IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 30, 2018, accepted October 28, 2018, date of publication November 19, 2018,
date of current version December 27, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2881470

EEG Feature Extraction for Person Identification Using Wavelet Decomposition and Multi-Objective Flower Pollination Algorithm

ZAID ABDI ALKAREEM ALYASSERI[1,2], (Member, IEEE), AHAMAD TAJUDIN KHADER[1], MOHAMMED AZMI AL-BETAR[3], JOÃO P. PAPA[4], (Senior Member, IEEE), AND OSAMA AHMAD ALOMARI[1]

[1]School of Computer Sciences, Universiti Sains Malaysia, Pulau Pinang 11800, Malaysia
[2]ECE Department-Faculty of Engineering, University of Kufa, Najaf, Iraq
[3]Department of Information Technology, Al-Huson University College, Al-Balqa Applied University, Irbid, Jordan
[4]Department of Computing, São Paulo State University, Bauru 13084-971, Brazil

Corresponding author: Zaid Abdi Alkareem Alyasseri (zaid.alyasseri@uokufa.edu.iq)

This work was supported by USM under Grant 1001/PKOMP/ 8014016. The work of Z. A. A. Alyasseri was supported in part by The World Academic Science (TWAS) and in part by the University Science Malaysia (USM) (TWAS-USM Postgraduate Fellowship 2015), under Grant 3240287134. The work of J. P. Papa was supported in part by FAPESP under Grants 2013/07375-0, 2014/12236-1, and 2016/19403-6, and in part by CNPq under grant 307066/2017-7.

ABSTRACT In the modern life, the authentication technique for any system is considered as one of the most important and challenging tasks. Therefore, many researchers have developed traditional authentication systems to deal with our digital society. Recently, several studies showed that the brain electrical activity or electroencephalogram (EEG) signals could provide robust and unique features that can be considered as a new biometric authentication technique, given that accurate methods to decompose the signals must also be considered. This paper proposes a novel method for extracting EEG features using multi-objective flower pollination algorithm and the wavelet transform. The proposed method was applied in two scenarios for EEG signal decomposition to extract unique features from the original signals. Moreover, the proposed method is compared with the state-of-the-art techniques using different criteria with promising results.

INDEX TERMS Biometric authentication, EEG, wavelet decomposition, feature extraction, flower pollination algorithm, multi-objective.

I. INTRODUCTION

Electroencephalogram (EEG) is a graphical recording of brain electrical activity that is captured from the scalp. This recording represents the voltage fluctuations resulting from ionic current flows within the neurons of the brain [30], [48]. Therefore, EEG signals can provide most of the required information about brain activity, and they are captured using *invasive* or *non-invasive* techniques [40]. 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 group-Brain Computer Interface *(ECOG-BCI)* for arm movement control [41]. Meanwhile, there are several techniques to record the brain activity other than EEG, such as magnetoencephalography for magnetic field fluctuations caused by electrical activity in the brain, and functional magnetic resonance imaging and functional near-infrared imaging for changes in the blood oxygenation level resulting from neural activity [41].

In [17], Berger introduced 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 of patients. Such 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, as reported in [1]. Recently, EEG signals have been

successfully used as a new biometric technique in security and authentication applications [1], [29], [30].

Ferdous [23] proposed a biometric method using the power spectral density estimates of EEG signals over the combined Alpha-Beta rhythm. The authors tested the proposed method using ANT Neuro device for capturing the EEG signals from two people, with the highest accuracy rate around 50%. Abo-Zahhad *et al.* [2] proposed to combine EEG and eye blinking electrooculography (EOG) to improve the performance of the brain signal for biometric authentication. The proposed method extracted several EEG features to achieve the best results such as autoregressive, canonical correlation, and EEG density. A Linear discriminant analysis (LDA) classifier has been applied to classify the correct classification rate, and the signals collected from the subjects were based on three different tasks. The proposed approach showed a significant improvement in all tasks.

Safont *et al.* [44] proposed a biometric authentication method using EEG signals using three EEG channels to capture a self-collected dataset from 70 subjects. The authors tested the proposed system using six classification methods, achieving accuracies nearly to 93.8%. Kumar and Vaish [30] proposed a user identification system based on EEG signals collected from six people using a headset with 14 channels. In the preprocessing phase, they employed a Butterworth 5th-order filter with a range as of $6-35$ Hz to achieve the highest signal-to-noise ratio (SNR) concerning the input EEG signal. In the feature extraction phase, a Wavelet Transform (WT) was used to perform feature extraction. Besides, three standard statistical measurements were computed and used as features for learning vector quantization-neural networks (LVQ-NN) classifier. Finally, the recognition rate has been calculated over the different scenarios to find the best combination of channels that can provide the highest accuracies.

Sharma and Vaish [48] investigated some cognitive tasks for building an individual identification system. They used standard EEG signals based on motor/movement and imaginary tasks [47]. Additionally, they also considered WT to decompose the signal into five levels to extract four different features from each sub-band: energy, the logarithm of the energy, absolute energy, and resting energy expenditure (REE) energy. For classification purposes, a neural network classifier was evaluated in four different train-test scenarios. The authors found that the highest identification rates can be obtained using the cognitive tasks based on motor imagination when compared with the results based on motor movement.

Jayarathne *et al.* [26] proposed a novel approach for EEG-based biometric authentication. The proposed method used a headset with 14 channels to collect the signals. The Common Spatial Patterns (CSP) and Linear discriminant analysis (LDA) have been used for feature extraction and classification phase, respectively. The proposed method was tested over 12 subjects with the highest accuracy rate as of 96.97%. Finally, the authors suggested using the EEG signal rather than the personal identification number (PIN) for authenticating a person while using an automated teller machine (ATM).

Ruiz-Blodent *et al.* [43] proposed a new cognitive event-related biometric recognition protocol called CEREBRE, which is designed to elicit individual responses from multiple functional brain systems (e.g., the primary visual, facial recognition, and gustatory/appetitive systems). The authors tested the proposed protocol using 50 users, where each one is asked to respond with a button press when they detect a color stimulus. The proposed approach obtained 100% of identification accuracy.

Kumar and Vaish [31] proposed a novel method for feature extraction from EEG signals based on Canonical Correlation Analysis. The proposed approach was tested on a standard EEG dataset [27] with five mental tasks, including a baseline, multiplication of two numbers, geometric figure rotation, letter composing, and visual counting. Additionally, each task was repeated several times for ten seconds, and the EEG signals were collected from seven subjects. The proposed method used three techniques for feature extraction, i.e., empirical mode decomposition, information theoretic measures, and statistical measurements. The authors employed a Linear Vector Quantization (LVQ) neural network classifier and an extension (LVQ2) for classification purposes.

Recently, several works suggested to using optimization methods for solving problems inherent to non-stationary signals [8]–[11], [13], [14], [42]. Rodrigues *et al.* [42] proposed the Binary Flower Pollination Algorithm (BFPA) to select the EEG channels that can provide the highest recognition rate for person identification. Their work was tested using a standard EEG dataset composed of motor/movement and imaginary tasks [47]. The proposed approach obtained recognition rates close to 86% using the Optimum-Path Forest classifier (OPF) [35], [36], while reducing the number of EEG channels to nearly half. Alyasseri *et al.* [10] proposed a hybrid approach for electrocardiogram (ECG) signal denoising that based on β-hill climbing [4] to obtain the suitable parameters for WT-based denoising. Later, the same group of authors applied the same approach in the context of EEG signal denoising [8], [9], [11]. The proposed method also has successfully used for ECG classification using several classification methods such as Neural Networks, naive Bayes, and Decision Trees.

Roughly speaking, a standard WT-based signal denoising approach comprises five parameters, which are crucial to the efficiency of the denoising process. Also, the selection of these parameters is usually performed according to experience or empirical evidence. In previous studies, the WT parameter configuration was formulated as an optimization problem using the minimum squared error (MSE) as the objective function [9], [11]. Despite the advantages of signal denoising using the Wavelet Transform and optimization techniques, most of the current approaches degrade the energy of the original signal when reducing its noise. This situation usually occurs since they consider only the

minimum squared error between the original and denoised signals.

For that reason, this work designs a multi-objective function that considers a balance between reducing the EEG noise and, at the same time, keeping the signal's energy. We proposed an extension of the work by Alyasseri *et al.* [13], which employed the Flower Pollination Algorithm (FPA) [55] to fine-tune WT parameters while denoising EEG signals using multi-objective criteria. In this paper, we deepened the wavelet decomposition process to 10 levels, as well as we proposed to extract a broader range of features, thus outperforming our previous results. Another main contribution of this work is to present a more detailed discussion and experimental section concerning the problem of WT-based denoising of EEG signals applied to biometric identification.

The remainder of this paper is organized as follows. Section II presents the theoretical background about EEG signal denoising using the Wavelet Transform. Section III describes the Flower Pollination Algorithm and its multi-objective variant, and Section IV presents the approach proposed in this work. The results and discussion are described in Section V, and conclusions are stated in Section VI.

II. WAVELET TRANSFORM

Wavelet Transform 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, as well as to address several problems, such as those related to signal compression, feature selection, and signal denoising [7], [10], [28], [37]. Recently, WT has been extensively tailored for non-stationary signals due to its robust performance in removing several EEG artifact noises that can corrupt the original signal during its recording time. These noises include eye blinking, eye movements, muscle activity noise, electromyogram (EMG) noise, and interference of electronic device signals [6], [19], [32].

A. WAVELET DENOISING PRINCIPLE FOR NON-STATIONARY SIGNALS

The Wavelet Transform is a powerful tool for time-frequency domain representation. This technique represents the signal by the correlation between the translation and the dilation of a mother wavelet function (MWF) [22], [37], [54]. In general, the problems solved by a WT can be categorized into two 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 [21] as the so-called Donoho's approach. In general, DWT decomposes a signal by using a set of low and high pass filters to produce the approximation and details coefficients, respectively.

