



Automatic classification of schizophrenia patients using resting-state EEG signals

Hossein Najafzadeh¹ · Mahdad Esmaeili¹ · Sara Farhang² · Yashar Sarbaz³ · Seyed Hossein Rasta^{1,4,5}

Received: 24 April 2021 / Accepted: 26 July 2021 / Published online: 9 August 2021
© Australasian College of Physical Scientists and Engineers in Medicine 2021

Abstract

Schizophrenia is one of the serious mental disorders, which can suspend the patient from all aspects of life. In this paper we introduced a new method based on the adaptive neuro fuzzy inference system (ANFIS) to classify recorded electroencephalogram (EEG) signals from 14 schizophrenia patients and 14 age-matched control participants. Sixteen EEG channels from 19 main channels that had the most discriminatory information were selected. Possible artifacts of these channels were eliminated with the second-order Butterworth filter. Four features, Shannon entropy, spectral entropy, approximate entropy, and the absolute value of the highest slope of autoregressive coefficients (AVLSAC) were extracted from each selected EEG channel in 5 frequency sub-bands, Delta, Theta, Alpha, Beta, and Gamma. Forty-six features were introduced among the 640 possible ones, and the results included accuracies of near 100%, 98.89%, and 95.59% for classifiers of ANFIS, support vector machine (SVM), and artificial neural network (ANN), respectively. Also, our results show that channels of alpha of O1, theta and delta of Fz and F8, and gamma of Fp1 have the most discriminatory information between the two groups. The performance of our proposed model was also compared with the recently published approaches. This study led to presenting a new decision support system (DSS) that can receive a person's EEG signal and separates the schizophrenia patient and healthy subjects with high accuracy.

Keywords Schizophrenia · Classification · Entropy · Decision support system · Feature selection

Introduction

Schizophrenia is a mental disorder that affects about one in 100 people. Also, among the psychological diseases, this is a serious one, suspending the person from all aspects of life based on the diagnostic criteria of the American Psychiatric Association (DSM-IV) [1], this disease often refers to personality rupture disorder because the thoughts and feelings of the patient won't have a natural and logical connection with each other. The patient is also affected by cognitive impairment. These deficiencies can reduce expression memory, short-term memory, language performance, and other executive functions such as occupational disorders. These deficiencies can help in the early diagnosis of schizophrenia [2]. It is very important to treat schizophrenia as soon as possible after the onset. With delay ineffective treatment, patients may be at increased risk for brain volume loss with harmful implications for long-term treatment outcomes.

Extensive studies have been performed using electroencephalogram (EEG) signals based on DSM-IV criteria for the diagnosis of schizophrenia and comparing schizophrenic

✉ Yashar Sarbaz
yashar.sarbaz@tabrizu.ac.ir

✉ Seyed Hossein Rasta
s.h.rasta@abdn.ac.uk

¹ Department of Medical Bioengineering, Faculty of Advanced Medical Sciences, Tabriz University of Medical Sciences, Golgasht Ave, 51666 Tabriz, Iran

² Department of Psychiatry, School of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran

³ Department of Biomedical Engineering, Faculty of Electrical and Computer Engineering, University of Tabriz, Daneshgah St., Tabriz, Iran

⁴ Department of Medical Physics, School of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran

⁵ Department of Biomedical Physics, School of Medical Sciences, University of Aberdeen, Aberdeen AB25 ZD, UK

patients with healthy subjects [3]. Several researchers have reported that using EEG irregularities and paroxysmal dysrhythmias may have a positive effect on the early diagnosis and prediction of schizophrenia [4].

The analysis of measured EEG signals from schizophrenic patients seems to have significant differences in the two situations of resting and applying tasks. In task-based studies, the responses of neurons with time-locked are specific to the designed event; all other involuntary activities are usually considered background noises. Many studies suggested that the brain is a system that acts intrinsically, with inherent resting-state integration. External sensory information communicates with, rather than assigns, the operation of brain systems. Using resting-state EEG for the diagnosis of schizophrenia seems to have acceptable results. Also, people who were diagnosed, with schizophrenia were referred to a specialist for more accurate management and diagnosis.

In recent years, nonlinear methods and machine learning have been used to classify healthy individuals and patients. Today, the DSS systems are widely used to diagnose a variety of diseases, such as hepatitis [5, 6], Parkinson's [7], etc., which show that intelligent systems that are used as DSS can save lives by helping to diagnose diseases. Boostani et al. [8], used support vector machine (SVM) and linear discriminant analysis (LDA) classifiers to classify resting EEG signals of 13 schizophrenic patients and 18 healthy controls. The extracted features of the autoregressive model coefficient (ARC), subbands frequencies (SF), fractal dimension (FD), and wavelet energy (WE) were used to classify the signals. Also, the genetic algorithm was used to select the best EEG signal recording channels and reduce the dimension of the extracted feature. Most of the channels from the temporal lobe regions of the brain were selected by the genetic algorithm, and the classification accuracy was improved. Boostani et al. [9] and Sabeti et al. [10] used direct linear discriminant analysis (DLDA), weighted distance nearest neighbor (WDNN), basic nearest neighbor (BsNN), naive Bayes (NB), DLDA, adaptive distance measure (ADM), LDA, Adaboost, fuzzy SVM (FSVM), and SVM classifiers to classify the both healthy and schizophrenic groups. The features extracted from the EEG signal included the AR model, SF, and FD. The cross-validation method of Leave-one-out was used to produce outputs with high accuracy and reliability. Their results showed that the SVM, WDNN, and DLDA classifiers in classification were highly accurate. In other two studies, Sabeti et al. [11] used Adaboost and LDA classifiers to classify the EEG signals of 20 schizophrenia patients and 20 healthy subjects in resting state. The extracted features of the measured EEG signals included Shannon entropy (ShEn), spectral entropy (SpEn), approximate entropy (ApEn), Lempel-Ziv complexity (LZC), and Higuchi fractal dimension (HFD). Channels with

discriminatory information, including Cz, C3, T3, T4, Fp2, F3, F4, T5, and O2 most of which were situated in the frontal, temporal, and limbic regions. Li et al. [12] chose artificial neural network (ANN) classifier to classify the EEG signals of depressed and schizophrenia patients as well as healthy individuals. The back-propagation on artificial neural networks (BP ANN) and Self-organizing competitive artificial neural networks (SOC ANN) were used. They extracted the power spectrum as a feature of the EEG signal. Darkhovsky et al. [13] focused more on reducing the dimensions of the extracted features, which is the power spectrum (PS). The autoregressive moving average (ARMA) model algorithm, which estimates PS coefficients, reduces the dimension of the features. After a significant decrease in size (from 96 to 4), SVM and random forest (RF) classifiers were used for classification. Boostani et al. [14] used three methods of phase-locking value (PLV), robust synchronization (RS), and synchronization probability (SL) to extract the feature from the EEG signal. The greedy overall relevancy (GOR) and Across-group variance (AGV) methods were employed to optimize the feature extraction. SVM, DLDA, and Modified nearest neighbor (BNN) classifiers were utilized to distinguish patients with schizophrenia and bipolar disorder. Jahmunah et al. [15] used non-linear entropy features to classify the EEG signal of healthy individuals and schizophrenic patients. The results of this study showed that the SVM classifier was more accurate than classifiers of the LD, *k*-nearest-neighbour (KNN), probabilistic-neural-network (PNN), and DT. Two other studies [16, 17] used a convolutional neural network (CNN) method based on in-depth learning to diagnose healthy individuals and schizophrenic patients. They converted the EEG signal to 2D using the short-time Fourier transform (STFT) to extract the useful features. A literature review on available techniques of schizophrenia classification has been provided in Table 1.

Here, a new classification approach with efficient features is introduced for the diagnosis of schizophrenic patients with healthy subjects. The proposed method is based on ANFIS to classify recorded EEG signals from 14 schizophrenic patients and 14 age-matched control participants. The 2nd order Butterworth filter is used to remove possible artifacts of 16 EEG channels. Four features, namely Shannon, spectral entropy, approximate entropy, and AVLSAC were extracted from each selected EEG channel in 5 frequency sub-bands, including Delta, Theta, Alpha, Beta, and Gamma. In addition, the AVLSAC method has a special initiative that does not seem to have been deployed. The results show that our proposed method is superior to the previously published methods.

