

EEG classification using recurrent adaptive neuro-fuzzy network based on time-series prediction

Hossein Komijani¹ · Mohammad Reza Parsaei² · Ebrahim Khajeh³ ·
Mohammad Javad Golkar⁴ · Houman Zarrabi⁵

Received: 8 November 2016 / Accepted: 4 October 2017 / Published online: 4 November 2017
© The Natural Computing Applications Forum 2017

Abstract Brain–computer interface (BCI) is a system that provides a way for brain and computer to communicate with each other directly. Electroencephalogram (EEG) is an important process in a BCI that can be used to determine whether the subject is doing action and/or imagination. This paper presents a motor imagery (MI) classification for BCI systems using recurrent adaptive neuro-fuzzy interface system (ANFIS). The classification system is based on time-series prediction where features are exploited from the EEG signals recorded from subjects imagining of the right hand, left hand, tongue, and foot movement. The classification system contains some recurrent ANFISes. Each recurrent ANFIS is trained on MI signals of one class and specializes in recognizing the signals of the same class from the signals of other categories. Recurrent ANFISes are employed to predict one step ahead for the EEG time-series data, and then, the classification is performed by mean square error (MSE) of the predicted signals. This approach is carried out on twelve subjects MI signals of four classes in online mode. Average prediction MSE of 0.0302 and average classification accuracy of 85.52% are obtained as results.

Keywords Classification · Electroencephalography (EEG) · Recurrent adaptive neuro-fuzzy interface system (recurrent ANFIS) · Brain–computer interface (BCI) · Motor imagery (MI) · Time-series prediction

1 Introduction

In recent years, the studies on the subjects of the brain are increased dramatically. These wide studies provide a new system for communicating brain and computer just with electroencephalogram (EEG) processing. This new approach is named brain–computer interface (BCI). BCI does not rely on muscular and neuromuscular inputs. In a BCI, the use of nerves, muscles, and movements is replaced with electrophysiological signals processing by means of some software and hardware [1]. This advantage enables the subjects with muscular and neuromuscular disabilities to communicate and control their ambience without any use of prerequisite movements. Therefore, BCIs can amend existing assistive technologies. Using BCIs has advantages for handicapped subjects [2] and applications such as neuro-feedback for stroke rehabilitation [3], medical usage such as sleep disorders heeling [6], epileptic seizure prediction [4, 7], alertness detection and cognitive load monitoring [8]. BCIs are also profitable as an additional technology in robotics [9], computer games [5], virtual reality [10], etc.

EEG signals of various motor imagery (MI) classes have distinct event-related brain potentials (ERP) which can further have specific attributes such as event-related desynchronization (ERD) which manifest a decrease in power, and event-related synchronization (ERS) which manifest an increase in power in mu and beta rhythms over the sensorimotor cortex [11].

✉ Hossein Komijani
h.komijani90@ms.tabrizu.ac.ir

¹ Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran

² Department of Engineering and IT, Shiraz University of Technology, Shiraz, Iran

³ Faculty of Computing, Universiti Teknologi Malaysia, Skudai, Johor, Malaysia

⁴ Department of Electrical and Computer Engineering, Zahedan Branch, Islamic Azad University, Zahedan, Iran

⁵ ICT Research Institute, Tehran, Iran

Classification is an important computational procedure in a BCI. A two-class EEG-based classification classifies fatigue state versus alert state using a combination of independent component by entropy rate-bound minimization analysis (ERBM-ICA) for the source separation, autoregressive (AR) modeling for the features extraction, and Bayesian neural network as the classifier algorithm is presented in [15]. An optimal probe selection for EEG classification in a BCI for three mental tasks using blind source separation (BSS) technique based on independent component analysis (ICA) with its back-projecting of the scalp map is presented in [14]. As a result, the four channels of P3, O1, C4, and O2 are cited as the prominent channels with more scattered features. EEG classification of pre- and post-mental load tasks for mental fatigue detection by the classification system combined of principal component analysis (PCA) as the dimension reduction method, power spectral density (PSD) as feature extractor, and Bayesian neural network (BNN) as classifier is presented in [16]. An EEG classification system for fatigue detection includes PSD as a feature extractor, fuzzy swarm-based artificial neural network (ANN) as a classifier, and component analysis of entropy rate-bound minimization (ICA-ERBM) for source separation technique, is presented in [17]. EEG classification is also studied for fatigue detection [15–17] and user identification [18–20] as biometric systems.

This paper presents a novel EEG classification procedure based on time-series prediction. A new class of recurrent adaptive neuro-fuzzy interface system (ANFIS) is developed for predicting EEG time series. Autoregressive moving average (ARMA) prediction models, typical feedforward ANFIS, and ANN are also used for prediction approach in a similar framework to have a comparison between the developed recurrent ANFIS and the existing systems.

ANNs are widely used as approaches for pattern recognition including MI classification [11–13]. Fuzzy systems can describe objects and processes that are not defined precisely or have some uncertainty in their description. Fuzzy systems perform a valuable role in handling with uncertainty. Therefore, fuzzy systems have a growing interest in modern applications such as production techniques medical studies, information technology, pattern recognition, decision making, data analysis, and diagnostics [3–5].

Neuro-fuzzy systems are a combination of fuzzy systems and ANNs and utilize the benefits of them concurrently. Adaptive neuro-fuzzy inference system (ANFIS) is a significant approach in neuro-fuzzy networks which has shown notable potential in modeling nonlinear systems. In ANFIS, the parameters of fuzzy membership functions are gained from a dataset defining the system's manner [9, 10].

Successful executions of ANFIS in biomedical applications have been announced such as EEG classification [21–25] and data analysis in biosignal processing [2, 26].

In this paper, a new recurrent structure based on ANFIS is performed for EEG classification. The presented approach that involves eight recurrent ANFISes classifies four classes of EEG signals when raw EEG data in the time domain are used as inputs. These recurrent ANFISes are trained with back-propagation gradient descent method in composition with least squares method. After the training phase, each recurrent ANFIS will be specialized in sort of signals which is trained with and can recognize the signals of the same class from the signals of other classes.

2 Experiment datasets

The dataset IIIa from BCI Competition III and the dataset 2a from BCI-IV are used in this paper [28, 29]. The dataset IIIa was recorded from three healthy subjects (S1–S3), and nine healthy subjects (S4–S12) were recorded in dataset 2a. These datasets consisted of four classes: left-hand movement imagination (class 1), right-hand movement imagination (class 2), foot movement imagination (class 3), and tongue movement imagination (class 4).

For dataset IIIa, signals were recorded in ten sessions for each subject. Subjects S1 fulfilled 360 trials, and subjects S2 and S3 fulfilled 240 trials [29, 30]. The signals for datasets 2a were recorded in two sessions on various days for each subject. Each session includes six recording parts which is cut off by short breathers. Each recording part contained 48 trials (12 trials for each class), leading to 288 trials per session and totally 576 trials.

