An Intelligent Method for Diagnosis of Breast Cancer

Pouya Derakhshan-Barjoei[1], Mojdeh Bahadorzadeh[2]

[1] *Assistant Professor, Young Researchers and Elite Club, Department of Telecommunication Engineering, Naein Branch, Islamic Azad University, Isfahan, Iran*

[2]*Assistant Professor, Department of General Surgery, Qom University of Medical Sciences, Qom, Iran*

*Corresponding Author: *Pouya Derakhshan-Barjoei, Assistant Professor, Young Researchers and Elite Club, Department of Telecommunication Engineering, Naein Branch, Islamic Azad University, Isfahan, Iran*

Abstract: *Diagnosis of breast cancer dramatically reduces the mortality rate in the female population. mamography image processing is a computer aided method which enhances the amount of detail visible on a digitalized image. The effect of this technique in the diagnoses of breast cancer, where the detection of early malignant tumors is essential for effective treatment, is reviewed in this paper. The mammograms display a small percentage of the information they detect and that is due to the minor difference between normal glandular tissues and malignant disease. The digital medical image processing uses denoising and image enhancement techniques so as to reveal any tumors that may not be obvious and help the surgeons decide. The idea is to transform the data into the wavelet basis, in which the large coefficients are mainly the signal and the smaller ones represent the noise and using particle swarm optimization for high precision and feature detection. In this paper we employ evolutionary method of image enhancement, Particle swarm optimization algorithm and wavelet decomposition accompany with denoising are implemented and the conclusion would be satisfied.*

Keywords: *Wavelet, Particle Swarm Optimization Glandular Tissue, Micrcoalcification, Mammogram,*

1. INTRODUCTION

Recently, the significance of image processing and machine vision is increasing in industrial automation, security, health, and traffic control in parallel with the developments in engineering technology. Breast cancer is one of the most common cancers among women, which is the second most common cause of death in women. Breast cancer is an uncontrolled growth of abnormal cells that occurs in different regions of the breast.[1-4]

As mentioned, the timely diagnosis of breast cancer dramatically reduces the mortality rate in the female population. Needle aspiration testing is a simple, inexpensive and non-invasive method for accurate and early diagnosis of this cancer, which today is attempted to be smart and machine-made. The steps to create an intelligent system for detecting breast cancer include: recording microscopic images from fine needle sampling, extracting numerical features from these images, selecting separation features, and designing and classifying tests.

Most applied data sets, including the breast cancer data set, include a large number of examples and many features. Since this set of data may be aggregated for non-performing reasons (such as categorization), they may include many additional or unnecessary features. Therefore, in this situation, diminishing the dimensions of the inputs is necessary. Extracting valuable information from such a collection requires a complete search of the collection, which itself creates other challenges, such as complexity management and computational time. A common practice to overcome these problems is to use dimension reduction techniques such as feature selection.[3-7] Selecting a subset of features can increase class accuracy by reducing the predictive error by reducing the number of final samples. Breast cancer may occur in the tissues of the breasts like the dairy that carry the milk, in a tissue that produces non-germs. Breast cancer after skin cancer is the second most common cancer in women. According to the National Cancer Institute of the United States, one out of every eight women suffers from breast cancer. It can be easily treated if diagnosed on a timely basis. The risk of breast cancer increases with age. About three quarters of the cases of breast cancer develop in women over 50 years

of age. Other risk factors include family history, lack of exercise life and low physical activity, high alcohol consumption, non-pregnancy or pregnancy after 30 years of age, and so on. Symptoms of breast cancer include a lump or thickening in the breast or axillary secretions, clear or bloody chest, and dive nipples inward, redness or swelling of the breasts, in short, any change in the shape of breasts that differ from it with bust be no longer. Breast cancer treatment depends on a variety of factors. Surgical treatment may be performed in the form of partial removal to the full extent of the chest. If the cancerous tissue is concentrated, a combination of surgery (lumpectomy), chemotherapy, and radiotherapy is usually used. Therefore, the timely diagnosis of breast cancer (up to 5 years after the first cancer cell division) increases the chance of survival from 56% to more than 86% . Therefore, there is a need for a precise and reliable system for timely diagnosis of chest tumor or malignant tumor. Traditionally, breast cancer diagnosis has always been based on morphological and clinical measurements and has been done through sampling and surgery. Among the existing methods, this method has the highest diagnostic accuracy, but an invasive method, It's time consuming and expensive [3-9]. The needle aspiration test is another method that does not have the disadvantages of the previous method and, with the aid of machine learning techniques, can provide an efficient system for the diagnosis of breast cancer.