Table 1 A summary of the properties of the methods available in EEG signal processing

	Advantages	Disadvantages	Ref
Methods			
Features			
ARC	1. Improved frequency resolution 2. Compatible with frequency spectrum	Sensitive to order models	[18]
WC	1. Ability to analyze all frequency range 2. Capability to analyze sudden changes 3. Ability to analyze complex signals	Sensitive to choosing mother wavelet	
SpEn	1. Describe the signal spectrum irregularities 2. Describe the signal power distribution	Requires stationary data	[9–11, 13, 15, 19–21]
ApEn	1. Measure to the signal irregularity 2. Describes the ratio of new information production 3. Proper performance for noisy data 4. Proper for non-stationary data	Sensitive to data length	
ShEn	1. Measurement of impurities in the signal 2. Proper for PDF estimation of signal	Requires stationary data	
LZC	1. Estimation of the signal complexity 2. Proper for non-stationary signals	Sensitive to data length	
HFD	Measurement of the ratio of increase amplitude	1. Requires stationary data 2. Sensitive to data length 3. Sensitive to noise	
Classifiers			
ANN	1. Saving information on the entire network 2. Ability to work with incomplete knowledge 3. The capability of fault tolerance 4. Ability to make machine learning 5. Ability to Parallel processing	1. Dependence on hardware 2. Unexplained behavior of the network	[15, 18]
LDA	1. Linear and stable 2. Simplicity and low calculations 3. Well-suited for multi-class problems 4. The supervised method	1. Sensitive to variables with higher-order 2. Sensitive to noise	[9, 11, 15, 21, 22]
Adaboost	1. Non-linear classifier 2. Ability to maximize margins 3. Very simple implementation	1. Sensitive to noise and outliers 2. Not support the null rejection	
DLDA	1. Strong and stable 2. Compatibility with short data	1. High sensitivity to the location of samples 2. Sensitive to overlap data	[9, 10, 14]
SVM	1. Linear and nonlinear classifier 2. Stable 3. Independent of the feature dimension space 4. Strong against overtraining	Sensitive to noise	[15, 20, 22–24]
FSVM	1. Suitable for short data	1. Weak and inaccurate 2. Sensitive to data length	
WDNN	1. Compatibility with diverse data 2. Low error	Removing low value data	[10, 15, 22]
NB	1. Easy to implement 2. Nonlinear and stable	1. Sensitive to data length 2. Low speed	
ADM	1. Nonlinear and stable 2. Simplicity	1. Sensitive to data type 2. Sensitive to feature dimensions	
BNN	1. Nonlinear and stable 2. Simplicity	1. Sensitive to feature dimensions 2. Assigning inappropriate coefficients	
ROC	1. High ability to organize classes 2. Assigning the optimal cut-off amount	Sensitive to cut-off selection	[13, 19, 24]
RF	Ability to specify variable importance	1. Sensitive to data length 2. Sensitive to noise	[20, 22]
DT	1. Simplicity 2. Ability to assign specific values to samples 3. Compatibility with short data	1. Unstable 2. Inappropriate prediction 3. Inappropriate for data over fitting	[15, 22]

Materials and methods

Clinical data

We tested our algorithm on 14 healthy controls (7 females: 28.7 ± 3.4 years, 7 males: 26.8 ± 2.9) and 14 patients with paranoid schizophrenia (7 females: 28.3 ± 4.1 years, 7 males: 27.9 ± 3.3 years). The patients were hospitalized at the Institute of Psychiatry and Neurology in Warsaw, Poland and the study protocol was approved by the Ethics Committee of the Institute of Psychiatry and Neurology in Warsaw [25]. The EEG signals were recorded in resting state during eyes-closed (EC) and were measured based on the 10–20 standard for 15 min with a sampling frequency of 250 HZ. All 19 EEG channels were constituted Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, O2 and the electrode FCz was considered as the reference electrode.

The selected areas included frontal, temporal, and occipital as well as 16 channels of Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, T6, O1, and O2 from the 19 main channels which have the greatest overlap with the involved areas are considered. The channels have been selected based on previous studies [11, 26, 27]. The position of the electrodes and brain areas selected for EEG signal processing are shown in Fig. 1.

Pre-processing

Initially, we preprocessed the EEG signals of 16 selected channels which were divided into 32 sections of 25 s to remove the artifacts with the second-order Butterworth filter.

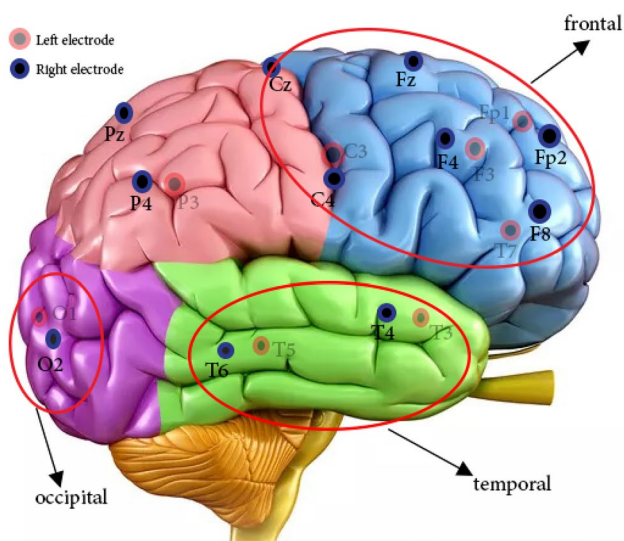


Fig. 1 The position of the electrodes and brain areas selected for EEG signal processing

These segments were without overlap. In the second step, each of the 16 channels was divided into 5 sub-frequency regions of Delta, Theta, Alpha, Beta, and Gamma by the wavelet algorithm with the mother wavelet of Daubechies 2 (Dp2).

Five frequency regions of delta (Δ), theta (θ), alpha (α), beta (β), gamma (γ) were used for more accurate analysis of the EEG signal. The regions of selected frequency filters included 2–4 Hz, 4.5–7.5 Hz, 8.5–12.5 Hz, 13–30 Hz, and 30–45 Hz, respectively. Figure 2 shows the frequency regions of the F4 channel of a healthy individual and schizophrenic individual.

Feature extraction

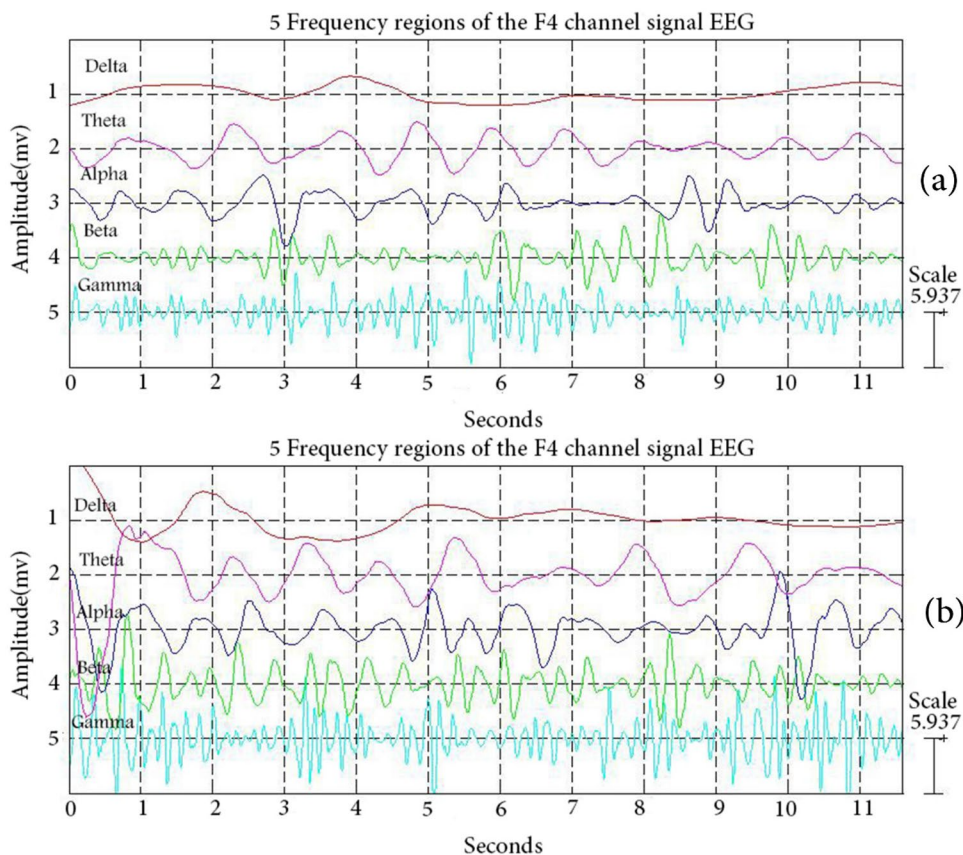
Feature selection is of paramount importance since different types of brain activity patterns can be decoded from EEG signals of patients. Four types of features, namely ShEn, SpEn, ApEn, and the absolute value of the highest slope of autoregressive coefficients (AVLSAC) are used for differentiation of healthy individuals and schizophrenic patients. These proposed features have an excellent reflection of the main nature of the EEG signal. The three entropy methods of spectral, Shannon and approximate, which are compatible with non-stationary and stationary signals, are described in detail [13, 28, 29]. In addition, the ALVSCA seems to be a suitable feature extraction method because it can increase the frequency resolution by using the AR method, which is a powerful algorithm in signal modeling [30].

