Biomedical Signal Processing and Control 31 (2017) 108–115

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Electroencephalogram (EEG)-based computer-aided technique to diagnose major depressive disorder (MDD)☆

Wajid Mumtaz[a], Likun Xia[d], Syed Saad Azhar Ali[a], Mohd Azhar Mohd Yasin[b], Muhammad Hussain[c], Aamir Saeed Malik[a,*]

[a] Centre for Intelligent Signal and Imaging Research (CISIR),Universiti Teknologi PETRONAS, 32610, Bandar Seri Iskandar, Perak, Malaysia
[b] Department of Psychiatry, Universiti Sains Malaysia, Jalan Hospital Universiti Sains Malaysia,16150 Kubang Kerian, Kota Bharu, Kelantan, Malaysia
[c] Department of Computer Science, College of Computer and Information, Sciences, King Saud University, 12372, Riyadh, Saudi Arabia
[d] Beijing Institute of Technology, Beijing, 100081, China

ARTICLE INFO

Article history:
Received 4 December 2015
Received in revised form 15 July 2016
Accepted 21 July 2016
Available online 30 July 2016

Keywords:
Major depressive disorder
EEG-based diagnosing of depression
EEG-based linear features
Machine learning techniques for depression
Supervised classification
EEG alpha interhemispheric asymmetry
EEG spectral power

ABSTRACT

Recently, Electroencephalogram (EEG)-based computer-aided (CAD) techniques have shown their promise as decision-making tools to diagnose major depressive disorder (MDD) or simply depression. Although the research results have motivated the use of CAD techniques to help assist psychiatrists in clinics yet their clinical translation has been less clear and remains a research topic. In this paper, a proposed machine learning (ML) scheme was tested and validated with resting-state EEG data involving 33 MDD patients and 30 healthy controls. The EEG-derived measures such as power of different EEG frequency bands and EEG alpha interhemispheric asymmetry were investigated as input features to the proposed ML scheme to discriminate the MDD patients and healthy controls, and to prove their feasibility for diagnosing depression. The acquired EEG data were subjected to noise removal and feature extraction. As a result, a data matrix was constructed by the columns-wise concatenation of the extracted features. Furthermore, the z-score standardization was performed to standardize each column of the data matrix according to its mean and variance. The data matrix may have redundant and irrelevant features; therefore, to determine the most significant features, a weight was assigned to each feature based on its ability to separate the target classes according to the criterion, i.e., receiver operating characteristics (roc). Hence, only the most significant features were used for testing and training the classifier models: Logistic regression (LR), Support vector machine (SVM), and Naïve Bayesian (NB). Finally, the classifier models were validated with 10-fold cross-validation that has provided the performance metrics such as test accuracy, sensitivity, and specificity. As a result of the investigations, most significant features such as EEG signal power and EEG alpha interhemispheric asymmetry from the brain areas such as frontal, temporal, parietal and occipital were found significant. In addition, the proposed ML framework proved automatic identification of aberrant EEG patterns specific to disease conditions and provide high classification results i.e., LR classifier (accuracy = 97.6%, sensitivity = 96.66%, specificity = 98.5%), NB classification (accuracy = 96.8%, sensitivity = 96.6%, specificity = 97.02%), and SVM (accuracy = 98.4%, sensitivity = 96.66%, specificity = 100%). In conclusion, the proposed ML scheme along with the EEG signal power and EEG alpha interhemispheric asymmetry are proved suitable as clinical diagnostic tools for MDD.

1. Introduction

Major depressive disorder (MDD) is a chronic, recurrent, and a life-threatening mental illness commonly termed as depression. In the USA, it has been reported as highly prevalent (13.2% to 16%) and has been found as a leading cause of functional disability [1]. According to World Health Organization (WHO), in the year 2020, it has become the 2nd most leading cause of disease burden worldwide [2]. An accurate and early diagnosis is recommended by clinicians in order to avoid increased risk of treatment failure with

☆ Financial support for this study was provided by NSTIP strategic technologies programs, grant number (12-INF2582-02), in the Kingdom of Saudi Arabia. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

* Corresponding author.
E-mail address: aamir.saeed@petronas.com.my (A.S. Malik).

http://dx.doi.org/10.1016/j.bspc.2016.07.006
1746-8094/

an increase in the time of onset. In addition, an early diagnosis will further help during the treatment process and helps to improve patients quality-of-life [3]. Currently, the diagnosis includes structured questionnaires that are administered as an interview session between health practitioners and the MDD patients. However, the subjectivity involved due to the heterogeneity of MDD and the potential errors incurred by human factors may not be ignored and could result into a misdiagnosis.

Recently, machine learning (ML) techniques have received considerable attention due to their capability to mine non-invasive neuroimaging data to establish the computer-aided (CAD)-based solutions that facilitate during the diagnosis [4–6]. For example, mining functional magnetic resonance imaging (fMRI) data with ML methods has shown promising research results [7,8]. Specifically, the support vector machine (SVM) has been emphasized as a method of choice for the diagnosis of depression [9]. On the other hand, the automated EEG-based ML methods are proved feasible to discriminate the depressed patients from healthy controls [10–14]. In addition to depression, the classification algorithms are found useful for neurological diseases such as schizophrenia, and Alzheimer's disease [15]. In the context of depression, classifier such as artificial neural networks (ANN) is trained to classify the depressed and healthy controls [16,17]. Recently, a depression diagnostic index is proposed based on nonlinear features extracted from EEG data [11]. Electroencephalogram (EEG) offers high temporal resolution and lower cost than fMRI which makes it suitable for portable and remote clinical applications involving monitoring epileptic patients [18,19], quantifying different sleep stages [20], indexing for anesthesia monitoring [21], and diagnosing patients with alcohol addiction [22]. Furthermore, EEG data acquisition is faster than fMRI and a trained nursing staff could handle the EEG-based CAD system.

In the literature, various nonlinear features such as detrended fluctuation analysis (DFA), Higuchi fractal, correlation dimension and Lyapunov exponent are extracted from EEG signal and have shown promises during the MDD diagnosis, e.g., a recently performed study has achieved 90% accuracy for discriminating the MDD patients and healthy controls [12]. MDD has been associated with cognitive deficits and functional impairments [23,24] localized to areas such as frontal and temporal. According to a recent review, abnormalities such as MDD tends to exhibit decreased left frontal activity (measured as increased interhemispheric alpha power/amplitude) [25]. In a study, greater left frontal activity is associated with less depressive symptoms [26]. In addition, EEG alpha interhemispheric asymmetry is concluded as a risk marker for MDD because of the finding that the study participants with lifetime depressive symptoms has shown less relative frontal activity when compared with subjects with no depression [27].