Another way of accessing data on suspected infections is to conduct mammograms. Extremely digital mammography used in treatment centers uses precision digital imaging with minimal radiation exposure. When digital images are obtained and sent to a digital mammography station. Radiologists analyze images to determine the precise location of the area and the extent of the disease. But sometimes it's difficult and time consuming to detect malignant or tumorous tumors even for the best practitioners in this area. That's why intelligent systems have been introduced to help clinicians diagnose as quickly and accurately as possible. Breast cancer accounts for one third of all cancers detected and 18% of all cancer deaths. Untill recently breast cancer was the leading cause of death among women but since 1985 it has ranked second lung cancer. Prevention of this disease is not possible since its cause is not fully understood. However, current methods of Treatment are very effective against breast cancer in its early phase [10]. Therefore, the most promising way to achieve a change in the current breast cancer situation is to remove the cancer in its early stages. The early detection of breast cancer is most reliably achieved with mammography. However 10% to 30% of women who have breast cancer and who undergo mammography have negative mammograms [9][12]. In approximately two thirds of these false negative mammograms the radiologist failed to detect the cancer that was evident retrospectively [11]. It is difficult for the human radiologist to maintain interest in interpreting large number of images in which only a small number show abnormalities. Hence the need to construct computer aided systems to diagnose breast cancer in mammograms becomes apparent. Micro calcifications usually come in clusters, having very sharp edges, and usually irregular shape of very small size. Micro calcifications are usually noncancerous, but certain patterns can be a sign of cancer. If calcifications are suspicious, further testing may be necessary, including additional mammograms with magnification views or a breast biopsy. While some calcifications may indicate breast cancer, there are many noncancerous (benign) conditions in the breast that can cause calcifications to form. Causes of breast calcifications include:

Breast cysts, Cell secretions or debris, Ductal carcinoma in situ (DCIS), Fibro adenoma, Mammary duct ectasia, Mastitis, Previous injury to the breast, Previous radiation therapy for cancer, Skin (dermal) or blood vessel (vascular) calcification. Mammograms are initially enhanced by either increasing the contrast of suspicious area or by removing background noise. Various mathematical methods are then applied to detect the individual tumors depending on whether the tumor appears as a micro calcification cluster or a mass, in mammography the interesting characteristics of an image are malignant masses, micro calcifications and skin thickening of which the last two are said to be indirect signs of malignancy[13-22].

2. MICRO CALCIFITION AND IMAGE FEATURE

The detection of micro calcifications in mammograms since this is the first indication of the presence of breast cancer is highly recommended. The first deals with enhancing the contrast of suspicious area in the image while the second technique involves the removal of background noise from the image. The favored method of image enhancement of mammograms in the removal of background noise while preserving the edge information of suspicious areas in the images. This can be achieved using

three different methods, a) selective averaging schemes) median filtering and c) a modification of median filtering. generally, the median filter allows a great deal of high spatial frequency detail to pass while remaining very effective at removing noise on images where less than half of the pixels in a smoothing neighborhood have been effected.

