Received November 24, 2019, accepted December 23, 2019, date of publication December 27, 2019,
date of current version January 17, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2962658

EEG Signals Denoising Using Optimal Wavelet Transform Hybridized With Efficient Metaheuristic Methods

ZAID ABDI ALKAREEM ALYASSERI[1,2], (Senior Member, IEEE), AHAMAD TAJUDIN KHADER[1], MOHAMMED AZMI AL-BETAR[3], AMMAR KAMAL ABASI[1], AND SHARIF NASER MAKHADMEH[1]

[1]School of Computer Sciences, Universiti Sains Malaysia, Pulau Pinang 11800, Malaysia
[2]ECE Department, Faculty of Engineering, University of Kufa, Najaf 54001, Iraq
[3]Department of Information Technology, Al-Huson University College, Al-Balqa Applied University, Irbid 19117, Jordan

Corresponding author: Ahamad Tajudin Khader (tajudin@usm.my)

This work was supported by the University Science Malaysia (USM) under Grant 1001/PKOMP/8014016. The work of Zaid Abdi Alkareem Alyasseri was supported by The World Academic Science (TWAS) and USM through the TWAS-USM Postgraduate Fellowship 2015 under Grant 3240287134.

ABSTRACT **Background**. The most common and successful technique for signal denoising with non-stationary signals, such as electroencephalogram (EEG) and electrocardiogram (ECG) is the wavelet transform (WT). The success of WT depends on the optimal configuration of its control parameters which are often experimentally set. Fortunately, the optimality of the combination of these parameters can be measured in advance by using the mean squared error (MSE) function. **Method**. In this paper, five powerful metaheuristic algorithms are proposed to find the optimal WT parameters for EEG signal denoising which are harmony search (HS), β-hill climbing (β-hc), particle swarm optimization (PSO), genetic algorithm (GA), and flower pollination algorithm (FPA). It is worth mentioning that this is the initial investigation of using optimization methods for WT parameter configuration. This paper then examines which efficient algorithm has obtained the minimum MSE and the best WT parameter configurations. **Result**. The performance of the proposed algorithms is tested using two standard EEG datasets, namely, Kiern's EEG dataset and EEG Motor Movement/Imagery dataset. The results of the proposed algorithms are evaluated using five common criteria: signal-to-noise-ratio (SNR), SNR improvement, mean square error (MSE), root mean square error (RMSE), and percentage root mean square difference (PRD). Interestingly, for almost all evaluating criteria, FPA achieves the best parameters configuration for WT and empowers this technique to efficiently denoise the EEG signals for almost all used datasets. To further validate the FPA results, a comparative study between the FPA results and the results of two previous studies is conducted, and the findings favor to FPA. **Conclusion**. In conclusion, the results show that the proposed methods for EEG signal denoising can produce better results than manual configurations based on ad hoc strategy. Therefore, using metaheuristic approaches to optimize the parameters for EEG signals positively affects the denoising process performance of the WT method.

INDEX TERMS EEG, signal denoising, wavelet transform, metaheuristic algorithms, optimization, flower pollination algorithm.

I. INTRODUCTION

Electroencephalogram (EEG) is a graphical recording of brain electrical activity that is recorded from the scalp. This recording represents the voltage fluctuations resulting from ionic current flows within the neurons of the brain [1], [2]. Therefore, EEG signals can provide most of the required information about brain activity. EEG signals from the brain are captured using *invasive* or *non-invasive* techniques [3]. The main difference between these techniques is that the invasive approach involves the use of electrode arrays implanted inside the brain, such as the eastern cooperative oncology

The associate editor coordinating the review of this manuscript and approving it for publication was Peng Xu.

group-Brain Computer Interface *(ECOG-BCI)* for arm movement control [4]. Meanwhile, there are several techniques to record the brain activity can also be captured using different types of signal capturing devices, including EEG for electrical activity from the scalp, MEG for magnetic field fluctuations caused by electrical activity in the brain, and functional magnetic resonance imaging (FMRI) and functional near-infrared spectroscopy (fNIRS) for changes in blood oxygenation level resulting from neural activity [4]. In [5], Berger proposed for the first time the use of EEG signals as a non-invasive technique for capturing brain activities. Over the past several decades, researchers have developed Hans's technique to suit multiple applications. For instance, EEG signals have been used in medical applications for prevention, detection diagnosis, rehabilitation and restoration. This technique has also been used for non-medical applications, such as education and self-regulation, neuromarketing and advertisement, neuroergonomics and smart environment, games and entertainment, and learning and education [6], [7]. Recently, EEG signals have been used as a new biometric technique in security and authentication applications [1], [6].

Several artifact noises can corrupt the original EEG signal during its recording time, such as eye blink, eye movements, muscle activity, and interference of electronic device signals [8]. Therefore, the EEG signal must be processed to reduce such noise. Several EEG noise removal techniques have been proposed in the literature, such as filtering and adaptive thresholding. Recently, wavelet transform (WT) has been successfully applied for denoising non-stationary signals, including ECG and EEG [9]–[12].

In general, WT has five parameters with each parameter having different types (Table 1). The efficiency of EEG signal denoising depends on the selection of the best combination of WT parameters. The selection is usually performed based on experience or empirical evidence. In previous research, the WT parameter configuration is formulated as an optimization problem with MSE as its objective function [11]. As aforementioned, WT has five parameters, namely, (i) mother wavelet function (MWF) Φ, (ii) decomposition level $\boldsymbol{L}$, (iii) thresholding function β, (iv) threshold selection rules λ, and (v) threshold re-scaling methods ρ. Each of these parameters has several values and is used for a specific denoising level. The optimal values of these parameters are required to empower WT in the denoising process. For ECG signals, El-Dahshan in [9] attempted to obtain the optimal configuration using GA, the results were better than those that were produced experimentally. Alyasseri *et al.* [10], [13] proposed a hybrid scheme for non-stationary signals denoising, such as ECG and EEG that is based on β-hill climbing (βhc) optimization algorithm [14] with WT to obtain the optimal wavelet parameters. The proposed method (βhc-WT) was tested using an MIT-BIH dataset [15], where the original ECG signal was corrupted with white Gaussian noise (WGN) using different input SNR noises that corrupted the ECG from 0 dB to 40 dB. The performance of the βhc-WT method was evaluated using minimum squared error (MSE) and SNR.

TABLE 1. The ranges of the wavelet denoising parameters.

WT denoising parameters	Method (range)
Mother Wavelet function Φ	**Symlet** (sym1..sym45), **Coiflet** (coif1..coif5), **Daubechies** (db1..db45), and **Biorthogonal** (bior1.1.. bior1.5&bior2.2 .. bior2.8& bior3.1..bior3.9).
Thresholding function β	soft or hard threshold
Decomposition level $\boldsymbol{L}$	5
Thresholding selection rule λ	Heursure, Rigsure, Sqtwolog, and Minimax
Re-scaling approach ρ	one, sln, mln

The proposed method successfully removed WGN from the ECG and EEG signals [10]–[13]. Several metaheuristic optimization algorithms have also been employed. Metaheuristic is a general optimization framework for several kinds of optimization problems that uses efficient learning operators to explore the search space regions and to exploit the accumulative search controlled by certain parameters. Metaheuristic algorithms are conventionally categorized into: i) evolutionary algorithms (EAs), including GA [16], harmony search (HS) [17], and genetic programming (GP) [18]; ii) swarm-based intelligence algorithms (SI), including particle swarm optimization (PSO) [19], artificial bee colony (ABC) [20], flower pollination algorithm (FPA) [21], and iii) trajectory-based algorithm (TAs), including β-hill climbing (βHC) [14], simulating annealing (SA) [22], tabu search (TS) [23], greedy randomized adaptive search procedure (GRASP) [24], variable neighborhood search (VNS) [25], iterated local search (ILS) [26] meta-heuristic. The main research question of this paper is what is the best choice optimization algorithm that can be find the optimal parameters values for WT to empower its denoising process for EEG signals? The main objective of this paper is to propose five metaheuristic algorithms from different metaheuristic categories for optimal settings of WT parameters. This paper also attempts to determine the most efficient metaheuristic algorithm for finding those parameters and that can help WT denoise EEG signals efficiently. These metaheuristic algorithms include some SI algorithms (i.e., FPA and PSO), some EA algorithms (i.e., GA and HS), and one TA algorithm (i.e., β-HC). These algorithms are carefully chosen based on their performance in successfully solving a wide variety of signal and image processing problems [10]–[12], [27]–[29]. Therefore, five versions of WT, namely, FPA-WT, GA-WT, HSA-WT, PSO-WT, and βHC-WT are tested in an experiment. The original EEG signal benchmark taken from two *de facto* EEG datasets, namely, Kierns[1] and *Motor Movement/Imagery dataset*[2] are used for the evaluation process [15], [30]. To evaluate the performance of the meta-heuristic algorithms, EEG signals are corrupted using three different noise mechanisms, including power line noise (PLN), electromyogram (EMG), and white Gaussian noise (WGN) [9], [31], [32]. Initially, each proposed meta-heuristic algorithm generates optimal parameter settings for WT to denoise the EEG signal of each dataset. Afterward, the denoising results are evaluated using five measurement

[1]http://www.cs.colostate.edu/eeg/main/data/1989_Keirn_and_Aunon
[2]https://www.physionet.org/physiobank/database/eegmmidb/

factors, namely, SNR, SNR improvement, MSE, RMSE, and PRD. For comparative evaluation, the denoising results of the five proposed methods are compared with one another. Interestingly, FPA-WT achieves efficient EEG signal denoising for EMG and WGN datasets. In addition, FPA-WT and GA-WT obtain the best denoising levels for PLN dataset. In conclusion, FPA is the best algorithm that can be incorporated with WT to achieve an efficient EEG signal denoising.

This paper is organized as follows. Section III provide a background to Wavelet Transform (WT). Section III-A presents a Wavelet denoising principle for EEG signal denoising. The selected meta-heuristic algorithms presents in Section IV. The hybrid scheme between meta-heuristic algorithms and WT explains in Section V. The results and discussion presents in section VI. Finally, the conclusions and future works describes in Section VII.

II. RELATED WORKS

Kumar and Vaish in [1] proposed a user identification system on the basis of EEG signal collected from six users using EMOTIVE EPOC headset with 14 channels. These researchers used wavelet transform (WT) for EEG signal denoising where a *db4* mother wavelet function (MWF) is used with five levels of signal decomposition. They tested their method using the EEG dataset established in [30]. Afterwards, the same authors investigated several cognitive tasks to design an individual identification system [2]. These researchers used standard EEG datasets related to motor/movement and imaginary tasks [15] with only one channel (i.e. Cz) to obtain an input signal. In addition, the authors used WT to decompose the EEG signal into five levels and then extract four features from each EEG sub-band.

Al-Qazzaz *et al.* [33], [34] conducted a comparative study to determine the efficient MWFs that can provide high signal characteristics for an EEG channel. These authors tested 45 MWFs that are categorized into Daubechies, Symlets and Coiflets families. An MWF called 'sym9' showed efficient results in nearly all brain regions. The same team of researchers applied WT with independent component analysis to decompose the EEG signals for obtaining an efficient feature for discriminating stroke-related mild cognitive impairment and vascular dementia [35].

Reddy [36] proposed WT for processing the EEG signal. These authors applied WT to EEG signal denoising and used *db8* as an MWF with eight EEG signal decomposition levels. Furthermore, these authors classified the EEG signal on the basis of the features that are extracted from the WT signal denoising process [37].

Padmaja *et al.* [38] proposed a method for removing ocular artifacts from EEG signal by using Hilbert-Huang transform (HHT). The proposed method has two steps. Firstly, input EEG signals are decomposed using empirical mode decomposition; secondly, HHT is applied to obtain the frequency of each intrinsic mode functions. The proposed method was tested using an EEG signal recorded from six patients. The results of HHT showed better performance than fast Fourier transform in terms of signal-to-noise ratio (SNR) criteria.

Yang *et al.* in [7] proposed an artificial method for removing the EOG artifacts from the EEG raw. The proposed method (CCA-EEMD) involves three steps. In the first step, the input EEG signal proposed using CCA to spread the EOG. In the second step, the EOG will be decomposed into multi-level and apply intrinsic mode functions (IMFs) using EEMD approach. Finally, the clear EEG data are ready to use and extract more features. The (CCA-EEMD) tested using seven subjects. The results show that the (CCA-EEMD) method it is not only EOG removal method but also it can keep the EEG features to the maximum extent. Torabi *et al.* in [39] introduced a combining method between nonlinearity EEG features and wavelet coefficients for improving the performance of the recognition rate classification. The proposed method applied a linear SVM classifier and the effect of the combining technique shown significant improvement in the classification results from (54%) to (73%). Furthermore, the proposed method has been also applied for feature selection for the same problem, while it is selected up (44%) for nonlinear features.

III. WAVELET TRANSFORM

Wavelet Transform (WT) is a common and powerful tool for representing signals in the time-frequency domain. WT has been successfully used for non-stationary signals, such as ECG and EEG, to address several problems, such as those related to signal compression, feature selection, and signal denoising [10], [40], [41]. Recently, WT has been extensively tailored for non-stationary signals because of its powerful performance in removing several EEG artifact noises that can corrupt the original EEG signal during its recording time. These noises include eye blinking noise, eye movement noise, muscle activity noise, electromyogram (EMG) noise, and interference of electronic device signals [42], [43].