The AR model with p -order predicts the input S via minimizing forward and backward prediction errors using the Burg method [31]. While the AR coefficients are finite, c_i , to satisfy the Levinson–Durbin returns:

$$s(t) = - \sum_{i=1}^p c_i s(t-i) \quad (1)$$

Value p represents the order of the model. There are several methods to determine the model order, such as decision rules based on the Bayesian approach [32], the number of information measures [33], and the maximum likelihood approach [34]. In this study, the minimum error (threshold) method [35] is used to obtain the best order of the model. Also, we calculated the absolute value of the highest slope between the autoregressive coefficients. The order utilized in this study is 20. Figure 3 shows an example of the absolute value of the maximum slope in a specific window of the F4 channel of a healthy individual and a schizophrenia patient.

Fig. 2 The 5 frequency regions of the F4 channel EEG signal. **a** Healthy individual, **b** schizophrenia patient



Classifiers

Accurate classification and labeling of different groups are executed by proper classifiers. Three types of classifiers, SVM [36, 37], ANN [38], and ANFIS [39], were selected to classify the two groups, and it seems that these three classifiers have better power and better compatibility with the EEG signal of patients with schizophrenia.

SVM structure

Nonlinear SVM_ radial basis function (RBF): given a training set of instance-label pairs (x_i, y_i) , $i = 1, \dots, l$ where $x_i \in \mathcal{R}^n$ and $y_i \in \{-1, +1\}^l$, support vector machines (SVM) require the solution of the following (primal) optimization problem:

$$\min_{w,b,\xi} \frac{1}{2} (w^T w) + c \sum_{i=1}^l \xi_i \tag{2}$$

subject to $y_i (w^T z_i + b) > 1 - \xi_i, \xi_i \geq 0, i = 1, \dots, l$

Here, training vector x_i is mapped into a higher- (maybe infinite-) dimensional space by the function \varnothing as $z_i = \varnothing(x_i)$. $C > 0$ is the penalty parameter of the error term.

Usually, Eq. 2 were used to solve the following dual problem:

$$\min_{\alpha} F(\alpha) = \frac{1}{2} \alpha^T \rho \alpha - e^T \alpha \tag{3}$$

subject to $0 \leq \alpha_i \leq C, i = 1, \dots, l, y^T \alpha = 0$.

where e is the vector of all ones and ρ is an l by l positive semidefinite matrix.

The (i, j) th element of ρ is given by $\rho_{ij} \equiv y_i y_j k(x_i, x_j)$, where $k(x_i, x_j) \equiv \varnothing(x_i)^T \varnothing(x_j)$ is called the kernel function. Then $w = \sum_{i=1}^l \alpha_i y_i \varnothing(x_i)$ and the decision function is as follows:

$$\text{sgn}(w^T \varnothing(x) + b) = \text{sgn} \left(\sum_{i=1}^l \alpha_i y_i k(x_i, x) + b \right) \tag{4}$$

The Gaussian radial basis function (RBF) kernel, two samples of \tilde{x} and \bar{x} samples are defined as feature vectors in some input spaces as follows:

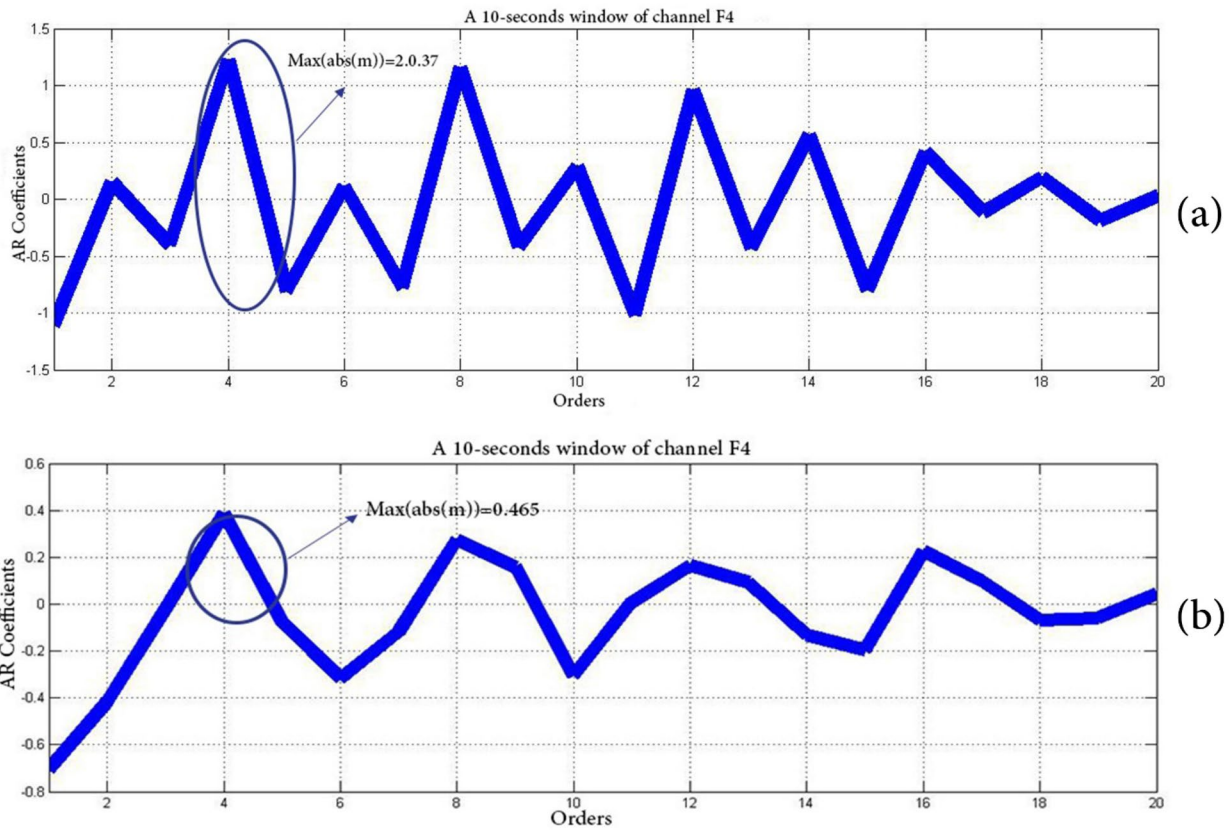


Fig. 3 Example of the absolute value of the maximum slope in a specific window of the F4 channel EEG

$$k(\tilde{x}, \tilde{x}) = \exp\left(-\frac{\tilde{x} - \tilde{x}^2}{2\sigma^2}\right), \gamma = \frac{1}{2\sigma^2} \tag{5}$$

The best value for the kernel function coefficient, gamma (γ)=0.6, was selected. As the gamma value increases, the algorithm tries to perform the fitting exactly based on the training data set, which leads to the generalization of the error and the over-fitting problem. The C parameter trades off the correct classification of training examples against the maximization of the decision function’s margin. The optimal value of C was considered 0.1. Figure 4 shows the architecture of the SVM and the structure of the network in detail.

ANN structure

A multi-layer perceptron neural network (MLPNN) network, which is one of the most generally used neural networks, has been a for the post-classification procedure and trained by a supervised method of learning known as backpropagation [40]. It is a feed-forward Artificial Neural network used for the classification of data. It has several layers that consist of nodes in a graph that is directed. Every layer is made

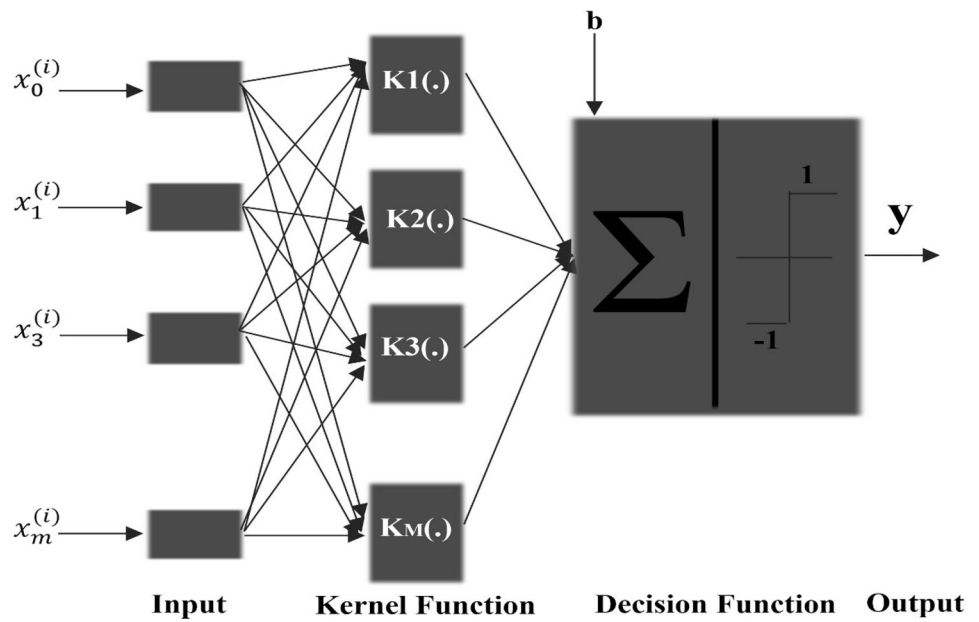
to be fully connected to the subsequent layer. It is a variation of a linear-perceptron that possesses the capability to differentiate samples not linearly separable. Based on the error percentage in the given output obtained via comparing with the desired result, learning occurs inside the network by updating the connecting weights post-processing of every piece of data.

