature selection method by assigning weight and a rank-ordered to an EEG feature matrix. Then, the features were ranked-ordered according to the scores calculated by the area under the curve (AUC). And finally, the parts with the highest scores were selected. Afterwards, Wajid et al. [18] further improved their feature selection method. They first ranked the extracted features in descending order according to receiver operating characteristics (ROC), then split the features into subgroups to train the classifiers, and ultimately selected the subgroup whose features could make the classifier achieve the optimal result. However, these methods select features based on AUC and ROC metrics, which would overemphasize the importance of ranking, and thus ignore the information of the recall rate. In addition, these wrap-based feature selection methods have high feature searching space. Especially for high-dimensional data, the computational complexity is quite high.

To solve the above challenges, in this study, we propose a novel EEG-based framework for MDD detection and severity assessment with a two-stage feature selection. First of all, 92 participants at Shenzhen Traditional Chinese Medicine Hospital with a signed consent form were recruited to collect EEG signals with assessing MDD severity with the HAM-D-17 scale by a physician. Then, highly MDD-correlated features are derived based on the ratio of extracted features from EEG signals at frequency bands between  $\beta$  and  $\alpha$ . Subsequently, a two-stage feature selection method called PAR with the combination of Pearson correlation coefficient (PCC) and recursive feature elimination (RFE) is presented with the advantage of minimizing the feature searching space. Finally, we employ widely used machine learning methods of LR, SVM, and linear regression (LNR) for MDD detection and MDD severity assessment with the merit of well feature interpretability. Experiment results show that our proposed framework can obtain promising results, where the best  $F_1$  score and regression determination coefficient  $R^2$  are 0.9846 and 0.9479, respectively. Differed in the multi-modal method [19] requiring multiple inputs such as facial expressions, heart rate, and posture, the input of our framework is based on EEG signals, which is much more convenient for community/township hospitals to implement. The main contributions of this study can be summarized as follows:

1. We derive highly MDD-correlated features called  $\beta/\alpha$  ratio features calculated by the ratio of extracted features from EEG signals at frequency bands between  $\beta$  and  $\alpha$ , which can greatly improve the MDD detection performance and strong medical interpretation.

2. To further minimize the feature searching space, we present a two-stage feature selection method with the combination of PCC and RFE.
3. We utilize widely used machine learning methods of LR, SVM, and LNR for MDD detection and MDD severity assessment with well interpretability.
4. Extensive experiment results show that the proposed MDD detection framework has superiority to the state-of-the-art methods with a big margin.

The remained contents of this paper are organized as follows. In “Related work” section, the literature review of MDD detection is presented. In “Methods” section, the details of the proposed framework are briefly introduced, including the data collection, data preprocessing, two-stage feature selection, and MDD detection model. In “Results” section, MDD detection experiment results are presented. In “Discussion” section, hyper-parameters, limitations and future work are discussed. Finally, this study is summarized in “Conclusion” section.

### Related work

In recent years, physiological signals, including EEG [6], ECG [20, 21] and EMG [22], have been widely used in disease detection. With the development of machine learning (ML) technology, many researchers have made outstanding contributions to automatic MDD detection using ML. Various deep learning methods [23, 24] have been proposed. Specifically, Sharma et al. [25] presented an EEG-based network for depression screening with CNN and LSTM. Seal et al. [26] proposed a designed deep CNN framework for detecting depression using EEG signals. Although these deep learning methods can achieve high accuracy, their selected features lack interpretability, which is a big defect in medical applications where interpretability is highly valued. Although traditional machine learning methods are slightly inferior to deep learning methods in the accuracy of MDD detection, MDD-related features for both machine use and human analysis are selected through feature engineering combined with prior knowledge of humans, and these features are accessible to humans to understand.

There have been several machine learning methods employing feature engineering to detect MDD. Akbari et al. and Mohammadi and Moradi [27, 28] utilized the reconstructed phase space (RPS) of EEG signals and geometrical features for training the SVM classifier and KNN classifier for depression recognition. Jiang et al. [29] extracted features from spatial information of EEG signals to detect MDD. Sadiq et al. [30] introduced a unified algorithm to classify neural diseases from two distinct EEG domains. However, most of their proposed models have only used mathematical or signal processing

methods to extract EEG-related features, which are less correlated with MDD, resulting in poor performance of MDD detection. Therefore, how to extract MDD-related features from EEG signals for better detection remains to be solved.

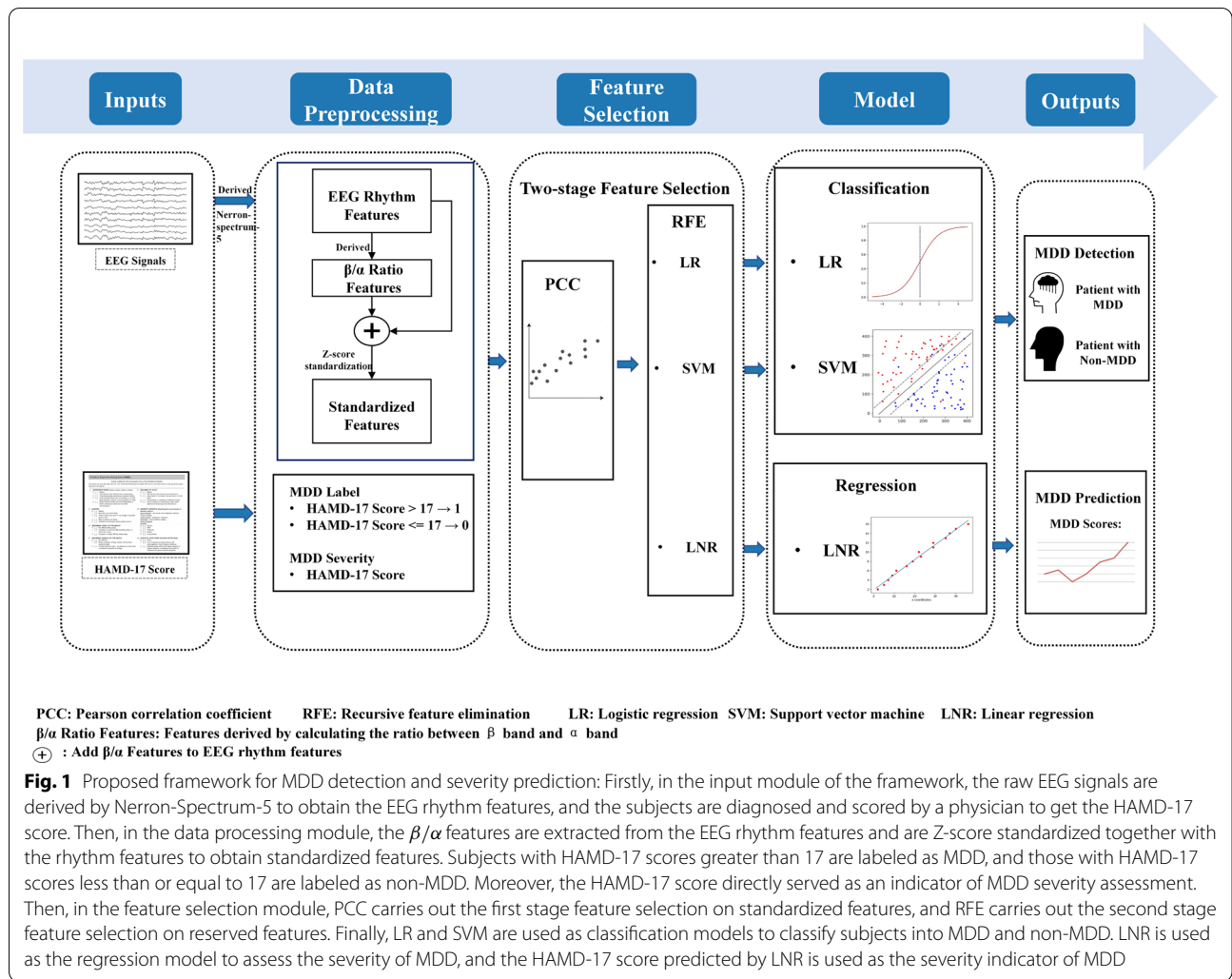
Some previous researchers [24, 31, 32] attempt to improve the performance of ML methods by extracting MDD-related features from EEG signals. Ciarleglio et al. [31] derived features from EEG by a brain function measure named Frontal power asymmetry (FA). Avots et al. [32] tried different linear and nonlinear EEG features, including relative band power, spectral asymmetry index, Higuchi fractal dimension, and Lempel–Ziv complexity. Bailey et al. [33] extracted features by exploring the resting EEG  $\theta$  connectivity or  $\alpha$  power to the predictors of response to rTMS treatment. Although the aforementioned MDD detection methods could extract MDD-related features, they generally employ one-stage feature selection, which tends to search local optimal selected features from high dimensional feature search space, whose MDD detection performance still exists an improvement margin.