The main objective of using DWT for denoising is to decompose the input signal via different coefficient levels to manipulate the high frequency of the input signals [51]. In other words, DWT decomposes the EEG signal into several frequency bands because it assumed that the artifacts would have large amplitudes in the respective frequency bands. Figure 1 shows the wavelet denoising procedure with decomposition level $L = 3$. Normally, the denoising process involves three phases:

- Decomposition phase: we assume the original EEG signal with n samples $s(t) = [s^1(t), s^2(t), \ldots, s^n(t)]$ will be divided into three levels (just for the sake of explanation), and each level will be decomposed into two parts, namely, approximation coefficients (cA) and detail coefficients (cD). In this case, cD will be processed using a high-pass filter, while cA will continue to be decomposed for the next level, as follows:

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

  where $cA_i(t)$ and $cD_i(t)$ denote the approximation and detail coefficients of level i, and ϕ and Ψ refer to the scaling and shifting parameters, respectively.
- Thresholding phase: a threshold value is defined for each level according to the noise level of the coefficient.
- Reconstruction phase: the denoised signal $\hat{x}$ is reconstructed using the iDWT, as follows [34]:

$$\hat{x}(t) = \sum_{k=-\infty}^{\infty} cA_L(k)\phi'_i(t-k) + \sum_{i=1}^{L}\sum_{k=-\infty}^{\infty} cD_{i+1}(k)\Psi'_i(t-k). \quad (2)$$

The WT has five parameters, each one having a different meaning (Table 1). Therefore, the success of EEG signal denoising relies on the selection of those parameters properly. 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 function (i.e., β)), can be divided into hard and soft thresholding [20], [21]. Figure 2 shows the difference between these thresholding methodologies. Such types (soft or hard) must be selected in the second phase along with the fourth parameter, namely the selection rule (λ),

TABLE 1. 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 L	5
Thresholding selection rule λ	Heursure, Rigsure, Sqtwolog, and Minimax
Re-scaling approach ρ	one, sln, mln

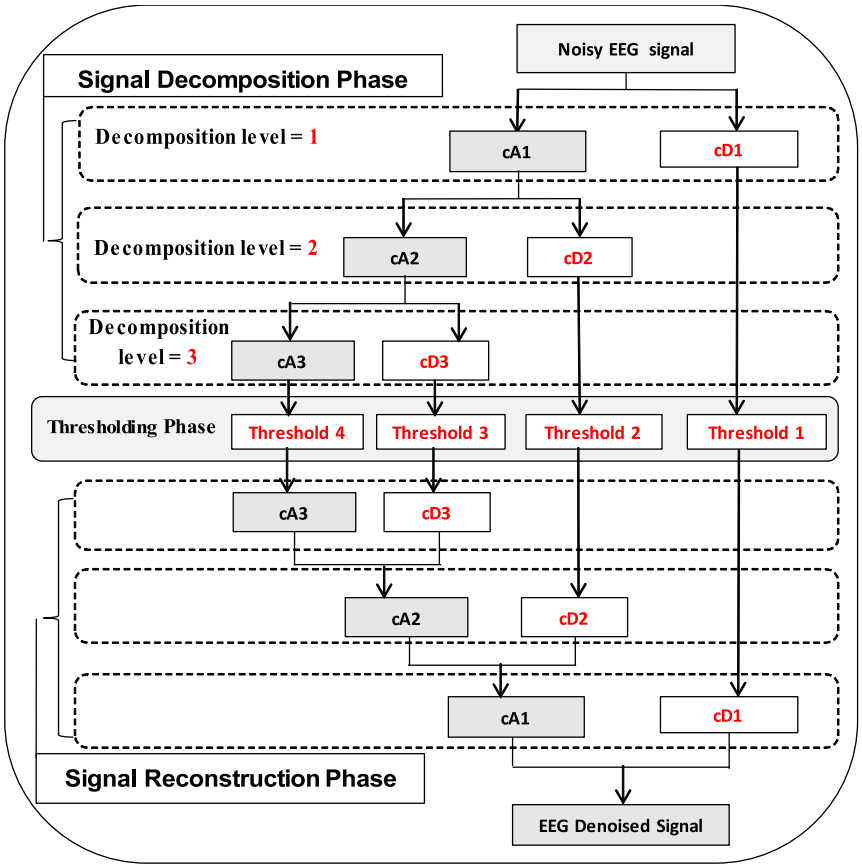


FIGURE 1. EEG denoising process employed in this work.

and the fifth parameter concerns the rescaling method (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 [22].

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 (3), as follows:

$$s_{noisy}(t) = s(t) + \sigma e(t) \tag{3}$$

where $s(t)$ stands for the original EEG signal, e denotes the noise, σ concerns the amplitude of the noise, and n the number samples. The wavelet parameters (i.e., 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 using the iDWT (Equation 2).

III. BACKGROUND

This section provides a background about the Flower Pollination Algorithm and its multi-objective technique. Section III-A introduces the standard Flower Pollination algorithm, while Section III-B explains the concepts of the multi-objective optimization.

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

A. STANDARD FLOWER POLLINATION ALGORITHM

In the recent optimization review, the meta-heuristic algorithms can be classified into: evolutionary algorithm [5], [15], swarm intelligence [12], [18], [49], and trajectory algorithms [3], [4], [10].

Flower pollination algorithm (FPA) is a successful swarm-based intelligence technique inspired from the pollination behaviour of the flowering plants. FPA was introduced by Yang and Ren [55], and successfully applied for many optimization problems [12].

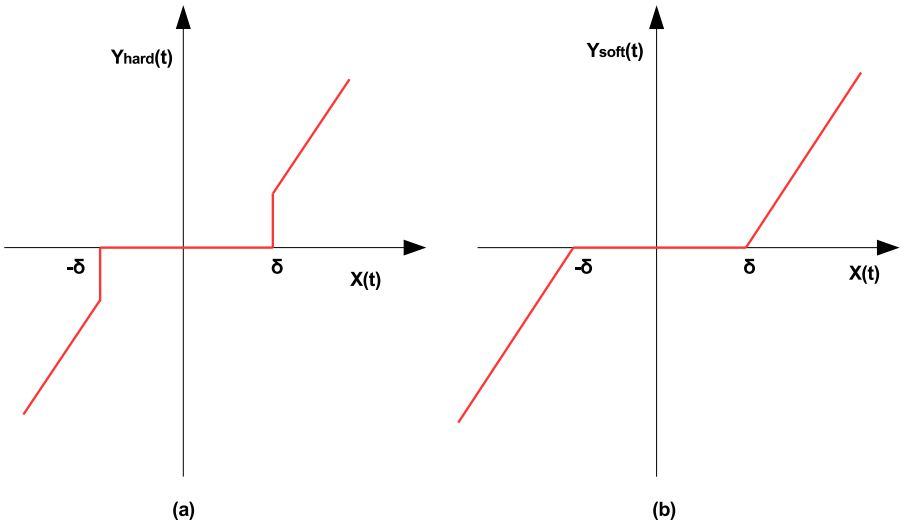


FIGURE 2. Soft and hard thresholding methods. (a) Hard threshold. (b) Soft threshold.

In general, the procedure of FPA is summarized through five steps which are shown as follows. Note that the flowchart of the FPA is displayed in Figure 3, and Algorithm 1 implements its pseudo-code.

Step 1 (Initialization of FPA and Problem Parameters): For solving any optimization problem using a global optimization technique, the first step initializes the population solution within possible range parameters value $\boldsymbol{x}$, as well as initialized problem parameters. Therefor, the formulation of the initialization can be generalized as follows:

$$\min\{f(\boldsymbol{x}) \,|\, \boldsymbol{x} \in X\},$$

where $f(\boldsymbol{x})$ is the objective function; $\boldsymbol{x} = \{x_i \,|\, i = 1, \ldots, d\}$ is the set of decision variables, and $X = \{\boldsymbol{X}_i \,|\, i = 1, \ldots, d\}$ is the possible value range for each decision variable, where $\boldsymbol{X}_i \in [LB_i, UB_i]$, where LB_i and UB_i are the lower and upper bounds for the decision variable x_i, respectively, and d is the number of decision variables.

On the other hand, the parameters of FPA should be initialized in this step as well, where these parameters can be summarized as follows:

- *Fsize*: number of flowers (population size).
- Determining g^*: selecting the best current solution from the initialized population size.
- P: switch probability, where the value of P will determine which path will follow either local or global pollination.
- L: is the strength of the pollination, which refers to a step size.

Step 2 (Initialize FPA Population Memory): The flower population memory (FPM) is a 2-dimensional matrix with size $Fsize \times d$ which contains sets of flower location vectors as many as *Fsize* (see Eq. (4)). Notice that these flowers are randomly generated as follows: $x_i^j = LB_i + (UB_i - LB_i) \times U(0, 1)$, $\forall i = 1, 2, \ldots, d$ and $\forall j = 1, 2, \ldots, Fsize$,

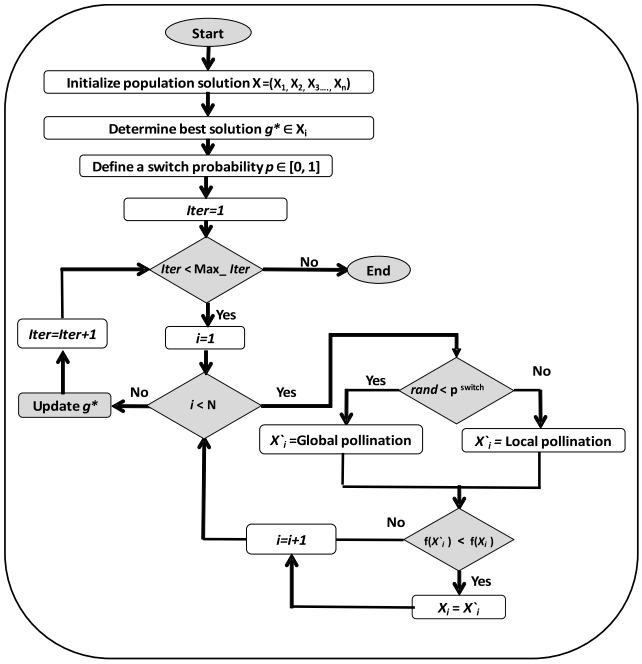


FIGURE 3. Flower pollination algorithm flowchart.

and $U(0, 1)$ generates a uniform random number between 0 and 1. The generated solutions are stored in the FPM in ascending order according to their objective function values, where $f(x^1) \leq f(x^2) \leq \ldots \leq f(x^{Fsize})$.

$$\text{FPM} = \begin{bmatrix} x_1^1 & x_2^1 & \cdots & x_d^1 \\ x_1^2 & x_2^2 & \cdots & x_d^2 \\ \vdots & \vdots & \cdots & \vdots \\ x_1^{Fsize} & x_2^{Fsize} & \cdots & x_d^{Fsize} \end{bmatrix}. \quad (4)$$

In this step, the global best flower location g^* is memorized where $g^* = x^1$.

