Evolutionary Algorithms Based RBF Neural Networks For Parkinson s Disease Diagnosis

Yavuz DELİCAN[1], Lale ÖZYILMAZ[2], Tülay YILDIRIM[2]

[1]Department of Digital Electronic Design, TUBITAK SAGE, Ankara, Turkey
yavuz.delican@sage.tubitak.gov.tr
[2]Department of Electronics and Communication Eng., Yildiz Technical University, Istanbul, Turkey
ozyilmaz@yildiz.edu.tr, tulay@yildiz.edu.tr

Abstract

Parkinson's Disease (PD) is the second most common neurodegenerative action and expected to increase in the next decade with accelerating treatment costs as a consequence. This situation leads us towards the need to develop a Decision Support System for PD. In this paper we propose different methods based on evolutionary algorithms and RBF neural networks for diagnosis of PD. Three different evolutionary algorithms; genetic algoritm, particle swarm optimization and artificial bee colony algorithm (ABC) ; are used for training different structures of RBF neural networks. The experimental results show that the usage of ABC algorithm based RBF networks results better than the other methods, either in terms of accuracy or speed for PD diagnosis.

1. Introduction

Parkinson's disease (also known as Parkinson disease or PD) is a degenerative disorder of the central nervous system that often impairs the sufferer's motor skills, speech, and other functions. PD causes cognitive and mood disturbances, being in many cases related.

Having so many factors to analyze to diagnose PD, specialist normally makes decisions by evaluating the current test results of their patients. Moreover, the previous decisions made on other patients with a similar condition are also done by them. These are complex procedures, especially when the number of factors that the specialist has to evaluate is high (high quantity and variety of these data). For these reasons, PD diagnosis involves experience and highly skilled specialists [1].

The use of classifier systems in medical diagnosis is increasing gradually. Recent advances in the field of artificial intelligence have led to the emergence of expert systems and Decision Support Systems for medical applications. Moreover, in the last few decades computational tools have been designed to improve the experiences and abilities of doctors and medical specialists in making decisions about their patients. Without doubt the evaluation of data taken from patients and decisions of experts are still the most important factors in diagnosis. However, expert systems and different Artificial Intelligence techniques for classification have the potential of being good supportive tools for the expert. Classification systems can help in increasing accuracy and reliability of diagnoses and minimizing possible errors, as well as making the diagnoses more time efficient [2].

An artificial neural network (ANN), usually called neural network (NN), is a mathematical or computational model that is inspired by the structure and/or functional aspects of biological neural networks. A Radial basis function (RBF) network which was introduced by Broomhead and Lowe [3], is a special type of ANN that uses a radial basis function as its activation function. RBF is a real-valued function whose value depends only on the distance from the origin. RBF networks are very popular for function approximation, curve fitting, time series prediction, control and classification problems. The radial basis function network is different from other neural networks, possessing several distinctive features. Because of their universal approximation, more compact topology and faster learning speed, RBF networks have attracted considerable attention and they have been widely applied in many science and engineering fields.

In the literature, several studies have been reported focusing on PD diagnosis. In these studies, different methods were applied to the given problems[4-9]. This paper deals with the application of evolutionary algorithms based RBF neural networks to a medical dataset concerning PD with the aim of automatically classifying patients in PD or non-PD depending on their medical attributes. Three different evolutionary algorithms; genetic algoritm, particle swarm optimization and artificial bee colony algorithm (ABC) ; are used for training different structures of RBF as a decision support system for PD. The experimental results show that the usage of ABC algorithm results in better learning than the other methods, either in terms of accuracy or speed

2. Radial Basis Function Networks

Neural networks are non-linear statistical data modeling tools and can be used to model complex relationships between inputs and outputs or to find patterns in a dataset. RBF network is a type of feed forward neural network composed of three layers, namely the input layer, the hidden layer and the output layer. Each of these layers has different tasks [10]. A general block diagram of an RBF network is illustrated in Fig. 1.

