Journal of Academic and Applied Studies
(Special Issue on Applied Sciences)
Vol. 4(1) January 2014, pp. 28-37

ISSN 1925-931X

Available online @www.academians.org

Accepted January 2014

Research Article

Mammography Image Segmentation Using Combined Fuzzy Logic and Neural Networks for Breast cancer Detection

Aioub Zeinvand Lorestani [1], Amir Massoud Bidgoli [2], Mashallah Abbasi Dezfoli[1]

[1] *Department of Computer, Science and Research Branch, Islamic Azad University, Khouzestan, Iran*
[2] *Department of Computer Engineering, Islamic Azad University, Tehran North Branch Tehran, Iran*

Abstract
Breast cancer is one of the major causes of death and the most common type of cancer among women worldwide. Breast cancer risk increases with age. Women within the age of 15-54 have more risk of breast cancer. Treatment is more efficient when detected early, as the evolution into a more severe stage is avoided. Screening mammography has been recommended as the most effective method for early detection of breast cancer. Computed aided detection intends to provide assistance to the mammography detection, reducing breast cancer misdiagnosis, and consequently allowing better treatment and prognosis.
Several research works have tried to develop computer aided diagnosis tools. They could help radiologists interpret mammograms and could be useful for an accurate diagnosis.
The purpose of this study is to introduce a new method based on the combination of neural networks and fuzzy logic in order for better detection and prognosis of breast cancer in mammography images. Results show that the proposed method is better than previous works and can facilitate the doctor to detect breast cancer in early stages of diagnosis process.

Key words: Mammography, Neural Network, Fuzzy Logic, Neuro-Fuzzy systems

Introduction

In recent years, many studies have been done on mammography images to detect cancerousmass by computer programming and image processing methods without the interference of an individual detector to avoid the individual’s fatigue, inaccuracy and vision problems. Some of the implemented algorithms segmented the image using classic method and image derivativeto detectabnormalities. Another groupapplied neural networks and fuzzy logic to segment and detect abnormal areas.The other one used the algorithms implemented based on neural networks. In the present study, an attempt has been made to extract cancerous massfrom mammography images by synthesizing neural networks and fuzzy logic in less time and with more accuracy.

Journal of Academic and Applied Studies
(Special Issue on Applied Sciences)
Vol. 4(1) January 2014, pp. 28-37

Int. Association Academians
ISSN 1925-931X

Available online @www.academians.org

1. **Preprocessing**

Mammography image may be of low contrast due to inadequate brightness and low dynamic range of imaging sensors. The increase of contrast may result in the increase of the dynamic range of the image. Often, before displaying the image it is necessary to make some changes. The purpose of these changes is to increase the image quality and highlight the desired patterns in order for better extraction at segmentation stage.

2. Extracting candidate areas

After standardizing image for preprocessing, suspicious areas (tumor region) are extracted from the image so that from then on instead of processing all the image points, only these areas would be processed in order for the image matrix to be seen smaller and the extraction speed of tumor region to get higher.

3. Labling candidate area

After segmenting candidate areas, the light pixels that are connected together get the same label. This labeling may cause the features of each area to be classified separately and the related patterns of each area will be analyzed separately.

4. Extracting features

In this section, features of the candidate areas are extracted from mammography image. The features should be selected in a way that they can separate tumor areas from other existing candidate areas.

5. **Classifying candidate areas**

After extracting features, the features of each candidate area are put into the desired class so that training and testing could be done. Moreover, it can be detected whether it is of cancerous or non-cancerous class.
In order for classification, usually the following categories are applied:
1. Neural networks
2. Fuzzy logic
3. Neuro-fuzzy system
In this study, neuro-fuzzy method was applied.