Step 3 (Intensification of the Current Flower Population): In this step, the pollinator will fly to find the best flower based on switch probability (p) value. The p value will determine which path will follow either local or global pollination as follows:

- Global Search of FPA (biotic): in this type of pollination, the flowers pollens are transferred by pollinators such as bees, bats, and birds. to long distances. This ensures the pollination and reproduction of the most fittest. Therefore, we can represent the procedure of biotic FPA as follows:

$$x_i^{(t+1)} = x_i^{(t)} + L(g^* - x_i^{(t)}) \tag{5}$$

  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, i.e. 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 [12], [42], [55]. That is, we draw L > 0 from a Levy distribution

$$L \sim \frac{\lambda\Gamma(\lambda)sin(\pi\lambda/2)}{\pi}\frac{1}{s^{1+\lambda}}, \quad (s \gg s_0 > 0) \tag{6}$$

  Where $\Gamma(\lambda)$ denotes the standard gamma function, and this distribution is valid for large steps s > 0. We set ($\lambda = 1.5$).
- Local Search of FPA (abiotic): the pollination of this type occurs without any pollinators. That means, it occurs based on the wind and diffusion to transfer the pollen. The local pollination and flower constancy are represented as follows:

$$x_i^{(t+1)} = x_i^{(t)} + \epsilon(x_j^{(t)} - x_k^{(t)}) \tag{7}$$

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

Step 4 (Updating the Flower Population Memory): In this step, for each pollinator in FPM, the new pollinator replaces the current flower location based on the objective function value of the new solution furthermore, the global best flower location $\boldsymbol{g}^*$ is also updated.

Step 5 (Stop Criterion): In this step, the Flower Pollination Algorithm repeats step 3 to 4 until the termination criterion is met. Normally, the criteria depends on some factors, such as the quality of the final outcomes, number of generations, and the computational time constraint.

Algorithm 1 Flower Pollination Algorithm Pseudo-Code

```
Objective min f(x), x ∈ ℜ^d
Initialize a population of n flowers/pollens with random solution
Find the best solution g* in the initial population
Define a switch probability p ∈ [0, 1]
Calculate all (f(x)) for n solutions
t=0
while t ≤ MaxGeneration do
  for i = 1, .., N do
    if rnd ≤ p then
      Draw a (d-dimensional) step vector 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 ∈ [0,1]
      Randomly choose j and k among all solution
      Do local pollination via x_i^(t+1) = x_i^(t) + ∈ (x_j^(t) − x_k^(t))
    end if
    Calculate(f(x'))
    if f(x') ≤ 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
```

B. MULTI-OBJECTIVE OPTIMIZATION

This section describes a briefly introduction about multi-objective optimization techniques. In general, the multi-objective optimization refers to solve any optimization problem using more than one objective function [33], [56], [57]. The multi-objective optimization problem for m objective functions can be formulated as follows:

$$Mimimze\ F(x) = f_1(x), f_2(x), \ldots, f_m(x), \tag{8}$$

Subject to:

$$g_i(x) \leq 0, \quad i = 1, 2, \ldots, k \tag{9}$$

$$h_i(x) = 0, \quad i = 1, 2, \ldots, N \tag{10}$$

$$L_i \leq (x)_i \leq U_i, \quad i = 1, 2, \ldots, d \tag{11}$$

where m refers to number of objective functions, N refers to number of variables, K denotes number of inequality constraints, M denotes number of equality constraints.

The main difference between single and multi-objective optimization problems known that solutions with single objective function can be compared easily with one objective function. For instance, for minimization problem, a solution **A** is better than **B** if, and only if, **X1** < **X2**. Inconstant, the solution space of multi-objective optimization problem can not be compared with single objective technique because both of them using different way to evaluate the solution. In this case, the best solution of a multi-objective optimization problem will be selected if, and only if, the solution

dominates another solution for all or at least one objective function [33], [57].

C. DEFINITIONS

Concept 1 (Pareto Dominance): Suppose we have two vectors, such as: $\overrightarrow{V} = (v_1, v_2, \ldots, v_n)$ and $\overrightarrow{Z} = (z_1, z_2, \ldots, z_n)$ vector V dominates vector Z if and only if:

$$\forall i \in \{1, 2, .., n\}, [f(v_i) \leq f(z_i)] \bigwedge [\exists i \in \{1, 2, .., n\} : f(v_i)] \tag{12}$$

Concept 2 (Pareto Optimality): The solution $\overrightarrow{v} \in V$ is called Pareto Optimal if and only if

$$\nexists \overrightarrow{z} \in V \mid F(\overrightarrow{z}) < F(\overrightarrow{v}) \tag{13}$$

Concept 3 (Pareto Optimal Set): The set of all optimal pareto solution is called Pareto optimal set and defined as follows:

$$Pareto_{set} = \{v, z \in V \mid \exists F(z) < F(v)\} \tag{14}$$

The set that is containing the corresponding value of the Pareto optimal solution in Pareto optimal set is called *Pareto front*, while the Pareto front defined as follows:

Concept 4 (Pareto Front):

$$Pareto_{front} = \{F(v) \mid v \in Pareto_{set}\} \tag{15}$$

The FPA has been extended to multi-objective optimization technique by Yang *et al.* [56], while the author adapted multi-objective flower pollination algorithm (MOFPA) for solving engineering optimization problems. MOFPA was evaluated using several engineering optimization problems to produce optimal results. Also, MOFPA has been applied for several real-world problems, such as engineering optimization problems [57], radial distribution system [52], Dynamic Economic Dispatch Considering Emission [16], transmission loss and power plant emission, minimization of generating cost, and improvement of voltage stability [50], power loss reduction [39], and Power Flow Problem [38].

MOFPA is implemented according to the weighted sum approach to combine two objectives into a composite one objective function. MOFPA defined the multi-objective optimization as follows:

$$F = \sum_{k=1}^{N} W_k f_k \tag{16}$$

with

$$\sum_{k=1}^{N} W_k = 1, \quad W_k > 0, \tag{17}$$

where N refers to number of objective functions and W_k are the non-negative weights. The essential idea of the weighted sum approach is that these weighting coefficients consider the preferences for these multi-objectives. For example, given a set of weights, such as $(w_1, w_2, \ldots, w_k)$ for each weight (w_i) there is a single Pareto front point will be generated. With a sufficiently large number of different weights, the proposed method able to obtain the true Pareto front such as in Eq 15.

IV. PROPOSED METHOD: NEW EEG FEATURES EXTRACTION USING MOFPA-WT

In this section, we introduce the proposed system for EEG-based user authentication, which comprise four phases, in which the output of each stage works as an input to the consecutive one. Figure 4 depicts the proposed approach, which is detailed next.

A. EEG SIGNAL ACQUISITION

We used a standard EEG signal datasets namely the 'Motor Movement/Imagery' [25]. Section V-A will provide full details and more explanation about these datasets.

B. EEG SIGNAL DENOISING USING MOFPA-WT

Despite WT has many advantages and has been successfully used for denoising non-stationary signals such as ECG and EEG [9], [11], most of the current approaches degrade the energy of the original signal when reducing its noise. This situation usually occurs because they consider only the minimum squared error between the original and denoised signals. For that reason, this work designs a multi-objective function that considers a balance between reducing the EEG noise and keeping its signal energy.

In this paper, we propose to estimate the optimum/near-optimum set of parameters concerning the Wavelet Transform for EEG signal denoising as a multi-objective optimization task. In our approach, the set of WT parameters is represented as a vector $x = (x_1, x_2, \ldots, x_d)$ where d is the number of parameters used for the Wavelet Transform.[1] In this context, x_1 represents the value of the mother wavelet function parameter Φ, x_2 stands for the value of the decomposition level parameter L, x_3 refers to the thresholding method β, x_4 represents the value of the thresholding selection rule parameter λ, and x_5 represents the re-scaling approach ρ (the possible ranges for these parameters are described in Table 1). Figure 5 depicts the representation of a possible solution using the proposed approach.

The proposed MOFPA-WT evaluates each solution using the multi-objective framework applying two objective functions: ***min(MSE)*** and ***max(SNR)***, as formulated below:

$$f = W_1 f_1 + W_2 f_2 \tag{18}$$

$$= W_1 * min(MSE) + W_2 * max(SNR), \tag{19}$$

where the weight vector is initialized as follows:

$$W_1 \sim U(0, 1),$$
$$W_2 = 1 - W_1. \tag{20}$$

The two objective functions which are mean squared error (MSE) and signal-to-noise ratio (SNR) are formulated as below:

$$MSE = \frac{1}{N} \sum_{i=1}^{N} [x_i - \widehat{x}_i]^2 \tag{21}$$

[1] In this paper, $d = 5$.

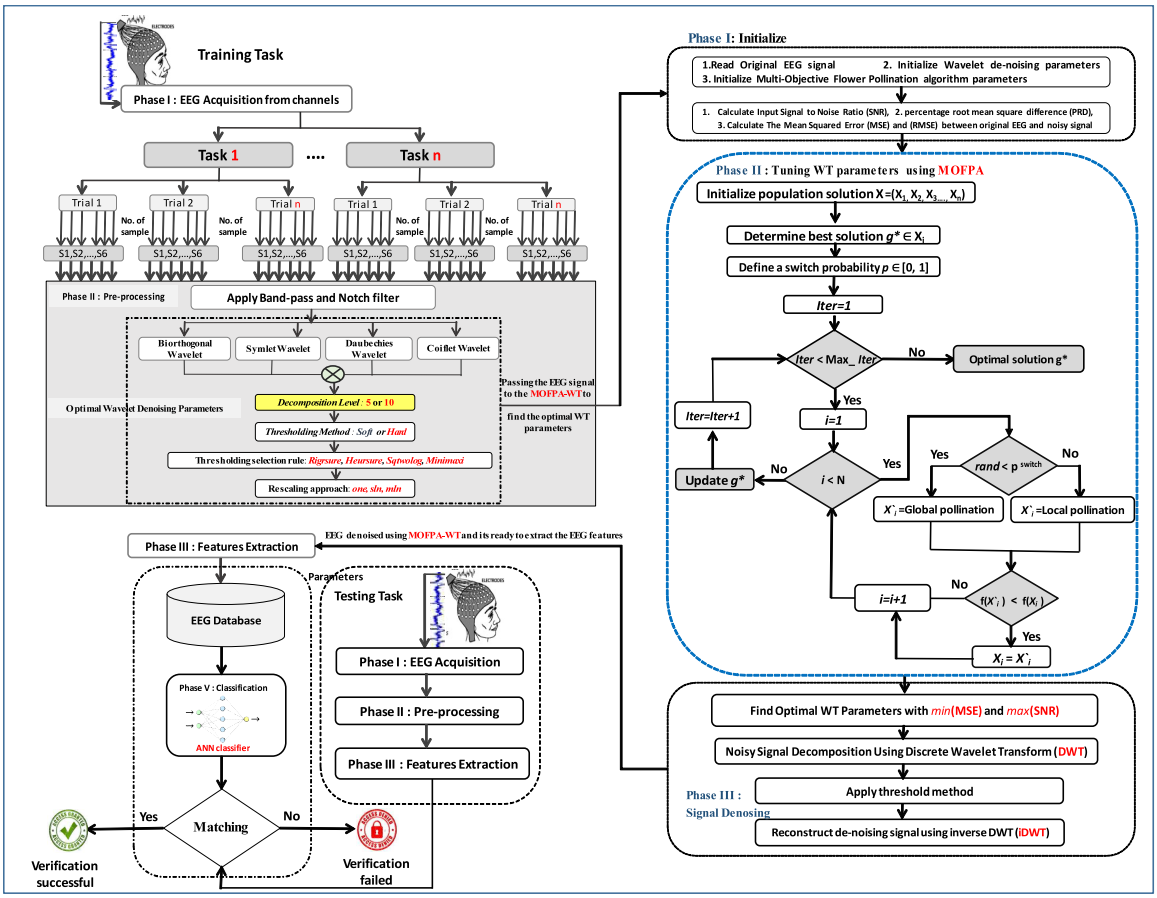


FIGURE 4. EEG signals based user identification system.

and

$$SNR = 10\log_{10}\left\{\frac{\sum_{i=1}^{N}[x_i]^2}{\sum_{i=1}^{N}[x_i-\widehat{x}_i]^2}\right\}, \tag{22}$$

where x_i and $\widehat{x}_i$ denote the original and denoised EEG signals, respectively. Notice that $\widehat{x}_i$ is obtained using the Wavelet Transform tuned by the proposed MOFPA-WT.

