Computers in Biology and Medicine 112 (2019) 103348

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Utilizing hybrid functional fuzzy wavelet neural networks with a teaching learning-based optimization algorithm for medical disease diagnosis

Jamal Salahaldeen Majeed Alneamy[a], Zakaria A. Hameed Alnaish[b,*], S.Z. Mohd Hashim[c], Rahma A. Hamed Alnaish[a]

[a] Department of Software Engineering, Computer and Mathematics Science College, University of Mosul, Mosul, Iraq
[b] College of Science, University of Mosul, Mosul, Iraq
[c] Faculty of Computer Science and Information Systems, Universiti Teknologi Malaysia, 81310, Skudai, Johor, Malaysia

ARTICLE INFO

Keywords:
Fuzzy wavelet neural network (FWNN)
Functional link neural network (FLNN)
Teaching learning-based optimization algorithm (TLBO)
Functional fuzzy wavelet neural network (FFWNN)

ABSTRACT

Accurate medical disease diagnosis is considered to be an important classification problem. The main goal of the classification process is to determine the class to which a certain pattern belongs. In this article, a new classification technique based on a combination of The Teaching Learning-Based Optimization (TLBO) algorithm and Fuzzy Wavelet Neural Network (FWNN) with Functional Link Neural Network (FLNN) is proposed. In addition, the TLBO algorithm is utilized for training the new hybrid Functional Fuzzy Wavelet Neural Network (FFWNN) and optimizing the learning parameters, which are weights, dilation and translation. To evaluate the performance of the proposed method, five standard medical datasets were used: Breast Cancer, Heart Disease, Hepatitis, Pima-Indian diabetes and Appendicitis. The efficiency of the proposed method is evaluated using 5-fold cross-validation and 10-fold cross-validation in terms of mean square error (MSE), classification accuracy, running time, sensitivity, specificity and kappa. The experimental results show that the efficiency of the proposed method for the medical classification problems is 98.309%, 91.1%, 91.39%, 88.67% and 93.51% for the Breast Cancer, Heart Disease, Hepatitis, Pima-Indian diabetes and Appendicitis datasets, respectively, in terms of accuracy after 30 runs for each dataset with low computational complexity. In addition, it has been observed that the proposed method has efficient performance compared with the performance of other methods found in the related previous studies.

1. Introduction

In recent decades, the utilization of classification systems in the field of medical diagnosis has gradually increased. While there is no doubt that the experience of specialists is very important in the diagnosis process, expert systems have the ability to support specialists and facilitate accurate diagnosing in shorter periods of time. In addition, such systems help to reduce potential mistakes by doctors with limited experience [1]. Various classification approaches, such as Support Vector Machine (SVM) [2], Artificial Immune System (AIS) [3] and Artificial Neural Network (ANN) [4,5], are quite popular for the diagnosing of many diseases, for instance, breast cancer, heart disease, hepatitis, diabetes and appendicitis [6–8]. ANN has gained wide attention in various fields, including classification, due to its ability to learn or adapt from examples and generalize data. However, ANN is only considered an effective artificial intelligence method when enough training data is available, and it can be difficult to describe its structure [9,10]. Alternately, fuzzy logic [11] has the ability to deal with cognitive uncertainties in a manner similar to that of humans [12]. Fuzzy logic is used to improve the ability of the neural network and increase its learning rate [13]. Thus, the combination of a neural network and fuzzy logic in one system leads to the creation of another computational tool called a fuzzy neural network (FNN), which combines the advantages of both approaches [12,13]. In spite of the fact that the FNN/NN has many advantages, it still suffers from some drawbacks, such as slow training speed, high approximation error, and poor convergence problems [14]. Thus, it does not seem suitable for use as a powerful tool for accurate diagnosis. Moreover, the wavelet neural network (WNN), which has drawn much research interest, combines both the wavelet function and neural network [14,15]. The main advantages of the WNN are better generalization capability [16], faster learning property [14,16], and smaller approximation errors and sizes of networks than NN [14]. Due

* Corresponding author.
E-mail addresses: jamal_alneamy@yahoo.com (J.S. Majeed Alneamy), zakriahamoalnaish@gmail.com (Z. A. Hameed Alnaish), sitizaiton@utm.my (S.Z. Mohd Hashim), postgraduatstudent@yahoo.com (R.A. Hamed Alnaish).

https://doi.org/10.1016/j.compbiomed.2019.103348
Received 26 January 2019; Received in revised form 30 June 2019; Accepted 1 July 2019

to these great advantages, it is able to overcome the obstacles of FNN/NN, especially in highly nonlinear systems [14]. Therefore, the Fuzzy Wavelet Neural Network (FWNN), which merges the wavelet function, neural network and fuzzy system, has been considered in different application areas [17–19]. In addition, the Functional Link neural network (FLNN) [20] is characterized by simple architecture and low computational complexity of the network. Consequently, it effectively achieves fast convergence rate. This encourages the utilization of the FLNN in classification tasks [21,22]. Considering the aforementioned discussion, in this paper, a new hybrid method called Functional Fuzzy Wavelet Neural Network (FFWNN) is proposed. This method utilizes the functional link neural network (FLNN) and the wavelet function in the consequent part of the fuzzy rules.

The training of the FFWNN, to find the best values of FFWNN parameters and the global optimal solution, is considered to be a very crucial process. The functional fuzzy wavelet neural network has many parameters, including the weights, dilation, translation, center and width of the membership function parameters. This process requires robust optimization techniques to make the performance of the FFWNN more accurate and efficient. According to a comparison study [23–26] of various recent nature-inspired optimization algorithms, including Particle Swarm Optimization (PSO) [27,28], Differential Evolution (DE) [29,30], Teaching Learning-Based Optimization (TLBO) [31], Artificial Bee Colony (ABC) [32,33], and the Firefly Algorithm (FA) [34,35], the TLBO algorithm is more accurate, effective and efficient, thus exhibiting superior performance. Moreover, TLBO has balanced exploration and exploitation abilities, so it does not encounter the problem of becoming trapped in local minima [36].). In addition, TLBO is utilized as an optimization algorithm for continuous nonlinear large-scale problems [23]. Thus, in this study, the teaching learning-based optimization (TLBO) algorithm is utilized to adjust the weights, dilation and translation parameters in FFWNN.

The rest of this paper is organized as follows. In Section 2, previous related studies are presented. Sections 3 through 5 provide a brief introduction to the concepts of the structure of the fuzzy wavelet neural network (FWNN) and functional link neural network (FLNN) and an overview of the teaching learning-based optimization algorithm (TLBO), respectively. Section 6 depicts the framework of the proposed method (TLBO-FFWNN). In Section 7, accuracy, complexity, sensitivity, specificity and kappa analysis are given. Section 8 explains and discusses the experimental results. Finally, a conclusion is given in Section 9.

2. Related works

Throughout recent decades, several studies have been concerned with solving classification problems, especially the classification of medical datasets. In 2007, Socha and Blum proposed a hybrid method based on the Ant Colony Optimization (ACO) algorithm and Back Propagation (BP) algorithm to train feed-forward neural networks for medical classification problems. The results showed that the performance of the ACO algorithm is comparable with that of the BP algorithm and better than that of the Genetic Algorithm (GA) [37]. [38] designed a hybrid system of a fuzzy neural network (FNN) and artificial neural network (ANN) for diagnosing diabetes and heart diseases, where the classifier parameters were optimized by the Back Propagation (BP) algorithm. The results showed that the proposed method is comparable with other methods [38]. In addition [39], proposed a system for heart disease diagnosis by using a neural network ensemble model that combines several individual neural networks, but this method increased the complexity and the execution time [39]. [4] presented a classification approach for diagnosing heart disease using Multi-Layer Perceptron (MLP) with the Back Propagation learning algorithm, as well as a feature selection algorithm that used 8 features instead of 13. The results showed that the accuracy rate was enhanced by 1.1% in the training data set and 0.82% in the testing data set [4]. [6] applied centripetal accelerated particle swarm optimization (CAPSO) to enhance the learning of an Artificial Neural Network (ANN) to classify the data of various standard medical datasets [6].

Then [40], utilized the Ant Colony Optimization (ACO) algorithm to train a feed-forward neural network for pattern classification. Moreover, the ACO training algorithm is hybridized with gradient descent training. The results showed that the ACO algorithm can be a useful technique in neural network training for pattern classification, especially when it is hybridized with gradient descent training methods [40]. [41] proposed the use of a GA algorithm for optimizing ANFIS to solve physical work rate classification [41]. Moreover [42], developed a new ensemble of classifiers that enables the efficient recognition of heart disorders based on support vector machines classifier genetic and optimization of classifiers parameters and novel genetic training utilized for combining classifiers [42]. [43]; proposed a new emotion recognition system based on facial expression images, which employed biorthogonal wavelet entropy for feature extraction and utilized fuzzy multiclass support vector machine for classification. Statistical analysis showed that this method is efficient [43]. In addition, in 2017, Wang S. et al. utilized wavelet contour analysis for backbone detection and then merged wavelet packet entropy (WPE) with fuzzy support vector machine (FSVM) for spine classification. The results showed that this method is promising and has some advantages that can reduce the noise, error rate and computing time, in addition to increasing the computing efficiency [44].

All previous studies have shown that the classification accuracy has not yet reached the desirable level. However, because the main objective of all previous studies was to increase the classification accuracy, thereby making the disease diagnosis process more accurate and efficient, it is, therefore, highly necessary to propose a new hybrid method to increase the classification accuracy and decrease the error rate.

3. Fuzzy wavelet neural network (FWNN)

The structure of FWNN, described in Fig. 1, contains seven layers [18,45–47], as described below:

Layer 1: there is no computation in this layer, wherein each node in this layer directly transmits the values of the input variables to the second layer.

Layer 2: the values of input variables are fuzzified in this layer, wherein each node represents one fuzzy set. Thus, the output of layer 2 represents the membership value of the input variable to the corresponding fuzzy set based on the selected membership function.

Layer 3: this layer receives the membership degrees of the inputs from the previous layer to create the antecedent part of each fuzzy rule, wherein each node in this layer represents one fuzzy rule. The output of this layer can be calculated using equation (1).

$$\mu_i = (x) = \prod_{j=1}^{q} A_j^i(x_j) \tag{1}$$

where Π refers to the product operation and A_j^i acts as a membership function that is used to calculate the membership degrees of the input variables. In this study, a Gaussian membership function is used, which can be represented by equation (2).

$$A_j^i(x_j) = \exp\left[-\left(\frac{x_j - c_j^i}{\sigma_j^i}\right)^2\right] \tag{2}$$

In equation (2), c_j^i and σ_j^i represent the center and width of the membership function, respectively.

Layer 4: nodes in this layer are called consequent nodes because they represent the consequent part of the fuzzy rules. Each node contains the wavelet function (ψ_{ij}). The wavelet function is a waveform that has limited duration, and the mean value of this duration is zero. This function has two parameters, i.e., the dilation parameter and translation parameter.

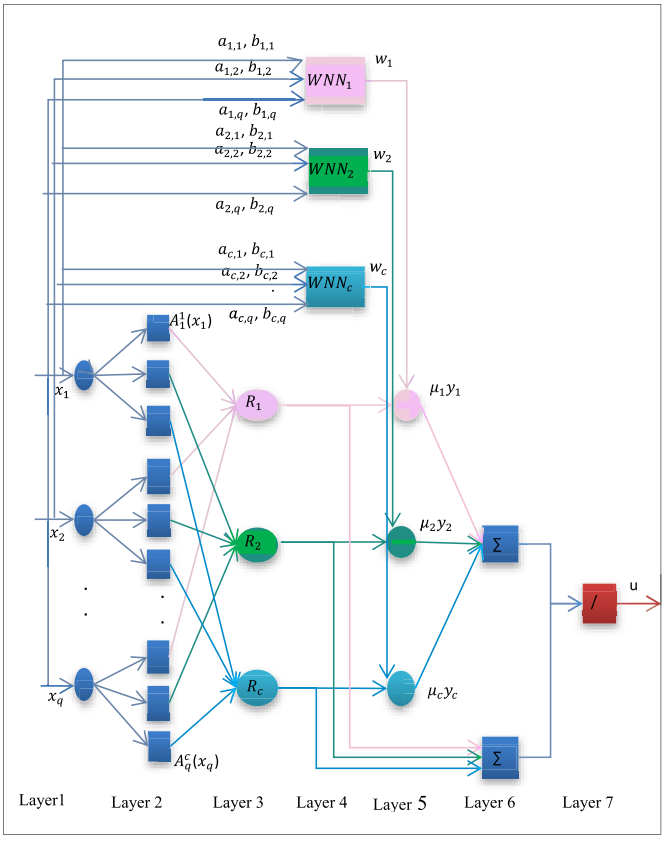


Fig. 1. The general structure of a Fuzzy Wavelet Neural Network [18].

