ORIGINAL ARTICLE

Investigation Capability of EEG-Based Non-Linear Features in Depression Detection

Mehdi Dehghani, Vahid Asayesh *, Majid Torabi Nikjeh, Sepideh Akhtari Khosroshahi

NPC Index Research Company, Tehran, Iran

*Corresponding Author: Vahid Asayesh
Received: 22 September 2022 / Accepted: 29 January 2023
Email: vahid.asayesh@gmail.com

Abstract

Purpose: The purpose of this study is to investigate the potential of non-linear electroencephalography-based features in depression detection.

Materials and Methods: First, the Electroencephalography (EEG) signal was recorded from 25 normal and 25 depressed subjects. After preprocessing these signals, non-linear features including the Sum of Logarithmic (SL) and second-order spectral Moment (2M) of the amplitudes of diagonal elements in the bispectrum, the normalized Entropy (En) of bispectrum in beta and gamma frequency bands, Katz Fractal Dimension (KFD), and Lempel-Ziv Complexity (LZC) are extracted from them. Then, the ability of these features in depression detection was investigated using Mann-Whitney statistical test. Also, the classification performance of significant features was evaluated using a support vector machine (SVM) classifier.

Results: The results of the statistical analysis demonstrate that bispectral 2M, SL, and KFD features show significant differences between depressed and healthy groups in the Eyes-Closed (EC) condition. Also, bispectral 2M and SL in the gamma frequency band show significant differences between the two groups in parietal and temporal regions in the EC condition and only in the temporal region in the Eyes-Open (EO) condition. Bispectral En does not show a significant difference in the whole 19 channel, but it shows significant differences in the frontal region and beta frequency band. Between these features, gamma bispectral 2M in the temporal region and EO condition shows the highest classification result with 78.6±7.2% accuracy.

Conclusion: Findings confirm that bispectral 2M in the gamma frequency band and EO condition can classify depression and healthy subjects.

Keywords: Depression Diagnosis; Electroencephalography; Bispectrum; Katz Fractal Dimension; Lempel-Ziv Complexity; Classification.

DOI: https://doi.org/10.18502/fbt.v12i4.19809

1. Introduction

According to the World Health Organization (WHO) report, more than 300 million people all over the world suffer from depression [1] and approximately 800,000 people die due to it every year [2]. Indigence, physical disorders, tragic events in life, problems with alcohol or drug consumption, and recently, the COVID-19 pandemic are considered factors leading to depression [3, 4]. This mental disorder causes loss of interest, lack of motivation, feeling of sadness, hopelessness, and energy changes. In severe cases, some patients will experience paranoia, illusion, and suicide [2]. Considering the fact that depression causes a burden on public health and has some adverse effects on depressive patients, the development of a reliable and efficient method for depression diagnosis is highly important. Nowadays, the most common way of depression detection is based on clinical interviews conducted by a medical professional [5]. The results of this human-intensive method are dependent on the doctor’s experience; therefore, besides subjective questionnaires, the brain activity of the patients can be monitored objectively using imaging techniques such as Electroencephalography (EEG), Computed Tomography (CT), and functional Magnetic Resonance Imaging (fMRI) [6]. Among these modalities, EEG is simple, inexpensive, and widely available that shows human brain activities, mental states, and diseases such as anxiety, psychosis, and depression [7-9].

Machine learning algorithms have achieved remarkable improvement in the learning and diagnosis of neural disorders. Applications of these methods in depression detection using EEG signals have been increased in recent times [10]. Traditionally, EEG power spectrum analysis has been widely used in classifying depressed and normal subjects [11, 12]. Due to the dynamic and nonlinear nature of the EEG signal, it has been proven that non-linear features have more classification accuracy than linear features. One of the first studies using non-linear features of EEG to categorize depressed and healthy subjects was conducted by [13]. They compared the potential of Katz Fractal Dimension (KFD) and Higuchi Fractal Dimension (HFD) features for the classification of 12 depressed and 12 healthy adults using a Probabilistic Neural Network (PNN) after evaluating the ability of these features to discriminate the groups using ANOVA test. Based on the results, the HDF of the beta band is more discriminative than the HFD of the gamma band while the KFD did not show any significant difference between the two groups. A high accuracy of 91.3% is achieved for the classification of depressed and healthy groups based on frontal brain beta sub-bands. The results of this paper should be considered tentative because of limited data. Although this paper achieved high classification accuracy using HFD, studies show that this feature is not robust to noise [14]. Authors in [15] utilized non-linear analysis of EEG to classify healthy and depressed persons. They extracted EEG signal standard sub-bands and used their power as linear features. They also, extracted Detrended Fluctuation Analysis (DFA), HFD, Correlation Dimension (CD), and Lyapunov exponent as non-linear characteristics of the signal. After feature extraction was performed, each feature vector was applied to K-Nearest Neighbors (KNN), Linear Discriminant Analysis (LDA), and linear regression classifiers. The highest classification accuracy of 83.3% is obtained by CD and linear regression classifier. Also, a classification accuracy of 90% is achieved by the combination of non-linear features. The results of [16] indicate that the sampling rate significantly affected the CD estimation. Besides the effects of sampling rate, the estimation can be influenced by several factors such as low-pass and high-pass filters and the length of the signal segment. Moreover, algorithms of CD computation are time-consuming and data-demanding [17]. The conclusion of this study was that non-linear features give better results in the classification term contrary to linear features. In [18], several entropy measures such as Renyi entropy, approximate entropy, and sample entropy were used to automate the detection of depression. In this paper, only Fp1, T3, Fp2, and T4 electrode positions were used. After applying Student’s t-test to evaluate features, several classification algorithms including KNN, PNN, Gaussian Mixture Model (GMM), and Support Vector Machine (SVM) were applied to them. The results showed that the PNN classifier with 99.5% accuracy performed better than the rest of the classifiers in discriminating between normal and depression EEG signals. Most of the entropy measures (except sample entropy) are sensitive to data length and they are less accurate for shorter segments. Also, approximate and sample entropy are affected by their parameters [19]. Bairy *et al.* [20] used sample entropy, CD, fractal dimension, Lyapunov exponent, Hurst exponent, and DFA features extracted from Discrete Cosine Transform (DCT) coefficients as input for SVM and KNN. They also ranked features using a t-value test. In this research, significant features are selected and the consistency of cross-validation technique results in a classification accuracy of 93.8%, but as this system is tested in the modeling environment it may not yield high accuracy in practice. Bachman *et al.* [5] performed a classification task using alpha power variability, relative gamma power, HFD, and Lempel-Ziv Complexity (LZC) as features and logistic regression as a classifier. In this study, maximal testing accuracy using a single measure was 77% for non-linear measures. A combination of HFD and LZC measures provides the accuracy of 85%. This study has limitations due to a small number of subjects that affect the results. Authors in [21] evaluated the effectiveness of HFD and sample entropy in the diagnosis of depression disorder. HFD and sample entropy were used as input for MLP, logistic regression, SVM, decision tree, and random forest classifiers. Average accuracy among classifiers ranged from 90% to 97.5%. Most of these studies had very modest samples, which is affecting the generalizability of the model.

