Original Article

Breast Cancer Recognition Using a Novel Hybrid Intelligent Method

Jalil Addeh, Ata Ebrahimzadeh

Department of Electrical and Computer Engineering, Babol University of Technology, Babol, Iran

ABSTRACT

Breast cancer is the second largest cause of cancer deaths among women. At the same time, it is also among the most curable cancer types if it can be diagnosed early. This paper presents a novel hybrid intelligent method for recognition of breast cancer tumors. The proposed method includes three main modules: the feature extraction module, the classifier module, and the optimization module. In the feature extraction module, fuzzy features are proposed as the efficient characteristic of the patterns. In the classifier module, because of the promising generalization capability of support vector machines (SVM), a SVM-based classifier is proposed. In support vector machine training, the hyperparameters have very important roles for its recognition accuracy. Therefore, in the optimization module, the bees algorithm (BA) is proposed for selecting appropriate parameters of the classifier. The proposed system is tested on Wisconsin Breast Cancer database and simulation results show that the recommended system has a high accuracy.

Key words: *Bees algorithm, breast cancer, fuzzy clustering, support vector machine, Wisconsin*

INTRODUCTION

Breast cancer is the first current cancer and the second element of death among women. In 2010, there were reported approximately 207,090 newly diagnosed cases and 30,840 deaths in the United States.[1] Since the cause of breast cancer is unknown, the methods for preventing this disease are not specified, thus really recognizing the existence of tumor and the type of cancerous tumor would have a very important role in getting decision of doctors for applying the methods of true treatment and therefore reclaiming the life of people (more than 40%).[2]

In recent years the mammography method has been used widely for early recognizing of cancerous tumor.[3] Usually mammography is away for detecting the presence of cancerous tumors; however, determining the type of the tumor is much more challenging. Some of the characteristics of malignant tumors are clustered calcification, isolated ducts, poorly defined mass, etc. Experts (doctors) physically look at mammograms to detect deformation that may be taken as an indicator of cancerous changes.[4] It is clear that these methods of recognition due to human mistakes and errors of medical devices are not reliable methods. Many investigators believe that automation of

mammogram screening analysis increases the rate of early detection.

With this aim, several approaches have been proposed for breast cancer recognition. Some of the researchers used the expert systems.[5,6] The advantage of an expert system or a rule-based system is that it contains the information explicitly. If required, the rules can be modified and updated easily. However, the use of rules based on statistical properties has the difficulty that similar statistical properties may be derived for some patterns of different classes, which may create problems of incorrect recognition. Also, artificial neural networks (ANNs) have been widely applied for breast cancer recognition. ANNs can be simply categorized into two groups comprising supervised and unsupervised. Most researchers[7-10] have used supervised ANNs, such as multilayer perceptron (MLP), radial basis function (RBF), and learning vector quantization (LVQ), to classify breast cancer. Furthermore, unsupervised methods, e.g., self-organized maps (SOM), have been applied to fulfill the same objective in other studies.[11] The advantage with neural network is that it is capable of handling noisy measurements requiring no assumption about the statistical distribution of the monitored data. It learns to recognize patterns directly through typical example patterns during a training phase.

One disadvantage with neural network is the difficulty in understanding how a particular classification decision has been reached and also in determining the details of how a given pattern resembles with a particular class. In addition, there is no systematic way to select the topology and architecture of a neural network. In general, this has to be found empirically, which can be time consuming. Some of the researchers used the support vector machine to breast cancer recognition.[12,13] Using Support Vector Machines (SVMs) is the method that is receiving increasing attention, with remarkable results recently. The main difference between ANNs and SVMs is the principle of risk minimization. An ANN implements empirical risk minimization to minimize the error on the training data, whereas an SVM implements the principle of structural risk minimization in place of experiential risk minimization, which makes it have excellent generalization ability in the situation when there is a small sample.

In designing of computer-aided diagnostic (CAD) system, the most important is the integration of suitable feature extraction and pattern classifier such that they have can operate in coordination to make effective and efficient CAD system. This paper is proposed using fuzzy c-mean (FCM) clustering algorithm to make a SVM system more effective. The structure of the proposed system is composed of two subnetworks: fuzzy classifier and SVM. The fuzzy self-organizing layer performs the preclassification task and the following SVM works as the final classifier. The fuzzy stage is responsible for analysis of the data distribution and grouping them into clusters with different membership values. Based on these membership values, the SVM classifies the applied input vector. By this act, a number of segments in training set are reduced using FCM clustering in self-organizing layer before inputs are presented to SVM.