In FWNN, fuzzy rules can be represented by the following equation:

$$R^i: \text{if } x1 \text{ is } A_1^i, \text{ and } x2 \text{ is } A_2^i \ldots \text{and } xq \text{ is } A_q^i$$
$$\text{then } y_i = w_i \sum_{j=1}^{q} \psi_{ij}(x_j) \tag{3}$$

where R^i refers to the *ith* rule, and ($1 \le i \le c$), where c is the number of the rule. x_j refers to the *jth* input, and ($1 \le j \le q$), where q is the number of input variables, while y_i refers to the *ith* output from the wavelet neural network (WNN). In the wavelet neural network, which is described in Fig. 2, the output can be calculated using equation (4).where w_i refers to the weight coefficients and ψ_{ij} represents a set of wavelet functions, which are called the wavelet family, and is defined computationally using equation (5).

$$\psi_{ij}(x_j) = \psi\left(\frac{x_j - b_{ij}}{a_{ij}}\right), \ a_{ij} \neq 0 \tag{5}$$

where a_{ij} and b_{ij} represent the dilation and translation parameters of the wavelet function, respectively, and ψ refers to the mother wavelet, which can be calculated using the Mexican hat wavelet function, which can be represented as in equation (6).

$$\psi(x) = \frac{1}{\sqrt{|a|}}(1 - 2x^2) \tag{6}$$

Layer 5: this layer integrates layer 3 with layer 4. The output of this layer is calculated by multiplication of the outputs of the third layer with the outputs of the fourth layer using equation (7).

$$L_5(i) = \mu_i{}^* y_i \tag{7}$$

where $1 \le i \le c$ and c refers to numbers of fuzzy rules and wavelet functions.

Layer 6: the output of this layer involves two parts. The first part (L_6^a) can be calculated using a sum operation of the outputs of the fifth layer, which can be represented using equation (8)

$$L_6^a = \sum_{i=1}^{c} \mu_i{}^* y_i \tag{8}$$

The second part (L_6^b)can be calculated using a sum operation of the outputs of the third layer, which can be represented using equation (9)

$$L_6^b = \sum_{i=1}^{c} \mu_i \tag{9}$$

Layer 7: this layer acts as a defuzzifier, and the overall output of FWNN is calculated in this layer using equation (10)

$$u = \frac{\sum_{i=1}^{c} \mu_i y_i}{\sum_{i=1}^{c} \mu_i} \tag{10}$$

4. A brief overview of functional link neural network (FLNN)

FLNN has been conveniently used for function approximation, and it has also emerged as an important tool for solving classification problems due to its single-layer property, i.e., it does not have hidden layers [48]. Thus, FLNN has low computational complexity and fast training speed due to it optimizing fewer weight parameters [21,48,49]. In FLNN, the values of the input variables are mapped to a higher dimensional space by functional expansion using trigonometric functions [50,51], which enhances the input representation and achieves linear separability in the extended space [48]. For instance, if there are two input variable samples $x = [x_1, x_2]^T$, the result of functional expansion using trigonometric functions will be

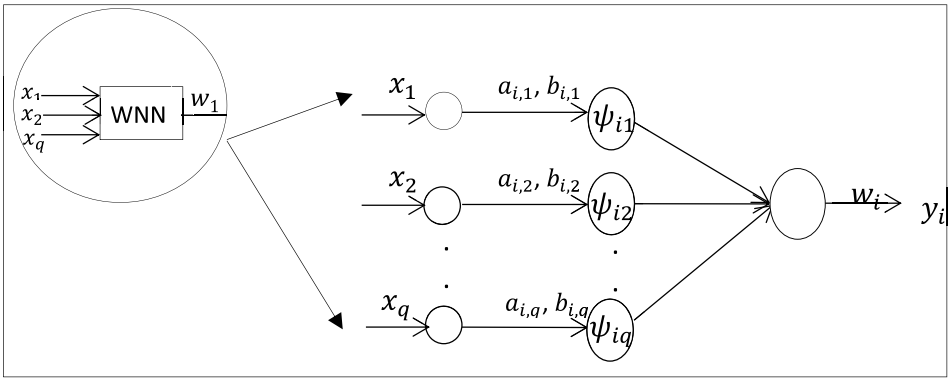


Fig. 2. The structure of a Wavelet Neural Network [18].

$$y_i = w_i \sum_{j=1}^{q} \psi_{ij}(x_j) \tag{4}$$

$\varphi = [x_1, \sin(pi*x_1), \cos(pi*x_1), x_2, \sin(pi*x_2), \cos(pi*x_2)]^T$. Therefore, the number of the basis function (φ) will be 3*N, where N represents the number of input variables. The local output of the FLNN, y_j, is given by using equation (11) [48].

$$y_j = \sum_{k=1}^{M} w_{kj}\varphi_k(x) \tag{11}$$

where x is the input variable and w_{kj} represents the weight that connects the basis function number k (φ_k) with the jth output node of the FLNN. M represents the total number of the basis function.

5. Teaching learning-based optimization algorithm (TLBO)

The teaching learning-based optimization algorithm (TLBO), proposed by Ref. [31]; is a learner teaching process-inspired algorithm [31]. This optimization algorithm depends mainly on two vital components, which are the teacher and the learners, where the learners can get knowledge via the teacher, who is usually considered to be a highly learned person (teacher phase), and via interacting with other learners (learner phase) [31,47,52,53].

In the TLBO algorithm, the group of n learners in a class is considered a population. Each learner represents a solution for the given optimization problem and the number of dimensions (D) of each solution, which are considered as different subjects offered to the learners and actually represent the parameters involved in the objective function of the given optimization problem. The learners' grades are considered to be a fitness value of the optimization problem, and the best solution, which has the best fitness value, is considered to be a teacher [47,52–55].

The attractive characteristic of the TLBO algorithm is that it is a simple mathematical model because it does not contain any specific parameter [56,57], and it is one of the most powerful tools for finding the optimal solution in a shorter computational time period [26].

The implementation of the TLBO algorithm is divided into two phases, namely, the "teacher phase" and "learner phase." The implementation of these two phases will be explained in the following subsection [31,47,58]:

1) Define the optimization problem and initialize the common parameters, which are population size (*ps*), which represents the number of learners (n), and the subjects (m) offered to the learners. In addition, set the value of the maximum number of iterations and the values of the constraints variables (*lb*, *ub*), which denote lower and upper boundaries, respectively.
2) Generate the initial population randomly with (n) rows and (m) columns within [*lb*, *ub*] and then calculate the objective function value of each solution using f(x), where x = 1, 2, 3, …, n. The results are sorted in ascending order corresponding to population size (*ps*), as shown in matrix (1) (ascending order is convenient for finding the minimum value; the maximum value can be obtained by multiplying by −1 before the objective).

$$\text{Ps}=\begin{bmatrix} A_1^1 & A_2^1 & \cdots & A_{m-1}^1 & A_m^1 \\ A_1^2 & A_2^2 & \cdots & A_{m-1}^2 & A_m^2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ A_1^{n-1} & A_1^{n-1} & \cdots & A_{m-1}^{n-1} & A_m^{n-1} \\ A_1^n & A_2^n & \cdots & A_{m-1}^n & A_m^n \end{bmatrix} \begin{matrix} \rightarrow \\ \rightarrow \\ \\ \rightarrow \\ \rightarrow \end{matrix} \begin{matrix} f(1) \\ f(2) \\ \vdots \\ f(n-1) \\ f(n) \end{matrix} \tag{1}$$

where f(1) < f(2) … < f(n-1) < f(n). Therefore, the first learner $A^1 = (A_1^1 A_2^1 \ldots A_m^1)$ is considered to be the best solution (teacher).

3) In the teacher phase, calculate the mean of the population column-wise where:

$$A^{mean} = \left\{ mean\left(\sum_{i=1}^{n} A_1^i\right) mean\left(\sum_{i=1}^{n} A_2^i\right) \ldots mean\left(\sum_{i=1}^{n} A_m^i\right) \right\} \tag{12}$$

4) The teacher tries to improve the average grade of the students using:

$$A^{new,i} = A^i r_i (A^1 - T_F A^{mean})(i = 1,2,3,\ldots,n) \tag{13}$$

where $A^{new,i}$ represents the improved learners, A^i represents the current learners, r_i is a random number in the interval [0,1], A^1 is the desired mean, A^{mean} is the current mean [59], and T_F is a teaching factor that is not a parameter of the TLBO algorithm: it is calculated randomly using equation (14), which decides the value of the mean to be changed [53].

$$T_F = round[1 + rand(1)] \tag{14}$$

In $A^{new,i}$, if the value of any variable is less than *lb* or greater than *ub*, it is equal to *lb* or *ub*, respectively [60].

If the objective function value of the new solution $f(A^{new,i})$ is better than the objective function value of the current solution $f(A^i)$, then the current solution is replaced by the new solution using equation (15).

$$If\ f(A^{new,i}) < f(A^i),\ then\ A^i = A^{new,i}\ else\ A^i = A^i \tag{15}$$

5) In the learner phase, a learner interacts randomly with other learners to enhance his or her knowledge.
6) Select two learners randomly, such as A^i *and* A^j, where $i \neq j$.

$$A^{new,i} = A^i + r_i(A^i - A^j)\ if\ (A^i) < f(A^j) \tag{16}$$

$$A^{new,i} = A^i + r_i(A^j - A^i)\ if\ f(A^j) < f(A^j) \tag{17}$$

In $A^{new,i}$, if the value of any variable is less than *lb* or greater than *ub*, it is equal to *lb* or *ub*, respectively.

If the objective function value of the new solution $f(A^{new,i})$ is better than the objective function value of the current solution $f(A^i)$, then the current solution is replaced by the new solution using equation (15).

7) Duplicate solutions are modified to avoid becoming trapped in local optima by using a mutation process on randomly selected dimensions of the duplicate solutions before executing the next iteration.
8) Sort the results in ascending order corresponding to (*ps*).
9) Repeat (3) through (5) until the termination condition is satisfied.

6. The proposed method of TLBO-FFWNN

The learning process of FFWNN is achieved by iteratively adjusting the FFWNN parameters. Due to the previously mentioned characteristics of the TLBO algorithm, this algorithm will be utilized for updating the linkage weights, dilation parameters and translation parameters of the FFWNN for five different datasets, which are Breast Cancer, Heart, Hepatitis, Pima-Indian diabetes and Appendicitis. The structure of FFWNN is predefined according to each type of dataset. The TLBO algorithm will find the optimal values of FFWNN parameters for the fixed structure.

The fuzzy rule for the proposed method FFWNN_TLBO can be represented in equation (18):

$$\begin{aligned} &R^i: if\ x1\ is\ A_1^i,\ and\ x2\ is\ A_2^i \ldots .\ and\ xq\ is\ A_q^i \\ &then\ u_i = w_i \sum_{j=1}^{q} \psi_{ij}(x_j) * \sum_{k=1}^{m} w_{kj}\,\varphi_k(x) \end{aligned} \tag{18}$$

where the consequent part of the fuzzy rule is the result of multiplying the output of the wavelet function (equation (4)) with the output of FLNN (equation (11)), and u_i represents the output of layer 5.

To conduct the training process, a sample of data, which is called training data, is used as the input variables to the FFWNN. Then, the mean square error (MSE), which represents the objective function of the TLBO algorithm and the output of the FFWNN, is calculated using equation (19) [61].

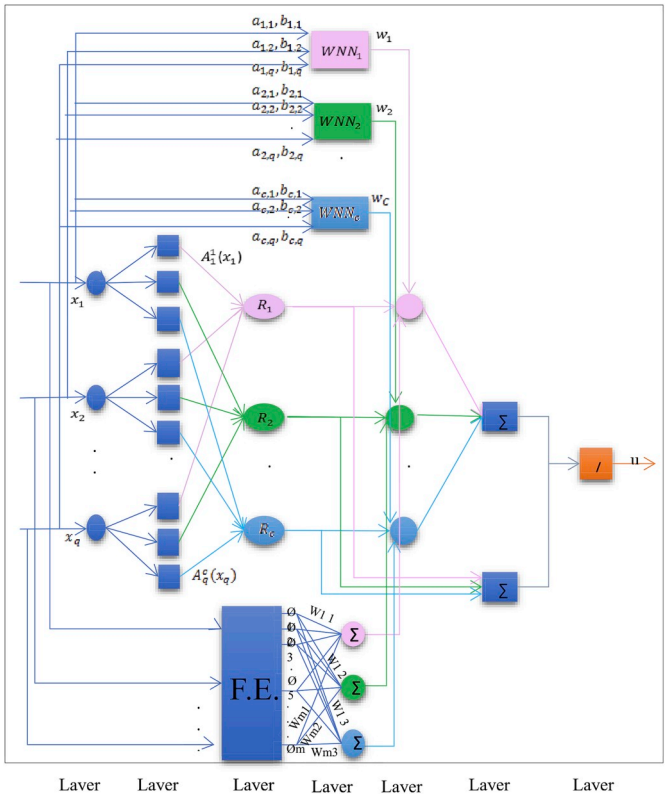


Fig. 3. The general structure of the proposed hybrid Functional Fuzzy Wavelet Neural Network (FFWNN).

$$MSE = \frac{1}{2PN_O}\sum_{K=1}^{N_O}\sum_{j=1}^{p}(t_j - O_j)^2 \quad (19)$$

where $1 \le p \le P$ represents the number of input patterns, N_O refers to the number of nodes in the out put layer, t_j represents the desired output, and O_j represents the actual output of the FFWNN.

