ISSN 1330-3651 (Print), ISSN 1848-6339 (Online) https://doi.org/10.17559/TV-20240728001884
Original scientific paper

Enhanced Motion Artifact Mitigation in ECG Signals using Nonlinear Autoregressive Networks with Cat Swarm Optimization and Fractional Calculus

Ibrahim ALMOHIMEED, Ahmed Farag Salem BABETAT, Mohammed Nasser ALMRAIKHI, Mohamed Yacin SIKKANDAR*, Bader Dhaidan OWAYYID OWAYJAH, Ali Abdullah ALMUKIL, Abdulrhman Abdulaziz ALMAJHAD

Abstract: The Motion artifacts in ECG data can lead to inaccurate circulatory state analysis. Recently, there has been growing interest in methods to mitigate motion distortion in ECG signals. One major challenge is to reduce motion artifacts without affecting the underlying ECG signal, as the motion artifacts and the ECG signal often overlap. Adaptive noise cancellers have proven effective in reducing motion artifacts, provided that an appropriate noise reference, which correlates with the noise in the ECG signal, is available. However, the correlation between motion distortion and the noise reference is not always consistent, and using an inappropriate noise reference can contaminate the ECG data. In this research, we present NARX_CSOFC, an advanced approach that significantly enhances the capabilities of the nonlinear autoregressive network with exogenous inputs (NARX) architecture by incorporating nonlinear combinations of input variables to globally estimate any nonlinear function. The method employs cat swarm optimization with fractional calculus (CSOFC) to find optimal solutions. The proposed NARX_CSOFC demonstrates substantial improvements in artifact reduction: achieving a 12 dB improvement for ECG artifacts, a 16 dB improvement for EMG artifacts, and a 15 dB improvement for random noise artifacts, as measured by Signal-to-Noise Ratio (SNR) and Mean Squared Error (MSE) compared to existing techniques.

Keywords: cat swarm optimization; ECG signal processing; fractional calculus; motion artifacts; nonlinear autoregressive network

1 INTRODUCTION

Electroencephalogram (EEG) signals are commonly used to study the electrical processes of the human brain. For effective diagnosis of neurological disorder present in EEG signal effective analysis of distortion present in the EEG signal is required for denoising [1]. The main problem that arises in removing artefacts from EEG signals is that the amplitude of the signal is less than the amplitude of the artefacts signal and these two spectrums overlap means they are at the same frequency. Removing artefact components without disturbing signal components is a great challenge. Various heuristic or regressive methods are used to search for the optimal coefficients of adaptive filters to improve the quality of EEG signals. Depending upon the type of noise, the number of channels and application suitable techniques has to be chosen. Ocular, muscular artefacts were denoised efficiently by ICA [2] and Detrended [3] based methods respectively. Heuristic methods are more efficient than traditional methods [4]. Empirical mode decomposition combination methods are more suitable for single-channel as well as multi-channel applications [5].

There are several methods suggested to eliminate eye artefacts from the contaminated EEG data. The wavelet transform approach, filtration, regression, and blind source separation are among the standard methods used to remove eye artefacts in general. Unlike other techniques, independent component analysis (ICA), one of the most popular contaminant elimination algorithms, does not require a reference channel [6]. These approaches can be contrasted in specific based on the reference source, process automation, and adaptability in an online environment. There is not a widely accepted traditional method for the elimination of eye abnormalities, despite mixed methods that incorporate several methods to lessen these disadvantages. With the common use of machine learning (ML), numerous studies in the area of eliminating EEG artefacts have been documented in addition to traditional techniques. The most popular methods for handling EEG artefacts are Support Vector Machines (SVM) and Neural Networks (NN) [7]. Many researches combined the use of ICA and SVM. A MATLAB Toolbox was created by Lawhern et al. that employs SVM categorization of EEG data. This toolkit uses an autoregressive model to extract and describe features from EEG signals and effectively separates fake data [8]. Additional ML research includes the Bayesian model [10], flexible reasoning system [9], and genetic programming [11].

Although ICA is a widely used technique, item detection takes professional observation. In terms of efficiency and effectiveness, ML methods may be regarded as superior to traditional methods in this situation. However, before creating an ML model, the feature vector needs to be carefully removed. These processes are time-consuming and necessitate the advice of a specialist [12]. As a result, these methods are unable to handle unprocessed data without pre-processing and expert assistance. The long-standing issues in machine learning studies have seen substantial development due to deep learning (DL). DL techniques use models that have been built up in various levels while analyzing the data to find the pertinent information. Consequently, the following are the accomplishments of this work:

In the present paper, using the ANC technique, a novel algorithm is suggested to eliminate heart, muscular, and eye artefacts from EEG data.

The construction of the nonlinear autoregressive network with exogenous inputs (NARX) design, which receives nonlinear combinations of input variables, significantly improves its ability to approximate any nonlinear function universally. The single-layer neural structure can produce nonlinear decision boundaries with nonlinear functional growth to create infinitely complicated decision regions.

Cat optimization algorithm plays a vital role in producing a better optimal solution and fractional calculus is included to incorporate previous weights in the updated solution of weights.

2 RELATED WORKS

In the last five years, there have been numerous publications. It suggests that the most widely used algorithms are those based on Blind Source Separation (BSS), particularly ICA. In recent years, many scholars have tended to favour the combination approach in order to improve the efficacy of tactics due to the limitations of singular methods like Regression and BSS. There is no agreed-upon best practice for all kinds of artefacts, despite the substantial study on artefact identification and elimination of EEG data that has been documented in the literature to date.