The power spectrum analysis performs power distribution calculation and it assumes that the signal arises from a linear process thus ignoring any possible interaction between components (phase coupling) [22]. However, the brain signal is part of the central nervous system with many nonlinear and non-Gaussian sources, so it is highly likely to have a phase coupling between signals. Therefore, bispectrum analysis is superior to the power spectrum analysis, because, in its mathematical formula, there is a correlation calculation between the frequency components and the phase coupling components of the EEG signals could be revealed [23]. According to the authors' knowledge, the capability of bispectrum analysis was not investigated in depression detection. Therefore, in this study, the performance of bispectrum analysis was evaluated in depression diagnosis and compared with the common frequency-based (linear) features such as absolute and relative powers in alpha, beta, gamma bands, alpha/beta, alpha/gamma, beta/gamma power ratios, and other most used non-linear features such as KFD, and LZC, in statistical and classification perspectives.

The rest of the paper can be summarized as follows: In section 2, EEG signal recording, preprocessing, feature extraction methods, statistical analysis, and machine learning algorithm are explained. In section 3, the results of statistical analysis and classification results are described and the capability of proposed features is compared with other popular non-linear and frequency-based features. In section 4, discussion and conclusion are presented.

2. Materials and Methods

2.1. Participants and Data Collection

In this study, the EEG signals were recorded from 25 normal and 25 depressed subjects that were referred to the Asayesh rehabilitation clinic, Tabriz, Iran. All of the subjects were medication-free and the diagnosis for them was made by an expert psychiatrist based on DSM-V criteria. The DSM-V outlines the following criterion to make a diagnosis of depression: depressed mood, markedly diminished interest in almost all activities, significant weight loss, slowing down of thought and a reduction of physical movement, fatigue or loss of energy, feelings of worthlessness or excessive or inappropriate guilt nearly every day, the diminished ability of thinking or concentration, recurrent thoughts of death or suicidal ideation, and insomnia or hypersomnia. The individual must be experiencing five or more symptoms during the same 2-week period and at least one of them should be either depressed mood or loss of interest. Also, symptoms must cause significant distress or impairment and the individual should not have manic or hypomanic behavior [24]. Individuals who had comorbidities with depression such as anxiety were not included in the experiment, and all depressed patients have pure depression. The DSM-V identifies depression in all or no fashion. In contrast, depression severity is a continuous variable. Therefore, it is commonly assessed with scales such as Beck Depression Inventory (BDI). The BDI contains 21 self-reported items which individuals complete using multiple-choice response format [25]. In this study, the severity of depression has been investigated using the BDI

Table 1. Demographic information of participants

	Gender			Age		
	Male	Female	Total	Male	Female	Total
Depression	13 (52%)	12 (48%)	25	35.84±10.93	38.25±14.40	37±12.50
Healthy	13 (52%)	12 (48%)	25	32.30±6.88	33.16±8.03	32.72±7.31

questionnaire, and to make the research comprehensive, no specific range of depression has been considered, so the depressed group is a combination of all severities. Table 1 shows the demographic information of the subjects. According to this table, there are no significant differences in gender and age between the normal and depressed groups. All subjects expressed their consent to participate in this research. The EEG signal was acquired from subjects for 10 minutes (5 minutes while their Eye was Closed (EC) and 5 minutes while their Eyes were Open (EO)) using Mitsar 19-channel system according to the 10-20 system. Then, recorded signals pass through the low-pass and high-pass filters with 50 and 0.1 Hz cut-off frequencies. Also, a notch filter with cut-off frequencies of 45 and 55 Hz was used to remove the power line noise. The Infomax Independent Component Analysis (ICA) decomposition was used to remove eye movements such as blinking or saccades [26]. Recordings were further cleaned with an automated z-score-based method, using the FASTER plugin [27].

2.2. Feature Extraction

After preprocessing EEG signals non-linear features including bispectral SL, bispectral 2M (as new features in depression detection), bispectrum normalized entropy, KFD, LZC (as usual features in depression detection), and mentioned frequency-based features are extracted from them. These features and their extraction methods are described below.

