Introduction of a New Diagnostic Method for Breast Cancer Based on Fine Needle Aspiration (FNA) Test Data and Combining Intelligent Systems

Mohammad Fiuzy[1], Javad Haddadnia[1], Nasrin Mollania[2], Maryam Hashemian[3], Kazem Hassanpour[3]

Abstract

Background: Accurate Diagnosis of Breast Cancer is of prime importance. Fine Needle Aspiration test or "FNA", which has been used for several years in Europe, is a simple, inexpensive, noninvasive and accurate technique for detecting breast cancer. Expending the suitable features of the Fine Needle Aspiration results is the most important diagnostic problem in early stages of breast cancer. In this study, we introduced a new algorithm that can detect breast cancer based on combining artificial intelligent system and Fine Needle Aspiration (FNA).

Methods: We studied the Features of Wisconsin Data Base Cancer which contained about 569 FNA test samples (212 patient samples (malignant) and 357 healthy samples (benign)). In this research, we combined Artificial Intelligence Approaches, such as Evolutionary Algorithm (EA) with Genetic Algorithm (GA), and also used Exact Classifier Systems (here by Fuzzy C-Means (FCM)) to separate malignant from benign samples. Furthermore, we examined artificial Neural Networks (NN) to identify the model and structure. This research proposed a new algorithm for an accurate diagnosis of breast cancer.

Results: According to Wisconsin Data Base Cancer (WDBC) data base, 62.75% of samples were benign, and 37.25% were malignant. After applying the proposed algorithm, we achieved high detection accuracy of about "96.579%" on 205 patients who were diagnosed as having breast cancer. It was found that the method had 93% sensitivity, 73% specialty, 65% positive predictive value, and 95% negative predictive value, respectively.

If done by experts, Fine Needle Aspiration (FNA) can be a reliable replacement for open biopsy in palpable breast masses. Evaluation of FNA samples during aspiration can decrease insufficient samples. FNA can be the first line of diagnosis in women with breast masses, at least in deprived regions, and may increase health standards and clinical supervision of patients.

Conclusion: Such a smart, economical, non-invasive, rapid and accurate system can be introduced as a useful diagnostic system for comprehensive treatment of breast cancer. Another advantage of this method is the possibility of diagnosing breast abnormalities. If done by experts, FNA can be a reliable replacement for open biopsy in palpable breast masses. Evaluation of FNA samples during aspiration can decrease insufficient samples.

Keywords: Breast cancer; Fine needle aspiration; Artificial intelligent

Please cite this article as: Fiuzy M, Haddadnia J, Mollania N, Hashemian M, Hassanpour K. Introduction of a New Diagnostic Method for Breast Cancer Based on Fine Needle Aspiration (FNA) Test Data and Combining Intelligent Systems. Iran J Cancer Prev.2012; 5(4): 169-77.

1. Dept. of Bioengineering, Faculty of Electrical and Computer, Hakim Sabzevari University, Sabzevar, Iran
2. Dept. of Biology, Faculty of Basic Sciences, Hakim Sabzevari University, Sabzevar, Iran
3. Dept. of Clinical Sciences, Sabzevar University of Medical Sciences, Sabzevar, Iran

Corresponding Author:
Javad Haddadnia, Ph.D;
Associate Professor of Electronic, Artificial Intelligent Field
Tel: (+98) 57 14 00 35 90
Email: Haddadnia@sttu.ac.ir

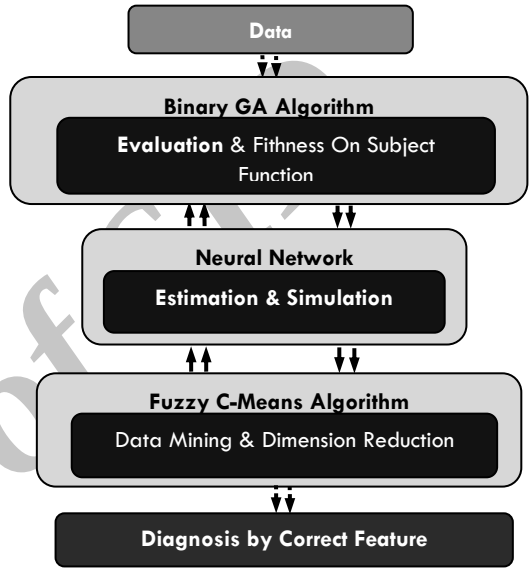
Received: 14 May 2012
Accepted: 22 Aug.2012
Iran J Cancer Prev 2012; 4:169-77

