




Schizophrenia diagnosis using innovative EEG feature-level fusion schemes

Atefeh Goshvarpour¹ · Ateke Goshvarpour² 

Received: 28 November 2019 / Accepted: 21 December 2019 / Published online: 2 January 2020
© Australasian College of Physical Scientists and Engineers in Medicine 2020

Abstract

Electroencephalogram (EEG) has become a practical tool for monitoring and diagnosing pathological/psychological brain states. To date, an increasing number of investigations considered differences between brain dynamic of patients with schizophrenia and healthy controls. However, insufficient studies have been performed to provide an intelligent and accurate system that detects the schizophrenia using EEG signals. This paper concerns this issue by providing new feature-level fusion algorithms. Firstly, we analyze EEG dynamics using three well-known nonlinear measures, including complexity (Cx), Higuchi fractal dimension (HFD), and Lyapunov exponents (Lya). Next, we propose some innovative feature-level fusion strategies to combine the information of these indices. We evaluate the effect of the classifier parameter (σ) adjustment and the cross-validation partitioning criteria on classification accuracy. The performance of EEG classification using combined features was compared with the non-combined attributes. Experimental results showed higher classification accuracy when feature-level features were utilized, compared to when each feature was used individually or all fed to the classifier simultaneously. Using the proposed algorithm, the classification accuracy increased up to 100%. These results establish the suggested framework as a superior scheme compared to the state-of-the-art EEG schizophrenia diagnosis tool.

Keywords Schizophrenia · Electroencephalogram · Nonlinear dynamics · Feature-level fusion · Classification

Introduction

Schizophrenia is a lifelong drastic and debilitating mental illness. The quality of life of schizophrenic patients is substantially influenced at various levels of social, cognitive, emotional, physiological, and psychological functions. They suffer from “positive” and “negative” symptoms. The former includes paranoia and hallucinations and the latter comprises cognitive deficiency, reduced emotions, and muddled

thinking. The prevalence of the disease has been reported in about 1% of the world’s population [1]. The intriguing features of the disorder, as well as its pervasiveness, have captured the attention of many researchers.

To evaluate brain activity, one of the most effective modalities is electroencephalography (EEG). This method measures the electrical activity generated by a tremendous number of neurons over the scalp by applying electrodes on some standard positions of the head. Formerly, the signal was used to study the psychological states of the brain in emotions [2, 3], cognitive tasks [4, 5], meditation [6], and pathological conditions like epilepsy [7], autism [8], attention-deficit/hyperactivity disorder (ADHD) [9], and schizophrenia [10, 11].

EEG is a nonlinear, complex, and nonstationary signal and traditional linear EEG measures cannot explain these dynamical brain behaviors and processes. Consequently, a growing number of nonlinear-based theories have been used and developed to interpret the signal. One of the most salient characteristics of dynamic systems is their dependence on the initial conditions. To examine this characteristic, the “Lyapunov exponent” is commonly appraised. Additionally,

✉ Ateke Goshvarpour
ak_goshvarpour@imamreza.ac.ir;
ateke.goshvarpour@gmail.com

Atefeh Goshvarpour
af_goshvarpour@sut.ac.ir

¹ Department of Biomedical Engineering, Faculty of Electrical Engineering, Sahand University of Technology, Tabriz, Iran

² Department of Biomedical Engineering, Imam Reza International University, PO. BOX: 91735-553, Rezvan Campus (Female Students), Phalestine Sq., Mashhad, Razavi Khorasan, Iran

to describe a set of apparently random events over time, the analysis of complexity is an appropriate approach. The complexity of a signal can be estimated by the complexity measures. Several complexity measures have been studied in schizophrenia research. However, the results of all studies have not been consistent. Some literature has reported an increment in the complexity of the brain signal in the disease [11, 12], some have informed a reduction [13–16], and the others have described both reduction and incremental values [17].

The complexity of EEG signals in schizophrenic patients was assessed during sleep [17]. The results revealed a reduction in EEG's dimensionality during sleep stages II and REM and an increment in the complexity during slow-wave sleep. Afterwards, Hoffman et al. [13] showed that the complexity of cortical processes was depressed in schizophrenia. This reduction was prominent in some brain regions, including left frontal and right cortical. By calculating the correlation dimension using spatial embedding, it was manifested that schizophrenics' EEG complexity reduced during a resting closed eyes state [14]. Li et al. [12] computed the Lempel–Ziv complexity of EEGs in two groups of healthy and schizophrenia at resting and during mental arithmetic tasks. Their results explained a higher complexity of the patients' EEGs. However, a greater value of EEGs complexity was reported under the resting state compared to those during the mental arithmetic task. Akar et al. [16] calculated the EEG complexity by means of approximate entropy, Shannon entropy, Kolmogorov complexity, and Lempel–Ziv complexity. The showed lower complexity values in schizophrenia. Additionally, the highlighted a significant complexity difference between schizophrenics and healthy groups in the left frontal (F3) and parietal (P3). Recently, Ibáñez-Molina et al. [11] assessed the complexity of schizophrenic EEGs at rest and during cognitive processing (the subjects were carrying out a naming task). To estimate complexity, they performed the Lempel–Ziv algorithm. Their results revealed a higher complexity of EEGs in schizophrenic patients compared with the controls, specifically in frontal regions at rest. Contradictory of the results of the literature may be due to the nature of complexity estimation algorithms or dissimilar conditions/protocols of the experiments. Fernández et al. [18] have studied several investigations and outlined the theories about possible causes.

In most literature, brain dynamic differences have been studied in healthy and patient groups. However, the purpose of some studies was to classify and distinguish between schizophrenic and control participants. Hornero

et al. [19] analyzed central tendency measure, approximate entropy, and Lempel–Ziv complexity of EEGs in schizophrenic subjects and controls. The highest rates of 80 and 90% for sensitivity and specificity were obtained with the complexity. In another investigation [20], a combination of autoregressive (AR) model parameters, band power, and fractal dimension with a Boosted version of Direct Linear Discriminant Analysis resulted in the maximum classification rate of 87.51%. Sabeti et al. [15] extracted some nonlinear features of EEG, including some entropy measures, Lempel–Ziv complexity, and Higuchi fractal dimension. After applying genetic programming as a feature selection algorithm, the selected features were fed to the linear discriminant analysis (LDA) and adaptive boosting (AdaBoost). The highest classification rates of 89% and 91% were reported by LDA and AdaBoost, respectively. Later, the team [21] performed the channel selection procedure before using the feature selection, using mutual information techniques. The accuracy improved up to 91.94%. Shim et al. [22] examined the role of feature fusion in the machine-learning-based diagnosis of schizophrenia. Their results emphasized the overall improvement of the classification accuracy with simultaneous applying of sensor-level and source-level feature sets, where a maximum accuracy of 88.24% was obtained.

