The Combination of Mammography and MRI for Diagnosing Breast Cancer Using Fuzzy NN and SVM

Elham Gohariyan[1], Mansour Esmaeilpour[2] *, Mohammad Mehdi Shirmohammadi[2]

[1]Department of Computer Engineering, Buin-Zahra Branch, Islamic Azad University, Buin-Zahra (Iran)
[2]Department of Computer Engineering, Hamedan Branch, Islamic Azad University, Hamedan (Iran)

Abstract

Breast cancer is one of the common cancers among women so that early diagnosing of it can effectively help its treatment in this study, considering combination of Mammography and MRI pictures, we will try to recognize glands in existing pictures which identify all around of gland complete and precisely and separate it completely. In this method using artificial intelligence algorithm such as Affine transformation, Gabor filter, neural network, and support vector machine, image analysis will be carried out. The accuracy of proposed method is 98.14. In this work a special framework is presented which simplifies cancer diagnosis. The algorithm of proposed method is tested on z16 images. High speed and lack of human error are the most important factors in proposed intelligent system.

I. Introduction

BREAST cancer is one of the common illnesses [2]. According to calculations of the national cancer institute of America, from each 8 women one is suffering of breast cancer. Breast cancer is the most common cancer among women and includes 32 percent of all women – related cancer and 19 percent of death reason of women with cancer, in other hand. It is the second reason of death with cancer [3]. If this cancer be diagnosed in time, it can be cured easily. Cancers are the third reason of death in our country and among 101 to 105 persons die due to it every day [3]. Gynecologist's emphasis that if women with breast cancer take action early, possibility of their treatment is more and their costs will be less [1]. In accordance with survey of Breast cancer research Centre (SID). The majority of patients (about 57 percent) are in second step of illness when they are in initial step [4]. The incidence of this cancer in Iran is z6 cases per 100/000 population of women over 30 years and in western countries it is 130 cases per 100/000, however, age of Iranian women who are infected with cancer is at least a decade less than women in developed countries [1]. Statistics show that incidence of cancer is increasing percent in the world annually. Despite of increments of it death amount caused by it has been decreased which indicates the development of notification methods thereby early detection of it becomes possible.

In past articles different methods of cancer detection are used. Technique using a new hybrid approach based on cuckoo algorithm and support vector machine classifier advantages of which are increasing computation performance and accuracy of SVM parameters that for improving it we have applied combination of fuzzy neural network and SVM. In another article namely the combination, decision tree

* Corresponding author.

E-mail address: esmaeilpour@iauh.ac.ir

and support vector machine have been used for breast cancer detection advantages of which are high accuracy and sensitivity and disadvantage of which is that its accuracy is relevance to the size of the training set [4]. However in this work we have attempted to train input data with evolutionary algorithm and then give its output to support machine that can be categorized easily. Fig. 1 indicates complete framework of this study.
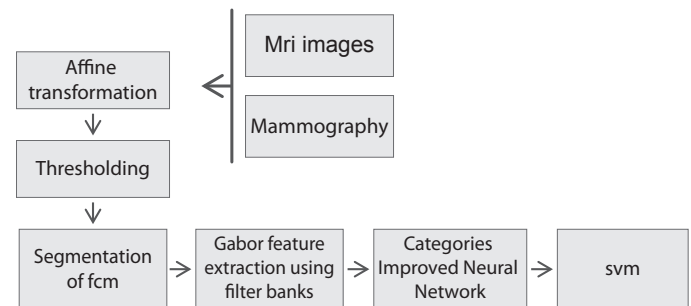

Fig. 1. The proposed method.

II. Affine Transformation and Fencing and Threshold Images

Affine transformation leaps either linearly (all points on a line will stay on a line after transformation) or the distances. In affine transformation straight lines stay straight and parallel lines will stay parallel [10].