The largest problems encountered in setting up the SVM model are how to select the kernel function and its parameter values. The parameters that should be optimized include the penalty parameter (C) and the kernel function parameters such as the value of gamma (γ) for the radial basis function (RBF) kernel. Turning back to breast cancer recognition systems, it can be found that the selection of the best free parameters of the adopted classifier is generally done empirically. On the other hand using of SVM has some difficulties, which are how to select the optimal kernel function type, most appropriate hyperparameters values for SVM training and testing stages. Therefore in this study, we used an efficient optimizer for finding the optimum values of hyperparameters, i.e., the kernel parameter and classifier parameters.

The rest of paper is organized as follows. The next section presents the Wisconsin database. The Needed Concepts section describes the concepts needed, including the feature extraction, the support vector machine and optimization algorithm concepts. The General Structure of the Proposed Method section presents the proposed model. The Simulation Results section shows some simulation results and finally the last section concludes the paper.

WISCONSIN BREAST CANCER (WBC) DATABASE

Breast cancer is the most common cancer among women; excluding nonmelanoma skin cancers. This cancer affects one in eight women during their lives. It occurs in both men and women, although male breast cancer is rare. Breast cancer is a malignant tumor that has developed from cells of the breast. Although scientists know some of the risk factors (e.g., aging, genetic risk factors, family history, menstrual periods, not having children, obesity) that increase a woman's chance of developing breast cancer, they do not yet know what causes most breast cancers or exactly how some of these risk factors cause cells to become cancerous. Research is underway to learn more and scientists are making great progress in understanding how certain changes in DNA can cause normal breast cells to become cancerous. In this study, the Wisconsin Breast Cancer (WBC) database was used and analyzed. They have been collected by Dr. William H. Wolberg (1989–1991) at the University of Wisconsin–Madison Hospitals. There are 699 records in this database. Each record in the database has nine attributes. The nine attributes detailed in Table 1 are graded on an interval scale from a normal state of 0.1–1, with 1 being the most abnormal state. In this database, 241 (34.5%) records are malignant and 458 (65.5%) records are benign.[8]

NEEDED CONCEPTS

Fuzzy C-Mean Clustering Algorithm and Fuzzy Features

Fuzzy C-Mean (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters. This method was developed by Dunn[14] and improved by Bezdek[15] FCM is frequently used in pattern recognition. It is based on minimization of the following objective function:

Table 1: Wisconsin breast cancer data description of attributes

Attribute description	Value of attributes	Mean	Standard deviation
Clump thickness	0.1-1	0.442	0.282
Uniformity of cell size	0.1-1	0.313	0.305
Uniformity of cell shape	0.1-1	0.320	0.297
Marginal adhesion	0.1-1	0.280	0.286
Single epithelial cell size	0.1-1	0.321	0.221
Bare nuclei	0.1-1	0.346	0.364
Bland chromatin	0.1-1	0.343	0.244
Normal nucleoli	0.1-1	0.287	0.305
Mitoses	0.1-1	0.159	0.171

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^m \left\| x_i - c_j \right\|^2, \qquad 1 \le m < \infty \qquad (1)$$

where m is any real number greater than 1, u_{ij} is the degree of membership of x_i in the cluster j, x_i is the ith of d-dimensional measured data, c_j is the d-dimensional center of the cluster, and $\|.\|$ is any norm expressing the similarity between any measured data and the center. Fuzzy partitioning is carried out through an iterative optimization of the objective function shown above, with the update of membership u_{ij} and the cluster centers c_j given by:

$$u_{ij} = \cfrac{1}{\displaystyle\sum_{k=1}^{C} \left(\cfrac{\left\| x_i - c_j \right\|}{\left\| x_i - c_k \right\|} \right)^{\frac{2}{m-1}}} \qquad (2)$$

$$c_j = \cfrac{\displaystyle\sum_{i=1}^{N} u_{ij}^m . x_i}{\displaystyle\sum_{i=1}^{N} u_{ij}^m}. \qquad (3)$$