Figure 1 shows the timing diagram of the routine of recording signals [27–30]. The routine of recording signals started with sitting the subjects in a comfortable armchair. Then, a cue arrow was shown on the screen, and the subjects were asked to imagine a movement according to the orientation of the arrow. The orientation of the arrow was selected randomly, and no feedback to the subjects was provided. The sampling frequency was 250 Hz. EEG in

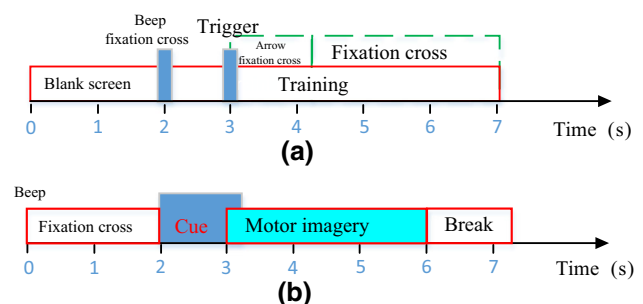


Fig. 1 Timing of the paradigm **a** dataset IIIa, **b** dataset 2a [28, 29]

dataset IIIa was filtered with a 1–50-Hz band-pass filter and 50-Hz notch filter on. EEG in dataset 2a was filtered with a 0.5–100-Hz band-pass filter and an additional 50-Hz notch filter to remove the line noise [27–30]. The EEG recording systems for the datasets IIIa and 2a were a sixty-four EEG-channel amplifier from Neuroscan and a twenty-two Ag/AgCl electrodes system, respectively, as shown in Fig. 2. The recording systems use the right mastoid as ground and the left mastoid as a Ref. [28, 29].

The two electrodes, named C3 and C4, are used in this paper. These electrodes are, respectively, placed over the left and right sensorimotor areas which are dominantly the main active areas in the imagination of motions [31]. These locations are signalized in Fig. 2. Usually, there is a pre-processing before EEG classification for artifacts and noise removal [32–34]. In this paper, the signals that highly affected by the artifacts are removed from the database by an expert, and no other artifact removal procedure is used.

3 Recurrent ANFIS architecture

ANFIS is a kind of neural network in a combination of Takagi–Sugeno fuzzy inference system. As ANFIS amalgamates both neural networks and fuzzy systems, it has the ability to acquire the advantages of both in a single structure [9]. ANFIS has pioneering results in recent years of neuro-fuzzy networks. Indeed, it is also considered to be one of the best systems in function approximation among the existing neuro-fuzzy models [35]. To express the ANFIS operating process, two first-order Takagi–Sugeno fuzzy if–then rules are assumed as:

$$\begin{aligned} \text{Rule 1: if } x \text{ is } A_1 \text{ and } y \text{ is } B_1 \text{ then } f_1 &= p_1x + q_1y + r_1 \\ \text{Rule 2: if } x \text{ is } A_2 \text{ and } y \text{ is } B_2 \text{ then } f_2 &= p_2x + q_2y + r_2 \end{aligned} \tag{1}$$

where x and y are the inputs, A_i and B_i are the fuzzy sets, f_i are the outputs within the fuzzy region designated by the fuzzy rules, and p_i , q_i , and r_i are the design parameters designated within the training phase. Figure 3 shows the ANFIS architecture to execute these two if–then rules. This network consists of five layers. If the output of node i of layer L is shown by $o_{L,i}$, the function of this network is described as follows:

The first layer calculates the degrees of inputs membership in fuzzy membership functions:

$$\begin{aligned} o_{1,i} &= \mu_{A_i}(x) \quad i = 1, 2 \\ o_{1,i} &= \mu_{B_{i-2}}(y) \quad i = 3, 4 \end{aligned} \tag{2}$$

where $\mu_{A_i}(x)$ and $\mu_{B_{i-2}}(y)$ are the fuzzy membership functions. For example, by employing Gaussian function, $\mu_{A_i}(x)$ is written as:

$$\mu_{A_i}(x) = e^{-\left(\frac{x-c_i}{\sigma_i}\right)^2} \tag{3}$$

where c_i and σ_i are the membership function parameters computed in training procedure.

The output of the second layer is gained with multiplying the output values of the first layer:

$$o_{2,i} = w_i = \mu_{A_i}(x) \cdot \mu_{B_i}(y) \quad i = 1, 2. \tag{4}$$

The output of the third layer is calculated by normalizing the output values of the second layer:

$$o_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad i = 1, 2. \tag{5}$$

The output of the fourth layer is computed by multiplying the output values of the previous layer with a first-order polynomial. For the first-order Takagi–Sugeno fuzzy if–then rules, it can be calculated as:

$$o_{4,i} = \bar{w}_i f_i = \bar{w}_i(p_i x + q_i y + r_i) \quad i = 1, 2 \tag{6}$$

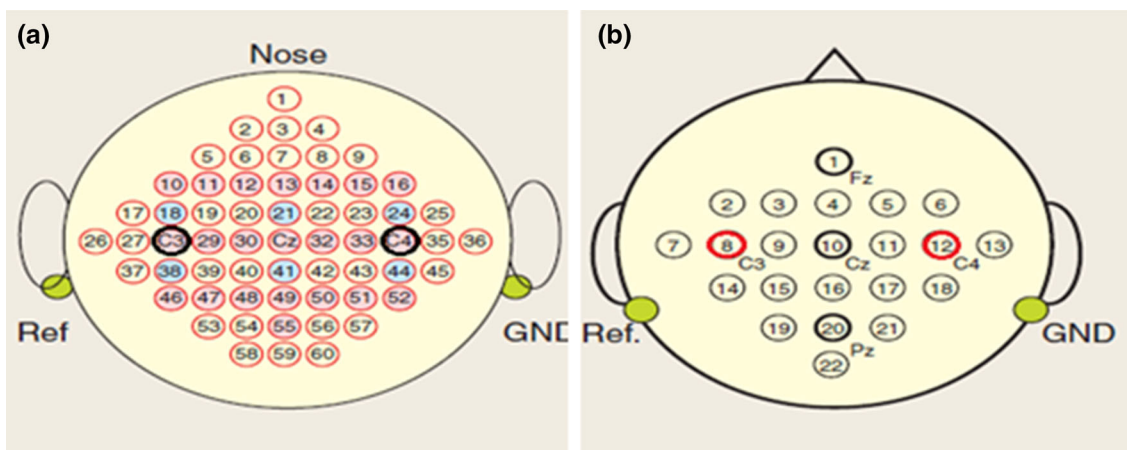
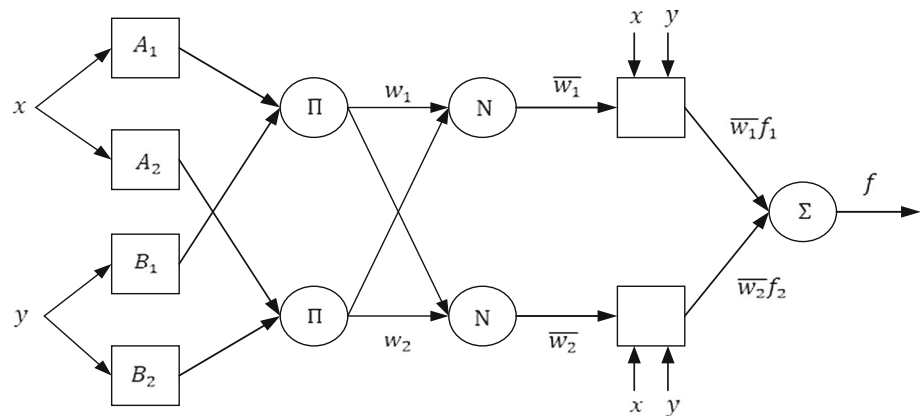


Fig. 2 Position of EEG electrodes a dataset IIIa, b dataset 2a [28, 29]

Fig. 3 ANFIS architecture

Finally, the output of network is calculated in layer five as:

$$o_{5,i} = o = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad i = 1, 2 \quad (7)$$

where o is the output of network, and \bar{w}_i is the output of third layer.