Using a k-nearest neighbor (kNN) algorithm and a long short-term memory (LSTM) network, [13] presents a stable technique that can autonomously identify and eliminate eyeblink and muscle distortions from EEG. The LSTM network receives the appropriate EEG frame after the kNN algorithm has identified the existence of artefacts, and it then removes the artefacts. The Competitive Swarm Dragonfly Algorithm (CSDA) is developed in [14] for categorizing the EEG data. Data enhancement is done to prepare the data for processing further. The muscular imaging EEG data is categorized using Deep Residual Network (DRN).

For the purpose of eliminating baseline noise in EEG data, various deep learning optimization techniques, structures, and designs with varying layer configurations were suggested in [15]. ReMAE, a MATLAB toolkit, is outlined in [16] as a method to eliminate muscle artefacts. This toolkit displays the results of each stage of the denoising process for EEG data. Detrended Fluctuation Analysis is used to accurately identify the mode of an EEG signal from a melancholy patient. In [17], an Adaptive Line Enhancer-based LMS method for denoising EEG data is developed in Verilog HDL and has effective results in terms of design and efficiency. An adaptive noise canceller with optimum weights determined by the oppositional whale algorithm (OWOA) and compared to LMS, RLS, GA, PSO, and WOA in [18] effectively removes white Gaussian noise-corrupted EEG data.

Its large database, poor SNR, and filtration are the major shortcomings in the study of EEG signal processing in many uses. A more effective BPNN method can be used to solve the analysis's issues [19]. ECG, white, EOG, and EMG appearing at various SNR levels were removed from EEG signals with high SNR using a combination of the Wide-Deep Echos State Networks (WDESN) and Uniform Search Particle Swarm Optimization (UPSO) methods [20]. Eye movement locations are located in EEG data using Variational Mode Extraction (VME), and these areas are further filtered using DWT to remove EOG artefacts [21]. Fisher information and the objective Firefly Algorithm (FA) were combined to create a composite channel rating method in [22] Artificial with metaheuristic selection in biomedical selection.

First, the regularized common spatial pattern with aggregation (RCSPA) technique is used to derive spatial-temporal characteristics from the preprocessed brain signals. In order to rate channels using their individual weighted scores and Fisher information, a new Channel Set Relevance Index (CSRI) is created. The Regularized Support Vector Machine (RSVM) algorithm is used to distinguish between various MI-tasks using the RCSPA characteristics of highly rated channels. Principal component analysis (PCA) and independent component analysis (ICA) methods are used to identify the characteristics in [24]. An optimized deformable convolutional network (ODCN) is used to recognize eye abnormalities from EEG input data after gathering the characteristics. When artefacts are detected, the empirical mean curve decomposition (EMCD) and ODCN for noise minimization in EEG data are applied as part of the balancing technique to address the automatic multi-channel EEG signal categorization for various illnesses, in [25]. The various standard databases are first used to collect the EEG readings. Furthermore, the Haar DWT and spike detection take features from data. These characteristics are then submitted to the "Enhanced Deep Belief Network" with RBM layers, which uses the Adam-based Coyote Optimization Algorithm to tune the parameters of the DBN in order to categorize the signal.

For the detection and categorization of mental tasks, the majority of the currently used methods employ traditional machine learning techniques and subject-specific handmade characteristics. Due to the high anatomical and functional variability of individuals and the non-stationarity of EEG, these methods might not apply across subjects and datasets. Deep learning methods, where the features are autonomously learned from the raw data during the training process, can be used to solve this problem.

3 PROPOSED WORK

The suggested method for reducing EEG signal artefacts through effective pre-processing and feature extraction is depicted in the methodology flow chart shown in Fig. 1 below.

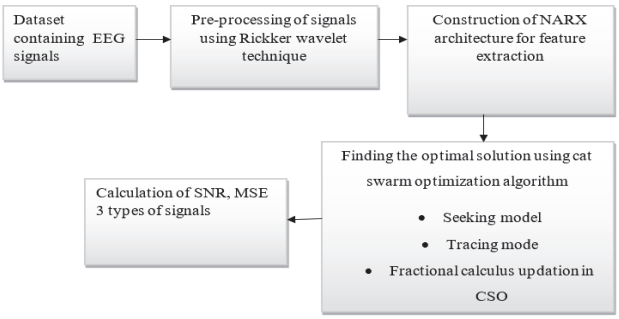

Figure 1 Steps involved in artifacts reduction

In order to preprocess the signals, the input dataset was first learned using the Rickker wavelet technique. The data has first been purified during preparation after being rebuilt and the noise has been eliminated. By building a NARX design, the noise-removed signal is provided to the feature extraction module, where convolution layer and scaling layer are present. Then, to integrate earlier weights in the revised solution of weights, we utilized fractional calculus and cat optimization to achieve an improved optimum solution. In the proposed system Ricker wavelets are employed to generate artificial signals along with white Gaussian noise. This helps to simulate the noise environment even if the actual motion distortion noise

reference is not available. This artificially generated reference signal can be used to train the models and to test the various denoising algorithms such as NARX (Nonlinear Autoregressive Network with Exogenous Inputs) systems.

ECG (Electrocardiogram) signal motion artifacts are undesirable signal changes that may be due to patients' movement or other disturbances that influence the signal. These artefacts are more likely to appear during the change of patient position, muscular contractions or even improper connection of electrodes to the patient and could seriously interfere with ECG recording. Finally, frequency-dependent motion artefacts are in phase with the ECG signal in the time and frequency domains, and hence, the isolation of the true cardiac signal from background noise is difficult. During most forms of patient movement including basic activities such as walking, coughing as well as even speaking, the muscle generated electrical signals (EMG signals) that interfere with the ECG are produced. These signals normally occupy a band of frequencies similar to that of the ECG signal and thus, identified motion artifacts can be seen to be originating from the same frequency range as the heart electrical activity. Another artifact involves the coverage of electrodes on the skin; the movement of electrodes alters the base line of the signal for a while (baseline wander). This baseline drift can undermine the usual appearance of ECG features such as the P-wave, QRS complex and T-wave.