Iteratively, the randomly generated solutions undergo refinement using the MOFPA-WT. The final result of this phase is an optimized solution $x^* = (x_1^*, x_2^*, \ldots, x_d^*)$ that will be passed to the denoising phase, which involves three main steps that are depicted in Figure 7 and described in more details below:

- Phase I: ***Initialization*** . This phase involves three steps: firstly, reading the input EEG signal from the source. The WT denoising approach was developed based on find the optimal WT parameters for EEG signals. Secondly, initialize WT denoising parameters (Φ, $\boldsymbol{L}$, β, λ, ρ) which are shown in Table 1, as well as the operators of MOFPA are also initialized as shown in Table 6.
- Phase II: ***Tuning WT parameters by using MOFPA***. Initially, the solution of WT parameters configuration is represented as a vector $Sol = (x_1, x_2, \ldots x_d)$ where $d = 5$, and x_1 represents 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 rescaling approach ρ, where the possible range for these parameters are selected from Table 1.

Figure 6 shows an example of the selection the optimal solution of WT parameters for denoising EEG signals using MOFPA. The final result of this phase is an optimized solution $Sol'_{opt} = (s'_1, s'_2, \ldots s'_D)$, which will be passed onto the next phase.

The process of tuning WT parameters by using MOFPA is summarized as follows.

1) Initialize the set of solutions from the possible ranges of WT parameters using MOFPA. Moreover, initialize the switch probability in FPA (P).
2) Calculate the two objective functions (i.e., *min(MSE)*, *Max(SNR)*) for all solutions and determine the current best solution (***g****) to be used for the global pollination later. For the local pollination, the solution will be selected randomly from the population.

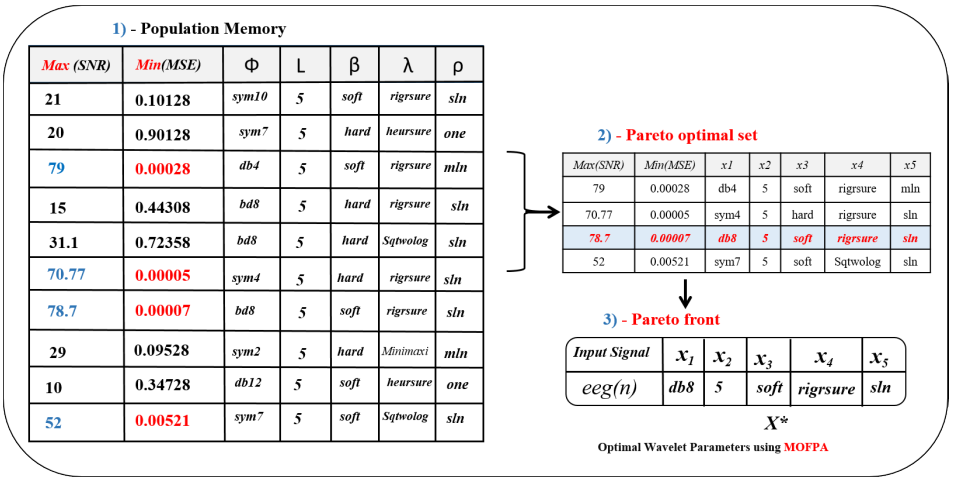


FIGURE 5. Modeling the problem of WT configuration for EEG signal denoising using MOFPA-WT.

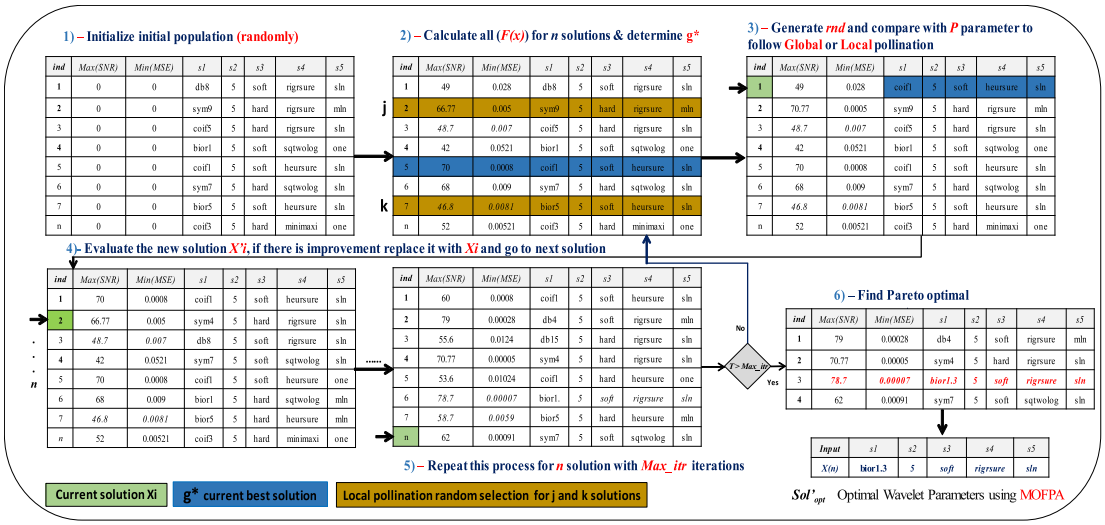


FIGURE 6. Solution of WT parameters for denoising EEG signals using MOFPA-WT.

3) Generate a random value and compare it with the switch probability in FPA P to manipulate the current solution (x_i) with global or local pollination to create the new solution (x_i').
4) Evaluate the new solution (x_i'); if an improvement is observed, replace it with the current solution (x_i) and proceed to the next solution (x_{i+1}). Repeat Steps 3 and 4 for all solution. The proposed MOFPA-WT evaluates the solution using the multi-objective function, which is formulated in Eq. (18). This method is applied using two objective functions namely, ***min(MSE)*** and ***max(SNR)*** to achieve the best combination of WT parameters for EEG signal denoising.
5) Update the current best solution (***g****).
6) Repeat Steps 3-5 based on Max_{itr}
7) The Pareto optimal set contains the set of solution with the best value for two objective functions (i.e., *min(MSE)*, *Max(SNR)*) or at least one objective function. Finally, the Pareto front solution will selected from the Pareto optimal set (best solution in the Pareto optimal set).

Iteratively, the randomly generated solutions undergoes refinement using the MOFPA-WT. The final result of this phase is an optimized solution $Sol'_{opt} = (x'_1, x'_2, \ldots x'_d)$ which will be passed to the next phase.

- Phase III: ***EEG denoising using WT based on*** Sol'_{opt}. As aforementioned in Section II-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. In the decomposition process, we must use the first two Sol'_{opt} parameters, namely, the mother wavelet furcation ρ and the decomposition level $\boldsymbol{L}$). The EEG signal is convolved using the high-pass and low-pass filters, while the block ($\downarrow 2$),

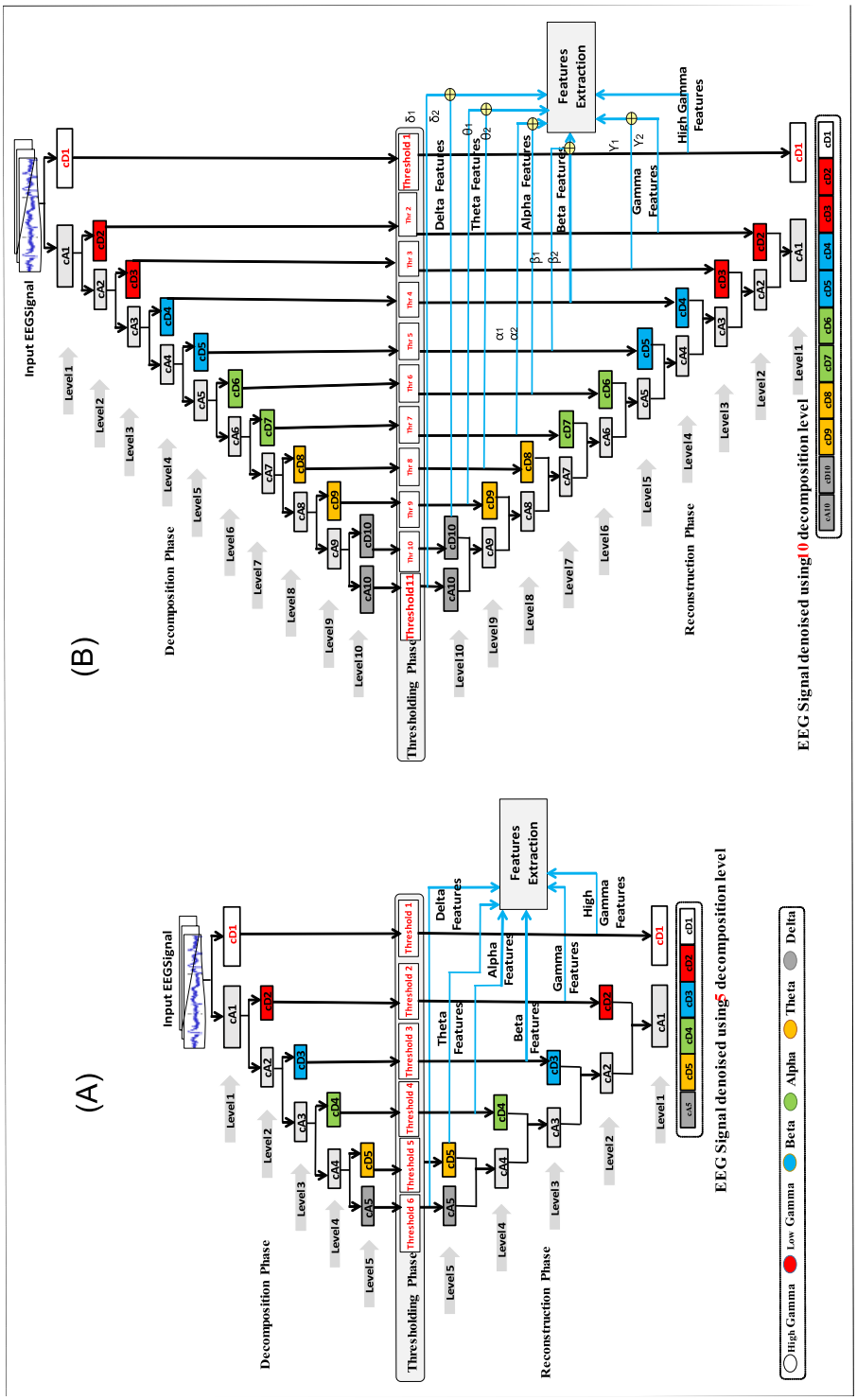


FIGURE 7. EEG feature extraction based WT decomposition with 5 and 10 levels.