A. WAVELET DENOISING PRINCIPLE FOR NON-STATIONARY SIGNALS

As aforementioned in Section III, WT is a powerful tool for time-frequency domain representation. This technique represents the signal on the basis of the correlation between the translation and the dilation of mother wavelet function (MWF) [9], [44], [45]. In general, the problems solved by WT can be categorized into two WT versions, namely, continuous wavelet transform (CWT) and discrete wavelet transform (DWT) [46]. In this paper, DWT has been proposed for EEG signal decomposition whereby inverse DWT (iDWT) is used for EEG signal reconstruction. DWT was originally established in [47] as the so-called Donoho's approach. In general, DWT decomposes a signal by using set of filtering (i.e., low pass and high pass filters) to product the approximation and details coefficients, respectively. The main objective of using DWT is to decompose the input signal via different coefficient levels to correct the high frequency of the input signals [48]. In other word, DWT decomposes the EEG signal

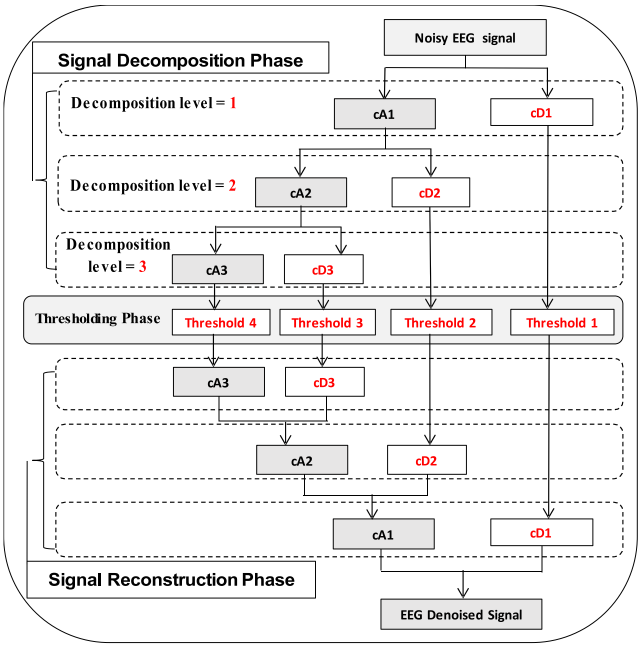


FIGURE 1. EEG denoising process taken from [50].

into several frequency bands because it assumed that the artifacts will have large amplitudes in the respective frequency bands. Fig. 1 shows the wavelet denoising procedure with decomposition level $L = 3$. Normally, the denoising process involves three phases:

- EEG signal ***decomposition*** phase: Assuming the original EEG signals with n samples $x(t) = [x(1), x(2), \ldots, x(n)]$ will be divided into three levels, and each level will be decomposed into two parts, namely, approximation coefficients (cA) and detail coefficients (cD). cD will be processed using a high-pass filter, while cA will continue to be decomposed for the next level.

$$cA_i(t) = \sum_{k=-\infty}^{\infty} cA_{i-1}(k)\,\phi_i(t-k) \quad (1)$$

$$cD_i(t) = \sum_{k=-\infty}^{\infty} cD_{i-1}(k)\,\Psi_i(t-k) \quad (2)$$

where $cA_i(t)$, $cD_i(t)$ denotes the approximation and detail coefficients of level i, Ψ, ϕ refers to scaling and shifting, respectively.
- ***Applying thresholding*** phase: A threshold value is defined for each level according to the noise level of the coefficient.
- ***Reconstruction*** phase: The EEG denoised signal is reconstructed using ***iDWT***. The formula of ***iDWT*** as follows [49]:

$$EEG_{clean} = \sum\nolimits_{k=-\infty}^{\infty} cA_L(k)\,\phi_i'(t-k) + \sum\nolimits_{i=1}^{L} \sum\nolimits_{k=-\infty}^{\infty} cD_{i+1}(k)\,\Psi_i'(t-k) \quad (3)$$

where $EEG_{clean}(t)$ denotes the reconstructed EEG signal, i refers to decomposition level,

Signal noise removal is considered a challenging task in signal processing [51]. Therefore, researchers have developed several approaches to solve this problem, such as using the filtering technique [52]–[54], thresholding technique [55], [56], and other techniques [57], [58]. WT is one of the powerful techniques for non-stationary signal denoising [44], [59]–[61]. WT has five parameters, with each parameter having different types (Table 1) the success of EEG signal denoising relies on the selection of WT parameters. As shown in Fig. 1, the wavelet denoising parameters are defined in three phases. In the decomposition phase, the first parameter, namely, MWF (Φ), is used in the EEG signal decomposition task. The second WT parameter, namely, the decomposition level (***L***), is also selected in the decomposition phase based on the EEG signal and experience.

The third parameter, namely, thresholding functions (i.e, β)), can be divided into **hard** and **soft** thresholding [47], [62]. Figure 2 shows the difference between hard and soft thresholding. The thresholding types (soft or hard) in the second phase must be selected along with the fourth parameter, namely, the selection rules (λ), and the fifth parameter, namely, the rescaling methods (ρ). These threshold mechanisms must be applied because the selection will affect the global denoising performance. The thresholding value is generally defined based on the standard deviation (σ) of the noise amplitude [9]. Tables 2 and 3 provide the different types of parameters for the thresholding selection rule and rescaling methods. The thresholding rules are selected according to Equation (4).

$$EEG_{noisy}(n) = x(n) + \sigma e(n) \quad (4)$$

where *x(n)* is the original EEG signal, e is the noise, σ is the amplitude of the noise, and n is the number samples. The wavelet parameters (β, λ, and ρ) must be separately applied for each wavelet coefficient (approximation and details) level.

In the last phase, the denoised EEG signal is reconstructed by ***iDWT*** as shown in Eq (3).

TABLE 2. Thresholding selection rules.

Thresholding selection rule	Description
Rule 1: Rigrsure	Threshold is selected using the principle of Stein's Unbiased Risk Estimate (SURE)
Rule 2: Sqtwolog	Threshold is selected equal to $\sqrt{(2logM)}$
Rule 3: Heursure	Threshold is selected according to mixture (***Rigrsure*** and ***Sqtwolog***)
Rule 4: Minimaxi	Threshold is selected equal to ***Max(MSE)***

TABLE 3. The wavelet thresholding rescaling methods.

Wavelet threshold rescaling methods ρ	rescaling
one	No scaling
sln	Single level
mln	Multiple level

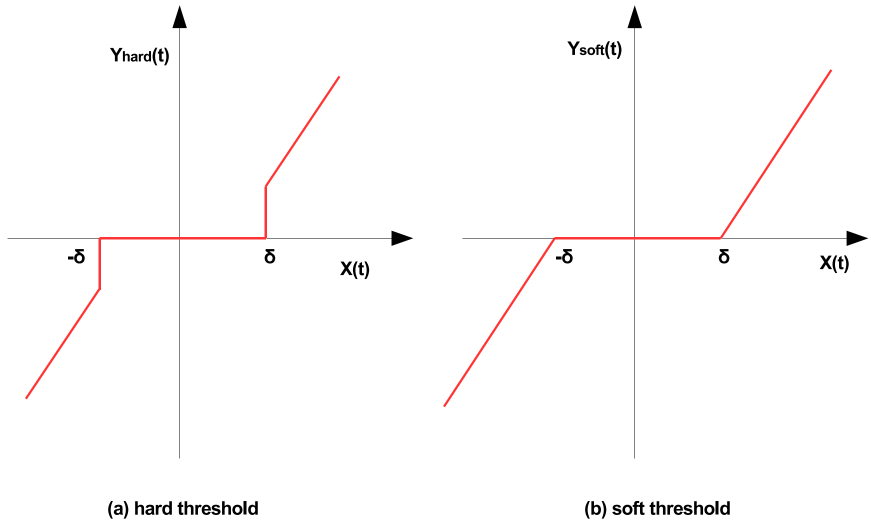


FIGURE 2. Soft and hard thresholding methods.

IV. METAHEURISTIC ALGORITHMS

As previously mentioned, metaheuristic-based approaches are conventionally classified into: evolutionary algorithm [27], [28], swarm intelligence [63], and trajectory algorithms [14]. In this paper, five metaheuristic algorithms are adopted to find the optimal WT parameters for the EEG signal denoising problem. These algorithms are carefully selected based on their performance in solving the signal and image processing problems. These algorithms have shown an excellent ability in solving various signal and image processing problems [10], [11], [28], [29]. The selected metaheuristic-based algorithms are described as follows:

A. GENETIC ALGORITHM

GA was developed in [16] to mimic the natural phenomenon of Darwin evolution theory. Based on the '*survival of the fittest*' principle, GA starts with many solutions, with each solution being a vector of decision variables and each decision variable having a specific range of values. In evolution context, the set of solutions is equivalent to *population*, each solution is analogous to *chromosome*, each decision variable is analogous to *gene*, and each value of the decision variables is analogous to *allele*.

In order to apply a successful GA to COPs, both the objective function and problem representation must be properly adjusted together with parameter tuning. GA typically has a set of parameter, including the size of the population P_{size}, the number of generations P_{no}, the crossover rate $P_{crossover}$, and the mutation rate $P_{mutation}$. In order to build an efficient and robust GA, the parameter settings of each COP must be closely examined. Algorithm 1 shows the high-level schematic pseudo-code of GA that starts with a population of candidate solutions X_{chrom}, where X_{chrom} is an augmented matrix of size $P_{size} \times N$ and N is the number of decision variables in each solution. Initially, the population X_{chrom} is filled with random candidate solutions across the problem search space, that is, $X_{chrom}{=}X_{chrom}{}^1$, $X_{chrom}{}^2, \ldots, X_{chrom}{}^{P_{size}}$. Each candidate solution $X_{chrom}{}^i$ is evaluated based on an objective function. The improvement loop in GA (see Algorithm 1, line 3 to 9) repeats the following steps until a termination criterion is met: select the parents (new population X_{chrom}') that will be used to generate the next population which will pairwise crossover with a probability of $P_{crossover}$ to come up with a new population X_{chrom}''. Afterward, each pairwise solution will be checked if it must be mutated with probability $P_{mutation}$ to come up with X_{chrom}'''. The new population will be reevaluated, and the X_{chrom}''' will be substituted with the population X_{chrom} based on such selection method. This procedure is followed to determine whether the offsprings are fit or not. This process will be repeated several times until an optimal solution is reached.

Algorithm 1 Genetic Algorithm Pseudo-Code

1: $X_{chrom} \leftarrow$ *Generate_Initital_Population*
2: Evaluate(X_{chrom})
3: **while** (Stopping criterion is not met) **do**
4:     $X_{chrom}' \leftarrow$ *Selection*(X_{chrom})
5:     $X_{chrom}'' \leftarrow$ Crossover (X_{chrom}')
6:     $X_{chrom}''' \leftarrow$ Mutation (X_{chrom}'')
7:     Evaluate(X_{chrom}''')
8:     $X_{chrom} \leftarrow$ Replacement ($X_{chrom}''' \cup X_{chrom}$)
9: **end while**

B. HARMONY SEARCH ALGORITHM

Harmony search (HS) is an evolutionary algorithm established by Geem Z.L. [17]. HS has five main procedural steps that are summarized in Algorithm 2 and described as follows:

Step 1: Initialize HS parameters: The HS parameters that are required for solving the optimization problem are specified in this step. These parameters

Algorithm 2 Harmony Search Algorithm Pseudo-Code

```
Set N, X_i, ∀i = 1, 2, ..., N,
Initialize harmony search parameters HMCR, PAR, NI, HMS, FW.
Generate HM solutions
Evaluate (f(x^j)), ∀j = (1, 2, ..., HMS)
while Stop criterion is not met do
  if (U(0, 1) ≤ HMCR) then
    x' = Memory consideration(x)
    if (U(0, 1) ≤ PAR) then
      x'' = Pitch adjustment(x')
    end if
  else
    x'' = Random consideration(x')
  end if
  if (f(x'') < f(x'worst)) then
    Update the HM by include x'' and exclude x'worst.
  end if
end while
```

include the harmony memory consideration rate (HMCR), which determines the rate of selecting the value from the memory, the harmony memory size (HMS), which is similar to the population size in other EAs, pitch adjustment rate (PAR), which determines the probability of local improvement, the fret width (FW), which determines the distance of adjustment, and the number of improvisations (NI) or the number of iterations.

Step 2: Initialize the harmony memory: The harmony memory (HM) is a repository of the population individuals, where $HM = \left[\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^{\text{HMS}}\right]^T$. In this step, these individuals are randomly generated as follows: $x_i^j = LB_i + (UB_i - LB_i) \times U(0, 1), \forall i = 1, 2, \ldots, N$ and $\forall j = 1, 2, \ldots, \text{HMS}$, and $U(0, 1)$ generates a uniform random number between 0 and 1.

Step 3: Improvise a new harmony: A new harmony vector is generated as $\boldsymbol{x}' = (x_1', x_2', \cdots, x_N')$ based on three operators, namely, (1) memory consideration (MC), (2) pitch adjustment (PA), and (3) random consideration (RC). These operators assign a value for each decision variable x_i' in the new harmony as formulated in equation (5):

$$x_i' \leftarrow \begin{cases} x_i' \in \{x_i^1, x_i^2, \ldots, x_i^{\text{HMS}}\} \\ \qquad \text{w.p. HMCR} \times (1 - \text{PAR}) \quad \{\text{MC}\} \\ x_i' = x_i' + U(-1, 1) \times \text{FW} \\ \qquad \text{w.p. HMCR} \times \text{PAR} \quad \{\text{PA}\} \\ x_i' \in \boldsymbol{X}_i \\ \qquad \text{w.p. 1 - HMCR} \quad \{\text{RC}\} \end{cases} \tag{5}$$

Step 4: Update the harmony memory: If better, the new harmony vector, $\boldsymbol{x}' = (x_1', x_2', \cdots, x_N')$, replaces the worst harmony $\boldsymbol{x}^{\text{worst}}$ stored in HM.

Step 5: Check the stop criterion: Repeat steps 3 and 4 of the HS algorithm until the stop criterion (which usually depends on NI) is met.

C. PARTICLE SWARM OPTIMIZATION

The application of PSO in optimization was initially studied in [19]. The PSO algorithm is initialized with a population of candidate solutions called *swarm*. Each candidate solution is called a *particle*, and each particle iteratively vacillates across the search space. In each iteration, each particle is influenced by the position of the best solution that is found in terms of the objective function achieved earlier by itself (*local best*) and by the best solution among the neighbors of the particle (*global best*). Each particle which performance is decided by an objective function is continually attracted to the local and global best. This situation mimics the social behavior of bird flocks [64].

Algorithm 3 shows the pseudo-code of PSO. in which each particle is basically represented by the following characteristics: (*i*) x_i - the current position of particle i; (*ii*) v_i- the current velocity of particle i; (*iii*) y_i - the local best of particle i; (*iv*) $\hat{y}_i$- the global best of particle i. During the improvement loop (see Algorithm 3, lines 6 to 16), these four characteristics are updated for each particle at each time t as follows:

$$y_i(t+1) \leftarrow \begin{cases} y_i(t) & \text{if } f(x_i(t+1)) \geq f(y_i(t)) \\ x_i(t+1) & \text{if } f(x_i(t+1)) < f(y_i(t)) \end{cases} \tag{6}$$

$$\hat{y} = \{y_i(t) | i = \arg \min_{i=1,\ldots,N} f(y_i(t))\} \tag{7}$$

where N is the number of particles in the swarm. In order to update the velocity for each dimension $j \in [1, N_d]$, in Eq. (8), $v_{i,j}$ refers to element j of the velocity vector of particle i. Eq. (8) also combines the following factors: (*i*) $\omega v_{i,j}(t)$: where $v_{i,j}$ is the previous velocity and ω controls the impact of the previous velocity. The larger the value of ω, the greater the concern with exploration. By contrast, the smaller

Algorithm 3 Particle Swarm Optimization Pseudo-Code

```
1:  for i = 1, ..., N do
2:    x_i = GenerateNewPositionForParticle (i)
3:    v_i = 0
4:    y_i = x_i
5:  end for
6:  repeat
7:    for i = 1, ..., N do
8:      f(i) = EvaluateParticle (i)
9:      y_i = Update Using Eq.(6)
10:     ŷ_i = Update Using Eq.(7)
11:     for j = 1, ..., N_d do
12:       v_i = Update Velocity Using Eq.(8)
13:     end for
14:     x_i= Update Using Eq.(9)
15:   end for
16: until (Stopping Criteria is met)
```

the value of ω, the greater the concerns with exploitation. (*ii*) $y_{i,j}(t) - x_{i,j}(t)$ means that the particle i is attracted to the local best direction. (*iii*) $\hat{y}_j(t) - x_{i,j}(t)$ means that the particle i is attracted to the global best direction.

$$v_i(t+1) = \omega v_{i,j}(t) + c_1 r_{1,j}(t)(y_{i,j}(t) - x_{i,j}(t)) + c_2 r_{2,j}(t)(\hat{y}_j(t) - x_{i,j}(t)) \quad (8)$$

where ω is called the 'intra weight' that controls the historical velocity, c_1 and c_2 are two acceleration constants, and r_1 and r_2 generate a uniform distribution random number between 0 and 1, that is, r_1 and $r_2 \sim U(0, 1)$.