3. SELECTION OF MEDIAN FILTER

Median filtering has been found to be very powerful in removing noise from 2-D signals without blurring edge. This makes it particularly suitable for enhancing images. To apply median filtering to a digital picture, we replace the value at a pixel by the median of the values in a neighborhood of the pixel. One of the major problems with the median filter is that it is relatively expensive and complex to compute. To find the median it is necessary to sort all the values in the neighborhood into numerical order and this is relatively slow, even with fast sorting algorithms[10][16]. Two dimensional median filters can be defined for arbitrary sizes and shapes of filter windows W (i,j), such as line segmenta, squares , circles and crosses. The edge preservation power of the standard median filter is not sufficient for enhancing mammogram   images due to the fuzziness of the boundaries of suspicious areas. A modification of the filter selective median filter was defined by Lai et al.,[18][19][20]. For a window W (i, j) centered at image coordinates (i, j) the output of the selective median filter is:
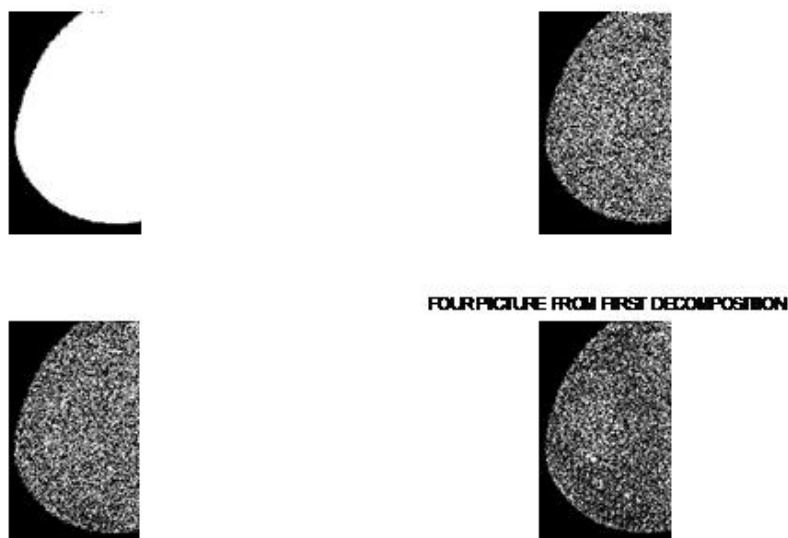
$$Xij = \text{median} \{ Xrs : (r,s) \in N(i,j) \text{ and } | Xrs - Xij | < T \} \tag{1}$$

Where $(i,j) \in Z2$,N(i,j) is the area in the image covered by window W (i,j) and T is a threshold.
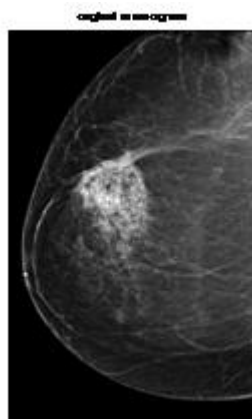
Fig1. *a) Original Mammogram; white specks*

Our proposed method using Wavelet decomposition for two times the result of frequency components of mammogram image is shown in Fig. 2. and Fig. 3.

Fig2. *Four images of the first step of decomposition from original mammogram image*
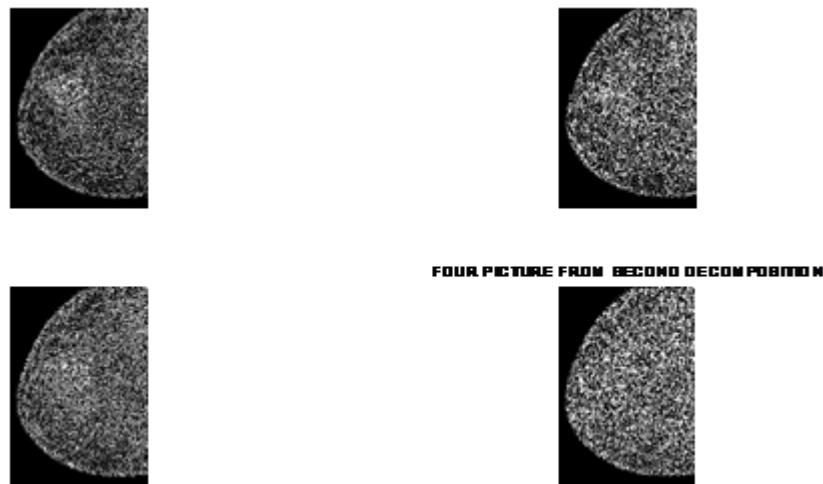
FOUR PICTURE FROM SECOND DECOMPOSITION

Fig3. *Four images of the second step of decomposition from mammogram image*

4. USING EVOLUTIONARY ALGORITHM AND FILTER

In the filter model, we use mutual information analysis based on the binary particle swarm optimization algorithm to measure the separation capability of the features.