This function is used to calculate the fitness value for each individual in the population. The population in the TLBO algorithm consists of multiple solutions, and each solution represents one learner. Each learner contains a group of values, and the number of these values is determined based on the number of FFWNN parameters that must be adjusted. In this study, the parameters are weights for both WNN (w_{WNN})and FLNN (w_{FLNN}), dilation and translation. The number of parameters depends on the network structure for each dataset, which are shown in the following table:

Each learner is initialized randomly and then updated in each iteration to get the optimal solution (learner with minimum MSE). This optimal solution will be used in the testing process with the testing dataset as an input to the trained network. The actual outputs of the network will be compared with the desired outputs to find the ability of FFWNN to diagnose diseases.

Fig. 3 represents the general structure of the proposed hybrid Functional Fuzzy Wavelet Neural Network (FFWNN).

The main steps of training FFWNN by using the TLBO algorithm are:

1) Set the values of the selected dataset as inputs of FFWNN.
2) Initialize each learner (weights, dilation parameters, translation parameters) randomly within [-1,1]. Then, initialize the common parameters of the TLBO algorithm, which are the population size, dimension and maximum iteration number.
3) Let Cycle = 1.
4) Calculate the fitness value of each learner based on a predefined fitness function based on equation (19).

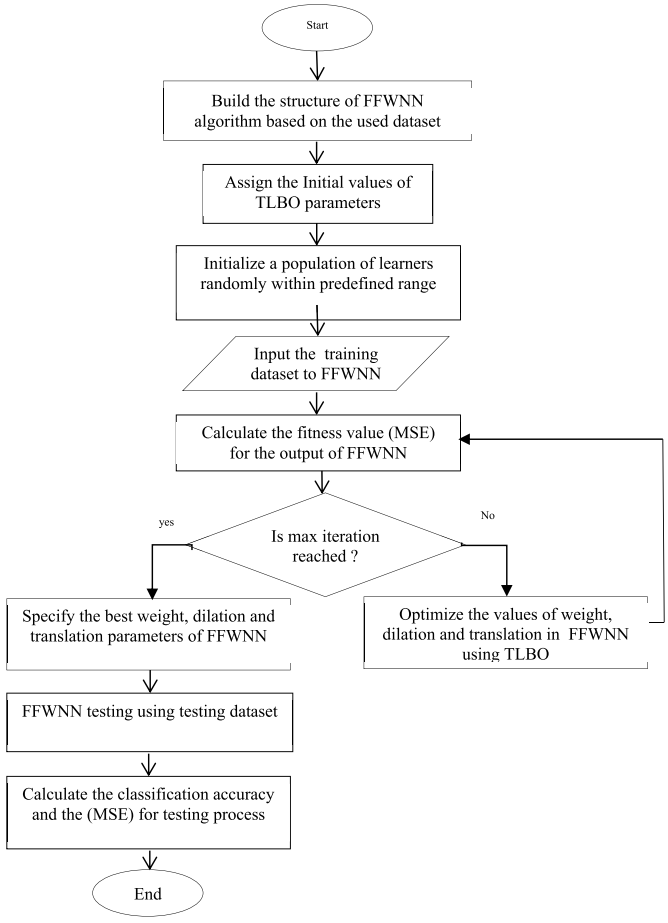


Fig. 4. Flowchart of the proposed FFWNN-TLBO method.

Table 1
The number of parameters for FFWNN that must be updated.

Dataset	w_{WNN}	w_{FLNN}	Dilation parameter	Translation parameter	Total number
Breast Cancer	3	81	27	27	138
Heart	3	117	39	39	198
Hepatitis	3	171	57	57	288
Pima-Indian diabetes	3	72	24	24	123
Appendicitis	3	63	21	21	108

5) Update the weight, dilation, and translation parameters using the TLBO algorithm.
6) Choose the best learner with best fitness value as a global solution, which represents the best values of weight, dilation, and translation parameters.
7) Let Cycle = Cycle +1.
8) Repeat steps (3) through (6) until the prespecified maximum iteration number is reached. Fig. 4 shows the flowchart of the proposed FFWNN-TLBO method.

7. Accuracy, complexity, sensitivity, specificity, and kappa analysis

The computational efficiency is based on several factors, which are computational cost, memory requirements and structural cost. The cost of learning strategies considers the most affected factor in the computational cost. Additionally, the cost of learning strategies can be affected by input parameter adaption, antecedent parameter adjusting fuzzy

Table 2
A summary of the datasets.

Dataset	No. of instances	No. of features	No. of classes	Data type
Breast Cancer	683	9	2	Numerical
Heart	297	13	2	Numerical & categorical
Hepatitis	80	19	2	Numerical & categorical
Pima-Indian diabetes	768	8	2	Numerical
Appendicitis	106	7	2	Numerical

Table 3
Numbers of patterns in both training and testing with both 5-CV and 10-CV.

Dataset	5-CV		10-CV	
	Training	Testing	Training	Testing
Breast Cancer	546	137	614	69
Heart	237	60	267	30
Hepatitis	64	16	72	8
Pima-Indian diabetes	614	154	691	77
Appendicitis	84	22	95	11

rule growing, consequent parameter adjusting, rule recruitment and rule pruning, which are based on the number of rules, number of data samples and input features [62,63]. The complexity load of the proposed (TLBO-FFWNN) is influenced by input parameter adaption, antecedent parameter adaption and consequent parameter adaption, which are of the order O(q(u*r), O(0) and O(m + (2u*r) + r), respectively, where m = 3u*r, r represents the number of fuzzy rules, u represents the number of features of the training samples, and q is the number of data samples. Hence, the total computational load of (TLBO-FFWNN) is of the order O(q(u*r) + m + (2u*r) + r). In addition, the memory requirement of (TLBO-FFWNN) is of the order O(m + (2u*r) + (u*r) + r + q). Moreover, the overall adjusting parameters, which are stored in the memory, represent the structural cost. Accordingly, the structure cost of (TLBO-FFWNN) is of the order O(m + (2u*r) + r).

Additionally, **Accuracy,** sensitivity, specificity, and kappa are common metrics that are utilized to measure the performance of the classifier. These metrics are calculated as follows [6,38]:

$$\text{Classification accuracy} = (1\text{-MSE})*100 \tag{20}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}\% \tag{21}$$

$$\text{Specificity} = \frac{TN}{TN + FP}\% \tag{22}$$

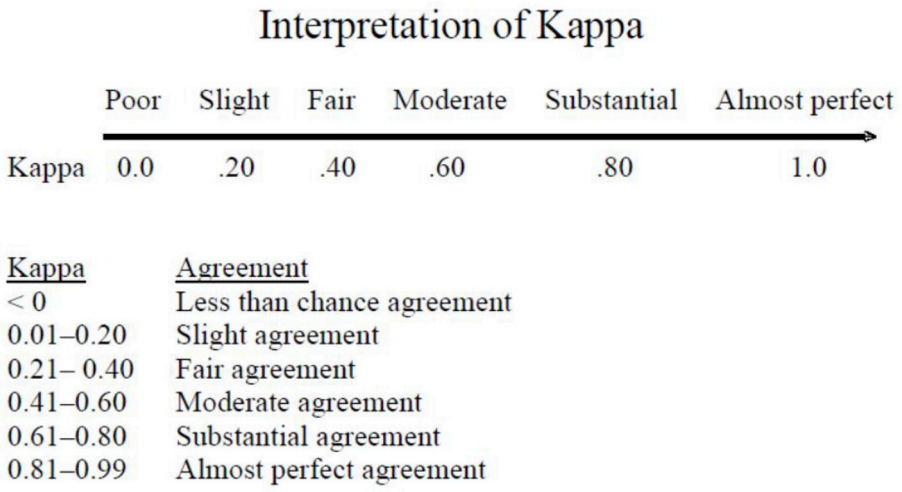
where TP, which represents True Positive, and TN, which represents True Negative, are classified correctly as (disease or no disease), respectively. Additionally, FP, which represents False Positive, and FN, which represents False Negative, are classified incorrectly as (disease or no disease), respectively.

Kappa is a measure of "true" agreement as shown in the interpretation of Kappa [64], which is calculated using equation (23)

$$K = \frac{OA - CA}{1 - CA} \tag{23}$$

where:

OA is the relative observed agreement among raters.
CA is the hypothetical probability of chance agreement



8. Experimental results and discussion

To investigate the performance of the new hybrid TLBO-FFWNN, five medical datasets, Breast Cancer, Heart Disease, Hepatitis, Pima-Indian diabetes and Appendicitis, are utilized (see Table 1). These datasets include two mixed types of data: numerical and categorical, with different numbers of patterns and input variables. All these datasets were obtained from the UCI machine learning repository [65] except the Appendicitis dataset, which was obtained from the KEEL dataset repository [66,67]. Table 2 presents a summary of these datasets.

Breast cancer is an irregular growth of cells that appears in the breast tissue. These cells may divide quickly and create masses called tumors. There are two types of tumors: noncancerous (benign) and cancerous (malignant). A malignant tumor damages healthy tissues. The naming of breast cancer came from the malignant tumor that originated from breast cells. One of the most common causes of cancer deaths among women is breast cancer [1,68], especially among women

Table 4
Classification accuracy and the error rate of TLBO-FFWNN for all datasets with 5-CV.

Datasets		Exp.1		Exp.2		Exp.3		Exp.4		Exp.5		Avg.	
		Error rate	Acc.	Error rate	Acc.	Error rate	Acc.	Error rate	Acc.	Error rate	Acc.	Error rate	Acc.
Breast Cancer (ps = 50)	train	0.015	98.48	0.013	98.62	0.014	98.55	0.011	98.81	0.009	99.05	0.012	98.70
	test	0.004	99.58	0.011	98.91	0.017	98.22	0.021	97.85	0.034	96.51	0.017	98.21
Breast Cancer (ps = 100)	train	0.016	98.37	0.013	98.66	0.01	98.97	0.011	94.28	0.009	99.05	0.011	97.86
	test	0.004	99.57	0.012	98.78	0.018	98.16	0.021	97.90	0.032	96.74	0.017	98.23
Heart (ps = 50)	train	0.051	94.94	0.046	95.32	0.054	94.54	0.052	94.72	0.058	94.17	0.052	94.73
	test	0.221	77.87	0.069	93.01	0.074	92.53	0.079	92.05	0.071	92.88	0.102	89.66
Heart (ps = 100)	train	0.049	95.02	0.053	94.66	0.048	95.12	0.05	94.99	0.051	94.83	0.050	94.92
	test	0.218	78.13	0.067	93.24	0.079	92.01	0.084	91.52	0.073	92.62	0.104	89.50
Hepatitis (ps = 50)	train	0.003	99.61	0.011	98.85	0.009	99.09	0.009	99.07	0.014	98.54	0.009	99.03
	test	0.096	90.39	0.177	82.26	0.039	96.08	0.048	95.11	0.007	99.21	0.073	92.61
Hepatitis (ps = 100)	train	0.003	99.67	0.004	99.53	0.013	98.60	0.006	99.34	0.011	98.83	0.007	99.19
	test	0.108	89.18	0.275	72.48	0.022	97.77	0.076	92.35	0.021	97.85	0.100	89.92
Pima-Indian diabetes (ps = 50)	train	0.078	92.17	0.073	92.66	0.073	92.62	0.084	91.53	0.078	92.11	0.077	92.21
	test	0.1	89.95	0.09	90.84	0.16	83.48	0.11	88.54	0.081	91.89	0.108	88.94
Pima-Indian diabetes (ps = 100)	train	0.077	92.29	0.072	92.7	0.073	92.62	0.073	92.65	0.074	92.54	0.073	92.56
	test	0.09	90.62	0.086	91.35	0.12	87.2	0.082	91.72	0.092	90.77	0.094	90.33
Appendicitis (ps = 50)	train	0.017	98.21	0.026	97.32	0.022	97.70	0.034	96.58	0.027	97.25	0.025	97.41
	test	0.058	94.12	0.096	90.36	0.078	92.18	0.054	94.51	0.136	86.36	0.084	91.5
Appendicitis (ps = 100)	train	0.026	97.37	0.021	97.85	0.026	97.35	0.034	96.59	0.027	97.26	0.026	97.28
	test	0.051	94.8	0.090	90.95	0.059	94.07	0.047	95.29	0.063	93.61	0.062	93.74

between 40 and 55 years old. Breast cancer is considered the first cause of death among women after lung cancer [1]. The breast cancer dataset was generated by Ref. [69]. The aim of this dataset is to predict whether a tumor is malignant or benign. This dataset contains 238 instances with a benign tumor and 445 instances with a malignant tumor.