Introduction

Breast cancer is a type of cancer with high mortality rates among women, and it is one of the most common causes of death in women [1]. According to National Cancer Institute statistics in America, one out of eight women suffers from breast cancer and 6 % of all deaths worldwide are caused

by this type of cancer [2]. In Iran, approximately 7500 new cases of breast cancer are diagnosed each year. In 2005, standardized incidence rate of breast cancer was 23.56 and 0.66 in females and males, respectively. Its standardized mortality rate was 3.31 in females which was the 4th leading cause of females' death due to cancer [3, 4]. It should be noted that the age of breast cancer in Iran is one decade lower than developed countries [5]. Early diagnosis is one of the major factors in the treatment of this disease [6]. Early diagnosis of breast cancer (maximum 5 years after the first cancer cell division) will increase survival chances from 56% to 86% [3]. Thus, a precise and reliable system is essential for the timely diagnosis of benign or malignant breast tumors [3]. Diagnosis of this cancer is done by samples with the help of surgery, which has the highest recognition accuracy among available methods, but it is an aggressive, time consuming and expensive procedure [4]. Fine Needle Aspiration (FNA) test is another method that can accurately identify breast cancer by using machinery learning techniques. Using Machinery Learning techniques, many studies have been done on diagnosing breast cancer based on WDBC database (back ground). A group of researchers, using the neural network [10, 11, 12, 42], Support Vector Machine (SVM) classifier with the Radial Basis Function (RBF) kernel [7- 9] and a fuzzy classifier [10-12] attempted to resolve Wisconsin Breast Cancer issue, and reached about 95% accuracy in identifying breast cancer. A recent study shows that classification based on core is superior to other existing methods such as linear classification based on wavelet and neural networks [12]. In [16] 93% identifying accuracy was achieved using a combination of Decision Tree (DT) and Neural Network (NN) Algorithm, in [17] 94.4% identifying accuracy was achieved by combining GA and Decision Tree (DT). In [18] 93.6% identifying accuracy was reached by using expert systems. In [19] 94% identifying accuracy was obtained using genetic algorithms and artificial immune algorithm, in [20] 92% identifying accuracy was achieved using Fuzzy Principal Component Algorithm (FPCA) algorithm. In all references [21-24] by using the expert system in [19] 94% identifying accuracy was achieved. Artificial neural networks and fuzzy systems, using Principal Component Algorithm and Dimension Reduction Method or by their combination, attempted to identify breast cancer. However, they had a fairly weaker accuracy than other methods. Many machine learning systems, due to inability to remove waste and undesirable features have low accuracy. Therefore, in order to select the best features and achieve the highest accuracy, a new and efficient algorithm was proposed in this study for feature selection, classification of samples, identification and modelling. Firstly, in this paper, the general method is described. Then the outline from Flowchart 1 will be discussed. Therefore, it is hoped that the study results despite being short could be significant in improving methods, treatment and diagnosis of breast cancer.

Proposed Method and Algorithm



Flowchart 1. Proposed system

At first, data are extracted from the database, and then delivered to Binary GA with neural network in its cost function. Binary genetic algorithm, based on optimal features selection at database, generates optimal answers (0 or 1) due to the cost function. The network output is then delivered to the classification algorithm "FCM" for reducing dimension, estimating accuracy of binary genetic algorithm and correcting the classification of patient from impatient. The output of this algorithm returns to the network again and this process is repeated in the cost function of genetic binary algorithm until the most optimal features for increasing maximum accuracy of feature selection are achieved. Flowchart 1 represents all stages in this study.

Materials and Methods

The FNA test involves extracting fluid from the breast tissue, and a visual examination of this sample under a microscope [8]. In smart detection, at first, this image turns to a gray level (it does not require

color info). Then, the related software defines the cell core boundary based on the image processing techniques, and calculates the features for each core such as radius, texture, perimeter, area, compaction (square perimeter divided by area), flat (mean difference in length of lines radially adjacent), concavity, symmetry and fractal (border kernel of approximation coastline [27]). Finally, it calculates the mean square error and the mean of the three largest achieved values. Respectively, 30 features by factual values are gained for each sample. WDBC database is the data set involving a few FNA tests provided in the above description, which have been performed on 569 patients in Wisconsin hospitals and identified 212 malignant and 357 benign cases [28]. One sample of FNA test is shown in Figure 1.