The particle swarm optimization algorithm was inspired by the behavior and movement of birds, bees and fish by Kennedy and Eberhart in 1995 [24]. In fact, a powerful and random method for evolutionary computing based on the intelligent movement of animal categories in the search for food, which is becoming increasingly commonplace since, in contrast to the genetic algorithm, especially in issues involving continuous design variables, in Finding optimal global solutions has higher returns [19]. This algorithm follows the conventional method of computational evolution. A: With random populations, possible answers begin. B: By updating the generations, it searches for the optimal answer. C. The population evaluation is based on previous generations. The algorithm of optimizing the congestion of particles The probable solutions (particles) in the solution space of the problem are displaced by looking for the optimal current particles. This transition is effected by a fitness function that evaluates the quality of each particle. If the search space is D-dimensional then the position of the ammunition of the particle of the category can be represented as a D-dimensional vector: $X_i = (X_{i1}, X_{i2}, ..., X_{iD})$ Also, the particle velocity (position change) is also represented by another D-dimensional vector Can be displayed: $V_i = (V_{i1}, V_{i2}, ..., V_{iD})$. The best fit for an amine particle and its corresponding position are displayed with 〖pbest〗_i and 〖xpbest〗_i, respectively.

$$V_i(t+1) = \omega V_i(t) + c_1 r_1 (xpbest_i(t)\_X_i(t)) + c_2 r_2 (xgbest_i(t)\_X_i(t)) \tag{2}$$

$$X_i(t+1) = X_i(t) + V_i(t+1) \tag{3}$$

The optimality of the best particle in the category (Optimal Global) should also be stored in the memory of the algorithm. It has been proven that the use of weighted weight w improves the performance of the algorithm [23-25]. In the particle swarm optimization algorithm, taking into account the inertia weight (inertia), particle update equations are performed with the following equations. Where $i = 1,2, ..., N$ and N are the population of the group and c_1 and c_2 are the learning coefficients that have positive values. The coefficients r_1 and r_2 are random numbers in the interval [0,1], or the inertial weight is usually between 0.4 and 0.9. According to [23-26], large values at the beginning of the search lead to a higher priority of global discovery than local discovery and to a more gradual search for more local issues, thus increasing the amount at the beginning of the search and gradually increasing the amount It slows down and slowly moves to zero. t also specifies the number of repetitions. The binary version of the particle swarm optimization algorithm was presented by Kennedy and Eberhart in 1997 [23-25] and, contrary to the standard, can be optimized in discrete spaces. The choice of a feature is a binary discretization optimization problem. By considering a binary string for each particle, the optimization algorithm for binary particle swarm begins. In this field, 0 represents the deletion of the attribute and 1 indicates the selection of that attribute. The only difference that exists between the particle swarm optimization algorithm and the particle position update is the following:

$$S(V(t+1)) = \frac{1}{1+\exp(-V(t+1))} \tag{4}$$

$$X(t+1) = \begin{cases} 1, if \ \frac{1}{1+\exp(-V(t+1))} > rand \\ \quad 0, \ otherwise \end{cases} \tag{5}$$

Where rand is a random number distributed uniformly over the interval [1,0]. The best values for the coefficients c_1 and c_2 can be calculated as follows:

$$\begin{cases} c_1 = \frac{1}{2Ln2} = 0.721 \\ c_2 = 0.5 + Ln2 = 1.193 \end{cases} \tag{6}$$

In the proposed algorithm, binary particle swarm optimization algorithm is used as a filter to select the characteristics with the best mutual information. The concept of mutual information was introduced for the first time to detect the relationship between these features of a data set. By having two random variables x,y the mutual information, namely, I(x,y), the function of the probability density functions, $p(x)$, $p(y)$, $p(x,y)$ is introduced:

$$I(x,y) = \int \int \frac{p(x,y)\log(p(x,y))}{p(x)p(y)} dxdy \tag{7}$$

$$I(x,y) = \sum_{X=x}\sum_{Y=y} \frac{p(x,y)\log(p(x,y))}{p(x)p(y)} \tag{8}$$

Formulas 7 and 8 are for continuous and discrete variables, respectively. However, even if one of the variables has a continuous value, it will be very difficult to calculate the amount of mutual information. To solve this problem, they often use discrete values of the continuous value. In order to obtain the simulation results, the Binary Particle Swarm optimization algorithm has been used as the filter for selection of characteristics with the best mutual information (MI). The parameters of this algorithm are presented in Table 1. The result of applying this algorithm to data sets is to specify a set of properties that provide the highest value of the fit function. The results of applying the filter algorithm for binary particle swarm optimization and mutual information are given for reference mammogram photographs

