2022-10-05

Intelligent fuzzy system for automatic artifact detection and removal from EEG signals

Agounad, Said

Elsevier

Agounad, Said, Soukaina Hamou, Ousama Tarahi, Mustapha Moufassih, and Md Kafiul Islam. "Intelligent fuzzy system for automatic artifact detection and removal from EEG signals." Journal of King Saud University-Computer and Information Sciences 34, no. 10 (2022): 9428-9441.

https://ar.iub.edu.bd/handle/123456789/557

Journal of King Saud University – Computer and Information Sciences 34 (2022) 9428–9441

HOSTED BY

Contents lists available at ScienceDirect

Journal of King Saud University – Computer and Information Sciences

journal homepage: www.sciencedirect.com

Intelligent fuzzy system for automatic artifact detection and removal from EEG signals

Said Agounad [a], Soukaina Hamou [a], Ousama Tarahi [a], Mustapha Moufassih [a], Md Kafiul Islam [b,*]

[a] Laboratory of Metrology and Information Processing, Department of physics, Faculty of Sciences, Ibn Zohr University, Agadir, Morocco
[b] Biomedical Instrumentation and Signal Processing Lab (BISPL), Department of Electrical and Electronic Engineering, Independent University, Bangladesh

ARTICLE INFO

Article history:
Received 3 June 2022
Revised 2 September 2022
Accepted 28 September 2022
Available online 5 October 2022

Keywords:
Fuzzy logic system
Stationary wavelet transform
EEG signal
Artifact removal algorithms
Performance evaluation

ABSTRACT

The EEG signals were used in many medical and technological applications such as diagnosis of diseases, rehabilitation of disabled peoples, preventive healthcare, BCI (brain computer interface) systems. EEG signal is prone to the physiological and non-physiological artifacts which severely affect them and lead to its misinterpretation. An automatic method and/or algorithm; for handling EEG artifacts; is proposed. The proposed method is based on three statistical parameters (entropy, kurtosis and skewness), fuzzy inference system (FIS) and stationary wavelet transform (SWT). Each incoming EEG epoch is described using these three statistical parameters. Based on the extracted statistical parameters, the designed FIS decides if an epoch is artifactual or not. Then SWT is used to decompose the EEG epoch into detail and approximation coefficients. To reduce the effect of artifact removal, we propose to use other fuzzy inference systems, which allow to select the contaminated wavelet coefficients. The universal thresholding method is then applied to the corrupted coefficients. Finally, the inverse SWT applies to the thresholded and non-corrupted coefficients to restore the cleaned EEG signal. The performance of the proposed method in terms of amount of artifact removal and signal distortion is evaluated in three scenarios: fully simulated, semi-simulated, and real artifactual EEG data. The comparison of our method with some existing state-of-the-art methods shows the superiority of our method over others in terms of performance and computational time.

1. Introduction

The EEG signal (Electroencephalogram) is one of the most used biosignals in a wide range of medical and non-medical applications. In the medical field, the EEG signal is used in diagnosis and preventive of neurological diseases and in the study of the brain function. Brain computer interface (BCI) is one of the most known applications of EEG signal in non-medical field. A BCI system is an intelligent system that allows the user to control the external devices, such as robotic limbs, exoskeletons, and computers, without passing through a neural pathway. At the beginning, the BCI systems are developed to help disabled people to improve their life. Recently, with technological and medical development, i.e. improved sensor, wireless data transfer, cloud computing, flexible and wearable electronics, new applications of EEG signals are emerged such as, preventive health care, patient health monitoring, robotics, etc. (Abdulkader et al., 2015).

Electroencephalography is the used recording technique to measure the EEG signal. It consists to measure the electrical activity of brain by placing electrodes on the scalp. EEG technique is one of the most attractive signals in several applications due to non-invasive, portable, acceptable temporal resolution, and low cost. Unfortunately, the EEG signals are commonly prone to the artifacts that originate from other sources other than brain. The artifacts in the EEG signal are classified into physiological and non physiological artifacts. The physiological artifacts known also as internal artifacts originate from the physiological activities of the subject. These artifacts include ECG artifact (electrocardiogram), EOG artifact (electrooculogram) which is due to eye movement and eye blink, EMG artifact (electomyogram) which is due to the muscle activity. Non-physiological artifacts are mainly due to the different sources of the electrical field, environment and movement. These artifacts comprise, electrode pop, 50/60 Hz power line artifact, subject movements. The different artifacts overlap with EEG signal in temporal, spectral and sometimes in spatial domains. It leads to misinterpretation of the EEG signal and take the incorrect

* Corresponding author at: Dept. of Electrical and Electronic Engineering, Independent University, Bangladesh, Plot 16, Block B, Bashundhara R/A, Dhaka 1229, Bangladesh.
E-mail address: kafiul_islam@iub.edu.bd (M.K. Islam).

https://doi.org/10.1016/j.jksuci.2022.09.020
1319-1578/

decisions. In technological field, the artifact can lead to unintentional control of the devices. For example, in the case of BCI application, the artifacts can lead to morphological change of a neural activity that drives the BCI system to make a false decision. In the medical field, artifacts can lead to prompting incorrect treatment. For example, in the case of preventive epileptic seizure, the artifacts can mimic the shape of the epileptic seizure that lead to false alarms (Seneviratne et al., 2013).

In the early applications of the EEG signal, the experts identify visually the EEG segments corrupted by the artifacts and reject them. This approach leads to lost the neural activity. Recently, with the emergence of new applications of the EEG signal, several methods have been developed to detect and remove the artifacts from the EEG signal (Islam et al., 2016). However, there is no an efficient method that provides a complete solution for handling EEG artifacts. Some methods are specified in removing of particular artifacts, for example EOG (Nguyen et al., 2012), EMG (Chen et al., 2019), EMG and EOG (Hu et al., 2015), ECG and eye blinks (Shoker et al., 2005), motion artifacts (Kilicarslan and Vidal, 2019), 50/60 power line artifacts (La Rosa et al., 2021). Other methods such as adaptive filtering (Molla et al., 2012; Borowicz, 2018) and regression (Bengtsson and Cavanaugh, 2006) methods can not perform if the reference channel or prior information are not available. Blind source separation methods like ICA (independent component analysis) and CCA (canonical component analysis) can work only in multi-channel EEG recording (Kanoga et al., 2020; Mahapatra et al., 2018). Empirical mode decomposition (EMD) method is used only in single EEG signal recording and is not suitable for online application since it costs in term of computational time (Islam et al., 2021; Martis et al., 2012). Thus, it is needed to develop new methods that are able to overcome the limitations of the existing artifact handling methods. There is a strong urge of that specially with the emerging trends of the EEG signal such as BCI and continuous ambulatory monitoring applications.

In this paper, we proposed an automatic method to detect and remove the artifacts from the EEG signal. This method is independent of artifact types, can perform for both single and multi-channel recordings, does not require a reference channel. In addition, it does not need prior information, performs fully automatically and does not require any user parameter adjustment. Thus, these characteristics make this method more suitable in pre-processor phase for different technological and medical applications of the EEG signals.

The proposed method is based on the fact that the artifact characteristics are different from those of the neural activity. In general the artifact is a random signal with higher amplitude than EEG background, tends to be transient unlike of EEG which is rhythmic, makes a distribution of data asymmetrical instead of symmetric as in the case of EEG background. In the proposed algorithm, for each epoch of EEG signal, we calculate three statistical parameters, namely entropy (a measure of disorder), skewness (a measure of symmetry) and kurtosis (a measure of peakedness). These parameters are used as the inputs of a developed fuzzy inference system which makes a decision if an epoch is artifactual or not. If an epoch is artifactual, then the stationary wavelet transform (SWT) is used to decompose the EEG epochs into its components. To reduce the amount of distortion brings to the EEG signal, we propose to use other fuzzy inference systems which allow to identify the contaminated SWT components. The universal thresholding method is then used in order to remove the artifacts in the contaminated components. Finally, the inverse SWT is used to reconstruct the artifact-free EEG segments. The performance of the proposed method is evaluated using fully simulated artifactual data (scenario 1), semi-simulated artifactual data (scenario 2) and fully real artifactual data (scenario 3). Qualitative and quantitative metrics are used to demonstrate the performance of the proposed method both in terms of artifact removal and signal distortion. Moreover, the proposed method is compared with some widely known methods in the literature both in terms of performance and computational time.

The rest of this paper is organized as follows. Section 2 presents the related work. The proposed method is described in Section 3. Section 4 presents the performance evaluation of the proposed algorithm. Finally, Section 5 includes the concluded remarks.

2. Related work

Several methods have been proposed to hand with the EEG artifacts. Blind source separation based algorithms are one of the most widely known methods especially, ICA, (Hyvarinen et al., 2001; Zhang and Sanderson, 2007; Klados et al., 2011; Kanoga et al., 2020), CCA (De Clercq et al., 2006; Gao et al., 2010) and MCA (Yong et al., 2009; Mahapatra et al., 2018). The ICA aims to separate multichannel EEG signals into their independent sources. Then, the artifactual sources are manually removed. The principle of CCA is based on second order statistics. It aims to separate the multiple EEG signals into maximally uncorrelated components and then the artifactual components are manually rejected before reconstructing the artifact corrected EEG signals. These methods are not suitable for online applications and cannot operate on few or single-channel EEG. The MCA requires a prior morphological characteristics on the artifacts. EMD and its version EEMD are other popular methods to remove artifacts from EEG signals (Safieddine et al., 2012; Molla et al., 2012; Martis et al., 2012). It consists to decompose the EEG recording into multiple components called IMFs. The computational complexity is one of the main drawbacks of EMD or EEMD. The linear regression (Bengtsson and Cavanaugh, 2006; Harrison et al., 2003) and adaptative filtering (Molla et al., 2012; Borowicz, 2018) based methods are also used to handle the EEG artifacts especially ECG and EOG. These methods need a reference channel, e.i., an extra electrode or an extra algorithm. In several applications, the reference electrode is not available and then the regression and adaptive based methods are not suitable for such applications. On the other hand, an extra algorithm means more time consuming, a thing which is not suitable especially in on-line application. Wavelet transform especially DWT and SWT are currently widely used in several medical and non-medical applications (Sharma et al., 2015; Islam et al., 2015; Ocak, 2009; Ghorbanian et al., 2012). It consists to decompose the EEG signal into components known as approximation and detail coefficients.