### Methods

The proposed framework mainly consists of data preprocessing, feature selection, MDD detection and severity prediction. Figure 1 shows the whole pipeline of our framework.

#### Data collection

92 participants with depression lasting more than two weeks, who were conscious, and did not get aphasia or mental retardation, were recruited from the Shenzhen Traditional Chinese Hospital and signed informed consent forms (IRB No. 2017-8) to collect EEG signals and corresponding HAM-D-17 scores. A 10–20 system of electrode placement was placed on the scalp of the participants, and smooth EEG signals were acquired with the Nerron-spectrum-5 EEG device for 30 seconds under the condition of eyes closed. The Nerron-spectrum-5 EEG device has a total of 19 channel leads, which are FP1-A1, FP2-A2, F3-A1, F4-A2, FZ-A2, C3-A1, C4-A2, CZ-A1, P3-A1, P4-A2, PZ-A2, O1-A1, O2-A2, F7-A1, F8-A2, T3-A1, T4-A2, T5-A1, and T6-A2. Each channel-lead can automatically obtain EEG-related features based on the information of rhythm waveform amplitude indexes and rhythm indexes. Both indexes possess frequency bands of  $\delta$ ,  $\theta$ ,  $\alpha$ , and  $\beta$ . Therefore, 152 EEG-related features are obtained from each participant. With additional demographic features of gender and age, a total of 154 features are derived for subsequent processing and analysis.



### Data preprocessing

Clinically, frequency band  $\beta$  in EEG is associated with an alert or excited state of mind and frequency band  $\alpha$  is more dominant in a relaxed state, both of them are associated with brain inactivation. Previous studies [34, 35] were also shown that the  $\beta/\alpha$  ratio features are potentially correlated with MDD, which are derived with EEG frequency bands of  $\alpha$  and  $\beta$ . Thus, in this work, the  $\beta/\alpha$  ratio features are derived by calculating the ratio between frequency bands of  $\beta$  and  $\alpha$  on each channel-lead in terms of EEG waveform amplitude indexes and rhythm index information, which are defined to be:

$$F_A = \frac{F_{A-\beta}}{F_{A-\alpha}} \tag{1}$$

$$F_R = \frac{F_{R-\beta}}{F_{R-\alpha}} \tag{2}$$

where features of  $F_A$  and  $F_R$  are the  $\beta/\alpha$  ratio features, derived from extracted features of EEG waveform amplitude indexes and rhythm indexes respectively.  $F_{A-\beta}$  and  $F_{A-\alpha}$  are features from EEG waveform amplitude indexes on frequency bands of  $\beta$  and  $\alpha$  in EEG signals, respectively.  $F_{R-\beta}$  and  $F_{R-\alpha}$  are features from EEG rhythm indexes on  $\beta$  and  $\alpha$  frequency bands, respectively. After calculation, a total of 38  $\beta/\alpha$  ratio features are derived, which are combined with the extracted feature set of EEG waveform amplitude index and rhythm index. Totally, there are 192 features obtained for feature selection.

In the MDD detection task, participants are divided into two groups according to the HADM-17 score to reflect the severity of MDD. Those with a HADM-17 score greater than 17 are defined as MDD, labeled as 1; those with a HADM-17 score no more than 17 are defined as mild depression or without depression, labeled as 0. In the MDD severity assessment task, the real HAMD-17 scores are compared with the predicted ones in the LNR-related

model. As reported in [10, 17, 18], z-score standardization is widely used in EEG-based MDD detection task. Therefore, in order to minimize the impact of different feature scales, we standardize the data on each feature using z-score standardization for numerical data, making different features to be in a uniform distribution. Since the obtained EEG features might not be centered and unequally distributed, z-score standardization is more suitable than min-max normalization, which can be affected by possible outliers in the dataset. Z-score standardization is defined to be:

$$z = \frac{x - \mu}{\sigma} \quad (3)$$

where  $\mu$  is the mean value of a feature and  $\sigma$  is the variance of it. In addition, for the non-numerical feature of gender, male is marked to be 1 and female is to be 0.

### Two-stage feature selection

After data preprocessing, directly using the obtained high-dimensional features as input usually cause poor performance in the classification task and the regression task. In addition, in feature selection, if the features with the highest correlation with MDD can be selected while reducing the feature searching space, the model can achieve the best results. Therefore, selecting an appropriate algorithm to conduct feature selection is essential before building a specified model. Typical single-stage feature selection methods usually use filters [36] to set thresholds to remove low correlation features or use wrappers [37] to optimize the given learner directly. Specifically, the filters have the advantages of simplicity and efficiency, and the wrappers have the advantages of high accuracy in the task of ML. However, during the selection process, the former can not obtain adequate information from the learner, resulting in poor performance of the learner, while the latter requires multiple iterations in all the input features to train the learner, and many low-correlation features are also iterated, resulting in high computational cost. Therefore, it is appropriate to employ the wrappers model on the features retained by the filters, which can effectively minimize the feature searching space.

To take advantage of the low resource consumption of the filter and the high performance of the wrapper simultaneously, in this work, we present a two-stage feature selection method named PAR with the combination of filter method PCC [38], and wrapper method RFE [39]. Specifically, we first define the PCC threshold as  $\tau$ , and RFE selected feature number as  $\zeta$ , then the PCC method, which is a simple tool to calculate the correlation coefficient between the EEG features and HAMD-17 scores, is selected for the first stage of feature selection. We employ PCC to calculate the absolute correlation coefficient (values between 0 and 1) between the features and the labels

(the values of labels are 1 for MDD, and 0 for non-MDD) as PCC values, and then sort the corresponding features according to the PCC values in reverse order. Then, the features with PCC values less than  $\tau$  are removed, and the features with PCC values greater than  $\tau$  will be retained for the second stage of feature selection. The second-stage feature selection method is employed with the RFE, which has the advantage of utilizing the supervision information to select optimal features. In the second stage, RFE first recursively eliminates the features selected by PCC and deletes the number of features to the specified number  $\zeta$ . Then it builds the model based on the remaining features and calculates the average score ( $F_1$  score in MDD detection and  $R^2$  in MDD severity assessment) through five-fold cross-validation. Finally, the feature combination that contributes the most to the prediction results and its corresponding hyper-parameters  $\tau$  and  $\zeta$  are saved. The final selected features and hyper-parameters  $\tau$  and  $\zeta$  are determined by the grid search technique described in the Algorithm 1. In general, in grid search, we determine the hyper-parameters  $\tau$  and  $\zeta$  that can make the model perform best in five-fold cross-validation through a two loop. The former hyper-parameter is the threshold value used by PCC to select features, and the latter is the number of features used by RFE to select features. Employing RFE feature selection on the features selected by PCC first can improve computing efficiency. Finally, the features obtained through two-stage feature selection are not only MDD high related, but also can be better utilized by the learners (LR, SVM, and LNR). To better describe our feature selection algorithm PAR, we write a pseudo-code which is described in detail in Algorithm 1.