6**. Methodology**

As it was discussed earlier, the purpose of this study is to extract tumor regions in breast mammography images in order to detect breast cancer. To extract tumor regions the following stages are used:
1) Preprocessing
2) Segmenting breast images and selecting candidate areas
3) Processing candidate areas to detect whether it is cancerous or not
4) Extracting tumor area

6.1. **Preprocessing**

Journal of Academic and Applied Studies
(Special Issue on Applied Sciences)
Vol. 4(1) January 2014, pp. 28-37

Int. Association Academians

ISSN 1925-931X

Available online @www.academians.org

Before segmenting images they should be standardized so that they can be processed and candidate areas can be detected. Usually for edge detection of images common algorithms of displaying edges are used, but these algorithms are not suitable for medical noisy images. In this method, mathematical morphology operator is used for preprocessing and noise removing. First, Top-hat operator and then Bottom-hat operator have been implemented on the original image.Then, all Top-hat operator image pixels are subtracted from Bottom-hat operator image pixels. Figure 1 shows the result of the operation on mammography image.

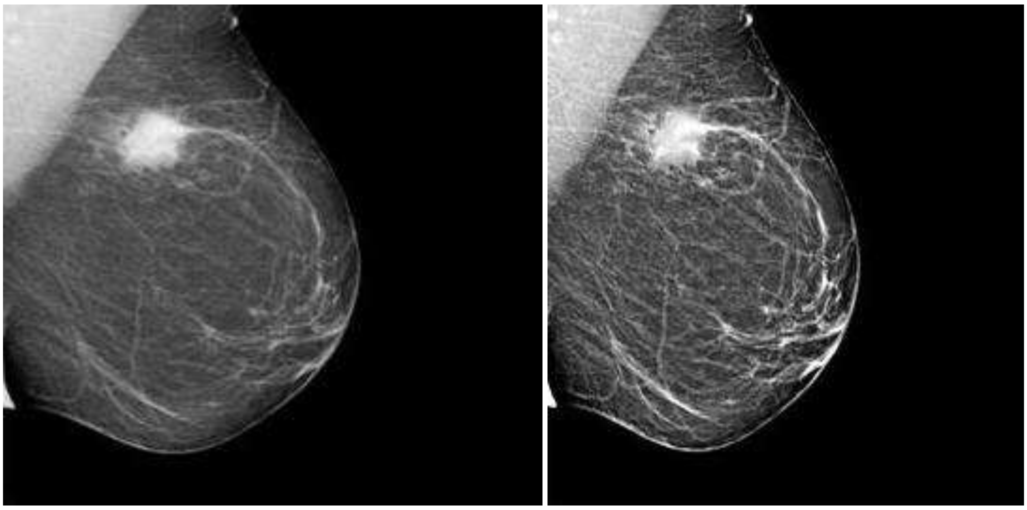


Figure 1: Converting Top-hat to an image. On the left: the original image, on the right: the improved image.

In the following, in order to remove bright calcium particles (benign areas of the image) by a disc-shaped structural element, with a diameter of 2 pixels, the image is eroded. The result of this conversion has been shown in the following figure.