2.2.1. Lempel-Ziv Complexity

One of the more explored properties of EEG is complexity which has been related to desynchrony, randomness, or irregularity [28]. LZC is an algorithm for computing complexity that has been proposed by Lempel and Ziv [29]. According to Equation 1, for computing this feature, first, the EEG signal $x(n)$is converted to a binary sequence $s(n)$by comparing it with a threshold that is mostly the median.

$$s(i) = \begin{cases} 0, & x(i) < median, \\ 1, & \text{otherwise.} \end{cases} \tag{1}$$

Then, by scanning the resultant EEG sequences, different patterns in the signal are counted by$c(n)$. Based on Equation 2, LZC is obtained by normalizing $c(n)$ to the upper limit of the complexity measure $b(n)$.

$$\begin{aligned} b(n) &= \lim_{n\to\infty} c(n) = \frac{n}{\log_2(n)}, \\ LZC &= \frac{c(n)}{b(n)}. \end{aligned} \tag{2}$$

2.2.2. Katz Fractal Dimension

The fractal dimension is a novel non-linear feature that shows the self-similarity of EEG signals based on the number of repetitions of a pattern in it [30]. One of the known algorithms for computing fractal dimension was proposed by Katz. This algorithm computes the fractal dimension of the signal x as Equation 3 [31]

$$\begin{aligned} KFD &= \frac{log\left(\frac{l}{a}\right)}{log\left(\frac{d}{a}\right)}; \\ l &= \sum_{i=2}^{N} \text{distance}\big(x(i), x(i-1)\big);\ a = \frac{l}{N-1} \end{aligned} \tag{3}$$

Where a is the average distance of consecutive points, N is the length of the time series, l is the sum of distances between all successive points, and d is the maximum distance between $x(1)$ and other points of the time series.

2.2.3. Bispectrum Features

The bispectrum is defined as the 2D Fourier transform of the skewness (third-order momentum) of the signal that shows the 2D mapping of the interaction level between all frequency pairs in the specific band of the signal [32]. The bispectrum quantifies the oscillatory relationship between basic functions of f_1,f_2 and their harmonic components according to Equation 4:

$$Bis(f_1, f_2) = \lim_{T\to\infty}\left(\frac{1}{T}\right)E[X(f_1+f_2)X^*(f_1)X^*(f_2)] \qquad (4)$$

Where $X(f)$ is the Fourier transform of $x(t)$, $(*)$ denotes the complex conjugate, and E is the expectation operator.

In order to characterize time series, quantitative features must be extracted from the bispectrum. Here, three measures are derived from the bispectrum of the EEG signal of every channel in the beta and gamma frequency bands. Two of these features are related to moments of the bispectrum that include the sum of logarithmic amplitudes of diagonal elements in the bispectrum (bispectral SL) and the second-order spectral moment of the amplitudes of diagonal elements in the bispectrum (bispectral 2M). The third measure was based on the entropy of the bispectrum (bispectral En) [33]. Due to the asymmetry of the bispectrum in the frequency range of calculation, the non-redundant region of it can characterize the whole bispectrum. Therefore, measures obtained from the bispectrum are computed in this region. The formula of measures which are extracted in this paper are as follows:

Bispectral SL

$$BispectralSL = \sum_{f_k \in \Omega} log(|B(f_k, f_k)|)\,. \qquad (5)$$

Where $B(f_k, f_k)$ is the bispectrum of the signal in diagonal elements in the non-redundant region Ω.

Bispectral 2M

$$Bispectral2M = \sum_{j=1}^{N}\left(j - \sum_{k=1}^{N} k\, log(|B(f_k, f_k)|)\right)^2 log\left(\left|B\left(f_j, f_j\right)\right|\right) \qquad (6)$$

Where N is the number of diagonal elements in the region Ω.

Normalized Bispectral En

$$BispectralEn = -\sum_n p_n\, log(p_n)\,; \qquad (7)$$

$$p_n = \frac{|B(f_1, f_2)|}{\sum_{\Omega}|B(f_1, f_2)|}$$

2.3. Statistical Analysis

In this study, to assess non-linear features' capability in distinguishing between depressed and healthy EEG recordings, a statistical analysis was performed to measure the level of statistically significant differences between various groups. Since some of the data did not have a normal distribution, we applied the Mann-Whitney test ($p \leq 0.01$) to compare every measure between the two groups of healthy and depressed. In addition to 19 EEG channels, to investigate brain regions' influences on depression diagnosis, the statistical analysis was performed at the brain regions level, too. For this aim, 19 channels of EEG signals are divided into 5 brain regions according to Table 2 [34] and the Mann-Whitney test is applied to these electrodes for investigating the related brain region's role in depression detection.

Table 2. Channel clustering for each brain region

Frontal	F7, F3, Fz, F4, F8, Fp1, Fp2
Central	C3, C4, Cz
Parietal	P3, Pz, P4
Temporal	T3, T4, T5, T6
Occipital	O1, O2

2.4. Machine Learning Approach

The appropriate selection of classification method is very crucial for the automated diagnosis of patients having neurological diseases [35]. Among the various classification algorithms, SVM has emerged as a powerful classification technique that solves a convex optimization problem which leads to a globally optimal solution [36]. In the simplest type of SVM, the training data set can be classified linearly with at least one hyper-plane. Linear classifiers are inefficient for real problems that have a non-linear structure. One of the abilities of SVM is to be converted to a non-linear learner, which is done by mapping the features to a higher-dimensional space [37]. In this study, the ability of SVM with quadratic, cubic, and Gaussian kernels was examined for automatic depression diagnosis using each feature set, and the results of the kernel which achieved the best performance were reported.

3. Results

3.1. Statistical Analysis

The results of the comparison of non-linear features extracted from the EEG signals of depressed and healthy subjects in EO and EC conditions by the Mann-Whitney test have been illustrated in Table 3. According to this table, in the EC condition, bispectral SL and bispectral 2M in beta and gamma frequency bands and KFD are statistically different between the two groups. None of the features in the EO condition had significant differences between the two groups. Figure 1 shows the mean, p-value distribution, and effect size of significant features on individuals of each group. According to this figure, the bispectral features in beta and gamma frequency bands are significantly higher in the depressed group. Also, depressed subjects have significantly greater KFD values than healthy subjects. P-values distribution in Figure 1, panel (b), shows that mostly temporal and parietal regions show significant differences between the two groups. For investigation in detail, the results of region-based statistical analysis are illustrated in Figure 2 and Table 4.