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 ***The 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 Sol'_{opt}.

– ***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:

$$z[x] = iDWT[\widehat{x}] \tag{23}$$

The reconstruction convolves the EEG signals using upsampling ($\uparrow 2$), which involves the insertion of zeros at the even index elements of EEG signals.

C. EEG FEATURE EXTRACTION

The third phase in the proposed system is the EEG feature extraction, where extracting efficient features considers a significant phase in any authentication system, because it will increase the performance of the proposed system to get good results in the correct classification step [45], [48]. Therefore, the main purpose of this phase is to find the unique characteristics features from each sub-band (i.e., high gamma, gamma, alpha, beta, theta, and delta). The proposed method (MOFPA-WT) was applied for extracting EEG features with 5 and 10 wavelet decomposition levels. There are several features that can be extracted from the denoised EEG signal. In this paper, we applied the variation of energy as well as the standard measurements to extract the EEG features from the six sub-bands of EEG signal. These features are formulated as follows:

$$feature_1 = EEG_{Energy} = \sum_{j=1}^{W} \mid w_{ij} \mid^2, \quad i = 1, 2, 3, \ldots, L \tag{24}$$

where W stands for the number of coefficients w_{ij}. The next features are calculated as follows:

$$feature_2 = \frac{Energy\ of\ EEG\ Rhythm}{EEG_{Energy}} * 100, \tag{25}$$

$$feature_3 = \log(feature_2), \tag{26}$$

and

$$feature_4 = Abs(feature_3). \tag{27}$$

Figure 7 shows an EEG feature extraction step based on WT decomposition with 5 and 10 levels.

1) EEG FEATURE EXTRACTION BASED ON 5 DECOMPOSITION LEVEL

This is a traditional way for extracting EEG feature using 5 decomposition level because there are five EEG sub-bands which are namely, gamma, alpha, beta, theta, and delta. The easiest way to extract the EEG features in to determine the frequency range for each sub-band, then extract any features from these rhythms. Figure 7 (A) shows EEG feature extraction based WT decomposition with 5 levels. Table 4 shows the characteristics of EEG rhythms.

2) EEG FEATURE EXTRACTION BASED ON 10 DECOMPOSITION LEVEL

According to [24], the 10 decomposition levels for the wavelet transform can provide a better classification accuracy rate rather than 5 decomposition levels. Therefore, in this section the EEG signal will be denoised based on 10 decomposition levels as well as a new features will be extracted from the input EEG signal to reduce the similarity between each class to increase the accuracy rate. The proposed method suggested a new method to extract more features from the input EEG signal. The new features are extracted from the sub-bands and it formulated as follows:

$$feature_{new-\delta} = \frac{\delta 1 + \delta 2}{2} \tag{28}$$

$$feature_{new-\theta} = \frac{\theta 1 + \theta 2}{2} \tag{29}$$

$$feature_{new-\alpha} = \frac{\alpha 1 + \alpha 2}{2} \tag{30}$$

$$feature_{new-\beta} = \frac{\beta 1 + \beta 2}{2} \tag{31}$$

$$feature_{new-\gamma} = \frac{\gamma 1 + \gamma 2}{2} \tag{32}$$

Figure 7 (B) shows feature extraction based WT on decomposition with 10 levels. Table 5 shows the characteristics of EEG rhythms.

TABLE 4. EEG rhythms characteristics.

Rhythm	Frequency Range	Frequency bandwidth (Hz)	Decomposition level
Delta(δ)	0-4(Hz)	4	Level 5 (cA 5)
Theta(θ)	4-8(Hz)	4	Level 5 (cD 5)
Alpha(α)	8-13(Hz)	5	Level 4 (cD 4)
Beta(β)	14-30(Hz)	16	Level 3 (cD 3)
Low Gamma(γ)	30-64(Hz)	32	Level 2 (cD 2)
High Gamma(γ)	64-128(Hz)	64	Level 1 (cD 1)

TABLE 5. EEG rhythms characteristics.

Rhythm	Frequency Range	Frequency bandwidth (Hz)	Decomposition level
Delta1($\delta 1$)	0-4(Hz)	4	Level 10 (cA 10)
Delta2($\delta 2$)	0-4(Hz)	4	Level 10 (cD 10)
Theta1($\theta 1$)	5-8(Hz)	4	Level 9 (cD 9)
Theta2($\theta 2$)	5-8(Hz)	4	Level 8 (cD 8)
Alpha1($\alpha 1$)	9-13(Hz)	5	Level 7 (cD 7)
Alpha2($\alpha 2$)	9-13(Hz)	5	Level 6 (cD 6)
Beta1($\beta 1$)	14-30(Hz)	16	Level 5 (cD 5)
Beta2($\beta 2$)	14-30(Hz)	16	Level 4 (cD 4)
Low Gamma1($\gamma 1$)	31-63(Hz)	32	Level 3 (cD 3)
Low Gamma2($\gamma 2$)	31-63(Hz)	32	Level 2 (cD 2)
High Gamma(γ)	64-128(Hz)	64	Level 1 (cD 1)

D. NEURAL NETWORK CLASSIFIER

The last phase in our proposed system is the EEG feature classification. To classify the extracted features from the denoised EEG signal into correct person an artificial neural network (ANN) classifier has been applied. We used ANN and a pattern recognition tool using Matlab R2015b for classification task, and designed a network with $N - input$ feature vector of each subject (i. e., No of features * 6 sub-bands) and 32 hidden layers and S output layers where S must be equal the number of subjects which are used. Figure 8 shows the Multi-layer back propagation ANN used in this work.

V. RESULTS AND DISCUSSIONS

In this section, we will explain the EEG datasets which are used in Sec. V-A. the FPA and WT parameters setting is

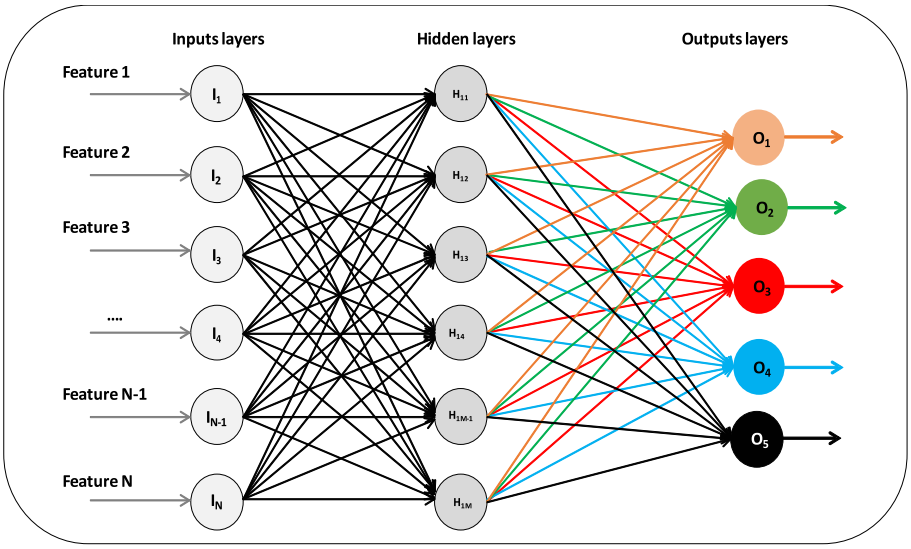


FIGURE 8. Multi-layer back propagation ANN.

described in Sec. V-B. Results of EEG feature extraction using WT with 5 and 10 decomposition level discussed in Sec. V-C and Sec. V-D, respectively. The results of MOFPA-WT are compared with state-of-Art in Sec. V-F.

A. DATASET

The multi-objective flower pollination algorithm (MOFPA) is tested using a standard EEG signal datasets, namely, the 'Motor Movement/Imagery'[2] [25]. The EEG signals are collected from 109 healthy subjects using a brain-computer interface software called BCI2000 system [47]. The EEG signals are recorded using 64 electrodes (EEG channels), each user 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 two tasks according to the position of a target that appears on the screen placed in front of them.

- **Task 1:** 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. Then the volunteer relaxes.
- **Task 2:** If the target appears on the right or left side of the screen, then the volunteer imagines open and close his/her fist corresponding to the position of the target on the screen. Then the volunteer relaxes.

In this paper, the input EEG signal collected only from the cerebral signal (Cz channel) because the cerebral region has a high activity when the user performs any motor-movement task [48]. The acquisition of the input EEG signal $x(n)$ repeated for each user three time with one mint for each trail after that the input EEG signal will divided into six samples with 10s for each sample. Figure 9 shows the distribution of electrodes in the EEG Motor Movement/Imagery Dataset.

[2]https://www.physionet.org/physiobank/database/eegmmidb/

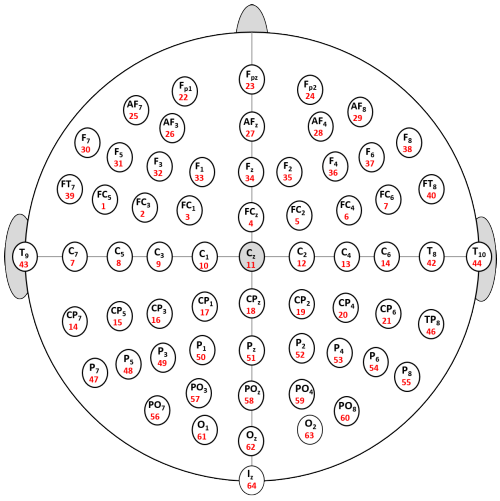


FIGURE 9. Distribution of electrodes in EEG motor movement/imagery dataset.