It is designed with three layers: an input layer, a hidden layer, and an output layer. Neurons in the input layer act as buffers for distributing the input signals z_p to neurons in the hidden layer. Each neuron p in the hidden layer sums up its input signal z_p after weighting them with the strengths of the respective connections w_{np} from the input layer. Then it calculates its output y_n by passing the sum through a nonlinear activation function, namely binary sigmoidal function, as presented in the following Equation.

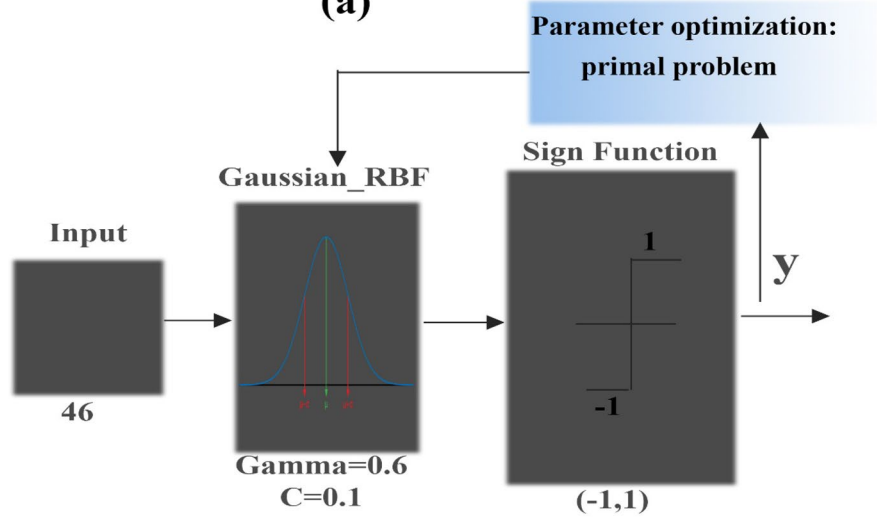
$$y_n = \text{sigm}\left(\sum_p w_{np}z_p\right) \tag{6}$$

The backpropagation algorithm is a gradient descent algorithm [41]. When the momentum term is added, the algorithm

Fig. 4 The architecture of the SVM (a) and the structure of the network in detail (b)



(a)



(b)

gives the change $\Delta w_{pn}(k)$ in the weight of a connection between neurons m and p as follows:

$$\Delta w_{pn}(k) = \gamma \delta_p z_p + \alpha \Delta w_{pn}(k - 1) \tag{7}$$

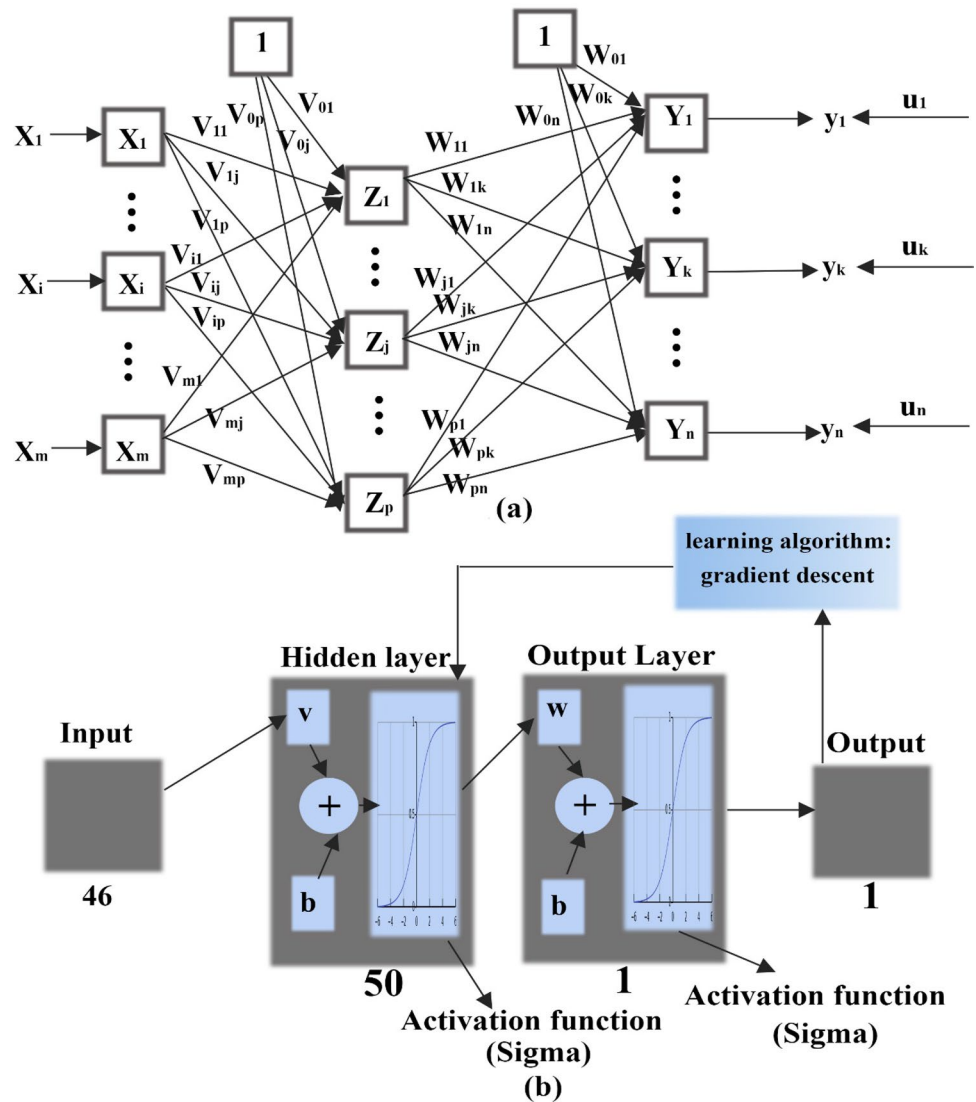
where δ_n is an agent depending on neuron n is a hidden neuron or an output neuron, γ is a learning rate parameter and α is the momentum coefficient. The output neuron is obtained as follows:

$$\delta_n = \left(\frac{\partial f}{\partial net_n} \right) (u_n^{(k)} - y_n), k = 1, \dots, n \tag{8}$$

where $u_n^{(k)}$ is the desired output for neuron n and $\frac{\partial f}{\partial net_n} = y_n(1 - y_n)$. Here f denotes the *sigma* function. Thus, iteratively, beginning with the output layer, the δ term is computed for neurons in all layers and weight updates for all connections according to Eq. 7 [38].

Figure 5 shows the architecture of the perceptron neural network and the structure of the network in detail. The MLPNN is trained several times using different numbers of hidden neurons until achieving the best performance. The most optimal neuron with excellent performance was 50 and the learning rate was $(\gamma) = 0.01$.