The current position of the particle i, is updated as Eq.(7)

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (9)$$

D. FLOWER POLLINATION ALGORITHM

Flower pollination algorithm (FPA) is classified as nature-inspired algorithm which inspired from the pollination behaviour of the flowering plants. FPA is established by Yang in 2012 [21]. In addition, we can summarize the procedure of FPA in four main rules which are describing as follows:

1) Global pollination involves the biotic and cross-pollination where the pollinators are carrying the pollen based on Levy flights.
2) Local pollination involves abiotic and self-pollination.
3) The reproduction probability can be considered as the flower constancy is proportional to the similarity between any two flowers.
4) The switch probability $p \in [0, 1]$ can be controlled between local pollination and global pollination Due to some external factors such as wind, local pollination will be a significant fraction p in the overall pollination activities.

Procedurally, FPA is a swarm-based optimization initiated with a set of provisional solutions. At each iteration, either one of the two operators is invoked: local pollination operator and global pollination operator. In a local pollination operator, the decision variables of the current solution attract the other two randomly selected variables from two population members. In a global pollination operator, the decision variables of the current solution attract the globally best solution found. The switch operator is responsible for exchanging the improvement loop either locally or globally. This process repeats until a stagnation point is obtained.

To illustrate the mechanism of the FPA based on these four rules, three key steps can be described in the following three subsections.

1) GLOBAL SEARCH OF FPA (BIOTIC)

As mentioned previously, in this type of pollination the flowers pollens are transferred by pollinators such as bees, bats, birds..etc. to long distances. This ensures the pollination and reproduction of the most fittest. Therefore, we can represent the first and third FPA rules mathematically as follows:

$$x_i^{t+1} = x_i^t + L(g^* - x_i^t) \quad (10)$$

where x_i^{t+1} the pollen i or solution vector x_i at iteration t, and g^* is the current best solution found among all solutions at the current iteration. The parameter L is the strength of the pollination, which essentially is a step size. Since insects may move over a long distance with various distance steps, we can use a Levy flight to mimic this characteristic efficiently [21], [29], [65]. That is, we draw L>0 from a Levy distribution

$$L \sim \frac{\lambda\Gamma(\lambda) sin(\pi\lambda/2)}{\pi} \frac{1}{s^{1+\lambda}}, \quad (s >> s_0 > 0) \quad (11)$$

$\Gamma(\lambda)$ denotes the standard gamma function, and this distribution is valid for large steps s>0. In all our simulations below, we have used $\lambda = 1.5$.

2) LOCAL SEARCH OF FPA (ABIOTIC)

In this kind of pollination occurs without any pollinators. Where it based on the wind and diffusion to transfer the pollen. The local pollination (rule 2) and flower constancy (rule 3) can we represented as follows:

$$x_i^{t+1} = x_i^t + \epsilon(x_j^t - x_k^t) \quad (12)$$

where x_j^t and x_j^k are pollens from the different flowers of the same plant type. This essentially mimic the flower constancy in a limited neighborhood. Mathematically, if x_j^t and x_j^k comes from the same species or selected from the same population, this become a local random walk if we draw ϵ from a uniform distribution in [0,1].

3) SWITCH PROBABILITY IN FPA

The third key steps that affects in the performance of the FPA is switch probability (rule 4). Where the value of p will determine which path will follow either local or global pollination. To start with, the value of p = 0.5 is initialing used then the author found p = 0.8 is the best value for most applications.

The three key steps can be summarized in the pseudocode of the FPA shown in Algorithm 4.

E. β-HILL CLIMBING ALGORITHM

Hill climbing is a simple trajectory-based method which is an iterative approach that starts with an arbitrary solution to a problem and then progressing the search by means of trying a trajectory in the problem space to find a better solution. If the previous step produced a better solution, an incremental change will continue to find a new solution. This process is repeated until the solution can no longer be improved. The problem with hill climbing algorithm is that only uphill movements are accepted, which leads to getting easily stuck in the local optima [14]. Several extensions have been proposed to overcome this problem. The most recent extension is proposed by Al-Betar in 2016 called β-hill climbing [14], wherein a learning stochastic operator is adapted in hill climbing to strike an efficient balance between both exploration and exploitation during the search.

As aforementioned, the β-hill climbing algorithm is a trajectory search technique that begins with single random

Algorithm 4 Flower Pollination Algorithm Pseudo-Code

Objective *min* $f(x)$, $x \in \Re^d$
Initialize a population of n flowers/pollens with random solution
Fins the best solution g^* in the initial population
Define a switch probability $p \in [0, 1]$
Calculate all (*f(x)*) for n solutions
t=0
while (Stopping criterion is not met) **do**
  for $i = 1, .., N$ **do**
    if *rnd* $\leq p$ **then**
      Draw a (d-dimensional) step vector $\boldsymbol{L}$ which obeys a Levy distribution
      Global pollination via $x_i^{t+1} = x_i^t + L * (g^* - x_i^t)$
    else
      Draw from a uniform distribution $\in$ [0,1]
      Randomly choose j and k among all solution
      Do local pollination via $x_i^{t+1} = x_i^t + \in (x_j^t\text{-}x_k^t)$
    end if
    Calculate(*f(x')*)
    if $f(x') \leq f(x)$ **then**
      $x = x'$
    end if
  end for
  Find the current best solution g^* among all x_i^t
  $t = t + 1$
end while

solution, $x = (x_1, x_2, \ldots, x_N)$. During the searching space, the new solution, $x' = (x'_1, x'_2, \ldots, x'_N)$, will be initiated by updating the current solution using two operators namely: $\mathcal{N}$-operator and β-operator, where these operators represents the sources for exploitation and exploration, respectively. Specifically, the $\mathcal{N}$-operator works as neighbourhood search, while β-operator works as similar to mutation operator. At each iteration, the new solution can be improved by $\mathcal{N}$-operator stage or β-operator stage until the optimal solution is reached.

The algorithm begins to generate the solution randomly, then the solution is evaluated using the objective function $f(\boldsymbol{x})$. The solution is then modified using $\mathcal{N}$-operator, which employs the *improve*($\mathcal{N}(\boldsymbol{x})$) function within a random range of its neighbors. The solution $\boldsymbol{x}$ is as follows:

$$x'_i = x_i \pm U(0, 1) \times bw \quad \exists i \in [1, N]$$

where i is randomly selected from the space range, $i \in [1, 2, \ldots, N]$. The parameter bw represents the bandwidth between the current value and the new value.

In β-operator, within the β range where $\beta \in [0, 1]$, variables of new solution will be assigned based on selected randomly from available range or from the existing values of the current solution as follows:

$$x'_i \leftarrow \begin{cases} x_r & rnd \leq \beta hc \\ x_i & otherwise. \end{cases}$$

where *rnd* generates a uniform random number between 0 and 1 and $x_r \in X_i$ is the possible range for the decision variable x'_i.

Algorithm 5 shows the pseudocode of the β-Hill Climbing Algorithm.

Algorithm 5 β-Hill Climbing Algorithm Pseudo-Code

$\boldsymbol{x}$ = Build Initial Solution
f($\boldsymbol{x}$) = Evaluate the initial Solution
while Stop criterion is not met **do**
  $x' = \mathcal{N} - Operator(x)$
  $x'' = \beta - Operator(x')$
  if $f(x'') \leq f(x)$ **then**
    replace x by x''
  end if
end while

Finally, the β-hill climbing has successfully achieved optimal results in many global problems such as sudoku problem, feature selection, and signal processing [10], [11].

V. META-HEURISTIC ALGORITHMS AND WAVELET TRANSFORM FOR EEG SIGNAL DENOISING: PROPOSED METHOD

This section provide a full discussion for the proposed methodology of the meta-heuristic algorithms with wavelet transform to solve EEG signal denoising problem. Algorithm 6 shows the pseudocode of the proposed method framework. The proposed methodology run through four phases where the result of each phase is an input to the consecutive one. The four phases are presented in Figure 3 and thoroughly described as follows:

Algorithm 6 Tuning WT Parameters Using a Meta-Heuristic Algorithms for EEG Signal Denoising

Initialize noisy EEG signal (nEEG), calculate the SNR, MSE, RMSE, and PRD for input EEG signal.
Initialize meta-heuristic operators, initialize solution(s) $X_i(i = 1, 2, .., N)$ N = 5 wavelet parameters, the initial solution $X_i(\Phi, \boldsymbol{L}, \beta, \lambda, \rho)$
X'_{opt} = Metheuristic (X, X_i)
EEG Denoise Signals = WT (X'_{opt},nEEG)
EEG Out Signals = Evaluate(EEG Denoise Signals, SNR_{out}, SNR_{imp}, MSE, RMSE, PRD).

- ***Phase I: Initialization*** . This phase involves three steps: firstly, reading the input EEG signal $x(n)$ from its source. The WT denoising approach was developed based on the original EEG signal being corrupted with white Gaussian noise (WGN), Power Line Noise (PLN), and Electromyogram (EMG) estimation [9], [31], [32]. Where these noises are exactly simulating the noises which will corrupt the original EEG signal during the recording

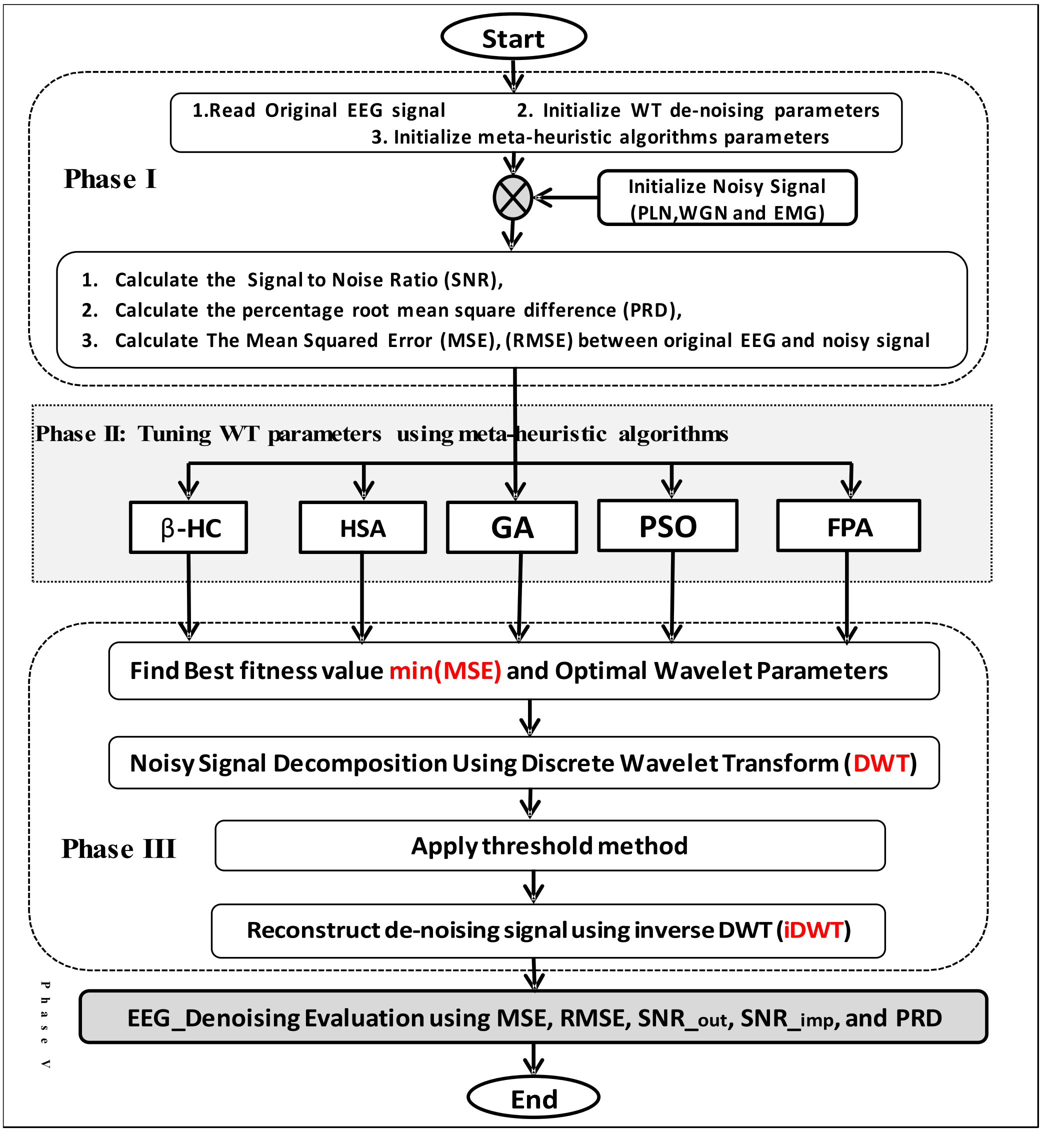


FIGURE 3. Proposed method for EEG denoising.

time such as eye blink noise, eye movement noise, electro signal distortion. The original EEG signals are provided then the signals corrupted by PLN using Eq. (13) followed by signals corrupted by EMG using Eq. (14) followed by signals corrupted by WGN using Eq. (15)) are given. These three types of noises corruption EEG signals are used as a dataset to evaluate the performance of proposed methods.

$$N(t) = A * sin(2 * \pi * f * t) \tag{13}$$

$$N(t) = E * rand(t) \tag{14}$$

$$N(t) = x(t) + \sigma \tag{15}$$

where $A = 60$ uV, $E = (0\text{-}10)$ uV, $f = 60$ Hz,e is the noise, σ is the amplitude of the noise in this work $\sigma = 15\ \mu$V. The N signal is added to the original EEG signal x to simulate PLN, EMG, and WGN respectively.

Secondly, initialize WT denoising parameters (Φ, $\boldsymbol{L}$, β, λ, ρ) which are shown in Table 9, as well as the parameter for each meta-heuristic algorithm is also initialized as shown in Table 4. Finally, compute the signal to noise ratio (SNR) by Eq.(25), percentage of root mean square difference (PRD) by Eq.(24), mean square error (MSE) by Eq.(16), and root mean square error (RMSE) by Eq. (27). This is to record the results of EEG signals before and after denoising process.