---

#### Algorithm 1 PAR

---

**Input:** Features  $X$ , labels  $Y$

- 1: Define  $M_T$  as the final selected feature matrix by the PAR;
- 2: Define  $S_{Max} = -\infty$ , as the maximum score of the current iteration;
- 3: Calculate the absolute value of correlation coefficient between  $X$  and  $Y$  as PCC values, obtaining correlation matrix  $M_A$ , in which the PCC values correspond to features  $X$ ;
- 4: **for**  $\tau = 0; \tau < 1; \tau = \tau + 0.005$  **do**
- 5:   Iterate through the PCC values in  $M_A$  and select the corresponding features with PCC value  $\geq \tau$  to form  $M_P$ , the PCC selected features matrix;
- 6:   **for each**  $\zeta \in [1, \text{len}(M_P)]$  **do**
- 7:     Select  $\zeta$  features in  $M_P$  through RFE and form  $M_R$ , the RFE selected feature matrix;
- 8:     Train the model with  $M_R$  using five-fold cross validation, calculate the average  $F_1$  score or  $R^2$  and obtain the current calculated score  $S_{Curr}$ ;
- 9:     **if**  $S_{Curr} > S_{Max}$  **then**
- 10:        $S_{Max} = S_{Curr}$ ;
- 11:        $M_T \leftarrow M_R$ ;
- 12:     **end if**
- 13:   **end for**
- 14: **end for**

**Output:**  $M_T$

---

In Algorithm 1, we first define the final selected feature matrix as  $M_T$ , and the maximum measurement score of the current iteration as  $S_{Max}$ . Then, we calculate each PCC value between input EEG features  $X$  and their corresponding label  $Y$  (MDD or non-MDD) and combine the PCC values to form the PCC matrix  $M_A$ . And then, in a two loop, the hyperparameter  $\tau$  of the outer loop is tuned from 0 to 1 with a step of 0.01, controlling the correlation of the retained features filtered by PCC in the first stage of feature selection, and the selected features are saved as matrix  $M_P$ . The hyperparameter  $\zeta$  of the inner loop is tuned from 192 to 1 with a step of 1, controlling the number of features selected by RFE in the second stage of feature selection, and the selected features are saved as matrix  $M_R$ . The retained EEG features are then used to train the model (SVM and LR in MDD detection, and LNR in MDD severity assessment), and the average score  $S_{Curr}$  ( $F_1$  score in MDD detection, and  $R^2$  in MDD severity assessment) is calculated using 5-fold cross-validation. Next, the calculated  $S_{Curr}$  is compared with  $S_{Max}$ , and if the value of  $S_{Curr}$  is greater than the value of  $S_{Max}$ , it will be assigned to  $S_{Max}$ , and the features of  $M_R$  will be assigned to  $M_T$ . The next iteration of the loop is then carried out. Finally, the EEG features that make the model perform best are saved as matrix  $M_T$  for MDD detection or severity assessment, and feature analysis.

**MDD detection**

MDD detection task can be essentially transformed into a binary classification problem, by classifying patients with MDD and non-MDD. After two-stage feature selection, the widely used machine learning methods LR and SVM are used for MDD detection in this work.

LR [40] is a widely used classification method, which utilizes the Sigmoid function as a posteriori probability distribution to identify patients with MDD or not. Meanwhile, LR is with merits of less computation, interpretability, and easy implementation, the equation [41] is defined to be:

$$P(y = 1|x) = \frac{\exp(\theta^T x)}{1 + \exp(\theta^T x)} \tag{4}$$

$$P(y = 0|x) = \frac{1}{1 + \exp(\theta^T x)} \tag{5}$$

where  $\theta$  is the weight importance of the LR model,  $y = 1$  denotes patients with MDD as well as  $y = 0$  denotes patients not with MDD.

SVM [42] is a binary classification method originally, and its decision boundary is the maximum-margin hyperplane, the equation [43] is defined to be:

$$w^T x + b = 0 \tag{6}$$

where  $w$  is weight importance and  $b$  is bias. The SVM uses the hyperplane to discriminate patients into MDD or not. Owing to the merit of human-interpretability, the SVM used in this work is with a linear kernel function [43], which is defined to be:

$$k(x_i, x_j) = x_i^T x_j + c \tag{7}$$

where  $x_i$  and  $x_j$  are vectors of different EEG rhythm signal samples,  $c$  is an optional constant.

**MDD severity assessment**

The MDD severity assessment can be transformed into a linear regression problem, and the HAMD-17 score is used as an indicator of depression severity, because the more severe the depression, the higher the corresponding HAMD-17 score will be. Therefore, we need to use the features obtained through two-stage feature selection to fit the measurement metric, to realize the assessment of the depression severity in patients. LNR [44] has a rapid modeling ability and is very effective for input data with a small sample size. In addition, it has a good explanation for each variable, which has been widely used in the medical field. LNR uses the minimum square function called linear regression equation [45] to model the relationship between one or more independent variables and dependent variables, which is defined as follows:

$$\hat{y} = \beta^T x + d \tag{8}$$

where  $\beta$  is the parameter of the model,  $x$  is the vector of the EEG rhythm signal sample, and  $d$  is an optional constant.

**MDD detection performance metrics**

Typical classification metrics, including sensitivity ( $Sen$ ), specificity ( $Spec$ ), precision ( $Prec$ ), recall ( $Rec$ ), accuracy ( $Acc$ ), and  $F_1$  score are used for each class. They are defined as:

$$Sen = \frac{TP}{TP + FN} \tag{9}$$

$$Spec = \frac{TN}{TN + FP} \tag{10}$$

$$Acc = \frac{TP + TN}{TP + FP + FN + TN} \tag{11}$$

$$Prec = \frac{TP}{TP + FP} \tag{12}$$

$$Rec = \frac{TP}{TP + FN} \tag{13}$$

$$F_1 = \frac{2 * (Prec * Rec)}{Prec + Rec} \tag{14}$$

where TP refers to the number of actual MDD patients classified to be MDD, FP refers to the number of actual non-MDD patients classified to be MDD, FN refers to the number of actual MDD patients classified to be non-MDD, and TN refers to the number of actual non-MDD patients classified to be non-MDD.

**MDD severity assessment performance metrics**

Typical regression metrics, including mean absolute error ( $L_{MAE}$ ) [46], mean squared error ( $L_{MSE}$ ) [47], determine coefficient  $R^2$  [48] are employed to evaluate the performance of LNR. They are defined as:

$$L_{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{15}$$

$$L_{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \tag{16}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \tag{17}$$

where  $y_i$  is the actual HAMD-17 score for each individual,  $\hat{y}_i$  is the corresponding predicted score, and  $\bar{y}$  is the mean value of all HAMD-17 scores.