In this study, three nonlinear features were selected, which are complexity (Cx), Higuchi fractal dimension (HFD), and Lyapunov exponents (Lya). The Lempel–Ziv complexity algorithm was used to estimate the complexity. These features have been frequently implemented in previous studies on schizophrenia [12, 15, 23–25]. Initially, we examine the nonlinear dynamics of the EEG signals in different areas of the scalp using these methods. Additionally, the proficiency of each index in the classification of schizophrenia is appraised. We also scrutinize the ability of a combined feature in the performance of the disease classification. Precisely, different fusion rules are proposed, for the first time. We hypothesize these simple integration techniques can improve the classification performances while reducing the dimension of the classifier inputs.

Materials and methods

In this experiment, we intended to develop an automated system for schizophrenia diagnosis using EEG signals. To this effect, we initially examined some nonlinear features of

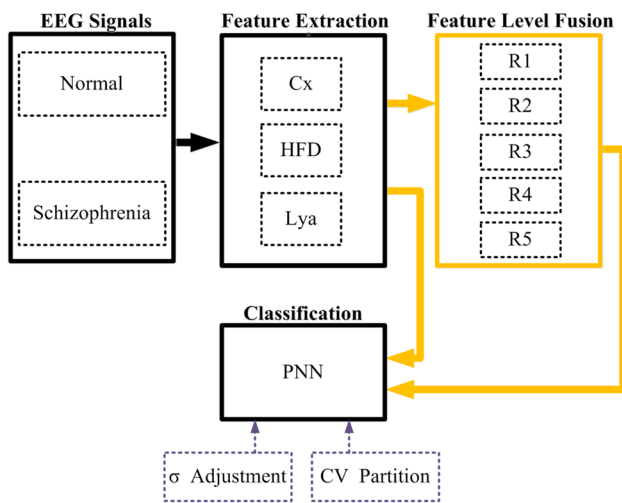


Fig. 1 Proposed scheme

EEG signals, including Cx, HFD, and Lya. Then, we proposed some novel feature-level fusion strategies to combine the nonlinear EEG characteristics and provide new and more efficient features. In addition, we applied the probabilistic neural network (PNN) to classify healthy and pathologic states using different cross-validation (CV) schemes and the classifier adjustment parameter.

The proposed schizophrenia detection framework is presented in Fig. 1.

Data

In this study, the EEG data were selected from the publicly available database at RepOD [26]. It contains EEG signals of 14 paranoid schizophrenias (7 males and 7 females, with the mean age of 27.9 ± 3.3 years and 28.3 ± 4.1 years, correspondingly). The patients met ICD-10 benchmarks (International Classification of Diseases) in the category F20.0 with a minimum age of 18 and a medication washout term of at least 7 days. In addition, the EEG time-series of 14 healthy controls were included in the database (7 males and 7 females, with the corresponding mean age of 26.8 ± 2.9 and 28.7 ± 3.4 years) [27].

The duration of all signals was 15 min, while participants were asked to be in an eyes-closed resting state. The EEG was digitized at the rate of 250 Hz. A 19-channel EEG recording was acquired using the standard 10–20 system. The EEG channels were Fp1, Fp2, F7, F3, Fz, F4, F8, T3,

C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, O2, with the reference electrode of FCz. More information about the dataset is accessible in [27].

Feature extraction

Complexity

In this study, we execute the complexity offered by Lempel and Ziv [28]. The procedure was realized as explained in [29].

In general, this algorithm can be summarized as follows. First, the numerical information is transformed into a symbolic one. Subsequently, distinct words are created by decomposing symbolic sequences, which are encoded by the length of $L(n)$. Lastly, the computation of Lempel–Ziv complexity (Cx) is realized as:

$$Cx = \frac{L(n)}{n} \quad (1)$$

HFD

A fractal dimension is a measure that quantifies the complexity of a time-series in the time domain. A number of fractal dimension algorithms have been proposed theoretically and empirically. One of these procedures called “Higuchi’s fractal dimension (HFD)”, which provided by Higuchi [30]. It serves as a measure of self-similarity and irregularity of a time-series. HFD is estimated using the slope of the linear fit over the log–log plot of the size and scales of the time-series [30]. Its value ranges from 1 to 2. When the dimension of a simple curve is quantified, the HFD is near 1. In contrast, when an arbitrarily dispersed curve that just about fills the Euclidean two-dimensional space, the HFD is about 2.

Lyapunov exponents

Briefly, Lyapunov exponents (Lya) show the average growing ratio of the primary distance between two neighboring points in the phase space. Assume the distance between the points at time 0 is $\|\delta x_i(0)\|$ and at a time t is $\|\delta x_i(t)\|$. Lya (λ_i) is calculated as follows:

$$\frac{\|\delta x_i(t)\|}{\|\delta x_i(0)\|} = 2^{\lambda_i t} \quad (t \rightarrow \infty) \quad (2)$$

$$\lambda_i = \lim_{t \rightarrow \infty} \frac{1}{t} \log_2 \frac{\|\delta x_i(t)\|}{\|\delta x_i(0)\|}$$

This measure is usually applied to determine chaotic or periodic behavior of a time-series. A chaotic signal is a sensitive dependence on initial conditions. When slight deviations in the state variables occur, a great alteration in the behavior at some future points is produced. Consequently, two neighboring points of the trajectory at time 0 broadly diverge from one another at time t .

On behalf of $Lya < 0$, a common fixed point is exposed. When $Lya = 0$, a stable attractor is established, and else a chaotic attractor is recognized.

In this study, the method developed by Rosenstein et al. [31] was adopted for the actual estimation of Lya .

Feature-level fusion

In this study, we examined some feature level fusion approaches. Assume a feature vector of **A** for Cx, **B** for HFD, and **C** for Lya . Different combining rules were proposed as follows.

- Rule 1 (R1): summation rule.

$$R1_i = \mathbf{A}_i + \mathbf{B}_i + \mathbf{C}_i \quad (3)$$

where $i = 1, 2, \dots, 19$, corresponds to 19 electrodes of EEGs.