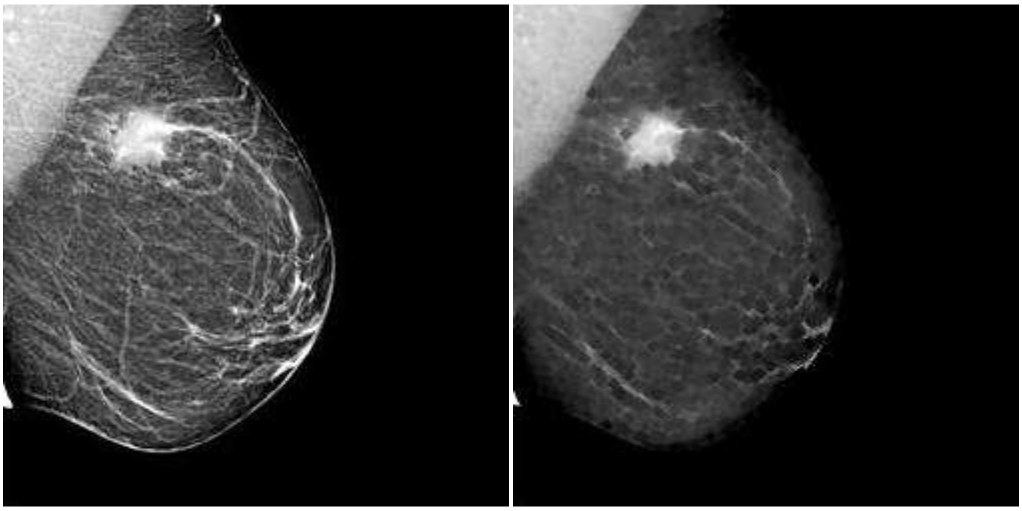


Figure 2: converting erosion to an image, on the left: improved image, on the right: erosive image.

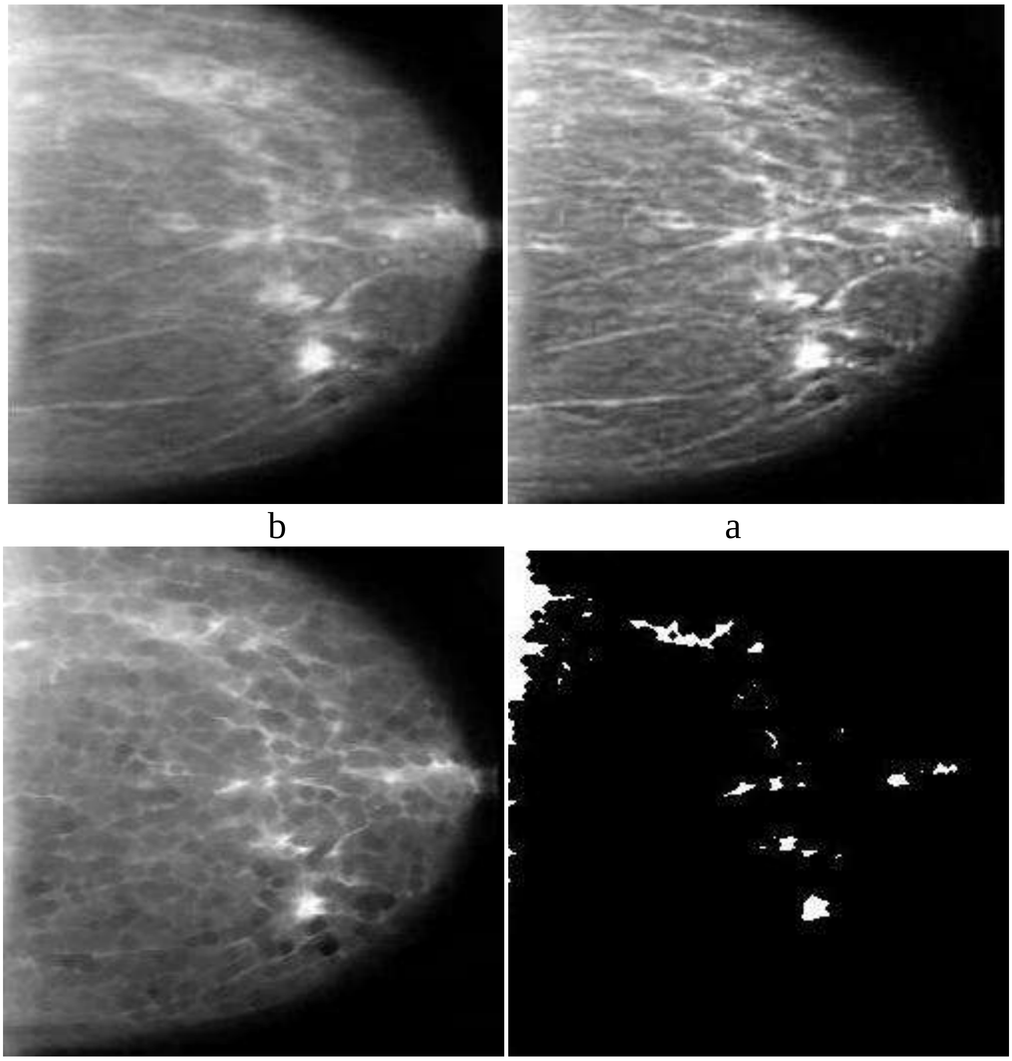
Journal of Academic and Applied Studies
(Special Issue on Applied Sciences)
Vol. 4(1) January 2014, pp. 28-37