**Results**

**Experimental environment**

The proposed framework is implemented with python 3.6.11 and scikit-learn 0.21.3. All the experiments are performed on a laptop equipped with an AMD Ryzen 7 4800U CPU and 16 GB memory.

**MDD detection performance**

In this work, two widely used machine learning methods of LR and SVM, with the merit of good feature interpretability, are employed to build the MDD detection model. Furthermore, to evaluate the MDD detection model generality, the five-fold cross-validation technique is used. With the hyperparameters of LR and SVM keeping default, experiment results show that the LR with derived  $\beta/\alpha$  features (LR-DF for short) achieved 0.5677, 0.6182, 0.5984, and 0.5050 in terms of sensitivity, specificity, accuracy, and  $F_1$  score, respectively. For the SVM with derived  $\beta/\alpha$  features (SVM-DF for short), its MDD detection performance is comparable to the LR-DE, which

achieved a sensitivity of 0.5333, specificity of 0.6894, accuracy of 0.6318, and  $F_1$  score of 0.5144. With the two-stage feature selection, the MDD detection performance of both LR-DF and SVM-DF are improved greatly. The SVM-DF with the two-stage feature selection obtained the best MDD detection performance, which are 0.9714, 1.0000, 0.9895, and 0.9846 in terms of sensitivity, specificity, accuracy, and  $F_1$  score, respectively. It means that the SVM-DF with two-stage feature selection can be potentially deployed into a medical system to provide an automatic MDD detection service.

**MDD severity assessment performance**

To further help physicians to obtain patients' MDD severity, the specified HAMD-17 score should be provided. In this work, the LNR is used for MDD severity assessment with default hyperparameters in the machine learning framework scikit-learn. As shown in Table 1, the LNR with the derived  $\beta/\alpha$  features (LNR-DF for short) achieves better performance for MDD severity assessment than that without the derived  $\beta/\alpha$  features, which means that the derived features can boost the performance of MDD severity assessment as well as MDD detection. Since the input parameter matrix required by LNR fitting is a nonsingular matrix, and we only have 92 data, there are 154 features of LNR and 192 features of LNR-DF without feature selection, which leads to the result that input rank is higher than the number of experimental data. Therefore, the  $L_{MAE}$ ,  $L_{MSE}$ , and  $R^2$  value of the LNR and LNR-DF model without feature selection cannot be calculated. Compared with the LNR, the LNR-DF achieved better MDD severity assessment performance, whose determine coefficient is 0.0927 with PCC selection, and 0.8962 with RFE selection. In PCC feature selection, the features selected by LNR are  $F_{A-\delta}^{FZ-A2}$ ,  $F_{A-\delta}^{CZ-A1}$ , and Gender, and the features selected by LNR-DF are  $F_{R-\beta/\alpha}^{FP2-A2}$ ,  $F_{A-\delta}^{F3-A1}$ ,  $F_{A-\delta}^{FZ-A2}$ ,  $F_{A-\delta}^{CZ-A1}$ ,  $F_{A-\delta}^{O2-A2}$ ,  $F_{A-\delta}^{T4-A2}$ ,  $F_{R-\beta/\alpha}^{FP2-A2}$ , and Gender. These features generally have a strong correlation with each other, resulting in that the learning region of the model can only be limited to a small range, and cannot make good use of other features with

**Table 1** MDD severity assessment performance

Feature selection	Model	$L_{MAE}$	$L_{MSE}$	$R^2$
PCC	LNR	3.6917	21.6649	0.0644
	LNR-DF	3.5198	20.9566	0.0927
RFE	LNR	1.5306	3.4942	0.8474
	LNR-DF	1.2450	2.3995	0.8962
PCC and RFE	LNR	1.2799	2.7388	0.8812
	LNR-DF	0.9123	1.2012	0.9479

less close correlation, eventually lead to poor performance. Furthermore, with the two-stage feature selection, MDD severity assessment performance of both LNR and LNR-DF is improved a lot. It can be observed that the LNR-DF has much superiority over the LNR on the task of MDD severity assessment, whose determine coefficient can be up to 0.9479. It means that LNR-DF with two-stage feature selection can regress the ground truth of the HAMD-17 score very well and can be a reference for physicians to assess the patients' MDD severity.

**Discussion**

**MDD detection performance analysis**

To verify the effectiveness of the derived  $\beta/\alpha$  features and two-stage feature selection, extensive experiments are performed in this study. As shown in Table 2, both LR-DF and SVM-DF, with the derived  $\beta/\alpha$  features, achieve better MDD detection performance than those methods of LR and SVM without the derived  $\beta/\alpha$  features. It demonstrates that the derived  $\beta/\alpha$  features have high correlations with the MDD detection task. Meanwhile, it can also be observed that MDD detection methods of LR, SVM, LR-DF and, SVM-DF raise their performance greatly no matter single-stage or two-stage feature selection. More importantly, the MDD detection methods with the RFE achieve better performance than those with the PCC, which is the reason why we choose the RFE as the core feature selection in the presented two-stage feature selection method.

Specifically, as shown in Table 2, the derived  $\beta/\alpha$  features can help MDD detection methods to improve their

performance more or less. Take MDD detection models of LR, SVM, LR-DF, and SVM-DF with the two-stage feature selection for example, the LR and SVM which are not with the derived  $\beta/\alpha$  features achieved  $F_1$  scores of 0.7803 and 0.9581, respectively. The LR-DF and SVM-DF which are with the derived  $\beta/\alpha$  features obtained  $F_1$  scores of 0.8713 and 0.9846, which outperform the LR and SVM with margins of 11.66% and 2.77%, respectively. What's more, it can be observed that MDD detection methods of LR, SVM, LR-DF, and SVM-DF improve their performance a lot with feature selection, especially with the two-stage feature selection. For example, the SVM-DF with single-stage feature selection achieved  $F_1$  scores of 0.5766 and 0.9667 when the PCC and RFE are employed respectively, which obviously outperform the SVM-DF without feature selection. Regarding the presented two-stage feature selection, the MDD detection performance of SVM-DF increases significantly to the SVM-DF without feature selection or with single-stage feature selection, which is up to  $F_1$  score of 0.9846. It means that the presented two-stage feature selection is effective in the task of MDD detection.