Preprocessing of signal

Even with complicated and time-varying waveforms, the most important non-linear features must be extracted in order to effectively remove errors from the EEG signal. Therefore, these issues must be resolved by implementing the Rickker wavelet technique of EEG distortion elimination. A wavelet transform is used to describe the time-frequency region. Through this method, the entire frequency range of a signal can be split up into a number of small regions, improving time-frequency localization and making it easier to spot low-energy frequency components. Presume that the initial EEG sensor output, $g_j^n(t) \epsilon U_j^n$, can be given by:

$$g_j^n(t) = \sum_k d_k^{j,n} u_n(2^j t - k) \tag{1}$$

As a result, the decomposition method for the following subset U_{j+1}^{2n} and U_{j+1}^{2n+1} can be written as follows.

$$\begin{cases} d_l^{j+1,2n} = \sum_{k-2l} d_k^{j,n} \\ d_l^{j+1,n+1} = \sum_{k-2l} d_k^{j,n} \end{cases} \tag{2}$$

The recovered information can be represented as

$$d_k^{j,n} = \sum_l [h_{k-2l} d_l^{j+1,2n} + g_{k-2l} d_l^{j+1,2n+1}] \tag{3}$$

where $d_l^{j+1,2n}$ and are the wavelet coefficients at level $j+1$ after decomposition, g_{k-2l} and h_{k-2l} are a combination of low and high pass filters, and $d_k^{j,n}$ are the wavelet coefficients at level j. In the time domain, the resulting Ricker wavelet is described as

$$f(t) = \left(1 - 2\pi^2 f_p^2 t^2\right) \exp(-\pi^2 f_p^2 t^2) \tag{4}$$

where f_p is the highest frequency, t is the duration, and $f(t)$ is the amplitude. Here, we decide to create artificial signals along with white Gaussian noise using Ricker wavelets. It is demonstrated that the fundamental Ricker wavelet, with a frequency range of 0-75 Hz, has a uniform trough-crest-trough structure in the time domain. Following that, the artificial signals are constantly analyzed and sampled at a frequency of 1000 Hz for denoising processing.

Construction of NARX architecture for feature extraction

After pre-processing, the filtered signals are fed to NARX system model, which consists of the multiple layers of feed forward networks, time delays, and recurrent loops which is used as a nonlinear system in many applications for the analysis of time series prediction problems. Three vector layers, including input, hidden, and output layers, make up the NARX system architecture as shown in Fig. 2.

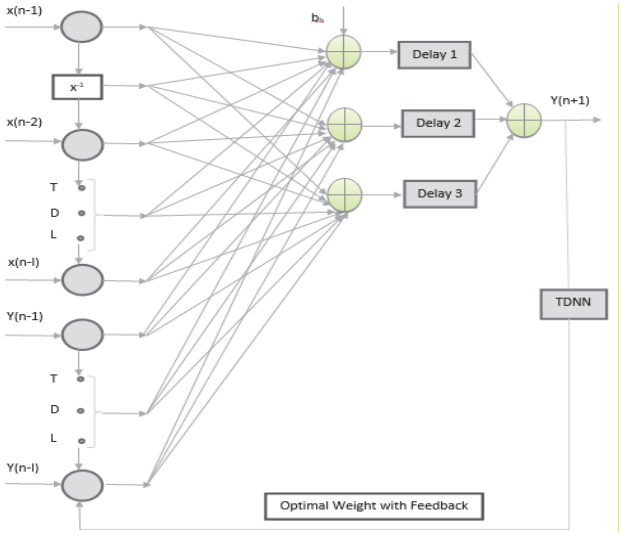

Figure 2 Architecture of NARX neural network

Due to the fact that the real one is made accessible during network training, it can be utilized in the modelling of nonlinear dynamic systems as the output transmitted back to the input of the feedforward neural network. Alternately, a series-parallel design could be developed in which the true output is employed rather than the predicted output being fed back. The NARX model's algebraic version is as follows:

$$y(n+1) = f[x(n-d_x), \ldots x(n-1); x(n); y(n-d_y) \ldots y(n)] \tag{5}$$

where d_{x1} and d_{y1}, $d_x d_y$, are the input and output memory values, accordingly, and $f(.)$ is a nonlinear mapping function. $x(n)$ and $y(n)$ are the system's input and output at discrete time step n. The NARX-NN's data can be completely expressed as:

$$y(n+1) = f_0[b_0 + f_h(b_h + \sum_{i=0}^{d_x} w_{ih} x(n-i) + \sum_{j=0}^{d_y} w_{jh} y(n-j)] \quad (6)$$

where w_{ho}, w_{ih}, w_{jh}; i = 1, 2, … d_x; j = 1, 2, … d_j; $h = 1, 2, ... Nh$ are the network weight vectors, b_h and b_0 are the biases, and $f_h(.)$ and $f_0(.)$ are the activation functions of the hidden and output layers, respectively.