Contractions, expansions, congruence, reflections, rotations and shear of similarity transformations, transfer and their combinations are all affine transformation keeps ratios on the line however it doesn't need to keep angles and length too. In Fig. 2 a sample of MRI and mammography pictures are shown.
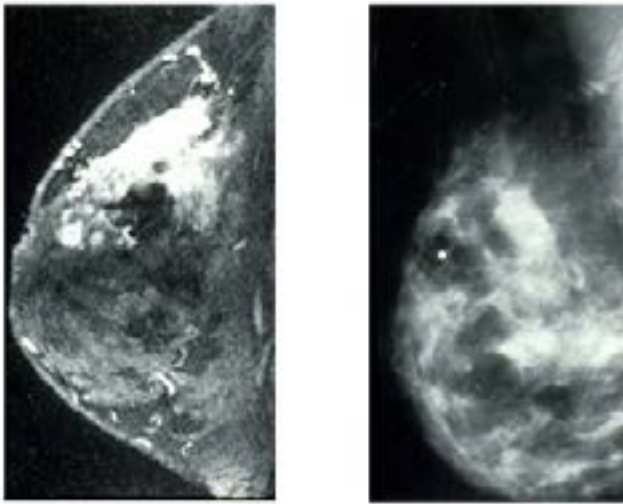
Fig. 2. Mamo right image and the left image is an MRI image.

In Fig. 3 these pictures are combined and in Fig. 4 an affine transformation has been applies to Fig. 3.
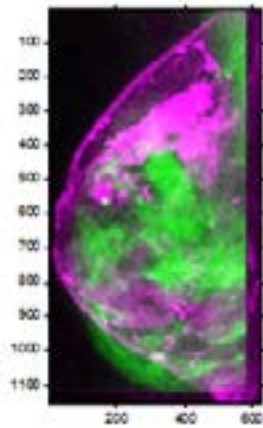

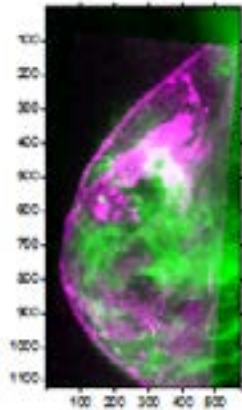Fig. 3. Combined mammography and MRI.

Fig. 4. Affine transformation conversions on images.

The main objective of fencing medical images is dividing it to different anatomy structures such as blood useless and tumors from background. For fencing first we use threshold for making dork points to light points in order to increase accuracy. Using threshold method in to transform images with gray scale to black and white images, these methods are accompanied with errors. In cases with threshold that pixels of background will became pixels of image will cause that object becomes thicker or thinner. For solving this problem we will use fuzzy

logic. FCM classification method is based on fuzzy theory [2]. In the Fig. 5 threshold method is used for improving injured points.
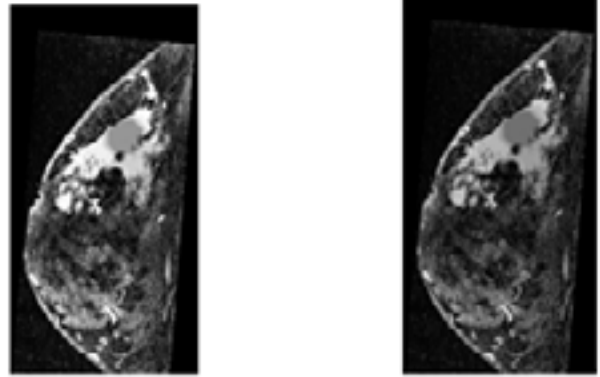

Fig. 5. Image segmentation with FCM.

In Fig. 5 is an example of image segmentation with FCM method is shown. To improve the image and isolation of cancer using noise removal methods can be performed better Isolated a sample of the image we have shown in Fig. 6.

Fig. 6. Image isolated tumor with FDM after noise.

III. FEATURE EXTRACTION WITH GABOR FILTER BANKS

While working on texture of image, features such as local spatial frequency and in formations related to pixel directions will be addressed. A certain approach for determining these features is calculating the power spectrum of image however Fourier transformation of an image calculates just a total estimation of image frequency in formation. A Gabor filter can be considered as a sinusoidal part with certain directions and frequencies that have been modeled by a Gaussian cover.