Fig. 5 The architecture of the perceptron neural network (a) and the structure of the network in detail (b)



ANFIS structure

The adaptive neuro-fuzzy inference system is a type of artificial neural network, which learns features in the data set and regulates the system parameters according to a given error standard. It is widely used in biological signal analysis. To improve the generalization, ANFIS classifiers were trained with the hybrid back-propagation of gradient descent (GD) method in combination with the least-squares estimation (LSE) technique. Feature vectors were used as inputs for ANFIS. Binary values of (1, 0) and (0, 1) were set as the target outputs for patients with schizophrenia and control subjects, respectively. For simplicity, we assume the fuzzy inference system under consideration with two inputs k_1 and k_2 and one output u . Figure 6

shows our implemented ANFIS architecture. This system is based on the Takagi–Sugeno fuzzy inference system and the fuzzy rule structure of the ANFIS classifiers were created by utilizing a generalized bell-shaped membership function defined as follows:

$$\beta A_i(k_1) = \frac{1}{1 + \left\{ \left(\frac{k_1 - t_i}{a_i} \right)^2 \right\}^{b_i}} \tag{9}$$

where a_i , b_i , and t_i are adaptable parameters, k_1 is the input to node i , and A_i is the linguistic label. Next, two first-order Sugeno-type ANFIS models are implemented with inputs of feature vectors and one output. The first-order Sugeno fuzzy models have the following rules:

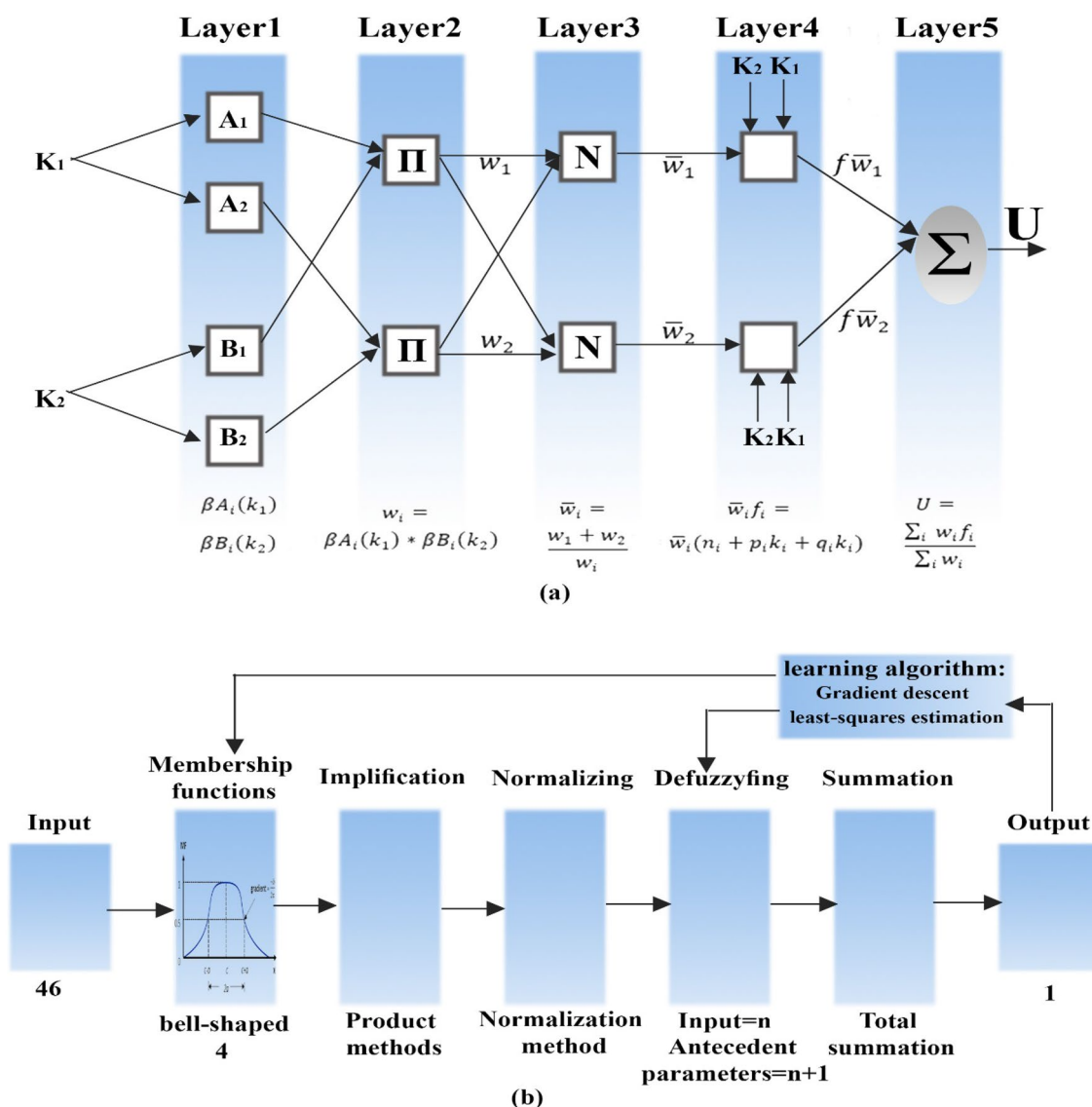


Fig. 6 The architecture of ANFIS and (a) The structure of the network in detail (b)

$$\begin{aligned}
 R_1 &: \text{if}(k_1 \text{ is } A_1 \text{ and } k_2 \text{ is } B_1) \\
 \text{then } u \text{ is } f_1 &= n_1 + p_1 k_1 + q_1 k_2 \\
 R_2 &: \text{if}(k_1 \text{ is } A_2 \text{ and } k_2 \text{ is } B_2) \\
 \text{then } u \text{ is } f_2 &= n_2 + p_2 k_1 + q_2 k_2
 \end{aligned}
 \tag{10}$$

where R_i is the i the rule of the fuzzy system, $k_i(i = 1, 2)$ are the inputs to the fuzzy system, and u is the output of the fuzzy system; $\{n_i, p_i, \text{ and } q_i\}(i = 1, 2)$ are membership function parameters.

Every node in layer 2 is a square node labeled Π which multiplies the incoming signals and sends the product out. For instance:

$$w_i = \beta A_i(k_1) \times \beta B_i(k_2) \tag{11}$$

Every node in layer 3 is a circle node labeled N . The i th node calculates the ratio of the i th rule’s firing strength to the sum of all rules’ firing strengths:

$$\bar{w}_i = \frac{w_1 + w_2}{w_i} \quad (12)$$

Every node i in layer 4 is a square node with a node function:

$$\bar{w}_i f_i = \bar{w}_i (n_i + p_i k_i + q_i k_i) \quad (13)$$

The following formula produces the ANFIS output:

$$U = \sum_{i=1,2} \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (14)$$

ANFIS learns and updates all its modifiable parameters by using a two-pass learning algorithm, namely forward pass and backward pass. ANFIS trains its parameters, such as the number of linear parameters as well as n_i , p_i , and q_i (antecedent parameters) for minimizing error between actual and the desired outputs using a hybrid of GD and LSE. For our data, bell-shaped member functions with number of 4 and epoch = 100 were the best.

Evaluation parameters

Cross-validation is a model evaluation method that determines the extent of generalizability and independency of the results that the statistical analysis on a data set suggests based on training data. The data set is divided into k equal subsets. From this k subset, n/k data for training and m/k data as a model test randomly were selected ($m + n = k, n \geq m$). This procedure is repeated k times. Sixty percent of individuals' feature vectors are used as training samples and 40 percent as test sets. The accuracy

of individuals' feature vectors was calculated five times randomly for further validation, then the average of the five iterations was reported.

Figure 7 shows the confusion matrix. The accuracy and error of the results were obtained through the following mathematical formulas:

$$Accuracy = \frac{TP + TN}{FP + FN + TP + TN} = 1 - Error \quad (15)$$

Here true positives (TP) are the number of positive samples that the model correctly classified as positive; true negatives (TN) are the number of negative samples that the model correctly classified as negative; false positives (FP) mean the number of negative samples that the model wrongly classified as positive; false negatives (FN) are the number of positive samples that the model wrongly classified as negative.

Result

Sixteen channels of 19 main EEG signal channels of 14 schizophrenia patients and 14 healthy controls were selected and divided into 5 frequency bands (delta, theta, alpha, beta, and gamma) because of the complex and non-stationary nature of the EEG signal [42]. Then these frequency regions were embedded in ShEn, ApEn, SpEn, and AVLSAC for extracting features. To calculate ApEn, ShEn, and AVLSAC respectively, the frequency region of each channel was divided into windows of six, four, and ten seconds. With an available signal length of 900 s, we divided them into 150 windows for ApEn, 225 windows for ShEn, and 90 windows for AVLSAC. Then the average of the extracted feature of each window was reported as a feature. To obtain the feature SpEn, the whole signal was used, and according to four feature extraction methods and five frequency regions (the length of each feature vector is 20), a feature vector was calculated for all channels of each person.