In RBF networks, the outputs of the input layer are determined by calculating the distance between the network inputs and hidden layer centers. The second layer is the linear hidden layer and outputs of this layer are weighted forms of the input layer outputs. Each neuron of the hidden layer has a parameter vector called center. Therefore, a general expression of the network can be given as [11]:

$$y'_j = \sum_{i=1}^{I} w_{ij}\phi(\|\mathrm{x}-\mathrm{c}_i\|) + \beta_j \tag{1}$$

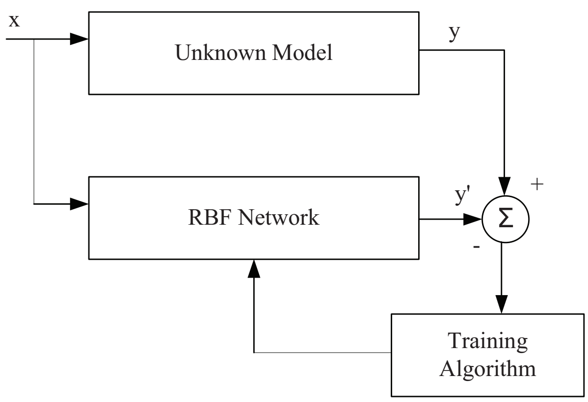

Fig. 1. Block diagram of a RBF network

The norm is usually taken to be the Euclidean distance and the radial basis function is also taken to be Gaussian function and defined as follows:

$$\varphi(r) = e^{(-\alpha_i . \|\text{x-c}_i\|^2)} \tag{2}$$

where,

I	Number of neurons in the hidden layer	$i \in \{1,2,...,I\}$
J	Number of neurons in the output layer	$j \in \{1,2,...,J\}$
w_{ij}	Weight of the i^{th} neuron and j^{th} output	
ϕ	Radial basis function	
α_i	Spread parameter of the i^{th} neuron	
x	Input data vector	
c_i	Center vector of the i^{th} neuron	
β_j	Bias value of the output j^{th} neuron	
y'_j	Network output of the j^{th} neuron	

Fig. 2 shows the detailed architecture of an RBF network. *M* dimensional inputs $(x_1,...,x_m)$ are located in the input layer, which broadcast the inputs to the hidden layer. The hidden layer includes *I* neurons and each neuron in this layer calculates the Euclidean distance between the centers and the inputs. A neuron in the hidden layer has an activation function called the basis function. In the literature, the Gaussian function is frequently chosen as the radial basis function and it has a spread parameter to shape the curve $(\alpha_1,...,\alpha_i)$. The weighted $(w_{11},...,w_{ij})$ outputs of the hidden layer are transmitted to the output layer. Here, $I(i \in \{1,2,...,I\})$ denotes the number of neurons in the hidden layer and $J(i \in \{1,2,...,J\})$ denotes the dimension of the output. The output layer calculates the linear combination of hidden layer outputs and bias parameters $(\beta_1,...,\beta_j)$. Finally, the outputs of the RBFnetwork are obtained $(y'_1,...,y'_j)$ [11].

The design procedure of the RBF neural network includes determining the number of neurons in the hidden layer. Then, in order to obtain the desired output of the RBF neural network *w*, *α*, c and *β* parameters might be adjusted properly. Reference based error metrics such as mean square error (MSE) or sum square error (SSE) can be used to evaluate the performance of the network. Error expression for the RBF network can be defined as follows:

$$\text{E}^{\text{SSE}}(w,\alpha,c,\beta) = \sum_{j=1}^{J}(y_j - y'_j)^2 \tag{3}$$

Here y_j indicates the desired output and y'_j indicates the RBF neural network output. The training procedure of the RBF neural network involves minimizing the error function [12].