The NARX network is built as a feedforward time delay neural network (TDNN) without delayed feedback loops. In this situation, the NARX neural network's output memory sequence d_y is decreased to zero. A TDNN in mathematics depicts a function with the following formula:

$$y(n+1) = f[x(n-d_{x,} \ldots x(n-1), x(n)] \quad (7)$$

Here, $y(n + 1)$ denotes the output time-series' one-step move ahead anticipated value using the historical values of the original time-series $x(n)$. Consequently, the Eq. (6) simplifies to:

$$y(n+1) = f_0[b_0 + \sum_{h=1}^{Nh} w_{ho} . f_h[b_h + \sum_{i=0}^{d_x} w_{ih} . x[n-i)] \quad (8)$$

It should be mentioned that the TDNN is powered by the historical values of foreign or external input $x(n)$, as opposed to the previously employed RNNs employing the past values of the identical time series that are to be forecasted. Due to the lack of an established standard for depicting TDNNs, in this paper, TDNNs are only referred to as NARX-NNs in order to set them apart from traditional FFNN configurations.

Network design is proven to have an impact on prediction accuracy. The number of neurons in the buried layer and the amount of the integrated memory of inputs I'_f of input O'_f determine the NARX model design. The process of figuring out these factors is not simple. By reducing the network error, it is possible to calculate the number of delay lines for input I'_f of input O'_f. For network training, the Bayesian regularization method is employed.

This method changes the network neurons' weights and shifts in accordance with the Levenberg-Marquardt optimization. With the help of the Bayesian regularization method, the hidden layer's neuron and weight numbers can be combined to give the network the greatest degree of generalization. At the training step of the network, the regularization function reduces the linear mixture of the squared mistakes and weight factors. As a result, it is feasible to maximize the buried layer's cell count and prevent the overtraining impact. The compressed output is fed into the dense/fully linked layer, the following layer, which creates the categorization output with dimension $M \times 1$, where M is the number of groups, as an input. The layer process is typically depicted as:

$$output = \sigma(< input, w_d > + b_d) \quad (9)$$

where, b_d stands for the bias vector for this layer, σ is the activation function, and $\langle input, w_d \rangle$ denotes the dot product between the input and the weight vector w_d used in this layer.

Both sigmoid and softmax activation are used in this study for binary and multi-class categorization, accordingly. The formula for the sigmoid activation function is:

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (10)$$

This method generates a binary result that represents the probability for a binary categorization, based on which the class name is either "0" or "1". The softmax activation function can additionally be expressed as

$$Soft\max(z)_i = p_i = \frac{\exp(z)_i}{\sum_{j=1}^{m} \exp(z)_j} \quad (11)$$

where z_i stands for the i-th member of the preceding layer z's output vector. To place the value of pi between 0 and 1, the numerator is adjusted by the total of all exponential components from 1 to M.

The category class names for multi-class categorization are generated by this layer. The weights are set using the glorot uniform initialized in this work, which does not use a bias vector for this layer.

Finding the optimal solution using cat swarm optimization algorithm

While extracting features, the finding of optimal solution is a must for updating weights for feedback. Here, this process is done by using cat optimization algorithm. Chu and Tsai [22] introduced the cat swarm optimization method in 2007. They were inspired by cat behaviour. The seeking mode and the tracking mode are two search logics that are combined in this method. Before the rounds, all of the cats will be split into two categories. The designed cat symbolizes an answer for the problem to be solved, and each designed cat has its own position and velocity. The Particle Swarm Optimization (PSO) algorithm is very similar to its update techniques. The globally best will be selected and will lead all of the cats to pursue its next position and velocity. The answers are revised by the cat's position and velocity, and approximated based on its degree of flexibility for the project. It is clear that this differs from the particle swarm optimization method from the perspective of biological group behavior. The following part will provide an introduction to the CSO algorithm's specifications.

Seeking mode - The GR (group rate) setting, which is typically set to 0.98, determines the proportion of cat groups that are in the seeking mode. The GR is used to

slightly adjust the cats' location while they are in seeking mode and does affect their velocity. The actions listed below will be conducted. First, each cat will make numerous copies of its own location, based on the amount of the SMP parameter, and place each duplicate in the associated seeking mode pool (SMP) unit. Then, a mutation operator will re-evaluate each point in the SMP, selecting a dimension of the anticipated variable x_i to modify, and determining the range of variance, up to 20% of x_i . The mutation operator is denoted by Eq. (12):

$$x_i = x_i + \Delta x_i \tag{12}$$

It will be decided to update x_i from the SMP location with the highest level of fitness.

Tracing mode - The particle in the PSO methods and the cats' evolving process in tracking mode are comparable. They are still somewhat unique, though. When in tracing mode, each cat will only change its own motion and location by tracking the cat with the greatest fitness globally. The finest persons on a worldwide and local scales are both tracked by the PSO particles. The Eqs. (13) and (14) can be used to describe the changing algorithm for the tracking method:

$$v_k(t+1) = \omega.v_k(t) + const.random.[x_{gb}(t) - x_k(t)] \tag{13}$$

$$x_k(t+1) = x_k(t) + v_k(t+1) \tag{14}$$

where x_k is the location of "cat_k" the generation for iterations, and x_{gb} is the position of the cat with the greatest fitness. *random* is a random integer in the range [0, 1], and *const* is a constant. The winning cat would be selected to change the variable after comparisons between the pursuing cat and the cat with the highest fitness in tracking mode. The necessary response is the ultimate x_{gb} .

Triangle Walk Strategy- The sand cats use the triangular walk technique to circle as they get close to their target. The first thing to do is to measure the distance $L1$ between the cat and its target. Then get the cat's step size range L2. Determine the way of the cat's strolling (β) using Eq. (17). Eqs. (15) and (16) indicate $L1$ and $L2$ respectively.

$$L1 = pos_b(t) - pos_c(t) \tag{15}$$

$$L2' = rand(h) \times L1' \tag{16}$$

$$\beta = 1 \times \pi \times rand() \tag{17}$$

$$pos_{new} = pos_b(t) + r \times p \tag{18}$$

Among them, "pos_{new}" refers to the location reached using the strolling strategy.