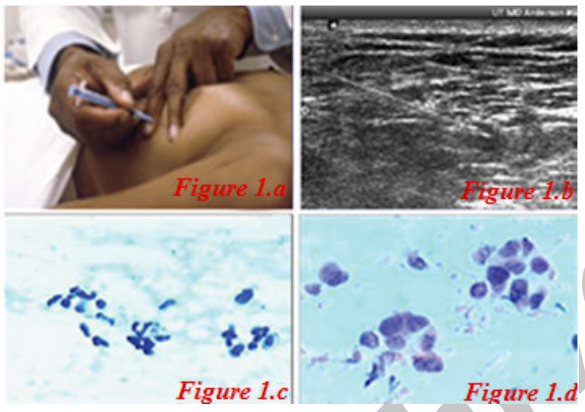


Figure 1. FNA Test and Samples, Figure1 a. Biopsy, Figure 1 b By Ultrasound, Figure1 c. Benign Sample, Figure1 d. Malignant Sample

Pattern Recognition

In fact, feature selection is the choice of features that have maximum power at output prediction [29]. To solve the problems which depend on optimal subset [26], feature selection algorithms are divided in to two major categories. If features selection is done independent of any type of learning algorithm, it will be called Filter method, in which, the result of the selected features are determined before processing. If the assessment process is associated with a classified algorithm, the method of feature selection will be called Wrapper or Closed loop. This is a common pattern recognition system which consists of 4 sections: feature extraction and selection; designing and training classifier; and testing. In this study, feature extraction and testing by neural networks selected best features by GA algorithm, and FCM classifier was used for class and sample classification.

Genetic Algorithm

Genetic algorithm was developed for the first time by John Holland [30]. During 1960-1970 [31], the genetic algorithm was used to find approximate answers in optimization and search problems derived from natural concepts like: heredity, mutation and recombination. The major objective in this algorithm is the chance to have good samples, to continue their life in the next generation, and improve the answers. The procedure of algorithm device is determined in pseudo – code [31]:

Procedure: Applied Genetic Algorithm
  Initialization
    t = 0;
    Set population size or pop size, number of generation or max_gen, probability of crossover or p_c and probability of mutation or p_m
    Initialize parent population ***P (t)***:
    Evaluate *P (t) and select the best solution* σ^* *with the* Optimum objective function among ***P (t)***
    While (no termination criteria) **do**
    Regenerate C (t) from P (t) by applying the crossover And mutation operations:
    Evaluate C (t) and select the current best solution σ With the minimum objective value among C (t)
    Update the best solution σ^* solution, I, e, if $\sigma < \sigma^*$, **then** $\sigma^* = \sigma$;
    Select P (t+1) from P (t) and C (t);
    t = t+1;
  End while;
End procedure

Here, the genetic algorithm was used from n-gene chromosomes to illustrate the problem space based on 30 features, for example, as a flowing vector. Many genes are "1" (effective in generation or effective features) and many are "0" (uneffective in generation or uneffective features).

0 1 0 1 1 1 0 1 0 1 1 0 1 0 1 0 1 0 1 0 1 0 1 1 1 1 0 1 0 1

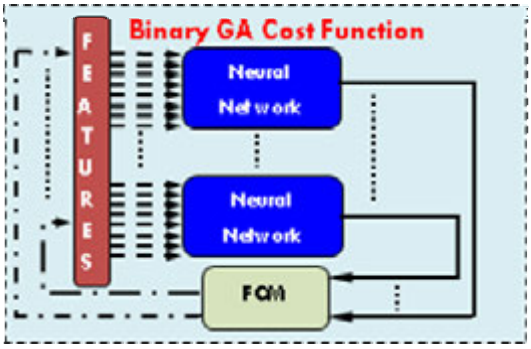
The rotating selection (roulette wheel) is used to select chromosomes to combine and generate [32]. The method applied for mutation is in the inverse way. In this method, a sub-collection of chromosome's gene is selected and then it is reversed in replacement. To combine chromosomes with the above mentioned structure, the Partially Mapped cross over (PMX) method is used [33]. The objective function is the equation that gets merits for each generation. It seems that accuracy in identifying is more important than the small selected subsets. Although the two series have the same accuracy, the smaller set is preferred, so we suggest the following fitness function.

$$Fitness = \alpha . Accuracy + \beta . \frac{|n| - |s|}{|n|} \qquad \text{(Eq.1)}$$

Which |n| is the number of the total features and |s| is the number of selective features. First sentence is the carefully identified coefficient and the second sentence is the rate coefficient of reduced features.

Table 1. GA Properties

ALG	Binary GA	POPsize	200
Maxiter	250	Varlow	-10
Var high	+10	Max generation	600
Recom Percent	0.1	Mut Percent	0.4
Cross Percent	0.5	Selection	Roulet



Flowchart 2. GA Cost Function

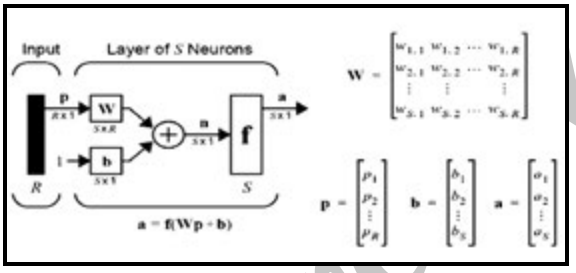


Figure 2. MLP Neural Network [Matlab® Help]