- ***Phase II:Tuning WT parameters by meta-heuristic algorithm***. In the proposed methodology, such meta-heuristic algorithm discussed in Sec.IV is adapted to find the optimal WT parameters which can be used for EEG signal denoising problem. Initially, the solution of WT parameters configuration is represented as a vector $x = (x_1, x_2, \ldots x_n)$ where n is the total number of parameter used for WT which is normally equal to 5. x_1 represent

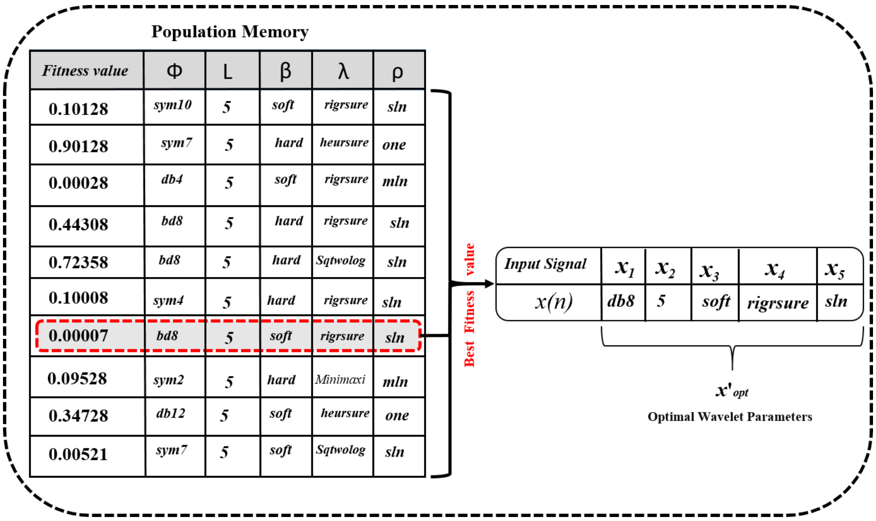


FIGURE 4. Optimal solution of WT parameters for denoising EEG signals using meta-heuristic.

TABLE 4. Meta-heuristic algorithms parameters.

meta-heuristic method	Parameters
β-HC-WT	β=0.5,$\mathcal{N}$=5, and D=1, FEN=10000
HSA-WT	HMCR=0.95,PAR=0.1, $\mathcal{N}$=5, and D=20, FEN=1000
PSO-WT	c_1=2,c_2=2, $\mathcal{N}$=5, and D=20, FEN=1000
GA-WT	$P_{mutation}$=0.2, $P_{crossover}$=0.9, $\mathcal{N}$=5, and D=20, FEN=1000
FPA-WT	P=0.8, $\mathcal{N}$=5, and D=20, FEN=1000

the value of mother wavelet function parameter Φ, x_2 denotes the value of decomposition level parameter $\boldsymbol{L}$, x_3 refers to the thresholding method β, x_4 represents the value of thresholding selection rule parameter λ, and x_5 represents the re-scaling approach ρ, where the possible range for these parameters are selected from Table 1. Fig. 4 shows an example solution of WT parameters for denoising EEG signals. The selected metaheuristic algorithm evaluates the solution using the MSE objective function which is formulated in Eq.(16).

$$MSE = \frac{1}{N}\sum_{n=1}^{N}[x(n) - \widehat{x}(n)]^2 \tag{16}$$

where $x(n)$ denotes the original EEG signal and $\widehat{x}(n)$ is the denoised EEG signal obtained by tuning the wavelet parameters using the meta-heuristic algorithm.
Iteratively, the randomly generated solution(s) undergoes refinement using the selected meta-heuristic algorithm. The final output of this phase is an optimized solution $x'_{opt} = (x'_1, x'_2, \ldots x'_n)$ which will be passed to the next phase.

- ***Phase III:EEG denoising using WT based on x'_{opt}.*** As aforementioned in Sec. III-A, the denoising process of WT involves three main steps that are visualized in Figure 1 and described in more details below:
  - EEG signal ***decomposition*** using DWT. In this step the DWT is applied to decompose the noise of the input EEG signals $x(n)$. In decomposition process, we must use the first two x'_{opt} parameters, namely, the mother wavelet furcation ρ and the decomposition level $\boldsymbol{L}$). Figure 5(a) shows the DWT procedure for two levels as example, where the noisy EEG signal is divided at each level into cA and cD. The latter is processed using a high-pass filter, while the former is processed using a low-pass filter and is decomposed for the next level. The EEG signal is convolved using the high-pass and low-pass filters, while the block($\downarrow$2), which is represented by the downsampling operator, is used to keep the even index elements of the EEG signal. The EEG signals are separated into cA and cD based on their frequency and amplitude.
  - The second step of EEG denoising is ***Thresholding*** which is applied based on the noise level of the coefficients. In this step, the last three wavelet parameters, namely, the thresholding type (β), the thresholding selection rules (λ), and the re-scaling methods (ρ), must be selected from x'_{opt}. According to [66], using a thresholding operation on the input noisy non-stationary signal $\widehat{X}$ can estimate the denoised EEG signal as follow:

$$Z = THR(\widehat{X}, \delta), \tag{17}$$

where the THR denotes a thresholding function, while δ denotes a threshold value. The EEG denoising performance in the wavelet domain depends on the estimation of δ. Therefore, several methods have been proposed for estimating δ. Donoho and Johnstone [47] calculated the threshold δ on an

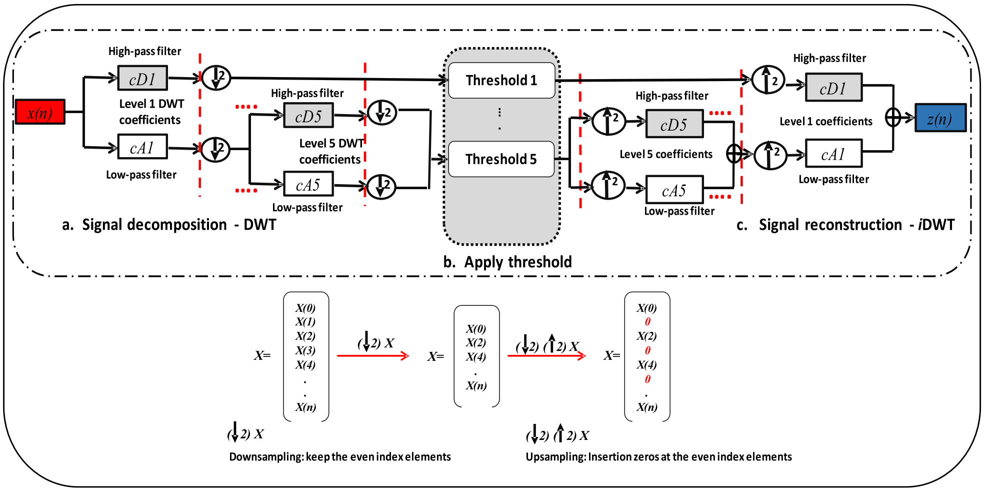


FIGURE 5. **EEG denoising procedure.**

orthonormal basis as follows

$$\delta = \sigma\sqrt{2logM} \tag{18}$$

where σ represents the standard deviation of DWT detail coefficients, while M denotes the length vector of the DWT coefficients. Given that the threshold value δ only depends on cD and that cA has a low frequency EEG signal and the highest amount of energy. We estimate the value of δ based on the coefficients level as follows:

$$\widehat{x}_d(l) = THR(\widehat{x}_d(l), \delta_l), \quad l = 1, 2, \ldots. \tag{19}$$

where $\widehat{x}_d$ represents a vector of threshold DWT detail coefficients, l denotes a wavelet decomposition level, and δ_l denotes the threshold value determined for that level. The wavelet generally provides two standard types of thresholding functions (β), namely, **hard** and **soft** thresholding [47], [62]. As shown in Figure 2, different between hard and soft thresholding are described as follows:

$$\widehat{x}_{di}(l) = \begin{cases} |\widehat{x}_{di}(l)| - \delta_l & |\widehat{x}_{di}(l)| \geq \delta_l \\ 0 & |\widehat{x}_{di}(l)| < \delta_l \end{cases} \tag{20}$$

$$\widehat{x}_{di}(l) = \begin{cases} \widehat{x}_{di}(l) & |\widehat{x}_{di}(l)| \geq \delta_l \\ 0 & |\widehat{x}_{di}(l)| < \delta_l \end{cases} \tag{21}$$

where i denotes the index of the DWT details coefficients at a level l. The thresholding DWT coefficients can be expressed as follows:

$$\widehat{X} = [\widehat{x}_d(1) \quad \widehat{x}_d(2) \quad \widehat{x}_a(2)] \tag{22}$$

- ***Reconstruction of the denoising EEG signal by* iDWT**. We estimate the value of the original EEG signals $\widehat{X}$ by applying *i*DWT on $\widehat{X}$ as follows:

$$z[n] = \sum\nolimits_{k=-\infty}^{\infty} cA_L(k)\, \phi_i'(n-k) + \sum\nolimits_{i=1}^{L} \sum\nolimits_{k=-\infty}^{\infty} cD_{i+1}(k)\, \Psi_i'(n-k) \tag{23}$$

The reconstruction convolves the EEG signals using upsampling ($\uparrow 2$), which involves the insertion of zeros at the even index elements of EEG signals. Figure 5(c) shows the *i*DWT procedure for five levels as an example.

- ***Phase V: EEG Denoising Evaluation*** . The final phase is evaluating the EEG output of WT. The evaluation will done based on five criteria which are: Signal-to-Noise-Ration (SNR), SNR improvement, Mean Square Error (MSE) eq. (16), Root Mean Square Error (RMSE), and percentage root mean square difference (PRD).

$$PRD = 100 * \sqrt{\frac{\sum_{n=1}^{N}[x(n) - \widehat{x}(n)]^2}{\sum_{n=1}^{N}[x(n)]^2}} \tag{24}$$

$$SNR_{out} = 10\log_{10}\left\{\frac{\sum_{n=1}^{N}[x(n)]^2}{\sum_{n=1}^{N}[x(n) - \widehat{x}(n)]^2}\right\} \tag{25}$$

$$SNR_{imp} = 10\log_{10}\left\{\frac{\sum_{n=1}^{N}[\delta(n) - x(n)]^2}{\sum_{n=1}^{N}[x(n) - \widehat{x}(n)]^2}\right\} \tag{26}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}[x(n) - \widehat{x}(n)]^2} \tag{27}$$

where *x(n)* denotes the original EEG signal, $\widehat{x}(n)$ is the denoised EEG signal obtained by tuning the wavelet parameters through the selected meta-heuristic algorithms, and N is the sampling number.
The final decision about the denoise results are decided by comparing the original criteria (i.e., SNR, MSE, RMSE, PRD) with improved one (i.e., SNR_{out}, SNR_{imp}, MSE, RMSE, PRD).

VI. RESULTS AND DISCUSSIONS

In order to check the performance validity of the proposed WT-based metaheuristic approaches, an extensive and exhaustive evaluation procedure, which will be described in this section, is applied. The EEG datasets used in this study are fully explained in Section VI-A. A comparative analysis

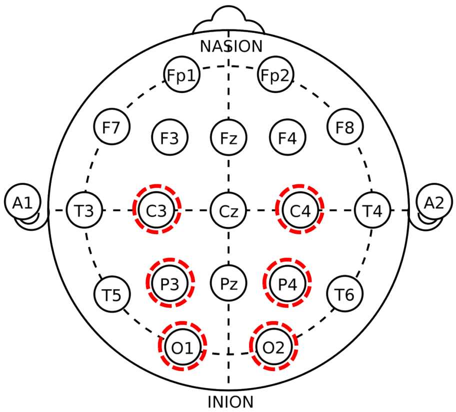


FIGURE 6. Distribution of electrodes in Keirn's EEG dataset.

among the proposed metaheuristic algorithms is provided in Section VI-B. The results of the metaheuristic with the best performance are compared with those of other well-established methods. The results of the EEG signals denoising using WT with and without metaheuristic algorithms are compared i) by comparing the results obtained by several approaches without optimization such as [1], [33] as described in Section VI-C. ii) by comparing the results of the metaheuristic algorithms to determine the best algorithm for EEG signal denoising using WT parameters as described in Section VI-B, and iii) by comparing the results of the best metaheuristic algorithms for EEG signal denoising using wavelet with those of algorithms without optimization as described in Section VI-C.

A. EEG DATASET

The meta-heuristic algorithms are tested using two standard EEG signal datasets, namely, Keirnś EEG dataset[3] [30] and the 'Motor Movement/Imagery'[4] [15].

Keirn's EEG dataset recorded EEG from seven subjects, and the EEG signals were recorded from six electrodes, namely, C3, C4, P3, P4, O1, and O2. Figure 6 shows the distribution of these EEG electrodes as recorded in [30]. During the recording time, each volunteer is given five mental tasks, while each task is repeated thrice for two sessions. The recording is performed with both rest eyes closed (REC) and rest eyes open (REO) in each session. The period of recording for each task was 10 seconds with a sampling rate of 250Hz per second. This dataset can be considered a small dataset with seven subjects (males and females between the ages of 21 and 48). The relevance of this database lies in the multi-task recording paradigm.

The Motor Movement/Imagery dataset [15] collected the EEG signals from 109 healthy subjects using a brain-computer interface software called BCI2000 system. The EEG signals are recorded using 64 Electrodes (EEG channels) with sampling rate of 160Hz per second, where each signal is stored in a separate EDF file. Each volunteer performs several motor/imagery tasks that are mainly used in different fields, such as neurological rehabilitation and brain-computer interface applications. In general, these tasks consist of imagining or simulating a given action, such as opening and closing the eyes. The EEG signals are recorded from each volunteer by asking them to perform four tasks according to the position of a target that appears on the screen placed in front of them. If the target appears on the right or left side of the screen, then the volunteer must open and close his/her fist corresponding to the position of the target on the screen. If the target appears on the top or bottom of the screen, then the volunteer must open and close his/her fists or feet. Fig. 7 shows the distribution of electrodes in the EEG Motor Movement/Imagery Dataset.

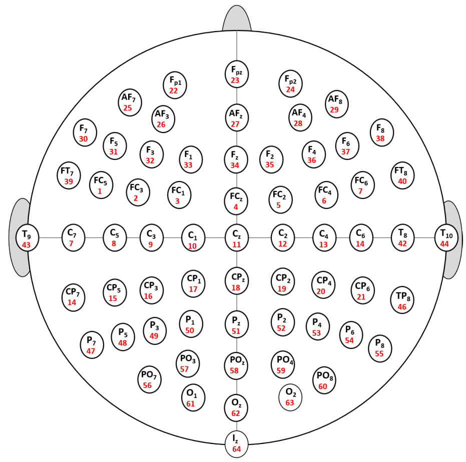


FIGURE 7. Distribution of electrodes in EEG motor movement/imagery dataset.

