Original Research

Applied Medical Informatics
Vol. 34, No. 1 /2014, pp: 22-30

Comparison of Back propagation neural network and Back propagation neural network Based Particle Swarm intelligence in Diagnostic Breast Cancer

Farahnaz SADOUGHI[1], Mustafa GHADERZADEH[2,*], Rebecca FEIN[3], Arran STANDRING[3]

[1] Department of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, Iran

[2] Department of Medical Informatics, Tehran University of Medical Sciences, Tehran, Iran

[3] Applied Health Informatics, Bryan University, Tempe, Arizona 85281, USA

Emails:msadoughi.f@gmail.com; Ir.medicalinformatics@yahoo.com*; rebeccaa.fein@gmail.com; arran.standring@bryanuniversity.edu

* Author to whom correspondence should be addressed; Tel.:+98-914-980-6771.

Received: 21.12.2013 /Accepted: 28.3.2014/ Published online: 31.3.2014

Abstract

Breast cancer is the most commonly diagnosed cancer and the most common cause of death in women all over the world. Use of computer technology supporting breast cancer diagnosing is now widespread and pervasive across a broad range of medical areas. Early diagnosis of this disease can greatly enhance the chances of long-term survival of breast cancer victims. Artificial Neural Networks (ANN) as mainly method play important role in early diagnoses breast cancer. This paper studies Levenberg Marquardet Backpropagation(LMBP) neural network and Levenberg Marquardet Backpropagation based Particle Swarm Optimization(LMBP-PSO) for the diagnosis of breast cancer. The obtained results show that LMBP and LMBP based PSO system provides higher classification efficiency. But LMBP based PSO needs minimum training and testing time. It helps in developing Medical Decision System (MDS) for breast cancer diagnosing. It can also be used as secondary observer in clinical decision making.

Keywords: Breast cancer; Artificial Neural Network; Levenberg Marquardet Backpropagation; Particle Swarm Optimization; Medical Decision System

Introduction

Breast cancer is a major cause of death by cancer in the female population. Most breast cancer cases occur in women aged 40 and above but certain women with high-risk characteristics, often hereditary, may develop breast cancer at a younger age. Cancer is a disease in which cells become abnormal and replicate forming more cells in an uncontrolled way. With breast cancer, the cancer begins in the tissues that make up the breast. The cancer cells may form a mass called a tumor. They may also invade nearby tissue and spread to lymph nodes and other parts of the body. Most breast cancer are detected by the patient as a lump in the breast. The majority of breast lumps are benign so it is the physician's responsibility to diagnose breast cancer, that is, to distinguish benign lumps from malignant ones[1, 2]. There are a number of different methods for diagnosing breast cancer, though the most widely used are: mammography, Fine-needle Aspiration (FNA) with visual interpretation, and surgical biopsy. The reported sensitivity (i.e., ability to correctly diagnose cancer when the disease is present) of mammography varies from 68% to 79%, of FNA with visual interpretation from 65% to 98%, and of surgical biopsy close to 100%. Therefore, mammography lacks sensitivity, FNA sensitivity varies widely, and surgical biopsy, although accurate, is invasive, time consuming, and costly. In order to develop a non-invasive way to diagnosis breast cancer artificial intelligence has been introduced [1-5].

Material and Method

Various artificial intelligence techniques have been used to improve the diagnostic procedures and to aid the physician's efforts. The most commonly intelligence techniques is Artificial Neural Networks.