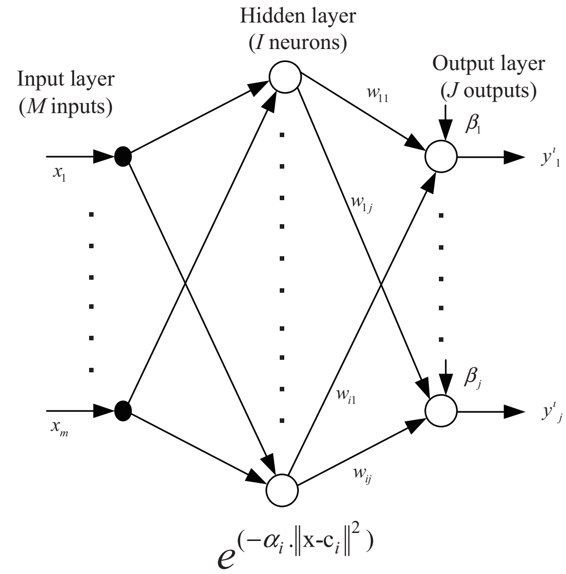

Fig. 2. Network architecture of the RBF

3. Evolutionary Algorithlms

In this paper; three different evolutionary algorithms ; genetic algoritm, particle swarm optimization and artificial bee colony algorithm.; are used for training RBF networks.

3.1. Genetic Algorithm (GA)

GA is an optimization method used in many research areas to find exact or approximate solutions to optimization and search problems. Inheritance, mutation, selection and crossover are the main aspects of GA that inspired from evolutionary biology. The population refers to the candidate solutions. The evolution starts from a randomly generated population. For all generations, the fitness (typically a cost function) of every individual in the population is evaluated and the genetic operators are implemented to obtain a new population. In the next iteration the new population is then used. Frequently, GA terminates when either a maximum number of generations has been reached, or a predefined fitness value has been achieved [12].

3.2. Particle Swarm Optimization (PSO)

PSO algorithm is inspired by the social behavior of bird flocking or fish schooling developed by Kenedy in 1995 [13]. In PSO, a set of particles (NP) of swarm is defined. Each particle represents a potential solution in the solution space and is characterized by its position and velocity. The number of parameters to be optimized determines the dimension of the problem. The position and velocity of i^{th} particle (i = 1, 2,... ,NP) in the D^{th} dimension are represented as $X_i = (x_{i1}, x_{i2}, x_{i3}, ..., x_{iD})$ and $V_i = (v_{i1}, v_{i2}, v_{i3}, ..., v_{iD})$ respectively. Each particle updates its position and velocity based on its own best position, (*pbest*) as well as the best position of the entire swarm (*gbest*).

3.3. Artificial Bee Colony Algorithm (ABC)

The Artificial Bee Colony algorithm is a heuristic optimization algorithm proposed by Karaboga in 2005 [14]. The ABC algorithm has been inspired by honey bees' intelligent foraging behavior. In the ABC model, the colony consists of three different bee groups, namely worker bees, onlooker bees and scout bees. For each food source there is only one associated employed bee. So, the number of the worker bees indicates the number of the food sources. Honey bees' intelligent foraging behavior can be explained as follows: worker bees go to their food sources, after determining the nectar amount for that food source they explore new neighboring food sources. Then, they come back and dance around the hive. Onlooker bees which are watching the dance of the worker bees, choose a food source according to the worker bees' dances. Probability of the food source that will be chosen is related with the quality of the food nectar and the left food amount. If a food source cannot be improved further through a predefined number of cycles, then the source is abandoned. Subsequently, randomly produced new sources are replaced with the abandoned ones by scouts. The best food source is determined and position of that food source is memorized. This cycle is repeated until requirements are met [15].

4. Training of RBF Neural Networks by Using Evolutionary Algorithms

Training of an RBF neural network can be obtained with the selection of the optimal values for the following parameters:

- weights between the hidden layer and the output layer (w)
- spread parameters of the hidden layer base function (α)
- center vectors of the hidden layer (c)
- bias parameters of the neurons of the output layer (β)

The number of neurons in the hidden layer is very important in neural networks. Using more neurons than that is needed causes an overlearned network and moreover, increases the complexity. Therefore, it has to be investigated how the numbers of neurons affect the network's performance. The individuals of the population of GA, PSO and ABC include the parameters of the weight (w^i), spread (α^i), center (c^i) and bias (β^i) vectors. An individual of the population of GA, PSO and ABC algorithm can be expressed as:

$$P_i = \left[w^i \alpha^i c^i \beta^i \right] \tag{4}$$

The quality of the individuals (possible solutions) can be calculated using an appropriate cost function. In the implementation, SSE between the actual output of the RBF network and the desired output is adopted as the fitness function [12]:

$$f = E^{SSE} \tag{5}$$

Fig. 3 shows the propesed approach to RBF network training with evolutionary algorithms. This algorithm starts by reading the Parkinson's dataset. This is followed by setting the desired RBF parameters (number of hidden neuorons and the maximum number of generations). The next step is to determine the control parameters of evolutionary algorithms. In each generation, every particle is evaluated based on the equation (5). Finally, the algorithm outputs the simulation results depending on equation (4).