Gabor filter can be formulated as below

$$\varphi(x,y) = exp\left(i\left(f_x x + f_y y\right)\right) \times exp\left(-\frac{(f^2)(x^2 + y^2)}{2\sigma^2}\right) \tag{1}$$

IV. IMPROVED NEURAL NETWORK BUNDLE

The classification of improved neural network Fig. 7 indicates a total structure of RFNN model where N is input. And the classifier algorithm of support vector machine.
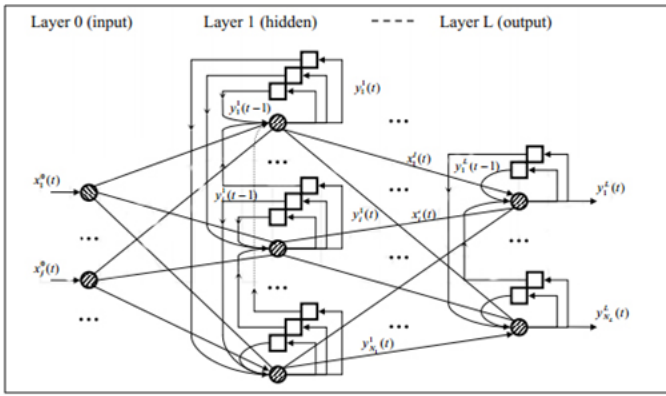
Fig. 7. Image isolated tumor with FDM after noise.

M is tern for each variable and p is output variable. Considering to Fig. 7 it is obvious that RFNN method has 6 layers. Each date function has a return relations ship with its structure. Kith rule with M inputs and N outputs in RFNN method is defined as below

$$Rule\ K:\ If\ x_1\ is\ A_1^K\ and\ x_2\ is\ A_2^K\ and\ ...\ x_N\ is\ A_N^K\ then$$

$$y_j^K = a_{j0}^K + a_{j1}^K x_1 + ... + a_{jN}^K x_N \quad j = 1,...,M \tag{2}$$

A. RFNN Network Training Based on Descending Gradient

RFNN network training is based on descending Gradient the cost function used in RFNN method training is de fined as below in formula

$$E(k) = \frac{1}{2}\sum_j (T(k) - O^5(k))^2 \tag{3}$$

T(k) is target vector of K_{th} sample in RFNN method rules of linear combination consists of inputs in addition a constant. The derivative of w_{st} function considering to parameters of date function is so important. When normalization comes in every output of 4th layer entails a function of all rules in RFNN method that this issue will complicate calculations.

B. Network Training Based on Learning by Combinatorial Optimization Algorithm

RFNN network training is based on learning by combinatorial optimization algorithms. In this part RFNN network training based on learning by combinatorial optimization algorithms will be addressed.

Since cost function with parameters of first part is complex, we use 4 optimization algorithms for optimizing these parameters.

C. Discussion

Technique using a new hybrid approach based on cuckoo algorithm and support vector machine classifier advantages of which are increasing computation performance and accuracy of SVM parameters that for improving it we have applied combination of fuzzy neural network and SVM.

The main advantage of it is to avoid over training and if it occurs, model will be able to predict similar data, which are used in learning step, be applied, system will have improper operation. Each data function in first part from RFNN model has 3 parameters for training. The centers of Fuzzy data functions are shown with m and standard deviation is determined by Q. O weight provides mobility for Fuzzy data functions should be trained. Parameters of teleport are trained with training algorithm of descending gradient. Every components of each part, includes fuzzy return data function, m, q, o, in this algorithm and in every stage of combinatorial optimization algorithm parameters are completed in the first interaction and mean square error is calculated. Semi code [1] indicates steps of training of RFNN network based on learning by combinatorial optimization algorithms.