B. EXPERIMENTS AND PARAMETERS SETTINGS

The EEG dataset which used in this paper have been separated into four different scenarios based on training and testing approach for each task, such as (11 instances for training and 7 for testing or 10 instances for training and 8 for testing) [13], [48]. To evaluate the performance of the MOFPA-WT method five criteria: (i) accuracy, (ii) sensitivity, (iii) specificity, (iv) false acceptance rate, and (v) F-score which can formulated as follows:

$$Accuracy = \frac{TA + TR}{TA + FA + TR + FR} * 100 \quad (33)$$

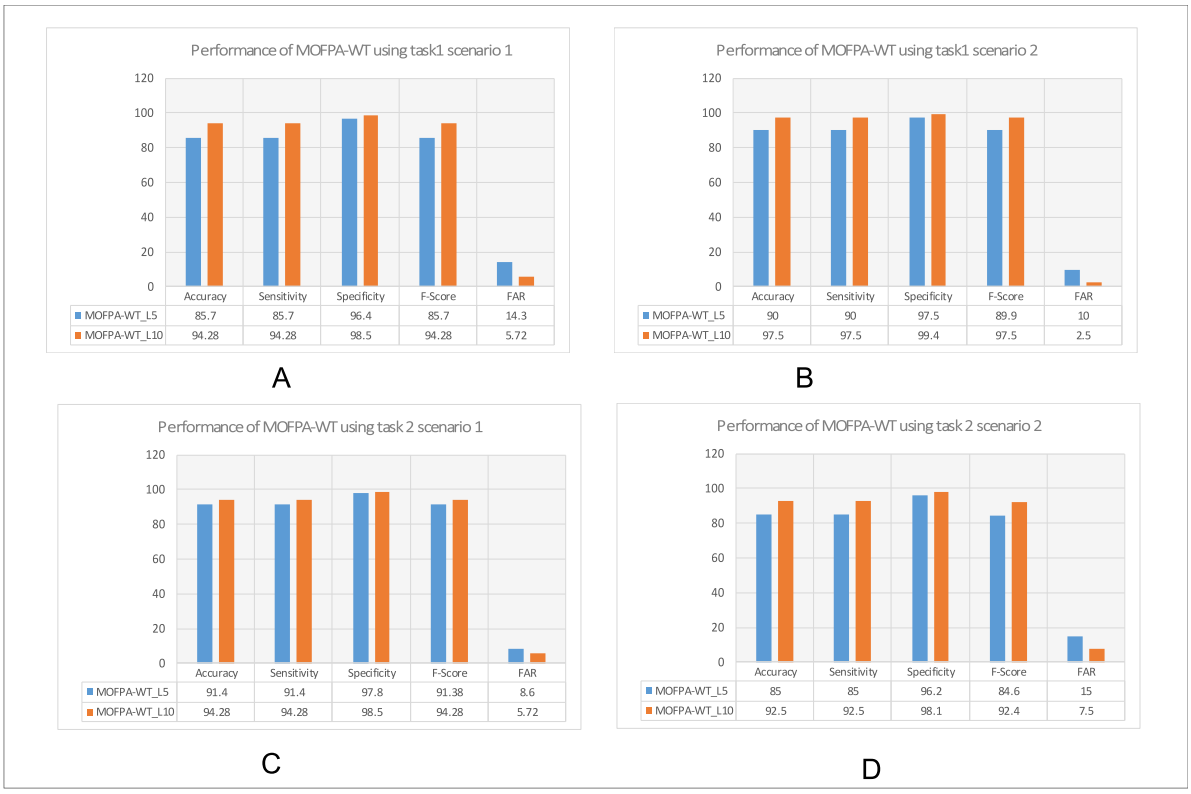


FIGURE 10. Comparison results of MOFPA-WT for 5 and 10 decomposition levels.

TABLE 6. Parameters setting for multi objective FPA.

FPA parameters	Value
Switch probability (p)	0.8
No. of iterations	100
Population size	20
Dimension of the search variables (d)	5

$$Sensitivity(TAR) = \frac{TA}{TA + FR} \tag{34}$$

$$Specifity(TFR) = \frac{TR}{TR + FR} \tag{35}$$

$$FAR = 1 - TFR \tag{36}$$

$$F - Score = \frac{2TA}{2TA + FA + FR}, \tag{37}$$

where TA,TR,FA, and FR represent true acceptance, true reject, false acceptance, and false reject, respectively. The results of classification phase are represented as a confusion matrix that tabulates whether they fall into one of four categories: true acceptance (TA), true reject (TR), false acceptance (FA) and false reject (FR).

Table 6 shows the parameters range of MOFPA-WT.

The details the EEG datasets which are used in this paper explain as follows:

$$EEG_{Data\ size} = Sub * p * K * T * S * Sub - band * F \tag{38}$$

where *sub* refers to number of subjects, P is the number of samples, K denotes number of task, T represents the number of trial, S is number of session, $Sub - band$ refers number of sub-band, and F denotes number of features which are extracted from the original EEG signal.

Table 7 shows the total number of EEG features used in this paper.

C. RESULTS OF FEATURE EXTRACTION USING MOFPA-WT BASED ON 5 DECOMPOSITION LEVEL

As aforementioned above, the EEG dataset used in this paper was divided into four different scenarios based on the standard training-and-testing approach for each task. This section discuss the results of proposed method (MOFPA-WT) using 5 decomposition level.

Table 8 presents the experiment concerning the training step with 11 persons and testing with 7 individuals (task 1). In this case, we obtained results of 85.71%, 14.28%, 0.964, 0.857, and 0.857 for accuracy, FAR, specificity, sensitivity, and F-score, respectively, which are pretty much interesting and suitable for EEG-based people identification. Table 9 presents the experiment concerning the training step with 10 persons and testing with 8 individuals (task 1). In this case, we obtained results of 90%, 10%, 0.975, 0.9, and 0.899 for accuracy, FAR, specificity, sensitivity, and F-score, respectively, i.e., we obtained better recognition rates when compared to the previous experiment but using less training cases.

TABLE 7. Total number of EEG features using 5 and 10 WT decomposition levels.

Dataset	$WT-level$	Sub	P	K	T	S	$Sub-band$	F	Dimension size
Motor Movement/Imagery	5	5	6	2	3	2	6	4	8640 features
Motor Movement/Imagery	10	5	6	2	3	2	11	9	35640 features

TABLE 8. Confusion matrix concerning the experiment of task 1 based on 5 decomposition levels.

training 11 testing 7	Sub1	Sub2	Sub3	Sub4	Sub5	Specificity	Sensitivity	F-Score
Sub1	**5**	0	2	0	0	0.928	0.714	0.714
Sub2	0	**7**	0	0	0	1	1	1
Sub3	2	0	**5**	0	0	0.928	0.714	0.714
Sub4	0	0	0	**7**	2	1	1	1
Sub5	0	0	0	1	**6**	1	0.857	0.923
FAR	2/7	0	2/7	0	1/7	Ave= 0.964	Ave=0.857	Ave=0.857

Training 11 Testing 7, Overall accuracy=**85.71%** and FAR=**14.29%**.

TABLE 9. Confusion matrix concerning the experiment of task 1 based on 5 decomposition levels.

Training 10 Testing8	Sub1	Sub2	Sub3	Sub4	Sub5	Specificity	Sensitivity	F-Score
Sub1	**7**	0	1	0	0	**0.968**	0.875	0.875
Sub2	0	**8**	0	0	0	**1**	1	0.933
Sub3	1	0	**7**	0	0	**0.968**	0.875	0.941
Sub4	0	0	0	**7**	1	**0.968**	0.875	0.875
Sub5	0	0	0	1	**7**	**0.968**	0.875	0.875
FAR	1/8	0	1/8	1/8	1/8	Ave=0.975	Ave=0.9	Ave=0.899

Training 10 Testing 8, Overall accuracy=**90.00%** and FAR=**10.00%**.

TABLE 10. Confusion matrix concerning the experiment of task 2 based on 5 decomposition levels.

training 11 testing 7	Sub1	Sub2	Sub3	Sub4	Sub5	Specificity	Sensitivity	F-Score
Sub1	**6**	0	1	0	0	**0.928**	0.857	0.8
Sub2	0	**7**	0	0	0	**1**	1	1
Sub3	2	0	**5**	0	0	**0.964**	0.714	0.769
Sub4	0	0	0	**7**	0	**1**	1	1
Sub5	0	0	0	0	**7**	**1**	1	1
FAR	0	0	1/7	1/7	1/7	Ave=0.978	Ave=0.914	Ave=0.9138

Training 10 Testing 8, Overall accuracy=**85.00%** and FAR=**15.00%**.

With respect to the task 2, Table 10 presents the experiments using 11 individuals for training and 7 for testing purposes. In this case, we obtained a TAR value as of 85%, FAR as of 15%, 0.978, 0.914, and 0.9138 for specificity, sensitivity, and F-score, respectively. Table 11 presents the experiments using 10 individuals for training and 8 for testing purposes. In this case, we obtained a TAR value as of 91.42% and FAR as of 8.58%, and 0.962, 0.85, 0.8469 for specificity, sensitivity, and F-score, respectively.

A summarizing for performance of MOFPA-WT with 5 decomposition level, the best recognition rates were obtained in both tasks using scenario 10 individuals for training and 8 for testing purposes. For scenario, 11 individuals for training and 7 for testing purposes task 1 obtained better results compared with task 2.

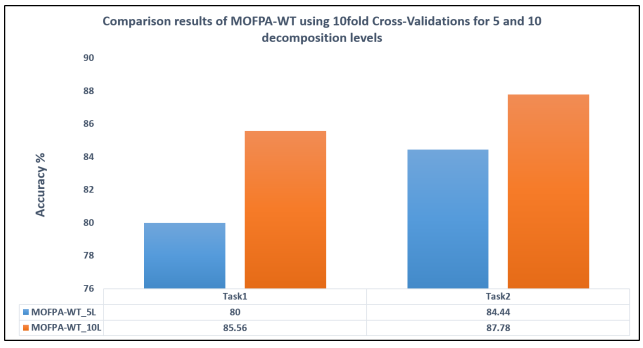


FIGURE 11. Accuracy results for task 1 and task 2 using a 10-fold cross-validation.