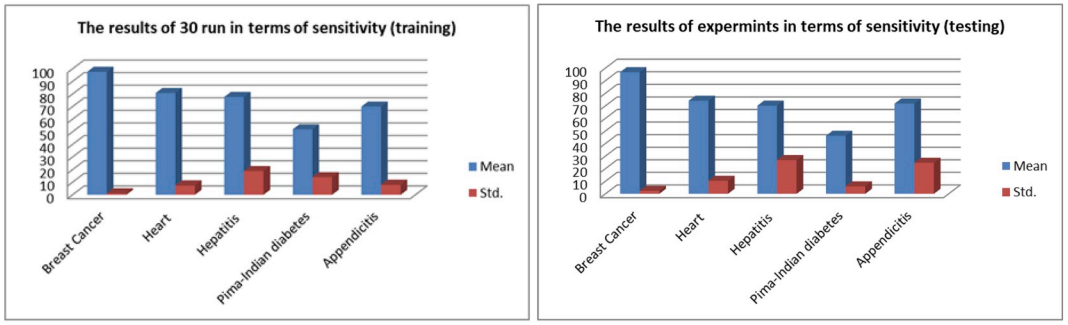
Heart disease is a term that refers to any disturbance that makes the heart function abnormally [47,70]. When the coronary arteries, which are responsible for supplying oxidized blood to the heart, are narrowed or blocked, the blood flow to the myocardium is decreased. This represents the main reason for the emergence of heart disease in humans. There are several factors that may increase the risk of this disease, including diabetes, smoking, obesity, a family history of heart disease, high cholesterol, and high blood pressure [39,70,71]. The Cleveland Heart dataset was collected by Detrano (1988). This dataset is used to predict the presence or the absence of heart disease. It contains 160 normal instances and 137 abnormal instances.

Table 5
The results of 30 runs in terms of accuracy.

Experiments	Training					Testing				
	Breast cancer	Heart	hepatitis	diabetes	appendicitis	Breast cancer	Heart	hepatitis	diabetes	appendicitis
1	97.96	90.66	97.15	90.69	96.04	97.77	91.5597	88.03	88.14	91.04
2	97.77	87.4486	96.004	87.37	95.06	98.49	94.02	94.81	84.57	96.69
3	98.23	88.6888	95.92	87.99	96.08	99.46	90.58	89.4	89.49	89.56
4	98.01	89.5939	96.72	89.48	96.63	99.13	86.69	94.64	90.6	91.53
5	98.22	90.0428	96.01	88.47	94.28	99.15	89.78	98.34	88.76	97.57
6	97.87	89.117	96.88	88.99	96.63	99.19	89.97	93.44	89.81	84.36
7	97.64	88.3912	96.96	92.4	94.68	98.37	92.04	93.17	92.06	93.75
8	97.63	87.217	97.65	92.09	94.5	98.43	92.22	87.72	88.54	96.08
9	97.84	94.8472	99.19	87.13	94.49	96.02	92.08	87.59	83.97	95.44
10	98.57	86.9854	98.98	92.6	95.23	98.24	91.84	89.91	86.95	99.27
11	98.21	90.95	96.78	91.3	95.96	97.21	92.35	87.91	88.56	90.1
12	97.52	88.12	95.84	87.21	95.36	98.11	93.75	95.5	84.1	94.25
13	98.11	89.23	96.17	88.1	95.25	98.15	89	90.1	86.45	91.23
14	97.99	90.1	96.15	89.78	97.32	99.23	87.41	95.11	88.6	93.56
15	98.5	89.9	95.99	88.15	93.19	99.69	90	99.24	90.2	95.5
16	98.12	89	98.1	88.56	96.22	99.2	89.35	93.45	88.52	86.73
17	96.99	87.45	95.21	92.4	94.36	97.89	92.92	92.15	93.41	93.75
18	97.52	88.1	97.15	91.98	94.25	98.65	91.58	88.21	89.41	94.67
19	98.02	95.12	99.02	87.12	95.2	96.03	92.1	88.5	85.97	96.5
20	98.23	87.12	98.76	92.24	95.71	98.64	92.23	90.1	86.95	98.57
21	96.41	88.98	98.54	89.41	95.99	98.71	89.35	87.15	89.15	91.04
22	98.59	88.68	96.91	88.15	93.65	97.56	94.11	93.4	85.21	96.69
23	97.15	85.37	94.65	86.44	97.35	98.86	90.41	92.65	90.41	89.56
24	96.45	90.21	96.81	90.15	97.65	96.41	88.26	92.67	89.97	91.53
25	99.21	89.62	95.99	88.47	95.31	99.75	89.68	95.52	88.62	97.57
26	98.65	89.57	96.54	88.99	97.67	98.51	87.52	90.89	90.61	84.36
27	97.12	88.57	97.63	90.95	93.97	98.53	94.68	90.65	90.5	93.75
28	96.13	87.81	96.36	89.95	95.78	96.51	93.54	86.24	88.26	96.08
29	95.89	95.24	98.75	89.57	95.83	98.65	94.36	88.23	83.97	95.44
30	99.19	87.65	98.67	92.6	94.17	98.73	89.65	87.12	98.57	99.27
Mean	97.7913	89.3260	97.0494	89.6243	95.460333	98.309	91.1009	91.3946	88.6776	93.514666
Std.	0.81718	2.30432	1.22599	1.90511	1.1859594	1.01391	2.18791	3.48492	3.02353	3.9925186

Table 6
The results of 30 runs in terms of sensitivity.

Experiments	Training					Testing				
	Breast cancer	Heart	hepatitis	diabetes	appendicitis	Breast cancer	Heart	hepatitis	diabetes	appendicitis
1	96.8	84.21	66.66	64.66	84.61	93.13	62.5	66.66	44.44	50
2	97.67	78.94	66.66	35.41	71.42	100	76.92	100	40	100
3	97.68	74.66	66.66	27.95	70.58	97.77	61.53	33.33	50	50
4	97.62	80.3	71.42	43.13	81.25	95.55	75	100	44.44	50
5	98.94	89.65	80	48.88	57.14	97.72	81.81	100	50	66.66
6	97.9	68.13	100	44.68	72.22	97.82	70	50	48.48	50
7	95.5	92.59	66.66	66.86	63.15	97.72	72.72	50	57.14	100
8	98.4	76.92	80	62.5	64.28	97.77	90.9	50	44.11	100
9	99.73	87.27	100	61.36	66.66	93.33	58.82	100	36.11	50
10	96.73	75.4	100	64	66.66	100	75	50	45.45	100
11	96.1	85.21	66.66	65.45	83.42	93.45	61.89	66.66	44.32	50.68
12	96.9	78.65	66.75	35.62	73.14	99.87	77.65	99.73	41.56	99.363
13	97.63	74.42	66.66	28.01	69.15	96.78	62.45	33.53	50	50
14	96.91	82.51	72.12	42.92	81.46	96.13	75.99	99.76	43.91	51.18
15	99	90.21	79.92	48.95	57.47	98.12	83.45	99.98	50.13	66.63
16	98.1	71.2	99.98	44.32	72.87	97.24	71.19	51.12	48.12	52.13
17	95.11	94.56	66.54	67.45	63.98	97.35	73.14	52.18	57.21	100
18	98.4	75.92	82.51	62.98	65.84	98.23	91.24	50	45.36	99.97
19	98.92	88.12	100	60.45	66.98	93.12	58.14	100	36.12	49.98
20	96.43	76.89	99.97	63.78	66.66	100	77.56	51.24	46.53	100
21	97.13	83.45	65.73	65	84.39	93.24	93.56	65.73	43.56	49.89
22	98.17	78.94	66.66	35.97	71.89	99.63	77.52	100	39.99	100
23	97.95	73.89	67.31	26	70.45	96.48	66	34.09	51.23	50.47
24	97.87	81.56	7.93	45.37	80.79	95.13	78.02	100	44.91	49.34
25	98.67	88.63	79.58	49.51	58.47	96.48	81.67	99.89	52.02	67.01
26	97.25	69.76	99.34	44.71	73.45	98.73	69.41	50	48.69	50.76
27	96.32	90.45	67	67.11	62.91	94.62	73.14	49.76	58.24	99.91
28	99.89	74.12	79.76	62.45	66.37	96.78	92.83	50.64	43.91	99.74
29	98.24	86.45	99.47	61.75	66.66	91.87	59.43	99.96	35.91	51.13
30	95.68	73.68	99.87	65.34	67.02	99.94	73.21	49.76	45.15	99.93
Mean	97.588	80.8896	77.7273	52.0856	70.0446666	96.8	74.0896	70.134	46.2346	71.82577
Std.	1.19253	7.21472	19.354	13.4920	7.72642743	2.40436	10.1245	25.8535	5.76069	23.81917



(a) training (b) testing

Fig. 5. (a,b). The experiment results in terms of Sensitivity.

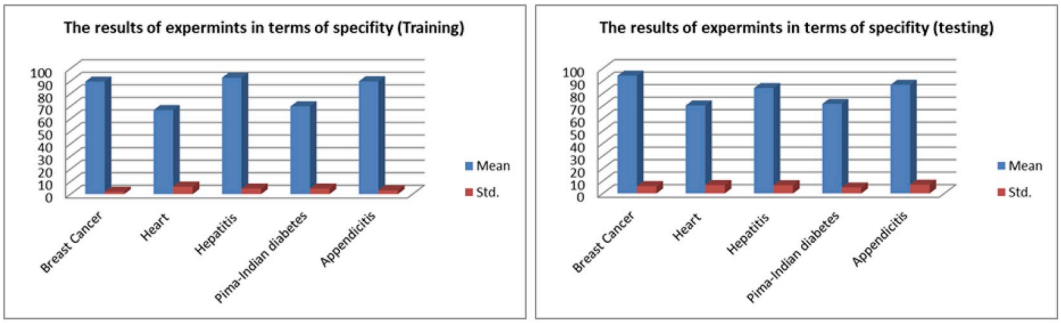
Hepatitis is an inflammation of the liver, and the most common causes of this disease in the world are liver viruses. Scientists have discovered five unique viruses that cause hepatitis: Hepatitis A, B, C, D and E, all of which lead to liver disease, but there are great differences between them [72,73]. There are several major risk factors for hepatitis: transfusion of contaminated blood, tattoos and piercing, drug abuse, and sexual contact [72]. The objective of the Hepatitis dataset [74] is to predict whether a patient with hepatitis will live or die. This dataset contains 80 instances, where 67 instances of them are in the LIVE class and 13 instances are in the DIE class.

The term diabetes includes a number of metabolic disorders that lead to increased blood sugar concentration [75,76]. Diabetes occurs when the body fails to produce insulin properly, which is necessary to regulate glucose [77]. There are several factors that cause diabetes, the most prominent being genetic and environmental factors [75]. However, this disease can be controlled by healthy nutrition, exercise programs and weight loss [77]. The Pima-Indian diabetes dataset [65] is used to diagnose whether a person has diabetes or not. This dataset contains 268 diabetes positive class examples and 500 diabetes negative class examples.

Appendicitis is an acute inflammation of the appendix, which is the most common abdominal surgical emergency [78]; Petroianu 2012). In typical cases, the pain of appendicitis appears in the lower right area of the abdomen [78]. Appendicitis can occur at any age, although it rarely occurs in elderly people [78,79]. Moreover, the incidence of this disease in men is higher than that in women [79]. The Appendicitis dataset [66,67] includes 106 patients, where 21 patients have appendicitis and 85 patients do not. To make the results of the testing process more reliable, K-Fold Cross-Validation [80] is used because it guarantees that all data are used for training and testing. According to this method, each dataset in this study is divided equally into 5 and 10 folds. Then, one of them will be selected for testing, and the rest will be utilized for

Table 7
The results of 30 runs in terms of Specificity.