The architecture of ANFIS implementing these two rules is described in details in [9].

However, a major deficiency of available neuro-fuzzy networks is that their efficiency is limited to static problems because of their feedforward structure. Consequently, they have less efficiency for representing dynamic functions in comparison with recurrent networks [36–38]. By upgrading typical feedforward ANFIS to recurrent type, its ability to handle time-series patterns recognition will be promoted. A considerable issue in this structure evolution is that the recurrent structure shows greater prediction capacities compared with the feedforward ones. Therefore, in this paper, time-series prediction is performed by an altered structure of ANFIS with output-to-input feedback loop.

For converting the network architecture to the recurrent type, the error of prediction is exerted to the network as an input. Indeed, the prediction error is feedback to the network, and by doing so, a closed-loop system is acquired. This architecture is shown in Fig. 4. If the output of typical ANFIS is considered as a function of inputs: $o = f(x, y)$, the output of this recurrent ANFIS will be changed to: $o = f(x, y, e)$. By applying this change, the output of network which was the function of inputs will become a function of inputs and errors of previous estimations.

There are two groups of parameters that should be calculated to match the ANFIS output to the training data: fuzzy membership function parameters called premise parameters and linear parameters called design parameters. The least squares and the gradient descent methods are

integrated into a hybrid algorithm to calculate these parameters for training the network. The hybrid algorithm includes backward pass and forward pass. The least squares method in forward pass is employed to adjust the design parameter values optimally when the premise parameters are fixed. When the optimum design parameters are defined, the back-propagation gradient descent method in the backward pass is employed to determine the optimal premise parameter values [9].

There is a difference in applying the network in the training and in the test phases. The error in the training phase is different with the error in the test phases. This problem is explained as follows:

Consider the consequence of inputs in the training phase to be as (x_1, y_1, e_1) and (x_2, y_2, e_2) and the outputs as o_1 and o_2 , respectively, and in test phase, the inputs consequence to be as (x'_1, y'_1, e'_1) and (x'_2, y'_2, e'_2) and the outputs as o'_1 and o'_2 , respectively. If $x_2 = x'_2$ and $y_2 = y'_2$, it is assumed that $o_2 = o'_2$, but if $x_1 \neq x'_1$ and/or $y_1 \neq y'_1$, as a result $o_1 \neq o'_1$, and $e_2 \neq e'_2$ (because the errors are affected by the previous level). As a result, although $x_2 = x'_2$ and $y_2 = y'_2$, because of $e_2 \neq e'_2$ the outputs $o_2 \neq o'_2$. Therefore, there is a problem in time-series prediction in recurrent form. To resolve this problem, an error estimation system (EES) is used in test phase. EES is trained with the network training data: x and y as inputs and e as output, and then, in level i in test phase, EES can find the proper e_i (e_i is affected by $i - 1$ level in training phase) according to x_i and y_i .

The network can have a good performance of prediction in test phase without using the EES, but by using EES, the error of prediction is decreased, whereas the classification is based on MSE, as a result, the classification accuracy (CA) is increased. The network is configured according to Fig. 4 in the training phase. In the test phase, the e input is gained by EES, and the whole network is configured according to Fig. 5 in the test phase.

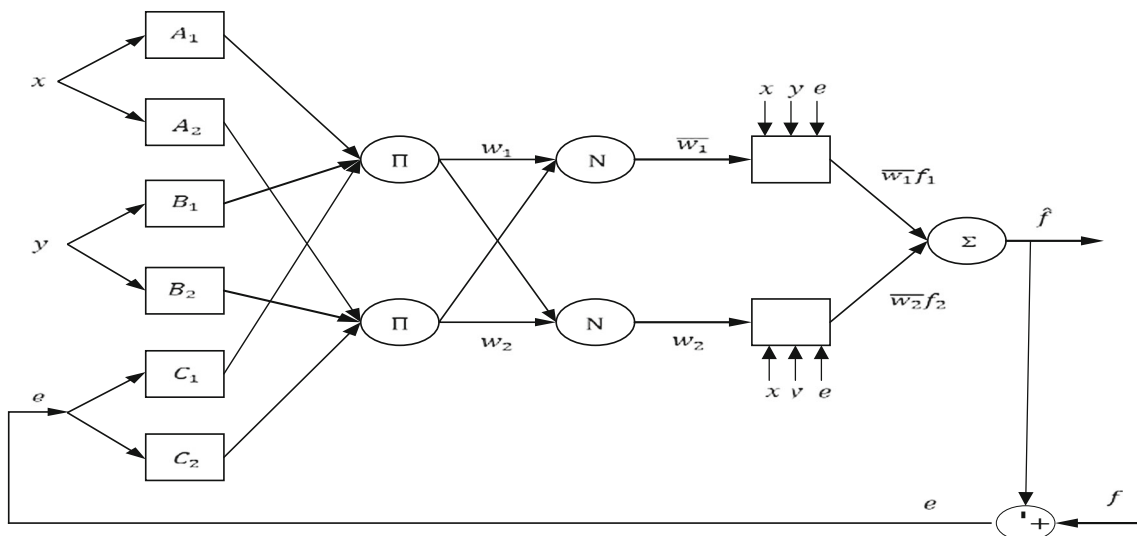
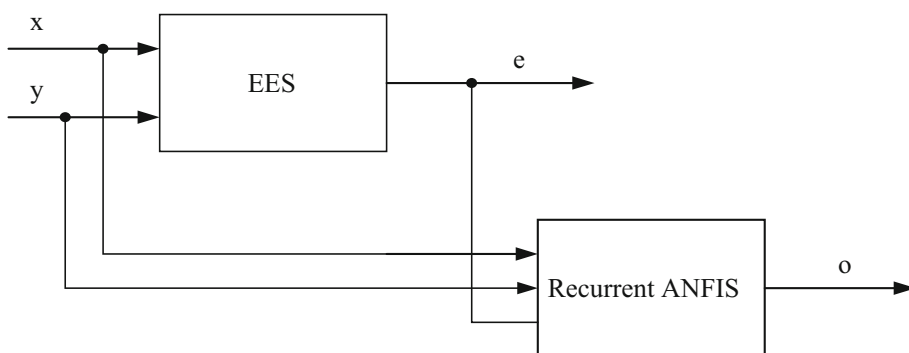


Fig. 4 Recurrent ANFIS architecture

Fig. 5 Recurrent ANFIS predictor architecture



4 Methodology

4.1 Network configuration

In [39, 40], feedforward neural network is used for EEG time-series prediction and EEG classification. By exploiting feedforward network, the previous value of the signal is used to predict the future value of time series. Therefore, the signal values from sample indexes $t - 1$ to $t - n$ are utilized to predict the signal value at sample index t , so it can be written as:

$$\hat{x}(t) = f[x(t - 1), x(t - 2), \dots, x(t - n)] \tag{8}$$

where $x(t)$ is the signal value at sample index t , $\hat{x}(t)$ is the predicted value of signal at sample index t , and f is a feedforward network function that predicts $\hat{x}(t)$ from prior values of signal.