This iteration will stop when $max\left\{ \left| u_{ij}^{k+1} - u_{ij}^k \right| \right\} \prec \varepsilon$, where ε is a termination criterion between 0 and 1, and k is the iteration step. The algorithm is composed of the following steps:

1. *Initialize $U = [u_{ij}]$ matrix, $U^{(0)}$*
2. *At k-step: calculate the centers vectors $C^{(k)} = [c_j]$ with $U^{(k)}$*

$$c_j = \cfrac{\displaystyle\sum_{i=1}^{N} u_{ij}^m . x_i}{\displaystyle\sum_{i=1}^{N} u_{ij}^m}$$

3. *Update $U^{(k)}$, $U^{(k+1)}$*

$$u_{ij} = \cfrac{1}{\displaystyle\sum_{k=1}^{C} \left(\cfrac{\left\| x_i - c_j \right\|}{\left\| x_i - c_k \right\|} \right)^{\frac{2}{m-1}}}$$

4. *If $\left\| U^{(k+1)} - U^{(k)} \right\| < \varepsilon$, then STOP; otherwise return to step 2.*

In this study, we used the membership values (fuzzy features) as the input vector. The membership values of each tumor for both two groups illustrated in Figure 1. As depicted in this figure, the amount of fuzzy features for benign and malignant tumors are completely different, so separation will be easier.

Support Vector Machine

Linear SVM
Consider the problem of separating the set of training vectors belonging to two linearly separable classes,
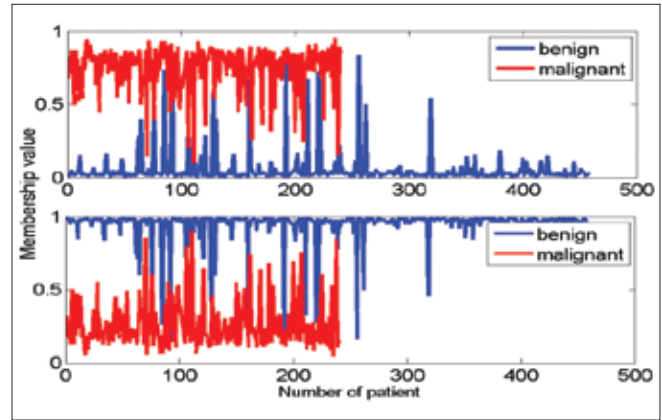

Figure 1: Fuzzy features for WBC dataset: (a) calculated based on malignant cluster, (b) calculated based on benign cluster.

$$(x_i, y_i), \quad x_i \in R^n, \quad y_i \in \{+1, -1\}$$
$$i = 1, 2, \dots, n \qquad (4)$$

where x_i is a real-valued n-dimensional input vector and y_i is a label that determines the class of x_i. A separating hyperplane is determined by an orthogonal vector w and a bias b, which identifies the points that satisfy

$$w.x + b = 0 \qquad (5)$$

The parameters w and b are constrained by

$$\min_i |w.x_i + b| \ge 1. \qquad (6)$$

A separating hyperplane in canonical form must satisfy the following constraints:

$$y_i(w.x_i + b) \ge 1, \quad i = 1, 2, \dots, n \qquad (7)$$

The hyperplane that optimally separates the data is the one that minimizes,

$$\Phi(w) = \frac{1}{2}(w.w). \qquad (8)$$

Relaxing the constraints of (7) by introducing slack variables $\xi_i \ge 0$, $i = 1, 2, \dots, n$, (7) becomes

$$y_i(w.x_i + b) \ge 1 - \xi_i, \quad i = 1, 2, \dots, n. \qquad (9)$$

In this case, the optimization problem becomes

$$\Phi(w, \xi) = \frac{1}{2}(w.w) + C\sum_{i=1}^{n} \xi_i \qquad (10)$$

with a user-defined positive finite constant C. The solution to the optimization problem in (10), under the constraints of (9), could be obtained in the saddle point of Lagrangian function:

$$L(w,b,\alpha,\xi,\gamma) = \frac{1}{2}(w.w) + C\sum_{i=1}^{n}\xi_i -$$

$$\sum_{i=1}^{n}\alpha_i\left|y_i(w.x_i+b)-1+\xi_i\right| - \sum_{i=1}^{n}\gamma_i\xi_i. \qquad (11)$$