Experiments	Training					Testing				
	Breast cancer	Heart	hepatitis	diabetes	appendicitis	Breast cancer	Heart	hepatitis	diabetes	appendicitis
1	88.5	68.58	93.65	72.22	91.46	92	63.63	80	67.79	77.77
2	90.74	62.85	86.95	65.15	90.12	96	76.47	85.71	69.04	88.88
3	91.11	64.58	89.39	64.04	92.3	100	70.58	80	68.85	77.77
4	86.8	66.16	90.76	66.55	92.4	100	61.11	85.71	70	88.88
5	89.74	66.98	88.05	68..94	82.95	96	63.15	100	82.85	100
6	89.22	65.34	91.04	66.66	93.5	100	65	83.33	75	77.77
7	91.5	65.72	90.9	75.47	90.78	96	68.42	75	69.84	90
8	87.44	61.39	95.16	74.56	87.65	95.83	78.94	83.33	72.09	90
9	88.84	82.16	100	66.92	88.75	83.33	76.92	85.71	65.85	88.88
10	93.51	62.13	98.38	75	88.75	88.88	77.77	83.33	72.72	90
11	88.1	67.35	92.85	73.42	92.34	91.66	64.51	79.99	68.14	78
12	90.64	63.14	87.54	66.14	89.75	95.32	77	86	70.5	88.61
13	89.96	65.49	90.27	63.89	93.15	99.89	71.13	79.99	68.1	78
14	86.32	67.57	91.13	66.43	91.46	100	62.25	86.02	69.57	89.23
15	90.02	67.08	87.63	69.02	83.51	95.45	63.78	99.9	83.4	99.9
16	89.35	66.13	91	66.25	92.51	100	65.9	84.2	74.89	77.62
17	89.46	67.45	91.87	76.33	89.41	95.87	69.41	74.89	70	89.81
18	93.2	62.73	96.81	74.23	88.06	96.51	79.51	83.19	72.98	89.73
19	87.14	83.24	99.32	67.42	89.01	84.51	77.65	86.23	66	89
20	87.42	63.51	99.51	74.19	89.35	89.35	78	84.12	72.89	89.71
21	89.21	67.06	94.15	73.51	92.41	91.48	63.63	79.51	68.91	78.61
22	91	63.14	87.52	66.51	89.58	95.99	76.47	86.52	70.86	89.11
23	91.78	65.71	90.45	65.21	93.15	99.68	70.58	79.51	69.51	78.37
24	86.32	66.89	89.84	67	91.45	100	61.11	84.61	69.31	88.33
25	90.1	67.02	87.5	70	82.14	95.78	63.15	99.9	83.51	99.12
26	89.91	66.13	91.75	67.35	92.15	99.98	65	84.5	74.95	77.51
27	91.26	66.32	91.35	76.19	90.81	95.45	68.42	74.31	70.32	89.66
28	88	60.56	96.48	75.81	88.15	79.51	78.94	84.55	73.41	89.73
29	88.14	83.21	99.99	67.06	88.52	83.33	77.32	86.61	66.34	89.01
30	92.45	62.75	99.9	74.12	89.32	88.88	77.19	84.55	73.21	89.91
Mean	89.5726	66.9456	92.7046	69.8844	89.829666	94.2226	70.4313	84.374	71.6943	86.964
Std.	1.93598	5.77880	4.27710	4.25471	2.9137231	5.79251	6.56412	6.31927	4.65887	6.7667460



(a) training (b)testing

Fig. 6. (a,b). The experiment results in terms of Specificity.

training.

Table 3 shows a number of patterns that have been selected in both the training and testing process for each experiment with both 5-CV and 10-CV.

As mentioned previously, the TLBO algorithm has common parameters, which are population size (ps) and the dimension of each solution (*D*).

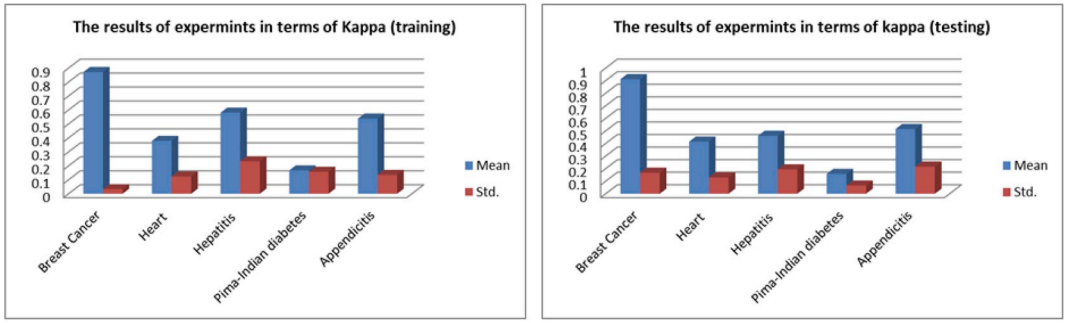
In this research, the value of the parameter *D*, which represents the weight, dilation and translation parameters of FFWNN, is varied for each dataset. For the Breast Cancer dataset, the value of the *D* parameter equals 138. For the Heart dataset, the value of the *D* parameter equals 198. The value of the *D* parameter equals 288 for the Hepatitis dataset. For the Pima-Indian diabetes dataset, the value of *D* equals 123. However, the value of the *D* parameter for the Appendicitis dataset equals 108.

In addition, the value of the (ps) parameter is varied because the user of this algorithm does not have adequate knowledge regarding the appropriate value of this parameter. In this study, the training process is repeated twice. First, it is repeated in five separate experiments for each dataset with two different (ps) values, which are 50 and 100. Second, the training process is repeated in ten separate experiments for each dataset with (ps = 50). The maximum iteration number, representing the stopping condition, is set to 500 for all experiments. Additionally, in FFWNN, the numbers of fuzzy sets, fuzzy rules, wavelet functions, and basis functions were each equal to three.

The maximum classification accuracy and the minimum error rate, which are obtained in training and testing FFWNN based on the TLBO algorithm on the Breast Cancer, Heart, Hepatitis, Pima-Indian diabetes and Appendicitis datasets with 5-CV, are demonstrated in Table 4. In Table 4, Exp. indicates the experiment, Acc. represents the classification accuracy value, and Avg. refers to the average value. From Table 4, some observations can be derived. In the case of the Breast Cancer

Table 8
The results of 30 runs in terms of Kappa.

Experiments	Training					Testing				
	Breast cancer	Heart	hepatitis	diabetes	appendicitis	Breast cancer	Heart	hepatitis	diabetes	appendicitis
1	0.86	0.44	0.57	0.28	0.65	0.84	0.21	0.46	0.1	0.2
2	0.89	0.294	0.23	0.002	0.5	0.96	0.52	0.6	0.09	0.74
3	0.89	0.3273	0.406	−0.04	0.61	0.96	0.32	0.14	0.15	0.23
4	0.85	0.3631	0.49	0.06	0.68	0.93	0.34	0.6	0.14	0.38
5	0.9	0.4068	0.41	0.14	0.22	0.93	0.41	1	0.31	0.74
6	0.88	0.3085	0.58	0.06	0.65	0.96	0.31	0.33	0.24	0.23
7	0.8	0.3949	0.44	0.37	0.53	0.93	0.38	0.25	0.19	0.62
8	0.87	0.2525	0.72	0.33	0.45	0.93	0.66	0.33	0.16	0.62
9	0.9	0.6814	0.94	0.09	0.5	0.77	0.34	0.6	0.01	0.38
10	0.9	0.2754	0.94	0.34	0.5	0.9	0.52	0.33	0.18	1
11	0.87	0.45	0.59	0.31	0.66	0.85	0.23	0.46	0.1	0.19
12	0.9	0.32	0.26	0.003	0.52	0.97	0.53	0.59	0.08	0.75
13	0.88	0.33	0.41	−0.041	0.62	0.96	0.36	0.15	0.16	0.24
14	0.86	0.37	0.5	0.05	0.69	0.94	0.35	0.59	0.13	0.39
15	0.88	0.41	0.42	0.15	0.24	0.95	0.42	0.99	0.3	0.76
16	0.87	0.31	0.59	0.07	0.69	0.96	0.33	0.34	0.23	0.25
17	0.79	0.4	0.46	0.38	0.54	0.94	0.39	0.27	0.18	0.61
18	0.88	0.27	0.75	0.34	0.46	0.93	0.68	0.34	0.15	0.59
19	0.91	0.69	0.95	0.08	0.51	0.78	0.35	0.61	0.02	0.36
20	0.89	0.28	0.95	0.35	0.52	0.95	0.55	0.32	0.16	0.99
21	0.85	0.43	0.58	0.29	0.66	0.83	0.23	0.45	0.12	0.19
22	0.88	0.3	0.24	0.003	0.51	0.95	0.58	0.59	0.08	0.75
23	0.87	0.33	0.41	−0.05	0.62	0.91	0.33	0.12	0.17	0.22
24	0.85	0.37	0.5	0.07	0.7	0.93	0.35	0.7	0.14	0.37
25	0.89	0.41	0.42	0.15	0.23	0.95	0.43	0.9	0.32	0.75
26	0.87	0.31	0.58	0.07	0.66	0.91	0.33	0.32	0.28	0.24
27	0.79	0.4	0.45	0.39	0.54	0.92	0.4	0.23	0.19	0.64
28	0.86	0.26	0.73	0.34	0.47	0.92	0.67	0.31	0.16	0.63
29	0.89	0.69	0.95	0.08	0.51	0.78	0.35	0.59	0.02	0.39
30	0.91	0.28	0.95	0.35	0.49	0.89	0.58	0.31	0.19	1
Mean	0.871	0.37846	0.58053	0.16723	0.5376666	0.911	0.415	0.46066	0.15833	0.515
Std.	0.03133	0.11916	0.22426	0.15192	0.1305078	0.05743	0.12984	0.23368	0.08030	0.261478



(a) training (b) testing

Fig. 7. (a,b). The experiment results in terms of kappa.

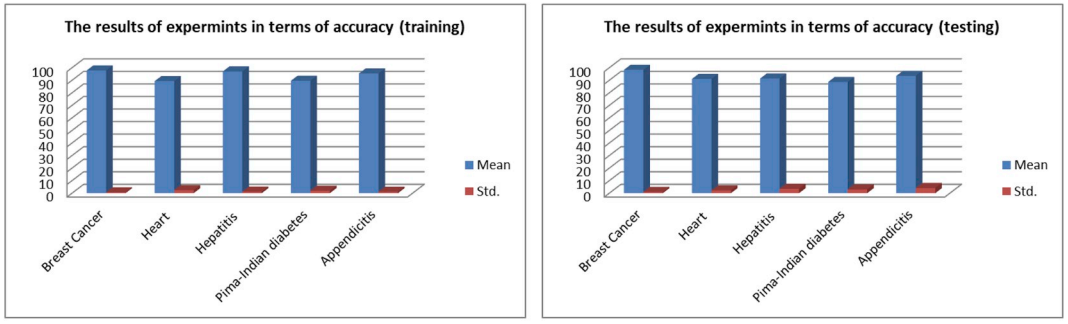
dataset, the average error rate of testing FFWNN in both population sizes of 50 and 100 is the same. However, the highest average classification accuracy for testing FFWNN (98.23) is obtained when the population size is equal to 100. In the case of the Heart dataset, the minimum average error rate of testing FFWNN with 0.102 is obtained when the population size is equal to 50, and the maximum average classification accuracy for testing FFWNN (89.66) is acquired when the population size is equal to 50. Moreover, in the case of the Hepatitis dataset, the minimum average testing error rate of 0.073 is achieved when the population size equals 50. Additionally, the highest average classification accuracy for testing FFWNN (92.61) is obtained when the population size is equal to 50. In addition, the minimum average testing error rate (0.094) and maximum classification accuracy (90.33) in the case of Pima-Indian Diabetes dataset are obtained when the population size equals 100. Finally, the minimum average testing error rate (0.062) and maximum classification accuracy (93.74) in the case of the Appendicitis dataset are reached when the population size equals 100.

In conclusion, the classification accuracy and the error rate are close to each other in both population sizes (50, 100). Additionally, increasing the population size leads to increasing the training duration time. Therefore, in 10-CV experiments, the value of the population size (ps) is set to 50.

For more accurate performance evaluation, all of the experiments are repeated 30 times based on 10-CV. In Table 5 and Fig. 8, TLBO-FFWNN provides the average maximum classification accuracy (mean) and standard deviation (std.): Breast Cancer produced 98.30% classification accuracy with 1.01 std., Heart produced 91.10% classification accuracy with 2.18 std., Hepatitis produced 91.39% classification accuracy with 3.48 std., Pima-Indian diabetes produced 88.67% with 3.02 std, and finally, Appendicitis produced 93.51% classification accuracy with 3.99 std. As seen in Table 6 and Fig. 5, the average sensitivity values for testing the proposed method on Breast Cancer, Heart,

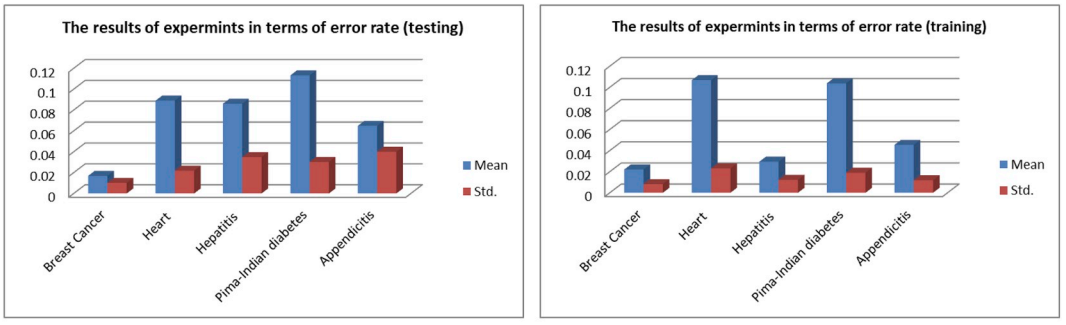
Table 9
The results of 30 runs in terms of error rate.