The importance of EEG alpha interhemispheric asymmetry in the diagnosis of depression is evident from various studies [28–30]. For example, psychomotor retardation during depression is linked with EEG alpha interhemispheric asymmetry [28]; EEG frontal asymmetry has been considered as a marker for vulnerability of depression [29]; decreased interhemispheric alpha waves are reported during depression [30]. In addition, altered structure of EEG oscillatory pattern is reported in MDD [25]. Hence, these evidences add to the confidence on the EEG alpha interhemispheric asymmetry, as a feature, to be used for automatic diagnosis of depression.

In addition to alpha band, activity in other bands such as theta band has shown relevance such as a decreased frontal theta activity has also been reported [31–33]. Moreover, hypo-activation of the left frontal [34,35] and hyperactivation in the right frontal regions [31] have been observed during MDD. Despite all research findings, the clinical applications of the frontal EEG alpha interhemispheric asymmetry and frontal midline theta activity have been largely unclear [36]. Hence, in context of diagnosing MDD, a further investigation involving features such as EEG alpha interhemispheric asymmetry, and EEG spectral power of different frequency bands is required.

To fill-up this gap, this study proposes a novel ML technique including EEG-derived features (spectral power computations for different EEG frequency bands and EEG alpha interhemispheric asymmetry) as input data. In this research, it has been hypothesized that the linear features such as EEG power computed form different frequency bands and EEG alpha interhemispheric asymmetry can discriminate the MDD patients and healthy controls with high classification efficiencies even in the absence of nonlinear EEG features. For this purpose, the proposed ML method is validated with resting-state EEG data acquired from the MDD patients and healthy controls. The feature selection involved 2 steps, 1) ranking features in descending order according to receiver operating characteristics (roc) criterion, 2) selecting subsets including top-ranked features to train and test classifier models, and to determine the upper limit of a minimum number of features necessary to provide highest accuracies among other subsets. As a result, a reduced set of features is used to test and train the classifiers such as logistic regression (LR), support vector machine (SVM) and naïve Bayesian (NB). The classifiers provide the functional model to relate the EEG significant features to the outcome target groups, i.e., MDD patients and healthy controls.

2. Materials and methods

2.1. Study participants

In this study, two groups of participants were recruited: 1) 33 MDD patients (Age, Mean = 40.33, SD = ±12.861), and 2) 30 age-matched healthy subjects as control group (Age, Mean = 38.227, SD = ±15.64). The participants were recruited from the outpatient clinic of hospital Universiti Sains Malaysia (HUSM), Malaysia. The experiment design was approved by the ethics committee, HUSM. The participants agreed to sign the consent forms of participation and were fully aware of the experimental procedure adopted for experimental data acquisition. Furthermore, according to the recruitment criteria of this study, the MDD participants should have confirmed diagnosis based on the symptoms of depression as mentioned in the Diagnostic and Statistical Manual for depression (DSM-IV) [37]. The diagnosis was performed by senior psychiatrists in the psychiatric clinic, HUSM. In this study, the MDD participants with psychotic symptoms, pregnant patients, alcoholics, smokers and patients having epileptic problems were excluded. On the other hand, the healthy controls were screened for possible mental or physical illness and were found disease naive.

2.2. Experimental data acquisition

The EEG data acquisition involved vigilance-controlled monitoring during the recordings, i.e., 5 min EEG data recordings during eyes closed (EC) and 5-min recordings during eyes open (EO) conditions involving a 19-channel EEG cap with linked-ear (LE) reference. The experimental data were recorded at the same time of day and the participants were instructed to abstain from coffee intake, and other drug abuses. The EEG sensors were placed on the scalp according to the international 10–20 electrode placement standard [38]. The 19-electrodes covering the scalp included the frontal (Fp1, Fp2, F3, F4, F7, F8, Fpz), temporal (T3, T4, T5, T6), parietal (P3, P4, P7, P8), occipital (O1, O2), and central (C3, C4) regions. Moreover, the 19-channels EEG cap was attached to an amplifier from Brain Master Systems with configurations such as 0.5 Hz to 70 Hz filter, with a 50 Hz notch filter, and a sample rate of 256 samples per

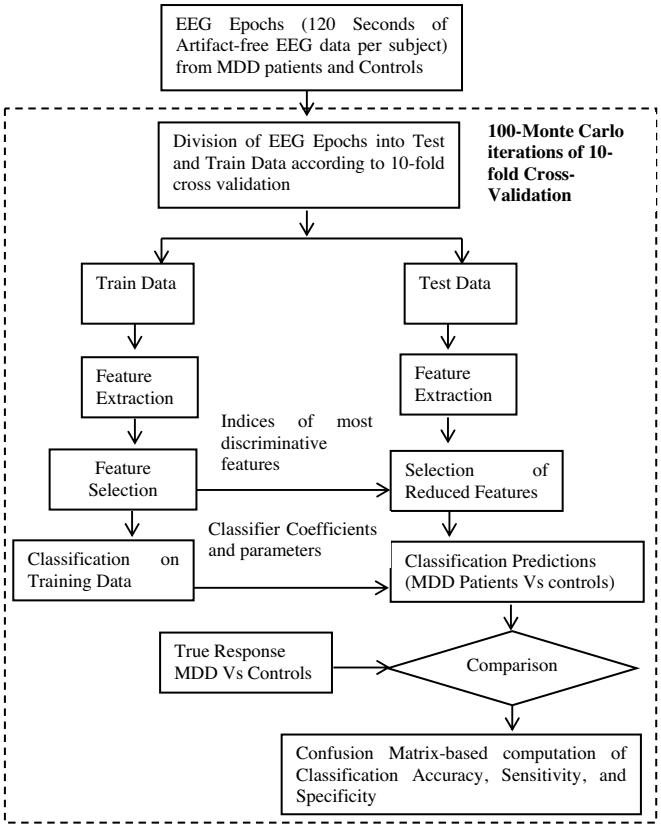

Fig. 1. The Proposed ML framework to diagnose patients with depression.

second. The settings were used to discretize the EEG data before storage to a computer disk for later analysis such as preprocessing and EEG analysis. For the sake of EEG analysis, the EEG data were re-referenced to infinity reference (IR) [39]. The severity of MDD was assessed using clinical questionnaires: Beck depression inventory-II (BDI-II) [40,41] and hospital anxiety and depression scale (HADS) [42]. In this study, the clinical questionnaires were considered as gold standards that were utilized during validation of the proposed method.