By using the errors of former predictions, the power of estimation is improved, and the errors of prediction in the next steps of prediction will be diminished; hence, a superior prediction will be done. By this consideration, $\hat{x}(t)$

has become a function of prior values of signal and the previous values of prediction error:

$$\hat{x}(t) = f[x(t - 1), \dots, x(t - n), e(t - 1), \dots, e(t - m)] \tag{9}$$

where $e(t)$ is the error of prediction at sample index t which is obtained as:

$$e(t) = x(t) - \hat{x}(t) \tag{10}$$

n and m values are chosen by empirical approach to gain the high and acceptable CA. In the beginning, they were chosen with the smaller values, and they were increased step by step to reach to a high classification accuracy. In this paper, m and n are set as $n = 3$ and $m = 3$. By using these values of parameters, the high accuracy of estimation is acquired. Of course, by choosing higher values of n and m , the higher classification accuracy is gained, but the training time is increased dramatically (typically, the neuro-fuzzy networks training is time-consuming), i.e., having two fuzzy membership functions for each input, and increasing just one unit of each parameter, the fuzzy if-then rules are increased twofold causing the twofold amount of computation. In this work, we try to have an

acceptable classification accuracy by choosing the minimum values of n and m .

Choosing $m = 3$ and $n = 3$, (9) is converted to (11); therefore, the signal and error values from time instants $t - 3$ to $t - 1$ are used to predict the measurement at time t :

$$\hat{x}(t) = f[x(t-1), x(t-2), x(t-3), e(t-1), e(t-2), e(t-3)] \quad (11)$$

To implement (11) for EEG time-series prediction, a network is configured in the training phase according to Fig. 6. After training phase and acquiring the values of prediction error, the EES is trained by $x(t-1)$, $x(t-2)$, and $x(t-3)$ as the inputs and $e(t-1)$, $e(t-2)$, and $e(t-3)$ as the outputs. Then, in the test phase, the final system to do the time-series prediction for the unknown signal is configured according to Fig. 7. However, without using EES, the prediction can be done in a good mode, but by exploiting EES, the power of prediction is increased that influences CA properly. In this paper, the EES is a two-layer MLP with 8 hidden neurons in the hidden layer.

Training just one predictor on all EEG signals of different classes is not ideal consistently because of the complexity and nonstationary characteristics of EEG data on the signals of various classes. Based on the supposition underlying the neural time-series prediction framework, if more than one channel is used for signals of each class, useful supplemental information pertinent to the differences of signals of each class can be extracted to enhance the separability of the overall features, in consequence amending BCI efficiency [40].

The number of EEG channels and the number of classes specify the number of predictor networks as:

$$N = C \times E \quad (12)$$

where N is the number of predictor networks, E is the number of selected EEG channels, and C is the number of classes. Two electrodes of C3 and C4 are chosen, and four classes of right-hand movement, left-hand movement, tongue movement, and foot movement exist, so eight networks are employed as predictors for EEG time-series data, and each network is trained corresponding to one-channel-class time-series EEG data.

4.2 Classification procedure

The classification procedure is configured in three phases. The first phase is configuration and training eight recurrent ANFISes separately to execute one-step-ahead prediction, using the sequence of $[x(t-1), x(t-2), x(t-3), e(t-1), e(t-2), e(t-3)]$ for each channel-class time-series EEG data as described in the previous section (Sect. 4.1). This phase (the training phase of the networks) is presented in Fig. 8.

After the training stage, the networks are ready to perform classification of test signals for unlabeled signals in two other phases as shown in Fig. 9. In Fig. 9, the recurrent ANFISes are labeled 'C3L' and 'C4L' for left-hand data-electrode 'C3' and left-hand data-electrode 'C4,' respectively, and 'C3R' and 'C4R' for right-hand data-electrode 'C3' and right-hand data-electrode 'C4,' respectively, and 'C3F' and 'C4F' for foot data-electrode 'C3' and foot data-electrode 'C4,' respectively, and 'C3T' and 'C4T' for tongue data-electrode 'C3' and tongue data-electrode 'C4,' respectively, corresponding to the type of each channel-class time-series EEG data on which they are trained with.

In the second stage, according to Fig. 9, to find the class of the unknown signal, the unknown signals recorded from

Fig. 6 Illustration of recurrent ANFIS architecture for EEG time-series prediction