Based on this table, in EC condition, bispectral SL and bispectral 2M in the gamma frequency band are significantly different between the two groups in temporal and parietal brain regions. Also, bispectral En in the beta frequency band shows significant differences. In the EO condition, these features show significant differences between depressed and normal groups only in the temporal region. Figure 2 shows the mean and p-values of significant region-based features in each group. These features have higher values in depressed patients than normal group.

Table 3. Results of the statistical analysis between two groups of depressed and healthy by Mann-Whitney test

	Features	P-value	
		EC	EO
1	Bispectral SL-Beta	**0.01***	0.14
2	Bispectral SL-Gamma	**2.24e-4***	0.42
3	Bispectral 2M-Beta	**0.01***	0.14
4	Bispectral 2M-Gamma	**2.47e-4***	0.43
5	Bispectral En-Beta	0.11	0.15
6	Bispectral En-Gamma	0.96	0.46
7	KFD	**0.01***	0.03
8	LZC	0.16	0.05

* P-value ≤ 0.01

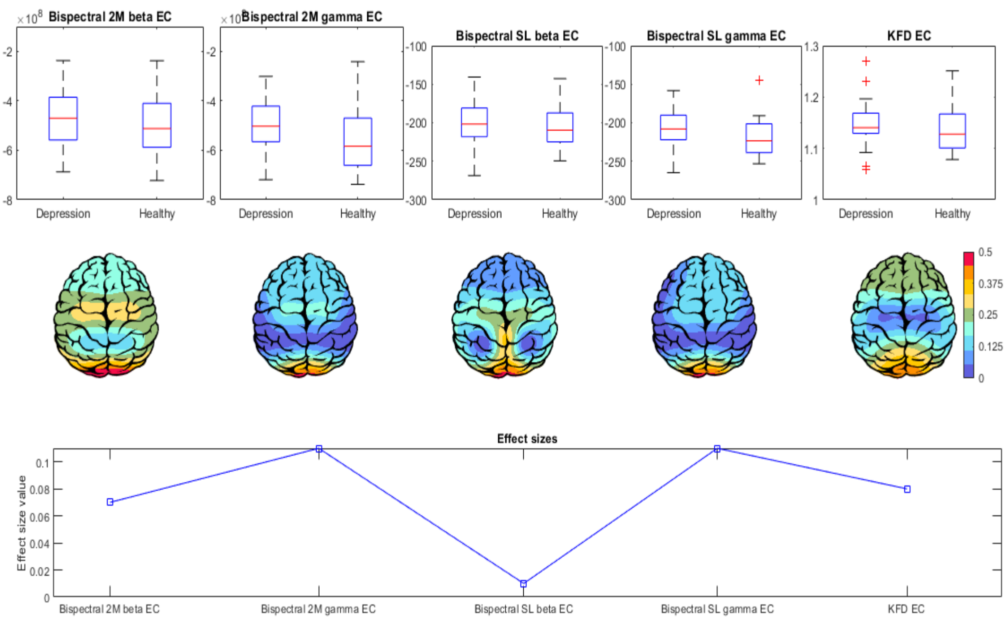

Figure 1. (a) Mean, **(b)** p-value distribution, and **(c)** effect size for bispectral 2M and SL in beta and gamma frequency bands and KFD features in EC condition

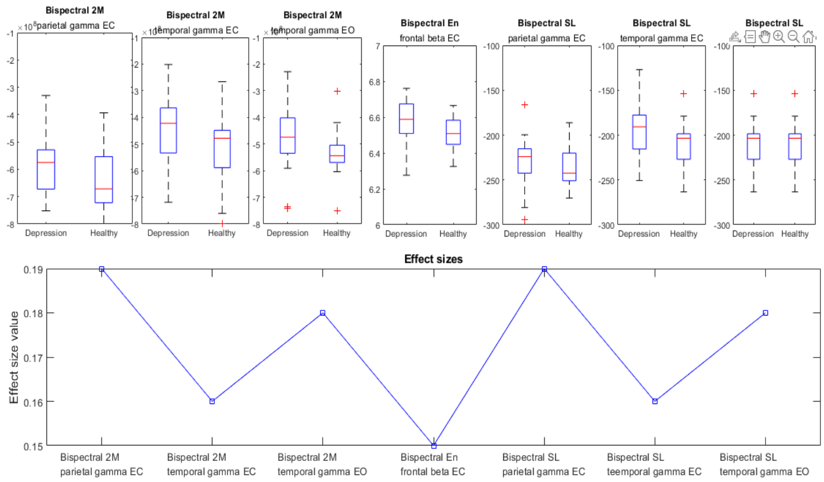

Figure 2. Mean and effect size for bispectral 2M parietal gamma EC, bispectral 2M temporal gamma EC, bispectral 2M temporal gamma EO, bispectral En frontal beta, bispectral SL parietal gamma EC, bispectral SL temporal gamma EC, bispectral SL temporal gamma EO

Table 4. Results of the region-based statistical analysis between two groups of depressed and healthy by Mann-Whitney test