To benefit from the advantages of different methods, recently several hybrid methods are proposed. Some authors proposed to combine the BSS methods and wavelet transform, for example, wavelet ICA (wICA) (Azzerboni et al., 2004; Mammone et al., 2011) and wavelet enhanced ICA (Castellanos and Makarov, 2006), CCA-SWT (Mowla et al., 2015; Raghavendra and Dutt, 2011). The works (Chen et al., 2014; Mijović et al., 2010; Sweeney et al., 2012) proposed artifact handling hybrid methods based on the association of BSS and EMD or EEMD. To remove ECG and EOG automatically in online applications some authors proposed to combine the BSS and regression techniques (Klados et al., 2011; Mannan et al., 2016; Wang et al., 2014; Guarnieri et al., 2018). In recent times, several works are proposed to combine the performance of machine learning methods and BSS techniques or wavelet transform. These methods have been introduced to detect and remove automatically the artifacts from the EEG signals. BSS-machine learning techniques combines one of the BSS methods and one of machine learning methods (e.g. SVM, ANN or fuzzy logic) (Yasoda et al., 2020; Radüntz et al., 2017). In Phadikar et al. (2020), Nguyen et al. (2012) and Tibdewal et al.

(2015), authors proposed to associate the wavelet transform with machine learning methods. To remove cardiac artifacts from sleep EEG recordings, authors of Ranjan et al. (2022) proposed a hybrid method which includes an adaptive threshold-based nonlinear Teager-Kaiser energy operator (TEO), empirical wavelet transforms (EWT), customized morphological filter and modified ensemble average subtraction (MEAS). They showed that the proposed method is efficient and removes the cardiac artifacts without affecting much the EEG activity. In Shukla et al. (2021) the authors presented an automatic algorithm to eradicate motion artifacts from EEG signals. This algorithm is developed using Gaussian elimination CCA (GECCA) and EEMD. Shahbakhti et al. (2021) proposed a novel method to remove the eye blink artifact from Fp1 EEG electrode: firstly, the authors used variational mode extraction to identify and derive blink artifact from Fp1 electrode. Secondly, they projected the identified eye blink artifact to the rest of electrodes and then filtered by a combination of DWT and PCA (principal component analysis). The authors of Placidi et al. (2021) developed an automatic framework for removing artifacts from EEG signals. They used ICA to decompose the EEG signal into ICs (independent components) whose re-projection on Topoplots (2D topographies of the scalp). Thereafter, the obtained Topoplots are used as inputs of the Convolutional Neural Network (CNN) which divides them into artifact components and free-artifact EEG components.

Another efficient and promising way to detect if an epoch of EEG recording is artifactual or not is to use the statistical parameters. Once an artifactual epoch is detected, the denoising method like wavelet based thresholding can be used to remove the artifacts. In Islam et al. (2021) and Shahbakhti et al. (2019) authors proposed to combine the statistical parameters and wavelet transform to detect the artifact and remove them from the EEG signal. Authors of Hussein et al. (2022) proposed to use EEG signal statistics and ICA to conceal eye and muscle movement artifacts from EEG signals. They showed qualitatively the performance of the proposed approach in EEG artifact rejection. SWT and kurtosis are used in Shahbakhti et al. (2021) to reject electrical shifts and linear trend artifacts in EEG recordings. The authors showed that the proposed method performed well than wICA and Enhanced wICA (EwICA) algorithms.

3. Proposed method

The flowchart of the proposed algorithm is presented in Fig. 1. This algorithm can be decomposed into three blocks, detection, denoising and reconstruction. In the following, we discuss the principle and the methods used in each block.

3.1. Detection block

3.1.1. Segmentation

The first step of our proposed algorithm is the segmentation of the EEG signal, S_n, which is sampled at frequency f_s. This aims to divide the incoming EEG signal into m non-overlapped windows or epochs, of size N. Lets x_i denotes the ith epoch and is defined as:

$$x_i = S_n((i-1)N+1, (i-1)N+2, \ldots, iN). \quad (1)$$

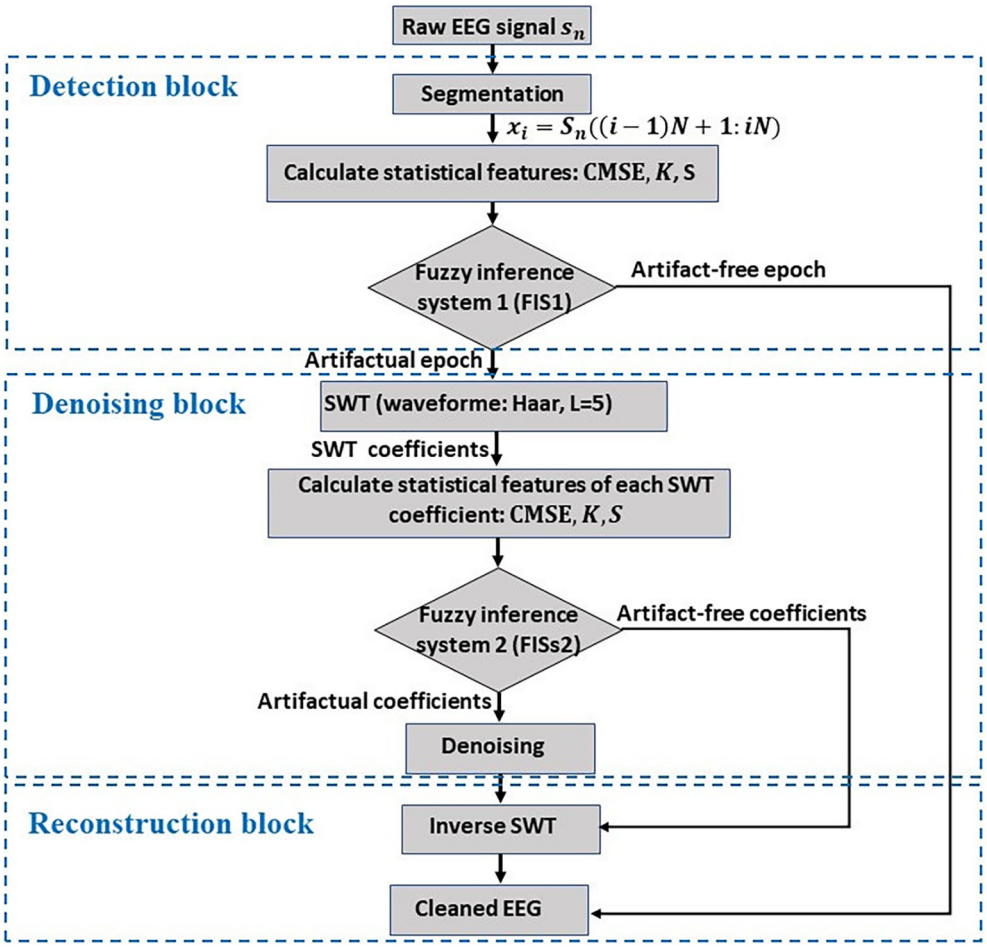

Fig. 1. Flowchart of the proposed method for detecting and removing artifacts from EEG recording.

It is worth mentioning that it is important to choose the duration of the segment, N/fs, carefully since it plays a key role in detection of artifacts and removing them. Indeed, short duration, e.g. $N/fs < 1s$, may not represent the neural activities properly and thus mimic the waveform of the artifact. Therefore, denoising such segment can increase the amount of distortion made to the EEG signal. On the other hand, long duration, i.e. $N/fs >$ several second (e.g. 4s), may lead to artifacts being missed. Therefore, the amount of artifact removal will be low. In addition, the choice of an optimal epoch duration is important, since the EEG signal can be considered as stationary during such epoch. Therefore, the statistical features in this epoch can be considered stationary too. In our case, after trying different values empirically, we found that the optimal N/fs is 1 s. In Islam et al. (2015) authors have considered epochs of three seconds in order to remove artifacts from seizure data. In Islam et al. (2021) authors showed that one second duration epoch is sufficient to detect and remove artifacts in the case of BCI applications.

3.1.2. Statistical features

Based on statistical characteristics, it is possible to detect if an EEG signal is artifactual or not. This can be done based on the fact that the characteristics of artifacts are different from those of the neural activity. In our algorithm, we verified that three statistical parameters, Composite multiscale entropy (CMSE, a measure of randomness), kurtosis (a measure of peakedness) and skewness (a measure of symmetry), are sufficient to characterize different artifacts.

- Composite multiscale entropy (CMSE)
  *CMSE*is used to quantify the randomness of the samples in an epoch. It takes lower values for an artifactual epoch and higher values for artifact-free epoch. That means that the artifact is more random than the neural activity. Lets $x = x_1, x_2, \ldots x_N$ a discrete time signal of length N. In the CMSE method, at a scale factor of τ, the sample entropies of all coarse-grained time signal x, associate with different starting points of the coarse-graining process, are calculated. The CMSE of x is then defined as the mean of τ sample entropy values (Wu et al., 2014):

$$CMSE(x, \tau, m, r) = \frac{1}{\tau}\sum_{k=1}^{\tau}\left(-\ln\frac{n_{k,\tau}^{m+1}}{n_{k,\tau}^{m}}\right). \tag{2}$$

  where m is a length of template vectors used to compute the sample entropy of x, r is predefined tolerance, which is used to decide if two template vectors are matched or not, $n_{k,\tau}^{m}$ represents the total number of m-dimensional matched vector pairs, is calculated from the *kth*coarse-grained time signal at a scale factor τ, and $n_{k,\tau}^{m+1}$ represents the total number of $(m+1)$-dimensional matched vector pairs.
- Skewness (S)
  S is used to measure the degree of symmetry or skewed of a signal around its mean. S is higher for artifactual epoch, while is lower for clean epoch. Therefore artifacts have an asymmetrical distribution with longer tail on one side of signal while the background EEG has symmetrical distribution. It can be computed using the following formula:

$$S = \frac{\frac{1}{N}\sum_{k=1}^{N}(x_k - \bar{x})^3}{\left(\frac{1}{N}\sum_{k=1}^{N}(x_k - \bar{x})^2\right)^{3/2}}. \tag{3}$$

  where N is the length of signal x and $\bar{x}$ is its mean. S can be positive, signal skewed right, negative, signal skewed left, or null, the signal is perfectly symmetrical.
- Kurtosis (K)
  $K$is a measure of peakedness of a signal. An artifacual epoch has higher kurtosis than the artifact-free EEG. That means that the artifact is characterized by higher amplitude than the EEG signal. It is defined as:

$$K = \frac{\frac{1}{N}\sum_{k=1}^{N}(x_k - \bar{x})^4}{\left(\frac{1}{N}\sum_{k=1}^{N}(x_k - \bar{x})^2\right)^{2}} - 3. \tag{4}$$

  The reference value of K is 3, which is the value of K for a normal distribution. If $K > 3$, compared to a normal distribution, the central peak of the signal is higher and sharper. While, if $K < 3$, the central peak is lower and broader.