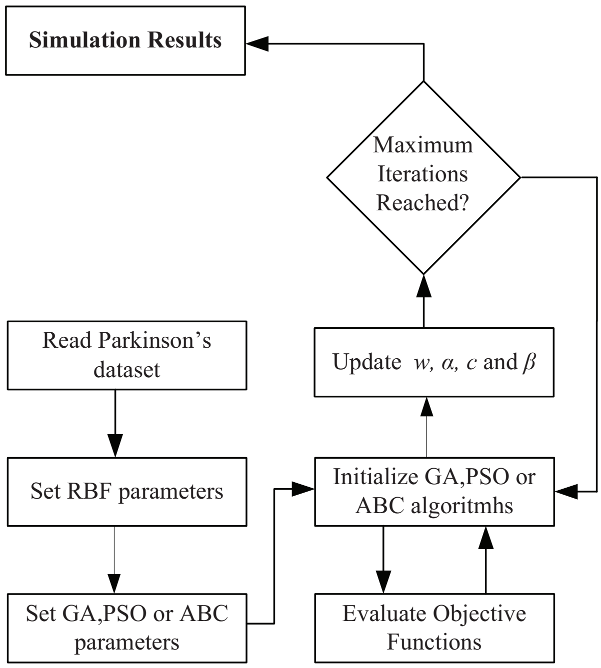

Fig.3. Evolutionary algorithms based RBF procudure

One of the most important design issue of an RBF network is the number of the neurons in the hidden layer. Therefore, the experiments are conducted on different RBF networks which have 1 neuron to 6 neurons located in the hidden layer.

5. Experimental Studies

Several experiments were conducted on the Parkinson's dataset to evalute the performance of the proposed algorithms. The Parkinson database used in this study is taken from the University of California at Irvine (UCI) machine learning repository [16]. It was used for training and testing experiments. The reason to use these sets of data is that the data sets of this website have been donated from hospitals. These data have been studied by many professionals of artificial intelligence departments. The dataset is composed of a range of biomedical voice measurements from 31 people, 23 with PD. Each column in Table 1 is a particular voice measure, and each row corresponds to one of 195 voice recordings of these individuals. The main aim of processing the data is to discriminate healthy people from those with PD, according to the "status" attribute which is set to non-PD for healthy and PD for people with Parkinson's disease, which is a two-decision classification problem. Table 1 shows the fields of this database and a brief description of each input variable.

Table 1. Description of the features of the Parkinsons's dataset

Field name	*Description*
MDVP: Fo(Hz)	Average vocal fundemental freq.
MDVP: Fi(Hz)	Max. vocal fundemental freq.
MDVP: Flo(Hz)	Min. vocal fundemental freq.
MDVP: Jitter(%)	Several measures of variation in fundamental frequency.
MDVP: Jitter(Abs)	
MDVP: RAP	
MDVP: PPQ	
Jitter: DDP	
MDVP: Shimmer	Several measures of variation in amplitude.
MDVP: Shimmer(dB)	
Shimmer: APQ3	
Shimmer: APQ5	
MDVP: APQ	
Shimmer: DDA	
NHR	Two measures of ratio of noise to tonal components in the voice
HNR	
RPDE	Two nonlinear dynamical complexity measures
D2	
DFA	Signal fractal scaling exponent
Spread1	Three nonlinear measures of fundamental frequency variation
Spread2	
PPE	

6. Simulation Results

In the experiments, the performance of RBF network trained by using GA, PSO and ABC is evaluated by the percent of correctly classified samples (PCCS) metric as a performance measurement.

$$PCCS = \frac{Correctly\ Classified\ Samples}{Total\ Samples} \times 100 \qquad (6)$$

For all datasets, experiments are repeated 20 times. For each run, datasets are randomly divided into train and test subsets. 70% of the data set is randomly selected as the training data and remained data set is selected as the testing data. Afterwards, average PCCS results of the 20 independent runs are calculated.