	Features	P-value									
		EC					EO				
		F	C	P	T	O	F	C	P	T	O
1	Bispectral SL-Beta	0.20	0.32	0.13	0.26	0.42	0.69	0.21	0.39	0.22	0.73
2	Bispectral SL-Gamma	0.07	0.09	**0.01***	**0.01***	0.23	0.95	0.93	0.89	**0.008***	0.99
3	Bispectral 2M-Beta	0.20	0.31	0.13	0.26	0.42	0.69	0.21	0.39	0.22	0.73
4	Bispectral 2M-Gamma	0.07	0.10	**0.01***	**0.01***	0.23	0.93	0.95	0.89	**0.008***	0.99
5	Bispectral En-Beta	**0.005***	0.20	0.55	0.70	0.51	0.26	0.40	0.17	0.96	0.10
6	Bispectral En-Gamma	0.55	0.15	0.55	0.13	0.12	0.62	0.85	0.11	0.81	0.47
7	KFD	0.25	0.05	0.22	0.16	0.39	0.33	0.07	0.55	0.04	0.51
8	LZC	0.76	0.32	0.77	0.04	0.75	0.16	0.66	0.98	0.02	0.95

- P-value $\leq$ 0.01, F: frontal brain region, C: central brain region, P: parietal brain region, T: temporal brain region, O: occipital brain region.

3.2. Classification

In this section, the classification ability of statistically significant features was investigated. As mentioned in section 2.4 the ability of SVM with quadratic, cubic, and Gaussian kernels was examined in each feature set and the results of the kernel with the highest performance were reported. The criteria applied for evaluating the classification method include accuracy, sensitivity, and specificity. Based on Equation 8, the classification accuracy is the augment of correctly predicted data points out of all the data points.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

Here, TP, TN, FP, and FN are true positive, true negative, false positive, and false negative values, respectively. According to Equation 9, sensitivity (TPR) is the measure that considers a classifier's ability to predict the true positives of each available class.

$$sensitivity = \frac{TP}{TP + FN}. \tag{9}$$

Based on Equation 10, specificity (TNR) is the parameter that measures a learner's ability to predict the true negatives of each available class.

$$specificity = \frac{TN}{TN + FP} \quad (10)$$

The classifier was trained using a 5-fold cross-validation algorithm on 70% of the dataset and tested on the remaining data. Then, an average and standard deviation of 5 results are presented in Table 5. The results show that temporal bispectral 2M in the gamma frequency band in EO condition achieved the best performance with 78.6±7.2% accuracy in the test stage using the Gaussian kernel.

4. Discussion and Conclusion

In this study, the ability of bispectrum features was examined in depression diagnosis. The results show that in the EO condition, bispectral SL and bispectral 2M in the gamma frequency band and parietal and temporal brain regions show significant differences between MDD and healthy groups. Also in the EO condition, these features show significant differences only in the temporal region. In addition to these features, bispectral En in the beta frequency band and frontal region show significant differences between the two groups just in the EC condition. In some studies, it is demonstrated that beta activity is related to depression levels and rTMS responsibility [33]. Findings reported by [38, 39] support that beta activity is increased in depressed patients. Also, the frontal lobe is one of the highly attended brain regions in depression research. Authors in [40, 41] found that the forehead beta band achieved the best results in depression detection. The gamma frequency band is related to sensory systems and mood swings [42]. There is some evidence indicating that in monopolar-depressed subjects, gamma activity is increased [43]. Furthermore, depression can be related to the dysfunction of tempoparietal mechanisms [44]. These results are in good agreement with the present study findings which demonstrate that bispectral features in beta and gamma frequency bands and parietal and temporal brain regions are good biomarkers for MDD detection from statistical and classification points of view.

Previous studies have shown the ability of the LZC and KFD features in depression detection [6, 45]. In our dataset, these features do not show any significant differences in brain regions analysis. In line with our findings, Bachmann *et al*. [5] do not find LZC feature statistically significant in depression detection. But in

Table 5. Classification results for significant features using SVM classifier

Feature	Kernel	Accuracy(%)		Sensitivity(%)		Specificity(%)	
		Train	Test	Train	Test	Train	Test
BispectralSL-beta-19 channel-EC	Cubic	50.6±5.7	55.5±5.6	51.1±7.5	56.2±12.5	50.1±7.0	51.2±10.9
BispectralSL-gamma-19 channel-EC	Gaussian	63.2±10.9	58.5±7.8	61.3±10.2	58.5±8.0	63.4±12.1	59.4±7.8
BispectralSL-gamma-Parietal-EC	Gaussian	66.1±4.9	52.8±3.9	65.2±4.4	52.8±3.9	67.2±5.8	52.8±3.9
BispectralSL-gamma-Temporal-EC	Linear	45.6±4.6	57.1±7.2	49.3±5.2	61.4±7.6	49.5±4.0	57.3±5.4
BispectralSL-gamma-Temporal-EO	Gaussian	62.8±5.9	71.4±8.7	63.7±7.6	82.5±7.6	61.2±5.3	72.6±7.1
Bispectral2M-beta-19 channel-EC	Quadratic	43.2±9.9	44.0±8.2	44.9±6.8	44.3±11.7	43.0±9.5	43.7±7.2
Bispectral2M-gamma-19 channel-EC	Gaussian	66.1±4.9	63.4±3.3	70.6±8.6	61.4±3.9	63.4±4.5	60.1±3.3
Bispectral2M-gamma-Parietal-EC	Gaussian	60±5.1	58.6±3.9	64.4±7.6	61.4±9.5	58.2±6.1	63.4±12.1
Bispectral2M-gamma-Temporal-EC	Gaussian	54.4±6.9	54.1±6.5	56±6.5	52.8±8.1	51.2±11.4	53.5±8.1
Bispectral2M-gamma-Temporal-EO	**Gaussian**	**80.0±4.8**	**78.6±7.2**	**71.4±5.3**	**78.6±8.4**	**67.7±6.3**	**71.2±9.3**
BispectralEn-beta-Frontal-EC	Linear	52.2±9.5	58.5±13.7	51.8±9.5	56.4±11.5	52.5±9.6	68.0±9.4
KFD-19 channel-EC	Gaussian	53.3±3.6	52.8±8.1	53.6±3.8	49.9±4.4	57.8±10.3	53.9±9.4