Assume that we have data point sets and we want to segregate them pair wise. Every vector is consists of real numbers that are variables of software behavior. Linear classification methods try to segregate data with super surface [5]. Support vector machine classifier method that is the best linear classification, will find this super surface [4].

```
BIGIN
        DEFINE the structure of RFNN
        CONSTRUCT template parameter vector X for holding RFNN weights and biases;
        DEFINE the cost function as function of error function of RFNN parameters;
            a. INITIALISE population with random candidate solutions
            b. COMPUTE_FITNESS of each candidate
            c. while not STOP_CRITERION do
                    i. SELECT parents
                    ii. RECOMBINE pairs of parents
                    iii. MUTATE the resultant offspring
                    iv. EVALUATE of new candidates
                    v. SELECT individuals for the next generation;
        SELECT the parameter vector (RFNN parameter set) with best cost (training error)
        function from population and defining weights and thresholds for RFNN;
END
```

Fig. 8. RFN algorithm pseudo-code training.
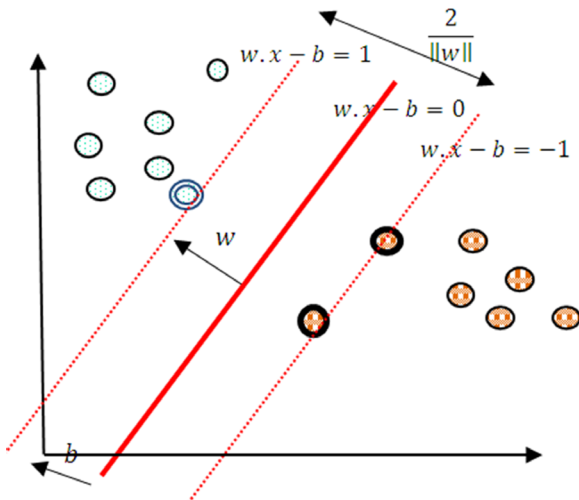
Fig. 9. How it classifies the SVM.

Goal of these classifications understand if a gland is benign or malignant. A common criterion for rating used is information gain 9 criterions defined as below.

$$IG(t_j) = \sum_{c=1}^{r} p(L_c) \, log_r \frac{1}{p(L_c)} - \sum_{m=1}^{r} p(t_j = m) \sum_{c=1}^{r} p(L_c|t_j = m) log_r \frac{1}{p(L_c|t_j = m)} \quad (4)$$

Where P(lc) is fractions of documents that has class LC. P(tj=1) and P(tj=0) means number of documents that have term tj or don't have it. IG(ti=0) is calculated for all terms and terms with vary law IG will be skipped.

V. Evaluation Criteria and Proposed Method Evaluation

Compare to digital and normal radiographies, digital radiographies are more a accurate [10]. Level of accuracy or measurement is determined by accurate. Goal of this study is to verify if this proposed method will increase accuracy in breast cancer detection or not.

$$Accuracy = \frac{(TP + TN)}{TP + TN + FP + FN}$$

$$Sensitivity = \frac{TP}{(TP + FN)}$$

$$Specificity = \frac{TN}{(TN + FP)} \quad (5)$$

In this section we compare our proposed method with other methods. In table we can conclude that results from Fuzzy C-Mean are better than network neural so this algorithm will be chosen as basic algorithm in other steps.

TABLE I.
COMPARE IMAGE SEGMENTATION ERROR

RMSE	Testing Error	RMSE	Learning Error	
0.30	0.14	0.17	4.12*e-5	NN
0.29	0.15	0.19	-1.08*e-7	FCM

Also for input and output variables training we will use particle swarm optimization Algorithm (PSO), Genetic Algorithm (GA), Imperialist competitive Algorithm (ICA) and simulated annealing Algorithm (SA). In accordance with results of performance, simulation and evaluation and also comparing classification errors we can conclude that PSD method is better.
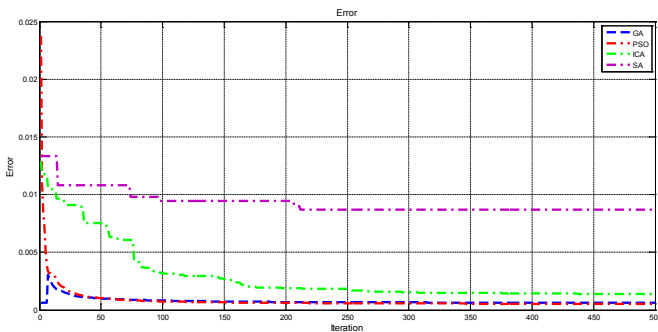

Fig. 10. Compare simulated fault.