where $\alpha_i \geq 0$, $\xi_i \geq 0$, $i=1,2,...,n$ are the Lagrange multipliers. The Lagrangian function has to be minimized with respect to $w, b,$ and ξ_i. Classical Lagrangian duality enables the primal problem, (11), to be transformed into its dual problem, which is easier to solve. The dual problem is given by

$$\max_{\alpha}\left[\sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i=1}^{n}\alpha_i\alpha_j\gamma_i\gamma_j(x_i.x_j)\right] \qquad (12)$$

with constraints

$$\sum_{i=1}^{n}\alpha_i\gamma_i = 0, \quad 0 \leq \alpha_i \leq C, \; i=1,2,...,n. \qquad (13)$$

This is a classic quadratic optimization problem, for which there exists a unique solution. According to the Kuhn–Tucker theorem of optimization theory,[16] the optimal solution satisfies

$$\alpha_i\left[y_i(w.x_i+b)-1\right] = 0, \quad i=1,2,...,n \qquad (14)$$

(14) has nonzero Lagrange multipliers if and only if the points x_i satisfy

$$y_i(w.x_i+b) = 1. \qquad (15)$$

These points are termed SV. The hyperplane is determined by the SV, which is a small subset of the training vectors. Hence if α_i^* is the nonzero optimal solution, the classifier function can be expressed as

$$f(x) = \text{sgn}\left\{\sum_{i=1}^{n}\alpha_i y_i(x_i.x) + b^*\right\} \qquad (16)$$

where b^* is the solution of (14) for any nonzero α_i^*.

Nonlinear SVM

When a linear boundary is inappropriate SVM can map the input vector into a high-dimensional feature space. By defining a nonlinear mapping, the SVM construct an optimal separating hyperplane in this higher dimensional space. Usually nonlinear mapping is defined as

$$\varphi(.) : R^n \rightarrow R^n. \qquad (17)$$

In this case, optimal function (12) becomes (18) with the same constraints

$$\max_{\alpha}\left[\sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i=1}^{n}\alpha_i\alpha_j\gamma_i\gamma_j K(x_i.x_j)\right] \qquad (18)$$

where

$$K(x_i, x_j) = \left\{\varphi(x_i).\,\varphi(x_j)\right\} \qquad (19)$$

is the kernel function performing the nonlinear mapping into feature space. The kernel function may be any of the symmetric functions that satisfy the Mercel conditions.[17] The most commonly used functions are the radial basis function (RBF):

$$K(x_i, x_j) = \exp\left\{-\gamma\left|x_i - x_j\right|^2\right\} \qquad (20)$$

and the polynomial function

$$K(x_i, x_j) = (x_i.\,x_j)^q \qquad (21)$$

where q is the polynomial order.

The performance of SVM can be controlled through the term C and the kernel parameter which are called hyperparameters. These parameters influence on the number of the support vectors and the maximization margin of the SVM. The suitable selection of parameters of SVM plays an important role on the classification performance. In this paper, BA is applied to select the parameters of SVM.

Bees Algorithm

The bees algorithm is an optimization algorithm inspired by the natural foraging behavior of honey bees to find the optimal solution. Figure 2 shows the pseudocode for the algorithm in its simplest form. The algorithm requires a number of parameters to be set, namely number of scout bees (n), number of sites selected out of n visited sites (m), number of best sites out of m selected sites (e), number of bees recruited for best e sites (nep), number of bees recruited for the other ($m-e$) selected sites (nsp), initial size of patches (ngh) which includes site and its neighborhood and stopping criterion. The algorithm starts with the n scout bees being placed randomly in the search space. The fitnesses of the sites visited by the scout bees are evaluated in step 2.

In step 4, bees that have the highest fitnesses are chosen as "selected bees" and sites visited by them are chosen for

1. Initialize the solution population
2. Evaluate the fitness of the population
3. While (stopping criterion is not met)
   //Forming new population
4. Select sites for neighborhood search
5. Recruit bees for selected sites (more bees for the best e sites) and evaluate fitnesses.
6. Select the fittest bee from each site
7. Assign remaining bees to search randomly and evaluate their fitnesses
8. End while

Figure 2: Pseudocode of the bees algorithm

neighborhood search. Then, in steps 5 and 6, the algorithm conducts searches in the neighborhood of the selected sites, assigning more bees to search near to the best e sites. The bees can be chosen directly according to the fitnesses associated with the sites they are visiting. Alternatively, the fitness values are used to determine the probability of the bees being selected. Searches in the neighborhood of the best e sites which represent more promising solutions are made more detailed by recruiting more bees to follow them than the other selected bees. Together with scouting, this differential recruitment is a key operation of the bees algorithm.