Table 6. Results of the region-based statistical analysis between two groups of depressed and healthy by Mann-Whitney test and common frequency-based features

Features	P-value											
	EC						EO					
	Total	F	C	P	T	O	Total	F	C	P	T	O
Absolute power-Alpha	0.33	0.35	0.66	0.74	0.31	0.74	0.05	0.02	0.27	0.50	0.34	0.54
Absolute power-Beta	**0.002***	0.08	0.16	0.12	0.07	0.48	0.09	0.95	0.17	0.18	0.19	0.96
Absolute power-Gamma	**6.05e-5***	0.04	0.06	0.02	0.03	0.06	**0.005***	0.66	0.06	0.15	0.02	0.21
Relative power-Alpha	**0.01***	0.13	0.43	0.21	0.06	0.82	0.05	0.15	0.50	0.47	0.27	0.41
Relative power-Beta	**0.002***	0.23	0.14	0.25	**0.002***	0.88	0.02	0.20	0.24	0.69	0.05	0.82
Relative power-Gamma	**0.002***	0.15	0.20	0.23	**0.01***	0.51	**0.001***	0.05	0.09	0.32	0.05	0.29
Alpha/Beta power ratio	**3.40e-4***	0.06	0.12	0.19	**0.003***	0.93	**0.001***	0.03	0.21	0.27	0.07	0.56
Alpha/gamma power ratio	**0.002***	0.12	0.19	0.20	0.02	0.62	**0.003***	0.05	0.16	0.51	0.08	0.44
Beta/gamma power ratio	0.20	0.41	0.99	0.99	0.26	0.35	0.10	0.22	0.51	0.78	0.16	0.83

* P-value ≤ 0.01, F: frontal brain region, C: central brain region, P: parietal brain region, T: temporal brain region, O: occipital brain region.

19-channel EEG analysis, the KFD feature shows a significant difference between MDD and control groups in the EC condition. For comparing linear and non-linear features ability in depression detection, several common frequency-based features were extracted from EEG signals. Based on Table 6, in our dataset, absolute power in beta and gamma bands, relative power in alpha, beta, and gamma bands, alpha/beta and alpha/gamma power ratios in EC condition and gamma absolute and relative powers, alpha/beta and alpha/gamma power ratios in EO condition show a significant difference between the two groups. Also, in region-based analysis, beta and gamma relative power, and alpha/beta power ratio in the EO state show a significant difference in the temporal brain region. Classification results using significant linear features are illustrated in Table 7. Based on the results, gamma absolute and relative powers in the EO state achieved the best classification performance with 60.0±6.4% accuracy. Results show that non-linear features are suitable for depression detection in our dataset. It seems that the reason for low accuracies compared to previous works is the complexity and heterogeneity of the depression dataset that includes a wide range of severities.

In this study, the merit of EEG non-linear features in the automatic diagnosis of MDD has been investigated. For this aim, SVM with different kernels was used as a classifier. Based on the results, bispectral 2M in the gamma frequency band and temporal region in EO mode achieves the best classification performance with 78.6±7.2% accuracy using the Gaussian kernel of SVM. According to Figure 2, between MDD and healthy groups, the distribution of this feature has more different ranges than other features. Therefore, it is reasonable that this feature shows higher performance than other features in the classification. Also, due to the non-linear nature of this feature, the Gaussian kernel is suitable for this classification task. Because the focus of this paper is on introducing new nonlinear features to diagnose normal and depressed subjects, the classification result has been examined with SVM only. Different classifiers can be used in this field and their ability can be tested. Especially if the amount of data is large, the use of deep learning methods can remarkably increase the results of classification.

Table 7. Classification results for significant frequency-based features using an SVM classifier

Feature	Kernel	Accuracy(%)		Sensitivity(%)		Specificity(%)	
		Train	Test	Train	Test	Train	Test
Absolute power-Beta-All-EC	Gaussian	57.8±94	51.4±7.8	54.2±7.4	50.7±4.7	62.8±12.6	56.7±5.8
Absolute power-Gamma-All-EC	Linear	55.9±3.6	58.6±13.7	57.9±4.3	64.3±18.3	54.8±3.2	56.2±11.4
Relative power-Alpha-All-EC	Linear	55.3±9.1	48.5±11.7	55.6±9.7	50.5±12.8	55.0±8.7	55.6±27.2
Relative power-Beta-All-EC	Gaussian	53.6±6.4	54.3±6.4	52.2±7.0	54.8±7.3	54.1±7.9	53.9±5.7
Relative power-Gamma-All-EC	Gaussian	51.2±8.9	57.1±7.2	50.3±12.5	60.3±10.8	51.3±7.4	55.6±5.7
Alpha/beta power ratio-All-EC	Gaussian	55.3±12.2	55.7±13.7	54.8±10.4	55.7±12.8	55.9±14.9	55.7±15.0
Alpha\gamma power ratio-All-EC	Cubic	48.8±6.4	47.1±14.8	48.9±5.8	55.7±12.8	48.7±7.3	55.7±15.0
Relative power-Beta-Temporal-EC	Linear	61.7±5.9	48.6±10.6	64.9±7.9	51.3±16.7	58.6±4.6	50.6±9.2
Relative power-Gamma-Temporal-EC	Linear	45.3±10.7	58.6±16.3	46.1±14.7	67.5±24.3	45.1±9.1	56.8±18.4
Alpha\beta power ratio-Temporal-EC	Gaussian	63.0±10.8	47.1±10.8	64.6±6.2	52.8±9.7	62.1±5.3	46.2±12.1
Absolute power-Gamma-All-EO	Quadratic	55.3±5.3	60.0±6.4	58.0±5.5	71.5±17.6	56.5±4.1	58.7±3.6
Relative power-Gamma-All-EO	Gaussian	55.3±7.3	60.0±6.4	58.4±10.5	63.4±9.4	53.8±5.9	59.8±4.3
Alpha\beta power ratio-All-EO	Gaussian	53.5±4.3	50.0±7.1	53.5±3.8	49.7±7.6	54.7±7.4	49.7±7.6
Alpha\gamma power ratio-All-EO	Gaussian	50.0±6.9	57.1±8.7	48.9±5.3	55.7±6.9	49.9±12.0	60.7±12.7

References

1- Atefeh Safayari and Hamidreza Bolhasani, "Depression diagnosis by deep learning using EEG signals: A systematic review." *Medicine in Novel Technology and Devices,* Vol. 12, p. 100102, (2021).