Levy Flight Walk Strategy- The sand cat stalks its target very closely before walking. Levy flight is a powerful statistical technique for generating random variables. Levy distribution-conforming walking can be accomplished using Levy flying. However, Levy's flight can occasionally have too lengthy of a stride. The constant C = 0.35 is increased in Levy flight to better match the behavior of sand cats approaching food. This enables the sand cat to approach its target as closely as feasible.

$$pos_{new} = pos_b(t) + (pos_b(t) - pos_c(t)) \times C \times levy \tag{19}$$

Updating fractional calculus in CSO

An innovative approach to managing the CSO algorithm is presented in this section. To change the sequence of the velocity derivative, the initial velocity Eq. (13) is first changed, as follows:

$$v(t+1) = v_t + \varphi 1(b-x) + \varphi 2(g-x) \tag{20}$$

This equation can also be expressed as:

$$v(t+1) - v_t = \varphi 1(b-x) + \varphi 2(g-x) \tag{21}$$

Given that $\propto = 1$ (assuming T = 1), the left side $v(t+1) - v_t$ is the discontinuous derivative with order = 1. This results in the expression:

$$D^{\propto}\left[v(t+1)\right] = \varphi 1(b-x) + \varphi 2(g-x) \tag{22}$$

If the FC perspective is used, the rank of the velocity derivative can be extended to a real number $0 \le \propto \le 1$, resulting in better fluctuation and a prolonged memory impact. An array of models are performed on evaluating values of $\propto$ varying from $\propto = 0$ up to $\propto = 1$, with steps of $\propto = 0.1$, in order to analyze the behaviors of this novel CSO approach.

As a result, Eq. (22) can be reformulated as Eq. (23) when r = 4 is considered in terms of differential derivative, generating

$$\begin{aligned} &v(t+1) - \propto v_t - \frac{1}{2} \propto v(t-1) - \frac{1}{6}(1-\propto)v(t-2) - \\ &-\frac{1}{24}(1-\propto)(2-\propto)v(t-3) = \varphi 1(b-x) + \varphi 2(g-x) \end{aligned} \tag{23}$$

The same findings were obtained when testing larger numbers. The program finally comes to a close at the conclusion of the iteration, and the best outcome offers the optimal weight.

In the study, CSO with fractional calculus is used to inquire optimal parameters of a NARX model. NARX is employed to solve problems of complicated nonlinear dynamics, for instance the reduction of a motion artifact with biomedical signal processing. Nevertheless, classical NARX models require accurate settings of the number of neurons, learning rate, and time delays to exhibit high performance. The hyperparameter tuning was done in the traditional approach or manually, which is time-consuming when it is being applied in larger model. This is where the CSO algorithm comes into the play.

CSO is used in this study to perform the automatic search for the best parameters for the NARX model in the search space. The algorithm mimics the natural behaviors of cats through two modes: exploratory and expolitive, which are represented by the two subprocesses of seeking and tracking. Mode of search enables rough tuning of all the parameters of the algorithm broadly, so that the algorithm does not converge to local optimum, while tracking mode narrows down the search to best candidates only. This two-step approach makes the search for the best hyperparameters complete as well as sensible and hence the best.

4 RESULTS AND DISCUSSION

In the proposed research, the time complexity of the overall model is calculated. When combining the NARX network, CSO optimization, and fractional calculus, the overall computational complexity can be represented as: $O(P{\cdot}I{\cdot}((d_x + d_y)N + n{\cdot}r))$ where P is the population size in CSO, I is the number of iterations, d_x and d_y are the input and output memory delays, N is the number of neurons in the hidden layer of the NARX model, n is the number of data points and r is the order of the fractional derivative.

Here, the modeling findings of the suggested artefact elimination method using the NARX_CSOFC network are quickly examined. The simulation is performed by comparing the outcomes with a variety of known methods, including the Competitive Swarm Dragonfly Algorithm (CSDA) + Deep Residual Network (DRN) [14], the Oppositional Whale Algorithm (OWOA) [18], and K-means Nearest Neighbour + Long Short Term Memory (KNN+LSTM) [13]. These are done for reducing EEG from ECG, EEG from EMG, EEG from random noise.

Dataset description-EEG denoise Net is a standard EEG dataset that can be used to train and evaluate DL-based denoising models and to compare model results. Users can combine contaminated EEG segments with the ground-truth clean using EEG denoise Net's 4514 clean EEG segments, 3400 eye artefact segments, and 5598 muscle artefact segments.

The goal of the suggested NARX_CSOFC technique is to eliminate the artefacts from the EEG data. Metrics used in this study for assessment include Mean Squared Error (*MSE*) and Signal to Noise Ratio (*SNR*).

Mean Square Error (*MSE*): The *MSE* measure describes the difference between the real reaction and the intended response:

$$MSE = \frac{1}{N}(O^n - D^n)^2 \tag{24}$$

Signal to Noise Ratio (*SNR*): As shown below, the *SNR* measurement is computed,

$$SNR = 10\log_{10}\frac{S(t)}{l(t)} \tag{25}$$

Fig. 3 shows the performance of EEG signal with ECG artefacts based on *SNR* and *MSE* of the suggested method with existing models. When analyzing, the existing KNN+LSTM, CSDA+DRN, OWOA achieve 45 dB, 57 dB and 32 dB of *SNR*, whereas the proposed NARX_CSOFC achieves 12 dB of *SNR* which is 33 dB,45 dB,20 dB better than aforementioned existing methods. When analyzing *MSE*, the existing methods achieve 9000, 7000, 10000 of *MSE* whereas the proposed NARX_CSOFC achieves 13500 of *MSE*, which is 4500, 6500, 3500 better than aforementioned existing methods.