Table1. *BPSO Parameters*

Value	Parameters
25	Population
50	Iteration
2	C_2 - C_1
6	Max speed
-6	Min speed
0-1	Inertia

Our original mammogram image and final result of our method as described is shown in Fig,4. We can compare them and see how the microcalcification detection in clear, so the threshold setup and noise reduction for each image should be set intelligently.
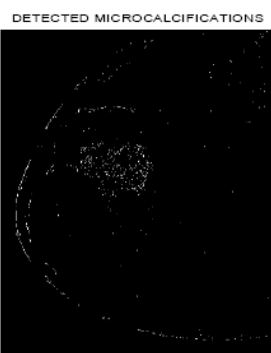
Fig4. M*icrocalcification detected image of our method.*

A 45-year-old female with history of breast reduction surgery two years prior presents with focal abnormality on routine screening mammogram. Mammograms obtained prior to surgery and one year after surgery are available for comparison. No significant family history of breast cancer. Kindly, We used image that prepared by Peter Bergquist, MD and Erini Makariou, MD. Noise suppression in medical images is a particularly delicate task. A tradeoff between noise reduction and the preservation of actual image features has to be made in a way that enhances the diagnostically relevant image content. Our method evaluates two-dimensional denoising procedures using medical test images corrupted with additive Gaussian noise. Our results, using the peak-signal-to-noise ratio as a measure of the quality of denoising, show that the proposed method outperforms well.

5. CONCLUSION

A new algorithm for detecting breast cancer cases is presented in this article. The algorithm uses a new hybrid system based on BPSO algorithm to find the proper features for classifying the process in breast cancer data sets. Finally, the proposed algorithm was tested on breast cancer data to demonstrate efficacy. Simulation results show the efficiency of the proposed algorithm in noise elimination and improvement. Also, the efficiency of the proposed algorithm is not limited to these two sets of data, and this method is applicable to other larger data sets, and especially to microarrays.

The most significant criterion for this applications used in these fields is to make sure the scheme to run at high speed and real time with precision. In our method a digital medical image processing uses denoising and image evolutionary and wavelet enhancement techniques to reveal any tumors that may not be obvious. Extensively for denoising medical images wavelet thresholding has been used. The idea of using evolutionary algorithm approached a good feature detection. We employ this joint method of image enhancement and the results show good performace. The proposed technique in the diagnoses of breast cancer enhanced by either increasing the contrast of suspicious area or by removing background noise and precise detection. the binary particle swarm optimization technique (BPSO) is employed for feature selection in order to maximize the accuracy of the decision-making system.

REFERENCES

[1] Przemyslaw Kupidur,"Semi-automatic method for built-up area intensity survey using morphological granulometry"IJIREC,pp271-277, 2010.

[2] P. Alamelu, P. Pragadeshkumar, V.Annapoorani " Study of Recent Development in Medical Image Processing Using Computer Visualization Based Method"IJIREC,pp20-23, dx.doi.org/10.20431/2349-4050.0305003

[3] Kleihues, P., & Sobin, L. H. (2000). World Health Organization classification of tumors. Cancer, 88(12), 2887-2887.

[4] Luo, Z., Wu, X., Guo, S., & Ye, B. (2008, June). Diagnosis of breast cancer tumor based on manifold learning and support vector machine. In Information and Automation, 2008. ICIA 2008. International Conference on (pp. 703-707). IEEE.

[5] Listgarten, J., Damaraju, S., Poulin, B., Cook, L., Dufour, J., Driga, A., ... & Zanke, B. (2004). Predictive models for breast cancer susceptibility from multiple single nucleotide polymorphisms. Clinical Cancer Research, 10(8), 2725-2737.

[6] Mangasarian, O. L., Street, W. N., & Wolberg, W. H. (1995). Breast cancer diagnosis and prognosis via linear programming. Operations Research, 43(4), 570-577.