**Comparison of MDD detection performance**

In addition, the proposed framework is compared with state-of-the-art methods. Table 3 shows the performance comparison of MDD detection between the proposed framework and cutting-edge methods published in recent years. Here, we have to point out that the test datasets used in this work and cutting-edge methods are different because the datasets in the medical domain are often private and difficult to be accessed publicly. Even

**Table 2** MDD detection performance of the proposed framework

Feature selection	Model	Sen	Prec	Spec	Acc	$F_1$
-	LR	0.4476	0.3776	0.6167	0.5557	0.4005
	SVM	0.4524	0.4809	0.7061	0.6139	0.4510
	LR-DF	0.5667	0.4728	0.6182	0.5984	0.5050
	SVM-DF	0.5333	0.5000	0.6894	0.6318	0.5144
PCC	LR	0.5619	0.5061	0.6530	0.6207	0.5235
	SVM	0.5381	0.5444	0.7591	0.6782	0.5306
	LR-DF	0.6238	0.6083	0.7394	0.6961	0.5988
	SVM-DF	0.5619	0.6442	0.7727	0.6969	0.5766
RFE	LR	0.7381	0.8571	0.9106	0.8466	0.7803
	SVM	0.9143	0.9667	0.9833	0.9579	0.9385
	LR-DF	0.8524	0.8298	0.8955	0.8800	0.8386
	SVM-DF	0.9429	1.0000	1.0000	0.9789	0.9667
PCC and RFE	LR	0.7381	0.8571	0.9106	0.8466	0.7803
	SVM	0.9714	0.9464	0.9652	0.9678	0.9581
	LR-DF	0.8238	0.9417	0.9636	0.9127	0.8713
	SVM-DF	0.9714	1.0000	1.0000	0.9895	0.9846

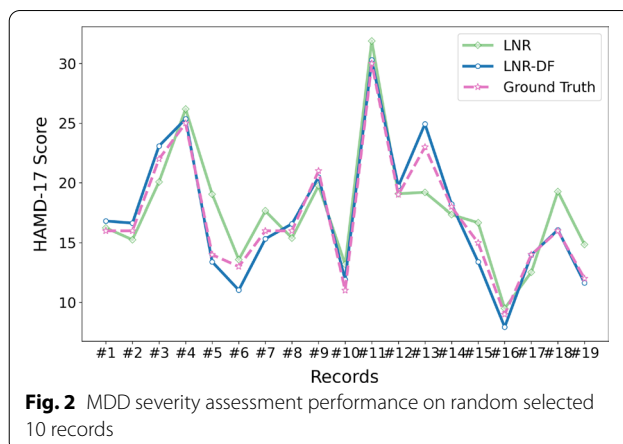


so, it still makes significant sense when cutting-edge methods are compared on a general level. As shown in Table 3, the best model SVM-DF of the proposed framework has obvious superiority to cutting-edge methods in terms of sensitivity, accuracy, and  $F_1$  score. Regarding the specificity, the SVM-DF achieves 0.98, which is still comparable to the best specificity of 1.0 reported by [17].

**MDD severity assessment performance analysis**

To further verify the effectiveness of the MDD severity assessment model, 19 records from individuals are randomly selected for validation. As shown in Fig. 2, it is the MDD severity fitting performance chart with the two-stage feature selection, where the x-axis and y-axis refer to the record and HAMD-17 score, respectively. Most of the cases, the LNR and LNR-DF can well predict the MDD severity represented by the HAMD-17 score. However, there exists a big difference in MDD severity assessment performance on the records 18 and 19. Throughout the whole Fig. 2, the results of LNR-DF are closer to the ground truth than those of the LNR. It demonstrates that the presented two-stage feature selection has positive effects on the task of MDD severity assessment and the MDD severity assessment model can improve its performance with the derived  $\beta/\alpha$  features.

Meanwhile, we draw a scatter plot (Fig. 3) of LNR and LNR-DF with the two-stage feature selection, where the x-axis and y-axis refer to predicted and ground truth HAMD-17 score, respectively. As shown in Fig. 3, most of the scatter points are nearby the diagonal line which means the presented two-stage feature selection method improves the MDD severity assessment performance very well, particularly with the combination of the derived  $\beta/\alpha$  features.



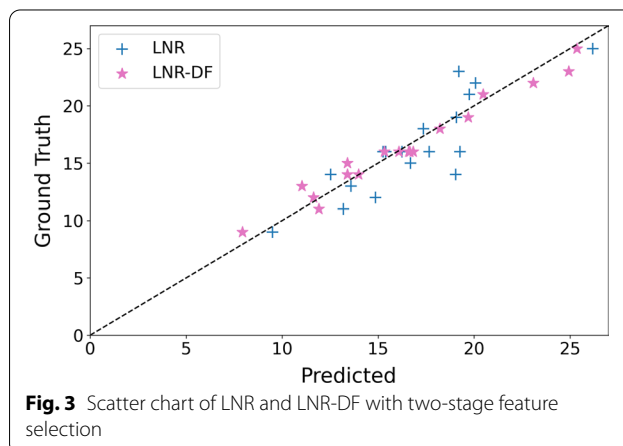
**Fig. 2** MDD severity assessment performance on random selected 10 records

**Hyperparameters tuning**

In this section, the tuning of two key hyperparameters of the proposed framework is introduced briefly, where the two hyperparameters of the proposed framework are the threshold value  $\tau$  for PCC feature selection and the number of features  $\zeta$  selected by RFE. To obtain the optimal hyperparameters, the grid search technique is utilized for searching the optimal combination of  $\tau$  and  $\zeta$ . Specifically, the hyperparameter  $\tau$  is tuning from 0.01 to 0.07 with a step of 0.01 while the hyperparameter  $\zeta$  from 192 to 1 with a step of 1. As shown in Fig. 4, take LR-DF, SVM-DF, and LNR-DF for example, and it is noted that the hyperparameter  $\tau$  is 0.01 for LR-DF and LNR-DF, and 0.015 for SVM-DF, while the hyperparameter  $\zeta$  is different among them. The optimal hyperparameters  $\zeta$  are 26, 36, and 63 in terms of LR-DF, SVM-DF, and LNR-DF, respectively, when they achieved the best performance. Furthermore, the key hyperparameters of  $\tau$  and  $\zeta$  of all MDD detection models and MDD severity assessment models are listed in Table 4 used in this work.

**Table 3** Performance comparison with cutting-edge MDD detection methods

Method	Sen	Spec	Acc	$F_1$
[9]	-	-	0.9000	-
[10]	0.9490	0.8090	0.8790	-
[11]	0.9444	-	0.8912	-
[12]	-	-	0.8500	-
[17]	0.9666	1.0000	0.9840	-
[18]	0.9990	0.9500	0.9800	0.9700
[13]	-	-	0.7927	-
[49]	-	-	0.8833	-
SVM-DF	0.9714	1.0000	0.9895	0.9846



**Fig. 3** Scatter chart of LNR and LNR-DF with two-stage feature selection

**Table 4** Hyper-parameter of feature selection

Feature selection	Method	$\tau$	$\zeta$	$F_1$	$R^2$
-	SVM	-	154	0.4510	-
	SVM-DF	-	192	0.5144	-
	LNR	-	-	-	-
	LNR-DF	-	-	-	-
PCC	SVM	0.055	-	0.5306	-
	SVM-DF	0.055	-	0.5766	-
	LNR	0.210	-	-	0.0644
	LNR-DF	0.185	-	-	0.0927
RFE	SVM	-	43	0.9385	-
	SVM-DF	-	53	0.9667	-
	LNR	-	53	-	0.8474
	LNR-DF	-	66	-	0.8962
PCC and RFE	SVM	0.040	38	0.9581	-
	SVM-DF	0.015	36	0.9846	-
	LNR	0.010	59	-	0.8812
	LNR-DF	0.010	63	-	0.9479

**Statistical analysis**

In order to more reasonably explain the significance of the  $\beta/\alpha$  features and two-stage feature selection to model improvement, we repeat the experiment 5 times and conduct statistical power analysis in MDD detection and MDD severity assessment by calculating the  $P$  value between the method with the best performance and other methods in our framework respectively. Here, we leverage three levels to mark the significance (\*, \*\*, \*\*\*), which represent the  $P$  value greater than or equal to 0.05, the  $P$  value between 0.05 and 0.01, and the  $P$  less than 0.01. As shown in Table 5, it is demonstrated that the SVM-DF with two-stage feature selection is superior to the compared methods significantly, especially