In the proposed algorithm, these statistical parameters are used in two stages:

- In the first stage, we used these parameters to calculate the statistical features of an epoch, which will be used later in the first decision stage to decide if an epoch is artifactual or not.
- In the second stage, to remove the artifacts without affecting the neural activity, we propose to calculate the statistical features of different approximation and detail coefficients of the EEG signal. These features will be used in the second decision stage to decide if an approx. or detail coefficient is artifactual or not.

3.1.3. Fuzzy inference system (FIS)

As described in the flowchart of the proposed algorithm (see Fig. 1), the FIS is used in two stages:

- In the first stage, the FIS is used to decide automatically if an epoch is artifactual or not. Therefore, this method is an alternative of several methods which need the intervention of the user or prior information to determine the artifactual epoch (Islam et al., 2021; Klados et al., 2011; Kanoga et al., 2020).
- In the second stage, the FIS is used to make a decision regarding denoising or not of an approx. or detail coefficient. In several works once an epoch is marked artifactual, denoising operation is applied to whole wavelet coefficients even if some of them are not contaminated (Islam et al., 2021; Mammone et al., 2011; Mowla et al., 2015; Raghavendra and Dutt, 2011). Therefore, the amount of distortion is increased. Our approach is based on the fact that almost all artifacts are localized in the frequency domain. Thus, the proposed FIS allows to preserve intact the clean frequency bands (approx. and detail coefficients).

At the level of the FIS each signal epoch, x_i, is represented by a set of features, $x_i^f = \{cmse_i, k_i, s_i\}$ extracted in the previous block. The goal of the FIS is to assign a class C_i to the input signal, x_i, from the predefined class set $C = \{artifact, artifact - free\}$, based on the set of features x_i^f. The first class, *artifact*, includes the artifactual incoming epochs. While the second class, *artifact – free*, includes the non artifactual epochs. The principle of FIS is based on a fuzzy logic theory which aims to manipulate data that are not precise based on the concept of fuzzy sets. In the classical theory of sets, each element in the set is represented by a binary statement (0 or 1), e.i. either belong or not belong to the set and the transition from one set to its neighbors is abrupt. While, in fuzzy set, an element can belong to more than one set with different partial degrees and the transition from one set to another may be gradual. Lets A be a fuzzy set (or fuzzy variable) on the universe of discourse

(or linguistic variable) U. A can be described as a set of ordered pairs:

$$A = \{(u, \mu_A(u)) | u \in U\} \quad \text{where} \quad 0 \leqslant \mu_A(u) \leqslant 1 \tag{5}$$

where $\mu_A(u)$ is the membership function of u in A. The membership function defines the degree with which each input u belongs to the set A. It varied between 0 and 1. Several types of membership functions can be considered such as trapezoidal, Gaussian, triangular, etc. The choice of the appropriate membership function either in term of form or boundary is one of the most important steps in the designing of the FIS. This function has a direct influence on the results of the classification since it defines the relation between the input variable and the universe of discourse. In this work, we consider three universes of discourses, $CMSE, K$ and S. The variation range of each of these three variables is determined based on the calculation of, entropy (CMSE), Kurtosis, and Skewness for different epochs of simulated clean EEG signal and simulated artifactual EEG signal. After deep studying of the defined variation ranges, we have decided to use trapezoidal membership functions. We have divided *CMSE* into two fuzzy sets which are low and large. Low values of *CMSE* mean that there are the artifacts, however, higher values mean that the EEG epoch is free of artifacts. S is represented by two fuzzy sets: low and large. Low values of S mean that the EEG signal is symmetric and it is free of artifacts, however, higher values mean that the signal is antisymmetric and it is probable to be artifactual. K is a collection of three fuzzy sets: low, medium, and large. Low and high values of K signify that the distribution of EEG epoch tends to be either horizontal line (non-Gaussian) or Gaussian function with high amplitude. In these two cases, it is possible that the EEG epoch is corrupted. However, medium values of K (standard amplitude of Gaussian function) mean that the signal is free of artifacts. Mathematically, the trapezoidal function is defined as follows:

$$\mu(x; a, b, c, d) = max\left(min\left(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c}\right), 0\right), \tag{6}$$

where, a, b, c, and d are the x-coordinates of the four corners of the trapezoidal function. Fig. 2 shows the membership functions for each of the input variables (CMSE, K, or S) and the output classes, $C = \{artifact, artifact - free\}$, which are used in the first decision stage to decide if an epoch is artifactual or not. FIS takes a decision based on a set of fuzzy rules. To build a FIS, it is needed to define two main things: knowledge base and fuzzy reasoning method. The knowledge base aims to divide the input space into a number of fuzzy sets with appropriate fuzzy membership function and to build a set of $if - then$ fuzzy rules based on fuzzy sets. The fuzzy reasoning method defines the procedure followed by the FIS to make a decision. Different types of fuzzy rules were used. In our case, the general form of fuzzy rules is:

$$R_j: \ if \ u_1 \text{ is } A_{j1} \text{ and } u_2 \text{ is } A_{j2} \text{ and} \ldots \text{and } u_m \text{ is } A_{jm} \text{ then Class is } C_j \text{ with } \omega_j. \tag{7}$$

where R_j, $j = 1, 2, \ldots, L$, is the jth rule, $u = \{u_1, u_2, \ldots, u_N\}$ is a feature vector, $A_{jn}, n = 1, 2, \ldots, m$ is fuzzy sets in antecedents, and ω_j is the weight of the jthr ule. Table 1 shows the fuzzy rule base used to decide if an epoch is artifactual or not. Several FIS were proposed such as Mandani fuzzy model and Takagi–Sugeno fuzzy model. In our case, we aim to develop a fuzzy classifier system that operates on the vote. Here are the steps followed to classify an incoming EEG segment, x_i, which is represented by its feature vector, $x_i^f = \{cmse_i, k_i, s_i\}$:

- The first step aims to determine, for each rule R_j, $j = 1, 2, \ldots, 8$,

  - the satisfaction degree of the clauses ($x_i^f(n)$ is A_{nj}), $n = 1, ..3$ and $j = 1, 2, \ldots, 8$. This degree defines the degree with which an input variable, $x_i^f(n)$, belongs to the fuzzy set A_{nj} and is determined by the membership function associated with $A_{nj} : \mu_{A_{nj}}(x_i^f(n))$.

  - the matching degree, β_j, which determines the activation strength of the if-part of the rule R_j with the incoming segment, x_i, and is defined as:

$$\beta_j(x_i) = sum(\mu_{A_{1j}}(cmse_i), \mu_{A_{2j}}(s_i), \mu_{A_{3j}}(k_i)). \tag{8}$$

  - the association degree, α_{jk}. This step consists to evaluate the association degree of the segment, x_i, with the output class of the rule R_j. As described in Table 1, the output of each rule is either *artifact* or $artifact - free$ class. That means that the weight of a class, ω_{jk}, is either one or zero. Thus the associate degree is none other than the matching degree:

$$\alpha_{jk} = \beta_j(x_i)\omega_{jk}, \ k = 1, 2. \tag{9}$$

- Aggregation step. This step consists to aggregate the outputs of the rules in order to define the winner class. Several aggregation operators can be used, in this work, the maximum operator is used as the aggregation operator. The aggregation degree of each class C_k is defined as:

$$\gamma_k = max(\alpha_{jk}), \ j = 1, 2, \ldots, 8, \ k = 1, 2. \tag{10}$$

- Decision step is the final step which consists to define the output class associate with the incoming EEG segment. The winner output class (*artifact*or $artifact - free$) is the class which corresponds to the highest aggregation degree.

3.2. Denoising block

As discussed in the related work several authors proposed to denoise the EEG signal based on the wavelet thresholding method. This method consists to decompose the EEG signal using wavelet transform into approximation and detail coefficients. After that, these coefficients are thresholded to remove the artifactual components. However, in this work the wavelet transform is used to decompose the EEG signal into approximation and detail coefficients, then the fuzzy inference system FIS2 (see Fig. 1) is used to select the artifactual wavelet coefficients. Finally, the thresholding method is applied to the contaminated coefficients. This approach allows to preserve intact the cleaned coefficients.

3.2.1. Wavelet decomposition

There are several types of wavelet transforms such as continuous wavelet transform (CWT), packet wavelet transform (PWT), discrete wavelet transform (DWT), stationary wavelet transform (SWT), etc. In the proposed algorithm, we chose to use SWT since it is translational invariant. That means that slow variation in a signal properties can't create a large variation in wavelet coefficients and large changes in energy distribution in the different wavelet scales (Safieddine et al., 2012). To use the wavelet transform it is necessary to set two things: mother wavelet and decomposition level. In the present work, we use the Haar wavelet as a mother wavelet due to its capacity to follow the transient artifactual activities better than the other mother wavelets. The choice of decomposition level is a function of the bandwidth of the EEG signal (i.e. 0.05–128 Hz) and the useful frequency band for the considered application, e.g, for seizure study the useful frequency band is 0–32 Hz, for BCI applications, the interesting bands are *delta*(< 4 Hz) for *ERP*-based BCI applications, e.g. *P*300 (Schalk and Mellinger, 2010) and for SCP-based BCI (Hou et al., 2017), *mu*(7–13 Hz) and *beta*(14–30 Hz) for motor imagery based BCI (Pfurtscheller et al., 2006) and *SSVEP*(12–18 Hz) (Muller-Putz and Pfurtscheller, 2007). In the case of BCI applications based EEG, as