D. RESULTS OF FEATURE EXTRACTION USING MOFPA-WT BASED ON 10 DECOMPOSITION LEVEL

As one mentioned in Sec IV-C2 the WT with 10 decomposition levels can provide a better recognition rate [24]. Since it decomposes the EEG signal with 10 decomposition levels, one can extract new features from the input EEG signal in order to reduce the similarity between each class, as well will increasing accuracy rate.

Table 12 presents the experiment concerning the training step with 11 persons and testing with 7 individuals (task 1). In this case, we obtained results of 94.28%, 5.72%, 0.985, 0.942, and 0.942 for accuracy, FAR, specificity, sensitivity,

TABLE 11. Confusion matrix concerning the experiment of task 2 based on 5 decomposition levels.

Training 10 Testing8	Sub1	Sub2	Sub3	Sub4	Sub5	Specificity	Sensitivity	F-Score
Sub1	**6**	0	2	0	0	**0.968**	0.75	0.8
Sub2	0	**8**	0	0	0	**1**	1	1
Sub3	1	0	**7**	0	0	**0.937**	0.875	0.823
Sub4	0	0	0	**5**	3	**1**	0.625	0.769
Sub5	0	0	0	0	**8**	**0.906**	1	0.842
FAR	2/8	0	1/8	3/8	0	Ave=0.962	Ave=0.85	Ave=0.8469

Training 11 Testing 7, Overall accuracy=**91.42%** and FAR=**8.58%**.

TABLE 12. Confusion matrix concerning the experiment of task 1 based on 10 decomposition levels.

training 11 testing 7	Sub1	Sub2	Sub3	Sub4	Sub5	Specificity	Sensitivity	F-Score
Sub1	**6**	0	1	0	0	0.964	0.857	0.857
Sub2	0	**7**	0	0	0	1	1	1
Sub3	1	0	**6**	0	0	0.964	0.857	0.857
Sub4	0	0	0	**7**	0	1	1	1
Sub5	0	0	0	0	**7**	1	1	1
FAR	0.142	0	0.142	0	0	Ave=0.985	Ave=0.942	Ave=0.942

Training 11 Testing 7, Overall accuracy=**94.28%** and FAR=**5.72%**.

TABLE 13. Confusion matrix concerning the experiment of task 1 based on 10 decomposition levels.

Training 10 Testing8	Sub1	Sub2	Sub3	Sub4	Sub5	Specificity	Sensitivity	F-Score
Sub1	**8**	0	0	0	0	0.969	1	0.941
Sub2	0	**8**	0	0	0	**1**	**1**	**1**
Sub3	1	0	**7**	0	0	**1**	**1**	**1**
Sub4	0	0	0	**8**	0	1	0.875	0.933
Sub5	0	0	0	0	**8**	**1**	**1**	**1**
FAR	0	0	0.125	0	0	Ave=0.994	Ave=0.975	Ave=0.975

Training 10 Testing 8, Overall accuracy=**97.5%** and FAR=**2.5%**.

TABLE 14. Confusion matrix concerning the experiment of task 2 based on 10 decomposition levels.

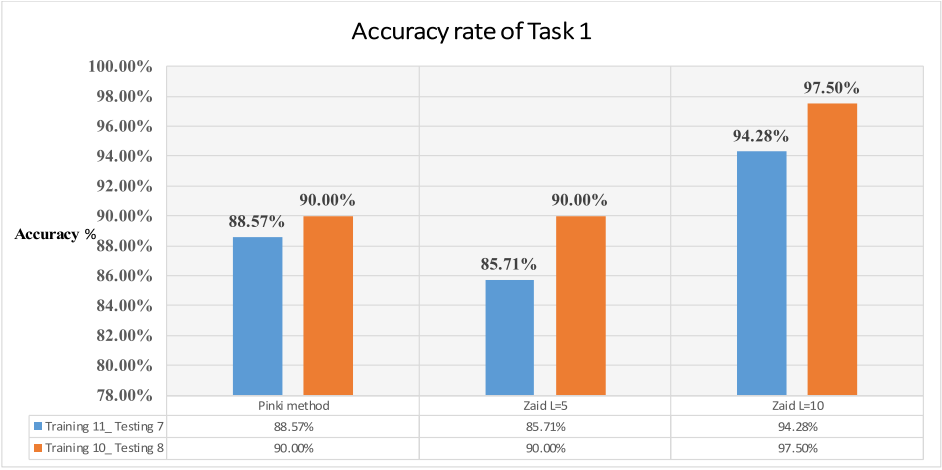
training 11 testing 7	Sub1	Sub2	Sub3	Sub4	Sub5	Specificity	Sensitivity	F-Score
Sub1	**6**	0	1	0	0	**0.964**	0.857	0.857
Sub2	0	**7**	0	0	0	**1**	1	1
Sub3	1	0	**6**	0	0	**0.964**	0.857	0.857
Sub4	0	0	0	**7**	0	**1**	1	1
Sub5	0	0	0	0	**7**	**1**	1	1
FAR	1/7	0	1/7	0	0	Ave=0.985	Ave=0.942	Ave=0.942

Training 11 Testing 7, Overall accuracy=**94.28%** and FAR=**5.72%**.

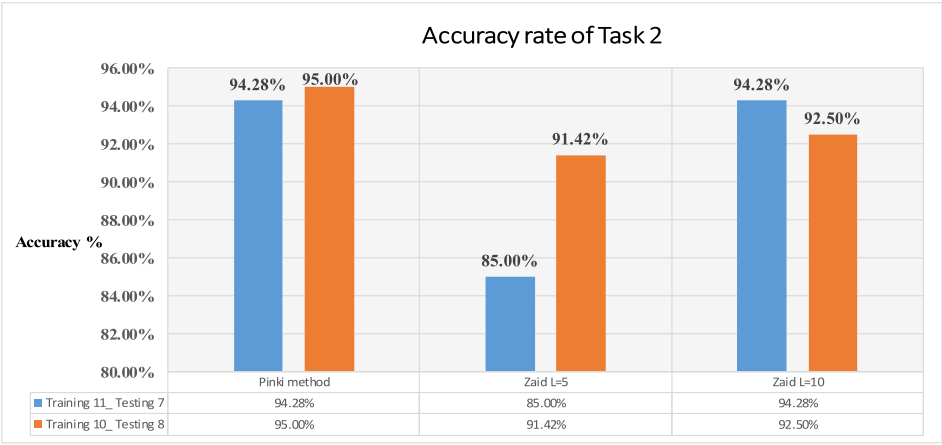
and F-score, respectively. which are pretty much interesting and suitable for EEG-based people identification. Table 13 presents the experiment concerning the training step with 10 persons and testing with 8 individuals (task 1). In this case, we obtained results of 97.5% and FAR as of 2.5%, 0.994, 0.975, and 0.975 for accuracy, FAR, specificity, sensitivity, and F-score, respectively.

With respect to the task 2, Table 14 presents the experiments using 11 individuals for training and 7 for testing purposes. In this case, we obtained results of 94.28%, 5.72%, 0.985, 0.942, and 0.942 for accuracy, FAR, specificity, sensitivity, and F-score, respectively. Table 15 presents the experiments using 10 individuals for training and 8 for testing purposes. In this case, we obtained results of 92.5%, 7.5%, 0.981, 0.925, and 0.924 for accuracy, FAR, specificity, sensitivity, and F-score, respectively.

In a nutshell, the new approach for EEG feature extraction using 10 decomposition levels shown significant improvements compared with the approach MOFPA-WT using 5 decomposition level. Figure 12 shows the comparison results between Five and Ten WT decomposition levels using five criteria: (i) accuracy, (ii) sensitivity, (iii) specificity, (iv) false acceptance rate, and (v) F-score. The new approach for EEG feature extraction using MOFPA-WT with 10 decomposition levels achieves the best results according to all criteria measures for all tasks. Also, the task 1 using 11 individuals for training and 7 for testing obtained the highest accuracy rate with 97.5%.



A) Accuracy rate of task 1



B) Accuracy rate of task 2

FIGURE 12. Comparison of the accuracy results for task 1 and task 2 with [13], [48].

TABLE 15. Confusion matrix concerning the experiment of task 2 based on 10 decomposition levels.

Training 10 Testing8	Sub1	Sub2	Sub3	Sub4	Sub5	Specificity	Sensitivity	F-Score
Sub1	**7**	0	1	0	0	**0.937**	0.875	0.823
Sub2	0	**8**	0	0	0	**1**	1	1
Sub3	2	0	**6**	0	0	**0.968**	0.75	0.8
Sub4	0	0	0	**8**	0	**1**	1	1
Sub5	0	0	0	0	**8**	**1**	1	1
FAR	1/8	0	2/8	0	0	Ave=0.981	Ave=0.925	Ave=0.924

Training 10 Testing 8, Overall accuracy=**92.5%** and FAR=**7.5%**.

E. RESULTS OF ACCURACY RATE USING MOFPA-WT BASED ON 5 AND 10 DECOMPOSITION LEVEL USING 10 FOLD CROSS-VALIDATION

We able performed the Wilcoxon signed-rank statistical test [53] to verify whether there is a significant difference between MOFPA-WT with 5 and 10 decomposition levels. Table 16 shows the comparison among them for task1 and task2 based on MOFPA-WT 5 and 10 decomposition levels.

F. COMPARING WITH STATE-OF-ART RESULTS

In order to compare the result performance of MOFPA-WT with state-of-art, Figure 12 shows the comparison between the proposed method with [13], [48]. The accuracy rate are used to measure the performance of the best methods, where the proposed approach with 10 decomposition levels shown a powerful result on task 1 in both training and testing scenarios. For the task 2, the performance of MOFPA-WT approach still achieves good results using 11 individuals for training and 7 for testing where it obtained an accuracy of 94.28%. However, for the scenario was 10 individuals for training and 8 for testing; Pinki method [48] achieved better results than our proposed method. That has happened due to the input EEG signals have been separated randomly.

TABLE 16. Wilcoxon signed-rank test evaluation of MOFPA-WT 5 and 10 levels.

Dataset	P-Value	Mean Difference	MOFPA-WT 10 levels
Task 1	0.05	5.56	Significant
Task 2	0.05	2.34	Significant

VI. CONCLUSIONS AND FUTURE WORK

In this paper, a novel technique for EEG feature extraction based on multi-objective flower pollination algorithm with multi-level wavelet decomposition is proposed. The main task of the proposed method (MOFPA-WT) is to find the efficient decomposition of the input EEG signal which can provide unique features from each sub-band. MOFPA-WT is tested using a standard EEG signal dataset, namely, EEG motor movement/imagery dataset. The performance of MOFPA-WT is evaluated using three criteria, namely accuracy, TAR, and FAR. The proposed method achieves the highest accuracy using the cognitive tasks based on motor movement compared with the results based on motor imagination.