B. A COMPARATIVE ANALYSIS OF THE PROPOSED METAHEURISTIC METHODS

This paper aims to find the optimal wavelet EEG denoising parameters by using five metaheuristic algorithms. The optimal results of these algorithms will be passed to WT to denoise the EEG signals with the least MSE value. Therefore, the results of these algorithms are compared in this section to determine the most efficient metaheuristic algorithm that can feed WT with the right parameter configurations and empower its final outcomes. As mentioned above, the proposed metaheuristic algorithms include FPA-WT, GA-WT, HSA-WT, PSO-WT, and βHC-WT. The parameter values of these algorithms are reported in Table 4. These parameter values are set by performing several trial-and-error processes to find the best configurations that resemble what has been suggested in the literature [1], [33].

As aforementioned, WT parameters selection is typically performed based on experience or empirical evidence. Therefore, the metaheuristic algorithms are proposed as optimization methods for finding the optimal wavelet EEG denoising parameters. Table 4 shows the metaheuristic

[3] http://www.cs.colostate.edu/eeg/main/data/1989_Keirn_and_Aunon
[4] https://www.physionet.org/physiobank/database/eegmmidb/

TABLE 5. The optimal wavelet denoising parameters obtained by selected meta-heuristic Algorithms for PLN, EMG, and WGN noise for [30] EEG dataset.

Algorithm	Noise	Best $F(x)$	W.Rescaling	Thr. Sel	Thr. type	Decom. Level	Wavelet Fun.
β-HC-WT	PLN	0.0187	heursure	soft	one	5	db17
HSA-WT	PLN	1.6490	rigrsure	soft	sln	5	bior3.9
PSO-WT	PLN	0.0226	rigrsure	soft	mln	5	coif5
GA-WT	PLN	0.0161	heursure	hard	one	5	demy
FPA-WT	**PLN**	**0.0144**	**heursure**	**hard**	**one**	**5**	**db27**
β-HC-WT	EMG	0.1234	heursure	soft	one	5	bior6.8
HSA-WT	EMG	10.5265	heursure	soft	mln	5	coif3
PSO-WT	EMG	0.0132	heursure	soft	one	5	bior2.4
GA-WT	EMG	0.0103	rigrsure	hard	mln	5	sym8
FPA-WT	**EMG**	**0.0098**	**rigrsure**	**hard**	**one**	**5**	**db1**
β-HC-WT	WGN	26.7800	heursure	soft	sln	5	db10
HSA-WT	WGN	30.5982	rigrsure	soft	one	5	db39
PSO-WT	WGN	26.3133	heursure	soft	mln	5	bior3.3
GA-WT	**WGN**	**24.7403**	**heursure**	**hard**	**mln**	**5**	**bior5.5**
FPA-WT	**WGN**	**24.7403**	**heursure**	**soft**	**sln**	**5**	**db35**
βHC-WT	EOG	0.960	minimaxi	hard	One	5	db41
HSA-WT	EOG	0.960	minimaxi	hard	One	5	db41
PSO-WT	EOG	0.0011	minimaxi	hard	mln	5	sym7
GA-WT	EOG	0.0012	sqtwolog	hard	sln	5	db9
FPA-WT	**EOG**	**0.0010**	**heursure**	**soft**	**one**	**5**	**bior3.9**

Bold value indicates best solution achieved ($F(x)$ lowest is best)

TABLE 6. The optimal wavelet denoising parameters obtained by selected meta-heuristic Algorithms for PLN, EMG, and WGN noise for Motor_Imaging EEG dataset.

Algorithm	Noise	Best $F(x)$	W.Rescaling	Thr. Sel	Thr. type	Decom. Level	Wavelet Fun.
βHC-WT	PLN	0.0115	heursure	hard	one	5	coif4
HSA-WT	PLN	0.0538	heursure	soft	one	5	db25
PSO-WT	PLN	0.0113	rigrsure	hard	one	5	db45
GA-WT	PLN	0.0112	heursure	hard	one	5	db19
FPA-WT	**PLN**	**0.0111**	**rigrsure**	**hard**	**one**	**5**	**coif3**
βHC-WT	EMG	0.0166	rigrsure	hard	one	5	bior2.2
HSA-WT	EMG	46.9164	rigrsure	hard	mln	5	db28
PSO-WT	EMG	0.0073	rigrsure	hard	one	5	sym2
GA-WT	EMG	0.0072	heursure	hard	one	5	db12
FPA-WT	**EMG**	**0.0071**	**rigrsure**	**hard**	**one**	**5**	**db13**
βHC-WT	WGN	27.2010	rigrsure	soft	mln	5	sym2
HSA-WT	WGN	31.3605	heursure	hard	one	5	bior3.5
PSO-WT	WGN	24.7363	rigrsure	hard	sln	5	demy
GA-WT	WGN	24.4624	heursure	hard	mln	5	db24
FPA-WT	**WGN**	**23.8184**	**rigrsure**	**soft**	**sln**	**5**	**db31**

Bold value indicates best solution achieved ($F(x)$ lowest is best)

algorithms parameters. FEN is a parameter to determine the number of iterations used in the experiments.

The WT denoising parameters are typically selected based on experience or empirical evidence. Therefore, to find the optimal WT EEG denoising parameters, the five metaheuristic algorithms are adapted. Table (5, 6) shows the optimal wavelet denoising parameters that are selected x'_{opt} using five metaheuristic algorithms for PLN, EMG, and WGN noise datasets [15], [30].

The results above are obtained by implementing the selected meta-heuristic Algorithms on a ***LENOVO Ideapad 310, Intel Core i7, RAM 8G***, using MATLAB R2014a.

To evaluate the final results, five criteria are used, including SNR, SNR improvement, PRD, MSE, and RMSE. Table 7 shows the results of the WT EEG signal denoising using five metaheuristic algorithms for the PLN, EMG, and WGN noise datasets [30]. The results show that FPA has successfully achieved an efficient EEG signal denoising based on the measurement factors for PLN, EMG, and WGN noises. For PLN noise, FPA obtains values of 0.0144, 0.1200, 30.5449, -3.7858, and 2.9700 for MSE, RMSE, SNR_{out}, SNR_{imp}, and PRD, respectively. For EMG noise, FPA obtains values of 0.0098, 0.0990, 33.6418, -2.4149, and 2.0793 for MSE, RMSE, SNR_{out}, SNR_{imp}, and PRD, respectively. For WGN noise, FPA obtains values of 24.7403, 4.3497, 2.2045, 2.0730, and 78.7682 for MSE, RMSE, SNR_{out}, SNR_{imp}, and PRD, respectively. Figure 8 visualizes the results of the selected metaheuristic algorithms based on MSE, RMSE, SNR_{out}, SNR_{imp}, and PRD. Figures (9, 10, 11) shows the EEG signal denoising results for PLN, EMG, and WGN respectively, where the proposed methods have obtained efficient EEG signal denoising results. However, in some cases,

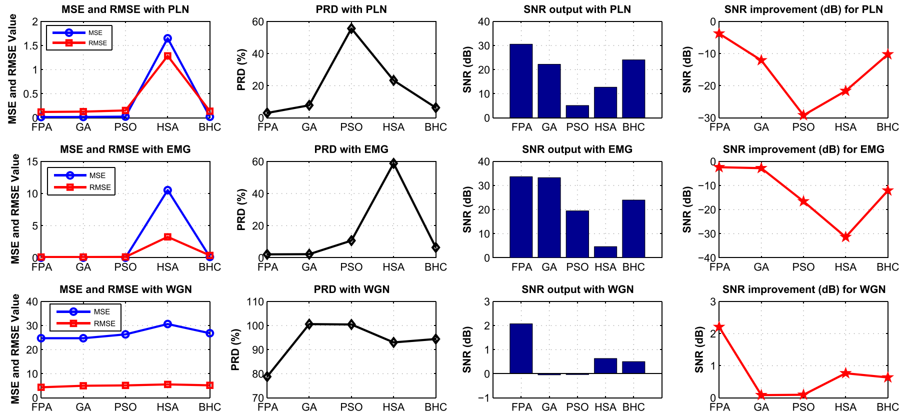


FIGURE 8. Performance of meta-heuristic algorithms for EEG signal denosing.

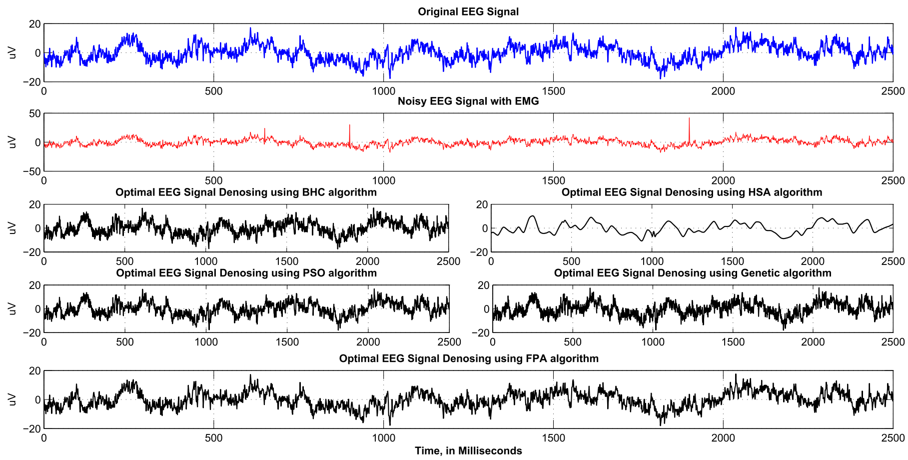


FIGURE 9. EEG signal denoising using meta-heuristic algorithms.

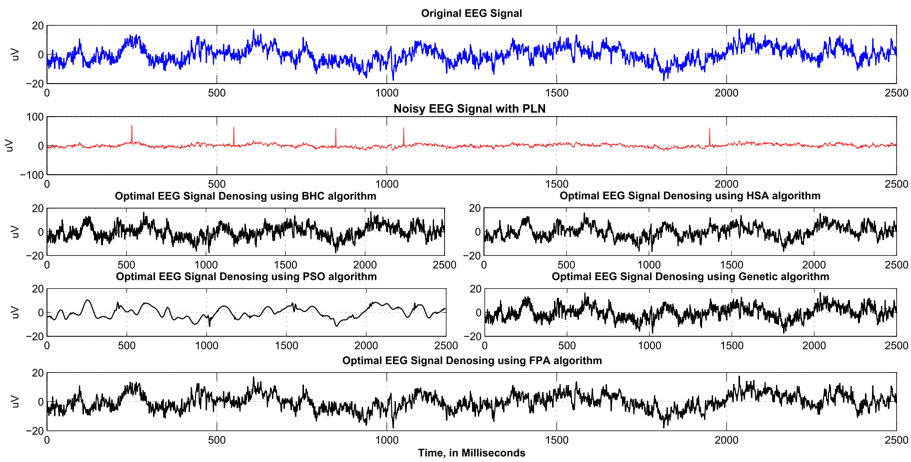


FIGURE 10. EEG signal denoising using meta-heuristic algorithms.

these methods have omitted most of the signal energy during the denoising process. These methods include the HS algorithm in Figure (10), PSO in Figure (9), and PSO and GA in Figure (11).

Table (8) shows the WT EEG signal denoising results using the five metaheuristic algorithms for PLN, EMG, and WGN, respectively [15]. FPA has successfully denoised EEG signals based on the measurement factors for PLN, EMG, and WGN noises. For PLN noise, FPA obtains values of 0.0111, 0.1054, 54.2624, -0.0397, and 0.1920 for MSE, RMSE, SNR_{out}, SNR_{imp}, and PRD, respectively. For EMG noise, FPA obtains values of 0.0071, 0.0843, 56.1710, -0.0240, and 0.1554 for

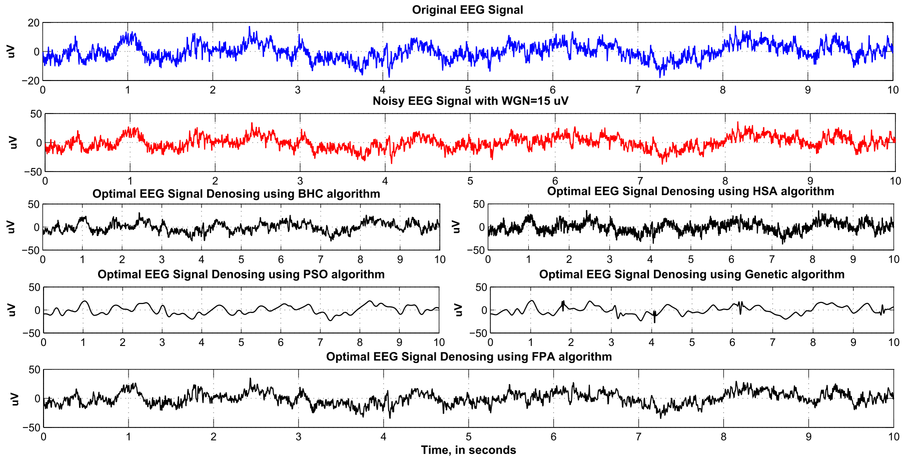


FIGURE 11. EEG signal denoising using meta-heuristic algorithms.

TABLE 7. Performance of denoising the EEG signals for 5 meta-heuristic Algorithms for PLN, EMG, and WGN for Kiern's Dataset.

Algorithm	Noise	SNR_{out}	SNR_{imp}	PRD	MSE	RMSE	Time (Sec)
βHC-WT	PLN	24.0674	-10.2633	6.2608	0.0187	0.1367	6.26E+02
HSA-WT	PLN	12.6701	-21.6605	23.2537	1.6490	1.2841	5.34E+02
PSO-WT	PLN	5.1023	-29.2284	55.5760	0.0226	0.1503	8.54E+02
GA-WT	PLN	22.2106	-12.1201	7.7530	0.0161	0.1269	6.80E+01
FPA-WT	**PLN**	**30.5449**	**-3.7858**	**2.9700**	**0.0144**	**0.1200**	5.66E+02
βHC-WT	EMG	23.9283	-12.1283	6.3619	0.1234	0.3513	7.50E+02
HSA-WT	EMG	4.6194	-31.4373	58.7532	10.5265	3.2445	5.10E+02
PSO-WT	EMG	19.4463	-16.6103	10.6582	0.0132	0.1149	6.92E+02
GA-WT	EMG	33.2292	-2.8275	2.1804	0.0103	0.1015	8.20E+01
FPA-WT	**EMG**	**33.6418**	**-2.4149**	**2.0793**	**0.0098**	**0.0990**	6.62E+02
βHC-WT	WGN	0.5005	0.6321	94.4007	26.7800	5.1749	5.88E+02
HSA-WT	WGN	0.6313	0.7628	92.9900	30.5982	5.5316	4.10E+02
PSO-WT	WGN	-0.0376	0.0940	100.4340	26.3133	5.1296	5.83E+02
GA-WT	WGN	-0.0490	0.0825	100.5659	24.7403	4.9740	7.73E+01
FPA-WT	**WGN**	**2.2045**	**2.0730**	**78.7682**	**24.7403**	**4.3497**	6.81E+02
βHC-WT	EOG	28.575	14.5599	3.7261	0.9600	0.9797	5.86E+02
HSA-WT	EOG	28.575	14.5599	3.7261	0.9600	0.9797	4.41E+02
PSO-WT	EOG	15.2342	11.8282	17.3097	0.001	0.0329	5.82E+02
GA-WT	EOG	17.8881	12.5256	12.7525	0.001	0.0329	7.00E+01
FPA-WT	**EOG**	**36.3513**	**15.6052**	**1.5221**	**0.001**	**0.0329**	5.69E+02

Bold value indicates best results where for SNR, SNRimp, highest is best and for MSE, RMSE, and PRD, lowest is best

MSE, RMSE, SNR_{out}, SNR_{imp}, and PRD, respectively. For WGN noise, FPA obtains values of 23.8184, 4.8804, 0.6180, 0.4677, and 94.7577 for MSE, RMSE, SNR_{out}, SNR_{imp}, and PRD, respectively. Figure (12) visualizes the results of the selected metaheuristic algorithms based on MSE, RMSE, SNR_{out}, SNR_{imp}, and PRD. In addition, Figures (13,14,15) shows the EEG signal denoising results according to PLN, EMG, and WGN respectively. Notably, the proposed WT-based metaheuristic algorithms have obtained efficient EEG signal denoising results, but not in all cases. The main shortcoming of the denoising process is that this process neglects the most useful energy signal for HS algorithm as shown in Figure (14) and for β-hc and GA as shown in Figure (15).