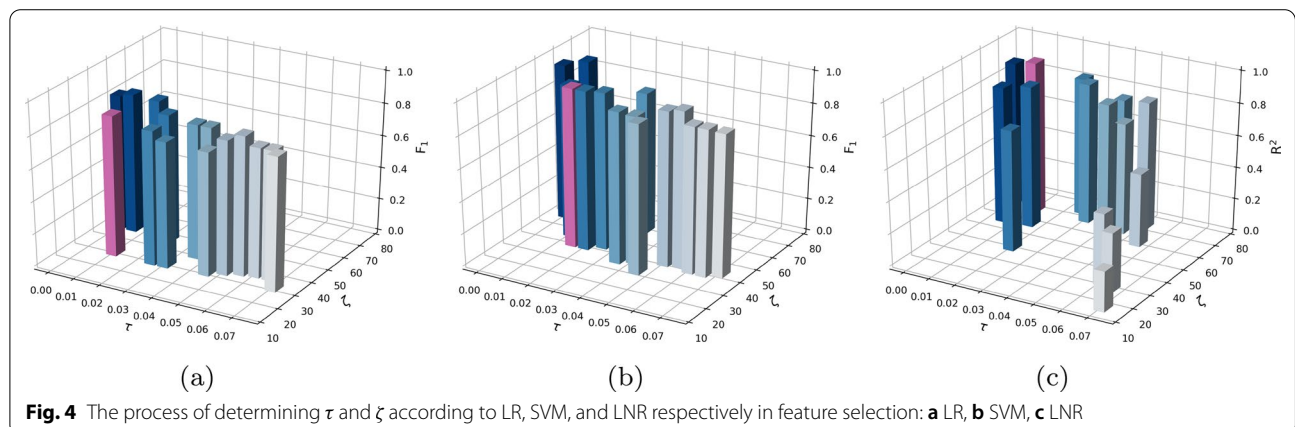
in  $F_1$  score, where  $P$  value between methods are much less than 0.01. Similarly, Table 6 shows that the LNR-DF with two-stage feature selection is superior to other methods that do not employ  $\beta/\alpha$  features and two-stage feature selection simultaneously. The statistical analysis results mean that the  $\beta/\alpha$  features and two-stage feature selection employed in our framework are effective and necessary.

**Selected feature analysis**

Figure 5 shows the shared selected features by MDD detection model SVM-DF and MDD severity assessment model LNR-DF. Obviously, we can see that the weight importance of selected features is quite different between SVM-DF and LNR-DF due to the difference in their principles for defining models. However, the direction of weight importance is identical between SVM-DF and LNR-DF, which means that the selected features have similar effects on the performance of MDD detection and MDD severity assessment. Meanwhile, it is observed that three derived  $\beta/\alpha$  features of  $F_{R-\beta/\alpha}^{P4-A2}$ ,  $F_{R-\beta/\alpha}^{CZ-A1}$ , and  $F_{R-\beta/\alpha}^{FP2-A2}$  are selected by both MDD detection and MDD severity assessment models, which are more than 21 %. It further verifies our findings of derived features are quite effective in the models of MDD detection and severity assessment.

**Limitations**

Since the size of the dataset used in this study is 92, our study is a pilot study, and we employ the SVM, LR, and LNR as the base model for MDD detection and MDD severity assessment, with the merits of small-size building models. To further conquer the over-fitting risk, the cross-validation technique is also utilized. Even so, it still potentially has the over-fitting risk for MDD detection and MDD severity assessment models.



**Fig. 4** The process of determining  $\tau$  and  $\zeta$  according to LR, SVM, and LNR respectively in feature selection: **a** LR, **b** SVM, **c** LNR

**Table 5** Statistical power analysis between SVM-DF with two-stage feature selection and other MDD detection methods in the framework

Feature selection	Model	P value				
		Sec	Prec	Spec	Acc	F <sub>1</sub>
-	LR	1.47e-05***	9.73e-06***	7.91e-06***	1.58e-06***	9.40e-06***
	SVM	1.52e-05***	7.80e-06***	1.67e-04***	6.50e-06***	2.80e-06***
	LR-DF	8.90e-07***	6.39e-06***	2.43e-05***	3.95e-06***	4.27e-06***
	SVM-DF	2.72e-04***	4.82e-05***	3.28e-05***	1.70e-06***	3.67e-05***
PCC	LR	4.81e-05***	7.46e-06***	2.64e-05***	1.53e-06***	3.66e-06***
	SVM	6.10e-06***	2.76e-06***	3.39e-04***	1.11e-05***	7.80e-07***
	LR-DF	8.37e-06***	1.44e-05***	1.90e-04***	5.78e-06***	1.30e-07***
	SVM-DF	4.61e-04***	2.99e-06***	2.30e-04***	2.44e-06***	1.26e-05***
RFE	LR	3.88e-04***	1.42e-03***	3.99e-04***	2.88e-04***	4.41e-04***
	SVM	8.64e-03***	2.59e-03***	4.07e-03***	2.55e-03***	2.61e-03***
	LR-DF	1.97e-03***	1.88e-02**	2.18e-02**	7.23e-05***	1.28e-04***
	SVM-DF	1.61e-02**	3.74e-01*	3.85e-01*	5.37e-02*	2.86e-02**
PCC and RFE	LR	1.44e-04***	5.34e-04***	1.78e-04***	5.68e-05***	1.21e-04***
	SVM	2.08e-01*	1.70e-01*	1.84e-01*	9.50e-03***	3.60e-03***
	LR-DF	3.16e-03***	4.47e-03***	3.99e-03***	7.74e-04***	9.18e-04***

\*P ≥ 0.05; \*\*0.05 > P ≥ 0.01; \*\*\*P < 0.01

**Table 6** Statistical power analysis between LNR-DF with two-stage feature selection and other MDD severity assessment methods in the framework

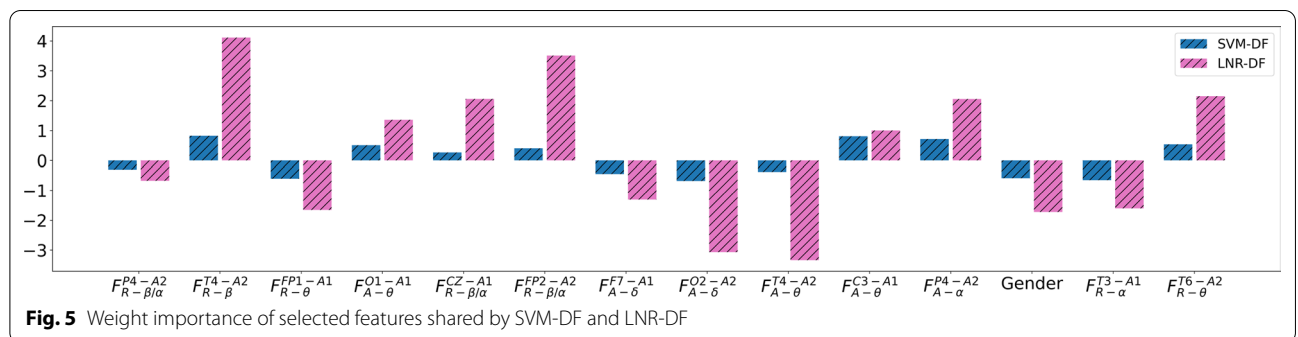
Feature selection	Model	P value		
		LMAE	LMSE	R <sup>2</sup>
-	LNR	7.43e-06***	4.61e-05***	2.43e-04***
	LNR-DF	4.07e-06***	5.48e-05***	9.15e-05***
PCC	LNR	4.00e-08***	6.00e-08***	4.00e-08***
	LNR-DF	7.90e-07***	1.40e-07***	3.00e-08***
RFE	LNR	1.60e-02**	1.81e-02**	1.63e-02**
	LNR-DF	3.79e-03***	7.15e-03***	2.78e-03***
PCC and RFE	LNR	1.25e-02**	1.44e-02**	1.10e-02**

\*P ≥ 0.05; \*\*0.05 > P ≥ 0.01; \*\*\*P < 0.01

On the other hand, the original EEG signals cannot be obtained from the commercial EEG device named Nerron-spectrum-5 as the data format for medical devices is usually commercial confidence. Herewith, the features used in this study are extracted by the EEG device directly. The noise would be introduced into extracted features due to the uncertainty of EEG signals quality, which would reduce the generality of the proposed framework potentially.