Experiments	Training					Testing				
	Breast cancer	Heart	hepatitis	diabetes	appendicitis	Breast cancer	Heart	hepatitis	diabetes	appendicitis
1	0.0204	0.0934	0.0285	0.0931	0.0396	0.0223	0.084403	0.1197	0.1186	0.0896
2	0.0223	0.125514	0.03996	0.1263	0.0494	0.0151	0.0598	0.0519	0.1543	0.0331
3	0.0177	0.113112	0.0408	0.1201	0.0392	0.0054	0.0942	0.106	0.1051	0.1044
4	0.0199	0.104061	0.0328	0.1052	0.0337	0.0087	0.1331	0.0536	0.094	0.0847
5	0.0178	0.099572	0.0399	0.1153	0.0572	0.0085	0.1022	0.0166	0.1124	0.0243
6	0.0213	0.10883	0.0312	0.1101	0.0337	0.0081	0.1003	0.0656	0.1019	0.1564
7	0.0236	0.116088	0.0304	0.076	0.0532	0.0163	0.0796	0.0683	0.0794	0.0625
8	0.0237	0.12783	0.0235	0.0791	0.055	0.0157	0.0778	0.1228	0.1146	0.0392
9	0.0216	0.051528	0.0081	0.1287	0.0551	0.0398	0.0792	0.1241	0.1603	0.0456
10	0.0143	0.130146	0.0102	0.074	0.0477	0.0176	0.0816	0.1009	0.1305	0.0073
11	0.0179	0.0905	0.0322	0.087	0.0404	0.0279	0.0765	0.1209	0.1144	0.099
12	0.0248	0.1188	0.0416	0.1279	0.0464	0.0189	0.0625	0.045	0.159	0.0575
13	0.0189	0.1077	0.0383	0.119	0.0475	0.0185	0.11	0.099	0.1355	0.0877
14	0.0201	0.099	0.0385	0.1022	0.0268	0.0077	0.1259	0.0489	0.114	0.0644
15	0.015	0.101	0.0401	0.1185	0.0681	0.0031	0.1	0.0076	0.098	0.045
16	0.0188	0.11	0.019	0.1144	0.0378	0.008	0.1065	0.0655	0.1148	0.1327
17	0.0301	0.1255	0.0479	0.076	0.0564	0.0211	0.0708	0.0785	0.0659	0.0625
18	0.0248	0.119	0.0285	0.0802	0.0575	0.0135	0.0842	0.1179	0.1059	0.0533
19	0.0198	0.0488	0.0098	0.1288	0.048	0.0397	0.079	0.115	0.1403	0.035
20	0.0177	0.1288	0.0124	0.0776	0.0429	0.0136	0.0777	0.099	0.1305	0.0143
21	0.0359	0.1102	0.0146	0.1059	0.0401	0.0129	0.1065	0.1285	0.1085	0.0896
22	0.0141	0.1132	0.0309	0.1185	0.0635	0.0244	0.0589	0.066	0.1479	0.0331
23	0.0285	0.1463	0.0535	0.1356	0.0265	0.0114	0.0959	0.0735	0.0959	0.1044
24	0.0355	0.0979	0.0319	0.0985	0.0235	0.0359	0.1174	0.0733	0.1003	0.0847
25	0.0079	0.1038	0.0401	0.1153	0.0469	0.0025	0.1032	0.0448	0.1138	0.0243
26	0.0135	0.1043	0.0346	0.1101	0.0233	0.0149	0.1248	0.0911	0.0939	0.1564
27	0.0288	0.1143	0.0237	0.0905	0.0603	0.0147	0.0532	0.0935	0.095	0.0625
28	0.0387	0.1219	0.0364	0.1005	0.0422	0.0349	0.0646	0.1376	0.1174	0.0392
29	0.0411	0.0476	0.0125	0.1043	0.0417	0.0135	0.0564	0.1177	0.1603	0.0456
30	0.0081	0.1235	0.0133	0.074	0.0583	0.0127	0.1035	0.1288	0.0143	0.0073
Mean	0.022086	0.10673	0.02950	0.10375	0.0453967	0.01691	0.08899	0.08605	0.11322	0.064853
Std.	0.008171	0.02304	0.01226	0.01905	0.0118596	0.01013	0.02187	0.03484	0.03023	0.039925



(a) training (b) testing

Fig. 8. (a,b). The experiments result in terms of accuracy.



(a) training (b) testing

Fig. 9. (a,b). The experiment results in terms of error rate.

Table 10
Number of rules, number of parameters, running time.

Datasets	Rule Number	Time$_{(training)}$ (*minutes*)	Time $_{(testing)}$ (*seconds*)	No. of parameters
Breast Cancer	3	89.87	0.0011	138
Heart	3	46.95	0.0015	198
Hepatitis	3	23.50	0.0013	288
Pima-Indian diabetes	3	99.62	0.0016	123
Appendicitis	3	18.13	0.0012	108

Hepatitis, Pima-Indian diabetes and Appendicitis are 96.8, 74.08, 70.13, 46.23 and 71.82, with standard deviations of 2.40, 10.12, 25.85, 5.76 and 23.81, respectively. Additionally, as shown in Table 7 and Fig. 6, the average specificity values for testing the proposed method on Breast Cancer, Heart, Hepatitis, Pima-Indian diabetes and Appendicitis are 94.22, 70.43, 84.37, 71.69 and 86.96 with standard deviations of 5.79, 6.56, 6.31, 4.65 and 6.76, respectively. In Table 8 and Fig. 7, the averages of kappa values for testing the proposed method on Breast Cancer, Heart, Hepatitis, Pima-Indian diabetes and Appendicitis are 0.91, 0.41, 0.46, 0.15 and 0.51 with standard deviations of 0.05, 0.12, 0.23, 0.08 and 0.26, respectively. Table 9 and Fig. 9 presents the minimum error rate that is obtained in training and testing the proposed TLBO-FFWNN method for all datasets:

In addition, the training and testing duration times of the TLBO-FFWNN method for all datasets are shown in Table 10.

For more clarity of the experiment results, a boxplot is used for all benchmark parameters in terms of Sensitivity, Specificity, Kappa, Accuracy and error rate. As shown in Fig. 10, the proposed method achieved the highest values for the breast cancer dataset in terms of Sensitivity, Specificity, Kappa, and Accuracy and the lowest value in terms of error rate. Alternately, the proposed method achieved the worst values for the Pima-India diabetes dataset in terms of Sensitivity, Specificity, Kappa, and Accuracy and the highest value in terms of error rate.

Moreover, to show the effectiveness of the proposed method, the performance of TLBO-FFWNN is compared with the performance of other classification methods from the literature entries that

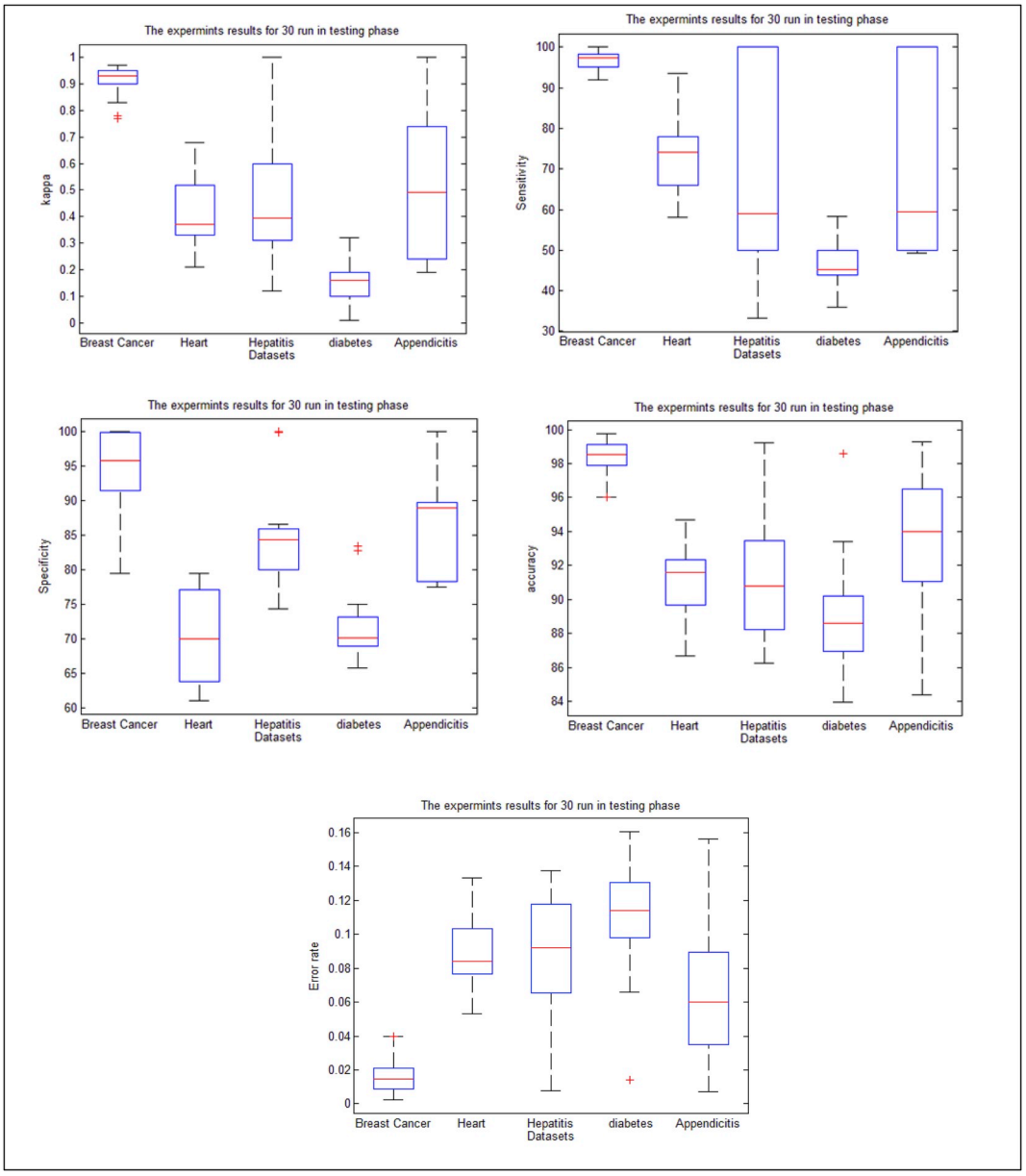


Fig. 10. Boxplot of all datasets for all metrics.

Table 11
A comparison of the proposed method and other classification methods based on average classification accuracy.

Methods	Breast Cancer	Heart	Hepatitis	Pima-Indian diabetes	Appendicitis
Proposed (TLBO-FFWNN)	98.30	91.10	91.39	88.67	93.51
TLBO-FWNN	/	90.29	/	/	/
neural network ensembles	/	89.01	/	/	/
ANN + FNN + BP	/	86.8	/	84.24	/
BSFP-PSO-FNFN-CUI	92.55	76.33	/	/	/
ISO-FLANN	/	/	75.72	/	/
CAPSO-MLP	/	81.85	71.29	72.99	/
PSO-MLP	/	71.11	68.71	72.40	/
GSA-MLP	/	56.67	67.74	56.43	/
ICA-MLP	/	66.67	64.25	64.61	/
NB	95.99	/	/	/	/
MLP	95.27	/	/	/	/
J48	95.13	/	/	/	/
SMO	96.99	/	/	/	/
IBK	94.56	/	/	/	/
SMO, IBK, NB, MLP	97.13	/	/	/	/
SMO, IBK, NB, J48	97.28	/	/	/	/
AR + NN	95.6	/	/	/	/
NN	95.2	/	/	/	/
V-ELM	/	/	/	/	87.80
V-ELM_AP	/	/	/	/	89.04
K-NN	97.6	/	/	77	89.4
MLP + BP	/	/	/	/	85.8
SMLP	/	/	/	/	88.2
SVM	/	/	/	/	88.1
LVQ	96.0 ± 2	82.8 ± 3.7	/	75.0 ± 2.7	87.8 ± 7.8
WLVQ	97.4 ± 1.3	83.2 ± 4.3	/	77.4 ± 4.3	88.7 ± 8.4
LION + SVM	98 ± 1.9	98.15 ± 1.7	/	83.72 ± 4.3	/
LION(M1) +SVM	98.29 ± 1.8	97.41 ± 2.7	/	83.60 ± 4.5	/
LION(M2) +SVM	97.71 ± 2.3	96.30 ± 1.5	/	82.94 ± 4.5	/
LION(M1 + M2) + SVM	98 ± 2.13	97.04 ± 2.1	/	84.24 ± 4.2	/

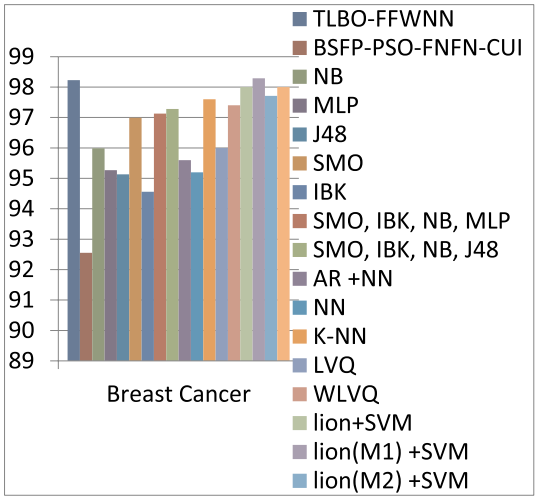


Fig. 11. Classification Accuracy on the Breast Cancer dataset.

experimented on the same datasets.