Figure 8 shows the accuracy of the ANFIS classifier containing 16 channels with SpEn, ShEn, ApEn, and AVLSAC as feature extraction methods. The range of accuracy shown in this figure is between 86 and 97. Accuracy outputs were calculated in five frequency bands. The best accuracy was obtained 97% for the ANFIS classifier in the alpha channel O1 with the AVLSAC method. The best available features for each channel in 5 frequency regions have been summarized in Table 2. In this table, five frequency regions, and four methods of feature extraction are separated for each of the 16 channels. The check mark in Table 2 indicates the best feature for each channel with the highest accuracy. A total of 46 of 640 possible features for healthy subjects

	Actually positive(1)	Actually negative(0)
Predicted positive (1)	True positives (TP)	False positives (FP)
Predicted negative (0)	False negatives (FN)	True negatives (TN)

Fig. 7 The confusion matrix

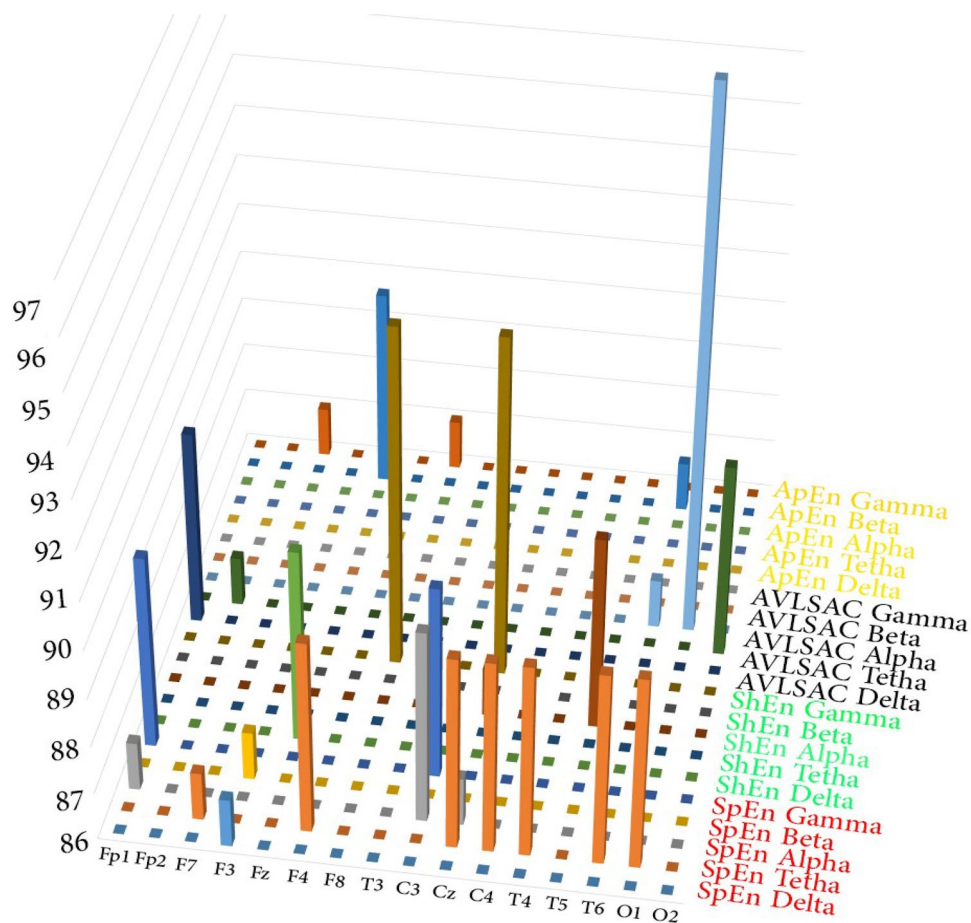


Fig. 8 The accuracy of the ANFIS classifier containing 16 channels with SpEn, ShEn, ApEn, and AVLSAC as feature extraction methods. This figure covers the accuracy range from 86 to 97

and schizophrenic patients were selected. The accuracy of ANFIS, SVM, and ANN classifiers with the best features with five-folds were near 100%, 98.89, and 95.56 respectively.

The highest difference in standard deviation was obtained between the patient and healthy group in the channel O1, which can be seen in Fig. 9. The absolute value of this difference is 1.3, which is related to ALVSCA of alpha. Figure 9 shows a statistical chart of the mean and standard deviation of the channel O1 in 5 sub-frequencies and 4 feature extraction methods. This chart includes the channel O1 of 14 patients and 14 healthy individuals.

Discussion

Today, the use of DSS in the diagnosis of various diseases is relatively common; therefore, recent studies have endeavored to design and present DSS with acceptable accuracy [5–7]. DSS requires signals that contain useful information for proper functioning; for this reason, the EEG signal contains useful information on brain function and disease diagnosis that has been considered by the majority of researchers. The authors in this study attempted to provide an acceptable DSS for schizophrenia, using the information contained in the EEG signal. An ideal classifier should distinguish patients

Table 2 The selection of the best available features for each channel in 5 frequency regions

Ch	Fp1	Fp2	F7	F3	Fz	F4	F8	T3	C3	Cz	C4	T4	T5	T6	O1	O2
Delta																
SpEn	-	-	-	-	-	✓	-	-	-	✓	✓	-	-	-	-	-
ShEn	-	-	-	✓	-	-	-	✓	-	-	-	-	✓	-	-	-
ApEn	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
AVLSAC	-	-	-	-	✓	-	-	-	-	✓	-	-	-	-	-	✓
Theta																
SpEn	-	✓	✓	-	-	-	✓	-	-	-	✓	✓	-	✓	-	-
ShEn	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ApEn	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-
AVLSAC	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓
Alpha																
SpEn	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ShEn	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-
ApEn	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
AVLSAC	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-
Beta																
SpEn	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-
ShEn	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ApEn	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
AVLSAC	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Gamma																
SpEn	✓	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-
ShEn	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	✓
ApEn	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
AVLSAC	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

at a high rate while identifying all healthy controls with no diseases. Therefore, both the selection of the appropriate feature and the classifier must be simultaneous, which will ultimately lead to high accuracy.

The four features SpEn, ApEn, ShEn, and AVLSAC were selected as explained in Section “Feature extraction”. These features show the momentary changes and the complexity of the signal and are compatible with the non-stationary property of data. In addition, the ALVSCA has the ability to present the hidden nature of the signal due to its special property in increasing the frequency resolution. Figure 10 shows the performance of the features in the three classifiers of SVM, ANN, and ANFIS with a dataset for each channel presented in Section “Clinical data”. The accuracy range

shown in this figure is selected between 92 and 97. According to Fig. 10, the most common area involved among the three classifiers is the frontal area in delta sub-frequency and the ALVSCA feature has acceptable performance and the highest accuracy among classifiers.

We compared our results with certain different available methods. Table 3 presents a series of classification accuracy of the ANFIS method and earlier methods with common feature dimensions. It is usually used 5-fold or 10-fold for validation. Figure 11 shows a comparison of the accuracy results obtained with different k values from 5 to 10 in k-fold cross-validation for the three classifiers. These results show us that our method using 5-fold cross-validation obtains the highest classification accuracy (99.92%).

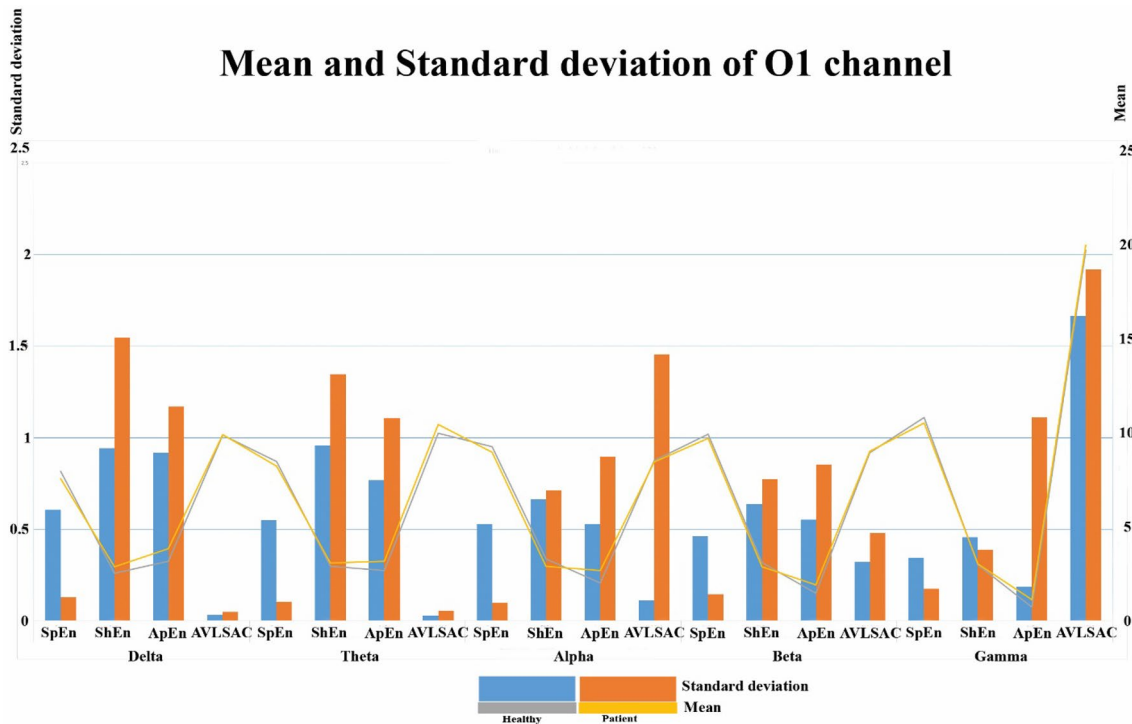


Fig. 9 Statistical chart of the mean and standard deviation of the channel O1 in 5 sub-frequencies and 4 feature extraction methods