- Rule 2 (R2): product rule.

$$R2_i = \mathbf{A}_i \times \mathbf{B}_i \times \mathbf{C}_i \quad (4)$$

- Rule 3 (R3): division rule.

$$R3_i = \mathbf{A}_i / \mathbf{B}_i / \mathbf{C}_i \quad (5)$$

- Rule 4 (R4): weighted sum rule 1.

$$R4_i = F_{Ai} \times \mathbf{A}_i + F_{Bi} \times \mathbf{B}_i + F_{Ci} \times \mathbf{C}_i \quad (6)$$

where F_A , F_B , and F_C denote the corresponding F-values of the statistical test for each feature.

- Rule 5 (R5): weighted sum rule 2.

$$R5_i = IGR_{Ai} \times \mathbf{A}_i + IGR_{Bi} \times \mathbf{B}_i + IGR_{Ci} \times \mathbf{C}_i \quad (7)$$

where IGR_A , IGR_B , and IGR_C denote the information gain ratio (IGR) of each feature.

IGR shows the quantity of “information” included in each feature about the class [32]. Consequently, there is no information about unrelated features. If the information gain (IG) is applied for the classification it can be moderate the uncertainty of each attribute, where the uncertainty is well-defined by the entropy. Accordingly, the greater the IGs signify the better and more favorable outcomes. The entropy (E) of the variable X is formulized as:

$$E(X) = - \sum_{x \in X} p_X(x) \log_2 p_X(x) \quad (8)$$

where $p_X(x)$ is the marginal probability density function for the X. Assume the training set termed as “S” and the observed values of X in the S are fragmented based on the values of a second variable Y. If the $E(X)$ with reference to the partitions convinced by Y is less than the $E(X)$ in advance of partitioning, then an association exists between X and Y. The $E(X)$ after observing Y is as follows:

$$E(X|Y) = - \sum_{y \in Y} p_Y(y) \sum_{x \in X} p_{X|Y}(x|y) \log_2 p_{X|Y}(x|y) \quad (9)$$

where $p(x|y)$ is the conditional probability of x given y .

Now, IG, a symmetrical measure, can be defined as:

$$IG = E(X) - E(X|Y) = E(Y) - E(Y|X) \quad (10)$$

The main drawback of this index is that it is biased for attributes with more values even when they are not more informative. The solution is to normalize the IG by dividing by the $E(Y)$. This ratio known as the IGR and well-defined as follows:

$$IGR = \frac{IG}{E(Y)} \quad (11)$$

In weighted sum rules, it was assumed that the involvement of some measures is more than the other features. Therefore, the features were weighted regarding their significance in each class. This significance was defined based on two measures: (1) the statistical F values of ANOVA test (which shows in Table 1) and (2) the corresponding IGR values.

Figure 2 demonstrates the feature level fusion rules, schematically.

Classification

The classification was performed using a probabilistic neural network (PNN) to separate EEG features of healthy and

Table 1 Statistical differences between nonlinear measures of healthy and schizophrenia states in different EEG electrode positions

Nonlinear feature	EEG channel	<i>F</i> value	<i>p</i> value	<i>HSD</i>
HFD	Fpz	9.29	0.005	*
	F8	14.45	0.0007	*
	T4	10.52	0.003	*
	T6	6.57	0.01	*
	O2	3.80	0.06	Ns
	Fp1	2.17	0.15	Ns
	F7	2.80	0.1	Ns
	T3	7.40	0.01	*
	T5	3.26	0.08	Ns
	O1	4.90	0.03	*
	F4	9.64	0.004	*
	C4	13.87	0.0009	*
	P4	4.70	0.03	*
	F3	1.42	0.2	Ns
	C3	2.49	0.12	Ns
	P3	3.31	0.08	Ns
	Fz	5.01	0.03	*
	Cz	4.37	0.04	*
	Pz	6.32	0.01	*
Lya	Fpz	5.25	0.03	*
	F8	8.58	0.006	*
	T4	5.45	0.027	*
	T6	3.82	0.061	Ns
	O2	1.11	0.30	Ns
	Fp1	0.73	0.40	Ns
	F7	0.35	0.55	Ns
	T3	3.50	0.07	Ns
	T5	0.73	0.39	Ns
	O1	3.10	0.09	Ns
	F4	6.75	0.01	*
	C4	10.15	0.003	*
	P4	4.07	0.054	Ns
	F3	0.67	0.42	Ns
	C3	1.27	0.26	Ns
	P3	1.59	0.21	Ns
	Fz	4.01	0.055	Ns
	Cz	3.05	0.09	Ns
	Pz	7.10	0.01	*
Cx	Fpz	1.55	0.22	Ns
	F8	1.14	0.29	Ns
	T4	3.36	0.07	Ns
	T6	0.24	0.62	Ns
	O2	1.55	0.22	Ns
	Fp1	2.04	0.16	Ns
	F7	1.49	0.23	Ns
	T3	4.94	0.03	*
	T5	0.52	0.47	Ns
	O1	1.18	0.28	Ns

Table 1 (continued)

Nonlinear feature	EEG channel	<i>F</i> value	<i>p</i> value	<i>HSD</i>
	F4	5.57	0.02	*
	C4	3.51	0.07	Ns
	P4	4.24	0.04	*
	F3	3.02	0.09	Ns
	C3	0.89	0.35	Ns
	P3	5.40	0.02	*
	Fz	6.19	0.01	*
	Cz	9.46	0.004	*
	Pz	7.93	0.009	*

Ns not significant

**p* < 0.05

schizophrenic participants. The PNN is assigned to the set of feed-forward neural networks which handles the radial basis function using the sigma adjustment parameter (σ). The kernel function is used for the nonlinear distance metric. To realize optimum results, the shape of the kernel functions must be tuned. This is done by adjusting the σ parameters. The subsequent stages indicate the PNN mechanism.

1. In the primary layer of the network, the classifier computes how close the attribute vector to the training vector is. Then, it saves the results in a new vector.
2. The second layer of the network calculates the summation of the contributions for each class. Then, it puts the outcomes in a probability vector.
3. The last layer of the network is the competing transfer function, which selects the highest probabilities. It saves “1” for the class with the highest probabilities and a “0” for the other classes [33].