Table 11 compares the testing classification accuracy of the proposed method with that of other methods (TLBO-FWNN) [47]; neural network ensembles [39]; hybrid system of ANN and FNN and back propagation algorithm (ANN + FNN + BP) [38]; hybrid Particle Swarm Optimization and Black Stork Foraging Process for learning

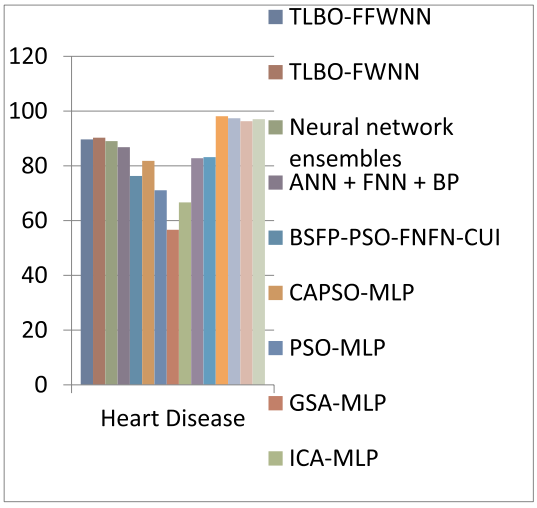


Fig. 12. Classification Accuracy on the Heart Disease dataset.

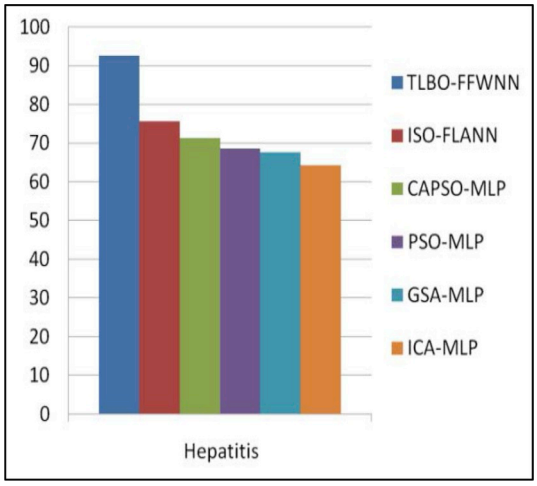


Fig. 13. Classification accuracy on the hepatitis dataset.

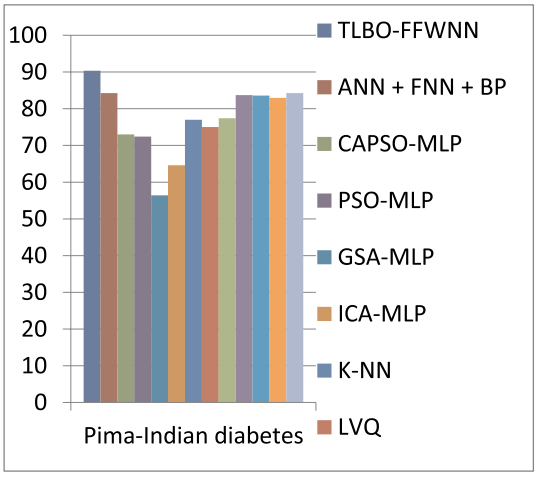


Fig. 14. Classification Accuracy on the Pima-Indian diabetes Dataset.

Functional Neural Fuzzy Network (BSFP-PSO- FNFN-CUI) [81]; Improved PSO and learning Functional Neural Fuzzy Network (ISO-FLANN) [50]; Centripetal Accelerated PSO and Multi-Layer Perceptron (CAPSO-MLP) [6]; (PSO-MLP) [6]; Gravitational Search Algorithm and MLP (GSA-MLP) [6]; Imperialist Competitive Algorithm with MLP (ICA-

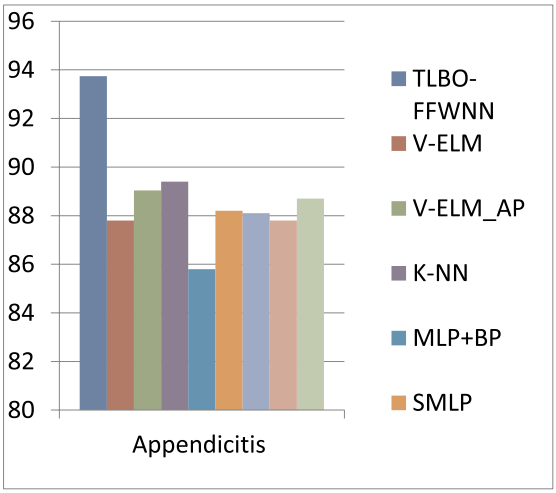


Fig. 15. Classification Accuracy on the Appendicitis dataset.

MLP) [6]; decision tree (J48) [82]; Multi-Layer Perception (MLP) [82]; Naive Bayes (NB) [82]; Sequential Minimal Optimization (SMO) [82]; Instance Based for K-Nearest neighbor (IBK) [82]; association rules and neural network (AR + NN) [83]; Voting-based Extreme Learning Machine (V-ELM), Voting-based Extreme Learning Machine with accuracy-based ensemble pruning (V-ELM_AP) [84]; k-nearest neighbor (K-NN), Multi-Layer perceptron + Back Propagation, search-based algorithm for Multi-Layer Perceptron (SMLP), Support Vector Machine (SVM) [8]; Learning Vector Quantization (LVQ) algorithm, Weighted Learning Vector Quantization (WLVQ) algorithm [7]; and lion's algorithm + support vector machine [85]. Table 11, Fig. 11, Fig. 12, Fig. 13, Fig. 14 and Fig. 15 show a comparison of the classification accuracy (testing phase) for the proposed method (TLBO-FFWNN) and other existing methods on the Breast Cancer, Heart Disease, Hepatitis, Pima-Indian diabetes and Appendicitis datasets, respectively.

The TLBO algorithm is used to manipulate the parameters of the FFWNN (weight, dilation, and translation). Various experiments have been conducted to ensure that all patterns in each dataset are used in the learning and testing process based on 5- and 10-fold cross-validation. From the obtained results, it can be concluded that TLBO-FFWNN has superior performance for all datasets, except the Heart dataset, with low computational complexity. For future research, the feature selection technique can be utilized to obtain more accurate results.

The proposed method (TLBO-FFWNN) has the best performance among all other methods across all datasets, except the LION + SVM method on the Heart dataset because the author used feature selection.

9. Conclusion

In this paper, the Teaching Learning-Based Optimization (TLBO) algorithm has been applied to evolve the learning of the proposed FFWNN method for classifying five medical datasets: Breast Cancer, Heart Disease, Hepatitis, Pima-Indian diabetes and Appendicitis.

The TLBO algorithm is used to manipulate the parameters of the FFWNN (weight, dilation, and translation). Various experiments have been conducted to ensure that all patterns in each dataset are used in the learning and testing process based on 5- and 10-fold cross-validation. From the obtained results, it can be concluded that TLBO-FFWNN has superior performance for all datasets, except the Heart dataset, with low computational complexity. For future research, the feature selection technique can be utilized to obtain more accurate results.

10. Compliance with ethical standards

The authors declare that there are no conflicts of interest regarding the publication of this paper. Additionally, this article does not contain any studies with human participants or animals performed by any of the authors.

Conflicts of interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] M.F. Akay, Support vector machines combined with feature selection for breast cancer diagnosis, Expert Syst. Appl. 36 (2) (2009) 3240–3247.
[2] S. Ghumbre, C. Patil, Ghatol, A Heart disease diagnosis using support vector machine, International Conference on Computer Science and Information Technology, Pattaya, 2011.
[3] S. Şahan, K. Polat, H. Kodaz, S. Güneş, The medical applications of attribute weighted artificial immune system (AWAIS): diagnosis of heart and diabetes diseases, Artificial Immune Systems, Springer, 2005, pp. 456–468.
[4] A. Khemphila, V. Boonjing, Heart disease classification using neural network and feature selection, Systems Engineering (ICSEng), 2011 21st International Conference, IEEE, 2011, pp. 406–409.
[5] K.U. Rani, Analysis of Heart Diseases Dataset Using Neural Network Approach, (2011) arXiv preprint arXiv:11102626.
[6] Z. Beheshti, S.M.H. Shamsuddin, E. Beheshti, S.S. Yuhaniz, Enhancement of Artificial Neural Network Learning Using Centripetal Accelerated Particle Swarm Optimization for Medical Diseases Diagnosis Soft Computing, (2013), pp. 1–18.
[7] M. Blachnik, W. Duch, LVQ Algorithm with Instance Weighting for Generation of Prototype-Based Rules Neural Networks vol. 24, (2011), pp. 824–830.
[8] M. Kordos, M. Blachnik, D. Strzempa, Do we need whatever more than k-NN? Artificial Intelligence and Soft Computing, Springer, 2010, pp. 414–421.
[9] D.K. Chaturvedi, A.P. Sinha, O.P. Malik, Short term load forecast using fuzzy logic and wavelet transform integrated generalized neural network, Int. J. Electr. Power Energy Syst. 67 (2015) 230–237.
[10] H. Malek, M.M. Ebadzadeh, M. Rahmati, Three new fuzzy neural networks learning algorithms based on clustering, training error and genetic algorithm, Appl. Intell. 37 (2012) 280–289.
[11] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (1965) 338–353.
[12] M. Gupta, D. Rao, On the Principles of Fuzzy Neural Networks Fuzzy Sets and Systems vol. 61, (1994), pp. 1–18.
[13] T. Hassanzadeh, K. Faez, G. Seyfi, A speech recognition system based on structure equivalent fuzzy neural network trained by firefly algorithm, Biomedical Engineering (ICoBE), 2012 International Conference, IEEE, 2012, pp. 63–67.
[14] Y. Wang, T. Mai, J. Mao, Adaptive Motion/force Control Strategy for Non-holonomic Mobile Manipulator Robot Using Recurrent Fuzzy Wavelet Neural Networks Engineering Applications of Artificial Intelligence vol. 34, (2014), pp. 137–153 https://doi.org/10.1016/j.engappai.2014.05.009.
[15] E. Karatepe, T. Hiyama, Fuzzy wavelet network identification of optimum operating point of non-crystalline silicon solar cells, Comput. Math. Appl. 63 (2012) 68–82.
[16] C.M. Lin, A.B. Ting, C.F. Hsu, C.M. Chung, Adaptive control for MIMO uncertain nonlinear systems using recurrent wavelet neural network, Int. J. Neural Syst. 22 (01) (2012) 37–50.
[17] Y. Bodyanskiy, O. Vynokurova, Hybrid Adaptive Wavelet-Neuro-Fuzzy System for Chaotic Time Series Identification Information Sciences vol. 220, (2013), pp. 170–179.
[18] M. Davanipoor, M. Zekri, F. Sheikholeslam, Fuzzy wavelet neural network with an accelerated hybrid learning algorithm Fuzzy Systems, IEEE Trans. 20 (2012) 463–470.
[19] D.W. Ho, P.-A. Zhang, J. Xu, Fuzzy wavelet networks for function learning Fuzzy Systems, IEEE Trans. 9 (2001) 200–211.
[20] Y.-H. Pao, Y. Takefji, Functional-link net computing, IEEE Comput. J. 25 (1992) 76–79.
[21] S. Dehuri, S.-B. Cho, Evolutionarily optimized features in functional link neural network for classification, Expert Syst. Appl. 37 (2010) 4379–4391.
[22] D.M. Sahoo, S. Chakraverty, Functional link neural network learning for response prediction of tall shear buildings with respect to earthquake data, IEEE Trans. Syst. Man Cybern.: Systems 48 (1) (2018) 1–10.
[23] R. Rao, V. Savsani, D. Vakharia, Teaching–learning-based optimization: an optimization method for continuous non-linear large scale problems, Inf. Sci. 183 (2012) 1–15.
[24] R. Rao, G. Waghmare, Solving composite test functions using teaching-learning-based optimization algorithm, Proceedings of the International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA), Springer, 2013, pp. 395–403.
[25] J. Salah Aldeen, R.A. Wahid, A comparative study among some natural-inspired optimization algorithms, J. Educ. Sci. Accepted (2014).
[26] S.C. Satapathy, A. Naik, K. Parvathi, Teaching learning based optimization for neural networks learning enhancement, Swarm, Evolutionary, and Memetic Computing, Springer, 2012, pp. 761–769.
[27] A. Karami, M. Guerrero-Zapata, A fuzzy anomaly detection system based on hybrid pso-kmeans algorithm in content-centric networks, Neurocomputing 149 (2015) 1253–1269.