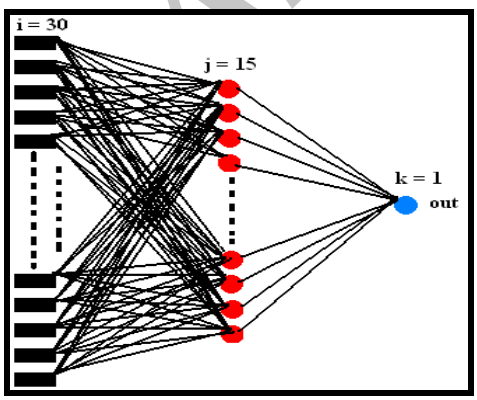


Figure 3. Proposed Neural Network (NN)

We supposed that the sum of α and β coefficient is fixed and equal to 100. Now, in regard to importance of the identified accuracy and fewer features used in the diagnosis, α and β coefficients are set. Undoubtedly, detection accuracy is more important in this problem, and therefore, α (here 99) will be greater than β (here 1) [26]. Genetic algorithm parameters are set according to Table 1.

Clearly, GA introduces the number of the subset which has the most and useful information. Respectively, GA searches in the 2- power 30 subsets to find the optimal solution. The suggested algorithm located in the genetic algorithm cost function is observed in Flowchart 2.

Multi layer progressive algorithm located in genetic algorithm cost function

Progressive (feed forward) multi-layer neural networks are composed of many processing elements that are in connection with each other, which are called neurons [34]. Generally, the input neurons (P) are added and their weights are multiplied (w), and applied by bias (b) (width from sources), and result in activation function (f), and then the results (a) will be transferred to the next layer (Figure 2). Feed Forward Multi-Layer Neural Network is one of the most important types of static neural networks, and used in many applications such as system identification problems, control, etc [34]. Error Back Propagation algorithm or BP model was introduced as a training algorithm for network by this model.

This algorithm provides a solution for reducing gradient to minimize errors in the network. In this project, we used 3 layers of feed forward network with sigmoid activation function for input layer (i) and output layer (k) with linear function for input layer (i) Figure 3 with adoption of [34].

At first, we selected 30 neurons for input in [26]. Then, we selected the middle layer of network for the best selection of the structure. Then, to achieve the best result, 15 neurons were selected in the middle layer. The number selection of the middle layer neurons is very important. If their number is low, Network will be faced with a shortage of learning resources for solving complex and nonlinear problems and if their number is high, the two following problems are caused: first, learning takes much time. Second, unimportant training data may be learned and the system network would be weak to solve the problems [35, 41, 42]. In this method, the error is passed to the back. In each layer, necessary reforms are carried out on weights. In this process, the quite long error coverage reaches to minimum amount or Error Global coverage becomes equal to times of maximum repetition, which is pre-determined. Error Global should be set so that it

Table 2. Neural Network (NN) Properties

Neurons	30/ 15 /1
Input / Output	30/1
Learning Algorithm	BP
MSE	00.02
Error Global	00.02
Maximum iteration	1000
1th Activation Function	Sigmoid
2th Activation Function	Linear

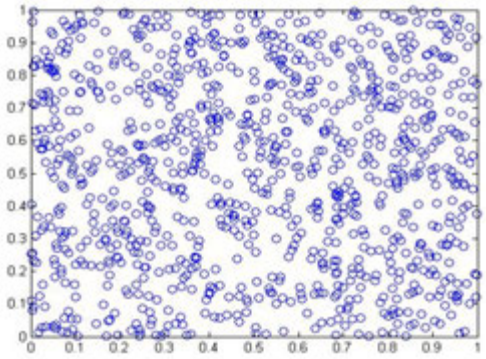


Figure 4. 1000 Rand Data before Clustering

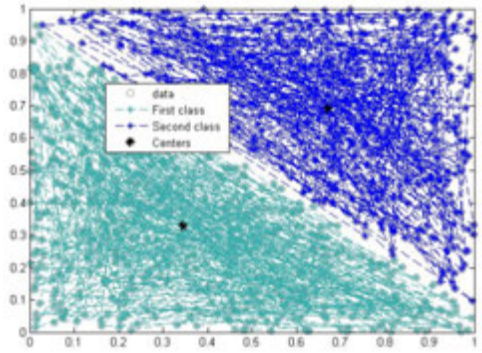


Figure 5. Two Cluster By 1000 Rand data

avoids the Network from learning. In this project, 250 has been considered as the number of repetition. Error global value has also been determined 0.02. Properties of the suggested neural network are expressed in Table 2.