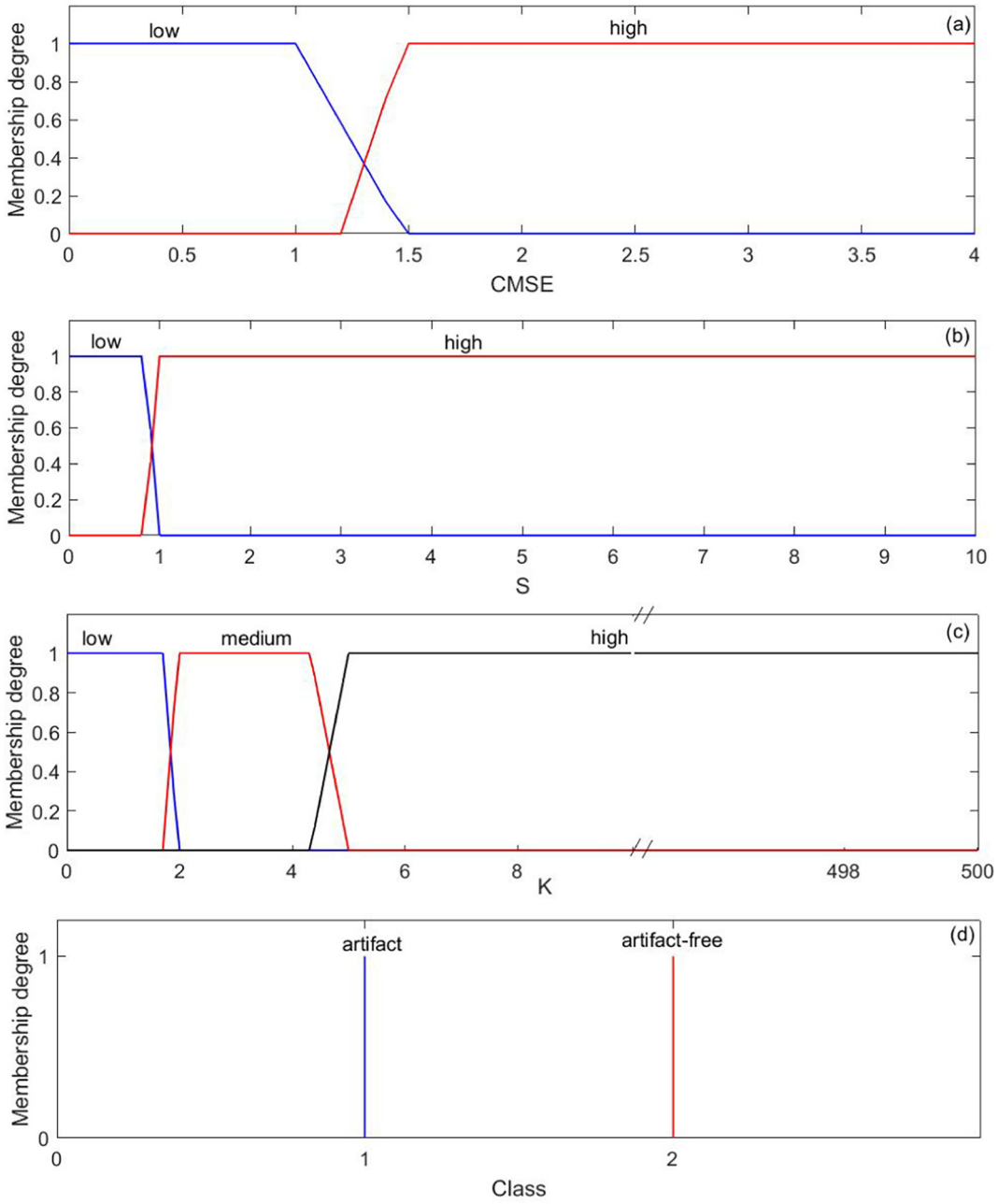

Fig. 2. Membership functions of FIS input variables (composite multiscale entropy (a), skewness (b), and kurtosis (c)) and membership function of the output classes (d).

Table 1
Rule base used in the FIS1 to decide if an epoch is artifactual or not.

Fuzzy rule base
R_1: **if** *cmse* is high **and** *k* is medium **and** *s* is low **then** class is *artifact – free* with 1
R_2: **if** *cmse* is high **and** *k* is medium **and** *s* is high **then** class is *artifact – free* with 1
R_3: **if** *cmse* is high **and** *k* is low or high **and** *s* is low **then** class is *artifact – free* with 1
R_4: **if** *cmse* is high **and** *k* is low or high **and** *s* is high **then** class is *artifact* with 1
R_5: **if** *cmse* is low **and** *k* is medium **and** *s* is low **then** class is *artifact – free* with 1
R_6: **if** *cmse* is low **and** *k* is medium **and** *s* is high **then** class is *artifact* with 1
R_7: **if** *cmse* is low **and** *k* is low or high **and** *s* is low **then** class is *artifact* with 1
R_8: **if** *cmse* is low **and** *k* is low or high **and** *s* is high **then** class is *artifact* with 1

considered in this work, level-5 decomposition is usually selected since this allows to cover different EEG frequency bands relevant to the BCI applications. The application of SWT on the incoming EEG segment, x_i, with level-5 decomposition and Haar as a mother wavelet provides the final approximation coefficient, a_{i5}, and detail coefficients $\{d_{i1}, d_{i2}, \ldots, d_{i5}\}$. Table 2 illustrates the frequency bands of the SWT coefficients, EEG rhythms, relevant BCI study and artifact type.

3.2.2. Identify statistical features of different SWT coefficients

As in the case of the incoming EEG segment, each SWT coefficient is described by a feature vector which include its statistical characteristics, i.e. cmse, s and k. The obtained feature vectors will be used further to feed the FISs2 (see Fig. 1).

Table 2
Frequency bands of SWT coefficients, EEG rhythm, relevant BCI studies and artifact type.

Frequency bands (HZ)	64–128	32–64	16–32	8–32	4–8	0–4
SWT coefficients	d_{i1}	d_{i2}	d_{i3}	d_{i4}	d_{i5}	a_{i5}
EEG rhythms	gamma	gamma	beta	mu alpha	theta	delta
Relevant BCI studies			motor imagery			P300
				SSVEP		SCP
Artifact type			EMG (8–64 Hz)		ECG	EOG
			electrode pop (8–64 Hz)		movement	

3.2.3. Identify the artifactual SWT coefficient

The same procedure presented in Section 3.1.3 is followed to build 6 FISs2 systems, 5 for detail coefficients and 1 for approx. coefficient. The feature vector which described each coefficient is applied to its corresponding FIS2 to decide if this coefficient is artifactual or not. The artifactual coefficients are transferred to the denoising stage while the non artifactual coefficients are preserved intact.

3.2.4. Denoising

Denoising block consists to denoise the artifactual SWT coefficients. In the proposed algorithm, we chose to use non-negative garrote shrinkage function (NNGSF) due to its interesting properties such as its low sensitivity to the input fluctuation, and low bias (Hong-Ye, 1998). NNGSF allows to make a nice trade off between the hard and soft threshold function in terms of amount of signal distortion and artifact removal. Let's denote $B_i \triangleq \{a_{i5}, d_{i1}, d_{i2}, \ldots, d_{i5}\}$. NNGSF for a given coefficient $b_{il} \in B_i$, with $l = 1, 2, \ldots, 6$ is defined as:

$$g(b_{il}) = \begin{cases} b_{il} & \text{if } |b_{il}| < T_{il} \\ \frac{T_{il}^2}{b_{il}} & \text{if } |b_{il}| > T_{il} \end{cases} \tag{11}$$

where T_{il} is the threshold value for the coefficient b_{il} and is defined as Safieddine et al. (2012):

$$T_{il} = c_{il}\sqrt{2\ln(N)}, \tag{12}$$

where N is the length of epoch and c_{il} is the estimated noise variance for b_{il} and is given by Castellanos and Makarov (2006)

$$c_{il} = \frac{median(|b_{il}|)}{0.6745}. \tag{13}$$

3.3. Reconstruction block

The final block consists to reconstruct the EEG epoch, x_i, by applying the inverse SWT on the denoised coefficients and non-artifactual coefficients. Finally, the cleaned EEG signal may be obtained by combining different denoised and non-artifactual epochs.

4. Performance evaluation

4.1. Data description

To evaluate the performance of the proposed method, we have considered three types of data, fully simulated artifactual EEG (scenario 1), semi-simulated artifactual EEG (scenario 2) and fully real artifactual EEG (scenario 3).

4.1.1. Scenario 1

In this scenario, to make fully simulated data, we have linearly combined the simulated EEG background activity with widely known six simulated artifact templates (muscle atifact, eye blink artifact, EOG artifact, slow movement artifact, electrode pop and cable movement) of different random amplitude and random duration. Fig. 3. (a) illustrates the process to synthesize the fully simulated artifactual EEG signal. The simulated background EEG signals are obtained based on the theory of the event related potentials (ERPs) (Yeung et al., 2004).

4.1.2. Scenario 2

The database used in this scenario is obtained by adding linearly the six artifact templates to the pure real EEG background activity. The characteristics of the used real EEG data base are presented in Table 3. Fig. 3(b) illustrates the process to generate the semi-simulated EEG signals.

4.1.3. Scenario 3

In the third considered scenario, we have evaluated the performance of our proposed algorithm using the fully real contaminated EEG data. These data are collected from the BCI competition IV: dataset-1 (Blankertz et al., 2007) and dataset-2a (Brunner et al., 2008). Table 3 shows the description of these data sets.

4.2. Metrics for performance evaluation

Two categories of metrics are used to evaluate the performance of the proposed algorithm. The first category allows to measure how much artifacts have been removed. This category includes: $\lambda(\%)$ which represents the amount of artifact removal in percentage and ΔSNR which indicates the improvement in signal to noise ratio (SNR). The second category is used to measure how much distortion it brings into the EEG signal. This category includes: the improvement in the root mean square error (RMSE), $\Delta RMSE$, improvement in power spectral density, ΔPSD, improvement in signal to noise and distortion ratio (SNDR), $\Delta SNDR$, difference in correlation, ΔCor, difference in coherence, ΔCoh. Correlation function is used to measure the similarity between two signals in time domain. While coherence is used to measure similarity in the frequency domain. It is worth mentioning that these metrics can be used only in the case of synthesized data (fully simulated and semi simulated). To calculate the different considered metrics, let denote x_{ref} the reference signal (clean signal) of length, N, x_{art} the synthesized artifactual signal, and x_{rec} the reconstructed signal (denoised signal). In addition, we defined the error signal before and after artifact removal respectively as follows:

$$e_{br}(n) = x_{art}(n) - x_{ref}(n), \tag{14}$$

$$e_{ar}(n) = x_{rec}(n) - x_{ref}(n), \tag{15}$$

- $\lambda(\%)$: the reduction in artifact in %, can be calculated as follows:

$$\lambda(\%) = 100\left(1 - \frac{R_{ref} - R_{ref-rec}}{R_{ref} - R_{ref-art}}\right), \tag{16}$$

where $R_{ref}, R_{ref-rec}$, and $R_{ref-art}$ are the autocorrelation coefficient of the reference signal at time lag 1, the cross-correlation coefficient between x_{ref} and x_{rec}, and the cross-correlation coefficient between x_{ref} and x_{art}, respectively.