However, in step 6, for each patch only the bee with the highest fitness will be selected to form the next bee population. In nature, there is no such restriction. This restriction is introduced here to reduce the number of points to be explored. In step 7, the remaining bees in the population are assigned randomly around the search space scouting for new potential solutions. These steps are repeated until a stopping criterion is met. At the end of each iteration, the colony will have two parts to its new population representatives from each selected patch and other scout bees assigned to conduct random searches.[18]

GENERAL STRUCTURE OF THE PROPOSED METHOD

The proposed system is the combination of the fuzzy self-organizing layer and the optimized SVM (OSVM) connected in cascade, named the fuzzy clustering optimized SVM (FCOSVM). Figure 3 shows the general scheme of this method.

The self-organizing layer is responsible for clustering of the input data. However, it is fuzzy clustering, in which the input vector x is preclassified to all sets with the different membership values. The penetration of the data space is better and the localization of the input vector x in the data space is more precise. The outputs of the self-organizing

subnetwork (membership values or fuzzy features) form the input vector to the second subnetwork (OSVM). The OSVM subnetwork is responsible for the final classification of the breast cancer tumor.

In the OSVM sub-network the most important parameters of the SVM, e.g., the penalty parameter (C) and the kernel function parameters such as the value of gamma(γ) for the radial basis function (RBF) kernel, are subjected to evolution using BA.

SIMULATION RESULTS

In this section we evaluate the performance of the proposed recognizer. This study has used WBC database. In order to compare the performance of classifiers, the k-fold cross validation technique is used. The k-fold cross validation technique proposed by Salzberg[19] was employed in the experiments, with $k=4$. The dataset was thus split into four portions, with each part of the data sharing the same proportion of each data class. Three data portions were used in the training process, while the remaining part was used in the testing process. The SVM-training methods were run four times to allow each slice of the data to take turn as a testing data. The classification accuracy rate is calculated by summing the individual accuracy rate for each run of testing, and then dividing the total by 4.

Several experiments were done to verify the effectiveness of the proposed method.

Experiment 1: Performance of Proposed System (FCOSVM)

First we have evaluated the performance of the recognizer. Tables 2 and 3 show the recognition accuracy (RA) of classifiers. Results imply that the proposed features have effective properties in breast cancer diagnosis. For example, SVM (GRBF) with WBC database has 95.75% recognition accuracy, while its performance increases with using fuzzy
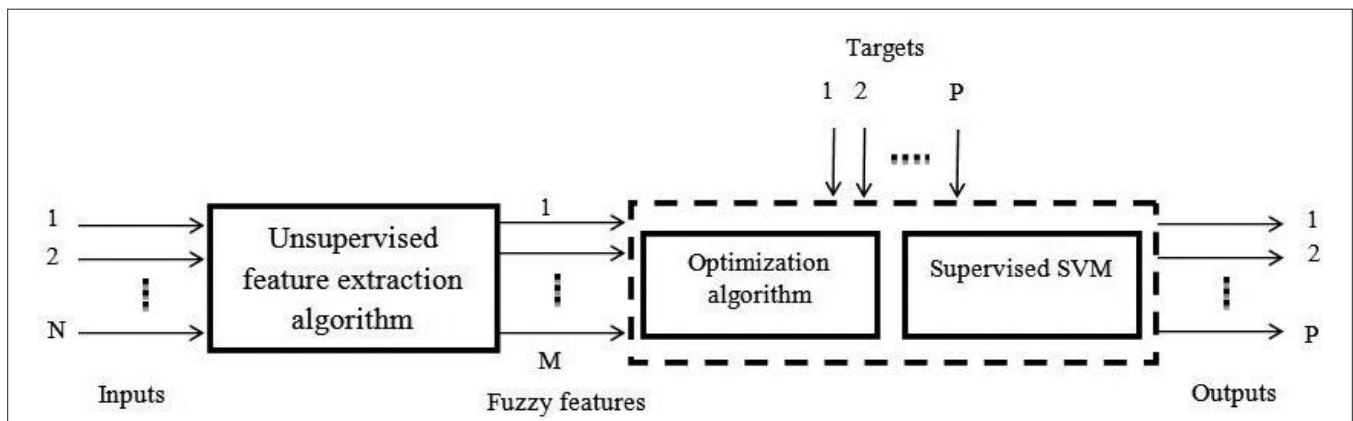