[28] J. Kennedy, Particle swarm optimization, Encyclopedia of Machine Learning, Springer, 2010, pp. 760–766.
[29] R. Mallipeddi, P.N. Suganthan, Q.-K. Pan, M.F. Tasgetiren, Differential evolution algorithm with ensemble of parameters and mutation strategies, Appl. Soft Comput. 11 (2011) 1679–1696.
[30] R. Storn, K. Price, Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces, J. Glob. Optim. 11 (1997) 341–359.
[31] R. Rao, V. Savsani, D. Vakharia, Teaching–learning-based optimization: a novel method for constrained mechanical design optimization problems, Comput. Aided Des. 43 (2011) 303–315.
[32] D. Karaboga, An Idea Based on Honey Bee Swarm for Numerical Optimization Techn Rep TR06, Erciyes Univ Press, Erciyes, 2005.
[33] M.S. Mahmod, Z.A.H. Alnais, I.A.A. Al-Hadi, hybrid intrusion detection system using artificial bee colony algorithm and multi-layer perceptron, Int. J. Comput. Sci. Inf. Secur. 13 (2015) 1.
[34] A.H. Gandomi, X.-S. Yang, A.H. Alavi, Mixed variable structural optimization using firefly algorithm, Comput. Struct. 89 (2011) 2325–2336.
[35] X.-S. Yang, Firefly algorithms for multimodal optimization, Stochastic Algorithms: Foundations and Applications, Springer, 2009, pp. 169–178.
[36] E. Uzlu, M. Kankal, A. Akpınar, T. Dede, Estimates of energy consumption in Turkey using neural networks with the teaching–learning-based optimization algorithm Energy, http://doi.org/10.1016/j.energy.2014.07.078.
[37] K. Socha, C. Blum, An Ant Colony Optimization Algorithm for Continuous Optimization: Application to Feed-Forward Neural Network Training Neural Computing and Applications vol. 16, (2007), pp. 235–247.
[38] H. Kahramanli, N. Allahverdi, Design of a hybrid system for the diabetes and heart diseases, Expert Syst. Appl. 35 (2008) 82–89.
[39] R. Das, I. Turkoglu, A. Sengur, Effective diagnosis of heart disease through neural networks ensembles, Expert Syst. Appl. 36 (2009) 7675–7680.
[40] M. Mavrovouniotis, S. Yang, Training neural networks with ant colony optimization algorithms for pattern classification, Soft Comput. 19 (2014) 1511–1522.
[41] E. Habibi, M. Salehi, G. Yadegarfar, A. Taheri, Optimization of ANFIS using a genetic algorithm for physical work rate classification, Int. J. Occup. Saf. Ergon. (2018 Feb 1) 1–4 just-accepted.
[42] P. Pławiak, Novel genetic ensembles of classifiers applied to myocardium dysfunction recognition based on ECG signals, Swarm Evol. Comput. (2017 Oct 20).
[43] Y.D. Zhang, Z.J. Yang, H.M. Lu, X.X. Zhou, P. Phillips, Q.M. Liu, S.H. Wang, Facial emotion recognition based on biorthogonal wavelet entropy, fuzzy support vector machine, and stratified cross validation, IEEE Access 4 (2016) 8375–8385.
[44] S. Wang, Y. Li, Y. Shao, C. Cattani, Y. Zhang, S. Du, Detection of dendritic spines using wavelet packet entropy and fuzzy support vector machine, CNS Neurol. Disord. - Drug Targets 16 (2) (2017) 116–121.
[45] R.H. Abiyev, O. Kaynak, Fuzzy wavelet neural networks for identification and control of dynamic plants—a novel structure and a comparative study Industrial Electronics, IEEE Trans. 55 (2008) 3133–3140.
[46] R.H. Abiyev, O. Kaynak, Identification and control of dynamic plants using fuzzy wavelet neural networks, Intelligent Control, 2008. ISIC 2008. IEEE International Symposium, IEEE, 2008, pp. 1295–1301.
[47] J.S. Majeed Alneamy, R.A. Hameed Alnaish, Heart Disease Diagnosis Utilizing Hybrid Fuzzy Wavelet Neural Network and Teaching Learning Based Optimization Algorithm Advances in Artificial Neural Systems 2014, (2014).
[48] C.-F. Wu, C.-J. Lin, C.-Y. Lee, A functional neural fuzzy network for classification applications, Expert Syst. Appl. 38 (2011) 6202–6208.
[49] C.-H. Chen, Y.-C. Liu, C.-J. Lin, C.-T. Lin, A hybrid of cooperative particle swarm optimization and cultural algorithm for neural fuzzy networks, Fuzzy Systems, 2008. FUZZ-IEEE 2008.(IEEE World Congress on Computational Intelligence). IEEE International Conference, IEEE, 2008, pp. 238–245.
[50] S. Dehuri, R. Roy, S.-B. Cho, A. Ghosh, An improved swarm optimized functional link artificial neural network (ISO-FLANN) for classification, J. Syst. Softw. 85 (2012) 1333–1345.
[51] Z.A. Hamed, S.Z.M. Hashim, Hybrid PSO-black stork foraging for functional neural fuzzy network learning enhancement, Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference, IEEE, 2012, pp. 1339–1344.
[52] A. Baykasoğlu, A. Hamzadayi, S.Y. Köse, Testing the performance of teaching–learning based optimization (TLBO) algorithm on combinatorial problems: flow shop and job shop scheduling cases, Inf. Sci. 276 (2014) 204–218.
[53] R. Venkata Rao, V. Kalyankar, Parameter Optimization of Modern Machining Processes Using Teaching–Learning-Based Optimization Algorithm Engineering Applications of Artificial Intelligence, (2012).
[54] M.Y. Cheng, D. Prayogo, Fuzzy adaptive teaching–learning-based optimization for global numerical optimization, Neural Comput. Appl. 29 (2) (2018) 309–327.
[55] R. Rao, V. Savsani, J. Balic, Teaching–learning-based Optimization Algorithm for Unconstrained and Constrained Real-Parameter Optimization Problems Engineering Optimization vol. 44, (2012), pp. 1447–1462.
[56] G. Sharma, A. Kumar, Improved DV-Hop localization algorithm using teaching learning based optimization for wireless sensor networks, Telecommun. Syst. 67 (2) (2018) 163–178.
[57] G. Waghmare, Comments on "A Note on Teaching-Learning-Based Optimization Algorithm" Information Sciences, (2012).
[58] W. Cheng, F. Liu, L. Li, SIZE AND GEOMETRY OPTIMIZATION OF TRUSSES USING TEACHING-LEARNING-BASED OPTIMIZATION, Int. J. Optim. Civ. Eng. 3 (2013) 431–444.
[59] P.V. Babu, S.C. Satapathy, M.K. Samantula, P. Patra, B.N. Biswal, Teaching learning based optimized mathematical model for data classification problems, Proceedings of the International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA), Springer, 2013, pp. 487–496.
[60] V. Toğan, Design of Planar Steel Frames using Teaching–Learning based Optimization Engineering Structures, 34 (2012), pp. 225–232.
[61] T. Ince, S. Kiranyaz, J. Pulkkinen, M. Gabbouj, Evaluation of global and local training techniques over feed-forward neural network architecture spaces for computer-aided medical diagnosis, Expert Syst. Appl. 37 (2010) 8450–8461.
[62] Y. Pan, M.J. Er, X. Li, H. Yu, R. Gouriveau, Machine Health Condition Prediction via Online Dynamic Fuzzy Neural Networks Engineering Applications of Artificial Intelligence vol. 35, (2014), pp. 105–113.
[63] M. Pratama, M.J. Er, X. Li, R.J. Oentaryo, E. Lughofer, I. Arifin, Data driven modeling based on dynamic parsimonious fuzzy neural network, Neurocomputing 110 (2013) 18–28.
[64] A.J. Viera, J.M. Garrett, Understanding interobserver agreement: the kappa statistic, Fam. Med. 37 (2005) 360–363.
[65] D.J. Newman, A. Asuncion, UCI Machine Learning Repository, (2007).
[66] J. Alcalá-Fdez, L. Sánchez, S. García, M.J. del Jesus, S. Ventura, J.M. Garrell, J. Otero, C. Romero, J. Bacardit, V.M. Rivas, J.C. Fernández, F.K.E.E.L. Herrera, A software tool to assess evolutionary algorithms to data mining problems, Soft Comput. 13 (3) (2009) 307–318, https://doi.org/10.1007/s00500-008-0323-y.
[67] J. Alcalá-Fdez, A. Fernandez, J. Luengo, J. Derrac, S. García, L. Sánchez, F. Herrera, KEEL data-mining software tool: data set repository, integration of algorithms and experimental analysis framework, J. Mult.-Valued Log. Soft Comput. 17 (2–3) (2011) 255–287.
[68] A.A. AbuBaker, R.S. Qahwaji, M.J. Aqel, H. Al-Osta, M.H. Saleh, Efficient Preprocessing of Usf and Mias Mammogram Images, (2006).
[69] H.W. William, O.L. Mangasarian, Multisurface method of pattern separation for medical diagnosis applied to breast cytology, Proc. Natl. Acad. Sci. U.S.A. 87 (1990) 9193–9196.
[70] K. Polat, S. Güneş, S. Tosun, Diagnosis of heart disease using artificial immune recognition system and fuzzy weighted pre-processing, Pattern Recogn. 39 (11) (2006) 2186–2193.
[71] P. Gayathri, N. Jaisankar, Comprehensive study of heart disease diagnosis using data mining and soft computing techniques, Int. J. Eng. Technol. 5 (3) (2013) 2947–2958.
[72] M. Nilashi, H. Ahmadi, L. Shahmoradi, O. Ibrahim, E. Akbari, A predictive method for hepatitis disease diagnosis using ensembles of neuro-fuzzy technique, J. Infect. Public Health 12 (1) (2019) 13–20.
[73] K. Prasidthrathsint, J.T. Stapleton, Laboratory diagnosis and monitoring of viral hepatitis, Gastroenterol. Clin. (2019).
[74] P. Diaconis, B. Efron, Computer-Intensive Methods in Statistics Scientific American 248, (1983).
[75] P. Bowman, S.E. Flanagan, A.T. Hattersley, Future roadmaps for precision medicine applied to diabetes: rising to the challenge of heterogeneity, J. Diabetes Res. (2018) 2018.
[76] K. Ogurtsova, J.D. da Rocha Fernandes, Y. Huang, U. Linnenkamp, L. Guariguata, N.H. Cho, L.E. Makaroff, IDF Diabetes Atlas: global estimates for the prevalence of diabetes for 2015 and 2040, Diabetes Res. Clin. Pract. 128 (2017) 40–50.
[77] K. Polat, S. Güneş, A. Arslan, A cascade learning system for classification of diabetes disease: generalized discriminant analysis and least square support vector machine, Expert Syst. Appl. 34 (1) (2008) 482–487.
[78] S.K. Okur, Y.S. Koca, İ. Yıldız, İ. Barut, Right Hydronephrosis as a Complication of Acute Appendicitis. *Case Reports in Emergency Medicine*, (2016) 2016..
[79] A. Petroianu, Diagnosis of acute appendicitis, Int. J. Surg. 10 (3) (2012) 115–119 2012.
[80] N.A. Baykan, N. Yilmaz, A mineral classification system with multiple artificial neural network using K-fold cross validation, Math. Comput. Appl. 16 (2011) 22.
[81] Z.A. Hamed, S.Z.M. Hashim, Hybrid Particle Swarm Optimization and Black Stork Foraging for Functional Neural Fuzzy Network Learning Enhancement Master Thesis, Universiti Teknologi Malaysia, 2012.
[82] G.I. Salama, M. Abdelhalim, MA-e Zeid, Breast cancer diagnosis on three different datasets using multi-classifiers Breast Cancer, WDBC) 32 (2012) 2.
[83] M. Karabatak, M.C. Ince, An expert system for detection of breast cancer based on association rules and neural network, Expert Syst. Appl. 36 (2009) 3465–3469.
[84] S. Shukla, R. Yadav, Voting based Extreme learning machine with accuracy based ensemble pruning, Int. J. Comput. Appl. 115 (2015).
[85] K.C. Lin, J.C. Hung, J.T. Wei, Feature selection with modified lion's algorithms and support vector machine for high-dimensional data, Appl. Soft Comput. (2018).