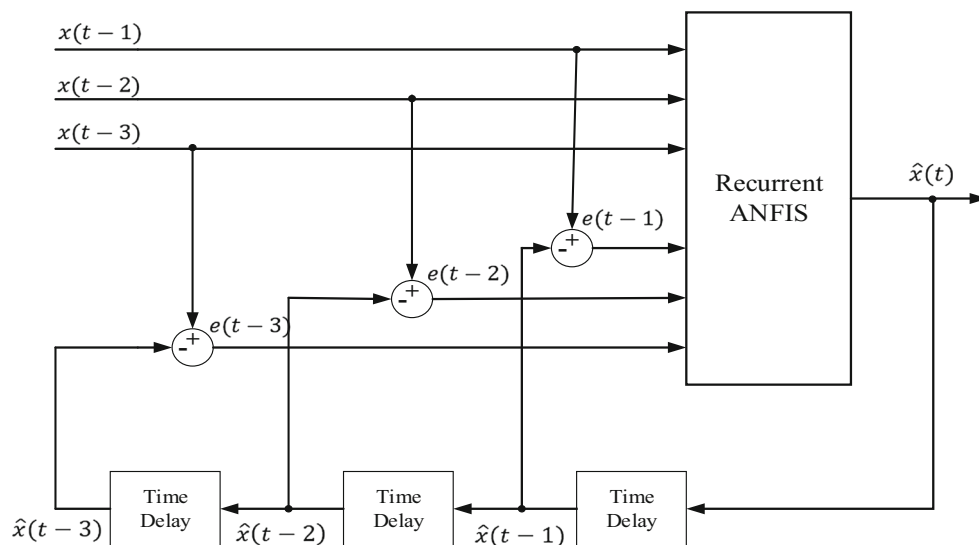


Fig. 7 Recurrent ANFIS predictor architecture for EEG time-series prediction

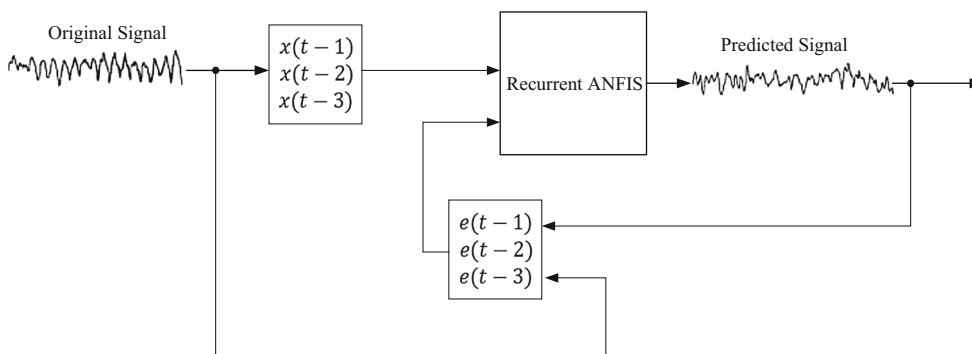
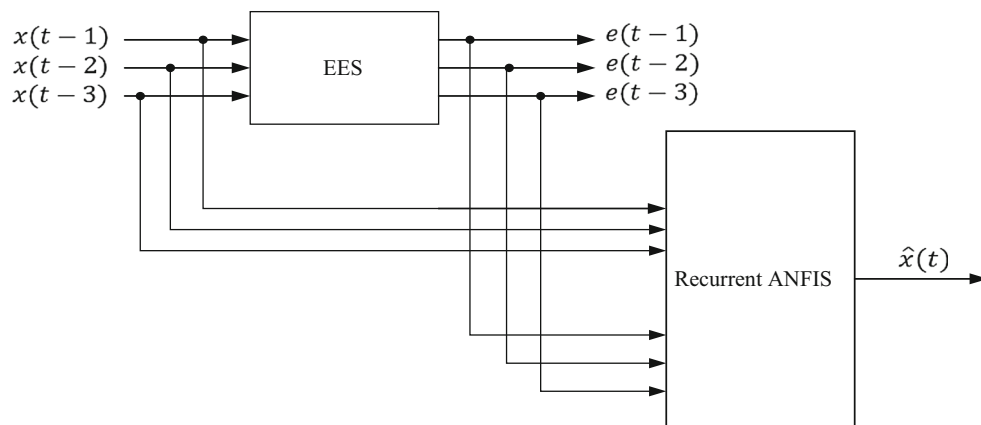


Fig. 8 Classification procedure in the training phase

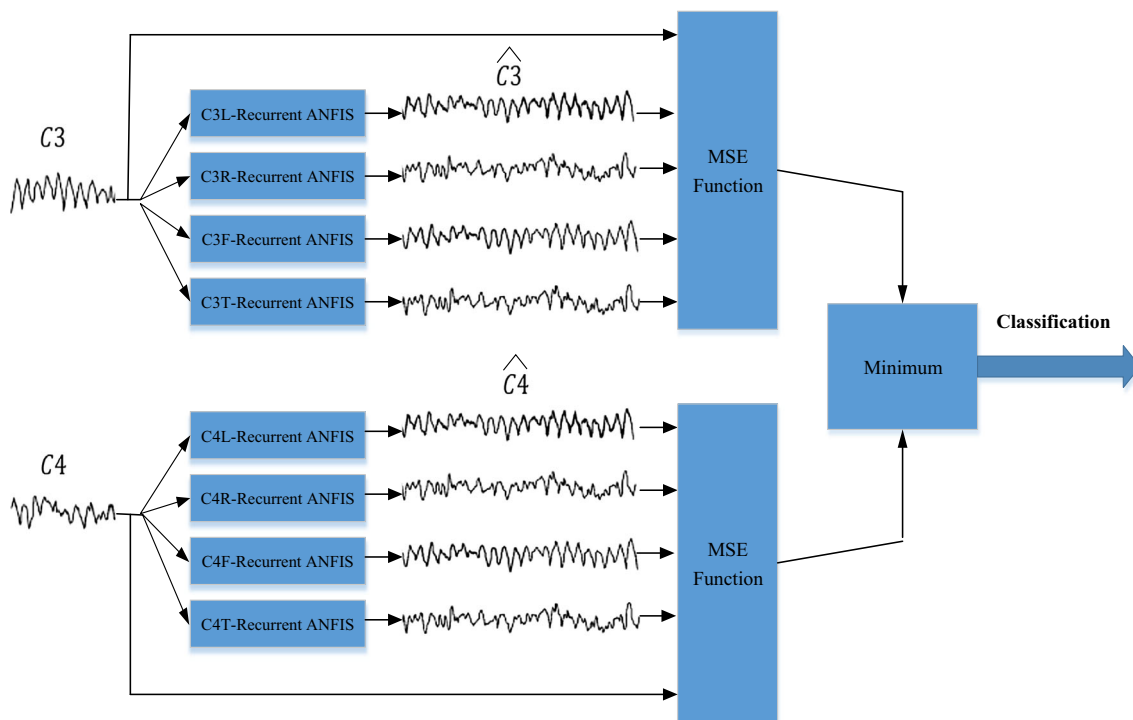


Fig. 9 Classification procedure in the test phase

C3 to C4 electrodes are entered to the relevant networks (i.e., recorded signal of C3 is entered to all networks labeled with C3, and symmetric behavior is done for recorded signal of C4). Each network performs one-step-forward prediction for the input signal in each trial. Then, in the third stage, the MSE of the prediction according to the difference of the original test signal and the predicted signal is calculated for the trial as follows:

$$\text{MSE} = \frac{1}{M} \sum_{k=1}^M [x(k) - \hat{x}(k)]^2 \quad (13)$$

where $x(k)$ is the values of actual signal and $\hat{x}(k)$ is the values of predicted signal, k index is the counter of the length of trial that increases from 1 to M , and M is the number of samples in the input trial. Then, the unlabeled signal is classified in agreement with the network which performs lowest MSE of prediction. The MSE is a criteria of the difference of the original test signal and the predicted signal. The supposition underling this classification procedure is that because each network is trained on signals of just one class, each network can predict the signals which are the same as its training signals with low error, but the other networks have the larger error on prediction of that kind of signals, because the structure of their training signals was different with that kind of signals. Thus, the network that has the same training signal with the test signal can perform the prediction with lower error than the other networks and can identify the similarity of the test signal with the training signals. Therefore, the class of test signal is specified according to the class of the network which has the minimum error of prediction. In this paper, the common feature extraction methods are not used. Since the classification is based on time-series prediction, each network shows the lower prediction error has chosen as the winner, and the class of test signal is selected similar to the network with lower prediction error. For example, if the C3T network shows the lower prediction error, the class of signal is tongue movement. In other words, the errors of prediction (there are 8 networks, so there are 8 prediction errors) can be assumed as the features and the minimum function as the classifier.

This procedure can be used for every trial time length utilizing a sliding window method. To do so, k in (13) ranges from $k = \alpha$ to M where α and M are specified before the next classification procedure (initially $k = 1$ and $M = \text{window size}$). It means that the previous samples placed before the index $k = \alpha$ are discarded as the window slides away from the index $k = \alpha$ to the end of trial. The benefit of utilizing the sliding window for classification is that there is no need to know the start point of the communication by the user; therefore, online classification can be configured. An important point in this regard is selecting

the optimal window size. By selecting window size properly, maximum CA is obtained as the signal pass through the window. In [39], the optimal window is appointed by means of an automated iterative search procedure. The optimal window size ranges between 2.3 and 3.2 s, and this means that the optimal M with considering the sampling frequency of 250 Hz is appointed in interval of [575, 800].