TABLE II.
COMPARE CLASSIFICATION ERROR

	Train			Test		
	MSE	RMSE	R	MSE	RMSE	R
RFNN+PSO	3.316e-04	0.0251	0.9988	0.0024	0.0498	0.9647
RFNN+GA	5.414e04	0.0233	0.9989	0.0018	0.0427	0.9739
RFNN+ICA	0.0624	0.2498	0.8868	0.0289	0.1700	0.8201
RFNN+SA	0.2109	0.4593	0.2530	0.1323	0.3637	0.2321

Table II indicates that error caused by combination of RFNN and PSO is less than other methods. Since verifying it we use confusion matrix.

TABLE III.
CONFUSION MATRIX BEFORE EVALUATING SVM

Non- Cancerous	Cancerous	TrueForecast
2	23	Cancerous
29	0	Non- Cancerous

Values of accuracy, sensitivity a features are formulated as below.

$$Accuracy = \frac{23+29}{23+29+0+2} \times 100 = 96.30$$

$$Sensitivity = \frac{23}{23+2} = 0.92$$

$$Specificity = \frac{29}{29+0} = 1$$

It is obvious that step of algorithm accuracy is 97%.

TABLE IV.
CONFUSION MATRIX AFTER EVALUATING SVM

Non-Cancerous	Cancerous	True Forecast
1	23	Cancerous
30	0	Non- Cancerous

In accordance with values in table IV we have.

$$Accuracy = \frac{23+30}{23+30+0+1} \times 100 = 98.14$$

$$Sensitivity = \frac{23}{23+1} = 0.95$$

$$Specificity = \frac{30}{30+0} = 1$$

You can see that after applying SVM final result has been improved and is 98.14.

VI. Conclusion and Future works

Breast cancer is a common cancer among women. In this study in order to early breast cancer detection we have proposed a method based on processing mammography and MRI images presented here. This method uses classification methods and different so that detection and determine the exact place of lump. This method effectively decease human errors. Proposed algorithms applied to 216 cases of images and all results verified efficiency of it. According to results and comparing them, ability of Gabor bank fitter in producing feature vector has been proved. Tests and performance evaluations are done on images such as MRI and mammography. A sample image which had cancer lumps gotten from processing images and can be seen as segregated parts. Then inputs entered into improved neural network and results calculate precisely. In order to comparing proposed method evaluation with other methods, results shown in table V.

TABLE V.
COMPARE CLASSIFICATION ERROR

Proposed (RFNN+SVM)	Proposed (RFNN)	PSOWNN (22)	DEOWNN (21)	SONN (20)	
95	92	94.167	93.333	90.984	Sensitivity%
100	100	92.105	89.474	86.111	Specificity%
98.14	96.3	93.671	92.405	89.873	Accuracy%

This study is a unique work in breast cancer detection realm. So results of which can be expanded. For this purpose suggestions are told. Multi – objective evolutionary algorithms and classification technology so that achieving better results are proposed.

REFERENCES

[1] Amin R., "Categorizing normal and cancer patterns in mammography to detect breast cancer", Eighth conference on vision machine and image processing, 2013.

[2] Daroogar M., "Detection of breast cancer using an integrative approach new feature selection algorithm based on Cuckoo algorithm and support vector machine", fifth national conference of electrical and electronic engineering of Iran, 2012.

[3] Falahnejad M., "Mammography images segmentation using Fuzzy Nero for automatic detection of breast cancer", second national conference on new ideas in electrical engineering, 2012.

[4] Kazemi A., "Combination of decision trees and support vector machine for detecting breast cancer", First national conference on intelligent systems applications in science and technology, 2013.

[5] Abbasa, Q., Celebic, M.E. and Garcíad, I.F., "Reast mass segmentation using region-based and edge-based methods in a 4-stage multiscale system",Biomedical Signal Processing and Control, 8(2): 204–214,March 2013.