Int. Association Academians
ISSN 1925-931X

Available online @www.academians.org

6-2. segmenting images and selecting candidate areas

After preprocessing (masking and increasing contrast),images should scroll all the image pixels and the threshold limit should be selected for them considering the tumor pixel value. The pixels that are close to tumor threshold limit can be
extracted as candidate area to the tumor and from them on only processing operation should be implemented on the areas. Threshold limit for image database has been considered as 190 and all pixels that are more than 190 considered as candidate pixel.



b a

d c

Figure 3: Extracting candidate areas from the mammography image, a: the original image, b: the improved image, c: the erosive image, d: extractive candidate areas

6-3. Labeling candidate areas

Journal of Academic and Applied Studies
(Special Issue on Applied Sciences)
Vol. 4(1) January 2014, pp. 28-37

ISSN 1925-931X

Available online @www.academians.org

After separating candidate areas, the light pixels that are connected together get the same label.This labeling cause the features of each area to be categorized separately and patterns related to each area are analyzed separately as well.

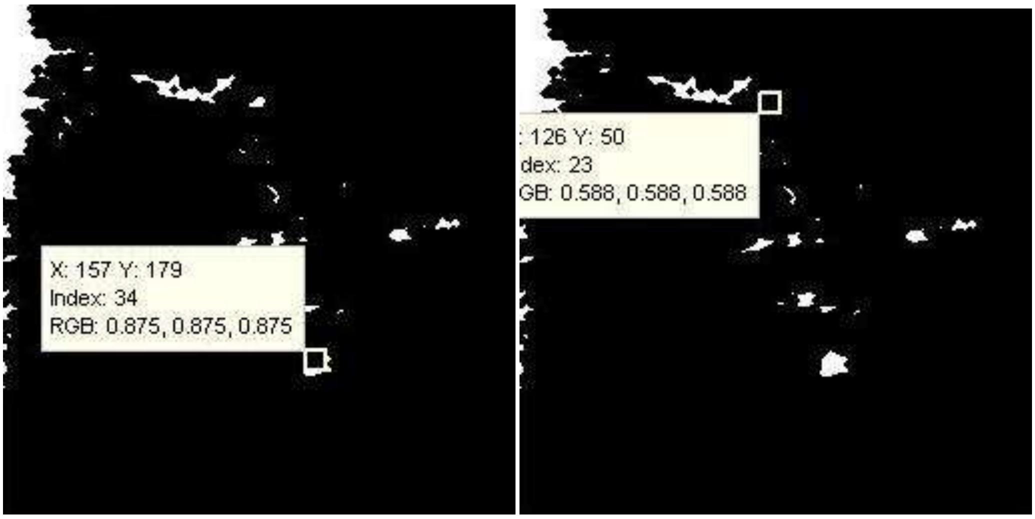


Figure 4: Labeling extractive candidate areas

6-4. Extracting the features of each candidate area

After labeling candidate areas, features of each area are extracted to send to the classifier and the training stage is implemented. The desired features that answered appropriately in training section are as the following:

- The total intensity of the pixels around the candidate area
- The average intensity of pixels of candidate area
- The intensity standard deviation of the pixels of candidate area
- Maximum intensity of the pixels of candidate area
- Minimum intensity of the pixels of candidate area
- The entropy intensity of the pixels of candidate area
- Total pixels of candidate area
- Total gradient in X direction
- Total gradient in Y direction

6-5. Classification based on neuro-fuzzy system

The applied neuro-fuzzy system of this study is Takagi-Sugeno. The system possesses parameters, rules and functions of compatible fuzzy member which is optimized during training bymeans of back propagation method.