FCM or fuzzy c-means algorithm

In 1969, Ruspini, proposed the first clustering model with fuzzy idea [36]. In this method, the amount of any membership or data belonging to each data to any cluster in matrix member is determined.

$$U = \lfloor u_{i,j} \rfloor_{c \times n} = (\vec{u}_1, \vec{u}_2, ..., \vec{u}_n) \quad \text{(Eq.2)}$$

Which, c is number of clusters and n is the data number. This method suffers from two main limitations: the first limitation was that no cluster should be empty.

$$(\sum_{j=1}^{u} u_{ij} > 0 \quad \forall_i \in \{1, ..., c\}) \quad \text{(Eq.3)}$$

The second limitation, called normalization constraints, was that the total membership of all clusters must be equal to "1" in each data class.

$$(\sum_{j=1}^{u} u_{ij} = 1 \quad \forall j \in \{1, ..., n\}) \quad \text{(Eq.4)}$$

FCM tries for any data set, finds the parts that minimize the following cost Equation 1 or objective function.

$$J_f(X, U_f, C) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^m d_{ij}^2 \quad \text{(Eq.5)}$$

Where in it, d_{ij} is data distance between X_i and the i^{th} cluster center. $m \in [1,\infty)$ is degree of fuzziness. If "M" goes to one, clustering will be more difficult or crisp. On the contrary, if "M" goes to infinity, clustering will be fuzzier. Supposedly as demonstrated in Figure 4 we supposedly classified 1000 random data in 2 classes.

Fuzzy –C- Means or "FCM" methods consist of 4 stages:

1. If function cannot be minimized directly, repeat algorithm could be used. To solve this problem, the optimal replacement scheme was used as follows: select the proper values for m, C and small positive number for "ε". The matrix C is randomly filled (middle or center of clusters) finally, set t = 0.

2. In (t = 0) membership matrix is calculated, and in (t > 0) updated membership matrix is determined. This means that the degree of membership for fixed parameters of clusters is optimized such as Equation 5 in the flowing:

$$u_{ij}^{(t+1)} = \frac{d_{ij}^{-2/(m-1)}}{\sum_{i=1}^{c} d_{ij}^{-2/(m-1)}} = \frac{1}{\sum_{l=1}^{c} (\frac{d_{ij}}{dij})^{1/(1-m)}}$$

$$for \quad i = 1,...c \quad and \quad j = 1,..... N \quad \text{(Eq .6)}$$

3. The final step is updating the center of the clusters with optimized membership matrix. In addition to these parameters, how the distance is measured is of prime importance.

4. Repeating Steps 2 and 3 until $\left\|C^{(t+1)} - C^{(t)}\right\| < \varepsilon$; or $\left\|U^{(t+1)} - U^{(t)}\right\| < \varepsilon$ by all above steps are applied on the random test data (Figure 4) and the second cluster is created. The result can be observed in Figure 5.

Shown in practice researches, this method does not get stuck on local peaks, and today FCM probable method is used as an inceptor in many clustering ways. More details exist on this method [36- 40].

Results

MATLAB® is used for simulation. In this implementation, trained neural network is performed at one time. All weights will change for each input or for one row of data collection. The gained results are presented in Table 3. There were 202 patients during the evaluation process in [7], in the second evaluation [10-12] there were 202 patients, 198 patients in evaluation in [16], in evaluation [17] 201 patients, in evaluation [18] approximately 194 patients, in evaluation [19] 199 patients, in evaluation [20] 195 patients. Approximately 204 patients were identified in the proposed method.

Table 3. Compare and Result

Patient identify	Best answer	Accuracy test	Method
202	Not Available	95.61%	SVM [7]
202	---	> 95%	RBF+FCM and SVM [10,11,12]
~198	---	> 93%	DT + NN [16]
~201	---	94.4%	GA + DT [17]
~194	---	> 93.6%	Exp Sys [18]
~195	---	> 94%	GA+ AI [19]
~195	---	> 92%	FPRCA [20]
~ 205	96.579%	96 %	Proposed algorithm

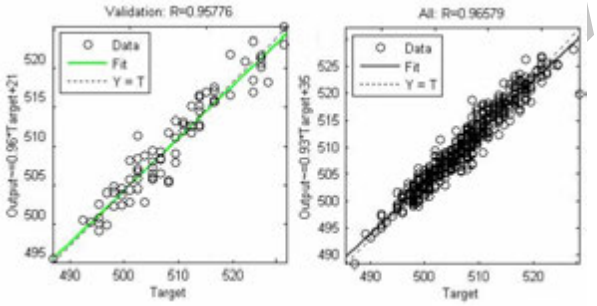


Figure 6. Regression plot for output and target in validation and final result (train, test, valid) or all

Therefore, the suggested method is precise and optimal. The best and most effective selected features in this suggested method are obtained in the following vector: this vector represents seven effective features (3, 6, 12, 18, 21, 25, 27) from all the features involved in the diagnosis of breast cancer.