2- Hanshu Cai *et al.*, "A pervasive approach to EEG-based depression detection." *Complexity,* Vol. 2018(2018).

3- WHO Depression, "Other common mental disorders: global health estimates." *Geneva: World Health Organization,* Vol. 24 (No. 1), (2017).

4- Mario Gennaro Mazza *et al.*, "Anxiety and depression in COVID-19 survivors: Role of inflammatory and clinical predictors." *Brain, behavior, and immunity,* Vol. 89pp. 594-600, (2020).

5- Maie Bachmann *et al.*, "Methods for classifying depression in single channel EEG using linear and nonlinear signal analysis." *Computer methods and programs in biomedicine,* Vol. 155pp. 11-17, (2018).

6- Egils Avots, Klāvs Jermakovs, Maie Bachmann, Laura Päeske, Cagri Ozcinar, and Gholamreza Anbarjafari, "Ensemble approach for detection of depression using EEG features." *Entropy,* Vol. 24 (No. 2), p. 211, (2022).

7- Dagmar Kr Hannesdóttir, Jacquelyn Doxie, Martha Ann Bell, Thomas H Ollendick, and Christy D Wolfe, "A longitudinal study of emotion regulation and anxiety in middle childhood: Associations with frontal EEG asymmetry in early childhood." *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology,* Vol. 52 (No. 2), pp. 197-204, (2010).

8- Scott R Sponheim, Brett A Clementz, William G Iacono, and Morton Beiser, "Clinical and biological concomitants of resting state EEG power abnormalities in schizophrenia." *Biological psychiatry,* Vol. 48 (No. 11), pp. 1088-97, (2000).

9- Ryan Thibodeau, Randall S Jorgensen, and Sangmoon Kim, "Depression, anxiety, and resting frontal EEG asymmetry: a meta-analytic review." *Journal of abnormal psychology,* Vol. 115 (No. 4), p. 715, (2006).

10- Betul Ay *et al.*, "Automated depression detection using deep representation and sequence learning with EEG signals." *Journal of medical systems,* Vol. 43 (No. 7), p. 205, (2019).

11- Subha D Puthankattil and Paul K Joseph, "Classification of EEG signals in normal and depression conditions by ANN using RWE and signal entropy." *Journal of Mechanics in Medicine and biology,* Vol. 12 (No. 04), p. 1240019, (2012).

12- Wajid Mumtaz, Likun Xia, Syed Saad Azhar Ali, Mohd Azhar Mohd Yasin, Muhammad Hussain, and Aamir Saeed Malik, "Electroencephalogram (EEG)-based computer-aided technique to diagnose major depressive disorder (MDD)." *Biomedical Signal Processing and Control,* Vol. 31pp. 108-15, (2017).

13- Mehran Ahmadlou, Hojjat Adeli, and Amir Adeli, "Fractality analysis of frontal brain in major depressive disorder." *International Journal of Psychophysiology,* Vol. 85 (No. 2), pp. 206-11, (2012).

14- Rosana Esteller, George Vachtsevanos, Javier Echauz, and Brian Litt, "A comparison of waveform fractal dimension algorithms." *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications,* Vol. 48 (No. 2), pp. 177-83, (2001).

15- Behshad Hosseinifard, Mohammad Hassan Moradi, and Reza Rostami, "Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal." *Computer Methods and Programs in Biomedicine,* Vol. 109 (No. 3), pp. 339-45, (2013).

16- Hongkui Jing and Morikuni Takigawa, "Low sampling rate induces high correlation dimension on electroencephalograms from healthy subjects." *Psychiatry and clinical neurosciences,* Vol. 54 (No. 4), pp. 407-12, (2000).

17- Anna Krakovská and Svorad Štolc Jr, "Spectral decay vs. correlation dimension of EEG." *Neurocomputing,* Vol. 71 (No. 13-15), pp. 2978-85, (2008).

18- Oliver Faust, Peng Chuan Alvin Ang, Subha D Puthankattil, and Paul K Joseph, "Depression diagnosis support system based on EEG signal entropies." *Journal of mechanics in medicine and biology,* Vol. 14 (No. 03), p. 1450035, (2014).

19- Steven M Pincus and Ary L Goldberger, "Physiological time-series analysis: what does regularity quantify?" *American Journal of Physiology-Heart and Circulatory Physiology,* Vol. 266 (No. 4), pp. H1643-H56, (1994).

20- G Muralidhar Bairy, Shreya Bhat, Lim Wei Jie Eugene, UC Niranjan, Subha D Puthankattil, and Paul K Joseph, "Automated classification of depression electroencephalographic signals using discrete cosine transform and nonlinear dynamics." *Journal of Medical Imaging and Health Informatics,* Vol. 5 (No. 3), pp. 635-40, (2015).

21- Milena Cukic, David Pokrajac, Miodrag Stokic, Vlada Radivojevic, and Milos Ljubisavljevic, "EEG machine learning with Higuchi fractal dimension and Sample Entropy as features for successful detection of depression." *arXiv preprint arXiv:1803.05985,* (2018).