Neural Network Techniques For Diagnosis Of Breast Cancer

Neural Networks are currently a 'hot' research area in medicine, particularly in the fields of radiology, urology, cardiology, and oncology. Keeping in view the significant characteristics of Neural Network (NN) and its advantages for implementation of the classification problem, Neural Network technique is highly used in the classification of data related to a medical field. Owing to their wide range of applicability, their ability to learn complex and non linear relationships; including noisy or less precise information, Neural Networks (NN) technique is used to solve problems in biomedical engineering[3, 6-9]. By their nature, Neural Networks are capable of high-speed parallel signal processing in real time. They have an advantage over conventional technologies because they can solve problems that are too complex and that do not have any algorithmic solution; or for which an algorithmic solution is too complex. The applications of neural networks in biomedical computing are numerous. Various applications of ANN techniques in the medical field like; medical expert system, cardiology, neurology, rheumatology, mammography, and pulmonology were studied [3, 8, 10]. The diagnosis of breast cancer in this study was performed by employing a Multilayer Feed Forward Neural Network (MFNN) with 2 inputs. The NN was trained by using the steepest descent with a momentum back propagation algorithm with logsig and purelin transfer function in a MATLAB environment. The back propagation algorithm is the most commonly used algorith[3]m in medical computational application as were experimented by[11, 12]. Present research focuses solely on neural networks and is produced in order to analyze different Neural Networks and their precision when it used Particle Swarm Optimization. The analysis is done on Levenberg-Marquardet Back Propagation Neural Network (LMBPNN) and Levenberg-Marquardet Back Propagation Neural Network based PSO (LMBPNN-PSO) with a Wisconsin Breast Cancer dataset and a conclusion is formed on the basis of their performance and efficiency.

Levenberg-Marquardt Back Propagation Algorithm

The application of Levenberg-Marquardt to neural network training is described in [8-9]. This algorithm has been shown to be the fastest method for training moderate-sized feed-forward neural networks (up to several hundred weights). It also has an efficient implementation in MATLAB software, since the solution of the matrix equationuation is a built-in function, so its attributes become even more pronounced in a MATLAB environment [10]. The network trainlm can train any network as long as its weight, net input, and transfer functions have derivative functions. Back propagation is used to calculate the Jacobian jX of performance with respect to the weight and bias variables X. Each variable is adjusted according to Levenberg-Marquardt equation (1);

$$\begin{gathered} \mathrm{jj} = \mathrm{jX} * \mathrm{jX} \\ \mathrm{je} = \mathrm{jX} * \mathrm{E} \\ \mathrm{dX} = -(\mathrm{jj} + \mathrm{I} * \mathrm{mu})/\mathrm{je} \end{gathered} \tag{1}$$

where E is all errors and **I** is the identity matrix. The adaptive value mu is increased until the change results in a reduced performance value [8, 13, 14].

Particle Swarm Optimization

Particle Swarm Optimization algorithm (PSO) is a randomly optimal algorithm based on swarm intelligence. The algorithm can be used to solve optimization problems [15]. One of the first implementations of Particle Swarm Optimization (PSO) was that of training Neural Networks and one key advantage of PSO over other optimization algorithms in training neural networks is its comparative simplicity. As described by Eberhart and Kennedy, the PSO algorithm is an adaptive algorithm based on a social psychological metaphor; a population of individuals adapts by returning stochastically toward previously successful regions in the search space, and is influenced by the successes of their topological neighbors [16-18]. PSO is a population-based search process where individuals initialized with a population of random solutions, referred to as particles, are grouped into a swarm. Each particle in the swarm represents a candidate solution to the optimization problem and if the solution is made up of a set of variables the particle can correspondingly be a vector of variables. In a PSO system each particle is "flown" through the multidimensional search space, adjusting its position in search space according to its own experience and that of neighboring particles. The particle therefore makes use of the best position encountered by itself and that of its neighbors to position itself toward an optimal solution. The performance of each particle is evaluated using a pre-defined fitness function, which encapsulates the characteristics of the optimization problem. The main operators of the PSO algorithm are the velocity and the position of the each particle. In each iteration particles evaluate their positions according to a fitness function. Then the velocity and the position of the each particle are updated according to the below Equation 2:

$$V_i(t+1) = wv_i(t) + c_1r_1\big(p_{id} - x_i(t)\big) + c_2r_2(p_{gd} - x_i(t)) \quad (2)$$