Training the system is supervisory. Even though all the rules and parameters are optimized automatically, but if necessary, parameters can be set up manually in every stage. This is one of the best features of neuro-fuzzy systems.

First, basic structure of the system should be identified. The basic structure includes number of inputs, outputs, fuzzy conditional rules and form of membership functions. Number of inputs are considered as equal as the dimensions of applied feature vector.

Journal of Academic and Applied Studies
(Special Issue on Applied Sciences)
Vol. 4(1) January 2014, pp. 28-37

Int. Association Academians

ISSN 1925-931X

Available online @www.academians.org

Number of outputs of the system equals 1 in every state. This means that only one neuron is considered in outer layer and the system is taught in a way that the output figure of the neuron equals to the number of desired class. This choice has two basic features:

1. Reducing the number of outputs may result in reducing the number of connections and the corresponding weights which leads to the decrease of number of calculations and increase of the speed of work.
2. Considering that the number of outputs does not depend on the number of classes, the output structure in cases with different classes does not need change.This feature increases generalizability of the method.

Block structure of the system has been illustrated in figure 5. As discussed earlier, only the number of the inputs of the system aresame as the number of features.

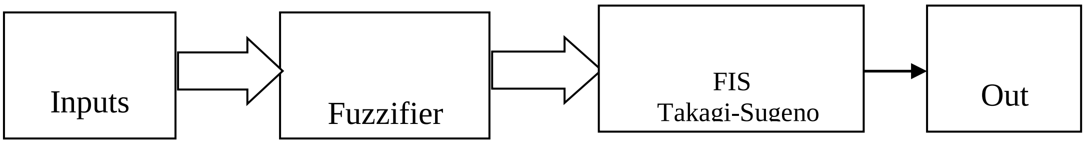


Figure 5: the proposed neuro-fuzzy system block structure

More accurate structure of the system have been illustrated in figures6 and 7. In this figure, in order for simplicity, the first and second blocks (i.e. Inputs and fuzzy blocks) have been synthesized.