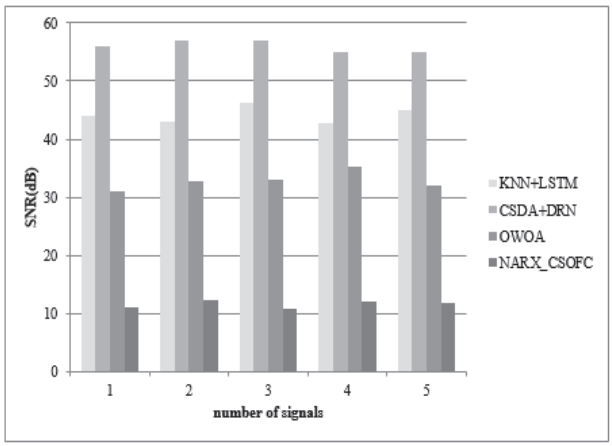

Figure 3a EEG signals with ECG artefact (*SNR*)

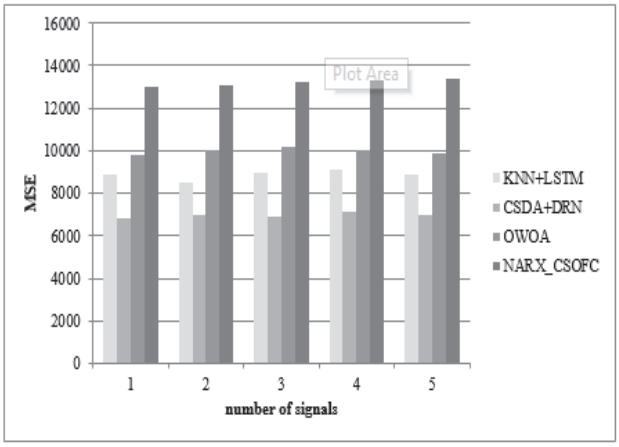

Figure 3b EEG signals with ECG artefact (*MSE*)

Fig. 4 shows the performance of EEG signal with EMG artefacts based on *SNR* and *MSE* of the suggested method with existing models. When analyzing, the existing KNN+LSTM, CSDA+DRN, OWOA achieve 54 dB, 59 dB and 38 dB of SNR, whereas the proposed NARX_CSOFC achieves 16 dB of *SNR* which is 38 dB, 43 dB, 22 dB better than aforementioned existing methods. When analyzing *MSE*, the existing methods achieve 8000, 6000, 11000 of *MSE* whereas the proposed NARX_CSOFC achieves 15000 of *MSE*, which is 7000, 9000, 4000 better than aforementioned existing methods.

Fig. 5 shows the performance of EEG signal with random noise based on *SNR* and *MSE* of the suggested method with existing models. When analyzing, the existing KNN+LSTM, CSDA+DRN, OWOA achieve 43 dB, 54 dB and 35 dB of *SNR*, whereas the proposed NARX_CSOFC achieves 15 dB of *SNR* which is 38 dB, 43 dB, 22 dB better than aforementioned existing methods. When analyzing *MSE*, the existing methods achieve 8500, 7800, 12000 of *MSE* whereas the proposed NARX_CSOFC achieves 14000 of *MSE*, which is 7000, 9000, 4000 better than aforementioned existing methods.

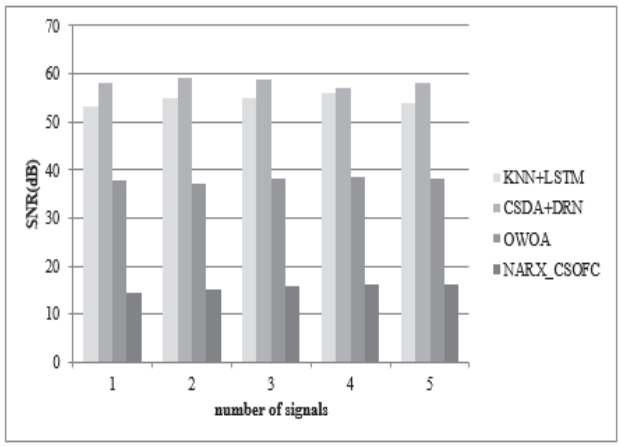

Figure 4a EEG signals with EMG artefact (*SNR*)

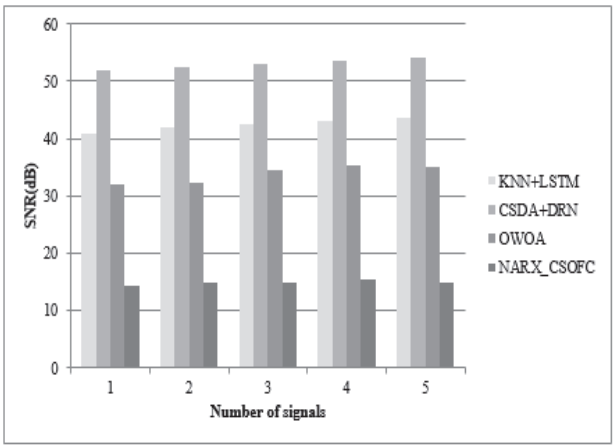

Figure 5a EEG signals with EMG artefact (*SNR*)