where it is the current step number, w is the inertia weight. Researchers have shown that for large values of the inertia weight, the global search ability of the algorithm increases. Nevertheless, once the algorithm converges to the optimum solution, it can be considered as a disadvantage to select a large value for the inertia weight. For this reason the methods which offer to adjust the inertia weight adaptively, have been proposed. c1 and c2 are the acceleration constants, r_1 and r_2 are two random numbers in the range [0, 1], $x_i(t)$ is the current position of the particle, p_{id} is the best one of the solutions this particle has reached, p_{gd} is the best one of the solutions all the particles have reached. After calculating the velocity, the new position of each particle can be calculated according to Equation 3:

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (3)$$

The PSO algorithm performs repeated applications of the update equationuations above until a specified number of iterations has been exceeded, or until the velocity updates are close to zero [16, 19].

Levenberg - Marquardet Back Propagation Neural Network (BPNN)-Particle Swarm Optimization algorithm

The PSO algorithm is a global algorithm, which has a strong ability to find global optimistic result. This PSO algorithm does however have a disadvantage in that the search around global optimum is very slow. This means that the particle swarm optimization algorithm was shown to converge rapidly during the initial stages of a global search, but around global optimum, and the search process will become very slow. On the contrary, Back Propagation Neural Network (BPNN) has a strong ability to find the local optimistic result, but its ability to find the global optimistic result is weak. In other words, it can achieve faster convergent speed around global optimum, and at the same time, the convergent accuracy can be higher. By combining the PSO with the BPNN, a new algorithm referred to as PSO–BPNN hybrid algorithm is formulated. The BP algorithm has a strong ability to find the local optimistic result. Some researchers have used PSO to train neural networks and found that PSO-based ANN has a better training performance, faster convergence rate, as well as a better predicting ability than BP-based ANN does. The fundamental idea for this hybrid algorithm is that at the beginning stage of searching for the optimum the PSO is employed to accelerate the training speed. When the fitness function value has not changed for some generations, or value changed is smaller than a pre-defined number, the searching process is switched to gradient descending searching according to this heuristic knowledge. PSO builds a set number of ANN, initializes all network weights to random values, and starts training each one. Once each pass through a data set, PSO compares each networks fitness. The network with the highest fitness is considered the global best. The other networks are updated based on the global best network rather than on their personal error or fitness.

Dataset

This breast cancer database was obtained from the University of Wisconsin Hospitals, Madison (WBCD) from Dr. William H. Wolberg. The database contains 699 samples with 683 complete data and 16 samples with missing attributes. The WBCD database consists of nine features obtained from Fine needle aspirates, each of which is ultimately represented as an integer value between 1 and 10. The nine attributes detailed in Table 1 are graded on an interval scale from a normal state of 1–10, with 10 being the most abnormal state. In this database, 241 (65.5%) records are malignant and 458 (34.5%) records are benign.

Table 1. Wisconsin breast cancer data description of attributes

Attribute number	Attribute description	Values of attributes	Mean	Standard deviation
1	Clump thickness	1-10	4.42	2.82
2	Uniformity of cell size	1-10	3.13	3.05
3	Uniformity of cell shape	1-10	3.20	2.97
4	Marginal adhesion	1-10	2.80	2.86
5	Single epithelial cell size	1-10	3.21	2.21
6	Bare nuclei	1-10	3.46	3.64
7	Bland chromatin	1-10	3.43	2.44
8	Normal nucleoli	1-10	2.87	3.05
9	Mitoses	1-10	1.59	1.71

N = 599 observations, 357 malignant and 212 benign.