The effectiveness of PNN in biomedical signal classification, detection, and pattern recognition has been proven in the foregoing studies [5, 34–37].

In advance, the features were normalized (Norm_F) as follows and then they input the PNN:

$$Norm_F = 2 \left(\frac{F - F_{\min}}{F_{\max} - F_{\min}} \right) - 1 \quad (12)$$

The classifier was tested by different cross-validation schemes, including k-fold cross-validation with $k=2, 5$, and 10 , and a holdout method with the proportion of observations to hold out for the test set (p) was fixed to $0.15, 0.2$, and 0.25 . The cross-validation processes fulfill the advantages of avoiding over-fitting and producing reliable network performances. In addition, different σ values were assessed applied

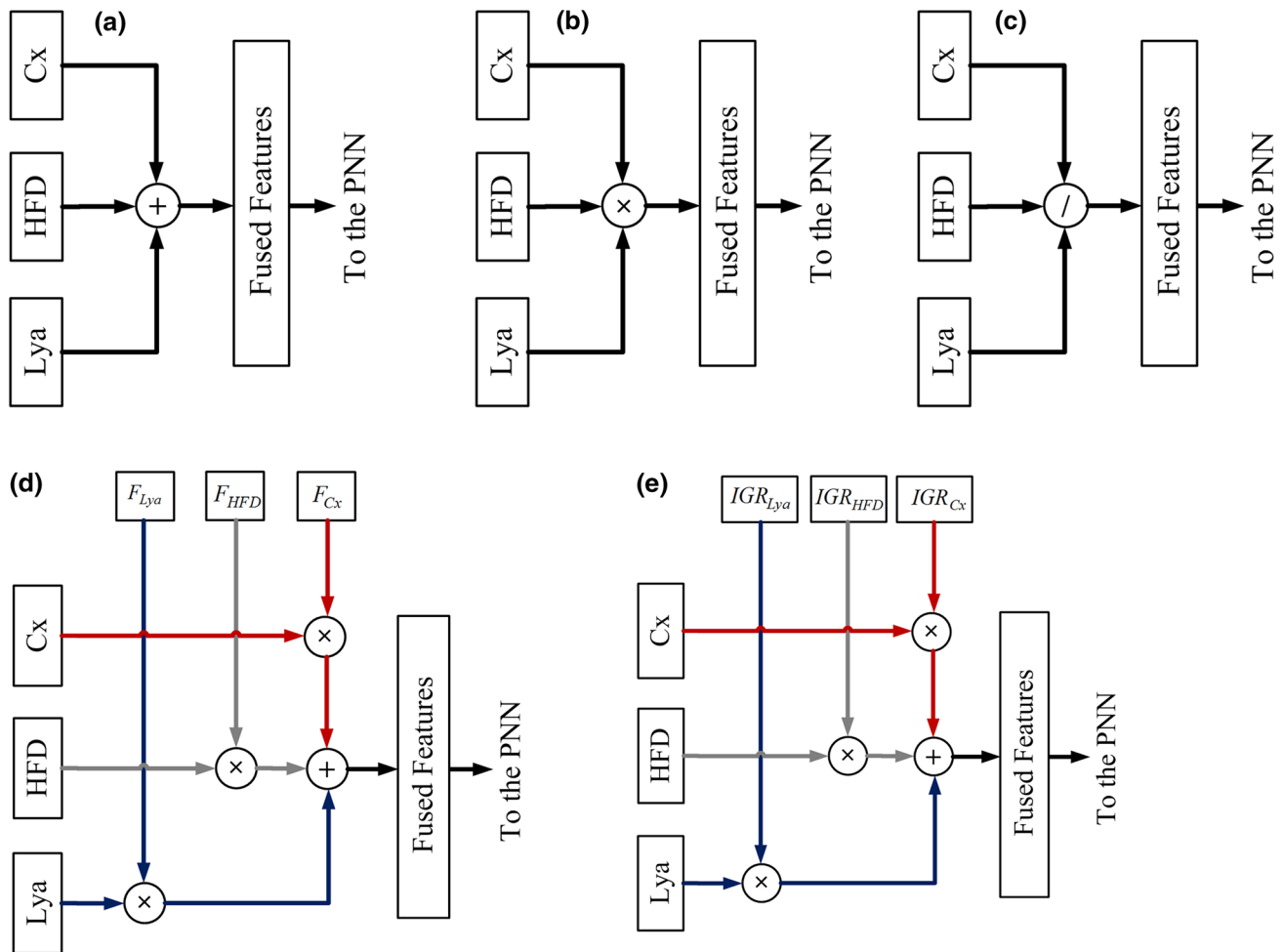


Fig. 2 Feature level fusion rules **a** R1; **b** R2; **c** R3; **d** R4; and **e** R5

to evaluate the performance of the classifier. The best classification performance results were obtained when the σ was in the range of 0.001 to 0.1; i.e. $\sigma = 0.001, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1$). To appraise the network performance, accuracy, sensitivity, and specificity were calculated.