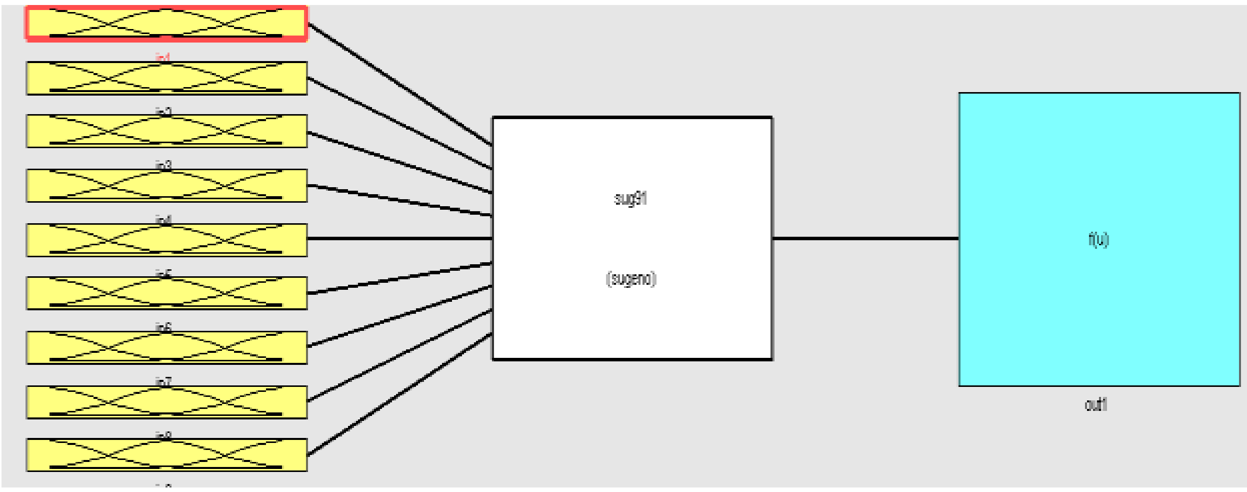


Figure 6: structure of the proposed neuro-fuzzy system

Journal of Academic and Applied Studies
(Special Issue on Applied Sciences)
Vol. 4(1) January 2014, pp. 28-37

Int. Association Academians

ISSN 1925-931X

Available online @www.academians.org

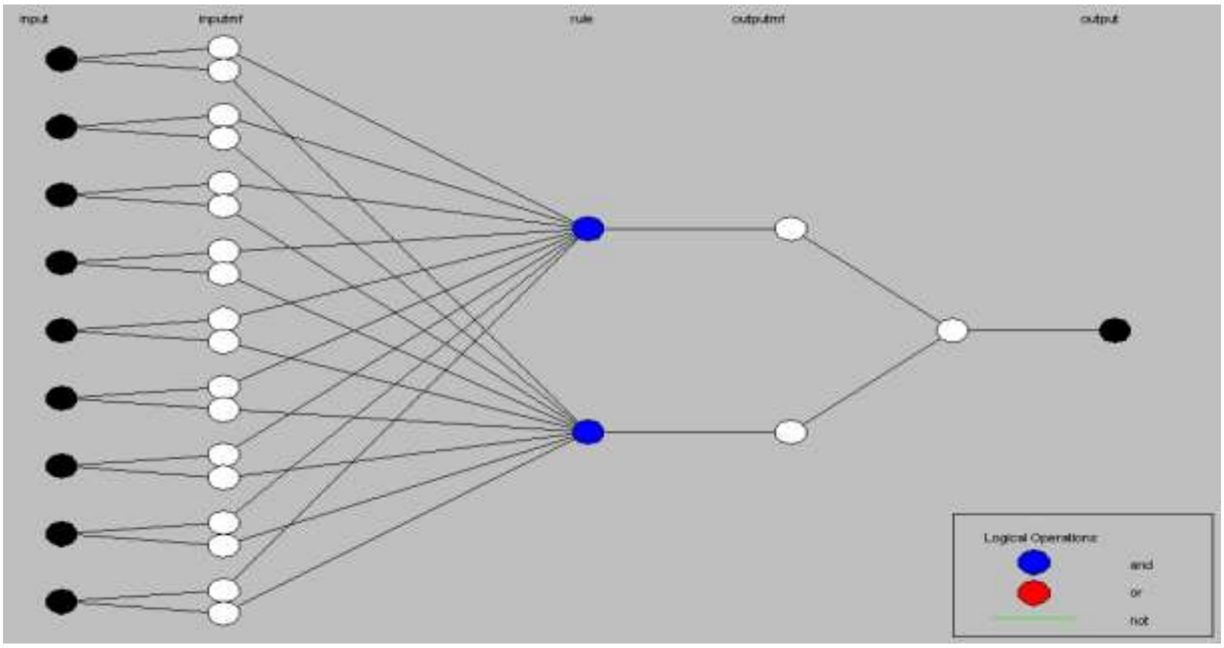


Figure 7: the implemented neuro-fuzzy structure

As discussed earlier, training this system is supervised learning. Therefore, for training the system a goal vector is established which corresponds to the number of training classes. Since one neuron has been considered, the target vector is defined as follows:

Target Vector T :{ 1, 2, 3… n+1} n: number of classes

As it was mentioned, the number of classes is considered as one class more than the actual number of classes so that unknown classes can be attributed to them. However, in the images tested in the present study, this class is not much of application, because in these images every class is predetermined and the system is trained to all of the classes.Now we can start training the system. The program is designed in a way that to start training two input values are requested from the user, one is the number of training samples and the other is the number of periods or iterations to reach the goal. To compare the achievements, we use 207 and the rest 105 are used for the test. The number of iterations of reaching the goal indicate that the system at the training stage in each training attempts several times to reach its goal. In other words, the parameter indicates the number of optimizing cycles and coefficient adjustment in back propagation mechanism of the system. This works the same for MLP as well. The value should be chosen in a way that root mean square error of the training is minimized. By increasing the iterations, the required time for training increases exponentially. On the other hand, we cannot consider the number of iterations as being too large because it takes too much time for training.

7. **Conclusion**

7-1.Validation of the results

This section consists of the results obtained from implementing the algorithms in the proposed method in order to extract tumor area from mammography images. The results of

Journal of Academic and Applied Studies
(Special Issue on Applied Sciences)
Vol. 4(1) January 2014, pp. 28-37

Int. Association Academians

ISSN 1925-931X

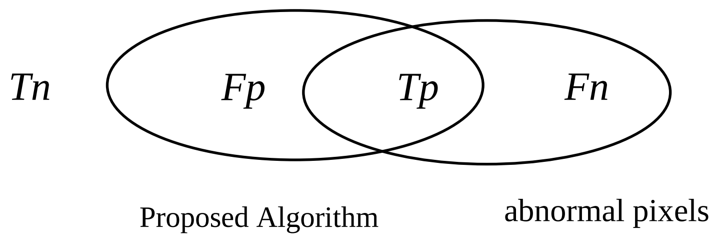
Available online @www.academians.org

implementing algorithm have been presented based on 2 criteria: sensitivity and specificity that are defined as follows:
Sensitivity: the ratio of the pixels that have identified as injury to the total pixels of injury
Feature: the ratio of the number of pixels that have been detected as healthy to the total pixels of injury
High level of the two criteria indicate better performance of the methods. The two criteria can be computed applying the following formulas:

$$sensitivity = \frac{T_p}{T_p + F_n}$$

$$specificity = \frac{T_n}{T_n + F_n}$$

Tp: the number of abnormal pixels that are detected as abnormal
Tn: the number of normal pixels that are detected as normal
FP: the number of normal pixels that are detected as abnormal
Fn: the number of abnormal pixels that are detected as normal



The results of the proposed method on the 5 images has been illustrated in Table1.

Table 1: The results of implementing the proposed algorithm on the 5 images

The image results	Tp	Tn	Fp	Fn	sensitivity	Feature
First	115	191153	9	12	.9055	.9799
second	586	190676	7	9	.9848	.9898
Third	136	191888	8	16	.8947	.9898
Fourth	106	191901	4	11	.9059	.9999
Fifth	780	191201	23	8	.9898	.9998

Journal of Academic and Applied Studies
(Special Issue on Applied Sciences)
Vol. 4(1) January 2014, pp. 28-37

Int. Association Academians

ISSN 1925-931X

Available online @www.academians.org

Table 2: Comparing the proposed algorithm with the previous methods

	The proposed algorithm	sensitivity	feature
Mammography images	Matosi2	0.90	.91
	Arfan	.93	0.96
	The prposed method	0.95	0.98