These attributes measure the external appearance and internal chromosome changes in nine different scales. There are two values in the class variable of breast cancer: benign (non cancerous) and malignant (cancerous). A total of 212 samples of the data set belong to benign, and remaining 357 data are malignant. The original data can be presented in the form of analog values with values ranging from 0-10. These attributes measure the external appearance and internal chromosome changes in nine different scales. There are two values in the class variable of breast cancer:
Benign (non-cancerous) and Malignant (cancerous),
Descriptions of Database:

- Number of instances 569
- Number of attributes: 10 plus the class attribute
- Attributes 2 through 10 will be used to represent instances
- Each instance has one of 2 possible classes: benign or malignant

Conversion of the given data sets into binary can be done based on certain ranges, which are defined for each attribute [20].

Result and Discussion

In this Section, PSO-NN is applied to diagnostic breast cancer. The architecture of multi-layered feed forward neural network is shown in Figure. 1. It consists of one input layer, one output layer, and a hidden layer. It may have one or more hidden layers. All layers are fully connected and of the feed forward type. The outputs are nonlinear function of inputs, and are controlled by weights that are computed during the learning process. The learning process used is a supervised type and the learning paradigm is the back propagation. Figure. 1 shows the structure of a two layered feed forward neural network. For Levenberg-Marquardt (LM) training process, a bounded differentiable activation function is required. The most commonly known function known as the tan-sigmoid has been used. It is bounded between the minimum (-1) and maximum (1). Before a signal is passed to the next layer of neurons, the summed output of each neuron is scaled by this function.

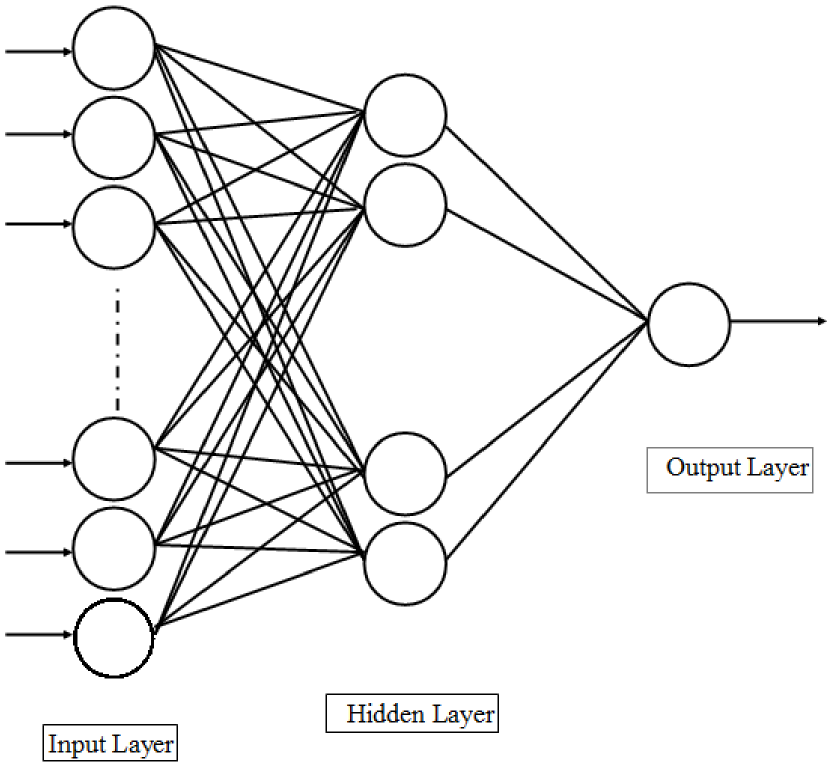

Figure 1. A two-layered feed forward neural network structure

The algorithm which is used in this study to evolve RBFNN is as follows[21, 22].

1. Initialize swarm of N particles. Each particle defines a network and the associated centers and bandwidths. Set the number of iterations as Max Iteration. Set count = 0.
2. Decode each particle into a network. Compute the connection weights between the hidden layer and the output of the network by the pseudo–inverse method. Compute the fitness of each particle.
3. Update pi for each particle and pg for whole swarm.
4. Update the velocity of each particle according to Equation (10). Limit the velocity in [Vmax, Vmax].
5. Update the position according to the Equation (11).
6. Set count = count +1; if count < MaxIterations, go to step 2, otherwise terminate the algorithm.