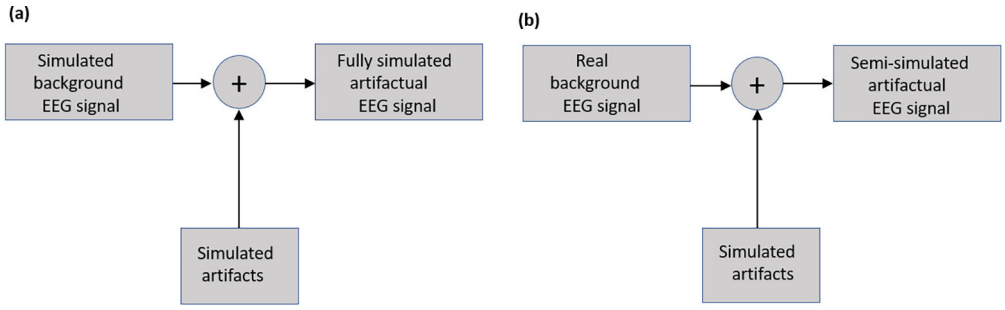

Fig. 3. Process to simulate the artifactual EEG signals: fully simulated artifactual EEG signal (a) and semi-simulated artifactual EEG signal (b).

Table 3
Description of the used databases.

	No. of EEG channels	Sampling frequency (Hz)	Bandpass filtered bandwidth	No. of subjects	Additional channels	No. of classes
EEG/EOG (Klados and Bamidis, 2016)	19	200	0.5–40	27	EOG	-
Dataset-1 (Blankertz et al., 2007)	64	1000	0.05–200 Hz	7	None	2
Dataset-2a (Brunner et al., 2008)	22	250	0.5–100	9	3 EOG	4

- ΔSNR: the improvement in SNR, is defined as the difference between SNR before artifact removal and SNR after artifact removal:

$$\Delta SNR = 10\log_{10}\left(\frac{\sigma^2_{ref}}{\sigma^2_{e_{br}}}\right) - 10\log_{10}\left(\frac{\sigma^2_{ref}}{\sigma^2_{e_{ar}}}\right), \tag{17}$$

where $\sigma^2_{ref}, \sigma^2_{e_{br}}$, and $\sigma^2_{e_{ar}}$ are the variances of x_{ref}, e_{br}, and e_{ar}, respectively.

- $\Delta RMSE$: the improvement in RMSE, is calculated from the RMSE before artifact removal, $RMSE_{br}$, and RMSE after artifact removal, $RMSE_{ar}$, as follows:

$$\begin{aligned} \Delta RMSE(\%) &= 100\left(\frac{RMSE_{br} - RMSE_{ar}}{RMSE_{br}}\right), \\ RMSE_{br} &= \sqrt{\frac{1}{N}\sum_{n=1}^{n=N}(e_{br}(n))^2}, \\ RMSE_{ar} &= \sqrt{\frac{1}{N}\sum_{n=1}^{n=N}(e_{ar}(n))^2}. \end{aligned} \tag{18}$$

- $\Delta PSD - dis$: improvement in PSD distortion, is calculated using the following formula:

$$\Delta PSD - dis(\%) = 100\left(\frac{PSD_{dis_{br}} - PSD_{dis_{ar}}}{PSD_{dis_{br}}}\right), \tag{19}$$

where $PSD_{dis_{br}}$ and $PSD_{dis_{ar}}$ are the PSD before and after artifact removal respectively, and are calculated from $PSD_{x_{ref}}, PSD_{x_{art}}$, and $PSD_{x_{rec}}$ which are the PSD for x_{ref}, x_{art} and x_{rec}:

$$PSD_{dis_{br}} = \frac{\sum (PSD_{x_{art}})^2}{\sum \left(PSD_{x_{ref}}\right)^2}, \tag{20}$$

$$PSD_{dis_{ar}} = \frac{\sum (PSD_{x_{rec}})^2}{\sum \left(PSD_{x_{ref}}\right)^2}. \tag{21}$$

- $\Delta SNDR$, improvement in SNDR, is calculated from $SNDR$ before and after artifact removal, $SNDR_{br}$ and $SNDR_{ar}$:

$$\Delta SNDR = \frac{1}{N}\sum (SNDR_{ar} - SNDR_{br}), \tag{22}$$

with

$$SNDR_{ar} = 10\log_{10}\left(\frac{PSD_{x_{rec}}}{PSD_{e_{br}}}\right), \tag{23}$$

$$SNDR_{ar} = 10\log_{10}\left(\frac{PSD_{x_{rec}}}{PSD_{e_{ar}}}\right), \tag{24}$$

where $PSD_{e_{br}}$ and $PSD_{e_{ar}}$ are the PSD of the error signal e_{br} and e_{ar}, respectively.

- ΔCor, difference in correlation is calculated from R_{br} which is the cross-correlation coefficient between x_{ref} and x_{art}, and R_{ar} which is the cross-correlation coefficient between x_{ref} and x_{rec}:

$$\Delta Cor(\%) = 100\left(\frac{R_{br} - R_{ar}}{R_{br}}\right), \tag{25}$$

- ΔCoh, difference in coherence is obtained from Coh_{br} which is the coherence coefficient between x_{ref} and x_{art}, and Coh_{ar} which is the coherence coefficient between x_{ref} and x_{rec}:

$$\Delta Coh(\%) = 100\left(\frac{Coh_{br} - Coh_{ar}}{Coh_{br}}\right), \tag{26}$$

where Coh_{br} and Coh_{ar} are calculated using the following formula:

$$Coh_{br} = \frac{G^2_{x_{ref}x_{art}}}{G_{x_{ref}x_{ref}}G_{x_{art}x_{art}}}, \tag{27}$$

$$Coh_{ar} = \frac{G^2_{x_{ref}x_{rec}}}{G_{x_{ref}x_{ref}}G_{x_{rec}x_{rec}}}, \tag{28}$$

where $G_{x_{ref}x_{ref}}, G_{x_{art}x_{art}}$ and $G_{x_{rec}x_{rec}}$ are the auto-spectral density of x_{ref}, x_{art} and x_{rec} respectively, and $G_{x_{ref}x_{art}}$ and $G_{x_{ref}x_{rec}}$ are the cross-spectral density between x_{ref} and x_{art} and between x_{ref} and x_{rec} respectively.

4.3. Results

4.3.1. Scenario 1

In this scenario the proposed method was applied to fully simulated artifactual EEG data. Fig. 4 shows the artifactual signal, the calculated statistical parameters in each epoch and the recon-

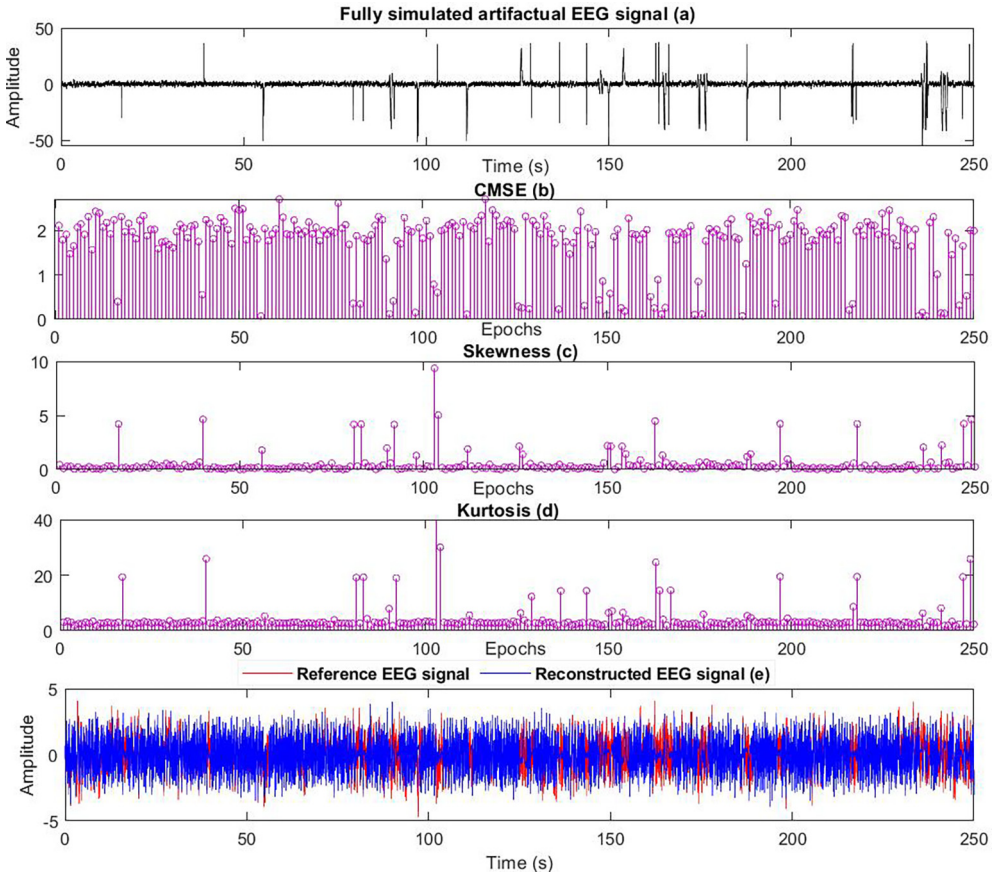

Fig. 4. Results of the proposed method on the scenario 1: Fully simulated artifactual EEG signal (a), composite multiscale entropy (CMSE) of different epochs (b), Skewness of different epochs (c), kurtosis of different epochs (d), and reference (red line) and reconstructed (blue line) EEG signals (e). Note data the time window of an epoch is 1 s and its length is N = fs = 256.

structed EEG signal along with the reference signal. The duration of an epoch is 1 *s*. Figs. 4(b-c) show that the three considered statistical features are capable to detect the six considered artifacts. An epoch is artifactual if *cmse* is low, skewness (s) and kurtosis (k) are high. The first proposed FIS1 uses these features to decide if an epoch is artifactual or not. Once an epoch is artifactual, the SWT is applied on this epoch, then the use of the proposed FISs2 in the second stage allows to detect the contaminated SWT coefficients. The universal threshold is used to denoise these coefficients. Finally, the use of the inverse SWT allows to obtain the artifact-free epoch. Fig. 4(e) shows that the proposed algorithm can significantly detect and remove all different types of the six considered artifacts. For more detail, Fig. 5 shows some artifactual epochs with the six types of the considered artifacts. Each plot illustrates the type of artifact with which an epoch is contaminated, the statistical features of the artifactual EEG segment, the reference and reconstructed EEG segments. These plots show that an artifact can be present in more than one epoch. In addition, these illustrations demonstrate that the proposed algorithm can significantly reduce all different types of artifacts. The proposed algorithm is also evaluated quantitatively using several metrics (Section 4.2). As discussed in Section 4.2 these metrics allows to quantify the obtained results in terms of artifact removal and its effect on the neural activities (in term of distortion). Table 4 shows the quantitative performance of the proposed algorithm.

4.3.2. Scenario 2

Fig. 6 illustrates the obtained results in the second scenario. This figure shows the semi-simulated artifactual EEG signal, statistical features at different epochs, reference and reconstructed EEG signals. It is obvious from these plots that artifacts are detected and corrected. Table 4 shows the quantitative performance of the proposed algorithm in the second scenario.