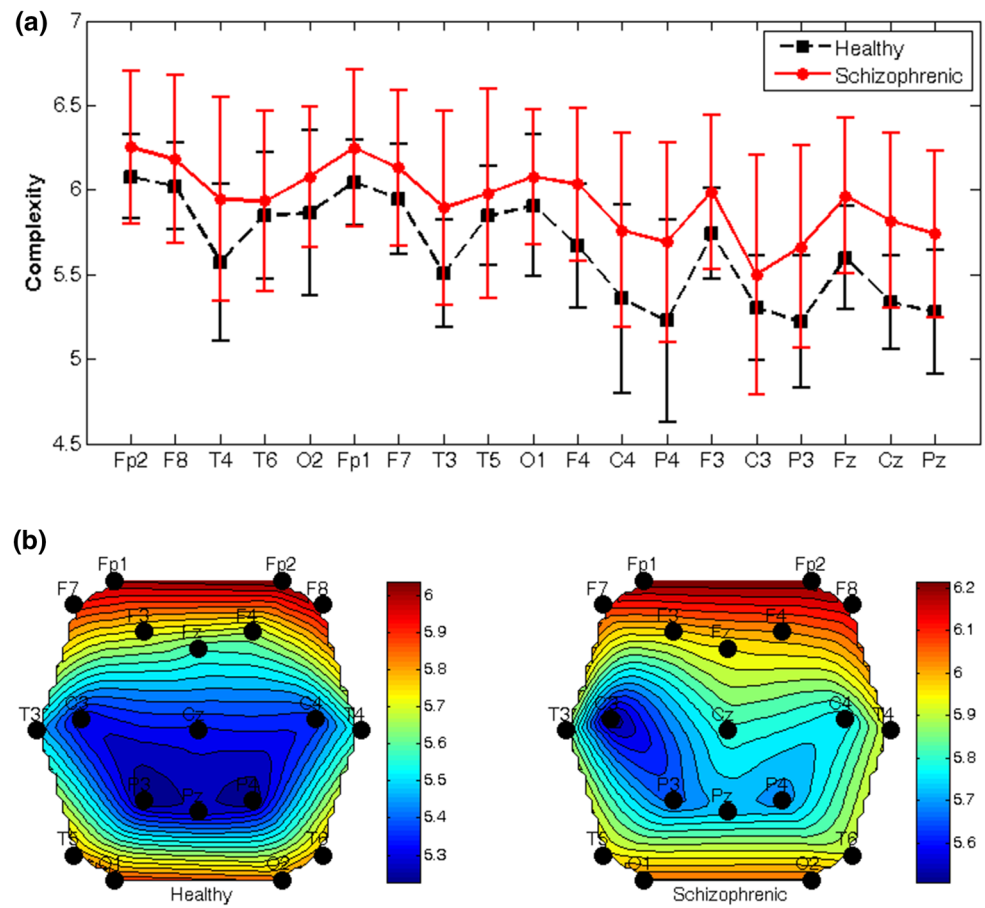
Results

Figures 3, 4, and 5 show the variation of nonlinear indices (i.e. Cx, HFD, and Lya) in different electrode positions in healthy and schizophrenic patients. In addition, the

distribution of the features across the scalp has been provided in the figures.

The results showed not only the values of nonlinear parameters are different in healthy and pathologic conditions for all electrodes, but also the intensity of these variables varies over the scalp. The maximum value of the Cx and HFD was concentrated in the frontal area. However, for Lya, the maximum value was distributed over the central and occipital regions. Additionally, the values of Cx and Lya were higher for schizophrenic patients compared to healthy participants. However, for HFD, this pattern was reversed, where lower values of HFD were observed in the schizophrenia state.

Fig. 3 Variations of Cx in healthy and schizophrenia; **a** Cx by the electrode, **b** the Cx distribution across the scalp



To examine significant differences between the conditions, Tukey's HSD (honestly significant difference) test in conjunction with an ANOVA (post hoc analysis) was performed. Precisely, the purpose of this algorithm is to find means that are significantly different from each other. Corresponding p value, F value, and the HSD results are reported in Table 1.

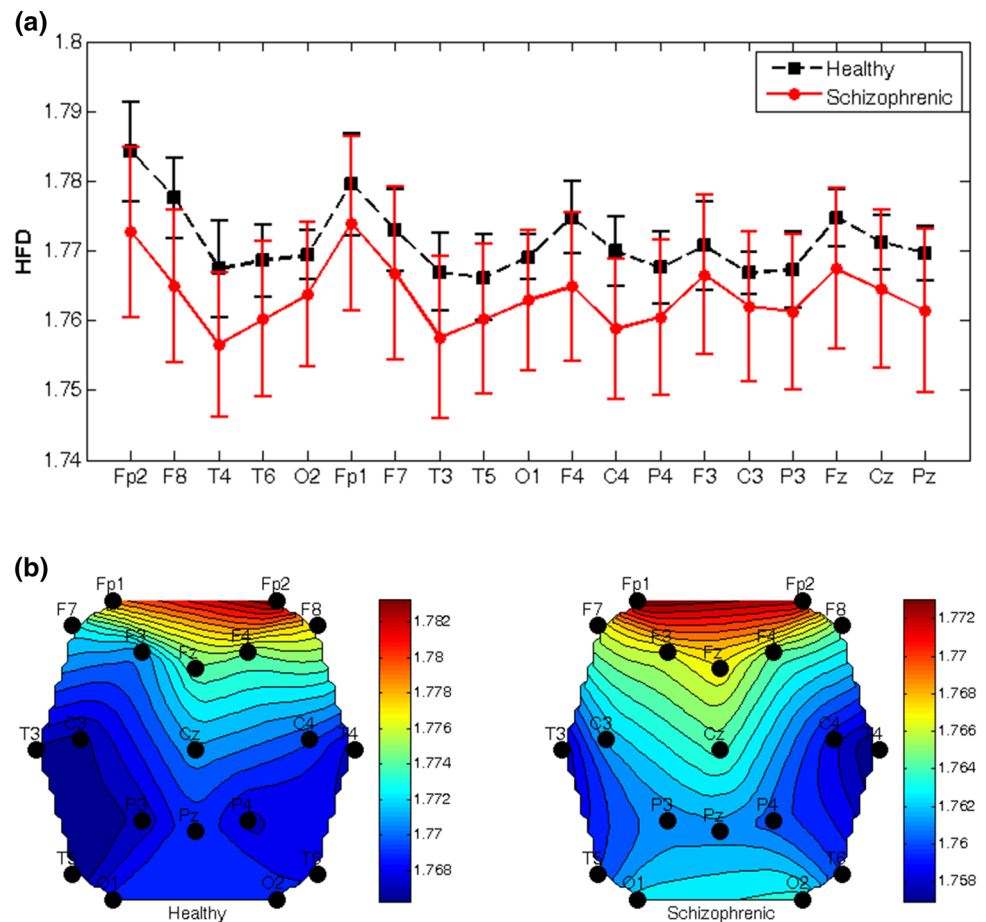
The most significant differences between the values of the brain channels in the two groups were found for the HFD. In this case, HFD values of 12 out of 19 channels were significantly different between healthy and schizophrenic subjects. Surprisingly, the results showed there were no significant differences in the nonlinear properties of the following channels: O2, FP1, F7, T5, F3, and C3, which mostly located in the left hemisphere.

Table 2 summarizes the performance of the PNN classifier in schizophrenia diagnosis. First, each feature was fed

to the classifier separately (Cx, HFD, and Lya). Then, all features were inputted the PNN (Tot). Finally, the efficiency of the proposed feature level fusion rules was examined (R1 to R5).