In general, the FPA has many advantages comparing with other metaheuristic algorithms such that FPA is easier to implement, it has fewer parameters which need to tuning as well as it has a stronger ability in the exploitation of the problem space. Overall, The FPA shows a robust ability and dominance for solving real-world optimization problems rather than other metaheuristic techniques [65].

Also, the computational time for the proposed metaheuristic algorithms has been computed in seconds, as listed in Tables (7) and (8). The GA-WT achieved the optimum results for the overall EEG noises in which reaching the optimal WT parameters consumed less than 2 min. This result is due to the GA does not require processing all solutions individually on the basis of the approach to solving the optimization problem and on the basis of the GA operations. Other proposed metaheuristic algorithms have achieved the optimal WT parameters in approximately 10 min, thereby implying feasible future enhancements to improve competitional time. Figure 16 depicts the competitional time for the proposed methods using different EEG datasets.

TABLE 8. Performance of denoising the EEG signals for WT-based meta-heuristic Algorithms according to PLN, EMG, and WGN for EEG motor imaging dataset.

Algorithm	Noise	SNR_{out}	SNR_{imp}	PRD	MSE	RMSE	Time (Sec)
βHC-WT	PLN	54.2216	-0.0805	0.1945	0.0115	0.1072	2.65E+02
HSA-WT	PLN	47.5091	-6.7930	0.4213	0.0538	0.2319	1.37E+02
PSO-WT	PLN	51.3788	-2.9232	0.2698	0.0113	0.1063	2.10E+02
GA-WT	PLN	54.1121	-0.1899	0.1930	0.0111	0.1054	2.83E+01
FPA-WT	**PLN**	**54.2624**	**-0.0397**	**0.1920**	**0.0111**	**0.1054**	2.60E+02
βHC-WT	EMG	52.6254	-3.5697	0.2337	0.0166	0.1288	2.75E+02
HSA-WT	EMG	6.3358	-49.8593	48.2184	46.9164	6.8496	1.61E+02
PSO-WT	EMG	55.3772	-0.8179	0.1703	0.0073	0.0854	2.21E+02
GA-WT	EMG	56.1401	-0.0550	0.1560	0.0072	0.0849	2.48E+01
FPA-WT	**EMG**	**56.1710**	**-0.0240**	**0.1554**	**0.0071**	**0.0843**	2.25E+02
βHC-WT	WGN	0.1851	0.3354	97.8912	27.2010	5.2155	1.04E+02
HSA-WT	WGN	-0.1503	-1.1413e-06	101.7451	31.3605	5.6000	1.44E+02
PSO-WT	WGN	0.2031	0.3533	97.6893	24.7363	4.9736	2.34E+02
GA-WT	WGN	0.0045	0.1548	99.9481	24.4624	4.9459	2.77E+01
FPA-WT	**WGN**	**0.4677**	**0.6180**	**94.7577**	**23.8184**	**4.8804**	2.40E+02

Bold value indicates best results where for SNR, SNRimp, highest is best and for MSE, RMSE, and PRD, lowest is best

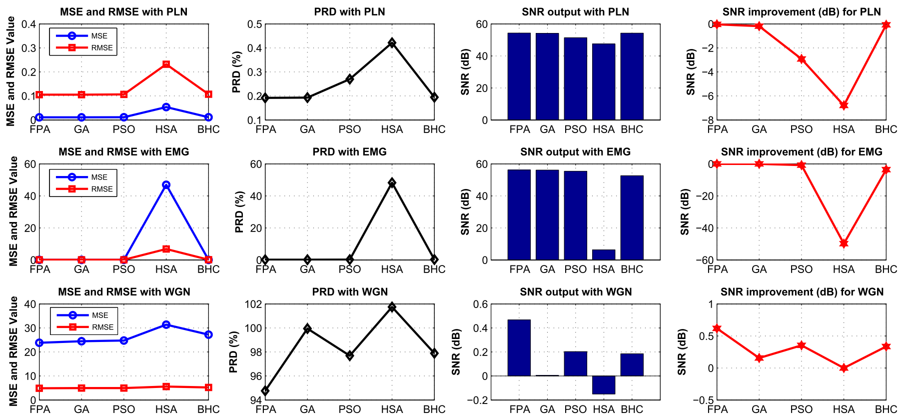


FIGURE 12. EEG signal denoising using meta-heuristic algorithms.

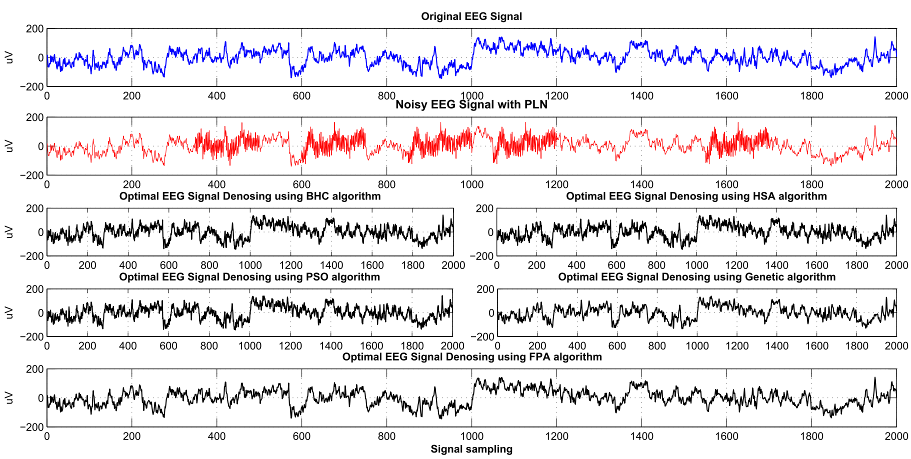


FIGURE 13. EEG signal denoising using meta-heuristic algorithms.

C. COMPARING THE PROPOSED METHOD WITH STATE-OF-THE-ART METHODS

In this section, two state-of-the-art methods for EEG signal denoising are discussed, namely, the Al-Qazzaz method [33] and the Kumari method [1]. These methods use WT for solving EEG signal denoising problems in which the WT parameters are set based on a comparative study. The best parameter configurations for WT as identified by these two methods are shown in Table 9.

We compare the results of these two methods with those generated by our proposed FPA-WT method, which outperforms the other methods as shown in the previous subsection.

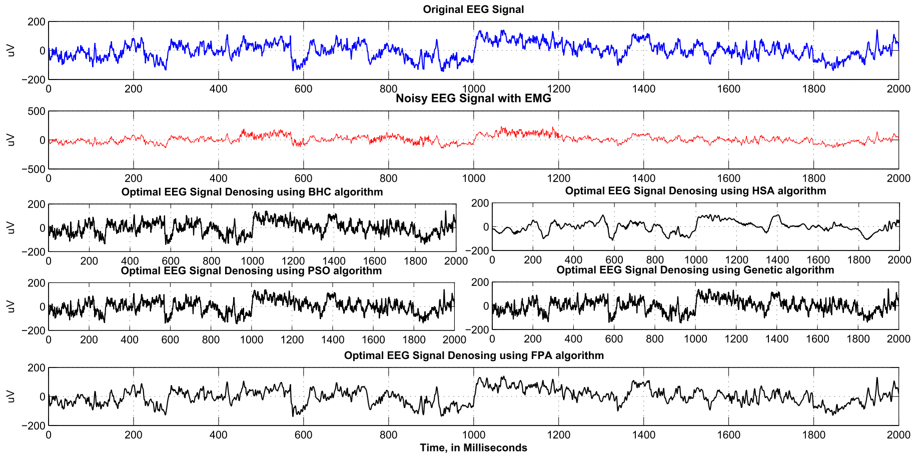


FIGURE 14. EEG signal denoising using meta-heuristic algorithms.

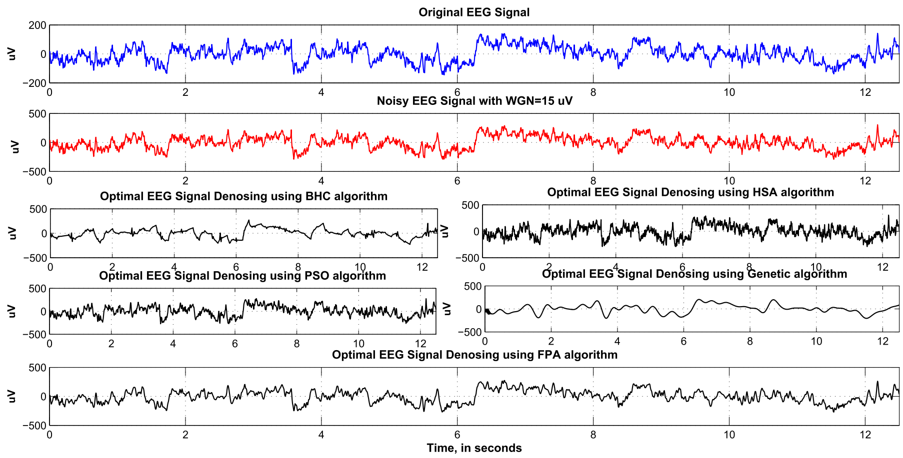


FIGURE 15. EEG signal denoising using meta-heuristic algorithms.

TABLE 9. Wavelet parameters range for Al-Qazzaz and kumari methods.

Wavelet parameters	Al-Qazzaz method	Kumari method
Mother wavelet (ϕ)	**Symlet** (sym9)	**Daubechies** (db4)
Decomposition level ($\boldsymbol{L}$)	5	5
Thresholding type (β)	soft and hard	soft and hard
Selection method (λ)	Rigrsure	Rigrsure
Rescaling approach (ρ)	sln, one	sln, one

The comparison is performed based on Kiern's dataset [30], where the original EEG signal is corrupted with WGN, PLN, and EMG [9], [31], [32]. The final results are evaluated using five criteria, namely, MSE, RMSE, SNR, SNR_{imp}, and PRD. Table 10 shows the EEG signal denoising results of the Al-Qazzaz, Kumari, and FPA-WT methods. The first column presents the ranking of each method based on the evaluation criteria adopted.

According to [7], the EOG is a major EEG artifact which can corrupt the original signal during recording time. The efficient method for removing the EOG artifacts from the original EEG will help obtain useful feature extraction and enhance classification rate accuracy. Therefore, the proposed method is applied to remove the EOG artifacts for the original signal where these artifacts were recorded for the same subjects. Tables 5 and 7 list the results of testing the proposed method using EOG artifacts. The results were evaluated using five measures, namely, MSE, RMSE, SNR_Out, SNR_imp, and PRD). The performance of the proposed method (FPA-WT) has been compared with two state-of-the-art methods [1], [33]; the results show that the proposed method achieves better outputs than [1], [33], as summarized in Table (10), in terms of the overall EEG signal denoising criteria.

Figure 17 proves that the proposed FPA-WT method outperforms both the Al-Qazzaz and Kumari methods for EEG signal denoising based on different noises. FPA-WT obtains the best results for WGN and EMG based on MSE, RMSE,

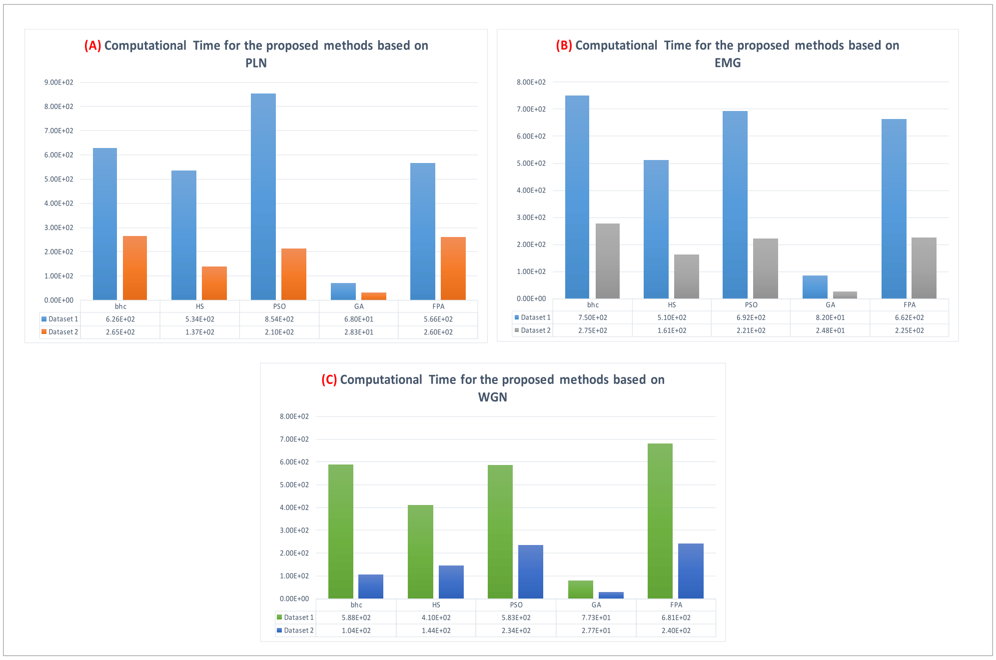


FIGURE 16. Computational Time for meta-heuristic algorithms using different EEG datasets.

TABLE 10. Comparing the proposed FPA-WT method with state-of-the-art methods for EEG signals denoising with different noises .