**Conclusions**

In this study, we propose an automatic EEG-based MDD detection framework. In the proposed framework, the high MDD correlation features named β/α features are derived; a two-stage feature selection method is employed



for well-selected correlated features to improve the model performance; three sample-sized machine learning methods of LR, SVM, and LNR as base models are utilized for MDD detection and MDD severity assessment. Experiment results show that the proposed framework achieves an optimal  $F_1$  score of 0.9846 in MDD detection, and achieves an optimal  $R^2$  of 0.9479 in MDD severity assessment. Meanwhile, the findings of the derived EEG-based  $\beta/\alpha$  features can greatly improve the performance of MDD detection and MDD severity assessment. For MDD detection, the derived features would improve the  $F_1$  score by over 2% and 11% in terms of SVM and LR, respectively. As for MDD severity assessment, the derived features would improve the  $R^2$  of LNR from 0.8812 to 0.9479. Meanwhile, the base models used in this study are LR, SVM with linear kernel, and LNR, all of which are linear and with the merits of small sample size and interpretability. It means that the proposed framework can be potentially deployed into a medical system to provide MDD detection services to help physicians to screen out MDD patients.

Concerning the sample size of our experimental data is not big enough, in future work, we are intended to continue collecting more samples to expand our dataset and time retrain the model in the case of data added incrementally [50], which will better improve the generalization of our framework. In addition, more samples can be generated using sophisticated augmentation methods like Generative Adversarial Networks (GANs) [51]. Furthermore, we will consider using interpretable deep learning models [52] to improve the interpretability of our framework, to better assist clinical diagnosis.

#### Author contributions

XMf, YSS, and GSZ conceived and designed the study. YSS and GSZ prepared the experimental equipment and resources. XXH and HBY collected the data. YJL, XXH and HBY analyzed the data. XMf, YSS and WJM interpreted the results. YJL, WJM and XMf wrote the manuscript. All authors read and approved the manuscript.

#### Funding

This work is supported partially by Philosophy and Social Sciences of Guangdong Province during the 13th Five-Year Plan period (GD18CJY05), Humanities and Social Sciences Research Foundation of Ministry of Education in 2021 (21YJAZH072).

#### Availability of data and materials

The data that support the findings of this study are available from Shenzhen Traditional Chinese Medicine Hospital but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the author named Xingxian Huang upon reasonable request and with permission of Shenzhen Traditional Chinese Medicine Hospital.

#### Declarations

##### Ethics approval and consent to participate

The study protocol was approved by the Institutional Review Board at the Shenzhen Traditional Chinese Medicine Hospital, Shenzhen, Guangdong,

China. All methods were carried out in accordance with relevant guidelines and regulations. All participants provided informed consent.

##### Consent for publication

Not applicable.

##### Competing interest

The authors declare that they have no competing interests.

##### Author details

<sup>1</sup>School of Computer Science, South China Normal University, Guangzhou, China. <sup>2</sup>College of Big Data and Internet, Shenzhen Technology University, Shenzhen, China. <sup>3</sup>Department of Acupuncture and Moxibustion, Shenzhen Traditional Chinese Medicine Hospital, Shenzhen, China.

Received: 20 November 2021 Accepted: 29 July 2022

Published online: 06 August 2022

#### References

- Otte C, Gold SM, Penninx BW, Pariante CM, Etkin A, Fava M, et al. Major depressive disorder. *Nat Rev Dis Primers*. 2016;2(1):1–20.
- Lozano R, Fullman N, Abate D, Abay SM, Abbafati C, Abbasi N, et al. Measuring progress from 1990 to 2017 and projecting attainment to 2030 of the health-related sustainable development goals for 195 countries and territories: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*. 2018;392(10159):2091–138.
- Rohan KJ, Rough JN, Evans M, Ho SY, Meyerhoff J, Roberts LM, et al. A protocol for the Hamilton Rating Scale for Depression: item scoring rules, rater training, and outcome accuracy with data on its application in a clinical trial. *J Affect Disord*. 2016;200:111–8.
- Chakladar DD, Chakraborty S. Feature extraction and classification in brain-computer interfacing: future research issues and challenges. In: *Natural computing for unsupervised learning*, Springer; 2019. p. 101–131.
- Chakladar DD, Chakraborty S. EEG based emotion classification using “Correlation Based Subset Selection”. *Biol Inspired Cognit Archit*. 2018;24:98–106.
- Ullah I, Hussain M, Aboalsamh H, et al. An automated system for epilepsy detection using EEG brain signals based on deep learning approach. *Expert Syst Appl*. 2018;107:61–71.
- Zhang Y, Guo Y, Yang P, Chen W, Lo B. Epilepsy seizure prediction on EEG using common spatial pattern and convolutional neural network. *IEEE J Biomed Health Inform*. 2019;24(2):465–74.
- Ieracitano C, Mammine N, Hussain A, Morabito FC. A novel multi-modal machine learning based approach for automatic classification of EEG recordings in dementia. *Neural Netw*. 2020;123:176–90.
- Hosseinifard B, Moradi MH, Rostami R. Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal. *Comput Methods Programs Biomed*. 2013;109(3):339–45.
- Khodayari-Rostamabad A, Reilly JP, Hasey GM, de Bruin H, MacCrimmon DJ. A machine learning approach using EEG data to predict response to SSRIs treatment for major depressive disorder. *Clin Neurophysiol*. 2013;124(10):1975–85.
- Erguzel TT, Ozekes S, Tan O, Gultekin S. Feature selection and classification of electroencephalographic signals: an artificial neural network and genetic algorithm based approach. *Clin EEG Neurosci*. 2015;46(4):321–6.
- Bachmann M, Lass J, Suhhova A, Hinrikus H. Spectral asymmetry and Higuchi's fractal dimension measures of depression electroencephalogram. *Comput Math Methods Med*. 2013. <https://doi.org/10.1155/2013/251638>.
- Cai H, Han J, Chen Y, Sha X, Wang Z, Hu B, et al. A pervasive approach to EEG-based depression detection. *Complexity*. 2018. <https://doi.org/10.1155/2018/5238028>.
- Acharya UR, Sudarshan VK, Adeli H, Santhosh J, Koh JE, Puthankatti SD, et al. A novel depression diagnosis index using nonlinear features in EEG signals. *Eur Neurol*. 2015;74(1–2):79–83.