4.3.3. Scenario 3

In this scenario the proposed method is applied to fully real EEG data. Figs. 7 shows the results obtained by the proposed method in the case of EEG data from BCI competition dataset IV. This data are contaminated with the ocular artifacts. It is clear from this figure that the proposed method can detect and significantly reduce the EEG artifacts.

4.4. Comparison of the proposed method with other methods

As discussed previously, the artifact removal from EEG signal is an active area of research since several methods have been developed to remove the artifacts from the EEG signal without much affecting the neural activity. Table 5 shows the comparison between the proposed algorithm with some artifact removal methods founded in the literature mainly ICA-Th (independent component analysis-Thresholding), wICA (wavelet ICA) (Castellanos and Makarov, 2006), wCCA (wavelet canonical correlation analysis) (Mowla et al., 2015; Raghavendra and Dutt, 2011), EMD-Th (empirical mode decomposition-Thresholding), EMD-ICA (Mijović et al., 2010), EMD-CCA (Chen et al., 2014), wPM (probability mapping-wavelet) (Islam et al., 2021).

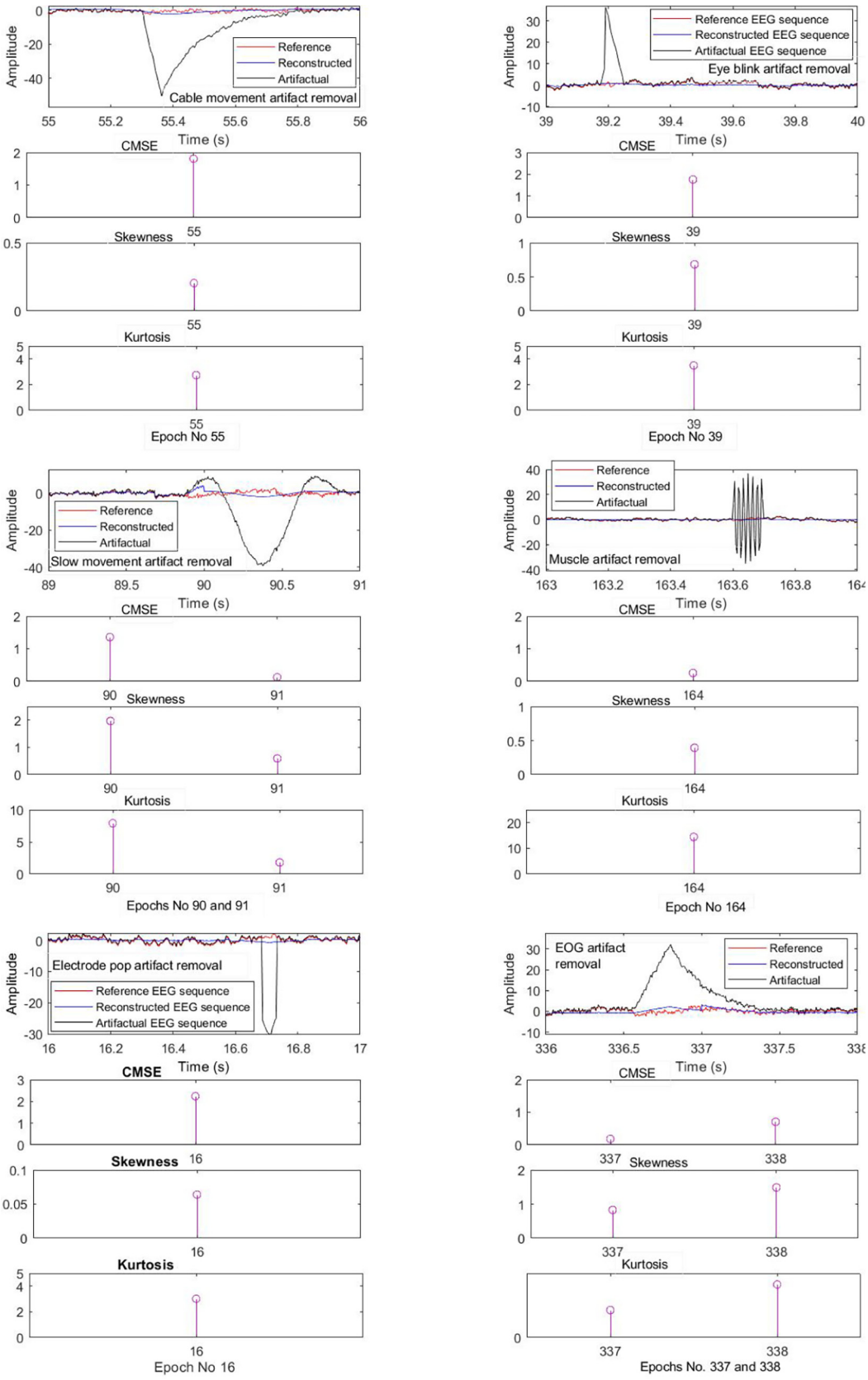

Fig. 5. EEG epochs contaminated with six different types of simulated artifacts that mimicking real artifacts in EEG recordings (Eye blink, cable movement, muscle artifact, slow movement artifact, EOG artifact, and electrode pop). These plots show also the statistical features of the corrupted EEG segment, the reference and the reconstructed EEG segments.

Table 4
The performance of the proposed algorithm in the first and the second scenarios.

Scenarios	λ (%)	ΔSNR (dB)	$\Delta RMSE$ (%)	ΔPSD (%)	$\Delta SNDR$ (dB)	ΔCor (%)	ΔCoh (%)
Scenario 1	77.94	6.97	54.02	83.15	6.71	264.79	85.27
Scenario 2	67.93	5.43	44.86	−70.48	8.88	229.75	82.73

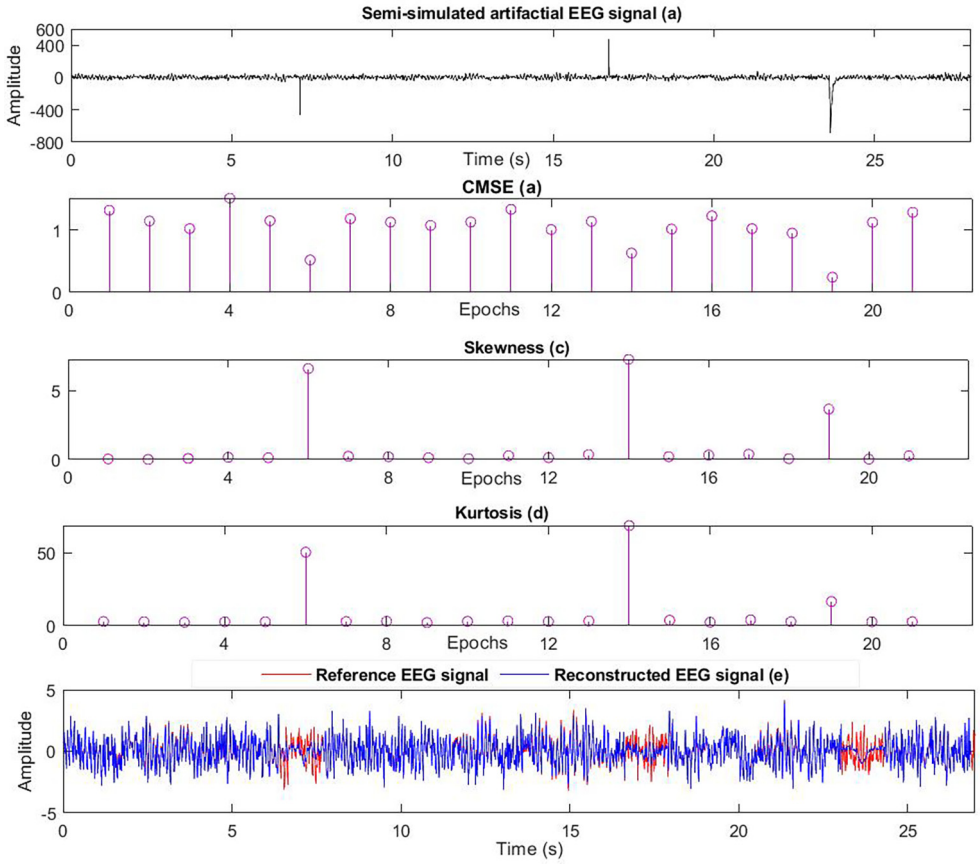

Fig. 6. Performance of the proposed method in the second scenario: contaminated EEG signal (a), statistical features of different epochs of the Contaminated EEG (b-d), reconstructed along with the reference EEG signal.

4.5. Discussion

The artifact handling is a real problem, especially with modern applications of EEG signal in medical and BCI domains. In this paper, we propose an automatic algorithm for detecting and removing artifacts from EEG signal. This algorithm is able to perform for both single and multiple channels, does not need prior information, does not need reference channel, is independent of artifact types, and does not require any user parameter adjustment. The presented results show that the proposed method is very efficient in the three considered scenarios. The performance comparison of this method with some of state-of-the-art artifact removal methods shows (see Table 5):

- That the proposed method outperforms some existing ones for artifact handling from single channel EEG signals (EMD-Th, EMD-ICA, EMD-CCA and wPM), in scenario 1 and 2. While, comparing the performance of our algorithm with multichannel artifact handling methods shows that the proposed method performs better than wICA, and wCCA. One can notice a slight difference in the performance between our algorithm and ICA-Th. The proposed method can be then proposed as an efficient and robust method for artifact handling especially in the case of single BCI applications.
- In term of computational time, Table 5 shows that our algorithm required more time than ICA-Th, wICA, wCCA and wPM. But it is worth mentioning that the proposed algorithm process the signal epoch by epoch. Then the time needs for this algorithm to processes an epoch of EEG signal is about 0.16 s in the first scenario and 0.11 s in the second scenario. This makes the proposed method feasible to be implemented in an FPGA or DSP chip to perform digital signal processing.
- The negative value means a negative change in the quality of the EEG signal while the positive value means that there is an improvement in the quality of the signal. In the first scenario our algorithm provides an improvement in the quality of the signal, e.g. in term of correlation, our method provides an improvement of 264.79. However, in the second scenario the proposed method brings a low deterioration in term of PSD compared to ICA-Th, wICA, wCCA, EMD-Th, EMD-ICA, EMD-CCA.
- One of the main advantages of the proposed method is its high value of λ which means that this method can detect and remove an important amount of artifacts.