Table 2 shows that for a twofold CV, the maximum accuracy was 85.71% using R3 and $\sigma=0.09$ and 0.1. For fivefold and tenfold CV, the maximum accuracy rate was increased to 100%. Precisely, for the former, the accuracy of 100% was achieved using R3 and $\sigma \geq 0.07$ and for the latter, the recognition rate of 100% was obtained using R1, R2, R3, and R4. For all k-fold CV strategies, the performance of the classifier was minimal when all the features were applied to the classifier (Tot). The best results for the holdout strategy were similarly matched to the k-fold strategy. In this case, again the highest performances were realized using R3. In addition, for $p=0.2$ and 0.25, the lowest classification rates were

Fig. 4 Variations of HFD in healthy and schizophrenia; **a** HFD by the electrode, **b** the HFD distribution across the scalp



reached using Tot. Typically, k-fold strategy performances were superior to those of the holdout strategy.

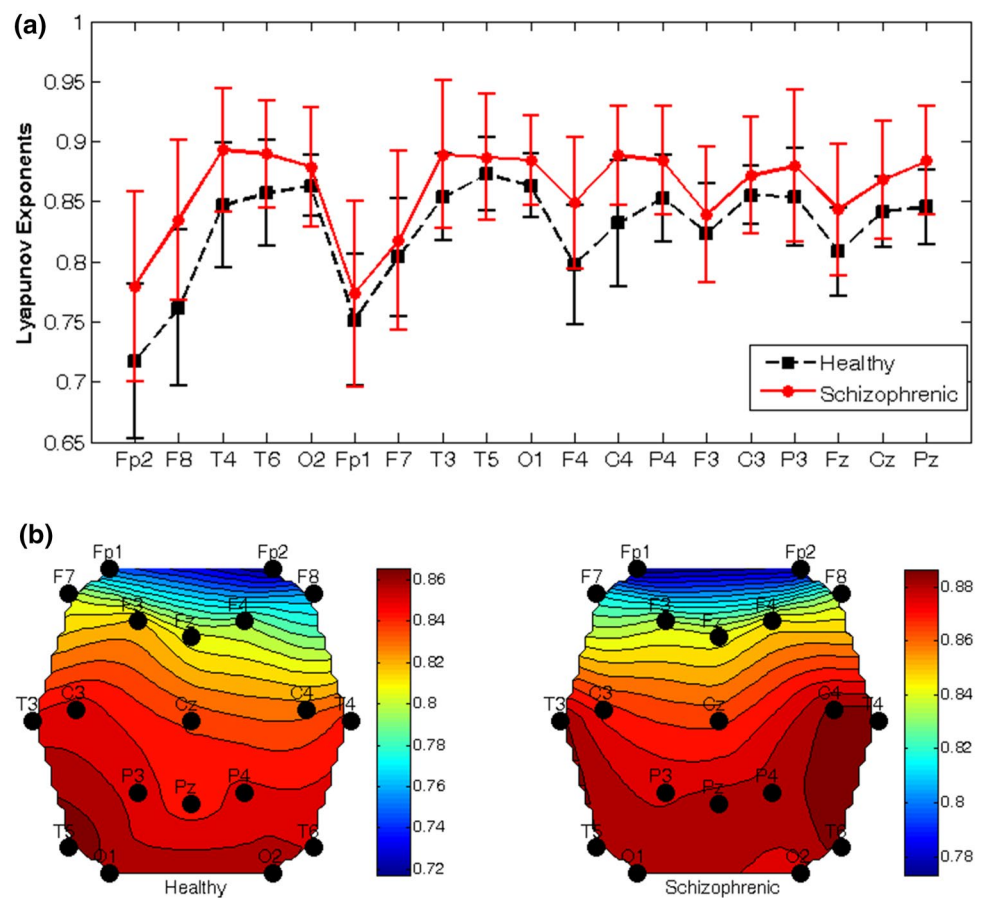
Discussion

In this paper, a comparison among nonlinear EEG indices was performed to detect schizophrenic state. In addition, several feature-level fusion strategies were evaluated. To this effect, various innovative rules, including R1 to R5 were proposed. Then, the discriminative ability of the combined features was determined. The system performances were

evaluated by changing the sigma parameter of the classifier and for different CV approaches.

Our findings stated higher values of Cx and Lya for schizophrenia compared to healthy participants. In addition, lower HFD values were obtained in the schizophrenia condition. These findings are in the line with some previously published articles. Earlier, it has been shown that EEG complexities are affected by the disease. A greater EEG complexity was also reported in [11, 12]. Besides, we found different intensities of the indices over the scalp, with higher Cx and HFD values in the frontal area and higher values of Lya over the central and occipital regions. These results were in contradiction with the findings of [13], in which a

Fig. 5 Variations of L_{ya} in healthy and schizophrenia; **a** L_{ya} by the electrode, **b** the L_{ya} distribution across the scalp



reduction of complexity in the frontal area of the brain was achieved. Our results also showed that there were no significant differences between the nonlinear dynamics of healthy and schizophrenic patients in O2, FP1, F7, T5, F3, and C3, which mostly located in the left hemisphere.

The results highlighted a substantial role of parameter adjustment in classification performances. Both CV and σ influenced the classification accuracy rates. Broadly, k-fold strategy performances were superior to those of the holdout approach. Moreover, the advantage of feature-level merging policies was established in EEG classification. Earlier, a beneficial role of fusion in the enhancement of the classifier performances for other psychological data as well as for schizophrenia diagnosis has been presented [22, 35, 38, 39].

By merging the nonlinear indices, valuable information of EEG time-series was obtained which provoked superior classification rates. Each non-linear index provides different information about signal dynamics. For example, the divergence rate of two adjacent points over time in phase space is obtained by the Lyapunov exponents. While the complexity does not include this information. Subsequently, the information about these measures is complementary. The highest classification performances (100%) were accomplished by fusion policies, exclusively division rule (R3) obtained a higher amount of maximum accuracy in different CVs and for various σ values (Table 2).