15. Mahato S, Paul S. Detection of major depressive disorder using linear and non-linear features from EEG signals. *Microsyst Technol.* 2019;25(3):1065–76.
16. Mohammadi M, Al-Azab F, Raahemi B, Richards G, Jaworska N, Smith D, et al. Data mining EEG signals in depression for their diagnostic value. *BMC Med Inform Decis Mak.* 2015;15(1):1–14.
17. Mumtaz W, Xia L, Ali SSA, Yasin MAM, Hussain M, Malik AS. Electroencephalogram (EEG)-based computer-aided technique to diagnose major depressive disorder (MDD). *Biomed Signal Process Control.* 2017;31:108–15.
18. Mumtaz W, Ali SSA, Yasin MAM, Malik AS. A machine learning framework involving EEG-based functional connectivity to diagnose major depressive disorder (MDD). *Med Biol Eng Comput.* 2018;56(2):233–46.
19. Walambe R, Nayak P, Bhardwaj A, Kotecha K. Employing multimodal machine learning for stress detection. *J Healthc Eng.* 2021. <https://doi.org/10.1155/2021/9356452>.
20. Fan X, Yao Q, Cai Y, Miao F, Sun F, Li Y. Multiscaled fusion of deep convolutional neural networks for screening atrial fibrillation from single lead short ECG recordings. *IEEE J Biomed Health Inform.* 2018;22(6):1744–53.
21. Fan X, Wang H, Xu F, Zhao Y, Tsui KL. Homecare-oriented intelligent long-term monitoring of blood pressure using electrocardiogram signals. *IEEE Trans Industr Inf.* 2019;16(11):7150–8.
22. Ma B, Li C, Wu Z, Huang Y, van der Zijp-Tan AC, Tan S, et al. Muscle fatigue detection and treatment system driven by internet of things. *BMC Med Inform Decis Mak.* 2019;19(7):1–9.
23. Yasin S, Hussain SA, Aslan S, Raza I, Muzammel M, Othmani A. EEG based major depressive disorder and bipolar disorder detection using neural networks: a review. *Comput Methods Programs Biomed.* 2021;202: 106007.
24. Kaur C, Bisht A, Singh P, Joshi G. EEG signal denoising using hybrid approach of variational mode decomposition and wavelets for depression. *Biomed Signal Process Control.* 2021;65: 102337.
25. Sharma G, Parashar A, Joshi AM. DepHNN: a novel hybrid neural network for electroencephalogram (EEG)-based screening of depression. *Biomed Signal Process Control.* 2021;66: 102393.
26. Seal A, Bajpai R, Agnihotri J, Yazidi A, Herrera-Viedma E, Krejcar O. DeprNet: a deep convolution neural network framework for detecting depression using EEG. *IEEE Trans Instrum Meas.* 2021;70:1–13.
27. Akbari H, Sadiq MT, Rehman AU, Ghazvini M, Naqvi RA, Payan M, et al. Depression recognition based on the reconstruction of phase space of EEG signals and geometrical features. *Appl Acoust.* 2021;179: 108078.
28. Mohammadi Y, Moradi MH. Prediction of depression severity scores based on functional connectivity and complexity of the EEG signal. *Clin EEG Neurosci.* 2021;52(1):52–60.
29. Jiang C, Li Y, Tang Y, Guan C. Enhancing EEG-based classification of depression patients using spatial information. *IEEE Trans Neural Syst Rehabil Eng.* 2021;29:566–75.
30. Sadiq MT, Akbari H, Siuly S, Yousaf A, Rehman AU. A novel computer-aided diagnosis framework for EEG-based identification of neural diseases. *Comput Biol Med.* 2021;138: 104922.
31. Ciarleglio A, Petkova E, Harel O. Elucidating age and sex-dependent association between frontal EEG asymmetry and depression: an application of multiple imputation in functional regression. *J Am Stat Assoc.* 2021;117:1–15.
32. Avots E, Jermakovs K, Bachmann M, Päske L, Ozcinar C, Anbarjafari G. Ensemble approach for detection of depression using EEG features. *Entropy.* 2022;24(2):211.
33. Bailey NW, Krepel N, van Dijk H, Leuchter AF, Vila-Rodríguez F, Blumberger DM, et al. Resting EEG theta connectivity and alpha power to predict repetitive transcranial magnetic stimulation response in depression: a non-replication from the ICON-DB consortium. *Clin Neurophysiol.* 2021;132(2):650–9.
34. Ramirez R, Vamvakousis Z. Detecting emotion from EEG signals using the emotive epoc device. In: *International conference on brain informatics.* Springer; 2012. p. 175–184.
35. Rajpurkar P, Yang J, Dass N, Vale V, Keller AS, Irvin J, et al. Evaluation of a machine learning model based on pretreatment symptoms and electroencephalographic features to predict outcomes of antidepressant treatment in adults with depression: a prespecified secondary analysis of a randomized clinical trial. *JAMA Netw Open.* 2020;3(6):e206653–e206653.
36. Bommert A, Welchowski T, Schmid M, Rahnenführer J. Benchmark of filter methods for feature selection in high-dimensional gene expression survival data. *Brief Bioinform.* 2022;23(1):bbab354.
37. Maldonado J, Riff MC, Neveu B. A review of recent approaches on wrapper feature selection for intrusion detection. *Expert Syst Appl.* 2022;198: 116822.
38. Zhou H, Deng Z, Xia Y, Fu M. A new sampling method in particle filter based on Pearson correlation coefficient. *Neurocomputing.* 2016;216:208–15.
39. Huang X, Zhang L, Wang B, Li F, Zhang Z. Feature clustering based support vector machine recursive feature elimination for gene selection. *Appl Intell.* 2018;48(3):594–607.
40. Austin PC, Merlo J. Intermediate and advanced topics in multilevel logistic regression analysis. *Stat Med.* 2017;36(20):3257–77.
41. Hosmer Jr DW, Lemeshow S, Sturdivant RX. *Applied logistic regression*, vol. 398. Wiley; 2013.
42. Uddin S, Khan A, Hossain ME, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. *BMC Med Inform Decis Mak.* 2019;19(1):1–16.
43. Suthaharan S. Support vector machine. In: *Machine learning models and algorithms for big data classification.* Springer; 2016. p. 207–235.
44. Schmidt AF, Finan C. Linear regression and the normality assumption. *J Clin Epidemiol.* 2018;98:146–51.
45. Pandis N. Linear regression. *Am J Orthod Dentofac Orthop.* 2016;149(3):431–4.
46. Qi J, Du J, Siniscalchi SM, Ma X, Lee CH. On mean absolute error for deep neural network based vector-to-vector regression. *IEEE Signal Process Lett.* 2020;27:1485–9.
47. Ohno S, Shiraki T, Tariq MR, Nagahara M. Mean squared error analysis of quantizers with error feedback. *IEEE Trans Signal Process.* 2017;65(22):5970–81.
48. Wang X, Jiang B, Liu JS. Generalized R-squared for detecting dependence. *Biometrika.* 2017;104(1):129–39.
49. Mahato S, Paul S. Classification of depression patients and normal subjects based on electroencephalogram (EEG) signal using alpha power and theta asymmetry. *J Med Syst.* 2020;44(1):1–8.
50. Sanghani G, Kotecha K. Incremental personalized E-mail spam filter using novel TFDCR feature selection with dynamic feature update. *Expert Syst Appl.* 2019;115:287–99.
51. Chaudhari P, Agrawal H, Kotecha K. Data augmentation using MG-GAN for improved cancer classification on gene expression data. *Soft Comput.* 2020;24(15):11381–91.
52. Joshi G, Walambe R, Kotecha K. A review on explainability in multimodal deep neural nets. *IEEE Access.* 2021.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