Figure 6 includes the obtained results in the final stage drawn in terms of factual result.

To review the performance of the proposed system, the below parameters are defined which have an important role in the final decision:

Tp = True positive = Cytological and suspicious positive diagnosis, which are positive in pathological tests.

Fp = False positive = Cytological and suspicious positive diagnosis, which are negative in pathological tests.

Tn = True negative = Cytological and suspicious negative diagnosis, which are negative in pathological tests.

Tp = True positive = Cytological and suspicious positive diagnosis, which are positive in pathological tests.

$$\text{Sensitivity} = \frac{T_p}{T_p + F_n} = 93\% \quad \text{(Eq.7)}$$

$$\text{Specialty} = \frac{T_n}{T_n + F_p} = 73\% \quad \text{(Eq.8)}$$

$$\text{Positive Predict value} = \frac{T_p}{T_p + F_p} = 65\% \quad \text{(Eq.9)}$$

$$\text{Negative Predictive value} = \frac{T_n}{T_n + F_n} = 95\% \quad \text{(Eq.10)}$$

$$\text{Accuracy} = \frac{T_{p+}T_n}{total} = 96\% \text{ number of true diagnosis} \quad \text{(Eq.11)}$$

For better comparison of the proposed algorithm and the medical method, their results are expressed in Table 4.

Table 4. Table for Comparing MD (Medical result) method with proposed ALG

	Medical result (MD)	Proposed ALG
SEN (%)	~99	93
SPE (%)	~99	73
ACC (%)	~99	96
PPv (%)	85	65
NPv (%)	98	95

The value of the diagnostic methods is in their ability to diagnosis the disease. Two important parameters are defined: Sensitivity or cytological sensitivity in malignancy diagnosis, and Specificity or cytological specificity in malignancy diagnosis. Given the sensitivity and specificity values presented in this review, the obtained cases are reliable. Finally, for comparing the proposed and actual values diagnosed by a physician, the Receiver Operating Characteristics (ROC) Chart may be used. In this chart, the chart with more areas indicates that the

system has better performance. In each of the 2 charts of the proposed system and real results (MD system), we used 2 new parameters: True Positive Fraction: Expresses the proportion of false positives to benign or

$$TPF = \frac{F_p}{A}, \quad \text{(Eq.12)}$$

and False Positives Fraction: Expresses whether the ratio of false negative is a true malignancy or

$$FPF = \frac{F_n}{B}. \quad \text{(Eq.13)}$$

This chart is displayed in Figure 7.

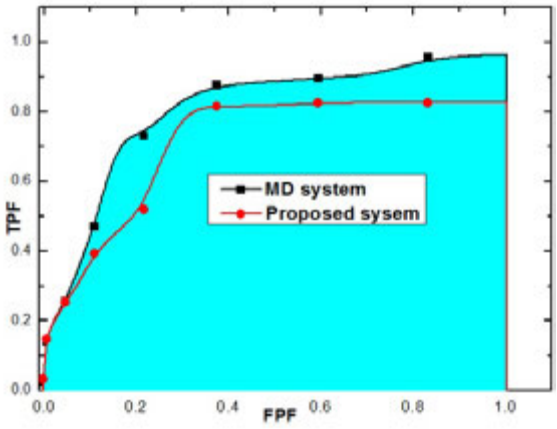


Figure 7. Receiver Operating Characteristics (ROC) chart between proposed algorithm and medical system

Given the two graphs, it is clear that the proposed system has a good performance. The results of the system are close to the real results.

Coclusion

Breast FNA is a quick test, which takes 10 to 20 seconds for each sample, and this procedure may be repeated several times until the doctor is convinced that a good sample is collected. The examination and FNA procedure will generally take around 20 to 30 minutes for ultrasound guidance and 45-60 minutes for mammographic guidance. Breast FNA should be done after careful medical examination and imaging tests such as mammogram or ultrasound as part of the "triple test" to ensure the correct diagnosis is made. Breast FNA is done by specialist doctors experienced in breast needle biopsy procedures. Breast cancer is the most common disease that leads to the women's deaths around the world. Due to necessity of early and timely disease diagnosis, a new method based on a combination of intelligent systems is presented in this research. Many researchers are interested in using this tool, but the challenge of neural networks training is their most important concern. Combination of evolutionary algorithms and data mining systems could be used to train these networks to identify and select properties, and therefore lead to their accurate performance. Simulation results show that using data mining process can increase accuracy and efficiency of the network training. The main goal of this research is to achieve 100% accuracy, so to guarantee those artificial intelligence systems and furnish them with practical aspects, we need adequate knowledge, and satisfactory control over intelligent systems, data mining, and breast cancer.