Fig. 10 The performance of the features in the three classifiers of SVM, ANN, and ANFIS with a common dataset. The accuracy range between 92 and 97

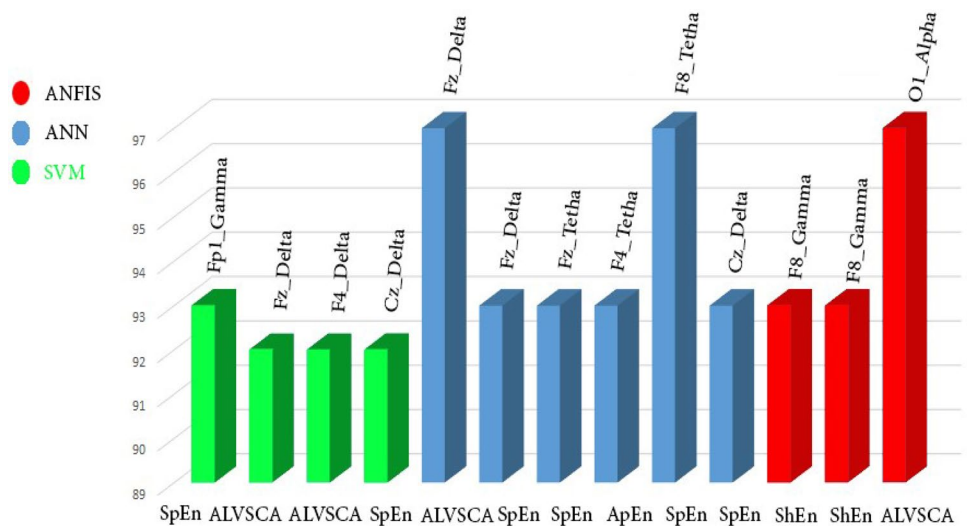


Table 3 The comparison results of the ANFIS accuracy with other available methods

Classifiers	Accuracy	Table 2	Features dimensions
ANFIS [39]	99.92		
SVM [36, 37]	98.89		
ANN [38]	95.59		
DT [43]	89.09		
LDA [44]	87.16		
KNN [45]	85.45		
Adboost [46]	68.25		
ROC [47]	53.37		

Our dataset was previously evaluated by Shu Lih Oh et al. [17] who reported achieving a classification accuracy of 98.07%. Their method was a CNN model that used raw EEG signals as inputs. The method used an eleven-layered CNN model to accurately assess schizophrenic patients versus controls by using the extraction, selection, and classification processes automatically. The previous dataset was also evaluated against a CNN model proposed by ZülfikarAslan et al. [16]. In their method, using 2-D time–frequency features for automatic diagnosis of schizophrenic patients predicts the class value, which obtained 97% accuracy. The second dataset was previously evaluated by Santos-Mayo et al.

[48] who reported obtaining 93.42% classification accuracy, their method was a Multi-Layer Perceptron (MLP) that used EEGlab and J5 as feature extraction methods.

In comparison to these previous literature methods related to our work, our method outperforms the majority of them. Our improved performance seems to be largely due to three methods implemented on raw EEG signals: 1. the use of converting EEGs to sub-frequency regions that explicitly represent the signal frequency information, 2. using the ALVSCA feature extraction method, 3. and selecting the best feature. The ALVSCA seems to be a suitable feature extraction method because it can increase the frequency resolution by using the AR method, which is a powerful algorithm in signal modelling [30]. Table 4 summarizes the relevant literature methods according to the methodology and the accuracy obtained. As a result of the table, it is obvious that the method proposed in this paper outperforms all the methods mentioned in the table.

Conclusion

It is beneficial to present a new decision support system (DSS) as an interactive information system to handle and analyze large volumes of data by allowing for better informed decision makings, solve problems in a timely

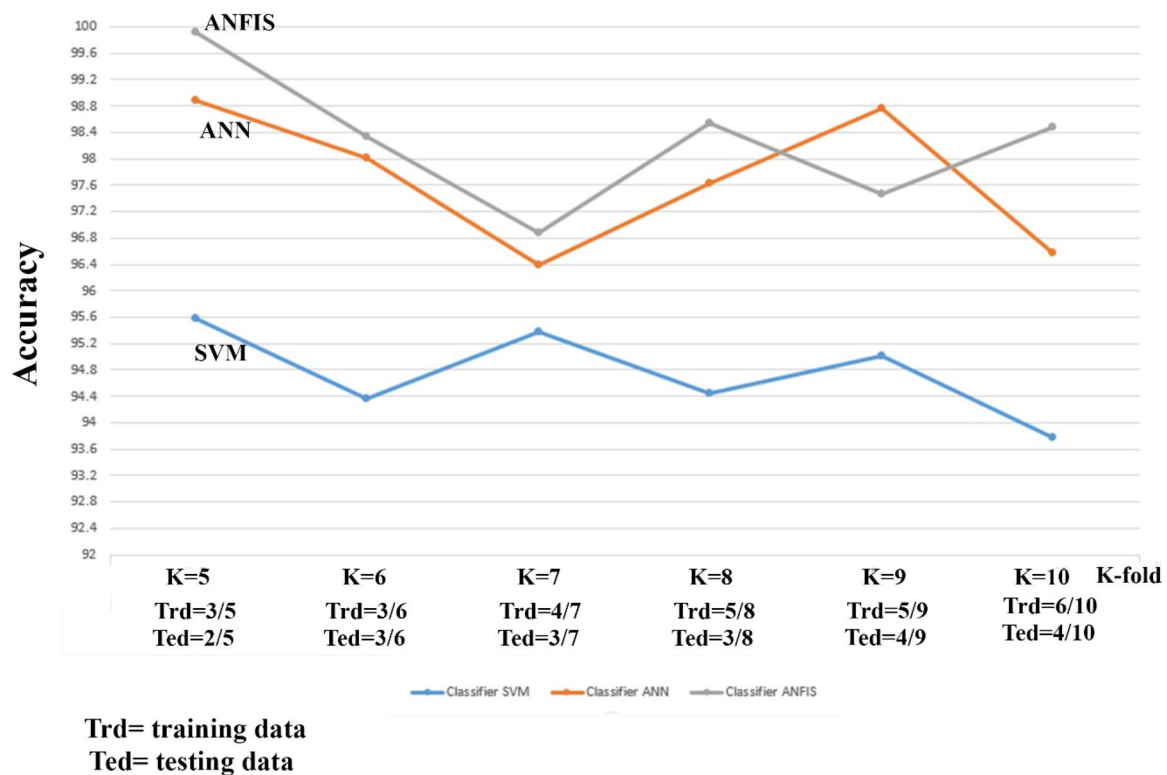
**Fig. 11** The performance of three implemented classifiers with different k values from 5 to 10 in k -fold cross-validation

Table 4 The comparison results of our method and the different methods available on the common dataset

Dataset	Classifiers	Best accuracy	Techniques
Normal: 14 healthy Schizophrenia: 14 patients	ANFIS	99.92	<ul style="list-style-type: none"> ■ Dividing each EEG channel into 5 frequency regions Delta, Theta, Alpha, Beta, and Gamma ■ Non-linear feature extracting ■ Feature selection using best accuracy ■ Classification using ANFIS
	SVM [15]	92.91	<ul style="list-style-type: none"> ■ EEG series splitting ■ Non-linear feature mining ■ Feature selection using t-test ■ Classification using DT, LD, kNN, PNN, SVM classifiers
	CNN [16]	97.00	<ul style="list-style-type: none"> ■ Multi-domain connectome CNN model ■ 2D time and frequency domain connectivity features gauged from EEG signals ■ Feature maps obtained for classification via features fed into CNN
	CNN [17]	98.07	<ul style="list-style-type: none"> ■ 11-layered deep CNN model ■ Subject base testing using 14-fold ■ Non-subject base testing using 10-fold

manner, and improve the efficiency in dealing with problems related to screening schizophrenia patients. In this paper, we present a new computerized information system that can receive a person's EEG signal and distinguish schizophrenic patients and healthy subjects with high accuracy. Individuals who were diagnosed with schizophrenia were referred to a specialist for more accurate management and diagnosis.

Acknowledgements The authors would like to thank the Medical Bioengineering Department, School of Advanced Medical Sciences, Tabriz University of Medical Sciences, Tabriz, Iran.