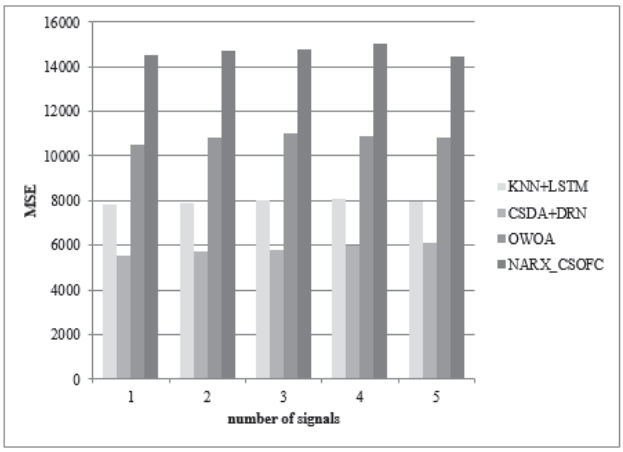

Figure 4b EEG signals with EMG artefact (*MSE*)

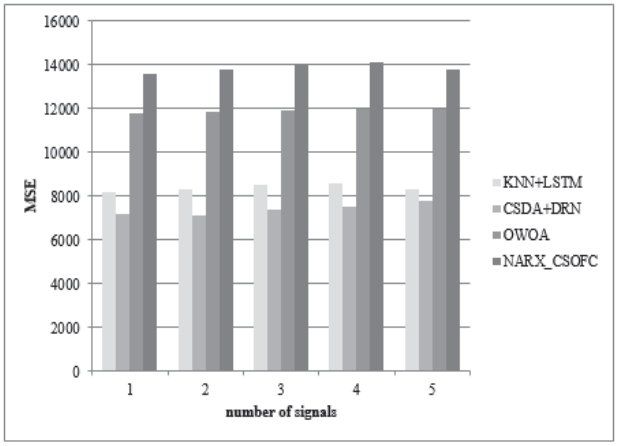

Figure 5b EEG signals with EMG artefact (*MSE*)

Table 1 Comparison of the EEG signals with Different methods

Signal with artefacts	KNN+LSTM		CSDA+DRN		OWOA		NARX_CSOFC	
	SNR / dB	*MSE*	*SNR* / dB	*MSE*	*SNR* / dB	*MSE*	*SNR* / dB	*MSE*
EEG from ECG	45	9000	57	7000	32	10000	12	13500
EEG from EMG	54	8000	59	6000	38	11000	16	15000
EEG from random noise	43	8500	54	7800	35	12000	15	14000

5 CONCLUSION

In this study, an efficient EEG motion distortion identification and elimination method is proposed for developing an accurate analysis and diagnosis of neurological diseases. The suggested design consists of a few-layer network assisted by a phase for contaminant suppression or feature extraction. As shown by different activation weights that produce higher SNR, our proposed NARX_CSOFC design has successfully learned discriminative feature maps for baseline and various classifications of mental tasks. Future work will focus on enhancing the adaptive artifact elimination method's ability to identify different types of neural abnormalities, and a better feature selection technique will be used to maximize the outcomes.

Acknowledgement

The author extends the appreciation to the Deanship of Postgraduate Studies and Scientific Research at Majmaah University for funding this research work through the project number R-2024-1369.

6 REFERENCES

[1] Manali, S. & Udit, S. (2020). Wavelet based waveform distortion measures for assessment of denoised EEG quality with reference to noise-free EEG signal. *IEEE Signal Processing Letters*, *99*, 1-1. https://doi.org/10.1109/LSP.2020.3006417

[2] Mahajan, R. & Morshed, B. I. (2015). Unsupervised eye blink artifact denoising of EEG data with modified multiscale sample entropy, kurtosis, and wavelet-ICA. *IEEE Journal of Biomedical and Health Informatics*, *19*(1), 158-165. https://doi.org/10.1109/JBHI.2014.2333010

[3] Ding, M. (2024). Seismic signal denoising based on Surelet transform for energy exploration. *Technical Gazette*, *31*(4), 1130-1142. https://doi.org/10.17559/TV-20230926000964

[4] Stephe, S., Jayasankar, T., & Kumar, K. V. (2022). Motor Imagery EEG Recognition using Deep Generative Adversarial Network with EMD for BCI Applications. *Tehnički vjesnik*, *29*(1), 92-100. https://doi.org/10.17559/TV-20210121112228

[5] Zhao, X., Liu, D., Ma, L. et al. (2020). EEG signals de-noising with wavelet by optimizing threshold based on fruit fly optimization. *ICNCC*. https://doi.org/10.1145/3447654.3447665

[6] Xun, C., Xueyuan, X., Aiping, L., McKeown, M. J., & Wang, Z. J. (2018). The use of multivariate EMD and CCA

for denoising muscle artefacts from few-channel EEG recordings. *IEEE Transactions on Instrumentation and Measurement*, *67(*2), 359-370. https://doi.org/10.1109/TIM.2017.2759398

[7] Gorjan, D., Gramann, K., Pauw, K. D., & Marusic, U. (2022). Removal of movement-induced EEG artefacts: current state of the art and guidelines. *Journal of Neural Engineering*, *19*(1), 011004. https://doi.org/10.1088/1741-2552/ac3e6b

[8] Kaur, C., Singh, P., & Sahni, S. (2021). EEG artifact removal system for depression using a hybrid denoising approach. *Basic and Clinical Neuroscience*, *12*(4), 465-476. https://doi.org/10.32598/bcn.2021.1388.2

[9] Lawhern, V., Hairston, W. D., & Robbins, K. (2013). DETECT: A MATLAB toolbox for event detection and identification in time series, with applications to artefact detection in EEG signals. *PLOS ONE*, *8*(4). https://doi.org/10.1371/journal.pone.0062944