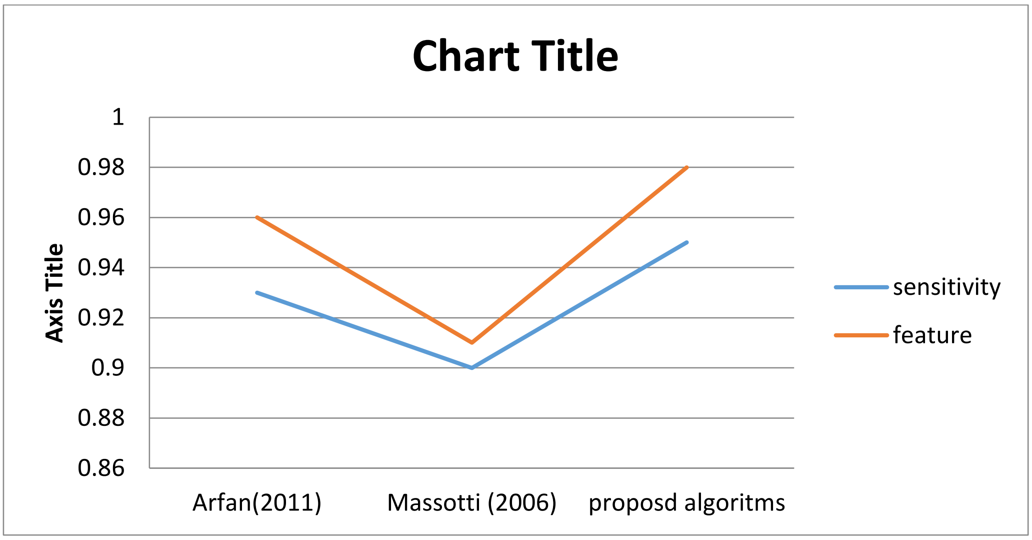


Figure 8: comparative chart of the proposed algorithm with the previous methods.

The required average time for in order to find the situation using CPU 1.8 GHz and 1G RAM was calculated in less than 80 second. During extracting the tumor area, choosing high threshold value (higher than 190) may result in increasing sensitivity and decreasing feature. The optimal value of the threshold experimentally was determined 190 for the desired image base.

7-2. Discussion

Radiological imaging plays a significant role in medical diagnosis by increasing image quality, upgrading imaging systems and computer technology. Image interpretation by a radiologist is affected by nonsystematic research of disease patterns and it can be varied in different individuals. Computer-based systems could be applied as the second option in pathological diagnosis. They improve the quality of medical images interpretation with a better review of different disease patterns and the accurate decision in diagnosis. An important factor which analyzes the record of computer-based systems is sensitivity parameter. It is a term which expresses percentage of diagnosed injuries in each image. In the present study, an attempt has been made to decrease mammography image features for processing and detecting cancer applying neuro-fuzzy system. By subtracting the value of feature matrix, the algorithm complexity decreases and we can achieve the desired answer with less number of training. Since the number of training in this system compared with fuzzy systems or multi-layer neural networks has been less and the time of diagnoses has decreased.

References

Arfan.J.M(2011),Mamdani.E.H(1975),Massotti.M(2006),Mencattini.A(2008),Menhaj.M.B(2009),Mohamed M.E(2012)

Arfan Jaffar M. , Naveed N., Zia S., Ahmed B., Tae-Sun Choi(2011) "DCT features based malignancy and abnormality type detection method for mammograms" International Journal of Innovative Computing, Information and Control, Volume 7, Number 9

Mamdani E. H. and Assilian S.,(1975), "An experiment in linguistic synthesis with a fuzzy logic controller", International Journal of Man-Machine Studies, Vol. 7, No.1, pp. 1-13.

Massotti, M. (2006). A ranklet-based image representation for mass classification in digital mammograms. Medical Physics, 33(10), 3951–3961.

Mencattini A., Salmeri M.,(2008)" Mammographic Images Enhancement and Denoising for Breast Cancer Detection Using Dyadic Wavelet Processing",IEEE transactions on instrumentation and measurement, VOL. 57, NO. 7.

Menhaj M. B.(2009) "Fundamentals of Neural Networks, Amirkabir University of Technology, (in Farsi)

Mohamed M.E. ,Ibrahima F., Brahim B. S.(2012)"A statistical based feature extraction method for breast cancer diagnosis in digital mammogram using multi resolution representation" Computers in Biology and Medicine 42 ,pp. 123–128