Rank	Method	Noise	MSE	SNR	SNRimp	PRD	RMSE	(ϕ)	L	β	λ	ρ
1	**Proposed method FPA-WT**	WGN	**24.7403**	**2.2045**	**2.0730**	**78.7682**	**4.3497**	**db35**	**5**	**Soft**	**heursure**	**sln**
2	Al-Qazzaz method [33]	WGN	26.186927	0.661388	0.792952	92.668167	5.117316	sym9	5	Soft	rigrsure	sln
3	Kumari method [1]	WGN	27.006156	0.527605	0.6592	94.106513	5.196744	db4	5	Soft	rigrsure	sln
1	**Al-Qazzaz method** [33]	PLN	0.025316	**30.808240**	**-3.522428**	**2.881296**	0.1591	**sym9**	**5**	**hard**	**rigrsure**	**one**
2	**Proposed method FPA-WT**	PLN	**0.0144**	30.5449	-3.7858	2.9700	**0.1200**	**db27**	**5**	**hard**	**heursure**	**one**
3	Kumari method [1]	PLN	0.030888	29.944328	-4.386341	3.182610	0.196744	db4	5	hard	rigrsure	one
1	**Proposed method FPA-WT**	EMG	**0.0098**	**33.6418**	**-2.4149**	**2.0793**	**0.0990**	**db1**	**5**	**hard**	**rigrsure**	**one**
2	Kumari method [1]	EMG	0.015076	33.059211	-2.99741	2.223511	0.122786	db4	5	hard	rigrsure	one
3	Al-Qazzaz method [33]	EMG	0.019144	32.021900	-4.034729	2.505561	0.138361	sym9	5	hard	rigrsure	one
1	**Proposed method FPA-WT**	EOG	**0.001**	**36.3513**	**15.6052**	**1.5221**	**0.0329**	**bior3.9**	**5**	**soft**	**heursure**	**one**
2	Al-Qazzaz method [33]	EOG	3.8699	22.4421	13.5106	7.5491	1.9672	sym9	5	hard	rigrsure	one
3	Kumari method [1]	EOG	4.6352	21.6583	13.3562	8.262	2.153	db4	5	hard	rigrsure	one

Bold value indicates best results where for SNR, SNRimp, highest is best and for MSE, RMSE, and PRD, lowest is best

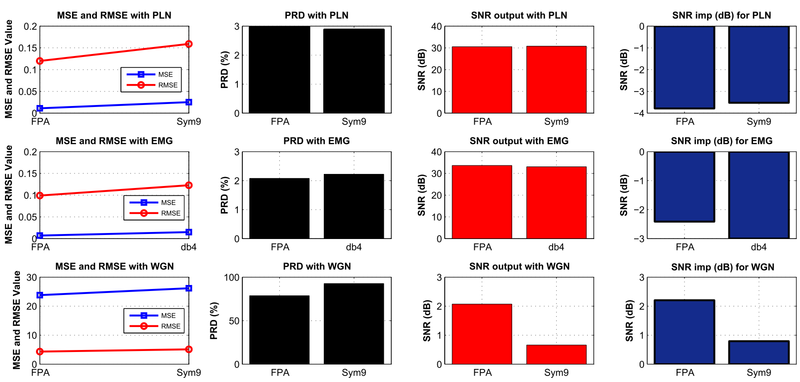


FIGURE 17. Comparative analysis between FPA-WT, Sym9 and db4.

SNR_{out}, SNR_{imp}, and PRD. For PLN, FPA-WT outperforms the Al-Qazzaz method [33] in terms of MSE (0.0144) and RMSE (0.1200). Meanwhile, the SNR_{out}, SNR_{imp}, and PRD values of these two methods are very close. In general, finding optimal parameter configurations for WT by using metaheuristic-based algorithms especially FPA, can directly improve the performance of WT in the EEG signal denoising process.

The results show that the proposed methods (βhc-WT, HS-WT, PSO-WT, GA-WT and FPA-WT) for EEG signal denoising can produce better results than manual configurations based on ad hoc strategy. Therefore, using metaheuristic

TABLE 11. Effect of selected optimal thresholding on EEG signal denoising with different noises.

EEG Signal	Method	(ϕ)	L	ρ	β	λ	MSE	RMSE	SNR_{Out}	SNR_{imp}	PRD
WGN	Optimal (FPA-WT)	db35	5	sln	Soft	heursure	24.7403	4.349	2.2045	2.0730	78.7682
WGN	Random selection	db35	5	mln	Hard	Minimax	30.5523	5.527	(-)0.0082	(-)20.8596	100.094
EMG	Optimal (FPA-WT)	db1	5	one	Hard	rigrsure	0.0098	0.099	33.6418	(-)2.4149	2.0793
EMG	Random selection	db1	5	one	Soft	Sqtwolog	5.7428	2.396	7.2532	(-)8.6053	43.3853
PLN	Optimal (FPA-WT)	db27	5	one	Hard	heursure	0.0144	0.120	30.5449	(-)3.7858	2.9700
PLN	Random selection	db27	5	sln	Soft	Minimax	5.9778	2.444	7.0769	(-)28.4984	44.2748

Bold value indicates best results where for SNR, SNRimp, highest is best and for MSE, RMSE, and PRD, lowest is best

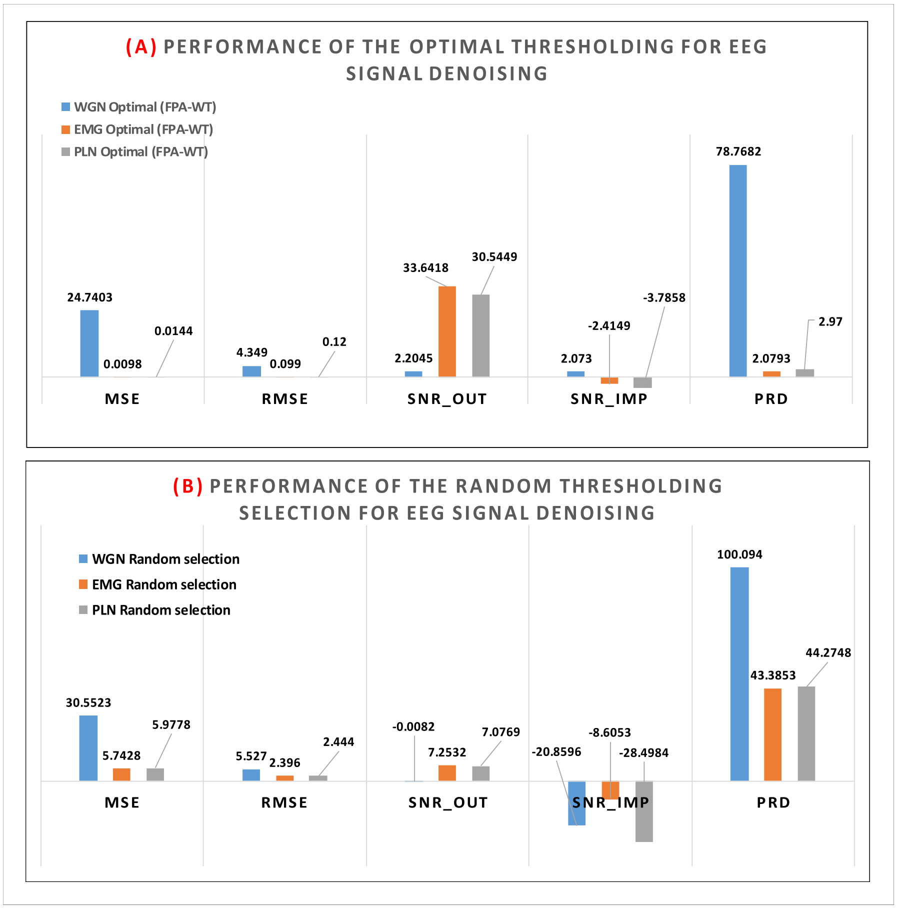


FIGURE 18. Comparative analysis between Optimal (FPA-WT) and random thresholding parameters selection.

approaches to optimize the parameters for EEG signals positively affects the denoising process performance of the WT method.

In real case, we require to apply the proposed algorithm in real-world applications. Therefore, we must first use some filters (preprocessing phase) to remove some famous noises that often corrupted the original EEG signal during the recording time such as high-pass, low-pass, band-pass, and notch filters. Then the literature normally used the wavelet transform (WT) method to denoise the EEG signal and extract the features from the denoised EEG signal based on its sub-band frequency such as Delta, Theta, Beta, Alpha, and Gamma. However, in the literature [1], [2], [33] they used WT to denoise the input EEG signal with applying the specific WT parameters such as mother wavelet function (db4 and sym9). Therefore, some important features from the input EEG signals will be lost during the WT denoising process. Consequently, the proposed FPA-WT method try to achieve better results compared with other literatures of [1], [33] according to five criteria which are (MSE, RMSE, SNR_out, SNR_improvement, and PRD) please see tables (7 and 8). In FPA-WT, although the results of the five measurements criteria is better, the important features of EEG signals are preserved.

D. THE EFFECT OF SELECTED OPTIMAL THRESHOLDING ON EEG SIGNAL DENOISING

As mentioned in Section III-A, the WT has the following five parameters. The first parameter (Φ) selects the denoising method, and the second parameter ($\boldsymbol{L}$) determines the number of levels of EEG input signal that will be decomposed. The three remaining parameters, namely,

β, λ and ρ, address the thresholding. Therefore, the optimal threshold value is selected through the proposed method (FPA-WT) after searching in the different value ranges. The selection of the optimal thresholding values is determined using the objective function (*min(MSE)*). This process is performed by compromising the EEG signals before and after denoising to confirm that the performance of the proposed method (FPA-WT) expectedly achieves the optimal tuning for the thresholding of WT parameters compared with random selection for these parameters. Table (11) summarizes the results of the proposed method (FPA-WT) in comparison with random selection for the thresholding parameters with different EEG noises. The results show the significance of using the FPA-WT method in comparison with random selection. The proposed method can obtain optimal results for all the EEG noises in accordance with MSE, RMSE, SNR_Out, SNR_imp, and PRD measures. Figure 18 shows comparative analysis between Optimal (FPA-WT) and random thresholding parameters selection.

VII. CONCLUSION AND FUTURE WORK

This paper proposes several variations of wavelet transform (WT) method for EEG signal denoising based on several meta-heuristic algorithms, including FPA-WT, GA-WT, HSA-WT, PSO-WT, and βHC-WT. As previously mentioned, the denoising performance of WT depends on its five main parameters, with each parameter having different types. Selecting the suitable WT parameters is a challenging task that is usually performed based on empirical evidence or experience. The proposed meta-heuristic algorithms aim to find the optimal WT parameters that can obtain the minimum MSE between the original and denoised EEG signals.

The proposed WT-based metaheuristic methods are evaluated using two standard EEG datasets, namely Kiern EEG dataset and the EEG Motor Movement-Imagery dataset. These dataset contain 7 and 109 volunteers respectively, and capture EEG signals from 6 and 64 EEG channels based on different mental tasks. These EEG signals are corrupted using three different noises namely, PLN, EMG, and WGN [9], [31], [32]. Five evaluation criteria are used, namely, SNR, SNR improvement, MSE, RMSE, and PRD. Several experiments are conducted to compare the performance of the proposed WT-based metaheuristic methods and to determine which of these methods can support WT in producing efficient EEG signal denoising outcomes. Interestingly, FPA-WT outperforms the other proposed methods (i.e., FPA-WT, GA-WT, HSA-WT, PSO-WT, and βHC-WT) in almost all datasets with different noise types (i.e, PLN, EMG, and WGN) based on the five measurement criteria (i.e., MSE, RMSE, SNR, SNR_{imp}, and PRD). For further validation, two well-established WT methods with the best WT parameter configurations are used for comparative evaluation. Again, FPA-WT outperforms these methods in almost all datasets with different noise types based on the five measurement criteria.

WT demonstrates many advantages and has been successfully used for denoising the non-stationary signals, such as ECG and EEG [11], [13]; however, most of the current proposed methods degrade the energy of the original signal when reducing its noise. This situation typically occurs because these approaches consider only the MSE between the original and the denoised signals. Thus, an optimum set of parameters in terms of the WT for EEG signal denoising as a multi-objective optimization task will be considered in the future. The multi-objective framework shall be applied with two objective functions, namely, min(MSE) and max(SNR), to achieve minimum noise and maintain the EEG signal energy of the maximum SNR.

Another limitation of the current version of the proposed method is time complexity. The proposed method (i.e. FPA-WT) takes 10 min to obtain the optimum WT parameters. Although the current version is not proposed for real-time applications, this problem can be overcome by reducing the search space of the WT parameters and modifying the (FPA-WT) algorithm for improved efficiency in the future. In order to apply the FPA-WT for real applications, the band-pass and notch filter in the preprocessing phase can be initially trigged. Thereafter, the output of this phase is used as input for FPA-WT to denoise the EEG signal and to extract the most important features. These extracted features can be useful to manipulate several applications such as medical applications or person identification-based EEG.

REFERENCES

[1] P. Kumari and A. Vaish, "Brainwave based user identification system: A pilot study in robotics environment," *Robot. Auton. Syst.*, vol. 65, pp. 15–23, Mar. 2015.

[2] P. Kumari Sharma and A. Vaish, "Individual identification based on neuro-signal using motor movement and imaginary cognitive process," *Optik*, vol. 127, no. 4, pp. 2143–2148, Feb. 2016.

[3] R. A. Ramadan and A. V. Vasilakos, "Brain computer interface: Control signals review," *Neurocomputing*, vol. 223, pp. 26–44, Feb. 2017.

[4] R. P. Rao, *Brain-Computer Interfacing: An Introduction*. Cambridge, U.K.: Cambridge Univ. Press, 2013.

[5] H. Berger, "Uber das elektrenkephalogramm des menschen," *Eur. Arch. Psychiatry Clin. Neurosci.*, vol. 87, no. 1, pp. 527–570, 1929.

[6] S. N. Abdulkader, A. Atia, and M.-S. M. Mostafa, "Brain computer interfacing: Applications and challenges," *Egyptian Inform. J.*, vol. 16, pp. 213–230, Jul. 2015.

[7] B. Yang, T. Zhang, Y. Zhang, W. Liu, J. Wang, and K. Duan, "Removal of electrooculogram artifacts from electroencephalogram using canonical correlation analysis with ensemble empirical mode decomposition," *Cogn. Comput.*, vol. 9, no. 5, pp. 626–633, Oct. 2017.

[8] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, "A wavelet-chaos methodology for analysis of EEGs and EEG subbands to detect seizure and epilepsy," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 2, pp. 205–211, Feb. 2007.

[9] E.-S.-A. El-Dahshan, "Genetic algorithm and wavelet hybrid scheme for ECG signal denoising," *Telecommun. Syst.*, vol. 46, no. 3, pp. 209–215, Mar. 2011.

[10] Z. A. A. Alyasseri, A. T. Khader, M. A. Al-Betar, and L. M. Abualigah. "ECG signal denoising using β-hill climbing algorithm and wavelet transform," in *Proc. 8th Int. Conf. Inf. Technol. (ICIT)*, 2017, pp. 1–7.