EEG classification in schizophrenia has not been well-documented. A few studies attempted diagnosing

Table 2 PNN results in classification of healthy and schizophrenia states using different cross validation methods and fusion strategies

Cross validation	Strategy	Ac	Sp	Se	σ
k-fold, k = 2	Cx	78.57	71.43	85.71	≥ 0.07
	HFD	71.42	71.43	71.42	≥ 0.05
	Lya	78.57	57.15	100	≥ 0.07
	Tot	57.15	57.15	57.15	≥ 0.09
	R1	64.29	57.15	85.71	≥ 0.03
	R2	57.17	57.15	71.43	≥ 0.08
	R3	85.71	100	71.43	≥ 0.09
	R4	57.15	57.15	71.43	≥ 0.07
	R5	71.43	71.43	71.43	≥ 0.07
k-fold, k = 5	Cx	80	66.67	100	≥ 0.04
	HFD	60	66.67	50	≥ 0.05
	Lya	40	60	100	≥ 0.001
	Tot	40	60	100	≥ 0.001
	R1	80	66.67	100	≥ 0.09
	R2	60	60	100	≥ 0.07
	R3	100	100	100	≥ 0.07
	R4	60	66.67	50	0.04
	R5	80	100	100	0.04 and ≥ 0.07
k-fold, k = 10	Cx	50	70	100	≥ 0.001
	HFD	50	70	100	≥ 0.001
	Lya	50	100	100	≥ 0.06
	Tot	50	70	100	≥ 0.001
	R1	100	100	100	0.03 and 0.04
	R2	100	100	100	≥ 0.05
	R3	100	100	100	≥ 0.04
	R4	100	100	100	≥ 0.07
	R5	50	70	100	≥ 0.001
Holdout, p = 0.15	Cx	50	50	100	≥ 0.001
	HFD	75	50	100	≥ 0.05
	Lya	75	100	100	≥ 0.06
	Tot	75	50	100	≥ 0.08
	R1	75	100	100	≥ 0.07
	R2	50	50	100	≥ 0.001
	R3	100	100	100	≥ 0.05
	R4	75	100	100	≥ 0.07
	R5	75	50	100	0.03–0.06
Holdout, p = 0.2	Cx	80	50	100	0.04–0.07
	HFD	80	66.67	100	≥ 0.05
	Lya	60	50	100	≤ 0.02
	Tot	40	50	100	≥ 0.001
	R1	80	66.67	100	0.03
	R2	80	100	66.67	≥ 0.05
	R3	100	100	100	≥ 0.04
	R4	80	100	100	≥ 0.05
	R5	60	66.67	50	≥ 0.07

Table 2 (continued)

Cross validation	Strategy	Ac	Sp	Se	σ
Holdout, p = 0.25	Cx	57.15	75	66.67	≥ 0.07
	HFD	85.71	100	75	≥ 0.08
	Lya	57.15	50	66.67	≥ 0.04
	Tot	42.86	25	66.67	≥ 0.001
	R1	71.42	66.67	75	≥ 0.07
	R2	71.42	50	100	≥ 0.05
	R3	85.71	100	100	≥ 0.09
	R4	42.86	50	66.67	≤ 0.03
	R5	57.14	50	66.67	0.04 and 0.05

Cx complexity, *HFD* Higuchi fractal dimension, *Lya* Lyapunov exponents, *Tot* total features, *R1* feature level fusion using Rule 1 (Summation), *R2* feature level fusion using Rule 2 (Product), *R3* feature level fusion using Rule 3 (Division), *R4* feature level fusion using Rule 4 (Weighted Sum using F-value), *R5* feature level fusion using Rule 5 (Weighted Sum using IGR), *Ac* accuracy, *Sp* specificity, *Se* sensitivity, σ sigma parameter adjustment in PNN

schizophrenia using EEG data. A comparison of previous EEG-based techniques used for schizophrenia classification with our framework has been provided in Table 3.

Concisely, our results established that innovative combined nonlinear features of EEG signals could be served as an applicable measurement to correctly discriminate between schizophrenics and controls.

One of the drawbacks of the present experiment is the limited number of EEG signals. It is recommended the scheme be evaluated on big data in the future. Furthermore, the data were related to paranoid patients. The system's performance in differentiating the different levels of the disease should be evaluated in future work. We used the EEG database available at RepOD [26], where the data were recorded in a medication washout term of at least 7 days. It is suggested that the effect of the drug on the EEG dynamics and classification results be investigated in future work.

Conclusions

In this work, we endeavored to characterize the EEG dynamics of schizophrenic patients and classify them from the EEG dynamics of healthy participants. To address this problem, three nonlinear measures were extracted, including *Cx*, *HFD*, and *Lya*. Additionally, several innovative feature level fusion methods were proposed. The experimental results

Table 3 Summary of automated schizophrenia diagnosis frameworks using EEG

Literatures	No. of control/ patient	Methodology	Results
[19]	20/20	CTM, approximate entropy, complexity	Sensitivity: 80% Specificity: 90%
[20]	18/13	AR parameters, band power, FD; BDLDA	Accuracy: 87.51%
[15]	20/20	Entropy, complexity, HFD, genetic programming, Adaboost	Accuracy: 91%
[21]	20/20	Channel selection, entropy, complexity, HFD, genetic programming, Adaboost	Accuracy: 91.94%
[22]	34/34	Combined sensor-level and source-level EEG features	Accuracy: 88.24%
Present work	14/14	Cx, HFD, Lya, feature-level fusion	Accuracy: 100% Sensitivity: 100% Specificity: 100%

CTM central tendency measure, AR autoregressive, FD fractal dimension, BDLDA: boosted version of direct linear discriminant analysis

established the ability of the proposed fusion strategies for schizophrenia diagnosis. Definitely, the results advocated that amalgamation of nonlinear dynamical information of the EEGs based on division rule (R3) can increase the accuracy of the classifier up to 100%. The constructive role of parameter tuning and CV selection on classification performances has been also demonstrated.

Funding This research did not receive any specific Grant from funding agencies in the public, commercial, or not-for-profit sectors.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants performed by any of the authors.