In present study the flowchart of LMBPNN based PSO is depicted as below Figure 2.

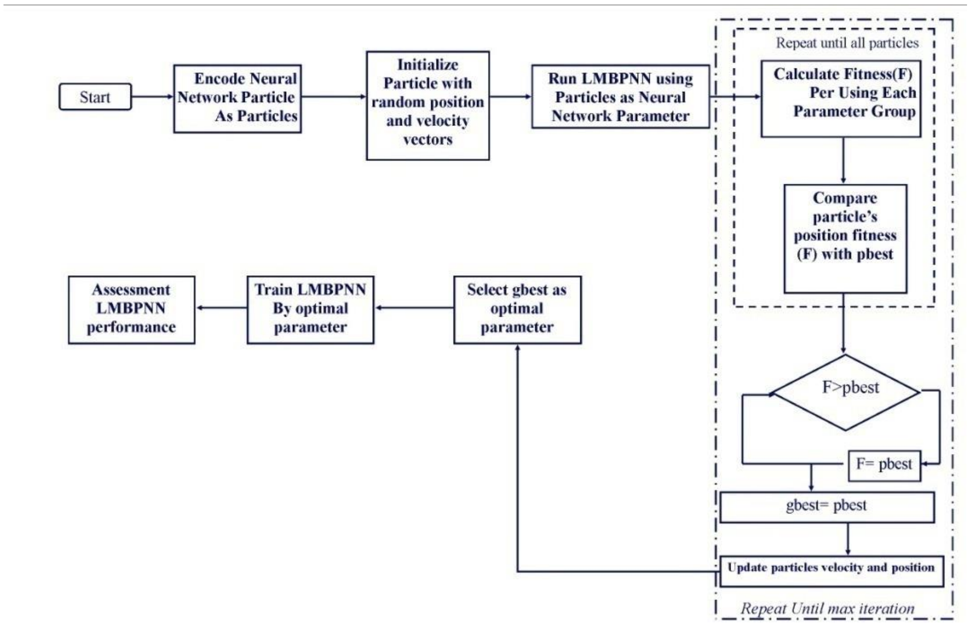

Figure 2. Strategy of using PSO in optimization LMBPNN parameters

In design algorithm fitness function of LMBPNN based PSO is the mean squared error (MSE). The algorithm begins with the random generation of an initial population of particle as a network parameter.

In this paper, an efficient modified Levenberg-Marquardet algorithm based PSO method is provided. Simulation study and comparisons with basic neural network algorithm and neural network based PSO algorithm show that NN-PSO improves the searching efficiency and searching quality effectively. The proposed algorithms have been implemented using MATLAB. In order to compare the performance of neural network techniques, firstly, each data set is split into 75% training set and 25% testing set. After the test data is classified by each method, the average of the ten results of the classification accuracy will be used for comparing the performance of LMBPNN and LMBPNN based PSO algorithm.

In the testing phase, the testing dataset is given to the proposed system to diagnose breast cancer and the obtained results are evaluated with the evaluation metrics namely; sensitivity, specificity, and accuracy. Sensitivity, specificity, and accuracy are the commonly used statistical measures to illustrate the medical diagnostic test and especially used to enumerate how the test was good and consistent. Sensitivity evaluates the diagnostic test correctly at detecting a positive disease. Specificity measures how the proportion of patients without disease can be correctly ruled out. Accuracy can be concluded with the aid of the sensitivity and specificity measures in the presence of prevalence. Accuracy measures correctly figured out diagnostic tests by eliminating a given condition. In order to find these metrics, we first compute some of the terms like: True positive, True negative, False negative and False positive based on the definitions given in Table 2.