[11] Z. A. A. Alyasseri, A. T. Khader, and M. A. Al-Betar. "Optimal electroencephalogram signals denoising using hybrid β-hill climbing algorithm and wavelet transform," in *Proc. Int. Conf. Imag., Signal Process. Commun.*, 2017, pp. 106–112.

[12] Z. A. A. Alyasseri, A. T. Khader, and M. A. Al-Betar, "Electroencephalogram signals denoising using various mother wavelet functions: A comparative analysis," in *Proc. Int. Conf. Imag., Signal Process. Commun.*, 2017, pp. 100–105.
[13] Z. A. A. Alyasseri, A. T. Khader, M. A. Al-Betar, and M. A. Awadallah, "Hybridizing β-hill climbing with wavelet transform for denoising ECG signals," *Inf. Sci.*, vol. 429, pp. 229–246, Mar. 2018.
[14] M. A. Al-Betar, "β-Hill climbing: An exploratory local search," *Neural Comput. Appl.*, vol. 28, no. S1, pp. 153–168, Dec. 2017.
[15] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, Jun. 2000. [Online]. Available: http://circ.ahajournals.org/content/101/23/e215.full, doi: 10.1161/01.CIR.101.23.e215.
[16] J. H. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. Cambridge, MA, USA: MIT Press, 1992.
[17] Z. Woo Geem, J. Hoon Kim, and G. Loganathan, "A new heuristic optimization algorithm: Harmony search," *Simulation*, vol. 76, no. 2, pp. 60–68, Feb. 2001.
[18] J. R. Koza, *Genetic Programming as a Means for Programming Computers by Natural Selection*, vol. 1. Cambridge, MA, USA: MIT Press, 1992.
[19] J. Kennedy, "Particle swarm optimization," in *Encyclopedia of the Sciences of Learning*. Heidelberg, Germany: Springer, 2011, pp. 760–766.
[20] D. Karaboga, "An idea based on honey bee swarm for numerical optimization," Comput. Eng. Dept., Erciyes Univ., Kayseri, Turkey, Tech. Rep.-tr06, 2005.
[21] X.-S. Yang, "Flower pollination algorithm for global optimization," in *Proc. Int. Conf. Unconventional Comput. Natural Comput.* Cham, Switzerland: Springer, 2012, pp. 240–249.
[22] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.
[23] F. Glover, "Future paths for integer programming and links to artificial intelligence," *Comput. Oper. Res.*, vol. 13, no. 5, pp. 533–549, Jan. 1986.
[24] T. A. Feo and M. G. Resende, "Greedy randomized adaptive search procedures," *J. Global Optim.*, vol. 6, no. 2, pp. 109–133, 1995.
[25] P. Hansen and N. Mladenović, "Variable neighborhood search: Principles and applications," *Eur. J. Oper. Res.*, vol. 130, no. 3, pp. 449–467, May 2001.
[26] F. Glover and G. Kochenberger, "Iterated local search," in *Handbook of Metaheuristics*. Norwell, MA, USA: Kluwer, 2002.
[27] Z. A. A. Alyasseri, I. Venkat, M. A. Al-Betar, and A. T. Khader, "Edge preserving image enhancement via harmony search algorithm," in *Proc. 4th Conf. Data Mining Optim. (DMO)*, 2012, pp. 47–52.
[28] M. A. Al-Betar, Z. A. A. Alyasseri, A. T. Khader, A. L. Bolaji, and M. A. Awadallah, "Gray image enhancement using harmony search," *Int. J. Comput. Intell. Syst.*, vol. 9, no. 5, pp. 932–944, Sep. 2016.
[29] D. Rodrigues, G. F. Silva, J. P. Papa, A. N. Marana, and X.-S. Yang, "EEG-based person identification through binary flower pollination algorithm," *Expert Syst. Appl.*, vol. 62, pp. 81–90, Nov. 2016.
[30] Z. Keirn and J. Aunon, "A new mode of communication between man and his surroundings," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 12, pp. 1209–1214, Dec. 1990.
[31] J. Wang, Y. Ye, X. Pan, and X. Gao, "Parallel-type fractional zero-phase filtering for ECG signal denoising," *Biomed. Signal Process. Control*, vol. 18, pp. 36–41, Apr. 2015.
[32] W. Jenkal, R. Latif, A. Toumanari, A. Dliou, O. El B'Charri, and F. M. Maoulainine, "An efficient algorithm of ECG signal denoising using the adaptive dual threshold filter and the discrete wavelet transform," *Biocybern. Biomed. Eng.*, vol. 36, no. 3, pp. 499–508, 2016.
[33] N. Al-Qazzaz, S. Hamid Bin Mohd Ali, S. Ahmad, M. Islam, and J. Escudero, "Selection of mother wavelet functions for multi-channel EEG signal analysis during a working memory task," *Sensors*, vol. 15, no. 11, pp. 29015–29035, Nov. 2015.
[34] N. K. Al-Qazzaz, S. Ali, S. A. Ahmad, M. S. Islam, and M. I. Ariff, "Selection of mother wavelets thresholding methods in denoising multi-channel EEG signals during working memory task," in *Proc. IEEE Conf. Biomed. Eng. Sci. (IECBES)*, Dec. 2014, pp. 214–219.
[35] N. K. Al-Qazzaz, S. H. B. M. Ali, S. A. Ahmad, M. S. Islam, and J. Escudero, "Discrimination of stroke-related mild cognitive impairment and vascular dementia using EEG signal analysis," *Med. Biol. Eng. Comput.*, vol. 56, no. 1, pp. 137–157, Jan. 2018.
[36] C. S. P. Reddy, "Analysis of EEG signal for the detection of brain abnormalities," *Int. J. Res.*, vol. 4, no. 17, pp. 1947–1950, 2017.
[37] P. Kumari and A. Vaish, "Feature-level fusion of mental task's brain signal for an efficient identification system," *Neural Comput. Appl.*, vol. 27, no. 3, pp. 659–669, Apr. 2016.
[38] N. Padmaja, M. Bharathi, and E. Sujatha, "A GUI based EEG signal denoising using hilbert huang transform," *J. Electron. Eng.*, vol. 7, no. 1, pp. 25–30, 2016.
[39] A. Torabi, F. Z. Jahromy, and M. R. Daliri, "Semantic category-based classification using nonlinear features and wavelet coefficients of brain signals," *Cogn. Comput.*, vol. 9, no. 5, pp. 702–711, Oct. 2017.
[40] S. Poornachandra and N. Kumaravel, "Hyper-trim shrinkage for denoising of ECG signal," *Digit. Signal Process.*, vol. 15, no. 3, pp. 317–327, May 2005.
[41] H. Kumar, S. P. Pai, G. Vijay, and R. Rao, "Wavelet transform for bearing condition monitoring and fault diagnosis: A review," *Int. J. COMADEM*, vol. 17, no. 1, pp. 9–23, 2014.
[42] M. Mamun, M. Al-Kadi, and M. Marufuzzaman, "Effectiveness of wavelet denoising on electroencephalogram signals," *J. Appl. Res. Technol.*, vol. 11, no. 1, pp. 156–160, Feb. 2013.
[43] M. Al-Kadi, M. Reaz, M. Ali, and C. Liu, "Reduction of the dimensionality of the EEG channels during scoliosis correction surgeries using a wavelet decomposition technique," *Sensors*, vol. 14, no. 7, pp. 13046–13069, Jul. 2014.
[44] R. Yang and M. Ren, "RETRACTED: Wavelet denoising using principal component analysis," *Expert Syst. Appl.*, vol. 38, no. 1, pp. 1073–1076, Jan. 2011.
[45] Z. A. A. Alyasseri, A. T. Khader, M. A. Al-Betar, J. P. Papa, O. A. Alomari, and S. N. Makhadme, "An efficient optimization technique of eeg decomposition for user authentication system," in *Proc. 2nd Int. Conf. BioSignal Anal., Process. Syst. (ICBAPS)*, Jul. 2018, pp. 1–6.
[46] C. Sawant and H. T. Patii, "Wavelet based ECG signal de-noising," in *Proc. 1st Int. Conf. Netw. Soft Comput. (ICNSC)*, 2014, pp. 20–24.
[47] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, Sep. 1994.
[48] B. N. Singh and A. K. Tiwari, "Optimal selection of wavelet basis function applied to ECG signal denoising," *Digit. Signal Process.*, vol. 16, no. 3, pp. 275–287, May 2006.
[49] M. R. Mowla, S.-C. Ng, M. S. Zilany, and R. Paramesran, "Artifacts-matched blind source separation and wavelet transform for multichannel EEG denoising," *Biomed. Signal Process. Control*, vol. 22, pp. 111–118, Sep. 2015.
[50] Z. A. A. Alyasseri, A. T. Khader, M. A. Al-Betar, J. P. Papa, and O. A. Alomari, "EEG feature extraction for person identification using wavelet decomposition and multi-objective flower pollination algorithm," *IEEE Access*, vol. 6, pp. 76007–76024, 2018.
[51] P. Mcsharry, G. Clifford, L. Tarassenko, and L. Smith, "A dynamical model for generating synthetic electrocardiogram signals," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 3, pp. 289–294, Mar. 2003.
[52] J. Feng, Z. Wang, and M. Zeng, "Distributed weighted robust Kalman filter fusion for uncertain systems with autocorrelated and cross-correlated noises," *Inf. Fusion*, vol. 14, no. 1, pp. 78–86, Jan. 2013.
[53] X.-J. Sun, Y. Gao, Z.-L. Deng, C. Li, and J.-W. Wang, "Multi-model information fusion Kalman filtering and white noise deconvolution," *Inf. Fusion*, vol. 11, no. 2, pp. 163–173, Apr. 2010.
[54] M. Chakraborty and S. Das, "Determination of signal to noise ratio of electrocardiograms filtered by band pass and savitzky-golay filters," *Procedia Technol.*, vol. 4, pp. 830–833, 2012.
[55] M. Alfaouri and K. Daqrouq, "ECG signal denoising by wavelet transform thresholding," *Amer. J. Appl. Sci.*, vol. 5, no. 3, pp. 276–281, Mar. 2008.
[56] M. Üstündağ, M. Gökbulut, A. Şengür, and F. Ata, "Denoising of weak ECG signals by using wavelet analysis and fuzzy thresholding," *Netw. Model. Anal. Health Inform. Bioinf.*, vol. 1, no. 4, pp. 135–140, Dec. 2012.
[57] K. Zeng and M. Dong, "A novel cuboid method with particle swarm optimization for real-life noise attenuation from heart sound signals," *Expert Syst. Appl.*, vol. 41, no. 15, pp. 6839–6847, Nov. 2014.
[58] M. Cyriac and P. Sankar, "Denoising of ECG signals using the framelet transform," *Int. J. Comput. Appl.*, vol. 108, no. 7, pp. 24–29, Dec. 2014.
[59] M. Lagha, M. Tikhemirine, S. Bergheul, T. Rezoug, and M. Bettayeb, "De-noised estimation of the weather Doppler spectrum by the wavelet method," *Digit. Signal Process.*, vol. 23, no. 1, pp. 322–328, Jan. 2013.
[60] S. Poornachandra,"Wavelet-based denoising using subband dependent threshold for ECG signals," *Digit. Signal Process.*, vol. 18, no. 1, pp. 49–55, Jan. 2008.

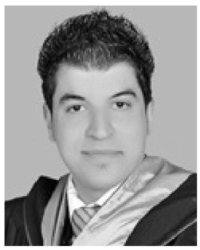
[61] T. Gandhi, B. K. Panigrahi, and S. Anand, "A comparative study of wavelet families for EEG signal classification," *Neurocomputing*, vol. 74, no. 17, pp. 3051–3057, Oct. 2011.
[62] D. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.
[63] A. L. Bolaji, M. A. Al-Betar, M. A. Awadallah, A. T. Khader, and L. M. Abualigah, "A comprehensive review: Krill Herd algorithm (KH) and its applications," *Appl. Soft Comput.*, vol. 49, pp. 437–446, Dec. 2016.
[64] R. Poli, J. Kennedy, and T. Blackwell, "Particle swarm optimization," *Swarm Intell.*, vol. 1, no. 1, pp. 33–57, Jun. 2007.
[65] Z. A. A. Alyasseri, A. T. Khader, M. A. Al-Betar, M. A. Awadallah, and A. Xin-She, "Variants of the flower pollination algorithm: A review," in *Nature-Inspired Algorithms and Applied Optimization*. Cham, Switzerland: Springer, 2018, pp. 91–118.
[66] M. A. Kabir and C. Shahnaz, "Denoising of ECG signals based on noise reduction algorithms in EMD and wavelet domains," *Biomed. Signal Process. Control*, vol. 7, no. 5, pp. 481–489, Sep. 2012.

ZAID ABDI ALKAREEM ALYASSERI received the B.Sc. degree in computer science from Babylon University, in 2007, and the M.Sc. degree in computer science from University Science Malaysia (USM), in 2013, where he is currently pursuing the Ph.D. degree in the field of artificial intelligence (brain-inspired computing) with the Computational Intelligence Research Group, School of Computer Sciences. He is also a Senior Lecturer with the University of Kufa, Iraq. His research interests are optimization, pattern recognition, EEG, brain–computer interface, signal and image processing, machine learning, and deep learning.

AHAMAD TAJUDIN KHADER received the B.Sc. and M.Sc. degrees in mathematics from the University of Ohio, USA, in 1982 and 1983, respectively, and the Ph.D. degree in computer science from the University of Strathclyde, U.K., in 1993. He is currently a Professor with the School of Computer Sciences, Universiti Sains Malaysia, where he is also the Dean. His research interest mainly focuses on optimization and scheduling.

MOHAMMED AZMI AL-BETAR received the B.Sc. and M.Sc. degrees from the Computer Science Department, Yarmouk University, Jordan, in 2001 and 2003, respectively, and the Ph.D. degree from the School of Computer Sciences, USM, in 2010. He is currently an Associate Professor with the Department of Computer Science and the Deputy Dean of scientific research with the Al-Huson University College, Al-Balqa Applied University, Jordan. His research interests are mainly directed to metaheuristic optimization methods and hard combinatorial optimization problems.

AMMAR KAMAL ABASI received the B.Sc. degree in computer information system from the Jordan University of Science and Technology and the M.Sc. degree in the international business from the University of Jordan. He is currently pursuing the Ph.D. degree with the School of Computer Sciences, Universiti Sains Malaysia. His research interests include evolutionary algorithms, nature-inspired computation, and their applications to optimization problems.



SHARIF NASER MAKHADMEH received the M.Sc. degree in computer science from Universiti Utara Malaysia (UUM), Malaysia, in 2015. He is currently pursuing the Ph.D. degree with the School of Computer Sciences, Universiti Sains Malaysia (USM). His research interests include optimization algorithms, artificial intelligence, power scheduling problem, and smart home.