[7] S.Karthik, M. Sathishkumar, N. Kannan, "Study of Implementation of Image Analysis with Hardware and Software Co-Design on the Xilinx Platform", IJIREC, pp 1-5,dx.doi.org/10.20431/2349-4050.0305001 Volume-3 Issue-5, 2016

[8] P. Derakhshan-Barjoei, M. Bahadorzadeh, "Enhancement in Medical Image Processing for Breast Calcifications and Tumor Detection", Research Journal of Applied Sciences, Engineering and Technology, 4(12): 1696-1700,2012.

[9] Holland T et al, 1982. "So-called interval cancers of the breast: pathologic and radiographic analysis", Cancer; 49:2527-2533.

[10] Stamatia Detounis, MD, 2004. "Computer –Aided Detection and second reading utility and implementation in a high-volume Breast clinic", Applied Radiology,pp:8-15.

[11] Nodes TA and Gallagher NC, 1982. "Median filters: some modifications and their properties", IEEE Trans. Acoust., Speech Signal Processing, vol. ASSP-30,Oct.

[12] Smith RA, 1993. "Epidemiology of breast cancer"syllabus:79th Scientific Assembly of the Radiological society of North America, pp:21-33.

[13] Eberhart, R., & Kennedy, J. (1995, October). A new optimizer using particle swarm theory. In Micro Machine and Human Science, 1995. MHS'95., Proceedings of the Sixth International Symposium on (pp. 39-43). IEEE.

[14] Abeloff MD, et al. Cancer of the breast. Abeloff's Clinical Oncology. 4th ed. Philadelphia, Pa.: Churchill Livingstone Elsevier; 2008

[15] Adam A, et al. Grainger & Allison's Diagnostic Radiology: A Textbook of Medical Imaging. 5th ed. Philadelphia, Pa.: Elsevier Churchill Livingstone; 2008.

[16] Bovik AC et al, 1987. "The effect of median filtering on edge estimation and detection", IEEE Trans. pattern Anal. Machine Intell.vol PAMI-9,pp.181-194,Mar.

[17] John J. Heine, Stanely R. Deans, Senior Member, IEEE, D. Kent Cullers, 1997. Richard stauduhar, member IEEE and Laurence p, Clarke Member IEEE "Multiresolution statistical Analysis of high-Resolution digital mammograms" IEEE Trans on medical Imaging,pp:603-515

[18] Lai S et al, 1989. "On techniques for detecting circumscribed masses in mammograms", IEEE Trans. on medical imaging,vol,8,no 4,pp:377-386,Dec

[19] Rosenfeld R and Kak AC, 1982. "Digital picture processing" New York Academic.

[20] Ted C. Wand and Nicolas B, Karayiannis, member IEEE, 1998. "Detection of Micro calcifications in digital mammograms using wavelets", IEEE Trans on medical Imaging pp:498-505.

[21] Sun, W., Zhang, Y. X., & Li, F. T. (2006, August). The neural network model based on PSO for short-term load forecasting. In Machine Learning and Cybernetics, 2006 International Conference on (pp. 3069-3072). IEEE.

[22] Parsopoulos, K. E., & Vrahatis, M. N. (2004). On the computation of all global minimizers through particle swarm optimization. Evolutionary Computation, IEEE Transactions on, 8(3), 211-224.

[23] Clerc, M., & Kennedy, J. (2002). The particle swarm-explosion, stability, and convergence in a multidimensional complex space. Evolutionary Computation, IEEE Transactions on, 6(1), 58-73.

[24] Kennedy, J., & Eberhart, R. (1995, November). Particle swarm optimization. In Proceedings of IEEE international conference on neural networks (Vol. 4, No. 2, pp. 1942-1948).

[25] Eberhart, R. C., & Shi, Y. (2000). Comparing inertia weights and constriction factors in particle swarm optimization. In Evolutionary Computation, 2000. Proceedings of the 2000 Congress on (Vol. 1, pp. 84-88). IEEE.

[26] Shi, Y., & Eberhart, R. C. (2001). Fuzzy adaptive particle swarm optimization. In Evolutionary Computation, 2001. Proceedings of the 2001 Congress on (Vol. 1, pp. 101-106). IEEE.