4.3 The comparison between the presented recurrent ANFIS and the other similar predictor systems like ANFIS, MLP, and ARMA

A comparison between the presented recurrent ANFIS and the other similar predictor systems like typical feedforward ANFIS, MLP, and ARMA model is considered. A parametric model named ARMA model consists of autoregressive (AR) part, and a moving average (MA) part is described in the following:

$$\hat{x}(t) = \sum_{i=1}^m a_i x(t-i) + \sum_{j=0}^n b_j e(t-j) \quad (14)$$

where a_i and b_j are the AR and MA parameters [41, 42].

For comparison, the classification is done as the same method described in Sect. 4.2. All stages of classification are the same, but only the time-series prediction system in our work which is a recurrent ANFIS is replaced by MLP, feedforward ANFIS, and ARMA model. These similar systems predict the EEG time series in a similar framework, and then, based on the errors of prediction, the CA of each one is gained. The comparison is made by setting $m = 3$ and $n = 3$ in (14) for ARMA model, setting $n = 6$ in (8) for the MLP, and setting $n = 6$ in (8) for the feedforward ANFIS. So, the MLP and feedforward ANFIS have 6 values of EEG signals as the inputs, and the ARMA model has 3 values of EEG signals and 3 values of prediction errors as the inputs. The MLP utilized here has one hidden layer with 8 neurons in the hidden layer, and the feedforward ANFIS has two membership functions in the first layer.

5 Results and discussion

The proficiency of the presented procedure is demonstrated in two parts: the effectiveness of the presented recurrent ANFIS in the prediction of time series applied on EEG signals and the CA of the classifier based on time-series prediction. The accuracy of proposed classification procedure is assayed by using a fivefold cross-validation.

Figure 10 shows an instance of the efficiency of the recurrent ANFIS in predicting EEG time-series data. According to Fig. 10, the proposed network can model the

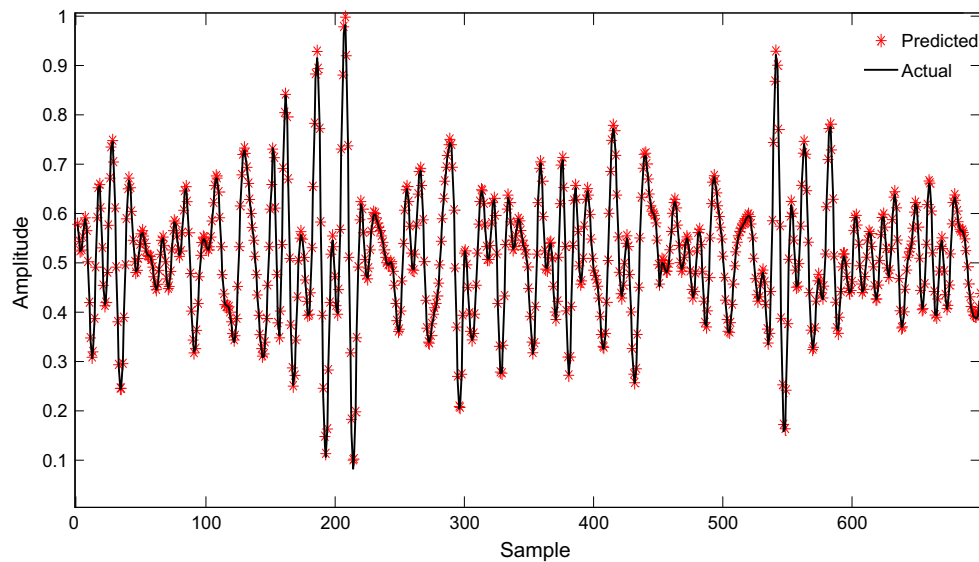


Fig. 10 Predicted signal versus actual EEG

EEG signal with negligible error. Although EEG signals are highly nonstationary and time variant, the network can track it fittingly.

The comparison of MSE of the proposed network, ARMA model, the MLP-ANN, and the feedforward ANFIS is presented in Table 1. In Tables 1 and 2, subjects (S1–S3) are from dataset IIIa and subjects (S4–S12) are from dataset 2a. Based on Table 1, it is obvious that our suggested network has smaller MSE in comparison with the other methods. The classification system acts based on the MSE of each network, and the unlabeled signal will be classified according to the network with lowest MSE, so the accuracy of the classifier is enhanced by the system whoever can exploit the time-series prediction with lower MSE. According to Table 1, the MSE of recurrent ANFIS is lower than the others, and it means

that the power of identifying the structure of EEG signals by recurrent ANFIS is higher than the others. The MSE rates of prediction influence the CA directly. However, the MSE rates decrease, and the CA would increase subsequently.

The percentage of CA for twelve persons from S1 to S12 is presented in Table 2. The percentage of CA is calculated as:

$$CA\% = \frac{\text{the number of true classified signals}}{\text{the number of all test signals}} \times 100. \tag{15}$$

The basis of classification is established on the morphological discrepancy of EEG signals in variant action imaginations because EEG signals are classified in the time domain and differences in their shapes are the basis of

Table 1 MSE comparison analysis for 12 subjects (S1–S12) of EEG time-series prediction for each framework included recurrent ANFIS, feedforward ANFIS, ARMA model, and neural network

	Recurrent ANFIS	Feedforward ANFIS	ARMA model	Neural network
S1	0.632 e–004	0.972 e–003	0.441 e–002	0.126 e–002
S2	0.278 e–003	0.285 e–001	0.491 e–001	0.372 e–001
S3	0.154 e–003	0.135 e–001	0.268 e–001	0.219 e–001
S4	0.447 e–002	0.508 e–001	0.815 e–001	0.687 e–001
S5	0.697 e–002	0.672 e–001	0.991 e–001	0.759 e–001
S6	0.255 e–003	0.305 e–001	0.402 e–001	0.443 e–001
S7	0.683 e–002	0.602 e–001	0.874 e–001	0.632 e–001
S8	0.769 e–002	0.874 e–001	0.105	0.992 e–001
S9	0.331	0.542	0.753	0.610
S10	0.230 e–003	0.268 e–001	0.337 e–001	0.288 e–001
S11	0.312 e–003	0.295 e–001	0.426 e–001	0.489 e–001
S12	0.441 e–002	0.564 e–001	0.663 e–001	0.672 e–001
Average	0.0302	0.0828	0.1158	0.0972

Table 2 Comparison analysis of average percentage classification accuracy rates for fivefold cross-validation for 12 subjects (S1–S12) of EEG time-series prediction for each framework included recurrent ANFIS, feedforward ANFIS, ARMA model, and neural network

	Recurrent ANFIS	Feedforward ANFIS	ARMA model	Neural network
S1	97.5 ± 3.2	85.6 ± 4.0	83.8 ± 4.1	84.1 ± 4.1
S2	89.3 ± 4.5	79.4 ± 5.1	76.8 ± 6.7	77.6 ± 5.5
S3	92.7 ± 2.8	82.5 ± 3.1	81.4 ± 3.0	82.3 ± 2.9
S4	83.2 ± 5.7	75.9 ± 6.2	72.8 ± 6.4	74.2 ± 5.9
S5	80.3 ± 6.1	74.6 ± 6.5	71.0 ± 6.5	73.6 ± 6.5
S6	89.1 ± 4.9	80.4 ± 6.6	78.5 ± 7.2	78.1 ± 6.8
S7	81.8 ± 5.2	79.1 ± 5.7	73.3 ± 5.3	75.4 ± 5.6
S8	79.7 ± 6.3	74.2 ± 7.0	70.8 ± 5.9	73.5 ± 6.5
S9	68.6 ± 5.4	65.4 ± 4.7	60.3 ± 5.7	64.3 ± 4.8
S10	91.5 ± 4.8	81.3 ± 5.0	78.1 ± 4.9	81.0 ± 5.1
S11	88.7 ± 7.0	79.2 ± 7.3	78.6 ± 8.8	77.3 ± 7.7
S12	83.9 ± 6.7	75.5 ± 6.9	74.9 ± 8.0	73.6 ± 7.4
Average	85.5 ± 5.2	77.7 ± 5.6	75.0 ± 6.0	76.2 ± 5.7

classification. According to Table 2, the CA of the recurrent ANFIS is more than the others which is acquired because of the lower MSE of prediction. The lower rate of MSE is gained because of the feedback structure of the designed network.