Acknowledgment

We would like to thank Dr. Mohammad Esmaeil Akbari, the Head of Cancer Research Center of Shahid Beheshti University of Medical Sciences for his help and support.

Conflict of Interest

The authors have no conflict of interest in this article.

Authors' Contribution

Mohammad Fiuzy conceived, designed the study, interpreted the results, drafted the manuscript and carried out the data analyses. Javad Haddadnia, Maryam Hashemian, Kazem Hassanpour; participated in writing and revising the manuscript, while Javad Haddadnia revised the final manuscript. All authors read and approved the final manuscript.

References

1. Horner MJ, Ries LAG, Krapcho M, Neyman N, Aminou R, Howlader N, et al. SEER Cancer Statistics Review, 1975-2006, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2006/.

2. Wun LM, Merril RM, Feuer EJ. Estimating lifetime and age-conditional probabilities of developing cancer. Lifetime Data Anal. 1998; 4(2): 169-86.

3. Akbari ME and co-workers, Iran Cancer Report. 1st edition, Cancer Research Center .Shahid Beheshti University of Medical Science, 2008.

4. Hashemzadeh SH, Kumar PV, Malekpour N, Hashemi Z, Fattahi F, Malekpour F. Diagnostic Accuracy of Fine Needle Aspiration Cytology: Comparison of Results in Tabriz Imam Khomeini Hospital and Shiraz University of Medical Sciences. Irania Journal of Cancer Prevention. 2009; 3: 133-6.

5. Harirchi I, Karbakhsh M, Kashefi A, Momtahen AJ. Breast Cancer in Iran: results of a multicenter

study. Asian Pakistan jornal of Cancer Prev, Pakistan. 2004; 5(1): 24-70.

6. Kazerooni I, Ghayumizade H, Haddadnia J. Introduction of an Intelligent System for the Accurate Determination of Masses in Mammographic Images. IJBD, Iran .2009; 2(3, 4).

7. Choi YD, Choi YH, Lee JH, Nam JH, Juhng SW, Choi C. Analysis of Fine Needle Aspiration cytology of the breast: a review of 1,297 cases and correlation with histologic diagnoses. Acta Cytol. 2004; 48(6):801-6.

8. Maglogiannis I, Zafiropoulos E, Anagnostopoulos I. An intelligent system for automated breast cancer diagnosis and prognosis using SVM based classifiers. Journal of Appiedl Intelligence. 2007; 30(1): 24-36.

9. Wang Y, Wan F. Breast cancer diagnosis via support vector machines. Proc. 25th Chinese Control Conf. 2006: 1853-6.

10. Yang Z, Lu W, Yu R. Detecting false benign in breast cancer diagnosis. IEEE NNS-ENNS International, Joint Conf, Neural Networks. 2000; 3: 3655-8.

11. Mu T, Nandi A. Breast cancer detection from FNA using SVM with different parameter tuning systems and SOM-RBF classifier. J Franklin Institute. 2007; 344: 285-311.

12. Fuentes-Uriarte J, Garcia M, Castillo O. Comparative study of fuzzy methods in breast cancer diagnosis. Annual Meeting of the NAFIPS. 2008: 1-5.

13. Yang Lwei Y, Nishikwa R, Jiang Y. A study on several machine learning methods for classification of malignant and benign clustered micro calcification. IEEE Trans Med Imaging. 2005; 24(3): 371-80.

14. Sewak M, Vaidya P, Chan C, Duan Z. SVM approach to breast cancer classification. 2nd International, Multi-Symp, Computer and Computational Sciences .2007: 32-7.

15. Anagnostopoilosi I, Anagnostopoilosi C, Vergados D, Rouskas A, Kormentzas G. The Wisconsin Breast Cancer Problem: Diagnosis and TTR/DFS time prognosis using probabilistic and generalized regression information classifiers. J Oncol Reports .2006; 15: 975-81.

16. Jose´ M, Jerez-Aragone´sa, Jose´ A Go´mez-Ruiza, Gonzalo Ramos-Jime´neza, Jose´ Mun˜oz-Pe´reza. A combined neural network and decision trees model for prognosis of breast cancer relapse. Journal of Artificial Intelligence in Medicine. 2003; 27 (1): 45-63.

17. Chin-Yuan Fana, Pei-Chann Changb, Jyun-Jie Linb, JC Hsiehb. A hybrid model combining case-based reasoning and fuzzy decision tree for medical data classification. Applied Soft Computing. 2011; 11(1): 632-44.

18. Murat Karabatak M. An expert system for detection of breast cancer based on association rules and neural network. Expert Systems with Applications. 2009; 36 (2): 3465 - 9.