In all experimental studies, population size, limit of iterations and target value of GA, PSO and ABC are taken same as 40, 300 and 10^{-6}, respectively. In addition; mutation rate and crossover ratio of GA are taken as 0.06 and 0.8. Velocity update parameters of PSO c_1, c_2 and w are 2, 2 and 0.9 respectively.

The average PCCS values of 20 executions of the proposed algortihms changing with hidden layer neurons for the PD database are illustrated in Table 2.

Table 2. The average PCCS results

		Hidden Layer Neurons 1	2	3	4	5	6
GA	**Train**	57.5	86.6	87.4	88.8	86.3	88.6
	Test	51.4	80.2	83.4	85.4	84.1	85.3
PSO	**Train**	62.3	88.1	89.6	91.9	93.5	93.1
	Test	58.8	83.8	85.7	89.0	90.9	90.5
ABC	**Train**	**65.4**	**89.3**	**90.1**	**93.7**	**91.6**	**90.8**
	Test	**60.0**	**85.5**	**87.0**	**91.2**	**89.9**	**88.9**

Receiver Operating Characteristic (ROC) analysis is also commonly used in medicine and healthcare to quantify the accuracy of diagnostic test. The basic idea of diagnostic test interpretation is to calculate the probability a patient has a disease under consideration given a certain result. Without ROC analysis, it is difficult to summarize the performance of a test with a manageable number of statistics and to compare the performance of different tests.

The diagnostic performance is usually evaluated in terms of sensitivity and specificity. Sensitivity is the proportion of patients with disease whose tests are positive. Specificity is the proportion of patients without disease whose tests are negative. The measures are defined as:

$$sensitivity = \frac{number\ of\ true\ positives}{number\ of\ true\ positives + number\ of\ false\ negatives} \qquad (7)$$

$$specificity = \frac{number\ of\ true\ negatives}{number\ of\ true\ negatives + number\ of\ false\ positives} \qquad (8)$$

where number of true positives and number of false negatives are the number of PD correctly classified and incorrectly classified as normal case, respectively. Similarly, number of true negatives and number of false positives are the number of normal case correctly classified and incorrectly classified as PD case [17].

ROC Analysis is applied to the PD data set and the results are presented in Table 3 for evolutionary algorithms based RBF neural networks with four hidden neurons structure at test stage.

Table 3. ROC analysis results

	Sensitivity	*Specificity*
GA	0.82	0.88
PSO	0.86	0.92
ABC	0.88	0.94

As can be seen from Table 2 and Table 3, from best to worst, the algorithms can be ordered as ABC, PSO, and GA, respectively. In addition, the number of neurons affects the network performance. As can be seen from Table 2, as the number of neurons increases, the performance of the network increase up to some point. So, as can be seen from Table 2, using ABC with four neurons RBF structure gives the best result about 93.7% training and 91.2% testing PCCS results. Other algorithms can reach the same performance by using more than four neurons. Since the number of neurons directly influences the time complexity of the algorithm, the required minimum number of neurons has to be used in the applications. In this context, the ABC algorithm is better than the others. As a result, it can be said that ABC is better than the others from the point of view of higher average PCCS results and higher speed.

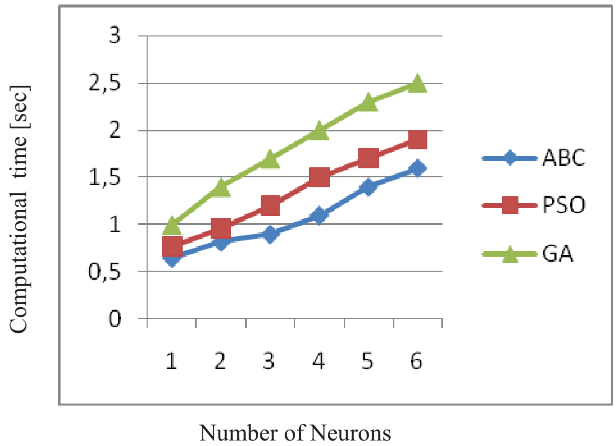

Fig. 4. Evaluation CPU time of ABC, PSO and GA based RBF neural network

Fig. 4 shows the evaluation CPU time of the evolutionary algorithms based RBF network after training stage. As can be seen from the figure, number of neurons located in the hidden layer of the RBF network affects the CPU time of the RBF network in real-time applications. While the number of neurons increase, computational time increase proportionally. Also it can be seen that ABC algorithm is better than the other ones considering computational times. The proposed algorithms are constituted using MATLAB 2009 with Pentium dual-core CPU @2.60GHz and 1GB RAM.