Time course CA related to the proposed recurrent ANFIS is shown in Fig. 11. Because the starting point of imagination in recording procedures is started from 3 to 7 s, this period is considered for the aim of this paper. CA rates have started from 0% at 3 s and have risen sharply, and after a while, no tangible changes are observed. According to Fig. 11, the CA rates become stable after a while despite the intrinsic time-to-time variation in EEG, which is a significant issue for a BCI system.

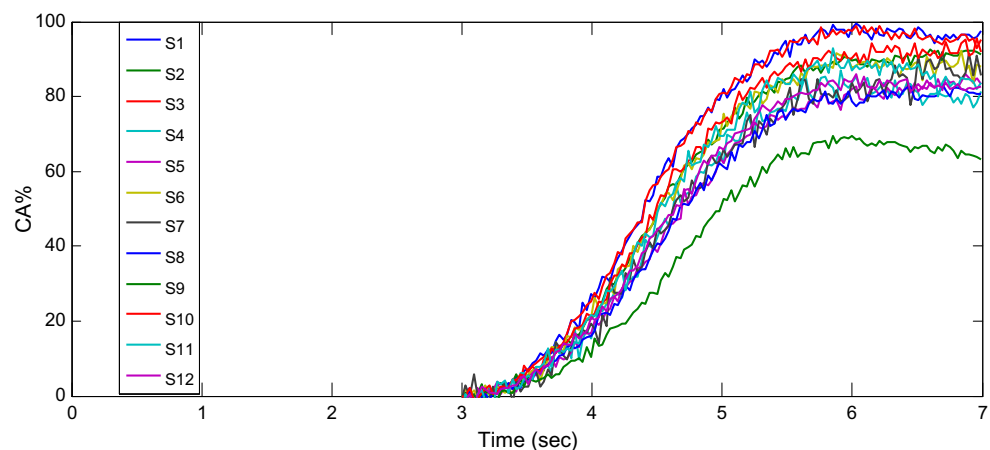
This approach shows a significant potential for the classification of such complex and nonstationary EEG signals. The designed network is obtained by feedbacking the errors of the former predictions. Therefore, a closed-loop system is configured. By using the errors of former predictions, the power of estimation is improved, and the errors of prediction in the next steps of prediction will be

diminished, and hereby a superior prediction will be performed. As a result, using the error information, the system can handle the uncertainty and the time-variant nature of the EEG signals. The system can model the EEG signals despite the nonlinearity and nonstationary identity. By increasing the modeling accuracy, the higher CA is obtained. Other potential advantages that can be mentioned here are the classification that can be carried out without knowing the initiate point of communication. Thus, this method has a potential for online EEG processing. Moreover, because the signal processing and classification are done in the time domain, there is no need to map signals to the other domains (e.g., frequency domain), typical feature extraction methods, and dimension reduction methods.

6 Conclusion

This paper presents a new class of recurrent adaptive fuzzy neural network for utilizing time-series prediction to classify EEG signals in brain–computer interface systems. The

Fig. 11 Time course of CA % for recurrent ANFIS based on time-series prediction for 12 subjects (S1–S12)



recurrent ANFIS is structured by feedbacking errors of previous predictions as inputs to make a closed-loop system. With feedback loops, this recurrent ANFIS can exploit the memorized past information to strengthen the capability of handling temporal prediction problems, such as modeling complex and nonstationary EEG signals. Overall, it is shown that the developed recurrent structure of ANFIS is a high-performance device regarding modeling and tracking.

For signals of each class, a network is considered. Each recurrent ANFIS is trained on EEG signals of one class and has the ability to recognize the similar structure on test signal by prediction with less error in comparison with other networks. Classification is established on the morphological discrepancy of MI signals in variant tasks. Results of online classification for twelve subjects are presented that demonstrate the potential of this prediction and classification method using recurrent ANFIS to be applied in a BCI.

A new class of recurrent ANFIS is developed that can be used for EEG modeling. The proposed network can also be used in biosignal modeling and processing properly too. Due to the inherent day-to-day variability in EEG signals, the feature extraction process needs to be cautious. The proposed classification method acts in the time domain and does not need the typical feature extraction and dimension reduction methods. Besides, classification based on the proposed time-series prediction does not need any knowledge of the initial point of signal recording. As a result, it can be useful for online asynchronous classification.

Acknowledgements The authors would like to acknowledge the Laboratory of Brain–Computer Interfaces, Institute of Biomedical Engineering, Graz University of Technology, Austria, for preparing the EEG data.

Compliance with ethical standards

Conflict of interest All authors have no potential conflict of interest.