In the present study previous research parameters were prepared and arranged. Changes in the velocity are stochastic; a particle can diverge from the solution space. So, a method is implemented to limit the velocity. At each iteration, after the velocity of the ith particle is updated, if the velocity is greater or smaller than from a given $[-v_{max}, v_{max}]$ interval, it is limited to $-v_{max}$ or v_{max}. This prevents the particle to diverge from the solution space. If the solution space boundary can be predicted, the v_{max} value can be chosen as $v_{max} = k \times x_{max}$, $0.1 \leq k \leq 1.0$ [21] .Parameters of PSO are also fine tuned in order to get the best results. Table I shows the range in which each parameter is searched, as well as the optimum values for each parameter. The optimum values for the PSO are derived after an extensive search over the ranges defined in Table 3.

Table 2. Terms used to define sensitivity, specificity, and accuracy

Outcome of the diagnostic test	Condition (e.g. disease) as determined by the Standard of Truth Positive	Negative	Row total
Positive	TP	FP	TP+ FP (total number of subjects with positive test)
Negative	FN	TN	FN+TN (total number of subjects with negative test)
Column total	TP+ FN (total number of subjects with given condition)	FP +TN (total number of subjects without given condition)	N =TP+TN+ FP+FN (Total number of subjects in study)

Table 3. PSO algorithm parameters

Parameter	Optimum Value
Number of Particles	**30**
c1	**0.3**
c2	**1.9**
[$-V_{max}$, V_{max}]	[-0.1 , 0.1]
MaxIteration	1000

After the test data is classified by each neural network algorithm, the average of the ten results of the classification accuracy will be used for comparing the performance of these Methods.

Table 4. Classification Accuracies Obtained by LMBPNN and LMBPNN-PSO Technique for Breast Cancer Diagnosis

NO	Number of Node in hidden layer	LMBPNN		LMBPNN-PSO	
		Sensitivity	Specificity	Sensitivity	Specificity
1	2	97	92	94.3	96.3
2	4	97	93.5	97.2	99.1
3	5	100	97	97.14	94.45
4	7	94.2	98	97.15	97.25
5	9	100	92.5	100	98.2
6	11	100	97.2	100	95.4
7	13	94.2	95.3	94.3	99.1
8	14	97.1	99	100	98.2
9	17	97	99	100	98.14
10	20	100	97.2	97.2	99.1

Conclusions

This paper presents the comparison of two Algorithms of Artificial intelligence and swarm intelligence. Two different learning algorithms were applied in this paper for training a Levenberg-Marquardet Back Propagation Neural Network (LMBPNN): Back Propagation (BP) and Particle Swarm Optimization techniques (PSO) were two training algorithms applied for updating and optimizing the output synaptic weight matrix. The Back Propagation learning algorithm is the most commonly used technique for updating neural network weight parameters. PSO is the most important technique of swarm intelligence. By implementation of this method we concluded that BP has a slow convergence speed and might at times diverge. It also requires extensive calculations if the size of the network increases and it may be difficult to implement when no gradient information is available for all activation functions. The obtained results show that LMBP and LMBP based PSO system provides higher classification efficiency. On the other hand the PSO algorithm has shown to have several advantages, both in terms of robustness and the efficiency in finding the optimal weights for the LMBPNN. Statistical results are provided that confirm PSO as a reliable algorithm for training such a neural network. It is clear that Artificial Neural Networks are a very powerful and accurate tool for diagnosis of breast cancer. LMBP based PSO needs minimum training and testing time. It helps in developing Medical Decision System (MDS) for breast cancer diagnosing. It can also be used as secondary observer in clinical decision making.

Conflict of Interest

The authors declare that they have no conflict of interest.

Acknowledgments

This study has been funded and supported by Iran university of medical sciences (IUMS); Grant no 18054.