2.3. Noise removal from the EEG data

In this study, the recorded EEG data were confounded with composite signals generated by sources other than the neuronal activity and were considered as noises. For example, the EEG data may be confounded with different type of noises such as eye blinks, movements and muscular activates (e.g., the heart beats). Unfortunately, the EEG data confounded with these artifacts may not truly represent the underlying brain activities. Hence, to clean the EEG data was an essential pre-processing step to continue further for data analysis. In this paper, removal of these noises included multiple source techniques [43] implemented in the standard brain electric source analysis (BESA) software [44]. According to the technique, the EEG artifact topographies were constructed based on the artifact vectors learned from the recorded EEG data. In addition, a head model integrated with the EEG artifact topographies were utilized to correct the EEG data confounded with eye blinks, or other artifacts as mentioned.

Fig. 1 shows the proposed ML scheme involving the clean EEG input data from the MDD patients and healthy controls. The input data were randomly sub-divided into test and train sets including 100 iterations of the ten-fold cross validation (10-CV). Each block in Fig. 1 is described in the subsequent section.

2.4. Feature extraction

For feature extraction, two minutes of artifact-free EEG epochs were selected from EC and EO data per study subject. As a result, a large number (*Nc*) of candidate features (power computation at different frequency bands and EEG alpha asymmetry) were achieved and formed a data matrix. In the literature, MDD has been associated with a dysfunction in different brain areas, especially, frontal and temporal regions. Hence, it would be useful to utilizing the spatial locations due to the spread of the EEG channels over the scalp. Averaging across the channels would reduce the spatial information; hence, it was not performed.

2.4.1. EEG alpha interhemispheric asymmetry

The difference of an observation such as EEG signal power between the left and right hemisphere can be observed and computed with a quantity such as the interhemispheric asymmetry [45]. Eqs. (1) and (2) describe mathematical formulas to compute powers in the desired frequency bands:

$$W'_{Lmn} = \sum_{f=f_1}^{f_2} s_{Lmn} \Big/ \sum_{f=0.5Hz}^{30Hz} s_{Lmn} \tag{1}$$

$$W'_{Rmn} = \sum_{f=f_1}^{f_2} s_{Rmn} \Big/ \sum_{f=0.5Hz}^{30Hz} s_{Rmn} \tag{2}$$

where the lower and upper frequencies of the band limited EEG signal power computation were represented by f_1 and f_2, respectively. Power spectral densities were represented as S_{Lmn} and S_{Rmn} such as they represent the left and right hemispheres, respectively. Finally, the Eq. (3) showed formula for calculating the interhemispheric EEG asymmetry:

$$A_{mn}(f_1, f_2) = \frac{W'_{Lmn} - W'_{Rmn}}{W'_{Lmn} + W'_{Rmn}} \times 100 \tag{3}$$

The EEG alpha interhemispheric asymmetry was computed for each channel pair involving frontal (Fp1, Fp2, F3, F4, F7, F8, Fpz), temporal (T3, T4, T5, T6), parietal (P3, P4, P7, P8), occipital (O1, O2), and central (C3, C4). For example, EEG alpha asymmetry computed for Fp1 included channel pairs such as Fp1-Fp2; Fp1-F4; Fp1-F8; Fp1-T4; Fp1-T6; Fp1-P4; Fp1-P8; Fp1-O2; Fp1-C4. In addition to computing features for classification, investigation of an overall trend such as either increasing or decreasing values between left and right hemisphere of the depressed and healthy controls, the averaging of EEG alpha interhemispheric asymmetry was performed region-wise, e.g., only averaging frontal EEG channels pairs located in the frontal region.

2.4.2. EEG spectral power

In this study, the Welch periodogram method was adopted to compute EEG signal powers involving the Hanning window [46]. The periodogram method computed EEG signal power by first segmenting the signal into small datasets with a 50% overlap. Furthermore, the spectrum was computed for each segment and averaged over all the segments to finally get power of that EEG signal. The spectral power computed was absolute and the computed statistical significance for the difference was less than 0.01 ($p < 0.01$). In the literature, MDD has been associated with a dysfunction in different brain areas, especially, frontal and temporal regions.

In order to perform classification, the features were arranged column-wise, each column was denoted as x_i, where $i = 1 \ldots N_c$. The rows of the matrix represented MDD patients. The feature space denoted by $L = [(x_i, y_i),\ i = 1 \ldots N_c]$ included both the feature

space matrix and the corresponding output class labels, $y_=$ [*MDD, Controls*]. Finally, the resulting matrix was rectangular in shape including high- dimensional datasets, i.e., the number of rows (*data points* = 126). The feature space matrix achieved so far might not be centered and also unequally distributed. Therefore, in order to eliminate the possible outliers, the data standardization based on z-scores was performed. The z-transformations were computed column-wise by subtracting each element value with its column-wise mean and divided by corresponding standard deviation.

In addition to computing features for classification, the averaging of EEG power in different frequency bands was performed for each region to show an overall trend such as increasing or decreasing values between the depressed and healthy controls. Moreover, it is possible that most of the features extracted by this procedure might be either redundant or irrelevant. The feature selection is desirable to reduce feature space, from N_c to a lower dimension, i.e., N_r.

2.5. Feature selection

For high dimensional datasets, feature selection remains as challenging research topic and carries critical importance during data analysis involving a typical ML methodology. From the classification point of view, this high dimensionality may easily over-fit or underfit a classification model. To enhance the classification performance and to reduce the irrelevancy and redundancy of the feature space, most significant features were ranked according to the specified criterion.

In this study, the features were ranked according to criterion i.e., receiver operating characteristics (ROC) [47]. According to the criterion, the area between the empirical ROC curve and the random classifier slope was computed for individual features. Each feature was assigned a weight according to the covered area and arranged in descending order of the weight values (z-values), accordingly. The values were computed based on their areas under the curve (AUC) for individual features, i.e., higher values of the metric z directly reflect the ability of a feature to discriminate the MDD patients and controls. The z-values may vary between the range 0 and 0.5 indicating bad to good classification ability, respectively. Different subsets of these ranked features, e.g., top ranked 5, 10, 15, and 19 features were iteratively used during classification. The optimal number of reduced set of features was determined based on highest classification performance among others.