Regarding future works, MOFPA-WT will be applied to extract all the EEG features such as time domain, frequency domain, and time-frequency domain features to find the best features for EEG which can achieve the highest accuracy rate. On the other hand, testing multi-classifiers techniques are recommended. Furthermore, the real-world applications are required to show the efficiency of MOFPA-WT. Other possible improvements is to apply one of features selection techniques to increase the accuracy rate as well as to reduce the redundancy of the extracted EEG features.

REFERENCES

[1] S. N. Abdulkader, A. Atia, and M.-S. M. Mostafa, "Brain computer interfacing: Applications and challenges," *Egyptian Inform. J.*, vol. 16, no. 2, pp. 213–230, 2015.

[2] M. Abo-Zahhad, S. M. Ahmed, and S. N. Abbas, "A new multi-level approach to EEG based human authentication using eye blinking," *Pattern Recognit. Lett.*, vol. 82, pp. 216–225, Oct. 2016.

[3] L. M. Abualigah, A. T. Khader, M. A. Al-Betar, Z. A. A. Alyasseri, O. A. Alomari, and E. S. Hanandeh, "Feature selection with β-hill climbing search for text clustering application," in *Proc. Palestinian Int. Conf. Inf. Commun. Technol. (PICICT)*, May 2017, pp. 22–27.

[4] M. A. Al-Betar, "β-hill climbing: An exploratory local search," *Neural Comput. Appl.*, vol. 28, no. 1, pp. 153–168, Dec. 2017.

[5] 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, 2016.

[6] M. I. Al-Kadi, M. B. I. Reaz, M. A. M. Ali, and C. Y. 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, 2014.

[7] N. K. Al-Qazzaz, S. H. B. M. Ali, S. A. Ahmad, M. S. 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, 2015.

[8] 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. ICISPC*, Penang, Malaysia, 2017, pp. 100–105.

[9] Z. A. A. Alyasseri, A. T. Khader, and M. A. Al-Betar, "Optimal EEG signals denoising using hybrid β-hill climbing algorithm and wavelet transform," in *Proc. ICISPC*, Penang, Malaysia, 2017, pp. 106–112.

[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)*, May 2017, pp. 1–7.

[11] 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.

[12] Z. A. A. Alyasseri, A. T. Khader, M. A. Al-Betar, M. A. Awadallah, and X.-S. Yang, "Variants of the flower pollination algorithm: A review," in *Nature-Inspired Algorithms and Applied Optimization*. Cham, Switzerland: Springer, 2018, pp. 91–118.

[13] Z. A. A. Alyasseri, A. T. Khader, M. A. Al-Betar, P. P. Joao, and A. A. Osama, "EEG-based person authentication using multi-objective flower pollination algorithm," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jul. 2018, pp. 15–22.

[14] Z. A. A. Alyasseri, A. T. Khader, M. A. Al-Betar, J. P. Papa, A. A. Osama, 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. 25–31.

[15] 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)*, Sep. 2012, pp. 47–52.

[16] M. F. Azis, A. Ryanta, D. F. U. Putra, and O. Fenno, "Dynamic economic dispatch considering emission using multi-objective flower pollination algorithm," in *Proc. ASEAN/Asian Acad. Soc. Int. Conf. Series*, 2015, pp. 6–10.

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

[18] 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.

[19] S. Borse, "EEG de-noising using wavelet transform and fast ica," *Int. J. Innov. Sci., Eng., Technol.*, vol. 2, no. 7, pp. 200–205, 2015.

[20] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.

[21] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

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

[23] J. Ferdous, "Assessing self-similarity and cross-similarity between EEG patterns for biometrical applications," M.S. thesis, Dept. Elect. Eng., Lamar Univ., Beaumont, TX, USA, 2016.

[24] 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, 2011.

[25] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[26] I. Jayarathne, M. Cohen, and S. Amarakeerthi, "BrainID: Development of an EEG-based biometric authentication system," in *Proc. IEEE 7th Annu. Inf. Technol., Electron. Mobile Commun. Conf. (IEMCON)*, Oct. 2016, pp. 1–6.

[27] Z. A. Keirn and J. I. Aunon, "A new mode of communication between man and his surroundings," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 12, pp. 1209–1214, Dec. 1990.

[28] 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.

[29] P. Kumari and A. Vaish, "Brainwave based authentication system: Research issues and challenges," *Int. J. Comput. Eng. Appl.*, vol. 4, nos. 1–2, pp. 89–108, 2014.

[30] 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.

[31] 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, 2016.

[32] 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, 2013.

[33] S. Mirjalili, S. Saremi, S. M. Mirjalili, and L. D. S. Coelho, "Multi-objective grey wolf optimizer: A novel algorithm for multi-criterion optimization," *Expert Syst. Appl.*, vol. 47, pp. 106–119, Apr. 2016.

[34] M. R. Mowla, S.-C. Ng, M. S. A. 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.
[35] J. P. Papa and A. X. Falcão, "A new variant of the optimum-path forest classifier," in *Proc. Int. Symp. Vis. Comput.* Berlin, Germany: Springer, 2008, pp. 935–944.
[36] J. P. Papa, A. X. Falcão, V. H. C. De Albuquerque, and J. M. R. S. Tavares, "Efficient supervised optimum-path forest classification for large datasets," *Pattern Recognit.*, vol. 45, no. 1, pp. 512–520, 2012.
[37] S. Poornachandra and N. Kumaravel, "Hyper-trim shrinkage for denoising of ECG signal," *Digit. Signal Process.*, vol. 15, no. 3, pp. 317–327, 2005.
[38] K. Rajalashmi and S. U. Prabha, "A hybrid algorithm for multiobjective optimal power flow problem using particle swarm algorithm and enhanced flower pollination algorithm," *Asian J. Res. Social Sci. Humanities*, vol. 7, no. 1, pp. 923–940, 2017.
[39] R. Rajaram and K. S. Kumar, "Multiobjective power loss reduction using flower pollination algorithm," *Int. J. Control Theory Appl.*, vol. 8, no. 5, pp. 2239–2245, 2015.
[40] R. A. Ramadan and A. V. Vasilakos, "Brain computer interface: Control signals review," *Neurocomputing*, vol. 223, pp. 26–44, Feb. 2017.
[41] R. P. Rao, *Brain-Computer Interfacing: An Introduction*. Cambridge, U.K.: Cambridge Univ. Press, 2013.
[42] D. Rodrigues, G. F. A. 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.
[43] M. V. Ruiz-Blondet, Z. Jin, and S. Laszlo, "CEREBRE: A novel method for very high accuracy event-related potential biometric identification," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 7, pp. 1618–1629, Jul. 2016.
[44] G. Safont, A. Salazar, A. Soriano, and L. Vergara, "Combination of multiple detectors for EEG based biometric identification/authentication," in *Proc. IEEE Int. Carnahan Conf. Secur. Technol. (ICCST)*, Oct. 2012, pp. 230–236.
[45] N. D. Sarier, "Improving the accuracy and storage cost in biometric remote authentication schemes," *J. Netw. Comput. Appl.*, vol. 33, no. 3, pp. 268–274, 2010.
[46] C. Sawant and H. T. Patii, "Wavelet based ECG signal de-noising," in *Proc. 1st Int. Conf. Netw. Soft Comput. (ICNSC)*, Aug. 2014, pp. 20–24.
[47] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000: A general-purpose brain-computer interface (BCI) system," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1034–1043, Jun. 2004.
[48] P. K. 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, 2016.
[49] M. Shehab, A. T. Khader, and M. A. Al-Betar, "A survey on applications and variants of the cuckoo search algorithm," *Appl. Soft Comput.*, vol. 61, pp. 1041–1059, Dec. 2017.
[50] C. Shilaja and K. Ravi, "Multi-objective optimal power flow problem using enhanced flower pollination algorithm," *Gazi Univ. J. Sci.*, vol. 30, no. 1, pp. 79–91, 2017.
[51] 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, 2006.
[52] V. Tamilselvan and T. Jayabarathi, "Multi objective flower pollination algorithm for solving capacitor placement in radial distribution system using data structure load flow analysis," *Arch. Elect. Eng.*, vol. 65, no. 2, pp. 203–220, 2016.
[53] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bull.*, vol. 1, no. 6, pp. 80–83, 1945.
[54] R. Yang and M. Ren, "RETRACTED: Wavelet denoising using principal component analysis," *Expert Syst. Appl.*, vol. 38, no. 1, pp. 1073–1076, 2011.
[55] X.-S. Yang, "Flower pollination algorithm for global optimization," in *Proc. Int. Conf. Unconventional Comput. Natural Comput.* Berlin, Germany: Springer, 2012, pp. 240–249.
[56] X.-S. Yang, M. Karamanoglu, and X. He, "Multi-objective flower algorithm for optimization," *Proc. Comput. Sci.*, vol. 18, pp. 861–868, Jan. 2013.
[57] X.-S. Yang, M. Karamanoglu, and X. He, "Flower pollination algorithm: A novel approach for multiobjective optimization," *Eng. Optim.*, vol. 46, no. 9, pp. 1222–1237, Apr. 2013.

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 Universiti Sains Malaysia (USM), in 2013, where he is currently pursuing the Ph.D. degree in the field of artificial intelligence (brain-inspired computing). Since 2016, he has been with the Computational Intelligence Research Group, School of Computer Sciences, USM. He is currently a Senior Lecturer with the University of Kufa, Iraq. His research interests are optimization, pattern recognition, EEG, brain-computer interface, and signal and image processing.

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 (USM). He is also the Dean of the School of Computer Sciences, USM. 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, Al-Huson University College, Al-Balqa Applied University, Jordan. His research interests are mainly directed to metaheuristic optimization methods and hard combinatorial optimization problems.

JOÃO P. PAPA received the B.Sc. degree in information systems from São Paulo State University, Brazil, the M.Sc. degree in computer science from the Federal University of São Carlos, Brazil, in 2005, and the Ph.D. degree in computer science from the University of Campinas, Brazil, in 2008. From 2008 to 2009, he was a Post-Doctoral Researcher with the University of Campinas, and from 2014 to 2015, he was a Visiting Scholar with Harvard University. He has been a Professor with the Computer Science Department, São Paulo State University, since 2009. His research interests include machine learning, pattern recognition, and image processing. He is also a recipient of the Alexander von Humboldt fellowship and the Brazilian representative of the International Association for Pattern Recognition.

OSAMA AHMAD ALOMARI received the M.Sc. degree in computer science from the Universiti Kebangsaan Malaysia, Malaysia, in 2014. 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.