Figure 3: General scheme of the proposed method (FCOSVM)

features value up to 97.34%. Also it can be seen that the optimization improves the performance of recognizer significantly (98.85%).

In order to indicate the details of recognition for each pattern, the confusion matrix of the recognizer for best result is shown by Table 4. As we know, the values in the diagonal of confusion matrix show the correct performance of recognizer for each pattern. In other words, these value show that how many of considered patterns are recognized correctly by the system. The other values show the mistakes of the system. For example, look at the second row of this matrix. The value of 98.50% shows the percentage of correct recognition of the malignant pattern and the value of 1.5% shows that this type of pattern is wrongly recognized as a benign pattern. In order to achieve the recognition accuracy (RA) of the system, it is needed to compute the average value that appears in the diagonal.

Experiment 2: Performance Evaluation with Optimization in Different Runs for FCOSVM

In this subsection, for evaluating the performance of the BA, five different runs have been performed. In this experiment, we trained the SVM classifier based on the Gaussian kernel, which proved in the previous experiments to be the most appropriate kernel for breast cancer classification. Figure 4 shows a typical increase of the fitness (classification accuracy) for the best individual fitness of population obtained from the proposed system for different runs. As indicated in this figure, its fitness curves gradually improved from iteration 0 to 100, and exhibited no significant improvements after iteration 40 for the five different runs. The optimal stopping iteration to get the highest validation accuracy for the five different runs was around iterations 30–40.

In order to compare the performance of BA with other optimization algorithm, we have used a genetic algorithm (GA),[20] particle swarm optimization (PSO) algorithm[21] and imperialist competitive algorithm (ICA)[22] to evolve the
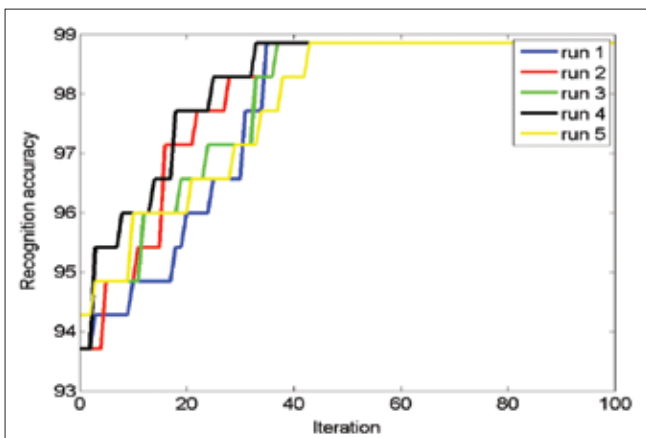

Figure 4: Evolution of fitness functions for different runs

SVM. Table 5 shows the obtained results. It can be seen that the success rate of BA is higher than the performance of other algorithms.

Experiment 3: Performance Comparing Performances of the Classification Techniques

The performance of the proposed classifier has been compared with other classifiers for investigating the capability of the proposed classifier, as indicated in Table 6. In this respect, probabilistic neural networks (PNN)[23] and multilayered perceptron (MLP) neural network with different training algorithm such as back propagation (BP) learning algorithm[24] and with Resilient propagation (RP) learning algorithm[25] are considered. They comprise parameters which should be readjusted in any new classification. Furthermore, those parameters regulate the

Table 2: Performance of different SVMs with WBC data

Classifier	RA (%)
SVM (Poly)	94.32
SVM (Linear)	94.22
SVM (GRBF)	95.75
OSVM (GRBF)	97.42

SVM – Support vector machines; OSVM – Optimized support vector machines

Table 3: Performance of different SVMs with fuzzy features

Classifier	RA (%)
SVM (Poly)	96.54
SVM (Linear)	96.48
SVM (GRBF)	97.34
FCOSVM (GRBF)	98.85