References

- Buccino AP, Keles HO, Omurtag A (2016) Hybrid EEG-fNIRS asynchronous brain–computer interface for multiple motor tasks. *PLoS ONE* 11(1):e0146610
- Wasserman D, Aas S (2016) BCIs and disability: enhancement, environmental modification, and embodiment. *Brain Comput Interfaces* 3(3):126–132
- Cano-Izquierdo JM, Ibarrola J, Almonacid M (2012) Improving motor imagery classification with a new BCI design using neuro-fuzzy S-dFasArt. *IEEE Trans Neural Syst Rehabil Eng* 20(1):2–7
- Li M, Chen W, Zhang T (2017) Classification of epilepsy EEG signals using DWT-based envelope analysis and neural network ensemble. *Biomed Signal Process Control* 31:357–365
- Lotte F, Lécuyer A, Arnaldi B (2009) FuRIA: an inverse solution based feature extraction algorithm using fuzzy set theory for brain–computer interfaces. *IEEE Trans Signal Process* 57(8):3253–3263
- Choubisa M, Trivedi P (2015) Analysing EEG signals for detection of mind awake stage and sleep deprivation stage. In: 2015 International conference on green computing and internet of things (ICGCIoT), IEEE, pp 1209–1211
- Komijani H, Nabaei A, Zarrabi H (2016) Classification of normal and epileptic EEG signals using adaptive neuro-fuzzy network based on time series prediction. *Neurosci Biomed Eng J* 4(4):273–277
- Abu-Alqumsan M, Ebert F, Peer A (2017) Goal-recognition-based adaptive brain–computer interface for navigating immersive robotic systems. *J Neural Eng* 14:036024
- Jang JS (1993) ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans Syst Man Cybern* 23(3):665–685
- Luu TP, He Y, Brown S, Nakagame S, Contreras-Vidal JL (2016) Gait adaptation to visual kinematic perturbations using a real-time closed-loop brain–computer interface to a virtual reality avatar. *J Neural Eng* 13(3):036006
- Saa JFD, Çetin M (2013) Discriminative methods for classification of asynchronous imaginary motor tasks from EEG data. *IEEE Trans Neural Syst Rehabil Eng* 21(5):716–724
- Coyle D, Prasad G, McGinnity TM (2009) Faster self-organizing fuzzy neural network training and a hyperparameter analysis for a brain–computer interface. *IEEE Trans Syst Man Cybern B Cybern* 39(6):1458–1471
- Mohammadi MR, Khaleghi A, Nasrabadi AM, Rafeivand S, Begol M, Zarafshan H (2016) EEG classification of ADHD and normal children using non-linear features and neural network. *Biomed Eng Lett* 6(2):66–73
- Chai R, Naik GR, Nguyen TN, Ling SH, Tran Y, Nguyen HT (2016) Selecting optimal EEG channels for mental tasks classification: an approach using ICA. In: 2016 IEEE congress on evolutionary computation (CEC), Vancouver, BC, pp 1331–1335
- Chai R, Naik GR, Nguyen TN, Ling SH, Tran Y, Craig A, Nguyen HT (2017) Driver fatigue classification with independent component by entropy rate bound minimization analysis in an EEG-based system. *IEEE J Biomed Health Inform* 21(3):715–724
- Chai R et al (2016) Classification of EEG based-mental fatigue using principal component analysis and Bayesian neural network. In: 2016 IEEE 38th annual international conference of the engineering in medicine and biology society (EMBC), IEEE
- Chai R et al (2015) Classification of driver fatigue in an electroencephalography-based countermeasure system with source separation module. In: 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC), Milan, pp 514–517
- La Rocca D, Campisi P, Scarano G (2012) EEG biometrics for individual recognition in resting state with closed eyes. In: 2012 BIOSIG-Proceedings of the international conference of the biometrics special interest group (BIOSIG), IEEE
- Kumari P, Kumar S, Vaish A (2014) Feature extraction using empirical mode decomposition for biometric system. In: 2014 international conference on signal propagation and computer technology (ICSPCT), IEEE
- Das R, et al (2015) EEG biometrics for user recognition using visually evoked potentials. In: 2015 International conference of the biometrics special interest group (BIOSIG), IEEE
- Vatankhah M, Yaghubi M (2009). Adaptive neuro-fuzzy inference system for classification of eeg signals using fractal dimension. In: 2009. EMS'09. Third UKSim European symposium on computer modeling and simulation, IEEE pp 214–218
- Yu W, Feng H, Feng Y, Madani K, Sabourin C (2015) Non-holonomic mobile system control by combining EEG-based BCI with ANFIS. *Bio Med Mater Eng* 26(s1):S1125–S1133

23. Begum D, Ravikumar KM, Mathew J, Kubakaddi S, Yadav R (2015) EEG based patient monitoring system for mental alertness using adaptive neuro-fuzzy approach. *J Med Bioeng* 4(1):59–66
24. Homri I, Yacoub S, Ellouze N (2015) Combined approaches of features selection for EEG classification. *Int Rev Comput Softw IRECOS* 10(3):256–264
25. Hekim M (2016) The classification of EEG signals using discretization-based entropy and the adaptive neuro-fuzzy inference system. *Turk J Electr Eng Comput Sci* 24(1):285–297
26. Adedayo OO, Oladipo F, Ijeh-Ogboi C (2015) Softcomputing techniques for improved electroencephalogram signal analysis. *Int J Curr Eng Technol* 5(3):2181–2186
27. Blankertz B, Muller KR, Krusienski DJ, Schalk G, Wolpaw JR, Schlögl A, Birbaumer N (2006) The BCI competition III: validating alternative approaches to actual BCI problems. *IEEE Trans Neural Syst Rehabil Eng* 14(2):153–159
28. Blankertz B et al (2005) BCI competition III and IV [Online]. <http://www.bbc.de/competition/>
29. Schlögl A et al (2005 and 2008) BCI-competitions III (Dataset IIIa) and IV (Dataset 2A) [Online]. http://www.bbc.de/competition/iii/#data_set_iii and http://www.Bbc.de/competition/iv/desc_2a.pdf
30. Schlögl A, Lee F, Bischof H, Pfurtscheller G (2005) Characterization of four-class motor imagery EEG data for the BCI-competition 2005. *J Neural Eng* 2(4):L14
31. Pfurtscheller G, Neuper C (2001) Motor imagery and direct brain–computer communication. *Proc IEEE* 89(7):1123–1134
32. Bhardwaj S et al (2015) Online and automated reliable system design to remove blink and muscle artefact in EEG. In: 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC), Milan, pp 6784–6787
33. Kwon Y et al (2015) Efficient learning of image super-resolution and compression artifact removal with semi-local Gaussian processes. *IEEE Trans Pattern Anal Mach Intell* 37(9):1792–1805
34. Jadhav PN et al (2014) Automated detection and correction of eye blink and muscular artefacts in EEG signal for analysis of Autism Spectrum Disorder. In: 2014 36th annual international conference of the IEEE engineering in medicine and biology society, Chicago, IL, pp 1881–1884
35. Jahani Fariman H, Ahmad SA, Hamiruce Marhaban M, Ali Jan Ghasab M, Chappell PH (2015) Simple and computationally efficient movement classification approach for EMG-controlled prosthetic hand: ANFIS vs. artificial neural network. *Intell Autom Soft Comput* 21(4):559–573
36. Dash R, Dash P (2016) Efficient stock price prediction using a self evolving recurrent neuro-fuzzy inference system optimized through a modified technique. *Expert Syst Appl* 52:75–90
37. Lin YY, Chang JY, Lin CT (2013) Identification and prediction of dynamic systems using an interactively recurrent self-evolving fuzzy neural network. *IEEE Trans Neural Netw Learn Syst* 24(2):310–321
38. Frolov AA, Husek D, Polyakov PY (2009) Recurrent-neural-network-based Boolean factor analysis and its application to word clustering. *IEEE Trans Neural Netw* 20(7):1073–1086
39. Coyle D, Prasad G, McGinnity TM (2005) A time-series prediction approach for feature extraction in a brain–computer interface. *IEEE Trans Neural Syst Rehabil Eng* 13(4):461–467
40. Coyle D (2009) Neural network based auto association and time-series prediction for biosignal processing in brain–computer interfaces. *IEEE Comput Intell Mag* 4:47–59
41. Shiman F, Safavi SH, Vaneghi FM, Oladazimi M, Safari MJ, Ibrahim F (2012) EEG feature extraction using parametric and non-parametric models. In: Proceedings of 2012 IEEE-EMBS international conference on biomedical and health informatics
42. Zhao Y, Billings SA, Wei HL, Sarrigiannis PG (2013) A parametric method to measure time-varying linear and nonlinear causality with applications to EEG data. *IEEE Trans Biomed Eng* 60(11):3141–3148