References

1. Tschacher W, Giersch A, Friston K (2017) Embodiment and schizophrenia: a review of implications and applications. *Schizophr Bull* 43:745–753
2. Goshvarpour A, Goshvarpour A (2019) EEG spectral powers and source localization in depressing, sad, and fun music videos focusing on gender differences. *Cogn Neurodyn* 13(2):161–173
3. Goshvarpour A, Abbasi A, Goshvarpour A (2016) Combination of sLORETA and nonlinear coupling for emotional EEG source localization. *Nonlinear Dyn Psychol Life Sci* 20(3):353–368
4. Gardony AL, Eddy MD, Brunyé TT, Taylor HA (2017) Cognitive strategies in the mental rotation task revealed by EEG spectral power. *Brain Cogn* 118:1–18
5. Goshvarpour A, Goshvarpour A (2018) Automatic EEG classification during rapid serial visual presentation task by a novel method based on dual-tree complex wavelet transform and Poincare plot indices. *Biomed Phys Eng Express* 4:065022
6. Goshvarpour A, Rahati S, Goshvarpour A, Saadatian V (2012) Estimating the depth of meditation using electroencephalogram and heart rate signals. *ZUMSJ* 20(79):44–54 (in Persian)
7. Vytautas A, Misiūnas M, Meškauskas T, Samaitienė R (2019) Algorithm for automatic EEG classification according to the epilepsy type: benign focal childhood epilepsy and structural focal epilepsy. *Biomed Signal Process Control* 48:118–127
8. Kang J, Chen H, Li X, Li X (2019) EEG entropy analysis in autistic children. *J Clin Neurosci* 62:199–206
9. Karimui RY, Azadi S, Keshavarzi P (2019) The ADHD effect on the high-dimensional phase space trajectories of EEG signals. *Chaos, Solitons Fractals* 121:39–49
10. Dvey-Aharon Z, Fogelson N, Peled A, Intrator N (2015) Schizophrenia detection and classification by advanced analysis of EEG recordings using a single electrode approach. *PLoS ONE* 10(4):e0123033
11. Ibáñez-Molina AJ, Lozano V, Soriano MF et al (2018) EEG multiscale complexity in schizophrenia during picture naming. *Front Physiol* 9:1213
12. Li Y, Tong S, Liu D et al (2008) Abnormal EEG complexity in patients with schizophrenia and depression. *Clin Neurophysiol* 119:1232–1241
13. Hoffmann RE, Buchsbaum MS, Jensen RV et al (1996) Dimensional complexity of EEG waveforms in neuroleptic-free schizophrenic patients and normal control subjects. *J Neuropsychiatry Clin Neurosci* 8:436–441
14. Lee Y-J, Zhu Y-S, Xu Y-H et al (2001) Detection of non-linearity in the EEG of schizophrenic patients. *Clin Neurophysiol* 112:1288–1294
15. Sabeti M, Katebi S, Boostani R (2009) Entropy and complexity measures for EEG signal classification of schizophrenic and control participants. *Artif Intell Med* 47:263–274
16. Akar SA, Kara S, Latifoglu F, Bilgi V (2016) Analysis of the complexity measures in the EEG of schizophrenia patients. *Int J Neural Syst* 26:1650008
17. Roschke J, Aldenhoff JB (1993) Estimation of the dimensionality of sleep-EEG data in schizophrenics. *Eur Arch Psychiatry Clin Neurosci* 242:191–196
18. Fernández A, Gómez C, Hornero R, López-Ibor JJ (2013) Complexity and schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry* 45:267–276
19. Hornero R, Abasolo D, Jimeno N et al (2006) Variability, regularity and complexity of time series generated by schizophrenic patients and control subjects. *IEEE Trans Biomed Eng* 53(2):210–218

20. Boostani R, Sadatnezhad K, Sabeti M (2009) An efficient classifier to diagnose of schizophrenia based on the EEG signals. *Expert Syst Appl* 36(3):6492–6499
21. Sabeti M, Katebi SD, Boostani R, Price GW (2011) A new approach for EEG signal classification of schizophrenic and control participants. *Expert Syst Appl* 38(3):2063–2071
22. Shim M, Hwang H-J, Kim D-W, Lee S-H, Im C-H (2016) Machine-learning-based diagnosis of schizophrenia using combined sensor-level and source-level EEG features. *Schizophr Res* 176(2–3):314–319
23. Kim DJ, Jeong J, Chae JH et al (2000) An estimation of the first positive Lyapunov exponent of the EEG in patients with schizophrenia. *Psychiatry Res* 98:177–189
24. Kotini A, Anninos P (2002) Detection of non-linearity in schizophrenic patients using magnetoencephalography. *Brain Topogr* 15:107–113
25. Raghavendra BS, Dutt DN, Halahalli HN, John JP (2009) Complexity analysis of EEG in patients with schizophrenia using fractal dimension. *Physiol Meas* 30:795–808
26. Olejarczyk E, Jernajczyk W (2017) EEG in schizophrenia. *RepOD*. <https://doi.org/10.18150/repod.0107441>
27. Olejarczyk E, Jernajczyk W (2017) Graph-based analysis of brain connectivity in schizophrenia. *PLoS ONE* 12(11):e0188629
28. Lempel A, Ziv J (1976) On the complexity of finite sequences. *IEEE Trans Inf Theory* 22(1):75–81
29. Zhang XS, Zhu YS (1999) Detecting ventricular tachycardia and fibrillation by complexity measure. *IEEE Trans Biomed Eng* 46(5):548–555
30. Higuchi T (1988) Approach to an irregular time series on the basis of the fractal theory. *Phys D* 31(2):277–283
31. Rosenstein MT, Collins JJ, DeLuca CJ (1993) A practical method for calculating largest Lyapunov exponents from small data sets. *Phys D* 65:117–134
32. Mitchell TM (1997) *Machine learning*. McGraw-Hill, New York
33. Demuth H, Beale M (2000) *Neural network toolbox*. The MathWorks Inc., Natick
34. Goshvarpour A, Goshvarpour A (2019) Human identification using a new Matching Pursuit-based feature set of ECG. *Comput Methods Programs Biomed* 172:87–94
35. Goshvarpour A, Goshvarpour A (2018) A novel feature level fusion for HRV classification using correntropy and Cauchy-Schwarz divergence. *J Med Syst* 42:109
36. Goshvarpour A, Abbasi A, Goshvarpour A (2017) An accurate emotion recognition system using ECG and GSR signals and matching pursuit method. *Biomed J* 40:355–368
37. Goshvarpour A, Abbasi A, Goshvarpour A (2017) Indices from lagged Poincare plots of heart rate variability: an efficient nonlinear tool for emotion discrimination. *Australas Phys Eng Sci Med* 40(2):277–287
38. Goshvarpour A, Abbasi A, Goshvarpour A (2017) Fusion of heart rate variability and pulse rate variability for emotion recognition using lagged Poincare plots. *Australas Phys Eng Sci Med* 40:617–629
39. Goshvarpour A, Abbasi A, Goshvarpour A, Daneshvar S (2016) A novel signal-based fusion approach for accurate music emotion recognition. *Biomed Eng Appl Basis Commun* 28:1650040

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