Author contributions Dr. SHR and Dr. YS conceived of the presented idea and designed the study. Dr.SF supervised the findings of this study. Dr. ME and HN carried out the experiments. All authors discussed the results and contributed to the final manuscript.

Funding This work is partially supported by vice-chancellery for research and technology of Tabriz University of Medical Sciences under Grant No. 62321 and ethical code number IR.TBZMED.REC.1398.545. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical approval This study was performed in accordance with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of Tabriz University of Medical Sciences, Tabriz, Iran (August 27, 2019; REC.1398.545).

Informed consent For this type of study formal consent is not required.

References

1. DSM I (1994) Diagnostic and stabilital manual of mental disorders. APA, Washington
2. Kaplan HI, Sadock BJ (1988) Synopsis of psychiatry: behavioral sciences clinical psychiatry. Williams & Wilkins Co., Philadelphia
3. A. P. Association (2013) Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub, Washington
4. Niedermeyer E (1999) The normal EEG of the waking adult. *Electroencephalogr Basic Princ Clin Appl Relat fields* 20(4):149–173
5. Bascil MS, Oztekin H (2012) A study on hepatitis disease diagnosis using probabilistic neural network. *J Med Syst* 36(3):1603–1606
6. Bascil MS, Temurtas F (2011) A study on hepatitis disease diagnosis using multilayer neural network with Levenberg Marquardt training algorithm. *J Med Syst* 35(3):433–436
7. Er O, Cetin O, Bascil MS, Temurtas F (2016) A Comparative study on Parkinson's disease diagnosis using neural networks and artificial immune system. *J Med Imaging Health Inform* 6(1):264–268
8. Sabeti M, Boostani R, Katebi S, Price G (2007) Selection of relevant features for EEG signal classification of schizophrenic patients. *Biomed Signal Process Control* 2(2):122–134
9. 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
10. Parvinnia E, Sabeti M, Jahromi MZ, Boostani R (2014) Classification of EEG signals using adaptive weighted distance nearest neighbor algorithm. *J King Saud Univ-Comput Inf Sci* 26(1):1–6
11. Sabeti M, Katebi S, Boostani R (2009) Entropy and complexity measures for EEG signal classification of schizophrenic and control participants. *Artif Intell Med* 47(3):263–274
12. Li Y et al (2008) Abnormal EEG complexity in patients with schizophrenia and depression. *Clin Neurophysiol* 119(6):1232–1241
13. Piryatinska A, Darkhovsky B, Kaplan A (2017) Binary classification of multichannel-EEG records based on the ϵ -complexity of continuous vector functions. *Comput Methods Progr Biomed* 152:131–139
14. Alimardani F, Boostani R, Azadehdel M, Ghanizadeh A, Rastegar K (2013) Presenting a new search strategy to select

- synchronization values for classifying bipolar mood disorders from schizophrenic patients. *Eng Appl Artif Intell* 26(2):913–923
15. Jahmunah V et al (2019) Automated detection of schizophrenia using nonlinear signal processing methods. *Artif Intell Med* 100:101698
 16. Aslan Z, Akin M (2020) Automatic detection of schizophrenia by applying deep learning over spectrogram images of EEG signals. *Traitement du Signal* 37(2):235–244
 17. Oh SL, Vinesh J, Ciaccio EJ, Yuvaraj R, Acharya UR (2019) Deep convolutional neural network model for automated diagnosis of schizophrenia using EEG signals. *Appl Sci* 9(14):2870
 18. Hazarika N, Chen JZ, Tsoi AC, Sergejew A (1997) Classification of EEG signals using the wavelet transform. *Signal Process* 59(1):61–72
 19. Kim JW, Lee YS, Han DH, Min KJ, Lee J, Lee K (2015) Diagnostic utility of quantitative EEG in un-medicated schizophrenia. *Neurosci Lett* 589:126–131
 20. Won GH, Kim JW, Choi TY, Lee YS, Min KJ, Seol KH (2018) Theta-phase gamma-amplitude coupling as a neurophysiological marker in neuroleptic-naïve schizophrenia. *Psychiatry Res* 260:406–411
 21. Winterer G et al (2000) Frontal dysfunction in schizophrenia—a new electrophysiological classifier for research and clinical applications. *Eur Arch Psychiatry Clin Neurosci* 250(4):207–214
 22. Laton J et al (2014) Single-subject classification of schizophrenia patients based on a combination of oddball and mismatch evoked potential paradigms. *J Neurol Sci* 347(1–2):262–267
 23. 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
 24. Taylor JA, Matthews N, Michie PT, Rosa MJ, Garrido MI (2017) Auditory prediction errors as individual biomarkers of schizophrenia. *NeuroImage Clin* 15:264–273
 25. Olejarczyk E, Jernajczyk W (2017) Graph-based analysis of brain connectivity in schizophrenia. *PloS One* 12(11):e0188629
 26. Kim D-J et al (2000) An estimation of the first positive Lyapunov exponent of the EEG in patients with schizophrenia. *Psychiatry Res Neuroimaging* 98(3):177–189
 27. Pressman A, Peled A, Geva A (2000) Synchronization analysis of multi-channel EEG of schizophrenic during working-memory tasks. In: 21st IEEE convention of the electrical and electronic engineers in Israel. Proceedings (Cat. No. 00EX377), IEEE, New York, pp 337–341
 28. Bruhn J, Röpcke H, Hoefl A (2000) Approximate entropy as an electroencephalographic measure of anesthetic drug effect during desflurane anesthesia. *Anesthesiology* 92(3):715–726
 29. Inouye T et al (1991) Quantification of EEG irregularity by use of the entropy of the power spectrum. *Electroencephalogr Clin Neurophysiol* 79(3):204–210
 30. Stoica P, Moses RL (1997) Introduction to spectral analysis. Prentice hall Upper Saddle River, NJ, New Jersey
 31. Kay S (1980) Noise compensation for autoregressive spectral estimates. *IEEE Trans Acoust Speech Signal Process* 28(3):292–303
 32. Annis J, Palmeri TJ (2018) Bayesian statistical approaches to evaluating cognitive models. *Wiley Interdiscip Rev Cogn Sci* 9(2):e1458
 33. Petter S, DeLone W, McLean E (2008) Measuring information systems success: models, dimensions, measures, and interrelationships. *Eur J Inf Syst* 17(3):236–263
 34. Myung IJ (2003) Tutorial on maximum likelihood estimation. *J Math Psychol* 47(1):90–100
 35. Chen J-F, Wang W-M, Huang C-M (1995) Analysis of an adaptive time-series autoregressive moving-average (ARMA) model for short-term load forecasting. *Electr Power Syst Res* 34(3):187–196
 36. Bishop CM (1995) Neural networks for pattern recognition. Oxford University Press, Oxford
 37. Kavzoğlu T, Çölkesen İ (2010) Destek vektör makineleri ile uydü görüntülerinin sınıflandırılmasında kernel fonksiyonlarının etkilerinin incelenmesi. *Harita Dergisi* 144(7):73–82
 38. Acır N, Güzelış C (2004) Automatic recognition of sleep spindles in EEG by using artificial neural networks. *Expert Syst Appl* 27(3):451–458
 39. Jang J-S (1993) ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans Syst Man Cybern* 23(3):665–685
 40. Medsker L, Jain LC (1999) Recurrent neural networks: design and applications. CRC Press, Boca Raton
 41. Khan S, Naseem I, Malik MA, Togneri R, Bennamoun M (2018) A fractional gradient descent-based rbf neural network. *Circuits Syst Signal Process* 37(12):5311–5332
 42. Andreas G (2000) Topics in nonlinear time series analysis, with implications for EEG analysis. World Scientific, Singapore
 43. Safavian SR, Landgrebe D (1991) A survey of decision tree classifier methodology. *IEEE Trans Syst Man Cybern* 21(3):660–674
 44. Mika S, Ratsch G, Weston J, Scholkopf B, K.-R. Mullers K-R (1999), Fisher discriminant analysis with kernels. In: neural networks for signal processing IX: proceedings of the 1999 IEEE signal processing society workshop (cat. no. 98th8468), IEEE, New York, pp 41–48
 45. Tan S (2006) An effective refinement strategy for KNN text classifier. *Expert Syst Appl* 30(2):290–298
 46. Vezhnevets A, Vezhnevets V (2005) Modest AdaBoost-teaching AdaBoost to generalize better. *Graphicon* 12(5):987–997
 47. Fawcett T (2006) An introduction to ROC analysis. *Pattern Recognit Lett* 27(8):861–874
 48. Santos-Mayo L, San-José-Revuelta LM, Arribas JI (2016) A computer-aided diagnosis system with EEG based on the P3b wave during an auditory odd-ball task in schizophrenia. *IEEE Trans Biomed Eng* 64(2):395–407

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