[6] Al-Shamlan and El-Zaart," Feature Extraction Values for Breast Cancer Mammography Images", Bioinformatics and Biomedical Technology (ICBBT). IEEE, Chengdu, pp. 335-340, 2010.

[7] Dom´Ænguez,A.R. and Nandi,A.K., "Enhanced Multi-Level Thresholding Segmentation and Rank Based Region Selection for Detection of Masses in Mammograms", Acoustics, Speech and Signal Processing. IEEE, pp. I-449- I-452, 2007.

[8] Kimori, Y.,"Morphological image processing for quantitative shape analysis of biomedical structures",effective contrast enhancement. NCBI, 1(1): 848–853, 2013.

[9] Li, L., Qian, W., Clarke, L.P., Clark, R.A. and Thomas, J.A., "Improving mass detection by adaptive and multiscale processing in digitized mammograms", SPIE--The International Society for Optical Engineering. Proceedings of the SPIE, pp. 490-498, 1999.

[10] Paramkusham, S., Kunda.M.M.Rao and Rao, B.V.V.S.N.P.," Early stage detection of breast cancer using novel image processing techniques, Matlab and Labview implementation", Advanced Computing Technologies (ICACT). IEEE, pp. 1-5, 2013.

[11] Rizzi, M., D'Aloia, M. and Castagnolo, B., "Computer aided detection of microcalcifications in digital mammograms adopting a wavelet decomposition", Integrated Computer-Aided Engineering, 16(2): 91-103, 2009.

[12] Aboul Ella Hassaniena, Hossam M. Moftahb, "Ahmad Taher Azar, Mahmoud Shoman, MRI breast cancer diagnosis hybrid approach using adaptive ant-basedsegmentation and multilayer perceptron neuralnetworks classifier", Applied Soft Computing 14 62–71, 2014.

[13] Dheeba J, Tamil Selvi S. "A swarm optimized neural network system forclassification of microcalcification in mammograms". J Med Syst;36(5):3051–61,2012.

[14] Dheeba J, Tamil Selvi S. "An improved decision support system for detection oflesions in mammograms using differential evolution optimized wavelet neuralnetwork". J Med Syst;36(5):3223–32,2012. approach", Journal of Biomedical Informatics 49 45–52,2014.

[15] Dheeba J, N. Albert Singh, S. Tamil Selvi, "Computer-aided detection of breast cancer on mammograms: Aswarm intelligence optimized wavelet neural network.

Elham Gohariyan

Elham Gohariyan received her B.S in Computer Engineering- Software and M.S in Computer Engineering-Artificial intelligent at Department of Computer Engineering, Islamic Azad University, Buin-Zahra Branch, Iran. He is currently lecturer in Department of Computer Engineering, Islamic Azad University, Hamedan Branch, Hamedan, Iran. His research interests are Learning Systems and Medical Image Processing.

Mansour Esmaeilpour

Mansour Esmaeilpour, received the B.S. and M.S. degrees in Computer Engineering-Software Engineering in Iran, in 2002 and 2004, respectively. He also received the Ph.D. degree in Computer Engineering from The National University of Malaysia, in 2012. Currently, he is an assistant professor in Computer Engineering Department at Islamic Azad University, Hamedan Branch, Hamedan, Iran. As well as, he worked from 2013 to 2014 as Postdoctoral Researcher at National University of Malaysia. His research interests include Data mining, Process Mining, Learning Systems and Machine Learning.

Mohammad Mehdi Shirmohammadi

Mohammad Mehdi Shirmohammadi received his B.S in Computer Engineering- Software and M.S in Information Technology from Islamic Azad University, Qazvin Branch, Qazvin, Iran and he is PhD student in Computer Engineering-Software systems in Islamic Azad University, Arak Branch, Arak, Iran. He is currently lecturer at Department of Computer Engineering, Islamic Azad University, Hamedan Branch, Hamedan, Iran. His main research interest is Sensor networks and decision systems.