To find minimum number of features that would be sufficient to train the classifier models without over-fitting, an empirical process was adopted. During feature selection stage, the features were ranked in descending order according to their z-values. The selection of minimum number of features was based on iteratively observing performance of the classification models for each feature subsets selected from top 5, 10, 15, and 19 features. For each feature subset, a 100 times run of the simulations was performed involving 10-fold cross-validation to achieve the box plot representations of the accuracies, sensitivities, and specificities. Since the individual iteration resulted in 100 different values of performance metrics (the accuracy, the sensitivity and the specificity), the final confusion matrix was computed by averaging over 100 times. Hence, the feature subset that provided highest accuracies was finally selected as a minimum number of features.

2.6. Classification models

In this study, the reduced set of features was considered as the independent variables and the corresponding treatment outcomes (MDD patients and controls) were the dependent variables. In the case of LR classifier [48], the relationship between the reduced set of features was modeled according to the logistic function as described in Eq. (4). For LR classifier, the coefficients estimation was based on maximum likelihood method. The LR classifier resulted in a likelihood value $l(x)$, where $0 \leq l(x) \leq 1$, which was an indication of subjects, associated either with MDD patients or controls. If $l(x)$ was greater than the threshold = 0.5, the subject was declared as MDD patients, and otherwise associated with the control group.

$$E(Y/x) = \frac{e^x}{1 + e^x} \tag{4}$$

The second classification model employed was SVM classifier with linear kernel. It can classify the feature space based on a 'hyperplane' that separated MDD patients and controls according to the class labels [49]. SVM is considered as a high-efficiency classifier model and used here for comparison purposes. According to SVM, a linear decision boundary can be found based on this high dimensional space. Use of a linear kernel instead of a non-linear kernel reduced the risk of over-fitting the data and improves the performance for our data and significantly reduces the overall model complexity. In summary, the LR classifier generated probability values to cater MDD patients as either MDD patients or controls and the SVM concluded a hyperplane to achieve the maximum classification accuracy. The 3rd classification model was the NB classification [50], based on generating the conditional posterior probabilities for each sample involving the target groups, i.e., MDD patients and healthy controls. The classifier was formed by assigning the sample to the class for which the sample has higher posterior probabilities.

2.7. Validation

After classifier design, a fair evaluation requires assessment of its performance over a range of selected features and classifier design (with suitable coefficient values until convergence) that corresponds to large number of subjects. To address this consideration, we evaluated classification performance based on 10-fold cross validation by dividing the selected EEG features into 10 segments of equal sizes. During each round, 9 of the segments were utilized as training subset and the remaining 1 as test subset. Ten-fold cross-validation provides a fair test of validation in cases where the data points are limited while utilizing features for both test and train classifier models.

The features extracted from the EEG data were arranged in data matrices. The data matrices, from the study participants, were divided into training and testing sets according to the 10-fold cross-validation; which ensures independence between the testing and training sets. In addition, the procedures such as feature extraction, and feature selection were performed separately on the train and test sets. Only the indices of most significant features learned from training set were utilized to select features from the testing set. Hence, the leakage between the testing and training sets was controlled. In addition, the 10-fold cross- validation was repeated for 100 times that ensures the validity of the testing results. Finally, the Results section included only testing set accuracies, sensitivities, and specificities.

A confusion matrix provided the basis to calculate the classifier performance metrics such as accuracy, sensitivity, and specificities. A brief description of these performance metrics are provided in this subsection. Based on the confusion matrix, the Eqs. (5)–(7) were used to compute the metrics such as classifier sensitivity, specificity, and accuracy, respectively. The sensitivity of a binary classifier involving classes such as MDD patients and healthy controls, the sensitivity was defined as to the percentage of MDD patients (true positive: TP) correctly classified as MDD patients, and was defined by Eq. (5). The specificity was defined as the percentage of healthy controls (true negative: TN) which are correctly classified as healthy controls and defined mathematically by Eq.

Table 1
Comparison of EEG alpha Interhemispheric asymmetry between the MDD patients and healthy controls ($p < 0.01$).

Brain Regions	MDD Patients	Healthy Controls
Frontal	Right < left	Left < right
Parietal, Central, Temporal, Occipital	Left < Right	Right < left

(6). Finally, the accuracy of the models was defined as the percentage of correctly classified MDD patients and healthy controls among the study participants and defined mathematically in Eq. (7). The false positives (FP) and false negatives (FN) were defined as misclassifications and considered as errors.

$$Sensitivity = \frac{TP}{TP + FN} \quad (5)$$

$$Specificity = \frac{TN}{TN + FP} \quad (6)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

3. Results

3.1. EEG signal power

Fig. 2 shows EEG signal power differences between MDD patients and healthy controls including frontal, temporal, parietal, occipital and central region. According to the results, the MDD patients exhibited less theta and alpha signal powers in the all regions when compared with healthy controls. In addition, the frontal and occipital regions have shown less delta signal power in depressed patients as compared to healthy controls. However, for brain regions such as central, temporal, and parietal, the delta and beta bands have shown a slight increased EEG signal powers when compared between the depressed and healthy controls. The computed statistical significance for the difference of power between the MDD patients and healthy controls was less than 0.01 ($p < 0.01$) and expressed with an asterisk '*'.

3.2. EEG alpha interhemispheric asymmetry

According to Table 1, the depressed individuals tend to exhibit greater relatively right frontal activity (less alpha) when compared with healthy controls. However, regions such as temporal, parietal, central, and occipital did not show such behavior. In short, the depressed individual showed greater anterior EEG activity.

3.3. Parameters for classification

Table 2 shows specific values of the parameter assigned for each classifier during training and testing. Regarding the LR classifier, the link function showed the relationship between the EEG features and clinical outcomes. The value was set as 'logit' since the classifier used was logistic regression. Binary classification assumes binomial distribution; this was set as binomial because the data were supposed to originate from two classes, i.e., MDD patients and healthy controls. The offset value was set to 1, whereas the mathematical model of LR classifier has included a constant term. Regarding the SVM classifier, the C values were assigned as 0.787 for MDD patient's class and 1.3684 for the healthy control's class, accordingly. The values were computed with formula $(N/2 \times N_1)$ and $(N/2 \times N_2)$, respectively. The 'N' denoted total number of study participants; N_1 indicated the number of MDD patients and N_2 indicated the number of healthy controls. Other parameters such as 'Degree of polynomial', 'No. of classes', and 'Kernal function' were assigned as 1, 2, and 'Linear', respectively. The parameters for Naïve Baysian were assigned with normal uniform distribution for MDD patient's class and the healthy control class.