[10] Vijila, C. K. S., Kanagasabapathy, P., Johnson, S., & Ewards, V. (2007). Artefacts removal in EEG signal using adaptive neuro fuzzy inference system. 2007 International Conference on Signal Processing, Communications and Networking. *IEEE*, 589-591. https://doi.org/10.1109/ICSPCN.2007.353146

[11] Schetinin, V. & Maple, C. (2007). A Bayesian model averaging methodology for detecting EEG artefacts. *2007 15th International Conference on Digital Signal Processing. IEEE*, 499-502. https://doi.org/10.1109/ICDSP.2007.4288600

[12] Fairley, J., Georgoulas, G., Stylios, C., & Rye, D. (2010). A hybrid approach for artefact detection in EEG data. International Conference on Artificial Neural Networks. Springer, 436-441. https://doi.org/10.1007/978-3-642-15825-4_55

[13] Nguyen, H. A. T., Musson, J., Li, F. et al. (2012). EOG artefact removal using a wavelet neural network. *Neurocomputing*, *97*, 374-389. https://doi.org/10.1016/j.neucom.2012.05.030

[14] Ghosh, R., Phadikar, S., Deb, N. et al. (2023). Automatic eye-blink and muscular artefact detection and removal from EEG signals using k-nearest neighbour classifier and long short-term memory networks. *IEEE Sensors Journal*. https://doi.org/10.1109/JSEN.2023.3060213

[15] Kumar, T. R., Mahalaxmi, U. S. B. K., Ramakrishna, M. M., & Bhatt, D. (2023). Optimization enabled deep residual neural network for motor imagery EEG signal classification. *Biomedical Signal Processing and Control*, *80*, 104317. https://doi.org/10.1016/j.bspc.2022.104317

[16] Aquino-Brítez, D., Ortiz, A., Ortega, J. et al. (2021). Optimization of deep architectures for EEG signal classification: an AutoML approach using evolutionary algorithms. *Sensors (Basel)*, *21*, 2096. https://doi.org/10.3390/s21062096

[17] Chen, X., Liu, Q., Tao, W. et al. (2019). ReMAE: A user-friendly toolbox for removing muscle artefacts from EEG. *IEEE Transactions on Instrumentation and Measurement*. https://doi.org/10.1109/TIM.2019.2934482

[18] Nagal, R., Kumar, P., & Bansal, P. (2019). An optimal approach for EEG/ERP noise cancellation using adaptive filter with oppositional whale optimization algorithm. *Biomedical Engineering: Applications, Basis and Communications*, *31*(5), 1950035. https://doi.org/10.4015/S1016237219500352

[19] Liu, L. (2019). Recognition and analysis of motor imagery EEG signal based on improved BP neural network. *IEEE Access*, *7*, 47794-47803. https://doi.org/10.1109/ACCESS.2019.2909177

[20] Sun, W., Su, Y., Wu, X. et al. (2021). EEG denoising through a wide and deep echo state network optimized by UPSO algorithm. *Applied Soft Computing*, *105*, 107149. https://doi.org/10.1016/j.asoc.2021.107149

[21] Shahbakhti, M., Beiramvand, M., Nazari, M. et al. (2021). A review on machine learning algorithms in handling EEG artefacts. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *29*, 408-417. https://doi.org/10.1109/TNSRE.2021.3056871

[22] Chu, S. C., Tsai, P., W. & Pan, J. S. (2006). Cat swarm optimization. *Proceedings of the 9th Pacific Rim International Conference on Artificial Intelligence*, 854-858. https://doi.org/10.1007/978-3-540-36668-3_87

[23] Selvam, R. P., Oliver, A. S., Mohan, V., Prakash, N. B., & Jayasankar, T. (2022). Explainable Artificial Intelligence with Metaheuristic Feature Selection Technique for Biomedical Data Classification. *Biomedical Data Analysis and Processing Using Explainable (XAI) and Responsive Artificial Intelligence (RAI)*, 43-57. https://doi.org/10.1007/978-981-19-1476-8_4

[24] Sahoo, S. K. & Mohapatra, S. K. (2022). Recognition of ocular artifacts in EEG signal through a hybrid optimized scheme. *Biomedical Research International*, *2022*, 4187167. https://doi.org/10.1155/2022/4187167

[25] Reddy, V. K. & Av, R. K. (2022). Multi-channel neuro signal classification using Adam-based coyote optimization enabled deep belief network. *Biomedical Signal Processing and Control*, *77*, 103774. https://doi.org/10.1016/j.bspc.2022.103774

Contact information:

Dr. Ibrahim ALMOHIMEED
College of Applied Medical Sciences,
Majmaah University, Department of Medical Equipment Technology,
Al Majmaah, Saudi Arabia

Ahmed Farag Salem BABETAT
Dr. Mohammad Alfagih Hospital, Riyadh

Mohammed Nasser ALMRAIKHI
Internship Unit, College of Medicine,
Majmaah University,
Al Majmaah, Saudi Arabia

Dr. Mohamed Yacin SIKKANDAR
(Corresponding author)
College of Applied Medical Sciences,
Majmaah University, Department of Medical Equipment Technology,
Al Majmaah, Saudi Arabia
E-mail: m.sikkandar@mu.edu.sa

Bader Dhaidan OWAYYID OWAYJAH
Internship Unit, College of Medicine,
Majmaah University,
Al Majmaah, Saudi Arabia

Ali Abdullah ALMUKIL
College of Applied Medical Sciences,
Majmaah University, Department of Medical Equipment Technology,
Al Majmaah, Saudi Arabia

Abdulrhman Abdulaziz ALMAJHAD
Clinical Skills Center Unit, College of Medicine,
Majmaah University,
Al Majmaah, Saudi Arabia