19. Seral, Sahana Kemal Polata Halife Kodazb, Salih Gune. A new hybrid method based on fuzzy-artificial immune system and k-nn algorithm for breast cancer diagnosis. Computers in Biology and Medicine. 2007; 37 (3): 415 -23.

20. Pasi Luukka. Classification based on fuzzy robust PCA algorithms and similarity classifier. Expert Systems with Applications.2009; 36 (4): 7463-8.

21. Boris Kovalerchuk, Evangelos Triantaphyllou, James F Ruiz, Jane Clayton. Fuzzy logic in computer-aided breast cancer diagnosis: analysis of lobula. Artificial Intelligence in Medicine.1997; 11 (1): 75-85.

22. PJG Lisboa, H Wong, P Harris, R Swindell. A Bayesian neural network approach for modeling censored data with an application to prognosis after surgery for breast cancer. Artificial Intelligence in Medicine.2003; 28 (1): 1-25.

23. EYK Ng, U Rajendra Acharya, Louis G Keith, Susan Lockwood. Detection and differentiation of breast cancer Using neural classifiers with first warning thermal sensors. Information Sciences.2007; 177 (20):4526-38.

24. David B, Fogel, Eugene C, Wasson IIP, Edward M Boughtonc. Evolving neural networks for detecting breast cancer. Cancer Letters. 1995; 96(1); 4 9-53.

25. Khooei AR, Mehrabi Bahar M, Ghaemi M, Mirshahi M. Sensitivity and specificity of CNB in diagnosis of breast masses. Iranian journal of Basic Medical sciences. 2005; 8(2): 6-100.

26. Alipoor M, Haddadnia J. Introduction of an intelligent system for accurate diagnosis of breast cancer. IJBD. 2009; 2 (2): 33-40.

27. Kim A, Lee J, Choi JS, Won NH, Koo BH. Fine needle aspiration cytology of the breast. Experience at an outpatient breast clinic. Acta Cytol. 2000; 8 (3):361-7.

28. http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+ (Diagnostic).

29. Jensen R. Combining rough and fuzzy sets for feature selection. Ph.D Thesis, School of informatics, University Edinburgh, 2005.

30. Joaquen Abelln, Andrés R, Masegosa. An ensemble method using credal decision trees. European Journal of Operational Research. 2010; 96 (1): 218-26.

31. Chiung Moon, Jongsoo Kim, Gyunghyun Choi, Yoonho Seo. An efficient genetic algorithm for the traveling salesman, problem with precedence

constraints. European Journal of Operational Research. 2002; 140 (3): 606-617.

32. Ron Shelly. Roulette Wheel Study (1998). http://en.wikipedia.org/wiki/Roulette#History.

33. Galiasso, Pablo , Wainwright, Roger L. A Hybrid Genetic Algorithm for the Point to Multipoint Routing Problem White Single Splite Paths. Proceedings of ACM/SIGAPP Symposium on Applied Computing. (SAC'01). 11-14 March 2001: 327-32.

34. Madan M, Gupta, Liang Jin, Noriyasu Homma. Static and Dynamic Neural Networks from fundamental to advanced theory. Forwarded by Lotfi A, Zadeh. IEEE Press. 2003 .

35. Nirooe Mehyar, Parviz Ab, Giti M. Simulating a Hybrid Model By GA and NN for Segregation patterns of benign and malignant breast cancer in mammography. Bio Phsics Iranian Journal. 2006; 13 (3): 25- 33.

36. Baraldi A, Blonda P. A Survey of Fuzzy Clustering, Algorithms for Pattern Recognition, Part I and II. IEEE TRANSACTIONS ON SYSTEMS. MAIN, AND CYBERNETICS PART B: CYBERNETICS.1999; 29 (6): 786-801.

37. Jain AK, Murty MN, Flynn PJ. Data Clustering. A Review. ACM Computing Surveys. 1999; 31 (3): 264-323.

38. Rui Xu , Donald Wunsch. Survey of Clustering Algorithm. IEEE TRANSACTIONS ON NEURAL NETWORKS. 2005; 30 (14): 2826-41.

39. Dring C, Lesot MJ, Kruse R. Data analysis with fuzzy clustering methods. Computational Statistics & Data Analysis. 2006; 51 (1): 192-214.

40. Asgarian E, Memarzadeh H, Saryani M, Jafar H. A new View On Clustering By GA. 14th CSICC 2009.

41. Abdi N, Amirahmadi N, Tahami E, Rbbani. Diabet Diagnisis By SVM. 14th ISCEE Iran. summer 2011: 25-31.

42. Zhaohui L, Xiaoming W, Shengwen G, Binggang Y. Diagnosis of breast cancer tumor based on manifold learning and support vector machine. IEEE Internation Conf of Information and Automation. 2008; 703-707.