3.4. Classification results

Table 3 shows results of classification performance based on the LR classification model. According to the table, the highest performance (accuracy = 97.6%, sensitivity = 96.66% and specificity = 98.5%) was achieved with alpha asymmetry by utilizing 10 features only. The second highest accuracy was achieved with theta powers features. Among these features, the wavelet coefficients have achieved lowest performance. However, it was high enough to be considered as good. In the table, the last column shows the number of features that have shown highest efficiencies among other features sub-sets.

Table 4 shows results of classification performance based on the NB classification model. According to the table, the highest performance (accuracy = 96.8%, sensitivity = 96.66% and specificity = 97.02%) was achieved with interhemispheric alpha asymmetry by utilizing 5 features only. The second highest accuracy was achieved with beta powers features. Among these features, the theta power has achieved lowest performance.

Table 5 shows results of classification performance based on the NB classification model. According to the table, the highest performance (accuracy = 98.4%, sensitivity = 96.66% and specificity = 100%) was achieved with interhemispheric alpha asymmetry by utilizing 19 features only. The second highest accuracy was achieved with beta powers features. Among these features, the wavelet coefficients have achieved lowest performance. However, the accuracy is high enough to be considered as good classification.

4. Discussion and conclusion

In this paper, we have analyzed QEEG features as discriminants to identify MDD patients among a population of healthy controls. The choice of EEG features and the proposed ML scheme have shown robust and promising results. In addition, the incorporation of ML techniques allows automation of the diagnosis process. EEG is cheaper than fMRI and encourages a low-cost solution that would be feasible for remote applications such as in small clinics and health-care facilities. On the other hand, the existing clinical diagnostic approach involved well-structured questionnaires and interviews conducted from MDD patients. MDD has been considered as heterogeneous due to less known underlying pathophysiology and due to its comorbid nature. Therefore, during a diagnosis process, subjectivity in treatment conditions may not be avoided and hence vulnerable to human errors. Moreover, in case of failure, the conventional methods may become time-consuming and cannot be standardized because of heterogenetic behavior of the MDD. Therefore, objective techniques are required to fill up this gap as EEG data are considered.

The ability to translate previous QEEG-based methods to clinical purposes is lacking due to multiple factors including less efficient methods, small sample sizes that poses a hindrance to the generalization of the methods and methodological differences incorporated by the heterogeneity of sample populations and the disease conditions. In contrary, the proposed method has shown highest efficiencies indicating feasibility for clinical applicability. The advantage of employing ML technique is to incorporate objectivity during diagnosing MDD. According to the results, the EEG features such as power computation from the frequency bands such as delta, theta, alpha and beta bands and EEG alpha interhemispheric asymmetry have shown their promise. The relatively

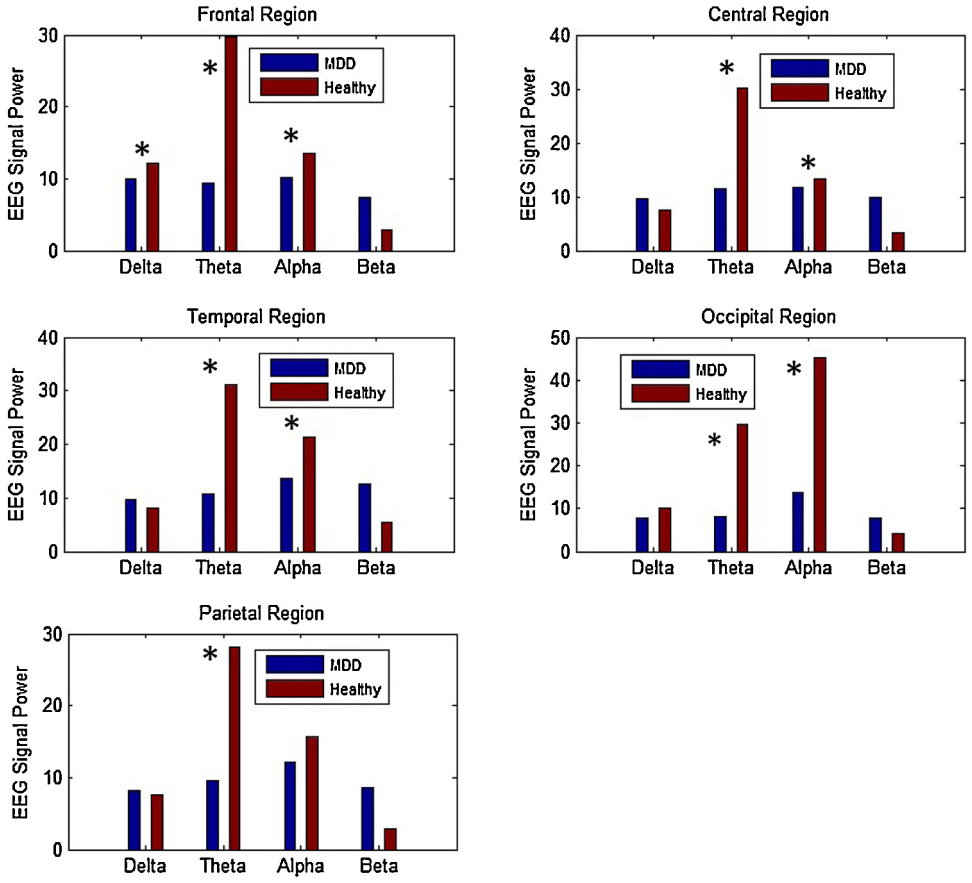

Fig. 2. Comparison between MDD patients and Healthy controls based on EEG Signal Power (Delta, theta, alpha, beta). The significance has been demonstrated with an asterisk, i.e., * ($p < 0.01$).

Table 2
Performance of extracted EEG (Infinity Ref. Data) features based on Logistic Regression (LR) classification while discriminating MDD patients and healthy controls.