Table 4: Confusion matrix for best result (98.85 %)

	Benign %	Malignant %
Benign	99.2	0.8
Malignant	1.50	98.50

Table 5: Comparison among the performance of different optimization algorithms

Classifier	Optimization algorithm	Recognition accuracy (%)
FCOSVM	GA	98.26
FCOSVM	PSO	98.17
FCOSVM	ICA	98.38
FCOSVM	BA	98.85

Table 6: Comparison the performance of proposed method with other classifiers

Classifier	Recognition accuracy (%)
PNN	96.65
MLP (BP)	95.32
MLP (RP)	97.68
OSVM	98.85

Table 7: Classification accuracies obtained with proposed method and other classifiers from literature

Ref. no	Year	Input	Method	RA (%)
[26]	1996	Row data	C4.5	94.74
[27]	1996	Row data	RAIC	95.00
[28]	1999	Row data	NEFCLASS	95.06
[29]	1999	Row data	Fuzzy-GA1	97.36
[30]	2000	Selected features	Neuro- rule 2a	98.10
[31]	2003	Row data	SFC	95.57
[8]	2009	Selected features	AR+NN	97.40
This study	-	Fuzzy feature	FCOSVM	98.85

classifiers to be best fitted in for classification task. In most cases, there is no classical method for obtaining the values of them, and therefore, they are experimentally specified through try and error. It can be seen from Table 6 that the proposed method has better recognition accuracy than other classifiers.

COMPARISON AND DISCUSSION

For comparison purposes, Table 7 gives the classification accuracies of our method and previous methods applied to the same database. As can be seen from the results, the proposed method obtains excellent classification accuracy.

CONCLUSION

Accurate recognition of breast cancer tumor is very important for sufficient treatment. This study has investigated the design of an automatic and accurate system for detection of the breast cancer tumor. Based on the experimented results, this paper recommends the use of a hybrid system (FCOSVM) for diagnoses of the breast cancer. The complexity of the recognition system is very low in comparison with other works. The highest level of accuracy ever obtained by various methods using Wisconsin Breast Cancer (WBC) database was 95.75%. The proposed method improves the accuracy up to 97.34% by using the fuzzy feature as the SVMs inputs. Furthermore, optimizing the structure of the SVM and using fuzzy feature as the input of optimized classifier significantly improves the accuracy of the proposed system up to 98.85%.