7. Conclusions

In this study, evolutionary algortihms based RBF neural networks are used for diagnosing of the Parkinson's disease. Three different evolutionary algorithms for training different structures of RBF networks were used. These are Genetic Algortihm, Particle Swarm Optimization and Artificial Bee Colony respectively. Several experimental studies were employed for calculating the performance score of the classifiers. Artificial Bee Colony based RBF neural networks with four neurons yielded the best score either in terms of accuracy or speed. The best simulation result gained 93.7% and 91.2% classification accuracy respectively as train and test classification stages for Parkinson's disease diagnosis.

8. References

[1] A.G. David, B. Johnson, "Diagnosing Parkinson by using Artificial Neural Networks and Support Vector Machines", *Global Journal of Computer Science and Technology*, Vol. 9, No.3, pp. 63-72, 2009.

[2] M.F. Akay, "Support vector machines combined with feature selection for breast cancer diagnosis", *Expert Systems with Applications*, Vol.36, pp. 3240-3247, 2008.

[3] D. Broomhead, D. Lowe, "Multivariable functional interpolation and adaptive networks", *Complex Systems* Vol. 2, pp. 321-355, 1988.

[4] E. Marius, "Neural network-based approach to discriminate healthy people from those with Parkinson's disease", *Mat. Comp. Sci. Ser.*, Vol.35,pp. 112-116, 2008.

[6] D. Resul, "A comparison of multiple classification methods for diagnosis of Parkinson disease", *Expert Systems with Applications*, Vol.37, pp. 1568-1572, 2010.

[7] C. Okan, O Kursun, "Telediagnosis of Parkinson's Disease Using Measurements of Dysphonia", *Journal of Medical Systems,* Vol. 34, pp. 591-599, 2009.

[8] K. Revett, F. Gorunescu and A.B. Salem, "Feature Selection in Parkinson's disease: A Rough Sets Approach", Proceedings of the IMCSIT, Vol. 4, pp. 425-428, 2009.

[9] B. Bolat, S. B. Sert, "Classification of Parkinson's disease by using voice measurements", *International Journal of Reasoning-based Intelligent Systems*, Vol.2, pp. 279 – 284, 2010.

[10] B. Yu, X. He, "Training Radial Basis Function Networks with Differential Evolution", *Proceedings of IEEE Conference on Granular Computing*, pp. 369-372, 2006.

[11] J. Robert, C. Howlett, "Radial Basis Function Networks 2: New Advances Design", *Physica-Verlag*, 2001.

[12] T. Kurban, E. Besdok, "A Comparison of RBF Neural Network Training Algorithms for Inertial Sensor Based Terrain Classification", *Sensors*, 2009.

[13] R. Eberhart, J. Kenedy, "Particle Swarm Optimization", *Proceedings of IEEE Int. Conference on Neural Networks*, pp. 1114-1121,1995.

[14] D. Karaboga, "*An Idea Based On Honey Bee Swarm for Numerical Optimization*", Technical Report-TR06, Erciyes University, 2005.

[15] B. Basturk, D. Karaboga, " A Powerful and Efficient Algorithm for Numerical Function Optimization: Artificial Bee Colony (ABC) Algorithm", *Journal of Global Optimization*, Vol. 39, pp. 459-471, 2007.

[16] Learning Repository, U.C.I.: http://archive.ics.uci.edu/ml/, June 2008.

[17] C. Senol, T. Yildirim, "Thyroid and Breast Cancer Diagnosis using Fuzzy-Neural Networks", *International Conference on Electrical and Electronics Engineering ELECO*, 2009.