Sr.	Classifier	Parameters	Value
1	Logistic Regression Classification	Link Function	Logit
		Distribution	Binomial
		Offset	1
		Constant term	A constant term is added in the model
2	Support Vector Machine (SVM)	C for class 1 (N/2xN1)	0.787
		C for class 2 (N/2xN2)	1.3684
		Degree of polynomial	1
		No. of classes	2
		Kernal function	Linear
3	Naïve Baysian Classification	Distribution	Normal
		Prior	Uniform distribution for all classes

Table 3
Performance of extracted EEG (Infinity Ref. Data) features based on Logistic Regression (LR) classification while discriminating MDD patients and healthy controls.

Sr.	Accuracy	Sensitivity	Specificity	Number of Features Used for Classification (5,10,15,19)
Alpha Power	87.2%	88.3%	86.4%	10
Theta Power	89.7%	91.6%	87.8%	5
Delta Power	84.44%	81.66%	86.6%	19
Beta Power	87.3%	90%	85%	15
Interhemispheric Alpha Asymmetry	**97.6%**	**96.66%**	**98.5%**	**10**
All Frequency Power	86.4%	90%	82.8%	10
Combining Power and Asymmetry Features	**98.33%**	**96.66%**	**100%**	**10**

Bold values indicates $P < 0.01$.

simplistic EEG features allow low cost solution for clinical applications and justify the feasibility of the proposed method as a potential CAD-solution to diagnose MDD.

According to the results presented in Table 1, the depressed individuals tend to exhibit greater relatively right frontal activity when compared with healthy controls. This finding is in line with the evi-

Table 4
Performance of extracted EEG (Infinity Ref. Data) features based on Naïve Bayesian (NB) classification while discriminating MDD patients and healthy controls.

Sr.	Accuracy	Sensitivity	Specificity	Number of Features Used for Classification (5,10,15,19)
Alpha Power	75.5%	80%	72.02%	5
Theta Power	61.2%	25%	94.2%	19
Delta Power	80.3%	66.6%	94%	19
Beta Power	**94.04%**	**86.6%**	**90.5%**	**19**
Interhemispheric Alpha Asymmetry	**96.8%**	**96.6%**	**97.02%**	**5**
All Frequency Power	86.28%	88.33%	83.8%	10
Combining Power and Asymmetry Features	**97.6%**	**96.6%**	**98.5%**	**10**

Bold values indicates $P < 0.01$.

Table 5
Performance of extracted EEG (Infinity Ref. Data) features based on Support Vector Machine (SVM) classification while discriminating MDD patients and healthy controls.

Sr.	Accuracy	Sensitivity	Specificity	Number of Features Used for Classification (5,10,15,19)
Alpha Power	88.2%	94.1%	83.01%	15
Theta Power	90.4%	93.3%	88.46%	15
Delta Power	86%	88.4%	84.1%	19
Beta Power	90.8%	92.9%	89.2%	19
Interhemispheric Alpha Asymmetry	**98.4%**	**96.66%**	**100%**	**19**
All Frequency Power	86.8%	88.3%	85.3%	19
Combining Power and Asymmetry Features	**98.4%**	**97.5%**	**100%**	**10**

Bold values indicates $P < 0.01$.

dence presented in huge set of older studies [51–54]. In addition, the EEG signal power has shown a decreasing trend in the theta and alpha bands in all the regions which proves that MDD has caused a dysfunction which can be observed by computing EEG signal powers in the theta and alpha power in the resting EEG data.

In this study, we have achieved the highest performance (97%) accuracies which are not shown previously, e.g., the study only concluded with 90% accuracy [12]. Knott et al. [33] has reported a classification of 91.3% involving 70 depressed patients and 23 normal subjects' based on linear features. However, it is clear that due to an imbalance between the numbers of samples in 2 groups the classification results may be biased towards the MDD patient's class. Moreover, in their study Lee et al. [13] analyzed EEG data using detrended fluctuation analysis (DFA) for 11 depressed and 11 controls. The study concluded higher DFA values in MDD patients when compared with controls. Due to small sample sizes, the generalization of the result could not be possible. In another study, EEG analysis based on wavelet entropy analysis has achieved 80% accuracy while involving 26 MDD participants only [55]. In comparison to all these studies, our proposed methodology has achieved the highest accuracy based on linear EEG features. This proves our hypothesis that a careful selection of ML techniques can achieve high performances even with linear features as input data. For example, it is evident from the results that not every frequency band is efficient enough for each classification method. In addition, a combination of the frequency bands resulted in lower classification efficiency than the individual frequency bands. This observation is in accordance with the research ideology that useful features may be discovered from EEG sub-bands as compared to EEG full-bands analysis [56]. Moreover, the highest classification performances are provided by SVM by using comparatively number of features as compared with LR and NB. This is because the SVM has more complex structure than the LR and NB classifiers. Hence, the SVM requires more data samples to train appropriately compared with LR and NB classifiers.

To remove the likelihood that the resulting classifier models are concluded due to noise present in the EEG data, we have adopted the following precautions. First, during preprocessing, the artifacts are carefully removed and tested by plotting their histograms plots to check the presence of any kind of outliers and found the data to be suitable for classification purposes. Second, we have selected equal sample sizes in both the groups. In addition to this, the gender distribution is equal among the groups as well. This is eliminating the gender biases from the final results. Third, the logistic regression and Naive Bayesian classifier models are relatively simpler than the support vector machine classifier. This is why the SVM takes more number of features to produce the classification results. The incorporation of classifier with 3 different structures has proved the validity of our data also. Fourthly, the over-fitting may happen; therefore, we have incorporated 100-time permutation test with 10-fold cross-validation to improve the robustness of the underlying models.

The study is confounded with few limitations. First, the MDD patients are asked to be medication free at least 2 weeks before the first EEG recording. However, the effect of mediation cannot be ruled out completely. Second, since the patients are out-patients, variable such as sleep patterns, appetite, and life style could not be controlled. The small sample size poses a constraint for generalization to a wider population.

In conclusion, despite such limitations the classifiers have shown promising discrimination abilities. Moreover, due to high specificities which are more important in terms of clinical application further provide a strong point that the methods could be utilized for clinical applications. Since the study has based on relatively small sample size, therefore, caution must be adopted while translating these findings to clinical utilities.

Conflict-of-interest statement

Regarding all authors of this paper, there is no conflict-of-interest found.