REFERENCES

1. National breast cancer foundation, inc. Available from: http://www.nationalbreastcancer.org [Last accessed on 2011].
2. Cheng HD, Shan J, Ju W, Guo Y, Zhang L. Automated breast cancer detection and classification using ultrasound images: A survey. Pattern Recognit 2010;43:299-317.
3. Kerhet A, Raffetto M, Boni A, Massa A. A SVM-based approach to microwave breast cancer detection. Eng Appl Artif Intell 2006;19:807-18.
4. Baguia SC, Baguib S, Palc K, Pald R. Breast cancer detection using ranknearest neighbor classification rules. Pattern Recognit 2003;36:25-34.
5. Keles A, Ugur Y. Expert system based on neuro-fuzzy rules for diagnosis breast cancer. Expert Syst Appl 2011;38:5719-26.
6. Marcano-Cedeno A, Quintanilla-Dominguez J, Andina D. WBCD breast cancer database classification applying artificial metaplasticity neural network. Expert Syst Appl 2011;38:9573-9.
7. Khosravi A, Addeh J, Ganjipour J. Breast Cancer Detection Using BA-BP Based Neural Networks and Efficient Features. Machine Vision and Image Processing (MVIP), 7th Iranian, 2011 .
8. Karabatak M, Ince M. An expert system for detection of breast cancer based on association rules and neural network. Expert Syst Appl 2009;36:3465-9.
9. Choua SM, Leeb TS, Shaoc YE, Chenb IF. Mining the breast cancer pattern using artificial neural networks and multivariate adaptive regression splines. Expert Syst Appl 2004;27:133-42.
10. Jerez-Aragonés JM, Gómez-Ruiz JA, Ramos-Jiménez G, Muñoz-Pérez J, Alba-Conejo E. A combined neural network and decision trees model for prognosis of breast cancer relapse. Artificial Intelligence in Medicine, 2003;27:45-63.
11. Mu T, Nandi AK. Breast cancer detection from FNA using SVM with different parameter tuning systems and SOM–RBF classifier. J Franklin Institute 2007;344:285-311.
12. Chen H, Yang B, Liu J, Liu D. A support vector machine classifier with rough set-based feature selection for breast cancer diagnosis. Expert Syst Appl 2011;38:9014-22.
13. Ubeyli ED. Implementing automated diagnostic systems for breast cancer detection. Expert Syst Appl 2007;33:1054-62.
14. Dunn JC. A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. J Cybern 1973;3:32-57.
15. Bezdek JC. Pattern Recognition with Fuzzy Objective Function Algoritms. New York.: Plenum Press; 1981.
16. Bertsekas DP. Nonlinear programming. Belmont: Athena Scientific; 1995.
17. Courant R, Hilbert D. Methods of mathematical physics. New York: Wiley Interscience; 1953.
18. Pham DT, Ghanbarzadeh A, Koç E, Otri S, Rahim S, Zaidi M. The Bees Algorithm – A Novel Tool for Complex Optimisation Problems. Intelligent Production Machines and Systems. Pham DT, Eldukhri EE, Soroka AJ, editors. Cardiff University, Manufacturing Engineering Centre, Cardiff, UK: Published by Elsevier Ltd. All rights reserved; 2006.
19. Salzberg SL. On comparing classifiers: pitfalls to avoid and a recommended approach. Data Min Knowl Discov 1997;1:317-28.
20. Tang KS, Man KF, Kwong S, He Q. Genetic algorithms and their applications. IEEE Signal Process Mag 1996;13:22-37.
21. Shi Y, Eberhart R. A modified particle swarm optimizer. In: Proc. IJCNN, 1999.
22. Rajabioun R, Hashemzadeh F, Atashpaz-Gargari E. Colonial competitive algorithm A novel approach for PID controller design in MIMO distillation column process. Int J Intell Comput Cybern 2008;3:337-55.
23. Specht DF. Probabilistic neural networks. Neural Netw, first edition. 1990. p. 109-18.
24. Haykin S. Neural Networks: A Comprehensive Foundation. 2nd ed. New York: MacMillan; 1999.
25. Riedmiller M, Braun H. A direct adaptive method for faster back propagation learning: the RPROP algorithm, in: Proceedings of the IEEE Int. Conf. On Neural Networks, SanFrancisco, CA, Mar 28,1993.
26. Quinlan JR. Improved use of continuous attributes in C4.5. J Artif Intell Res 1996;4:77-90
27. Hamiton HJ, Shan N, Cercone N. RIAC: A rule induction algorithm based on approximate classification. In International conference on engineering applications of neural networks, University of Regina; 1996.
28. Nauck D, Kruse R. Obtaining interpretable fuzzy classification rules from medical data. Artif Intell Med 1999;16:149-69.
29. Pena-Reyes CA, Sipper M. A fuzzy-genetic approach to breast cancer

diagnosis. Artif Intell Med 1999;17:131-55.

30. Setiono R. Generating concise and accurate classification rules for breast cancer diagnosis. Artif Intell Med 2000;18:205-17.

31. Abonyi J, Szeifert F. Supervised fuzzy clustering for the identification of fuzzy classifiers. Pattern Recognit Lett 2003;14:2195-207.

BIOGRAPHIES

Jalil Addeh was born in 1986, Iran. He received B.S. degree in electrical engineering from Babol University of Technology, Iran, in 2010, and the M.S. degree in electrical engineering (control) from Babol University of Technology, Babol, Iran, in 2012. His research interests include image processing, pattern recognition, medical imaging processing, fuzzy logic, and neural networks.

E-mail: jalil-addeh@stu.nit.ac.ir

Ata Ebrahimzadeh was born in Babolsar. He received his PhD degree in electrical engineering. Now he is an associate professor in the Faculty of Electrical and Computer Engineering at Babol University of Technology. His research interests are: General area of Signal Processing, Pattern Recognition, Artificial Intelligence, Modern Control. He is the reviewer of a large number of international journals and conferences.

E-mail: e_zadeh@ nit.ac.ir