22- Prima Dewi Purnamasari, Anak Agung Putri Ratna, and Benyamin Kusumoputro, "Development of filtered bispectrum for EEG signal feature extraction in automatic emotion recognition using artificial neural networks." *Algorithms,* Vol. 10 (No. 2), p. 63, (2017).

23- Maryam Mohebbi and Hassan Ghassemian, "Prediction of paroxysmal atrial fibrillation based on non-linear analysis and spectrum and bispectrum features of the heart rate variability signal." *Computer methods and programs in biomedicine,* Vol. 105 (No. 1), pp. 40-49, (2012).

24- J Lyness, "Unipolar depression in adults: Clinical features." *UpToDate.(internet),* pp. 1-15, (2019).

25- Julio C Tolentino and Sergio L Schmidt, "DSM-5 criteria and depression severity: implications for clinical practice." *Frontiers in psychiatry,* Vol. 9p. 450, (2018).

26- Tzyy-Ping Jung *et al.*, "Removing electroencephalographic artifacts by blind source separation." *Psychophysiology,* Vol. 37 (No. 2), pp. 163-78, (2000).

27- Hugh Nolan, Robert Whelan, and Richard B Reilly, "FASTER: fully automated statistical thresholding for EEG artifact rejection." *Journal of neuroscience methods,* Vol. 192 (No. 1), pp. 152-62, (2010).

28- Antonio J Ibáñez-Molina, Sergio Iglesias-Parro, María F Soriano, and José I Aznarte, "Multiscale Lempel–Ziv complexity for EEG measures." *Clinical Neurophysiology,* Vol. 126 (No. 3), pp. 541-48, (2015).

29- Abraham Lempel and Jacob Ziv, "On the complexity of finite sequences." *IEEE transactions on information theory,* Vol. 22 (No. 1), pp. 75-81, (2003).

30- Jisu Elsa Jacob, Gopakumar Kuttappan Nair, Ajith Cherian, and Thomas Iype, "Application of fractal dimension for EEG based diagnosis of encephalopathy." *Analog Integrated Circuits and Signal Processing,* Vol. 100 (No. 2), pp. 429-36, (2019).

31- Sarshar Dorosti and Reza Khosrowabadi, "Fractal dimension of EEG signal senses complexity of fractal animations." *bioRxiv,* (2021).

32- Elie Bou Assi, Laura Gagliano, Sandy Rihana, Dang K Nguyen, and Mohamad Sawan, "Bispectrum features and multilayer perceptron classifier to enhance seizure prediction." *Scientific reports,* Vol. 8 (No. 1), pp. 1-8, (2018).

33- Fatemeh Hasanzadeh, Maryam Mohebbi, and Reza Rostami, "Prediction of rTMS treatment response in major depressive disorder using machine learning techniques and nonlinear features of EEG signal." *Journal of affective disorders,* Vol. 256pp. 132-42, (2019).

34- Ahmad Shalbaf, Sara Bagherzadeh, and Arash Maghsoudi, "Transfer learning with deep convolutional neural network for automated detection of schizophrenia from EEG signals." *Physical and Engineering Sciences in Medicine,* Vol. 43 (No. 4), pp. 1229-39, (2020).

35- Bharat Richhariya and Muhammad Tanveer, "EEG signal classification using universum support vector machine." *Expert Systems with Applications,* Vol. 106pp. 169-82, (2018).

36- Vahid Asayesh, Mahdi Dehghani, Majid Torabi, Sepideh Akhtari, and Shahriar Gharibzadeh, "Predicting Mini-Mental State Examination Scores Using Electroencephalography Signal Features." *Frontiers in Biomedical Technologies,* Vol. 9 (No. 4), pp. 274-82, (2022).

37- William S Noble, "What is a support vector machine?" *Nature biotechnology,* Vol. 24 (No. 12), pp. 1565-67, (2006).

38- Hiie Hinrikus *et al.*, "Spectral features of EEG in depression." (2010).

39- Yu Sun, Yingjie Li, Yisheng Zhu, Xingshi Chen, and Shanbao Tong, "Electroencephalographic differences between depressed and control subjects: an aspect of interdependence analysis." *Brain Research Bulletin,* Vol. 76 (No. 6), pp. 559-64, (2008).

40- Hanshu Cai, Xiaocong Sha, Xue Han, Shixin Wei, and Bin Hu, "Pervasive EEG diagnosis of depression using Deep Belief Network with three-electrodes EEG collector." in *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, (2016): *IEEE*, pp. 1239-46.

41- Jian Shen, Shengjie Zhao, Yuan Yao, Yue Wang, and Lei Feng, "A novel depression detection method based on pervasive EEG and EEG splitting criterion." in *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, (2017): *IEEE*, pp. 1879-86.

42- Fernando Soares de Aguiar Neto and João Luís Garcia Rosa, "Depression biomarkers using non-invasive EEG: a review." *Neuroscience & Biobehavioral Reviews,* Vol. 105pp. 83-93, (2019).

43- Paul J Fitzgerald and Brendon O Watson, "Gamma oscillations as a biomarker for major depression: an emerging topic." *Translational psychiatry,* Vol. 8 (No. 1), pp. 1-7, (2018).

44- Wendy Heller, Marci A Etienne, and Gregory A Miller, "Patterns of perceptual asymmetry in depression and anxiety: implications for neuropsychological models of emotion and psychopathology." *Journal of abnormal psychology,* Vol. 104 (No. 2), p. 327, (1995).

45- Kaia Kalev, Maie Bachmann, Laura Orgo, Jaanus Lass, and Hiie Hinrikus, "Lempel-Ziv and multiscale Lempel-Ziv complexity in depression." in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (2015): *IEEE*, pp. 4158-61.