Acknowledgements

This research work was supported by NSTIP strategic technologies programs, grant number (12-INF2582-02), in the Kingdom of Saudi Arabia; National Natural Science Foundation of China (No. 61572076); China Postdoctoral Science Foundation Grant (No. 2015M570940); BIT Fundamental Research Grant (No. 20150442009).

References

[1] J. Volkert, H. Schulz, M. Härter, O. Wlodarczyk, S. Andreas, The prevalence of mental disorders in older people in Western countries– a meta-analysis, Ageing Res. Rev. 12 (2013) 339–353.
[2] W.H.O, The Global Burden of Disease 2004 Update, World Health Organization 2008, Geneva, Switzerland, 2008.
[3] P. Willner, J. Scheel-Krüger, C. Belzung, The neurobiology of depression and antidepressant action, Neurosci. Biobehav. Rev. 37 (2013) 2331–2371.
[4] W. Mumtaz, A.S. Malik, M.A.M. Yasin, L. Xia, Review on EEG and ERP predictive biomarkers for major depressive disorder, Biomed. Signal Process. Control 22 (2015) 85–98.
[5] S. Olbrich, M. Arns, EEG biomarkers in major depressive disorder: discriminative power and prediction of treatment response, Int. Rev. Psychiatry 25 (2013) 604–618.
[6] H. Alhaj, G. Wisniewski, R.H. McAllister-Williams, The use of the EEG in measuring therapeutic drug action: focus on depression and antidepressants, J. Psychopharmacol. (Oxf.) 25 (2011) 1175–1191.
[7] L.L. Zeng, H. Shen, L. Liu, D. Hu, Unsupervised classification of major depression using functional connectivity MRI, Hum. Brain Mapp. 35 (2014) 1630–1641.
[8] B. Magnin, L. Mesrob, S. Kinkingnéhun, M. Pélégrini-Issac, O. Colliot, M. Sarazin, B. Dubois, S. Lehéricy, H. Benali, Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI, Neuroradiology 51 (2009) 73–83.
[9] G. Orrù, W. Pettersson-Yeo, A.F. Marquand, G. Sartori, A. Mechelli, Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review, Neurosci. Biobehav. Rev. 36 (2012) 1140–1152.
[10] U.R. Acharya, V. Sudarshan, H. Adeli, J. Santhosh, J. Koh, A. Adeli, Computer-Aided diagnosis of depression using EEG signals, Eur. Neurol. 73 (2015) 329–336.
[11] U.R. Acharya, V.K. Sudarshan, H. Adeli, J. Santhosh, J.E. Koh, S.D. Puthankatti, A. Adeli, A novel depression diagnosis index using nonlinear features in EEG signals, Eur. Neurol. 74 (2015) 79–83.
[12] B. Hosseinifard, M.H. Moradi, R. Rostami, Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal, Comput. Methods Programs Biomed. 109 (2013) 339–345.
[13] J.-S. Lee, B.-H. Yang, J.-H. Lee, J.-H. Choi, I.-G. Choi, S.-B. Kim, Detrended fluctuation analysis of resting EEG in depressed outpatients and healthy controls, Clin. Neurophysiol. 118 (2007) 2489–2496.
[14] M. Mohammadi, F. Al-Azab, B. Raahemi, G. Richards, N. Jaworska, D. Smith, S. de la Salle, P. Blier, V. Knott, Data mining EEG signals in depression for their diagnostic value, BMC Med. Inform. Decis. Mak. 15 (2015) 1.
[15] S. Klöppel, A. Abdulkadir, C.R. Jack, N. Koutsouleris, J. Mourão-Miranda, P. Vemuri, Diagnostic neuroimaging across diseases, Neuroimage 61 (2012) 457–463.
[16] S.D. Puthankattil, P.K. Joseph, Classification of EEG signals in normal and depression conditions by ANN using RWE and signal entropy, J. Mech. Med. Biol. 12 (2012) 1240019.
[17] T.T. Erguzel, S. Ozekes, O. Tan, S. Gultekin, Feature selection and classification of electroencephalographic signals an artificial neural network and genetic algorithm based approach, Clin. EEG Neurosci. 46 (2015) 321–326.
[18] U.R. Acharya, S.V. Sree, A.P.C. Alvin, R. Yanti, J.S. Suri, Application of non-linear and wavelet based features for the automated identification of epileptic EEG signals, Int. J. Neural Syst. 22 (2012) 1–14.
[19] H. Adeli, S. Ghosh-Dastidar, N. Dadmehr, A wavelet-chaos methodology for analysis of EEGs and EEG subbands to detect seizure and epilepsy, IEEE Trans. Biomed. Eng. 54 (2007) 205–211.
[20] U.R. Acharya, O. Faust, N. Kannathal, T. Chua, S. Laxminarayan, Non-linear analysis of EEG signals at various sleep stages, Comput. Methods Programs Biomed. 80 (2005) 37–45.
[21] E. Olofsen, J. Sleigh, A. Dahan, Permutation entropy of the electroencephalogram: a measure of anaesthetic drug effect, Br. J. Anaesth. 101 (2008) 810–821.
[22] W. Mumtaz, P.L. Vuong, L. Xia, A.S. Malik, R.B.A. Rashid, Automatic diagnosis of alcohol use disorder using EEG features, Knowledge-Based Syst. 105 (2016) 48–59.
[23] R.S. McIntyre, D.S. Cha, J.K. Soczynska, H.O. Woldeyohannes, L.A. Gallaugher, P. Kudlow, M. Alsuwaidan, A. Baskaran, Cognitive deficits and functional outcomes in major depressive disorder: determinants, substrates, and treatment interventions, Depress. Anxiety 30 (2013) 515–527.
[24] J. Jaeger, S. Berns, S. Uzelac, S. Davis-Conway, Neurocognitive deficits and disability in major depressive disorder, Psychiatry Res. 145 (2006) 39–48.
[25] A.A. Fingelkurts, A.A. Fingelkurts, Altered structure of dynamic electroencephalogram oscillatory pattern in major depression, Biol. Psychiatry 77 (2015) 1050–1060.
[26] A.C. Deslandes, H. de Moraes, F.A. Pompeu, P. Ribeiro, M. Cagy, C. Capitão, H. Alves, R.A. Piedade, J. Laks, Electroencephalographic frontal asymmetry and depressive symptoms in the elderly, Biol. Psychol. 79 (2008) 317–322.