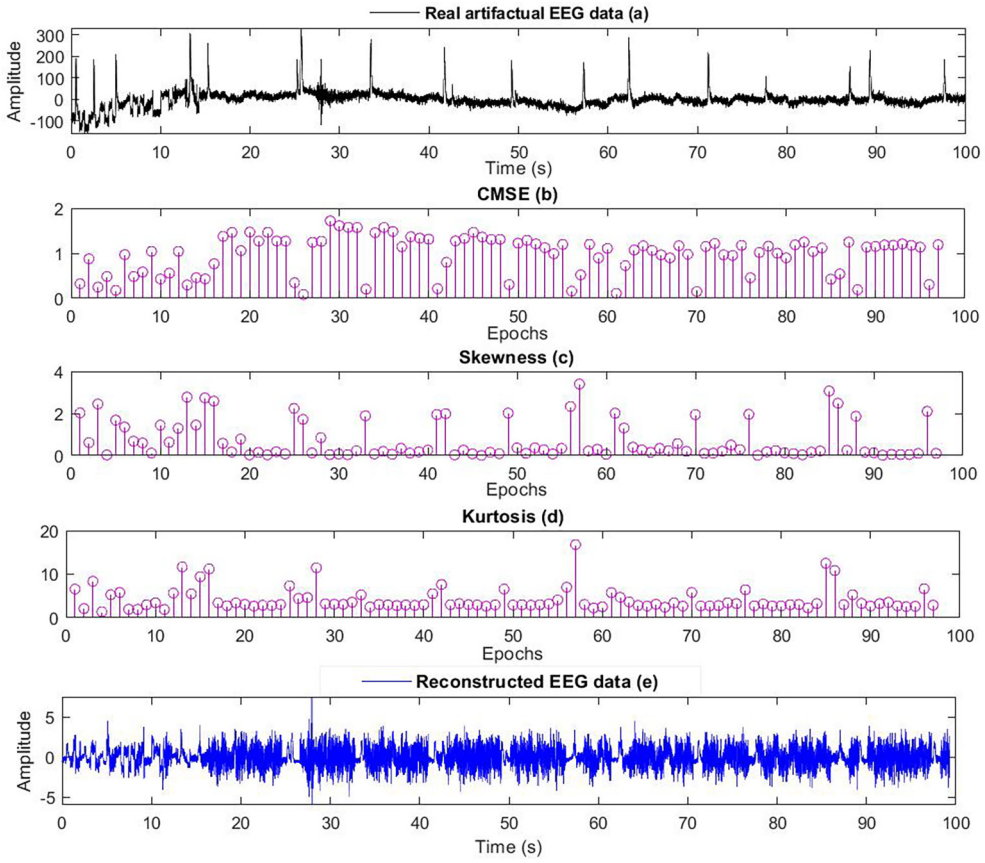

Fig. 7. Performance evaluation of the proposed method in the case of fully real contaminated EEG data from BCI competition IV dataset: EEG signal contaminated with ocular artifacts (a), statistical features of different EEG epochs (b-d), and reconstructed EEG data (e).

Table 5
The performance of the proposed method in the first and the second scenarios.

Scenarios	Methods	λ (%)	ΔSNR (*dB*)	$\Delta RMSE$ (%)	ΔPSD (%)	$\Delta SNDR$ (*dB*)	ΔCor (%)	ΔCoh (%)	time (s))	No. Of channels
Scenario 1	ICA-Th	79.19	7.57	56.19	86.53	5.44	265.29	65.88	0.20	14
	wICA	36.55	3.51	25.97	−95.50	3.44	145.10	62.15	0.42	14
	wCCA	30.04	2.02	18.43	−41.77	2.62	99.50	31.20	0.38	14
	EMD-Th	18.86	0.96	10.19	−251.04	3.21	58.05	52.15	324.01	1
	EMD-ICA	74.75	6.99	52.60	82	5.74	222.30	63.58	382.97	1
	EMD-CCA	36.02	2.13	20.87	−80.83	5.14	103.73	64.44	420.19	1
	wPM	−17.36	−0.63	−7.89	94.21	−0.92	−28.57	−68.80	1.55	1
	Proposed	77.94	6.97	54.02	83.15	**6.71**	264.79	**85.27**	5.53	1
Scenario 2	ICA-Th	68.13	5.45	45.09	−144.13	0.66	229.21	58.50	0.22	15
	wICA	35.40	2.89	23.63	−167.46	1.84	144.53	48.31	0.42	15
	wCCA	8.51	0.74	6.26	- 475.35	0.86	60.17	8.91	0.35	15
	EMD-Th	10.57	0.66	6.31	−2835.7	3.58	66.51	54.72	849.12	1
	EMD-ICA	54.84	4.37	36.04	−188.70	3.32	227.91	69.13	860.43	1
	EMD-CCA	−28.77	−1.05	−13.50	−739.65	6.78	−88.39	−14.74	918.02	1
	wPM	4.56	0.27	2.66	92.67	7.81	33.62	44.88	1.31	1
	Proposed	67.93	5.43	44.86	−70.48	**8.88**	**229.75**	**82.73**	3.74	1

- In term of stability, our method provides the same result from one run to another, however the performance and the computational time of ICA, EMD and other related methods are not stable. In addition, EMD, EMD-ICA and EMD-CCA can not perform on different kinds of data.

Despite the potential advantages of the proposed approach, there may be one major limitation that could be addressed in the future work. The complexity of the proposed method increases linearly with number of EEG channels due to the increase number of the stationary wavelet filters. It is then necessary to attempt to reduce the complexity of the proposed algorithm in term of wavelet filters in the case of a large number of EEG channels. In addition, there are some limitations of this study. We expect to test the proposed algorithm with more EEG data from different paradigms from different users in order to tune properly different parameters of our algorithm, namely the parameters of the fuzzy logic system and duration of an incoming EEG segment.

5. Conclusions

The artifacts are a serious problem in the medical field and BCI applications of the EEG signal. They are added to the neural activities and changed its morphology: 1) in the medical field, they lead to misinterpretation of the EEG signal. This can conduct to treat incorrectly the patients; and 2) in BCI applications, they cause unintentional control of devises. Several artifacts handling methods have been developed based on different advanced signal processing, machine learning and deep learning techniques. However, until the moment of writing of these lines, there is no an efficient method that allows to resolve completely the problem of EEG artifacts. This paper was concerned with artifact handling from EEG signals. An automatic algorithm for detecting and removing of artifacts from EEG signal was proposed. This algorithm was assessed in three scenarios: fully simulated artifactual EEG, semi-simulated artifactual EEG and fully real artifactual EEG recordings. The results obtained in these three scenarios show the superior efficacy of the proposed algorithm. This method was tested both in term of artifact removal and signal distortion. A comparative study was conducted between the performance of our algorithm and some widely known artifact removal methods. This comparison showed that the proposed method performs better than all the considered artifact removal methods from single channel EEG recording (i.e. wPM, EMD-Th, EMD-ICA and EMD-CCA) and also some artifact removal methods from multichannel EEG recordings (i.e. wICA and wCCA). However, a slight difference between the proposed method and ICA-Th was mentioned. The comparison of different methods in term of computational time proved that the proposed algorithm can achieve a high performance in an acceptable time. The necessary time to handle one second of data by our method is 0.16 s in the first scenario and 0.11 s in the second one. The proposed method can be useful to improve the quality of the EEG signal, especially in the new applications of BCI systems (i.g. neural prostheses) and also in the medical domain (i.g. health care monitoring), since this method can be implemented either for offline or online EEG artifact removal, with single channel or multi-channel EEG recording. This research is worth continuing further in order to improve and increase the EEG signal to noise ratio and overcome the limitations. Further work will focus on: 1) reduce the model complexity in the case of a large number of EEG recording channels; and 2) assessing the proposed approach with EEG data from different users and paradigms.

Declaration of Competing 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.

Acknowledgment

This work was supported by the Ministry of High Education, Scientific Research and Innovation, the Digital Development Agency (DDA) and the CNRST of Morocco under the number 14/FSA/2021. The APC is sponsored by IUB Sponsored Research Grant #2021-SETS-07)"

References

Abdulkader, S.N., Atia, A., Mostafa, M.-S.M., 2015. Brain computer interfacing: Applications and challenges. Egyptian Inform. J. 16, 213–230.

B. Azzerboni, M. Carpentieri, F. La Foresta, F. Morabito, Neural-ica and wavelet transform for artifacts removal in surface emg, in: 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541), volume 4, IEEE, 2004, pp. 3223–3228.

Bengtsson, T., Cavanaugh, J.E., 2006. An improved akaike information criterion for state-space model selection. Comput. Stat. Data Anal. 50, 2635–2654.

Blankertz, B., Dornhege, G., Krauledat, M., Müller, K.-R., Curio, G., 2007. The non-invasive berlin brain–computer interface: fast acquisition of effective performance in untrained subjects. NeuroImage 37, 539–550.

Borowicz, A., 2018. Using a multichannel wiener filter to remove eye-blink artifacts from eeg data. Biomed. Signal Process. Control 45, 246–255.

Brunner, C., Leeb, R., Müller-Putz, G., Schlögl, A., Pfurtscheller, G., 2008. Bci competition 2008–graz data set a, Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces). Graz University of Technology 16, 1–6.

Castellanos, N.P., Makarov, V.A., 2006. Recovering eeg brain signals: Artifact suppression with wavelet enhanced independent component analysis. J. Neurosci. Methods 158, 300–312.

Chen, X., He, C., Peng, H., 2014. Removal of muscle artifacts from single-channel eeg based on ensemble empirical mode decomposition and multiset canonical correlation analysis. J. Appl. Math. 2014, 1–10.

Chen, X., Xu, X., Liu, A., Lee, S., Chen, X., Zhang, X., McKeown, M.J., Wang, Z.J., 2019. Removal of muscle artifacts from the eeg: a review and recommendations. IEEE Sens. J. 19, 5353–5368.

De Clercq, W., Vergult, A., Vanrumste, B., Van Paesschen, W., Van Huffel, S., 2006. Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram. IEEE Trans. Biomed. Eng. 53, 2583–2587.

Gao, J., Zheng, C., Wang, P., 2010. Online removal of muscle artifact from electroencephalogram signals based on canonical correlation analysis. Clinical EEG Neurosci. 41, 53–59.

Ghorbanian, P., Devilbiss, D.M., Simon, A.J., Bernstein, A., Hess, T., Ashrafiuon, H., 2012. Discrete wavelet transform eeg features of alzheimer's disease in activated states. In: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, pp. 2937–2940.

Guarnieri, R., Marino, M., Barban, F., Ganzetti, M., Mantini, D., 2018. Online eeg artifact removal for bci applications by adaptive spatial filtering. J. Neural Eng. 15, 056009.

Harrison, L., Penny, W.D., Friston, K., 2003. Multivariate autoregressive modeling of fmri time series. Neuroimage 19, 1477–1491.