References

1. Arora N, Martins D, Ruggerio D, Tousimis E, Swistel AJ, Osborne MP, et al. Effectiveness of a noninvasive digital infrared thermal imaging system in the detection of breast cancer. The American Journal of Surgery. 2008;196(4):523-6.
2. Christoyianni I, Koutras A, Dermatas E, Kokkinakis G. Computer aided diagnosis of breast cancer in digitized mammograms. Computerized Medical Imaging and Graphics. 2002;26(5):309-19.
3. Delen D, Walker G, Kadam A. Predicting breast cancer survivability: a comparison of three data mining methods. Artificial Intelligence in Medicine. 2005;34(2):113-27.
4. Giger ML, editor. Computerized analysis of images in the detection and diagnosis of breast cancer. Seminars in Ultrasound, CT, and MRI; 2004: Elsevier.
5. Mangasarian OL, Street WN, Wolberg WH. Breast cancer diagnosis and prognosis via linear programming. Operations Research. 1995;43(4):570-7.
6. Khan J, Wei JS, Ringner M, Saal LH, Ladanyi M, Westermann F, et al. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. Nature medicine. 2001;7(6):673-9.
7. Floyd CE, Lo JY, Yun AJ, Sullivan DC, Kornguth PJ. Prediction of breast cancer malignancy using an artificial neural network. Cancer. 1994;74(11):2944-8.
8. Karabatak M, Ince MC. An expert system for detection of breast cancer based on association rules and neural network. Expert systems with Applications. 2009;36(2):3465-9.
9. Kiyan T. Breast cancer diagnosis using statistical neural networks. IU-Journal of Electrical & Electronics Engineering. 2011;4(2).
10. Akay MF. Support vector machines combined with feature selection for breast cancer diagnosis. Expert systems with Applications. 2009;36(2):3240-7.
11. Abbass HA. An evolutionary artificial neural networks approach for breast cancer diagnosis. Artificial Intelligence in Medicine. 2002;25(3):265-81.
12. Meinel LA, Stolpen AH, Berbaum KS, Fajardo LL, Reinhardt JM. Breast MRI lesion classification: Improved performance of human readers with a backpropagation neural network computer-aided diagnosis (CAD) system. Journal of magnetic resonance imaging. 2007;25(1):89-95.
13. Übeyli ED. Implementing automated diagnostic systems for breast cancer detection. Expert systems with Applications. 2007;33(4):1054-62.
14. Koay J, Herry C, Frize M, editors. Analysis of breast thermography with an artificial neural network. Engineering in Medicine and Biology Society, 2004 IEMBS'04 26th Annual International Conference of the IEEE; 2004: IEEE.

15. Zhou C, Gao L, Gao H-b. Particle swarm optimization based algorithm for constrained layout optimization. Control and Decision. 2005;20(1):36-40.
16. Kennedy J, Eberhart R, editors. Particle swarm optimization. Neural Networks, 1995 Proceedings, IEEE International Conference on; 1995: IEEE.
17. Settles M, Rylander B. Neural network learning using particle swarm optimizers. Advances in Information Science and Soft Computing. 2002:224-6.
18. Russell SJ, Norvig P, Canny JF, Malik JM, Edwards DD. Artificial intelligence: a modern approach: Prentice hall Englewood Cliffs; 1995.
19. Da Y, Xiurun G. An improved PSO-based ANN with simulated annealing technique. Neurocomputing. 2005;63:527-33.
20. Rani KU. Parallel approach for diagnosis of breast cancer using neural network technique. International Journal of Computer Applications. 2010;10(3).
21. Korürek M, Doğan B. ECG beat classification using particle swarm optimization and radial basis function neural network. Expert systems with Applications. 2010;37(12):7563-9.
22. Liang JJ, Qin AK, Suganthan PN, Baskar S. Comprehensive learning particle swarm optimizer for global optimization of multimodal functions. Evolutionary Computation, IEEE Transactions on. 2006;10(3):281-95.