[27] J.L. Stewart, J.A. Coan, D.N. Towers, J.J. Allen, Frontal EEG asymmetry during emotional challenge differentiates individuals with and without lifetime major depressive disorder, J. Affect. Disord. 129 (2011) 167–174.
[28] A. Cantisani, T. Koenig, H. Horn, T. Müller, W. Strik, S. Walther, Psychomotor retardation is linked to frontal alpha asymmetry in major depression, J. Affect. Disord. 188 (2015) 167–172.
[29] J.J. Allen, S.J. Reznik, Frontal EEG asymmetry as a promising marker of depression vulnerability: summary and methodological considerations, Curr. Opin. Psychol. 4 (2015) 93–97.
[30] D. Kan, P. Lee, Decrease alpha waves in depression: an electroencephalogram (EEG) study, in: BioSignal Analysis Processing and Systems (ICBAPS), 2015 International Conference on, IEEE, 2015, pp. 156–161.
[31] B. Saletu, P. Anderer, G. Saletu-Zyhlarz, EEG topography and tomography (LORETA) in diagnosis and pharmacotherapy of depression, Clin. EEG Neurosci. 41 (2010) 203–210.
[32] P. Coutin-Churchman, R. Moreno, Intracranial current density (LORETA) differences in QEEG frequency bands between depressed and non-depressed alcoholic patients, Clin. Neurophysiol. 119 (2008) 948–958.
[33] V. Knott, C. Mahoney, S. Kennedy, K. Evans, EEG power frequency, asymmetry and coherence in male depression, Psychiatry Res.: Neuroimaging 106 (2001) 123–140.
[34] J.B. Henriques, R.J. Davidson, Left frontal hypoactivation in depression, J. Abnorm. Psychol. 100 (1991) 535.
[35] A. Kemp, K. Griffiths, K. Felmingham, S. Shankman, W. Drinkenburg, M. Arns, C. Clark, R. Bryant, Disorder specificity despite comorbidity: resting EEG alpha asymmetry in major depressive disorder and post-traumatic stress disorder, Biol. Psychol. 85 (2010) 350–354.
[36] C. Gold, J. FACHNER, J. ERKKILÄ, Validity and reliability of electroencephalographic frontal alpha asymmetry and frontal midline theta as biomarkers for depression, Scand. J. Psychol. 54 (2013) 118–126.
[37] A.P. Association, Diagnostic and Statistical Manual of Mental Disorders (DSM), American psychiatric association, Washington DC, 1994, pp. 143–147.
[38] H.H. JASPER, The ten twenty electrode system of the international federation, Electroencephalogr. Clin. Neurophysiol. 10 (1958) 371–375.
[39] Y. Qin, P. Xu, D. Yao, A comparative study of different references for EEG default mode network: the use of the infinity reference, Clin. Neurophysiol. 121 (2010) 1981–1991.
[40] A.T. Beck, R.A. Steer, G.K. Brown, Manual for the Beck Depression Inventory-II, Psychological Corporation, San Antonio TX, 1996.
[41] W.M.R.W. Mahmud, A. Awang, I. Herman, M.N. Mohamed, Analysis of the psychometric properties of the Malay version of Beck Depression Inventory II (BDI-II) among postpartum women in Kedah, north west of peninsular Malaysia, Malays. J. Med. Sci.: MJMS 11 (2004) 19.
[42] N. Yusoff, W.Y. Low, C.-H. Yip, Psychometric properties of the Malay Version of the hospital anxiety and depression scale: a study of husbands of breast cancer patients in Kuala Lumpur Malaysia, Asian Pac. J. Cancer Prev. 12 (2011) 915–917.
[43] P. Berg, M. Scherg, A multiple source approach to the correction of eye artifacts, Electroencephalogr. Clin. Neurophysiol. 90 (1994) 229–241.
[44] K. Hoechstetter, P. Berg, M. Scherg, BESA research tutorial 4: Distributed source imaging, BESA Res. Tutorial (2010) 1–29.
[45] H. Hinrikus, A. Suhhova, M. Bachmann, K. Aadamsoo, Ü. Võhma, J. Lass, V. Tuulik, Electroencephalographic spectral asymmetry index for detection of depression, Med. Biol. Eng. Comput. 47 (2009) 1291–1299.
[46] A.V. Oppenheim, R.W. Schafer, J.R. Buck, Discrete-time Signal Processing, Prentice-hall Englewood Cliffs, 1989.
[47] H. Liu, H. Motoda, Computational methods of feature selection, CRC Press, 2007.
[48] D.W. Hosmer Jr., S. Lemeshow, Applied logistic regression, John Wiley & Sons, 2004.
[49] V.N. Vapnik, V. Vapnik, Statistical learning theory, Wiley New York, 1998.
[50] T.M. Mitchell, Machine Learning WCB, McGraw-Hill Boston, MA, 1997.
[51] G.E. Bruder, R. Fong, C.E. Tenke, P. Leite, J.P. Towey, J.E. Stewart, P.J. McGrath, F.M. Quitkin, Regional brain asymmetries in major depression with or without an anxiety disorder: a quantitative electroencephalographic study, Biol. Psychiatry 41 (1997) 939–948.
[52] I.R. Bell, G.E. Schwartz, E.E. Hardin, C.M. Baldwin, J.P. Kline, Differential resting quantitative electroencephalographic alpha patterns in women with environmental chemical intolerance, depressives, and normals, Biol. Psychiatry 43 (1998) 376–388.
[53] I.H. Gotlib, EEG alpha asymmetry, depression, and cognitive functioning, Cogn. Emotion 12 (1998) 449–478.
[54] G. Wiedemann, P. Pauli, W. Dengler, W. Lutzenberger, N. Birbaumer, G. Buchkremer, Frontal brain asymmetry as a biological substrate of emotions in patients with panic disorders, Arch. Gen. Psychiatry 56 (1999) 78–84.
[55] Y. Li, Y. Li, S. Tong, Y. Tang, Y. Zhu, More normal EEGs of depression patients during mental arithmetic than rest, in: Noninvasive Functional Source Imaging of the Brain and Heart and the International Conference on Functional Biomedical Imaging, NFSI-ICFBI 2007. Joint Meeting of the 6th International Symposium on, IEEE, 2007, pp. 165–168.
[56] M. Ahmadlou, H. Adeli, A. Adeli, Fractality analysis of frontal brain in major depressive disorder, Int. J. Psychophysiol. 85 (2012) 206–211.