Hong-Ye, G., 1998. Wavelet shrinkage denoising using the non-negative garrote. J. Comput. Graphical Stat. 7, 469–488.

Hou, H.-R., Meng, Q.-H., Zeng, M., Sun, B., 2017. Improving classification of slow cortical potential signals for bci systems with polynomial fitting and voting support vector machine. IEEE Signal Process. Lett. 25, 283–287.

Hu, J., Wang, C.-S., Wu, M., Du, Y.-X., He, Y., She, J., 2015. Removal of eog and emg artifacts from eeg using combination of functional link neural network and adaptive neural fuzzy inference system. Neurocomputing 151, 278–287.

Hussein, H.M., Abdalla, K.K., Mahmood, A.S., 2022. Eegs signals artifact rejection algorithm by signal statistics and independent components modification, in. In: Mobile Computing and Sustainable Informatics. Springer, pp. 275–290.

A. Hyvarinen, J. Karhunen, E. Ojal, Independent component analysis, in: Hilbert-Huang transform and its applications, Wiley- Interscience, Iune 2001.

Islam, M.K., Rastegarnia, A., Yang, Z., 2015. A wavelet-based artifact reduction from scalp eeg for epileptic seizure detection. IEEE J. Biomed. Health Inform. 20, 1321–1332.

Islam, M.K., Rastegarnia, A., Yang, Z., 2016. Methods for artifact detection and removal from scalp eeg: A review. Neurophysiologie Clinique/Clinical Neurophysiol. 46, 287–305.

Islam, M.K., Ghorbanzadeh, P., Rastegarnia, A., 2021. Probability mapping based artifact detection and removal from single-channel eeg signals for brain-computer interface applications. J. Neurosci. Methods 109249.

Kanoga, S., Hoshino, T., Asoh, H., 2020. Independent low-rank matrix analysis-based automatic artifact reduction technique applied to three bci paradigms. Front. Human Neurosci. 14, 173.

Kilicarslan, A., Vidal, J.L.C., 2019. Characterization and real-time removal of motion artifacts from eeg signals. J. Neural Eng. 16, 056027.

Klados, M.A., Bamidis, P.D., 2016. A semi-simulated eeg/eog dataset for the comparison of eog artifact rejection techniques. Data in brief 8, 1004–1006.

Klados, M.A., Papadelis, C., Braun, C., Bamidis, P.D., 2011. Reg-ica: a hybrid methodology combining blind source separation and regression techniques for the rejection of ocular artifacts. Biomed. Signal Process. Control 6, 291–300.

La Rosa, A.B., Pereira, P.T., Ücker, P., Paim, G., da Costa, E.A., Bampi, S., Almeida, S., 2021. Exploring nlms-based adaptive filter hardware architectures for eliminating power line interference in eeg signals. Circuits, Systems, Signal Process. 40, 3305–3337.

Mahapatra, A.G., Singh, B., Wagatsuma, H., Horio, K., 2018. Epilepsy eeg classification using morphological component analysis. EURASIP J. Adv. Signal Process. 2018, 1–12.

Mammone, N., La Foresta, F., Morabito, F.C., 2011. Automatic artifact rejection from multichannel scalp eeg by wavelet ica. IEEE Sens. J. 12, 533–542.

Mannan, M.M.N., Jeong, M.Y., Kamran, M.A., 2016. Hybrid ica–regression: Automatic identification and removal of ocular artifacts from electroencephalographic signals. Front. Human Neurosci. 10, 193.

Martis, R.J., Acharya, U.R., Tan, J.H., Petznick, A., Yanti, R., Chua, C.K., Ng, E.K., Tong, L., 2012. Application of empirical mode decomposition (emd) for automated detection of epilepsy using eeg signals. Int. J. Neural Syst. 22, 1250027.

Mijović, B., De Vos, M., Gligorijević, I., Taelman, J., Van Huffel, S., 2010. Source separation from single-channel recordings by combining empirical-mode decomposition and independent component analysis. IEEE Trans. Biomed. Eng. 57, 2188–2196.

Molla, M.K.I., Tanaka, T., Rutkowski, T.M., 2012. Multivariate emd based approach to eog artifacts separation from eeg. In: 2012 IEEE international conference on acoustics, speech and signal processing (ICASSP), Kyoto, Japan. IEEE, pp. 653–656.
Molla, M.K.I., Islam, M.R., Tanaka, T., Rutkowski, T.M., 2012. Artifact suppression from eeg signals using data adaptive time domain filtering. Neurocomputing 97, 297–308.
Mowla, M.R., Ng, S.-C., Zilany, M.S., Paramesran, R., 2015. Artifacts-matched blind source separation and wavelet transform for multichannel eeg denoising. Biomed. Signal Process. Control 22, 111–118.
Muller-Putz, G.R., Pfurtscheller, G., 2007. Control of an electrical prosthesis with an ssvep-based bci. IEEE Trans. Biomed. Eng. 55, 361–364.
Nguyen, H.-A.T., Musson, J., Li, F., Wang, W., Zhang, G., Xu, R., Richey, C., Schnell, T., McKenzie, F.D., Li, J., 2012. Eog artifact removal using a wavelet neural network. Neurocomputing 97, 374–389.
Ocak, H., 2009. Automatic detection of epileptic seizures in eeg using discrete wavelet transform and approximate entropy. Expert Syst. Appl. 36, 2027–2036.
Pfurtscheller, G., Brunner, C., Schlögl, A., Da Silva, F.L., 2006. Mu rhythm (de) synchronization and eeg single-trial classification of different motor imagery tasks. NeuroImage 31, 153–159.
Phadikar, S., Sinha, N., Ghosh, R., 2020. Automatic eyeblink artifact removal from eeg signal using wavelet transform with heuristically optimized threshold. IEEE J. Biomed. Health Inform. 25, 475–484.
Placidi, G., Cinque, L., Polsinelli, M., 2021. A fast and scalable framework for automated artifact recognition from eeg signals represented in scalp topographies of independent components. Comput. Biol. Med. 132, 104347.
Radüntz, T., Scouten, J., Hochmuth, O., Meffert, B., 2017. Automated eeg artifact elimination by applying machine learning algorithms to ica-based features. J. Neural Eng. 14, 046004.
Raghavendra, B., Dutt, D.N., 2011. Wavelet enhanced cca for minimization of ocular and muscle artifacts in eeg, World Academy of Science. Eng. Technol. 57, 1027–1032.
Ranjan, R., Sahana, B.C., Bhandari, A.K., 2022. Cardiac artifact noise removal from sleep eeg signals using hybrid denoising model. IEEE Trans. Instrum. Meas.
Safieddine, D., Kachenoura, A., Albera, L., Birot, G., Karfoul, A., Pasnicu, A., Biraben, A., Wendling, F., Senhadji, L., Merlet, I., 2012. Removal of muscle artifact from eeg data: comparison between stochastic (ica and cca) and deterministic (emd and wavelet-based) approaches. EURASIP J. Adv. Signal Processing 2012, 1–15.
Schalk, G., Mellinger, J., 2010. A practical guide to brain–computer interfacing with BCI2000: General-purpose software for brain-computer interface research, data acquisition, stimulus presentation and brain monitoring. Springer Science & Business Media.
Seneviratne, U., Mohamed, A., Cook, M., D'Souza, W., 2013. The utility of ambulatory electroencephalography in routine clinical practice: a critical review. Epilepsy Res. 105, 1–12.
Shahbakhti, M., Maugeon, M., Beiramvand, M., Marozas, V., 2019. Low complexity automatic stationary wavelet transform for elimination of eye blinks from eeg. Brain Sci. 9, 352.
Shahbakhti, M., Beiramvand, M., Rejer, I., Augustyniak, P., Broniec-Wójcik, A., Wierzchon, M., Marozas, V., 2021. Simultaneous eye blink characterization and elimination from low-channel prefrontal eeg signals enhances driver drowsiness detection. IEEE J. Biomed. Health Inform. 26, 1001–1012.
Shahbakhti, M., Rodrigues, A.S., Augustyniak, P., Broniec-Wójcik, A., Sološenko, A., Beiramvand, M., Marozas, V., 2021. Swt-kurtosis based algorithm for elimination of electrical shift and linear trend from eeg signals. Biomed. Signal Process. Control 65, 102373.
Sharma, R., Pachori, R.B., Acharya, U.R., 2015. An integrated index for the identification of focal electroencephalogram signals using discrete wavelet transform and entropy measures. Entropy 17, 5218–5240.
Shoker, L., Sanei, S., Chambers, J., 2005. Artifact removal from electroencephalograms using a hybrid bss-svm algorithm. IEEE Signal Process. Lett. 12, 721–724.
Shukla, P.K., Roy, V., Shukla, P.K., Chaturvedi, A.K., Saxena, A.K., Maheshwari, M., Pal, P.R., 2021. An advanced eeg motion artifacts eradication algorithm. Computer J.
Sweeney, K.T., McLoone, S.F., Ward, T.E., 2012. The use of ensemble empirical mode decomposition with canonical correlation analysis as a novel artifact removal technique. IEEE Trans. Biomed. Eng. 60, 97–105.
Tibdewal, M.N., Fate, R., Mahadevappa, M., Ray, A., 2015. Detection and classification of eye blink artifact in electroencephalogram through discrete wavelet transform and neural network. In: 2015 International Conference on Pervasive Computing (ICPC), Pune, India. IEEE, pp. 1–6.
Wang, Z., Xu, P., Liu, T., Tian, Y., Lei, X., Yao, D., 2014. Robust removal of ocular artifacts by combining independent component analysis and system identification. Biomed. Signal Process. Control 10, 250–259.
Wu, S.-D., Wu, C.-W., Lin, S.-G., Lee, K.-Y., Peng, C.-K., 2014. Analysis of complex time series using refined composite multiscale entropy. Phys. Lett. A 378, 1369–1374.
Yasoda, K., Ponmagal, R., Bhuvaneshwari, K., Venkatachalam, K., 2020. Automatic detection and classification of eeg artifacts using fuzzy kernel svm and wavelet ica (wica). Soft. Comput. 24, 16011–16019.
Yeung, N., Bogacz, R., Holroyd, C.B., Cohen, J.D., 2004. Detection of synchronized oscillations in the electroencephalogram: an evaluation of methods. Psychophysiology 41, 822–832.
Yong, X., Ward, R.K., Birch, G.E., 2009. Artifact removal in eeg using morphological component analysis. In: 2009 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, pp. 345–348.
Zhang, J., Sanderson, A.C., 2007. Jade: Self-adaptive differential evolution with fast and reliable convergence performance. In: 2007 IEEE congress on evolutionary computation. IEEE, pp. 2251–2258.